

Asia: VN/16660/2020

## **Lausuntopyyntö luonnoksesta hallituksen esitykseksi avoimen datan direktiivin täytäntöönpanoa koskevasta lainsäädännöstä**

### Tiedonhallintalain ja julkisuuslain muutokset

**Onko ehdotuksella vaikutuksia, joita ei ole vielä tunnistettu, erityisesti lakien soveltamisalassa oleviin organisaatioihin? Millaisista vaikutuksista on kyse tai miten niitä olisi mahdollista arvioida tarkemmin?**

Laissa tulisi selkeästi määritellä se, mitä tutkimusaineistoksi tulkitaan. Lisäksi Turun yliopisto pitää erityisen tärkeänä lähdekoodin sisällyttämistä osaksi tutkimusaineiston määritelmää. Nykymuodossaan lakiesitys ei huomioi lähdekoodia. Lähdekoodi on kuitenkin oleellinen osa tutkimusaineiston tulkintaa ja sen merkitys on nostettu esiin niin kansallisissa kuin Euroopan komission määrittelemissä ohjeissa. Lähdekoodin tärkeys nousi esiin Turun yliopiston tutkijan Leo Lahden kommentissa ja Turun yliopiston kanta on yhteinen hänen kommenttinsa kanssa.

**Ehdotuksessa ja direktiivissä ei velvoiteta rajapintojen, jotka olisivat nimenomaan avoimia, käyttöön. Mitä mieltä olette väitteestä, että viranomaiset oman harkintansa mukaan käyttäisivät avoimia rajapintoja joka tapauksessa silloin, kun se on tarkoituksenmukaista?**

Ei kommentoitavaa.

**Onko ehdotettujen muutosten johdosta tarpeen muuttaa myös jotain erityislainsäädäntöä, jota ei esitysluonnoksessa ole tunnistettu?**

Ei kommentoitavaa.

### Laki eräiden julkisten yritysten tietojen uudelleenkäytöstä

**Kommentit laista eräiden julkisten yritysten tietojen uudelleenkäytöstä kokonaisuutena.**

-

**Millaisiin yrityksiin toimialallanne tai alueellanne ehdotettu laki tulisi sovellettavaksi?**

-

**Tunnistatteko julkisia yrityksiä koskevaa sääntelyä, jota olisi ehdotetun lain johdosta muutettava tai joka on sen kanssa ristiriidassa?**

-  
**Onko tiedossanne sääntelyä, joka velvoittaa ehdotetun lain soveltamisalassa olevia julkisia yrityksiä antamaan niiden hallussa olevia asiakirjoja uudelleenkäytettäväksi?**

-  
**Onko tiedossanne sääntelyä, joka velvoittaa ehdotetun lain soveltamisalassa olevia julkisia yrityksiä julkaisemaan tai julkistamaan niiden hallussa olevia asiakirjoja, mutta jonka tosiasiallisena tarkoituksena tai vaikutuksena on tietojen uudelleenkäyttö kaupallisiin tai muihin tarkoituksiin?**

-  
**Onko ehdotetun lain soveltamisalassa oleviin yrityksiin tai muihin toimijoihin sellaisia vaikutuksia, joita ei ole tunnistettu? Millaisista vaikutuksista on kyse tai miten niitä olisi mahdollista arvioida tarkemmin?**

-  
**Onko asiakirjojen uudelleenkäytön salliminen ominaista tai olennaista jonkin ehdotetun lain soveltamisalassa olevan julkisen yrityksen liiketoiminnalle?**

Ei kommentoitavaa "Laki eräiden julkisten yritysten tietojen uudelleenkäytöstä" liittyviin kohtiin.

## Laki julkisin varoin tuotettujen tutkimusaineistojen uudelleenkäytöstä

### **Kommentit laista julkisin varoin tuotettujen tutkimusaineistojen uudelleenkäytöstä kokonaisuutena.**

Ehdotetusta laista puuttuu tietoaineiston julkaisijan pysyvää vastuuta käsittelevä kohta, erityisesti julkaistun tietoaineiston julkaisemisen päättäminen. Lakiehdotuksessa tulisi määrittää onko mahdollista ja millä ehdoilla voi julkaistun tietoaineiston julkaiseminen lopettaa. Onko aineiston julkaisijan aina pidettävä kerran julkaistua tietoaineistoa julkaistuna ts. uudelleenkäytettävissä?

Lisäksi pysyvään vastuuseen liittyy miten vastuu siirtyy tutkijan, tutkimusorganisaation ja rahoittajan välillä esim. tutkijan kuolemantapauksessa tai tutkijan organisaatiovaihdoksen yhteydessä. Aiheuttaako tämä ongelmatilanteita tekijäoikeudellisiin ja vastuukysymyksiin liittyen.

Millä varmistetaan julkaistun tietoaineiston ajantasaisuus esimerkiksi julkaisuformaatin teknisen vanhentumisen kannalta tai alkuperäisen julkaisijatahon lakkaaminen (avoimet repositoriot) esim. sulautumisen, konkurssin, kuoleman tms. syyn johdosta. Tallennusformaatin kehittyessä kenellä on velvollisuus päivittää tutkimusaineistot käytössä oleviin formaatteihin. Tämä vaatii myös resursseja.

Myös se, että asiat kehittyvät huimaa vauhtia ja mm. tallennusformaatit saattavat muuttua ja jotkut tallennusformaatit saattavat lakata olemasta. Miten silloin pitää huolehtia päivittämisestä, onko velvollisuus muuttaa toiseen muotoon? Tämä kaikki vaatii myös resursseja. Nykytilanteessa datan tallennusta ja säilömistä tehdään projektiluontoisesti erityisrahoituksella. Tästä syystä tulee huomioida mitä tapahtuu, kun projekti tai rahoitus päättyy.

**Onko ehdotuksella vaikutuksia tutkimusaineistoja tuottaviin tai julkaiseviin tahoihin, joita ei vielä ole tunnistettu? Millaisista vaikutuksista on kyse tai miten niitä olisi mahdollista arvioida tarkemmin?**

Vaikutuksia on käsitelty myös edellisessä kohdassa liittyen resursseihin ja tutkimusdatan avoimeen saatavuuteen pitkällä aikavälillä. Onko laissa huomioitu julkaisevien tahojen vastuu tutkimusaineiston oikeellisuudesta ja miten tämä vaikuttaa aineiston uudelleenkäyttöön, jos alkuperäiseen aineistoon on jäänyt esim. virheitä, joita ei ole uudelleenkäyttövaiheessa huomattu.

**Onko olemassa riskiä siitä, että tutkimusaineistoja ei enää asetettaisi julkisesti saataville sen vuoksi, että ehdotettu laki tulisi niihin julkaisun johdosta sovellettavaksi?**

Nykymuodossa epäselväksi jää, miten laki tulkitsee sitä mikä on julkisin varoin rahoitettua. Miten esim. säätiöiden ja apurahojen tuella tehtyjä tutkimuksia sovelletaan tämän lain piiriin? Mikäli tätä ei tarkenneta, saattaa tietyistä rahoitusmuodoista tulla sellaisia, ettei niiden avulla kerätä aineistoja, jos rahoittajan ehtoina on aineistojen välitön avoimuus ja tutkijalla on mahdollisuus kerätä ko. aineisto toisen rahoitusmuodon avulla (sellaisen, joka ei edellytä välitöntä avoimuutta).

On mahdollista, että laissa esitetyt avoimuuden vaatimukset rajoittavat tutkimusaineiston julkaisemista joillain tieteenaloilla, joissa tutkimusaineistojen julkaisemista ei pystytä toteuttamaan käytäntöjen puuttuessa, resurssien ollessa riittämättömät ja jos julkaiseminen ei itsessään vaadi tutkimusaineiston julkiseksi asettamista. Lisäksi tutkimuslaitosten ja tutkijoiden resurssien puute saattaa vaikuttaa siihen, että lain velvoitteet toteutuvat eri tieteenaloilla eri tavalla.

Lain valmistelussa tulisi huomioida, minkälaisia vaikutuksia välittömästi julkaisemisesta on tutkimusorganisaatioihin ja yksittäisiin tutkijoihin. Välitön julkaiseminen saattaa vaikuttaa tutkimuksen tekemiseen merkittävästi erityisesti silloin, kun kerätään laajoja aineistoja, joita on tarkoitus hyödyntää useissa eri tutkimuksissa, useiden tutkijoiden toimesta ja useamman vuoden ajan.

**Missä laajuudessa tutkimusaineistojen maksutonta saatavuutta edellytetään jo nykyään osana julkista tutkimusrahoitusta?**

Tällä hetkellä tutkimusaineistojen maksutonta saatavuutta edellytetään jo suuressa osassa tutkimusrahoitusta ja julkaisukanavia. Tässä on kuitenkin vielä tutkimusalakohtaisia eroja.

**Tunnistatko tutkimusaineistoja, niiden tuottamista tai hyödyntämistä koskevaa sääntelyä, jota olisi ehdotetun lain johdosta muutettava tai joka on ehdotetun sääntelyn kanssa ristiriidassa?**

Ei kommentoitavaa.

## Arvokkaat tietoaineistot

**Onko toimialallanne tai toiminnassanne tunnistettu tietoaineistoja, jotka mahdollisesti tai todennäköisesti tultaisiin komission toimesta säätämään arvokkaiksi tietoaineistoiksi?**

Arvokkaiden tietoaineistojen tunnistaminen on tällä hetkellä määritysten puuttuessa hankalaa ja arvioita arvokkaista tietoaineistoista ei ole tehty. Arvokkaita tietoaineistoja todennäköisesti sisältyy esitettyihin teemakategorioihin. Arvokkaiden tietoaineistojen määrittäminen tutkimusorganisaatioissa vaatii kuitenkin resursseja ja ne tulisi selkeästi määrittellä, jotta niiden määrittäminen onnistuu.

**Onko toimialallanne tai toiminnassanne varauduttu arvokkaiden tietoaineistojen sääntelystä johtuviin vaikutuksiin? Miten kuvailisitte vaikutuksia ja niihin varautumista? Antakaa mahdollisuuksien mukaan myös euro- tai lukumääräarvioita.**

Turun yliopiston ei tässä vaiheessa ole varautunut arvokkaita tietoaineiston sääntelyyn kohdistuviin vaikutuksiin.

## Yleiset kommentit

**Yleiset kommentit hallituksen esityksen luonnoksesta ja ehdotetun sääntelyn vaikutuksista kokonaisuudessaan.**

Tiheästi tuetettavan datan avaaminen edellyttää teknisiä vaatimuksia, jota varten tarvitaan taloudellisia resursseja ja henkilöstön erityisosaamista. Tutkimusorganisaatioissa tuotetaan moninaista tietoa, joka direktiivin mukaan tulisi olla avointa. Tämän tiedon tunnistaminen vaatisi resursseja ja on tutkimusorganisaatioille suuri haaste.

Lisäksi tutkimusorganisaatioiden keräämän reaaliaikaisen päivittyvän tiedon (dynaaminen data) ja edellytys tiedon asettamisesta julkiseksi on erittäin iso haaste vuoteen 2022 mennessä. Nykymuodossaan epäselväksi jää myös reaaliaikaisen päivittyvän tiedon määritelmä. Tällä hetkellä tutkimusorganisaatiot tuottavat esim. huomattavan määrän erilaisia tilastotietoja (mm. yliopiston sisäiset selvitykset ja tilastoinnit toiminnan kehittämiseksi), miten laki velvoittaa ja miten sitä sovelletaan näiden tietojen julkaisemiseen? Katsotaanko näiden tietojen kuuluvan dynaamiseen dataan ja tämän lain piiriin, mikäli eivät sisällä henkilötietoja?

Lain valmistelussa tulisi huomioida, ettei sen avulla ohiteta tekijänoikeudellisia kysymyksiä ja FAIR-periaatteita (niin avointa kuin mahdollista ja niin suljettua kuin tarpeellista).

Hakkila Kaisa  
Turun yliopisto