

Bitit ja politiikka: Tervetuloa laskennallinen yhteiskuntatieteen tutkimus

MATTI NELIMARKKA JA
SALLA-MAARIA LAAKSONEN

Tietoyhteiskuntakehitys ja teknologian muutokset ovat vaikuttaneet yhteiskuntatieteisiin, mukaan lukien politiikan tutkimukseen. Digitaalisissa toimintaympäristöissä tapahtuva poliittinen toiminta näyttäytyy houkuttelevana tutkimuskohteena, ja toisaalta esimerkiksi digitaalisten alustojen ja algoritmien tutkimus nostaa esille politiikan perimmäisiä kysymyksiä vallasta (esim. Beer 2017; Gillespie 2010; Neyland ja Möllers 2016). Monet kiinnostavista kysymyksistä kytkeytyvät poliittiseen viestintään: sosiaalinen media on jo haastanut perinteisiä viestinnän portinvartijateorioita (esim. Castells 2007; Chadwick 2014) ja uudet digitaaliset viestintävälineet muuttavat kansalaisosallistumisen tapoja (esim. Bennett ja Segerberg 2013; Juris 2012). Myös marxilainen pohdinta on tehnyt paluun alustatalouden myötä tapahtuneen pääoman jakautumisen seurauksena (esim. Spencer 2018). Jo tämä tutkimusnäkökulmien lyhyt lista osoittaa, että politiikan tutkimusperinteet ovat tärkeässä roolissa myös nykyisen digitaalisen yhteiskunnan aikana.

Digitaalisuus ei muuta vain tutkimuskohteita, vaan myös aineistoja ja menetelmiä. Digitaaliset jalanjäljet (*digital trace data*) ja massadata (*big data*) mahdollistavat uudenlaisten kysymysten esittämisen: aiemmin tutkijoilla ei ollut käytettävissä samankaltaisia yksityiskohtaisia ja laajoja aineistoja ihmisten, organisaatioiden ja liikkeiden toiminnasta, vaan tutkimuksessa on turvaututtu havainnointiin, haastatteluihin, kyselyaineistoihin ja rekisteriaineistoihin. David Lazerin ja kumppaneiden (2009) mukaan uudet digitaaliset aineistot ja niitä hyödyntävät laskennalliset menetelmät ovat kuin uusi mikroskooppi yhteiskuntatieteelliseen tutkimukseen. Sekä Rob Kitchen (2014) että Danah Boyd ja Kate Crawford (2012) kehottavat tutkijoita kuitenkin kriittisesti arvioimaan niitä tapoja, joilla tutkimusta tehdään massadatan aikana, ja sitä, kuinka laskennalliset menetelmät

muokkaavat yhteiskuntatieteellistä tutkimusta. Hyvä esimerkki peräänkuulutetusta kriittisyydestä on Justin Grimmerin ja Brandon M. Stewartin (2013) artikkeli, jossa he perinteisiin laadullisiin lähestymistapoihin verraten pohtivat, miten tekstianalyysiä voidaan toteuttaa esimerkiksi sanojen esiintymisfrekvenssejä tarkastelemalla.

Tämän symposiumin artikkelit ovat esimerkkejä tällaisesta yhteistyöstä ja menetelmäkehityksestä. Symposium koostuu kolmesta toisiaan täydentävästä tekstistä. Kaksi ensimmäistä esittelevät laskennallisten menetelmien käyttöä politiikan tutkimuksen kentällä, kolmas pohtii koneoppimiseen perustuvien laskennallisten menetelmien institutionalisoitumista politiikan tutkimukseen. Tekstit siis omalta osaltaan vastaavat Grimmerin (2015) ehdotukseen pyrkiä muodostamaan selkeämmin yhteiskuntatieteellisesti painottunut näkökulma laskennallisten menetelmien käyttöön ja kehitykseen.

Salla-Maaria Laaksosen ja Matti Nelimarkan artikkeli tutkii digitaalista vaalijulkisuutta vuoden 2015 eduskuntavaaleissa. Tutkimuksessa laskennallisesti analysoidaan vaalien julkisella agendalla olleet teemat ja yhdistetään saatua tietoa toisaalta poliittisen viestinnän agendatutkimuksen teorioihin ja puolueiden aiheomistajuuden analyysiin. Tuukka Ylä-Anttila, Veikko Eranti ja Anna Kukkonen taas käsittelevät katsauksessaan ilmastonmuutoksesta käytyä julkista keskustelua aihehallinnuksen avulla. Kirjoittajat käyvät läpi menetelmän reunaehdoja ja ehdottavat laadullista validointiprosessia, jonka avulla menetelmää voisi käyttää tekstien kehysanalyysinä.

Molemmat tekstit tarkastelevat agendan muodostumista laskennallisesti ja osoittavat samalla, että laskennalliset menetelmät voivat tarjota uusia työkaluja poliittisten argumenttien tutkimiseen ja sellaisiin politiikan ja poliittisen viestinnän polttaviin klassisiin kysymyksiin kuten agendan rakentaminen ja teemo-

jen kehystäminen. Ennen kaikkea menetelmät mahdollistavat tällaisen analyysin tekemisen paljon aiempaa laajemmilla aineistoilla. Molemmat tekstit käyttävät menetelmänä ohjaamatonta koneoppimista, tarkemmin aihehallinnusta, mutta sitovat valitun menetelmän perinteiseen yhteiskuntatieteelliseen kysymyksenasetteluun. Lisäksi tekstit käyvät keskustelua laskennallisia menetelmiä soveltavan yhteiskuntatieteen käsitteiden kanssa –nähdäksemme tämä ei ole vain tarpeellinen, vaan myös välttämätön keskustelu.

Professori Pertti Ahonen luo katsauksessaan näkyvän laskennallisten menetelmien institutionalisoitumiseen politiikan tutkimuksessa. Hän keskittyy nimenomaisesti laskennallisiin menetelmiin, joita on kehitetty politiikan tutkimuksen institutionalisoituneiden kysymysten tarkasteluun politiikan tutkijoiden toimesta. Ahonen päätyy toteamaan, että laskennallisten menetelmien käyttö politiikan tutkimuksessa on yhä sivupolku, ja varsinkin suomalaisessa politiikan tutkimuksessa melko vähäistä. Ahonen myös aiheellisesti peräänkuuluttaa syvällisempää keskustelua menetelmien filosofisista taustaoletuksista.

Menetelmäkeskustelua onkin yhä syytä käydä, ja sitä tulisi käydä poikkitieteellisesti. Poikkitieteellisyyden haasteeseen on herätty myös tietojenkäsittelytieteilijöiden joukossa (vrt. Wallach 2018). Oleellista on, että vaikka laskennallisia menetelmiä voi usein soveltaa suoraan “out of the box”, ne eivät ole taikalaitikoita, jotka ratkaisevat aiemmat tutkimukseen liittyvät ongelmat ja luotettavuuskysymykset. Laadullista tarkastelua ja teorialähtöisyyttä tarvitaan yhä rinnalle. Robotti ei vie politiikan tutkijan töitä, kuten Tuukka Ylä-Anttila ja kumppanit toteavat analyysin sa päätteeksi – eikä ehkä datatieteilijäkään.

LÄHTEET

- Beer, David. 2017. The social power of algorithms. *Information, Communication & Society* 20:1, 1–13.
- Bennett, Lance ja Segerberg, Alexandra. 2013. *The logic of connective action: Digital media and the personalization of contentious politics*. Cambridge: Cambridge University Press.
- Boyd, Danah ja Crawford, Kate. 2012. Critical questions for Big Data. *Information, Communication & Society* 15:5, 662–679.
- Castells, Manuel. 2007. Communication, power and counterpower in the network society. *International Journal of Communication* 1:29, 238–266.
- Chadwick, Andrew. 2013. *The hybrid media system: Politics and power*. Oxford: Oxford University Press.
- Gillespie, Tarleton. 2010. The politics of “platforms.” *New Media and Society* 12:3, 347–364.
- Grimmer, Justin. 2015. We are all social scientists now: How Big Data, machine learning, and causal inference work together. *PS: Political Science & Politics* 48:01, 80–83.
- Grimmer, Justin ja Stewart, Brandon M. 2013. Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis* 21:3, 267–297.
- Juris, Jeffrey. 2012. Reflections on #Occupy Everywhere: Social media, public space, and emerging logics of aggregation. *American Ethnologist* 39:2, 259–279.
- Kitchin, Rob. 2014. Big Data, new epistemologies and paradigm shifts. *Big Data & Society* 1:1, 1–12.
- Lazer, David, Pentland, Alex, Adamic, Lada, Aral, Sinan, Laszlo Barabasi, Albert, Brewer, Davon, Christakis, Nicholas, Contractor, Noshir, Fowler, James, Gutmann, Myron, Jebara, Tony, King, Gary, Macy, Michael, Roy, Deb ja Van Alstyne, Marshall. 2009. Life in the network: the coming age of computational social science. *Science* 323:5915, 721–723.
- Neyland, Daniel ja Möllers, Norma. 2016. Algorithmic IF ... THEN rules and the conditions and consequences of power. *Information, Communication & Society* 4462, 1–18.
- Spencer, David. 2018. Fear and hope in an age of mass automation: debating the future of work. *New Technology, Work and Employment* 33:1, 1–12.
- Wallach, Hanna. 2018. Computational social science ≠ computer science + social data. *Communications of the ACM* 61:3, 42–44.