

## **Online Information 99: Inga sensationer trots bransch i snabb förändring**

***Rapport från 23rd International Online Information Meeting, London, 7-9 december 1999.***

Lars Klasén

[lars.klasen@infodata.se](mailto:lars.klasen@infodata.se)

*Onlinebranschen är inne i en stark förändringsprocess. Detta till trots kantades årets upplaga av Online Information inte av några sensationer. Den största nyheten var att sammanslagningen "Dow Jones Reuters Business Interactive" nu fått sitt namn: Factiva. Kanske var det därför det kändes som ett "mellanår" för konferensen/mässan? Eller kanske var det för att branschen befinner sig i ett skede mellan det gamla och det nya? "Upphetsningen" över webben har vid det här laget har hunnit lägga sig, utan att det ännu går att se vad utvecklingen kommer att föra med sig vad gäller informationsförsörjningens villkor.*

Denna rapport från Online Information 99 inleds med några helhetsintryck (I). Därefter (II) ges diverse uppgifter från utställningen, som trots att den var något mindre än förra året, såväl ytmässigt som vad gäller antal utställare (strax under 300), naturligtvis var omöjlig att hinna att besöka mer än till en del, inte minst därför att föredragen tog en hel del tid i anspråk. Slutligen (III) refereras hälften av de drygt 50 föredragen (varav många märkligt nog inte finns i Proceedings!) samt informeras kort om andra.

### **I. Inledning**

Det är flera år sedan de stora onlinetjänsterna var "kungar". Databasproducenterna, som förra året hade en ganska framträdande roll genom att de dels var många, dels i många fall annonserade egna webbsatsningar, var i år mera i skymundan. Däremot hade förlag och förmedlare av tidskrifter en större roll än tidigare. Naturligtvis var det länkar av alla de slag till artiklar i fulltext som var huvudnumret här.

Allt fler aktörer på nätet erbjuder gratis sökning, även i sådan information som traditionellt varit förbehållen etablerade, avgiftsbelagda, databaser och onlinetjänster. Detta avspeglade sig även på mässan, i första hand genom närvaron av de två aktörerna som gjort sig mest bemärkta härvidlag, nämligen Northern Light och Powerize.com. Förutom att de var representerade på mässan med montrar så var de också företrädare på konferensdelen medelst medverkan i föredrag och paneldiskussioner i sessionen "Web search engines for professionals: the state to play". Om något tema skall pekas ut för årets konferens var det just portaler (inklusive sökmotorer), även inte ens det temat var särskilt framträdande. Andra "teman" var XML, Knowledge Management och åtkomst till fulltextinformation – där det senaste var det som märktes mest.

### ***Storbritannien och USA dominerade föredragen***

Konferensdelen var sig lik, med drygt 50 föredrag och som vanligt ca 1.000 deltagare. Att denna del inte växer har en rent praktisk orsak, berättade David Raitt, som vanligt ordförande för konferensen. Lokalerna rymmer helt enkelt inte fler. Kanske var det därför konferenskostnaden var så avskräckande hög, ca 10.000 kr? Åsikterna om kvaliteten på föredragen gick som vanligt

isär. Med en relativt hög andel inbjudna föredragshållare uppfattar jag personligen klassen som relativt hög. Intressant nog presenterades alla illustrationer via Powerpoint och "videokanon". Inte en OH-apparat så långt ögat kunde nå. Vilket i sig inte gjorde föredragshållarna bättre. Det är som med kläder: det är bäraren det hänger på. Och tyvärr frestar enkelheten att byta bild till överutnyttjande: 50 bilder på 20 minuter är inte kul! Och tänk att det fortfarande finns proffs som läser innantill från manus! Som vanligt smög sig också en del marknadsföring av den egna tjänsten/produkten in. Ändå kan man nog hålla med David när han säger att man lyckats hålla detta i stort sett i schack – till skillnad mot National Online Meeting (i USA) och andra konferenser, som mer eller mindre har urartat.

Tyvärr har strävan efter kvalitet på föredragen resulterat i en stor geografisk snedfördelning: Hälften av föredragen hölls av personer från UK, en fjärdedel från USA och en fjärdedel från övriga världen. Sverige var i alla fall representerat, tack vare Traugott Koch, NetLab, Lund, som dels ingick i konferensens "Advisory Committee", dels bistod med ett föredrag (tillsammans med Anders Ardö, DTV, Danmark). Antalet svenska deltagare minskar – kanske som en följd av nya Scandinavia Online? - i år var det inte ens 60, vilket till och med är mindre än vad våra nordiska grannar Norge och Finland vardera ställde upp med, nämligen något över 60. Danmark hade drygt 30 deltagare. Största antalet hade som sig bör UK, ca 250 deltagare, medan Nederländerna hade nära 140 och USA knappt 90. Öststatsinslaget var som vanligt litet, ca 10 delegater. Samma antal kom från Afrika (Sydafrika, Egypten och Zimbabwe) och Asien (Japan och Indien).

Som ett tidens tecken kan man väl se det faktum att Proceedings detta år förutom i pappersform (ska) publiceras på webben. Förra året var det ju papper och CD-ROM som gällde. Att CD-ROM är på nedgång märktes också på utställningen. Här passar det bra att nämna att det enda rent svenska inslaget var Optosof – vars produktioner numera till 90 procent är webbaserade medan CD-ROM-delen sjunkit i motsvarande mån. Den största nyheten beträffande Optosof rörde dock dess köp av norska Mogul och fyra andra nordiska Internetföretag, vilket annonserades under konferensens första dag. Antalet anställda i Optosof växer därmed från 160 till 400! Samtidigt avgår Lennart Fernheden som koncernchef. (Svenska PrioInfo var också närvarande, dock inte med egen monter men väl som svensk representant för ett flertal av utställarna samt som arrangör av seminarier kring ett antal produkter: EI, Bell & Howell, Bowker-Saur, Ovid, SilverPlatter, HighWire Press, CSA, Gale och Dialog Corp.).

Apropå "omvärlden" så spreds under mässans sista dag en nyhet av en helt annan dignitet, nämligen samverkansavtalet mellan Microsoft och Ericsson, något som lära snabba upp processen mot ett "portabelt Internet". Märkligt nog talades det annars mycket lite om just detta med portabilitet. Det var inte många gånger WAP nämndes t ex i föredragen, trots en del framtidsinriktade föredrag, och jag såg inte en enda utställare som förevisade någon WAP-baserad applikation, t ex för sökning i någon databas. Ändå står det ju helt klart att det inte kommer att ta många år innan vi kommer att komma åt det mesta, kanske hela webben, via mobila utrustningar och mobiltelefoner. Men det är väl som det sägs: vi i de nordiska länderna ligger steget före den övriga världen i detta avseende.

Något som också utmärker de nordiska länderna är vår väl utbyggda infrastruktur vad gäller offentlig information, inklusive myndigheter som är mycket aktiva på webben, inte minst vad gäller att göra sin information gratis tillgänglig där. Det här är naturligtvis något som oroar de kommersiella aktörer som tar betalt för att tillhandahålla sådan information. Men av denna oro märktes i princip ingenting alls på konferensen, där frågan trots allt berördes vid några av föredragen/sessionerna. Det är bara för aktörerna att möta utbudet av gratisinformation genom att "förädla", skapa mervärde! Derwent gör det vad gäller patent. Lexis-Nexis gör det vad gäller rättsinformation. Dun & Bradstreet gör det vad gäller bolagsinformation. Här märktes tydligt den geografiska snedfördelningen. Stora länder som UK och USA, med engelska som huvudspråk, ger helt andra förutsättningar för kommersiella aktörer än små länder och små språkområden.

Av de kommersiella aktörernas (t ex Context, UK) tidigare oro kring och opposition mot EU-

kommissionens frisläppande av allt mer information gratis, inklusive viss information från TED och CELEX, märktes ingenting alls. Det tycks som om man fallit till föga, delvis för att effekten på den egna försäljningen varit ringa. EU-kommissionens monter var f ö i år lika stor som förra året. Den verkade dock inte särskilt välbesökt. För övrigt har ju ECHO lagts ner under det år som gått.

Man fick över huvud taget inte intryck av att gratisinformation på webben utgör något hot. Alla måste ju försörja sig, och om det inte är via abonnemang eller sökintäkter så gäller det att kränga annonser. För övrigt blir det ju allt vanligare att erbjuda gratis sökning och möjligheten att betala funna dokument med kreditkort. Lexis-Nexis har haft det ett bra tag och Dialog iförde det nyligen i och med samarbetet med Netscape och sina tre ämnesinriktade portaler.

## II. Utställningen

Naturligtvis hann jag bara besöka en del av de nära 300 utställarna. Jag inriktade mig i första hand på online/webbtjänster och det kanske hetaste området, nämligen till åtkomst till artiklar mm i fulltext, men också i någon mån IR-området (och lite annat). Däremot uteslöt jag enstaka databaser – de var för övrigt inte särskilt många jämfört med tidigare år: det kändes som om luften gått ur många av förra årets stolta egna webbsatsningar från databasproducenternas sida. Här kan f ö noteras att tyska Hoppenstedt saknades; alltså de som i somras deklarerade att de inte bara tänkte satsa på sin egen webbplats utan även dra tillbaka sin databas från DataStar, Reuters och de andra onlinetjänsterna. (Anm: Nu i slutet av december meddelas för övrigt att Hoppenstedt köpts av svenska Bonnier Affärsinformation!). Inga konsumentinriktade tjänster deltog i år. Slutligen kan sägas att utställningen attraherade fler besökare än någonsin, preliminärt över 20.000.

Notera att länkar till samtliga utställares webbplatser finns på [www.online-information.co.uk](http://www.online-information.co.uk) under "Exhibitors", "Exhibitor List"! Så saknar du något kanske du finner det där!

### *Online/webbtjänster*

Samtidigt som förlagen märktes mer än någonsin på mässan så tyckte jag mig notera en viss nedgång vad gäller online/webbtjänster. Antalet "traditionella" aktörer som agerar internationellt är lätt räknade; ett dussintal beskrivs nedan. Jag rangordnar dem ungefär efter hur "stora" de är. Därefter beskriver jag ungefär lika många "nya" aktörer som också agerar internationellt, som EI, Ovid, etc. Så tar jag upp några av de allt färre online/webbtjänster med i huvudsak nationell marknad som deltar på Online. Bland dem jag saknade var tyska Genios, som "alltid" varit med tidigare, och norska Lovdata, som varit med de senaste åren. Slutligen nämner jag lite, lite om tjänster för nyhetsbevakning

Men allra först måste jag naturligtvis nämna nykomlingarna *Northern Light* och *Powerize.com*, vilka båda erbjuder gratis sökning i tusentals publikationer med möjlighet att köpa funna artiklar via kreditkort. Trots sin gratisprincip uppfattade jag åtminstone Northern Lights framtoning på mässan som vilken annan online/webbtjänst som helst, med kunnig personal som gärna svarade på frågor om innehåll och sökteknik. *För beskrivningar: Se nedan under rubriken "Web search engines for professionals, där jag refererar de föredrag som hölls av David Seuss, Northern Light, och Mike Gallagher, Powerize.com!"*

*Dialog Corp.* lanserade på mässans första dag något de kallar *WebTop.com*. Det är en sökmotor för webben som är specifikt inriktad mot yrkesmässig användning. WebTop.com nyttjar Dialogs *InfoSort* för att automatiskt kategorisera information på webben. I början av år 2000 har 200 miljoner webbsidor analyserats och gjorts tillgängliga på detta sätt. Sökfrågor kan ställas i naturligt språk. Lingvistiska metoder ombesörjer att sökfrågan blir rätt ställd. En funktion kallad

k-check gör det möjligt att nyttja WebTop.com genom att bara "drag and drop" ett textavsnitt från ett dokument, ett e-mail eller en webbsida i in en speciell "k-check-ikon" på skärmen. Därefter analyseras texten och en sökning utförs automatiskt. Dialog berättade också om sin *PowerPortal*, vilken ger möjlighet för ett företag att själv skräddarsy en portal som dels utnyttjar Dialog, dels webbresurser. En nyhet var också att Dialog nu profilerar *TradStat* som en separat tjänst. Man annonserade också satsningar på att länka till elektroniska tidskrifter på webben. De tre portaler på webben som Dialog tillsammans med Netscape nyligen lanserade nämndes, dock med mindre bokstäver. Inriktningen är ju här helt klart slutanvändare, kanske främst privatpersoner. De erbjuder gratis sökning och möjlighet att köpa funna dokument via kreditkort. Avslutningsvis ryktas det om att Dialog möjligen kan vara till salu snart. Det behövs mer kapital för att finansiera nysatsningarna, dels vad gäller webben, dels inom e-handel.

Allt fler användare nyttjar **LEXIS-NEXIS** via dess webbtjänst, kallad *LEXIS-NEXIS Universe*. Det lär röra sig om över 50 %. Bland nyheterna fanns liksom hos Dialog satsningar på sökningar på den fria webben, här materialiserade genom att ge möjlighet att duplicera sina sökningar även på webben, med utnyttjandet av HotBot eller InfoSeek. Man annonserade också att databaserna från RDS, Responsive Database Services, nu gjorts tillgängliga. En reflektion är att utbudet vad gäller affärsinformation från "de stora", dvs LEXIS-NEXIS, Factiva och Dialog, blir alltmer lika ... och konkurrensen hårdnar i samma mån. Men trots de aggressiva satsningar på Europa som LEXIS-NEXIS annonserade för några år sedan tror jag inte att man ännu lyckats särskilt bra. Hur som helst har man nu fått en svensk representant, nämligen Sema Group InfoData, som integrerar tjänsten i InfoTorg. (Se även referatet av föredraget av Hans Gieskes, *Lexis-Nexis, nedan under "Going global – a new information world order?"*).

**Factiva**, det sammanslagna Dow Jones Reuters Business Interactive, hade en framträdande roll på mässan. Det är intressant att konstatera eftersom det inte är så många år sedan varken Dow Jones eller Reuters var med på mässan – och åtminstone Dow Jones hade mycket få abonnenter utanför USA. Nu säger man sig ha 6.000 kunder i 30 länder, och dess information från 7.000 källor inom affärsinformation finns tillgänglig via "more than a million desktops". Naturligtvis visade man speciellt sina "alert"-tjänster. I sommar ska sammanslagningen vara helt avslutad.

**Westlaw**, här representerat av *Westlaw UK*, som drivs av *Sweet & Maxwell* och *West Group*, informerade om sin nya satsning på rättsinformation från UK och EU. Ägaren *Thomson* berättade också om sitt senaste förvärv, nämligen nederländska *ELLIS Publications*. Thomson äger ju också svenska Fakta Info Direkt, varför man väl kan anta att Westlaw på sikt kommer att säljas via dess nya (kommande) webbtjänst.

**STN International** berättade bland annat om att sökspråket Messenger nu kan användas via "STN on the Web". Man erbjuder nu också potentiella kunder att pröva sina sökfrågor gratis för att på det sättet "locka" dem att bli kunder. Därtill presenterade man möjligheten att med hjälp av tjänsten FIZ AutoDoc erhålla köpa dokument i fulltext från över 100.000 tidskrifter, bland annat tack vare länkar till British Library Document Supply Centre (BLDSC), Springer och, sedan i november, Swetsnet (med ca 2.000 elektroniska tidskrifter från ca 50 förlag mm). STN representeras i Sverige av KTHB. (Se också referatet av föredraget av Eileen Shanbrom, *CAS, nedan under "Full text – bringing it all together"*. Där beskrivs bl a ChemPort.)

**Questel-Orbit** profilerar sig numera som en tjänst inriktad mot "intellectual property". Och här är man ju verkligen starka. Vad sägs t ex om 25 databaser med över 80 miljoner referenser till patent från 66 länder (inklusive 6 miljoner bilder), och nära 20 databaser med 12 miljoner varumärken (inklusive logotyper). Men man fortsätter också med databaser som EI, INSPEC, Pascal, PROMT, etc; totalt ett hundratal.

**FT Profile** var naturligtvis också med på mässan. Man erbjuder numera information från över 6.000 källor, varav merparten i fulltext. Märkligt nog är man dock fortfarande inte ute på webben. Det kommer dock att ske i och med FT Profile ver. 3 i början av år 2000. Svensk representant för

tjänsten är Affärsdata. I sammanhanget kan nämnas att *Kompass*, som ju finns tillgänglig via just Affärsdata, hade en egen monter där man marknadsförde sin egen onlinetjänst. Vad gäller UK ingår 190.000 företag och tjänsten finansieras till 20 % av annonser.

**OCLC**s nya version av *FirstSearch* innehåller i dagsläget 80 av de 85 databaserna, men kommer att vara komplett i januari 2000. Under år 2000 kommer man också att ge möjlighet till "per-article-purchase" från tidskrifter som ingår i *Electronic Collection Online*, dvs de 7 miljoner artiklar i PFD/fulltext (från ca 2.000 tidskrifter) som det finns länkar till från databaserna. Därutöver tillhandahålls artiklar i fulltext (enbart text) från över 2.000 tidskrifter. Beträffande Medline erbjuds länkar till artiklar i fulltext från 4.000 tidskrifter, varav ca 600 i PDF. Över 80 % av användningen av FirstSearch är nu webbaserad. FirstSearch representeras i Sverige av Bibliotekstjänst.

Tyska **DIMDI** arbetar oförtrutet vidare. Man genomför nu gradvis migreringen av sina över 100 främst (bio)medicinska databaser till IR-systemet BASIS Plus. Man har ca 100 personer anställda. Ca 50 % av verksamheten är intäktsfinansierad; det övriga är subventionerat av staten. Ett flertal av databaserna är gratis att söka. DIMDI underhåller fortfarande CIME, det EU-projekt som hade som målsättning att kytta samman fyra europeiska onlinetjänster med hjälp av Z39.50.

**H. W. Wilson** satsar liksom andra mer och mer på fulltext, under namnet *OmniFiles*. Så förutom referenser och abstracts från 2.700 tidskrifter innehåller Wilsons databaser nu också fulltext från 1.300 av dessa. På min fråga om man också skulle börja länka till fulltext fick jag svaret att man så smått funderade på det. Däremot hade man inga planer på att erbjuda gratis sökning a la Powerize.com etc. och erbjuda köp av artiklar, "eftersom vi är inriktade på användare inom universiteten".

Den enda finansiella realtidstjänst som deltog var **Bloomberg**. Bland några andra tjänster, mer inriktade på finansiell information i form av företagsinformation mm fanns **Thomson Financial** (med bl a First Call med 5 miljoner dokument om företag och marknader) och **Thomson Financial Securities Data** (med bl a nya ThomsonDirect.com), **Mergent FIS/FISonline** (med bl a EDGAR filings och Moodys Manuals) och **Dun & Bradstreet** (med information om 55 miljoner företag i 230 länder). Däremot saknades ICC i år.

**ISI**, som ingår i *Thomson Science and Technology*, är fortsatt mycket aktivt. Dess *Web of Science* erbjuder idag sökning i referenser till artiklar i över 8.000 tidskrifter. I och med avtal med 15 förlag kommer uppemot 70 % av referenserna kunna länkas till artiklarna i fulltext. Idag finns 3.3 miljoner sådana länkar. Man uppger att tjänsten nu har 2.5 miljoner användare. Ett av de senaste avtalen är med konsortier inom universitetssektorn i UK.

**EI Village** beskrevs utförligt i TLS Information 6/7, december 1999. Compendex med 5 miljoner referenser är nu försedd med länkar till artiklar i 600 tidskrifter från Elsevier. Själva torget ger tillgång till 24.000 utvalda webbplatser och 150 databaser. Speciella torg har förutom ingenjörsvetenskap också byggts upp inom papper/material-, IT- och olje/gas-områdena.

**Ovid** ger idag tillgång till 90 databaser och, via *Journals@Ovid*, artiklar i fulltext i 350 tidskrifter från 66 förlag. Man lanserar nu också *Books@Ovid*, dvs elektroniska versioner av "reference textbooks". Ännu rör det sig om 14 böcker från 2 förlag, nämligen Lippincott Williams & Wilkins and McGraw-Hill, men antalet ska öka till 35 om ett år. Bland de förlag man förhandlar med finns Oxford University Press. Ovid lanserade också *Ovid OpenLinks Toolkit*, med vilken det är möjligt för kunder att lokalt länka referenser i Ovid-databaser till artiklar i fulltext vid förlagens webbplatser. Färdiga definitioner för detta kommer att finnas för IDEAL, Springer-Verlag och några flera när produkten börjar säljas i början av år 2000. (Se även referatet av föredraget av Debbie Hull, vd vid Ovid, nedan under "Going Global – a new information world order?").

**Cambridge Scientific Abstracts** har nu ca 35 databaser, egna och andras, tillgängliga via sin

webplats. De egna databaserna finns också tillgängliga vid 11 onlinetjänster.

**SilverPlatter** har nu 220 databaser sökbara via webben. Man satsar nu hårt på att inom ramen för *SilverLinker* ordna länkar till artiklar i fulltext i tidskrifter som finns lokalt eller är tillgängliga vid förlag och förmedlare. Bland de som hittills är kontrakterade finns Academic Press, Blackwell Information Services, Catchword, EBSCO, Kluwer Academic Publishers, NRC Research Press, Project MUSE, SIAM, Springer-Verlag, SwetsNet och Walter De Gruyter. Elseviers ScienceDirect är också på gång, enligt Elsevier. Ännu är dock "bara" 5 % av materialet i databaserna tillgängligt på detta sätt. SilverPlatter har också startat en ny tjänst, *WebMedLit*, som presenterar referenser till kommande artiklar i ett 30-tal tidskrifter inom medicin, vilket uppges kunna ske ibland flera månader innan motsvarande dyker upp i Medline m fl databaser.

**Gale Group**, som ingår i *Thomson Business Information*, satsar mycket på information från dagstidningar. I och med de senaste tillskotten Daily Telegraph and Sunday Telegraph är antalet snart uppe i 100. Sammanlagt har man nu information från 6.000 källor i sina databaser.

**EBR, European Business Register**, tycks ha kommit igång på allvar, även om tjänsten ännu bara täcker in uppgifter från 11 länders aktiebolagsregister. Det rör sig alltså om uppgifter som utgör en slags "minsta gemensamma nämnare" vad gäller dessa register, dvs i princip "basuppgifter". Companies House, UK (se nedan), ingår fortfarande inte. Tjänsten är inte gratis men ändå relativt billig. Basuppgifter om ett företag kan kosta i storleksordningen 20 kr. Svenska PRV tillhör de mest aktiva vad gäller att marknadsföra systemet. I sammanhanget kan nämnas att EBR kommer att bli tillgänglig via Sema Group InfoDats InfoTorg inom kort, i och med ett avtal med PRV.

**Bell & Howell Information and Learning**, f d UMI, "återintroducerade sig" efter namnbytet och köpet av Chadwyck-Healey. Bland dess produkter märks *ProQuest* med en rad databaser tillgängliga online och/eller på CD-ROM, t ex ABI/INFORM, Dissertation Abstracts och INSPEC.

**EU-kommissionen** var naturligtvis också representerad, liksom **European Patent Office**, som visade sin gratis webbtjänst *esp@cenet*.

**Context**, UK, med *Justis.com*, dvs juridisk information från UK och EU, lanserade sin PC-produkt *Automatic J-Link*, en produkt med vars hjälp man kan länka direkt från funna referenser (på webben, i e-mail såväl som i ett Word-dokument) till dokumentet i fulltext på webben.

**Lawtel**, UK, som tillhandahåller databaser med rättsinformation från UK – t ex rättsfall och tidskriftsartiklar (delvis i fulltext), köpte i början av året Consultancy Europe. Som en följd av detta tas Consultancy Europes välkända databas *Spicers*, som refererar till det mesta som produceras av EU-kommissionen, bort från DataStar, Reuters, Lexis-Nexis, Sema Group InfoData/InfoTorg och de andra onlinetjänsterna som hittills haft dem online. Lawtel anser att man får för lite ersättning från dessa och tänker nu erbjuda den själva via webben.

**Companies House**, dvs UKs motsvarighet till vårt PRV bolagsavdelning, ger tillgång till grundläggande uppgifter om bolag gratis. Ett "paket" med utförliga uppgifter kostar ca 70 kr. Man erbjuder numera också gratis sökning i sitt register över företag på obestånd (insolvency).

Bland de få online/webbtjänsterna på mässan med i huvudsak nationell marknad, till vilka man möjligen också får räkna de två sist nämnda ovan, fanns också tyska **GBI, German Business Information**, som numera tillhandahåller ca 170 databaser, och finska **Helecon**, med 10 databaser med ekonomisk och affärsinformation.

**NewsEdge** säger sig vara det ledande företaget vad gäller nyhetsbevakning – och med sina 670.000 abonnenter och en omsättning på sådär 600 miljoner kr per år så tror man dem gärna. Man bevakar ca 1.000 publikationer från 75 informationsproducenter och säljer sina tjänster i

olika "paket", beroende på kundernas behov, med namn som *NewsPage*, *NewsEdge Objects* och *NewsEdge Notes*. Precis hemkommen från Online fick jag reda på att NewsEdge just köpts av *RoweCom*, ett amerikanskt företag inom e-commerce, specialiserat på lösningar för att hantera tidskrifter, artiklar och böcker. Köpeskillingen var svindlande 2 miljarder kr. Det ryktas f ö att NewsEdge funderar på att sälja av tjänsten *Individual.com*, dvs den tjänst som tidigare hette *NewsPage.com* och som är inriktad mot konsumenter (idag har man över en miljon abonnenter) och som står för mindre än en tiondel av omsättningen.

Nyhetsbevakning, ja ... För tre år sedan var push ett av de allra hetaste ämnena på mässan. Så icke nu. Men det fanns alltså några undantag. Ett annat var ***infoPager***, Israel, som efter två års arbete nu lanserade *newZPrint*. Det är en produkt med vars hjälp man kan automatiskt kan ladda ner nyheter i eget urval från 120 affärsinformationskällor på nätet; allt från dagstidningar till Bloomberg. Resultatet är avsett att skrivas ut på den egna skrivaren för att t ex läsas vid frukostboret eller på resan till jobbet, varför layouten också är anpassad till detta.

### ***Förlag, tidskriftsförmedlare, länkningsar ...***

Ett av de områden där det händer mest är länkningsar till (och mellan!) artiklar/tidskrifter i fulltext. I centrum står förstas förlagen, andra producenter av källmaterial och tidskriftsförmedlare. Om man därtill lägger alla andra som också är aktiva inom området, t ex andra mellanhänder, online/webbtjänster, databasproducenter m fl så måste jag måste erkänna att området är något svårt att få grepp om, om man säger så ... Ännu mer komplext blir det om man också beaktar alla konsortier som är i farten med upphandlingar. Man kan i alla fall konstatera att förlagen och förmedlarna aldrig tidigare varit så aktiva på Online Information. Här lite av vad jag fångade upp (utom de online/webbtjänster som presenteras ovan). Flera konsortier m fl var också representerade. Det skulle dock leda för långt att här redovisa även dessa här.

***Springer-Verlag*** slog på stora trumman för sin produkt och tjänst *LINK*. Den ger tillgång till över 350 tidskrifter från Springer och andra förlag samt några vetenskapliga sällskap. Man har även länkar mellan publikationer och därtill en gratistjänst, *Online First*, som ger information via e-mail om innehållet i tidskrifterna ca 4 veckor innan de publiceras. Den jag pratade med i montererna uppgav att man måste vara prenumerant för att tillgång till enstaka artiklar. (Anm: Abonnenter på STN lär kunna köpa enstaka artiklar, funna i STNs databaser, i fulltext från 800 tidskrifter utgivna av Springer-Verlag.)

***IDEAL*** var förstas också med på mässan. Detta "International Digital Electronic Access Library" ger tillgång till 250 tidskrifter från Academic Press, W.B. Saunders och Churchill Livingstone. Vem som helst kan länka från sina databaser, system etc till artiklar i IDEAL. Man ger gratis tillgång till innehållsförteckningar och sammanfattningar. I och med den nya plattform som togs i bruk nyligen erbjuder man förutom vanlig sökning en "more like this"-funktion som kan nyttjas för att finna artiklar som "liknar" en man just funnit. IDEAL säger sig ha 8 miljoner användare vid de 1.200 organisationer i 15 länder med vilka man tecknat avtal, oftast via konsortier.

***Kluwer Academic Publishers*** monter var mest inriktad på "fysiska" publikationer. Man informerade dock om *Kluwer Online*, som är en "electronic gateway" till 400 tidskrifter i fulltext, vilket snart kommer att öka till nära 700 efter att man införlivat även tidskrifterna från Plenum Publishers, vilken man nyligen köpt.

***EBSCO*** med bl a *EBSCO Online* ger tillgång till mer än 3.600 elektroniska tidskrifter. Idag ingår ca 31.000 häften med 370.000 artiklar. Länkar till tidskrifternas hemsidor ingår i alla artiklar. EBSCO erbjuder också över 20 databaser inom medicin, vetenskap och teknik. EBSCO informerade också om att bibliotek i 39 länder (varav 27 i central- och östeuropa och 10 i södra Afrika) inom ramen för Electronic Information for Libraries Direct (EIFL Direct), "by far the largest information consortium in the world", nyligen tecknat kontrakt med EBSCO om tillgång

till ett stort antal elektroniska tidskrifter.

**Science Direct** från *Elsevier Science* annonserade på Online 99 att man i vår kommer att ge tillgång även till patentlitteratur i fulltext, tack vare länkar till patentdatabaser vid Lexis-Nexis (som ju också ingår i *Reed Elsevier*). Totalt har nu ScienceDirect 750.000 artiklar i fulltext från 1.100 tidskrifter. För en månad sedan informerade man om att de som abbonerar på ScienceDirect Abstracts Database (2 miljoner referenser från 2.300 tidskrifter) tack vare ett avtal med Academic Press kommer att kunna länka direkt till artiklar i fulltext i IDEAL.

**Swets Blackwell** är namnet på det företag som bildades strax före Online 99 genom en sammanslagning av Swets Subscription Services och Blackwell's Information Services. Praktiskt kommer det att genomföras succesivt under år 2000. Företaget kommer att omsätta över 8 miljarder kr. Huvudkontoret placeras i Nederländerna. Tjänsten *SwetsNet* ger nu åtkomst till över 2.100 tidskrifter från 50 förlag i elektronisk form, vilket innebär över en halv miljon artiklar. Men tjänsten fortsätter såklart att expandera, och i början av år 2000 tror man att antalet tidskrifter passerar 3.000. Nyligen annonserade man också möjligheten att länka till artiklar i IDEAL från SwetsNet. För år 2000 planerar Swets Blackwell också en pay-per-view-möjlighet för enstaka artiklar.

**Adonis** var ett av de tidigaste initiativen vad gäller åtkomst till tidskrifter och artiklar i fulltext. Man började 1991 med 350 tidskrifter inom (bio)medicin och kemi. Idag rör det sig om 900 från 80 förlag. Antalet artiklar har nått över 1 miljon. Ännu levereras materialet för uppläggning i det egna intranätet (totalt krävs 150 GB disk per årgång!), men planen är att det ska kunna nås online/webb från år 2001. Företaget har bara 6 anställda!

**Ingenta**, UK, var en intressant bekantskap. Företaget startade 1998 men har redan skaffat sig en framträdande plats vad gäller informationssökning och åtkomst till i tidskriftsartiklar i fulltext. Trots att man inte har annonser låg man i oktober på 16:e plats i "webbligan" i UK. Ingenta erbjuder gratis sökning i artiklar i över 2.000 publikationer. Om man inte är prenumerant kan funna artiklar köpas via kreditkortsbetalning till en kostnad av i snitt ca 130 kr på basis av principen "pay per view". Av detta får förlaget 27 kr. Affärsidén är att låta förlagen betala för att vara med. Bland de ca 30 förlag som hittills medverkar finns Academic Press, Blackwell, Elsevier, John Wiley och Scandinavian University Press. Företaget köpte nyligen BIDS, Bath Information Data Services, UKs största onlinetjänst för universiteten, vilket tycks ha blivit det verkliga "lyftet" i och med att man plötsligt nådde ut till 2 miljoner personer vid universitet i UK. Ingenta erbjuder också sökning i databaser som Pascal, Inspec, Embase m fl. (Anm: Se även mitt referat av föredraget av Mark Rowse, vd vid Ingenta, nedan under Full text – bringing it all together!).

En person i Ingentas monter gav mig också mer kött på benen vad gäller en kommande "reference linking service" jag läst om. Tjänsten är ännu namnlös. Det rör sig om förlag som på icke kommersiell basis ska erbjuda länkar mellan referenser i varandras artiklar. Tjänsten ska vara gratis, men vad gäller åtkomsten till fulltexten måste man vara prenumerant. En "pay-per-view"-möjlighet ska etableras så småningom. Bland de 12 förlag som hittills ingår finns Academic Press, Blackwell Science, Elsevier Science, IEEE, Kluwer Academic, Oxford University Press, Springer och John Wiley. När tjänsten lanseras i början av år 2000 kommer den att länka över 3 miljoner artiklar, med ett årligt tillskott om 500.000. Om jag förstod det rätt baseras allt på vad

**International Document Object Identifier (DOI) Foundation** åstadkommit. Varje objekt, t ex artikel, tilldelas ett unikt nummer. I och med att detta DOI-nummer alltid åtföljer objektet ifråga så går det också att söka upp via t ex sökmotorer. Den långsiktiga ambitionen är att alla objekt över huvud taget så småningom ska kunna märkas på detta sätt! För mera information se [www.doi.org](http://www.doi.org). I sammanhanget kan nämnas att ett komplett ISSN-register nu finns online vid [www.issn.org](http://www.issn.org). Det kan testas gratis.



## ***IR/textsökningssystem mm***

IR-systemen återfinns numera under beteckningar som dokumenthanteringssystem, KM, företagsportaler, webb- och publishinglösningar mm. Nåväl, på mässan fann jag bland annat följande: Open Text/BASIS Plus, Dataware, Excalibur, NextPage/Folio, Optosof, Verity och ZyLAB. Däremot saknades PLWeb, ODARS och PC DOCS/Fulcrum, åtminstone vad jag fann. Beträffande den sistnämnda kan i alla fall sägas att det går rykten om att dess nuvarande ägare Hummingbird, Kanada, ska köpa Open Text, dvs BASIS ägare, något som dock nu i slutet av december dementeras. Vi får väl se ... IR-marknaden har länge varit ganska kärv och omstruktureringen fortsätter.

Även om ***Open Text*** förevisade ***BASIS*** (som man numera tycks föredra att kalla produkten) är dess huvudnummer ***Livelink***, dvs dess KM-koncept. Apropos KM så annonserade ***Dataware*** bara några dagar efter Online '99 att man blivit partner till företaget ***Knowledge Management***, UK. Man informerade såklart också om ***BRS-NetAnswer***, dess traditionella IR-produkt i ny skepnad, ***Query Server***, en produkt för att indexera och söka interna såväl som externa webbsidor, och ***InQuery***, sitt "data-mining" sökverktyg. ***Excalibur*** betonade som vanligt sin produkts bakgrund i intelligenta system och användningen av semantiska nätverk och mönsterigenkänning, som bland annat möjliggör "tolerant sökning".

Bland senaste nytt från ***Verity*** var att HighWire Press (Stanford University) valt Verity K2 Toolkit för sökning i de över 500.000 artiklar i fulltext från 150 tidskrifter man f n erbjuder, totalt 1/2 terabyte data, med en tillväxt om 100.000 artiklar per år.

***NextPage*** informerade om ***LivePublish***, som man väl kan kalla webbefterföljaren till ***Folio*** – åtminstone tycks det vara den produkt som man ska migrera användarna av Folio (CD-ROM) till. Till skillnad mot Folio, som använder ett eget format/en egen databas, Infobase, för lagringen så arbetar LivePublish med XML. Nu tycks det som om LivePublish ännu inte är riktigt "klar". I alla fall ryktas det om att just detta är orsaken till att svenska Fakta Info Direkt ännu inte är ute på webben med sina CD-ROM-produktioner inom juridik. (Anm: NextPage förvärvade i juli exklusiv rätt till, och köpeoption på, Folio-produkterna från Open Market, som köpte dem från Reed Elsevier 1994).

Svenska ***Optosof*** (eller Optosof/Clockwork) var som vanligt representerat på mässan. Men till skillnad mot tidigare år betonade man inte IR för CD-ROM utan snarare generellt: för CD-ROM och DVD såväl som för Internet och intranet. Så är också produktionerna numera till 90 % webbaserade. Produktnamn i fokus var ***OptoEngine***, ***OptoFind*** och ***OptoMake***. Man betonade att man arbetar med "alla" format, inkl XML (som använts vid uppläggningsen av Gula Sidorna). Produkterna kan också samverka med SQL-databaser. (Anm: Se även nyheterna om Optosof i inledningen!).

***ZyLAB*** betonade speciellt möjligheten att med hjälp av ***ZyIMAGE*** läsa in och lagra dokument (TIFF) och samtidigt automatiskt identifiera och indexera ingående texter och göra dem sökbara. Produkterna körs fortfarande endast i Windows.

En av de produkter som profilerade sig specifikt på KM var ***Knowledge Insight*** från ***Insight***, UK. Den är webbaserad och kan förse med och indexera/kategorisera information från en lång rad miljöer, inklusive MS Exchange, Lotus Notes, MS Office, ODMA-dokument, SQL-databaser, Internettjänster, realtidskällor mm. Den innehåller också "personliga agenter" och alert-funktioner och mycket annat. Bland andra produkter inriktade på KM fanns ***Autonomy*** och ***Headfast***.

Publicering och XML representerades av många, t ex ***Stilo***, med bland annat publiceringsverktyget ***Stilo WebWriter***, och ***Abbey Information Systems***. Den sistnämndas ***ActiveTEXT*** framstår som en mycket intressant produkt vad gäller att smidigt kunna granska olika versioner av dokument, t ex äldre lydelse av författningar. Där ingår också en avancerad Word till SGML/XML/HTML-

konverterare.

### III. Referat från föredrag

*Du som är intresserad av något eller några föredrag får gärna höra av dig! Med hjälp av delegatförteckningen kan jag förmedla kontakt med lämplig person i "närheten" som har tillgång till Proceedings!*

**Brian Lang**, chef för *British Library*, inledde konferensen. Lite märkligt, kan det tyckas, eftersom BL är mest känt för handskrifter och media. Men det fanns en tanke med det, nämligen att påpeka att vi fortfarande lever i en hybrid värld. På Internet finner vi en salig blandning av media. Och det produceras mer böcker idag än någonsin tidigare: 100.000 titlar per år. Och Brian trodde inte att vi kommer att få en helt elektronisk miljö. Se till exempel på TV, som inte tog död på radion. Videon tog inte död på biografen. Etc. Men samtidigt är information en "konstig vara": den kan ju användas om och om igen. Därmed ger den utrymme för att skapa mervärde, t ex för informationsproducenterna att se till att den hittas.

#### *Going global – a new information world order?*

Med sina föredragshållare från några ledande online/webbtjänster verkade sessionen "Going global – a new information world order?" mycket lovande. Tyvärr infriades det inte, utan det vi fick höra var i stort sett marknadsföring. **Tim Andrews**, vd vid *Factiva*, berättade om sammanslagningen och hur man försöker få respektive Dow Jones- och Reuters-kontor runt om i världen att behålla sin kultur och särart, mödosamt uppbyggda och väl anpassade till den miljö de verkar i. Allra viktigast: att behålla nyckelpersonerna. Han berättade också om att Factiva utvecklar gränssnitt på 7 språk ("alla pratar inte engelska!"), att man tillhandahåller information på 20 språk och att man tittar på system för automatisk översättning.

**Debbie Hull**, vd vid *Ovid* (som ingår i Wolters Kluwer Health and Science Group), inledde med att berätta om företagets utveckling. 1988 hade man Medline på CD-ROM, två kunder och en anställd. Nu 1999 sker försäljningen till 80 % på Internet, man tillhandahåller 90 bibliografiska databaser, snart 500 tidskrifter och 35 böcker i fulltext, och har 5.500 kunder. 40 % av omsättningen genereras utanför Nordamerika med målet 50 % inom 3 år. Precis som Factiva är man mån om en stark lokal förankring i de länder där man agerar: lokala kontor med tonvikt på lokal personal (speciellt tekniker), lokalt anpassade affärsstrategier, lokal information, nationella servrar och inte minst integration av sin information till lokala resurser, t ex bibliotek, och möjligheten att nyttja den direkt vid arbetsplatserna, t ex på sjukhusen, ja varför inte direkt vid sjuksängen! Debbie betonade speciellt vikten av länkar mellan olika typer av information: idag erbjuds 5 miljoner länkar mellan referenser/sammanfattningar/tidskrifter/böcker. Hon avslutade med att notera den trend mot att nyttja information via webben snarare än att ha den på CD-ROM för åtkomst i intranät.

Vd nummer tre var **Hans Gieskes**, *Lexis-Nexis*. Företaget, som ägs av Reed-Elsevier, tillhandahåller totalt 2.5 miljarder dokument. Sin storlek och sina stora ambitioner till trots tycks man ha haft svårt med sina internationella satsningar: fortfarande genereras 95 % av intäkterna i Nordamerika. 50 % av sökningarna sker idag via webben. Man har haft stor framgång vad gäller spridning på universitet och skolor i USA, där konsortieavtal innebär att 5 miljoner studerande numera kan nå Lexis-Nexis (något som jag i och för sig har hört också orsakat en del problem i och med en starkt ökat belastning på servrarna: 25 % av resurserna härrör från just denna användning – trots att avtalet bara inbringar sådär 1 % av de totala intäkterna).

### ***Portals: all change or all the same?***

Mer givande var sessionen "Portals: all change or all the same". **Martin White**, *Intranet Focus*, inledde med "Portals: here today, gone tomorrow?". Veteran som han är kunde han inte låta bli att påpeka att portaler minsann inte är något nytt. Det är ju snart tre decennier sedan portaler som Lockheed-Dialog och SDC-Orbit startade! Den första "konsument-webb-portalen" startades av Yahoo. Publika portaler i alla ära: det som betyder mest är företagsportaler. År 2001 kommer sådana att vara den generellt sett viktigaste ingången till information. En orsak är den stadigt ökande betydelsen av CRM, dvs att knyta kunder till sig i och med avtal av olika slag. Martin urskiljde 3 olika intressentgrupper vad gäller portaler: a) Informationsindustrin (Dialog, Factiva, Financial Times, etc), b) Vortaler (dvs "vertical portals, "nischade portaler") som Verticalnet, Powerize.com, m fl, där informationen närmast är att betrakta som ett medel att stöda e-handel, c) Affärssystemaktörer (SAP, Baan, etc), en ny intressant grupp för vilken slaget om "the single desktop" är i fokus. Alltså finns det aktörer med helt olika kompetenser och intressen på arenan! Vissa är systemintegratörer, andra kommer från "content"-sidan. För informationsindustrin betyder detta att man bör sträva efter att teckna exklusiva avtal med kunder om informationsleveranser, att man ska satsa på en egen, stark profil vad gäller innehåll och att man ska tillhandahålla informationen i olika format. Det är vidare viktigt att leverera kunskap, inte bara information, till varje "desktop". Slutligen ställde Martin frågan om det verkligen finns en marknad i Europa för alla de portaler som behövs, i och med de många men små nationella marknaderna: hur många portaler kan en vertikal marknad egentligen bära? Hans slutsats var att det som kommer att betyda allra mest är företagets egna portaler - och de närmaste åren kommer att innebära oerhört stora förändringar vad gäller informationsförsörjningen, generellt sett.

**John McNerney** från *NewsEdge* gjorde en sträng åtskillnad mellan portaler, som han menade var "media business", och "Enterprise services", som är "business service". Portalernas egentliga kunder är annonsörerna. De är gratis att använda, men så finner du inte heller vad du söker utan det du råkar finna. 850 företag säger sig vara "portal companies". 50 miljoner "business people" använder Internet på jobbet. De använder Internet 3 gånger per vecka 40 minuter per gång och arbetar vid företag med mer än 7.000 anställda. "Enterprise services" har 8 miljoner användare. En studie av den egna tjänsten Individual.com visar att de till över 50 % har högre examen. 50 % av dem är "manager" eller högre och 40 % har en årsinkomst över 100.000 USD. Generellt sett rör det sig om personer som är måna om att spara tid.

Titeln på föredraget av **Chris Beckett**, *CatchWord*, var "Whose brand is it anyway? Is there a life after portals for content producers?" Han trodde på en framtid för vortaler, som inom sitt ämnesområde utvidgas med realtidskonferenser, medlemsservice, egen publicering, riktade annonser och sökmotorer. Men eftersom det är ett måste för dem att bli stora kommer de dels att samverka med andra (gäller speciellt förlagen!) men också att breddas. Chris tog Ovid som exempel. De började med medicin men har nu mycket mer än så. Att etablera sitt varumärke och därmed få direktkontakt med sina kunder är allra viktigast i denna utveckling. Utan ett namn är man ingenting. Det gäller inte minst vad gäller tidskriftspublicering, eftersom ett känt namn och förtroende för källan där är A och O.

**Frederick Bowes**, vd vid *PubList.com*, uppmanade förlagen att sätta sig in i kundernas egentliga behov och önskemål och anpassa sig därefter. Kunderna är intresserade av artiklar, inte av publikationer eller förlag. Slutsatsen är att förlagen dels måste ordna så det är enkelt att söka i deras material, gratis, och så omedelbar tillgång till fulltexten, t ex via länkar, gärna via pay-per-article. Krångel med prenumerationer etc duger inte. De enskilda förlagen har en tuff tid framför sig – samtidigt är deras material eftertraktat av portalerna, vilket de såklart bör utnyttja.

Jag ställde frågan hur förlagen ska klara sig nu när även artiklar i allt högre utsträckning blir gratis på nätet. Ja, denna trend såg föredragshållarna inte som något hot. Det behövs alltid "aggregators" och andra som kan ordna smidig åtkomst från t ex intranät.

## ***Fee or free?***

Frågan om gratisinformationen fick en fortsättning i paneldiskussionen "*Fee or free – does the Web promise an unlimited source of free sales and marketing information or do you still get what you pay for?*" med **Michael Colson** (Manning and Napier), **David Green** och **Randall K Smith** (IMR Mall). De ansåg att mycket hårda tider väntar förlagen. Portaler och andra aktörer på nätet gör vad som helst för att skapa trafik. I detta ingår att erbjuda gratis information. Men hur blir det då med kvaliteten? Ja, för de företag som satsar på överlevnad ingår god kvalitet - och förstås ett känt varumärke, som skapar förtroende. Ett ytterligare starkt skäl för just "branding" är att vi vet så lite av vad framtiden har i sitt sköte. Tjänster och produkter i all ära – ett starkt varumärke består. Bland de förtroendeskapande åtgärderna ingår också relationer med kända tjänster och produkter (från t ex förlag). Men går det då inte att först satsa på gratisinformation för att etablera sig och sedan, när man blivit känd, börja ta betalt? Detta trodde inte paneldeltagarna på. För övrigt menade de att det är dumt att försöka leva på betaltjänster och avstå från de mycket stora pengar som finns i reklamen. Samtidigt måste de som satsar vara uthålliga: det kan ta årtal innan annonsörerna kommer. Mycket och seriös information är ett bra dragplåster.

I många länder upplevs gratis information från myndigheterna som ett hot mot kommersiella tjänster. Detta ville paneldeltagarna inte alls kännas vid. I USA lyckas man minsann sälja både företagsinformation som patent trots att det flödar av gratis sådan information. Tiden är ute för enbart "basal" myndighetsinformation; nu krävs det mervärde! Det kan bara de kommersiella tjänsterna erbjuda; myndigheterna alltför långsamma. Själv tror jag att slutsatsen möjligen gäller för stora länder med engelska som gemensamt språk men är mer tveksam i de flesta mindre länder i t ex EU.

## ***Full text – bringing it all together***

En intressant session var också "*Full text – bringing it all together*" – inte konstigt med tanke på den centrala roll som just förlag och länknings har. Den inleddes av **George Plosker** från **Gale Group**. Han pratade om att biblioteken måste vara "proaktiva"; det duger inte längre att bara svara på frågor i referensdisken. 90 % av de potentiella användarna besöker aldrig biblioteket! Exempel på aktiviteter är att via intervjuer ta reda på informationsbehoven, att etablera samarbete med IT-avdelningen, att aktivera sig i intranätsfrågor, att föreslå och skaffa verktyg, produkter och tjänster för åtkomst till extern information, etc. Och inte minst att sätta sig in i, skaffa tillgång till och administrera den komplexa värld som består av resurser på webben, sökmotorer, förlag, gränssnitt, länkar ... etc.

Mer substans gav **Mark Rowse**, vd vid **Ingenta**. Han pratade om "*Only connect: The role of linkage in information search and retrieval*". Totalt finns det idag 168.000 tidskriftstitlar från 12.000 förlag m fl. Antalet "core journals" är 25.000. Ännu är bara en liten del online/webb: ca 6.000 st. Men antalet växer mycket snabbt: prognosen för år 2000 är 18.000 st, år 2001 33.000 st, år 2002 49.000 st och år 2003 53.000 st! Ingenta ansvarar för driften av BIDS, Bath Information Data Services, som är UKs största onlinetjänst för universiteten. I och med avtal med Elsevier Science, John Wiley, OUP och Academic Press kan man erbjuda en halv miljon artiklar i fulltext från över 2.000 publikationer. Tjänsten används 500.000 gånger per månad, varvid 20 miljoner referenser "visas". 2 miljoner studenter m fl kan nå tjänsten, varav 350.000 är aktiva. Mest aktiva är de inom "life sciences" (46 %), "physical sciences" (20 %), "social sciences" (14 %), "health sciences" (6 %) och ingenjörsvetenskap/teknik (6 %). De största användargrupperna är studenter (61 %), forskare (17 %), "lecturers" (13 %), professorer (3 %) och bibliotek (3 %). Som mycket viktigt anses bibliografisk sökning (det tycker 75 %), länkning till fulltext från referenser (65 %) granskning av innehållsförteckningar (56 %), länkar till lokala beståndet (56 %) och länkar från citeringar till artiklar (49 %). Apropos det sista utvecklar man nu möjligheten att inte bara länka "bakåt" utan även "framåt", dvs hur en artikel citerats sedan dess publicering. Ingentas mål är att

kunna erbjuda 30 miljoner artiklar i fulltext (t ex PDF) i början av år 2000. Men länkar kan även vara länkar mellan personer! Som ett bra exempel nämnde Mark den e-mail tjänst "ask a professor" som EI Village erbjuder sina abonnenter.

**Steve Dale, Reuters**, talade om "Web site content management – a case study". "Hans" site används av 55.000 kunder och består av 800 sidor och 11.000 länkar. I genomsnitt finns 30 länkar per sida. Alla länkar generas automatiskt, vilket gör att de som bidrar med material kan koncentrera sig på innehållet. Man tillämpar decentraliserad uppdatering, men sådant material som ska vara externt tillgängligt kontrolleras av en redaktion innan det släpps.

Sist ut i denna session var **Eileen Shanbrom, CAS**, som tillsammans med tre andra från CAS och FiZ Karlsruhe författat "Getting it together on the Web: a working, collaborative approach to the full-text solution". Situationen för förlagen förändras snabbt. Det traditionella prenumerationerna blir färre och färre. Databasproducenterna ser sin chans på webben – men de möter ett hot från sökmotorerna. De senare visar dock många brister. En studie i Nature visar att 5 % av länkarna inte leder någonstans, att till och med de bästa bara indexerar 16 % av alla sidor och att sidorna var i snitt 186 dagar gamla. En frikostig användning av XML och Dublin Core skulle kunna lösa situationen – men dit är det flera år. Tidskriftsagenterna möter också ett antal hot – för att inte tala om de traditionella onlinetjänsterna, som får slåss för sin överlevnad. Men nu växer ett ny "informationsmiljö" fram. Databasproducenterna, t ex CAS, möter hoten genom att satsa på snabb uppdatering och hög kvalitet. Onlinetjänsterna, t ex STN, satsar på att bli att bli den "självklara sökmotorn", alltid enkelt tillgängliga via den egna PC:n. Så mycket fulltextinformation som möjligt är självklart, liksom länkar till förlagens bestånd. Här nämnde Eileen förstås ChemPort, som idag erbjuder tre sorters länkar: a) länkar till artiklar och patent online från STNs producenter och tjänster, b) länkar från referenser i artiklar till 15 miljoner dokument i Caplus och c) länkar till det egna beståndet. I början av år 2000 kommer man att ha 2.8 miljoner länkar. STN har idag 400 miljoner referenser, varav 11 miljoner patent och 50 miljoner i CA, och man arbetar tillsammans med 28 partners, bland dem förlag. På en fråga från auditoriet om hur informationen i STN säkras för framtiden berättade Eileen om ett projekt, "COBRA", som tar över när/om förlagen lägger ner sin verksamhet. Hur blir det med åtkomsten om man inte är abonnent då? Jo, man kommer att ta fram en pay-per-view-lösning som alla kan nyttja.

### ***Web search engines for professionals***

Konferensens mest spännande session var "Web search engines for professionals: the state to play". Den inleddes av **David Seuss**, vd för *Northern Light*, med "Making the Web meaningful – improving Web searching." Northern Light (NL) är annonsfinansierat och erbjuder gratis sökning i snart 7.000 källor – men det spelar ingen roll: dagens användare förväntar sig ändå snabb, perfekt service och de är noll-toleranta. David skräddade inte orden vad gäller NLs mål, nämligen att "indexera och klassificera all mänsklig kunskap och erbjuda den på ett enhetligt sätt". NL riktar sig både mot konsumenter och den professionella marknaden. David ansåg att de traditionella aktörerna betraktar NL och andra liknande tjänster som pirater. Men man ska minsann visa dem att man är kapabel att ha allt från bra täckning, kvalitet och uppdatering till relevansrankning och mycket avancerade sökfunktionaliteter. Enligt David är NL världens största sökmotor, idag med 200 miljoner sidor. Som källa nyttjade han välbekante Greg Notess, och var inte nådig mot de aktörer som påstår att de är ännu större än NL. "De är inte ärliga; uppgifterna sprider de därför att det är billigare att bygga pressreleaser än databaser – och pressfolk är inte sena att sprida uppgifterna utan att kontrollera dem". Så inbjöd han de journalister som tvivlade att komma och detaljstudera NL. NL indexerar hela databasen i ett svep, till skillnad mot vissa andra sökmotorer som bryter ner databasen i många små delar; något som inte visat sig effektivt (t ex Inktomi, som använder sig av 150 PCs i lika många kluster – som därför ger olika sökresultat från gång till annan beroende på att vissa kluster är överbelastade). NL indexerar hela databasen, totalt 3 TB (!), på ett dygn. Enligt David har NL också mycket färre "döda" sidor än andra sökmotorer. Vad gäller antal fulltextkällor är NL bäst med 6816 st. Lexis-Nexis News Library har 5701 st, Dow

Jones Interactive Publications Library 4366 st och Powerize.com 2914 st. Som om detta inte var nog stack David ut huvudet ännu mer och deklarerade att NL inom två år har en databas på en miljard dokument. Man arbetar också med att få sitt söksystemet att bli så intelligent att det kan "läsa dina tankar". Ett del i detta är en naturligt-språkmodul som sades fungera helt suveränt. David pekade också på den kompetens som finns inom företaget: hela 3 av 6 personer i ledningen har högre biblioteksutbildning. Som ett resultat satsar man t ex på avancerad klassificering, t ex dokumenttyp, språk, källa, etc – samt en indelning i olika ämnesområden. Automatisk relevansbedömning är mycket viktigt, och här nyttjar man statistiska metoder, länkars popularitet, metadata, mm. Slutligen berättade David om den gratis alertfunktion för dygnsvis avsökning av de senaste uppdateringarna som har fått så stor uppmärksamhet nyligen samt om att NL kan bearbeta innehåll oberoende av format, källa och medium.

**Mike Gallagher** från *Powerize.com* hade en mjukare framtoning men en lika stark tro på grattistjänsternas framtid. Tjänsten emanerar från InfoMarket och Lotus Newsstand, vilka köptes från IBM 1998 och är byggd bl a på erfarenheter från Excalibur. Detta gav en snabb start, nämligen våren 1999. Visionen är "en plattform för att göra affärer på" och till skillnad mot NL satsar man enbart på den professionella marknaden. Man har idag 100 informationsleverantörer och erbjuder 11.000 källor, varav 2.400 är "fria". Dokumenten kan nyttjas "pay-per-view" eller via abonnemang. Varför är då Powerize.com gratis när Dilaog, Lexis-Nexis etc kostar? Orsaken är att Internetanvändarna förväntar sig det! Inkomsterna kommer från reklam, provision på e-handelstjänster (inkl dokumentleverantörernas) och företagsabonnemang. Viss information kommer att vara gratis, viss kommer att kosta. Mike påstod att Powerize.com är webbens största "business research site" i och med att man nu har mer än 125.000 "medlemmar". Man erbjuder också gratis e-mail och nyheter i realtid. Bland de partners man har finns Netscape Netcenter, VerticalNet, Lotus, IBM, I-syndicate och Inktomi.

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

Sist i denna session fick vi ta del av konferensens i mitt tycke bästa föredrag. Det hölls av veteranen **Steve Arnold**. Han ville "ta oss med in i framtiden". Komponenter i denna utveckling är bland annat mobil utrustning och trådlösa förbindelser, XML och MP3. Steve gav oss följande "sökhistorielektion": 1975-1985: Onlinetjänster med kommandostyrd sökning och hårt strukturerade databaser – och en tid när 90 % av det som var tillgängligt elektronisk också fanns i databaserna, 1986-1991: Expansion även till optiska skivor i interna nät. Onlinetjänster i stark tillväxt. Antalet BBSer växer till 25.000, 1992-1995: Onlinetjänsterna förändras starkt; man jagar "slutanvändare" och har det svårt med lönsamheten; nya företag, som t ex Dow Jones, träder in, 1996-2000: Webbexplosionen. Internettekniken i nätverk leder till nya tjänster och information – samtidigt som bara 25 % av det som finns i elektronisk form kan sökas i databaser och sökmotorer. Och nu kommer perioden 2000-2003: "The reengineering revolution", där online och avancerad sökning är något som tillhör vardagen. En stor skillnad mot "fortiden" är att så mycket är gratis. När Arnold studerar de olika sökmotorerna och väger popularitet mot kraftfullhet och komplexitet finner han att "simpler and easier are better than more powerful and complex systems". De tjänster som inte klarar av att förändra sitt sätt att göra affärer inom 12 månader kommer att få stora problem. Användarna kommer att kunna välja mellan ett flertal bra tjänster – men samtidigt finna sig i att ingen sökmotor/sökplats ger allt.

## *SMEs and the information revolution*

Beträffande sessionen “SMEs and the information revolution” refererar jag bara de två första föredragen. Inledde gjorde **Sheila Webber**, *University of Strathclyde*, UK, med “A view from the barricades”. Hon redogjorde för en studie i UK av småföretags informationsanvändning och -behov. Resultaten lät bekanta: SMEs är “diverse and demanding”, att bara få dem att inse att de behöver information är svårt, formerna för stöd för sökning av information är förvirrande (med EICs, handelskammare, etc bakom), etc. En intressant slutsats var att de projekt etc som rapporteras och skrivs om är “success stories”, medan man talar tyst om misslyckaden. Därmed ges en alltför ljus bild av den roll olika former av stöd haft. Sheilas slutsats var att det i många fall var bättre att erbjuda handgripligt stöd i form av hjälp från informatiker eller bibliotekarier än teknik (PC, kommunikation etc). Orsaken till SMEs svårigheter är inte brist på intresse eller kompetens utan brist på tid!

**Bob Stenbridge**, *Derwent*, talade om “Commercial information: how the market is changing”. Han menade att förändringstakten inte är större idag än tidigare. Dessutom tas gammal visdom till heders gång efter annan: “refereeing”, sammanfattningar författade av “människor”, traditionell indexering och klassificering. Se bara på Yahoo, som använder mer mänskliga och maskinella resurser än någonsin. Han trodde definitivt inte på att gratisinformation, som kan användas utan kunskaper, kan frälsa världen. Vi är långt från gratis medicinsk information som kan göra oss till läkare, långt från gratis patentinformation som kan göra oss till lyckosamma uppfinnare, etc. Visst är det bra att myndigheterna lägger ut information gratis, men att säga att småföretagen därför blir så lyckliga är att sprida illusioner. Bob betonade risken för att gratisinformationen slår ut kommersiella aktörer och därmed försämrar informationsförsörjningen. Helt gratis information kan och bör aldrig bli verklighet. Bob nämnde slutligen att Derwents databaser växer med 20.000 dokument i veckan, eller mer än 1 miljon per år. Man arbetar mer och mer med fastprisavtal och “mass customisation”.

## *Moving past the millenium*

**Paul Maidment**, *ft.com*, inledde denna session med “The shape of things to come – new audiences and new content”. Vad är det vi talar om: medium eller distributionskanal? Paul såg två inriktningar, a) Fast linje: Trend mot bredband, PCn och TVn växer ihop (här har USA ledningen!) b) Trådlös: Mobiltelefoner och annan “handhållen” utrustning (här är Europa först!). Ju närmare vi har till “apparaten” desto större interaktivitet: 12 feet: TV och fjärrkontroll, 12 inches: PC och tangentbord, 0 feet: telefon. (Paul spekulerade också om “möjligheten” att vi får ett telefonnummer vid födseln, som sedan följer oss livet igenom). Massmarknaden utvecklas till “en-till-en-kommunikation” i global skala. Makten över informationen flyttas över till konsumenterna; information kommer att sökas, inte “distribueras”. Användarna kommer att bli allt bättre på att veta var/hur de ska söka information men de kommer inte att vara lojala utan nyttja de tjänster etc som är bäst för tillfället. Användarna är med och skapar innehåll (“communities”). Informationen kommer att skräddarsys för respektive användare. Det senare kommer delvis att ske automatiskt medelst “intelligent dynamisk inläring”. Här är just “en-till-en” nyckelordet. Detta talar för behov av att bygga upp en relation mellan t ex förlag och användare. För informationsproducenterna kommer det inte att vara frågan om att slåss om marknadsandelar utan om att skapa uppmärksamhet hos användarna.

**Chris Nuttall**, Internetredaktör vid tidningen “*Sunday Business*” (om jag uppfattade det rätt!), talade om “Envisioning the future – where will the new media take us?”. Enligt en studie från BT kommer 95 % av befolkningen att vara online år 2010. Vårt liv kommer att styras av en “intelligent personlig assistent” och speciella programvaror kommer att hjälpa oss att prioritera våra liv. Virtuellt grannsamverkan och hela världen som klassrum är andra komponenter i utvecklingen. Men andra studier visar att anarki kommer att råda, att vi får en “info underclass” och att handeln via Internet gör att skattesystemet brakar samman. Och så förde han naturligtvis

det trådlösa samhället på tal, där vi inom 4-5 år kommer att nyttja 2 Mbs mobilt istället för dagens 9.600 bps. Det möjliggör till och med högkvalitetsvideo! En rad nya tjänster baserade på positionering kommer att utvecklas, t ex "varning för vägarbeten om en km" när du kör bil.

I den efterföljande paneldiskussionen luftades naturligtvis en rad åsikter. Internet, med "en-till-en-kommunikation" som gäller, gör att journalisternas/redaktörernas roll minskar. Men de kommer inte att försvinna, ty det behövs alltid sådana som sammanställer information vi kan lita på. Tidningen, denna underbara uppfinning, printas ut hemma för att läsas på väg till jobbet. Men om nu allt blir "en-till-en", hur går det då med den allmänna samhällsdebatten? Kan den fungera utan ett gemensamt medium, med ett fåtal tongivande tidningar etc? Panelen trodde på en framtid för ämnesinriktade medier med information från auktoritativa källor och med god kvalitet. Men först måste "marknaden" mogna; vi måste komma ifrån dagens situation, med tidningar, TV-kanaler etc som innehåller en salig mix av kvalitet och skräp. I denna process kommer många medier att "flyta samman": alla blir "publicister", de må sedan kallas förlag, onlinetjänster eller internetoperatörer. Den vetenskapliga publiceringen kommer att förändras starkt. Men tidskrifterna kommer att finnas kvar; som en slutstation, där det "slutliga" resultatet dokumenteras. Och återigen pratades det om vikten av välrenommerade varumärken och "branding" ...

### *Is the price right?*

Den sista session jag åhörde inleddes av **Angus Carroll, Dialog Corp.**, med "Moving online pricing models to the Web". "En gång i tiden var allt en idyll för onlinetjänsterna – men så kom Internet", med nya aktörer, massor av gratisinformation och databasproducenter som tillhandahåller sin information själva, till lågpris. Men det har ställt till problem för användarna: sämre kvalitet, fullständighet och sökeffektivitet; problem som det helt klart är värt att betala för att slippa. Och även om gratisinformationen har god kvalitet och fullständighet så skulle det vara värt pengarna att få god sökeffektivitet - ty tid är pengar! Det man vinner i rena pengar genom att använda Internets sökmotorerna förlorar man med råge på extra söktid; något som sökmotorer aldrig kommer att kunna lösa. Men självklart måste Dialog anpassa sig till utvecklingen. Idag baseras prissättningen på grovt sett 10 % "anslutning", 25 % "kostnad för sökning" och 65 % "kostnad för information". (Här nämnde Angus att det förkättrade Dialunit är på väg bort – han höll med om att det var en miss!). "I morgon" kommer priset att bestå av två delar, X % "kostnad för sökning" och Y % "kostnad för information". Men så småningom kommer bara en sak att kosta, nämligen sökning – naturligtvis förutsatt att informationen är gratis. Och detta är något som Dialog förbereder sig på! (Allt detta är synnerligen intressant, eftersom det signalerar en återgång till onlinetjänsternas ursprungliga roll: att erbjuda sökning snarare än information i sig!). Men hur rimmar detta med att Dialog nyligen startat sina tre portaler för gratis sökning, med möjlighet att köpa dokument med via kreditkortsbetalning? Jo, det är ju för att nå även slutanvändarna.

**Peter Scott, Hemmington Scott**, talade om hur hans företag gått ut på webben med sin företagsinformation. Starten skedde 1995 och resultatet var inte bara en besvikelse, det var en katastrof! Oerhört liten användning. Men så i början av 1999 började användarna plötsligt upptäcka tjänsten och sedan dess har användningen stigit som en raket! Allt är gratis trots att samma information säljs i bokform (vilket upprört en del köpare av boken ...). Och faktum är att intäkterna från boken inte täcker även Internetsatsningen. Men hur klarar man sig då? Jo, genom annonsintäkter! Peters råd var följaktligen: Satsa gärna på Internet – men gör allt vad ni kan för att få användarna att använda er webbplats (något som dock är väsentligt svårare idag än för några år sedan, då det inte var så många som slogs om uppmärksamheten). Och ta in annonser – men bara sådana som har anknytning till den information nu erbjuder.

**Nicolas Kind, The Arden Shakespeare**, talade om något så speciellt som en tjänst kring Shakespeares samlade verk, Ardenonline, vilken öppnade i somras. Den innehåller dels verken i sig, dels en mängd relaterat material, och tillkom som ett projekt i vilket 50 personer arbetade. Arbetet pågick i 2.5 år. Resultatet hittills är 12.000 html-sidor och ett system som medger en rad



avancerade funktioner för sökning, sammanställningar, länkningar etc. Användarna är universitet, skolor och bibliotek och priset varierar från 75 GBP per år för ett litet folkbibliotek till 1.200 GBP för ett stort universitet. Kan det verkligen bli lönsamt? Ja, om två år ska det gå runt, trodde Nicolas. Tjänsten har blivit mycket populär, inte minst därför att varumärket Arden har ett så gott rykte. ("Branding" igen!)

**Ian Gailbraith**, OMS Services, talade om "Scots Origins: a case study". OMS driver sedan sedan 1998 en "genealogisk" databas, Scots Origins, på webben. Den innehåller 30 miljoner poster med äldre (minst 75 år) uppgifter om Skottlands befolkning, hämtade från motsvarigheten till våra kyrkböcker. Tjänsten har 2 miljoner användare varje månad. Man tar betalt "pay-per-view", via kreditkort. Som en av de allra första att ta betalt på detta sätt var man tvungen att pröva sig fram. Med utgångspunkt från att "fysisk" access till samlingarna kostar 17 GBP satte man priset till 6 GBP för 24 timmars användning, inklusive uttag av 30 sidor (vilket innebär i genomsnitt 30 sökningar). Egentligen skulle man kunna ha en högre gräns, men eftersom användarna alltid strävar efter att utnyttja sin "pott" så satte man den "lagom" högt. Kopior av originaldokument (10 GBP) kan beställas direkt. Erfarenheterna visar att användarna tycker det är helt OK att betala för tjänsten eftersom det trots allt blir billigare än vad "fysisk" access skulle kosta (inkl reskostnad) – för att inte tala om tidsbesparingen. Den konkurrens man har från gratistjänster på webben har hjälpt snarare än stjälpt – användarna kan själva jämföra och konstatera vilken skillnad på kvalitet det är. Problem med kreditkortsbetalning? Ja, men inte med missbruk utan med de rutiner som finns kring kreditkortsbetalning. Bankerna klarar detta dåligt; de är helt enkelt inte tillräckligt erfarna än.

## Övrigt

Bland teman för sessioner jag inte hann bevista fanns XML, rättsliga frågor och klassificering av resurser på nätet. Sessionen om XML, "Working med XML", inleddes med ett föredrag som behandlade XML "från början" och fortsatte med ett föredrag som behandlade XML och länkningar (XLink, Xpath, Xpointer, etc), därefter ett som beskrev en modell för XML och IR, baserad på ett objektorienterat synsätt, och slutligen ett som behandlade fördelar med och framtidsutsikterna för XML.

"Keeping it legal" behandlade frågor om upphovsrätt, persondata och integritet. Med utgångspunkt från Proceedings uppfattar jag att **David Mirchin**: "Protecting and using intellectual property on the Internet: exploding a myth" kan vara av stort värde för den specialintresserade, inte minst därför föredraget refererar och kommenterar en lång rad prejudicerande rättsfall. Som mycket auktoritativa uppfattar jag också **John Warchus**: "The protection of personal data online" och **Johanthan Raper m fl**: "Information privacy in a digital era".

Sessionen "Classifying Internet resources" leddes av **Traugott Koch**. Han bidrog också tillsammans med **Anders Ardö** med föredraget "Automatic classification applied to full-text Internet documents in a robot-generated subject index". Det beskriver ett arbete som ingår i EUs projekt DESIRE och går ut på att integrera manuellt utvalda webbresurser med sådana som är insamlade och klassificerade med hjälp av en sökrobot. Man har bland annat utnyttjat EIs (Engineering Index) thesarus för matchningen. Experiment har visat att i 57-66 % av fallen är den automatiska klassificeringen lika bra eller bättre än den manuella. Man ska nu gå vidare med bl a klassificering enligt Dewey. **Gerhard Möller m fl**: "Automatic classification of the World Wide Web using Universal Decimal Classification" beskriver i detalj en robotbaserad insamling och automatisk klassificering av material från den tyska webben. **T. Hickey & D. Vizine-Goetz**: "The role of classification in CORC" beskriver det arbete som bedrivs inom OCLC med att komplettera WorldCats 42 miljoner referenser med material insamlat från webben. Det sker inom projektet CORC, Cooperative Online Resource Catalog, vilket nyttjar automatisk klassificering och Dewey.

Jag vill också nämna några föredrag från övriga sessioner. **Robin Murray**: "The digital library

jigsaw: fitting the pieces together” är ett försök att beskriva och sammanfoga till en helhet de flesta av de begrepp (från portaler till metadata) och standarder (från Z39.50 till XML) som florerar.

**Tim Owen:** “Building the UK’s knowledge community: progress on the public library network” berättar om det imponerande projekt som har som mål att till år 2002 med bredbandsförbindelser sammanlänka samtliga 4.500 folkbibliotek i UK, förse nätverket med information anpassad för deras 34 miljoner besökare samt utbilda 35.000 anställda i att använda detta nätverk. Om jag förstår det rätt är budgeten på gigantiska 800 miljoner GBP!

Föredraget av **Mogens Sandfaer, Anders Ardö m fl**, “The architecture of DADS, a large digital library of scientific journals”, beskriver utförligt danska Digital Article Database Service. Det är ett stort system som innehåller 13 miljoner referenser till artiklar mm från förlag som Elsevier, Springer, Academic Press, Kluwer m fl. 10-15 % av detta finns tillgängligt i fulltext online och resten i pappersform. Webbåtkomst och Z39.50 är självklara ingredienser. I samma session, benämnd “STM Publishing – who’s invited to the revolution?” ingick också ett föredrag som beskriver SPARC, det kontroversiella initiativ till samarbete för att motverka förlagens allt dyrare tidskrifter som nu stöds av 175 medlemmar, mest universitetsbibliotek i USA.

Slutligen måste jag för rättvisans skull nämna att det också fanns sessioner inom KM (knowledge management), CI (corporate intelligence) - och en hel del annat. Det är dock inte möjligt att ta upp allt detta här.

*Jag vill upprepa att du som är intresserad av något föredrag gärna får höra av dig - jag kan då tipsa om lämplig person i “närheten” som har tillgång till Proceedings!*

[lars.klasen@infodata.sema.se](mailto:lars.klasen@infodata.sema.se)