

Strategi for digital bevaring

Birgit Hansen, IT-chef i Statens Arkiver
Oktober 2001

Indledning

Statens Arkivers påbegyndte som et af de første nationale arkiver i verden tidligt i 1970'erne arbejdet med langtidsbevaring af e-arkiver. Som det var tilfældet i National Archives (USA) og i det svenske Rigsarkiv valgtes der en strategi, der byggede på bevaring af data i systemuafhængig form og på kopiering og konvertering til nye medier, når der var behov derfor.

Da Statens Arkiver i 1995 som de første i verden udarbejdede en standard for myndigheders aflevering af fuldt elektroniske sags- og dokumenthåndteringssystemer (så det ikke længere var nødvendigt at aflevere papirudskrifter) førte dette til, at Statens Arkiver måtte videreudvikle sin strategi for at sikre langtidsbevaringen. På det tidspunkt hjalp det ikke meget at skele til arkivvæsenet i andre lande, idet Danmark var/er pioner på området.

Valg af bevaringsstrategi

Valg af strategi for bevaring af e-arkiver indebærer blandt andet overvejelser omkring, hvor mange ressourcer arkivvæsenet har til rådighed og arkivlovens beføjelser. På verdensplan diskuteres ud fra 4 mulige strategier:

- den museale strategi, hvor der etableres et computermuseum, hvor alle udgaver af maskiner, programmer, styresystemer, processorer etc., opbevares så man er sikker på, at alle data kan læses med det programmel, de var udarbejdet på.
- emuleringsstrategi, hvor man emulerer – efterligner – de systemer, der har været anvendt til læsningen af de pågældende data. I stedet for at bevare de gamle computere, udvikler man programmer der kan virke ligesom dem.
- Non custodial archiving, hvor bevaringen bliver lagt på de enkelte myndigheder, der selv vælger og gennemfører en bevaringsstrategi.
- migrationstrategi, hvor data konverteres – og ikke applikationer (programmer) - med passende tidsinterval så de til stadighed lever op til den gængse teknologi – kaldes også for konverteringsstrategi

Den museale strategi er teknisk umulig og økonomisk meget dyr. Emuleringsstrategien er teknisk set svær, stor i omfang og tung i IT-kompetencer, og strategien er derfor meget dyr. non custodial archiving er risikofyldt, idet de enkelte institutioner ikke kan forventes at kunne have hverken de nødvendige IT-kompetencer eller det nødvendige udstyr for at sikre bevaringen. Migrationsstrategien er teknisk og økonomisk overkommelig og hører i øjeblikket til én af de mindst ressourcekrævende og mest effektive metoder, hvor e-arkivers bevaring og tilgængeliggørelse over tid samtidig er sikret. Statens Arkiver har derfor valgt migrationsstrategien som bevaringsstrategi for digitalt eller digitaliseret materiale.

Migreringsstrategien indebærer, at arkivvæsenet stiller krav til, hvordan myndighederne skal aflevere deres e-arkivsystemer, samt hvordan de bruger og opbevarer dem i forvaltningen. Det er nødvendigt med præcise krav til arkiveringsversionen for at kunne forstå – dokumentere – brugen af det elektroniske arkivsystem. Et andet vigtigt mål er, at systemerne afleveres på en måde, der sikrer at data ikke går tabt, når de konverteres, og at de kan blive tilgængelige for eftertiden.

Migreringsstrategien kræver, at arkivvæsenet hele tiden arbejder på at finde formater, der er

systemuafhængige, og som ikke benytter en uønsket tabsgivende kompressionsmetode og lignende. Selv det mindste tab af tilsyneladende ubetydelige data kan, efterhånden som de konverteres, medføre at betydningsbærende data mistes. Strategien kræver at hele arkivets elektroniske data skal konverteres løbende, men det er samtidig meget sikkert at data overlever.

Principper for valg af standarder for dataformater og medie

For at kunne bevare de afleverede e-arkiver forsvarligt på en måde, så de kan læses og bruges af eftertiden, har Statens Arkiver udpeget få (men tilstrækkelig mange) veldefinerede og velkendte formater, som myndighederne skal aflevere e-arkiver i. De udvalgte formater skal med sikkerhed kunne konverteres til fremtidige formater. Statens Arkiver har ligeledes udpeget ét elektronisk medie til aflevering. Derfor stilles der følgende krav til medier og formater:

- De skal være standardiserede – helst internationale de jure standarder - eller som minimum veldokumenterede
- De skal være bredt understøttede, platformuafhængige og åbne
- De må ikke være tabsgivende så kvaliteten af data forringes, og de skal have en lang forventet levetid
- De skal være system- og programuafhængigt, så de kan konverteres til kommende formater
- Formatets fremtrædelse skal være uafhængig af fonte mv., såfremt fremtrædelsen har betydning. Fremtrædelsen har ikke betydning i en tekstversion af et digitalt materiale, der udelukkende bruges til søgning.
- Mediets pris pr. MB og søgetiderne skal være lave.

Gennemførelse af migrationstrategien

Tidsintervallet for konverteringen bestemmes af tests af mediets kvalitet samt dets fysiske og teknologiske levetid. En systematisk og kontinuerlig test og kontrol af samlingen af e-arkiver gør langtidsbevaring og dermed tilgængeliggørelse mulig. Alle e-arkiver testes ved modtagelse fra myndighederne for at sikre at alle krav er overholdt. Desuden testes arkivalierne løbende for at teste og dokumentere den samlede bestand af e-arkiver for i tide at kunne konvertere til nyt datamedie.

Det er derfor nødvendigt med gode testværktøjer og testprocedurer for at sikre arkivaliernes overlevelse. Disse værktøjer må Statens Arkiver selv udvikle, og det er derfor nødvendigt at afsætte ressourcer til dette dels for at gøre test og konverteringer teknisk mulige, og dels for at effektivisere og automatisere test- og konverteringsopgaverne i takt med opgavernes stigende omfang.

Der er endelig vigtigt at sikre, at Statens Arkivers krav ikke bliver en hæmsko for den teknologiske udvikling i den offentlige forvaltning i Danmark. Også dette betinger, at Statens Arkiver nøje følger udviklingen på området. Migreringsstrategien hører i øjeblikket til én af de mindst ressourcekrævende og mest effektive metoder, hvor arkivaliernes opbevaring og tilgængeliggørelse samtidig er sikret, også i fremtiden.

Forudsætninger for bevaring af digitalt materiale

Bevaring af digitalt eller digitaliseret materiale kræver i modsætning til konventionelle papirbaserede materialer en aktiv bevaringsindsats fra skabelse af det digitale/digitaliserede materiale til bevaringsindsatsen over tid til formidling eller tilgængeliggørelse over tid af det digitale materiale. Indsatsen er løbende og kræver etablering og fastholdelse af en stab af kompetente, engagerede og bevaringsbevidste IT-medarbejdere samt løbende investeringer i det nødvendige udstyr, således at bevaring over tid bliver mulig.

Hvis indsatsen i en periode reduceres væsentligt er der alvorlige risiko for tab af data, såfremt de nødvendige konverteringer ikke finder sted. Datatabet kan i så fald være uopretteligt og en del af den

bevaringsværdige digitale kulturarv kan således være tabt for evigt.