

# Massadigitoinnin palvelukuvaus

*Valtionhallinnon analogisten tietoaaineistojen siirtäminen Kansallisarkistoon ja muuttaminen digitaaliseen muotoon*

## 1 Yleiskuvaus palvelusta

Kansallisarkisto ottaa vastaan valtionhallinnon viranomaisten pysyvästi säilytettäviä eli arkistoitavia tietoaaineistoja ensisijaisesti digitaalisessa muodossa.

Kansallisarkisto tarjoaa viranomaisille arkistoitavan tietoaaineiston vastaanottoa, digitaaliseen muotoon muuttamista, säilytystä sekä käyttöä koskevaa massadigitoinnin palvelua. Palveluun kuuluu myös digitaaliseen muotoon muutetun analogisen tietoaaineiston hävittäminen, palvelun kehittäminen sekä digitaalisten tietoaaineistojen laadun ja saatavuuden edistäminen.

Tehtävät perustuvat lakiin Kansallisarkistosta (1145/2016) 2 § sekä arkistolaisissa (831/1994) Kansallisarkistolle säädettyihin tehtäviin.

## 2 Palvelun tarjoaminen

### 2.1 Palvelun osapuolet

Palvelun osapuolina toimivat palvelua tarjoava Kansallisarkisto sekä palvelun asiakkaina valtionhallinnon viranomaiset (jatkossa siirtäjä), joiden hallussa on pysyvästi säilytettäviä eli arkistoitavia tietoaaineistoja.

Osapuolten välille tehdään aina vakiomuotoinen sopimus palvelusta.

### 2.2 Massadigitointiin siirrettävät arkistoitavat tietoaaineistot

Massadigitoinnin tiekartassa on määritelty ylätasolla massadigitoinnin palveluun sisältyvät tietoaaineistot tiekarttaryhmittäin. Siirtäjille laaditussa massadigitoinnin siirto-oppaassa on esitetty kriteerit, joiden perusteella tiekarttaryhmät ja niiden digitointijärjestys on laadittu. Tiekarttaa päivitetään tarpeiden mukaan, ja sitä tarkennetaan vuosikohtaisilla digitointisuunnitelmillä.

Massadigitoinnin tiekartan ja digitointisuunnitelmien tavoitteena on mahdollistaa valtiosektoritasolla saavutettava kokonaishyöty eli tietoaaineistojen mahdollisimman tehokas digitointi sekä tietoaaineistojen käytettävyyden parantaminen ja muiden hyötyjen, kuten viranomaistoiminnan tehostamisen, toteutuminen. Massadigitoinnin tiekarttaan ja digitointisuunnitelmiin sisällytettyjen tietoaaineistojen soveltuvuus massadigitointiin varmistetaan viimeistään osana siirrosta sopimista massadigitoinnin palvelussa.

Massadigitoinnin palveluun siirretään arkistoitavia tietoaaineistoja, jotka on Kansallisarkiston antamissa seulontapäätöksissä määrätty arkistoitavaksi yksinomaan digitaalisessa muodossa. Massadigitoinnin siirto-oppaassa on määritelty aineistotyyppit, jotka eivät sovellu massadigitointiin. Tietyt ominaisuudet, kuten tietoaaineiston huono kunto, rajaavat osan tietoaaineistoista massadigitoinnin ulkopuolelle.

Massadigitointiin ei voida siirtää Valtioneuvoston asetuksen (1101/2019) mukaisia turvallisuusluokiteltuja tietoaineistoja. Palveluun voidaan siirtää julkista tai käyttörajoitettua tietoaineistoa. Tietoaineisto voi sisältää henkilötietoja, erityisiä henkilötietoryhmiä koskevia tietoja tai rikostuomioihin ja rikkomuksiin liittyviä henkilötietoja.

### **3 Analogisen tietoaineiston siirtäminen**

Massadigitointiin tulevan analogisen tietoaineiston siirtoprosessi muodostuu neljästä vaiheesta, joita ovat: siirrosta sopiminen, viranomaisvalmistelu, kuljettaminen ja siirron hyväksyminen.

Siirroissa noudatetaan aina voimassa olevaa Kansallisarkiston antamaa siirto-ohjeistusta.

#### **3.1 Siirrosta sopiminen**

Kansallisarkisto tekee siirtäjän kanssa sopimuksen massadigitoinnin palvelun käytöstä. Sopimus määrittää osapuolten väliset vastuut ja oikeudet tietoaineiston siirtämiseen, digitaaliseen muotoon muuttamiseen ja käyttöön liittyen. Sopimisen yhteydessä siirtäjä määrittää, onko tietoaineisto säilytys- vai arkistointivaiheessa sekä milloin säilytysvaiheessa oleva tietoaineisto siirtyy arkistointivaiheeseen.

#### **3.2 Viranomaisvalmistelu**

Siirtäjän vastuulle kuuluu tietoaineiston viranomaisvalmistelu Kansallisarkiston määrittämään siirtokuntoon. Viranomaisvalmisteluun kuuluvat välttämättömät toimet, jotka varmistavat tietoaineiston todistusvoimaisuuden ja käytettävyyden digitaalisessa muodossa. Siirtäjän muodostamien, luetteloinnin ja kuvailun yhteydessä syntyvien, metatietojen avulla mahdollistetaan tietoaineistojen käyttö digitaalisessa toimintaympäristössä käyttörajoitukset huomioon ottaen.

Viranomaisvalmisteluun kuuluvat muun muassa määräaikaisten asiakirjojen seulominen, siirrettävän tietoaineiston järjestäminen sekä luettelointi ja kuvailu. Massadigitoinnin palvelun kautta siirrettävien tietoaineistojen suojaamisessa ei tarvitse ottaa huomioon arkistokelpoisuutta.

#### **3.3 Kuljetus**

Siirtäjä vastaa analogisen tietoaineiston kuljettamisesta Siirtäjältä Kansallisarkistoon. Siirtäjä vastaa tietoaineiston kuljettamisesta todistusvoimaisuuden, säilyvyyden ja käytettävyyden sekä tietosuoja- ja tietoturva vaatimusten edellyttämällä tavalla ja sitoutuu toteuttamaan tietoaineiston kuljettamisen yhdessä Kansallisarkiston kanssa sovitun aikataulun sekä Kansallisarkiston kuljettamiselle ja käsittelylle tiloissaan määrittämien ehtojen mukaisesti. Siirtäjä vastaa myös kuljetuskustannuksista.

Kansallisarkisto järjestää Siirtäjän pyynnöstä tietoaineiston kuljetuksen Siirtäjältä Kansallisarkistoon. Kansallisarkisto laskuttaa järjestämistään 1.1.2023 lähtien tapahtuvista kuljetuksista Siirtäjältä omakustannehinnan. Laskutettavasta kuljetuksesta sovitaan Kansallisarkiston ja Siirtäjän kesken erikseen.

### 3.4 Siirron hyväksyminen

Kun analoginen tietoaineisto on kokonaisuudessaan vastaanotettu, Kansallisarkisto tarkastaa tietoaineiston. Siirtotarkastuksella varmistetaan, että tietoaineisto on Kansallisarkiston edellyttämässä kunnossa. Kansallisarkisto toimittaa hyväksymisestä siirtäjälle kirjallisen vahvistuksen (luovutustodistus).

Kansallisarkisto edellyttää, että siirtäjä korjaa tarkastamisen yhteydessä havaitut puutteet ja virheet. Jos havaitut puutteet ja virheet ovat merkittäviä ja niiden korjaaminen tulisi viemään paljon aikaa, sovitaan toimenpiteiden laajuudesta ja toteuttamistavasta siirtäjän kanssa erikseen.

Siirtäjän tulee huomioida, että vaikka siirto hyväksytään jo ennen tietoaineiston digitointia, siirtäjällä säilyy vastuu viranomaisvalmistelun virheistä ja puutteista, jotka huomataan vasta digitoinnin yhteydessä.

## 4 Digitointi

Kansallisarkistoon massadigitoinnin palvelussa vastaanotettu analoginen tietoaineisto muutetaan digitaaliseen muotoon. Digitointiprosessi muodostuu digitoinnin valmistelusta, skannauksesta, validoinnista sekä digitaalisen että analogisen tietoaineiston loppukäsittelystä. Koko digitointiprosessissa noudatetaan Kansallisarkiston vaatimuksia analogisten kappaleiden hävittämiseen tähtäävästä digitoinnista. Kansallisarkisto vastaa kokonaisuudessaan digitoinnin aikana tehtävistä toimenpiteistä, eikä niistä peritä siirtäjältä maksua.

Massadigitoinnissa skannaus tapahtuu arkistoyksiköittäin, joiden sisältämät asiakirjat digitoidaan siirtäjän luomassa sisäisessä järjestyksessä. Digitointi toteutetaan tietoaineiston sisältöä tai järjestystä muuttamatta. Skannauksessa muodostuneet kuvat (digitaaliset ilmentymät) validoidaan, jolla varmistetaan, että kuvat ovat virheettömät ja niistä on nähtävissä kaikki informaatio. Validoinnin lisäksi laadunvarmistusta tehdään massadigitoinnin palvelussa kaikissa vaiheissa erilaisin toimenpitein, mikä varmistaa digitoinnin lopputuloksen luotettavuuden.

Digitointiprosessi sisältää kuvatiedostojen tuottamisen lisäksi myös sisällöntunnistuksen, joka suoritetaan OCR-tekniikalla (*Optical Character Recognition*) ja jonka tarkoituksena on tunnistaa kuvatiedostojen sisältämät konekirjoitetut tekstit. Tekstintunnistuksen tarkkuus riippuu esimerkiksi asiakirjan sisältämän tekstin visuaalisesta laadusta, kirjaisintyypeistä sekä siitä, sisältääkö asiakirja konekirjoitettua ja käsinkirjoitettua tekstiä sekaisin. Toistaiseksi käsinkirjoitettua tekstiä ei tunnisteta osana massadigitointia.

### 4.1 Digitaalisen tietoaineiston siirto pitkäaikaissäilytykseen

Massadigitoinnin palvelussa muodostunut digitaalinen tietoaineisto metatietoineen siirretään Kansallisarkiston sähköisen arkistoinnin palvelujärjestelmään edelleen käsiteltäväksi säilyttämistä ja käyttöä varten. Siirron yhteydessä digitaaliselle tietoaineistolle tehdään teknisiä tarkastuksia, joilla varmistutaan, että tietoaineisto täyttää pitkäaikaissäilytyksen vaatimukset ja tietoaineisto noudattaa Kansallisarkiston hävittämiseen tähtäävän digitoinnin vaatimuksia. Mikäli tietoaineistossa havaitaan siirron yhteydessä virheitä, vastaa Kansallisarkisto tietoaineiston korjaamisesta ja sen uudelleensiirtämisestä.

## **4.2 Analogisen tietoaaineiston hävittäminen**

Kansallisarkisto vastaa digitoinnin jälkeen digitaaliseen muotoon muutettujen tietoaaineistojen analogisten ilmentymien hävittämisestä varoajan jälkeen.

Hävittäminen voidaan tehdä, kun digitaaliset ilmentymät ovat pitkäaikaissäilytyksessä ja niiden säilyvyys ja käytettävyys on varmistettu. Hävittämisestä informoidaan siirtäjää, mikäli tietoaaineistot ovat säilytysvaiheessa.

## **4.3 Tietojen antaminen digitoinnin aikana**

Siitä hetkestä, kun analoginen tietoaaineisto on kuljetettu Kansallisarkiston tiloihin ja ennen kuin tietoaaineisto on siirtäjälle digitaalisesti käytettävissä siihen tarkoitetun Astia-käyttöliittymän kautta, Kansallisarkisto vastaa tietopyyntöjen käsittelystä siirtäjälle. Edellä mainittuna aikana tietoaaineistoon kohdistuvat tietopyynnöt ovat siirtäjälle maksuttomia.

Tietojen antamisen vastuut ja toteuttaminen on kuvattu tarkemmin massadigitoinnin palvelua koskevan sopimuksen liitteessä.

## **4.4 Tietoaaineiston hyödyntäminen**

Kansallisarkisto voi hyödyntää massadigitoinnin palvelussa tietoaaineistoja palvelun kehittämiseen, tietoaaineistojen laadun ja säilyvyyden varmistamiseen sekä saatavuuden ja käytettävyyden edistämiseen. Kansallisarkiston tavoitteena on kehittää koneoppimiseen pohjautuvia menetelmiä, joiden avulla tietoaaineistoista voidaan automaattisesti tunnistaa esimerkiksi skannauksesta johtuvat virheet.

Kehitystyön mahdollistamiseksi Kansallisarkisto voi hyödyntää tarkoitukseen soveltuvia tietoaaineistoja koneoppimismallien kouluttamisessa. Tietoaaineistoa käsitellään tässä yhteydessä samassa tietoteknisessä ympäristössä ja samojen tietoturvaperiaatteiden mukaisesti kuin muutakin massadigitoinnissa käsiteltävää tietoaaineistoa.

# **5 Digitaalisen tietoaaineiston säilyvyyden varmistaminen ja käyttö**

## **5.1 Digitaalisen tietoaaineiston säilytys**

Kansallisarkisto vastaa massadigitoinnissa muodostuneen digitaalisen tietoaaineiston säilyttämisestä sähköisen arkistoinnin palvelussa, jonka tarkoituksena on digitaalisten tietoaaineistojen todistusvoimaisuuden ja käytettävyyden varmistaminen, tietoturvallinen säilyttäminen ja arkistointi sekä viranomaistoiminnan ja tutkimuksen tarpeiden huomiointi.

## **5.2 Digitaalisen tietoaaineiston käyttö**

Digitaalisten tietoaaineistojen käyttö tapahtuu Kansallisarkiston tarjoamassa Astia-käyttöliittymässä itsepalveluperiaatteella. Astia tarjoaa siirtäjälle tietoaaineistojen haku- ja käyttöpalveluita, joiden perusteella siirtäjä voi hakea Astiasta tarvitsemansa tietoaaineiston. Astia tukee yleisimpiä selaimia, ja sen käyttö on maksutonta.

Käyttörajoitettujen tietoaaineistojen käyttö Astiassa edellyttää siirtäjältä sähköistä tunnistautumista (Virtu) sekä henkilökohtaista käyttöoikeutta tietoaaineistoon. Siirtäjällä tulee olla käytössään Virtu-tunnistautuminen sekä virtuPersonEntitlement-attribuutin välitys.

Tarkemmin Astian käytöstä on kerrottu Astia-kuvauksessa viranomaisasiakkaille.

### 5.3 Tietoaaineiston käytön palveluaika ja vasteajat

Astia on siirtäjille jatkuvasti käytettävissä lukuun ottamatta säännöllisiä huoltokatkoja sekä erikseen ilmoitettavia huolto- ja ylläpitokatkoja.

Kansallisarkisto käsittelee tietoaaineiston käyttöön liittyvät virheilmoitukset ja palvelupyynnöt tässä kuvatussa vasteajan puitteissa. Vasteaika noudatetaan kaikkien niiden siirtäjien kanssa, joilla on sopimus massadigitoinnin palvelusta.

Palvelupyynnöiden ja vikailmoitusten vasteaika tarkoittaa aikaa, jonka kuluessa Kansallisarkiston tulee ottaa asia käsiteltäväksi. Vasteaika alkaa normaalin palveluajan sisällä silloin, kun palvelupyynnö tai vikailmoitus on vastaanotettu ja päättyy siihen hetkeen, jolloin asian ratkaisu Kansallisarkistossa aloitetaan ja palvelun tekninen tuki kuittaa palvelupyynnön tai vikailmoituksen vastaanotetuksi. Normaali palveluaika on Kansallisarkiston virka-ajan puitteissa arkipäivisin.

Vasteajan pituus riippuu palveluajan tai vikailmoituksen kiireellisyysluokituksesta (ks. taulukko 1). Palveluajan ja palveluvasteen määrittelyssä on sovellettu JHS 212 -suositusta.

#### Taulukko 1. Vasteajan kesto.

| Luokitus          | Kuvaus   | Vasteaika palveluaikana |
|-------------------|--|-------------------------|
| Laajavaikutteinen | Palvelun toimintoa tai ominaisuutta ei voi käyttää tai palvelun käyttö on muuten olennaisella tavoin häiriintynyt.                 | 1 työpäivä              |
| Kohtalainen       | Palvelun osa tai toiminto ei toimi sovitun mukaisesti ja aiheuttaa vakavaa hidastumista palvelussa.                                | 2 työpäivää             |
| Vähäinen          | Toiminnallinen häiriö tai palvelun toiminnallinen puute, jonka merkitys on vähäinen ja se ei vaikuta palvelun normaaliin käyttöön. | 5 työpäivää             |

## 6 Palvelun tietosuoja ja tietoturva

Tarvittaessa Kansallisarkisto ja siirtäjä laativat osana siirroista sopimista riskiarvion, jossa kartoitetaan ja arvioidaan tietoaaineiston analogisesta muodosta digitaaliseen muotoon muuttamiseen sisältyviä riskejä.

Palvelun ylläpidossa ja kehittämisessä huomioidaan riittävät suojatoimet, joita ovat esimerkiksi tekniset tietoturvatyökalut, kuten salaus ja lokitiedot. Hallinnollisina suojatoimina palvelussa toteutetaan muun muassa tietojen käsittelyn rajaaminen vain henkilöille, joille se tehtäviensä puolesta kuuluu, sekä huolehditaan salassapitovelvollisuuksien toteutumisesta. Kansallisarkisto huolehtii tarvittavan koulutuksen ja ohjeistuksen antamisesta tietojen käsittelyyn osallistuvalla henkilöstöllä. Massadigitoinnin palvelussa hyödynnettävistä järjestelmistä on laadittu tietosuojaa koskeva vaikutustentative. Kansallisarkisto toteuttaa säännöllisesti ulkopuolisen tahon tietoturva-auditointeja massadigitoinnissa hyödynnettäville tietojärjestelmille.

## 7 Maksulliset lisäpalvelut

### 7.1 Maksullinen kuljetus Kansallisarkistoon

Kansallisarkisto järjestää Siirtäjän pyynnöstä tietoaineiston kuljetuksen Siirtäjältä Kansallisarkistoon todistusvoimaisuuden, säilyvyyden ja käytettävyyden sekä tietosuoja- ja tietoturva-vaatimusten edellyttämällä tavalla sekä yhdessä Siirtäjän kanssa sovitun aikataulun mukaisesti. Kansallisarkisto laskuttaa järjestämistään 1.1.2023 lähtien tapahtuvista kuljetuksista Siirtäjältä omakustannehinnan. Laskutettavasta kuljetuksesta sovitaan Kansallisarkiston ja Siirtäjän kesken erikseen.

### 7.2 Aikaistettu siirto Kansallisarkistoon maksulliseen säilytykseen

Jos siirtäjä siirtää tietoaineistoa massadigitoinnin tiekarttaa ja sitä tarkentavia digitointisuunnitelmia aikaisemmin Kansallisarkistoon, laaditaan aikaistetusta siirrosta maksullista säilyttämistä ja tietojen antamista koskeva sopimus. Kansallisarkisto voi ottaa vastaan viranomaisen massadigitoitavan analogisen tietoaineiston massadigitoinnin tiekartan aikataulua aikaisemmin maksulliseen säilytykseen odottamaan digitointia. Tällöin siirtäjä maksaa tietoaineiston säilytyksestä ja siihen kohdistuvien tietopyyntöjen käsittelykustannuksista Kansallisarkiston hinnaston mukaisesti, kunnes tietoaineiston digitointi käynnistyy. Lisäksi siirtäjä vastaa kuljetuksista ja niiden kustannuksista.

### 7.3 Puutteellisesti valmistettujen analogisten tietoaineistojen korjaaminen

Jos digitoinnin yhteydessä huomataan, että viranomaisvalmistelu on toteutettu puutteellisesti, voidaan tarvittavat korjaavat toimenpiteet tehdä siirtäjän kustannuksella, jolloin Kansallisarkisto veloittaa työstä Kansallisarkiston hinnaston mukaisesti. Erikseen laskutettavasta työstä sovitaan aina erikseen siirtäjän ja Kansallisarkiston kesken.

### 7.4 Digitaalisten tietoaineistojen toimittaminen erikseen siirtäjälle

Kansallisarkisto voi siirtäjän pyynnöstä toimittaa siirtäjälle digitaalisen tietoaineiston tai sen osan fyysisellä tallennusvälineellä. Kansallisarkisto veloittaa toimitettavan digitaalisen tietoaineiston muodostamiseen käytettävästä työstä erikseen laadittavan sopimuksen perusteella.

Kansallisarkisto voi toimittaa siirtäjälle vain säilytysvaiheessa olevaa digitaalista tietoaineistoa. Arkistointivaiheessa olevaa tietoaineistoa ei toimiteta siirtäjälle erikseen.

## 8 Vastuunjako siirtäjän ja Kansallisarkiston välillä

Taulukossa kaksi on kuvattuna tiivistetysti edellä kuvatut massadigitoinnin palveluun sisältyvät vaiheet ja vastuun jakaminen siirtäjän ja Kansallisarkiston välillä.

**Taulukko 2. Vastuunjakotaulukko.**

| Tehtävä   | Siirtäjä | Kansallisarkisto |
|---|----------|------------------|
| Siirrosta sopiminen   | X        | X                |
| Siirron ohjaus  |          | X                |
| Siirrettävän tietoaineiston viranomaisvalmistelu (kuten inventointi, järjestäminen, seulonta, luettelointi ja kuvailu, nimiöinti) | X        |                  |
| Kuljetuksen suunnittelu ja toteutus   | X        |                  |
| Siirron tarkastaminen ja hyväksyntä   |          | X                |
| Digitointi  |          | X                |
| Digitaalisen tietoaineiston siirto säilytykseen   |          | X                |
| Digitaalisen tietoaineiston säilyttäminen   |          | X                |
| Tietojen antaminen analogisesta tietoaineistosta siirtäjälle digitoinnin aikana   |          | X                |
| Tietojen antaminen digitaalisesta tietoaineistosta säilyttämisen vaiheessa  | X        |                  |
| Tietojen antaminen digitaalisesta tietoaineistosta arkistoinnin vaiheessa   |          | X                |
| Digitaaliseen muotoon muutetun tietoaineiston analogisen säilytysmuodon hävittäminen  |          | X                |

## 9 Lisätietoja

Massadigitoinnin palvelusta saa lisätietoja Kansallisarkiston kotisivuilta viranomaisille suunnatusta osiosta [Massadigitoinnin palvelu](#).