

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?

Tilastokeskuksesta annetun lain 2 a §:n 1 mom. mukainen lisäys tutkija- ja aineistopalvelun ylläpitämisestä Tilastokeskuksessa on kannatettava. Tilastokeskus on jo pitkään tarjonnut tilastointitarkoitukseen keräämiään tietoja tieteelliseen tutkimus- ja selvitystyöhön, ja kehittänyt tähän liittyviä palveluita ja järjestelmiä. Lakiehdotus selkeyttää tätä roolia, ja toivottavasti avaa uusia mahdollisuuksia tutkijapalveluiden rahoituksen järjestämiseen ja sujuvampaan yhteistyöhön tutkimuslaitosten kanssa. Tieteellisen tutkimuskäytön lisäksi SYKE tarvitsee tilastoaineistoja myös lakisääteisten tehtäviensä ja asiantuntijapalveluidensa hoitamiseen. Lakiehdotuksen sanamuodosta jääkin epäselväksi, mitä kaikkea sisältyy ilmaisun ”yhteiskuntaoloja koskevat tilastolliset selvitykset” alle.

Hallituksen esityksestä ilmenee, että aineistojen käyttöluvahakuprosessissa ja tilastokeskuksen etäkäyttäjärjestelmään liittyvissä yksityiskohdissa (isojen aineistojen käsitteleminen, aineistojen valmistelu, tulosten tarkastus yms.) on ollut käytännön haasteita. Esimerkiksi esityksessä mainittu käyttöluvahakemusten 10–12 kuukauden jonotusaika on kestävä tutkimuksen ja päätöksenteon näkökulmasta. Jotta tilastokeskus voisi ”ylläpitää, kehittää ja parantaa” kustannustehokasta, laadukasta ja tietoturvallista etäkäyttäjärjestelmää, tarvitaan resursseja teknisten aineistojen jakamis- ja analyysiratkaisuiden kehittämiseen. Henkilötyöstä riippuvaisten työvaiheiden vähentäminen toisi säästöjä, tehokkuutta ja yhdenvertaisuutta. Lisäksi aineistojen käsittely-ympäristön tulisi vastata resurssitehokkaasti nyky-yhteiskunnan tietotarpeisiin ja tietojenkäsittelyn haasteisiin, mukaan lukien isojen aineistojen, sekä paikkatietopohjaisten aineistojen yhdistely ja analysointi.

SYKE toteaa, että tällä hetkellä Tilastokeskuksen tutkijapalvelu ja sen etäkäyttäjärjestelmä eivät ole riittävän tehokas ja helppokäyttöinen väline SYKEssä tehtävään tieteelliseen tutkimustyöhön. Esimerkiksi tilastokeskuksen nykyinen etäkäyttäjärjestelmä ei riittävästi palvele isojen paikkatietoaineistojen tehokasta hyödyntämistä. Ehdotettu lakimuutos ei ratkaise tilastokeskuksen

tarjoamassa etäkäyttöjärjestelmässä olevia haasteita. Siten SYKE katsoo, että paikkaan sidottuja tietoaineistoja tulisi olla mahdollisuus käyttää erityisesti spatiaalisessa mallinnuksessa myös muissa toimintaympäristöissä kuin tilastokeskuksen tällä hetkellä tarjoamassa FIONA:ssa. SYKE katsoo, että ongelman voisi ratkaista määrittelemällä tietoturvalliset käyttöympäristöt ja luoda niille sertifiointimenettely. Tämä olisi tutkimuslaitosten tarpeiden kannalta tällä hetkellä toimivampi ja mukautuvampi ratkaisu kuin mitä lakimuutoksessa esitetään.

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

Käyttölujarahakemusprosessin pitäisi olla lähtökohtaisesti niin sujuva, että käyttöluorien myöntäminen ja aineistojen valmistelu ei jonoudu kohtuuttomasti.

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

SYKEN mukaan on kannatettavaa, että ”Tilastokeskuksen on säilytettävä kuvaus luovuttamiensa tietoaineistojen ja aggregoitujen tilastotietojen muodostamisesta, käyttämistään pseudonymisointi- ja anonymisointimenetelmistä sekä luovutetuista lopputuloksista.” Näiden prosessikuvausten ja metatietojen tuottamisen tulee olla mahdollisimman pitkälle automatisoitu, jotta tästä ei aiheudu turhia kustannuksia tutkijapalveluiden ylläpidossa.

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?

Hallituksen esitys selkeyttää Tilastokeskuksen roolia tutkimustiedon tuottamisessa, mutta ei ratkaise ongelmia, jotka liittyvät tilastoaineistojen hyödyntämiseen tieteellisessä tutkimuksessa ja yhteiskuntaoloja koskevissa selvityksissä. Tilastolain 13 §:n 3 mom. mukaan tilastoviranomainen ei saa luovuttaa tieteellistä tutkimusta ja yhteiskuntaoloja koskevaa tilastollista selvitystä varten sellaisia tietoja, joiden perusteella tilastoyksikkö voidaan tunnistaa suoraan. Tilastoviranomainen voi kuitenkin antaa käyttöoikeuden välillisesti tunnistettaviin salassa pidettäviin tietoihin.

SYKE katsoo, että tilastolaki on suoriin ja epäsuoriin tunnistetietoihin liittyvän terminologian osalta epäselvä eikä vastaa nykyisen tietosuojalainsäädännön mukaisia periaatteita tai turvaa tieteellistä vapautta. Tietosuoja-asetuksen myötä tieteellisen tutkimuksen asema ja tarpeiden huomioon ottaminen lainsäädännössä ovat selkiytyneet ja samalla vahvistuneet. Tämän tulisi heijastua myös tilastoaineistojen käyttöön tieteellisessä tutkimuksessa. Tilastokeskuksessa on suorien tunnistetietojen osalta omaksuttu liian suppea tulkinta, joka estää ja vaikeuttaa tietojen käyttöä tieteellisessä tutkimuksessa. Tulkinnan mukaan ruudun koordinaattitieto on suora tunnistetieto, jos se sisältää alle kolme tai 10 havaintoyksikköä tietolajista riippuen. SYKE katsoo, että suorien tunnistetietojen määritelmää tulisi tarkentaa lakimuutoksen yhteydessä, jolloin tilastotietojen saanti tutkimustarkoituksiin ei perustuisi lakia soveltavan viranomaisen liian kapeaan tulkintaan yksityisyyden suojan osalta.

SYKE katsoo, että tilastolaki ja sen sisältämä epätasallinen suorien tunnistetietojen määritelmä yhdessä tiukan tulkinnan kanssa kaventavat tietosuoja-asetuksen ja tietosuojalain mukaista mahdollisuutta käyttää henkilötietoja yleisen edun mukaisessa tieteellisessä tutkimuksessa. Kuten HE:ssä on todettu, perustuslakivaliokunta katsoo, että tietosuoja-asetuksen yksityiskohtainen sääntely, jota tulkitaan ja sovelletaan EU:n perusoikeuskirjassa turvattujen oikeuksien mukaisesti, muodostaa yleensä riittävän säännöspohjan myös perustuslain 10 §:ssä turvattun yksityiselämän ja henkilötietojen suojan kannalta. Valiokunta painottaa, että erityislainsäädännön tarpeellisuutta on arvioitava tietosuoja-asetuksenkin edellyttämän riskiperusteisen lähestymistavan mukaisesti. Tilastolain säännöksiä tulisikin arvioida riskiperusteisesti, mitä nyt ei olla tehty. Erityisesti julkiset tutkimusorganisaatiot ja laitokset ovat sitoutuneet hyvin tiukoihin tutkimuseettisiin periaatteisiin liittyen henkilötietojen käsittelyyn, mikä osaltaan varmistaa henkilötietojen asianmukaista käsittelyä. SYKE katsoo, että analysoitaessa ruututietoja tietoturvalisessa käsittely-ympäristössä hyväksytyyn tutkimussuunnitelman mukaan, merkittävää riskiä henkilötietojen näkökulmasta ei aiheudu.

Lopuksi SYKE toteaa, että nykyisellään tilastolaki asettaa erityyppisiä aineistoja käyttävät tutkimusotteet eriarvoiseen asemaan. Tilastolaki mahdollista hyvinkin tarkalle tasolle menevän tieteellisen tutkimuksen henkilön iän, sukupuolen, koulutuksen, ammatin ja sosioekonomisesta aseman osalta (tilastolain 19 §) esimerkiksi taloustieteen alan tutkimuksessa ja sovelluksissa kuten HGSE:n tilannehuoneessa, mutta vastaavia tietoja ei voi käyttää alueita koskevassa spatiaalisessa tutkimuksessa. Paikkatietojen osalta tilastolain nykytulkinta rajaa yhteiskunnan kannalta keskeisten kysymysten, kuten haja-asutusalueiden ja kaupunkiseutujen reuna-alueiden kehittymisen tarkastelua. Paikkatietojen avulla on mahdollista tuottaa korkealaatuista tieteellistä tutkimusta ja alueellista tilannekuvatietoa päätöksenteon tueksi, mikäli parhaat mahdolliset aineistot ovat tutkimus- ja selvitystyön hyödynnettävissä.

Tilastokeskuksen tuottamat paikkatietoaineistot, kuten 250 m x 250 m ruututiedot ja niihin linkitetyt tilastotiedot ovat keskeisessä roolissa kestävän yhdyskuntarakenteen tutkimuksessa. Nykyisen laintulkinnan mukaan tilastoruudut, joissa on alle 3 tai 10 havaintoyksikköä (muuttujasta riippuen) tulee salata ennen aineistojen luovuttamista tutkimuskäyttöön. Tämä johtaa siihen, että harvaan asutut alueet putoavat pois aineistosta. Mikäli tilastoruudun sijainti tulkittaisi epäsuoraksi tunnistetiedoksi, näitä tietoja voisi analysoida tietoturvalisessa käsittely-ympäristössä ja jatkojalostaa karkeampaan muotoon. SYKE esittää, että tällainen tulkinta tulee ottaa harkittavaksi.

Wainio-Biese Terhi
Suomen ympäristökeskus SYKE - Hallintopalvelut- yksikkö