

# Dokumentvejledningen

---

*Vejledning i håndtering af dokumenter ved produktion af arkiveringsversioner.  
Supplement til Vejledning til Bekendtgørelse om arkiveringsversioner.*

*Rigsarkivet december 2015, BETA-version*

Beta-version

## Indhold

<b>0. Læsevejledning til Dokumentvejledningen</b> .....	<b>2</b>
A. Vejledningens målgruppe og anvendelse.....	2
B. Henvisning til øvrig vejledning.....	2
C. Lovgivning og retsfor skrifter.....	3
D. Definitioner.....	3
<b>1. Håndtering af digitale dokumenter</b> .....	<b>4</b>
A. Kvalitetssikring i driftsperioden.....	4
B. Flere versioner af samme dokument .....	4
C. Indlejrede dokumenter og vedhæftede filer.....	4
D. Pakkede og komprimerede filer (Zip-filer) .....	5
E. Kassation af dokumenter.....	5
<b>2. Arkivformater</b> .....	<b>6</b>
A. Hvad er et dokument?.....	6
B. Tekst og billeder .....	6
C. Lyd og video.....	6
D. Geografiske data.....	7
E. Databaser.....	7
F. Indlejrede elementer i dokumenter.....	7
G. Regneark.....	8
H. Binary large object (dokumenter i BLOB-felter).....	8
I. XML og anden opmærket tekst .....	9
J. Øvrige dokumentformater .....	9
<b>3. Konvertering af dokumenter til arkivformat</b> .....	<b>10</b>
A. Tidsforbrug og ressourcer ved konvertering .....	10
B. Navngivning og placering af dokumenter i arkiveringsversionen .....	10
C. Konvertering af dokumenter til TIFF .....	10
D. Konvertering af lyd og video.....	11
<b>4. Problemer ved konvertering</b> .....	<b>12</b>
A. Konverteringsfejl .....	12
B. Erstatningsdokumenter .....	12
C. Manglende dokumenter i arkiveringsversionen .....	13
D. Filer med adgangskoder .....	13
E. Tomme dokumenter.....	13
<b>5. Bilag 1: Filer der altid kan kasseres</b> .....	<b>15</b>
<b>6. Bilag 2: Fejlliste</b> .....	<b>15</b>

## 0. Læsevejledning til Dokumentvejledningen

Offentlige myndigheder er forpligtet til at aflevere en kopi af data og dokumenter i bevaringsværdige it-systemer, kaldet arkiveringsversioner, til et offentligt arkiv. Statslige myndigheder skal aflevere til Rigsarkivet. Kommuner og regioner kan vælge, om de vil aflevere til Rigsarkivet eller oprette deres eget arkiv.

Rigsarkivet har fastsat en række bestemmelser for arkiveringsversionerne af hensyn til bevaring og fremtidig brug af data, som alle myndigheder skal overholde. Disse bestemmelser er beskrevet i Rigsarkivets bekendtgørelse om arkiveringsversioner (Bekendtgørelse nr. 1007 af 20. august 2010, herefter benævnt bekendtgørelse 1007/2010).

**Dokumentvejledningen handler om de særlige regler, der gælder for håndtering af dokumenter i arkiveringsversioner.**

Statslige myndigheder skal anmelde deres it-systemer til Rigsarkivet før ibrugtagning. Bevaringsværdige systemer skal endvidere godkendes af Rigsarkivet, for at det kan sikres, at data og dokumenter kan konverteres til en arkiveringsversion. Bestemmelserne er fastsat i Rigsarkivets cirkulære om anmeldelse og godkendelse (Cirkulære nr. 9290 af 21. juni 2013, herefter benævnt cirkulære 9290/2013).

Kommuner og regioner skal ikke anmelde deres systemer. Systemerne skal heller ikke godkendes, men de skal leve op til bestemmelserne i bekendtgørelse 1007/2010.

### A. Vejledningens målgruppe og anvendelse

Dokumentvejledningen henvender sig til dem, som producerer arkiveringsversioner af it-systemer med dokumenter, der afleveres til Rigsarkivet. Det kan være myndighedens egen it-afdeling eller en ekstern it-leverandør.

Dokumentvejledningen beskriver enkelte steder de situationer, hvor myndigheden skal kontakte Rigsarkivet og drøfte konkrete problemstillinger. It-leverandører bør gøre deres kunder opmærksomme på situationer, hvor der skal laves aftaler eller indhentes tilladelser fra Rigsarkivet, hvis myndigheden ikke selv er opmærksom på forholdet.

Vejledningen er relevant for arkiveringsversioner fra såvel statslige som regionale og kommunale myndigheder.

### B. Henvisning til øvrig vejledning

Foruden Dokumentvejledningen har Rigsarkivet udarbejdet andre vejledninger, der har betydning for produktion og aflevering af arkiveringsversioner:

- Vejledning til bekendtgørelse 2007/2010 om arkiveringsversioner
- Vejledning til cirkulære 9290/2013 om anmeldelse og godkendelse
- Kort og godt om aflevering af it-systemer
- Kort og godt om kassation af dokumenter
- Kort og godt om test af arkiveringsversioner
- Oversigt over de mest almindelige fejl ved test
- Eksempelaflevering med dokumenter AVID.SA.18000
- Introduktion til Ada, Rigsarkivets program til test af arkiveringsversioner

Alt vejledningsmateriale kan tilgås fra Rigsarkivets hjemmeside:

- Statslige myndigheder: <https://www.sa.dk/aflevering-arkivet/statslige/aflevering-it-systemer>
- Regionale og kommunale myndigheder: <https://www.sa.dk/aflevering-arkivet/kommunale/aflevering-af-it-systemer>

### C. Lovgivning og retsfor skrifter

Information om lovgivning m.v. findes på Rigsarkivets hjemmeside:

- Statslige myndigheder: <https://www.sa.dk/aflevering-arkivet/statslige/arkivlovgivning>
- Regionale og kommunale myndigheder: <https://www.sa.dk/aflevering-arkivet/kommunale/arkivlovgivning-kommunale>

### D. Definitioner

**Arkiveringsversioner af it-systemer med dokumenter** består overordnet set af tabeldata, der afleveres i en nærmere defineret xml-struktur (se bekendtgørelse 1007), samt af dokumenter, der skal afleveres i Rigsarkivets arkivformater.

**Arkivformater:** Rigsarkivet benytter 6 arkivformater: TIFF, JPEG2000, MP3, WAV, MPEG2 og MPEG4.

Derudover lagres eventuelle geodata som GML-filer efter nærmere fastsatte retningslinjer, som behandles i GML-vejledningen (se punkt **0.B Henvisning til øvrig vejledning** ovenfor). Yderligere om formater: se afsnit **2**

**Arkivformater.**

**Dokumenter** er filer fra kontorprogrampakker som tekstbehandling, præsentationer og regneark, foruden billeder, lyd og video. Databaser er ikke omfattet af Rigsarkivets og denne vejlednings dokumentbegreb. Om databaser i it-systemer med dokumenter, se punkt 2.F.

**Myndigheder** er alle offentlige myndigheder på alle administrative niveauer: kommuner, regioner og stat, herunder domstolene, særlige forvaltningsenheder, offentligt ejede virksomheder samt de selvejende institutioner, der har pligt til at aflevere til et offentligt arkiv. Hvis der er tvivl om, hvorvidt en institution falder ind under dette myndighedsbegreb, kan Rigsarkivet hjælpe med afklaring.

**Standardkassation** er kassation af filer, der optræder i Dokumentvejledningens Bilag 1.

# 1. Håndtering af digitale dokumenter

## A. Kvalitetssikring i driftsperioden

Statslige myndigheder skal løbende foretage kvalitetssikring af systemets dokumenter og dokumentregistrering (dokumenternes metadata). Når en statslig myndighed får godkendt et system efter cirkulære 9290/2013, stiller Rigsarkivet krav om, at der kun gemmes formater, der kan konverteres til arkivformat. Det er myndighedens ansvar at sikre, at dette overholdes i praksis, også selvom systemet teknisk set kan håndtere de ikke tilladte formater.

Rigsarkivet anbefaler, at også kommuner og regioner sikrer, at deres dokumenter kan konverteres, da deres arkiveringsversioner skal leve op til bekendtgørelse 1007/2010.

Myndighedens kvalitetskontrol skal sikre, at dokumenterne er læsbare, og at de er korrekt registreret i systemet (har de rigtige metadata), og at dokumenterne kan konverteres til arkivformaterne (Se Vejledning til bekendtgørelse 1007/2010, Bilag 5.E, foruden Dokumentvejledningen afsnit **2 Arkivformater**).

## B. Flere versioner af samme dokument

I mange it-systemer med dokumenter bliver der lagret flere versioner af samme dokument. Dette kaldes versionering. Versionering kan være udtryk for, at et dokument udarbejdes over flere dage. I disse tilfælde skal kun det færdige dokument indgå i arkiveringsversionen.

Versionering kan også anvendes i forbindelse med forelæggelse eller godkendelse af et dokument. Hvis dette er tilfældet, skal flere versioner af samme dokument ofte medtages, således at beslutningsgangen dokumenteres. Ved aflevering af et system, hvor versionering har været brugt på denne måde, skal myndigheden kontakte Rigsarkivet og lave en aftale om håndtering af disse dokumenter.

## C. Indlejrede dokumenter og vedhæftede filer

I mange programmer kan man indlejre en fil i et dokument, dvs. bruge én fil som container for en anden. Det gælder eksempelvis:

- et billede, animation (gif-fil), eller video, der indlejres i et tekstdokument, et regneark eller en præsentation
- et billede, en tekstfil, et regneark eller lignende, der er indlejret i en mail i stedet for at være vedhæftet

Et billede vises normalt i den fil, det er indlejret i. I sådanne tilfælde kommer billedet normalt med ved konvertering til arkivformat, men det er noget, som myndigheden skal være opmærksom på ved produktion af arkiveringsversionen. Indlejrede regneark og mediefiler skal oftest eksekveres (startes, åbnes), for at indholdet bliver synligt. Det betyder, at indholdet ikke automatisk kommer med i konvertering til arkivformat. Det kan i stedet være nødvendigt at håndtere filerne manuelt på samme måde som filer med konverteringsfejl eller få programmet særlige rutiner til dette. Se afsnit **4 A Konverteringsfejl**.

I mange mails er der vedhæftede eller indlejrede filer, som skal medtages i arkiveringsversionen. Referencen mellem dokumenterne skal registreres ved at tilføje et parent-ID (PID) til child-dokumentet i docIndex.xml. Dokumenternes indbyrdes relation registreres som en parent/child-relation. Herved er det muligt at genkende dokumenters oprindelige tilhørsforhold. Se Vejledning til bekendtgørelse 1007/2010, afsnit 4.C.6.a, særligt figur 4.4.

I mange mails indgår der logo-filer, visitkort og digitalsignaturer som vedhæftede eller indlejrede filer. Filer af denne type kan som udgangspunkt standardkasseres, se Bilag 1. Det vil derfor være en fordel at frasortere filerne i det omfang, det er muligt, forud for konvertering og aflevering.

#### **D. Pakkede og komprimerede filer (Zip-filer)**

Komprimerede filer skal pakkes ud og registreres enkeltvis med selvstændigt dokument-ID. Når filerne er pakket ud og konverteret til det respektive arkivformat, skal den oprindelige zip-fil kasseres. Filen behandles efter samme procedure som filer, der slettes ved standardkassation, se afsnit **4 B Erstatningsdokumenter**.

Pakkede og komprimerede filer kan håndteres, inden der laves dataudtræk til arkiveringsversionen. I denne forbindelse skal myndigheden tage stilling til, om de udpakkede filer kan standardkasseres i henhold til Bilag 1.

#### **E. Kassation af dokumenter**

Som udgangspunkt udsteder Rigsarkivet bestemmelser om bevaring eller kassation af hele it-systemer. I it-systemer med dokumenter kan Rigsarkivet dog i nogle tilfælde gives tilladelse til kassation af visse dokumenter. Der gives kun tilladelse til kassation af større, afgrænsede grupper af dokumenter, der falder i en af disse tre grupper:

- standardbreve eller andre standardiserede (ens) dokumenter
- dokumenter lagret i forbindelse med rutinesagsbehandling
- dokumenter, der udelukkende indeholder oplysninger, som også findes som registerdata i systemet

Dokumenterne skal kunne udskilles på en klar og entydig måde. I et ESDH-system eller andet system med dokumenter kan det dreje sig om alle dokumenter i en specifik saggruppe. Et eksempel på af det kasserede dokument skal indgå i kontekstdokumentationen. De kasserede dokumenters registreringsdata (metadata) må ikke slettes. Der skal oprettes erstatningsdokumenter for slettede dokumenter, se afsnit **4 B Erstatningsdokumenter**.

Hvis en statslig myndighed ønsker at kassere ens typer af dokumenter af den ovenfor beskrevne type, fx blanketter eller standardbreve, skal den ansøge Rigsarkivet herom. I ansøgningen skal myndigheden kunne

- udpege klart afgrænsede grupper af dokumenter
- redegøre for udsøgningen af de pågældende dokumenter
- argumentere for, hvorfor disse kan kasseres

Kommunale myndigheder skal følge kassationsreglerne i cirkulæreskrivelse om kassation af dokumenter i kommunernes it-systemer af 12. juni 2014 (herefter cirkulæreskrivelse 2014).

Hvis en statslig myndighed har fået tilladelse til at kassere dokumenter i et it-system, skal godkendelsen indgå i kontekstdokumentationen. For kommunale systemer skal der være en angivelse af om der er kasseret i henhold cirkulæreskrivelse 2014, hvis de ønsker at foretage kassation af dokumenter.

Der er i øvrigt en række filtyper, som umiddelbart kan standardkasseres. Se bilag 1.

## 2. Arkivformater

### A. Hvad er et dokument?

Dokumenter er som udgangspunkt alle former for dokumentation, der har været anvendt af myndigheden, og som ikke er lagret som rå data i en database. Det er åbenlyst, at tekstfiler, regneark og præsentationer er dokumenter, men det gælder også filer, der indeholder lyd og billedgengivelser, som er skabt eller har været anvendt hos myndigheden. Rigsarkivets brug af begrebet **dokumenter** er dermed bredere end den almene forståelse.

Hvis myndigheden er i tvivl om, hvorvidt en filtype skal bevares, skal myndigheden spørge Rigsarkivet for at få en afklaring og eventuel løsning på udfordringer for konverteringen. Foruden filtypen lægger Rigsarkivet især vægt på dokumenternes indhold ved vurdering af bevaring.

### B. Tekst og billeder

Der er som udgangspunkt to arkivformater til tekst, regneark, præsentationer og billeder: TIFF og JPEG2000. Det er dokumenternes indhold, der er afgørende for valg af arkivformat til den aktuelle dokumentindsamling.

#### TIFF

TIFF bruges fortrinsvis til tekstdokumenter. Der kan vælges mellem flere forskellige TIFF-formater alt efter tekstdokumentets indhold.

Dokumenttype	Egnet kompression
Sort/hvide dokumenter	TIFF 1 bit med kompression grp. 4 samt LZW
Tegninger og fotos, hvor antallet af farver varierer i størrelse og antal fra side til side, eller hvor farven i dokumenterne ikke er betydningsbærende	TIFF 4 til 24 bit med kompression LZW
Tegninger, kort, fotos etc., hvor den eksakte bevaring af dokumentet, herunder dets farver, er af betydning	TIFF 24 bit

#### JPEG 2000

JPEG2000 er velegnet til store tegninger, kort, fotos etc., hvor formålet er grafisk formidling eller gengivelse, men farven på de enkelte pixels er af mindre betydning. Som eksempel kan nævnes skannede byggesager eller affotograferede skibstegninger. Kvaliteten skal svare til den kvalitet, som myndigheden har lagret og anvendt i driftsversionen. For skannet eller affotograferet materiale skal kvaliteten i arkiveringsversionen i udgangspunktet modsvare kvaliteten af det originale materiale.

Imidlertid afspejler kvaliteten af billeder, hvor formålet er grafisk formidling eller gengivelse, ikke nødvendigvis den administrative anvendelse. Der kan være tilfælde, hvor det er hensigtsmæssigt at fortage en reduktion af kvaliteten for disse billeder i forbindelse med konvertering til arkivformat. Hvis it-systemet indeholder sådanne billeder, skal myndigheden kontakte Rigsarkivet for at træffe aftale herom.

Kvaliteten i arkiveringsversionen bør **ikke** være bedre, end den som myndigheden har anvendt i driftsversionen af systemet.

### C. Lyd og video

Hvis myndigheden har anvendt lyd og video, skal det i udgangspunktet afleveres i en kvalitet, der svarer til den, som oprindeligt er anvendt. Lyd skal konverteres til MP3 eller WAVE. Video skal konverteres til MPEG-2,

hvor lyden er indkodet som MP3, eller som MPEG-4, hvor lyden er indkodet som AAC. Konverteringen skal foretages, så den oprindelige kvalitet i størst muligt omfang bibeholdes.

Imidlertid afspejler kvaliteten af optagelser ikke nødvendigvis den administrative anvendelse. Der kan være tilfælde, hvor det er hensigtsmæssigt at fortage en reduktion af kvaliteten i forbindelse med konvertering til arkivformat. Hvis it-systemet indeholder lyd eller video, skal myndigheden kontakte Rigsarkivet for at træffe aftale herom.

#### D. Geografiske data

Geografiske data (geodata) kan ikke konverteres til de formater, der er beskrevet i bekendtgørelse 1007/2010, uden væsentlige tab af data.

Hvis der i et it-system findes enkelte filer, der indeholder geodata, eksempelvis i et ESDH-system, skal myndigheden kontakte Rigsarkivet for at aftale, hvordan de pågældende geodata skal håndteres.

Hvis det pågældende it-system derimod indeholder store mængder geodata, eller hvis systemets primære formål har været at håndtere geodata, skal data afleveres efter det format, der er fastsat i Bekendtgørelse 1007/2010. Formatet følger standarden GML 3.1.1.

#### E. Databaser

Hvis der identificeres mindre databaser i et system i forbindelse med produktion af arkiveringsversionen, og data herfra har haft betydning i sagsbehandlingen, skal data fra databasen konverteres til et godkendt filformat. Det kan eksempelvis være som rapporter eller tabeludskrifter, der derefter konverteres til TIFF.

Hvis ovenstående løsning giver myndigheden problemer, skal Rigsarkivet kontaktes. Rigsarkivet vil vurdere, om der skal udarbejdes en selvstændig arkiveringsversion fra databasen, eller om data kan kasseres.

Kommunale og regionale myndigheder skal aflevere eller kassere store databaser i henhold til bekendtgørelse nr. 995/2010. I tvivlstilfælde kontaktes Rigsarkivet.

Hvis der er tvivl om behandlingsmetoden, eller hvis antallet af databaser har et betydeligt omfang, skal Rigsarkivet træffe beslutning om den videre proces.

#### F. Indlejrede elementer i dokumenter

Et dokument, der er konverteret til TIFF, svarer kort fortalt til en udskrift. Hvis et dokument indeholder et indlejret element – fx skjult tekst i form af noter i et præsentationsprogram – vil elementet ikke blive synligt i arkivformat.

Rigsarkivet ser løbende problemer i forbindelse med konverteringer af dokumenter fra kontorprogrammer som Office-pakken og lignede programmer. Da disse dokumenter udgør en væsentlig del af de dokumenter, der afleveres, er der her en række udfordringer, hvor myndigheden skal være særlig opmærksom på konverteringsfejl. Det drejer sig primært om tekstbehandling, præsentationer og regneark (om regneark, se også afsnit **2 G Regneark**).

Kontorprogrammer giver i mange tilfælde brugeren mulighed for at indlejre elementer i dokumenter. Det drejer sig primært om følgende:

- skjult tekst (fx noter i præsentationsprogrammer)
- kommentarer



- korrektur

Hvis ovenstående er anvendt, skal det sikres, at elementerne er synlige, inden konvertering til arkivformat, så oplysningerne bevares ved konvertering.

## G. Regneark

Et regneark, der er konverteret til TIFF, svarer kort fortalt til en udskrift. Hvis et regneark har skjult indhold – fx kommentarer ("gule sedler") eller skjulte faneblade – vil elementet ikke blive synligt i arkivformat.

Regneark kan helt eller delvis have skjult indhold. Vær opmærksom på følgende:

- skjulte faneblade
- skjulte rækker/kolonner
- kommentarer
- brugerdefinerede visninger
- formelark
- labels
- kæder mellem faneblade eller andre regneark

Alle ark, rækker og kolonner af betydning for sagers afgørelse skal konverteres til arkivformat. Kolonnernes bredde skal svare til længden af den tekst, der står i kolonnen. Ellers vil man ikke kunne se hele teksten i det konverterede dokument. Har der eventuelt været anvendt brugerdefinerede visninger af betydning for sagers afgørelse, skal de ligeledes konverteres til arkivformat. Har der været anvendt beregninger/formler i regnearket, der har haft betydning for sagers afgørelse, skal formelark ligeledes konverteres til arkivformat.

Hvis der er anvendt betydningsbærende farvekoder, må regnearket ikke konverteres til sort/hvid. Når regneark konverteres til tiff, vil det ofte fylde mange sider. For at sikre overblik anbefales det, at der tilføjes kolonne- og rækkeangivelser på alle sider.

Har myndigheden anvendt store, komplekse regneark, skal konvertering af disse drøftes med Rigsarkivet.

En hyppigt forekommende fejl er, at brugeren har skrevet et enkelt ubetydeligt tegn langt nede i regnearket, eller at en kolonne er autoudfyldt. Det betyder, at der i konverteringen af arket vil blive medtaget mange kolonner og rækker som er uden indhold. Det tager uforholdsmæssigt meget plads og vil vise sig ved test af arkiveringsversionen. Dette kan imødegås ved at tjekke dokumentmængden for meget store regneark inden konvertering. Det vil spare tid og plads i den sidste ende.

## H. Binary large object (dokumenter i BLOB-felter)

I nogle it-systemer lagres dokumenter i databasen som Binary Large Objects (BLOB). Hvor dette er tilfældet, skal dokumenterne trækkes ud og registreres på samme måde som øvrige dokumenter i arkiveringsversionen. Det skal ske således, at de pågældende dokumenter fremstår i det layout, der har været tilgængeligt i systemets driftsversion.

Hvis et blob-felt udelukkende indeholder tekst, som formateres i applikationslaget, skal denne formatering også ske i forbindelse med udtrækket.

Hvis et blob-felt indeholder filer med egen formatering, skal producenten være opmærksom på, at der kan opstå samme problemer ved konvertering til arkivformat, som gælder for andre dokumenter. Se afsnit 4

### Problemer ved konvertering

## I. XML og anden opmærket tekst

Arkiveringsversioner af it-systemer må ikke indeholde xml-filer. De skal konverteres eller kasseres, alt efter deres formål i driftsversionen. Xml-filerne skal håndteres som vist i Figur 1:

Xml-filens funktion i driftsversionen	Håndtering	Bemærkning
Xml-følgesedler til e-mails, der er journaliseret sammen med den pågældende mail	Standardkasseres, jf. Bilag 1	
Xml-filer, der er vedhæftet e-mail, men som har en selvstændig betydning som bilag til e-mailen, dvs. <b>ikke</b> er en følgeseddel	Konverteres almindeligvis til TIFF	Hvis it-systemet indeholder mange filer af denne type, skal myndigheden kontakte Rigsarkivet og aftale, hvordan de skal håndteres
Xml-filer, der bruges til at lagre tekstdokumenter i enkelte felter i databasen som Binary Large Object (BLOB)	Dokumenterne i BLOB-felterne skal trækkes ud, konverteres til arkivformat og lagres som dokumenter.	Se afsnit <b>2 H Binary large object (dokumenter i BLOB-felter)</b>
Xml-filer, der er brugt som bærende datastruktur	Myndigheden skal kontakte Rigsarkivet og aftale, hvordan xml-data skal håndteres.	Fx xml-databaser, spørgeskemaer, indberetningskemaer, logfiler

Figur 1

## J. Øvrige dokumentformater

Andre filtyper kan vise sig umulige at konvertere uden væsentligt datatab. Hvis et it-system indeholder filer af denne type, skal myndigheden kontakte Rigsarkivet for at drøfte mulige løsninger. Figur 2 viser eksempler på sådanne filtyper og på mulige løsninger:

Filtype	Håndtering
Microsoft Project-filer med mange visningsmuligheder så som gant-diagrammer, ressourcer og kalender	Microsoft Project-filer skal udskrives som dokumenter og medtages i arkiveringsversionen, hvis dokumenterne har haft betydning for pågældende sag/projekt
CAD-filer	2D-vektorformater skal afleveres i formatet TIFF eller JPEG2000 3D-vektorformater skal afleveres i formatet TIFF eller JPEG2000 i form af et vertikalt og et horisontalt billede
SVG-filer	3D-vektorformater skal afleveres i formatet TIFF eller JPEG2000 i form af et vertikalt og et horisontalt billede

Figur 2

Hvis der i systemet i øvrigt er filtyper, som ikke kan konverteres, skal Rigsarkivet kontaktes.

### 3. Konvertering af dokumenter til arkivformat

#### A. Tidsforbrug og ressourcer ved konvertering

Når der skal produceres en arkiveringsversion af et it-system med dokumenter, skal dokumenterne konverteres til de tilladte formater. Arkiveringsversionen skal være færdig cirka 6 måneder efter afslutningen af journalperioden (systemets driftsperiode). Afleveringsfristen fastsættes af Rigsarkivet efter forhandling med myndigheden.

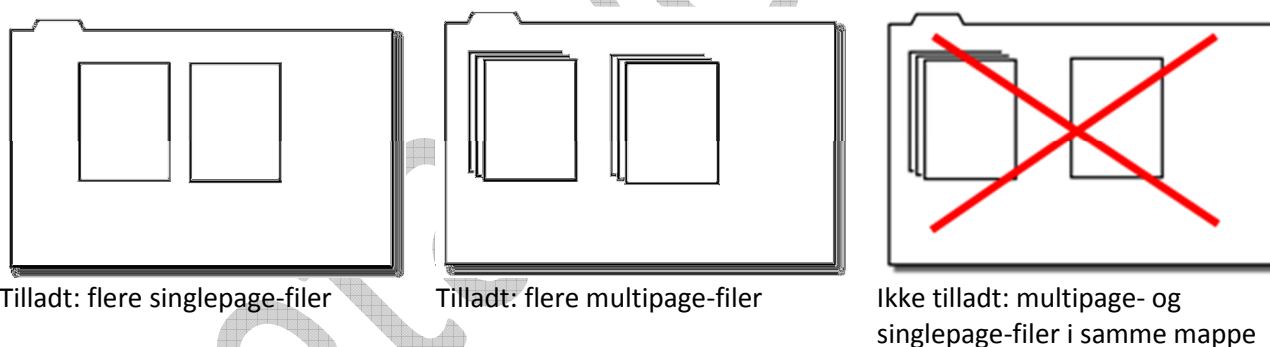
Myndigheden skal sikre sig – eventuelt i samarbejde med leverandøren – at samtlige dokumenter i systemet kan konverteres til et arkivformat inden for ovenstående tidsramme. Ved større it-systemer vil det i en række tilfælde betyde, at forberedelser til konvertering af dokumenter skal påbegyndes i god tid inden afslutningen af journalperioden.

#### B. Navngivning og placering af dokumenter i arkiveringsversionen

Hvert dokument skal placeres i en individuel dokumentfolder under en dokumentssamlingsmappe, som navngives "docCollection[fortløbende nummer]". Dokumentmapperne navngives med et fortløbende nummer. En dokumentssamlingsmappe med kontekstdokumentation må indeholde op til 10.000 dokumentmapper. Se i øvrigt vejledningen til bekendtgørelse 1007, afsnit 4.G.

Dokumenter i arkiveringsversionen skal navngives 1.[ekstension] (eksempelvis 1.TIFF eller 1.MP3).

Det enkelte tiff-dokument skal være **enten** singlepage **eller** multipage. Tiff-dokumenter, der er multipage, må gerne bestå af flere multipage-filer, se figur 3.



Figur 3

Hvis et dokument består af flere filer, skal de placeres i samme dokumentmappe og navngives med et fortløbende nummer. Se i øvrigt vejledningen til bekendtgørelse 1007, afsnit 4.C.6.a.

Det er nummeret på den dokumentmappe, dokumentet ligger i, der fungerer som dokument-identifikator. Referencen mellem dokumentet og den øvrige arkiveringsversion etableres i docIndex.xml. Se Vejledning til bekendtgørelse 1007/2010 afsnit 4.C.6.a og frem.

For alle dokumenter skal der i docIndex.xml angives oplysninger om dokumentets oprindelige filtype, og det skal angives, hvilken lagringsform dokumentet havde i driftsversionen. Endelig skal der i systemer, hvor det er relevant, angives parent/child-relationer mellem eksempelvis mails og vedhæftninger. Se også afsnit **1.C Indlejrede dokumenter og vedhæftede filer**.

#### C. Konvertering af dokumenter til TIFF

## Komprimering

Rigsarkivet anviser overordnede regler for komprimering af dokumenter ved konvertering til arkivformat jf. bekendtgørelse nr. 1007 om arkiveringsversioner. I visse tilfælde anviser Rigsarkivet derudover yderligere retningslinjer for optimeringen af dokumenter for at nedbringe pladsforbruget i arkiveringsversionen, fx ved at anviser anvendelse af optimal bitdybde.

Man kan optælle antallet af unikke farver og dermed bestemme den bedste bitdybde. En bitmap indeholder et antal pixels, som hver især repræsenterer en farve. Hver pixel optager plads afhængig af, hvordan man gemmer. En pixel kan gemmes i form af index til palette (1 til 8 bit) eller direkte som en binær værdi (typisk 16 eller 24 bit). En side med 15 unikke farver behøver kun 4 bits (en halv byte), hvorimod en side med 34.243 unikke farver har brug for 16 bit (2 bytes). Hvis man gemmer som 24 bit, vil der blive anvendt 3 bytes, uanset om siden indeholder mange, få eller ingen data. Hvis man optæller antallet af unikke farver anvendt på en side, kan man bestemme, hvor mange bits der er nødvendige for at udtrykke netop dette antal farver, og derefter gemme siden i denne bitdybde.

TIFF-formatet tillader blandt andet bitdybderne 1-8 samt 12, 15, 16 og 24, men de understøttes ikke alle af TIFF-viewere. De bedst understøttede bitdybder er 1, 2, 4, 8 og 24, hvorfor Rigsarkivet anbefaler, at kun disse bitdybder benyttes. Anvendelse af ekstrakanaler (alpha-kanaler) bør undlades, med mindre der er tale om specifik anvendelse (fx afmaskning af bitmap).

TIFF-angivelse af X-resolution og Y-resolution (DPI) er betydningsbærende og vil blive benyttet til skalering i forbindelse med visning/konvertering. Angivelse af X-resolution og Y-resolution bør som udgangspunkt være identisk (fx 300x300), men kan være forskellig for TIFF-filer, som benytter grp. 3- og grp. 4-kompression.

Hvis man kender filformatet, kan man ofte bestemme den optimale bitdybde. Tekstdokumenter (txt) kan gemmes som TIFF grp. 4, og JPG-filer kan gemmes som TIFF 24 bit LZW uden yderligere analyse. Dokumentsider med bitmap af meget begrænset størrelse må antages at være uden betydning for forståelsen af sidens indhold (logo, streger, ikoner). De kan derfor gemmes som 8 bit uanset antallet af unikke farver.

Valg af kompressionstype og bitdybde (antal mulige farver) må gerne variere fra side til side inden for samme dokument. Myndigheden skal sikre sig, at dokumenterne også efter konvertering til arkivformatet er læsbare fx. ved at kontrollere, at kommaer fremstår korrekt i dokumentet og ikke ligner punktummer. Hvis det er tilfældet, er opløsningen for lav, og dokumentet skal lagres med en højere opløsning. Rigsarkivet anbefaler, at der anvendes grp. 4 og LZW-kompression, hvor det er velegnet, eftersom dokumenter komprimeret med disse algoritmer fylder mindre, end hvis der benyttes andre kompressioner. Se i øvrigt afsnit **3 Konvertering af dokumenter til arkivformat**.

## D. Konvertering af lyd og video

Omfanget af lyd og video i myndighedernes it-systemer er endnu begrænset, og det samme er arkivernes erfaring med konvertering af disse filer. Derfor er der endnu ikke grundlag for en fastsættelse af generelle krav til konverteringen. Ved aflevering af lyd og video skal format og komprimering derfor altid drøftes med Rigsarkivet.

- Lydfiler skal konverteres til MP3. Rigsarkivet kan give tilladelse til at lydfiler afleveres i formatet WAVE, hvis myndigheden ansøger om dette. Valg af format og kvalitet indenfor disse formater afhænger af lydfilernes indhold og deres anvendelse i driftssystemet.

- Videofiler konverteres til MPEG-2 med et lydspor i enten MP3 eller AAC.

## 4. Problemer ved konvertering

### A. Konverteringsfejl

Ved konvertering viser det sig ofte, at dokumenter, der skal indgå i arkiveringsversionen, ikke kan konverteres automatisk. Producenten skal efter automatisk konvertering gennemgå alle de dokumenter, der ikke er konverteret korrekt, og konvertere disse manuelt i det omfang, det er muligt.

Myndigheden skal sende en liste over de dokumenter, som det hverken er muligt at konvertere automatisk eller manuelt, til Rigsarkivet. I fejllisten angives hvilke fejl, der knytter sig til det enkelte dokument. Fejllisten skal udformes som beskrevet i bilag 2. Rigsarkivet godkender fejllisten, hvorefter godkendelsen og fejllisten skal indgå som kontekstdokumentation i arkiveringsversionen.

De fejlbehæftede dokumenter medtages ikke i arkiveringsversionen, men erstattes af ét erstatningsdokument, se afsnit 4 B **Erstatningsdokumenter**.

### B. Erstatningsdokumenter

Der kan oprettes et erstatningsdokument for:

- Dokumenter kasseret efter aftale med Rigsarkivet (statslige myndigheder)
- Dokumenter kasseret efter cirkulæreskrivelse 2014 (kommunale myndigheder)
- Dokumenter kasseret på grund af konverteringsfejl

Om dokumenter, der er standardkasserede i henhold til Bilag 1: Se nedenfor.

Erstatningsdokumentet skal forklare årsagen til kassation. Der oprettes ét erstatningsdokument per kassationsårsag.

Dokumentreferencen fra hvert kasseret dokument skal pege på det erstatningsdokument, der forklarer dokumentets kassationsårsag. Hvis mere end et dokument er kasseret af samme årsag, skal deres dokumentreferencer pege på et og samme erstatningsdokument, nemlig det, der forklarer årsagen til kassation.

Erstatningsdokumenter skal tildeles dokument-id som angivet i Figur 4.

	Årsag til kassation	Dokument-id
	Standardkassation, jf. Rigsarkivets retningslinjer	[99999991]
	Dokument kasseret på grund af fejl i konverteringen	[99999992]
<b>Statslige myndigheder</b>	Dokument kasseret fordi Rigsarkivet har givet tilladelse i henhold til ansøgning fra myndigheden	[99999993]
<b>Kommuner</b>	Dokument kasseret i henhold til gældende bestemmelse om bevaring og kassation i kommuner og regioners it-systemer	[99999993]

Figur 4

Hvis der i et system er yderligere årsager til kassation, skal producenten rådføre sig med Rigsarkivet.

Se også afsnit 4. A **Konverteringsfejl**.

## Særligt om standardkassation i henhold til Bilag 1

Ved standardkassation af dokumenter i henhold til Bilag 1 kan myndigheden vælge mellem to forskellige løsninger.

1. Oprettelse af erstatningsdokumenter efter samme retningslinjer som for øvrige kasserede dokumenter (ovenfor)
2. Sletning af dokumentreferencer fra de standardkasserede dokumenter. Dog skal alle metadata om de kasserede dokumenter bevares. Bemærk, at lagringsform for kasserede dokumenter er værdien 3 ("ikke relevant"), som skal fremgå af kolonnen til registrering af dokumenternes lagringsform.

### C. Manglende dokumenter i arkiveringsversionen

Nogle systemer tillader, at brugerne kan henvise til dokumenter uden for systemet. Eksempelvis kan der være mulighed for at henvise til dokumenter på et lokalt fildrev, på myndighedens intranet eller i dokumentsiloer uden for selve systemet. Hvis ikke det er muligt at medtage dokumenterne i afleveringen, så de lægges ind og erstatter det henvisende link, skal myndigheden kontakte Rigsarkivet, som i samråd med myndigheden træffer beslutning om, hvordan problemet løses. Hvis der har været sådanne eksternt henviste dokumenter, skal dette dokumenteres i arkiveringsversionen, så en fremtidig bruger kan orientere sig om, hvordan systemet har været anvendt. Rigsarkivet stiller krav om dokumentationens udformning.

Hvis der ved konvertering af systemet til arkiveringsversionen opstår problemer med dokumenter, der ikke kan konverteres til arkivformat, eller hvis der er dokumenter, som helt mangler i systemet, skal myndigheden forsøge at konvertere dokumenterne manuelt eller genskabe dem. Ved genskabelse menes, at dokumentet eftersøges i mailprogram, på personlige drev, fællesdrev og lignende og lægges i arkiveringsversionen på det fejlkonverterede dokumentets plads. Arbejdet kan foretages i dialog med Rigsarkivet.

### D. Filer med adgangskoder

Alle dokumenter i systemet skal indgå i arkiveringsversionen og afleveres, med mindre der er tale om de filtyper, der altid kan standardkasseres, eller hvis Rigsarkivet giver tilladelse til kassation, se punkt 1.D.

Afleveringspligten omfatter også filer, der er beskyttet med adgangskode (password). Rigsarkivet anbefaler alle myndigheder at forhindre, at adgangsbeskyttede filer gemmes i afleveringspligtige it-systemer. Dette gøres almindeligvis ved at skrive det ind i journalinstruksen.

Hvis der alligevel optræder adgangsbeskyttede filer i et system, når arkiveringsversionen skal produceres, skal myndigheden omgang forsøge at få adgang til dokumentets indhold gennem den konkrete medarbejder. Hvis myndigheden har ret til indholdet i filen, kan myndigheden i stedet forsøge at knække adgangskoden (bryde passwordet). Hvis dette ikke lykkes, skal myndigheden forsøge at genskabe dokumentet, jævnfør genskabelse i afsnit **4 A Konverteringsfejl**. Hvis ingen af disse løsninger giver positivt resultat, skal myndigheden anmode Rigsarkivet om tilladelse til at kassere dokumentet jævnfør afsnit **1 E Kassation af dokumenter**.

### E. Tomme dokumenter

Når dokumenter konverteres til arkiveringsformatet, kan der opstå fejl, så en eller flere sider bliver blanke. Rigsarkivets testværktøj **Ada** kan undersøge, om der er blanke dokumenter, eller dokumenter hvor første side er blank.

Hvis der er i forbindelse med test af arkiveringsversionen findes flere blanke dokumenter, tolkes dette som en konverteringsfejl. Er der i driftsperioden blanke dokumenter eller dokumenter, hvor første side er blank, skal dette dokumenteres i kontekstdokumentationen, for at arkiveringsversionen kan godkendes. Rigsarkivet anbefaler derfor, at denne test foretages før afleveringen, så eventuelle problemer kan håndteres på forhånd. Hvis der er helt tomme eller blanke dokumenter i driftsversionen på grund af fejljournalisering eller lignende, skal de medtages på listen over fejlede dokumenter.

Beta-version

## 5. Bilag 1: Filer der altid kan kasseres

Mange it-systemer indeholder dokumenter, som ikke har et betydningsbærende indhold med relevans for arkiveringsversionen. Følgende filtyper skal ikke konverteres og indgå i arkiveringsversioner:

- XML-følgesedler og GML-følgesedler
- Arkivfiler (zip, rar o.l.), hvor indholdet er pakket ud, og hver enkelt dokument er registreret med selvstændigt dokument-ID
- VCF-filer eller tilsvarende "visitkort", såfremt oplysningerne om afsender/modtager fremgår af den mail, visitkortet har været vedhæftet på
- Mailsignaturer, -logoer og lignede
- Thumbnails af dokumenter, hvor originalerne indgår i arkiveringsversionen
- Programfiler, eksempelvis exe-, class- eller sys- filer
- Interne/eksterne genvejsfiler (f.eks. til internettet eller brugerens skrivebord)

For samtlige af disse filtyper gælder, at de skal kunne identificeres sikkert i systemet for at måtte kasseres.

Myndigheden skal optælle antallet af filer, der kasseres i henhold til bilag 1, efter filtype og sende opgørelsen til Rigsarkivet.

## 6. Bilag 2: Fejlliste

Listen over dokumenter, der ikke kan konverteres, skal indeholde følgende felter i det omfang, felterne er relevante for systemet:

- journal-/sagsnummer
- saggruppe/facet/emneord
- sagstitel
- aktnummer
- akttitel
- dokumentnummer
- dokumenttitel
- dokumentdato
- filnavn
- filtype/filformat
- sagsbehandler/kontaktperson
- tekst, der kort beskriver fejlen. Hvor der er flere ens fejl, skal teksten være identisk

Filtyper, der er omtalt i Bilag 1, skal ikke indgå i fejllisten.

### Statslige myndigheder

Filer, der kasseres efter aftale med Rigsarkivet, skal ikke indgå i fejllisten.

### Kommunale og regionale myndigheder

Filer, der kasseres i henhold til cirkulæreskrivelse 2014, skal **ikke** indgå i fejllisten, hvis cirkulæreskrivelsen følges i sin helhed.

Filer, der kasseres i henhold til cirkulæreskrivelse 2014, **skal** indgå i fejllisten, hvis der kun kasseres delvis efter cirkulæreskrivelsen.