

Yhdistysten sähköisten asiakirjojen arkistointi Toimihenkilöarkistoon



1. Yleistä

Yhdistykset tuottavat asiakirjansa nykyään suurimmaksi osaksi sähköisinä. Ne kannattaa tallentaa sellaisessa muodossa, että ne ovat myös arkistoitavissa sähköisinä myöhempää käyttöä varten. Tämä ohje on tarkoitettu pienten sähköisten asiakirjaerien tuottamiseen ja siirtämiseen Toimihenkilöarkistoon (THA). Ohje keskittyy erityisesti tekstiasiakirjoihin. Se on laadittu ensisijaisesti THA:ssa tällä hetkellä käytössä olevaa sähköistä arkistopalvelua silmällä pitäen. Tavoitteenamme on, että toimihenkilöjärjestöjen sähköiset arkistoaineistot ovat löydettävissä myös arkistojen, museoiden ja kirjastojen yhteisestä hakupalvelusta Finnasta (<http://www.finna.fi>). Täten THA:n vaatimusten on noudatettava myös Kansallisen digitaalisen kirjaston ja arkistolaitoksen vaatimuksia.

Käy tutustumassa sähköiseen arkistoomme osoitteessa <http://www.th-arkisto.fi/sahkoinen-arkisto>

Pienimuotoinenkin sähköinen arkistointi edellyttää suunnittelua. Esimerkiksi sähköisen aineiston vastaanottamisen edellytys arkistoomme on, että yhdistyksen arkiston runko eli sarjarakenne on olemassa. Sen hoitaa THA. Kun alatte yhdistyksessä suunnitella asiakirjojen säilyttämistä sähköisinä, ottakaa jo alkuvaiheessa yhteys Toimihenkilöarkistoon (tha@th-arkisto.fi tai puhelin 040-773 4997). Näin kaikki sujuu mutkattomasti ja turhalta työtä puolin ja toisin välttyään.

2. Sanastoa

Tässä on muutamia digitoinnin yhteydessä toistuvia termejä, joihin on hyödyllistä tutustua. Niiden tunnistaminen auttaa esimerkiksi skannerin asetusten säätämisessä (kuvatyyppi, arvojen muuttaminen ym.).

Bittisyys. Värisyvyys, joka vaikuttaa kuvan laatuun (esimerkiksi 1 bit, 8 bit, 24 bit). Määritetään skannerin asetuksissa (kuvatyyppi).

Digitointi. Paperiasiakirjan kopioiminen sähköiseen muotoon useimmiten skannerilla, joskus valokuvaamalla.

Formaatti. Yleisimpiä kuvatiedostomuotoja ovat esimerkiksi BMP, JPEG, PNG ja TIFF. Digitoitava aineisto on ensisijaisesti tallennettava PDF/A -muotoon.

Metadata tai metatieto. Asiakirjoja ja tiedostoja kuvailevat tiedot. Kaikki metatiedot eivät tule käyttäjien nähtäville, mutta sähköinen arkistojärjestelmä tarvitsee niitä.

Resoluutio. Kuvan tarkkuus pikseliä (kuvapistettä) pituusyksikköä kohti ppi tai dpi. Tulostimet tuottavat pistetiheyden (dpi, *dots per inch*) ja skannerit tai digikamerat pikselitiheyden (ppi, *points per inch*). Määritetään skannerin asetuksissa.

Väriprofiili. Digitaalisten kuvien värit saattavat muuttua, kun kuvatiedosto siirtyy laiteympäristöstä toiseen. Väriprofiili on kokoelma dataa, joka määrittelee laitteen väriavaruuden (esimerkiksi sRGB). Skannerin asetuksissa pääset määrittämään ICC- tai ICM-profiilin kohteille eli skannatuille asiakirjoille (tiedostoille).

3. Vaatimukset vastaanotettavalle sähköiselle aineistolle

Hyväksyttävät tiedostomuodot

Tekstiasiakirjat hyväksytään sähköiseen arkistoon PDF/A-tiedostoina. Sähköisinä tuotettavat asiakirjat, joihin ei tule allekirjoituksia, voidaan sellaisinaan tallentaa PDF/A-muotoon¹. (Microsoft Word -ohjelmassa: Tallenna nimellä PDF -> Asetuksista valitse ISO 19005-1 -yhteensopiva

PDF/A). Jos PDF/A-muotoisten tiedostojen tuottaminen on yhdistykselle aivan mahdotonta, THA voi ottaa vastaan myös TIFF-muotoisia tiedostoja.

Muut vaatimukset

Kuvatiedostojen on oltava 1:1 -koossa alkuperäisen asiakirjan kanssa. Skannattaessa kuva-alan on oltava suurempi kuin kuvattu asiakirja. Kuvan reunan ja asiakirjan reunan välissä on oltava ”tyhjää”, jotta lukija voi olla varma siitä, että asiakirja on kuvattu kokonaan. Tämä on tärkeää etenkin silloin, kun asiakirjan marginaalit ovat hyvin kapeat tai niitä ei ole lainkaan.

Kuvatyyppin valinta

Asiakirjoja skannattaessa on otettava huomioon asiakirjan ulkoasu ja ikä. Ne vaikuttavat siihen, valitaanko väri-, harmaasävy- tai mustavalkodigitointi. Mustavalkodigitointi on 1-bittistä digitointia ja soveltuu vain uudehkojen kokonaan mustavalkoisten tulostettujen asiakirjojen digitointiin. Huomioi, että minimiresoluutio on 300–600 ppi (skannausresoluutio, pixel per inch).

Käytä **harmaasävydigitointia** aina, kun se on mahdollista. Mustavalkoisissa asiakirjoissa, erityisesti vanhoissa käsin ja koneella kirjoitetuissa, musteen ja paperin sävyt vaihtelevat. Mustejälki saattaa olla kuitenkin niin heikkoa, että mustavalkodigitoinnillakaan sitä ei saada näkyviin. Myös sellaiset värilliset asiakirjat, joissa väreillä ei ole merkitystä sisällön ymmärtämisen kannalta, voidaan digitoida harmaasävydigitointina. Luettavuus ja asiakirjan toistuminen mahdollisimman alkuperäisenä varmistetaan kontrastin ja tummuuden säädöillä.

→Harmaasävydigitointi on 8-bittistä.

→Minimiresoluutio on 300 ppi.

→Minimiarvoilla A4-asiakirjan tiedostokoko harmaasävydigitoinnissa on noin 8 Mt.

Arkistolaitoksen vaatimukset täyttävä **väridigitointi** edellyttää monitorin kalibrointia mittausratkaisulla ja yksilöllisen väriprofiilin laatimista skannerille. Profiili on tallennettava kuvan mukana. Jos ammattimaiseen kalibrointiin ei ole mahdollisuutta, jokaisen työerän mukaan on skannettava ja tallennettava yleisesti saatavilla oleva värikartta, josta värien toistuminen voidaan tarkistaa (esimerkiksi <http://www.qpcard.com>). Koska väridigitointi on työlästä, THA ei suosittele sitä yhdistyksille.

→Minimiresoluutio on 300 ppi.

→Minimibittisyys 24 bit.

→Väridigitoinnissa minimiarvoilla A4-asiakirjan tiedostokoko on 25–30 Mt.

Metatiedot

Asiakirjojen metatietojen osalta noudatetaan Yksityiset Keskusarkisto ry:n digitointihankkeen vaatimuksia, jolloin THA:n sähköisessä arkistossa olevan aineiston metatieto on yhdenmukaista. Metatietojen tallentamista varten saat THA:sta esitäytetyn Excel-taulukon, johon on mahdollista syöttää minimivaatimuksia enemmän tietoja. Vaadittu metadata esitetään taulukossa 1.

¹ Suositellut PDF/A-1a, PDF/A-1b, (ISO 19005-1:2005), PDF/A-2a, PDF/A-2b, PDF/A-2u (ISO 19005-2:2011).

TAULUKKO 1. Esimerkki täytetystä metadatataulukosta

1. Tekninen metadata		Lisätietoja
Tiedostonimi		Esim. THA_100_Da_1_1 (THA_arkistonmuodostajan numero_arkistosarja_arkistöyksikön numero_juokseva numero). Sovi etukäteen THA:n kanssa.
Kuvan korkeus pikseleinä		Löytyy TIFF-tiedoston ominaisuuksista ja PDF/A-tiedostosta ainakin Adobe Photoshopin avulla.
Kuvan leveys pikseleinä		Löytyy TIFF-tiedoston ominaisuuksista ja PDF/A-tiedostosta ainakin Adobe Photoshopin avulla.
Kuvan resoluutio (pikselitiheys)		Numerotieto. Määritetään skannausasetuksista: tarkkuus (ppi).
Tiedostokoko		Löytyy tiedoston ominaisuuksista.
Väriprofiili		Skannausasetuksissa: määritykset -> väri. TIFF-tiedoston ominaisuudet. Usein RGB.
Tarkistesumma		MD5-tiiviste. THA hoitaa tarvittaessa.
Bittisyys		Numerotieto. Määritetään skannausasetuksista: kuvatyypin.
Formaatti		esim. TIFF, PDF/A
Laji		Viittaa alkuperäiseen dokumenttiin, ei tiedostotyyppiin. Valokuva, piirros=image; tekstiasiakirja=text
2. Hallinnollinen metadata		Lisätietoja
Digitointipaikka, digitoija		Yritys tai yhdistys, henkilö
Digitointiaika		Päiväys / kellonaika. Tiedoston ominaisuudet -> luotu
Laite		Käytetyn laitteen merkki
Malli		Käytetyn laitteen malli
Kuvankäsittelypaikka, käsittelijä		Yritys tai yhdistys, henkilö
Kuvankäsittelyaika		Päiväys / kellonaika. Tiedoston ominaisuudet -> muokattu.
Kuvankäsittelyohjelma		Jos kuvaa käsitelty, millä ohjelmalla.
3. Kuvaileva metadata		Lisätietoja
Nimike		Asiakirjan otsikko. Esim. Toimintakertomus vuodelta 1994, Liittokokouspöytäkirja 14.05.1966, Kertomus yhdistyksen retkestä Lappiin 16.06.2012. Tarvittaessa THA auttaa.
(Aihe)		Mikäli ei käy ilmi nimikkeestä. Kuvan tai asiakirjan aihe. Tarvittaessa kysy THA:lta.
(Kuvaus)		Mikäli ei käy ilmi nimikkeestä. Oleelliset tiedot. Kysy tarvittaessa THA:lta.
(Alueellinen kattavuus)		Paikka, mikäli merkityksellinen.
(Ajallinen kattavuus)		Aika, mikäli merkityksellinen.

Tiedostojen nimeäminen ja kansiorakenne

Arkistoitavat tiedostot kannattaa nimetä niin, että tiedoston nimestä käy ilmi yhdistyksen nimi (tai sen numero THA:n arkistojärjestelmässä) sekä se, mistä asiakirjasarjasta on kysymys (esimerkiksi 100_Cb_xxx tai 100_vuosikokousptk_xxx tai TEHYao151_toimkert_xxx). Tallenna siirrettävät tiedostot ulkoiseen massamuistiin. Jotta asiakirjojen etsiminen siirtomediasta sujuisi nopeasti, kansiorakenteen tulisi noudattaa yhdistyksen arkiston sarjarakennetta, samaan tapaan kuin tiedostonimi. Tästä saat tarkemmat ohjeet arkistoltamme.

Siirtotiheys ja siirron toteuttaminen

THA ottaa maksutta vastaan 20 vuotta vanhat ja sitä vanhemmat yhdistysten paperiasiakirjat. Sähköisen aineiston vastaanottosyklin on oltava huomattavasti tätä lyhyempi. Mitä lyhyempi säilytysaika yhdistyksessä on, sitä paremmin pystytään takaamaan sähköisen aineiston

muuttumattomuus ja sitä kattavampana aineisto saadaan päätearkistoon. Siirtotiheydestä sovitaan tapauskohtaisesti. Samalla voidaan sopia siirrossa käytettävästä mediasta ja siirron muista yksityiskohdista. Siirron kustannuksista (esimerkiksi siirtovälineen hankinta) vastaa luovuttava yhdistys.

TAULUKKO 2. Sähköisten asiakirjojen siirron vaiheet ja vastuut

Toiminto	Vastuuosapuoli
1. Yhteydenotto Toimihenkilöarkistoon	Yhdistys
2. Ohjeiden antaminen yhdistykselle ja yhdistyksen arkistorakenteen tarkastaminen	THA
3. Sähköisen asiakirjan ja sen metatietojen tuottaminen ohjeiden mukaisesti	Yhdistys
4. Tallentaminen ulkoiseen massamuistiin	Yhdistys
5. Aineiston toimittaminen Toimihenkilöarkistoon	Yhdistys
6. Aineiston vastaanottaminen ja tarkistaminen	THA
7. Luovutussopimuksen tekeminen siirretystä aineistosta	Yhdistys ja THA
8. Yhdistyksen arkiston rakenteen tarkistaminen ja täydentäminen tarvittaessa	THA
9. Paperiasiakirjojen liittäminen arkistoon ja arkistoyksiköiden lisääminen, jos luovutus sisältää sekä paperisia että sähköisiä asiakirjoja	THA
10. Sähköisten asiakirjojen ja niiden metatietojen liittäminen sähköiseen arkistoon	THA
11. Asiakirjojen julkaiseminen sähköisessä arkistopalvelussa siirron jälkeen tai salassapitoajan päätyttyä	THA