



NB-notat 12-08

Politik for langtidsbevaring af digitalt materiale på Det Kongelige Bibliotek

Dette dokument opdateres og godkendes af Det Kongelige Biblioteks (KB) direktion hvert andet år. Revisionsansvaret påhviler den leder, der har det funktionelle ansvar for aktiviteten 'digital bevaring', for nærværende afdelingslederen for Digital Bevaring i Nationalbiblioteksområdet.

Dokumentet publiceres på KB's hjemmeside i en dansk og en engelsk version med det formål, at det kan indgå i den almindelige internationale evaluering af politikker på dette særlige område.

1. Indledning

Dette dokument beskriver KB's politik for langtidsbevaring af digitalt samlingsmateriale. Det indledes med en gennemgang af de udfordringer, digitale materialer giver i forhold til bevaring og i forhold til at skabe vedvarende adgang til materialerne. Derefter beskrives formålet med politikken og de aktiviteter, den omfatter. Herefter følger et afsnit om de implementeringsprincipper, som politikken bygger på. Notatet afsluttes af en ordliste og en begrænset litteraturliste.

KB har siden 2001 haft en bevaringspolitik for de fysiske samlinger. Dette arbejde suppleres herved med nærværende politik for langtidsbevaring af bibliotekets digitale samlingsmateriale.

Store dele af bibliotekets samlinger bliver digitale og skaber dermed en række nye udfordringer for KB. Mht. bevaring af det fysiske materiale har KB således allerede magasinplads, en række gennemtestede og vedtagne politikker samt en økonomi afsat til området. Tilsvarende forhold skal etableres for at kunne bevare adgangen til det digitale materiale.

De udfordringer, som i den sammenhæng i særlig grad møder os, fordi der er tale om digitalt materiale, er:

- Der er tale om en ny materialetype, der vil være i stærk vækst de kommende år, men hvor vi pt. ikke kender det konkrete volumen



- Digitalt materiale kan ikke, som det f.eks. er tilfældet med bøger, henlægges passivt til det skal anvendes, da systemer, lagringsmedier, dataformater og programmer er i konstant forandring. Manglende handling – også i de tidlige faser - vil derfor med sikkerhed føre til tab af evnen til at give adgang til de bevarede data
- Der foreligger ikke nogle færdige metoder og løsninger til at klare problemerne, men der arbejdes internationalt i sektoren med at udvikle strategier og teknikker, der håndterer nogle af problemerne for delmængder af materialet
- Der foreligger derfor heller ikke nogen færdige modeller for, hvad aktiviteterne koster, hverken årligt eller samlet i materialets levetid, men der arbejdes internationalt på at etablere nogle første bud på omkostningsmodeller
- Lovgivningen omkring lagring og adgang til disse materialetyper er mere restriktiv end for det fysiske materiale

Der findes pt. ikke en national politik på området. Nærværende politik forholder sig således udelukkende til KB's egne behov på området.

Bevaringspolitikken for pligtafleveret internetmateriale fastlægges i tæt samarbejde med Statsbiblioteket og har sin egen bevaringspolitik.

Dette dokument er en del af en samling af dokumenter, der yderligere omfatter:

- *"KBs accessions- og kassationspolitik for digitalt samlingsmateriale"*
- *"Strategi for langtidsbevaring af digitalt materiale på Det Kongelige Bibliotek"*, der beskriver, hvorledes den vedtagne politik forsøges indfriet.

KB har normalt sine data anbragt i stabile IT systemer, som er underlagt backup. En backup er en kortsigtet løsning til at genetablere nyligt tabte data i de samme strukturer. Digital langtidsbevaring har til hensigt at sikre data over en meget langt tidshorisont og er ikke det samme som en backup.

2. Formål

Politikkens overordnede mål på KB er at indføre en risikostyring i forhold til de digitale samlinger som KB vil bevare over lang tid, ved

- at sikre, at digitalt samlingsmateriale, som KB anskaffer eller oparbejder inden for sin samlingsprofil, kan indgå i en tidssvarende formidling, i form af adgang og brug, til forskere og almindeligt interesserede brugere på såvel kort som langt sigt (> 100 år)
- at sikre, at de digitale bevaringsprocesser gennemføres på en sådan måde, at der fortsat er stor troværdighed til ægtheden (autenticiteten) af KB's samlinger
- at sikre at KB's løsninger følger internationale standarder og best practice



- at sikre, at KB har den nødvendige organisation og finansiering til løsning af opgaven

3. Aktiviteter omfattet af nærværende bevaringspolitik

For at sikre adgang til digitalt materiale over tid gennemføres en række bevaringsaktiviteter som går under fællesbetegnelsen ”langtidsbevaring”. Langtidsbevaring omfatter to hovedaktiviteter:

- Bitbevaring – bevaring af bitsekvensen
- Logisk bevaring – bevaring af evnen til at fortolke bitsekvensen

Bitbevaring består af sikker lagring af den fysiske bitsekvens, her forstået som lagring i systemer med velkendte risici, samt kontrol af integritet og genoprettelse af evt. fejl.

Logisk bevaring skal sikre de digitale materialer mod teknologisk forældelse, så man både nu og i fremtiden kan læse, forstå og vise/afspille materialerne med gængse programmer og udstyr.

3.1. Indlemmelse af materiale

Indlemmelse og kassation af samlinger, der skal omfattes af digital bevaring, følger dokumentet: ”*KB’s accessions- og kassationspolitik for digitalt samlingsmateriale*”.

Bibliografiske, bevarings-, formidlings- og forvaltningsmæssige beskrivelser (herunder ejerskabsforhold og adgangsrettigheder) - under ét kaldet metadata – udformes efter lokalt beskrevne normer, som ligger så tæt på internationale standarder og ’best practice’ som muligt.

3.2. Bevaring

For hver samling KB indlemmer tages der stilling til hvad der skal bevares både mht. indhold, struktur, funktionalitet og fremtræden. For hver samling tages der desuden stilling til bevaringsniveau, herunder bla. bitsikkerhed og, kryptering.

Indtil en finansiering er tilvejebragt, begrænses omfanget af digitale samlinger, der skal bevares, til fortrinsvis at omfatte digitale bevaringskopier og født-digitalt samlingsmateriale, jfr. ”*KBs accessions- og kassationspolitik for digitalt samlingsmateriale*”.

Der opstilles bevaringsstrategier for alle KB’s digitale samlinger, der er taget ud til bevaring, jfr. ”*Strategi for bevaring af digitalt materiale på KB*”.

Der sker en løbende overvågning af disse samlings sammensætning og volumen og der udarbejdes konkrete bevaringsaktionsplaner for de (del)samlinger, som måtte være truet.

Hvis det pga. tekniske eller økonomiske forhold ikke er muligt at bevare al funktionalitet i et digitalt objekt i forbindelse med gennemførelse af en bevaringsaktivitet, prioriteres en bevaring af det digitale objekts informationsindhold som hovedregel højest.



Tilsvarende sker en løbende overvågning i forhold til det eksterne miljø mht. den teknologiske udvikling af relevante værktøjer, standarder mm., samt i forhold til udviklingen i brugernes krav til funktionalitet.

3.3. Adgang til bevarede digitale samlingsdata

Som hovedregel er adgangen til de bevarede digitale master kopier (samlinger) reguleret, restriktiv og fuldt registreret.

Der skal være adgang til KB's bevarede digitale samlinger til følgende formål:

- On-demand fra brugere til materiale lokaliseret f.eks. via REX grænsefladen
- Reproduktion af tidssvarende formidlingskopier
- Forvaltning af samlingerne for den ansvarlige samlingsafdeling
- Udførelsen af automatiserede processer i forbindelse med gennemførelsen af bevaringspolitikken

4. Implementationsprincipper

4.1. Anvendelsen af standarder

Overalt i den internationale biblioteks- og arkivverden anvendes standarden "Reference Model for an Open Archival Information System (OAIS)" som referencemodel ved opbygningen af løsninger til at varetage de opgaver, som er forbundet med forvaltning og bevaring af digitale objekter. KB anvender tilsvarende denne OAIS-model som vores referencemodel i forbindelse med beskrivelse og opbygning af vores løsninger.

Også for andre væsentlige parametre som løsningsmodeller, omkostningsbeskrivelser og certificering som "Trusted Digital Repository" vælger KB at ligge på internationalt niveau.

KB støtter og praktiserer anvendelsen af open source, der i særlig grad er nødvendig for dette område, hvor ingen færdige løsninger foreligger endnu.

4.2. Data og Dataformater

Ved valg af dataformater vælger KB åbne, standardiserede formater med en begrundet formodning om, at de vil kunne håndteres inden for en af de bevaringsstrategier, som biblioteket vælger at benytte sig af.

Hvor det af praktiske eller økonomiske grunde besluttes at anvende komprimering af data, anvendes ikke-tabsgivende komprimeringsalgoritmer.

Samlingsdata lagres ukrypteret af hensyn til bevaringssikkerheden. I de få tilfælde hvor adgangssikkerheden vejer højere end bevaringssikkerheden, tages der stilling til om kryptering skal anvendes. Ved datatransmission af fortrolige samlingsdata krypteres disse, hvor det er muligt.



4.3. Lovgivning

Digitalt materiale er for nærværende underlagt de samme love som fysisk samlingsmateriale, men med flere restriktioner for benyttelse. I det omfang, biblioteket vælger at bevare digitalt samlingsmateriale, sker det på en måde, så de skærpede restriktioner, der primært vedrører adgangsforhold, kan sikres.

4.4. Teknisk infrastruktur

KB etablerer den nødvendige tekniske infrastruktur til at kunne gennemføre en bevaring af sine digitale samlinger.

Infrastrukturen understøtter den nødvendige overvågning af systemer, medier og formater for at sikre en proaktiv indsats.

Infrastrukturen omfatter værktøjer til såvel at understøtte beslutningsprocesser som at overvåge gennemførelse af konkrete bevaringsaktiviteter.

4.5. Finansiering

Bevaring af digitalt samlingsmateriale er en ny kerneydelse for biblioteksfunktionen på KB. En væsentlig del af opgaven er at sikre den nødvendige finansiering. Optimering af samlingernes brug gennem partnerskaber og samarbejde kan være et led i dette.

4.6. Samarbejde med andre institutioner

Arbejdsdeling (primært mht. ansvar for valg af strategier for dataformater) og driftsfællesskaber mellem danske, statslige bevaringsinstitutioner, der besidder den nødvendige kompetence, skal så vidt mulig aftales.

KB skal gennem strategiske samarbejdsprojekter sikre sig en central placering i det internationale netværk, således at institutionen juridisk, politisk, videnskabeligt, metodemæssigt og værktøjsmæssigt er på niveau med de mellemstore toneangivende nationalbiblioteker i Europa, vi normalt ønsker at sammenligne os med på dette område.

4.7. Bemanding

Det skal sikres, at biblioteket til stadighed har tilstrækkeligt, opdateret, fagkyndigt personale til, at bevaringspolitikken kan vedligeholdes og gennemføres forsvarligt, hvad enten den gennemføres internt eller ved outsourcing.

5. Ordforklaring

Ved **digital bevaring** forstås de processer, der indgår i at sikre vedvarende adgang til digital information på tværs af skiftende teknologier (lagringsmedier, dataformater og programmer til adgang) over tid. Adgang betyder i denne sammenhæng både, at data er fejlfrit opbevaret, men også at de kan fortolkes af kommende brugergrupper i et perspektiv på minimum 100 år. Digital bevaring er en opgave med mange forskellige aspekter, hvor der ikke findes færdige løsninger og svar på nuværende tidspunkt. Digital bevaring er ikke det samme som en backup. En backup er en kortsigtet løsning til at genetablere nyligt tabte data i de samme strukturer.



Der er grundlæggende 2 hovedaktiviteter, som begge skal gennemføres, for at man kan tale om digital bevaring:

- Man skal gennemføre en **bitbevaring**, dvs. at man skal bevare den helt korrekte sekvens af 0'er og 1'taller for eftertiden
- Man skal gennemføre en **logisk bevaring**, dvs. at man løbende i al fremtid skal have metoder til at omforme filerne til noget, der er forståeligt for mennesker. Det sidste kan grundlæggende ske ud fra tre forskellige bevaringsstrategier:
 - Den **museale**, hvor man forsøger at gemme den oprindelige hardware og software i form af de elektroniske læseapparater, der løbende udvikles til formålet.
 - **Formatkonvertering (også kaldet formatmigrering)**, hvor man sikrer, at man til stadighed har materialet på en form, som kan fortolkes med de til enhver tids gængse læseapparater/fortolkere. Ved formatkonvertering er der en risiko for at man mister noget – typisk ikke noget, der har med dokumentets intellektuelle indhold at gøre, men f.eks. noget, der kan knytte sig til det funktionelle, som look-and-feel. Formatkonvertering er lettest at gå til, fordi man kan påbegynde opgaven med det samme for de formater, hvor man kender en farbar løsning.
 - **Emulering**, hvor man enten i software eller i hardware sikrer en emulering af de oprindelige omgivelser, således at man ved afspilning af de originale programmer på de originale data kan genskabe den oprindelige oplevelse. Emulering er vanskeligst at gå til, fordi man skal løse alle udfordringerne som et samlet hele.

En digital **masterkopi**, er den kopi af et digitalt objekt, der anvendes som arkivkopi. Den er typisk mere rig på informationer end kopier til andre formål og fylder som sådan også normalt mere. En masterkopi danner oftest grundlag for produktion af afledte formater til f.eks. formidling.

I det omfang, der anvendes forskellige formater til formidling og langtidsbevaring, kan der etableres en **formidlingskopi**, der er en kopi af et digitalt objekt, der alene anvendes til en formidlingssituation. Den er typisk knap så rig på informationer som dens masterkopi og er oftest tilpasset et bestemt formidlingsformål. Levetiden for en formidlingskopi vil ofte være væsentligt kortere end en masterkopi, idet kravet til tidssvarende formidling er i konstant bevægelse.

6. Litteraturliste:

Erpanet Guidance: Digital Preservation Policy Tool, September 2003:
<http://www.erpanet.org/guidance/docs/ERPANETPolicyTool.pdf>

National Library of Australia: Digital Preservation Policy, 3rd Edition (2008)
<http://www.nla.gov.au/policy/digpres.html>

Neil Beagrie, Najla Semple, Peter Williams, and Richard Wright : Digital Preservation Policies Study, 30 October 2008, <http://www.jisc.ac.uk/publications/publications/jiscpolicyfinalreport.aspx>



Consultative Committee for Space Data Systems (CCSDS), Reference Model for an Open Archival Information System (OAIS), Magenta Book, Recommended Practice, CCSDS 650.0-M-2, 2012 (ISO14721:2003), <http://public.ccsds.org/publications/archive/650x0m2.pdf>

Hvordan laver man en bevaringspolitik og -strategi? <http://digitalbevaring.dk/bevaringspolitik-strategi/>

Statsbibliotekets politik for digital bevaring: http://www.statsbiblioteket.dk/om-statsbiblioteket/nyheder/filer/Politik_for_digital_bevaring_v%201%200.pdf