



Kuva: Canva.

Anna-Sofia Ruth ja Päivi Kanerva

Dataviittausten vakiintuminen edellyttää muutoksia työtapoihin ja tutkijan arviointiin

Jokainen tutkimus nojaa aina jollakin tavalla aiemmin tuotettuun tutkimukseen. Tutkijalla on eettinen velvollisuus tuoda esille aiempi tieteellinen panos, jolle hän tutkimuksensa rakentaa. Viittaaminen muiden tekemiin tutkimustuotoksiin on itsestään selvyyttä, kun puhutaan tieteellisistä julkaisuista, mutta tutkimusaineistoihin viittaaminen sen sijaan ei ole vielä vakiintunutta.

Tähän problematiikkaan pureutui Suomen tiedekustantajien liiton, Tietoarkiston ja Tampereen yliopiston kirjaston 8. 3. 2023 järjestämä hybridiseminaari *Dataan viittaaminen julkaisutoiminnassa*. Osallistujia oli noin sata, joista valtaosa etäyhteydellä. Paikan päälle Tampereen yliopiston pääkampukselle saapui kuitenkin kolmisenkymmentä osallistujaa, mikä vielä luo toivonkipinän koronanjälkeisten livetapahtumien tulevaisuudesta.

Seminaarin avasi Susanna Nykyri,

joka on Tampereen yliopiston avoimen tieteen johtava asiantuntija sekä Tampere University Pressin kustannuspäällikkö. Nykyri totesi alustuksessaan, että dataan viittaaminen julkaisutoiminnassa ei periaatteena ole mitään uutta ja ihmeellistä, vaan lähtökohtana on hyvä tieteellinen käytäntö, jonka keskeisiä elementtejä ovat tunnustuksen antaminen muiden tutkijoiden työlle sekä eettisesti kestävä tiedonhankinta-, tutkimus- ja arviointimenetelmät.

Tutkimuksen vastuullista arviointia

Ohje tieteellisille julkaisukanaville vastuullisen aineisto- ja datapolitiikan laatimiseksi

Vastuullisen tieteen julkaisusarjassa on ilmestynyt ohjeistus, jota tieteelliset julkaisijat voivat hyödyntää oman data- ja aineistopolitiikkansa laatimisessa. Malli sisältää seitsemäntoista kohtaa:

1. Yleinen kommentti avoimen datan ja aineiston hyödyistä, joiden avulla perustellaan politiikassa esitetyt linjaukset.
2. Poliittikadokumentin omistajan nimeäminen.
3. Muistutus aineistojen ja datan alkuperäisten julkaisujen nimeämisestä.
4. Poikkeustapaukset, joissa linjauksista on mahdollista poiketa.
5. Kannanotto datan säilytyspalveluun tallentamisen pakollisuudesta tai vapaaehtoisuudesta.
6. Datat ja aineiston säilytyspalveluita koskevat suositukset.
7. Datat tai aineiston linkittämistä tai liittämistä koskevat ohjeet.
8. Linkittämisen muotoa koskevat ohjeet.
9. Liitetiedostona olevaa dataa ja aineistoa koskevat ohjeet.
10. Laajemmasta datavarannosta valikoituja osia koskevat ohjeet.
11. Tutkimusdataan viittaamista koskevat ohjeet.
12. Datat ja aineiston saatavuus vertaisarviointiprosessin aikana.
13. Datat avaamisen embargoa koskeva linjaus.
14. Eettisiä ja juridisia kysymyksiä koskevat kannanotot.
15. Sensitiivisen datat ja aineistojen jakamista koskevat ohjeet.
16. Datat tai aineiston omistus- ja käyttöoikeuksien selvittämistä koskevat ohjeet.
17. Ohjeet datat analyysin kuvaamisesta.

(Avoimen tieteen ja tutkimuksen koordinaatio 2020).

koskevissa linjauksissa korostetaan entistä painokkaammin tutkimustuotosten moninaisuuden huomioimista, joten data tulisi nähdä yhtä konkreettisena ja merkityksellisenä tuotoksena kuin vertaisarvioidut tutkimusjulkaisutkin. Tuoreita esimerkkejä tällaisista vastuullisen arvioinnin linjauksista ovat esimerkiksi CoARA:n *The Agreement on Reforming Research Assessment* ja *Tutkijanarvioinnin hyvät käytännöt: Vastuullisen tutkijanarvioinnin kansallinen suositus*.

Vastuuta dataviittausten vakiinnuttamisesta on niin tutkimusorganisaatioilla, tutkijoilla, tiedepoliittisilla toimijoilla, infrastruktuurien tarjoajilla, rahoittajilla



Kannattaa hyödyntää olemassa olevia palveluinfrastruktuureja, kuten Tietoarkistoa ja Kielipankkia.

kuin julkaisijoilla. Seuraavassa puheenvuorossa Tietoarkiston johtaja Helena Laaksonen esitteli dataan viittaamista tiedekustantajien ohjeistuksissa.

Dataan viittaaminen toistaiseksi harvinaista

Vuonna 2022 Helena Laaksonen tarkasteli yhdeksän kotimaisen ihmistieteellisen kausijulkaisun kirjoittajaohjeita selvittääkseen, löytyykö niistä mainintoja dataan viittaamisesta. Ainakaan tuolloin ohjeita ei vielä löytynyt. Haastatteluiden ja sähköpostikeskusteluiden perusteella osassa tapauksista syynä oli pelko, että lehden pitäisi pystyä tarjoamaan

tallennuspaikka datalle, mikä ei muutenkin vähäisten resurssien kanssa kamppailevia toimituksia innostanut.

Laaksonen totesi puheenvuorossaan, että tiedejulkaisijat olisi kuitenkin tärkeää saada mukaan. Tämä ei tarkoita, että jokainen kotimainen tiedejulkaisija ryhtyisi ylläpitämään omia data-arkistoja, vaan kannattaa hyödyntää olemassa olevia palveluinfrastruktuureja, kuten Tietoarkistoa ja Kielipankkia.

Muutakin tukea ja ohjeistusta on tarjolla. Esimerkiksi Tietoarkisto antaa kaikille aineistoilleen malliviitetiedon, jonka kyseistä aineistoa hyödyntävä kirjoittaja voi lisätä lähdeluettelonsa suoraan kopioituna tai halutun viittaus-tyylin mukaiseksi muokattuna (Tieto-

arkisto 2023). Malliviittausten rakenne perustuu SFS-standardiin, joka määrittää, mitä elementtejä sen tulee sisältää.

Laaksonen mukaan yhteiskuntatieteiden julkaisuissa näkee harvoin malliviitteitä, ja omaan dataan viittaaminen on vielä harvinaisempaa. Tyypillisesti tutkimusaineistosta kerrotaan tai sitä kuvaillaan tekstissä. Tällöin lukija ei yleensä saa tietoa siitä, mistä datan saa käyttöönsä. Jos dataa lisäksi ajatellaan meritoivana tutkimustuotoksena, sen pitäisi olla viitattavassa muodossa.

Kustantaja tasapainoilee suosittelun ja edellyttämisen välissä

Kustannuspäällikkö Susanna Nykyri toi tapausesimerkiksi Tampere University Pressin, jonka julkaisuprofiili keskittyy juuri ihmistieteisiin, erityisesti sosiaalitieteisiin. TUP on julkaissut vuonna 2021 oman data- ja aineistopolitiikan hyvän aineistonhallinnan edistämiseksi. Kustantamo ei edellytä, mutta suosittelee datan ja vähintäänkin metadatan avaamista. Lisäksi toiveena on, että tutkijat lähettäisivät aineistonhallintasuunnitelman kirjakäsikirjoituksensa ohessa.

Koska aineistonhallinta ja dataan viittaaminen saattavat olla osalle kirjoittajista suhteellisen tuntemattomia asioita, kustantajan pitää tasapainoilla ohjeistuksissaan edellyttämisen ja suosittelun välillä. Tukipalvelujen pariin ohjaaminenkaan ei ole yksiselitteistä: Tampereen yliopiston datatukipalvelut ovat vain oman organisaation tutkijoi-

den käytössä, mutta kustantamoon voi tarjota käsikirjoitusta mihin tahansa organisaatioon affilioitunut tutkija.

Kansainvälisten DORA- ja CoARA -julistusten myötä tutkimuksen ja tutkijan arviointia on alettu kehittää kohti monimuotoisempaa osaamisen tunnustamista. Vastuullisten arviointikäytäntöjen jalkautuminen yliopistoihin ja muualle tiedeyhteisöön on oleellista myös tiedekustantajan näkökulmasta. Jos paine julkaisujen tuottamiseen vähenisi, voisivat kirjoittajat

ehkä panostaa enemmän laatuun, kuten huolelliseen aineiston- ja viitteidenhallintaan sekä aineiston avaamiseen.

Näkymätön työ näkyväksi

Tutkijapuheenvuoron seminaarisessa antoi yhteiskuntatieteiden tohtori Josefina Sipilinen. Hän työskentelee tutkijatohtorina Suomen Akatemian rahoittamassa hankkeessa *Education, Political Efficacy and informed Citizenship* (EPIC) sekä Kunnallissalan kehittämissäätiön rahoittamassa tut-

Kuva: Canva.

Selvitys tutkimusaineistojen saatavuusilmoituksista tieteellisissä artikkeleissa

Niin sanottu *Data Availability Statement* on julkaisuun lisätty ilmoitus, jossa kerrotaan tutkimuksessa käytetyn aineiston saatavuudesta ja käyttömahdollisuuksista. Tampereen yliopiston kirjaston asiantuntijat Tomi Toikko ja Kaisa Kylmälä analysoivat yhteensä 2 085 tieteellistä vertaisarvioitua lehtiartikkeliä vuosilta 2020 ja 2021. Analysoiduista artikkeleista noin 26 prosenttiin sisältyi *Data Availability Statement*.

Selvityksen johtopäätökset:

- DAS-ilmoitusten sisältö ja sijainti artikkeleissa vaihtelee.
- Lyhyen aikavälin tarkastelussa DAS-ilmoitusten määrä näyttää olevan nousussa.
- Kustantajien ohjeilla ja vaatimuksilla on merkitystä käytäntöjen yhdenmukaisuuden kannalta.
- Avoimiin julkaisuihin liittyy useammin avointa dataa ja DAS-ilmoituksia.
- Tässä tarkastelussa avointa dataa oli melko vähän, mutta kansainvälisellä yhteistyöllä ja yritys-yhteistyöllä oli myönteinen vaikutus datan avoimuuteen.

(Toikko & Kylmälä 2023.)

kimushankkeessa *Tulevaisuuden kuntapäätäjät*.

Sipinen nosti esiin ristiriidan, joi- ta suuria kyselyaineistokeräyksiä toteuttavat tutkijat kokevat arkipäivän tutkimustyössään: pelkästään kyselylo- makkeen ja tutkittavien informointi- ja suostumusdokumenttien suunnitte- luun, saati itse aineiston keruuseen menee tutkijan työaika huomattavia määriä. Käytännön esimerkkinä Sipi- nen kertoi 4 000 kirjeen postituksesta printtaamisineen, postimerkkien lii- maamisineen ja paperiviiltoineen. Sil- ti aineistot eivät tuota samankaltaista meriittiä kuin tieteellisten julkaisujen kirjoittaminen.

Tuotos per panos -suhde ei siis var- sinaisesti kannusta näkemään vaivaa. Vähintä, mitä tutkijat haluavat, on että

heidän tekemänsä työ tulee näkyväksi kaikessa monimuotoisuudessaan. Myös tutkimusaineistojen keruuseen käytet- ty aika ja siitä hankittu kokemus tulisi huomioida. Tutkimustyön monipuoli- nen palkitseminen mahdollistaisi sen, että tutkimusryhmän aineistonkeruus- sa taitavat ja siitä kiinnostuneet tutki- jat voisivat halutessaan keskittyä tähän tutkimusprosessin osa-alueeseen.

Vaikka Sipinen itse kokee, että yksi parhaista tavoista oppia tutkimuksen teon moninaisuutta on koota oma ai- neisto, kannustaa hän työssään opis- kelijoita käyttämään myös valmiita aineistoja sekä yhdistelemään niitä it- se kerättyihin. Uudelleenkäyttö voisi vähentää paljon tutkittujen ryhmien kokemaa ”kyselyähkyä”, ja se on lisäksi kaikin puolin taloudellista. ♦

Lähdeviitteet

Avoimen tieteen ja tutkimuksen koordinaatio, ILVA, J., NYKYRI, S., MUSTAJOKI, H., PARLAND-VON ESSEN, J., SYRJÄMÄKI, S. (2020). Ohje tieteellisille julkaisukanaville vastuullisen aineisto- ja datapolitiikan laatimiseksi. Vastuullisen tieteen julkaisusarja 4/2020. Helsinki: Tiedonjulkistamisen neuvottelukunta ja Tieteellisten seurain valtuuskunta. <https://doi.org/10.23847/isbn.9789525995220>

Tietoarkisto (2023). Aineistoon viittaaminen. <https://www.fsd.tuni.fi/fi/aineistot/aineistojen-kaytto-ja-lataaminen/aineistoon-viittaaminen/> (viitattu 13. 2. 2023 klo 10.28)

TOIKKO, T. & KYLMÄLÄ, K. (2023). Data Availability Statements in the 2020 and 2021 scientific publications of Tampere University (slides). Zenodo. <https://doi.org/10.5281/zenodo.7655892>

Kirjoittajat

ANNA-SOFIA RUTH
Tampereen yliopisto
Tampere University Press
anna.ruth@tuni.fi
orcid.org/0000-0003-0662-4415

PÄIVI KANERVA
Tampereen yliopisto
paivi.kanerva@tuni.fi
orcid.org/0000-0001-9570-6177

