

Tampere University Pressin data- ja aineistopolitiikka

Versio 1

Julkaistu 27.1.2021

1 Johdanto

Tutkimuksessa ja tutkimusjulkaisuissa käytetyn datan ja muun aineiston vastuullinen avoin jakaminen edistää tutkijayhteisön yhteistyötä, tehostaa tulosten arviointia, helpottaa tutkimusten toistamista sekä kannustaa aineiston laajempaan hyödyntämiseen. Avoimuuden myötä tutkimus saa enemmän huomiota ja viittauksia. Lisäksi tutkimusprosessin laatu ja luotettavuus paranevat.

Jos aineistoa ei voi esimerkiksi sen sensitiivisyyden takia avata, tulee aineiston keskeisten metatietojen löydettävyydestä silti huolehtia. Hyvä datanhallinta on edellytys datan avaamiselle, mutta myös ei-avoimen datan hallinnan on oltava asianmukaista.

Tampere University Press suosittelee aineistohallintasuunnitelman (data management plan, DMP) laadintaa, sen päivittämistä projektin edetessä ja tuoreimman version liittämistä osaksi arviointiin toimitettavaa käsikirjoitusta (ensimmäinen kustantajalle lähetettävä käsikirjoitusversio).

Datanhallintaan saa tukea ja ohjeita omasta organisaatiosta sekä kansallisista ja kansainvälisistä palveluista.

Katso lisää esimerkiksi:

- Aineistohallintasuunnitelmatyökalu DMPTuuli <https://www.dmptuuli.fi/>
- CSC:n datanhallinnan ohjeita <https://research.csc.fi/data-management-planning>
- Tietoarkiston Aineistohallinnan käsikirja <https://www.fsd.tuni.fi/aineistohallinta/fi/>

Datan merkityksestä tutkimustuotoksena:

- Avoin tiede: avoin saatavuus ja aineistojen avaaminen (Suomen Akatemia) <https://www.aka.fi/tutkimusrahoitus/vastuullinen-tiede/avoin-tiede/>
- Tutkijan ansioluettelomalli: Tutkimuseettisen neuvottelukunnan suositus <https://tenk.fi/fi/ohjeet-ja-aineistot/tutkijan-ansioluettelomalli>

2 Tutkimusetiikka ja juridiikka

Datan ja aineiston hallinnassa ja avaamisessa tulee huomioida näihin liittyvät eettiset ja juridiset kysymykset. Tutkijan tulee seurata Tutkimuseettisen neuvottelukunnan ohjeita (<https://www.tenk.fi/fi/eettinen-ennakkoarviointi-ihmistieteissa>) ja ottaa huomioon tietosuojalainsäädännön vaatimukset (<https://tietosuoja.fi/tieteellinen-tutkimus>) sekä dataan ja aineistoon mahdollisesti liittyvät sopimukset.

3 Datan tallentaminen säilytyspalveluun

Tampere University Press suosittelee datan ja aineiston säilyttämistä palveluissa, jotka mahdollistavat [FAIR-periaatteiden](#) (löydettävyyden, saavutettavuuden, yhteentoimivuuden, uudelleenkäytettävyyden) toteutumisen tutkimusalakohtaisten periaatteiden mukaisesti. Hyviä vaihtoehtoja pitkäaikaissäilytykseen ovat esimerkiksi sertifioidut, korkeatasoiset, kotimaiset ei-kaupalliset palveluntarjoajat Tietoarkisto (<https://www.fsd.tuni.fi/fi/>) ja Kielipankki (<https://www.kielipankki.fi/>).

4 Sensitiivisen datan ja aineistojen jakaminen

Kirjoittajien tulee selvittää, millä ehdoin ja miten käsiteltynä sensitiivinen data voisi olla jaettavissa muiden tutkijoiden käyttöön. Niissä tapauksissa, joissa dataa ei voi sensitiivisyytensä vuoksi jakaa edes osittain, tulee keskeisten metatietojen olla julkisesti saatavilla. Sensitiivisenä pidetään dataa, joka sisältää esimerkiksi etnistä alkuperää, geenejä tai terveyttä koskevia tietoja.

Katso lisää:

- Henkilötiedon ja tunnisteellisen tiedon määrittelyä <https://www.fsd.tuni.fi/fi/palvelut/aineistonhallinta/tunnisteellisuus-ja-anonymisointi/>
- Henkilötiedot ja erityiset henkilötietoryhmät <https://www.tuni.fi/fi/tutkimus/vastuullinen-tiede/tutkimuksen-tietosuoja#expander-trigger--030f390c-4859-4586-913e-0310e77b8b3f>
- Tutkimusaineistojen kuvailutyökalu Qvain qvain.fairdata.fi

5 Datan tai aineiston käyttöoikeudet

Kirjoittajien tulee selvittää aineiston, datan tai niiden eri osien immateriaalioikeudellinen status eli selvittää omistusoikeudet ja käyttöehdot kaikkien tekijöiden, heidän instituutioidensa ja rahoittajiensa välillä. Jos dataa saa käyttää uudelleen, tulee kirjoittajien ilmoittaa käytön ehdot. Asian selkiyttämiseksi data on syytä lisensoida, tai jos data on hankittu muualta, tulee viitata sopimuksiin.

Katso lisää:

- Ohje tekijänoikeuksiin liittyvästä avoimesta lisensoinnista tutkijoille ja tieteellisille kustantajille <https://doi.org/10.23847/isbn.9789525995350>
- Tutkimusaineistojen hallinta: Creative Commons -lisenssit https://libguides.tuni.fi/tutkimusaineistojen-hallinta/creative_commons
- Mari Elisa Kuusniemi: *Kuka omistaa tutkimusdatan?* <https://vastuullinentiede.fi/fi/tutkimuksen-suunnittelu/kuka-omistaa-tutkimusdatan>

6 Datan analyysia koskevat tiedot

Datan analyysissa käytetyt ohjelmistot tulee mainita ja käytetyt menetelmät, algoritmit ja lähdekoodi tulee kuvata.

7 Vertaisarviointi, data ja aineistot

Vertaisarvioijalle on tutkimuksen luotettavuuden arvioinnissa paljon hyötyä siitä, että hän pystyy tarkastelemaan siinä käytettyä dataa ja aineistoja. Linkki asianmukaisesti tallennettuihin tutkimusaineistoihin (ks. kohdat 2–4) tulee liittää julkaisuun (ks. kohta 9), jos vähintään datan metatiedot ovat julkisesti saatavilla.

8 Datan avaaminen

Tampere University Press suosittelee, että vähintään datan keskeiset metatiedot (joita ovat esim. otsikko, kuvaus, tunnisteet, toimijat, oikeudet ja lisenssit) ovat julkaisuhetkellä avoimesti saatavilla. Kuvailutietojen tuottamiseen ja julkaisemiseen on suositeltavaa käyttää Qvain-palvelua (<https://www.fairdata.fi/qvain/>) ja Etsin-palvelua (<https://www.fairdata.fi/etsin/>).

9 Aineistojen ja datan alkuperäiset julkaisijat ja dataan viittaaminen

Dataan tulee suhtautua samoin kuin muihin tutkimusjulkaisuihin. Julkaisun kirjoittajien tulee nimetä tutkimuksessa käytetyt aineistot ja niiden alkuperäiset julkaisijat. Viittaaminen dataan tulee tehdä pysyviä tunnisteita (URN- tai DOI-tunniste) käyttäen, millä varmistetaan linkkien pysyvyys ja uudelleenohjautuvuus.

- Kansallinen ohjeistus *Tracing data : Data citation roadmap for Finland* <http://urn.fi/URN:NBN:fi-fe201804106446>
- Tietoarkiston ohje: Aineistoon viittaaminen <https://www.fsd.tuni.fi/fi/aineistot/aineistojen-kaytto-ja-lataaminen/aineistoon-viittaaminen/>

10 Laajemmasta datavarannosta valikoidut osat

Mikäli julkaistu tutkimus perustuu suuremman datavarannon osaan, suositellaan kyseisen erillisen osan saattamista avoimeksi toistettavuuden ja uudelleenkäytön mahdollistamiseksi.

Tämän ohjeistuksen laadinnassa on käytetty apuna Tiedonjulkistamisen neuvottelukunnan ja Tieteellisten seurain valtuuskunnan julkaisemaa Ohjetta tieteellisille julkaisukanaville vastuullisen aineisto- ja datapolitiikan laatimiseksi <https://doi.org/10.23847/isbn.9789525995220>

The Data and Materials Policy of Tampere University Press

Version 1

Published on 27 January 2021

1 Introduction

The responsible sharing of data and other materials, which are collected or created while undertaking research or writing a research paper, fosters collaboration in the scientific community, supports the validation of research results and the reproducibility of research, and promotes the wider reuse of data. Open access not only increases citation rate and the visibility of research but also improves the quality and reliability of scientific work.

In cases where data cannot be made openly accessible, for example, due to its sensitive nature, it is nevertheless important to publish key metadata. While good data management is the prerequisite for publishing data, data that cannot be made openly accessible must also be appropriately managed.

Tampere University Press recommends that researchers prepare a data management plan (DMP), keep the plan up to date throughout their book project, and deliver the latest version to the publisher alongside the manuscript submitted for peer review (first manuscript version).

With questions about data management, please contact your own organisation or national or international service providers.

Read more:

- Data management planning tool DMPTuuli <https://www.dmptuuli.fi/>
- CSC's instructions for planning data management <https://research.csc.fi/data-management-planning>
- Data management guidelines issued by the Finnish Social Science Data Archive <https://www.fsd.tuni.fi/aineistonhallinta/fi/>

Read more about data as research output:

- Open science: open access publishing and open data (Academy of Finland) <https://www.aka.fi/en/research-funding/responsible-science/open-science/>
- CV template for researchers: recommended template issued by the Finnish National Board on Research Integrity TENK <https://tenk.fi/en/advice-and-materials/template-researchers-curriculum-vitae>

2 Research ethics and applicable legislation

Data may be subject to legal and ethical obligations that impose restrictions on its management and publication. All researchers must adhere to the guidelines issued by the Finnish National Board on Research Integrity TENK (<https://tenk.fi/en/ethical-review/ethical-review-human-sciences>), the requirements set out in data protection laws (<https://tietosuoja.fi/en/scientific-research-and-data-protection>) and the terms and conditions of possible agreements relating to data.

3 Storing data in a repository service

Tampere University Press recommends storing data in repositories that comply with the [FAIR principles](#) (findable, accessible, interoperable, reusable) in accordance with discipline-specific guidelines. For the long-term storage of data, Tampere University Press recommends repositories offered by certified, domestic, non-commercial and recognised service providers, such as the Finnish Social Science Data Archive (<https://www.fsd.tuni.fi/en/>) and the Language Bank of Finland (<https://www.kielipankki.fi/language-bank/>).

4 Sharing sensitive data

Authors must look into the conditions that apply to the sharing of sensitive data and determine whether sensitive data can be prepared for sharing through techniques that prevent the identification of research subjects. If highly sensitive data cannot be made even partially open, the key metadata must still be published. Data is considered sensitive if it, for example, reveals a person's ethnic origin or contains genetic or health data.

Read more:

- Definitions of personal data and identifiable data <https://www.fsd.tuni.fi/en/services/data-management-guidelines/anonymisation-and-identifiers/>
- Personal data and special category data <https://www.tuni.fi/en/research/responsible-research/data-protection>
- The Qvain research dataset description tool qvain.fairdata.fi

5 Right to use data

The authors of research papers must look into the intellectual property rights relating to materials, data or parts thereof to determine not only their ownership but also the terms governing their use that apply to all authors, their affiliated institutions and funders. If data may be reused, the authors must describe the terms of use. For the avoidance of doubt, it is best to license your data or, if data is acquired elsewhere, to include a reference to the related agreements.

Read more:

- Open licensing guidelines for researchers and scientific publishers (in Finnish only) <https://doi.org/10.23847/isbn.9789525995350>
- Research data management: Creative Commons licenses <https://libguides.tuni.fi/researchdatamanagement/cc>
- Mari Elisa Kuusniemi: *Who owns research data?* (in Finnish only) <https://vastuullinentiede.fi/fi/tutkimuksen-suunnittelu/kuka-omistaa-tutkimusdatan>

6 Information about data analysis

Researchers must identify their data analysis software and the methods, algorithms, and source code used in their research.

7 Peer review, data and materials

Having access to data and research materials helps peer reviewers in the evaluation of the credibility of research. Researchers must include a link to appropriately stored research data (see sections 2–4 herein) in their publications (see section 9) if at least the metadata are openly available.

8 Making data accessible

Tampere University Press recommends making at least the key metadata (such as title, description, unique identifier, authors, rights and access information, licenses) openly accessible at the time of publication.

Tampere University Press recommends using the Qvain research dataset description tool (<https://www.fairdata.fi/en/qvain/>) and the Etsin research dataset finder (<https://www.fairdata.fi/en/etsin/>) to create and publish metadata.

9 Authorship and citing data

As any research output, data must be appropriately cited. When you are writing a research paper, you must cite the data associated with your paper and acknowledge the authors who generated the data. Cite data using persistent identifiers (URN or DOI) that provide a permanent link to the resources regardless of where they are stored.

- National guidelines: *Tracing data: Data citation roadmap for Finland* <http://urn.fi/URN:NBN:fi-fe201804106446>
- Guidelines issued by the Finnish Social Science Data Archive: Citing archival data <https://www.fsd.tuni.fi/en/data/downloading-and-using-data/citing-data//>

10 Segments of large dataset

If published research is based on a segment of a large dataset, Tampere University Press recommends making the separate segment openly accessible to enable reproducibility and reuse.

This policy is based on the document titled Guidelines for preparing a responsible data and materials policy for a scientific publication channel (only available in Finnish) published by the Committee for Public Information and the Federation of Finnish Learned Societies <https://doi.org/10.23847/isbn.9789525995220>