

SAMARBEJDE MELLEM UDGIVERE OG BIBLIOTEKER – TIL ALLES FORDEL?

En katalogadgang, hvor søgningen i metadata suppleres med søgning i fuldtekst – og med visning af snippets fra de relevante steder – burde være til at håndtere. Men sådan er det slet ikke. Et DEFF-projekt med deltagelse af udgivere og biblioteker samler op på problemerne og erfaringerne og konkluderer, at der potentiale i udvidede partnerskaber.

AF BIRTE CHRISTENSEN-DALSGAARD, BCD@KB.DK, DET KONGELIGE BIBLIOTEK, CARSTEN FENGER-GRØNDAHL, CFG@UNIPRESS.AU.DK, AARHUS UNIVERSITET, TONNY SKOVGÅRD JENSEN TSJ@STATSBIBLIOTEKET.DK, STATSBIBLIOTEKET OG STEFFEN SØRENSEN, SCS@PUBLIZON.DK, PUBLIZON A/S

En ting er at få en lang række resultater af en søgning, en anden er at afgøre hvilke bøger, artikler eller andet, som er relevant for ens brug. I de seneste år har bibliotekerne lært meget af kommercielle aktører som Amazon og Google.

Google har vænnet vores brugere til at kunne søge i et simpelt felt og få relevant materiale og samtidig lært vores brugere, at man kan se lidt af, hvad der er årsagen til, at man har fundet dette eller hint. Amazon har vist brugerne, hvordan en søgning kan resultere i så rigt materiale, at man ikke er et sekund i tvivl om, hvad det er, man er i gang med at købe – om det er relevant, hvad det indeholder osv.

Bibliotekerne vil selvfølgelig gerne stille sammenlignelige tilbud til rådighed for deres slutbrugere – vi vil gerne tage det bedste fra forskellige websteder – og bruge det sammen med information fra vores bibliotekskatalog.

Man kan tage fat mange steder, men et helt oplagt emne er at arbejde med såvel introduktion af fuldtekstsøgning i bøger som supplement til søgning i de bibliografiske data samt mulighed for at vise en snippet, dvs. en lille stump tekst, hvori søgeordet indgår, som illustrerer overfor brugeren, hvorfor netop denne post bliver vist.

En katalogadgang, hvor søgningen i metadata suppleres med en søgning i fuldteksten, og hvor visning inkluderer snippets fra de relevante steder, burde være til at håndtere. Men så nemt er det nu ikke, som et forsøg mellem to udgivere – Aarhus Universitetsforlag og e-bogsproducenten og distributøren, Publizon – samt Statsbiblioteket og Det Kongelige Bibliotek har vist. Både teknisk og juridisk var det meget kompliceret og krævede faktisk en lovændring. DEFF bevilgede i 2007 penge til projektet, som også skulle se på en mulig forretningsmodel for de kommercielle udgivere, ligesom det skulle undersøge rettighedssituationen for at tilbyde en sådan tjeneste. Projektet havde groft sagt fire dimensioner:

- Den rettighedsmæssige dimension, hvor lovgrundlaget for den foreslåede service skulle håndteres, herunder at få afklaret en række juridiske problemstillinger i forhold til copyright.

- Den teknologiske dimension, hvor de to biblioteker skulle undersøge, om de kunne håndtere datamængderne, om de kunne opnå relevant ranking og om visningen af snippets kunne gennemføres, og at disse udvidelser ville resultere i en merværdi.
- Den forretningsmæssige dimension, hvor især Publizon skulle se på om der kunne laves en forretning ud af digitaliseringsaktiviteten. Et mål var en forretningsmodel, som gør det muligt for både forlag og biblioteker at få glæde af de samme skanninger.
- Den styringsmæssige dimension, hvor omkostningerne og kompleksiteten ved en total digitalisering af et forlags produktion skulle undersøges. Materialet til skanning kom både fra Statsbiblioteket og Aarhus Universitetsforlag, og skanningen blev foretaget af Publizon.

Gennemførelsen af projektet var noget mere kompliceret, end vi havde forestillet os. Det meget lange forløb skyldes primært udfordringen omkring rettighedssituationen, men andre forhold viste sig også mere besværlige end forventet, da projektet blev planlagt. Nedenfor gives en kort version af historien; den interesserede læser kan med fordel læse hele rapporten.

Aftalegrundlaget for brugen af det digitale materiale

Da projektet blev defineret, var alle godt klar over, at man ikke bare kunne påberåbe sig citatretten, og at et sådant projekt skulle forankres i forlagsverdenen. Der blev derfor arrangeret møder med repræsentanter for udgiverne og bestilt en juridisk udredning for at afklare, hvordan Publizon kunne etablere en service overfor bibliotekerne.

Målet var oprindeligt at etablere privatretslige aftaler mellem forlag og leverandøren af de skannede bogsider på den ene side, bibliotekerne på den anden side. Ved projektets start var der ikke lovhjælp for aftalelicens for denne type materiale, hvorfor projektet skulle undersøge andre muligheder for at få hjemmel til denne form for brug. Til sikring af disse aftaler hyrede projektet Advokatfirmaet Bender, Von Haller, Dragsted, hvis udredning langt fra gav et tilfredsstillende grundlag for det videre arbejde.

Kulturministeriet havde i mellemtiden igangsat arbejdet med at udvide den kollektive licensmodel således at den også dækkede for digitaliserede tekster. Denne blev vedtaget i sommeren 2008 således at ophavsretsloven åbnede mulighed for at indgå en aftale med aftalelicensvirkning, dvs. at den også omfattede rettighedshavere, som ikke er medlemmer af en rettighedshaverorganisation.

I lyset af den manglende privatretslige hjemmel og i lyset af aftalelicensens nye muligheder blev det besluttet, at projektet ville arbejde på at klare rettighedssituationen på denne måde. Resultatet blev, at projektet udarbejdede to sæt aftaler:

- mellem Projektet og Copydan som repræsentant for autorettighederne om klarering af rettighederne i forhold til autorerne, dvs. forfattere, illustratører m.v.
- mellem Projektet og Aarhus Universitetsforlag som har udgiverrettighederne, hvor Aarhus Universitetsforlag gav samtykke til, at projektet kunne anvende deres udgivelser.

Aftalen med Copydan blev underskrevet af Statsbiblioteket på vegne af projektet i maj 2009 og afventede herefter Kulturministeriets godkendelse i et år.

Aftalen omfatter retten til at indeksere materialet og retten til at vise såkaldte snippets. Det viste sig umuligt at få forhandlet aftalen til at gælde alle medlemmer af DEFF, en begrænsning som blev accepteret efter forhandling med DEFF. Aftalen er tidsbegrænset til 2 år og skal derfor meget snart genforhandles.

Perspektiv og forretning

En evaluering af projektet skal naturligvis ses i forhold til virksomhedernes mission og strategi. Bibliotekernes mål er klart en serviceforbedring, hvis pris står mål med merværdien for brugerne, men Aarhus Universitetsforlag har ud over en indtjening også en mission om at støtte udbredelsen af viden fra AU, og distributøren har et klart indtjeningskrav.

Der er altså tale om tre forretningsmodeller, hvor udbyttet måles med forskellige målestokke. En ide i projektet var, at de tre forretningsmodeller kunne støtte hinanden, så en øget omsætning kunne reducere prisen for de to offentlige institutioner. Konklusionen er imidlertid, at dette ikke kan lade sig gøre.

Publizon

Fra starten har projektet været tilrettelagt sådan, at Publizon skulle skanne billigt mod til gengæld at kunne tjene penge på en senere viderebearbejdning. Der har været ideer fremme om udnyttelse, men endnu er ingen af dem blevet realiseret. Publizon

har således endnu ikke tjent penge på sekundær udnyttelse og har tabt penge på at deltage i projektet. Publizon ville, i tilfælde af fremtidige projekter, enten sikre sig fuld betaling for arbejdet, eller forhåndsaftaler med forlaget om et vist minimumskøb til forud aftalte priser.

Aarhus Universitetsforlag

Forlaget har arbejdet med forskellige muligheder for at udnytte de digitaliserede bøger, men endnu ikke fundet en attraktiv model. Forlaget havde ved projektstart en forestilling om, at andre akademiske forlag også ville digitalisere i stor stil, og at der kunne etableres en fælles platform og dermed en markant afsætningskanal for akademiske e-bøger. Flere forlags deltagelse ville kunne nedbringe forlagets nødvendige investering i en platform til præsentation og evt. salg af det digitale materiale, og det større antal titler ville give mulighed for at tiltrække stor opmærksomhed fra potentielle købere.

På plussiden er dog potentialet. Forlaget har nu mulighed for at få adgang til alle filerne billigt, men i den anvendte model kun gennem digitaliseringsfirmaet. De digitaliserede bøger kunne genudnyttes til udgivelse som e-bøger, som book-on-demand eller andre former for udnyttelse, men det er ikke sket endnu.

Statsbiblioteket og Det Kongelige Bibliotek

De to biblioteker havde flere mål med projektet, herunder at forbedre brugeroplevelsen. For at få lov til at bruge materiale til indeksering og til snippet-visning kræves en aftale med Copydan, som groft sagt er de enere vindere i dette projekt.

Modellen med aftalelicens fungerer godt nok, men de konkrete vilkår i aftalen vil skulle tilpasses forholdene i hvert enkelt tilfælde, hvis andre aktører vil bruge en lignende model. Såvel Copydan som bibliotekerne har da også eksplicit gjort opmærksom på, at aftalen er "uden præjudice" – at vilkår og priser ikke nødvendigvis kan overføres til andre projekter og aftaler. I øjeblikket overvejer de to biblioteker, om merværdien for brugerne er omkostningerne værd.

Erfaringer med søgning

Merværdien for brugerne skulle være en forbedret søgning og vurdering af materialet, og både Statsbiblioteket og Det Kongelige Bibliotek har gennemført forsøg med at inkludere fuldteksterne i genereringen af indekser og med at vise snippets.

Statsbiblioteket valgte at hæfte fuldteksterne op på de eksemplarer (metadata), som allerede



Statsbiblioteket og Det Kongelige Bibliotek havde flere mål med projektet, herunder at forbedre brugeroplevelsen.

eksisterede i katalogen. Dermed er fuldteksterne indekseret sammen med de relevante metadata og dermed søgbare. Da Statsbibliotekets søgeværktøj, Search, på forhånd er tænkt til at kunne håndtere fuldtekster, har dette ikke voldt nogen problemer.

Proceduren på Det Kongelige Bibliotek var noget mere kompliceret, men efter nogle forsøg blev det besluttet at betragte fuldteksten som en berigelse til de eksisterende bogposter, hvorpå disse berigede poster blev fuldtekstindekseret. For at gøre dette blev der udviklet et modul til Primo, et såkaldt Enrichment Plug-in.

Begge biblioteker havde også den erfaring, at for at de få fuldtekstbøger ikke totalt skulle dominere mange søgeresultater, var det nødvendigt at give disse en negativ vægt i rankeringen. Endvidere er en fuldtekstsøgning, som ikke følges op af en form for snippet, forvirrende for brugerne, som ikke forstår, hvorfor en given bog vises.

Samlet er det teknisk muligt at integrere fuldteksterne, men de begrænsede erfaringer med, hvordan fuldteksterne påvirker kvaliteten af søgefunktionerne, peger imidlertid på, at fuldteksterne er svære at vægte korrekt og giver mere værdi end støj.

En mulig forklaring kan være den meget lille mængde fuldtekstmateriale i forhold til de store mængder materiale, hvortil der er knyttet relativt få ord i form af den bibliografiske post. Forekommer søgeordet ofte i fuldteksten, vil denne post få en stor vægt og

blive placeret øverst på resultatlisten. For at få et tilfredsstillende resultat har det været nødvendigt at give en relativt lille vægt på fuldtekstmaterialet.

Dette forhold forventes at ændres, når mere materiale er tilgængeligt for fuldtekstindeksering. Denne forklaring er konsistent med at søgninger i Google Book Search fungerer godt.

Nye partnerskaber?

Projektet har givet mange erfaringer, men har ikke resulteret i den forventede serviceforbedring eller identificeret den gode forretningsmodel. Men trods de umiddelbare negative konklusioner, har projektet dog også identificeret et potentiale i udvidede partnerskaber.

Projektet havde fokus på at forbedre brugerens mulighed for at finde og vurdere materiale og ikke på eksempelvis at understøtteudgivers forretning i form af større eksponering af deres materiale og dermed en større brug. Man kunne forestille sig forretningsmodeller, som bygges op omkring en fælles interesse for at give adgang til materialet – med andre ord, hvor digitaliseringen betales via en andel i en afgift.

Så selv om projektet ikke i sig selv kan udråbes til en stor succes, så har processen peget på fælles interesse og mulige nye samarbejder – med det formål at sikre udbredelsen af dansk faglitteratur indenfor forskningsverdenen. ●