

Asia: VN/7791/2021

Hallituksen esitys eduskunnalle laeiksi tilastokeskuksesta annetun lain ja tilastolain muuttamisesta

Lausunnonantajan lausunto

1. Onko Tilastokeskukselle hallituksen esityksessä kirjattu uusi tehtävä selkeästi kuvattu?

CSC kiittää valtiovarainministeriötä mahdollisuudesta lausua hallituksen esityksestä laeiksi tilastokeskuksesta annetun lain ja tilastolain muuttamisesta. CSC – Tieteen tietotekniikan keskus Oy on Suomen valtion ja korkeakoulujen omistama, voittoa tavoittelematon erityistehtäväyhtiö. Palvelemme laajasti koko yhteiskuntaa tuottamalla teknologiapalveluja ja -ratkaisuja TKI-toiminnalle, koulutukselle, kulttuurille ja julkishallinnolle. CSC tuottaa palveluja myös tutkimuksen arkaluonteisen datan käsittelyyn, ja on merkittävästi osallisena eurooppalaisissa infrastruktuuri- ja tutkimushankkeissa.

CSC yhtyy huoleen tutkijoille suunnattujen palveluiden ja tutkimusinfrastruktuurien resursoinnista. Monet palvelut käynnistyvät määräaikaisina pilotteina tai projekteina. Ylläpito ja kehitys vaativat kuitenkin rahoitusta myös käynnistysvaiheen jälkeen, mihin tulee varautua etukäteen. Budjettirahoituksella toimivien organisaatioiden kohdalla kyse on osin myös poliittisesta tahdosta ja organisaation prioriteeteistä. Tutkimusinfrastruktuurien ja ekosysteemien rahoitukseen tulee panostaa kansallinen kokonaiskuva ja kansainvälinen yhteistyö huomioiden.

Tilastokeskus on merkittävä tutkimusekosysteemin toimija sekä oman tutkimustoimintansa, että muille tutkijoille tarjoamiensa resurssien kautta. Korkealaatuisen ja riippumattoman tutkimuksen edellytyksenä on muun muassa tutkijoiden mahdollisuus saada ja yhdistellä tutkimusmateriaalia ja -palveluita eri lähteistä.

Julkisten resurssien käytön kannalta on järkevää, että niitä kohdennetaan tarkoituksenmukaisesti ja eri toimijoiden ja palveluiden välillä on työnjako, jolloin ne toiminnallisesti tukevat toisiaan, eivätkä kilpaile keskenään, vaikka toimisivatkin samalla kentällä. Työnjaon tulee määrittyä vastuista, ja huomioida eri toimijoiden erilaiset tavoitteet. Tutkimuksen ja siihen pohjautuvan päätöksenteon

näkökulmista on tärkeää mahdollisuus yhdistellä ja käsitellä eri lähteistä peräisin olevia aineistoja yhdessä. Tämä tulee huomioida lakiesityksessä myös siitä näkökulmasta, että Tilastokeskuksen luovuttamia aineistoja tulee voida käsitellä muissakin palveluissa, kuten esimerkiksi toisiolain puitteissa Findatan luovuttamia aineistoja.

Nyt lausuttavana olevan lakiehdotuksessa Tilastokeskuksen lakisääteisiä tehtäviä ollaan ehdotetun 2 a § perusteella lisäämässä tutkija- ja aineistopalvelun tuottamiseen. Tutkimusaineistojen ja -palveluiden lisääminen tutkijoille on lähtökohtaisesti kannatettava asia. Koska Tilastokeskus on julkisesti rahoitettu organisaatio ja osa kansallista tutkimusekosysteemiä, kansainvälisiä kytkentöjä unohtamatta, uusien tehtävien perustamisen yhteydessä tulee huomioida niiden vaikutus kokonaisuuteen. Lakiesityksen perusteluiden tulee sisältää arviointi, millainen vaikutus esityksen mukaisella tutkija- ja aineistopalvelulla on kansallisesti. Tehtävien tarkemman määrittelyn lisäksi esitykseen tulee sisältyä analyysi toimintakentästä sekä kuvaus työnjaosta muiden toimijoiden kanssa, mukaan lukien suhde muiden toimijoiden tehtäviin. Samoin tulee arvioida lakiehdotuksen vaikutukset tutkimukselle sekä selvittää tarkemmin tutkijoiden tarpeita. Korkeakoulut ja tutkimuslaitokset on mainittu lakiesityksen perusteluiden johdannossa, mutta niihin ei viitata myöhemmin pääasiallisten vaikutusten kuvauksessa.

Vaikutusarvioinnin puuttuessa Tilastokeskuksen tuleva rooli tutkimuskentällä jää epäselväksi. Tilastokeskuksen ja muiden toimijoiden roolien määrittely sekä vaikutusten arviointi selventäisi lakiesitystä myös tältä osin. Avoin mandaatti tietojen keräämiseen muihin kuin tilastotarkoituksiin nostaa esiin lukuisia kysymyksiä esimerkiksi siitä, kuinka paljon resursseja tähän toimintaan on varattava. Tutkimusnäkökulmasta laajatin aineistot sinänsä voivat olla jopa arvottomia, ellei niitä kuratoida asianmukaisesti. Muualta saatujen aineistojen käytettävyyteen liittyy teknisten näkökulmien (esimerkiksi datasettien integriteetti ja koodaaminen yhteentoimivasti) lisäksi lukuisia muita аспекteja, kuten yhdenmukaiset metatiedot ja pysyvät tunnisteet. Näiden seikkojen huomioiminen vaatii työtä yhdenkin organisaation sisällä ja varsinkin siirrettäessä aineistoja toimijalta toiselle.

Tilastokeskuksen uuden mandaatin selventämiseksi lakiehdotuksen 2 a § muotoilua ”Tilastokeskus voi lisäksi ottaa vastaan, koostaa ja säilyttää tutkija- ja aineistopalvelua varten sellaisia tietoja, joita ei kerätä tilastointitarkoitukseen” voisi tarkentaa esimerkiksi muotoon ”Tilastokeskus voi lisäksi tilastolaissa määritellyiltä tahoilta ottaa vastaan, koostaa ja säilyttää tutkija- ja aineistopalvelua varten sellaisia tietoja, joita ei kerätä tilastointitarkoitukseen”. Tämä tarkennus selventäisi työnjakoa, mahdollistaisi resurssien tehokkaamman käytön ja olisi myös tutkijoiden näkökulmasta selventävä, millaisia aineistoja ja niihin liittyviä palveluita Tilastokeskus tarjoaa.

Eryteisesti ulkopuolisille tutkijoille tarjottavien palveluiden osalta olisi toimivuuden ja kustannustehokkuuden näkökulmasta perusteltua tehdä ero niihin rajoitettua käyttöä vaativiin palveluihin, joissa käsitellään sensitiivistä dataa, sekä muita datoja hyödyntäviin palveluihin, joissa ei tarvita yhtä raskasta tietoturvaa käyttäjätunnistuksineen. Lisäksi lakiesityksessä tulee huomioida aineistojen uudelleenkäytettävyys tieteen validoinnin näkökulmasta. Näin ollen esimerkiksi kerran pseudonymisoitu tai anonymisoitu aineisto tulee olla muidenkin kuin sitä ensimmäisen kerran

pyytäneen tutkijan käytettävissä, eikä pseudonymisointia tai anonymisointia tule tehdä samalle aineistolle uudelleen toista tutkijaa varten.

Tieteen avoimuus ja sen myötä tulosten validointi toistettavuudella on noussut viime vuosina entistä tärkeämmäksi lähtökohdaksi tutkimukselle. Tutkijoille suunnattujen palveluiden lähtökohdana tulee olla FAIR-periaatteiden noudattaminen ja datan avaaminen mahdollisimman avoimesti. Aineistojen tulee siis olla niin avoimia kuin mahdollista ja niin rajoitettuja kuin tarpeellista. Tämä koskee myös arkaluonteisia aineistoja, jotka voidaan saattaa tutkijoiden saataville anonymisoimalla tai pseudonymisoimalla. On huomattava, että FAIR-periaatteet ovat lähtökohta kaikkien aineistojen käsittelyyn. Myös arkaluonteisia tietoja sisältävistä aineistoista voidaan jakaa tietoa esimerkiksi niiden olemassaolosta (metatiedot), vaikka aineiston käyttö sinänsä olisi luvanvaraista.

Tietojärjestelmien yhteentoimivuus on keskeinen edellytys laajalle tiedon hyödyntämiselle yhteiskunnassa. Palveluiden ja aineistojen tulee olla yhteentoimivia sekä kansallisesti että kansainvälisesti eurooppalaisen yhteentoimivuuskehityksen (European Interoperability Framework) mukaisesti. Datat yhteentoimivuutta tulee edistää niin oikeudellisella, organisatorisella, semanttisella kuin teknisellä tasolla.

Toisin kuin lain taustoituksessa mainitaan, CSC ei ole OKM:n, vaan Suomen valtion ja korkeakoulujen omistama erityistehtävayhtiö. Pyydämme korjaamaan tämän tiedon lakiluonnoksen lopulliseen versioon.

Kimmo Koski

Irina Kupiainen

Toimitusjohtaja

Yhteiskuntasuhdejohtaja

2. Miten arvioitte valtioneuvoston etusijaa käyttöluupahakemusten käsittelyssä suhteessa alaiseen hallinnon viranomaisiin?

-

3. Ovatko hallituksen esityksen ehdotusten tietosuojaa koskevat pykälät riittävän täsmällisiä?

-

4. Onko aineiston tietosuojaoikeudellinen käsittelyperuste riittävästi kuvattu?

-

5. Onko aineiston säilytystä koskeva tekstiosuus riittävän täsmällinen?

-

6. Onko tietojen käyttötarkoitus rajattu riittävän selkeästi?

-

7. Ovatko raaka-aineiston ja valmisaineiston käyttörajaukset riittävän selkeitä?

-

8. Mahdollistaako hallituksen esitys tutkijalle luovutettavien tietojen antamisen suoraan Tilastokeskuksen tutkijapalvelulle, kun tutkijan on tarkoitus käsitellä tietoja tutkijapalvelun etäkäyttöympäristössä?

-

9. Edistääkö hallituksen esitys tutkimuksessa hyödyllisen tiedon saamista tutkijoiden käytettäväksi?

-

Hyppölä Jenni