

24.8.2016

AL/11130/07.01.02.04.02/2008

ARKISTOLAITOKSEN VAATIMUKSET DIGITOIDULLE AINEISTOLLE

Sisältö	Arkistolaitoksen tekniset vaatimukset digitoidulle aineistolle
Tarkoitus	Varmistaa kansalliseen kulttuuriperintöön kuuluvien asiakirjojen säilyminen ja niiden käytettävyys. Vaatimukset täydentävät arkistolaitoksen digitoinnin laatukriteerejä (AL/11130/07.01.02.04.02/2008).
Kohderyhmä	Digitoinnin vaatimusmäärittelyt on tarkoitettu arkistolaitokselle ja digitointia arkistolaitokselle toteuttaville muistiorganisaatioille, yhteisöille ja yrityksille.
Voimassaoloaika	Toistaiseksi

24.8.2016

Sisällys

1	Johdanto	1
2	Pintaoriginaalit	2
2.1	Sidokset, asiakirjat, aktit ja kortistot	2
2.2	Kartat ja piirustukset	2
2.3	Valokuvavedokset	3
3	Läpivalaistavat originaalit	3
3.1	Lasinegatiivit, muut negatiivit ja diat	3
3.2	Mikrofilmi	4
4	Digitoinnin tekniset metatiedot	4
5	Siirtopaketit	5

1 Johdanto

Digitoinnilla edistetään asiakirjallisen kulttuuriperinnön käyttöä ja säilyttämistä. Digitoinnin tarkoituksena on alkuperäisen aineiston sisältämän informaation saattaminen laajemmin käyttöön. Samalla alkuperäiset aineistot saadaan pois aktiivikäytöstä ja suojattua. Digitoinnilla on oma roolinsa myös aineistojen säilytystilan hallinnassa. Digitointi edistää omalta osaltaan digitalisaatiota ja on osa kulttuuriperintöaineiston säilyvyyden ja käytettävyyden edistämisen kokonaisuutta. Tässä dokumentissa kuvataan arkistolaitoksen digitoinnin tuloksena muodostuvien aineistojen tekniset kriteerit.

Dokumentti on luonteeltaan velvoittava ja se täydentää Kansallisarkiston suositusta digitoinnin laatukriteereiksi (Suositus asiakirjallisen kulttuuriperinnön digitoinnin laatukriteereiksi, AL/11130/07.01.02.04.02/2008¹). Dokumentissa esitetyt kriteerit tulee noudattaa arkistolaitokselle toimitettavan aineiston digitoinnissa sekä muodostuneiden digitaalisten kappaleiden toimittamisessa. Digitoidun aineiston vastaanottaminen arkistolaitoksen tietojärjestelmiin edellyttää, että aineisto täyttää tässä dokumentissa esitetyt vaatimukset. Aineistoja, jotka eivät täytä tässä dokumentissa esitetyjä vaatimuksia, ei vastaanoteta arkistolaitoksen tietojärjestelmiin. Dokumentti kohdentuu asiakirjojen digitointiin kuvatiedostoiksi, eikä käsittele esimerkiksi äänen tai elävän kuvan digitointia.

Dokumentissa käytetty termistö perustuu Internet Engineering Task Force:n toimesta tehtyyn määrittelyyn [RFC 2119].

Dokumentissa käytetyt käännökset:

ENGLANTI	SUOMI
MUST	PITÄÄ
MUST NOT	EI SAA
REQUIRED	PAKOLLINEN
SHOULD	PITÄISI
SHOULD NOT	EI PITÄISI
MAY	SAA
OPTIONAL	VAPAAEHTOINEN

¹ <http://www.arkisto.fi/fi/digitoinnin-laatukriteerit/>

24.8.2016

2 Pintaoriginaalit

Tässä luvussa esitetyt vaatimukset soveltuvat heijastuvassa valossa digitoitaville pintaoriginaaleille, eli esimerkiksi asiakirjoille, kartoille, piirustuksille, valokuvavedoksille ja kirjoille, joista tehdään kuvatiedostoja. Aineistot on jaoteltu vaatimuksineen alalukuihin.

2.1 Sidokset, asiakirjat, aktit ja kortistot

Aineistot PITÄISI digitoida niin, että kuvankäsittelyssä niitä ei tarvitse kääntää/suoristaa. Aineiston PITÄÄ olla kuvaushetkellä tasainen. Eheyden varmistamiseksi aineisto PITÄÄ digitoida kokonaisuudessaan siten, että kuva-ala ylittää alkuperäisen aineiston reunat. Tästä poikkeuksena ovat dokumenttiskannerilla digitoidut aineistot. Tapauskohtaisesti päätetään toteutetaanko digitointi värillisenä (24bit) vai harmaasävyinä (8bit). Mustavalkoisia (1bit) tiedostoja EI SAA toimittaa arkistolaitokselle.

Asiakirjat PITÄÄ digitoida kuvatiedostoiksi siten, että syntyvä tiedosto on fyysisiltä mitoiltaan (n cm x n cm) 1:1 koossa alkuperäisen kanssa resoluutiolla 300 ppi (300 pikseliä tuuman janalla). Jos alkuperäiset asiakirjat ovat kooltaan alle A4, SAA korkeampaa resoluutiota käyttää.

Kortistoista PITÄÄ digitoida aineiston eheyden varmistamiseksi jokaisen kortin molemmat puolet siten, että sekä etupuoli että kääntöpuoli esiintyvät samassa kuvatiedostossa. Kääntöpuoli PITÄÄ digitoida aina, vaikka se ei sisältäisikään informaatiota. Korttien visuaaliseen esittämisen PITÄISI olla alla esitetyn kaltainen.

Etupuoli
Kääntöpuoli

Talletiedostolle soveltuvat profiilit:

eciRGB v2	ProPhoto RGB	AdobeRGB (1998)	sRGB	Gray gamma 2,2
x	x	x	x	x

2.2 Kartat ja piirustukset

Historialliset kartat ja piirustukset PITÄÄ digitoida kuvatiedostoiksi siten, että syntyvä kuvatiedosto on fyysisiltä mitoiltaan (n cm x n cm) 1:1 koossa alkuperäisen asiakirjan kanssa resoluutiolla 300 ppi. Digitoitavan aineiston eheyden varmistamiseksi aineisto PITÄÄ asiakirja digitoida kokonaisuudessaan siten, että kuva-ala ylittää sivujen reunat. Asiakirjan PITÄÄ olla kuvaushetkellä tasainen. Jos tasaisuuden saavuttamiseksi käytetään lasia tms., PITÄÄ se huomioida värinhallinnassa. Aineistot PITÄISI digitoida yhtenä tiedostona, mutta erittäin suurikokoiset alkuperäiset aineistot (>A0) SAA digitoida useana otoksena. Asiakirja PITÄÄ tässäkin tapauksessa digitoida siten, että alkuperäisestä asiakirjasta ei rajaudu informaatiota pois. Kartat ja piirustukset PITÄISI digitoida värillisinä (24bit), poikkeuksena värittömät piirustukset harmaasävyinä (8bit). Mustavalkoisia (1bit) tiedostoja EI SAA toimittaa arkistolaitokselle.

Talletiedostolle soveltuvat profiilit:

eciRGB v2	ProPhoto RGB	AdobeRGB (1998)	sRGB	Gray gamma 2,2
x	x	x		x

24.8.2016

2.3 Valokuvavedokset

Valokuvavedosten digitoinnissa resoluution valinnassa PITÄÄ hyödyntää tässä luvussa esitettyä resoluutiotaulukkoa. Eheyden varmistamiseksi aineisto PITÄÄ digitoida kokonaisuudessaan siten, että kuva-ala ylittää sivujen reunat. Riippuen alkuperäisen, digitoitavan kohteen, värillisyydestä, PITÄISI digitointi toteuttaa värillisenä (24bit) tai harmaasävyisenä (8bit).

Talletiedostolle soveltuvat profiilit:

eciRGB v2	ProPhoto RGB	AdobeRGB (1998)	sRGB	Gray gamma 2,2
x	x	x		x

Aineiston ollessa pienikokoista PITÄÄ resoluutiota nostaa minimivaatimuksista seuraavasti (tavoitteena noin A3 kokoinen tiedosto resoluutiolla 300 ppi):

Resoluutiotaulukko:

Originaalin koko	Resoluutio
6x9cm	1600 ppi
9x12cm	1200 ppi
13x18cm	900 ppi
18x24cm	600 ppi
A3 (tai suurempi)	300 ppi

3 Läpivalaistavat originaalit

Tässä luvussa esitetyt vaatimukset soveltuvat läpivalaisua vaativille originaaleille, eli esimerkiksi lasinegatiiveille, negatiiveille ja dioille.

3.1 Lasinegatiivit, muut negatiivit ja diat

Negatiiviaineistojen digitoinnissa PITÄISI käyttää laitteita, jotka eivät aiheuta aineistojen hetkellistä kuumenemistä. Negatiiviaineistoja ei myöskään PITÄISI altistaa äkillisille kosteudenmuutoksille tai suureen valaistukseen. Suositus on käyttää negatiivien digitoinnissa laadukasta järjestelmäkameraa ja ilmastoitua valopöytää.

Resoluution valinnassa PITÄÄ hyödyntää tässä luvussa esitettyä resoluutiotaulukkoa Aineiston ollessa pienikokoista, PITÄÄ resoluutiota nostaa minimivaatimuksista seuraavasti (tavoitteena noin A3 kokoinen tiedosto resoluutiolla 300 ppi):

Resoluutiotaulukko:

Originaalin koko	Resoluutio
6x9cm	1600 ppi
9x12cm	1200 ppi
13x18cm	900 ppi
18x24cm	600 ppi
A3 (tai suurempi)	300 ppi

Riippuen alkuperäisen, digitoitavan kohteen, värillisyydestä, PITÄISI digitointi toteuttaa värillisenä (24bit) tai harmaasävyisenä (8bit).

24.8.2016

Talletiedostolle soveltuvat profiilit:

eciRGB v2	ProPhoto RGB	AdobeRGB (1998)	sRGB	Gray gamma 2,2
x	x	x		x

3.2 Mikrofilmi

Mikrofilmiltä digitointi PITÄÄ toteuttaa aina harmaasävydigitoitina, jotta filmillä olevat harmaasävyt toistuvat ja tekstin himmeimmätkin detaljit eivät katoa. Mikrofilmin pienennyskerroin PITÄÄ tarkistaa filmirullan alusta. Mikäli pienennyskerrointa ei ole ilmoitettu, PITÄISI suurenuskerroin asettaa riittävän suureksi, jotta digitoitu asiakirja on luettavissa. Alkuperäisen aineiston oikean koon todentamisessa voi olla käytetty myös muita menetelmiä (esimerkiksi mittanauha) joita PITÄISI tarkastella tapauskohtaisesti, mutta kuitenkin niin, että tavoitellaan alla mainittua kuvakokoa.

Mikrofilmien osalta suurenuskertoimen avulla saadaan kuvattu aineisto toistumaan riittävän luettavana resoluutiolla 300 ppi. Skannauksessa käytettävä resoluutio on suurenuskerroin x haluttu resoluutio (esimerkiksi 12 x 300 = 3600). Mikrofilmi itsessään on varsin jyrkkä tallenneväline, joten digitoinnissa PITÄISI varoa jyrkkyyden korostamista.

Mikäli kyseessä on värillinen mikrofilmi, PITÄÄ siihen soveltaa kohdassa 4.1 mainittuja kriteerejä.

4 Digitoinnin tekniset metatiedot

Digitoinnin teknisillä metatiedoilla tarkoitetaan tallennetun talletiedoston (TIFF) sisään kirjoitettuja metatietoja digitointitapahtumasta, ts. siitä hetkestä, kun analoginen aineisto on muunnettu sähköiseen muotoon. Erilaisia formaatteja metatiedoille on muun muassa EXIF, IPTC ja XMP.

Kansallisarkiston järjestelmiin siirrettävissä tiedostoissa PITÄÄ olla seuraavat tiedot:

Metatieto	Tallennustapa
Arkistoyksikön tekninen tunniste (Vakka-arkistontietokannasta, mikäli tiedossa digitointihetkellä)	XMP-dc: Title
Digitointilaitteen merkki	TIFF Properties: Make
Digitointilaitteen malli	TIFF Properties: Model
Digitointiajankohta.	XMP-xmp: CreateDate
Skannauksen suorittanut henkilö, skannauspaikka (yritys/yhteisö)	XMP-dc: Creator
Kuvankäsittelyn ajankohta	XMP-xmp: ModifyDate
Kuvankäsittelyn suorittanut henkilö, kuvankäsittelypaikka (yritys/yhteisö)	XMP-dc: Contributor
Skannausohjelma	XMP-xmp: CreatorTool

Mikäli kuvia käsitellään skannaustapahtuman jälkeen, PITÄÄ kuvankäsittelytapahtumasta tallentaa kuvankäsittelyohjelma/t ja -tapahtumat. Jos aineistoa toimittava taho ei itse tallenna edellä mainittuja metatietoja tiedostoihin, PITÄÄ ne toimittaa Kansallisarkistolle .txt – tiedostona kuvatiedostojen toimittamisen yhteydessä (tästä PITÄÄ sopia aina erikseen).

24.8.2016

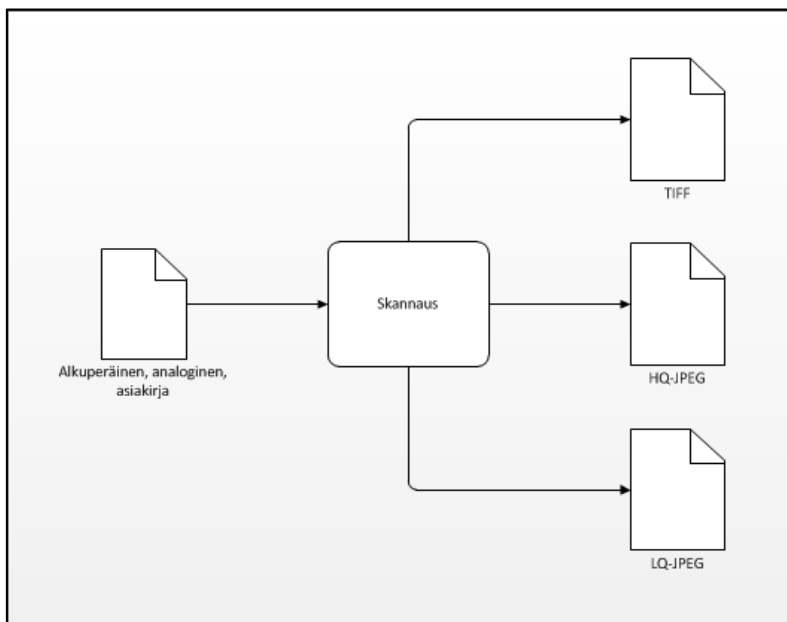
5 Siirtopaketit

Arkistolaitoksen digitoinnissa syntyvä siirtopaketti muodostuu talle- ja käyttötiedostoista jotka PITÄÄ tallentaa määrättyyn hakemistorakenteeseen. Jokaisesta analogisesta asiakirjasta PITÄÄ tallentaa alla kuvatut tiedostomuodot, joiden PITÄÄ olla tietosisällöltään identtisiä, mutta poikkeavat toisistaan teknisesti (Kuva 1: Analoginen yksittäinen asiakirja ja siitä tuotetut tiedostot).

Esimerkkipaketti:

- TIFF 6.0 = talletiedosto
300 ppi, 24/8bit, eciRGB v2/AdobeRGB(1998)/ProPhotoRGB/sRGB/Gray Gamma 2.2, pakkaamaton
 - poikkeuksen em. määrittelyyn muodostavat pienet originaalit, jotka ovat digitoitu suuremmalla resoluutiolla
- JPEG x 2 = käyttötiedostot
 - HQ-JPEG = 300 ppi, 24/8bit, sRGB/Gray Gamma 2.2, pakattu (pakkaustaso asteikolla 0-100% = 75%, asteikolla 1-12 = 9)
 - LQ-JPEG = 100 ppi, 24/8bit, sRGB/Gray Gamma 2.2, pakattu (pakkaustaso asteikolla 0-100% = 75%, asteikolla 1-12 = 9)

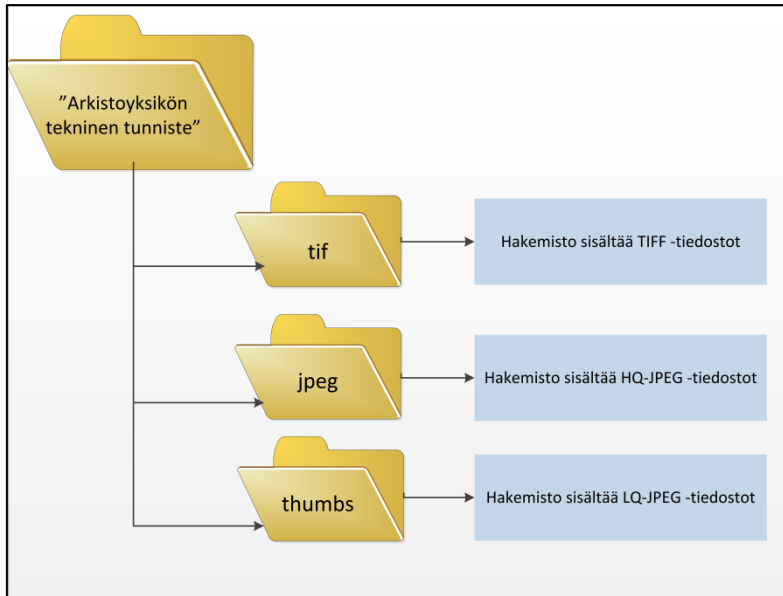
TIFF- ja HQ-JPEG – tiedostojen PITÄÄ olla bittisyvyydeltään ja resoluutioltaan samanlaisia. LQ-JPEG:n koon PITÄÄ olla yksi kolmasosaa TIFF ja HQ-JPEG – tiedostoista. Kaikki edellä mainitut tiedostoformaatit PITÄÄ tallentaa yksittäisinä tiedostoina. Arkistolaitos ei ota vastaan esimerkiksi monisivuisia TIFF – tiedostoja.



Kuva 1: Analoginen yksittäinen asiakirja ja siitä tuotetut tiedostot

Arkistoyksikkökansio PITÄÄ nimetä arkistotietokannasta noudettavalla arkistoyksikön teknisellä tunnisteella. Alikansioiden nimet ovat merkkikokoriippuvaisia ja ne PITÄÄ nimetä muodossa: tif, jpeg, thumbs. Alikansioissa olevat tiedostot PITÄÄ nimetä nelinumeroisina alkaen 0001.tif tai 0001.jpg.

24.8.2016



Kuva 2: Arkistolaitoksen vastaanottaman siirtopaketin kansiorakenne

Jokaisesta arkistolaitokselle toimitettavasta kuvatiedostosta PITÄÄ muodostaa tarkistussumma, jota käytetään arkistolaitoksessa vastaanoton yhteydessä kuvatiedostojen eheyden tarkistamiseen. Tarkistussumma PITÄÄ muodostaa erilliseen tiedostoon hakemistokohtaisesti, jokaiseen kuvaa sisältävään hakemistoon. Hyväksyttäviä algoritmeja ovat MD5 ja SHA-1.