

Katsaus

# Sosiaalisen median tutkija ja houkuttelevat numerot

Salla-Maaria Laaksonen

**Mittarit ja numerot ovat oleellinen osa sosiaalisen median kulttuuria ja sieltä ladattavia tutkimusaineistoja. Ne kannustavat järjestämään ja rajaamaan aineistoa sekä tekemään tulkintoja digitaalisesta sosiaalisesta toiminnasta. Metriikat ovat kuitenkin sosioteknisessä prosessissa syntynyttä kulttuurista varantoa, johon tulisi suhtautua kriittisesti.**

Metriikat ovat olennainen ja läpitukeva osa sosiaalista mediaa. Kaikissa laajasti käytössä olevissa sosiaalisen median palveluissa sisältöjen ohessa esitetään jotakin numeerista tietoa. Numerot kertovat, onko viesti nähty, onko se herättänyt reaktioita tai onko sitä jaettu eteenpäin. Mastodonissa jaetaan tähtiä, Twitterissä

ja Instagramissa sydämiä, Facebookissa ja perinteisemmällä keskustelupalstoilla peukutuksia. Monilla alustoilla numerot seuraavat myös käyttäjän profiilia: ne kertovat, montako seuraajaa tilillä on, montako viestiä tili on lähettänyt tai mikä on käyttäjän keräämien mainepisteiden määrä.

Kuten *Kulttuurintutkimuksen* tässä numerossa tuodaan erilaisten esimerkkien avulla esiin, elämme laajemminkin mittaamisen kulttuurissa. Yhteiskunnan ja talouden tilaa seurataan tiiviisti erilaisilla numeerisilla mittareilla, sovellukset ja älykellot mittaavat ihmiskehon tilaa, ja yliopistojen menestystä mitataan

julkaisumäärillä ja yliopistorankingeilla. Onkin oikeastaan niin, että sosiaalinen media mittareineen sopii neoliberaaliin mittausyhteiskuntaan kuin nenä päähän – se on tämän yhteiskunnan tuote, jota digitalisaatio on entisestään kiihdyttänyt (ks. Beer 2019; Bolin 2011). Sosiaalisen median palveluiden tai laajemminkin digitalisen median tuottamaa mitattavaa dataa luonnehtii myös se, että sitä kertyy huomaamatta. Numeerisia jälkiä jää paitsi tietoisesta ja aktiivisesta toiminnasta, kuten tykkäyksistä, myös klikkauksista, kulutuspoluista, hakutoiminnoista ja ylipäättään sisällön parissa viettämästämme ajasta.

Tutkimuksen näkökulmasta kyse on *digitaalista jäljestä* (*digital trace data*): jäljet ovat numeerista tietoa ja erilaisia korrelaatioita, jotka syntyvät, kun käytämme erilaisia digitaalisia mediatuotteita ja järjestelmiä (Hepp ym. 2018; Freelon 2014). Ne antavat jollakin tavalla kuvaa siitä, mitä yksittäinen käyttäjä tai käyttäjäjoukko tekee eri digitaalisissa palveluissa, ja samalla tuottavat mitattavan representaation sosiaalisesta toiminnasta. Tämän vuoksi ne tarjoavat varsin houkuttavan aineiston tai aineiston kerroksen myös ihmistieteiden tutkijalle.

Tässä katsauksessa pohdin sosiaalisen median metriikoiden merkitystä tie-teellisessä analyysissä. Millä tavalla so-

siaalisen median metriikat vaikuttavat ja näkyvät tutkijan – myös kulttuurintutkijan – työssä? Miten niitä voisi käyttää hyödyksi tutkimuksessa, kuitenkin kriittinen näkökulma säilyttäen?

### **Numerot sosiaalisen median tutkimusaineistoissa**

Sosiaalinen media on erilaisten viestintää mahdollistavien verkkopalveluiden muodostama teknologinen ympäristö, mutta myös kulttuurinen konteksti (Hine 2000; Van Dijck, Poell & De Waal 2018). Siellä tapahtuvaa toimintaa ohjaavat paitsi teknologisen alustan toimintalogiikka ja käyttöliittymämerkaisut, myös kullekin alustalle syntyvä kulttuuri ja sosiaaliset käytänteet (Hine 2000). Sosiaalisen median palveluiden ytimessä on myös piilevä tallentamisen logiikka, jota on kutsuttu datafikaatioksi tai dataistumiseksi (*datafication*): prosessi, jossa teknologisten järjestelmien avulla sosiaalista toimintaa muunnetaan laskettavissa olevaksi dataksi, digitaalisiksi jäljiksi, jonka avulla ihmistoimintaa voidaan seurata ja analysoida jopa reaaliajassa (Andrejevic 2013; Couldry & Yu 2018; Van Dijck 2014). Sosiaalisen median mitattavuus on keskeinen osa koko teknologiaa ja niiden arkista käyttökokemusta (Bolin & Velkova 2020).

Datafikaatioon ja mittauskulttuurin nousuun kytkeytyy ajatuksia ja toiveita myös ihmistoiminnan ennakoimisesta: toivomuksena on, että kaiken kerättävän datan avulla saataisiin niin tarkkaa tietoa ihmistoiminnasta, että sen avulla voitaisiin ennustaa vähintäänkin nykyisyyttä (ns. *nowcasting*), mutta ehkä myös tulevaa ihmistoimintaa (esim. Mayer-Schönberger & Cukier 2013; Pentland 2014). On myös pohdittu, voisiko digitaalisten aineistojen avulla tehdä väestötason tilastollisia analyyseja ilman kalliita ja aikaa vieviä kyselytutkimuksia (ks. Cesare ym. 2018; Isotalo ym. 2019). Nämä toiveet osaltaan kiihdyttävät kehitystä, joka kannustaa toimijoita – yrityksistä tutkijoihin – keräämään mahdollisimman paljon dataa varmuuden vuoksi, mahdollista tulevaa käyttöä varten (ks. Fourcaden ja Healyn [2017] käsite *data imperative* ja Andrejevicin [2019] käsite *data capture*).

Tutkimuskäytössä sosiaalisen median aineisto sisältää tyypillisesti paitsi tekstiä ja kuvia, myös numeerista informaatiota. Numeromuodossa olevaa tietoa ovat esimerkiksi aikaleimat, erilaiset vuorovaikutuksien metriikat ja tieto sisällön käyttömääristä. Näiden muuttujien käyttäminen analyysin apuna on houkuttelevaa ja hyödyllistä varsinkin silloin, kun aineiston koko on suuri. Tutkijat saattavat muuntaa metriikkaa käy-

täntöön ja tulkita niitä esimerkiksi suosion ja kiinnostuksen mittareina, kannatuksen tai erimielisyyden osoittajana tai ajallisten muutosten osoittajana (esim. Bandari, Asur & Huberman 2012; Bruns & Stieglitz 2013). Numeerisen tiedon avulla aineistoa voi järjestää ja suodattaa esimerkiksi keskittymällä suosituimpiin viesteihin tai tiettyihin ajanjaksoihin, tai kenties ajanjaksoihin, jolloin keskustelu on ollut vilkkainta (esim. Ojala, Pantti & Laaksonen 2019). Numeerisen tiedon avulla voidaan myös tehdä monimutkaisempia analyysejä, esimerkiksi verkostoaalyysejä, esimerkiksi pyrkiä ymmärtämään aineistossa esiintyvän vuorovaikutuksen rakennetta. Tällainen lähestymistapa on tyyppillinen esimerkiksi Twitter-vuorovaikutuksen analysoinnissa (Marttila ym. 2016; Vainikka & Huhtamäki 2015), mutta samalla tavalla on mahdollista tarkastella esimerkiksi ihmisten ja dokumenttien välisiä yhteyksiä digitaalisessa yhteistyötilassa (Blaschke, Schoeneborn & Seidl 2012).

Numeromuodossa oleva data siis helpottaa sosiaalisen median aineiston käsittelyä ja tarjoaa kiinnostavia keinoja tuoda rakennetta sotkuiselta vaikuttavaan aineistoon. Tällaiseen metriikoiden käyttöön kannustetaan monissa digitaalisten menetelmien akateemisissa oppaissa (esim. Rogers 2017; Laaksonen

& Matikainen 2013; Laaksonen 2021). Metriikoita hyödyntävän tutkijan olisi kuitenkin hyvä olla tietoinen siitä, mitä numerot oikeastaan ovat ja miksi niiden houkuttavuutta on syytä pohtia myös kriittisesti.

### Objektiivisuuden tenho

Lukuisat yhteiskuntatieteilijät ovat kirjoittaneet numeroiden ja niihin perustuvan datan merkityksestä arjen ja yhteiskunnan hallinnassa. Numerot ja numeerinen data ovat vähemmän viattomia kuin ensinäkemältä vaikuttaisi. Ensimmäkin olemme sosiaalistuneet siihen, että numerot ovat lähtökohtaisesti luotettavuutta ja faktatietoa henkivä viestinnän tapa. Esimerkiksi tieteenhistorioitsija Theodore Porter (1995) korostaa luottamuksen numeroihin olevan yhteiskunnallisen käytännön, joilla rakennetaan objektiivisuutta ja puolueettomuutta. Tieteenfilosofi Ian Hacking (1990) puolestaan kirjoittaa kriittisesti numerofetisistä: numeroita kerätään vain koska voidaan, mutta siitä huolimatta niitä aletaan käyttää ongelmanratkaisun välineenä. Numeerinen tieto ja sen mahdollistava analytiikka omalta osaltaan luovat mittauskulttuureja ja mittauksen politiikkaa ja tietynlaisia tietämisen tapoja (Ruppert & Savage 2011).

Toisekseen numerot ja mittaukset

ovat tyyppillisesti myös performatiivisia eli todellisuutta tuottavia ja rakentavia: ne eivät vain puolueettomasti mittaa todellisuutta, vaan samalla luovat yhteiskuntaa, jota mitataan (Desrosières 1998; Beer 2016). Tämä pätee erityisesti esimerkiksi väestötiedon ja taloudellisen tiedon tuottamiseen (MacKenzie, Muniesa & Siu 2007): mittauksen yhteydessä rakennetaan kansaa tai numeerisen tiedon avulla piirretään vaikeasti ymmärrettävä talous jollakin tavalla olevalle. Performatiiviset numerot myös luovat todellisuutta, sosiaalisen median tapauksessa sosiaalisen median julkisuutta, kulttuurisia ilmiöitä ja toisaalta myös sosiaalista mediaa tutkimuskohteenä. Uudistamme ja vahvistamme tätä mittauskulttuuria joka kerta, kun tukeudumme numeroihin argumentoinnissa ja analyysissa.

Kolmanneksi eräs keskeinen toiminto, johon numerotietoa tarvitaan, on asioiden vertailu ja yhteismitallistaminen (Espeland & Stevens 1998). Eriaiset arvojärjestykset ja numeeriset arvioinnit ovat yksi keskeinen sosiaalisen median palveluiden toimintalogiikka (Scott & Orlikowski 2012). Järjestämistä tarvitaan, jotta palvelun uutisvirta osaa näyttää julkaisut jossakin järjestyksessä, jotta hotellinvarauspalvelu osaa vastata käyttäjän hakupyynnöön relevantilla tuloksilla, ja jotta tuhansien tai

miljoonien kuluttajien mielipiteet saadaan esitettyä jollakin tavalla kootusti. Monet mitattavista asioista eivät kuitenkaan yksiselitteisesti taivu numeroksi. Hotellivieraan tai elokuvayleisön mielipide on perimmiltään laadullinen kuluttajan kokemus, mutta yhteismitallistamisen prosessiin kuuluu väistämättä yhteisen standardin luominen, jotta vertailu tai muut matemaattiset operaatiot olisivat mahdollisia (ks. Espeland & Stevens 1998, 315–316). Tiedon puristaminen numeeriseen muotoon vaatii poikkeuksetta jonkinlaista tiedon tiivistämistä ja redusointia, mihin liittyy aina valintoja (esim. Gitelman & Jackson 2013; Laaksonen ym. 2020).

Mittauskulttuurien puristuksessa samalla normalisoidaan ajatusta siitä, että mittauksen kohteet, esimerkiksi juuri arviot tai sosiaalisen median sisällöt olisivat jotenkin mitattavissa ja verrattavissa. Sosiologi David Beer (2016) muistuttaaakin numeroiden usein kehystyvän paitsi havaintojen luokitteluun, myös laajemmin kilpailun ajatukseen kapitalismissa. Yhteismitallisuutta tarvitaan, jotta toimijoita ja ilmiöitä voitaisiin asettaa järjestykseen (ks. Espeland & Stevens 1998; Bowker & Star 2000). Kummallisella tavalla kyse on samasta järjestämisen logiikasta, oli kyse sitten pörssistä, yliopistorankingista tai sosiaalisen median sisältövirrasta.

## Häilyvät numerot

Kuten kaikki numeerinen tieto, myös sosiaalisen median aineistojen numerot ovat syntyneet jossakin kontekstissa, jota määrittävät teknologia, viestin sisältö ja käyttäjän käyttötilanteen konteksti (esim. Gitelman & Jackson 2013; Nelimarkka 2022). Tilanne, jossa yksittäinen klikkaus tai tykkäys syntyy, voi siis olla varsin moninainen, eikä niitä analysoivalla tutkijalla ole kovin hyviä keinoja päästä käsiksi sen jättäneen käyttäjän kontekstiin saati intentioihin. Jäljellä on vain digitaalinen jälki käyttötilanteesta, tallennettuna numerona järjestelmän lokiin. Metriikoiden häilyvyys pakenee helposti myös tutkijalta. He tulkitsevat digitaalisia jälkiä usein eri tavalla, ja numerotiedon perusteella tuotettuja tutkimustuloksia eli välttämättä validoida eikä tehtyjä tulkintoja perustella (Freelon 2014). Vaikka sosiaalinen media alkaa olla jo vanha ilmiö ja käyetyt mittaukset ovat yllättävän vakiintuneita palvelusta toiseen, ei ole olemassa konsensusta siitä, miten mittareita pitäisi tulkita.

Sosiaalisen median metriikoiden tulkintaa hankaloittaa lisäksi myös se, että moninaiset sosiaaliset kokemukset ja reaktiot typistyvät käyttöliittymän tarjoamien mahdollisuuksien kautta niihin toiminnan muotoihin, joita on tar-

jolla. Siksi sosiaalisen median metriikoissa on vielä tavallistakin enemmän sitä ongelmaa, että ne yhteismitallistavat asioita, jotka eivät ole yhteismitallisia. Tutkimuksissa on esimerkiksi osoitettu, että käyttäjillä on kymmeniä eri tapoja käyttää Facebookin tykkäyspainiketta (Eranti & Lonkila 2015; Hayes, Carr & Wohn 2016; Sumner, Ruge-Jones & Alcorn 2018). Tykkäys voi olla kirjaimellisesti tykkäys, mutta myös ironinen viesti, kuittaus viestin näkemisestä tai sillä voi olla muita vuorovaikutuksellisia funktioita (Haeyes ym. 2016). Tämän vuoksi on mahdotonta tyhjentävästi sanoa, mitä oikeastaan tarkoittaa, jos jokin viesti on saavuttanut tuhansia reaktiota, tai jos politiikolla on kymmeniä tuhansia seuraajia. Esimerkiksi Yhdysvaltojen entisen presidentin Donald Trumpin yli 87 miljoonalla Twitter-seuraajalla on todennäköisesti varsin erilaisia syitä seurata häntä.

Reaktioita aidompi kiinnostuksen mittari olisi kenties tieto katsoja- tai lukijamääristä, mutta julkisesta sosiaalisen median aineistosta ei tyypillisesti saa täsmällistä tietoa kävijämääristä. Kävijämäärät ovat tarjolla profiilien ylläpitäjälle, mutta eivät aineistoa lataavalle tutkijalle. Poikkeuksen muodostavat YouTube ja Wikipedia, joissa näyttökerrat ovat suoraan saatavilla osana aineistoa, ja toisaalta Google, joka tar-

joaa indeksitasoista tietoa hakutermien suosiosta.

Sosiaalisen median aineistojen metriikoihin vaikuttavat myös alustojen sisältöä järjestävät algoritmit. Palveluiden algoritmit tyypillisesti järjestävät sisältöä käyttäjien syötteissä niin, että suosittu sisältö saa enemmän näyttökertoja ja sitä suositellaan yhä uusille käyttäjille (Gillespie 2014). Siksi suositut sisällöt ja suositut käyttäjät saavat yhä enemmän näkyvyyttä alustan informaatioarkkitehtuurin tukemana, eivät välttämättä siksi että ne olisivat aidosti kiinnostavia (ns. *Matthew-efekti*, Merton 1968). Tässä mielessä sosiaalisen median metriikat ovat aina sosiaalisen toiminnan ja teknologisen järjestelmän yhteistuotoksia.

Numeroihin vaikuttaa myös aineiston keräämistapa ja arkistoinnin prosessit (Pink ym. 2018). Esimerkiksi Ho (2020) osoitti, että Facebookin rajapinnasta kerätty aineisto on alustan puolelta esikarsittua, ja tutkijan saaman aineiston vuorovaikutusmetriikoiden lisäksi vinoutuneita ovat myös otoksen julkaisu-tyypit, aiheet ja sentimenttijaakauma. Omassa tutkimuksessani paljastui vastikään, että CrowdTangle-niminen analytiikkatyökalu, joka tarjoaa pääsyn Metan palveluiden sisältöjen lataamiseen, ei välttämättä päivitä numeroa vastaamaan sosiaalisen median palve-

luissa tapahtuneita muutoksia. CrowdTangle-datassa näkyvä kommenttien määrä on toistuvasti suurempi kuin itse palvelussa näkyvä numero. Ilmeisesti analytiikkatyökalu raportoi joko keräyshetken todellisuutta tai sitten kyseisen sisällön historian korkeinta arvoa. Varma vastaus kysymykseen ei selvinnyt edes CrowdTanglen tuen kanssa keskustelemalla.

Tärkeä rooli onkin myös sillä, minikälaisten työkalujen avulla aineistoja käytämme ja analysoimme ja miten niiden käyttöliittymät on suunniteltu (esim. Jia ym. 2022; Madsen 2015). CrowdTangle-esimerkki on erinomainen muistutus myös siitä, että usein päästäksemme käsiin aineistoihin käytämme erilaisia palveluita, joiden toimintalogiikka ei välttämättä ole selkeä tai avoin.

Palveluissa saatetaan hyödyntää erilaisia visualisointeja, jotka pohjautuvat numeerisiin aineistoihin, mutta joiden esitysmuotoa ohjaavat muut kuin tieteelliset kriteerit (Rupper & Savage 2011; Laaksonen & Pääkkönen 2020). Tämä pätee niin erilaisiin analytiikkatyökaluihin, sosiaalisen median palveluiden omiin työkaluihin ja rajapintoihin kuin erilaisiin arkistoihinkin. Palveluista saatavat numerot ja kuvaajat vaikuttavat kuitenkin yhtä luotettavilta ja objektiiviselta kuin numeerinen tieto yleensäkin.

## Lopuksi

Kuten sosiaalisen media itse, myös sosiaalisen median aineistot ovat sosioteknisen prosessin muovaamia, yhteisen rakentumisen tulosta (Gerlitz & Rieder 2018; Lievrouw & Livingstone 2006). Ne näyttäytyvät selkeinä ja objektiivisina totuuksina (Porter 1995), mutta – kuten kaikki data – ovat todellisuudessa kulttuurista varantoa, joka täytyy tuottaa, muovata ja tulkita (Gitelman & Jackson 2013). Muovaaminen ja tulkinta jatkuvat myös tutkijan työpöydällä, kun tutkija tekee omaan analyysiprosessiinsa liittyviä päätöksiä.

Myös sosiaalisen median aineistoja käyttävälle tutkijalle on oleellista kysyä Desrosièresin (1998) esittämä kysymys: mitataanko maailmaa sellaisena kuin se on, vai luodaanko mittaamalla tai olemassa olevia mittareita hyödyntämällä todellisuutta? Esimerkiksi suosion logiikka on yksi sosiaalisen median määrittelevimmistä piirteistä (Van Dijck & Poell 2013). Kiinnostavalla tavalla pelkästään numeroiden läsnäolo sosiaalisen median aineistossa kannustaa myös tutkijaa toisintamaan tätä logiikkaa aineistoa käyttäessään. Numeerisessa muodossa esitetty informaatio johdattaa tulkitsijansa helposti pelkistämään ja yksinkertaistamaan – esimerkiksi ajattelemaan tykkäyksiä juuri suosion mitta-

rina. Onko sosiaalisen median sisältöjä analysoivan tutkijan relevanttia ja validia keskittyä eniten reaktioita herättäneisiin viesteihin?

Tämän katsauksen viesti ei ole väittää, ettei numeroita pitäisi käyttää lainkaan sosiaalisen median aineiston analyysissä tai niiden rajaamisessa. Olisi kuitenkin tärkeää, että tutkijat suhtautuisivat numeroihin sosioteknisinä konstruktioina – kriittisellä otteella ja pyrkimällä ymmärtämään sitä sosioteknistä kontekstia, jossa ne ovat syntyneet. Ymmärrys sosiaalisen median teknologisista ja kulttuurisista mekanismeista auttaa asettamaan kontekstiinsa myös luotettavuutta ja varmuutta huokuvat numerot. Näitä reunaehtoja purkavaa tutkimusta tehdään paljon erikseen digitaalisten menetelmien ja kriittisen alustatutkimuksen kentällä (esim. Rieder ym. 2015; Gerlitz & Rieder 2018; Ho 2020), mutta omalta osaltaan siihen voi osallistua jokainen tutkimuksensa valintoja ja rajoituksia pohtiva tutkija.

#### LÄHTEET

Andrejevic Mark (2013) *Infoglut: How too Much Information is Changing the Way we Think and Know*. New York: Routledge.

Andrejevic Mark (2019) Automating surveillance. *Surveillance & Society* 17(1/2), 7–13.

Bandari, Roja, Asur, Sitaram & Huberman, Bernardo A. (2012) *The Pulse of News in Social Media: Forecasting Popularity*. <http://arxiv.org/abs/1202.0332> (Tarkistettu toukokuussa 2023)

Beer, David (2016) *Metric Power*. Basingstoke: Palgrave Macmillan. <https://doi.org/10.1057/978-1-137-55649-3>

Beer, David (2019) *The Data Gaze: Capitalism, Power and Perception*. London: SAGE. <https://doi.org/10.4135/9781526463210>

Blaschke, Steffen, Schoeneborn, Dennis & Seidl, David (2012) Organizations as networks of communication episodes: Turning the network perspective inside out. *Organization Studies* 33(7), 879–906. <https://doi.org/10.1177/0170840612443459>

Bolin, Göran (2011) *Value and the Media. Cultural Production and Consumption in Digital Markets*. Farnham: Ashgate.

Bolin, Göran & Velkova, Julia (2020) Audience-metric continuity? Approaching the meaning of measurement in the digital everyday. *Media, Culture & Society* 42(7–8), 1193–1209. <https://doi.org/10.1177/0163443720907017>

Bowker, Geoffrey & Star, Susan L. (2000) *Sorting Things Out: Classification and Its Consequences*. Cambridge, MA: MIT press.

Bruns, Axel & Stieglitz, Stefan (2013) Towards more systematic Twitter analysis: Metrics for tweeting activities. *International Journal of Social Research Methodology* 16(2), 91–108. <https://doi.org/10.1080/13645579.2012.756095>

Couldry, Nick & Yu, Jun (2018) Deconstructing datafication's brave new world. *New Media & Society* 20(12), 4473–4491. <https://doi.org/10.1177/1461444818775968>

Desrosières, Alain (1998) *The Politics of Large Numbers: A History of Statistical Reasoning*. Harvard University Press.

Eranti, Veikko & Lonkila, Markku (2015) The social significance of the Facebook Like button. *First Monday* 20(6). <https://doi.org/10.5210/fm.v20i6.5505>

Espeland Wendy N. & Stevens Mitchell L. (1998) Commensuration as a social process. *Annual Review of Sociology* 24(1), 313–343.

Fourcade, Marion & Healy, Kieran (2017) Seeing like a market. *Socio-Economic Review* 15(1), 9–29.

Gerlitz, Carolin & Rieder, Bernhard (2018) Tweets are not created equal: Investigating Twitter's client ecosystem. *International Journal of Communication* 12, 528–547.

Gitelman, Lisa & Jackson, Virginia (2013) Introduction. Teoksessa Lisa Gitelman (toim.) *“Raw Data” is an Oxymoron*. Cambridge, MA: MIT Press.

Hacking, Ian (1990) *The Taming of Chance*. Cambridge: Cambridge University Press.

Hayes, Rebecca A., Carr, Caleb T. & Wohn, Donghee Yvette (2016) One click, many meanings: Interpreting paralinguistic digital affordances in social media. *Journal of Broadcasting & Electronic Media* 60(1), 171–187.

Ho, Justin Chun-Ting (2020) How biased is the sample? Reverse engineering the ranking algorithm of Facebook's Graph application programming interface. *Big Data & Society* 7(1). <https://doi.org/10.1177/2053951720905874>

Isotalo, Veikko, Laaksanen, Salla-Maaria, Pöyry, Essi & Jallinoja, Piia (2019) Sosiaalisen median ennustekyky kaupan myynnissä – esimerkkinä vegaanisi ja vegaanisit ruuat. *Kansantaloudellinen aikakauskirja* 115(1), 91–112.

Jia, Haiyan, Miller, Larrisa I., Hicks, Jessica, Moscot, Ethan, Landberg, Alissa, Heflin, Jeff & Davison, Brian D. (2022) Truth in a sea of data: adoption and use of data search tools among researchers and journalists. *Information, Communication & Society* 27.11.2022. <https://doi.org/10.1080/1369118X.2022.2147398>

Laaksanen, Salla-Maaria (2021) Sosiaalinen media tutkimusaineistona. Teoksessa Jaana Vuori (toim.) *Laadullisen tutkimuksen verkkokäsikirja*. Tampere: Yhteiskuntatieteellinen tietoaristo (FSD). <https://www.fsd.tuni.fi/fi/palvelut/menetelmaopetus/kvali/laadullisen-tutkimusaineistot/sosiaalinen-media-tutkimusaineistona/> (Tarkistettu toukokuussa 2023)

Laaksanen, Salla-Maaria, Haapoja, Jesse, Kinnunen, Teemu, Nelimarkka, Matti & Pöyhkäri, Reeta (2020) The datafication of hate: Expectations and challenges in automated hate speech monitoring. *Frontiers in Big Data* 3/2020. <https://doi.org/10.3389/fdata.2020.00003>

Laaksanen, Salla-Maaria & Matikainen, Janne (2013) Tutkimuskohteena vuoro vaikutus ja keskustelu verkossa. Teoksessa Salla-Maaria Laaksanen, Janne Matikainen & Minttu Tikka (toim.) *Otteita verkosta. Verkon ja sosiaalisen median tutkimusmenetelmät*. Tampere: Vastapaino, 193–215.

Laaksanen, Salla-Maaria & Pääkkönen, Juhon (2020) Between automation and interpretation: Using data visualization in social media analytics companies. Teoksessa Helen Kennedy & Martin Engebretsen (toim.) *Data Visualization in Society*. Amsterdam: Amsterdam University Press, 95–110.

Lievrouw, Leah A. & Livingstone, Sonia (2006) Introduction to the updated student edition. Teoksessa Leah A. Lievrouw & Sonia Livingstone

(toim.) *Handbook of New Media*. London: Sage, 35–54.

MacKenzie, Donald; Muniesa, Fabian & Siu, Leung-Sea (2007) *Do Economists Make Markets? On the Performativity of Economics*. Princeton & Oxford: Princeton University Press.

Madsen, Anders Koed (2015) Tracing data – paying attention. Teoksessa Martin Kornberger, Lise Justesen, Anders Koed Madsen & Jan Mouritsen (toim.) *Making Things Valuable*. Oxford: Oxford University Press, 257–278. <https://doi.org/10.1093/acprof:oso/9780198712282.003.0013>

Marttila, Mari, Laaksonen, Salla-Maaria, Kekkonen, Arto, Tuokko, Mari & Nelimarkka, Matti (2016) Digitaalinen vaalitelta: Twitter politiikan areenana eduskuntavaaleissa 2015. Teoksessa Hanna Wass & Kimmo Grönlund (toim.) *Eduskuntavaalitutkimus 2015: Poliittisen osallistumisen erityyminen*. Oikeusministeriön Selvityksiä ja julkaisuja 28/2016.

Merton, Robert K. (1968) The Matthew Effect in science: The reward and communication systems of science are considered. *Science* 159(3810), 56–63.

Mayer-Schönberger, Viktor & Cukier, Kenneth (2013) *Big Data: A Revolution That Will Transform How We Live, Work and Think*. London: John Murray.

Nelimarkka, Matti (2022) *Computational Thinking and Social Science: Combining Programming, Methodologies and Fundamental Concepts*. London: Sage.

Ojala, Markus, Pantti, Mervi & Laaksonen, Salla-Maaria (2019) Networked publics as agents of accountability: Online interactions between citizens, the media, and immigration officials during the European refugee crisis. *New Media & Society* 21(2), 279–297

Pentland, Alex (2014) *Social Physics: How Good Ideas Spread – The Lessons From a New Science*. London: Penguin Press.

Pink, Sarah, Ruckenstein, Minna, Willim, Robert & Duque, Melisa (2018) Broken data: Conceptualising data in an emerging world. *Big Data & Society* 5(1). <https://doi.org/10.1177/2053951717753228>

Porter, Theodore M. (1995) *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton: Princeton University Press. <https://doi.org/10.1017/CBO9781107415324.004>

Rieder, Bernhard, Abdulla, Rasha, Poell, Thomas, Woltering, Robbert & Zack, Liesbeth (2015) Data critique and analytical opportuni-

ties for very large Facebook Pages: Lessons learned from exploring “We are all Khaled Said”. *Big Data & Society* 2(2). <https://doi.org/10.1177/2053951715614980>

Rogers, Richard (2017) Digital methods for cross-platform analysis. Teoksessa Jean Burgess, Alice Marwick & Thomas Poell (toim.) *The SAGE Handbook of Social Media*. London: Sage, 91–108. <https://doi.org/10.4135/9781473984066.n6>

Ruppert Evelyn & Savage Mike (2011) Trans-accional politics. *The Sociological Review* 59(2, supplement), 73–79. <https://doi.org/10.1111/j.1467-954X.2012.02057.x>

Sumner, Erin M., Ruge-Jones, Luisa & Alcorn, Davis (2018) A functional approach to the Facebook Like button: An exploration of meaning, interpersonal functionality, and potential alternative response buttons. *New Media and Society* 20(4), 1451–1469. <https://doi.org/10.1177/1461444817697917>

Vainikka, Eliisa & Huhtamäki, Jukka (2015) Tviittien politiikkaa – poliittisen viestinnän sisäpiirit Twitterissä. *Media & Viestintä* 38(3), 165–183.

Van Dijck, José (2014) Datafication, dataism and dataveillance: Big data between scientific paradigm and ideology. *Surveillance & Society* 12(2), 197–208.

Van Dijck, José & Poell, Thomas (2013) Understanding social media logic. *Media and Communication* 1(1), 2–14. <https://doi.org/10.17645/mac.v1i1.70>

Van Dijck, José, Poell, Thomas & De Waal, Martijn (2018) *The Platform Society: Public Values in a Connective World*. Oxford: Oxford University Press.

Dosentti, VIT **Salla-Maaria Laaksonen** työskentelee yliopistotutkijana Helsingin yliopiston Kuluttajatutkimuskeskuksessa. Laaksonen on viestinnän ja teknologian tutkija, joka on työskennellyt myös organisaatiotutkimuksen ja teknologian tutkimuksen parissa. Hän tutkii teknologisen julkisuuden valta-asetelmia sekä datan, algoritmien ja automaation käyttöä organisaatioissa. Laaksonen hyödyntää työssään paljon sosiaalisen median aineistoja sekä digitaalisia ja laskennallisia menetelmiä.