

Søgemuligheder i fremtiden

Taler vi samme sprog?

Ved Søren Kaae

Historisk tilbageblik

Så længe jeg kan huske, har spørgsmålet om et fælles sprog - med hensyn til at udtrykke dokumenters emner - været diskuteret i bibliotekssammenhæng. Der har været skrevet rapporter, nedsat udvalg og holdt konferencer.

I 1975 anbefalede Forskningsbibliotekernes Fællesråds Forretningsudvalg, at UDK blev indført som fælles klassifikationssystem for forskningsbibliotekerne - og da folkebibliotekerne var godt kørende med DK5, så skulle det jo være på skinner. Men forskningsbibliotekerne er ikke sådan at få has på. Så Dewey-systemet kom på banen i anbefalingen fra 1976. Men de fleste "egne" systemer eksisterer vel fortsat.

Samtidig diskuteredes i folkebiblioteksregi - men også med tanke på forskningsbibliotekerne - mulighederne for et fælles klassifikationssystem og et evt. verbalt system. I denne forbindelse udarbejdede Friis-Hansen og Mogens Weitemeyer en redegørelse til Emnekatalog-udvalget. Heri anbefales også UDK - Dewey - DK5 som klassifikationssystemer, mens det præ-koordinerede PRECIS-system anbefales som det verbale system (set i relation til fastforms-kataloger).

I 1982-83 var der igen udvalgsarbejde i BC-regi - denne gang om online-emesøgning. Man anbefalede at fortsætte med DK5 samt en udbygning med emneord - det vi i dag kender som DBCs emneord - som så startede i 1984 som udviklingsprojekt og gik i drift i 1986 for 10 år siden.

I 1984 holdt Danmarks Forskningsbiblioteksforening en konference om emnedata i online-alderen - hvorfra der også foreligger en publikation. I sig selv giver den ingen anbefalinger - men forskningsbibliotekerne vil også gerne igang med verbale emnedata - men der foreslås bl.a. at der arbejdes med en standard for indeksering.

I 1991 nedsættes endnu en arbejdsgruppe af Statens Bibliotekstjeneste, som skal komme med "forslag til langsigtede planer for udviklingen indenfor emnedataområdet i forskningsbibliotekerne, herunder klassifikation og indeksering".

Denne gang afholder man sig fra at anbefale systemer - og slet ikke klassifikationssystemer. Derimod anbefales, at der tilbydes verbal emnesøgning, samt at der udarbejdes fælles retningslinier for udformning af emneord.

Vi er altså langsomt gledet fra at snakke fælles sprog i klassifikationsforstand - til at snakke fælles sprog i emneordsforstand.

Og fordelen ved det er, at vi nu fokuserer på en emnemæssig udtryksform, som i hvert fald på overfladen, kan gøres forholdsvis ensartet - og forståelig for andre end bibliotekarer - hvis vi vil.

Nu ved jeg godt - og det gør I også - at selvom vi engang skulle blive enige om en fælles udtryksside (emneord - eller i gamle dage UDK f.eks.) - så vil den samme bog stadig kunne indekseres forskelligt i forskellige biblioteker - afhængigt af, hvordan den læses - fortolkes - og af hensigten med at anskaffe den. Men måden at læse og forstå på, kan vi heldigvis ikke kontrollere og ensrette.

Jo flere (kvalificerede) synsvinkler, der kan angives på et dokument, jo mere og bredere kan dokumentet bruges.

Nok om det.

Når vi snakker om sprog - og her specielt indekseringssprog - kan vi prøve at se på nogle elementer i sproget, som kan gøre det fælles - det vi kunne kalde kontrollen med sproget.

Hvis vi vil/skal tale samme sprog, og forstå hinanden, skal vi overholde nogle regler - nogle grammatiske regler for såvel ords udformning som sætningsopbygning (morfologisk kontrol og syntaktisk kontrol) - og vi skal også mene nogenlunde det samme med ordene - vi skal være nogenlunde enige om ordenes betydning - det er den semantiske kontrol.

Når jeg nu kunne finde på at sige, at vi taler samme sprog - i indekseringsmæssig forstand - så er det fordi vi arbejder med det danske sprog - og fordi de sproglige elementer vi arbejder med, er nogenlunde genkendelige for os alle sammen. Men alligevel går det også galt - vi taler ikke helt samme sprog.

Vi har ikke samme regler for:

- grammatik
 - bøjningsformer
 - termformer
 - syntaks
- mening

Vi kan starte med det sværeste: meningen - den semantiske kontrol.

Det er her I begynder at tænke på tesaurusser:

- kontrol med synonymer
- homonymer
- hierarkiske relationer
- idéassociative relationer

Eks.: Kommunerne og de sindslidende.

emneord 1

psykisk syge; social forsorg; kommuner; undersøgelser (formbetegnelse); levevilkår (ukontrollereret)

emneord 2

distriktpsychiatri; psykiatri; sindslidende; sindssyge

emneord 3

distriktpsychiatri; psykiatriplanlægning; psykiatriske patienter; psykiatriske patienters levevilkår

Hvis vi skulle tale samme sprog, burde vi vel opbygge en tesaurus, som vi alle kunne bruge - på et eller andet niveau. Men det har i lang tid været god latin at mene, at det stort set er umuligt og i hvert fald meget omkostningskrævende at opbygge sådan een. At det ikke er umagen værd - at ikke alle ville kunne bruge den o.s.v. Det kan diskuteres længe - og jeg tror vi her ryger ind i det samme problem som med klassifikationssystemerne: vi synes hver især, at vi er så specielle, at det ikke kan lade sig gøre. Og det er vi muligvis også.

Det der så kunne være lidt mindre vanskeligt kunne være at tage fat på grammatikken i sprog(ene).

Dvs. syntaksen, (sammen)sætninger af ord samt termformerne og morfologien.

Med hensyn til syntaksen - det at lave sætninger - sætte emneordene i relation til hinanden, bruger de fleste post-koordinerede emneord. Eller rettere - der er ikke mange der bruger rent præ-koordinerede indekseringssprog. I de præ-koordinerede systemer sætter indeksøren termene/emneordene sammen, så de danner "subject headings" eller strenge.

Eks.:

- kvinder i Europa: Sverige ("subject heading")
- kvinder - ligestilling - arbejdsmarkedet - Sverige (strenge med rækkefølge)

Her kræves der en syntaks - dvs. nogle regler for, hvordan emneordene skal komme efter hinanden i strengen eller i "subject heading".

Disse typer emneord er der ikke mange der laver - de hænger udviklingsmæssigt sammen med fastformskatalogerne. Selvfølgelig kan man til online-brug faktorer sætningerne, så man kan søge på det enkelte ord.

Men kort om det - det er der ikke mange der bruger konsekvent. Men der er en del der bruger termer eller fraser der ligner "subject headings" dvs. lidt syntaks bliver der brugt rundt omkring.

Samtidig må man sige, at mange af de emneord der bruges i Danmark er sammensatte og mindre adjektiviske fraser (det er lidt syntaks - lidt termformer).

- psykiatriske patienters levevilkår
 psykiatri + patienter + levevilkår
- psykiatriplanlægning
 psykiatri + planlægning
- ligestillingspolitik
 ligestilling + politik
- befolkningsstatistik
 befolkning + statistik
- kvindelige politikere
 kvinder + politik
- forældrevold
 forældre + vold

Kan vi finde nogle fælles regler for, hvilke sammensatte ord og fraser vi vil bruge? DBC har en formuleret politik - og jeg har ikke set andre herhjemme. Men det er et besværligt område, fordi de sammensatte ord er så stor en del af vores almindelige sprogbrug.

Hvis vi slet ikke har nogen syntaks i sproget, så har vi det opsplittede sprog:

Og den syntaks der er tilbage er så: den boolske søgelogik (hvor et normalt *og* er et logisk *eller* - eller hvordan det nu er).

Det sidste der er tilbage, og måske det enkleste, er morfologien:

- ental/flertal
- bestemt/ubestemt
- forkortelser mv.

[FABITA](#) | [Konferencen Søgemuligheder i fremtiden.](#)

Sidst opdateret 4.12.96. Informationer på disse sider må benyttes såfremt kilden angives.