

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?

-

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?

Tilastokeskus painottaa, että kun tietoaaineistot ovat avoimia, niiden tulee olla yhtenäisten laatukriteerien mukaisia, käyttää yhtenäisiä metatietoja ja ne tulee tarjota käyttöön yhtenäisten API-rajapintojen kautta.

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

-

Laki eräiden julkisten yritysten tietojen uudelleenkäytöstä

Kommentit laista eräiden julkisten yritysten tietojen uudelleenkäytöstä kokonaisuutena.

Tilastokeskus katsoo, että tietojen uudelleenkäyttö on periaatteena kannatettava.

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?

-

Laki julkisin varoin tuotettujen tutkimusaineistojen uudelleenkäytöstä

Kommentit laista julkisin varoin tuotettujen tutkimusaineistojen uudelleenkäytöstä kokonaisuutena.

-

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

-

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

-

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

Avoin data on yleisesti ottaen kannatettavaa ja tutkimusrahoitushauissa asetetaan usein toiveeksi, että tutkimuksessa käytetyt aineistot myös avattaisiin. On kuitenkin muistettava, että osa tutkimushankkeista käsittelee aineistoja, jotka ovat luonteelta sellaisia, että niitä ei voida julkaista esimerkiksi henkilöiden tietosuojan tai yrityssalaisuuksien takia. Tilastokeskus tarjoaa tutkijoille käyttöoikeuden suojattuihin yksikötasoisin aineistoihin tietoturvallisen etäkäyttöjärjestelmän kautta. Tutkijat eivät saa aineistoja ulos järjestelmästä. Tilastokeskus toivoo, että tietosuojasta huolehtiminen ei estä tutkijoita saamaan julkista tutkimusrahoitusta.

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

-

Arvokkaat tietoaineistot

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

Tilastokeskuksen hallussa on muun muassa paikkatietoja, tilastotietoaineistoja ja yritystietoaineistoja. Paljon näistä tiedoista on saatavilla avoimesti jo nyt taulukkomuodossa ja ilman tunnistetietoja.

Ehdotus liittyy arvokkaiisiin tietoaineistoihin ja tiheästi päivittyviin tietoaineistoihin. Tässä kohtaa on kuitenkin vielä epäselvää, mitä arvokkaat tietoaineistot ovat ja miltä osin tämä esitys koskee Tilastokeskusta. Euroopan tilastojärjestelmässä on valmisteltu ehdotus high-value data aineistoista. Ehdotusta valmistellaan eurooppalaisessa kontekstissa.

Epäselvää on tässä vaiheessa myös se, miten toimitaan, jos ”arvokkaiden tietoaineistojen” määrittely muuttuu nyt suunnitellusta. Kustannusvaikutuksien arviointi tulevaisuudessa on tästä syystä hankalaa. Lisäksi tärkeä olisi ymmärtää, millä prosessilla arvokkaat tietoaineistot tulevaisuudessa valitaan ja miten mahdolliset tulomenetykset ja avaamiskustannukset huomioidaan tällöin.

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

Toistaiseksi Tilastokeskuksen osalta on vaikea arvioida sääntelyn vaikutuksia, koska arvokkaista tietoaineistoista ei ole tämän täsmällisempää tietoa. Tälläkään hetkellä Tilastokeskus ei veloita varsinaisesti asiakkaiden tilaamista aineistoista mutta kun aineistot muodostetaan käyttäjän tarpeiden mukaan, aineistojen muodostamisesta ja käsittelystä syntyy kuluja, jotka laskutetaan aineiston tilaajalta. Tässä vaiheessa Tilastokeskus arvioi ettei nyt lausunnolla oleva avoimen datan direktiivin täytäntöönpanoa koskeva lainsäädäntö tulisi tilasto- ja paikkatietovaatimusten osalta laajentamaan Tilastokeskukselle kohdistuvia avoimen datan vaatimuksia.

Mikäli esimerkiksi tulevaisuudessa Tilastokeskuksen julkiset yritystiedot sekä aluetietopalvelun aineistot arvioidaan arvokkaiksi tietoaineistoiksi, joiden tulisi olla jatkossa maksutta ja avoimesti saatavilla, Tilastokeskuksen tulonmenetyks voi olla 2,5 milj €/ vuosi. Tilastokeskus on Sote-lakiesitykseen liittyen ilmoittanut tästä aluetietopalvelun avaamisen osuuden kustannukseksi 1,5 milj €. Kustannusarvioita tulisi tehdä tapauskohtaisesti, kun arvokkaat tietoaineistot ovat tiedossa.

Yleiset kommentit

Yleiset kommentit hallituksen esityksen luonnoksesta ja ehdotetun sääntelyn vaikutuksista kokonaisuudessaan.

Tilastokeskus kiittää mahdollisuudesta lausua. Kannatamme tiedon laajempaa avointa käyttöä ja pyrimme osaltamme julkaisemaan perustilastot julkisesti ja maksutta verkkopalvelussamme ja API-rajapinnan yli ladattavina.

Tilastokeskuksen maksuttomat tietokannat sisältävät runsaasti keskeistä perustietoa Suomesta ja maailmalta. Tilastokeskus on rakentanut valmiita rajapintoja avoimen datan käyttäjille, ja rajapintojen luominen on aloitettu StatFin-tietokannasta. Tilastokeskuksen Tutkijapalvelut tarjoaa anonymisoituja yksikkötason public use files (PUF) aineistoja myös opetuskäyttöön.

Tilastokeskus tuottaa näiden budjettirahoituksella tuotettujen perustilastojen lisäksi myös yhteiskunnan erityistarpeisiin tilastoaineistoja maksullisina toimeksiantoina. On selvää, että Tilastokeskuksen kehysrahoitteiset kustannukset kasvavat ja tulot vähenevät olennaisesti, jos nykykäytännöistä luovutaan.

Vertanen Ville
Tilastokeskus

Johnson Marianne
Tilastokeskus