

Johdanto: Yhteiskuntatieteiden ja informaatioteknologian rajapinnoilla

Salla-Maaria Laaksonen

Helsingin yliopisto

salla.laaksonen@helsinki.fi

<https://orcid.org/0000-0003-3532-2387>

Jesse Haapoja

Helsingin yliopisto, Aalto-yliopisto

jesse.haapoja@helsinki.fi

<https://orcid.org/0000-0001-6877-7957>

Thomas Olsson

Tampereen yliopisto

thomas.olsson@tuni.fi

<https://orcid.org/0000-0002-1106-2544>

Asiasanat: digitalisaatio, digitaaliset ihmistieteet, informaatioteknologia, tietotekniikka, tutkimusmenetelmät, yhteiskunta

Informaatioteknologian kehitys ja yhteiskunnan digitalisoituminen ovat kannustaneet tutkijoita eri tieteenaloilta tarkastelemaan erilaisten sosio-tekniisten järjestelmien ja digitalisoituvien käytäntöjen merkitystä yhteiskunnassa. Teknologiavälitteisyys muun muassa muovaa sosiaalista, poliittista ja kaupallista kanssakäymistä sekä muokkaa yhteiskunnan instituutioita kouluista mediaan. Samalla digitalisoituva yhteiskunta tuottaa uusia ihmisiä ja yhteiskuntatieteiden tutkimuskohteita, aineistoja ja menetelmiä. Tämän monitieteisen teemanumeron artikkelit kartoittavat ja pohtivat digitaalisen ihmis- ja yhteiskuntatieteen nykytilaa ja tulevaisuutta erityisesti muuttuvien aineistojen luomien haasteiden ja mahdollisuuksien näkökulmasta. Teemanumero on toimitettu yhteistyössä digitaalisen yhteiskuntatieteen tutkimuksen yhdistyksen Rajapinta ry:n kanssa.

Informaatioteknologian ja digitalisoitumisen yhteiskunnallisia vaikutuksia, mahdollisuuksia ja reunaehtoja onkin tutkittu lukuisilla eri tieteenaloilla esimerkiksi sosiaalitieteiden, viestintätieteiden, taloustieteiden ja informaatiotutkimuksen perinteistä käsin. Digitalisoitumisen vaikutukset ihmistieteelliseen tutkimukseen ovat monitasoisia. Ensinnäkin informaatioteknologian laajeneva käyttö kannustaa tutkimaan teknologian roolia erilaisissa yhteiskunnallisissa prosesseissa. Toisekseen informaatioteknologiset järjestelmät tuottavat tutkijoille yhteiskunnallisen toiminnan aidossa kontekstissa syntyneitä, laajoja aineistoja ihmisten, yhteisöjen ja yhteiskunnan toiminnasta. Sosiaalisen toiminnan jättämät digitaaliset jäljet mahdollistavat uudenlaisten kysymysten esittämisen. Samalla ne kuitenkin aiheuttavat uusia haasteita tutkimukselle: miten menetelmällisesti lähestyä aineistoja niin, että yhteiskuntatieteellinen ote säilyy? Miten tuottaa ymmärrystä teknologiavälitteisistä, digitaalisista yhteiskunnan ilmiöistä ilman, että päätyy tutkimaan pelkästään teknologiaa tai pelkästään yhteiskuntaa?

Vastataksaan näihin haasteisiin digitalisoituvan yhteiskunnan rajapinnoilla uraauurtavaa työtä tehneet tutkijat ovat kehittäneet uudenlaisia, digitaalisen yhteiskunnan tutkimukseen sovellettuja käsitteitä ja menetelmiä. Erillisinä, varsin metodologisesti orientoituneina aloina erottuvat jo digitaaliset ihmistieteet (esim. Terras ym. 2013), digitaalinen sosiologia (Marres 2017, Lupton 2014), digitaalinen tai laskennallinen yhteiskuntatiede (Marres 2012; Lazer ym. 2009; Nelimarkka & Laaksonen 2018) ja diginatiivit menetelmät (Rogers 2013). Näissä lähestymistavoissa on kehitetty erilaisia tapoja paitsi kerätä ja analysoida digitaalisia aineistoja, myös valjastaa digitaaliset järjestelmät tutkimuksen avuksi tukemaan perinteisempiä tutkimusotteita. Tarve tällaiselle tutkimusotteelle näkyy myös rahoitusinstrumenteissa: esimerkiksi vuosina 2016–2020 Suomen Akatemia pyöritti Digitaaliset ihmistieteet -akatemiaohjelmaa, jonka kautta rahoitettiin kaikkiaan 15 eri hanketta.

Ihmistieteen menetelmien ja aineistojen digitalisoituminen on herättänyt myös kriittistä keskustelua. Tutkijat ovat pyrkineet arvioimaan isojen digitaalisten aineistojen aikaansaamia muutoksia yhteiskuntatieteelliseen tutkimukseen (Kitchin 2014; boyd & Crawford 2012; Elish & boyd 2018; Frade 2016; Marres & Weltevrede 2013; Ruppert ym. 2013). Yksi merkittävä keskustelunaihe on ollut digitaalisten aineistojen edustavuus ja suhde tutkittavaan ilmiöön: Millä tavoin teknologiset alustat muokkaavat dataa, jota ne ihmistoiminnasta keräävät ja tuottavat? Mitä datasta jää puuttumaan (Hargittai 2020; Tromble 2019; Halford ym. 2018; Marres 2017)? Myös digitaalisten aineistojen käytön etiikasta käydään vilkasta keskustelua (Kosonen ym. 2018, Lazer ym. 2020; Zimmer & Kinder-Kurlanda 2017). Toiset ovat kritisoineet digitaalisen tutkimuksen keskittyvän liiallisesti tiettyihin alustoihin ja ympäristöihin, joista aineistoa on helposti saatavilla (Hargittai 2020; Marres 2017; Burgess & Bruns 2015).

Kehittyvä tutkimuskenttä kaipaakin paitsi menetelmäkehitystä, myös menetelmien reflektointia sekä pyrkimyksiä soveltaa perinteisiä menetelmiä uusissa konteksteissa – esimerkiksi vuorovaikutuksen tutkimista digitaalisilla alustoilla (esim. Herring 1996; Laaksonen & Matikainen 2013; Laitinen ym. 2021), verkkoympäristöön sovellettua etnografiaa (esim. Hine 2000; Laaksonen ym. 2017) tai digitaalisten palveluiden käyttöä tallentavia ja tarkastelevia menetelmiä (esim. McMillan ym. 2015; Haapoja ym. 2021). Lisäksi tarvitaan lisää tutkittuun tietoon perustuvaa ymmärrystä teknologisista konteksteista ja niiden kulttuureista. Tällaiset lähtökohdat vaativat tyypillisesti sekä oman tieteenalan vankkaa osaamista että teknologisen maailman tuntemusta. Tämän teemanumeron ydinsisältöä ovat monitieteiseen, teknologian ja yhteiskunnan rajapinnalla tehtävään tutkimukseen liittyvät mahdollisuudet ja haasteet sekä tutkimuksen uudenlaiset käytännönläheiset ja episteemiset kysymykset. Teemanumeron kahdeksan artikkelia ja kaksi katsausta tarjoavat monipuolisen ja monitieteisen kattauksen teknologian ja yhteiskunnan risteämäkohdan jännitteisiin.

Jäsenämme teemanumeron kokonaisuutta kolmen pääteeman kautta. Ensimmäinen teema “Muuttuva yhteiskunta, muuttuva teknologia” piirtää kuvastoa nopeasti muuttuvista digitaalisista yhteiskunnista, yhtäältä suomalaisesta ja toisaalta laajemmasta länsimaisesta näkökulmasta. Toinen teema “Uudet aineistot ja näkökulmat analyysiin” tarkastelee aineistoja sekä perspektiivejä niiden analyysiin tutkimuksen epistemologisina perustekijöinä, avaten mahdollisuuksia uudensuuntaamiseen ja ymmärtämisen tapoihin. Lopuksi teema “Yhteistyön infrastruktuurit” tarjoaa käytännönläheisen katsauksen uusiin tutkimus- ja tietoresursseihin infrastruktuurina, joka mahdollistaa monitieteistä yhteiskunnan ja teknologian törmäyskohtien tutkimista.

Muuttuva yhteiskunta, muuttuva teknologia

Kenties keskeisin digitaalisen yhteiskunnan tutkimusta kuvaava sana on muutos. Digitaaliset ympäristöt ja järjestelmät kehittyvät jatkuvasti suunnittelun ja alustojen keskinäisen kilpailun kautta (Savor ym. 2016; Bossetta, 2020). Jatkuva kehitys tekee digitaalisen yhteiskunnan ilmiöistä vaikeasti rajattavia ja hyvin kontekstuaalisia; verkostomaisia ilmiöitä, jotka saattavat näyttäytyä erilaisina eri käyttäjille (Laaksonen ym. 2013). Tutkijan tulisikin kyetä ottamaan huomioon paitsi tutkittavan aineiston sisältö, myös sen synty- ja käyttökonteksti (esim. Sumiala & Tikka 2011, Vaahensalo, tämä numero), mikä voi olla erityisen haasteellista isoilla digitaalisilla aineistoilla tehtävien tutkimusten kanssa. Keskeistä on ymmärtää sosiaalisen toiminnan ja teknologisten järjestelmien kytkeytyneisyys: teknologiat syntyvät sosiaalisen toiminnan ja ihmistyön tuloksena, ja synnyttyään ne vaikuttavat niiden avulla tapahtuvaan sosiaalisen toimintaan. Toisaalta sosiaalinen toiminta vastavuoroisesti muokkaa teknologioiden käyttöä ja niihin liitettyjä merkityksiä. Tätä vuorovaikutusta on tutkimuksessa kutsuttu teknologian sosiaalisesti muovautumiseksi tai teknologian ja sosiaalisen toiminnan vastavuoroisen rakentumisen periaatteeksi (esim. Lievrouw & Livingstone 2006, MacKenzie & Wajcman 1999). Teemanumeron ensimmäisen osion neljä artikkelia heijastelevat kaikki juuri tätä yhteispeliä.

Elisa Kannasto ja Ari Haasio tutkivat artikkelissaan “Tuntuu, että vietän liikaakin aikaa Internetissä.” – Verkon pieni maailma virtuaalisena kotiseutuna” empiirisesti opiskelijoiden virtuaalista kotiseutua verkkopäiväkirjaa hyödyntäen. Virtuaalinen kotiseutu on käsite, joka kattaa verkon ulkopuolisen maailman virtuaaliset representaatiot ja digitaalisen maailman kotoiset ympäristöt, joilla ei ole vastinetta verkon ulkopuolella (Riukulehto & Haasio, 2020). Kirjoittajat hyödynsivät verkkopäiväkirjoja kerätäkseen aineistoa osallistujien verkkopalveluiden käytöstä ja roolista heidän jokapäiväisessä elämässään, ja hahmottelivat näin edellä mainittua virtuaalista kotiseutua ja eri käyttäjätyyppejä. Aineistonsa pohjalta kirjoittavat esittelevät neljä eri virtuaalisen kotiseudun arkkityyppiä, joissa korostuvat eri tavoin esimerkiksi painotukset viihteelliseen ja tarvekäyttöön. Artikkelin tulokset muistuttavat digitaalisten ympäristöjen tutkijoita siitä, että erilaisten palveluiden ja teknologioiden käyttötavat ja merkitykset käyttäjien arjessa voivat olla hyvinkin moninaisia.

Liisa Kääntä esittelee artikkelissaan “Kannanottaminen Twitterissä sosiaalisen ja digitaalisen toiminnan rajapintana” keskusteluanalyysin soveltamista kannanottamisen tutkimuksessa tapausesimerkinään korona-aikana ajankohtainen etätyötä käsittelevä Twitter-keskustelu. Kääntä

osoittaa kannanottamisen Twitterissä olevan sekä sosiaalista että digitaalista toimintaa, jossa etätyön arvottaminen ja näkemykset siitä sulautuvat yhteen digitaalisen maailman vuorovaikutuksellisten ominaisuuksien, kuten emojiiden ja aihetunnisteiden kanssa. Ne erottavat vuorovaikutusta digitaalisessa ympäristössä kasvokkain tapahtuvasta. Kääntä korostaa kuinka vuorovaikutus verkossa on kytköksissä paitsi sosiaalisuuteen myös vuorovaikutusta välittävän alustan teknologiaan. Teknologian pienilläkin vuorovaikutusta tukevilla muodoilla kuten emojiilla on tärkeä rooli siinä, millaiseksi vuorovaikutus rakentuu verkossa.

Osion kolmas artikkeli liikkuu myös työn ja teknologian konteksteissa. Paul Jonker-Hoffrénin englanninkielinen artikkeli “Bridging the gap between critical algorithm studies and the social sciences: the case of platform work” käsittelee modernia työtä alustataloudessa. Hän soveltaa sosiologi Maurizio Lazzaraton sosiaalisen subjektiviteetin käsitettä tarkastellakseen kuinka alustojen teknologiset valinnat ja algoritmiset logiikat määrittelevät alustatyön ehtoja. Artikkelin problematisoi työntekijän toimijuuden ja itsensä määräämisoikeuden vakavaa heikentymistä alustatyössä. Metodologisena johtopäätöksenä Jonker-Hoffrén alleviivaa inhimillisten ja teknologisten näkökulmien kytkettyneisyyttä erityisesti alustatyön kaltaisten sosioteknisten ilmiöiden kontekstissa. Teemanumeron tematiikan kannalta artikkeli tarjoaa syväluotaavan analyysin yhdestä digitaalisen talouden merkittävimmistä, mutta yllättävän vähän tutkituista ilmiöistä.

Ensimmäisen osion viimeinen artikkeli tarjoaa näkökulman teknologian suunnitteluun. Veikko Isotalon englanninkielinen artikkeli “Improving candidate-based voting advice application design: The case of Finland” tarkastelee vaalikoneita suunnittelututkimuksen näkökulmasta. Isotalo analysoi vaalikoneiden käyttöliittymäsuunnittelun lähtökohtia ja suunnitteluavaruutta erityisesti Suomen vaalijärjestelmän ja poliittisen vaikuttamisen kontekstissa. Analyysi kartoittaa vaihtoehtoisia valintoja liittyen mm. kysymyksen asetteluun, vastausasteikkoihin, informaation visualisointiin ja ehdokkaita priorisoivan algoritmin logiikkaan. Teemanumeron metodologista kirjoa ajatellen Isotalo avaa erinomaisesti sitä kompleksisuutta, jota informaatiojärjestelmien suunnittelussa usein kohdataan sekä sen analysoinnin tärkeyttä, miten suunnittelussa – tiedostaen tai tiedostamatta – tehdyt arvovallinnat vaikuttavat toteutetun järjestelmän tarkoituksenmukaisuuteen ja poliittiseen luonteeseen.

Uudet aineistot ja näkökulmat analyysiin

Digitaalisen yhteiskunnan tutkimuksessa hyödynnetään paljon perinteisiä ihmistieteiden menetelmiä haastatteluista kyselytutkimukseen ja etnografiaan. Monet niistä ovat kuitenkin sopeutettuja soveltumaan juuri digitaalisten ympäristöjen ja aineistojen tutkimukseen (Laaksonen ym. 2013, Marres 2017). Viime vuosina myös ihmistieteilijät ovat ryhtyneet kehittämään erilaisia digitaaliseen maailmaan erikoistuneita metodologioita ja menetelmiä, usein monitieteisen yhteistyön kautta. Ajankohtaisia, suosittuja lähestymistapoja ovat esimerkiksi digitaaliset tai diginatiivit menetelmät (Rogers 2013; Marres & Gerlitz 2016; Gerlitz & Rieder 2018), laskennallisen yhteiskuntatieteen yleistyminen (Lazer ym. 2009; Bail 2014; Nelimarkka & Laaksonen 2018), tai verkkoetnografian erilaiset muodot (esim. Hine 2000; Isomäki ym. 2013; Knox & Nafus 2018). Tämän teemanumeron toinen osio keskittyy digitaalisen yhteiskunnan tutkimuksen menetelmällisiin kysymyksiin.

Osion ensimmäinen artikkeli jatkaa osaltaan muuttuvan kontekstin hahmottelua, mutta vahvasti tutkimusmenetelmien ja -prosessin näkökulmista. Artikkelissaan “Affordansien ja kulttuuristen kontekstien rajapinnoilla – Kontekstualisointimalli sosiaalisen median lähdekritiikin avaimena” Elina Vaahensalo tarjoaa kontekstualisointimallin sosiaalisen median lähdekritiikin avuksi sisältöjä tulkitsevalle tutkijalle. Hän erittelee mallissaan erilaisia konteksteja ja affordansseja, joita tutkijoiden olisi hyvä ottaa huomioon sosiaalisen median aineistoja tutkimusmateriaalina käyttäessään. Vaahensalo huomauttaa esimerkiksi, että sosiaalisen median aineistoja tulisi tulkita osana omaa aikaansa sekä sen hetkistä kulttuurista ja yhteiskunnallista taustaa. Hän myös muistuttaa osuvasti, että sosiaalisen median aineistoja analysoiva tutkija ei ole aineistostaan irrallinen tulkitsija, vaan aineistoon vaikuttavat kontekstit muovaavat myös tutkijan tekemiä tulkintoja.

Osion toinen artikkeli sukeltaa laskennallisen yhteiskuntatieteen maailmaan. Eräs varsin suosittu automaattisen tekstianalyysin menetelmä ihmistieteissä on ollut aihehallinnus, joka ohjaamattomasti etsii sanaryhmien avulla tunnistettavia aiheita tekstimassasta (Blei ym. 2012; Nelimarkka 2019). Artikkelissaan “Keskusteluaineistojen analyysi aihehallinnuksella ja aineiston esikäsittelyn vaikutukset sosiaalitieteelliselle tulkinnalle” Arho Toikka esittelee tutkimusprosessin, jossa suoritetaan toistettuja aihehallinnuksia eri parametrein, ja klusterianalyysin avulla ryhmitellään samankaltaisia aiheita eri mallinnuksista. Artikkelisi esittelee yhden keinon, jolla aihehallien parametrisonnaisuutta voi lähestyä ja hallita, ja samalla parantaa analyysin reliabiliteettia. Samalla Toikan artikkeli keskustelee kriittisesti yleisemmän

kysymyksen kanssa: miten eri menetelmiin erottamattomasti kuuluvat valinnat vaikuttavat analyysin tuloksiin.

Oman haasteensa digitaalisten aineistojen analyysiin luovat niiden muuttuvat esitystavat (esim. Paasonen 2013). Margareta Salosen, Elisa Kannaston ja Laura Paatelaisen katsaus “Sosiaalisen median kommenttien analyysi multimodaalisesta näkökulmasta digitaalisen journalismin tutkimuksessa” esittelee kiehtovan kirjallisuuskatsauksen, jossa multimodaalisuus on analyttisen viitekehyksen keskiössä. Katsauksessa analysoidaan 66 digitaalisen journalismin ja poliittisen viestinnän aloilla julkaistua artikkelia, joissa analysoidaan sosiaalisen median kommentteja. Kattavaa kirjallisuusaineistoa tarkastellaan kysymällä miten multimodaalisuus eli erilaisten kielellisten muotojen yhdistelmät (esim. tekstisisältö, kuva, linkki, emoji, aihetunniste) on otettu huomioon sosiaalisen median kommentointia tutkittaessa. Katsauksen keskeinen johtopäätös on, että eri moodeja on hyödynnetty verrattain vähän ja niiden valintaa ei useinkaan ole perusteltu, mikä yhtäältä kertoo laajasta hyödyntämättömästä potentiaalista sosiaalisen median kommenttien analysoinnissa. Artikkeliaineiston perusteella on vaikea arvioida, onko kyse tekstisisällön analyysin helpoudesta vai tietoisista valinnoista, mutta katsauksen perusteella tutkimuskentällä on selkeä tarve huomioida sosiaalisen median aineistojen eri muotoja kattavammin.

Yhteistyön infrastruktuurit

Teemanumeron viimeinen osio ottaa tarkastelun kohteeksi tutkijoiden näkökulmat ja käytänteet digitaalisten aineistojen ja niiden infrastruktuurien äärellä. Digitaalisen ihmistieteen kehitys on alkuvaiheissaan ollut vahvasti yksittäisten tutkijoiden ja tutkijaryhmien varassa, ja projekteissa on kehitetty työkaluja kulloisenkin tarpeen mukaan. Tutkijoita on kannustettu yhdistämään rohkeasti erilaisia menetelmiä (esim. Laaksonen ym. 2017; Geiger & Ribes 2011), hyödyntämään erilaisia sosiaalisen massadatan muotoja (Kitchin 2014; Olshannikova ym. 2017) ja kehittämään monitieteistä vuoropuhelua erilaisten tieteenalataustojen ja osaamisprofiilien välillä (Halford & Savage 2017; Moats & Borra 2018). Tällainen toimintatapa tuottaa kumuloituvaa tietoa ja osaamista, joka voi kuitenkin olla pirstaloitunutta ja vaikeasti löydettävää; onkin epäselvää, miten hyvin näin hankittu osaaminen siirtyy tutkimusyhteisössä eteenpäin.

Tämän huolen johdattelemana viime aikoina on keskustelu yhä enemmän digitaalisen tutkimuksen infrastruktuureista ja erilaisista keskitetyistä työvälineistä ja ratkaisuista, joita voitaisiin tarjota tutkimuksen tueksi.

Työkalut voivat olla esimerkiksi koodikirjastoja tai jopa valmiita ohjelmistoja, jolloin myös vähemmän teknisesti orientoituneet tutkijat voivat käyttää niitä itsenäisesti. Tämän teemanumeron päättävät kaksi artikkelia ja yksi katsaus pureutuvat digitaalisen ihmistieteen infrastruktuureihin ja niiden äärellä tehtävään tutkimustyöhön.

Erno Liukkonen, Liisa Näpärä ja Tuula Pääkkönen tarkastelevat Kansalliskirjaston datapakettien ja rajapintojen tutkimuskäytön mahdollisuuksia ja rajoitteita artikkelissaan “Digitaalisten aineistojen tekninen kehitys: Rajapinnat tutkijoiden työkaluna Kansalliskirjastossa nyt ja tulevaisuudessa”. Artikkeliki esittelee tarjolla olevia valmiita datapaketteja ja rajapintoja sekä niiden käyttöä tutkijoiden kokemuksiä kartoittaneiden kyselyiden ja haastattelujen avulla, sekä vertailemalla niitä muihin kirjastojen tekemiin ratkaisuihin. Empiirisen analyysin mukaan sekä datapaketit että rajapinnat ovat alikäytettyjä ja näyttäytyvät vaikeasti lähestyttävänä perinteiselle ihmistieteilijälle. Rajoittavia tekijöitä ovat niin taidot, ymmärrys kuin asenteetkin. Ratkaisuna näihin ongelmiin kirjoittajat esittävät parempaa dokumentaatiota sekä erilaisten valmiiden työkalujen tarjoamista erityisesti niille tutkijoille, joilla ei ole ohjelmointitaitoja.

Ikään kuin yhtenä vastauksena Liukkonen ja kollegoiden esittämään haasteeseen Eero Hyvönen kollegoineen esittelevät katsauksessaan työkalun eduskunnan aineistojen linkitetyn avoimen datan tutkimukseen. Työkalu tarjoaa uudenlaisen tavan päästä käsiksi rikkaaseen eduskunta-aineistoon. Hankkeessa kehitettävän semanttisen webin teknologioille perustuvan *Parlamenttisampo – eduskunta semanttisessa webissä* -nimisen työkalun tarkoituksena on tarjota avoimen datan palvelu tutkijoille, kansalaisille, medialle ja valtionhallinnolle, ja palvelussa eduskunta-dataa rikastetaan muulla avoimella datalla, kuten esimerkiksi lainsäädäntöä koskevalla tiedolla. Kattavuudessaan palvelu luo varsin ainutlaatuisen näkymän yhteen digitaalisen yhteiskunnan tutkimuksen teema-alueeseen ja mahdollistaa poliittisen päätöksenteon ja vuorovaikutuksen tutkimisen erilaisilla aineistoilla.

Heikki Keskustalo, Elina Late, Laura Korkeamäki, Sanna Kumpulainen ja Kimmo Kettunen tarkastelevat artikkelissaan “Informaatiovuorovaikutus historian tutkimuksessa” digitaalisten aineistojen ja työvälineiden käyttöä osana tutkimusprosessia yhdistämällä käyttäjakeskeistä ja järjestelmakeskeistä tiedonhaun tutkimusta tarkastelemalla niitä informaatiovuorovaikutuksena: miten tutkijat vuorovaikuttavat aineistojen kanssa tutkimusprosessin aikana? Näin ollen artikkeli tarkastelee sekä ihmisten tiedontarpeita että dokumenttien muodostamaa avaruutta niihin vastaajana; kirjoittajien termein sekä kognitiivista avaruutta että dokumenttiavaruutta. Artikkelin empiirinen aineisto koostuu haastatteluista ja havainnoista, joita on kerätty erilaisten

historiallisten aineistojen kanssa työskentelevien tutkijoiden parissa. Artikkelellä piirtää monipuolisen kuvan niistä vuorovaikutuksen tavoista, joilla tutkijat aineistonsa parissa toimivat. Kirjoittajat korostavat, että sillan rakentaminen kognitiivisen avaruuden ja dokumenttiavaruuden välille on haastavaa. Esteiksi muodostuvat sekä tekniset että sosiaaliset haasteet: aineisto ei suoraan vastaa tutkimuskysymykseen tai edes puhu vaadittavaa kieltä, ja oma haasteensa on etsiä yhteinen tiedonintressi usein monitieteisessä tutkimusryhmässä.

Kaikissa tämän osion artikkeleissa korostuu digitaalisen ihmistieteen tutkijoiden ja tutkimusprosessien moninaisuus. Liukkonen ja kumppanit kategorisoivat artikkelissaan tutkijat karkeasti kolmeen eri luokkaan: digitaalisia aineistoja perinteisin menetelmin analysoivat tutkijat; tutkijat, joille digitaaliset menetelmät tukevat muita analyysitapoja; ja tutkijat, joille digitaaliset menetelmät ovat pääosassa. Keskustalo ja kollegat nostavat esille, että tutkijat kaipaavat tutkimusongelmasta riippuen erilaisia kognitiivisia pääsykohtia aineistoihin. Näitä voivat olla esimerkiksi aineistoissa esiintyvät henkilöt ja paikat tai niiden ominaisuudet. Näiltä osin Hyvösen ja kollegoiden esittelemä Parlamenttisampo näyttäytyykin palveluna, joka tarjoaa varsin oivallisesti erilaisia pääsykohtia ja siten fasiltoi ja helpottaa tehtävää tutkimusta riippumatta tutkijoiden teknisestä taitotasosta. Toimivien infrastruktuurien vaatimukset eri taustaisille tutkijoille ja erilaisille tutkimustarpeille voivat kuitenkin olla varsin erilaisia.

Lopuksi

Kuten tämän teemanumeron artikkeleiden ja kirjoittajien taustojen kirjo osoittaa, digitaalisen yhteiskunnan tutkimus elää tieteidien ja konventioidien rajapinnoilla: se on usein monitieteistä työskentelyä yhteiskuntatieteellisten kysymysten äärellä. Digitaalisten ympäristöjen ja digitalisoituvan yhteiskunnan tutkiminen vaatii monikäisyyttä: aineiston käsittelyssä ja analysoinnissa tarvittavia teknisiä taitoja, teoreettista tietämystä yhteiskuntatieteellisistä keskusteluista sekä menetelmällistä osaamista uutta isoista aineistoista mielenkiintoisia havaintoja. Toisinaan nämä taidot yhdistyvät yhdessä tutkijassa, mutta kenties useammin tutkimusryhmän sisällä. Artikkelissaan Susan Halford ja Mike Savage (2017) käyttävät sinfonian vertauskuvaa kuvatakseen sitä tapaa, jolla isoja data-aineistoja voidaan käyttää yhteiskuntatieteellisesti kiinnostavan tutkimuksen tekemiseen yhteiskuntatieteen lähtökohdista. Sinfoniassa yhdistyvät erilaiset vaihtelevat teemat, jotka yhdessä muodostavat musiikkiteoksen. Sinfonisessa sosiaalitetieteessä yhdistyvät rikas teoreettinen

ymmärrys sekä huolellisesti valittu aineisto, joka piirtää kuvaa laajoista sosiaalisista kysymyksistä. Tällaiset kysymykset ovat yhä ihmistieteiden ytimessä myös digitaalisella aikakaudella.

Teemanumeron artikkelit osoittavat, että digitalisoituvan yhteiskunnan jäljet näkyvät ihmistieteen tutkimuksessa monella eri tasolla. Digitaalisuus muuttaa tutkimuskohteita, tutkittavia aineistoja ja niiden analyysimenetelmiä, sekä lopulta kutsuu ja kannustaa tiedeyhteisöä etsimään niin ihmis- kuin teknologiainfrastruktuureihin liittyviä pysyvämpiä ratkaisuja digitaalisten ihmistieteiden tutkimusprosesseihin. Tämä teemanumero valottaa näitä kysymyksiä, mutta myös avaa uusia. Reflektiivinen tarkastelu on yhä tarpeen esimerkiksi suhteessa digitaalisen ihmistieteen epistemologisiin, metodologisiin ja praktisiin reunaehtoihin: miten käytetyt työkalut ja aineistot vaikuttavat tutkimuksen suunnitteluun, rahoituksen hakuun ja käytännön toteutukseen, tutkimuksen tuloksiin ja päätelmiin sekä eri tieteenalojen kehityskulkuihin pidemmällä aikavälillä?

Teemanumeron toimittajat kiittävät lämpimästi kaikkia numeron kirjoittajia sekä artikkeleiden vertaisarvioija. Teemanumeron kokonaisuutta ja käsikirjoitusten ensimmäisiä versioita hiottiin vertaiskommentoinnin keinoin myös huhtikuussa 2021 järjestetyssä virtuaalisessa Skrivarstuga-työpajassa.

Viitteet

- Bail, C. A. (2014). The cultural environment: Measuring culture with big data. *Theory and Society*, 43(3), 465–524. <https://doi.org/10.1007/s11186-014-9216-5>
- Blei, D. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77–84. <https://doi.org/10.1109/MSP.2010.938079>
- Bossetta, M. (2020). Scandalous Design: How Social Media Platforms' Responses to Scandal Impacts Campaigns and Elections. *Social Media + Society*, 6(2), 205630512092477. <https://doi.org/10.1177/2056305120924777>
- Burgess, J., & Bruns, A. (2015). Easy data, hard data: the politics and pragmatics of Twitter research after the computational turn. Teoksessa G. Langlois, J. Redden, & G. Elmer (toim.), *Compromised Data: From Social Media to Big Data* (s. 93–111). Bloomsbury Publishing.
- boyd, d., & Crawford, K. (2012). Critical Questions for Big Data. *Information, Communication & Society*, 15(5), 662–679.
- Elish, M. C., & boyd, d. (2018). Situating methods in the magic of Big Data and AI. *Communication Monographs*, 85(1), 57–80. <https://doi.org/10.1080/03637751.2017.1375130>

- Frade, C. (2016). Social Theory and the Politics of Big Data and Method. *Sociology*, 50(5), 863–877. <https://doi.org/10.1177/0038038515614186>
- Geiger, R. S., & Ribes, D. (2011). Trace ethnography: Following coordination through documentary practices. *Proceedings of the 44th Hawaii International Conference on System Sciences 2011*. <https://doi.org/10.1109/HICSS.2011.455>
- Gerlitz, C., & Rieder, B. (2018). Tweets Are Not Created Equal: Investigating Twitter's Client Ecosystem. *International Journal of Communication*, 12, 528–547. Retrieved from <http://ijoc.org/index.php/ijoc/article/viewFile/5974/2252>
- Haapoja, J., Lampinen, A., & Vesala, K. M. (2021). Personalised Services in Social Situations: Principal-Agent Relationships in Account Sharing. *Proceedings of the ACM on human-computer interaction* 4(CSCW3), no. 219. <https://doi.org/10.1145/3432918>
- Halford, S., & Savage, M. (2017). Speaking Sociologically with Big Data: Symphonic Social Science and the Future for Big Data Research. *Sociology*, 51(6), 1132–1148. <https://doi.org/10.1177/0038038517698639>
- Halford, S., Weal, M., Tinati, R., Carr, L., & Pope, C. (2018). Understanding the production and circulation of social media data: Towards methodological principles and praxis. *New Media and Society*, 20(9), 3341–3358. <https://doi.org/10.1177/1461444817748953>
- Hargittai, E. (2020). Potential Biases in Big Data: Omitted Voices on Social Media. *Social Science Computer Review*, 38(1), 10–24. <https://doi.org/10.1177/0894439318788322>
- Herring, S. C. (toim.). (1996). *Computer-mediated communication: Linguistic, social, and cross-cultural perspectives*. John Benjamins Publishing.
- Hine, C. (2000). *Virtual ethnography*. Sage. <https://doi.org/10.4135/9780857020277>
- Isomäki, H., Lappi, T.-R., & Silvennoinen, J. (2013). Verkon etnografinen tutkimus. Teoksessa S.-M. Laaksonen, J. Matikainen, & M. Tikka (toim.), *Otteita verkosta: Verkon ja sosiaalisen median tutkimusmenetelmät* (s. 150–169). Vastapaino.
- Kitchin, R. (2014). Big Data, new epistemologies and paradigm shifts. *Big Data & Society*, 1(1), 2053951714528481. <https://doi.org/10.1177/2053951714528481>
- Knox, H., & Nafus, D. (2018). *Ethnography for a data-saturated world*. Manchester University Press. <https://doi.org/10.7765/9781526127600>
- Kosonen, M., Rydenfelt, H., Laaksonen, S.-M., & Terkamo-Moisio, A. (2018). Sosiaalinen media ja tutkijan etiikka. *Media & Viestintä*, 41(1). <https://doi.org/10.23983/mv.69924>
- Laaksonen, S.-M., & Matikainen, J. (2013). Tutkimuskohteena vuorovaikutus ja keskustelu verkossa. Teoksessa S.-M. Laaksonen, J. Matikainen, & M. Tikka (toim.), *Otteita verkosta: Verkon ja sosiaalisen median tutkimusmenetelmät* (s. 193–215). Vastapaino.
- Laaksonen, S.-M., Matikainen, J., & Tikka, M. (2013). Tutkimusotteita verkosta. Teoksessa S.-M. Laaksonen, J. Matikainen, & M. Tikka (toim.), *Otteita verkosta: Verkon ja sosiaalisen median tutkimusmenetelmät* (s. 9–33). Vastapaino.
- Laaksonen, S.-M., Nelimarkka, M., Tuokko, M., Marttila, M., Kekkonen, A., & Villi, M. (2017). Working the fields of big data: Using big-data-augmented online ethnography to study candidate–candidate interaction at election time. *Journal of Information Technology and Politics*, 14(2), 110–131. <https://doi.org/10.1080/19331681.2016.1266981>

- Laitinen, K., Laaksonen, S.-M., & Koivula, M. (2021). Slacking with the Bot: Programmable Social Bot in Virtual Team Interaction. *Journal of Computer-Mediated Communication*. <https://doi.org/10.1093/jcmc/zmab012>
- Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A. L., Brewer, D., . . . Van Alstyne, M. (2009). Social science: Computational social science. *Science*, *323*(5915), 721–723 <https://doi.org/10.1126/science.1167742>
- Lazer, D. M. J., Pentland, A., Watts, D. J., Aral, S., Athey, S., Contractor, N., . . . Wagner, C. (2020). Computational social science: Obstacles and opportunities. *Science*, *369*(6507), 1060–1062. <https://doi.org/10.1126/science.aaz8170>
- Lievrouw, L. A., & Livingstone, S. (2006). Introduction to the Updated Student Edition. Teoksessa L. A. Lievrouw & S. Livingstone (toim.), *Handbook of new media* (s. 35–54). Sage.
- Lupton, D. (2014). *Digital sociology*. Routledge.
- MacKenzie, D., & Wajcman, J. (1999). *The social shaping of technology*. Toinen painos. Open University Press.
- Marres, N. (2012). The redistribution of methods: On intervention in digital social research, broadly conceived. *Sociological Review*, *60*(SUPPL. 1), 139–165. <https://doi.org/10.1111/j.1467-954X.2012.02121.x>
- Marres, N. (2017). *Digital Sociology: The Reinvention of Social Research*. Polity Press.
- Marres, N., & Gerlitz, C. (2016). Interface methods: Renegotiating relations between digital social research, STS and sociology. *Sociological Review*, *64*(1), 21–46. <https://doi.org/10.1111/1467-954X.12314>
- Marres, N., & Weltevrede, E. (2013). SCRAPING THE SOCIAL? Issues in live social research. *Journal of Cultural Economy*, *6*(3), 313–335. <https://doi.org/10.1080/17530350.2013.772070>
- McMillan, D., McGregor, M., & Brown, B. (2015). From in the wild to in vivo: Video Analysis of Mobile Device Use. *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services* (s. 494–503).
- Moats, D., & Borra, E. (2018). Quali-quantitative methods beyond networks: Studying information diffusion on Twitter with the Modulation Sequencer. *Big Data & Society*, *5*(1), 205395171877213. <https://doi.org/10.1177/2053951718772137>
- Nelimarkka, M. (2019). Aihemallinnus sekä muut ohjaamattomat koneoppimismenetelmät yhteiskuntatieteellisessä tutkimuksessa: kriittisiä havaintoja. *Politiikka*, *61*(1), 6–33. Noudettu osoitteesta <https://journal.fi/politiikka/article/view/79629>
- Nelimarkka, M., & Laaksonen, S.-M. (2018). Bitit ja politiikka: tervetuloa laskennallinen yhteiskuntatieteen tutkimus. *Politiikka*, *60*(2), 130–131.
- Olshannikova, E., Olsson, T., Huhtamäki, J., & Kärkkäinen, H. (2017). Conceptualizing Big Social Data. *Journal of Big Data*, *4*(1). <https://doi.org/10.1186/s40537-017-0063-x>
- Paasonen, S. (2013). Ihmisiä, kuvia, tekstejä ja teknologioita. Teoksessa S.-M. Laaksonen, J. Matikainen, & M. Tikka (toim.), *Otteita verkosta: Verkon ja sosiaalisen median tutkimusmenetelmät* (s. 34–48). Vastapaino.
- Riukulehto, S., & Haasio, A. (2020) Kotiseutukokemuksia verkkoympäristössä. Teoksessa S. Riukulehto, & A. Haasio (toim.), *Virtuaalinen kotiseutu* (s. 7–11). Tietolipas 266. SKS.
- Rogers, R. (2013). *Digital Methods*. MIT Press.

- Ruppert, E., Law, J., & Savage, M. (2013). Reassembling Social Science Methods: The Challenge of Digital Devices. *Theory, Culture & Society*, 30(4), 22–46. <https://doi.org/10.1177/0263276413484941>
- Savor, T., Douglas, M., Gentili, M., Williams, L., Beck, K., & Stumm, M. (2016). Continuous deployment at Facebook and OANDA. *Proceedings - International Conference on Software Engineering* (s. 21–30). <https://doi.org/10.1145/2889160.2889223>
- Sumiala, J., & Tikka, M. (2011). Visualizing Globalized Fears: School Shooting Videos and Circulation of Violence on YouTube. *Social Anthropology/Anthropologie Sociale*, 19(3), 245–267.
- Terras, M. M., Nyhan, J., & Vanhoutte, E. (2013). *Defining digital humanities: a reader*. Routledge.
- Tromble, R. (2019). In Search of Meaning: Why We Still Don't Know What Digital Data Represent. *Journal of Digital Social Research*, 1(1), 17–24. <https://doi.org/10.33621/jdsr.v1i1.8>
- Zimmer, M., & Kinder-Kurlanda, K. (2017). *Internet research ethics for the social age: new cases and challenges*. Sage.