



Videncenter for Centraladministrationen

Deff

Slotsholmsgade 12
1216 København K

Telefon 7226 9891
7226 9895
Fax 7226 9899

Mail dab@dab.dk
Web www.dab.dk

EAN 5798 00041 6901
CVR 20 58 23 08

Status på projektet

30. oktober 2015

”Automatisk indhentning af oplysninger om statslige myndigheders digitale publikationer”

Overordnet beskrivelse af projektet, herunder ændringer ifht oprindelig projektbeskrivelse

Projektet drejer sig i korthed om en automatiseret model for indsamling og registrering af oplysninger om publikationer udgivet af statslige institutioner. Det drejer sig om udgivelser fra ca. 120 departementer, styrelser, direktorater og lignende institutioner. Disse institutioner udgiver skønsmæssigt ca. 2.000 digitale publikationer om året.

Flere aktører er beskæftiget med indsamling, registrering og opbevaring af statslige digitale publikationer. Den vigtigste er KB, der ifølge pligtafleveringsloven skal høste alle danske domæner og opbevare det høstede. Opbevaringen sker i Netarkivet, men fremfindning af materialet er uhyre vanskeligt. Dels fordi Datatilsynet har fastsat strenge betingelser for at få adgang til Netarkivet, og dels fordi der ikke foretages en registrering af det høstede.

DBC registrerer dele af det digitale materiale, der udgives. Men registreringen er dels ikke komplet mht. dækning af alle de publikationer, der udgives, og dels anvendes selektionskriterier, der udelukker store dele af yderst relevant materiale.

Som centraladministrationens bibliotek har DAB siden slutningen af 1970'erne registreret centraladministrationens egne publikationer og har de senere år været opmærksom på den stigende mængde publikationer, der udgives i digital form. De centraladministrative institutioner har mulighed for selv at udfylde og indsende såkaldte anmeldelsesblanketter til DAB på digitale publikationer, der efterfølgende bliver registreret i Statens Netbibliotek - et virtuelt bibliotek omfattende statslige publikationer. Det er dog kun få institutioner, der benytter blanketten. Der er formentlig mange forskellige årsager til, at de nuværende procedurer ikke fungerer, og det er vanskeligt præcist at sige, hvad årsagen er.

Der er fra DABs side gjort en stor indsats for at udbrede kendskabet til blanketten, men vi må konstatere, at det ikke har gjort den store forskel i institutionernes indberetningsadfærd. For at råde bod på dette bruger DAB mange personeres ressourcer på at gennemgå statslige hjemmesider for at tjekke for digitale publikationer, som ikke er blevet anmeldt.

De statslige udgivere ikke er forpligtet til at opbevare deres publikationer på deres hjemmesider. Det er helt op til den enkelte udgiver at beslutte, hvor længe publikationerne skal kunne findes på hjemmesiderne. Så hvis udgiverne mener, at en publikation er uaktuel eller har pladmangel på serveren, kan de slette publikationer. Sletning af publikationer forekommer ofte, når institutioner nedlægges, sammenlægges, eller ændrer navn og dermed også web-adresse. Institutioner skifter med mellemrum soft- eller hardware, og dette kan ligeledes medføre, at publikationer slettes.

Udgiverne er heller ikke forpligtet til at stille et søgesystem til rådighed til fremfindning af publikationer på deres hjemmeside. Der er stor variation inden for søgesystemer hos udgiverne; Efter DABs erfaring er det mest almindelige, at udgiverne nøjes med at gruppere publikationerne efter udgivelsesår. Nogle udgivere går videre end det og foretager en alfabetisk inddeling efter publikationernes titel under de respektive udgivelsesår. Nogle udgivere stiller derudover et søgefelt til rådighed, hvor der kan foretages fritekstsøgning i publikationernes titler. DAB er også stødt på hjemmesider, hvor publikationerne først er inddelt efter type, som fx på Justitsministeriets hjemmeside, hvor publikationstyperne Betænkninger, Redegørelser, Hvidbøger og Andre publikationer er benyttet. Dvs. at valg af den rette publikationstype er afgørende for, om en bruger finder det materiale, der søges efter. Andre udgivere – fx Sundhedsstyrelsen – stiller flere søgemuligheder til rådighed. Både en kronologisk indgang, en indgang efter publikationstype (Regler, Tal & analyse, Faglig rådgivning og Inspirationsmateriale) samt et fritekstsøgefelt.

Det DAB så som den centrale udfordring og projektets formål var at få udviklet system, der kunne overvåge statslige hjemmesider for digitale udgivelser – foretage en registrering – omforme registreringen til en post i Statens Netbibliotek.

Gennem de senere år har DAB drøftet projektet med en lang række leverandører; indtil nu, uden held.

Vi vurderer, at årsagen til dette formentlig er, at det ikke alene er vanskeligt at

udvikle det rette materialesporings- og indhøstningssystem; der er også en lang række udfordringer i registreringen, opbevaringen og tilgængeliggørelsen af materialet.

Der ud over, er der selvfølgelig nogle ophavsretsmæssige krav til den del af materialet som har en mere mudret proveniens.

Efter ganske mange overvejelser og drøftelser med både eksterne leverandører, og indhentning af erfaringer fra internationale aktører vurderer vi nu, at det ikke vil være hverken muligt eller ønskværdigt, at indhøste høste og registrere publikationerne automatisk.

Der er ganske enkelt for mange risikomomenter i denne proces, der har ført til, at projektet nu fokuserer på den automatiske indhøstning. Det er formentlig ikke muligt at undgå momenter af manuelt arbejde i forbindelse med selve registreringen. Projektet vil derfor indeholde en udarbejdelse af modellen for registrering og afdækning af omfanget den ønskede eliminering af manuelt arbejde.

Med DABs og KBs udvikling af PID-systemet, og den særlige struktur med anvendelse af KBs 'Netarkivet' som en slags repository, er der nu etableret et system, der er let at administrere og som har en meget høj grad af sikkerhed for publikationernes tilgængelighed over tid. En løsning, som flere andre institutioner viser interesse for nu.

Årsagen til, at dette projekt har trukket ud, er naturligvis mangfoldige, men en vigtig grund er formentlig, at det simpelt hen var forud for sin tid. DAB og KB flere gange i den nuværende afklaringsfase kigget på muligheden af, at indlede et udviklingssamarbejde. Det ser nu ud til at kunne realiseres, og vi håber på at kunne afklare det mere præcist i slutningen af 2015. Den helt store fordel for begge parter er, at vi både kan udnytte det allerede udviklede PID-system, og at vi kan gøre effektivt brug af 'Netarkivet' til selve høstningen.

Hvordan understøtter projektet implementeringen af DEFFs mål og virke?

Projektet vil styrke Deff's ambition om national 'open access', vha. sikring af bevaring og adgang til materiale der som udgangspunkt er frembragt og stillet gratis til rådighed for offentligheden af statslige institutioner.

Det færdige og gennemførte projekt vil stille digitalt materiale til rådighed for borgere, virksomheder og institutioner på en sammenhængende og enkel måde.

Projektet vil desuden styrke sammenhæng og samarbejde mellem flere aktører i biblioteksvæsnet og endvidere bygge videre på og styrke anvendelsen og udnyttelsen af allerede gennemførte Deff-projekter.

Har andre nationalt eller internationalt arbejdet med samme emne og i givet fald, hvad er sammenhængen med nærværende projekt?

Da vi gik i gang, fandtes der ikke nogen tilsvarende systemer. Det har vist sig, at der på internationalt plan er begyndt at ske noget. Fx har British Library arbejdet med et system til arkivering af offentlige publikationer; et system der på nogen måder kan minde om det der er beskrevet i DABs projektansøgning.

Erfaringen fra England er, at systemet er et både komplekst og dyrt at udvikle. Det skal dog siges, at det danske projekt vil kunne drage nytte af englændernes erfaringer.

Ud over de engelske erfaringer, er der en række andre projekter vi ser nærmere på. I OCLC-regi er man således begyndt at undersøge hele feltet omkring registrering af web ressourcer i OPAC regi.

Hvilken målgruppe vil få gavn af projektet, og er der evt. en sekundær målgruppe?

Vi vurderer, at projektet vil kunne skabe værdi for en lang række brugere: borgere, virksomheder, biblioteker, studerende m.fl.

Essensen er, at projektet vil kunne sikre, at borgere kan få et samlet overblik over, og adgang til, de publikationer, som statslige myndigheder udgiver. Det vil kunne sikre, at statslige publikationer samles ét sted og kan fremfindes ved benyttelse af et professionelt søgesystem.

I sammenhæng med det allerede gennemførte "PID-projekt" vil der også være sikret permanent adgang til alle publikationerne.

Hvilke målbare succeskriterier opstilles for projektet, og hvordan skal disse vurderes?

Successkriteriet er i og for sig enkelt nok: indhøstningssystemet skal være i drift, og der skal være etableret en funktionsduelig driftsmodel, så det sikres at så godt som alle de statslige publikationer der beskrives som bevaringsværdige, også registreres og tilgængeliggøres for eftertiden.

Beskrivelse af kommunikationsplan

Projektets resultater bør naturligvis formidles på sædvanlig vis til brancheaktører nationalt og internationalt, vha. nyheder, conferenceoplæg, artikler mv. Vi ved der er stor international interesse for problematikken, så det vil være med til at profilere en særlig dansk . Det vil uden tvivl føre til øget brug af både det kernemateriale der registreres, men mindst lige så vigtigt, at modellen og systemet kan bruges af andre med lignende behov. Endvidere vil det kunne føre til øget brug af KBs 'Netarkivet', en enorm resurse, der pt. ligger relativt uudnyttet hen.

I forhold til synlighed, vil det være helt naturligt at ... I øjeblikket har DAB indført flg. tekst i samtlige poster knyttet til projektet: " Det Administrative Biblioteks

digitale arkiv er støttet af DEFF Danmarks Elektroniske Fag- og Forskningsbibliotek.”

Revideret tidsplan

Da der endnu ikke er en endelig tidplan i forhold til projekts gennemførelse i samarbejde med KB, kan vi ikke komme det nærmere end, at det forventes, at projektet gennemføres i 2016.

Der er ikke nogen ændringer til projektets budget på i alt 355.000 kr., hvoraf der var ansøgt om 300.000 kr. i den oprindelige projektbevilling.

København, den 30. oktober 2015



Niels H. Jensenius
Bibliotekschef