

Online kataloger med indholdsfortegnelser af samleværker III:

Undersøgelse af Roskilde Universitetsbibliotek og samlet konklusion

Af Nils Bredsdorff og Claus Poulsen

Abstract

De forbedrede muligheder for automatisk indlæsning af tekst har bragt muligheden for at lægge dokumenters indholdsfortegnelser i bibliotekernes online kataloger nærmere.

Før bibliotekerne bestemmer sig til at gøre monografiens indholdsfortegnelser søgbare, bør de kende omfanget af katalogets forøgelse - om ikke andet for at kunne udvikle strategier, der gør brugeren i stand til at undgå "informationsdøden".

I denne artikel afsluttes temaet "online kataloger med indholdsfortegnelser af samleværker" med en undersøgelse af antallet af samleværker og antallet af bidrag i samleværkerne indenfor det naturvidenskabelige og det humanistisk-samfundsvidenskabelige område på Roskilde Universitetsbibliotek.

Disse resultater fra et dansk universalbibliotek sammenlignes med resultaterne fra et amerikansk universalbibliotek og et dansk specialbibliotek og resultaterne søges generaliseret.

Indledning

Der er tidligere argumenteret for, at bibliotekernes online kataloger bør gøre titler og forfattere til de enkelte artikler i samleværker søgbare (Poulsen 1992/Online kataloger med indholdsfortegnelser af samleværker I: Biblioteksarbejde, 13 (36), 1992). Effekten er ikke blot et spørgsmål om at tilgængeliggøre skjulte dele af bibliotekernes samlinger, men også at forbedre mulighederne for emnesøgning i online katalogerne (Poulsen 1994).

Men hvad sker der med et online katalog, hvis man begynder at indlæse indholdsfortegnelser af samleværker? Et er sikkert: De bliver større.

Og er det nu helt ubetinget godt, som man kunne få indtryk af ved at læse Paul Holt's indlæg på sidste årsmøde i Danmarks Forskningsbiblioteksforening? (Holt 1994)

Erfaringer med anvendelse af online kataloger viser, at ofte får brugerne alt for megen information ved emnesøgninger - information, de ikke er i stand til at behandle og nyttiggøre. Og hvem er glad for at få over 200 fund, som er det gennem-

snitlige antal fund ved en emnesøgning på et stort amerikansk online katalog, hvis man ikke er i stand til at præcisere sit spørgsmål yderligere? Derfor foretrækker brugeren oftest fritekstsøgning - blandt andet fordi det giver et meget mindre antal fund i gennemsnit end traditionel emnesøgning (Larson 1991).

Ved at forøge mængden af søgbare ord kan man altså risikere, at ødelægge en af de kvaliteter i fritekstsøgningen, der gør den så populær.

Derfor er det vigtigt ikke blot at indlæse indholdsfortegnelser i online kataloger, men også at have en kvalificeret strategi til at reducere antallet af fund, brugeren får gennem en sådan søgning. En sådan strategi kunne bestå i at begrænse antallet af dokumenter, hvis indholdsfortegnelse indlæses i kataloget - for eksempel ved at begrænse sig til at lade denne form for emnesøgning foregå i oversigtslitteratur eller samleværker (Lancaster et.al. 1991 og Poulsen 1994).

Spørgsmålet er herefter: Er denne forøgelse af nu også så stor?

Resumé af hidtidige undersøgelser

For at få et indblik i problemets omfang vil vi idet følgende udelukkende beskæftige os med effekten af indlæsning af indholdsfortegnelser af samleværker.

Hvor mange ekstra titler vil indholdsfortegnelser bidrage med?

Hoffman & Magner har foretaget en undersøgelse i et amerikansk college bibliotek, Santa Ana College Library (Hoffman & Magner 1985).

Dette bibliotek er et universal- og institutionsbibliotek for en højere læreanstalt med 70.000 bind. En stikprøve på 4.094 bøger viste at 871 var samleværker, svarende til 21.3%. Samleværkerne indeholdt i gennemsnit 31.2 titler.

For hver titel i ovenstående bibliotekskatalog gemmer der sig således 6 til i bogmassen. Biblioteket kan tilbyde 600% flere titler at søge i, end i dag, uden at købe én bog mere. Eller: lagt op i et online katalog svarer det til, at der i stedet for

70.000 er omkring 350.000 søgbare titler, stadig uden at købe flere titler.

Dette peger på en ganske voldsom forøgelse i antallet af fund ved en søgning i kataloget.

Nu er spørgsmålet om størrelsen af en sådan forøgelse er alment gældende for forskningsbiblioteker, eller om der er tale om et specielt fænomen for det undersøgte amerikanske college bibliotek.

Foreløbig er der foretaget tre undersøgelser af dokument sammensætningen i et dansk højt specialiseret forskningsbibliotek, nemlig i den pædagogiske specialsamling, bestående af 350.000 monografier indenfor pædagogik samt børne- og ungdomspsykologi, på Danmarks Pædagogiske Bibliotek (Poulsen & Lenshau Teglers 1992).

Første undersøgelse forgik i litteratur, publiceret før 1990.

Antallet af samleværker udgjorde $16\% \pm 3\%$, mens antallet af værker i hvert samleværk blev målt til 20 ± 4 .

Anden undersøgelse tog udgangspunkt i at undersøge om der gennem årene skulle være tale om en ændring af indkøbs- eller publiceringspraksis. Derfor blev litteratur, publiceret i perioden 1990 til primo 1992 undersøgt. Resultatet var, at antallet af samleværker var $17\% \pm 5\%$ og det gennemsnitlige antal værker i hvert samleværk var 17 ± 5 .

Tredie undersøgelse tog udgangspunkt i, at den stigende mængde engelsksprogede kunne forrykke billedet. Derfor blev engelsksproget litteratur, publiceret i perioden 1980 til 1989 undersøgt. Her viste det sig, at samleværkerne udgjorde $15 \pm 3\%$ af samlingen, mens der var gennemsnitlig 29 ± 9 artikler i hvert samleværk.

Det samlede resultat af de i alt fire undersøgelser er, at andelen af samleværker er knap 20% af bibliotekets monografier, og at antallet af titler i hvert samleværk er godt tyve - dog således, at engelsksproget litteratur synes at bidrage med et

overgennemsnitligt (30) antal artikler pr. samleværk.

En undersøgelse af et dansk universalbibliotek: Roskilde Universitetsbibliotek

Indtil nu er der undersøgt et amerikansk college bibliotek og et dansk forskningsbibliotek med en meget specialiseret profil. For at få et brugbart udgangspunkt til vurdering af konsekvenserne af at indlæse indholdsfortegnelser i bibliotekskataloger, må vi undersøge et mere bredt dækkende dansk forskningsbibliotek. Et sådant er Roskilde Universitetsbibliotek med ca. 230.000 monografier og en årlig tilvækst på knap 10.000 bind.

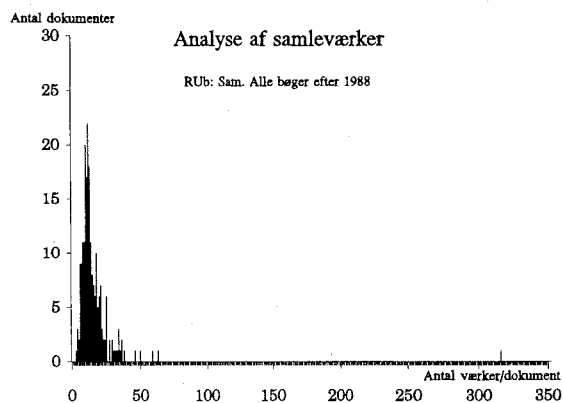
Vi vil se på om der skulle være fagspecifikke karakteristika - om der skulle forskel på antallet af samleværker og antallet af artikler i hvert samleværk indenfor det naturvidenskabelige og det samfundsvidenskabelige område.

For at belyse den aktuelle situation valgte dokumenter fra de to hovedområder, publiceret mellem 1988 og 1992.

Vi har valgt at få statistisk sammenlignelige resultater, hvilket kræver, at der blev valgt lige mange dokumenter (ca.700) ud fra de to hovedområder. Dokumenterne valgte fra kataloget, hvert tiende dokument indenfor det samfundsvidenskabelige område og hvert tredje indenfor det naturvidenskabelige område.

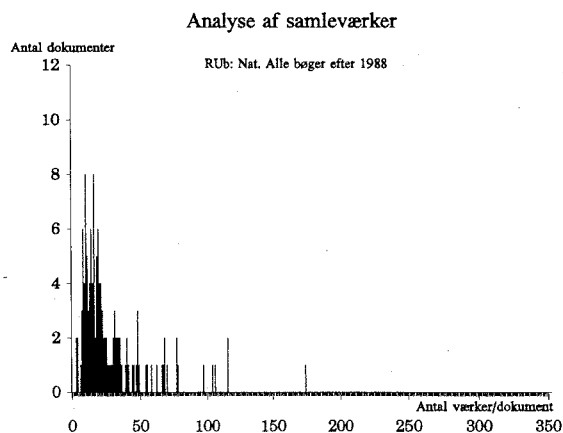
Prøven Figur 1 viser at der for de undersøgte 887 bog-poster fra det samfundsvidenskabelige område er 23,7% samleværker. Det betyder at vi med 95% sikkerhed kan sige, at der er $23,7\% \pm 2,9\%$ samleværker i Rub's samfundsvidenskabelige litteratur, publiceret mellem 1988 og 1992.

Prøven viser også, at hvert samleværk i gennemsnit indeholder $16,7 \pm 3,1$ artikler indenfor 95% grænsen.



Figur 1

Stikprøveanalyse af andelen af samleværker samt antallet af værker i hvert samleværk. Antal prøver: 887. Antal samleværker: 210.



Figur 2

Stikprøveanalyse af andelen af samleværker samt antallet af værker i hvert samleværk. Antal prøver : 698. Antal samleværker: 130.

Prøven Figur 2 viser at der for de undersøgte 698 bog-poster fra det naturvidenskabelige område er 18,3% samleværker. Det betyder at vi med 95% sikkerhed kan sige, at der er $18,3\% \pm 2,9\%$ samleværker i Rub's naturvidenskabelige litteratur, publiceret mellem 1988 og 1992.

Prøven viser også, at hvert samleværk i gennemsnit indeholder $28,0 \pm 4,7$ artikler indenfor 95% grænsen.

Bibliotek	Delmængde		Andel Samleværker	Antal titler pr. Samleværk
	Udvalgskriterier	Antal		
Santa Ana		4.098	21,00%	31
DPB	< 1990	718	16% ± 3%	20 ± 4
DPB	≥ 1990	211	17% ± 5%	17 ± 5
DPB	eng.spr. 80/90	496	15% ± 3%	29 ± 9
RUB	sam 88/92	887	24% ± 3%	17 ± 3
RUB	nat 88/92	698	18% ± 3%	28 ± 5

Tabel I

Resultaterne af de 6 refererede stikprøveundersøgelser på Santa Ana College Library, Danmarks Pædagogiske Bibliotek samt Roskilde Universitetsbibliotek.

Sammenfatning

Resultatet af de nu i alt seks præsenterede stikprøveundersøgelser, samlet i tabel I, viser at antallet af samleværker i gennemsnit udgør ca. 20% på de tre undersøgte biblioteker og at antallet af titler pr. samleværk i gennemsnit på de danske biblioteker er godt 20, mens det er markant højere (ca.30) for den engelsksprogede litteratur vedkommende på Danmarks Pædagogiske Bibliotek og på det amerikanske college bibliotek.

Går vi lidt mere i detaljer ser vi, at den litteratur, der markant bidrager til at øge antallet af titler i samleværkerne, er den engelsksprogede litteratur. Det kunne forklare det høje antal titler pr. samleværk i undersøgelsen af Rub's naturvidenskabelige litteratur, det markant høje antal titler i undersøgelsen af Danmarks Pædagogiske Biblioteks engelsksprogede litteratur, samt den engelsksprogede litteratur i Santa Ana College Library.

Alt i alt viste de foretagne stikprøver at forskningsbibliotekerne kunne forøge antallet af titler med mindst 400% uden at købe én eneste bog ekstra. Bibliotekerne skjuler altså 4 gange så mange titler som de gør tilgængelige via deres online katalog.

Referencer

Hoffman, Herbert H. and Jeruel L. Magner: Future Outlook: Better Retrieval through Analytic Catalogs. *The Journal of Academic Librarianship*, 11(2), 1985: 151-153.

Holt, Paul: Automatiser katalogisering og klassifikation. *DF-Revy 17(8)*, 1994: 217-221

Lancaster, F.W., Tschera Harness Connell, Nancy Bishop & Sherry McCowan: Identifying barriers to effective subject access in library catalogs. *Library Resources and Technical Services*, 35(4), 1991: 377-391.

Larson, Ray. R.: The decline of subject searching: long-term trends and patterns of index use in an online catalog. *Journal of American Society for Information Science*, 42(3), 1991: 197-215.

Poulsen, Claus: Informationens skygge og foran. Informationskvalitet, informationsekspllosion og online kataloger. Institut for Datalogi, Kommunikation og Uddannelsesforskning. Roskilde Universitetscenter. Roskilde. 1994 (Papirer om Faglig Formidling nr 36).

Poulsen, Claus: Online kataloger med indholdsfortegnelser af samleværker I: Kvalitetsforbedring for brugeren ved omlægning af ressourceforbruget. *Biblioteksarbejde*, 13(36), 1992: 31-42.

Poulsen, Claus & Anni Lenschau Teglers: Online kataloger med indholdsfortegnelser af samleværker II: Undersøgelse af Danmarks Pædagogiske Bibliotek. *Biblioteksarbejde*, 13(36), 1992: 43-47.