



ABM ARBEJDSGRUPPE

Afrapportering fra arbejdsgruppen om ABM-standarder til styregruppen, august 2003

Forord

Kulturministeriet nedsatte i januar 2003 en styregruppe og en arbejdsgruppe, der fik til opgave at udarbejde anbefalinger for samarbejdet mellem institutioner på arkiv-, museums- og biblioteksområdet. Styregruppen består af direktørerne for Statens Arkiver, Biblioteksstyrelsen og Kulturarvsstyrelsen, mens arbejdsgruppen udgøres af to personer fra hvert af de tre områder. Opgaven er at vurdere definitioner på dataindhold, dataformater og datatransport med henblik på at etablere fælles præsentation af informationer fra arkiver, biblioteker og museer på internettet.

Initiativet er taget i forlængelse af igangsatte projekter om udvikling af lokale og nationale, fælles søgedatabaser på ABM området, som Nordjyllands Kulturhistoriske Søgebase (NOKS) www.noks.dk, www.historiefyn.dk og www.danskebilleder.dk. Arbejdsgruppen har i sit arbejde derfor taget udgangspunkt i de regionale forsøg for at sikre, at kommende initiativer kan drage nytte af udvalgets anbefalinger, og at man i disse projekter derved kan minimere overvejelserne omkring dataindhold, dataformater, datatransport og præsentation. Samtidig undgås det, at kommende regionale forsøg udvikles i forskellige retninger til skade for integrationen. Det er således udvalgets ambition at stille en ”værktøjskasse” til rådighed, der både kan anvendes i forbindelse med de regionale forsøg, og som kan skaleres til også at omfatte et ABM samarbejde på landsplan.

Der lægges ved opgaveformuleringen vægt på, at der ved opbygningen af søgefaciliteter etableres et løbende samspil mellem forskellige typer af databaser dækkende de forskellige områder. Det sker ikke ved at opbygge en ny, sammenhængende database, som er skabt ved engangsudtræk fra andre databaser. En sådan model er velegnet i forbindelse med udvikling af prototyper og til pilotprojekter, men egner sig ikke til systemer i drift. Her er det nødvendigt, at alle interessenter overholder fælles standarder. Præsentationssystemer på ABM området kan således opbygges på flere måder, men det afgørende er, at der som grundlag anvendes fælles beskrivelser af dataindhold, dataformater og datatransport.

Da vi på dette område er i en tidlig fase, vil arbejdsgruppens anbefalinger naturligt være præget af en vis usikkerhed, som efterfølgende vil kunne mindskes ved et fortsat samarbejde.

Arbejdsgruppens kommissorium

”Der nedsættes en arbejdsgruppe af Biblioteksstyrelsen, Kulturarvsstyrelsen og Statens Arkiver med det formål at vurdere fælles definitioner på dataindhold, dataformater og datatransport med henblik på at kunne etablere fælles præsentation af data om materialer fra arkiver, biblioteker og museer og stille relevante forslag hertil.

Gruppen nedsættes af direktøren i Kulturarvsstyrelsen, direktøren i Biblioteksstyrelsen og rigsarkivaren, der tillige udgør en styregruppe, der refererer til Kulturministeriets departement. Arbejdsgruppen refererer til styregruppen.

Arbejdsgruppen består af 6 personer, to fra arkiverne, to fra museums- og to fra biblioteksområdet. Formand for arbejdsgruppen er kontorchef Henrik Jarl Hansen, Kulturarvsstyrelsen.

Arbejdsgruppen skal inden den 1. juli 2003 levere udkast til anbefalinger.”

Arbejdsgruppens sammensætning

Arbejdsgruppen har følgende sammensætning:

- Anders Sode-Pedersen, Statens Arkiver
- Jens Topholm, Aalborg Stadsarkiv
- Leif Andresen, Biblioteksstyrelsen
- Henning Midtgaard Hanssen, Statsbiblioteket i Århus
- Henrik Jarl Hansen, Kulturarvsstyrelsen
- Eske Wohlfahrt, Kulturarvsstyrelsen

Arbejdsgruppen har undervejs i forløbet inddraget Sebastian Hammer, Indeks Data og Hans Uwe Petersen, Arbejderbevægelsens Bibliotek og Arkiv, der begge deltog i et af gruppens i alt fem møder.

Baggrund og konkretisering af opgaven

Givet den korte tidsramme og det komplekse område er gruppens anbefalinger nødvendigvis relativt generelle og bør efterfølgende konkretiseres. Vi har således valgt at levere en skitse til en løsning baseret på en række hovedlinjer og generelle anbefalinger. Målet har ikke været at udvikle egne standarder, men at anvende og profilere internationale standarder. I et længere perspektiv kan det dog give mening at opløfte eventuelle egne fortolkninger og udfoldninger af internationale standarder til relevante nationale og internationale fora, i første omgang Dansk Standard.

For at konkretisere arbejdet blev det tidligt i forløbet besluttet, at gruppen ville give et bud på hvilke databaser fra de tre sektorer, det i første omgang er relevant at inddrage i arbejdet, og ud fra disse at identificere på hvilket niveau dataudveksling er realistisk og giver mening. Det drejer sig i første omgang om: DANPA, Arkibas og Daisy fra arkiverne, DanBib og bibliotek.dk fra bibliotekerne samt Museernes Samlinger, Kunst Indeks Danmark og Fund og Fortidsminder, som er centrale baser i museumsverdenen. I bilag I gives en kort beskrivelse af hver enkelt af de nævnte baser.

Gruppen har også ønsket at inddrage erfaringer fra allerede eksisterende ABM samarbejder i form af etablerede søgedatabaser og lokalhistoriske portaler såsom Nordjyllands Kulturhistoriske Søgebase (NOKS), Danske Billeder, Historie Fyn m.fl. De nævnte er alle eksempler på samle- eller udtræksdatabaser, hvor informationerne er hentet fra fagspecifikke databaser og derefter bearbejdet for at kunne fungere i de relevante sammenhænge. Gruppen har endvidere anvendt de nævnte projekter som inspirationskilde i forbindelse med en afklaring af forskelligheder imellem institutionernes data og en kortlægning af brugernes søgebehov og måder at finde informationer på. Det har givet anledning til at arbejdsgruppen har valgt også kort at beskæftige sig med begreberne selektion og emnesøgning som en del af den fælles præsentation.

Arbejdsgruppens hovedopgaver ifølge kommissoriet:

- dataindhold
- dataformater
- datatransport
- fælles præsentation (herunder selektion og emnesøgning)

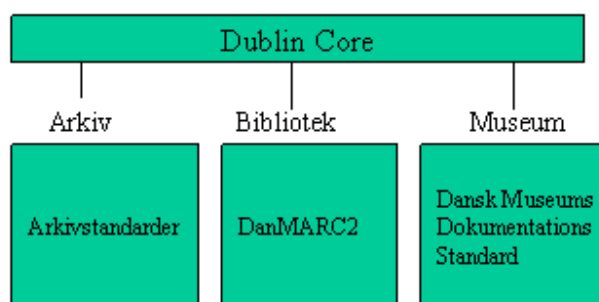
Arbejdsgruppens løsningsforslag

Dataindhold

Arbejdsgruppen har taget som udgangspunkt i videst mulige omfang at anvende gældende internationale standarder og også at lægge sig op ad generelle trends.

Den mest anvendte systematik til at beskrive data i sektoroverskridende sammenhæng er ”Dublin Core” (DC) (ISO 15836). Dublin Core består basalt set af 15 informationskategorier eller elementer, men har en overbygning med kvalificering dels i underelementer og dels i kvalificering af elementer (eksempelvis: Subject med et schema = et klassifikationssystem). Dublin Core tager historisk set sit udgangspunkt i beskrivelse af websites og elementer i tilknytning hertil. Der anvendes betegnelsen ”document-like objects”, hvilket stammer fra det tidspunkt, hvor Dublin Core blev udvidet fra 13 til 15 elementer. Udlagt på mere forståeligt dansk betyder det, at ikke-tekst materiale (ikke mindst billeder) kan beskrives.

Helt overordnet kan gruppens løsningsforslag visualiseres i nedenstående tegning, der viser, hvordan Dublin Core forventeligt vil kunne benyttes som et fælles, konkret indholdsformat i forbindelse med de enkelte sektorer specifikke databaseindhold



Til de enkelte DC elementer kan man henføre de meget forskelligartede oplysninger i de enkelte sektorer databaser. Man taler om at ”mappe” indholdet af en database op mod Dublin Cores 15 elementer.

Vi valgte i arbejdsgruppen at tage udgangspunkt i de tre sektorer egne registreringsskemaer og på basis af disse selekttere de dataelementer, som det blev vurderet ville være af betydning i forhold til præsentation og søgning set fra almenheden. Dernæst blev de udvalgte dataelementer sat i relation til Dublin Core. I første omgang alene til de 15 elementer, men det viste sig, at der er behov for en yderligere kvalificering for at kunne bevare betydningsindholdet for en hel del dataelementers vedkommende.

I bilag II ses arbejdsgruppens forsøg på at lægge eksempler på data fra hver af de tre sektorer ind i et Dublin Core (DC) format. Gruppen anser det for vigtigt, at der sker en yderligere vurdering af fordele og ulemper ved at benytte dette internationale format. Dublin Core er især kendt i biblioteksmiljøet og har indtil videre større udbredelse og accept her end i arkivverdenen og blandt museerne. Naturligvis sker der et informationstab, når man forsøger at repræsentere en databases indholdsunivers i kun 15 prædefinerede elementer. Det er derfor gruppens opfattelse, at det skal afklares, om ikke tabet af information opvejes af

den gevinst, der ligger i at benytte et internationalt standardformat for beskrivelse af dataindhold. Afhængig af databasernes kompleksitet vil man opleve, at det kun giver mening, eller i praksis kun er muligt at repræsentere en delmængde af de enkelte basers informationer i Dublin Core. Det er således vigtigt for hver enkelt database at identificere det realistiske niveau, som kan danne grundlag for dataudvekslingen.

Formålet er derfor at definere en fælles ramme, som gør, at data fra arkiver, biblioteker og museer over for almenheden kan præsenteres i sammenhæng. Dette formål afgrænser opgaven ud fra (mindst) to vinkler:

- ikke alle data fra databaser i ABM institutionerne skal kunne indgå i denne fælles ramme, men alene dem som har betydning i forhold til præsentation og søgning set fra almenheden
- krav til detailrigdom ligger ikke på samme niveau som i ABM institutionernes egne databaser til faglig brug, fordi målgruppen er bredere. Men det skal afklares, hvordan man efter at have fundet en relevant post kan komme videre til ABM institutionernes egne databaser, og der få adgang til de fulde informationer og eksempelvis kunne foretage en bestilling.

Arbejdsgruppen har ikke fundet alternativer til Dublin Core. Gruppen er opmærksom på, at der især knytter sig uløste problemer til beskrivelser af mere komplekse strukturer. Der udestår også en del arbejde i forhold til en færdig beskrivelse af en fælles udfoldet Dublin Core model og i denne sammenhæng også en del arbejde i forbindelse med den detaljerede omsætning af sektorspecifikke dataelementer til fælles Dublin Core dataelementer.

Vi mangler bl.a. at tage stilling til, om data skal udfoldes, således at der kan foretages specifikke søgninger (f.eks. på forskellige typer ophavsmænd), eller om vi i denne sammenhæng i stedet skal fokusere på at kunne vise ophavsmænds forskellige roller ved præsentationen på skærm.

Arbejdsgruppens indstilling er derfor, at der arbejdes videre med at undersøge Dublin Cores anvendelighed som basis med henblik på at kunne levere en fælles beskrivelse af dataindholdet, som er brugbar for de tre sektorer.

Dataformater (rammeformater)

Arbejdsgruppen er endnu ikke gået i detaljer med dette punkt. Det skyldes, at den generelle model efter vores opfattelse er enkel set ud fra betragtningen om i videst mulige omfang at anvende gældende internationale standarder og at lægge sig op ad generelle strømninger og anbefalinger.

Vi peger derfor på, at formatet skal være XML, som er en W3C standard, og som er adopteret i utallige sammenhænge. XML har ingen aktuelle konkurrenter som schema for udveksling af data mellem systemer.

Der skal derfor bygges videre på det internationale arbejde med XML-schemaer i forhold til Dublin Core og i denne sammenhæng også det danske arbejde i regi af XML-komiteen under Ministeriet for Videnskab, Teknologi og Udvikling. XML som fælles offentlig standard er med til at garantere, at informationer kan udveksles på tværs af den offentlige sektor. Samtidig sikres det, at allerede genererede data kan genanvendes optimalt.

Arbejdsgruppens indstilling er derfor, at der arbejdes videre med XML schemaer som basis.

Datatransport

Opgaven med datatransport skal ses i relation til udgangspunktet om at skabe en sammenhængende præsentation af data fra arkiver, biblioteker og museer over for almenheden i lokalt, regionalt og evt. nationalt regi. Arbejdsgruppen forholder sig ikke til spørgsmålet om, hvorvidt der regionalt eller på landsplan skal etableres ABM databaser – gruppen har alene ambition om at pege på anbefalelsesværdige værktøjer og løsningsmodeller i mulige, konkrete ABM samarbejder.

Det er vigtigt, at de endelige standarder ikke låser til en bestemt model, hvilket de foranstående anbefalinger om dataindhold og dataformat netop heller ikke gør.

Selve datatransporten kan principielt ske på to forskellige måder:

- Søgning i de sektorspecifikke databaser (evt. parallelt med søgning i andre databaser) efterfulgt af samlet præsentation = datatransport ”on the fly” af de ønskede dataposter
- Søgning i en fælles database = datatransport af alle relevante dataposter på forhånd

Decentral søgning i de originale baser i data, der ikke er konverteret til et fælles dataformat vil indebære nogle ekstra omkostninger. Der skal ske en parallel omsætning ”den modsatte vej” af søgetermer/metoder fra en fælles model til de sektorspecifikke modeller. Endvidere skal der ske konvertering ”on the fly” til det fælles dataformat. I det omfang der fokuseres på selektion af informationer fra databaser i de enkelte sektorer, vil der også være problemer med søgning i sådanne delmængder afspejlet ned i de sektorspecifikke databaser.

Central søgning i en database medfører i sagens natur, at der skal etableres en løbende opdateringsstrøm fra de sektorspecifikke databaser til den fælles database.

Parallelt hermed er der to forskellige ”skoler” med tilhørende standarder:

- Decentral søgning med Z39.50 (=ISO23950) som den etablerede model i biblioteksverdenen eller med SRW som en kandidat til bredere anvendelse baseret på http.
- Central søgning med mange forskellige opdateringsmodeller baseret på filtransport med ftp til mere standardiserede metoder baseret på OAI-PMH høst initieret fra den fælles service.

Det er arbejdsgruppens opfattelse, at den mest realistiske model for lokale implementeringer vil være en central model med en fælles database, etableret gennem ”harvesting”, der sikrer til stadighed aktuelle oplysninger. Det nære perspektiv her vil være at udarbejde en beskrivelse for brugen af OAI-PMH, men med det forhold for øje at det i projektfaser og indledende implementeringer vil være muligt at ”håndstyre” filtransport via ftp.

På lidt længere sigt bør SRW holdes under opsigt, da der kan ske et gennembrud i bredere sammenhæng af denne pt. eksperimenterende model. Det hører med, at det er arbejdsgruppens indstilling, at det vil være en forkert anvendelse af ressourcerne i arkiver og museer at implementere Z39.50 på nuværende tidspunkt i forbindelse med lokale projekter. Erfaringerne fra biblioteker viser, at det har kostet relativt lang tid at nå frem til brugbare implementeringer, og med SRW i horisonten vil det næppe være realistisk at indføre Z39.50 i de enkelte arkiver og på museerne.

Arbejdsgruppen ser ikke andre modeller for "remote procedure calls", idet udbredte søgesprog som SQL og XML Query netop baserer sig på, at den konkrete databaseopsætning er kendt af den forespørgende maskine i modsætning til, hvad der gælder for Z39.50 og SRW.

Fælles præsentation

Præsentation af data involverer spørgsmål om målgrupper, selektion, søgning, visning, design m.v. Arbejdsgruppen har i denne sammenhæng valgt at behandle selektion og søgning.

Selektion

Spørgsmålet om selektion kan være bestemt af databasens formål eller enkeltposters struktur og formål.

I regionale kulturhistoriske baser kan både sted, tid og emne begrunde afgrænsninger i alle tre typer af institutioners materialer. Samme overvejelser vil være relevante for eksempelvis nationale emnedatabaser.

Enkeltposter kan indeholde oplysninger af intern/forskningsmæssig/administrativ art, der ikke er beregnet for offentligheden.

Arbejdsgruppen ønsker at gøre opmærksom på, at selektion er en relevant problematik, der gælder alle ABM data og -databaser.

Søgning

Overvejelser omkring søgning og søgestrukturer er særligt relevante, fordi ABM baser retter sig mod almenheden, og fordi det er vanskeligt at samsøge i tre forskellige strukturer.

Mange brugere vil have specifikke formål, der kan opfyldes ved fritekstsøgning med trunkering m.m.

Der vil også være brugere, der stiller krav til emne, tid, sted og repræsentativitet. Hvis det skal tilfredsstilles, må gældende strukturer i ABM institutionerne udnyttes. Det kan gøres ved at udarbejde en struktureret ABM emneordsliste på 300-500 emneord med udgangspunkt i de eksisterende systemer og ved at etablere fælles standarder for tid og sted.

De forventeligt langt større søgeresultater skal kunne indsnævres ved gode faciliteter for raffinering, hvori kunne indgå system- eller tidstypiske begreber og termer.

Man kan forvente, at der ved en analyse af de forskellige systemer og en kortlægning af udvalgte institutioners reelle brug af samme kan skabes grundlag for at udarbejde en struktur, samt at der i forlængelse af et sådant arbejde kan skabes det første grundlag for en national standard.

Arbejdsgruppen indstiller, at det også indgår i det videre arbejde at afklare kriterier og principper for søgning og at udarbejde en emneordsliste med udgangspunkt heri.

Anbefalinger for det videre arbejde

Anbefalingerne i denne rapport er som nævnt holdt på et generelt niveau, der peger på den retning og de tekniske løsninger, vi gerne ser nye ABM samarbejdsprojekter følger. Der er tale om et udkast på overordnet niveau, der forsøger at afstikke en generel kurs. Men det forhindrer ikke, at vi vurderer det som ønskeligt at gå dybere ned i substansen på alle hovedområder. Det gælder ikke mindst inden for mapningen af informationer fra de relevante databaser op mod Dublin Core (se de sektorspecifikke eksempler i bilag II). Det er vigtigt, at dette arbejde bliver gjort på en ensartet måde for alle sektorer, således at sammenligneligheden understøttes mest muligt. Arbejdsgruppen peger derfor på, at der bør afsættes midler, der gør det muligt for gruppen at tilknytte ekstern ekspertise, som skal løse følgende konkrete opgaver:

- Raffinere håndtering af metadata, dvs. konvertering af sektorspecifikke standarder til Dublin Core
- Overføre det resulterende Dublin Core univers til et sammenhængende XML-schema
- Udarbejde en emneordstilgang og angive kriterier og principper herfor
- Sikre tværgående konsistens mellem sektorerne i de enkelte opgaveløsninger

Det vil være hensigtsmæssigt, at denne ekspertise ikke repræsenterer en enkelt af de deltagende styrelser.

Det foreslås samtidigt at den eksisterende ABM arbejdsgruppe fortsætter i denne periode, og at kommissoriet afspejler de nye opgaver.

Desuden vil det være frugtbart, hvis der kunne skabes mulighed for konkret afprøvning af de ”værktøjer”, som den foreliggende rapport anbefaler. Det kunne evt. ske gennem et nærmere konkret samarbejde med et af de regionale ABM forsøg, f.eks. i forbindelse med udviklingen af NOKS version 2 baseret på de i nærværende rapport anbefalede og beskrevne principper. Eller det kunne ske ved en test implementering i form af en landsdækkende, emneafgrænset ABM database eksempelvis med information om industrisamfundet.

Ordliste

Nedenstående ordforklaringer er ikke fagligt formelt dækkende, men er skrevet med henblik på at gøre rapportens tekst mere forståelig.

Dublin Core: (ISO 15836 The Dublin Core Metadata Element Set) Format til registrering af identificerbare informationsressourcer. Er oprindeligt udarbejdet til web-dokumenter m.v., men har de senere år fået en bredere anvendelse. Dublin Core er nu en ISO-standard i den enkle udgave med 15 elementer og sigter mod sektoroverskridende anvendelse.

Metadata: Data om data – og derfor en fælles betegnelse for alle registreringsformater og mere til. Anvendes til tider synonymt med metadata for webdokumenter.

OAI: (Open Archives Initiative) Standard for indhøstning af dokumenter og/eller metadata fra nettet. Oprindelig udviklet for at kunne stille preprintarkiver til fri rådighed. Med OAI-PMH (protocol for metadata harvesting) videreudviklet til indhøstning af metadata – primært i Dublin Core.

XML: (eXtensible Markup Language) Format til transport og lagring af data af alle typer, inkl. metadata, fuldtekst og meddelelser. Er en videreudvikling af HTML i W3C's regi. XML er ikke helt moden, men har allerede en bred anvendelse – og ikke mindst en meget bred planlagt anvendelse. Den formelle definition af et XML-format sker enten som et XML Schema eller som en XML DTD.

W3C: (World Wide Web Consortium) International organisation, som består bl.a. af førende IT-virksomheder og som de facto sætter standarderne for internettet, herunder i særdeleshed for World Wide Web.

Z39.50: (Information Retrieval (Z39.50): Application Service Definition and Protocol Specification = ISO23950) en standard udviklet før WWW blev opfundet. En metode for et program (klientprogram) til at bede et andet program (serverprogram) om at udføre en bestemt søgning eller anden handling i tilknytning hertil og svare med resultatet. Z39.50 har således ikke nogen brugergrænseflade, da denne er mellem bruger og klientprogram. Z39.50 er primært anvendt i bibliotekssektoren. Et klientprogram kan være et bibliotekssystem, et pc-program eller en web-gateway.

SRW: (Search and Retrieve on the Web) For at tilnærme Z39.50 til WWW-miljøet er udviklet en simplere metode for et program (klientprogram) til at bede et andet program (serverprogram) om at udføre en bestemt søgning eller anden handling i tilknytning hertil og svare med resultatet. Der benyttes http-transport på linie med en webbrowser. SRW er ikke nogen formel standard, men kan ses som en kandidat til en søgeprotokol på WWW.

Trunkering: Betegnelse for at man ved søgning kun skriver en del af det ord, man vil søge på. F.eks. ”skole?” for samtidig at søge skolelærere, skolevæsen etc.

Bilag I

I det følgende gives en kort beskrivelse af de sektorspecifikke databaser, som har dannet udgangspunkt og reference for gruppens arbejde. Det er også fra nogle af disse databaser, at eksemplerne i bilag II er taget.

Arkiverne

Inden for arkiververdenen findes der en landsdækkende database DANPA samt en lang række institutionelle databaser. I den nærmeste fremtid kommer Statens Arkiver med Daisy og Sømmenslutningen af Lokal Arkiver (SLA) med en ny version af ARKIBAS.

- **DANPA** www.danpa.dk er en landsdækkende database for privatarkivmateriale. Efter opdateringen i 2002 indeholder DANPA materiale fra 190 arkivinstitutter. Antallet af registreringer er 120.000, og der er 265.000 henvisninger at søge på. Dataleverandører er Statens Arkiver (Rigsarkivet og landsarkiverne), Sømmenslutningen af Lokal Arkiver (SLA), Det Kongelige Bibliotek, Arbejderbevægelsens Bibliotek og Arkiv, Udvandrerarkivet samt andre specialarkiver. Perspektiverne og ambitionerne for den videre færd frem er at udvide med offentlige arkivalier og andre kildetyper – først og fremmest billeder. DANPA er en udtræksdatabase, der søger at overkomme de enormt forskellige registreringspraksis, der er en realitet i arkiververdenen. Databasen er stærkt præget heraf – både i præsentation og søgemuligheder. Det er en klar konklusion, at der i kølvandet på DANPA skal besluttet standarder og anvendelse af samme.
- **ARKIBAS.dk**. Sømmenslutningen af Lokal Arkiver (SLA) er ved at udvikle en central ARKIBAS-4, hvorfra der skal spejles over i webudgaven kaldet ARKIBAS.dk. Sidstnævnte vil være en realitet i efteråret 2004 eller måske mere sandsynligt i foråret 2005. Alle ARKIBAS data, der ikke er med i DANPA, bliver hermed tilgængelige. Der er ca. 1,5 mio. registreringer, heraf eks. 200.000 scannede billeder med registreringer.
- **Daisy** er Statens Arkivers database til registrering af oplysninger om samtlige arkivalier og arkivskabere mv. i Statens Arkivers varetægt. Daisy er under udvikling, men har været i anvendelse som indtastningsdatabase siden efteråret 2000. Daisy indeholder til dels konverterede data fra en tidligere arkivdatabase, men der vil i de kommende år foregå en total omregistrering af samtlige arkivalier og arkivskabere. Når denne proces er overstået, vil der være registreret ca. 70.000 arkivskabere, ca. 400.000 heuristiske enheder samt ca. 3 mio. pakker og bind. Udover at være en indtastningsdatabase, vil Daisy ved udgangen af 2004 tillige være en søgedatabase, som kan tilgås via nettet. På længere sigt vil det være muligt at bestille arkivalier on-line via Daisy samt få direkte adgang til digitale arkivalier.

Bibliotekerne

Der findes to nationale databaser DanBib og bibliotek.dk. Derudover findes en lang række databaser over en eller flere bibliotekers bestande.

- **DanBib** består af katalogiseringer af alle medier på danske folke- og forskningsbiblioteker, den danske nationalbibliografi (inkl. tidsskriftsartikler og anmeldelser), nordiske tidsskrifter samt katalogposter fra Library of Congress og British Library. Det fælles registreringsformat er danMARC2, men basen er forberedt til parallel brug af Dublin Core. DanBib er et professionelt

værktøj og anvendelsen forudsætter et abonnement. Adgang via Z39.50 og http. Omfang: 20 mio. poster.

- **bibliotek.dk** rummer en delmængde af katalogposter i DanBib dækkende katalogiseringer af alle medier på offentlige danske folke- og forskningsbiblioteker, den danske nationalbibliografi (inkl. tidsskriftsartikler og anmeldelser). bibliotek.dk er frit tilgængelig og giver brugerne mulighed for at søge og bestille materiale. Adgang via http. Databasen har et omfang på ca. 7 mio. poster.

Museerne

På museumsområdet findes der i dag tre centrale databaser. Det drejer sig om Fund og Fortidsminder, hvori en delmængde af den ”faste kulturarv” er registreret samt om Kunst Indeks Danmark og Museernes Samlinger, der tilsammen dækker ”den løse kulturarv”. I den aktuelle fase har vi valgt at koncentrere os om kulturhistoriske data inden for den løse kulturarv registreret efter ”Dansk Museums Dokumentations Standard” og repræsenteret i Museernes Samlinger.

- **Kunst Indeks Danmark (KID)** er den centrale oversigt over danske kunstværker i offentligt eje. Alle statsanerkendte og statslige museer er i medfør af museumsloven pålagt at indberette deres værker til KID. Aktuelt rummer KID information om 38.000 kunstværker fordelt på knap 90 institutioner. KID er offentligt tilgængelig på www.kid.dk.
- **Museernes Samlinger**, som er offentligt tilgængelig på www.kulturhistorieonline.dk, er den tilsvarende centrale oversigt over museumssager og -genstande på danske statsanerkendte og statslige kulturhistoriske museer. I modsætning til Fund og Fortidsminder og KID er Museernes Samlinger helt ny og foreligger kun i en forsøgsversion med 20 deltagende museer og i alt knap 400.000 poster. I modsætning til de to andre databaser er Museernes Samlinger fra begyndelsen anlagt som en kompilering af decentralt registrerede oplysninger.
- **Fund og Fortidsminder** (tidligere Det Kulturhistoriske Centralregister) er tilgængelig på www.dkconline.dk. Databasen omfatter dele af den ”faste kulturarv” - dvs. den stedbundne fysiske kulturarv i landskabet. Dens datagrundlag er dels løbende indberetninger om udgravninger fra landets arkæologiske museer, dels information om fortidsminder ekstraheret fra arkiver, primært de såkaldte sognebeskrivelser. Fund og Fortidsminder rummer information om ca. 158.000 kulturhistoriske lokaliteter fra landjorden og fra det marine område.

Bilag II

I dette bilag gives eksempler på udtræk af poster fra tre af de databaser, der beskrives i bilag I. Disse eksempler er forsøgt ”mappet” mod Dublin Cores 15 elementer og ekstensions. Det er vigtigt at understrege, at eksemplerne har karakter af dokumenter under behandling. De er kun taget med her for at illustrere mulighederne for at beskrive indholdet i databaserne i relation til DC, ikke som et endegyldigt bud på sektorspecifikke mapninger. Eksemplerne har endnu ikke været genstand for en gennemgribende tværfaglig diskussion med henblik på at ensarte anvendelsen af Dublin Core i de tre sektorer. Dette arbejde ligger først i den næste fase af gruppens arbejde jfr. afsnittet om det videre arbejde.

Arkiverne

ARKIBAS post – sat i relation til Dublin Core

Dublin Core			ARKIBAS	
Element nr.	Label	Dansk navn	Felt	Definition
1	D.C. Title	Titel	Reg 12	Beskrivelse: A- og Æ-reg Arkivbetegnelse
2	DC.Creator	Forfatter eller ophav	Reg 12	Beskrivelse: A- og Æ-reg Arkivbetegnelse Kommentar: I ARKIBAS 4 kommer et selvstændigt Arkivskaberfelt
3	DC.Subject	Emne og nøgleord	Reg 14	Kategorikode: A- og Æ-reg, koder for person, forening, erhverv, institution, offentlig.
			Reg 15	Kategorinavn: A- og Æ-reg Personnavne, foreningstyper, erhverv
			Henvis 14	ARKIBAS DK5
			Henvis 15	Overordningsord
			Henvis 16	Underordningsord
			Reg 31	Foto Billedreg, fotografnavn
4	DC.Description	Beskrivelse	Reg 12	Beskrivelse: Billedreg, KTPreg, Diversereg. Sekvensreg.
			Reg 36	Indhold A- og Æ-reg, sekvensreg.
			Reg 17	Signaturer i indhold: A og Æ reg.

5	DC.Publisher	Udgiver	Reg 3	LHAnr. Opbevarende institutions navn og nummer Alle
6	DC.Contributor	Anden bidragyder	Reg. 33	Bemærkninger kan indeholde relevante oplysninger
7	DC.Date	Dato		-
8	DC.Type	Ressourcetype	Reg	Arkivfond, billede, film, tegning, udklip...
9	DC.Format	Format	Reg 12	Materiale: Billedreg. Pos. Neg. Nitrat...
			Reg 29	Art KTPreg
			Reg 25	TekniskOpl KTPreg
			Reg 37	Attachment Alle
			Reg 30	Negativ Billedreg, neg. Nummer
			Reg 19-24	Antal bd., pakker, læg, længde..
10	DC.Indentifier	Identifikator	Reg 3	Reg. Nr. Alle
			Reg 6	Sr.Nr. Alle særnummerrækker
			Reg 7	Journalnummer på afleveringssagen
11	DC.Source	Kilde (originalt værk)		-
12	DC.Language	Sprog		-
13	DC.Relation	Relationer	Reg 26	Reference KTPreg
			Reg 7	Journalnummer på afleveringssagen, der indeholder afleveringsrelationer.
			Reg. 23	Bemærkninger indeholder ofte relationer til andre a-fonds. I ARKIBAS 4 kommer et egentligt felt med relationer
14	DC.Coverage	Spatial	Reg 8	StedKode
			Reg 9	StedNnavn
			Reg 10	KomKode
			Reg 11	KomNavn

		Temporal	Reg 17	Yderår A- og Æ-reg
			Reg 27	Stiftet A- og Æ-reg
			Reg 28	Nedlagt A- og Æ-reg
			Henvis 12	Tid1 Fra
			Henvis 13	Tid2 Til
15	DC.Rights	Rettigheder	Reg 35	Klausul A- og Æ-reg., billedreg., KTP-reg og sekvensreg.

Statens Arkivers arkivaliebeskrivelsesmetode og Daisy – sat i relation til Dublin Core

Dublin Core			Statens Arkivers arkivaliebeskrivelsesmetode og Daisy	
Element nr.	Label	Dansk navn	Felt/delfelt	Definition
1	DC.Title	Titel	KortNavn	Betegnelse for Heuristisk enheder (arkivalier) valgt af Statens Arkiver.
			Navn Type	Originale betegnelser for Heuristiske enheder (arkivalier). Bemærk at der være flere navne og at de kan være af forskellige typer, f.eks.: Officielt, kaldenavn mv.

Dublin Core			Statens Arkivers arkivaliebeskrivelsesmetode og Daisy	
Element nr.	Label	Dansk navn	Felt/delfelt	Definition
3	DC.Subject	Emne og nøgleord	(Afventer)	

Dublin Core			Statens Arkivers arkivaliebeskrivelsesmetode og Daisy	
Element nr.	Label	Dansk navn	Felt/delfelt¹	Definition
4	DC.Description	Beskrivelse	PubBem	Note med supplerende oplysninger om arkivskaber
			”HIF”	Strukturel beskrivelse af format af Heuristiske identifikatorer. Angiver identifikation af enkelte elementer i en heuristisk enhed og henvisning til andre elementer i den aktuelle heuristiske enhed eller i andre heuristiske enheder.
			”RIF”	Strukturel beskrivelse af format af rækkefølgeidentifikator. Angiver fremfindingsystematik inden for et eksemplar af en heuristisk enhed.
7	DC.Date	Dato	Fra	Startår på heuristisk enhed
			Til	Slutår på heuristisk enhed

¹ Feltbetegnelser er i anførelsestegn, da det enten vil føre for vidt med en mere præcis beskrivelse, eller for disse dele er under udvikling.

Dublin Core			Statens Arkivers arkivaliebeskrivelsesmetode og Daisy	
Element nr.	Label	Dansk navn	Felt/delfelt²	Definition
13	DC.Relation	Relationer	”He2He”	Relationer mellem Heuristiske enheder (arkivalier).
			”He2EP”	Relationer mellem den heuristiske enhed og forskellige eksemplarer af denne.
14	DC.Coverage	Dækning	Afventer	

Arkivskaber

Dublin Core			Statens Arkivers arkivaliebeskrivelsesmetode og Daisy	
Element nr.	Label	Dansk navn	Felt/delfelt	Definition
2	DC.Creator ³	Forfatter eller ophav	Præfiks KortNavn	Betegnelse på overordnet aktør til arkivskaber valgt af Statens Arkiver, og den aktuelle arkivskabers betegnelse valgt af Statens Arkiver
			Navn Type	Originale betegnelser for arkivskaber. Bemærk at der være flere navne og at de kan være af forskellige typer, f.eks.: Officielt, kaldenavn mv.
3	DC.Subject	Emne og nøgleord	(Afventer)	

Dublin Core			Statens Arkivers arkivaliebeskrivelsesmetode og Daisy	
Element nr.	Label	Dansk navn	Felt/delfelt	Definition
4	DC.Description	Beskrivelse	PubBem	Note med supplerende oplysninger om arkivskaber
6	DC.Contributor	Anden bidrager ⁴	(Irrelevant)	
7	DC.Date	Dato	Fra	Startår for Præfiks og KortNavn på arkivskaber
			Til	Slutår for Præfiks og KortNavn på arkivskaber
			Fra	Startår for Navn på arkivskaber
			Til	Slutår på Navn for arkivskaber

² Feltbetegnelser er i anførelsestegn, da det enten vil føre for vidt med en mere præcis beskrivelse, eller for disse dele er under udvikling.

³ NB Der kan være flere arkivskabere (Creator) til en heuristisk enhed (Title).

⁴ Der skelnes ikke mellem bidrager (Contributor) og arkivskaber (Creator), jf. note 1.

Dublin Core			Statens Arkivers arkivaliebeskrivelsesmetode og Daisy	
Element nr.	Label	Dansk navn	Felt/delfelt	Definition
13	DC.Relation	Relationer	OverNGId	Relation til overordnet aktør.
			UnderNGID	Relation til underordnet aktør.
			Fra	Startår på relation til underordnet aktør.
			Til	Slutår på relation til overordnet aktør.
			NG2NGArtID	Type på relation mellem aktører, f.eks.: Intern, Ekstern mv.
14	DC.Coverage	Dækning	Afventer	

Dublin Core			Statens Arkivers arkivaliebeskrivelsesmetode og Daisy	
Element nr.	Label	Dansk navn	Felt/delfelt	Definition
5	DC.Publisher	Udgiver	<i>(Irrelevant)</i>	

Dublin Core			Statens Arkivers arkivaliebeskrivelsesmetode og Daisy	
Element nr.	Label	Dansk navn	Felt/delfelt	Definition
8	DC.Type	Ressourcetype	<i>(Irrelevant)</i>	
9	DC.Format	Format	<i>(Irrelevant)</i>	
10	DC.Identifier	Identifikator	<i>(Irrelevant)</i>	
11	DC.Source	Kilde (originalt værk)	<i>(Irrelevant)</i>	
12	DC.Language	Sprog	<i>(Irrelevant)</i>	

Dublin Core			Statens Arkivers arkivaliebeskrivelsesmetode og Daisy	
Element nr.	Label	Dansk navn	Felt/delfelt	Definition
15	DC.Rights	Rettigheder	<i>(Irrelevant)</i>	

Bibliotekerne

Fælles bibliografiske data i biblioteksposter af relevans for ABM samarbejdet – sat i relation til Dublin Core.

Dublin Core			danMARC2	
Element nr.	Label	Dansk navn	Felt/delfelt ⁵	Definition
1	DC.Title	Titel	239 *t *ø *u *v	Uniform titel for musikværker som opstillingselement - uniform titel, identificerende tilføjelse, tilnavn, uddragstitel
			245 *a *b *ø *æ *n *o *c *u *p *q *r *s	Titel - hovedtitel, resttitel/alternativ titel, identificerende tilføjelser, undertitler, paralleltitler
			247 *g *a *c	Sektionsspecifikation for flerbindsværk eller supplement - sektionsnummer, sektionstitel, sektionsundertitel og anden titelinformation
			248 *g *a *c	Bindspecifikation for flerbindsværk eller supplement - bindnummer eller supplementsbetegnelse, bindtitel eller supplementstitel, undertitel og anden titelinformation
			440 (minus *e *z)	Seriebetegnelse i materialets form
			740	Uniform titel
			745	Varianttitel eller anden titel
			795	Analytisk titel
			840 (minus *z)	Serietitel i normativ form
2	DC.Creator	Forfatter eller ophav	100	Personnavn som opstillingselement
			110	Korporationsnavn (navn på gruppe, institution, stat/kommune, møde) som opstillingselement
			239 *a *h *e *f *c	Navn på komponist/musiker/musikergruppe som del af uniform titel og som opstillingselement - enkeltstående navn, fornavne, tilføjelser, fødselsår
			513	Note om fælles ophavsangivelse (musikoptagelser)
			700	Personnavn som søgeelement
			710	Korporationsnavn (navn på gruppe, institution, stat/kommune, møde) som søgeelement

⁵ Hvor kun feltnummer er angivet, skal det opfattes som gældende alle delfelter i feltet

Dublin Core			danMARC2	
Element nr.	Label	Dansk navn	Felt/delfelt	Definition
3	DC.Subject	Emne og nøgleord	080	UDK (Det universelle decimalklassifikationssystem)
			086	Verbalisering af klassifikationsnotation
			600	Personnavn som emneord
			610	Korporationsnavn (navn på gruppe, institution, stat/kommune, møde) som emneord
			630	Kontrolleret emneord
			631	Ukontrolleret emneord
			645	Titel som emneord
			652	DK5-klassemærke, inkl. alfabetisk underdeling
			666	Kontrolleret DBC emneord
4	DC.Description	Beskrivelse	504	Indholdsbeskrivende note
			507	Note om anledning til udgivelse
			509	Note om besætning (musikaler)
			512	Note til beskrivelsen
			517	Note om målgruppe og anvendelighed
			530	Indholdsnote
			534	Note om partielt indhold
			559	Generel note
5	DC.Publisher	Udgiver	260 *a *b *g *t	Publicering, distribution el. lign. -hjemsted for forlag, navn på forlag, distributør el. lign., navn på producent
			440 *e	Serieudgiver
			710 ⁶	Korporationsnavn (navn på gruppe, institution, stat/kommune, møde) som søgeelement
6	DC.Contributor	Anden bidragyder	245 *f *k	Titel og ophavsangivelse - ophavsangivelse, som ikke giver anledning til opslag, medlemmer af gruppe
			247 *e	Sektionsspecifikation for flerbindsværk eller supplement - ophavsangivelse
			248 *e	Bindspecifikation for flerbindsværk eller supplement - ophavsangivelse
			250 *c	Ophavsangivelse, der hører til udgaven og som giver anledning til opslag

⁶ Kun hvis enkeltperson angives som forfatter eller ophav (element nr. 2)

			770	Personnavn som opslagsselement for analyse
			780	Korporationsnavn som opslagsselement for analyse

Dublin Core			danMARC2	
Element nr.	Label	Dansk navn	Felt/delfelt	Definition
7	DC.Date	Dato	008 *a *z	Generelle søgekoder for udgivelsesår
			260 *c	Publicering, distribution el. lign. – udgivelsesår, distributionsår el. lign.
8	DC.Type	Ressourcetype	005	Søgekoder for musikaler og lydoptagelser
			008 *d *j *k *o	Generelle søgekoder – koder for indholdets form, skønlitterær form, biografi, børne- og skolemateriale
			505	Note om art eller form
			506	Note om disputats eller anden afhandling
9	DC.Format	Format	008 *q	Generelle søgekoder – kode for filtype
			009	Koder for materialebetegnelse
			259	Filbeskrivelse
			300	Fysisk beskrivelse
			501	Note om systemkrav eller adgangsmåde
			856 *q	Elektronisk lokalisering og adgangsmåde - filoverførselsmodus
10	DC.Identifier	Identifikator	021	ISBN
			022	ISSN
			023	Producentens stregkode
			024	ISRC
			027	ISRN
			028	ISMN
			247 *z	Sektionsspecifikation for flerbindsværk eller supplement – sektionens ISBN
			248 *z	Bindspecifikation for flerbindsværk eller supplement – bindets ISBN
			440 *z	Seriens ISSN
			538	Note om numre, der indgår i materialet
			840 *z	Seriens ISSN
			856 *u	Elektronisk lokalisering og adgangsmåde – URI, URL eller PURL
11	DC.Source	Kilde (originalt værk)	520	Note om værkets bibliografiske historie
12	DC.Language	Sprog	008 *l	Generelle søgekoder – kode for hovedsprog
			041	Sprogkoder

			508	Note om sprog og alfabet
--	--	--	-----	--------------------------

Dublin Core			danMARC2	
Element nr.	Label	Dansk navn	Felt/delfelt	Definition
13	DC.Relation	Relationer	241	Originaltitel
			250 *a *x	Udgave- og oplagsbetegnelse
			502	Note om originaltitel eller oversættelse
			521	Note om udgavens oplag
			523	Note om udgivelse i andre materialetyper
			525	Note om komplekse forhold til andre periodica
			526	Note om sammenhæng med andre værker
			557	Periodicum som værtspublikation for I-analyse
			558	Monografi som værtspublikation for I-analyse
			860	Tidligere titel (periodica)
			861	Senere titel (periodica)
			863	Udgivet sammen med/som del af (periodica)
			865	Udgave i andet medium (periodica)
			866	Udgave på originalsprog (periodica)
			867	Hovedudgave på andet sprog og/eller med andet indhold
			868	Udgaver på andre sprog og/eller med ændret indhold (periodica)
			870	Hovedpublikation til supplement (periodica)
			871	Supplement til hovedpublikation (periodica)
			873	Hovedserie (periodica)
			874	Underserie (periodica)
			879	Andre relationer (periodica)
14	DC.Coverage	Dækning	033	Amtskommune eller kommune (Dansk lokalbibliografi)
			620	Overskrift til dansk kortfortegnelse
			621	Emnedata for kartografisk materiale
			622	Systematisk notation for kartografiske materialer
			633	Stednavn som emneord
			634	Periodebetegnelse som emneord
15	DC.Rights	Rettigheder	518	Note om anvendelsesklausul

Museerne

Museumssag 95 udtrykt i Dublin Core - skemaform

Element nr.	Label	Qualifier 1	Qualifier 2	Indhold	Eksempel fra "Museernes Samlinger"	Felt i DMI	Kommentar
1	DC.Title (Titel)			Sagsbetegnelse	Søfarten fra Fur 1850-1970, Etnologisk Undersøgelse 1994	15.3	
2	DC.Creator (Forfatter eller ophav)	Role (Rolle)	Navn	Responsible for work (ansvarshavende museum)	Limfjordsmuseet	Ikke eksisterende felt for Museumsakronym	
				Ansvarlig inspektør	Eske Wohlfahrt	15.9	
3	DC Subject (Emne og nøgleord)			Saglig Registrant for Kulturhistoriske Museer	Søfart	63.3	
				Saglig Registrant for Kulturhistoriske Museer	Leg	63.3	
4	DC.Description (Beskrivelse)			Fritekstbeskrivelse af museumssagens baggrund og indhold	Sagen opstået i forbindelse med bortsalg af det sidste skib hjemmehørende på Fur, hvilket gav anledning til en etnologisk lokalundersøgelse af søfartssamfundet Fur. Undersøgelsen dokumenterer søfartens betydning for øen ca. 1850-1970	15.0	

5	DC.Publisher (udgiver)			Navn på den institution, der stiller museumssagen til rådighed i dens nuværende form angivet. fx Kulturarvsstyrelsen	Kulturarvsstyrelsen	15.???	
				Navn på den institution, der stiller museumssagen til rådighed i dens nuværende form angivet. fx Kulturarvsstyrelsen	Limfjordsmuseet	15.???	
6	DC. Contributor			Bruges pt. ikke af ADS			Benyttes ikke
7	DC.Date (Dato)	Project start		Projekt start	?	?	
		Project end		Projekt slut	?	?	
		Record created		Oprettelse af museumssag	11/03/2003	15.6	
		Record modified		Seneste ændring af museumssag	20/03/2003	15.7	
8	DC.Type (Ressourcetype)				Collection (Museumssag)		ADS benytter UKOLN collections level descriptors which include "Collection" og "Digital Archive". Mon ikke vore museumssager er "Collections"?
9	DC Format			Bruges pt. ikke af ADS			Benyttes ikke
10	DC Identifier (Identifikator)			Akronym for pågældende museum & Journalnummer for pågældende museumssag.	LIM J.nr. 0095	15.2	Da KUAS formidler museumssager fra flere museer, tilføjer KUAS et præfix til Journalnummeret i form af det pågældende museums 3 bogstav akronym for derved at sikre nummerets unicitet

11	DC.Source (kilde el. originalt værk)			ISBN		15??	Benyttes kun hvis man kan angive en ressource hvoraf den aktuelle museumssag er udsprunget (hvilket man i praksis næppe kan for museer i DK)
12	DC.Language (Sprog)			Ikke eksisterende felt der angiver at sproget er dansk, el. grønlandsk el. hebraisk		15.???	ADS benytter 2 bogstavkoder (ISO 639) fx en, ro, pl osv.
13	DC Relation (Relationer)	has part		De registreringsobjekter der er indeholdt i sagen	Model af skonnert Ellen(1879)	22.3	
14	DC.Coverage (Dækning)						
		Spatial	Location (sted)	Place (stednavn under sognet)		74.2	
		Spatial	Location (sted)	Parish (Sogn)	Fur	63.6	
		Spatial	Location (sted)	Unitary (Herred)	Harre Herred		
		Spatial	Location (sted)	County (Amt)	Viborg Amt (før 1970)		
		Spatial	Location (sted)	Country (land)	Danmark		
		Spatial	Coordinate (koordinate r)	UTM reference		74.8	
		Temporal	Period	Periode	1850-nutid	15.4	
		Temporal	Datetime	Datering		82.3	
15	DC.Rights (Rettigheder)			Ikke eksisterende felt der angiver at museumssagen må publiceres på nettet		15.??	