

RANKING, BRUGERDATA OG MATERIALEINDKØB HØRER SAMMEN

Du undrer dig måske over hvilken sammenhæng der er mellem Google og dine indkøb af elektroniske licenser? Faktum er, at ranking efter relevans, empiriske brugerdata og udviklingen af vores lokale e-samlinger kommer til at hænge uløseligt sammen.

AF ROBERT BLEY, ROBERT.BLEY@EXLIBRIS.CO.UK, EX LIBRIS UK

Lad os starte med at se på hvilke udfordringer, som de danske fag- og forskningsbibliotekerne må se i øjnene, i en elektronisk verden med øget kompleksitet:

Indkøb af pakker vs. at udvælge indhold: Aggregerede pakker er attraktive, da de giver mange varer på hylderne til en relativ lav pris. De kan også – i noget omfang – lette bibliotekernes administrative besvær ved at reducere antallet af udgivere, som bibliotekerne skal besværes med.

Adgang til Open Access vs. at betale for indhold: Kvantiteten så vel som kvaliteten af gratis forskningsresultater øges dag for dag. Ikke bare er mange Open Access e-tidsskrifter og e-bøger nu frit tilgængelige, men tusinde af institutionale eller andre repositories giver gratis adgang.

“Just-in-time” vs. “just-in-case”: Når brugerne efterspørger indhold, som kun kan skaffes via en eller anden for betaling, hvilken er så bedst? Er brugen stor nok til, at et abonnement eller tilsvarende en ordning kan forsvares? Eller er det bedre at indkøbe, når behovet opstår?

I det hele taget at identificere hvad vores brugere har brug for: At få anbefalinger til indkøb fra forskere og undervisere kan være ganske udfordrende, lige som anskaffelsen skal foregå i en jungle af udgivere, platforme, aggregatorer, pakker og mange kommercielle modeller.

Heldigvis er det sådan, at den teknologiske udvikling og nye standarder gør det lettere for bibliotekerne at adressere disse udfordringer. Allerede i dag kan vi få fordel af, at brugere får søgeresultater, som er skræddersyet til deres interesseområder og af den faktiske brug ved hjælp af Data Mining kan facilitere discovery-systemer og fremtidig accession. Resultatet er ægte evidens-baseret accession og klogere brug af biblioteksbudgetterne.

Ranking efter relevans

Bibliotekernes mål har altid været at give adgang til relevant information så hurtigt og præcist som muligt. Til det formål har bibliotekarere blandt andet skabt sofistikerede klassifikationssystemer i egne online-kataloger.

Men det ser ud til at brugerne synes, at Google og tilsvarende søgesystemer gør et meget bedre job. Spørgsmålet er om dette er sandt?

Googles PageRank-algoritme tager for eksempel hensyn til – blandt flere parametre – antallet af og signifikansen af websites der linker til en bestemt webside. Men måden, som algoritmen afgør en sides relevans i forhold til en bestemt brugers interessefelter, er ikke klar.

En hovedudfordring for bibliotekernes nye generation af discovery-systemer er, hvordan man leverer relevante resultater i et stort hav af information. For at imødegå denne udfordring har udviklere beriget systemerne med nye værktøjer. Facetteret navigation kan for eksempel hjælpe brugerne til hurtigt at raffinere deres resultater og fokusere på en relevant delmængde.

Problemet er bare, at alle nu er så familiære med Google, at brugere af discovery-systemer derfor kun kigger på de øverste søgeresultater. Dermed overses relevante dokumenter længere nede på listen.

Svaret er derfor, at ranking efter relevans er blevet afgørende for at kunne opfylde brugernes forventninger og behov. Uanset det skønne ved andre metoder.

Data Mining

Bibliotekarere har aldrig fokuseret på at levere efterspurgte materialer. Hvad med al den gemte og glemte guld, som brugerne ikke en gang ved de burde kende?

Biblioteker har altid understøttet den slags opdagelsesrejser, for eksempel gennem velstrukturerede thesauri, vejledning

og bibliotekernes fysiske indretning. Igen, søgemaskiner har udviklet serendipitet, men de mangler afgørende viden om, hvad forskere og undervisere søger efter. Biblioteksleverandører har, på den anden side, adgang til lige netop den slags data fra logning af systemer, der kører på langs og på tværs af det akademiske landskab, for eksempel data genereret af både søgeværktøjer og link-resolvers.

Nye services, der f.eks. oplister månedens mest populære artikler inden for fagdisciplinerne, er kun begyndelsen på bedre udnyttelse af brugerdata og blandt de få, som kommer til offentlighedens kendskab.

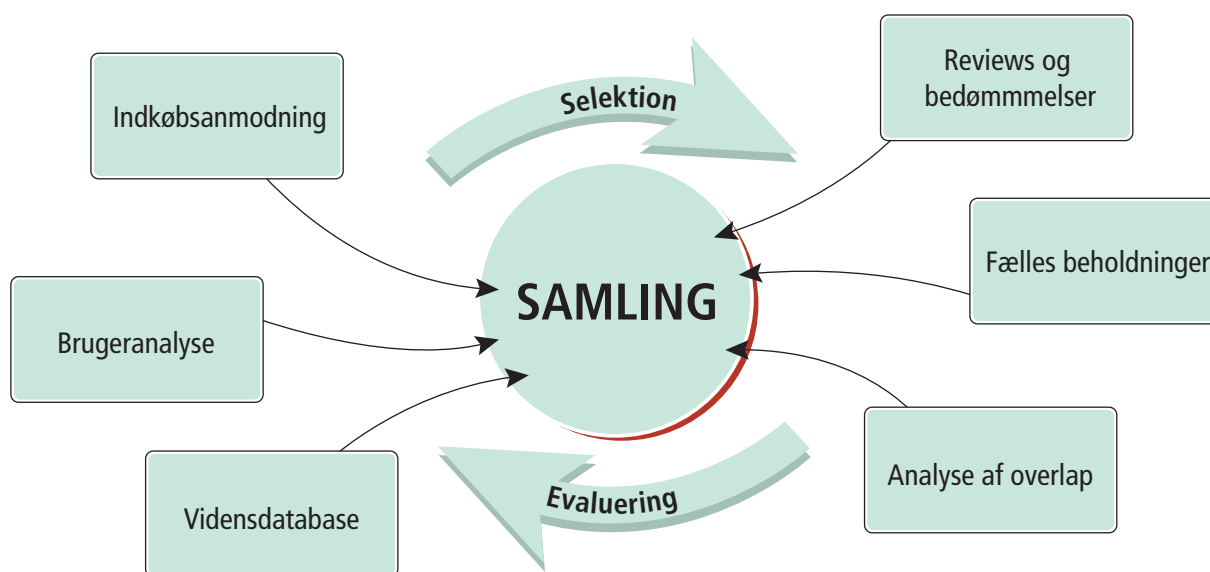
Mindst lige så vigtig er de interne data på bibliotekerne, dvs. rapporter og analyser som beskriver, hvad brugerne leder efter, hvordan de ser deres søgninger, hvad de downloader og ikke mindst, hvad de vælger ikke at bruge.

tilsvarende værktøjer til det digitale bibliotek. Den næste generation af bibliotekssystemer kan således ekstrahere realtidsdata fra forskellige medietyper, platforme osv. Og denne dataindsamling kan begynde allerede i anskaffelsesprocessen. Ved at integrere data, fra prøvelicenser og andre evalueringsforsøg, ind i den bredere accessionsproces, så kan næste generations bibliotekssystemer strømline accessionsprocessen og facilitere, at bibliotekerne adresserer mange flere behov.

Det stigende pres for at bevise bibliotekernes ROI (Return on Investments) over for moderinstitutionerne gør, at mange er gået i gang med at bruge de værktøjer der findes.

Den virkelige forandring sker, når bibliotekerne i fremtiden ikke længere kun bruger et lille udsnit af data fra forskellige systemer (som LMS, ERMS og link-resolvers), men i stedet forsøger at bruge *alt* data til noget meningsfuldt.

Endelig vil Data Mining med udveksling af data mellem biblioteker give mulighed for at skabe meningsfulde sammenligninger og endnu bedre beslutninger.



Evidenbaserede beslutninger

Langt hen ad vejen har bibliotekssystemer altid produceret meningsfulde data. Der er produceret masser af sofistikerede rapporter om søgeadfærd i de fysiske, print-baserede biblioteker. Information om brug af de trykte materialer har altid givet værdifuldt input i forbindelse med accession i den analoge verden.

Dette har naturligvis medført, at bibliotekernes leverandører har udviklet og fortsat udvikler

Fremtiden

Fremskridt i brugen af data og andre aspekter af e-ressource management i øvrigt, som f.eks. COUNTER (Counting Online Usage of Networked Electronic Resources), skal ses som en sluttet kredsløb af først analyse, derpå beslutningsproces og igen analyse osv.

I den ikke-alt-for-fjern fremtid kan vi forvente, at almindelige teknologivirksomheder udvikler mange flere services. Modsat indholdsleverandører og biblioteker, så er disse virksomheder gearede til at få meningsfulde sammenhænge ud af de mange data fra de mange platforme og ressourcer. ●