

# Minne menet, yksityinen aineisto?



**Taina Saarenpää**  
Tutkija  
Turun yliopisto



**Miia Kosonen**  
TKI-asiantuntija  
Xamk, Digitalia

**Y**ksityinen digitaalinen tieto on yksityisen henkilön tai yhteisön tuottamaa syntysähköistä tai digitoitua tietoa. Tässä artikkelissa pääpaino on yksittäisten ihmisten itse keräämissä ja tuottamissa digitaalisissa tiedoissa. Selvitimme, millaisia palveluja näiden aineistojen säilyttämiseen on saatavilla meillä ja maailmalla.

Elämme yksityisen tiedon tuottamisen kulta-aikaa. Samalla on vallalla hämmennys siitä, kuinka yksityinen digitaalinen tieto voidaan säilyttää, kenellä on vastuu sen keräämisestä ja säilyttämisestä, ja mikä on muistiorganisaatioiden rooli tässä kokonaisuudessa. Koska emme osaa hallita ja säilyttää tietojamme, nykyaikaa on luonnehdittu myös digitaaliseksi pimeäksi ajaksi (Bollacker, 2010). Tieto on moninaista ja hajanaista.

Myös muiden kuin viranomaisten ja merkkihenkilöiden aineistot nähdään kuitenkin nykyisin kiinnostavana ja olennaisena osana yhteistä kulttuuriperintöä. Historiantutkimuksessa on jo 1970-luvulta alkaen puhuttu niin sanottujen uusien historioiden noususta tutkimuskohteeksi. Lisämausteensa muutokseen ovat tuoneet sosiaalisen verkon maailmat, joissa merkittävät ilmiöt ja kansanliikkeet voivat syntyä ruohonjuuritasolla tavallisten ihmis-

ten toimesta. Emme voi etukäteen ennustaa, mikä tieto osoittautuu tulevaisuudessa arvokkaaksi. Siksi on erityisen tärkeää saada yksityistä aineistoa talteen, oli se sitten asiakirjoja, päiväkirjoja, sähköposteja tai sosiaalisen median päivityksiä.

## Missä yksityinen digitaalinen tieto on tällä hetkellä?

*Marshall* (2008) listasi tutkimuksessaan yleisimmät tavat yksityisten henkilöiden digitaalisen tiedon säilyttämiseen. Näitä ovat varmuuskopiot, tiedostojen siirto aktiivikäytössä olevalle laitteelle, ulkoiset tallennusvälineet (kovalevyt, muistitikut, CD, DVD), ulkoiset palvelut (sähköposti, sosiaalinen media, sittemmin yleistyneet pilvipalvelut) ja koko alustan säilyttäminen tietoineen. Mikään näistä yleisesti hyödynnetyistä vaihtoehdoista ei kuitenkaan turvaa tiedon pitkäaikaissäilytystä. Säilyttämisen luotettavia teknisiä ratkaisuja on kuvattu *Osmo Palosen* artikkelissa Liikearkistoyhdistyksen tuoreessa kirjassa *Säilyykö sähköinen – ja kuinka kauan?* (2018).

Toistuva ongelma yksityisissä digitaalisissa aineistoissa on myös kuvailevan metatiedon puuttuminen. Jokainen tietää tunteen, kun suvun valokuvista katsovat vieraat kasvot. Tieto kuvissa esiintyvistä henkilöis-

tä katoaa ajan myötä asianosaisten kuoltua tai sairastuttua. Kaakkois-Suomen ammattikorkeakoulussa tutkimuskeskus Digitalia on kehittänyt työnimellä ”Kansalaisarkisto” kulkevaa säilytysratkaisua, joka tukee kattavaa metatiedottamista, mutta aniharva tavallinen kansalainen on yhtä omistautunut aineistonsa kuvailulle kuin palvelun pilottikäyttäjät (Kausalainen & Uosukainen, 2017). Työ on aikaa vievää eikä mikään automaatio pysty ottamaan siitä vastuuta käyttäjän puolesta. Jokainen on omien tietojensa paras – ja usein ainoa – asiantuntija. Arkistoala ei ole vielä löytänyt viisastenkiveä siihen, kuinka aineiston kuvailusta tehdään riittävän helppoa ja näkymätöntä tavalliselle käyttäjälle menettämättä silti tietojen arvoa.

Puhe on tulevaisuuden käyttöliittymä, mutta olisiko se myös tulevaisuuden tapa metatiedottaa henkilökohtainen aineisto? Saisimmeko aineistot kuvattua riittävällä tarkkuudella, jos kone poimii tiedot suoraan juttuhetkestämme suvun vanhimpien kanssa ja tunnistaa jopa eri murteiden vivahteet? Kuten tekoälyn kehityksessä ylipäättään, kielentutkimuksen osaaminen ja sen soveltaminen yhdessä teknologian rinnalla on jälleen avainasemassa.

## Palveluja tiedon tallentamiseen ja esittämiseen

Yksityisille käyttäjille on tarjolla koko joukko erilaisia tiedon tallennus- ja järjestämispalveluja. Näissä tavoitteena on yleensä tiedon säilyttäminen ja hallinta lyhyellä aikavälillä. Pitkäaikaissäilytyksen näkökulma puuttuu, eikä oma kuolemakaan tunnu koskaan etukäteen ajankohtaiselta. Muita tunnusomaisia lähestymistapoja ovat digitaalisen minän rakentaminen, esittäminen ja jakaminen tuomalla tietoa yhteen eri kanavista. Esimerkiksi **Infi** jatkaa sosiaalisen median perusajatusta mahdollistamalla luodun ”digitaalisen minän” jakamisen ja esittämisen muille. Erilaiset aikajana- ja järjestämispalvelut ovat yleensäkin suosittuja. Sukututkijoille suunnattu ilmainen **Twile**-palvelu tukee oman elämäntarinan kokoamista, tallentamista ja jakamista.

**Digi.me**:n avulla käyttäjä voi kerätä eri sosiaalisen median palveluihin tuottamansa tiedon yhteen paikkaan ja tallentaa sen haluamalleen tallennusalustalle. **Digi.me** pyrkii antamaan käyttäjille paremmat mahdollisuudet hallita omia tietojaan. Palvelu on ilmainen. Ansainta perustuu siihen, että yritykset maksavat, mikäli käyttäjä luovuttaa tietojaan niiden käyttöön. Oma lähestymistapansa on myös digitaalinen perintö. Esimerkiksi suomalainen kaupallinen palvelu **Emonum** lupaa koota käyttäjien digitaalista aineistoa yhteen paikkaan ja antaa mahdollisuuden päättää tietojen omistusoikeudesta tilanteessa, jossa oma hallinta katoaa esimerkiksi käyttäjän kuoltua.

Seuraava taulukko tiivistää eri palveluiden lähestymistapoja yksityiseen tietoon.

Tavoite	Palvelun fokus	Esimerkkejä
Tiedon hallinta	Omien tietojen lyhytaikainen säilyttäminen, hallinta ja synkronointi, myös yhteiset työtilat	Google Drive Dropbox Evernote
Digitaalinen minä	Koosteet käyttäjän itse jakamista tiedoista eri digipalveluissa	Digi.me
Tiedon jakelu ja visualisointi	Aikajana esim. omista elämäntapahtumista	Infi Twile
Digitaalinen perintö	Kuolemaan varautuminen, läheisille pääsy omiin tietoihin	Everplans Emonum
Pitkäaikaissäilytys	Omien tietojen pitkäaikainen säilyttäminen	Heritage Werks

## Myös pysyvyyttä lupaavat palvelut katoavat

Yhdysvaltalainen **Chronicle of Life Foundation** on voittoa tavoittelematon, vuonna 2008 perustettu yksityiselle aineistolle pysyvää säilytystä tarjoava verkkopalvelu. Tieto on lähikäyttökohtaisesti yksityistä. Sitä voi jakaa läheisten ja ystävien kanssa, tai tehdä julkiseksi. Digitaalisen perinnön lähestymistavan mukaisesti palveluun voi myös rakentaa alustan esimerkiksi kuolleen läheisen muistamiseksi. Palvelulupauksen mukaan toteutus on ainoa laatuaan, mutta kysymys jatkuvuudesta ja rahoituksesta jää kuitenkin epäselväksi.

Iso-Britanniasta lähtöisin oleva **Life's Time Capsule** tarjoaa kuukausimaksullista palvelua käyttäjän digitaalisten tietojen kokoamiseen ja jakeluun valitulle yhteisölle. Tieto voi olla esimerkiksi kuvia, päiväkirjamerkintöjä ja sosiaalisen median profiileja. Tavoite on, että ihmiset muistetaan myös tulevaisuudessa. ”Kapseli” siis pyrkii laajentamaan digitaalisen minän lähestymistapaa kohti arkistoa. Säilytysalustana on yksi isoimmista pelureista eli Amazon. Myös Time Capsulella on kiinnostava palvelulupaus, mutta hämmennystä käyttäjissä voi herättää esimerkiksi se, ettei ylläpito

ole päivittänyt palvelun sosiaalisen median tilejä kesän 2017 jälkeen.

Selvityksen kohteena olevista palveluista vakuuttavimmaksi osoittautui yhdysvaltalainen **Heritage Werks**. Se tarjoaa kokonaispakettia erityyppisten aineistojen arkistointiin ja tiedon koko elinkaaren hallintaan. Monista muista toimijoista poiketen yrityksellä on palkkalistoillaan niin teknisiä asiantuntijoita kuin arkistointihoidattajia, tutkijoita ja historioitsijoita. Pääasiallista kohderyhmää ovat yritykset, urheiluseurat ja luksusbrändit, mutta myös julkisuuden henkilöt, tavalliset kansalaiset ja sukuyhteisöt voivat hyödyntää palvelua. Mielenkiintoinen yksityiskohta on hintatietojen puuttuminen: kuukausimaksun suuruus riippuu luonnollisesti siitä, tilaako käyttäjä yksittäisen kuva-albumin koostamisen vai laajalaisen sähköisen arkiston, mutta lähtöhintaa voi vain arvailla.

Edellä mainitut palvelut vertautuvat Digitaliassa kehitettävään Kansalaisarkistoon. Yhtäläisiä piirteitä on monia, mutta käytännössä tämän tyyppisiä pitkäaikaissäilytyksen kokonaispalveluja ei juuri ole kuluttajille tarjolla. Kyse on kuitenkin niin mittavasta ja teknis-juridisesti haastavasta ratkaisusta, ettei tätä voida pitää yllätyksenä.

Realistisempi liikkeellelähtö olisikin tuottaa pienempiä osaratkaisuja, joilla varmistetaan ainakin tietojen luettavuus, yhteensopivuus ja arkistointikelpoinen muoto. Tällaisia ovat Digitaliassa jo kehitetyt työkulut digitaalisten aineistojen käsittelyyn. Esimerkiksi tekstiä sisältävien kuvien sisältö analysoidaan, avainsanat tunnistetaan, dokumentteihin lisätään OCR-tieto ja lopuksi ne muunnetaan arkistokelpoiseen PDF/A-muotoon sisältäen OCR-tiedon ja sisältöä kuvailevan metadatan (Jääskeläinen & Uosukainen, 2018). Niin ikään vanhat sähköpostikanisiot liitetiedostoinen muuntuvat automaattisesti PDF/A-muotoon (Jääskeläinen, 2016). Työkulut perustuvat täysin avoimen lähdekoodin työkaluille ja niitä pääsee vapaasti soveltamaan ja testaamaan: digitalia.xamk.fi. Sähköpostien muuntotyökalua varten tarvittavat tunnukset saa Digitaliasta *Anssi Jääskeläiseltä*.

## Lopuksi

Kansallisarkiston strategiaan on kirjattu kansallinen toimintamalli yksityisten digitaalisten arkistojen käsittelyyn ja pitkäaikaissäilyttämiseen. Tietoa mahdollisista ratkaisuista ei selvitystä tehdessä vielä ollut saatavilla. Yhdessä henkilökohtaisia elämäntarinoita kokoavien digitaalisten palveluiden kehittämisen kanssa tämä luo kuitenkin toivoa siitä, että arkistoissa on tulevaisuudessa entistä vähemmän hiljaisuuksia ja aiempaa rikkaammat kokoelmat myös yksityisiä aineistoja.

Ongelmaksi useimmissa tällä hetkellä tarjolla olevissa palveluissa nousivat aidon pitkäaikaissäilytyksen puuttuminen sekä liiketoimintaan väistämättä liitettävä kysymys jatkuvuudesta ja pysyvyydestä. Selvityksessämme ei löytynyt ainuttakaan toteutusta, jossa muistiorganisaatio

olisi mukana tarjoamassa yksityisille käyttäjille digitaalisen tiedon säilytyspalvelua pitkäaikaissäilytystä ja jatkokäyttöä varten. Tällaisen mallin tarpeellisuuteen olisi hyvä havahtua, jotta vastuu yksityisten aineistojen säilymisestä ei jäisi pelkästään kaupallisten palveluntarjoajien tai yhtä lailla lyhytjänteisellä hankerahoituksella toimivien kehittämisorganisaatioiden varaan.

Tällä hetkellä muistiorganisaatioiden osuus rajoittuu ohjeistamiseen ja aineistojen vastaanottoon ilman tähän soveltuvia verkkopalveluita. Yksityisen tiedon säilyttäminen, ylläpito ja käytettäväksi asettaminen ei voi olla vain yksityishenkilöiden tai pienyhteisöjen resurssien ja tietotaidon varassa. Monimutkaisessa tietoympäristössä tarvitaan yhteistyötä, nykyistä laajempaa rahoitus pohjaa ja edellä mainittujen organisaatioiden osaamisen yhdistämistä kokonaan uudenlaiseksi toimintamalliksi.

*Artikkeli pohjautuu Taina Saarenpään yamk-opinnäytetyöhön "Yksityisen digitaalisen tiedon pitkäaikaissäilytyksen nykytila" (2018), Kaakkois-Suomen ammattikorkeakoulu.*

## Lähteet

Bollacker, K.D. 2010. Computing Science: Avoiding a Digital Dark Age. *American Scientist*, 98, 2, 106–110.

Heikkinen, J. P. 2018. Valemedia osaksi kulttuuriperintöä. *Scripta selecta* – kirjoituksia Kansalliskirjaston kokoelmista. Saatavilla: <https://blogs.helsinki.fi/scriptaselecta/2018/03/26/valemedia-osaksi-kulttuuriperintoa/>

Jääskeläinen, A. 2016. Sähköpostit Outlookista arkistoon. *Faili 4/2016*, 33–35.

Jääskeläinen, A. & Uosukainen, L. 2018. OCR: Unleash the hidden information. *Archiving*, Washington, 17–20 April, 2018.

Kausalainen, E. & Uosukainen, L. 2017. Kansalaisarkisto – sukuyhteisön aarteet talteen digitaaliseen arkistoon. Kosonen M. (toim.): *Digitaalinen tieto haltuun*. Ratkaisuja digitaalisten aineistojen hallintaan ja käyttöön, s. 40–47. Xamk kehittää, 11.

Marshall, C. 2008. Rethinking Personal Digital Archiving, Part 1. Four Challenges from the Field. *D-Lib Magazine*, 14, 3-4. Saatavilla: <http://www.dlib.org/dlib/march08/marshall/03marshall-pt1.html>

Palonen, O. 2018. Sähköinen säilyttäminen toteutetaan it-perustekniikkaa soveltamalla. Kirjassa 'Säilyykö sähköinen – ja kuinka kauan?' Liikearkistoyhdistyksen julkaisuja, 19, 57–68.

Chronicle of Life Foundation. <https://www.chronicleoflife.com/>

Digi.me. <https://digi.me/>

Emonum. <https://emonum.com/>

HeritageWerks. <https://www.heritage-werks.com/>

INFI. <https://www.infibond.com/>

Kansalaisarkisto. <https://osa.xamk.fi>

Life's Time Capsule. <https://www.lifetimecapsule.com/>

Twile: Family history timeline. <http://twile.com>