

Lokal registrering og indeksering af Internetdokumenter

Af Marianne Lykke Nielsen

I stadig stigende omfang lægges oplysninger af den type, som bibliotekarer traditionelt betjener deres lånere med, ud på Internettet. Flere typer af informationsressourcer: software, billeder, diskussionsgrupper, firmabeskrivelser, personoplysninger o.s.v. er blevet gjort tilgængelige via nettet. Ressourcer, der alle er anvendelige for bibliotekets brugere, og som det kan være relevant at gøre tilgængelige for biblioteksbrugerne. Et centralt spørgsmål er imidlertid, hvorledes bibliotekets brugere skal have adgang til de nye informationer. Der findes flere alternative strategier for, hvorledes et bibliotek kan vælge at formidle adgang til Internettet.

På Internettet findes både internationale og danske søgeredskaber, der stilles gratis til rådighed for Nettets brugere, og som bibliotekerne kan vælge at give biblioteksbrugerne adgang til. Biblioteket kan også vælge selvstændigt at udarbejde lokale, skræddersyede vejvisere, der giver adgang til et udvalg af dokumenter, som skønnes særligt relevante for bibliotekets brugere. Og endelig kan bibliotekerne vælge at registrere Internetdokumenter i det lokale bibliotekskatalog og derved formidle Internetdokumenter på lige fod med bibliotekets andre materialetyper.

Hver form har sine fordele og ulemper. I denne

artikel vil de forskellige søgeadgange blive diskuteret. Diskussionen vil koncentrere sig om de forskellige værktøjers søgemuligheder, og om de krav, som det stiller til katalogisering og emnemæssig beskrivelse af Internetdokumenter, når biblioteket selv vil formidle disse i lokale værktøjer.

Søgeværktøjer på Internettet

Internettets søgemaskiner er store databaser, der ofte registrerer flere millioner Internetdokumenter. Søgemaskinerne benytter sig af søgeroboter, der gennem søger Nettet for nye dokumenter, og lægger data om de fremsøgte Internetdokumenter ned i databasen. Katalogisering og indeksering foregår oftest automatisk ved at ord udtrækkes af computeren enten fra dokumentets fulde tekst eller fra udvalgte dele af dokumentet, f.eks. overskrifter og indledning. Den automatiske indeksering er effektiv og betyder, at en stor del af det enorme antal dokumenter på Internettet bliver registreret og gjort søgbar. Kvalitetsmæssigt medfører den automatiske indeksering en uensartet og ufuldstændig beskrivelse, der sammenlignet med intellektuelt udarbejdede beskrivelser stiller store krav til søgeren, der ved søgning skal tage højde for disse mangler i beskrivelsen (Winship 1995; Dong 1997). Søger må selv medtænke og tage højde for forskellige navneformer, synonymmer og stavemå-

der. Endvidere er det ikke sikkert, at alle gængse oplysninger er med i beskrivelsen. Der er ingen garanti for, at der vil blive udtrykt ord, der angiver, hvem der har udarbejdet Internetdokumentet, hvornår, hvor o.s.v. I den automatisk genererede beskrivelse savnes ofte oplysning om

- ophav
- udgiver
- dato
- sprog
- rettigheder

Mange søgemaskiner giver mulighed for at foretage kombinatoriske søgninger, hvor man kan fremsøge dokumenter ved hjælp af et sæt af søgekriterier, og derved foretage en forholdsvis specifik søgning. Databasernes omfang medfører imidlertid ofte, at søgninger resulterer i et stort antal hit, som det kan være vanskeligt at overskue og relevansbedømme, ikke mindst fordi beskrivelsen af de fremfundne dokumenter ofte er mangelfuld. Mange søgemaskiner vægter svarene efter formodet relevans, hvilket er en nyttig hjælp til at overskue de mange hit.

Søgemaskinerne er værdifulde redskaber for bibliotekets brugere, når det gælder fremfindning af Internetdokumenter, fordi databaserne dækker bredt og hyppigt opdateres, men de kan være vanskelige at bruge effektivt. Den analytiske søgeform, hvor de valgte søgekriterier kombineres ved hjælp af forskellige symboler, kræver både viden og rutine. Brugere skal dels tilegne sig viden om søgemaskinens specielle faciliteter, dels skal brugerne have en grundlæggende forståelse for kombinatorisk søgning og automatisk indeksering. Mange undersøgelser af slutbrugersøgning viser, at brugerne ikke har denne forståelse. De kan sagtens fremfinde information om det søgte emne, da databaserne ofte er store og indeholder flere relevante dokumenter, men undersøgelserne viser, at der er risiko for både lav recall og lav precision (Efthimiadis 1994; Lancaster 1994; Truelsen 1995).

Der arbejdes hele tiden på at gøre søgemulighederne bedre i søgemaskinerne. Metadata-konceptet

(Gerner 1997) har bl.a. til formål at forbedre mulighederne for at foretage mere præcise søgninger ved at øge og forbedre datagrundlaget for den automatiske indeksering. I nogle søgemaskiner tilbydes raffinerede søgefaciliteter som "find lignende dokumenter" og "automatisk spørgsmålsudvidelse", hvis formål er at hjælpe slutbrugeren med at formulere mere effektive søgeforespørgsler.

Trods forbedringer stiller søgemaskinerne store krav til brugeren. Det gør de strukturerede emneindekser ikke. Strukturerede emneindekser er systematisk eller alfabetisk ordnede lister over udvalgte Internetdokumenter. I modsætning til søgemaskinernes automatiske indeksering er de strukturerede indekser fremstillet ved, at nogle personer på baggrund af en vurdering har udvalgt et sæt af dokumenter til registrering og ordning. Indekserne kan være udarbejdet med et generelt sigte og registrere dokumenter, der omhandler mange forskellige emner. Indekserne kan også være afgrænsede og kun indeholde Internetdokumenter vedr. et bestemt emne, fra et bestemt land, af en bestemt form, o.s.v.

Fælles for de emnestrukturerede lister er, at de er nemme at anvende. Som søger browser man rundt i indeksets net af links. Der skal ikke formuleres nogen søgeforespørgsel. Søger vælger blot den emnebetegnelse, som han finder relevant, klikker og får præsenteret en liste af Internetdokumenter, der er blevet registreret under den pågældende betegnelse. Oftest er emnestrukturen simpel og består af få, mellem 10 - 30 overordnede betegnelser, der sjældent er underdelt mere end én gang. Der angives kun lidt information om de indekserede dokumenter, oftest angives kun titlen, og man må normalt klikke videre til Internetdokumentet for at se, om den er relevant.

De strukturerede emneindekser kan betragtes som en form for vejvisere eller bibliografier, og i forhold til søgemaskinernes mekaniske registreringsproces er de kendetegnet af, at den institution, der har udarbejdet listen, har foretaget en vurdering og et materialevalg - en form for kvalitetskontrol. De indeholder derfor også langt færre dokumenter end

søgemaskinerne - Yahoo er den største med omkring 50.000 dokumenter (Hoel 1996). Alligevel kan de som søgemaskinerne være uoverskuelige og især langsommelige at anvende. For hvis emnestrukturen skal være overskuelig, vil der blive registreret mange dokumenter under de enkelte kategorier. Som bruger skal man ofte browse ned gennem lange lister af Internet titler.

Derfor har flere biblioteker også valgt lokalt at udarbejde egne skræddersyede lister, der bedre passer til de lokale brugeres behov, både hvad angår indhold og søgemuligheder. Nogle biblioteker eksperimenterer endvidere med at registrere Internetdokumenter i den lokale bibliotekskatalog og på denne måde formidle adgang til udvalgte Internetdokumenter på lige fod med bibliotekets andre materialer. Effektiviteten af de sidste tiltag er i høj grad afhængig af, at der er web-adgang fra bibliotekskatalogen, således at brugerne med et klik kan gå direkte fra bibliotekskatalogen til Internetdokumentet.

Som det ser ud i øjeblikket, kan bibliotekerne altså vælge at registrere og synliggøre Internetdokumenter gennem to former for registrering:

- *Registrering i den lokale bibliotekskatalog*
 - registrering i bibliotekskatalogen
- *Registrering i lokale bibliografier og vejvisere*
 - registrering i strukturerede vejvisere
 - registrering i bibliografiske databaser

Registrering i den lokale bibliotekskatalog

I den lokale bibliotekskatalog registreres traditionelt de dokumenter, der indgår i bibliotekets samling, d.v.s. de dokumenter, som biblioteket enten har erhvervet sig eller som biblioteket abonnerer på. I fremtiden kan det blive vanskeligt at bibeholde den traditionelle opfattelse af katalogens funktion. Registreres udvalgte Internetdokumenter i bibliotekskatalogen, som biblioteket ikke decideret råder over, forandres katalogens funktion, og formålet med bibliotekskatalogen kan ikke som hidtil

begrænses til registrering af de dokumenter, der indgår i en samling.

Lokal registrering i bibliotekskatalogen kan imidlertid være en frugtbar løsning for bibliotekets brugere. Internetdokumenter vil på denne måde blive synliggjort og gjort søgbare på lige fod med bibliotekets andre materialer, og f.eks. kan en emnesøgning foretages ved hjælp af ét værktøj.

Andre forhold omkring Internetdokumenter vil endvidere røkke ved den traditionelle opfattelse af katalogens funktion. For eksempel vil flere dokumenter i fremtiden eksistere både i en trykt og en Internet version. Allerede nu findes flere eksempler, f.eks. findes tidsskriftet *Sygeplejersken* både i en trykt og en Internetversion. Dette rejser to spørgsmål, dels om der skal udarbejdes to selvstændige registreringer for henholdsvis den trykte og den elektroniske udgave, dels om man ved søgning i katalogen skal kunne henvises fra/hoppe fra en registrering af den trykte udgave til den elektroniske version. Direkte link til den elektroniske version kan medføre, at der i katalogen registreres dokumenter, som biblioteket hverken ejer eller abonnerer på.

Registrering i lokale bibliografier og vejvisere

Ønsker biblioteket ikke at ændre katalogens funktion, kan biblioteket vælge en anden strategi og udarbejde lokale bibliografier i form af emnestrukturerede, lokale vejvisere eller i form af lokale elektroniske databaser. Disse lokale produkter kan - på traditionel vis - registreres i bibliotekskatalogen, evt. med et direkte link til vejviseren uden at katalogens funktion forandres.

Normalt defineres en bibliografi som en afgrænset fortegnelse over dokumenter, hvor nationale, sproglige eller emnemæssige kriterier ligger til grund for udvælgelsen af dokumenter til bibliografien. Bibliografier kan ordnes efter forskellige principper, og i denne sammenhæng ligger forskellen mellem en elektroniske database og en vejviser netop i ordningen og søgemulighederne. I en data-

base kan de registrerede dokumenter fremsøges ud fra mange forskellige kriterier, der ydermere kan kombineres, mens de registrerede dokumenter i en vejviser typisk er ordnet under et eller få kriterier, der ikke kan kombineres på søgetidspunktet. Søgemæssigt stilles der altså større krav til beskrivelse og registrering i en database, hvis søgemulighederne skal udnyttes optimalt, end i en vejviser, hvor det enkelte dokument kun skal kunne fremsøges under nogle få beskrivelseselementer. Med henblik på relevansbedømmelse bør der stilles de samme krav til de to typer af søgeværktøjer om en fyldig karakteristik af de enkelte dokumenter. I de fleste emnestrukturerede vejvisere, der findes på nettet i dag, danner alene titlen baggrund for relevansbedømmelse. Titler, der iøvrigt ofte er konstruerede (Gerner 1997).

Strategier for registrering

Bibliotekerne står overfor at skulle formulere en strategi for registrering af Internetdokumenter. Overordnet skal bibliotekerne bestemme, hvorledes de vil give brugerne adgang til de mange dokumenter på Internettet. Skal det ske

- via Internettets eksterne søgeredskaber?
- via registrering i bibliotekskatalogen?
- via lokale vejvisere og bibliografiske databaser?

Dernæst skal de formulere en politik for, hvilke Internetdokumenter, som de ønsker at registrere og hvordan. Sluttelig skal de finde retningslinier for, hvorledes registrering skal foregå ved de forskellige registreringsformer. Registreringsarbejdet vil bestå af følgende delprocesser:

- evaluering og vurdering af Internetdokumenter
- "materialevalg"
- valg af registreringsform for det enkelte Internetdokument
- registrering

Registrering

Formålet med registrering er synliggørelse og fremfindning. De data, der fremkommer ved regi-

streringen, skal muliggøre, at brugerne kan

- søge/finde Internetdokumenter, der svarer til brugernes søgekriterier
- identificere, om det beskrevne dokument svarer til det, der blev efterspurgt
- vælge det/de dokumenter, der synes relevante
- få adgang til det/de valgte dokumenter

Registreringen skal indeholde dataelementer, der identificerer og karakteriserer de dokumenter, der skal gøres tilgængelige for brugerne. Typisk opdeles en registrering i deskriptive data, data i normativ form, emnedata og lokale data, hvor de deskriptive data anvendes som redskab til at identificere og karakterisere dokumentet, mens emnedata og de normative data skal anvendes til at genfinde dokumentet, og de lokale data angiver administrative forhold vedr. accession, beholdning, brug/udlån og opstilling.

Internetdokumenter er elektroniske dokumenter, der er kendetegnet af at være dynamiske dokumenter under stadig forandring både indholdsmæssigt og med hensyn til lokalisering. Et Internetdokument består ofte af en samling af medietyper, der kan omfatte tekst, grafik, billeder, animationer, lyd, videoklip og links til andre sider, og de kan også være sammensat af flere materialeformer og bestå af en kombination af præsentationer, artikler, henvisninger (links til andre dokumenter), databaser, ordforklaringer m.m. Mange Internetdokumenter er konstrueret som et net (web) af relaterede dokumenter, hvor de enkelte dokumenter kan have selvstændige adresser (URL'er). Endvidere kan nogle Internetdokumenter have tætte relationer til andre dokumenter, trykte eller elektroniske. Nogle dokumenter publiceres ofte i to identiske versioner: en trykt og en Internet version. Ligeledes bruges Internetdokumenter til opdatering af elektroniske opslagsbøger på cd-rom, og har på denne måde stærk relation til et andet dokument.

Ovennævnte karakteristika er væsentlige og bør beskrives, da de har betydning for brugen og relevansbedømmelsen af dokumentet.

Deskriptive data

Deskriptive data anvendes som redskab til identifikation og relevansbedømmelse, og dataene hentes traditionelt fra selve dokumentet og gengives i den ordlyd, som forekommer i dokumentet. Traditionelt opdeles de deskriptive data i en formaliseret del, som er en gengivelse af data, som findes i selve dokumentet og i noter, som er en uddybning eller supplement til den formelle del i form af f.eks. indholdsbeskrivende noter, angivelse af medietyper, struktur, fysisk beskrivelse og materialetyper.

Ved registrering af Internetdokumenter bør følgende datatyper søges identificeret og indlemmes i den deskriptive beskrivelse:

- oprettelsesdato og sidste opdateringsdato
- registreringsdato
- URL/PURL
- fysisk beskrivelse (herunder edb-tekniske karakteristika)
- medietyper
- struktur
- relation til andre dokumenter

Derudover er det selvfølgelig vigtigt at traditionelle deskriptive data registreres:

- titel
- ophav
- udgiver
- sprog
- rettigheder

Normative data

Normative data er data i vedtagen form. Normative data har en samlende og relaterende funktion og ved at lade f.eks. én bestemt forfatter optræde i samme navneform, skabes sikkerhed for entydig identifikation af forfatteren og høj grad af recall.

Traditionelt optræder følgende dataelementer i vedtagen form:

- ophav
- serietitler
- uniforme titler

Ved Internet indeksering kan det være hensigtsmæssigt at normalisere og standardisere den fysiske beskrivelse og anvende standardiserede betegnelser for medietyper, struktur, edb-tekniske karakteristika og relationer. Edb-teknikken udvikler sig hele tiden, og det kan være vanskeligt både som registrator og søger at vide, hvilke betegnelser man skal anvende. En hyppigt opdateret liste over normative betegnelse kan være et nyttigt redskab for både registrator og søger.

Emnedata

Formålet med emnedata er genfindning af dokumenter udfra deres emne(r). Formaliserede emnedata gives til dokumentet efter en analyse af dokumentets indhold. Emneindeksering består grundlæggende af tre faser:

1. undersøgelse af dokumentet og bestemmelse af dets emner
2. bestemmelse af de hovedbegreber, som emnet består af
3. at udtrykke emnets hovedbegreber i det valgte indekseringssprog

Ved registrering udtrykkes dokumentets emner og begreber gennem et eller flere indekseringssprog. Indekseringssprog kan deles i verbale indekseringssprog, hvor emnet udtrykkes i emneord, eller i systematiske indekseringssprog, hvor emneindholdet beskrives i form af notationer. Samtidig skal man være opmærksom på, at også deskriptive data og normative data kan betragtes som emnedata, f.eks. vil

- ord fra en titel (faglitteratur)
 - ord fra indholdsbeskrivende noter
 - ord fra indholdsfortegnelser
- ofte fortælle noget om dokumentets emne.

Det kan ofte være nyttigt at berige den emnemæssige beskrivelse med en angivelse af den måde, som dokumentets emner er behandlet på i dokumentet. Ved f.eks. at beskrive dokumentets målgruppe, angives indirekte det faglige niveau og den synsvinkel, hvormed emnet er behandlet i dokumentet. Emnedata, der beskriver dokumentets

form, angiver tilsvarende vigtige emnemæssige forhold ved dokumentet. Man vil ikke få det samme at vide om emnet 'engelsk sprog' i henholdsvis en 'grammatik', 'case study' og 'læsebog'.

Ekshaustivitet er et udtryk for, hvor udtømmende og dækkende man søger at beskrive et dokument emneindhold. Målet kan være at indeksere alle de emner og begreber, der kan identificeres i et dokument, d.v.s. man stiler mod en høj grad af ekshaustivitet, eller målet kan være kun at indeksere de emner, der behandles grundigt i dokumentet, hvilket medfører en lav grad af ekshaustivitet, men en høj grad af vigtighed.

Ved fastlæggelse af retningslinier for indeksering af Internetdokumenter kan det være relevant at indeksere mere indirekte emner ekshaustivt, f.eks. sværhedsgrad, målgruppe, synsvinkel, materialetype, kvalitet og forskningsmetode. På grund af Internettets omfang vil der for mange emner kunne findes et meget stort antal emnemæssigt relevante dokumenter. Medtager beskrivelsen også den tilgang og den form, som emnet er behandlet ud fra, vil det være muligt at adskille og frasortere dokumenter, der behandler det efterspurgte emne ud fra en forkert synsvinkel. Denne frasortering vil da kunne foretages ud fra formelle, kvalitative kriterier af betydning for emnet og for søger, og ikke ud fra en afgrænsning på publikationstidspunkt, et nemt og praktisk fravalgskriterie, der ofte anvendes, men som ikke direkte har relation til emnet.

Med samme argumentation kan det være hensigtsmæssigt ved Internet indeksering kun at indeksere emner af en vis vigtighed i dokumentet, eller bestemte aspekter af et emne. Ved vægtet indeksering af kun meget vigtige emner vil antallet af fremfundne dokumenter nedbringes, og der vil skabes en højere grad af sikkerhed for emnemæssig relevans.

Specificitet er et begreb, der som ekshaustivitet er knyttet til den emnemæssige beskrivelse. Specificitet er et udtryk for, hvor nuanceret og specifikt det er muligt at udtrykke et dokument emner og

begreber ved hjælp af indekseringssproget, d.v.s. den udstrækning et bestemt begreb lader sig udtrykke i indekseringssproget. Muligheden for at anvende specifikke emnedata øger mulighederne for at udføre præcise søgninger. Derfor, da antallet af Internetdokumenter er stort, kan det være hensigtsmæssigt at emnebeskrive ved anvendelse af specifikke emneord.

Behovet for henholdsvis ekshaustivitet og specificitet hænger dog i høj grad sammen med den valgte registreringsform og søgemuligheder. Ekshaustiv og specifik emneindeksering udnyttes bedst gennem analytisk søgning i bibliografiske databaser og online bibliotekskataloger, mens browsing i emnestrukturerede lister kræver en overskuelig struktur, der sjældent vil kunne kombineres med en meget specifik og ekshaustiv indeksering. Andre faktorer vil også have betydning for indekseringsniveauet, idet både emneområdets karakter og dokumenternes karakter vil påvirke indekseringen. Det er nemmere at etablere en overskuelig og dybtgående hierarkisk struktur for et mindre emneområde fremfor et mere komplekst emneområde. Ligeledes vil mindre komplekse dokumenter, der kan beskrives ved hjælp af få emnedata, være nemmere at systematisere i et hierarki fremfor mere komplekse dokumenter, der vil kunne ordnes under flere emnedata.

Lokale data

De lokale data beskriver administrative forhold ved dokumentet:

- er dokumentet udlånt
- hvor mange eksemplarer haves
- hvor befinder dokumentet sig i biblioteket

Også til Internetdokumenter knytter der sig nogle lokale data. Af registreringen bør det fremgå, om dokumentet kan bruges vederlagsfrit, eller om der f.eks. er knyttet nogle betalingsbetingelser og dertilhørende restriktioner til brugen. Det bør endvidere angives, hvorledes man kommer i kontakt med dokumentet. Dette kan gøres ved at angive dokumentets URL, men bedst er det, hvis der ved hjælp

af en klikbar adresse/link er direkte adgang til Internetdokumentet fra registreringen til det beskrevne dokument.

Som det fremgår kan der knyttes forskellige typer af beskrivende data til Internet dokumenter, og ved valg af registreringsniveau er det op til det enkelte bibliotek selv at bestemme registreringsniveau og standardiseringsgrad. Valget vil være afhængig af flere faktorer:

- registreringsform
- dokumenttyper
- emneområdet
- tilgængelige ressourcer

Ved registrering i katalogen vil det ofte være naturligt at følge de almindelige retningslinier og regler for registrering, mens der kan være behov for at udarbejde specielle, lokale retningslinier for registrering i bibliografiske databaser og strukturerede vejvisere. Når det gælder registrering i lokale online bibliotekskatalog og lokale bibliografiske databaser, vil mange biblioteker nok vælge at tage udgangspunkt i de retningslinier, der angives i eksisterende katalogiseringsregler og formatbeskrivelser, mens biblioteket i højere grad vil finde sine egne regler og standard for eventuelle lokale vejvisere.

Endnu er der kun udarbejdet få specielle retningslinier for, hvorledes en emneliste på Internettet emnebeskrives og struktureres hensigtsmæssigt (Strain 1996; Weinberg 1996; Rusch-Feja 1997). I de følgende afsnit vil forskellige regler og retningslinier fra henholdsvis klassifikationsteorien og teorierne vedr. tesauruskonstruktion blive inddraget og diskuteret med henblik på at finde brugbare designprincipper for emnestrukturerede vejvisere.

Opbygning af (emne)strukturerede vejvisere

De emnestrukturerede vejvisere er bygget op i hypertext, der er et system af elementer (tekstdele, billeder, video, lyd, grafik, animationer etc.). Disse elementer, der kaldes nodes, er maskinelt forbundne ved prædefinerede links, der gør det muligt at

navigere rundt i hypertext systemet, fra node til node. Springet fra en node til en anden udløses af et klik med en musetast på et såkaldt anker. Dette anker kan være et ikon, en tekststreng, et stykke grafik etc.

Hypertext muliggør en anderledes og mere uformel, interaktiv søgeform. Brugeren kan vælge sin egen vej til systemets forskellige nodes, uafhængig af systemets lineære struktur, men selvfølgelig afhængig af de prædefinerede links.

Søgning i hypertext stiller ikke samme krav om søgeviden til brugeren som analytisk søgning, idet søgning ikke kræver, at der udfærdiges en formaliseret søgeforespørgsel. Muligheden for at finde sit eget spor gennem hypertexten letter søgningen, men rummer samtidig en fare for, at man mister orienteringen på sin vej rundt i systemet.

Hypertext tillader ikke kombination af søgekriterier. Søgeresultatet er fastlagt på forhånd, og bruger kan ikke som ved analytisk søgning styre recall og precision. Søgeeffektiviteten er også afhængig af dokumenternes omfang og indhold. Det tager ofte længere tid at hente et dokument, bestående af multimedier, end et rent tekstbaseret dokument.

En vejviser struktureres traditionelt omkring følgende nodes:

- ordningselement (oftest i form af et emneord eller en materialebetegnelse)
- beskrivelsesenhed (oftest i form af en titel)
- link til Internetdokumentet

Ved opbygning af en emnestruktureret vejviser skal konstruktøren fastlægge retningslinier for henholdsvis ordningselementer, vejviserens struktur og beskrivelsen af Internetdokumentet:

- valg af ordningselementer, d.v.s. fastlægge graden af specificitet
- tildeling af ordningselementer, d.v.s. fastlægge graden af ekshaustivitet
- henvisninger mellem ordningselementer

- strukturering af ordningsselementer
- valg af beskrivelsesenheder
- standardisering af beskrivelsesenheder

Valg af ordningsselementer

Målet med de emnestrukturerede lister er at samle beslægtede Internetdokumenter. Derfor grupperes det enkelte Internetdokument under en eller flere betegnelser (emneord), der angiver et af de hovedemner, som Internetdokumentet omhandler, eller som beskriver den form, som karakteriserer dokumentet.

Disse betegnelser bør være kontrollerede og standardiserede og udvalgt i forhold til den sprogbrug, der kendetegner vejviserens målgruppe. Det er indlysende, at de valgte emneord bør være klare og præcise og umiddelbart forståelige for målgruppen. Det kan eventuelt være hensigtsmæssigt at anvende kvalifikatorer eller noter til at tydeliggøre emneordets mening. En kvalifikator kan eksempelvis angive forskellen mellem *medicin (faget)* og *medicin (medikament)*. Det kan være nyttigt at bruge emneord fra et velkendt indekseringssprog.

Brugernes sprogbrug, og også antallet af Internetdokumenter, skal bestemme emneordenes grad af specificitet. Mange dokumenter vil kræve flere og ofte mere specifikke betegnelser for få en overskuelig vejviser.

Kravet om overskuelighed påvirker endvidere graden af ekshaustivitet. Mange dokumenter vil kunne ordnes (registreres/indekseres) under flere ordningsselementer, men placering under flere betegnelser bør begrænses, hvis omfanget af listen skal være overskuelig.

Det kan endvidere være hensigtsmæssigt at berige vejviseren med henvisninger fra synonyme og relaterede betegnelser til det valgte ordningsselement, f.eks. fra *kræftsygdomme* til *cancersygdomme*. Igen vil antallet af Internetdokumenter og også den grafiske layout have indflydelse på, hvorvidt disse

henvisninger vil komme til at fungere efter hensigten og ikke blot øge disorienteringen.

Logisk og overskuelig struktur

Det siger sig selv, at dokumenterne bør samles i grupper, hvis struktur forekommer målgruppen logisk. Strukturering kan ske ud fra forskellige principper og underdelingskriterier, og strukturering bør ske i forhold til både målgruppen, fagområdet og antallet af Internetdokumenter.

En mulighed vil være at lave en struktur, baseret på emneområdets delemner og discipliner, f.eks. kan emneområdet *sygepleje* opdeles i følgende delområder og discipliner: *foreninger og institutioner, sygeplejeuddannelse, anatomi, fysiologi, sygdomslære, behandlingsformer og hospitalsvæsen*. En anden mulighed vil være at anvende facetaanalysen som grundlag for strukturering og opdele emneområdet i et antal begrebskategorier (facetter). En vejviser ordnet i facetter opdeler ikke emneområdet i delemner, men i typiske aspekter for emneområdet. *Sygepleje* kan f.eks. deles i følgende facetter: *aktiviteter, agenter og redskaber*:

Aktivitetesfacet

*pleje
forebyggelse
behandling
uddannelse*

Personalefacet

*sygeplejesker
sundhedsassistenter
laboranter*

Hjælpemiddelfacet

*kørestole
gangstole*

For at opnå forståelse og overblik bør en emnegruppe ikke være for stor. Det er bedre at underdele de enkelte grupper af emneord. Gennem underdeling skabes en hierarkisk struktur, hvori nogle emneord er overordnede og andre er underordnede. Hver gruppe kan indeholde flere sideordnede emneord.

Når den overordnede struktur og eventuel underdeling er fastlagt, skal der tages bestemmelse om præsentationsrækkefølgen for de enkelte emner. Dette kaldes i klassifikationsteorien for systemorden. Igen er det vigtigt at strukturen tilpasses den indfaldsvinkel, som er logisk og umiddelbart forståelig for brugergruppen. Forskellige principper for orden kan nævnes: det generelle før det specielle, det mere konkrete før det mindre konkrete, konsensus indenfor faget m.v.

For de grupper, der yderligere er underdelt i sideordnede delelementer, skal der endvidere fastsættes en nyttig orden for de sideordnede emner. Der er forskellige valgmuligheder, der bør bruges i forhold til målgruppen: kronologisk orden, orden efter tiltagende kompleksitet, geografisk orden m.v.

Hypertekst giver gode muligheder for at lave en overskuelig struktur. I og med, at man ved hjælp af links kan klikke sig dybere ned i et hierarki, er det ofte muligt at præsentere de enkelte niveauer i strukturen i ét skærm billede, hvorved strukturen bedre kan overskues. Mange niveauer i hierarkiet kan imidlertid påvirke søgeeffektiviteten og medføre, at der bliver for mange trin i fra øverste til nederste del af hierarkiet. Ved mange Internetdokumenter kan det være vanskeligt at finde den rette balance mellem overskuelighed og effektivitet.

De beskrivende data

I de fleste emnestrukturerede lister angives kun lidt information til beskrivelse af Internetdokumentet. Ofte angives kun titlen, og man må normalt klikke videre til Internetdokumentet for at bedømme, om det er et brugbart dokument. Relevansbedømmelse foretages selvfølgelig bedst med udgangspunkt i det egentlige dokument, men med henblik på at øge søgeeffektiviteten og mindske søgetiden kan det være hensigtsmæssigt at udvide beskrivelsen med flere data. Udover at berige beskrivelsen med deskriptive oplysninger om ophav, udgiver, konstruktions- og opdateringsdato kan det være nyttigt at tilføje oplysninger, der evaluerer og vurderer Internetdokumentet. Oplysninger, der både af-

grænser og kvalitetsvurderer dokumentets indhold og dets funktionalitet og effektivitet.

Indeholder en vejviser mange dokumenter, kan brugeren i en valgsituation mellem flere dokumenter i høj grad drage nytte af oplysninger, der beskriver dokumenternes

- indhold
 - aktualitet
 - kvalitet
 - autoritet
 - stabilitet
- funktionalitet
 - søge- og navigeringsmuligheder
 - orientering
 - overskuelighed
- teknisk format
- omkostninger

Dokumenternes karakter og målgruppens behov vil være afgørende for beskrivelsesniveauet. Mange oplysninger vil kun blive angivet, når Internetdokumentet har nogle specielle karakteristika. For eksempel kan det være hensigtsmæssigt at angive det i beskrivelsen, hvis der kræves en bestemt browser for at få det fulde udbytte ved læsning af dokumentet.

I figur 1a til 1d vises udsnit af en vejviser for sygeplejeelever. Vejviseren er et opdigtet eksempel, og Internetvejviser for Sundkøbing Sygeplejeskole eksisterer ikke i virkeligheden. Emneord og struktur er dog valgt med udgangspunkt i virkelighedens verden, idet et udvalg af relevante dokumenter, både Internetdokumenter og trykte lære- og fagbøger indenfor fagområdet er blevet konsulteret. Desuden er undervisningsplaner og en sygeplejeske med både praktisk erfaring og undervisningserfaring blevet kontaktet. En mindre domæneanalyse danner således grundlag for udarbejdelsen af eksemplet.

På øverste niveau er vejviseren opdelt i grupper af emneord, der hver repræsenterer et aspekt af faget. Grupperne er ordnet og præsenteret således, at det generelle kommer før det specielle - først *sundhedsvæsen, sygeplejefaget og sygeplejeuddannelse*, dernæst *sygepleje og omsorg, ernæring og kost, sygdomslære og -behandling og socialmedicin*. Tilsidst kommer det mest specielle: fag på sygeplejeskolen fra andre fagområder: *anatomi, fysiologi, psykologi, pædagogik og filosofi og etik*. I grupperne vedr. *sygepleje* og vedr. *andre fag* er de sideordnede emneord ordnet i en rækkefølge, der angiver vigtighed i forhold til det grundlæggende fag *sygepleje*. Gruppen vedr. *faglige forhold* er ordnet således, at det generelle kommer før det mere specielle.

De enkelte emneord er underdelt efter behov ved brug af forskellige underdelingskriterier. Medmindre der findes en logisk orden for målgruppen, er de sideordnede emneord indenfor de enkelte undergrupper ordnet alfabetisk. Underdeling sker kun, når antallet af Internetdokumenter og læsevenligheden påkræver det, da mange hierarkier kan medføre ineffektiv og langsommelig navigering. Er der ikke behov for at tydeliggøre den valgte opdeling eksplicit gennem opdeling i flere niveauer, vil Internetdokumenterne implicit blive ordnet efter de valgte opdelingskriterier og i den valgte orden, som det f.eks. er sket i vejviserens overordnede niveau (se figur 1a).

Listen er ikke fuldstændig, men tilpasset mængden af Internetdokumenter og målgruppens tilgang til emnet. Der mangler f.eks. en gruppe for *tandsygdomme*, som ikke skønnes relevant for sygeplejeelever. Ligeledes kan andre af sygdomsbetegnelserne falde væk på grund af manglende litterært belæg.

Emneordene er valgt i forhold til den sprogbrug, der anvendes på sygeplejeskolen, f.eks. er ordet *kræft* valgt i stedet for *cancer*. Til nogle af emneordene er knyttet en forklarende note, der præciserer emneindholdet.

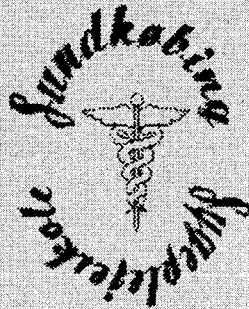
Også beskrivelsen er beriget. Der gives basisoplysninger om ophav, titel og dokumenttype. Foruden angives dokumentets emneindhold og eventuelle andre vigtige oplysninger i en note, f.eks. om dokumentet har relevante links til andre Internetdokumenter eller om der er specielle adgangsforhold.

Konklusion

Mange biblioteker ser det som deres opgave at formidle Internettets mange dokumenter på samme måde, som biblioteket bestræber sig på at stille bibliotekets øvrige materialer til rådighed. Præsentationen af Internetdokumenter kan ske ved anvendelse af flere forskellige strategier. På Nettet stilles forskellige værktøjer vederlagsfrit til rådighed i form af bibliografiske databaser og strukturerede emnelister. Disse kan gøres tilgængelige for bibliotekets lånere. De kan imidlertid være vanskelige og uoverskuelige at anvende, når man ønsker at finde få præcise oplysninger, og derfor vælger mange biblioteker at udarbejde specielle emnelister, der rummer et udvalg af relevante Interdokumenter. En tredje strategi er at formidle Internetdokumenter via bibliotekskatalogen.

Hver formidlingsform kræver forskellige beslutninger og retningslinier, der også skal afvejes i forhold til biblioteket og dets målgruppe. Ofte kan de almindelige regler for bibliografisk og emnemæssig registrering (gen)bruges ved registrering i lokale bibliotekskataloger og lokale bibliografiske databaser, mens klassifikationsteorien og teorien vedr. tesaurskonstruktion kan anvendes til konstruktion af emnestrukturerede vejvisere.

INTERNETVEJVISER FOR SUNDKØBING SYGEPLEJESKOLE



© Udarbejdet af Biblioteket ved Sundkøbing Sygeplejeskole
1996

Opdateret august 1997

Sundhedsvæsen
Sygeplejefaget
Sygeplejeuddannelse

Sygepleje og omsorg
Ernæring og kost
Sygdomslære og -behandling
Socialmedicin

Anatomi
Fysiologi
Psykologi
Pædagogik
Filosofi og etik

Figur 1a: Internetvejviser for Sundkøbing Sygeplejeskole - Overordnet struktur

Sygdomslære og -behandling - patologi er læren om sygdommene og deres indvirkning på organismen. Gruppen er delt op efter kroppens organer og alfabetisk ordnet. Derudover er der specielle grupper for allergiske, arvelige, aldersbetingede og andre sygdomme, der ikke karakteriseres af et bestemt organ.

Sygdomme i

Bevægeapparatet

Blodet

Brysterne

Centralnervesystemet

Fordøjelsesorganer

Hormoner

Hud- og kønssygdomme

Kredsløbet

Kønsorganer

Luftveje

Nyrer og urinveje

Øjet

Øre-næse-hals

Specielle sygdomme

Allergi

Arvelige sygdomme

Aldersbetingede sygdomme

Infektionssygdomme

Kræft

Psykiske lidelser

Figur 1b: Underdeling af emnegruppen *Sygdomslære og -behandling*

Psykiske lidelser - gruppen er ordnet alfabetisk efter sygdommens navn

Alzheimers sygdom

Autisme

Senil demens

Skizofreni

Figur 1c: Underdeling af gruppen *Psykiske lidelser*

Autisme

What is autism

Artiklen er udarbejdet for Autism Society of America af Catriona Johnson & Ben Dorman. Opdateres halvårligt.

Giver en kort og præcis beskrivelse sygdommen autisme: typer, årsager, diagnose og behandling.

Videnscenter og center for autisms hjemmeside

Organisationshjemmesiden er udarbejdet af Videnscenter for Autisme. Opdateres månedligt.

Indeholder en præsentation af centret og dets aktiviteter: projekter, kurser, konferencer, bibliotek og publikationer. Der findes en grundig beskrivelse af autisme. Hjemmesiden henviser til andre Internetdokumenter om emnet.

Autism resources page

Internetvejviseren er udarbejdet af John Wobus, far til autistisk dreng. Opdateres ugentligt.

Henviser til internetdokumenter om autisme og Asperger's Syndrom. Der er ikke anvendt specifikke optagelseskriterier udover, at dokumenterne skal indeholde information om emnet. Vejviseren henviser til dokumenter om behandlingsmetoder og -programmer, online artikler, personlige skildringer, organisationer, diskussionsgrupper, nyheder, bibliografier og vejvisere, biblioteker.

Figur 1d: Vejviser for emneordet *Autisme*

Litteratur

Dong, Xiaoying & Su, Louise T. (1997). Search engines on the World Wide Web and information retrieval from the Internet: a review and evaluation. i: Online & cdrom review, vol. 21, nr. 2, 1997, s. 67 - 81

Efthimiadis, Efthimis N. (1994). I: Proceedings of the fifteenth National Online Meeting. Medford, NJ : Learned Information, 1994, s. 113 - 120

Grauballe, Henning et al. (1997). Klassifikations-teori. København : Danmarks Biblioteksskole, 1997, 49 s. (Undervisningsnoter i Klassifikation ; 1)

Grauballe, Henning & Kaae, Søren (1997). Konstruktion af klassifikationssystemer. København : Danmarks Biblioteksskole, 1997, 79 s. (Undervisningsnoter i Klassifikation ; 2)

Hoel, Ivar A. L. (1996). Søgeværktøjer og indekser til Internet. i: Biblioteksskolenyt, nr. 5, 1996, s. 8-9

Lancaster, F. W. et al. (1994). Searching databases on CD-ROM: comparison of the results of end-user searching with results from two modes of searching by skilled intermediaries. i: RQ, vol. 33, nr. 3, 1994, s. 370 - 386

Rusch-Feja, Diann (1997). Subject-oriented collection of information resources from the Internet. i: Libri, vol. 47, nr. 1, 1997, s. 1 - 24

Strain, Helen & Berry, Pauline (1996). Better page design for the World Wide Web. i: Online & cdrom review, vol. 20, nr. 5, 1996, s. 227 - 237

Thorlund Jepsen, Erik (1997). Undervisningsnoter til katalogisering I. København : Danmarks Biblioteksskole, 1997, 29 s.

Truelssen, Birgit (1995). Selvbetjening på CD-ROM. I: Biblioteksarbejde, 16. Årgang, nr. 45, 1995, s. 55 - 65

Weinberg, Bella Hass (1996). Complexity in indexing systems - abandonment and failure implications for organizing the Internet. i: Proceedings of the 59th Annual Meeting of the American Society for Information Science, Baltimore, Maryland : ASIS, 1996, s. 84 - 90

Winship, Ian R. (1995). World-Wide Web searching tools: an evaluation. i: Vine, nr. 99, June, 1995, s. 49 - 54