

Ministerium: Kulturministeriet
Journalnummer: Kulturmin.,
Statens Arkiver, j.nr. 2010-018996

Senere ændringer til forskriften
Ingen

Bekendtgørelse om arkiveringsversioner

I medfør af § 5, stk. 1, § 13, stk. 1, og § 14 i bekendtgørelse nr. 591 af 26. juni 2003 om offentlige arkiver og offentlige arkivers virksomhed og efter drøftelse med de kommunale og regionale parter fastsættes:

Område

§ 1. Bestemmelserne i denne bekendtgørelse gælder for arkivalier skabt af den offentlige forvaltning og domstolene, og som af Statens Arkiver er bestemt til bevaring.

§ 2. Bevaring af data fra it-systemer og af lyd og video skal ske i form af arkiveringsversioner.

Stk. 2. Statens Arkiver kan af bevaringsmæssige grunde fastsætte, at bevaringen af andre arkivalier skal ske i form af arkiveringsversioner.

§ 3. En arkiveringsversion af bevaringsværdige data skal fremstilles efter de anvisninger, der udstedes af rigsarkivaren, jf. bilag 2-8.

Stk. 2. Rigsarkivaren kan anvise anden bevaringsform end en arkiveringsversion, fremstillet efter anvisningerne i bilag 2-8, hvis bevaringsmæssige hensyn tilsiger det.

Produktion og aflevering

§ 4. Produktion og aflevering af arkiveringsversioner af data fra den statslige forvaltning og domstolens it-systemer skal finde sted på tidspunkter, der fastsættes af Statens Arkiver.

§ 5. Produktion og aflevering af arkiveringsversioner af data fra kommunernes og regionernes it-systemer, som indeholder personoplysninger, skal finde sted, inden data skal slettes. Den pågældende kommune eller region kan træffe aftale om tidligere tidspunkter med det modtagende arkiv.

Stk. 2. Produktion af arkiveringsversioner af data fra øvrige it-systemer skal finde sted, inden data slettes, eller når it-systemet tages ud af brug.

Stk. 3. Statens Arkiver kan derudover fastsætte, at der skal produceres en arkiveringsversion af data fra et it-system, når det af bevaringsmæssige hensyn er nødvendigt.

§ 6. Arkiveringsversioner af statslige myndigheders data skal godkendes af Statens Arkiver. Arkiveringsversioner af kommunale og regionale myndigheders data skal godkendes af det modtagende arkiv. Arkiveringsversioner af bevaringsværdige kommunale og regionale data, som ikke er omfattet af afleveringspligt, skal godkendes af samme offentlige arkiv, som modtager myndighedens afleveringspligtige arkiveringsversioner.

Stk. 2. Data, som er overført til en arkiveringsversion, må ikke slettes hos myndigheden, før arkiveringsversionen er godkendt.

Ikrafttrædelse m.v.

§ 7. Bekendtgørelsen træder i kraft den 1. september 2010.

Stk. 2. Bekendtgørelse nr. 342 af 11. marts 2004 om arkiveringsversioner af bevaringsværdige data fra elektroniske arkivsystemer ophæves.

Stk. 3. Bekendtgørelse nr. 302 af 16. april 2009 om aflevering af lyd og video ophæves.

Stk. 4. Til og med 1. juli 2011 kan der indgås aftale med offentligt arkiv om, at arkiveringsversion af data udarbejdes efter reglerne i den i stk. 2 nævnte bekendtgørelse.

Stk. 5. Arkiveringsversioner af data efter reglerne i den i stk. 2 nævnte bekendtgørelse skal være afleveret til det pågældende offentlige arkiv inden 31. december 2011.

Stk. 6. Rigsarkivaren kan efter ansøgning meddele dispensation for reglerne i stk. 4 og 5.

Rigsarkivet, den 20. august 2010

ASBJØRN HELLUM

/ Kirsten Villadsen Kristmar

Bilag 1 Bilagsfortegnelse

Bilag 2 Grafisk oversigt over elementer og struktur i en arkiveringsversion

Bilag 3 Arkiveringsversion af it-systemets data og eventuelle dokumenter

- 3.A. Generelle regler om arkiveringsversioner
- 3.B. Datastruktur
- 3.C. Dataindhold
- 3.D. Information om arkiveringsversionen

Bilag 4 Datastruktur

- 4.A. Generelle regler om datastruktur
- 4.B. Placering af mapper og filer
- 4.C. Mappen *Indices*
- 4.D. Mappen *Tables*
- 4.E. Mappen *ContextDocumentation*
- 4.F. Mappen *Schemas*
- 4.G. Mappen *Documents*

Bilag 5 Dataindhold

- 5.A. Tabelindhold
- 5.B. Datatyper
- 5.C. Konvertering af tabelindhold til digitale dokumenter, lyd, video eller geodata
- 5.D. Tekstformat
- 5.E. Digitale dokumenter
- 5.F. Lyd og video
- 5.G. Geodata
- 5.H. Komprimering
- 5.I. Optimering
- 5.J. Ingen forringelse

Bilag 6 Information om arkiveringsversionen

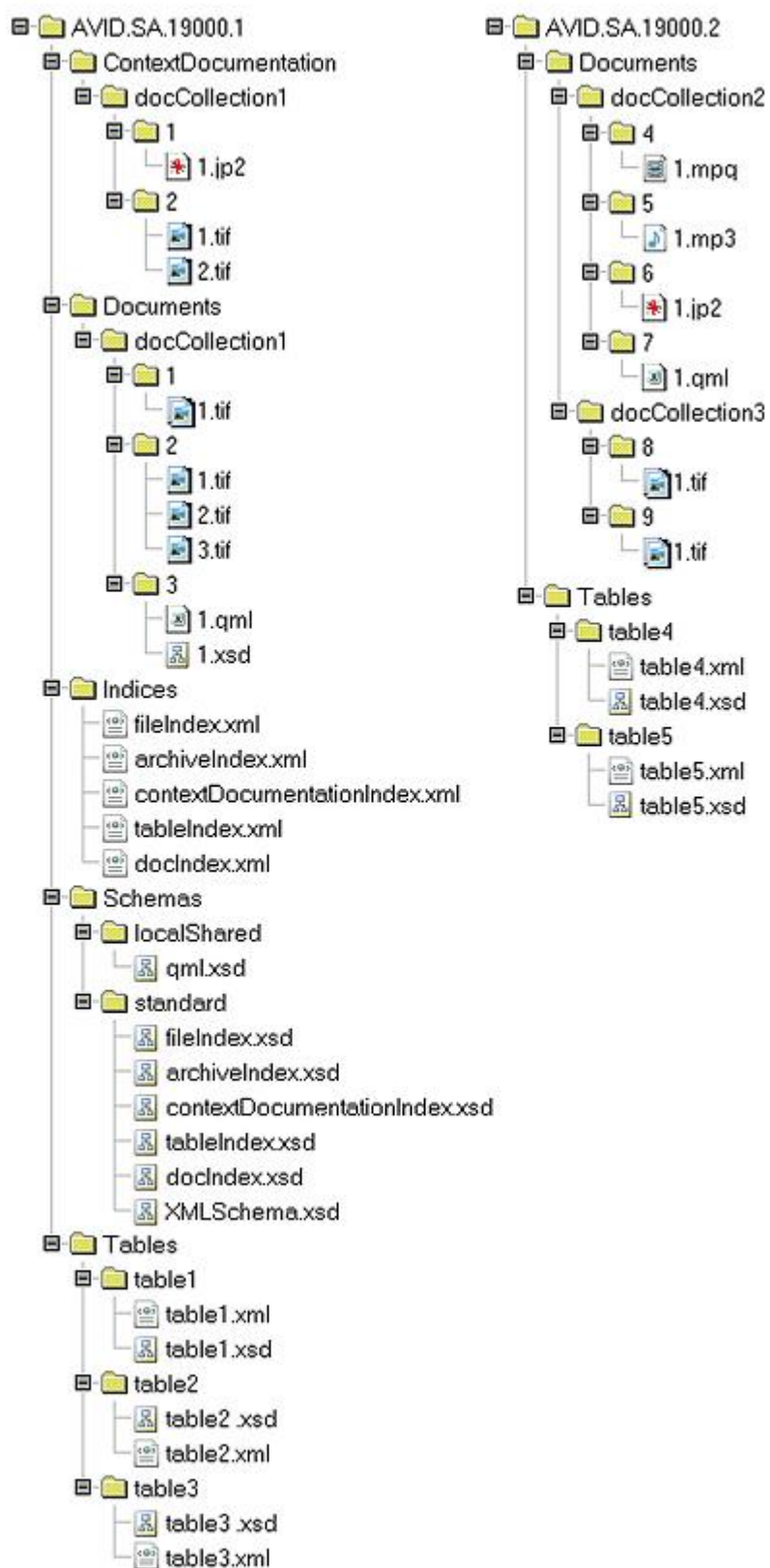
- 6.A. Arkivbeskrivelsesfil
- 6.B. Kontekstdokumentation
- 6.C. Data om arkiveringsversionens tabeller
- 6.D. SQL-forespørgsler

Bilag 7 Afleveringsmedie

Bilag 8 Skemaer

Grafisk oversigt over elementer og struktur i en arkiveringsversion

Figur 2.1



Arkiveringsversion af it-systemets data og eventuelle dokumenter

3.A. Generelle regler om arkiveringsversioner

3.A.1 En arkiveringsversion består af:

- datastruktur, jf. 3.B
- dataindhold, jf. 3.C
- information om arkiveringsversionen, jf. 3.D

3.A.2 En arkiveringsversion må ikke indeholde krypteret information.

3.A.3 En arkiveringsversion skal indeholde samtlige bevaringsværdige data og eventuelle dokumenter fra en afgrænset periode, hvor der ikke længere rettes i eller tilføjes data, eller udformes som et øjebliksbillede og indeholde samtlige bevaringsværdige data og eventuelle dokumenter på et bestemt tidspunkt.

3.A.4 Hvis arkiveringsversionen af et it-system med dokumenter udarbejdes uden skift af journalperiode, eller hvis der i forbindelse med skift af journalperiode sker overførsel af dokumenter til ny periode, skal de dokumenter, der indgår i arkiveringsversionen, markeres på en sådan måde i it-systemet, at de kan undtages fra efterfølgende arkiveringsversioner.

3.B. Datastruktur

3.B.1 Datastrukturen i arkiveringsversionen består af:

- en mappestruktur, jf. 4.B, illustreret i figur 2.1.
- en relationel databasestruktur på 1. normalform eller højere, angivet i indeksfilen tableIndex.xml, jf. 4.C.5.a.
- øvrige indeksfiler i XML, jf. 4.C, der strukturerer data om arkiveringsversionens overordnede indhold, samtlige filer i arkiveringsversionen, dens kontekstdokumentation og dens eventuelle digitale dokumenter.

3.C. Dataindhold

3.C.1 Dataindholdet består af:

- tabelindhold i standardiserede datatyper, jf. 5.B.
- eventuelle digitale dokumenter, lyd, video og geodata i standardiserede dataformater, jf. 5.E - 5.G.

3.D. Information om arkiveringsversionen

3.D.1 Information om arkiveringsversionen består af beskrivelser af administrativ brug, dataindhold og it-system, jf. bilag 6.

Datastruktur

4.A. Generelle regler om datastruktur

4.A.1 I en arkiveringsversion skal alle nøgler have en entydig identifikator. Der må ikke forekomme situationer, hvor det er nødvendigt at uddrage dele af nøglefelter for at forstå it-systemets indhold eller funktion.

4.A.2 Når en værdi i et felt er en kode som repræsentation for en fast og entydig værdi, skal koderne forklares. Såfremt værdien ikke findes i en kode- eller opslagstabel i it-systemet, skal der i arkiveringsversionen genereres én eller flere tabeller med angivelse af værdierne i kodede felter. Alternativt skal kodeværdien i tabeldata ved udtræk til arkiveringsversion udskiftes med det egentlige indhold.

4.A.3 Såfremt dokumenterne i et it-system med dokumenter er lagret i en betydningsbærende struktur, skal denne struktur i arkiveringsversionen omdannes til en eller flere tabeller.

4.B. Placering af mapper og filer

4.B.1 I roden af filsystemet på afleveringsmediet, jf. bilag 7, skal der være placeret en mappe navngivet med mediets navn. Mediets navn består af det unikke arkiveringsversionsID med tilføjelse af et suffiks »n«, der angiver rækkefølgen af medierne, hvor n er et fortløbende medieløbenummer begyndende med 1.

4.B.2 Arkiveringsversionens indhold fordeles i mapper, som angivet i figur 4.1.

Figur 4.1

Navn på mappe	Beskrivelse	Obligatorisk	Placering
<i>Indices</i>	Indeksfiler, jf. 4.C	Ja	Altid på første medie
<i>Tables</i>	Tabelindhold, jf. 4.D	Ja	Kan placeres frit og fordeles over flere medier
<i>ContextDocumentation</i>	Kontekstdokumentation, jf. 4.E	Ja	Altid på første medie
<i>Schemas</i>	XML-skemaer, jf. 4.F. Skemamappen skal have to undermapper: <i>standard</i> og <i>localShared</i>	Ja	Altid på første medie
<i>Documents</i>	Eventuelle dokumenter, lyd, video og geodata, jf. 4.G	Nej	Kan placeres frit og fordeles over flere medier

4.B.3 Mapperne skal navngives som angivet i figur 4.1.

4.B.4.a Et arkiveringsversionsID består af præfikset AVID, en kode på 2-4 bogstaver, som angiver det modtagende arkiv, samt et arkiveringsversionsløbenummer. Elementerne adskilles med punktum.

4.B.4.b ArkiveringsversionsID udleveres af Statens Arkiver.

4.B.5.a En arkiveringsversion, som ikke kan være på ét medie, jf. bilag 7, kan fordeles over flere medier.

4.B.5.b Mapperne *ContextDocumentation*, *Indices* og *Schemas* skal altid være placeret på afleveringens første medie.

4.B.5.c Mappen i roden på efterfølgende medier skal kun indeholde de mapper, hvis indhold kræver fordeling over flere medier. Disse mapper navngives altid som angivet i figur 4.1 uden brug af suffiks.

4.B.6 Der må kun være én af hver af de i figur 4.1 angivne mapper på hvert medie.

4.C. Mappen *Indices*

4.C.1.a Mappen *Indices* skal indeholde følgende indeksfiler med oplysninger om arkiveringsversionen og dens indhold:

- fileIndex.xml
- archiveIndex.xml
- contextDocumentationIndex.xml
- tableIndex.xml

4.C.1.b Hvis arkiveringsversionen indeholder digitale dokumenter, lyd, video eller geodata, skal mappen *Indices* endvidere indeholde følgende indeksfil:

- docIndex.xml

4.C.1.c Alle indeksfiler skal overholde deres tilhørende skema, jf. bilag 8.

4.C.2.a **fileIndex.xml** skal indeholde en komplet liste over samtlige filer, der findes i arkiveringsversionen. fileIndex.xml er dog undtaget fra denne regel.

4.C.2.b For hver enkelt fil i arkiveringsversionen angives de oplysninger, som fremgår af figur 4.2.

Figur 4.2

Elementnavn	Betegnelse	Beskrivelse	Forekomst pr fil	Obligatorisk
foN	Sti	Angivelse af sti til den mappe i arkiveringsversionen, hvor filen findes	1	Ja
fiN	Navn	Filens navn	1	Ja
md5	Kontrolsum	Filens kontrolsum af typen MD5 iht. <i>IETF RFC1321 - The MD5 Message-Digest Algorithm</i> : 128 bit (16 bytes) repræsenteret som 32 hexadecimale cifre, alle angivet med <i>enten</i> minuskler <i>eller</i> versaler (små eller store bogstaver). Krav om 32 hexadecimale cifre medfører således krav om, at der afhængigt af værdien anvendes foranstillede nuller	1	Ja

4.C.3 **archiveIndex.xml** skal indeholde de oplysninger, som fremgår af 6.A.

4.C.4.a **contextDocumentationIndex.xml** skal indeholde et indeks over de dokumenter, som findes i arkiveringsversionens kontekstdokumentation.

4.C.4.b For hvert dokument i kontekstdokumentationen angives de oplysninger, som fremgår af figur 4.3.

Figur 4.3

Elementnavn	Betegnelse	Udfaldsrum	Forekomst pr. dokument	Obligatorisk
documentID	DokumentID	ID på op til 12 cifre	1	Ja
documentTitle	Dokumenttitel	Fritekst	1	Ja
documentDescription	Dokumentbeskrivelse	Fritekst	0-1	Nej
documentDate	Dato	År, år-måned, år-måned-dag, iht. XML-standarden	0-1	Nej
authorName	Forfatternavn	Fritekst	0-m	Nej
authorInstitution	Forfatterinstitution	Fritekst	0-m	Nej
documentCategory	Dokumentkategori	Kategori iht. bilag 6, figur 6.2	1-m	Ja

4.C.5.a **tableIndex.xml** skal indeholde en angivelse af en relationel databasestruktur på 1. normalform eller højere. Samtlige tabeller i arkiveringsversionen skal angives.

4.C.5.b »tableIndex.xml« skal overholde det generelle XML-skema »tableIndex.xsd«, jf. 4.F.

4.C.5.c Hvis et felt må have værdien NULL, skal der i »tableIndex.xml« være angivet værdien »true« for kolonnens tilhørende element »nullable«.

4.C.6.a **docIndex.xml** skal danne forbindelsen mellem hvert dokument og dets placering. »docIndex.xml« skal desuden indeholde oplysninger om dokumenternes oprindelige filnavne, filtype i arkiveringsversionen samt eventuelle overordnede dokumenter. »docIndex.xml« skal ikke indeholde oplysninger om dokumenterne i kontekstdokumentationen.

4.C.6.b For hvert enkelt dokument i docIndex.xml angives de oplysninger, som fremgår af figur 4.4.

Figur 4.4

Elementnavn	Betegnelse	Beskrivelse	Udfaldsrum	Forekomst pr. dokument	Obligatorisk
dID	DocID	ID, som entydigt udpeger det enkelte dokument	ID på op til 12 cifre	1	Ja
pID	Parent	ID på overordnet dokument	DocID	0-1	Nej
mID	MediaID	Navn på det lagringsmedie, som dokumentet ligger på	Jf. 4.B.1	1	Ja
dCf	SubFolder	Dokumentsamlingsmappe, som dokumentet ligger i	Jf. 4.G	1	Ja
oFn	OrigFilename	Angivelse af filens navn (inkl. ekstension) i it-systemet	Fritekst	1	Ja
aFt	Archival-Format	Angivelse af filens format i arkiveringsversionen	Ekstension, jf. 4.G.8	1	Ja
gmlXsd	Schema	For GML filer angives reference til det skema, der skal	Fritekst	0-1	Nej

		bruges til validering af den pågældende GML-fil			
--	--	-------------------------------------------------	--	--	--

4.D. Mappen *Tables*

4.D.1 Mappen *Tables* skal indeholde én mappe for hver tabel i arkiveringsversionen.

4.D.2.a Mappen for en tabel navngives »table[fortløbende nummer]«.

4.D.2.b Den fortløbende nummerering begynder med 1. Foranstillede nuller må ikke anvendes.

4.D.3 Mappen for hver tabel skal indeholde to filer:

- table[fortløbende nummer].xsd
- table[fortløbende nummer].xml

4.D.4 »table[fortløbende nummer].xsd« er et XML-skema, der kun angiver strukturen for den enkelte pågældende tabel, og skal være i overensstemmelse med XML-instansen »tableIndex.xml«, jf. 4.C.5.a, der angiver strukturen for hele den pågældende relationelle database, herunder samtlige tabeller.

4.D.5 »table[fortløbende nummer].xml« er en XML-instans, der indeholder data for den pågældende tabel, og dens struktur skal være i overensstemmelse med det tilhørende XML-skema, »table[fortløbende nummer].xsd«.

4.D.6 Hvis et felt i en tabel kan have værdien NULL, skal den pågældende kolonne i det tilhørende skema (»table[fortløbende nummer].xsd«) indeholde attributten nillable=»true«. Tilsvarende skal XML-instansens (»table[fortløbende nummer].xml«) element indeholde attributten xsi:nil=»true« jf. W3C standarden for håndtering af NULL værdier i XML.

4.E. Mappen *ContextDocumentation*

4.E.1 Mappen *ContextDocumentation* skal indeholde én eller flere dokumentsamlingsmapper med kontekstdokumentation, jf. 6.B.

4.E.2 En dokumentsamlingsmappe med kontekstdokumentation må indeholde op til 10.000 dokumentmapper.

4.E.3 Dokumentsamlingsmapperne navngives »docCollection[fortløbende nummer]«, begyndende med 1. Navnet skal være unikt inden for *ContextDocumentation*.

4.E.4 Hvert dokument i kontekstdokumentationen skal tildeles et ID på op til 12 cifre. Dokumentets ID skal være unikt inden for *ContextDocumentation*.

4.E.5 En dokumentmappe skal indeholde ét dokument, som består af én eller flere filer af samme format, og navngives med dokumentets ID. Foranstillede nuller må ikke anvendes.

4.E.6 Et dokumentets fil (eller filer) navngives fortløbende med et nummer, begyndende med 1 samt formatets ekstension, jf. 4.G.8

4.F. Mappen *Schemas*

4.F.1 Mappen *Schemas* skal være opdelt i undermapperne *standard* og *localShared*.

4.F.2 Mappen *standard* skal indeholde skemaer for arkiveringsversionens indeksfiler, jf. bilag 8, samt W3C standard XML-skema, jf. <http://www.w3.org/2001/XMLSchema.xsd>.

4.F.3 For skemaerne *fileIndex.xsd*, *archiveIndex.xsd*, *contextDocumentationIndex.xsd*, *tableIndex.xsd*, *docIndex.xsd* samt W3Cs standard XML-skema gælder, at der altid skal anvendes de skemaer, som Statens Arkiver stiller til rådighed. Skemaerne og deres navngivning må ikke ændres i arkiveringsversionen.

4.F.4 Mappen *localShared* skal indeholde eventuelle GML-skemaer, som ikke er placeret sammen med det pågældende GML-dokument, jf. 4.G.7.a.

4.G. Mappen *Documents*

4.G.1 Mappen *Documents* skal indeholde én eller flere dokumentsamlingsmapper, dog maksimalt 10.000.

4.G.2 Dokumentsamlingsmapperne navngives »docCollection[fortløbende nummer]«, begyndende med 1. Navnet skal være unikt inden for *Documents*.

4.G.3 En dokumentsamlingsmappe må indeholde op til 10.000 dokumentmapper.

4.G.4 Hvert dokument skal i arkiveringsversionen tildeles et ID på op til 12 cifre. Dokumentets ID skal være unikt inden for *Documents*.

4.G.5 En dokumentmappe skal indeholde ét dokument, som består af én eller flere filer af samme format, og navngives med dokumentets ID. Foranstillede nuller må ikke anvendes.

4.G.6 Et dokument fil (eller filer) navngives fortløbende med et nummer, begyndende med 1 samt formatets ekstension. Foranstillede nuller må ikke anvendes.

4.G.7 For GML-filer lagres det relevante skema i samme mappe som GML-filen, og navngives med fortløbende nummer efterfulgt af .xsd, jf. dog 4.G.7.a. Foranstillede nuller må ikke anvendes.

4.G.7.a GML-skemaer kan alternativt lagres i den skema-mappe, som navngives *localShared*, jf. 4.F. GML-skemaer i mappen *localShared* navngives »localSchema[fortløbende nummer]«, begyndende med 1.

4.G.8 Anvendelse af ekstensions

4.G.8.a Dokumenter i formatet TIFF skal have ekstension tif.

4.G.8.b Dokumenter i formatet MP3 skal have ekstension mp3.

4.G.8.c Dokumenter i formatet MPEG-2 og MPEG-4 skal have ekstension mpg.

4.G.8.d Dokumenter i formatet JPEG-2000 skal have ekstension jp2.

4.G.8.e Dokumenter i formatet GML skal have ekstension gml.

4.G.8.f Dokumenter i formatet WAVE skal have ekstension wav.

4.G.9 Muligheden for at fremsøge sagligt sammenhørende dokumenter skal overføres til arkiveringsversionen efter det modtagende arkivs nærmere anvisning.

Dataindhold

Arkiveringsversionens dataindhold består af tabelindhold i standardiserede datatyper og af eventuelle digitale dokumenter, lyd, video og geodata i standardiserede formater.

5.A. Tabelindhold

5.A.1.a I overensstemmelse med den tabelstruktur, der er defineret for hver tabel i hvert sit tilhørende XML-skema, navngivet »table[fortløbende nummer].xsd«, jf. 4.D, skal hver tabel findes i en XML-instans navngivet »table[fortløbende nummer].xml«.

5.A.1.b Den fortløbende nummerering begynder med 1. Foranstillede nuller må ikke anvendes.

5.A.2 Indholdet af de enkelte felter skal renses for eventuelle foran- og efterstillede blanktegn.

5.B. Datatyper

5.B.1.a De standardiserede datatyper, som skal anvendes for tabelindhold, er angivet i figur 5.1. De er et uddrag af datatyper fra standarden SQL:1999 repræsenteret som datatyper i W3C XML Schema Language 1.0.

5.B.1.b Det er datatypen i W3C XML Schema Language 1.0, som skal anvendes.

Oversættelsen fra datatyper i SQL:1999 er angivet for at vise, hvorledes oversættelsen til datatyper i W3C XML Schema Language 1.0 skal foretages.

Figur 5.1

Datatype	SQL:1999	XML-datatype
Tekst	CHAR, CHARACTER, NCHAR, NATIONAL CHAR, NATIONAL CHARACTER, VARCHAR, CHARACTER VARYING, NATIONAL CHARACTER VARYING	string
Heltal	INTEGER, SMALL INTEGER	integer
Decimaltal	NUMERIC, DECIMAL, FLOAT, DOUBLE PRECISION, REAL	decimal
Boolsk	BOOLEAN	boolean
Hexadecimal	CHAR, CHARACTER, NCHAR, NATIONAL CHAR, NATIONAL CHARACTER, VARCHAR, CHARACTER VARYING, NATIONAL CHARACTER VARYING	hexBinary
Dato	DATE	date
Tidspunkt	TIME[WITH TIME ZONE]	time
Tidsstempel	TIMESTAMP[WITH TIME ZONE]	dateTime
Tidsperiode	INTERVAL	duration

5.B.2 Datatypen *string* må kun indeholde ikke-opmærket tekst, som umiddelbart kan fortolkes.

5.B.3 Datatypen *boolean* kan jf. W3C kun antage værdierne 1; 0 eller *true*; *false*.

5.B.4 Datatyperne *date*, *time* og *dateTime* kan bruges med eller uden *Time Zone*.

5.C. Konvertering af tabelindhold til digitale dokumenter, lyd, video eller geodata

5.C.1 Tabelindhold skal overholde de angivne datatyper, jf. 5.B. Det følger heraf, at dataindhold i tabelform fra et it-system, som skal overføres til en arkiveringsversion og som ikke umiddelbart kan overholde dette krav, skal have sit dataindhold konverteret således:

5.C.1.a til digitale dokumenter, lyd, video eller geodata, idet indholdet konverteres til de formater, som fremgår af 5.E - 5.G.

5.C.1.b til tabelindhold af datatypen *string*, jf. 5.B, idet andet indhold end det for datatypen tilladte slettes.

5.C.2 Det modtagende arkiv anviser, om et givet indhold skal håndteres iht. 5.C.1.a eller 5.C.1.b.

5.D. Tekstformat

5.D.1 Unicode indkodning

5.D.1.a Data i arkiveringsversionens indeksfiler og tabelindhold skal være indkodet som *well-formed* UTF-8, som angivet i ISO/IEC 10646:2003 Annex D og som beskrevet i *The Unicode Standard 5.1*, kapitel 3.

5.D.1.b De indkodede karakterer skal være gyldige Unicode skalarværdier. Surrogater/RC-elementer og Unicode *noncharacters* må ikke anvendes.

5.D.1.c Ingen karakterer i *Private Use Area* må indkodes.

5.D.1.d Styrekodetegnene fra og med #x00 til og med #x1F er ikke tilladt med undtagelse af styrekodetegnene tabulator (TAB) #x09, nylinje (LF) #x0A og vognretur (CR) #x0D.

5.D.2 XML indkodning

5.D.2.a I henhold til XML-standarden gælder de regler for angivelse af de pågældende tegn, som fremgår af figur 5.2.

Figur 5.2

Tegn	Entitet	Anvendelse
<	<	Obligatorisk
>	>	Anbefalet
&	&	Obligatorisk
'	'	Anbefalet
"	"	Anbefalet

5.D.2.b Tegnene fra og med #x7F til og med #x9F skal angives med deres decimale eller hexadecimale tegnreference.

5.D.2.c CDATA sektioner <![CDATA]]> må ikke anvendes.

5.E. Digitale dokumenter

5.E.1.a Et digitalt dokument, jf. dog 5.F og 5.G, skal lagres i ét af følgende formater:

- det grafiske bitmapformat TIFF, version 6.0 baseline.
- JPEG-2000 efter standarden ISO/IEC 15444-1:2004. Information technology - JPEG 2000 image coding system - Part 1: Core coding system.

5.E.1.b Det er tilladt at benytte begge formater inden for samme arkiveringsversion.

5.E.2 Dokumenter i TIFF skal komprimeres efter følgende kompressionsregler:

5.E.2.a Sort/hvide dokumenter skal komprimeres med CCITT/TSS grp. 3, grp. 4, PackBit eller LZW.

5.E.2.b Dokumenter med gråtoner eller farver skal komprimeres med PackBit eller LZW.

5.F. Lyd og video

5.F.1 Lydfiler skal lagres efter standarden MP3 DS/EN ISO/IEC 11172-3.

5.F.2 Det modtagende arkiv kan tillade, at lydfiler afleveres i formatet WAVE LPCM som specificeret i *Multimedia Programming Interface and Data Specifications 1.0. IBM Corporation and Microsoft Corporation, August 1991*. Dog begrænset til bitdybder, der er hele multipla af 8.

5.F.3 Videofiler skal lagres efter en af følgende standarder:

- MPEG-2 DS/EN ISO/IEC 13818-2. Eventuel lyd indkodes som MP3, som specificeret i ISO/IEC 13818-3.
- MPEG-4 AVC DS/EN ISO/IEC 14496-10 (ITU-T H.264). Video indkodes som specificeret i ISO/IEC 14496-10. Eventuel lyd indkodes som AAC, som specificeret i ISO/IEC 14496-3. Video og lyd indpakkes i MPEG-4 formatet som defineret i ISO/IEC 14496-14.

5.G. Geodata

5.G.1 Data fra geografiske informationssystemer samt øvrige geodata lagres som GML-filer iht. GML 3v3.DK, den danske profil af GML-standarden ISO 19136.

5.G.2 GML-filer, som er større end 1 GB, opdeles i enheder efter det modtagende arkivs anvisning.

5.G.3 Hver GML-fil, evt. efter opdeling jf. 5.G.2, behandles som et separat dokument iht. reglerne i 4.G.

5.G.4 De nødvendige XML-skemaer til validering af GML-filerne skal indgå i arkiveringsversionen.

5.H. Komprimering

5.H.1 En arkiveringsversion må ikke være komprimeret udover den kompression, som er anvist for eller følger af de dokumentformater, der kræves til brug for arkiveringsversionen, jf. 5.E og 5.F.

5.I. Optimering

5.I.1 Det modtagende arkiv kan i konkrete tilfælde træffe afgørelse om nedbringelse af dokumenters pladsforbrug, f.eks. ved at anvise anvendelse af optimal bitdybde.

5.J. Ingen forringelse

5.J.1 I forbindelse med generering af arkiveringsversionen må der ikke ske en kvalitetsmæssig forringelse af dokumenterne, herunder lyd og video, ud over hvad der måtte være en konsekvens af det krævede format til brug for arkiveringsversionen eller de anvisninger, som i øvrigt gives af det modtagende arkiv.

Information om arkiveringsversionen

Informationen om arkiveringsversionen består af følgende elementer:

- Arkivbeskrivelsesfil, jf. 6.A.
- Kontekstdokumentation, jf. 6.B.
- Data om arkiveringsversionens tabeller (tabelindeks), jf. 6.C.
- Evt. SQL-forespørgsler, jf. 6.D.

6.A. Arkivbeskrivelsesfil

6.A.1 Enhver arkiveringsversion skal indeholde en arkivbeskrivelsesfil med angivelse af oplysninger iht. figur 6.1.

6.A.2 Arkivbeskrivelsesfilen navngives archiveIndex.xml og skal overholde det tilhørende skema, jf. bilag 8.

6.A.3 Indholdet af arkivbeskrivelsesfilen fastlægges efter drøftelse mellem den afleverende myndighed og det modtagende arkiv.

Figur 6.1

Elementnavn	Betegnelse	Beskrivelse	Udfaldsrum	Forekomst	Obligatorisk
archiveInformationPackageID	ArkiveringsversionsID	Entydigt ID som tildeles arkiveringsversionen af Statens Arkiver	ArkiveringsversionsID, som defineret i denne bekendtgørelse	1	Ja
archiveInformationPackageIDPrevious	Tidligere aflevering	Entydigt arkiveringsversionsID på tidligere aflevering fra samme it-system	ArkiveringsversionsID, som defineret i denne eller tidligere bekendtgørelse	0-1	Nej
archivePeriodStart	Arkiveringsversion startdato	Startdato for de afleverede data	År, år-måned, eller år-måned-dag	1	Ja
archivePeriodEnd	Arkiveringsversion slutdato	Slutdato for de afleverede data	År, år-måned, eller år-måned-dag	1	Ja
archiveInformationPacketType	Slutaflevering	Angivelse af, om arkiveringsversionen er en slutaflevering	Boolsk værdi	1	Ja
creatorName	Arkivskaber	Den eller de myndigheder, der har skabt arkivaliet. Kan også være organisatoriske enheder inden for en myndighed	Fritekst	1-m	Ja

Figur 6.1

Elementnavn	Betegnelse	Beskrivelse	Udfaldsrum	Forekomst	Obligatorisk
		eller andre bidragydere til arkivaliet			
creationPeriod-Start	Arkivskaber startdato	For hver arkivskaber angives, hvornår arkivskaber er begyndt at tilføje data til it-systemet	År, år-måned, eller år-måned-dag	1-m	Ja
creationPeriod-End	Arkivskaber slutdato	For hver arkivskaber angives, hvornår arkivskaber er ophørt med at tilføje data til it-systemet	År, år-måned, eller år-måned-dag	1-m	Ja
archiveType	Arkiveringsversionsperiodetype	Angivelse af, om der er tale om en afsluttet arkivperiode/årgang, eller om der er tale om et øjebliksbillede	Boolsk værdi	1	Ja
systemName	Systemnavn	Det officielle navn på it-systemet, hvor alle forkortelser er opløst	Fritekst	1	Ja
alternativeName	Alternativt systemnavn	Alternative navne på it-systemet	Fritekst	0-m	Nej
systemPurpose	Systemformål	Beskrivelse af formålet med at oprette og drive it-systemet	Fritekst	1	Ja
systemContent	Systemindhold	It-systemets centrale population og centrale variabler	Fritekst	1	Ja
regionNum	Regionsnumre	Angivelse af, om der i systematisk form er registreret regionsnumre i it-systemet	Boolsk værdi	1	Ja
komNum	Kommunenumre	Angivelse af, om der i systematisk form er registreret kommune-numre i it-systemet	Boolsk værdi	1	Ja
cprNum	CPR-numre	Angivelse af, om der i systematisk form er registreret CPR-numre i it-systemet	Boolsk værdi	1	Ja
cvrNum	CVR-numre	Angivelse af, om der i systematisk form er re-	Boolsk værdi	1	Ja

Figur 6.1

Elementnavn	Betegnelsen	Beskrivelse	Udfaldsrum	Forekomst	Obligatorisk
		gistreret CVR-numre i it-systemet			
matrikNum	Matrikelnumre	Angivelse af, om der i systematisk form er registreret matrikelnumre i it-systemet	Boolsk værdi	1	Ja
bbrNum	BBR-numre	Angivelse af, om der i systematisk form er registreret BBR-numre i it-systemet	Boolsk værdi	1	Ja
whoSygKod	WHO-sygdomskoder	Angivelse af, om der i systematisk form er registreret WHO-sygdomskoder i it-systemet	Boolsk værdi	1	Ja
sourceName	Datakilde	Andre it-systemer som har leveret data til it-systemet (opslag, overførsel, samkøring osv.)	Fritekst	0-m	Nej
userName	Databrugere	Andre it-systemer som har brugt data fra it-systemet (opslag, overførsel, samkøring osv.)	Fritekst	0-m	Nej
predecessorName	Forgængersystemer	Systemer, der tidligere har varetaget samme funktion	Fritekst	0-m	Nej
formVersion	FORM-version	Angivelse af, hvilken version af FORM (den fællesoffentlige forretningsreferencemodel) nedenstående kategorisering er hentet fra	Gyldige FORM-versionsnumre	1	Obligatorisk for it-systemer fra offentlige myndigheder
formClass	FORM-klassifikation	Kategorisering af arkiveringsversionen iht. FORM	Enhver gyldig FORM-referencen iht. den angivne version af FORM	1-m	Obligatorisk for it-systemer fra offentlige myndigheder

Figur 6.1

Elementnavn	Betegnelse	Beskrivelse	Udfaldsrum	Forekomst	Obligatorisk
formClassText	FORM-klassifikation klartekst	FORM-klassifikation i klartekst	Fritekst	1-m	Obligatorisk for it-systemer fra offentlige myndigheder
containsDigitalDocuments	Digitale dokumenter indeholdt	Angivelse af, om der er digitale dokumenter i arkiveringsversionen, udover kontekstdokumentation	Boolsk værdi	1	Ja
searchRelatedOtherRecords	Søgemiddel til andre sager eller dokumenter	Angivelse af, om arkiveringsversionen er et nødvendigt søgemiddel til papirsager/dokumenter eller sager/dokumenter i et andet it-system	Boolsk værdi	1	Ja
relatedRecordsName	Adresserede arkivalier	Henvisning til de arkivalier, som arkiveringsversionen er søgemiddel til	Fritekst	0-m	Ja, hvis der er svaret ja på foregående spørgsmål
systemFileConcept	Eksistens af sagsbegreb i it-systemet	Angivelse af, at it-systemet har et sagsbegreb, forstået som en registreret sammenhæng mellem sagligt sammenhørende dokumenter	Boolsk værdi. Kan kun være <i>true</i> , hvis der er digitale dokumenter i it-systemet	1	Ja
multipleDataCollection	SOA arkitektur	Angivelse af, om it-systemet er sammensat af data og eventuelt dokumenter fra flere forskellige it-systemer i en serviceorienteret arkitektur	Boolsk værdi	1	Ja

Figur 6.1

Elementnavn	Betegnelse	Beskrivelse	Udfaldsrum	Forekomst	Obligatorisk
personalDataRestrictedInfo	Persondata	Angivelse af, om der i arkiveringsversionen findes følsomme personoplysninger iht. persondataloven	Boolsk værdi	1	Ja
otherAccessTypeRestrictions	Fristforlængende oplysninger	Angivelse af, om der i arkiveringsversionen findes oplysninger, der kan betinge længere tilgængelighedsfrist i øvrigt	Boolsk værdi	1	Ja
archiveApproval	Godkendelsesarkiv	Angivelse af, hvilket offentligt arkiv, der godkender arkiveringsversionen	Identifikation af det pågældende arkiv (2-4 tegn)	1	Ja
archiveRestrictions	Tilgængelighedsbegrænsninger	Angivelse af nærmere bestemmelser for adgang til materialet. Elementet kan anvendes efter det modtagende arkivs nærmere retningslinjer	Fritekst	0-1	Nej

6.B. Kontekstdokumentation

6.B.1 Enhver arkiveringsversion skal indeholde dokumenter, som dokumenterer it-systemets administrative funktion samt struktur og funktionalitet.

6.B.2 Det modtagende arkiv fastlægger efter drøftelse med den afleverende myndighed, hvilke dokumenter der skal afleveres, herunder hvilke punkter i figur 6.2, som ikke er relevante at dokumentere i den konkrete aflevering.

6.B.3.a Dokumenterne indplaceres i én eller flere af de kategorier, som fremgår af figur 6.2.

6.B.3.b Oplysninger om kategoriseringen registreres i indeksfilen contextDocumentationIndex, jf. 4.C.4.a.

6.B.4 Dokumenterne skal lagres i ét af de i arkiveringsversionen tilladte dokumentformater, jf. 5.E - 5.F.

Figur 6.2

1. Dokumentation vedrørende administrativ brug af it-systemet		
	Elementnavn	Beskrivelse
1.a	systemPurpose	It-systemets formål
1.b	systemRegulations	It-systemets lov- og regelgrundlag
1.c	systemContent	It-systemets indhold, population og særlige begreber

1.d	systemAdministrativeFunctions	It-systemets administrative funktioner
1.e	systemPresentationStructure	It-systemets præsenteringsstruktur
1.f	systemDataProvision	Tilvejebringelse af data
1.g	systemDataTransfer	Videregivelse af data
1.h	systemPreviousSubsequentFunctions	Data og funktioner fælles med forgænger- og efterfølgersystemer
1.i	systemAgencyQualityControl	Myndighedens egen kvalitetskontrol
1.j	systemPublication	Publikation af og om data
1.k	systemInformationOther	Andet
2. Dokumentation vedrørende it-systemets tekniske udformning, drift og udvikling		
	Elementnavn	Beskrivelse
2.a	operationalSystemInformation	Driftsversionens opbygning
2.b	operationalSystemConvertedInformation	Konvertering hos myndigheden
2.c	operationalSystemSOA	Dokumentation af sammensætning af data og eventuelle dokumenter fra flere forskellige it-systemer i en serviceorienteret arkitektur
2.d	operationalSystemInformationOther	Andet
3. Dokumentation vedrørende arkivskabers aflevering af data		
	Elementnavn	Beskrivelse
3.a	archivalProvisions	Arkivets bestemmelser, herunder afleveringsbestemmelse
3.b	archivalTransformationInformation	Dokumentation af konvertering fra driftsversion til arkiveringsversion
3.c	archivalInformationOther	Andet
4. Dokumentation vedrørende arkivets modtagelse af data (udfyldes af modtagende arkiv)		
	Elementnavn	Beskrivelse
4.a	archivistNotes	Arkivarnoter
4.b	archivalTestNotes	Testnoter
4.c	archivalInformationOther	Andet
5. Dokumentation vedrørende arkivets bevaring af arkiveringsversionen (udfyldes af det modtagende arkiv)		
	Elementnavn	Beskrivelse
5.a	archivalMigrationInformation	Konvertering hos arkivet
5.b	archivalInformationOther	Andet
6. Anden dokumentation		
	Elementnavn	Beskrivelse
6.a	informationOther	Andet

6.C. Data om arkiveringsversionens tabeller

6.C.1 En arkiveringsversion skal indeholde dokumentation af arkiveringsversionens tabeller og relationer (tabelindeks). Tabelindekset skal indeholde de oplysninger, som fremgår af figur 6.3 nedenfor.

Figur 6.3

1. Oplysninger om databasen				
	Elementnavn	Beskrivelse	Forekomst pr. database	Obligatorisk
1.a	version	Format version, altid "1.0"	1	Ja
1.b	dbName	Kort navn på databasen	0-1	Ja, hvis oplysningen findes i databasen
1.c	databaseProduct	Navn på og version af det databaseprodukt, hvorfra arkivering af data er sket	0-1	Ja, hvis oplysningen findes i databasen
2. Oplysninger om tabeller og views				
	Elementnavn	Beskrivelse	Forekomst pr. arkiveringsversion	Obligatorisk
2.a	tables	Liste over tabeller i databasen	1	Ja
2.b	views	Liste over views i databasen	0-1	Ja, hvis der indgår views i arkiveringsversionen
3. Oplysninger om de enkelte tabeller				
	Elementnavn	Beskrivelse	Forekomst pr. tabel	Obligatorisk
3.a	name	Tabelnavn. Må ikke begynde med et tal.	1	Ja
3.b	folder	Navn på mappen som indeholder tabel og tilhørende skema	1	Ja
3.c	description	Beskrivelse af tabellens indhold og betydning	1	Ja
3.d	columns	Liste over kolonner i tabellen	1-m	Ja
3.e	primaryKey	Tabellens primærnøgle	1	Ja
3.f	foreignKey	Fremmednøgle i tabel	0-m	Ja, hvis der findes fremmednøgler
3.g	rows	Angivelse af antal rækker i tabellen	1	Ja
4. Oplysninger om kolonner				
	Elementnavn	Beskrivelse	Forekomst pr. kolonne	Obligatorisk
4.a	name	Kolonnenavn	1	Ja
4.b	columnID	Entydig identifikation af kolonne. Begynder med bogstavet 'c' efterfulgt af nummeret på kolonnens plads i tabellen.	1	Ja
4.c	type	SQL:1999 datatype	1	Ja

4.d	typeOriginal	Original datatype	0-1	Ja, hvis oplysningen findes i databasen
4.e	defaultValue	Defaultværdi for felter i kolonnen	0-1	Ja, hvis oplysningen findes i databasen
4.f	nullable	Angivelse af, om felter i kolonnen kan være "NULL". Boolsk værdi.	1	Ja
4.g	description	Beskrivelse af kolonnens indhold	1	Ja
4.h	functionalDescription	Angivelse af kolonnens funktion iht. figur 6.4, figur 6.5 og figur 6.6	0-m	Ja, hvis kolonnen har en af de angivne funktioner

5. Oplysninger om primærnøgler

	Elementnavn	Beskrivelse	Forekomst pr. nøgle i tabel	Obligatorisk
5.a	name	Navn på primærnøgle. Navnet skal være unikt inden for arkiveringsversionen og være i overensstemmelse med reglerne for <i>constraint names</i> i standarden SQL1999	1	Ja
5.b	column	Oprindeligt navn på kolonner i primærnøglen	1-m	Ja

6. Oplysninger om fremmednøgler

	Elementnavn	Beskrivelse	Forekomst pr. nøgle	Obligatorisk
6.a	name	Navn på fremmednøgle. Navnet skal være unikt inden for arkiveringsversionen og leve op til reglerne for <i>constraint names</i> i standarden SQL:1999	1	Ja
6.b	referencedTable	Den tabel, som fremmednøglen refererer til	1	Ja
6.c	reference	Oprindeligt navn på de kolonner, der indgår i fremmednøglen	1-m	Ja
6.d	referenced	Oprindeligt navn på de kolonner, fremmednøglen refererer til	1-m	Ja

7. Oplysninger om views og forespørgsler

	Elementnavn	Beskrivelse	Forekomst pr. forespørgsel eller view	Obligatorisk
--	-------------	-------------	---------------------------------------	--------------

7.a	name	Navn på view	1	Ja
7.b	queryOriginal	Original SQL-forespørgsel, som definerer view eller SQL-forespørgsel, som er dannet specifikt til arkiveringsversionen	1	Ja
7.c	description	Indhold og betydning af view eller SQL-forespørgsel	0-1	Ja, hvis de kan trækkes fra databasen eller hvis der er tale om forespørgsler, der er dannet specifikt til arkiveringsversionen

6.C.2 Systemviews medtages ikke.

6.C.3 Det modtagende arkiv kan anvise, at de væsentligste views skal have en beskrivelse, jf. figur 6.3, 7.c.

6.C.4 Der skal opmærkes særlige oplysninger for it-systemer med registrering af oplysninger om dokumenter.

6.C.5 For arkiveringsversioner af it-systemer, som nævnt i 6.C.4, skal kolonner, der indeholder særlige informationer, jf. figur 6.4 - figur 6.6, identificeres via elementet functionalDescription, jf. figur 6.3, 4.h.

6.C.6 De særlige oplysninger, der er angivet i figur 6.6, skal opmærkes i det omfang, de er registreret i it-systemet. Hvis oplysningerne i figur 6.6 ikke findes registreret i it-systemet, skal der i forbindelse med afleveringen aftales eventuelle alternative opmærkninger, der sikrer identifikation af sagligt sammenhørende dokumenter, jf. gældende bestemmelser herom.

Figur 6.4

Opmærkning af kolonner med faste oplysninger i arkiveringsversioner af it-systemer med registrering af dokumenter, der deles af flere myndigheder	
Oplysning	Beskrivelse
Myndighedsidentifikation	Den eller de kolonner i arkiveringsversionen, som indeholder oplysninger om hvilken myndighed, der har registreret sagen eller dokumentet

Figur 6.5

Opmærkning af kolonner med faste oplysninger i arkiveringsversioner af it-systemer med dokumenter		
Oplysning	Beskrivelse	Udfaldsrum i tabellata
Dokumentidentifikation	Bruges til at angive den eller de kolonner i arkiveringsversionen, som beskriver dokumenternes entydige identifikation	

Lagringsform	Bruges til at angive den eller de kolonner i arkiveringsversionen, som beskriver, om dokumentet er lagret digitalt, på papir eller delvist på papir	Helt eller delvis digitalt = 1, papir = 2, ikke relevant = 3
Afleveret	Bruges ved aflevering af øjebliksbilleder m.v. til at angive den eller de kolonner i arkiveringsversionen, som beskriver, om dokumentet allerede er afleveret i en tidligere arkiveringsversion	Tidligere afleveret = 1, ikke tidligere afleveret = 2

Figur 6.6

Opmærkning af kolonner med oplysninger i arkiveringsversioner af it-systemer med registrering af dokumenter	
Oplysning	Beskrivelse
Sagsidentifikation	Den eller de kolonner i arkiveringsversionen, som beskriver sagernes entydige identifikation
Sagstitel	Den eller de kolonner i arkiveringsversionen, som indeholder sagernes titler
Dokumenttitel	Den eller de kolonner i arkiveringsversionen, som indeholder dokumenternes titler/beskrivelser
Dokumentdato	Den eller de kolonner i arkiveringsversionen, som indeholder oplysninger om dokumenternes afsendelses- og modtagelsesdatoer
Afsender/modtager	Den eller de kolonner i arkiveringsversionen, som indeholder oplysninger om et dokumentes afsender eller modtager
Digital signatur	Den eller de kolonner i arkiveringsversionen, som indeholder oplysninger, der er udtaget fra en digital signatur
FORM	Den eller de kolonner i arkiveringsversionen, som indeholder reference til FORM (Den fællesoffentlige forretningsreferencemodel)
Kassation	Den eller de kolonner i arkiveringsversionen, som indeholder oplysninger om bevaring og kassation

6.D. SQL-forespørgsler

6.D.1 Det modtagende arkiv kan fastsætte, at der til en arkiveringsversion skal defineres et antal SQL-forespørgsler til dokumentation af bestemte sammenhænge i arkiveringsversionen.

6.D.2 SQL-forespørgslerne udformes efter standarden SQL:1999 (core).

6.D.3 Forespørgslerne placeres i »Oplysninger om views og forespørgsler« i tabelindekset jf. figur 6.3, 7, og navngives efter myndighedens eget valg, dog således at navnet på de pågældende forespørgsler begynder med »AV«.

Afleveringsmedie

- 7.A.1 Arkiveringsversioner kan afleveres på CD-R, DVD-R eller USB-medie.
- 7.A.2 Antallet af CD-R og DVD-R i én aflevering må ikke overstige 10, med mindre andet aftales mellem afleverende myndighed og modtagende arkiv.
- 7.A.3 Den afleverende myndighed og det modtagende arkiv kan træffe aftale om aflevering på andre medier eller andre metoder til transport af data.
- 7.A.4 Det modtagende arkiv kan tillade, at en arkiveringsversion krypteres i forbindelse med transport.

Skemaer

Færdige skemaer til brug for skabelsen af en arkiveringsversion kan hentes fra Statens Arkivers hjemmeside