

Browsing i bibliotekskataloger ved hjælp af Topic Maps

Af Knut Anton Bøckman

Abstract

Artiklen begrundes og beskriver opbygningen af et emnekort (topic map) på basis af et større humanistisk forskningsbiblioteks katalogposter. Formålet er tredelt – materialeintegration i søgefladen, udvidet browseadgang til gavn for en broget brugergruppe og semantisk støtte til emnevokabularet. Der beskrives, hvordan en ontologi opbygges og udvikles til en første prototype, som vurderes med henblik på den grundlæggende funktionalitet. Endelig vurderes, hvordan emnekort kan benyttes på mindre skala i bibliotekssammenhæng.

Indledning

Denne artikel beskriver og vurderer et projekt, hvis formål var at udvikle en alternativ og mere brugervenlig præsentationsmetode til Danmarks Kunstbiblioteks katalog. Som del af projektet blev løsningsforslaget fremstillet som en første prototype, der skitserer hvordan katalogdataene giver anledning til nye funktioner, når de transformeres til et emnekort (topic map). Der er flere grunde til valget af emnekortteknologien, som især bruges til organisation af større informationskomplekser¹. For det første er emnekort opbyggede som et netværk af forbundne emnesider, der hver især præsenterer en informationsenhed sammen med dennes forbindelser til andre informationsenheder. Strukturen lægger til rette for browsing i emnekortets informationer, og netop browsing (i bibliotekskatalogens data) var den primære funktionalitet, som projektet efterspurgte. Ønsket om en browsefunktion udspringer fra særlige brugerforhold, der knytter sig til bibliotekets domæne. Formålet begrundes i næste afsnit, der også kort diskuterer, hvordan projektets strategi adskiller sig fra de to dominerende retninger inden for katalogudvikling. Den anden grund til brugen af emnekort ligger i deres evne til at håndtere den særlige sammensætning af humanistiske fags emnevokabular, der beskrives i tredje afsnit. Endelig havde det betydning, at emnekortteknologien (som beskrives i fjerde afsnit) er en forholdsvis enkel og billig teknologi med stor fleksibilitet, både i integration med anden teknologi og i videreudvikling. De grundlæggende elementer, der indgår i det aktuelle emnekort forkla-

Knut Anton Bøckman, cand.phil., MA & cand.scient. bibl., ansat ved Danmarks Kunstbibliotek, e-mail: knuant@kunstbib.dk

res i afsnit fem, og derefter fremføres prototypens foreløbige resultater. Fordele, problemer og mulige områder for biblioteksmæssig udnyttelse af teknologien diskuteres i de afsluttende afsnit.

Projektets karakter og formål

Bibliotekskatalogens udvikling de sidste 10-15 år har først og fremmest været præget af web-OPAC'ens uomgængelighed. Katalogen er gået fra at være et redskab for bibliotekarerne til at blive et værktøj for slutbrugernes selvbetjening, hvilket stiller ændrede krav til dens indhold (informationer og funktioner) og dens præsentation. For en stor del vil brugerne hverken forvente eller ønske, at en bibliotekar formidler mellem deres behov og katalogens oplysninger, men til gengæld forventes bibliotekets ressourcer at være intuitivt operable for selvbetjening.

'Intuition' er i denne sammenhæng selvsagt ingen antropologisk konstans, men et udtryk for almen brugtrolighed med web-brug. Med andre ord er den brugertilpassede web-OPAC-udvikling stærkt præget af andre populære web-tjenester.

Tydeligst er dette i udviklingen af søgefunktionen, som i stadig flere tilfælde 'skal ligne Google', hvor med som regel menes ét søgefelt, fritekstsøgning og enkelhed (Jf. Markey, 2007). Bibliotekskatalogens traditionelle styrke – kombination af specifikke søgefelter – relegeres til en 'avanceret søgning'. Visse steder stikker 'googliceringen' dog dybere end til søgebilledets opbygning, som i den integrerede søgning, der udføres af bl.a. Summa (www.statsbiblioteket.dk). På en præ-indeksert basis søger Summa i mange forskellige datakilder – herunder det brede internet. Endvidere forsøger Summa at lave en relevansrangering ud fra søgetermernes match i bestemte, vægtede felter i de bibliografiske poster – en øvelse, der i skrivende stund knapt kan siges at leve op til Google's PageRank-algoritme, hverken konceptuelt eller resultatmæssigt.

Katalogudviklere har selvsagt øje for, at web er mere end Google og søgemaskiner. Den anden overordnede udviklingstrend går således under betegnelsen Bibliotek 2.0 – i forlængelse af begrebet Web 2.0, der omtales som det brugerdrevne web. Også her findes der forskellige grader af dybde, hvormed man tager konceptet til sig, fra søgeterm- og materialeforslag genereret fra tidligere anonyme søgninger

(Summa har det første; det andet er den Amazon-inspirerede: "lånere, der har lånt denne bog, har også lånt følgende bøger...") over bidrag i form af tags, boganmeldelser eller blogs fra brugerne (et af de mere interessante projekter er Lamson Library² til oprettelse af egentlige sociale fora, hvor man kan se andre brugeres profiler. Mig bekendt er der endnu ingen biblioteker, der har lagt hele katalogen ud som en Wiki, der tillader brugerne at rette og tilføje i den, men måske er det blot et tidsspørgsmål.

Disse to webnavigationsstrategier – den algoritmisk betingede term-matchning og den sociale/brugerdrevne vejledning – har altså allerede gjort sit indtog i biblioteksverdenen. En tredje generel web-udviklingstendens, er det såkaldte semantiske web³. Grundtanken i det semantiske web, som den blev præsenteret af Tim Bernes-Lee & al. (2001) var, at web-robotterne ikke blot skulle kunne matche termer, men også forstå (eller snarere registrere) deres mening. Forudsætningen for den noget magre "meningsforståelse" er termernes *kontekst*, dvs. at en term i *en* sammenhæng betyder noget bestemt og andet en den samme term i *en anden* sammenhæng. I den oprindelige udvikling af semantisk web-teknologi søges denne kontekstualisering foretaget ved hjælp af datamodellen Resource Description Framework (RDF) (Antinou & van Harmelen, 2004, ch.3). RDF er i dag en alternativ teknologi til den parallelt udviklede emnekortteknologi (Topic Maps), der blev anvendt i det her beskrevne projekt. Uanset valget af teknologisk løsning må den grundlæggende mekanisme bygge på en kortlægning af de mulige sammenhænge (og tilhørende betydninger) som termer kan indgå i, hvilket kræver et større arbejde med udarbejdelse af ontologier og – ikke mindst – kommunikation og interoperabilitet mellem disse. Praktiske problemer af denne art er givetvis en hovedårsag til at det semantiske web ikke har kunnet indfri Bernes-Lee's snart gamle visioner, men er blevet overhalet af bl.a. det brugerdrevne web. Jeg vil nedenfor kort komme ind på et af de mange problemer, der knytter sig til en ekstensiv ontologi-opbygning; her skal blot nævnes, at kravet om at opfatte termers betydning i form af entydig reference kun i meget specielle og begrænsede tilfælde lader sig indfri⁴.

Projektet ved Danmarks Kunstbibliotek (DKB) afprøvede emnekort som en teknologi, der kunne bruges til at reorganisere og præsentere bibliotekskatalogens indhold ud fra et humanistisk forskningsbib-

lioteks særlige vilkår. Indledningsvis blev der identificeret tre hovedindsatsområder:

1. *Materialeintegration*: Biblioteket havde på projekttidspunktet fem søgedatabaser, af forskellig karakter (en bibliografi og fire kataloger) og repræsenterende forskellige materialer (bøger og tidsskrifter, analyseposter, dias, arkitekturtegninger, personarkivalier)
2. *Emnevokabular*: det faglige domæne har generelt en løs terminologisk struktur og bibliotekets egne indekseringsemneord var ikke tilgængelige i en systematisk oversigt. I tillæg var en stor del af især de ældre poster mangelfuldt indekseret.
3. *Brugerdiversitet*: bibliotekets brugergruppe springer fra højt specialiserede forskere med stor domæne-fortrolighed over fagligt mindre afklarede studerende, til den alment kunstintresserede borger. Dette stiller særlige krav til præsentation af kataloginformation og andre biblioteksfaciliteter: Hvor den faglige novice kan have behov for at ledes ved hånden gennem f.eks. emnevokabularates forskellige betydningslag og få opstillet forskellige alternative muligheder, vil den slags semantiske omveje blot være et irritationsmoment for specialisten, der ved præcis, hvad hun vil have.

Studier af informationsadfærd har vist, at en stor del af biblioteksbrugerne bruger biblioteket primært til at *lokalisere* litteratur, som de er blevet gjort opmærksom via andre kanaler (Akselbo et al., 2006). Det er belæg for at tro, at *opdagelsen* af ny litteratur gennem bibliotekskatalogen spiller en gradvis mindre rolle jo større domænekendskab og ekspertise brugeren har om emnet (Markey, 2007), og måske især for humanisters vedkommende (jf. Lönnqvist, 2003). Den store Getty-studie, under ledelse af Marcia Bates (Bates, 1994, 1996; Bates et al., 1993, 1995; Siegfried et al., 1993) frembragte bl.a. den informationsvidenskabeligt set temmelig overraskende observation, at de humanistiske forskere kendte praktisk talt al litteratur om deres specialeområde. Det fik undersøgelsens forfattere til bl.a. at spekulere over værdien af traditionelle relevansmålinger, når humanisterne betragtede en emnesøgning som en stor succes, hvis et søgesæt på 200 poster indeholdt 1 eller 2 ukendte artikler (Bates, 1996). Samme studie viste imidlertid, at på andre fagområder end deres eget speciale, blev emnesøgning vurderet som meget mere værdifuldt – dvs. at man kan udlede, at for det humanistiske domæne spiller emnesøgning hoved-

sagligt en rolle for faglige novicer (uagtet at de er eksperter på et andet felt).

Informationssøgning forstået som *lokalisering/be-stilling* var udmærket dækket i bibliotekets eksisterende søgefunktion, forudsat at brugeren valgte rigtig søgebase (pkt.1 ovenfor). Projektet ville i stedet koncentrere sig om at udvikle et tilbud, der letter muligheden for at *opdage* emnerelevante biblioteksressourcer. Selv om dette ikke er den primære aktivitet for den største gruppe af biblioteksbrugere, er den et væsentligt tilbud til dem, der i pkt. 3 ovenfor betegnes som domæne-novicer (jf. Crawford, 2007). Det er også overvejende denne gruppe, der har behov for (og problemer med) at benytte den eksisterende emnesøgning. Følgelig kom det afgørende fokus i projektet at ligge på udviklingen af en lettere tilgængelig emneorienteret adgang til bibliotekets materialesamling.

Begrundelse for projektets udformning

Det blev altså vurderet at muligheden for at browse i bibliotekets ressourcer, var den funktionalitet, der bedst kunne give denne gruppe af brugere en alternativ indgang til materialerne. Begrundelsen herfor var:

1. Biblioteket havde oplevet stor tilfredshed fra netop denne gruppe efter at bøgerne var blevet opstillet efter emnegrupper i åbne hylder. Hyldebrowsing har da også vist sig at være en yndet informations-søgningsstrategi blandt humaniser (Tibbo, 1993, p.6).
2. Såvel almindelig erfaring som informations-forskning tilsiger, at det er lettere at genkende et emne, vi bliver præsenteret for, end at formulere de relevante søgeord for det (Koch & al. (u.å.), Markey, 2007).
3. Emnesøgning på biblioteket var besværliggjort af en historisk set meget mager og vekslende emneordsindeksering. I poster fra kortkatalogets tid blev der fx ikke anført generelle emneord, hvis der var tilknyttet et personnavn som emneord.

En kort uddybning af disse punkter kan forklare tanken bag projektets udformning af en TM-baseret browseadgang til bibliotekets ressourcer:

Ad 1

Selv om hyldebrowsing er populært, er det ikke særligt effektivt, idet materialets fysiske begrænsninger

indsnævrer det fulde informationspotentiale, som burde kunne udtrykkes ved grupperingsprincippet:

- a. En bog kan kun stå ét sted, og den stilles derfor hen i den gruppe, hvor dens emnemæssige tilhørighed er størst. På den måde forties alle andre forbindelser, som den kunne indgå i: andre emner den omhandler, andre bøger af samme forfatter eller udgivet af samme korporation, andre bøger i samme serie osv.
- b. På det konkrete bibliotek betød samlingernes heterogenitet videre at materialer om samme emne heller ikke rent fysisk kunne stå sammen – idet arkitekturtegninger, bøger (i forskellige formater), dias, fotos og personlige arkivalier ikke praktisk kan opbevares sammen.
- c. Hyldebrowsing er begrænset til hyldens faktiske indhold, og går dermed blik af udlånte, fejlplacerede og magasinopstillede ressourcer.

Ønsket var derfor at lade browsingen foregå elektronisk, således at de fysiske begrænsninger blev ophævet.

Ad 2

Humanioras løse emne Terminologi: Håndtering af emnevokabularer og fagterminologi er aldrig nemt, men det er en almindelig opfattelse i BDI-faget at de humanistiske fag byder på ganske særlige udfordringer. En af hovedaktørerne bag *Art and Architecture Thesaurus*, Pat Mollholt, har fx udtalt om det kunsthistoriske felt, at: "One faces imprecise terminology with a natural language full of the pitfalls of varying interpretations from author to author."⁵ Betragter vi humanioras brug af generelle emneord, er opfattelsen af vokabularet mangler på stabilitet og præcision ikke ubegrundet. Men situationen skyldes hverken intellektuel dovenskab eller noget særligt ønske fra humanisters side om at udtrykke sig upræcist. Opfattelsen fremtræder i sammenligning med teknik- og naturvidenskaber, der i højere grad kan nyde godt af definatorisk entydighed i deres vokabular, fordi denne strategi stemmer overens med deres traditionelle vidensideal, hvor der er blot et begrænset rum for fortolkning. Men hvis et entydigt vokabular søger at indskrænke fortolkningsrummet, er det næppe særligt velegnet i humaniora, hvis hovedaktivitet og vidensideal netop er fortolkning. Humaniora producerer viden ved at studere menneskelige produkter og handlinger (herunder andre fortolkninger) for at fortolke dem, dvs. finde ud af, hvad de betyder. Da studieobjekterne for en stor del består af ord, er det ikke

hensigtsmæssigt, at betydningen af disse ord er fastlagt i forvejen, f.eks. gennem rigide definitioner i et kontrolleret vokabular. Humanioras terminologiske tilstand må derfor ses som en iboende konsekvens netop af den særlige vidensform, der er humaniora, nemlig hermeneutikken. Det betyder dog ikke at ord ikke har nogen betydning i humaniora; men denne betydning fastlægges sjældent ved entydige definitioner, men søges i stedet frembragt gennem en nøje opbygget kontekst, der får de ønskede betydninger og fortolkninger til at træde frem. Denne kontekst-opbygning kræver ord, plads og tid, og er en af grundene til, at monografien står så stærkt i humaniora (jf. Cullars, 1985, Bates, 1996). Også som short-hand for den betydning, man tillægger en term, benyttes henvisning til et forfatterskab som betydnings-stabiliserende kontekst (som fx i 'ikke Chomskysk, men Wittgensteinsk grammatik').

Ad 3

Sammensætning af humanisters emnevokabular. Henvisningen til forfatterskaber, epoker og navne som betydningsramme påpeger samtidig behovet for at nuancere den almindelige BDI-opfattelse af det humanistiske emnevokabular. Dette har nemlig en sammensætning der betyder, at det for en stor dels vedkommende har en usædvanligt høj grad af præcision og entydighed. I forhold til andre fagområder består det humanistiske emnevokabular (undersøgt både som indekserings- og søgesprog) for en stor del af navne, der i kraft af relativt entydige referencer har en meget præcis betydning (Wiberley, 1983). Foreløbige undersøgelser tyder på at navne og værktitler udgør omkring halvdelen af emnevokabularet i humaniora, mod hhv. 0 og 5% i natur- og samfundsvidenskab (Bates & al., 1993, p. 15ff). Baggrunden for betydningen af navne (i bred betydning: som navne på personer, grupper, steder, institutioner, perioder) er ganske forståelig: humaniora beskæftiger sig med værker, handlinger og begivenheder fremstillet af mennesker – disse mennesker har navne, ligesom det, de har frembragt, ofte har det. Det humaniora beskæftiger sig med, er det humanistisk litteratur er om, og denne form for navne vil derfor være et naturligt valg af emneord for meget humanistiske dokumenter. Selv om en stor del af humanisters emnevokabular således har meget høj præcision, skaber denne situation i sig selv problemer for det samlede emnevokabular. For disse navne-emneord eksisterer side om side med de generelle emneord, der er den terminologi, der typisk finder indpas i thesauri

o.l. Det er en selvstændig udfordring, som det ses nedenfor, at skulle integrere disse to hovedtyper af emneord i en samlet organisering, primært fordi navne-emneordenes evne til at indgå i fx taksonomiske relationer (BT/NT) er omvendt proportional med deres præcision (periodebetegnelser som 'klassicisme' og 'romantik' kan have 'nyere tid' som BT, som igen kan have 'historisk tid' som BT, men 'Abildgaard' kan hverken have NT eller BT).

Emnekort som vidensorganiserende system

Som teknologi har emnekort deres styrke i organisering af store informationsmængder i associative netværk. Den grundlæggende organisering har fokus på *emner* (topics⁶), ikke termer: organiseringen foregår ved at specificere forbindelsen mellem emner, som på en eller anden måde er relaterede. Et eksempel kunne være, at jeg i dag skal til *lønforhandling*, hvilket er en handel om *prisen på arbejde*, der er teoretiseret af *Karl Marx*, der var *filosof* og blev født i *Trier*, der ligger i *Tyskland*, der blev samlet i 1871, osv. Her repræsenterer de kursiverede ord emner, som der så består forskellige typer af forbindelser mellem: en kalenderrelation mellem dags dato og lønforhandling, en forklaringsrelation mellem lønforhandling og prisen på arbejde, en <teoretiseret-af>-relation mellem prisen på arbejde og Marx, en indeholdelses-relation mellem Trier og Tyskland osv. At det er emnerne (og ikke termerne) der er forbundne betyder at de samme emner kan repræsenteres af forskellige termer – Tyskland kan udskiftes med Preussen, eller DDR/BRD, for bestemte perioder, eller med Germany eller Tyskland for andre sproglige præsentationer. Helt basalt består emnekort således af emner (f.x. Trier), der er forbundet med andre emner (fx Tyskland); hvor både emner og forbindelser kan være af bestemte typer (her fx 'by', <indeholdt i>-relation, 'land'). Disse typer er, som det ses, selv emner. Et emne (fx 'land') kan også have forskellige navne (stat, nation, rige, osv.), og instanser (fx Tyskland, Ungarn), der også selv kan være emner. Det betimelige spørgsmål er hvordan der holdes styr på de forskellige emner, når deres repræsentations- og udtryksmuligheder er så åbne. Kontrollen sikres ved at fastsætte specifikke, unikke emneidentiteter, som entydig reference. Når et emne således identificeres, vil man samtidig fastsætte hvad slags emnetype(r) det tilhører ('Trier' <er af typen> 'by'). Dermed er emnets forbindelsesmuligheder også defineret, idet disse typisk knytter sig til emne- og forbindelsestyper

('by' kan være <indeholdt i> et emne af typen 'land', men ikke omvendt; 'by' kan være <fødested for> emner af typen 'person', men ikke omvendt; 'by' skal bestå af mindst en 'bygning', osv.). At fastlægge sådanne krav og specifikationer for emne- og forbindelsestyper, kaldes at opbygge en ontologi – det aktuelle emnekorts grundlæggende skelet, der derefter befolkes med konkrete emner. Netop fordi "et emne kan være hvad som helst" (Garshol, 2004; Pepper, 2002) og forbindelsen af en hvilken som helst karakter, er det væsentligt at ontologien opbygges med tanke på to forhold: hvilke informationsressourcer har man at arbejde med, og hvad er emnekortets formål. Man kan ikke sige, at formålet står særligt klar frem i det ovenfor anførte eksempel med lønforhandlingen, og forbindelserne forekom da også temmelig tilfældige. Til gengæld havde det emnekort, der skulle udvikles til Danmarks Kunstbibliotek et tydeligt formål: en browsefunktion på samlingsniveau, der kunne integrere bibliotekets forskellige materialetyper og afhjælpe de problemer der knytter sig til (især novicers) emnesøgning i et kunsthistorisk domæne.

Hvad angår informationsressourcer var projektsituationen også forholdsvis klar: de informationer, der skulle indgå i emnekortet lå allerede organiseret i et betydningsstrukturerende format, nemlig MARC. På den måde vidste man på forhånd, at hvis der stod 'Trier' i et 600-felt, var der tale om en person, i et 633-felt en by. De forhåndsstrukturerede data udgjorde en stor lettelse for arbejde med at udvikle ontologien og befolke emnekortet – og da både MARC-posterne og emnekortet er lagret i XML-formater, vil der forholdsvis nemt kunne oversættes mellem dem for en automatisk datahøst. Det afgrænsede formål og de strukturerede data lettede også opgaven med at lave entydig reference (emneidentitet) for emnerne. Dette punkt er ellers det helt store aber med ethvert ontologi-baseret vidensorganiserende system, og derfor også med emnekort: der forudsættes en entydig reference mellem ord og betydning. Derfor ivrer emnekort-miljøet for at man skal identificere ens emner ved hjælp af offentlige emneindikatorer (Published Subject Indicator (PSI) – typisk i form af en URL), der ikke er relative til det enkelte emnekort. Håbet er bl.a., at emnekort udviklet med forskellige, specifikke formål vil kunne fusioneres og fungere sammen i kraft af, at de har et eller (helst) flere fælles emner (identificeret ved den samme PSI); ideelt set vil stadig større mængder af verdens viden kunne indfan-

ges gennem emnekort. At emneindikatoren er of-fentliggjort (dvs. eksisterer udenfor emnekortet selv) skal derfor sikre stabilitet til emnekortenes emner og forbindelser, og sikre at emnekort-fusionen også får de identiske emner (fra hvert kort) til at fusionere, hvorved de kommer til at fremstå med forbindelser fra begge emnekort. (Samtidig vil en PSI-URL ty-pisk henvise til en ressource, der forklarer afgørende træk ved emnet; en god PSI for emnet 'Karl Marx' vil således være en autoritativ, balanceret netside om ham.)

Prospektet lyder måske fascinerende, men det er ikke desto mindre tale om en gammel, falsk melodi, nemlig en version af den logiske positivismes sprogrens-ningsstrategi. Der består en betydelig analogi mel-lem de logiske positivisters idé om et altomfattende videnskabeligt system baseret på logiske slutninger fra entydige empiriske udsagn og emnekort-miljøets forestillinger om et ekspansivt netværk af asser-tive forbindelser mellem entydigt definerede emner. Analogiens omdrejningspunkt er det ontologiske krav om, at der skal være en entydigt definerende reference mellem sprogets elementer (fx ordene) og elementerne i det, sproget er om (fx tingene). Det er næppe nogen tvivl om, at en entydig, fælles forståel-se af hvad ord betyder, i de fleste tilfælde er noget vi *efterstræber* i forsøg på at undgå misforståelser, som generelt vurderes som uhensigtsmæssige. Ej heller er der tvivl om, at vi i mange sammenhænge *faktisk ev-ner* at blive enige om stabile definitioner af centrale begreber⁷. Men blot fordi vi i visse tilfælde kan give ostensive definitioner på et ord ved at udpege dets ikke-sproglige reference ("det dér er en motorcykel") er denne augustinske sprogopfattelse ikke gyldig som en generel forståelse af hvordan ord får betydning. Som først og fremmest Wittgenstein påpegede, er selv den ostensive definition kun meningsbærende i kraft af allerede eksisterende regler for sprogbrug, der foreskriver, at dette er en måde at mene på (Wittgenstein, 1958, p.169). Pointen i vor sammenhæng er ikke at ostensive definitioner ikke er gyldige, men at de ikke er de grundlæggende betydningsbærende elementer; tvært i mod er de et specialtilfælde af sprogbrug, og det er generelt i kraft af vor sprogbrug, at ordene får betydning (jf. Olesen, 2002). Men da sprogbrug altid er kontekstuel, kan ords betydning ikke deponeres i en ekstrasproglig reference og for-ventes at kunne indgå restløst i en anden kontekst. Overført til emnekortet betyder det, at der ingen grund er til at forvente, at nogle der i en sammen-

hæng (et emnekort) behandler et emne, tillægger det samme betydning som andre, der behandler emnet i en anden sammenhæng (et andet emnekort). Benytter de to emnekort samme PSI som emneindikator, vil en fusion nok kunne gennemføres, men ikke nødven-digvis blive meningsfuld. Reference er ikke lig med betydning, og denne indsigt er særligt central for ter-minologiske funktioner, der skal understøtte huma-nioras hermeneutiske vidensform.

Efter denne principielle kritik af de mere luftige am-bitioner der knytter sig til emnekortteknologien (og anden ontologi-baseret semantisk web-tænkning) kan det måske undre, at teknologien blev anvendt i pro-jektet. Men kritikken knytter sig til forestillingen om at fastlægge betydning via entydig reference på tværs af diskursive sammenhænge, hvorimod projektet der-imod udforsker emnekorts anvendelse inden for en diskurs, nemlig den kunsthistoriske. Endnu vigtigere er, at man på biblioteket netop stod i den specielle situation at termers betydning (inden for denne dis-kurs) i nogen grad faktisk kunne bindes op som enty-dige referencer, nemlig til specificerede MARC-fel-ter. Dette *var* altså et specialtilfælde – men selvsagt ingen tilfældighed: det var resultat af bibliotekets arbejde med at indeksere og katalogisere materialet, hvorved de relevante betydninger kodificeres. Pro-jektet ville viderebearbejde og synliggøre værdien af dette arbejde.

Dermed også sagt, at et emnekort henter en stor del af dets styrke i bevidstheden om de diskursive be-grænsninger, det skal fungere inden for. De afgø-rende parametre er igen tilgængelige informations-ressourcer og formålsrelevans. Ovennævnte (tænkte) eksempel førte os fra 'hvad jeg skal i dag' til 'Tysk-land blev samlet i 1871'; det er svært at forestille sig sammenhænge, hvor det ville være en relevant for-bindelsesrække. Idet et emne principielt kan være re-lateret til et hvilket som helst emne på utallige måder er det vigtigste i opbygningen af et godt emnekort at lave begrænsninger: hvilke forbindelsestyper må be-stemte emnetyper indgå i, hvilken rolle må de spille i forbindelserne, hvor mange og af hvilken art må de være, osv. Dette er hvad der fastlægges i ontologien, der imidlertid kan ændres undervejs, efterhånden som emnekortet (dets formål og informationsressour-cer) udvikler sig.

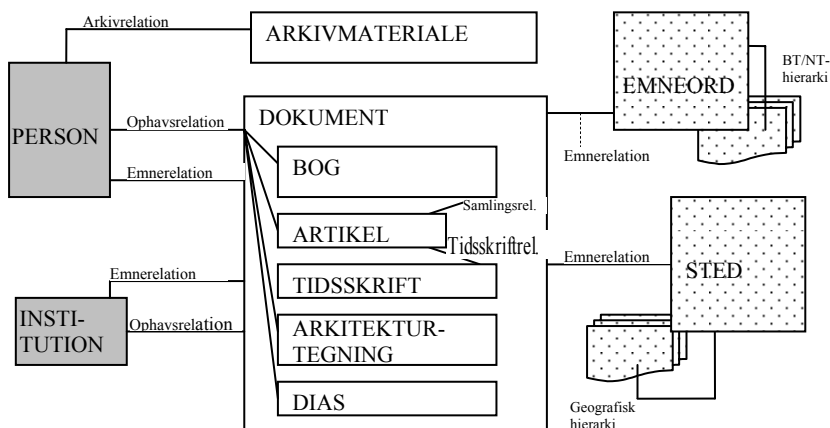
Obygning af en emnekortontologi

I vort tilfælde havde vi – netop i kraft af projektets formål og informationsressourcer – en ganske klar liste over grundlæggende emnetyper. De kan inddeles i tre hovedkategorier – det ene var de materialer, som brugerne skulle kunne opdage, dvs. emnetyper *dokument*, med undertyperne: arkitekturtegning, dias, arkivmateriale, bøger, tidsskrifter og artikler. Den anden kategori var *personer og institutioner*, der både kunne være ophav til og emne for dokumenterne – det betyder fx at emnetyperne 'person' kan indgå i forbindelsestyper som <er forfatter til> og <er emne for> med emnetyperne 'bog' og 'artikel'. (En fastlæggelse af hvilke *roller* i forbindelsen, de enkelte emnetyper kan spille, sikrer at forbindelserne set i modsat retning fremstår som <er skrevet af> og <omhandler>.) Endelig er der den tredje kategori af emnetyper, der kun kan indgå i emneforbindelser med dokumenterne, nemlig de *generelle emneord og steder* (bibliotekets praksis indebar ikke indeksering med periodebetegnelser som emneord). Principielt kunne man knyttet sted-emnetyper til også at kunne fungere som udgivelsested (MARC-felt 260 *a), men dette blev fravalgt fordi det blev vurderet som en usandsynlig association at ville gå fra en bog udgivet i Leipzig til en bog om Leipzig (hvorimod association fra en bog af Eckersberg til en bog om ham, er plausibel).

Opstillingen af relevante emnetyper er udtryk for 1) projektets formål om at bidrage til opdagelsen af dokumenter, ved at fremvise forbindelser mellem dem, 2) betydningen af navne som emneord i humanistisk litteratur, 3) behovet for at supplere den generelle emneordsliste. Informationsgrundlaget er reflekteret i, at de forskellige emnetyper henter em-

nernes navne i specifikke MARC-felter: dokumenter i 245-feltet, personer i 100, 600 og 700, institutioner i 610 og 710, emneord i 630 og steder i 633 (deres indbyrdes forbindelse til en post er sikret gennem postens sekventielfilsopmærkning).

Emnetyperne opfører sig forskelligt i deres relationer med emner af samme type: der kan opstilles hierarkier mellem generelle emneord (BT/NT-relationer) og mellem steder (indeholder/ligger i-relationer), som ikke giver mening for de øvrige emnetyper. (Oplysninger som 'Marion True' <ansat ved> 'Getty Museum', vil i mange andre emnekort være relevante, men ligger uden for dette bibliotekariske projekt.) Eneste undtagelse er dokument-undertypen 'artikel', der kan stå i en <indgår i>/<indeholder>-relation med emnetyperne 'bog' og 'tidsskrift' – et forbindelsesforhold, der administreres gennem indhøstning fra felt 558 og 557. Det blev besluttet; at udarbejde et selvstændigt hierarkisk emnekort over de generelle emneord ud fra bibliotekets kontrollerede emneordsliste, for derefter at fusionere dette med det emnekort, der indeholdt de MARC-høstede oplysninger. Ved visning af en bogs emneside; ville man således kunne se, hvilke emner den omhandler, og videre kunne se, hvilke nært tilknyttede emneord (BT/NT) der er til dette emne. Emneordshierarkiet blev dannet automatisk ud fra de UDK-klassifikationskoder, der var tilknyttet emneordene (således at smallere termer havde flere decimaler end bredere termer; termer i samme træ med samme antal decimaler blev tildelt forbindelsen 'tilstødende termer'). Et tilsvarende hierarki kunne laves for sted-emnerne, men projektets tidsplan tillod ikke realiseringen af dette på prototype-niveau. Idet de to emnekort skulle fusioneres, blev emneordene identificeret med PSIs med unikke URL-adresser.



Ontologien blev opbygget i programmet Ontopoly, der er en emnekort-editor med grafisk brugergrænseflade, hvor emnetyper, forbindelsestyper og andre grundelementer, kan oprettes, redigeres og kravsificeres. Ontopoly distribueres sammen med læse-/fremvisningsværktøjerne Omnigator og Vizigator i Ontopia Knowledge Suite (OKS) fra emnekort-selskabet Ontopia. Projektet benyttede OKS for samtidig at afprøve tilgængelige værktøjer for emnekortsopbygning, der ikke kræver specialviden, netop med tanke på at undersøge, om emnekort ville kunne anvendes i en bibliotekssammenhæng uden konstant (og omkostningstung) afhængighed af eksterne konsulenter. Et alternativ til at bruge en kommerciel⁸ slutbruger-orienteret løsning ville være at selv definere ontologiens elementer, enten i en særlig emnekortnotation som den Ontopia-udviklede LTM (Linear Topic Map Notation), eller i XML-formatet XTM, som også er det format, Ontopoly lagrer emnekort i. LTM er en betydeligt mere overskuelig syntaks end XTM, men indeholder ligesom de grafiske editorer begrænsninger i forhold til den fulde emnekortsfunktionalitet. XTM-syntaksen er mere omstændelig men ikke egentlig vanskelig; alligevel kan den for den almindelige bibliotekars anvendelse af teknologien udgøre en barriere, som en grafisk editor ville minimere. Det viste sig imidlertid at det i projektet var nødvendigt at skrive dele af ontologien i XTM for derefter at importere den til Ontopoly for videre redigering. Det drejede sig om emneordshierarkiet, der krævede at et emne havde mulighed for at være af to typer på samme tid; selv om der i emnekort-standarden ikke er noget i vejen for dette (og det dermed kan udtrykkes og valideres i XTM) går det ud over Ontopolys (og LTM's) begrænsninger. Ikke desto mindre er OKS-værktøjerne i stand til at operere med sådanne forhold i et importeret emnekort.

Afprøvning af en prototype

Emnekortene (dvs. katalogemnekortet, der omsætter MARC-posternes oplysninger til emner, og emneordshierarkiet) blev befolket i Ontopoly. I en reel implementering ville omsætningen skulle foregå gennem automatisk indhøstning og i praksis som en oversættelse fra en 'dialekt' af XML, nemlig databasernes MARCXML (Lee & al., 2006, p. 245f), til en anden, nemlig XTM (jf. Andersen, 2003, p.57). Transformationen til XTM fra en anden XML-dialekt vil samtidig være en konstruktion af et emnekort i XTM, og den foretages ved at anvende eXtensible

Stylesheet Language (XSL) til at definere de nærmere regler for denne transformation. Reglerne indskrives i hhv. en XSL Transformer og en XTM-procesor, i form af algoritmer ('hvis-så'; betingede processer med definerede enheder, variabler og værdier) (Librelotto & al., 2004, section 4; Andersen, 2003, p.58f). Udviklingen af denne procedure lå imidlertid udenfor projektets rammer, så i stedet blev prototypen sat op ved at indtaste oplysninger fra et antal poster, der blev fremfundet via en fritekstsøgning i bibliotekets baser. Fritekstsøgning på termen 'video' gav 179 dokumenter fordelt på alle fem baser; oplysningerne blev indtastet som emner og de korrekte forbindelser etableret. Derefter blev katalogemnekortet og emneordshierarkiet fusioneret.

På dette tidspunkt ville emnekortet indbefatte to typer information: emneordshierarkiets relationer og bibliografiske data, der allerede fremgik ved søgning i biblioteksbaserne. Eksempelvis ville emneordet 'installationer, kunst' have en emneside, der oplyser om dets UDK-klassifikation, hvilket emneordshierarki det indgår i, hvad som er dets bredere, smallere og tilstødende termer og hvilke dokumenter, der er tilknyttet emneordet. Ved at klikke på en af disse oplysninger kommer man videre til emnesiden til det pågældende emne, fx artiklen "Jesper Just: Midway Contemporary Art", der viser sig at indgå i tidskriftet Frieze, er udgivet i 2004, er skrevet af Max Andrews og udover 'installationer, kunst' handler om 'video' og 'Just, Jesper'. Alle disse oplysninger vil igen være klikbare og føre til nye emnesider. Derudover er der fra alle dokument-emnesider et direkte link til dokumentposten for bestilling fra biblioteket. Fra mange af person-emnesiderne er der direkte link til opslag om personen i Weilbachs kunstnerleksikon (som biblioteket i forvejen benytter til autoritetskontrol af danske kunstnernavne). Endelig fremvises forskellige noteoplysninger til dokumenterne, fx navn, sted og dato for den udstilling, ovennævnte artikel refererer til.

Emnekortteknologien åbner imidlertid for at man kan gøre mere med disse oplysninger, når man nu først har dem struktureret som et emnekort. Ud fra projektets formål om at støtte de faglige novicer, fandt jeg det nyttigt at prøve at udsige mere om emnerne end det, der ligger i den enkelte MARC-post. Jeg ønskede at opstille en profil for et emne A ved at trække oplysninger ind, der knytter sig til de emner, som A er relateret til på en bestemt måde. Princippet kendes

fra bl.a. co-citationsanalyser, men anvendes her på emner i et emnekort, hvor oplysninger hentes ud og præsenteres ved hjælp af *tolog*, et særligt søgesprog til emnekort (Garshol, 2006). Der blev etableret to typer af disse indirekte forbindelser: en ophavsprofil fremviser hvad en person har skrevet om, ved at hente alle emner, der er knyttet i en emnerelation til de dokumenter, som personen har en ophavsrelation til. En emneprofil fortæller hvilke emner, der gennem dokumenterne er knyttet til et hvilket som helst emne i emnerelationer, dvs. hvad fx emnet 'installationer, kunst' behandles sammen med. Disse *indirekte* forbindelser baserer sig altså på de bibliografiske posters oplysninger, men udnytter dem til at synliggøre nye forbindelser mellem de enkelte oplysninger som direkte kun har været knyttede til den enkelte post.

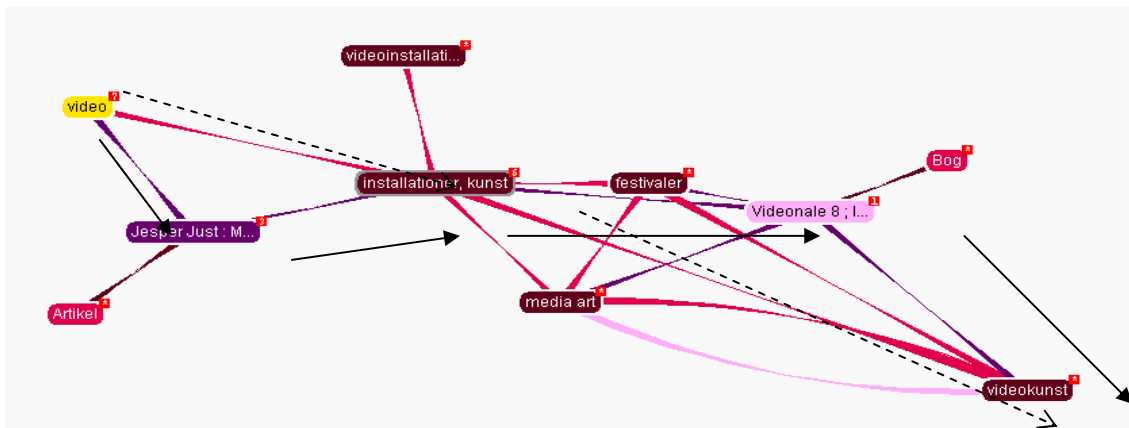
Resultater

En gennemgang af prototypen⁹ viste, at den grundlæggende funktionalitet var ganske god og opfyldte projektets formål, nemlig at

- integrere bibliotekets materialer i en søgeflade,
- skabe en browseadgang især med tanke på faglige novicer,
- samt synliggøre og videreudvikle resultatet af bibliotekets metadata-arbejde.

Projektets prioritering af områdets særlige emneordsproblematik viste sig særligt tilgodeset af løsningen. For det første betød selve opbygningen af et browsebart emneordshierarki, at emneordenes systematik blev synlig for brugerne på en måde som den hidtil eksisterende alfabetiske emneordsliste ikke tillod.

Den systematiske ordning har selvsagt ikke i sig selv noget med emnekortteknologien at gøre, men ved at udarbejde systematikken som et emnekort kunne den nemt fusioneres med det andet emnekort, der indeholdt katalogposternes oplysninger. Et emneord er dermed samtidig forbundet med andre emneord i en hierarkisk ordning, med dokumenter (i kraft af indekseringsarbejdet) og med andre emner (både emneord, personer, institutioner og steder) i kraft af den indirekte forbindelse emneprofil. Alle disse oplysninger udgør en kontekstualisering af det enkelte emneord (hvilke er dets bredere/smallerer termer, hvilke titler er det knyttet til, hvilke emner behandles det ofte sammen med), der kan bidrage til at stabilisere dets mening. For det andet viste det sig, at der i katalogposterne var brugt mange emneord, der ikke indgik i emneordslisten, og altså ikke var kontrollerede (det gjaldt fx det benyttede søgeord 'video'). Netop for disse ukontrollerede emne- eller søgeord (browseren Omnigator har også en søgefunktion) er det en meget god funktionalitet, at kunne ledes videre gennem dokumenter og andre emner hen til den kontrollerede form af emneordet, der typisk vil være tilknyttet flere dokumenter, foruden at det vil indgå i den systematiske struktur, måske have en tilknyttet forklaring osv. Figuren nedenfor beskriver hvordan en sådan 'opdagelsesfærd' kan foregå fra 'video' til den kontrollerede term 'videokunst'. De heltrukne pile viser vejen via direkte forbindelser (der principielt også kan gennemføres via videresøgning på web-OPAC), mens de stiplede pile angiver vejen via indirekte forbindelser (i dette tilfælde emneprofilen; <behandles sammen med>).



For det tredje tager emnekortet højde for det særlige humanistiske forhold at bl.a. personer og institutioner lige så selvfølgelig som generelle emneord er emner for dokumenter, men at de derudover også kan være ophav til dokumenter. I de bibliografiske poster repræsenteres dette ved brug af forskellige MARC-felter (6xx-serien respektive 1xx/7xx-serien), men i emnekortet får man synliggjort at det drejer sig om samme person (vedkommende har en emneside) der indgår i forskellige relationer (emnerelation, ophavsrelation) til dokumenterne. På den måde respekterer emnekortet humanistisk litteraturs tendens til at være centreret omkring navne.

Prototypen afslørede da også store mangler, der skal udbedres, inden emnekortet kan fremstå som den efterspurgte løsning. Der skal lægges betydeligt arbejde i udvikling af et hensigtsmæssig grænsefladedesign, der sandsynligvis bør ligge i forlængelse af bibliotekets øvrige web-design. Dette område indgik slet ikke i projektet, som i stedet undersøgte emnekortets overordnede funktionalitet. Men også på funktionsområdet er der betydeligt rum for forbedringer: der skal signaleres klarere hvilken emnetype et relateret emne tilhører, der bør indføres en vægtning eller rangordning (og tilhørende bagatelgrænse) for at undgå alt for lange lister over relaterede emner, der bør udvikles en funktion, der tillader mere sømløs veksling mellem søgninger i bibliotekets almindelige katalog og browsing i emnekortet. I emnekortet savnes desuden en funktion, så man kan benytte et link til et relateret emne *enten* til at gå til dette emnes emneside (som i dag) *eller* til at afgrænse visningen af emner (fx vis de emner, hvor 'Lars Morell' <har skrevet om> 'videokunst'). Denne sidste mulighed vil kunne udvikles via et søgesprog som tolog. En yderligere udviklingsmulighed ville være at inddrage data og informationsressourcer der går ud over de MARC-baserede oplysninger – fx Bibliotek 2.0-elementer som tags, brugerprofiler, udlånskoordinater o.l. Emnekort-modellen kan således med fordel kombineres med web-OPAC'ens øvrige udviklingstendenser. Netop for at sikre anvendeligheden i forhold til alment udbredt web-forståelse, bør reelle brugbarhedstest indgå i et evt. videre udviklingsarbejde.

Konklusioner

Projektet viste altså, at man ved at omsætte bibliotekets bibliografiske katalogoplysninger til emner i et

emnekort, ville kunne udvikle en browsefunktionelitet, som plausibelt vil medføre en betydelig lettelse af den emnebaserede materialesøgning. Ikke desto mindre rejser projektet spørgsmål om hvor og hvordan en sådan tilgang mest sandsynlig ville kunne komme til anvendelse. Det er især to anfægtelser, der knytter sig til det beskrevne projekt om en relativt nem og billig opbygget ontologi, befolket ved hjælp automatisk indhøstning af MARC-oplysninger og oversættelse til XTM. Den første anfægtelse knytter sig til det velkendte problem med automatiseret indhøstning, nemlig at resultatet er meget afhængigt af datagrundlagets kvalitet. Det kan derfor kun anbefales, at man holder hånd i hanke med oversættelsen, i hvert fald til at begynde med, for at kunne fange og evt. justere for uhensigtsmæssigheder. Dette taler for at omsætningen fra katalog til emnekort bør foregå i mindre portioner end projektets forestilling om at hele bibliotekskatalogen skal gøres browsebar. Den anden anfægtelse fører til samme konklusion, men på grundlag af skaleringsforholdet mellem poster og emner. Prototypen blev opbygget på grundlag af 179 dokumenter, der resulterede i 535 emner, plus 328 emneordsemner, der ikke vil mangfoldiggøres i samme grad, da de (ud fra en kontrolleret liste) vil gå igen i mange af posterne. Men hvis dette skaleringsforhold skulle holde for alle bibliotekets ca. 380 000 poster ville resultatet være et emnekort med omkring 1,2 millioner emner (plus de 1800 emneordsemner). Det er næppe sandsynligt, at et bibliotek vil binde an med administration af et emnekort af den størrelse i et hug.

Til gengæld kan det realiseres ved, at man tager egnede delmængder af biblioteksbase og opbygger relevante emnekort, der så senere kan fusioneres sammen, hvorved hele materiale-samlingen efterhånden bliver dækket. Samtidig vil man undervejs kunne afprøve og justere brugervenligheden. Af hensyn til realismen er det desuden afgørende, at emnekortene kan opbygges som led i forskellige selvstændige projekter, der har et specifikt tematisk fokus – fx som understøttelse til udstillinger, projektarbejder, undervisningsforløb eller institutionelle samlinger – og derefter samles og dermed både binde projekt-tematikkerne og biblioteksmaterialerne bedre sammen i en synlig, associativ relationsstruktur.

Noter

1. Se fx www.kulturnett.no og www.vimu.info. Topic Maps (TM) er en international standard med betegnelsen ISO 13250; i artiklen benyttes den danske oversættelse 'emnekort'.
2. Katalogen, som er integreret i et WordPress-program, er under udvikling på www.plymouth.edu/library/opac
3. For en generel introduktion til fænomenet henvises til Antinou & van Harmelen (2004).
4. Den succes som fx en ontologi som GeneOntology har kunnet opvise skyldes netop dets begrænsede område og helt specifikke formål, jf. Bada et al. (2004) og Bard (2003).
5. Citeret, blandt andre, i Wiberley (1983).
6. En terminologisk oversættelse af emnekortbegreber kan findes i Bøckman (2006), bilag 1.
7. På dette punkt slipper TM-miljøet – med dets PSI som en faktisk accepteret sproglig-social autoritet – lettere end de logiske positivister, hvis fundamentale empiriske primærudsagn viste sig at indeholde sprogfilosofiske selvmodsigelser (jf. Sellars, 1997, ch.8).
8. OKS er tilgængelig i en betalingsfri demo-version fra Ontopias netsider, www.ontopia.net. En Open-Source-løsning distribueres gennem www.tm4j.org
9. En udførlig gennemgang kan findes i Bøckman (2006), afsnit 7.2 og bilag 13.

Litteratur

Akselbo et al. (2006). *Det hybride bibliotek set med brugerens øjne*. (DEFF-rapport) Århus: Statsbiblioteket.

Andersen, Anett M. K. (2003). *Topic maps og Single sourcing. Automatisk konstruktion for single source modeller*. Cand.it-speciale. Kolding: Syddansk Universitet.

Antinou, G. & van Harmelen, F. (2004). *A Semantic Web Primer*. Cambridge, Mass.: MIT Press.

Bada et al. (2004). A Short Study on the Success of the Gene Ontology. *Journal of web semantics, vol. 1, no.2*, pp. 234-240.

Bard, Jonathan (2003). Ontologies: formalising biological knowledge for bioinformatics. *BioEssays, vol. 25, no. 5*, pp. 501-506.

Bates, Marcia J. (1994). The Design of Databases and Other Information Resources for Humanities Scholars: The Getty Online Searching Project Report No. 4. *Online & CDROM Review, vol. 18, no. 6*, pp. 331-340.

Bates, Marcia J. (1996). Document Familiarity, Relevance, and Bradford's Law: The Getty Online Searching Project Report No. 5. *Information Processing & Management, vol. 32, no. 6*, pp. 697-707.

Bates, Marcia J. & al. (1993). An Analysis of Search Terminology Used by Humanities Scholars: The Getty Online Searching Project Report Number 1. *Library Quarterly, vol. 63, no. 1*, pp.1-39.

Bates, Marcia J. & al. (1995). Research Practices of Humanities Scholars in an Online Environment: The Getty Online Searching Project Report No. 3. *Library and Information Science Research 17*, pp. 5-40.

Berners-Lee, T., Hendler, J. & Lassila, O. (2001). The Semantic Web. *Scientific American, vol. 284, no. 3*, pp. 29-37.

Bøckman, Knut Anton (2006). Topic Maps som browsemodel for humanistiske bibliotekers informationsformidling. Med særligt henblik på humanisters emnevokabular. Cand.scient.bibl.-speciale, Danmarks Biblioteksskole, København.

Crawford, Walt (2007). Patrons and the Library. *Cites & Insights, vol. 7, no.1, January 2007*, pp. 11-18.

Cullars, John (1985). Characteristics of the Monographic Literature of British and American Literary Studies. *College & Research Libraries, vol. 46, no. 6*, pp. 511-522.

- Garshol, Lars M. (2004). Metadata? Thesauri? Taxonomies? Topic Maps! Making sense of it all. *Journal of Information Science*, vol. 30, no. 4, pp. 378-391.
- Garshol, Lars M (2006). tolog – A Topic Maps Query Language. In: Maicher & Park (eds.): *Charting the Topic Maps Research and Applications Landscape. First International Workshop on Topic Maps Research and Applications, TMRA 2005; Leipzig, Germany, October 2005. Revised Selected Papers*. Berlin: Springer Verlag.
- Koch, Traugott & al. (u.å.). Provide browsing using classification schemes. Lokaliseret 20.05.2006 på: <http://www.lub.lu.se/desire/handbook/class.html>
- Lee, Hyun-Sil & al. (2006). MARCXTM: Topic Maps Modeling of MARC Bibliographic Information. In: Maicher & Park (eds.): *Charting the Topic Maps Research and Applications Landscape. First International Workshop on Topic Maps Research and Applications, TMRA 2005; Leipzig, Germany, October 2005. Revised Selected Papers*. Berlin: Springer Verlag.
- Librelotto, Giovanni Rubert & al. (2004). TM-Builder: An Ontology Builder based on XML Topic Maps. *CLEI Electronic Journal*, vol. 7, no. 2, paper 4.
- Lönnqvist, Harriet (2003). *Humanistiska forskningsprocesser och informationssökare*. Tammerfors: Tampere University Press.
- Markey, Karen (2007). The Online Library Catalog; Paradise Lost and Paradise Regained? *D-Lib Magazine*, vol. 13, no.13, January/February 2007.
- Pepper, Steve (2002). The TAO of Topic Maps. Finding the Way in the Age of Infoglut. Lokaliseret 20.05.2006 på <http://www.ontopia.net/topicmaps/materials/tao.html>
- Olesen, Charlotte (2002). Definitionsteori. En filosofisk og terminologisk tilgang til definitionsteoretiske problemstillinger. *Hermes. Journal of Linguistics*, 28, pp. 155-177.
- Sellars, Wilfred (1997). *Empiricism and the Philosophy of Mind*. Cambridge, Mass.: Harvard UP.
- Siegfried, Susan & al. (1993). A Profile of End-User Searching Behavior by Humanities Scholars: The Getty Online Searching Project Report No. 2. *Journal of the American Society of Information Science*, vol. 44, no.5, pp. 273-291.
- Tibbo, Helen R. (1993). *Abstracting, Information Retrieval and the Humanities. Providing Access to Historical Literature*. Chicago, Ill.: American Library Association.
- Wiberley, Stephen E., Jr. (1983). Subject Access in the Humanities and the Precision of the Humanist's Vocabulary. *Library Quarterly*, vol. 53, no. 4, pp. 420-433.
- Wittgenstein, Ludwig (1958). *Philosophical Investigations*. 2nd edition. Oxford: Basil Blackwell.