Elina Late

AINEISTOT AVOIMIKSI YHTEISKUNTATIETEISSÄ – MUTTA MISTÄ KÄYTTÄJÄT?

Yhteiskuntatieteellisillä aloilla on pitkät perinteet aineistojen uudelleen käytölle tutkimuksessa. Esimerkiksi Maailmanpankin ja OECD tarjoamia tilastoaineistoja on hyödynnetty vuosikymmeniä tutkimuksessa. Tutkimuksen digitalisoituminen ja vaatimus tieteen avoimuudesta ovat mahdollistaneet aineistojen jakamisen organisaatioiden ohella myös yksittäisille tutkijoille. Tänä päivänä tutkijoiden tallentamien aineistojen voisi olettaa muodostavan valtaosan tutkimusdatojen säilyttämiseen ja jakamiseen erikoistuneiden arkistojen sisällöstä.

Tampereen yliopiston yhteydessä toimivaan Yhteiskuntatieteelliseen tietoarkistoon on tallennettu vuodesta 1999 lähtien yhteensä noin 1700 tutkimusaineistoa, joista yli 80 prosenttia on yhteiskuntatieteille tyypillisiä kysely- ja tilastoaineistoja. Suomalaisen data-arkiston kaltaisia yhteiskuntatieteisiin keskittyneitä arkistoja on ympäri Eurooppaa (ks. CESSDA). Yksi esimerkki on Sveitsissä Lausannen yliopiston yhteydessä toimiva FORSbase.

Molempien arkistojen käyttötilastot osoittavat vuosittain kasvavia aineistojen latausmääriä. Kuitenkin tarkempi tarkastelu paljastaa, että merkittävä osa latauksista keskittyy vain pienelle osalle aineistoja. Esimerkiksi Sveitsissä 40 prosenttia latauksista kohdistui kymmenelle ladatuimmalle aineistolle. Nämä aineistot ovat tyypillisesti vakiintuneiden instituutioiden keräämiä karttuvia pitkittäisaineistoja, kuten European Social Survey.

Aineistojen käyttäjien profiilit paljastavat opiskelijoiden muodostavan suurimman käyttäjäryhmän sekä Suomessa että Sveitsissä. Sveitsiläinen aineisto myös osoittaa, että tutkimuskäyttöön ladatuista aineistoista vain noin viidennes ladataan julkaisua ajatellen. Vielä onkin epäselvää, mitä kaikkea aineistojen uudelleenkäyttö voi lopulta tarkoittaa.

Näiden tietojen pohjalta herää kysymys, löytääkö yksittäiseen tutkimushankkeeseen kerätty aineisto arkistosta käyttäjää ja voidaanko kaikkia aineistoja lopulta hyödyntää tutkimuksessa? Tietoarkiston tilastojen perusteella suurin osa aineistosta ladataan kuitenkin ainakin kerran. Tästä huolimatta aineistojen käyttö näyttää noudattelevan samaa logiikkaa julkaisujen jatkokäytön kanssa; viittaukset kasaantuvat pienelle joukolle julkaisuja ja tekijän tunnettuus lisää käyttöä.

Toisaalta tutkimusaineistojen laajempi uudelleen käyttö yhteiskuntatieteissä saattaa vaatia perustavanlaatuista muutosta tieteenalalle ominaiselle tavalle tuottaa uutta tietoa. Jo kerätty tieto saattaa odottaa löytäjäänsä. Tilanne herättää informaatiotutkimuksen perinteistäkin perinteisemmän kysymyksen; miten saadaan oikea tieto oikealle käyttäjälle? Kirjastoilla, jos kenellä on tähän tehtävään asiantuntemusta annettavana.

ELINA LATE

Kirjoittaja toimii Tampereen yliopiston Informaatioteknologian ja viestinnän tiedekunnassa yliopistotutkijana. Tällä hetkellä hän työskentelee Suomen Akatemian rahoittamassa ImAccess projektissa, jossa tutkitaan avointen kuva-aineistojen käyttöä tutkimusaineistona.