

Tutkimusaineistojen määrälliset indikaattorit AVOTT-seurannassa

Avoimen tieteen ja tutkimuksen kansallisessa (AVOTT-)seurannassa käytettäviä tutkimusaineistojen määrällisiä indikaattoreita on käsitelty useissa eri yhteyksissä. Tämä dokumentti kokoaa kirjallisissa lähteissä esitetyt indikaattorit sekä ne avoimuuden edistämisen tavoitteet, johon kyseinen indikaattori liittyy. Lisäksi esitetään aiheesta vuosina 2023–2024 järjestetyissä työpajoissa ehdotetut indikaattorit, joista kuvataan myös käytettävissä olevat tiedonlähteet ja niiden kattavuus, jos nämä tiedot ovat saatavilla.

Tarkastelu painottuu suomalaiseen tiedeyhteisöön. Kansainvälisiä indikaattoreita käsitellään, jos nämä indikaattorit sisältyvät

1. AVOTT-seurantaan tai sen yhteydessä tehtävään tiedonkeruuseen tai
2. sellaiseen avoimen tieteen kansainväliseen seurantaan, johon vastataan kokonaan tai osittain AVOTT-seurannan avulla.

Katsauksen alussa listataan tällä hetkellä voimassa olevat linjaustavoitteet. Ehdotettu indikaattori on kohdistettu linjauksen tavoitteeseen, jos se on ollut mahdollista. Yleisemmät ehdotukset uusiksi indikaattoreiksi esitetään dokumentin lopussa.

Tutkimusaineistojen ja -menetelmien avoimuus.

Korkeakoulu- ja tutkimusyhteisön kansallinen linjaus ja toimenpideohjelma 2021–2025.

Osalinjaus 1: Tutkimusdatan avoin saatavuus

Linjauksen tavoitteet

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.

- Tutkimuspalvelupäivillä 2024 ehdotettu indikaattori
 - DMP (ei tarkempaa kuvausta)
- Indikaattorin soveltamisessa otettava huomioon
 - Tutkimusorganisaatioissa ei tiedetä kaikkien tutkimushankkeiden määrää
 - Tarkistetut DMP:t: organisaatio tarkistanut, että DMP on toteuttamiskelpoinen > tämä on mahdollinen, mittariksi kaikki DMP:t, jotka ovat käyneet kirjaston tarkistuksen, vertailuluvuksi tutkijoiden määrä sisältäen väitöskirjatutkijat.
- **Mahdollinen käyttö vuonna 2026**
 - Kerätään pohjadata kevyellä pisteytyksellä, jonka perusteella täsmennetään tarkempi pisteytys vuoden 2028 seurantaan.

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

- Ei ehdotettuja indikaattoreita

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

- AVOTT-seurannassa 2024 kysyttiin metatietueiden määrää kolmessa organisaation itse valitsemissa repositorioissa.
- **Käyttö vuonna 2026**
 - Kysytään tätä, pyydetään käyttämään samojen repositorioiden tietoja kuin vuonna 2024, mutta aikarajauksella esim. 2024 alusta 2025 loppuun.
 - Kysymyksiä työpajassa 11.3.: Mikä on latauspäivä? Miten kumulatiivisia aineistoja käsitellään?
- Tutkimuspalvelupäivillä 2024 ehdotetut indikaattorit
 - Avoin metadata vs. avoin data
 - Sähköiset laboratoriopäiväkirjat
 - Julkaisujen data availability statement > kansallinen, ei organisaatiossa seurattava mittari
 - Kuinka paljon aineistoja on käytetty uudelleen?
- Indikaattorikehityksen työpajat kevät 2023
 - Hyvään datanhallintaan liittyvien asioiden: DOI, urn, lisenssi, seuranta.

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

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

- AVOTT-seurannassa 2024 kysyttiin datanhallinnan asiantuntijoiden määrää
- **Käyttö vuonna 2026**
 - Käytetään indikaattorina, mutta muokataan pisteytysasteikkoa, X % tutkimushenkilökunnan määrästä, selvitetään aihetta käsittelevistä julkaisuista
 - Huom. muistetaan lisätä kouluttaminen tehtäviin datanhallinnan asiantuntijan määritelmässä
- Tutkimuspalvelupäivillä 2024 ehdotetut indikaattorit
 - Avoimen tieteen (+data) koulutuksiin osallistujien määrä
 - Opiskelijoiden määrät RDM-kursseilla
 - Opiskelijoiden määrä datanhallinnan koulutuksissa
 - Koulutustuntien, joissa aiheena datanhallinta, määrä
 - Annettujen ja saatujen koulutusten (daha) mittaaminen?
 - **Halutaanko kehittää indikaattori seurantaan 2026 datakoulutuksista?**

Osalinjaus 2: Tutkimusmenetelmien ja -infrastruktuurien avoin saatavuus

TAVOITE 1: Vuoteen 2023 mennessä avoimet tutkimusmenetelmät ja -infrastruktuurit tunnistetaan tutkimuksen laadun ja vaikuttavuuden keskeisiksi tekijöiksi.

- Tutkimuspalvelupäivillä 2024 ehdotettu indikaattori
 - Infrojen DMPolien määrä (avointen)
 - Työ tutkimusaineistojen parissa myös tutkijalle meritoiva

TAVOITE 2: Vuoteen 2024 mennessä tutkimusorganisaatiot ja tutkimusinfrastruktuurien tarjoajat tukevat avoimia tutkimusmenetelmiä.

TAVOITE 3: Vuoteen 2025 mennessä tutkimusmenetelmiin ja -infrastruktuureihin liittyvistä oikeuksista ja ehdoista sovitaan selkeästi kaikkien tutkimustyöhön osallistuvien kesken.

Seurantamalli

Tutkimusaineistot ja -menetelmät

Organisaatioiden tasomäärittelyyn käytettävät indikaattorit aineistojen avoimen saatavuuden seuraamiselle aikaisintaan vuonna 2024:

1. Metatiedot. Tietolähde: Metax.
2. Menetelmien avoimuuden tukeminen: Tietolähde määritellään osalinjauksessa 2022.
3. Avointen aineistojen/datasettien määrä metatietojen perusteella TAI avointen materiaalien lataukset, uudelleenkäyttö (julkaisuissa) ja viittaukset. Tietolähde: Metax.

Tulevaisuuden näkymät ja ehdotetut kehitettävät indikaattorit:

1. DMP-seuranta: DMP-määrät, niiden arviointi organisaatioissa.
2. FAIR-periaatteiden noudattaminen.
3. Menetelmien ja ohjelmistojen avoimuuden seuranta.
4. Esirekisteröintien seuranta.
5. Kuvailevan metadatan avoimuus.
6. Linkitetty data ja/tai PID-graph.

Infrastruktuurit

Organisaatioiden tasomäärittelyyn käytettävät indikaattorit infrastruktuurien avoimen saatavuuden seuraamiselle vuonna 2024:

1. Tutkimusinfrastruktuurien ja käyttöpolitiikkojen näkyvyys tutkimustietovarannossa. Tietolähde: tutkimustietovaranto.
2. Tämä on jo hyväksytty indikaattoriksi avoimen tieteen ja tutkimuksen ohjausryhmässä seurantaan 2024. Silloin kuitenkin tiedonkeruuta ei ollut saatu käyntiin. Jos tiedonkeruun prosessit ovat valmiit ja toimivat vuonna 2026, tämä indikaattori otetaan käyttöön.

Tulevaisuuden näkymät ja ehdotetut kehitettävät indikaattorit:

1. Tutkimusinfrastruktuurien käyttöasteen seuranta: viittaukset ja viittauskäytänteet.

2. EOSC:issa kehitteillä olevat mittarit, standardit ja/tai suositukset sekä tietovarannot, jotka sisältävät tietoja kaiken kokoisista infrastruktuureista, esimerkkinä Open Iris.

AVOTT-seurannassa 2024 käytetyt indikaattorit

Metatiedot

Ilmoita metatiedot organisaatiosi itse valitsemista enintään kolmesta repositoriosta. Ilmoitetaan repositorion nimi ja metatietueiden määrä. Tiedot raportoidaan repositoriokohtaisesti.

Datan määrä on koko kumuloitunut data ilman aikarajausta. Ensimmäisen repositorion tiedoista saa yhden lisäpisteen. Jos annat tiedot useammasta kuin yhdestä repositoriasta, saat toisen lisäpisteen.

Ilmoitettu määrä täytyy voida tarkistaa. Repositoriossa täytyy olla haku, jossa voi tehdä tiedonhaun organisaation nimellä.

Vastaukseen täytyy sisältyä linkki repositorioon.

Datanhallinnan asiantuntijatyö

Kuinka paljon organisaatiossa käytettiin työaika datanhallinnan asiantuntijatyöhön vuonna 2023? Vastaus on arvio ja annetaan henkilötyövuosina yhden desimaalin tarkkuudella.

Datanhallinnan asiantuntija voi työskennellä myös muualla kuin kirjastossa tai tutkimuspalveluissa.

Kommentti Tutkimuspalvelupäivien 2024 työpajasta

- Kysymys datatukihenkilöistä oli liian kiltti, siitä sai liian helposti hyvät pisteet

Tutkimuspalvelupäivien 2024 työpaja

Yleisiä ehdotuksia, jotka eivät kohdistu tiettyyn linjaustavoitteeseen tai joissa ei ehdoteta määrällistä indikaattoria

- Laadullisia kuvauksia, esim. työkalun kehittäminen
- Avoimia laadullisia kuvauksia data rescue -projekteista – esim. mikä tärkeä aineisto on avattu ja saatettu käyttöön
- Miten työt on järjestetty > ohjeeksi muille
- Työn tekeminen näkyväksi
- Seuranta vipuvartena ”ei kannata tehdä priimaa”
- AMKeilla oma seuranta?
- Ollaan linjassa kv-linjausten ja kv-seurantojen kanssa
- Avoimuus ei ole vain readily available vaan löydettävissä & saavutettavissa
- Meillä on yhteiset kuvailutietokäytännöt/tietomallit.
- Tieteenalakohtaiset erot ja konventiot aineistojen jakamisessa – generalismin mittaaminen

Laadullisen indikaattorin käyttöä vuoden 2026 valmistellaan omana erillisenä kokonaisuutena. Asiantuntijayhteisölle järjestetään työpaja aiheesta kevään 2025 aikana.

AVOTT-indikaattorityöpaja 13.11.2024

- Tutkimusorganisaatiot saavat pyyntöjä tietoaaineiston avaamiseen, tätä pitäisi kannustaa esim. laadullisella mittarilla?

EOSC

EOSCin seurannassa kysytään käyttötapauksia datanhallinnalle, FAIRin mukaiselle datalle, avoimelle datalle, avoimen lähdekoodin ohjelmistoille, palveluiden tarjoamiselle EOSCin kautta, datanhallinnan asiantuntijoille, repojen yhdistämiselle EOSCiin, pitkäaikaissäilytykselle, avoimen tieteen koulutukselle ja taidoille, kannustimille sekä kansalaistieteelle. Käyttötapauksissa haetaan kuitenkin kansallisia tai kansallisessa yhteistyössä toteutettuja tms. isompia kokonaisuuksia, ei yksittäisen tutkimusorganisaation käyttötapauksia.

EOSCin seurannan aineistoihin liittyvät määrälliset kysymykset:

- Share of Country Data Management Plans
- Share of Country FAIR Data Sets
- Share of Country Open Data Sets
- Share of Country Open Source Software Sets
- Share of Country Services Offered through EOSC
- Share of Country Data Stewards
- Share of Country Repositories Connected to EOSC
- Share of Country Repositories Offering Long-term Data Preservation
- Share of Country Educational Curricula with an Open Science Dimension (ei liity vain aineistoihin)

UNESCO

UNESCOn ensimmäisessä seurannassa, jonka kysely tehtiin loppuvuonna 2024 ja alkuvuonna 2025, ei ollut tutkimusaineistojen, -menetelmien tai -infrastruktuurien määrällisiä kysymyksiä.