

Haastattelu

Minttu Tikka, Noora Kotilainen ja Jukka Huhtamäki

Datateoria yhdistää tulkinnallisen sosiologian ja laskennalliset metodit

Digitaalinen sosiologi Simon Lindgren haastattelussa

Simon Lindgren on sosiologian professori ja DIGISUM-tutkimuskeskuksen johtaja Uumajan yliopistossa. Hänen kiinnostuksen kohteenaan ovat digitaalisen teknologian ja yhteiskunnan suhteet: kuinka digitaalinen (media)teknologia muuttaa yhteiskuntiamme ja miten dataistuminen ja alustojen algoritmit vaikuttavat yhteiskuntiin, politiikkaan ja valtaan.

Koodaava sosiologi Lindgren operoi jälkistrukturalistisen yhteiskuntateorian ja suurten media-aineistojen laskennallisen analyysin risteyksessä. Hän haluaa tuoda kriittisen teorian digitaalisen median aikaan ja osoittaa, että teorialla on paikkansa myös nykyisten medioituneiden yhteiskuntien tarkastelussa. Samalla hän on tyyppiesimerkki siitä, kuinka nyky-yhteiskuntien tarkastelijassa yhdistyy parhaimmillaan sekä syvä ymmärrys yhteiskuntateoriasta että kyky työntää kädet konkreettisesti syvälle digitalisoituvan yhteiskunnan tuottamaan dataan.

Lindgren on julkaissut monen suomalaisenkin media-aineistoja ja yhteiskuntaa tutkivan käsissä kuluneen *Digital media and society* -kirjan (2017, Sage). Hänen uusin kirjansa *Data Theory: Interpretive Sociology and Computational Methods* (2020, Wiley) tarkastelee dataistumisen mahdollistaman laskennallisen tutkimusotteen roolia yhteiskuntatieteellisen teorian kehittämisessä, reflektoinnissa ja tulkinnassa medioituneissa ja dataistuneissa yhteiskunnissa. Kirjassaan Lindgren sekä haastaa että pyrkii tuomaan yhteen yhteiskunnallisiin teorioihin vihkiytyneitä ja laskennallisia menetelmiä käyttäviä tutkijoita. Hän myös ehdottaa etnografiasta nousevan tulkinnallisen ja reflektiivisen otteen tuomista datatieteen kentälle.

Tapasimme Simon Lindgrenin keväällä 2021 Suomen Akatemian rahoittaman *Hybrid Terrorizing* (HYTE) -tutkimushankkeen työpajassa ja keskustelimme yhteiskuntateorian ja suurten sosiaalisen median aineistojen yhdistämiseen liittyvistä kysymyksistä väkivaltaisten mediatapahtumien tarkastelussa. Palasimme hänen uusimman kirjansa aiheisiin haastattelun merkeissä keväällä 2021.

**Teet tutkimusta yhteiskuntatieteen ja datatieteen risteyskohdassa. Miten olet pääty-
nyt siihen missä nyt olet?**

Aloitin tutkimuksen tekemisen keskittyen diskurssianalyysiin ja tekstien lähilukuun. Tuolloin analysoin mediaesityksiä eri aiheista, kuten nuorisokulttuurista ja rikollisuudesta kriittisen, jälkistrukturalistisen teorian ohjaamana. Tutkimukseni fokuksen siirtyessä internettutkimuksen suuntaan kiinnostuin yhä enemmän suurten aineistojen analysoimisesta. Olen kuitenkin ollut aina uskollinen sosiaalisen konstruktivismin epistemologialle. Uskon siihen, että se minkä uskomme olevan totta, on sosiaalisesti ja historiallisesti rakentunutta pikemminkin kuin luonnon ja biologian muovaamaa. Tämä pohja on johtanut minut nykypolulleni, pyrkimykseen sovittaa yhteen kriittistä yhteiskuntateoriaa ja datatieteen käytäntöjä. Uskon, ettei tarinani ole ainutkertainen, vaan pikemminkin se kertoo siitä polusta, jota pitkin koko digitaalisen sosiaalitieteen tutkimus ylipäänsä on kulkenut. Alun perin kulttuuritutkimuksen näkökulmat dominoivat, mutta nyt laskennallisia metodeja käyttävät hybridilähestymistavat ovat yhä yleisempiä.

Toimit DIGISUM -tutkimuskeskuksen johtajana Uumajan yliopistolla. Kerro lyhyesti DIGISUMista, miksi tällaista tutkimuskeskusta tarvitaan juuri nyt?

DIGISUM on Uumajan yliopistossa toimiva yhteistyöalusta useille tutkimusryhmille, jotka toimivat digitaalisen sosiaalitieteen alalla. Yliopistollamme toimii useita tutkimusryhmiä, jotka tarkastelevat digitaalisen ja yhteiskuntatieteellisen tutkimuksen risteyskohtia erilaisin hedelmällisesti limittyvin tavoin. Monitieteisyys, myös sosiaalitieteiden ulkopuolinen, on tietysti myös keskeistä.

Miksi kirjoitit uusimman kirjasi *Data Theory*? Mitä haluat sanoa kirjalla ja keiden haluaisit nähdä lukevan sitä?

Kirjoitin kirjan, koska minulla oli tarve järjestellä ajatuksiani siitä, kuinka tulkinnallisia ja teoreettisia lähestymistapoja sekä laskennallisia metodeja voidaan yhdistellä yhteiskuntatieteellisessä tutkimuksessa. Kirja onkin hieman manifestimainen avaus siitä, miten kriittistä yhteiskunta-analyysiä voidaan käyttää dataistuneessa maailmassa ja ajassamme. Kirja on toivottavasti hyödyllinen myös muille kuin itselleni. Ainakin se liittyy useiden nyt käynnissä olevien keskustelujen kanssa, ja olenkin varma, että moni muukin miettii juuri nyt niitä asioita, joita kirjassa pohdin.

Millaisena näet yhteiskuntatieteen teorioiden tulevaisuuden datafikaation aikakaudella? Ovatko suuret klassiset yhteiskuntateoriat yhä sovellettavissa datafikoituneeseen maailmaan vai tarvitsemmeko uusia teorioita?

Mielestäni tärkeintä on jatkaa teoreettista ajattelua mutta samalla olla kunnioittamatta liikaa suuria teorioita. Uskon teorioiden luovaan yhdistämiseen muuttuvassa maailmassa ja näen teorioiden päivittämisen, pikemmin kuin niiden korvaamisen, olevan osa tätä pyrkimystä.

Kirjassasi nostat esille tarpeen kohdata ja purkaa metodologista ja teoreettista dogmatismia sekä kyseenalaistaa yhteiskuntatieteellisten oppiaineiden väliin rakennettuja rajoja. Olisitko siis valmis kaatamaan perinteiset tieteiden väliset raja-aidat?

En henkilökohtaisesti ole oikea henkilö kaatamaan näitä rajoja. Kirjani antaa kuitenkin oman panoksensa laajempaan käynnissä olevaan kehitykseen, jossa nähdäkseni ai-

kamme suuria kysymyksiä ratkotaan yhä enemmän tieteidenvälisyydellä sekä monitieteisyydellä. Luulen, että monitieteisyys on lopulta melko yleistä. En siis usko, että olen useiden eri tieteenalojen välissä operoivana tutkijana ainutlaatuinen, sillä samankaltaisia asioita tekeviä tutkijoita on todella pitkä lista. Ja tämä on mainio asia.

Viittaat kirjassasi datatieteeseen eräänlaisena etnografiana. Mitä tarkoitat tällä epistemologisesti ja käytännöllisesti?

Yksinkertaisesti sanottuna se tarkoittaa, että etnografista tutkiskelevaa ajattelutapaa ja tutkimusotetta sekä totunnaisten tutkimuskäytänteiden reflektointia voidaan hyödyntää myös datatieteissä ja laskennallisessa yhteiskuntatieteessä. Pidän myös kiehtovana, että pyytämättä annetut ja sotkuiset aineistojen muodot, joita datatieteilijät käyttävät, voivat välillä tulla lähemmäs antropologista tiheää dataa kuin esimerkiksi kyselytutkimusten vastauksia.

Miksi laadullisen tutkimuksen perinteestä aloittanut sosiologi opetteli koodaamaan? Onko ohjelmointitaito muuttanut tapaasi lähestyä tutkimusta?

Ohjelmoinnin opettelu on ollut oma yksityinen projektini työn ja vapaa-ajan rajapinnassa. Pidän tietokoneista ja nörteistä jutuista, joten olen käyttänyt paljon aikaa teknisten taitojen kehittämiseen. Olen oppinut paljon aineistojen käsittelystä viimeisen viiden vuoden aikana. En koe, että uusi osaaminen on muuttanut tapani tehdä tutkimusta vaan päinvastoin: mikäli haluat tehdä tutkimusta tietyllä tavalla, sinun on omaksuttava tiettyjä uusia taitoja.

Datateoriassa yhteiskunnallinen teoria yhdistyy laskennallisen otteen toteutettavaan ohjelmakoodiin. Käydään seuraavaksi läpi käytännön esimerkki tästä lähestymistavasta. Kertoisitko meille tarkemmin kirjassasi kuvattavasta toimijaverkkoteoreettisesta analyysistä vaihe vaiheelta?

Toimijaverkkoteoreettisessa (*actor-network theory*) tutkimuksessa on tavoitteena kartoittaa teknisten ja ei-teknisten toimijoiden välisiä yhteyksiä. Mainitussa analyysissä käytän luonnollisen kielen prosessoinnissa (*natural language processing, NLP*) kehitettyä nimettyjen entiteettien tunnistamista (*named entity recognition*) erottaakseni ilmastomuutosta käsittelevästä tviittiaineistosta näitä toimijoita. Seuraavassa vaiheessa hyödynsin koneoppimismallia tunnistaakseni toimijoiden väliset yhteydet ja lopulta visualisoin yhteyksistä muodostuvan kokonaisuuden verkostona. Tämä lähestymistapa on vain yksi lukemattomista vaihtoehdoista yhdistellä erilaisia työvälineitä ja teorioita luovasti toisiinsa.

Data Theory -kirjassa kirjoitat auki monia sellaisia teemoja, joiden käsitteellistämisen parissa olemme kamppailleet tutkiessamme väkivaltaisia hybridejä mediatapahtumia HYTE-hankkeen monitieteisessä tutkimusryhmässä. Mitä haluaisit sanoa tutkijoille tai opiskelijoille, jotka haluavat tehdä tulkinallista yhteiskuntatiedettä digitaalisista lähteistä?

Uskon teorioiden ja metodien luovaan käyttöön enkä huolehdi liikaa siitä, miten yksittäiset palaset sopivat yhteen perinteisillä mittareilla tarkastellen. Liiallinen monialaisuus voi ehdottomasti aiheuttaa ongelmia erityisesti uraansa rakentaville tutkijoille, joiden toimintaa arvioidaan usein nimenomaan tieteen perinteestä käsin. Kannustan joka

tapauksessa kokeilemaan rohkeasti ja kehittelemään uudenlaisia tapoja tarkastella yhteiskuntia. Näkemykseni mukaan kaikkien tutkimuksessa tehtyjen valintojen säntillinen kuvailu ja perustelu tukee uusien toimintatapojen kehittämistä niin, että emme ajaudu umpikujaan, jossa “millään ei ole mitään väliä”.

Suurten aineistojen yleistymisen myötä on esitetty näkemyksiä datan kaikenkattavasta voimasta selittää ja ennustaa yhteiskunnallisia ilmiöitä ilman tarvetta teorioille. Lindgrenin *Data Theory* -kirja puhuu vahvasti tätä ajattelutapaa vastaan. Hänen mukaansa idea siitä, että “data puhuu puolestaan” on yhtä huono kuin ajatus teoriasta, joka ei lainkaan kohtaa empiiristä todellisuutta. Lindgrenin argumentin voisikin tiivistää siihen, että dataa ja teoriaa ei kannata pitää erillään, sillä nyky-yhteiskunnan ilmiöiden ymmärtäminen ja analysoiminen vaatii erottamattomasti molempia: kykyä kerätä ja käsitellä aineistoja sekä kriittistä ja reflektiivää yhteiskuntateoreettista tulkintaa, joka tähtää teorian kehittämiseen dataistuneessa ajassa.