

Från Dialindex till Discovery: Gamla och nya drömmar om ”all information på ett ställe”

Lars Klasén

”All information på ett ställe”. En dröm ända sedan de första databaserna blev allmänt tillgängliga online 1972! Artikeln gör en historisk exposé över ett antal tjänster, samordningsansatser och företeelser, vilka var och en speglar sin tids verklighet och tro på framtiden vad gäller att på bästa sätt realisera drömmen, relaterad till behoven hos avsedd användargrupp. Bland de som redovisas är Dialogs Dialindex, Scannet, IDC-TRIP, nationella samordningsansatser, Rättsbanken, TESS, Aramis, videotex, Minitel, CCL, gateways, EasyNet, PC, IANI, Z39.50, CIME, NWI, DC, den semantiska webben, EUR-Lex, Web of Science, SCOPUS, federated search, discoverytjänster och förstås Google - som idag utgör själva sinnebilderna för drömmen. Men drömmen finns kvar, liksom ambitionerna. Avslutningsvis görs några reflektioner över utvecklingen och diskuteras om det är sökmotorer, den semantiska webben eller något annat som hör framtiden till vad gäller att nå och söka i ”all information på ett ställe”.

Inledning

”All information på ett ställe”. En dröm ända sedan de första databaserna blev allmänt tillgängliga online 1972! De första manifestationerna var förstås onlinetjänsterna, databasvärdarna som Dialog, Lexis och ESA/IRS, med samlad tillgång till många databaser. Men sedan dess har utvecklingen frambringat en mångfald tjänster, samordningsansatser och företeelser med drömmen i sikte, från Dialogs Dialindex till discoverytjänster och Google. Och drömmen finns kvar, liksom ambitionerna.

Följande är en historisk exposé i form av exempel, vilka vart och ett på sitt eget sätt, i stort eller i smått, manifesterar drömmen. Den är primärt inriktad mot ”sökning” snarare än ”information”, även om det är den sistnämnda som utgör målet för sökningen. Avslutningsvis görs några reflektioner över utvecklingen och diskuteras om det är

sökmotorer, den semantiska webben eller något annat som hör framtiden till vad gäller att nå och söka i "all information på ett ställe". Om det över huvud taget är möjligt.

Dialog – den första manifestationen

Den första allmänt tillgängliga onlinetjänsten var amerikanska Lockheed Dialog, som lanserades 1972. Man satsade på databaser inom vetenskap och teknik, och kunde efter bara några år erbjuda sökning bland referenser som täckte stora delar av den vetenskapliga litteraturen. Därmed kan Dialog betraktas som den första manifestationen av "all information på ett ställe" – och blev också den första tjänst som kallades just "database supermarket".

Dialog följdes snart av flera onlinetjänster, som ESA/IRS (Italien), SDC/Orbit (USA), BRS (USA), Data-Star (Schweiz) och STN (Tyskland), bara för att nämna några namnkunniga. Inom området rättsinformation kom två onlinetjänster att var och en representera "all information på ett ställe" vad gäller USA, nämligen Mead Data Central/LEXIS, som startade 1973, och West Publishing/Westlaw, 1975. Inom medicin blev NLMs databas MEDLINE allmänt tillgänglig online vid svenska MIC-KIBIC, 1972, och året efter vid NLM (USA).

Dialindex

Under de första åren var det möjligt för åtminstone erfarna onlineanvändare att söka i samtliga relevanta databaser vid de ännu ganska få onlinetjänsterna. Men när antalet ökade så blev detta också allt svårare. Till exempel fanns det 1979 nära 400 databaser vid globalt ca 60 onlinetjänster. Vid det laget var det till och med svårt att välja databaser vid de största onlinetjänsterna. Dialog, med över 100 databaser, lanserade då en funktion kallad Dialindex. En sökfråga ställd i Dialindex visade antalet träffar i de olika databaserna och nget mer. Det var således en, med dagens ögon sett, primitiv funktion. Den gav ändå vägledning till val av databas/er eftersom många träffar indikerade databaser värda att söka i.

Trots att Dialindex var en så enkel funktion så mötte den motstånd. Så här berättade Dialogs grundare och dåvarande chef Roger Summit i en sentida, 2003, intervju: *"Some of the database suppliers were not overly excited with Dialindex, and if we offered it, they wanted to be paid a royalty for its use. We explained that it encouraged usage on the databases it identified and so was in a sense a marketing tool. In the end, most agreed to go along with it."*

Motsvarande funktioner utvecklades gradvis vid andra ledande onlinetjänster. Exempel är Data-Stars "CROS" och ESA/IRSs "Questindex". Senare utvecklades mer avancerade funktioner, med möjlighet till reell samsökning i många databaser, inklusive dubblettminering. Dialog var först med sitt OneSearch. Andra följde efter med liknande funktioner, till exempel ESA/IRS med "Questcluster" och STN med "Multifile Search".

Tidiga svenska samordningsansatser

Om onlinebranschen globalt var väldigt expansiv fram till 1980 så övertrumpfades detta ändå under de följande åren. Sverige var inget undantag; tvärtom! Här fanns 1983 ca 90 databaser vid 20 onlinetjänster, vilket innebar 4-5 procent av världens databaser och 8-10 procent av världens onlinetjänster! Faktum är att den svenska onlinemarknaden aldrig varit mer splittrad och längre från "all information på ett ställe". Det var förstås besvärligt för användarna, men även för onlinetjänsterna eftersom så många tjänster i förhållande till en ännu låg onlineanvändning gav små förutsättningar för lönsam drift. Nyttan av samordning framstod allt klarare.

En ansats till nationell samordning var det projekt som manifesterades i onlinetjänsten IDC-TRIP vid IDC-KTHB. Det genomfördes kring 1979-1980 av bland annat ett antal databasproducenter, med medel från forskningsrådet SINFDOK. Som mest fanns här drygt tio databaser. Men tjänsten drevs på försöksbasis och lades ner efter projektets avslutning. En tidigare samordningsinsats var kommunikationsnätverket SCANNET för nordiska databaser, som togs i drift 1976. Dess funktion övertogs gradvis av de nationella televerkens nät. 1979 övertogs SCANNETs verksamhet av NORDINFO, Nordiska samarbetsorganet för vetenskaplig information, och 1983 övergick det till att bli ett organ för att sprida kännedom om nordiska databaser.

Behovet av nationell samordning diskuterades nu också livligt i många sammanhang, inte minst vid DFI, Delegationen för vetenskaplig och teknisk informationsförsörjning (SINFDOKs efterträdare), och Statskontoret. Det rörde allt från möjligheten att samordna databasproduktionen (eller åtminstone databasformaten) till standardisering av sökspråk, söksystem, datormiljö och, mest extremt, statlig styrning av var databaserna borde läggas upp. 1986 beskrev jag min egen syn på "*den ideala situationen*" så här: "*... ur användarsynpunkt kanske det bästa vore en enda stor svensk databasvärd med alla de databaser upplagda som nu är utspridda på olika värddar.*" Som ett mer realistiskt mål föreslog jag en "*uppdelad samordning*", baserad på de grupper av "*stora, växande databasvärdar*" som då utkristalliserat sig inom områden som affärsinformation, rättsinformation, miljö, medicin/biologi, samhälle/kultur/humaniora och teknik/industri/produktion.

Men 1980-talet var privatiseringens tid, och utvecklingen gick snarast i motsatt riktning. Till exempel bolagiserades 1986 den statliga dataservicebyrån DAFA och blev DAFA Data AB. Därmed avhände sig staten möjligheten till en fortsättning på den redan påbörjade samordningen av såväl rättsdatabaser (Rättsdata/Rättsbanken hade startat 1981) som de stora statliga registren i drift vid DAFA (som 1988 resulterade i den kommersiella onlinetjänsten InfoTorg). En faktor som också motverkade samordning var (och är) den svenska modellen för statsförvaltningen med självständiga myndigheter. Det betydde ett "stuprörstänk", där alla "gör som de vill", i detta fall egna beslut om uppläggning och tillhandahållandet online av databaser.

Samtidigt gjorde en del av marknadens aktörer egna ansatser till samordning. Ett exempel är ett brett upplagt försök 1983-1985 att samla 9 olika databaser inom bland annat teknik, ekonomi och miljö, utspridda på olika onlinetjänster, till en enda onlinetjänst. Av delvis organisationsegoistiska skäl blev det inte så. Till exempel ställde sig Byggdok utanför.

Istället kom den onlinetjänst som bildades, TESS, TekniskEkonomisk SökService, att bestå av bara 4 databaser, producerade av Jernkontoret, Mekanförbundet, Standardiseringskommissionen och Tekniska Nomenklaturcentralen. TESS lades ner efter ett par års drift eftersom den inte bar sig. Andra påtänkta aktörer, tre myndigheter, realiserade istället onlinetjänsten Aramis, med databaser inom miljöområdet. Men även Aramis fick låg användning, och lades ner 1989.

Utan insatser från staten fick marknaden fortsätta att formera sig själv. Mot slutet av 1980-talet hade också utvecklats en handfull onlinetjänster som både var större och växte mer än de övriga. En konsolidering av marknaden, helt enkelt, som i kombination med den ständigt ökande onlineanvändningen gav förutsättningar för lönsam drift, och därmed vidareutveckling. Denna utveckling speglade i stora drag skeendet på den internationella onlinemarknaden. Dock förekom stora skillnader mellan olika länder, med extremer som USA, där marknadskrafterna fick råda i stort sett fritt, och Frankrike, med nationella satsningar som den heltäckande databasen inom vetenskap och teknik Pascal, onlinetjänsten Questel och videotextsystemet Minitel.

Videotex, Minitel

Videotex är en teknik som utvecklades i slutet av 1970-talet för att kunna nå digital information med det som fanns i alla hem: telefon och TV. Det enda som också behövdes var en tillsats som kopplade ihop dem. Via telenätet gick det så att nå tjänster med text och grafik som rymdes på TV:ns 24 rader à 40 tecken. Världens första videotextsystem var engelska Prestel, som startade 1979. Sedan startades videotextsystem i många andra länder i Europa liksom i USA, Kanada och Japan. I Sverige introducerades videotext kommersiellt 1982, med televerket som central aktör. Under namn som Datavision, Telebild, Aktievisionen och AutoTel erbjöds här ett stort antal tjänster, som telefonkatalogen, Gula sidorna, nyheter, väder, börskurser, banktjänster, bilregistret, kreditupplysningar, resebyråinformation, fastighetsdata, e-handlare och mycket mer. När videotext i Sverige lades ner 1993 fanns ca 30 000 abonnenter med ca 47 000 användare. De kunde nå över 300 tjänster vid ca 60 datorsystem, i princip allt på ett bräde, ungefär som vi idag når "allt" via Internet och webben.

Men en ännu större likhet med webben stod den franska videotexttjänsten Minitel för. Den startades 1982 av franska televerket och riktade sig främst till hemmen. Den första tjänsten var telefonkatalogen, och för att slippa en del av den tryckta upplagan så gav man bort en speciell Minitelterminal till dem som så önskade. Allt fler tjänster tillkom: banktjänster, börskurser, väder, resor, studieresultat, e-handel, e-post, databassökning i Questel (via en gateway) och mycket mer, i slutet av 1990-talet så många som 26 000 tjänster! Då gick det också att nå webben via Minitel - och omvänt att nå Minitel via webben. Som mest hade Minitel 25 miljoner användare och lades inte ner förrän 2012. Succén med Minitel berodde på strikta regler för teknik, struktur etc, fastställda av televerket, på gratis (senare billig) terminal och på enkelheten: det behövdes inget extra abonnemang eller lösenord, utan kostnaden togs ut på telefonräkningen. En utrustning, en uppkoppling – för att nå "allt". Minitel var verkligen all information på ett ställe!

CCL – ett gemensamt sökspråk

Inför starten av Euronet, 1982, hade det inom EG-kommissionen tagits fram en standard för sökspråk, benämnd CCL, Common Command Language (senare ISO 8777 och ANSI/NISO 39.58). Man försökte nu få onlinetjänsterna i Europa att implementera denna, i syfte att användarna skulle kunna söka i alla deras databaser med bara ett sökspråk, det vill säga "all information med ett sökspråk". CCL fick aldrig något reellt genomslag i den meningen att "alla", eller ens flertalet, onlinetjänster i Europa anslöt sig. En orsak var att det var tekniskt svårt. Men viktigast var att det blev enkelt för kunder som behärskade CCL att gå över till en konkurrent med CCL. Därför blev CCL aldrig aktuellt för flertalet stora onlinetjänster i Europa och naturligtvis inte heller för de amerikanska. I Sverige erbjöd 1987 ett dussin av de drygt 30 onlinetjänsterna CCL. Jag bedömer CCLs roll i utvecklingen som ganska marginell. Med tiden blev ju också sökspråk en obsolet företeelse för de allra flesta onlineanvändare.

Gateways

Allteftersom antalet onlinetjänster ökade så blev det också, från slutet av 1970-talet, allt vanligare med olika former av samverkan. Det kunde röra allt från att en svensk onlinetjänst utsågs till representant för en amerikansk - med bara gemensam marknadsföring och kanske fakturering i svensk valuta - till teknisk sammankoppling, en "gateway", där det efter inloggning till den ena tjänsten också var möjligt att söka i den andra tjänstens databaser; i extrema fall med ett och samma sökspråk. Resultatet var i alla lägen en ökad attraktionskraft för den egna tjänsten eftersom den gav tillgång till "mer information på ett ställe". En form av samverkan som gynnade både tjänsten/rna och användarna. DataArkiv var den första svenska onlinetjänst som etablerade en gateway, nämligen till Data-Star, i samband med att man blev dess representant. Redan tidigare hade det gått att via Data-Star nå DataArkiv. 1987 hade det etablerats nära 60 gateways i världen.

EasyNet

När amerikanska TeleBase 1985 lanserade gatewaytjänsten EasyNet var det något av en revolution. Det fanns då nära 3000 databaser vid ca 400 onlinetjänster. EasyNet gjorde det möjligt för användarna att via en uppkoppling nå och söka i ett stort antal av dessa databaser utan att vara kund till den/de onlinetjänster de fanns vid; ja utan att ens behöva kunna deras sökspråk. EasyNet hade inga egna databaser utan fungerade som en intelligent knutpunkt med gateways till ett flertal stora onlinetjänster. Initialt ingick bland annat Questel, Dialog, BRS och SDC/Orbit. Ett par år senare var antalet 16 stycken, med totalt 900 databaser. EasyNet lades ner 2002 efter att Dialog dragit sig ur. Tjänsten hade då för länge sedan spelat ut sin roll eftersom de största onlinetjänsterna då själva hade ett så stort utbud av databaser att det för de flesta användare räckte att vara kund vid en av dessa. Varför då gå via en mellanhand?

Samlad åtkomst i all ära – det innovativa, och unika, med EasyNet var att den erbjöd ett enda, menybaserat, gränssnitt mot alla anslutna tjänster. Systemet skötte resten. Användaren behövde bara ange sin sökfråga i klartext så valde EasyNet lämplig databas, utförde sökningen i aktuell onlinetjänst och levererade ett sökresultat. Allt till ett enhetspris för de 10 första träffarna, oberoende av onlinetjänst.

Ett avtal, en uppkoppling, ett sökspråk, en faktura – för att nå och söka samordnat i 900 databaser! Ur användarens synpunkt alltså "all information på ett ställe"! Idealet? Nja, inte riktigt, eftersom användarvänligheten hade sitt pris i form av avsaknad av många sökmöjligheter. Men kanske var det tillräckligt bra! Som Dick Kollin, EasyNets skapare, uttryckte sig i en intervju 2003: *"I know it's crappy, but it's good enough. These people that are using it are used to getting nothing, and now they get something, and that's good enough. They don't need to get everything. That was a very profound observation. If you extend that statement to what the Internet is today, it's exactly right on. It's good enough."*

InfoTorg

När DAFA Data AB 1988 lanserade sin onlinetjänst InfoTorg var det en flygande start eftersom de stora statliga registren som byggts upp där antingen fortfarande var i drift där eller anslutna dit med gateways. Därför blev InfoTorg omedelbart den ledande onlinetjänsten i Sverige, med Rättsbanken, SPAR, Basun och KB/Libris i egen drift, samt med gateways till Fastighetsdata, Bilregistret och Bolagsverket. InfoTorg tog sig alltså, på kommersiell basis, an rollen som nationell knutpunkt för myndigheternas information. I många andra länder, till exempel Danmark, sköttes detta i central, statlig regi. InfoTorgs utbud utvidgades gradvis med allt fler tjänster, tillgängliga via gateways, så inom några år ingick även Affärsdata, Aktievisionen, D&B, Soliditet, UC med flera. InfoTorg benämndes därefter ofta som en online- och gatewaytjänst. InfoTorg ingår sedan 2005 i affärsinformationskoncernen Bisnode, som också äger flera av de tjänster som man tidigare hade avtal med. Apropos "all information på ett ställe" så använder InfoTorg sedan länge just detta som sin devis.

För att anknyta till Dialog och EasyNet så kan det vara intressant att nämna att EasyNet blev tillgänglig vid InfoTorg 1992. Detta hade sin bakgrund i kunders förfrågningar om att kunna söka även i internationella databaser. InfoTorg inledde då samtal med Dialog om en gateway, men eftersom Dialog var svårflörtad så valdes EasyNet istället – via vilken det ju gick att söka i Dialog!

Persondatorn som samordnare

Den enskilt viktigaste faktorn för tillväxten av online på 1980-talet var genomslaget för IBM PC, som lanserades 1981. Den, och även andra persondatorer, ordbehandlare, mikrodatorer etc., försedda med kommunikationsprogram som emulerade terminaler, kom då att i hög grad ersätta dumma terminaler. Nu utvecklades också programvaror för persondatorer av många onlinetjänster och databasproducenter, till exempel Dialoglink från Dialog och Search Helper från databasproducenten IAC (USA), vilka erbjöd enkel uppkoppling, anpassade hjälpfunktioner med mera till dessa. Men det kom också mer

generella produkter. De enklaste erbjöd i princip bara uppkoppling och inloggning till flera onlinetjänster medan de mest avancerade även gav möjlighet att föra en gemensam dialog med flera onlinetjänster, oberoende av deras ordinarie sökspråk. Mest känd var Pro-Search från amerikanska Menlo Corp, som var anpassad för sökning i Dialog och BRS.

Man kan betrakta denna typ av programvaror för lokala knutpunkter eftersom de gav användaren samlad åtkomst till många onlinetjänster, med i alla fall teoretisk möjlighet att erbjuda samma funktionaliteter som till exempel en extern knutpunktslösning i stil med EasyNet. De fick emellertid bara begränsad spridning och betydelse. En orsak var att uppdatering av PC-program på den tiden skedde från diskett, senare CD-ROM, vilket innebar besvärliga utskicksrutiner.

IANI, Intelligent Access to Nordic Databases

1986 beslutade NORDINFO att satsa på det som kom att bli dess största och dyraste (8 miljoner kronor) projekt, nämligen IANI, Intelligent Access to Nordic Databases. Målet var att göra databaser i Norden samlad sökbara med sökspråket CCL/ISO 8777, oberoende av onlinetjänsternas ordinarie sökspråk. Konceptet baserades på att användarna med PC kopplade upp sig till en central knutpunkt som utgjordes av en kraftfull DOS- eller UNIX-dator med en programvara som översatte CCL till de anslutna onlinetjänsternas sökspråk. Programvaran uppdaterades kontinuerligt med aktuell information om onlinetjänsternas sökspråk, databaser etc., vilken filöverfördes dit med hjälp av en egenutvecklad rutin (det här var långt före webben).

Just filöverföringen blev en av projektets akilleshälar, och den ersattes senare med leverans på disketter. Programmeringsbehovet blev också större än förutsett på grund av brist på ekvivalens mellan CCL och tjänsternas sökspråk. Detta ledde i sin tur den andra akilleshälen: ständigt otillräcklig kapacitet hos den centrala datorn, liksom allt större krav på användarnas utrustning. Först 1990 fanns en fungerande prototyp, som kunde användas mot bland annat KB/Libris. Senare kompletterades IANI med databaser vid 6 internationella onlinetjänster, som Dialog och ESA/IRS. När IANI-projektet avslutades 1993 fick programvaruutvecklaren själv vidareutveckla och marknadsföra produkten, som dock blev kortlivad.

Z39.50

Även om sökspråket CCL blev obsolet så fick det ändå en viktig betydelse, nämligen som inspiration för ANSI/NISO Z39.50, en standard som definierar interaktion mellan distribuerade IR-system (*Information Retrieval-system*) och databaser. Z39.50 kom i sin första version 1988 men tilldrog sig speciellt stort intresse i mitten av 1990-talet, då den på sina håll, i synnerhet inom den akademiska världen, utmålades som en framtidsväg för informationssökning, med decentraliserade system och databaser istället för centrala lösningar. Tänk att via en plattform, ett sökspråk, ett gränssnitt kunna söka i många onlinetjänster och databaser – utan att behöva anlita någon etablerad onlinetjänst ("dinosaurier")! Z39.50 implementerades i flera av marknadens ledande IR-system och fick ordentligt genomslag inom biblioteksvärlden (onlinekataloger, bibliotekssystem, OCLC etc.), däremot inte bland onlinetjänster - även om undantag fanns.

EINS och CIME – ”Europas information på ett ställe”

Ett sådant undantag var de onlinetjänster och databaser som ingick i EINS, European Information Network Services. EINS startade 1996 som en onlinetjänst av bland annat ESA/IRS och British Library, avsedd att ersätta ESA/IRS, som skulle läggas ner. EINS gjorde att kunderna till ESA/IRS kunde fortsätta att nå och söka i databaser med ESA/IRS sökspråk. Detta realiserades med hjälp av Z39.50 och ”transparent gateways” till databaser vid olika onlinetjänster. Det började med databasen Pascal vid Questel-Orbit och snart bland annat Medline vid DIMDI (Tyskland). EINS utökades gradvis under de kommande tio åren med ca 100 databaser som Compendex, Inspec, Energy etc., vid onlinetjänster som FT Profile (UK), LexisNexis med flera. EINS lades ner 2008.

En mer ambitiös ansats, med inriktning på ”alla Europas onlinetjänster på ett ställe”, var projektet CIME, Co-operative Information Marketplace in Europe. Det pågick 1996-1998 som en del av EUs ESPRIT-program. Med en budget på 28 miljoner kronor syftade CIME till att skapa ett underlag för ett europeiskt alternativ till USAs dominerande, typ Dialog, CAS-STN med flera. Det skulle ske genom att med hjälp av Z39.50 knyta samman europeiska onlinetjänster. I projektet deltog ledande onlinetjänster i fyra länder: DIMDI, ESA/IRS, InfoTorg och Questel. Orbit. Projektets ambition var att användarna till vilken som helst av dessa tjänster skulle kunna söka i alla deras sammanlagt ca 500 databaser med den valda tjänstens sökspråk. Det var alltså fråga om en dubbelriktad och därmed ännu mer avancerad sammankoppling än hos EINS. Väl i produktion skulle europeiska onlinetjänster i gemen erbjudas att delta. Så blev det aldrig, utan projektet resulterade bara i ett par bilaterala kopplingar.

Sök- och länkplatser. NWI, Nordic Web Index

Vi är nu framme vid tiden för webbens genomslag, vilket plötsligt hade öppnat möjlighet för alla och envar att starta tjänster på nätet. Speciellt aktiva här i Sverige vad gäller information och sökning var högskolor, bibliotek och myndigheter, vilka byggde allt från enkla länkplatser till sökplatser med genomtänkt kategorisering och/eller indexering och avancerade sökmöjligheter. Exempel är Svenska Miljönätet, Kulturnät Sverige, Skoldatanätet, Digitala Salongen, Safari, Elvil, Nätverk för kunskap, Sverige Direkt, Skånewebben, Svesök och Nordic Web Index (NWI), alla med ambitionen att så väl som möjligt täcka resurser på nätet inom sina respektive områden (”all information på ett ställe”). Alla är sedan länge nedlagda; sist lades Svesök ned 2004, den tjänst som vid starten 1998 kallades ”början till en nationalbibliografi för den svenska webben”.

NWI är värt att nämna speciellt. NWI ingick i det av NORDINFO stödda samnordiska Nordic metadata project (1996-1998). Det utvecklades främst vid Lunds universitets bibliotek och gick i produktion 1997. NWI stödde uppbyggande av bland annat sökplatsen Safari, senare även Svenska Miljönätet och Svesök. Ambitionen var att indexera och erbjuda sökning i hela den nordiska webben. En av resursskäl förståelig avgränsning, men ur söksynpunkt lite märklig eftersom den som söker information sällan har anledning att behöva avgränsa sig till just Norden. Så har också generella sökmotorer typ Google och Bing sedan länge övertagit dennas (och de flesta andra nämndas påtänkta) roll. De erbjuder ju också hyfsade möjligheter till geografisk/språklig avgränsning.

Dublin Core

Arbetet som lades ner på dessa sökplatser var ändå inte helt förgäves. I NWI/Nordic metadata project ingick ingredienser som Z39.50 och Dublin Core, DC, metadatastandarden med ursprung i ett av OCLC ordnat möte i Dublin 1995. Förutom stöd till sökplatser enligt ovan gav NWI erfarenheter och kompetensuppbyggnad som kom till nytta i andra sammanhang. För att anknyta till CIME så kom även det projektet till nytta i form av kompetensuppbyggnad inom Z39.50, kommunikation och söksystem hos de deltagande onlinetjänsterna.

I slutet av 1990-talet och främst inom biblioteksvärlden knöts stora förhoppningar till DC som ett medel för att katalogisera webben och därmed skapa underlag för bättre indexering och sökning hos till exempel sökmotorer. DC har dock aldrig fått den breda acceptans som skulle krävas för detta. Det fick inte heller Z39.50 (som för övrigt integrerade DC som en del), inte heller dess efterföljare SRU/CQL (i princip sökspråksdelen av Z39.50).

Den semantiska webben

CCL, Z39.50, DC, SRU/CQL – alla är de ansatser syftande till en mer enhetlig åtkomst till information. Dessa har i sin tur bäring på en rad standarder, som RDF (Resource Description Framework) och OWL (Web Ontology Language), och ytterst på det koncept som kallas den semantiska webben, föreslagen kring 1999 av Tim Berners-Lee, webbens skapare och chef för World Wide Web Consortium. Den semantiska webben kan sägas realiseras av de resurser på webben som på ett formaliserat sätt beskriver innehåll och relationer. Med en bred spridning kommer dessa, enligt visionen, att tillsammans utgöra den ultimata samordningen av all världens information. Därtill också medlet för sökning i allt, inte bara "på" utan även "från" ett ställe - där den senare möjligheten utgör nyckeln till en värld av information som är helt decentraliserad men ändå fullt sökbar.

När Berners-Lee presenterade sin idé så kallade han den en dröm. Och så har den förblivit. "Informationsvärlden" tycks vara, precis som den alltid varit, alltför mångfacetterad och fragmenterad och utvecklas alltför anarkistiskt för att kunna samordnas. Lite ironiskt då att drömmen innehåller ingredienser som syftar till att klara informationsförsörjningen just i en decentraliserad, splittrad, värld utan överordnade instanser, centralism och stora företag

Rättsinformation

Sett i nationellt perspektiv, oberoende av land, är rättsinformation ett relativt avgränsat område, som erbjuder möjlighet till tjänster som täcker stora delar av den information som yrkesverksamma inom området behöver. Inledningsvis berättade jag att det i USA startades två onlinetjänster som var och en representerade "all information på ett ställe". Gradvis, i synnerhet på 1980-talet, startades också i många andra länder separata tjänster för just rättsinformation. Men till skillnad mot i USA var de i allmänhet initierade av staten. Som en följd av en allmän privatiseringstrend har dock marknaden på de allra flesta håll

delvis kommersialiserats, även om denna utveckling på senare tid hämmats av att myndigheterna i allt högre grad själva gör sin information tillgänglig online.

I Sverige byggde ett antal myndigheter under 1970-talet upp egna databaser vid DAFA, för internt bruk. 1981 gjordes en del av dessa allmänt tillgängliga online under namnet Rättsdata, från 1985 Rättsbanken (efter att DAFA utvidgat med fler databaser). En gemensam formatstandard togs fram, och under det ambitiösa mottot "ett heltäckande svenskt rättsinformationssystem" blev uppbyggnadstakten snabb. Efter privatiseringen 1986 blev det svårare att övertyga myndigheterna att göra sina databaser tillgängliga vid Rättsbanken varför målet delvis fick överges. Å andra sidan tillkom databaser med nyheter, referenser till litteratur med mera, vilket gjorde att Rättsbanken 1990 hade ca 40 databaser.

Staten hade i och med privatiseringen av DAFA 1986 avhämt sig möjligheten till en fortsättning på den påbörjade samordningen av rättsdatabaser. Men 1993 skedde en nystart. Då lanserades Rixlex, Riksdagens informationssystem, med delvis samma information som redan fanns i Rättsbanken. 1998 presenterades planerna på ett nytt statligt rättsinformationssystem. Trots att det fanns praktiskt fungerande centrala tjänster och system i våra grannländer, till exempel norska Lovdata och danska Retsinformation, så valde man att tro på att framtiden låg i decentrala lösningar med distribuerade databaser, i närtid baserade på Z39.50, på sikt visionen om den semantiska webben. Men ganska snart efter starten insåg man att idén om Z39.50 var en alltför avancerad ansats, varför man skrinlade den. Det skulle vara svårt nog att få uppemot hundra myndigheter att anpassa sig till en gemensam struktur – men att också få dem att implementera Z39.50 i sina system skulle helt enkelt inte gå.

Rättsinformationssystemet är ännu idag långt ifrån klart. Det går inte att söka samlat i myndigheternas rättsinformation. Men nu finns en standardiserad struktur, och allteftersom myndigheterna anpassar sin information till denna ska Domstolsverket, som ansvarig myndighet, kunna hämta in denna till ett gemensamt "rättsdatalager". Via indexering av dennas information ska all information göras samlat sökbar.

EUR-Lex är EUs tjänst för EUs gällande lagstiftning, med över 3 miljoner dokument centralt lagrade och sökbara. Allteftersom de nationella lagarna i allt större utsträckning anpassas till EU-medlemskapet så har intresset ökat för att också kunna söka samordnat i *både* EU-lag och nationell lag. Det har inspirerat lösningar och projekt, från enkla till mer storvulna. N-Lex är en existerande "gemensam portal för nationell lagstiftning". Den ingår i EUR-Lex men erbjuder inte mer än ett gränssnitt på EU:s alla språk för sökning i vilket som helst lands system. Bland storvulna projekt finns EULEGIS, som föreslagit en decentraliserad lösning, baserad på sammanlänkning av alla nationella lagar och domar till EUs centrala. Mest populära är projekt inriktade på standardiserad struktur, uppmärkning och beskrivning av dokument. Mer avancerat var projekt Estrella (2006-2008), som med utgångspunkt i semantiska webb-standarder var inriktat på kommunikation mellan olika rättsinformationssystem.

Det är självklart möjligt att bygga en sökmotor typ Google för samlad sökning i EU-ländernas rättsinformation. Men då blir kvaliteten därefter. Ibland duger "good enough", men för att åstadkomma en seriös tjänst för yrkesverksamma så är minimikravet en

generell anpassning till en gemensam informationsstandard. Frågan är om det över huvud taget går. Kanske är en central lösning den enda framkomliga vägen för att inom rimlig tid åstadkomma "all EUs rättsinformation på ett ställe"? Det skulle innebära att kopior av de nationella systemens dokument tas in till EUR-Lex för att där på traditionellt vis konverteras till ett gemensamt format - och så göras sökbara där.

SCOPUS och Web of Science

Teknikutvecklingen har inneburit allt mer avancerade söksystem, möjligheter att bygga allt större databaser och att lagra alltmer data till allt lägre priser. "Big data", helt enkelt. Konkret ser vi det i form av gratistjänster som Google, Bing, Facebook, Youtube etc., med gigantiska datavolymer. Men utvecklingen har förstås gynnat även onlinetjänsterna, som kunnat expandera sina databaser både till antal och storlek, till exempel med fler poster, mer fulltext, grafik etc., i princip hur mycket som helst (problemet är att få tag på information!). LexisNexis och Dialog (numera ProQuest Dialog) har nämnts, men självklart gynnas även andra stora som Westlaw, Factiva, EBSCOhost, ProQuest Academic med flera, liksom mindre.

Speciellt intressant i sammanhanget "all information på ett ställe" är dels den plattform för databassökning som tillhandahålls av veteranen Web of Science (Thomson Reuters), dels och framför allt SCOPUS, denna "jättedatabas" inriktad på fulltexttäckning som lanserades av Elsevier 2004. SCOPUS består av (stora delar av) innehållet i databaser som EMBASE, Ei Compendex, MEDLINE och PsycINFO, men också mycket annat. SCOPUS innehöll vid starten också information hämtad från webben med hjälp av den "vetenskapliga sökmotorn" Scirus, vilken Elsevier lanserat 2001. Scirus lades ner 2014 eftersom den inte bar sig. Den gav då tillgång till 575 miljoner "science-related" webbsidor/-platser och utgjorde i många lägen en konkurrent till Google Scholar, också lanserad 2004.

Federerad sökning

Idag är det i stort sett en självklarhet att tjänster, databaser och system, interna och externa, som erbjuder "sökning" också ger möjlighet till detta via en enda sökruta. Tack vare avancerade rankingsmekanismer, med sortering efter bedömd relevans, ger också så kallad "single search" oftast tillräckligt bra sökresultat – förutsatt ett gemensamt index. Men till exempel bibliotek och advokatbyråer har i allmänhet tjänster, databaser och system från flera leverantörer, vart och ett med sitt eget gränssnitt, index etc. Det här accepterar inte användarna - de vill kunna söka i allt på en gång, precis som i Google. Ett sätt att lösa detta är förstås ett gemensamt index. Det är dock ofta svårt att genomföra, av tekniskt/praktiska (till exempel olika format) eller administrativa (till exempel olika avtal) orsaker.

Det är här system för så kallad federerad sökning kommer in. Begreppet började användas på 2000-talet men metoden användes tidigare för så kallade metasökmotorer, vilka söker i flera av webbens sökmotorer och sammanlägger träffarna till en enda träfflista. Och det är precis så de kommersiella systemen fungerar.

Att slippa ha ett samlat index i drift må vara skönt, men det innebär ett antal nackdelar. Bland dessa är att sökfrågan i sökrutan måste vidarebefordras till gränssnitt eller APlar med helt olika funktionaliteter, varför sökresultatet i princip inte blir bättre än vad den "sämsta" källan förmår. Detta gäller allt från sökning i olika fält till relevansbedömning och rankning. Och hur slår man samman och sorterar helt disparata träffmängder från olika källor? Ett annat problem är dubletter. Och även om systemen i princip kan användas oberoende av källorna så krävs ändå avtal med licensierade källor där sådana ska användas.

Discoverytjänster

På grund av nackdelarna hos system för federerad sökning används numera inom biblioteksvärlden istället så kallade discoverytjänster av den typ som baseras på ett centralt index, byggt av formaliserad metadata och/eller fulltext från de tjänster och system som ska sökas. Sökning kan också här ske via en sökruta, men tack vare att använt söksystem kan nyttja samma index går det att få en samlad relevansbedömning och rankning, inklusive dubbletteliminering, och därmed normalt bättre sökresultat.

Begreppet discovery användes första gången 2007. En av de första discoverytjänsterna var Summon från Serials Solutions (ingår i ProQuest), som sedan det lanserades 2009 har varit en av fyra ledande på marknaden. De andra är Primo från Ex Libris Group (som hösten 2015 är på gång att förvärfvas av ProQuest), EDS (EBSCO Discovery Service) från EBSCO, och WorldCat Discovery Services från OCLC. I produkterna ingår ett antal licensierade källor, men det finns ändå viss valfrihet för biblioteken både att exkludera och tillföra källor, inklusive egna kataloger och databaser.

Produkterna blir alltmer avancerade vad gäller såväl möjligheter att integrera fler källor (även till exempel OA) som sökmöjligheter, med fler gemensamma fält och anpassning till användarens ämnesområde och, på sikt, personliga preferenser och tidigare användning (typ Google). Allt i syfte att via en tjänst, ett ställe, tillhandahålla all information som användarna behöver i sin verksamhet. Men eftersom tjänsterna är bundna till sinsemellan konkurrerande leverantörer går detta inte att uppnå fullt ut. Angående leverantörer finns problem även inom respektive tjänst med vems referenser som ska sorteras först/sist, och vilka som ska dubblettelimineras.

Discoverytjänsterna har fått ett blandat mottagande. Jag tycker det illustreras väl av följande uttalande av en bibliotekarie som själv utrett hur universitetsstuderande använder dem: *"I kind of hate to say it, since I am a librarian /.../. We pay a lot of money for discovery tools. And then I go off and just use Google Scholar."*

Avslutning

Google utgör idag själva sinnebild för sökning av "all information på ett ställe". Visst förstod vi som var med i onlinetjänsternas barndom att den fortsatta teknikutvecklingen skulle leda till kraftfullare datorer, större lagringsmöjligheter, större databaser, bättre programvaror, användarvänliga gränssnitt, bättre kommunikationer och allt fler

användare. Och förstås allt mer information. Men vi förutsåg inte att denna information till stor del skulle bestå av gratisinformation och att en så stor del av vår informationsförsörjning skulle baseras på annonsintäkter. Och det är just dessa två faktorer som är grunden för Googles och andra sökmotorers etc. existens. Tekniken i sig är inte ny; den är baserad på samma tekniska koncept och filosofi som utvecklades redan för ett halvt sekel sedan, det vill säga *Information Retrieval* (IR), må vara med tillägg av allt mer sofistikerade rankningsmekanismer.

Artikeln redovisar ett antal tjänster, funktioner, system, samordningsansatser och projekt som, från och med starten av online 1972, var och en speglar sin tids verklighet och tro på framtiden vad gäller att på bästa sätt realisera "all information på ett ställe", relaterat till behoven hos avsedd användargrupp. Själv trodde jag under de två första decennierna på, och verkade stenhårt för, en nationell (och i viss mån internationell) samordning vad gäller databasproduktion, onlinetjänster och standarder. Men efter att ha erfärit hur sådana ansatser, hur vällovliga de än har varit, misslyckats, dra ut på tiden eller inte få genomslag, medan "marknaden" lyckats skapa tjänster, produkter och "standarder" som, även om de bara siktat på att tillgodose sina egna intressen, ändå kommit till nytta, så har jag i viss mån ändrat åsikt.

Det finns i princip inga tekniska hinder för att lagra och göra all världens information sökbar via en tjänst som Google. Men den stora frågan är om det går att göra samma sak på ett bättre sätt. Ty med en fortsatt mångfacetterad och fragmenterad informationsvärld kan sökningar med IR och en aldrig så intelligent rankning aldrig bli bättre än "good enough". När Tim Berners-Lee för över 15 år sedan presenterade sin idé om den semantiska webben så kallade han den förvisso en dröm – men han trodde ändå på att skulle kunna realiseras på sikt. Men i takt med att webben kommit att utvecklas alltmer vildsint, med ständigt nya typer av information och tjänster, så lär hoppet ha svalnat. Bloggar, tweets, Youtube, LinkedIn, Instagram, Facebook, ständigt uppkopplade mobiler med platsloggning, realtidspubliserad av foto/film, etc. – det var mycket som inte fanns när drömmen formulerades! Och mer blir det, inte minst när det som idag kan kallas "dokument" utvecklas till något nytt som kanske mycket väl passar in i Tim Berners-Lees dröm!

Finns det då någon tredje väg till den ideala informationsvärlden? Det finns det säkert; hur den ser ut får framtiden utvisa. Men en mycket mer brännande fråga är vilken information som över huvud taget kan bli föremål för åtkomst och sökning, med eller utan indexering, med eller utan formaliserade beskrivningar à den semantiska webben. Kommer all information att läggas ut fritt tillgänglig på webben? Eller kan man tänka sig licensierad information som tillhandahålls mot avgift? Eller i en kombination, där allt går att söka men bara blir tillgänglig i sin helhet mot avgift, typ Google Scholar eller Google Books? Vad menas för övrigt med "helhet", när allt mer "information" hänger samman mer eller mindre intimt? Vi kan bara gissa – men att de närmaste decennierna blir omvälvande står helt klart. Och lika klart är att "all information på ett ställe", relaterat till (nästan) varje behov, kanske mot avgift, kommer att realiseras.

Om författaren



Lars Klasén

Är civilingenjör, informationsspecialist och expert inom databaser, online och söksystem. 1976-1983 arbetade han vid Informations- och dokumentationscentralen vid KTHB med sökning i internationella databaser, utbildning mm. 1982 var han med och startade informationsmäklaren SVP-Interfact AB. 1983-2011 arbetade han vid InfoTorg AB (fd DAFA, InfoData, Sema Group, mm, nu Bisnode Information) med allt från utveckling av databaser, onlinetjänster och söksystem till projektledning och omvärldsbevakning. 1986 tog han initiativ till onlinetjänsten InfoTorg, som startade 1988 och nu är Sveriges största. 1999-2006 var han redaktör för SFIS nyhetsbrev "Online- och webbnyheter". Lars är nu seniorkonsult och nås på lars.klasen@outlook.com.

Litteratur

Berners-Lee, Tim (1999). *Weaving the Web. The Original Design and Ultimate Destiny of the World Wide Web by Its Inventor*. New York, NY.: HarperCollins

Bjørner, Susanne & Ardito, Stephanie C. (2003). *Online Before the Internet, Part 1-9: Early Pioneers Tell their Stories*.

http://www.infotoday.com/searcher/jun03/ardito_bjorner.shtml

Bourne, Charles P & Bellardo Hahn, Trudy (2003). *A history of Online Information Services 1963-1976*. Cambridge, Mass.: The MIT Press

CIME- Co-operative information marketplace in Europe (1999). European Commission.
http://cordis.europa.eu/project/rcn/35858_en.html

El profesional de la información (1997). *Esa-lrs desaparecerá como host*.
http://www.elprofesionaldelainformacion.com/contenidos/1997/enero/esairs_desaparecer_como_host.html

von Hertzen, Maria (1992). *SCANNET*. Nordinfo-Nytt, 1992:4, s 32-33

Klasén, Lars (1984). *Datakommunikationer med utlandet för informationssökning (IR)*. Ds C 1984:4B Stockholm: Liber

Klasén, Lars (1986). *Informationsdatabaser - ett område för samordning och standardisering?* Ett diskussionsunderlag inför HSFR:s och DFI:s symposium om bibliografier och databaser inom kulturforskningen. Rapport från symposium. Lidingö 1-2 oktober 1986, ss 21-25

Klasén, Lars (1992). *Teknikutvecklingens roll för samordning av databaser*. I Teknologi och kompetens. Proceedings från 8:e Nordiska konferensen för Information och Dokumentation. Helsingborg 19-21 maj 1992, ss 111-116

- Klasén, Lars (1996). Onlinemarknad i omdaning. Teknikutvecklingens roll för tillgång till och sökning av elektronisk information online. I Brinnen, Martin (red.) *Telekommunikation - rättsliga aspekter. Nordisk Årsbok i Rättsinformatik 1996*. Stockholm: Norstedts Juridik, ss 143-157
- Klasén, Lars (1998). *Online i Internetåldern - Fakta och tankar kring de etablerade onlinetjänsterna, Internet och informationsförsörjningen*, Tidskrift för Dokumentation 53 1998:1 ss 9-25
- Klasén, Lars (1992). *Teknikutvecklingens roll för samordning av databaser*. I Teknologi och kompetens. Proceedings från 8:e Nordiska konferensen för Information och Dokumentation. Helsingborg 19-21 maj 1992, ss 111-116
- Klasén, Lars & Olofsson, Anders (1995). *Den svenska marknaden för online, audiotex och CD-ROM – framväxt, nuläge, utveckling och trender*. Stockholm: TELDOK
<http://www.teldok.net/pdf/Teldok-rapport93.pdf>
- Klasén, Lars (2001). *InfoTorg from past to present: on a large Swedish online service and its role in the online market*. The New Review of Information Networking 7 2001 ss 239-249
- Klasén, Lars (2002). *Rätt på nätet. Om den rättsliga informationsförsörjningen i Sverige och statens roll för dess utveckling*. Tidskrift för Dokumentation 57 2002:4 ss 111-120
- La memoire du MINITEL et des reseaux X.25 (2015). <http://www.minitel.org/>
- Mickos, Elisabeth et al (red) (2007). *Online development in the Nordic countries*. HELDA, Helsingfors universitet. <https://helda.helsinki.fi/handle/1975/1490>
- Orton, David (1995). *Online searching in science and technology*. London: The British Library
- Parry, Marc (2014). *As Researchers Turn to Google, Libraries Navigate the Messy World of Discovery Tools*. The Chronicle of Higher Education <http://chronicle.com/article/As-Researchers-Turn-to-Google/146081/>
- Schröder, Maria (1992). *Evaluering av det nordiska projektet IANI*. Nordinfo-Nytt, 1992:3, s 4-30
- Scardilli, Brandi (2014). Getting the Most out of Discovery Service.
<http://newsbreaks.infotoday.com/Spotlight/Getting-the-Most-Out-of-Discovery-Service-97380.asp>
- Warren, Dennis (2007). *Lost in translation: The reality of federated search*. Australian Academic & Research Libraries 38 2007:4 ss 258-269
https://openlab.citytech.cuny.edu/thecoolfoxes/files/2014/05/Lost-in-Translation_The-Reality-of-Federated-Searching.pdf
- Williams, Martha E. (1986). *Transparent Information Systems Through Gateways, Front Ends, Intermediaries, and Interfaces*. Journal of the American Society for Information Science 37 1986:4 ss 204-214