

Google er Gud – myten om den gnidningsløse informationsøgning

Af Kirsten Larsen

Denne artikel er et forsøg på i nogenlunde kort form at beskrive, hvordan arbejdet med bibliotek.dk foregår og hvorfor vi har valgt de løsninger, vi bruger lige nu. Der findes en webversion¹ af artiklen, hvor nogle af de samme temaer er behandlet med links til afprøvning af nogle konkrete eksempler.

Kirsten Larsen
Afdelingschef, DBC
kl@dbc.dk



Det seneste halve år har der været en del debat om bibliotekskatalogernes udformning² – en herlig ting når man har beskæftiget sig med at lave kataloger i mange år, hvor udformning af katalogen ikke ligefrem har været et hot emne. Til gengæld er det faglige niveau i debatten forbløffende i betragtning af, at vi er netop informationsprofessionelle.

Jeg har selv både beskæftiget mig med formidling og med udformning af kataloger – dels i form af mange års undervisning i pædagogik for informationsprofessionelle, som skal lære studerende på alle niveauer at blive mere informationskompetente – dels i form af funktionalitet og grænseflade i produkter til informationsprofessionelle og slutbrugere, især netpunkt og bibliotek.dk.

Er der i bibliotekerne slet ingen kontakt mellem dem der arbejder med informationskompetence, og dem der arbejder med udformning af katalogerne? Det kan virke som om sidstnævnte øjensynlig mener, at vi kan glemme alt om informationskompetence. Vi skal blot lave et enkelt søgefelt, og så sortere resultaterne, så brugerne altid får det, de søger som de allerførste resultater. Så enkelt er det!

Fremstillingen er måske sat lidt på spidsen, men mange i biblioteksvæsnet har øjensynlig et problem med at respektere deres eget fag – måske fordi de ikke har haft tid til at vedligeholde det – måske fordi det som mange tværgående fag har og har haft problemer med at definere kernen i professionen.

”Hvorfor er der ikke bare ét søgefelt ligesom hos Google?”

”Brugerne skrider på enkle, google-lignende grænseflader.”

”Produktdesignet på bibliotek.dk er alt for komplekst. Med Amazon.com som benchmark, kan bibliotekerne i dag således ikke engang det, Amazon.com kunne i 1994.”

Jeg tror, at alle, der arbejder med grænseflader og informationssøgningssystem, ønsker at give flest mulige brugere det, de ønsker med mindst muligt besvær. Processen er i systematiseret form:

1. Brugeren forsøger at formulere sit informationsbehov på en eller anden måde i den grænseflade, han sidder ved.
2. Brugeren præsenteres for et resultat. Ofte indeholder grænsefladen også hjælp og handlemuligheder fx begrænsning/udvidelse af søgeresultatet, stavehjælp, materiale, der ligner osv.
3. Brugeren finder materiale, som (han mener) løser informationsbehovet.

Den investerede tid og indsats for brugeren fra 1 til 3 skal helst være så kort og ubesværet, som muligt. Parentesen i punkt 3 skal naturligvis antyde en situation, vi alle kender: Man finder noget, der umiddelbart ser brugbart ud, men senere viser sig ikke at være det. Erkendelsen kan være brugerens egen – eller en evt. lærers eller censors, som må anfægte dokumenternes lødighed eller relevans.

Bibliotek.dk har ikke valgt Google-modellen

Hvis vi troede på, at en grænseflade med ét søgefelt gav flest muligt det ønskede hurtigst, VAR der kun ét søgefelt. Det er – indtil videre – vores erfaring, at brugerne ville bruge mere tid på at finde det, de leder efter ved en ”Google-like-grænseflade” end den, vi har valgt.

”Google-modellen” er jo en kombination af ét søgefelt, søgning primært i fuldtekstdokumenter og ranking af resultaterne efter en algoritme, hvor blandt andet antal links til dokumenterne bestemmer rankingen. Det giver ofte et godt resultat, men det afhænger naturligvis af, hvor god brugeren er til at formulere sit informationsbehov, hvilken type dokumenter man leder efter og hvor mange dokumenter, der findes og kan bruges.

På konferencen ”Brugernes bibliotek” (januar 2007) gav Thomas Madsen-Mygdal³ en illustration af Google-søgemetodens (og rankingens) begrænsninger (uden dog at omtale det som et Google-problem). Hvis man leder efter Tor Nørretranders i Google for at få links til bibliotek.dk, så er Google ikke god nok – 3. februar 2007 er det første bibliotek.dk-link hit nr. 18 på Google.dk. Der er da nogle gode links til noget, der har med Tor Nørretranders at gøre, men hvis nu man var ude efter hans nyeste bog, er den der ikke. Det er derimod hit nr. 1, hvis man skriver ”tor nørretranders” i forfatterfeltet i bibliotek.dk (og nr. 2 hvis man skriver det i fritext). Bogen kan naturligvis også findes i Google, hvis man ved hvad den hedder – ellers skal man febr. 2007 blade hen til hit nr. 75.

¹ www.danbib.dk/index.php?doc=bibdkdebat

² *Summa - et bud på fremtidens søgesystem*. Af Birte Christensen-Dalsgaard. I: DF Revy. - Årg. 29, nr. 8, (2006). - S. 6-8, 10-11

³ <http://bootstrapping.net/wp-content/uploads/2007/01/tmmbrugernesbibliotek.pdf>

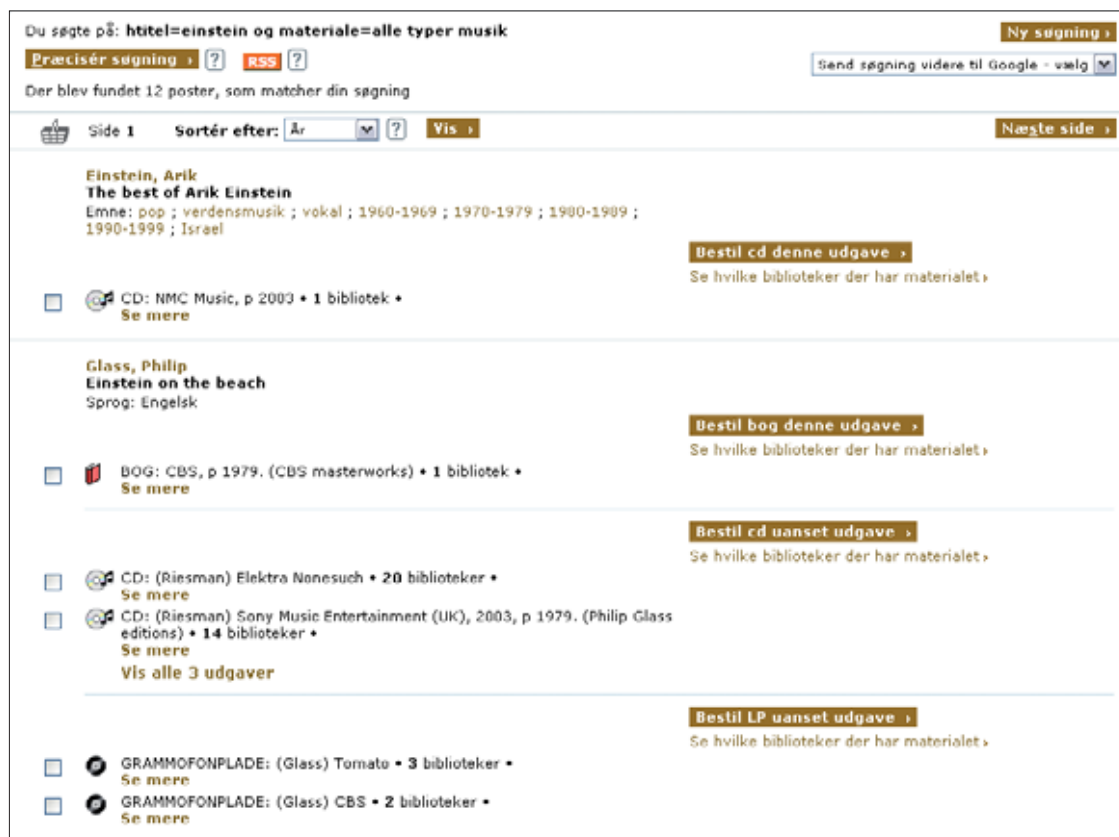


Fig.1 Søgning på musiktitel i bibliotek.dk

Man tilfredsstillter flest hurtigst ved straks at give mulighed for at præcisere sin søgning. Det kan man, hvis man har strukturerede metadata

Man kan godt i bibliotek.dk søge ”som i Google” – der er et fritestfelt, men det er ikke placeret som det første og eneste felt, fordi det ofte giver et stort søgeresultat. Og vi har så ingen viden om, hvad det er, brugeren ønsker, da der jo i fritest kan skrives alle typer data fra emner, ord i titler til isbn eller forfatternavne.

Google laver naturligvis også andre søgemuligheder, fordi de godt ved, at søgningen i det ene felt langtfra altid er lykken. Der er en ”avanceret søgning” (som dog ikke giver muligheder, som bibliotek.dk), og så findes der Google Scholar, Google Books osv. Google har valgt disse to grænseflader som dem, de lægger biblioteksmaterialer ind i. Når man også finder bibliotek.dk-links via Google, er det fordi vi (og mange andre) laver websider, der indekseres af Google – i bibliotek.dk’s tilfælde Emneoversigten.

Går man ind i Google Books og markerer, at man vil søge i bibliotekskataloger og bruger den bibliotek.dk-agtige grænseflade til at søge på forfatteren ”tor nørretranders”, ja så vælter titlerne frem med links til bibliotek.dk. Folkene på Google har altså samme opfattelse, som ligger bag de fleste bibliotekskataloger verden over: Man tilfredsstillter flest hurtigst ved straks at give mulighed for at præcisere sin søgning. Det kan man, hvis man har strukturerede metadata.

Man kan undre sig over, hvorfor der er tradition for at kalde formularsøgesider for avanceret søgning. Det kræver jo langt mere at søge præcist, hvis man kun har ét søgefelt, og derfor kan blive nødt til at bruge søgelogik og søgekoder, hvis man vil have et præcist svar. Læs evt. mere om bibliotek.dk i Google⁴.

Der er i øvrigt ganske mange der alligevel finder de danske biblioteker via Google. Af dem, der kommer ind på bibliotek.dk (januar 2007), kommer 20 % fra Google.dk og godt 5 % fra andre Google-varianter.

Hvornår er ”specialiserede grænseflader” relevante?

Vi bruger alle specialiserede grænseflader mange andre steder – og undres ikke over det. Når vi skal finde et tog til Hobro, går vi ikke til Google og skriver Hobro 3.3. kl. 9.45 og forventer at få et brugbart svar, ligesom vi også accepterer, at når vi skal finde et billigt fly til Malaga, skal vi ind i en særlig grænseflade. Også søgesider som ”Krak” og ”De gule sider” bruger specialiserede søgegrænseflader.

Biblioteksgrænseflader minder på mange måder om sådanne databaser eller andre specialiserede baser som fx beskrevet i den generelle

wikipedia beskrivelse af Vertical search⁵.

Det der er karakteristisk for biblioteks-baser og mange andre baser er: hovedparten af data er strukturerede ”metadata” og ikke fuldtekstdata. Det ville da være dejligt, hvis alt var digitaliseret, men grænsefladen må indrettes ud fra den virkelighed vi står i nu. Og vi har ikke adgang til et redskab i stil med ”linkpopularitet”!

”Relevansrangering”

Hver gang man laver en sorterings- eller rangeringsalgoritme, vælger man naturligvis, hvad man synes er vigtigst for brugeren. Ingen af os kan vide, hvad der er relevant for en bruger, som skriver et eller flere ord i et fritestfelt. Man kan vælge ”forståelige kriterier” som: dét der er hjemme vises først, det der er nyest kommer først, posterne kommer alfabetisk efter titlen e.l. Eller man kan vælge rangeringer, hvor brugeren ikke umiddelbart kan gennemskue eller ændre på rangeringen som fx Googles rangering efter blandt andet hvor mange, der linker til den enkelte side.

Skriver en bruger ”einstein” (for nu at blive ved Tor Nørretranders’ nyeste bog) i bibliotek.dk’s fritestfelt, kan det være for at finde

⁴ www.danbib.dk/index.php?doc=bibdkgoogle

⁵ http://en.wikipedia.org/wiki/Vertical_search

Basket (0) [Show](#)

Search for: einstein

1 to 20 of 12424. Previous 1 [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) ... [622](#) [Next](#) ▶

Sort by: [Year](#) | [Relevance](#)

	Title	Author	Year	Type	Basket
Institute Only	Einstein : (Transl. from the Russian by V. Talmy)	B. Kuznetsov	(1970)		
Institute Only	Einstein	Jeremy Bernstein	1978		
Institute Only	Einstein : His life and times	Philipp Frank, Shuichi Kusaka	1947		
2-8 weeks	Einstein	Jeremy Bernstein	1987		
2-8 weeks	Einstein : the life and times	Ronald W. Clark	1979		
2-8 weeks	Einstein	Peter Michelmore	1966		
Available	Einstein : hans liv og univers	Jim Breithaupt	2001		
Institute Only	Einstein : the life and times	Ronald W. Clark	1994		
Available	Einstein : as myth and muse	Alan J. Friedman 1928-, Carol C. Donley	1990		
2-8 weeks	Einstein : The life and times	Ronald W. Clark	1979		

Limit your search

▼ **Authors**

- [the pennsylvania state university citeseer archives](#) (1234)
- [einstein alfred](#) (110)
- [einstein albert](#) (94)
- [adhikari sadhan k](#) (39)
- [herbert h einstein](#) (38)
- [ketterle w](#) (34)
- [massachusetts institute of technology dept of civil and environmental engineering](#) (34)
- [csorgo t](#) (33)
- [ueda masahito](#) (32)
- [de andrade l c garcia](#) (32)
- [stoof h t c](#) (29)
- [kevrekidis p g](#) (29)
- [stringari s](#) (28)
- [visser matt](#) (28)
- [rendall alan d](#) (25)

▼ **Material types**

- [Internet document](#) (11822)
- [Open Archive](#) (11822)
- [Text](#) (9890)

Fig. 2 Søgning på einstein i Summa

netop Nørretranders bog. Det kan også være for at finde bøger af Einstein, og er det en skoleelev, der skriver, vil Nørretranders bog næppe være det mest relevante. Det giver 1813 hits (febr. 2007), og det er unægtelig en blandet landhandel. For et par år siden lavede vi undersøgelser af muligheden for at rankere efter emne; det resulterede i sorteringsmuligheden "sortér efter emne", men det er en tilvalgsmulighed, da vi ikke kan vide om brugeren søger efter emne, hvis de søger fritekst.

Grænsefladerne i bibliotek.dk indbyder brugere til at meddele lidt mere end blot "einstein" – dels ved valg af grænseflade dels ved at give mulighed for at skrive, hvilken slags oplysning "einstein" er. Vælger man musiksøgesiden og skriver det ind i "titel på cd" får man 12 hit – overskueligt, hvis det nu var Philip Glass' "Einstein on the beach", man var ude efter (se fig. 1).

Halvdelen af søgningerne på bibliotek.dk sker på en anden grænseflade end defaultgrænsefladen. Da det aldrig er let at få brugere til at vælge noget andet end default, må det være fordi, de kan se, at det er nyttigt at gøre det i en lang række søgesituationer.

Statsbiblioteket har forsøgt sig med en "relevanssortering" som default-sortering. Den kan sikkert give en fornuftig sortering af nogle søgeresultater, men generelt er det svært at se det som overbevisende – her igen søgningen på einstein (se fig.2).

DBC havde for en del år siden et udviklingsarbejde omkring en lignende rankeringsalgoritme, men fandt det ikke overbevisende til anvendelse i baser som DanBib på det tidspunkt.

Det andet alternativ kunne være at vælge en sortering efter, hvad der er mest "populært". Hvilket vil betyde, at uanset hvem man er og hvad man leder efter, vil man altid få præsenteret de samme hits øverst. Jeg har svært ved at se nytten af det generelt, men det kan da forsøges som et tilbud (når vi får adgang til en central base over udlånsdata).

I dag har brugeren i bibliotek.dk visse muligheder for personalisering som fx at ønske som default kun materialer på dansk og kun fra folkebiblioteker og ikke lydbøger m.v. Det reducerer "einstein"-søgningen til 295 poster og det kan nok for mange brugere være en fornuftig default-indstilling. Vi fortsætter med udvikling af flere personaliseringsmuligheder. Ligesom vi arbejder

videre med andre former for hjælp, når brugeren er kommet til visningen af første resultat: Find titler der ligner, Søg litteratur om forfatteren, Stavehjælp, Præciser søgning, Zoom, Best-match osv.

Værkvisning – se hvor det er hjemme – værkbestil

På en række områder kan vi ikke gå til hverken Google eller Amazon for at finde inspiration til løsninger. Vi skal også fokusere på, at den egentlige ydelse er den vare, som bibliotekerne leverer – helt i Library 2.0's ånd. Derfor har der i hele bibliotek.dk's levetid været arbejdet med områderne værkvisning og –bestilling samt let adgang til at se, hvor materialet er hjemme.

Arbejdet er på ingen måde "færdigt", men vi er kommet så langt, at vi kan samle hovedparten af udgaver (uanset materialetype) i én værkklynge, brugeren kan ved ét klik se udlånsstatus på sine favoritbiblioteker, og brugeren kan sende en bestilling på et "værk" til sit bibliotek, så vi i så høj grad som muligt udnytter, hvis det valgte bibliotek har én eller anden udgave af værket. Og om et par måneder går projektet "automatisering af bibliotek.dk-bestillinger" i drift – først og fremmest så brugerne hurtigere får det, de ønsker, men

Folkene på Google har altså samme opfattelse, som ligger bag de fleste bibliotekskataloger verden over: Man tilfredsstiller flest hurtigst ved straks at give mulighed for at præcisere sin søgning.

Du søgte på: **forfatter=nørretranders og titel=mærk verden** Nyt søgning >

Præcisér søgning > ? RSS ? Send søgning videre til Google - vælg ▾

Der blev fundet 19 poster, som matcher din søgning

Side 1 Sortér efter: År ▾ ? Vis > Næste side >

Nørretranders, Tor
Mærk verden : en beretning om bevidsthed
 Emne: videnskab ; filosofi ; psykologi

Bestil bog uanset udgave >
 Se hvilke biblioteker der har materialet >

 BOG: Gyldendal 1991 • 1 bibliotek •
 Se mere

 BOG: 3. udgave, 2. oplag, Gyldendal, 2005 • 47 biblioteker • 
 Se mere
Vis alle 8 udgaver

 LYDBOG (BÅND): Danmarks Blindebibliotek, 1993 • 1 bibliotek •
 Se mere

Forbeholdt lånere på Danmarks Blindebibliotek: Slå op

Fig. 3 Søgning på *nørretranders* i *bibliotek.dk*

Der er ikke én rigtig udformning af en katalog, og især er der ikke én løsning, der holder i ret lang tid.

også med den effekt, at bibliotekerne får rationaliseret arbejdet med bestillinger, og landets samlede materialebestand udnyttes bedre⁶ (se fig. 3).

Brugerinvolvering

En stor del af bibliotek 2.0-debatten handler jo om brugerinvolvering. Indtil videre har brugerne mest været involveret i bibliotek.dk i form af brugerundersøgelser og inddragelse i udvikling af funktioner og grænseflade. Der er givet delte meninger om brugerindvolvering i biblioteks-kredse – se opsamlingen af høringssvar i referat fra møde om udviklingsplanen⁷. Men der er afsat ressourcer til forsøg på Udviklingsplan for 2007. Brugere af bibliotek.dk er særdeles kommunikationsivrige. Det har vi glæde af både ved brugerundersøgelser, udvikling af ny funktionalitet og henvendelser om store og små ting i det daglige. Derfor ville det være spændende, om de også kunne berige bibliotek.dk på anden vis, fx i form af ratings, indholdsbeskrivelser, emneord m.v. Det ville også kunne give et input til, hvordan vi udvikler metadata, så de bliver mere rettet mod de slutbrugere, som i dag er den største brugergruppe i modsætning til dengang, hvor

katalogen var bibliotekarens. Læs evt. artiklen om "*Paradise lost and Paradise Regained?*" om hvordan brugertagging også kan virke tilbage på den professionelle indeksering, som gerne skulle have fokus på berigelse mere end deskriptiv katalogisering⁸.

Hele Library 2.0-debatten har i høj grad også været medvirkende til at bringe katalogen i fokus – det er adgangen til en kerneydelse. Der er ikke én rigtig udformning af en katalog, og især er der ikke én løsning, der holder i ret lang tid. Brugernes informationskompetence udvikler sig, og nye redskaber gør det muligt at lave andre løsninger end hidtil. Men de skal stadigvæk vurderes i forhold til nytteværdien for slutbrugere – en løsning, der er fin til ét formål, er det ikke nødvendigvis til et andet.

Man kan undre sig over, hvorfor der er tradition for at kalde formularsøgesider for avanceret søgning. Det kræver jo langt mere at søge præcist, hvis man kun har ét søgefelt, og derfor kan blive nødt til at bruge søgelogik og søgekoder, hvis man vil have et præcist svar...

Litteratur

Kampen om søgesystemerne. Af Birgitte Dalgaard, interview med Kasper Løvshall, Jens Hofman, Erik Thorlund Jepsen. I: Bibliotekspressen. – 2006, nr. 19. – S. 6-8.

Fra søg til find – Primo i Danmark. Af Jørgen Madsen. I: DF Revy. – Årg. 30, nr. 1 (2007). – S. 6-8.

Fremtidens søgeværktøjer. Af Erik Thorlund Jepsen. I: Bibliotekspressen. – 2007, nr. 2. – S. 18-19.

⁶ www.danbib.dk/index.php?doc=autfjern

⁷ www.bs.dk/content.aspx?itemid={86F4724C-1DE5-42BA-8B41-7E20EBB314C7}

⁸ Karen Markey: The Online Library Catalog
 www.dlib.org/dlib/january07/markey/01markey.html