

Forskellige syn på informationssøgning giver forskellige søgesystemer

Af Birte Christensen-Dalsgaard og Jens Hofman Hansen

Gnidningsløs informationssøgning er måske en myte – men i vores udvikling af fremtidens søgesystemer er det essentielt, at vi sætter brugernes forventninger og adfærd i centrum. Det er nemlig nødvendigt, hvis vi vil realisere den strategi, vi har for at få biblioteket til at spille en aktiv rolle i brugernes litteratursøgning.

Birte Christensen-Dalsgaard
SB
bcd@statsbiblioteket.dk

Jens Hofman Hansen
SB
jhh@statsbiblioteket.dk



I sidste nummer af DF Revy gav Kirsten Larsen¹ sit input til en efter vores mening meget sund debat om fremtidens søgesystem – og specielt hvad det er, man kan, skal eller vil understøtte. Artiklen synes dog at afspejle en vis irritation over, at den traditionelle søgegrænseflade bliver kritiseret fra flere sider, selvom bibliotek.dk besøgsmæssigt er en succes.

Vores mål med denne artikel er ikke at kritisere bibliotek.dk, for selvfølgelig tilfredsstillende det en stor brugerskares behov. Vores mål er at nuancere debatten ved at operere med forskellige målgrupper og forskellige behov, hvor det ikke er et ”enten eller” men et ”både og”, hvor målet er at komme frem med ideer, som kan føre til en bedre og mere nuanceret betjening af vores brugere, og hvor traditionelle bibliografiske dyder som viden om ressourcer og kvalitet bringes i spil i forhold til brugerens valg.

Hvad vil vi opnå, og hvordan arbejder vores brugere?

Efter vores mening er det vigtigt, at man, inden man påbegynder et design, gør sig nøje overvejelser over, hvem ens målgruppe er, og hvad man vil opnå. Dette lyder måske trivielt, men accepterer vi det, har det store konsekvenser for, hvordan vores interface kommer til at se ud – og bibliotekarer vil i øvrigt sandsynligvis ikke have samme favoritinterface i deres hverdag, som brugerne.

SUMMA er udviklet til drive-in-brugeren, som vi i feltstudier² har kunnet iagttage i form af brugere, der anser sig selv for at være informationsspecialister. Forventningen blandt disse brugere er et enkelt søgefelt og et meningsfuldt resultat. Uanset hvordan vi eller andre konstruerer vores system, vil det aldrig blive perfekt. Men vi accepterer, at brugernes forventninger, adfærd og tankegang i forhold til søgning er, som den er – sådan en accept er hele kernen i brugercentreret design. Det system, som vi designer, tager afsæt i denne viden, og vi konstruerer systemet på en måde, så det bedst opfylder bibliotekets strategi for formidling. For der er vel at mærke ikke tale om, at vi forfladiger informationssøgning eller gør brugerne dumme – vi ser blot på deres adfærd og forventninger og indretter vores formidling efter den.

En anden brugergruppe er biblioteksentusiastene, som kender og kan bruge bibliotekets tilbud om præcise, nuancerede søgninger. Disse er ikke målgruppen for SUMMA; de bliver jo allerede betjent, som de gerne vil. De får fortsat stillet en udvidet søgegrænseflade til rådighed. Der er i øvrigt intet i SUMMAS søgeteknologi, der forhindrer udvidede søgegrænseflader – men det er blot ikke fokus for vores udvikling, da det ikke er vigtigt for vores højt prioriterede drive-in-brugere.

For at kvalificere diskussionen yderligere vil vi i det efterfølgende skelne mellem den grænseflade, hvor søgninger formuleres, og den efterfølgende selektionsproces i søgeresultatet.

Konstruktion af søgestreng i forskellige brugssituationer

Formålet med søgefelterne i et interface som bibliotek.dk er at hjælpe brugeren til at formulere en søgning, så den ideelt set resulterer i præcis det, som brugeren leder efter. Kan man ikke formulere sig præcist nok, kan man ende med flere poster og skal gennem en selektionsproces.

En sådan virkemåde giver mening, når vi forestiller os en bruger, der gennemfører en verifikativ søgning – dvs. en søgning hvor brugeren, allerede inden han begynder at søge, ved, hvilket materiale han skal finde og bestille. Men vores feltstudier peger på, at denne adfærd ikke er den eneste blandt vores studerende og forskere. Brugere har nemlig også en eksplorativ adgang til informationssøgning, hvor de har en klar idé om, hvilken opgave de skal løse – men de ved blot ikke præcist, hvilke titler af hvilke forfattere de har brug for. Da vi som bibliotek har et strategisk ønske om at deltage aktivt i hele brugernes litteratursøgningsproces (og ikke kun den afsluttende, verifikative del), så skal vores system selvfølgelig understøtte denne brugeradfærd.

Bibliotekarer vil i øvrigt sandsynligvis ikke have samme favoritinterface i deres hverdag, som brugerne.

¹ *Google er Gud – myten om den gnidningsløse informationssøgning.* Af Kirsten Larsen. DF Revy Årg 30, nr. 2, marts 2007, s. 16 - 19

² *Det hybride bibliotek set med brugerne øjne.* Af Jeppe Lomholt Akselbo, Lise Arnfred, Sten Barfort, Gina Bay, Tine Bagger Christiansen, Jens Hofman Hansen, Henrik Tvermoes Jensen, Gitte Bach Markussen, Anna Mette Morthorst og Michael Poltorak Nielsen. Feb. 2006. Set 26/3 2007 på: www.statsbiblioteket.dk/publ/feltstudier.pdf

Vores mål er at nuancere debatten ved at operere med forskellige målgrupper og forskellige behov, hvor det ikke er et ”enten eller” men et ”både og”, hvor målet er at komme frem med ideer, som kan føre til en bedre og mere nuanceret betjening af vores brugere, og hvor traditionelle bibliografiske dyder som viden om ressourcer og kvalitet bringes i spil i forhold til brugerens valg.

Det er klart, at de fleste bibliotekssystemer er en smart online udgave af det traditionelle kortkatalog, som er optimeret til det formål at lokalisere og bestille materialer. Af samme grund vil brugertests typisk fokusere på at optimere denne proces. Derfor kan vi også se, at der over årene er kommet flere og flere søgefelter på forsiden af bibliotek.dk, som tillader større og større præcision tidligt i interfacet, og det giver god mening, da systemet netop skal understøtte lokalisering af kendt materiale, – en brug, som også undersøges fra f.eks. Californien³ bekræfter. Her siges der direkte: *”The current Library Catalogue is poorly designed for the tasks of finding, discovering, and selecting the growing set of available resources... it is best at locating and obtaining a known item”*. Vores feltstudier viste også denne opførsel, men fik kvalificeret brugerens opførsel, nemlig at materiale identificeres ud fra f.eks. Google eller Amazon, før det bestilles. Det er derfor klart, at mange Google-brugere også er biblioteksbrugere.

En anden ting, som vi fik ud af at undersøge, hvordan brugerne finder frem til information, er viden om, hvordan tillid hænger sammen med deres faglige netværk. Er man forsker, har man tillid til sine medforskere på tværs af kloden, og er man studerende, har man især tillid til sin vejleder. Disse sociale konstruktioner er centrale at indbygge i et system, som skal levere information med høj troværdighed. Et system, der giver en tungere vægt til materiale, som relaterer sig til den af vejlederen anbefalede litteratur, vil ganske givet opleves mere relevant af de studerende.

Et helt afgørende element i at finde litteratur er også at se på referencer til og fra litteraturen. Det er et element, som slet ikke er afspejlet i de danske bibliotekers systemer. Når man ser på, hvor meget denne semantik betyder for brugerne, virker meget arbejde med at fin-tune interfaces pludselig irrelevant. En rationel fokusering af bibliotekernes ressourcer ville helt klart betyde, at vi fik forhandlet ret til at bruge eksempelvis citationsindekser i de forskellige søgesystemer.

Når bibliotek.dk i øjeblikket kritiseres så kraftigt, er det fordi nogle biblioteksfolks og brugeres forventninger til online biblioteket grundlægges andre steder på nettet, hvor man kan se, at it-systemer også kan understøtte eksplorativ søgeadfærd. Forsøgene i bibliotek.dk har været at arbejde med emneord – men dette er kun en lillebitte og relativt uvæsentlig del af de store muligheder, der er for at understøtte det eksplorative.

Støtte til valg af materiale

Resultatet af en søgestreng er traditionelt set en liste over materiale, som på den ene eller anden måde matcher søgestrengen. Det er klart, at jo større mængde materiale, der søges i, jo vigtigere bliver det at sikre, at brugeren gives rigtig god hjælp til at udvælge det, som er relevant.

I analysen af søgeresultatet er det igen relevant at skelne mellem den verifikative og den eksplorative tilgang til søgning. I ingen af tilfældene er brugeren villig til at kigge mange sideres søgeresultater igennem. Sorterer vi søgeresultatet efter årstal eller anskaffelsesår, vil

det give mening for nogle brugere, mens det for andre vil virke meningsløst og helt irrelevant. Af den grund arbejder vi med at rangordne efter relevans.

Man kan med rette spørge, hvad der definerer relevans. I en tidligere artikel om SUMMA skrev vi⁴: ”I SUMMA bruger vi ordet ”relevans” til vores rangordning. Men hvad er relevans? Vi ser som udgangspunkt, at relevans er et produkt af tre faktorer: Hvor godt matcher posten en søgning, hvor god er kvaliteten af det materiale, som repræsenteres af posten og endelig, hvor godt matcher materialet brugerens behov?”

Relevans er altså et relativt begreb, som afhænger af personen og af sammenhængen. Men at kunne give et meningsfuldt svar bliver vigtigere og vigtigere, efterhånden som mængden af materiale, som man opererer på, vokser. En alfabetisk eller kronologisk liste giver mening, når der er 25 svar, men giver ikke mening, når der er 5.000 svar. Da det ikke er en farbar strategi at lægge den byrde, det er at reducere søgeresultatet til 25 eller færre hits på brugeren, er konsekvensen at arbejde aktivt med en relevans-rangordning i søgeresultatet. Desuden betragter vi i arbejdet med SUMMA ikke søgeresultatet som en simpel liste – det giver større mening at betragte søgeresultatet som en informationsstruktur, der indeholder en liste og nogle interaktionsmuligheder, som lader brugeren omstrukturere listen, begrænse listens omfang, etc.

Det er i øvrigt værd at bemærke, at så snart vi forlader princippet om, at et søgeresultat er en simpel liste sorteret efter år eller forfatter, går vi

Når bibliotek.dk i øjeblikket kritiseres så kraftigt, er det fordi nogle biblioteksfolks og brugeres forventninger til online biblioteket grundlægges andre steder på nettet, hvor man kan se, at it-systemer også kan understøtte eksplorativ søgeadfærd.

³ *The University of California Libraries: Final Report* Dec. 2005 Bibliographic Services Task Force set 26/3 2007 på <http://libraries.universityofcalifornia.edu/sopag/BSTF/Final.pdf>

⁴ *Summa – et bud på fremtidens søgesystem*. Af Birte Christensen-Dalsgaard. DF Revy Årg. 29, nr. 8, december 2006, s. 6-8, 10-11

Lad os se på brugeradfærd og brugerforventninger og ud fra denne basis på, hvordan vores strategi kan realiseres.



bevidst over til en rangordning, hvis algoritmiske virkemåde reelt er uigennemskuelig for brugeren. I Google ved de fleste godt nok, at en række parametre (f.eks. match med queryen og antal indgående links til websitet) er afgørende for, hvor højt det kommer i søgeresultatet. Men ingen (udover Googles ingeniører) kan redegøre for, hvorfor et website ligger som nr. 7 og et andet som nr. 8 i søgeresultatet. Det interessante er, at brugeren ikke behøver at kunne gennemskue algoritmen for at være trykke ved at bruge systemet – så længe relevansrangordningen virker anvendelig i langt de fleste brugstilfælde.

Men hermed er ikke sagt, at vi som systemdesignere er tilfredse. Arbejdet med relevans i SUMMAS søgeresultater er slet ikke færdigt – men vi er allerede nået et nødvendigt stykke hen ad vejen. For yderligere at understøtte den eksplorative søgning er der brug for materiale og informationer, som i dag ikke er en del af bibliotekskatalogen. Det er nødvendigt at vide noget om niveauet af en bog – er den for eksperter, eller kan den læses af en gymnasieelev? Det kan vi f.eks. sige noget om, hvis vi har information om, hvilke biblioteker der har bogen stående, som gjort af OCLC (de kalder det "audience level"). Det vil være en fordel med tilhørende information, som siger noget om indholdet, og det vil være rigtig godt at vide, hvordan andre ser på bogen. Er den klar i sin fremstilling, eller indeholder den mange illustrationer, som kan inspirere til en

grafisk fremstilling af resultater? På Statsbiblioteket har vi desuden en strategi om "digital fremskudt betjening", hvor vi meget gerne ser vores søgefelt indlagt i relevante omgivelser, som f.eks. universitetets e-læringssystem. I den forbindelse kan vi arbejde med højere relevans i søgeresultaterne ved, at vi kender til brugermassen og den litteratur, som f.eks. står på semesterhylderne. Vi får pludselig rigtig meget information, som kan bruges i en efterfølgende rangordning.

Da vores system er gearret til at arbejde med relevans, er det ikke så meget en diskussion af, om vi skal gøre det, men snarere en biblioteksfaglig og usabilityfaglig diskussion af, hvordan vi bruger data til relevansrangordning. Folkebibliotekerne kommer også til at benytte SUMMA til deres lokale bibliotekspræsentationer. For deres brugere vil man atter kunne forestille sig datagrundlag, som kan gøre de lokale præsentationer bedre målrettet til deres brugere.

I både den eksplorative og verificative fase vil brugeren være interesseret i at præcisere sin søgning. F.eks. bliver brugeren klar over, at det er en bestemt materialetype, vedkommende er interesseret i – men brugeren kan ikke altid på forhånd med mening formulere, om det er den ene eller den anden materialetype. Ved at flytte disse valg et trin hen i søgeprocessen og præsentere mulighederne ved siden af søgeresultatet sikrer vi os, at brugeren kun kan vælge ting, der giver mening

i forhold til den opgave, han er i gang med. Brugeren kan f.eks. kun vælge de materialetyper, som er indeholdt i den aktuelle søgning, og bliver således ikke på baggrund af dette søgekriterium ramt af 0-hit-sider.

Med sådan et design understøtter vi især den eksplorative tilgang. Man kan kalde det motiverende design, hvor brugeren præsenteres for informationer og muligheder på tidspunkter, hvor systemet kan identificere, at disse sandsynligvis er relevante for brugeren, frem for at alle muligheder præsenteres, før vi ved noget som helst om brugerens behov og adfærd.

Ingen søgesystemer er gode nok

Det skal slutteligt siges, at ingen bibliotekssøgesystemer endnu har leveret en relevansrangordning, som vi synes er tilfredsstillende – og det er derfor, at vores arbejde med SUMMA og relevansbegrebet er meget højt prioriteret på Statsbiblioteket. Vi vil meget gerne diskutere, hvordan vi kan tilpasse relevansalgoritmer fremadrettet, og vi vil meget gerne samarbejde med institutioner, virksomheder og brugere, der kan levere data, der kan bidrage til øget relevans i søgeresultaterne. Vores rettesnor vil hele tiden være: Lad os se på brugeradfærd og brugerforventninger og ud fra denne basis på, hvordan vores strategi kan realiseres.

