

Informationsordbogen - nogle kommentarer

Af Birger Hjørland

Informationsordbogen. Ordbog for informationshåndtering, bog og bibliotek. 2.udg. Udarbejdet af Jens B. Friis-Hansen, Torben Høst, Poul Sten Larsen & Henning Spang-Hanssen. [Hellerup]: Dansk Standardiseringsråd, 1991. 149 sider. (DS/INF 27).

Dansk Standardiseringsråd har udgivet "Informationshåndbogen. Ordbog for informationshåndtering, bog og bibliotek" i 1991. Jeg skal her give den nogle ord med på vejen og diskutere nogle af dens begreber. Dette må imidlertid ikke opfattes som en egentlig anmeldelse: Jeg har ikke tænkt mig at analysere alle bogens typer af begreber eller analysere den ud fra bredere anvendelsessigter. Min mission er her mere målrettet, at diskutere nogle overordnede problemer. Med hensyn til en egentlig anmeldelse vil jeg henvise til Finn Slentes udmærkede anmeldelse i Svensk Biblioteksforskning, 1992, nr. 2, som jeg i mange henseender er ret så enig med.

Jeg vil først understrege, at informationsordbogen er en værdifuld publikation og udtryk for et grundigt og sobert arbejde. Der er medtaget rigtigt mange ord og foruden ordbogsforklaring får man indbygget oversættelser til engelsk og et engelsk register, der henviser til dansk. Et klyngeregister tjener som oversigt over fagtermernes systematiske sammenhæng.

Status som standard?

Ordbogen skriver selv udtrykkelig side 5: "Det bemærkes, at denne ordbog ikke er nogen standard, men har vejledende karakter...". Denne holdning har jeg også fået bekræftet f.eks. i samtale med Henning Spang-Hanssen. Det skal dog siges, at jeg alligevel ikke finder ordbogens tilknytning til Dansk Standardiseringsråd helt uproblematisk. For det første har jeg ved begrebsdiskussioner med Poul Steen Larsen blevet mødt med argumenter for definitioner, der prioriterer andre "standarder" højere som kilder end f.eks. artikler i videnskabelige tidsskrifter. For det andet får bogen hos visse ledende personer i biblioteks-

verdenen et officielt præg, der nogle gange kan udgøre et (ganske vist mindre) problem, hvis man ikke er enig med ordbogen. Et arbejde som dette skulle gerne åbne op for debat og nyudvikling, ikke lukke i eller blokere på grund af misforståelser omkring dens status. Men denne bebrejdelse er det naturligvis lidt urimeligt at rette imod ordbogen, der så eksplicit siger, at den ikke er en standard. Disse bemærkninger sigter derfor også i højere grad til de personer, der alligevel tilskriver den en særlig status qua dette udgivelsesforhold.

Mere principielt kan jeg sige, at jeg i min egen nyligt publicerede bog "Informationsvidenskabelige grundbegreber" (Hjørland, 1992, side 174-175) skriver under opslagsordet "Standardisering":

"*BDI-sektoren er et område præget af standarder og standardisering. Der eksisterer f.eks. nationale og internationale standarder for vokabularier, principper og koordination af terminologi, karactersæt og translitteration, dokumentbeskrivelse, søgesprog (Common Command Language, CCL), papirformater, reproduktion m.v.

Et standardiseringsarbejde kan betragtes som en form for videnskabeligt arbejde. Som sådant må det søge en balance mellem to tendenser: På den ene side en tendens til vilkårlighed og unødvendig divergens, på den anden side en tendens til stivnen, til en lukken om normer, der virker hæmmende på en faglig og teknisk udvikling. Specielt på stærkt dynamiske områder kan standardisering let virke konserverende.

Nogle områder er mere velegnede til standardisering end andre. Enhver kan se fordelene ved standarder for f.eks. papirformater og alfabetiseringsregler. Mere problematisk er standardisering af grundlæggende videnskabelige begreber af kontroversiel natur, f.eks. begrebet "information". Synspunkter på og argumenter for bestemte definitioner af et sådant begreb er og bør være en del af den videnskabelige debat

i selve forskningslitteraturen. Forsøg på at "standardisere" en sådan terminologi er at udråbe sig selv til en overdommer, der tilmed ikke eksplicit via referencer og analyser behøver forholde sig eksplicit til andre definitioner".

I samme bog diskuterer jeg begrebet "information" side 71-74. Jeg citerer her såvel informationsordbogens definition som en ældre definition af Spang-Hanssen:

"Begrebet "information" benyttes i daglig tale ofte synonymt med "oplysning" eller "meddelelse". Informationsordbogen (DS/INF 27; 1991) definerer information: "Nedfældet eller formidlet viden".

I informationsvidenskaben har man - bl.a. inspireret af Shannon & Weavers informationsteori fra 1949 og videre via bl.a. cybernetikken og det *kognitive paradigme - forsøgt at definere begrebet mere præcist, og afvigende fra ovennævnte betydning. Eksempelvis forlanger Spang-Hanssen (bl.a. 1974, side 49) at tre kriterier bør være tilstede for at et dokument indhold kan kaldes information: Når der tænkes på en bestemt bruger må dokumentindholdet a) Være forståeligt for vedkommende; b) Komme ham/hende ved (være relevant for vedkommende c) Bibringe vedkommende noget nyt....."

Dette eksempel viser, at en af forfatterne til informationsordbogen har skiftet opfattelse med hensyn til definitionen af begrebet "information" - eller i hvert fald ikke har gjort sin tidligere opfattelse gældende. Når der er tale om et begreb af en så teoretisk karakter som bl.a. "information", så rækker en kort forklaring som informationsordbogens ikke så langt. Hvorfor har Spang-Hanssen enten skiftet mening eller bøjet af i dette definitionsspørgsmål? Hvilke begrundelser er der forbundet med at definere termerne på den ene eller den anden måde, og hvilke konsekvenser får definitionerne for den BDI-faglige teori og praksis? Det er jo den slags spørgsmål, der er virkeligt

interessante, når man beskæftiger sig med teoretiske spørgsmål. Det kan en ordbog dårligt give, det får man bedre i et leksikon, hvor der er plads til at referere og diskutere de forskellige definitioner.

Ordbog versus leksikon

Grænserne mellem begrebet "ordbog" og "leksikon" er naturligvis flydende, men kan, som jeg vil prøve at demonstrere her, være ret så markante. Informationsordbogen er typisk en ordbog. Det betyder, som jeg demonstrerede med begrebet "information" ovenfor, at det kun holder sig til en kort definition af ordet, ikke går ind i en længere teoretisk analyse af begrebet, ikke undersøger begrebets forhold til forskellige teoridannelser, praktiske og teoretiske konsekvenser af at definere det på forskellig måde, etc.

Som ordbog er Informationsordbogen værdifuld til simpelthen at forklare svære ord, opløse forkortelser, angive engelske penderter etc. Man kan få oplyst hvad "IFLA" er, hvad "imprimatur" betyder, hvad "ikon" betyder, at "nyhedsblad" er eller kan være synonymt med husorgan, informationsblad, newsletter og nyhedsbrev o.s.v. o.s.v.

En ordbog, der forklarer ord, man ikke kender, kan kaldes en form for "fremmedordbog". Mange mennesker har en tendens til i deres ordlister m.v. at koncentrere sig om ord, hvor man har en klar fornemmelse af noget svært eller vanskeligt. Det at have et "svært" fagsprog virker også afstivende i forhold til ens faglige selvvurdering. Denne tendens står i et vist modsætningsforhold til en mere filosofisk afklaring af de mere dagligdags begreber. Personligt finder jeg det betydeligt mere væsentligt at få afklaret almenbegreberne, f.eks. begreber som dokument, emne, informationssystem, kvalitet, medie, viden etc. end en definition af mere specielle termer, som f.eks. kanalmikrokort, løbeseddel, føljetonroman, postinkunabel, brødskrift, fonogram, melodiudgave o.s.v. Der er imidlertid tale om to forskellige behov: i tilfældet med de almene begreber er behovet især at få

udviklet en adkvat teori og adekvate begreber om fagets grundlag. I tilfældet med de mere specielle termer er behovet især muligheden for at kunne få en forklaring for specielle udtryk, man støder på i sit studium eller arbejde indenfor BDI-området. Forsøget på at udvikle en adkvat teori og et adkvat begrebsapparat forsøger jeg i min "Informationsvidenskabelige grundbegreber", mens informationsordbogen i højere grad er et værktøj til forklaring af specielle ord og udtryk, man støder på i sit faglige arbejde.

Naturligvis er Informationsordbogens ambitionsniveau større end den blotte forklaring af "svære ord", og i mange tilfælde er der taget stof med, der egentligt må kaldes leksikalsk, f.eks. når VINITI ikke blot forklares, men det også oplyses, at den udsender referatidsskriftet Referativnyj Zurnal i 18 sektioner og 200 delserier. Selvom sådanne oplysninger strengt taget er leksikalske, finder jeg det dog naturligt, at de tages med i en ordbog, fordi det typisk er en ret konkret oplysning, som er nyttig i forbindelse med beskrivelse af ordbetydningen, og som ikke involverer mere kontroversielle teoretiske betydnings.

Når det imidlertid drejer sig om BDI-faglige begreber af stor faglig betydning, og hvorom der foregår en løbende debat i den internationale faglitteratur, slår denne ordbogsform ikke til. (Der findes andre ordbogsformer, f.eks. de store nationalordbøger og fagordbøger, der bringer citater fra faglitteraturen, men det er en anden sag, og kan være en overgangsform mellem ordbog og leksikon).

Vi har allerede set på begrebet "information". Nedenfor vil jeg eksemplificere med en diskussion af nogle begreber vedr. emnedata m.v. (Dette eksempel er valgt, fordi jeg i forbindelse med arbejdet i Statens Bibliotekstjenestes emnedataarbejdsgruppe alligevel har måttet analysere Informationsordbogens definitioner på dette begrebskompleks).

Informationsordbogens begreber om emnedata m.v

Emnedata

Informationsordbogen, 1991: "Emnedata [subject representation data] Data som repræsenterer dokumenters emner i søgesystemer, fx emneord eller klassifikationsdata. (Modsat deskriptive data)".

Kommentar: Jeg er ikke enige med Informationsordbogen i, at emnedata og deskriptive data udgør modsætninger, hvilket hænger sammen med en anden opfattelse af begrebet emnesøgning end informationsordbogens. Titelord er således en datakategori, der typisk anvendes i emnesøgning, og derfor vil udgøre såvel deskriptive data som emnedata, d.v.s. emnedata og deskriptive data udgør ikke modsætninger, som informationsordbogen siger, men overlappende grupper.

Ved emnedata forstår jeg/vi (note 1) alle data i en bibliografisk post, der kan repræsentere de emner, som behandles i det registrerede værk (ved analyseposter: dele af et værk). Emnedata omfatter således klassemærker, "subject headings", kontrollerede emneord, ukontrollerede emneord, person- og institutionsnavne som emneangivelser m.v.

Titelord henregnes som sagt også til emnedata vel vidende, at hverken titlen som helhed eller dens enkelte ord nødvendigvis angiver værkets emne (og endog kan være direkte vildledende: "Skraldespanden tur retur" handler ikke om skraldespande, men om psykoterapi). I det omfang posterne indeholder **abstracts og/eller gengivelse af værkets indholdsfortegnelse** m.v., må man behandle ord og udtryk herfra som emnedata. (Emnedata af ovenstående type anvendes typisk i automatisk indeksering, herunder "KWIC", KWAC" & "KWOC"-indeksering, i "SAP-indeksering" m.v.).

Også **litteraturreferencer** kan fungere som emnedata (citationsindeksering), ligesom f.eks. **fuldtekstbasernes tekst**. Som led i en mere omfattende og adekvat teori om informationssøgning bliver sådanne datatyper vigtige at have med i emned-

atabegrebet, ligesom vi må forudse at data, der ikke er knyttet til en traditionel bibliografisk post, men f.eks. via hypertext forbundet med andre informationer, må betragtes som stadig vigtigere for emnesøgningens teori og praksis.

Emnesøgning

Informationsordbogen, 1991: "Emnesøgning [subject retrieval] Informationssøgning ved hjælp af *emnedata".

Kommentar: Ovenstående definition er en smule uklar, idet f.eks. titelord som fremhævet ovenfor ikke er medregnet i Informationsordbogens definition af *emnedata. Titelord anvendes ofte i emnesøgning. Det er således uhensigtsmæssigt at definere emnesøgning ud fra en bestemt type af data. **En emnesøgning vil jeg istedet for definere som en søgning efter på forhånd ukendte dokumenter, der skal bidrage til at tilfredsstille et bestemt informationsbehov.** Til emnesøgning kan også anvendes såkaldt deskriptive data, f.eks. kan en forlagsangivelse "Dansk psykologisk forlag" eller et tidsskriftnavn "Child Psychology" indeholde *implicit emneinformation* og således i nogle situationer spille en rolle i emnesøgning.

Emneanalyse/-beskrivelse

Informationsordbogen, 1991: "Emnebeskrivelse [subject analysis] Undersøgelse, analyse og formaliseret beskrivelse af et dokumentets emne, med den hensigt at muliggøre søgning efter dokumentet ud fra dets emne. Til den formaliserede emnebeskrivelse benyttes klassifikationssymboler eller emneord; man taler henholdsvis om at klassificere eller om at indeksere".

Kommentar: Begrebet emneanalyse er grundigere behandlet i Hjørland (1992a+b) hvori bl.a. standarden "ISO 5963" vedr. emneanalyse er diskuteret [Denne standard omtaler Informationsordbogen mærkeligt nok ikke, selvom standarder iøvrigt foretrækkes som kilder]. Det er vigtigt - som også påpeget af Langridge (1989) & Lancaster (1991) - at skelne mellem selve emneanalysen, der er den intellektuelle eller evt. automatiserede proces,

hvorved et dokument emner analyseres og bestemmes og den efterfølgende "oversættelsesproces", i hvilken et dokument emner udtrykkes i et givent emnedata-system. Den efterfølgende "oversættelsesproces" bør principielt ikke betragtes som en del af emneanalysen, men som en udtryksmåde, der kan være mere eller mindre formaliseret, formuleret i "naturligt sprog", i form af kontroleret vokabular, klassifikationskoder eller andet.

Når man anskuer det på denne måde, forsvinder den grundlæggende destinktion mellem klassifikation og indexering. I praksis spiller selve "oversættelsesprocessen" ofte tilbage på emneanalysen, således at man analyserer emnet, sådan som det anvendte emnedatasystem lægger op til, men hvis man blander tingene sammen på denne måde, kan man bl.a. ikke evaluere et emnedatasystems evne til at udtrykke dokumenters emner eller opstille kriterier for god emneanalyse, der er uafhængig af det anvendte emnedatasystem. Efter min opfattelse skelner informationsordbogen ikke tilstrækkeligt tydeligt mellem selve emneanalysen og selve beskrivelsen i form af eet eller andet emnedatasystem eller "IR-sprog".

Emne

Informationsordbogen, 1991: "I BDI-sammenhæng det/de tema(er), fænomen(er) eller begreb(er) som et dokument kan behandle".

Kommentar: Informationsordbogen definerer begrebet "emne", som dette forstås af lægfolk, men forholder sig ikke til den biblioteks- og informationsvidenskabelige forskning, der har fundet sted i dette århundrede, f.eks. Ranganathans definition af "emne" eller lingvisten Hutchins og andres begrundelse for at foretrække termen "aboutness". En fagterm, herunder "emne", skal defineres, således at definitionen er en rettesnor for det praktiske og videnskabelige arbejde.

Der er meget at sige om emnebegrebet, og det har jeg gjort andetsteds. På dette sted skal kun fremdrages, at et dokument emne som en BDI-mæssig fagterm er det, som en emneanalysen sigter mod at

identificere for dernæst evt. at blive udtrykt i een eller anden form for emnedata. **"Emnet" er altså emneanalysens objekt eller genstand.** Istedet for at hæfte sig for meget ved "dokumentets emne" som en bestemt ting (der evt. afspejler forfatterens hensigt), lægger moderne informationssøgeteori vægt på en analyse af, **hvilke potentielle spørgsmål dokumentet er i stand til at besvare.**

Konklusion

Jeg har ovenfor demonstreret, at jeg med hensyn til visse teoretiske begreber er **uenig** med informationsordbogens definitioner. Det er naturligvis relevant nok at få diskuteret disse specifikke definitioner i sig selv. I denne forbindelse har hensigten imidlertid mere været at få påpeget, at selve terminologiarbejdet ikke kan løsrives fra teoretisk, BDI-fagligt arbejde, at definitionerne ikke eksisterer isoleret, men indgår i en teoretisk argumentation og har store konsekvenser for fremtidigt arbejde.

Eksemplet med emnedatabegreberne viser, at begreberne er tæt sammenvævet med informationsvidenskabelig teori. Hvis begreberne er ufrugtbart defineret, da opstår der en teoretisk eller filosofisk blokering, der hindrer konstruktionen og evalueringen af effektive informationssystemer.

Derfor er det vigtigt, at fremstillingsformen giver muligheder for at præsentere en teoretisk skitse, almindelig praksis på området, kritiske analyser og teoretiske alternativer og ikke mindst: dokumentation i form af litteraturhenvisninger. En sådan fremstillingsform tillader leksikaet eller den concise encyclopædi i højere grad, og derfor arbejder jeg selv med bl.a. denne fremstillingsform. I år har jeg som sagt publiceret "Informationsvidenskabelige grundbegreber" (Hjørland, 1992b), som jeg selv mener dækker et meget stort hul i den BDI-faglige faglitteratur. Det er naturligvis et værk, jeg løbende vil videreudvikle. Hvorvidt det i fremtiden ville være en fordel at samarbejde de to værker til eet, skal jeg ikke her tage stilling til. Foreløbigt frem-

træder de ret så forskelligt og varetager måske også ret forskellige behov.

Noter

Note 1:

Formuleringen her er ikke min egen, men skyldes førstebibliotekar Niels Erik Wille, SBT, der er medlem af det omtalte emnedataudvalg. Dette udvalg vil senere fremkomme med en rapport, der vil afspejle udvalgets opfattelse eller opfattelser.

Referencer

Hjørland, Birger: The Concept of "Subject" in Information Science. **Journal of Documentation**, 1992a, vol. 48 (2), side 172-200.

Hjørland, Birger: **Informationsvidenskabelige grundbegreber**. København: Danmarks Biblioteksskole, 1992b. xiv + 200 sider. Pris: 150 kr + moms + porto.

Informationsordbogen. Ordbog for informationshåndtering, bog og bibliotek. 2.udg. Udarbejdet af Jens B. Friis-Hansen, Torben Høst, Poul Sten Larsen & Henning Spang-Hanssen. [Hellerup]: Dansk Standardiseringsråd, 1991. 149 sider. (DS/INF 27).

Slente, Finn: anmeldelse af Informationsordbogen. **Svensk Biblioteksforskning**, 1992, nr. 2., side 71-76.

Lancaster, F.W.: **Indexing and Abstracting in Theory and Practice**. London: Library Association, 1991. 328 sider.

Langridge, D.W.: **Subject analysis. Principles and procedures**. London: Bowker-Saur, 1989.

Spang-Hanssen, Henning: **How to teach about information as related to documentation**. upubl. manus. 10 sider. Sept. 1970.

Spang-Hanssen, Henning: Kunnskapsorganisasjon, informasjonsgjenfinning, automatisering og språk. (i: **Kunnskapsorganisasjon og informasjonsgjenfinning**. Oslo: Riksbibliotekstjenesten, 1974).