

Statens Arkivers strategi for modtagelse, bevaring og tilgængeliggørelse af elektroniske arkivalier

Formål

I henhold til arkivlovens § 4 har Statens Arkiver bl.a. til formål at sikre bevaring af arkivalier og at stille arkivalier til rådighed for borgere og myndigheder, herunder til forskningsformål. Formålet med dette dokument er at beskrive og begrunde den strategi, som Statens Arkiver har valgt for at kunne modtage, bevare og tilgængeliggøre elektroniske arkivalier, så de kan stilles til rådighed for offentlighed og forskning. Strategien danner grundlag for Statens Arkivers arbejde med elektroniske arkivalier. Statens Arkivers indsamler først og fremmest elektroniske arkivalier, skabt af offentlige myndigheder, men strategien bør som udgangspunkt også anvendes til bevaring af elektroniske arkivalier, skabt af private arkivskabere.

Der gælder de samme overordnede krav til Statens Arkivers arbejde med elektroniske arkivalier, som der gælder til Statens Arkiver arbejde med papir arkivalier, nemlig at Statens Arkiver så vidt muligt:

- Skal kunne modtage de arkivalier, som er bevaringsværdige.
- Skal kunne bevare de modtagne arkivalierne på en måde, så deres informationsindhold kan bevares over tid, og så de senere kan tilgængeliggøres med så minimal en indsats som muligt. Tilgængeliggørelse indebærer, at arkivalier skal kunne genfindes, og at deres informationsindhold skal kunne genanvendes.
- Skal minimere myndighedernes udgifter til aflevering.
- Skal minimere sine egne udgifter til modtagelse, bevaring og tilgængeliggørelse.

Hvad er elektroniske arkivalier?

Elektroniske arkivalier er informationer gemt i elektroniske arkivsystemer. Elektroniske arkivsystemer er Statens Arkivers betegnelse for de IT-systemer, som myndighederne anvender ved forvaltningen af deres ressortområder og dækker bl.a. registre og databaser samt elektroniske journaler og elektroniske sags- og dokumenthåndteringssystemer.

Elektroniske arkivalier adskiller sig fra papir arkivalier ved ikke umiddelbart at lade sig genanvende uden brug af et computer program. Computer programmer er som bekendt underkastet den løbende teknologiske udvikling, og data skabt til at blive læst og anvendt v.h.a. et givet computer program lader sig derfor ikke nødvendigvis læse og anvende v.h.a. et andet program eller bare en nyere version af det samme program. Elektroniske arkivalier adskiller sig endvidere fra papirarkivalier ved at være lagret på medier, som har en meget begrænset levetid. Elektroniske arkivalier forudsætter derfor en mere aktiv bevaringsindsats end papir arkivalier, hvis de også skal kunne læses og anvendes i fremtiden.

Historie

Statens Arkiver har modtaget elektroniske arkivalier siden midten af 1970'erne. Indtil omkring 1988 skete det på bånd og som flade filer, mens både datas struktur og datas formater kunne variere. I 1975 kom de første regler om statslige myndigheders aflevering til Statens Arkiver, og i 2000 kom det første cirkulære, som i 2004 blev omsat til en bekendtgørelse og også sat i kraft for kommunernes aflevering af data til offentlige arkiver¹. Samtidig blev teksten omdisponeret.

Valg af strategi

I den internationale arkivverden findes der fire forskellige strategier for bevaring af elektroniske arkivalier:

- DEN MUSEALE, hvor der etableres et computermuseum, hvor alle udgaver af maskiner, styresystemer, programmer etc. opbevares med henblik på, at data kan læses med de programmer og på de maskiner, hvor de blev skabt.
- EMULERING, hvor man emulerer/efterligner de systemer, der har været anvendt til læsning af data. I stedet for at bevare de gamle computere udvikler, man altså programmer på moderne computere, der virker ligesom de gamle.

¹ Begrebet offentlige arkiver omfatter såvel Statens Arkiver som de offentlige arkiver, som kommunerne kan oprette. Jf. arkivlovens § 7.

- OPBEVARING HOS ARKIVSKABER, hvor opbevaringsansvaret ligger hos arkivskaber, der selv vælger og gennemfører en bevaringsmetode.
- KONVERTERING, hvor data med passende mellemrum konverteres til nye formater så de til stadighed kan håndteres med den gængse teknologi.

De tre første strategier indebærer bevaring af både systemer og data, mens den fjerde og sidste kun indebærer bevaring af data.

Statens Arkiver har valgt konverteringsstrategien. Det skyldes først og fremmest en vurdering af, at den i øjeblikket er den eneste realistiske. Den er derfor også den billigste. Hertil kommer, at den danske arkivlovgivning giver mulighed for at stille krav til offentlige myndigheders aflevering af arkivalier, hvilket er en forudsætning for bevaring i systemuafhængige formater.

Valget af konverteringsstrategien indebærer, at Statens Arkiver løbende må påtage sig at håndtere den udfordring, der består i at bevare arkivalier ved at omforme dem, uden at der sker tab af information.

Valget af konverteringsstrategien indebærer endvidere, at elektroniske arkivalier i endnu mindre grad end papirarkivalier bevares i den form, som de blev skabt. En kirkebog fremtræder stort set i samme form for brugeren på læsesalen, som for den person, der i sin tid førte den. De applikationer, som danner grundlag for registreringer i den elektroniske kirkebog, vil derimod ikke blive bevaret sammen med data. Det stiller større krav til dokumentation af elektroniske arkivaliers informationsindhold og den sammenhæng, som de er blevet til i, end der stilles til dokumentationen af papir arkivalier.

Strategi

Statens Arkivers strategi for modtagelse, bevaring og tilgængeliggørelse af elektroniske arkivalier baserer sig på følgende elementer:

- Periodisk aflevering af arkiveringsversioner
- Arkivering i systemuafhængige formater
- Periodisk konvertering af data til nye formater
- Periodisk overføring af data til nye bevaringsmedier

Periodisk aflevering af arkiveringsversioner

Offentlige myndigheder skal producere og aflevere arkiveringsversioner af de elektroniske arkivsystemer, der skal bevares, inden der sker sletning af data, eller inden systemerne går ud af administrativt brug. For de fleste statslige systemer betyder det, at der sker arkivering ca. hvert 5. år. Herved undgås det, at data går tabt ved sletning eller ved teknologisk forældelse. Sletning sker som oftest som følge af persondataloven, mens teknologisk forældelse er et resultat af, at de computerprogrammer, der er nødvendig for at kunne læse data, går ud af brug. Produktion og aflevering af en arkiveringsversion sikrer også, at dokumentationen af systemet ikke går tabt, men kan arkiveres som en del af arkiveringsversionen.

Dokumentationen af data vedrører dels datas indhold (semantisk dokumentation), dels datas struktur (syntaktisk dokumentation). Den semantiske del af dokumentationen vedrører formålet med og anvendelsen af det elektroniske arkivsystem, som arkiveringsversionen er produceret på grundlag af. Den semantiske del af dokumentationen har til formål at gøre det muligt for senere brugere at forstå datas betydning og datas ophavssituation. Den syntaktiske del af dokumentationen vedrører struktureringen af data og har til formål at gøre det muligt at genindlæse data i proprietære programmer, udviklet af Statens Arkiver eller andre.

Arkivering i systemuafhængige formater

I henhold til arkivlovgivningen afholder offentlige myndigheder selv udgifterne til aflevering af arkivalier. Myndighederne har derfor også ansvaret for produktion af arkiveringsversioner i det arkiveringsformat, som Statens Arkiver til enhver tid anvender.

Det arkiveringsformat, som Statens Arkivers i øjeblikket anvender, baserer sig på en opfattelse af, at offentlige myndigheders elektroniske arkivsystemer består af strukturerede data, eventuelt suppleret med dokumenter, som kan indeholde tekst, tal, grafik, lyd og levende billeder.²

Arkivering af en arkiveringsversion af en offentlig myndigheds elektroniske arkivsystem indebærer:

- At data struktureret i felter:
 - konverteres til en relationel struktur
 - konverteres til et standardiseret tegnsæt
 - konverteres til standardiserede repræsentationer
- At tekst dokumenter konverteres til et standardiseret format, som i øjeblikket er grafisk
- At lyd og levende billeder ligeledes konverteres til et standardiseret format
- At data dokumenteres, så de kan genanvendes

Når der anvendes en relationel struktur skyldes det, at den relationelle struktur i årtier har været en de facto standard for håndtering af data, der skal genanvendes i struktureret form.

Når der anvendes standardiserede tegnsæt og formater i øvrigt, skyldes det, at standardiserede tegnsæt og formater er umiddelbart tilgængelige for de myndigheder og leverandører, der skal producere arkiveringsversionerne. Det skyldes endvidere, at standardiserede tegnsæt og formater er så velbeskrevne, at de kan danne grundlag for senere konverteringer og den senere tilgængeliggørelse af data, uden at der sker datatab.

Der anvendes så vidt muligt kun formater, der ikke benytter en tabsgivende komprimeringsmetode. Det skyldes, at selv det mindste tab af tilsyneladende ubetydelige data kan medføre, at betydningsbærende data mistes, efterhånden som data konverteres flere gange.

Periodisk konvertering af data til nye formater

Den teknologiske udvikling medfører per definition, at Statens Arkivers bevaringsformat på et forholdsvist kort sigt bliver forældet i forhold til de elektroniske arkivsystemer, som myndighederne anvender. Det skyldes, at formatet ikke kan sikre bevaringen af de data, som myndighederne skaber med de nye teknologier. Et nyt bevaringsformat vil dog baserer sig på de samme principper, som er beskrevet ovenfor, nemlig en systemuafhængig struktur, standardiserede tegnsæt og formater, og semantisk og syntaktisk dokumentation.

Selvom Statens Arkiver tager et nyt bevaringsformat i brug ved modtagelse af afleveringer af arkiveringsversioner, indebærer det ikke nødvendigvis, at den eksisterende samling skal konverteres. Det er først nødvendigt, når det bliver for dyrt at vedligeholde og udvikle værktøjer til tilgængeliggørelse af data i det gamle bevaringsformat.

Periodisk overføring af data til nye bevaringsmedier

Ingen af de bevaringsmedier, der i øjeblikket er tilgængelige på markedet, har nogen lang fysisk eller teknologisk levetid. Statens Arkivers strategi for bevaring af elektroniske arkivalier indebærer derfor:

- at data opbevares i flere eksemplarer på flere forskellige typer af bevaringsmedier
- at bevaringsmedierne opbevares på forskellige adresser
- at bevaringsmedierne løbende overvåges
- at data periodisk konverteres til nye bevaringsmedier
- at data på bevaringsmedier ikke anvendes til tilgængeliggørelse

Der anvendes flere forskellige bevaringsmedier for at imødegå risikoen ved, at det ene nedbrydes. Hvor mange forskellige afhænger af kvaliteten af de medier, der er til rådighed på markedet. Opbevaringen sker på to adresser for at

² I øjeblikket er dette også i overvejende grad korrekt, bl.a. fordi den offentlige forvaltning stadigvæk i høj grad baserer sig på databaser og dokumenter, og fordi Statens Arkiver skal godkende de elektroniske arkivsystemer, som statslige myndigheder anvender til arkivering af data om dokumenter (elektroniske journaler) og til arkivering af data om dokumenter og dokumenter som sådan (elektroniske sags- og dokumenthåndteringssystemer). Denne opfattelse af virkeligheden er til en vis grad under pres, dels af de teoretiske overvejelser om standardisering der finder sted i diverse fora i centraladministrationen, dels af den teknologiske udvikling. Indenfor en kortere årrække må Statens Arkiver derfor redefinere sit systemuafhængige arkiveringsformat.

imødegå risikoen for tyveri og brand. Overvågningen indebærer en løbende, systematisk kontrol af, at klimaet i de magasiner, hvor bevaringsmedierne opbevares og af, at data kan læses. Denne overvågning er nødvendig for, at Statens Arkiver kan reagere på tegn på, at bevaringsmedierne er under nedbrydning.

Bevaringsmedierne vælges på grundlag af pris og kvalitet. Det indebærer, at Statens Arkiver vælger nyt bevaringsmedie til opbevaring af nye afleveringer, når der kommer bevaringsmedier på markedet, som har en acceptabel kvalitet, og som er så meget billigere, end de der hidtil har været anvendt, at det kan berettige udgifterne til et skift. Hele samlingen konverteres, når bevaringsmedierne viser tegn på nedbrydning, eller at teknologisk forældelse gør, at det begynder at blive dyrt eller besværlig at udlæse data.

Implementering af strategi

En bevaringsstrategi baseret på et systemuafhængigt arkiveringsformat medfører, at Statens Arkiver skal godkende nye elektroniske journaler og ESDH-systemer, inden de tages i brug. Herved kan Statens Arkiver på forhånd sikre, at der kan produceres en arkiveringsversion af systemerne. En bevaringsstrategi baseret på et systemuafhængigt arkiveringsformat medfører endvidere, at Statens Arkiver skal teste, om de arkiveringsversioner, som myndighederne afleverer, faktisk overholder arkiveringsformatet.

Statens Arkivers strategi for bevaring af elektroniske arkivalier medfører også, at den informationsteknologiske udvikling løbende følges. Det gælder både viden om, hvilke teknologier der kommer på markedet, og viden om i hvilket omfang, og hvordan de anvendes. Statens Arkiver har derfor udviklet en særlig strategi for vedligeholdelse af viden på det informationsteknologiske område.

Som det fremgår, indebærer implementeringen af konverteringsstrategien, at der løbende skal vælges nye teknologier. De valgte teknologier skal være velafprøvede og udbredte, og udgifterne til dem skal ligge inden for de økonomiske rammer, der er til rådighed for Statens Arkiver.

Specielt for udarbejdelsen af nye bevaringsformater gælder:

- at ethvert tab eller enhver ændring af data så vidt som overhovedet muligt skal undgås
- at det skal være så let og så billigt som muligt for myndighederne at gennemføre en aflevering
- at Statens Arkivers udgifter til bevaring og tilgængeliggørelse skal minimeres

Tilgængeliggørelse

Formålet med at bevare arkivalier er at kunne stille dem til rådighed. Statens Arkivers bevaringsstrategi danner derfor grundlag for al senere tilgængeliggørelse af elektroniske arkivalier. Statens Arkivers strategi for tilgængeliggørelse af elektroniske arkivalier indebærer, at der udvikles og vedligeholdes programmer til indlæsning af arkiveringsversioner i proprietære programmer, som er almindelig udbredt, eller i programmer som er udviklet af Statens Arkiver. Målet er, at de programmer der stilles til rådighed, skal kunne anvendes af brugere uden særlige IT-forudsætninger.