



**DIGITALISERING AF
KULTURARVEN**
— ENDELIG RAPPORT FRA
DIGITALISERINGSUDVALGET

DIGITALISERING AF KULTURARVEN
– endelig rapport fra Digitaliseringsudvalget

Udgivet i 2009 af:

Kulturministeriet

Nybrogade 2

1203 København K

Rapporten udgives kun elektronisk på Kulturministeriets hjemmeside, www.kum.dk

ISBN, elektronisk version: 978-87-7960-130-7

Indhold

1	Indledning	3
1.1	Kommissorium	4
1.2	Udvalgets arbejde	6
2	Udvalgets overvejelser	7
2.1	Formål med digitalisering af kulturarven.....	8
2.2	Tilgængeliggørelse af digitaliseret materiale	10
2.3	Ophavsret	12
2.4	Formidling	13
2.5	Baggrund for udvalgets forslag	14
3	Forslag 1	17
3.1	Udvalgets forslag	17
3.2	Udvalgets bemærkninger	21
4	Forslag 2	24
4.1	Udvalgets forslag	24
4.2	Udvalgets bemærkninger	30
5	Forslag 3	32
5.1	Udvalgets forslag	32
5.2	Udvalgets bemærkninger	35
Bilag 1:	Skematisk oversigt over forslag	37
Bilag 2:	Teknisk redegørelse vedr. digitalisering af kulturarven	38

1 Indledning

I regeringsgrundlaget "Mulighedernes Samfund" fra november 2007 er digitalisering af kulturarven nævnt med følgende tekst:

"Regeringen nedsatte i efteråret 2006 et udvalg med henblik på at få udarbejdet forslag til digitalisering af udvalgte, prioriterede dele af den danske kulturarv. Regeringen vil fortsætte arbejdet med digitalisering af kulturarven".

Af regeringsgrundlaget fremgår det endvidere, at:

"Kulturarven har væsentlig betydning for danskernes identitetsfølelse i en globaliseret verden, og kunst og kultur får i disse år en stigende betydning. Regeringen vil derfor fortsætte arbejdet med formidling af den danske kulturarv nationalt og internationalt".

Området er yderligere understreget i Kulturministerens Arbejdsprogram fra august 2008.

Et initiativ, der har understøttet dette fokus, var nedsættelsen af et udvalg om digitalisering af kulturarven i oktober 2006 som led i den mediepolitiske aftale for 2007-2010. Nedsættelsen af udvalget udsprang igen af udredningen om bevaring af kulturarven fra 2003, og er i forlængelse af den indsats for digitalisering af kulturarven, der allerede er påbegyndt finansieret af UMTS-midler.

Udvalgets arbejde skal endvidere ses på baggrund af indsatsen på fælleseuropæisk plan, med startskud i kommissionens meddelelse i 2010: Digitale Biblioteker, der sætter fokus på onlineadgang til den europæiske kulturarv. Heraf er udsprunget en række initiativer, bl.a. den fælleseuropæiske kulturarvsportal Europeana, som er under opbygning, samt konklusionerne fra Det Europæiske Råd af 13. november 2006, hvor medlemsstaterne gensidigt opfordrede hinanden til at udarbejde nationale strategier for digitalisering af kulturarven, en opfordring som Danmark tilsluttede sig.

Den foreliggende rapport konkluderer på udvalgets overvejelser ved at fremlægge tre forskellige forslag til digitalisering af udvalgte, prioriterede dele af kulturarven. Denne afsluttende rapport skal igen ses på baggrund af udvalgets midtvejsrapport fra 5. september 2008, der i lyset af danske og internationale erfaringer gennemgik en række faglige, tekniske, organisatoriske og ophavsretslige problemstillinger i forbindelse med digitalisering af kulturarven.

Det digitaliseringsarbejde, som rapporten beskæftiger sig med, er således afgrænset til at omfatte kulturarven. I den forbindelse lægges det til grund, at arbejdet vedrører de samlinger, som statslige og statsstøttede institutioner under Kulturministeriet har opbygget i kulturbevaringsøjemed. Det drejer sig primært om de institutioner, der opbygger samlinger i henhold til arkivloven, filmloven, museumsloven og pligtafleveringsloven. Hertil kommer på det audiovisuelle område de programarkiver, der er opbygget af public service-institutionerne, det vil sige DR og TV 2, samt Statens Mediesamling i Statsbiblioteket. Andre institutioners samlinger af kulturbevaringsmæssig interesse er indgået i arbejdsgruppens overvejelser efter en konkret vurdering.

Rapporten vedrører primært digitalisering af kulturarvsmateriale, som foreligger i analog form¹, såkaldt retrodigitalisering (i modsætning til digitalt født materiale). Herunder kan der være tale om digitalisering af registre, kataloger mv. (digitalisering af metadata²) – en aktivitet, der ofte betegnes som retrokonvertering. Dette er en aktivitet, der på de forskellige kulturbevaringsområder er gennemført i varieret, men betydeligt omfang.

Der kan også være tale om digitalisering af selve de kulturelle objekter inden for de forskellige materialetyper, som levende billeder, bøger, arkivalier, lyd mv. På nuværende tidspunkt er denne objektdigitalisering kun gennemført i begrænset omfang, og det er en beskrivelse af denne opgave, herunder omfang, løsningsmuligheder og formidlingsperspektiver, der er rapportens hovedemne.

Rapporten vedrører som nævnt digitalisering af analogt materiale, men når det digitale resultat foreligger, vil dette materiale på de berørte institutioner blive behandlet parallelt med det digitalt fødte materiale, f.eks. i formidlingssammenhæng og i sammenhæng med bevaringsovervejelser.

1.1 Kommissorium

Udvalget om digitalisering af kulturarven blev nedsat i oktober 2006 med følgende kommissorium:

Kommissorium for arbejdsgruppe om digitalisering af kulturarven

Baggrund

Som led i den mediepolitiske aftale for 2007-2010, indgået den 6. juni 2006 mellem regeringen, DF, S, RV og SF, er det aftalt at nedsætte en arbejdsgruppe, der skal udarbejde forskellige forslag til digitalisering af udvalgte, prioriterede dele af kulturarven.

Af den mediepolitiske aftale fremgår:

I forlængelse af udredningen om bevaring af kulturarven fra 2003 nedsættes der en bredt sammensat arbejdsgruppe – under Kulturministeriets ledelse – med repræsentanter for de relevante offentlige myndigheder, institutioner og organisationer, som får til opgave at udarbejde forskellige forslag til digitalisering af udvalgte, prioriterede dele af kulturarven med henblik på bevaring, formidling og tilgængeliggørelse heraf; herunder digitalisering af DR's arkiver. I forbindelse med udnyttelsen af materialet i arkiverne er det vigtigt at overveje de ophavsretslige problemstillinger. Forslaget skal foreligge i løbet af 2007. Udgangspunktet er, at udgifterne til digitaliseringen af DR's arkiver afholdes af DR's egne midler, inkl. 75 mio. kr. af det forventede licensmerprovenu i årene 2005 og 2006 på i alt 150 mio. kr.

På den baggrund nedsættes en arbejdsgruppe med nedenstående sammensætning og opgaver.

¹ I enkelte tilfælde er der tale om materiale, der ikke er egentligt analogt, f.eks. DAT-bånd, hvor båndet indeholder digitalt materiale, men hvor det digitale tekniske format har kort levetid. I disse tilfælde skal der ske en konvertering af materialerne, så de sikres med henblik på langtidsbevaring.

² Metadata er data, der leverer information om andre data, f.eks. et register, hvori en genstand står beskrevet. Registeret indeholder metadata og selve den digitaliserede genstand betegnes som objektet.

Arbejdsgruppens sammensætning

Arbejdsgruppen skal bestå af én repræsentant for henholdsvis Statens Arkiver, Det Kongelige Bibliotek, Statsbiblioteket, Statens Museum for Kunst, Nationalmuseet, Biblioteksstyrelsen, Kulturarvsstyrelsen, Det Danske Filminstitut, DR som repræsentant for public service-institutionerne, Samrådet for Ophavsret og Ophavsretligt Forum, samt ét medlem udpeget af Det Strategiske Forskningsråd, ét medlem udpeget af Forskningsrådet for Kultur og Kommunikation og ét medlem udpeget af it-brancheforeningerne, samt af henholdsvis Finansministeriet, Videnskabsministeriet og Kulturministeriet (formand).

Arbejdsgruppen kan beslutte at nedsætte underudvalg, som skal tage stilling til enkelte problemstillinger som ophavsretlige spørgsmål, spørgsmål relateret til formidling af digitaliseret materiale mv.

Arbejdsgruppen kan beslutte at inddrage ekstern bistand og eksperter i arbejdsgruppens arbejde.

Kulturministeriet varetager sekretariatsfunktionen for arbejdsgruppen.

Arbejdsgruppens opgaver

Arbejdsgruppen skal udarbejde forskellige forslag til digitalisering af udvalgte, prioriterede dele af kulturarven. Det overordnede fokus for arbejdsgruppens arbejde og opgaveløsning er bevaring, formidling og tilgængeliggørelse af kulturarven.

Formålet er endvidere at fremme samarbejde og sikre en rational arbejdsdeling mellem de involverede parter nationalt og internationalt.

Arbejdsgruppen skal:

- Som tema have udvalgte prioriterede dele af kulturarven i form af tekst (bøger, tidsskrifter, aviser), fotos, museumsgenstande, arkivalier og audiovisuelt materiale
- Foretage en vurdering af, hvilket ikke digitalt født eller ikke digitaliseret materiale der evt. skal digitaliseres med henblik på bevaring, formidling og tilgængeliggørelse, samt give begrundelse herfor
- På baggrund af en bevaringsfaglig, formidlingsfaglig og samfundsøkonomisk bedømmelse overveje, hvilke selektionskriterier der kan benyttes ved digitalisering af udvalgte prioriterede dele af kulturarven og anbefale selektionskriterier til de respektive løsningsmodeller
- Foretage en kortlægning af igangværende og planlagte initiativer vedrørende digitalisering af kulturarven og tilgængeliggørelse heraf
- Koordinere sit arbejde i forhold til andre igangsatte eller påtænkte initiativer med digitalt formidlings- og bevaringsperspektiv
- Inddrage internationale erfaringer med digitalisering af kulturarven, i det omfang det skønnes nødvendigt
- Kortlægge ophavsretlige problemstillinger, som knytter sig til digitalisering og tilgængeliggørelse af kulturarven og, hvis der er behov derfor, fremkomme med løsningsforslag
- Udarbejde flere forskellige selvstændige, men kombinerbare løsningsmodeller og flere forskellige omkostningsniveauer for digitalisering af udvalgte, prioriterede dele af kulturarven. Arbejdsgruppens forskellige modeller skal indeholde en bedømmelse af omfanget af digitalisering samt en tidshorizont for de forskellige digitaliseringsmodeller. Mindst én af de overvejede løsningsmodeller skal kunne finansieres inden for kulturarvsinstitutionernes

ordinære driftsbevillinger. Udgangspunktet er, at udgifterne til digitaliseringen af DR's arkiver afholdes af DR's egne midler, inkl. 75 mio. kr. af det forventede licensmerprovenu i årene 2005 og 2006 på i alt 150 mio. kr.

- Overveje eventuelle lovgivningsmæssige og bevillingsmæssige konsekvenser af de fremlagte forslag.

Tekniske overvejelser relateret til bevaring af digitalt materiale håndteres i en allerede nedsat styregruppe for digital bevaring, der på tværs af arkiver, biblioteker og museer skal samarbejde om den fortsatte udvikling af hensigtsmæssige bevaringsmetoder og -strategier.”

1.2 Udvalgets arbejde

Udvalgets arbejde har i den første fase (2006-2007) primært vedrørt en kortlægning af det stedfundne og planlagte digitaliseringsarbejde på kulturarvsområdet og de bevarings- og formidlingsmæssige hensyn, der ligger bag aktiviteterne. Der er endvidere med bistand af it-analysevirksomheden Gartner foretaget en vurdering af de omkostninger, der er forbundet med dele af digitaliseringen. Ligeledes er der gennemført en undersøgelse af en række ophavsretlige forhold, der kan læses om i kapitel 7.

Den anden fase af udvalgets arbejde (2007-2008) har særlig vedrørt en undersøgelse af dels de faglige, økonomiske og organisatoriske spørgsmål, der knytter sig til prioriteringen af indsatsen, dels de tekniske og økonomiske forhold, der er knyttet til spørgsmålet om tilgængeliggørelsen af det digitaliserede materiale for publikum. Det drejer sig her om en tilgængeliggørelse på en sådan måde, at der skabes det bedst mulige grundlag for en formidlingsaktivitet, der imødekommer mange brugergrupperes behov. It-analysevirksomheden Gartner har også medvirket ved denne del af undersøgelsen.

Den tredje fase af udvalgets arbejde (2008-primus 2009) er i høj grad foregået via en nedsat underarbejdsgruppe og har haft til formål at konkludere på udvalgets hidtidige arbejde ved, via den foreliggende rapport, at fremlægge tre forslag til digitalisering af udvalgte, prioriterede dele af kulturarven.

2 Udvalgets overvejelser

Kulturarven er nationens hukommelse. Kulturarvens aktive videreførelse og brug er grundlaget for vor eksistens som en kulturnation. Kulturarven er fundamentet for Danmarks identitet som et videnssamfund og et helt centralt råstof for fremtidens videnøkonomi.

Kulturarv kan være et flygtigt begreb, men kulturarven er også helt fysisk og konkret, i form af arkiver, bogsamlinger, pladesamlinger, kunstsamlinger, filmsamlinger, museumssamlinger og meget mere. Kulturarven er således ikke blot fysisk til stede, men har som helt konkret fysisk materiale et overvældende omfang.

Det er kulturarvens store omfang, og de deraf følgende muligheder for læring, formidling og forskning, der gør, at kulturarven er en *arv* og ikke blot et dødt minde om fortiden. Takket være den omhu, som det danske samfund faktisk i flere hundrede år har udvist med at indsamle og bevare kulturarven, er vor kulturelle arv i dag et materiale langt større, end et menneske kan sætte sig ind i på et helt liv, og som vedblivende er en kilde til ny viden, inspiration og råmateriale til kulturel nyskabelse.

Men kulturarvens vældige fysiske omfang er naturligvis også en udfordring for den, der vil gøre brug af den, uanset om brugeren er forsker, uddannelsessøgende eller blot interesseret borger. Kulturarvsgenstandene i samlingerne på arkiver, biblioteker og museer kan være for skrøbelige til at håndtere eller for kostbare til at give fri adgang til, og det er under alle omstændigheder ressourcekrævende at finde genstandene frem fra store samlinger og arkiver. På den måde er en del af vor fælles kulturarv ofte kun praktisk tilgængelig for dedikerede forskere, og selv for dem kun med stor indsats.

Her åbner den digitale formidling et nyt, hidtil aldeles ukendt perspektiv for, hvor mange mennesker der kan få adgang til, og bruge, kulturarven. Når kulturarven digitaliseres, kan den pludselig gøres tilgængelig med et museklik og bliver dermed tilgængelig for enhver, døgnet rundt, uden begrænsninger, i hele landet – ja, principielt hele verden. Digitalisering af kulturarven åbner nye muligheder for enhver kategori af brugere. Alle vil kunne blive klogere på Danmarks kulturelle og historiske arv.

For den interesserede borger åbner sig et nyt univers med let adgang til langt mere detaljeret viden og til autentiske kilder vedrørende et utal af emner. Det giver stærke kulturelle oplevelser til utallige mennesker og et rigere og mere nuanceret indhold til begreber som national identitet og historisk forståelse. Den særligt interesserede borger fra udlandet, f.eks. den potentielle turist, får også helt nye muligheder for at erhverve sig et nuanceret billede af Danmark – ikke mindst gennem værktøjer som den nye europæiske kulturarvsplatform Europeana.

For den uddannelsessøgende betyder den digitaliserede kulturarv nye læringsoplevelser, med lejlighed til at omgås et materiale af høj videnkvalitet, som samtidig byder på reel frihed til at gå på opdagelse i et udtømmeligt materiale, anvendeligt i et utal af undervisningsforløb. Med den digitaliserede kulturarv til rådighed, med dens ægte dybde og detaljerighed at arbejde med, kan de uddannelsessøgende erhverve sig en helt ny *grad* af kompetencer i forhold til bearbejdning og aktivering af nuanceret viden; kompetencer, der er uundværlige for dem, både som mennesker, som borgere og som fremtidige aktører på arbejdsmarkedet.

For forskeren, der i forvejen bruger kulturarven som råmateriale i sin forskning, kan digitaliseringen ofte betyde et kvantespring. Den lette adgang til kilderne, og ikke mindst adgang til søgninger i dem på tværs af materialetyper med ny kombinationsmuligheder til følge, betyder et helt nyt videnniveau for den enkelte forsker. Det betyder nye muligheder for samarbejde, også på tværs af landegrænser. Hermed kan dansk kulturarv og kilderne til den for alvor træde i karakter som noget, der trækker udenlandske forskere til samarbejde med danske institutioner og styrker Danmarks placering i den globale videnproduktion og -formidling.

For turistene, eller blot den interesserede netbruger verden over, er den digitaliserede kulturarv et åbent vindue til Danmark som spændende ferieland, gammelt kulturland og moderne, innovativt samfund med dybe rødder.

For den professionelle kulturskaber og den kreative iværksætter er kulturarven, som for forskeren, et uundværligt råstof. Men for den nyskabende betyder digitaliseringen et endnu større løft til kulturarvens værdi, for lethed af den digitale formidling giver iværksætteren mulighed for hurtigt at skabe helt nye produkter på nettet.

2.1 Formål med digitalisering af kulturarven

Kulturarvsinstitutioner kan digitalisere materiale med flere formål – først og fremmest bevaring, beskyttelse og tilgængeliggørelse³. Med beskyttelse menes det værn mod slid, uheld og tyveri, som ligger i, at originalmaterialet ikke længere behøver håndteres ved brug. Mens beskyttelse og tilgængeliggørelse meget ofte går hånd i hånd, så er det kun en mindre del af kulturarvsmaterialet, hvor der er behov for egentlig bevaringsdigitalisering.

Størstedelen af kulturarvsmaterialet er, med den rette opbevaring og håndtering, i en sådan stand, at det ikke er akut nødvendigt at tage skridt til digital bevaring. For andre materialer er digitalisering i bevaringsformat endnu ikke teknisk og ressourcemæssigt praktisk (det gælder eksempelvis film på filmruller).

Der er dog også en række kulturarvsmaterialer, der allerede på nuværende tidspunkt befinder sig i en truet tilstand, eller som i det hele taget er meget besværlige og ressourcekrævende at bevare i originaltilstanden. Det er særligt audiovisuelt materiale, dette gør sig gældende for. Der er tale om lyd og levende billeder i et stort antal formater, på mange forskellige typer fysiske medier, som det byder på store vanskeligheder at bevare, både fordi medierne i sig selv er skrøbelige, og fordi udstyret til at afspille dem på udgår – eller allerede er udgået – af almindelig handel.

Bevaringsmæssige problemer eksisterer imidlertid også for andre materialer. Det gælder ikke mindst trykte materialer, der ikke i udgangspunktet er produceret med henblik på langtidsbevaring – aviser, blade, småtryk mv. Samtidig er disse materialer, der netop i kraft af, at de er skabt med henblik på øjeblikkeligt ”forbrug”, er kulturarvsmæssigt unikke, fordi de giver et dag til dag-billede af historien.

³ Institutioner kan tillige digitalisere med henblik på interne arbejdsprocesser. Dette hensyn ligger dog uden for udvalgets overvejelser.

Selvom bevaringshensynet således ikke er negligerbart, når man skal prioritere indsatsen på digitaliseringsområdet, så overskygges det alligevel af hensynet til tilgængeliggørelsen. Hvor store dele af kulturarven med den kendte teknologi bedst og billigst lader sig bevare i original form (som man selvfølgelig normalt også foretrækker at bevare, såfremt det er muligt), så vil en digital tilgængeliggørelse betyde et stort spring i borgernes mulighed for hurtig og nem adgang.

Derfor er dette også det afgørende fokus i udvalgets forslag til prioritering af digitaliseringen. Der er i høj grad taget udgangspunkt i materialer, hvor man af erfaring ved, at der er stor potentiel efterspørgsel, men hvor det er kompliceret at få adgang til originalmaterialet: arkiver, kortsamling, billedsamlinger, småtrykssamlinger, Statsbibliotekets samling af danske aviser mv. Der er tale om materialekategorier, der er samlet ét sted, og hvor søgning, bestilling og fremfinding i dag er besværlig og langsommelig, mens brugstiden ofte er kort, fordi brugeren måske blot har behov for at finde en enkelt reference. Et oplagt eksempel er kirkebøger, hvor folk med interesse for slægtsforskning ofte blot skal bruge én oplysning fra en given kirkebog. Det er materialer, der i digitaliseret form kan bruges på helt nye måder og i helt nye sammenhænge.

Det Kongelige Biblioteks store nationalbibliotekssamling, der omfatter al trykt litteratur udgivet i Danmark og om Danmark siden 1482, er derimod ikke prioriteret, selv ikke i den mest ambitiøse model. Det skyldes dels omfanget, dels at der er nem adgang til de fysiske eksemplarer gennem et funktionelt system for tilgængelighed af bøger med et rimeligt forhold mellem fremfindingstid og brugstid for det enkelte materiale, hvorimod en lang række andre samlingstyper med digitaliserings-teknologien for første gang kan gøres reelt tilgængelige for offentligheden.

Samtidig skal man, jf. afsnit 2.2 nedenfor, se den relativt lave prioritering af bogsamlingerne i lyset af, at store dele heraf typisk er omfattet af ophavsret og dermed relativt kostbare at stille til rådighed, ligesom man kan antage, at en hel del vil blive tilgængeliggjort digitalt af de kommercielle aktører.

Det skal dog understreges, at digitaliseringen også for bibliotekssamlingernes vedkommende gør en afgørende forskel i tilgængelighed og interesse (jf. faktaboks nedenfor), og på sigt er der næppe tvivl om, at digital tilgængelighed er forudsætningen for, at den danske litterære arv kan blive levende, også i den digitale tidsalder. Der er ingen tvivl om, at udvalgte dele af de ældre samlinger allerede inden for rapportens tidshorisont om muligt burde digitaliseres, såfremt der opstår økonomisk rum hertil.

Det drejer sig først og fremmest om de store viden-, kilde- og tekstkorpera, der ligger i faglitterære referenceværker, de centrale faglige og videnskabelige tidsskrifter, historiske og juridiske kildesamlinger, dokument- og brevdgaver samt evt. samlingsdele, der understøtter digitaliseringsinitiativer inden for andre materialetyper. Det er samlinger, der, jf. ovenfor, normalt er karakteriseret ved stort tidsforbrug til fremfinding og ofte relativt kort brugstid.

Generelt skal det bemærkes, at eftersom scenarierne i det følgende alle er bygget op om en tiårig periode, vil det utvivlsomt blive nødvendigt med tilpasninger og konkrete omprioriteringer undervejs. Der kan vise sig udviklinger i teknologien, eller der kan melde sig private samarbejdspartnere, som kan betyde, at andre materialesamlinger pludselig viser sig mere lovende at digitalisere. De indhøstede erfaringer, både fra selve digitaliseringen og fra den efterfølgende brug af materialet, kan også betyde nødvendige omprioriteringer.

Faktaboks: Hittidige erfaringer med digital tilgængeliggørelse

Styrelsen for Bibliotek og Medier har i Folke- og Forskningsbiblioteksstatistik 2007 set på forskellige udviklingstræk i de 16 største forskningsbiblioteker. Det fremgår heraf, at antallet af fysiske udlån eksklusive fornyelser har været faldende i perioden fra 2000 til 2007. I samme periode er de elektroniske "udlån" (download) vokset med 750 %, således at de herefter udgør næsten 12 millioner.

Fra 2002 har man opgjort antallet af download på interne og eksterne servere. Eksterne servere giver adgang til betalingsbelagte materialer, især tidsskrifter, medens materialer på interne servere overvejende er genereret af bibliotekerne og deres moderinstitutioner. Omfanget af download fra eksterne servere er steget med 181 % siden 2002, medens omfanget af download fra interne servere er steget med 406 %. Det er altså især tale om en kraftig vækst i antallet af download fra institutionernes egne digitale data.

For Det Kongelige Biblioteks vedkommende stiger antallet af download fra egne servere fra knap 540.000 i 2002 til godt 1,4 millioner i 2007, medens antallet af eksterne download stiger fra 0,7 million til 2,4 millioner i samme periode.

For Statsbibliotekets vedkommende stiger antallet af interne download fra knap 65.000 i 2002 til knap 780.000 i 2007, medens de eksterne stiger fra 0,6 million til 1,4 millioner. Det kan altså konstateres, at der er en meget stor efterspørgsel efter digitale data, ikke mindst fra bibliotekernes egne digitaliseringer.

For Statens Arkivers vedkommende kan nævnes, at antallet af registrerede bruger af Arkivalieronline steg fra 8.262 i 2004 til 96.216 ved udgang af 2007 og 113.280 ved udgangen af 2008. Der blev fra Arkivalieronline.dk i 2007 foretaget eksterne download af 105 mio. sider; i 2008 steg dette tal til over 151 mio. sider. Arkivalieronline har i gennemsnit ca. 4.500 unikke besøg om dagen. Dansk Demografisk Database, der drives af Statens Arkiver, havde i 2008 i gennemsnit ca. 5.000 unikke besøg om dagen.

2.2 Tilgængeliggørelse af digitaliseret materiale

Udvalgets forslag til digitalisering af kulturarven, som er skildret i de følgende kapitler, er bygget op omkring det helt centrale spørgsmål for en digitaliseringsplan – nemlig hvilket udsnit af det enorme eksisterende kulturarvsmateriale der skal prioriteres digitaliseret⁴. Men herefter skal brugerne naturligvis også have adgang til den digitaliserede kulturarv.

Udvalget skal dog også pege på en række tekniske og organisatoriske forhold, der skal være på plads for at danne grundlag for digitalisering og tilgængeliggørelse af det udpegede materiale. For en uddybning heraf henvises til Bilag: Teknisk redegørelse vedr. digitalisering af kulturarven.

⁴ Til udvalgets midtvejsrapport fra september 2008 (www.kum.dk/sw75986.asp) foretog it-konsulenterne Gartner en analyse, der viste, at en digitalisering af alt digitaliseringsværdigt materiale på kulturarvsinstitutionerne ville koste mellem 2 og 3 mia. kr. med den kendte teknologi.

Forudsætningerne er:

- En effektiv og konsolideret proces for selve digitaliseringen, så resultatet faktisk bliver brugbare digitale filer
- En pålidelig og langtidsholdbar bevaring af de digitale filer
- En konsekvent anvendelse af internationalt anerkendte standarder for filer og metadata, så materialet kan søges og anvendes på flest mulige platforme
- Eksponering af materialet, dvs. at det kan ”ses” af søgemaskiner som Google og Europeana. Effektiv eksponering kræver sandsynligvis, at der også laves en struktur for indsamling af metadata.

Disse forudsætninger kan forekomme indlysende og bør også være indlysende rettesnore i arbejdet. Når det alligevel skal fremhæves her, er det fordi man i praksis har set flere internationale eksempler på, at man i konkrete digitaliseringer *ikke* har fået opfyldt forudsætningerne.

At opfylde forudsætningerne er således ikke selvfølgelig, men vil kræve en infrastruktur mellem kulturarvsinstitutionerne, der sikrer, at de kompetencer, der allerede findes blandt institutionerne, kommer alle institutionerne til gavn, og at man anvender fælles standarder og procedurer. I den tekniske redegørelse peger udvalget derfor på *muligheden for* at oprette egentlige kompetencecentre for de enkelte materialetyper.

Det kan også blive relevant med en fælles bevaringsinfrastruktur og som minimum en fælles indsamling af metadata, som kan sikre, at materialet bliver søgbart og tilgængeligt på tværs af institutioner – dvs. som minimum et ”virtuelt” datawarehouse, gennem hvilket man kan fremsøge et hvilket som helst digitaliseret materiale fra kulturarvsinstitutionerne.

Bilag 2 indeholder en analyse af udgifterne til en minimal infrastruktur baseret på et fælles lager. De samlede udgifter er estimeret til 10,5 mio. kr. årligt. Estimatet for den samlede infrastruktur er udarbejdet under forudsætning af opbygning af en fælles infrastruktur og en vis mængde digitaliseret materiale. Infrastrukturen vil være relevant i forhold til scenarie 2 og 3.

Udgangspunktet for estimatet er en lagerstørrelse på 2500 TB, svarende ca. til scenarie to. Et fuldt udfoldet scenarie 2 vil således indebære udgifter på ca. 10 mio. kr. årligt, mens et fuldt udfoldet scenarie 3 vil indebære udgifter på 10-15 mio. kr. årligt. Det skal bemærkes, at estimeringen af driftsudgifterne bygger på et worst case-scenarie, idet prognoserne for omkostningen til lagringen af TB er meget usikre.

Det bemærkes som nævnt, at udgiftsskønnet er foretaget ud fra et fælles lager. Som det fremgår ovenfor, er et fælles lager ikke en forudsætning for at etablere søgbarhed på tværs, og udvalget ser sig ikke på det foreliggende grundlag i stand til at vurdere, om et fælles lager faktisk er den optimale løsning, og kan således ikke komme med nogen anbefaling herom.

Men med de i bilaget nævnte forbehold angående økonomiskønnet kan det dog konkluderes, at der findes i hvert fald én mulig løsning omkring bevaring og tilgængeliggørelse inden for de i scenarierne angivne økonomiske rammer, og at scenarierne således lader sig realisere som beskrevet.

I en konkret planlægningssituation vil det så kræve en mere grundig overvejelse, om et fælles lager er den optimale løsning, eller der kan peges på bedre, decentrale muligheder.

Infrastrukturen vil omfatte lagring og tilgængeliggørelse af det digitaliserede materiale. Den vil tilbyde tværgående søgning i metadata fra materialerne og understøtte eksisterende formidlingsinitiativer.

Den beskrevne infrastruktur er baseret på en større mængde digitaliseret materiale og vil derfor få en anden udformning i forhold til scenarie 1. Infrastrukturen for scenarie 1 vil hovedsagligt bestå i standardisering af metadata og eksisterende systemer, således at det bliver muligt at høste metadata og eksponere dem for andre systemer. I scenarie 1 vil metadata blive høstet og gjort tilgængeligt for søgning i en fælles national base og blive eksporteret til Europeana.

2.3 Ophavsret

En betydelig del af det materiale, som kulturarvsinstitutionerne ligger inde med, er fortsat omfattet af ophavsret. Dette materiale er normalt det nyeste og oftest det mest attraktive for brugerne, og det understreges ofte, hvor væsentligt det er, at der ikke i tilgængeliggørelsen af kulturarven bliver et såkaldt ”sort hul i det 20. århundrede”, som kan medvirke til at fremmedgøre store grupper over for kulturarvens tilbud.

Danmark er på dette område langt fremme, idet Folketinget i foråret 2008 vedtog en udvidelse af mulighederne for brug af de såkaldte *aftalelicenser* (lov nr. 231 af 8. april 2008)⁵. Denne ordning reducerer i høj grad de administrative omkostninger ved digital tilgængeliggørelse af ophavsretsbeskyttede materialer.

Disse regler giver et stærkt udgangspunkt for, at kulturarvsinstitutionerne kan formidle også ophavsretsbeskyttet materiale fra deres samlinger. Der findes allerede en række kanaler, der kunne være velegnede til formidling af sådant materiale, f.eks. e-museum.

Det er samtidig et område, hvor der tegner sig potentiale for samarbejde mellem kulturarvsinstitutioner og private virksomheder om at stille de fulde samlinger til rådighed i en kombineret offentlig/privat løsning (et godt eksempel er et kommende samarbejde mellem Statsbiblioteket og en privat udbyder om tilgængeliggørelse af Statsbibliotekets samling af danske aviser, se faktaboks nedenfor side 25).

Aftalelicensordningen ændrer dog ikke ved, at tilgængeliggørelse af ophavsretsbeskyttede materialer medfører væsentlige omkostninger for brugere og/eller institution, som ikke optræder, hvis materialet ikke er omfattet af ophavsret, eller hvis institutionen selv er indehaver af den fulde ophavsret.

⁵ En aftalelicens er en tilladelse til, at en organisation kan indgå aftale på vegne af rettighedshaverne til en bestemt art af værker. Sådanne organisationer skal være repræsentative for de berørte rettighedshavere og vil typisk være forvaltningsselskaber som f.eks. Copydan. Dette betyder, at en kulturarvsinstitution i forbindelse med digitalisering kan indgå aftaler om ophavsret med én organisation, hvor de ellers ville have skullet indgå aftaler med tusindvis af individuelle ophavsretshavere.

Derfor skal man i forhold til digitaliseringsforslagene i de følgende kapitler være opmærksom på, at hensynet til sådanne omkostninger har væsentlig betydning for, hvilket materiale der peges på til digitalisering, og en række valg og fravalg må forstås i lyset heraf.

Det skal også bemærkes, at omkostningerne ved tilgængeliggørelse af ophavsretsbeskyttet materiale afhænger af en konkret forhandling, og at det derfor i beskrivelsen af de enkelte scenarier kan være vanskeligt præcist at angive, hvor bredt det kan tilgængeliggøres. Som nævnt ovenfor kan det dog være muligt i sådanne tilfælde at finde kommercielle samarbejdspartnere, der kan gøre materialet tilgængeligt for den brede offentlighed mod vederlag.

2.4 Formidling

Det er udvalgets opfattelse, at hovedopgaven for en national plan for digitalisering af kulturarven må være at sikre, at materialet digitaliseres og gøres reelt tilgængeligt. Dette indebærer, jf. afsnit 2.2, også formidling i form af f.eks. produktion af metadata, der kan eksponere materialet i gængse søgemaskiner.

I forhold til forslag 3 (se afsnit 2.5) kunne det ydermere overvejes at få udarbejdet en form for engelsksproget indholdsoversigt til samlingerne med henblik på at skabe interesse hos udenlandske forskere for samarbejde med danske forskere om brug af danske kilder. Uden et sådant redskab kan det blive svært for potentielle internationale samarbejdspartnere at danne sig et overblik over, hvorvidt der er interessante muligheder. Ideelt burde samlingernes metadata være søgbare på engelsk, men muligheden herfor inden for de økonomiske scenarier afhænger af den teknologiske udvikling i forhold til automatisk oversættelse af metadata.

Spørgsmålet om formidling er naturligvis helt afgørende for den faktiske værdi af materialet for brugerne, men antagelsen må være, at det afgørende er at samordne den tværgående brugbarhed, mens formidlingen trives bedst med decentraliseret initiativ og kreativitet, skabt af institutioner, brugere og private virksomheder.

Derfor er fokus også flyttet fra tanken om en samlet portal, som tidligere har været fremme i udvalgets arbejde. Denne tanke var desuden i midtvejsrapporten forbundet med anbefalingen af at opbygge en kapacitet til integreret søgning, dvs. søgning via indeksering af filernes indhold og ikke blot via deres metadata (det man også kunne kalde fuldtekstsøgning).

Som teknologien er i dag, er det imidlertid vanskeligt at arbejde med sådanne søgninger i andet end netop tekst, og da vægten i udvalgets prioritering i sidste ende i højere grad er blevet lagt på andre typer materiale (billeder, lyd, levende billeder), er relevansen af at etablere en portal med integreret søgning også formindsket.

Der findes også allerede forskellige sektorspecifikke portaler til formidling til bestemte brugergrupper, f.eks. EMU og herunder e-museum, rettet mod bl.a. grundskoler og ungdomsuddannelser, som der kan samarbejdes med.

Udvalget har valgt at fremhæve forskellige eksempler på mulige formidlings-værktøjer, som kunne tage udgangspunkt i det tilgængeliggjorte digitale kulturarvsmateriale i de forskellige scenarier. Men

det afgørende fokus for udvalget har været selve digitaliseringen og tilgængeliggørelsen. Formidlingseksemplerne skal således alene ses som eksempler.

Faktaboks: Et afprøvet formidlingsredskab – dansk.kultur.arv.dk

På www.dansk.kultur.arv.dk gik syv af Danmarks store kulturinstitutioner sammen om (bl.a.) en undervisningscase på tværs af institutioner, samlinger, tid og rum, hvor man ønskede at eksperimentere med, hvordan den fælles kulturarv kan understøtte undervisning og indlæring i folkeskolerne. Alle syv institutioner bidrog med digitale materialer såsom billeder, video, tekst og lyd, udvalgt på baggrund af et bestemt tema. Der er i princippet ingen begrænsninger i temaer eller tilrettelæggelse af grænseflader, hvis der ligger nok digitaliseret materiale til grund.

Grundtanken var at bygge undervisningen op omkring onlineaktiviteter som eleverne kender fra deres fritid. Der blev derfor udviklet en undervisningsgrænseflade efter inspiration fra bl.a. iTunes og flickrs mange interaktive muligheder, hvor man kan "tagge" materiale, tilføje egne beskrivelser, læse andres beskrivelser, sammensætte egne samlinger (playlister) af udvalgt materiale samt lade sig inspirere af andres playlister. Derudover blev en lille håndfuld kendisser, som eleverne ville kunne relatere til, bedt om at lave en playlist til inspiration og opdagelse for eleverne.

Sammen med undervisningskonsulenter blev der tilrettelagt tre onlineundervisningsforløb for at gøre det let for lærerne at inddrage forløbene i deres undervisning. Hvert undervisningsforløb indeholdt opgaver og spørgsmål, som skulle besvares af eleverne vha. materiale fra hjemmesiden. En videreudvikling af ideen kunne omfatte muligheden for, at lærerne selv udvikler nye undervisningsforløb og lægger dem ud til fælles brug, eller eksport af indholdet ud i andre grænseflader, så det kan genbruges og mikses på nye måder.

For nærmere beskrivelse, download af undervisningsmateriale samt brugerundersøgelse, se: www.dansk.kultur.arv.dk (hjemmesiden er fortsat åben, men ikke længere aktiv).

2.5 Baggrund for udvalgets forslag

I de følgende tre kapitler skitseres tre forskellige forslag fra udvalget til, hvilke dele af kulturarven man, under forskellige økonomiske forudsætninger, kan prioritere at digitalisere.

De tre forslag bygger oven på hinanden, i den forstand at forslag 3 også indeholder materialet i forslag 2, som igen indeholder materialet i forslag 1. Udvalget har med forslagene valgt at tage udgangspunkt i tre forskellige ambitionsniveauer, dvs. i tre forskellige niveauer for statslig økonomisk indsats på området.

Udvalget skal bemærke, at der er stordriftsfordele i selve arbejdsprocesserne omkring digitalisering, ligesom tilgængeligheden af en "kritisk masse" af samlinger har betydning for borgernes tilfredsstillende brug af materialet, hvorfor der både hvad angår kvalitet og kvantitet er et stort spring i resultat fra forslag 1 til forslag 2 og igen fra forslag 2 til forslag 3.

Det første forslag kan, i overensstemmelse med udvalgets kommissorium, finansieres inden for institutionernes ordinære bevillinger. Udvalget har med forslag 2 og forslag 3 valgt at se på nogle forholdsvis ambitiøse økonomiske scenarier, hvortil udgifterne udgør henholdsvis i alt 300 mio. kr. og 570 mio. kr. i perioden 2010-2019, inkl. udgifter til tilgængeliggørelse. Man kan ikke på et fagligt grundlag pege på ét rigtigt ressourcemæssigt ambitionsniveau; det er et politisk prioriterings spørgsmål.

Afklaringen af finansieringen af udgifterne til gennemførelsen af forslag 2 og forslag 3 udestår, men vil skulle overvejes i sammenhæng med valg af model.

De økonomiske scenarier såvel som tidshorizonten skal naturligvis ses på baggrund af digitaliseringsopgavens væsentlighed og store omfang, jf. midtvejsrapportens analyse af det samlede digitaliseringspotentiale. Således skal man være opmærksom på, at man selv i den mest ambitiøse model kun kan håndtere en delmængde af det samlede digitaliseringspotentiale, og at mange ting ikke vil kunne indbefattes (f.eks. fag- og skønlitteraturen, jf. ovenfor).

Samtidig skal man være opmærksom på den store indsats, der foregår i mange lande, ikke mindst de engelsksprogede. Indsatsen i f.eks. USA skal selvfølgelig ses i forhold til landets størrelse og foregår i høj grad via kommercielle aktører som eksempelvis Google.

Men også i europæiske lande sker der en indsats. Norge har eksempelvis påbegyndt en digitalisering af den samlede trykte nationallitteratur. Frankrigs nationalbibliotek har indledt et program, der skal digitalisere 150.000 trykte bog- og tidsskriftsenheder om året (komparativt svarende til 15.000 enheder om året i Danmark). Og Nederlandene har afsat i alt 170 mio. euro til digitalisering af audiovisuelle arkiver (film, tv og radio).

Der kan være en risiko for, at dansk kulturarv i både synlighed og brug bliver marginaliseret i forhold til engelsksproget (og anden fremmedsproget) kulturarv på nettet, også for danske brugere, med negative konsekvenser for dansk sprog og identitet, dansk kulturproduktion og innovation og dansk forskning inden for samfundsvidenskabelige, humanistiske og æstetiske fag.

Faktaboks: KulturPerler – ny national portal, der beskriver og registrerer digitaliserede ressourcer og samlinger i Danmark

Er de unummerede betænkninger i Danmark digitaliserede? Hvad er der digitaliseret, der kunne være relevant for min undervisning i faget historie eller dansk i gymnasiet? Er der noget i naturvidenskab? Hvem har digitaliseret noder? Hvilke danske fagtidsskrifter kan findes online? Hvad har Dansk Folkemindesamling digitaliseret? Disse og lignende spørgsmål har været svære eller umulige at svare sikkert, udtømmende og autoritativt på indtil nu, eftersom de ikke kan besvares inden for vore traditionelle katalogbaser, registre og søgesystemer, bl.a. fordi en fast katalognorm kun findes på biblioteksområdet.

Som første resultat af Kulturministeriets arbejde med en planlægning af digitalisering af kulturarven giver Det Kongelige Bibliotek på sin hjemmeside www.kb.dk fra 2009 adgang til "KulturPerler" en helt ny digital, nationalbibliografisk tjeneste, der har til opgave at beskrive og registrere den digitaliserede kulturarv i Danmark, dvs. ressourcer og samlinger af alle arter, efterhånden som disse retrodigitaliseres. Tjenesten er tværsektoriel og -institutionel og omfatter, hvad der er retrodigitaliseret af både biblioteker, arkiver, museer og andre institutioner, f.eks. Det Danske Sprog- og Litteraturselskab, Kort- og Matrikelstyrelsen og Danmarks Radio. Det er den første af sin art internationalt set, og konceptet har allerede vakt interesse i udlandet.

"KulturPerler.kb.dk" sigter på at give både et overblik over digitaliseringssituationen i Danmark og systematiske, faktuelle oplysninger om de enkelte digitaliserede samlinger. Det Kongelige Bibliotek har stået for udarbejdelsen i samarbejde med de relevante institutioner, men fremover skal tjenesten være interaktiv, i den forstand at institutioner, personer eller organisationer, der digitaliserer eller er ansvarlig for digitalisering, skal kunne indberette (selvangive) til tjenesten i et fast format med beskrivelsesforslag til en beskrivelsesdel og oplysninger til en katalogdel. Det endelige redaktionelle ansvar ligger dog hos Det Kongelige Bibliotek.

Første del, oversigterne, sigter især på *det politisk-administrative* niveau, brugere/borgere og fagfolk, der vil have et overblik over, *hvad* der er digitaliseret i Danmark i enten *institutionelt* regi, ud fra en *faglig* eller en *materialetypologisk* synsvinkel, mens anden del, katalogdelen, sigter på de sædvanlige brugere af bibliotekernes ressourcer, herunder bibliotekarer m.fl.

En sådan fortegnelse, der kan give såvel overblik som systematiske oplysninger om den enkelte digitaliserede samling, findes ikke i dag og falder uden for rammerne af de eksisterende nationalbibliografiske fortegnelser, biblioteks-, museums- og andre institutionskataloger. Den skal være instrumentet til at følge med i udviklingen af digitaliseringssituationen i Danmark, der i de forløbne år har fået flere betegnelser i den offentlige debat, f.eks. "dansk på nettet", "kulturarven på nettet" mv. Fortegnelsen skal i første omgang *ikke* omfatte *digitalt fødte*, herunder kommercielt producerede og publicerede, ressourcer.

"KulturPerler.kb.dk" skal således omfatte både allerede digitaliseret kulturarvsmateriale og fremtidig digitaliseringsindsats, både i form af massedigitalisering og selektiv digitalisering uanset form og begrænsninger i Danmark. Allerede nu indgår ca. 170 samlinger af meget forskellig størrelse.

3 Forslag 1

3.1 Udvalgets forslag

Inden for de økonomiske rammer, som sættes af institutionernes ordinære bevillinger, foreslår udvalget digitalisering af nedenstående materialer i perioden 2010-19⁶:

Trykte tekster, arkivalier, håndskrifter mv.

- 20.000 tekstsider fra artikel- og dokumentsamling (DFI), som belyser dansk filmhistorie, og som kan bruges af både uddannelsessøgende, interesserede borgere og forskere som kontekstmateriale for instituttets øvrige samlinger.
- Folketællingsskemaer (ca. 4 mio. sider) fra tællingerne i 1901, 1906, 1911, 1921 og 1925 og kirkebøger 1925-1960 (ca. 2 mio. sider). Materialet er relevant for såvel demografiske som regionale og lokale studier, hvad enten udgangspunktet er en lokalitet eller en population. Der er tale om komplette digitaliseringer med meget brede anvendelsesmuligheder for forskere. Samtidig er materialet stærkt efterspurgt af den meget store gruppe af slægtsforskere. De digitaliserede kirkebøger og folketællinger, der allerede findes på arkivalieronline.dk, tiltrækker dagligt 4.-5.000 unikke besøg.
- 370.000 avissider fra *Adresseavisen* 1759-1909 (SB). En digitalisering vil gøre dette efterspurgte materiale bedre tilgængeligt for forskere og tilgængeligt i det hele taget for kulturinteresserede borgere, ligesom det har oplagte anvendelsesmuligheder i en undervisningssammenhæng.
- 93.000 sider atlas, musiktidsskrifter, teaterprogrammer og småtryk (KB). Der er tale om et materiale, der i analog form er besværligt at finde frem, men som fra talrige vinkler giver et godt billede af tilværelsen i Danmark gennem årene. Inden for rammerne af forslag 1 bliver der tale om et relativt begrænset udpluk, der kan hæmme den forskningsmæssige værdi, men interessen for den almindelige borger og værdien i en undervisningssituation vil være betydelig.
- 110.000 sider manuskripter og breve mv. (KB). Ligesom med småtryk er der tale om et materiale, som kan være besværligt at bruge i analog form, men som i alment tilgængelig form kan have stor interesse i mange sammenhænge.

Billeder

- 50.000 stills fra danske film (DFI); digitaliseres løbende (bl.a. efter forespørgsel). Billederne gøres tilgængelige både i lav opløsning frit på internettet og i høj opløsning via

⁶ Oplysningerne vedr. DR's samlinger angår dog den indeværende medieforligsperiode 2007-10, hvor DR har en særlig licensbevilling på 75 mio. kr. til digitalisering.

betalingsservice, til professionelle brugere som medier og forlag mv. Billederne vil blive digitaliseret i fuld bevaringskvalitet.

- 20.000 stills fra danske kort- og dokumentarfilm (DFI); digitaliseres løbende, især ud af den centrale samling fra statslige producenter (Statens Filmcentral mv.). Den digitale samling af stills kan f.eks. i forhold til undervisningssektoren virke som supplement til brugen af kort- og dokumentarfilm i undervisningen (eksempelvis som råstof til elevproduceret materiale), ligesom samlingen også kan være relevant for forskere (i denne model er der dog kun tale om en del af samlingen på ca. 64.000 stills fra kort- og dokumentarfilm, hvorfor forskere fortsat vil være nødt til i høj grad at bruge de analoge faciliteter).
- Udvalgte udpluk af Bygningsfredningsarkivet og Arkivet over Fredede Fortidsminder (KUAS). Der er tale om Kulturarvsstyrelsens arbejdsarkiv, som rummer en enestående dokumentation af historiske bygninger og lokaliteter i Danmark. Inden for rammerne af forslag 1 vil der kunne digitaliseres et mindre udvalg af materiale, primært kort, tegninger og billeder, som er mest interessant formidlingsmæssigt, og som også er det mest bevaringsmæssigt udsatte.
- 15.000 af Den Kongelige Kobberstikssamlings værker (SMK – i øjeblikket mangler 218.000 af samlingens i alt 245.000 værker at blive digitaliseret). Der er tale om museets i international målestok bedste og mest efterspurgte samling. Digitalisering af 15.000 værker vil tilgængeliggøre mindre, udvalgte dele af samlingen for offentligheden, mindske slitage på nogle af de mest efterspurgte værker og samtidig gøre de mest efterspurgte dele af samlingen internationalt tilgængelige, til gavn for museets placering i internationale forskningssammenhænge (se også faktaboks side 27).
- 52.000 fotografier (KB). Det Kongelige Bibliotek er nationalbibliotek for fotografier og ligger dermed inde med fotos i en lang række kategorier. Inden for rammerne af forslag 1 er der tale om et udpluk, som ikke desto mindre kan give et indblik i en lang række historiske, geografiske og samfundsmæssige emner til stor interesse for den almindelige borger.
- 50.000 billeder fra analoge billedsamlinger (NATMUS). Billederne er overvejende topografiske og stammer fra museets forskellige arbejdsområder. Både billeder af museumsgenstande, ekspeditionsbilleder fra f.eks. Thuleekspeditionerne og kulturhistorisk dokumentation fra f.eks. Købstadsundersøgelserne vil blive digitaliseret som punktnedslag.
- 1.000 billeder om Jellingmonumentets historie (NATMUS). I forbindelse med forsknings- og formidlingsprojektet *Jellingprojektet – et kongelige monument i dansk og europæisk belysning* digitaliserer NM de gamle analoge billeder, som ikke blot viser de arkæologiske udgravninger, men også monumenternes rolle som besøgsål og nationalt mindesmærke.
- 5.000 optagelser af museumsgenstande fra Danmarks Oldtid (NATMUS). Genstandene er såkaldte ENB-genstande (enestående national betydning), og ca. 10 % af billederne vil blive offentliggjort på museets hjemmeside som 3D-genstande, dvs. som billeder, der kan roteres og zoomes. På den måde vil både forskere og den almindelige kulturinteresserede person på nært hold kunne opleve oldtidens håndværksmæssige kunnen.

- Kalkmalerier i 327 af landets sognekirker (NATMUS). Formålet med projekt *Bevaringsplan for kalkmalerier i de danske kirker* er at skabe et styringsredskab i form af en database, som giver et samlet overblik over kalkmaleriernes bevaringstilstand. Målgruppen er primært de kirkelige myndigheder med ansvar for restaurering; men databasen vil også være interessant for alle, som er interesseret i den enestående kulturarv, som kalkmalerierne i de danske kirker repræsenterer, idet det vil være muligt at søge f.eks. på motiv. Omfanget svarer til ca. 30 % af samtlige danske kirkers kalkmalerier.

Musik og lyd

- 15.000 pligtafleverede cd'er (SB). Der er tale om løbende digitalisering af indkomne materialer, som i overensstemmelse med internationale anbefalinger for bevaring af cd'er skal sikres allerede ved modtagelsen pga. mediets ringe holdbarhed. Der er derfor primært tale om digitalisering med henblik på at bevare den komplette samling for eftertiden. En del af den efterspurgte nyudgivne musik bliver under alle omstændigheder gjort digitalt tilgængelig på kommercielle vilkår. Men tilgængeliggørelse af den fulde musiksamling, herunder også den mere smalle musik, vil være relevant, eksempelvis i en udbygning af Netmusik.dk, og kan håndteres inden for de nye regler om aftalelicens.
- 3.000 pligtafleverede lydbøger (SB). Der er tale om både løbende afleveringer og om indhentning af efterslæb i digitaliseringen. Ligesom hvad angår musik-cd'er, er der tale om en høj prioritet i bevaringshenseende, med henblik på at bevare en komplet samling for eftertiden.
- 600 unikke danske vokswalser med privatoptagelser fra perioden 1900-1930 (SB). Der er tale om et mindre, men interessant materiale fra lydmediets barndom i Danmark, der i sin nuværende form er utilgængeligt for formidling, da der er tale om skrøbelige fysiske medier, som kræver helt specielt afspilningsudstyr.
- 4.000 dialektoptagelser, primært spolebånd (SB). Et bevaringsmæssigt truet materiale, der erfaringsmæssigt har meget stor interesse blandt den almene befolkning, og som også har meget stor relevans både i uddannelsessituationer og for forskere, og som vil kunne stilles til rådighed via en aftale efter den nye aftalelicensordning.

Formidlingseksempel: Dialektoptagelser i undervisningen

Det danske sprog er under konstant forandring. I dag går udviklingen endog meget hurtigt. Der er bekymring for, om det danske sprog overlever, i lyset af mediernes påvirkning og globaliseringen.

Det har sat fokus på nødvendigheden af at styrke danskundervisningen og give eleverne en forståelse af det danske sprogs historie. Dansk er et mangfoldigt sprog med mange dialekter, og hver dialekt har knyttet sin egen kulturhistorie til sig.

Hvordan talte man i gamle dage i Sønderjylland og på Bornholm? Og hvorfor talte man lige sådan?

Det kan man få svar på ved at høre de dialektoptagelser, som Statsbiblioteket i samarbejde med Københavns Universitet og Aarhus Universitet ønsker digitaliseret og stillet til rådighed for forskning og undervisning. Optagelserne dækker dialekter fra hele landet med eksempler på, hvordan folk, der er født i begyndelsen af 1900-tallet, talte. Det er hverdagen i gamle dage, der fortælles om, og vi får derfor også et indblik i livsvilkårene dengang.

DIALEKT-OMRÅDER I DANMARK



Kilde: *dialekt.dk*

Ved klik på et Danmarkskort kan eleverne samtidig med at de lytter til dialekten klikke videre til uddybende forklaringer på de enkelte egenes dialekttræk og karakteristika.

- 3.300 musikoptagelser med uropførelser af ny dansk kompositionsmusik og jazz fra perioden 1979-2002 (SB). En unik samling med stor potentiel værdi for musikforskere og evt. via samarbejde med kommercielle partnere også for almindeligt musikinteresserede.
- 54.000 sider noder (KB). Med tværgående søgning kan udpluk af nodesamlingen have stor værdi i samspil med samlinger af musik for både forskere og almindeligt kulturinteresserede borgere.
- Ca. 355.000 timer lydoptagelser fra arkiv (DR). Materialet ligger på spolebånd og DAT-bånd og indeholder samtlige radioprogrammer fra 1989 og frem til i dag. Der er således tale om et materiale med meget stor samtidshistorisk interesse for både specialister og alment interesserede.

Levende billeder

- 500 stumfilm fra 1896-1930 (DFI), størstedelen dokumentariske. Der er tale om dokumentariske optagelser fra starten af 1900-tallet samt især kortere fiktionsfilm, som tilsammen tilbyder et vindue til Danmark for 100 år siden. Mange af filmene er ikke længere omfattet af ophavsret og kan dermed stilles frit til rådighed på internettet. Flere film portrætterer kendte danskere fra datidens kulturliv, politik, samfund og kongehus. Der er eksempelvis optagelser med Konseilspræsident C. Th. Zahle (1912), Maleren Kristian Zahrtmann (1913), samt Zar Nikolai den 2. (1901). Filmene vil indgå i projektet European Film Gateway.
- Skønsmæssigt 28.000 pligtafleverede VHS/dvd'er (SB). Der er tale om både løbende afleveringer og om indhentning af efterslæb i digitaliseringen. Der er tale om digitalisering med henblik på at bevare den komplette samling. Bevaring af billedmedierne i deres nuværende form er en stor udfordring, både fysisk og formatmæssigt, og bevaringsdigitalisering har derfor meget høj prioritet.
- Tv-udsendelser *on demand* (SB). Der er ikke inden for rammerne af forslag 1 ressourcer til en fuld digitalisering af Statsbibliotekets store samling af komplette danske sendeflader fra DR, TV 2 og regionalstationerne, men der vil kunne ske digitalisering af i alt ca. 10.000 enheder, i takt med at de efterspørges.
- Ca. 52.000 timer timers videooptagelser fra arkiv (DR). Materialet findes på en række forskellige videoformater og udgør en stor del af DR's sendeflade fra slutningen af firserne og frem, plus en del ældre udsendelser digitaliseret *on demand*.

3.2 Udvalgets bemærkninger

Digitaliseringen inden for rammerne af forslag 1 tager udgangspunkt i en udvælgelse af særlige materialer, og der vil kun i begrænset omfang være tale om digitalisering af komplette samlinger. I en del tilfælde må bevaringshensynet overskygge formidlingshensynet, fordi kulturarvsinstitutionernes første prioritet er at bevare samlingerne, så de efterfølgende kan tilgængeliggøres og formidles.

Fordelen ved forslag 1, når man taler om forholdet mellem indsats og udbytte, er, at man fokuserer på det mest efterspurgte og efter en forhåndsvurdering mest interessante.

Ulempen ligger omvendt i en manglende skala. Der udnyttes således ikke den fulde effektivitet i selve digitaliseringen og den tekniske bevaring og tilgængeliggørelse, men især har det betydning for den kritiske masse på selve indholdssiden.

Som det fremgår af det tekniske bilag, kan man, også inden for forslag 1, gøre rigtig meget for en forbedret tværgående søgbarhed og brugbarhed, blot ved den rette anvendelse af standarder og metadata. Men den begrænsede volumen i digitaliseringen betyder, at der kan mangle kritisk masse i forhold til at give brugeren mulighed for at bevæge sig frit i materialet. Man kan komme til at gå forgæves, og det er vanskeligt for materialet at give fuld belysning af samlede temaer – uanset

tværgående søgemuligheder vil materialet set fra brugerens synspunkt i nogen grad fortsat fremstå som enkeltsamlinger.

For forskere kan dette nedsætte brugbarheden – hvilket kan have betydning i forhold til at bruge digitaliseringen til at styrke den danske position i den internationale forskning inden for kulturelle og æstetiske fag.

Men også i en undervisningssituation kan det have betydning, at volumen mangler. Muligheden for at bevæge sig frit rundt i dansk kulturarv vil være indskrænket til udvalgt materiale.

Sammenfattende er der tale om et forslag, der giver et spændende og værdifuldt indblik i kulturarven og sikrer enkelte værdifulde samlinger for eftertiden, idet forslag 1 dog samtidig indebærer betydelige fravalg samt begrænset mulighed for fuldt ud at realisere de stordriftsfordele, der er ved model 2 og 3.

Faktaboks: Indhøstede erfaringer – DR's Kulturarvsprojekt

DR modtog 75 mio. kr. i forbindelse med den seneste mediepolitiske aftale. Disse midler anvendes til at påbegynde sikringen og digitaliseringen af dele af DR's audiovisuelle arkiver. Samtidig indgår DR i en række afklarende projekter for at undersøge perspektiverne i den fælles danske digitale kulturarv.

DR's kulturarvsprojekt har således ageret pilotprojekt i samarbejde med RUNAS institutionerne og DFI i forbindelse med både undersøgelser og test af formidlingsløsninger, metadata og filformater. Se mere herom i faktaboks side 14 eller på www.danskkulturarv.dk.

Overordnet kan aktiviteterne i DR's Kulturarvsprojekt inddeles i to kategorier:

Massedigitalisering og sikring

DR's arkiv består af mange forskellige samlinger og formater og har en volumen og beskaffenhed, der i nogle tilfælde muliggør, at eksterne leverandører kan byde på selve digitaliseringsopgaven på konkurrencedygtige vilkår. DR er i gang med to sådanne eksterne samarbejder, der begge vurderes at medføre økonomiske besparelser.

Samtidig oplever DR, at den teknologi, der er brug for, bliver billigere over tid, og omkostningerne reduceres derfor løbende.

Formidling og samarbejde

I forbindelse med projektet har DR samarbejdet med en hel række andre kulturarvsinstitutioner om www.danskkulturarv.dk (se faktaboks side 14 ovenfor).

Et andet og meget succesfuldt tiltag har været **Bonanza** – *det bedste og det værste* fra DR's arkiver som blev lanceret med udgangspunkt i rettighedsaftalen mellem DR og Copydan. Ideen var at inddrage brugerne i en kæmpe afstemning om, hvor digitaliseringen og tilgængeliggørelsen skulle starte.

Alle danskerne blev inviteret med, og brugerne fik mulighed for at vælge deres favoritter inden for 10 kategorier, hvor der samlet blev vist smagsprøver fra 1.000 forskellige programrækker, shows, nyheder, dramaserier og meget mere. Over 100.000 mennesker deltog i afstemningen, og hjemmesiden har nu haft mere end 12 mio. visninger på under et år. Interessen har været enorm, og hjemmesiden er en kulturperle, der med al tydelighed viser potentialer ved at flytte materialerne op fra kælderen og ud til brugerne gennem en digitaliserings- og sikringsproces.

Resultaterne af DR's undersøgelser og analyser, og erfaringerne fra alle formidlings- og samarbejdsprojekterne, viser, at brugerne meget gerne vil have adgang til materialerne. De ønsker at gense eller opleve dem og gerne på tværs af samlinger. Materialerne beriger hinanden, og samlingerne rummer et nærmest uendeligt socioøkonomisk potentiale, hvis synergieffekterne udnyttes.

Materialerne og brugerne kan skabe relationer på tværs af samlingerne, hvilket udvider oplevelserne for brugerne. Brugerne er aktive og vil gerne bidrage gennem forskellige former for deltagelse. Hvis den moderne teknologi indarbejdes rigtigt, kan denne deltagelse fremadrettet både give brugerne adgang og oplevelser i en positiv spiral.

4 Forslag 2

4.1 Udvalgets forslag

Inden for de økonomiske rammer af en øremærket bevilling på i alt 250 mio. kr. i 2010-2019 foreslår udvalget digitalisering af nedenstående materialer i perioden 2010-19.

Hertil kommer udgifter til tilgængeliggørelse, der løbende vil stige op til ca. 10 mio. kr. årligt ved udløbet af perioden, jf. afsnit 2.2.

Samlet indebærer forslag 2 således merudgifter på i alt 300 mio. kr. i perioden 2010-19⁷.

Der er tale om følgende materiale ud over det, der er nævnt i forslag 1:

Trykte tekster, arkivalier, håndskrifter mv.

- 30.000 tekstsider fra særsamlinger om markante danske filmpersonligheder (DFI). Der er tale om samlinger vedr. Carl Th. Dreyer, Asta Nielsen og Theodor Christensen, nogle af de, ikke mindst i international sammenhæng, mest interessante danske filmpersonligheder, og digitaliseringen vil dermed være forskningsmæssig væsentlig. Der er også tale om filmpersonligheder, som er oplagte at beskæftige sig med i undervisningssammenhæng.
- Arkivalier vedr. skiftevesen (ca. 7,5 mio. sider). Skiftearkivalierne knytter sig til de kirkebøger og folketællinger, som i dag er tilgængelige via nettet. Skiftearkivalier indeholder især for 1900-tallets vedkommende oplysninger, som stadig kan have betydning i forbindelse med dokumentation af arveretlige forhold. Desuden spiller skiftearkivalier en rolle for slægtsforskere, idet skifterne indeholder oplysninger om afdøde personers familiemæssige og økonomiske forhold. Skiftearkivalier findes flere forskellige steder i arkiverne, hvor de er arkiveret efter arkivskaber, først og fremmest
 - (Kriminal)retter
 - By-, herreds- og birkefogeder
 - Amtmænd
 - Godsejere
 - Provster
 - Særlige skifteretter

Der er tale om et meget omfattende materiale, i alt ca. 7,5 mio. optagelser. Heraf påregnes ca. 5 mio. at kunne skannes på grundlag af eksisterende mikrofilm.

- Udvalgte myndighedsskabte søgemidler fra centraladministration om amter. (ca. 2,5 mio. sider) Under betegnelsen ”myndighedsskabte søgemidler” gemmer sig myndighedernes

⁷ Dette er baseret på den forudsætning, at omkostningerne til digital bevaring og tilgængeliggørelse – der afhænger af det digitale materiales omfang – er jævnt stigende over perioden, i takt med at materiale digitaliseres; dvs. i gennemsnit er 5 mio. kr. pr. år. Det bemærkes, at omkostninger til bevaring og tilgængeliggørelse er vanskelige at skønne på længere sigt, da den teknologiske udvikling løbende reducerer omkostningerne pr. lagerenhed. Hvilken tilgængelighedskvalitet man vil kunne få for de i forslaget afsatte midler, er derfor ikke muligt at sige præcist på nuværende tidspunkt.

fortegnelser og registre over deres sager og dokumenter. Søgemidlerne er indgangen til arkivernes indhold, fremstillet af de offentlige myndigheder til eget brug. Søgemidlerne består bl.a. af journaler, navnregistre, postlister og brevregistre. Udtrykt på anden vis omfatter ”søgemidler ” de myndighedsskabte metadata til de dokumenter, der indgår i den offentlige forvaltnings arkivalier i Statens Arkiver. I dag kan de myndighedsskabte søgemidler kun benyttes på læsesalen ved den arkivinstitution, hvortil de er afleveret. Søgemidler gøres af en række bevaringsmæssige og praktiske hensyn ikke til genstand for fjernlån. Ved at lægge udvalgte myndighedsskabte søgemidler på nettet kan det opnås, at en langt større del af publikum får mulighed for at gennemføre den søgeproces, der er nødvendig for at identificere netop de oplysninger, der efterspørges. Gøres journaler, journalregistre, brevregistre mv. tilgængelige digitalt, får brugerne mulighed for hjemmefra at fremfinde konkrete sager, som derefter via Statens Arkivers elektroniske søgedatabase Daisy vil kunne forudbestilles. Kun søgemidler, der er umiddelbart tilgængelige, dvs. over 75 år gamle, skannes.

- 11 mio. avissider; 11 landsdækkende og 10 regionale aviser fra perioden 1749-1990 (SB). Projektet planlægges af Statsbiblioteket og ønskes gennemført i samarbejde med en privat udbyder.

Faktaboks: Avisdigitaliseringsprojektet

11 millioner avissider fra de sidste 250 år ligger i dag omhyggeligt bevaret men vanskeligt tilgængelige i Statsbibliotekets magasin i Århus. Aviserne afspejler dag for dag – helt tilbage fra 1749 – tidens internationale, nationale og lokale begivenheder og fortæller samtidig detaljeret om det danske demokratis udvikling og om dagliglivet på landet og i byerne gennem tiderne.

Disse 11 millioner avissider indgår i projektet *Avisbibliotek Danmark*, som Statsbiblioteket i fællesskab med en privat udbyder ønsker at gennemføre. Det drejer sig om 11 landsdækkende og 10 regionale dagblade fra deres respektive start. Aviserne skal digitaliseres og tilgængeliggøres på nettet til glæde for forskning, undervisning, erhvervsliv, herunder medierne selv, samt for alle danske borgere. Anslåede projektudgifter er 58 mio. kr.

Forretningsmodellen

Projektet rummer store værdier, der kan realiseres både ud fra en samfundsmæssig og en kommerciel synsvinkel. Statsbiblioteket ønsker at digitalisere det historiske materiale med henblik på at kunne stille det til rådighed for den danske befolkning som ’public service’, mens den private udbyder skal kunne kapitalisere på indhold, dvs. en kommerciel model, hvor indholdet kan sælges til både private, erhvervslivet, biblioteker og undervisningsinstitutioner. Markedet skal kunne udnytte den del af materialet, som har kommercielt potentiale, mens materiale med lidt eller intet kommercielt potentiale skal stilles gratis til rådighed for alle.

Det kan ske ved at den kommercielle partner efter aftale med rettighedshaverne får kommerciel råderet over aviser fra nyere tid, mens Statsbiblioteket får råderet over det ældre historiske materiale, som stilles gratis til rådighed for hele den danske befolkning.

Med en statslig finansiering, kombineret med det kommercielle potentiale og bidrag fra fonde, vurderes det, at projektet vil kunne realiseres inden for rammerne af udvalgets forslag 2.

- 40.000 sider dansk fotografisk litteratur (KB). Der er tale om litteratur, som kan supplere og perspektivere det store billedmateriale, som Det Kongelige Bibliotek vil kunne digitalisere inden for rammerne af forslag 2.
- 1,4 mio. sider dansk topografisk litteratur (KB). Materialet vil kunne perspektivere det store topografiske billedmateriale, som kan digitaliseres inden for rammerne af forslag 2.
- 650.000 sider tekst, mindst 325.000 tegninger og mere end 300.000 fotografier, som tilsammen udgør Antikvarisk-Topografisk Arkivs hovedarkiv (NATMUS). Det er et af museets største arkiver, hvad angår oplysninger om og billeder af landets middelalderlige og nyere fortidsminder såsom kirker, borge og voldsteder, herregårde og købstæder. Alt materialet er placeret topografisk i sogne (2.130), herreder og amter efter inddelingen i Trap Danmarks 4. udg. Arkivet rummer beretningsmateriale, tegninger, akvareller, stik og fotografier. Arkivet omfatter ikke alene det nuværende Danmark, men har også materiale fra Færøerne, Grønland, Island, Skåne og Slesvig-Holsten. Til de ældre beretninger om f.eks. kalkmalerier hører akvareller udført af anerkendte konservatorer og kunstnere som J. Magnus Petersen og J. Kornerup. I enkelte tilfælde er disse akvareller eneste vidnesbyrd om kalkmalerier, som i dag ikke længere eksisterer. Arkivet er meget benyttet både af interne og eksterne brugere.
- 410.000 sider erindringer nedskrevet af ca. 41.000 almindelige mennesker indsamlet siden 1939 (NATMUS). Erindringerne findes i beretningsarkivet *Nationalmuseets Etnologiske Undersøgelser*. Erindringerne kan dels være komplette levnedbeskrivelser, dels besvarelser af spørgelister om specielle emner. Materialet er et enestående fond af både sproglig, kulturel og historisk interesse.

Billeder

- 44.000 stillbilleder fra kort- og dokumentarfilm (DFI). Jf. forslag 1 betyder dette, at hele filminstitutets samling af stills fra kort- og dokumentarfilm digitaliseres. Den fulde digitalisering betyder en væsentlig forøgelse af værdien for forskere, og også en betydelig forøgelse af mulighederne for brug i undervisningssammenhæng.
- Dele af Bygningsfredningsarkivet og Arkivet over Fredede Fortidsminder (KUAS), jf. forslag 1. Inden for rammerne af forslag 2 vil det være muligt at digitalisere væsentlige dele af materialet (om end materialets ukurante karakter af arbejdsarkiv gør det vanskeligt at vurdere præcist hvor meget). En større volumen i det digitaliserede materiale vil gøre, at det bedre kan spille sammen med samlinger fra f.eks. Nationalmuseet og Det Kongelige Bibliotek.
- 70.000 værker fra Den Kongelige Kobberstikssamling (SMK). Inklusive forslag 1 og allerede foretagne digitaliseringer vil i alt ca. 110.000 af samlingens 245.000 værker dermed være digitaliseret, og den digitale samling vil for alvor få værdi som brugbar samling i forskningøjemed.

Faktaboks: Den Kongelige Kobberstiksamling

Kobberstiksamlingen er en af de ældste grafiksamlinger i verden. Den kan formentlig føres tilbage til 1500-tallet, hvor de danske konger lagde grunden til den store samling, der i dag omfatter ca. 245.000 værker. Allerede i 1843 blev den tilgængelig for offentligheden, og i 1896 flyttede den ind på Statens Museum for Kunst, da museet stod færdigt.

Kobberstiksamlingen rummer dansk og international tegnekunst og grafik, der spænder fra 1400-tallet over det moderne til samtidskunsten. En betragtelig del af samlingen udgøres af den ældre, europæiske grafik med eksempler på meget sjældne tryk, som kun findes på få museer i verden. Ud over repræsentative samlinger af især danske tegninger findes der en stor bestand af skitsebøger, kunstnerbøger og kunstnermapper samt en samling af fotografier, der er opbygget i de seneste årtier.

Fra gammel tid er samlingen løbende blevet registreret i inventarprotokoller og i de såkaldte arkkataloger. Men sammenlignet med udenlandske grafiksamlinger, der systematisk registrerer deres værker elektronisk, er det endnu kun en lille del af Kobberstiksamlingen, der er blevet gjort tilgængelig for brugerne via museets kunstdatabase. Set i det store perspektiv er samlingen således relativt ukendt for internationale forskere, der ikke har mulighed for gennem det digitale medie at skaffe sig kendskab til samlingens værker og derfor afskæres fra at udnytte det potentiale, som samlingen rummer i forhold til forskningen og den udstillingsvirksomhed, museet meget gerne vil bidrage til. En digitalisering af samlingen vil omvendt give museet lejlighed til at indgå i et udvidet, internationalt samarbejde, der vil få betydning for Kobberstiksamlingens egen forskning og udstillingsaktiviteter.

- 5.000 danske kort (KB). Denne værdifulde samling er i analog form besværligt tilgængelig, bl.a. fordi sådanne samlinger er blandt de mest udsatte for tyveri, men med lettere adgang vil kortsamlingen utvivlsomt være blandt de mest interessante for brugerne, ligesom den er et godt udgangspunkt for brug af bibliotekets store samling af topografisk materiale i øvrigt.
- Knap 700.000 luftfotos (KB). Sylvest Jensen-samlingen består af luftfotografier fra perioden 1936 til 1990. Luftfotografierne er dels optagelser af fritliggende ejendomme, oftest optaget 2-4 gange i perioden, optagelser af landsbyer samt en lang række optagelser fra større byer såvel oversigtsbilleder som detailbilleder af områder. Materialet giver således en enestående beskrivelse af såvel bygningers som landskabers udvikling i perioden. KB har fuld ophavsret til materialet. 700.000 fotos udgør ca. halvdelen af samlingen.
- 200.000 topografiske fotos (KB). Der er tale om et stort fotografisk materiale, som kan supplere kort og luftfotos. Materialet er generelt også meget efterspurgt i forlagsbranchen. Sammen med kort og luftfotos vil materialet også betyde, at man som turist på forhånd vil kunne gå på opdagelse i det danske landskab.
- 300.000 danske postkort (KB). Samlingen af danske postkort rummer en stor mængde information om dansk geografi og kulturhistorie og har stor anvendelighed i forskningsøjemed, som et spændende redskab i undervisningssituationer eller simpelthen som et materiale, hvor borgeren kan gå på opdagelse i sit lokalområde fra en ny vinkel.

- 600.000 portrætfotos (KB). Der er tale om broderparten af bibliotekets portrætsamling på i alt 900.000 fotos. Fotografierne er forsynet med basale metadata om navn mv. og kan oplagt indgå i forskellige sammenhænge, hvor der etableres brugerskabte metadata på tværs af forskellige materialer.

Faktaboks: Det Kongelige Biblioteks portrætsamling

Samlingen af portrætter findes primært i Hovedsamlingen, hvor alle personer er registreret på navn i portrætregranteren. Ud over navn findes korte oplysninger om fødselsår, evt. erhverv og giftermål. Samlingen omfatter omkring 100.000 personer.

Ud over Hovedsamlingen findes en lang række mindre særsamlinger (Pressehusets samling, Dagens Nyheder, Aftenbladet, Tidens Kvinder). Herudover findes negativarkiver i Albumsamlingen samt i de hundreder af protokoller fra en lang række forskellige fotografer.

Portrætterne kan f. eks. indgå i undervisningsforløb inden for historie, samfundsfag eller kunsthistorie, hvor man kan vise billeder af centrale aktører. De kan indgå i samspil med lydoptagelser af den pågældende fra f.eks. Statsbiblioteket, kombineres med arkivdata fra Statens Arkiver eller med fotografier fra Kulturarvsstyrelsen af bygninger, de måtte have tegnet eller lade opføre, manuskripter fra Det Kongelige Biblioteks håndskriftsafdeling eller kunstværker, de måtte have fremstillet.

Herudover kan portrætterne bruges til vise træk af fotografiets udvikling, fortælle noget om mode og beklædningsudvikling, hårmode og andre historisk oplysninger, der kan aflæses af personfotografier, fotograferingsmetoder og andre sociale fænomener.

- 100.000 ENB-genstande, dvs. museumsgenstande af enestående national betydning (NATMUS). Billedmateriale af museets genstande er meget efterspurgt af både forskere, undervisningssektoren og den almindelige borger med interesse for kulturhistorien. I dag kender alle Solvognen og Gundestrupkedlen, men Nationalmuseet har mange genstande, som på hver deres måde er lige så enestående, men i dag ganske ukendte uden for de snævre faglige kredse. Digitaliseret som 3D-genstande, der kan drejes, så genstandene kan ses fra alle sider, giver digitalisering mulighed for fordybelse og inspiration for både lægdpersoner, forskere og designinteresserede med mulighed for skabelse af ny kreativitet.
- 190.000 glasnegativer (NATMUS). Billederne er dels billeder af museets egne genstande optaget med henblik på udgivelse af museets statelige publikationer, dels egentlige arkivbilleder, som stammer fra Nationalmuseets forskellige arbejdsområder: arkæologiske udgravningsbilleder, ekspeditionsbilleder og andre former for dokumentationsbilleder af kulturhistorisk interesse. Materialet supplerer indholdet i Nationalmuseets beretningsarkiver. Materialet er i sin analoge form stort set utilgængeligt.

Faktaboks: Nationalmuseets Billedsamlinger

I 1870'erne begyndte Nationalmuseets arkæologer at medbringe fotoudstyr, når de skulle på udgravning. Hidtil havde man haft tegnere til at dokumentere udgravningsarbejdet. Sådan foregik det f.eks. i Nydam Mose i Sønderjylland ved udgravningen af det berømte våbenmosefund.

Mange af de senere så berømte udgravninger blev dokumenteret med den moderne fototeknik. F.eks. Ertebølle-stenalderens køkkenmøddinger og bronzealderens gravhøje. I 1891 udgravede Vilhelm Boye lokaliteten Guldhøj, og både fundet og fundomstændighederne af de berømte egekister blev nøje fotograferet. Senere i 1922 blev den enestående Hjortspringsbåd fra førromersk jernalder udgravet og dokumenteret på glasplader. Dette enestående dokumentationsmateriale viser arkæologernes arbejde i marken, og det er vigtigt, at disse billeder bliver sikret for eftertiden og gjort tilgængelige, så de kan bidrage til vores forståelse af det særlige ved Danmarks Oldtid og formidlingen af den.

Kulturhistorikeren Hugo Matthiessen arbejdede i en længere årrække på Nationalmuseet. Med støtte fra en række byråd fotograferede han gamle huse og gadebilleder fra danske købstæder. Købstæderne var på den tid under kraftig forandring. Disse fotos er i dag meget benyttede som dokumentation af købstædernes udvikling og ændrede vilkår og har stor lokalhistorisk interesse.

Derudover har Nationalmuseet topografisk og emneordnede fotosamlinger, som dokumenterer almindelige danskeres dagligdag gennem tiden: arbejdet på land og i by, de forskellige håndværk, mænd og kvinder, unge og gamle. Nationalmuseet har desuden en stor samling billeder fra besættelsestiden (ca. 60.000).

NM har også fotodokumentation fra mere fremmede himmelstrøg. I Etnografisk Samling findes således billeder fra de store ekspeditioner: Thuleekspeditionerne 1912-33, Henning Haslund Christensens Centralasiatiske Ekspeditioner 1936-39 og 1947; ekspeditioner hvor man ud over billedmaterialet også tilvejebragte nogle af Nationalmuseets fineste genstandssamlinger.

I Antiksamlingens billedsamling findes dokumentation fra museets deltagelse i Carlsbergfondets Rhodosekspedition 1902-14. Ud over at dokumentere kendte og ukendte danskeres tankevækkende tilstedeværelse overalt i Verden, inviterer disse billeder til formidling på mange niveauer, og de supplerer andre kulturinstitutioners kilder, så det samlede billede af tider, personer og begivenheder får flere facetter, som giver mulighed for nye tolkninger for flere brugere af institutionernes kulturudbud.

I Nationalmuseets fotosamlinger er der helt unikke billeder, som viser lokaliteter og forhold, som i dag er helt forandrede eller forsvundet som f.eks. museets billeder fra den tsunamiramte atol Nias i Indonesiens Acehprovins. Billeder som i dag har stor betydning for den lokale befolknings identitet og historieopfattelse, idet øens eget arkivmateriale er gået tabt.

Levende billeder

- 3.000 timer (6.000 titler) kort- og dokumentarfilm på videomasters (DFI). Der er tale om en samling af unik national værdi, som rummer både undervisningsfilm og vigtige kunstneriske værker. Digitalisering i bevaringsformat er påkrævet, hvis ikke denne vigtige kulturarv skal gå tabt på sigt. Samtidig har Filminstituttet også i mange tilfælde de ikke-kommercielle rettigheder, hvorfor filmsamlingen, der er en unik kilde til nyere dansk historie, vil kunne gøres tilgængelig for forskere, uddannelsessøgende og via bibliotekerne. Der vil evt. også kunne etableres tilgængelighed i fuld kvalitet for filmbranchen, hvilket kraftigt vil øge samlingens værdi som råstof for nyproduktioner.
- Optagelser af sendeflader på dansk tv i perioden 1987-2005 (SB). I egenskab af Statens Mediesamling har Statsbiblioteket indsamlet komplette danske sendeflader for hele perioden fra 1987, i alt svarende til 256.000 timer. Samlingen omfatter de landsdækkende og regionale kanaler under DR og TV 2, samt i udvalg kommercielle tv-stationer som TV3, TV-sport og diverse lokal-tv. I Statsbibliotekets nuværende aftale med Copydan indgår allerede en hensigtserklæring om indgåelse af en aftale om digital tilgængeliggørelse for målgrupper som videregående uddannelser og journalister, for hvem den potentielle værdi af samlingen er meget betydelig. Det skal tilføjes, at digitaliseringen også er bevaringsmæssigt påtrængende, da samlingen ligger på VHS-bånd.

4.2 Udvalgets bemærkninger

Inden for rammerne af forslag 2 kan man opnå en mere omfattende digitalisering end i forslag 1. Forslag 2 indebærer samtidig merudgifter på i alt 300 mio. kr. i forhold til forslag 1 i periode 2010-2019.

Det kvalitative spring ligger i, at man inden for rammerne af forslag 2 i højere grad digitaliserer hele samlinger eller væsentlige dele af dem. Hermed forøges materialets potentiale til tværgående brug, eftersom der bliver flere og flere emner og problemstillinger, som man kan belyse med materiale fra flere samlinger, samtidig med at muligheden for at grave sig dybt ned i helt specifikke problemstillinger inden for enkelte samlinger også bliver større.

Herved opnår de allerede i sig selv værdifulde samlinger forøget værdi for alle grupper af brugere, måske mest for to kategorier – nemlig de *meget* interesserede (professionelle) brugere, og de i udgangspunktet måske lidt mindre interesserede.

For de professionelle brugere, ofte forskere af forskellig art, betyder kompletheden af samlingerne en stor lettelse i materialesøgningen, fordi de i større omfang undgår at skulle finde analogt materiale. Det er dog først med et digitaliseringsomfang som i forslag 3, at fordelene for forskergruppen høstes fuldt.

For de i udgangspunktet lidt mindre interesserede – det kan være skoleelever i en undervisningssituation eller blot den lidt løst søgende almene bruger uden nogen specifik interesse – betyder den store bredde sandsynligvis også meget. Den giver en øget frihedsgrad og betyder, at man sjældnere søger forgæves i materialet: giver forøgede muligheder for spontant at søge på tværs.

Det indfanger den tilfældige forbipasserende på nettet og fanger interessen for kulturarven, ligesom den gør brugen af materialet til en mere engagerende oplevelse for eleven i en undervisningssituation. Ikke mindst i undervisningssituationen tillader den større volumen også en mere selvstændig og kreativ brug af materialet, som kan øge indlæringspotentialet.

Også for elever i de yngre klasser, for hvem det brede søgeudvalg måske er svært at forholde sig til, kan det have betydning, at man med den bredere digitalisering får adgang til materiale fra stort set ethvert sted i landet, således at man præsenterer denne aldersgruppe for kilder, som direkte relaterer sig til deres hverdag og lokalområde.

Formidlingseksempel: Tre tværgående temaer

Med udgangspunkt i det digitaliserede materiale i forslag 2 kunne man skabe en formidlingsportal rettet mod folkeskoler og ungdomsuddannelser, der kan fungere som indgang til materialet i perspektiv af tre temaer: *Fra Landbrugssamfund til Vidensamfund, Historie og Erindring* samt *Det 20. Århundredes Kultur*.

Under temaet *Fra Landbrugssamfund til Vidensamfund* kan samles kort, luftfotos, andre topografiske fotos og til dels postkort fra KB og tinglysningsdokumenter mv. fra Statens Arkiver til belysning af kulturlandskabets udvikling gennem de sidste 150 år, sammen med avissamlingen fra Statsbiblioteket, som belyser den sideløbende samfundsudvikling. Man vil kunne fremsøge oplysninger om stort set et hvilket som helst sted og en hvilken som helst begivenhed i den historiske overgang til det moderne samfund, således at brugeren oplever virkelig frihed i materialet og ikke bare oplever at arbejde med en lærebog på internettet.

Under temaet *Historie og Erindring* kan sammenkobles en mængde materiale om samfundsmæssig og personlig erindring: arkiver om fortidsminder og arkæologi, erindringer, portrætfoto, tv-udsendelser mv. Materialet vil være en udtømmelig kilde til belysning af spørgsmål om historie og identitet: uhyre relevante spørgsmål i undervisning på alle niveauer.

Under temaet *Det 20. Århundredes Kultur* kan samles et unikt dokumentationsmateriale om kulturen i de sidste 100 år, i form af film, tv-sendeflader, musik, aviser, og samlinger af postkort og småtryk – et materiale som vil kunne bruges i forbindelse med stort set ethvert emne i nyere dansk kulturhistorie.

En fordel ved forslag 2 er, at rigdommen og frihedsgraden i materialet kan trække nye brugergrupper til på en måde, som det nøje udvalgte materialeudvalg under forslag 1 ikke kan – fordi man her risikerer at blive skuffet eller frustreret, når det, man spontant søger, ikke er der, og den i udgangspunktet mindre motiverede bruger kan dermed hurtigt give op.

På den måde kan forslag 2 i højere grad bidrage til at gøre kulturarven til alle mands eje.

Ulempen ved forslag 2 i forhold til forslag 1 er derimod, at man med den større volumen og mere konsekvente samlingsdigitalisering i højere grad vil digitalisere materiale, der kun vil blive meget lidt brugt. Til gengæld falder marginalomkostningerne pr. inkluderet enhed formodentlig relativt meget.

5 Forslag 3

5.1 Udvalgets forslag

Inden for de økonomiske rammer af en øremærket bevilling på i alt 500 mio. kr. i 2010-2019 foreslår udvalget digitalisering af nedenstående materialer i perioden 2010-19.

Hertil kommer udgifter til tilgængeliggørelse, der løbende vil stige op til 10-15 mio. kr. årligt ved udløbet af perioden, jf. afsnit 2.2.

Samlet indebærer forslag 3 således merudgifter på op til i alt 570 mio. kr. i perioden 2010-19⁸.

Der er tale om følgende materiale ud over det, der er nævnt i forslag 1 og 2:

Trykte tekster, arkivalier, håndskrifter mv.

- 100.000 sider danske atlas (KB). Der er tale om en mindre del af bibliotekets samling af danske atlas, som vil tjene til at supplere den kortsamling, som digitaliseres under forslag 2. Hermed vil kortsamlingen blive nærmere komplet, hvilket vil have stor værdi for forskere mv., da det vil reducere behovet for at benytte de analoge faciliteter.
- 250.000 sider småtryk af særlig topografisk interesse (KB). Der er tale om samlinger af lokale politiske blade, lokale idrætsblade, kirkeblade, skoleberetninger mv., som har stor interesse i samspil med de øvrige topografiske samlinger og kan levere interessant lokal- og personalhistorisk materiale både til professionelle historikere, slægtshistorikere og almindeligt interesserede borgere.
- 1,4 mio. sider topografisk litteratur (KB). Hermed vil en meget omfattende samling topografisk litteratur (2,8 mio. sider i alt) være digitaliseret, og et meget dybdegående materiale vil være tilgængeligt til at levere kontekst og perspektiv til det topografiske billedmateriale.
- 1 mio. sider dansk-norsk personalhistorie (KB). Der er tale om en samling, der har oplagt interesse også for ikke-professionelle, og som er skræddersyet til at spille sammen med samlingen af portrætfoto.
- Arkivalier af særlig slægtshistorisk interesse, ca. 1 mio. optagelser (SA):
 - Lægdsruller og søruller
 - Testamentprotokoller
 - Fæsteprotokoller
 - Borgerskabsprotokoller

Disse arkivalier hører til blandt dem, der efterspørges af slægtsforskere, og knytter sig således til kirkebøger, folketællinger og skiftearkivalier.

⁸ Jf. note 7, side 22, regnes der med 7 mio. kr. pr. år i gennemsnit.

- Udvalgte søgemidler (SA). Som 2: yderligere 2 mio sider.
- Forsikrings- og taksationsprotokoller vedr. brandforsikring (SA) (3,5 mio sider). Materialet indeholder oplysninger om forsikrede bygninger helt tilbage fra 1700-tallet, med oplysninger som ejernavne, vurderinger og forsikringsmæssige beskrivelser af ejendommens bygninger. Materialet har således stor interesse i bygningshistorisk, slægtshistorisk og lokalhistorisk perspektiv, men er i analog form ret vanskeligt anvendeligt.
- 100.000 sider fra Nordisk Films kontorarkiv fra 1906-70 (DFI). Der er tale om et materiale, der vil have stor værdi for forskere og andre med interesse for at studere dansk films historie, og som vil supplere øvrigt digitaliseret materiale på filmområdet.
- 500.000 sider tekst fra Antikvarisk-Topografisk Arkivs korrespondancearkiv, samt ca. 325.000 sider fra arkivets øvrige sidearkiver (NATMUS). Dette materiale fortsætter tidsmæssigt Bygningsfredningsarkivet og Arkivet over Fredede Fortidsminder i KUAS.
- 200.000 sider tekst fra mere end 50.000 almindelige danskere, som skrev dagbog den 2. september 1992 efter opfordring fra Nationalmuseet og DR (NATMUS). Dagbøgerne er et enestående øjebliksbillede med fortællinger fra barndom og voksenliv, som det blev erindret en bestemt dag i 1992. Denne enestående dokumentation af danskeres holdninger til tilværelsen vil set i sammenhæng med avisartikler, tv-produktion mv. fra samme dag give mulighed for et aldrig tidligere set punktnedslag i Danmarks historie. Materialet er en del af Nationalmuseets Etnologiske Undersøgelser.
- 125.000 sider tekst fra to af Nationalmuseets store bogværker: Værket Danmarks Kirker supplerer Antikvarisk-topografisk Arkiv med den nyeste forskning, og Nationalmuseets Arbejdsmark er trods sin karakter af årbog stedet, hvor mange af museets genstande er publiceret. Digitale udgaver med de moderne søgefunktioner vil øge tilgængeligheden af indholdet betragteligt

Billeder

- 700.000 luftfotos (KB). Jf. forslag 2 vil den fulde Sylvestsamling så være digitaliseret, og den professionelle anvendelse af samlingen for forskere og erhvervsliv vil hermed blive betragteligt lettet, ligesom det vil have stor betydning for borgeres og uddannelsessøgendes oplevelse af materialet, at man meget sjældent vil komme til at gå forgæves i den digitale samling.
- 100.000 øvrige topografiske fotografier (KB). Hermed vil i alt 300.000 øvrige topografiske fotos være digitaliseret, og en meget omfattende samling vil hermed være tilgængelig digitalt.
- 300.000 portrætfotografier (KB). Hermed vil hele bibliotekets samling på 900.000 portrætfotografier være digitaliseret. Den komplette digitale adgang til samlingen vil selvsagt betyde en lettelse for professionelle brugere, men vil også forbedre oplevelsen for almindelige interesserede borgere.

- De formidlingsegneede dele af Bygningsfredningsarkivet og Arkivet over Fredede Fortidsminder (KUAS), jf. forslag 1 og 2. Inden for rammerne af forslag 3 vil det være muligt at digitalisere alle væsentlige dele af materialet, ligesom der evt. også vil kunne inddrages relevant materiale fra statsanerkendte museer. Herved vil et komplet billede af Danmarks fredede bygninger og fortidsminder være digitalt tilgængeligt.
- 95.000 værker fra Den Kongelige Kobberstikssamling (SMK). Inklusive forslag 1 og 2, samt allerede foretagne digitaliseringer vil i alt ca. 205.000 af samlingens 245.000 værker dermed være digitaliseret. En så nær komplet digitalisering vil betyde, at samlingen vil være næsten fuldt tilgængelig online, hvilket kraftigt vil øge samlingens værdi for udenlandske forskere og dermed forstærke museets placering i internationale forsknings- og museumsnetværk.
- 40.000 portrætter (DFI). Der er tale om fotografiske portrætter af primært danske instruktører og skuespillere, der kan komplementere det øvrige digitaliserede materiale fra instituttets samlinger.
- 60.000 stillbilleder fra danske tonefilm (DFI). Der er tale om en særsamling med stor interesse for forskere og andre filminteresserede, men også i forhold til medier og forlag.
- 150.000 billeder af genstande fra Danmarks Nyere Tid, som vil komplettere og supplere museets database over museumsgenstande (NATMUS).
- 100.000 nyoptagelser af formidlingsegneede museumsgenstande (NATMUS).
- 60.000 billeder fra besættelsestiden (NATMUS).
- 500.000 billeder fra museets mange billedsamlinger (NATMUS).

Musik og lyd

- Musik indspillet i Danmark 1903-1960 (SB). Statsbiblioteket har så godt som alle danske musikudgivelser fra de sidste 100 år, og 78'-pladesamlingen, som dækker perioden 1903-1960, udgør derfor en komplet indgang til den danske musikskat i perioden. Adgangen til denne musik er i dag stærkt begrænset, da de gamle plader ikke kan udlånes af hensyn til samlingens bevaring, og fordi brugerne ikke har afspilningsudstyret. Størstedelen af den musik, der indgår i samlingen, er i dag kun tilgængeligt ved personligt fremmøde på Statsbiblioteket i Århus, men vil i digital form kunne give et vigtigt og spændende indblik i dansk kulturhistorie.

Levende billeder

- Danske reklamefilm 1900-1955 (SB). Der er tale om en samling på skønsmæssigt 10.000 reklamefilm, der giver et unikt indblik i dansk historie i perioden, men som i analog form er stort set fuldstændig utilgængelig, og som samtidig har et stærkt behov for en bevaringsindsats. Samlingen vil kunne have stor værdi for forskere og i undervisningssituationer, stor interesse for borgerne generelt og vil også kunne være værdifuld som inspirationsværktøj for reklamefilmbranchen i dag. Samlingen vil supplere de

nyere danske reklamefilm, som Statsbiblioteket i dag stiller online til rådighed for forskere og studerende ved de videregående uddannelser i hele landet i henhold til en aftale med Copydan.

- 2.000 kort- og dokumentarfilm på filmbase fra 1939 til i dag (DFI). Jf. forslag 2 vil hele samlingen på ca. 8.000 kort- og dokumentarfilm (i alt ca. 4.000 timer film) hermed være digitaliseret, og en unik kilde til dansk samfunds- og kulturhistorie vil hermed være fuldt digitalt tilgængelig. Samlingen i sin helhed vil formodentlig kunne formidles til videregående uddannelser og folkeskoler.

Faktaboks: Kort- og dokumentarfilmsamlingen

Det Danske Filminstituts kort- og dokumentarfilmsamling består primært af film produceret for eller gennem støtte fra statslige filminstitutioner som Ministeriernes Filmudvalg, Dansk Kulturfilm og Statens Filmcentral. Samlingen indeholder oplysende, opdragende, og i nogle tilfælde eksperimenterende film, der beskriver alt fra hjemlig hygiejne til verdensrummets sammensætning.

Det vil være svært at finde emner, der ikke er beskrevet i disse over 5.000 film, der er produceret fra 1930'erne til i dag. Det er en ressource, der ikke alene indeholder information om utallige emner, men også er en historisk kilde til, hvordan staten og samfundet historisk har belyst disse emner. Der er tale om film produceret til den danske befolknings glæde og opdragelse, der i dag kan belyse ikke alene de historiske forhold, men også hvordan det offentlige Danmark har henvendt sig til befolkningen, fra rene umedierede optagelser af kunstnerisk optræden som "Mary Wigman danser (1929)" til oplysningsfilm som "Enlige mødre (1954)" eller portrætfilm om danske kunstnere som "Inger Christensen – Cikaderne findes (1998)".

Ud over selve filmene besidder DFI supplerende samlinger i form af fotos og dokumentsamlinger, der kan berige samlingens forskningsmæssige og samfundshistoriske værdi ved sammenstilling i digital form.

5.2 Udvalgets bemærkninger

Forslag 3 byder på en digitalisering, der giver meget stor bredde og dybde i det digitaliserede materiale. Fokus er kort sagt *komplethed* i de digitaliserede samlinger – selvom der naturligvis fortsat kun er tale om en delmængde af den samlede, digitaliseringsværdige kulturarv og i den forstand selvfølgelig ikke er tale om komplethed i total forstand, kun i forhold til de enkelte samlinger.

Men også dette byder naturligvis på store, ekstra muligheder for alle brugere, men især giver det en ekstra dimension for de brugere, der arbejder meget seriøst med det digitaliserede materiale.

Dette har betydning for forskere inden for især de samfundsvidenskabelige, humanistiske og æstetiske fag. Den digitale tilgængelighed af så mange forskellige ressourcer betyder en stor lettelse i adgangen til kildematerialet, fordi man i meget vid udstrækning ikke længere behøver tilgå de besværlige analoge materialer.

I forhold til den stigende internationalisering af forskningsverdenen er det også af stor betydning for danske forskere, at internationale kolleger kan opnå fjernadgang til dansk materiale og således lettere kan engageres i samarbejde med udgangspunkt i danske materialesamlinger af høj kvalitet. Sådant samarbejde kan være afgørende for at kunne udnytte den eksplosive vækst i tværdisciplinært samarbejde, som man oplever i disse år, og som utvivlsomt vil blive styrket fremover.

Man skal heller ikke være blind for, at muligheden for at deltage i internationale netværk er væsentlig for danskere forskeres muligheder for at få del i internationale finansieringskilder som f.eks. EU-midler. Her må man også have i baghovedet, at lande som eksempelvis Frankrig, Nederlandene og Norge også satser kraftigt på digitalisering i disse år, og at det dermed vurderes at skulle sikres, at danske forskere og forskningsinstitutioner har adgang til materiale svarende til deres udenlandske kollegaer.

Det er imidlertid ikke kun i forskningssammenhænge, at der er en stor værdi ved at gøre meget komplette samlinger tilgængelige digitalt. Der er også en værdi for uddannelsessektoren heri.

Uddannelsessektoren står, på alle niveauer fra folkeskole til universitet, over for store udfordringer i disse år. Fremtidens samfund og arbejdsmarked kræver, at uddannelsessystemet ikke blot bibringer de uddannelsessøgende viden, men også kyndighed i at erhverve og udnytte viden.

Den hidtidige erfaring viser imidlertid, hvor let arbejdet med disse færdigheder i realiteten bliver en useriøs substitut for reel læring, hvis man ikke arbejder med et ordentligt kildemateriale og fyldigt kildemateriale, med dybde og kvalitet, og hvor bredden er så stor, at man har reel frihed til at definere sit emne.

Her tilbyder den mest ambitiøse plan for digitalisering af kulturarven et kildemateriale, der kan bruges til behandling af stort set ethvert historisk, kulturhistorisk og geografisk emne i dansk kontekst, og som kan engagere elever og studerende gennem muligheden for konstant at finde nye materialer og vinkler.

Endelig skal man også erindre den store værdi, som et komplet kildemateriale har for kulturerhvervene, f.eks. medievirksomheder som tv-produktionsselskaber og forlag. Det største erhvervsmæssige potentiale ligger dog utvivlsomt i virksomheder, der udvikler helt nye netbaserede forretningskoncepter, byggende på digitaliseret kulturarvsmateriale.

Ulempen ved forslag 3 i forhold til forslag 1 er, at man med den større volumen og mere konsekvente samlingsdigitalisering i endnu højere grad end forslag 2 vil digitalisere materiale, der kun vil blive meget lidt brugt.

Samlet må udvalget således bemærke om forslag 3, at det sikrer digitalisering af en meget stor volumen materiale, som vil skabe kritisk masse i forhold til alle brugergrupper. Til gengæld vil der naturligvis være tale om en meget vidt favnende digitalisering, hvor man ikke kun fokuserer på det udsøgte, men inden for nogle kategorier (ikke mindst fotos) i høj grad får digitaliseret meget bredt.

Bilag 1: Skematisk oversigt over forslag

	Forslag 1	Forslag 2	Forslag 3
Materiale	<ul style="list-style-type: none"> • 0,6 mio. sider tekst mv. • 0,2 mio. stk. fotos mv. • 6 mio. arkiv-optagelser • 31.300 timer lyd • 96.200 timer levende billeder 	<ul style="list-style-type: none"> • 14,1 mio. sider tekst mv. • 3,0 mio. stk. fotos mv. • 16 mio. arkivoptagelser • 31.300 timer lyd • 355.000 timer levende billeder 	<ul style="list-style-type: none"> • 17,8 mio. sider tekst mv. • 5,1 mio. stk. fotos mv. • 22,5 mio. arkiv-optagelser • 32.500 timer lyd • 357.000 timer levende billeder
Tilgængeliggørelse	<ul style="list-style-type: none"> • Obligatorisk fælles standard for metadata 	<ul style="list-style-type: none"> • Obligatorisk fælles standard for metadata • Mulighed for fælles bevaringsinfrastruktur • Mulighed for fælles streamingserver 	<ul style="list-style-type: none"> • Obligatorisk fælles standard for metadata • Mulighed for fælles bevaringsinfrastruktur • Mulighed for fælles streamingserver
Formidling	<ul style="list-style-type: none"> • Fælles eksponering til søge-maskiner 	<ul style="list-style-type: none"> • Fælles eksponering til søgemaskiner 	<ul style="list-style-type: none"> • Fælles eksponering til søgemaskiner • Engelsksproget indholdsoversigt
Økonomiske konsekvenser	<ul style="list-style-type: none"> • afholdes inden for institutionernes ordinære økonomiske rammer 	<ul style="list-style-type: none"> • Merudgifter på i alt 300 mio. kr. i perioden 2010-19. 	<ul style="list-style-type: none"> • Merudgifter på i alt 570 mio. kr. i perioden 2010-19.

Materialeopgørelsen omfatter de finanslovsfinansierede kulturarvsinstitutioner. For hver model er der tale om en samlet opgørelse, dvs. ikke om, hvad der digitaliseres ekstra i forhold til den foregående model.

Bilag 2: Teknisk redegørelse vedr. digitalisering af kulturarven

Indhold

1	Indledning	40
2	Digitaliseringsprocessen	42
2.1	Indledning	42
2.2	Kompetencecentres rolle og opgaver	42
2.3	Digitalisering af lyd	44
2.3.1	Indledning	44
2.3.2	Digitaliseringsprocessen	44
2.4	Tekst- og billeddigitalisering	45
2.4.1	Indledning	45
2.4.2	Digitalisering af tekst og billeder	45
2.5	Digitalisering af levende billeder	47
2.5.1	Indledning	47
2.5.2	Film	47
2.5.3	Video	48
3	Tilgængeliggørelse og formidling	49
3.1	Indledning	49
3.2	Portaler og servicere	49
3.3	Tilgængeliggørelse, metadata og it-systemer	50
3.4	Metadata	51
3.4.1	Behovet for metadata	51
3.4.2	Anvendelse af metadata	51
3.4.3	Standarder for metadata	53
3.5	Udveksling af metadata	54
3.6	Fælles infrastruktur	55
3.6.1	Indledning	55
3.6.2	Lagring og håndtering – bitarkiv	56
3.6.3	Lagring og håndtering – objektniveauet	57
3.6.4	Arkitektur for tilgængeliggørelse	58
3.6.5	Formidling	59
3.6.6	Estimat for minimale udgifter til infrastruktur	61
4	Sammenfatning	62

Forord

Efter offentliggørelsen af Digitaliseringsudvalgets midtvejsrapport i september 2008 blev der nedsat en arbejdsgruppe, som skulle levere konkrete forslag til, hvilke dele af kulturarven der kunne digitaliseres som led i en mere omfattende strategisk indsats for digitalisering af kulturarven. Sideløbende med dette arbejde har en undergruppe drøftet de tekniske afspejter af denne arbejdsgruppes forslag⁹. Formålet med arbejdet i denne tekniske undergruppe var at nuancere midtvejsrapportens beskrivelse af digitaliseringsprocessen og infrastrukturen i lyset af de mere detaljerede forslag til, hvilke specifikke samlinger der kunne digitaliseres.

Den tekniske undergruppe har således sammenholdt anbefalinger og beskrivelser fra midtvejsrapporten med arbejdsgruppens konkrete forslag til digitalisering. Resultatet er en mere nuanceret og detaljeret beskrivelse af digitaliseringsprocessen og et mere specifikt økonomisk estimat for omkostningerne ved etablering af den infrastruktur, der skal understøtte formidling af den digitaliserede kulturarv.

1 Indledning

Digitaliseringen af kulturarv er primært interessant på grund af de nye muligheder, som den giver danskerne for at opleve kulturarven. Befolkningen får adgang til en kulturarv, som de ellers kun ville møde, hvis de fysisk opsøgte arkiverne og samlingerne, og den digitale form giver helt nye muligheder for at præsentere og formidle kulturarven. Samtidig kan digitalisering bevare materiale, der ellers ville være gået til grunde og bevare adgangen til audiovisuelt materiale, der bliver utilgængeligt i sin originale form, fordi afspilningsapparatet forsvinder.

Selve digitaliseringen er således et middel til at gøre kulturarven tilgængelig på lige fod med de øvrige digitale tilbud, der bliver en stadig vigtigere del af hverdagen, ligesom digitaliseringen også tjener et bevaringsmæssigt formål. Selv om digitaliseringsprocessen – i et tilgængelighedsperspektiv – kan synes at spille en underordnet rolle, er en fornuftig tilrettelæggelse og gennemførelse af processen en forudsætning for, at brugerens udbytte af den digitale kulturarv bliver optimalt. Ud over en effektiv digitaliseringsproces er det afgørende, at det digitaliserede materiale efterfølgende kan bevares og tilgængeliggøres gennem en effektiv infrastruktur, der kan understøtte de mange formidlingsinitiativer, der gør brug af materialerne.

Rapporten er opdelt i to overordnede afsnit: en beskrivelse af digitaliseringsprocessen samt en beskrivelse af den efterfølgende tilgængeliggørelse og formidling. Afsnittet om digitaliseringsprocessen beskriver problemstillingerne ved digitalisering af forskellige materialekategorier samt forslag til, hvordan de større institutioners kompetencer og erfaringer kan udnyttes i relation til specifikke materialekategorier.

⁹ Medlemmerne af den tekniske undergruppe var: it-chef Arne Sørensen (Statsbiblioteket), vicedirektør Birte Christensen-Dalsgaard (Det Kongelige Bibliotek), projektchef Tobias Golodnoff (DR Medier, Kulturarvsprojektet), museumsinspektør Thomas C. Christensen (Museum & Cinematek, Det Danske Filminstitut), fuldmægtig Jakob Broberg Lind (Kulturministeriet) og kontorchef Jakob Heide Petersen (Styrelsen for Bibliotek og Medier). Rapporten er udarbejdet af gruppens medlemmer med bidrag fra institutionernes medarbejdere.

Afsnittet om tilgængeliggørelse indledes med en beskrivelse af de generelle problemstillinger vedrørende tilgængeliggørelse på internettet. Derudover redegøres for, hvordan materialet tilgængeliggøres gennem tilføjelse af metadata, hvilke krav der kan stille til it-systemer, og hvordan en infrastruktur for bevaring og tilgængeliggørelse af den digitaliserede kulturarv kan etableres.

2 Digitaliseringsprocessen

2.1 Indledning

En optimal digitaliseringsproces forudsætter et indgående kendskab til den relevante materialetype og den teknologi, der med fordel kan anvendes ved digitaliseringen. Det er afgørende, at processen resulterer i filer i den rigtige kvalitet og det rigtige format, så resultatet både kan anvendes umiddelbart og er fremtidssikret.

Kulturministeriets institutioner benytter ofte udbud i forbindelse med større digitaliseringsopgaver, hvilket medfører, at en stor del af det praktiske arbejde udføres af eksterne aktører, ganske ofte placeret i udlandet. Større digitaliseringsprojekter forudsætter således i den forbindelse betydelige kompetencer, både hvad angår tekniske aspekter, arbejdsprocesser, udbudsforretninger samt opfølgning og kvalitetssikring.

Som beskrevet i det følgende har de større institutioner på Kulturministeriets område allerede disse kompetencer inden for de centrale materialekategorier, og institutionerne vil således kunne fungere som kompetencecentre, såfremt der etableres en større tværgående digitaliseringsindsats. Udvalget har dog ikke taget stilling til, hvorvidt der skal udpeges egentlige kompetencecentre; en sådan beslutning vil bedst kunne tages i forbindelse med den konkrete implementering.

Et væsentligt spørgsmål i forhold til alle digitaliseringsprocesser er komprimering af data. Som udgangspunkt er ukomprimerede formater (eventuelt pakkede ukomprimerede formater) at foretrække, men på grund af de større datamængder og dertil knyttede udgifter ved ukomprimerede formater er der tale om en konkret afvejning fra digitalisering til digitalisering (ved digitalisering med bevaringsformål må ukomprimerede formater være et krav). Men afgørende er det under alle omstændigheder, at digitaliseringen finder sted via åbne, internationalt anerkendte standarder, således at materialet kan anvendes uafhængigt af platform.

2.2 Kompetencecentres rolle og opgaver

Som det fremgår nedenfor, har de institutioner, der råder over store samlinger inden for materialekategorierne lyd, tekst og billeder samt levende billeder, også omfattende erfaring med digitalisering på området. Disse institutioner kan således i princippet fungere som kompetencecentre inden for en af de tre materialekategorier. Kompetencecentrene vil fungere inden for de eksisterende økonomiske rammer. Denne opgave indgår ikke på nuværende tidspunkt i institutionernes opgavebeskrivelse, men man anbefalede en lignende ansvarsfordeling for bevaring i forbindelse med udredning af bevaring om kulturarven¹⁰.

¹⁰ Udredning om bevaring af Kulturarven, Kulturministeriet 2003, side 207, se: http://www.kb.dk/export/sites/kb_dk/da/kb/nb/bev/bevaring_af_kulturarven.pdf

Kompetencecentret inden for en given materialekategori skal kunne rådgive andre offentlige institutioner om bevaring af materialet og indgå i partnerskaber samt evt. insource digitaliseringsopgaver.

Kompetencecentret skal have ekspertviden om alle forhold omkring bevaring af materialet:

- Viden om de forskellige typer analoge medier inden for kategorierne:
 - opbevaring af medierne (magasinforhold, klima mm.)
 - tilstandsvurdering og monitorering af samlingernes tilstand
 - håndtering af medierne ved brug
 - restaurering af skadede medier (f.eks. bånd under nedbrydning)

- Viden om afspilningsudstyr til alle medietyper for lyd og levende billeder samt bevaring af dette udstyr:
 - opbyggelse en samling af afspilningsudstyr, herunder reservedele for udstyr som er udgået af handel, der dækker i princippet alle medietyper
 - ekspertise i brug og vedligehold af afspilningsudstyr

- Viden om digitalisering af medierne:
 - optimal overføring af materialet fra analogt medie til digital fil, herunder inspektion, klargøring/rensning af mediet, valg af udstyr (hvor det er relevant), kalibrering mv.
 - formater til såvel langtidsbevaring som formidling
 - optimale arbejdsgange
 - dokumentation af digitaliseringsprocessen
 - metadata til såvel langtidsbevaring som formidling
 - for lyd: efterbehandling ('rensning') med henblik på formidling
 - effektive digitaliseringsprocesser i form af f.eks. udstyr til masseproduktion og højhastighedsafspilning.

Kompetencecenteret kan være den institution, hvor man på nationalt plan har professionel viden om bevaring af materialet som kulturarv. Opgaven er helt overordnet at sikre, at museale og bevaringsmæssige opgaver på medieområdet varetages i overensstemmelse med åbne, internationalt anerkendte arkivstandarder.

Ovenstående opgaver bør i videst muligt omfang varetages iht. åbne, internationalt anerkendte standarder og løbende udviklet i takt med nye tekniske muligheder. Dette kræver, at standardiseringsarbejdet følges, samt at viden om tekniske forhold løbende ajourføres, bl.a. ved deltagelse i relevante internationale organer.

Langtidsbevaring af en given materialetype bør være en opgave, der løses i fællesskab med andre kompetencecentre på nationalt og internationalt plan, idet digital langtidsbevaring generelt rummer problemstillinger, der er fælles for stort set alle medietyper. Danmark deltager i øjeblikket i to europæiske projekter omkring digital langtidsbevaring – repræsenteret ved Det Kongelige Bibliotek og Statsbiblioteket.

2.3 Digitalisering af lyd

2.3.1 Indledning

Som det fremgår af digitaliseringsudvalgets midtvejsrapport (side 22), befinder hovedparten af kulturarens lydmateriale sig på Statsbiblioteket og i DR. Statsbiblioteket huser Statens Mediesamling, som er det nationale center for audiovisuelle materialer med ansvar for bevaring af danske lydoptagelser. Lydsamlingen rummer 150.000 danske enheder og dækker indholdsmæssigt alt, hvad der er udgivet i Danmark på grammofonplader og bånd mv., ligesom en lang række unikke lydoptagelser fra samfunds- og kulturlivet i det 20. årh. indgår – enkelte helt tilbage fra lydhistoriens barndom i slutningen af 1900-tallet.

Hertil kommer komplette sendeflader med radiooptagelser fra DR fra de sidste 25 år (landsdækkende og regionalt), samt en omfattende lokalradiosamling. Lydoptagelserne i Statens Mediesamling er lagret på medier af alle typer fra de sidste 100 år – fra fonografvalser, 78'-plader, trådruller, spolebånd, vinylplader, kassettebånd og DAT-bånd til cd'er. Mange af de nævnte medietyper findes i varierende undertyper, som hver især kræver særligt afspilningsudstyr. Der modtages fortsat donationer, idet Statsbiblioteket i kraft af sin medieopgave anses for det naturlige sted at aflevere historiske lydoptagelser til.

Historiske lydmedier findes også i andre kulturinstitutioner, der ikke har lydmedier som kerneområde, men hvor de indgår som et vigtigt supplerende kulturarvselement i tilknytning til de centrale samlinger. Blandt disse institutioner kan nævnes Det Kongelige Bibliotek, Nationalmuseet, Statens Arkiver, Dansk Folkemindesamling og Arbejderbevægelsens Bibliotek og Arkiv (se Kulturarvsrapportens bilag 11). DR har endvidere i sine arkiver en del af de programmer, der gennem DR's historie har været udsendt på de landsdækkende programmer. Der er for de ældre udsendelsers vedkommende overvejende tale om arkivering af enkeltudsendelser (ikke hele sendeflader).

2.3.2 Digitaliseringsprocessen

Digitalisering af lyd kræver specialviden om samtlige medietyper, der har været på markedet igennem det 20. årh. Lydoptagelserne i Statens Mediesamling og i de andre offentlige institutioner er for en stor dels vedkommende lagret på konsummedier (ikke-professionelle formater), som ofte er slidte, fordi de er modtaget som donation (pligtaflevering af lydoptagelser blev først indført i 1998).

Lydfæstninger er derfor det medie, som findes i flest forskellige formater. Blandt særligt skrøbelige typer, kan nævnes:

- Voksvalser, som i princippet kun kan afspilles optimalt én gang
- Acetatplader, hvor overfladen pludselig kan krakelere
- Spolebånd, som har været udsat for uprofessionel håndtering
- DAT-bånd i proprietære formater.

I digitaliseringsprocessen er det vigtigt at have kompetencer og viden om optimal overførsel af lyden fra disse forgængelige medier til et digitalt format. Det drejer sig med andre ord om at kunne 'trække' så meget af den originale lyd ud af det analoge medie som overhovedet muligt. Dette er specielt

problematisk for de mange lydmedier, som igennem årene har været udsat for kemisk og fysisk nedbrydning under opbevaring i privat eller offentligt regi uden kendskab til korrekte opbevaringsforhold.

Statsbiblioteket har igennem de senere år arbejdet på etableringen af et kompetencecenter for lyd som et resultat af anbefalingen i Kulturarvsrapporten, dels ved rekruttering af sagkyndigt personale, dels ved opbygning af erfaring igennem konkrete digitaliseringsprojekter. Biblioteket har sideløbende i et vist omfang udført rådgivningsopgaver – og i nogle tilfælde foretaget digitalisering af lydoptagelser for andre offentlige institutioner. Biblioteket får stadigt hyppigere henvendelse udefra om mulighed for overspilning/rådgivning i digitalisering af historisk lyd fra f.eks. Naturhistorisk Museum, Kvindemuseet og Niels Bohr-arkivet.

2.4 Tekst- og billeddigitalisering

2.4.1 Indledning

Der er en lang række institutioner, som har kulturarvsmateriale i form af tekst og billeder. Langt størstedelen af de relevante samlinger inden for disse materialekategorier findes dog på Det Kongelige Bibliotek.

Det Kongelige Bibliotek er nationalbibliotek og universitetsbibliotek for Københavns Universitet. Nationalbiblioteksfunktionen blev etableret i 1648, medens universitetsbiblioteksfunktionen har rødder tilbage til 1482. Bibliotekets samlinger af tekst og billeder er meget omfattende. Biblioteket har 6,9 mio. katalogiserede enheder tekst (bøger, tidsskrifter mv.) svarende til 175 hyldekilometer. Bibliotekets samlinger af fotografisk materiale mv. rummer 16,3 mio. enheder, svarende til 12,2 hyldekilometer. Der findes også store samlinger af tekstmateriale i andre kulturinstitutioner og i mindre omfang også fotografisk materiale. Blandt disse institutioner kan især nævnes Det Danske Filminstitut, Kulturarvsstyrelsen, Nationalmuseet, Statens Arkiver, Statens Museum for Kunst og Statsbiblioteket.

2.4.2 Digitalisering af tekst og billeder

Digitalisering af tekst og billeder falder i en række grupper:

Digitalisering af forskellige kategorier af materiale, der kræver særlige kompetencer, eksempelvis:

- Sjældne dokumenter og værker, som ikke kan forlade institutioner
- Skrøbelige dokumenter og værker, som kræver meget stor viden om håndtering
- Bøger, som må eller ikke må skæres op i ryggen
- Skanning baseret på mikrofilm.

Arbejdsgangene vil naturligvis afspejle hensyn til sådanne materialer. Arbejdsgangene må ligeledes afspejle mængden af allerede eksisterende metadata til materialet, samt afspejle beslutninger om, hvorvidt man vil foretage tekstgenkendelse af tekstmateriale og i bekræftende fald, om man tillige vil lave en opmærket version af materialet.

Kravene og arbejdsgangene vil også afhænge af, om man efter digitaliseringen stadig har rådighed over den analoge original. De tilfælde, hvor originalen ikke er brugbar efterfølgende, taler man om såkaldte substitutskanninger. Her skannes et objekt i stedet for at introducere ekstremt kostbare bevaringsforanstaltninger i forhold til originalen.

Det er typisk for tekstdigitalisering, at man ofte arbejder med mange aktører: Materiale skannes et sted, tekstgenkendes et andet, opmærkes et tredje og kvalitetstjekkes et fjerde. Oven i disse kommer arbejdet med at beskrive og indlemme materialet. Der skal træffes en lang række beslutninger i forbindelse med digitaliseringen, herunder valg af opløsning, farvedybde og format (råt, komprimeret eller pakket). Disse valg vil have stor betydning for den samlede pris for den efterfølgende bevaringsaktivitet.

Et tekst- og billedkompetencecenter bør have den relevante viden i forhold til digitalisering og bevaring af den del af kulturarven, som er udkommet på tryk eller forefindes som billeder (herunder også negativer). Kompetencecentret kan bidrage til at sikre, at materiale digitaliseres i henhold til åbne, internationalt anerkendte standarder, og rådgive om, hvordan man bedst og billigst gennemfører en konkret digitaliseringsopgave samt den digitale efterbehandling og kvalitetskontrol.

Inden for tekstbaseret materiale er der allerede mange erfaringer både på Det Kongelige Bibliotek og ved en række andre biblioteker og institutioner i Europa og USA. Disse spænder over modeller med en meget arbejdskraftsintensiv, selektiv digitalisering af enestående værker til mere automatiseret massedigitalisering af hele samlinger. Inden for massedigitalisering er en alliance med Google en mulighed, som flere større forskningsbiblioteker har benyttet sig af. Samarbejdet finder sted under den forudsætning, at materialet efterfølgende kun kan benyttes af den institution, der råder over samlingen. Det ville i det Kongelige Biblioteks tilfælde indebære, at kun lånere ved KB kunne benytte materialet efterfølgende.

Flere firmaer er begyndt at tilbyde alliancer, herunder Internet Archive, som tilbyder at lave den efterfølgende digitale behandling, hvis institutionen selv står for skanningen. Det tyske firma CSS har en arbejdsgang, hvor digitalisering også gennemføres lokalt, men hvor den efterfølgende kvalitetskontrol er outsourcet til Asien.

Det Kongelige Bibliotek har i mere end 10 år arbejdet med digitalisering af egne materialer, i de sidste 4 år har KB digitaliseret for andre institutioner. Biblioteket har derved opbygget viden og erfaring vedrørende skannere, software, applikationsudvikling, arbejdskraft, definition og styring af opgaver samt outsourcing af opgaver. Fokus har været på digitalisering af materialer inden for tekst og billeder. Der er stor spredning og kompleksitet i bibliotekets digitaliseringsprojekter, hvilket har medvirket til opbygning af omfattende ekspertise på området. Der arbejdes med alle kvaliteter af materialer både i form af fysisk materiale, der skal digitaliseres, og i form af det output, som er resultatet af digitaliseringsprocessen.

Det færdige resultat afhænger af formålet. Biblioteket har opgaver, der varierer meget i størrelse fra få hundrede sider til over 100.000 sider, hvilket har medvirket til opbygning af viden omkring workflow og optimering af disse.

Det Kongelige Bibliotek har gennem 5 år benyttet sig af outsourcing på næsten alle områder af digitaliseringsprocessen, undtagen selve skanningen. Derved har biblioteket høstet erfaringer med, hvilke faktorer der er afgørende for at opnå den ønskede effekt af outsourcingen. Det vigtigste er, at udbyder har et indgående kendskab til opgaven, dvs. institutionen har tilstrækkelig viden til selv at

løse opgaven, så den kan udarbejde de rette kravspecifikationer. Hvis ikke denne ekspertise er til stede, risikerer man, at besparelsen ved outsourcing udebliver. Det Kongelige Bibliotek har outsourcet opgaver til Rumænien, Iran og Indien og tidligere også benyttet Rusland. Erfaring med tidligere projekter gør det muligt at placere nye opgaver de steder, der kan tilbyde den rette pris og de relevante kompetencer til netop den opgave.

Uanset institutionens nuværende kompetencer og erfaringer er det afgørende, at man løbende holder sig orienteret ikke bare om de teknologiske muligheder, men også om, hvordan andre institutioner løser opgaven og til hvilken pris. Det Kongelige Bibliotek har et meget veludviklet kontaktnet til såvel de andre nationalbiblioteker som til de andre institutioner, som har store bog-, dokument- eller billedsamlinger. Blandt dem, som biblioteket aktivt udveksler erfaringer med, kan nævnes nationalbibliotekerne i Frankrig, Holland, England og Østrig, som alle er i gang med omfattende digitaliseringsopgaver

2.5 Digitalisering af levende billeder

2.5.1 Indledning

Det Danske Filminstitut (DFI) er den nationale bevaringsinstitution for film. Samlingen omfatter filmmaterialer på filmbase, såsom de brandfarlige 35 mm nitratfilm, til moderne produktioner på digitale masters. Filminstituttet er kompetencecenter for bevaring og restaurering af film. Inden for de sidste 10 år er faciliteterne til opbevaring af film blevet opgraderet til tidssvarende niveau gennem konstruktion af nye arkiver med klimaforhold, der sikrer en levetid på 500 år til analoge filmmaterialer. Da driften af filmlaboratorier kræver en vis volumen, benytter Filminstituttet sig inden for filmkopiering og film til video-transfer af outsourcing til eksterne speciallaboratorier. Kopiering mellem videoformater er en aktivitet, både DFI og DR besidder faciliteter til, men større kopieringsprogrammer udliciteres typisk til eksterne leverandører.

Det er nødvendigt at kategorisere levende billeder i to undergrupper: film og video. Ligeledes er det på grund af de levende billeders fysiske omfang, i endnu højere grad end for andre materialetyper, hensigtsmæssigt at operere med forskellige formater til master-/bevaringsbrug og formidlingsbrug.

2.3.2 Film

Der anslås ikke i dag at være noget digitalt alternativ til bevaring af analoge filmbilleder. Biograffilm postproduceres ofte digitalt, men distributionen forgår endnu typisk på 35 mm analog kopi, ligesom et udskudt analogt filmmateriale fra de digitale filer anses som det eneste langtidsholdbare medie. Filminstituttet restaurerer udvalgte titler digitalt, hvor slutresultatet er et analogt 35 mm negativ samt en HD-version til digital brug. Filmmateriale i god stand kan opbevares i århundreder under forudsætning af optimale fysiske forhold (5 grader Celsius, eller mindre, samt relativ fugtighed på 35 %).

Der er ikke i skrivende stund et standardiseret digitalt filmformat. Det forventes, at 2K (1920x1080) eller 4K (4000x2000) mpeg2000-filer vil blive fremtidens biografformat. Udfordringerne ved fremtidens biograffilmformater understreges ved, at en 4K ukomprimeret digital filmmaster fylder 4TB per time. Filminstituttet digitaliserer typisk film til HDSR, HDCAM og Digital Betacam, der alle har lavere

opløsning end den originale film. I denne rapport vil der for digitalisering af levende billeder blive beskrevet processer, der overfører film eller videoformater til et videomasterformat, samt digitalisering af disse videomastere til filbaserede videosystemer. Da den primære brug forventes at være til webbrug og broadcast, opereres med en forventning om at overføre film til High Definition tape (HD), 1080x1920 pixel, med en efterfølgende digitalisering til filbaserede systemer.

Scanning/digitalisering af film er en håndteringsmæssigt krævende proces. Det originale materiale skal behandles med varsomhed, da det er originalt bevaringsmateriale. Der er tale om skrøbelige originaler, der kræver kompetent fysisk håndtering, ligesom kopieringen kræver en farvekorrigering, der er specialarbejde. Både udstyr og mandskab er kostbart. DR har in house-faciliteter og kompetence til overførsel af film i god stand til video, mens DFI benytter sig af eksterne laboratorier til både film til filmkopiering samt til digitalisering af film til video. Især for ældre materialer er det svært at effektivisere kopieringsprocesserne, der er arbejdsintensive. Da indholdet skal lyssættes i forbindelse med kopiering, påregnes i omegnen af 4 timer per times indhold til denne proces alene. Hertil kommer fysisk klargøring af materialet før kopiering og efterfølgende digitalisering og konvertering til relevante brugsformater.

2.5.3 Video

Analog video er under hastig udfasning efter ca. 40 år på markedet. Da ethvert videoformat rent fysisk har en holdbarhed på 10-20 år, er mange bånd under akut fysisk nedbrydning. Apparaturet til afspilning er desuden sjældent og svært at vedligeholde i optimal stand. Overførsel til digitale videoformater er den eneste bevaringsmulighed. Filminstituttet anser i dag Digital Betacam som det mest holdbare medie til Standard Definition (SD) video. Der er mange faldgruber i digitalisering af analog video. Ikke alene er det en teknisk kompliceret opgave, men typisk er analog video i formatet 4:3, mens moderne digital video er 16:9, hvilket kræver formatmæssige valg, der aldrig er optimale. Et alternativ er digitalisering med bibeholdelse af det originale format (4:3). Der er endnu ikke sket en modning eller de facto-standardisering af digitale videoformater på masterniveau. Eksisterende løsninger er derfor i vid udstrækning enten proprietære eller baseret på individuelle løsninger.

Filminstituttet benytter digital video som bevaringsformat for SD- og HD-video, herunder båndformaterne HDSR og Digital Betacam, der af DFI betragtes som de mindst ringe lagermedier. Digitale videobånd har en forventet levetid på ca. 15 år under gode opbevaringsforhold (8C/35%RF). Som for analog video vil en fremtidig løsning formodentlig blive migrering/konvertering til jpeg2000 eller lignende filformat. Som bevaringsformat vil ukomprimerede billeder blive foretrukket, hvilket kræver ca. 100 mbs, svarende til 60GB per time for SD samt ca. 4 gange så meget for HD.

Det er værd at bemærke, at DFI har en høj grad af fokus på de bevaringsmæssige aspekter, mens Danmarks Radio også lader produktionsmæssige forhold spille ind. DR har eksempelvis valgt et digitalt SD-format, DVCPro50, der ikke indebærer den højest mulige kvalitet. DR har ved valget af et filbaseret system optimeret produktionsgangene og forventer at fremtidig migrering af indhold vil kunne ske automatisk og uafhængigt af individuel håndtering af materiale.

Digitalisering til lettere håndterbare intermediate "masters", såsom AVI 6mbs, hvorfra flash-filer eller lignende kan genereres til webbrug, er ikke som sådan en kompliceret proces, men vil logistiskmæssigt være vigtigt at få beskrevet, så de samarbejdende institutioner kan digitalisere video på en homogen og effektiv måde på tværs af institutionerne.

3 Tilgængeliggørelse og formidling

3.1 Indledning

Internettet er det vigtigste medie til at give brugerne adgang til den digitaliserede kulturarv. Nettes alsidighed og udbredelse giver helt nye muligheder for at forbinde brugere og kulturarv. Internettets omfang og overflod af information giver dog samtidig en række udfordringer, hvis man skal sikre, at den digitaliserede kulturarv er synlig for brugeren, og at brugeren finder frem til det relevante materiale.

Den første forudsætning, for at brugerne kan anvende den digitale kulturarv, er at den er tilgængelig via internettet. Tilgængeliggørelse kan i denne sammenhæng defineres som en eksponering af materialet på en måde, der giver slutbrugeren mulighed for at få adgang til materialet via internettet. Der er tale om en mere passiv præsentation af materialet end decideret formidling, hvor materialet præsenteres i en kontekst, der formodes at være relevant for brugeren eller for et givent emne.

Tilgængeliggørelsen kan bestå i, at brugeren gennem en søgning f.eks. via Google kan fremfinde nogle få oplysninger om materialet med et link til selve det digitaliserede billede, lydclip, video eller lignende. Formidlingen kan eksempelvis være formidling af materialet gennem en portal eller en skræddersyet præsentation til en given målgruppe.

3.2 Portaler og servicier

Sondringen mellem tilgængeliggørelse og formidling er blevet mere relevant, efterhånden som brugere i de senere år i højere grad anvender nogle få søgemaskiner til at finde information på internettet. En stor del af den tidlige webformidling bestod i opbygningen af portaler, der ofte risikerede at blive forbigået af brugerne, hvis ikke brugerne i forvejen havde kendskab til portalerne, eller hvis ikke indholdet fra portalerne dukkede op i søgemaskinerne.

Med søgemaskinerne og især Googles udbredelse blev det efterhånden set som vigtigere, at materialet var tilgængeligt og synligt for søgemaskinerne frem for gennem en portal, som brugerne måske ikke fandt frem til. Den udvikling skyldes i høj grad den øgede informationsmængde på internettet: Der var en overflod af information, som var vanskelig at formidle i et struktureret portalunivers. Portalerne mistede efterhånden deres monopol som indgange til information, og man talte om en bevægelse fra "content is king" til "context is king". Man satsede i højere grad på at gøre materialet tilgængeligt i den kontekst, brugeren befandt sig i, frem for at forsøge at lokke brugeren til et bestemt sted på internettet.

Desuden havde mange portaler et bestemt formidlingsperspektiv, der ofte var meget afsenderorienteret og tog udgangspunkt i et bestemt emne eller en bestemt institution. Portalerne havde i mange tilfælde også en ret lukket struktur både i forhold til andre tilbud på nettet og i forhold til interaktionen med brugerne. Sådanne portaler er blevet beskrevet som "walled gardens" med proprietært indhold – de tilbød brugerne en helt bestemt service inden for meget afgrænsede rammer.

En stor del af den nyere webformidling fremhæver vigtigheden af, at brugeren kan møde materialet i sin daglige adfærd på nettet, og at formidlingen inddrager brugeren i størst muligt omfang. Det betyder, at man fortsat kan arbejde med portaler, der tager afsæt i bestemte brugergrupper eller emner, men at disse portaler ofte er mere åbne i deres struktur. Man stiller således ofte indholdet til rådighed som servicier, som andre websites kan anvende til at supplere egen formidling (f.eks. via RSS). I den forbindelse har institutioner f.eks. på kulturarvsområdet ofte gennem deres kendskab til brugerne og deres specifikke behov mulighed for at sikre, at formidlingen tilrettelægges på en måde, der tilgodeser brugerens behov.

Ofte bestræber man sig på at sikre, at metadata optræder i flest mulige sammenhænge, således at brugerne har flere chancer for at møde materialet gennem sine foretrukne websites eller søgemaskiner. Denne nyere webformidling er i høj grad understøttet af en række nye teknologier som webservicer, RSS-feeds og mere lette og dynamiske webprogrammeringsværktøjer og hele det nyt paradigme omkring web 2.0.

Udviklingen betyder, at digitalt kulturarvsmateriale vil kunne anvendes på helt nye og uventede måder, når det gøres tilgængeligt på internettet. Den betyder dog også, at det er vanskeligt at udfolde en samlet formidlingsvision – f.eks. i form af én bestemt portal eller ét færdigt brugerscenarie. Det er baggrunden for, at skildringen af formidlingen består af eksempler.

Den teknologiske udvikling har dog medført, at det er langt lettere og billigere end tidligere at lave formidling via internettet, og at der er langt flere eksisterende formidlingskanaler, hvor materialet kan finde anvendelse og skabe værdi. Det er baggrunden for, at der i det følgende er forholdsvis stor vægt på tilgængeliggørelse frem for formidling.

3.3 Tilgængeliggørelse, metadata og it-systemer

Den konkrete tilgængeliggørelse af den digitaliserede kulturarv stiller både krav til it-systemerne og den måde, materialet er beskrevet på:

- De systemer, der indeholder det digitaliserede materiale og oplysningerne om materialet (metadata), skal være forbundet med internettet.
- Systemerne skal kunne udveksle information om materialet med andre systemer via åbne standardprotokoller.
- Materialet skal være beskrevet med metadata og/eller kunne indekseres af søgemaskiner.

Den mest simple model for fælles tilgængeliggørelse består i, at de lokale systemer understøtter standardprotokoller for dataudveksling, at der anvendes standardiserede metadata, og at metadata høstes til en fælles database, som brugeren kan søge i. Det er den model, som i midtvejsrapporten blev beskrevet som simpel søgning (side 64).

I midtvejsrapporten blev modellen fravalgt til fordel for en mere omfattende model med integreret søgning. Den infrastruktur, der beskrives her, vil som minimum tilbyde samme funktionalitet som den simple model for fælles tilgængeliggørelse.

Der arbejdes i nærværende rapport med en nuancering af anbefalingerne i midtvejsrapporten. Den væsentligste forskel mellem anbefalingerne i denne rapport og midtvejsrapportens anbefalinger er vægtningen af løsningen med integreret søgning. Der er i denne rapport mindre fokus på integreret søgning, fordi en del af de materialer, som foreslås digitaliseret i hovedrapporten, med fordel kan formidles på andre måder. Integreret søgning er baseret på indeksering af digitalt indhold – på nuværende tidspunkt primært tekst. Billeder og levende billeder vil på nuværende tidspunkt ikke kunne formidles væsentligt bedre med integreret søgning, og da en stor del af det foreslåede indhold er omfattet af disse kategorier, er fokus flyttet til andre formidlingsformer.

3.4 Metadata

3.4.1 Behovet for metadata

Metadata er ”data om data” og altså i denne sammenhæng en beskrivelse af det digitaliserede materiale. En åben og struktureret beskrivelse af viden om alle objekterne i den digitaliserede kulturarv ville give optimale forudsætninger for tilgængeliggørelse. Således ville en beskrivelse af metadata ud fra ”topic maps” give gode forudsætninger for at søge på tværs af samlinger, ligesom det vil udgøre et godt grundlag for blandt andet tematiske formidlingsinitiativer.

Det er desværre særdeles omkostningskrævende at udføre manuelle beskrivelser af de mange digitale objekter. Det ville således formentlig være mere omkostningskrævende manuelt at beskrive de 700.000 luftfotos fra Det Kongelige Biblioteks Sylvest Jensen-samling, end det ville være at digitalisere dem. Omvendt kan det være meget vanskeligt og måske endda umuligt for en bruger at finde frem til et billede, hvis der ikke er tilknyttet beskrivende data. Metadata kan altså være forudsætningen for, at det digitale objekt overhovedet bliver anvendt, men samtidig være omkostningskrævende.

Hvis metadata skal udarbejdes som led i en manuel proces, bliver det meget vigtigt at afgrænse metadata til de mest nødvendige data for ikke at gøre digitaliseringen prohibitivt dyr. Der skal altså i nogle tilfælde foretages en afvejning mellem ønsket om righoldige metadata og udgifterne ved at producere dem.

I mange tilfælde kan man imidlertid tilknytte beskrivende data til de digitaliserede objekter uden at skulle gøre det manuelt. Man kan således hente metadata fra andre registre, f.eks. bibliotekernes kataloger og museers og arkivers registre. Digitaliseringen vil her udnytte de beskrivende data, der allerede findes om de fysiske genstande. Disse metadata er dog ikke oprindeligt skabt med henblik på digitalisering eller tværgående tilgængeliggørelse. Det betyder, at de ikke er ensartede på tværs af samlinger og på tværs af sektorer (f.eks. biblioteker, arkiver og museer). Generelt bør eksisterende metadata (protokoller og kartoteker mv.) digitaliseres så fyldigt som muligt med henblik på at undgå tab heraf.

3.4.2 Anvendelse af metadata

Når anvendelsen af metadata varierer på tværs af samlinger og sektorer, skyldes det naturligvis, at metadata ofte kræver økonomiske ressourcer, og at man derfor ikke anvender flere metadata end formålet tilsiger. Det oprindelige formål med beskrivelsen af kulturarvsobjekter har oftest været

afgørende for valg af metadata. Afhængigt af formålet med at tildele metadata til fysiske objekter i en samling kan der således anvendes mange forskellige typer metadata, f.eks.:

- bibliografiske (f.eks. "billede af Eiffeltårnet taget af NN")¹¹
- administrative (f.eks. "taget den 19/04/2005")
- juridiske (f.eks. "ophavsretligt beskyttet")
- bevaringsmæssige (f.eks. "skal vises som jpeg")
- tekniske (f.eks. "jpeg-format"; "85.8kb fil")
- uddannelsesmæssige (f.eks. "En illustration af en jernkonstruktion"; Folkeskolen 8. klassetrin")
- strukturelle (f.eks. "enkeltstående billede").

Afhængigt af, hvilken type metadata der anvendes, kan man anvende forskellige standarder. Der er således eksempelvis en fælles standard for bibliografiske metadata på biblioteksområdet. Da anvendelsen af metadata er bestemt af formålet med beskrivelsen, kan man naturligvis ikke forvente, at de metadata, som kulturarvsinstitutioner anvender til at beskrive genstande i deres samlinger, er helt ens eller har indtænkt alle fremtidige benyttelser af genstandene.

Et godt eksempel er anvendelsen af digitaliseret kulturarv i undervisningssammenhænge. På undervisningsområdet anvender man en standard, SCORM (Sharable Content Object Reference Model), som er en standard for, hvilke informationer og data et læringsobjekt skal indeholde, før det kan eksporteres mellem SCORM-kompatible uddannelsessystemer. Man kan naturligvis anvende digitaliseret kulturarv til undervisningsbrug uden for disse systemer, men i nogle sammenhænge vil det være en fordel at følge standarden. I de tilfælde vil det formentlig være hensigtsmæssigt at lade de institutioner, der ønsker at udnytte objekterne, tilføje flere og andre metadata. Det vil imidlertid være hensigtsmæssigt, at kulturarvsinstitutionerne overholder fælles standarder og minimumskrav til metadata, jf. nedenfor.

Anvendelse af eksisterende metadata for fysiske objekter og manuel beskrivelse af digitale objekter er imidlertid ikke de eneste kilder til metadata. En tredje mulighed er at udnytte selve digitaliseringen i beskrivelsen. Når materialet digitaliseres, er der i nogle tilfælde mulighed for at gøre brug af automatisk dannede metadata såsom informationer trukket ud af de digitaliserede filer eller af konteksten ved høstning. Det kan dels være tekniske og dels indbyggede beskrivende metadata (f.eks. i html, xml, word og pdf). En digitaliseret tekst kan således både fungere som digitalt objekt og som kilde til metadata. Søgmaskiner kan f.eks. indekserer en digital tekst og derved gøre det muligt at søge i selve teksten frem for blot i beskrivende elementer. Dette medfører dog ofte problemer med ensartethed i stil med de problemer, der knytter sig til uensartede metadata. Der arbejdes med at udvikle automatiseret indeksering og søgning i lyd og billeder, men langt fra med de søgemuligheder, der findes for tekst.

Metadata kan også hentes fra den sammenhæng, som materialet optræder i. Der kan tilføjes flere oplysninger om et givent materiale ved at sammenholde det med materialer, der har beskrivende elementer tilfælles med dette materiale, men derudover mere righoldige beskrivelser på andre områder.

¹¹ Eksemplet er hentet fra: <http://mandate.cdlr.strath.ac.uk/metadata.htm>

En fjerde mulighed for at få digitaliseret materiale beskrevet med metadata er at få brugerne til at medvirke ved beskrivelsen af materialet. Flyfotos, hvor man blot kender den geografiske lokalitet, eller portrætter, hvor man blot kender et navn, kan typisk tilføres mere righoldige beskrivelser ved at lade brugeren bidrage.

3.4.3 Standarder for metadata

Der findes en række standarder for metadata, både mere generelle og helt specifikke. I Danmark anvendes således en række nationale og institutionsspecifikke specifikationer for metadata. For at understøtte en fælles præsentation af data fra arkiver, biblioteker og museer blev der i 2007 publiceret et sæt specifikationer, som definerer et Dublin Core-baseret fælles format¹². Udgangspunktet var nogle regionale initiativer for samarbejde om metadata, og specifikationerne er sideløbende med udviklingen anvendt af www.arkibas.dk. I forbindelse med EDLnet (forberedelse af <http://www.europeana.eu>) har Styrelsen for Bibliotek og Medier i et notat beskrevet en enkel omsætning fra dette fælles danske ABM-format til basis-eksportformatet til Europeana¹³.

Anvendelse af fælles format vil gøre det lettere at eksponere data i både nationale systemer, søgemaskiner og i Europeana. Desuden vil en standardiseret anvendelse af data fremme tværgående formidlingsinitiativer. Der bør derfor som minimum stilles krav om anvendelse af Dublin Core-baseret

¹² Det mest anvendte format for metadata kaldes Dublin Core og omfatter som udgangspunkt 15 elementer som f.eks. titel, forfatter/ophav, emne/nøgleord, beskrivelse, udgiver og anden bidrager. Betegnelsen Dublin Core skyldes, at formatet blev udarbejdet i byen Dublin, Ohio, og core referer til, at der er tale om en kerne af beskrivende metadata, der kan udvides.

¹³ Danske standarder

Specifikationer for fælles præsentation af data fra arkiver, biblioteker og museer på internettet ved ABM standard-arbejdsgruppen nedsat af Statens Arkiver, Biblioteksstyrelsen og Kulturarvsstyrelsen:

<http://www.bs.dk/standards/abm/> (omtales ofte som ABM standarden)

ARKIBAS 4: <http://www.danskearkiver.dk/Arkibas/Oma4.htm>

DAISY. Dansk Arkivalieinformationssystem: <http://daisy.sa.dk/>

danMARC2. Edb-format til inddatering og udveksling af bibliografiske data i maskinlæsbar form: <http://www.kat-format.dk/danMARC2/>

Regin: Baggrunden for Dansk Museums Dokumentations Standard:

<http://www.kulturarv.dk/forvaltning/museumsdrift/vejledninger/registrering/standard/baggrund.jsp>

Internationale standarder

Relevante internationale standarder:

EAD – Encoded Archival Description: <http://www.loc.gov/ead/>

EBU Core Metadata Set: http://www.ebu.ch/metadata/documentation/EBUCore/tec_doc_t3293_2008_FinalDraft.pdf (radio- og tv-arkiver)

DCMI Metadata Terms: <http://dublincore.org/documents/dcmi-terms/> (Dublin Core 1.1. suppleret med refinements mv.)

ISO 15836:2003 Information and documentation - The Dublin Core metadata element set. Se også: DCES, version 1.1: <http://dublincore.org/documents/dces/>

ISO 21127:2006 Information and documentation – A reference ontology for the interchange of cultural heritage information. Se også: CIDOC CRM – Conceptual Reference Model: <http://cidoc.ics.forth.gr/>

MARC21: <http://www.loc.gov/marc/marc.html>

PREMIS Preservation Metadata: <http://www.loc.gov/standards/premis/>

UNIMARC: <http://www.unimarc.net/unimarc-formats.html>

VRA Core: <http://www.vraweb.org/projects/vracore4/index.html> (visuelt materiale)

fælles format ved anvendelse af strukturerede metadata. Som nævnt ovenfor vil der i nogle tilfælde kunne hentes flere metadata, og det bør naturligvis udnyttes i det omfang, hvor det er muligt. I andre tilfælde vil det være hensigtsmæssigt at lade brugerne bidrage til beskrivelsen af objekterne. I de tilfælde kan det derfor være fornuftigt at fravige kravet om anvendelse af et standardiseret format. Man kunne forestille sig, at anvendelse af fælles ABM-format blev en bevillingsforudsætning, som kun fraviges, hvis institutionen kan begrunde en undtagelse.

3.5 Udveksling af metadata

Udveksling af metadata mellem digitale arkiver bør foregå over standardprotokoller og i standardformater. Den nok mest udbredte standard for udveksling af metadata er den såkaldte OAI-PMH-standard¹⁴. OAI-PMH 2.0 anbefales af Europeana som basis for udveksling af simple metadata mellem indholdsleverandører og Europeana¹⁵.

Standard OAI-PMH er udviklet til at høste forholdsvis simple metadata i Dublin Core (unqualified), men standarden tillader også at udstille andre metadata i andre formater (andre såkaldte metadataPrefix). OAI-PMH kan således også bruges til udveksling af metadata i andre formater, f.eks. METS, MPEG21, MARCXML, PB-CORE. Bruges OAI-PMH til dette, kræves det naturligvis at afsender og modtager er enige om, hvilke metadataformater der udveksles

Standard unqualified Dublin Core rummer 15 felter, og disse kan beskrive de fleste data til de fleste brugsscenarier. Der er bibliografiske metadata som f.eks. ophav, medvirkende, emneord og tid samt administrative metadata som f.eks. rettigheder. Der bør skrives en anbefaling til, hvorledes flere af Dublin Core-felterne anvendes, og en standard for, hvordan eksempelvis rettigheder angives.

Google stoppede desværre i begyndelsen af 2008 support for OAI-PMH til fordel for en ny standard, XML Sitemaps¹⁶. Begrundelsen for at stoppe denne understøttelse var, at Google vurderede, at OAI-PMH var utilstrækkelig til Googles behov som søgemaskine. Desuden mente Google, at OAI-PMH ikke var tilstrækkelig udbredt uden for de miljøer, hvor man arbejder med repositories (digitale arkiver) til, at de kunne fortsætte understøttelsen. XML Sitemaps er en langt simplere standard, udviklet mere specifikt til at udstille indhold (udelukkende udtrykt i URL'er) over for søgemaskiner. Denne standard er således ikke egnet til udveksling af metadata mellem arkiver, men såfremt søgemaskinerne er mål for metadata, bør arkiver kunne udstille deres indhold også i XML Sitemaps.

OAI-PMH har en standardmodel til simpel udveksling af simple metadata i form af Dublin Core simple (15 kernelementer) – eksempelvis et sæt metadata til en digital fil. Moderne repositorysystemer og arkiver bruger mere avancerede og ofte objektorienterede datamodeller til at beskrive sammenhænge mellem grupper af metadata og filer samt relationer mellem såvel grupper som enkeltobjekter. Relationer kan godt beskrives i unqualified Dublin Core vha. DC:Relations, men her er udelukkende

¹⁴ OAI-PMH – Open Archives Initiative Protocol for Metadata Harvesting:
<http://www.openarchives.org/OAI/openarchivesprotocol.html>

¹⁵ <http://www.europeana.eu>

¹⁶ <http://www.sitemaps.org/protocol.php>

mulighed for at angive links (URL'er) til andre objekter, og det kan således ikke angives, hvilken relation der er tale om. Men OAI-PHM kan også anvendes til udveksling af andre metadatamodeller og dermed også for avancerede datamodeller.

For at understøtte udveksling af metadata mellem digitale arkiver, der bruger avancerede datamodeller, har samme organisation, som stod bag OAI-PMH, de seneste år udviklet en ny standard, OAI-ORE. OAI-ORE er netop i oktober 2008 udkommet i version 1.0 og findes således nu i en stabil version.

Til indsamling af metadata til Europeana forventes fortsat anvendt OAI-PMH. I prototypen anvendes internt aktuelt et format med knap 50 elementer hentet både fra de 15 Dublin Core-kerneelementer og fra metadata fra supplerende indberetninger i faglige formater som f.eks. bibliotekernes MARC-formater. I den kommende operationelle udgave af Europeana planlægges anvendt en intern datamodel, der er udviklet over samme koncepter som OAI-ORE med brug af begreber som surrogates og aggregations. Der er aftalt et samarbejde mellem OAI-ORE og Europeana om udviklingen af standarder. Europeana skriver selv: *"This internal recursive structure of EDL surrogates is similar to the OAI-ORE [80] surrogate model and is intended to evolve in line with it."*¹⁷

Med OAI-ORE kan man således udstille grupper af data og metadata som samlede helheder i såkaldte Ressource Maps. Et Ressource Map har metadata om hele gruppen af objekter samt strukturerede links til delementer og i øvrigt mulighed for links til andre grupper såvel internt i samme arkiv som eksternt. Relationer kan i OAI-ORE beskrives f.eks. med RDF – et udbredt framework til relationsbeskrivelse udviklet til semantisk web.

Da OAI-ORE således synes at kunne blive standarden for beskrivelse af komplekse data, bør standarden undersøges yderligere og udviklingen følges nøje.

Konkluderende er anbefalingen således pt. 3-strengt:

1. Brug OAI-PMH for at være på linje med Europeana
2. Brug XML Sitemaps, såfremt metadata skal udstilles for Google
3. Følg med i udviklingen af OAI-ORE for at være fremtidssikret.

3.6 Fælles infrastruktur

3.6.1 Indledning

Som beskrevet i midtvejsrapporten behøver en fælles infrastruktur ikke at indebære, at der etableres ét centralt datacenter; infrastrukturen kan være fælles men distribueret på flere geografiske lokaliteter og institutioner. I det følgende skelnes der imidlertid ikke mellem en distribueret og en central infrastruktur. Dog er de økonomiske estimater udarbejdet på grundlag af en central infrastruktur, der ud fra et umiddelbart skøn vurderes at være den mest omkostningseffektive. Udvalget ser sig dog ikke i stand til på det foreliggende grundlag at komme med en anbefaling om,

¹⁷ EDLnet - D2.2 - Initial Semantic and Technical Interoperability Requirements

hvorvidt der bør etableres en central infrastrukturløsning – det må vurderes gennem en nøjere overvejelse i forhold til et specifikt digitaliserings scenarie.

Det bemærkes, at udgifterne til infrastruktur på grund af den løbende teknologiske udvikling under alle omstændigheder er meget vanskelige at fremskrive, især over en længere årrække.

Beskrivelsen af en mulig fælles infrastruktur tager udgangspunkt i Gartner Groups anbefaling til systemarkitektur for løsningsdesign i midtvejsrapporten (side 67). Her deles systemarkitekturen i 3 niveauer:

- 1) Et nederste niveau bestående af lagring og håndtering
- 2) Et mellemniveau bestående af tilgængeliggørelse
- 3) Et øverste niveau bestående af formidling.

3.6.2 Lagring og håndtering – bitarkiv

Statens Arkiver, Det Kongelige Bibliotek og Statsbiblioteket samarbejder allerede om løsninger, der sikrer den fysiske bevaring af de digitale materialer, de hver især har ansvaret for: Statens Arkiver har data på egne installationer og er i færd med at etablere en geografisk og mediemæssig redundans af disse data på Statsbiblioteket i Århus. Netarkivet ved Statsbiblioteket og Det Kongelige Bibliotek har replikeret sine data på geografisk adskilte installationer i hhv. København og Århus.

Statsbiblioteket har sine øvrige data replikeret mellem Statsbibliotekets forskellige adresser i Århus, mens Det Kongelige Bibliotek er i færd med at replikere sine øvrige data til Statsbiblioteket efter samme model, som anvendes i Netarkivet.

De tre institutioner har igangsat et fælles initiativ, som skal undersøge rammerne for at etablere en fælles løsning, hvor data kan bevares med forskellige niveauer for administrativ, organisatorisk og fysisk sikkerhed. Initiativet vil løfte dette samarbejde op på et sikkert fagligt niveau af international standard. Det er hensigten, at strategien efterfølgende skal danne grundlag for indgåelse af et forpligtende samarbejde mellem parterne omkring etablering og drift af et nationalt bitarkiv.

Infrastrukturen og omkostningerne ved et fælles bitarkiv for den digitaliserede kulturarv kan beskrives på baggrund af erfaringerne med en lignende løsning, som er etableret hos Statsbiblioteket, og som i et vist omfang også anvendes af Det Kongelige Bibliotek og Statens Arkiver. Løsningens hovedkarakteristika er:

- Hver enhed (fil) findes i 3 fysisk adskilte kopier hvoraf mindst én er en onlineversion.
- Sikkerhedskonceptet håndteres efter DS484, skærpet på enkelte punkter.
- Det anvendte udstyr er traditionelt datacenterudstyr med diske og båndrobotter.
- Udstyret er anskaffet gennem EU-udbud i 2006, som løbende giver mulighed for at supplere med kapacitet frem til udgangen af 2009 (kan forlænges til 2012).
- Størrelsen pr. 1. oktober 2008: 600 Terabyte.

Som indikation for omkostningerne vælges den løsning, der anvendes i Statsbibliotekets og Det Kongelige Biblioteks fælles projekt, Netarkivet. Løsningen anvender ovennævnte udstyr og arkitektur på Statsbiblioteket og er derudover udvidet med flere af de ekstra sikkerhedskrav, der forudses at

være relevante for et egentligt fælles bitarkiv. Løsningen for Netarkivet består således af 3 eksemplarer af hver fil, hvor de 2 er onlineversioner, med den ene placeret på Det Kongelige Bibliotek.

<i>KB+SB priser, prognose for TB priser i Netarkivet</i>				
	år 2009	år 2010	år 2011	år 2012
Pr. TB i arkiv med 2 disk- og 1 tapekopi	kr. 12.945	kr. 7.389	kr. 5.121	kr. 3.619

Estimatet for størrelsen af Netarkivet er 823 Terabyte (TB) i 2012. Statsbibliotekets samlede løsning (som ud over Netarkivet består af flere andre digitale samlinger med radio/tv som en dominerende del) når op på 2.624 TB til en omkostning på kr. 1.863 pr. TB i 2012. Denne løsning er imidlertid uden institutionel redundans, og det vurderes derfor, at det er mest hensigtsmæssigt at anvende Netarkivets tal som vejledende for omkostninger ved et egentlig fællesarkiv. Det skal bemærkes, at Netarkivets tal bygger på et økonomisk worst case-scenarie, idet prognoserne for omkostningen til lagringen af TB er meget usikre.

Omkostningerne i både Det Kongelige Biblioteks og Statsbibliotekets løsning inkluderer nyanskaffelse af lager i det femte år efter den oprindelige anskaffelse og inkluderer udgifter til køling, strøm og driftspersonale. Til sammenligning anslog Gartner Group i sin simulering en udgift på 1.910 kr. pr. TB ved en central løsning¹⁸. Gartner Groups tal var uden bemanning og uden indregning af teknologisk fornyelse. Ved en forøgelse af kapaciteten til Gartnerrapportens ca. 5.000 TB (i 2012) vil der være yderligere stordriftsfordel, og den anførte pris pr. TB må anses for worst case-eksempel, som vil kunne gøres billigere pr. TB, hvis alle deltager i løsningen.

Både Statsbibliotekets beregninger og Gartner groups beregninger peger således på et tal på ca. 2.000 kr. pr. TB som et minimumstal ved fuld størrelse af arkivet på ca. 5.000 TB. I opbygningsfasen anbefales det dog at anvende tabellens tal som vejledende.

Tabellen for omkostningerne ved netarkivet medfører altså et estimat på ca. 3.600 kr. pr. TB ved 823 TB i 2012. Dette sammenholdes med den noget mere nødtørftige løsning på Statsbiblioteket på 1.863 kr. pr. TB ved 2.624 TB i 2012 og Gartner Groups estimat på 1.910 kr. pr. TB ved 5000 TB i 2012.

3.6.3 Lagring og håndtering – objektniveauet

Håndteringen af objekter vil i meget høj grad afhænge af disses form og kompleksitet. Håndteringen af digitaliseret film, som er meget store filer, er fundamentalt anderledes end håndteringen af eksempelvis websider, som består af mange små linkede objekter. Der er derfor forskellige systemer til beskrivelse, formidling og bevaring af disse objekter.

Det kan etableres et fælles lagersystem for de digitale objekter. Et sådant system kaldes ofte for et Digital Object Management System eller kort et DOMS. Et DOMS kan være distribueret, således at det består af et netværk af servere, der f.eks. håndterer forskellige filtyper i den samlede infrastruktur. Et DOMS er designet til at understøtte en række processer, herunder aktiviteter i forbindelse med bevaring af objekter og de nødvendige ændringer, som gennemføres på objekter. Eksempler på sådanne operationer er, at man kopierer objektet fra et disksystem til et andet, eller at

¹⁸ Gartner Group: Finansielle analyser i forbindelse med digitalisering af kulturarven, 10. maj 2007, side 28. Rapportens specifikke angivelse er 5,6 petabyte til 1,91 kr. pr. gigabyte pr. år

man opdaterer formen, eksempelvis fra en gammel Word-version til en nyere eller fra et gammelt filmformat til et nyere.

Et DOMS kan også varetage håndtering af rettigheder, således at materiale kun udleveres i henhold til gældende aftaler. Endelig kan systemet styre, hvordan de forskellige objekter vises.

Det er vigtig at planlægge, hvordan materialer indlemmes i systemet, og specielt, hvordan man automatisk kan generere metadata som en integreret del af processen. Det har ikke nødvendigvis noget med selve systemet at gøre. I tilknytning til DOMS kan der etableres et workflowsystem, som tillader systemet at håndtere i praksis alle typer digitale objekter med tilhørende metadatastrukturer, og som understøtter digitaliseringsarbejdsgange med metadatatilføjelser.

De forskellige institutioner har forskellige tilgange til, hvordan objekter skal beskrives og håndteres. I modsætning til den ovenfor beskrevne løsning for bitarkivet er det mindre oplagt, at der kan skabes en fælles løsning på dette område.

Inden for bibliotekssektoren har de to store biblioteker valgt at prioritere forskelligt – en prioritering som i høj grad afspejler de to institutioners fokus på forskellige opgaver på nuværende tidspunkt. Der er imidlertid påbegyndt en dialog, som kan føre frem til et fælles system. De to forskellige tiltag har således fokus på forskellige relevante aspekter og kan dermed måske forenes: Statsbiblioteket fokuser på en dynamisk beskrivelse og etablering af workflowsystem som ovenfor beskrevet, og Det Kongelige Bibliotek har fokus på bevaringsaspektet.

Et DOMS kan med andre ord anvendes til at håndtere en samling bestående af digitale objekter og fungere som både lager og administrativt værktøj til at styre eksempelvis rettigheder til objekterne. Det kan endvidere fungere som grundlag for en tilgængeliggørelse over for brugerne.

3.6.4 Arkitektur for tilgængeliggørelse

Tilgængeliggørelsen af materialet kan foregå på mange måder. De oplagte er et system baseret på en søgeadgang (som Google) eller systemer, som baserer sig på geografiske interfaces. Førstnævnte system baserer sig i dag på de strukturerede og ustrukturerede ord, som findes i DOMS'er, det andet vil kræve, at materialet tilføjes metadata om lokalitet (geokodes).

Det er vigtigt, at materialet kan eksponeres så bredt som muligt, herunder Europeana (som understøtter begge tilgange), Google og andre søgemaskiner.

Gennem anvendelse af DOMS'er kan der tilbydes en snitflade mod de digitale objekter, som betyder, at disse kan synliggøres gennem andre produkter og Portaler som Google og Europeana (jf. ovenfor om udveksling af metadata).

De to store biblioteker arbejder begge på at etablere effektive søgesystemer af "integrated search"-typen (Primo og Summa)¹⁹.

En løsning baseret på DOMS i kombination med et sådant søgesystem vil kunne tilbyde en effektiv tilgængeliggørelse af indhold af objektsamlinger. Anvendeligheden vil selvfølgelig afhænge af kvalitet

¹⁹ Se Midtvejsrapporten side 67-72.

og mængde af metadata. Statsbiblioteket har allerede etableret en Summa-løsning, som fungerer i sammenhæng med DOMS. Omkostningerne til tilføjelse af yderligere digitale samlinger til denne infrastruktur afhænger af omfang og struktur af de metadata, der forefindes for de enkelte samlinger.

Et særligt aspekt i tilgængeliggørelse er adgangsstyringen. *Der kan være materiale, hvor adgangen og benyttelsen er begrænset til en nærmere defineret kreds af brugere. For dette materiale findes der i princippet to forskellige løsninger, som begge vil skulle understøttes.*

Den ene er, at rettighedshaverne til materialet indgår som aktiv part i tilgængeliggørelsen og etablerer separat adgangsstyring og afregningssystem. Eksempler på dette er Statsbibliotekets aftale med musikindustrien i Netmusik.dk og en lignende aftale om tilgængeliggørelse af aviser med en privat udbyder. I disse tilfælde er der ingen statslige ekstraudgifter til tilgængeliggørelse. Rettighedshaveren vil skulle dække procesudgifterne i tilfælde af, at rettighedshaveren ønsker at trække materiale fra arkivet. Brugeren vil skulle betale for servicen.

Det andet eksempel er brug af aftalelicenser eller lignende til at sikre, at kun bestemte brugere har adgang til materialet. Eller tilfælde hvor lovgivningen tilsiger at bestemte brugergrupper har adgang til materialet (som f.eks. medieforskere til radio/tv-arkivet).

I sidstnævnte tilfælde vil det være naturligt at etablere et adgangsstyringssystem som del af infrastrukturen. Et godt bud på et sådant, der dækker store dele af den relevante brugergruppe, er repræsenteret i WAYF-samarbejdet mellem 3 ministerier. Statsbiblioteket har senest anvendt denne løsning til at tilgængeliggøre 30.000 digitaliserede reklamefilm for forskning og uddannelse i Danmark. Ud over selve WAYF-strukturen krævedes i dette tilfælde, at der i forbindelse med hvert udtræk bliver indlagt vandmærke i objektet med brugerens identitet til beskyttelse af misbrug.

I forbindelse med digitaliseringen vil det for noget materiale være relevant at undersøge perspektiverne i at anvende licenstyper som Creative Commons. Denne licenstype kan f.eks. oplyse om, at det digitaliserede materiale gøres tilgængeligt på ikke-kommercielle vilkår, og at det kan anvendes på bestemte betingelser. Sådanne standardiserede beskrivelser af vilkårene for benyttelse af materialer kan omsættes til maskinlæsbare licenser og dermed være særlig relevante i en distribueret infrastruktur. Hvis et digitaliseret objekt eksempelvis optræder i en helt ny formidlingssammenhæng, kan det være hensigtsmæssigt, at vilkårene for benyttelse er beskrevet i metadata.

3.6.5 Formidling

Det øverste niveau i midtvejsrapportens systemarkitektur er formidling. Som tidligere nævnt betragtes mere målrettede formidlingsløsninger, som eksempelvis DR's Bonanza og Kulturarvsportalen, ikke som en del af infrastrukturen og beskrives derfor ikke nærmere.

Dog vil en minimal formidling, herunder eksponering mod Google og Europeana, kunne ske direkte oven på infrastrukturen, såfremt der etableres en adgangsstyring og on the fly-levering af digitale objekter fra arkivet. I forbindelse med etablering af prototypen af Europeana er der indsamlet metadata fra et begrænset antal bidragsydere i Europa med en stor manuel indsats, stort set uden nyere automatiske værktøjer som OAI-PMH. I den planlagte operationelle version af Europeana forventes det, at hvert land etablerer en national aggregator, som står for at samle og klargøre

metadata til Europeana. Det er oplagt at løse opgaven som national dansk aggregator til Europeana i sammenhæng med tilgængeliggørelse af dansk digitaliseret materiale.

I det omfang materialet eksponeres på internettet og formidles mere aktivt, er der behov for, at infrastrukturen kan betjene mange samtidige brugere med eksemplarer af det efterspurgte materiale.

Det anslås, at en pulje på tre hurtige servere tilknyttet et centralt bitbevaringsarkiv vil kunne betjene i størrelsesordenen af 100 samtidige brugere. I en sådan løsning vil de digitale objekter blive tilgængeliggjort via on the fly-konvertering til formidlingsversion, således at denne kun optager midlertidig ekstra lagerplads.

Der vil være behov for specialiserede servere for nogle samlingstyper som video og radio. Omkostningerne vil afhænge af den endelige arkitektur, målgruppens størrelse og omfanget af benyttelsen. En løsning kunne være at anvende en løsning, der ligner den løsning, som DR er ved at udarbejde for at betjene sine brugere på nettet.

DR er ved at opgradere det eksisterende streaming setup og analysere flere forskellige løsninger. Løsningen skal kunne håndtere mange samtidig brugere. En af løsningerne er en meget kraftig server, der kan håndtere ca. 10.000 samtidige brugere. Denne kan indeholde 5 TB materiale. Disse vil udgøre de mest sete og/eller senest anvendte filer. Efterspørger en bruger en anden fil end dem, der ligger på onlineserveren, hentes den fra et internt lager eller transkodes on the fly som i bitarkivsløsningen fra et ”billige” bagvedliggende diskarkiv. Filer der ikke er efterspurgte, fjernes automatisk fra onlineserveren.

Serveren koster ca. 2 mio. kr., og det antages, at DR skal etablere en redundantløsning for at sikre fuld adgang til institutionens onlinesatsninger som dr.dk/tv, Bonanza, netradiokanaler osv. Den samlede omkostning vil være på 4 mio. kr. uden reinvesteringer og driftsomkostninger. Løsningen vurderes at kunne kobles med bitarkivet arkitektur, uden at de nærmere detaljer har været analyseret konkret. En sådan investering vil efter statslige regler skulle afskrives over en periode, typisk lineært afskrevet. Hvis investeringen betragtes som en ordinær hardwareinvestering, er afskrivningsperioden tre år med en årlig udgift på ca. 1,3 mio. kr. Såfremt investeringen betragtes som led i et udviklingsprojekt, er der større råderum for at fastlægge afskrivningsperioden, men den kunne f.eks. være fem år. Det ville give en årlig udgift på 800.000 kr. Det er næppe sandsynligt, at de samlinger, der udvælges til digitalisering i hovedrapporten, vil give samme behov for serverkapacitet som hos DR. På den baggrund burde en årlig udgift på ca. 800.000 kr. være tilstrækkeligt til at sikre tilgængeliggørelse af video og radio.

I praksis vil der skulle allokeres mandskab til support og drift til arkivet ved direkte anvendelse af arkivet til tilgængeliggørelse. Det indebærer, at omkostningerne til tilgængeliggørelse omfatte ca. 1 årsværk samt de nævnte servere, der vil andrage en årlig udgift på 750.000 kr. I praksis vil det også være nødvendigt at gennemføre visse udviklingsopgaver (som f.eks. den nævnte vandmærkning i reklamefilmseksemplet). Omkostninger til udvikling er dog ligesom servere til specialiseret tilgængeliggørelse udeladt af estimatet. Omkostningerne til en adgangsstyringsinfrastruktur kan estimeres som Kulturministeriets nuværende bidrag til WAYF, som anslås til 25 % af WAYF-årsbudgettet på 3 mio. kr., dvs. 750.000 kr. pr. år.

3.6.6 Estimat for minimale udgifter til infrastruktur

Beskrivelsen af den fælles infrastruktur anvendes til at fremhæve, at der allerede findes systemer og kompetencer, som kan bidrage til en fælles infrastruktur for digitaliseret kulturarv. Desuden kan beskrivelsen anvendes til at estimere omkostningerne ved en fælles infrastruktur. Estimaterne for den fælles infrastruktur er samlet i nedenstående tabel.

<i>Årlige omkostninger til elementer i infrastrukturen</i>	år 2009	år 2010	år 2011	år 2012
<i>Pr. TB i arkiv med 2 disk- og 1 tapekopi</i>	kr. 12.945	kr. 7.389	kr. 5.121	kr. 3.619
Samlet udgift til adgangsstyring	kr. 750.000	kr. 750.000	kr. 750.000	kr. 750.000
Samlet udgift til medieservere	kr. 800.000	kr. 800.000	kr. 800.000	kr. 800.000
Samlet udgift til tilgængeliggørelse	kr. 750.000	kr. 750.000	kr. 750.000	kr. 750.000

Som det fremgår af tabellen, arbejdes der med en omkostning pr. TB som i Netarkivet. Enhedsomkostningerne falder kraftigt med tiden især pga. stordriftsfordele. De årlige omkostninger vil imidlertid være relativt konstante, fordi mindre mængder i de første år skal håndteres til højere enhedspriser end de større mængder, der senere skal håndteres af infrastrukturen, efterhånden som digitaliseringen skrider frem.

Estimaterne indebærer, at hvis der i 2012 er 2.500 TB i arkivet, vil de årlige driftsudgifter til infrastrukturen således være: 2500 TB á 3.619 kr. pr. TB + 800.000 kr. + (2 x 750.000 kr.) = 11.347.500 kr. Det bemærkes, at der er tale om årlige omkostninger.

Estimatet for de årlige omkostninger til fælles infrastruktur kan specificeres og kvalificeres yderligere, når mere præcise lagerstørrelser for de prioriterede digitaliseringsprojekter foreligger. Derudover er der en række variabler, hvis værdi vil kunne fastlægges nøjere i budgetteringssituationen. Der kan således være behov for en nærmere analyse af lønomkostningerne ved tilgængeliggørelse samt løbende udviklingsomkostninger. Omvendt kan adgangsstyring og dedikerede medieservere vise sig at være mindre relevant eller kunne dækkes af eksisterende løsninger på institutionerne. Desuden kan priser pr. TB blive lavere som konsekvens af den tekniske udvikling. Endelig kan der være dele af infrastrukturen, der kan dækkes af eksisterende bevillinger.

Der er således faktorer, der tilsiger, at estimatet kunne være højere, og andre faktorer, der måske kunne tale for at nedjustere. Der er på baggrund af beskrivelserne af scenarierne for digitalisering enighed i arbejdsgruppen om at tage udgangspunkt i en lagerstørrelse på 2500 TB, svarende ca. til scenarie 2. Et fuldt udfoldet scenarie 2 vil således indebære udgifter på ca. 10 mio. kr. årligt, mens et fuldt udfoldet scenarie 3 vil indebære noget højere årlige udgifter.

4 Sammenfatning

En vigtig målsætning for gruppens arbejde har været at beskrive en infrastruktur, der kan understøtte tværgående formidling af kulturarven. Infrastrukturen omfatter ikke en portal som indgang til kulturarven. Den foreslåede infrastruktur vil omfatte en søgeadgang til det samlede digitaliserede materiale og andre interfaces til andre dele af materialerne (f.eks. et Danmarkskort til at fremfinde relevant materiale). Den vil ligeledes understøtte mange andre formidlingsinitiativer, som dog ikke udgør en del af den grundlæggende infrastruktur.

Infrastrukturen vil understøtte præsentation af metadata i mange formidlingssammenhænge. Det kan finde sted ved at metadata høstes, at de stilles til rådighed gennem webservicer eller gennem direkte søgning i metadata. I de tilfælde, hvor metadata beskriver et digitalt objekt, vil der være et stabilt link til selve objektet. Det giver mulighed for at præsentere den digitaliserede kulturarv i eksisterende formidlingssammenhænge f.eks. inden for uddannelsessektoren

Som beskrevet i kapitel 2 er digitaliseringsprocesserne for de forskellige materialetyper temmelig komplicerede. Der findes imidlertid allerede kompetencer og samarbejdsstrukturer, der kan håndtere udfordringerne for de forskellige materialetyper. Nogle af de større institutioner fungerer allerede som kompetencecentre, men der kan være grund til at formalisere og udbygge denne rolle med nogle af de funktioner, der er beskrevet i afsnit 2.2. Der vil under alle omstændigheder være behov for et tættere samarbejde mellem institutionerne både for at koordinere formidlingen af den digitaliserede kulturarv og anvendelsen af de standarder, som er forudsætningen for en tværgående formidling.

Som beskrevet i afsnittet om metadata kan disse fremskaffes fra flere kilder:

- Metadata fra andre registre: bibliotekernes kataloger samt museers og arkivers registre
- Automatisk dannede metadata: informationer trukket ud de digitaliserede filer eller af konteksten ved høstning
- Manuel registrering i forbindelse med digitaliseringen
- Brugerskabte data: supplerende data fra brugere.

Udgangspunktet vil i alle tilfælde være anvendelse af eksisterende strukturerede metadata fra eksisterende registre og et krav om at anvende det fælles ABM-format eventuelt som en comply or explain-strategi. Derudover bør man naturligvis udnytte andre kilder til metadata under skyldig hensyntagen til omkostningerne. En række af de ønsker, der måtte være til metadata som led i en bestemt anvendelse af formidling, må opfyldes af de institutioner, der ønsker at anvende den digitaliserede kulturarv.

Afhængigt af, hvordan infrastrukturen opbygges, vil der blive stillet krav til de lokale systemer. Som beskrevet i afsnit 3.5 bør disse systemer kunne anvendes OAI-PMH for at følge Europeana og XML Sitemaps, såfremt metadata skal udstilles for Google. Desuden må institutionen følge udviklingen af OAI-ORE.

Som det er fremgået, er der grundlag for at overveje en fælles infrastruktur særligt i relation til bitbevaring. Det er derimod muligt, at infrastrukturen skal være fælles, men distribueret på objektniveauet. Som det fremgår, kan der imidlertid etableres en infrastruktur for tilgængeliggørelse i

umiddelbar tilknytning til bitarkivet. Såfremt der etableres en mere omfattende indsats for digitalisering af kulturarven, bør disse forhold analyseres og drøftes nærmere, idet der ikke kun er tale om tekniske overvejelser.

På baggrund af en analyse af udgifterne til en minimal infrastruktur baseret på et fælles lager estimeres de samlede udgifter til 10,5 mio. kr. baseret på en lagerstørrelse på 2500 TB. Som beskrevet ovenfor kan yderligere information for en række nøglevariabler både øge og sænke estimatet. Det er arbejdsgruppens opfattelse, at man med en årlig investering på 10,5 mio. kr. vil kunne etablere en infrastruktur, der både kan bevare og lagre den digitaliserede kulturarv og understøtte tværgående formidling på en sådan måde, at materialet bliver tilgængeligt via institutionerne, i en fælles arkiv med DC-metadata, et geografisk interface og i Europeana og søgemaskiner som Google.