

Langtidsopbevaring af digitale data for Det Kongelige Bibliotek.

Forslag til handlingsplan for fastlæggelse af metoder

Indholdsfortegnelse

Del I:

Afsnit 1: Indledning

Afsnit 2: Målsætning

Afsnit 3: Elementer i handlingsplanen

- 3.1: Fuldstændighed og genanvendelighed
- 3.2: Hvilke materialegrupper er omfattet af handlingsplanen
- 3.3: Hvilke datatyper og formater er omfattet af handlingsplanen
- 3.4: Hvilke datamedier er omfattet af handlingsplanen og hvilke krav stilles mht. de fysiske mediers langtidsofbevareling
- 3.5: Hvilken metode er teknisk realistisk og økonomisk rentabel at anvende til langtidsofbevareling af hvilke materialegrupper og datatyper
- 3.6: Sikring, kontrol og organisering. Organisering og administration af langtidsofbevareling af digitale data
- 3.7: Tilvækst i dataomfang
- 3.8: Autenticitet og autentifikation
- 3.9: Metadata for bevareling af digitalt materiale (digitale objekter)
- 3.10: Rettighedsoplysninger og adgangsstyrling
- 3.11: Identifikation i forbindelse med det arkiverede materiale
- 3.12: Tilpasning af KBs egen produktion af digitalt materiale og lagringsstrategi

Del II:

(udarb. af Fischer & Lorenz)

Afsnit 4: Praksis på dette område i andre lande og i andre organisationer

- 4.1: Status fra andre biblioteker, arkiver og virksomheder
- 4.2: Status fra samarbejdsprojekter, interessegrupper og lign.
- 4.3: Open Archive Information System (OAIS)

Afsnit 5: Indstilling om realisering af handlingsplan

- 5.1: Tidsplan og budget for projekteringsprojektet
- 5.2: Projektforberedelse
- 5.3: Analysearbejde
- 5.4: Udarbejdelse af handlingsplan

Del I

Afsnit 1

Indledning

I gennem 1990'erne er der set en stærk stigning i produktionen af digitalt materiale. I takt med denne stigning vokser behovet for at udvikle systemer til håndtering af materialet i de institutioner, som har arkiv-forpligtelserne for det.

Systemer til indsamling og langtidsopbevaring samt udvikling af metoder til at sikre tilgængelighed til materialet på langt sigt står derfor i stigende grad på dagsordenen i biblioteker og arkiver verden over – ikke blot på det teoretiske plan i form af konferencer, artikelskrivninger og løbende debatter, men også i form af praktiske implementeringer.

Således er praktiske implementeringer i gang på nationalbibliotekerne i Australien, Holland, England og Norge. Der er her tale om store investeringer i størrelsesordenen kr. 20-30 mill.

I USA har The Library of Congress har netop modtaget \$ 100 mill fra Kongressen til at udvikle et landsdækkende program til bevaring af specielt det materiale, der er født digitalt, med henblik på at sikre det for efterfølgende generationer.

På Det Kongelige Bibliotek er der på tilsvarende vis behov for en plan for, hvorledes vi vil tage hånd om det digitale materiale, vi har bevaringsforpligtelserne for i Danmark.

Væsentlige dele af kulturarven, som det skønnes vigtigt at bevare for eftertiden af kulturelle eller forskningsmæssige grunde, foreligger allerede nu alene i digital form. Det er et nyt område, hvor informationsmængden vokser eksplosivt, samtidig vides det med sikkerhed, at informationen vil forsvinde, hvis der ikke gøres en indsats for at indsamle den og gøres en indsats for at udvikle metoder til at bevare den for eftertiden. Som nævnt ovenfor har lande, som Danmark normalt sammenligner sig med, allerede iværksat projekter til løsning af disse opgaver. Det er nødvendigt at et tilsvarende arbejde iværksættes her i landet.

Nærværende dokument er et forslag til, hvordan arbejdet med langtidsopbevaring af digitale data igangsættes.

Sammenfatning

Forslaget er udarbejdet af Det Kongelige Bibliotek i samarbejde med konsulentfirmaet Fischer og Lorenz og har til formål at fremsætte en handlingsplan for fastlæggelse af egnede metoder til varetagelse af langtidsopbevaring af digitale data for Det Kongelige Bibliotek.

Forslaget indeholder derfor

- en beskrivelse af de problemstillinger, som er knyttet til langtidsopbevaring af digitale data
- en status for arbejdet med denne problemstilling internationalt og i norden
- et emnekatalog over de problemstillinger, som handlingsplanen skal omfatte
- en handlingsplan for udarbejdelse af en operationel handlings-plan for langtidsopbevaring af digitale data og dets implemen-tering på Det Kongelige Bibliotek samt
- et budget for omkostningerne ved udarbejdelsen af en handlingsplan for langtidsbevaring af digitale data.

Hovedsynspunktet er, at det er bydende nødvendigt at iværksætte arbejdet med udarbejdelse af en operationel handlingsplan.

Konsulentfirmaet Fischer og Lorenz har udarbejdet et budget for opgavens løsning på i alt 2,1 mio kr., idet der kalkuleres med 1,2 mio kr. til konsulentassistance og dertil anbefales, at KB kalkulerer med en udgift på ca. 0,9 mio. kr (= 2 år) til projektarbejde og projektledelse.

Opgaven er tidsmæssigt skønnet til 9 måneder.

Handlingsplanen kan både være udgangspunkt for et selvstændigt arbejde på Det Kongelige Bibliotek og for tilrettelæggelse af en handlingsplan, der inddrager flere nationale institutioner, herunder især Statsbiblioteket.

Afsnit 2

Målsætning for langtidsoptbevaring af digitale data for Det Kongelige Bibliotek

Biblioteket erhverver værker og information til sine samlinger på grundlag af sine forpligtelser som nationalbibliotek, universitetsbibliotek og hovedfagbibliotek og skal derfor kunne tilgængeliggøre og bevare information på alle relevante medier. IT-udviklingen betyder, at information i stigende grad publiceres på digitale medier, således at man kan tale om digitale værker på linje med konventionelle værker, d.v.s. værker udgivet på konventionelle medier som papir.

Biblioteket skal, når værkernes karakter og indhold berettiger til det kunne bevare værkerne på langt sigt uanset om de er konventionelle eller digitale. Biblioteket ønsker endvidere, at disse værker bevarer deres autenticitet i form og indhold i videst muligt omfang.

Medens der foreligger velkendte metoder for langtidsoptbevaring af værker i konventionel form er det samme ikke tilfældet for langtidsoptbevaring af digitale værker. Digitale værker er kendetegnet ved en række specielle forhold betinget af, at værkets benyttelse er helt afhængig af de bærende computersystemer. Denne systemafhængighed giver også værkerne dynamiske og interaktive funktioner, som er nye og komplekse i forhold til de konventionelle værker. En høj grad af autenticitet i bevaringen af digitale værker kræver således, at man også er i stand til at bevare de digitale værkers funktionelle kvaliteter.

Et nyt system for langtidslagring af digitale data skal tilføjes følgende nye funktionaliteter til biblioteket:

- lagring af alle typer digitalt materiale i KBs samlinger
- en facilitet til administration og "konservering" af de digitale objekter, som biblioteket har en forpligtelse til at give adgang til for kommende generationer
- en funktion til at danne og vedligeholde metadata, der knytter sig til de digitale objekter
- en facilitet til at udveksle metadata med andre af bibliotekets systemer

- en facilitet, der gør det muligt for brugerne at identificere, søge, fremtage og benytte de digitale objekter og gøre det på en måde, der sikrer opretholdelsen af de restriktioner, der gælder for benyttelsen af digitale værker.

Afsnit 3

Elementer i handlingsplanen

3.1

Samlingsopbygning, fuldstændighed og genanvendelighed

Samlingsopbygning

Det Kongelige Bibliotek erhverver værker og information til sine samlinger på grundlag af sine forpligtelser som nationalbibliotek, universitets-bibliotek og hovedfagbibliotek og skal derfor kunne tilgængeliggøre og bevare information på alle relevante medier.

Det Kongelige Bibliotek følger principielt de samme principper for samlingsopbygning, selektion og bevaring for digitale værker som for værker i konventionel form.

Erhvervelsen af værker sker på 3 måder: ved køb, ved gave og ved pligtaflevering.

Køb

Biblioteket råder over et budget til materialeanskaffelser og en stab af fagkyndige medarbejdere, som varetager køb af materialer til bibliotekets samlinger med udgangspunkt i bibliotekets anskaffelsespolitik for de enkelte anskaffelsesområder. Der er her tale om en selektiv, målrettet, langsigtet samlingsopbygning. Sker det ved modtagelsen, at det købte værk ikke svarer til forventningen, indlemmes det ikke, men returneres eller frasorteres. Denne fremgangsmåde gælder for såvel konventionelle som digitale værker.

Gave

Ved gaver gælder det, at der ved modtagelsen af gaven vurderes, om det er af interesse at erhverve gaven. Hvis kun nogle enkelte værker i gaven har interesse, erhverves alene disse. Hvis det drejer sig om en større samling af værker, hvor hovedparten af gaven har interesse, erhverves hele gaven, men altid med den klausul, at ikke-relevante dele kan frasorteres, og at gaven ikke vil blive opbevaret samlet som en særsamling. I forbindelse med indlemmelse af gaven sker der en frasortering af værker, som det vurderes ikke er af betydning at bevare. De samme principper gælder for digitale værker.

Pligtaflevering

I pligtafleveringen foregår en indsamling af værker på basis af pligtafleveringsloven. Målet er at skabe en så fuldstændig opfyldelse af pligtafleveringslovens intentioner som muligt. Det er dog ikke således, at alt, hvad der udgives, bevares. Visse kategorier af værker er i henhold til bekendtgørelsen undtaget for aflevering. Desuden foretages der yderligere en frasortering efter principper, der er godkendt af Kulturministeriet (i pligtafleveringslovens terminologi kaldet kassation). Frasortering sker bl.a. for at sikre, at der ikke opbevares værker af helt ensartet karakter, hvor der alene gør sig f.eks. lokale forskelle gældende, eller hvor det pågældende værk ikke tilføjer noget nyt i forhold til allerede eksisterende information.

For digitale værker følges de samme principper, og digitale værker, der ikke falder ind under pligtafleveringsloven, bliver frasorteret.

Registrering og tilgængeliggørelse af indlemmede værker

I forbindelse med indlemmelse af værker skal det sikres, at disse kan genfindes. Det kan ske enten ved registrering af værket i bibliotekets kataloger eller ved sortering og opstilling efter arkivalske systemer. Registrering til katalogen kan afhængigt af formålet variere fra national-bibliografisk niveau til simple registreringer.

For værker i digital form vil der desuden være mulighed for at udnytte maskinel indeksering og tilgængeliggørelse af værkerne via dertil svarende søgeværktøjer.

De digitale værkers tilgængelighed og anvendelighed

Medens konventionelle værker som f.eks. bøger altid vil kunne genanvendes, såfremt de er intakte, og der er det fornødne lys, stiller sagen sig anderledes for digitale værker. Digitale værker stiller særlige krav til den langsigtede opbevaring på grund af værkernes afhængighed af de bærende edb-systemer.

Det er derfor en særskilt opgave at analysere vilkår og muligheder for fremvisning og funktionalitet hos de digitale værker. Dette spørgsmål er taget nærmere op i afsnit 3.5 vedrørende langtidsofopbevaring.

For at en indsamling og bevaring af værker skal være meningsfuld, skal værkerne også efter lang tids opbevaring kunne anvendes efter deres formål. Man kan her skelne mellem forskellige niveauer af anvendelighed. Denne har en nøje relation til værkernes autenticitet.

A. Fuldstændighed i anvendelsen af værket

1. Værket kan bruges med fuld autenticitet, d.v.s. både med hensyn til informationsindholdet og med hensyn til funktionalitet. Med funktionalitet kan man f.eks. tænke på de dialog- og søgefunktioner, som er knyttet til et digitalt leksikon, således at man fuldt ud kan bruge det pågældende værk, som man kunne, da værket blev publiceret.

2. Fuld autenticitet i anvendelsen som ovenfor og med fuld dækning, også af de referencer, der indgår i værket, men som rækker ud over dette. Dette indebærer, at også mål for referencer uden for værket er bevaret og tilgængelige enten internt på egne servere eller eksternt i henhold til særlige aftaler herom.

B. Ufuldstændig genanvendelse

Der kan tænkes en lang række forskellige former for ufuldstændighed fra mindre dysfunktioner i funktionaliteten til, at man overhovedet ikke er i stand til at få fremvist værkets data.

Som beskrevet i afsnit 3.5 er der en række metoder til at sikre værkets anvendelighed på sigt.

Et af de væsentlige bidrag til langsigtet bevaring er systematisk dekomponering af værkerne, således at man har alle data til rådighed i struktureret og opmærket form og har dækkende beskrivelser af funktionaliteten i det edb-system, som gav værket dets dynamiske egenskaber. Hertil kommer naturligvis, at man har bevaret applikationerne, som er i stand til sammen med data at genskabe funktionalitet, herunder at man har varianter af applikationerne i form af nye applikationer, der kan emulere de tidligere applikationers funktionalitet. Edb-systemernes meget hurtige udviklingstakt betyder, at der er en særlig udfordring med hensyn til sikringen af bevaringen af de digitale værkers dynamiske egenskaber.

Da de dynamiske egenskaber er et væsentligt træk ved de digitale værker, er det vigtigt, at man ved væsentlige værker også dokumenterer de dynamiske egenskaber, når der er risiko for, at applikationen mister sin funktionalitet, og der ikke er udsigt til etablering af emulatorer for den pågældende applikation. En sådan dokumentation kan f.eks. ske ved, at man på video optager karakteristiske brugerdialoger med det pågældende system og lagrer denne video om funktionaliteten sammen med funktionsbeskrivelsen og den originale applikation.

Denne form for dokumentation af et digitalt værks dynamiske egenskaber kan forventeligt kun gennemføres meget selektivt.

Det er indlysende, at de økonomiske omstændigheder ved gennemførelse af bevaring og sikring af anvendeligheden af de digitale værker indgår som en helt afgørende faktor i forbindelse med tilrettelæggelse af metoder for langtidsbevaring og for sikring af værkeres genanvendelighed. Det er endvidere indlysende, at sammenhængen mellem økonomi og metoder til langtidsopbevaring og genanvendelighed har indflydelse på graden af fuldstændighed ved bevaring af digitale værker på langt sigt.

3.2

Hvilke materialegrupper er omfattet af handlingsplanen

Det digitale materiale, der findes på Det Kongelige Bibliotek kan grupperes således:

- a) Digitalt materiale der indgår i bibliotekets samlinger gennem pligtafleveringen, hvad enten det er netbåren eller ej (eks: netpublikationer, cd-rom'er, GIS-applikationer, multi-medie-applikationer)
- b) Digitalt materiale der indgår i bibliotekets samlinger gennem accession eller gave, hvad enten det er netbårent eller ej (eks: digitale brevsamlinger i form af en e-mail samling, digitale billedsamlinger, digitale noder eller kulturinstitutionelle intranet).
- c) Digitale dokumenter der er resultat af egen digitalisering af materiale i konventionel form (eks: billederne i billeddatabasen, Håndskrift-afdelingens og Musikafdelingens digitaliserede manuskripter, billed- eller tekstversioner af f.eks. danske litteraturudgivelser i f. eks. ascii eller opmærket i f.eks. XML eller til brug i e-bøger)
- d) Digitalt materiale, der fremstår som digitale publikationer oparbejdet af Det Kongelige Bibliotek (eks: online katalogen REX og bibliotekets paralleludgivelse på latin, engelsk og dansk af Tycho Brahes *Astronomiæ instauratæ mechanica* (1598), Arkiv for Dansk Litteratur, virtuelle studiesamlinger, 'e-learning' kurser, e-bøger til offline brug)
- e) Digitalt materiale i form af cd-rom baser eller adgang til licenseret materiale (bl.a. elektroniske tidsskrifter eller databaser (eks. adgang til *Journal of Social History* eller *PsycLIT 1887 – 1995*).
- f) Bibliotekets administrative digitale dokumenter (eks: administrativ korrespondance).

Kun grupperne a) til e) omfatter således det, vi traditionelt benævner KBs samlinger. Det licenserede webmateriale adskiller sig fra materiale-grupperne a) til d) ved ikke at være unikt dansk eller bundet af en dansk lovgivning. Det findes eller anvendes på en lang række forsknings-biblioteker verden over, og det vil derfor være mere relevant at se på det projekt eller prototype, der kaldes LOCKSS (Lots Of Copies Keep Stuff Safe) og er udviklet af Stanford University og Sun Microsystems, til sikring af dette materiale (jfr. <http://lockss.stanford.edu>). Filosofien i

LOCKSS er, at findes og anvendes materialet til stadighed på 30-40 andre servere i verden, vil dette være tilstrækkeligt til at sikre dets eksistens og dermed den permanente adgang til det, hvorfor hovedopgaven er at monitorere, hvor de produkter man selv ønsker at bevare, anvendes og opbevares. Prototypen skal i 2001 beta-testes på bl.a. Lunds og Bergens Universitet.

På nationalt plan skal der tages stilling til, hvorledes en langsigtet tilgængelighed til det licenserede materiale skal sikres i Danmark, hvorfor vi her vil antage at dette bliver en opgave i Danmarks Elektroniske Forskningsbiblioteks regi, der evt. kan blive relevant for KB på et senere tidspunkt. Materialet vil derfor ikke blive omfattet af denne handlingsplan.

Med hensyn til KBs samlinger vil der således blive fokuseret på materialegrupperne a) til d) i handlingsplanen og en af de ting, der skal behandles, er således hvorvidt og i givet fald hvordan, en selektion skal finde sted.

Punkt f) vedrører KBs arkiv virksomhed, der falder ind under Statens Arkivers arkivforpligtelser og derfor ikke vil blive omfattet af denne handlingsplan.

3.3

Hvilke datatyper og formater er omfattet af handlingsplanen

Formelt kan man opdele alle digitale ressourcer i følgende datatyper:

	Datatype	Eksempler på formater	Relevant for materialegruppe omfattet af KB's handlingsplan
1	Data sæt /databaser	f.eks. DanMARC2 poster i online katalogen REX, evt. formateret som kommasepareret ASCII	a) b) c) d)
2	Struktureret tekst	f.eks. tekst i HTML opmærkning, tekst i XML TEI xlite opmærkning, tekst opmærket til e-bøger	a) b) c) d)
3	Office dokumenter	f.eks. korrespondance	-
4	Design data	f.eks. vektor grafik fra CAD udstyr	-
5	Præsentationsgrafik	f.eks. publikationer i pdf	a) b) måske c) d)
6	Billeder	f.eks. JPEG, TIFF, GIF	a) b) c) d)
7	Lyd	f.eks. DVD-Audio eller komprimerede former som MP3 (MPEG-1 Audio Layer-3, RA (Real Audio), WAVE (Microsoft)	a) b)
8	Video	f.eks. MPEG-1	a) b)
9	Geografiske data	Data til GIS-applikationer (vektor grafik og alfanumeriske data)	a)
10	Interaktive multimedie publikationer	f.eks. komplekse cdrom-publikationer, der sammenknytter mindst 3 datatyper	a) d)
11	Eksekverbare programmer	f.eks. perl scripts, plug-ins eller cgi-programmer der indgår i en web publikation	a) b) c) d)

Det vil være en del af arbejdet med handlingsplanen at færdiggøre denne tabel, således at der skabes et fuldstændigt overblik over hvilke formater, der skal omfattes af KBs handlingsplan og i hvilke materialegrupper de optræder.

3.4

Hvilke datamedier er omfattet af handlingsplanen og hvilke krav stilles mht. de fysiske mediers langtidsoopbevaring

De fysiske datamedier, som skal bevares, er af to typer – magnetiske og optiske. Det bærende materiale er enten af metal (hard disks), polykarbonater (CD, DVD mm) eller polyester (båndsystem og disketter). På eller i disse bærere findes et lag, som holder informationen – enten med magnetiske eller optiske egenskaber.

Bevaringsproblematikken er, præcis som med traditionelle kulturgenstande, hovedsagligt koblet til påvirkning af varme, fugt, lys og luftforureninger samt materialernes kemiske sammensætning. Alle disse faktorer kan i mere eller mindre grad påvirke mediet, hvilket derefter kan lede til at informationen bliver ulæselig gennem de strukturelle forandringer, som sker i basematerialet eller informationslaget. Disse mediers langtidsholdbarhed i et konventionelt kontormiljø beregnes som bedst til 5 – 10 år.

Et eksempel. I et af de optiske medier – CD-R-plader (compact disc – recordable) bruges farvestoffer for at skabe et optisk reflekterende lag. Disse farvestoffer er nært beslægtede med dem som bruges i farvefotografi og har også samme dårlige ældningsegenskaber. Den praktiske forskel ligger i, at et farvebillede bleges langsomt og til at begynde med er det kun farveinformationen som forandres, ikke motivet. Små forandringer i farvestoffet på en CD-R-plade gør den hurtigt ulæselig, hvorved motivet bliver direkte utilgængeligt.

Både magnetiske og optiske datamedier skal bevares mørkt. Den relative luftfugtighed skal ligge mellem 30% og 40 % (+/- 5%/24h) og temperaturen være under 18°C (+/-2°C/24h).

Da det digitale medium nedbrydes hurtigere end konventionelle medier som papir eller s/h film, betyder dette, at der skal investeres ressourcer i løbende at konvertere de medier, hvorpå det digitale materiale opbevares. I forbindelse med en medie-konvertering på KB bør der tages hensyn til flg. forhold:

- Digitalt materiale bør findes parallelt på flere digitale medier
- Mediekonvertering bør finde sted i forskudte tidsintervaller, således at det digitale materiale ikke pludselig kun findes på en række helt nye medietyper.

Den hurtige udvikling indenfor computerteknologien udgør også et problem for langtidsbevaring. Derfor er der udviklet systemer med henblik på langtidsbevaring af typen SAN – storage area network. Denne systemtype lagrer informationen samtidig med at den findes i hurtig tilgængelig form. Et sådant system har selvkontrollerende funktioner og er forberedt til migrering til kommende teknikgenerationer. Dette borger for den bedst mulige langtidsbevaring af digital information.

3.5

Hvilken metode er teknisk realistisk og økonomisk rentabel at anvende til langtidsopbevaring af hvilke materialegrupper og datatyper

Målet er at kunne bevare og tilgængeliggøre digitalt materiale over tid, med henblik på såvel nutiden som kommende generationer. Et bud på bevaringsperspektiver er:

kortsigt:	op til 5 år
mellemsigt:	5 – 10 år
langsigte:	> 10 år

Ved det langsigtede perspektiv skal KB sikre, at materialet til enhver tid er opdateret til en læsbar form og klargjort, således at fremtidige generationer kan få adgang til det ved hjælp af de værktøjer, som må være til rådighed til den tid.

Bevaring frembyder to hovedproblemer: *Nedbrydning og teknologiforældelse*, hvoraf det første allerede er behandlet i det foregående afsnit.

Både maskiner og programmer har udviklet og ændret sig hastigt i de sidste 20 år, og uden at der er fastsat endegyldige standarder. Dette kan medføre, at selv om et digitalt værk er bevaret, er det ikke nødvendigvis muligt at formidle det, således som det var tænkt. Det er derfor nødvendigt at se på, hvilke modeller, der findes til at

afhjælpe denne teknologiforældelse for sikre en tilgængelighed for det digitale materiale, man har valgt at bevare.

Der findes overordnet 3 mulige løsningsmodeller:

a) Teknologibevaring

Løsningsmodellen bygger på, at man gemmer et nødvendigt antal maskiner og programmer (såvel operativsystemer som applikationsprogrammer) med det formål at kunne vise det digitale objekt i et miljø, der er kompatibelt med det, objektet er udviklet til. Denne løsningsmodel kan kun anvendes, når der er tale om et kortsigtet bevaringsperspektiv, idet løsningen besværliggøres af en række praktiske problemer med dels at holde maskinerne kørende (at skaffe reservedele) dels de pladsproblemer en sådan tilgang vil afstedkomme for en institution, der vil give adgang til sit arkiv.

b) Teknologiemulering

Denne løsningsmodel bygger på at det digitale materiale forbliver i deres oprindelige form, og at man i stedet bygger programmer (emulatorer), der efterligner de gamle edb-miljøer (maskiner og operativsystemer) eller i en mindre ambitiøs model blot efterligner, hvordan det program, der direkte interagerede med det digitale materiale, virkede.

Den løbende udvikling består derefter i at udvikle nye emulatorer, der netop efterligner det nyeste edb-miljø på det man netop har forladt. Skal man således give adgang til meget gamle data, gør man dette ved at tage de oprindelige data og applikationsprogrammer og afvikle dem oven på den stak af emulatorer, der skal til for at de tilsammen emulerer det oprindelige edb-miljø.

Det gode ved en sådan tilgang er, at kan man emulere et bestemt edb-miljø, kan man afvikle alle de programmer på denne emulator, som var udviklet til dette edb-miljø og man har derved løst tilgængelighedsproblemet for alt det digitale materiale, som virkede i dette miljø, og ikke kun for en enkelt datatype. Adskillige firmaer har gennem de sidste 20 år løbende udviklet emulatorer til ældre edb-miljøer, men de praktiske erfaringer med emulering indenfor bibliotekssektoren er meget sparsomme, jvf.

<http://www.rlg.org/preserv/diginews/diginews3-6.html>.

Det Kongelige Bibliotek i Holland har gennem de seneste år fokuseret på emulering som en farbar løsningsmodel, senest gennem NEDLIB projektet. Af deres website ses desuden, at IBM er gået med ind i den aktuelle problemstilling, og allerede har indgivet patentansøgninger.

http://www.kb.nl/kb/resources/frameset_kb.html?/kb/ict/dea/ibm/ibm-summary-en.html

Hvor det ved den teknologibevarende løsningsmodel er de praktiske problemer, der nedbringer dens anvendelighed, er det de teknologiske problemer, der er i fokus, når der er tale om teknologiemulering. Det er vanskeligt at udvikle gode emulatorer, ligesom der heller ikke er sikkerhed for, at nogen hverken kan eller vil gøre det ud i fremtiden.

Det må derfor anses for risikabelt udelukkende at basere sig på denne løsningsmodel.

c) Migration af digital information

Denne løsningsmodel bygger på at det i det mindste skal være muligt at få adgang til den digitale information, hvorfor man i en løbende proces migrerer data fremad for at sikre, at de er tilgængelige i nye edb-miljøer.

Migrationen kan gennemføres på flere måder:

1. Ved brug af brug af programmer, der er bagud kompatibelt, således at det digitale materiale i det forældede format kan hentes ind i programmerne og gemmes i det nye, fremadrettede format. Dette er vel den mest anvendte migrationsform og anvendes hvert år på KB, når vi går fra én forældet version af f.eks. Word til den nyeste version.
2. Ved brug af konkurrerende programmer, der benytter sig af fælles udvekslingsformater. Det var f.eks. således vi gik fra WordPerfect til Word til vores administrative data på KB.
3. Ved at konvertere det digitale materiale til et standardformat, som er velegnet i forhold til de digitale samlinger, som skal bevares. På KB anvender vi f.eks. p.t. XML og TIFF.
4. Ved at foretage et medieskift. Man forsøger udelukkende at bevare informationen i det digitale materiale, og accepterer at form og funktionalitet går tabt. Dette kan f.eks. ske ved at udprinte data på papir og bevare dette som konventionelle medier, og dermed berøve informationerne bl.a. deres søgbare

kvalitet. Denne migrations-strategi må anses for at være en sidste udvej.

For samtlige migrationsmetoder vil gælde, at mens informationerne for det meste vil være bevaret efter migrationen, kan andre dele være gået tabt, f.eks. layout, metadata eller farverum.

Da løsningsmodellen forudsætter, at man migrerer det digitale materiale format for format, vil den være arbejdskrævende at anvende for den enkelte institution. Til gengæld er det en metode vi allerede kender, og som vi både kan anvende nu og i fremtiden. På KB er den således senest anvendt ved overgangen fra bibliotekssystemet RCLIB til ALEPH, hvor grunddata er tilstede i struktureret form med danMARC2.

Nedenfor findes en oversigt over de mest hensigtsmæssige bevaringsstrategier set i forhold til KBs forskellige digitale datatyper:

Digitalt materiale omfattet af KB's handlingsplan	Mest hensigtsmæssige bevaringsstrategi	Del af strategi	Bemærkninger i forhold til arbejdet på KB
1. Data sæt	Migrere digitale information	Konvertering til standard format	Valg af standard format skal foretages
2. Struktureret tekst	Migrere digital information	Konvertering til standard format	Valg af standard format skal foretages
5. Præsentationsgrafik	Migrere digital information	Brug af bagud kompatible programmer kombineret med konvertering til standard formater	Valg af standard format skal foretages
6. Billeder	Migrere digital information	Brug af bagud kompatible programmer kombineret med konvertering til standard formater og medieskift	Valg af standard format skal foretages. Egen digitaliseringsindsats skal ske, således at der digitaliseres til lagringsformatet.
9. Geografiske data	Migrere digital	Brug af bagud	Valg af standard

	information	kompatible programmer kombineret med konvertering til standard formater og brug af konkurrerende programmer	format skal foretages. Her findes ingen færdige løsninger p.t.
10.Multimediepræsentationer	Teknologi bevaring kombineret med emulering		Her findes ingen færdige løsninger p.t.

Kilde : National Preservation Office. *Digital culture: maximising the nation's investment*. Ed. Mary Feeney. 1999
<http://www.ukoln.ac.uk/services/elib/papers/other/jisc-npo-dig/intro.html>

Handlingsplanen skal tage stilling til, hvorledes KB skal håndtere problemerne omkring teknologiforældelsen, således at man opnår en teknisk realistisk og økonomisk rentabel plan for det materiale, man ønsker at bevare. Der vil blive tale om en pragmatisk og selektiv tilgang til bevaring af den digitale information med et ønske om at lade såvel emulering som migration af information indgå i strategien.

3.6

Sikring, kontrol og organisering. Organisering og administration af langtidsopbevaring af digitale data

Etablering af et nyt system til langtidsopbevaring af digitale data, betyder også etablering af en række nye funktioner, hvortil der skal etableres arbejdsgange og rutiner:

- at håndtere administration og digital bevaring af det digitale materiale. Hvordan skabes der sikring for, at det materiale (typer og medier) KB modtager konverteres/migreres til de formater og medier KB ønsker at langtidslagre i og på.
- at lagre alle typer digitalt materiale i KBs samlinger

- at danne og vedligeholde de metadata, der knytter sig til det digitale materiale
- at give adgang (identificere, søge, fremtage og benytte) for brugere i og uden for læsesale
- at kontrollere den valgte strategi for langtidslagring.
Hvordan sikres og kontrolleres det, at den information man valgte at bevare, faktisk er bevaret intakt og kan læses logisk?

Disse funktioner vil berøre de fleste afdelinger på tværs af biblioteket. Dette betyder bl.a. at alle berørte medarbejdergrupper skal identificeres med henblik på inddragelse, ligesom der efterfølgende skal etableres uddannelse for de medarbejdere, der skal i berøring med systemet.

Ved indpasning af det nye system i den eksisterende IT-infrastruktur, skal denne gennemgås med henblik på sikkerhedsaspektet, både med hensyn til sikring af adgangsforhold samt datasikkerhed, således at de lagrede data ikke ændres utilsigtet.

3.7

Tilvækst i dataomfang

Det er forbundet med meget stor usikkerhed at skulle forsøge at estimere væksten i volumen i det digitale materiale på KB over den kommende årrække, også selv om der kun er tale om en 5-årig periode.

Nedenstående skal derfor betragtes som et kvalificeret bud på denne udvikling inden for hver af de materialegrupper, der må anses for relevante. For hver materialegruppe er det i skemaform forsøgt opgjort, hvor stort volumen er i dag, og hvor stort det skønnes at blive i hhv. 2001 og 2005. Efter skemaet følger yderlige oplysninger vedrørende forud-sætningerne for skemaets tal. Som det vil fremgå, er der ikke i dette afsnit taget hensyn til eventuelle konsekvenser af en selektionspolitik.

Det fremgår af skemaet, at det først og fremmest er pligtafleveret digitalt materiale, der fylder nu og om 5 år.

Det må være en del af handlingsplanen, at der etableres et mere sikkert estimat over omfanget af det materiale, der skal sikres for eftertiden.

Volumen KB's digitale materiale i Gbyte

Materialegruppe	Total ved udgangen af 2000	Tilvækst 1999	Tilvækst 2000	Forventet årlig tilvækst i 2001	Forventet total i 2001	Forventet total 2005
1. Pligtafleverede cdrom'er	6119	0	0	2075	8184	16484
2. Pligtafleverede netpublikationer	19	5,6	10,4	15	34	94
3. KB-initieret digitalisering	65	13,1	21,2	30	95	457
4. Digitalisering initieret udefra	245	0	0	120	200	1840
5. CDROM over netværk	71,2	0	0	0	70	50
6. E-tidsskrifter i abon.	0	0	0	0	0	0
7. Web-harvesting	0	0	0	0	0	9830

Ad 1) Pligtafleverede musik cd og cdrom produktioner

Danske Afdeling: Har ca. 300 fra før 2000, tilvækst i 2000 : 2400 cdrom, forventet årlig tilvækst på 2000.

Musikafdelingen: Har ca. 10.000, tilvækst 2000 om året i de næste 5 år

Kort- og Billedafdelingen: Har 128, forventet tilvækst på 2-300 pr. år. Beregningerne er gennemført med en tilvækst på 250 om året. CD'erne vil have data af forskelligt omfang fra producenternes side. En cdrom kan rumme op til 650 Mbyte. Vi har for lethedens skyld gennemført beregningerne med et gennemsnit på 500 Mbyte pr. cdrom

Totalt i 2000 haves således ca. 12530 * 500 Mbyte = 6.0 TByte

Der forventes en årlig tilvækst på 4250 *500 Mbyte = 2 Tbyte

Ad 2) *Pligtafleverede netpublikationer.*

Det skønnes, at tilvæksten er 15 Gbyte pr. år for 2001 og hvert af de kommende år.

Ad 3) *Digitalisering foranlediget af KB selv.*

Produktionen fra 1996-1998 har været ca. 10 Gbyte pr. år og stigende til 21 Gbyte i 2000.

Der forventes en kraftig vækst i denne digitalisering i 2001 og de følgende år. Vi vil forvente en stigning med 50% om året hvert af de kommende 5 år, 2001 – 2005. Omfanget er naturligvis helt afhængig af finansieringsmulighederne.

Ad 4) *Digitalisering foranlediget af henvendelser udefra*

Den samlede produktion over de sidste 3 år skønnes til 245 Gbyte. Der produceres et relativt stort volumen til dette formål, idet der digitaliseres med meget høj opløsning, hvorfor hver enkelt optagelse fylder tilsvarende. Der forventes en stigning i den digitale produktion med 50% om året hvert af de kommende år.

Ad 5) *Erhvervede Cdrom over netværk.*

Der forventes en nedgang over den kommende 5-årsperiode som følge af en omlægning til abonnementsordninger.

Ad 6) *e-tidsskrifter i abonnement.*

Der forventes en kraftig vækst i dette område, men disse data er p.t.

holdt uden for KBs handlingsplan, jfr. afsnit 3.2, hvorfor de ikke indgår med en mængde i denne oversigt.

Ad 7) *Web-harvesting*

Forventes at kunne gennemføres fra 2002.

Vi vil umiddelbart antage, at det danske net i størrelse ligner det svenske, hvis vi som de kun indsamler og arkiverer fra den åbne del af det danske net og kun indsamler web-sider.

Ved sidste arkiveringsrunde i Sverige indsamledes 30.8 mill digitale objekter på i alt 1.2 TByte.

Det antages, at 80% af alle siderne ændres i løbet af et år, dvs. at arkivering 2 gange årligt i volumen kun er lidt større, end hvis der arkiveres akkumulerende. Hvis vi antager, at nettets størrelse

forbliver uændret vil dette svare til en årlig tilvækst på 2.4 TByte. Vi må imidlertid forvente, at nettets størrelse er voksende, således at disse tal må anses for et absolut minimum.

3.8

Autenticitet og autentifikation

Ved autenticitet forstås, at værket er det, det angives at være. Ved værker, som er masseproducerede, betyder autenticitet, at værket er identisk med den master, hvoraf det er en kopi.

For digitale værker kræves, at et værk for at være autentisk både rummer det samme sæt af informationer og tilbyder det samme sæt af funktioner, som masteren, jvf. også afsnit 3.1.

Ved autentifikation forstås, at det kan godtgøres, at den aktuelle digitale kopi af værket er identisk med masteren.

Autentifikationen kan gennemføres ved, at der ved kopiering fra masteren tildeles kopien en identifikationskode, som med stor sikkerhed kan bevares i kopien og kan fremkaldes til dokumentation for værkets autenticitet.

I forbindelse med pligtaflevering ydes en særskilt indsats for at sikre, at den kopi af værket, der opbevares på Det Kongelige Bibliotek, er identisk med det værk, den er en kopi af.

Tildelingen af autentifikationskoder er en af de opgaver, som skal analyseres nærmere i forbindelse med det kommende arbejde.

I denne sammenhæng kan det oplyses, at Det Kongelige Bibliotek som led i et projekt i Kulturnet Danmark har udviklet et digitalt vandmærke til brug i forbindelse med billeder i farver, hvor det digitale vandmærke kan tilføres on the fly i forbindelse med brugernes fremtagning af de pågældende værker. Et sådant vandmærke tjener også den funktion, at biblioteket kan dokumentere oprindelsen af et billede.

Tildelingen af autentifikationskoder kan også tænkes som en af de serviceydelse, som Nationalbiblioteket udfører i forbindelse med

oparbejdelsen af digitale arkiver, beregnet på langtidsbevaring, dels som en del af standardaktiviteterne i forbindelse med lagringen af de enkelte værker, og dels som en service overfor brugere, der måtte ønske at sikre deres værkers autenticitet.

3.9

Metadata for bevaring af digitalt materiale (digitale objekter)

Metadata er data om data. Der findes en række standarder for metadata til hvert sit specialiserede formål. Dette afsnit omhandler metadata med henblik på bevaring af digitale objekter.

Et digitalt objekt er en samling af bits. Et af dets karakteristika er, at det ikke i sig selv er forståeligt for omverdenen, men kun kan afkodes og forstås i en bestemt sammenhæng ved brug af hardware og software. Digital information kræver således i særlig grad detaljerede metadata for at sikre, at både vi og fremtidige generationer kan få adgang til informationen.

Foruden det digitale objekt og de grundlæggende oplysninger om dets indhold og tilblivelse er det således f. eks. også nødvendigt at registrere information om, hvordan objektet er kodet til brug for computer-behandling, hvordan bevaringstiltag er gennemført (migrering, emulering..), hvordan ophavsrettigheder skal håndteres samt gemme information til at sikre autenticiteten over tid. I forbindelse med digitale ressourcer har der tidligere været mest fokus på de beskrivende metadata (de førstnævnte) gennem f.eks. Dublin Core initiativet, men i forbindelse med opbevaring og langtidslagring af disse ressourcer vil det være de sidstnævnte, dvs de oplysninger, der omhandler forvaltningen af samlingen samt understøtter beslutninger og handlinger vedrørende bevaring, der bliver de centrale.

Open Archive Information System Reference Model (OAIS, jfr. afsnit 4) er en konceptuel disposition eller struktur for et digitalt arkiv. Dens metadata model er en overordnet beskrivelse af de typer af information, som er nødvendige for at arkivets forskellige funktioner kan udføres.

OAIS opererer med følgende overordnede disposition for metadata:

'Archive Information Package' består af:

'Content information': (den information arkivet er betroet at bevare ud over selve objektet, f.eks. krav til installation (bl.a. HW og SW), dataformat og forskellige repræsentationer)

'Preservation description information': (den information der er nødvendig for at 'content information' kan bevares. Opdeles i:

'Reference information': Objektets forskellige identifikatorer (f.eks. ID inden for arkivet eller URL), med henblik på eentydig identifikation samt dato for opståen

'Context information': De informationer der vedrører forholdet til omgivelserne – hvordan er objektet dannet, og hvordan det står i forhold til andre digitale objekter

'Provenance information': Information vedrørende det digitale objekts oprindelse og alle forandringer der senere har fundet sted

'Fixity information': Autentifikationsmekanismer og dvs nøgler til at sikre at det digitale objekt er intakt, f.eks. checksum og digital signatur)

'Packaging information': (Det der binder informationerne sammen til en identificerbar enhed på et specifikt medie, f.eks. en iso 9600 filstruktur på en CDROM)

'Descriptive Information': (De informationer der udbyder adgang til det digitale objekt gennem arkivets søge- og adgangsveje.

Denne information vil ofte være dannet ud fra 'content information' og 'preservation description Information')

Flere institutioner har allerede anvendt OAIS modellen i forbindelse med deres arbejde med metadata til bevaring og publiceret deres arbejde på dette område, således det engelske CEDAR projekt, det australske nationalbibliotek og en række europæiske nationalbiblioteker gennem EU-projektet NEDLIB.

- Kelly Russell, Derek Sergeant, Andy Stone, Ellis Weinberger and Michael Day, Metadata for Digital

Preservation: the Cedars Outline Specification. Leeds: Cedars project, March 2000.
<http://www.leeds.ac.uk/cedars/OutlineSpec.htm> . Denne specifikation danner grundlag for implementationen hos British Library.

- National Library of Australia Preservation Metadata Working Group, Preservation Metadata for Digital Collections: Exposure Draft. Canberra: National Library of Australia, 15 October 1999. <http://www.nla.gov.au/preserve/pmeta.html>
- Catherine Lupovici and Julien Masanès, Metadata for Long Term Preservation. NEDLIB Report series, 2. The Hague: Koninklijke Bibliotheek, July 2000.
<http://www.kb.nl/coop/nedlib/results/D4.2/D4.2.htm>

Primo 2001 har OCLC/RLG publiceret et 'white-paper',

http://www.oclc.org/digitalpreservation/presmeta_wp.pdf

der sammenligner disse tre initiativer samt et fjerde fra Harvard Universitet, der endnu ikke er publiceret, med det formål at

- udvikle en omfattende disposition for metadata
- identificere de vigtigste metadata elementer med henblik på digital bevaring

- identificere og udvikle alternative implementationer
- anwise 'best practice' på dette område

Resultatet af dette arbejde vil danne et godt udgangspunkt for KBs eget arbejde på dette område.

I december 2000 blev den første internationale workshop afholdt om metadata for bevaring af digitalt materiale. Der findes kortfattede referater af workshoppen ved Robin Dale, RLG og Neil Beagrie, JISC <http://www.rlg.org/preserv/diginews/diginews4-6.html> og Michael Day <http://www.ariadne.ac.uk/issue26/metadata> foruden de enkelte indlæg fra workshoppen <http://www.rlg.org/events/pres-2000/infopapers.html> .

KBs handlingsplan skal på metadataområdet

- Fastlægge det nødvendige minimum af metadata for KB med udgangspunkt i de udenlandske erfaringer
- Anvise, hvorledes der kan tilvejebringes metoder til at generere så mange af disse metadata som muligt på automatisk vis, således at opgaven kan holdes inden for det praktisk og økonomisk realistiske, det meget store antal digitale objekter, som der skal arkiveres, in mente.

3.10

Rettighedsoplysninger og adgangsstyring

Ophavsretslige spørgsmål er så dominerende i forbindelse med information i digital form, at der må tages hensyn hertil i en politik for langtidsoptbevaring. Det er derfor nødvendigt, at der etableres systemer for adgangsstyring og adgangskontrol, således at de indsamlede værker kan benyttes samtidig med at de ophavsretlige hensyn tilgodeses. Hvorledes dette skal løses konkret må fremgå af den endelige handlingsplan, men fra andre projekter af tilsvarende karakter vides at forhandlinger om og afklaring af ophavsretslige spørgsmål kan være meget tidskrævende.

I gældende lov om pligtaflevering er det ophavsretlige spørgsmål løst ved at det i lovens §8, stk. 2 hedder: "Eksemplarer af værker i

digital form må dog alene stilles til rådighed for enkeltpersoner til personligt gennemsyn eller studium på stedet ved hjælp af teknisk udstyr. Eksemplarfremstilling i digital form er ikke tilladt". I praksis er det løst ved, at der er adgang til de pligtafleverede værker på en stand-alone-maskine på Det Kongelige Bibliotek og en stand-alone-maskine på Statsbiblioteket, hvor det ikke er muligt at foretage digital kopiering.

3.11

Identifikation i forbindelse med det arkiverede materiale

Alle arkiverede objekter skal kunne identificeres entydigt. Til det formål tænkes URN (Uniform Resource Name) anvendt i lighed med den praksis, der er ved at tage form på andre national-biblioteker

(bl.a. Helsinki University Library og Nationalbiblioteket i Norge), således at KB til arkivformål udtager sit eget URN namespace og selv forestår tildelingen af identifikationer på objektniveau, dvs. hvert billede, lydfil, eller dokument, der lægges i arkivet, får sin egen identifikation i form af en URN.

Det forventes, at Nationalbibliografisk Nævn fastlægger, at URN inden for DK namespace knytter sig til en service (i modsætning til et digitalt objekt), idet dette er den eneste operationelle i forhold til nationalbibliografien.

Det er derfor nødvendigt, at KB udtager sit eget URN namespace med henblik på arkiveringsopgaven, således at det bliver muligt at identificere det enkelte digitale objekt entydigt.

Ved etableringen af et digitalt arkiv skal KBs eget URN namespace fastlægges og den tilhørende tildelings- og resolutionsservice skal specificeres og implementeres.

3.12

Tilpasning af KBs egen produktion af digitalt materiale og lagringsstrategi

Handlingsplanen vil omfatte fastlæggelse af de standardformater, som KB skal følge, jfr. afsnit 3.3: Hvilke datatyper og formater er omfattet af handlingsplanen. Ved KBs egen produktion af digitalt materiale lettes migrationen ved tilpasning af produktionen til de valgte standardformater. Det vil derfor være en helt naturlig del af arbejdet med udformningen af handlingsplanen at vurdere, hvorvidt de allerede anvendte formater skal ændres.

Del II

Afsnit 4

Praksis på dette område i andre lande og i andre organisationer

og

Afsnit 5

Indstilling om realisering af handlingsplan

se vedlagte bidrag fra Fischer & Lorenz.

Afsnit 4 og 5:

Fischer & Lorenz's bidrag til KB's handlingsplan for etablering af en strategi og plan for langtidslagring af elektronisk materiale.
Januar 2001

Indholdsfortegnelse

4.....	Praksis i andre lande og i andre organisationer	29
4.1.....	Status fra Andre biblioteker, arkiver og virksomheder	31
4.1.1	National Library - Australien.....	31
4.1.2	Library of Congress - USA.....	32
4.1.3	Kongelige Bibliotek - Holland.....	33
4.1.4	Kongelige Bibliotek - Sverige.....	34
4.1.5	Finland.....	35
4.1.6	National Biblioteket - Norge.....	35
4.1.7	Island.....	36
4.1.8	British Library - England.....	36
4.1.9	Statens Arkiver - Danmark.....	37
4.1.10	Patent og Varemærke direktoratet - Danmark.....	37
4.2.....	Status fra samarbejdsprojekter, interessegrupper og lign.	38
4.2.1	NEDLIB - Networked European Digital Libraries.....	38
4.2.2	CEDAR.....	39
4.2.3	Research Library Group.....	39
4.3.....	Open Archive Information System (OAIS)	40
5.....	Indstilling om realisering af handlingsplan	43
5.1.....	Tidsplan og budget for projekterings-projektet	44
5.2.....	Projektforberedelse	46
5.3.....	Analysearbejde	47
5.4.....	Udarbejde handlingsplan	47

4. Praksis i andre lande og i andre organisationer

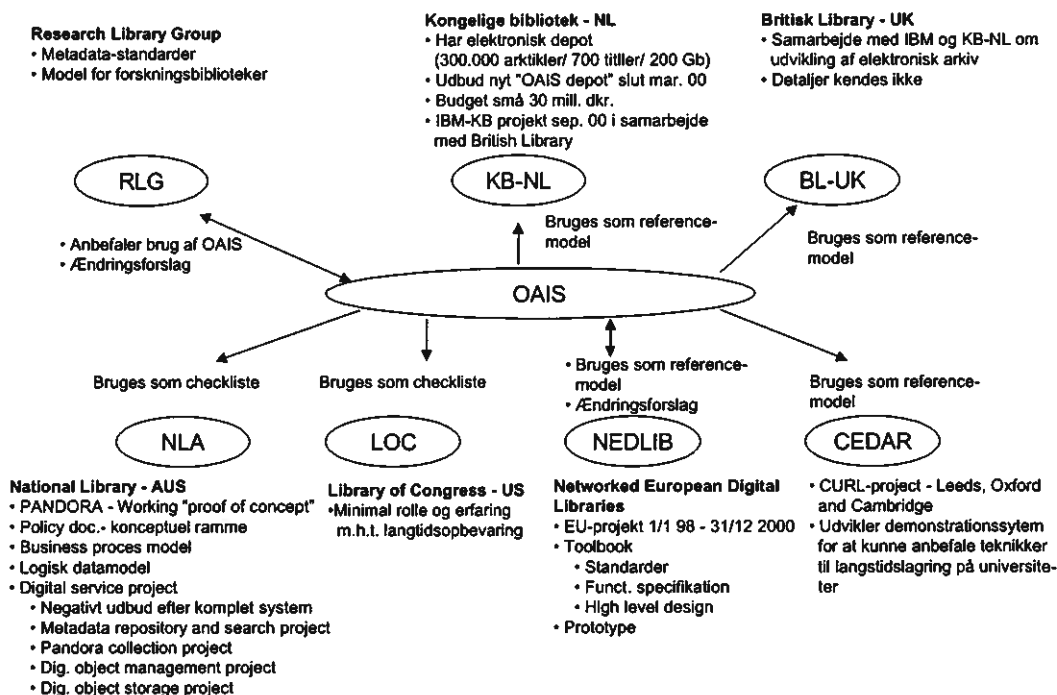
I forbindelse med det indledende arbejde har projektgruppen forsøgt at gøre rede for status hos en række andre biblioteker, institutioner og virksomheder med hensyn til langtidsopbevaring af elektroniske publikationer. Dette er i sær foregået ved at søge oplysninger på deres web-sider og gennem samtaler med relevante personer, f.eks. i forbindelse med seminarer og kongresser eller over telefon.

Det er vidt forskelligt, hvor langt national biblioteker og andre institutioner med arkiv forpligtelser er kommet med hensyn til at analysere, hvilke typer information man skal indsamle, opbevare og gøre tilgængelige for eftertiden, og hvad der kan lade sig gøre. Situationen er kort fortalt som følger:

- Det er især den hurtige teknologiske udviklingen og manglende brug af standardformater for tekstdokumenter, billeder etc., der gør det svært at sikre, at f.eks. elektroniske publikationer kan gøres tilgængelige, vises og forstås om f.eks. 100 år.
- Nogle få biblioteker og institutioner har været i gang i mere end 5 år med at analysere problemstillingen omkring langtidsopbevaring og finde fremgangsmåder, som passer til dem. Nogle er ved at gå i gang, og andre er slet ikke begyndt at overveje dette seriøst.
- Der er ingen, der har fundet det endelige svar på, hvordan man opbevarer informationer i lang tid og sikrer, at de kan gøres tilgængelige og forstås om f.eks. 100 år. Der tales i dag om tre strategier, som anvendes eller påtænkes anvendt i forskellige kombinationer:
 1. *Det tekniske museum*, her skal et antal af alle de typer udstyr og programmel, der gennem tiderne bliver brugt til at vise elektroniske publikationer, bevares og vedligeholdes, så det til sin tid kan anvendes til at vise publikationerne.
Der er ingen, der betragter denne strategi som en løsning på lang sigt. Det vil være en uoverkommelig opgave at vedligeholde og bevare udstyret funktionsdygtigt i årtier og århundreder. Der vil ligeledes være risiko for flaskehalse, da kun et begrænset omfang af udstyr vil blive bevaret.
 2. *Migrering eller konvertering*, her skal elektroniske publikationer med passende intervaller konverteres til formater, der kan vises af den gængse teknologi på konverteringstidspunkterne.
Der er risiko for, at der går informationer tabt i forbindelse med konvertering. Det kan f.eks. resultere i, at publikationerne får et andet udseende, at der mangler funktioner i en multimedie publikation, eller at der mangler bibliografiske informationer.
 3. *Emulering*, som går ud på at bevare den elektroniske publikation i sin oprindelige for og simulere det teknisk miljø, der oprindeligt blev anvendt til at vise den elektroniske publikation, når den skal vises for eftertiden.
Der foretages forsøg med emulering forskellige steder, men der findes ikke nogen veletableret standarder eller systemer til emulering. IBM hævder at have en patentansøgning under behandling på en lovende metode til emulering.

- Der er ingen biblioteker eller andre organisationer i Danmark eller i andre lande, der har et færdigt systemkompleks med sammenhængende systemer, som håndterer alle aspekter i forbindelse med indsamling og langtidsopbevaring af elektroniske publikationer og andre digitale informationer samt adgangen til disse. Specielt er problemstillingen om langtidsopbevaringen ikke løst.
- Nogle biblioteker har for at opbygge viden og erfaring etableret testsystemer, eller første generation af systemer til at indsamle, kategorisere, arkivere og stille elektroniske publikationer til rådighed for forskellige brugere.
- Nationalbiblioteket i Australien, det kongelige bibliotek i Holland og British Library, virker som de biblioteker, der er længst fremme med hensyn til etablering af sammenhængende systemer til indsamling og langtidsopbevaring af elektroniske publikationer samt adgang til disse.
- KB, national biblioteker og andre institutioner med arkiv forpligtelser i andre lande, står tilsyneladende alene med kravene om at kunne modtage publikationer i vilkårlige formater efter udgiverens og forfatterens eget valg.
- Der er tilsyneladende ved at blive etableret en bred accept i biblioteksverdenen af en reference model for elektroniske arkiver til opbevaring af digitale informationer. Modellen betegnes som Open Archive Information Systems (OAIS) og er beskrevet nærmere i et senere afsnit.

Nedenstående figur viser, hvorledes de og nogle andre biblioteker, projekter og interessegrupper forholder sig til OAIS modellen, og deres aktiviteter mht. langtidsopbevaring.



Figur 0-1: Status og anvendelse af OAIS modellen

4.1 Status fra Andre biblioteker, arkiver og virksomheder

4.1.1 National Library - Australien

4.1.1.1 Generel Status

National Library i Australien (NLA) er et af de biblioteker, som er længst fremme med etablering af et elektronisk arkiv for elektroniske publikationer. National library kører et projekt kaldet PANDORA med det formål at etablere en national løsning til indsamling og bevaring af elektroniske publikationer i Australien. Man har valgt migrering og konvertering som strategi for langtidsopbevaring.

NLA har tidligere været i udbud efter et komplet system, som kunne håndtere alle aspekter omkring indsamling og lagring i et elektronisk arkiv. Man konkluderede, at ingen kunne tilbyde et sådan komplet system, stoppede udbudet og gik i gang med en trinvis implementering af et elektronisk arkiv. Der er nu etableret et kørende elektronisk arkiv, men endnu ikke en løsning til langtidsopbevaring.

NLA er p.t. ved at vurdere tilbud dels på et data management system og dels på et arkiv, der kan opfylde arkiveringsbehovene de næste 5-6 år. Det vides ikke, hvilke filformater, og metoder der skal anvendes til arkivering, adgang visning.

4.1.1.2 Forløb og planer

- 1996
- Etablering af PANDORA projektet
 - Etablering af PANDORA arkivet der indsamler og giver adgang til publikationer på Internettet. Arkivet er etableret som en arbejdende "proof of concept" model og er fortsat under udvikling. Indsamling af publikationer på nettet og langtidsopbevaring er to af de områder, man fokuserer på p.t.
 - Policy documents.- konceptuel ramme
 - Business proces model
 - Logisk datamodel
- 1998
- Etablering af **Digital services projektet**, som skal frembringe tekniske løsninger og installationer til integreret indsamling, lagring, administration af elektroniske publikationer. Der skal ligeledes etableres tekniske løsninger som tillader hurtig udvikling af elektroniske ydelser med øget adgang til digitale samlinger, f.eks. i samarbejde med andre institutioner.
- Dec. 1998
- Digital services - Information paper
- 1999
- Digital services - Metadata Repository and Search System
 - Request For Quotation (RFQ)
 - Køb af MetaStar Enterprise software fra Blue Angle Technologi.
- 1999
- Digital services - Digital Collection and Management system
 - Request For Tender (RFT) - på et komplet integreret system til indsamling, datamanagement og arkivering.

- RFT afbrudt i erkendelse af, at der ikke eksisterede et "hyldevare" system, som kunne opfylde alle kravene. Beslutning om følgende tre nye indkøbs og udviklingsprojekter:
 - Pandora Collection System Project*
 - Digital Object Management System Project*
 - Digital Object Store Project*
- 1999 - 2000
- Pandora Collection System Project
Det eksisterende system til indsamling af web-sider, der er baseret på softwareproduktet Harvest, skal erstattes af et mere effektivt system. National library vil selv udvikle det nye system og anvende nyere produkter til indsamling af web-steder.

Projektet skal være afsluttet december 2000.
- 1999 - 2000
- Digital Objekt Management System Project
Etablering af et datamanagement system til registrering af bibliografiske og tekniske informationer som beskriver de lagrede publikationer, samt til administration af de lagrede publikationer. Der har været udsendt RFQ og modtaget tilbud, som nu evalueres.
- 1999-2000
- Digital Object Management System Project
Indkøb og etablering af elektroniks arkiveringsudstyr som kan opfylde bibliotekets behov i perioden 2001 - 2006. der har været udsendt RFQ og modtaget tilbud, som evalueres.

4.1.1.3 Budget

Ikke offentliggjort

4.1.1.4 Referencer

- Denne Web-side giver oversigt over og referencer til National Library's aktiviteter, dokumenter, systemer og lignede, som er relateret til bevaring af digitale informationer:
<http://www.nla.gov.au/padi>
- Der findes en kort beskrivelse af Pandora projekts indhold og status på denne Web-side:
<http://pandora.nla.gov.au/pandora/documents.html>
- Der findes en længere beskrivelse af Pandora projektet på denne Web-side:
<http://pandora.nla.gov.au/pandora/bpm.html>
- Man kan få adgang til Pandora arkivet på denne Web-side:
<http://pandora.nla.gov.au/pandora>

4.1.2 Library of Congress - USA

4.1.2.1 Generel Status

Library of Congress spiller efter eget udsagn en minimal rolle og har meget begrænset erfaring m.h.t. langtidsofbevaring. De vil tilsyneladende ikke være blandt de første, som etablerer et komplet elektronisk arkiv for elektroniske

publikationer, hvor alle problemer om langtidsopbevaring er løst. De vil derimod gerne indgå i forskellige samarbejder og følge andres erfaringer.

Ovenstående er bl.a. udledt fra de to sidste afsnit på side 4-7 i LOC's digitale strategi ref:

<http://books.nap.edu/books/0309071445/html/79.html>

4.1.2.2 Forløb og planer

Kendes ikke

4.1.2.3 Projekter

Kendes ikke

4.1.2.4 Budget

Kendes ikke

4.1.2.5 Referencer

Library of Congress' hjemmeside:

<http://www.loc.gov>

4.1.3 Kongelige Bibliotek - Holland

4.1.3.1 Generel status

Det kongelige bibliotek i Holland (KB-NL) er et af de biblioteker, der er længst fremme med hensyn til analyse, planlægning og etablering af et depot system for elektroniske publikationer. Systemet kaldes DNEP (Depot van Nederlandse Elektronische Publicaties).

KB-NL har siden 1995 lavet analyser og opbygget praktiske erfaringer gennem diverse pilot projekter. De har bl.a. etableret et elektronisk depot med ca. 300.000 artikler fra 700 titler 200 Gb. Hertil anvender de IBM-Digital library programmet på en AIX platform. Depotet dækker bl.a. ikke behovene i forbindelse med langtidsopbevaring.

KB-NL har nu indgået en kontrakt med IBM om udvikling af et elektronisk depot, som følger OAIS modellen. Modellen har været anvendt som referencemodel i forbindelse med udbud og kontraktforhandlinger, men er blevet suppleret mht. langtidsopbevaring. British Library indgår også i dette samarbejde, men det vides ikke hvordan.

Der har ikke været adgang til aftalen med IBM, men det nye system skal efter sigende anvende IBM-Digital library, som KB-NL har i forvejen. Det vides heller ikke, hvor mange, eller hvordan funktionerne i OAIS modellen vil blive implementeret.

4.1.3.2 Forløb og planer

1995 - 1999

- Analyser
- Diverse projekter
- Etablering af det første elektroniske depot som beskrevet ovenstående.

- | | |
|------------------------|---|
| Primo 1999 | • Undersøgelse af løsningsmuligheder på markedet (Request for Information) med besvarelser fra 4 leverandører |
| Ult. 1999 - Medio 2000 | • Udbud af nyt system i henhold til OAIS modellen |
| Sep. 2000 | • Valg af IBM som leverandør |
| Jan. 2000 | • Projektstart |
| Jan. 2001 - Juli 2001 | • Detaljeret design afsluttet |
| Sep. 2001 | • Study preservation and webarchiving: Indhold ikke kendt |
| Dec. 2001 | • Release 1 afsluttet: Indhold ikke kendt |
| Maj. 2002 | • Release 2 afsluttet: Indhold ikke kendt |
| | • Release 3 afsluttet: Indhold ikke kendt |

4.1.3.3 Budget

4 mill. euro eller små 30 mill. kr.

Det nøjagtige budget kendes ikke, men på KB-NL's hjemmeside, er der anført at budget-rammen for DNEP udbudet var 4 mill. euro. IBM har måske kunne give et lavere tilbud, hvis man bygger på KB-NL's investering i IBM-Digital Library.

4.1.3.4 Referencer

Informationer om DNEP kan findes på denne Web-side:

http://www.kb.nl/kb/resources/frameset_funciemenu-en.html

4.1.4 Kongelige Bibliotek - Sverige

4.1.4.1 Generel Status

Kulturärw3 er tilsyneladende det eneste projekt, som det kongelige bibliotek i Sverige (KB-S) har kørende, der vedrører indsamling og opbevaring af elektroniske publikationer. Kulturärw3 er afgrænset til udvikling af et system til indsamling, opbevaring og adgang til svenske web-sider, KB-S vil i princippet indsamle alle svenske web-sider.

Systemet er stadig under udvikling. Der er f.eks. endnu ikke taget hul på problematikken omkring langtidsofopbevaring, men det er erkendt som et problematisk område.

4.1.4.2 Forløb og planer

Projektet startede i 1996, da der blev ansat en projekt ingeniør. Siden er der blevet anskaffet dataudstyr til testmiljø, søgerobot og senest et arkivsystem.

KB-S har nu prøvet at indsamle svenske web-sider i flere omgange fra .se, .com, .org, .net, .nu domæner. Seneste omgang gav ca. 15 millioner URL'er fra 58 500 web-steder. to tredjedele findes under .se. Den totale datamængde er ca. 280 Gbyte medtekst, billeder, lyd m.m..

Parallelt med indsamlingen arbejdes der med at gøre den lagrede information tilgængeligt. Man er efter sigende langt fremme med et program, som gør det muligt at surfe i arkivet. Det ser nu ikke ud til at virke over nettet.

Der er modelleret, men ikke implementeret en database over de indsamlede dokumenter. Det indgår også i planerne at etableret et søgeindeks til arkivet.

KB-S mangler også at sikre, at de lagrede dokumenter kan bevares og gøres tilgængelige på lang sigt. Hertil behøves der mere personale til udviklingsarbejde og kraftigere udstyr.

4.1.4.3 Budget

Kendes ikke.

4.1.4.4 Referencer

Yderligere informationer om Kulturärw3 kan findes på denne Web-side:

<http://kulturarw.kb.se>

4.1.5 Finland

4.1.5.1 Generel Status

Finland deltager i 2 projekter, der omhandler arkivering af elektronisk materiale: EVA- Et finsk projekt vedrørende indsamling og arkivering af online dokumenter (parallelt til det svenske Kulturärw3 projekt) og NEDLIB. Så vidt vides finder p.t. ingen drift eller udvikling af et system til langtidslagring af digitalt materiale sted.

4.1.5.2 Forløb og planer

Kendes ikke

4.1.5.3 Budget

Kendes ikke

4.1.5.4 Referencer

<http://www.lib.helsinki.fi/hyk/hul/proje/indexpe.html>

4.1.6 National Biblioteket - Norge

4.1.6.1. Generel Status

I Norge udvikler National Biblioteket - Mo I Rana selv et digitalt sikringsmagasin til håndtering af lagring, identifikation, metadata, migrering og ophavsret. Adgang for brugere vil først komme på et senere tidspunkt. De anvender ikke en reference model som f.eks. OAIS, men udvikler systemer og funktioner successivt.

På Nationalbibliotekets hjemmeside omtales 2 projekter vedrørende lagring af digitale dokumenter: Pligtaflleveringen af elektroniske dokumenter samt deltagelsen i EU-projektet NEDLIB. Projektet 'Digitalt sikringsmagasin', der står i centrum for denne sammenfatning, omtales ikke.

4.1.6.2. Forløb og planer

Kendes ikke

4.1.6.3 Budget

2000: 10 mill NKR. og 3 årsværk

2001: 10 mill NKR. og 3 årsværk

4.1.6.4 Referencer

<http://www.nb.no/html/litteratur.html>

<http://www.nb.no/rapporter/preserv.html>

Svein Arne Solbakk og Svein Arne Brygfjeld, National Biblioteket, Mo I Rana.

4.1.7 Island

4.1.7.1 Generel Status

Der er pr. 1.12. 2000 ansat en person ved det islandske nationalbibliotek, der vil have langtidslagring af digitale data som sit ansvarsområde.

4.1.7.2 Forløb og planer

Kendes ikke

4.1.7.3 Budget

Kendes ikke

4.1.7.4 Referencer

Thorsteinn Hallgrímsson

4.1.8 British Library - England

4.1.8.1 Generel Status

British Library har i september 2000 indgået en aftale med IBM om levering af et nyt 'Digital Library System' til langtidsoptbevaring af digitale dokumenter.

Systemet, der anvender OAIIS som referencemodel, omfatter samtlige modellens funktioner inkl. håndtering af migration/emulering af de digitale data. CEDARS Outline Specification vil danne basis for DLS metadata definition.

Der samarbejdes desuden med Det Kongelige Bibliotek i Holland, der ligeledes i september har indgået kontrakt med IBM om levering af et system til håndtering af lignende opgaver. De to projekter er ikke identiske.

4.1.8.2 Forløb og planer

Samarbejdet med IBM blev initieret d. 18. september 2000 og systemet skal være fuldt implementeret over en periode på 20 måneder; dvs. marts 2002.

4.1.8.3 Budget

Kendes ikke, men anføres i pressemeddelelsen til at være på flere millioner (multi-million) pund, dvs. mindst 30 mill. DKK, men formodentlig væsentligt højere.

4.1.8.4 Referencer

<http://www.bl.uk/diglib/dlp/dls.html>

<http://www.bl.uk/information/pr2000/85.html>

<http://www.rlg.org/events/pres-2000/shenton.html>

4.1.9 Statens Arkiver - Danmark

4.1.9.1 Generel Status

Statens arkiver har valgt en strategi for langtidsopbevaring af digitale informationer som er baseret på migrering, dvs. lagrede informationer med passende intervaller konverteres til tidssvarende formater.

Statens Arkiver kan i modsætning til KB stille krav til afleveringsformater og til brug af dokument og journal systemer i de forvaltninger som skal aflevere arkivalier, det vil sige at de kan nøjes med få udvalgte formater. Til dokumenter med grafisk indhold skal TIFF eller CCITT/TSS (gruppe 4 fax) anvendes og til ren tekst skal Tegnsæt DS/ISO 8859-1:1987 (Latin 1) anvendes.

4.1.9.2 Forløb og planer

Kendes ikke, men IT-chef Birgit Hansen, vil gerne være vært ved et møde i begyndelsen af 2001, hvor man gennemgår de forskellige aspekter og overvejelser man har gjort hos Statens arkiver i forbindelse med langtidsopbevaring.

4.1.9.3 Budget

Kendes ikke.

4.1.9.4 Referencer

Cirkulære om aflevering af elektroniske arkivsystemer:

<http://www.sa.dk/sa/statamtkom/arklov/nyaflicirk.htm#b4>

IT-chef Birgit Hansen

4.1.10 Patent og Varemærke direktoratet - Danmark

4.1.10.1 Generel Status

Patentbeskrivelser og ansøgninger skal opbevares i mange år og vil i fremtiden blive afleveret elektronisk. Patent og Varemærke Styrelsen (PVS) vil derfor ligesom KB have en problemstilling om langtidsopbevaring af elektronisk materiale. (PVS) modtager endnu ikke patent- og varemærke-ansøgninger elektronisk, bl.a. på grund af manglende elektronisk signatur. Ansøgninger og patentbeskrivelser modtages på papir, skannes og lagres på CD-ROM. Man har tilsyneladende ikke formuleret nogle strategier eller planer for langtidsopbevaring.

4.1.10.2 Forløb og planer

Man vil på et tidspunkt starte med at modtage varemærke-ansøgninger elektronisk og senere patentbeskrivelser og ansøgninger. Der vil højst sandsynligt blive stille krav om afleveringsformater.

4.1.10.3 Budget

Kendes ikke

4.1.10.4 Referencer

Dokumentationschef Anne Katrine Skibelund

4. 2 Status fra samarbejdsprojekter, interessegrupper og lign.

4.2.1 NEDLIB - Networked European Digital Libraries

4.2.1.1 Generel Status

EU projektet NEDLIB er et samarbejde mellem et antal europæiske nationalbiblioteker (se nedenstående liste) med det formål at få etableret en basis infrastruktur som kan binde europæiske elektroniske biblioteker sammen.

Der er ikke blevet etableret en infrastruktur, der binder europæiske elektroniske biblioteker sammen. En af årsagerne er nok, at de enkelte biblioteker stadig er ved at opbygge deres elektroniske biblioteker.

Projektet har i det væsentlige handlet om at definere nogle fælles referencer, rammer og standarder for digitale biblioteker.

Biblioteker som deltager i NEDLIB:

- Agentschap Rijksarchiefdienst (NSA)
- Biblioteca Nacional do Portugal (BNP)
- Biblioteca Nazionale Centrale di Firenze (BNCF)
- Bibliothèque Nationale de France (BnF)
- Die Deutsche Bibliothek (DDB)
- Helsinki University Library (HUL)
- Koninklijke Bibliotheek (KB)
- National Library of Norway - Mo I Rana (NBR)
- National Library of Switzerland (SNL)

4.2.1.2 Forløb og planer

Projektet startede den 1/1 - 98 og skal være færdig 31/12 2000. Projektet afsluttets til nytår, og der har netop medio december været afholdt en række møder i Haag, hvor projektets resultater er demonstreret.

De væsentligste kendte resultater er:

- Toolbook (værktøjskasse til brug for analyse, design og indkøb) med følgende dokumenter
 - Functional specifikation (specifikation af funktionelle krav)
 - High level design (overordnet design af et elektronisk arkiv)
 - Standarder (samling af henvisninger til standarder som er relevante i forbindelse med implementering af et elektronisk arkiv. Der er ikke foretaget endelige valg af standarder, som skal vælges under hensyn til eksisterende systemer, IT-arkitektur, etc.)
 - Metadata for long term preservation - definition af de minimumsdata det er nødvendigt at gemme om elektronisk materiale i forbindelse langtidsopbevaring og adgang til elektronisk materiale.
- Prototype (test af emulering i forbindelse med langtidsopbevaring af og adgang til digitale publicationer)

- Kommentarer og ændringsforslag til OAIS modellen. NEDLIB har tilsyneladende taget OAIS modellen til sig og indsendt kommentarer og ændringsforslag.

4.2.1.3 Referencer

NEDLIB's hjemmeside:

<http://www.konbib.nl/coop/nedlib/homeflash.html>

Resultater fra NEDLIB projektet:

<http://www.konbib.nl/coop/nedlib/homeflash.html>

Oversigt over status hos de deltagende biblioteker dateret 20. juli 2000:

http://www.konbib.nl/coop/nedlib/results/local_situations_v2.htm

4.2.2 CEDAR

4.2.2.1 Generel Status

CEDAR projektet gennemføres i samarbejde mellem forskningsbiblioteker i Storbritannien. Projektet har det formål at finde frem til fælles måder at indsamle og bevare elektroniske publikationer i Storbritannien og at give input til British Library arbejde inden for samme område.

Man har mest beskæftiget sig med standarder for bibliografiske data (f.eks. 'Metadata For Digital Preservation: The Cedars Outline Specification' jfr. British Library) samt bidraget med input og kommentarer til arbejdet med OAIS modellen.

Der er ikke fundet nogen testinstallationer eller konkrete anbefalinger med hensyn til problematikken omkring langtidsopbevaring og adgang til elektroniske publikationer.

4.2.2.2 Forløb og planer

Projektet startede 1998 og forventes afsluttet marts 2001.

4.2.2.3 Referencer

CEDAR projektets hjemmeside

<http://www.leeds.ac.uk/cedars/>

4.2.3 Research Library Group

4.2.3.1 Generel Status

RLG er et samarbejdsforum for forskningsbiblioteker i USA. Der er nedsat en arbejdsgruppe som skal arbejde med aspekterne omkring langtidsopbevaring af elektronisk materiale. Hensigten er at udvikle en model for elektronisk opbevaring, der passer til forskningsbibliotekernes formål.

Den største fokus har indtil nu været på at definere standarder for typen af bibliografiske informationer, der skal opbevares. Der er ikke fundet nogle konkrete anbefalinger eller installationer der håndterer problemerne omkring langtidsopbevaring og adgang til elektroniske publikationer.

RLG og OCLC (Online Computer Library Center) har i foråret 2000 indledt et samarbejde med henblik på i fællesskab at skabe infrastrukturer for digital arkivering. Arbejdet fokuserer i første række på at producere 2 dokumenter med henblik på at etablere 'best practices': Et vedrørende det som karakteriserer en stabil arkiveringsfunktion eller -service, og et vedrørende de metadata der skal til for at kunne håndtere og langtidsopbevare de digitale filer.

4.2.3.2 Forløb og planer

Dokumentet vedrørende metadata, som udarbejdes med OCLC, forventes at blive udgivet som white-paper i januar 2001.

4.2.3.3 Referencer

Hjemmeside for RLG's arbejdsgruppe der arbejder med langtidsopbevaring af elektroniske publikationer.

<http://www.rlg.org/longterm/>

Nyhedsbrev om RLG's og OCLC's samarbejde:

<http://www.oclc.org/oclc/press/20000310.htm>

Præsentation af planer og indholdet i RLG's og OCLC's samarbejde:

<http://www.rlg.org/events/pres-2000/lavoie/sld001.htm>

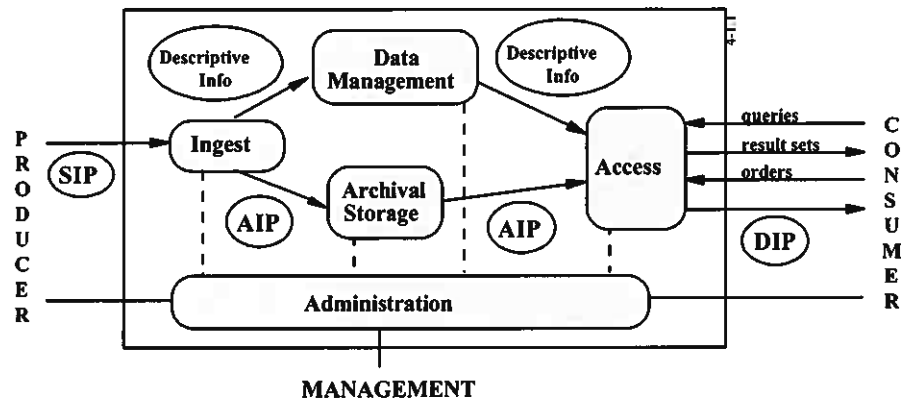
4.3 Open Archive Information System (OAIS)

OAIS er en referencemodel for et åbent elektronisk arkiv system. Den beskriver problemstillingen og nødvendige funktioner, når man skal indsamle, opbevare, og give adgang til digitale informationer i meget lang tid. OAIS modellen er udarbejdet som et udkast til en ISO standard under opsyn af CCSDS (Consultative Committee for Space Data Systems). CCSDS er en officiel organisation etableret af forskellige rumforsknings institutter.

Modellen kan udgøre et godt reference grundlag, når bibliotekerne skal beskrive deres behov og nødvendige funktioner, f.eks. når der skal opstilles strategier, indhentes og vurderes tilbud på systemer til indsamling, opbevaring og adgang til elektroniske publikationer og lignende.

Modellen er dog ikke fyldestgørende vedrørende bibliotekernes problematik med langtidsopbevaring af informationer, der kan afleveres i vilkårlige formater. Modellen omfatter ikke et komplet sæt af tekniske standarder for alle tekniske og administrative aspekter for et elektronisk arkiv, disse standarder er ikke valgt eller defineret endnu.

Nogle biblioteker, der anvender modellen, har indsendt forslag om udvidelse af OAIS modellen med et "preservation planning" modul. OAIS modellen med førnævnte udvidelse forventes ISO certificeret i 2001.



SIP - Submission Information Packet
AIP - Archival Information Packet
DIP - Dissemination Information Packet

Figur 0-1: OAIS overordnet funktions model

Ovenstående figur viser i firkanterne med afrundede hjørner de overordnede funktioner, der skal til for at indsamle, modtage, registrere, lagre og stille informationer til rådighed for brugere. Betegnelserne er engelske og beskrives kort i det følgende.

Ingest omfatter funktioner til at:

- modtage informations pakker (SIP) med elektroniske publikationer o.lign. med tilhørende informationer fra udgiverne (producer)
- udføre kvalitetskontrol, funktioner til at generere arkiv informations pakker (AIP) som overholder arkivets standarder
- udtrække beskrivende oplysninger om informationspakkerne og koordinere opdateringer af datamangement og archival storage funktionerne.

Archival Storage omfatter funktioner til at:

- lagre,
- vedligeholde,
- hente,
- kvalitetskontrollere og
- genopfriske - informationspakker på lagringsmedier.

Datamangement omfatter funktioner

- til at opbevare og gøre bibliografiske og tekniske informationer tilgængelige der beskriver, hvad der er gemt i arkivet,
- til den daglige administration og drift af arkivsystemet.

Access omfatter funktioner til at:

- hjælpe brugere med at finde ud af, hvilke informationer der er lagret i systemet,
- give brugere mulighed for at forespørge på og modtage informationer,
- begrænse brugernes adgang til de informationer, som de har lov til at se.

Administration omfatter funktioner til

- at varetage den overordnede administration og drift af et elektronisk arkiv, som f.eks. forhandling med udgivere om betingelser for forsendelse af informations pakker,
- kvalitetskontrol af forsendelser og indhold i arkivet,
- vedligeholdelse og udbygning af arkivsystem.

OAIS Modellen er beskrevet i en Red Book (ref.

<http://www.ccsds.org/documents/pdf/CCSDS-650.0-R-1.pdf> , som et udkast til en anbefaling. Der kan ikke redegøres nærmere for, hvornår den kommer som en endelig anbefaling, CCSDS bruger følgende farver i forbindelse med behandling af anbefalinger.

Blue Book = CCSDS Recommendation

Pink Book = CCSDS Draft Recommendation (update to current Recommendation)

Red Book = CCSDS Draft Recommendation (a Recommendation has NOT been adopted)

White Book = CCSDS Draft Recommendation (while being developed into Red Book)

4.3.1.1 Referencer

Yderligere informationer kan findes på:

WWW.CCSDS.ORG

http://ssdoo.gsfc.nasa.gov/nost/isoas/workplan_oaisrm.html

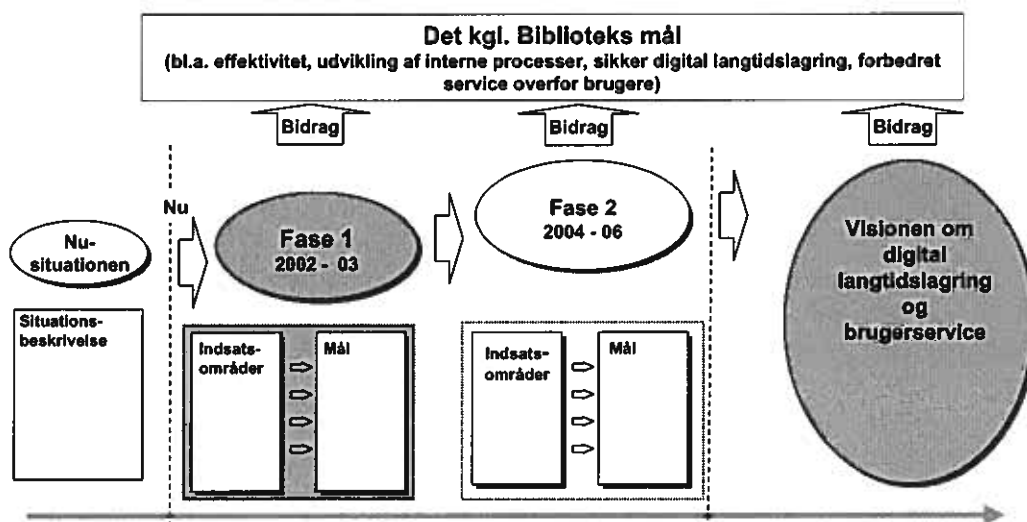
<http://www.rlg.org/events/pres-2000/sijtsma/sld009.htm>

5. Indstilling om realisering af handlingsplan

Det Kgl. Bibliotek skal i løbet af 2001 udarbejde en plan for hvorledes biblioteket mest hensigtsmæssigt kan indsamle, langtidslagre og videregive elektronisk materiale.

Det drejer sig om at lægge en plan således at KB i år 2002 kan starte med at iværksætte implementeringen af hensigtsmæssige arkiveringsfunktioner på det digitale område.

Det er ønsket at udarbejde en strategi for digital langtidslagring med tilhørende brugerservice funktioner. Strategien vil i princippet indeholde de på figur 1 viste elementer.



Figur 1 Elementer i en strategiproces

Der skal udarbejdes en vision for KBs brug af digital teknologi i langtidslagring og brugerservice. Udviklingen kan med udgangspunkt i nu-situationen opdeles i et antal faser. Der skal udarbejdes en *handlingsplan*, som giver mulighed for at sætte konkrete mål og bestemme de nødvendige indsatser for en første fase, der fx. kunne dække år 2002 – 03.

Handlingsplanen skal dels inddrage erfaringer med og forventninger til teknologien, dels inddrage planlægning af de ændringer teknologien vil give anledning til i bibliotekets procedurer omkring modtagelse, produktion og registrering af materiale, lagring af materiale samt tilgængeliggørelse af materiale.

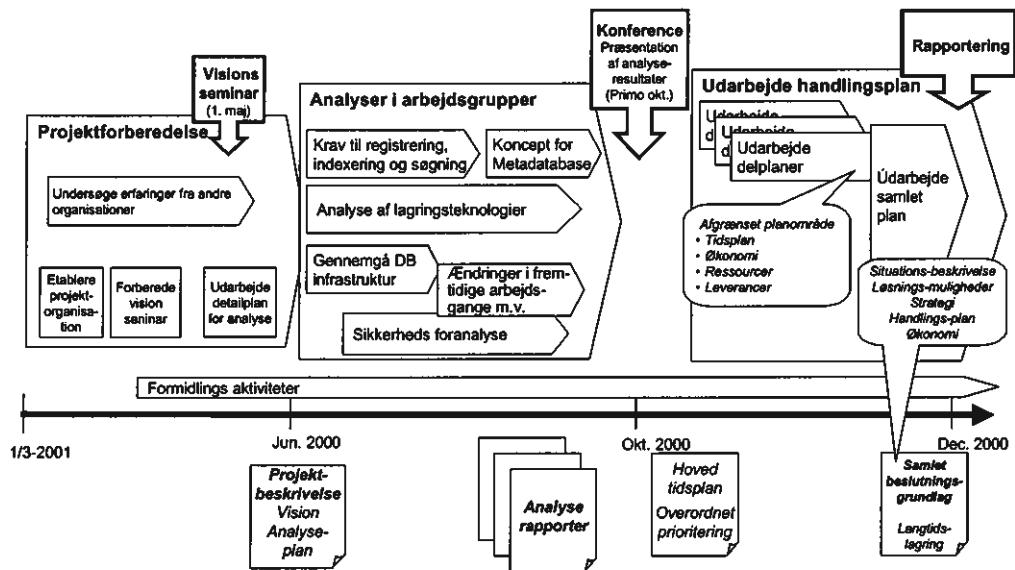
Det forventes, at udarbejdelsen af handlingsplanen kan startes marts 2001 og afsluttes december 2001. Resultatet af arbejdet skal være et samlet beslutningsgrundlag for digital langtidslagring fase 1, herunder indstilling om

- Strategi og plan for implementering af langtidslagring (konvertering, migrering etc.)

Endvidere skal der indgå investeringsbehov i hardware og software, behov for programudvikling og behov for udvikling af bibliotekets arbejdsgange og kompetence.

5.1 Tidsplan og budget for projekteringsprojektet

Et udkast til tids- og aktivitetsplan for udvikling af denne handlingsplan fremgår af figur 2. I det følgende vil vi kalde dette for *projekteringsprojektet*, idet det drejer sig om projektering af den fremtidige digitale langtidslagring.



Figur 2 Tids- og aktivitetsplan for projekteringsprojektet

Ideen i aktivitetsplanen er at gennemføre projekteringsprojektet i tre trin og på en sådan måde at et samlet beslutningsgrundlag foreligger ultimo 2002:

- *Projektforberedelse*, hvor projektet endelig afgrænses, projektorganisation og ressourcer bringes på plads og hvor der starter en formidling af ideen med projektet bredt i Det kgl. Bibliotek.
- *Analyser i arbejdsgrupper*, hvor en række analyser gennemføres og delrapporter udarbejdes
- *Udarbejdelse af handlingsplan*, hvor der udarbejdes delplaner og der etableres en samlet handlingsplan for langtidslagring.

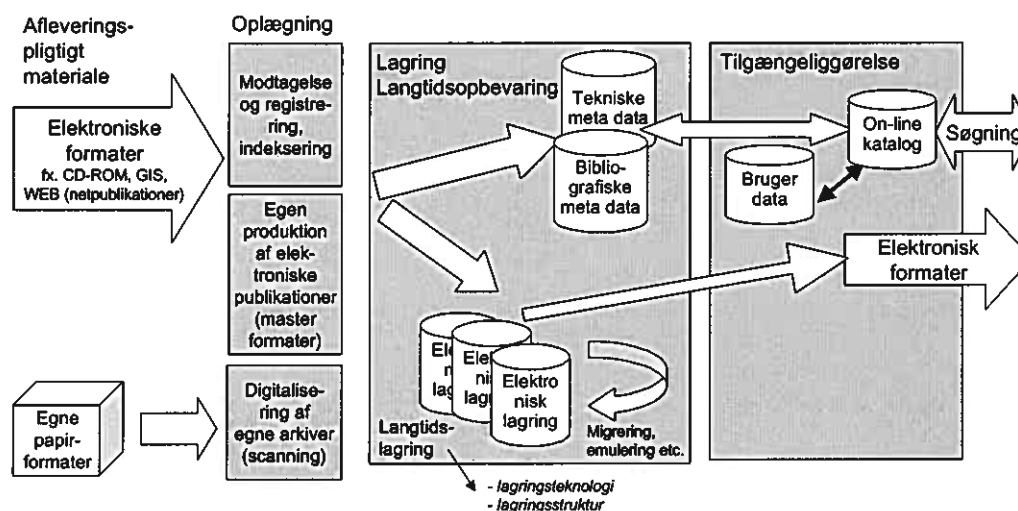
Nedenstående ses et budget for konsulentassistance til projekteringsprojektet. Forudsætningen for budgettet er, at alle projektaktiviteter gennemføres i samarbejde med projektdeltagere fra KB, og at KB tager ansvaret for leverance af resultater fra nogle af aktiviteterne som f.eks. krav til registrering, indeksering og søgning, gennemgang af KB's infrastruktur og ændringer i arbejdsgange. Ansvarsfordelingen vil blive fastlagt i med opstart af projektet.

Projektfase/-aktivitet	Konsulentass.
Projektforberedelse ca.	180
Projektmøder	
Undersøg erfaringer fra andre org.	
Konkluder og forbered visionsseminar	
Visionsseminar	
Udarbejd detailplan for analyse	
Analyser - ca.	310
Projektmøder	
Krav - registrering, indeks, søgning	
Koncept - metadatabase	
Lagringsteknologier	
Gennemgå KB's infrastruktur	
Sikkerhedsforanalyse	
Identificér ændringer arbejdsgange	
Samlet analyse	
Forbered konference	
Konference - analyseresultater	
Handlingsplan - ca.	220
Projektmøder	
Udarbejd delplaner på 5 områder	
Udarbejd samlet plan	
Forbered seminar	
Rapportering-seminar	
Konsulenttimer i alt - ca.	710
Konsulentassistance i alt ca.	1.200.000

Med hensyn til KB's egne ressourcer anbefaler Fischer & Lorenz, at KB afsætter minimum 2 årsværk til projektledelse og projektarbejde.

5.2 Projektforberedelse

Projektet vil berøre mange processer og arbejdsgange i KB. I figur 3 er skitseret de væsentligste processer, som skal indgå i planlægningen af digital langtidslagring.



Figur 3 Processer og arbejdsgange

Digitalt afleveringspligtigt materiale (fx. i form af CD-ROM, GIS-kort materiale og net-publikationer) skal modtages, registreres og indekseres. På tilsvarende vis skal egenproduceret digitalt materiale og digitaliseret (scannet) materiale af egne papirformater oplægges, så det kan indgå i det digitale lager.

I forbindelse med digital langtidslagring skal problemer omkring den rette lagringsteknologi og lagringsstrukturen behandles. Strategien for sikring af fremtidig tidssvarende langtidslagring skal ligeledes behandles. Der tænkes her eksempelvis på systemer og procedurer til migrering eller emulering.

Endelig skal muligheder og procedurer for tilgængeliggørelse behandles, herunder sikkerhed, bruger kategorier, brugerrettigheder og betaling af ydelser i denne forbindelse.

Den projektorganisation, der skal skabes under projektforberedelsen skal afspejle at alle disse organisatoriske og tekniske faktorer skal indgå.

Det vil være hensigtsmæssigt at starte projekteringsprojektet med et *visionsseminar*, hvor visionen for digital langtidslagring og tilgængeliggørelsen drøftes og hvor udfordringerne i forbindelse med denne udvikling behandles. Visionsseminaret kunne ligeledes indeholde eksterne indslag fra andre organisationer, der arbejder med emnet og har erfaringer med planlægning eller realisering af løsninger.

Formålet med visionsseminaret er således at give alle projektdeltagere et fælles udgangspunkt og et fælles begrebsapparat omkring digital langtidslagring.

Resultatet af dette trin vil være en projektbeskrivelse med en detaljeret analyseplan, herunder kommissorier for de analysegrupper, som skal i gang med de detaljerede analyser, samt endelig en skitse af de visioner, KB har på området.

Dette trin forventes at vare ca. 3 måneder.

5.3 Analysearbejde

Analysearbejdet vil omfatte en lang række temaer. Nogle af disse er relativt overskuelige, andre vil være mere omfattende. På nuværende tidspunkt springer følgende temaer i øjnene:

- fremtidige krav til registrering, indeksering og søgning,
- udarbejdelse af koncept for metadatabase, herunder samspil med andre organisationer
- analyse af fremtidige lagringsteknologier med hensyn til kapacitet, levetid, software understøttelse og lagringsøkonomi
- analyse af fremtidige lagringsformater
- gennemgang af KBs elektroniske infrastruktur med henblik på at afgøre hvilke krav de fremtidige systemer vil stille
- sikkerhedsanalyse bl.a. omkring lagring og tilgængeliggørelse
- nødvendige ændringer i fremtidige arbejdsgange omkring oplægning, herunder egenproduktion og scanning, vedligeholdelse af digitale lagre og styring af tilgængeliggørelse,

De forskellige analysegrupper udarbejder rapporter, som præsenteres på en konference. Formålet med konferencen er at skabe et samlet overblik over analyseresultaterne, således at det vil være muligt at foretage en overordnet prioritering af indsatsområderne og efterfølgende udarbejde en hovedtidsplan.

Dette trin forventes at vare ca. 4 måneder.

5.4 Udarbejde handlingsplan

Udarbejdelsen af handlingsplanen kan foregå ved, at der udarbejdes en række delplaner. For hver delplan, der omfatter et afgrænset planområde, fastlægges tidsplan, økonomi herunder ressourceindsats og de leverancer planen vil resultere i.

Planerne kan herefter samles til et beslutningsgrundlag, der indeholder en kort situationsberetning (KBs udgangspunkt), en strategi for digital langtidslagring, en vurdering af løsningsmulighederne samt en mere detaljeret beskrivelse af en første fase incl. handlingsplan og budget.

Dette trin forventes at vare ca. 2 måneder.

Det må fremhæves, at tidsplanen er stram og forudsætter, at KB dels stiller en fuldtidsprojektleder til rådighed for projektet i perioden, dels frigør de nødvendige ressourcer til deltagelse i arbejdsgruppearbejdet.