

# Laskennallinen yhteiskuntatiede

Matti Nelimarkka

Helsinki 5.5.2011

LuK -tutkielma

HELSINGIN YLIOPISTO

Tietojenkäsittelytieteen laitos

Tiedekunta — Fakultet — Faculty		Laitos — Institution — Department	
Matemaattis-luonnontieteellinen tiedekunta		Tietojenkäsittelytieteen laitos	
Tekijä — Författare — Author			
Matti Nelimarkka			
Työn nimi — Arbetets titel — Title			
Laskennallinen yhteiskuntatiede			
Oppiaine — Läroämne — Subject			
Työn laji — Arbetets art — Level		Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages
LuK –tutkielma		5.5.2011	
Tiivistelmä — Referat — Abstract			
<p>Työssä esitellään laskennallisten menetelmien vaikutusta yhteiskuntatieteelliseen tutkimukseen kolmantena laajana menetelmällisenä perheenä laadullisen ja määrällisen tutkimuksen lisäksi. Laskennallinen yhteiskuntatiede (<i>computational social science</i>) soveltaa moderneja laskennallisia menetelmiä, kuten kone-oppimista ja tilastollista analyysiä, yhteiskuntaa koskevien kysymysten esittämiseen ja tulkintaan.</p> <p>Erityisesti työssä keskitytään n menetelmään, lista. Menetelmät on valittu esittämään tällä hetkellä käytössä olevia laskennallisen yhteiskuntatieteen sovelluskohteita. Kunkin menetelmän kohdalta esitetään lyhyesti sen laskennallinen teoria sekä tutkimuksia, missä kyseistä menetelmää on sovellettu.</p> <p>Yksittäisten tutkimusten tarkastelun kautta voidaan nostaa esille suurempia teemakokonaisuuksia laskennallisen yhteiskuntatieteen kehityksen kannalta. Ensimmäisenä käsittelen laskennallisen yhteiskuntatieteen validiteettia ja reliabiliteettia, eli tulosten oikeellisuutta ja uskottavuutta. Toinen merkittävä teema on monitieteellinen työskentely, minkä koen välttämättömänä laskennallisen yhteiskuntatieteen alalla, mutta tunnistan siihen liittyvät ongelmat. Viimeisenä keskustelun osiona pohdin, voiko laskennallinen tiede edustaa kuhnilaisessa mielessä paradigman muutosta, millä olisi laajempi vaikutus yhteiskuntatieteen kehitykseen ja tulevaisuuteen.</p>			
Avainsanat — Nyckelord — Keywords			

# Sisältö

<b>1</b>	<b>Johdanto</b>	<b>1</b>
<b>2</b>	<b>Agenttipohjainen simulaatio</b>	<b>2</b>
<b>3</b>	<b>Menetelmä n</b>	<b>2</b>
<b>4</b>	<b>Keskustelu</b>	<b>2</b>
4.1	Validiteetti ja reliabiliteetti haasteena . . . . .	2
4.2	Työskentely monitieteellisesti . . . . .	2
4.3	Paradigman muutos? . . . . .	2
<b>5</b>	<b>Johtopäätökset</b>	<b>2</b>

# 1 Johdanto

Tässä työssä yhteiskuntatiede nähdään laajasti yhteiskuntaa tutkivien tieteiden joukkona, joihin kuuluvat esimerkiksi politiikan tutkimus, taloustiede, sosiologia, sosiaalipsykologia ja viestinnän tutkimus<sup>1</sup>. Tutkimusaloille tyypillistä on samankaltaisten tutkimuskysymysten lähestyminen kunkin alan oman oppihistorian sekä käsitteistön kautta [?]. Esimerkiksi poliitikkojen sosiaalisen median käyttöä voidaan lähestyä niin politiikan tutkimuksen, viestinnän tutkimuksen kuin esimerkiksi sosiologian näkökulmasta. Esimerkki.

Vaikka tutkimuskysymykset ja niiden painotukset eroavat toisistaan, niin yhdistävänä tekijänä yhteiskuntatieteissä voidaan – yhteiskunnallisten ilmiöiden tutkimuksen lisäksi – pitää kahden laajemman lähestymistavan soveltaminen tutkimuskysymyksiin. Tutkimuksessa sovelletaan yleensä laadullista tai määrällistä tutkimusotetta, missä ensimmäisessä pääpaino on enemmän kuvailevassa aineistossa ja sen tulkinnessa, jälkimmäinen taas pyrkii mitattaviin aineistoihin ja selittämiseen matemaattisia menetelmiä hyödyntäen [??]. Sellaisenaan voimakasta menetelmäjakoa on kritisoitu, koska useiden ilmiöiden kohdalla olisi mielekästä soveltaa molempia tutkimusmenetelmiä: laadullisen tutkimuksen perinteinen haaste on tutkimuksen yleistettävyyden koko otosjoukkoon, kun taas määrällisen tutkimuksen tapauksessa varsinaisen kausaalisen tekijän voi olla haastava esittää [???].

Esimerkkejä määrällisestä lähestymistavasta ovat tilastollisten menetelmien soveltaminen aineiston käsittelyssä, mutta myös formaalit menetelmät, kuten peliteoria voidaan laskea osaksi tätä lähestymistapaa [?]. Myös tarkastelun kohteena oleva laskennallinen yhteiskuntatiede (*computational social science*) edustaa määrällistä lähestymistapaa. Laskennalliset tieteet määritellään [???]. Vastaava esitys on myös nähtävissä laskennallisen yhteiskuntatieteen määritelmässä:

**Cioffi-Revilla [2010]** määrittelee laskennallisen yhteiskuntatieteen laajasti seuraavien laskennallisten menetelmien soveltamisena: tiedon uuttamisen menetelmät (*information extraction*), sosiaalisten verkostojen analyysi, paikkatietojärjestelmien käyttämisen ja mallinnuksen sekä simulaation menetelmät. Hän myös arvioi, että tiedon visualisointimenetelmät voivat myöhemmin laajentua käytettäväksi laskennallisen yhteiskuntatieteen menetelmänä. Hän siis näkee, että laskennallista tiedettä voidaan määritellä tiettyjen menetelmien käyttönä.

**Lazer et al. [2009]** taas esittää laskennallisen yhteiskuntatieteen laajojen aineistomäärien keräämisenä ja käsittelynä määrällisten menetelmien avulla. He nostavat esille, että laskennallinen yhteiskuntatiede on aineistolähtöistä (*data-driven*), minkä tulkitsemisen tarkoituksena, että perinteinen hypoteeseihin ja niiden tarkasteluun perustuvat tilastolliset menetelmät eivät olisi

---

<sup>1</sup>Tässä työssä esimerkit on erityisesti valittu politiikan tutkimuksen näkökulmasta, mutta esitetyt menetelmät ovat laajemmin sovellettavissa laajemminkin yhteiskuntatieteisiin.

laskennallista yhteiskuntatiedettä.

**Bankes et al. [2002]** kuvaavat laskennallisia epistemologioita (*computational epistemology*), eli käsityksiä hyvästä tieteestä. Toisaalta he ehdottavat, että laskennallinen yhteiskuntatieteen tarkoitus on luoda selittäviä malleja, joiden avulla yhteisöä voidaan tutkia. Toisaalta, he nostavat esille myös mahdollisuuden käyttää näitä malleja ei vain yhteisön tutkimiseen vaan myös laskennallisten koeasetelmien suorittamiseen. Tässä määritelmässä siis tärkeintä on tutkittavan ilmiön mallinnus laskennallisilla menetelmin.

Kuten havaitaan, tutkimusalan määritelmät eivät ole kaikilla tutkijoilla samankaltaisia, johtuen myös erilaista lähestymistavoista laskennallisen yhteiskuntatieteen määritelmään. Kuten Cioffi-Revilla [2010] huomauttaa, laskennallisen tieteen osalta voidaan erottaa laskennallisuus menetelmänä ja laskenta teoreettisena lähtökohtana. Jaottelu ei kuitenkaan ole ongelmaton: esimerkiksi nykyisin tilastollisten menetelmien kohdalla käytetään poikkeuksetta laskennallista välineistöä. Toisaalta, sosiaalisen verkostojen analyysin kohdalla taustalla on graafiteoria ja sen sovellukset, vaikkakin niitä suoritetaan laskennallisesti.

Tietenkin luonnollinen syy laskennallisten menetelmien yleistymiseen on tietoteknisten resurssien laajempi saavutettavuus. Lisäksi uudet laajat tietovarastot, niin Internetin [Adamic and Glance, 2005] kuin elinympäristöstä kerättyjen sensoriaineistojen [Eagle and Pentland, 2006, Oulasvirta et al., 2012], yleistyminen ja käyttöönotto on helpottanut suurten aineistojen keräämistä. Ensinnäkin, tällöin perinteiset ei-laskennalliset menetelmät eivät välttämättä toimi, ja samaan aikaan laajat aineistot mahdollistavat aiemmin haastavina pidettyjen menetelmien käytön – esimerkiksi verkosto-analyysin ongelma on ollut aineiston vaikea keräys [?].

## 2 Agenttipohjainen simulaatio

### 3 Menetelmä n

### 4 Keskustelu

#### 4.1 Validiteetti ja reliabiliteetti haasteena

#### 4.2 Työskentely monitieteellisesti

#### 4.3 Paradigman muutos?

## 5 Johtopäätökset

## Lähteet

- L. A. Adamic and N. Glance. The political blogosphere and the 2004 u.s. election: divided they blog. In *Proceedings of the 3rd international workshop on Link discovery*, LinkKDD '05, pages 36–43, 2005.
- S. Bankes, R. Lempert, and S. Popper. Making computational social science effective. *Social Science Computer Review*, 20(4):377–388, 2002.
- C. Cioffi-Revilla. Computational social science. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(3):259–271, 2010.
- N. Eagle and A. S. Pentland. Reality mining: sensing complex social systems. *Personal Ubiquitous Computing*, 10(4):255–268, 2006.
- D. Lazer, A. Pentland, L. Adamic, S. Aral, A.-L. Barabási, D. Brewer, N. Christakis, N. Contractor, J. Fowler, M. Gutmann, T. Jebara, G. King, M. Macy, D. Roy, and M. Van Alstyne. Life in the network: the coming age of computational social science. *Science*, 323(5915):721–723, 2009.
- A. Oulasvirta, A. Pihlajamaa, J. Perkiö, D. Ray, T. Vähäkangas, T. Hasu, N. Vainio, and P. Myllymäki. Long-term effects of ubiquitous surveillance in the home. In *Proceedings of The 14th International Conference on Ubiquitous Computing*, 2012.