



K U L
T U R
A R V

Fælles præsentation af data fra arkiver, biblioteker og museer på internettet

Rapport fra arbejdsgruppen om ABM-
standarder, februar 2006

Indholdsfortegnelse

1. Indledning.....	3
Perspektiver. Hvad kan det bruges til?.....	4
2. Udvikling af fælles standarder for indhold, formater, transport og repræsentation	5
Dataindhold	5
Dataformater (rammeformater)	6
Datatransport.....	7
Fælles præsentation.....	8
3. Fælles format for indhold.....	9
Grundlæggende problemstillinger	10
Dublin Core generelt.....	10
Administrative Components	11
Generisk model for mapning ABM -> DC.....	12
Funktionelle formål med mapning.....	12
Model for mapning.....	13
Gennemgang af mapningsformat	13
Dublin Core i relation til mapning.....	13
Eksempler på mapning.....	17
Eksempler på mapning af ”Collection-level” post	17
Eksempler på mapning af poster på Item-level	20
Diskussion af sammenligning af data på tværs	20
Format oversigt	22
4. XML Schema for dansk metadata udveksling.....	26
ABM schema.....	26
dkdcplus	28
Teknisk beskrivelse af felter i ABM schema	28
XML schema dkabm	31
XML schema dkdcplus	31
Eksempler på XML formater	33
5. Ordliste	38
6. Referencer.....	39
7. Resume.....	40
Bilag A: Sektorspecifikke databaser	41
Bilag B. Mapning fra Regin	43
Bilag C. Mapning fra Daisy	65
Bilag D. Mapning fra Arkibas4.....	68
Bilag E. Mapning fra danMARC2	75
Bilag F. Vurdering af eksport af data.....	88

1. Indledning

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. Formålet er således at definere en fælles ramme, som gør, at data fra arkiver, biblioteker og museer over for almenheden kan præsenteres i sammenhæng.

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, først og fremmest derved at man kan minimere overvejelserne om dataindhold, dataformater, datatransport og præsentation. Samtidig undgås det, at kommende regionale forsøg udvikles i forskellige retninger til skade for integrationen. Det er således arbejdsgruppens 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 er ved opgaveformuleringen lagt 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.

Arbejdsgruppen afleverede i august 2003 en foreløbig rapport til styregruppen. Denne rapport indeholdt forslag til rammer for standardisering indenfor ABM-området af:

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

Der er siden arbejdet videre med at raffinere håndtering af metadata, dvs. konvertering af sektorspecifikke standarder til Dublin Core samt at overføre det resulterende Dublin Core univers til et sammenhængende XML-schema. Der blev indgået en aftale med Danmarks Biblioteksskole og Index Data om at udføre denne opgave. De afleverede juni 2005 en rapport, som danner grundlag for nærværende rapport afsnit 4 med tilhørende bilag.

Hermed foreligger en samlet beskrivelse af dataindhold og dataformater i form af en overordnet model og udfoldet i konverteringsbeskrivelser og et fælles XML-format. Endvidere nogle overordnede overvejelser om datatransport og fælles præsentation.

Arbejdsgruppen har følgende sammensætning:

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

Perspektiver. Hvad kan det bruges til?

Givet de rette betingelser rummer ABM-samarbejdet nogle relevante muligheder for udvikling af mere publikumsrettede adgange til materialer i arkiver, biblioteker og museer. Det gælder også i relation til søgebaser, hvor der kan gives den almindelige slutbruger en første indgang til materiale, der ligger på ABM institutioner, på tværs af eksisterende systemer og beliggenhed, baseret på internationale standarder og med fokus på let tilgængelighed.

Med det foreliggende forslag til en DC mapning af ABM-formater er der givet et bud på et værktøj, der kan anvendes i en fortsat udvikling og implementering af regionale ABM søgebaser, således som det er sket med www.ArBiMus.dk (NOKS2 og VAKS).

Samtidig med at ABM institutioner og styrelser giver støtte til udvikling af regionale ABM baser, arbejdes der fortsat med nationale A, B og M søgebaser med DAISY, ARKIBAS, Bibliotek.dk, Museernes Samlinger, m.fl.

Udviklingen af de landsdækkende A, B og M baser justerer sig ikke i forhold til en fælles national ABM søgning. Der eksisterer ikke et sæt centralt vedtagne elementer, der skal udgøre kernen i denne ultimative ABM søgning.

Udviklingen af de landsdækkende sektorspecifikke baser er for nærværende ikke rettet mod en fælles national ABM søgning. De regionale ABM søgebaser, der anvender data fra de landsdækkende, vil give et grundlag for at afklare behovet for at udvikle fælles ABM beskrivelsesmetoder. Samtidig er det dog værd at overveje den mulighed, der nu er til stede, for at afprøve en national driftssituation med udvalgte data fra de nationale baser. Dermed mindskes også den dataklargøring, der i dag er nødvendig, når data skal optræde i ABM sammenhæng.

ABM-arbejdsgruppens Dublin Core schema kan således bruges både i praktiske forsøg med regionale baser og udvalgte data fra nationale baser, og dermed lægge op til en mere præcis afklaring af muligheder og perspektiver for ABM søgning på nationalt plan.

Det skal også bemærkes, at regionale baser som f.eks. NOKS ikke er en sammenlægning af alle ABM-data. NOKS er netop defineret emnemæssigt som en kulturhistorisk database. Det hænger bl.a. sammen med, at der fra biblioteksside alene er medtaget lokalhistoriske poster og ikke fag- og skønlitteratur fra andre områder.

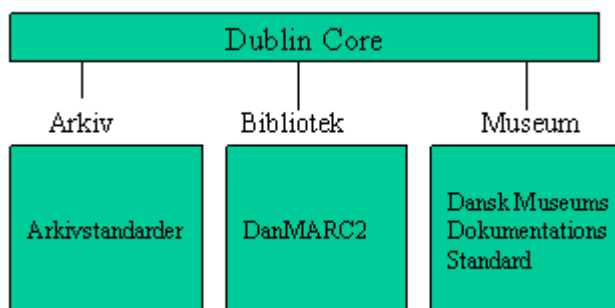
2. Udvikling af fælles standarder for indhold, formater, transport og repræsentation

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 som dels omfatter en specificering i underelementer, dels en kvalificering af elementerne (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 sektorerers 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 sektoreres egne registreringskemaer og på basis af disse selektede 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.

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 var derfor gruppens opfattelse, at det skulle 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 de, 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 er vigtigt, at 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. Da arbejdsgruppen var opmærksom på, at der især knytter sig uløste problemer til beskrivelser af mere komplekse strukturer, indgik dette forhold i den analyse, som Danmarks Biblioteksskole blev bedt om at gennemføre.

Dataformater (rammeformater)

Den generelle model 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.

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.

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.

Derfor er udarbejdet et XML schema for de fælles ABM dataelementer. Formatet kan rumme de fælles definerede data, men er udformet, således at schemaet også kan bruges generelt indenfor arkiver, biblioteker og museer til beslægtede formål. XML schemaet indgår i afsnit 5 i nærværende rapport.

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 web servicen SRW, som bygger videre på Z39.50.
- 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 nuværende tidspunkt vil det være en forkert anvendelse af ressourcerne i arkiver og museer at implementere Z39.50 på trods effektive implementeringer i bibliotekssektoren. Erfaringerne fra

biblioteker viser, at det har kostet relativt lang tid at nå frem til brugbare implementeringer. Derfor peges på SRW som web service model for søgning på tværs af arkiver, biblioteker og museer. 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. - forhold som knytter sig til den enkelte database. Arbejdsgruppen har i denne sammenhæng valgt at behandle selektion og søgning, da der knytter sig nogle generelle aspekter til disse.

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 eksempelvis gøres ved at udarbejde en struktureret ABM emneordliste og ved at etablere fælles standarder for beskrivelse af tid og sted. Men det vil altid være en overvejelse, hvornår at brugere med mere avancerede behov vil være bedre stillet med at bruge de bagved liggende databaser.

Men på den anden side vil en analyse af de forskellige systemer og en kortlægning af udvalgte institutioners reelle brug af samme kunne skabe grundlag for at udarbejde en struktur, samt at der i forlængelse af et sådant arbejde kan skabes grundlag for mere udviklede nationale standarder. Dette ligger dog udenfor nærværende rapport.

3. Fælles format for indhold

Dette afsnit bygger på en rapport om mapning fra ABM sektorenes registreringsformater til Dublin Core med beskrivelse af eventuelle problemområder udarbejdet af Danmarks Biblioteksskole ved konsulenterne Haakon Lund og Henning Grauballe til ABM standard arbejdsgruppen. Konsulenternes forslag er blevet justeret af arbejdsgruppen.

Formålet var at gennemarbejde de i ABM arbejdsgruppens ”*Af rapportering til styregruppen, august 2003*” beskrevne mapninger fra formater i de tre sektorer til Dublin Core med henblik på præsentation i en fælles brugergrænseflade.

Mapning til Dublin Core er foretaget fra følgende formater:

- Arkiver: Arkibas 4 og Daisy
- Biblioteker: DanMARC2
- Museer: Dansk Museums Dokumentations Standard

Mapning fra de 4 ovenstående formaters komplekse indholdsunivers til Dublin Cores simple 15 prædefinerede elementer kan potentielt føre til informationstab.

I 2003-rapportens analyse identificeres det hvilket niveau der menes at være realistisk for dataudtræk fra de 4 formater til Dublin Core. Der tages her udgangspunkt i, at målgruppen, der er defineret som almenheden, ikke vil have samme krav til detaljerighed som ABM institutionernes faglige brugere. Dette medfører, at kun data, som har relevans i forhold til søgning og præsentation for almenheden, skal indgå i den fælles ramme.

Ved identifikationen af niveauet for dataudtræk fra de 4 formater har det tillige været vigtigt at sikre, at eventuelt informationstab i konverteringen ikke vil have meningsforstyrrende karakter.

Danmarks Biblioteksskoles bidrag til projektet omfatter følgende:

1) Mapning af udvalgte data, der skønnes relevante for ”alment interesserede brugere”, i følgende fire standarder for

- Biblioteker (DanMARC)
- Museer (Dansk Museums Dokumentations Standard, dvs. den standard for registrering af løs kulturarv som ligger til grund for registrering i Regin)
- Arkiver (2 standarder, heraf 1 for Statens Arkiver og 1 for de lokalhistoriske arkiver)

2) Identifikation og analyse af problemstillinger vedr. konvertering af data mellem de fire formater og Dublin Core.

Grundlæggende problemstillinger

Ved mapning fra de tre ovennævnte sektorer til et fælles format (Dublin Core) tages der udgangspunkt i registreringsformater, der er specifikke for de enkelte sektorer systemer.

Disse formater er udviklet med vægt på kravene indenfor de enkelte ABM-sektorer inspireret af internationale standarder som f.eks. AACR2 og ISAD(G).

Konsekvensen af dette er, at man i de enkelte sektorer har forskellige traditioner for hvad og hvordan data registreres, hvilket kan forårsage problemer, når disse data skal konverteres til et fælles system.

Disse forskellige registreringstraditioner kommer bl.a. til udtryk i niveauet for registrering. For arkiver og museers vedkommende vil registreringen ofte være på et samlingsniveau (arkiv-fonds eller sager), således at en registrering er et aggregat af enheder. For bibliotekssektorens vedkommende vil en registrering i højere grad være på enhedsniveau, hvor undtagelsen dog er flerpoststruktur, der kunne betragtes som samlingsniveau analogt med arkiver og museer.

Problemet er hvorledes denne hierarkiske struktur, som samlingerne er udtryk for, mappes til Dublin Cores flade struktur, således at det sikres at et minimum af funktionalitet bevares.

Opretholdelse af den originale funktionalitet af data elementer ved konvertering til fælles format kan generelt være problematisk grundet den relative simplicitet i Dublin Core formatet. Der vil således ofte være en konflikt mellem optimering af søgeindgange og optimering af præsentation.

Disse grundlæggende problemstillinger er behandlet efterfølgende med anvisninger på løsningsforslag.

Dublin Core generelt

Dublin Core Metadata Element Set (DC) er et simpelt registreringsformat som grundlæggende består af 15 metadata-elementer:

Title, Creator, Subject, Description, Publisher, Contributor, Date, Type, Format, Identifier, Source, Language, Relation, Coverage, Rights.

Registreringsformatet er oprindeligt udviklet med henblik på ophavsgenererede beskrivelser af dokumentlignende objekter (tekster i internet-regi) med fokus på genfinding (resource discovery). Fokus har således ikke været på f.eks. relevansbedømmelse.

Generelle principper for Dublin Core:

1. Kernen i element-sættet kan udvides efter behovet i konkrete domæner
2. Alle elementer er frivillige
3. Alle elementer kan gentages
4. Elementer kan modificeres/præciseres ved hjælp af ”refinements”
5. Data kan angives at stamme fra standardiserede lister, såkaldte ”schemas”

Administrative Components¹

Dublin Core DCMI Administrative Metadata er udviklet som et værktøj til administration af metadata. Administrative Components (AC) indeholder dataelementer, der er nødvendige for administration og håndtering af poster i DC-metadata, såvel i enkelte baser som i forbindelse med dataudveksling mellem forskellige baser.

De administrative metadata opdeles i tre kategorier:

- A. Data, som knytter sig til den samlede metadata-post
- B. Særlige elementer, der angår transaktioner/ændringsdata
- C. Headerinformation knyttet til batchfil

Dansk elementbetegnelse	DKAC name	Bemærkninger
Gruppe A: Data, som knytter sig til den samlede metadata-post		
Post-identifikator	identifier	
Registrerende entitet	source	Kilde / registrerende enhed
Område	scope	Ukvalificeret tænkes elementet anvendt til angivelse af det bredere formål med at registrere den pågældende ressource, fx. Nationalbibliografi, Netkatsamarbejde mv.
Kommentar	comment	Skal også kunne anvendes i tilknytning til AC.activity (dvs. gruppe B)
Metadata-placering	location	Bruges kun hvis metadata ligger et andet sted
Sprog	language	Bruges ikke pt.
Rettigheder	rights	
Gyldighed	dateRange	
Håndtering	handling	Anvendes til beskeder om fremtidige handlinger vedr. posten
Aktivitet	activity	Anvendes til oplysninger om hvad der er sket med posten
Gruppe B: Særlige elementer, der angår transaktioner/ændringsdata		
Navn	name	Person eller korporation
Email-adresse	email	
Kontakt	contact	Fx. Postadresse, tlf.nr. eller URL
Dato for aktivitet	date	Anvendes til oplysninger om

¹ Nærværende korte introduktion til Dublin Core DCMI Administrative Components er baseret på

<http://www.bs.dk/standards/AdministrativeComponents.htm> og

”DC/AC udvekslingsformatet”. PDF dokument tilgængelig via <http://www.bs.dk> hvor yderligere information kan indhentes.

		hvad der er sket med posten
Ansvarlig institution	affiliation	

Gruppe C: Headerinformation knyttet til batchfil		
Database	database	Navn/kode for base som batchfil skal sendes til
Indsender	transmitter	Navn/kode for person/korporation
Filnavn	filename	Batchfilens navn
Teknisk format	technical format	Fx ISO2709, XML, HTML

Til identifikation af dataleverende system anvendes elementet post-identifikator i AC. Som identifier anvendes et fast ID for den enkelte ressource hentet fra det oprindelige system suppleret med en værdi for registrerende entitet i elementet Source. Ved bibliografiske poster med oprindelse i DANBIB angives DANBIB som registrerende entitet.

Eksempel:

AC.Identifier = [identifier]

AC.Source = [id for entitet]

Generisk model for mapning ABM -> DC

Funktionelle formål med mapning

Et af de oprindelige formål med udvikling af Dublin Core som et Meta-data format var som nævnt at understøtte ”Resource discovery”. Der har således i udviklingen af Dublin Core implementationer i mapningsprojekter ikke været lagt særligt vægt på en understøttelse af de funktionelle krav, der har været lagt til grund for udviklingen af de registreringssystemer / registreringsformater, som data oprindeligt er registreret i følge. Der er lagt vægt på at understøtte identifikationen af eksistensen af et objekt.

Bevæggrunden for dette valg er, at man derved opnår en relativ simplicitet i mapningsformaterne. Dette betyder, at det umiddelbart vil være mere simpelt at udveksle data med et fællessystem uagtet hvilket oprindeligt registreringsformat, der har været anvendt. Konsekvensen af dette valg er, at brugerne af systemet kan få problemer med fortolkningen af de konverterede data, da den semantik, der er iboende det oprindelige registreringsformat, ikke længere er til stede.

For at sikre et vist niveau mht. understøttelse af genfinding i systemet er det vigtigt, at dele af de oprindelige systemers semantik bibeholdes. Brugernes muligheder for ved præsentation af data at kunne bruge disse som baggrund for en vurdering og forståelse af kildens relevans i forhold til brugerens behov vil også i væsentlig grad være afhængig af, at dele af de oprindelige systemers semantik bibeholdes.

Det vil dog være sådan, at en mapning af data fra fire meget forskellige faglige domæner med hver deres tradition for dataregistrering som udgangspunkt vil være inkonsistent. Dette skyldes, at der er tale om en registrering af forskellige typer af artefakter, hvor det kan være svært at finde fælles karakteristika. Man må derfor acceptere en hvis ujævnhed i konverterede data, og dette må understreges overfor brugerne af et kommende system. En måde at indikere dette kan være ved at beholde oplysninger om hvilket system data er leveret fra. En metode til at sikre denne information er ved at anvende AC til håndtering af systemadministrative data, herunder angivelse af dataproducerende system. Det vil hermed være muligt at tage hensyn til domæne specifikke forhold ved etablering af søgeregistre og ved dannelse af præsentationsformater, hvorved det vil være muligt for brugerne at tilpasse søgningen efter dataproducent. Tillige vil det ved præsentation af søgeresultater være muligt at understrege de forskelligheder, der vil / kan være i dataelementernes betydning.

Eksempler:

- Ved angivelse af DC.creator vil der være tale om forskellige typer af ophav. Det kan være ophav til f.eks. en SAG hvilket kan være et Museum. I danMARC-poster vil der være tale om et ophav til det pågældende værk/dokument, der er registreret, og ved Statens Arkiver og Lokalararkiverne vil der være tale om den institution, der har skabt de pågældende arkivalier.
- Ved angivelse af DC.title vil der ligeledes være tale om forskellige typer af titler. I danMARC-poster vil det være en titel på det værk/dokument, der er registreret. I forhold til museer og arkiver er der ingen garanti for, at en titel i traditionel bibliografisk forstand er til stede. Det kan derfor være nødvendigt at generere en titel på baggrund af f.eks. en betegnelse eller beskrivelse.
- Ved angivelse af DC.type og DC.format kan der forventes en del inkonsistens mellem de oprindelige registreringsformater. Det bør overvejes om denne inkonsistens skal elimineres ved udarbejdelse af fælles standardiserede betegnelser.

Generelt vil der være en række DC elementer som f.eks. DC.publisher og DC.language, der vurderes ikke at være relevante for museernes og arkivernes vedkommende. Årsagen til dette skal findes i Dublin Cores oprindelige udspring i det bibliografiske univers. Problemstillingen vurderes dog ikke at frembyde nogen egentlige vanskeligheder.

Model for mapning

I modellen for mapning af data fra de forskellige systemer er defineret to niveauer for beskrivelse: et samlende niveau (Collection) og et niveau for enkelt del (Item). Denne opdeling er valgt, da der i de museale systemer og i arkivsystemerne arbejdes med en sådan skelnen i de respektive registreringssystemer, og det er herved muligt at bibeholde det oprindelige udtryk i registreringerne. Angivelse af niveau findes i DC.Type med anvendelse af DCMI Type Vocabulary.

Gennemgang af mapningsformat

Dublin Core i relation til mapning

Udover de nedenstående forslag til udvidelser af Dublin Core, kan de Elementer, Element Refinements, Schemas m.m. som er beskrevet i DC, anvendes. Samtlige Elements er repeterbare som defineret i DC.

Ved anvendelse af kontrollerede lister i de forskellige systemer bør det angives hvilken standard liste, der er brugt. Det kan være allerede definerede Dublin Core Encoding Schemes, men også systemspecifikke så som ”Den saglige registrant for kulturhistoriske museer” populært betegnet som ”Den grønne registrant”, der anvendes i Regin og DK5, som anvendes i flere systemer. Vi har i nogle få tilfælde angivet hvilke schemas, der er anvendt, men navnekonventionen er ikke endelig, ligesom vi ikke har det fulde overblik over alle tilfælde af anvendte lister. I de tilfælde hvor der er anvendt et allerede kendt Encoding Scheme, anvendes den navngivning, som er defineret jvf. DCMI Metadata Terms eller lokale danske tilføjelser til DC. Der er i nedenstående givet nogle få eksempler på anvendelse af de enkelte elementer.

DC.Title

Anvendes til angivelse af ressourcens titel. En titel kan være givet til en ressource hvilket ofte vil være tilfældet i arkiverne og museerne ved deres registrering af sager / genstande / arkivalier. En titel kan også være uddraget fra ressourcen hvilket er tilfældet ved bibliotekerne.

DC.Creator

Anvendes til angivelse af ophav til dokumenter (bibliografisk ophav) og til angivelse af sagsoprettende / sagsansvarlige institutioner / personer. Det foreslås, at der oprettes følgende element refinements :

- **PreferredName.** Anvendes til angivelse af vedtagen navneform for ophav
- **AlternativeName.** Anvendes til angivelse af andre navnformer for ophav (DAISY)
- **ActPeriod.** Anvendes til angivelse af tid for et ophavs virke. Der angives startår og slutår.

Eksempel

DC.Creator.PreferredName = Abel Catrines stiftelse

DC.Creator.ActPeriod = 1741 - 1949

DC.Subject

Angivelse af emnedata. Hvis der er anvendt et velbeskrevet system, skal dette angives som ”Encoding Scheme”. I de tilfælde hvor der er anvendt et Scheme, som ikke er beskrevet i DC, skal der etableres et fast navn for dette.

Eksempel:

”SRKM” (Den saglige registrant for kulturhistoriske museer).

DC.Description

Anvendes til beskrivelse af ressourcen. Dette kan være i form af abstract, gengivelse af indholdsfortegnelse m.m. Ved registreringer fra Regin vil den tekstuelle beskrivelse af en sag / genstand være angivet her.

DC.Publisher

Udgiver, der er ansvarlig for at stille ressourcen til rådighed, som f.eks. forlag.
Anvendes kun ved bibliografiske ressourcer.

DC.Contributor

Bidragydere, udover primære ophav angivet i DC.Creator, som f.eks. illustratører eller oversættere.
Anvendes kun ved bibliografiske ressourcer.

DC.Date

Anvendes til angivelse af ”udgivelsesdato”.

Eksempler:

Ved genstande i Regin kan der være identificeret et bestemt år / dato for skabelse af genstanden.

DC.Type

Anvendes til angivelse af ressourcens genre. I de tilfælde hvor der ikke findes data i den oprindelige registrering, angives dette jvf. DCMI Type Vocabulary².

Værdien Collection anvendes for alle poster på samlende niveau. Dette af hensyn til præsentation af poster således at den oprindelige sammenhæng i posterne kan bevares.

Eksempel:

Collection anvendes som betegnelse for ressourcestype sag (Regin) og PhysicalObject for ressourcetypen skib (Regin).

Collection anvendes som betegnelse for hovedpost (danMARC2 felt 004 *a hvis værdien h), Text for bindpost (danMARC2 felt 004 *a hvis værdien b).

² <http://dublincore.org/documents/2004/06/14/dcmi-type-vocabulary/>

Collection anvendes som betegnelse for A-fond og beskrivelse på enhedsniveau angives jvf. DCMI Type Vocabulary på baggrund af værdi i [Registrering.RegType] (Arkibas 4).

DC.Format

Anvendes til angivelse af den fysiske udstrækning eller varighed af ressourcen.

Eksempler:

Varighed af en film, antal bind og pakker i et arkiv.

DC.Identifier

Identifikation af den enkelte ressource. Ved Regin-poster angives det ID, som den enkelte sag eller genstand er identificeret med i det oprindelige system, hvis et sådan findes.

DC.Source

Anvendes til angivelse af værk, som ressourcen evt. er baseret på.

Anvendes kun ved bibliografiske ressourcer.

DC.Language

Anvendes til angivelse af sproget for det intellektuelle indhold af dokumentet.

Anvendes kun ved bibliografiske ressourcer.

DC.Relation

hasPart / isPartOf Anvendes til angivelse af relationer til dele i f.eks. en sag eller ved hoved/bindposter i bibliografiske ressourcer. Der henvises ved hjælp af identifikation angivet i AC Identifier.

Dette kan f.eks være fra Sag til genstande indeholdt i sagen (Regin).

DC.Coverage

I refinement Spatial angives en udfoldning af musernes h_Topografinummer, hvor dette 4-cifrede nummer anbefales udfoldet til den fulde tekst, som beskriver sogn, herred og før-1972-amt.

Endvidere anvendes refinement Temporal til at angive tidsperioden som beskrevet i *DCMI Period Encoding Scheme: specification of the limits of a time interval, and methods for encoding this in a text string*, se: <http://dublincore.org/documents/dcmi-period/>

DC.Rights

Anvendes til oplysninger om eventuelle klausuler vedrørende adgang til materialet (Arkibas).

Eksempler på mapning

I nedenstående eksempel vises mapning af eksempelposter fra de fire ABM-systemer. Der er givet eksempler på poster på samlende niveau (Collection) samt på Item niveau.

Sammenhæng mellem Collection og Item niveau skabes ved anvendelse af DC.Relation. Poster fra DAISY vil kun optræde på Collection-level.

Eksempler på mapning af "Collection-level" post

AC - element	Regin
Identifier	[C123]
Source	xxx
DC - element	
Title	Div. porcelæn og bøger
Creator	Århus Bymuseum
Subject	Bordstel og borddækning
Description	Syltetøjsskål, kande og glas, Raadvad knive i etui, kogebog, manuskript til en bog og julekort lavet som et personligt brev. Alle genstandene er fra o. 1850-1900. Derudover var der en postkvitteringsbog til sagen, denne skal sendes til Bornholm, da den stammer fra Rønne. Til sagen var der også 20 bøger og hæfter. Bøgerne var hovedsageligt skønlitterære med nogle enkelte sangbøger og et mindehæfte. Disse er nu opstillet på biblioteket. Størstedelen af bøgerne var også fra 1800'tallet. Der kommer muligvis flere genstande til sagen.
Type	Collection
Identifier	10/00
Relation.hasPart	I123
Coverage.Spatial	2114
Coverage.Temporal Scheme=Period	start=1850 end=1900

AC - element	Daisy
Identifier	[DAISY id]
Source	SA
DC - element	
Title	Justitsprotokol
Title	Tingbog
Creator.PreferredName	Retten i Helsingør
Creator.PreferredName	Helsingør byfoged
Creator.AlternativeName	Dommeren i Helsingør
Creator.AlternativeName	Helsingør ret
Creator.ActPeriod	1549-1919
Subject	Justitsprotokol
Subject	Tingbog
Date	1549-1920
Type	Collection
Format	Original, papir mm.
Format	105 fysiske enheder
Coverage.Spatial	Retten i Helsingør
Coverage.Spatial	Helsingør byfoged
Coverage.Spatial	Dommeren i Helsingør
Coverage.Spatial	Helsingør ret
Coverage.Temporal Scheme=Period	start=1549 end=1920

AC - element	Arkibas
Identifier	[A 699]
Source	851-01
DC - element	
Title	Dansk Arbejdsmands- og Specialarbejder Forbund, Vårst Afdeling
Creator.PreferredName	Dansk Arbejdsmands- og Specialarbejder Forbund, Vårst Afdeling
Creator.ActPeriod	1918-1973
Subject	33.129
Subject	Fagforeninger
Subject	Dansk Arbejdsmands- og Specialarbejder Forbund, Vårst Afdeling
Subject	Vårst og Omegns Landarbejderforening
Subject	F-22-42- Lønmdt.org. uden særligt branchetilhørsforhold
Description	Forhandlingsprtk. Nemlig:

	1: 1918-1931 2: 1931-1938 3: 1938-1964 4: 1965-1973 Kassebog 1963-1972
Type	Collection
Format	5 bind
Identifier	A 699
Coverage.Spatial	Nordjyllands Amt
Coverage.Spatial	Aalborg Kommune
Coverage.Spatial	851-10-0 Gunderup Sogn
Coverage.Temporal Scheme=Period	start=1918 end=1973
Rights	Klausul: Almindelige regler

AC - element	DanMARC2
Identifier	43528076 [ID fra 001*a]
Source	Danbib
DC - element	
Title	The Viking Age buildings of Dublin
Title.Alternative	Medieval Dublin Excavations 1982-81 ; vol. 1
Creator	Patrick F. Wallace
Subject Scheme=DK5	91.155
Subject	vikingetiden
Subject	bygninger
Subject	arkæologi
Publisher	Dublin : Royal Irish Academy
Date	1992
Type	Collection
Type	tekst
Format	papirform
Format	2 bind
Identifier	1-874045-01-1
Language	engelsk
Coverage.Spatial	Irland
Coverage.Spatial	Dublin

Eksempler på mapning af poster på Item-level

AC - element	Regin
Identifier	I123
Source	Museumsnummer
DC - element	
Title	Glas og kande
Subject	Bordstel og borddækning
Description	Glaskande og tilhørende glas. Genstandene er pyntet med rødt og mønster. Genstandene har stået på et natbord og skulle være fra o. 1850.
Type	Physical Object
Identifier	10/00x1
Relation.IsPartOf	C123
Coverage.Temporal Scheme=Period	start=1850 end=1859

AC - element	DanMARC2
Identifier	43528092 [ID fra 001*a]
Source	Danbib
DC - element	
Title	Part 2 : Illustrations
Type	Text
Format	vii, 215 sider
Relation.IsPartOf	43528076 [AC_Identifier på hovedpost]

Diskussion af sammenligning af data på tværs

Der er ovenfor behandlet de særlige forhold og problemer der gør sig gældende i forbindelse med udviklingen af et fælles dataudvekslingsformat mellem de fire forskellige faglige domæner med afsæt i Dublin Core.

Der fremdrages eksempler på inkonsistens i dataregistreringerne, nemlig i Dublin Core elementerne: Creator, Title, Type, Format, Publisher og Language.

Hvad angår Publisher og Language er disse ikke relevante for museernes og arkivernes vedkommende, men dette udgør ikke noget problem, da disse data netop ikke forekommer. Hvad angår Type og Format anvendes forskellig terminologi, men dette udgør næppe heller noget problem, da det blot for det konkrete materiale afspejler forskellighederne.

Uegaliteten, for så vidt angår Title, udspringer af det forhold at biblioteksmaterialer og dermed biblioteksposter er født med en titel, mens det samme ikke nødvendigvis gør sig gældende for arkivalier og museumsgenstande. Titler kan være genereret af de pågældende institutioner selv. Det er i denne sammenhæng værd at bemærke, at der er tale om forskellige typer af Poster. I Regin repræsenterer posten en museumssag, skabt af et museum, og altså ikke selve museumsgenstanden, hvilket kommer til udtryk i titlen, f.eks. *Div. porcelæn og bøger*. I Arkibas repræsenterer posten en arkivfond, svarende til et ophavs samlede arkivdannelse, hvorfor Title og Creator typisk vil være identisk. I Daisy repræsenterer posten en arkivserie, dvs. en klart afgrænset delmængde af et ophavs arkivdannelse, hvor titlen f.eks. kan være *Justitsprotokol*.

Det forhold, at posterne er forskelligartede, slår også igennem i forhold til Creator. Der er i denne sammenhæng ikke stor forskel på biblioteker ("forfatter") og arkiver (arkivskaber), anderledes forholder det sig med museerne, hvor Creator er det pågældende museum selv.

Af andre eksempler på inkonsistens, som kan fremdrages, er Subject og Coverage.

Mens poster i Regin, Arkibas og DanMARC 2 er indekseret med emneord og klassifikationskode, gør noget tilsvarende sig ikke gældende for Daisy endnu, hvorfor arkivseriens titel foreslås anvendt som subject.

Hvad angår Coverage.Spatial, gør det samme forhold sig gældende, dog foreslås arkivseriens titel for poster i Daisy anvendt i stedet for.

Hvad angår Coverage.Temporal, anvendes disse konsekvent i Regin, Daisy og Arkibas, mens noget tilsvarende ikke gør sig gældende i DanMARC 2.

Ovennævnte inkonsistens i dataregistreringerne burde ikke give anledning til større problemer i forhold til præsentation af data, især ikke hvis kilden til data (ophavssystem) vises. Anderledes forholder det sig dog ved søgning, hvor forskellene mellem de fire faglige domæners registreringsformater bliver mere mærkbar. Dette gælder både for almindelig fritekstsøgning og i endnu højere grad for struktureret søgning, hvor søgeresultaterne tydeligt vil afspejle den overfor anførte inkonsistens i dataregistreringerne. Det fremgår af tidligere afsnit, at anførelse af det dataproducerende system muliggør at tage hensyn til domæne specifikke forhold ved etablering af søgeregistre og ved dannelse af præsentationsformater, hvorved det vil være muligt for brugerne at tilpasse søgningen efter dataproducent. Ved præsentation af søgeresultater vil det være muligt at understrege de forskelligheder, der vil / kan være i dataelementernes betydning.

Denne mulighed foreligger, men herved begrænses også brugernes muligheder for at foretage søgning på tværs af de fire forskellige domæner i en og samme søgning, eller også risikerer man, at søgeresultaterne omfatter enten for lidt eller for meget fra enkelte af de fire forskellige domæner. Der ligger i håndteringen heraf en klar udfordring ved udformning af brugergrænseflader. Derfor anbefales, at der for hvert enkelt projekt med søgning på tværs i ABM-systemers data vurderes, om det i det konkrete tilfælde er formålstjenligt at etablere en fælles søge- og præsentationsbrugergrænseflade.

Format oversigt

Der er i denne oversigt medtaget alle de p.t. mulige felter for Dublin Core (dc), udvidet Dublin Core (dcterms), Administrative Components (ac) og den danske udvidelse for ABM-området (dkdplus). Der er anført "Anvendes ikke for nærværende ved udveksling af ABM-data", hvor dette ikke er beskrevet som tilfældet i denne rapport.

Det bemærkes, at oversigten er opbygget logisk og ikke afspejler det tekniske forhold, at f.eks. dcterms elementet *Alternative* rent teknisk "står alene" og ikke er pakket ind i dc elementet *Title*.

Felt	Refinement	Beskrivelse
DC.Title		Anvendes til angivelse af ressourcens titel. En titel kan være givet til en ressource, hvilket ofte vil være tilfældet i arkiverne og museerne ved deres registrering af sager / genstande / arkivalier. En titel kan også være uddraget fra ressourcen, hvilket er tilfældet ved bibliotekerne.
	Alternative	Anvendes til angivelse af alternativ titel.
DC.Creator		Anvendes til angivelse af ophav til dokumenter (bibliografisk ophav) og til angivelse af sagsoprettende / sagsansvarlige institutioner / personer.
	PreferredName	Anvendes til angivelse af vedtagen navneform for ophav
	AlternativeName	Anvendes til angivelse af andre navnformer for ophav (DAISY)
	ActPeriod med scheme Period	Anvendes til angivelse af tid for et ophavs virke. Der angives startår og slutår.
DC.Subject		Angivelse af emnedata. Hvis der er anvendt et velbeskrevet system, skal dette angives som "Scheme". I de tilfælde hvor der er anvendt et Scheme, som ikke er beskrevet i DC, skal der etableres et fast navn for dette.
DC.Description		Anvendes til beskrivelse af ressourcen. Dette kan være i form af abstract, gengivelse af indholdsfortegnelse m.m. Ved registreringer fra Regin vil den tekstuelle beskrivelse af en sag / genstand være angivet her.
	tableofContents	Anvendes ikke for nærværende ved udveksling af ABM-data

	abstract	Anvendes ikke for nærværende ved udveksling af ABM-data
	version	Version eller udgave
DC.Publisher		Udgiver som er ansvarlig for at stille ressourcen til rådighed, som f.eks. forlag. Anvendes kun ved bibliografiske ressourcer.
DC.Contributor		Bidragydere, udover primære ophav angivet i DC.Creator, som f.eks. illustratører eller oversættere. Anvendes kun ved bibliografiske ressourcer.
DC.Date		Anvendes til angivelse af ”udgivelsesdato”. F.eks. ved genstande i museer kan der være identificeret et bestemt år/dato for skabelse af genstanden.
	created	Anvendes ikke for nærværende ved udveksling af ABM-data
	valid	Anvendes ikke for nærværende ved udveksling af ABM-data
	available	Anvendes ikke for nærværende ved udveksling af ABM-data
	issued	Anvendes ikke for nærværende ved udveksling af ABM-data
	modified	Anvendes ikke for nærværende ved udveksling af ABM-data
	dateAccepted	Anvendes ikke for nærværende ved udveksling af ABM-data
	dateCopyrighted	Anvendes ikke for nærværende ved udveksling af ABM-data
	dateSubmitted	Anvendes ikke for nærværende ved udveksling af ABM-data
DC.Type		Anvendes til angivelse af ressourcens genre. I de tilfælde hvor der ikke findes data i den oprindelige registrering, angives dette jvf. DCMI Type Vocabulary 2. Værdien Collection anvendes for alle poster på samlende niveau. Dette af hensyn til præsentation af poster således at den oprindelige sammenhæng i posterne kan bevares.

	Anvendte værdier i Encoding Scheme dcmi-type: Collection PhysicalObject Text	Se i øvrigt http://dublincore.org/documents/dcmi-type-vocabulary/ for flere typer
DC.Format		Anvendes til angivelse af den fysiske udstrækning eller varighed af ressourcen.
	Extent	Anvendes ikke for nærværende ved udveksling af ABM-data
	Medium	Anvendes ikke for nærværende ved udveksling af ABM-data
DC.Identifier		Identifikation af den enkelte ressource.
DC.Source		Anvendes til angivelse af værk, som ressourcen evt. er baseret på. Anvendes kun ved bibliografiske ressourcer.
DC.Language		Anvendes til angivelse af sproget for det intellektuelle indhold af dokumentet. Anvendes kun ved bibliografiske ressourcer.
DC.Relation		hasPart / isPartOf Anvendes til angivelse af relationer til dele i f.eks. en sag eller ved hoved/bind-poster i bibliografiske ressourcer. Der henvises ved hjælp af identifikation angivet i AC Identifier.
	isVersionOf	Anvendes ikke for nærværende ved udveksling af ABM-data
	hasVersion	Anvendes ikke for nærværende ved udveksling af ABM-data
	isReplacedBy	Anvendes ikke for nærværende ved udveksling af ABM-data
	replaces	Anvendes ikke for nærværende ved udveksling af ABM-data
	isRequiredBy	Anvendes ikke for nærværende ved udveksling af ABM-data
	requires	Anvendes ikke for nærværende ved udveksling af ABM-data
	isPartOf	Anvendes til angivelse af relationer til dele i f.eks. en sag eller ved hoved/bind-poster i bibliografiske ressourcer
	hasPart	Anvendes til angivelse af relationer til dele i f.eks. en sag eller ved hoved/bind-poster i bibliografiske ressourcer

	isReferencedBy	Anvendes til angivelse af relationer til dele i f.eks. en sag eller ved hoved/bind-poster i bibliografiske ressourcer
	references	Anvendes til angivelse af relationer til dele i f.eks. en sag eller ved hoved/bind-poster i bibliografiske ressourcer
	isFormatOf	Anvendes ikke for nærværende ved udveksling af ABM-data
	hasFormat	Anvendes ikke for nærværende ved udveksling af ABM-data
	conformsTo	Anvendes ikke for nærværende ved udveksling af ABM-data
DC.Coverage		Anvendes til angivelse af materialets emnemæssige dækning mht. til tidsperiode og sted.
	Spatial	Anvendes til angivelse af materialets emnemæssige dækning mht. til sted.
	Temporal med scheme Period	tekstlig beskrivelse for periode med mulighed for start og slut
DC.Rights		Anvendes til oplysninger om eventuelle klausuler vedrørende adgang til materialet.
AC.Identifier		Post-identifikator
AC.Source		Registrerende entitet
AC.Scope		Anvendes ikke for nærværende ved udveksling af ABM-data
AC.Comment		Anvendes ikke for nærværende ved udveksling af ABM-data
AC.Location		Anvendes ikke for nærværende ved udveksling af ABM-data
AC.Language		Anvendes ikke for nærværende ved udveksling af ABM-data
AC.Rights		Anvendes ikke for nærværende ved udveksling af ABM-data
AC.DateRange		Anvendes ikke for nærværende ved udveksling af ABM-data
AC.Handling	Action Activity Name Email Contact Date Affiliation	Anvendes ikke for nærværende ved udveksling af ABM-data
AC.Database		Anvendes ikke for nærværende ved udveksling af ABM-data

AC.Transmitter		Anvendes ikke for nærværende ved udveksling af ABM-data
AC.FileName		Anvendes ikke for nærværende ved udveksling af ABM-data
AC.Technicalformat		Anvendes ikke for nærværende ved udveksling af ABM-data

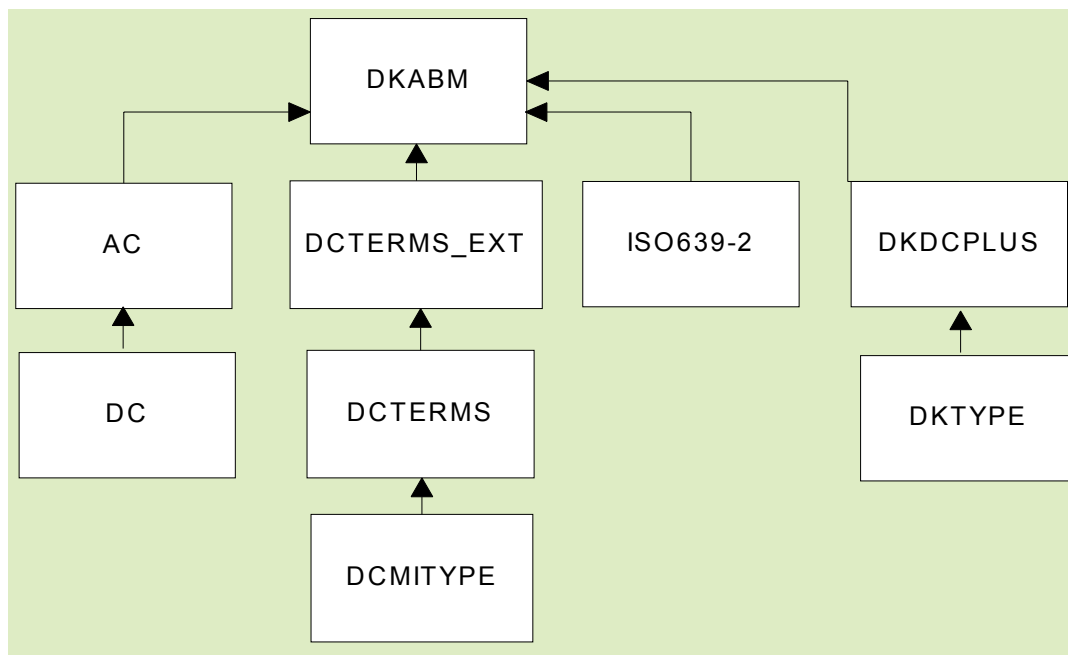
4. XML Schema for dansk metadata udveksling

ABM schema

Der er defineret et XML schema, der beskriver det fælles ABM-udvekslingsformat, som hedder *dkabm*. Dette schema er bygget op ved at tage dels internationale og anerkendte schemas og dels schemas, der er konstrueret til danske forhold og inkludere dem i et fælles schemas. Det drejer sig om følgende:

dcabm.xsd (http://www.bs.dk/standards/schemas/dkabm_2006-01-13.xsd) – som samler de nedenstående schemaer.

- **dc.xsd** (<http://purl.org/dc/elements/1.1/dc.xsd>) - definerer de 15 standard Dublin Core elementer, som er gennemgået ovenfor.
- **ac.xsd** (http://www.bs.dk/standards/schemas/ac_2005-09-01.xsd) - definerer Dublin Core Administrative Components, som er blevet gennemgået ovenfor.
- **dkdcplus.xsd** (http://www.bs.dk/standards/schemas/dkdcplus_2006-01-13.xsd)- indeholder de forskellige emneordslister til brug i f.eks. dc.subject.
- **dcterms_ext.xsd** (http://www.bs.dk/standards/schemas/dcterms_ext_2006-01-13.xsd) – som indeholder ekstra element, der bruges i dkdcplus.xsd
- **dcterms.xsd** (<http://purl.org/dc/terms/dcterms.xsd>) - indeholder ekstra Dublin Core elementer, refinements og typer. F.eks. er det her, at ”temporal” og ”spatial” defineres for dc:coverage.
- **dcmitype.xsd** (<http://purl.org/dc/dcmitype/dcmitype.xsd>) - beskriver en Dublin Core type indeholdende liste over ressource typer
- **dktype.xsd** (http://www.bs.dk/standards/schemas/dktype_2003-05-01.xsd) - danske udvidelser til ressource typer
- **ISO639-2.xsd** (<http://www.ddb.ds/standards/xmetadiss/ISO639-2.xsd>) – definerer gyldige sprogkoder.



Strukturen af schemas er som følger: dkabm er hovedskemaet og heri inkluderes dc, dcterms_ext, dcterms, dcmitype, ac, ISO639-2, dktype og dkdcplus.

DKABM Schema er udformet, således at det kan fungere som et fælles schema for arkiv-, biblioteks- og kulturarvssektoren beregnet for data til udveksling med andre systemer. Derfor er ikke indbygget det ellers beskrevne check af tilstedeværelse af to felter, nemlig titel (DC.Title) og postidentifikator (AC.Identifier) som en del af det tekniske XML schema, men overlades til de enkelte applikationer at håndtere.

Hensigten er ikke at lave en løsning, som alene kunne anvendes i én model for ABM samarbejde, men en løsning, som kan anvendes mere generelt indenfor ABM-sektorerne ved udveksling af data på generelt niveau (dvs. når det ikke sker i de sektorspecifikke schemaer). Derfor er defineret schemaet dkdcplus ikke blot for at rumme udvidelserne nødvendige for denne opgave, men som også dynamisk kan rumme andre fælles dataelementer.

Der skal etableres en model for vedligeholdelse af dkdcplus, således at det kan fungere dynamisk. Dvs. både være en samlet ressource for identifikation og beskrivelse af elementer for fælles beskrivelse af data til eksport til andre systemer og samtidig løbende kunne opdateres ud fra nye behov.

Det er vores vurdering, at der ikke er uoverkommelige problemer forbundet med konverteringen af nogle af de fire her omtalte dataformater til ABM-formatet. Det vil dog kræve en vis arbejdsindsats at udarbejde de nødvendige programmer, dels fordi dataene er så forskellige, og dels fordi der skal en del databerigelse af dataene til, for at de egner sig til brug i ABM-sammenhæng.

Referencer til indhold:

DC, dcterms m.v. - Dublin Core er dokumenteret i *DCMI Metadata Terms*,

se: <http://dublincore.org/documents/dcmi-terms/>

Administrative Components er dokumenteret i *AC - Administrative Components*,

se: <http://www.bs.dk/standards/AdministrativeComponents.htm>

Dansk udvekslingsformat for metadata. DC/AC formatet,

se: <http://www.bs.dk/publikationer/andre/dcac/index.htm>

dkdcplus

Udvidelsen til Dublin Core *dkdcplus* for danske forhold indeholder for nærværende:

DC.Creator	PreferredName	Anvendes til angivelse af vedtagen navneform for ophav
DC.Creator	AlternativeName	Anvendes til angivelse af andre navnformer for ophav
DC.Creator	ActPeriod med scheme Period	Anvendes til angivelse af tid for et ophavs virke. Der angives startår og slutår.
DC.Description	Version	Version, udgave m.v.

Til DC.Subject knyttes liste SubjectType over mulige værdier:

DK5	DK5-klassemærke
DBCS	DBC's emneordssystem: skønlitterært emneord
DBCF	DBC's emneordssystem: faglitterært emneord
DBCM	DBC's emneordssystem: musik emneord
SRKM	Den saglige registrant for kulturhistoriske museer (også benævnt: den grønne registrant) http://www.kulturarv.dk/forvaltning/museum/sdrift/vejledninger/registrering/kulturhistorie/saglig_registrant.xls

Denne liste skal suppleres med evt. andre klassifikationssystemer fra arkiver og museer, inden arbejdet med ABM schemaet er afsluttet.

Teknisk beskrivelse af felter i ABM schema

De lovlige felter i ABM schema er listet i nedenstående oversigt.

Det bemærkes, at oversigten er opbygget logisk og ikke afspejler det tekniske forhold, at f.eks. dcterms elementet *Alternative* rent teknisk "står alene" og ikke er pakket ind i dc elementet *Title*.

Element	Refinement	Namespace
DC.Title		dc
	alternative	dcterms
DC.Creator		dc
	preferredName	dkdcplus
	alternativeName	dkdcplus
	ActPeriod med scheme Period (Period fra dcterms)	dkdcplus
DC.Subject	Lister over values SubjectType med værdierne DK5, DBCS, DBCF, DBCM og SRKM i dkdcplus.	dc
DC.Description		dc
	tableofContents	dcterms
	abstract	dcterms
	version	dkdcplus
DC.Publisher		dc
DC.Contributor		dc
DC.Date		dc
	created	dcterms
	valid	dcterms
	available	dcterms
	issued	dcterms
	modified	dcterms
	dateAccepted	dcterms
	dateCopyrighted	dcterms
	dateSubmitted	dcterms
DC.Type	Lister over values fra dcmi-type og dansk version i dkdcplus	dc
DC.Format		dc
	Extent	dcterms
	Medium	dcterms
DC.Identifier		dc
DC.Source		dc
DC.Language		dc
DC.Relation		dc
	isVersionOf	dcterms
	hasVersion	dcterms
	isReplacedBy	dcterms
	replaces	dcterms
	isRequiredBy	dcterms
	requires	dcterms

	isPartOf	dcterms
	hasPart	dcterms
	isReferencedBy	dcterms
	references	dcterms
	isFormatOf	dcterms
	hasFormat	dcterms
	conformsTo	dcterms
DC.Coverage		dc
	Spatial	dcerms
	Temporal med scheme Period	dcterms (scheme Period fra dcterms)
DC.Rights		dc
AC.Identifier		ac
AC.Source		ac
AC.Scope		ac
AC.Comment		ac
AC.Location		ac
AC.Language		ac
AC.Rights		ac
AC.DateRange		ac
AC.Handling	Action Activity Name Email Contact Date Affiliation	ac
AC.Database		ac
AC.Transmitter		ac
AC.Filename		ac
AC.Technicalformat		ac
AC.characterSet		ac
AC.bibliographicFormat		ac
AC.resultFile		ac

XML schema dkabm

```
<?xml version="1.0" encoding="UTF-8" ?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"
xmlns:dkabm="http://www.bs.dk/standards/#dkabm" xmlns:ISO639-
2="http://lcweb.loc.gov/standards/iso639-2/" xmlns:ac="http://www.bs.dk/standards/#ac"
xmlns:dkdcplus="http://www.bs.dk/standards/#dkdcplus" xmlns:dcterms="http://purl.org/dc/terms/"
targetNamespace="http://www.bs.dk/standards/#dkabm" elementFormDefault="qualified"
attributeFormDefault="unqualified">
<xs:annotation>
<xs:documentation>XML Schema for http://www.bs.dk/standards/#dkabm namespace This XML schema
shall support the Archives-Libraries-Museum standard working group Dublin Core based format for common
presentation of records from archives, libraries and museums. Created in accordance with the naming
convention used in the Dublin Core Metadata Initiative (DCMI) Created 2005-09-09 by Per Moerkegaard
Hansen (perhans@indexdata.dk) with assistance from Soeren Pedersen
(sorenpedersen@makeitright.dk)</xs:documentation>
</xs:annotation>
<xs:import namespace="http://purl.org/dc/terms/"
schemaLocation="dcterms_ext_2006-01-13.xsd" />
<xs:import namespace="http://www.bs.dk/standards/#ac"
schemaLocation="ac_2005-09-01.xsd" />
<xs:import namespace="http://www.bs.dk/standards/#dkdcplus"
schemaLocation="dkdcplus_2006-01-13.xsd" />
<xs:import namespace="http://lcweb.loc.gov/standards/iso639-2/"
schemaLocation="http://www.ddb.de/standards/xmetadiss/iso639-2.xsd" />
<xs:element name="metadata" type="dcterms:elementOrRefinementContainer" >
</xs:schema>
```

XML schema dkdcplus

```
<?xml version="1.0" encoding="UTF-8" ?>
- <xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"
targetNamespace="http://www.bs.dk/standards/#dkdcplus"
xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:dkdcplus="http://www.bs.dk/standards/#dkdcplus"
xmlns:dcterms="http://purl.org/dc/terms/"
xmlns:dktype="http://www.bs.dk/standards/#dktype" elementFormDefault="qualified"
attributeFormDefault="unqualified">
- <xs:annotation>
<xs:documentation xml:lang="en">XML Schema for http://www.bs.dk/standards/#dkdcplus namespace
Created 2005-11-13 by Per M. Hansen, Index Data.</xs:documentation>
</xs:annotation>
```

```

<xs:import namespace="http://www.w3.org/XML/1998/namespace"
  schemaLocation="http://www.w3.org/2001/03/xml.xsd" />
<xs:import namespace="http://purl.org/dc/elements/1.1/"
  schemaLocation="http://dublincore.org/schemas/xmls/qdc/2003/04/02/dc.xsd" />
<xs:import namespace="http://purl.org/dc/terms/"
  schemaLocation="dcterms_ext_2006-01-13.xsd" />
<xs:import namespace="http://www.bs.dk/standards/#dktype"
  schemaLocation="dktype_2003-05-01.xsd" />
<xs:element name="preferredName" substitutionGroup="dc:creator" />
<xs:element name="alternativeName" substitutionGroup="dc:creator" />
<xs:element name="actPeriod" substitutionGroup="dc:creator" />
<xs:element name="version" substitutionGroup="dc:description" />
- <xs:complexType name="DKType">
- <xs:simpleContent>
  - <xs:restriction base="dc:SimpleLiteral">
    - <xs:simpleType>
      <xs:restriction base="xs:string" />
    </xs:simpleType>
    <xs:attribute ref="xml:lang" use="prohibited" />
  </xs:restriction>
</xs:simpleContent>
</xs:complexType>
- <xs:complexType name="DK5">
- <xs:simpleContent>
  - <xs:restriction base="dc:SimpleLiteral">
    - <xs:simpleType>
      <xs:restriction base="xs:string" />
    </xs:simpleType>
    <xs:attribute ref="xml:lang" use="prohibited" />
  </xs:restriction>
</xs:simpleContent>
</xs:complexType>
- <xs:complexType name="DBCS">
- <xs:simpleContent>
  - <xs:restriction base="dc:SimpleLiteral">
    - <xs:simpleType>
      <xs:restriction base="xs:string" />
    </xs:simpleType>
    <xs:attribute ref="xml:lang" use="prohibited" />
  </xs:restriction>
</xs:simpleContent>
</xs:complexType>
- <xs:complexType name="DBCF">
- <xs:simpleContent>

```

```

- <xs:restriction base="dc:SimpleLiteral">
  - <xs:simpleType>
    <xs:restriction base="xs:string" />
  </xs:simpleType>
  <xs:attribute ref="xml:lang" use="prohibited" />
</xs:restriction>
</xs:simpleContent>
</xs:complexType>
- <xs:complexType name="DBCM">
  - <xs:simpleContent>
    - <xs:restriction base="dc:SimpleLiteral">
      - <xs:simpleType>
        <xs:restriction base="xs:string" />
      </xs:simpleType>
      <xs:attribute ref="xml:lang" use="prohibited" />
    </xs:restriction>
  </xs:simpleContent>
</xs:complexType>
- <xs:complexType name="SKRM">
  - <xs:simpleContent>
    - <xs:restriction base="dc:SimpleLiteral">
      - <xs:simpleType>
        <xs:restriction base="xs:string" />
      </xs:simpleType>
      <xs:attribute ref="xml:lang" use="prohibited" />
    </xs:restriction>
  </xs:simpleContent>
</xs:complexType>
</xs:schema>

```

Eksempler på XML formater

Eksempel XML format for sag (Regin)

```

<?xml version="1.0" encoding="ISO-8859-1" ?>
- <metadata
  xmlns="http://www.bs.dk/standards/#dkabm"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:ac="http://www.bs.dk/standards/#ac"
  xmlns:dkdcplus="http://www.bs.dk/standards/#dkdcplus"
  xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"

```

```

xsi:schemaLocation="http://www.bs.dk/standards/#dkabm dkabm_2006_01_13.xsd">
<ac:identifier>[C123]</ac:identifier>
<ac:source>xxx</ac:source>
<dc:title>Div. porcelæn og bøger</dc:title>
<dc:creator>Århus Bymuseum</dc:creator>
<dc:subject>Bordstel og borddækning</dc:subject>
<dc:description>Syltetøjsskål, kande og glas, Raadvad knive i etui, kogebog, manuskript til en bog og
julekort lavet som et personligt brev. Alle genstandene er fra o. 1850-1900. Derudover var der en
postkvitteringsbog til sagen, denne skal sendes til Bornholm, da den stammer fra Rønne. Til sagen var
der også 20 bøger og hæfter. Bøgerne var hovedsageligt skønlitterære med nogle enkelte sangbøger
og et mindehæfte. Disse er nu opstillet på biblioteket. Størstedelen af bøgerne var også fra
1800 ´tallet. Der kommer muligvis flere genstande til sagen.</dc:description>
<dc:type>Collection</dc:type>
<dc:identifier>10/00</dc:identifier>
<dcterms:hasPart>I123</dcterms:hasPart>
<dcterms:spatial>2114</dcterms:spatial>
<dcterms:temporal>start=1850 end=1900</dcterms:temporal>
</metadata>

```

Eksempel XML format for sag (Daisy)

```

<?xml version="1.0" encoding="ISO-8859-1" ?>
- <metadata
  xmlns="http://www.bs.dk/standards/#dkabm"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:ac="http://www.bs.dk/standards/#ac"
  xmlns:dkdcplus="http://www.bs.dk/standards/#dkdcplus"
  xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.bs.dk/standards/#dkabm dkabm_2006_01_13.xsd">
<ac:identifier>[DAISY id]</dc:identifier>
<ac:source>SA</dc:source>
<dc:title>Justitsprotokol</dc:title>
<dc:title>Tingbog</dc:title>
<dkdcplus:preferredName>Retten i Helsingør</dkdcplus:preferredName>
<dkdcplus:preferredName>Helsingør byfoged</dkdcplus:preferredName>
<dkdcplus:alternativeName>Dommeren i Helsingør</dkdcplus:alternativeName>
<dkdcplus:alternativeName>Helsingør ret</dkdcplus:alternativeName>
<dkdcplus:actPeriod>1549-1919</dkdcplus:actPeriod>
<dc:subject>Justitsprotokol</dc:subject>
<dc:subject>Tingbog</dc:subject>
<dc:date>1549-1920</dc:date>
<dc:type>Collection</dc:type>
<dc:format>Original, papir mm.</dc:format>

```

```

<dc:format>105 fysiske enheder</dc:format>
<dcterms:spatial>Retten i Helsingør</dcterms:spatial>
<dcterms:spatial>Helsingør byfoged</dcterms:spatial>
<dcterms:spatial>Dommeren i Helsingør</dcterms:spatial>
<dcterms:spatial>Helsingør ret</dcterms:spatial>
<dcterms:temporal>start=1549; end=1920</dcterms:temporal>
</metadata>

```

Eksempel XML format for sag (Arkibas)

```

<?xml version="1.0" encoding="ISO-8859-1" ?>
- <metadata
  xmlns="http://www.bs.dk/standards/#dkabm"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:ac="http://www.bs.dk/standards/#ac"
  xmlns:dkdcplus="http://www.bs.dk/standards/#dkdcplus"
  xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.bs.dk/standards/#dkabm dkabm_2006_01_13.xsd">
<ac:identifier> [A 699]</dc:identifier>
<ac:source>851-01</dc:source>
<dc:title>Dansk Arbejdsmands- og Specialarbejder Forbund, Vårst Afdeling</dc:title>
<dc:title></dc:title>
<dkdcplus:preferredName> Dansk Arbejdsmands- og Specialarbejder Forbund, Vårst
  Afdeling</dkdcplus:preferredName>
<dkdcplus:actPeriod>1918-1973</dkdcplus:actPeriod>
<dc:subject>33.129</dc:subject>
<dc:subject>Fagforeninger</dc:subject>
<dc:subject>Dansk Arbejdsmands- og Specialarbejder Forbund, Vårst Afdeling</dc:subject>
<dc:subject>Vårst og Omegns Landarbejderforening</dc:subject>
<dc:subject>F-22-42- Lønmdt.org. uden særligt branche- tilhørsforhold</dc:subject>
<dc:description>Forhandlingsprtk. Nemlig:
  1: 1918-1931
  2: 1931-1938
  3: 1938-1964
  4: 1965-1973
  Kassebog 1963-1972</dc:description>
<dc:type>Collection</dc:type>
<dc:format>5 bind</dc:format>
<dc:identifier>A 699</dc:identifier>
<dcterms:spatial>Nordjyllands Amt</dcterms:spatial>
<dcterms:spatial>Aalborg Kommune </dcterms:spatial>
<dcterms:spatial>851-10-0 Gunderup Sogn</dcterms:spatial>
<dcterms:temporal>start= 1918; end= 1973</dcterms:temporal>

```

```
<dc:rights> Klausul: Almindelige regler</dc: rights>
</metadata>
```

Eksempel XML format for genstand (Regin)

```
<?xml version="1.0" encoding="ISO-8859-1" ?>
- <metadata
  xmlns="http://www.bs.dk/standards/#dkabm"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:ac="http://www.bs.dk/standards/#ac"
  xmlns:dkdcplus="http://www.bs.dk/standards/#dkdcplus"
  xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.bs.dk/standards/#dkabm dkabm_2006_01_13.xsd">
  <ac:identifier>I123</ac:identifier>
  <ac:source>Museumsnummer</ac:source>
  <dc:title>Glas og kande</dc:title>
  <dc:subject>Bordstel og borddækning</dc:subject>
  <dc:description>Glaskande og tilhørende glas. Genstandene er pyntet med rødt og mønster.
  Genstandene har stået på et natbord og skulle være fra o. 1850.</dc:description>
  <dc:type>Physical Object</dc:type>
  <dc:identifier>10/00x1</dc:identifier>
  <dcterms:isPartOf>C123</dcterms:isPartOf>
  <dcterms:temporal>start=1850 end=1859</dcterms:temporal>
</metadata>
```

Eksempel XML format for enkeltpost (biblioteksbase)

```
<?xml version="1.0" encoding="ISO-8859-1" ?>
- <metadata
  xmlns="http://www.bs.dk/standards/#dkabm"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:ac="http://www.bs.dk/standards/#ac"
  xmlns:dkdcplus="http://www.bs.dk/standards/#dkdcplus"
  xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.bs.dk/standards/#dkabm dkabm_2006_01_13.xsd">
  <ac:identifier>43528076</ac:identifier>
  <ac:source>Danbib</ac:source>
  <dc:title>The Viking Age buildings of Dublin</dc:title>
  <dcterms:alternative>Medieval Dublin Excavations 1982-81 ; vol. 1</dcterms:alternative>
  <dc:creator>Patrick F. Wallace</dc:creator>
  <dc:subject xsi:type="dkdcplus:DK5">>91.155</dc:subject>
  <dc:subject>vikingetiden</dc:subject>
```

```

<dc:subject>bygninger</dc:subject>
<dc:subject>arkæologi</dc:subject>
<dc:publisher>Dublin : Royal Irish Academy</dc:publisher>
<dc:date>1992</dc:date>
<dc:type>Collection</dc:type>
<dc:type>tekst</dc:type>
<dc:format>papirform</dc:format>
<dc:format>2 bind</dc:format>
<dc:identifier>1-874045-01-1</dc:identifier>
<dc:language>engelsk</dc:language>
<dcterms:spatial>Irland</dcterms:spatial>
<dcterms:spatial>Dublin</dcterms:spatial>
</metadata>

```

Eksempel XML format for bindpost (biblioteksbase)

```

<?xml version="1.0" encoding="ISO-8859-1" ?>
- <metadata
  xmlns="http://www.bs.dk/standards/#dkabm"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:ac="http://www.bs.dk/standards/#ac"
  xmlns:dkdcplus="http://www.bs.dk/standards/#dkdcplus"
  xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.bs.dk/standards/#dkabm dkabm_2006_01_13.xsd">
  <ac:identifier>43528092</ac:identifier>
  <ac:source>Danbib</ac:source>
  <dc:title>Part 2 : Illustrations</dc:title>
  <dc:type>Text</dc:type>
  <dc:format>vii, 215 sider</dc:format>
  <dcterms:isPartOf>43528076</dcterms:isPartOf>
</metadata>

```

5. 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. Oprindeligt 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.

Schema eller scheme: En systematisk kombination af elementer eller termer. *XML Schemas* er maskinlæsbare specifikationer, som definerer struktur og syntaks af metadataspecifikationer. Et Encoding Scheme er et sæt af regler for kodning af information – enten som f.eks. et klassifikationssystem jf. DK-5 eller en konkret list, jf. DCMI Type.
Se også: <http://dublincore.org/documents/usageguide/glossary.shtml#S>

SRW: (Search and Retrieve Web service) For at tilnærme Z39.50 til WWW-miljøet er udviklet en simple 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 er formelt registreret af den amerikanske standardiseringsorganisation NISO.

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 på skolelærer, skolevæsen etc.

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.

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.

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.

6. Referencer

AC – administrative components : Dublin Core DCMI administrative metadata / Jytte Hansen og Leif Andresen (2003)

<http://www.bs.dk/standards/AdministrativeComponents.htm>

Collections and collection description focus /Pete Johnston & Bridget Robinson (2002)

<http://www.ukoln.ac.uk/cd-focus/briefings/bp1/bp1.pdf>

Dansk udvekslingsformat for metadata : DC/AC formatet / Susanne Thorborg og Leif Andresen (2003)

<http://www.bs.dk/publikationer/andre/dcac/index.htm>

DCMI metadata terms / DCMI Usage Board (2005)

<http://dublincore.org/documents/dcmi-terms/>

Discovering Online Resources: Unifying Resource Discovery Metadata for the Humanities: An Application Based upon the Dublin Core / Paul Miller (1997)

http://ahds.ac.uk/public/metadata/disc_05.html

Dublin Core metadata schema: ADS interpretation and mapping for the Arena project partners / Tony Austin (2003)

<http://ads.ahds.ac.uk/arena/DCmeta.pdf>

Guide to best practice: Dublin Core / Consortium for the Computer Interchange of Museum Information (2000)

http://www.cimi.org/public_docs/meta_bestprac_v1_1_210400.pdf

ISAD(G) : General International Standard Archival Description / adopted by the Committee on Descriptive Standards (2000)

http://www.ica.org/biblio/cds/isad_g_2e.pdf

XML skemaer til udveksling af Dublin Core metadata om netpublikationer / af Niels Erik Gilvad og Martin Everhøj(2003). København, IT-højskolen

7. Resume

Formater

Indholdsformat - Dublin Core anbefales som et fælles, konkret indholdsformat, hvortil man kan mappe de udvalgte dele af de enkelte sektors specifikke databaseindhold, som set fra almenhedens synsvinkel er relevante for præsentation og søgning.

Det må anses for påkrævet at præcisere eller kvalificere nogle af Dublin Core's 15 grundelementer med såkaldte "refinements". Desuden bør der inddrages visse elementer (i første række "identifier" og "source") fra Dublin Core DCMI Administrative Components.

Rammeformat – Rammeformatet eller udvekslingsformatet bør være XML (W3C-standard). Der er i tilknytning til rapporten udarbejdet et XML-schema, der kan rumme de fælles ABM-dataelementer. Dette schema kaldes *dkabm* og inkluderer dels internationale og anerkendte schemas, dels schemas, der er konstrueret til danske forhold. Sidstnævnte omfatter *dkedplus* (danske udvidelser til Dublin Core, bl.a. forskellige emneordslister til brug i elementet "subject") og *dktype* (danske udvidelser til ressource typer).

Datatransport

Som den mest realistiske model for lokale ABM-implementeringer foreslås en central model med en fælles database, der til stadighed opdateres med aktuelle oplysninger via høstning (harvesting). Den standard for indhøstning af dokumenter og/eller metadata fra nettet, som det på det nuværende udviklingsstade vil være mest nærliggende at anbefale, er OAI-PMH.

Som alternativ til den centrale model kan man forestille sig en web service model baseret på den nye protokol SRW, der tillader søgning på tværs af arkiver, biblioteker og museer. (Det må derimod anses at være for ressourcekrævende at implementere den fra bibliotekerne kendte protokol Z39.50 i arkiver og museer).

Fælles præsentation af data

Selektion – Selektion er en meget relevant problemstilling i arbejdet med ABM-data og -databaser. Der bør i alle tre institutionstypers materialer kunne foretages afgrænsninger i forhold til sted, tid og emne. Det bør også være muligt at afgrænse data af intern/forskningsmæssig/administrativ art, der ikke er beregnet for offentligheden.

Søgning. - Der peges på nytten af udarbejdelse af en struktureret ABM-emneordsliste og af udvikling af fælles standarder for beskrivelse af tid og sted.

Bilag A: Sektorspecifikke databaser

I det følgende gives en kort beskrivelse af de sektorspecifikke databaser, som har dannet udgangspunkt og reference for gruppens arbejde.

Arkiverne

Inden for arkivverdenen findes der en landsdækkende database DANPA samt en lang række institutionelle databaser. Endvidere har Statens Arkiver lanceret Daisy på Internettet, og Sammenslutningen af Lokal Arkiver (SLA) arbejder med en ny version af ARKIBAS.

- **DANPA** www.danpa.dk er en landsdækkende database for privatarkivmateriale. Efter opdatering i 2002 indeholder DANPA materiale fra 190 arkivinstitutioner. Antallet af registreringer er 120.000, og der er 265.000 henvisninger at søge på. Dataleverandører er Sammenslutningen af Lokal Arkiver (SLA), Statens Arkiver (Rigsarkivet og landsarkiverne), Det Kongelige Bibliotek, Arbejderbevægelsens Bibliotek og Arkiv, Udvandrerarkivet samt andre specialarkiver. Der er planlagt endnu en opdatering i 2006. DANPA er en udtræksdatabase, der søger at overkomme de forskellige registreringspraksis, der er en realitet i arkivverdenen. Databasen er stærkt præget heraf – både i præsentation og søgemuligheder. Det er en klar konklusion, at der i kølvandet på DANPA skal besluttes standarder og anvendelse af samme.
- **Daisy** (Dansk ArkivalieinforMationsSYstem) er Statens Arkivers IT-system til registrering og præsentation af oplysninger om samtlige arkivalier i Statens Arkiver varetægt. Daisy har været i anvendelse som registreringssystem siden efteråret 2000 og er under fortsat videreudvikling. Siden december 2004 har Daisy været tilgængelig på Internettet (daisy.sa.dk). Internetversionen er under fortsat videreudvikling, og på længere sigt vil det blive muligt at bestille arkivalier online via Daisy. I efteråret 2000 påbegyndte Statens Arkiver om- og/eller nyregistrering af samlingerne, og i august 2005 var omfanget af tilgængelige data i daisy.sa.dk: ca. 110.000 hyldemeter arkivalier, svarende til ca. en tredjedel af Statens Arkivers samlinger (Rigsarkivet, Landsarkiverne for Sjælland, Fyn, Nørrejylland, Sønderjylland samt Erhvervsarkivet).
- **ARKIBAS.dk**. Sammenslutningen af Lokal Arkiver (SLA) er ved at afslutte udviklingen af en central ARKIBAS-4, hvorfra det er planen at spejle over i en web-udgave kaldet ARKIBAS.dk. Arbejdet hermed vil blive påbegyndt i 2006. Alle ARKIBAS data, der ikke er med i DANPA, bliver hermed tilgængelige. Det er i alt ca. 1,5 mio. registreringer, heraf mere end 200.000 scannede billeder med registreringer.

Bibliotekerne

Der findes to nationale databaser DanBib og bibliotek.dk. Derudover findes en lang række databaser over en eller flere bibliotekers bestande, som vil være de relevante kilder ved regionale fælles baser. Regionalt har bibliotekerne også registreret lokalbibliografiske poster, som ikke indgår i DanBib/bibliotek.dk.

- **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: 25 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. 8 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.

- **Museernes Samlinger**, er den centrale database over museumssager og -genstande på danske statsanerkendte og statslige kulturhistoriske museer. Museerne indtaster online til Museernes Samlinger via nettet i et interface kaldet **Regin**. Museernes Samlinger (Regin) baserer sig på Dansk Museums Dokumentationsstandard, dvs. en kollegialt vedtaget norm for registrering af ”løs kulturarv” på danske museer. Teknisk set er Museernes Samlinger og Regin begge klienter på nettet i forhold til den centrale database. Godt 100 museer benytter dagligt Regin og i alt ca. 200.000 genstande er i sommeren 2005 registreret i den centrale database. Museernes Samlinger er offentligt tilgængelig på <http://kulturarv.dk/mussam>.
- **Kunst Indeks Danmark (KID)** er den centrale database 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. I 2007 forventes KID implementeret som en ”klient” på Museernes Samlinger.
- **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. 160.000 kulturhistoriske lokaliteter fra landjorden og fra det marine område.

Bilag B. Mapning fra Regin

Som grundlag for dette forslag til mapning af Regin er anvendt datamodellen for DMI-konvertering udarbejdet af UNI-C³.

Datamodellen er opdelt i primære entiteter, som beskriver genstande eller begreber indenfor domænet og i sekundære entiteter, som beskriver yderligere egenskaber eller aspekter ved de primære entiteter.

Endvidere består modellen af en række relationstabeller, som beskriver sammenhænge mellem entiteter både primære og sekundære samt en række hjælpetabeller indeholdende standardværdier m.v.

I mapningen er der lagt vægt på at opnå en 1:1 afbildning af de eksisterende primære entiteter med en overførsel af relevante attributter jvf. kommissoriet for mapningens dataindhold. Dette gælder for entiteterne: Arkivfond, Magnetmedier, Sag, Foto, Skibe, Genstand, Litteratur, Rapport og Storformat. De primære entiteter: Kunstner og Aktør optræder som attributter i de ovennævnte entiteter. Værdier fra sekundære entiteter samt hjælpetabeller er overført som attributter til de primære entiteter jvf. datamodellens angivelser af fremmednøgler i de primære entiteter samt de opregnede relationer i relationstabellerne. Primær entitet : Kommunikation og Værk er ikke medtaget i denne beskrivelse efter aftale.

I mapningen skelnes der i mellem de primære entiteter, således at entiteterne p_Sag og p_Arkivfond opfattes som samlende entiteter, der beskriver en helhed (Collection-level) og de resterende primære entiteter som beskrivende enkelte enheder (Items). Dette er identificeret ved hjælp af DC.Type med anvendelse af DCMI Type Vocabulary.

Når data er hentet fra en standardiseret liste eller vokabular, bør dette være identificeret ved anvendelse af DC Encoding Schemes. Dette vil f.eks. være tilfældet ved ”Den saglige registrant”, men der kan være anvendt andre, som ikke er angivet i konverteringsbeskrivelsen. For ”Den saglige registrant” anvendes forkortelsen SRKM.

Kodet information i f.eks. hjælpetabellerne bør overføres i sin opløste form ved konvertering. Dette af hensyn til både søgning og præsentation af data.

Entydig identifikation af de enkelte poster forslås angivet i AC element **identifier**, således at værdien fra den primære nøgle [p_entitet].[entitetId] + museumsId danner identifier i fællessystemet. Det er herved muligt at angive, dels hvilket system den enkelte post stammer fra, og i hvilken institution den fysiske genstand måtte befinde sig. Dette kræver dog, at entitets ID er fast knyttet til den enkelte indførsel. Placering indenfor institutionen angives ikke her, men kræver opslag i det oprindelige system. Dette kan evt. være via førnævnte entitets ID.

³ Relationel datamodel : DMI-konvertering for Kulturarvsstyrelsen. Version 1.3. UNI-C, marts 2004

DC-elementet **DC.Creator** anvendes til angivelse af egentligt ophav til det registrerede, dvs. ved sager er det den sagsoprettende institution, der angives.

For p_rapport og p_litteratur gælder, at de indgår som en del af p_Sag ved anvendelse af DC.Relation.hasPart hhv. isPartOf. Hvor p_Rapport og p_Litteratur er relateret til ligeværdige entiteter anvendes DC.relation.References hhv. isReferencedBy

Ved entiteten p_Litteratur anvendes eventuelt ISBN som DC.Identifier. Denne kan derved anvendes til at pege på biblioteksposter og derved danne en sammenhæng mellem data fra forskellige dataleverandører.

Ved konvertering af h_Topografnummer, der er et 4-cifret nummer, bør nummeret foldes ud til en tekst svarende til sogn, herred og gammelAmt.

Konvertering	Regin -> Dublin Core				
Administrative Components (gælder for alle entitetstyper)					
Label	Qualifier	Scheme	Værdi	Tabel	Feltnavn
AC.Identifier				p_[Entitet]	[entitet]Id
AC.Scource				p_[Entitet]	museumId
p_Sag				Regin	
Label	Qualifier	Scheme	Værdi	Tabel	Feltnavn
DC.Title				p_Sag	betegnelse
DC.Creator				a_Museum	navn
DC.Subject		"Den grønne registrant"		h_StedregistrantKlassifikation	
				h_StedregistrantGruppe	
				h_StedregistrantUndergruppe	
DC.Description				p_Sag	eksternNote
DC.Publisher					
DC.Contributor					
DC.Date					
DC.Type		DCMI Type Vocabulary	Collection		
DC.Format					
DC.Identifier				p_Sag	nummer
DC.Source					
DC.Language					

Label	Qualifier	Scheme	Værdi	Tabel	Felt navn
DC.Relation	hasPart			p_Genstand	genstandId
	hasPart			p_FotoFilm	fotoFilmId
	hasPart			p_Magnetmedium	magnetmediumId
	hasPart			p_Rapport	rapportId
	hasPart			p_Skib	skibId
	hasPart			p_Storformat	storformatId
	hasPart			r_Sag_Litteratur	litteraturId
	is part of			r_Sag_Litteratur	litteraturId
	hasPart			r_Sag_Arkivfond	arkivfondId
	is part of			r_Sag_Arkivfond	arkivfondId
DC.Coverage	Spatial			s_Stedregistrant	lokalitetsnummer
	Spatial			h_Topografinummer	sogn/herred/gammel Amt
	Spatial			s_LokalitetUdland	geografiskOmraade
				s_LokalitetUdland	etniskNavn
	Temporal	Period	name	h_Periode	tekst
	Temporal	Period	start	h_Periode	startAar
	Temporal	Period	end	h_Periode	slutAar
DC.Rights					

p_Arkivfond					
Label	Qualifier	Scheme	Værdi	Tabel	Felt navn
DC.Title				p_Arkivfond	betegnelse
DC.Creator				a_Museum	navn
DC.Subject		"Den grønne registrant"		h_StedregistrantKlassifikation	
				h_StedregistrantGruppe	
				h_StedregistrantUndergruppe	
DC.Description				p_Arkivfond	eksternNote
				p_Arkivfond	beskrivelse
DC.Publisher					
DC.Contributor					
DC.Date				p_Arkivfond	oprettelsesdato
					nedlæggelsesdato
DC.Type		DCMI Type Vocabulary	Collection		
DC.Format				p_Arkivfond	bindantal
					laegantal
					pakkeantal
DC.Identifier				p_Arkivfond	nummer
DC.Source					
DC.Language					
DC.Relation	hasPart			r_Sag_Arkivfond	sagId
	isPartOf			r_Sag_Arkivfond	sagId
	isReferencedBy			r_Arkivfond_Litteratur	litteraturId
	Reference			r_Arkivfond_Litteratur	litteraturId

Label	Qualifier	Scheme	Værdi	Tabel	Felt navn
DC.Coverage	Spatial			s_Stedregistrant	lokalitetsnummer
	Spatial			h_Topografinummer	sogn/herred/gammel Amt
	Spatial			s_LokalitetUdland	geografiskOmraade
				s_LokalitetUdland	etniskNavn
	Temporal	Period	name	h_Periode	tekst
	Temporal	Period	start	h_Periode	startAar
	Temporal	Period	end	h_Periode	slutAar
DC.Rights					

p_Genstand					
Label	Qualifier	Scheme	Værdi	Label	Felt navn
DC.Title				p_Genstand	betegnelse
DC.Creator					
DC.Subject		"Den grønne registrant"		h_StedregistrantKlassifikation	
				h_StedregistrantGruppe	
				h_StedregistrantUndergruppe	
DC.Description				p_Genstand	eksternNote
DC.Publisher					
DC.Contributor					
DC.Date				s_Datering	datering
DC.Type		DCMI Type Vocabulary	PhysicalObject		
DC.Format					
DC.Identifier				p_Genstand	nummer
DC.Source					
DC.Language					
DC.Relation	isReferencedBy			r_Genstand_Litteratur	litteraturId
	Reference			r_Genstand_Litteratur	litteraturId
	hasPart			s_MIMEtypeobjekt	filename
	isPartOf			p_Genstand	sagId

Label	Qualifier	Scheme	Værdi	Tabel	Felt navn
DC.Coverage	Spatial			s_Stedregistrant	lokalitetsnummer
	Spatial			h_Topografinummer	sogn/herred/gammel Amt
	Spatial			s_LokalitetUdland	geografiskOmraade
				s_LokalitetUdland	etniskNavn
	Temporal	Period	name	h_Periode	tekst
	Temporal	Period	start	h_Periode	startAar
	Temporal	Period	end	h_Periode	slutAar
DC.Rights					

p_Skib					
Label	Qualifier	Scheme	Værdi	Tabel	Felt navn
DC.Title				p_Skib	navn
	Alternative			p_Skib	kendingsbogstaver
	Alternative			p_Skib	tidligereKendingsbogstaver
DC.Creator					
DC.Subject		"Den grønne registrant"		h_StedregistrantKlassifikation	
				h_StedregistrantGruppe	
				h_StedregistrantUndergruppe	
DC.Description				p_Skib	eksternNote
				h_skibTypeId	tekst
				p_Skib	byggeaar
				p_Skib	byggested
DC.Publisher					
DC.Contributor					
DC.Date					
DC.Type		DCMI Type Vocabulary	PhysicalObject		
DC.Format					
DC.Identifier					
DC.Source					
DC.Language					
DC.Relation	isPartOf			p_Sag	sagId

Label	Qualifier	Scheme	Værdi	Tabel	Feltnavn
DC.Coverage	Spatial			s_Stedregistrant	lokalitetsnummer
	Spatial			h_Topografinummer	sogn/herred/gammel Amt
	Spatial			s_LokalitetUdland	geografiskOmraade
				s_LokalitetUdland	etniskNavn
DC.Rights					

p_FotoFilm					
Label	Qualifier	Scheme	Værdi	Tabel	Feltnavn
DC.Title				p_FotoFilm	beskrivelse
DC.Creator					
DC.Subject		"Den grønne registrant"		h_Stedregistrant Klassifikation	
				h_Stedregistrant Gruppe	
				h_Stedregistrant Undergruppe	
DC.Description				p_FotoFilm	eksternNote
DC.Publisher					
DC.Contributor					
DC.Date				p_FotoFilm	optagelsesdato
DC.Type		DCMI Type Vocabulary	StillImage [hvis h_FotoFilmTy pe [kode] ikke lig FLM]	h_FotoFilmType	tekst

		DCMI Type Vocabulary	MovingImage [hvis h_FotoFilmType pe [kode] er lig FLM]	h_FotoFilmType	tekst
DC.Format				h_FotoFilmFormat	tekst
	Extent			p_FotoFilm	spilletid
DC.Identifier				p_FotoFilm	nummer
DC.Source					
DC.Language					
DC.Relation	isPartOf			p_FotoFilm	sagId
	isReferencedBy			r_FotoFilm_ Litteratur	litteraturId
	Reference			r_FotoFilm_ Litteratur	litteraturId
	hasPart			s_MIMEtype objekt	filename
DC.Coverage	Spatial			s_Stedregistrant	Lokalitetsnummer
	Spatial			h_Topografinummer	sogn/herred/gammel Amt
	Spatial			s_LokalitetUdland	geografiskOmraade
				s_LokalitetUdland	etniskNavn
	Temporal	Period	name	h_Periode	tekst
	Temporal	Period	start	h_Periode	startAar

	Temporal	Period	end	h_Periode	slutAar
				s_Datering	datering
DC.Rights					

p_Storformat					
Label	Qualifier	Scheme	Værdi	Tabel	Felt navn
DC.Title				p_Storformat	beskrivelse
DC.Creator				a_Museum	navn
DC.Subject		"Den grønne registrant"		h_StedregistrantKlassifikation	
				h_StedregistrantGruppe	
				h_StedregistrantUndergruppe	
DC.Description				p_Storformat	eksternNote
DC.Publisher					
DC.Contributor					
DC.Date				p_Storformat	fremstillingsdato
DC.Type		DCMI Type Vocabulary	StillImage		
DC.Format				h_StorformatType	tekst
	Extent			h_StorformatFormat	tekst
DC.Identifier				p_Storformat	nummer
DC.Source					
DC.Language					

Label	Qualifier	Scheme	Værdi	Tabel	Felt navn
DC.Relation	isPartOf			p_Storformat	sagId
	isReferencedBy			r_Storformat_Litteratur	litteraturId
	Reference			r_Storformat_Litteratur	litteraturId
	hasPart			s_MIMEtypeobjekt	filename
DC.Coverage	Spatial			s_Stedregistrant	lokalitetsnummer
	Spatial			h_Topografinummer	sogn/herred/gammel Amt
	Spatial			s_LokalitetUdland	geografiskOmraade
				s_LokalitetUdland	etniskNavn
	Temporal	Period	name	h_Periode	tekst
	Temporal	Period	start	h_Periode	startAar
	Temporal	Period	end	h_Periode	slutAar
DC.Rights					

p_Litteratur						
Label	Qualifier	Scheme	Værdi	Tabel	Felt navn	Note
DC.Title				p_Litteratur	titel	
DC.Creator				p_Litteratur	forfatter	
DC.Subject				h_Litteratur Decimalklassifikation	kode	
				p_Litteratur	alternativKlassifikat	
DC.Description				p_Litteratur	eksternNote	
DC.Publisher				p_Litteratur	forlag	
				p_Litteratur	udgivelsessted	
DC.Contributor						
DC.Date						
				p_Litteratur	udgivelsesaar	
DC.Type		DCMI Type Vocabulary	Text			
DC.Format						
DC.Identifier				p_Litteratur	isbn	
DC.Source						
DC.Language						

Label	Qualifier	Scheme	Værdi	Tabel	Feltnavn	Note
DC.Relation						Fremmednøgle til
	Reference			r_Sag_Litteratur	sagId	p_Sag(sagId)
	isReferencedBy			r_Sag_Litteratur	sagId	p_Sag(sagId)
	Reference			r_Genstand_Litteratur	genstandId	p_Genstand(genstandId)
	isReferencedBy			r_Genstand_Litteratur	genstandId	p_Genstand(genstandId)
	Reference			r_FotoFilm_Litteratur	fotoFilmId	p_Fotofilm(fotofilmId)
	isReferencedBy			r_FotoFilm_Litteratur	fotoFilmId	p_Fotofilm(fotofilmId)
	Reference			r_Storformat_Litteratur	storformatId	p_Storformat(storformatId)
	isReferencedBy			r_Storformat_Litteratur	storformatId	p_Storformat(storformatId)
	Reference			r_Arkivfond_Litteratur	arkivfondId	p_Arkivfond(arkivfondId)
	isReferencedBy			r_Arkivfond_Litteratur	arkivfondId	p_Arkivfond(arkivfondId)
	Reference			r_Rapport_Litteratur	rapportId	p_Rapport(rapportId)
	isReferencedBy			r_Rapport_Litteratur	rapportId	p_Rapport(rapportId)
	Reference			r_Magnetmedium_Litteratur	magnetmediumId	p_Magnetmedium(magnetmediumId)

	isReferencedBy			r_Magnetmedium_ Litteratur	magnetmediumId	p_Magnetme dium (magnetmedi umId)
DC.Coverage						
DC.Rights						

p_Rapport					
Label	Qualifier	Scheme	Værdi	Tabel	Felt navn
DC.Title				p_Rapport	beskrivelse
DC.Creator				a_Museum	navn
DC.Subject		"Den grønne registrant"		h_Stedregistrant Klassifikation	
				h_Stedregistrant Gruppe	
				h_Stedregistrant Undergruppe	
DC.Description				p_Rapport	eksternNote
DC.Publisher					
DC.Contributor					
DC.Date				p_Rapport	udfaerdigelsesdato
DC.Type		DCMI Type Vocabulary	Text		
DC.Format					
DC.Identifier				p_Rapport	nummer
DC.Source					
DC.Language					

Label	Qualifier	Scheme	Værdi	Tabel	Felt navn
DC.Relation	isReferencedBy			r_Rapport_ Litteratur	Litteraturlid *1)
	Reference			r_Rapport_ Litteratur	Litteraturlid *1)
	isReferencedBy			p_Sag	nummer
DC.Coverage	Spatial			s_Stedregistrant	lokalitetsnummer
	Spatial			h_Topografinum mer	sogn/herred/gammel Amt
	Spatial			s_LokalitetUdland	geografiskOmraade
				s_LokalitetUdland	etniskNavn
	Temporal	Period	name	h_Periode	tekst
	Temporal	Period	start	h_Periode	startAar
	Temporal	Period	end	h_Periode	slutAar
DC.Rights					
*1) Fremmednøgle til p_Litteratur(litteraturlid)					

p_Magnetmedium					
Label	Qualifier	Scheme	Værdi	Tabel	Felt navn
DC.Title				p_Magnetmedium	beskrivelse
DC.Creator					
DC.Subject		"Den grønne registrant"		h_StedregistrantKlassifikation	
				h_StedregistrantGruppe	
				h_StedregistrantUndergruppe	
DC.Description				p_Magnetmedium	eksternNote
DC.Publisher					
DC.Contributor					
DC.Date					
DC.Type		DCMI Type Vocabulary			
DC.Format	Medium			h_MagnetmediumType	tekst
	Extent			p_Magnetmedium	spilletid
DC.Identifier				p_Magnetmedium	nummer
DC.Source					
DC.Language					

DC.Relation	Reference			r_Magnetmedium_Litteratur	LitteraturId *1)
	isReferencedBy			r_Magnetmedium_Litteratur	LitteraturId *1)
	isPartOf			p_Magnetmedium	sagId
Label	Qualifier	Scheme	Værdi	Tabel	Felt navn
DC.Coverage	Spatial			s_Stedregistrant	lokalitetsnummer
	Spatial			h_Topografinummer	sogn/herred/gammel Amt
	Spatial			s_LokalitetUdland	geografiskOmraade
				s_LokalitetUdland	etniskNavn
	Temporal	Period	name	h_Periode	tekst
	Temporal	Period	start	h_Periode	startAar
Temporal	Period	end	h_Periode	slutAar	
DC.Rights					

*1) Fremmednøgle til p_Litteratur(litteraturId)

Bilag C. Mapning fra Daisy

Som grundlag for udarbejdelse af mapning er anvendt formatbeskrivelse som beskrevet i Daisy dokumentation samt konverteringsbeskrivelse i ABM-projekt. DAISY datamodel er opdelt i 3 hoveddele bestående af Aktører (NG), Heuristiske enheder (HE) og Magasinenheder (ME). Beskrivelsen her er med udgangspunkt i en konvertering af data beskrivende NG og HE. ME-data beskriver forhold vedrørende lokalisering/opstilling og indhold af de enkelte arkiver og vil være tilgængelige ved opslag i DAISY.

Efter aftale med Statens Arkiver vil konverteringsbeskrivelsen være på et overordnet niveau i forhold til beskrivelse af dataudtræk fra de enkelte data-felter.

Forekomsten af en HE vil ved konverteringen betyde dannelse af en resources post. Den/de tilknyttede aktører vil optræde som DC.Creator med mulighed for at kvalificere en aktør, således at det navn aktøren er indført under kan udskilles fra eventuelle alternative former for samme aktør.

Statens Arkiver arbejder for tiden med at etablere et emnesystem samt system for geografiske navne. Dette betyder, at der i DAISY ikke findes oplysninger om emnemæssig dækning og geografisk dækning / virkeområde for de registrerede entiteter (både HE og Aktører). Efter aftale med Statens Arkiver foreslås det derfor, at HE Navn betragtes som ukontrollerede emneord, og derfor også placeres i DC.Subject (samt i DC.Title). For ligeledes at sikre en søgemulighed på geografisk dækning/virkeområde kan aktørnavne angives i DC.Coverage (udover at de findes i DC.Creator). Dette fordi aktørnavne i en del tilfælde vil indeholde en betegnelse for deres geografiske virkeområde. Når definitionen af felter og systemer til emneord og geografi ligger fast, skal disse inkluderes i konverteringen i stedet for den angivne løsning.

Det er pt. ikke muligt at lænke tilbage til Daisy fra eventuelt fællessystem. Der overvejes pt. andre løsninger for dette.

Konvertering	Daisy -> Dublin Core			Daisy		
Label	Qualifier	Scheme	Værdi	Tabel	Feltnavn	
AC.Identifier						[Daisy.ID]
AC.Source						Her kan angives SA
Label	Qualifier	Scheme	Værdi	Tabel	Feltnavn	Note
DC.Title				HeNavn	Navn	Hvor artID er officielt navn
	Alternative			HeNavn	Navn	Hvor artID er kaldenavn og/eller forkortelse
DC.Creator						
				NG	PrefixNavn + NGNavn[Navn]	Yngste navn.
	ActPeriod					Datoer for aktørens virkeperiode. Der skal angives yderår
	AlternativeName			NgNavn	Navn	(Hvis et sådan kan identificeres ?)
DC.Subject				HeNavn		Når emnedatasystem er fastlagt angives disse
DC.Description				He	PubBem	
DC.Publisher					Ikke relevant	

DC.Contributor						Ikke relevant
DC.Date				HeTid	Fra	
				HeTid	Til	
Label	Qualifier	Scheme	Værdi	Tabel	Felt navn	Note
DC.Type		DCMI Type Vocabulary	Collection			
DC.Format						Her angives medietype ved flere versioner af samme HE
	Extent					Antal fysiske enheder
DC.Identifier						Ikke relevant
DC.Source						Ikke relevant
DC.Language						Ikke relevant
DC.Relation						Ikke relevant
DC.Coverage	Spatial			NG	PrefixNavn + NGNavn[Navn]	Når system for geografisk betegnelse er fastlagt skal disse angives
	Temporal	Period				Her angives værdi for arkivets tidsmæssige dækning
						Der skal angives yderår
DC.Rights						Ikke relevant

Bilag D. Mapning fra Arkibas4

Mapning er foretaget på baggrund af Arkibas 4 version 1.1 dateret 30-01-03. Det har ikke været muligt at fremskaffe endelig dokumentation på datamodellen for Arkibas 4, da der stadig udvikles, og dataudtræk fra Arkibas 4 har ikke været til rådighed. Der må derfor tages forbehold for mapningens validitet i forhold til den endelige version af Arkibas 4.

Data er i visse tilfælde hentet fra en standardiseret liste eller vokabular, som bør identificeres ved hjælp af DC Scheme. Dette vil f.eks. være tilfældet i ”Decimal”, der angives ved DK5, men der kan være andre tilfælde, som ikke er angivet i konverteringsbeskrivelsen eksempelvis i forhold til ”Emneord”

Liste over relevante tabeller

Nummer	Tabelnavn	Bemærkning
1	Amt	Amter i Danmark – opslagstabel
2	ArkivFondIndhold	Indholdsoplysninger til arkivfonds. Inddeles i sektioner, som kan niveauinddeles hierarkisk i forhold til hinanden.
3	Bilag	Bilag til registrering og journal. Tilknyttes hele reg/jour eller bestemt del af indhold. Bilag kan være OLE objekt eller link til fil, web o.a.
4	Decimal	Tablet med oplysninger om DK5 klassifikationssystematik og benævnelse af de enkelte decimalklasser
5	DecimalStikord	Stikord til Decimalklassetabellen som anvendes til at fremsøge decimalklasserne ved henvisning.
6	Emneord	Emneord som er underordnet decimalklasserne på henvisningerne.
7	Henvisning	Henvisninger til registreringer danner grundlag for struktureret udsøgning.
8	Institution	Institutionsliste med tilmeldte arkiver.
9	Kategori	Kategorier til brug ved kategorisering af Arkivfonds.
10	KategoriStikord	Stikord til kategorier.
11	Kommune	Liste med kommuner i Danmark.
12	RefKategori	Referencetablet som knytter kategorier til arkivfondsregistreringer.
13	Registrering	Registreringer af alle typer – superklasse.
14	RegType	Opslagstabel med journal- og registreringstyper
15	Sekvensindhold	Indhold til sekvensregistreringer.
16	SrNr	Link af sænr til registreringer. Der kan tilknyttes flere sænr til hver registrering.
17	Sted	Stedkoder og stedbetegnelser.
18	StedStikord	Stikord til stedoplysninger.
19	StedGen	Stedkoder og betegnelser til fællessystematik
20	StedStikordGen	Stikord til fællessystematisk stedregister
21	DecimalLokal	Tablet med oplysninger om arkivernes lokale DK5 klassifikationssystematik og benævnelse af de enkelte klasser
22	DecimalLokalStikord	Stikord til arkivernes lokale Decimalklassetabel som anvendes til at fremsøge decimalklasserne

Administrative Components				Arkibas 4	
Label	Qualifier	Scheme	Værdi	Tabel	Felt navn
AC.Indentifier				Registrering	Id
AC.Source				Institution	LHAnr

Konvertering Arkibas 4 -> Dublin Core						
Label	Qualifier	Scheme	Værdi	Tabel	Felt navn	Note
DC.Title				Registrering	Betegnelse	Registreringer af alle typer – superklasse.
				Henvisning	Henvisning	Henvisninger til registreringer danner grundlag for struktureret udsøgning.

Label	Qualifier	Scheme	Værdi	Tabel	Felt navn	Note
DC.Creator				Registrering	Arkivskaber	
				Registrering	Fotograf	Registreringer af alle typer – superklasse.
	ActPeriod		start		Stiftet	
	ActPeriod		end		Nedlagt	

Label	Qualifier	Scheme	Værdi	Tabel	Feltnavn	Note
DC:Subject		DK5		Decimal	Decimal	Tabel med oplysninger om DK5 klassifikationssystematik og benævnelse af de enkelte decimalklasser
					DecimalTxt	
				DecimalStikord	Decimal	Stikord til Decimalklassetabellen som anvendes til at fremsøge decimalklasserne ved henvisning.
					Stikord	
				Emneord	Emneord	Emneord som er underordnet decimalklasserne på henvisningerne
				Kategori	Kategori	Kategorier til brug ved kategorisering af Arkivfonds.
					KategoriTxt	
					Kat 1-4	
				KategoriStikord	Stikord	Stikord til kategorier.
				DecimalLokal	Decimal	Tabel med oplysninger om arkivernes lokale DK5 klassifikationssystematik og benævnelse af de enkelte klasser
					DecimalTxt	
					CommonDecimal	
				DecimalLokalStikord	Decimal	Stikord til arkivernes lokale Decimalklassetabel som anvendes til at fremsøge decimalklasserne
					Stikord	

Label	Qualifier	Scheme	Værdi	Tabel	Feltnavn	Note
DC.Description				ArkivFondIndhold	Indhold	Indholdsoplysninger til arkivfonds. Inddeles i sektioner, som kan niveauinddeles hierarkisk i forhold til hinanden.
				Registrering	Betegnelse	Registreringer af alle typer – superklasse.
					Bem	
				Sekvensindhold	Indhold	Indhold til sekvensregistreringer.

Label	Qualifier	Scheme	Værdi	Tabel	Feltnavn	Note
DC.Publisher						

Label	Qualifier	Scheme	Værdi	Tabel	Feltnavn	Note
DC.Contributor						

Label	Qualifier	Scheme	Værdi	Tabel	Feltnavn	Note
DC.Date						

Label	Qualifier	Scheme	Værdi	Tabel	Feltnavn	Note
DC.Type				Registrering	RegType	Registreringer af alle typer – superklasse.
					Art	

Label	Qualifier	Scheme	Værdi	Tabel	Feltnavn	Note
DC.Format				ArkivFondIndhold	Bind	Indholdsoplysninger til arkivfonds. Inddeles i sektioner, som kan niveauinddeles hierarkisk i forhold til hinanden.
					Pakker	
					Laeg	
				Registrering	Varighed	Registreringer af alle typer – superklasse.
					laltBind	
					laltPakket	
					laltLaeg	
					Mael	
					Materiale	

Label	Qualifier	Scheme	Værdi	Tabel	Feltnavn	Note
DC.Identifier				Registrering	Regnr	Registreringer af alle typer – superklasse.
					Journalnr	
					Negativ	
					FotografNegativ	
				Srn	Srn	Link af særn til registreringer. Der kan tilknyttes flere særn til hver registrering.

Label	Qualifier	Scheme	Værdi	Tabel	Feltnavn	Note
DC.Source						

Label	Qualifier	Scheme	Værdi	Tabel	Feltnavn	Note
DC.Language						

Label	Qualifier	Scheme	Værdi	Tabel	Feltnavn	Note
DC.Relation	haspart			Bilag	Betegnelse	Bilag til registrering og journal. Tilknyttes hele reg/jour eller bestemt del af indhold. Bilag kan være OLE objekt eller link til fil, web o.a.

Label	Qualifier	Scheme	Værdi	Tabel	Feltnavn	Note	
DC.Coverage	Spatial			Amt	AmtNr	Amter i Danmark – opslagstabel	
					AmtNavn		
				Kommune	KommuneNr	Liste med kommuner i Danmark.	
					KommuneNavn		
					AmtNr		
				Sted	Stedkode	Stedkoder og stedbetegnelser.	
					StedTxt		
					Sted 1-3		
					StedStikord	Stikord	
	Temporal				ArkivFondIndhold	Aar	Indholdsoplysninger til arkivfonds. Inddeles i sektioner, som kan niveauiindeles hierarkisk i forhold til hinanden.
					Registrering	Fra	Registreringer af alle typer –

						superklasse.
					Til	
					Datering	
					Yderaar	

Label	Qualifier	Scheme	Værdi	Tabel	Felt navn	Note
DC.Rights				Registrering	Klausul	Registreringer af alle typer – superklasse.

Bilag E. Mapning fra danMARC2

Mapningen fra danMARC 2 til DC inkluderer alle MARC data, som umiddelbart lader sig overføre til DC. Denne prioritering er foretaget for ikke at lægge begrænsninger i den potentielle anvendelse af mapningen.

Udover denne fuldstændige mapning er også foretaget en mere simpel der, med udgangspunkt i FORMKAT-niveau, Folkebibliotekernes minimumsniveau og anvendelsesfrekvens for felter, skitserer hvilke data som primært kunne være relevante for almenheden.

Schemes

I konverteringsbeskrivelsen er identificeret en række schemes til angivelse af standardiserede lister og vokabularer f.eks. i forhold til materialetyper, ISBN, emneord og DK5. Herudover vil der kunne identificeres en række yderligere schemes, især i forhold DC.Subject. Disse er i en vis udstrækning allerede defineret i DC: <http://dublincore.org/documents/dcmi-terms/#H4>

Kodede data

En række felter/delfelter i DanMARC2 vil indeholde kodede data, f.eks. 009 *a *g. Det anbefales, at disse opløses ved konvertering af hensyn til præsentation og søgning.

Præsentation/Søgning

Generelt konverteres delfelter til de tilsvarende specifikke DC elementer, således at data i eksempelvis felt 245 opsplittes og konverteres til henholdsvis DC.Title og DC.Creator.

Dette gøres i tilfælde hvor søgbarheden af data prioriteres højt.

I visse tilfælde fraviges denne fremgangsmåde. Dette gør sig især gældende hvor ”del-helheds relationer” kan identificeres. Eksempelvis konverteres felt 248 og 558 uden opsplitning til DC.Relation. Dette gøres fordi den samlede præsentation af disse data prioriteres højt og det vurderes at være den bedste måde at undgå meningsforstyrrende informationstab omkring kontekst for disse data. Præsentationen bliver altså i disse tilfælde prioriteret på bekostning af, at søgbarheden bliver mindre optimal.

Felt 440 kunne logisk set også betragtes som tilhørende ”del-helheds relationer” og således mappes til DC.relation med de fordele dette vil give i forhold til angivelse af strukturer i præsentationsformat. På trods af dette er feltet mappet til DC.title, da det vurderes at være mest hensigtsmæssigt, at disse data kan søges på lige fod med andre titeldata.

Søgbarheden prioriteres altså i dette tilfælde på bekostning af præsentationen.

Det kan overvejes, at anvende qualifier for PreferredName i DC.Creator (danMARC2 felt 100, 110, 700, 710). Dette vil være relevant hvis henvisninger i danMARC2 (felt 9xx) ønskes mappet til fællessystem.

Administrative Components			DanMAR2		
Label	Qualifier	Scheme	Værdi	Felt/delfelt	
AC.Indentifier				001 *a	
AC.Source			DanBib		

Konvertering DanMARC2 -> Dublin Core							
Label	Qualifier	Scheme	Værdi	Felt/delfelt simpel	Felt/delfelt fuldstændig	Note	
DC.Title					210	Forkortet nøgletitel	
					222	Nøgletitel	
				239 *t *u	239 *t *ø *u *v	Uniform titel for musikværker som opstillingselement - uniform titel, identificerende tilføjelse, tilnavn, uddragstitel	
					240	Uniform titel som opstillingselement	
				245 *a *ø *c *n *o *g	245 *a *b *ø *n *o *c *u *p *q *r *s *x *g	Titel - hovedtitel, resttitel/alternativ titel, identificerende tilføjelser, undertitler, paralleltitler, hovedtitel på tilføjet værk af andet ophav, numerisk eller alfabetisk betegnelse for bind	
					440 *a *v	440 (minus *e *z)	Seriebetegnelse i materialets form
						739 *t *ø *u *v	Uniform titel (musikværker) - uniform titel, identificerende tilføjelse, tilnavn, uddragstitel
						740	Uniform titel
	Alternative				745 *i *a	745 (minus *æ)	Varianttitel eller anden titel

				795 Opsplittes	Analytisk titel
				840 (minus *z)	Serietitel i normativ form

Label	Qualifier	Scheme	Værdi	Felt/delfelt simpel	Felt/delfelt fuldstændig	Note
DC.Creator				100 *a *h *e *f	100	Personnavn som opstillingselement
				110 *a *e *s	110	Korporationsnavn (navn på gruppe, institution, stat/kommune, møde) som opstillingselement
				239 *a *h *e *f	239 *a *h *e *f *c	Navn på komponist/musiker/musikergruppe som del af uniform titel og som opstillingselement - enkeltstående navn, fornavne, romertal, tilføjelser, fødselsår
					245 *e *æ	Ophav – ophavsangivelse som giver anledning til opslag, identificerende ophavsangivelse til titlen
					250 *c	Ophavsangivelse, der hører til udgaven og som giver anledning til opslag
					513	Note om fælles ophavsangivelse (musikoptagelser)
				700 *a *h	700	Personnavn som søgeelement
				710 *a *c *s *i *k *j	710	Korporationsnavn (navn på gruppe, institution, stat/kommune, møde) som søgeelement
					739 *a *h *e *f *c	Navn på komponist/musiker/musikergruppe som del af uniform titel - enkeltstående navn, fornavne, romertal, tilføjelser, fødselsår
					745 *æ	Varianttitel - identificerende ophavsangivelse
					770	Personnavn som opslagselement for analyse
					780	Korporationsnavn som opslagselement for

						analyse
Label	Qualifier	Scheme	Værdi	Felt/delfelt simpel	Felt/delfelt fuldstændig	Note
DC.Subject					050 (MultiMARC)	Library of Congress klassifikation
					060	National Library of Medicine klassifikation
					070 (MultiMARC)	National Agricultural Library klassifikation
					079 (MultiMARC)	Deutsche Bibliothek klassifikation
					080 (MultiMARC)	UDK (Det universelle decimalklassifikationssystem)
					082 (MultiMARC, ISSN Danmark)	Dewey Decimal Classification
					085 (MultiMARC)	British Catalogue of Music klassifikation
					086	Verbalisering af klassifikationsnotation
					087-089	Lokal klassifikation
				600	600	Personnavn som emneord
				610	610	Korporationsnavn (navn på gruppe, institution, stat/kommune, møde) som emneord
				630	630	Kontrolleret emneord
				631	631	Ukontrolleret emneord
				645	645	Titel som emneord
					650	Library of Congress emneord
		DK5		652 *m *p *a *h *e *f	652	DK5-klassemærke, inkl. alfabetisk underdeling
					660	MeSH emneord
					661	Europæisk Pædagogisk Tesaurus
					662	AGROVOC

		DBC's emneordliste		666 (minus *e *q * i)	666 (minus *e *q * i)	Kontrolleret DBC emneord
Label	Qualifier	Scheme	Værdi	Felt/delfelt simpel	Felt/delfelt fuldstændig	Note
					670	NAL emneord
					690	COMPASS emneord
Label	Qualifier	Scheme	Værdi	Felt/delfelt simpel	Felt/delfelt fuldstændig	Note
DC.Description				504	504	Indholdsbeskrivende note
					507	Note om anledning til udgivelse
					509	Note om besætning (musikalier)
				512	512	Note til beskrivelsen
				517	517	Note om målgruppe og anvendelighed
				530	530	Indholdsnote
				559	559	Generel note
Label	Qualifier	Scheme	Værdi	Felt/delfelt simpel	Felt/delfelt fuldstændig	Note
DC.Publisher				260 *a *b	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 (Kun hvis enkeltperson angives som forfatter eller ophav, element nr. 2)	Korporationsnavn (navn på gruppe, institution, stat/kommune, møde) som søgeelement
Label	Qualifier	Scheme	Værdi	Felt/delfelt simpel	Felt/delfelt	Note

					fuldstændig	
DC.Contributor					245 *f *k	Titel og ophavsangivelse - ophavsangivelse, som ikke giver anledning til opslag, medlemmer af gruppe
					250 *d	Ophavsangivelse, der hører til udgaven og som ikke giver anledning til opslag

Label	Qualifier	Scheme	Værdi	Felt/delfelt simpel	Felt/delfelt fuldstændig	Note
DC.Date				008 *a *z	008 *a *z	Generelle søgekoder for udgivelsesår
				260 *c	260 *c	Publicering, distribution el. lign. – udgivelsesår, distributionsår el. lign.

Label	Qualifier	Scheme	Værdi	Felt/delfelt simpel	Felt/delfelt fuldstændig	Note	
DC.Type		DCMI Type Vocabulary	Collection hvis h Text hvis b	004 *a	004 *a	Kode for bibliografisk posttype Anvendes hvor data er h eller b	
					005	Søgekoder for musikaler og lydoptagelser	
						008 *d *j *k *o	Generelle søgekoder – koder for indholdets form, skønlitterær form, biografi, børne- og skolemateriale
				009 *a	009 *a *b	Kode for materialebetegnelse – generelle materialebetegnelser	
						038	Søge/opstillingskode for børnelitteratur
						039	Søge/opstillingskode for musikoptagelser mv.
						505	Note om art eller form
						506	Note om disputats eller anden afhandling

Label	Qualifier	Scheme	Værdi	Felt/delfelt simpel	Felt/delfelt fuldstændig	Note
DC.format					008 *q	Generelle søgekoder – kode for filtype
				009 *g	009 *g *h	Kode for specifikke materialebetegnelser
					259	Filbeskrivelse
				300 *n *a *d	300	Fysisk beskrivelse
				501	501	Note om systemkrav eller adgangsmåde
						856 *q

Label	Qualifier	Scheme	Værdi	Felt/delfelt simpel	Felt/delfelt fuldstændig	Note
DC.Identifier		DS/ISO 2108		021 *a *x	021	ISBN
		DS/ISO 3297		022 *a *x *z	022	ISSN
					023	Producentens stregkode
		ISO 3901			024	ISRC
		DS/ISO 10444			027	ISRN
		DS/ISO 10957		028	028	ISMN
		CODEN			030	CODEN
					245 *z	Titel og ophavsangivelse - ISRC
					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	856 *u	Elektronisk lokalisering og adgangsmåde – URI, URL eller PURL	

Label	Qualifier	Scheme	Værdi	Felt/delfelt simpel	Felt/delfelt fuldstændig	Note
DC.Source				241 *a	241	Originaltitel
					502	Note om originaltitel eller oversættelse
					520	Note om værkets bibliografiske historie

Label	Qualifier	Scheme	Værdi	Felt/delfelt simpel	Felt/delfelt fuldstændig	Note
DC.Language		ISO 639-2		008 *l	008 *l	Generelle søgekoder – kode for hovedsprog
					041	Sprogkoder
					508	Note om sprog og alfabet

Label	Qualifier	Scheme	Værdi	Felt/delfelt simpel	Felt/delfelt fuldstændig	Note	
DC.Relation	isPartOf				013	ID-nummer på post på højeste niveau	
	isPartOf			014	014	ID-nummer på post på højere niveau	
	hasPart				015	ID-nummer på post på laveste niveau	
	isPartOf				016	ID-nummer på værtspublikation	
						017	ID-nummer på første udgave
	hasPart					247 *g *a *c *e *f	Sektionsspecifikation for flerbindsværk eller supplement - sektionsnummer, sektionstitel, sektionsundertitel og anden titelinformation, ophavsangivelse, sekundær ophavsangivelse
	hasPart				248 *g *a	248 *g *a *c *e *f	Bindspecifikation for flerbindsværk eller supplement - bindnummer eller supplementsbetegnelse, bindtitel eller supplementstitel, undertitel og anden titelinformation, ophavsangivelse, sekundær ophavsangivelse
					250 *a	250 *a *x	Udgave- og oplagsbetegnelse
					521	Note om udgavens oplag	

	isFormatOf				523	Note om udgivelse i andre materialetyper
Label	Qualifier	Scheme	Værdi	Felt/delfelt simpel	Felt/delfelt fuldstændig	Note
					525	Note om komplekse forhold til andre periodica
					526	Note om sammenhæng med andre værker
	isReferenced By				529	Note om referencer til værket
	hasPart				534	Note om partielt indhold
	isPartOf			557 *a *v *k	557	Periodicum som værtspublikation for I-analyse
	isPartOf			558 *a *e *g	558	Monografi som værtspublikation for I-analyse
	Replaces			860 *i *t	860	Tidligere titel (periodica)
	isReplacedBy			861 *i *t	861	Senere titel (periodica)
					863	Udgivet sammen med/som del af (periodica)
					865	Udgave i andet medium (periodica)
	isVersionOf				866	Udgave på originalsprog (periodica)
	isVersionOf				867	Hovedudgave på andet sprog og/eller med andet indhold
	hasVersion				868	Udgaver på andre sprog og/eller med ændret indhold (periodica)
					870	Hovedpublikation til supplement (periodica)
					871	Supplement til hovedpublikation (periodica)
	isPartOf				873	Hovedserie (periodica)
	hasPart				874	Underserie (periodica)
					879	Andre relationer (periodica)

Label	Qualifier	Scheme	Værdi	Felt/delfelt simpel	Felt/delfelt fuldstændig	Note	
DC.Coverage	spatial				033	Amtskommune eller kommune (Dansk lokalbibliografi)	
					620	Overskrift til dansk kortfortegnelse	
					621	Emnedata for kartografisk materiale	
					622	Systematisk notation for kartografiske materialer	
				633	633	Stednavn som emneord	
	temporal			634	634	Periodebetegnelse som emneord	
	spatial					651	Library of Congress geografisk emneord
				666 *e *q	666 *e *q		Kontrolleret DBC emneord – stednavn som emneord
	temporal			666 *i	666 *i	Kontrolleret DBC emneord – tidsangivelse	

Label	Qualifier	Scheme	Værdi	Felt/delfelt simpel	Felt/delfelt fuldstændig	Note
DC.Rights					008 *n	Kode for netdokumenters tilgængelighed
					518	Note om anvendelsesklausul

Bilag F. Vurdering af eksport af data

Index Data fik til opgave at vurdere implementering af den i nærværende rapport beskrevne model for eksport af data. Vurderingen gengives som modtaget fra Index Data.

Index Data har fået til opgave at undersøge hvilke vanskeligheder der evt. kunne ligge i at rent praktisk eksportere data fra de enkelte sektorspecifikke formater og at konvertere dataene til det fælles ABM-format. Dette indebærer at finde ud af om det i dag er muligt at få data ud af de enkelte systemer, vurdere hvilken kvalitet de eksporterede data har i forhold til de oprindelige data, dvs. om der er et datatab i forbindelse med eksporten, og endelig se om der er andre problemer i forbindelse med eksporten og konverteringen. Denne analyse er beskrevet i denne rapport.

De sektorspecifikke formater er Arkibas, Regin, Daisy og DanMarc. Hvert enkelt format vil blive gennemgået i det følgende med henblik på at klarlægge hvilke vanskeligheder der ligger i en sådan eksport og konvertering.

Arkibas

Arkibas er det software de lokalhistoriske arkiver benytter til registrering af arkivalier. Arkibas er i skrivende stund midt i et udviklingsforløb hvor man går væk fra en egentlig applikationsløsning, dvs. hvor de enkelte arkiver har installeret et program på de enkelte Pc'er, til en web baseret løsning, hvor alle data ligger samlet i én central database. Den tidligere løsning hedder Arkibas 3 og den kommende kommer til at hedde Arkibas 4. Vi har diskuteret dels med Danmarks Biblioteksskole, som står for udarbejdelsen af konverteringsnøglen fra Arkibas formatet til ABM formatet og dels med organisationen bag Arkibas, om hvilken version vi skulle gå ud fra. Valget er faldet på Arkibas 4. Dette skyldes flere forskellige forhold, men hovedårsagen er at selv om det vurderes, at der vil gå nogle år før størstedelen af de lokalhistoriske arkiver konverterer til Arkibas 4, så er det trods alt fremtidens system

Vi har holdt et møde med organisationen bag Arkibas, hvor de beskrev det nye system og fortalte, at det nye dataformat vil være kraftigt inspireret af de internationale arkivstandard ISAD(G) og ISAAR. Arkibas 4 er sat i drift; men eksportfunktionen er ikke færdigimplementeret, og vi har derfor ikke kunnet få udleveret Arkibas 4-data, som vi har kunnet teste. Vi har dog haft lejlighed til at se på Arkibas 3-poster, og det er på baggrund af disse, vi har lavet vores analyse. Ligeledes har vi haft lejlighed til at studere en foreløbig datamodel for Arkibas 4.

Selvom der er visse usikkerheder omkring Arkibas 4-dataene, er det dog vores vurdering ud fra den nye struktur i Arkibas samt ud fra det, vi har fået fortalt, at det vil være en overkommelig opgave at eksportere Arkibas 4-poster til ABM-formatet.

Regin

Regin er brugerinterface til museernes webbaserede registrering i den centrale database kaldet "Museernes Samlinger". Regin og Museernes Samlinger er sat i drift, men eksportfunktionen er ikke færdigimplementeret (sommer 2005), og vi har derfor ikke kunnet få udleveret Regin-data, som vi har kunnet teste. Vi har derimod haft lejlighed til dels at drøfte eksportfunktionen med Christian Ertmann-Christiansen, som er den teknisk ansvarlige hos KUAS, og dels har vi fået udleveret et XML skema, der beskriver eksportformatet. Ud fra dette er det vores vurdering, at det vil være en forholdsvis overkommelig opgave at konvertere Regin-posterne til ABM-formatet.

Der er dog et forhold, der komplicerer konverteringen, idet KUAS har valgt kun at medtage referencer til diverse kontrollerede vokabularer, men ikke selve dataene i eksportdataene. Eksempel: Genstandene i Regin katalogiseres efter "Den grønne registrant" og en skjorte, der har tilhørt Chr. 4., bliver måske katalogiseret som "dragt". I de eksporterede data står der imidlertid ikke "dragt", men i stedet et id, der refererer til denne genstandstype i "Den grønne registrant".

For at ordet "dragter" skal være søgbart og kunne vises for brugerne, skal det efterfølgende lægges ind i posterne i forbindelse med at de konverteres til ABM-formatet. Dette er kun muligt hvis der findes en fil eller database, hvori sammenhængen mellem referencerne og ordene er angivet, og det gør der efter Christian Ertmann-Christiansens udsagn. Dette forhold gør dog konverteringen noget mere kompleks og dyr, specielt hvis der er mange af disse filer eller databaser, der skal slås op i og hvis de har forskelligt format, således at der skal laves mange specialløsninger for at genskabe nogle brugbare data.

Daisy

Daisy er Statens Arkivers dataformat til registrering af deres arkivalier. Daisy kan eksportere posterne i et velbeskrevet XML-format, som dog er udfordrende, idet selve filerne kan være meget store, selv om dette ikke burde give anledning til væsentlige problemer.

Daisy-formatet kan også give problemer, som beskrevet under Regin-formatet, mht. data der kun bliver refereret til via et id, eller fordi dataene af andre grunde ikke egner sig til direkte præsentation for brugere. Det er dog uklart hvor relevante, de data der refereres til er i ABM-sammenhæng. Det må dog forventes at dataene i et eller andet omfang skal beriges fra eksterne kilder, før de er egnede til at blive vist i ABM-sammenhæng.

Derudover er der problemer med, at nogle felter indeholder en sammenblanding af dels data, der skal vises for brugerne, og dels data der ikke skal vises. Et eksempel er Proveniensnumre hvor selve nummeret har et foranstillet afdelingsnummer. Under en konvertering skal disse numre skilles ad og behandles separat.

De data i Daisy, der er relevante i ABM-sammenhæng, er dog så få at det ikke vurderes, at en konvertering til ABM-formatet vil volde væsentlige problemer.

danMARC2

DanMARC2 er bibliotekernes dataformat til håndtering af bibliografiske oplysninger og er beskrevet på <http://www.kat-format.dk/danMARC2/>. DanMARC2 er det ældste, formentligt det mest veldokumenterede og mest gennemtestede format og dette sammenholdt med at de bibliografiske poster er forholdsvis simple i forhold til de andre formater, der er blevet gennemgået i denne rapport, så forventer vi ikke væsentlige problemer med konverteringen af danMARC2 -poster til ABM- formatet. Der er dog et par områder, man skal være opmærksom på:

- DanMARC2 er et binært ISO2709 format (se <http://www.danbib.dk/index.php?doc=iso2709format> for yderligere oplysninger), men det er også almindeligt at se danMARC2 poster i det såkaldte linjeforamt (se <http://www.danbib.dk/index.php?doc=linjeforamt> for yderligere oplysninger). Ydermere er det for nyligt blevet muligt at repræsenterer danMARC2 poster i XML vha. MarcXchange-formatet (se <http://www.bs.dk/MarcXchange/> for yderligere oplysninger). Der er således i princippet 3 forskellige formater, som et bibliotek kan levere danMARC2 -poster i, men da vi i alle andre henseender går efter en XML-løsning, har projektgruppen besluttet, at poster kun kan eksporteres i MarcXchange hvis de skal konverteres til ABM-formatet.
- I danMARC2 -formatet er der i princippet de samme problemer, som beskrevet under Regin-formatet, mht. data der kun bliver refereret til via et id eller fordi dataene af andre grunde ikke egner sig til direkte præsentation for brugerne. Materiale type er et godt eksempel på sådanne data, hvor der f.eks. bare står ”tb” i stedet for ”Cd-rom”. Det er dog vores vurdering, at de data, som evt. skal substitueres i forbindelse med en

konvertering fra danMARC2 til ABM-formatet, ikke er så kritiske som i Regin, og det er derfor nemmere at lave brugbare data ud af danMARC2 - poster.

Konklusion

Det er vores vurdering, at der ikke er uoverkommelige problemer forbundet med konverteringen af nogle af de fire her omtalte dataformater, til ABM-formatet.

Det vil dog kræve en vis arbejdsindsats at skrive de nødvendige konverteringsscripts, dels fordi dataene er så forskellige, og dels fordi der skal en del databerigelse af dataene til, for at de egner sig til brug i ABM-sammenhæng.

Endeligt skal det tilføjes, at denne vurdering er behæftet med en vis usikkerhed, idet vi kun har haft mulighed for at se på eksempeldata fra to af de fire involverede dataformater, og det er erfaringsmæssigt her man opdager de små detaljer, som kan give problemer i konverteringssammenhæng.