

Bibliografi 3.0 – några reflektioner

Följer man *Oxford English Dictionary* lär den något besynnerliga termen ”webliography” dykt upp för första gången under sommaren 1995. En webliografi var uppenbarligen ett slags ”lista av elektroniska verk eller dokument”, speciellt sådana hävdades det ”som relaterar till ett specifikt ämne”. Strukturen var den enklast tänkbara; en anspråklös hemsida med hyperlänkar till olika dokumentkällor på webben. Att termen dök upp i mitten av nittiotalet är inte ägnat att förvåna; i västvärlden hade ”the World Wide Web” då så sakteliga börjat att fått sitt publika genomslag. Användandet av detta globala informationsmedium hade exempelvis väsentligen underlättats genom introduktionen av nya, snabbare webbläsare som Mosaic och Netscape Navigator.

Utseendemässigt dominerades den tidiga webben av en hypertextuell länkstruktur av skrift. En webliografi skilde sig dock från en bibliografi i det att den både var en förteckning och ett mer eller mindre omfattande hyperlänkat dokument av förbindelser till andra dokument. Den bokstavliga katalogartade informationen var med andra ord *inte* sig själv nog. Med det nya mediets hjälp kunde det som listades på ett enkelt sätt digitalt synkroniseras med de dokument som utgjorde själva förteckningen. Till skillnad från bokmediet gjorde webben det möjligt att konvergera index och innehåll.

Nuförtiden förefaller emellertid termen webliografi mer eller mindre skattat åt förgängelsen; sajten webliography.org har exempelvis inte uppdaterats på flera år. En webliografi framstår där-

till som något av en tautologi i nätverksreproduktionens tidsålder. Kod för de kommunikationsprotokoll som sköter datatrafiken på internet – liksom på webben – är ju till sin binära natur uppbyggd av ständiga förbindelser. Google skulle exempelvis kunna betraktas som en enda gigantisk webliografisk maskin. En sökning genererar ju blixtsnabbt träffar i form av en närmast oändlig lista av länkar. Som kommunikationsnätverk är webben – och i synnerhet den så kallade web 2.0 – därtill helt baserad på länkar och kopplingar, anslutningar och anknäpningar på alla möjliga (och omöjliga) sätt. Samtidens snabba nätverk av nätverk – där dator och internet alltmer vuxit samman till en oskiljaktig helhet – ligger därför till grund för den sociala webb 2.0, med ledord som ”samarbete” och ”deltagande”, ”användarvänlighet” och ”konvergens”. Vidare utgör fildelning ett exempel på det kanske mest sofistikerade sätt som internet *delar* upp information på. När distribution av information via en hemsida blir trög och sävlig av för många användare, så förhåller det sig tvärtom med ett fildelningsprotokoll som BitTorrent. Ju fler användare – desto snabbare distribution. Så fort du laddar ned, blir du *per definition* också en uppladdare av de delar av filen du i samma sekund tagit hem. E-böcker lär för övrigt utgöra ungefär en procent av den totala trafikmängden data som överförs i de globala fildelningssystemen.

Även om termen webliografi håller på att försvinna från den virtuella agendan, råder det knappast något tvivel om att det bibliografiska tänkandet väsentligt upptraderas av den binära

informationsteknologin. Inte minst är det intressant att resonera om bibliografins grundläggande egenskap som systematisk förteckning i ljuset av de nya förbindelser som webbmediet erbjuder. En bibliografi kan givetvis ha många syften. Redan den lexikala betydelsen antyder att termen både refererar till forskning om böcker som föremål – ”the study of the bookness of books” som någon fyndigt formulerat det – och till en systematisk beskrivning av böcker, det vill säga ett slags objektorienterad förteckningspraktik med månghundraåriga anor.

Man bör emellertid lägga på minnet att alla böcker som idag produceras *a priori* är immateriella digitala filer; så även splitternya nationalbibliografier – det vill säga innan de trycks och blir till analoga och materiella objekt. Om man för enkelhetens skull lämnar beständighetsfrågan därhän – även om det för ett nationalbibliotek nog vore bättre att spara somliga böcker som masterfiler än som ting; limningen av exempelvis pocketböcker är ju inte direkt en bevarandemetod *sub specie aeternitatis*, medan digitaltekniken har framtiden för sig – så är denna bokens digitala ursprung intressant när man reflekterar över ett slags digitalt bibliografiskt tänkande, en ”bibliografi 3.0” om man så vill givet den semantiska webbens tilltagande betydelse. I det följande skall dock inte alls ordas om bibliografiska databaser, utan snarare om det sätt som webbens mediala informationsnätverk är i färd med att förändra hur information om böcker presenteras – eller för all del medier överlag. Om webben på nittio-talet kännetecknades av *hypertextualitet*, så har ju denna numera uppdaterats till att istället handla om ett slags *hypermedialitet*.

En traditionell bibliografisk beskrivning förutses ofta innehålla uppgifter om en boks författare och dess titel, tryckort och tryckår, liksom förlag och eventuella andra uppgifter nödvändiga för att omisskännligen identifiera boken. Om bibliografer tidigare var upptagna med att sammanställa rullor, förteckningar och register över böcker med slik information, är det idag med digital teknik möjligt att länka en fullständigt exempellös mängd upplysningar till en bok. Att den traditionella informationen ofta är till fromma för forskare – och inte minst för biblioteken själva – är det få som betvivlar, men att den information som man *annu dazumal* kom

överens om att inkludera skall fortsätta att gälla framöver är inte lika självklart. Att bibliografiska uppgifter ser ut som de gör beror förstås på historiska omständigheter; precis som ett katalogkort på 7,5 gånger 12,5 centimeter en gång i tiden gjorde det nödvändigt med till exempel olika förkortningar, så har den bibliografiska genrens sina egna mediala förutsättningar – och därmed begränsningar. Här finns inte plats att i detalj gå in på dem, poängen är snarare att i en tid som erbjuder helt nya informationsmöjligheter så bör man vara uppmärksam på den relativa godtyckligheten i standardiseringen av uppgifter som ett bibliotek *de facto* producerar – även om en längre tids praktiska verksamhet förstås givit systemet stadga.

Bibliografiska förbindelser mellan förteckning och bok förefaller till exempel vara ett område som med fördel skulle kunna uppgraderas både en och två gånger. Här erbjuder webbens sätt att organisera information genom framför allt ”taggning” en modell. Taggar – ett slags identifikatorisk och ofta ämnesorienterad metadata kopplad till ett enskilt objekt – har kommit att bli helt centralt för webb användare i upplänkandet av såväl böcker som medier. Elaka tungor påstår till exempel att Googles köp av videosajten Youtube för många miljarder kronor enbart var ett sätt att få tillgång till sofistikerad medial användarinformation – det vill säga hur användare taggat olika videos – information som genom feed-back förs tillbaka in i Googles system.

Enligt Wikipedia – en oöverträffad källa över det nya informationslandskapet – är taggning ”en klassificeringsmetod som innebär att man märker något med ett kategoriserande begrepp. I en ordlista kan exempelvis ordet ’vara’ taggas med begreppen ’Verb, infinitivform’. Avsikten med taggningen är oftast att man sedan ska kunna inlemma den taggade entiteten i ett klassificeringssystem i någon form.” Taggning har också på kort tid blivit det dominerande sätt som webbens användare kommit att organisera data på. På fotosajten Flickr finns till exempel idag kanske uppemot 10 miljoner olika slags taggar som användare kategoriserat sina fotografier med. Varje bild kan ges 75 stycken taggar, och den mest populära taggen är i skrivande stund ”wedding” – klickar man på den får man träff på mer än 8,5 miljoner fotografier! Olika webbapplikationer

kan också kombinera taggning på sofistikerade sätt. Google Maps/Earth utgör ett exempel bland många. Genom öppenhet i programkoden är det bland annat möjligt att med en mobilkamera direkt länka upp fotografier av en plats till Google Maps/Earth via 3G-nätet. Har kameran en GPS (Global Positioning System) präglas varje fotografi också av en rumskoordinat, liksom förstås den tidpunkt då bilden togs.

Den här typen av automatgenererad taggning av ett medieobjekt – ett foto, en bok eller en film – menar många experter är den riktning som webben rör sig mot. Den så kallade semantiska webben – eller webb 3.0 – innebär i korthet ett slags förenklad process för informationsutbyte mellan datasystem, det vill säga att ett informationsobjekt på webben själv kommunicerar med exempelvis en sökmotor vad objektet innehåller för information. Kommunikationen underlättas av till exempel taggar, och tanken med den semantiska webben är att om någon söker efter, säg ”bibliografi”, så letar exempelvis Google *inte* upp sajter med information om det ordet som det hyperlänkas till (vilket ju är fallet idag). Den semantiska webben kommer snarare att göra det möjligt för Google att leta fram ”ren information” om termen ”bibliografi”, vilken sedan sekundsnabbt direktgenereras till sökgränssnittet, som alltså drar samman diverse ”bibliografi”-informationsflöden i cyberrymden och presenterar dem för användaren. För att detta skall vara möjligt måste dock informationen *per se* ha ett slags metataggar – det vill säga information om sig själv. Antingen kräver det taggning till förbannelse av användare själva. Eller, mer sannolikt, kommer mediematerial framöver (text, bild, ljud) i allt högre grad att automatindexeras och så att säga tagga sig självt. Som exemplet ovan med GPS-koordinater på ett fotografi antyder, sker det redan i en ökande omfattning. En Nokia mobiltelefon är faktiskt idag den apparat som genererar mest medial information alla kategorier. Och inom en inte alltför avlägsen framtid kommer förmodligen också en sådan här artikel om nya former av bibliografi att kunna självgenerera taggar om sitt eget innehåll. Google ligger här givetvis i framkant och företaget satsar stora summor på att göra sina system så artificiellt intelligenta som möjligt. Bland annat lär deras enorma inskanningsprojekt av världslitteratur inte främst vara avsett för oss användare

– den primäre läsaren är istället Googles eget datasystem som studerar flitigt för att bli allt mer semantiskt intelligent.

Eftersom alla böcker idag är digitalt producerade skulle en bibliografi över ett speciellt ämnesområde med andra ord framöver kunna tänkas genereras av de böcker den själv består av – *per automatik*. Det gäller även äldre litteratur, förutsatt att den digitaliserats. OCR-skanning av böcker – och här utgör applikationen Google books en intressant förstudie – skulle alltså inte enbart ha den fördelen att det för forskare blev enklare att söka i dem. Det skulle också göra det möjligt att via webben integrera den information de innehåller, förutsatt att innehållet indexerats och taggats. Min poäng här är dock främst att lyfta fram hur taggning är ett sätt att organisera data på ett annorlunda sätt än till exempel det klassiska Dewey-systemet. Vad internet erbjuder är helt enkelt nya sätt att strukturera och ordna information, där den kanske mest grundläggande skillnaden mot exempelvis Deweys universella biblioteksklassifikation är möjligheten att upphäva gränsen mellan metadata och medium. En som gått till storm mot Deweys alltmer antikvarerade system är bland annat David Weinberger. I hans bok, *Everything is Miscellaneous: The Power of the New Digital Disorder* – en bok som för övrigt är dedicerad till ”alla bibliotekarier” – skriver han att den digitala världen ”aldrig stött på information som den inte tycker om, och som den inte kan använda till något.” *Allt* är metadata har han vidare hävdad, vilket just antyder hur traditionell bibliografisk metod blivit mer eller mindre anakronistisk. På kommersiella boksajter som Amazon.com hittar man ju idag automatgenererad information om hur mycket en bok sålts, dess omslag och en stor del av dess innehåll, rättigheter, recensioner av både professionella kritiker och vanliga läsare, vad de som köpt boken införskaffat för andra boktitlar, vilka taggar boken förknippas med, vilka filmer eller spel den utgjort förlaga till, och så vidare *ad nauseum*. All denna information är med nödvändighet inte intressant – och den är framför allt där för att få oss att konsumera flera böcker. Det nya informationslandskapet kännetecknas förvisso av redundans, men information i överflöd är ändå alltid bättre än knapphet av densamma. I en bibliotekskontext kan man dessutom tänka sig en helt

annan upplänkad bibliografisk data än den som Amazon.com erbjuder: hur mycket en bok lånats ut, hur ofta den citeras, hur många kurslistor den figurerar på, vad experter inom området anser om boken och inte minst vilka digitala tillgänglighetskriterier som är förbundna med den.

Avslutningsvis bör det just noteras att för att förbättra och effektivisera informationsförsörjningen inom högre utbildning och forskning, så kommer det framöver att bli alltmer centralt att indexera texters och böckers juridiska status. Wikimedia commons och dess databas för fria bilder, videofilmer och musik utgör här en förlaga som bibliotekssektorn borde titta närmare på. Syftet med ny informationsteknologi kan faktiskt inte längre reduceras till att bara handla om förenklade eftersökningar av information *om* – en modell som i bibliotekssektorn varit råd-

de i mer än två decennier. Wikimedia commons erbjuder knappast katalogposter av sitt material – de ger användaren (eller forskaren) tillgång till ”the real thing”, nämligen materialet i sig. För att vara uppriktig så tror jag att som webbanvändare är de flesta forskare idag måttligt intresserade av bibliografisk information (även om det givetvis finns undantag), de vill ju naturligtvis se materialet – även om boktexten består av en inskannad binär bild. Det är ju något av en kunskapsmässig anomali att även om en bok digitaliserats, så kan man som användare ändå bara ta del av information *om* den. Givetvis beror det på juridiska omständigheter, men ett uppdaterat och webbanpassat bibliografiskt tänkande bör likväl se det som en självklarhet att information om en bok, i görligaste mån bör kombineras med en digital version av verket ifråga.

NYHETER & NOTERAT

Dewey nästa

Kungl. biblioteket övergår från det svenska klassifikationssystemet SAB till det internationellt spridda DDC (Dewey Decimal Classification).

– Detta är ett spännande steg i rätt riktning, säger riksbibliotekarien Gunnar Sahlin som fattat beslutet om övergången. Det effektiviserar katalogiseringsarbetet och är ett gott exempel på den internationella samordningen.

Beslutet fattades efter det att Nationella referensgruppen, som består av bibliotekschefer från ett tiotal svenska forskningsbibliotek, ställt sig positiv till byte av system. Flera forskningsbibliotek väntas göra samma sak. Själva övergången kommer dock att dröja några år eftersom delar av Dewey måste översättas.

– Motiveringen är att internationalisera och rationalisera klassificeringen av böcker, säger Magdalena Svanberg på KB som utrett konsekvenserna av en övergång till Dewey. I dagsläget har runt 80 procent av de utländska böcker som

forskningsbiblioteken köper in redan en Deweykod.

– Övergången till Dewey skulle också göra svenska böcker mer synliga i utländska bibliotekskataloger, tillägger Magdalena Svanberg.

Övergången förväntas inte medföra något krångel för låntagarna. Man behöver inte kunna klassifikationskoderna för att göra ämnessökningar. På sikt kommer man ta fram sökverktyg som gör det möjligt att söka litteratur klassad med SAB och Dewey tillsammans.

– Biblioteken behöver inte nödvändigtvis flytta hyllor och märka om böcker för att anpassa sig till Dewey, menar Magdalena Svanberg. Skälet är att de rubriker som är knutna till Deweys sifferkoder stämmer bra överens med SAB-systemets rubriker.

Skytningen behöver dock ändras så att de nya koderna finns med. En hylla som idag är märkt med ”Historia (K)” ändras till ”Historia (900)”, då sifferkombinationen 900 står för historia.