

Ministerium: Kulturministeriet
Journalnummer: Kulturmin.,
Rigsarkivet, j.nr. 17/09504

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 om offentlige arkivers virksomhed og efter drøftelse med de kommunale og regionale parter fastsættes:

Område

§ 1. Bekendtgørelsen gælder for arkivalier som hidrører fra den offentlige forvaltning og domstolene, og som af Rigsarkivet er bestemt til bevaring.

§ 2. Bevaring af digitale data skal ske i form af arkiveringsversioner.

Stk. 2. Bevaringsværdig lyd og video digitaliseres inden aflevering til offentligt arkiv og bevares som arkiveringsversioner, jf. stk. 1.

Stk. 3. Hvor det af bevaringsmæssige grunde er nødvendigt, kan Rigsarkivet anvise, at bevaringsværdige arkivalier digitaliseres inden aflevering til offentligt arkiv og bevares som arkiveringsversioner, jf. stk. 1.

§ 3. En arkiveringsversion af bevaringsværdige data skal fremstilles efter de retningslinjer, der fremgår af bilag 2-8 i denne bekendtgørelse.

Stk. 2. Data, som er skabt i forbindelse med forskning med anvendelse af videnskabelig metode, og som er skabt eller bearbejdet i statistikprogrammer eller tilsvarende, afleveres som en afleveringspakke efter reglerne i bilag 9.

Stk. 3. Rigsarkivaren kan anvise andre bevaringsformer end arkiveringsversioner og afleveringspakker, fremstillet efter bilag 2-9, hvis bevaringsmæssige hensyn tilsiger det.

Produktion og aflevering

§ 4. Produktion og aflevering af arkiveringsversioner og afleveringspakker af digitale data fra den statslige forvaltning og domstolene skal finde sted på tidspunkter, der fastsættes af Rigsarkivet.

§ 5. Produktion og aflevering af arkiveringsversioner og afleveringspakker af digitale data fra kommuner og regioner, som indeholder personoplysninger, skal finde sted, inden data slettes. Den pågældende kommune eller region kan træffe aftale om tidligere tidspunkter med det modtagende arkiv.

Stk. 2. Produktion af arkiveringsversioner af øvrige data skal finde sted, inden data slettes, eller når det it-system, data er skabt i, tages ud af brug.

Stk. 3. Rigsarkivet kan derudover fastsætte, at der skal produceres en arkiveringsversion af data, når det af bevaringsmæssige hensyn er nødvendigt.

§ 6. Såfremt den afleverende myndighed i forbindelse med produktion af arkiveringsversion konstaterer fejl eller mangler i dokumentkonverteringen eller i konverteringen af tabelindhold, skal der træffes aftale med det modtagende arkiv om udbedring heraf, inden arkiveringsversionen afleveres.

§ 7. Arkiveringsversioner af statslige myndigheders data skal godkendes af Rigsarkivet. Arkiveringsversioner af kommunale og regionale myndigheders data skal godkendes af det modtagende arkiv.

Stk. 2. 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. 3. Data, som er overført til en arkiveringsversion, må ikke slettes hos myndigheden, før arkiveringsversionen er godkendt.

Stk. 4. For forskningsdata afleveret som en afleveringspakke efter bilag 9, må data først slettes, når det modtagende arkiv har produceret og godkendt en arkiveringsversion af afleveringspakken.

Ikrafttræden m.v.

§ 8. Bekendtgørelsen træder i kraft den 1. marts 2020.

Stk. 2. Bekendtgørelse nr. 1007 af 20. august 2010 om arkiveringsversioner ophæves.

Rigsarkivet, den 12. februar 2020

ANNE-SOFIE JENSEN

/ Kirsten Villadsen Kristmar

Bilagsfortegnelse

Bilag 1 Bilagsfortegnelse

Bilag 2 Grafisk oversigt over elementer og struktur i en arkiveringsversion

Bilag 3 Arkiveringsversion af digitale 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

Bilag 9 Afleveringspakke for visse typer af forskningsdata

- 9. A. Aflevering af forskningsdata
- 9. B. Afleveringspakkens mappestruktur

- 9. C. Mappen Indices
- 9. D. Mappen ContextDocumentation
- 9. E. Mappen Data
- 9. F. Tekstformat
- 9. G. Datafil
- 9. H. Datatyper
- 9. I. Metadatafil

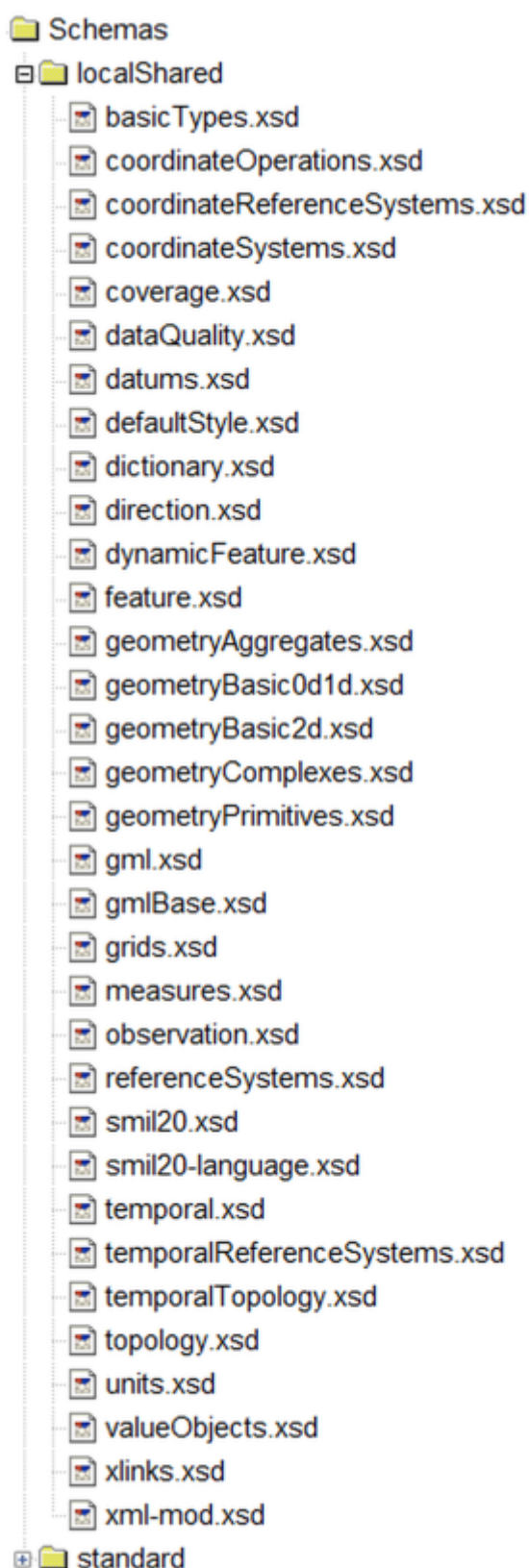
Grafisk oversigt over elementer og struktur i en arkiveringsversion

Figur 2.1 Grafisk oversigt over mulige elementer og struktur i en arkiveringsversion

Elementer, som ikke er gråmarkerede, er obligatoriske i en arkiveringsversion.



Figur 2.2 Placering af skemaer fra OGC til validering af GML version 3.1.1



Arkiveringsversion af digitale 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.a 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, jf. dog punkt 3. A. 3.b.

3. A. 3.b Under forudsætning af at der er gennemført kontrol med, at afsluttede sager er lukket korrekt, kan det modtagende arkiv godkende, at en arkiveringsversion kun indeholder dokumenter fra afsluttede sager. Metadata om samtlige sager og dokumenter, både afsluttede og uafsluttede, skal dog indgå i arkiveringsversionen.

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. A. 5 Mindre dokumentsamlinger, som er skabt i forbindelse med forskning med anvendelse af videnskabelig metode, kan afleveres som en del af institutionens ESDH-system efter aftale med det modtagende arkiv.

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. Et felt, der indgår i en primærnøgle må ikke udelukkende bestå af blanktegn eller være NULL.

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 en eller flere tabeller med angivelse af værdierne i kodede felter. Alternativt skal der enten tilføjes et felt i data med den værdi, som koden dækker, eller kodeværdien i tabellen skal 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 Arkiveringsversionens mapper

Navn på mappe	Beskrivelse	Obligato- risk	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. Skema-mappen 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 Rigsarkivet.

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 en 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 Hvis arkiveringsversionen indeholder data, som er skabt i forbindelse med forskning med anvendelse af videnskabelig metode og er afleveret efter reglerne i bilag 9, skal mappen *Indices* endvidere indeholde følgende indeksfil:

- researchIndex.xml

4. C. 1.d 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 Oplysninger i fileIndex.xml

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).	1	Ja

	Krav om 32 hexadecimale cifre medfører således krav om, at der afhængigt af værdien anvendes foranstillede nuller		
--	---	--	--

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 Oplysninger i contextDocumentationIndex.xml

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 Oplysninger i docIndex.xml

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	ArchivalFormat	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 bruges til validering af den pågældende GML-fil	Fritekst	0-1	Nej

4. C. 7.a **researchIndex.xml** skal indeholde angivelse af hovedtabeller og koder for manglende værdier, jf. figur 4.5:

Figur 4.5 Oplysninger i researchIndex.xml

1. Oplysninger om arkiveringsversionen				
	Elementnavn	Beskrivelse	Forekomst pr. aflevering	Obligatorisk
1.a	mainTables	Liste over hovedtabeller	1	Ja
2. Oplysninger om liste over hovedtabeller				
	Elementnavn	Beskrivelse	Forekomst pr. liste over hovedtabeller	Obligatorisk
2.a	table	Hovedtabel	1-m	Ja
3. Oplysninger om tabel				
	Elementnavn	Beskrivelse	Forekomst pr. tabel	Obligatorisk
3.a	tableID	Hovedtabellens ID, som angivet i elementet »folder«) i tableIndex.xml, jf. regler for navngivning af tabeller i 5. A.	1	Ja
3.b	source	Datas oprindelige format eller navn på program hvorfra data stammer	1	Ja

3.c	specialNumeric	Boolsk værdi, som angiver, om der er specialkoder for manglende værdier	0-1	Ja, hvis der er anvendt koder for manglende værdier
3.d	columns	Liste over kolonner	0-1	Ja, hvis der er anvendt koder for manglende værdier

4. Oplysninger om kolonneliste

	Elementnavn	Beskrivelse	Forekomst pr. liste over kolonner	Obligatorisk
4.a	column	Kolonne	1-m	Ja

5. Oplysninger om kolonne

	Elementnavn	Beskrivelse	Forekomst pr. kolonne	Obligatorisk
5.a	columnID	Kolonnens ID som angivet i tableIndex.xml, jf. regler for navngivning af kolonner i figur 6.3, 4.b	1	Ja
5.b	missingValues	Liste over koder for manglende værdier	1	Ja

6. Oplysninger om liste over koder for manglende værdier

	Elementnavn	Beskrivelse	Forekomst pr. liste over koder for manglende værdier	Obligatorisk
6.a	value	Kode for manglende værdi	1-m	Ja

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 en fil: table[fortløbende nummer].xml, jf. dog 4. D. 5

4. D. 4 »table[fortløbende nummer].xml« er en XML-instans, der indeholder data for den pågældende tabel, og dens struktur skal være i overensstemmelse med XML-instansen »tableIndex.xml«, således at datatyper i XML instansen er transformeret fra SQL:1999 til XML-datatyper i overensstemmelse med bilag 5 figur 5.1.

4. D. 5 Det er tilladt i mappen for hver tabel at inkludere et XML-skema, der kun angiver strukturen for den pågældende tabel. Skemaet navngives »table[fortløbende nummer].xsd«. Skemaet skal være i overensstemmelse med XML-instansen »tableIndex.xml«, der angiver strukturen for hele den relationelle database herunder samtlige tabeller.

4. D. 6 Hvis et felt i en tabel har værdien NULL, skal XML-instansens (»table[fortløbende nummer].xml«) element have attributten xsi:nil="true" for dette felt, jf. W3C-standarden for håndtering af nil-værdier i XML.

4. E. Mappen *ContextDocumentation*

4. E. 1 Mappen *ContextDocumentation* skal indeholde en 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*, *researchIndex.xsd* samt W3Cs standard XML-skema gælder, at der altid skal anvendes de skemaer, som Rigsarkivet 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 en eller flere filer af samme format, og navngives med dokumentets ID. Foranstillede nuller må ikke anvendes.

4. G. 6 Et dokumentets 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 i XML-instansen »tableIndex.xml« er defineret for hver tabel, 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 Tilladte datatyper for tabelindhold

Datatype	SQL:1999 (angives i tableIndex.xml)	XML-datatype (angives i table[fortløbende nummer]. xsd)
Tekst	CHARACTER, CHAR, CHARACTER VARYING, CHAR VARYING, VARCHAR, NATIONAL CHARACTER, NATIONAL CHAR, NCHAR NATIONAL CHARACTER VARYING, NATIONAL CHAR VARYING, NCHAR VARYING	string
Heltal	INTEGER, INT, SMALLINT	integer
Decimaltal	NUMERIC, DECIMAL, DEC	decimal
	FLOAT	float
	REAL, DOUBLE PRECISION	double
Boolsk	BOOLEAN	boolean
Hexadecimal	CHARACTER, CHAR, CHARACTER VARYING, CHAR VARYING, VARCHAR, NATIONAL CHARACTER, NATIONAL CHAR, NCHAR NATIONAL CHARACTER VARYING,	hexBinary

	NATIONAL CHAR VARYING, NCHAR VARYING	
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 Angivelse af tegn i XML-standarden

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 Et digitalt dokument, jf. dog 5. F. og 5. G., skal lagres i et 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.a 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 Grp3, Grp4, PackBit eller LZW.

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

5. E. 3 Dokumenter i TIFF RGB må udelukkende benytte følgende bitdybder: 1, 2, 4, 8, 24 og 32.

5. E. 3.a Dokumenter i TIFF RGB må maksimalt benytte 3 farvekanaler med en maksimal bitdybde på 24 bit (8x8x8 bit) evt. suppleret med maksimalt en alfakanal (8 bit), således at den samlede bitdybde for en billedfil ikke kan overstige 32 bit.

5. E. 4 Dokumenter i TIFF CMYK må udelukkende benytte følgende bitdybder: 1, 2, 4, 8, 32 og 40.

5. E. 4.a Dokumenter i TIFF CMYK må maksimalt benytte 4 farvekanaler med en maksimal bitdybde på 32 bit (8x8x8x8 bit) evt. suppleret med maksimalt én alfakanal (8 bit), således at den samlede bitdybde for en billedfil ikke kan overstige 40 bit.

5. E. 5 TIFF dokumenters anvendelse af XResolution og YResolution (TIFF Tag 282 og 283), skal ske på en sådan måde, at forholdet mellem bredde og højde ved anvendelse af disse værdier (skalering) svarer til sidernes dimensioner i det oprindelige dokument.

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. specifikationerne, der er angivet af OGC (Open Geospatial Consortium) i ISO 19136 GML version 3.1.1.

5. G. 1.a Geodatafiler journaliseret på sager i systemer med dokumenter kan konverteres til TIFF eller JPEG-2000, jf. 5. E., og behandles som et separat dokument iht. reglerne i bilag 4, 4. G. samt bilag 6, 6. C. 4 - 6. C. 6 efter aftale med det modtagende arkiv.

5. G. 2 I undermappen *localShared* til mappen *Schemas*, placeres skemasamlingen fra Open Geospatial Consortium til GML version 3.1.1 i uændret form. Skemaerne kan hentes i en samlet zip-pakke på Rigsarkivets hjemmeside.

5. G. 3 GML-filer, som er større end 1 GB, opdeles i mindre enheder efter det modtagende arkivs anvisning.

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

5. G. 4.a Hver GML-fil opmærkes i *docIndex.xml* og elementet *gmlXsd* opmærkes med en reference til det lokale XML-skema, der angiver GML-filens syntaksregel, jf. bilag 4, punkt 4. C. 6.a og 4. C. 6.b.

5. G. 4.b Kolonner i arkiveringsversionens øvrige tabeldata, der indeholder informationer om *Dokumentidentifikation* og *Lagringsform*, samt øvrige relevante særlige oplysninger, skal identificeres via elementet *functionalDescription* i *tableIndex.xml*, jf. bilag 6, punkt 6. C. 4 - 6. C. 6.

5. G. 5 Lokale XML-skemaer

5. G. 5.a De nødvendige lokale XML-skemaer til dokumentation og validering af GML-filernes indhold skal indgå i arkiveringsversionen.

5. G. 5.b Lokale XML-skemaer skal overholde syntaksen for skemaer i henhold til *W3C XML Schema Definition Language (XSD) 1.1 Part 1*.

5. G. 5.c Lokale XML-skemaer skal indkodes i tegnsættet UTF-8, hvilket skal fremgå af XML-deklarationen. Anvendelsen af tegn er begrænset, jf. bilag 5 punkt 5. D.

5. G. 5.d Et lokalt XML-skema placeres i samme mappe, som den GML-fil skemaet validerer, jf. reglerne i 4. G. 7. Et lokalt XML-skema kan også fungere globalt for flere GML-filer i en arkiveringsversion, såfremt strukturen for arkiveringsversionens GML-filer er indeholdt i skemaets syntaksregler, og placeres i så tilfælde i undermappen *localShared* til mappen *Schemas*, jf. 4. G. 7.a.

5. G. 5.e I det lokale XML-skemas rodelement skal følgende navnerum opmærkes: *gml*, *xlink*, *xsi* og den afleverende myndigheds navnerum, jf. figur 5.3. Navnerummet »xxx« skal substitueres af den afleverende myndigheds *URI*.

Figur 5.3 Navnerum i XML-skemaets rodelement

```
<schema
xmlns="http://www.w3.org/2001/XMLSchema"
xmlns:gml="http://www.opengis.net/gml"
xmlns:xlink="http://www.w3.org/1999/xlink"
xmlns:xxx="http://www.xxx.dk/gml/xxx"
targetNamespace="http://www.xxx.dk/gml/xxx" elementFormDefault="qualified">
```

- 5. G. 5.f) Det lokale XML-skema skal indeholde et importelement til navnerummet for GML version 3.1.1.
- 5. G. 5.g) Det lokale XML-skema skal indeholde en generel beskrivelse af GML-filens indhold dokumenteret ved brug af XML-dokumentationselementerne *annotation* og *documentation*. Omfanget og indholdet af dokumentationen aftales nærmere med det modtagende arkiv.
- 5. G. 5.h) et lokalt XML-skema skal der altid være opmærket mindst en lokal GML-feature med tilhørende GML-geometri. Omfanget og indholdet af features aftales nærmere med det modtagende arkiv.
- 5. G. 5.i) Det lokale XML-skemas GML-features og geometrier opmærkes som en sekvens af features med elementet *sequence* og omslutes i GML-filen af et XML-samleelement, der navngives »GEOMETRI«.
- 5. G. 5.j) Samleelementet *GEOMETRI* defineres i det lokale XML-skema og nedarver fra klassen *gml:_Feature* opmærket ved brug af XML-attributten *substitutionGroup* samt opmærkes med elementerne *complexType* og *complexContent* og placeres i den afleverende myndigheds navnerum.
- 5. G. 5.k) Det lokale XML-skema skal nedarve egenskaberne fra GML-klassen *gml:AbstractFeatureType*.
- 5. G. 5.l) Myndighedens lokale GML-features samt geometrier skal beskrives og dokumenteres ved brug af XML-dokumentationselementer *annotation* og *documentation*. Omfanget og indholdet af dokumentationen aftales nærmere med det modtagende arkiv.
- 5. G. 5.m) Der skal til den enkelte *gml:featureMember* i GML-filen etableres en entydig sammenkædning til arkiveringsversionens øvrige tabeldata, hvis en sådan eksisterer. Sammenkædningen mellem tabeldata og GML-instansen skal dokumenteres i det lokale XML-skema og i *tableIndex.xml* for de lokale features og kolonner, som udgør referencen. Mulige løsningsmodeller for etablering af referencer aftales nærmere med det modtagende arkiv.
- 5. G. 6 GML filer
 - 5. G. 6.a) Alle GML-filer i en arkiveringsversion skal overholde de gældende syntaksregler for henholdsvis XML version 1.0 og GML version 3.1.1.
 - 5. G. 6.b) GML-filer skal indkodes i tegnsættet UTF-8, hvilket skal fremgå af XML-deklarationen. Anvendelsen af tegn er begrænset, jf. bilag 5 punkt 5. D.
 - 5. G. 6.c) Geometrier og features fra samme kortmateriale/tabel skal så vidt muligt holdes samlet inden for den samme GML-fil. Afvigelse herfra aftales nærmere med det modtagende arkiv.
 - 5. G. 6.d) Som rodelement skal anvendes et *gml:FeatureCollection* element, jf. figur 5.4.

5. G. 6.e I rodelementet skal følgende navnerum opmærkes: gml, xlink, xsi og den afleverende myndigheds navnerum, jf. figur 5.4. Navnerummet »xxx« skal substitueres af den afleverende myndigheds *URI*.

Figur 5.4 Navnerum i GML-filens rodelement

```
<gml:FeatureCollection
xmlns:gml="http://www.opengis.net/gml"
xmlns:xlink="http://www.w3.org/1999/xlink"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:xxx="http://www.xxx.dk/gml/xxx"
xsi:schemaLocation="http://www.xxx.dk/gml/xxx . /1.xsd">
```

5. G. 6.f I rodelementet skal angives en lokal skemareference *xsi:schemaLocation* med angivelse af en relativ sti for placeringen af det lokale skema, jf. figur 5.4 og 5. G. 5.d.

5. G. 6.g Alle GML-filer skal opmærkes med et koordinat- og referencesystem. »EPSG:25832« alias »ETRF89 / ETRS89 UTM zone 32N« skal anvendes som opmærkning, hvis de afleverede geodata er placeret i Jylland, på Fyn eller Sjælland samt farvandet heromkring, dvs. inden for punkterne angivet i *lowerCorner* og *upperCorner*, jf. punkt 5. G. 6.i. Er dette ikke tilfældet, kan anvendes et andet koordinat- og referencesystem efter aftale med det modtagende arkiv.

5. G. 6.h Koordinat- og referencesystemet angives ved brug af GML-elementet *gml:boundedBy* med underelementerne *Envelope* og *lower-* og *upperCorner*. Denne opmærkning er gældende globalt for hele GML-filen og skal ikke angives lokalt i de enkelte *gml:featureMember* elementer.

5. G. 6.i Geografierne i GML-filen skal placeres inden for punkterne *lowerCorner* (212481.60 6019669.40) og *upperCorner* (961440.75 6510422.51). En anden eller mere præcis angivelse af disse yderpunkter kan anvendes efter aftale med det modtagende arkiv.

5. G. 6.j Geografiens dimension skal opmærkes ved brug af GML-attributten *srsDimension* i GML-elementet ”*Envelope*” sammen med GML-attributten *srsName*, der angiver kortmaterialets koordinat- og referencesystem.

5. G. 6.k En GML-fil skal indeholde mindst et *gml:featureMember* element. I *gml:featureMember* placeres samleelementet *GEOMETRI*, der indeholder GML-filens lokale features og geometrier. Elementet *GEOMETRI* skal være inkluderet i den afleverende myndigheds navnerum.

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 Oplysninger i archiveIndex.xml

Elementnavn	Betegnelse	Beskrivelse	Udfaldsrum	Forekomst	Obligatorisk
archiveInformationPackageID	ArkiveringsversionsID	Entydigt ID som tildeles arkiveringsversionen af Rigsarkivet	ArkiveringsversionsID, som defineret i denne bekendtgørelse	1	Ja
archiveInformationPackageIDPrevious	Tidligere aflevering	Entydigt arkiveringsversionsID på tidligere aflevering fra samme it-system eller datakilde	ArkiveringsversionsID, som defineret i denne eller tidligere bekendtgørelse	0-m	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
documentPeriodStart	Startdato for dokumenter	Startdato for digitale dokumenter i arkiveringsversionen	År, år-måned, eller år-måned-dag	0-1	Ja, hvis arkiveringsversionen indeholder digitale dokumenter

documentPeriodEnd	Slutdato for dokumenter	Slutdato for digitale dokumenter i arkiveringsversionen	År, år-måned, eller år-måned-dag	0-1	Ja, hvis arkiveringsversionen indeholder digitale dokumenter
archiveInformationPacketType	Slutaflevering	Angivelse af, om arkiveringsversionen er en slutaflevering	Boolsk værdi	1	Ja
creatorName	Arkivskaber	Den eller de myndigheder, institutioner, organisationer eller personer, der har skabt data. Kan også være organisatoriske enheder inden for en myndighed eller andre bidragydere	Fritekst	1-m	Ja
creationPeriodStart	Arkivskaber startdato	For hver arkivskaber angives, hvornår arkivskaber er begyndt at tilføje data	År, år-måned, eller år-måned-dag	1-m	Ja
creationPeriodEnd	Arkivskaber slutdato	For hver arkivskaber angives, hvornår arkivskaber er ophørt med at tilføje data	År, år-måned, eller år-måned-dag	1-m	Ja
archiveType	Arkiveringsversionsperiodetype	Angivelse af, om der er tale om en afsluttet arkivperiode (årgang)	Boolsk værdi	1	Ja
archiveTypeClosedFiles	Arkiveringsversion med lukkede sager	Angivelse af, om der for it-systemer med dokumenter, er tale om et øjebliksbillede af metadata men kun dokumenter fra lukkede sager	Boolsk værdi	0-1	Ja, hvis arkiveringsversionen indeholder digitale dokumenter
systemName	Systemnavn	Det officielle navn på it-system eller datakilde, hvor alle forkortelser er opløst	Fritekst	1	Ja
alternativeName	Alternativt systemnavn	Alternative navne på it-system eller datakilde	Fritekst	0-m	Nej
systemPurpose	Systemformål	Beskrivelse af formålet med at oprette og an-	Fritekst	1	Ja

		vende it-systemet eller indsamle data			
systemContent	Systemindhold	Beskrivelse af den centrale population og centrale variable i data	Fritekst	1	Ja
regionNum	Regionsnumre	Angivelse af, om der i systematisk form er registreret regionsnumre i data	Boolsk værdi	1	Ja
komNum	Kommunenumre	Angivelse af, om der i systematisk form er registreret kommunenumre i data	Boolsk værdi	1	Ja
cprNum	CPR-numre	Angivelse af, om der i systematisk form er registreret CPR-numre i data	Boolsk værdi	1	Ja
cvrNum	CVR-numre	Angivelse af, om der i systematisk form er registreret CVR-numre i data	Boolsk værdi	1	Ja
matrikNum	Matrikelnumre	Angivelse af, om der i systematisk form er registreret matrikelnumre i data	Boolsk værdi	1	Ja
bbrNum	BBR-numre	Angivelse af, om der i systematisk form er registreret BBR-numre i data	Boolsk værdi	1	Ja
whoSygKod	WHO-sygdomskoder	Angivelse af, om der i systematisk form er registreret WHO-sygdomskoder i data	Boolsk værdi	1	Ja
sourceName	Datakilde	Det officielle navn på andre datakilder, som har leveret data til datasættet (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	0-1	Ja, for it-systemer fra offentlige myndigheder
formClass	FORM-klassifikation	Kategorisering af arkiveringsversionen iht. FORM	Enhver gyldig FORM-reference iht. den angivne version af FORM	0-m	Ja, for it-systemer fra offentlige myndigheder
formClassText	FORM-klassifikation klartekst	FORM-klassifikation i klartekst	Fritekst	0-m	Ja, 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
containsGeodata	GML-filer indeholdt	Angivelse af, om der er GML-filer i arkiveringsversionen	Boolsk værdi	1	Ja
containsResearchData	Forskningsdata indeholdt	Angivelse af, om arkiveringsversionen indeholder forskningsdata	Boolsk værdi	1	Ja
researchSIP	Forskningsdata afleveret iht. bilag 9	Angivelse af, om data er afleveret iht. til bilag 9	Boolsk værdi	1	Ja
documentsDisposal	Kassation	Angivelse af, om der er foretaget systematisk kassation af dokumenter inden eller i forbindelse med produktion af arkiveringsversion	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 foregåen-

					de 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
personalDataRestrictedInfo	Persondata	Angivelse af, om der i arkiveringsversionen findes følsomme personoplysninger iht. databeskyttelsesforordningen	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 eller forskningsdatas ophav og opbygning.
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 en 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 et af de i arkiveringsversionen tilladte dokumentformater, jf. 5. E. - 5. F.

Figur 6.2 Dokumentkategorier i contextDocumentationIndex.xml (kontekstdokumentationsfilen)

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 præsentationsstruktur
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
1.l	systemTaxonomy	Registreringssystematik
1.m	systemInstruction	Instruks for anvendelse af systemet
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
7. Dokumentation af forskningsdata		
	Elementnavn	Beskrivelse
7.a	researchProjectDescription	Projektbeskrivelse gældende for de afleverede data
7.b	researchQuestionnaire	Spørgeskema, interviewguide og/eller registreringskema anvendt til at indsamle og analysere de afleverede data
7.c	researchProtocol	Protokoller og metoderapporter
7.d	researchPublication	Publikationer, som er udgivet på basis af de afleverede data
7.e	researchInformationOther	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 Oplysninger i tableIndex.xml

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 SQL:1999	1	Ja	
5.b	column	Oprindeligt navn på kolonner i primærnøglen	1-m	Ja	
6. Oplysninger om fremmednøgler					
	Elementnavn	Beskrivelse	Forekomst pr. nøgle/reference	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	Samleelement, der parvis binder de refererende og de refererede kolonner sammen. Ved sammensatte fremmednøgler gentages elementet.	1-m	Ja	
6.d		column	Navn på kolonnen, der refereres fra	1	Ja
6.e		referenced	Navn på kolonnen, der refereres til	1	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 functionalDescriptions i IT-systemer, der deles af flere myndigheder

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 functionalDescriptions i IT-systemer med dokumenter

Opmærkning af kolonner med faste oplysninger i arkiveringsversioner af it-systemer med dokumenter		
Oplysning	Beskrivelse	Udfaldsrum i tabeldata
Dokumentidentifikation	Bruges til at angive den eller de kolonner i arkiveringsversionen, som beskriver de digitale dokumenters entydige identifikation i arkiveringsversionens struktur	ID på op til 12 cifre, som findes blandt de navngivne dokumentmapper og er unikt inden for mappen <i>Documents</i> , jf. 4. G. 4

Lagringsform	Bruges til at angive den eller de kolonner i arkiveringsversionen, som beskriver, om dokumentet er helt eller delvist digitalt, på papir eller ikke har tilknyttet en fil	Helt eller delvis digitalt = 1, papir = 2, ikke relevant = 3
Afleveret	Bruges ved aflevering af øjeblicsbilleder 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 functionalDescriptions i IT-systemer med registrering af dokumenter

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 uddraget 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, BD-R eller USB-medie.
- 7. A. 2 Antallet af CD-R, DVD-R eller BD-R i en aflevering må ikke overstige 10, med mindre andet aftales mellem afleverende myndighed og modtagende arkiv.
- 7. A. 3 Ved aflevering på USB-medie af arkiveringsversioner over 4 GB skal der anvendes USB 3.0 (USB 3.1 Gen1) eller hurtigere USB-standard.
- 7. A. 4 Den afleverende myndighed og det modtagende arkiv kan træffe aftale om aflevering på andre medier eller andre metoder til transport af data.
- 7. A. 5 Medier med følsomme personoplysninger krypteres efter anvisning fra det modtagende arkiv.

Skemaer

Færdige skemaer til brug for skabelsen af en arkiveringsversion kan hentes fra Rigsarkivets hjemmeside.

Afleveringspakke for visse typer af forskningsdata

9. A. Aflevering af forskningsdata

9. A. 1 Reglerne i dette bilag gælder kun for data, som er skabt i forbindelse med forskning med anvendelse af videnskabelig metode, og som er skabt eller bearbejdet i statistikprogrammer eller tilsvarende.

9. A. 2 Afleveringspakken konverteres til en arkiveringsversion, jf. reglerne i bilag 1-8, af det modtagne arkiv.

9. B. Afleveringspakkens mappestruktur

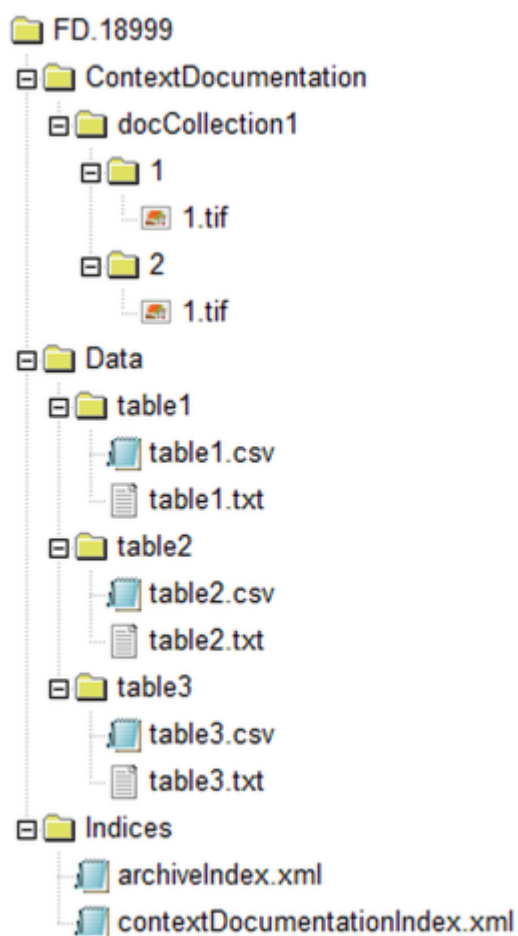
9. B. 1 I roden af filsystemet på afleveringsmediet, jf. bilag 7, skal der være placeret en mappe navngivet med afleveringspakkens navn. Afleveringspakkens navn består af præfikset »FD.« samt et unikt løbenummer for afleveringspakken.

9. B. 2 Løbenummeret for afleveringspakken udleveres af Rigsarkivet.

9. B. 3 Afleveringspakkens indhold fordeles i mapper som angivet i figur 9.2.

9. B. 4 Mapperne skal navngives som angivet i figur 9.2.

Figur 9.1 Grafisk oversigt over elementer og struktur i en afleveringspakke



Figur 9.2 Afleveringspakkens mapper

Navn på mappe	Indhold
<i>ContextDocumentation</i>	Kontekstdokumentation, jf. 9. D.
<i>Data</i>	Datafil og metadatafil, jf. 9. E.
<i>Indices</i>	Indeksfiler, jf. 9. C.

9. C. Mappen Indices

9. C. 1 Mappen *Indices* skal indeholde følgende indeksfiler med oplysninger om afleveringspakken og dens indhold:

- archiveIndex.xml
- contextDocumentationIndex.xml

9. C. 2 Indeksfilerne skal overholde deres tilhørende skema, jf. bilag 8.

9. C. 3 Arkivbeskrivelsesfilen archiveIndex.xml skal overholde reglerne i bilag 6 punkt 6. A.

9. C. 4 Kontekstdokumentationsfilen contextDocumentationIndex.xml skal overholde reglerne i bilag 4 punkt 4. C. 4.a samt 4. C. 4.b.

9. D. Mappen ContextDocumentation

9. D. 1 Mappen *ContextDocumentation* skal indeholde kontekstdokumentation, jf. 4. E.

9. E. Mappen Data

9. E. 1 En afleveringspakke skal indeholde et eller flere datasæt. Hvert datasæt skal bestå af en datafil og en tilhørende metadatafil.

9. E. 2 Datafil og metadatafil placeres i mappen *Data* i en undermappe, der navngives efter bilag 4 punkt 4. D. 2.a samt 4. D. 2.b.

9. E. 2.a Datafil navngives med undermappens navn efterfulgt af ekstensionen ». csv«.

9. E. 2.b Metadatafil navngives med undermappens navn efterfulgt af ekstensionen ». txt«.

9. F. Tekstformat

9. F. 1 Tegnsættet for henholdsvis datafil og metadatafil skal være indkodet som UTF-8, som angivet i ISO/IEC 10646:2003 Annex D og som beskrevet i *The Unicode Standard 5.1*, kapitel 3, og skal i øvrigt overholde bestemmelserne i bilag 5 punkt 5. D. 1.b – 5. D. 1.d.

9. G. Datafil

9. G. 1 Datafil afleveres som en semikolonsepareret tekstfil i henhold til RCF 4180 Common Format and MIME Type for CSV Files, der overholder syntaksen beskrevet som EBNF i figur 9.12.

9. G. 1.a Første linje i datafilen skal altid angive alle variabelnavne, angivet i samme rækkefølge som i metadatafilen.

9. G. 1.b Hvis semikolon indgår i en værdi for en variabel, skal hele værdien omslutes med dobbelt apostrof »"« (U+0022). Hvis dobbelt apostrof indgår i en værdi for en variabel, skal dobbelt apostroffen foranstilles med en dobbelt apostrof, og hele værdien omslutes med dobbelt apostrof.

9. G. 1.c Som linjeseparator skal anvendes en af følgende metoder for linjeskift: »CR+LF« (U+000D) samt (U+000A) eller »CR« (U+000D) eller »LF« (U+000A). Indhold i variable i datafilen må ikke indeholde linjeskift.

9. G. 2 En manglende værdi kan være en af følgende tre typer: Manglende værdi (tom) (jf. 9. G. 2.a), specialkode for manglende værdi (jf. 9. G. 2.d) eller brugerdefineret kode for manglende værdi (jf. 9. I. 6).

9. G. 2.a Manglende værdier (tom) i datafilen skal enten repræsenteres som ingen værdi »« eller et mellemrum » « (U+0020).

9. G. 2.b I en datafil må der konsekvent kun anvendes enten specialkoder eller brugerdefinerede koder for manglende værdier.

9. G. 2.c Specialkoder for manglende værdier må kun anvendes for kategoriske og numeriske variable.

9. G. 2.d Specialkoder for manglende værdier må kun anvendes for heltal og decimaltal og skal angives enten som en værdi fra A-Z eller . a- z.

9. G. 3 Indholdet af de enkelte variable skal renses for eventuelle foran- og efterstillede blanktegn.

9. H. Datatyper

9. H. 1 De seks standardiserede datatyper, som skal anvendes i datafilen, fremgår af figur 9.3.

9. H. 2 Dataformatnotationer for anvendte dataformater skal angives i metadatafilen, jf. figur 9.3. Dataformatnotationerne er case sensitive.

9. H. 2.a Værdier for bogstaverne »w« og »d« i figur 9.3 skal konsekvent angives. »w« angiver datatypens totale bredde. »d« angiver decimaler eller præcision i fraktioner af sekunder.

Figur 9.3 Tilladte datatyper

Datatype	Dataformat i datafil	Dataformatnotation i metadatafil	
Tekst	UTF-8 tegnsæt, jf. punkt 9. F. 1.	xml	string
		Stata	%ws
		SAS	\$w.
		SPSS	aw
Numerisk heltal	Repræsentation af et heltal med eller uden fortegn i henhold til DS/ISO 6093:1985 (NR1) standard, jf. syntaksregel i figur 9.6.	xml	int
		Stata	%w.Of
		SAS	fw.
		SPSS	fw
Numerisk decimaltal	Repræsentation af et decimaltal med eller uden fortegn i henhold til DS/ISO 6093:1985 (NR2) standard, jf. syntaksregel i figur 9.7.	xml	decimal
		Stata	%w.df eller %w.dg
		SAS	fw.d
		SPSS	fw.d
Dato	Angivelse af kalenderdato i henhold til DS/ISO 8601:1993 udvidet format:	xml	date
		Stata	%tdCCYY-NN-DD

	CCYY-MM-DD	SAS	yymmdd10.
	Alternativt kan følgende format anvendes, jf. syntaksregel i figur 9.8:	SPSS	sdate10
	CCYY/MM/DD		
Tidspunkt	Angivelse af tidspunkt i henhold til DS/ISO 8601:1993 udvidet format, jf. syntaksregel i figur 9.9: hh:mm:ss	xml	time
		Stata	%tcHH:MM:SS
		SAS	time. eller time8.
		SPSS	time8
Tidsstempel	Angivelse af dato og tidspunkt i henhold til DS/ISO 8601:1993 udvidet format: CCYY-MM-DDThh:mm:ss.sss eller CCYY-MM-DD hh:mm:ss.sss Alternativt kan følgende format anvendes, jf. syntaksregel i figur 9.10: dd-mmm-yyyy hh:mm:ss Fraktioner af sekunder er valgfrit og tilladt med en præcision på op til 6 cifre. Tidszone i tidsangivelser er ikke tilladt, jf. syntaksregel i figur 9.10.	xml	datetime
		Stata	%tcCCYY-NN-DD! THH:MM:SS %tcCCYY-NN-DD! THH:MM:SS.sss
		SAS	e8601dt19. e8601dtw. d
		SPSS	datetime20 eller ymdhms19 ymdhmsw.d

9. I. Metadatafil

9. I. 1 Metadata fra en datafil afleveres som en struktureret tekstfil, der overholder syntaksen beskrevet som EBNF i figur 9.11.

9. I. 1.a Metadatafilen udformes som anvist i figur 9.4, hvor otte etiketter opdeler metadata i specifikke kategorier.

9. I. 1.b Hver etiket skal forekomme en gang i metadatafilen. Forekomsten af indholdet af etiketten fremgår af kolonnerne »Forekomst« og »Obligatorisk« i figur 9.4.

9. I. 1.c Etiketnavne er reserverede ord og må ikke benyttes til navngivning af metadatafilens øvrige indhold.

Figur 9.4 Metadatafilens struktur

Etiket	Beskrivelse af etiketindhold	Udfaldsrum	Forekomst	Obligatorisk
SYSTEMNAVN	Navn på det program, som data udtrækkes fra, eller datas oprindelige format.	SPSS SAS Stata Excel eller fritekst	1	Ja
DATAFILNAVN	Navnet på datafilen, som den benævnes i brugssammenhæng. Datafilnavn skal overholde krav i punkt 9. I. 2.	ISO/IEC 9075:1999 - Database Language SQL (SQL-99)	1	Ja
DATAFILBESKRIVELSE	Beskrivelse af datafilens indhold.	Fritekst	1	Ja
NØGLEVARIABLE	Datafilens unikke nøglevariabel angivet med navne på den/de variable, nøglevariablen består af.	ISO/IEC 9075:1999 - Database Language SQL (SQL-99)	0-1	Ja, hvis unik nøglevariabel findes
REFERENCE	Referencer til andre datafiler i afleveringspakken angivet som anvist i punkt 9. I. 3.	ISO/IEC 9075:1999 - Database Language SQL (SQL-99)	0-m	Ja, hvis reference til anden datafil i afleveringspakken findes
VARIABLE	Variable i datafilen. En variabel angives som et sæt bestående af variabelnavn efterfulgt af notation for variabelens dataformat, jf. figur 9.3, samt en eventuel kodelistereference, jf. 9. I. 5.f. Variable skal overholde krav i punkt 9. I. 4.	Navngivning af variable skal overholde ISO/IEC 9075:1999 - Database Language SQL (SQL-99)	1-m	Ja
VARIABLEBESKRIVELSE	Beskrivelser af variabelnes indhold.	Fritekst	1-m	Ja
KODELISTE	Kodelister angivet med kodelistens navn efterfulgt af flere sæt bestående af kode og kodebeskrivelse.	Navngivning af kodelister skal overholde ISO/IEC 9075:1999 - Da-	0-m	Ja, hvis der findes kodelister

	Kodelister skal overholde krav i punkt 9. I. 5.	Database Language SQL (SQL-99)		
BRUGERKODE	Brugerdefinerede koder for manglende værdier angivet med navnet på den variabel, hvor koderne anvendes, efterfulgt af de brugerdefinerede koder. Brugerdefinerede koder for manglende værdier skal overholde krav i punkt 9. I. 6.	Navngivning af variable skal overholde ISO/IEC 9075:1999 - Database Language SQL (SQL-99)	0-m	Ja, hvis der findes brugerdefinerede koder for manglende værdier

9. I. 2 Datafilnavne skal være unikke inden for samme afleveringspakke.

9. I. 3 Reference

9. I. 3.a En reference til en anden datafil i afleveringspakken angives med navnet på den datafil, der refereres til (fremmeddatafil), efterfulgt af variabelnavnet for fremmeddatafilens nøglevariabel (fremmedvariabel) efterfulgt af variabelnavnet for den variabel (referencevariabel) i datafilen, der refererer til fremmedvariablen i fremmeddatafilen.

9. I. 3.b Der skal være fuld overensstemmelse mellem datatype og længde i nøglevariablene, der indgår i referencen.

9. I. 4 Variabelnavne skal være unikke inden for samme metadatafil.

9. I. 5 Kodeliste

9. I. 5.a Kun kategoriske og numeriske variable må have henvisninger til en kodeliste.

9. I. 5.b Det er kun tilladt at medtage kodelister for variable, der indeholder data af datatypen numerisk heltal, decimaltal eller tekst, jf. datatyper i figur 9.3.

9. I. 5.c Alle koder i data skal defineres i en kodeliste og forklares med en kodebeskrivelse.

9. I. 5.d Hvis koder er defineret som et interval, hvor ikke alle koder har kodebeskrivelser, skal dette dokumenteres i variabelbeskrivelsen.

9. I. 5.e Koder skal være unikke inden for samme kodeliste.

9. I. 5.f Referencen mellem en kodeliste og den variabel, som refererer til kodelisten, angives under etiketten VARIABEL, jf. figur 9.4, som anvist i 9. I. 5.g og 9. I. 5.h.

9. I. 5.g Hvis datatypen er numerisk heltal eller decimaltal, angives referencen med det valgte kodelistenavn og et efterstillet punktum ».« (U+002E).

9. I. 5.h Hvis datatypen er tekst, angives referencen med det valgte kodelistenavn med et foranstillet dollartegn »\$« (U+0024) og et efterstillet punktum ».« (U+002E).

9. I. 6 Brugerkode

9. I. 6.a En brugerdefineret kode for en manglende værdi er kun tilladt for kategoriske og numeriske variable.

9. I. 6.b En brugerdefineret kode for en manglende værdi angivet i metadatafilen, skal altid fremgå af kodelisten, som den tilhører.

Figur 9.5 EBNF tegnforklaring

::= defineret som

() angiver en gruppering, der skal udføres samlet

[] angiver muligheden for (0 eller 1)

{ } angiver mulig gentagelse (0 eller flere)

| angiver et valg (enten eller)

... angiver et fortløbende interval

!! angiver en beskrivende forklaring på alm. dansk

""" omslutter faktiske værdier, der skal skrives uden fortolkning

Figur 9.6 EBNF for heltalstype DS/ISO 6093:1985 (NR1) standard

Nonterminal	Terminal
INT ::=	NR1
NR1 ::=	[FORTEGN] CIFFER {CIFFER}
FORTEGN ::=	"+" "-" !! <i>Det er valgfrit at anvende fortegn for positive heltal</i>
CIFFER ::=	"0" "1" ... "9"

Figur 9.7 EBNF for decimaltalstype DS/ISO 6093:1985 (NR2) standard

Nonterminal	Terminal
DECIMAL ::=	NR2
NR2 ::=	[FORTEGN] CIFFER {CIFFER} [FORTEGN] CIFFER {CIFFER} DECIMALMÆRKE CIFFER {CIFFER} DECIMALMÆRKE CIFFER {CIFFER}
FORTEGN ::=	"+" "-" !! <i>Det er valgfrit at anvende fortegn for positive decimaltal !! Det er ikke tilladt at anvende negativt fortegn foran værdien 0</i>
CIFFER ::=	"0" "1" ... "9"
DECIMALMÆRKE ::=	". " "," "

Figur 9.8 EBNF for datotyper

Nonterminal	Terminal
DATO ::=	ISO-8601-DATE ALTERNATIV-DATE
ISO-8601-DATE ::=	CC YY BINDESTREG MM BINDESTREG DD
CC ::=	CIFFER CIFFER !! <i>årtusinde angivet med to heltal</i>
BINDESTREG ::=	"-" !! <i>bindestreg (U+002D)</i>

YY ::=	CIFFER CIFFER !! årstal angivet med to heltal
MM ::=	CIFFER CIFFER !! måned angivet med to heltal
DD ::=	CIFFER CIFFER !! dag angivet med to heltal
CIFFER ::=	"0" "1" ... "9"
ALTERNATIV-DATE ::=	CC YY SKRÅSTREG MM SKRÅSTREG DD
SKRÅSTREG ::=	"/" !! skråstreg U+002F

Figur 9.9 EBNF for tidstyper

Nonterminal	Terminal
TIDSPUNKT ::=	TIME KOLON MINUT KOLON SEKUND
TIME ::=	CIFFER [CIFFER] !! time angivet med et eller to heltal !! udfaldsrum er 0-23
KOLON ::=	":" !! kolon (U+003A)
MINUT ::=	CIFFER CIFFER !! minut angivet med to heltal !! udfaldsrum er 00-59
SEKUND ::=	CIFFER CIFFER !! sekund angivet med to heltal !! udfaldsrum er 00-59
CIFFER ::=	"0" "1" ... "9"

Figur 9.10 EBNF for datetimetyper

Nonterminal	Terminal
DATETIME ::=	ISO-8601-DATETIME ALTERNATIV-DATETIME IBM-DATETIME
ISO-8601-DATETIME ::=	CC YY BINDESTREG MM BINDESTREG DD TIDSTEMPELMÆRKE TIME KOLON MINUT KOLON SEKUND [PUNKTUM MILLISEKUNDER]
CC ::=	CIFFER CIFFER !! årtusinde angivet med to heltal
CIFFER ::=	"0" "1" ... "9"
BINDESTREG ::=	"-" !! bindestreg (U+002D)
YY ::=	CIFFER CIFFER !! årstal angivet med to heltal
MM ::=	CIFFER CIFFER !! måned angivet med to heltal
DD ::=	CIFFER CIFFER !! dag angivet med to heltal
TIDSTEMPELMÆRKE ::=	"T" MELLEMRUM
MELLEMRUM ::=	" " !! mellemrumstegn (U+0020)
TIME ::=	CIFFER CIFFER !! time angivet med to heltal !! udfaldsrum er 00-23
KOLON ::=	":" !! kolon (U+003A)
MINUT ::=	CIFFER CIFFER !! minut angivet med to heltal !! udfaldsrum er 00-59
SEKUND ::=	CIFFER CIFFER !! sekund angivet med to heltal !! udfaldsrum er 00-59
PUNKTUM ::=	". " !! punktum (U+002E)
MILLISEKUNDER ::=	CIFFER {CIFFER} !! millisekunder angivet med op til max 6 cifre

IBM-DATETIME ::=	DD BINDESTREG MÅNED BINDESTREG ÅRSTAL MEL- LEMRUM TIME KOLON MINUT KOLON SEKUND
MÅNED ::=	"Jan" "Feb" "Mar" "Apr" "May" "Jun" "Jul" "Aug" "Sep" "Oct" "Nov" "Dec" !! <i>der skal anvendes engelske for- kortelser for månednavn, og de er ikke case sensitive.</i>
ÅRSTAL ::=	CIFFER CIFFER CIFFER CIFFER !! <i>årstal angivet med fire heltal</i>

Figur 9.11 EBNF Syntaksregler for metadatafil

Nonterminal	Terminal
METADATAFIL ::=	SYSTEMNAVN DATAFILNAVN DATAFILBESKRIVELSE NØGLEVARIABLE REFERENCE VARIABLE VARIABLEBESKRIVELSE KODELISTE BRUGERKODE
SYSTEMNAVN ::=	"SYSTEMNAVN" LINJESKIFT "SAS" "Stata" "SPSS" "Excel" FRITEKST LINJESKIFT LINJESKIFT {LINJE- SKIFT}
DATAFILNAVN ::=	"DATAFILNAVN" LINJESKIFT TITEL LINJESKIFT LINJE- SKIFT {LINJESKIFT}
LINJESKIFT ::=	(CR LF) CR LF
CR ::=	!! <i>vognretur (U+000D)</i>
LF ::=	!! <i>linjeskift (U+000A)</i>
TITEL ::=	(BOGSTAV {BOGSTAV CIFFER}) (DOBBELTAPOSTROF BOGSTAV {BOGSTAV CIFFER} DOBBELTAPOSTROF) !! <i>En TITEL må aldrig begynde med et tal, men må godt være en blanding af bogstaver og tal på max 128 tegn, og hvis titel er et reserveret ord i SQL:1999, skal titel omkranses af dobbeltapost- rof, jf. ISO/IEC 9075:1999 - Database Language SQL (SQL:1999)</i>
DOBBELTAPOSTROF ::=	"" !! <i>dobbeltapostrof (U+0022)</i>
BOGSTAV ::=	"A" "B" ... "Z" "a" "b" ... "z" "_ " !! <i>Samt andre nati- onale karakterer, der er tilladte i constraint names i standarden SQL:1999 !! underscore (U+005F)</i>
CIFFER ::=	"0" "1" ... "9"

DATAFILBESKRIVELSE ::=	"DATAFILBESKRIVELSE" LINJESKIFT FRITEKST LINJESKIFT LINJESKIFT {LINJESKIFT}
FRITEKST ::=	BOGSTAV CIFFER ANDRETEGN {BOGSTAV CIFFER ANDRETEGN} !! <i>FRITEKST kan være en blanding af bogstaver, tal og andre tegn, så længe det giver semantisk mening</i>
ANDRETEGN ::=	!! <i>alle tilladte tegn i UTF-8, som ikke er bogstaver og tal jf. punkt 9. F</i>
NØGLEVARIABEL ::=	"NØGLEVARIABEL" LINJESKIFT {VARIABELNAVN MELLEMRUM} LINJESKIFT LINJESKIFT {LINJESKIFT}
VARIABELNAVN ::=	TITEL
MELLEMRUM ::=	!! <i>mellemrumstegn (U+0020)</i>
REFERENCE ::=	"REFERENCE" LINJESKIFT {FREMMEDDATAFIL MELLEMRUM FREMMEDVARIABEL MELLEMRUM REFERENCEVARIABEL LINJESKIFT} LINJESKIFT
FREMMEDDATAFIL ::=	TITEL !! <i>navn på den datafil, der refereres til</i>
FREMMEDVARIABEL ::=	APOSTROF VARIABELNAVN {MELLEMRUM VARIABELNAVN} APOSTROF !! <i>navn på nøglevariablen i den datafil, der refereres til</i>
REFERENCEVARIABEL ::=	APOSTROF VARIABELNAVN {MELLEMRUM VARIABELNAVN} APOSTROF !! <i>navn på variablen i datafilen, der refereres fra, som refererer til nøglevariablen i den datafil, der refereres til</i>
APOSTROF ::=	"'" !! <i>enkelt apostrof (U+0027)</i>
VARIABEL ::=	"VARIABEL" LINESKIFT VARIABELSÆT {VARIABELSÆT} LINJESKIFT {LINESKIFT}
VARIABELSÆT ::=	VARIABELNAVN MELLEMRUM DATAFORMATNOTATION MELLEMRUM [KODELISTEREFERENCE] LINJESKIFT
KODELISTEREFERENCE ::=	[DOLLAR] TITEL PUNKTUM
DATAFORMATNOTATION ::=	INTEGERNOTATION DECIMALNOTATION DATONOTATION DATETIMENOTATION TIMENOTATION TEKSTNOTATION
INTEGERNOTATION ::=	!! <i>se figur 9.3 for dataformatnotationer for numerisk heltal</i>
DECIMALNOTATION ::=	!! <i>se figur 9.3 for dataformatnotationer for decimaltal</i>
DATONOTATION ::=	!! <i>se figur 9.3 for dataformatnotationer for datotyper</i>
DATETIMENOTATION ::=	!! <i>se figur 9.3 for dataformatnotationer for tidsstempler</i>
TIMENOTATION ::=	!! <i>se figur 9.3 for dataformatnotationer for tidspunkter</i>
TEKSTNOTATION ::=	!! <i>se figur 9.3 for dataformatnotationer for tekst</i>
DOLLAR ::=	"\$" !! <i>dollartegn (U+0024)</i>
PUNKTUM ::=	". " !! <i>punktum (U+002E)</i>
VARIABELBESKRIVELSE ::=	"VARIABELBESKRIVELSE" LINJESKIFT BESKRIVELSE {BESKRIVELSE} LINJESKIFT {LINJESKIFT}
BESKRIVELSE ::=	VARIABELNAVN MELLEMRUM APOSTROF FRITEKST APOSTROF LINJESKIFT

KODELISTE ::=	"KODELISTE" LINJESKIFT {KODESÆT} LINJESKIFT {LINJESKIFT}
KODESÆT ::=	KODELISTENAVN LINJESKIFT {KODE MELLEMRUM KODEBESKRIVELSE LINJESKIFT}
KODELISTENAVN ::=	TITEL
KODE ::=	APOSTROF HELTAL DECIMAL FRITEKST APOSTROF
KODEBESKRIVELSE ::=	APOSTROF FRITEKST APOSTROF
HELTAL ::=	!! se figur 9.6 EBNF for heltalstype
DECIMAL ::=	!! se figur 9.7 EBNF for decimaltalstype
BRUGERKODE ::=	"BRUGERKODE" LINJESKIFT {VÆRDISÆT} LINJESKIFT {LINJESKIFT}
VÆRDISÆT ::=	VARIABELNAVN MELLEMRUM VÆRDI {MELLEMRUM VÆRDI} LINJESKIFT
VÆRDI ::=	APOSTROF HELTAL DECIMAL FRITEKST APOSTROF

Figur 9.12 EBNF Syntaksregler for datafil

Nonterminal	Terminal
DATAFIL ::=	DATAINDHOLD
DATAINDHOLD ::=	OVERSKRIFT LINJESKIFT RÆKKE {RÆKKE}
OVERSKRIFT ::=	VARIABELNAVN {SEPARATORTEGN VARIABELNAVN} !! <i>Der anvendes ikke et separatortegn efter det sidste variabelnavn i overskriftslinjen</i>
VARIABELNAVN ::=	TITEL
TITEL ::=	(BOGSTAV {BOGSTAV CIFFER}) (DOBBELTAPOSTROF BOGSTAV {BOGSTAV CIFFER} DOBBELTAPOSTROF) !! <i>En TITEL må aldrig begynde med et tal, men må godt være en blanding af bogstaver og tal, på max 128 tegn, og hvis titel er et reserveret ord i SQL:1999 skal titel omkranses af dobbeltapostrof, jf. ISO/IEC 9075:1999 - Database Language SQL (SQL:1999)</i>
BOGSTAV ::=	"A" "B" ... "Z" "a" "b" ... "z" "_ " !! <i>Samt andre nationale karakterer, der er tilladte i constraint names i standarden SQL:1999 !! underscore (U+005F)</i>
CIFFER ::=	"0" "1" "..." "9"
SEPARATORTEGN ::=	";" !! <i>semikolon (U+003B)</i>
LINJESKIFT ::=	(CR LF) CR LF
CR ::=	!! <i>vognretur (U+000D)</i>
LF ::=	!! <i>linjeskift (U+000A)</i>
RÆKKE ::=	VÆRDI (DOBBELTAPOSTROF VÆRDI DOBBELTAPOSTROF) {SEPARATORTEGN VÆRDI (DOBBELTAPOSTROF VÆRDI DOBBELTAPOSTROF)} LINJESKIFT !! <i>Der anvendes ikke et separatortegn efter den sidste værdi i en række. Se 9. G. 1.b for omslutning af en værdi, som indeholder separatortegn og/eller dobbeltapostrof</i>

VÆRDI ::=	BOGSTAV CIFFER ANDRETEGN {BOGSTAV CIFFER ANDRETEGN} !! <i>VÆRDI kan være en blanding af bogstaver tal og andre tegn !! Foran- og efterstillede blanktegn er ikke tilladt i værdier</i>
ANDRETEGN ::=	!! <i>alle tilladte tegn i UTF-8, som ikke er bogstaver og tal, jf. 9. F.</i>
DOBBELTAPOSTROF ::=	"""" !! <i>dobbeltapostrof (U+0022)</i>