



TUTKIMUSAINEISTOJEN JA -MENETELMIEN AVOIMUUS. KORKEAKOULU- JA TUTKIMUSYHTEISÖN KANSALLINEN LINJAUS JA TOIMENPIDEOHJELMA 2021–2025.

OSALINJAUS 1: TUTKIMUSDATAN AVOIN SAATAVUUS

VASTUULLISEN TIETEEN JULKAISUSARJA

Vastuullisen tieteen julkaisusarjassa julkaistaan julistuksia, linjauksia, selvityksiä, suosituksia ja muita asiakirjoja, jotka liittyvät tieteen ja tutkimuksen avoimuuteen, vastuullisuuteen, saavutettavuuteen, tiedeviestintään sekä tieteen ja yhteiskunnan vuorovaikutukseen. Julkaisusarja ei ole tieteellinen vertaisarvioitu julkaisu. Julkaisusarjaa julkaisevat Tiedonjulkistamisen neuvottelukunta ja Tieteellisten seurain valtuuskunta.

Tutkimusaineistojen ja -menetelmien avoimuus. Korkeakoulu- ja tutkimusyhteisön kansallinen linjaus ja toimenpideohjelma 2021–2025. Osalinjaus 1: Tutkimusdatan avoin saatavuus

Vastuullisen tieteen julkaisusarja 5:2021
3. vuosikerta

Tekijä: Avoimen tieteen koordinaatio, Tieteellisten seurain valtuuskunta.

Toimituskunta: Susanna Nykyri, Seliina Päällysaho, Tomi Rosti (pj.), Anne Sunikka, Anssi Neuvonen, Mari Elisa Kuusniemi

Julkaisija: Tiedonjulkistamisen neuvottelukunta ja Tieteellisten seurain valtuuskunta

Helsinki, 2021

ISSN 2670-062X

ISBN 978-952-5995-46-6

DOI <https://doi.org/10.23847/isbn.9789525995466>

Tämä teos on lisensoitu Creative Commons Nimeä 4.0 Kansainvälinen -lisenssillä. Tarkastele lisenssiä osoitteessa <https://creativecommons.org/licenses/by/4.0/deed.fi>.



TUTKIMUSAINEISTOJEN JA -MENETELMIEN AVOIMUUS. KORKEAKOULU- JA TUTKIMUSYHTEISÖN KANSALLINEN LINJAUS JA TOIMENPIDEOHJELMA 2021–2025.

OSALINJAUS 1: TUTKIMUSDATAN AVOIN SAATAVUUS



Hyväksytty
Avoimen tieteen ja tutkimuksen kansallisessa
ohjausryhmässä
13.4.2021.

SISÄLLYS

Linjauksen päämäärä	5
Johdanto	5
Strategiset periaatteet	11
OSALINJAUS: TUTKIMUSDATAN AVOIN SAATAVUUS	13
Kirjallisuutta	17
Liite 1: Sanasto	19
Liite 2: Toimenpiteiden 2021–2025 aikajana	22
Liite 3: Linjaustyöryhmän jäsenet	24

LINJAUKSEN PÄÄMÄÄRÄ

Tutkimusaineistot ja -menetelmät ovat niin avoimia kuin mahdollista ja niin suljettuja kuin välttämätöntä. Aineistoja hallitaan tarkoituksenmukaisesti FAIR-periaatteiden toteuttamista tavoitellen.¹ Tutkimusmenetelmät ja -aineistot, mukaan lukien tutkimusdata, tunnistetaan itsenäisiksi tutkimustuotoksiksi.

JOHDANTO

Tämä linjaus koostuu osalinjauksista, joista ensimmäisenä valmistuu tutkimusdataa koskeva osalinjaus keväällä 2021. Tutkimusmenetelmiä koskevan osalinjauksen on arvioitu valmistuvan vuonna 2022. Tässä osalinjauksessa on tarkoitus käsitellä koodeja ja ohjelmistoja.

Yleisellä tasolla tutkimusaineistojen ja -menetelmien avoimuus ja hyvä hallinta parantavat mahdollisuuksia tutkimusten todentamiseen sekä edistävät olemassa olevien aineistojen ja menetelmien tarkoituksenmukaista uudelleenkäyttöä². Avoimuus lisää globaalia yhdenvertaisuutta, koska tutkimusaineistoja tarjotaan maailmanlaajuisesti tutkijoiden käyttöön. Tämä puolestaan edesauttaa uusien innovaatioiden syntymistä. Aineistojen ja menetelmien avoimuuteen liittyy merkittäviä tutkimuseettisiä ja lainsäädännöllisiä kysymyksiä sekä rajoitteita, jotka tutkijoiden ja muiden aineistohallinnan parissa työskentelevien tulee tuntea.

Jotta yhdenvertaisuus toteutuisi sekä kansallisella että kansainvälisellä tasolla, tutkimusaineistojen ja -menetelmien avoimuutta tulee edistää taloudellisesti kestäväällä tavalla. Taloudellisesti kestävä avoimuuden edistäminen merkitsee sitä, että keskeiset avoimen tieteen infra-

¹ FAIR-periaatteet viittaavat tutkimusdatan laatuun sen jatkokäytön näkökulmasta. Periaatteiden tavoitteena on tehdä tutkimusaineistosta löydettävää (Findable), saavutettavaa (Accessible), yhteentoimivaa (Interoperable) uudelleenkäytettävää (Re-usable). Ks. Liite 1: Sanasto sekä <https://www.force11.org/fairprinciples>.

² Vastuullisen tutkimus- ja innovaatiotoiminnan tavoitteena on saada toimijat tuottamaan eettisesti hyväksyttäviä, kestäviä ja yhteiskunnallisesti kiinnostavia tutkimus- ja innovaatiotuotoksia, ks. <https://www.rrl-practice.eu/about-rrl-practice/what-is-rrl/>

struktuurit ovat voittoa tavoittelemattomia³. Suomessa julkisilla varoilla tuotettujen tutkimusaineistojen tulee olla kaikkien avoimesti hyödynnettävissä.

TUTKIMUKSEN VAPAAUS JA TUTKIJAN VASTUU

Linjaus perustuu tutkimusaineistojen ja -menetelmien avoimen saatavuuden kansallisiin ja kansainvälisiin linjauksiin. Sen tavoitteena on tukea avoimen tieteen päämäärää ja lisätä tutkimuksen vapautta sekä tutkijan mahdollisuuksia jakaa ja hyödyntää tutkittua tietoa. Tutkijan vastuulla on tehdä laadullisesti parasta mahdollista tutkimusta, johon sisältyy vastuullinen tutkimusaineistojen ja -menetelmien käsittely. Tutkimusaineistolle ja -menetelmille valitaan sellainen repositorio, joka parhaiten sopii kullekin kokonaisuudelle ja tarjoaa FAIR-periaatteiden mukaisen palvelun. Avoimen saatavuuden tarkoituksena on mahdollistaa tutkimusaineistojen ja -menetelmien mahdollisimman vapaa jatkokäyttö. Avoimuuden astetta voidaan rajoittaa perustelluista syistä.

Tutkimusyhteisön tulee taata tutkimusaineistojen ja -menetelmien avaamiselle osaaminen, kannustimet, resurssit ja rakenteet (kuten infrastruktuurit ja palvelut) tavalla, joka tukee ja kunnioittaa tutkijoiden työtä ja tasavertaisuutta.

³ First draft of the UNESCO Recommendation on Open Science, s. 5, <https://unesdoc.unesco.org/ark:/48223/pf0000374409.locale=en>. page=14

KANSAINVÄLINEN VIITEKEHYS

Tämän linjauksen laatimisessa on hyödynnetty sitä laajaa kotimaista ja kansainvälistä työtä, jota tutkimusaineiston avoimuuden edistämiseksi tehdään ja on tehty. Hyviä yhteenvedoja tästä työstä ovat esimerkiksi Euroopan tutkimusyliopistojen liiton laatima suunnitelma *Open Science and its Role in Universities: A Roadmap for Cultural Change* (2018) ja Euroopan komission julkaisema loppuraportti *Turning FAIR into Reality. Final report and action plan from the European Commission expert group on FAIR data* (2018)⁴. Toimintaympäristöä muovaa myös datan avoimuutta koskeva nykyinen ja tuleva EU-sääntely, esimerkiksi avoimen datan direktiivi.

Tiedekustantajat vaativat tutkimusjulkaisujen kirjoittajilta tutkimusaineistojen ja -menetelmien avointa saatavuutta. Tiedekustantajat myös tarjoavat tai edellyttävät kirjoittajilta tutkimusaineistojen tai -menetelmien tallentamista repositorioihin, jotka voivat olla suljettuja, avoimia tai maksumuurien takana^{5,6}.

Rahoittajat edellyttävät tutkimusaineistojen ja -menetelmien avaamista, ja se kytkeytyy läheisesti hyvään aineistohallintaan, josta esimerkkinä Science European ohje *Practical Guide to the International Alignment of Research Data Management*⁷.

Huolehtimalla siitä, että suomalainen linjaus vastaa kansainvälistä kehitystä, kotimainen tutkimusyhteisö on mukana luomassa yhteisiä kansainvälisiä käytäntöjä tällä nopeasti laajenevalla ja kehittyvällä alueella. Suomalaiset

4 Ayris, Paul; López de San Román, Alea; Maes, Katrien; Labastida, Ignasi: *Open Science and its Role in Universities: A Roadmap for Cultural Change*. League of European Research Universities, 2018. <https://www.leru.org/publications/open-science-and-its-role-in-universities-a-roadmap-for-cultural-change>;

Turning FAIR into Reality. Final report and action plan from the European Commission expert group on FAIR data. Tutkimuksen ja innovoinnin pääosasto, Euroopan komissio, 2018. doi.org/10.2777/1524. Ks. myös Lawrence, Rebecca; Mendez, Eva: *Progress on open science: Towards a shared research knowledge system. Final report of the open science policy platform*. Tutkimuksen ja innovoinnin pääosasto, Euroopan komissio, 2020. doi.org/10.2777/00139

5 Springer Nature: "Data Availability Statements" <https://www.springernature.com/gp/authors/research-data-policy/data-availability-statements/12330880>

6 *First draft of the UNESCO Recommendation on Open Science*, <https://unesdoc.unesco.org/ark:/48223/pf0000374409.locale=en.page=10>

7 Science Europe: *Practical Guide to the International Alignment of Research Data Management*, 2018. https://www.scienceeurope.org/media/jezkhnoo/se_rdm_practical_guide_final.pdf

toimijat osallistuvat aktiivisesti eurooppalaiseen ja kansainväliseen työhön, esimerkiksi European Open Science Cloud (EOSC) ja Research Data Alliance (RDA) -yhteisöissä. Suomalainen tutkimusyhteisö osallistuu vuoropuheluun, josta voi omaksua hyviä käytäntöjä ja jossa voi olla mukana luomassa toimivia ratkaisuja.

RISKIT JA UHAT

Linjaukseen sisältyy tunnistettuja riskejä. Tutkimusaineistojen ja -menetelmien avaaminen pohjautuu hyvälle aineistohallinnalle. Hyvän tieteellisen käytännön ja hyvän aineistohallinnan toteuttaminen vaatii kuitenkin resursseja ja tutkimusorganisaatioille suunnattuja taloudellisia kannustimia koko tutkimusaineiston ja -menetelmien elinkaaren aikana. Jotta resursseja olisi käytettävissä koko elinkaaren aikana, erilaisissa rooleissa toimivien organisaatioiden olisi sitouduttava resurssien ylläpitämiseen. Toisaalta palveluihin panostaminen edistää resurssien tehokkaampaa hyödyntämistä ja tuosäästöjä.

Tutkijat kantavat huolta arkaluontoisten ja luottamuksellisten (esimerkiksi henkilötiedot) tutkimusaineistojen ja -menetelmien avaamisesta. Tämän vuoksi tulee ottaa huomioon erilaiset avaamisen asteet ja tutkijan ensikäyttöoikeus. Joissakin tapauksissa riittää kuvailutiedon avaaminen. Tutkimusaineistojen ja -menetelmien avaamisessa nousevat esille niihin liittyvät juridiset kysymykset (esimerkiksi sopimukselliset) sekä tutkimusaineistoihin ja -menetelmiin liittyvä tutkijan oikeusturva.

Tutkimusrahoittajat ja tiedekustantajat edellyttävät yhä enenevässä määrin tutkimusaineistojen ja -menetelmien avaamista. Rahoituksen hakijoiden osalta riskinä on, että tutkimusrahoituksen määrä laskee tai rahoitusta peritään takaisin, jos aineistoja ei voi avata. Puuttuvat palvelut ja infrastruktuurit sekä puutteellinen osaaminen ovat syitä siihen, että tutkimusaineistoja ja -menetelmiä ei avata. Myös lainsäädännölliset, eettiset ja sopimukselliset rajoitteet voivat estää avaamisen.

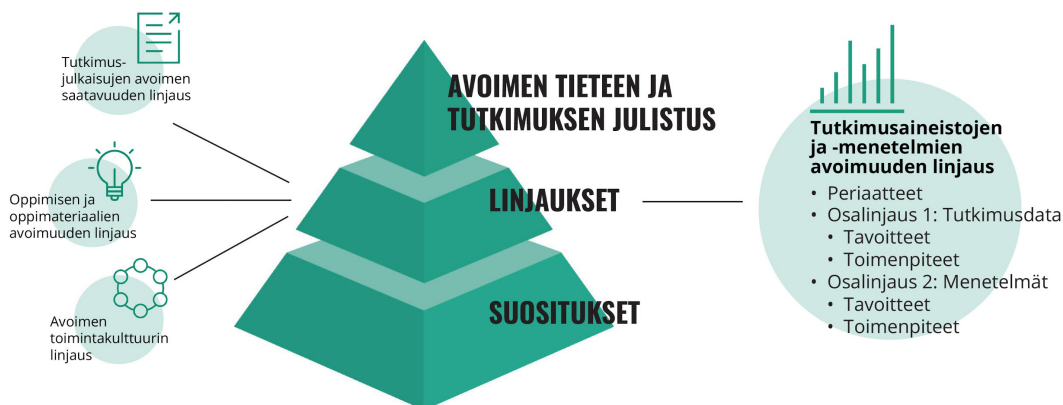
LINJAUKSEN RAKENNE JA TAUSTA

Tämän linjauksen on laatinut suomalainen tutkimusyhteisö. Työn edistymisestä on vastannut Tieteellisten seurain valtuuskunnan Avoimen tieteen koordinaation tutkimusaineistojen avoimuuden asiantuntijaryhmän erikseen kokoama työryhmä. Työtä on ohjannut Avoimen tieteen ja tutkimuksen kansallinen ohjausryhmä. Linjaus tukee *Avoimen tieteen ja tutkimuksen julistusta 2020–2025*.

Tämä tutkimusaineistojen ja -menetelmien avoimen saatavuuden linjaus koostuu koko linjaukselle yhteisistä strategisista periaatteista sekä kahdesta osalinjauksesta, joissa määritellään tavoitteita ja toimenpiteitä kullekin osa-alueelle. Strategisissa periaatteissa määritellään yleiset reunaehdot tutkimusaineistojen ja -menetelmien avoimen saatavuuden tavoittelulle. Ne sanoittavat tutkimusyhteisölle tärkeitä periaatteita, joista on pidettävä kiinni avoimuutta toteutettaessa. Osalinjauksiin kirjatut tavoitteet ovat enemmän aikaan sidottuja, spesifejä ja mitattavia tavoitetasoja avoimuudelle. Tavoitteille listataan myös niiden saavuttamiseen tarvittavia konkreettisia toimenpiteitä. Muuttuva kansainvälinen ympäristö vaikuttaa tavoitteisiin ja niiden toimenpiteisiin nopeammin kuin periaatteisiin.

Linjauksessa käytetty sanasto (ks. Liite 1) tukee dokumentin luettavuutta ja riittävää ymmärrettävyyttä. Aihepiirin sanasto on erilaista eri tieteenaloilla, ja myös lainsäädännössä termit ja niiden käännökset vaihtelevat paljon. Käytetyt termit määritellään eri yhteyksissä toisistaan poikkeavilla tavoilla. Tieteenalojen erilaisuudesta johtuen yhtenäistä sanastoa on haastavaa luoda.

Kuva 1. Linjauksen suhde muihin avoimen tieteen dokumentteihin.



TOIMEENPANO JA SEURANTA

Tutkimusaineistojen ja -menetelmien avoimen saatavuuden linjauksen toimeenpanosta vastaa koko suomalainen tutkimusyhteisö, ja linjaus astuu voimaan asteittain. Tutkimusdatan avoimen saatavuuden osalinjaus valmistuu keväällä 2021, ja tutkimusmenetelmiä koskeva osalinjaus keväällä 2022. Koko linjaus osalinjauksineen tarkistetaan viimeistään vuonna 2025.

Linjauksen toimeenpanon seurannasta vastaa Tieteellisten seurain valtuuskunnan Avoimen tieteen koordinaatio. Koordinaation vastuulla on tukea ja edistää jatkuvaa keskustelua tavoitteiden saavuttamiseksi sekä varmistaa linjauksen ajantasaisuus. Tutkimusaineistojen ja -menetelmien avoimuuden tai toimeenpanon seurantamekanismit tulevat olemaan osa kehitteillä olevaa kansallista avoimen tieteen seurantaa.

STRATEGISET PERIAATTEET

Periaate 1: Tutkimusaineistoa ja -menetelmiä hallitaan, avataan ja käytetään vastuullisesti ja tarkoituksenmukaisesti.

Periaatteen toteutumisen varmistaminen ja seuranta:

- A) Lähtötilanne: Vastuullinen tutkimusaineistojen ja -menetelmien hallinta on keskeisin edellytys niiden avaamiselle. Tutkimusyhteisö on tunnistanut puutteita toimintatavoissa.
- B) Jatkuva seuranta: Osana avoimen tieteen ja tutkimuksen kansallista seurantaa.

Periaate 2: Tutkijoilla on käytettävissään vastuullisen aineistonhallinnan mahdollistavat infrastruktuurit ja palvelut, ja niitä kehitetään tutkijoiden tarpeet huomioiden taloudellisesti kestävällä tavalla.

Periaatteen toteutumisen varmistaminen ja seuranta:

- A) Lähtötilanteen kartoitus: Vuoden 2021 aikana Avoimen tieteen koordinaatio määrittelee yhteistyössä tutkimusorganisaatioiden ja palveluntuottajien kanssa minimitason aineistonhallinnan infrastruktuureille ja palveluille.
- B) Jatkuva seuranta: Avoimen tieteen koordinaatio laatii viimeistään vuonna 2022 arviointimallin⁸, jota voidaan jatkossa hyödyntää arviointivälineenä organisaatioiden säännöllisessä itsearvioinnissa. Jatkuva seuranta selvitetään osana avoimen tieteen ja tutkimuksen kansallista seurantaa.

8 Rans, J and Whyte, A. (2017). *Using RISE, the Research Infrastructure Self-Evaluation Framework*. Digital Curation Centre, Edinburgh. www.dcc.ac.uk/resources/how-guides

Periaate 3: Tutkijan ansioita hyvän aineistohallinnan edistämisessä, tutkimusaineistoon ja -menetelmiin liittyvää työtä sekä tutkimusaineiston ja -menetelmien tarkoituksenmukaista avaamista arvostetaan ja se voi tukea tutkijan meritoitumista.

Periaatteen toteutumisen varmistaminen ja seuranta:

- A) Lähtötilanne: Viimeistään vuonna 2022 Avoimen tieteen koordinaatio laatii suosituksen hyvistä käytännöistä eli siitä, kuinka hyvän datanhallinnan edistäminen, tutkimusdataan liittyvä työ ja tutkimusdatan avaaminen huomioidaan tutkijan työssä ja kuinka näitä ansioita arvostetaan.
- B) Jatkuva seuranta: Osana avoimen tieteen ja tutkimuksen kansallista seurantaa.

OSALINJAUS: TUTKIMUSDATAN AVOIN SAATAVUUS

Tämä kansallinen osalinjaus on suomalaisen tutkimusyhteisön⁹ yhteinen suunnannäyttävä tutkimusdatojen avoimen saatavuuden edistämiseksi. Osalinjaus ei koske muita tutkimusaineistoja eikä -menetelmiä, vaan linjausrakenteen mukaisesti niitä voidaan käsitellä omissa osalinjauksissaan.

Tutkimusdatan avoimen saatavuuden osalinjaus koskee ensisijaisesti tutkimusdataa, joka on tuotettu tai käytetty osana tutkimus- tai kehittämisprosessia 1.7.2021 tai sen jälkeen ja

- jossa tutkija on suomalaisessa tutkimusorganisaatiossa työskentelevä tai siihen affilioitunut ja/ tai suomalaisen tutkimusrahoittajan rahoituksella työskentelevä tutkija

tai

- jossa tutkimus- tai kehittämisprojekti, jonka yhteydessä tutkimusdata kootaan tai käytetään, sijoittuu suomalaiseen tutkimusorganisaatioon ja/ tai on suomalaisen tutkimusrahoittajan rahoittama.

Toimenpiteissä ei oteta kantaa siihen, miten organisaatiot ratkaisevat asiat, joten tutkimusorganisaatioille on jätetty tehtäväksi suunnitella kokonaisuus omista lähtökohdistaan käsin.

TAVOITTEET JA TOIMENPITEET

Tavoite 1: Viimeistään vuonna 2023 kaikille aloittaville tutkimus- ja kehittämisprojekteille on osana laadunhallintaa laadittu datanhallintasuunnitelmat, joissa on huomioitu eri tieteenalojen tarpeita ja tutkimusdatan elinkaaren vaiheet.

⁹ Tutkimusyhteisön määritelmässä noudatetaan *Avoimen tieteen ja tutkimuksen julistuksen 2020–2025* (2020, 5) määritelmää. <https://doi.org/10.23847/isbn.9789525995237>.

Tavoitteen vaatimat toimenpiteet:

- Viimeistään vuonna 2022 korkeakoulut tarjoavat ohjeet, käytänteet ja koulutusta datanhallinnan suunnitteluun opiskelijoille, tutkijoille ja muulle henkilöstölle.
- Viimeistään vuonna 2023 tutkimusorganisaatiot sisällyttävät datanhallintasuunnitelman laadinnan ja ylläpidon osaksi tutkimus- ja palveluprosesseja aineiston elinkaaren jokaisessa vaiheessa.
- Viimeistään vuonna 2024 korkeakoulut huolehtivat, että opinnäytetöiden ohjaajat osaavat osana ohjaustyötään arvioida ja kommentoida datanhallintasuunnitelmia.
- Viimeistään vuonna 2022 tutkimusorganisaatiot tukevat tutkimuksen suunnittelua siten, että datanhallinnan ja sen tuen riittävään resursointiin ja kustannuksiin osataan varautua.
- Korkeakoulut sisällyttävät datanhallinnan koulutuksen opetussuunnitelmiin perus- ja jatkokoulutuksessa opetussuunnitelmien päivittämisen yhteydessä.
- Viimeistään vuonna 2024 tutkimusorganisaatiot ovat kehittäneet seurannan hyvän ja vastuullisen datanhallintakäytännön varmistamiseksi.
- Viimeistään vuonna 2022 tutkimusorganisaatioilla on olemassa datapolitiikka, jota päivitetään säännöllisesti.

Tavoite 2: Viimeistään vuonna 2022 tutkimusorganisaatioissa on käytössä toimintamallit, joilla tutkimusdatan käyttöön liittyvät oikeudet, ehdot ja lisenssit voidaan sopia selkeästi kaikkien tutkimustyöhön osallistuvien kesken.

Tavoitteen vaatimat toimenpiteet:

- Avoimen tieteen koordinaatio tuottaa vuoden 2021 aikana selvityksen avoimen tieteen juridisista kysymyksistä. Yhteistyössä suomalaisen tutkimusyhteisön kanssa luodaan suosituksia sopimusten sisällöiksi.
- Vuonna 2021 tutkimusorganisaatioilla on perusperiaatteet tutkimusdatojen oikeuksien ja vastuiden sopimisesta. Erikoistapauksiin on tarjolla neuvontaa ja opastusta.
- Vuonna 2021 tutkimusrahoittajat ja -organisaatiot ohjeistavat tutkijoita, että joko avattavat tutkimusda-

tat tulee lisensoida tarkoituksenmukaisella lisenssillä tai tutkimustuotosten yhteiskunnallinen vaikuttavuus tulee saavuttaa kaupallistamisella.

- Vuonna 2021 tutkimusrahoittajat ja -organisaatiot ohjeistavat lisensoinnissa tutkijoita julkisin varoin tuotettujen tutkimusaineistojen uudelleenkäytöstä annetun lain vaatimusten mukaisesti.
- Tutkimusorganisaatioissa tarjotaan vuonna 2021 opetusta ja/tai koulutusta ja neuvontaa tutkimusdataan kohdistuvista oikeuksista ja avoimen tieteen lisensseistä.

Tavoite 3: Viimeistään vuonna 2022 aloittavissa tutkimus- ja kehittämisprojekteissa tuotetut tutkimusdatat on dokumentoitu siten, että dokumentaatio tukee tutkimusdatan avaamista, jatkokäyttöä, löydettävyyttä, yhteentoimivuutta ja saatavuutta.

Tavoitteen vaatimat toimenpiteet:

- Viimeistään vuonna 2022 tutkimusorganisaatiot tukevat tutkijaa metadatan tuottamisessa tukipalveluiden ja kannustimien avulla.
- Viimeistään vuonna 2022 tutkimusorganisaatiot tarjoavat eri kohderyhmille suunnattua koulutusta ja tukea sekä työkaluja tutkimusdatan dokumentointiin huomioiden kohderyhmien osaamisen kehittämisen, eri tieteenalojen tarpeet ja tutkimuksen elinkaaren.
- Vuonna 2021 tutkimusorganisaatiot tukevat tutkimuksen suunnittelua siten, että tutkimushankkeissa osataan varautua aineistojen dokumentoinnin kustannuksiin.
- Viimeistään vuonna 2022 avoimen tieteen ja tutkimuksen seurantaryhmä määrittelee avoimen tieteen indikaattorit.
- Viimeistään vuonna 2023 tutkimusorganisaatiot hyödyntävät soveltuvia indikaattoreita, joiden avulla ne seuraavat edistymistään hyvässä datanhallinnassa osana laatujärjestelmää.

Tavoite 4: Viimeistään vuonna 2022 tutkimusdatalle on mahdollistettu hyvän datanhallinnan kannalta asianmukainen, tutkimusdatan elinkaaren huomioiva, tallennus- tai säilytysratkaisu, infrastruktuuri ja palvelut.

Tavoitteen vaatimat toimenpiteet:

- Vuonna 2021 suomalaiset tutkimusorganisaatiot tarjoavat olemassa olevat palvelut osana kansallista arkkitehtuurityötä. Organisaatiot käynnistävät omalta osaltaan tarvittavat kehittämistoimenpiteet sekä edistävät kansallisten ja kansainvälisten palvelujen hyödyntämistä.
- Viimeistään vuonna 2022 tutkimusorganisaatiot tuottavat ja tarjoavat yhteistyössä kansallisten ja kansainvälisten toimijoiden kanssa tutkijoille ja tutkimusryhmille ohjeistuksen sekä riittävät tukipalvelut hyvän datanhallinnan periaatteiden mukaiseen tiedon tallennukseen ja julkaisemiseen.
- Tutkimusyhteisössä huolehditaan, että keskeiset avoimen tutkimusdatan infrastruktuurit ovat voittoa tuottamattomia¹⁰.

Tavoite 5: Viimeistään vuonna 2022 tutkimusorganisaatioissa on käytössä toimintamalli, jolla asiantuntijuutta ja moniammatillista yhteistyötä hyödynnetään koulutusten, osaamisen ja tarvittavien datanhallinnan palveluiden kehittämiseksi.

Toimintamallia ei voida rajoittaa vain avattavaan tutkimusdataan, vaan hyvä datanhallinnan kokonaisuus edellyttää koko elinkaaren hallintaa lopputuloksesta riippumatta.

Tavoitteen vaatimat toimenpiteet:

- Tutkimusyhteisö kehittää osaamisen ja palveluiden rakenteita sekä sisältöjä kansallisessa ja kansainvälisessä yhteistyössä.
- Tutkimusorganisaatiot tarjoavat itsenäisesti tai yhteistyössä yhteisönsä tarpeita vastaavasti hyvän datanhallinnan koulutusta kaikkien uravaiheiden tutkijoille, opettajille, ohjaajille sekä tukihenkilöstölle vuonna 2021.
- Viimeistään vuonna 2023 tutkimusorganisaatiot kansallisessa yhteistyössä kuvaavat, millaisia asiantuntijarooleja tarvitaan ja luovat näille mahdollisia urapolkuja.
- Korkeakoulut kehittävät asiantuntijakoulutusta osana opiskelijoiden opetustarjontaa.

¹⁰ First draft of the UNESCO Recommendation on Open Science, <https://unesdoc.unesco.org/ark:/48223/pf0000374409.locale=en.page=10>

KIRJALLISUUTTA

KANSAINVÄLISIÄ LÄHTEITÄ

- Ayris, Paul; López de San Román, Alea; Maes, Katrien; Labastida, Ignasi (2018): *Open Science and its Role in Universities: A Roadmap for Cultural Change*. League of European Research Universities (LERU). <https://www.leru.org/publications/open-science-and-its-role-in-universities-a-roadmap-for-cultural-change>.
- Ball, A. (2014). *DCC How-to Guides: How to License Research Data*. Digital Curation Centre, Edinburgh. <http://www.dcc.ac.uk/resources/how-guides>.
- Euroopan parlamentin ja neuvoston direktiivi 2019/1024 avoimesta datasta ja julkisen sektorin hallussa olevien tietojen uudelleenkäytöstä. <https://eur-lex.europa.eu/legal-content/FI/TXT/?uri=celex%3A32019L1024>.
- FORCE11 Group: *Guiding Principles for Findable, Accessible, Interoperable and Re-usable Data Publishing*. <https://www.force11.org/fairprinciples>.
- Rans, J and Whyte, A. (2017). *Using RISE, the Research Infrastructure Self-Evaluation Framework*. Digital Curation Centre, Edinburgh. www.dcc.ac.uk/resources/how-guides.
- RDA FAIR Data Maturity Model Working Group (2020): *FAIR Data Maturity Model: specification and guidelines*. Research Data Alliance. doi.org/10.15497/RDA00045.
- RRI-Practice: "What is RRI?". <https://www.rri-practice.eu/about-rri-practice/what-is-rri/>.
- Science Europe (2018): "Practical Guide to the International Alignment of Research Data Management". https://www.scienceeurope.org/media/jezkhn00/se_rdm_practical_guide_final.pdf.
- Springer Nature: "Data Availability Statements". <https://www.springernature.com/gp/authors/research-data-policy/data-availability-statements/12330880>.
- Tutkimuksen ja innovoinnin pääosasto, Euroopan komissio (2018): *Turning FAIR into Reality. Final report and action plan from the European Commission expert group on FAIR data*. doi.org/10.2777/1524.
- Tutkimuksen ja innovoinnin pääosasto, Euroopan komissio (2018): *Prompting an EOSC in Practice. Final report and recommendations of the Commission 2nd High Level Expert Group on the European Open Science Cloud*. doi.org/10.2777/112658.
- Tutkimuksen ja innovoinnin pääosasto, Euroopan komissio (2016): *Guidelines on Fair Data Management in Horizon 2020*. https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf.
- UK Concordat on Open Research Data (2016): <https://www.ukri.org/files/legacy/documents/concordatonopenresearchdata-pdf/>.
- UNESCO (2020). *First draft of the UNESCO Recommendation on Open Science*. <https://unesdoc.unesco.org/ark:/48223/pf0000374409.locale=en.page=10>.

KOTIMAISIA LÄHTEITÄ

- Aineistohallinnan käsikirja*. Tampere: Yhteiskuntatieteellinen tietoarkisto (ylläpitäjä ja tuottaja). <https://www.fsd.tuni.fi/fi/palvelut/aineistohallinta/>. urn:nbn:fi:fsd:V-201504200001.
- Avoimen tieteen koordinaatio, Tieteellisten seurain valtuuskunta (2019): *Tutkimusjulkaisujen avoin saatavuus – Tutkimusyhteisön kansallinen linjaus ja toimenpideohjelma 2020–2025 (1). Lehti- ja konferenssiartikkeleiden avoimen saatavuuden osalinjaus*. Vastuullisen tieteen julkaisusarja 1:2019. doi.org/10.23847/isbn.9789525995206.
- Avoimen tieteen koordinaatio, Tieteellisten seurain valtuuskunta (2020): *Avoimen tieteen ja tutkimuksen julistus 2020–2025*. Vastuullisen tieteen julkaisusarja 1:2020. doi.org/10.23847/isbn.9789525995237.
- Avoin tiede ja tutkimus -hankkeen (2014–2017) keskeiset linjaukset ja julkaisut: <https://avointiede.fi/fi/linjaukset-ja-aineistot/kotimaiset-linjaukset/att-hankkeen-keskeiset-linjaukset-ja-julkaisut>.
- Finnish Committee for Research Data (2018): *Tracing Data: Data citation roadmap for Finland*. <http://urn.fi/URN:NBN:fi-fe201804106446>.
- Fuchs, S. & Kuusniemi, M. E. (2018): *Making a research project understandable – Guide for data documentation*. doi.org/10.5281/zenodo.1914401
- Parland-von Essen, J., Fält, K., Maalick, Z., Alonen, M., & Gonzalez, E. (2018). *Supporting FAIR data: Categorization of research data as a tool in data management*. Informaatiotutkimus, 37(4). doi.org/10.23978/inf.77419.
- Tuuli-projekti (2019): *Täydentävä ohje arkaluonteisia ja luottamuksellisia tietoja sisältävän datan hallinnan suunnitteluun*. [Zenodo] doi.org/10.5281/zenodo.3247282.
- UNIFI (2018): *Avoimen tieteen ja datan toimenpideohjelma*. https://www.unifi.fi/wp-content/uploads/2019/04/UNIFI_Avoimen_tieteen_ja_datan_toimenpideohjelma.pdf.

LIITE 1: SANASTO

Tutkimusaineisto (*research material*) on tutkijan tai tutkimusryhmän tutkimusprosessin aikana käyttämä resurssi, eli digitaalisessa, analogisessa tai fyysisessä muodossa olevaa tieteellisen ja taiteellisen tutkimuksen perusaineistoa. Tutkimusaineisto on laajempi käsite kuin tutkimusdata, kattaen esimerkiksi lähdekirjallisuuden (esimerkiksi asiakirja-aineisto) ja näytteet (esimerkiksi verinäytteet, sammalet).

Tutkimusdata (*research data*) on tutkimusaineistoa, joka on kerätty, havaittu, mitattu tai luotu hypoteesien vahvistamiseksi ja tutkimustulosten todentamiseksi. Tutkimusdata voi yleensä olla digitaalisessa muodossa, mutta se voi esiintyä myös analogisessa tai fyysisessä muodossa (esimerkiksi laboratoriopäiväkirjat).

Metadata: *Metatiedolla, metatiedolla tai kuvailutiedolla* tarkoitetaan tutkimusaineiston kontekstia, sisältöä ja rakennetta, hallintaa ja/tai käsittelyä sekä sen kokoamista kuvaavaa ja tiivistettyä tietoa.¹¹

Tutkimusdatan avoimuudella tarkoitetaan tässä linjauksessa sitä, että tutkimusdata on toisten tutkijoiden löydettävissä, saatavissa, käytettävissä ja jaettavissa. Tutkimusdatan avoin saatavuus edellyttää hyvää ja vastuullista datan hallintaa. Tutkimusdatan uudelleenkäyttöä tuetaan lisensseillä¹² tai sen käyttö voi edellyttää asianmukaista tutkimuslupaa.

Henkilötietoja sisältävän, sensitiivisen tai luottamuksellisen tutkimusdatan vastuullinen käsittely edellyttää tutkijalta niin lainsäädännön kuin hyvien tutkimuseettisten käytäntöjen noudattamista. Tällaisissa tapauksissa on usein kuitenkin mahdollista avata tutkimusdatan kuvailutiedot eli metadata sekä tarjota muille tutkijoille pääsy aineistoon erillisellä aineiston luovutus-sopimuksella.

Tutkimusdatan hyvä ja vastuullinen hallinta: Tässä linjauksessa tutkimusdatan *hyvällä hallinnalla* tarkoitetaan sitä, että tutkimusdata ja siihen liittyvä kuvaileva tieto (metatieto, metadata) on "luotu, tallennettu ja järjestetty siten, että tutkimusdata säilyy käyttökelpoisena ja luotettavana ja että tietosuoja ja tietoturva on varmistettu tutkimusdatan koko elinkaaren ajan".¹³ Lisäksi nimenomaan avoimen tieteen mukainen tutkimusdatan hyvä hallinta tarkoittaa sitä, että tutkimusprosessin kaikissa vaiheissa huomioidaan "niin avointa kuin mahdollista, niin suljettua kuin välttämätöntä" -periaate, FAIR-periaatteet sekä vastuullisuuden periaate. **Tutkimusdatan hyvä hallinta on sen avoimen saatavuuden välttämätön edellytys.** Tutkimusdatan hyvää hallintaa arvioidaan seuraavien näkökulmien pohjalta:

11 https://tieteentermipankki.fi/wiki/Nimitys:tutkimusaineiston_metadata

12 Ball, A. (2014). "How to License Research Data". DCC How-to Guides. Edinburgh: Digital Curation Centre. <http://www.dcc.ac.uk/resources/how-guides>

13 Aineistohallinnan käsikirja. Tampere: Yhteiskuntatieteellinen tietoaarkisto. <https://www.fsd.tuni.fi/fi/palvelut/aineistohallinta/>. urn:nbn:fi:fsd-V-201504200001 (Haettu 11.1.2021)

1. *Niin avointa kuin mahdollista, niin suljettua kuin välttämätöntä* -periaate¹⁴ edellyttää, että kaikki, mikä voidaan avata julkiseksi ja uudelleen käytettäväksi, myös avataan. Vastaavasti sellainen, mitä ei voida julkistaa ja jakaa, suojataan ja säilytetään turvallisesti. Se kumpaan ryhmään tutkimusdata kuuluu, on viime kädessä tutkijan vastuulla. Avaamatta jättäminen edellyttää aina perusteen.
2. *FAIR-periaatteet* (Findable, Accessible, Interoperable, Reusable)¹⁵ viittavat tutkimusdatan laatuun sen jatkokäytön näkökulmasta. Periaatteiden tavoitteena on tehdä tutkimusaineistosta:
 - löydettävää (**F**indable)
 - saavutettavaa (**A**ccessible)
 - yhteentoimivaa (**I**nteroperable)
 - uudelleenkäytettävää (**R**eusable).

FAIR-periaatteiden mukainen tutkimusdata ja metadata ovat semanttisesti yhteentoimivia, eli niin hyvin jäsenneiltyjä, kuvattuja, tunnistein merkittyjä, lisensoituja ja turvallisesti tallennettuja, että ne ovat koneellisesti löydettäviä ja luettavia. Täysin FAIR-periaatteiden mukainen data ei ole vielä useimmilla aloilla mahdollista osaamisen ja palveluiden puutteista tai datan luonteesta johtuen.

3. *Vastuullisella datan hallinnalla* tarkoitetaan tässä linjauksessa:
 - tutkimusalan periaatteiden tuntemusta ja noudattamista
 - tutkimuseettisten periaatteiden tuntemusta ja noudattamista
 - lainsäädännön tuntemusta ja noudattamista
 - tietoturvan ja tietosuojan periaatteiden tuntemusta ja noudattamista.

Aineistojen avoimuuden kaikkien muotojen edellytyksenä on vastuullinen tutkimusaineistojen hallinta. Vastuullinen aineistohallinta edellyttää tutkimusaineistojen, tutkimusdatan ja metatietojen tietoturvallista, tietosuojan mukaista ja tutkimuseettisesti kestävää hallintaa. Tutkimusdatan vastuullinen hallinta ja sitä kautta syntyvä mahdollisuus tutkimusaineiston avaamiseen ovat osa hyvää tieteellistä käytäntöä.

¹⁴ Tutkimuksen ja innovoinnin pääosasto, Euroopan komissio (2016). *Guidelines on Fair Data Management in Horizon 2020*. https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

¹⁵ FORCE11 Group: *Guiding Principles for Findable, Accessible, Interoperable and Re-usable Data Publishing*. <https://www.force11.org/fairprinciples>.

LIITE 2: TOIMENPITEIDEN 2021–2025 AIKAJANA

2021

- **Avoimen tieteen koordinaatio** tuottaa selvityksen avoimen tieteen juridisista kysymyksistä ja sen jälkeen **tutkimusyhteisö** luo suositukset sopimusten sisällöiksi, jolloin **tutkimusorganisaatiolla** on perusperiaatteet tutkimusdatojen oikeuksien ja vastuiden sopimisesta (lisensointi, kaupallistaminen ja oikeuksiensiiirto).
- **Tutkimusorganisaatiot** tarjoavat itsenäisesti tai yhteistyössä yhteisönsä tarpeita vastaavasti hyvän datanhallinnan koulutusta sekä neuvontaa ja opastusta.
- **Tutkimusorganisaatiot** tukevat tutkimuksen suunnittelua siten, että tutkimushankkeissa osataan varautua aineistojen dokumentointiin kustannuksiin.
- **Tutkimusyhteisö** käynnistää kansallisten palveluiden ja tutkimusorganisaatioiden palveluiden kartoituksen osana kansallista arkkitehtuurityötä.

2022

- **Tutkimusorganisaatioilla** on olemassa datapolitiikka, jota päivitetään säännöllisesti.
- **Tutkimusorganisaatiot** tarjoavat ohjeet, käytänteet ja koulutusta datanhallinnan suunnitteluun yhteisönsä tarpeita vastaavasti hyödyntäen moniammatillista yhteistyötä.
- **Tutkimusorganisaatiot** tukevat tutkijaa metadatan tuottamisessa tukipalveluiden ja kannustimien avulla, huomioiden tutkimusdatan avaamisen, jatkokäytön, löydettävyyden, yhteentoimivuuden ja saatavuuden.
- Avoimen tieteen koordinaation alla toimiva **avoimen tieteen ja tutkimuksen seurantaryhmä** määrittelee avoimen tieteen indikaattorit.
- **Tutkimusorganisaatiot** tuottavat ja tarjoavat yhteistyössä kansallisten ja kansainvälisten toimijoiden kanssa tutkijoille ja tutkimusryhmille ohjeistuksen sekä riittävät tukipalvelut hyvän datanhallinnan periaatteiden mukaiseen tiedon tallennukseen ja julkaisemiseen.

2023

- **Tutkimusorganisaatiot** sisällyttävät datanhallintasuunnitelman laadinnan ja ylläpidon osaksi tutkimus- ja palveluprosesseja aineiston elinkaaren jokaisessa vaiheessa.
- **Tutkimusorganisaatiot** hyödyntävät soveltuvia indikaattoreita, joiden avulla organisaatiot seuraavat edistymistään hyvässä datanhallinnassa osana laatujärjestelmää.
- **Tutkimusorganisaatiot** kuvaavat kansallisessa yhteistyössä, millaisia asiantuntijarooleja tarvitaan ja luovat asiantuntijarooleille mahdollisia urapolkuja.

2024

- **Korkeakoulut** huolehtivat, että opinnäytetöiden ohjaajat osaavat osana ohjaustyötään arvioida ja kommentoida datanhallintasuunnitelmia.
- **Tutkimusorganisaatiot** ovat kehittäneet seurannan hyvän ja vastuullisen datanhallintakäytännön varmistamiseksi.

LIITE 3: LINJAUSTYÖRYHMÄN JÄSENET

Minna Ala-Mantila, CSC

Nina Edgren-Henrichson, Svenska litteratursällskapet i Finland

Pamela Gustavsson, Svenska litteratursällskapet i Finland

Kaisa Hakkila, Turun yliopisto

Anne Holappa, Suomen tekijänoikeudellinen yhdistys

Outi Hupaniittu, Suomalaisen kirjallisuuden seura

Nina Järviö, Tieteellisten seurain valtuuskunta (siht. -03/2020)

Meri-Tuulia Kaarakainen, Turun yliopisto

Miki Kallio, Oulun yliopisto

Anu Kantola, Luonnonvarakeskus

Liisa Karlsson, Helsingin yliopisto

Olli-Pekka Kaurahalme, Turun yliopisto

Joona Koironen, Metropolia Ammattikorkeakoulu

Tuija Korhonen, Helsingin yliopisto

Mari Elisa (MEK) Kuusniemi, Helsingin yliopisto/Tuuli

Hanna Lahdenperä, Tieteellisten seurain valtuuskunta (siht. 05/2020-)

Heidi Laine, CSC

Kristina Linnovaara, Svenska litteratursällskapet i Finland

Riku Louhimo, Työterveyslaitos

Juuso Marttila, Jyväskylän yliopisto

Juhani Naskali, Turun yliopisto

Anssi Neuvonen, Teknologian tutkimuskeskus VTT

Maria Niku, Suomalaisen kirjallisuuden seura

Tiina Nokkala, Turun ammattikorkeakoulu

Elina Nurminen, Metropolia Ammattikorkeakoulu

Susanna Nykyri, Tampereen yliopisto

Turkka Näppilä, Tampereen yliopisto

Rainer Oesch, Helsingin yliopisto

Pekka Orponen, Aalto-yliopisto & FCRD (pj -05/2020)

Seliina Päällysaho, Seinäjoen ammattikorkeakoulu
Antti Pursula, CSC
Maria Rehbinder, Aalto-yliopisto
Mari Riipinen, Turun yliopisto
Sulevi Riukulehto, Helsingin yliopiston Ruralia-instituutti
Tomi Rosti, Itä-Suomen yliopisto (pj 05/2020–)
Matti Ruuskanen, Turun yliopisto
Toni Saari, Itä-Suomen yliopisto
Jussi Salmi, Turun yliopisto
Nina-Mari Salminen, Luonnonvarakeskus
Hannele Seppälä, Kansallinen koulutuksen arviointikeskus Karvi
Janne Seppänen, Jyväskylän yliopisto
Tiina Sipola, Oulun yliopisto
Anne Sunikka, Aalto-yliopisto
Anna Suorsa, Oulun yliopisto
Ville Tenhunen, Helsingin yliopisto, tietotekniikkakeskus
Heidi Troberg, Vaasan yliopisto, Tritonia
Antti Tuomi-Nikula, Terveysten ja hyvinvoinnin laitos



Tieteellisten seurain
valtuuskunta



TIEDONJULKISTAMISEN NEUVOTTELUKUNTA
DELEGATIONEN FOR INFORMATIONSPRÖNING