

Sami Mikkola

**YKSILÖN DIGITAALISEN JALANJÄLJEN MERKITYS  
JA HALLINTA**



JYVÄSKYLÄN YLIOPISTO  
INFORMAATIOTEKNOLOGIAN TIEDEKUNTA  
2022

# TIIVISTELMÄ

Mikkola, Sami

Yksilön digitaalisen jalanjäljen merkitys ja hallinta

Jyväskylä: Jyväskylän yliopisto, 2022, 36 s.

Tietojärjestelmätiede, kandidaatintutkielma

Ohjaaja(t): Marttiin, Pentti

Digitaaliset jalanjäljet ovat tärkeä osa jokaisen elämää. Yksilö jättää niitä enemmän jälkeensä sitä mukaa, kun elämän eri osa-alueet muuttuvat digitaalisiksi. Digitaalisten jalanjälkien suurempi määrä tarkoittaa, että yksilöstä on enemmän tietoa tarjolla. Niille on siten myös enemmän käyttökohteita, jotka lisäävät myös niiden merkitystä. Tämän tutkielman tarkoituksena oli selvittää, minkälainen merkitys yksilön digitaalisilla jalanjäljillä on sekä miten yksilön digitaalisia jalanjälkiä voidaan hallita. Tutkimus toteutettiin kirjallisuuskatsauksena. Merkityksen osalta päädyttiin sellaiseen johtopäätökseen, että yksilön digitaalisten jalanjälkien merkitys on hyvin moninainen. Merkitys syntyy digitaalisten jalanjälkien hyödyistä ja haitoista. Hallinnasta todettiin, että yksilön digitaaliseen jalanjäljen hallintaan eniten vaikuttava tekijä on yksilön taidot. Lisäksi hallintaan ja taitoihin voidaan vaikuttaa koulutuksella ja lainsäädännöllä. Hallinnan osalta päädyttiin myös luokittelemaan kolme erilaista hallintastrategiaa.

Asiasanat: digitaalinen jalanjälki, profilointi, yksityisyys, digitaalisen jalanjäljen hallinta

## **ABSTRACT**

Mikkola, Sami

The significance and management of an individual's digital footprints

Jyväskylä: University of Jyväskylä, 2022, 36 pp.

Information Systems, Bachelor's Thesis

Supervisor(s): Marttiin, Pentti

Digital footprints are an integral part of everyone's everyday life. An individual leaves them behind in an increasing manner as different aspects of their life become digital. An increase in digital footprints also means that more information is available about the individual. More information equals more uses for the digital footprints, which also adds to their significance. The purpose of this bachelor's thesis was to discover the significance of an individual's digital footprints and how an individual's digital footprints can be managed. This research was conducted as a literature review. Regarding the significance of an individual's footprints, it was concluded that the significance is a sum of various benefits and harms caused by the different uses of digital footprints. The most important factors in the management of an individual's footprints were the individual's skills. It was also stated that education and legislation can be key to successful management. Three different management strategies were also presented.

Keywords: digital footprint, profiling, data privacy, digital footprint management

## TAULUKOT

TAULUKKO 1	Hallintastrategioiden luokittelussa hyödynnetyt lähteet	
strategioittain	.....	29

# SISÄLLYS

TIIVISTELMÄ

ABSTRACT

TAULUKOT

1	JOHDANTO.....	6
2	DIGITAALINEN JALANJÄLKI.....	8
	2.1 Määritelmä.....	8
	2.2 Kehitys.....	10
	2.3 Muodostumistavat.....	11
3	YKSILÖN DIGITAALISEN JALANJÄLJEN MERKITYS.....	14
	3.1 Profilointi.....	14
	3.2 Verkkominä.....	17
	3.3 Yksityisyys.....	19
	3.4 Käyttökohteita tutkimuksessa.....	20
4	YKSILÖN DIGITAALISEN JALANJÄLJEN HALLINTA.....	22
	4.1 Taidot, koulutus ja lainsäädäntö.....	22
	4.2 Hallintastrategiat.....	24
	4.2.1 Digitaalisten jalanjalkien välttäminen.....	25
	4.2.2 Positiivinen digitaalinen jalanjälki.....	26
	4.2.3 Digitaalisten jalanjalkien keskittäminen.....	27
	4.2.4 Luvun yhteenveto.....	28
5	YHTEENVETO.....	31
	LÄHTEET.....	33

# 1 JOHDANTO

Digitaalisuus on osa kaikkien elämää. Lähes kaikki omistavat älypuhelimien, joka kulkee mukana päivän mittaan. Viimeisimpänä koronapandemian myötä myös työ ja koulunkäynti ovat muuttuneet digitaalisiksi. Lähes kaikista elämän osa-alueista jää digitaalinen jälki. Digitaaliset jalanjäljet ovat digitaalisia jälkiä, joita jätämme jälkeemme jokapäiväisessä elämässämme (Weaver & Gahegan, 2007). Teknologisen edistyksen ja siten myös digitaalisten sensoreiden lisääntymisen myötä digitaalisten jalanjälkien määrä on kasvanut viime vuosikymmenien aikana merkittävästi (Zhang, Guo, Li & Yu, 2010). Suurempi digitaalisten jalanjälkien määrä mahdollistaa tutkimukselle pääsyn lähemmäksi ihmistä.

Tämän tutkielman tarkoituksena on havainnollistaa yksilön digitaalisten jalanjälkien laajuutta ja merkitystä. Toisekseen pyritään tutkimaan, mitkä asiat vaikuttavat yksilön digitaalisen jalanjäljen hallintaan sekä voidaanko hallintaa luokitella jotenkin. Merkityksen on tarkoitus pohjustaa hallintaa. Tutkimusta rajataan keskittymällä yksilöön.

Tämä eroaa aikaisemmasta tutkimuksesta ainakin siten, että tässä tutkielmassa ei keskitytä ainoastaan yksilöiden asenteisiin merkityksen perustana. Merkityksen oletetaan syntyvän lisäksi yksilön digitaalisten jalanjälkien käyttökohdeiden hyödyistä ja haitoista. Hallinnan suhteen ero aikaisempaan tutkimukseen on taas pyrkimys luokitella eri hallintastrategioita.

Edellä mainittujen asioiden kautta päädyttiin seuraaviin tutkimuskysymyksiin, joihin tässä tutkielmassa pyritään vastaamaan:

- 1) *Minkälainen merkitys yksilön digitaalisella jalanjäljellä on?*
- 2) *Miten yksilön digitaalista jalanjälkeä voidaan hallita?*

Tutkimus toteutetaan kuvailevana kirjallisuuskatsauksena. Kirjallisuutta haetaan Scopuksesta ja Google Scholarista aiheeseen liittyvin hakusanoin. Olenaisimpia hakusanoja ovat muun muassa "digital footprint", "digital footprints" ja "digital traces". Näiden hakusanojen tulokset toimivat pääaineistona. Tämän lisäksi hyödynnetään joidenkin artikkelien lähdeluettelojen mielenkiintoisia aineistoja, jotka eivät välttämättä sisällä "digital footprint" -termiä. Aineiston

hakemisen haasteena on termin monet eri merkitykset ja mahdollinen laajuus. Aineiston valintakriteerinä pyritään käyttämään viittausten määrää.

Tutkielmassa havaittiin, että digitaaliset jalanjäljet ovat yllättävän laajoja. Yksilön digitaalisten jalanjalkien merkitykseen vaikuttavat myös monet tekijät, koska niitä hyödyntävät yksilön lisäksi myös muutkin osapuolet. Pääasiassa yksilön digitaalisten jalanjalkien hallintaa motivoivat riskit. Hallinnalle tärkein yksittäinen asia on yksilön omat taidot, joihin voidaan kuitenkin vaikuttaa koulutuksen avulla. Lainsäädäntö taas pystyy vaikuttamaan koulutukseen sekä yksilöiden tietosuojaan. Tutkielmassa myös tunnistettiin kolme erilaista digitaalisten jalanjalkien hallintastrategiaa.

Tämä tutkielma rakentuu johdannosta, kolmesta sisältöluvusta sekä yhteenvedosta. Ensimmäisessä sisältöluvussa määritellään digitaalinen jalanjalki, keskitytään käsitteen kehitykseen sekä havainnollistetaan sen laajuutta kertomalla sen muodostumistavoista. Toisessa sisältöluvussa avataan yksilön digitaalisen jalanjäljen merkitystä kertomalla sen käyttökohteista ja riskeistä. Toinen sisältöluke pyrkii myös vastaamaan ensimmäiseen tutkimuskysymykseen. Kolmannessa sisältöluvussa käydään läpi yksilön digitaalisen jalanjäljen hallintaan liittyviä tekijöitä, esitetään kolme hallintastrategiaa sekä pyritään vastaamaan toiseen tutkimuskysymykseen. Viimeisessä luvussa esitetään johtopäätökset, yhteenvedo sekä jatkotutkimusaiheita.

## 2 DIGITAALINEN JALANJÄLKI

Tässä luvussa määritellään digitaalinen jalanjälki, käsitellään termin kehitystä sekä havainnollistetaan digitaalisen jalanjäljen muodostumistapoja. Ensimmäisessä alaluvussa pohditaan digitaalinen jalanjälki -termin tarkoitusta sekä määritellään se tutkimuksen kannalta. Toisessa alaluvussa käsitellään termin historiaa ja kehitystä. Kolmannessa alaluvussa käsitellään digitaalisen jalanjäljen muodostumista Thatcheria (2014) sekä Maddenia, Foxia, Smithiä ja Vitakia (2007) mukaillen jakamalla digitaalinen jalanjälki aktiiviseen ja passiiviseen niiden muodostumistavan mukaan. Sisältöluvun päätavoitteena on kartoittaa, mikä digitaalinen jalanjälki on ja mistä eri lähteistä digitaalinen jalanjälki muodostuu.

### 2.1 Määritelmä

Digitaaliset jalanjäljet ovat digitaalisia jälkiä, joita jätämme jälkeemme jokapäiväisessä elämässämme (Weaver & Gahegan, 2007). Laajimman määritelmän mukaan digitaalinen jalanjälki on mikä tahansa jälki, joka on tallennettu sensorein (Cheng, 2018). Tämä määritelmä on kuitenkin tutkielman kannalta turhan laaja. Tässä tutkielmassa ollaan enemmän kiinnostuneita digitaalisista jalanjäljistä, jotka ovat saatavilla esimerkiksi verkossa. Digitaaliset jalanjäljet ovat siis myös jäljitettävissä olevaa dataa ja informaatiota, jota tuotetaan verkossa toimiessa (BOUDLAIE, 2019; Buchanan, Southgate, Smith, Murray & Noble, 2017). Termi itsessään koostuu kahdesta sanasta: digitaalisuudesta ja jalanjäljestä. Seuraavissa kappaleissa pohditaan näiden sanojen merkitystä, miten eri lähteiden määritelmät vertautuvat toisiinsa nähden sekä mikä on tämän tutkielman kannalta sopivin määritelmä.

Digitaalinen jalanjälki -termin ensimmäinen puolisko on digitaalisuus. Brennenin ja Kreissin (2016) mukaan digitaalisuus tarkoittaa datamuotoa, jossa data on bitteinä eli nollina ja ykkösinä. Digitaalisuuden vastakohta on analogisuus, jossa data saa liukuvia arvoja. Bittimuodossa data on tarkempaa ja helpommin toistettavaa. Tämä mahdollistaa helpomman datan käsittelyn, jonka vuoksi



suurien datamäärien käsittely on myös digitaalisessa muodossa helpompaa kuin analogisessa. (Brennen & Kreiss, 2016) Miksi sitten ei kuitenkaan puhuta analogisesta jalanjäljestä? Analoginen jalanjälki on epäkäytännöllinen verrattuna digitaaliseen jalanjälkeen datan toistettavuuden vaikeuden takia. Juuri digitaalinen jalanjälki on tutkimusaiheena mielenkiintoinen, koska se on helposti saatavilla toisin kuin analoginen data. Digitaalinen jalanjälki on siis käytännössä dataa tietokonein luettavassa muodossa.

Jalanjälki on kirjaimellisesti jalkapohjan jättämä jälki, useimmiten jossain materiaalissa, josta se erottuu. Se on merkki, että olemme olleet jossain. Digitaalisen jalanjäljen tapauksessa olemme olleet kyber-fyysisessä tilassa (Zhang ym., 2010). Jalanjälkiä voidaan seurata, kuten metsästäjätkin seuraavat saaliinsa jälkiä. Kuitenkin kun useat jalanjäljet sekoittuvat keskenään seuraamisesta tulee haastavaa. Samoin verkon suuren käyttäjämäärän seurauksena yksilön jalanjälki on usein mitätön. Metsästäjät haluavat seurata vain tietyn saaliin jälkiä. Jalanjälkiä seurattaessa on pakollista kiinnittää huomiota jalanjäljen ominaisuuksiin. Samoin digitaalisista jalanjäljistä tulee merkityksellisiä, kun niiden jättäjästä saadaan enemmän tietoa. Herääkin kysymys, että ovatko sellaiset digitaaliset jalanjäljet, joita ei voida yhdistää yksilöön edes jalanjälkiä? Mikä tahansa sensorin keräämä data voi tuottaa arvoa sen kerääjälle, mutta tämän tutkielman osalta sellaiset digitaaliset jalanjäljet eivät ole kiinnostavia, jotka eivät ole yhdistettävissä jättäjänsä.

Osassa tutkimuksissa viitataan, että digitaalinen jalanjälki on ikään kuin profiili tai kokonaisuus, joka on löydettävissä verkosta (Feher, 2021; Micheli, Lutz & Büchi, 2018; Weaver & Gahegan, 2007). Tämä on ristiriitaista aikaisempiin kappaleisiin nähden, mutta se johtuu jalanjälki sanan kaksoismerkityksestä. Yksittäinen jalanjälki eli luonnollisesti yhden jalan jättämä jälki on yhtä paljon jalanjälki sanan merkityksessä kuin useiden jalanjälkien vana. Periaatteessa verkon kanssa toimiessa yksilö jättää kronologisesti jatkuvan jäljen omista toimistaan samoin kuin hän jättäisi luonnossa kulkiessaan.

Yksilön kokonaiseen digitaaliseen jalanjälkeen on kuitenkin lähes mahdollonta päästä käsiksi. Yksilö ei yleensä pääse käsiksi omaan digitaaliseen jalanjälkeensä, koska se sijaitsee hajallaan useiden osapuolien tietokannoissa, jotka eivät ole yhteydessä toisiinsa (Sjöberg ym., 2017). Weaverin ja Gaheganin (2007) mukaan jo 1960-luvulla Frank Horton sanoi tiedon pirstaleisuuden olevan yksi tärkeimmistä yksityisyyden varmistajista. Samoin kuin oikeista jalanjäljistä myöskään digitaalisista jalanjäljistä ei voida saada koko kuvaa ihmisestä (Micheli ym., 2018). Mitä enemmän digitaalisista jalanjäljistä voidaan päätellä, sitä kiinnostavampia ne kuitenkin ovat. Suurempi kokonaisuus on parempi kuin yksittäinen jalanjälki.

Ottaen huomioon tämän tutkielman tutkimuskysymykset, on tärkeää määrittää mistä digitaalisista jalanjäljistä on kyse, jotta aihealuetta saadaan rajattua. Kuten aikaisemmista kappaleista käy ilmi, tutkielmassa digitaalisille jalanjäljille on kriteerejä. Tutkielmalle mielenkiintoisimmat digitaaliset jalanjäljet ovat verkossa, jäljitettävissä sekä suurempia kokonaisuuksia. Tämä perustuu siihen, että ne kertovat enemmän yksilöstä. Näitä kuvastaa melko hyvin jo aiemmin mainitut

Boudlalien (2019) ja Buchananin ym. (2017) määritelmät. Vielä paremmin tätä mielenkiintoa kuvastaa kuitenkin Maddenin ym. (2007) määritelmä, jonka mukaan digitaalinen jalanjälki on verkossa saatavilla olevaa henkilökohtaista dataa.

## 2.2 Kehitys

Tässä alaluvussa käsitellään digitaalisen jalanjäljen kehitystä itsessään sekä termin kehitystä tieteellisessä kirjallisuudessa. Pohditaan milloin ensimmäiset digitaaliset jalanjäljet ovat syntyneet ja miten ne ovat kasvaneet vuosien saatossa eri teknologioiden johdosta. Tehdään myös katsaus termin historiaan tieteellisessä kirjallisuudessa.

Digitaaliset jalanjäljet ovat saaneet alkunsa, kun analogisista tietolähteistä siirryttiin digitaalisiin lähteisiin. Tätä tiedon muuttamista digitaaliseen muotoon kutsutaan digitoinniksi (Brennen & Kreiss, 2016). Varhaisimmat esimerkit digitaalisista jalanjäljistä ovat henkilötietojen digitointi 1970-luvulta alkaen. Aluksi digitaaliset jalanjäljet ovat siis olleet hajautettuna lähinnä yrityksissä ja virastoissa eri tietokannoissa. Digitaaliset jalanjäljet ovat olleet harvassa, ja vain valituilla on ollut niihin pääsy.

Ensimmäinen maininta ”digital footprint” -termistä Google Scholarista on Leen (2003) artikkeli, jossa tutkitaan jättääkö työyhteisössä toimiminen digitaalisia jalanjälkiä. Digitaalisia jalanjälkiä on ollut huomattavasti ennen vuotta 2003, joten herää kysymys siitä, että miksi vasta silloin digitaaliset jalanjäljet ovat tulleet ajankohtaiseksi. Syitä tähän ovat luultavasti digitaalisten jalanjälkien määrän kasvu sekä tarve seurata digitaalisten jalanjälkien jättäjiä.

Ajan myötä digitointi on yleistynyt, joten myös digitaalisten jalanjälkien määrä on kasvanut. Tätä digitoinnin yleistymistä kutsutaan digitalisaatioksi, jossa digitaaliset laitteet yleistyvät eri toimialoilla. Digitalisaation seurauksena dataa on tänä päivänä saatavilla lähes joka elämän osa-alueelta (Harjumaa ym., 2016). Seuraavaksi pohditaan syitä sille, miksi digitaalisten jalanjälkien määrä on kasvanut niin suureksi, että siitä on tullut tutkimuksen arvoinen aihe.

Mullistava teknologinen kehitys digitaalisen jalanjäljen kannalta on internet. Toki ennen internetiä henkilökohtaiset tietokoneet ovat antaneet kuluttajille hyvän mahdollisuuden tuottaa digitaalista dataa, mutta internetin nähdään tuoneen digitaalisen median massoille. Internetin myötä digitaaliset jalanjäljet eivät ole ainoastaan kasvaneet, vaan osa niistä on myös siirtynyt suljetuista verkoista avoimiin verkkoihin, toisin sanoen avoimiin tietolähteisiin. Aluksi internetin käyttäjät lähinnä kuluttivat sisältöä, jota vain hyvin harvat tuottivat ja hallitsivat (Hengstler, 2011). 1990-luvulla yksilön digitaaliset jalanjäljet ovat olleet kuitenkin sen verran harvassa, että tutkimukselle ei ole ollut vielä tarvetta. Digitaalisen tiedon saatavuus internetin välityksellä on mahdollistanut digitaalisten jalanjälkien seuraamisen, mikä lienee myös Leen (2003) artikkelin motiivi.

Sen sijaan yksilön digitaalisten jalanjälkien kannalta mullistava teknologinen kehitys on niin sanottu web 2.0. Web 2.0 on internetin kehitysvaihetta kuvaava termi, jolle ominaista on käyttäjien osallistumisen ja sisällöntuottamisen

kasvaminen (Hengstler, 2011). Maddenin ym. (2007) mukaan web 2.0 on muuttanut henkilökohtaisten tietojen luonnetta. Julkisen tiedon muuttuminen digitaalisiksi ja hakukoneiden tarkkuuden paraneminen on mahdollistanut kenelle tahansa pääsyn henkilökohtaisiin tietoihin, jotka olivat aiemmin vain ammattilaisten saatavilla. (Madden ym., 2007) Web 2.0:n myötä digitaalisten jalanjälkien määrä on kasvanut merkittävästi (Muhammad, Dey & Weerakkody, 2018). Tämä ilmiö on mahdollistanut yksilön digitaalisten jalanjälkien seuraamisen ja tutkimisen.

Tämän jälkeen termin käyttö on aluksi hiljalleen kasvattanut suosiotaan teellisessä kirjallisuudessa. Scopusin ”digital footprint” -haun analyysin mukaan julkaisujen määrä on kasvanut sitten Leen (2003) artikkelin vuosi vuodelta kymmeniä prosenttiyksiköitä pois lukien vuodet 2012 ja 2013, jolloin julkaisuja oli vain muutama vähemmän kuin vuonna 2011. Tämän jälkeen julkaisujen määrässä ei ole näkynyt hidastumisen merkkejä.

Teknologisen edistyksen ja siten myös digitaalisten sensoreiden lisääntymisen myötä digitaalisten jalanjälkien määrä on kasvanut viime vuosikymmenien aikana merkittävästi (Zhang ym., 2010). Älypuhelimet, jotka ovat sensoririkkaita tietokoneita, ovat olleet portti ihmisten arjen seuraamiseen, sillä lähes jokainen kantaa sellaista mukanaan (Stachl ym., 2020). Kasvokkain käytävät ystävien väliset kanssakäymiset ovat korvautuneet digitaalisella kommunikaatiolla. (Buchanan, Southgate, Scevak & Smith, 2018)

Muhammadin ym. (2018) mukaan sosiaalisen median käyttö on kasvanut eksponentiaalisesti, ja siitä on tullut keskeinen osa kuluttajien elämää. Muun muassa sen seurauksena ihmiset käyttävät älylaitteita, kuten älypuhelimia, tabletteja, älykelloja sekä älykotilaitteita, jotka voivat olla yhteydessä internetiin, mikä taas jättää palveluntarjoajille entistä enemmän seurattavia digitaalisia jalanjälkiä. (Muhammad ym., 2018) Älykellot ja älykoti ovat esimerkkejä esineiden internetistä eli laitteista, jotka ovat yhdistettynä internetiin. Esineiden internet tulee lopulta yhdistämään lähes jokaisen digitaalisen jalanjäljen verkkoon eli siitä on tulossa ubiikki (Cheng, 2018; Vervier, Zeissig, Lidynia & Ziefle, 2017).

Digitalisaation seurauksena dataa on saatavilla lähes joka elämän osa-alueelta (Harjumaa ym., 2016). Tämä myös tarkoittaa, että digitaalisen jalanjäljen tutkimuksella on paremmat mahdollisuudet levittäytyä laajemmalle. Weaverin ja Gaheganin (2007) artikkelissa seurataan yksilön toimintaa päivän mittaan kartalle havainnollistettuna. Tänä päivänä sensoreita on vielä enemmän, joten yksilön toiminnan seuraaminen on vieläkin tarkempaa, mikä on myös mahdollistanut tutkimuksen siirtymisen kartalta käyttäytymiseen. Digitaaliset jalanjäljet eivät ehkä anna koko kuvaa ihmisestä (Micheli ym., 2018), mutta ne kertovat kuitenkin enemmän kuin voisi kuvitella.

## 2.3 Muodostumistavat

Tässä alaluvussa kartoitetaan digitaalista jalanjälkeä sen muodostumistavan mukaan. Digitaalinen jalanjälki koostuu useista datajäljistä, joilla on myös hyvin

erilaisia lähteitä. Digitaalisen jalanjäljen muodostumistavoista on tärkeää olla tietoinen, koska se on olennainen osa oman digitaalisen jalanjäljen hallintaa. Lisäksi tässä alaluvussa käydään läpi joitain yleisiä datalähteitä, joiden avulla pyritään havainnollistamaan digitaalisen jalanjäljen laajuutta.

Digitaalinen jalanjälki voidaan jakaa aktiiviseen ja passiiviseen sen muodostumistavan mukaan (Madden ym., 2007; Thatcher, 2014). Näiden ero on Thatcherin (2014) mukaan datan jakamisen tarkoituksellisuus. Digitaalinen jalanjälki on aktiivinen silloin, kun käyttäjä jakaa tietoa tarkoituksellisesti. Passiivisen digitaalisen jalanjäljen osalta tiedon jakaminen ei ole tarkoituksellista, eikä siitä olla välttämättä edes tietoisia. (Thatcher, 2014) Seuraavaksi pohditaan tarkemmin näitä molempia muodostumistapoja.

Thatcherin (2014) mukaan aktiivinen digitaalinen jalanjälki syntyy, kun käyttäjä jättää jäljen tarkoituksellisesti. Maddenin ym. (2007) mukaan aktiivinen digitaalinen jalanjälki on taas henkilökohtaista dataa, jonka käyttäjä tarkoituksellisesti jakaa verkossa muiden saataville. Näille molemmille on siis yhteistä datajäljen tarkoituksellisuus. Yksi hyvin selkeä aktiivinen digitaalinen jalanjälki on sosiaalisen median julkaisut. Niiden tekeminen vaatii käyttäjältä harkintaa ja konkreettisia tekoja.

Aktiiviset digitaaliset jalanjäljet ovat itseilmaisua, omien tietojen jakamista. Muita aktiivisia digitaalisia jalanjälkiä ovat esimerkiksi jaetut kuvat, videot sekä sellaiset käyttäjätietojen jakamiset, joiden tiedetään olevan julkisia. Vaikkapa sosiaalisen median käyttäjän ylläpitäminen on usein tarkoituksellista toimintaa.

Thatcherin (2014) mukaan passiivinen digitaalinen jalanjälki syntyy muun muassa älypuhelimien jakaessa sijaintidataa ilman käyttäjän suoranaista aktiivista jakamista, ikään kuin taustalla. Maddenin ym. (2007) mukaan kyseessä on henkilökohtainen data, joka päättyy verkkoon muiden saataville ilman yksilön tarkoituksellista toimintaa. Passiivinen digitaalinen jalanjälki tarkoittaa siis datajälkeä, joka jää jälkeemme ilman, että olemme siitä välttämättä tietoisia tai ainakaan emme tietoisesti yritä jakaa sitä. Seuraavaksi käymme läpi erilaisia passiivisia lähteitä digitaaliselle jalanjäljelle.

Digitaaliset jalanjäljet eivät ole ainoastaan aktiivisen osallistumisen seurausta, vaan sitä kertyy myös algoritmien ja muiden internetin käyttäjien ansiosta (Micheli ym., 2018). Muiden käyttäjien julkaisut, joissa joku muu esiintyy tietämättään, ovat myös osa digitaalista jalanjälkeä. Esimerkki tällaisesta voi olla vaikkapa kuva, johon käyttäjä on merkitty.

Verkkoalustat keräävät käyttäjätietoja huomaamatta algoritmien avulla (Büchi, Lutz, ym., 2017; Micheli ym., 2018). Näitä ei kuitenkaan yleensä edes mielletä digitaalisiksi jalanjäljiksi, eikä näistä usein olla tietoisia (Vervier ym., 2017). Yksi tällainen huomaamaton digitaalisen jalanjäljen ilmentymismuoto on valvonta. Tätä kuvastaa muun muassa hyvin Husseinin (2018) yrityksen sisäisen tietoturvan ratkaisu. Kyseessä on lokitiedoston kerääjä, joka seuraa yrityksen verkossa tapahtuvaa toimintaa. Se kerää muun muassa käyttökellonajan, IP-osoitteen, käyttäjänimen, kirjautumisen todentamistavan, käytetyt resurssit ja selaintyyppin, eikä se vaadi käyttäjältä mitään tietoisuutta. (Hussein, 2018)

Digitaalisen jalanjäljen jakaminen aktiiviseen ja passiiviseen on ongelmallista. Kun puhelin luo passiivisesti digitaalista jalanjälkeä esimerkiksi sijaintidatan muodossa, käyttäjä kuitenkin yleensä vapaaehtoisesti pitää puhelinta mukanaan. Jos samainen henkilö olisi tietoinen, että puhelin kerää sijaintidataa olisiko digitaalinen jalanjälki silloin aktiivinen? Tällöin sama datalähde voi olla eri henkilöiden kohdalla joko aktiivinen tai passiivinen. Tällä ei kuitenkaan onneksi ole loppuen lopuksi mitään merkitystä, koska jaotteluidean takana on ehkä ajatus siitä, että olisi hyvä olla tietoinen etenkin passiivisista digitaalisista jalanjäljistä.

Aktiivinen ja passiivinen digitaalinen jalanjälki ovat myös ristiriitaisia tutkimuksessa. Sosiaalisen median toimintoja, kuten tykkäämistä, seuraamista tai kommentointia ei välttämättä koeta aktiiviseksi osallistumiseksi (Büchi, Lutz, ym., 2017). Monesti sosiaalisessa mediassa julkaistuihin kuviin on merkattu identiteetit, sijainti ja aika, joten niiden avulla voidaan päätellä, kuka tekee mitä, missä ja milloin (Golder & Macy, 2014). Tällainen metadata voi luultavasti paljastaa enemmän tietoa kuin käyttäjä välttämättä haluaa jakaa.

Digitaaliset jalanjäljet muodostuvat monista eri datalähteistä. Monet voivat olla yllättäviä, ja paljon dataa kerätään ilman käyttäjän tietoisuutta. Passiiviset digitaaliset jalanjäljet kasvavat muun muassa esineiden internetin myötä (Büchi, Lutz, ym., 2017). Samoin kasvavat myös aktiiviset digitaaliset jalanjäljet, koska ihmisten identiteetit pohjautuvat koko ajan enemmän ja enemmän siihen, mitä he julkaisevat internetissä. On kuitenkin selvää, että passiiviseen digitaaliseen jalanjälkeen on vaikeampi vaikuttaa.

### 3 YKSILÖN DIGITAALISEN JALANJÄLJEN MERKITYS

Tässä luvussa käsitellään digitaalisen jalanjäljen merkitystä yksilölle. Pohditaan mihin yksilön digitaalisia jalanjälkiä käytetään, minkälaisia sosiaalisia vaikutuksia digitaalisella jalanjäljellä voi olla yksilölle sekä mitä yksilöt ajattelevat omista digitaalisista jalanjäljistään. Lisäksi kartoitetaan tämänhetkisen tutkimuksen trendejä, joilla voi olla merkitystä yksilön näkökulmasta.

Sisältöluvussa hyödynnetään Feherin (2021) ajatusta digitaalisten jalanjälkien jakamisesta kahteen: data ja verkkominä. Data kuvastaa digitaalisten palveluiden oheistuotteena syntynyttä digitaalista jalanjälkeä, joka ei ilmene välttämättä käyttäjänä verkossa. Verkkominä taas kattaa muun muassa digitaalisessa ympäristössä käyttäjää kuvastavat profiilit. Tämä jakotapa on myös toisaalta rinnastettavissa aktiivisiin ja passiivisiin digitaalisiin jalanjälkiin (ks. alaluku 2.3).

Profilointi -alaluvussa keskitytään enemmän yksilön digitaalisiin jalanjälkiin datana. Verkkominä -alaluvussa taas yksilön digitaalisen jalanjäljen merkitys on enemmän sosiaalinen. Yksityisyys -alaluvussa pohditaan mitä yksityisyys merkitsee yksilölle digitaalisten jalanjälkien kannalta. Viimeisessä alaluvussa käydään läpi joitain tutkimussuuntauksi, jotka voivat olla merkittäviä yksilön kannalta sekä mitä yksilön digitaaliset jalanjäljet merkitsevät tutkimukselle.

#### 3.1 Profilointi

Tässä alaluvussa käydään läpi mitä profilointi on. Alaluku aloitetaan käsittelemällä profiilia eli profiloinnin tuotosta. Profiilit ovat osa yksilön digitaalista jalanjälkeä. Sen jälkeen pohditaan profilointia ja sen merkitystä. Profilointiin liittyy lähinnä uhkakuvia yksilön näkökulmasta, vaikka siitä seuraa myös toki hyviäkin asioita.

Profiili on kooste yksilöön viittaavasta datasta (Büchi ym., 2020). Jotta data saadaan yhdistettyä yksilöön, sillä täytyy olla jokin yhdistävä tekijä. Tietokantojen tapauksessa kyseessä voisi olla vaikkapa perusavaimen ja viiteavaimien

hyödyntäminen. Usein perusavaimena toimii yksilöllinen käyttäjänimi, joka taas toimii viiteavaimena muuhun käyttäjään liittyvään dataan. Mitä enemmän viitteitä yhdistyy toisiinsa, sitä enemmän tietoa saadaan käyttäjästä.

Irani, Webb, Li ja Pu (2009) toteavat, että profiilin luominen on usein vaatimuksena sosiaalisen median palvelujen käyttämiseksi. Profiilia luodessaan käyttäjä syöttää itsestään perustietoja sekä sivuston käyttötarkoituksen mukaisia tietoja. (Irani ym., 2009) Esimerkki tällaisesta sivustosta, jolla on jokin erityinen käyttötarkoitus, on musiikin suoratoistopalvelu Spotify, jonka käyttäjädata liittyy luonnollisesti musiikkiin. Tällöin tästä profiilista saadaan muun muassa käyttäjän musiikkimakuun liittyviä tietoja. Käyttäjät olettavat, että heidän tietonsa pysyvät ainoastaan asianomaisella alustalla ja että tietosuojakäytännöt ovat standardoidut (Irani ym., 2009). Todellisuudessa eri palveluiden tietojen yhdistäminen on kuitenkin usein mahdollista, ja tiedot voivat levitä alustojen välillä.

Sosiaalisen median kasvavan suosion takia ihmisillä on nykyisin useita käyttäjiä eri alustoilla (Irani ym., 2009; Malhotra, Totti, Meira, Kumaraguru & Almeida, 2012), kuten Facebook, LinkedIn, Twitter ja Youtube. Facebook profiilit sisältävät muun muassa arvokasta väestöstä kertovaa dataa, kuten nimet, syntymäajat, sijainti, kaveriyhteydet, organisaatiot, poliittiset ja sosiaaliset järjestöt sekä kulttuurilliset arvot (Golder & Macy, 2014). Malhotra ym. (2012) keräsivät puolestaan tutkimukseensa tietoaaineiston, joka koostuu käyttäjänimestä, näyttönimestä, kuvauksesta, sijainnista, profiilikuvasta sekä yhteyksien määrästä. Tämä kuvastaa hyvin profiilin sisältämiä datakenttiä.

Irani ym. (2009) yhdistelivät eri palveluiden julkisia profiileja keskenään ja keskimäärin yhdestä profiilista ilmeni neljä eri kenttää tietoa käyttäjästä. Jos käyttäjällä oli taas käyttäjä yli kahdeksassa eri palvelussa, kenttiä paljastui keskimäärin kahdeksan eli kaksinkertainen määrä. Pseudonyymeillä tai nimellä sosiaalisen median alustoilta hakemalla voi saada selville jopa 40 % käyttäjän digitaalisen jalanjäljestä. (Irani ym., 2009) Malhotra ym. (2012) ovat myös yhtä mieltä, että käyttäjänimi ja nimi ovat parhaimmat profiileja yhdistävät tekijät. Käyttäjien useiden verkkoidentiteettien yhdistäminen voi johtaa yksityisyys- ja turvallisuusuhkiin, kuten identiteettivarkaus tai vainoaminen (Malhotra ym., 2012).

Aikaisemmissa kappaleissa keskityttiin lähinnä avoimissa tietolähteissä oleviin profiileihin, joista kuka tahansa voi periaatteessa kerätä dataa. Seuraavaksi keskitytään kuitenkin järjestelmällisempään, usein organisaatioiden harjoittamaan toimintaan, jossa pyritään hyötymään yksilöiden digitaalisista jalanjäljistä.

Profilointi on järjestelmällistä ja tarkoituksellista yksilöön liittyvän datan tallentamista ja luokittelua (Büchi ym., 2020). Profiloinnissa luodaan uutta tietoa vanhan tiedon pohjalta analysoimalla yhtäläisyyksiä (Büchi ym., 2020; Nawi, Hussin, Ren, Norsaidi & Mohd Pozi, 2020). Suurin osa profiloinnista tapahtuu algoritmien avulla automatisoidusti. Ennen yksilön data ei ollut koostettuna vaan pieninä palasina, ja se oli haastavaa koostaa yhteen profiiliksi (Hengstler, 2011). Nykyisin sen sijaan oikeilla työkaluilla saatavilla olevista digitaalisista

jalanjäljistä voidaan koostaa erittäin tarkka profiili tai kuvaus käyttäjästä (Micheli ym., 2018).

Verkossa oleva data on usein ilmaista, helposti kerättävää ja sitä on paljon verrattuna perinteisiin datalähteisiin (BOUDLAIE, 2019). Käyttäjät usein syöttävät henkilökohtaisia tietojaan itse vapaaehtoisesti, mutta automatisoinnin avulla kootut päätelmät kertovat usein enemmän kuin käyttäjät osaavat kuvitellakaan (Büchi ym., 2020). Tätä automatisoitua toimintaa kutsutaan tiedonlouhinnaksi. Digitalisaatio trendinä on lisännyt tiedonlouhinnan avulla saatavan tiedon määrää huomattavasti (Cheng, 2018). Ihmiset eivät ole usein edes tietoisia, että tiedonlouhintaa on tapahtunut (Leonetti, 2014).

Muhammadin ym. (2018) mukaan sosiaalisen median palvelut, kuten Facebook ja Twitter, ovat uudelleen määrittäneet tavan, jolla yrityksen arvoa voidaan luoda. Teknologijätit hyödyntävät käyttäjien henkilökohtaista dataa, digitaalisia jalanjälkiä seuraamalla asiakkaitaan ja heidän käyttäytymistään tunkeilevasti ja ubiikisti. (Muhammad ym., 2018) Verkkoalustat keräävät käyttäjistä dataa ja koostavat niistä profiileja, jotka voidaan myydä kolmansille osapuolille, kuten mainostajille (Micheli ym., 2018; Nawi ym., 2020). Palveluiden tuottama arvo tulee niiden käyttäjäkunnasta ja sitä koskevasta datasta (Hengstler, 2011). Etenkin sosiaalinen media ja ilmaisapelit tarjoavat palveluitaan vastineeksi käyttäjädatasta (Montag & Rumpf, 2021). Käyttäjät antavat henkilökohtaisia tietojaan vastineeksi ilmaisista palveluista tajuamatta edes, kuinka paljon taloudellista hyötyä tiedot tuottavat muun muassa palveluntarjoajille, datankerääjille ja -välittäjille (Cheng, 2018). Jos jokin on ilmaista, niin silloin käyttäjä on tuote.

Profiloinnissa hyödynnettävät tietokonealgoritmit voivat ennustaa yksilön piirteitä, käyttäytymistä ja tarpeita digitaalisista jalanjäljistä (Hinds & Joinson, 2019; Sjöberg ym., 2017). Profiloinnin käyttökohteita voivat olla esimerkiksi psykologinen maalittaminen tai vaikuttaminen ihmisten toimintaan, kuten osto- ja äänestyskäyttäytymiseen (Stachl ym., 2020). Seuraavaksi käydään läpi aineistoa, jossa näitä käyttökohteita käsitellään.

Thatcher (2014) tuo ilmi artikkelissaan, että yritykset haluavat luoda lisäarvoa käyttäjilleen. Büchin ym. (2020) mukaan yksityinen sektori on kiinnostunut muun muassa ostokäyttäytymisestä, verkkosurffaamiskäyttäytymisestä ja sijaintidatasta. Näiden pohjalta voidaan luoda profiileja, joiden avulla mainoksia ja tuotesijoittelua on helpompi kohdentaa. (Büchi ym., 2020) Käyttäjä hyötyy tästä siten, että hän saa juuri hänelle suunnattua sisältöä, ja näin voidaan vaikuttaa yksilön ostokäyttäytymiseen.

Psykologinen maalittaminen taas tarkoittaa digitaalisten jalanjälkien pohjalta koottujen psykologisten profiilien avulla tapahtuvaa vaikuttamista asenteisiin, tunteisiin ja käytökseen psykologisia keinoja hyödyntäen (Matz, Appel & Kosinski, 2020). Esimerkiksi yksilön tunteisiin vetoamalla voidaan vaikuttaa äänestyskäyttäytymiseen.

Digitaalisista jalanjäljistä voidaan päätellä paljon yksilön taipumuksia. Berg, Burg, Gombović ja Puri (2020) selvittivät tutkimuksessaan, että pelkästään verkkokaupan asiakkaiden käyttöjärjestelmien ja sähköpostiosoitteiden perusteella voidaan arvioida heidän luottoluokitustaan. Esimerkiksi sellaiset asiakkaat,



joiden sähköpostiosoitteessa oli heidän nimensä, jättivät 30 % vähemmän todennäköisemmin maksamatta. Tämä tutkimus voi vähentää eriarvoisuutta, koska luottotiedottomilla mutta hyvillä digitaalisilla jalanjäljillä varustetuilla asiakkailla voi olla mahdollisuus luottoon. (Berg ym., 2020)

Büchin ym. (2020) mukaan datamäärien kasvaessa profilointi luo merkittävän voimaepätasapainon käyttäjien ja alustojen välille. Profilointi on rinnastettavissa valvontaan. Tästä on esimerkkinä Kiinan valtion suorittama datavalvonta ja sosiaalinen luottoluokitusjärjestelmä. Tällaiset profiloititoimet voivat vaikuttaa ihmisten käyttäytymiseen. Ne voivat rajoittaa itseilmaisua ja ihmisten monimuotoisuutta, jos yhä suurempi osa ihmisistä alkaa mukailemaan profiloitijärjestelmiä. Profiloinnista tietämättömät voivat kärsiä esimerkiksi siten, että he eivät saa lainaa tai töitä. (Büchi ym., 2020)

## 3.2 Verkkominä

Tässä alaluvussa tarkastellaan digitaalisten jalanjälkien sosiaalista näkökulmaa eli sitä, miten yksilön digitaalinen jalanjälki ilmenee muille verkossa. Digitaaliset jalanjäljet ovat merkittäviä esimerkiksi yksilön itseilmaisun sekä tulevaisuuden mahdollisuuksien kannalta. Seuraavaksi käydään läpi tarkemmin, miksi ne ovat merkityksellisiä.

Iänikuinen filosofinen kysymys ”kuka minä olen?” on muuttunut muotoon ”kuka minä olen verkossa?” (Feher, 2021). Verkkominä on se kokonaisuus, jona käyttäjä ilmenee muille verkossa. Se sisältää käyttäjän profiilit, häneen liittyvän tiedon sekä toiminnan eri alustoilla ja yhteisöissä. Käyttäjälle eduksi oleva verkkominä on ikään kuin brändi, joka kuvastaa hänen mielenkiinnon kohteita, saavutuksia ja taitoja (Buchanan ym., 2017). Digitaalinen esitys eli verkkominä ja tietoaineisto eli data ovat merkki olemassaolosta (Feher, 2021).

Fyysisissä ympäristöissä teemme päätelmiä muista ihmisistä esimerkiksi heidän äänensä, pukeutumisensa ja käytöksensä perusteella, kun taas verkossa päätelmämme pohjautuvat muun muassa Facebook profiileihin, kuva-albumeihin, twiitteihin ja siihen miten he kirjoittavat (Hinds & Joinson, 2019). Facebook profiileista käy ilmi sosiaalisesti merkittäviä asioita, kuten ystävyysuhteet, organisaatiot, poliittiset ja sosiaaliset liikkeet sekä kulttuuriset mieltymykset (Golder & Macy, 2014).

Digitaalinen jalanjälki mahdollistaa muun muassa nopean ”Google” -identiteetin ja pätevyyden varmistamisen (Buchanan ym., 2017). Hindsin ja Joinsonin (2019) mukaan luotamme usein internet-pohjaisiin päätelmiin tehdessämme muita koskevia päätöksiä, kuten kenen kanssa haluamme olla tekemisissä, kenet valitsemme kumppaniksemme tai kenet palkkaamme. Digitaaliset jalanjäljet saattavat olla ainoa lähde, jonka pohjalta saamme vaikutelman jostain henkilöstä. (Hinds & Joinson, 2019) Siispä puuttuva digitaalinen jalanjälki voi olla yhtä haitallinen kuin huonosti hoidettu sellainen (Buchanan ym., 2017).

Lapset ja nuoret rakentavat itselleen laajempaa verkkominää kuin mikään muu ihmisryhmä aiemmin, koska he ovat olleet verkossa pienestä asti (Buchanan

ym., 2018). Siispä heidän digitaaliset jalanjälkensä ovat hyvin laajat. Mitä enemmän verkkoon jättää jälkeä sitä todennäköisemmin se on jonkun löydettävissä ja hyödynnettävissä (Hengstler, 2011). Digitaaliset jalanjäljet voivat mahdollisesti vaikuttaa lasten tulevaisuuteen koulutus- ja työllistymismahdollisuuksien kautta (Buchanan ym., 2018). Joskus henkilöstöhallinto saattaa esimerkiksi tarkastaa työnhakijan sosiaaliset mediat sen varalta, että siellä on syitä jättää palkkaamatta (BOUDLAIE, 2019).

Feherin (2021) mukaan verkkominän hallinnan tavoitteena on optimaalinen itsensä mainonta verkossa sekä itseilmaisuuksien hallinta. Tämä verkkominän ilmentymän hallinta vaatii jatkuvaa vaivannäköä, ehkä jopa enemmän kuin verkon ulkopuolisen minän hallinta. Yksittäiset käyttäjät luovat visuaalisia esityksiä, henkilökohtaisia narratiiveja ja tarinoita osana digitaalista jalanjälkeään sosiaalisissa mediassa ja verkkoyhteisöissä, (Feher, 2021) toisin sanoen kuvin, videoin, tekstein. Verkkominä ei ole kuitenkaan pelkästään käyttäjän itsensä luomaa sisältöä. Muiden tuottaman datan merkitys on erityisen ilmeistä nettijulkaisuille, kuten vaikuttajille ja youtubettajille (Micheli ym., 2018).

Monilla ihmisillä on nykypäivänä eri verkkoidentiteettejä. Monet kuuluvat esimerkiksi eri alakulttuurien verkkoyhteisöihin. Jopa ympäristöissä, joissa käyttäjät pysyvät nimimerkkien takana, he usein muodostavat pitkäaikaisia ja pidettyjä identiteettejä sekä mainetta, josta heidän on vaikea luopua (Golder & Macy, 2014). Hindsin ja Joinsonin (2019) mukaan käyttäjät voivat esittää itsensä verkossa tavoin, jotka tekevät heistä houkuttelevampia kuin he oikeasti ovat. Internet tarjoaa yksilöille paljon mahdollisuuksia esittää itsensä eri tavoin vähäisistä itsensä paranteluista pahantahtoisiin identiteettihuijauksiin. (Hinds & Joinson, 2019)

Digitaalisten jalanjälkien sosiaalinen ja ammatillinen merkitys on kasvanut huomattavasti (Buchanan ym., 2018). Hengstlerin (2011) mukaan on tärkeää hallita osallistumista sosiaalisessa mediassa, jotta voidaan minimoida virheet ja maksimoida henkilökohtaiset sekä ammatilliset mahdollisuudet. Digitaalisten jalanjälkien sisällön hallinnasta on tullut ammatin kannalta kriittinen taito, jonka avulla ammatillista mainetta voidaan suojella. On esimerkiksi kannattavaa erottaa henkilökohtaiset ja ammatilliset profiilit, käyttäjänimet sekä verkostot toisistaan. Saman käyttäjänimen käyttäminen useilla alustoilla on hyvä asia ammatillisen profiilin näkökulmasta. (Hengstler, 2011) Tällä pyritään siihen, että ammatilliset profiilit ovat helposti löydettävissä.

Michelin ym. (2018) mukaan erilaiset sosiaaliset ryhmät voivat kokea järjestelmällistä hyötyä tai haittaa digitaalisten jalanjälkiensä takia. Ne, jotka käyttävät internetiä tehokkaasti voivat lisätä sosiaalista, taloudellista, kulttuurista ja henkistä pääomaansa. Hyvin sosiaalisen median alustat tuntevat käyttäjät osaavat paremmin räätälöidä viestinsä saavuttaakseen haluamansa yleisön ja maksimoidakseen näkyvyytensä verkossa. (Micheli ym., 2018) Yksilöt, joilla taas ei ole tietämystä digitaalisista jalanjäljistä ja suotuisasta verkko-olemuksesta, voivat joutua epäedulliseen asemaan (Wook, Mohamed, Noor, Muda & Zairon, 2019).

Siispä verkkominällä voi olla yhteiskunnallisia vaikutuksia uutena epätasavertaisuuden lisääjänä. Bergin ym. (2020) mukaan digitaalinen jalanjälki saattaa kehittyä

digitaaliseksi vastineeksi kalliille puvulle, jota ihmiset pitivät käydessään pankissa. Pääero on siinä, että digitaalisen jalanjäljen hallinta vaikuttaa huomattavasti enemmän päivittäiseen elämään kuin kalliin puvun pitäminen. (Berg ym., 2020)

### 3.3 Yksityisyys

Tässä alaluvussa käsitellään yksityisyyttä eli yksilön näkemystä siitä, mikä häntä koskeva tieto ei kuulu muille. Pohditaan myös yksityisyyteen vaikuttavia tekijöitä, kuten arkaluontoisen tiedon vuotamisesta koituvia seurauksia.

Yksityisyys yhdistetään yleensä lapsiin, vanhempiin, sukulaisiin, ystäviin ja muihin meille tärkeisiin ihmisiin (Feher, 2021). Maxwellin (2018) mukaan yksityisyys ja anonymiteetti eivät enää koske ainoastaan rikollisia tai rikkaita ja kuuluisia, vaan ne koskevat kaikkia. Henkilökohtaiset tiedot ovat helposti saatavilla verkossa riippumatta iästä tai väestöryhmästä. (Maxwell, 2018)

Yksityisyys on muun muassa oikeus jäädä rauhaan, oikeus hallita yksilöä koskevaa tietoa (Baruh, Secinti & Cemalcilar, 2017) ja oikeus tulla unohdetuksi (Büchi, Lutz, ym., 2017). Harjumaan ym. (2016) tutkimuksen haastateltavat halusivat pystyä päättämään omasta datastaan. He haluavat jakaa arkaluontoista dataa ainoastaan luotettavien organisaatioiden kanssa. (Harjumaan ym., 2016) Yksityisyyttä voisi sanoa hallinnan tunteeksi. Yksilö haluaa hallita imagoansa sekä tietoa, jonka kokee arkaluontoiseksi suojellakseen itseänsä.

Vervierin ym. (2017) mukaan ihmiset ovat huolissaan yksityisyydestään verkossa. Ihanteellisessa tilanteessa jakaessaan tietojaan käyttäjät ovat tietoisia kaikista riskeistä ja hyödyistä, jolloin he voivat tehdä järkipäisiä päätöksiä. Päätökset tehdään kuitenkin yleensä välittömän hyödyn pohjalta mahdolliset pitkäaikaiset riskit ja seuraamukset unohtaen. Ainoastaan vain kaikista ilmeisimmät riskit huomioidaan. (Vervier ym., 2017) Siispä pelkkä yksityisyydestä välittäminen ei takaa merkittäviä itsensä suojelutoimia (Büchi, Just, ym., 2017). Tämä risitiriittaisuus asenteiden ja varsinaisen käytöksen välillä tunnetaan yksityisyysparadoksina (Matz ym., 2020; Vervier ym., 2017).

Yksityisyysparadoksia selittävä tekijä on käyttäjien arviot yksityisyysriskeistä suhteessa tiedonjakamisen hyötyihin – riskejä vähätellään (Baruh ym., 2017). Vervierin ym. (2017) mukaan ihmiset painottavat haluavansa suojella dataansa sekä hallita kenellä siihen on pääsy, mutta silti sosiaaliset ja tekniset hyödyt vaikuttavat enemmän datan jakamiseen. Ryhmään kuulumisen hyöty on tärkeämpää kuin arkaluontoisen datan jakaminen. Lisäksi sovellusten kätevyys ja houkuttelevuus ovat myös tärkeämpiä. Sen sijaan mahdollinen rahallinen korvaus ei ole houkutteleva tutkimuksen mukaan. (Vervier ym., 2017)

Digitaalisiin jalanjälkiin liittyy yksityisyyden kannalta kielteisiä ilmiöitä. Ne liittyvät osaltaan myös ihmisten haluun pitää huolen yksityisyydestään eli siitä, mitä he jakavat itsestään. Esimerkkejä negatiivisista ilmiöistä ovat vainoaminen ja häirintä (Tengku Wook ym., 2019). Häirintää on myös esimerkiksi doksaaminen (engl. doxing) eli ihmisen tietojen luvaton vuotaminen usein

uhkailumielessä (Micheli ym., 2018). Uhkailun tarkoituksena voi olla verkkominnan maineen tuhoaminen.

Vainoamisen välttämiseksi on tärkeää osata arvioida tietojen arkaluontoisuutta. Pelletin, Shialesin ja Stavroun (2019) mukaan sosiaalisen median julkaisujen ja koneoppimisen avulla voidaan määrittää käyttäjän sijainti ilman GPS:ää. Lisäksi julkaisuihin voidaan huomata kaavamaisuuksia, joiden avulla käyttäjien sijainti voidaan ennustaa. (Pellet ym., 2019) Klassinen esimerkki sosiaalisen median julkaisujen arkaluontoisuudesta on lomamatkasta julkaiseminen, jolloin julkaisijan koti on tyhjillään. Tämän seurauksena murtoavarkaiden työ tehdään huomattavasti helpommaksi.

Lomamatkasta ilmoittamista tehdään kuitenkin, vaikka riskistä ollaan tietoisia. Sosiaalisen median käyttäjät sietävät yksityisyysriskejä, koska he hyötyvät itsensä jakamisesta esimerkiksi itseilmaisun ja ihmissuhteiden kehittämisen kautta (Baruh ym., 2017). Lisäksi henkilökohtaisten tietojen jakaminen ja paljastaminen on ihmissuhteiden kannalta olennaista, joten yksityisyyttä suojeleva käyttäytyminen voi olla sosiaalisesti haitallista (Büchi, Just, ym., 2017).

### 3.4 Käyttökohteita tutkimuksessa

Tässä alaluvussa käsitellään digitaalisten jalanjälkien nykyisiä ja tulevia käyttökohteita tieteellisestä kirjallisuudesta. Digitaalisilla jalanjäljillä on paljon käyttökohteita ja mahdollisuuksia. Merkittävin mahdollisuus on käyttäjien käytöksen tutkiminen ja ennustaminen. Lisäksi pohditaan yleisesti digitaalisia jalanjälkiä tutkimuksen näkökulmasta.

Golderin ja Macyn (2014) mukaan verkkoyhteisöistä kerätyt digitaaliset jalanjäljet mahdollistavat ihmisten käyttäytymisen ja sosiaalisen kanssakäymisen ymmärtämisen eri tavalla kuin aikaisemmin. Kerätty verkkodata voi tarjota tarkan kuvan päivittäisistä aktiviteeteista sekä sosiaalisten kanssakäymisten määrästä ja intensiteetistä. (Golder & Macy, 2014) Siispä Fernandezin, Levinsonin ja Rodebaughin (2012) mukaan digitaalisten jalanjälkien avulla voidaan havaita sellaisia henkilökohtaisia kokemuksia, joita on esitetty vain osittain mahdolliseksi huomata muiden toimesta. Muun muassa sosiaalista ahdistusta on mahdollista arvioida Facebook profiilien perusteella, sillä sosiaalisesta ahdistuksesta kärsiviä yhdistää esimerkiksi jotkin profiileista löydettävät piirteet, kuten suurempi määrä henkilökohtaisia tietoja. (Fernandez ym., 2012)

Stachlin ym. (2020) artikkeli taas demonstroi mahdollisuutta saada tietoa yksilön yksityisistä piirteistä passiivisesti älypuhelimella kerättyjen käyttäytymismallien pohjalta. Näitä ovat esimerkiksi sovellusten käyttö, musiikin kuuntelu, sosiaalisuus, kommunikaatio, liikkuminen sekä päivä- ja yöaktiivisuus. (Stachl ym., 2020) Hindsin ja Joinsonin (2019) mukaan digitaaliset jalanjäljet käytöksen ilmentymänä ovat helpommin saatavilla analysoitavaksi kuin muut käytöksen ilmentymät. Ne mahdollistavat tutkijoille persoonallisuusmittareiden yhdistämisen varsinaiseen käytökseen. (Hinds & Joinson, 2019)

Henkilöiden toimintaa voidaan ennustaa ja erittäin henkilökohtaisia tietoja voidaan louhia (Sjöberg ym., 2017). Tietokoneiden analyysit voisivat tarjota mahdollisuuksia netin tai laitteiden välityksellä kannustaa ihmisiä elämään terveemmin, löytämään tukiryhmiä ja vaikkapa kierrättämään (Hinds & Joinson, 2019). Tietokonealgoritmit voivat automaattisesti tunnistaa persoonallisuuspiirteitä käyttäjän digitaalisesta jalanjäljestä (Deeva, 2019; Hinds & Joinson, 2019). Niiden avulla on muun muassa mahdollista ymmärtää mielenterveyttä ja muita riskitekijöitä (Deeva, 2019).

Digitaalisista jalanjäljistä voi olla myös esimerkiksi hyötyä henkilöstön valitsemisessa, jos itse täytettyjen kyselyiden sijaan hyödynnettäisiin käyttäytymisdataa (Stachl ym., 2020). Etu olisi siinä, että käyttäytymisdata on luotettavampaa. Itseraportoinnin kautta saadun datan ongelmia luotettavuuden näkökulmasta ovat muun muassa taipumus vastata sosiaalisesti toivotulla tavalla, muistiharhat sekä kykenemättömyys itsetutkiskeluun (Montag & Rumpf, 2021). Lisäksi digitaaliset jalanjäljet mahdollistavat tutkimuksen suorittamisen missä vain ja milloin vain. Toisin kuin ohimenevät kasvokkain käydyt keskustelut, verkko-haastattelut tallentuvat täydellisellä tarkkuudella ja voivat säilyä ikuisesti (Golder & Macy, 2014).

Deevan (2019) mukaan tutkimuksen kannalta ongelmallista digitaalisten jalanjälkien käyttämisessä on se, että luokitellun datan kerääminen on aikaa vievää ja kallista. Lisäksi tällainen tieto voi olla myös luottamuksellista, mikä lisää datan käsittelyn vaatimuksia. (Deeva, 2019) Golder ja Macy (2014) toteavat, että pääsy yksityiseen dataan voi olla merkittävä haaste. Suurin osa verkossa olevasta datasta on yksityisten yritysten omistuksessa, ja he saattavat rajoittaa pääsyä suojellakseen asiakkaidensa yksityisyyttä. (Golder & Macy, 2014)

Harjumaa ym. (2016) tutkivat digitaalisen jalanjäljen mahdollisuuksia terveysdatan lähteenä. Tutkimuksen osallistujat pyysivät terveystietoja käsitteleviltä tahoilta omaa dataansa. Tutkimuksessa todetaan, että henkilökohtaisen datan kerääminen on haastavaa ja hidasta. Data on enimmäkseen huonossa muodossa, turhaa sekä ehkä joskus jopa väärää. Data ei ollut tarpeeksi sopivaa eikä uudelleen käytettävässä muodossa. Tutkimuksessa myös todetaan, että terveysdata on luotettavampaa ja parempaa, jos useamman datalähteen yhdistäminen on mahdollista. (Harjumaa ym., 2016)

Digitaalisen jalanjäljen käyttö luo mullistavan mahdollisuuden ymmärtää ihmisten käyttäytymistä (Nawi ym., 2020). Perinteisiin datalähteisiin verrattuna verkossa oleva data on usein ilmaista, helposti kerättävää ja sitä on paljon (BOUDLAIE, 2019). Tietokoneista tulee väijäämättä parempia tulevaisuuden käytöksen ennustajia kuin ihmisistä (Hinds & Joinson, 2019). Kaiken kaikkiaan digitaalisen jalanjäljen datalla on paljon potentiaalia, mutta työkalut ja palvelut vaativat lisää kehittämistä (Harjumaa ym., 2016).

## 4 YKSILÖN DIGITAALISEN JALANJÄLJEN HALLINTA

Tässä luvussa käsitellään yksilön digitaalisen jalanjäljen hallintaa. Aikaisempi sisältöluke pohjustaa tavallaan syitä miksi yksilö haluaa hallita omaa digitaalista jalanjälkeänsä. Ensimmäisessä alaluvussa pohditaan yksilön digitaalisen jalanjäljen hallintaan vaikuttavia tekijöitä sekä tarjotaan ratkaisuja hallintaan. Toisessa alaluvussa käsitellään läpikäydyn aineiston pohjalta koottuja yksilön digitaalisen jalanjäljen hallintastrategioita.

### 4.1 Taidot, koulutus ja lainsäädäntö

Digitaalisen jalanjäljen hallinta kattaa sekä oman datan suojaamisen että verkkominän ylläpitämisen. Se on yksityisyydestä huolehtimista. Se vaatii tämän vuosisadan taitoja, tietämystä ja tapoja, joita tarvitaan osallistumiseksi ja menestymiseksi teknologisessa maailmassa (Buchanan ym., 2018). Näihin taitoihin voidaan vaikuttaa esimerkiksi koulutuksen avulla. Lainsäädännön avulla taas taitojen merkitystä voidaan vähentää tai koulutusta voidaan lisätä. Tässä alaluvussa keskitytään hallinnan kannalta olennaisiin taitoihin, siihen liittyvään koulutuksen tarpeeseen sekä lainsäädännön mahdolliseen rooliin digitaalisen jalanjäljen hallinnassa.

Tietoturvariskit ovat merkittävin syy hallita digitaalisia jalanjälkiä. Hallinnan avulla voi vähentää ainakin riskiä joutua verkkorikollisuuden kohteeksi. Tietoturva ei ole ainoastaan datan suojelemisen kannalta tarpeellista, vaan esimerkiksi huono salasana voi paljastaa arkaluontoista tietoa, mikä voi vaikuttaa negatiivisesti verkkominään ja maineeseen (Feher, 2021). Oma apu on usein riittävä toimi tietoturvariskejä vastaan (Büchi, Just, ym., 2017). Tämä tarkoittaa yksilön itsensä suorittamaa digitaalisten jalanjälkien hallintaa.

Büchin, Justin ym. (2017) mukaan yleiset nettitaidot selittävät parhaiten tietoturvakäyttäytymistä. Nämä taidot mahdollistavat käyttäjille hyötymisen internetistä samalla minimoiden siihen liittyvät riskit. Sen sijaan heikommassa

asemassa olevat internetin käyttäjät, jotka eivät ole kehittäneet nettitaitojaan tai joilla on rajoittuneita käyttötaipumuksia, ovat todennäköisemmin alttiita tietoturvauhille. (Büchi, Just, ym., 2017)

Tietoturvakäyttäytymisessä on eroja väestön kesken. Sosiaalista mediaa käyttäessään nuoret eivät ole kovinkaan huolissaan siitä, miten yritykset ja kolmannet osapuolet käyttävät heidän tietojaan (Micheli ym., 2018). Nuoret eivät ole jatkuvasti tietoisia digitaalisista jalanjäljistään (Vervier ym., 2017). Vanhempien aikuisten nettitaidot ovat heikommät ja he harrastavat vähemmän yksityisyydensä suojelemista (Büchi, Just, ym., 2017). Nuoremmat henkilöt ovat taipuvaisempia jakamaan dataa voidakseen olla yhteydessä ikäistensä kanssa verrattuna keski-ikäisiin käyttäjiin, jotka ovat vastahakoisempia tämän suhteen (Vervier ym., 2017).

Diginatiiveille on koulutettu jo jonkin aikaa nettitaitoja. Buchananin ym. (2017) tutkimuksessa lapset ovat hyvin tietoisia digitaalisten jalanjalkien huonoista puolista, mutta he eivät ole tietoisia niiden myönteisistä mahdollisuuksista. Lapset ovat tietoisia heidän toimintansa ja julkaisujensa mahdollisista seuraamuksista. Tämän seurauksena he ovat muokanneet käyttäytymistään vähemmän näkyväksi ja alttiiksi verkossa. Tutkimuksessa lapset ovat tyytymättömiä siihen, että he eivät voi hallita mitä muut julkaisevat heistä. Esimerkiksi vanhemmat julkaisevat kuvia heistä vasten heidän tahtoaan. Lisäksi yksikään lapsista ei ollut saanut opastusta vanhemmiltaan siitä, miten he voisivat saada aikaan positiivisen digitaalisen jalanjäljen. (Buchanan ym., 2017) Tästä voitaisiinkin päätellä, että vanhempien sukupolvien nettikäyttäytyminen on ongelmallista diginatiivien näkökulmasta.

Büchin, Justin ym. (2017) mukaan pelkät asenteet ja tietoisuus eivät takaa sopivaa tietoturvakäyttäytymistä. Tämä on linjassa yksityisyysparadoksin (ks. alaluku 3.3) kanssa. Siispä pelkkä tietoisuus ei riitä, vaan käyttäjille tulee kouluttaa tietotaitoa, joka siirtyy alustalta toiselle. Digitaalisten taitojen kehittämisen tulee olla jatkuvaa alustojen vaihtumisen takia. Tarve koskee myös kaikkia, joten yhteiskunnan tulisi ottaa koulutuksesta vastuuta. (Büchi, Just, ym., 2017)

Pelkät nettitaidot eivät ole kuitenkaan välttämättä tarpeeksi digitaalisia jalanjalkia hallitessa. Ihmiset harvoin tajuavat minkälaista tietoa he itse asiassa jatkavat (Vervier ym., 2017). He saattavat esimerkiksi luulla arkaluontoista tietoa harmittomaksi, koska he eivät tiedä tarpeeksi tietoon liittyvistä prosesseista ja siinä mukana olevista osapuolista (Matz ym., 2020). Tiedon kaikki tulevaisuuden käyttökohteet eivät välttämättä ole jakamisen hyväksymisvaiheessa tiedossa, joten edes taitavat ja järkevät käyttäjät eivät voi täysin ymmärtää, mitä he hyväksyvät. (Matz ym., 2020) Tämän takia lainsäädännölle on tarvetta.

Tietotekniikka hallitsee ihmisten elämää jatkuvasti enemmän ilman, että huomaamme tilanteen vakavuutta ja mahdollista haittaa tietoturvalle (Cheng, 2018). Siitä kuka hallitsee digitaalista jalanjalkia, on tulossa merkittävä teknologinen ja yhteiskunnallinen ongelma, joka on kaikkien vastuulla (Buchanan ym., 2018; Sjöberg ym., 2017).

Cheng (2018) toteaa, että digitaalisen jalanjäljen tietoturvan pitäisi olla ihmisoikeus, koska digitaalisilla jalanjäljillä on korkea kaupallinen arvo, joka

saavutetaan tiedonlouhinnalla. Tiedonlouhinta on herättänyt huolta perusteettomasta valvonnasta, syrjinnästä ja vaikeasti havaittavasta tungettelusta, jotka ovat välttämättömiä internetin vaatimalle luottamukselle. Internetin käyttäjät eivät halua jakaa tätä tiedonlouhinnan avulla koostettua tietoa muiden, etenkin kolmansien osapuolien kanssa. (Cheng, 2018)

Chengin (2018) mukaan viranomaisten pitäisi puuttua tietoturvan suojelemiseen tai koulutukseen, jotta ihmiset voivat oppia pitämään itse huolen tietoturvastaan internetissä. Esimerkiksi Euroopan unionissa asiaan on puututtu ja sisäänrakennettu tietosuojaja on perusoikeus. Sisäänrakennetusta tietosuojasta esimerkki on älä seuraa -toiminto, joka käytettäessä viestittää palvelimelle, että käyttäjä ei halua hänestä kerättävien evästeitä. Toiminnossa on kaksi vaihtoehtoista asetusta: käytön mahdollistaminen sekä käyttö oletuksena. Arkaluontoisimmat tiedot vaativat toiminnon oletusasetusta. (Cheng, 2018)

Lainsäädäntö vaihtelee paljon maailmalla. Chengin (2018) mukaan Euroopan unionissa sisäänrakennettu tietosuojaja on perusoikeus, kun taas Yhdysvalloissa se ei ole. Tästä seuraa se, että EU:ssa yksilön digitaalisten jänjälkien kaupallinen käyttö ei ole sallittua ilman lupaa, kun taas Yhdysvalloissa se on. Älä seuraa -toiminnon suhteen EU:ssa on käytössä oletusasetus ja Yhdysvalloissa käytön mahdollistaminen. (Cheng, 2018)

Ylipäänsä asiat ovat Yhdysvalloissa toisin kuin Euroopassa. Leonetti (2014) mukaan datavalvontaa harrastaa yritysten lisäksi viranomaiset. Yhdysvaltain hallitus hyödyntää itse tiedonlouhinta. Datavalvontaa ja tiedonlouhinta perustellaan terrorisminvastaisena toimena. Se valvoo pääasiassa mahdollisten terroristien toimia, mutta samalla se tarkkailee myös muidenkin kansalaistensa tekemisiä. Tämän seurauksena kaikki valvonnan kohteena olevat kansalaiset ovat mahdollisia epäiltyjä. (Leonetti, 2014) Viranomaiset voivat siis toimillaan myös olla haitaksi yksilön yksityisyydelle ja digitaalisten jänjälkien hallinnalle.

## 4.2 Hallintastrategiat

Digitaalisten jänjälkien hallinnassa pyritään täyttämään yksilön sosiaaliset tarpeet sekä pitämään yllä tunnetta yksityisyydestä. Olennaista on siis, että käyttäjälle arkaluontoinen data on turvassa ja verkkominä kuvastaa käyttäjää hänen näkökulmastaan edullisesti.

Tämän alaluvun alle on koottu kolme aineistosta tunnistettua digitaalisten jänjälkien hallintastrategiaa. Ne eroavat muun muassa niiden toteuttamiseen vaadittavien taitojen osalta sekä yleisesti niissä suhtaudutaan teknologiaan eri tavoin. Digitaalisten jänjälkien välttäminen ei vaadi välttämättä juuri minkäänlaisia taitoja ja siinä teknologiaan suhtaudutaan yleensä uhkana. Positiivinen digitaalinen jänjälki taas vaatii enemmän teknologisia taitoja ja siinä teknologiaan suhtaudutaan mahdollistajana. Digitaalisten jänjälkien keskittäminen vaatii yksilöltä ainakin tällä hetkellä paljon toimia. Lisäksi siihen liittyy ehkä teknokratista ajattelua siinä mielessä, että keskittäjien oletetaan käytännössä olevan asian tuntijoita.



Nämä kolme strategiaa voidaan nähdä myös etenevän kronologisesti. Välttäminen on ollut pitkän aikaa perinteinen tapa hallita digitaalisia jalanjälkiä. Digitaalisten jalanjälkien yleistyessä teknologian hyödyntämisestä on tullut olennainen taito, joten positiivisen digitaalisen jalanjäljen voitaisiin nähdä olevan tätä päivää. Digitaalisten jalanjälkien keskittäminen on taas tulevaisuutta, kenties ehkä jopa utopistista tällä hetkellä.

#### 4.2.1 Digitaalisten jalanjälkien välttäminen

Digitaalisten jalanjälkien välttäminen on kaikista konservatiivisin hallintastrategia. Tämän strategian toteuttajat ovat usein huolestuneita yksityisyydestään (Baruh ym., 2017). Sen takana on usein ajatus teknologiasta uhkana. Mediassa toistuu huonosti hallittujen digitaalisten jalanjälkien seuraamukset (Buchanan ym., 2018). Näitä ovat esimerkiksi työpaikkojen menettämiset jonkin ristiriitaisen julkaisun takia tai verkkorikollisuuden kohteeksi joutuminen. Lisäksi aikaisemmat kokemukset tietoturvaloukkauksista ennustavat suojelevaa, pelokasta käytöstä (Büchi ym., 2017). Nämä ruokkivat ihmisten pelkoa teknologiaa kohtaan. Ajatellaan, että mitään pahaa ei voi tapahtua, jos digitaalisia jalanjälkiä ei jätä.

Pahimmillaan digitaalisten jalanjälkien välttäminen on kieltäytymistä digitaalisen teknologian käytöstä kokonaan. Tämä muuttuu kuitenkin päivä päivältä mahdollisemmaksi, koska digitaaliset teknologiat kasvattavat merkitystään arjessa. Lievempi ilmiö on muutosvastarinta, jossa uusiin teknologioihin suhtaudutaan kriittisesti ja vastahakoisesti. Muutosvastarinta on osaltaan tervettä, koska teknologiaan liittyy lähes aina riskejä. Kätevien mutta turvattomien digitaalisten palveluiden käytön sietäminen on kuitenkin usein osa arkea, koska käyttäjien on lähes mahdotonta olla tietoturvatietoisempia, jos he haluavat olla yhteydessä muiden kanssa (Feher, 2021).

Digitaalisten jalanjälkien välttäminen on perinteinen ja helpoin tapa hallita digitaalista jalanjälkeä. Buchanan ym. (2017) tuo artikkelissaan ilmi, että jopa lapset hyödyntävät tätä strategiaa. He ovat tietoisia tietoturvariskeistä, mutta heille ei ole vielä opetettu digitaalisten jalanjälkien hallinnan kannalta hyödyllisiä taitoja. (Buchanan ym., 2017.) Tämän strategian toteuttaminen on siis järkevää, jos käyttäjällä ei ole hallinnalle sopivia taitoja.

Tavoitteena digitaalisten jalanjälkien välttämässä on, että halutut tiedot ovat minimissään tai vain halutun yleisön nähtävillä. Tähän ratkaisuna ovat käytön välttämisen lisäksi myös tekniset ratkaisut, kuten evästeiden hallinta ja älä seuraa -toiminto (Büchi, Just, ym., 2017). Käyttäjät eivät kuitenkaan hallitse digitaalista jalanjälkeään ainoastaan yksityisyysasetuksilla ja harkitsevalla julkaisemisella, vaan myös poistamalla itsensä merkkaukset kuvista, jotka ovat ristiriitaisia tai vähemmän mairittelevia tai poistamalla julkaisut, jotka kuvastavat heitä negatiivisesti (Micheli ym., 2018).

Pysyvästi verkossa olevan yhteiskunnan ja verkkoympäristössä työnteke-  
misen vaatimuksen vuoksi ei ole järkevästi mahdollista vetäytyä verkossa ole-  
mista kokonaan (Montag & Rumpf, 2021). Henkilökohtaisten tietojen jakaminen  
ja paljastaminen on ihmissuhteiden kannalta olennaista, joten digitaalisten jalan-  
jälkien välttäminen voi olla sosiaalisesti haitallista (Büchi, Just, ym., 2017).

Yleisesti verkossa osallistumisen vähyys nähdään lähinnä epäedullisena toimintana (Lutz & Hoffmann, 2017). On kuitenkin hyvä muistaa, että mitä enemmän verkkoon jättää jälkeä sitä todennäköisemmin se on jonkun löydettävissä ja hyödynnettävissä (Hengstler, 2011).

#### 4.2.2 Positiivinen digitaalinen jalanjälki

Digitaalisten jalanjälkien hallinta on myös tiedon julkaisemista ja julkaistun tiedon seuranta (Büchi, Just, ym., 2017). Tässä strategiassa teknologiaan suhtaudutaan mahdollistajana. Digitaalisesta jalanjäljestä pyritään hyötymään mahdollisimman paljon. Positiivinen digitaalinen jalanjälki havainnollistaa muun muassa yksilön taitoja sekä saavutuksia ja signaloi teknologiaosaamista (Buchanan ym., 2017). Olennaista strategian toteuttamiselle ovat digitaaliset taidot, joiden avulla käyttäjä pystyy esittämään itsensä mahdollisimman hyvässä valossa.

Buchananin ym. (2018) mukaan positiivinen digitaalinen jalanjälki ei ole ainoastaan verkossa tai digitaalisessa muodossa oleva ansioluettelo. Se on konsepti verkko-olemuksesta, joka painottaa yhtenäisen, myönteisen ja tarkoituksellisen olemuksen luomisen tärkeyttä. Se on kommunikointi- ja hallintataitoja, jotka muuttavat digitaalisen jalanjäljen jostain, jota pitää välttää sellaiseen, jota yksilöt voivat hallita. Tämä sisältää tarkoituksellista verkko-olemuksen luomista ja ylläpitämistä, jonka tarkoituksena on esittää yksilön saavutuksia, taitoja, identiteettiä ja mielenkiinnon kohteita. (Buchanan ym., 2018)

Yhtenäinen verkko-olemus saavutetaan esimerkiksi käyttämällä eri alustojen profiilien kohdalla samoja käyttäjänimiä, jolloin ne ovat helposti löydettävissä. Toisaalta voi olla myös järkevää erottaa esimerkiksi vapaa-ajan profiilit ammatin kannalta tärkeistä profiileista käyttämällä eri käyttäjiä ja sähköposteja (Hengstler, 2011). Tällöin profiilien sisällöt pysyvät varmemmin yhtenäisinä niiden käyttötarkoitukseen perustuen.

Tarkoituksellinen ja myönteinen verkko-olemus saavutetaan taas harkintaa käyttämällä. Olennaista on tietää, mitkä asiat ovat myönteisiä. Esimerkiksi huojojen elintapojen (Büchi ym., 2020) tai rikollisen toiminnan näyttäminen ei ole kovin järkevää (Hengstler, 2011). Sen sijaan myönteisempien ominaisuuksien, kuten sellaisten harrastusten näyttäminen on toivottavaa, jotka voivat olla eduksi vaikkapa töitä hakiessa. On hyvä olettaa, että julkaisut tulevat kaikkien nähtäville (Hengstler, 2011).

Internet tarjoaa yksilölle laajat mahdollisuudet itsensä parantelulle eri työkaluin (Hinds & Joinson, 2019). Joidenkin digitaalisten jalanjälkien osien muokkaaminen on kallista tai ne vaativat perustavanlaatuisia tapojen muuttamista, mutta on myös olemassa osia, joita on helppo manipuloida (Berg ym., 2020). Toivottu verkko-olemus voidaan saavuttaa näiden avulla myös epärehellisesti, kuvaamalla itseään todellisuudesta poikkeavasti esimerkiksi kuvanmuokkauksella.

Verkko-olemuksen kannalta tärkeää on sen ylläpito. Pitää valvoa mitä jaa itsestään tai mitä sinusta jaetaan (Hengstler, 2011). On hyvä tarkastaa aika ajoin, mitä itsestä on saatavilla verkossa esimerkiksi hakemalla omia tietoja sekä tutkimalla omia profiileja (Feher, 2021).

On kannattavaa haalia nettitaitoja, joiden avulla omaa sisältöään voi esittää haluamallaan tavalla. Hyvin sosiaalisen median alustat tuntevat käyttäjät osaat paremmin räätälöidä viestinsä saavuttaakseen haluamansa yleisön ja he osaat maksimoida näkyvyytensä verkossa (Micheli ym., 2018). Bergin ym. (2020) mukaan digitaalinen jalanjälki saattaa kehittyä digitaaliseksi vastineeksi kalliille puvulle, jota ihmiset pitivät käydessään pankissa. Pääero on siinä, että digitaalisen jalanjäljen hallinta vaikuttaa huomattavasti enemmän päivittäiseen elämään kuin kalliin puvun pitäminen. (Berg ym., 2020)

### 4.2.3 Digitaalisten jalanjälkien keskittäminen

Viimeinen kirjallisuudesta tunnistettu hallintastrategia on tällä hetkellä utopistinen. Datan omistajuudesta on tullut digitaalisen jalanjäljen hallinnan kannalta olennainen kysymys (Matz ym., 2020). Strategialle olennaista on näkemys siitä, että käyttäjä omistaa itse tuottamansa datan, ja hänellä tulee olla oikeus hallita sitä. Kirjallisuudessa ratkaisuksi ehdotetaan datanhallintajärjestelmiä, joihin käyttäjät voivat keskittää kaiken oman datansa (Sjöberg ym., 2017).

Vähän tämän tyyppisiä digitaalisen jalanjäljen hallintapalveluita on kuitenkin ollut jo jonkin aikaa olemassa. On esimerkiksi palveluita, jotka hakevat eri tietokannoista tietoa halutuilla hakusanoilla. Tämän avulla yksilö pystyy seuraamaan tietojensa vuotamista. Sähköposti ja puhelinnumero riittävät paljastamaan yksilön tietojen altistumisen tällaisissa palveluissa (Maxwell, 2018). Malhotra ym. (2012) ehdottavat myös julkisten digitaalisten jalanjälkien koostamista, jotta käyttäjä pystyy paremmin ymmärtämään, mitä kaikkea hänestä on verkossa.

Keskittämiseen on myös muita muotoja, mutta tärkeintä on kuitenkin harkinnan käyttäminen palveluja käytettäessä. On järkevää käyttää ainoastaan tarpeellisia palveluita, jolloin profiileja on vähemmän ja niitä on helpompi hallita. Lisäksi keskittäminen on muun muassa ainoastaan yhden suuren palveluntarjoajan asiakkaana olemista. Tämä on kuitenkin toisaalta nykyisen trendin vastaista. Nykyisin ollaan siirtymässä niin sanottuun web 3.0:an, jossa käyttäjien väliset kolmansien osapuolten palvelut vaihdetaan käyttäjien hallitsemiin palveluihin.

Oman datan pitäminen yhdessä paikassa lisää hallinnan tunnetta (Harjuma ym., 2016). Sjöbergin ym. (2017) ratkaisu tähän on palvelu, johon yksilö voi tallentaa kaiken datansa. Siispä tavoitteena on vaihtaa organisaatiokeskeisyydestä ihmiskeskeiseen järjestelmään. Palvelun avulla käyttäjä voi esimerkiksi vaikuttaa datansa käyttöoikeuksiin. (Sjöberg ym., 2017)

Tämä ei ole kuitenkaan tällä hetkellä kovinkaan käytännöllistä. Sjöbergin ym. (2017) mukaan yksilön data ei ole yleensä yksilön saatavilla, vaan se sijaitsee eri yritysten tietokannoissa. Käyttäjät eivät voi helposti ottaa dataansa ja vaihtaa kilpailevalle palveluntarjoajalle. (Sjöberg ym., 2017) Ongelmallista on datan vaikean saatavuuden lisäksi sen käytettävyyttä, koska data on yleensä hyvin vaihtelevassa ja vaikeasti hyödynnettävässä muodossa.

Tämä ratkaisu vaatii toimiakseen muun muassa lainsäädännöllisiä muutoksia, minkä takia se tuskin tulee toteutumaan hetkeen. Lisäksi on kyseenalaista, että haluaako kovinkaan moni edes nähdä vaivaa hallita omaa dataansa, sillä tällainen ratkaisu vaatii käyttäjältä paljon tietämystä ja vaivaa. Massat vaikuttavat

tällä hetkellä olevan tyytyväisiä tilanteeseen, jossa suuryritykset hyödyntävät heidän dataansa. Tätä voidaan perustella yksityisyysparadoksilla (ks. alaluku 3.3).

#### 4.2.4 Luvun yhteenveto

Tässä alaluvussa tehdään yhteenveto alaluvun 4.2 hallintastrategioista. Luokiteltuja yksilön digitaalisen jalanjäljen hallintastrategioita vertaillaan yleisesti sekä taulukkoon (taulukko 1) pohjautuen. Taulukon hyödyntämisellä pyritään huomaamaan muun muassa hallintastrategioiden välisiä eroja ja yhtäläisyyksiä lähteiden avulla.

Yleisesti ottaen välttäminen on tehokkain, helpoin ja perinteisin hallintastrategia. Tätä strategiaa on hyödynnetty esimerkiksi jo ennen kuin digitaaliset laitteet ovat tulleet massojen käyttöön. Myös taulukko tukee tätä väitettä perinteisyydestä, sillä Hengstlerin (2011) artikkeli on taulukon vanhin, ja käsittelee hyvin paljon välttämismenetelmiä. Digitalisaation myötä digitaaliset palvelut ovat kuitenkin kasvattaneet merkitystään, joten niitä on välttämätöntä osata hyödyntää. Tämä tarkoituksellinen hyödyntäminen tarkoittaa positiivisten digitaalisten jalanjälkien luomista. Siispä nykyaikana yhteiskunnassa menestyminen vaatii positiivisen digitaalisen jalanjäljen.

Taulukosta on havaittavissa, että monet lähteet käsittelevät sekä välttämistä että positiivista digitaalista jalanjälkeä. Oikeastaan yksikään lähde ei käsittele ainoastaan jompaakumpaa hallintastrategiaa, vaan ne risteävät keskenään. Siispä itse asiassa osa positiivisen digitaalisen jalanjäljen luomista voi olla myös joidenkin jälkien tietoinen välttäminen. Nämä kaksi hallintastrategiaa voidaan kuitenkin erottaa luokittelussa melko selkeästi toisistaan digitaalisen jalanjäljen suhtautumistapoja vertailemalla. Olennainen ero on, että välttämässä digitaalisiin jalanjälkiin suhtaudutaan skeptisesti ja positiivisessa mahdollistajana.

Digitaalisiin jalanjälkiin suhtautumiseen vaikuttavat muun muassa yksilön digitaaliset taidot. Taitavat internetin käyttäjät osaavat hyödyntää digitaalisia jalanjälkiään itselleen edullisesti, joten he suhtautuvat niihin myös myönteisemmin. Hallintastrategian valintaan vaikuttavat siis sekä edellä mainitut taidot että myös kokemuksiin perustuvat asenteet, joiden pohjalta yksilön digitaalisiin jalanjälkiin suhtautumistapa muodostuu.

Digitaalisten jalanjälkien keskittäminen tavallaan yhdistää molempia aikaisempia hallintastrategioita. Keskittämisellä pyritään välttämään digitaalisten jalanjälkien päätymistä ei-halutuille osapuolille. Lähteissä teknologiaan suhtaudutaan myös ylipäänsä myönteisesti, koska digitaalisten jalanjälkien keskittämisen uskotaan tuottavan enemmän hyötyä. Lisäksi keskittämiseen liittyy individualistisia näkemyksiä, jotka ovat kenties keskeisimmässä osassa tätä hallintastrategiaa valitessa.

Aineisto ei ole kovinkaan kattavaa ottaen huomioon, että kutakin strategiaa kohti on hyödynnetty ainoastaan vain muutamia lähteitä. Lähteet jakautuvat kuitenkin melko tasaisesti hallintastrategioiden välille. Suoranaisesti hallintastrategioita käsitteleviä artikkeleita on hyvin vähän, joten pelkästään sellaisten artikkeleiden vertailu ei olisi kovinkaan mielekäästä. Siispä hallintastrategioiden

vertailu on haastavaa ja rajallista tällaisen taulukon muodossa, jonka artikkelit vain sivuavat strategioita. Taulukosta ei oikeastaan löydy eroavaisuuksia strategioiden välillä julkaisuvuosien perusteella, mikä on hieman yllättävää. Taulukon avulla strategioiden vertaaminen ei siis tue juurikaan väitettä strategioiden mahdollisesta kronologisuudesta. Jokainen hallintastrategia on vielä ajankohtainen, vaikka niiden kehitysasteissa on eroavaisuuksia, jotka eivät käy ilmi taulukosta. Aineisto lienee kuitenkin sopiva tämän laajuiselle tutkimukselle, koska hallintastrategioita onnistuttiin luokittelemaan ja kuvailemaan.

TAULUKKO 1 Hallintastrategioiden luokittelussa hyödynnetyt lähteet strategioittain

Lähde	Strategia	Näkökulma / Väite
Baruh ym., 2017	Välttäminen	Yksityisyys / Välttäminen on edellytys yksityisyydelle
Buchanan ym., 2017	Välttäminen	Lasten digitaalisten jalanjalkien hallintastrategiat
Montag & Rumpf, 2021	Välttäminen	Psyykinen diagnosointi digitaalisesti / Kaiken välttäminen on mahdotonta
Lutz & Hoffmann, 2017	Välttäminen	Verkko-osallistuminen / Välttäminen koetaan epäedulliseksi
Hengstler, 2011	Välttäminen / Positiivinen	Yksilön taidot vaikuttavat digitaalisten jalanjalkien hallintastrategian valintaan
Büchi ym., 2020	Välttäminen / Positiivinen	Profiloinnin vaikutus käyttäytymiseen / Huonoja elintapoja ei kannata näyttää
Büchi, Just ym., 2017	Välttäminen / Positiivinen	Taidot ovat tietoisuutta tärkeämpiä
Feher, 2021	Välttäminen / Positiivinen	Digitaalinen identiteetti ja verkko-minä / Kaiken välttäminen on mahdotonta, oman verkko-olemuksen seuraaminen on olennaista
Micheli ym., 2018	Välttäminen / Positiivinen	Digitaalinen eriarvoisuus
Buchanan ym., 2018	Positiivinen	Digitaalisten jalanjalkien hallinnan opettaminen osaksi opetussuunnitelmia
Hinds & Joinson, 2019	Positiivinen	Persoonallisuuden ennustaminen digitaalisista jalanjäljistä / Itsensä parantelu digitaalisia jalanjalkia muokkaamalla
Berg ym., 2020	Positiivinen	Digitaaliset jalanjäljet luottopisteytyksen perustana

Sjöberg ym., 2017	Keskittäminen	"Digital Me" -palvelu, johon koko digitaalisen jalanjäljen voi siirtää
Malhotra ym., 2012	Keskittäminen	Yksilön dataa koostava palvelu voisi antaa yksilölle paremman käsityksen tämän jalanjäljestä
Matz ym., 2020	Keskittäminen	Yksityisyys ja psykologinen maailmankuva / Datan omistajuuden vaikutus hallintaan
Maxwell, 2018	Keskittäminen	Digitaalisen jalanjäljen kooste voi vaikuttaa hallintaan
Harjumaa ym., 2017	Keskittäminen	Yksilön digitaalisten jalanjälkien koostaminen terveystietojaksi

## 5 YHTEENVETO

Tässä kandidaatintutkielmassa pyrittiin kirjallisuuskatsauksen avulla havainnollistamaan yksilön digitaalisten jalanjälkien laajuutta, merkitystä sekä hallintaa. Aihe on ajankohtainen, koska digitaaliset jalanjäljet vaikuttavat kaikkien elämään. Digitaalisten ratkaisujen merkitys on korostunut viime aikoina etenkin koronapandemian seurauksena. Elämän eri osa-alueet ovat muuttuneet entistä enemmän digitaalisiksi. Tämä tarkoittaa myös suurempia digitaalisia jalanjälkiä, joiden todettiin olevan yllättävän laajoja. Siispä tutkielmassa pyrittiin löytämään vastauksia seuraaviin tutkimuskysymyksiin:

- 1) *Minkälainen merkitys yksilön digitaalisella jalanjäljellä on?*
- 2) *Miten yksilön digitaalista jalanjälkeä voidaan hallita?*

Jokaisella yksilöllä on omat digitaaliset jalanjälkensä, joten niiden merkitys voi myös vaihdella henkilöittäin. Yksilön digitaalisten jalanjälkien merkitys on siis hyvin laaja konsepti. Tutkielmassa pyrittiin kuitenkin tunnistamaan kirjallisuudesta syitä, jotka voivat vaikuttaa merkitykseen. Näitä ovat profiloinnin tuomat uhkakuvat, verkkominän sosiaaliset vaikutukset, näkemykset yksityisyydestä sekä tutkimuksen tuottamat uudet hyödyt.

Varsinainen vastaus tutkimuskysymykseen voisi olla, että yksilön digitaalisen jalanjäljen merkitys on hyvin moninainen. Siihen vaikuttavat monet tekijät, mutta kaikille näille se on kuitenkin merkittävä hyödyn lähde. Ainoa kärsijä voi olla yksilö itse, mikä voi taas lisätä niiden merkitystä yksilön näkökulmasta.

Tutkielmassa keskityttiin nimenomaan yksilön digitaalisiin jalanjälkiin eikä siihen, mitä digitaaliset jalanjäljet merkitsevät yksilölle. Jälkimmäisestä löytyy ainakin aikaisempia laadullisia tutkimuksia. Kontribuutio tämän tutkimuskysymyksen osalta on ainakin se, että tiedettävästi missään muussa artikkelissa ei olla päädytty täsmälleen samanlaiseen kuvaelmaan yksilön digitaalisten jalanjälkien merkityksestä. Tutkielman tulos ei ole kuitenkaan erityisen mullistava, ja se tulee lähinnä aikaisempaa tutkimusta.

Tutkielmassa päädyttiin toisen tutkimuskysymyksen osalta siihen tulokseen, että yksilön digitaalista jalanjälkeä voidaan hallita pääasiassa yksilön

omilla toimilla. Yksilön digitaalisten jalanjälkien hallintaan vaikuttaviin taitoihin voidaan kuitenkin vaikuttaa koulutuksen avulla. Lisäksi lainsäädäntö voi tukea koulutusta sekä yksilön digitaalisten jalanjälkien hallintaa esimerkiksi tietosuojasetusten muodossa.

Tutkielman varsinainen kontribuutio on kuitenkin hallintastrategioiden tunnistamisessa. Tutkielmassa onnistuttiin luokittelemaan kolme erilaista hallintastrategiaa: digitaalisten jalanjälkien välttäminen, positiivinen digitaalinen jalanjälki sekä digitaalisten jalanjälkien keskittäminen. Digitaalisten jalanjälkien hallintaa on käsitelty aikaisemmissa tutkimuksissa, mutta hallintastrategioita ei ole luokiteltu tällä tavalla. Hallintastrategioita on mahdollista tunnistaa enemmänkin, koska teknologiaan suhtautumistapoja on myös enemmän kuin tutkielmassa on hyödynnetty. Tutkielman yksinkertainen jakotapa on hyvin vajavainen, ja tarvetta voisi olla tarkemmalle tarkastelulle.

Yksi tutkielmaa rajoittava tekijä on se, että digitaalinen jalanjälki on hyvin epämääräinen ja laaja käsite. Sitä voisi sanoa yläkäsitteeksi, jonka alle jää muun muassa tutkielmassakin hyödynnetty jakotapa passiiviseen ja aktiiviseen digitaaliseen jalanjälkeen. Esimerkiksi datanhallinta voitaisiin rinnastaa passiivisen digitaalisen jalanjäljen hallinnaksi. Siispä digitaalisen jalanjäljen tutkimuksessa voidaan hyödyntää todella montaa muuta tutkimussuuntausta.

Tässä kirjallisuuskatsauksessa hyödynnettiin myös sellaisia artikkeleita, joissa ei mainittu digitaalisia jalanjälkiä, mutta ne käsittelivät digitaalisen jalanjäljen alle jääviä aiheita. Selkeämpi kontribuutio olisi voitu saada tekemällä systemaattinen kirjallisuuskatsaus pelkän digitaalinen jalanjälki -termin pohjalta. Ongelmallista tässä olisi kuitenkin se, että muiden termien pohjalta voitaisiin saada jo jotain sellaista tietoa, mitä digitaalinen jalanjälki -termin alla ei olla vielä tutkittu. Lisäksi aineiston laatuun olisi voitu käyttää tiukempaa seulaa. Muutamat lähteet ovat heikkolaatuisia julkaisuja, mutta niitä on kuitenkin tarkasteltu kriittisesti.

Nämä molemmat tutkimuskysymykset ovat saman tutkielman alla myös rajoittavia tekijöitä, koska ne ovat niin laajoja. Lisärajukselle olisi ollut tarvetta. Jatkotutkimusaiheita voisi olla pelkästään jompaankumpaan tutkimuskysymykseen keskittyminen esimerkiksi systemaattisen kirjallisuuskatsauksen muodossa siten, että useammat selkeät teemat onnistuttaisiin luokittelemaan.



## LÄHTEET

- Baruh, L., Secinti, E., & Cemalcilar, Z. (2017). Online Privacy Concerns and Privacy Management: A Meta-Analytical Review: Privacy Concerns Meta-Analysis. *Journal of Communication*, 67.
- Berg, T., Burg, V., Gombović, A., & Puri, M. (2020). On the Rise of FinTechs: Credit Scoring Using Digital Footprints. *The Review of Financial Studies*, 33(7), 2845–2897.
- BOUDLAIE, H. (2019). DIGITAL FOOTPRINT IN WEB 3.0: SOCIAL MEDIA USAGE IN RECRUITMENT. *Ad-Minister : Revista de La Escuela de Administración, Universidad EAFIT*, 34, 131–148.
- Brennen, J. S., & Kreiss, D. (2016). Digitalization. Teoksessa *The International Encyclopedia of Communication Theory and Philosophy* (ss. 1–11). John Wiley & Sons, Ltd.
- Buchanan, R., Southgate, E., Scevak, J., & Smith, S. P. (2018). Expert insights into education for positive digital footprint development. *Scan: The Journal for Educators*, 37, 49–64.
- Buchanan, R., Southgate, E., Smith, S. P., Murray, T., & Noble, B. (2017). Post no photos, leave no trace: Children’s digital footprint management strategies. *E-Learning and Digital Media*, 14(5), 275–290.
- Büchi, M., Fosch-Villaronga, E., Lutz, C., Tamò-Larrieux, A., Velidi, S., & Viljoen, S. (2020). The chilling effects of algorithmic profiling: Mapping the issues. *Computer Law & Security Review*, 36, 105367.
- Büchi, M., Just, N., & Latzer, M. (2017). Caring is not enough: The importance of Internet skills for online privacy protection. *Information, Communication & Society*, 20(8), 1261–1278.
- Büchi, M., Lutz, C., & Micheli, M. (2017). Life online: The digital footprint gap. *In International scientific conference for the Partnership for Progress on the Digital Divide*.
- Cheng, F.-C. (2018). THE DO NOT TRACK MECHANISM FOR DIGITAL FOOTPRINT PRIVACY PROTECTION IN MARKETING APPLICATIONS. *Journal of Business Economics and Management*, 19(2), 253–267.
- Deeva, I. (2019). Computational Personality Prediction Based on Digital Footprint of A Social Media User. *Procedia Computer Science*, 156, 185–193.
- Feher, K. (2021). Digital identity and the online self: Footprint strategies – An exploratory and comparative research study. *Journal of Information Science*, 47(2), 192–205.

- Fernandez, K. C., Levinson, C. A., & Rodebaugh, T. L. (2012). Profiling: Predicting Social Anxiety From Facebook Profiles. *Social Psychological and Personality Science*, 3(6), 706–713.
- Golder, S. A., & Macy, M. W. (2014). Digital Footprints: Opportunities and Challenges for Online Social Research. *Annual Review of Sociology*, 40(1), 129–152.
- Harjumaa, M., Saraniemi, S., Pekkarinen, S., Lappi, M., Similä, H., & Isomursu, M. (2016). Feasibility of digital footprint data for health analytics and services: An explorative pilot study. *BMC Medical Informatics and Decision Making*, 16(1), 1–9.
- Hengstler, J. (2011). Managing Your Digital Footprint: Ostriches v. Eagles. *Education for a Digital World*, 2, 89–139.
- Hinds, J., & Joinson, A. (2019). Human and Computer Personality Prediction From Digital Footprints. *Current Directions in Psychological Science*, 28(2), 204–211.
- Hussein, H. A. (2018). A passive Digital footprint for Extranet environment. *Al-Qadisiyah Journal Of Pure Science*, 23(1), 152–158.
- Irani, D., Webb, S., Li, K., & Pu, C. (2009). Large online social footprints--an emerging threat. In 2009 International conference on computational science and engineering (Vol. 3, pp. 271-276). IEEE.
- Lee, L. L. (2003). Does your Community leave a Digital Footprint. Haettu 26.3.2022 osoitteesta [http://www.optimice.com.au/upload/Does\\_your\\_Community\\_Leave\\_a\\_Digital\\_Footprint.pdf](http://www.optimice.com.au/upload/Does_your_Community_Leave_a_Digital_Footprint.pdf)
- Leonetti, C. (2014). Bigfoot: Data Mining, the Digital Footprint, and the Constitutionalization of Inconvenience. *J. High Tech. L.*, 15, 260.
- Lutz, C., & Hoffmann, C. P. (2017). The dark side of online participation: Exploring non-, passive and negative participation. *Information, Communication & Society*, 20(6), 876–897.
- Madden, M., Fox, S., Smith, A., & Vitak, J. (2007, joulukuuta 16). Digital Footprints. *Pew Research Center: Internet, Science & Tech*. Haettu 22.3.2022 osoitteesta <https://www.pewresearch.org/internet/2007/12/16/digital-footprints/>
- Malhotra, A., Totti, L., Meira Jr, W., Kumaraguru, P., & Almeida, V. (2012). Studying user footprints in different online social networks. In 2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (pp. 1065-1070). IEEE.
- Matz, S. C., Appel, R. E., & Kosinski, M. (2020). Privacy in the age of psychological targeting. *Current Opinion in Psychology*, 31, 116–121.

- Maxwell, W. N. (2017). *An Analysis of the Size and Impact of Digital Footprints*. Brigham Young University.
- Micheli, M., Lutz, C., & Büchi, M. (2018). Digital footprints: An emerging dimension of digital inequality. *Journal of Information, Communication and Ethics in Society*, 16(3), 242–251.
- Montag, C., & Rumpf, H.-J. (2021). The Potential of Digital Phenotyping and Mobile Sensing for Psycho-Diagnostics of Internet Use Disorders. *Current Addiction Reports*, 8(3), 422–430.
- Muhammad, S. S., Dey, B. L., & Weerakkody, V. (2018). Analysis of Factors that Influence Customers' Willingness to Leave Big Data Digital Footprints on Social Media: A Systematic Review of Literature. *Information Systems Frontiers*, 20(3), 559–576.
- Nawi, A., Hussin, Z., Ren, C. C., Norsaidi, N. S., & Mohd Pozi, M. S. (2020). Identifying the Types of Digital Footprint Data Used to Predict Psychographic and Human Behaviour. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 12504 LNCS, 287–296.
- Pellet, H., Shiaeles, S., & Stavrou, S. (2019). Localising social network users and profiling their movement. *Computers & Security*, 81, 49–57.
- Sjöberg, M., Chen, H.-H., Floréen, P., Koskela, M., Kuikkaniemi, K., Lehtiniemi, T., & Peltonen, J. (2017). Digital Me: Controlling and Making Sense of My Digital Footprint. Teoksessa L. Gamberini, A. Spagnolli, G. Jacucci, B. Blankertz, & J. Freeman (Toim.), *Symbiotic Interaction* (Vsk. 9961, ss. 155–167). Springer International Publishing.
- Stachl, C., Au, Q., Schoedel, R., Gosling, S. D., Harari, G. M., Buschek, D., Völkel, S. T., Schuwerk, T., Oldemeier, M., Ullmann, T., Hussmann, H., Bischl, B., & Bühner, M. (2020). Predicting personality from patterns of behavior collected with smartphones. *Proceedings of the National Academy of Sciences*, 117(30), 17680–17687.
- Wook, T. S., Mohamed, H., Noor, S., Muda, Z., & Zairon, I. Y. (2019). Awareness of digital footprint management in the new media amongst youth. *Jurnal Komunikasi: Malaysian Journal of Communication*, 35(3), 407–421.
- Thatcher, J. (2014). Living on Fumes: Digital Footprints, Data Fumes, and the Limitations of Spatial Big Data. *International Journal of Communication (19328036)*, 8, 1765–1783.
- Vervier, L., Zeissig, E.-M., Lidynia, C., & Ziefle, M. (2017). Perceptions of Digital Footprints and the Value of Privacy: *Proceedings of the 2nd International Conference on Internet of Things, Big Data and Security*, 80–91.
- Weaver, S. D., & Gahegan, M. (2007). Constructing, Visualizing, and Analyzing a Digital Footprint. *Geographical Review*, 97(3), 324–350.

Zhang, D., Guo, B., Li, B., & Yu, Z. (2010). Extracting social and community intelligence from digital footprints: An emerging research area. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6406 LNCS, 4–18.