

Brugen af web-statistik i danske folke- og forskningsbiblioteker

Af Kasper Løvschall

Vi tæller i disse år til den helt store guldmedalje. På udlån, indlån, besøgende, udførte informationsvejledninger, årsværk, og så videre. Fysiske indikatorer, hvor statistikkerne bruges blandt andet til at evaluere vores aktivitetsniveau, målrette vores indsatsområder og kort og godt som dokumentation af vores berettigelse over for bevillingsgiverne.

Kasper Løvschall

Teknisk-naturvidenskabelig udviklingskonsulent
Aalborg Universitetsbibliotek
kl@aub.aau.dk



I den senere tid er vi begyndt at sætte øget fokus på web-statistikken; blandt andet fordi bibliotekernes web-sites efterhånden er blevet en meget central del af vores samlede tilbud til brugerne. På den baggrund investerer vi fortsat mange midler til drift og systemudvikling og den arbejdskraft, der skal til for at forfatte indhold og i stigende omfang vedligeholde den etablerede informationsmængde.

Web-statistik er såmænd ikke et nyt påfund. Den har eksisteret lige så længe, vi har anvendt web-servere. Blot har statistikken tidligere været knyttet til de driftsmæssige aspekter. Men i stigende omfang har web-udviklerne og ledelsen fået øjnene op for de muligheder, der ligger i anvendelsen. Samtidig er web-statistik indgået som element i den samlede biblioteksstatistik.

De store spørgsmål

Et er at tælle, noget andet er at få svar på de spørgsmål, der virkelig er interessante: Hvem er vores brugere, og finder de den information, de søger? Sker det på den mest optimale måde? Hvordan kommer de til mit web-site, og hvordan forlader de det? Og endelig: Får vi nok ud af vores investering i web-sitet?

Der er for så vidt ikke noget i vejen med at lede efter disse svar i statistikkerne, så længe man kender til de begrænsninger, der ligger i dens anvendelse. Man skal først og fremmest huske på, at selve den

statistiske opsamling ikke har ændret sig synderligt, siden man fandt på at lagre web-serverens transaktioner. Derfor er mange af de spørgsmål særdeles komplicerede at finde et kvalificeret svar på. Blandt andet fordi store dele af materialet er flertydigt, og fordi der ikke lagres de nødvendige oplysninger om brugerens besøg.

Søgen efter disse svar får den konsekvens, at man må omdirigere sit web-site til en struktur, der bedre faciliterer datamining. Ofte bliver man også nødt til at ændre i de rutiner på webserveren, der opsamler datagrundlaget for at få alle nødvendige oplysninger med.

Helt galt kan det gå, når man med sin statistik i hånden begynder at foretage sammenligninger institutioner eller web-sites imellem. Den gængse statistik kan nemlig i bedste fald sige noget om en selv, men sammenlignet med andre, bliver den ofte ubrugelig.

Jagten efter den unikke bruger

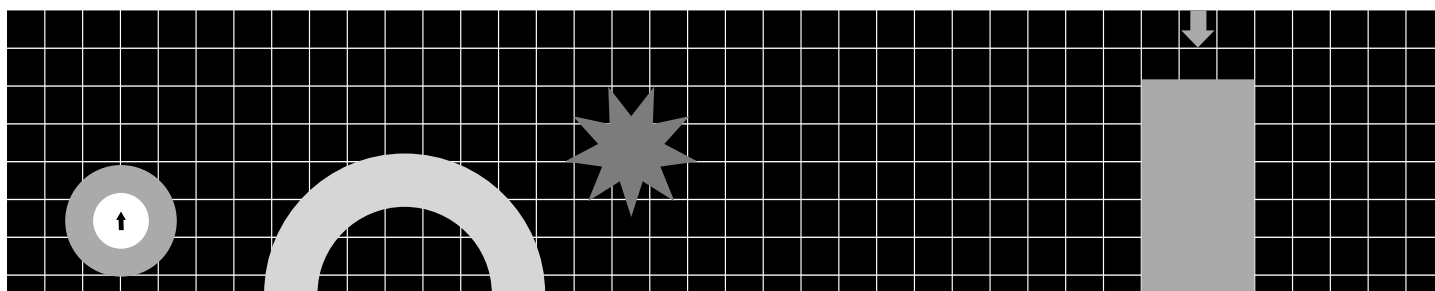
Et af de oftest stillede spørgsmål til web-statistikken er identifikationen af den unikke bruger. Det er et håndgribeligt tal, der for eksempel kan sammenlignes med besøgstal i det fysiske bibliotek. Kendskabet hertil er samtidig et basisparameter for svar på de fleste andre spørgsmål, vi stiller.

Desværre er det på ingen måde enkelt at identificere den unikke bruger. De fleste statistikprogrammer baserer deres beregning ud fra en identifikation af browserens ip-adresse (computerens unikke "telefon"-nummer på Internettet), men da f.eks. universiteterne i stor udstrækning anvender såkaldte proxy-servere, kan flere tusinde brugere være gemt under samme identifikation (lidt ligesom at gå igennem en omstilling med sine udgående telefonsamtaler). De fremstår altså som den samme, unikke bruger. På grund af firewalls i private virksomheder og offentlige forvaltninger er dette problem ikke kun knyttet til universiteter og andre uddannelsesinstitutioner. Konklusionen er således, at vi har langt flere unikke besøgende end der opgives. Men nøjagtigt hvor mange, er fuldstændig afhængigt af den institution, man hører til, og hvilken slags brugere man har.

Der findes flere muligheder for at tackle dette problem. Sagen er blot den, at der ikke findes en standard for måden, hvorpå dette skal gøres, og på samme måde har statistikprogrammerne forskellig tilgangsvinkel til at kompensere i beregningerne på baggrund af de løsninger, vi måtte implementere. Endelig vil en løsning typisk være bundet til den software, webserveren kører på.

Af konkrete muligheder kan man arbejde med cookie og session spring. En cookie er en lille pakke med oplysninger, som

Web-statistikken er i sig selv nok ikke det orakel, vi ofte efterspørger. Men kombineret med andre analysemetoder såsom brugerundersøgelser og fokusgruppeinterviews, kan vi nok få en bedre afklaring af de spørgsmål, vi stiller.



webserveren afleverer til ens browser. Cookien ligger i ens browser, til man lukker vinduet – på den måde kan webserveren holde styr på brugeren, mens man navigerer rundt på bibliotekets web-site. Sessionsstyring er forholdsvis analogt til cookies, men lagres ofte i browserens adressefelt. Begge metoder er dog ikke helt fejlfri, men implementeret korrekt, kan de bidrage til et bedre statistisk grundlag.

Sammenlignelighed, pærer og bananer

Som jeg har nævnt tidligere, kan det gå helt galt, når man forsøger at sammenligne web-statistikker. Et af de grundlæggende problemer ligger i, hvilke forudsætninger og beregningsmetoder det anvendte statistikprogram benytter. Da de fleste programmer fungerer efter ”black-box” princippet, hvor vi putter data ind i den ene ende og får en detaljeret rapport ud i den anden, så aner vi ofte ikke, hvad der skete i den mellemliggende proces. Derfor kan man ikke foretage sammenligninger, medmindre man anvender samme program.

Selv om vi besluttede os for at anvende samme statistiske beregningsmetode, ville vi med stor sandsynlighed alligevel løbe ind i problemer. Da vi anvender vidt forskellige strukturer og teknologier på de enkelte web-sites, kan for eksempel sammenligninger af sidevisninger være vanskelige. Hvad der på et bibliotek kan formidles på en enkelt side, bliver måske formidlet over flere sider på et andet. Så en håndgribelig indikator som sidevisnin-

ger bliver lige pludselig reduceret til en sammenligning mellem pærer og bananer. Så hvad er værdien?

Vi skal altid huske på, at selvom vi måske selv godt kender svaghederne i vores statistikker, så vil der med stor sandsynlighed være mange, der ikke gør det. Derfor er det op til os at sikre, at der ikke opstår mistolkninger. Eller udvikle metoder, hvor sammenligneligheden er garanteret.

Forsøg med at skabe sammenlignelighed

Biblioteksstyrelsen arbejder i øjeblikket sammen med konsulentfirmaet TWINS A/S om udviklingen af en platform til sammenlignelig statistik mellem folkebibliotekerne.

Tanken i metoden er at reducere bibliotekets sidevisninger til ikke-tælbare og tælbare ”publikationer”, monografier og serier. Der er altså tale om en biblioteksorienteret tilgang til web-statistik. Siderne, der ikke tælles med, er for eksempel navigationssider, søgesider med videre, som i sig selv ikke bidrager med et egentligt tekstindhold. De sider, der tælles med, indeholder selve massen, som kan bidrage til brugerens egentlig orienterings- eller informationsbehov.

Metoden forudsætter en form for katalogisering af indholdet på ens web-site, hvor siderne inddeles i de to kategorier, der skal tælles med, og den kategori, der

skal holdes udenfor. Det foregår via et web-interface, hvor den enkelte institution kan gennemgå sine sider og efter endt klassifikation udtrække de faktiske tal for sidevisninger af publikationer.

Fordelen herved er, at alle anvender samme software, som samtidig er dokumenteret og ikke virker efter ”black-box” princippet. Ulempen er, at en direkte sammenligning mellem biblioteker stadigvæk ikke er entydig, da det enkelte bibliotek kan vælge at formidle samme slags information på forskellig vis og derved med forskellig påvirkning af statistikken. På den anden side bliver det enklere for det enkelte bibliotek at sammenligne sin aktivitet med sig selv over tid. Også på en måde, hvor man kan dokumentere sin aktivitet over for en udenforstående part.

Hvad så?

Web-statistikken er i sig selv nok ikke det orakel, vi ofte efterspørger. Men kombineret med andre analysemetoder såsom brugerundersøgelser og fokusgruppeinterviews, kan vi nok få en bedre afklaring af de spørgsmål, vi stiller.

Web-statistikken har helt bestemt også nogle rigtig gode styrker, når man kigger på det enkelte biblioteks websider. Man kan lynhurtigt evaluere, om enkelte sider benyttes tilstrækkeligt i forhold til andre, om navigationsstrukturen virker hensigtsmæssig og leder brugeren det rigtige sted hen.

Vi har også mulighed for at implementere mere direkte feedback muligheder på websiderne, hvor en bruger kan give os "karakterer" for det materiale, vi stiller til rådighed – var det denne oplysning, du ledte efter? Hvilken karakter fra 1 - 6 vil du give indholdet på denne side? På den måde er vejen fra bruger til os noget kortere i "cyberspace" og skal ikke vurderes gennem matematiske formler og forudsætninger. I stedet for at tælle så meget, må vi kigge på indholdet, brugbarheden og kvaliteten. Det er trods alt de faktorer, vi alligevel skal ind og pille ved, når vi ønsker en større aktivitet på bibliotekernes web-sites.

I stedet for at tælle så meget, må vi kigge på indholdet, brugbarheden og kvaliteten.

Dødens Triumf på Statsbiblioteket

Statsbibliotekets 102 års fødselsdag den 17. juni blev traditionen tro fejret med gaver til biblioteket og legater til nogle af dets medarbejdere fra Emanuel Sejrs Mindefond.

En af gaverne til Statsbiblioteket var en indspilning af Savage Rose's balletmusik til Dødens Triumf, udgivet på plademærket Polydor i 1979. Da Savage Rose nedlagde fagedforbud mod denne udgivelse, blev de fleste eksemplarer makuleret, og kun få er tilbage.

Musikken til Dødens Triumf blev oprindeligt skrevet til en ballet med udgangspunkt i Eugène Ionescos skuespil "Jeux de Massacre" fra 1970. Daværende balletmester ved Det Kongelige Teater, Flemming Flindt, lavede koreografien til Thomas Koppels musik, der blev udført af Savage Rose. Danmarks Radios TV uropførte balletten i 1971, og Det Kongelige Teater havde premiere på balletten i 1972. Balletten vakte stor opmærksomhed på grund af nøgenscener, et politisk budskab og ikke mindst den iørefaldende moderne musik



Mindefondens formand, Bjarne Dalsgaard, med Dødens Triumf.

leveret af et pop-rock-orkester. Balletten blev en af Det Kongelige Teaters største publikumssucceser. I 1979 blev balletten genopført i Cirkusbygningen i København,

men på grund af uoverensstemmelser omkring musikken blev der nedlagt fagedforbud, som blandt andet betød, at den samtidige pladeudgivelse blev stoppet.