

# **BEDRE BRUGEROPLEVELSE MED DATA SCIENCE**

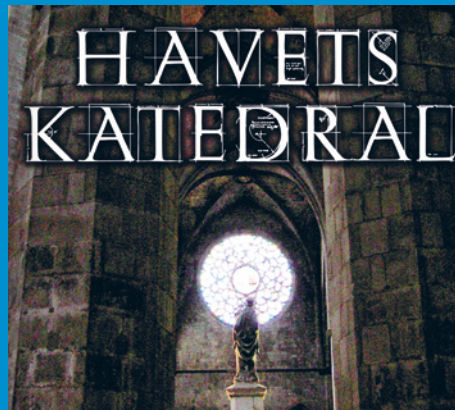


Data science er en videnskabelig analysemetode, hvor man via automatiseret bearbejdning af større datamængder kan udtrække viden om sammenhænge, mønstre og adfærd.

I bibliotekssammenhæng kan det fx være viden om, hvilke materialer der ud fra udlånsdata og indhold ligner hinanden, eller hvilke emneord, der typisk anvendes sammen i metadata.



“Havets katedral” af Ildefonso Falcones (udsnit)



Bibliotekernes position som kulturformidlere udfordres af kommercielle aktører som Netflix, Amazon, Saxo og andre steder, hvor personalisering og tilbud baseret på adfærdsdata for længst er blevet standarden og noget, mange brugere ønsker og forventer.

Er data science et spor, bibliotekerne bør følge, eller en 'kommercialisering', som man bør holde sig langt væk fra?

På DBC er vi ikke i tvivl.



"Biroller" af Erling Jepsen (udsnit)



“Hundehoved” af Morten Ramsland (udsnit)

## DET HANDLER OM AT GIVE BRUGERNE DET, DE FORVENTER – OG LIDT TIL

Der bør gøres en indsats for at fastholde og udvikle bibliotekets rolle på oplevelses- og læringsmarkedet. Inddragelsen af data science i den digitale brugeroplevelse er en værdiskaber for brugerne og bibliotekerne og bør være en naturlig del af en fremtidig søgning i bibliotekets tilbud.

Med anbefalinger, filtrerede søgninger og forbedret ranking i svarsæt er data science med til at give brugerne bedre overblik over bibliotekets materialer.

Anbefalinger kan også inspirere os til nytænkning. For når det kommer til stykket, er vi så ikke alle sammen mentalt mere eller mindre allerede i vores egen selvkonstruerede ”boble”, hvor vi har svært ved at tænke uden for egne ”filtre” og har brug for inspiration til opdagelsen af andre materialer end dem, vi selv kan komme i tanke om?

## FANGET I THE WALKING DEAD

Når du går på Amazon, eBay eller Netflix, så kender de dig og din adfærd. De ved, hvad du søgte efter sidst, og giver dig forslag til, hvad de tror, du også vil kunne lide.

Nogle gange fanger forslagene din interesse. Andre gange bliver du irriteret, for det er jo ikke sikkert, at du, fordi du så et enkelt afsnit af ’The Walking Dead’, har lyst til, at din verden skal indskrænkes til et Walking Dead-lignende univers i al evighed.

Hvor ved de overhovedet fra, hvad du egentlig leder efter, og hvad du har brug for? Og hvad går du glip af, fordi du slet ikke får det præsenteret?

## SKJULT MANIPULATION

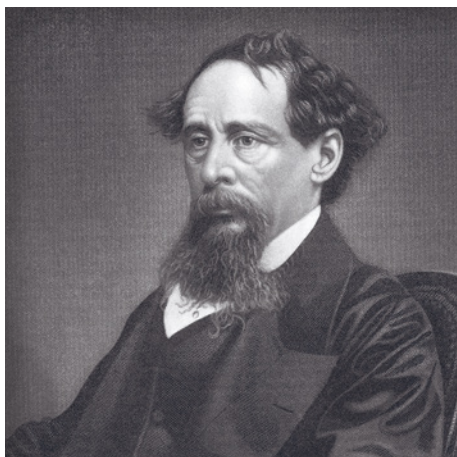
Sådan er data science i den påtvungne, uigennemskuelige udgave. Når den er frakoblet valgfrihed og transparens, og du ikke kan gennemskue, hvad der ligger bag de stillede forslag.

Sat på spidsen udsættes du for en diskret men dog umiskendelig intellektuel manipulation, hvor nogen og noget har valgt at indskrænke din personlige frihed og dit udsyn.

Men at data science kan bruges til at fremme kommercielle interesser, er ikke det samme som, at der er noget galt med teknologien.

## BIBLIOTEKET HAR IKKE NOGET TIL SALG – OG SÆLGER IKKE UD

Til forskel fra kommercielle platforme, der holder skarpt øje med salgspotentialet, har bibliotekerne det lettere. Her er ingenting til salg, og alle er velkomne til det hele.



Charles Dickens

Her er et kurateret, biblioteksfagligt tilbud efter brugerens egne valgkriterier et positivt tilvalg, som man kan benytte eller lade være.

Når bibliotekerne tager data science i brug, skal det være med respekt for den enkeltes ret til at vælge til og fra og med fokus på, at de systemer, brugeren møder i bibliotekssammenhænge, skal være transparente og gennemskuelige.

## ER FREMTIDEN EN BIBLIOTEKAR-ROBOT?

Absolut ikke. Ingenting kan erstatte den personlige, biblioteksfaglige formidling, men i takt med at mange, ud over den personlige biblioteksbetjening, også bruger bibliotekstilbuddet digitalt, vokser behovet for at kvalificere de digitale visningsformer, så brugeren også her kan få et biblioteksfagligt, kurateret tilbud. Både så bibliotekets forskellige materialer bliver synlige og som hjælp til den digitale biblioteksbruger, der sidder hjemme og gerne vil vide, hvad biblioteket anbefaler.





## ANBEFALINGER – FIRE INDFALDSVINKLER

I øjeblikket arbejder vi med fire indfaldsvinkler til anbefalingsfiltre, som en bruger vil kunne anvende for i sin søgning at se det udsnit af bibliotekets tilbud, han/hun vil præsenteres for:

**MEST POPULÆRE** præsenterer titler på baggrund af udlånsdata og hvad andre har lånt, kombineret med presseomtale og omtaler på sociale medier som f.eks. Facebook og Twitter.

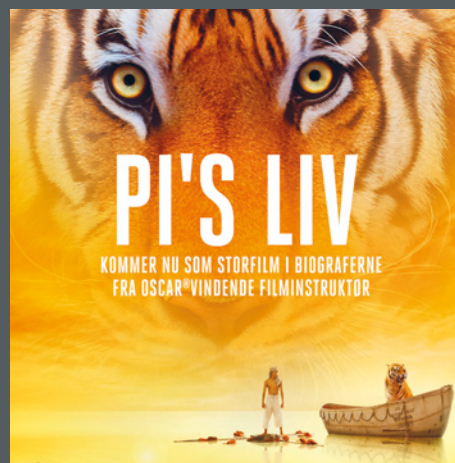
**BIBLIOTEKAREN ANBEFALER** præsenterer de materialer i søgesættet, der har tilknyttet en anmeldelse eller er anbefalet af bibliotekarer i form af omtaler, top 10-liste, materialevurderinger o.l.

Dette filter giver mulighed for at bringe flere af bibliotekets tilbud i spil. Biblioteket kan fx gøre opmærksom på materialer, der har høj kvalitet men ikke bliver meget udlånt.

**NOGET DER LIGNER** præsenterer materiale, der ligner en konkret titel. Det kan både være afledt af brugeradfærd (brugere, der låner Bog A, låner typisk også Bog B og Bog C) eller afledt af lighed i metadata (Bog A og Bog B handler begge om slægtsforskning; Bog A og Bog C har samme forfatter).

**BEDSTE FORSLAG TIL MIG** kombinerer de tre øvrige filtre. Det tager udgangspunkt i brugerens personlige profil og præsenterer de bedst matchende titler på baggrund af metadata, "Mest populære", "Bibliotekaren anbefaler" og noget, der ligner brugerens tidligere lån. Vi eksperimenterer i øjeblikket med dette filter, der skal afprøves og fintunes i samarbejde med bibliotekarer og brugere.

"Pi's liv" af Yann Martel (udsnit)



## FORDELENE I COMPUTERASSISTERET KATALOGISERING ER TIL AT FÅ ØJE PÅ

Data science er ikke kun relevant for brugerne men også i forhold til den bibliografiske registrering af fysiske materialer og af fuldtekster, hvor systemet, ud fra en analyse af teksten, stiller forslag til metadata.

Siden 2015 har vi brugt data science til at optimere og kvalificere den bibliografiske registrering, så vi nu får automatisk genererede forslag til emneord og DK5-klassemærker. Ud fra et manuelt registreret emneord foreslår den computer-assisterede indeksering de emneord, der oftest optræder sammen med det pågældende emneord. Systemet kan også ud fra emneord foreslå klassemærker og omvendt – foreslå emneord ud fra klassemærker.

Computerassisteret katalogisering er ikke et udtryk for, at registreringsarbejdet nu kan erstattes af en computer. Tværtimod bygger metoden på at finde mønstrene i de tusindvis og atter tusindvis af katalogiseringer, der er menneskeskabt gennem årene. Uden alle disse katalogiseringer ville der ikke være data at arbejde med.

Men fordelene – i form af kvalificerede forslag og øget datakonsistens og tidsbesparelse er til af få øje på.

## TRANSPARENS OG VALGFRIHED

Princippet om valgfrihed og muligheden for at kunne gennemskue kriterierne bag udvælgelsen af, hvad man bliver præsenteret for, gør data science i bibliotekssammenhæng til et positivt tilvalg, der hverken tilstræber at begrænse eller lede brugeren i en bestemt retning – med mindre brugeren ønsker det.

Hverken i princippet eller i virkeligheden er der jo noget i vejen med, at en bruger foretrækker at få præsenteret titler efter bestemte selektionskriterier. Den afgørende forskel er, at man har mulighed for at vælge til og fra.

Formålet er alene at give brugeren værktøjer, der gør det lettere at søge målrettet i bibliotekets tilbud.

## RESPEKT FOR PERSONFØLSOMME DATA

De data science-modeller, vi arbejder med, omfatter ikke personfølsomme data.

Når vi anvender udlånsdata til at koble materialer, frasorteres sjældent lånte materialer, idet titler, der har mindre end 100 lån, ekskluderes fra datasættet. Og da udlånsdata altid anvendes i kombination med andre modeller som metadata og generel popularitet, er det umuligt at udtrække personfølsom information fra de data science-baserede anbefalinger.

## OPSUMMERING

Anbefalinger ved brug af data science er en værdiskaber i den digitale formidling, der:

- inspirerer og stiller kvalificerede forslag
- sætter fokus på meget andet end det mest populære
- giver brugerne det, de forventer og det, de ikke vidste fandtes
- bringer bibliotekernes metadata og formidlingsarbejde i spil digitalt
- er et positivt tilvalg, fordi brugeren selv bestemmer.



## BESØG DBC'S Udstilling på NETTET



### Scan QR-koden og gå til:

- e-version af denne folder
- billeder og videoklip fra årsmødet
- Indtryk fra DBC's workshop om data science: 'Bibliotekerne – den demokratiske søgefunktion'
- Artikler om data science.

## DBC DRIVER OG UDVIKLER DEN BIBLIOGRAFISKE OG IT-MÆSSIGE INFRASTRUKTUR TIL BIBLIOTEKERNE I DANMARK

Det gør vi blandt andet ved at:

- udfærdige den danske nationalbibliografi, der er en fortegnelse over bøger, film, musik, internetdokumenter mm., der udkommer i Danmark
- udarbejde en bibliotekskatalogisering, hvor materialer af særlig relevans for de kommunale biblioteker forsynes med yderligere registrering af data, målrettet brugen i bibliotekerne
- drive og udvikle de danske bibliotekers fælles base, Danbib, der blandt andet understøtter det fælles lånesamarbejde mellem bibliotekerne
- drive og videreudvikle bibliotek.dk, der er borgernes fælles indgang til alle danske udgivelser og til de danske bibliotekers samlinger og digitale ressourcer, og hvor man kan søge og bestille materialer fra alle landets biblioteker
- drive og udvikle den nationale databrønd og et stort antal webservices, der er rygraden i biblioteksinfrastrukturen
- drive og udvikle artikeltjenesterne Faktalink og Forfatterweb og filmtjenesten Filmstriben
- producere lektørudtalelser/materialevurderinger til brug for bibliotekernes materialeanskaffelse og formidling.

DBC er et aktieselskab ejet af KL (61,5 %) og staten (38,5 %).