

Arkistoitavien asiakirjojen digitointia koskevat vaatimukset ja ohjeet



KANSALLISARKISTO

Ohje

17.10.2024

KA/20919/30.01.00/2024

1	Sisältö ja tarkoitus	1
1.1	Kohderyhmä	1
1.2	Rajaukset	2
1.3	Keskeiset muutokset ja viittaukset muihin asiakirjoihin	2
1.4	Termit ja käsitteet	4
2	Johdanto	8
2.1	Vaatimusten soveltaminen eri asiakirjatyyppeihin kohdalla.....	9
2.2	Digitoinnin yleiset hyvät käytänteet ja suositukset.....	10
2.2.1	Digitoinnin laadunvarmistus	12
3	A4- tai foliokokoisia aineistoja sekä sidoksia ja erikoisaineistoja koskevat vaatimukset.....	13
3.1	Vaatimusten mukainen digitointiprosessi ja digitointitulokset.....	14
3.1.1	Kuvatiedostoa ja sen ominaisuuksia koskevat vaatimukset.....	16
3.1.2	Kuvatiedoston prosessointia kuvaavat MIX-metatiedot ja niiden rakenne	21
3.2	Aineistotyyppikohtaiset vaatimukset	24
3.2.1	A4-kokoiset ja foliokokoiset asiakirjat sekä sidokset.....	24
3.2.2	Kartat ja piirustukset.....	24
3.2.3	Sidokset, joissa tietosisältö esitetään aukeamittain.....	25
3.2.4	Valokuvat	27
3.2.5	Lasinegatiivit, muut negatiivit ja diat.....	28
3.2.6	Mikrofilmit ja -kortit	29
4	Audiovisuaalisia aineistoja koskevat vaatimukset.....	30
4.1	Ääninauhat	30
4.1.1	Äänitiedostojen vaaditut tekniset metatiedot	31
4.2	Kuvanauhat.....	31
4.2.1	Kuvanauhoista digitoitujen videotiedostojen tekniset metatiedot.....	32
5	Digitointiprosessissa muodostettava hakemistorakenne	33
5.1	Kuva-aineiston hakemistorakenne	33
5.2	Audiovisuaalisen aineiston hakemistorakenne	34
5.3	Esimerkkipaketit	35
5.4	Digitoidun aineiston siirtäminen Kansallisarkistoon	36
6	Muutokset	36



KANSALLISARKISTO

Säädökset, johon ohjeen antaminen perustuu

Laki Kansallisarkistosta (1145/2016, 2 §)

Arkistolaki (831/1994, 11 §)

Voimassaoloaika

Toistaiseksi, ohjeen päiväyksestä eteenpäin

Tämä ohje korvaa ja kumoaa ohjeet

- Kansallisarkiston vaatimukset hävittämiseen tähtäävään digitointiin 12.4.2019 (KA/3357/07.01.02.04.02/2019)
- Kansallisarkiston määrytykset arkistoitavien erikoisaineistojen digitointiin 19.1.2021 (KA/15780/07.01.02.04.02/2020)

1 Sisältö ja tarkoitus

Asiakirja sisältää arkistoitavaksi määrättyjen eli arkistolain (831/1994) mukaan pysyvästi säilytettävien asiakirjojen digitointia koskevat Kansallisarkiston vaatimukset ja ohjeet.

Kansallisarkiston digitointivaatimusten tavoite on varmistaa kansalliseen kulttuuriperintöön kuuluvien arkistoitavien asiakirjojen tietosisällön säilyminen ja käytettävyys digitoinnin jälkeisessä digitaalisessa muodossa. Vaatimusten mukainen digitointiprosessi ja digitointitulos mahdollistavat analogisten asiakirjojen hävittämisen digitoinnin jälkeen, mikäli asiakirjojen analogista säilytysmuotoa ei ole arvioitu kulttuurihistoriallisesti arvokkaaksi. Asiakirjat, joilla on kulttuurihistoriallista arvoa, arkistoidaan digitoinnin jälkeen sekä analogisessa että digitaalisessa muodossa.

Kansallisarkistoon siirrettävien asiakirjojen tulee olla muutettu digitaaliseen muotoon tässä ohjeessa esitettyjen vaatimusten mukaisesti.

1.1 Kohderyhmä

Ohje on tarkoitettu Kansallisarkistolle sekä niille toimijoille, jotka digitoivat arkistolain (831/1994) nojalla arkistoitavaksi määrättyjä analogisia asiakirjoja.



KANSALLISARKISTO

1.2 Rajaukset

Asiakirjassa esitetyt digitointivaatimukset koskevat toimijalle jo kertyneiden arkistoitavien asiakirjojen ja muodostuneiden arkistojen digitointia. Vaatimukset eivät koske niitä toimenpiteitä, joilla asiakirja muutetaan sähköiseen muotoon silloin, kun se saapuu viranomaiselle muussa kuin sähköisessä muodossa.

1.3 Keskeiset muutokset ja viittaukset muihin asiakirjoihin

Taulukko 1. Viittaukset muihin asiakirjoihin

ASIAKIRJA	SISÄLTÖ
Arkistolaitoksenvaatimukset digitoidulle aineistolle 24.8.2016 (AL/11130/07.01.02.04.02/2008)	Arkistolaitoksen tekniset vaatimukset digitoidulle aineistolle. Kumottiin 19.1.2021 ohjeella Kansallisarkiston määräykset arkistoitavien erikoisaineistojen digitointiin (KA/15780/07.01.02/04.02/2020).
Kansallisarkiston vaatimukset hävittämisen tähtävään digitointiin 12.4.2019 (KA/3357/07.01.02.04.02/2019)	Kansallisarkiston vaatimukset hävittämiseen tähtävään digitointiin.
Kansallisarkiston määräykset arkistoitavien erikoisaineistojen digitointiin 19.1.2021	Kansallisarkiston määräykset arkistoitavien erikoisaineistojen digitointiin. Kumosi ohjeen Arkistolaitoksen vaatimukset digitoidulle aineistolle 24.8.2016 (AL/11130/07.01.02.04.02/2008)
Määräys arkistoitavien asiakirjojen muodosta 22.12.2021 (KA/15906/07.01.01.00/2021)	Määräyksen mukaan julkishallinnon toimijoille (pl. tuomioistuimet ja muut lainkäyttöelimet) 1.1.2022 alkaen muodostuville asiakirjoille ei erikseen haeta sähköistä arkistointia mahdollistavaa seulontapäätöstä, vaan se mahdollistetaan kyseisessä määräyksessä 1.1.2022.
Julkishallinnon pysyvästi säilytettävien digitoitujen ja digitoitavien asiakirjojen säilytysmuoto ja analogisen säilytysmuodon kulttuurihistoriallisen arvon arviointitietojen pysyvä säilytys 8.12.2023 (KA/15927/28.02.01/2023)	Päätöksessä määrätään ennen 1.1.2022 muodostuneet pysyvästi säilytettävät asiakirjat arkistoitaviksi yksinomaan digitaalisessa muodossa (lukuun ottamatta analogiselta säilytysmuodoltaan kulttuurihistoriallisesti arvokkaita asiakirjoja). Päätös mahdollistaa ennen 1.1.2022 muodostuneiden arkistoitavien analogisten asiakirjojen hävittämisen silloin, kun digitointi on jo tehty ennen 1.1.2022 Kansallisarkiston ajantasaisen digitointi-



KANSALLISARKISTO

	<p>vaatimusten mukaisesti tai arkistolaitoksen digitointivaatimusten 24.8.2016 (AL/11130/07.01.02.04.02/2008) perusteella (arkistolaitoksen 24.6.2016 digitointivaatimukset on sittemmin kumottu, eikä vaatimuksia voi enää soveltaa 1.1.2022 jälkeen tehtyyn tai tehtävään digitointiin).</p> <p>Lisäksi päätöksessä määrätään kulttuurihistoriallisen arvon kriteereistä. Toimijat vastaavat 1.1.2024 alkaen kulttuurihistoriallisen arvon arvioinnista noudattaen päätöksen kriteereitä ja dokumentoivat arviointitiedot.</p> <p>Arkistoitavat analogiset asiakirjat digitoidaan päätöksessä esitettyjen reunaehtojen mukaisesti, mikä mahdollistaa niiden korvaamisen digitaalisilla ilmentymillä.</p>
--	--

Nykytila valmistelun aikana

Kansallisarkiston ajantasaisia digitointivaatimuksia noudatetaan joko 22.12.2021 annetun määräyksen (KA/15906/07.01.01.00/2021) tai 8.12.2023 annetun päätöksen (KA/15927/28.02.01/2023) perusteella. Niissä määrätään, mitä Kansallisarkiston digitointivaatimuksia kulloinkin noudatetaan, kun asiakirjoja muutetaan digitaaliseen muotoon.

Kansallisarkistolla ei ole erillisiä ohjeita audiovisuaalisten aineistojen digitaaliseen muotoon muuttamiseen.

Kansallisarkiston nykyisissä digitointivaatimuksissa¹ kuvataan digitointiprosessia ja sen lopputulosta. Ohjeissa on vaatimukset muun muassa kuvatiedoston pakollisista metatiedoista, mutta varsinaista digitointiprosessia kuvataan ohjeissa vain pääpiirteittäin. Tämä johtuu siitä, että toimijat voivat tehdä määritysten mukaista digitointia käytännössä monin eri tavoin ja menetelmin.

Digitointivaatimuksissa on kohtia, jotka eivät vastaa nykyistä seulontapäätösmenttelyä tai Kansallisarkiston keskeisiä päätöksiä ja määräyksiä.

Vaatimuksissa on lisäksi kohtia ja muotoiluja, joita tulisi saatujen kysymysten perusteella selkeyttää, jotta vaatimusten soveltaminen on mahdollisimman yksiselitteistä ja helppoa.

¹ Kansallisarkiston vaatimukset hävittämiseen tähtävään digitointiin (KA/3357/07.01.02.04.02/2019), Kansallisarkiston määritykset arkistoitavien erikoisaineistojen digitointiin (KA/15780/07.01.02.04.02/2020).



KANSALLISARKISTO

Keskeiset muutokset

Tämä asiakirja (KA/20919/30.01.00/2024) korvaa ja kumoaa Kansallisarkiston aiemmin antamat vaatimukset hävittämiseen tähtävään digitointiin ja erikoisaineistojen digitointiin.² Ne on tässä asiakirjassa yhdistetty yhdeksi ohjeeksi, joka sisältää kootusti Kansallisarkiston digitointia koskevat vaatimukset ja ohjeet. Ohjeeseen on lisätty myös audiovisuaalisten aineistojen digitointivaatimukset. Ohje on lisäksi ajantasaistettu ja yhdenmukaistettu Kansallisarkiston digitointia koskevien keskeisten määräysten ja päätösten kanssa.

Päivittämisen yhteydessä on parannettu vaatimusten ymmärrettävyyttä ja luettavuutta. Digitointia koskevat tekniset vaatimukset ovat säilyneet pääosin ennallaan. Muutokset suhteessa kumottuihin ohjeisiin ja vaatimuksiin on dokumentoitu luvussa 6.

1.4 Termit ja käsitteet

Asiakirjassa on käytetty vaatimusten merkitsemiseen Internet Engineering Task Forcen määrittelemää termistöä³. Termien suomenkieliset vastineet on esitetty taulukossa 2.

Taulukko 2. Termit ja niiden käännökset

Englanti	Suomi
Must	Pitää
Must not	Ei saa
Required	Pakollinen
Should	Pitäisi
Should not	Ei pitäisi
May	Saa
Optional	Vapaaehtoinen

Asiakirjassa käytetyt käsitteet selitteineen on koottu taulukkoon 3.

Taulukko 3. Käsitteistö

Käsite	Selite
Analoginen asiakirja	Paperisessa tai muussa käsin kosketeltavassa muodossa laadittu, säilytetty ja käytettävä asiakirja (vrt. digitaalinen tieto tai asiakirja, joka muodostuu bittijonoista). Ks. Digitaalinen asiakirja .

² Kansallisarkiston vaatimukset hävittämiseen tähtävään digitointiin (KA/3357/07.01.02.04.02/2019), Kansallisarkiston määräykset arkistoitavien erikoisaineistojen digitointiin (KA/15780/07.01.02.04.02/2020).

³ RFC 2119, <https://www.ietf.org/rfc/rfc2119.txt>, viitattu 9.9.2024.



KANSALLISARKISTO

Käsite	Selite
Analogisen asiakirjan säilytysmuodon kulttuurihistoriallinen arvo	Arkistoidulle tai arkistoitavalle analogiselle asiakirjalle annettu arvo, joka on sidoksissa asiakirjan aineettomiin tai aineellisiin ominaisuuksiin. Kulttuurihistoriallisesti arvokkaaksi arvioitu asiakirja on sellaisenaan korvaamaton ja pitää siksi säilyttää ja arkistoida myös digitoinnin jälkeen sekä analogisessa että digitaalisessa muodossa. Asiakirjan analogista ilmentymää ei siis saa hävittää. Ks. Digitaalinen asiakirja, Digitointi, Hävittäminen, Ilmentymä.
Arkistointi	Arkistointi tarkoittaa osaksi asiakirjallista kulttuuriperintöä määriteltyjen asiakirjojen ja tietoaineistojen säilyttämistä sekä niiden säilyvyyden ja käytettävyyden turvaamista. Arkistoituja asiakirjoja ja tietoaineistoja säilytetään pysyvästi eli käytännössä ilman aikarajaa. Arkistointi perustuu Kansallisarkiston seulantapäätöksiin tai lainsäädäntöön. Arkistointi tapahtuu säilytysajan päättymisen jälkeen (Tiedonhallintalaki 906/2019 21 §). Arkistointi tarkoittaa myös arkistolain mukaista pysyvää säilyttämistä (8 § 3 momentti).
Arkistoyksikkö	Sarja jakautuu pienempiin osiin eli arkistoyksiköihin. Analogisissa aineistoissa arkistoyksikön tulisi olla fyysisesti hallittava kokonaisuus. Se voi olla yksittäinen asiakirja, asiakirjanippu, sidos, kartta jne. Ks. Sarja.
Asiakirja/asiakirjatieto	Asiakirja on todistusvoimaista tietoa. Toimija on tuottanut tai vastaanottanut tiedon osana tehtäviään ja säilyttää sitä tietovarantona sekä todisteena. Asiakirja on sisällön, rakenteen/ilmiänsun ja kontekstin kokonaisuus. Metatiedot sisältyvät tai liittyvät asiakirjatietoihin. (SFS-ISO15489-1: Tieto ja dokumentointi. Asiakirjahallinto. Osa 1. Käsitteet ja periaatteet). Lainsäädännössä asiakirjalla tarkoitetaan kirjallisen ja kuvallisen esityksen lisäksi sellaista käyttönsä vuoksi yhteen kuuluviksi tarkoitetuista merkeistä muodostuvaa tiettyä kohdetta tai asiaa koskevaa viestiä, joka on saatavissa selville vain automaattisen tietojenkäsittelyn tai äänen- ja kuvantoistolaitteiden taikka muiden apuvälineiden avulla. (Julkisuuslaki 621/1999, 5 §).
Audiovisuaaliset analogiset aineistot	Analogiset audiovisuaaliset aineistot ovat magneettinauhatekniikkaan pohjautuvia aineistoja, joiden analoginen käyttäminen edellyttää kyseiselle nauhatyypille (kuva tai ääni) soveltuvia laitteita. Niiden käyttö on analogisina rajoitettua myös materiaalien fyysisen haurauden vuoksi. Ovat ns. itsetuhoutuvia aineistoja, joiden elinkaari on huomattavasti lyhyempi kuin arkistokelpoisille materiaaleille tuotettujen asiakirjojen. Ks. Analoginen asiakirja, Digitaalinen asiakirja.
Digitaalinen asiakirja	Tässä ohjeessa digitaalisella asiakirjalla tarkoitetaan Kansallisarkiston digitointiohjeiden mukaan analogisesta asiakirjasta tuotettua samansisältöistä digitaalista asiakirjaa, eli analogisen asiakirjan digitaalista ilmentymää. Digitaalinen asiakirja voi koostua useista tiedostoista. Ks. Analoginen asiakirja, Ilmentymä, Tiedosto.



KANSALLISARKISTO

Käsite	Selite
Digitointi	Tässä ohjeessa digitoinnilla tarkoitetaan analogisessa muodossa säilytettävien tai arkistoitujen asiakirjojen tietosisällön muuntamista digitaaliseen muotoon. Digitointi on prosessi eli joukko toimenpiteitä, joita tehdään manuaalisesti, laittein sekä tietojärjestelmäympäristössä. Esimerkiksi skannaaminen on yksi digitointiprosessin toimenpiteistä. Digitointiprosessissa tuotetaan digitoinnin tuloksensa tallekappaleetiedosto. ⁴ Ks. Analoginen asiakirja, Ilmentymä, Digitointilaite, Skannaus, Tallekappale.
Digitointilaite	Laite, jolla analoginen asiakirja muunnetaan digitaaliseen muotoon. Ks. Digitointi.
Erikoisaineisto	Tässä ohjeessa erikoisaineistoilla tarkoitetaan karttoja, piirustuksia, valokuvia, negatiiveja, mikrofilmejä, mikrokortteja ja sidoksia, joissa sidosten tietosisältö esitetään aukeamittain Erikoisaineistojen digitointi vaatii esimerkiksi niiden koon, sisällön, materiaalien tai muiden ominaisuuksien vuoksi erilaista esivalmistelua tai erilaisia digitointilaitteita ja -määrittäjiä kuin määrämutoisten A4- tai foliokokoisten aineistojen digitointi. Erikoisaineistot on usein arkistoitu omina kokonaisuuksinaan tai kokoelminaan.
Hakemistorakenne	Tässä ohjeessa digitoinnissa muodostettujen tiedostojen järjestämiseen ja tallentamiseen käytetty rakenne. Kun tämän ohjeen mukaan digitoituja asiakirjoja siirretään Kansallisarkistoon, hakemistorakenteesta pitää pystyä muodostamaan Kansallisarkiston tietojärjestelmiin siirrettävä siirtopaketti. Ks. Siirtopaketti.
Hävittäminen	Määräajan säilytettävän asiakirjan fyysinen (ml. tekninen) hävittäminen (tuhoaminen) sekä arkistoitavan asiakirjan analogisen muodon eli analogisen ilmentymän, hävittäminen. Ks. Ilmentymä, Tuhoaminen.
Ilmentymä	Asiakirjan digitaalinen tai analoginen ilmenemismuoto. Aineistolla voi olla elinkaarensa aikana useita ilmentymiä. Perinteinen esimerkki eri ilmentymien välisestä suhteesta on aineiston analoginen ilmentymä ja siitä digitoimalla tuotettu digitaalinen ilmentymä. Ks. Analoginen asiakirja, Digitaalinen asiakirja, Digitointi.
Jakelukappale⁵	Ks. Käyttökappale.
Kuva-ala	Kuvatiedostosta silmämääräisellä tarkastelulla näkyvä informaatio.
Kuvatiedosto	Digitointiprosessissa tuotettu bittikarttakuva. Ks. Tallekappale.

⁴ Yleisemmän määritelmän mukaan digitointi on prosessi, jossa analogisessa muodossa oleva objekti muunnetaan digitaaliseen muotoon. Teknisemmin ilmaistuna kyse on analogisen signaalin tallentamisesta digitaaliseen muotoon: digitoitavasta objektista (valosta) peräisin olevat analogiset signaalitiedot muunnetaan digitaalisesti koodattuun muotoon.

⁵ FADGI, käyttökappale: <http://www.digitizationguidelines.gov/term.php?term=derivativefile> Viitattu 17.10.2024.



KANSALLISARKISTO

Käsite	Selite
Käyttökappale⁶	Käyttökappaleet ovat tallekappaleesta erilaisia käyttötarkoituksia ja käyttöliittymiä varten luotoja tiedostoja. Käytetään myös käsitettä jake-lukappale. Ks. Tallekappale .
Nauhakokonaisuus	Ääninauhan yksi puoli kokonaisuudessaan.
Pysyvä säilyttäminen	<p>Pysyvä säilyttäminen tarkoittaa henkilötietojen käsittelyssä sitä, että tie-toja säilytetään alkuperäiseen käyttötarkoitukseensa siten, ettei niitä tu-hota tai siirretä varsinaisesti arkistoon lainkaan, vaan niiden säilyttämi-selle on olemassa pysyvä peruste etujen, oikeuksien tai velvollisuuksien määrittäjänä. (Tiedonhallintalaki, HE 284/2018 vp, s. 98).</p> <p>Arkistolaissa käytetty termi pysyvä säilyttäminen tarkoittaa EU:n tieto-suoja-asetuksen (EU) 2016/679 ja tiedonhallintalain mukaista arkistoin-tia.</p>
Päälukusuunta	Mahdollistaa asiakirjan tietosisällön tulkitsemisen kuvatiedostoa kään-tämättä. Kun asiakirjassa on tietosisältöä useammassa lukusuunnassa, päälukusuunta tarkoittaa sitä suuntaa, jossa pääosa tietosisällöstä on luettavissa.
Sarja	Arkistoinnissa aineistokokonaisuus voidaan jakaa useisiin pienempiin osiin, joita kutsutaan sarjoiksi. Sarja on muodoltaan, tehtävältään ja/tai sisällöltään yhteenkuuluvien aineistojen muodostama aineiston luokka tai osakokonaisuus. Sarja jakautuu pienempiin osiin eli arkistoyksiköi-hin. Ks. Arkistoyksikkö .
Seulontapäätös	Kansallisarkiston päätös arkistolain 8, 11 ja 14 a §:ien nojalla pysyvästi säilytettävistä (arkistoitavista) asiakirjoista ja niiden säilytysmuodosta.
Siirtopaketti	Kansallisarkisto ottaa digitoitua tietoaineistoa vastaan määrämuotoi-sissa siirtorakenteissa. Mikäli tässä asiakirjassa esitettyjen vaatimusten mukaan digitoitua tietoaineistoa siirretään Kansallisarkistoon, on se tal-lennettava siirtoa varten Kansallisarkiston siirto-ohjeiden mukaiseen ha-kemistorakenteeseen ja paketoitava yhdeksi tiedostoksi. Ks. myös Hakemistorakenne .
Skannaus	Skannaus on digitoinnin vaihe, joka toteutetaan joko skannereilla tai kameroilla kuvaamalla. Ks. Digitointi .
Tallekappale⁷	Teknisiltä ominaispiirteiltään laadukkain digitoinnissa tuotettu tiedosto. Tallekappale viedään yleensä pitkäaikaissäilytysjärjestelmiin. Ks. Di-gi-tointi, Käyttökappale, Tiedosto
Tiedosto	Tiedosto muodostaa joko yksin tai muiden tiedostojen kanssa digitaali-sen asiakirjan. Tiedostot voivat olla monentyyppisiä. Tässä ohjeessa esi-merkiksi bittikarttakuva, XML-tiedosto tai video-/äänitiedosto. Ks. Il-mentymä, Digitaalinen asiakirja, Digitointi, Tallekappale .

⁶ FADGI, käyttökappale: <http://www.digitizationguidelines.gov/term.php?term=derivativefile> Viitattu 17.10.2024.

⁷ FADGI, tallekappale: <http://www.digitizationguidelines.gov/term.php?term=productionmasterfile> Viitattu 17.10.2024.



KANSALLISARKISTO

Käsite	Selite
Tuhoaminen	Ks. Hävittäminen
Tuotantovuorokausi	Vuorokausi, jonka aikana digitointilaitteella tuotetaan tiedostoja. Ks. Digitointilaitte, Tiedosto.

2 Johdanto

Tämän asiakirjan digitointivaatimukset koskevat arkistoitavaksi määrättyjen asiakirjojen digitointia. Sen tavoite on varmistaa kansalliseen kulttuuriperintöön kuuluvien asiakirjojen tietosisällön säilyminen ja käytettävyys digitoinnin jälkeisessä digitaalisessa muodossa. Vaatimuksissa on huomioitu arkistosektorilla yleisesti käytössä olevat standardit sekä muiden kansallisarkistojen laatuvaatimukset digitoinnille sekä pitkäaikaissäilytyspalveluiden (PAS-palvelut) kansalliset vaatimukset.

Analogisten asiakirjojen tietosisällön arkistointi yksinomaan digitaalisessa muodossa edellyttää sitä, etteivät digitaaliseen muotoon muuttamisessa käytetyt menetelmät heikennä asiakirjan todistusvoimaisuutta, eheyttä ja autenttisuutta. Vaatimusten mukainen digitointiprosessi ja digitointitulos mahdollistavat analogisten asiakirjojen hävittämisen digitoinnin jälkeen. Hävittäminen on peruuttamaton toimenpide, jota ennen on varmistettava digitaalisten ilmentymien todistusvoimaisuus, säilyvyys ja käytettävyys. Tämän asiakirjan vaatimuksia pitää noudattaa myös silloin, kun digitoidaan aineistoja, joiden analogisella säilytysmuodolla julkishallinnon toimija on arvioinut olevan kulttuurihistoriallista arvoa tai joille Kansallisarkisto on antanut seulontapäätöksen säilytysmuodon kulttuurihistoriallisesta arvosta.⁸ Asiakirjat, joilla on kulttuurihistoriallista arvoa, arkistoidaan digitoinnin jälkeen sekä analogisessa että digitaalisessa muodossa.

Ohjeen vaatimusten velvoittavuus perustuu Kansallisarkiston yleispäätökseen 8.12.2023 (KA/15927/28.02.01/2023) ja määräykseen 22.12.2021 (KA/15906/07.01.01.00/2021), joissa määrätään arkistoitavien asiakirjojen digitaaliseen muotoon muuttamisen vaatimuksista. Muulla tavoin digitoidut asiakirjat ovat arkistoinnin näkökulmasta toimijan omaan käyttöön tuottamia käyttökappaleita. Arkistoitavien asiakirjojen analogisia ilmentymiä ei tällöin voi digitoinnin jälkeen hävittää tai digitaaliseen muotoon muutettuja asiakirjoja siirtää Kansallisarkistoon. Ennen digitoinnin aloittamista tulisi aina varmistaa, että asiakirjoilla on Kansallisarkiston seulontapäätös, jossa on

⁸ Vuoden 2024 alusta lähtien julkishallinnon toimijat arvioivat analogisten asiakirjojen kulttuurihistoriallisen arvon ja arkistoitavan säilytysmuodon itsenäisesti noudattaen Kansallisarkiston antamaa yleispäätöstä ja arviointikriteereitä. Kansallisarkisto on antanut aiemmin seulontapäätöksiä analogisten asiakirjojen kulttuurihistoriallisesta arvosta ja säilytysmuodosta. Tästä toimintatavasta luovuttiin säilytysmuotoa koskevan yleispäätöksen (KA/15927/28.02.01/2023, 8.12.2023) myötä.



KANSALLISARKISTO

määrätty niiden arkistoinnista (pysyvästä säilyttämisestä). Tällöin arkistoitavien asiakirjojen digitoimista koskevat vaatimukset pystytään huomioimaan digitoinnin suunnittelussa ja toteutuksessa.

2.1 Vaatimusten soveltaminen eri asiakirjatyypin kohdalla

Tämä asiakirja on jaettu pääsääntöisesti A4- tai foliokokoisia sekä sidoksia, erikoisaineistoja ja audiovisuaalisia aineistoa (AV-aineistot) koskeviin vaatimuksiin ja ohjeisiin. Erikoisaineistot koostuvat kartoista, piirustuksista, valokuvista, negatiiveista, mikrofilmeistä ja mikrokorteista sekä sidoksista, joissa sidosten tietosisältö esitetään aukeamittain. Ne on usein arkistoitu omina kokonaisuuksinaan/kokoelmina arkistoiduista. Erikoisaineistoksi ei luokitella yksittäisiä asiakirjoja, jotka on arkistoitu A4- tai foliokokoisien asiakirjojen joukkoon. Audiovisuaaliset aineistot jaetaan ohjeessa ääninauhoihin ja kuvanauhoihin

Asiakirjojen kuulumista eri aineistotyyppisiin on jäsennetty taulukkoon 3. Asiakirjatyypikohtaiset vaatimukset ja niiden tarkennukset on koottu lukuun 3 (A4- tai foliokokoiset sekä sidokset ja erikoisaineistot) ja lukuun 4 (audiovisuaaliset aineistot).

Taulukko 4. Vaatimusten soveltaminen eri asiakirjatyypin yhteydessä

Aineisto/Aineistotyyppi	Pääsääntöisesti A4- tai foliokokoiset ja sidokset	Erikoisaineistot	Audiovisuaaliset aineistot
A4- ja foliokokoiset asiakirjat			
A4- ja foliokokoisien asiakirjojen joukossa olevat yksittäiset valokuvat, piirustukset ja kartat			
Sidokset			
Karttakokoelma			
Piirustuskokoelma			
Valokuvakokoelma			
Lasinegatiivit, muut negatiivit ja diat			



KANSALLISARKISTO

Aineisto/Aineistotyyppi	Pääsääntöisesti A4- tai foliokokoi- set ja sidokset	Erikoisaineistot	Audiovisuaaliset aineis- tot
Mikrofilmit ja -kortit			
Sidokset, jossa sidosten tietosisältö esitetään aukeamittain			
Ääninauhat (äänikelanauhat, C-kasetti ja minikasetti)			
Kuvanauhat (VHS-, H8- ja Betacam-vi- deonauhat)			

2.2 Digitoinnin yleiset hyvät käytänteet ja suositukset

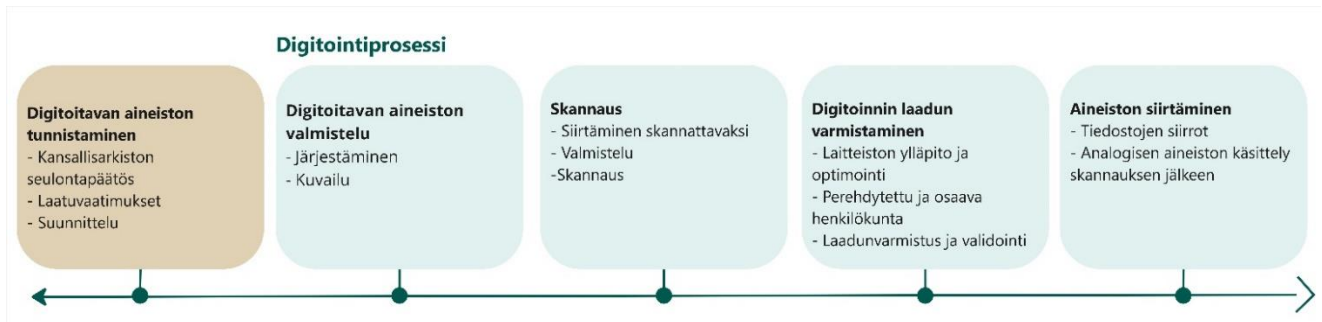
Digitoinnin tavoitteena on tuottaa analogisesta tiedosta autenttisia ja todistusvoimaisia digitaalisia ilmentymiä. Digitointi ymmärretään yleisesti prosessiksi. Tässä luvussa kuvataan digitointiprosessin ja digitoinnin laadunvarmistuksen yleisiä hyviä käytänteitä. Vaatimusten mukaista digitointia voidaan käytännössä tehdä monin eri tavoin ja menetelmin. Prosessia kuvataan luvussa siksi vain pääpiirteittäin, eikä siinä käytetä taulukon 1 termejä.

Digitointiprosessilla ei tarkoiteta esimerkiksi ainoastaan skannausta, vaan skannaus on yksi prosessin vaihe. Ennen asiakirjan skannaamista tai kuvaamista digitoitava aineisto pitää tunnistaa, valmistella ja siirtää digitoitavaksi.⁹ Skannaamisen jälkeen huolehditaan sekä digitoidun analogisen aineiston käsittelystä että digitoinnissa tuotettujen tiedostojen ja tietojen laadusta, tallentamisesta, säilymisestä sekä saattamisesta käyttöön.

⁹ Varsinaisen digitoinnin aikana aineiston alkuperäistä järjestystä ei korjata. Digitoija voi olettaa, että aineisto on ennen digitointia järjestetty riittävällä tarkkuudella. Digitoinnissa havaitut puutteet tai epäjärjestys on kuitenkin syytä dokumentoida ja kuvata metatietojen avulla.



KANSALLISARKISTO



Kuva 1. Digitaliseen muotoon muuntamisen yleisen prosessin vaiheita ovat muun muassa digitoitavan aineiston tunnistaminen, digitoitavan aineiston valmistelu, skannaus, digitoinnin laadun varmistaminen eri menetelmin ja aineiston siirtäminen.

Digitointi toteutetaan käytännössä arkistoyksiköittäin. Ennen digitoinnin aloittamista ja skannausta pitää varmistaa, että asiakirjat on luetteloitu ja kuvailtu. Digitointiprosessissa muodostuvat digitaaliset ilmentymät pitää pystyä yhdistämään luettelointi- ja kuvailutietoihin. Myös analogisen ilmentymän käsittelyketju sekä digitointivalmistelua koskevat tiedot tulisi aina dokumentoida¹⁰.

AV-aineistoista esimerkiksi ääninauhat digitoidaan laadukkailla studiokäyttöön tarkoitetuilla laitteistoilla. Niiden osalta pitää ennen digitointia varmistaa myös muun muassa nauhojen eheys sekä niin sanotun kuivakäsittelyn tarve.¹¹

Digitointilaitteet ja ohjelmistot muodostavat digitointiprosessin yhteydessä tiedostokohtaisia metatietoja, jotka dokumentoivat digitointiprosessia ja digitoinnissa muodostuneen tiedoston ominaisuuksia. Skannauksen lopputuloksena muodostunutta digitaalista tiedostoa pitää usein vielä rikastaa erilaisin metodein. Skannausohjelmistot¹² ovat yleensä kaupallisia, mutta digitoinnin jälkiprosessointi on esimerkiksi Kansallisarkistossa toteutettu hyödyntämällä myös avoimen lähdekoodin ohjelmia sekä omaa sovelluskehitystä.

Skannattuihin kuviin pitää aina liittää oikea ICC-väriprofiili, joka vastaa käytettyä värivaruutta. Tämä varmistaa, että kuvat näkyvät oikein eri laitteilla ja ohjelmistoilla. Jos väriprofiili lisätään tai

¹⁰ Jos digitoinnin valmistelun yhteydessä aineisto vaurioituu ja informaatio osittain tai kokonaan menetetään, pitää tämä dokumentoida asianmukaisesti. Esimerkiksi on hyvä dokumentoida sellainen aineisto, jota ei voida digitoida, koska sen informaatio on täysin tuhoutunut (esim. paperi haurastunut silpuksi/jauheeksi). Täysin tuhoutunut arkki, josta mitään tietosisältöä ei ole pelastettavissa digitoinnilla, korvataan A4-kokoisella huomioarkilla, jossa on teksti: "Arkistoyksikössä on tällä paikalla ollut vaurioitunut asiakirja, jonka tietosisältö on täysin tuhoutunut ennen digitointia.". Arkki asetetaan tuhoutuneen paikalle arkistoyksikköön. Tuhoutuneen arkin silput laitetaan tietoturvaroskikseen.

¹¹ Kelanauhojen kuivatuskäsittelyä käytetään tahmaavien nauhojen käsittelyyn. Lämpökäsittelyn jälkeen nauha yleensä saadaan toistettua ilman toistopään tukkeutumista. Käsittely pitää aina toteuttaa ammattilaisilla. Englanniksi aiheesta löytyy tieto termillä "Sticky Shed Syndrome".

¹² Digitointilaitteistoja voidaan yleensä ohjata useilla eri ohjelmistoilla tai käyttää laitteen mukana tulevaa ohjelmistoa.



KANSALLISARKISTO

sitä muunnetaan jälkikäteen, on varmistettava, että tämä tehdään oikein, jotta värintoisto vastaa alkuperäistä kuvaa.¹³

Migraatiot ja tiedostokonversiot, joissa tiedoston tallennusformaattia muutetaan tai versioidaan, ovat aina riskejä tiedon säilymisen näkökulmasta. On tärkeää digitoida analoginen aineisto aina nimenomaan ohjeen mukaiseen tallennusformaattiin, eikä digitointiprosessissa tuotettava tiedosto saa missään käsittelyvaiheessa olla esitettyjä vaatimuksia heikkolaatuisempi. Muunlainen toimintatapa on tämän ohjeen vaatimuksissa myös kielletty. Heikompilaatuisesta tallennusformaattista ei voi tuottaa parempilaatuista versiota myöhemmässä käsittelyssä.

Esimerkki 1: Digitointiprosessissa tuotetaan ensin sRGB-väriavaruudessa oleva tiedosto, joka myöhemmin tallennetaan Adobe RGB -väriavaruudessa. Vaikka Adobe RGB kattaa enemmän värisävyjä, siitä ei tässä tapauksessa ole hyötyä, koska kuvassa on edelleen vain ne värisävyt, jotka sRGB kattaa.

Esimerkki 2: Digitointiprosessissa tuotetaan ensin häviöllinen JPEG-kuva, joka myöhemmin tallennetaan häviöttömäksi TIFF-kuvaksi. JPEG-pakkauksen aiheuttama laadun heikkenemä ja pakkausvirheet säilyvät edelleen TIFF-kuvassa.

Yhteenvetona: Molemmissa esimerkeissä toimintatapa aiheuttaa sen, että lopullista kuvaa saatetaan luulla parempilaatuiseksi kuin se on. Oikea tapa toimia on tuottaa kuva jo skannausvaiheessa oikeilla määrityksillä. Parempilaatuisen kuvan saa aina tallennettua huonompilaatuiseksi, mutta ei toisin päin.

2.2.1 Digitoinnin laadunvarmistus

Digitoinnin laadunvarmistusta ei voi kohdentaa digitointiprosessissa tiettyihin kohtiin, vaan sitä on tehtävä prosessin kaikissa vaiheissa. Esimerkiksi skannausvaiheen laadunvarmistus voidaan jakaa vielä ennen skannausta tehtäviin toimenpiteisiin ja skannauksen jälkeiseen laadunvarmistukseen.

Skannaus voidaan tehdä joko kameralla kuvaamalla tai skannerilla skannaamalla. Luettavuuden takia tässä ohjeessa käytetään molemmista tavoista termiä skannaus. Skannausinfrastruktuurin tuottama kuvatiedosto ei saa olla epätarkka, vääristynyt tai muuten mittasuhteiltaan muuttunut, jotta

¹³ Väriavaruus on joukko värejä, jossa on määritelty kunkin värin sijainti tietyssä värjärjestelmässä (esim. sRGB tai Adobe RGB). Väriprofiili puolestaan kertoo, miten tätä väriavaruutta tulisi tulkita, jotta kuva toistuu oikein eri laitteilla ja ohjelmistoilla. Oikein määritetty väriprofiili varmistaa, että värit näkyvät mahdollisimman tarkasti ja todenmukaisesti riippumatta käyttöympäristöstä.



KANSALLISARKISTO

tietosisällön tulkinta ei vaarannu. Kuvan laadun lisäksi laiteinfrastruktuurissa on varmistettava, että analogiset asiakirjat muunnetaan digitaaliseen muotoon tietosisällöltään täydellisinä.

Skannausprosesseissa pitää välttää esimerkiksi kahden tai useamman päällekkäisen analogisen asiakirjan menemistä laitteen läpi. Tämä on vaarana, kun käytetään läpisyöttöskannereita, avoraskannereita tai muita skannausratkaisuja, joissa asiakirjoja skannataan "massana". Tämän tyyppisissä skannereissa pitää olla ns. tuplasyötön tunnistus (engl. multifeed detection), jonka toiminta on varmistettu. Yleisin tuplasyötön tunnistuksen tapa on ultraäänisensorit, jotka havaitsevat hyvin herkästi päällekkäisten asiakirjojen arkit.

Ennen skannauksen aloittamista optimoidaan skannausinfrastruktuurin suorituskyky. Tavoite on, että skannauksessa tuotetaan aina niin laadukas tiedosto kuin käytössä olevalla teknisellä kokoonpanolla on mahdollista tuottaa. Optimoinnin jälkeen infrastruktuurin suorituskykyä tulisi tarkkailla suunnitellusti, jotta prosessissa tuotettavien digitaalisten ilmentymien laatu säilyy tasaisena. Laadun tarkkailua varten tarvitaan yleensä mittataulu¹⁴, mittataulun referenssiarvot¹⁵ ja analysointiohjelmisto.

Skannauksen jälkeisen validoinnin tavoitteena on varmistua siitä, että digitointi ja digitointitulos täyttävät niille määritellyt vaatimukset. Arkistoitavaksi määrättyjen asiakirjojen kohdalla kyse on tämän ohjeen vaatimuksista. Validointi voidaan toteuttaa otannoilla, joiden määrä ja koko riippuvat skannausprosessin luotettavuudesta ja laadunvalvonnan kokonaistoteutuksesta eli siitä, missä prosessin vaiheissa ja millä tavoilla valvontaa toteutetaan ja miten luotettavaa se kokonaisuutena arvioiden on. Laatupoikkeamat pyritään havaitsemaan silmämääräisellä tarkastelulla.

3 A4- tai foliokokoisia aineistoja sekä sidoksia ja erikoisaineistoja koskevat vaatimukset

Tämä luku sisältää digitointiprosessissa tuotettua kuvatiedostoa ja sen prosessointia koskevat vaatimukset silloin, kun digitoidaan pääsääntöisesti A4- tai foliokokoisia aineistoja sekä sidoksia tai erikoisaineistoja. Erikoisaineistoiksi ei tässä ohjeessa luokitella yksittäisiä asiakirjoja, jotka on arkistoitu A4- tai foliokokoisien asiakirjojen joukkoon.

Aineistotyyppit on eritelty [taulukossa 4](#).

¹⁴ Mittataululla tarkoitetaan esimerkiksi Universal Test Target (UTT) -mittataulua tai IT8-mittataulua. Niillä mitataan kuvasta muun muassa sen todellista resoluutiota, värejä, vääristymiä, sävyntoistoa jne.

¹⁵ Referenssiarvoilla viitataan esimerkiksi niihin arvoihin, joihin mittaustulosta verrataan. Usein myös mittataulut voivat olla värimitattuja, jolloin saadaan mittataulukohdaiset arvot. Analysointiohjelmisto voi olla automaattinen ohjelma, joka osaa lukea mittataulusta skannatun kuvan ja siitä eri mitta-alueet.



KANSALLISARKISTO

Luvussa 3.1.1 esitetään aineistotyyppien yhteiset vaatimukset. Aineistotyyppiin riippuvaiset vaatimukset on koottu lukuun 3.2.

3.1 Vaatimusten mukainen digitointiprosessi ja digitointitulos

Digitointiprosessissa tuotetaan tallekappaletiedosto eli kuvatiedosto, joka on aineistotyyppikohtaisten vaatimusten mukaisesti joko häviöttömästi pakatussa TIFF-muodossa tai häviöllisesti pakatussa JPEG-muodossa.¹⁶

Digitointiprosessissa tuotettu tiedosto ja sen metatiedot dokumentoivat prosessia. Digitointiprosessi **pitää** dokumentoida tässä ohjeessa kuvatuilla tavoilla ja tarkkuudella. Dokumentointi tarkoittaa sitä, että skannauksesta ja kuvatiedostojen mahdollisesta käsittelystä tallennetaan näitä toimenpiteitä dokumentoivat metatiedot. Metatietona **saa** tallentaa myös muita prosessiin kuuluneita toimenpiteitä.¹⁷

Ennen skannausta digitoinnissa käytetyn infrastruktuurin suorituskyky **pitää** optimoida käyttämällä tarkoitukseen soveltuvia mittatauluja ja ohjelmistoja. Optimoinnin jälkeen infrastruktuurin tuottamien kuvatiedostojen laatu **pitäisi** todentaa kerran tuotantovuorokaudessa käyttämällä tähän tarkoitukseen tarkoitettuja mittatauluja.

Digitointiprosessissa **pitää** varmistua siitä, että digitoitavaksi tarkoitettu kokonaisuus digitoidaan:

- kokonaisuutena
- arkiston alkuperäisen järjestyksen mukaisesti
- sisällöllisesti täydellisenä.

Digitoitavaan kokonaisuuteen kuuluvat analogiset asiakirjat **pitää** digitoida siten, ettei mitään informaatiota¹⁸ jää teknisen tai toiminnallisen virheen takia muuntamatta digitaaliseen muotoon. Toisin sanoen, jokaisesta digitoitavaan kokonaisuuteen liittyvästä yksittäisestä kuvatiedostosta **pitää** silmämääräisellä tarkastelulla olla saatavissa kaikki sama informaatio kuin sen analogisesta vastineesta.

¹⁶ Tallekappale on teknisiltä ominaispiirteiltään laadukkain digitointiprosessissa tuotettu kuvatiedosto. Se vie usein yleisesti pitkäaikaissäilytysjärjestelmiin. Vrt. tallekappaleesta muodostetut käyttökappaleet eli erilaisia käyttötarkoituksia varten muodostetut tiedostot.

¹⁷ Digitointia ja muodostuvan kuvatiedoston ominaisuuksia dokumentoivia metatietoja muodostavat pääasiassa digitointiprosessin yhteydessä prosessissa käytettävät laitteet ja ohjelmistot.

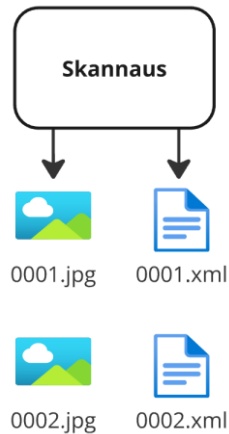
¹⁸ Tarkoitus on, ettei mitään tietoja katoa, jotta digitoitu ilmentymä on autenttinen, todistusvoimainen ja eheä. Autenttisuus, todistusvoimaisuus ja eheys vaativat, että analoginen ilmentymä digitoidaan tietosisällöltään sellaisena kuin se on.



KANSALLISARKISTO

Monisivuisen asiakirjan jokainen sivu **pitäisi** digitoida molemmilta puolilta¹⁹. Kummastakin puolesta muodostuu oma kuvatiedostonsa.

Kuvan metatiedot²⁰ **pitää** tallentaa kuvatiedostoon. Niistä sekä kuvatiedoston muodostumista ja prosessointia koskevista teknisistä metatiedoista **pitää** lisäksi tuottaa kutakin kuvatiedostoa vastaava tiedostokohtainen MIX-metatietoskeeman (versio 2.0) mukainen XML-tiedosto.



Kuva 2. Jokaisesta kuvatiedostosta tuotetaan MIX-metatietoskeeman (versio 2.0) mukainen tiedostokohtainen XML-tiedosto.

Kuvatiedosto **ei saa** sisältää mitään sellaisia elementtejä, joita ei ole analogisessa vastineessa. Poikkeuksen muodostavat samaan kuvaan tarpeen mukaan skannattavat kuvatiedoston värejä, harmaasävyjä, mittasuhteita ja resoluutiota todentavat skannaustekniset mittataulut sekä aineiston käytön kannalta oleelliset tietolaput. Tietolapuilla voidaan kertoa asiakirjan tietosisältöön tai digitoinnissa tehtyihin ratkaisuihin liittyvistä asioista.

Esimerkki. Jos asiakirja on huomattavasti A0-kokoa suurempi, se voidaan joutua skannaamaan osissa. Kun esimerkiksi suurikokoinen kartta on skannattu kahdessa osassa, tietolapuilla oleva jaottelutapa kertoo, että kyseessä on osa 1/2, osa 2/2 jne.

Digitointiprosessissa **ei saa** poistaa merkintöjä sisältäviä sivuja.

Digitointiprosessissa tuotettujen kuvatiedostojen **pitää** olla käännetty päälukusuuntaan.

¹⁹ Mikäli toimija on täysin varma, että digitoinnin kohteena olevan tietoaineistokokonaisuuden kääntöpuolet ovat kokonaisuudessaan tyhjiä, voidaan ne jättää digitoimatta. Kansallisarkisto digitoi tietoaineistot lähtökohdaisesti sekä etu- että kääntöpuolilta.

²⁰ Esimerkiksi EXIF-tiedot ja muut vastaavat tiedot.



KANSALLISARKISTO

Digitointiprosessissa tuotettuja kuvatiedostoja **saa** kääntää niiden skannauksen jälkeen vain 90 asteen välein. Jos kuvatiedostoja käsitellään skannauksen jälkeen, yksityiskohtainen kuvankäsittelyhistoria **pitäisi** tallentaa ainakin kuvatiedostojen metatietoihin ja mahdollisuuksien mukaan myös kuvatiedoston muodostumista kuvaileviin XML-muotoisiin tietoihin (MIX-metatiedot).

3.1.1 Kuvatiedostoa ja sen ominaisuuksia koskevat vaatimukset

Digitointiprosessissa tuotettava kuvatiedosto **pitää** tallentaa 24-bittisenä RGB-kuvana. Aineistotyyppikohtainen poikkeus ovat harmaasävyiset mikrofilmit, joiden kuvatiedostot **pitää** tallentaa 8-bittisenä.

Kuvatiedosto **ei saa** missään käsittelyvaiheessa olla tässä luvussa esitettyjä vaatimuksia heikkolaatuisempi.

Kuvatiedosto **pitää** aineistotyyppikohtaiset vaatimukset huomioiden tallentaa joko häviöttömästi pakatussa TIFF-muodossa tai häviöllisesti pakatussa JPEG-muodossa. Sitä **ei saa** tallentaa molemmissa muodoissa.

Mikäli prosessissa ensimmäisenä tuotettu kuvatiedosto on pakattu, lähtökohtaisesti sitä **ei pitäisi** käsitellä ensimmäisen tallennuskerran jälkeen ja tallentaa tämän jälkeen uudelleen. Toisen tallennuksen **saa** tehdä vain, kun se on digitointiprosessissa välttämätöntä, eikä pakkaamatonta formaattia esimerkiksi laitteiston kapasiteetin takia pystytä käyttämään.

Jos toinen tallennuskerta joudutaan tekemään, kuvatiedosto **ei saa** missään vaiheessa olla näissä vaatimuksissa määriteltyä heikkolaatuisempi.²¹

Jos tuotettua kuvatiedostoa käsitellään skannaamisen jälkeen, kuvatiedoston metatietoihin **pitää** dokumentoida ohjelmat, joilla sitä on käsitelty siinä järjestyksessä, missä käsittely on tapahtunut. Tiedot merkitään kuvatiedoston EXIF-metatietojen IFD0:ProcessingSoftware (0x000b) -tagiin. Arvot erotetaan toisistaan puolipisteellä (;):

²¹ Tiedosto pitäisi toisella tallennuskerralla tallentaa muun muassa ilman subsampling-menetelmää²¹, jotta laadun heikkeneminen on mahdollisimman vähäistä. Subsampling (engl. jpeg subsampling) tarkoittaa prosessia, jossa kuvaa pakatessa vierekkäisten pikselien värejä (ei luminanssia, joka säilyy entisellään) sulautetaan toisiinsa. Tällöin osa alkuperäisestä väritarkkuudesta katoaa. Esimerkiksi ImageMagick-ohjelmassa subsampling-toiminnon saa pois päältä "sampling-factor 4:4:4" -parametrin avulla.



KANSALLISARKISTO

Esimerkki. <IFD0:ProcessingSoftware>Photoshop 5.0; ImageMagick 12.4; Kasittelyohjelma v2 </IFD0:ProcessingSoftware>.

MIX-tiedoston ChangeHistory-elementtiin pitäisi tällöin lisäksi päivittää kuvatiedoston alkuperäiset tiedot, kuvatiedostolle tehdyt toimenpiteet ja kuvatiedoston nykyiset tiedot. Tämän asiakirjan liitteenä on toimintatapaa avaava esimerkkiedosto (ChangeHistory.xml).

Kuvatiedostossa **pitää** olla tiedostomuodosta riippuen taulukossa 5 tai 6 esitetyt pakolliset tiedot koneluettavassa muodossa. Kuvatiedosto **saa** sisältää myös sellaisia metatietokenttiä, joita ei ole esitetty taulukoissa.

Vaadittu arvo -sarakkeen arvojen selitykset:

- **Aineistotyyppi riippuvainen:** Vaadittu arvo tarkistetaan oikean aineistotyyppin kohdalta (luku 3.2).
- **Kuvatiedostokohtainen:** Arvot vaihtelevat kuvatiedostoittain. Muut määrytykset saattavat kuitenkin heijastua näihin arvoihin. Esimerkiksi "Kuvan leveys" ja "Kuvan korkeus" ovat kuvatiedostokohtaisia, vaikka ne ovatkin riippuvaisia käytetystä resoluutiosta.
- **Tuotantoprosessi riippuvainen:** Arvot pysyvät samoina koko prosessin ajan eli ne eivät ole riippuvaisia aineistotyypeistä tai vaihtelevat kuvatiedostoittain. Esimerkkejä tällaisista arvoista ovat "Digitaalisen kuvatiedoston tekijä" ja "Käytetty skannerimalli".

Taulukko 5. JPEG-kuvatiedoston vaatimukset

ELEMENTTI	TARKENNE	VAADITTU ARVO	METATIETOSKEEMA	Metatietokenttä (ID) / Exiftool Group:Tag
Formaatti	JPEG	image tai jpeg		MIME Type / File:MIMETYPE
Versio	JPEG part 1 versio	1.00 tai 1.01 tai 1.02		
Kuvan nimi	Kuva nimetään neljällä numerolla 0001.jpg, 0002.jpg. tai viidellä jos kuvia on 10 000 tai enemmän. Ks. luku 5.			
Kuvatiedoston koko	Tiedoston koko tavuina ja kokonaislukuina. Kansallisarkistoon siirrettävien kuvien enimmäiskoko	Kuvatiedostokohtainen		FileSize / System:FileSize



KANSALLISARKISTO

	pikseleissä on 65535x65535.			
Väritila	Kuvatiedoston väritila	Aineistotyyppiiripuvainen	Exif.Image	PhotometricInterpretation (262) / IFD0:PhotometricInterpretation
ICC-profiili	Kuvatiedoston metatietoihin tallennettu (embedded) väriprofiili.	Aineistotyyppiiripuvainen	ICC	profileDescription / ICC_Profile:ProfileDescription
Bittisyvyys	Bittien määrä pikselin kanava-arvossa	8	Exif.Image	BitsPerSample (258) / IFD0:BitsPerSample
Näytteiden määrä	Kanava-arvojen määrä pikselissä	Aineistotyyppiiripuvainen	Exif.Image	SamplesPerPixel (277) / IFD0:SamplesPerPixel
Tiedoston pakkaaminen	JPEG	6 = JPEG (old-style) tai 7 = JPEG	Exif.Image	Compression (259) / IFD0:Compression
JPEG-laatu	JPEG-pakkauksen laatu asteikolla 0–100 % (tiedolle ei ole exifkenttää, mutta jpeg:n PITÄÄ olla tallennettu tällä laadulla.)	90 %		
Kuvan leveys	Kertoo kuvan leveyden pikselien määrällä per rivi	Kuvatiedostokohtainen	Exif.Image	ImageWidth (256) / IFD0:ImageWidth
Kuvan korkeus	Kertoo kuvan korkeuden pikselirivien määrällä kuvassa	Kuvatiedostokohtainen	Exif.Image	ImageLenght (257) / IFD0:ImageHeight
Digitaalisen kuvatiedoston tekijä	Organisaatio, joka on luonut kuvatiedoston analogisesta ilmentymästä	Tuotantoprosessiiripuvainen	Exif.Image	Artist (315) / IFD0:Artist
Digitointilaitte (skannaus tai kuvaus)	Kertoo minkä valmistajan laitteella analoginen ilmentymä on muutettu digitaaliseen muotoon (valmistajan nimi)	Tuotantoprosessiiripuvainen	Exif.Image	Make (271) / IFD0:Make
Digitointilaitteen malli (skannaus tai kuvaus)	Tarkentaa digitointilaitetta kertomalla valmistajan mallin nimen	Tuotantoprosessiiripuvainen	Exif.Image	Model (272) / IFD0:Model
Digitoinnissa käytetyn laitteen sarjanumero	Yksilöi käytetyn laitteen	Tuotantoprosessiiripuvainen	Exif.Image	CameraSerialNumber (50735) / IFD0:CameraSerialNumber
Kuvatiedoston luomisessa käytetty ohjelma	Sovellus ja versio, jolla digitaalinen tiedosto on luotu	Tuotantoprosessiiripuvainen	Exif.Image	Software (305) / IFD0:Software



KANSALLISARKISTO

Kuvatiedoston luontipäivämäärä ja aika	Ilmaistaan muodossa: YYYY:MM:DD HH:MM:SS	Kuvatiedostokohtainen	Exif.Image	DateTime (306) / IFD0:ModifyDate
Lukusuunta	Tiedoston lukusuunta. Lukusuunta ei ota kantaa kuvan tietosisällön lukusuuntaan, vaan tässä ilmaistaan tiedoston lukusuunta.	Tuotantoprosessi-riippuvainen/ Kuvatiedostokohtainen	Exif.Image	Orientation (274) / IFD0:Orientation
Resoluution mittayksikkö	Mittayksikkö, jota käytetään tulkittaessa X ja Y resoluutiota	2 = inch	Exif.Image	ResolutionUnit (296) / IFD0:ResolutionUnit
XResoluutio	Pikselien määrä resoluution mittayksikköä kohden kuvan leveys-suunnassa.	Aineistotyyppi-riippuvainen	Exif.Image	Image.XResolution (282) / IFD0:XResolution
YResoluutio	Pikselien määrä resoluution mittayksikköä kohden kuvan korkeussuunnassa.	Aineistotyyppi-riippuvainen	Exif.Image	Image.YResolution (283) / IFD0:YResolution
Kuvatiedoston käsittelyohjelma	Mikäli digitointiprosessissa luodaan ensin pakkaamaton tiedosto, jota käsitellään skannauksen jälkeen, tallennetaan käsitelyohjelman nimi ja versio. Jos käsittelyohjelmia on useita, kirjoitetaan ne siinä järjestyksessä, kun niitä on käytetty. Erotusmerkkinä käytetään puolipistettä.	Tuotantoprosessi-riippuvainen	Exif.Image	Image.Processing-Software (11) / IFD0:ProcessingSoftware

Taulukko 6. TIFF-kuvatiedoston vaatimukset

Elementti	Tarkenne	Vaadittu arvo	Metatieto-skeema	Metatietokenttä (ID) / Exiftool Group:Tag
Formaatti	TIFF	image/tiff		MIME Type / File:MIMEType
Versio	TIFF-tiedoston vaadittu versio	6.0		
Kuvan nimi	Kuva nimetään neljällä numerolla esim. 0001.tif, 0002.tif tai viidellä jos kuvia on 10 000 tai enemmän Ks. luku 5.			
Kuvatiedoston koko	Tiedoston koko tavuina ja kokonaisluokuna. Kansallisarkistoon siirrettävien kuvien enimmäiskoko pikseleissä on 65535x65535.	Kuvatiedostokohtainen		FileSize / System:FileSize



KANSALLISARKISTO

Elementti	Tarkenne	Vaadittu arvo	Metatieto-skeema	Metatietokenttä (ID) / Exiftool Group:Tag
Väritila	Kuvatiedoston väritila	Aineistotyyppi riippuvainen	Exif.Image	PhotometricInterpretation (262) / IFD0:PhotometricInterpretation
ICC-profiili	Kuvatiedoston metatietoihin tallennettu (embedded) väriprofiili.	Aineistotyyppi riippuvainen	TIFF tag, private	ICC Profile (34675) / ICC_Profile
Bittisyvyys	Bittien määrä pikselin kanava-arvossa	8	Exif.Image	BitsPerSample (258) / IFD0:BitsPerSample
Näytteiden määrä	Kanava-arvojen määrä pikselissä	Aineistotyyppi riippuvainen	Exif.Image	SamplesPerPixel (277) / IFD0:SamplesPerPixel
Tiedoston pakkaaminen	Käytetty pakkaus-tyyppi	Aineistotyyppi riippuvainen	Exif.Image	Compression (259) / IFD0:Compression
Kuvan leveys	Kertoo kuvan leveyden pikselien määrällä per rivi	Kuvatiedostokohtainen	Exif.Image	ImageWidth (256) / IFD0:ImageWidth
Kuvan korkeus	Kertoo kuvan korkeuden pikselirivien määrällä kuvassa	Kuvatiedostokohtainen	Exif.Image	ImageLenght (257) / IFD0:ImageHeight
Digitaalisen kuvatiedoston tekijä	Organisaatio, joka on luonut kuvatiedoston analogisesta ilmentymästä	Tuotantoprosessi riippuvainen	Exif.Image	Artist (315) / IFD0:Artist
Digitointilaitte (skannaus tai kuvaus)	Kertoo minkä valmistajan laitteella analoginen ilmentymä on muutettu digitaaliseen muotoon (valmistajan nimi)	Tuotantoprosessi riippuvainen	Exif.Image	Make (271) / IFD0:Make
Digitointilaitteen malli (skannaus tai kuvaus)	Tarkentaa digitointilaitetta kertomalla valmistajan mallin nimen	Tuotantoprosessi riippuvainen	Exif.Image	Model (272) / IFD0:Model
Digitoinnissa käytetyn laitteen sarjanumero	Yksilöi käytetyn laitteen	Tuotantoprosessi riippuvainen	Exif.Image	CameraSerialNumber (50735) / IFD0:CameraSerialNumber
Kuvatiedoston luomisessa käytetty ohjelma	Sovellus ja versio, jolla digitaalinen tiedosto on luotu	Tuotantoprosessi riippuvainen	Exif.Image	Software (305) / IFD0:Software
Kuvatiedoston luontipäivämäärä ja aika (skannauspäivämäärä)	Ilmaistaan muodossa: YYYY:MM:DDTHH:MM:SS	Kuvatiedostokohtainen	Exif.Image	DateTimeOriginal (36867) / ExifIFD:DateTimeOriginal
Lukusuunta	Tiedoston lukusuunta. Lukusuunta ei ota	Tuotantoprosessi riippuvainen/	Exif.Image	Orientation (274) / IFD0:Orientation



KANSALLISARKISTO

Elementti	Tarkenne	Vaadittu arvo	Metatieto-skeema	Metatietokenttä (ID) / Exiftool Group:Tag
	kantaa kuvan tietosivallön lukusuuntaan, vaan tässä ilmaistaan tiedoston lukusuunta.	Kuvatiedostokohtainen		
Resoluution mittayksikkö	Mittayksikkö, jota käytetään tulkitessa X ja Y resoluutiota	2 = inch	Exif.Image	ResolutionUnit (296) / IFD0:ResolutionUnit
XResoluutio	Pikselien määrä resoluution mittayksikköä kohden kuvan leveysuunnassa.	Aineistotyyppi riippuvainen	Exif.Image	Image.XResolution (282) / IFD0:XResolution
YResoluutio	Pikselien määrä resoluution mittayksikköä kohden kuvan korkeusuunnassa.	Aineistotyyppi riippuvainen	Exif.Image	Image.YResolution (283) / IFD0:YResolution
Tavujärjestys	Määrittelee järjestyksen, jossa tavut luetaan, kun tiedosto avataan.	big endian tai little endian		ByteOrder
Kuvatiedoston käsittelyohjelma	Mikäli digitointiprosessissa luodaan ensin pakkaamaton tiedosto, jota käsitellään skannauksen jälkeen, tallennetaan käsittelyohjelman nimi ja versio. Jos käsittelyohjelmia on useita, kirjoitetaan ne siinä järjestyksessä, kun niitä on käytetty. Erotusmerkkinä käytetään puolipistettä.	Tuotantoprosessi riippuvainen	Exif.Image	Image.Processing-Software (11) / IFD0:ProcessingSoftware

3.1.2 Kuvatiedoston prosessointia kuvaavat MIX-metatiedot ja niiden rakenne

Tässä luvussa määritellyt pakolliset tekniset metatiedot dokumentoivat kuvatiedoston syntyhistoriaa, joka todentaa osaltaan myös prosessissa muodostuvan digitaalisen ilmentymän autenttisuutta eli alkuperää ja aitoutta

Tekniset metatiedot **pitää** tuottaa MIX-metatietoskeeman (versio 2.0) mukaiseksi XML-tiedostoksi.²² Jokaisella tuotetulla kuvatiedostolla **pitää** olla sitä vastaava XML-tiedosto.

MIX-skeeman kentät ovat joko **dataelementtejä**, jotka sisältävät tietyn arvon tai **säiliöitä**, jotka sisältävät yhden tai useamman dataelementin. Säiliöt voivat sisältää myös toisia säiliöitä dataelementteineen. Kaikki tässä ohjeessa vaaditut MIX-metatiedot ovat dataelementtejä eli ne sisältävät

²² The Library of Congress » Standards » MIX. Kongressin kirjaston verkkosivu <http://www.loc.gov/standards/mix/> Viitattu 17.10.2024.



KANSALLISARKISTO

vain arvon. Metatiedot on eritelty taulukossa 7, jossa on kerrottu MIX-kentän nimen lisäksi sen tarkoitus vapaasti suomennettuna. Velvoite-sarake kertoo kentän ja sen skeeman mukaisen pakollisuuden seuraavasti:

- P = pakollinen – tämä tieto **pitää** kuvata
- V = Vapaaehtoinen – tämä tieto **pitäisi** kuvata, mutta se ei ole pakollista

Taulukko 7. Kuvatiedostoa ja sen prosessointia kuvaavat metatiedot

MIX-kentän nimi	Kentän tarkoitus	Velvoite
objectIdentifierType	Dataelementti, joka määrittää sen järjestelmän tai verkkotunnuksen, jossa digitaalisen objektin yksilöivä ID (ks. objectIdentifierValue) on uniikki.	P
objectIdentifierValue	Digitaalisen objektin identifioiva merkkisarja.	P
fileSize	Tiedoston koko tavuissa, esimerkiksi 72839.	P
formatName	Tiedoston formaatti. Vaadittu arvo: image/jpeg tai image/tiff	P
formatVersion	Tiedoston versio, esimerkiksi 1.01	V
byteOrder	Dataelementti, joka määrittää tavujen tallennusjärjestyksen. Vaadittu arvo: big endian tai little endian.	P
compressionScheme	Käytetty pakkaus. Vaadittu arvo: JPEG (kun tallekappale on jpeg-formaatissa) tai LZW (kun tallekappale on TIFF-formaatissa).	P
compressionRatio	Dataelementti, joka kertoo käytetyn pakkauksen tason. Ilmaistaan käyttäen numerator "90" ja denominator "100".	P (vain JPEG)
messageDigestAlgorithm	Dataelementti, joka identifioi algoritmin, jolla messageDigest-kentän arvo on luotu. Kentän arvo on joko: MD5, SHA-1, SHA256, SHA384 tai SHA512	P
messageDigest	messageDigestAlgorithm-kentän määrittämän algoritmin tuottama merkkisarja, esimerkiksi e8064dc0	P
imageWidth	Kuvan leveys pikseleissä, esimerkiksi 1330.	P
imageHeight	Kuvan korkeus pikseleissä, esimerkiksi 1600.	P
colorSpace	Dataelementti, joka määrittää kuvan väriavaruuden. Vaadittu arvo: RGB.	P
iccProfileName	Dataelementti, joka määrittää yleisesti käytetyn ICC-profiilin nimen. Vaadittu arvo JPEG-tiedostossa: sRGB. Vaadittu arvo TIFF-tiedostossa: sRGB, ecRGB v2, ProPhoto RGB tai AdobeRGB (1998)	P



KANSALLISARKISTO

MIX-kentän nimi	Kentän tarkoitus	Velvoite
iccProfileVersion	Dataelementti, joka kertoo käytetyn ICC-profiilin version, esimerkiksi v4 [eli sRGB v4]	P
iccProfileURI	Jos ICC-profiili ei ole hyvin dokumentoitu, profiilin URI/URN tallennetaan tähän kenttään.	V
dateTimeCreated	Dataelementti, joka kertoo kuvatiedoston luontiajan. Ilmaistaan muodossa: YYYY-MM-DDTHH:MM:SS	P
imageProducer	Dataelementti, joka identifioi digitaalisen objektin luoneen organisaation.	P
scannerManufacturer	Dataelementti, joka kertoo skannauksessa käytetyn laitteen valmistajan nimen.	P
scannerModelName	Dataelementti, joka kertoo käytetyn digitointilaitteen mallin nimen.	P
scannerModelNumber	Dataelementti, joka tarkoittaa digitointilaitteen mallin nimeä sen tyyppinumeroilla.	P
scannerModelSerialNo	Digitointilaitteen sarjanumero, jonka avulla tietty laite on mahdollista yksilöidä.	P
scanningSoftwareName	Käytetyn skannausohjelmiston nimi.	P
scanningSoftwareVersionNo	Käytetyn skannausohjelmiston version numero.	P
orientation	Dataelementti, joka kertoo kuvan lukusuunnan.	P
samplingFrequencyUnit	Dataelementti, joka kertoo mittayksikön, jota käytetään tulkittaessa X ja Y resoluutiota. Vaadittu arvo "in."	P
xSamplingFrequency	Pikselien määrä resoluution mittayksikkö kohden leveysuunnassa. Vaadittu arvo: 300	P
ySamplingFrequency	Pikselien määrä resoluution mittayksikkö kohden pystysuunnassa. Vaadittu arvo: 300	P
bitsPerSampleValue	Dataelementti, joka määrittelee jokaisessa kanavassa olevien bittien määrän. Vaadittu arvo: 8	P
bitsPerSampleUnit	Dataelementti, joka määrittää bittien tulkintatavan. Arvo on joko integer tai floating point.	P
samplesPerPixel	Dataelementti, joka määrittää kanava-arvojen määrän pikselissä. Vaadittu arvo: 3	P



KANSALLISARKISTO

MIX-kentän nimi	Kentän tarkoitus	Velvoite
targetType	Dataelementti, joka kertoo, onko skannauksen laatua todentava mittataulu osa kuvaa vai skannattu erilliseen kuvaan.	V
targetManufacturer	Dataelementti, johon merkitään mittataulun valmistaja.	V
targetName	Dataelementti, joka identifioi käytetyn mittataulun nimen.	V
targetNo	Dataelementti, joka sisältää käytetyn mittataulun sarjanumeron.	V
externalTarget	Dataelementti, joka kertoo, mistä TargetID-säiliön yksilöidyn mittataulun digitaalinen kuva löytyy.	V
performanceData	Dataelementti, joka kertoo mistä TargetID-säiliön yksilöimän mittataulun mittausdata löytyy.	V

3.2 Aineistotyyppikohtaiset vaatimukset

Tästä luvusta löytyvät taulukoihin 5 ja 6 aineistotyyppiin riippuvaisiksi merkityt arvot.

3.2.1 A4-kokoiset ja foliokokoiset asiakirjat sekä sidokset

Taulukko 8. A4-kokoisten ja foliokokoisten asiakirjojen sekä sidosten määrittäykset

Elementti	Vaadittu arvo
Tallekappaleen formaatti	TIFF 6.0 tai JPEG part 1 versio 1.00 tai 1.01 tai 1.02
Tallekappaleen pakkaaminen	TIFF = LZW tai JPEG = JPEG-laatusaso 90 %
Resoluutio	300 ppi
Väritila	2 = RGB
Bittisyvyys	8 bittiä kanava, 3 kanavaa (24 bit)
Tallekappaleen ICC-profiili	sRGB, eciRGB v2, ProPhoto RGB tai AdobeRGB (1998)

3.2.2 Kartat ja piirustukset

Kartat ja piirustukset **pitää** digitoida kuvatiedostoiksi siten, että syntyvä kuvatiedosto on fyysisiltä mitoiltaan (n cm x n cm) 1:1-koossa analogisen asiakirjan kanssa resoluutiolla 300 ppi.



KANSALLISARKISTO

Digitoitavan analogisen asiakirjan tietosisällön eheyden varmistamiseksi asiakirja **pitää** digitoida siten, että kuva-ala ylittää sivujen reunat. Asiakirjan **pitää** olla kuvaushetkellä tasainen. Mikäli tasaisuuden saavuttamiseksi käytetään lasia tai muuta apuvälinettä, tämä **pitää** huomioida värinhallinnassa.

Kartat ja piirustukset **pitää** digitoida värillisiksi kuvatiedostoiksi (RGB, 24 bit).

Skannauslaitteisto **pitäisi** lähtökohtaisesti mitoittaa aineistolle sopivaksi, siten että kuva-ala on riittävä. Jos skannattava asiakirja on erittäin suuri eikä mahdu kokonaan esimerkiksi A0-kokoiseen skanneriin, sen **saa** skannata useassa osassa. Skannatut osat **pitää** tällöin merkitä tietolapulla, jossa jakotapa kerrotaan selkeästi.²³ Tietolappu **ei saa** peittää tietosisältöä. Jokaiseen osaan **pitää** jättää lisäksi riittävä marginaali, jotta asiakirjaa on helppo lukea myös jakokohdasta.

Osiin jakaminen voi olla tarpeen myös silloin, kun kuvatiedostosta tulee teknisesti liian suuri. Kansallisarkistoon siirrettävien kuvien enimmäiskoko pikseleissä on 65535x65535.

Karttojen ja piirustusten digitointia koskevat määrytykset on koottu taulukkoon 9.

Taulukko 9. Karttojen ja piirustusten määrytykset

Elementti	Vaadittu arvo
Tallekappaleen formaatti	TIFF 6.0
Tallekappaleen pakkaaminen	Pakkaamaton tai LZW
Resoluutio	300 ppi
Väritila	2 = RGB
Bittisyvyys	8 bittiä kanava, 3 kanavaa (24 bit)
Tallekappaleen ICC-profiili	eciRGB v2, ProPhoto RGB, AdobeRGB (1998)

3.2.3 Sidokset, joissa tietosisältö esitetään aukeamittain

Mikäli sidosten sisältämä tietosisältö esitetään aukeamittain, **pitää** digitoinnissa soveltaa tässä luvussa esitettyjä vaatimuksia. Tällaisia tietosisältöjä voivat olla esimerkiksi sivun ylittävät taulukot tai muu sisältö, joka on alun perin tuotettu analogiseen asiakirjaan niin, että analogisen asiakirjan yksittäisiä sivuja on ymmärrettävyyden vuoksi tarkasteltava rinnakkain eli aukeamittain (ks. kuva 4).

Tietosisällön eheyden varmistamiseksi analogiset asiakirjat **pitää** digitoida kokonaisuudessaan siten, että kuva-ala ylittää sivujen reunat.

²³ Esimerkiksi: "Osa 1/2" ja "Osa 2/2".



KANSALLISARKISTO

Aukeama										
Sivu					Sivu					
<i>Tietosisältö</i>					<i>laajissa vuorossa 1907.</i>					
Perintökirjan		Kilisekunta	Pitäjä	Kylä	Tila	Perintökirjasta				
No.	Astajin				Nimi	Talokas	Masturi	Alk. jälle on lähin ma- lotta.	Koko per- intökirja- litti.	% on pe- rillönnä- tönnä.
1907								1907		
1.	Huhtik. Naakun	Liedon	Ingöiden	2.	1.	13				
2.	Huhtik. Naakun	Liedon	Ingöiden	3.	1.	67				
D					D					

Kuva 3. Tietosisältö esitetään aukeamittain / Tietosisältö ulottuu sidoksessa koko aukeamalle.

Alla olevassa taulukossa (Taulukko 10) esitetään aukeamittain digitoitavien sidosten digitointia koskevat määritykset.



KANSALLISARKISTO

Taulukko 10. Aukeamittain digitoitavien sidosten määrittymiset

Elementti	Vaadittu arvo
Tallekappaleen formaatti	TIFF 6.0 tai JPEG part 1 versio 1.00 tai 1.01 tai 1.02
Tallekappaleen pakkaaminen	TIFF = LZW tai JPEG = JPEG-laatuso 90 %
Resoluutio	300 ppi
Väritila	2 = RGB
Bittisyvyys	8 bittiä kanava, 3 kanavaa (24 bit)
Tallekappaleen ICC-profiili	eciRGB v2, ProPhoto RGB, AdobeRGB (1998), sRGB

3.2.4 Valokuvat

Tässä valokuvalla voidaan tarkoittaa yksittäisiä valokuvia tai useamman valokuvan kokonaisuuksia, esimerkiksi valokuva-albumin sivuja.

Valokuvien sisältämän tietosisällön eheyden varmistamiseksi analogiset asiakirjat **pitää** digitoida kokonaisuudessaan²⁴ siten, että kuva-ala ylittää skannattavan kohteen reunat. Valokuvat **pitää** digitoida värillisiksi kuvatiedostoiksi (RGB, 24 bit).

Jokaisen digitoinnin yhteydessä **pitäisi** kuvata referenssikohde eli mittataulu. Värien ja sävyjen säätö **pitää** tehdä mittataulun avulla.

Alla olevassa taulukossa (Taulukko 11) esitetään valokuvien digitoimista koskevat määrittymiset.

Taulukko 11. Valokuvakokoelmien määrittymiset

Elementti	Vaadittu arvo
Tallekappaleen formaatti	TIFF 6.0
Tallekappaleen pakkaaminen	LZW
Resoluutio	Analogisen ilmentymän mukaan, ks. Taulukko7 Resoluutio- taulukko
Väritila	2 = RGB
Bittisyvyys	RGB = 8 bittiä kanava, 3 kanavaa (24 bit)
Tallekappaleen ICC-profiili	RGB: eciRGB v2, ProPhoto RGB, AdobeRGB (1998)

Kun digitoitava asiakirja on pienikokoinen, **pitää** resoluutiota nostaa alla olevan resoluutiotaulukon (Taulukko 12) mukaisesti. Mikäli analoginen asiakirja on fyysisiltä mitoiltaan pienempi kuin taulukossa 11 esitetään, **pitää** digitoinnissa käytettävää resoluutiota kasvattaa samaa laskentakaavaa soveltaen. Resoluutiota **ei saa** muuttaa kuvatiedostoa käsiteltäessä, jotta analogisen asiakirjan fyysisten mittojen arviointi on digitaalisen tallekappaleen perusteella mahdollista. Tavoitteena on aina ~A3-kokoinen (cm x cm) tallekappale resoluutiolla 300 ppi.

²⁴ Tarkoittaa myös valokuvien säilytysvälineitä ja kehyksiä, joissa on tietoja tai merkintöjä.



KANSALLISARKISTO

Taulukko 12. Resoluutiotaulukko

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

3.2.5 Lasinegatiivit, muut negatiivit ja diat

Lasinegatiivien, muiden negatiivien ja diojen tietosisällön eheyden varmistamiseksi analogiset asiakirjat **pitää** digitoida kokonaisuudessaan siten, että kuva-ala ylittää skannauksen kohteena olevan analogisen negatiivin tai dian reunat. Mikäli negatiivi tai dia on asetettu kehyksiin, rajauksessa **pitää** noudattaa äärimmäistä huolellisuutta, jotta tietosisältöä ei rajaudu pois kuvatiedostosta. Diat ja negatiivit²⁵ **pitää** digitoida värillisiksi kuvatiedostoiksi (RGB, 24 bit).

Lasinegatiivien, muiden negatiivien ja diojen digitointia koskevat määritykset on koottu alla olevaan taulukkoon (Taulukko 13).

Taulukko 13. Lasinegatiivien, muiden negatiivien ja diojen määritykset

Elementti	Vaadittu arvo
Tallekappaleen formaatti	TIFF 6.0
Tallekappaleen pakkaaminen	LZW
Resoluutio	Analogisen ilmentymän mukaan, ks. Taulukko 7 Resoluutiotaulukko
Väritila	2 = RGB
Bittisyvyys	RGB = 8 bittiä kanava, 3 kanavaa (24 bit)
Tallekappaleen ICC-profiili	RGB: eciRGB v2, ProPhoto RGB, AdobeRGB (1998)

Kun digitoitava analoginen asiakirja on pienikokoinen, **pitää** resoluutiota nostaa resoluutiotaulukon (Taulukko 12) mukaisesti. Mikäli analoginen asiakirja on fyysisiltä mitoiltaan pienempi kuin taulukossa 13 esitetään, **pitää** digitoinnissa käytettävää resoluutiota kasvattaa samaa laskentakaavaa soveltaen. Tallekappaleetiedoston resoluutiota **ei saa** muuttaa kuvatiedostoa käsiteltäessä, jotta analogisen asiakirjan fyysisten mittojen arviointi on digitaalisen tallekappaleen perusteella mahdollista.

Tavoitteena on aina ~A3-kokoinen (cm x cm) tallekappale resoluutiolla 300 ppi.

²⁵ Negatiivit digitoidaan suorassa ja oikein päin, ei peilikuvina. Kuvaamalla digitoitaessa negatiivit asetetaan emulsiopuoli alaspäin, kiiltävä puoli kameraa kohden. Lasinegatiivit on suositeltavaa kuvata emulsiopuoli ylöspäin, jolloin ne digitoituvat peilikuvana. Varmista, että digitoitu kuva tulee lopulta tallennettuna oikean lukusuunnan mukaisesti.



KANSALLISARKISTO

3.2.6 Mikrofilmit ja -kortit

Mikrofilmin pienennyskerroin **pitää** tarkistaa filmirullan alusta. Mikäli pienennyskerrointa ei ole ilmoitettu, suurennuskerroin **pitää** asettaa riittävän suureksi, jotta tuotetun digitaalisen asiakirjan tietosisältö on luettavissa. Mikrofilmien kohdalla filmille kuvattu aineisto saadaan suurennuskertoimen avulla toistumaan riittävän luettavana resoluutiolla 300 ppi. Skannauksessa käytettävä resoluutio on suurennuskerroin kertaa (x) haluttu resoluutio (esimerkiksi 12 x 300 = 3600). Mikrofilmi itsessään on varsin jyrkkä tallenneväline, joten digitoinnissa pitää varoa jyrkkyyden korostamista, jotta luvussa 3 esitetyt määritykset toteutuvat.

Alkuperäisen aineiston oikean koon todentamisessa on voitu käyttää myös muita menetelmiä (esimerkiksi mittanauha). Käytettyjä menetelmiä **pitää** tarkastella tapauskohtaisesti. Samalla **pitää** pyrkiä arvioimaan alkuperäisen, mikrofilmille kuvatun analogisen asiakirjan fyysisiä mittoja, vaikka mitään edellä mainituista menetelmistä ei olisi mikrokuvauksen yhteydessä sovellettu.

Alla olevassa taulukossa (Taulukko 14) esitetään mikrofilmien digitointia koskevat määritykset.

Taulukko 14. Mikrofilmien määritykset

Elementti	Vaadittu arvo
Tallekappaleen formaatti	TIFF 6.0 tai JPEG part 1 versio 1.00 tai 1.01 tai 1.02
Tallekappaleen pakkaaminen	TIFF = LZW tai JPEG = JPEG-laatusaso 90 %
Resoluutio	300 ppi
Väritila	Harmaasävy
Bittisyvyys	8 bittiä kanava, 1 kanava (8 bit)
Tallekappaleen ICC-profiili	Gray gamma 2.2

Värilliseen mikrofilmiin **pitää** soveltaa tässä luvussa esitettyjä vaatimuksia, mutta tallekappaleen väritilan **pitää** olla RGB, bittisyvyys pitää olla 24 bit ja tallekappaleen ICC-profiilin **pitää** olla eciRGB v2, ProPhoto RGB, AdobeRGB (1998) tai sRGB (ks. Taulukko 15).

Taulukko 15. Värillisten mikrofilmien määritykset

Elementti	Vaadittu arvo
Tallekappaleen formaatti	TIFF 6.0 tai JPEG part 1 versio 1.00 tai 1.01 tai 1.02
Tallekappaleen pakkaaminen	TIFF = LZW tai JPEG = JPEG-laatusaso 90 %
Resoluutio	300 ppi
Väritila	RGB
Bittisyvyys	24 bit
Tallekappaleen ICC-profiili	eciRGB v2, ProPhoto RGB, AdobeRGB (1998) tai sRGB



KANSALLISARKISTO

4 Audiovisuaalisia aineistoja koskevat vaatimukset

Analogiset audiovisuaaliset aineistot ovat ominaisuuksiltaan sekä laiteriippuvaisia että itsetuhoutuvia. Niiden fyysinen säilyvyys ja käyttöikä on huomattavasti lyhyempi kuin muilla asiakirja-aineistoilla. Laiteriippuvuus tuo myös omat käytön haasteensa. Audiovisuaalisten aineistojen sisällön säilyttämisen näkökulmasta digitointi on keskeistä.

Audiovisuaaliset aineistot jaetaan tässä ohjeessa kahteen ryhmään, ääninauhoihin ja kuvanauhoihin. Aineistojen digitointiprosessissa tuotetaan tallekappaletiedosto, joka on aineistotyyppin mukaan joko ääni- tai videotiedosto.

Audiovisuaalisten aineistojen digitoinnin tavoitteena on taata aineistojen pysyvä säilyminen digitaalisessa muodossa. Digitoitujen aineistojen **pitää** sisältää mahdollisimman tarkoin sama sisällöllinen kuva- ja ääni-informaatio kuin alkuperäisissä ääninauhoissa tai kuvanauhoissa.

4.1 Ääninauhat

Ääninauhoihin lasketaan muun muassa:

- äänikelanauha (avokelanauha, kelanauha)
- C-kasetti
- minikasetti

Avokela- ja C-kasettinauhureiden äänipäiden atsimuuttikulma²⁶ **pitää** säätää nauhakohtaisesti mitaustekniikkaa apuna käyttäen. Nauhat **pitää** digitoida toistamalla jokainen monoraita tai stereoraitapari kokonaisuudessaan alusta loppuun alkuperäisellä nopeudella.

Analoginen äänisignaali muunnetaan digitaaliseksi erillisellä AD-muuntimella. Digitointiparametreina käytetään 24 bitin resoluutiota ja 48 000 hertsin näytteenottotaajuutta, mutta 96 000 hertsin näytteenottotaajuutta voidaan käyttää tapauskohtaisesti. Nauhoissa oleva ääni **pitää** tallentaa mahdollisimman muuttamattomana, joten ääntä **ei saa** parannella digitoinnin yhteydessä. Normalisointia tai kompressoiteja **ei saa** tehdä.

Digitoitujen tiedostojen muotona **pitää** käyttää häviötöntä LPCM-koodattua WAV-muotoa.

Äänenvoimakkuuden taso **pitää** säätää digitoitaessa tallennekohtaisesti niin, että huippuarvo asetuu 0- ja 6 dB:n välille. Tästä vaatimuksesta **saa** yksittäistapauksissa poiketa, jos sen noudattaminen johtaa lopputuloksen heikkenemiseen. Äänen tasoa **ei saa** säätää ohjelmallisesti digitoinnin jälkeen.

²⁶ Atsimuuttikulmalla tarkoitetaan äänipään kulmaa suhteessa ääninauhaan. Jos kulma ei ole kohtisuora, on eri kanavien välillä taajuudesta riippuva vaihe-ero.



KANSALLISARKISTO

Jokainen nauhakokonaisuus **pitää** digitoida kokonaisuudessaan pääsääntöisesti samoilla asetuksilla. Nauhakokonaisuudella tarkoitetaan tässä yhteydessä nauhan yhtä puolta kokonaisuudessaan.

Taulukko 16. Ääninauhojen määrittelyt

Elementti	Vaadittu arvo
Tallekappaleen formaatti	LPCM-koodattu WAV
Tallekappaleen pakkaaminen	Pakkaamaton
Näytteenottotaajuus	48 kHz / 96 kHz
Bittisyvyys	24 bit

4.1.1 Äänitiedostojen vaaditut tekniset metatiedot

Äänitiedostojen pakollisten teknisten metatietojen metatietoskeema on AudioMD²⁷.

Taulukko 17. Äänitiedostojen teknisen metatiedon skeema

Termi	Metatieto	Tuntematon arvo sallittu
Äänikoodauksen tyyppi	audio_data_encoding	Kyllä
Kesto	duration	Kyllä
Datavirran nopeus	data_rate	Kyllä
Datavirran tila	data_rate_mode	Ei
Bittisyvyys	bits_per_sample	Kyllä
Ohjelmisto (koodekki)	codec_creator_app	Kyllä
Versio (koodekki)	codec_creator_app_version	Kyllä
Nimi (koodekki)	codec_name	Kyllä
Pakkauksen laatu (koodekki)	codec_quality	Ei
Näytteenottotaajuus	sampling_frequency	Kyllä
Kanavien määrä	num_channels	Kyllä

4.2 Kuvanauhat

Kuvanauhoihin lasketaan muun muassa:

- VHS-nauhat
- Betacam-nauhat
- DVcam nauhat
- miniDV-nauhat.

²⁷ The Library of Congress » Standards » AudioMD/VideoMD. Kongressin kirjaston verkkosivu <http://www.loc.gov/standards/amdvmd/> Viitattu 19.9.2024.



KANSALLISARKISTO

Kuvanauhoista digitoitujen videotiedostojen **pitää** sisältää mahdollisimman tarkoin sama sisällöllinen kuva- ja ääni-informaatio kuin alkuperäisissä nauhoissa. Kuvanauhat **pitää** digitoida laadukkailla studiokäyttöön tarkoitetuilla laitteistoilla. Kuvanauhojen määrytykset on esitetty alla olevassa taulukossa (Taulukko 18).

Taulukko 18. Kuvanauhojen määrytykset

Elementti	Vaadittu arvo
Tallekappaleen formaatti	FFV1 versio 3 [FFV1] fmt/569
Tallekappaleen pakkaaminen	Pakkaamaton

4.2.1 Kuvanauhoista digitoitujen videotiedostojen tekniset metatiedot

Videotiedostojen pakollisten teknisten metatietojen metatietoskeema on VideoMD²⁸. Pakolliset tekniset metatiedot on esitetty alla olevassa taulukossa (Taulukko 19).

Taulukko 19. Kuvanauhatiedostojen teknisen metatiedon skeema

Termi	Metatieto	Tuntematon arvo sallittu
Ääni	sound	Ei
Kesto	duration	Kyllä
Datavirran nopeus	data_rate	Kyllä
Datavirran tila	data_rate_mode	Ei
Bittisyvyys	bits_per_sample	Kyllä
Väritila	color	Ei
Ohjelmisto (koodekki)	codec_creator_app	Kyllä
Versio (koodekki)	codec_creator_app_version	Kyllä
Nimi (koodekki)	codec_name	Kyllä
Pakkauksen laatu (koodekki)	codec_quality	Ei
Kuvataajuus	frame_rate	Kyllä
Kuvan korkeus	height	Kyllä
Kuvan leveys	width	Kyllä
Pikselin kuvasuhde (PAR)	par	Kyllä
Näytön kuvasuhde (DAR)	dar	Kyllä
Värien alinäytteistys	sampling	Kyllä
Signaalin muoto	signal_format	Kyllä

²⁸ The Library of Congress » Standards » AudioMD/VideoMD. Kongressin kirjaston verkkosivu <http://www.loc.gov/standards/amdvmd/>. Viitattu 19.9.2024.



KANSALLISARKISTO

5 Digitointiprosessissa muodostettava hakemistorakenne

Kansallisarkisto ottaa vastaan vain ennalta määriteltyihin hakemistorakenteisiin tallennettua digitoitua tietoaineistoa. Mikäli tässä asiakirjassa esitettyjen vaatimusten mukaan digitoitu tietoaineisto siirretään Kansallisarkistoon, se **pitää** tallentaa Kansallisarkiston siirto-ohjeiden mukaiseen hakemistorakenteeseen.

Siirto-ohjeissa määritellyt ja tässä luvussa tiivistetyt kuvatut hakemistorakenteet koskevat Kansallisarkistoon siirrettäviä tietoaineistoja eivätkä sulje pois muiden tallennusrakenteiden käyttöä. Organisaatio voi tallentaa digitoidusta kuva-aineistosta aina esimerkiksi myös PDF-muotoiset käyttökapaleet omiin järjestelmiinsä haluamassaan tietorakenteessa.

Kuva-aineistolle ja AV-aineistolle on omat hakemistorakenteensa. Kuhunkin juurihakemistoon tallennetaan pääsääntöisesti yhden arkistoyksikön sisältö. AV-aineistoa sisältävä arkistoyksikkö **saa** sisältää useiden nauhojen tiedostot.

5.1 Kuva-aineiston hakemistorakenne

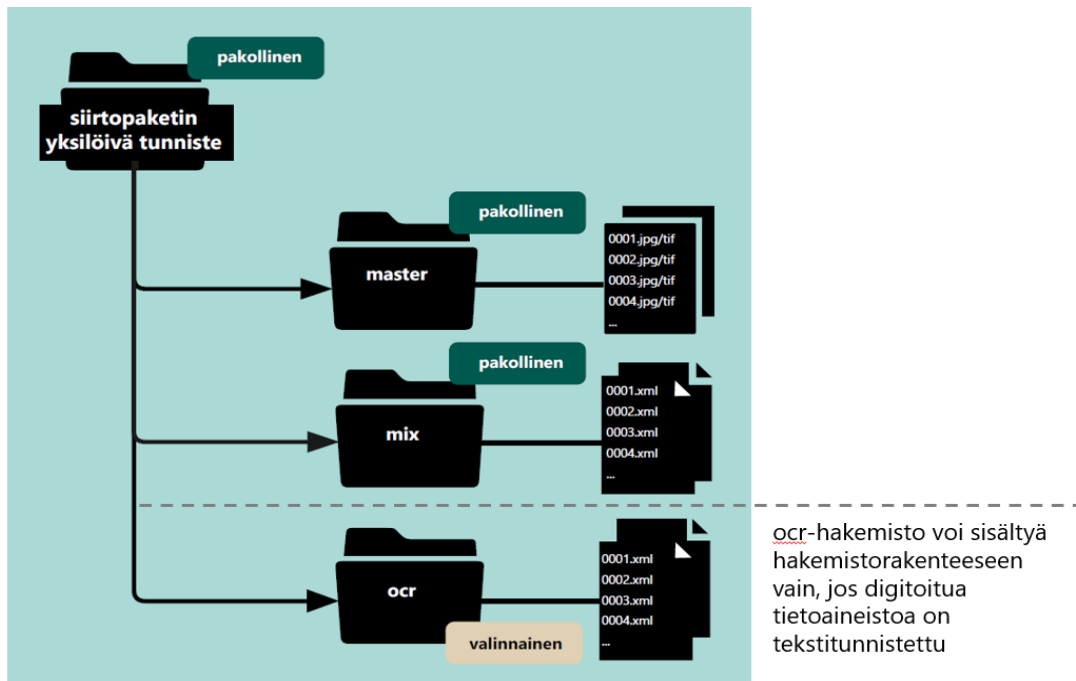
Digitointiprosessissa tuotetut kuvatiedostot, niiden MIX-metatietotiedostot ja tekstitunnistuksessa mahdollisesti tuotetut tiedostot **pitää** tallentaa kuvan 5 mukaiseen hakemistorakenteeseen riippumatta siitä, milloin aineisto siirretään Kansallisarkistoon. Hakemistorakenteen luominen jälkikäteen saattaa olla työlästä.

- Juurihakemiston nimen **pitää** olla yksilöivä, ja se saa sisältää ainoastaan seuraavia merkkejä: a-z, A-Z ja 0–9.
- Kuvatiedostot, niiden tekniset metatiedot ja mahdolliset tekstitunnistuksessa luodut tiedostot tallennetaan omiin alihakemistoihinsa.
- Alihakemistot **pitää** nimetä seuraavasti:
 - kuvatiedostojen hakemisto: **master**
 - MIX-metatietotiedostojen hakemisto: **mix**
 - tekstintunnistuksessa luotujen tiedostojen hakemisto: **ocr**

Kaikkien alihakemistojen tiedostot **pitää** nimetä lähtökohtaisesti nelinumeroisella lukujonolla. Lukujono on juokseva niin, että ensimmäinen tiedosto saa nimen 0001, toinen 0002 jne. Mix- ja ocr-hakemistojen sisältämien tiedostojen tiedostonimien **pitää** vastata master-hakemiston kuvatiedostojen nimiä. Kuvatiedoston nimeä vastaava lukujono siis linkittää digitoinnissa luodun kuvatiedoston ja siihen liittyvät mix-metatiedot toisiinsa.



KANSALLISARKISTO



Kuva 4. Tuotettava hakemistorakenne

Juurihakemisto **ei saa** sisältää muita kuin kuvan 5 mukaisia alihakemistoja. Kun aineisto siirretään Kansallisarkistoon, jokainen juurihakemisto **pitää** paketoita TAR-tiedostoksi. Tätä TAR-tiedostoa kutsutaan siirtopaketiksi. TAR-tiedoston **saa** lisäksi pakata häviöttömään GZIP- (.tar.gz) tai BZIP2- (.tar.bz2) muotoon.

Hakemistorakenteen ja siirtopaketin tarkempi kuvaus on Kansallisarkiston ajantasaisissa ohjeissa.

5.2 Audiovisuaalisen aineiston hakemistorakenne

AV-aineiston digitointiprosessissa tuotetut tiedostot **pitää** tallentaa Sähköisen arkistoinnin palvelun Muut digitaaliset tietoaineistot – siirtopaketin muodostaminen -ohjeessa kuvattuun hakemistorakenteeseen ennen kuin ne siirretään Kansallisarkistoon.

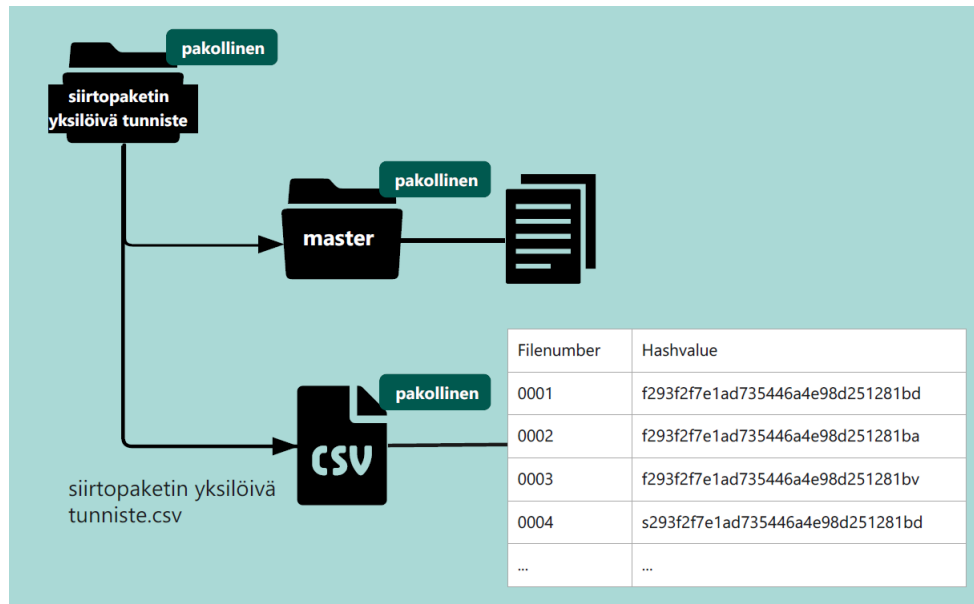
- Juurihakemiston nimen **pitää** olla yksilöivä ja se saa sisältää ainoastaan seuraavia merkkejä: a-z, A-Z ja 0–9.
- Digitoinnissa tuotetut tiedostot **pitää** tallentaa master-nimiseen alihakemistoon.
 - Tiedostot **pitää** nimetä nelinumeroisella lukujonolla. Lukujono on juokseva niin, että ensimmäinen tiedosto saa nimen 0001, toinen 0002 jne.
- Juurihakemistoon tallennettavan csv-tiedoston **pitää** sisältää master-hakemiston sisältämien tiedostojen tiedostonimet (sarake Filenumber) ja tiedostoille lasketut tarkistussummat



KANSALLISARKISTO

(sarake Hashvalue). Filenumber-sarakkeeseen **ei saa** sisällyttää tiedostopäätettä. Tarkistus-summa **pitää** antaa muodossa MD5.

- Tiedosto **pitää** nimetä siirtopaketin yksilöivällä tunnisteella, eli täsmälleen samalla merkkijonolla kuin juurihakemisto.
- Tiedoston merkistön **pitää** olla UTF-8.
- Kenttien ympärillä **ei saa** olla lainausmerkkejä.



Kuva 5. AV-aineistoista tuotettava hakemistorakenne

Juurihakemiston **pitää** sisältää vain kuvan 6 mukainen hakemisto tiedostoineen. Kun aineisto siirretään Kansallisarkistoon, jokainen juurihakemisto **pitää** paketoita TAR-tiedostoksi. Tätä TAR-tiedostoa kutsutaan siirtopaketiksi. TAR-tiedoston **saa** lisäksi pakata häviöttömään GZIP- (.tar.gz) tai BZIP2- (.tar.bz2) muotoon.

5.3 Esimerkkipaketit

Kuvatiedostojen siirrossa käytettävistä siirtopaketeista on muodostettu ohjeen liitteeksi kaksi esimerkkipakettia:

1. EsimerkkipakettiJPEG2020.tar
2. EsimerkkipakettiTIF2020.tar

Master-hakemistojen sisältämät kuvatiedostot eivät esimerkkipaketeissa ole kuvanlaadullisia referenssejä. Kuvatiedostot sisältävät tässä asiakirjassa pakollisiksi määritellyt metatiedot. Mix-hakemistojen tiedostot ovat esimerkkejä liitteen paketissa olevista master-tiedostoista, ellei elementtien kohdalla todeta toisin.



KANSALLISARKISTO

5.4 Digitoidun aineiston siirtäminen Kansallisarkistoon

Kun digitoituja aineistoja siirretään Kansallisarkistoon, noudatetaan kulloinkin voimassa olevia siirto-ohjeita. Ajantasaiset ohjeet ovat aina Kansallisarkiston [ohjepankissa](#).

Kansallisarkistoon siirrettävien kuvien enimmäiskoko pikseleissä on 65535x65535.

6 Muutokset

Taulukko 20. Tehdyt muutokset.

Koskee ohjetta: Kansallisarkiston vaatimukset hävittämiseen tähtäävään digitointiin (KA/3357/07.01.02.04.02/2019)	Koskee ohjetta: Kansallisarkiston määritykset arkistoitavien erikoisaineistojen digitointiin (KA/15780/07.01.02.04.02/2020).	Muutos selitettynä
Kyllä	Kyllä	Seulontapäätösmentettelyyn liittyvä ohjeistus on päivitetty.
Kyllä	Kyllä	Termeistä ja käsitteitä koskevat taulukot on siirretty samaan lukuun. Käsitteitä on lisätty ja joitain käsitteitä on tarkennettu.
Kyllä	Kyllä	Yleiset luvut (esimerkiksi johdanto) on yhdistetty.
Kyllä	Kyllä	Lisätty aiemmista ohjeista puuttuneet AV-aineistojen digitointia koskevat ohjeet ja vaatimukset.
Kyllä	Kyllä	Kuvatiedostoa koskevat pakolliset metatiedot on jaoteltu kahteen taulukkoon (taulukko 5, JPEG ja taulukko 6, TIFF). Lisäksi taulukoissa esitettyjä vaatimuksia on pyritty avaamaan aiempaa selvemmin.
Kyllä	Kyllä	Luku 3.1, taulukot 5 ja 6. "Käsittelyohjelma"-metatietoon on lisätty tarkenne, että siihen tallennetaan kaikki käytetyt käsittelyohjelmat.
Ei	Kyllä	Muutettu taulukosta 5 "Kuvatiedoston luontipäivämäärä ja aika" -metatiedon



KANSALLISARKISTO

		vaatimus DateTimeOriginal muotoon DateTime (306).
Kyllä	Kyllä	Aineistotyyppikohtaiset vaatimukset on koottu lukuun 3.2. (sekä A4- ja foliokokoiset, sidokset ja erikoisaineistot). Taulukoissa 5 ja 6 aineistotyyppi-riippuvaiset "Vaadittu arvo" -sarakkeen arvo "Aineistotyyppi-riippuvainen" tarkoittaa sitä, että vaadittu metatietoelementti riippuu aineistotyyppistä ja se on tarkistettava luvusta 3.2.
Kyllä	Kyllä	Taulukot kuvatiedostoa ja kuvatiedoston prosessointia kuvaavaista metatiedoista ja rakenteesta on yhdistetty samaan taulukkoon (taulukko 7).
Kyllä	Kyllä	Taulukkoon 7: MIX-metatietokentän nimi iccProfile URL on korjattu muotoon iccProfileURI. MIX-metatietojen kansainvälisten määritysten mukaisessa skeemassa (versio 2) on iccProfileURI-niminen elementti. Skeeman kuvauksessa ja Kansallisarkiston digitointimäärittelyissä nimi on ollut aiemmin version 1 mukaisessa muodossa iccProfileURL. Kansallisarkistoon siirrettävissä digitoiduissa aineistossa tulee käyttää skeeman mukaista muotoa iccProfileURI.
Ei	Kyllä	Värillisten mikrofilmien vaatimukset on taulukoitu omaan taulukkoonsa (taulukko 15).
Kyllä	Kyllä	Lisätty poikkeus (luku 3.1.), siitä, milloin pakattu kuvatiedosto voidaan tallentaa toisen kerran. Lisäksi on määritelty, miten ProcessingSoftware -kenttään tulee kirjata käsittelyohjelmat. Määritelty myös MIX tiedoston ChangeHistory -elementin päivitys, kun kuvatiedostoa muokataan.



KANSALLISARKISTO

Kyllä	Kyllä	Lisätty tarkennuksia digitoidun kuva-aineiston siirtopaketin muodostamisen ohjeeseen. Lisätty kuvaus siirtopakettirakenteesta, jossa digitoitu AV-aineisto siirretään Kansallisarkistoon.
Kyllä	Kyllä	Tarkennettu ohjetta suurista kuvatiestoista, sekä lisätty tarkat maksimit pikseleissä kuville. Luku 3.1
Kyllä		Poistettu kappale 5.2 Tunnistetun tekstin tallennusformaatti.