

18.9.2020

Kommentit tutkimusdatan avoimen saatavuuden osalinjauksesta

Tieteellisten seurain valtuuskunta on pyytänyt Tieteentekijöiden liittoa lausumaan Tutkimusdatan avoimen saatavuuden osalinjauksesta 25.9.2020 mennessä.

Johdanto

Osaluonnoksen johdannossa ja myös muualla asiakirjassa todetaan, että linjaus on *suomalaisen tutkimusyhteisön yhteistyön tulos*. Suhtaudumme kriittisesti tähän lausumaan. Tosiasiassa osalinjauksen laatinut työryhmä koostuu ensisijaisesti yliopistojen hallintohenkilökunnasta sekä informaatioalan ammattilaisista. Esimerkiksi väitöskirjatutkijoiden, professorien ja post doc -tutkijoiden näkökulmaa osalinjauksessa ei ole. Vaikka osalinjaus on avoimesti kommenttikierroksella, on aika rohkeaa todeta, että asiakirja olisi koko tutkimusyhteisön yhteistyön tulos. Väite pitäisi siis joko pystyä perustelemaan, tai muuttaa tai tarkentaa tätä lausetta.

On pidettävä ansiona, että johdannossa mainittuja asioita on kypsyystasoselvityksen muodossa konkretisoitu myöhemmin asiakirjassa. On hyvä, että kansainvälinen konteksti huomioidaan alusta alkaen. Lisäksi pidämme ansiokkaana, että johdannossa on todettu hyvien tieteellisten käytäntöjen ja hyvän aineistohallinnan toteuttamisen vaativan resursseja.

Riskit ja uhat -otsikon alle on kirjoitettu tutkijan oikeusturvasta. Tutkijan oikeusturvalla haasteita aiheuttavat muun muassa määräaikaikaisuudet sekä liikkuvuus. Tutkimusaineiston osalta on hyvin ongelmallista se, että tutkijan neuvotteluasema on erittäin huono eikä kollektiivista neuvotteluelintä ole olemassa. Sopimusten osalta on huomioitava, mikä on osapuolten neuvotteluasema ja todellinen mahdollisuus sopia. Tutkijan oikeusturvaa sekä oikeutta keräämäänsä aineistoon on syytä korostaa läpi koko asiakirjan nykyistä enemmän. Riskit ja uhat -otsikon alla mainitun tutkijan oikeusturvan lisäksi on huomioitava myös tutkittavien oikeusturva.

On huomioitava, että tieteenaloissa on todella suuria eroja koskien tutkimusaineiston ja -datan avoimuutta. Osalla tieteenaloista ei välttämättä edes kerry juurikaan tutkimusaineistoa. Jos tutkimuksessa ei käytetä dataa (esimerkiksi matematiikan tutkimus) tai tutkija käyttää arkiston dataa, jota hän ei saa avata, tämäkin on huomioitava eivätkä datan avoimuuden kriteerit saa johtaa tutkimuksen vinoutumiseen.

Sanasto-otsikon alla todetaan seuraavasti: *"Tieteenalojen erilaisuudesta johtuen, vakiintunutta sanastoa ei tule koskaan syntyään ja jokainen aihetta käsittelevä joutuu määrittelemään käyttämänsä termit parhaaksi katsomallaan tavalla"*. Olemme eri mieltä tästä virkkeestä ja pyydämme sen poistamista kokonaisuudessaan. Vaikka käsitteistöä ei vielä tässä dokumentissa pystytä luomaan, sen tarve on tunnustettava. On pyrittävä siihen, että löydetään yhteiset termit eri aloille.

Sanastosta puuttuvat lähdeviitteet ja myös oikoluku on ilmeisesti sen osalta vielä tekemättä. Tutkimusaineisto-kappaleessa on kirjoitusvirhe, se sisältänee yhden "tai"-sanon liikaa. Huomautamme, että *laboratoriopäiväkirja* on huono esimerkki tutkimusdatasta, sillä on tulkinnanvaraista mihin kategoriaan laboratoriopäiväkirjat luokitellaan.

Johdannossa viitataan FAIR-periaatteisiin. Olisi kuitenkin tärkeä huomioida, että myös niillä aloilla, joilla kokonaisvaltainen FAIR-periaatteiden noudattaminen ei ole mahdollista, tavoiteltaisiin nykyistä avoimempaa tutkimusdataa. On hyvä periaate, että data on koneiden saavutettavissa ja luettavissa, mutta yhtä aikaa on tunnustettava, että joillakin aloilla tutkimusdata ei ole vielä edes ihmisten saavutettavissa. Olisi hyödyllistä määritellä erikseen nk. koneellinen FAIR-periaatteiden alainen tutkimusdata ja metadata sekä niin sanotusti ihmisten luettavissa oleva data (*human fair* ja *machine fair*).

Yleiset kommentit

Osalinjaus on sisällöllisesti varsin hyvä sekä perusteltu. Dokumentin rakenne on kuitenkin haastava ja vaikeuttaa sen ymmärtämistä. Muu muassa toimeenpano ja seuranta on erikoisessa kohdassa dokumenttia. Tavoitteiden osalta on väljyyttä ja epämääräisyyttä, mikä saattaa johtua myös asiakirjan rakenteesta.

Pidämme erittäin tärkeänä, että tieteentekijöiden ja erityisesti haavoittuvimmassa asemassa olevien tieteentekijöiden tulee päästä vaikuttamaan sekä nykyisen lähtötilanteen kartoitukseen että seurantaan koskien kaikkia periaatteita. Tutkijoiden äänen tulee kuulua paitsi osalinjauksen konkreettisessa toimeenpanossa myös datan avaamisen jatkokehittelyssä ja jatkokeskusteluissa. Avoimuutta ei ole syytä kehittää vain yhteistyössä tutkimusorganisaation kanssa, vaan tulee miettiä, miten avoimuutta kehitetään yhteistyössä tutkijoiden kanssa.

Kommentoi linjauksen periaatetta 1 ja sen toteutumisen vastuita:

Periaate 1, tutkimusaineistoa avataan vain vastuullisesti, on sinänsä hyvä, mutta konkretia jää puuttumaan. Toteutumisen vastuut herättävät ihmetystä, sillä Tutkimuseettisellä neuvottelukunnalla tuskin on varsinaista vastuuta konkreettisessa tutkimusaineiston avaamisessa. Tutkimusaineiston avaamisessa vastuullisia ovat ainakin tutkijat, tutkimushankkeen vastuullinen johtaja, väitösohjaaja, tutkimusorganisaatiot sekä valtiovalta ja OKM. Näitä vastuita tulisi tunnistaa ja eritellä huomattavasti paremmin. Vastuullisuuden roolitusta on selkeytettävä (ks. esim. <https://openscience.jyu.fi/fi/tutkimusdata/tutkimusdatapolitiikka>).

Kommentoi linjauksen periaatetta 2 ja sen toteutumisen vastuita:

Pidetään periaatetta 2, jonka mukaan tutkijoilla on käytettävissään aineistonhallinnan infrastruktuurit ja palvelut, ja niitä kehitetään tutkijalähtöisesti, hyvänä ja tutkijalähtöisyyttä tärkeänä asiana.

Kommentoi linjauksen periaatetta 3 ja sen toteutumisen vastuita:

Periaate 3, jonka mukaan tutkijan ansioita hyvän aineistonhallinnan edistämässä, tutkimusaineistoon liittyvää työtä sekä tutkimusaineiston avaamista arvostetaan meritoitumiskriteereissä, on erittäin tärkeä. Pidämme hyvänä, että meritoitumista nostetaan esille. Meritoivuus on erittäin tärkeää ja edellyttää palkitsemisjärjestelmän laatimista. Meritoivuudessa on kuitenkin huomioitava todella suuret tieteenalakohtaiset erot datan avoimuudessa ja ylipäättään aineistojen käyttämisessä. Yleisellä tasolla tutkijoiden datan ja

aineistojen keräämiseen käyttämää työtä olisi tunnistettava, ja siitä olisi annettava tunnustusta, huomattavasti nykyistä paremmin.

Kommentoi osalinjauksen tavoitetta 1 ja sitä täsmentäviä toimenpiteitä:

Tavoite 1 koskien datanhallintasuunnitelmia on hyvä, mutta se jää hieman epämääräiseksi, sillä aloitettavia tutkimus- ja kehittämisprosesseja ei ole määritelty mitenkään. On epäselvää, miten mitataan tavoitteen toteuttamista. Kustannuksiin varautumisessa tulee huomioida myös henkilöstöresurssit.

Kommentoi osalinjauksen tavoitetta 2 ja sitä täsmentäviä toimenpiteitä:

Tavoitteen kaksi alla mainittujen kansallisen sopimusmallien laatimisessa on erittäin tärkeää, että työntekijäosapuoli on edustettuna. Vastuullinen aineistojen käyttö edellyttää sitä, että neuvotteluasema tutkijoiden ja tutkimusorganisaatioiden välillä on tasapainossa.

Kommentoi osalinjauksen tavoitetta 3 ja sitä täsmentäviä toimenpiteitä:

Tavoite 3 koskien tutkimusdatojen dokumentointia on todella hyvä ja tärkeä, mutta aikataulua pidetään mahdollisesti epärealistisena. Toimenpiteistä ei käy ilmi, tuotetaanko tukipalvelut keskitetysti esimerkiksi OKM:n rahoituksella vai tuottaako jokainen korkeakoulu palvelut itse.

Kommentoi osalinjauksen tavoitetta 4 ja sitä täsmentäviä toimenpiteitä:

Tavoite 4 on haastava, sillä tavoitteessa puhutaan ainoastaan tutkimusdatasta, mutta datanhallinta ja metatietojen tallentaminen jää epäselväksi. Osalinjauksessa esitettyjä tavoitteita olisi hyvä kytkeä yhteyteen keskenään. Sanamuodosta jää sellainen kuva, että kyseessä olisi suuri yksi datasiilo. Tavoitteen sanamuoto olisi syytä muuttaa monikkomuotoon. Lisäksi tavoitteeseen olisi hyvä lisätä *tallennus- ja säilytyspaikan* lisäksi myös *käsittelypaikan* termi.

Kommentoi osalinjauksen tavoitetta 5 ja sitä täsmentäviä toimenpiteitä:

Tavoitteessa 5 koskien ammatillisia rooleja tulisi mainita erikseen informaatioalan osaaminen, joka on aivan keskiössä tavoitteessa. Termit ”ammatillinen rooli” ja ”ammatillinen koulutus” ovat tavoitteen yhteydessä harhaanjohtavia. Tulisi puhua asiantuntijarooleista sekä asiantuntijoiden osaamisen kehittamisestä. Samoin tieteidenvälisyys olisi tarpeen huomioida tavoitteessa.

Lisätietoja: Salla Viitanen, lakimies, salla.viitanen@tieteentekijat.fi