

NORDTERM 16

Ontologier og taksonomier

REDAKTØRER: BODIL NISTRUP MADSEN & HANNE ERDMAN THOMSEN

NORD TERM

NORDTERM 2009
København 9.-12. juni 2009



Personaliseret information

Emnesystematik bag kommunikation af videnskabelig information på Internettet

Jørgen Burchardt

seniorforsker, Nyt om Arbejdsliv, Nyborgvej 13, 5750 Ringe www.burchardt.name.

Abstract. Med relativt små midler kan en emnesystematik udvikles til brug hos en videnskabelig nyhedstjeneste. Med eksempel fra Nyt om Arbejdsliv fortælles hvorledes man "quick and dirty" kan udarbejde en emnesystematik. Den er blevet en central del af hjemmesidens systemarkitektur, hvorved mere end 100.000 personer hvert år finder oplysninger blandt nettets mange milliarder sider. En systematik er samtidig nødvendig, idet folk ikke interesserer sig for et fag, men oftest kun for et afgrænset emne.

HVORFOR PERSONALISERET INFORMATION

Mængden af viden er de seneste årtier nærmest eksploderet. Digitaliserede arkiver er blevet fyldt med information, og Internettet har gjort dette tilgængeligt på en fantastisk effektiv måde. Søgmaskiner kan give adgang til dette næsten uendeligt store materiale. Med få tastetryk står man med links til det materiale man ønsker. Google har i skrivende stund efterår 2009 indekseret mere end 17 milliarder sider.

Som bruger kan det imidlertid godt være vanskeligt at få adgang til den ønskede information. Det var lettere i 1472, hvor størstedelen af verdens bøger fandtes på et bibliotek som Queens College med dets i alt 199 bøger. Der er nemlig et meget stort problem: Der er så store mængder, at søgeresultaterne oftest er så omfattende med så mange henvisninger, at det er blevet vanskeligt og tidskrævende at finde de rigtige oplysninger. "Information overload" hedder fænomenet, hvor man bliver overbelastet af for mange oplysninger, så man ikke kan finde den ønskede nål i høstakken.

Dette problem vil naturligt vokse, idet mængderne af information fortsat vil vokse.

Problemet eksisterer ikke alene på hele nettet. Selv på hjemmesider fra informationsudbydere er mængden af information vokset således at det bliver umuligt at overskue alt materialet. På vores nyhedstjeneste Nyt om Arbejdsliv har vi mere end 8.000 sider. Kun en målrettet søgning kan forhindre spild af tid.

Personaliseret information er løsningen på dette "information overload" problem. Der skal kun bringes de informationer, som jeg er interesseret i, og ikke alle de øvrige, som ville spille min tid.

Denne måde at organisere og vise oplysninger på er ikke ny. Kunderne hos Amazon.com kender til de anbefalede bøger, som firmaet anbefaler ud fra lighed med de tidligere købte bøger. Datamining har længe været benyttet af store ressourcerige firmaer.

Denne arbejds måde kan også benyttes af fattigere organisationer, og informationerne kan ordnes på en endnu mere kvalificeret måde, end det er tilfældet hos Amazon.com.

Emnesystematik er et vigtigt redskab i opbygningen af den interne infrastruktur, hvor det ønskes, at brugerne vælger lige præcis de emner, de er interesserede i, og som de til gengæld kan blive optimalt informeret om.

ERFARINGSMATERIALET

Jeg vil fortælle om erfaringerne fra et tværvideenskabeligt forskningsområde om arbejdsliv, arbejdsmarked og arbejdsmiljø. Mere end 500 danske forskere er aktive på feltet både på universiteter og private forskningsinstitutioner, og deres forskning bliver i stor udstrækning gjort tilgængelig for samfundet – lægmænd som andre forskere – gennem Internetportalen "Nyt om Arbejdsliv". En del af de samme oplysninger findes på engelsk "Working Life Info" i en automatisk "spejlet" udgave.

En emnesystematik på dansk og engelsk med en specifik udarbejdet liste over emneord er central i den etablerede IT-plattform. Alle videnskabelige produkter – artikler, anmeldelser og bøger – bliver kategoriseret sammen med tilbud om møder og konferencer. Hvert år er der omkring 168.000 såkaldte unikke besøgende, hvilket i realiteten betyder, at ca. 100.000 forskellige personer benytter denne tjeneste gennem deres mere end 250.000 besøg.

Vi startede op i 2003 og kan således give erfaring fra mere end 6 år videre.

Brugernes ønsker – et emne

Det er brugernes ønsker, som vi må tage udgangspunkt i, når vi designer en nyhedstjeneste på nettet. Vi har naturligvis vores egne ønsker og idéer, men når vi kommer til den endelige udformning, er det brugerne, vi i stor udstrækning skal tage hensyn til.

Det var et forhold, vi meget hurtigt blev klar over, da vi havde lavet de første udgaver af vores hjemmeside og efterfølgende så på hvorledes brugerne i virkeligheden udnyttede de muligheder, vi havde stillet til rådighed for dem. Det blev vist gennem alle vores undersøgelser via interviews og gruppesamtaler med brugere samt ved analyse af logfiler.

Vi må konstatere, at brugerne i stor udstrækning opførte sig helt anderledes, end vi havde forventet. Vores udgangspunkt er en forening for alle forskere inden for området arbejdsliv, som dækker alle områder af "arbejdsliv". Det viste sig hurtigt, at brugerne ikke interesserede sig for faget som sådan, men kun for et eller få emner.

Det blev derfor naturligt for os, at tage udgangspunkt i brugernes behov, nemlig det at de ville vide noget om et begrænset emne, mens de øvrige emner på det nærmeste var irrelevante. Vores nyhedstjeneste skulle derfor gøre det muligt at levere personaliseret information.

Emnefeltet arbejdsliv

Vores område er "arbejdsliv". Det er faktisk ikke nærmere defineret, men er blevet afgrænset af de aktiviteter, som vores organisation Center for Studier i Arbejdsliv står for.

Forskningsfeltet er tværfagligt, idet der er aktive forskere fra så forskellige fagområder som sociologi, psykologi, medicin, etnografi, økonomi, politologi, jura, historie samt en del yderligere fagområder. Fagets emner er således ikke udviklet inden for en enkelt videnskab, men benyttes af en række forskellige forskere fra forskellige fag. Dette gør tillige, at der ikke findes en 100 % holdbar systematik for forskerne fra alle forskningsfelter.

Af praktiske årsager har vi ladet aktiviteterne i Tidsskrift for Arbejdsliv definere områdets omfang og indhold.

Tidsskriftet havde eksisteret siden 1999, og der var på det tidspunkt tale om omkring 100 artikler i tidsskriftet. Hertil kom omkring 1.000 bøger, som tidsskriftets anmelderredaktion havde udvalgt til at være relevante for anmeldelse i tidsskriftet.

Dette har været en pragmatisk måde til at definere feltet på. Metoden vil derfor kunne bruges til fortsat at udvikle emneordlisten i takt med at området fagligt ændrer sig.

Denne quick and dirty måde at definere emneområdet på har selvfølgelig en stor fordel ved at kunne etableres på meget kort tid. Svagheden er, at den er afhængig af relativt få menneskers beslutninger, nemlig redaktionens optagelse af artikler og en enkelt anmelderredaktørs beslutning om hvilke bøger, der skal medtages på tidsskriftets anmeldelsesliste.

Disse begrænsninger er dog ikke særlig alvorlige, da det netop har drejet sig om at få indekseret bidragene til tidsskriftet samt de løbende boglister. Hvis der løbende bliver justeret, vil emneordlisten til stadighed være ajour i forhold til det aktuelle behov.



FIGUR 1. Denne hjemmeside er en nyhedstjeneste, som Center for Studier i Arbejdsliv udgiver med den nyeste forskning fra områdets mere end 500 forskere. Artikler og anmeldelser fra Tidsskrift for Arbejdsliv er unikt materiale, der suppleres med nyheder om bøger og konferencer. Mere end 1.700 modtager CSA's nyhedsmail, og årligt besøges siden mere end 250.000 gange.

Emneordslistens opbygning

Da vi startede arbejdet, havde vi ingen erfaringer på området. Vi kunne se, at det ville være ønskeligt med en systematik på flere hundrede emneord, således at mængden ved søgning via et emneord kunne reducere resultatet til en delmængde på i gennemsnit en hundrededel eller mindre.

I første omgang blev det undersøgt, om der i forvejen fandtes et bestående system, som blot kunne overtages evt. efter en simpel revision. Det viste sig desværre ikke at være tilfældet.

De første systemer var systemerne fra de danske folkebiblioteker og et officielt system fra IT- og Telestyrelsen. De viste sig ikke at kunne opfylde vores krav til detaljeringsgrad. Samtidig havde de begge en form for systematik, som i hvert fald for vores system ikke ville fungere optimalt. Bibliotekernes system kunne således ikke bruges i sin helhed, men de fundne ord blev lagt i en database for efterfølgende behandling. Det officielle danske IT-væsen havde selv et ønske om at benytte

systematik til sin sagsbehandling og brug af automatisk journalisering. Vi så nærmere på systemet, men måtte konstatere, at det var alt for simpelt.

Større held var der med bibliotekssystemet hos det nu nedlagte Arbejdslivsinstitutets bibliotek, som dengang var Europas mest omfattende forskningsinstitution på området. Der var ikke tale om en systematik, men mange gode emneord blev samlet sammen og supplerede den oparbejdede database.

Jagten fortsatte, og mange andre udenlandske systemer blev gennemgået. Et af de mest gennemarbejdede systemer var det finske Työterveyslaitos. Det havde sin begrænsning i kun at medtage en del af det emneområde, vi gerne ville dække. Til gengæld var det på mange måder et optimalt udført arbejde. Mange ord var defineret, og der fandtes samtidig et tilsvarende engelsk begreb i mange tilfælde.¹

Arbejde udført fra grunden

Efter det forberedende arbejde havde vi opstillet en længere liste af ord. Vi vidste dog, at hele emneområdet ikke var dækket. Nu kom vi til det tidskrævende arbejde at få listen suppleret med de manglende ord.

Her valgte vi at få suppleret ved at gennemgå 1.000 bøger, artikler og konferencer. Hver enkelt blev vurderet for hvad man kunne kategorisere dem at omhandle.

Af praktiske årsager var det kun muligt at kategorisere ud fra titel, undertitel samt for bøger og tidsskriftsartiklers vedkommende fra det korte resumé leveret af forfatter eller udgiver.

Dette arbejde gav mange nye begreber, som ikke havde været omfattet i de hidtidige emneordslister og systematikker. Der opstod dog samtidig et nyt problem, der skulle afklares, nemlig at mange ord dækkede et for detaljeret begreb, at det ikke ville være praktisk at benytte det i en systematik. Vi skulle ikke udarbejde en udførlig fagterm, men kun vælge de ca. 20 % mest brugbare og dækkende termer. I en artikel om arbejdsmiljø, kunne der til eksempel være omtalt begrebet ”sikkerhedsudvalgsmøde”. Det dækker over et reelt begreb, men det ville nok give en for detaljeret ordliste, hvis alle disse ord skulle medtages. ”Sikkerhedsarbejde” og ”sikkerhedsudvalg” ville nok være begreber, der dækker møderne. Der var mange af denne type afgørelser, der skulle træffes. Af hensyn til de få ressourcer i projektet, måtte der hurtigt træffes en beslutning, hvilket skete uden større argumentering men blot ud fra en fornemmelse af, at dette eller hint begreb nok blev for specielt.

¹ Vi har bl.a. fået inspiration fra Mannheimer Zentrum für Europäische Sozialforschung, Library of Congress, Outline of Cultural Material, Universal Decimalklasse Klassifikation, Arbejdslivsinstitutet, AMI, Svenska ämnesord, UNESCO Thesaurus ULCC, EUROVOC thesaurus, Cambridge Scientific Abstracts, Sociology Thesaurus, Työterveyslaitos (Työsuojelusanasto).



FIGUR 2. Emnet "arbejdsliv" består af begreber fra en lang række fag. Det er derfor kun folk med føling i alle disse områder, der kan bedømme termer.

HVEM KAN UDARBEJDE EN SYSTEMATIK?

Vores erfaringer fra et fattigt forskningsfelt viser, at det kan lade sig gøre at etablere en fungerende emneordliste med relativt få midler. Hvis bare metoderne er gennemtænkte, kan tingene udformes rimelig let.

I praksis har der været få mennesker involveret i arbejdet. Det har ikke været nødvendigt at involvere mange mennesker i den første forberedende fase. Det har dog været et udgangspunkt, at mindst én person med et kendskab til hele forskningsfeltet har skullet deltage. Det praktiske arbejde omkring programmering og tilretning kunne udføres af personer uden for fagmiljøet.

Det har i mange forbindelser været ønsket, at få tilbagemelding fra forskningsmiljøet. Sådanne tilbagemeldinger har det ikke været muligt at få i særligt omfattende eller gennemgribende omfang. Derimod er der modtaget mange enkeltstående kommentarer om fortrinsvis enkelte begrebs brugbarhed.

At benytte en systematik

Da vi startede op i 2003 ønskede vi ikke alene at have en emneordliste. Listen skulle samtidig være systematisk. Det var et optimalt ønske, idet man herved kunne give brugerne mulighed for at søge på et meget snævert emne, men kunne samtidig give en alternativ mulighed for at vælge et større emne. Til eksempel kan man vælge emnet "lærlingeuddannelse", men man kan også vælge det mere omfattende "uddannelse", som omfatter mange andre former for uddannelse.

Det var et fint ideal. Virkeligheden viste sig at være mere problematisk. Vi startede med en usystematisk liste på 400 ord, og øvelsen var så at gruppere dem i et system. I

lighed med selve emneordene, fandtes der ingen systematik at bygge videre på. Det hele skulle udvikles fra grunden.

Mange ord var lette at organisere. Størstedelen af ordene faldt nærmest af sig selv ind i en systematik. Arbejdet bestod ”blot” i at sætte etiketter på et hierarki, som man kunne forestille sig kun ventede på at dette meget indlysende arbejde skulle blive foretaget.

Der var desværre en relativ stor gruppe ord tilbage, som ikke egentlig havde nogen systematik at høre til eller – hvilket var værre – kunne høre til to eller flere andre dele af den udarbejdede systematik.

Det første kunne godt løse sig ved at lave nogle overbegreber, som blot dækkede disse emner. Det andet problem var vanskeligere – for ikke at sige umuligt - at løse, og det måtte man lære at leve med, at der var denne synlige svaghed.

En systemarkitektur for personaliseret information

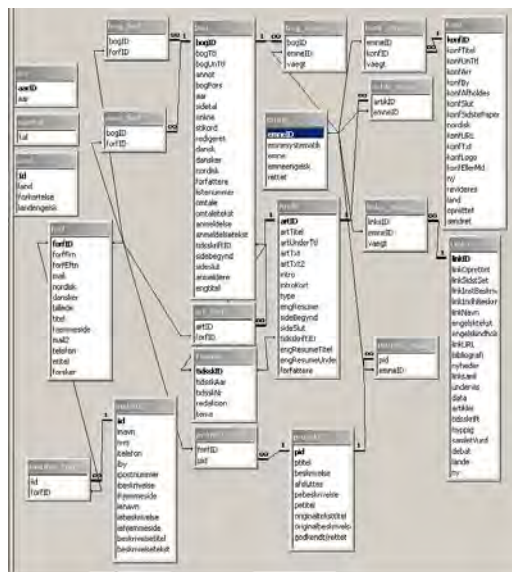
I sommeren 2003 var arbejdet med systematikken afsluttet. Der var lavet en systematik med knap 400 emneord, som var blevet inddelt i et mindre antal overordnede begreber. Nu kunne systemet for alvor blive indført.

Der blev valgt at arbejde med en relationelt opbygget database. Ønsket var at starte med at arbejde i den simple og billige database Access, som de fleste i forvejen har liggende på deres computer. Det var meningen senere at skalere til et større databasesystem. Der viste sig da også nogle kapacitetsproblemer med Access, da systemet kom i drift. Hvis der var for mange samtidig brugere, kunne systemet blokere, og kun en genstart af serveren kunne få systemet i gang igen. Det var et problem, som vi skulle identificere, og som vi efterfølgende skulle bede web-hotellet om at løse. Inden vi havde opdaget, at der var en fejl, kunne der gå dage, og hvis det skete ved en weekend, kunne der gå yderligere en del timer mens hjemmesiden ikke var tilgængelig.

En omlægning ville tage ret lang tid, og da vi endelig havde besluttet at starte at omlægge til det forbedrede databasesystem, var der imidlertid ingen problemer mere på IT-siden. Web-hotellet havde på grund af forbedrede opsætninger gjort systemet mindre sårbart, samtidig med at Microsoft leverede bedre programmer både ved selve Access programmet, men også de servere og de web værktøjer vi havde valgt (hvilket var .Net tilsvarende fra Microsoft). Vi kører således med stor tilfredshed på en Access database, hvor opetid er på over 99,9%

Vi valgte at arbejde med referencefiler til hver enkelt af de kategorier, der blev indtastet. Det har været en god løsning, idet det har fungeret robust, og hvor det har været let at udnytte emneordene gennem programmering af web-systemet.

Det må nævnes, at vi havde valgt selv at opbygge vores CMS, hvorved de nævnte fordele let har kunnet implementeres. Det må dog også nævnes, at et standard CMS ville have løst os fra nogle af de opgaver, der løbende har været med opdatering til nye forhold.



FIGUR 3. Allerede ved planlægningen af nyhedstjenestens infrastruktur blev den systematiske emneordliste indtænkt. Illustrationen viser tabellerne i den relationelle database.

IGANGSÆTNING

Da arbejdet skulle igangsættes, havde vi allerede omkring 1.000 dokumenter i systemet. Det var rimelig let at udføre indekseringen, idet der på indtastningskærmen var blevet placeret et vindue, hvor man kunne vælge de ønskede emneord ud fra en menu.

I praksis viste det sig, at det var relevant at give en bog, artikel eller konference mellem to eller tre emneord. Kun sjældent var der kun tale om et så smalt emne, at kun et enkelt ord var dækkende. Omvendt var der sjældent tale om at give mere end seks ord, og som nævnt var det mest almindelige to til tre ord.

Der var naturligvis i f.eks. en bog berørt mange hundrede begreber. Imidlertid måtte vi som registratorer identificere bogens hovedemner. Som nævnt blev arbejdet udført med meget få ressourcer, hvorfor dette arbejde måtte udføres med en meget kort beslutningstid for hver enkelt enhed. Til en slags trøst vidste vi, at alle kategoriseringer kunne ændres, hvis de viste sig at være forkerte eller hvis de manglede emner. Dette har det dog ikke været tale om.

Ved dette arbejde viste der sig imidlertid også en række svagheder ved det udarbejdede system, og en justering af ordlisten blev udført løbende i et vist omfang. Det kunne ske ved tilføjelse af nye ord til de etablerede begreber. Derimod viste det sig heldigvis ikke nødvendigt at flytte rundt på begreber eller ændre hierarkiet.

ORDLISTE PÅ ENGELSK

Det var naturligt at udarbejde en engelsk ordliste, da vi skulle etablere en nyhedstjeneste på engelsk. I stor udstrækning var denne hjemmeside en "spejling" af den danske udgave, således at indekseringen af bøger, artikler og begivenheder i den danske udgave automatisk blev overført til den engelske udgaves bøger, artikler og begivenheder (der var en del bøger og begivenheder på dansk, norsk og svensk, som ikke var medtaget i den engelske udgave).

Vi fandt hurtigt ud af, at mange af de engelske begreber ikke kunne findes via en standard ordbog. Dels var der mange ord, som ikke fandtes i ordbogen, eller også fandtes der flere muligheder, hvor vi ikke var i stand til at bedømme, hvilke der var de rigtige. Det gav – i hvert fald i begyndelsen – en ordliste med visse svagheder.

WIKI OG ORDBOG

Svaghederne i den engelske ordliste blev dog efterhånden begrænset, og ordlisten fik herved nye brugsmuligheder.

Vi søgte at finde de rigtige engelske begreber til at bekræfte eller supplere ordlisten. Det kunne blandt andet ske ved at se brugen af begreber i bøger, som vi tilfældigvis fik lejlighed til at gennemlæse.

En anden og meget vellykket metode var at lade læserne kommentere og supplere ordlisten. Denne Wiki-metode fungerede især ved at folk påpegede fejl og svagheder ved ordlisten, hvor det engelske ord ikke svarede til det danske eller ikke helt havde den samme betydning. Der kom dog også nogle enkelte tilføjelser via denne metode.

I lighed med den folkelige deltagelse, blev ordlisten samtidig en lille ordbog mellem dansk og engelsk. Herved har folk fået et gratis opslagsværk på dette fagområde.

6 ÅRS GODE ERFARINGER

Systemet har nu været i funktion i 6 år i den danske udgave, og det fungerer stadigvæk. Det bør justeres, som de angives i næste kapitel, men det har trods alt fungeret og fungerer stadigvæk godt.

Emneordene har været nyttige på to helt forskellige områder.

Det ene er den direkte personalisering, hvor hjemmesidens brugere har kunnet søge via emneord. De har haft mulighed for at vælge de overordnede begreber, eller de har kunnet vælge mere smalle begreber.

Den anden brug har faktisk vist sig at være den vigtigste. Hjemmesiden er blevet opbygget således, at søgemaskinernes robotter har høstet hjemmesiden efter emne (samt navn og andre dimensioner). Denne opbygning har betydet, at til eksempel Google har indekseret vores hjemmeside meget høj på mange af de 400 emneord. Den

seneste opgørelse er fra 2007, hvor 70 af vores emneord lå inden for 20 af Googles bedste pladser. Af dem var 20 førstepladser. Derved har brugen af emneord været en forudsætning for at vores hjemmeside har haft en søgemaskineoptimering, som må betegnes at være tæt på det optimale.

De mange mangler

Jeg har ikke påstået, at systemet er perfekt. Det vil en emneordliste eller en systematik vistnok aldrig kunne blive.

Bortset fra enkelte tilføjelser af ord til listen, har der ikke været udført en egentlig revision. En sådanne står umiddelbart foran sin effektivering.

Igennem alle årene er der blevet skrevet en "fejlrapport", når der har været en bog eller artikel, som ikke havde sit emne på listen. Der er derved blevet udarbejdet en liste på efterhånden mere end 100 ord.

Årsagen er delvist, at den oprindelige ordliste allerede ved fødslen havde svagheder, således at den ikke kunne tage højde for videnskabelige resultater omkring visse specielle smalle emner, der kun opstod meget sjældent.

En anden årsag er, at faget har udviklet sig. I tidsskriftets efterhånden mere end 10 års forløb har videnskaben ændret sig både i takt med at samfundet har ændret sig, men sandelig også, fordi der i videnskaben selv har været en udvikling, som ikke nødvendigvis er parallel med samfundets udvikling

Alle disse mangler skal der tages højde for ved den snarlige revision.

Systematikens svagheder

Selvom systematikken benyttes, er der nogle svagheder i systemet, som ikke sådan lige er til at komme udenom. Det er tilfældet, når nogle ord benyttes ualmindeligt meget eller når de ikke udnyttes særlig meget.

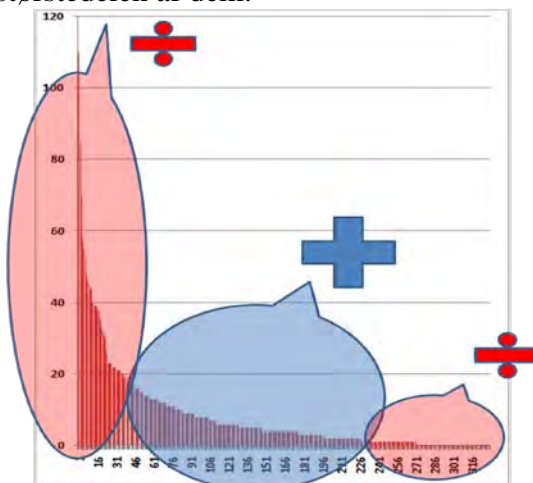
Nogle ord benyttes så meget, at de ikke er særlig velegnede til at lave en delmængde. F.eks. behandler mange bøger "globalisering" eller de kan i det mindste have et aspekt omkring emnet. En søgning på dette emne i bøger på vores hjemmeside vil give 110 henvisninger, hvilket er et meget stort antal at forholde sig til.

Af de mange søgeord havde 12 flere end 40 hits. Det må ved systematikens videre udvikling søges at bringe dette antal ned. Ved begrebet globalisering må der til eksempel opfindes nye begreber, der underinddeler og findeler emnet. Det kunne til eksempel være "globalisering, økonomi", "globalisering, ledelse" og tilsvarende opdeling af globalisering.

Denne svaghed findes udelukkende ved brug af systematikken. Hvis man søger på emneord, vil man kunne kombinere "globalisering" med et andet begreb og herved reducere antallet af hits.

En anden svaghed er hvor emneord har få eller ingen resultater. 130 af emneordene havde faktisk kun én eller ingen bøger at henvise til. Med så få tilfælde kan man sige,

at disse emneord er for specifikke, og at de på en måde er ”støj”, der fylder op og gør tingene uoverskuelige. Det er ikke et mål i sig selv, at udvikle et ”optimalt” system, der dækker alle begreber. Der skal være en pragmatisk fornuft, og der vil derfor ikke ske noget ved at fjerne størstedelen af dem.



FIGUR 4. Begreber med for få eller for mange henvisninger var u hensigtsmæssige.

DEN MÅLBARE SUCCES – ELLER FIASKO

Ved Internettet er vi så heldig, at vi i stor udstrækning kan se hvordan vores brugere benytter vores tjeneste. Mens man ved en udgivelse i bogform ikke kan se, om læserne benytter et indeks eller stikordsregister, er det helt anderledes let at se, hvorledes brugerne navigerer rundt på en hjemmeside. Faktisk er det muligt at registrere hvert eneste tastetryk. Det er ren Big Brother.

Brugen af emneord viser sig således kun at være en betinget succes, eller i hvert fald, at kun en mindre del af brugerne udnytter de nye muligheder til at navigere via systematikken. Kun 50 ud af 1.000 besøgende udnyttede denne mulighed.

Om dette nu er en fiasko, skal der en nærmere vurdering til for at afklare. Mange ikke-brugere navigerer efter andre metoder. De ser f.eks. alle bøger igennem på de aktuelle boglister, hvorved de bliver orienteret om hele forskningsfeltet. Andre søger efter en enkelt forfatter eller en bestemt artikel og er tilsyneladende tilfreds med dette.

Den relativt ringe brug af systematikken kan have flere årsager – hvis det altså er en ringe brug. Det kunne være, at navigationen til systematikken var for besværlig. Det kunne man få bekræftet eller afkræftet ved at gøre navigationen i en periode ualmindelig let at komme til og herefter se, om det skulle være årsagen.

Det kan også være, at en systematik kun benyttes, når en forsker skal i gang med et forskningsprojekt og her ønsker at få et overblik over den totale litteratur på et område. At det kun er de rigtige ”professionelle”, der benytter systematikken, men altså også, at almindelige mennesker måske ikke kender til fordelene ved at benytte en sådan mulighed.

De største succeser

Inden vi kaster systematikken bort, vil jeg dog berette om dens brug på to områder, hvor den virkelig har vist at have berettigelse.

Det ene er på muligheden for målrettet information, en aktiv push fra indholdsleverandørens side, til folk, der er interesseret i at følge med i bestemte områder. Det har vi organiseret via RSS, hvor brugerne har mulighed for at vælge mellem 216 emner. Når der på hjemmesiden kommer en ny bog, vil der blive givet sendt besked til brugernes RSS læser om nyheden.

Den tekniske konstruktion af RSS bygger på emneordene, og her må vi konstatere, at det er en succes i stort omfang. Der er således udvalgt mere end 5.000 RSS-emner, hvilket vi er rimelig tilfredse med. Vores sammenligningsgrundlag er, at vi har 500 abonnenter på den bogtrykte udgave af vores tidsskrift, og da vi har konstateret, at brugerne i gennemsnit kun vælger et par få emner at abonnere på, er der tale om flere tusinde brugere.

Største succes for Google søgerne

Den anden – og vel nok langt den største - succes for systematikken er for brugerne af søgemaskiner. Når man søger efter et ord, vil man typisk kun se efter søgeresultater på den første side og oftest kun efter de første få øverste resultater.

Her har vores hjemmeside en ualmindelig god stilling. En analyse af hjemmesidens placering af søgeord viste i 2007, at stillingen for vores søgeord var blevet væsentlig forbedret siden indførelse af den nye teknik for søgemaskineoptimering baseret på systematikken. 71 begreber lå på de øverste 9 pladser i Google, den altdominerende søgemaskine, og heraf havde vi endda 21 ord på førstepladsen.

Det er helt sikkert den vigtigste årsag bag vores hjemmesides succes. En meget stor del af de besøgende på vores hjemmeside kommer ind via søgemaskiner – mere end en tredjedel er der tale om ved de konkrete målinger, og da mange af de øvrige brugere, som linker direkte til hjemmesiden, i første omgang er blevet opmærksom på hjemmesiden på grund af en vellykket søgning, er resultatet efter brug af søgeord langt mere end halvdelen af hjemmesidens brug i det hele taget.

Succes for det interne arbejde

Ud over den nævnte direkte brug for de besøgende på hjemmesiden, har vi internt i redaktionen yderligere haft udbytte af systematikken. Ved at analysere brugernes anvendelse af systematikken, er vi blevet klogere og kan derved lave et bedre produkt.

Vi har til eksempel set på hvilke begreber, som brugerne har benyttet og hvilke de ikke har benyttet. Det giver os herved en mulighed for at justere vores udvælgelse. Selvfølgelig skal det – efter vores mening – ikke været styret fuldt ud af brugernes

interesser. Det er os i redaktionen, som har forstand på kvalitet og som redigerer ud fra denne selvforståelse. Det vil dog være ufornuftigt, hvis vi ikke lader os inspirere af de ord, som brugerne i særdeleshed efterspørger, idet de afslører et behov for viden på disse områder hos brede kredse, hvilket vi kan lade indgå i vores fremtidige arbejde.

Der må dog nævnes et negativt aspekt ved brug af denne type analyser: De tager tid. Det er i hvert fald for vores nyhedstjeneste ikke muligt at udføre denne type analyser mere end én gang om året.

KONKLUSION

Det må konstateres, at vores nyhedstjeneste har haft et stort udbytte af at benytte en faglig systematik inden for det tværfaglige forskningsfelt, arbejdsliv. Alle vores oplysninger er blevet indekseret, således at hver bogomtale, artikel, anmeldelse eller præsentation af konference har fået et eller oftest flere emner koblet til dig.

Emnesystematikken er blevet udviklet af vores egne redaktører, som er de bedste (og vel også eneste) med et indgående kendskab til fagområdets begreber. Vi har ikke haft økonomi til at få erfarne emneordsudviklere til at hjælpe os med systemet, men kunne dog godt have tænkt os at have fået hjælp undervejs, hvilket ville have lettet os arbejdet og sikkert givet et bedre resultat.

Selvom vi har gjort brug af og blevet inspireret af samtlige nationale og internationale emneordslister i og omkring vores område, har vi måtte udføre arbejdet fra grunden af. Vi kunne dog godt tænke os, at arbejdet med emneord på bedre vis bliver koordineret internationalt, således at andre vil kunne udnytte vores systematik og vi omvendt få hjælp af andre. Det ville være en god forudsætning for en fremtidig udveksling af oplysninger baseret på emneord og systematik.

REFERENCER

Nyt om Arbejdsliv, systematik www.nyt-om-arbejdsliv.dk/systematik.aspx

Working Life Info, systematik www.workinglife.info/esystematik.aspx