



Vejledning til aflevering af dokumenter fra ESDH-systemer

Marts 2008

					0				
	@							1	
				§					
							0		
			1						

Indholdsfortegnelse

1. TIDSRISTER FOR KONVERTERING AF DOKUMENTARKIVET	4
2. ANALYSE AF DOKUMENTFORMATER I ARKIVET	4
A. INDLEDENDE MØDE MED STATENS ARKIVER	4
B. RISIKO FOR DATATAB OG DATAFORVANSKNING	4
C. OPLYSNINGER, DER HAR BETYDNING FOR AFGØRELSE AF SAGER.....	5
D. DATABASER	6
E. ØVRIGE DOKUMENTFORMATER	6
F. INDLEJREDE DOKUMENTER, BILAG OG E-MAILS MED VEDHÆFTEDE FILER.....	6
G. FLERE VERSIONER AF ET DOKUMENT	6
H. LAGRINGSFORM FOR DOKUMENT/AKT	7
I. TEKSTVERSIONER AF EGENPRODUCERED E DOKUMENTER.....	7
J. LYD OG VIDEO.....	7
K. FILER STØRRE END 700 MB.....	7
3. KONTROL AF DOKUMENTER I ARKIVERINGSVERSIONEN	7
A. MANGLENDE DOKUMENTER I ARKIVERINGSVERSIONEN	7
B. DOKUMENTER, DER IKKE KAN KONVERTERES TIL TIFF	8
C. DOKUMENTER, DER IKKE ER INDSKANNET.....	8
BILAG – PROBLEMER MED LÆSNING/KONVERTERING AF FILER OG LØSNING HERPÅ	9

Indledning

Arkiveringsversioner af ESDH-systemer fra offentlige myndigheder skal opfylde kravene i Statens Arkivers Bekendtgørelse nr. 342 af 11. marts 2004. Hertil er udarbejdet en generel "Vejledning til Bekendtgørelse om arkiveringsversioner af bevaringsværdige data fra elektroniske arkiveringsversioner" i maj 2004. Statens Arkiver har i oktober 2005 yderligere udgivet "Vejledning om produktion af arkiveringsversioner", som imidlertid ikke omfatter arkiveringsversioner af ESDH-systemer.

Arkiveringsversioner af ESDH-systemer adskiller sig fra arkiveringsversioner af registre og e-journaler ved, at de tilhørende sagsdokumenter (dokumentarkivet) skal afleveres i elektronisk form i TIFF-format og for egenproducerede dokumenter endvidere i en tekstversion. De hidtidige erfaringer med afleveringer af arkiveringsversioner af ESDH-systemer har afdækket nogle generelle problemer i forbindelse med konverteringen af dokumentarkivet til TIFF-format, som denne vejledning adresserer. Denne vejledning skal derfor læses som et supplement til de to ovennævnte vejledninger.

Regioner og kommuner har ikke som statslige myndigheder afleveringspligt til Statens Arkiver, ligesom Statens Arkiver ikke godkender deres ESDH-systemer inden ibrugtagning. Det betyder blandt andet, at de ikke er forpligtet til at anvende arkivperioder og til aflevering af en arkiveringsversion af ESDH-systemet senest 6 måneder efter arkivperiodens afslutning. De tidsfrister, der er nævnt nedenfor gælder således ikke for regioner og kommuner, ligesom Statens Arkiver heller ikke tager initiativ til afholdelse af et møde for nærmere drøftelse af produktionen af en arkiveringsversion af ESDH-systemet. Regioner og kommuner skal selv tage initiativ til produktionen af arkiveringsversioner af deres ESDH-systemer for at opfylde arkivlovens krav om arkivmæssige hensyn. Statens Arkiver har udarbejdet en vejledning – "Varetagelse af e-arkivmæssige hensyn – en vejledning til kommunale myndigheder", hvor aflevering af e-arkivalier er nærmere beskrevet i vejledningens pkt. 6. Imidlertid skal, som nævnt ovenfor, alle arkiveringsversioner fra alle offentlige myndigheder opfylde kravene i Statens Arkivers bekendtgørelse nr. 342, og de nedenstående konkrete problemområder vil derfor også have relevans for konvertering af dokumentarkiver i ESDH-systemer i regioner og kommuner.

Bekendtgørelse nr. 342 om arkiveringsversioner af bevaringsværdige data fra elektroniske arkivsystemer og de ovennævnte tre vejledninger –

- Vejledning til Bekendtgørelse om arkiveringsversioner af bevaringsværdige data fra elektroniske arkiveringsversioner
- Vejledning om produktion af arkiveringsversioner
- Varetagelse af e-arkivmæssige hensyn – en vejledning til kommunale myndigheder

kan hentes på Statens Arkivers hjemmeside www.sa.dk.

Denne vejledning er senest opdateret d.11. december 2009.

1. Tidsfrister for konvertering af dokumentarkivet

ESDH-systemer skal opdeles i 5-årige arkivperioder; ved afslutningen af en arkivperiode kan myndigheden i op til 6 måneder benytte en såkaldt overlappingsperiode, hvor der stadig kan registreres dokumenter på sager fra den afsluttede periode, således at så mange sager som muligt kan lukkes inden arkiveringsversionen produceres.

For ESDH-systemer gælder, at ca. 6 måneder efter afslutningen af perioden (+ en eventuel overlappingsperiode), skal der afleveres en arkiveringsversion til Statens Arkiver.

De hidtidige erfaringer med afleveringer af ESDH-systemer har vist, at specielt konverteringen af større dokumentarkiver til TIFF-format kan tage så lang tid, at ovenstående tidsfrist ikke i alle tilfælde har kunnet overholdes.

Myndigheden skal derfor sikre sig – eventuelt i samarbejde med sin leverandør - at samtlige dokumenter i dokumentarkivet kan konverteres til TIFF inden for ovenstående tidsramme. Ved større dokumentarkiver vil det i en række tilfælde betyde, at konvertering af allerede arkiverede dokumenter til TIFF-format skal påbegyndes i god tid inden afslutningen af arkivperioden.

Myndigheden er i den forbindelse velkommen til at kontakte Statens Arkiver for at få fremrykket tidspunktet for nedenstående møde med Statens Arkiver.

Statens Arkiver anbefaler i øvrigt at konverteringsprocessen implementeres i ESDH-systemet, så der foretages en løbende og umiddelbar konvertering af alle dokumenter til TIFF, efterhånden som de arkiveres i systemet. En række af de problemfelter, der beskrives nedenfor, vil så på forhånd være løst.

2. Analyse af dokumentformater i arkivet

ESDH-systemer skal godkendes af Statens Arkiver inden ibrugtagning, og i den forbindelse skal myndigheden oplyse, hvilke filtyper der tillades arkiveret i systemet. Statens Arkiver anbefaler, at så få filtyper som muligt arkiveres i dokumentarkivet, da alle dokumenter, der har haft betydning for sagsdannelsen, skal kunne konverteres til TIFF-format i forbindelse med produktionen af arkiveringsversionen.

Hvis en myndigheds dokumenter på afleveringstidspunktet ikke kan konverteres til TIFF, er der tale om tilsidesættelse af arkivmæssige hensyn. Statens Arkiver indberetter væsentlig tilsidesættelse af arkivmæssige hensyn til det pågældende ressortministerium, jf. § 1 og § 3, stk. 4 -5, i bekendtgørelse nr. 591 af 26. juni 2003 om offentlige arkivalier og om offentlige arkivers virksomhed.

A. Indledende møde med Statens Arkiver

Statens Arkiver kontakter myndigheden ca. ½ år inden afslutningen af arkivperioden for at aftale et møde, hvor de nærmere forhold vedrørende produktionen af arkiveringsversionen af den afsluttede periode kan drøftes. Som forberedelse til mødet skal myndigheden fremsende følgende til Statens Arkiver:

- en liste over de filtyper, som p.t. findes i systemet, med angivelse af antal for hver filtype
- en liste over filer, der er større end 200 MB

Listen over filtyper danner – sammen med denne vejledning - udgangspunkt for en drøftelse af, hvordan konverteringen af dokumentarkivet kan foretages.

B. Risiko for datatab og dataforvanskning

Der kan bl.a. være risiko for datatab og dataforvanskning på følgende områder:

1. anvendelse af flerdimensionelle datastrukturer (fx "gule sedler" i elektronisk form, som enten ikke vil komme med i en udgave af TIFF eller som kommer til at dække over en del af teksten)
2. anvendelse af regneark, hvor formler og referencer mellem regneark går tabt ved konvertering til TIFF

3. anvendelse af indlejrede objekter (fx regneark indlejrede i tekstbehandlingsdokumenter)
4. anvendelse af dynamiske dokumenter og dokumentkæder (fx underdokumenter og hyperlinks)
5. ændring eller skift af applikationer, som påvirker tilgangen til dokumenter og sammensatte dokumenter
6. skift af internt lagringsformat (fx skift fra Word 97 til XML)
7. udskiftning af skrifttyper (fonte) og/eller ændringer af den i systemet anvendte tegntabel (codepage) (fx ved konvertering). En ofte forekommende fejl vil være udskiftning af nationale tegn og specialtegn med andre tegn (fx "Søren" udskiftes til "S]ren").
8. Dokumenter beskyttet med adgangskoder

Procedurer til sikring mod ovenstående datatab og dataforvanskning skal, såfremt systemet er blevet godkendt af Statens Arkiver efter marts 2002, være beskrevet i instruksen for anvendelsen af systemet. Myndigheden skal inden konvertering af dokumentarkivet sikre sig, at instruksen har været fulgt for pkt. 1- 4 og afklare hvilke konsekvenser et eventuelt skift i perioden af applikationer, lagringsformater, fonte og tegntabeller eventuelt har haft. Eventuelle dokumenter beskyttet med adgangskoder skal låses op af brugeren eller på anden vis (jf. bilag).

C. Oplysninger, der har betydning for afgørelse af sager

Alle oplysninger, der har betydning for afgørelse af sager, skal medtages i arkiveringsversionen. Der er derfor yderligere en række forhold, man skal være opmærksom på ved konvertering til TIFF-format. Har man eksempelvis anvendt Microsoft Office (men gælder principielt alle kontorpakker), skal man være opmærksom på at der forekomme følgende skjulte oplysninger:

I Word:

- skjult tekst
- kommentarer
- korrektur
- dokumentegenskaber

Hvis man har anvendt ovenstående, skal der tages hensyn til dette inden udskrivning via TIFF-printerdriver. 'Skjult tekst' kan udskrives ved markering under **Funktioner/Indstillinger/Udskriv**. Hvis man har indtastet og anvendt supplerende oplysninger om dokumentet med funktionen **Filer/Egenskaber**, kan udskrift af disse markeres på samme måde.

Er der kommentarer i dokumentet (ex. rettelser med 'forfatternavne'), der har betydning for afgørelse af sager, kan disse medtages under menupunktet **Vis/Markering**.

PowerPoint

- noter
- skjulte dias

Har der været anvendt noter, skal de ligeledes konverteres til TIFF. Skjulte dias medtages, hvis de har været anvendt.

Excel

- skjulte ark
- skjulte rækker/kolonner
- kommentarer
- brugerdefinerede visninger
- formelark

Alle ark, rækker og kolonner af betydning for sagers afgørelse, skal udskrives/konverteres til TIFF. Har der eventuelt været anvendt brugerdefinerede visninger i af betydning for sagers afgørelse, skal de ligeledes udskrives/konverteres til TIFF. Har der været anvendt beregninger/formler i regnearket, der har haft betydning for sagers afgørelse, skal formelark ligeledes udskrives/konverteres til TIFF.

Microsoft Project

- mange visningsmuligheder

De visningsmuligheder (gant-diagrammer, ressourcer, kalender etc.), der har haft betydning for projektet/sagens/sagernes afgørelse skal udskrives som dokumenter og medtages i TIFF-format i arkiveringsversionen.

Har myndigheden anvendt mange store/komplekse regneark eller Microsoft Project-filer, skal TIFF-konverteringer heraf forinden drøftes med Statens Arkiver.

D. Databaser

I forbindelse med godkendelsen af statslige ESDH-systemer, tillader Statens Arkiver normalt ikke, at der arkiveres databaser i dokumentarkivet. Findes der databaser i dokumentarkivet gælder – analogt til regneark ovenfor – at oplysninger i databasen, der har haft betydning for sagers afgørelse, skal udskrives så præcist og fyldestgørende som muligt i et dokument (rapport), der erstatter databasen.

For registrerede databaser, hvor ovenstående løsning ikke er mulig, skal Statens Arkiver tage stilling til eventuel bevarening eller kassation af databasen. For de databaser, der skal bevares, skal der afleveres en selvstændig arkiveringsversion.

E. Øvrige dokumentformater

Hvis systemet er blevet godkendt af Statens Arkiver efter marts 2002, indeholder instruksen en liste over de filformater, der må arkiveres i dokumentarkivet. Det er myndighedens ansvar, at der ikke er arkiveret filformater i ESDH-systemet, der ikke på afleveringstidspunktet kan konverteres til TIFF-format. Myndigheden skal inden mødet med Statens Arkiver have gennemgået den udskrevne liste over filtyper arkiveret i systemet, så eventuelle problemer kan blive drøftet. I vedlagte bilag er listet en række problemer, der kan forekomme ved konverteringen og forslag til løsning (jf. senere under pkt. 3 B om TIFF-filer, der ikke kan konverteres).

F. Indlejrede dokumenter, bilag og e-mails med vedhæftede filer

Alle dokumenter, som i ESDH-systemets driftsversion, kan identificeres som separate dokumenter (altså f.eks. kan klikkes frem) skal i arkiveringsversionen lagres separat og have hver sit entydige dokument-ID. Det gælder eksempelvis dokumenter med indlejrede, vedlagte eller vedhæftede bilag og e-mails med vedhæftede dokumenter. De indlejrede, vedlagte eller vedhæftede dokumenter skal tildeles hvert sit dokument-ID. Man skal samtidig være opmærksom på, at driftsversionens basale funktionalitet til genfinding af sagligt sammenhørende dokumenter, skal bevares i arkiveringsversionen - f.eks. at e-mail og vedhæftede dokument knyttes til samme sags- eller aktregistrering. Indeholder driftsversionen funktionalitet til hierarkisering af dokumenterne, f.eks. dokument/e-mail og tilhørende bilag, skal denne hierarkisering også bevares i arkiveringsversionen (f.eks. ved brug af rekursive referencer/nøgler).

G. Flere versioner af et dokument

I mange ESDH-systemer kan man ikke oprette et dokument uden at dette bliver registreret i ESDH-systemet. Et dokument kan derfor findes i mange forskellige versioner, hvilket gør afleveringen unødvendigt meget større, hvis alle dokumentversioner medtages. Hvis ESDH-systemet giver mulighed for at adskille de dokumentversioner, der repræsenterer egentlige sagskridt (som man normalt vil journalisere) fra de øvrige versioner, kan man undlade de mellemliggende kladder m.v.

F.eks.: Et notat er oprettet og bearbejdet flere gange, og findes i 10 versioner i ESDH-systemet. Der er imidlertid kun journaliseret de tre versioner, der udgør de egentlige sagskridt:

20061201 N Fremsender notat til kommentering
20061205 N Notat retur med kommentarer fra afd.chef NN
20061206 N Endeligt notat

Det er kun disse tre dokumenter, der skal medtages i arkiveringsversionen, mens de syv kladder ikke medtages. Kan man ikke lave denne adskillelse, skal der træffes separat aftale om håndtering.

H. Lagringsform for dokument/akt

I arkiveringsversionen skal der for hvert dokument/akt være registreret lagringsform. I nogle ESDH-systemer er det imidlertid muligt at anvende dokumentregistreringstabellen til at registrere hændelser/akter, der ikke har et tilknyttet dokument (også kaldt "tom" dokumentregistrering). Det kan f.eks. være muligt at registrere, at der har været en telefonsamtale ("Dd. ringet til klienten uden at træffe vedkommende hjemme") etc. I de tilfælde, hvor der ikke er – og ikke skal være – tilknyttet et dokument til en dokumentregistrering, skal feltet til angivelse af lagringsform for dokumentet være blankt.

I. Tekstversioner af egenproducerede dokumenter

Arkiveringsversionen skal indeholde en tekstversion af alle de dokumenter, som myndigheden har produceret elektronisk samt af de dokumenter, som myndigheden har OCR-behandlet. Det vil sige tekstversioner af alle dokumenter produceret af eller anvendt i myndighedens egne office-værktøjer (tekstbehandlings-, regnearksfiler mv.) samt PDF-dokumenter, hvorimod grafiske formater som TIFF, JPEG, CAD/CAM etc. er friholdt for kravet om konvertering til tekst.

J. Lyd og video

Hvis myndigheden har anvendt lyd og video skal disse filer afleveres i den kvalitet, som de oprindeligt er skabt og anvendt i hos myndigheden. Lyd og video skal konverteres til henholdsvis MP3- og MPEG2-formatet og skal konverteres på en sådan måde, at den oprindelige kvalitet i størst muligt omfang bibeholdes (MP3- og MPEG2-formatet er i sig selv tabsgivende; er der særlige behov for at aflevere lyd- eller video-sekvenser i et ikke tabsgivende format, vil Statens Arkiver kunne anvisne en løsning).

K. Filer større end 700 MB

Arkiveringsversionens enkelte dokumentfiler må ikke opdeles. I nogle ESDH-systemer er der imidlertid konstateret dokumenter, der er større end 700 MB, og som dermed ikke kan indgå i en arkiveringsversion.

Det har vist sig, at de aktuelle problemer har kunnet løses ved optimering af TIFF-konverteringen, således at dokumenternes størrelse er bragt ned under de 700 Mb. På Statens Arkivers hjemmeside findes et eksempel på hjælpeprogram, der kan anvendes til optimering af TIFF-dokumenter.

Hvis det viser sig, at der er tale om dokumenter, som ikke på en acceptabel måde kan optimeres, så filstørrelsen bliver mindre end 700 Mb (f.eks. meget store regneark) kontaktes Statens Arkiver for anvisning af en løsning.

3. Kontrol af dokumenter i arkiveringsversionen

Inden fremsendelse til Statens Arkiver testes arkiveringsversionen med Statens Arkivers hjælpeprogram TEA (Test Elektroniske Arkivalier), der kan downloades fra www.sa.dk.

A. Manglende dokumenter i arkiveringsversionen

Alle dokumenter i arkiveringsversionen, hvor lagringsform er angivet med enten værdien 1 = elektronisk dokument eller værdien 3 = delvist på papir, skal have filreference i dokumenttabellen og DOKMAB.TAB tabellen. TEA tester blandt andet for, at de angivne filreferencer er gyldige, dvs. at der til hver filreference også er lagret et elektronisk dokument.

Konstateres det under testen, at der mangler elektroniske dokumenter, skal myndigheden undersøge, om dokumenterne kan fremskaffes på anden vis. Hvis ikke, udarbejdes en liste over de manglende dokumenter med angivelse af eventuel årsag til, at dokumenterne ikke findes i dokumentarkivet. Listen udarbejdes i et regneark og skal indeholde følgende oplysninger:

DokumentID, journal/sagsnummer, sagstitel, akttitel, dokumenttitel, dokumentdato

Listen sendes til Statens Arkiver. Arkiveringsversionen kan ikke færdiggøres før Statens Arkiver har taget stilling til en løsning.

B. Dokumenter, der ikke kan konverteres til TIFF

TEA tester også for, at TIFF-dokumenterne ikke er tomme. Konstateres der under testen, at der er TIFF-filer, der er tomme eller andre dokumenter, der af en eller anden årsag ikke (jf. bilag for eventuelle løsninger) kan konverteres til TIFF, skal myndigheden forsøge at konvertere dokumenterne enten automatisk eller manuelt. Hvis det ikke er muligt, udarbejde en liste over filerne. Listen udarbejdes først, når det er konstateret, at det ikke er muligt at konvertere dokumenterne enten automatisk eller manuelt.

Listen udarbejdes i et regneark og skal indeholde følgende oplysninger:

DokumentID, journal/sagsnummer, sagstitel, akttitel, dokumenttitel, dokumentdato samt eventuel årsag til manglende konvertering.

Listen sendes til Statens Arkiver. Arkiveringsversionen kan ikke færdiggøres før Statens Arkiver har taget stilling til en løsning.

C. Dokumenter, der ikke er indskannet

I ESDH-systemets instruks skal der være beskrevet en procedure for arkivering af dokumenter, der ikke er indskannet eller kun indskannet delvist. Disse papirdokumenter skal afleveres efter Statens Arkivers almindelige regler for papirarkivalier, det vil sige *inden* de er 30 år gamle. Der kan dog dispenseres fra denne frist, f.eks. hvis myndigheden har administrativ brug for dokumenterne ud over denne periode.

Hvis en myndigheds dokumenter på afleveringstidspunktet ikke kan konverteres til TIFF, er der tale om tilsidesættelse af arkivmæssige hensyn. Statens Arkiver indberetter væsentlig tilsidesættelse af arkivmæssige hensyn til det pågældende ressortministerium, jf. § 1 og § 3, stk. 4 -5, i bekendtgørelse nr. 591 af 26. juni 2003 om offentlige arkivalier og om offentlige arkivers virksomhed.

Bilag – Problemer med læsning/konvertering af filer og løsning herpå

Statens Arkiver benytter fire dokumentformater (TIFF, MP3, MPEG2 og TXT). Fælles for disse formater er, at de kan rammes af en række generelle fejl, hvor årsagen skal findes i det generelle operativsystem eller hardware. Hertil kommer typer af fejl, som er mere specifikke, og hvor årsagen kan skyldes formatets opbygning eller fejl/mangler i den software, som skal håndtere (vise/konvertere) dokumentet. Måden hvorpå disse problemer kan håndteres afhænger fuldstændig af, hvad der er årsagen.

Eksempler på problemer:

Generelle

- Fejl i filsystem (fx FAT tabel indeholder urigtige oplysninger). *Kan evt. løses ved at indlæse fra backup.*
- Fejl i filsystem (manglende eller overskrevet data). *Kan evt. løses ved at indlæse fra backup.*
- Manglende ekstension (fx operativsystem ved ikke hvilken applikation som skal benyttes til at håndtere den pågældende fil). *Kan evt. løses ved at analysere data.*
- Ukorrekt ekstension (fx operativsystem prøver at åbne dokumentet med den forkerte applikation). *Kan evt. løses ved at analysere data.*
- Ukorrekt registrering af ekstension (fx ekstension er korrekt, men operativsystem prøver at åbne WORD dokument med forkert applikation). *Kan evt. løses ved at analysere data.*
- Konvertering fra gammelt format til nyt (fx fejl grundet manglende fonte i forbindelse med konvertering fra WP til Word som udprintes til TIFF). *Find de manglende fonte og installer dem.*
- Password beskyttede filer. *Kan i visse tilfælde åbnes/brydes med frit tilgængelig software.*
- Krypteret data. *Ingen umiddelbar realistisk løsning*
- Hardware ikke skaleret korrekt (fx for lidt hukommelse som gør at software fejler og i værste fald konverterer ukorrekt). *Opgrader hardware.*
- Filen kan læses og udskrives af systemet, men ikke konverteres til TIFF. *Dokumentet udskrives og indskannes i TIFF-format.*
- Fejl i printerdriver fra 3. part som gør at udskrivning (til papir eller TIFF) fejler (fx tomme sider, forvansket indhold eller forkert skalering af tekst). *Kan evt. løses ved opdatering af driver.*

Specifikke

- TIFF format ukorrekt gemt (fx ukorrekt angivelse af side offset som visse applikationer evt. vil kunne håndtere, medens andre vil fejle ved enten at give en meddelelse, gå ned eller i værste fald konvertere ukorrekt). *Kan ofte løses ved at benytte en alternativ applikation.*
- TIFF format korrekt men ikke kendt af applikation (fx visse afarter af JPEG kompression som visse applikationer evt. vil kunne håndtere, medens andre vil fejle ved enten at give en meddelelse, gå ned eller i værste fald konvertere ukorrekt). *Kan typisk løses ved at tilføje plugin eller DLL.*
- MP3/MPEG2 ukendt dekoder. *Kan løses ved at installere dekoder.*
- MP3/MPEG2 fejl i dekoder. *Kan evt. løses ved at installere alternativ dekoder.*