

Forskrift til arkivloven av 1. desember 1999 nr. 1566 om utfyllende tekniske og arkivfaglige bestemmelser om behandling av offentlige arkiver, kapittel VIII:

BESTEMMELSER OM ELEKTRONISK ARKIVMATERIALE SOM AVLEVERES ELLER OVERFØRES SOM DEPOSITUM TIL ARKIVVERKET

Fastsatt av Riksarkivaren 11. mai 2000 med hjemmel i forskrift av 11. desember 1998 nr. 1193 om offentlige arkiv (arkivforskriften) § 5-8.

1. Generelle bestemmelser

§ 1-1 Virkeområde

Bestemmelsene gjelder for statlige organer.

Bestemmelsene har gyldighet for elektroniske journal- og arkivsystemer og for bevaringsverdig arkivinformasjon i andre elektroniske systemer.

Gjennom egne, generelle eller spesifikke bestemmelser skal Riksarkivaren på forhånd ha fastsatt hvilke systemer som skal produsere arkivversjoner for avlevering eller deponering, og hvilke konkrete opplysninger som skal bevares fra det enkelte system, jf. arkivforskriften §§ 3-20 og 3-21. Jf. også generelle bestemmelser i "Felles bevarings- og kassasjonsbestemmelser for statsforvaltningen", gitt i kapittel III i denne forskriften.

§ 1-2 Definisjoner

Med *avlevering* menes en fysisk overføring av arkivmateriale og råderetten over dette materialet.

Med *deponering (depositum)* menes en fysisk overføring av arkivmateriale uten at råderetten over det overføres.

Med et systems *arkivversjon* menes den samlede leveransen av data og medfølgende dokumentasjon som skal mottas av Arkivverket ved avlevering eller deponering.

Med *datauttrekk* menes det datainnhold som skal hentes ut av et system og inngå som arkivversjonens hoveddel.

Med *tabelluttrekk* menes datauttrekk fra databasetabeller, eventuelt andre uttrekk bestående av strukturerte data som kan presenteres i tabellform.

Med *elektroniske dokumenter* i datauttrekk menes dataobjekter i ustrukturert, ikke-tabellarisk form.

Med en *rapport* menes en logisk sammenstilling av utvalgte data fra en database. Teknisk kan en rapport være en avledet databasetabell eller et elektronisk dokument.

§ 1-3 Ansvarsforhold

Ved *avlevering* til Arkivverket overtar Riksarkivaren råderetten over arkivmaterialet og ansvaret for alt videre vedlikehold, jf. arkivloven § 10.

Når kopi av elektronisk arkivmateriale *deponeres* i Arkivverket, jf. arkivforskriften §§ 3-17 og 5-2, beholder vedkommende organ råderetten og det fulle ansvaret for materialet. Organet skal betjene og vedlikeholde materialet inntil avlevering senere har funnet sted. Arkivverket skal bevare den deponerte kopien utilgjengelig for bruk, og bare utføre teknisk vedlikehold.

Med mindre organet eller Arkivverket har meldt behov for en ny vurdering, skifter en deponering automatisk status til avlevering når det fastsatte avleveringstidspunktet er nådd.

§ 1-4 Valg av metode for å produsere arkivversjon

En arkivversjon med et uttrekk av de data som skal bevares fra et elektronisk system, skal produseres og overføres til Arkivverket med definerte mellomrom, jf. også § 2-2. Riksarkivaren fastsetter hvilken metode som skal legges til grunn, og hvilken periodeinndeling som skal brukes ved produksjon av datauttrekk fra det enkelte system, jf. vedlegg 1.

For journal- og arkivsystemer gjelder egne bestemmelser, jf. vedlegg 2.

2. Fremgangsmåte ved avlevering eller deponering

§ 2-1 Arkivversjonen

Ved avlevering eller deponering skal data eksporteres fra vedkommende system, og konverteres til et spesifisert, standardisert format, jf. § 4-1. Dette datauttrekket skal overføres sammen med nødvendig dokumentasjon, og bevares uavhengig av det opprinnelige systemet som systemets arkivversjon.

Arkivversjoner fra alle elektroniske systemer, unntatt journal- og arkivsystemer som er basert på Noark-standarden, skal bestå av følgende:

- a) uttrekk fra databasetabeller (tabelluttrekk), jf. §§ 4-1 – 4-4,
- b) beskrivelse av tabelluttrekkes struktur og innhold, jf. § 6-1,
- c) eventuelle tilknyttede elektroniske dokumenter, jf. §§ 5-1 – 5-3,
- d) dokumentasjon av det opprinnelige systemet og bruken av dette, jf. §§ 6-2 og 7-1,
- e) informasjon om avleveringen eller deponeringen, jf. § 7-2.

Arkivversjoner fra Noark-systemer skal følge spesifikasjonene i vedlegg 2. I dette tilfellet kreves verken struktur- og innholdsbeskrivelse eller dokumentasjon av det opprinnelige systemet, men det skal følge med informasjon om avleveringen eller deponeringen, jf. § 7-2.

§ 2-2 Tidspunkt, tidsintervall og form for overføring

Riksarkivaren bestemmer tidspunkt og tidsintervall for overføring av arkivversjoner fra det enkelte system.

Riksarkivaren bestemmer om materialet skal overføres med status som avlevering eller deponering, og kan kreve deponering tidligere enn tidspunktet for avlevering. *Avlevering* skal ordinært skje når materialet er ca. 25-30 år gammelt, men Riksarkivaren kan også fastsette et tidligere eller senere avleveringstidspunkt, jf. arkivforskriften § 5-2.

Dersom et system går ut av administrativ bruk uten at datainnholdet flyttes til et nytt system, eller dersom et organs virksomhet opphører, skal det foretas en samlet *avlevering*, jf. arkivforskriften § 5-3.

Systemer ("registre") som er underlagt konsesjon fra Datatilsynet, skal *avleveres* senest når organets konsesjon opphører.

Fra elektroniske journal- og arkivsystemer skal en arkivversjon *deponeres* straks en arkivperiode (vanligvis en 5-årsperiode) er avsluttet, jf. arkivforskriften § 3-17. Eventuelle tilknyttede elektroniske dokumenter skal medfølge ved deponering. Tilhørende dokumenter på papir skal følge som ledd i den ordinære avleveringen når materialet er ca. 25-30 år gammelt, jf. vedlegg 2.

§ 2-3 Gjennomføring

Avlevering eller deponering av arkivmateriale skal på forhånd være avtalt med Arkivverket, jf. arkivforskriften § 5-9. Av praktiske grunner kan Arkivverket kreve at det først leveres en prøve av arkivversjonen.

Kostnadene ved avlevering og deponering skal dekkes av vedkommende organ, jf. arkivloven § 10.

Et organ som avleverer elektronisk arkivmateriale, skal ikke slette grunnlagsdataene for arkivversjonen i sitt system før avleveringen er godkjent av Arkivverket. Grunnlagsdataene kan aldri slettes så lenge et materiale er *deponert*, jf. § 1-3.

§ 2-4 Godkjenning

Ved avlevering og deponering skal Arkivverket godkjenne arkivversjonen, og kvittere for det mottatte materialet.

Arkivverket kan nekte å godkjenne arkivversjonen dersom:

- innholdet ikke omfatter de deler som er fastsatt i § 2-1, punkt a-e,
- tabelluttrekkene ikke omfatter de opplysningene som er fastsatt,
- kravene til lagringsmedium, tegnsett, formater eller fil- og katalogstruktur ikke tilfredsstilles,
- feil eller mangler i dataene kan tilbakeføres til feil ved selve produksjonen av uttrekket,
- eller dersom den medfølgende struktur- og innholdsbeskrivelsen er feilaktig eller ufullstendig.

Hvis materialet ikke godkjennes, kan Arkivverket kreve at organet produserer en ny arkivversjon. Alternativt kan Arkivverket kreve refusjon for nødvendig arbeid for å bringe materialet i samsvar med bestemmelsene, jf. arkivloven § 11.

§ 2-5 Tilbake lån

Avlevert eller deponert elektronisk arkivmateriale kan ikke lånes tilbake til vedkommende organ. I stedet kan det leveres kopi av slikt materiale mot at organet dekker kostnadene forbundet med dette, jf. arkivforskriften § 5-10.

Kopi av *avlevert* materiale leveres så sant dette ikke er i strid med bestemmelser gitt i medhold av lov som tidsbegrenser vedkommende organs tilgang til materialet.

3. Krav til lagringsmedium

§ 3-1 Lagringsmedium med tilhørende filsystem

Ved avlevering og deponering skal det benyttes en type lagringsmedium som er godkjent av Riksarkivaren. På alle lagringsmedier skal det benyttes et godkjent filsystem som bevarer arkivversjonens opprinnelige filnavn. Godkjente lagringsmedier og filsystemer er spesifisert i vedlegg 3. Andre typer medier kan kun benyttes etter særskilt avtale med Arkivverket.

Lagringsmedier skal merkes etter reglene i vedlegg 3.

§ 3-2 Navngiving og organisering

Kataloger og filer på et lagringsmedium skal være navngitt som fastsatt i vedlegg 4. Lagringsmediets katalogstruktur skal også følge bestemmelsene i vedlegg 4.

Dersom en datafil har en referanse til en annen datafil, skal den være navngitt og plassert i en katalogstruktur på en slik måte at de relative referansene blir bevart.

4. Formatkrav mm. for uttrekk fra databasetabeller

§ 4-1 Organisering av tabelluttrekk

Tabelluttrekk fra databaser skal bestå av sekvensielle tekstfiler. Et databaseuttrekk kan bestå av en eller flere tabeller. En tabell vil kunne være identisk med én tabell eller et sett (join) av tabeller i den opprinnelige databasen, jf. vedlegg 1. Relasjoner mellom tabellene/filene i uttrekket skal være beskrevet i den medfølgende dokumentasjonen, jf. § 6-1 og vedlegg 6.

Tabelluttrekk fra eldre nettverksdatabaser og hierarkiske databaser skal være normalisert til relasjonell form, som spesifisert i vedlegg 5. Uttrekk kan bare tas ut i annen form etter nærmere avtale med Arkivverket.

Tabelluttrekk fra Noark-baserte systemer skal produseres i det format som er spesifisert i den aktuelle versjonen av Noark-standard, jf. vedlegg 2.

§ 4-2 Tegnsett

Tabelluttrekk og medfølgende struktur- og innholdsbeskrivelser, jf. § 6-1, skal overføres som ren tekst i ukryptert og ukomprimert form, og benytte godkjent tegnsett.

Godkjente tegnsett er Unicode UTF-8 (ISO/IEC 10646-1:2000 Annex D) og ISO 8859-1: 1998, Latin 1. I det siste tilfellet brukes alternativt ISO 8859-4: 1998, Latin-4 for samiske tegn.

Andre tegnsett aksepteres bare etter avtale med Arkivverket.

§ 4-3 Feltlengder

Feltlengder i tabelluttrekk kan være inntil 2047 tegn. Et felt som er lengre enn 2047 tegn, skal lagres som et elektronisk dokument i format ren tekst, jf. § 5-3. Dokumentet skal tilknyttes tabelluttrekket som en selvstendig fil etter bestemmelsene i § 5-1.

§ 4-4 Tekstfilformater

Tekstfiler med tabelluttrekk skal ha ett av følgende formater:

- a) "*Fast format*", dvs. at alle felt kommer i faste posisjoner innen postene. Lengde og/eller startposisjon for hvert felt skal spesifiseres i den medfølgende struktur- og innholdsbeskrivelsen, jf. § 6-1 og vedlegg 6. Alle lengder og posisjoner angis utifra antall tegn.
- b) "*Tegnseparert format*" ("kommaseparert" format), dvs. at poster og felt har variabel lengde og er atskilt med faste, entydige skilletegn. Postskille skal være <CR><LF>. Uttrekket skal ikke inneholde noen post med feltnavn (Heading). Feltskilletegnet og feltnavn skal spesifiseres i den medfølgende struktur- og innholdsbeskrivelsen, jf. § 6-1 og vedlegg 6.
- c) *XML-format*.

Spesifiserte bestemmelser om formater for tabelluttrekk er fastsatt i vedlegg 5.

5. Formatkrav mm. for elektroniske dokumenter

§ 5-1 Tilknytning av elektroniske dokumenter til tabelluttrekk

Elektroniske dokumenter skal være tilknyttet tabelluttrekk fra et journalføringssystem eller annet elektronisk system for registrering av dokumenter, jf. § 1-2 i Bestemmelser om elektronisk arkivering av saksdokumenter i kapittel IX i denne forskriften.

De elektroniske dokumentene skal lagres som separate filer, og overføres sammen med tabelluttrekket. Tabelluttrekket skal inneholde referanse til filnavn for hvert enkelt dokument. Hvis et dokument består av mer enn én fil, skal filene samles i en egen underkatalog. Hvis fil- eller katalognavn endres ved eksport, må endringene gjenspeiles i filreferansene for å opprettholde koblingene mellom tabeller og dokumenter, jf. vedlegg 4.

§ 5-2 Bruk av kryptering og digital signatur

Elektroniske dokumenter skal overføres ukryptert med mindre det er gjort særskilt avtale om kryptering. En slik avtale må inkludere regler om disponeringen av nøkler for dekryptering.

Elektroniske dokumenter kan være påført digitale signaturer basert på offentlig nøkkeltografi, fortrinnsvis etter mønster av Noark-4, kap. 10.2.

§ 5-3 Godkjente dokumentformater for arkivformål

Elektroniske dokumenter skal ved deponering og avlevering være lagret i dokumentformater som Riksarkivaren har godkjent for arkivformål. Med de presiseringer og begrensninger som fremgår av vedlegg 7, aksepteres følgende dokumentformater:

- Ren tekst: UTF-8 (ISO/IEC 10646-1:2000 Annex D) eller ISO 8859-1: 1998, Latin-1. I det siste tilfellet brukes alternativt ISO 8859-4: 1998, Latin-4 for samiske tegn
- TIFF – Tag Image File Format, versjon 6
- PNG – Portable Network Graphics
- XML – Extensible Markup Language, versjon 1.0
- PDF/A-1a - ISO 19005-1:2005 eller (inntil 01.01.2008): Adobe PDF.

- For fotografier, kart, video-, og lydsekvenser aksepteres følgende formater:
- for digitale fotografier samt for store digitale bilder: JPEG (10918-1:1994)
 - for digitale kart: SOSI versjon 2.2 (1995) og nyere
 - for digital video: MPEG2 (ISO 13818-2)
 - for digital lyd: MP3 (ISO 11172-3), Flac, PCM eller PCM-basert Wave.

Bestemmelser om tidligere aksepterte dokumentformater som ikke lenger godkjennes ved avlevering og deponering, er fastsatt i vedlegg 7.

§ 5-4 Elektroniske dokumenter fra Noark-systemer

Elektroniske dokumenter og elektroniske rapporter fra Noark-systemer skal ved avlevering eller deponering være lagret som fastsatt i vedlegg 2. Jf. også § 5-1.

6. Krav til teknisk dokumentasjon

§ 6-1 Struktur- og innholdsbeskrivelse av datauttrekket

For å muliggjøre gjenskaping og senere bruk av det elektroniske arkivmaterialet skal det medfølge en fullstendig struktur- og innholdsbeskrivelse av tabelluttrekkene. Beskrivelsen skal omfatte tabellene og feltene i uttrekket som avleveres eller deponeres, *ikke* tabellstrukturen i den databasen som uttrekket er eksportert fra.

Struktur- og innholdsbeskrivelsen skal overføres i elektronisk form som fastsatt i vedlegg 6. Beskrivelsen skal være i ren tekst i godkjent tegnsett.

Beskrivelsen skal fortrinnsvis også benytte XML-syntaks. Nærmere bestemmelser om XML-beskrivelser for de alternative tekstfilformater i tabelluttrekk, jf. § 4-4, er gitt i vedlegg 6. Riksarkivaren kan i det enkelte tilfelle bestemme at beskrivelsen *skal* være XML-basert og følge reglene i den standard som Riksarkivaren har utarbeidet og publisert spesifikt for dette formålet: ADDML - *Archival Data Description Markup Language*.

Struktur- og innholdsbeskrivelse kreves ikke for uttrekk fra Noark-systemer.

§ 6-2 Dokumentasjon av systemet som datauttrekket er eksportert fra

Det skal gis en oversikt over det opprinnelige systemets overordnede struktur, herunder relasjoner mellom tabeller og primær- og sekundærnøkler i de enkelte tabellene i systemets database. Denne oversikten kan ha form av et ER-diagram. Arkivverket kan dessuten be om at utvalgte skjermbilder og utskrifter fra systemet blir dokumentert.

For å beskrive søk, datakombinasjoner og uttrekksmetoder som ble anvendt i det opprinnelige systemet, kan Arkivverket fastsette at et antall SQL-spørringer fra systemet skal medfølge som dokumentasjon. Disse skal følge standarden SQL-92 (ISO/IEC 9075: 1992), og være tilpasset tabelluttrekket. SQL-setningene skal inngå i struktur- og innholdsbeskrivelsen dersom denne er utformet i XML.

Uttrekk fra Noark-systemer krever ikke dokumentasjon av det opprinnelige systemet.

7. Krav til administrativ dokumentasjon

§ 7-1 Det opprinnelige systemets funksjon og brukssammenheng

Følgende opplysninger skal gis om bruken av systemet som har produsert datauttrekket:

- a) systemets formål og administrative brukssammenheng,
- b) systemets hovedfunksjoner,
- c) navn på arkivskaperen, eventuelt på de organer som har hatt roller som arkivskapere ved å oppdatere systemet. Når et system har hatt flere arkivskapere, skal også de ulike rolletypene som arkivskaper beskrives,
- d) systemets inn- og ut-data, herunder datautveksling med andre systemer,
- e) eventuell hjemmel i lov eller forskrift for å opprette systemet.

En håndbok som beskriver bruken av systemet bør i tillegg følge som vedlegg når den finnes.

Dokumentasjonen skal være i elektronisk form, i ett eller flere av formatene som er fastsatt i § 5-3. Filnavn og filplassering på lagringsmediet skal være som fastsatt i vedlegg 4.

Uttrekk fra Noark-systemer krever bare opplysningene under punkt c, jf. også § 7-2.

§ 7-2 Overordnede opplysninger om avleveringen eller deponeringen

Følgende overordnede opplysninger om avleveringen eller deponeringen skal ligge ved:

- a) navn på organet som er ansvarlig for avleveringen eller deponeringen,
- b) navn og eventuelt versjonsnummer på systemet som datauttrekket kommer fra,
- c) tidsrommet som datauttrekket omfatter,
- d) navn på arkivskaper (en eller flere). Dersom datauttrekket omfatter flere arkivskapere, skal start- og sluttdato oppgis for hver arkivskaper. Dersom uttrekket kommer fra et Noark-system som har vært delt av flere organer, skal det beskrives hvordan hver arkivskaper identifiseres,
- e) antall poster i datauttrekket (for de viktigste tabellene) og annen tallfesting av omfang,
- f) antall lagringsenheter som medfølger,
- g) navn på den ansvarlige for overføringen hos det avleverende eller deponerende organet.

Dersom deler av materialet er gradert, konsesjonsbelagt, underlagt taushetsplikt eller av andre grunner unntatt fra offentlighet, skal det informeres særskilt om dette.

Informasjonen skal være i elektronisk form, i ren tekst i godkjent tegnsett. Filnavn og filplassering på lagringsmediet skal være som fastsatt i vedlegg 4.

Vedlegg

Vedlegg 1:

Uttreksmetode ved produksjon av arkivversjon for avlevering eller deponering (jf. § 1-4)

Før avlevering eller deponering må det være fastsatt hvordan datauttrekk for arkivering skal foretas fra vedkommende system. Uttreksmetoden må baseres på den oppdateringsmåte som er anvendt i det aktuelle systemet. Oppdateringen av opplysninger kan alternativt være foretatt ved:

- overskriving av eksisterende data,
- overskriving kombinert med en endringslogg (historisk informasjon),
- tilføyning av nye poster (akkumulering),
- eventuelt ved en kombinasjon av disse metodene.

På grunnlag av den oppdateringsmåten som er brukt, bestemmer Riksarkivaren om tabelluttrekket fra en database skal produseres som et *snitt* (øyeblikksbilde eller sistestatus, vanligvis i form av en årsversjon), eller om det skal omfatte *akkumulerte opplysninger* for nærmere bestemte tidsrom.

Riksarkivaren bestemmer i hvert tilfelle om en arkivversjons tabelluttrekk skal omfatte:

1. alle (eller de fleste) tabeller fra systemet,
2. en eller flere avledede (konsoliderte) tabeller.

Alternativt kan Riksarkivaren fastsette at arkivversjonens datauttrekk skal produseres som en eller flere rapporter i et av dokumentformatene som er spesifisert i vedlegg 7.

Vedlegg 2:

Bestemmelser om uttrekk fra Noark-systemer og andre journal- og arkivsystemer (jf. § 1-4)

1. Fellesbestemmelser for Noark-systemer

I henhold til arkivforskriften § 3-17 skal uttrekk fra elektroniske journaler, inkludert eventuelle tilknyttede elektroniske dokumenter, *deponeres* i Arkivverket straks en journalperiode er avsluttet. Den del av sakarkivet som består av dokumenter på papir, skal først overføres til Arkivverket som ledd i den ordinære *avleveringen* når materialet er ca. 25-30 år gammelt. Elektroniske dokumenter kan bare medfølge uttrekk fra Noark-4-systemer som er godkjent for elektronisk arkivering ("O2"-godkjente systemer).

Journalopplysningene skal eksporteres fra databasen som tabelluttrekk fra hele arkivperioder, jf. reglene for periodisering og bortsetting i Noark-standarden. Uttrekket skal omfatte:

- 1) alle saker som ble *avsluttet* i perioden med samtlige journalposter i disse sakene,
- 2) eventuelle elektroniske saksdokumenter som er tilknyttet journalpostene.

Sammen med uttrekket skal det følge 2 elektroniske rapporter:

- 3) en samlet, saksordnet rapport som inneholder alle saker som ble *avsluttet* i perioden og samtlige journalposter som er tilknyttet disse sakene,
- 4) en samlet kronologisk ordnet rapport som inneholder samtlige journalposter med journaldato innenfor perioden (i Noark-4: de aktuelle arkivdelene), dvs. samtlige journalposter som ble *registrert* i perioden.

2. Systemer basert på Noark-3 (1994) og Koark (1995)

Uttrekk basert på Noark-3 og Koark skal omfatte alle avsluttede saker fra en *journalperiode*. Uttrekk skal foretas i henhold til spesifikasjonene i Noark-3 (1994), kapittel 15.5.1, *alternativ 1*, dog med følgende endringer:

- Mediet skal ikke være magnetbånd, men som spesifisert i § 3-1.
- Journalsystemets arkivnøkkel skal vedlegges. Arkivnøkkeltabellen skal ha navnet ARKIV.DAT, og tas ut i følgende format: Arkivkode på fast 16 posisjoner og Gruppetekst (kodeteksten) på fast 255 posisjoner. Det skal ikke brukes postskille.

Følgende rapporter i elektronisk form skal deponeres sammen med tabelluttrekket:

1. *Endelig saksordnet utskrift* (jf. Noark-3, kap. 14.3.4).
2. *Endelig kronologisk journalutskrift* (jf. Noark-3, kap. 14.3.1). For å omfatte alle journalposter som ble registrert i perioden, må denne produseres *før* avsluttede saker skilles ut av den aktive databasen ved periodesaning.

Rapportene skal overføres i dokumentformat, enten som ren tekst, som PDF-dokumenter eller som XML-dokumenter med vedlagt DTD eller XML Schema etter formatkravene i vedlegg 7.

To *elektroniske* rapporter skal senere følge det papirbaserte sakarkivet når dette *avleveres*:

- *Avleveringsliste*
- *Avgraderingsliste*, i de tilfeller denne funksjonen er anvendt.

Disse tilleggssrapportene skal overføres i dokumentformat, enten som ren tekst, som PDF-dokumenter eller som XML-dokumenter med vedlagt DTD eller XML Schema etter formatkravene i vedlegg 7.

For uttrekk basert på Noark-3 og Koark kreves ikke annen medfølgende dokumentasjon enn filene med informasjon om selve avleveringen eller deponeringen, jf. § 7-2.

3. Systemer basert på Noark-4 (1999/2002)

Uttrekk basert på Noark-4 skal omfatte en eller flere avsluttede hele *arkivdeler*. Uttrekket skal produseres i det avleveringsformat som er spesifisert i Noark versjon 4.1 (Riksarkivaren, 2002) med eventuelle senere tilleggsspesifikasjoner fra Riksarkivaren.

Attributtens format, deres fordeling på tabeller og de spesifikke attributt- og tabellnavn som brukes i Noark-4, er obligatoriske ved eksport av data for avlevering og deponering. Uttrekk fra Noark-4-systemer som eventuelt bruker en annen intern struktur eller avvikende tabell- og attributtnavn, må være konvertert til dette formatet før avlevering eller deponering.

Uttrekk fra tabellen DOKVERSJON skal inneholde referanser til filnavn for samtlige medfølgende elektroniske dokumenter.

De elektroniske rapportene som skal følge tabelluttrekk fra Noark-4-baserte systemer, er:

1. *Saks- og dokumentoversikt* (jf. spesifikasjoner i Noark 4.1, kap. 15.3.4).
2. *Journal* (jf. spesifikasjoner i Noark 4.1, kap. 15.3.3)

Følgende elektroniske rapporter skal senere følge det papirbaserte arkivet ved *avlevering*:

- *Liste for bortsetting, avlevering og overføring*.
- *Arkivoversikt* (bare i de tilfeller en avlevering eller deponering omfatter data fra flere arkivdeler).
- *Avgraderingsliste* (bare i de tilfeller denne funksjonen er anvendt).
- *Kassasjonsliste* (bare i de tilfeller denne funksjonen er anvendt).

Disse tilleggssrapportene skal overføres som tabelluttrekk i ren tekst eller som elektroniske dokumenter i ett av formatene som er fastsatt i § 5-3.

For uttrekk basert på Noark-4 skal det følge spesifikk dokumentasjon fra systemleverandøren som beskriver hvilke deler av Noark-4 som er inkludert i systemet. For øvrig kreves ikke annen medfølgende dokumentasjon enn filene med informasjon om selve avleveringen eller deponeringen, jf. § 7-2.

4. Systemer som ikke er godkjent som Noark-systemer

Journal- og arkivsystemer som ikke er godkjent som Noark-systemer, skal produsere uttrekk som i størst mulig grad følger spesifikasjonene for Noark-systemer. Uttrekk fra systemer som ikke følger Noark-standarden, skal være fullstendig dokumentert som fastsatt i §§ 6-1 – 6-2 og §§ 7-1 – 7-2, med mindre Riksarkivaren har fastsatt noe annet for det enkelte system.

Vedlegg 3:

Bestemmelser om lagringsmedier og filsystemer (jf. §§ 3-1)

Følgende lagringsmedier med tilhørende filsystem er godkjent av Riksarkivaren til bruk ved avlevering og deponering av arkivversjoner:

1. CD-R

- Maksimal kapasitet skal være 700 MB. "Overburn" tillates ikke.
- Krav til filsystem: *ISO 9660: 1988, Level 1* eller *Microsoft Joliet*.
- Krav til merking og oppbevaring: CD-plater skal ikke pålimes etikett, men merkes med CD-penn. Hver plate skal oppbevares i lukket plastkassett som skal ha pålimt, merket etikett.

2. DVD-R eller DVD+R

- Data skal bare være lagret i ett lag på én side.
- Krav til filsystem: *UDF - Universal Disk Format*, versjon 1.02 eller senere. Versjonsnummer av UDF skal oppgis.
- Krav til merking og oppbevaring: som for CD-R.

3. Harddisk med USB 2.0-tilkobling

- Krav til filsystem: *FAT32* eller *NTFS* eller *UDF* versjon 1.02 eller senere. Versjonsnummer skal oppgis når UDF brukes.

4. Minnepinne/minneplugg med USB 2.0-tilkobling

- Krav til filsystem: *FAT32* eller *NTFS*.

Vedlegg 4:

Bestemmelser om filorganisering på lagringsmedier (jf. §§ 3-2, 5-1, 6-1, 7-1 og 7-2)

Alle navn på filer som beskriver arkivversjonen og som dokumenterer datauttrekket, skal være som vist nedenfor. Det er ingen spesielle krav til navn på de filene som utgjør selve datauttrekket, men det skal være en konsekvent bruk av dem i hele arkivversjonen når det gjelder store og små bokstaver, både for datafiler og for beskrivelses- og dokumentasjonsfiler.

Dersom en arkivversjon består av mer enn én lagringsenhet, skal det på hver enhet opprettes en rotkatalog med et entydig navn, f.eks. ARKIV1, ARKIV2 osv.

På *rotnivå på første lagringsenhet* skal det finnes en fil med navn INFO.TXT. Den skal inneholde de opplysninger om avleveringen eller deponeringen som er spesifisert i § 7-2.

En fil med navn INNHOLD.TXT skal finnes på *1. nivå under rotkatalogen på hver lagringsenhet*. Den skal gi oversikt over katalog- og filstrukturen på vedkommende enhet.

På *1. nivå under rotkatalogen på første lagringsenhet* skal følgende dokumentasjon plasseres:

- *Struktur- og innholdsbeskrivelsen* som medfølger arkivversjonen (jf. § 6-1). Denne filen skal ha navnet *metadata.aml* dersom beskrivelsen er basert på ADDML, og navnet *metadata.txt* hvis det er en ren tekst-fil uten XML-syntaks. Når beskrivelsen bygger på ADDML (*metadata.aml*), skal *addml.dtd* eller *addml.xsd* ligge ved i samme katalog som strukturbeskrivelsen. Tilhørende DTD eller XML Schema skal ligge ved dersom beskrivelsen for tabelluttrekk i XML-format ikke følger ADDML, jf. vedlegg 6.
- *System- og bruksdokumentasjon* (jf. §§ 6-2 og 7-1) dersom denne foreligger i elektronisk form. Det skal opprettes en *katalog* med navnet SYSDOK (på 1. nivå under rotkatalogen), og filene med dokumentasjon skal ligge under denne (på 2. nivå under rotkatalogen).

Filene som inneholder tabelluttrekk, eventuelle rapporter og elektroniske dokumenter, skal være organisert på følgende måte:

- *Tabelluttrekk* skal ligge på 1. nivå under rotkatalogen på vedkommende lagringsenhet. Filnavnene skal være entydige, og ha filendelse DAT når det er brukt ”fast format” eller ”tegnseparert format”, og filendelse XML når formatet er XML. Det forutsettes at det brukes et lagringsmedium som ikke gjør det nødvendig å dele filer med tabelluttrekk.
- For *elektroniske dokumenter* skal det opprettes en *katalog* med navnet DOKUMENT på 1. nivå under rotkatalogen på vedkommende lagringsenhet. Dokumentfilene skal samles under denne (på, eller fra og med 2. nivå under rotkatalogen). Filene kan eventuelt struktureres i flere underkataloger. Denne organiseringen kan fravikes dersom det er nødvendig for å bevare de relative filreferansene mellom tabelldata og elektroniske dokumenter.
- *Rapporter* (elektroniske utskrifter) fra datasystemet skal lagres under en *katalog* med navnet RAPPORT på 1. nivå under rotkatalogen på vedkommende lagringsenhet. Rapportfilene skal samles under denne katalogen (på 2. nivå under rotkatalogen). Denne organiseringen kan fravikes for rapporter med kobling til elektroniske dokumenter dersom dette er nødvendig for å bevare de relative filreferansene til de elektroniske dokumentene.

Filformat for medfølgende elektroniske dokumenter og dokumenter med system- og bruksdokumentasjon skal angis ved filendelse, f.eks. TXT, XML, PDF eller TIF. Dersom et enkelt dokument består av flere filer, skal disse samles i én katalog med et entydig navn. Fil- og katalognavn må samsvare med de aktuelle referansene i tabelluttrekkene.

Vedlegg 5:

Bestemmelser om formater for tabelluttrekk (jf. §§ 4-1 og 4-4)

1. Normalisering av tabelluttrekk fra eldre databaser

I uttrekk fra en database av eldre type (nettverksdatabase eller hierarkisk database), hvor de interne filene kan inneholde poster av flere typer, skal data tas ut på følgende måte:

- Hver posttype skal eksporteres til én enkelt sekvensiell fil. Opprinnelige fysiske pekere skal omgjøres til unike logiske nøkler (primær- og fremmednøkler).
- Repeterende grupper av felt (dvs. at ett eller flere felt er repetert flere ganger innenfor en post) skal skilles ut fra den opprinnelige posten, og eksporteres som egne poster i en ny fil med tilhørende nøkler.

2. Eksport av ulike datatyper

Ved eksport til tekstfiler i godkjent tegnsett skal ulike datatyper i den opprinnelige databasen legges ut slik:

- Ved bruk av ”fast format” legges tekststrenger ut venstrejustert i feltene. Etterfølgende posisjoner fylles opp med blanke tegn. Heltall og desimaltall legges ut høyrejustert uten ledende nuller, og posisjonene foran fylles opp med blanke tegn.
- Negative heltall og desimaltall legges ut med et minustegn før første siffer. Positive tall skal ikke representeres med plustegn.
- Desimaltall legges ut med komma som desimalskilletegn. Skilletegn for hver tusen skal ikke benyttes.
- Dato legges ut som ÅÅÅÅMMDD (ISO 8601).

3. Bruk av pakkede og binære tallfelt

Pakkede og binære tallfelt tillates ikke i tabelluttrekk.

4. Tabelluttrekk i XML-format

Tabelluttrekk i XML-format skal følge versjon 1.0 av XML-standarden.

Vedlegg 6:

Bestemmelser om dokumentasjon av tabelluttrekk (jf. § 6-1)

Struktur- og innholdsbeskrivelse av tabelluttrekk

For alle datafiler som inneholder tabelluttrekk, skal det ved avlevering eller deponering medfølge en fullstendig struktur- og innholdsbeskrivelse for å muliggjøre senere bruk av materialet.

Beskrivelsen skal følge i elektronisk form. Den skal være i ren tekst i godkjent tegnsatt (jf. § 4-2). Beskrivelsen skal omfatte tabellene (dvs. filene/posttypene) og feltene i det uttrekket som deponeres eller avleveres, *ikke* tabellstrukturen i den databasen som uttrekket er eksportert fra.

Dokumentasjonen skal omfatte:

- *Fil- og postbeskrivelse*
Hver enkelt tabell (fil/posttype) skal spesifiseres. Det skal gis en kort beskrivelse av innholdet. Relasjoner mellom tabeller angis, og primærnøkler og fremmednøkler samt kardinalitet defineres.
- *Feltbeskrivelse*
Felt navn i samtlige tabeller (filer/posttyper) beskrives. Feltenes datatype oppgis. Ved bruk av "fast format" oppgis hvert felts startposisjon og lengde. Det kreves også en utfyllende forklaring av hvert felts innhold i den grad dette ikke fremgår av feltnavnet.
- *Kodeforklaringer*
I de tilfeller felt inneholder kodede verdier, kreves fullstendige kodeforklaringer. Dersom systemet lagrer kodene i egne hjelpetabeller, skal uttrekk fra disse inngå i arkivversjonen.

Riksarkivaren har utformet en egen standard for struktur- og innholdsbeskrivelser som benytter XML-syntaks: ADDML – *Archival Data Description Markup Language*. Når XML-syntaks benyttes i beskrivelsen av tabelluttrekk med tekstfiler i fast eller tegnseparert format, skal den være basert på reglene i ADDML.

For tabelluttrekk med tekstfiler i XML-format behøver ikke struktur- og innholdsbeskrivelsen å følge ADDML. I så tilfelle skal en tilhørende DTD eller et tilhørende XML Schema følge med.

Struktur- og innholdsbeskrivelsen skal finnes i en egen fil, jf. bestemmelsene i vedlegg 4. Når beskrivelsen er basert på ADDML, skal filen ha navnet *metadata.aml*. Når beskrivelsen er en ren tekstfil uten XML-syntaks, skal filen ha navnet *metadata.txt*.

Vedlegg 7:

Bestemmelser om dokumentformater (jf. § 5-3)

1. Ordinære dokumenter med tekst- og bilde/grafikk-objekter kan ved avlevering eller deponering være lagret i følgende formater:

For ren tekst:

- 1.1 *UTF-8 (8-bit Unicode Transformation Format) - ISO/IEC 10646-1:2000 Annex D (2000) og RFC 3629 (<http://www.ietf.org>.)*

eller:

- 1.2 *ISO 8859-1: 1998, Latin-1. For samiske tegn: ISO 8859-4: 1998, Latin-4.*

For rastergrafikk med bilde- og tekstobjekter (herunder skannede dokumenter):

- 1.3 *PNG - Portable Network Graphics (ISO/IEC15948:2003 [E] og W3C 10.11.2003)*

Bare Deflate/LZ77-basert komprimering kan benyttes. Proprietære tillegg ("private chunks") i PNG-filer aksepteres ikke.

eller:

- 1.4 *TIFF - Tag Image File Format, versjon 6 (Aldus/Adobe, 1992)*

TIFF-dokumenter kan være "single page" eller "multiple page". Ved "single page" må hver side av dokumentet være lagret som en egen fil under en felles katalog.

TIFF-dokumenter med egendefinerte "tagger" aksepteres ikke.

Tapsfri komprimering iht. CCITT/TSS gruppe 4, LZW eller PackBit skal ordinært brukes for dokumenter i svart/hvitt, og LZW eller PackBit for dokumenter med farger/gråtoner.

Tapsfri Deflate/LZ77-komprimering generelt og tapsbringende komprimering med JPEG (ISO 10918-1:1994) i store TIFF-dokumenter med fotografier eller kart aksepteres, men under den grunnleggende forutsetning at man da anvender de endrede komprimeringsverdiene som fastsettes i TIFF TechNote2 (1995):

- Ved Deflate-komprimering skal komprimeringsverdi 8 brukes i stedet for den opprinnelige komprimeringsverdien 32946 (80b2 HEX) i TIFF versjon 6 (1992).
- Ved JPEG-komprimering (ISO 10918-1:1994) skal komprimeringsverdi 7 brukes i stedet for den opprinnelige komprimeringsverdien 6 i TIFF versjon 6 (1992).

For tekst- og bildeobjekter i kombinasjon (universalformater):

- 1.5 *XML - Extensible Markup Language, versjon 1.0 (W3C, 1998)*

Tegnsett i XML-dokumenter skal være UTF-8 eller ISO 8859-1/8859-4.

For XML-dokumenter skal en tilhørende DTD (Document Type Definition) og malfiler (f.eks. for stiloppsett) ligge ved som egne filer i de tilfeller DTD'er og malfiler er brukt. Alternativt til en DTD kan et tilhørende XML Schema ligge ved.

Alle filer eller objekter som et XML-dokument refererer til, skal være samlet i en underkatalog, jf. § 5-1.

eller:

- 1.6 *PDF/A-1a - ISO 19005-1:2005, Use of PDF 1.4, Level A conformance*

eller:

1.7 *Adobe PDF, versjon 1 og senere (inntil 01.01.2008)*

Adobe PDF aksepteres bare brukt til arkivering på midlertidig basis inntil 01.01.2008 i påvente av at den nye PDF/A-standard, som er utviklet spesifikt for langtidsarkivering, blir allment tilgjengelig. *Adobe PDF* må da brukes slik at dokumentene selv inneholder den informasjon og de hjelpemidler som kreves for å gjengi dem på alternative teknologi-plattformen. For at dokumentene skal være selvberørende, skal *Adobe PDF* brukes med følgende restriksjoner:

- alle brukte fonter skal være innbygd ("embedded"),
- dokumentene skal ikke ha innbygde filer, herunder lyd, video og andre multimedia-objekter, og ikke ha innbygd funksjonalitet for fileksekvering,
- dokumentene skal ikke benytte Java Script eller inneholde skjemadata,
- dokumentene skal ikke ha aktive linker til andre dokumenter, filer eller nettsteder,
- alle sikkerhetsfunksjoner som er mulige i PDF-dokumenter, så som passord og kryptering, må være deaktivert ved overføring til Arkivverket,
- skannede dokumenter som er konvertert til PDF, skal overføres i original versjon, ikke i en versjon hvor originalen helt eller delvis er erstattet med OCR-lest tekst.

2. Fotografier, enkeltvis eller i fotosamlinger, kan være lagret i:

2.1 *JPEG (ISO 10918-1:1994)*

JPEG-bilder må ikke være endret og lagret flere ganger fordi dette kan resultere i gjentatt komprimering og en ytterligere svekket bildekvalitet.

Fotografiske bilder som inneholder tegninger eller lesbar tekst, bør av kvalitetsgrunner fortrinnsvis lagres og komprimeres i PNG eller TIFF.

JPEG aksepteres ikke som et selvstendig dokumentformat for saksdokumenter pga. forringelsen ved gjentatt lagring. Innbygd JPEG-komprimering aksepteres imidlertid i PDF- og XML-dokumenter, og på nærmere angitte betingelser i TIFF, jf. punkt 1.4, ovenfor.

JPEG2000 (ISO/IEC 15444-1:2000) aksepteres ikke.

3. Digitale kartdata (geodata) kan være lagret i:

3.1 *SOSI - Samordnet opplegg for stedfestet informasjon (Statens kartverk), versjon 2.2 (1995) og nyere*

Denne norske nasjonale standarden for geografisk informasjon er under videreutvikling med sikte på en konvergering mot den kommende XML-baserte GML-standard for koding av geodata: ISO 19136. Det tas sikte på å godkjenne ISO 19136 for langtidsbevaring av kartinformasjon når den foreligger.

eller:

3.2 *TIFF versjon 6* (jf. punkt 1.4, ovenfor)

I dette tilfellet lagres kartdata som rene kartbilder, uten SOSI-formatets innbygde egenskapsinformasjon (metadata).

4. Digital video kan være lagret i:

4.1 *MPEG-2 (ISO/IEC 13818-2)*

MPEG-2 benytter komprimering. Fremstilling av versjoner for deponering og avlevering må foretas slik at det ikke medfører en ytterligere komprimering og svekket bildekvalitet.

5. Digital lyd kan være lagret i følgende formater:

For komprimert digital lyd:

5.1 *MP3 (ISO/IEC 11172-3)*

MP3 aksepteres når lyddokumentene er, eller inngår i saksdokumenter. Fremstilling av versjoner for deponering og avlevering må foretas slik at det ikke medfører en ytterligere komprimering og svekket lyd kvalitet.

For høykvalitets digital lyd:

5.2 *PCM - Pulse Code Modulation audio data /WAVE form audio format*

Ukomprimert PCM eller ukomprimert, PCM-basert WAVE aksepteres for høykvalitets digital lyd i lydarkiver.

eller:

5.3 *Flac - Free Lossless Audio Codec*

Tapsfritt komprimert Flac-format aksepteres for høykvalitets digital lyd i lydarkiver.

Valget mellom komprimert MP3 og høykvalitets PCM/Wave eller Flac skal i det enkelte tilfelle være avtalt før materialet deponeres eller avleveres.

6. Formater for web-innhold:

Saksdokumentasjon i HTML-format som er hentet fra World Wide Web, skal før avlevering eller deponering være konvertert til XML (jf. punkt 1,5), PDF/A-1a (jf. punkt 1,6) eller Adobe PDF (jf. punkt 1.7).

Tidligere godkjente dokumentformater som ikke aksepteres

Riksarkivaren foretar en løpende justering av bestemmelsene om dokumentformater for arkivformål. Følgende dokumentformater, som var godkjent i de tidligere avleveringsbestemmelsene av 11. mai 2000, er ikke lenger godkjent:

- TIFF/IT (ISO 12639)
- SGML
- HTML.

Etter 1. januar 2007 vil det ikke bli akseptert at disse formatene brukes ved arkivering av dokumenter som skal deponeres eller avleveres.