



HANKKEEN TULOSTEN AVAAMINEN

Tavoitteena on, että tutkimusaineistot ja -menetelmät ovat mahdollisimman avoimia.

- Tutkimusaineistojen ja -menetelmien avoimuus Korkeakoulu- ja tutkimusyhteisön kansallinen linjaus ja toimenpideohjelma 2021–2025:
 - [Osalinjaus 1: Tutkimusdatan avoin saatavuus sekä tähän linjaukseen liittyvät suositukset.](#)
 - [Osalinjaus 2: Tutkimusmentelmien ja infrastruktuurien avoin saatavuus](#)
- Hankkeen aineistojen avoimella saatavuudella edistetään niiden jatkokäyttöä. Avoimet aineistot tuovat myös näkyvyyttä ja meriittiä tekijälleen.
- Avoimuus lisää läpinäkyvyyttä ja uskottavuutta, toistettavuutta ja todennettavuutta sekä sovellettavuutta uusiin käyttötarkoituksiin.
- Tutkimuksen julkaisijat ja rahoittajat edellyttävät yhä useammin, että tutkimuksen tausta-aineistot tuodaan mahdollisimman avoimesti saataville viimeistään tutkimuksen päätyttyä.
- Aineistoilla voi perustelluista syistä olla erilaisia avoimuuden asteita vaihdellen kaikille täysin avoimesta salaiseen. Vähintään aineiston metadatan eli kuvailutiedon tulisi olla avointa.
- Tutkimusdatan avaamisen suhteen on hyvä pitää ohjenuorana Horizon 2020:n Open Research Data -pilotin periaate ”As open as possible, as closed as necessary”.

Tutkimusjulkaisujen avoin saatavuus – Tutkimusyhteisön kansallinen linjaus ja toimenpideohjelma 2020–2025: [Lehti- ja konferenssiartikkeleiden avoimen saatavuuden osalinjaus \(2019\)](#) sekä tähän linjaukseen liittyvät [suositukset](#).

Avoimet oppimateriaalit

Oppimisen ja oppimateriaalien avoimuus. Korkeakoulu- ja tutkimusyhteisön kansallinen linjaus ja toimenpideohjelma. [Osalinjaus 1 – Oppimateriaalien avoin saatavuus \(2020\)](#) sekä tähän linjaukseen liittyvät [suositukset](#).

Datan avaaminen

Datan avaaminen tarkoittaa, että vähintään metatiedot (kuvailutiedot) tarjotaan myös muiden käytettäväksi.

Lähtökohtaisesti

- vain sellaista dataa voi avata, johon on suostumus / sopimus
- avaaminen ei saa olla ristiriidassa lainsäädännön, tietosuojan, tutkimusetiikan tai sopimusten kanssa
- datan säilytystavoista ja mahdollisesta avaamisesta on informoitava tutkittavia
- datan voi avata myös osittain
- datan voi avata embargolla (avataan tietyn ajan jälkeen).

Rahoittaja voi vaatia rahoitusehdoissaan tutkimusdatan avaamista tai saattamista yleisesti hyödynnettäväksi, mutta mikäli datan avaaminen ei ole mahdollista, se hyväksyy pätevät perustelut.

Osa tutkimuslehdistä on ryhtynyt vaatimaan tutkimusartikkelin pohjana olevan datan tallentamista artikkelin yhteyteen.

Ammattikorkeakoululle kriittistä tai arvokasta dataa ei avata arvioimatta ensin mahdollisen avoimuuden tuottamaa lisäarvoa.

Datan avaaminen jatkokäyttöön

- tuo tekijälle näkyvyyttä ja on osa meritoitumista
- lisää tutkimuksen läpinäkyvyyttä ja luotettavuutta
- voi auttaa rahoituksen saamisessa
- nopeuttaa tutkimusta ja vähentää päällekkäistä työtä
- mahdollistaa uusia käyttötapoja
- lisää kaupallisia innovaatioita
- edistää yhdenvertaisuutta.

Tutkimusaineistojen jatkokäytön edistämiseksi datan tulisi olla nk. [FAIR-periaatteiden mukaista](#). FAIR-periaatteita noudattava data on löydettävissä

(Findable), saavutettavissa (Accessible), yhteentoimiva (Interoperable) ja uudelleenkäytettävissä (Re-usable).



Kuva: [Jessica Parland-von Essen & CSC](#), CCO

Tunnisteellinen aineisto

Henkilötietoja tulee kerätä vain siinä määrin kuin on välttämätöntä tutkimuksen toteuttamiseksi.

EU:n tietosuoja-asetuksen määritelmän mukaan henkilötiedoilla tarkoitetaan kaikkia tunnistettuun tai tunnistettavissa olevaan luonnolliseen henkilöön liittyviä tietoja.

Tunnistettavissa olevana pidetään luonnollista henkilöä, joka voidaan suoraan tai epäsuorasti tunnistaa erityisesti tunnistetietojen, kuten nimen, henkilötunnuksen, sijaintitiedon, verkkotunnistetietojen tai hänelle tunnusomaisen esimerkiksi fyysisen, taloudellisen tai kulttuurillisen tekijän perusteella. Tutkimusaineistoihin voi myös sisältyä tunnistetietoja tutkittavien lähipiiristä tai muista kolmansista henkilöistä. Myös heitä tunnistettavasti käsittelevät tiedot ovat aina henkilötietoja.

Tunnistetietoja ovat:

- Suorat tunnisteet: koko nimi, henkilötunnus, henkilönimen mukainen sähköpostiosoite ja biometriset tunnisteet (sormenjälki, kasvokuva, ääni,

silmän iiris, kämmenen muoto, käsin tehty allekirjoitus).

- Vahvat epäsuorat tunnisteet: esimerkiksi postiosoite, puhelinnumero, auton rekisteri, harvinainen ammattinimike, hyvin harvinainen sairaus tai erilaiset yksilöivät koodit kuten opiskelijatunnus.
- Epäsuorat tunnisteet: tiedot, jotka yksin eivät riitä tunnistamiseen, mutta yhdistettynä voivat mahdollistaa henkilön tunnistamisen. Tavallisimpia epäsuoria tunnisteita ovat esimerkiksi sukupuoli, ikä, koulutus, ammattiasema, etninen tausta, työpaikka tai koulu.

Tunnisteellisia aineistoja voi käyttää tieteelliseen tutkimukseen silloin, kun se on tarkoituksenmukaista, suunniteltua, asiallisesti perusteltua ja tietojen käsittelyyn on laillinen käsittelyperuste (esimerkiksi tutkittavan suostumus tai yleisen edun mukainen tutkimus).

Henkilötietojen käsittelyn yleisiä suojatoimia ovat pseudonymisointi, anonymisointi ja säilytyksen rajoittaminen.

Pseudonymisointi

Pseudonymisointi on aineiston tunnisteellisten tietojen poistamista tai korvaamista peitetiedolla tai koodeilla, jotka prosessin jälkeen säilytetään erillään aineistosta organisatorisesti ja teknisesti. Organisatorisilla toimenpiteillä tarkoitetaan tietojen suojattua fyysistä käyttöympäristöä ja hallinnollisesti rajattua ja valvottua käyttöoikeutta. Teknisillä toimenpiteillä viitataan tietoturvallisiin tallennusratkaisuihin. Pseudonymistamista aineistosta tulee anonyymi, kun erillään säilytettävät tunnistetiedot (koodiavain, henkilötiedot ja tiedot muutettujen arvojen muodostamistavoista) hävitetään.

Anonymisointi

Täysin anonyymiä tietoa ei ole olemassa. Anonymisoinnilla voidaan kuitenkin päästä sellaiseen tulokseen, jossa yksittäisiä henkilöitä ei voi annettujen tietojen perusteella tai tietoja muihin tietoihin yhdistämällä tunnistaa. Aineisto on siis anonyymi, jos sitä ei voi kohtuullisin keinoin enää yhdistää alkuperäisiin henkilötietoihin.

Tutkimusaineiston anonymisointiin ei ole olemassa valmista kaikkiiin aineistoihin soveltuvaa menettelytapaa. Anonymisointi tulee suunnitella aina

aineistokohtaisesti ottaen huomioon aineiston ominaisuudet, käyttöympäristö ja käytettävyys.

Anonymisointiprosessin hahmottamiseksi niin kvantitatiivisissa kuin kvalitatiivisissa aineistoissa voi käyttää apuna seuraavia kysymyksiä:

- Mitä suoria tai epäsuoria tunnisteita aineisto sisältää?
- Sisältääkö aineisto ainutlaatuisia tai harvinaisia havaintoja?
- Mitä aineiston tietoja yhdistelemällä henkilö saattaa henkilö voi olla tunnistettavissa?
- Onko saatavilla ulkopuolisia tietoja, jotka voidaan yhdistää aineistoon niin, että havainnot/tutkittavat voivat olla tunnistettavissa?
- Mihin aineistoa tullaan käyttämään ja mitkä ovat juuri ne aineiston ominaisuudet, jotka halutaan säilyttää ja mitkä voidaan ”uhrata” anonymisointiprosessissa.

Säilytyksen rajoittaminen

Tutkimuksen toteuttamiselle tarpeettomat henkilötiedot poistetaan heti, kun se on mahdollista. Esimerkiksi aineiston keruuvaiheessa tarvittut nimitiedot ja osoitteet hävitetään heti, kun ne eivät ole enää välttämättömiä tutkimuksessa.

[Tietoarkisto: Aineistohallinnan käsikirja, Tunnisteellisuus ja anonymisointi](#)

Dataan viitataan samojen periaatteiden mukaisesti kuin julkaisuihinkin

Dataviitteen tärkeimmät osat ovat:

- Tekijä(t)
- Nimi – tutkimusaineiston nimi
- Päiväys – vuosi, jona tutkimusaineisto lisättiin arkistoon
- Julkaisija / Jakaja
- Version numero – käytetään tapauksissa, joissa on saatavilla useita versioita
- Linkki – suositellaan pysyvää digitaalista tunnistetta, esimerkiksi DOI tai URN

Perusmalli:

Tekijän sukunimi, Etunimen alkukirjain tai alkukirjaimet. [Julkaisuaika].

Tutkimusaineiston nimi. [sähköinen tietoaineisto]. Julkaisija/jakaja. URN/URL

Esimerkiksi:

Rainó, P., Alanne, K., Tupi, E., Tammenmaa, C., Widberg, S., Wenman, L., Palo, M., Manunen, J., Nisula, M. (2020). Viittomakielibarometri 2020. [sähköinen tietoaaineisto]. Tietoarkisto. <https://urn.fi/urn:nbn:fi:fsd:T-FSD3684>