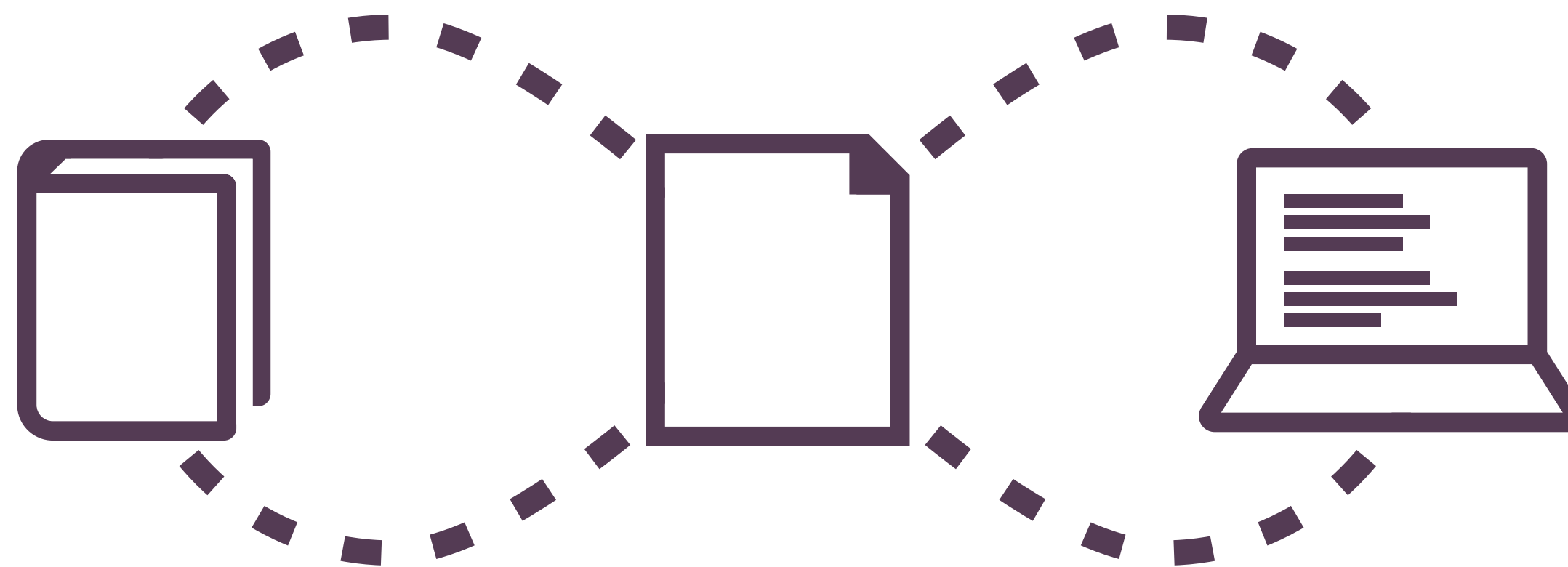




Kirsi Hänninen,  
Miia Kosonen &  
Kiti Venäläinen  
(toim.)



# DIGISTI PAREMPI

Kehitä digitointia vastuullisesti

**XAMK**

Kaakkois-Suomen  
ammattikorkeakoulu



Euroopan unionin  
osarahoittama



Julkaisu perustuu Joustavat koulutus- ja työelämäpolut tulevaisuuden digitointiosaajille (JoDi) -hankkeen Digitoinnin kehittäjä -koulutukseen, joka toteutettiin Kansallisarkiston ja Xamkin yhteistyönä 2024. JoDi-hanketta rahoitti Etelä-Savon ELY-keskus Uudistuva ja osaava Suomi 2021–2027 EU:n alue- ja rakennepolitiikan ohjelmasta (ESR+). Hankkeen toteuttivat Kaakkois-Suomen ammattikorkeakoulu Oy, Itä-Savon koulutuskuntayhtymä ja Kansallisarkisto 1.5.2023–31.5.2025.



© **Tekijät** ja Kaakkois-Suomen  
ammattikorkeakoulu

Tämä teos on kuvia lukuun ottamatta lisensoitu  
[Creative Commons Nimeä-JaaSamoin 4.0](#)  
[Kansainvälinen](#) -lisenssillä (CC BY-SA).

**XAMK KEHITTÄÄ 243**

**KAAKKOIS-SUOMEN AMMATTIKORKEAKOULU**  
**MIKKELI 2025**

**Kannen kuva:** Petri Hurme, Vinkeä Design Oy

**Graafinen suunnittelu ja kuvitus:**  
Petri Hurme, Vinkeä Design Oy

ISBN 978-952-344-604-5 (PDF)  
ISSN 2489-3102  
julkaisut@xamk.fi



# TIIVISTELMÄ

Digisti parempi -julkaisu käsittelee digitoinnin kehittämistä ja vastuullista toteuttamista. Se tarjoaa katsauksen digitoinnin prosesseihin, laadunvalvontaan, riskeihin, kestävyyteen ja tekoälypohjaisten menetelmien soveltamiseen erityisesti massadigitoinnin näkökulmasta. Julkaisun artikkelit ovat Kansallisarkiston, Kansalliskirjaston ja Kaakkois-Suomen ammattikorkeakoulu Xamkin asiantuntijoiden laatimia.

Julkaisu perustuu Joustavat koulutus- ja työelämäpolut tulevaisuuden digitointiosaajille (JoDi) -hankkeen Digitoinnin kehittäjä -koulutuksen sisältöihin. Koulutuksen toteuttivat yhteistyössä Kansallisarkisto ja Xamk.

Suomessa ei ole ollut tarjolla yhtenäistä eikä pysyvää digitointikoulutusta. Tähän tarpeeseen Xamk ja sen kumppanit ovat kehittäneet TKI-hankkeissa työelämälähtöistä digitointikoulutusta ja kouluttaneet digitoinnin osaajia ja asiantuntijoita. Koulutuksiin on osallistunut kaikkiaan 145 henkilöä vuosina 2020–2025.

Koulutuksissa havaittiin myös tarve verkostoitumiselle ja kokemusten jakamiselle. JoDi-hankkeessa rakennettiin syksyllä 2024 toimintansa aloittanut digitoinnin osaamisverkosto. Verkosto toimii suomalaisella Skillhive-yhteistyöalustalla ja on avoin kaikille digitoinnista kiinnostuneille.



# ABSTRACT

The publication addresses the development and sustainable implementation of digitization. It provides an overview of digitization processes, quality control, risks, sustainability, and the application of Artificial Intelligence (AI) based methods, particularly from the perspective of mass digitization. The articles in the publication are authored by experts from the National Archives of Finland, the National Library of Finland, and the South-Eastern Finland University of Applied Sciences (Xamk).

The publication is based on the Digitization Developer training from the Flexible Education and Career Paths for Future Digitization Experts (JoDi) project. The training was carried out in collaboration between the National

Archives of Finland and the South-Eastern Finland University of Applied Sciences (Xamk).

Previously, there was no unified or permanent digitization training available in Finland. To meet this need, Xamk and its partners developed digitization training oriented towards working life in RDI projects. The training was then used to train digitization experts and specialists. A total of 145 people participated in the training between 2020 and 2025.

The training also identified the need for networking and sharing experiences. As a result, a digitization competence network was established in the JoDi project in autumn 2024. The network operates on the Finnish Skillhive collaboration platform and is open to everyone interested in digitization.

Keywords: digitization, training, sustainability, networks, RDI projects



# ALKUSANAT

Arkistoaineistojen digitoinnin on sanottu alkaneen vuonna 1989, kun Kongressin kirjasto käynnisti Amerikan muisti -projektin Yhdysvalloissa. Digitoinnista on siis jo yli 35 vuoden kokemus. Sinä aikana siitä on kehittynyt oma erityisalueensa, jonka toimivista ja parhaista käytännöistä on nykyisin hyvä käsitys.

Maallikko voisi naivisti ajatella, että kyse on vain aineiston syöttämisestä skanneriin, mutta näin ei ole. Digitoinnin arvoiset aineistot on jollain tavalla valittava. Aineistojen erityispiirteet voivat olla haasteellisia. Laadunvalvonnasta on pidettävä huolta. Digitoinnin tulokset on tar-

jottava käyttäjille niin, että tieto on käytettävää ja löydettävää. Jotta kyse ei ole hetkellisestä hukkainvestoinnista, vaaditaan pitkäjänteistä suunnittelua ja resurssointia. Tämä kaikki ei onnistu ilman koulutusta ja osaamista.

Onnistuminen on entistä tärkeämpää, koska elämme muuttuvassa ajassa. Joskus digikopiot olivat ehkä vain vähäarvoisia käyttökopioita tärkeistä alkuperäisaineistoista. Nykyisin ne ovat yhä useammin kaikki, mitä jää jäljelle, koska alkuperäiset hävitetään. Ja vaikka näin ei olisikaan, digitoitujen aineistojen kautta tieto kuitenkin tavallisesti saavuttaa käyttäjänsä.

— Pekka Henttonen, dosentti, YTT —

28.3.2025 Tampere



# TEKIJÄT

## **Pekka Henttonen**

YTT, dosentti, Tampereen yliopisto

## **Kirsi Hänninen**

Ph.D., tiedonhallinnan suunnittelija,  
TKI-asiantuntija, Kaakkois-Suomen  
ammattikorkeakoulu

[orcid.org/0000-0003-3126-6242](https://orcid.org/0000-0003-3126-6242)

## **Tiina Hölttä**

FM, tradenomi yamk, tiiminvetäjä,  
ylitarkastaja, Kansallisarkisto

## **Anssi Jääskeläinen**

TkT, tutkimuspäällikkö, Kaakkois-  
Suomen ammattikorkeakoulu

## **Ville Kajanne**

FT, yksikönpäällikkö, Kansallisarkisto

## **Sami Kangasniemi**

Tradenomi, suunnittelija, Kansallisarkisto

## **Miia Kosonen**

KTT, tutkimuspäällikkö, Kaakkois-Suomen  
ammattikorkeakoulu

[orcid.org/0000-0002-7367-4810](https://orcid.org/0000-0002-7367-4810)

## **Johanna Lilja**

FT, palvelujohtaja, Kansalliskirjasto,  
Kokoelma- ja datapalvelut

[orcid.org/0000-0002-3905-0035](https://orcid.org/0000-0002-3905-0035)

## **Tuomo Räisänen**

FT, ohjelmistosuunnittelija, Kaakkois-  
Suomen ammattikorkeakoulu

## **Noora Talsi**

YTT, tutkimusjohtaja, Kaakkois-  
Suomen ammattikorkeakoulu

## **Anna-Elina Ukkonen**

YTM, projektipäällikkö, Kansalliskirjasto,  
Kokoelma- ja datapalvelut

[orcid.org/0000-0003-3666-6376](https://orcid.org/0000-0003-3666-6376)



# SISÄLLYS

Tiivistelmä.....	3
Abstract.....	4
Alkusanat.....	5
Tekijät.....	6
Digitointikoulutusten ja -osaamisen kehittämishankkeet.....	8
Digitoinnin osaamisverkosto jatkuvan oppimisen palveluksessa.....	15
Laadun käsite massadigitoinnin prosessissa.....	22
Miten aineiston elinkaari näkyy digitoijan työssä.....	30
Tekoälyllä tehokkuutta digitointiprosessiin.....	36
Digitoinnin riskit.....	42
Kestävyyšnäkökulma Kansalliskirjaston digitointiohjelmassa.....	54
Loppusanat.....	66





Kirsi Hänninen

# DIGITOINTI- KOULUTUSTEN JA -OSAAMISEN KEHITTÄMISHANKKEET

**D**igitointi on laaja ja monipuolinen ilmiö, jonka ymmärtäminen ja jonka parissa työskentely edellyttävät tietoa ja taitoa. Sana tarkasti ottaen digitointi tarkoittaa analogisen tiedon muuttamista sähköiseksi erilaisilla tekniikoilla ja välineillä, mutta prosessina ja projektina digitointi on paljon muutakin, eikä sitä toteuteta muusta ympäristöstä erillään.

Digitointiin vaikuttavat niin organisaation, ympäröivän yhteisön kuin ympäristönkin arvot, tarpeet, mahdollisuudet ja rajat. Digitointiin liittyvät työtehtävät ovat yhtä moninaisia kuin ilmiö itsekin: yhtäällä tarvitaan esimerkiksi laitteiden ja ohjelmistojen osaajia, toisaalla lain-säädännön ja projektinhallinnan taitajia.

Suomessa ei ole tarjolla pysyvää eikä yhtenäistä digitointikoulutusta, ja digitointitehtävissä työskennellään erilaisilla koulutustaustoilla. Muistiorganisaatiot, viranomaiset, yhdistykset ja yritykset ohjaavat ja kouluttavat omat työntekijänsä digitointityöhön. Osaaminen karttuu, mutta samalla myös siiloutuu. Yhdessä työpaikassa opitut standardit, ohjeet, laitteet ja ohjelmistot voivatkin olla soveltumattomia seuraavassa digitointityössä. Toisaalta digitointijien ja digitointihankkeiden toteuttajien liikkuaessa työmarkkinoilla he vievät osaamistaan uusiin organisaatioihin.



Kun virallista digitointikoulutusta ei ole, kuinka yksittäiset ihmiset ja organisaatiot saavat digitointikoulutusta ja kokemusta? Kaakkois-Suomen ammattikorkeakoulu (Xamk) ja sen kumppanit Mikkelin kehitysyhtiö Miksei Oy, Kansallisarkisto ja Itä-Savon koulutuskuntayhtymä Samiedu ovat vastanneet tähän kysymykseen suunnitteleamalla ja toteuttamalla työelämälähtöistä digitointikoulutusta TKI-hankkeissa. Xamkissa hankkeet ovat sijoittuneet Digitaliaan, joka on Kaakkois-Suomen ammattikorkeakoulun, Kansalliskirjaston ja Helsingin yliopiston yhteinen tutkimusyksikkö. Eri hankekumppanikoonpanoilla toteutetut hankkeet ajoittuvat vuosille 2018–2025, ja yhteensä niissä koulutettiin 145 digitoinnin osaajaa, ammattilaista, taitajaa, toteuttajaa ja kehittäjää. Hankkeet

on rahoitettu sekä ulkopuolisella että toteuttajien omalla rahoituksella.

Digitointia on opetettu Mikkelissä aiemminkin. Mikkelin ammattikorkeakoulu (Mamk) järjesti pilotoivaa digitointikoulutusta jo vuosina 2004–2005 ja uudelleen vuosina 2009–2010. Mamkilla oli muitakin väyliä digitoinnin opettamiseen. Oppilaitos tarjosi sähköisen aineistohallinnan ja arkistoinnin erikoistumisopintoja vuosina 2005–2007 ja 2009. Vuonna 2008 Mamk aloitti YAMK-tutkintoon johtavan sähköisen asioinnin ja arkistoinnin koulutuksen, nykyisen tiedonhallinta ja sähköinen arkistointi -koulutuksen. Osaan näistäkin koulutuksista on kuulunut digitoinnin koulutusta. (Saltiola ym. 2019, 24–27.)

## Massadigitoinnin osaajakoulutuksen suunnitteluhanke

Käsittelimieni koulutushankkeiden ketjun ensimmäinen etappi oli *Massadigitoinnin osaajakoulutuksen suunnitteluhanke* (1.11.2018–31.8.2019). Suunnitteluhanketta hallinnoi Mikkelin kehitysyhtiö Miksei Oy, osatoteuttajana oli Kaakkois-Suomen ammattikorkeakoulu ja hanketta rahoitti Etelä-Savon elinkeino-, liikenne- ja ympäristökeskus ESR-rahoituksella.

Taustalla vaikutti valtioneuvoston vuonna 2017 tekemä periaatepäätös siitä, että viranomaisten hallussa olevat paperiaineistot tulee digitoida vuoteen 2030 mennessä. Päätöksen myötä Kansallisarkisto ryhtyi suunnittelemaan ja pilotoimaan massadigitointia. *Massadigitoinnin osaajakoulutuksen suunnitteluhankkeessa* varauduttiin siihen, että Kansallisarkiston massadigitointikeskus sijoi-

tettaisiin Mikkeliin. Hankkeessa kartoitettiin digitoijien ja digitoinnin asiantuntijoiden osaamistausta- ja koulutusvaatimuksia, työllistymismahdollisuuksia, koulutusvaihtoehtoja sekä Mikkelin seudun organisaatioiden työntekijä- ja koulutustarpeita ja mahdollisuuksia olla mukana koulutuksessa.

Nostan hankkeen tuloksista esimerkiksi mielenkiintoisen havainnon: Hankkeessa kysyttiin Kansallisarkistolta, Kansalliskirjastolta, Suomen Elinkeinoelämän Keskusarkistolta (ELKA) ja Disec Oy:ltä, millaisia tietoja ja taitoja digitoijalla tulisi olla. Vastaus oli, että digitoijien pitäisi hallita tietotekniikan perustaidot, heillä tulisi olla hyvät suomen ja englannin kielen taidot, heidän tulisi olla oma-aloitteisia, itseohjautuvia, yhteistyökykyisiä ja huolellisia sekä pärjätä välillä yksitoikkoisessa ja fyysisestikin raskaassa



työssä. Koulutuksen kohderyhmiä puolestaan arvioitiin yhdessä TE-toimiston kanssa. Vahvimpana työnhakijaryhmänä TE-toimisto piti työttömiä datanomeja, merkonomeja ja toimistotyöntekijöitä. (Saltiola ym. 2019, 12–16.) Siinä missä digitoijan työelämätaito-vaatimukset ovat yhä tänäänkin tunnistettavissa, niin digitoititöihin ja -koulutuksiin hakeutuvien ammattitaustat vaihtelevat.

Seuraavaksi hankkeessa selvitettiin yhdessä Kansallisarkiston ja Kansalliskirjaston kanssa, millaista koulutusta digitoijilla ja digitoinnin asiantuntijoilla tulisi olla. Digitoijilla painottuivat digitointiprosessin hallinta ja laadunvarmistuksen osaaminen molemmilla organisaatioilla. Asiantuntijoilta odotettiin lisäksi lainsäädännön, tietoturvan sekä järjes-

telmien ja ohjelmistojen osaamista. Digitoitinkoulutuksesta toivottiin sekä KAM-sektorille eli arkistoille, kirjastoille ja museoille että yrityksille sopivaa yleistä digitoitinkoulutusta. (Saltiola ym. 2019, 33–36.)

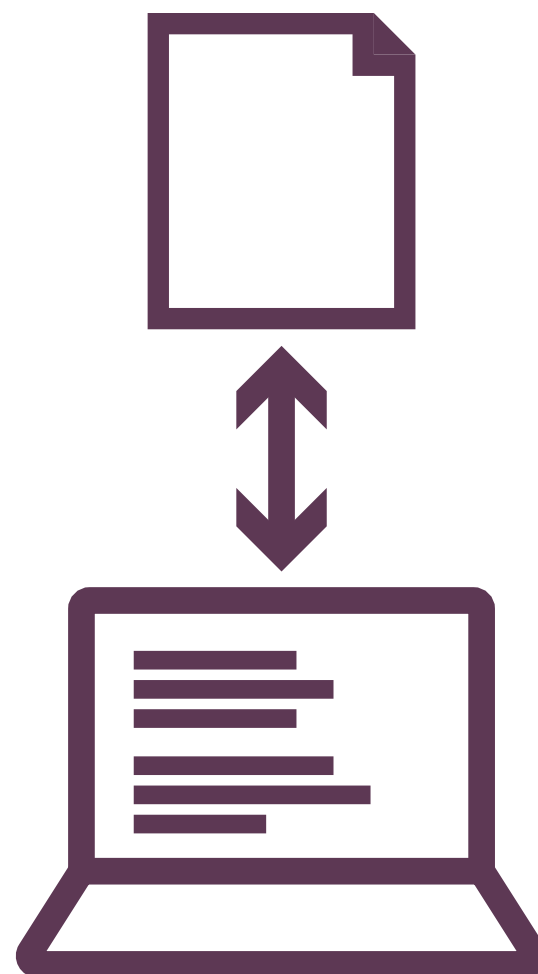
Kartoituksen tuloksien pohjalta hankkeessa luotiin koulutussuunnitelmat sekä digitoijille että digitoinnin asiantuntijoille. Digitoinnin ammattilaiseksi -koulutus rakennettiin sel-laiseksi, että opiskelija pystyisi koulutuksessa saamallaan osaamisella työllistymään digitointitehtäviin. Digitoinnin asiantuntijakoulutus puolestaan suunniteltiin täydennyskoulutukseksi jo alalla toimiville arkistoalan ammattilaisille. Koulutus antaisi opiskelijoille tietoja ja taitoja toimia digitoinnin asiantuntija- ja ohjaustehtävissä. (Saltiola ym. 2019, 37–42.)

## Digitoinnin ammattilaiseksi

Päätös Kansallisarkiston massadigitoinnin toimintojen sijoittumisesta Mikkeliin tuli keväällä 2020. Kun toiminta käynnistyi keväällä 2021, päästiin toteuttamaan aiemmin tehtyjä koulutussuunnitelmia. *Digitoinnin ammattilaiseksi* -hankkeessa (1.9.2020–31.12.2021) järjestettiin kaksi Digitoinnin ammattilaiseksi (digitoija) -koulutusta ja yksi Digitoinnin ammattilaiseksi (digitoinnin asiantuntija) -koulutus. Digitoija-koulutuksista valmistui 30 henkilöä ja asiantuntijakoulutuksesta kahdeksan. Hanketta hallinnoi Kaakkois-Suomen ammattikorkeakoulu ja osatoteuttajana toimi Mikkelin kehitysyhtiö Miksei Oy. Hanketta rahoitti Etelä-Savon elinkeino-, liikenne- ja ympäristökeskus ESR-rahoituksella. (Digitoinnin ammattilaiseksi -hankkeen loppuraportti 2022.)

Digitoinnin ammattilaiseksi (digitoija) -koulutukset ajoitettiin Kansallisarkiston massadigitoinnin tuotantoympäristön perustamisen ja tuotannon käynnistämisvaiheiden mukaisesti keväälle ja syksylle 2021. Koulutuksen laajuus oli 15 opintopistettä, ja se koostui kolmesta opintojaksosta: Digitoinnin perusosaaminen, Digitoinnin syventävä osaaminen ja Digitointiprojektin toteutus.

Lisäksi opiskelijoilla oli mahdollisuus ottaa täydentäviä opintoja Avoimen AMK:n koulustarjonnasta. Digitoinnin ammattilaiseksi (asiantuntija) -koulutus oli täydennyskoulutusta alalla jo toimiville arkistoalan ammattilaisille. Tämänkin koulutuksen laajuus oli 15 opintopistettä. Koulutukseen kuului kolme opintojaksoa, joissa perehdyttiin digitoinnin laadunvarmistukseen ja aineiston käytettä-



vyytteen, prosessien kuvaamiseen ja kehittämiseen sekä sähköiseen arkistointiin ja pitkäaikaissäilyttämiseen. Kaksi opintojaksoa integroitiin Tiedonhallinnan ja sähköisen arkistoinnin yamk-opintoihin. (Digitoinnin ammatilaiseksi -hankkeen loppuraportti 2022.)

Koulutukset toteutettiin verkko- ja monimuoto-opintoina. Molempien koulutusten opiskelijoilla oli käytettävissään Xamkin digitoinnin harjoitteluympäristö, jonne hanke toimitti skannereita, kameroita, objektiivaja ja ohjelmiston elinkeinoelämän suositukset huomioiden. Lisäksi tarjolla oli kilpailutetut työssäoppimispaikat, joissa opiskelijat suorittivat 2–4 opintopisteen laajuisen työssäoppimisen koulutuskokonaisuuden. Hankkeen päätyttyä koulutusmateriaalit avattiin Avointen oppimateriaalien kirjastossa. (Digitoinnin ammatilaiseksi -hankkeen loppuraportti 2022.)

## Digitoinnin taitajaksi

*Digitoinnin taitajaksi* -hanke (1.1.–31.12.2022) luotiin edeltäneissä hankkeissa nousseiden työelämätarpeiden pohjalta. Hanke toteutti täydennyskoulutustyyppisen koulutuksen, joka keskittyi digitointiprojektin ja digitointityön suunnitteluun ja valmisteluun sekä lisäsi opiskelijoiden digikyvykkyyttä. Koulutus oli suunnattu arkistoalan, asiakirjahallinnan ja tiedonhallinnan työntekijöille sekä muille digitointityöstä kiinnostuneille. Hanketta hallinnoi Kaakkois-Suomen ammattikorkeakoulu ja osatoteuttajana toimi Mikkelin kehitysyhtiö Miksei Oy. Hanketta rahoitti Etelä-Savon elinkeino-, liikenne- ja ympäristökeskus REACT EU ESR -rahoituksella. Vuonna 2022 Memory Campuksen koordinointi siirtyi Mikseiltä Xamkille, ja siirtymävaiheeseen osunut hanke tuki ja vahvisti yhteistyötä. Memory

Campuksen toimijat tarjosivat case-esimerkkejä ja osallistuivat koulutuksen ja oppaan tuottamiseen. (Digitoinnin taitajaksi -hankkeen loppuraportti 2023.)

Hanke järjesti vuoden mittaan kaksi koulutusta, ja koulutuksista valmistui 33 opiskelijaa. Koulutukset olivat kymmenen opintopisteen laajuisia. Sisällöiltään koulutukset kattoivat johdannon digitointiin ja digitaitoihin, digitointiprojektin suunnittelun ja hallinnan, digitoinnin määräykset ja ohjeet, digitoinnin laadunvarmistuksen, kokoelmanhallinnan ja sähköisen arkistoinnin. Koulutus toteutettiin verkko-opintoina. Teoriaopetusta seurasi ryhmätöitä, ja opiskelijat tekivät omaa projektisuunnitelmaansa koko koulutuksen ajan. Lähtökohtana oli tehdä projektisuunnitelma omalle työpaikalle. (Digitoinnin taitajaksi -hankkeen loppuraportti 2023.)



Hanke julkaisi *Digiksi – digitointiprojektin suunnittelijan ja toteuttajan oppaan*, joka on suunnattu kaikille digitointiprojektia suunnitteleville ja valmisteleville. Oppaan tekstit käsittelevät erityisesti julkisen sektorin digitointia, mutta julkaisu soveltuu myös yksityisen ja kolmannen sektorin käyttöön. Opas koostuu digitointiprojektin suunnittelun ja hallinnan perustiedoista, näkökulma-artikkeleista ja erilaisten digitointiprojektien case-esimerkeistä. Hankkeen toteuttama koulutus mallinnettiin mahdollista jatkokäyttöä varten, ja koulutusmateriaalit (digitoinnin taitajien opintojaksomateriaali, koulutusmalli, oheislukemisto) avattiin Avointen oppimateriaalien kirjastossa. (Digitoinnin taitajaksi -hankkeen loppuraportti 2023.)

Digitoinnin koulutushankkeet saivat vuonna 2022 myös kansallista tunnustusta. Arkistojen

päivän toimikunta myönsi Digitalian ja Mikkelin kehitysyhtiö Miksei Oy:n yhdessä toteuttamille digitoinnin koulutushankkeille Vuoden arkistoteko 2022 -kunniamaininnan. Perusteluissa todettiin, että koulutukset ovat lisänneet merkittävällä tavalla arkistoinnin ja asiakirjahallinnan digitalisaatiossa tarvittavaa tietämystä ja ammattitaitoa. (Arkistoyhdistys 2022.)

## Digitoinnin osaamiskartoitus

*Digitoinnin osaamiskartoitus* -hankkeessa (1.1.–30.4.2023) selvitettiin, millaisia osaamistarpeita digitoinnin toimialalla oli kartoitushetkellä ja millaista osaamista ennakoidaan tarvittavan tulevaisuudessa. Hankkeen toteutti Kaakkois-Suomen ammattikorkeakoulu, ja hanketta rahoitti Etelä-Savon

maakuntaliitto Alueiden kestävän kasvun ja elinvoiman määrärahasta (AKKE). Kartoitus toteutettiin Etelä-Savossa, ja siihen osallistui myös organisaatioita, jotka toimivat kansallisella tasolla.

Osaamiskartoituksen kohderyhmäksi valittiin Memory Campus -yhteisön jäseniä, Digitalian yhteistyökumppaneita sekä edellisten digitoinnin koulutushankkeiden aikana kertyneitä kontakteja. Kutsu kartoitukseen lähetettiin 54 eri organisaatiolle, ja kutsuun vastasi ja kartoitukseen osallistui 12 organisaatiota eri aloilta.

Kartoitus toteutettiin haastatteluilla. Kartoituksen menetelmänä käytettiin kahta sähköisen arkistoinnin ja pitkäaikaissäilytyksen osaamisen selvittämiseen suunniteltua viitekehystä: Digital Preservation Coalition Rapid

Assessment Model (DPC RAM) ja Digital Preservation Competency Framework (DPC CAT). Viitekehysten sisällä on osaamisalueita ja kyvykkyyksiä, kuten viestintä ja edunvalvonta, lainsäädäntö ja vastuullisuus, aineistojen vastaanotto ja siirto sekä metatiedon hallinta. Haastatteluissa käytetty kyselylomake perustui näihin viitekehyksiin. (Järn 2023, 4–5.)

Kartoitus osoitti, että mukana olleissa organisaatioissa oli digitoinnin osaamista, mutta myös tarvetta lisätä ja päivittää osaamista. Vastaajat pitivät tärkeinä digitointiosaamisen alueina aineiston pitkäaikaissäilytyksen vaatimusten hallintaa sekä tietoja ja taitoja hallinnosta ja resursoinnista, viestinnästä, edunvalvonnasta, lainsäädännöstä ja vastuullisuudesta.



Informaatioteknologiaan liittyvissä näkemyksissä oli enemmän hajontaa, sillä digitoinnin ulkoistaneet vastaajat pitivät aluetta epärelevanttina. Toisaalta organisaatioissa, joissa digitointi on ulkoistettu, korostui hankintojen osaaminen.

Digitoinnin teknologian ymmärtämisen tarve oli kaikille yhteistä. Tätä ymmärrystä tarvitaan erityisesti palvelujen, laitteistojen ja järjestelmien hankinnoissa. Moni kyselyyn vastaaja toivoi digitointiin liittyvän verkostoyhteistyön lisäämistä. Tämän lisäksi vastaajat toivoivat räätälöityjä koulutuksia, jotka vastaisivat työelämän tarpeisiin. (Järn 2023, 6–9.)

Kartoituksen jälkeen hanke laati suositukset digitoinnin koulutustarjonnasta ja osaamissisällöistä, työelämälähtöisen digitoinnin

osaamisen ennakoitavimallista, digitointi-osaamisen kuvaamisesta sekä digitointiosaamisen tunnistamisesta ja tunnustamisesta. (Järn 2023, 9–10.)

### Joustavat koulutus- ja työelämäpolut tulevaisuuden digitointiosaajille

Digitointikoulutusten viimeisin hanke, *Joustavat koulutus- ja työelämäpolut tulevaisuuden digitointiosaajille* (JoDi, 1.5.2023–31.5.2025), laajensi koulutustarjonnan myös ammatilliseen koulutukseen ja loi edellisen hankkeen tuloksissa toivotun digitointiverkoston. Hanketta hallinnoi Kaakkois-Suomen ammattikorkeakoulu, ja osatoteuttajina olivat Itä-Savon koulutuskuntayhtymä (Samiedu) ja Kansallis-

arkisto. Hanketta rahoitti Etelä-Savon elinkeino-, liikenne- ja ympäristökeskus ESR+ -rahoituksella.

Hankkeessa toteutettiin työelämälähtöiset koulutukset kahdelle koulutusasteelle: ylemmän ammattikorkeakoulututkinnon tasoinen, viiden opintopisteen laajuinen Digitoinnin kehittäjä -koulutus Xamkin ja Kansallisarkiston yhteistyönä sekä toisen asteen perustutkinnon tasoinen, 30 osaamispisteen laajuinen Digitoinnin toteuttaja -koulutus Samiedun ja Kansallisarkiston yhteistyönä.

Digitoinnin kehittäjä -koulutus järjestettiin kahdesti, keväällä ja syksyllä 2024. Koulutuksen kohderyhmä oli tiedonhallinnan, asiakirjahallinnan ja arkistoalan asiantuntijat, mutta koulutus oli avoin myös muille digi-

toinnin kehittämistä kiinnostuneille. Koulutukseen kuului itsenäistä työskentelyä sekä kontaktiopetusviikonloppuja verkossa. Asiantuntijaluennot, ryhmäkeskustelut ja tehtävät käsittelivät aineiston elinkaarta, digitointi-projektin hallintaa, digitoinnin sääntelyä, ohjeistusta, laatua ja riskejä, digitoinnin vastuullisuutta ja ympäristövaikutuksia, tekoälyn vaikutuksia digitointiin sekä yhteistyöverkostoja. Koulutuksen lopputehtävänä opiskelijat laativat digitointi-projektin suunnitelman. Koulutuksen suoritti 38 opiskelijaa.

Digitoinnin toteuttaja -koulutus oli suunnattu digitointialalla jo työskenteleville sekä sinne haluaville. Koulutuksen aikana opiskelija suoritti Samiedun media-alan ja kuvallisen ilmaisun perustutkinnon valinnaisen osan nimeltään Tallennetuotannossa työskentely.



Käytännönläheisessä koulutuksessa opiskelijat tutustuivat digitointiin, sen merkitykseen ja käyttökohteisiin, muistiorganisaatioihin, digitoinnin vaiheisiin, digitoinnin tulosten laadunvarmistukseen ja -valvontaan sekä digitointityötä säätelevään lainsäädäntöön, standardeihin ja suosituksiin. Opiskelijat opetettiin myös käyttämään digitoinnissa tarvittavia laitteita ja ohjelmistoja. Koulutus toteutettiin monimuoto-opintoina, joihin kuului itsenäisiä verkko-opintoja, lähiopetuspäiviä sekä työelämässä oppimisen jakso eli työharjoittelu. Koulutus oli tarkoitus hankesuunnitelman mukaisesti järjestää kaksi kertaa, mutta sen suosio oli niin suuri, että koulutus toteutettiin kolmannenkin kerran. Koulutuksen suoritti 36 opiskelijaa.

Hankkeen kolmas saavutus oli digitoinnin osaamisverkoston rakentaminen. Kyseessä on foorumi, joka tukee digitoinnin jatkuvaa oppimista ja ongelmanratkaisua ja joka on tarkoitettu digitoinnin asiantuntijoille, kaikille aiheesta kiinnostuneille ja digitointipalveluiden tarjoajille. Verkosto lanseerattiin syyskuussa 2024, ja se toimii suomalaisella Skillhive-alustalla. Miia Kosonen käsittelee artikkelissaan tässä julkaisussa laajemmin verkoston syntyä ja toimintaa.

Hankkeessa myös julkaistiin tämä nyt käsillä oleva *Digisti parempi – Kehitä digitointia vastuullisesti* -julkaisu. Hankkeen tulokset jatkavat elämäänsä hankkeen päätyttyäkin, sillä Digitoinnin kehittäjä -koulutukseen perustuva opintojakso tulee Xamkin avoimen amk:n tarjontaan syksyllä 2025.

## Lopuksi

TKI-hanketoiminnalla on pystytty vastaamaan nopeasti ja kohdennetusti digitointiosaamisen kehittämisen tarpeisiin suunnittelemalla ja toteuttamalla koulutuksia. Työelämän koulutustarpeista kertoo esimerkiksi se, että Digitoinnin kehittäjä -koulutuksen paikat täyttyivät 2–3 minuutissa ilmoittautumisen avautumisesta ja hankkeiden kaikkiin koulutuksiin on ollut enemmän hakijoita kuin paikkoja (Järn 2023, 2). Koulutetut ovat olleet ympäri Suomea ja eri sektoreilta, ja he ovat voineet vertaisoppia ja jakaa kokemuksiaan toistensa kanssa koulutuksissa sekä viedä uutta osaamistaan organisaatioihinsa. Tämän turvaaminen on jatkossakin tarpeen. Laadukkaan koulutuksen tarjoaminen turvaa vastuullista kulttuuriperinnön digitointia, säilytystä ja jakamista nyt ja tulevaisuudessa.

## Lähteet

Arkistoyhdistys. 2022. Vuoden arkistoteko 2022 -palkinto Romano Missiolle. WWW-dokumentti. Saatavissa: <https://arkistoyhdistys.fi/index.php/2022/11/11/vuoden-arkistoteko-2022-palkinto-romano-missiolle/> [viitattu 4.4.2025].

Digitoinnin ammattilaiseksi -hankkeen loppuraportti. 2022. Kaakkois-Suomen ammattikorkeakoulu.

Digitoinnin taitajaksi -hankkeen loppuraportti. 2023. Kaakkois-Suomen ammattikorkeakoulu.

Järn, S. 2023. Digitoinnin osaamiskartoitus. Kaakkois-Suomen ammattikorkeakoulu. PDF-dokumentti. Saatavissa: <https://www.xamk.fi/app/uploads/sites/2/2024/08/digitoinnin-osaamiskartoitus.pdf> [viitattu 4.4.2025].

Saltiola, K., Ristolainen H. & Jurvanen K. 2019. Digitointikoulutuksen suunnitteluraportti. Mikkelin Kehitysyhtiö Miksei Oy. Kaakkois-Suomen ammattikorkeakoulu.



Miia Kosonen

# DIGITOINNIN OSAAMISVERKOSTO JATKUVAN OPPIMISEN PALVELUKSESSA

Digitoinnin osaamisverkosto on syyskuussa 2024 avattu maksuton ja Suomen mitakaavassa ainutlaatuinen foorumi kaikille digitoinnin asiantuntijoille, aiheesta kiinnostuneille ja digitointipalveluiden tarjoajille. Verkosto tukee digitoinnin jatkuvaa oppimista ja ongelmanratkaisua yli organisaatio- ja sektorirajojen. Sen tavoitteena on tarjota tietoa, yhteyksiä ja vertaistukea. Digitoinnin osaamisverkosto toimii suomalaisella Skillhive-yhteisöalustalla osoitteessa [digitointi.skillhive.com](https://digitointi.skillhive.com).

Osaamisverkosto rakennettiin osana *Joustavat koulutus- ja työelämäpolut tulevaisuuden digitointiosaajille (JoDi)* -hanketta. Hanketta on toteutettu ajalla 1.5.2023–31.5.2025 ja sitä

rahoitti Etelä-Savon ELY-keskus Euroopan sosiaalirahastosta (ESR+).

TKI-hankkeiden perustehtävä on tarjota toimijoille mahdollisuus kokeilla ja kehittää uutta. JoDi-hankkeessa Xamk on tarttunut haasteeseen luomalla perustan digitointiin keskittyvälle asiantuntijaverkostolle, jollainen Suomesta on aiemmin puuttunut. Kaikilla yhteisöillä ja verkostoilla on neljä vaihetta: esivaihe, varhainen vaihe, kypsyysvaihe ja luopuminen ("kuolemavaihe"). Tässä artikkelissa kuvaan digitoinnin osaamisverkoston kahta ensimmäistä vaihetta, esivaihetta suunnitteluineen sekä varhaisia askeleita toimivana verkostona.



## Käsitteistä

Yhteiskuntatieteissä verkosto on sosiaalinen rakenne, joka muodostuu toimijoista – esimerkiksi yksilöistä tai organisaatioista – ja näiden välisistä sidoksista, jotka voivat olla vahvoja tai heikkoja. Verkkoyhteisöt taas ovat sosiaalisia verkostoja, jotka hyödyntävät viestintäteknologiaa pitääkseen yhteyttä muihin jäseniin liittyen tiettyyn yhteiseen aihepiiriin tai intressiin (Wasko & Faraj 2000) ja joissa jäsenet kokevat yhteenkuuluvuutta juuri kyseiseen yhteisöön (Blanchard & Markus 2004). Toisin sanoen kaikkien yhteisöjen ytimessä on sosiaalinen verkosto, mutta jokainen verkosto ei ole automaattisesti yhteisö.

Osaaminen-käsitteen sisältö on vahvasti kulttuurisidonnaista. Suomessa osaaminen on vakiintunut käännökseksi englannin kielen termeille *competence* ja *competency*, joista ensin mainittu painottaa oppimisen tuloksia

ja jälkimmäinen prosessia. Osaaminen voidaan määritellä yhdistelmäksi tietoa, taitoa ja asenteita. Monet määritelmät lisäävät listaan myös kokemukset, joiden avulla tietoja ja taitoja käytännössä kerrytetään.

Jyväskylän ammattikorkeakoulussa kehitetty *Osaamiskäsi* (kuva 1) huomioi kokemusten lisäksi verkostot ja osallisuuden erilaisissa yhteisöissä: emme opi sosiaalisessa tyhjiössä, vaan vuorovaikutuksessa muiden kanssa, samalla aiemmalle tiedolle rakentaen.

Myös digitoinnin osaamisverkostossa osamista lähestytään laaja-alaisesti: sillä tarkoitetaan yksilöiden tietotaitoa ja kokemusta digitoinnista, alan jaettua hiljaista tietoa, kuten arvoja, asenteita ja yhteisiä toimintatapoja, sekä jaettua näkyvää (eksplisiittistä) tietoa, kuten digitointiin liittyviä ohjeita ja



Kuva 1. Osaamiskäsi. Kuva: Jyväskylän ammattikorkeakoulu (CC BY) <https://oppimateriaalit.jamk.fi/fi/portfolio/portfolio-jamkissa/osaamisen-tunnistaminen/>



## Käytäntöyhteisöt digiajassa

standardeja. Tämän alakohtaisen erityisosaamisen lisäksi tarvitaan yleistä osaamista, kuten taitoa hyödyntää digitaalisia foorumeja tiedon jakamiseen.

Entä miten määritellään jatkuva oppiminen? Sillä tarkoitetaan ”koko elämänkaaren aikaista, eri elämänaalueille ulottuvaa oppimista” (Työ- ja elinkeinoministeriö). Tavoitteena jatkuvan oppimisen edistämiseksi on vastata työelämän muutoksista aiheutuviin osaamistarpeisiin erityisesti työikäisen väestön keskuudessa. Digitoinnin osaamisverkostossa kohtaavat kasvava digitointiosaajien tarve ja työikäisten tarve kouluttautua uudelleen tai hankkia täydentävää oppia oman osaamisen kasvattamiseksi.

Digitoinnin osaamisverkosto on osa laajempaa käytäntöyhteisöjen verkkoa, jota yhdistää pyrkimys luotettavan tiedon säilyttämiseen ja hallintaan. Jäsenet edustavat esimerkiksi kirjasto-, arkisto- ja museoaloja. Lave ja Wenger (1991) kuvailevat käytäntöyhteisöä toimintajärjestelmänä, jossa osallistujilla on yhteinen käsitys toiminnastaan ja sen merkityksestä heille ja yhteisölleen. Ihmiset jakavat tietyn aihealueen, haasteen tai intohimon kohteen ja pyrkivät jatkuvasti syventämään tietotaitoaan.

Virtuaaliset tai digitaaliset käytäntöyhteisöt laajentavat olemassa olevia sosiaalisia käytäntöjä verkkoon. Ne tarjoavat jäsenilleen erottuvan rakenteen keskinäiseen vuorovaikutukseen ja tiedon jakamiseen. Verkon

käytäntöyhteisöt eivät ole erillisiä sosiaalisia todellisuuksia, vaan usein jatkoa jo olemassa oleville käytäntöyhteisöille ja niissä tapahtuvalle kanssakäymiselle. Ensisijainen syy verkkofoorumien olemassaololle on yhteisen tiedon jakaminen ja syventäminen. Alkuvaiheessa ryhmä on rakenteeltaan hyvin löyhän ja hajanaisen verkoston kaltainen, mutta aikaa myöten se voi saada yhteisöllisempiä piirteitä. (Kosonen 2008.)

Virtuaalisissa käytäntöyhteisöissä tiedon jakamista voivat helpottaa useat tekijät (Kosonen 2008):

- ◆ yksilöllinen motivaatio: halua auttaa, mahdollisuus oppia, tunnustus, maineen kasvu

- ◆ sosiaaliset tekijät: luottamus, vastavuoroisuuden normit, jaettu ammattikieli, yhteinen kulttuuri
- ◆ teknologian ominaisuudet: sovellusten laatu, käytettävyys, viestinnän helppous, joustavuus, yksityisyys, tietojen hallinta ja suojaaminen.

2010-luvulla alettiin puhua yhteisömanagerista ammattikuntana. Yhteisömanageri vastaa yhteisön rakentamisesta, ylläpidosta ja jäsenten palvelemisesta yhteisön tavoitteita vastaavalla tavalla. Fokus on jäsenten välisten suhteiden ja oman aktiivisuuden ravitsemisessa, erotuksena esimerkiksi sosiaalisen median managereihin, joiden ensisijainen tehtävä on tuottaa yleisölle – usein tietyn brändin tai yrityksen asiakkaille – kiinnostavaa ja sitouttavaa sisältöä.



Seuraavat yhdeksän suunnittelustrategiaa luonnehtivat Kimin (2000) mukaan menestyviä ja pysyviä verkkoyhteisöjä:

1. tarkoituksen määrittely
2. kokoontumispaikkojen rakentaminen
3. jäsenprofiilien luominen
4. roolien suunnittelu
5. johtajuuskäytäntöjen kehittäminen
6. etikettisääntöjen laatiminen
7. tapahtuminen järjestäminen
8. rituaalien integrointi yhteisöön
9. jäsenten vetämien alaryhmien (subgroup) muodostaminen

Kuten tästä listasta havaitaan, yhteisömanagerin rooli korostuu alkuvaiheessa, kun jäsenten omia alaryhmiä ja yhteisiä rituaaleja ei vielä ole ehtinyt muodostua. Vain kokeilemalla ja aktiivisella tekemisellä nähdään, mistä

verkostoaihioista lopulta kasvaa itsenäisiä ja pysyvämpiä verkkoyhteisöjä.

### Miksi digitoinnin osaamisverkosto?

Epäviralliset verkostot ovat etenkin kokemustiedon lähteenä ylivoimaisia. Digitointi-projektin ei tarvitse enää olla organisaation yksilösuoritus eikä jokaisen projektinvetäjän yksinäinen päänsärky. Digitaalisia verkostoja on kuitenkin osattava soveltaa tavalla, joka rakentuu digitaalisuuden vahvuuksille, ei heikkouksille. Esimerkiksi käsillä tehtävien monimutkaisten työvaiheiden opettelu on helpompaa fyysisesti samassa tilassa, vaikka opasvideot hyvänä tukimateriaalina toimivatkin.

Verkossa taas on helppo tavoittaa laaja jäsenyhteisö yhdellä kertaa. Keskusteluissa välittyy käden käänteessä tieto siitä, ketkä tietävät aiheesta, kuten myös tieto siitä, millaisia ohjeistuksia ja periaatteita tietyssä tehtävässä tulee soveltaa. On hyvä muistaa, että hiljainen tieto on läsnä kaikessa, mitä teemme ja sanomme – sitä ei voi irrottaa ihmisestä erilleen, vaikka kohtaaminen tapahtuisikin digitaalisissa kanavissa (Kosonen 2023).

Aikaisemmissa digitointiosaamisen kehittämishankkeissa (kts. Kirsi Hännisen artikkeli tässä julkaisussa) ja niiden opiskelijaryhmissä on noussut vahvasti esiin tarve jakaa tietoa onnistuneista ja ei niin onnistuneista digitoineista. Tässä on samaan aikaan sekä osaamisverkoston vahvuus että sen potentiaalinen sudenkuoppa: yhteinen tarve on kyllä selkeä, mutta samaan aikaan organisaatiokulttuurit ja

avoin tiedonjako voivat ajautua törmäyskursseille keskenään. Saako myös epäonnistuneista digitointiprojekteista puhua ääneen? Millaisilla foorumeilla? Ketkä niistä saavat kertoa?

Verkkoyhteisöissä oppimisen avoimuus ja organisaatorajattomuus vaativat vastapainokseen turvallisuuden kokemusta. Tätä on pyritty digitoinnin osaamisverkoston alkuvaiheessa edistämään kahta eri reittiä: toisaalta luottaen KAM-sektorin asiantuntijoiden asiakeskeisyyteen ja sen myötä rakentavaan keskustelukulttuuriin, ja toisaalta tarjoamalla palvelu suomalaisella luotetulla yhteisöalustalla (Skillhive), mikä mahdollistaa sen, ettei verkosto ole algoritmien armoilla vaan hallinnoi itse omia tietojaan ja keskustelujaan. Alustalla voi myös muodostaa pienempiä ja rajatumpia keskusteluryhmiä.



Tärkein yksittäinen tekijä digitaalisten verkostojen tiedonjakoa ajatellen on kuitenkin se, miten ihmiset tunnetasolla kiinnittyvät ryhmään ja kuinka jäsenet kohtelevat toinen toisiaan. Esimerkiksi onko verkostossa turvallista olla sellaistenkin, jotka ovat parhaillaan kouluttautumassa tiedonhallinta- tai arkistointitehtäviin, vaikka mukana on myös kokeneita alan asiantuntijoita? Tässä yksi ratkaisu on tarjota erityyppisiä keskustelualueita yhden yhteisen kanavan sijaan ja mahdollistaa edellä jo mainittu rajatumpien pienryhmien perustaminen. Rohkaisuna voi toimia myös se, että kaikki verkoston ylläpitäjät ja aktiivit eivät ole digitoinnin taitajia, vaan jokaisella on aina uutta opittavaa.

Lisäksi verkoston on tietenkin oltava jäsenilleen hyödyllinen, jotta kynnys monille uuden verkostopalvelun käyttöönottoon ja säännölliseen käyttöön pysyy riittävän matalana.

Kognitiivisen kuormituksen ja informaatiotulvan aikakaudella on erityisen tärkeää katalysoida verkostoja, jotka helpottavat elämää, sen sijaan että niiden koetaan asettavan asiantuntijoille vain lisää vaatimuksia.

### Mitä osaamisverkosto tarjoaa?

Digitoinnin osaamisverkostoon on sen alkuvaiheissa rakennettu viisi eri osiota:

- ◆ **Yleinen keskustelu** vapaasti valittavista aiheista
- ◆ **Organisaatiot** ja niiden esittelyt, ml. digitointipalveluita tarjoavat yritykset
- ◆ **Projektit:** eri organisaatioiden digitointiprojektit, digitoinnin kehittämishankkeet

- ◆ **Oppiminen:** digitointikurssit, koulutukset, tapahtumat
- ◆ **Rekryt:** työpaikkailmoitukset tiedonhallinnan ja digitoinnin alalta.

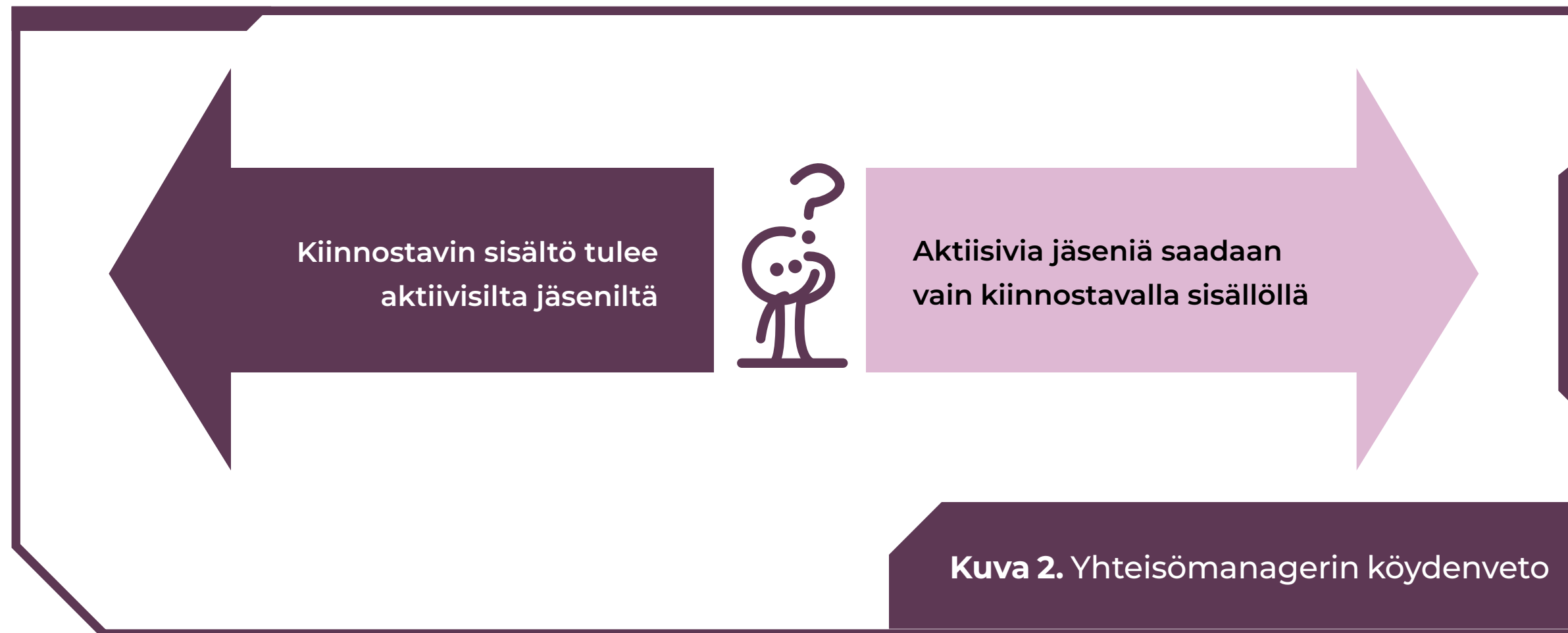
Nämä osiot palvelevat kaikkia jäseniä. Esimerkiksi yritykset ja digitointia toteuttavat organisaatiot ovat tehneet omia esittelyitään Organisaatiot-osioon ja Rekryt-osiossa on sekä jaettu työpaikkailmoituksia että etsitty harjoittelupaikkaa osana korkeakouluopintoja.

Säännöllisin väliajoin verkostossa julkaistaan myös yhteisiä haasteita tai tehtäviä. Ensimmäiset on jaettu ylläpidon toimesta, mutta kuka tahansa jäsen voi avata uuden yhteisen työkohteen tärkeäksi katsotusta asiasta. Ensimmäiset haasteet ovat liittyneet muun muassa tämän julkaisun nimikilpailuun (johon saatiin peräti 35 ehdotusta), digitointia teke-

vien organisaatioiden esittelyihin sekä kiinnostaviin digitointiprojekteihin.

Ensimmäisen puolen vuoden perusteella voidaan todeta, että osaamisverkoston jäsenmäärä on ylittänyt tavoitteet, mutta samaan aikaan jäsenten aktiivisuuden herättelyssä on ollut haasteita (kts. kuva 2). Siinä missä osa jäsenistä odottaa valmista pöytää, josta voi käydä poimimassa hyödyllisen tiedon lautaselleen, osa jäsenistä on ryhtynyt avaamaan itse aiheita ja keskusteluja, joita verkostolta toivovat. Valmista sisältöä on lanseerausvaiheessa oltava tarpeeksi, mutta ei kuitenkaan liikaa, jotta siihen perehtyminen ei käy liian työlääksi. Lisähaasteena on monien jäsenten eteen tullut uuden alustan käytön opettelu.

Jo yhden verkoston ravitseminen käy yhteisömanagereilta kokopäivätyöstä, mihin ei



reaalitodellisuudessa ole kenelläkään mahdollisuutta oman työn ohessa. Yhteisömanageroinnin näkökulmasta tarvitaan paitsi aikaa myös tasapainoilua siinä, kuinka paljon verkoston aktiivisuutta ruokitaan ylläpidon toimesta ja paljonko jää jäsenten oman aktiivisuuden varaan. Kuva 2 havainnollistaa tätä köyden vetoa.

Arkisena esimerkkinä uusien jäsenten toivottamisessa tervetulleeksi on oltava kärsivällinen: jos ylläpitäjä on aina se, joka reagoi heti ensimmäisenä, muut jäsenet tottuvat tähän eivätkä

välttämättä rohkaistu avaamaan keskustelua, vaikka uusi tulija olisi heille vanha tuttu.

### Osaamisverkoston tulevaisuus

Verkosto- tai yhteisömanageri on kuin arkkitehti, joka voi kyllä suunnitella taloon hienot rakenteet, seinän ja katon, mutta vasta taloon muuttavat ihmiset ratkaisevat, millainen koti siitä oikeasti tulee. Todellinen tarve ja kysyntä siis ratkaisee myös digitoinnin osaamisverkoston tulevaisuuden. Se voi jäädä lyhytaikaiseksi

kokeiluksi tai toimia vuosikymmeniä – tai jotain siltä väliltä. Ennustamattomuus kuuluu inhimillisen yhteistoiminnan perusluonteeseen.

Digitoinnin osaamisverkoston aktiivisuutta seurataan ja yhteisiä tehtäväkohteita kehitetään jatkuvasti, jolloin verkostolla on paremmat edellytykset kantaa itsensä alun ”kuolemanlaakson” eli käynnistysvaiheiden yli. Nämä ovat tunnusomaisia mille tahansa verkkoyhteisölle: toiminta aktivoituu ja passivoituu kunkin yhteisön tarpeiden mukaisesti. Jos osaamisverkostolle havaitaan olevan

pysyvä tarve ja kysyntää, sen on mahdollista toimia osana Memory Campus -klusteria ja sen jatkuvan oppimisen palveluita.

### Lopuksi

Julkaisumme on suunnattu digitoinnin kehittämisestä kiinnostuneille organisaatioille ja yksilöille. Tässä artikkelissa on käyty lyhyesti läpi myös verkostojen rakentamisen teoriaa ja käytäntöä sekä yhteisömanagerin roolia verkostojen alkuvaiheessa. Miksi näin? Ver-



kosten ja käytäntöyhteisöjen teoria auttaa ymmärtämään digitoinnin osaamisverkostoa, ja se puolestaan auttaa meitä ymmärtämään ja kehittämään digitointia.

Vertaisoppiminen ja -organisointumisen täydentävät perinteisempiä tapoja oppia ja jakaa tietoa erityisesti silloin, kun käsiteltävät ongelmat ovat monimutkaisia, niiden ratkaisemisessa tarvitaan aineettomia resursseja ja luovaa ajattelua, eikä kysymyksiin ole olemassa yhtä ainoaa oikeaa vastausta (Kosonen & Ruohisto 2017). Maaperä on nyt otollinen seuraavan sukupolven digitaalisille yhteistyöalustoille. Ne ovat luotettavia ja turvallisia ympäristöjä, joissa ihmisillä säilyy hallinta omaan tietoonsa

ja joissa työskennellään yhdessä muiden asiantuntijoiden kanssa. Tässä digitoinnin osaamisverkosto voi näyttää suuntaa muille vastaaville jatkuvan oppimisen toteutuksille.

## LÄHTEET

Blanchard, A. & Markus, M. L. 2004. The Experienced "Sense" of a Virtual Community: Characteristics and Processes. *The DATA BASE for Advances in Information Systems* 1, 65–79.

Digitoinnin osaamisverkoston jäsenhakemus. <http://memorycampus.fi/digitoinnin-osaamisverkosto> [viitattu 2.12.2024].

JAMK avoimet oppimateriaalit. Osaamisen tunnistaminen. <https://oppimateriaalit.jamk.fi>

[fi/fi/portfolio/portfolio-jamkissa/osaamisen-tunnistaminen/?gad\\_source=1](https://memorycampus.fi/fi/fi/portfolio/portfolio-jamkissa/osaamisen-tunnistaminen/?gad_source=1) [viitattu 10.12.2024].

Kim, A. J. 2000. *Community Building on the Web: Secret Strategies for Successful Online Communities*. Berkeley, CA: Peachpit Press.

Kosonen, M. 2008. Knowledge sharing in virtual communities. *Acta Universitatis Lappeenrantaensis*, 335.

Kosonen, M. & Ruohisto, J. 2017. Parviäly ja ongelmanratkaisu digitaalisilla alustoilla. *Työn tuuli* 2/2017, 26–35.

Kosonen, M. 2023. Yhteisöllisyys ja sen johtaminen etätöiden aikakaudella. *Työn tuuli* 2/2023, 47–51.

Kukkonen, H. & Raudasoja, A. 2018. Osaamisperusteinen ammatillinen koulutus. Teoksessa H. Kukkonen & A. Raudasoja (toim.)

Osaaminen esiin: ammatillisen koulutuksen reformi ja osaamisperusteisuus. Tampere: Tampereen ammattikorkeakoulu, 9–14.

Lave, J., & Wenger, E. 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge: Cambridge University Press. Saatavissa: <http://dx.doi.org/10.1017/CBO9780511815355>.

Työ- ja elinkeinoministeriö. Jatkuva oppiminen. Saatavissa: <https://tem.fi/jatkuva-oppiminen> [Viitattu 10.12.2024].

Wasko, M. & Faraj, S. 2005. Why should I share? Examining social capital and knowledge contribution in electronic networks of practice. *MIS Quarterly*, 35–57.



Tiina Hölttä

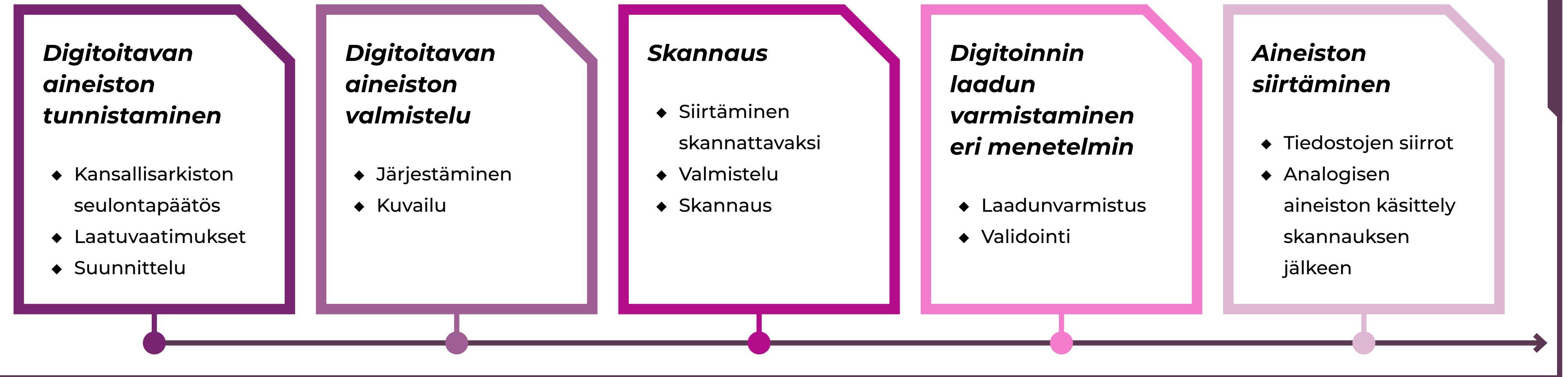
# LAADUN KÄSITE MASSADIGITOINNIN PROSESSISSA

Tämä artikkeli käsittelee Kansallisarkistossa tehtävää massadigitointia ja sen laatukäsitteitä digitointiprosessin aikana. Massadigitointi on Kansallisarkiston palvelu, jonka tehtävänä on tuottaa digitointipalveluja valtion viranomaisille. Tarkoituksena on, että valtionhallinnon pysyvästi säilytettävät analogiset tietoaineistot muutetaan keskitetysti digitaaliseen muotoon hävittämiseen tähtäävän digitoinnin vaatimusten mukaisesti. Digitoinnin jälkeen vastaava analoginen aineisto hävitetään tiettyjen kriteerien ja varoajan täyttyessä. Täten asiakirjasta jää pysyvästi säilytettäväksi ainoastaan digitaalinen ilmentymä.

Massadigitoinnille on tunnusomaista keskitetty toiminta, aineiston mittava määrä ja analogisen aineiston hävittäminen. Laajamittaisuus ja tavoitteet ovat kansainvälisesti kiinnostavia, koska arkistoaineistoja ei tyypillisesti hävitetä. On hyvä samalla todeta, että Kansallisarkiston massadigitoitavat aineistot ovat arkistomielessä nuorempia aineistoja, pääsääntöisesti 1960-luvulta lähes nykyhetkeen saakka. Tämän vuoksi ne ovat myös suhteellisen hyväkuntoisia, eikä niillä ole vastaavaa kulttuurihistoriallista perintöarvoa kuin vuotta 1921 vanhemmalla asiakirja-aineistolla on.



Kuva 1. Digitointiprosessi  
(Kansallisarkisto 2024)



Kansallisarkiston massadigitoinnin tuotantoa tehdään tällä hetkellä kahdella paikakunnalla Keskusarkistossa Mikkelissä ja Siltavuoren toimipisteessä Helsingissä. Tiimin koko vaihtelee noin 20 ja 35 henkilön välillä rahoituksen vaihtelujen mukaisesti. Lisäksi massadigitoinnin palvelun prosessiin osallistuu Kansallisarkiston siirtotimin (viranomaisohjaus) henkilöiden lisäksi lukuisa määrä eri viranomaisia riippuen siitä, kenen aineistoa tuotannossa kyseisinä toimintavuosina massadigitoidaan.

Seuraavaksi käsittelem digitointiprosessia ja sen tavoitteita valottaakseni aineiston matkaa viranomaiselta massadigitointiin: mitä pääpiirteissään tapahtuu tämän matkan varrella sekä miten analogisen ja digitaalisen aineiston tiet erkanevat. Sen jälkeen kuvaan laadun eri ulottuvuuksia massadigitoinnin prosessissa. Artikkelin lopussa tuon esille aineiston hävittämistä laatu näkökulmasta tarkasteltuna. Esitän myös muutamia ajatuksia massadigitoinnin yhteiskunnallisesta vaikuttavuudesta ja merkityksestä hävittämiseen tähtäävässä digitoinnissa.

## Digitointiprosessi ja sen tavoitteet

Massadigitoinnin prosessi mukailee mitä tahansa digitointiprosessia (kuva 1), jonka ”tavoitteena on tuottaa autenttisia ja todistusvoimaisia digitaalisia ilmentymiä alun perin analogisessa muodossa syntyneestä tiedosta. Analogisen asiakirjan tietosisältö pitää siirtää uuteen muotoon eheästi – kokonaisuudessaan ja virheettömästi. Tiedon säilymistä eheänä on voitava seurata tiedon elinkaaren kaikissa vaiheissa.” (Kansallisarkisto 2024.)

Massadigitoinnin palveluprosessi on pitkä. Se on käytännössä päättymätön ketju erilaisia tehtäviä ja toimenpiteitä eri toimijoiden tekemänä. Kuvan 1 mukaisesti ensin on tunnistettava massadigitoinnin kriteerit täyttävä viranomaisaineisto, jolle tehdään Kansallisarkiston seulontapäätös (aineiston säilytysmuoto vaihdetaan analogisesta digitaaliseksi). Kansallisarkiston digitoinnin laatuvaatimukset määrittävät viranomaisyhteistyötä ja ohjausta sekä digitoitavan luovutuskokonaisuuden valmistelua. Viranomaisjärjestäjä, seuloo ja kuvailee digitointiin tulevat aineistot



### Viranomaisohjaus

- ◆ Tiekartta 135 hkm:n tuottavuushankkeina v. 2024 →
- ◆ Digitointisuunnitelma
- ◆ Sopimukset: viranomaiset ja kuljetus

### Siirtokuljetus

- ◆ Kuljetuksen suunnittelu ja kuljetus
- ◆ Vastaanoton valvonta
- ◆ Vastaanottotarkistus
- ◆ Siirron lopputoimet metatietojärjestelmässä

### MD logistiikka

- ◆ Aineiston tuonti tuotannon logistiikkaohjelmaan
- ◆ Eteneminen tuotanto- ja viikkosuunnitelman mukaan
- ◆ Aineistohaut ja palautukset makasiiniin

### MD valmistelu

- ◆ Digitointiohje
- ◆ Skannerien esivalinta
- ◆ Niitit, liittimet tms. pois

### MD skannaus

- ◆ Erillisskannaus (S): dokumenttiskannerit (vain tarvittaessa)
- ◆ Erikoisskannaus (yli A3, postimerkkikoot, sidokset): suurteho, masto ja mankeli

### MD valmistelu + skannaus

- ◆ Yhtäaikainen valmistelu ja skannaus (V + S) yleisin ja nopein tapa digitoida dokumenttiskannereilla

### MD Validointi

- ◆ 100 % laaduntarkistus
- ◆ Vakioidut hylkäyssyyt
- ◆ Tekoälykomponentit avustavat (koneoppivat API:t)

### MD uudelleenskannaus ja validointi

- ◆ Skannataan yksittäiset asiakirjat / koko säilytysyksikkö
- ◆ Uudelleenvalidointi

### Analogisen ja sähköisen aineiston tiet erkanevat

Kansallisarkiston luettelointi- ja metatietojärjestelmään Kansallisarkiston viranomaisohjauksesta vastaavan siirtotiimin ohjauksella vuodesta kahteen ennen aineiston siirtämistä massadigitointiin, skannaukseen ja validointiin. Digitoinnin jälkeen analogisen ja digitaalisen aineiston tiet erkanevat. Analoginen aineisto palautetaan makasiiniin odottamaan hävitystä ja digitaaliset tiedostot siirtyvät käyttö- ja pitkäaikaissäilytysjärjestelmiin.

Massadigitoitavan aineiston prosessi on pääpiirteiltään kuvan 2 mukainen. Prosessikuvaus on tarkempi kuin kuvassa 1 ja sisältää enemmän sekä prosessitehtäviä että toimenpiteitä.

Ennen massadigitoinnin tuotantovaihetta on viranomaisten kanssa kyselyllä selvitetty massadigitointiprosessiin soveltuvan aineiston määrää hyllykilometreinä. Lähtövaiheessa

Kuva 2. Massadigitoinnin prosessi (Hölttä 2024)



arvioidaan viranomaisilla olleen noin 135 hyllykilometriä massadigitointiin soveltuvaa aineistoa. Vuosina 2019–2023 tiekartta-aineistojen massadigitointiin suunnattiin korvamerkittyä valtionrahoitusta. Se muuttui vuoden 2024 alussa tuottavuushankerahoitukseksi, jossa digitoidaan vain muutaman digitoitiohjelman valitun viranomaisen aineistoja kerrallaan. Tähän mennessä on digitoitu lähes yhdeksän hyllykilometriä viranomaisaineistoja (vajaat 10 %), joten lukuisa määrä viranomaisaineistoja on vielä digitoimatta.

Viranomaisohjausvaiheessa tehdään digitoitisuunnitelmat, tällä hetkellä kolmeksi vuodeksi kerrallaan, ja solmitaan erilaisia määrämuotoisia sopimuksia viranomaisen kanssa sekä kilpailutetaan aineistojen siirtokuljetukset tietyiksi määräajoiksi. Viranomaisohjauksesta vastaava siirtotiimi huolehtii siitä, että

viranomaisen valmistelee ajallaan ja ohjeiden mukaisesti luovutettavan viranomaisaineiston siirtokuntoon, kuten esimerkiksi huolehtii nimiötarrojen ja skannauksen ohjausarkkien muodostamisesta ja niiden liittämistä jokaiseen kotelo-, sidos- tai kortistomuotoiseen säilytysyksikköön (kuva 2).

Siirtokuljetusvaihe on siirtotiimin ja massadigitoinnin tuotannon yhteinen vaihe. Aineiston siirtovaihe toteutetaan tietoturvanäkökulmat huomioon ottaen. Aineisto vastaanotetaan valvotusti, luovutuskokonaisuus tarkistetaan kokonaisuudessaan, hyllymetrit mitataan sekä siirron lopputoimet tehdään metatietojärjestelmässä (kuva 2).

Kun tuodaan MD-logistiikkaan eli massadigitoinnin logistiikkaohjelmaan aineistotiedot metatietojärjestelmästä rajapinnan

kautta, alkaa analogisen aineiston matka käytännössä kohti tuotantoa ja aineiston käsittely siirtyy kokonaan massadigitoinnin tuotannon vastuulle. Tuotantosuunnitelman ja logistiikkaohjelman avulla ohjataan ja seurataan tuotannon etenemistä päivittäin, viikoittain, kuukausittain tai vuosittain. Tuotannolle tehdään päiväkohtaiset viikkosuunnitelmat työpisteittäin ja henkilöittäin. Analogisia aineistoja haetaan tuotantoon ja palautetaan makasiiniin. Tuotantotilojen infonäytöt kertovat tuotantoaineiston etenemisestä: erillisestä skannauksesta yhtäaikaiseen valmisteluun ja skannaukseen sekä validointiin asti. Laajaan digitoitiohjeeseen on dokumentoitu sekä kirjallisesti että kuvaesimerkein eri massadigitoinnin vaiheisiin liittyvä ohjeistus, jolla pyritään varmistamaan yhdenmukainen arkistoaineistojen käsittely paikkakunnasta tai henkilöstä riippumatta ja muuttamaan

analogisen aineiston tietosisältö digitaaliseen muotoon eheästi asiakirja-aineistojen todistusvoimaisuus säilyttäen (kuva 2).

Aineisto määrittää, mitä toimenpiteitä sille on tehtävä ja missä vaiheessa näitä toimenpiteitä tehdään. Voidaanko digitoida yleisimmällä ja taloudellisesti tehokkaimmalla tavalla yhtäaikaisesti valmistelua ja skannausta vuorotellen peräkkäin jaksotellen? Jaksottelulla tarkoitetaan sitä, että ensin valmistellaan (tekemällä useita toimenpiteitä, esimerkiksi poistamalla liittimet ja niitit) pieniä asiakirjanippuja. Tämän jälkeen skannataan juuri valmisteltu pieni asiakirjanippu skanneriin kerralla. Työvaiheet etenevät valmistelun ja skannauksen vaihdellessa niin kauan, että koko kotelon sisältö on valmisteltu ja skannattu kokonaisuudessaan. Toisinaan on välttämätöntä valmistella aineisto ensin: esimerkiksi sidokset



on purettava, kannet irrotettava, sivut leikat-  
tava yksittäisiksi arkeiksi ja vasta sen jälkeen  
skannattava aineisto omana vaiheenaan. Yli  
A3-kokoiset asiakirja-aineistot tai yksittäiset  
asiakirjat on skannattava erikoisskannereilla  
käyttäen joko masto- tai mankeli (laajakoko)  
-skanneria (kuva 2).

Skannauksen jälkeen seuraa massadigitoin-  
nin prosessissa tärkeä laadunvalvontavaihe.  
Kaikki massadigitoidut kuvat validoidaan  
ohjelmallisesti ja ihmissilmin 100-prosentti-  
sesti. Näin toimitaan, koska aineiston pysyvä  
säilytysmuoto muuttuu sähköiseksi ja ana-  
logiset ilmentymät tuhoataan digitoinnin  
päättymisen sekä tietyn varoajan jälkeen.  
Validoinnissa hylätyt kuvat skannataan ja  
validoidaan uudelleen, kunnes hyväksytyt  
kuvapaketit ovat täydellisiä ja siirtovalmiita  
etenemään automaatioprosessissa sähköisen

palvelun järjestelmän kautta joko aineiston  
käyttöjärjestelmään tai kulttuuriperinnön pit-  
käaikaisjärjestelmään pysyvään säilytykseen.

Analogisen ja sähköisen aineiston tiet ovat  
erkaantuneet. Analoginen aineisto odottaa  
varoajan täyttymistä makasiinissa, ja vuo-  
sittaisten hävityssuunnitelmien mukaisesti  
aineisto kuljetetaan tietoturva- ja tietosuoja-  
periaatteita noudattaen jätteenkäsittelylai-  
tokseen, jossa aineisto tuhoetaan polttamalla.  
Polton yhteydessä vapautunut energia hyö-  
dynnetään kiertotalouden mukaisesti jätteen-  
käsittelyn lopputuloksena.

### Laadun eri ulottuvuuksia massadigitoinnissa

Digitoinnin laadun perusta rakentuu monita-  
soisesti, ja se koostuu lukuisista eri tekijöistä  
digitointiprosessin aikana. Käytännössä se  
alkaa jo asiakirjallisen aineiston syntymisestä  
viranomaistehtävän yhteydessä. Kuten Häk-  
kinen ym. (2023) kuvaavat, laadunhallinnan  
fokuksen tulisi olla laadun optimoinnissa ja  
virheiden estämisessä – ei virheiden korjaa-  
misessa jälkikäteen. Tältä ei pitkässä massa-  
digitoinnin ketjussa pystytä aina välttymään,  
mutta tavoitteena on saada tietosisältö mah-  
dollisimman kattavasti ja muuttumattomasti  
tallentumaan alkuperäisen objektin mukai-  
sena. Digitaalisen tallenteen tulee säilyttää  
todistusvoimaisuus ja tiedon säilyä luotetta-  
vana. Toisaalta alkuperäisen asiakirja-aineis-  
ton tie digitointiin on pitkä ja erilaista käyttöä

sekä vaurioita aineistolle on voinut jo tapah-  
tua matkan aikana ennen digitointia.

Tiivistettynä digitoinnin laatu perustuu erilai-  
siin teknisiin suosituksiin ja digitoinnin laa-  
dun mittareihin sekä parhaisiin käytäntöihin,  
kuten dokumentoituihin työnkulkuihin, suun-  
nitelmallisiin prosesseihin ohjeistuksineen  
ja perehdytykseen. Digitointituotannossa  
jatkuva, eheä ja johdonmukainen laa-  
dunarviointi on arkipäivää. Tueksi tarvitaan  
erilaista teknistä välineistöä, kuten kuvatar-  
getien käyttöä, kuvan laadun testausta ja  
analysointiohjelmien avustusta ihmistyön  
lisäksi. Ammattitaitoinen, monipuolisesti  
osaava henkilöstö on mielestäni tärkein  
voimavara kaikissa digitoinnin vaiheissa.

Kuvaan 3 on koottu laadukkaan, tehokkaan  
ja laajamittaisen digitoinnin edellyttämiä



Kansallisarkiston  
hävittämiseen  
tähtäävän  
digitoinnin  
vaatimukset

Viranomais-  
valmistelu ja  
ennakko-ohjaus

Ennakkotiedot  
digitoitavasta  
aineistosta

Perehdytys ja  
ammattitaitoinen  
henkilökunta

Digitointiohje:  
yhteisen tiedon  
luominen

Kuvanlaadun  
määritykset, UTT

Toimivat laitteet  
ja toimintaan  
soveltuvat tilat

Tekninen  
ympäristö, infra  
ja käytettävät  
sovellukset

Työsuunnittelu,  
ohjaus ja  
johtaminen

**Kuva 3.** Laadun eri ulottuvuudet  
massadigitoinnissa (Hölttä 2024)

laadun ulottuvuuksia. Hävittämiseen tähtäävään digitoinnin vaatimukset ja erilaiset tarkastuskriteerit (manuaaliset ja automaattiset) digitointiprosessin vaiheissa tuovat järjestelmällisyyttä toteutukseen.

Kansallisarkisto on myös kehittänyt omia sovelluksiaan tukemaan tehtävää digitointityötä. Aineiston etenemistä voidaan seurata makasiinista tuotantoon ja takaisin makasiiniin arkisto- ja säilytysyksikön tarkkuudella. Skannereiden kuvanlaatua tarkkaillaan ja mitataan päivittäin, ja tuotantoluvan saamisen edellytyksenä on laatukriteerien analyysien onnistunut läpäisy. Prosessi sisältää kaiken kaikkiaan lukuisia automaattisia ja ihmisen tekemiä tarkastusvaiheita. Lisäksi ohjelmiin on kehitetty ja kehitetään edelleen koneoppivia algoritmeja avustamaan ihmistä.



Validointivaiheessa jokainen digitoitu kuva tarkistetaan ohjelmallisesti ihmistyönä. Arkistoyksikkö joko hyväksytään tai hylätään. Yksittäisiä kuvatiedostoja (tai erittäin harvoin koko arkistoyksikkö) hylätään valmiin hylkäyssyyarvolistan vakioituja hylkäyssyitä käyttäen ja hylätyt sekä skannataan että validoidaan uudestaan.

## Analogisen aineiston hävittäminen, laatu ja massadigitoinnin yhteiskunnallinen vaikuttavuus

Tarkastelen lopuksi analogisen aineiston hävittämistä massadigitointiprosessin päätteeksi. Millaisia kriteerejä ja varotoimia on asetettu, että analogiset aineistot voidaan hävittää tietyn ajan jälkeen? Ja mitä yhteiskunnassa tästä hyödymme – millainen yhteiskunnallinen vai-

kuttavuus Kansallisarkiston massadigitoinnilla saavutetaan?

Kansallisarkisto on määrittänyt massadigitoinnin hävittämisen kriteereille useita edellytyksiä. Kuvassa 4 on nähtävissä lisää analogisen aineiston hävittämistä tukevia laatutekijöitä jo aiemmin esille tulleiden lisäksi. Hävittäminen perustuu Kansallisarkiston ja viranomaisen välisiin sopimuksiin, joissa annetaan lupa analogisten ilmentymien hävittämiselle. Aineistolle ei ole määritelty kulttuurihistoriallista arvoa, ja riittää, että se arkistoidaan pysyvästi digitaalisessa muodossa.

Mitä hyötyä ja yhteiskunnallista vaikuttavuutta massadigitointi tuottaa? Vuosittain noin kaksi kilometriä aineistoa massadigitoidaan, mikä tarkoittaa sitä, että viranomaisen käyttöön tuotetaan helposti käytettävää ja saavutet-

tavaa digitaalista tietoa aineistoa noin 25–30 miljoonan kuvatiedoston vauhdilla vuosittain. Viranomaisten tehtävien hoito digitoinnin myötä helpottuu ja tehostuu. Lisäksi saavutettavuus paranee. Kansalainen puolestaan saa palvelua nopeammin ja tiedontarpeisiinsa vastauksia viranomaisen välityksellä.

Samanaikaisesti massadigitointiin sijoitettu rahoitus käytännössä tuottaa merkittäviä säästöjä valtiolle, koska se digitoinnin myötä vapauttaa valtion viranomaisten arkistotiloja toisiin käyttötarkoituksiin ja vähentää kiinteistö kustannuksia. Toisaalta sähköinen säilyttäminen maksaa myös, mutta olisiko digitointi tämän päivän uhkatekijöiden maailmassa myös mahdollisuus siirtää aineistoja turvaan helpommin kuin satoja tai tuhansia hyllykilometrejä arvokasta analogisessa muodossa olevaa kulttuuriperintöaineistoamme?

Massadigitoinnin prosessi on pitkä, ja siihen sisältyy lukuisa määrä eri vaiheita ja laatutarkistuspisteitä. Laatu alkaa jo asiakirjan syntymävaiheessa ja jatkuu koko elinkaaren ajan. Kokonaislaatu määrittyy pitkän ajan kuluessa eri tekijöiden toteuttaessa omia tehtäviään ja tarkastuksiaan koko laadunvalvonnan pitkän prosessiketjun ajan aina asiakirjan synnystä ja käytöstä, viranomaisen siirtovalmistelusta, digitoinnin tuotannosta ja jälkikäsitteystä digitaalisen aineiston käyttöön ja pitkäaikais säilytysvaiheen toimiin asti.

Katson, että merkityksellisintä on meidän kaikkien ketjuun osallistuvien toiminta. Ihmiset viime kädessä määrittävät ja myös toteuttavat laatutoimia, mutta myös ymmärtävät laadun merkityksen laajamittaisessa massadigitoinnin kaltaisessa toiminnassa.



### Massadigitoinnin hävittämisen kriteerit – edellytykset

- ◆ Kansallisarkiston ja viranomaisen välinen sopimus, lupa hävittää analogiset ilmentymät
  - ▶ Toteutus arkistolain 14 a § edellyttämin perustein
- ◆ Aineisto määrätty arkistoitavaksi yksinomaan digitaalisessa muodossa (ei kulttuurihistoriallista arvoa)
- ◆ Luovutuskokonaisuuden metatiedot on hyväksytty KA:n metatietojärjestelmässä ja aineiston hallinnoijana KA
- ◆ Luovutustodistus on toimitettu siirtäjälle, kun aineisto on kuljetuksen jälkeen tarkastettu
- ◆ Digitointiprosessissa noudatetaan KA:n hävittämiseen tähtäävän digitoinnin vaatimuksia ja sisäistä ohjeistusta, jossa määritellään digitoinnin toteutustapa ja laadunvarmistuksen menetelmät
  - ▶ Useita automatisoituja ja manuaalisia tarkastusvaiheita
- ◆ Digitointiprosessissa muodostuneet kuvatiedostot on tarkastettu ja hyväksytty järjestelmätasolla osana pitkäaikaissäilytys- ja käyttöliittymäjärjestelmiin toimittamista (SAPA, KP-PAS)
- ◆ Arkistoyksiköille on metatietojärjestelmässä automaattisesti muodostunut tieto digitaalisista ilmentymistä sekä elinkaaritapahtumiin metatiedot aineiston digitoinnista ja metatietojen ja digitaalisten ilmentymien tallekappaleiden siirtymisestä pitkäaikaissäilytykseen (PAS-AIP-ID)
- ◆ Digitaalisten ilmentymien (käyttökappaleiden) käytettävyys asiointipalvelussa (Astia) tarkastettu digitoinnin valmistumisen jälkeen
  - ▶ Varmistettu aineiston käytettävyys ja digitointi otannalla
- ◆ Analoginen aineisto on ollut edellisten kriteerien toteutumisen jälkeen säilytyksessä vähintään kuuden (6) kuukauden ajan (massadigitoinnin varoaika)
- ◆ Lisäedellytykset: Hävittäminen luovutuskokonaisuuksina, prosessi, työvaiheet ja viestintä määritelty, hävittäminen huomioitu riskienhallintadokumentaatioissa ja -menettelyissä (sähköisen säilyvyyden ja käytettävyyden varmistaminen häiriö- ja poikkeustilanteissa)

Kuva 4. Hävittämisen kriteerit massadigitoinnissa (Kansallisarkisto 2024)

### LÄHTEET

Arkistoitavien asiakirjojen digitointia koskevat vaatimukset ja ohjeet (A/20919/30.01.00/2024). Kansallisarkisto 2024. PDF-dokumentti. Saatavissa: <https://kansallisarkisto.fi/documents/141232930/151500284/Arkistoitavien+asiakirjojen+digitointia+koskevat+vaatimukset+ja+ohjeet.pdf/9ea50bdd-ad59-300c-a6ce-1c7bbe5f9d3f/Arkistoitavien+asiakirjojen+digitointia+koskevat+vaatimukset+ja+ohjeet.pdf?t=1734607700682> [viitattu 14.1.2025].

Häkkinen, H., Jaakkola, S., Järvinen I., Kettunen, J. & Kilponen, M. 2023. Digitoinnin laadunhallinta. Tekninen ohjeistus kaksiolotteisten kulttuuriperintöaineistojen digitointiin museoissa. Museovirasto, Versio 1.1. PDF-dokumentti. Saatavissa: [Digitoinnin laadunhallintaohjeistus](#) [viitattu 16.1.2025].

Hölttä, T. 2024. Laadun käsite massadigitoinnin prosessissa. Luentomateriaali Xamkin Digitoinnin kehittäjä -koulutuksessa 5.10.2024.

Kansallisarkisto. 2024. Digitointia koskevat Kansallisarkiston vaatimukset ja ohjeet. WWW-sivu. Saatavissa: <https://kansallisarkisto.fi/kansallisarkiston-digitointimaaritykset> [viitattu 14.1.2025].



Sami Kangasniemi

# MITEN AINEISTON ELINKAARI NÄKYY DIGITOIJAN TYÖSSÄ

Kerron tässä artikkelissa asiakirja-aineiston digitoinnista Kansallisarkiston massadigitoinnin palvelussa. Kirjoituksessani käsittelen digitointi- ja suunnitteluprosessia digitointiin saapuvan aineiston ja digitointityötä tuotannossa tekevän digitoijan näkökulmasta sekä seikkoja, joita kyseisen aineiston digitoinnissa tulisi ottaa huomioon.

## Massadigitoinnin tuotanto

Massadigitoinnin palvelu on Kansallisarkiston tuottama palvelu, jossa valtion viranomaisten massadigitointiin digitoitavaksi soveltu-

vat, pysyvästi säilytettävät analogiset aineistot muutetaan digitaaliseen muotoon. Vuosittainen digitointimäärä massadigitoinnissa on ollut noin 2000 hyllymetriä (hm), mikä tarkoittaa noin 20–30 miljoonaa digitaalista kuvaa. (Kansallisarkisto 2025)

Massadigitoinnissa käytettävä tuotantovälineistö koostuu erityyppisistä skannereista ja aineiston valmisteluvälineistä. Merkittävin osa digitoitavaksi tulevasta aineistosta on kotoitua paperiaineistoa, ja sen vuoksi käytetyin skannerityyppi on nopea dokumenttiskanneri. Näiden lisäksi käytettävissämme on myös taso-, kirja- ja suurkuvaskannereita. Skan-



**Kuva 1.** Massadigitoinnin aineisto-  
valmistelun työkaluja (Kangasniemi 2024)

nausta edeltävässä aineiston valmistelussa käytämme monipuolisesti erilaisia valmistelu-  
välineitä, kuten niitinpoistajia ja erilaisia leik-  
kureita (kuva 1). Vaikka suurin osa aineistosta  
on arkistoitua paperiaineistoa, on se silti hyvin  
monimuotoista, ja siksi valmisteluvälineitä on  
oltava monipuolisesti digitoijan pöydällä jo  
pelkästään ergonomisista syistä.

### Digitointiin valmistautuminen

Ennen digitoinnin aloitusta massadigitoin-  
nissa on tapana pitää digitoijien ja digitoinnin-  
ohjaajan yhteinen aineistokatselmus, jossa  
tutustutaan tarkemmin tulevaan aineistoon  
valmistelun ja digitoinnin näkökulmasta.  
Aineistokatselmuksen aikana käydään läpi  
aineistoa kuvaavat tiedot ja katsotaan arkis-

tokoteloiden sisältöä. Mikäli aineisto on poik-  
keuksellista tai luonteeltaan sellaista, jota ei  
aiemmin ole digitoitu, on mahdollista tehdä  
pilotointi pienelle määrälle aineistoa, jotta  
saamme mahdolliset digitointiprosessia häi-  
ritsevät ongelmat selvitettyksi ennen varsina-  
isen tuotannon alkua. Pilotointi voi olla tarpeen  
esimerkiksi, jos aineisto sisältää kortistoja,  
sidoksia tai erilaisia paperilaatuja tai -kokoja.

Digitoijan apuna on myös massadigitoinnin  
digitointiohje, joka on jatkuvasti päivittyvä  
ohje, jossa kerrotaan erilaisten asiakirjojen  
valmistelusta, skannaamisesta ja validoin-  
nista. Ohjeen päivittämiseen voivat johtaa  
esimerkiksi ennen näkemättömät aineiston-  
kokoamisen tai arkistoinnin tavat, joita arkis-  
tonmuodostaja (massadigitoinnin kohdalla  
viranomaisen) on käyttänyt.



## Digitointia edeltävä aineiston elinkaari

Kuten aiemmin mainitsin, on massadigitointiin saapuva aineisto suurilta osin arkistokoteloihin ladottuja paperisia asiakirjoja. Vaikka kotelot ovat ulkoisesti saman näköisiä, ei digitoija voi ennen kotelon avaamista tietää, millaista ja minkä kuntoista aineistoa kotelo sisältää, vaikka aineisto kuuluisikin samaan arkistoon tai sarjaan edellisen kotelon kanssa.

Aineiston valmistelun näkökulmasta aineiston aiemmalla elinkaarella on iso vaikutus siihen, kuinka nopeasti ja laadukkaasti se voidaan digitoida. Valmistelun tehokkuuteen voivat vaikuttaa muun muassa aineiston ikä, säilytysolosuhde, säilytystapa, käsittelyn aiheuttama rasite ja käsittelyn sekä aineistokokoamisen käytännöt.

Aiemman elinkaaren tapahtumat voivat näkyä digitointiprosessin aikana erilaisina ongelmoina sekä ajankäytön ja virheiden lisäänty-

misenä. Paljon käytetty ja vuosikymmeniä vanha aineisto on usein erittäin pölyistä ja huonokuntoista. Skannattaessa tällaista aineistoa siitä irtoaa usein hienoa pölyä tai isompia paperinpaloja, jotka jäävät dokumenttiskannerin sisään. Näistä osa kulkeutuu kuvanlukualueelle ja aiheuttaa enemmän tai vähemmän raitoja digitaaliseen kuvaan. Tämä saattaa heikentää kuvanlaatua visuaalisesti: teksti voi näyttää yli- tai alleviivatulta. Pahimmillaan niillä voi olla vaikutusta tekstintunnistuksen laatuun, mikä voi vaikuttaa digitaalisen aineiston saavutettavuuteen. Pölyistä aineistoa käsiteltäessä dokumenttiskannerin puhtaanapito korostuu, mikä lisää ajankäyttöä puhtaanapitoon, kun skanneria on puhdistettava kuvanlukualueelta useammin.

Ajan myötä arkistokoteloissa olevat paperiarakit ovat saattaneet painua mutkalle, kuten kuvassa 2 on nähtävillä. Tämä johtuu siitä, että



**Kuva 2.** Asiakirjat painuneet mutkalle vajaatäyttöisessä arkistokotelossa (Kangasniemi 2024)

vajaatäyttöisiä koteloida on säilytetty pystyasennossa, kuten arkistoissa on tapana. Tällainen aineisto saattaa aiheuttaa jonkin verran ongelmia digitoijalle, koska skannattaessa arkit eivät asetu dokumenttiskannerista tullessaan nätisti pinoon, vaan ne voivat törmäillä toisiinsa. Ongelmien ehkäisemiseksi digitoija voi tehdä arkkiniipulle niin sanotun vastarullauksen, jonka avulla arkkien rullautumista saadaan korjattua skannauksen sujuvoitta-

miseksi. Muita keinoja ongelmien ehkäisyyn ovat skannaussuunnan muuttaminen tai pienempien nippujen skannaaminen kerrallaan.

Arkistokoteloissa olevat asiakirjat ovat usein liitetty toisiinsa erilaisin tavoin, joista tyypillisin tapa on niittaus yhdellä niitillä arkkiniipun yläkulmasta. Myös mm. klemmareita, nauhoja ja ns. Velobind-sidoksia on käytetty. Näistä klemmarit ja nauhat ovat helppoja ja nopeita



**Kuva 3.** Sähköinen niitinpoistaja (Kuva: Kansallisarkisto)



**Kuva 4.** Massadigitoinnin tuotantovälineitä (Kuva: Kansallisarkisto)

irrottaa, kun taas niitit ja Velobindit ovat haasteellisempia.

Aineistovalmistelun aikana digitoija näkee erilaisia liittämisen käytäntöjä, jotka voivat vaihdella suurestikin saman viranomaisen aineistoissa. Niittauksen osalta aineistoa on voitu liittää useilla niiteillä joko kulmista tai sivuilta, niittien päälle on voitu laittaa teippiä tai niitit on saatettu nitoa ristiin. Poikkeavat tavat hidastavat digitoijan työtä ja saattavat

aiheuttaa ongelmia, kuten repeämiä skannauksen aikana, jos digitoijalta on jäänyt nidonta huomaamatta, kun se on ollut muusta aineistosta poikkeavassa kohdassa.

Digitoijalla on hyvä olla monipuolisesti erilaisia niitinpoistajia käytössään, joista voi valita kuhunkin tilanteeseen sopivan. Kuvan 1 oikeassa alakulmassa näkyy viisi erilaista massadigitoinnissa käytettävää niitinpoistajaa, joista voi valita eri tilanteisiin sopivan

esimerkiksi niitin leveyden, arkkinipun pakisuuden tai niitin sijainnin mukaan. Niitinpoistajat ovat myös ergonomisesti erilaisia, koska työpäivän aikana voidaan poistaa satoja niittejä, joten saman liikkeen toistaminen voi aiheuttaa käsiin kohdistuvaa raskautta. Digitoijien käytössä on myös sähköinen niitinpoistaja, joka irrottaa niitin napin painalluksella ja on ergonomialtaan suositeltava niittien irrotukseen, vaikka ei välttämättä ole nopein tapa (kuva 3). Huomioitavaa on se, että laite

soveltuu vain niiteille, jotka sijaitsevat arkkinipun kulmissa tai reunoissa. Laite ei siis sovellu niittien irrotukseen sanoma- ja aikakauslehdistä tai muista aineistoista, jotka on nidottu keskitaiteesta.

### Digitointiprosessin suunnittelu

Digitointiprosessin suunnittelun osuus on merkittävä koko prosessin kannalta. Digitoitavan aineiston määrästä riippuen voi olla



tarvetta sekä pitkän että lyhyen tähtäimen suunnitelmille. Massadigitointiin soveltuva aineistomäärä on niin suuri (135 hkm), että sitä varten on tehty pitkän tähtäimen digitointisuunnitelman lisäksi lyhyen tähtäimen vuosittainen tuotantosuunnitelma. Digitoijan näkökulmasta suunnitteluvaiheen tärkeimpiä asioita on digitoitavaan aineistoon tutustuminen etukäteen, jotta sen sujuvaa digitointia varten on osattu hankkia oikeanlaisia skannereita, valmistelutyökaluja sekä muita välineitä (kuva 4).

## Muutamia massasta poikkeavia aineistoja

### Velobind-sidonta

Joskus aineistossa törmätään Velobind-sidontaan (kuva 5), jossa arkinpun reuna on

rei'itetty ja reikien läpi on painettu muovipiikit, jotka on sitten sulatettu. Tällöin sidonnasta tulee kestävä. Tällaisen sidonnan purkaminen on hankalaa ilman siihen soveltuvaa työkalua (kuva 6). Työkalulla vedetään sidontaa pitkin samalla työkalua kääntäen, jolloin terä leikkaa muovipiikit poikki ja arkit voidaan erottaa toisistaan. Purkaessa irtoaa muovinpalloja, joiden päätymistä aineiston mukana skanneriin tulee välttää mahdollisen skannerin rikkoutumisen vuoksi.

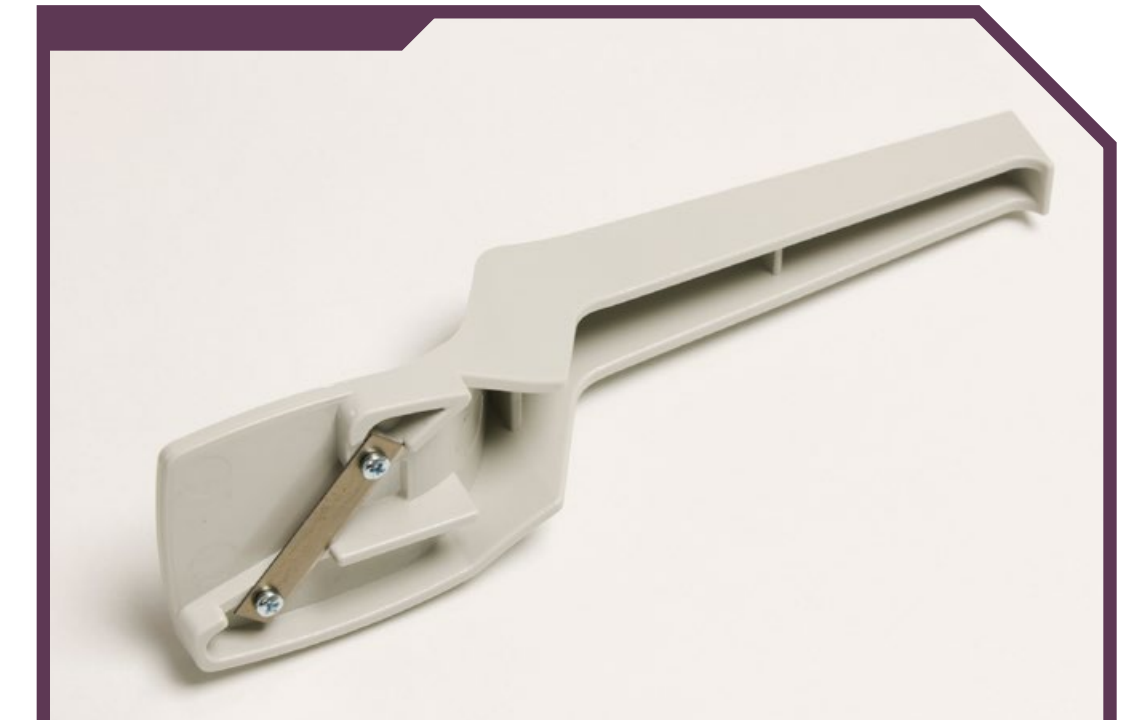
### Kortistot

Kortistot ovat myös yksi tavanomaisesta paperiaineistosta eroava aineistotyyppi, joiden digitoinnissa on otettava huomioon joitain seikkoja. Kortistoissa näkyy usein runsaasta käytöstä johtuvia jälkiä, kuten rispautuneita reunoja. Sen takia ne aiheuttavat pölyongelmaa skannereihin, mikä voi johtaa raitoihin



**Kuva 5.** Velobind-sidonta  
(Kuva: Kansallisarkisto)

digitaalisissa kuvissa. Huonokuntoiset reunat saattavat myös taittua skannatessa tietosivun päälle. Tämä ongelma voidaan välttää skannaamalla pino kortteja ehjin reuna edellä. Kortistot voivat olla kaksipuoleisia ja toisen puolen teksti voi olla eripäin, jolloin skannerilta vaaditaan ominaisuus, joka kääntää kuvat oikeinpäin riippumatta siitä, kummalta puolelta korttia kuva on otettu.



**Kuva 6.** Velobind-sidoksen purkaja  
(Kuva: Kansallisarkisto)

### Sidokset

Massadigitointiin saapuvan aineiston mukana tulee ajoittain erilaisia sidoksia, jotka pääsääntöisesti puretaan, jotta irtoarkit voidaan skannata läpisyöttöskannerilla. Tämä käytäntö on todettu nopeimmaksi tavaksi, vaikka purkamiseen kuluukin aikaa. Sen sijaan sidokset, joita ei voida purkaa esimerkiksi aukeamien



**Kuva 7.** Kortisto (Kuva: Kansallisarkisto)

yli menevän tietosisällön vuoksi, digitoidaan mastoskannerilla. Siinä on kirjahissi lasikannella, jonka alle sidos asetetaan.

Sidosten digitoinnissa yleinen ongelma on tiukan sidonnan aiheuttama tiedon häviäminen taitteeseen. Tätä ongelmaa voidaan helpottaa katkomalla muutamia nidontalankoja, jolloin sidos löystyy ja taite aukeaa paremmin. Jos sidoksessa on irtoselkä, voidaan selän ja

arkkien väliin työntää kapula, kun sidos asetetaan kirjahissiin ja selkä on kirjahissin tasojen välissä. Kapulan avulla sidos aukeaa paremmin ja se saadaan väännettyä tiukemmin lasiasten. Sidos tulisi asemoida kohtisuoraan kameraan nähden, jotta tietosisältö tulisi kuvattua myös taitteen keskeltä. Tarvittaessa sidoksen sijaintia voidaan muuttaa, jos ongelmana on lehden taittuminen taitteen päälle, kun sidos painetaan lasia vasten.

## Lopuksi

Lyhytkin kokemus massadigitoinnista on osoittanut, että asiakirja-aineistojen arkistoinnin käytänteet eri arkistonmuodostajilla saattavat vaihdella huomattavasti toisistaan. Myös arkistonmuodostajan sisäisissä käytännöissä on usein ollut eroavaisuuksia esimerkiksi toimipaikkojen välillä. Näiden eroavaisuuksien vuoksi digitoijan työpöydältä on löydettävä erilaisia työvälineitä, jotta aineistovalmistelu sujuisi tehokkaasti. On oltava käytettävissä

erilaisia skannereita, koska arkistot sisältävät perinteisen paperiaineiston lisäksi vaihtelevasti myös sidoksia, suurtulosteita, valokuvia ja läpivalaistavaa aineistoa, kuten diakuvia.

Digitointiprojektia suunniteltaessa on digitoitavaan aineistoon tutustuttava mahdollisuuksien mukaan hyvinkin tarkasti, jotta voidaan hankkia juuri oikeat työvälineet ja laitteet digitointia varten. Näin voidaan välttää mahdollisimman monet ongelmat projektin aikana, eikä digitointiin tule katkoksia puutteiden vuoksi.

## Lähteet

Kansallisarkisto. 2025. Massadigitoinnin palvelu. WWW-sivu. Saatavissa: <https://kansallisarkisto.fi/massadigitointi> [viitattu 15.4.2025].

Kangasniemi, S. 2024. Miten aineiston elinkaari näkyy digitoijan ja digitoinnin kehittäjän työssä. Luentomateriaali Xamkin Digitoinnin kehittäjä -koulutuksessa 5.10.2024.



**Kuva 8.** Sidos purettu valmistelukulmaan odottamaan skannausta. Paperiliuskoilla merkitty skannaajalle huomioitavia kohtia. (Kangasniemi 2024)



Anssi Jääskeläinen  
& Tuomo Räisänen

# TEKOÄLYLLÄ TEHOKKUUTTA DIGITOINTIPROSESSIIN

Tekoäly (AI, Artificial Intelligence) on ollut nouseva trendi viimeistään ChatGPT:n myötä. Tekoälyn hyödyntäminen on noussut kehityssuunnaksi myös digitoinnissa, jossa tekoälyltä haetaan muun muassa prosessien tehostamista, aineistojen laadun parantamista ja sujuvampaa käytettävyyttä. Digitointiprosessin läpivieminen on perinteisesti vaatinut aina ihmisen läsnäoloa monessa vaiheessa. Tämä luonnollisesti vie aikaa ja resursseja.

Tekoälyn avulla prosessin osia voidaan suorittaa automaattisesti, jolloin ne tehostuvat huomattavasti (Räisänen ym. 2023). Esimerkiksi tehostettu tekstintunnistus (OCR, Opti-

cal Character Recognition) ja automaattinen skannausvirheiden tunnistus mahdollistavat monien manuaalisten työvaiheiden automatisoinnin. Lisäksi tekoäly voi myös korjata skannausvirheitä sekä lisätä asiakirjoihin metatietoja ja asiasanoja, mikä helpottaa tiedon löytämistä.

Vaikka tekoälyn hyödyntämiseen digitoinnissa liittyy monia etuja, on tärkeää huomioida myös siihen liittyvät haasteet. Yksi keskeinen haaste on riittävän laajan ja laadukkaan opetusaineiston kerääminen tekoälyalgoritmien kouluttamiseksi. Lisäksi tekoälysovellusten käyttöönotto ja ylläpito vaativat investoin-



teja ja osaamista, mikä voi olla haaste erityisesti pienemmille organisaatioille (Chen ym. 2022). Xamkissa digitaalisen tiedonhallinnan tutkimuskeskus Digitalia on yhdessä kumppaneidensa kanssa kehittänyt avoimia tekoälypohjaisia ratkaisuja myös digitointiin.

Tässä artikkelissa tarkastelemme, miten tekoälyä voidaan hyödyntää digitointiprosessin tehostamisessa ja mitä etuja ja haasteita siihen liittyy. Lisäksi esittelemme esimerkkejä tekoälyn soveltamisesta digitointiprosesseissa ja pohdimme tulevaisuuden näkymiä tällä nopeasti kehittyvällä alalla. Jos kansalainen haluaa digitoida oman aineistonsa, siihen tuskin tekoälyä tarvitaan. Tilanne muuttuu kuitenkin, kun puhutaan niin sanotusta massadigitoinnista. Ennen massadigitoinnin aloittamista Kansallisarkistossa (KA) tehdyn

kyselyn mukaan valtion viranomaisilla oli 135 hyllykilometriä digitoitavaa aineistoa. Nykyiset skannerit voivat kyllä digitoida tätä aineistomassaa rivakkaan tahtiin, mutta laadunvalvonta ja tarkastaminen ihmisresurssein on mahdotonta. Tässä tekoäly tulee avuksi.

## Digitointiprosessi

Yleisellä tasolla digitointiprosessit sisältävät työvaiheita, joilla pyritään varmistamaan, että fyysiset aineistot saadaan muutettua korkealaatuisiksi digitaalisiksi käyttökelpoisiksi tiedostoiksi. Taulukko 1 esittää digitointiprosessin yleiset vaiheet sekä ne kohdat, joita tämän julkaisun taustalla tapahtuneessa työssä on kehitetty, testattu ja pilotoitu tekoälyavusteisesti.

**Taulukko 1. Digitointiprosessin vaiheet ja tekoälyn (AI) rooli niissä**

Vaihe	Selite	AI
<b>Aineiston valmistelu</b>	Digitoitavat aineistot lajitellaan ja järjestetään ennen skannausta, jonka jälkeen ne tarvittaessa puhdistetaan pölystä ja korjataan esimerkiksi repeämät.	
<b>Skannaus</b>	Jos mahdollista, valitaan sopiva skanneri aineiston tyyppin ja laadun mukaan ja määritetään soveltuvat skannausasetukset mm. resoluutio, tiedostotyyppi, tallennussijainti jne.	
<b>Virheiden tunnistaminen</b>	Pyritään tunnistamaan skannauksessa tulleet virheet, joita voivat olla mm. väärinpäin olevat kuvat, raidat, taittuneet kulmat, piiloon jääneet kohteet jne.	X
<b>Kuvanparannus</b>	Skannattuja kuvia voidaan parantaa esimerkiksi terävöittämällä tai poistamalla kohinaa. Monet kuvankäsittelyohjelmat, mm. Photoshop ja Gimp, kuten myös erilaiset ohjelmointikirjastot tarjoavat tähän työkaluja.	X
<b>Tekstintunnistus</b>	OCR-tekniikan avulla skannatut kuvat muunnetaan tekstiksi, mikä mahdollistaa asiakirjojen sisällön hakemisen ja muokkaamisen.	X
<b>Metatiedot</b>	Skannattuihin aineistoihin lisätään metatietoja, kuten otsikot, tekijät ja päivämäärät, mikä helpottaa niiden hakemista ja järjestämistä. Lisäksi aineistoihin voidaan liittää asiasanoja, jotka mahdollistavat helpomman löydettävyyden.	X
<b>Tallennus</b>	Digitoitu aineisto metatietoineen viedään käyttöön olevaan tallennus-/arkistojärjestelmään.	



## Virheiden tunnistaminen

Tekoälypohjaiset laadunvalvontajärjestelmät voivat analysoida suuria määriä digitoituja asiakirjoja nopeasti ja tarkasti. Tämä vähentää inhimillisten virheiden mahdollisuutta ja parantaa koko prosessin tehokkuutta.

Mikä on virhe digitoinnissa? Tätä voidaan lähestyä siltä kannalta, että informaatiota ei saa hukata. Toisaalta ne sivut, jotka eivät sisällä lainkaan informaatiota, ovat resurssien hukkaamista. Ensimmäisessä vaiheessa tarvitaan ihmiskättä ja -silmää. Jos ja kun digitointiprosessissa havaitaan systemaattisia virheitä, ne kerätään ja luokitellaan.

Paperiaineistoa digitoidaan useimmiten kaksipuoleisesti. Näin ollen vajaa puolet aineistosta sisältää tyhjiä sivuja. Vaikka tätä ei voi varsinaisesti virheeksi luokitella, on arvokasta

seuloa tämä materiaali pois jatkokäsittelystä. Tehtävä on ihmissilmälle helppo, mutta tekoälylle ei, ellei sitä erikseen tähän opeteta.

Tekoälyn virhealttiudesta on esimerkkinä läpi-kuultava skannaus. Virhe on kuitenkin siinä mielessä oikeansuuntainen, että informaatiota ei katoa. Piilossa olevaa informaatiota voi löytyä esimerkiksi taittuneiden kulmien tai muistilappujen takaa. Näitä tarkoituksia varten on kerätty aineistoa *Dalai*-hankkeessa ([kansallisarkisto.fi/dalai](https://kansallisarkisto.fi/dalai)) ja luotu malleja, jotka ovat testattavissa sivustolla [arkkiivi.memorilab.fi/](https://arkkiivi.memorilab.fi/). Mallien taustoista voi lukea lisää Arkkiivin sivuilta [arkkiivi.fi](https://arkkiivi.fi).

## Kuvanparannus

Skannerien tuottama kuvanlaatu voi vaihdella huomattavasti riippuen skannerista,

asetuksista, digitoitavan materiaalin laadusta ja monista muistakin tekijöistä (Hänninen 2019). Skannerin asetuksia on hidasta säätää jokaiselle materiaalille optimaaliseksi, minkä vuoksi kuvan jälkikäsittely ja parannustoimenpiteet ovat tärkeitä hyvän lopputuloksen saavuttamiseksi.

Digitoitujen kuvien parantamiseen on olemassa monia erilaisia menetelmiä. Esimerkiksi kirkkauden säätö parantaa kuvan yleistä valotusta, jolloin tummat alueet tulevat näkyvämmiksi ja vaaleat alueet kirkkaammiksi. Kontrastin säätö lisää eroa kuvan tummien ja vaaleiden alueiden välillä, mikä tekee kuvasta selkeämmän ja yksityiskohtaisemman.

Tasojen säätö parantaa kuvan dynamiikkaa säätämällä tummien ja vaaleiden alueiden tasoja, mikä lisää kuvan syvyyttä ja kontrastia. Terävyyden lisääminen parantaa kuvan

yksityiskohtia ja selkeyttä, jolloin pienet yksityiskohdat tulevat paremmin esiin. Kohinan vähentäminen puolestaan poistaa kuvasta ylimääräistä kohinaa ja häiriöitä, mikä parantaa kuvan laatua ja tekee siitä selkeämmän. Histogrammin tasoitus parantaa kuvan kontrastia ja tekee siitä tasapainoisemman.

Näitäkään menetelmiä ei voida kaikkia hyödyntää yhtä aikaa, vaan soveltuvimmat kuvanparannusmenetelmät on valittava tapauskohtaisesti. Tässäkin tekoälyä voidaan hyödyntää päätöksenteossa. Kaikkiin näihin ja moniin muihinkin parannusmenetelmiin löytyvät ohjelmalliset ratkaisut muun muassa Python-ohjelmointikielen tarjoamista kirjastoista, joita kehittämishankkeissamme on ahkerasti hyödynnetty (Jääskeläinen ym. 2023).



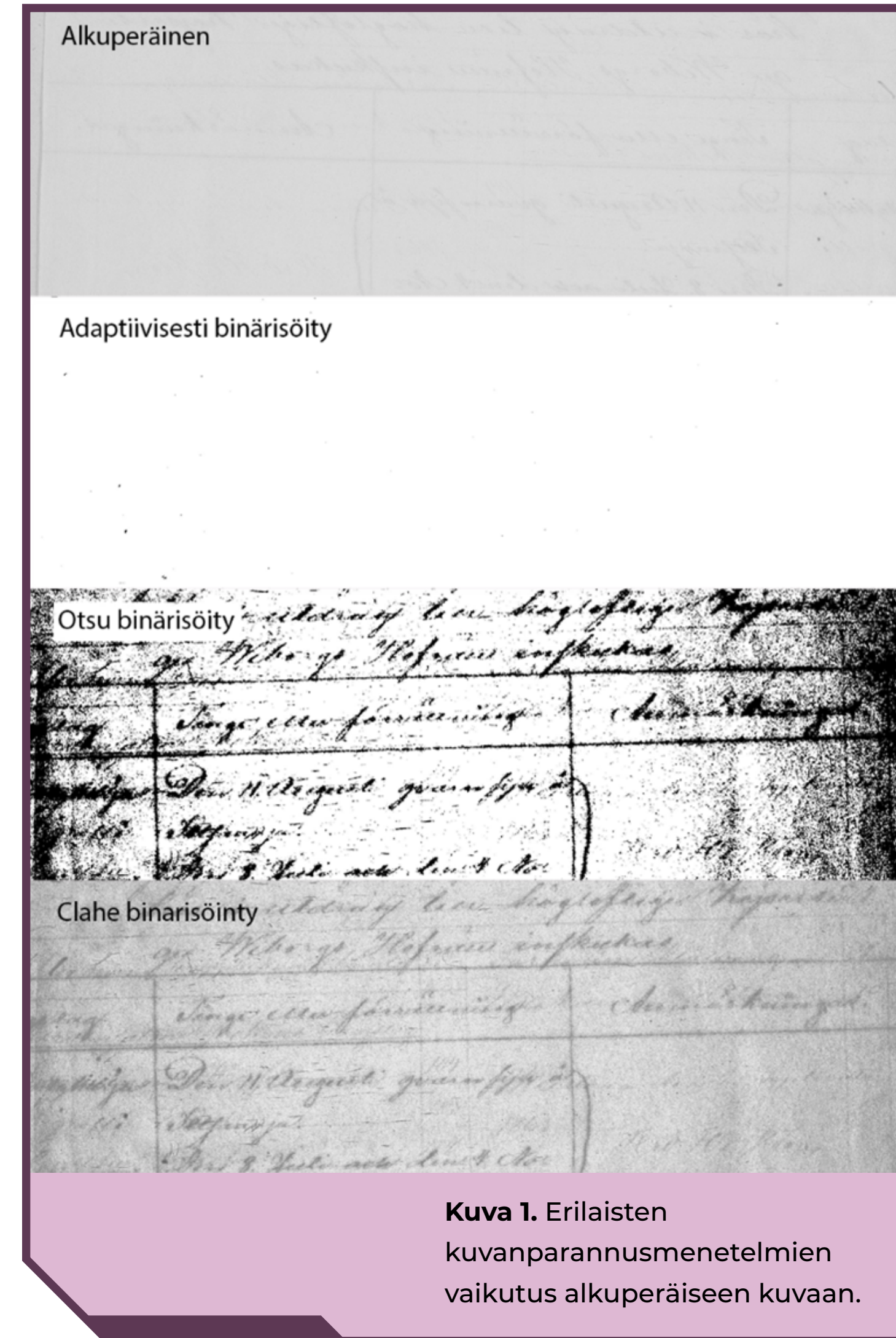
## Tekstintunnistus

Tekstintunnistus voidaan jakaa kahteen luokkaan: OCR (Optimal Character Recognition) ja HTR (Handwritten Text Recognition), jotka molemmat ovat digitoinnin kannalta merkittäviä. Ikävämpi seikka on se, että OCR-ohjelmat eivät tunnista käsin kirjoitettua tekstiä ja HTR-ohjelmat eivät puolestaan pidä konekirjoitetusta tekstistä (Clematide & Ströbel 2018). Tämä johtaa tilanteeseen, jossa on ensin tunnistettava, sisältääkö digitoitu sivu tekstiä, ja jos sisältää, onko se kirjoitettu käsin vai koneella.

Xamkin, Kansallisarkiston ja Elkan (Suomen Elinkeinoelämän Keskusarkisto) yhteisessä *Dalai*-hankkeessa kehitettiin ratkaisu myös tähän haasteeseen. Ratkaisu on osa hankkeessa kehitettyä Arkkiivi-palvelua, jota voi testata osoitteessa [arkkiivi.memorylab.fi](https://arkkiivi.memorylab.fi). Ark-

kiivin taustalla olevat toiminnallisuudet on rakennettu pääasiassa arkistojen käyttöön, mutta niitä voivat hyödyntää kaikki sellaiset toimijat, jotka voivat hyötyä automaattisesta asiasanoittamisesta ja tekstintunnistuksesta.

OCR:n de-facto työkalu kaupallisella puolella on pitkään ollut ABBYY FineReader, jonka uusimmissa versioissa hyödynnetään myös tekoälyä (Abbyy 2023). Avoimen lähdekoodin puolella vastaava työkalu on Tesseract, jota hyödyntävät muun muassa Google, IBM ja Adobe osana omia tuotteitaan. Tesseractin suurimpia ongelmia on sen jatkokouluttamisessa tarvittava tekninen osaaminen sekä koulutuksen haastavuus. Toinen suuri ongelma Tesseractissa on sen puuttuva GPU eli näytönohjaimen hyödyntämisen tuki, joka tarkoittaa



**Kuva 1.** Erilaisten kuvanparannusmenetelmien vaikutus alkuperäiseen kuvaan.



sitä, että sen kouluttaminen ja prosessoin-  
tinopeus ovat huomattavasti hitaampia kuin  
OCR-moottorilla, josta GPU-tuki löytyy.

Edellä mainitusta syystä *Dalai*-hankkeessa  
tekstintunnistuksen työkaluksi ja jatkokou-  
luttamiseen valittiin PaddleOCR, joka on sekä  
helpompi kouluttaa että tukee rinnakkaista-  
mista niin monelle GPU:lle kuin on saatavilla.

Jatkokoulutuksen tarve johtuu siitä, että  
OCR-mallit on kehitetty lähinnä englannin-  
kielistä materiaalia varten. Kansallisarkiston  
ja Elkan kanssa kerättyä aineistoa käytettiin  
PaddleOCR:n hienosäätöön hyvin tuloksin  
[xamk.fi/hanke/tekoalylla-lisaarvoa-digiarkis-  
tojen-asiakkaille](https://xamk.fi/hanke/tekoalylla-lisaarvoa-digiarkis-tojen-asiakkaille). Tässä yhteydessä kävi ilmi las-  
kentakapasiteetin tarve. Perustietokoneella ja  
käyttämällämme aineistolla hienosäätö (fine  
tuning) olisi kestänyt noin 284 vuorokautta.

Koska Digitalia-tutkimusyksiköllä on suurteho-  
laskennan kapasiteettia omasta takaa Memory  
Labissa ([memorylab.fi/laitteisto](https://memorylab.fi/laitteisto)), mallin jatko-  
koulutus onnistui alle kahdessa tunnissa.

Tekoälylle on haasteellista tunnistaa, onko  
materiaali kirjoitettu käsin vai koneella. Mikäli  
se onnistuu, tekstintunnistus paranee mer-  
kittävästi. Tähän suuntaan ollaan menossa ja  
lupaavia tuloksiakin on jo saatu. Esimerkiksi  
kuvan1 Otsu binärisöity -kohdasta Kansallisar-  
kiston käsin kirjoitetun tekstin tekoälypoh-  
jainen multicentury-htr-malli ([huggingface.  
co/Kansallisarkisto/multicentury-htr-model](https://huggingface.co/Kansallisarkisto/multicentury-htr-model))  
tunnisti seuraavaa tekstiä: *Förståndringh, som  
köpteligen och Till För och Wiborgs Hofrum  
inskickas, till tillskap hafwa, sådane först han  
är honom ifrånwarande, att från tillförende  
Bengt effter eller förordningar.*

Mainittu malli pohjautuu Hugging Face -pal-  
velussa vapaasti saatavilla olevaan micro-  
soft/trocr-large-handwritten-malliin, joka ei  
samaisesta kuvasta löytänyt lainkaan tekstiä.  
Kansallisarkistossa on jatko-opetettu malli  
ymmärtämään paremmin vanhaa suomen ja  
ruotsin kieltä, ja tulokset puhuvat puolestaan.

## Metatiedot

Tekoäly voi tunnistaa ja poimia asiakirjoista  
tärkeitä tietoja, kuten nimet, päivämäärät ja  
osoitteet. Tämä helpottaa tietojen käsittelyä ja  
analysointia. Lisäksi tekoäly voi verrata poimit-  
tuja tietoja ennalta määriteltyihin sääntöihin  
tai tietokantoihin, mikä auttaa tunnistamaan  
virheitä ja epäjohdonmukaisuuksia. Tähän  
käytämme *Dalai*-hankkeessa hyväksi havait-  
tuja komponentteja. Tiedonhakua helpottaa

esimerkiksi Annif ([kansalliskirjasto.fi/fi/palve-  
lut/annif](https://kansalliskirjasto.fi/fi/palvelut/annif)) -ohjelmisto, joka pyrkii kokoamaan  
asiasanat käsiteltävästä dokumentista. Tämä  
komponentti on liitetty osaksi edellä esiteltyä  
Arkkiivi.fi-palvelua.

## Tulevaisuuden näkymät

Tekoälyn hyödyntäminen digitoinnissa ja  
arkistoissa on vasta alkutekijöissään, mutta  
sen potentiaali on valtava. Tulevaisuudessa  
voimme odottaa tekoälyn roolin kasvavan  
entisestään, kun teknologia kehittyy ja uusia  
sovelluksia otetaan käyttöön. Yksi merkittä-  
vimmistä kehityssuunnista on tekoälyn kyky  
oppia ja mukautua jatkuvasti. Tämä tarkoittaa,  
että tekoälyjärjestelmät voivat parantaa suo-  
rituskykyään ajan myötä, kun ne saavat lisää  
dataa ja palautetta.



Toinen tärkeä kehityssuunta on tekoälyn integroituminen olemassa oleviin teknologioihin. Tekoälyä käytetään jo monissa arkipäiväisissä askareissa. Se voi jo nyt analysoida ja tehdä todella nopeasti puheesta tekstiksi -muunnosta monilla kielillä, myös muilla kuin valtakielillä. Se voi myös tunnistaa ja analysoida tunteita sekä tunnistaa kasvoja ja esineitä reaaliaikaisesta videokuvasta.

Kaiken kaikkiaan tekoälyn tulevaisuus digitoinnissa vaikuttaa valoisalta. Sen avulla voidaan tehostaa prosesseja, tehdä aineistoista helpommin saavutettavia ja hyödyntää niitä uusilla ja innovatiivisilla tavoilla. Toimijoilta tarvitaan ennen kaikkea uskallusta siirtyä hetkeksi epämurkkualueelle ja mahdollisuutta

testata erilaisia tekoälyratkaisuita omassa toiminnassaan. Xamkissa Digitalia-tutkimuskeskus ja sen tekoälyn kouluttamiseen ja hyödyntämiseen erikoistunut Memory Lab ovat pyrkineet omassa toiminnassaan tuomaan tällaisia pilotteja ja kokeiluita kaikkien hyödynnettäväksi.

Uusimpana kehitystyön tuloksena [memorylab.fi/demot/asr](https://memorylab.fi/demot/asr)-sivustolta löytyy automaattinen puheesta tekstiksi -moottori. Työtä on jatkettu ETKOT-hankkeessa, [xamk.fi/hanke/etkot](https://xamk.fi/hanke/etkot), jossa edellä kuvattuja komponentteja edelleen kehitetään ja yhdistetään yhden käyttöliittymän alle. Tämä mahdollistaa aineiston käytettävyyden kaikille käyttäjille.

## LÄHTEET

ABBYY. 2023. ABBYY FineReader PDF: Powered by AI. FineReader Blog. Saatavissa: <https://pdf.abbyy.com/blog/finereader-powered-by-ai/> [Viitattu 14.1.2025].

Chen, Z., Wu, M., Chan, A., Li, X. & Ong, Y.-S. 2022. A Survey on AI Sustainability: Emerging Trends on Learning Algorithms and Research Challenges. arXiv. Saatavissa: <https://doi.org/10.48550/arXiv.2205.03824>.

Clematide, S. & Ströbel, P. 2018. Improving OCR quality of Historical Newspapers with Handwritten Text Recognition Models. DARIAH-CH Workshop, Neuchâtel, 29 November 2018 - 30 November 2018, University of Neuchâtel.

Hänninen, T. 2019. Digitointiopas. Xamk Kehittää 82. Mikkelä: Kaakkois-Suomen ammattikorkeakoulu. Saatavissa: <https://www.theseus.fi/handle/10024/227340>.

Jääskeläinen, A., Lipsanen, M., Föhr, A. & Räisänen, T. 2023. OCR quality: Key to enhanced Data Mining. 2023 3rd International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), Tenerife, Canary Islands, Spain, 1-6. Saatavissa: <https://doi.org/10.1109/ICECCME57830.2023.10252214>.

Räisänen, T., Lipsanen, M. & Föhr, A., 2023. AI Powered Tools for Improving Usability in Digital Archiving. Archiving Conference, 71-74. Saatavissa: <https://doi.org/10.2352/issn.2168-3204.2023.20.1.14>.



Ville Kajanne

# DIGITOINNIN RISKIT

Tässä artikkelissa käsittelen pysyvästi säilytettävien tietoaaineistojen digitointiin liittyviä riskejä ja niiden hallintaa. Vaikka viittaa tekstissä jonkin verran yleiseen riskienhallinnan ohjeistukseen, tarkastelen aihetta nimenomaan digitoinnin näkökulmasta. Käyn läpi sellaisia riskienhallinnan näkökulmia, joita digitoinnin tuotannossa on noussut esiin omassa työssäni. Artikkelini ei siten ole riskienhallinnan yleinen kuvaus eikä myöskään tuo esiin kaikkia digitointiin liittyviä riskejä.

Artikkelin ensimmäinen luku käsittelee riskienhallintaa yleisesti. Toisessa luvussa käyn läpi digitointiin liittyviä riskejä. Viimeisessä eli kolmannessa luvussa tarkastelen aihetta laadunvalvonnan näkökulmasta.

## Riskienhallinnan yleisiä periaatteita

Tietoturvan ja tietosuojan merkitys on viimeisen vuosikymmenen aikana noussut yhä tärkeämmäksi aiheeksi organisaatioiden tiedonhallinnassa. Digitalisaation jatkuvasti yleistyessä sekä EU-tasolla että kansallisesti aihetta koskevan lainsäädännön ja muiden normien määrä on lisääntynyt asettaen organisaatioille suoria velvoitteita. Viime vuosien kansainväliseen ja kansalliseen turvallisuuteen liittyvät uhkakuvat ovat entisestään vahvistaneet tätä kehitystä: henkilötietojen ohella keskiöön on noussut kysymys julkisten ja muiden yleisen turvallisuuden kannalta kriittisten organisaatioiden toiminnan vakaudesta, luotettavuudesta ja jatkuvuudesta.

Riskienhallinnan tavoitteena on varmistaa strategisten ja operatiivisten tavoitteiden saavuttaminen sekä turvata toiminnan jatkuvuus ja toimintaedellytysten säilyminen. Riskienhallinta on koko organisaation eikä vain sen jonkin yksittäisen toiminnon työtä. Yksittäisten, tiettyihin tehtäviin liittyvien riskien huolellinenkin tunnistaminen valuu hukkaan, jos siihen liittyvissä muissa prosesseissa vastaava työtä ei ole tehty. Tarvitaan kokonaisvaltaista, koordinoitua riskienhallintapolitiikkaa, jossa on selkeät johtamis- ja vastuuroolit niihin nimettyine henkilöineen. (Sääskilahti & Mustonen 2023, 15–17.)

Riskienhallinnan uskottavuus kulkee käsi kädessä siitä laaditun dokumentaation



kanssa. Perinteisesti riskienhallinta on lähtenyt liikkeelle yksittäisinä tiedostoina laadituista riskirekistereistä. Leikkisästi voisi sanoa, että dokumentaation pitäminen ajantasaisena on yksi riskienhallinnan suurimpia riskejä. Tosiasia taitaa olla, että yksittäisten tiedostojen ylläpitäminen tai toimivuus kokonaisen organisaation riskienhallinnan kannalta ei ole uskottavaa.

Järkevintä onkin ottaa käyttöön erillisiä riskienhallinnan (verkkoympäristön) työkaluja, jotka tiedon säilyttämisen ohella antavat rakenteet työlle ja ohjaavat sitä. Nämä valmiit työkalut pohjautuvat alan standardeihin (ISO 31000 ym.). Välineiden ja niihin laadittujen tarkempien kuvausten rinnalla riskienhallinnan dokumentaation toisen haaran muodostavat organisaation muut sisäiset riskienhallinnan ohjeet, kuten riskienhallinnan käsikirjat ja

aiheesta yleisemmällä tasolla laadittu sisäinen viestintä. (Sääskilahti & Mustonen 2023, 14–15.)

Puretaan vielä hiukan riskienhallinnan vastuita ja työtä. Vastuita voidaan jakaa ainakin neljälle eri taholle. Kokonaisvaltaista riskienhallintaa johtaa organisaatiossa aina joku, yleensä johtoon kuuluva henkilö. Riskienhallinnan käytännön työtä, kuten siinä käytettävien välineiden hankintaa ja niiden käyttöönottoa, johtaa toinen, esimerkiksi tietoturvan parissa toimiva henkilö. Riskien tunnistaminen, määrittely, dokumentaatio ja hallinta vastuutetaan mielellään koko henkilöstölle. Tällöin päällikkörooleissa olevat henkilöt vastaavat työstä omilla vastuualueillaan ja osallistavat siihen tiimejään. Osana tätä työtä nimetään riskien omistajat. Omistaja on vastuussa riskistä ja seuraa riskin käsittelyn edistymistä. Omistaja esimerkiksi ylläpitää

riskin tietoja ja huolehtii siitä, että riskille on määritelty riittävät hallintatoimenpiteet.

Riskienhallinta käynnistyy riskien tunnistamisella (Sääskilahti & Mustonen 2023, 20). Riskit dokumentoidaan käytettävään järjestelmään. Tyypillisesti riskit jaotellaan sisäisiin ja ulkoiisiin. Sitten riskit arvioidaan yleensä luokittelemalla niiden todennäköisyys ja vaikutus eli yhteisvaikuttavuus. Kun riskien peruskehikko (riskirekisteri) on olemassa, siirrytään niiden hallintaan. Jatkuvuudenhallinnassa tarkastellaan riskien toteutumisen vaikutusta organisaation toimintaympäristöön, varaudutaan mallintamalla erilaisten poikkeamatilanteiden skenaarioita ja keinoja niistä palautumiseen tai toiminnan organisoimiseen väliaikaisesti muulla tavoin sekä harjoitellaan skenaarioiden toteutumista ja henkilöstön kykyä toimia suunnitellun mukaisesti poikkeustilanteissa.

Näin riskienhallinta nivotaan osaksi organisaation jokapäiväisiä prosesseja ja vastuurooleja.

Riskit elävät ja muuttuvat ajassa organisaation tavoitteiden, painopisteiden, tehtävien ja henkilöiden vaihtuessa. Vähintään riskien arvioinnissa tyypillisesti tapahtuu muutoksia: joidenkin riskien kriittisyys laskee, toisten kasvaa. Vakiintuneen riskien seurannan lisäksi onkin hyvä käydä riskejä aika ajoin läpi yhdessä esimerkiksi työpajoissa. Riskienhallintatyössä keskeistä vastuuta kantavien tulee saada säännöllisesti koulutusta tehtävään. Vertaistuki ja sisäinen keskustelu on välttämätöntä.



## Strategisia ja taloudellisia riskejä

Riskit voidaan luokitella eri alueiden riskeihin ja edelleen tunnistaa keskeiset riskityypit sekä riskien lähteet. Tässä artikkelissa tarkastelen digitoinnin riskejä riskialueiden avulla, joista käsittelen kolmea tyypillisintä: strategisia, taloudellisia ja operatiivisia riskejä. Tässä luvussa käsittelen strategisia ja taloudellisia riskejä ja viimeisessä luvussa suuremmin laatuun liittyviä operatiivisia riskejä. (Sääskilahti & Mustonen 2023, 20–24.)

Riskien tunnistaminen ja hallinta edellyttävät, että digitoinnin prosessi on määritelty kokonaisvaltaisesti, jotta voidaan prosessivaiheittain lähteä arvioimaan niihin liittyviä riskejä. Jos digitointia toteutetaan palveluntarjoajan toimesta ja digitoinnin kohteena oleva analoginen

aineisto hävitetään digitoinnin valmistuttua, voidaan tunnistaa ainakin seuraavat digitointiin kokonaisuudessaan liittyvät prosessivaiheet:

- ◆ digitoinnista (palvelusta, aineiston käyttöehdoista ym.) sopiminen
- ◆ aineiston palveluntarjoajalle siirtämiseen liittyvät toimenpiteet (fyysinen valmistelu, metatietotyöt tmv.)
- ◆ aineiston siirtäminen
- ◆ digitointi (useita aliprosessivaiheita)
- ◆ jälkikäsittely (esim. sisällöntunnistusprosessi tai vastaavia käytön parantamiseen liittyviä prosesseja, digitaalisen tietoaineiston tarkistusprosesseja ja siirtämistä säilytys- ja käyttöympäristöihin)

- ◆ säilytys (digitaalisen, ja jos edelleen säilytetään, myös analogisen tietoaineiston säilytysjärjestelmät tai -tilat)
- ◆ hävittäminen (analogisen, ja jos osin hävitetään, digitaalisen tietoaineiston hävittämisprosessi).

Prosessit voi esittää myös Tiina Höltän artikkelissa (kuva 1) kuvatulla tavalla. Tässä artikkelissa ei käsitellä kaikkiin prosessivaiheisiin liittyviä riskejä.

Jokaisella organisaatiolla on tulevien vuosien toimintaa viitoittava strategia. On hyvä pohtia, miten suurta roolia digitointi strategian toimeenpanossa esittää: Voidaanko digitoinnin

riskejä pitää strategisina riskeinä? Entä onko digitointi ja sen toteutustapa linjassa strategian kanssa ja tukee strategisten tavoitteiden saavuttamista? Mitä epävarmuuksia digitointiin tästä näkökulmasta voi liittyä? Mitä digitoinnissa tulee ottaa huomioon, jotta se edistää ja toteuttaa strategisia tavoitteita? Alla muutama hypoteettinen esimerkki tällaisista riskeistä.

Jos organisaation tavoitteena on siirtää tietoaineistojensa säilytysmuotoa merkittävästi analogisesta digitaaliseen ja toteuttaa tätä siirtymää suurimittakaavaisella digitoinnin tuotannolla (digitoida suuria määriä aineistoja vuositasolla sen sijaan, että digitoitaisiin pitkällä tähtäimellä vähän kerrallaan vuosittain), voi riskiksi muodostua esimerkiksi aineiston



soveltuvuus digitointiin. Huonokuntoisen, hauraan tai fyysiseltä muodoltaan hankalan aineiston digitointi voi vaatia huomattavasti enemmän aikaa, erityisosaamista ja resursseja (henkilöstöä, erityislaitteistoa) kuin on osattu ottaa huomioon. Tuloksena voi olla digitoinnin tyrehtyminen ja strategisen tavoitteen (keskeiset tietoaineistot pääasiassa digitaalisessa säilytysmuodossa) saavuttamatta jääminen.

Jos digitointia tehdään merkittävässä laajuudessa, voi strategisena tavoitteena olla digitoinnin vaikuttavuuden lisääminen. Tällöin riskiksi voivat osoittautua digitointivalinnat ja digitoinnin käyttöä parantavien teknologioiden hyödyntäminen.

Helpoimmin digitoitavissa olevat ja käytön kannalta merkittävimmät tietoaineistot eivät välttämättä ole samoja. Digitointi voidaan

keskittää nopeimmin digitoitavissa olevaan aineistoon muun muassa kustannussyistä. Jos sen käyttöarvo on hyvin vähäinen, ei strateginen tavoite toteudu. Jos digitointi toteutetaan tavalla, jossa digitaalisen tietoaineiston käyttö on hyvin työlästä (hankalat tiedostorakenteet, puutteelliset käyttöliittymät, olemattomat käyttöä edistävät sisällöntunnistus- ja vastaavat tekoälypohjaiset ratkaisut), voi tuloksena olla tarve jatkaa analogisen tietoaineiston käyttöä kuten ennenkin ja jättää digitaalinen tietoaineisto käytännössä hyödyntämättä. Tuloksena on strategisen tavoitteen toteutumatta jääminen ja hyödyttömät taloudelliset panostukset.

Strategisten riskien arvioinnissa voi tarkastella digitointia hiukan yleisemminkin irrallaan eksakteista strategisista tavoitteista kysymällä, mikä on digitoinnin vaikutusarvo koko organisaation kannalta. Miten se suhteutuu orga-

nisaation ydintehtäviin, julkiseen imagoon, rahoituspohjaan, osaamisen painopisteisiin ja rooliin toimintakentällään? Ovatko digitoinnin riskit tältä pohjalta merkittäviä organisaation olemassaolon, uskottavuuden ja jatkuvuuden kannalta? Jos vastaus on myönteinen, on asia syytä kirjata strategisena riskinä. Tällöin se tunnustetaan paremmin kuin jos riskiin sisältyvät eri asiat esitettäisiin hajautetummin ja sirpaleisemmin muun muassa yksittäisinä operatiivisina riskeinä.

Taloudelliset riskit voivat limittyä strategisiin riskeihin. Erilaisia taloudellisia riskejä syntyy siitä, että digitointiin ei ole ymmärretty (tai voitu) panostaa riittävästi rahoitusta. Taloudelliset riskit voivat siten muodostua myös operatiivisiksi riskeiksi, ja jos kokonaisvaikutukset ovat merkittäviä, ne voivat kumuloitua strategisiksi riskeiksi.

Niukka (tai väärin painotettu) rahoitus vaikuttaa yleensä ensin digitoinnin toteutustapaan: hankitaan esimerkiksi vääränlaisia, suorituskyvyltään tai laatuominaisuuksiltaan riittämättömiä laitteita. Rahoitus voidaan kohdentaa kokonaan varsinaiseen tuotantoympäristöön, jolloin sen kehittäminen on erittäin vaikeaa – laajemmassa tuotannossa kehittämisympäristö on yhtä tärkeä kuin tuotantoympäristö. Jos henkilöstön rekrytointiin osoitetaan liian vähän rahaa, voi tuotanto laahata tavoitemääristään. Tai jos yksittäiset palkkasummat ovat liian matalia, on tuloksena henkilöstön heikko sitoutuminen, mistä puolestaan seuraa merkittäviä osaamisen kertymisen riskejä. Osaava henkilöstö hakeutuu siis muualle töihin, jolloin digitoinnin merkittävin resurssi on menetetty.

Taloudellinen niukkuus voi myös asettaa kohtuuttomia ehtoja digitointituotannolle. Jos



rahoitus on määritelty kunnianhimoisen tuotantonopeuden mukaan, voivat pienetkin häiriötilanteet aiheuttaa aikatauluista jälkeen jäämistä. Tyypillisiä tilanteita, joissa tiukkaan lasketut budjetit muodostavat huomattavaa taloudellista riskiä, ovat erilaiset käyttöönoton tilanteet. Esimerkiksi uuden laitteiston hankinta ja käyttöönotto voi itsestä riippumattomista syistä viivästyä merkittävästi suunnitellusta. Vielä suurempia taloudellisia riskejä liittyy koko tuotannon ensi vaiheen käynnistämiseen. Aikataulut ja niiden edellyttämä budjetti pitäisi laskea realistisesti niin, että käyttöönottovaiheisiin on varattu aikaa testaukseen, pilotointiin ja varsinaiseen käyttöönottoon. Jatkuva budjetin rajoilla tapah-

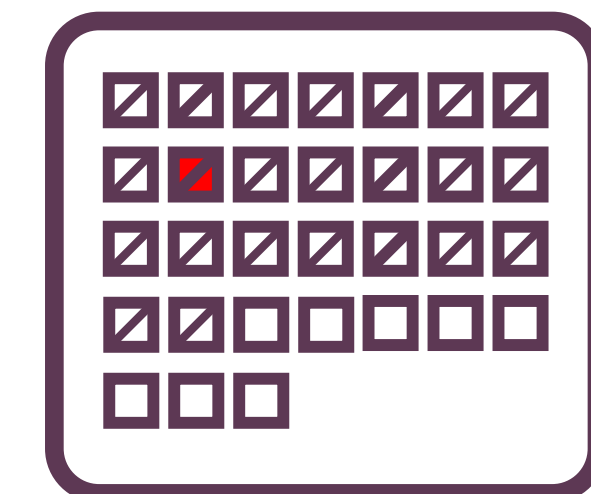
tuva tuotanto, ”yleinen niukkuus”, vaikeuttaa kaiken kaikkiaan palautumista häiriötilanteista ja nostaa riskejä yleisesti.

Laajemmin tarkasteltuna taloudelliset riskit ovat pahimmillaan silloin, kun on pystytetty sinällään toimiva tuotantoympäristö, jolla ei ole jatkuvuutta taloudellisten riskien realisoituessa. Digitointiin liittyy aina jonkinlaisia ylläpito- tai muita jatkuvia kustannuksia. Jos näihin ei varauduta taloudellisesti, voidaan joutua ajamaan alas toimiva tuotanto. Pahimmillaan joudutaan erilaisten sitoumusten vuoksi maksamaan kustannuksia tuotannon jo päätyttyä.

## Operatiiviset riskit ja laatu

### *Osapuolten roolit ja aineistotuntemus*

Jos digitointi on annettu palveluntarjoajan toteutettavaksi, liittyy sopimisvaiheeseen merkittäviä riskejä. Aina osapuolet eivät sisäistä roolejaan ja vastuitaan digitoinnin toteuttamisessa ja muun muassa siihen liittyvässä tietoaineiston siirtelyssä ja käsittelyssä. Toimeksiantaja tyypillisesti yliarvioi digitoinnin toteuttajan kyvykkyudet ja palveluntarjoaja toimeksiantoon liittyvien tavoitteiden ja taustatietojen (esimerkiksi kilpailutukseen liittyvät tavoitekuvaukset ja taustadokumentaatiot) riittävyden.





Osapuolilla ei välttämättä ole riittävää asiantuntijuutta, luottamusta tai halua molemmiin puolin ymmärrettävään kommunikaatioon toimeksiannon tarkasta sisällöstä tai vastuidensa toteuttamiseen, vaikka ne olisikin ymmärretty. Osapuolilta voi myös puuttua riittävät resurssit vastuidensa toteuttamiseen. Puutteellisesti laadittu sopimus tai kommunikatiovajeesta johtuva sen puutteellinen noudattaminen muodostaa suuren riskin koko projektin ajautumiseen väärille raiteille heti käynnistysvaiheessa.

Digitointiprojektien suurimpia riskejä ovat toimitusvaikeudet, jotka yleensä liittyvät juuri sopimuksellisten vastuiden huonoon sisäistämiseen: tietoaineistoja ei toimiteta ajoissa tuotantoon tai digitoinnin päätyttyä digitaalisia tietoaineistoja ei pystytä toimittamaan toivo-

tulla tavalla tilaajan haltuun (esimerkiksi kiinnitettyinä järjestelmätasolle). Toimitusvaikeudet voivat pysäyttää koko tuotannon tai estää tyystin valmistuneen tietoaineiston käytön.

Valtaosa operatiivisista riskeistä liittyy tietoaineiston käsittelyyn. Analogisen tietoaineiston sisällön tai fyysisen muodon (edellyttämien työskentelytapojen) tuntemus voi olla hyvin puutteellista. Digitaalisen tietoaineiston tiedostotyyppien, rakenteiden tai niihin kytkettävien metatietojen vaatimukset ovat vieraat. Yleisesti voidaan todeta, että tietoaineistojen käsittelyyn digitointiprosessin eri vaiheissa liittyy erilaisia menettelytapoja ja vaatimuksia, joiden noudattamatta jättäminen riittää kokonaisuuden kannalta tekemään turhiksi muiden vaiheiden oikeat toimenpiteet.

### **Hankinnat**

Operatiiviset riskit tulee huomioida jo niissä toimenpiteissä, jotka valmistelevat tai edeltävät varsinaista operatiivista vaihetta. Tyypillinen tällainen osa-alue ovat operatiivisen toiminnan edellyttämät hankinnat. Tyypillisesti digitointiin hankitaan skannereita ja niiden apuna käytettävää laitteistoa. Niiden kilpailutuksessa on ehdottomasti varmistettava, että hankittavan laitteiston suorituskyvykkyys riittää vaatimusten mukaiseen digitointiin sekä lisäksi tuotannolle asetettuihin määrällisiin tavoitteisiin.

Esimerkiksi kuvanlaadussa olevia puutteellisia kyvykkyksiä on lähes mahdoton paikata muuten kuin vaihtamalla laitteistoa. Kyvykkyudet tulee myös varmistaa rauhallisella testaus-





ja käyttöönottovaiheistuksella, etenkin kun kyse on kustomoiduista laitteista (siis muista kuin valmiista hyllytuotteista, esimerkiksi räätälöidyistä kamerapöytäyhdistelmistä).

### **Aineistovalmistelu ja skannaus**

Analoginen valmistelu on erittäin kriittinen vaihe varsinkin tietosisällön kannalta. Väärin järjestetty kokonaisuus, tietohävikki tai tietoa-aineiston turmeleminen valmisteluvaiheessa käytännössä mitätöi digitoinnin myöhemmät vaiheet. Myöhemmissä prosessivaiheissa ei myöskään ole mahdollisuuksia palata pohti-maan tai tarkistamaan järjestykseen liittyviä kysymyksiä, vaan valmistelun tasoon on voi-tava luottaa.

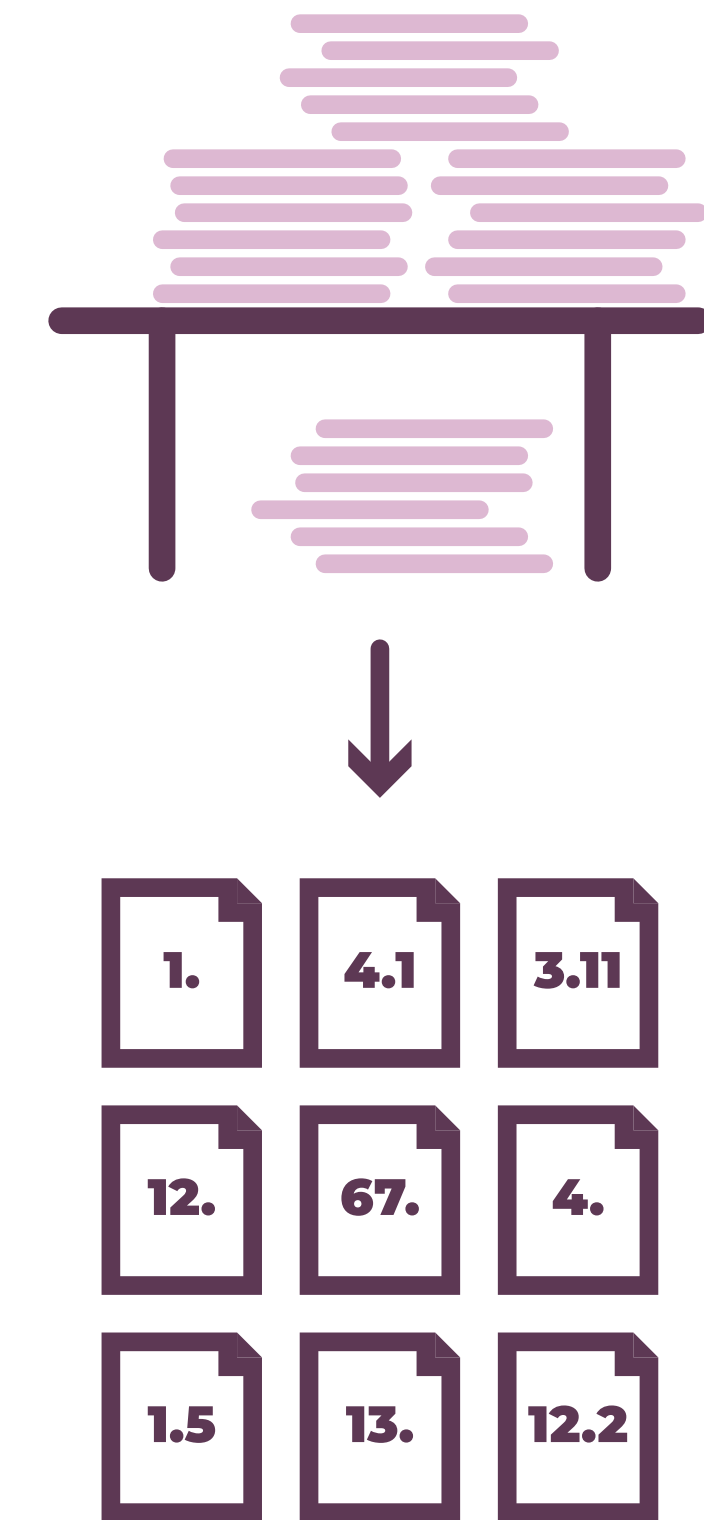
Toisaalta valmisteluvaiheessa havaitut puu-teet tulisi kuvata (käytännössä joko digitointiin sisällytettävillä tietolomakkeilla tai kuvaa-malla asia tietoa-aineiston metatiedoissa), jotta digitoitun tietoa-aineiston käyttäjällä ei herää perusteettomia epäilyjä digitoinnin yhtey-dessä tapahtuneesta tietohävikistä. Hiukan prosessista tai järjestelmäkokonaisuudesta riippuen myös tietoa-aineistoon liittyvien meta-tietojen oikeellisuuden varmistaminen voi kuulua valmisteluvaiheeseen.

Skannausvaiheeseen liittyy useita sekä kuvan-laatuun että tietosisältöön liittyviä laadullisia riskejä. Usein digitoinnin yhteydessä kuul-tava kysymys on, miten varmistetaan, ettei tietoa-aineistoa jää skannaamatta esimerkiksi tuplasyötön tai skannaajan virheen vuoksi. Todellisuudessa tämä riski on verrattain pieni,

ja sen vaikutus kokonaisuudessa on minimaalinen, koska aineiston skannaamatta jäämi-nen suuremmassa mittakaavassa on edes pääpiirteittäin asiallisesti toteutetussa digi-tointiprosessissa erittäin epätodennäköistä. Skannereiden optiset lukijat (joiden määrää on hyvä arvioida hankintavaiheessa) tunnistavat päällekkäiset lehdet luotettavasti, ja asiantun-tevat skannaajat huomaavat puutteet hyvin. Edelleen virheet ovat todennettavissa digitaalisen tietoa-aineiston prosessointi- ja siirtovai-heissa ja jälkikäteisessä laadunvalvonnassa.

### **Laatuvaatimukset**

Edellä on valmistelun ja skannauksen kohdalla jo käsitelty riskejä, jotka liittyvät tietoa-aineiston eheyden varmistamiseen. Pysyvästi säilytet-





tävän tietoaineiston todistusvoimaisuuden ja säilyvyyden turvaaminen onkin analogisen tietoaineiston hävittämiseen päättyvän digitoinnin keskeinen kysymys, joten myös siihen liittyvät riskit ovat korostetussa asemassa. Laatuun liittyvillä operatiivisilla riskeillä on selkeitä yhtymäkohtia taloudellisiin ja strategisiin riskeihin. Laadussa ilmenevät merkittävät poikkeamat ovat lyhyesti sanottuna vaikutukseltaan perustavanlaatuisia.<sup>1</sup>

Todistusvoimaisuuden ja säilyvyyden edellyttämän laadun vaatimusten – ja niiden riskien

<sup>1</sup> Vuoden 2022 jälkeen muodostuneet tietoaineistot arkistoidaan sähköisessä muodossa ja ennen vuotta 2022 laaditut tietoaineistot arkistoidaan digitoinnin jälkeen ensisijaisesti yksinomaan digitaalisessa muodossa lukuun ottamatta analogisia asiakirjoja, joiden analoginen muoto on kulttuurihistoriallisesti arvokas. Kansallisarkiston päätökset 22.12.2021, KA/15906/07.01.01.00/2021 ja 8.12.2023, KA/15927/28.02.01/2023.

– voisi luokitella liittyvän joko kuvanlaatuun, tietosisältöön, metatietoihin tai säilyvyyteen (käytettävyyteen). Vaatimuksista ja yleisesti niiden varmistamiseen liittyvästä laadunvalvonnasta ohjeistetaan Kansallisarkiston tuoreessa ohjeessa.<sup>2</sup>

Edellä käsiteltyä tietoaineiston jonkin osan skannaamatta jäämistä huomattavasti suurempi riski ovat kuvanlaadun puutteet. Tätä riskiä vähennetään optimoimalla kuvanlaadullinen suorituskyky laitteiston kalibroinnilla ja asianmukaisella huollolla sekä seuraamalla kuvanlaatua säännöllisesti ja käyttämällä luotettavia mittausvälineitä ja analysointiohjelmia. Viime kädessä niin sanottujen referenssiarvojen asettamisessa ja toteutumien seurannassa

<sup>2</sup> Arkistoitavien asiakirjojen digitointia koskevat vaatimukset ja ohjeet, 19.12.2024, KA/20919/30.01.00/2024.

on inhimillinenkin ulottuvuutensa; tarvitaan käytännölliset ohjeet siitä, sallitaanko niiden tulkinnassa jonkinlaista vaihteluväliä ja miten toimitaan tilanteissa, joissa vaadittuihin arvoihin ei päästä. Kuvanlaadun seuranta vaatii pitkällä aikavälillä merkittävää ajallista panostusta, jonka laadusta tinkimätön tuotanto on valmis siihen sijoittamaan.

Tiedostotyyppeihin, rakenteisiin ja muuhun digitaalisen tietoaineiston prosessointiin (esimerkiksi rikastamiseen erilaisilla tunnistusteknologioilla) liittyvät laadulliset riskit ovat luonnollisesti fundamentaalisia. Käsiteltävän aineiston määrä on lähes aina niin suuri, että prosessointiin liittyvät tarkistukset on ehdottomasti järkevintä toteuttaa ohjelmallisesti tietojärjestelmän automatisoiduissa prosesseissa. Tästä taas seuraa, että perustavanlaatuiset järjestelmävirheet tuottavat virheitä suuressa



mittakaavassa. Koska automatisoidun prosessin luotettavuutta ei ole mielekästä tiheästi alkaa varmistaa manuaalisesti ihmistyönä, olisi parempi vaihtoehto ajaa sama prosessointi toiseen kertaan (esimerkiksi johonkin toiseen säilytysjärjestelmään siirron yhteydessä), jolloin virheet voidaan helpommin havaita.

Prosessointiin kokonaisuudessaan liittyy lukuisia vaiheita, joiden kaikkia riskejä ei ole mahdollista käsitellä tässä yhteydessä. Lyhyesti voidaan todeta, että huomiota tulisi kiinnittää digitointiprosessia koskevien metatietojen oikeelliseen keräämiseen sekä siihen, että pakattua tallettiedostoa ei saisi tallentaa missään tapauksessa kahta kertaa enempää ja lähtökohtaisesti vain kerran.<sup>3</sup>

<sup>3</sup> Tiedostojen, tiedostorakenteiden ja digitointiproessin vaatimusten osalta tulee tutustua Kansallisarkiston ohjeen luvussa 4 esitettyihin vaatimuksiin: Digitointia koskevat vaatimukset.

Tietoaineiston tietosisältö pitäisi varmistaa niin, että jokaisesta yksittäisestä kuvatiedostosta pitää silmämääräisellä tarkastelulla olla saatavissa kaikki sama informaatio kuin sen analogisesta vastineesta (Kansallisarkisto 2023). Tämä on autenttisuuden, todistusvoimaisuuden ja eheyden edellytys. Tämä sisällön laadunvalvonta, validointi, voidaan toteuttaa monin tavoin, mutta oleellista on, että sitä tehdään systemaattisesti ja jatkuvasti.

Virhetyypit ovat yleensä samankaltaisia ja toistuvia – raitoja kuvassa, peittyneitä tai muusta syystä ei-alkuperäistä vastaavassa muodossa luettavissa olevaa sisältöä – ja niitä voidaan tunnistaa ihmistyönä tarkastelemalla silmämääräisesti, koneellisin tunnistusmenetelmin tai niiden yhdistelminä.

Koneellisen tunnistuksen haasteena on luotettavuus (erityisesti, jos tietoaineistot vaihtuvat jatkuvasti digitoinnissa), ja ihmistyön haasteena on epätasaisuus.

Jälkimmäisessä tapauksessa kyse on siis tulkinnoista, joissa ei useampien henkilöiden toteuttamassa validoinnissa voida koskaan saavuttaa täydellistä yhdenmukaisuutta. Esimerkiksi miten suhtaudutaan muun kuin varsinaisen tietosisällön, kuten vakiolomaketehtojen, tyhjien lomakesivujen, vesileimojen tai sivunumeroiden marginaalisiin laatupoikkeamiin? Johdonmukaisuuden saavuttaminen edellyttää jatkuvaa kouluttamista ja validoinnin johtamista.

### **Tuotanto-olosuhteet**

Digitointituotannon olosuhteisiin ja puitteisiin liittyy mittava määrä operatiivisia riskejä, joihin voidaan tässä yhteydessä vain lyhyesti viitata. Tuotantotyötilojen tulee tietenkin olla sellaisia, että työtä voidaan tehdä turvallisesti. Laitteistojen ja teknologioiden käyttöikään ei kohdistu tarpeetonta riskiä. Tietoaineiston säilyvyys ja digitoitavuus ei saa olosuhteiden takia vaarantua. Esimerkiksi tilojen huono ilmanvaihto tai -laatu ja huono siivoaminen aiheuttavat todellisia riskejä niin henkilöstölle, tietoaineistoille kuin laitteille: pöly ja lika kulkeutuvat laitteisiin niitä vaurioittaen ja laatua heikentäen. Käytettävät teknologiat ja ohjelmistot tai lisenssit luovat osaamisen, soveltuvuuden tai kehittämisen riskejä. Kel-



paako esimerkiksi valmis skannausohjelmisto sellaisenaan tuotannon tarpeisiin, vai tulisiko siihen tehdä joihinkin ominaisuuksiin liittyviä muokkauksia?

Merkittävä riski liittyy erilaisiin ylläpitopalveluihin tai jopa koko digitoinnin tietojärjestelmän vakauteen, laatuun ja riippuvuuksiin. Laitteistoa hankittaessa ei välttämättä kiinnitetä huomiota siihen, onko sen ylläpitoon saatavissa ammattitaitoisia, saavutettavia ja mielekkään hintaisia palveluita. Jos järjestelmäkokonaisuuteen on rakennettu palveluntarjoajien ylläpitämiä ominaisuuksia tai käyttöoikeusrajoituksia, voidaan päätyä kehittämisen kannalta hankalaan vendor lock-asetelmaan. Prosessoinnin kannalta keskei-

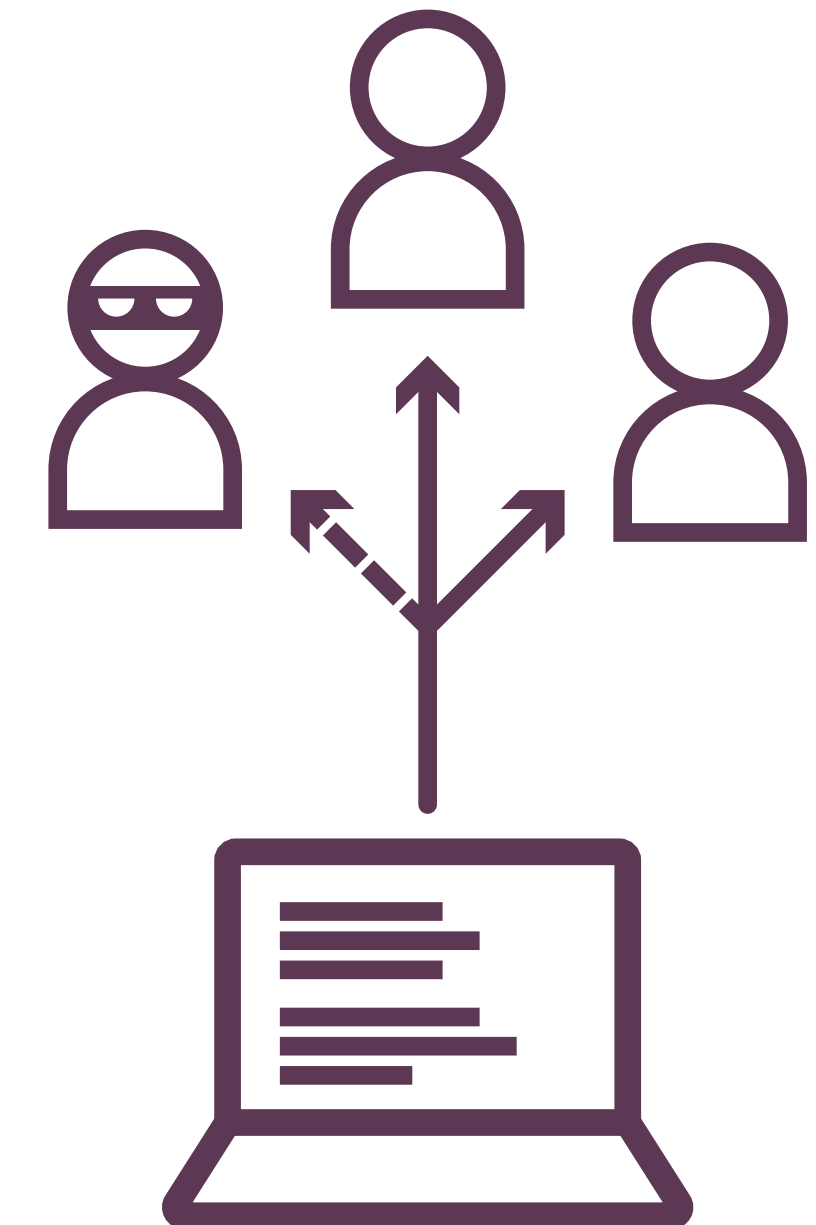
siä riskejä liittyy kapasiteetin riittävyteen, modulaarisuuteen ja optimoitavuuteen. Tuotantonopeuteen liittyvä merkittävä riski on tiedonsiirron vakaus ja nopeus. Tämä korostuu entisestään silloin, jos digitointiprosessi tapahtuu pääasiassa virtualisoiduissa verkkoratkaisuissa.

Tuotantoympäristön ja digitointiprosessin merkityksen riskienhallinnassa voisi tiivistää esimerkiksi näin: digitointiprosessi ja tuotanto tulee rakentaa niin, että toimintamalli on dokumentoitu ja toimenpiteet ovat jälkikäteen rekonstruoitavissa, ja mieluiten niin, että keskeiset tiedot tallennetaan järjestelmälähtöisesti ja automaattisesti (lokitiedot, elinkaaritiedot, tietojärjestelmäkuvaukset ym.).

### Tietoturva ja tietosuoja

Merkittäviä riskejä liittyy tietoturvan ja tietosuojan vaatimuksiin.<sup>4</sup> Riskit kohoavat entisestään, jos tietoa sisältää luottamuksellista tai käyttörajoitettua tietoa tai henkilötietoja. Tietosuoja-asetuksen mukainen vaikutustenarviointi ja erilaiset auditoinnit sekä tietosuojamenettelyiden ulkopuoliset selvitykset ovat tapoja, joilla vahvistetaan sekä itse turvallisuutta että henkilöstön tietoisuutta vaatimuksista.

<sup>4</sup> Tiedonhallintalain mukaan ”Tiedonhallintayksikön on seurattava toimintaympäristönsä tietoturvallisuuden tilaa ja varmistettava tietoa sisältävien tietojärjestelmien tietoturvallisuus koko niiden elinkaaren ajan Tiedonhallintayksikön on selvitettävä olennaiset tietojenkäsittelyyn kohdistuvat riskit ja mitoitettava tietoturvaluustoimenpiteet riskiarvioinnin mukaisesti”. Laki julkisen hallinnon tiedonhallinnasta 906/2019, § 13.





Menettelytapojen riskit tulee tunnistaa ja määrittää niille hallintakeinot sekä tunnistaa ja hyväksyä todennäköiset jäännösriskit. Tietoturvapoikkeamien menettelytavat tulee määrittellä ja testata. Tietoturvariskeissä on yhtäältä kyse lainsäädännön asettamista vaatimuksista – henkilöiden oikeusturvasta ja organisaatioiden turvallisuudesta – mutta myös digitoinnin sopimusosapuolten toisilleen asettamista vaatimuksista ja lainsäädännön käytännön tulkinnoista, jotka yleensä liittyvät aineiston omistajuuteen ja siihen, mitä tietoaineiston käytöstä ja säilyttämisestä on sovittu.

Parhaiten riskejä voidaan vähentää rakentamalla mahdollisimman turvallinen ympäristö suurimman haitan riskien mukaisesti. Digitointi voidaan toteuttaa muista verkoista

eriytettynä, tiukan kulunvalvonnan tiloissa. Henkilöstöriskit voidaan minimoida turva- ja taustaselvitysprosessien avulla. Tuloksena voi olla kustannuksiltaan ja olosuhteiltaan raskas ratkaisu, joka voi samalla lisätä taloudellisia ja operatiivisia riskejä.

### **Henkilöriskit**

Lopuksi on syytä parilla sanalla käsitellä henkilöstöön liittyviä riskejä. Vaikka digitoinnin suunnittelussa ja julkisessa keskustelussa korostuvat usein teknologiset ja tekniset yksityiskohdat, on asiantunteva ja riittävä henkilöstö hyvinkin mahdollisesti digitoinnin tärkein voimavara, joten siitä tulee pitää pitkäjänteisesti huolta.

Tyypillinen riski on henkilöstön vaihtuvuus: palkan, työviihtyvyyden ja yleisten olosuhteiden vuoksi vaihtuvuus voi muodostua suureksi, jolloin osaamista ja luotettavuutta ei kerry organisaatioon. Lisäksi joudutaan jatkuvasti osoittamaan paljon resursseja rekrytointiin ja perehdytykseen. Vaihtuvuus kuormittaa henkilöstöä merkittävästi ja nostaa ainakin kaikkia operatiivisia riskejä. Riskiä voidaan vähentää rakentamalla urapolkuja ja osaamisen kehittämistä, suosimalla tehtäväkiertoa ja varmistamalla oikeudenmukainen ja läpinäkyvä palkkamalli.

Muita henkilöstöön liittyviä riskejä ovat liialliset riippuvuudet ja asiantuntijakapeikot: kriittinen osaaminen voi kohdentua liian harvoille henkilöille, mikä lisää riskejä varsinkin





poikkeusolosuhteissa. Riskien pienentäminen tapahtuu jakamalla rohkeasti vastuuta laajemmalle ja varmistamalla varajärjestelyt kaikissa olosuhteissa. Toinen tyypillinen riski on liian vähäinen henkilöstö. Riskiä voidaan pienentää tekemällä rekrytointeja vaiheittain samalla kun rauhassa selvitetään, millaisia resursseja suunniteltu tuotanto tosiasiallisesti esimerkiksi tietoaaineistoon ja laitemäärään suhteutettuna edellyttää.

Henkilöstöön liittyvät riskit monella tapaa tiivistyvät myös sen ensimmäiseen vaiheeseen eli rekrytointiin. Usein kiireessä ja suunnittelemta toteutettavan rekrytointiprosessin

riskit voivat olla vaikeasti korjattavissa jälkikäteen, jos rekrytoinnissa päädytään epäsopiviin valintoihin. Rekrytointi kannattaakin toteuttaa vaiheittain, osallistaa siihen useampia henkilöitä johdosta suorittavaan portaaseen ja noudattaa saatavilla olevia henkilöstöasiantuntijoiden laatimia ohjeita.

## Lähteet

Arkistoitavien asiakirjojen digitointia koskevat vaatimukset ja ohjeet (A/20919/30.01.00/2024). Kansallisarkisto 2024. PDF-dokumentti. Saatavissa: <https://kansallisarkisto.fi/documents/141232930/151500284/Arkistoitavien+asiakirjojen+digitointia+koskevat+vaatimukset+ja+ohjeet.pdf/9ea50bdd-ad59-300c-a6ce-1c7bbe5f9d3f/Arkistoitavien+asiakirjojen+digitointia+koskevat+vaatimukset+ja+ohjeet.pdf?t=1734607700682> [viitattu 24.1.2025].

Julkishallinnon pysyvästi säilytettävien digitoitujen ja digitoitavien asiakirjojen säilytysmuoto ja analogisen säilytysmuodon kulttuurihistoriallisen arvon arviointitietojen pysyvä säilytys (KA/15927/28.02.01/2023). Kansallisarkisto 2023. PDF-dokumentti. Saatavissa: <https://kansallisarkisto.fi/documents/141232930/149322637/S%C3%A4ilytysmuoto+yleisp%C3%A4%C3%A4t%C3%B6s.pdf/5fe6ff04-22c8-dca6-8f4e-19d4f8d4d924/S%C3%A4ilytysmuoto+yleisp%C3%A4%C3%A4t%C3%B6s.pdf?t=1703075500567> [viitattu 24.1.2025].

Laki julkisen hallinnon tiedonhallinnasta 906/2019.

Määräys arkistoitavien asiakirjojen muodosta (KA/15906/07.01.01.00/2021). Kansallisarkisto 2021. PDF-dokumentti. Saatavissa: <https://kansallisarkisto.fi/documents/141232930/151497947/M%C3%A4%C3%A4r%C3%A4ys+arkistoitavien+asiakirjojen+muodosta+2021.pdf/> [viitattu 23.1.2025]

SFS-ISO 31000:2018. Riskienhallinta. Ohjeet. Sääskilahti, T. & Mustonen, E. 2023. Valtionvarainministeriön Riskienhallinnan käsikirja valtionhallinnon toimijoille. Valtiovarainministeriön julkaisuja 2023:54.



Johanna Lilja &  
Anna-Elina Ukkonen

# KESTÄVYYS- NÄKÖKULMA KANSALLIS- KIRJASTON DIGITOINTI- OHJELMASSA

Kansalliskirjasto on sitoutunut edistämään kestävä kehitystä ohjelmassaan *Rohkeasti kestävään muutokseen – Kestävä kehitys Kansalliskirjastossa 2021–2030*. Ohjelma nojaa YK:n Agenda 2030:n tavoitteisiin, joissa kestävyttä tarkastellaan kolmesta näkökulmasta: ekologinen, taloudellinen ja sosiaalinen kestävyys. Kulttuuriperintöorganisaationa Kansalliskirjasto haluaa nostaa teemaksi myös kulttuurisen kestävyden, onhan tehtävänäme säilyttää omana aikamme julkaistua tietoa ja kulttuuria ja välittää se käytettäväksi tuleville vuosisadoille (Eklund ym. 2022, 5, 12–13). Tätä tukeakseen Kansalliskirjasto on allekirjoittanut Climate Heritage Networkin Culture at the heart of

Climate action -aloitteen, jossa organisaatiot ympäri maailman peräänkuuluttavat kulttuuriperintö-, taide- ja luovien alojen ottamista mukaan ilmastopolitiikkaan.

Ekologisen kestävyden osalta Kansalliskirjasto on sitoutunut emo-organisaationsa Helsingin yliopiston tavoitteeseen olla hiilineutraali vuonna 2030. Digitaalisuuden edistäminen sekä tukee että heikentää tätä tavoitetta. Digitaalinen pääsy aineistoihin ja palveluihin lisää aineistojen vaikuttavuutta tuomalla ne kaikkien ulottuville, mutta digitaalisen aineiston tuottaminen lisää päästöjä sekä tuotantoprosessin aikana että kasvattamalla palvelimilla säilytettävän datan määrää.



## KESTÄVÄN KEHITYKSEN TAVOITTEET



**Kuva 1.** YK:n Agenda 2030 sisältää 17 tavoitetta, joista monia voidaan toteuttaa kulttuuriperintöä digitoimalla. <https://mediabank.finland.fi/l/DXQf8r7DHRv2/f/kzBQ> Ulkoministeriön mediapankki (vapaa käyttö)

Digitalisaation jarruttaminen ekologisista syistä ei kuitenkaan ole perusteltua, sillä se heikentäisi tuntuvasti taloudellista ja sosiaalista kestävyttä. Siksi on olennaista pohtia, miten digitaalisia palveluita ja sisältöjä tuotetaan mahdollisimman vastuullisesti eli kestävä kehitys huomioiden.

Kansalliskirjaston tuore digitointiohjelma 2025–2028 pyrkii huomioimaan Agenda 2030 -tavoitteet aineiston valinnassa, digitointiprosessin eri vaiheissa sekä aineiston käytön ja jatkokäytön suunnittelussa (Ukkonen & Lilja 2025, 1–6). Agenda 2030 tavoitteet on esitelty kuvassa 1.

Kestävyys toteutumisen seuraaminen edellyttää mittareita. Kestävyttä kuvaavia mittareita on jo olemassa, mutta niitä käytetään vaihtelevasti. Tarvitaan kokonaisvaltaisempia

mittaamisen lähestymistapoja, jotka perustuvat kansainvälisiin, tietopohjaisesti tuotettuihin mittaristoihin (Kestävän kehityksen toimikunnan toimenpidesuunnitelma vuosille 2025–2027, versio 11.2.2025).

Tässä artikkelissa kerrotaan, mitä kestävyys merkitsee digitointiprosessin kussakin vaiheessa Kansalliskirjastossa. Kestävän kehityksen periaatteet on tarpeen ulottaa kaikkeen digitaaliseen kulttuuriperintöön. Eri kulttuuriperintöorganisaatioilla on kuitenkin erilaisia aineistoihin, prosesseihin, järjestelmiin ja lainsäädäntöön perustuvia ratkaisuja, joten kaikki Kansalliskirjaston linjaukset eivät sellaisenaan sovi museoille, arkistoille tai perinnearkistoille. Toivottavasti artikkelimme toimii kuitenkin keskustelunavauksena digitaalisen kulttuuriperinnön tallentajille ja sen käyttäjille.



Kulttuuriperintöorganisaatioiden toimintaympäristö on vahvassa ja nopeassa muutoksessa. Tekoälyn kehittyminen tuo niin uusia mahdollisuuksia kuin uusia uhkiakin (Litvinets 2024). Yhteiskunnan polarisaatio vaikuttaa digitaalisten palvelujen käyttäjien tarpeisiin ja odotuksiin sekä myös omana aikanamme syntyvään kulttuuriperintöön.

Avoimuus on 2000-luvulla noussut keskeiseksi arvoksi, mutta joudutaanko siitä tulevaisuudessa tinkimään esimerkiksi kansallisen turvallisuuden vuoksi? Ehkä tämäkin artikkeli täytyy jo lähivuosina kirjoittaa uusiksi. Siihen asti edetkäämme kuitenkin niin kestävästi kuin nykyisen tietomme varassa pystymme!

### **Kestävyyisperiaatteet aineistojen valinnassa ja digitointiprojektien rahoitus**

Kestävyyden näkökulmasta on syytä välttää tekemästä asioita vain koska voimme, vaan toiminnan laajemmat vaikutukset on huomioitava. Kulttuuriperinnön digitointi on tärkeää muun muassa hauraiden ja ainutlaatuisten aineistojen turvaamiseksi sekä aineiston paremman saavutettavuuden ja käytettävyyden mahdollistamiseksi, mutta digitoitavat aineistot on syytä valita perustellusti. Kaikkea ei voida eikä ole tarpeen digitoida.

Kansalliskirjastossa fyysinen kappale säilytetään digitoinnin jälkeen, joten digitointi ei vähennä varastoinnin hiilijalanjälkeä tai säilytyskuluja. Digitointi vaatii myös taloudellisia panostuksia. Näin ollen ei ole aivan sama, mitä

digitoidaan. Kestävyyden eri ulottuvuudet, ekologinen, taloudellinen sekä sosiaalinen ja kulttuurinen kestävyys, painottuvat digitoitavan aineiston valinnassa. Onko aineistolle aktiivista, ajankohtaista yleisöä tai voidaanko digitoida lisää aineiston käytettävyyttä? Onko alkuperäinen, fyysinen aineisto mahdollisesti saavutettavissa? Voidaanko aineistoa avata laajasti yleisölle? Mitä yhteiskunnallista merkitystä digitoitavalla aineistolla on?

Ekologisen ja taloudellisen kestävyden näkökulmasta digitoidut sisällöt kannattaa rajata uniikkeihin aineistoihin ja sellaisiin aineistoihin, joita ei ole digitoitu muualla, varmistaa niiden näkyvyys myös erilaisia aineistoja kokoavissa kulttuuriperintöportaaleissa, kuten Finna.fi:ssä ja Europeanassa. (Ukkonen & Lilja 2025, 5)

Digitointivalinnoilla voi tukea sosiaalista kestävyttä, jos aineistoja valitaan siten, että ne edustavat laajasti yhteiskunnan eri ryhmiä. Erilaisten vähemmistöjen ja etnisten ryhmien aineistojen digitointi vähentää globaalia eriarvoisuutta ja tuo monimuotoisia sisältöjä kaikkien ulottuville (Pinto 2018, 23–24). Kansalliskirjaston digitointiohjelmassa esimerkiksi lehtiaineisto edustaa kaikkia suomalaisia, inarin- ja koltansaamenkieliset aineistot taas osaa saamelaisvähemmistöstä (Ukkonen & Lilja 2025, 11–13).

Kulttuurista kestävyttä uhkaavat sodat ja poliittiset kriisit, jotka tekevät kulttuuriperinnöstä haavoittuvaa ja siten uhkaavat pyyhkiä pois kansojen tai etnisten ryhmien menneisyyttä ja kulttuuria. Kansalliskirjaston digitointiohjelman yksi kokonaisuus on suunniteltu turvaamaan ukrainalaisen kulttuuri-



perinnön säilymistä digitaalisena. Hanke tuo digitoitavaksi Kansalliskirjaston kokoelmissa säilynyttä Hersonin alueen aineistoa. (Ukkonen & Lilja 2025, 16.)

Digitoitu kulttuuriperintö on merkittävä tutkimuksen ja opetuksen tietovaranto, jolla voidaan tukea YK:n Agenda 2030:n tavoitetta 4 hyvästä koulutuksesta sekä kestävästä teollisuudesta, innovaatioista ja infrastruktuureista. Eheidän kokonaisuuksien digitointi tukee kattavan tiedon tarjontaa ja luo edellytyksiä tutkimukselle ja opetukselle. (Ukkonen & Lilja 2025, 4.)

Taloudellisen kestävyuden kannalta on tärkeää, että digitoitava aineisto tuottaa mahdollisimman suuren hyödyn yhteiskunnalle. Digitoidun aineiston käyttäjiä ovat tutkijat ja historian harrastajat, joista sukututkijat ovat suurin ryhmä, mutta myös koulut, järjestöt ja yritykset voivat hyödyntää digitaalista kulttuu-

riperintöä. Sitä voidaan käyttää myös vaikkapa luoviin aloihin liittyvässä toiminnassa. Kansalliskirjaston digitointiohjelmaan 2025–2028 nostetut aineistot valikoituivat yhteistyöllä. Esitykset tulivat kirjaston asiantuntijoilta, jotka tuntevat niin aineistot kuin niihin kohdistuvat asiakastarpeetkin. Nämä aineistoehdotukset käsiteltiin ja priorisoitiin eri tieteenaloja edustavien neuvottelukuntien työpajassa. Myös yleisön toiveita pyritään ottamaan huomioon. (Ukkonen & Lilja 2025, 10.)

Varsinaisen digitointiprojektin toteuttamiseen tarvitaan taloudellisia resursseja. Kansalliskirjasto digitoi perusrahalta enimmäkseen sanomalehtiä, jotka ovat hyvin käytettyä aineistoa, sekä pelastusdigitointina äänitteitä, joita halutaan suojata käytössä kulumiselta. Muun aineiston digitoinnissa on merkittävä rooli täydentävällä rahoituksella, kuten julkaisijoiden, säätiöiden tai yksityishenkilöiden tuella.

Lahjoitukset ja digitointikumppanuudet, joissa digitoitu aineisto tarjotaan avoimeen tai kumppanin sisäiseen käyttöön, vahvistavat paitsi taloudellista myös ekologista kestävyttä. Keskittämällä digitointi, aineiston käyttöjärjestelmä ja pitkäaikaissäilytys Kansalliskirjastoon kumppanit säästävät omissa kustannuksissaan. Lisäksi vältetään aineistojen moninkertaiselta tallentamiselta.

Täydentävän rahoituksen hankintaa suunnataan digitointiohjelmassa määriteltyihin aineistoihin, jotka on valittu kestävä kehityksen periaatteiden mukaisesti. Varainhankinta ja kumppanuusdigitoinnit suunnitellaan siten, että ne tukevat niin kumppanin kuin Kansalliskirjastonkin tavoitteita, ja rahoitukselta hankkeista viestitään lahjoittajan tai kumppanin toiveita kunnioittaen. Kansalliskirjaston digitointiohjelmassa kerrotaan läpinäkyvästi, mitä hankkeita Kansalliskirjasto

voi toteuttaa perusrahoituksella ja mihin se tarvitsee täydentävää rahoitusta. (Ukkonen & Lilja 2025, 12.)

## Kuvailun merkitys kestävälle kehitykselle

Digitoidun aineiston käyttäjä ei useinkaan tule ajatelleeksi, kuinka paljon kuvailu eli metadata työ helpottaa aineiston hakua ja käyttöä, eikä sitä, kuinka paljon kuvailun eteen on tehty työtä ennen aineiston julkaisemista. Kuvailulla on laajamittainen vaikutus kaikista kestävyuden näkökulmista tarkasteltuna.

Ekologista kestävyttä vahvistaa osuvien hakutulosten saaminen. Laadukas kuvailu vähentää turhien hakujen ja väärin latausten määrää ja sitä kautta energian kulutusta. Jos kuvailua ei tehtäisi, haut kohdistuisivat



miljoonien sivujen sisältöön, ja mitä yleisempi hakutermin olisi käytössä, sitä pitempi olisi hakutulostila ja sitä enemmän tulisi myös turhia latauksia. (Ukkonen & Lilja 2025, 5.)

Taloudellisen kestävyden näkökulmasta kuvailu tuottaa ajansäästöä käyttäjälle sekä laajempia käyttömahdollisuuksia aineistoille. Kuvailu auttaa kohdentamaan haun tiettyyn elementtiin, esimerkiksi tekijään. Lisäksi kuvailutiedot kertovat, miten aineistoa voi käyttää oikeuksien tai teknisten ominaisuuksien näkökulmasta. Kuvailu mahdollistaa viitetaustoinnot, jolloin käyttäjän on helppo siteerata aineistoa omassa tekstissään. Kun kuvailutieto on ladattavissa datana aineiston mukana, voi käyttäjä helposti lajitella aineistomassoina. (Ukkonen & Lilja 2025, 4–5.)

Kuvailussa on elementtejä, jotka auttavat hahmottamaan julkaisun ajallista tai paikallista tai jopa poliittista kontekstia. Tämä on tärkeää varsinkin silloin, kun historiallinen aineisto poikkeaa vahvasti nykyisestä maailmankuvasta – onhan vanhoissa teksteissä ja kuvissa toisinaan ajatuksia, kuvia ja ilmaisuja, jotka voidaan kokea loukkaavina. Julkaisupaikka, -aika ja julkaisija tai julkaisusarja auttavat käyttäjää näkemään, missä kontekstissa teos on tuotettu. Digi.kansalliskirjasto.fi-palvelussa on lisätty sanomalehtien metatietoihin lehden puoluekanta tai poliittinen suuntaus silloin, kun se on ollut löydettävissä lehdestä itseltään tai luotettavista hakuteoksista. Tämäkin voi auttaa lukijaa tulkitsemaan sisältöjä.

Kirjojen ja muiden yksittäisten teosten sisälönkuvailulla niiden sisältö voidaan linkittää

**AJAN SUUNTA**  
N:o 278  
Sunnuntai 15. p.  
PÄÄTOIMITTAJA: R. G. KALLIA  
JULKAISIJA: KANSAN SANAN OSAYHTIÖ.  
TOIMITUSNEUVOS: PAAVO RAUTALA, BRUNO A. SALMIÄLÄ, V.  
Isänmaallisen kansanliikkeen pää-äänenkannattaja

**MINISTERI PAASIKIVI AVUSTAA**  
**palaa huomenna Helsinkiin selostamaan Neuvostoliiton ehdotuksia.**  
Moskovassa keskusteltiin eilen kahteen otteeseen ja neuvotteluja jatketaan jonkin päivän kuluttua.

**Englantilainen taidemaalari "Royal Oak" upotettiin 1,000-lukuisesta sukellusvenästä hukkuu useita viikkoja.**

Nimekkeen tiedot	
Päänimeke	Ajan Suunta
ISSN	fk10009
Ulkoasu	A2
Alkupyvm	02.12.1932
Loppupyvm	03.10.1944
Kustantaja	IKL:n Sanomalehdistön Kannatusyhdistys
Julkaisupaikka	Helsinki 02.12.1932
Näytenumerot	Ajan Suunta 02.12.1932 Ajan Suunta 13.12.1932 Ajan Suunta 02.12.1932 Ajan Suunta 02.12.1932
Edeltäjät	Ajan Sana
Lehden tiedot Suomen kansallisbibliografiassa.	
Puoluekanta	IKL:n äänenkannattaja <sup>1</sup>

**Kuva 2.** Ajan suunta -lehti 15.10.1939 ja sen kuvailutiedot, joista ilmenee, että lehti on IKL:n (Isänmaallinen kansanliike) äänenkannattaja. Kuvakaappaus digi.kansalliskirjasto.fi-palvelusta. <https://digi.kansalliskirjasto.fi/sanomalehti/binding/1957810?page=1&marc=true>

nykypäivään ja ilmaista ajallemme korrektein termein, mitä teos käsittelee. Tosin on hyvä huomioida, että kuvailukin tehdään oman aikansa termein ja käsityksin, jotka tulevaisuudessa voidaan kokea väärinä tai loukkaavina. Siksi on tärkeää, että myös kuvailun ajankohta tallentuu ja on käyttäjälle nähtävissä.

Jotta kuvailu täyttäisi tehtävänsä, sen on oltava virheetöntä ja vastattava käytettyjä kuvailustandardeja. On tärkeää, että kuvailutieto on koneluettavaa ja yhteentoimivaa järjestelmästä riippumatta. Parhailaan on meneillään siirtymä MARC21-formaatista linkitettyyn dataan (Hyvärinen ym. 2023, 2–8). On myös tärkeää, että metatiedot julkaistaan



avoimella lisensillä, jotta ne ovat monipuolisesti hyödynnettävissä.

Kansalliskirjaston digitointiohjelmassa laadukas kuvailu sisältyy kaikkiin hankkeisiin. Kuvailutiedot ovat käytettävissä sekä Kansalliskirjaston hakupalvelussa [kansalliskirjasto.finna.fi](https://kansalliskirjasto.finna.fi) että [digi.kansalliskirjasto.fi](https://digi.kansalliskirjasto.fi)-palvelussa, jossa asiakas pääsee tekemään hakuja myös tekstisisältöihin. Parhaillaan on meneillään hanke, jossa tuotetaan käytäntöjä kokoelmien kuvailuun. Kokoelmakuvaus tuottaa käyttäjälle listauksen kokoelman sisällöstä ja lisäksi lyhyen taustoituksen siitä, miten kokoelma on syntynyt. Tämän on tarkoitus helpottaa aineiston tutkimuskäyttöä sekä kokoelmien lataamista datana. (Ukkonen & Lilja 2025, 6, 23.)

Kuvailu on tarkkuutta vaativaa ja työvoimavaltaista. Tekoäly on kuitenkin jo vahvasti

tulossa kuvailun tueksi. Sisällönkuvailua voi jo nyt tehdä Annif-robotilla, joka ehdottaa analysoimilleen teksteille asiasanoja käyttäjän valitsemasta sanastosta (Inkinen ym. 2023). Bibliografisen kuvailun osalta tekoälyratkaisuja vasta kehitetään ja testataan. Kuvailuosaamisen ja aineistoihin liittyvän asiantuntemuksen tarve ei kuitenkaan ole katoamassa, sillä tekoälyäkin pitää kehittää, ymmärtää ja valvoa. (Suominen 2024.)

## Digitoinnin ympäristövaikutukset

Digitointiprosessi ja digitoitavan aineiston käyttö ovat vahvasti riippuvaisia tieto- ja viestintäteknikasta (ICT, information and communication technologies) aina skannauksesta pitkäaikaissäilytyksen infrastruktuuriin ja lop-

pukäyttäjien laitteisiin. Näin ollen vastuullisilla ICT-ratkaisuilla on suuri merkitys. Riippuvuus ICT:stä on haaste ekologisen kestävyuden kannalta, mutta harkituilla teknisillä valinnoilla voidaan vaikuttaa digitointiprosessin synnyttämään ympäristökuormaan.

Merkittävin laitteisiin liittyvä kestävyysshaaste on niiden lyhenevä elinkaari, koska laitteita uusitaan yhä kiihtyvällä tahdilla (GreenICT 2022). Kansalliskirjasto on huomionnut laitevalinnoissaan esimerkiksi skannereiden huollettavuuden ja siten niiden elinkaaren pidentämisen. Olemassa olevan skannerikapasiteetin käyttöä optimoidaan ja laitteistoa hyödynnetään mahdollisimman monipuolisesti.

Toisaalta uudempi teknologia voi olla joskus perusteltua. Uudemmissa skannereissa esimerkiksi sähkönkulutus voi olla optimoidum-

paa kuin vanhemmissa laitteissa tai viimeisin palvelinteknologia voi olla edeltäjiään ekologisempaa. Yleisesti IT-laitteita käytetään Kansalliskirjastossa mahdollisimman kauan ja ne hankitaan ekologisin periaattein sekä kierrätetään Helsingin yliopiston IT-laitteiden kierrätyspalvelun kierrätysohjeistuksia noudattaen.

Digitointiprosessi tuottaa dataa, ja datalla on aina koti. Palvelimien osalta ekologista kestävyttä voidaan tukea esimerkiksi siten, että vain välttämätön määrä palvelimia on käytössä ja tarvittaessa vain valmiustilassa. Keskitetyillä palvelinratkaisuilla voidaan myös saavuttaa ympäristökuormaa keventäviä vaikutuksia. Pilvipalveluissa on otettava selvää pilvipalvelun tuottamisperiaatteista (esim. sähkönkulutus) ja arvioitava, soveltuuko käytön tai tilan mukaan laskutus omaan käyttö-



tilanteeseen. Kaikissa palvelinratkaisuihin on myös huomiotava varmuuskopiointi, tietoturva ja tietosuojat.

Digitaalisen ympäristökuorman vaikutukset kasvavat kopioiden määrän lisääntyessä, joten digitoidun aineiston tarpeettomien kopioiden määrää on syytä tarkastella (Pendergrass ym. 2019, 191; Paschalidou ym. 2022, 1075). Kestävyyden kannalta olennaisen kopioiden määrän rajaamisen suhteen tasapainoillaan sen seikan kanssa, että vastuullinen digitointi on myös digitaalisen aineiston ja jo tehdyn työn sekä aineiston käytettävyyden turvaamista. Digitaalisen sisällön noutamisen vasteaika pitkäaikaissäilytyksestä on jopa tunteja, eikä se sovellu siten verkkopalvelussa käytettäväksi, joten Kansalliskirjastossa on vielä tarvetta omille kopioille.

Digitoidut alkuperäiskappaleet, joita ammattikielessä kutsutaan master-versioiksi, ovat korkeimman laadun säilytyskappaleita ja siten myös tallennustilaa vaativia. Onkin mielekästä säilyttää master-tiedostot pienemmän ympäristövaikutuksen arkistosäilytyksessä ja tehdä käyttöversioista keveämpiä. Kansalliskirjaston digi.kansalliskirjasto.fi-julkaisujärjestelmässä aineisto on käyttäjälle tarjolla JPEG- ja PDF-formaateissa raskaamman TIFF-master-version sijaan. Käyttökuvat vastaavat noin 10:tä prosenttia TIFF:ien koosta.

Kestävän kehityksen kannalta digitointi on valintoja. Esimerkiksi värigitointi vaatii huomattavasti enemmän tallennustilaa ja sen käsittely on hitaampaa kuin harmaasävyisenä digitointi. Sanomalehtiaineistot 1980-luvulle saakka digitoidaan Kansalliskirjastossa harmaasävyisenä, sillä värien merkityksellinen

käyttö lisääntyi merkittävästi vasta myöhemmin. Digitoidun aineiston pienempi resoluutio vähentää levytilan tarvetta. Resoluution tuplaantuessa tilantarve nelinkertaistuu. Digi.kansalliskirjasto.fi-palvelun lehtiaineistojen käyttöversioissa on master-kuvia pienempi, mutta käytön kannalta riittävä resoluutio 200 dpi. Kirjoissa ja pienpainatteissa käyttöversion resoluutio on 300 dpi. Valintoja tehdessä on syytä muistaa, että on kuitenkin tärkeää digitoida kerralla riittävän korkeaa laatua, jotta uudelleendigitoinnille ei tule tarvetta.

Digitoinnin laatua voidaan parantaa tekoälypohjaisilla ratkaisuilla. Kansalliskirjaston digitointiprosessissa tekoäly avustaa eri vaiheissa, esimerkiksi jälkikäsitteilyn palstoituksessa. Toimiva palstoitus taas parantaa tekstintunnistuksen laatua. Kansalliskirjastossa sanomalehtiaineiston jälkikäsitteily on

täysin automatisoitu ja tätä mahdollisuutta voidaan laajentaa tietyille muillekin aineistotyypeille. Tekoälyn käyttö mahdollistaa ihmistyön suuntaamisen niihin työvaiheisiin, joita tekoäly ei voi suorittaa, kuten skannaamiseen ja logistiikkaan.

Kestävyyden kannalta tekoälyn hyödyntämisessä on punnittava hyödyt suhteessa haittoihin. OECD:n (2022, 5–6) mukaan tekoälyjärjestelmät voivat käyttää valtavia laskennallisia resursseja ja siten suuren määrän energiaa ja datakeskusten jäähdytykseen käytettävää vettä. Tekoälyn avulla voidaan tukea kestävyystavoitteiden saavuttamista, mutta samaan aikaan tekoälyjärjestelmien koulutuksella ja käytöllä on ympäristövaikutuksia.

Tekoälyä on hyödynnetty monissa yhteyksissä kulttuuriperintökentällä, mutta yhtenäistä



näkemyistä sen hyödyistä ja toisaalta sen aiheuttamasta ympäristökuormasta ei ole. Tärkeää on kerätä kokemuksia, analysoida niitä yhdessä ja etsiä ratkaisuja.

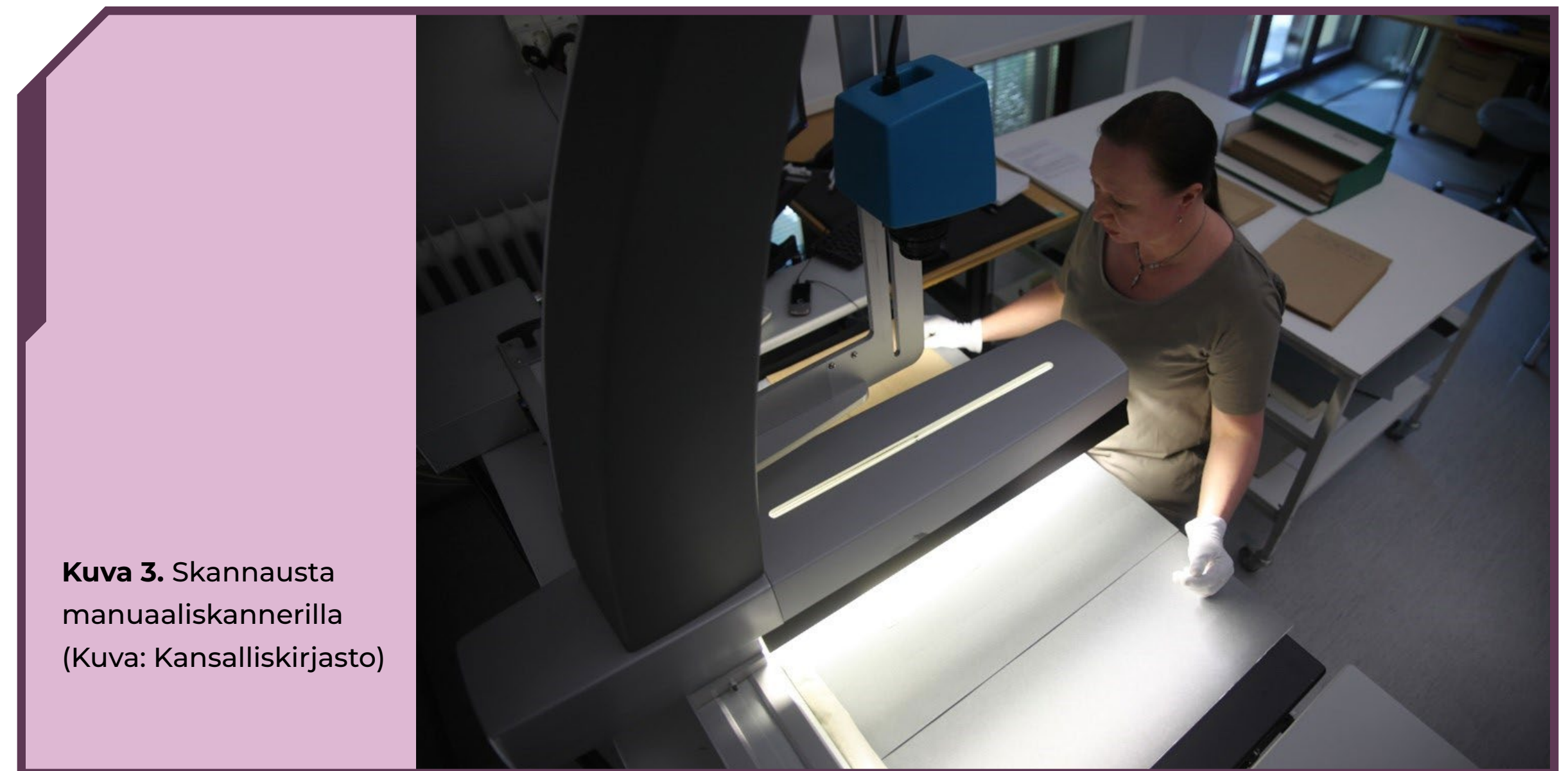
### Muita näkökulmia vastuullisuuteen

Digitointi on muutakin kuin laitteita ja dataa. Tarvitaan tilat digitointiprosessin eri vaiheisiin ja logistiikkaa aineiston kuljettamiseen sekä ennen kaikkea asiantuntemusta ja osaamista digitoinnin toteuttamiseen – eli digitointialan osaavia ammattilaisia. Agenda 2030 -ohjelman tavoite 8 edistää ihmisarvoista työtä ja talouskasvua. Kansainvälinen työjärjestö ILO (2015) on määritellyt ihmisarvoisen työn muun muassa mahdollisuudeksi tehdä tuottavaa,

turvallista ja henkilökohtaista kehittymistä tukevaa työtä tasa-arvoisessa työympäristössä.

Kansalliskirjastossa henkilöstön hyvinvoinnista pidetään huolta yhdessä työterveyshuollon kanssa. Työhyvinvointikysely järjestetään joka toinen vuosi ja niiden tuloksia hyödynnetään työhyvinvointisuunnitelmien pohjana. Digitointiprosessissa on toisinaan melko toisteisia työvaiheita, joten hyvä ergonomia on ensiarvoisen tärkeää. Pitkät projektit ja vakioidut työprosessit tuovat ennakoitavuutta työn arkeen kuten myös vakinaiset työsuhteet. Työkiertoja toteutetaan mahdollisuuksien mukaan. Kuvassa 3 digitoidaan hauraita aineistoja manuaaliskannerilla.

Tila- ja logistiikkaratkaisulla on vaikutusta erityisesti ekologiseen ja taloudelliseen kestävy-



**Kuva 3.** Skannausta manuaaliskannerilla (Kuva: Kansalliskirjasto)



teen. Aineistokuljetusten osalta vastuullisia tekoja voidaan tehdä kuljettamalla aineistoa kerralla suurempia määriä – kuitenkin siten, että aineistoturvallisuus ei vaarannu – ja siirtymällä vähäpäästöisiin ajoneuvoihin siinä vaiheessa, kun niiden vaihto on ajankohtainen, tai asettamalla hiilidioksidipäästöt yhdeksi kriteeriksi kuljetuspalveluita kilpailutettaessa. Aineiston turhan siirtelyn ja käsittelyn välttämiseksi Kansalliskirjaston digitointiprosessissa huomioidaan kaikki aineiston tarvitsemat vaiheet, kuten esimerkiksi konservointi.

## Vastuullinen julkaiseminen ja käyttö

Ekologisen, taloudellisen ja sosiaalisen kestävyden vuoksi digitoitu aineisto kannattaa julkaista mahdollisimman avoimena, sillä näin

sen tuottama hyöty on suurin. Tämä tukee myös YK:n Agenda 2030:n tavoitetta 10 Eriarvoisuuden vähentäminen. Verkon kautta käytettävissä oleva aineisto mahdollistaa yhä useammalle pääsyn kulttuuriperintöön ajasta ja paikasta riippumatta lisäten näin tiedonsaannin tasa-arvoa.

YK:n Agenda 2030:n tavoite 16 muistuttaa hyvästä hallinnosta, mikä tarkoittaa, että digitoitujen aineistojen julkaisemisen ja käytön pitää olla lakien mukaista. Käytännössä tämä tarkoittaa tietosuojan ja tekijänoikeuteen liittyvien rajoitusten huomioimista. Arkistojen kannalta tietosuoja vaatii usein enemmän rajoituksia kuin tekijänoikeus, joka taas on kirjastoaineistojen osalta huomioitava tarkkaan, kun aineistoja halutaan avata. Museoiden kuvakokoelmiin saattaa liittyä molempia rajoitteita.

Tekijänoikeus on voimassa 70 vuotta teoksen kaikkien tekijöiden kuolemasta. Monografioiden osalta on useimmiten mahdollista tarkastaa tekijöiden kuolinvuodet. Lehti- ja pienpainateaineistot ovat tässä suhteessa ongelmallisempia, koska tekijöitä on hyvin paljon eikä kaikkien nimiä ole tiedossa. Aineiston avaaminen on kuitenkin mahdollista tekijänoikeusjärjestöjen kanssa sopimalla (Tekijänoikeuslaki 404/1961 16d§ 26d§). Sopimukset ovat yleensä määräaikaista ja edellyttävät vuosittain maksettavaa lisenssimaksua. Sopimuksia solmiessa on pyrittävä ehtoihin, jotka eivät ole liian rajaavia.

Tekijänoikeudenalainen digitoitu aineisto, jonka käytöstä ei ole sovittu tekijänoikeusjärjestöjen kanssa, on käytettävissä vain vapaakappalekirjastoissa, Eduskunnan kirjastossa ja Kansallisessa audiovisuaalisessa instituutissa.

Koska vapaakappalekirjastoja on Suomessa vain kuusi, edellyttää suojatun aineiston käyttö usein matkustamista, eikä tasa-arvoisen pääsy tietoon tältä osin toteudu. (Vapaakappaleaineistojen käyttö.)

Paitsi aineistojen avoimuudella myös niiden saavutettavuudella voidaan edistää yhdenvertaisuutta. Julkaisujärjestelmän käyttöliittymän tulee palvella erilaisia käyttäjiä ja olla käytettävyydeltään ja ymmärrettävyydeltään sellainen, että kaikki pystyvät aineistoja käyttämään. Digitaalinen aineistomuoto mahdollistaa fyysistä aineistoa paremmin aineiston käytettävyyden erilaisille käyttäjäryhmille – esimerkiksi ruudunlukulaitteen käytön tai tekstin suurentamisen.

Taloudellisen kestävyden näkökulmasta aineiston monipuolinen käyttö on tärkeää.



Pelkät hakutoiminnot eivät riitä, vaan aineisto kuvailutietoineen pitäisi pystyä lataamaan myös datana. Kansalliskirjaston digitointiohjelma huomioi myös aineiston käytön datana silloin, kun se tekijänoikeuksien puolesta on mahdollista. Digi.kansalliskirjasto.fi-palvelun lataustyökalulla voi jokainen luoda aineistopaketteja haluamastaan tekijänoikeusvapaasta aineistosta. Kansalliskirjasto on myös avan- nut eri hankkeissa tuottamia datapaketteja muun muassa Data.nationallibrary.fi-palve- lussa. (Ukkonen & Lilja 2025, 9.)

Kansalliskirjasto on mukana edistämässä aineistojen yhteiskäyttöä. Tämän mahdollis- tavat IIF (International Image Interoperability Framework) -standardien mukaiset tekniset ratkaisut, joiden avulla voidaan esimerkiksi esittää eri organisaatioiden tuottamia kuvia ja tekstejä rinnakkain samassa katseluohjel-

massa. Niitä voi myös muokata ja hyödyntää monipuolisilla tavoilla. Keväällä 2025 käynnis- tyvä hanke toteuttaa IIF-tuen Finnaan ja digi. kansalliskirjasto.fi-palveluun.

Kestävyyttä on mahdollista edelleen vahvis- taa ohjeistamalla myös aineiston käyttäjiä toimimaan vastuullisesti. Viittausohjeet sekä tekijänoikeuksia ja tietosuojaa huomioivat ohjeet ovat kulttuuriperintöorganisaatioissa jo arkipäivää, mutta ohjeita voitaisiin laajentaa huomioimaan myös ekologinen kestävyys liit- tyen esimerkiksi käyttäjän päätelaitteisiin tai datanhallintaan. Tutkimuskäytäntöjen ympä- ristövastuullisuuden tueksi on olemassa erilai- sia työkaluja, esimerkiksi Digital Humanities Climate Coalitionin työkalupakki (DHCC. The Digital Humanities Climate Coalition Toolkit).

## Digitoidun aineiston elinkaari

Digitoidun aineiston elinkaari ei pääty siihen, että se viedään julkaisujärjestelmään tai joku lataa sen. Digitoidun aineiston pitkäaikaissäi- lytys vaatii erillisen ratkaisun. Se ei ole vain varmuuskopion tallentamista, vaan myös aineiston säilyvyyteen ja käytettävyyteen liit- tyvien ratkaisujen kehittämistä. Aineisto pitää pystyä löytämään, avaamaan ja lukemaan vuosikymmenten, jopa vuosisatojen päästä. Tämä edellyttää digitoinnilta ja kuvailulta tiet- tyjä laatustandardeja. Teknisten ratkaisujen osalta Kansalliskirjasto on noudattanut Kan- sallisarkiston määräyksiä (2021).

Pitkäaikaissäilytys takaa aineistolle mahdolli- simman pitkäaikaisen käytön, mutta sillä on oma hiilijalanjälkensä, mistä syystä aineiston valinnassa täytyy olla kriittinen ja välttää pääl- lekkäisyyksiä. Suomessa kulttuuriperintöorga- nisaatioiden PAS-palvelut tarjoaa ja omistaa

opetus- ja kulttuuriministeriö. Ne toteuttaa Tieteen tietotekniikan keskus CSC, joka on laatinut vuonna 2023 pitkäaikaissaatavuuden palveluille kestävyystiekartan vuoteen 2026. Palvelulle laskettiin myös vuosittainen hiilija- lanjälki (CSC Vastuullisuus ja kestävä kehitys).

Pitkäaikaissäilytyksen lisäksi digitoitu aineisto voi jatkaa elämänsä tutkimusaineistona. Avoimen tieteen käytäntöjä seuraava tutkija haluaa mahdollisuuksien mukaan tallentaa ja avata käyttämänsä datan ja näin varmis- taa, että kerran tehty työ on hyödynnettävissä uudelleen osana toista tutkimusta. (Avoin tiede. Tutkimusaineistojen avoimuus.) Vastuul- lisesti digitoiva organisaatio ottaa jatkokäytön huomioon ja laatii aineiston käyttöehdot tai lisensiointisopimukset siten, että jatkokäytön mahdollisuudet on selkeästi ilmaistu.



## Yhteenveto

Kestävä kehitys on moniulotteinen käsite. Ekologinen, taloudellinen, sosiaalinen ja kulttuurinen kestävyys korostavat erilaisia, joskus myös keskenään ristiriitaisia arvoja. Kestävän kehityksen edistäminen digitoinnissa merkitseekin jatkuvaa arvopohdintaa: Painotetaanko ekologista kestävyttä, joka kannustaa tyytymään määrällisesti vähempään, vai mahdollisimman laajaa käyttöä tavoittelevaa taloudellista kestävyttä? Vai pitäisikö noudattaa vähemmistöjen tarpeet huomioivaa sosiaalista kestävyttä tai katoamassa olevan kieliyhmän aineistojen kulttuurista kestävyttä?

Kaikkia kestävyiden aspekteja palvelee korkean laadun ylläpitäminen digitointiprosessin

vaiheissa: eri käyttäjäryhmien huomioiminen aineiston valinnassa, korkealaatuinen kuvailu, tehokkaat ja vihreät teknologiset ratkaisut, hyvä työympäristö sekä lainsäädännön puitteissa mahdollisimman avoin käyttö. Laatu tuottamalla varmistetaan, että digitointi tarvitsee toteuttaa vain kerran ja että aineiston käyttö on tehokasta.

Kestävä kehitys muuttaa toimintaamme. Muutosten päämääränä on turvata nykyisille ja tuleville sukupolville hyvät elämisen mahdollisuudet. Kestävyyden periaatteet huomioivalla digitoinnilla voimme tehdä osamme tavoitteiden saavuttamiseksi ja turvata myös kulttuuriperinnön säilymisen tuleville sukupolville.

## LÄHTEET

Avoin tiede. Tutkimusaineistojen avoimuus. Saatavissa: <https://avointiede.fi/fi/asiantuntijaryhmat/tutkimusaineistojen-avoimuus> [viitattu 5.2.2025].

CSC. Vastuullisuusraportti. 2023. Saatavissa: <https://csc.fi/app/uploads/2024/04/CSC-vastuullisuusraportti-2023.pdf> [viitattu 8.2.2025].

CSC. Vastuullisuus ja kestävä kehitys. Saatavissa: <https://csc.fi/tietoa-meista/vastuullisuus-ja-kestava-kehitys/> [viitattu 5.2.2025].

Culture at the heart of Climate action. Saatavissa: <https://www.climateheritage.org/global-call> [viitattu 21.2.2025].

DHCC. The Digital Humanities Climate Coalition Toolkit. Saatavissa: <https://sas-dhrh.github.io/dhcc-toolkit/> [viitattu 19.2.2025].

Eklund, S. ym. 2022. Rohkeasti kestävään muutokseen – Kestävä kehitys Kansalliskirjastossa 2021–2030. Helsinki: Kansalliskirjasto. Saatavissa: <https://urn.fi/URN:ISBN:978-951-51-7423-9> [viitattu 29.1.2025].

Green ICT. 2022. Green ICT: hankkijan opas. Saatavissa: <https://greenict.fi/tietopankki/hankkijanopas/> [viitattu 7.2.2025].

Hiilineutraali Helsingin yliopisto 2030: Tiekartta. Saatavissa: [https://www.helsinki.fi/assets/drupal/2023-04/HY\\_Tiekartta\\_tauptapaperi.pdf](https://www.helsinki.fi/assets/drupal/2023-04/HY_Tiekartta_tauptapaperi.pdf) [viitattu 29.1.2025].

Hyvärinen, K. ym. 2023. Kuvailun laatusuositus 2023. Helsinki: Kansalliskirjasto. Saatavissa: <https://urn.fi/URN:ISBN:978-951-51-8974-5> [viitattu 31.1.2025].

ILO (International Labour Organization). 2015. Decent Work Agenda. Saatavissa: <https://www.ilo.org/topics/decent-work> [viitattu 13.2.2025].

Inkinen J., Lehtinen M. & Suominen O. 2023. Annifin ehdotusten osuvuus on parantunut Theseus-julkaisuarkistossa. Tietolinja 1. Saatavissa: <https://urn.fi/URN:NBN:fi-fe2023062865282>. [viitattu 4.2.2025].



Kansallisarkiston määrittämät arkistoitavien erikoisaineistojen digitointiin 19.1.2021. Saatavissa: <https://kansallisarkisto.fi/documents/141232930/151500284/KA+m%C3%A4%C3%A4ritykset+erikoisaineistojen+digitointiin.pdf/059812a8-c50b-4753-727e-887d4650644e/KA+m%C3%A4%C3%A4ritykset+erikoisaineistojen+digitointiin.pdf?t=1677499069673> [viitattu 5.2.2025].

Kestävän kehityksen toimikunnan toimenpidesuunnitelma vuosille 2025–2027. Saatavissa: [https://kestavakehitys.fi/documents/205289881/236229875/Kest%C3%A4v%C3%A4n%20kehityksen%20toimikunnan%20toimenpidesuunnitelma%202025-2027\\_1.pdf/9b252ae7-4a73-98b6-07ca-825f552c8a0e/Kest%C3%A4v%C3%A4n%20kehityksen%20toimikunnan%20toimenpidesuunnitelma%202025-2027\\_1.pdf?t=1742302052164](https://kestavakehitys.fi/documents/205289881/236229875/Kest%C3%A4v%C3%A4n%20kehityksen%20toimikunnan%20toimenpidesuunnitelma%202025-2027_1.pdf/9b252ae7-4a73-98b6-07ca-825f552c8a0e/Kest%C3%A4v%C3%A4n%20kehityksen%20toimikunnan%20toimenpidesuunnitelma%202025-2027_1.pdf?t=1742302052164) [viitattu 7.5.2025].

Laki viranomaisten toiminnan julkisuudesta. 621/1999 Saatavissa: <https://www.finlex.fi/fi/lainsaadanto/1999/621> [viitattu 4.2.2025].

Litvinets, V. & Pijselman, M. 2024. AI and Sustainability: Opportunities, Challenges, and Impact. EY: Shape the future with confidence 14-11-2024. Saatavissa: [https://www.ey.com/en\\_nl/insights/climate-change-sustainability-services/ai-and-sustainability-opportunities-challenges-and-impact](https://www.ey.com/en_nl/insights/climate-change-sustainability-services/ai-and-sustainability-opportunities-challenges-and-impact). [viitattu 14.2.2025].

OECD 2022. Measuring the environmental impacts of artificial intelligence compute and applications: The AI footprint. OECD Digital Economy Papers 341. Paris: OECD Publishing. Saatavissa: <https://doi.org/10.1787/7babf571-en> [viitattu 14.2.2025].

Paschalidou, E., Fafet, C. & Milios, L. 2022. A Strong Sustainability Framework for Digital Preservation of Cultural Heritage: Introducing the Eco-Sufficiency Perspective 20-05-2022.

Saatavissa: [https://www.researchgate.net/publication/360753371\\_A\\_Strong\\_Sustainability\\_Framework\\_for\\_Digital\\_Preservation\\_of\\_Cultural\\_Heritage\\_Introducing\\_the\\_Eco-Sufficiency\\_Perspective](https://www.researchgate.net/publication/360753371_A_Strong_Sustainability_Framework_for_Digital_Preservation_of_Cultural_Heritage_Introducing_the_Eco-Sufficiency_Perspective) [viitattu 12.2.2025].

Pendergrass, K., Sampson, W., Walsh, T. & Alagna, L. 2019. Toward Environmentally Sustainable Digital Preservation. The American Archivist 1. Saatavissa: <https://dash.harvard.edu/bitstream/handle/1/40741399/360-9081-82.1.165.pdf?sequence=1&isAllowed=y> [viitattu 12.2.2025].

Pinto, R. À. 2018. Digital sovereignty or digital colonialism? Internet and Democracy 15:27.

Suominen, O. 2024. Käteviä kratteja vai Vanhanpakanan juonitteluja? Automaattista kuvailua Viron kansalliskirjastossa. Tietolinja 2. Saatavissa: <https://urn.fi/URN:NBN:fi-fe20241217103953> [viitattu 4.2.2025].

Tekijänoikeuslaki 404/1961. <https://www.finlex.fi/fi/lainsaadanto/1961/404> [viitattu 4.2.2025].

Ukkonen, A.-E. & Lilja, J. (toim.) 2025. Kansalliskirjaston digitointiohjelma 2025–2028. Saatavissa: <https://urn.fi/URN:ISBN:978-952-84-0742-3> [viitattu 17.2.2025].

Uusi hanke monipuolistaa visuaalisten kulttuuriaineistojen esittämistä. 2024. Kansalliskirjaston verkkosivujen uutinen 5.4.2024. Saatavissa: <https://www.kansalliskirjasto.fi/fi/uutiset/uusi-hanke-monipuolistaa-visuaalisten-kulttuuriaineistojen-esittamista> [viitattu 12.2.2025].

Vapaakappaleaineistojen käyttö. Saatavissa: <https://www.kansalliskirjasto.fi/fi/vapaakappalepalvelut/vapaakappaleaineistojen-kaytto> [viitattu 17.2.2025].



Noora Talsi

## LOPPUSANAT

**D**igisti parempi -julkaisu on suunnattu digitoinnin kehittäjille ja tekijöille, jotka haluavat edistää aineistojen säilymistä ja saavutettavuutta. Digitointi ei ainoastaan suojaaa kulttuuriperintöämme, vaan se on myös keskeinen osa tiedonhallintaa ja varmistaa aineistojen säilymisen tuleville sukupolville.

Suuret digitointiprojektit vaativat aikaa ja asiantuntemusta, ja niiden merkitys kasvaa jatkuvasti arkistoissa ja muissa muistiorganisaatioissa. Tarve tehokkaalle ja asiantuntevalle digitoinnille on suurempi kuin koskaan. Digitaalinen tallennus parantaa aineistojen

säilyvyyttä, vähentää fyysistä kulumista ja suojaaa katastrofeilta, kuten tulipaloilta ja vesivaingoilta. Lisäksi se mahdollistaa paremman saavutettavuuden ja käytettävyyden, sillä digitoitut aineistot eivät ole sidottuja aikaan tai paikkaan.

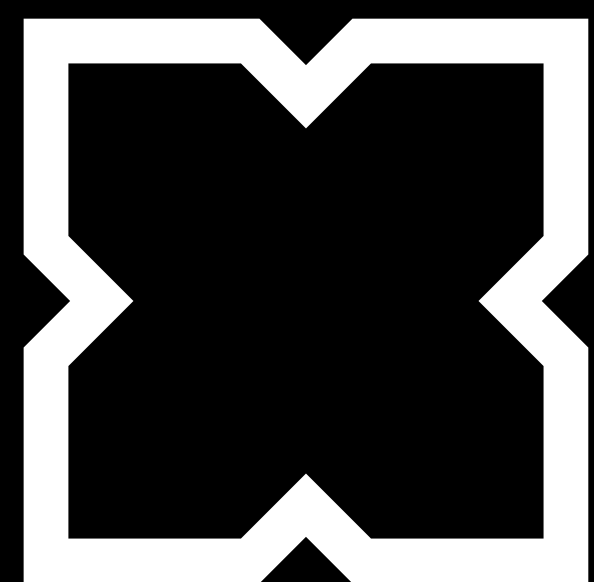
Tekstintunnistuksen (OCR) kaltaiset työkalut helpottavat tiedon hakua ja käyttöä, ja digitoitut aineistot säästävät fyysistä tilaa sekä ympäristöä. Digitaaliset tiedostot voidaan myös suojata ja hallita käyttöoikeuksin, mikä lisää tietoturvaa ja estää tietojen menetyksen varmuuskopioinnin avulla.

Digitointiosaajien koulutukset ovat tarjonneet kattavaa oppia koko digitointiprosessista – suunnittelusta ja aineiston valmistelusta aina laadunvarmistukseen ja tallennukseen asti. Erityisesti arkistosektori on hyötynyt täsmäkoulutuksesta, joka vastaa alan ajankohtaisiin tarpeisiin.

Eurooppalainen Time Machine -organisaatio tähtää koko Euroopan kulttuuriperinnön digitointiin elinaikanamme, ja myös Xamkin digitointiprojektit ovat herättäneet kansainvälistä kiinnostusta. Digitointiosaamisen kehittämiseen kaivataan lisää koulutusmahdollisuuksia,

ja Suomessa kehitetyt ketterät digitointikoulutusmallit voisivat toimia pohjana kansainväliselle yhteistyölle. Kiinnostusta on jo herännyt eri puolilla Eurooppaa, ja tulevaisuudessa tavoitteena on rakentaa vahvempaa kansainvälistä verkostoa digitoinnin koulutuksen ja kehittämisen ympärille.

Digitointi on avain menneisyyden säilyttämiseen ja tulevaisuuden rakentamiseen. Nyt on aika tehdä kulttuuriperinnöstä digisti parempi.



**XAMK**  
**KEHITTÄÄ**