

KANSALLISARKISTON MÄÄRITYKSET ARKISTOITAVIEN ERIKOISAINESTOJEN DIGITOINTIIN

- Sisältö** Kansallisarkiston määritykset arkistoitavien erikoisaineistojen digitointiin.
- Tarkoitus** Tämä ohje määrittelee digitointiprosessin ja lopputuloksen, jotka varmistavat arkistoitavien erikoisaineistojen tietosisältöjen säilymisen ja käytettävyyden digitaalisessa muodossa.
- Kohderyhmä** Ohje on tarkoitettu Kansallisarkistolle sekä niille toimijoille, jotka digitoivat arkistolain (831/1994) nojalla arkistoitavaksi määrättyjä analogisia erikoisaineistoja.
- Rajaukset** Ohje ei koske skannausta, joka tehdään asiakirjan laatimisen tai asiankäsittelyn yhteydessä tai silloin kun organisaatiolle saapuvia analogisia asiakirjoja muutetaan digitaaliseen muotoon. Erikoisaineistot koostuvat karttakokoelmista, piirustuskokoelmista, valokuvakokoelmista, negatiiveista, dioista, mikrofilmeistä sekä sidoksista, joissa sidosten tietosisältö esitetään aukeamittain (Taulukko 1). Ohje ei koske erikoisaineistoiksi luokiteltavia yksittäisiä asiakirjoja, jotka on arkistoitu A4- ja foliokokoisten asiakirjojen joukkoon.

Taulukko 1, Ohjeiden soveltaminen eri asiakirjatyypin yhteydessä.

Aineisto	Kansallisarkiston määritykset arkistoitavien erikoisaineistojen digitointiin (2021)	Kansallisarkiston vaatimukset hävittämiseen tähtäävään digitointiin 2019
Karttakokoelma	x	
Piirustuskokoelma	x	
Valokuvakokoelmat	x	
Lasinegatiivit, muut negatiivit ja diat	x	
Mikrofilmit	x	
Sidokset, jossa sidosten tietosisältö esitetään aukeamittain	x	
Sidokset		x
A4- ja foliokokoiset asiakirjat		x
A4- ja foliokokoisten asiakirjojen joukossa olevat yksittäiset valokuvat, piirustukset ja kartat		x



Tämän ohjeen mukaisen digitointiprosessin lopputulos ei ole pitkäaikaissäilytyspaketti, joka tallennetaan pitkäaikaissäilytysjärjestelmään. Tallennettava pitkäaikaissäilytyspaketti on kuitenkin mahdollista muodostaa ohjeen mukaisesti digitoidusta aineistosta. Ohje ei ota kantaa aineiston analogisen ilmentymän hävittämisprosessiin, vaan siinä kuvataan ne kriteerit, joiden täytyttyä analogisen asiakirjan hävittäminen on mahdollista.

Säädökset, johon ohjeen antaminen perustuu

Laki Kansallisarkistosta 1145/2016 § 2.

Voimassaoloaika

Toistaiseksi, ohjeen päiväyksestä eteenpäin.

Kumoaa ohjeen ja suosituksen

Arkistolaitoksen tekniset vaatimukset digitoidulle aineistolle 24.8.2016 ja 15.8.2008, AL/11130/07.01.02.04.02/2008.

Sisällysluettelo

1	Termit ja käsitteet	1
2	Johdanto.....	3
3	Yleiset digitointiprosessin suositukset ja hyvät käytänteet.....	4
4	Yleiset digitointiprosessin määritykset	5
5	Aineistotyyppikohtaiset määritykset	5
5.1	Kartta- ja piirustuskokoelmat	5
5.2	Sidokset, joissa tietosisältö esitetään aukeamittain	6
5.3	Valokuvakokoelmat	8
5.4	Lasinegatiivit, muut negatiivit ja diat	8
5.5	Mikrofilmit.....	9
6	Hyväksyttävät formaatit.....	10
6.1	Kuvatiedosto.....	10
6.2	Kuvatiedostoa ja kuvatiedoston prosessointia kuvaavat metatiedot ja rakenne	14
7	Digitointiprosessissa muodostettava siirtopaketti	15
8	Allekirjoitukset	17

1 Termit ja käsitteet

Ohjeessa käytetty termistö perustuu Internet Engineering Task Forcen tekemään määrittelyyn [RFC 2119].¹ Alla olevassa taulukossa (Taulukko 2) kuvataan tässä ohjeessa käytetyt termistön suomenkieliset käännökset.

Taulukko 2, Tässä asiakirjassa käytetyt käännökset

ENGLANTI	SUOMI
must	pitää
must not	ei saa
required	pakollinen
should	pitäisi
should not	ei pitäisi
may	saa
optional	vapaaehtoinen

Alla olevassa taulukossa (Taulukko 3) on esitetty asiakirjassa käytetyt käsitteet selitteineen:

Taulukko 3, Käsitteistö

KÄSITE	SELITE
Analoginen asiakirja	Paperisessa tai muussa käsin kosketeltavassa muodossa laadittu, säilytetty ja käytettävä asiakirja.
Analoginen ilmentymä	Digitoitavaksi päätetyn analogisen asiakirjakokonaisuuden analoginen olomuoto.
Analogisen säilytysmuodon kulttuurihistoriallinen arvo	Arkistoitavan tai arkistoidun analogisen asiakirjan (asiakirjatiedon) ominaisuus, kulttuurihistoriallinen arvo, jonka vuoksi analoginen asiakirja säilytetään digitoinnista huolimatta myös analogisessa muodossa.
Arkistointi	Arkistointi tapahtuu säilytysajan päättymisen jälkeen (Tiedonhallintalaki 906/2019 21 §). Arkistointi tarkoittaa myös arkistolain mukaista pysyvää säilyttämistä (8 § 3 momentti).
Arkistointiperuste	Arkistointiperuste määrää analogisten asiakirjojen keskinäisen järjestyksen yksittäisen sarjan sisällä.
Asiakirja	Asiakirja on todistusvoimaista tietoa. Toimija on tuottanut tai vastaanottanut tiedon osana tehtäviään ja säilyttää sitä tietovarantona sekä todisteena. Asiakirja on sisällön, rakenteen/ilmiasun 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 tarkoi-

¹ <https://www.ietf.org/rfc/rfc2119.txt> Viitattu 11.1.2021.

KÄSITE	SELITE
	tetuista 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 §).
Digitaalinen ilmentymä	Digitoitavaksi päätetyn analogisen asiakirjakokonaisuuden digitaalinen olomuoto.
Digitointi	Analogisessa muodossa säilytettävien tai arkistoitujen asiakirjojen digitaaliseen muotoon muuttaminen. Laajempi joukko toimintoja, joissa skannaaminen on yksi vaihe.
Digitointilaitte	Laite, jolla analoginen asiakirja muunnetaan digitaaliseen muotoon.
Digitointiprosessi	Joukko toimintoja, joiden avulla analoginen ilmentymä muunnetaan digitaalseksi ilmentymäksi.
Digitoitu asiakirja	Analogisesta asiakirjasta digitointiprosessilla tuotettu digitaalinen versio, joka voi koostua n-määrästä tiedostoja.
Erikoisaineisto	Koostuu omina kokonaisuuksina/kokoelmina arkistoiduista kartoista, piirustuksista, valokuvista, negatiiveista, mikrofilmeistä ja mikrokorteista sekä sidoksista, joissa sidosten tietosisältö esitetään aukeamittain.
Hävittäminen	Määräajan säilytettävän asiakirjan fyysinen (ml. tekninen) hävittäminen (tuhoaminen) sekä arkistoitavan asiakirjan analogisen muodon, ilmentymän, hävittäminen.
Kuva-ala	Kuvatiedostosta visuaalisella tarkastelulla näkyvä informaatio.
Kuvatiedosto	Digitointiprosessissa tuotettu bittikarttakuva. Tarkentaa digitaalisen tiedoston muodon bittikarttakuvaksi.
Käyttökappale²	Tiedosto, joka tarjotaan käyttöön erilaisiin käyttöliittymiin. Vertaa tallekappale.
Pysyvä säilyttäminen	Pysyvä säilyttäminen tarkoittaa henkilötietojen käsittelyssä sitä, että tietoja säilytetään alkuperäiseen käyttötarkoitukseensa siten, ettei niitä tuhota tai siirretä varsinaisesti arkistoon lainkaan, vaan niiden säilyttämiselle on olemassa pysyvä peruste etujen, oikeuksien tai velvollisuuksien määrittäjänä. (Tiedonhallintalaki, HE 284/2018 vp, s. 98). Arkistolaisissa käytetty termi pysyvä säilyttäminen tarkoittaa EU:n tietosuoja-asetuksen (EU) 2016/679 ja tiedonhallintalain mukaista arkistointia.
Päälukusuunta	Mahdollistaa asiakirjan tietosisällön tulkitsemisen kuvatiedostoa kääntämättä. Mikäli asiakirjassa esiintyy tietosisältöä useampaan lukusuuntaan, tarkoittaa päälukusuunta sitä suuntaa, jossa suurin osa asiakirjan tietosisällöstä on luettavissa.
Seulontapäätös	Kansallisarkiston päätös arkistolain 8, 11 ja 14 a §:ien nojalla pysyvästi säilytettävistä (arkistoitavista) asiakirjoista ja niiden säilytysmuodosta.
Siirtopaketti	Digitointiprosessissa muodostettu kokonaisuus, joka on siirrettävissä Kansallisarkiston tietojärjestelmiin.

² FADGI, käyttökappale: <http://www.digitizationguidelines.gov/term.php?term=derivativefile> Viitattu 11.1.2021.

KÄSITE	SELITE
Skannaus	Tässä ohjeessa skannaus käsitetään osana digitointiprosessia. Skannausta voidaan suorittaa joko kameralla kuvaamalla tai skannerilla skannaamalla. Luettavuuden takia käytetään tässä ohjeessa termiä skannaus.
Tallekappale³	Digitointiprosessissa tuotettu tiedosto (bittikartakuva), joka on teknisiltä ominaispiirteiltään laadukkain digitointiprosessissa tuotettu kuvatiedosto. Tallekappale yleisesti viedään pitkäaikaissäilytysjärjestelmiin.
Tiedosto	Joko yksin tai muiden tiedostojen kanssa muodostaa digitoidun asiakirjan. Tässä ohjeessa tiedosto on joko bittikartakuva tai XML-tiedosto.
Tuotantovuorokausi	Vuorokausi, jonka aikana laitteella tuotetaan digitaalisia tiedostoja.

2 Johdanto

Digitointi edistää osaltaan digitalisaatiota ja kulttuuriperintöaineistojen säilyvyyttä ja käytettävyyttä. Tässä ohjeessa kuvatut määritykset koskevat erikoisaineistojen digitointiprosessia. Näitä määrityksiä **pitää** noudattaa, kun julkishallinnon toimija digitoi arkistoitavaksi määrättyjä analogisia erikoisaineistoja. Ohjetta **pitää** noudattaa myös silloin, kun digitoidaan erikoisaineistoja, joiden analogiset säilytysmuodot ovat kulttuurihistoriallisesti arvokkaita.

Digitoinnin kohteena olevilla asiakirjoilla **pitää** olla Kansallisarkiston seulontapäätös jo ennen digitoinnin aloittamista. Seulontapäätöksessä määrätään analogisten asiakirjojen kulttuurihistoriallisesta arvosta ja digitoinnin jälkeisestä säilytysmuodosta. Se määrittelee, onko analogisella säilytysmuodolla kulttuurihistoriallista arvoa, jonka vuoksi analogista muotoa ei saa hävittää digitoinnin jälkeen. Ellei tällaista kulttuurihistoriallista arvoa ole ja jos digitaalisten ilmentymien laatu, säilyvyys ja käytettävyys on varmistettu, alkuperäiset analogiset asiakirjat voidaan hävittää. Analogisia asiakirjoja **ei saa** hävittää ennen kuin seulontapäätös on tehty.

Analogisten asiakirjojen tietosisällön arkistointi yksinomaan digitaalisessa muodossa edellyttää, että digitaaliseen muotoon muuttaminen on toteutettu menetelmillä, jotka eivät heikennä asiakirjan todistusvoimaisuutta, eheyttä ja autenttisuutta. Tämän ohjeen noudattaminen turvaa osaltaan digitoitavien analogisten asiakirjojen tietosisällön säilyvyyden yksinomaan digitaalisena.

Digitaalisten ilmentymien vastaanottaminen Kansallisarkiston tietojärjestelmiin edellyttää, että ne täyttävät tässä ohjeessa esitetyt määritykset. Tämän ohjeen laadinnassa on huomioitu arkistosektorilla yleisesti käytössä olevat standardit sekä muiden kansallisarkistojen laatuvaatimukset digitoinnille. Lisäksi luvuissa 6 Hyväksyttävät formaatit ja 7 Digitointiprosessissa muodostettava siirtopaketti on huomioitu kansalliset pitkäaikaissäilytyspalveluiden (PAS-palvelut) asettamat määritykset säilytettävälle aineistoille.⁴

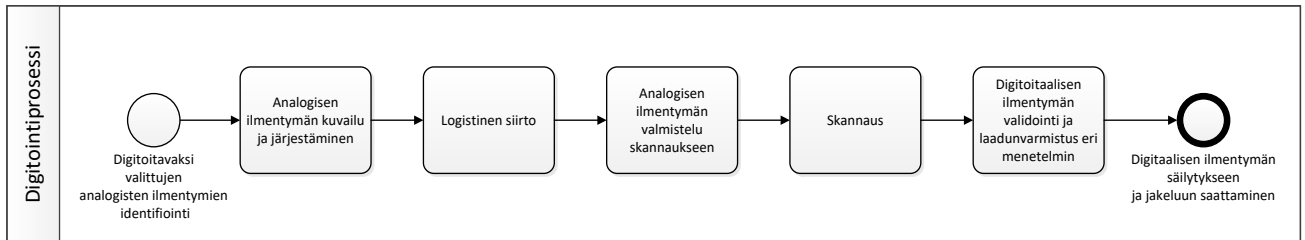
³ FADGI, tallekappale: <http://www.digitizationguidelines.gov/term.php?term=productionmasterfile> Viitattu 11.1.2021.

⁴ <http://digitalpreservation.fi/files/PAS-tiedostomuodot-1.8.0.pdf> Viitattu 11.1.2021.

3 Yleiset digitointiprosessin suositukset ja hyvät käytänteet

Tässä luvussa kuvataan yleisiä hyviä käytänteitä. Luvussa ei käytetä taulukossa 2 esitettyjä termejä.

Digitointi ymmärretään yleisesti prosessina, joka sisältää alla olevassa kuvassa (Kuva 1) esitetyt vaiheet.



Kuva 1, Digitaaliseen muotoon muuntamisen prosessi - yleinen.

Digitointiprosessissa laadunvarmistusta ei voi kohdentaa vain tiettyyn prosessin vaiheeseen, vaan laadunvarmistusta on tehtävä prosessin eri vaiheissa. Tässä ohjeessa keskitytään erityisesti skannauksen laadunvarmistukseen.

Digitoitavat asiakirjat on hyvä kuvailla metatietojen hallintaan tarkoitettuun järjestelmään ennen niiden digitaaliseen muotoon muuntamisen aloittamista. Tällöin voidaan dokumentoida kattavammin analogisen ilmentymän käsittelyketjua sekä esimerkiksi tietoja aineiston valmistelusta. Skannauksen jälkeen aineiston metatietoja voidaan rikastaa joko manuaalisesti tai automaattisin menetelmin.

Skannauksen laadunvarmistus voidaan karkeasti jakaa ennen skannausta tapahtuvaan toimintaan ja sen jälkeiseen laadunvarmistamiseen eli validointiin, jossa visuaalisella tarkastelulla pyritään havaitsemaan laatu-epäkohdat.

Skannausinfrastruktuurin suorituskyky optimoidaan ennen skannausta siten, että sen tuottama digitaalinen ilmentymä on laadultaan paras mahdollinen, joka kyseisellä teknisellä kokoonpanolla voidaan tuottaa. Optimoinnin jälkeen infrastruktuurin suorituskykyä tulisi tarkkailla suunnitellusti, jotta prosessissa tuotettavien digitaalisten ilmentymien laatu säilyy tasaisena. Tarkkailua varten tarvitaan yleisesti mittataulu, mittataulun referenssiarvot ja analysointiohjelmisto. Kuvan laadun lisäksi laiteinfrastruktuurissa pitäisi varmistaa se, että analogiset asiakirjat muunnetaan digitaaliseen muotoon tietosisällöltään täydellisinä. Skannausprosesseissa pitää välttää esimerkiksi kahden päällekkäisen analogisen asiakirjan menemisestä laitteen läpi (läpisyöttöskannerit, avoratasannerit ja muut skannausratkaisut, joissa asiakirjoja skannataan ”massana”). Infrastruktuurin tuottama kuvatiedosto ei myöskään saa olla vääristynyt tai muutoin mittasuhteiltaan muuttunut, jolloin tietosisällön tulkinta vaarantuu.

Skannauksen jälkeinen validointi voidaan toteuttaa otannoilla. Otannan määrä on riippuvainen skannausprosessin luotettavuudesta. Validoinnin tavoitteena on varmistua siitä, että tässä ohjeessa esitetyt määritykset täyttyvät.

Jos kuvatiedostoja käsitellään skannauksen jälkeen, pitäisi yksityiskohtainen kuvankäsittelyhistoria tallentaa ainakin kuvatiedostojen metatietoihin, mahdollisuuksien mukaan myös kuvatiedoston syntyä kuvaileviin XML-tietoihin (MIX).

Huomioitavaa on myös se, että jokainen migraatio ja tiedostokonversio (tiedoston tallennusformaatin muuttaminen) sisältää riskejä liittyen tiedon säilymiseen eheänä. Tämän takia on tärkeää digitoida analoginen aineisto tässä ohjeessa esitettyyn tallennusformaattiin. Heikompileaattisesta tallennusformaattista ei voi tuottaa parempileaattista versiota myöhemmässä käsittelyssä. Se on myös tämän ohjeen määrityksissä kielletty.

4 Yleiset digitointiprosessin määritykset

Analogisten ilmentymien digitaaliseen muotoon muuttamisen vaiheet (digitointiprosessi) **pitää** dokumentoida tässä ohjeessa kuvatuilla tavoilla ja tarkkuudella. Prosessin dokumentoinnilla tarkoitetaan, että skannauksesta ja kuvien mahdollisesta käsittelystä tallennetaan näitä toimenpiteitä dokumentoivat metatiedot. Prosessista **saa** tallentaa metatiedoiksi myös muita toimenpiteitä.

Digitointiprosessissa **pitää** varmistua siitä, että digitoitavaksi tarkoitettu kokonaisuus on digitoitu kokonaisuutena, arkistointiperusteisesti ja sisällöllisesti täydellisenä. Tämä tarkoittaa sitä, että kaikki digitoitavaan kokonaisuuteen kuuluvat analogiset asiakirjat **pitää** digitoida siten, ettei mitään informaatiota jää teknisen tai toiminnallisen virheen takia muuntamatta digitaaliseen muotoon.

Jokaisesta digitoitavaan kokonaisuuteen liittyvästä yksittäisestä kuvatiedostosta **pitää** olla visuaalisella tarkastelulla saatavissa sama informaatio kuin sen analogisesta vastineesta. Kuvatiedosto **ei saa** sisältää mitään elementtejä, joita ei ole analogisessa vastineessa. Tästä poikkeuksen muodostavat mahdolliset samaan kuvatiedostoon skannattavat tai kuvattavat kuvatiedoston värejä, harmaasävyjä, mittasuhteita ja resoluutiota todentavat skannaustekniset mittataulut. Kyseiset mittataulut **pitää** asetella siten, että ne eivät peitä digitoitavaa kohdetta.

Digitointiprosessissa **ei saa** poistaa merkintöjä sisältäviä sivuja. Digitointiprosessissa tuotettujen kuvatiedostojen **pitää** olla käännetty päälukusuuntaan. Digitointiprosessissa tuotettuja kuvatiedostoja **saa** kääntää niiden skannauksen jälkeen vain 90 asteen välein.

5 Aineistotyyppikohtaiset määritykset

Tässä luvussa kuvataan aineistotyyppikohtaisia määrityksiä digitointiprosesseille sekä prosessissa tuotettaville tallekappaleille. Digitaalisten tiedostojen yleiset määritykset esitetään luvussa 6. Aineistotyyppit on kuvattu taulukossa 1. Luvussa 5 aineistotyyppinä on yhdistelty, mikäli aineistotyyppiin kohdistetut määritykset ovat keskenään samanlaisia.

5.1 Kartta- ja piirustuskokoelmat

Erillisinä kokoelmina olevat kartta- ja piirustuskokoelmat **pitää** digitoida kuvatiedostoiksi siten, että syntyvä kuvatiedosto on fyysisiltä mitoiltaan (n cm x n cm) 1:1-koossa analogisen asiakirjan kanssa resoluutiolla 300 ppi. Digitoitavan analogisen asiakirjan tietosisällön eheyden varmistamiseksi analoginen 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 digitoidaan värillisiksi kuvatiedostoiksi (RGB, 24 bit).

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

Alla olevassa taulukossa (Taulukko 4) esitetään kartta- ja piirustuskokoelmien digitoinnin määrittäykset.

Taulukko 4, Kartta- ja piirustuskokoelmien määrittäykset

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)

5.2 Sidokset, joissa tietosisältö esitetään aukeamittain

Mikäli sidosten sisältämä tietosisältö esitetään aukeamittain, **pitää** sidosten digitoinnissa soveltaa tässä luvussa esitettyjä vaatimuksia. Tällaisia sisältöjä voivat olla esimerkiksi sivun ylittävät taulukot tai muu tietosisältö, joka on alun perin tuotettu analogiseen asiakirjaan niin, että ymmärrettävyyden takia analogisen asiakirjan yksittäisiä sivuja on tarkasteltava rinnakkain eli aukeamittain (Kuva 2). Tietosisällön eheyden varmistamiseksi analogiset asiakirjat **pitää** digitoida kokonaisuudessaan siten, että kuva-ala ylittää sivujen reunat.

19.1.2021

KA/15780/07.01.02.04.02/2020

Aukeama										
Sivu					Sivu					
<i>Turun ja Pöytä</i>					<i>laaisia vuonna 1907</i>					8
Periätkirja	Kirkkukuja	Pitäjä	Kylä	Tila	Tilan nimi	Periätkirjoittajan nimi	Periätkirjoittajien			
N:o	Astapani			Numero, Talokki, Maattali			Alku jolloin se on tilalle rekisteröity	Koko periätkirjoittajien	$\frac{1}{2}$ osan periätkirjoittajien	
1907					1907					
1.	Huhtala	Machon	Siedon	Ingoisten	2.	1.	13			
					Ingoisten	Talokki	Kaarle Tuoma	1907	93.58	
2.	Mäntylä	Thorsenmaan	Semlantiin	Teterängän	3.	$\frac{1}{2}$	$\frac{1}{2}$			
					Storgård	Stor Karl Johan Edholm ja	Lohakall	1907	3.78	

Kuva 2, Tietosisältö esitetään aukeamittain

Alla olevassa taulukossa (Taulukko 5) esitetään määritykset aukeamittain digitoitavien sidosten digitointiin.

Taulukko 5, Sidosten määritykset

Elementti	Vaadittu arvo
Tallekappaleen formaatti	TIFF 6.0 tai JPEG part 1 versio
Tallekappaleen pakkaaminen	TIFF = LZW tai JPEG = JPEG-laatuso 90 %
Resoluutio	300 ppi

⁵ Kuvatiedosto ei ole laadullinen referenssi, vaan esimerkki tietosisällöstä, joka on jakautunut aukeamalle.

Elementti	Vaadittu arvo
Väritila	2 = RGB
Bittisyvyys	8 bittiä kanava, 3 kanavaa (24 bit)
Tallekappaleen ICC-profiili	eciRGB v2, ProPhoto RGB, AdobeRGB (1998), sRGB

5.3 Valokuvakokoelmat

Valokuvien sisältämän tietosisällön eheyden varmistamiseksi analogiset asiakirjat **pitää** digitoida kokonaisuudessaan siten, että kuva-ala ylittää sivujen reunat. Valokuvat **pitää** digitoida värillisiksi kuvatiedostoiksi (RGB, 24 bit). Alla olevassa taulukossa (Taulukko 6) esitetään valokuvakokoelmien digitointiin kohdistuvat määritykset.

Taulukko 6, Valokuvakokoelmien määritykset

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 asiakirjan on pienikokoinen, **pitää** resoluutiota nostaa alla olevan taulukon (Taulukko 7) mukaisesti. Mikäli analoginen asiakirja on fyysisiltä mitoiltaan pienempi kuin alla olevassa taulukossa 7 esitetään, **pitää** digitoinnissa käytettävää resoluutiota kasvattaa samaa laskentakavaa 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.

Taulukko 7, Resoluutiotaulukko

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

5.4 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 negatiivit **pitää** digitoida värillisiksi kuvatiedostoiksi (RGB, 24 bit). Alla olevassa taulukossa (Taulukko 8) esitetään lasinegatiivien, muiden negatiivien ja diojen digitoinnin määritykset.

Taulukko 8, 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 digitoinnin kohteena oleva analoginen asiakirja on pienikokoinen, **pitää** resoluutiota nostaa resoluutiotaulukon (Taulukko 7) mukaisesti. Mikäli analoginen asiakirja on fyysisiltä mitoiltaan pienempi kuin taulukossa 7 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.

5.5 Mikrofilmit

Mikrofilmin pienennyskerroin **pitää** tarkistaa filmirullan alusta. Mikäli pienennyskerrointa ei ole ilmoitettu, suurennuskerroin **pitää** asettaa riittävän suureksi, jotta tuotetun digitaalisen asiakirjan tietosisältö on luettavissa. Mikrofilmien osalta saadaan filmille kuvattu aineisto suurennuskertoimen avulla toistumaan riittävän luettavana resoluutiolla 300 ppi. Skannauksessa käytettävä resoluutio on suurennuskerroin x haluttu resoluutio (esimerkiksi 12 x 300 = 3600). Mikrofilmi itsessään on varsin jyrkkä tallenneväline, joten digitoinnissa **pitää** varoa jyrkkyyden korostamista, jotta luvussa 4 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 9) esitetään mikrofilmien digitoinnin määritykset.

Taulukko 9, 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-laatusuhteus 90 %
Resoluutio	300 ppi
Väritila	Harmaasävy
Bittisyvyys	8 bittiä kanava, 1 kanava (8 bit)
Tallekappaleen ICC-profiili	Gray gamma 2,2

Mikäli kyseessä on värillinen mikrofilmi, siihen **pitää** soveltaa tässä luvussa esitettyjä vaatimuksia, mutta tallekappaleen väritilan **pitää** olla RGB, bittisyvyyden **pitää** olla 24 bit ja tallekappaleen ICC-profiiliin **pitää** olla eciRGB v2, ProPhoto RGB, AdobeRGB (1998) tai sRGB.

6 Hyväksyttävät formaatit

Formaattiosio on jaettu kahteen alaluokkaan:

1. Kuvatiedosto
2. Kuvatiedostoa ja kuvatiedoston prosessoinnin metatiedot ja rakenne

6.1 Kuvatiedosto

Digitointiprosessissa tuotettava kuvatiedosto **ei saa** missään käsittelyvaiheessa olla tässä luvussa esitettyjä vaatimuksia heikkolaatuisempi. Tallekappale **pitää** tallentaa joko pakkaamattomassa/häviöttömästi pakatussa TIFF-muodossa tai häviöllisesti pakatussa JPEG-muodossa (ei molemmissa).⁶ Tallekappaleen tallennusformaatti riippuu analogisen aineiston olomuodosta. Aineistotyyppikohtaiset määritykset esitetään luvussa 5 ja sen alaluvuissa. Mikäli prosessissa ensimmäisenä tuotettu kuvatiedosto (tallekappale), on häviöllisesti pakattu, sitä **ei saa** käsitellä ensimmäisen tallennuskerän jälkeen ja tallentaa tämän jälkeen uudelleen tallekappaleeksi.

Alla olevissa taulukoissa (Taulukko 10 ja Taulukko 11) esitetään tiedostoformaateittain pakolliset tiedot, jotka kuvatiedostossa **pitää** olla koneymmärrettävässä muodossa. Mikäli taulukoiden "Elementti"-saraketta ei tarkenneta, **pitää** tieto ilmaista, mutta tiedolle ei ole tässä yhteydessä määritelty vaadittavaa kenttää. Taulukossa esitettyjen tietojen lisäksi kuvatiedosto **saa** sisältää muita metatietokenttiä. Kuvatiedostoihin **ei saa** tallentaa metatietoja, jotka määritellään henkilö tiedoiksi.

Taulukoissa 10 ja 11 esitetään kaikki mahdolliset arvot, joita tässä ohjeessa määriteltyjen kuvatiedostojen tallennusformaattien osalta **pitää** ilmaista. Taulukoiden 10 ja 11 arvoja **pitää** tulkita yhdessä luvussa 5 ja sen alaluvuissa esitettyjen aineistotyyppikohtaisten vaatimusten kanssa. Edellä mainitut kohdat ilmaistaan sarakkeessa "Vaadittu arvo" arvolla "Aineistotyyppi riippuvainen".

Mikäli sarakkeessa "Vaadittu arvo" lukee "Kuvatiedostokohtainen", on jokaisella kuvatiedostolla yksilöllinen arvo. Mikäli sarakkeessa lukee "Tuotantoprosessiriippuvainen", on arvo riippuvainen prosessista, miten digitaaliset kuvatiedostot on kulloinkin tuotettu.

Taulukko 10, Kuvatiedoston (TIFF) pakolliset metatiedot

Elementti	Tarkenne	Vaadittu arvo	Metatieto-skeema	Metatietokenttä
Formaatti	TIFF	image/tiff		MIME Type
Versio	TIFF-tiedoston vaadittu versio	6.0		
Kuvan nimi	Kuvan nimi ilmaistaan neljällä numerolla 0001.tif,			

⁶ JPEG: <https://jpeg.org/jpeg/index.html> Viitattu 11.1.2021

TIFF: <https://www.itu.int/itudoc/itu-t/com16/tiff-fx/docs/tiff6.pdf> Viitattu 11.1.2021

OHJE

19.1.2021

KA/15780/07.01.02.04.02/2020

Elementti	Tarkenne	Vaadittu arvo	Metatieto- skeema	Metatietokenttä
	0002.tif, 0003.tif jne. Ks. luku 7.			
Kuvatiedoston koko				
Väritila	Kuvatiedoston väritila	Aineistotyyppi riippuvainen. 0 = WhitesZero tai 1 = BlacksZero tai 2 = RGB	Exif.Image	PhotometricInterpretation (262)
ICC-profiili	Kuvatiedoston metatietoihin tallennettu (embedded) väriprofiili.	Aineistotyyppi riippuvainen. sRGB, eciRGB v2, ProPhoto RGB, AdobeRGB (1998) tai Gray gamma 2,2	ICC	profileDescription
Bittisyvyys	Bittien määrä pikselin kanava-arvossa	8	Exif.Image	BitsPerSample (258)
	Kanava-arvojen määrä pikselissä	Aineistotyyppi riippuvainen. 3 tai 1	Exif.Image	SamplesPerPixel (277)
Tiedoston pakkaaminen		Aineistotyyppi riippuvainen: 1 = Uncompressed tai 5 LZW	Exif.Image	Compression (259)
Kuvan leveys	Kertoo kuvan leveyden pikselien määrällä per rivi	Kuvatiedostokohmainen	Exif.Image	ImageWidth (256)
Kuvan korkeus	Kertoo kuvan korkeuden pikselirivien määrällä kuvassa	Kuvatiedostokohmainen	Exif.Image	ImageLength (257)
Digitaalisen kuvatiedoston tekijä	Organisaatio, joka on luonut kuvatiedoston analogisesta ilmentymästä	Tuotantoprosessi riippuvainen	Exif.Image	Artist (315)
Digitointilaitte (skannaus tai kuvaus)	Kertoo minkä valmistajan laitteella analoginen ilmentymä on muutettu digitaaliseen muotoon (valmistajan nimi)	Tuotantoprosessi riippuvainen	Exif.Image	Make (271)
Digitointilaitteen malli (skannaus tai kuvaus)	Tarkentaa digitointilaitetta kertomalla valmistajan mallin nimen	Tuotantoprosessi riippuvainen	Exif.Image	Model (272)
Digitoinnissa käytetyn laitteen sarjanumero	Yksilöi käytetyn laitteen	Tuotantoprosessi riippuvainen	Exif.Image	CameraSerialNumber (50735)
Kuvatiedoston luomisessa käytetty ohjelma	Sovellus ja versio, jolla digitaalinen tiedosto on luotu	Tuotantoprosessi riippuvainen	Exif.Image	Software (305)
Kuvatiedoston luontipäivämäärä ja aika (skannauspäivämäärä)	Ilmaistaan muodossa: YYYY:MM:DDTHH:MM:SS	Kuvatiedostokohmainen	Exif.Image	DateTimeOriginal (36867)
Lukusuunta	Tiedoston lukusuunta (vaaka tai pysty). Lukusuunta ei ota kantaa kuvan tietosisällön lukusuuntaan, vaan tässä ilmaistaan tiedoston lukusuunta.	Tuotantoprosessi riippuvainen/ Kuvatiedostokohmainen	Exif.Image	Orientation (274)

OHJE

19.1.2021

KA/15780/07.01.02.04.02/2020

Elementti	Tarkenne	Vaadittu arvo	Metatieto- skeema	Metatietokenttä
Resoluution mittayksikkö	Mittayksikkö, jota käytetään tulkittaessa X ja Y resoluutiota	2 = inch	Exif.Image	ResolutionUnit (296)
XResoluutio	Pikselien määrä resoluution mittayksikköä kohti leveyssuunnassa.	Aineistotyyppi riippuvainen	Exif.Image	XResolution (282)
YResoluutio	Pikselien määrä resoluution mittayksikköä kohti pystysuunnassa.	Aineistotyyppi riippuvainen	Exif.Image	YResolution (283)
Tavujärjestys		big endian tai little endian		ByteOrder
Kuvatiedoston käsittelyohjelma	Mikäli digitointiprosessissa luodaan ensin pakkaamaton tiedosto, jota käsitellään skannauksen jälkeen, tallennetaan käsittelyohjelman nimi ja versio	Tuotantoprosessi riippuvainen	Exif.Image	Image.Processing-Software (11)

Taulukko 11, Kuvatiedoston (JPEG) pakolliset metatiedot

Elementti	Tarkenne	Vaadittu arvo/se- lite	Metatieto- skeema	Metatietokenttä
Formaatti	JPEG	image/jpeg		MIME Type
Versio	JPEG part 1 versio	1.00 tai 1.01 tai 1.02		JFIF Version
Kuvan nimi	Kuvan nimi ilmaistaan neljällä numerolla 0001.jpg, 0002.jpg, 0003.jpg jne, Ks. luku 7.			
Kuvatiedoston koko				
Väritila	Kuvatiedoston väritila	Aineistotyyppi riippuvainen. 0 = WhitesZero tai 1 = BlacksZero tai 2 = RGB	Exif.Image	PhotometricInterpretation (262)
ICC-profiili	Kuvatiedoston metatietoihin tallennettu (embedded) väriprofiili.	Aineistotyyppi riippuvainen. sRGB, eciRGB v2, ProPhoto RGB, AdobeRGB (1998) tai Gray gamma 2,2	ICC	profileDescription
Bittisyvyys	Bittien määrä pikselin kanava-arvossa	8	Exif.Image	BitsPerSample (258)
	Kanava-arvojen määrä pikselissä	Aineistotyyppi riippuvainen. 3 tai 1	Exif.Image	SamplesPerPixel (277)
Tiedoston pakkaaminen	JPEG	6 = JPEG (old-style) tai 7 = JPEG	Exif.Image.	Compression (259)
JPEG-laatu	JPEG-pakkauksen laatu asteikolla 0-100 %	90 %		
Kuvatiedoston tekijä	Organisaatio, joka on luonut kuvatiedoston	Tuotantoprosessi riippuvainen	Exif.Image	Artist (315)

Elementti	Tarkenne	Vaadittu arvo/se- lite	Metatieto- skeema	Metatietokenttä
	analogisesta ilmenty- mästä			
Kuvan korkeus	Kertoo kuvan korkeuden pikselirivien määrällä kuvassa	Kuvatiedostokoh- tainen	Exif.Image	ImageLength(257)
Kuvan leveys	Kertoo kuvan leveyden pikselirivien määrällä kuvassa	Kuvatiedostokoh- tainen	Exif.Image	ImageWidth(256)
Digitointilaitte	Kertoo minkä valmista- jan laitteella analoginen asiakirja on muutettu di- gitaaliseksi (valmistajan nimi)	Tuotantoprosessi- riippuvainen	Exif.Image	Make (271)
Digitointilaitteen malli	Tarkentaa digitointilai- tetta kertomalla valmis- tajan mallin nimen	Tuotantoprosessi- riippuvainen	Exif.Image	Model (272)
Digitoinnissa käytetyn lait- teen sarjanumero	Tarkentaa mallia ja yksi- löi laitteen, jonka avulla analoginen asiakirja on muunnettu digitaaliseksi	Tuotantoprosessi- riippuvainen	Exif.Image	CameraSerialNum- ber (50735)
Kuvatiedoston luomisessa käytetty ohjelma	Sovellus ja versio, millä digitaalinen tiedosto on luotu	Tuotantoprosessi- riippuvainen	Exif.Image	Software (305)
Kuvatiedoston luontipäivä- määrä ja aika	Ilmaistaan muodossa: YYYY:MM:DD HH:MM:SS	Kuvatiedostokoh- tainen	Exif.Image	DateTimeOriginal (36867)
Lukusuunta	Tiedoston lukusuunta (horisontaalinen tai ver- tikaalinen)	Tuotantoprosessi- riippuvainen/ Kuvatiedostokoh- tainen	Exif.Image	Orientation (274)
Resoluution mittausyksikkö	Mittayksikkö, jota käytetään tulkittaessa X ja Y resoluutiota	2 = inch	Exif.Image	Image.Resolu- tionUnit (296)
XResoluutio	Pikselien määrä resoluution mittayksikköä kohden kuvan leveyssuunnassa.	Aineistotyyppiiriip- puvainen/ Kuvatiedostokoh- tainen	Exif.Image	Image.XResolution (282)
YResoluutio	Pikselien määrä resoluution mittayksikköä kohden kuvan korkeussuunnassa.	Aineistotyyppiiriip- puvainen/Kuvatie- dostokohtainen	Exif.Image	Image.YResolution (283)
Kuvatiedoston käsittelyoh- jelma	Mikäli digitointiprosessi- ssa luodaan ensin pak- kaamaton tiedosto, jota käsitellään skannauksen jälkeen, tallennetaan kä- sittelyohjelman nimi ja versio	Tuotantoprosessi- riippuvainen	Exif.Image	Image.Processing- Software (11)

6.2 Kuvatiedostoa ja kuvatiedoston prosessointia kuvaavat metatiedot ja rakenne

Tässä luvussa määritellyt metatiedot kuvaavat kuvatiedoston syntyhistoriaa ja todentavat osaltaan myös prosessissa syntyneen digitaalisen ilmentymän autenttisuutta. Kuvatiedostojen pakolliset tekniset metatiedot **pitää** esittää MIX-metatietoskeeman version 2.0 mukaisesti.⁷

Alla olevassa taulukossa (Taulukko 12) on merkitty vasemmalta oikealle MIX-kentän nimi, kentän tarkoitus vapaasti käännettynä ja velvoite. Velvoite-kentässä ilmaistaan kyseisen kentän ja sen skeeman mukaisen tiedon pakollisuus seuraavalla tavalla:

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

MIX-metatietoskeemassa on kahdenlaisia kenttiä: säiliöitä ja dataelementtejä. Dataelementit sisältävät tietyn arvon, kun taas säiliöt sisältävät yhden tai useamman dataelementin ja ne voivat sisältää toisia säiliöitä dataelementteineen. Taulukossa 12 on kuvattuna vain tietyn arvon sisältäviä kenttiä eli dataelementtejä.

Taulukko 12, Kuvatiedostoa ja sen prosessointia kuvaavat metatiedot (taulukossa on ilmaistu vain tietoa sisältävät kentät, jotka PITÄÄ esittää MIX-metatietoskeeman version 2.0 mukaisessa rakenteessa)

MIX-kentän nimi	Kentän tarkoitus	Velvoite
objectIdentifierType	Dataelementti, joka määrittää järjestelmän tai verkkotunnuksen, jossa digitaalisen asiakirjan yksilöivä ID on uniikki.	P
objectIdentifierValue	Digitaalisen tiedoston 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 TAI 6.0	V
byteOrder	Dataelementti, joka määrittää tavujen tallennusjärjestyksen. Vaadittu arvo on joko 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 jokin seuraavista: MD5, SHA-1, SHA256, SHA384, SHA512.	P
messageDigest	messageDigestAlgorithm kentän määrittämän algoritmin tuotama merkki sarja, 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 tai WhitelsZero tai BlackIsZero.	P

⁷ The Library of Congress » Standards » MIX. Kongressin kirjaston verkkosivu <http://www.loc.gov/standards/mix/> ja ANSI/NISO Z39.87-2006 (R2017) Data Dictionary - Technical Metadata for Digital Still Images <https://www.niso.org/publications/ansiniso-z3987-2006-r2017-data-dictionary-technical-metadata-digital-still-images>
Viitattu 11.1.2020

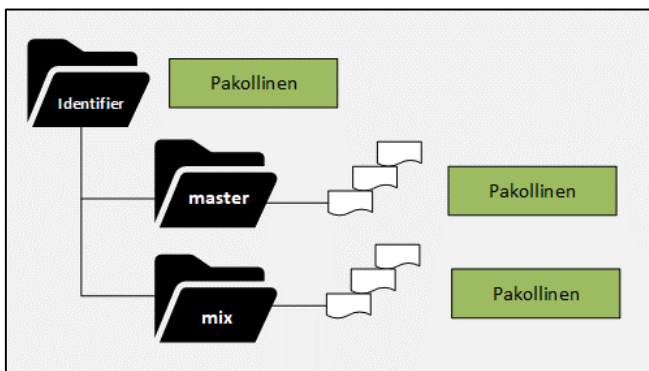
MIX-kentän nimi	Kentän tarkoitus	Velvoite
iccProfileName	Dataelementti, joka määrittää yleisesti käytetyn ICC-profiilin nimen. Vaadittu arvo sRGB, eciRGB v2, ProPhoto RGB, AdobeRGB (1998) tai Gray gamma 2,2.	P
iccProfileVersion	Dataelementti, joka kertoo käytetyn ICC-profiilin version, esimerkiksi v4 [eli sRGB v4].	P
iccProfileURL	Dataelementti, johon tallennetaan ICC-profiilin URL/URN, mikäli ICC-profiili ei ole hyvin dokumentoitu.	V
dateTimeCreated	Dataelementti, joka kertoo kuvatiedoston luontiajan. Ilmaistaan muodossa: YYYY-MM-DDTHH:MM:SS.	P
imageProducer	Dataelementti, joka identifioi digitaalisen tiedoston 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 tarkentaa digitointilaitteen mallin nimeä sen tyyppinumerolla.	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 leveyssuunnassa. Vaadittu arvo on aineistotyyppiin riippuvainen.	P
ySamplingFrequency	Pikselien määrä resoluution mittayksikkö kohden pystysuunnassa. Vaadittu arvo on aineistotyyppiin riippuvainen.	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 joko 3 tai 1.	P
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 yksilöidyn mittataulun mittausdata löytyy.	V

7 Digitointiprosessissa muodostettava siirtopaketti

Luvussa 6 ja sen alaluvuissa mainitut digitointiprosessissa tuotetut tiedostot **pitää** tallentaa alla olevassa kuvassa (Kuva 3) esitetyn hakemistorakenteen mukaisesti, jotta ne voidaan ottaa vastaan

Kansallisarkistoon. Digitaalinen ilmentymä **pitää** tuottaa hakemistorakenteeseen riippumatta siitä, milloin se siirretään Kansallisarkistoon. Aineistoja luovutettaessa siirtopaketti **ei saa** sisältää mitään muuta kuin kuvassa 3 esitettyjä hakemistoja.

Tässä määritellyn hakemistorakenteen lisäksi organisaatio **saa** tallentaa esimerkiksi käyttökappaleet omiin tietojärjestelmiinsä siinä tietorakenteessa, jota kyseinen järjestelmä edellyttää. Tässä asiakirjassa määritelty rakenne ei siis sulje pois muiden mahdollisten tallennusrakenteiden käyttöä.



Kuva 3, Digitointiprosessin vaadittu siirtopakettirakenne

Alla olevassa taulukossa (Taulukko 13) kuvaillaan, miten digitaaliset tiedostot **pitää** nimetä esitetyin hakemistorakenteen sisällä. Prosessissa tuotettujen digitaalisten tiedostojen **pitää** kohdata keskenään, toisin sanoen MIX-metatietoskeeman mukaisen 0001.xml-tiedoston **pitää** sisältää kuvatiedostoa 0001.jpg tai 0001.tif kuvailevia metatietoja.

Taulukko 13, Siirtopaketin hakemistojen sisältö

Hakemisto	Selite
identifier	Tarkoittaa digitaalisen ilmentymän yksilöivää tunnusta, jonka avulla pitää pystyä tunnistamaan, mistä asiakirjakokonaisuudesta on kyse (esimerkiksi arkistoyksikkö). ⁸ Hakemisto sisältää ”digitaalisten tiedostojen hakemistot”.
master	Hakemistoon pitää tallentaa taulukoissa (Taulukko 10, Kuvatiedoston (TIFF) pakolliset metatiedot tai Taulukko 11, Kuvatiedoston (JPEG) pakolliset metatiedot) esitetyt kuvatiedostot yksittäisinä tiedostoina. Tiedostot pitää nimetä nelinumeroisina alkaen 0001.tif tai 0001.jpg.
mix	Hakemistoon pitää tallentaa taulukossa 12 esitetyt pakolliset tiedot koskien kaikkia master-hakemiston sisällä olevia kuvatiedostoja. Tiedostoon SAA tallentaa myös muita MIX-metatietoskeeman mukaisia tietoja, skeeman mukaisessa rakenteessa. Tiedostot pitää nimetä nelinumeroisina alkaen 0001.xml.

Mikäli aineisto toimitetaan Kansallisarkistolle, jokainen siirtopaketti **pitää** paketoita TAR- tai ZIP-paketiksi. TAR- tai ZIP-paketin sisältöä **ei saa** tässä vaiheessa pakata. TAR- tai ZIP-paketille **pitää** laskea tarkistesumma MD5-muodossa ja se **pitää** toimittaa siirron yhteydessä. Kansallisarkistoon aineistoa toimitettaessa tarkoitetaan identifier-hakemistolla AHAA-järjestelmän aineiston tunnistetta Ai01 (roolissa ahaa tekninen).

⁸ Digitoitavaksi päätetyn analogisen ilmentymän pitäisi olla kuvailtuna (kuvaileva metatieto tuotettu) ennen sen digitointia. Identifierin avulla PITÄÄ pystyä yhdistämään digitointiprosessissa syntyneet digitaaliset ilmentymät edellä mainittuun kuvailevaan metatietoon.

19.1.2021

KA/15780/07.01.02.04.02/2020

Esimerkkipaketit

1. EsimerkkipakettiJPEG2020.tar
2. EsimerkkipakettiTIF2020.tar

Master-hakemistojen sisältämät kuvatiedostot eivät ole kuvanlaadullisia referenssejä. Kuvatiedostot sisältävät tässä asiakirjassa pakollisiksi määritellyt metatiedot. Mix-hakemistojen tiedostot ovat esimerkkejä liitteen paketissa olevista master-tiedostoista lukuun ottamatta mahdollisia elementtejä, joiden kohdalla toisin todetaan.

8 Allekirjoitukset

Pääjohtaja

Jussi Nuorteva

Tutkimusjohtaja

Päivi Happonen