



15.8.2008

ARKISTOLAITOKSEN SUOSITUS DIGITOINNIN LAATUKRITEEREIKSI:

SUOSITUS AL/11130/07.01.02.04.02/2008

Sisältö Suositus asiakirjallisen kulttuuriperinnön digitoinnin
laatukriteereiksi

Säännökset, joihin toimivalta suosituksen antamiseen perustuu

Arkistolain 831/94 4 § mukaan arkistolaitoksen tehtävänä on mm.
varmistaa kansalliseen kulttuuriperintöön kuuluvien asiakirjojen
säilyminen ja niiden käytettävyys.

Kohderyhmä Digitointia toteuttavat muistiorganisaatiot ja yhteisöt

Voimassaoloaika Toistaiseksi

Lisätietoja Suosituksen soveltamista koskevia lisätietoja antaa
Kansallisarkiston arkistotekninen yksikkö

JOHDANTO

Digitoinnilla edistetään asiakirjallisen kulttuuriperinnön käyttöä ja säilyttämistä. Suosituksen tarkoituksena on antaa muistiorganisaatioiden ja yhteisöjen käyttöön yhtenäiset ja riittävän korkeatasoiset käytänteet kulttuuriperintöaineiston digitointiin. Suositus painottuu asiakirjallisen aineiston, karttojen ja mikrofilmien digitointiin. Suositusta tullaan yhdessä muiden muistiorganisaatioiden kanssa laajentamaan siten, että se koskee myös ääntä, liikkuvaa kuvaa, multimediaa sekä muuta digitoitua kulttuuriperintöä.

Suositus pohjautuu Kansallisarkiston arkistoteknisen yksikön digitoinnin käytänteisiin, sisäiseen ohjeeseen 19.5.2006 ”Pysyvästi säilytettävät sähköiset audiovisuaaliset aineistot”, vastaaviin ulkomaisiin ohjeistuksiin sekä muistiorganisaatioilta ja maakunta-arkistoilta saatuihin kommentteihin ja lausuntoihin.

Suosituksessa huomioidaan digitoinnin teknisten kriteerien lisäksi digitaalisen aineiston pitkäaikaissäilytyksen tarpeita sekä aineiston verkkokäyttö. Mikään tekninen laatuohjeistus ei kuitenkaan varmista, että tiedostot automaattisesti säilyisivät pitkään. Digitaalisen aineiston pitkäaikaissäilytys vaatii omat säilytysstrategiansa. Suosituksessa ei anneta ohjeistuksia metatietojen¹ kriteereistä.

Muulla kuin arkistolaitoksessa digitoidun aineiston tuominen Digitaaliarkistoon <http://digi.narc.fi/digi/> edellyttää suosituksen noudattamista digitointitoiminnassa. Teknisesti huonompilaatuista ulkopuolista aineistoa ei oteta vastaan Digitaaliarkistoon.

SÄHKÖISESSÄ MUODOSSA SYNTYNYT AINEISTO

Asianhallintajärjestelmien metatietoja, toiminnallisia ja teknisiä ominaisuuksia koskeva määräys ”Asiankäsittelyjärjestelmiin sisältyvien pysyvästi säilytettävien asiakirjallisten tietojen säilyttäminen yksinomaan sähköisessä muodossa” (KA 1486/40/2005) ja SÄHKE -määritykset (<http://www.narc.fi/Arkistolaitos/sahkoinen>) poikkeavat digitoinnin laatuksiteereistä. Tiedostojen alkuperäisyys ja eheys varmistetaan siinä vaiheessa, kun tiedosto on asianhallintajärjestelmässä luotu tai kun se on sinne tallennettu. Metatiedot siirretään tiedostojen mukana asianhallintajärjestelmään, jolloin alkuperäisyys ja eheys sekä tiedostomuotoihin kohdistuneet toimenpiteet ovat myöhemminkin todennettavissa.

Tämä suositus ei koske asiankäsittelyjärjestelmissä tuotettuja sähköisiä aineistoja.

¹ Metatiedoista mm. julkaisuissa:

- Arkistolaitos, SÄHKE -määritykset <http://www.narc.fi/Arkistolaitos/sahkoinen/>

- Dublin Core metatietosanakstostandardi <http://dublincore.org/>

- Salminen, A. 2005. Metatiedot organisaatioiden sisällönhallinnassa. Artikkelit saatavilla pdf-muodossa: <http://www.cs.jyu.fi/~airi/papers/Metatietoartikkeli-2005.pdf>

ASIAKIRJALLISEN KULTTUURIPERINNÖN DIGITOINTI

Analogiselle aineistolle on ominaista sen informaation jatkuvuus. Värisävyt ja tummuusasteet muuttuvat portaattomasti. Aineiston kopioimisessa analogisesti kuvasta tulee jokaisella kopiointikerralla hieman epätarkempi. Digitointi on analogisen aineiston muuttamista digitaaliseen muotoon. Digitaalinen informaatio on tallennettu numeroina, ykkösinä ja nolliina. Digitaalinen kuva koostuu tietyin välein otetuista näytteistä. Jos näytteiden ottotaajuus eli resoluutio on riittävän suuri, saadaan vaikutelma sävyjen jatkuvuudesta. Digitaalista informaatiota voidaan monistaa ilman, että sen laatu heikkenee.

Digitoinnin tarkoituksena on alkuperäisen aineiston sisältämän informaation saattaminen laajemmin käyttöön. Samalla alkuperäiset aineistot saadaan suojattua. Digitointi on osa kulttuuriperintöaineiston säilyvyyden ja käytettävyyden edistämisen prosessia. Tätä prosessia ei tässä suosituksessa esitellä tarkemmin, koska kullakin muistiorganisaatiolla ovat omat käytänteensä. Aineiston esikäsittelyyn kuuluvat konservointitoimenpiteet, kuten puhdistus, niittien poisto, suoristus, on syytä tehdä huolellisesti.

Asiakirjat digitoidaan kuvatiedostoiksi siten, että syntyvä tiedosto on 1:1 koossa alkuperäisen kanssa. Aineiston ollessa pienikokoista, kuten valokuvien kohdalla, on harkittava resoluution nostamista minimivaatimuksista. Mikrofilmien osalta suurennuskertoimen avulla saadaan kuvattu aineisto toistumaan riittävän luettavana myös minimiresoluutiolla. Asiakirja-aineiston luotettavuuden varmistamiseksi aineisto digitoidaan kokonaisuudessaan siten, että kuva-ala ylittää sivujen reunat.

DIGITOINNIN OSA-ALUEET:

Valokuvaaminen digitaalikameralla:

Valokuvaaminen digitaalikameralla ei aina täytä arkistolaitoksen laatukriteerejä. Kameran objektiivi saattaa vääristää kohteen mittasuhteita ja linjoja. Tiedoston resoluutio ja tiedostomuoto saattavat poiketa suosituksesta. Mikäli digitointityö on tarkoitus toteuttaa digitaalikameralla, on syytä ensin ottaa yhteyttä Kansallisarkiston arkistotekniseen yksikköön laadun ja yhteensopivuuden varmistamiseksi.

Väridigitointi:

Väridigitointi tulee tehdä siten, että digitaalisten kuvien värit ja sävyt vastaavat mahdollisimman tarkasti alkuperäisaineistoa. Tämä on mahdollista vain kalibroidussa ja profiloidussa työympäristössä. Monitori on kalibroitava mittalaitetta käyttäen ja skannerille on tehtävä yksilöllinen profiili. Skannerin profiili on tallennettava kuvan mukana, ja/tai kuvan väriavaruus olisi muunnettava johonkin laajatoistaiseen kuva-avaruuteen (esim. Adobe98, tai eciRGB). Jos ammattimaiseen kalibrointiin ei ole mahdollisuutta, jokaisen työerän yhteydessä on skannettava ja tallennettava mukaan jokin yleisesti saatavilla oleva värikartta, josta värien ja sävyjen toistumisen voi tarkistaa. Kuvan vaaleimmat ja tummimmat värisävyt eivät saa leikkautua.

Väridigitoinnin minimiresoluutio on 300 ppi² ja minimibittisyys 24 bit. Tallennusmuotona pakkaamaton TIFF. Minimiarvoilla esimerkiksi A-4 kokoisen arkin tiedostokoko on noin 25-30 Mt. Pienemmissä kohteissa, kuten valokuvavedoksissa ja diakuvissa, tarvitaan huomattavasti suurempaa resoluutiota. Ns. mustavalkoiset valokuvavedokset tulee digitoida värillisinä.

Harmaasävydigitointi:

Harmaasävydigitointi tulee käsitteenä erottaa mustavalkodigitoinnista. Harmaasävydigitoinnissa sävyskaala on jatkuva mustan ja valkoisen välillä. Kontrastin sekä tummuuden säädöillä varmistetaan tekstin luettavuus ja kohteen toistuminen mahdollisimman alkuperäisenä. Paperin vaaleimmilla ja tekstin tummimmilla alueilla on oltava näkyvissä harmaasävyjä.

Mikrofilmiltä digitointi toteutetaan aina harmaasävydigitointina (ainoana poikkeuksena harvinaiset värilliset mikrofilmit), jotta filmillä olevat harmaasävyt toistuvat ja tekstin himmeimmätkin detaljit eivät katoa. Mikrofilmin pienennyskerroin on tarkistettava filmirullan alusta. Mikäli pienennyskerrointa ei ole ilmoitettu, tulee suurennuskerroin asettaa riittävän suureksi, jotta digitoitu asiakirja on luettavissa. Kuva-alalta rajataan pois kaikki ylimääräinen.

Harmaasävydigitoinnin minimiresoluutio on 300 ppi ja bittisyys 8-bit. Tallennusmuotona pakkaamaton TIFF. Minimiarvoilla esimerkiksi A-4 kokoisen arkin tiedostokoko on 8 Mt.

Mustavalkodigitointi:

Mustavalkodigitointi on 1-bittistä digitointia, jossa tuloksena on vain joko mustaa tai valkoista, ei harmaasävyjä. Kuva on hyvin graafinen. 1-bittinen digitointi ei sovellu historiallisen käsinkirjoitetun asiakirja-aineiston tai mikrofilmien digitointiin. Tämänlaisessa aineistossa tekstin tummuus ja paperin pohjan väri vaihtelevat, jolloin riski informaation katoamiseen on suuri. Vanhan asiakirjan autenttisuus ja luettavuus menetetään 1-bittisellä digitoinnilla. Modernin tulostetun asiakirja-aineiston kohdalla digitointi 1-bittisenä saattaa olla mahdollista, tällöin suositellaan 300-600 ppi minimiresoluutiota. Soveltuvuus on arvioitava etukäteen aineistokohtaisesti.

Äänitiedostojen digitointi:

Äänitiedostojen osuus arkistolaitoksen digitaaliarkistossa on hyvin vähäinen. Seuraavana esitellään vain digitoinnin tekniset kriteerit.

Talletiedostot: WAV, 24 bit, 48-96 kHz

Käyttötiedostot: MPEG-1 Audio Layer 3 (tiedostotunnus MP3), 128kb/s, kiinteä näytetaajuus

² ppi (pixel per inch) tarkoittaa skannausresoluutiota, ei tulostusresoluutiota dpi (dotch per inch)

DIGITOIDUN AINEISTON HALLINTA:

Arkistolaitoksen Digitaaliarkistossa olevan digitoidun aineiston hallinnassa noudatetaan arkistolaitoksen digitointistrategian linjauksia. Pitkäaikaissäilytyksen ratkaisut tulevat lähivuosina täsmentymään muistiorganisaatioiden yhteistyön Kansallinen Digitaalinen Kirjasto -hankkeessa.

Arkistolaitoksella on omat käytänteensä, mutta ei virallista suositusta pitkäaikaissäilytykseen käytettävistä tallennusmedioista. On huolehdittava siitä, että tiedot tuoreutetaan eli kopioidaan uudelle medialle riittävän usein. Sopivana tuoreutusaikana voidaan pitää optisilla ja magneettisilla medioilla noin viittä vuotta, mikäli mediat säilytetään valmistajan ohjeiden mukaisesti. Lisäksi on huolehdittava siitä, että tiedot ovat vähintään kahdella mielellään eri valmistajan erillisellä medialla.

Arkistolaitoksen Digitaaliarkistossa korkealaatuinen pakkaamattomassa TIFF -tallennusmuodossa oleva digitoitu aineisto talletetaan kahtena identtisenä talletiedostona nauhatallennusjärjestelmään. Talletiedostot kopioidaan uudelle nauhalle tai nauhamedialle viiden vuoden välein. Nauhamediaa vaihdetaan tarvittaessa uudemman nauhatyyppin tultua korvaamaan aiempaa. Ohjelmallisesti TIFF- tiedostoista tuotetaan kaksi pienempää jpg-tiedostoa käyttötiedostoiksi, jotka talletetaan palvelimelle selailukuviksi internetkäyttöön. Käyttötiedostot ovat erikokoisia; toista voidaan näytöllä suurentaa tarkemman luettavuuden saavuttamiseksi.

Kuvatiedostoina Digitaaliarkistoon otetaan vastaan vain pakkaamaton TIFF v. 6.0 muodossa. Tiedostonimi saa sisältää vain merkit A-Z sekä numerot 0-9 ja alaviivan ilman välilyöntejä. Erikoismerkkejä tai skandinaavisia kirjaimia, kuten ÅÖ, ei hyväksytä.

Pääjohtaja

Jussi Nuorteva

Kehittämispäällikkö

István Kecskeméti