

BRUGERSPOR FORBEDRER BRUGEROPLEVELSEN

Store webtjenester som Facebook og Google har gjort anvendelsen af brugerspor til en selvfølge. Brugerne forventer ligefrem, at også forskningsbibliotekerne forbedrer brugeroplevelsen via brugen af brugerspor. Men inden forskningsbibliotekerne går i gang, er der nogle etiske og juridiske problemstillinger, som skal adresseres.

AF JENS HOFMAN HANSEN, JHH@STATSBIBLIOTEKET.DK OG MICHAEL NIELSEN, MN@STATSBIBLIOTEKET.DK, STATSBIBLIOTEKET, KASPER BØGH, KASB@DTIC.DTU.DK, DTIC, KASPER LØVSCHELL, KL@AUB.AAU.DK OG KARSTEN KRYGER HANSEN, KKH@AUB.AAU.DK, AALBORG UNIVERSITETSBIBLIOTEK

Anvendelsen af brugerspor i det web-baserede forskningsbibliotek blev for nylig belyst på endagskonferencen OnTracks. På konferencen gjorde fire speakere fra ind- og udland det klart, hvor langt udviklingen inden for brugerspor er kommet, og hvad vi som forskningsbiblioteker kan gøre for at fremme en bedre brugeroplevelse.

En konklusion er, at anvendelsen af brugerspor er blevet konventionel, og at det ligefrem forventes af brugerne, at der anvendes brugerspor. Det er især store webtjenester som Facebook, Google, iTunes og Amazon, som er årsagen til, at det er blevet et spørgsmål om, hvordan teknikken skal anvendes – i stedet for om den skal anvendes, idet de sætter standarden for brugeroplevelse på nettet.

Passive og aktive data

Når vi taler om brugerspor, tænkes der primært på passivt afgivne data, som 'bag kulisserne' kan anvendes til at forbedre f.eks. søgeoplevelsen. Der er her tale om data, som brugerne afgiver ubevidst og uden ekstra indsats.

Fordelen er, at det er relativt let at få skabt disse data og komme i gang med at skabe en bedre brugeroplevelse. Brugen og eksponeringen af passivt afgivne data kan være første skridt på vejen mod et stærkere engagement med brugerne, hvilket typisk er forudsætningen for en mere aktiv brugerskabelse af data. Eksempler på aktivt skabte data er vurderinger, tags og anmeldelser.

Brugerens sti rundt på siden

Potentielt set er mængden af brugerspor endnu mere overvældende, omfattende og ikke mindst kompleks end de egentlige data, som bibliotekerne stiller til rådighed for deres brugere.

En bruger afgiver, når han besøger bibliotekets website, oplysninger om websted, brug af browser, skærmstørrelse, operativsystem og lokalitet. Oplysninger, som i sig selv måske ikke er særlig interessante, men sammen med et session id udgør de grundlaget for en sporing af brugeren, og en tilpasset og forbedret brugeroplevelse kan potentielt indledes med disse oplysninger. Hvis en bruger desuden ender med at logge ind med sit brugernavn, kan hele brugersporet kobles til den registrerede bruger.

Hvis brugeren forsvinder for at vende tilbage til bibliotekssiden på et senere tidspunkt, kan den fortsatte adfærd på websiden kobles til den tidligere registrering via cookie-teknologi. Det er her, at anvendelsen af brugersporerne for alvor begynder at adskille sig fra en almindelig implementering af f.eks. Google Analytics, for brugersporsfilosofien går et skridt videre og knytter serveraktiviteter til den enkelte bruger – anonymiseret eller ej.

Filtrede søgeresultater

Brugerens låne- og downloadhistorik kan gemmes og anvendes til filtrering af f.eks. søgeresultater, så brugeren slipper for at se data, som vedkommende formentlig slet ikke vil være interesseret i. Google praktiserer i dag denne teknik på en måde, så søgeresultater for forskellige brugere fremtræder differentieret afhængigt af deres geografiske tilhørsforhold og deres søgehistorik. Sender du et Google-søgeresultat til en ven, er det altså ikke nødvendigvis det identiske søgeresultat, han får at se.

Kendskabet til brugeren kan udover søgedifferentiering anvendes til at give målrettede anbefalinger som "Brugere som dig, der kigger på dette, er typisk også interesserede i dette" eller "Brugere, der gjorde dette, gjorde det her bagefter".

Brugergenereret udvikling

Den måske allervigtigste årsag til at beskæftige sig med brugerspor i formidlingen af bibliotekets materialer og digitale tilbud er, at brugerne efterhånden forventer en brugeroplevelse, der er så målrettet, at den kun kan opnås via anvendelsen af brugerspor. Med andre ord forventer brugerne i stigende grad, at vi udnytter vores kendskab til dem.

Det er en påstand, der i sig selv virker ganske plausibel, men dokumentationen begynder også at blive mere og mere entydig. Blandt andet skulle en rapport fra MOASIC-projektet under JISC dokumentere den.

En lignende udvikling skete for 5-6 år siden, hvor de første biblioteker begyndte at gå over til et enkelt søgefelt med efterfølgende relevansrangordning. Det var ikke en biblioteksoptfindelse, men også resultatet af en brugerforventning. Hvis vi ikke anvender brugerspor aktivt i vores service af brugerne, får det den yderste konsekvens, at vi som biblioteker kommer til at servicere under niveau.

Den gode nyhed er, at vi som forskningsbiblioteker har mulighed for at komme til at vide en masse om brugerne – især fordi det tætte samarbejde og fortroligheden med universiteterne giver os en unik mulighed for at udveksle data, som Google og andre virksomheder aldrig kommer i besiddelse af.

Første opgave:

Definer målet med brugerdata

Det professionelle arbejde med brugerspor kræver, at man holder sig relevante problemer for øje.

David Kay, som har været i spidsen for syntesen af ni JISC-projekter, som alle handler om brugerspor, mener, at mange af de problemer, som dukker op i diskussionerne om brugerspor, er 'second order challenges'.

Den primære udfordring er nemlig at få defineret, hvad vi egentlig vil opnå med anvendelsen af brugerdata. Det handler om at få fastlagt de bruger- og forretningsmæssige mål med anvendelsen. Diskussioner om kvalitet, jura, etik, teknik og økonomi må nødvendigvis komme bagefter som noget sekundært.

Korrekt anvendelse af brugerspor findes ikke

Bekymringer om, hvorvidt datagrundlaget er sandfærdigt og komplet, dukker ret hurtigt op. Forskellige systemer leverer data på vidt forskellige måder, og det er umuligt at få adgang til 'perfekte' data. Derfor må diskussionen vedrørende kvalitet ændres til et spørgsmål om nytte og opfyldelse af målsætning. For eksempel kunne bibliotekarerne på Huddersfield University identificere en markant stigning i udlån, efter brugergrænsefladen i 2006 fik tilføjet "Borrowing suggestions" som anbefaling til brugerne.

Tim Spalding fra LibraryThing mener, at anbefalinger baseret på brugerdata ikke handler om at gengive en sandhed eller om præcise forudsigelser. Anbefalinger handler derimod om at matche brugerbehov og brugerforventninger med en veltilrettelagt algoritme. I onlinesupermarkedet er det for eksempel ikke nødvendigvis babymad, man skal forsøge at sælge til den far, der for første gang køber bleer – den skal han nok købe alligevel.

Derimod skulle man snarere forsøge at sælge ham en sixpack øl. Efter han er blevet far, kan det lokale værtshus ikke længere besøges med samme frekvens, og derfor vil han være tilbøjelig til at købe øl, som, som han kan drikke derhjemme.

Ligeledes er det matematisk set et godt forslag at anbefale folk at læse Harry Potter nr. 6, hvis de har læst Harry Potter nr. 1-5. Men forslaget er ikke overraskende og byder ikke på noget, som brugeren ikke selv med rimelig sandsynlighed har overvejet. Konklusionen er, at den algoritme bør justeres til at blive lidt mere overraskende i sine forslag.

Et andet spørgsmål går på kritisk masse. Kan man som et mellemstort universitetsbibliotek overhovedet foretage sig noget anvendeligt på den trods alt begrænsede mængde brugerdata, man har til rådighed? Svaret er, at det kan lade sig gøre. Der er dog forskel på, hvor avancerede anbefalingerne kan blive. Hvis der er få brugere, kan forslag som "Andre brugere, der kiggede på dette, kiggede også på dette" være et realistisk ambitionsniveau. Hvis der er mange brugere og mange transaktioner, kan du lave anbefalingen mere målrettet som "Andre brugere som dig, der kiggede på dette, kiggede også på dette".

Er det lovligt?

En klassisk barriere for at gå i gang med at anvende brugerdata er den juridiske. I Danmark har vi en meget stram persondatalov, som foreskriver, at vi ikke må gemme personhenførbare data. Imidlertid kan loven næppe diktere, at alle brugersporsinitiativer må skrottes.

Data kan anonymiseres i forskellige grader, men det er stadig uklart, hvorvidt man kan lade brugerne give lov til at data gemmes, hvis de selv finder det nyttigt.

Det kan i den forbindelse undre, at adskillige virksomheder, som opererer i Danmark under dansk lovgivning (Google, Amazon, Apple, etc.), og som sætter standarden for anvendelsen af brugerdata, tilsyneladende ikke har problemer med Datatilsynet. Der er helt sikkert nogle jura-spøgelser i biblioteksverdenen, som gør det bekvemt at undlade at beskæftige sig med brugerspor.

Og hvad med det etiske?

Udover det juridiske er der selvfølgelig også den etiske problemstilling. At indsamlingen og anvendelsen af de forskellige data er lovlig, gør den ikke nødvendigvis etisk forsvarlig. Bo Sannung fra SAS Institute har givet et eksempel på, hvordan for eksempel pensionselskaber på deres hjemmesider samler informationer fra indtastningsformularer, som brugeren ikke aktivt indsender, fordi de fortryder på halvvejen. Disse informationer kan være nyttige at have om brugeren, der måske ender med at få et personligt opkald fra en 'rådgiver'.

Eksemplet er grænseoverskridende for de fleste biblioteksfolk, men det illustrerer, hvor omverdenen i visse tilfælde befinder sig på dette område. Det kan være en etisk udfordring at levere en god kundeservice.

Det er i etisk forstand en god strategi at vægte transparens, når biblioteket tager brugerdata ind som aktivt element i formidlingen. Det skal være muligt at gennemskue, hvordan og hvornår data bruges – og måske skal det også være muligt at operere med et personligt filter, der justerer brugen af data.

Samtidig er det dog værd at huske på, at brugerne er meget kritiske. Hvis bibliotekerne anvender brugerspor til at forbedre en relevansalgoritme eller til at foreslå relateret materiale i lister, betyder det ikke, at brugerne tvinges til middelmådighed eller til en bestemt ideologi. I mange tilfælde kan de vælge blot at ignorere forslagene.

Her presser begrebet 'filterboble' sig i øvrigt på. Egentlig har vi altid levet i filterbobler. For eksempel abonnerede ens



Bibliotekerne er "followers" og ikke "leaders" hvad angår anvendelse af brugerspor.

forældre på Politiken i løbet af ens opvækst, og det abonnement blev til ens filterboble.

Det nye og lidt luskede er imidlertid, at de digitale filterbobler er personlige (så vi er alene i dem), de er ofte usynlige, og brugerne vælger dem som regel ikke selv. Det betyder, at der ligger et stort etisk ansvar hos de systemudviklere og de biblioteker, der designer filterboblerne.

'Followers' og ikke 'leaders'

Hvad med det økonomiske og det tekniske problem? Ofte hænger de to størrelser sammen. Mange vil ikke have ressourcer til rådighed til at påvirke de systemer, som bliver stillet til rådighed for brugerne, og i mange situationer hvor der er stillet udviklere til rådighed for bibliotekerne, vil der ikke stilles tilstrækkeligt med data til rådighed fra de eksterne udbydere.

Ovenstående er selvfølgelig et praktisk problem. Bibliotekerne er allerede 'followers' og ikke 'leaders', hvad angår anvendelse af brugerspor, og det må blive det muliges kunst at tage brugerspor i anvendelse. Første trin er dog at have kendskab til feltet, så man kan stille krav til systemleverandørerne.

Anbefaling: Start nu og start småt

Hvad skal bibliotekerne så gøre? Hvis man beslutter sig for at prioritere brugerdata, er det første man skal sørge for, at data rent faktisk indhentes. Det er ikke altid en selvfølge. Det er derfor tilrådeligt at starte en

basal dataindhentning og sideløbende arbejde med at få anvendelsen af data sat på dagsordenen. Det vil være meget svært i praksis at allokere ressourcer til et udviklingsprojekt, hvis der slet ikke er nogen data at eksperimentere med.

Det anbefales desuden at starte lokalt for hurtigt at påvise effekten af et initiativ. Sideløbende og/eller efterfølgende kan det give mening at aggregere data til et nationalt niveau, hvor det også bliver muligt at øge den kritiske masse med efterfølgende højere kvalitet til følge.

Sidst, men ikke mindst, er det en god ide at starte simpelt. Man kan hurtigt blive overvældet af den kompleksitet, som et system baseret på brugerdata kan resultere i, og det er slet ikke sikkert, at et meget kompliceret eksperiment vil give særligt anvendelige resultater. ●

Slides og videoptagelser fra konferencen OnTracks er tilgængelige på konferencesitet www.ontracks.dk

Konferencen blev finansieret af DEFF