Danmarks Elektroniske
Fag- og Forskningsbibliotek

Programområdet
Informationsforsyning

DEFF databrønde for artikler og e-bøger

Rapport med anbefalinger til Styregruppen fra de to afklaringsprojekter

---

Projektdeltagere:

AALBORG UNIVERSITETSBIBLIOTEK:
Per Steen Hansen, Kasper Løvschall
AARHUS SCHOOL OF BUSINESS BIBLIOTEK:
Anna Mette Morthorst
COPENHAGEN BUSINESS SCHOOL BIBLIOTEK:
Lars Nondal, Henriette Jakobsen, Birgit Brejnebøl
DANMARKS TEKNISKE INFORMATIONSCENTER:
Mogens Sandfær, Projektleder for artikler, Hans Meulengracht-Madsen, Christian Tønsberg, Tove Nielsen, Kirsten Møller-Nielsen
DANSK BIBLIOTEKSCENTER
Tommy Schomacker, Per Mogens Petersen, Kurt Bonde Poulsen, Jacob Poulsen
DEFF SEKRETARIATET
Bo Öhrström, Lise Mikkelsen, Anette Schneider, Janne Vendt
DET KONGELIGE BIBLIOTEK:
Ole Holm, Projektleder for e-bøger, Anders Christensen, Ingbritt Butina, Barbara Melchior, Christian Vandel, Anne Skov
PROGRAMGRUPPEN FOR NYE INSTITUTIONER:
Niels Jensenius, DAB, og Lars Kofoed-Jensen, Frederiksberg Seminarium
STATSBIBLIOTEKET:
Birte Christensen-Dalsgaard, Hans Lauridsen, Vibeke Christensen
SYDDANSK UNIVERSITETSBIBLIOTEK:
Ingelise Stæhr, Bjarne Christensen

København den 4. november 2008

1. INDLEDNING

DEFF igangsatte i april 2008 to afklaringsprojekter på området nationale databrønde – med fokus på henholdsvis artikler og e-bøger. De to projekter afrapporteres her samlet ligesom yderligere information kan studeres på projekternes fælles wiki-site – bl.a. meget udførlige detaljer om det relevante/ønskelige dataindhold i form af licenspakker, tidsskrifter, bibliografier mv.: https://infoshare.dtv.dk/twiki/bin/view/DEFFDataWells/WebHome

I løbet af 2007 blev spørgsmålet om DEFF databrønde (eller et nationalt samarbejde om hjemtagelse og udnyttelse af globale datasæt af biblioteksmæssig interesse) aktualiseret af en bred interesse i servicekonceptet "Integrated Search and Delivery". Dette koncept forudsætter netop databrønde, der på enkel vis stiller metadata til rådighed for bibliotekernes søgemaskiner.

I FC's henvendelse til DEFF af 29.5.2007 kommer denne aktualisering klart til udtryk, idet FC henstiller, at DEFF i de kommende år entydigt satser på integrated search samt udvikling og drift af databrønde som højeste prioritet (se bilag H).

DEFF Styregruppen havde allerede tidligere på året nedsat en ad-hoc arbejdsgruppe på området "Fælles data til Integrated Search". Denne afsluttede sit arbejde den 30.8.2007 og referatet fra det sidste møde er vedlagt som bilag I. Heri beskrives en køreplan for området, som ind til videre har været styrende for aktiviteterne:

Arbejdet bør organiseres i to klart adskilte faser:

1. ***Afklaringsfase**, der besvarer generiske spørgsmål som: Hvad er en DEFF databrønd – koncept og profil? Hvilke datastrømme, hvilke potentielle leverandører? Hvilket aftalemæssigt grundlag for dataleverancer? Hvilke aftagere i DEFF landskabet? Hvilke behov og ønsker har disse? Hvilke typer bidragsmodeller til finansiering af driften kan tænkes, vil være ønskelige/realistiske? Hvilke formater og standarder findes/kan/bør følges? Hvad er de internationale erfaringer? -- Denne fase kan inddrage alle relevante parter – også sådanne, der evt. måtte overveje at deltage i et senere udbud af operatøropgaven.*

2. ***Udbudsgrundlag**, hvor afklaringen overfor omformes til et reelt udbudsgrundlag med kravspecifikation mv. – antagelig bistået at et udbudsvant konsulentfirma. Denne fase kan ikke inddrage parter (eller personer med nære relationer til sådanne), der ønsker at deltage i det kommende udbud af operatøropgaven.*

DEFF sekretariatet må sørge for at den præcise grænsedragning mellem fase 1 og 2 er korrekt og klart kommunikeret. Efter fase 2 kan DEFF sekretariatet gennemføre en udbudsforretning, antagelig bistået af et konsulentfirma.

Med denne afrapportering afsluttes ovennævnte fase 1, og de videre tiltag afventer stillingtagen i DEFF Styregruppe og Sekretariat.

2. INDHOLDSFORTEGNELSE

Arbejdet blev indledt af en fælles workshop i maj, hvorefter en række deludredningsgrupper arbejdede frem med en afsluttende workshop i september. Som det fremgår nedenfor, er rapporten disponeret med et hovedafsnit og 9 bilag, hvor hovedafsnittet stort set svarer til deludredningen fra gruppen for "koncept og organisation" og de fleste bilag er opsummeringer af de øvrige deludredninger, der kan studeres fuldt udfoldet på projektets wiki-site: https://infoshare.dtv.dk/twiki/bin/view/DEFFDataWells/WebHome

1. Indledning 2
2. Indholdsfortegnelse 3
3. Executive summary 4
4. Hovedrapport 5
   - Baggrund 5
   - Formål 5
   - Databrøndens produkter 6
   - Databrøndens kunder 6
   - Databrøndens indhold 6
   - Leverandører 7
   - Aftagerprofiler 7
   - National databrønd – serviceniveauer 8
   - Databrøndens organisation 10
   - Databrøndens styring 10
   - Databrøndens finansiering 11
   - Databrøndens etablering 13
   - Tidsplan 13
5. Bilag A: Andres erfaringer – danske og udenlandske 14
6. Bilag B: Spørgeskemaundersøgelse af bibliotekernes behov 17
7. Bilag C: Business case 23
8. Bilag D: Datastrømme, leverandører, licenser - Artikler 25
9. Bilag E: Datastrømme, leverandører, licenser – E-bøger 27
10. Bilag F: Metadata, emnedata og værkklynger 29
11. Bilag G: Krav til licenser og forlagenes dataleverancer 31
12. Bilag H: Forskningsbibliotekernes Chefkollegiums henvendelse 32
13. Bilag I: Referat af sidste møde i "Fælles data til Integrated Search" 33

3. EXECUTIVE SUMMARY

Afklaringsprojektet har gennem sine undersøgelser

- bekræftet behovet for en DEFF databrønd for tidsskrifter, artikler og e-bøger
- bekræftet og stipuleret det økonomiske rationale i en fælles løsning på dette område
- klarlagt de internationale erfaringer på området
- klarlagt de potentielle kunder og aftagerprofiler
- klarlagt de potentielle leverandører og de licensmæssige udfordringer
- afdækket bibliotekernes forventninger og præferencer ifm. en sådan databrønds
    - organisering og styring
    - finansieringskoncepter
    - indhold og dækning
    - produkter og serviceydelser

Afklaringsprojektet anbefaler DEFF Styregruppen

- at beslutte at implementere en sådan DEFF databrønd hurtigst muligt
- at iværksætte beslutningen som en højt prioriteret opgave for DEFF Sekretariatet – evt. støttet af eksterne konsulenter
- at udstyre DEFF sekretariatet med de nødvendige ressourcer og bemyndigelser til hurtigt og effektivt at udføre opgaven
- at gøre plads i de kommende års budgetter til en væsentlig DEFF-støtte til databrøndens udvikling og drift

4. HOVEDRAPPORT

Baggrund

I forskningsbiblioteksverdenen har der længe været fokus på, hvordan man får formidlet de meget omkostningsfulde elektroniske ressourcer til brugerne, således at brugen af e-materialer bliver endnu højere. Et centralt problem har været, at søgemulighederne i de licensbelagte elektroniske materialer ikke er gode nok. Brugerundersøgelser har samtidig vist, at brugerne ønsker ét søgefelt til alt materiale ligesom Google kombineret med præsentation af søgeresultaterne ligesom Amazon.

Nogle af universitetsbibliotekerne har arbejdet med integrated search som en mulig løsning. I integrated search indsamles og indekseres alle informationsressourcer samlet, hvorved bibliotekerne kan leve op til kravet om ét søgefelt. Nogle af de problemer, som skal løses for at få en optimal løsning, er;

1) at det er vanskeligt at få adgang til alle relevante data,
2) at det stiller store krav til hardware,
3) at det stiller store krav til it-kompetencen,
4) at det stiller store krav til dokumentlevering og adgangsstyring.

Nationale databrønde vil være et fælles initiativ, hvor man med fælles styrke ville have bedre muligheder for at fremskaffe de nødvendige data fra forlagene. Et fælles initiativ vil også være en fordel i forhold til de problemer, som skal løses i form af it-kompetencer, søgemaskiner, services og dokumentlevering.

Der er ingen tvivl om, at hele projektet omkring organisering og implementering af nationale databrønde på alle relevante materialetyper, den efterfølgende tilgængeliggørelse og formidling m.m. bliver en meget stor opgave for biblioteksektoren i årene fremover.

Af hensyn til sandsynligheden for at kunne implementere forslagene i forprojektet bliver det nødvendigt at gøre klart hvilke opgaver, der hører til en national databrønd, og hvilke der ligger udover, og som bør håndteres i projekter sideløbende med udviklingen af nationale databrønde for e-bøger og e-tidsskriftsartikler. Disse er de materialetyper, som i første omgang er i fokus.

Formål

Formålet med nationale databrønde for elektroniske materialer er at få etableret et lager for et nationalt datasæt for forskellige digitale materialer, som på ensartede vilkår i videst muligt omfang stilles til rådighed for alle institutioner i Danmarks Elektroniske Fag- og Forskningsbibliotek (DEFF). Datasættet skal f.eks. kunne anvendes i forbindelse med integreret søgning.

I de kørende forprojekter er det databrønde for henholdsvis e-bøger og e-tidsskriftsartikler, som er i fokus, men på sigt skal modellen kunne udvides til andre digitale materialer. En databrønd skal kunne tilbyde import og eksport af data samt forskellige grader af databehandling, og i tillæg til dette og afhængig af efterspørgsel funktioner som søgning og fremvisning af data således, at det enkelte bibliotek kan trække på en national løsning på de områder, hvor dette ønskes.
I denne udredning gives et bud på, hvordan en national databrønd skal styres, organiseres, finansieres mv. Udgangspunktet er e-bøger og e-tidsskriftsartikler, men modellen kan udvides med flere datatyper.

Blandt deltagere i de to analyseprojekter har der entydigt været interesse for at DEFF, på vegne af styregruppen for de nationale databrønde, bliver leverandør af den nationale databrønd for e-bøger, e-tidsskriftsartikler, databaser, encyklopædier m.m., hvilket derfor er anvendt som udgangspunkt.

Databrøndens produkter

Databrøndens produkter er overordnet:

- Lager for et nationalt datasæt for fortrinsvis e-bøger og e-tidsskriftsartikler
- Funktioner til import og eksport af data
- Funktioner til forskellige grader af databehandling
- Mulighed for tillægsfunktioner udenfor den egentlige brønd såsom søgning og fremvisning af data, således at det enkelte bibliotek kan trække på en national løsning på de områder, hvor dette ønskes.

Databrøndens kunder

Potentielt er alle ikke-kommercielle institutioner i Danmark kunder. Der tænkes her primært på DEFF's fag- og forskningsbiblioteker, men også folkebibliotekerne kan på sigt tænkes at være interesserede i dele af databrønden. DEFF databrøndens kunder vil i udgangspunktet maksimalt svare til DEFF licenssekretariatets kunder. En kunde betaler for brug af databrønden via et årligt abonnement.
Mens den fælles licensadministration fuldt ud dækkes af DEFF bevillingen, forudses det, at databrønden i et vist omfang vil være finansieret af bidrag fra de deltagende biblioteker. Som følge af licensadministrationen, råder DEFF sekretariatet allerede over instrumenter til fakturering af deltagerne såvel som forhandling med og betaling af data- og teknik-leverandører.

Databrøndens indhold

Til opstart af databrønden defineres et data startsæt, som så senere kan udvides med flere pakker efter ønske og prioritering fra den nationale databrønds styregruppe. Spørgsmålet om databrøndens indhold er behandlet i Bilag D og E om Datastrømme, leverandører, licenser.

I praksis vil aftalerne om leverance af data indgå som en del af licensforhandlingerne. Når det er muligt og ønskeligt, vil licensaftalerne blive udvidet med paragraffer der regulerer leverancen af data såvel som DEFF databrøndens udnyttelse og videredistribution af data. Mens det i udgangs-

punktet er ønskeligt, at "alle" må søge i og fremvise alle metadata – lokalt som centralt - vil det først som en del af licensforhandlingerne blive klart, hvorvidt leverandørerne vil tillade dette og hvorvidt dette vil medføre krav om yderligere licensbetalinger. I Bilag G: "Krav til licenser og forlagenes dataleverancer" behandles dette aspekt lidt mere detaljeret. Metadata anses for at tilhøre de biblioteker, som er tilknyttet den pågældende licens via en brugsret i licensaftalen. Hvis et bibliotek fratræder databrøndskonsortiet tager man en kopi af "sine" metadata med.

Leverandører

Både private firmaer og offentlige institutioner kan være leverandører til databrønden. Kunderne bestemmer indholdet af databrønden. I nærværende projekt er forlag og mellemhandlere med e-bøger og e-tidsskriftsartikler leverandører til databrønden.

Aftagerprofiler

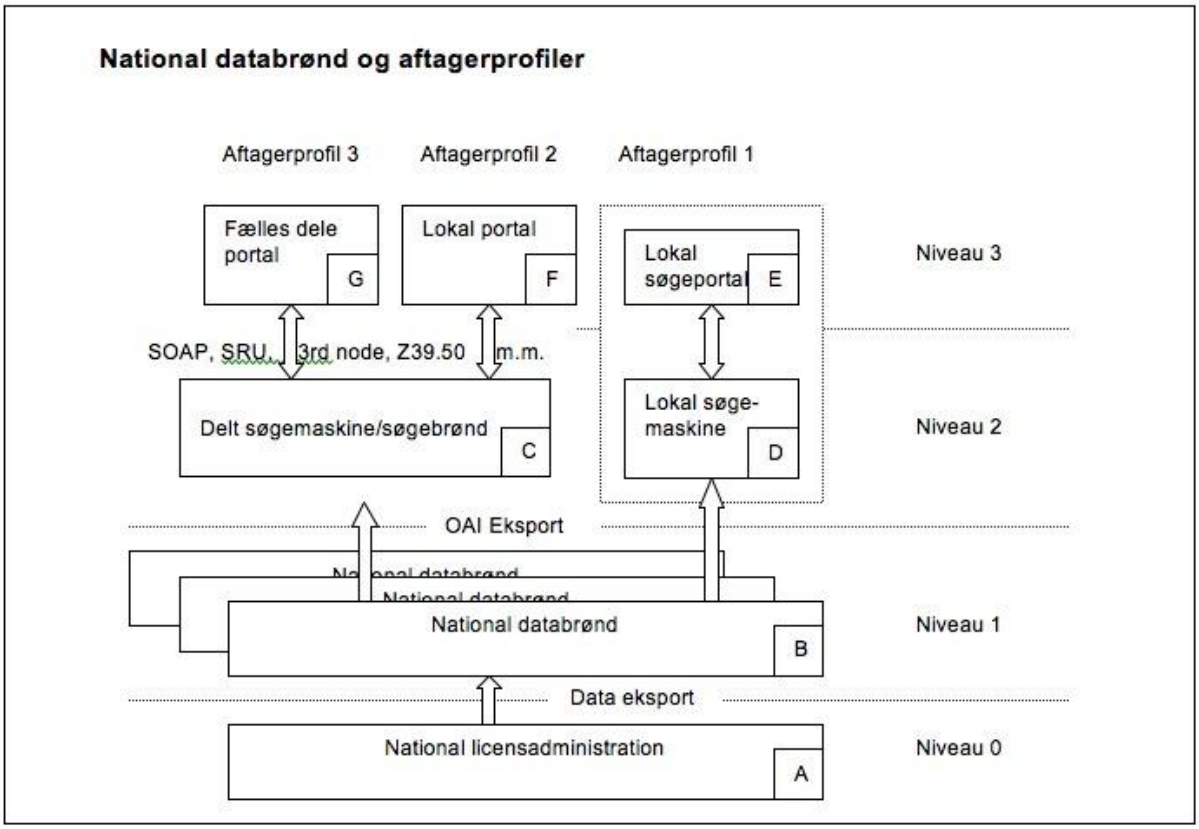
I forprojektet og i spørgeskemaundersøgelsen har man arbejdet med 3 aftagerprofiler. I profil 1 står biblioteket selv for lagring af data, søgemaskine og brugergrænseflade. Dette stiller naturligvis meget store krav til økonomi og tekniske ressourcer. I profil 2 har man valgt en central lagring og central søgning af data, mens visning af data foregår i eget system. I profil 3 arbejdes der med en central løsning på alle tre områder, dvs. central lagring, søgning og grænseflade.
I nedenstående tabel er der givet et bud på nogle af de overvejelser, man bør gøre i forbindelse med valg af hovedprofil og dermed løsning.

Oversigt over aftagerprofiler

	Profil 1 lokal lagring, søgning og visning af data			**Profil 2** central lagring, central søgning, visning af data i eget system			**Profil 3** central lagring, central søgning samt central webgrænseflade		
Emner som kan overvejes ifm valg af hovedprofil:	**A** artikler	**B** tidsskrifter	**C** ebøger	**A** artikler	**B** tidsskrifter	**C** ebøger	**A** artikler	**B** tidsskrifter	**C** ebøger
Lagring af data	lokalt	lokalt	lokalt	Centralt	centralt	centralt	centralt	centralt	centralt
Søgemaskine	lokalt	lokalt	lokalt	Centralt	centralt	central	centralt	centralt	centralt
Brugergrænseflade	lokalt	lokalt	lokalt	lokalt	lokalt	lokalt	centralt	centralt	centralt
Krav til bibliotekets tekniske ressourcer	Meget store			Mellem			Små		
Fleksibilitet	Meget høj			Mellem			Lille, standard løsning		
Krav til lokale systemer	Meget store			Mellem, skal kunne anvende webservice eller Z39.50			Små		
Søgning	Store krav, da søgning foregår lokalt			Ingen lokale krav, da søgning foregår ifm databrønd			Ingen lokale krav, da søgning foregår ifm databrønd		

National databrønd – serviceniveauer

En national databrønd indgår i en model med 4 niveauer. Det nederste niveau 0 er DEFF licensadministration, niveau 1 er den nationale databrønd, niveau 2 er søgemaskiner og niveau 3 er søgeportaler. Arkitekturen er skitseret i nedenstående model. De enkelte elementer i modellen er markeret med bogstaver, og hvert element gennemgås i den efterfølgende tekst.



A) National licensadministration

National licensadministration henviser til de licenser, som DEFF varetager på vegne af bibliotekerne. Dette arbejde omfatter på nuværende tidspunkt licensforhandlinger, kontraktarbejde, administration m.m. Projektet har dog tillige vist, at der er interesse for et samarbejde omkring vedligeholdelse af en national knowledge base i forbindelse med link resolver, der skal spille sammen med en base indeholdende information om adgangsrettigheder. Et sådant projekt ville skulle køres som et sideløbende projekt, og eventuelle udbudsforpligtelser skal afklares. Der er naturligvis også mulighed for at uploade data fra lokale licenser til en national databrønd. Dette sker under forudsætning af, at man lokalt har sikret sig de nødvendige rettigheder både nationalt og lokalt på vegne af hele fællesskabet. Man skal dog være opmærksom på, at hver dataupload vil være forbundet med en udgift i forbindelse med databehandling i element B National databrønd.

B) National databrønd

Der vil være en national databrønd for hver datatype f.eks. e-bøger og e-tidsskriftsartikler. Dette skyldes hensynet til datas forskellighed og mulighederne for at kunne trække præcist de data ud, man ønsker til en bestemt sammenhæng. F.eks. vil nogen måske gerne have e-bogs poster i deres katalog, mens e-tidsskriftsartikler kræver en anden søgeindgang. I en national databrønd samles metadata m.m. fra mange forskellige kilder og i flere forskellige og mere eller mindre standardiserede formater. Der udføres forskellige typer af databehandling, såsom deduplikering, navneformsproblematikker m.m. Databrønde vil kunne eksportere data i forskellige formater og indtil videre viser en spørgeskemaundersøgelse, at der er interesse for mindst 3 formater, 1) "råt" format, 2) Databrønd super format og 3) DanMarc 2 format.

Krav og ønsker til databehandlingen og specifikation af de enkelte formater bliver en væsentlig opgave at få specificeret nærmere i forbindelse med tilrettelæggelsen af det videre arbejde.

C) Delt søgemaskine/søgebrønd

Data fra databrønden kan høstes via OAI-PMH til en fælles søgemaskine. Her kunne en gruppe af Primo kunder gå sammen om en fælles søgemaskine og dermed opnå driftsfordele.

D) Lokal søgemaskine

Der er også mulighed for at OAI eksportere data til en lokal søgemaskine og foretage indeksering og søgninger lokalt.

E) Lokal søgeportal

Den lokale søgeportal er slutbrugernes søgegrænseflade, hvor data udstilles. Punkt D Lokal søgemaskine og E lokal søgeportal svarer til Summa løsningen på nuværende tidspunkt.

F) Lokal søgeportal

Punkt F lokal portal er en lokal søgegrænseflade, som trækker på en fælles søgemaskine. Kombinationen af delt C søgemaskine og F lokal søgeportal svarer på nuværende tidspunkt til Primo løsningen.

G) Fælles dele portal

At drifte og vedligeholde en lokal søgeportal kræver økonomiske såvel som tekniske ressourcer og er sandsynligvis forbeholdt nogle af universiteterne eller andre leverandører. Tanken med en fælles dele portal er, at biblioteker kunne slå sig sammen i et eller flere konsortier afhængigt af interesser og dermed dele udgifterne til udvikling og drift af en søgeportal. Man kunne forestille sig en meget standardiseret løsning med begrænsede muligheder for at lave ændringer, ligesom DEFF på nuværende tidspunkt har med DEFFnet. Hvilken teknisk løsning, man så vælger, er op til det enkelte konsortium og interesserede leverandører.

Sammenfatning
Projektet for en national databrønd er afgrænset til at behandle det 1. niveau i modellen, dvs. selve den nationale databrønd. Men der er en meget stor sammenhæng til de andre niveauer i modellen, og man må forvente, at der er en vis interesse i, at der efterfølgende følges op i forhold til de andre niveauer, således at der kan oprettes et Primo eller Summa konsortium. Man bør ligeledes få undersøgt interessen blandt mindre og mellemstore biblioteker for løsningen. Denne opfølgning kan foregå i DEFF-sekretariatet i samarbejde med Programområdet Informationsforsyning og Programområdet Nye Institutioner. Derudover skal det undersøges, hvordan man kan etablere dokumentlevering og adgangsstyring i forbindelse med en national databrønd, hvor alle søger i alles data.

Databrøndens organisation

Styrelsen for Bibliotek og Medier ved DEFF er leverandør af tjenesten den nationale databrønd til bibliotekerne. I praksis er DEFF her identificeret ved DEFF-sekretariatet, der kontraktstyrer og betaler den tekniske tjenesteleverandør. DEFF-sekretariatet kan indgå kontrakter i kraft at sin institutionsmæssige placering i Styrelsen for Bibliotek og Medier.

Den tekniske tjenesteleverandør kan eksempelvis være en enhed i et forskningsbibliotek, en kommerciel driftsleverandør eller en markedsmæssigt beslægtet leverandør. Tjenesteleverandøren udfører driften og udviklingen af den af den nationale databrønd på basis af de ønsker om ydelser, som bibliotekerne har interesse for at anvende og betale. Omkostningerne søges løbende minimeret muligvis gennem økonomiske incitamenter for operatøren, således at eksempelvis flere tilfredse kunder giver større gevinst for operatøren.

DEFF-sekretariatet kontraherer med biblioteker, der ønsker at anvende databrønden og fakturerer disse i følge de indgåede kontrakter. Deltagerne, som anvender data, har alle indgået en kontrakt med DEFF vedrørende, hvordan metadata m.m. må anvendes.

DEFF-sekretariatet kontraherer fortsat via det løbende licensarbejde med leverandører af data i dette tilfælde forlagene. Her sikrer DEFF ved indgåelse af nye licenskontrakter mulighederne for løbende dataleverancer til databrønden, mens leverancekontakter vil regulere de tekniske leverancer. Disse leverancekontrakter er enten dele af licenskontrakten eller separate kontrakter, og disse indgås ligeledes med DEFF-sekretariatet.

Databrøndens styring

Der oprettes én styregruppe med én repræsentant for hver kunde eller kundegruppe i konsortiet. Dette betyder, at der er én styregruppe uafhængigt af, hvor mange databrønde man efterhånden får etableret. Styrelsen for Bibliotek og Medier ved DEFF-sekretariatet udfylder rollen som styregruppeformand, og styregruppen sekretariatsbetjenes af DEFF-sekretariatet.

Styregruppen indrettes således, at dem, der betaler for databrøndens drift og ydelser, styrer databrønden. Der fastlægges tillige en beslutningsstruktur, hvor der er en nær sammenhæng

mellem størrelsen af betalingen og den tilhørende indflydelse (stor betaling medfører passende stor indflydelse).

En af de opgaver, som er forbundet med en udgift for databrønden, er tilførsel af flere data. Som udgangspunkt må databrønden starte med et datasæt af passende bred interesse. Dette må svare til de nuværende DEFF licenser, eftersom disse karakteriseres ved at flere biblioteker har tegnet abonnement på dem. Dernæst er der naturligvis spørgsmålet, hvorvidt de ønskede data kan fremskaffes fra forlag mv. på acceptable vilkår. Det er nødvendigt, at man klart definerer det første datasæt, så bibliotekerne såvel som den tekniske leverandør ved, hvad de går ind til.

Den videre udbygning af databrønden gennem løbende tilførsel af nye datakilder bliver et centralt spørgsmål for licenssekretariatet, den tekniske operatør og databrønds-styregruppen. Databrønds-styregruppen, der repræsenterer konsortiets betalende partnere/kunder, vil skulle fastsætte retningslinjer for, hvilke kilder der skal medtages, i hvilken rækkefølge og hvordan dette skal finansieres. Indlægning af data er ikke kun forbundet med øgede hardware- og drifts-udgifter, men indebærer også at formatproblemer skal løses, data skal deduplikeres m.m., og derudover skal man naturligvis kunne modtage opdateringer.

Databrøndens finansiering

Databrønden ses som en naturlig forlængelse af den service, organisation og finansieringsmodel, der har etableret sig på licensområdet. Således vil en model for finansiering være, at bibliotekerne finansierer dette med et væsentligt DEFF-tilskud, der aftrappes over et antal år. DEFF kan som på licensområdet addere til finansieringen gennem personale til sekretariatsbetjening af styregruppe, licensarbejde, daglig rekvirentfunktion for databrønden og lignende. Sidstnævnte skal ligeledes godkendes af DEFF styregruppen, og forventningen er, at der mindst skal tilføres én fuldtidsstilling.

Fordelingsmodellen mellem partnerne/kunderne må ligeledes være afklaret inden implementeringen sættes i værk. Basisudgiften inkluderer dataimport, -bearbejdning, -lagring, -eksport m.m. Basisudgiften kan fordeles efter årsværk eller lignende, og efter om der er tale om et universitet, specialbibliotek eller en anden uddannelsesinstitution. Umiddelbart synes der at være to modeller, der anvendes i beslægtede tilfælde:

- Fordelingsnøgle baseret på moderinstitutionernes FTE. Denne anvendes allerede i flere licenssammenhænge og er for universiteterne normalt klart defineret som VIP+STÅ, der autoritativt kan slås op på Danske Universiteters hjemmeside. For øvrige institutioner er der udviklet en praksis for beregning af kompatible FTE-tal – bl.a. i licenssekretariatet.
- Fordelingsnøgle baseret på bibliotekernes samledes informationsindkøb/accessionsbudget. Denne anvendes så vidt vides ikke i Danmark i dag, men er for nylig blevet præsenteret af non-profit "bevarings-databrønden" Portico, der fordeler udgifter efter LME "Library's Total Materials Expenditure", således som defineret af Association of Research Libraries (ARL) - http://portico.org/libraries/lme.html - Det er uklart, om den danske biblioteksstatistik tilbyder et tilstrækkeligt grundlag for en sådan nøgle. Det er ligeledes

uklart, om det vil frembyde vanskeligheder at aggregere LME-tal for en given institutions hoved- og institutbiblioteker mv.

Basisudgiften dækker også en mulighed for at downloade rå data, hvilket vil kræve, at man laver den nødvendige databearbejdning m.m. lokalt. Dette for at sikre at brøndens kunder kan (kilde for kilde) kan vælge at forestå raffineringen lokalt. Håbet er dog, at behovet for dette minimeres ved opbygningen af en national databrønd, som udfører denne opgave til alles tilfredshed.

Søgemaskine og søgeportal ligger udover en national databrønd og finansiering af disse skal afklares i de eventuelle konsortier, som opbygges. Man må forvente, at den samlede omkostning for tjenesten budgetteres årligt, og at en separat årlig abonnementsordning skal dække denne omkostning.

Der bør i priserne indregnes et tilskud på x%, som skal anvendes til fælles udviklingstiltag, tilførsel af data og lignende. Det præcise tilskud bør fastsættes i Styregruppen for den nationale databrønd.

Økonomi i opstartsfasen
Man kan på sigt forestille sig en række kundegrupper til en national databrønd: a) universiteter, b) specialbiblioteker, c) Primo konsortium, d) Summa konsortium, e) Professionshøjskoler. Men i opstartsfasen er det tvivlsomt, om alle er med. Man må derfor forvente, at de deltagere, der er med i opstartsfasen har en forholdsvis stor udgift, som så vil blive mindre, når flere deltagere melder sig. Dette forventes balanceret af et DEFF opstartstilskud, der aftrappes i løbet af nogle år.

Regneeksempel
I nedenstående regneeksempel tages der udgangspunkt i omkostningsestimaterne i Bilag C: Business Case, der anfører følgende forbehold:

"Indsamling og håndtering af data i en national databrønd er forbundet med mange personale ressourcer. Økonomisk set er der derfor besparelser at hente ved at gøre dette i fællesskab, således at opgaverne kan løses en gang for alle. Som det kan ses af nedenstående overslag, så skal der ikke være mere end 2 lokale løsninger, førend en fælles løsning er billigere. Mens forholdet mellem lokal og fælles løsning anses for velestimeret, tages der forbehold for selve udgiftsniveauet. Af gode grunde er dette ind til videre et gæt, hvad angår behovet for medarbejdere og udgifterne til løn, overhead og teknik."

Ligeledes gættes der i dette regneeksempel på antallet af deltagere/kunder, der hver især kan være et enkelt bibliotek eller et konsortium af biblioteker (f.eks. ved professionshøjskolerne). Der gættes på 4 deltagere det første år, 6 det andet og 8 det tredje og frem. Deltagernes samlede betaling forventes fordelt iht. en nøgle, der afspejler deres størrelse/tyngde, men denne er ikke kendt pt., hvorfor regneeksemplet anvender en ren gennemsnitsfordeling.

DEFF-tilskuddet stipuleres med 2/3 i år 1 og derefter faldende til 1/2, 1/3 og 1/6.

Drift i 2008 kroner	År 1	År 2	År 3	År 4
Omkostninger	3.000.000	3.000.000	3.000.000	3.000.000
DEFF tilskud	2.000.000	1.500.000	1.000.000	500.000
Deltagerbetaling	1.000.000	1.500.000	2.000.000	2.500.000
Antal deltagere	4	6	8	8
Gns. pr. deltager	250.000	250.000	250.000	312.500

Databrøndens etablering

Efter gennemførelse af denne foranalyse skal resultaterne diskuteres i biblioteksektoren, og opbakning til det videre forløb sikres formentlig gennem underskrift af bindende aftale. Dette gælder både etablering og efterfølgende løbende finansiering af databrønden.

Etableringen forventes at ske på basis af et udbud af den tekniske udvikling af den nationale databrønd samt driften af denne. Dette udbud forventes foretaget af DEFF for at muliggøre bud fra biblioteker og sikre uafhængighed. Det er en mulighed at gennemføre et EU-udbud efter modellen konkurrencepræget dialog.

Midler til gennemførelse af udbud kan søges skaffet gennem ansøgning af DEFF-styregruppen. DEFF-sekretariatet sikrer oplæg til DEFF-styregruppen med mulige scenarier og udgifter for henholdsvis udbud og direkte kontakt og udpegning til mulige leverandører, såfremt sidstnævnte muliggøres af lovgivningen.

Tidsplan

Den foreløbige tidsplan er:

Afslutning og afrapportering fra afklaringsprojekterne	30. oktober 2008
Notat vedr. national databrønd til DEFF-styregruppen	7. november 2008
Beslutning vedr. næste trin på DEFF-styregruppemøde	17. november 2008
Diskussion af databrøndsløsninger i programgrupperne	okt. - dec. 2008
Sikring af opbakning fra biblioteksektoren senest	31. december 2008
Skrivning af kravspecifikation, gennemførelse af udbud samt indgåelse af kontrakt	jan. - juli 2009

5. BILAG A: ANDRES ERFARINGER – DANSKE OG UDENLANDSKE

Den nationale databrønd og dens slægtninge

En tjeneste med egenskaber svarende til en dansk national databrønd (= et fælles, normaliseret datagrundlag) eksisterer næppe i dag, hverken i udlandet eller i Danmark.
Ser vi på eksisterende og planlagte tjenester rundt om i verden og i Danmark med lighedspunkter til den danske idé er udvalget stadig fåtalligt og forskelligartet.

Billedet udviser et spænd fra etablerede tjenester med bred indholdsmæssig dækning som OhioLink, ELIN@, DanBib og DADS som alle har en del egenskaber i tråd med en national databrønd i form af bred indholdsmæssig dækning, over tjenester med et snævrere indholdsmæssig og faglig fokus som f.eks. HighWire hvor vægten er lagt på not-for-profit materiale og et fagligt fokus på medicin, til planlagte eller igangværende projekter med varierende ambitionsniveau som eksempelvis konsolideringen af de Islandske elektroniske bibliotekstjenester og det tyske projekt til lokal lagring af backfiles.

De ovennævnte tiltag er blevet undersøgt af arbejdsgruppen.
Herudover har arbejdsgruppen undersøgt en række tjenester som i sig selv ikke kan betragtes som nære slægtninge til en databrønd men som dels kunne inspirere i det videre arbejde frem mod en dansk løsning, dels kunne tænkes at indgå som medspillere f.eks. i form af supplerende datakilder i landskabet omkring den nationale databrønd, eksempelvis Infomedia og OCLC tjenesterne FirstSearch og ArticleFirst.
De undersøgte tjenester omfatter således:

Databrøndsbeslægtede tjenester

- Ohio Link
- HighWire
- ELIN@/
- DanBib
- DADS

Udviklingsprojekter med databrøndslignende intentioner

- Det Islandske projekt: Konsolidering af de Islandske elektroniske bibliotekstjenester
- Et Tyske projekt til lokal lagring i Tyskland *(og de Tyske Nationallicenser)*

Tjenester med relevans som f.eks. supplerende datakilder til en national databrønd

- Infomedia
- FirstSearch
- ArticleFirst
- WorldCat
- NetLibrary
- Forsideservice
- E-bog.dk

Databrøndsbeslægtede tjenester

Koncept og organisation

Et fælles træk for de databrøndsbeslægtede tjenester er at tjenesterne, ud over et normaliseret datagrundlag, omfatter overbygninger mhp. at integrere flere serviceelementer med tjenestens grundlæggende adgangsfunktioner. Således indgår, til forskel fra konceptet for en databrønd, alerting, dokumentlevering, personalisering m.m. typisk i tjenesterne.

En sammenlignende analyse af enkelte tjenesters organisation, drift- og udviklings-omkostninger m.m. vil derfor kræve indsamling af flere detaljer om de enkelte tjenesters baggrund end indeværende forprojekt tillader.

Omkostninger

Oplysninger om omkostningsniveauet for de undersøgte tjenester er få, kortfattede og forskelligartede.

Tjenesterne er endvidere kvantitativt og kvalitativt yderst uensartede, eksempelvis betjener Ohio Link 87 uddannelsesinstitutioner med i alt 600.000 brugere med et indhold på 9,8 mio. artikler og 36.000 e-bøger medens DADS betjener 4 uddannelsesinstitutioner med tilsammen under 100.000 brugere og med et indhold på 100 mio. artikler og ingen e-bøger.

I det lys siger sig selv, at en brugbar omkostningsanalyse af så forskellige tjenester vil kræve en omfattende indsamling og grundig gennemgang af relevant baggrunds-materiale, også her altså en større indsats end indeværende forprojekt kan bære.

Udviklingsprojekter med databrøndslignende intentioner

Det Islandske projekt befinder sig i en overgangsfase hvor man dels overvejer et gennemgribende teknologiskift fra forældede biblioteksystemer til et integreret adgangssystem, dels en opgradering af datagrundlaget fra en traditionel bogcentreret fælleskatalog til et normaliseret datagrundlag omfattende "alle" materialer, herunder også e-tidsskrifter.

Det Islandske projekt er måske af særlig interesse set fra Dansk side takket være de særlig gunstige nationale licenser som Island har opnået med en del af de større internationale - tidsskriftforlag.

I øjeblikket er der kontakt mellem Det Islandske projekt og Det Kongelige Bibliotek og DTIC gennem et afklaringsprojekt mhp et evt. samarbejde omkring samsøgning og datagrundlag.

Det foreslås at man fra Dansk side fortsat følger det Islandske projekt, f.eks. ved en kontakt gennem DEFF.

Et Tysk projekt til lokal lagring i Tyskland.

Enkelte tiltag er undervejs på delstatsniveau, således er TIB Hannover er i gang med at etablere en lokal lagring af data. Første produkt vil blive en lokal lagring af arkivet fra American Physical

Society, købt som national licens. Dette vil være et pilotprojekt fra 1.1.2009. Lagringen vil blive udliciteret og på kontrakt.

De Tyske Nationallicenser

Med en bevilling på 21,5 mio. euro i hånden har Deutsche Forschungsgemeinschaft forskaffet adgang på nationalt niveau til elektroniske arkiver fra et antal større internationale forlag som Oxford University Press, Elsevier, Springer og Wiley-VCH såvel som henimod 30 A&I databaser. Der er i første omgang ikke tale om lokal lagring, hovedsagelig fordi teknikken hertil endnu ikke er tilstede. De opnåede licensaftaler ligger på den anden side tæt på en model for tilsvarende danske aftaler i forbindelse med Den Danske Databrønd.

Det foreslås at man fra Dansk side følger udviklingen vedr. De Tyske Nationallicenser , f.eks. ved en kontakt gennem DEFF.

Tjenester med relevans som f.eks. supplerende datakilder til en national databrønd

De nævnte tjenester ligger har i sig selv få egenskaber tilfælles med databrøndstanken. Deres potentiale som datakilder til en National Databrønd bør overvejes i forbindelse med Databrøndens etableringsfase.

Yderligere baggrundsmateriale på:
https://infoshare.dtv.dk/twiki/bin/view/DEFFDataWells/Artikler_erfaringer

6. BILAG B: SPØRGESKEMAUNDERSØGELSE AF BIBLIOTEKERNES BEHOV

Opsummering af undersøgelsen

Af den foretagne spørgeskemaundersøgelse fremgår det, at knap halvdelen af de biblioteker (6 ud af 14, der har besvaret spørgeskemaet), angiver at de allerede har et Integreret Søgesystem, primært PRIMO, - men også Summa/Infonet nævnes. For de øvrige biblioteker gælder det, at de udtrykker ønske om at få et sådan. Flere har planer på kort sigt, men alle er de (om end i varierende grad) usikre på konceptet, og alle er usikre på økonomien.

I spørgeskemaundersøgelsen blev bibliotekerne spurgt til deres forventninger til omkostningerne ved etablering af nationale databrønde. og Integreret Søgning. Der kan næppe udledes alt for håndfaste konklusioner af dette, da det drejede sig om et ubestemt formuleret spørgsmål, om 'hvor meget man er indstillet på at betale' for en ikke yderligere specificeret service. Men af de kun 8 biblioteker der har tilkendegivet hvor meget de vil være i stand til at betale, har 6 svaret 100.000 eller derunder, og 2 ligger i intervallet 3-500.000. Og ingen har svaret at de slet ikke vil være i stand til at betale noget som helst!

I kommentarerne til dette spørgsmål gøres der opmærksom på nogle centrale problemstillinger omkring økonomien, og dermed indirekte for strategien for hele databrøndsprojektet, som bibliotekerne opfatter som uafklarede på nuværende tidspunkt. Det drejer sig bl.a. om, hvorvidt der er tale om ekstraudgifter til metadata oven i udgifterne til selve ressourcerne, og om hvorvidt omkostningerne til at behandle/ reformatere både data og metadata, bevirker at det enkelte bibliotek ender med at kunne stille færre ressourcer til rådighed for slutbrugerne?

Den foretagne spørgeskemaundersøgelse giver ikke noget entydigt svar på, hvilke aftagerprofiler de forskellige biblioteker ønsker sig. Profilerne ser ud til at ramme de aftagerbehov som bibliotekerne forestiller sig pt. Og der er umiddelbart aftager-biblioteker til alle de skitserede profiler. Der kan iagttages en vis usikkerhed omkring hvad man egentligt vil have. Dette er ikke overraskende på nuværende tidspunkt, og det er selvfølgelig noget man skal være opmærksom på i processen fremover.

Forprojektet handler om databrønde for hhv. e-artikler og e-bøger. Men spørgeskemaet afslører, navnlig i kommentarerne, derudover en interesse for løsninger vedr. både artikler, tidsskrifter og e-bøger. Der udtrykkes interesse for en tidsskriftsliste/databrønd med tidsskrifter, måske baseret på en national konsolidering af lokale SFX-'knowledge-baser'.

Af spørgeskemaundersøgelsen fremgår det, at det meget naturligt især er de store forskningsbiblioteker, der udtrykker interesse for at lagre og indeksere data lokalt. Og de er alle interesserede i at komme til at arbejde med både fuld-tekst og berigede meta-data (med abstracts og emnedata). Og med en enkelt undtagelse tilkendegiver de samme biblioteker, at de gerne vil modtage rå, ubehandlede data fra databrønden – dog med enkelte kommentarer om, at rå-data i givet fald er en 'fall-back option', i tilfælde af, at databrønden ellers leverer behandlede data i en uhensigtsmæssig form.

Formål
I forbindelse med forprojekterne for en national databrønd for e-bøger og e-tidsskrifter har der været et ønske om, at få afprøvet interessen blandt bibliotekerne for de potentielle aftagerprofiler som er blevet skitseret i projektet. Formålet med denne spørgeskemaundersøgelse er, at indhente informationer fra bibliotekerne vedrørende deres ønsker til den nationale databrønds aftagerprofiler, økonomi, formater, protokoller m.m. Resultaterne har givet input til den arbejdsgruppe som ser på organisation og styring, og arbejdsgruppen som arbejder med formater og dataudveksling.

Metode
Arbejdsgruppen bag spørgeskemaet består af Lars Nondal (CBS – tovholder), Lise Mikkelsen (DEFF), Birgit Brejnebøl (CBS), Janne Vendt (DEFF), Lars Kofod-Jensen (NI, Frederiksberg Seminarium) og Anne Skov (KB).
Denne arbejdsgruppe skulle oprindeligt have lavet en spørgeskemaundersøgelse, der kun fokuserede på e-bøger. Af forskellige grunde bevægede vi os væk fra dette udgangspunkt, og spørgeskemaet er endt med at fokusere på den nationale databrønds overordnede arkitektur og koncept, i form af de nævnte aftagerprofiler. I arbejdsgruppen har man udarbejdet et spørgeskema, der er blevet testet af flere personer fra andre arbejdsgrupper i databrøndsprojektet. Denne test er udført for at sikre, at spørgeskemaet spørgeteknisk fungerer, og at der indhentes brugbar information til de andre grupper.

Spørgeskemaet er blevet udsendt til nedenstående deltagere:

Universitetsbibliotekerne (typisk bibliotekets IT-chef)
Professionshøjskoler (biblioteksschef)
Specialbiblioteker (5 repræsentanter i licensarbejdet)
I alt 21 institutioner

Spørgeskemaet indeholder en række lukkede spørgsmål med mulighed for sætte flere krydser, samt mulighed for at kommentere og uddybe.

Følgende biblioteker har besvaret spørgeskemaet:

University College Lillebælt	Det Kongelige Bibliotek, KUB, NB
UC Vest Biblioteket	Statsbiblioteket
University College Nordjylland, Biblioteket	Syddansk Universitetsbibliotek
Kunstakademiets Arkitektskoles Bibliotek	Aalborg Universitetsbibliotek
Medicinsk Bibliotek, Aalborg Sygehus	Danmarks Tekniske Informationscenter
Forsvarsakademiet	Roskilde Universitetsbibliotek
UC Syd Bibliotekerne	Det Administrative Bibliotek

Universitetsbibliotek	7
Professionshøjskole	4
Specialbibliotek	3
Andet	0
I alt besvarelser	14

De tre aftagerprofiler
I undersøgelsen er defineret 3 forskellige 'aftagerprofiler'. Disse profiler er forsøgt designet således, at de både tager højde for behovet hos store universitetsbiblioteker med store tekniske og økonomiske ressourcer, og behovet hos mindre og mellemstore biblioteker, som forventes at være mere interesserede i centrale (mere eller mindre færdigpakkede og standardiserede) løsninger.

I aftagerprofil 1 foregår hele processen fra lagring af data, indeksering, søgemaskine og til at stille en slutbrugergrænseflade til rådighed hos det lokale bibliotek. Man har så haft mulighed for at afkrydse om man ønsker denne profil for 1) tidsskriftsartikler 2) tidsskrifter (titler) 3) e-bøger. Denne adskillelse mellem de forskellige materialetyper er bl.a. lavet fordi, der er meget store forskelle i mængden af data. For tidsskriftsartikler taler vi om 100.000-vis af titler, mens antallet af tilgængelige e-bøger nok er i stigning, men stadig på et lavere niveau.

Og for hhv. artikler og e-bøgers vedkommende bliver man bedt om at tage stilling til, hvor store mængder data man i givet fald vil behandle lokalt – dvs. om man ønsker at download/importere f.eks. e-bogsposter (bibliografiske metadata) til sit lokale system eller om man også ønsker den fulde tekst.

I profil 2 har man en central lagring og en central søgemaskine, mens søgning og grænseflade styres lokalt. I den tredje profil foregår alt, dvs. lagring, søgning og grænseflade centralt. Og ligeledes for begge disse profiler har det været muligt at afkrydse om man ønsker denne løsning el. profil for hhv. 1) tidsskriftsartikler 2) tidsskrifter (titler) 3) e-bøger.

Oversigt over hovedprofiler

	Profil 1 lokal lagring, søgning og visning af data			**Profil 2** lokal lagring, central søgning, visning af data i eget system			**Profil 3** central lagring, søgning samt central webgrænseflade		
Emner som kan overvejes ifm valg af hovedprofil:	**A** artikler	**B** tidsskrifter	**C** ebøger	**A** artikler	**B** tidsskrifter	**C** ebøger	**A** artikler	**B** tidsskrifter	**C** ebøger
Lagring af data	lokalt	lokalt	lokalt	centralt	centralt	centralt	centralt	centralt	centralt
Søgemaskine	lokalt	lokalt	lokalt	centralt	centralt	central	centralt	centralt	centralt
Brugergrænseflade	lokalt	lokalt	lokalt	lokalt	lokalt	lokalt	centralt	centralt	centralt
Krav til bibliotekets tekniske ressourcer	Meget store			Mellem			Små		
Fleksibilitet	Meget høj			Mellem			Lille, standard løsning		
Krav til lokale systemer	Meget store			Mellem, skal kunne anvende webservice eller Z39.50			Små		
Søgning	Store krav, da søgning foregår lokalt			Ingen lokale krav, da søgning foregår ifm databrønd			Ingen lokale krav, da søgning foregår ifm databrønd		

Aftagerprofilerne er mere detaljeret beskrevet på wiki'en.
Se: https://infoshare.dtv.dk/twiki/bin/view/DEFFDataWells/AftagerProfil

Resultater
Besvarelserne (med kommentarer) findes i filen undersøgelse_samlet.doc, der ligeledes ligger på wiki'en.

Visse af resultaterne fra spørgeskemaet har givet forskellige input til hhv. den arbejdsgruppe som

ser på organisation og styring, og til arbejdsgruppen som arbejder med formater og dataudveksling. Derfor gives der her kun en præsentation af de tilkendegivelser, der kom ang. nogle af de mere overordnede spørgsmål omkring databrøndenes profiler, services og økonomi.

For information om bibliotekernes tilkendegivelser ang. mere specifikke, tekniske forhold omkring formater, dataudveksling, protokoller og de systemer som data fra databrønden i givet fald skal integreres med – se den samlede rapport fra SurveyXact (filen undersøgelse_samlet.doc)

Der blev oprindeligt udsendt mails til 21 institutioner med opfordring til at deltage i undersøgelsen. Og på trods af flere rykkere, nåede vi ikke op på mere end 14 besvarelser. Selve svarprocenten er hæderlig, men under alle omstændigheder er materialet så begrænset, at der ikke er grundlag for alt for skråsikre konklusioner.

Helt overordnet situationsbillede:

Knap halvdelen af de deltagende biblioteker har allerede et Integreret Søgesystem, primært PRIMO – men også Summa/Infonet nævnes. For den øvrige halvdel gælder det, at de udtrykker at de gerne vil have et sådan. Flere har planer på kort sigt, men alle er de (i varierende grad) usikre på konceptet, og alle er usikre på økonomien.

I denne rapport har vi valgt at fokusere på følgende 3 hovedspørgsmål:

A) Økonomien
B) Hvilke aftagerprofiler foretrækkes?
C) Ved lokal lagring og databehandling: hvilke former for data ønskes stillet til rådighed fra databrønden?

Tema A) Økonomien
For at illustrere spørgsmålet om økonomiens betydning, vises nedenfor både svarfordelingen på det meget enkle (og totalt ubestemte/ ukvalificerede) spørgsmål om, hvor meget man er indstillet på at betale – og de kommentarer, der indkom i forbindelse med dette spørgsmål.

Af de 8 biblioteker der har tilkendegivet hvor meget de vil være i stand til at betale, har 6 svaret 100.000 eller derunder, og 2 ligger i intervallet 3-500.000. Ingen har svaret at de slet ikke vil være i stand til at betale noget som helst!

I kommentarerne er med gult markeret nogle udsagn, der peger på nogle centrale problemstillinger/dilemmaer omkring økonomien – og dermed strategien for udviklingen af databrøndenes indhold. Er der tale om ekstraudgifter oven i udgifterne til selve ressourcerne, og vil omkostningerne til at forbedre søgemulighederne bevirke at det enkelte bibliotek få stille færre ressourcer til rådighed?

Spm: Hvor meget vil I betale for at kunne trække på en national databrønd? Man må forvente, at adgangen til en databrønd bliver betalingsbaseret. En betaling kunne for eksempel afhænge af det enkelte biblioteks ønsker til funktionalitet, datamængde og bearbejdning.

	Svar:
10-50.000,-	3
51-100.000,-	3
101-300.000,-	0
301-500.000,-	2
over 500.000,-	0
Vil ikke kunne betale	0
Kommentar	7
I alt	11

Kommentarerne:

Jeg forstår ikke spørgsmålet, er betalingen for oplægning og administration af basen eller dækker det også adgangen til ressourcerne
Jeg kan p.t. ikke sige noget om vores økonomiske muligheder
For vores eget vedkommende - og for de små og mindre biblioteker generelt - er prisen en meget vigtig faktor. Her er det i vis grad bekymrende, at prisen kunne komme til at afhænge af bearbejdningsgrad, når netop de små biblioteker kan forventes at have behov for relativt mere bearbejde data end 'de store'. Det eneste sted man så vil kunne kompensere er i datamængde, og så er man (groft sagt) lige vidt. Så vil der selvfølgelig (forhåbentlig) blive mulighed for at lave konsortier eller lign., men.
Direktionsanliggende
Hvis vi får det hele.....
Det kan vi ikke sige noget om på nuværende tidspunkt. Det er imidlertid vigtigt, at udgiften til brug af databrønden ikke får den konsekvens, at SDUB må opsige eksisterende vigtige abonnementer. I en sådan situation foretrækker vi adgang til ressourcerne frem for databrønden.
Størrelsen af betalingen vil selvfølgelig afhængig meget af hvilke ydelser og funktioner, der tilbydes ovenstående er en indikation af max. betaling, hvis alle data, ønsker og krav til en national databrønd er opfyldt
Betalingen må basere sig på en fordeling af omkostningerne til de services man abonnerer på - efter DEFF's sponsorering (50%)

Tema B.) Hvilken aftagerprofil foretrækkes?

Spørgeskemaet var ikke opbygget, så det enkelte bibliotek definitivt skulle vælge en af de 3 aftagerprofiler – og dermed fravælge de 2 andre. Ud fra den antagelse, at det kunne være relevant for et specifikt bibliotek at gå efter forskellige profiler (hhv. lokale el. centrale løsninger), blandt andet afhængigt af materialetyperne.

For alligevel at forsøge, at få et overblik over hvilke profiler der typisk ønskes, er i nedenstående tabel sammenfattet hvordan de indledende ja/nej-svar til hver profil fordeler sig.

	Artikler		Tidsskrifter		E-bøger	
Profil 1: lokal lagring og lokal søgning	J	N	J	N	J	N
	4	4	4	4	5	4
Profil 2: central lagring og lokal søgning	J	N	J	N	J	N
	5	4	6	3	5	3
Profil 3: central lagring og central søgning	J	N	J	N	J	N
	3	6	3	6	3	6

Det er næppe muligt at udlede entydige og klare svar på baggrund af denne fordeling. Men nogle forsigtige konklusioner kunne være:

1) Der er umiddelbart aftagere til alle de skitserede profiler. Profilerne ser ud til at ramme de aftagerbehov som bibliotekerne forestiller sig pt.

2) Flere kan have "helgarderet", dvs. sagt JA til mere en 1 profil. Det kan sikkert tolkes som en visusikkerhed omkring hvad man egentligt vil have. Dette er ikke overraskende på nuværende tidspunkt, og det er selvfølgelig noget man skal være opmærksom på i processen fremover.

3) Forprojektet handler om databrønde for hhv. e-artikler og e-bøger. Men spørgeskemaet afslører interesse for løsninger vedr. både artikler, tidsskrifter og e-bøger. Der udtrykkes interesse for en tidsskriftsliste/databrønd med tidsskrifter, måske baseret på en national konsolidering af lokale SFX 'knowledge-baser'. Dette rejser spørgsmålet om hvor langt kan man nå med tidsskrifter i en databrønd for e-tidsskriftsartikler? Og er det en ny databrønd?

Tema C:) I tilfælde af lokal lagring (og indeksering) (d.v.s. Profil 1) – hvilke typer data er man interesseret i at modtage fra databrønden?

	Artikler	Tidsskrifter	E-bøger
Rå metadata/bibliografiske poster uden efterbehandling (den oprindelige kildes format)	4	4	4
Metadata med efterbehandling (i databrøndens format)	4	3	4
De-dublerede metadata	2	3	3
Harmoniserede (f.eks. ensrettede navneformer) de-dublerede metadata	3	2	2
Berigede metadata (f.eks. emnedata og abstract)	5	5	5
Fuldtekst	4		5
Andet	1	0	0
Kommentar	4	2	1
I alt	5	5	5

Forsigtige konklusioner: alle der er interesserede i at lagre og behandle data lokalt, er interesserede i: fuld-tekst og berigede meta-data (med abstracts og emnedata), næsten lige så mange tilkendegiver, at de gerne vil modtage de rå, ubehandlede data – dog med enkelte kommentarer om, at rå-data er en 'fall-back option', i tilfælde af, at databrønden leverer behandlede data i en uhensigtsmæssig form.

7. BILAG C: BUSINESS CASE

Titel :	National databrønd for digitale materialer
Version:	0.1
Dato:	6. September 2008
Udarbejdet af:	Bo Öhrström & Lise Mikkelsen, Styrelsen for Bibliotek og Medier

Løsningsbeskrivelse

1.1. Kort beskrivelse af løsningen

National databrønd for elektroniske materialer til anvendelse i forbindelse med blandt andet integreret søgning

1.2 Formål

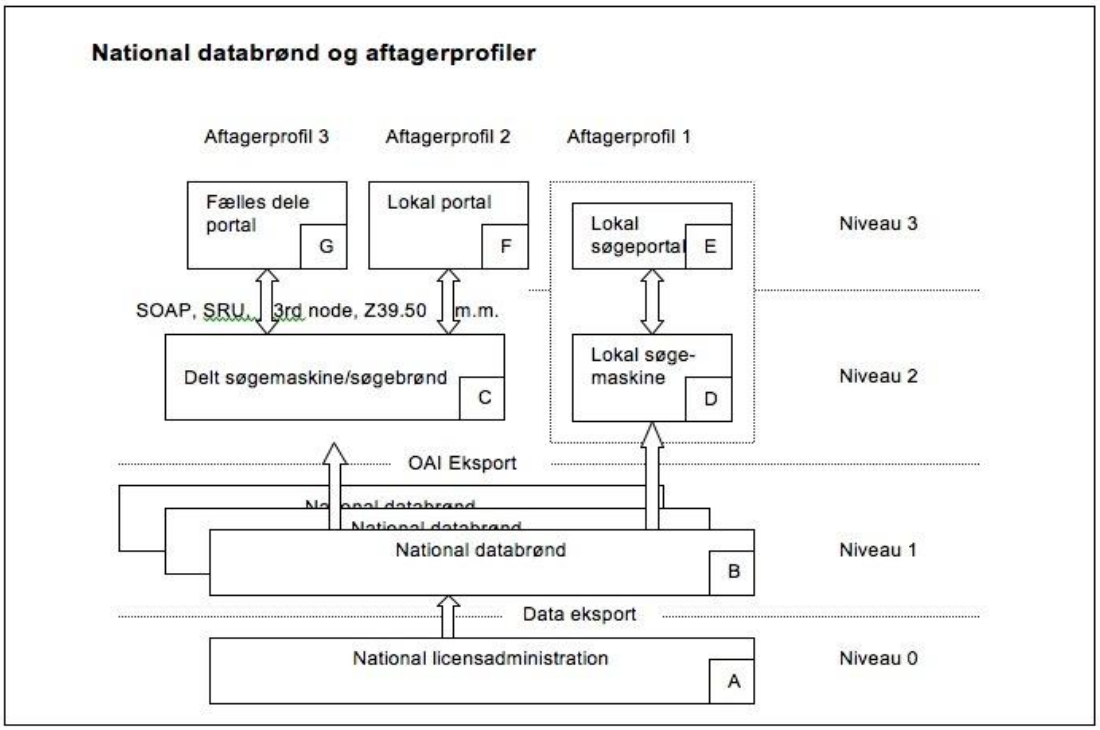
Skabe det nødvendige fundament og samarbejde for at kunne lave nationale databrønde
Skabe et datagrundlag og IT- infrastruktur som kan give væsentlig bedre anvendelse af bibliotekernes licenser
Forbedre servicen overfor bibliotekernes slutbrugere
Forbedre omfanget af metadata om digitale materialer i bibliotekernes søgesystemer
Udbrede bibliotekernes muligheder for at anvende metadata om digitale materialer i søge- og formidlingsmæssig sammenhæng
Forbedre omfanget af fuldtekstmateriale om elektroniske materialer til indeksering, søgning og i formidling af søgeresultater
Opbygge effektiv og fælles IT-infrastruktur
Reducere driftsomkostninger
Øge samarbejdet omkring licenser og udbrede omfanget af fælles nationale licenser i det omfang der er interesse

1.3 Forretningsmæssig baggrund

Bibliotekerne har stor fokus på integrated search. Integrated search ses som en mulig løsning på brugernes ønske om ét søgefelt med adgang til alle materialer, samt ønsket om en bedre formidling af søgeresultaterne. Integreret søgning kræver i første omgang adgang til metadata, ToC, fuldtekst m.m., hvilket blandt andet skal sikres gennem licenskontrakterne. DEFF varetager ca. 200 nationale licensaftaler, og det er ønsket fra bibliotekernes side, at der i disse kontrakter blandt andet sikres adgang til metadata, indeksering af fuldtekst m.m. Forventningen er, at bibliotekernes samlede krav overfor forlagene via DEFF giver bedre muligheder for at få adgang til disse data og bedre anvendelsesmuligheder. Den datamæssige og tekniske håndtering af disse mange data er voldsomt ressourcekrævende, hvorfor en national løsning er nødvendig for at løse opgaven.

IT-mæssig løsningsbeskrivelse

En national databrønd indgår som 1. niveau i en trelags-arkitektur, hvor 1. niveau er databrønd. På det andet niveau udføres søgning og det tredje niveau er grænseflade/portal. I nedenstående illustration er indført et niveau 0, som er den nationale licensadministration samt en national SFX løsning, som endnu ikke er etableret. Nøglekomponenterne i løsningen er en eller flere nationale databrønde, lokale eller fælles søgemaskiner samt fælles eller lokale portaler. De nationale databrønde vil være nye komponenter, mens behovet for søgemaskiner og portaler formodes at kunne dækkes af Summa eller Primo.



1.5 Beskrivelse af identificerede delprojekter

Undersøgelse af mulighederne for at lave national løsning omkring SFX og Verde	Mulige deltagere: Programområdet for Informationsforsyning, LUB/LUF, DEFF licenssekretariatet
Undersøgelse af interessen og økonomien blandt biblioteker for at lave fælles konsortier til niveau 2 og 3, dvs. Summa eller Primo løsninger. Mulige konsortier: Universitetsbiblioteker, Specialbiblioteker, Professionshøjskoler m.m.	Mulige deltagere: Programområdet for Informationsforsyning, Programområdet Nye Institutioner, DEFF sekretariatet
Undersøgelse af mulighederne for at lave dokumentlevering og/eller Pay Per View	Mulige deltagere: DEFF licenskontoret, Programområdet Informationsforsyning, Programområdet Nye Institutioner
Yderligere analyse af økonomi i opstartsfase og drift	DEFF sekretariatet

1.6 Økonomi

Indsamling og håndtering af data i en national databrønd er forbundet med mange personale ressourcer. Økonomisk set er der derfor besparelser at hente ved at gøre dette i fællesskab, således at opgaverne kan løses en gang for alle. Som det kan ses af nedenstående overslag, så skal der ikke være mere end 2 lokale løsninger, førend en fælles løsning er billigere. Mens forholdet mellem lokal og fælles løsning anses for velestimeret, tages der forbehold for selve udgiftsniveauet. Af gode grunde er dette ind til videre et gæt, hvad angår behovet for medarbejdere og udgifterne til løn, overhead og teknik.		
Fælles løsning:	4 medarbejdere Overhead Teknik I alt	2.000.000,- 500.000,- 500.000,- 3.0000.000,-
Per lokal løsning	2 medarbejdere Overhead Teknik I alt	1.000.000,- 250.000,- 300.000,- 1.550.000,-

8. BILAG D: DATASTRØMME, LEVERANDØRER, LICENSER - ARTIKLER

Hvad skal med i Databrønden:
Det er vigtigt, at indholdet i databrønden kommer til at omfatte alle DEFF-licenserne. Det anbefales også, at databrønden suppleres med metadata til tidsskrifter udenfor DEFF-pakkerne. Dette skal være til tidsskrifter som de enkelte biblioteker selv har anskaffet, enten som en pakke eller som enkelttidsskrifter.

Udover ovennævnte tidsskrifter, anbefales det at også at Open archives fra starten bliver en integreret del af databrønden. Det anbefales endvidere, at lægge sig tæt op af de beslutninger, der kommer fra gruppen omkring bøger for at forhindre, at noget falder mellem stolene. Det er svært at afgrænse bogserier/konferencer. Bedste måde at afgrænse på er, at skelne mellem ISBN og ISSN. Det skal dog påpeges, at flere DEFF-licenser indeholder både bøger og tidsskrifter, ligesom der kan optræde både ISBN og ISSN på samme publikation.

Det anbefales, at Bibliografier som er indgangen til bl.a. papirtidsskrifterne, med tiden bliver inkluderet, men at disse bliver prioriteret lavere end fuldteksttidsskrifter.

Encyclopædier/referenceværker bør afklares i samarbejde med arbejdsgruppen for bøger, men bør også blive en del af databrønden.

Avisdatabaser kunne være godt at få ind i en databrønd. Men erfaringer med både forhandlinger og linking til aviser gør, at dette skal afvente alle øvrige produkter.

Metadata til tidsskrifter udenfor DEFF-licenserne, kan forhandles enten direkte fra de enkelte biblioteker, via den centrale forhandling eller ved at data anskaffes v.h.a. TOC-databaser fra Swets, Ebsco eller lignende leverandører. Sidstnævnte løsning vil overflødiggøre en lang række henvendelser til mindre forlag og societies, samt minimere evt. dobbelt-henvendelser fra bibliotekerne.

Model/procedurer/skabelon for dataleverancer.
Erfaringer viser, at data (næsten altid) vil blive leveret i det format som forlaget har mulighed for. Men et ønske er, at metadata leveres i XML, og at fuldtekst leveres i PDF eller XML, hvis alt andet er udelukket så HTML. Der bør skaffes så mange metadata fra den enkelte leverandør som muligt, herunder også fra backfiles, article first m.v.

Levering af data kan ske på op til 4 niveauer:

- TOC
- TOC + abstract
- TOC + abstract + citationer
- TOC + abstract + citationer + fuldtekst

Da erfaringer viser, at det kan være problematisk at forpligte udgiveren på et bestemt metadataformat, skal det indgå i kravspecifikationen til operatøren af databrønden, at der skal kunne håndteres flere forskellige formater.

Forhandling.
Forhandlinger bør være centralt koordineret, da det skal være klart for leverandørerne hvem de forhandler med, og i hvilken sammenhæng data vil blive brugt.

Der er udarbejdet en beskrivelse, der indeholder de nødvendige informationer, dels om projektet, integrated search, databrønd samt de klausuler som der bør indgå i kontrakten med forlaget.

Det er vigtigt, at opdateringer skal være ajour og up to date. Dette gælder ikke kun data fra forlagene, men også i den lokale lagring skal der være aktualitet. Derfor anbefales det, at krav omkring aktualitet indgår i kravet til forlaget, og at det også indgår i en kravspecifikation til operatør af databrønden. I beskrivelsen til forlagene lægges der derfor op til daglige leveringer af data.

Skal der være adgang til fuldtekst?
Der er et ønske om, at der bliver adgang til fuldtekstdata for at kunne indeksere, søge i fuldtekst, datamine m.v.

Hvis lokal lagret fuldtekst bliver brugt til levering af fuldtekst, så har konsortiet ansvaret for at der kun er adgang for autoriserede brugere, og også for i tvivls tilfælde at kunne bevise det overfor forlagene. Forhandlinger indtil nu viser, at forlagene er bekymrede ved at lade fuldteksten lagre lokalt. De mister kontrollen, og har ikke mulighed for at kontrollere brugen m.v. Det må formodes, at lokal lagring af data mere end et sted i Danmark vil give problemer med forlagene. Der bør derfor, så vidt det er muligt leveres fuldtekstdata, men det må ikke have hverken afgørende eller opsættende virkning. Evt. kunne det aftales med forlagene, at den leverede fuldtekst kun bruges til indeksering m.v., og at levering af fuldtekst til slutbrugerne sker v.h.a. link til forlagets server, eller fra databrøndsoperatøren.

9. BILAG E: DATASTRØMME, LEVERANDØRER, LICENSER – E-BØGER

Datastrømme

Til overvejelserne omkring datastrømme i en databrønd for e-bøger, er der taget udgangspunkt i at e-bøger nu såvel som fremadrettet repræsenteres ved katalogposter i danMARC2 formatet, samt at der i disse poster som hovedregel, linkes til fuldtekst e-bogen hos en eksterne udbyder.

I relation til e-bøger og datastrømme, rettes fokus på datastandarder, datatyper og ikke mindst det enkelte biblioteks tilgang til søgning og fuldtekst.

Datastandarder

Der er behov for udarbejdelse af fælles nationale krav til datastandarder i en databrønd for e-bøger. I dag anvender hver enkelt bibliotek ressourcer på konvertering af e-bogsposter til egen katalog og eget katalogformat - ressourcer der bør kunne rationaliseres med etablering af en fælles databrønd.

Efterbehandling af data afhænger af det enkelte biblioteks behov, hvor nogle ønsker rå data og andre bearbejdede. Det må være målet med databrønden, at efterbehandling til egen katalog er en mulighed men ikke en nødvendighed.

Datatyper

Databrønden skal som minimum kunne levere e-bogsdata i danMARC2 format. Alt efter aftagerbehov kan andre formater også komme på tale, særligt i forhold til supplerende data som TOC, uddrag, fuldtekst og supplerede medietyper.

Konvertering af e-bogsdata fra andre formater, herunder MARC21, XML og Dublin Core, skal håndteres af brøndens operatør.

Central adgang til e-bogskatalogposter i et standardiseret format, vil sikre en vis ensartethed og at der vil skulle anvendes langt færre ressourcer på behandling af e-bogsdata.

Fysisk placering af supplerende datatyper, stiller krav til rammerne for leverandøraftalerne, men afhænger også af valg af brøndprofil og tilgangen til data. Ved høstning af data til lokal katalog kan det diskuteres, om det er hensigtsmæssigt at indhøste supplerende data, frem for at tilgå disse via eksterne links i e-bogsposterne.

Tilgang

Udgangspunktet er, at alle DEFF biblioteker har adgang til alle data i databrønden for e-bøger. Hvorvidt det enkelte bibliotek har adgang til den konkrete pakke, titel eller fuldtekst, håndteres ikke i databrønden, men skal som i dag administreres af bibliotekets egen rettigheds- og adgangsstyring.

Alt afhængig af valg af brøndprofil, vil e-bøgerne i databrønden enten skulle søges separat i forhold til bibliotekernes egne kataloger, eller data høstes og indekseres i bibliotekernes lokale søgesystemer. Dette gælder såvel traditionelle biblioteks OPACs, samt for integrerede søgesystemer.

Behov for integreret søgning i e-bøger, flettet med søgning i lokale katalogdata, kræver høstning af data og lokal indeksering. Dette stiller særlige krav til rammerne for de licensaftaler der indgås med dataleverandørerne, da høstning og lokal indeksering i realiteten er en videredistribution af data.

Leverandører og licenser

DEFF indtager en naturlig rolle som licensforhandler på e-bogspakker i en national databrønd. Det forudsættes som minimum, at brønden rummer de pakker og titler DEFF i forvejen forhandler adgang til. Der er brug for en model, der tager højde for at der fremover vil være stigende behov for køb af enkelttitler frem for hele pakker. Forhandling om adgang til enkelttitler og delmængder af pakker, kunne være en DEFF opgave.

Herudover tilstræbes det, at brønden samlet set indeholder foreningsmængden af de pakker de tilsluttede biblioteker abonnerer på. Såfremt det enkelte bibliotek herudover ønsker specifikke enkelttitler tilføjet i brønden, må dette ske ved eget initiativ og arbejde.

Såfremt DEFF skal varetage opgaven, må der være tale om begrænsede antal og titler af en vis "brøndværdi" for de involverede biblioteker. Der bør i den forbindelse etableres fælles politikker for hvad det enkelte bibliotek kan forvente og forlange at berige databrønden med.

10. BILAG F: METADATA, EMNEDATA OG VÆRKKLYNGER

Metadata, emnedata, værkklynger

Databehov

Data og metadata i brønden skal understøtte følgende behov:

- Høstning : Udbyder- og pakkeinformationer, identifikatorer
- Deduplikering: Identifikatorer, titel- og forfatteroplysninger, ...
- Søgning: Artikel- og tidsskriftsoplysninger, emneord, klassifikationer, abstracts, fuldtekst...
- Fremvisning: Artikel- og tidsskriftsoplysninger, abstract
- Levering: (oplysninger til) openurl, url, fuldtekst, …

Mht. deduplikering og værkklynger kan det konstateres at ønskerne vil variere fra anvendelse til anvendelse, så frem for at forsøge at tilvejebringe en fælles definition anbefales det at man sikrer at de for deduplikering og værkklyngning fornødne oplysninger forefindes i metadata. Ligeledes skønnes det urealistisk inden for rammerne af det aktuelle databrøndsprojekt at definere et fælles leverandøruafhængigt emnesystem, om end interessante problematikker blev vendt under analysen. F.eks. kunne informationer fra *Bibliometrisk Forskningsindikator* tænkes anvendt til emnetildeling på tidsskriftniveau.

Lagrings- og eksponeringsformat

Der har gennem projektarbejdet vist sig to "skoler" m.h.t. krav og ønsker til lagrings- og eksponeringsformater. Den ene ønsker et centralt vedligeholdt, normaliseret metadata-superformat, som på en enkel måde kan høstes ind i eget system og/eller tilgås i en søge-overbygning på brønden(e). Den anden "skole" foretrækker data i en form så tæt de enkelte leverandørers som overhovedet muligt og vil selv stå for ekstraktion af de fornødne metadata. Et kompromis vil være at et superformat suppleres med data i "rå" form. Et minimum af forarbejdede metadata vil under alle omstændigheder være nødvendigt af hensyn til høstning og vedligehold af brønden. Der er ikke her taget stilling til hvorvidt aftagerbehovene m.h.t. oparbejdede metadata skal have indflydelse på betalingsmodellen.

Krav/ønsker til det interne format

For så vidt angår et superformat, skønnes dataelementerne i det nuværende DADS-format at dække ovennævnte/kendte behov. Eksponeringsformatet bør dog være XML, og standardiserede metadataskemaer som f.eks. MODS vil formentlig også være anvendelige, men et forsøg på en crosswalk har afsløret at det næppe kan benyttes uden et vist tab af specificitet i forhold til et dedikeret dataskema. De originale data fra den enkelte udbyder medsendes i så rå form som muligt indlejret i et CDATA-element, således at aftagere får mulighed for at overstyre eller berige metadata i forbindelse med høstning ind i eget system.

De indhøstede data vil have en meget svingende kvalitet, både hvad angår struktur, omfang og datakvalitet. Dette stiller særlige krav til det interne lagringsformat, da det på den ene side er ønskværdigt at man ikke "forfladiger" data af høj kvalitet, mens det på den anden side er

nødvendigt at kunne søge i sammenlignelige data på tværs af dataleverandører. Et eksempel er personnavne, hvor man ofte kun benytter forbogstav og efternavn. Ønskerne til metadata kan opsummeres således:

- Intet eller minimalt informationstab ved mapning af de fra leverandøren modtagne metadata
- Indlejring (eller parallel opbevaring) af originaldata
- Det bør indikeres vha. xml-attributter el., når givne oplysninger er beriget ind fra anden kilde.

Tilgængeliggørelse og adgangsstyring

Artikeldata i brønden tilgængeliggøres vha. OAI-PMH. I forbindelse med initiel load kan der suppleres med dataoverførsler på f.eks. usb-diske.

Da det må forventes at adgangen til visse metadataleverancer vil være begrænset til licenserede aftagere og forskellige licensaftaler vil give brøndens aftagere forskellig adkomst til fuldtekster vil det være nødvendigt med et element af adgangskontrol. Det skal sikres, at den enkelte aftager opnår adgang til netop de metadata og det fuldtekstmateriale, som denne institution har adkomst til. Dette kræver som minimum, at operatøren af databrønden til stadighed har adgang til opdateret og præcis information om aftagernes metadataleveranceaftaler og fuldtekstlicenser. Dette kan kræve etablering af indrapporteringssystem (webbaseret, mailbaseret, el.lign.), hvormed aftagerne løbende effektivt og præcist kan orientere operatøren om ændringer i aftale- og licensforhold. Med adgang til disse oplysninger, kan OAI høstadgangen til metadata rettighedsstyres langs to dimensioner: Per aftager og per leverandør. Dette vil sikre, at aftageren af brønddata opnår adgang til netop det metadata, som dennes metadataleveranceaftaler giver denne adkomst til. Den enkelte aftager kan således hjemtage data per leverandør (eventuelt suppleret med en "all suppliers" foreningsmængde) i sikker forvisning om kun at få udleveret metadata, som svarer til de indrapporterede aftaleforhold.

Udleveringen af fuldtekster - både i forbindelse med eventuel hjemhøstning og udlevering til slutbrugere - vil på tilsvarende vis skulle adgangsstyres langs samme to dimensioner, baseret på indrapporterede licensforhold og på identifikation af aftageren. En sådan løsning fritager dog ikke de enkelte institutioner for selv i fornødent omfang at skulle adgangsstyre, såfremt der er differentieret adgang for slutbrugerne.

Tidsskrift overblik

En strøm af artikelmetadata som den, der vil opstå i en databrønd, giver mulighed for at opbygge et overblik over tidsskriftsmetadata (hierarkisk ordnet data over ISSN, år, volume, issue).
Dette overblik har oplagte interne og administrative anvendelser for operatøren, men bør tillige gøres tilgængeligt for andre - som minimum aftagerne.
At tilbyde aftager-specifikke tidsskrift overblik (baseret på aftagerens metadataaftaler og/eller fuldtekstlicenser) adderer kompleksitet til operatørens opgave, og vil derfor være fordyrende.

11. BILAG G: KRAV TIL LICENSER OG FORLAGENES DATALEVERANCER

Som udgangspunkt stilles nedenstående krav/spørgsmål til ethvert forlag (”Publisher”) om brugen af forlagets dataleverancer. ”Institution” er en institution med licens til forlagets produkt, og ”Authorized Users” er institutionens brugere, der ifølge licenskontrakten må anvende produktet til forskning, studier mv. ”Third party” er en kommerciel driftsleverandør.

Who may store delivered data from Publisher?	**Metadata**	**Full text**
Third party under agreement with DEFF or	x	x
One named Institution	x	x

Who may access stored delivered data from Publisher?	**Metadata**	**Full text**
Institutions’ and their Authorized Users	x	x
All libraries in higher and further education in Denmark and their users	x	-

Who may index data?	**Metadata**	**Full text**
Third party under agreement with DEFF	x	x
Institutions	x	x

Who may store indexed data?	**Metadata**	**Full text**
Third party under agreement with DEFF	x	x
Institutions	x	x

Who may access indexed data?	**Metadata**	**Full text**
Institutions’ and their Authorized Users	x	x
All libraries in higher and further education in Denmark and their users	x	-

Gennem disse krav søges maksimale, realistiske dataleverancer opnået til gavn for den størst mulige, realistiske danske kundeskare.

Hvis der fra folkebiblioteksside viser sig interesse for at gå ind i databrøndskonsortiet, vil det være relevant at udvide de ovenstående brugerkategorier med en eller flere.

12. BILAG H: FORSKNINGSBIBLIOTEKERNES CHEFKOLLEGIUMS HENVENDELSE

FC Forskningsbibliotekernes Chefkollegium

Formandskabet

J.nr.: FC-2007-007

DEFF
Biblioteksstyrelsen
H.C. Andersens Boulevard 2
1553 København V

Att. Formand Kim Østrup

Aalborg, den 29. maj 2007

Kære Kim

FC har den 21. maj diskuteret DEFF's kommende strategi på baggrund af notatet "Sekretariatets skitseplaner og rammer for programområder" (DEFF 07/201a v.2). Vi skal hermed tilkendegive, hvordan vi gerne ser den strategiske indsats i de kommende 2 år.

FC anbefaler at man klart fokuserer indsatsen omkring udviklingen af **Integrated Search and Delivery.** Dette indsatsområde, der er helt centralt for såvel DEFF som for fag- og forskningsbibliotekerne, omfatter følgende elementer:

- Integrerede søgeportaler (søgemaskiner og brugergrænseflader)
- Fælles datalicenser (forhandling og administration)
- Fælles databrønd (import, drift og deling af nationalt datasæt)
- Fælles dokumentleverancetjenester (digitalisering on demand)
- Fælles serviceorienteret arkitektur til sammenbinding af områdets elementer

Der er således bidrag fra alle Programområder og et tilsvarende behov for fokusering af disses arbejde. Vi mener følgelig, at området skal have højeste prioritet i såvel Styregruppen som i Programområderne, og at der bør etablere en effektiv tværgående styring, der kan sikre sammenhængende fremdrift og resultater indenfor den tidshorisont, den nuværende Styregruppe arbejder under.

FC er opmærksom på, at den tidlige udvikling på området **Integrated Search and Delivery** i nogen grad har været præget af parallelle indsatser. Men FC er i forhold til den videre udvikling af området helt afklaret. På trods af at der på søgeportal-fronten har været forfulgt tre forskellige udviklingsveje indtil nu, er det FC's klare holdning, at man fremover vil søge et tæt samspil i realiseringen af en samlet løsning i DEFF-regi.

Denne konsoliderede løsningsmodel er allerede ønsket og delvis realiseret i forbindelse med det nationale arbejde med datalicenser og udviklingen frem mod en fælles national databrønd. Initiativet om at søge midler hertil fra infrastrukturpuljen hilses derfor særdeles velkomment. Skulle det mislykkes, må der gennem DEFF's primære fokus på dette centrale indsatsområde være mulighed for at anvende DEFF-midler til databrøndens etablering og drift.

FC's vurdering er også, at et sats på Integrated Search and Delivery vil komme alle institutioner og brugergrupper i DEFF-samarbejdet til gode. Uanset hvem det enkelte biblioteks brugere er, vil de gennem den brugervenlige og integrerede tilgang de mange informationsressourcer få et meget bedre udbytte af søgning end i dag.

FC's institutioner er klar til at bidrage på afgørende vis til et fremadrettet DEFF-samarbejde på området Integrated Search and Delivery til glæde for alle DEFF's brugeregrupper. Vi mener følgelig, at DEFF klart bør prioritere dette område de næste 2 år.

På FC-mødet var der i øvrigt enighed om, at der fortsat bør sættes en mindre pulje af som seed money til særlige udviklingsprojekter, som måtte opstå undervejs udenfor det strategiske indsatsområde, ligesom konsolidering af bibliotekerne under de nye professionshøjskoler er et vigtigt særligt område.

Med venlig hilsen

Niels-Henrik Gylstorff

13. BILAG I: REFERAT AF SIDSTE MØDE I "FÆLLES DATA TIL INTEGRATED SEARCH"

Mødetid: 30.08.07 kl. 10:00 – 12:30
Mødested: H. C. Andersens Boulevard 2
Deltagere: Carsten Andersen, Hans Meulengracht, Mogens Sandfær, Bo Öhrström, Jakob Heide Petersen, Anders Christensen, Ingbritt Butina, Birte Christensen-Dalsgaard
Afbud: Annette Winkel Schwartz, Svend Larsen, Erland Kolding Nielsen, Niels-Henrik Gylstorff, Per Mogens Petersen

1. Godkendelse af dagsorden (referat er godkendt i marts 2007)
Dagsordenen blev godkendt med den kommentar, at det ikke var givet, at organisationsideen i dagsordenens punkt 3 umiddelbart kunne gennemføres i den simple form. Den lød "placering af ledelsesansvar i programgruppe Informationsforsyning og bevilling i DEFF-sekretariatet v. Bo"

Bo erindrede om at gruppen var nedsat for at rådgive omkring disponeringen af den million, der som en del af 2006-budgettet blev dedikeret formålet data til integrated search. Siden var DEFF blevet reorganiseret og Programgruppen for Informationsforsyning oprettet med dette område som en af sine hovedopgaver. Det drejer sig derfor om at finde ud af, hvorledes denne gruppe og de reserverede midler bedst muligt kan bidrage til fremskridt på området indenfor DEFF's nye ramme.

2. Overblik over materiale
E-bøger
Det bør afklares, hvad der er relevante metadata for e-bøger – herunder informationer som ikke hører med til alm. bibliografiske data såsom licensbetingelser. Skal e-bøger behandles som e-tidsskrifter eller bøger? Hvorledes ønsker bibliotekerne at integrere e-bogsdata i deres søgesystemer? Hvilke datakilder og –leverandører er til rådighed – og hvad er omkostningerne? En afklarende udredning efterspørges.
Det forventes at Programgruppen for Informationsforsyning adresserer disse punkter – antagelig i form af et afklarende projekt og/eller konsulentundersøgelse.

Tidsskrifter
Til mødet var udarbejdet en analyse af DADS dækning af dels DEFF-licenserne og dels øvrige licenser fra KB, SB og DTV. DADS etablerer en bred basisdækning ved hjælp af "minimumsdata" fra SwetsScan og uddyber så denne dækning, forlag for forlag, ved at importere "maksimumsdata" direkte fra forlagene. En del forlag kræver afgift for håndtering/levering af data. Ebsco er en tilsvarende leverandør af "minimumsdata" (foruden også at være en licenspakke-aggregator). Det vil være nyttigt at få nærmere undersøgt og beskrevet forskellene på hhv. minimumsløsningen og maksimumsløsningen for de enkelte forlag.

Der findes internationalt løsninger svarende til DADS, som f.eks. Elin i Lund, Los Alamos, OhioLink, University of Toronto og der er et nationalt system på vej i Tyskland i forbindelse med nationallicenserne. Endvidere beskæftiger systemer som TDNet og Article First (OCLC) sig også med tidsskrift- og artikeldata på forskellig vis.

Hvis artikeldata (som dem i DADS) skal distribueres til lokale integrated search systemer i DEFF landskabet, vil det kræve tillægsaftaler til DEFF licenserne. I forbindelse med DADS er der gode erfaringer med tillægsaftaler, der placerer data og søgemaskine ét sted i Danmark. Hvis data skal leveres videre fra den nationale databrønd til lokale søgemaskiner i DEFF bibliotekerne, skal dette aftalekoncept udvides, da forlagene ønsker præcis kontrol med udbredelsen og administrationen af deres data. Det "DADS som Primo/Lucene 3rd node" koncept, som DTV og Ex Libris pt. udvikler til KB, falder inden for det kendte koncept, da søgemaskinen fortsat befinder sig på DTV.

Skal der defineres et nationalt format? Vi kan ikke definere, hvad forlagene leverer. Forlagene (og øvrige dataleverandører) leverer stort set i hvert deres eget format. Harmoniserende fællesformater (til samsøgning, deduplicering, clustering mv.) må nødvendigvis tage udgangspunkt i forlagenes realiteter.

3. Organisation af fælles data projektet (den fælles databrønd)

Etableringen af en fælles national databrønd bestående af høstede/leverede metadata, indholdsfortegnelser, resumeer og fuldtekstdata på artikelniveau fra tidsskrift- og databaseudbyderne fordrer et udredningsarbejde, som bør i gang sættes hurtigst muligt.

Overordnet var der enighed om at etableringen og den videre drift af en national databrønd afstedkom opgaver som bedst lod sig løse og håndtere i et samarbejde mellem 2 forskellige organisatoriske enheder: (1) en databrøndoperatør og (2) DEFF.

(1) Databrøndsoperatørens opgaver vil overordnet set være at indsamle, opbevare, normalisere og formidle de relevante data (metadata/fuldtekster mv.)

(2) DEFF's opgave vil være at varetage kontakten til forlagene vedr. rettigheder og aftaler om levering af metadata mv. til den nationale databrønd.

Der var endvidere enighed om, at det ville være mest hensigtsmæssigt, at gennemføre en udbudsforretning for at finde den rette operatør. På mødet nævntes følgende - måske -mulige leverandører: DADS, DBC, Swets, Ebsco og CSA, men der ikke gennemført nogen undersøgelse heraf.

For at gennemføre det nødvendige udredningsarbejde, som i princippet omfatter både en kravs specifikation til en databrøndsoperatør og til DEFF, anbefales det at nedsætte en arbejdsgruppe (eller et projekt) i DEFF regi under programområdet informationsforsyning. Det bør overvejes at

tilknytte ekstern konsulentbistand til arbejdsgruppen, ligesom DEFF sekretariatet bør være repræsenteret på højt niveau samt bidrage substantielt til udredningsarbejdet.

Arbejdet bør organiseres i to klart adskilte faser:

(1) **Afklaringsfase**, der besvarer generiske spørgsmål som: Hvad er en DEFF databrønd – koncept og profil? Hvilke datastrømme, hvilke potentielle leverandører? Hvilket aftalemæssigt grundlag for dataleverancer? Hvilke aftagere i DEFF landskabet? Hvilke behov og ønsker har disse? Hvilke typer bidragsmodeller til finansiering af driften kan tænkes, vil være ønskelige/realistiske? Hvilke formater og standarder findes/kan/bør følges? Hvad er de internationale erfaringer? -- Denne fase kan inddrage alle relevante parter – også sådanne, der evt. måtte overveje at deltage i et senere udbud af operatøropgaven.

(2) Udbudsgrundlag, hvor afklaringen overfor omformes til et reelt udbudsgrundlag med kravspecifikation mv. – antagelig bistået at et udbudsvant konsulentfirma. Denne fase kan ikke inddrage parter (eller personer med nære relationer til sådanne), der ønsker at deltage i det kommende udbud af operatøropgaven.

DEFF sekretariatet må sørge for at præcise grænsedragning mellem fase 1 og 2 er korrekt og klart kommunikeret. Efter fase 2 kan DEFF sekretariatet gennemføre en udbudsforretning, antagelig bistået af et konsulentfirma.

Udover at der var enighed om, at DEFF skulle varetage den ovenfor beskrevne opgave i forbindelse med etableringen af den nationale databrønd blev der peget på to relaterede opgaver, som ville komplettere løsningen og som eventuelt kunne løses i DEFF-regi.

(1) DEFF kunne varetage licensstyringen af samtlige licenser på forskningsbiblioteksområdet (f.eks. i en konsortieløsning baseret på et system som VERDE fra Ex Libris)

(2) DEFF kunne drive og vedligeholde en fælles link-resolver på forskningsbiblioteksområdet (f.eks. i en konsortieløsning baseret på SFX-systemet, der allerede er meget udbredt i sektoren)

Opgave (1) bør forelægges Licensgruppen, mens opgave (2) allerede er sat i værk i Programgruppen for Informationsforsyning, der behandler emnet på sit næste møde.

4. Præcisering af udeståender og planlægning af det videre arbejde
Se punkt 3.

5. Evt.
Der var intet under eventuelt.