

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 velloiteta 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?

Yleisesti SYKE kannattaa, että viranomaisten tietoaineistoja olisi saatavilla laajemmin avointen rajapintojen kautta. Toisaalta SYKE katsoo, että avointen rajapintojen rakentamisen tulisi perustua tarveharkintaan, jolloin myös kustannustehokkuus voitaisiin arvioida. Esimerkiksi aineistot, joiden käyttö on vähäistä, on usein kustannustehokkaampaa toimittaa aineiston pyytäjälle muulla tavoin kuin toteuttamalla avoin rajapinta. SYKE katsookin, että viranomaisten hallussa olevien aineistojen avaamista tulisi edistää myös muilla tavoin. Aineiston käytön tarpeen lisääntyessä, avoin rajapinta voitaisiin tuki rakentaa myös tällaiselle aineistolle.

Avointen rajapintojen kustannusten osalta SYKE toteaa, että SYKellä on jo käytössä teknisiä ratkaisuja, joiden avulla avointen rajapintojen rakentaminen on mahdollista eikä uusien rajapintojen rakentaminen aiheuta huomattavia kustannuksia. Merkittäviä kustannuksia voi puolestaan aiheuttaa se, mitä vaatimuksia asetetaan rajapintojen käyttövarmuudelle (vrt. INSPIRE direktiivin vaatimukset saatavuudelle, kapasiteetille ja suorituskyvyille). Lisäksi joidenkin aineistojen muokkaaminen, jotta aineistot voidaan ylipäättään avata, voi aiheuttaa merkittäviäkin kustannuksia.

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.

-

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.

Lähtökotaisesti tavoite julkisilla varoilla tuotettujen tutkimusaineistojen avoimuudesta ja sen mahdollistamasta uudelleenkäytöstä on kannatettava. SYKE huomauttaa kuitenkin, että osa tutkimusaineistoista ovat sellaisia, että niiden avaaminen ei ole mielekästä, koska aineistot eivät esimerkiksi ole uudelleenkäytettävissä muodossa tai tutkimusaineistolle ei ole muulla perusteella tarvetta uudelleenkäytölle. Toisaalta ehdotettu laki jättää harkinnan tutkimusaineistojen uudelleenkäytöstä tutkimusta harjoittaville organisaatioille, joten kaikkien tutkimusaineistojen avaamista ei ole tarkoitus säätää velvoittavaksi.

SYKE korostaa meta- ja kuvailutietojen merkitystä tutkimusaineistojen jakamisessa. Hallituksen esityksessä on todettu, että tutkimusaineistojen olisi myös oltava saatavilla koneluettavassa muodossa, jos se olisi tarkoituksenmukaista ja suoritettavissa ilman kohtuutonta vaivaa. SYKE

katsoo, että metatietojen osalta niiden saatavuus koneluettavassa muodossa voisi olla jopa velvoittavaa. Tällöin itse tutkimusaineistot olisivat löydettävissä metatietojen avulla. Tutkimusaineistojen koneluettavaan muotoon saattaminen voi puolestaan aiheuttaa merkittäviä kustannuksia, koska tutkimusaineistojen kokonaisuus on varsin heterogeeninen. Esimerkiksi SYKEssä tutkimusaineistot ovat avoimesti jaossa metatietokuvauksesta löytyvinä ladattavina tiedostoina. SYKE on arvioinut, että em. tapa on kustannustehokkain tapa avata tutkimusaineistoja.

SYKE korostaa, että kaikelle julkisesti tarjolle asetetuille tutkimusaineistoille on kuvailutietoihin ja/tai metatietoihin määriteltävä käyttöehto, esim. sopiva avoin lisenssi. Lisäksi tutkimusaineiston kuvailu- ja/tai metatiedoista tulisi käydä ilmi aineiston käyttöön mahdollisesti kohdistuvat lainsäädännön tai kolmannen oikeuksien aiheuttamat rajoitteet (kuten aineiston sisältämät salassa pidettävät tiedot tai henkilötiedot) ja se, mikä osa tutkimusaineistosta on saatavissa tutkimuskäyttöön rajoitukset huomioiden. Mikäli tutkimusaineistoon ei kohdistu immateriaalioikeuksia, tulisi myös tämä tieto kirjata kuvailu- ja/tai metatietoihin, koska tällöin käyttäjä voisi käyttää aineistoa rajoituksetta (pl. yllä mainitut lainsäädännön käytölle asetetut rajoitteet). Ilman kuvailu- ja/tai metatiedoissa kerrottua käyttöehtoa aineisto ei ole käyttökelpoista uudelleenkäyttöön edes teknisesti avoimesti saatavana, sillä käyttäjän tulee voida viitata käyttämäänsä tutkimusaineistoon ja pystyä näyttämään, että on käyttänyt tutkimusaineistoa asianmukaisesti.

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

Hallituksen esityksessä on todettu, että lakia tutkimusaineistojen uudelleenkäytöstä sovellettaisiin tutkimusaineistoihin, jotka niiden julkaisija on jo asettanut julkisesti saataville digitaalisessa muodossa. Tietyt kustantajat mahdollistavat tutkimusaineiston julkaisun kustantajan alustalla tutkimusjulkaisun yhteydessä. Tietyillä kustantajilla myös tutkimusaineistot saattavat ilmeisesti olla maksumuurin takana. Tällaista tutkimusaineistojen julkaisemista pidettänee lain mukaisena saataville asettamisena digitaalisessa muodossa. Tutkimusaineiston osalta tulisivat selvittää tai varmistaa, että tutkimusaineisto voidaan tällöin myös julkaista rinnakkain avoimella alustalla. Tutkimusaineistolla tulisi olla pysyvä tunniste, joka löytyisi myös avoimesti saatavilla olevan tutkimusaineiston metatiedosta. Ehdotetulla säädöksellä saattaisi siten olla vaikutusta kustantajan valintaan.

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

Niin kauan kuin tutkimusaineistojen uudelleenkäyttö perustuu tutkimusta harjoittavan organisaation harkintaan, SYKE ei näe ehdotetussa lainsäädännössä ongelmia. SYKE toteaa kuitenkin, että tutkimuksen mielekkään tekemisen kannalta tulee varmistua siitä, että tutkimusta harjoittavalla organisaatiolla on mahdollisuus tiettyihin embargoihin ja pitkittäisaineistojen keräämiseen ilman painetta niiden avaamisesta liian varhaisessa vaiheessa. Lisäksi SYKE korostaa, että kvantitatiivisten aineistojen avaaminen on usein helpompaa kuin laadullisten aineistojen. Siten tutkimusta harjoittaville organisaatioille tulee jättää harkintavalta sekä sen suhteen, mitä ja missä määrin tutkimusaineistoja avataan, mutta myös, miten niitä voidaan avata.

Lisäksi SYKE toteaa, että yhteiskuntatieteellisessä tutkimuksessa vahva pyrkimys tutkimusaineistojen avaamiseen voi johtaa tiettyjen tutkimusotteiden vahvistumiseen toisten kustannuksella. Tällä taas voi olla tutkimuseettisiä seurauksia, kuten evidenssin kapeutuminen, joihin tulee kiinnittää huomioita.

Tietyt rahoittajat (esim. EU, Business Finland) edellyttävät lisääntyvissä määrin myös yritysten kanssa yhteistyössä tehtävää tutkimusta, jolloin tutkimukseen liittyy usein vahva taloudellinen intressi. Usein tällaisen 'public-private' -tutkimuksen osatavoitteena on kehittää liiketoimintaa, jolloin tulosten ja tutkimusaineistojen saattaminen julkiseksi ei palvele em. tavoitetta. Suomalaisten tutkimusta harjoittavien organisaatioiden osallistuminen myös tällaiseen tutkimukseen tulee varmistaa jatkossa.

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

SYKEN näkemyksen mukaan tutkimusaineistojen maksutonta saatavuutta edellytetään rahoittajien puolesta enenevässä määrin. Varsinkin SYKEN kannalta merkittävät rahoittajat, kuten Suomen Akatemia ja EU, edellyttävät tutkimusaineistojen avaamista. SYKEN hankkeissa tuottamia tutkimusaineistoja on saatavilla maksutta esimerkiksi yhteiskuntatieteellisestä tietoarkistosta.

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?

SYKellä on useita aineistoja, jotka todennäköisesti liittyvät teemakategoriaan "maan havainnointi ja ympäristö" (earth observation and environment). Ympäristöön liittyviä paikkatietoaineistoja on jo nyt asetettu saataville mm. INSPIRE-direktiivin mukaisina WMS-katselurajapintoina ja ladattavina tietotuotteina (ATOM-syötteen kautta ladattavat GML-tiedostot). Lisäksi SYKellä on satelliittihavaintoihin perustuvia tietoaineistoja (maan havainnointi –teema), joita on käytettävissä OGC-standardin mukaisina WMS-katselurajapintoina, sekä erilaisia ympäristöön liittyviä havaintotietoja, joita on asetettu saataville oData API-rajapintoina. Kuten lausuntopyyntöissäkin todetaan, arvokkaiden tietoaineistojen teemakategorioiden sisältö on vielä määrittelemättä tarkemmin, joten arvokkaiden tietoaineistojen tunnistaminen on tässä vaiheessa haasteellista.

Suurin osa em. aineistot on asetettu saataville avoimina aineistoina, joten tiedot aineistojen uudelleenkäytöstä ovat rajalliset ja perustuvat esimerkiksi asiakaskyselyillä saataviin tietoihin. Tämän vuoksi on myös haasteellista tunnistaa arvokkaita tietoaineistoja tästä näkökulmasta: "uudelleenkäyttöön liittyy yhteiskunnallisia tai ympäristöön tai talouteen liittyviä hyötyjä, erityisesti siksi, että ne soveltuvat lisäarvopalvelujen, sovellusten sekä uusien laadukkaiden ja ihmisarvoisten

työpaikkojen luomiseen, ja näihin tietoaaineistoihin perustuvista lisäarvopalveluista ja -sovelluksista hyötyvien joukko voi olla suuri”.

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

Teknisen kehityksen kustannukset voivat olla SYKEssä kohtalaisen suuret, jos arvokkaiden tietoaaineistojen saataville asettamisen osalta tulee uusia vaatimuksia esimerkiksi INSPIRE-direktiivin toimeenpanovelvoitteiden, tai SYKEN avoimien aineistojen nykyisten jakelumuotojen lisäksi. Lisäksi jos toimeenpanosäännöksiin tullaan sisällyttämään avoimen datan harmonisointia unionin tasolla, kuten INSPIRE-direktiivissä, on kyseessä mittava ja resursseja vaativa työ. Lähtökohtana tulisikin olla, että eri direktiiveistä ei koituisi tiedontuottajille päällekkäisiä tai toisistaan poikkeavia vaatimuksia, jos samat aineistot kuuluvat eri direktiiveistä aiheutuvan sääntelyn piiriin.

Lisäksi erityisesti jatkuvatoimisten (tiheästi päivittyvien), arvokkaiksi katsottavien, mittaustietojen osalta käytännöissä ja teknisessä ympäristössä on paljon kehitettävää, joten niiden osalta kustannusvaikutuksia on odotettavissa. Tiheästi päivittyvien tietojen saataville asettamisessa tulee huomioida myös esimerkiksi laadunvarmennus, jotta voidaan varmistaa tiedon luotettavuus ja käytettävyys.

Kustannusvaikutusten arviointi on kuitenkin vaikeaa, koska tarkemmat toimeenpanosäännökset, mukaan lukien mahdolliset saatavuus- ja käytettävyysvaatimukset, puuttuvat. Suurempia kustannusvaikutuksia syntyy, jos asetetaan tiukat kriteerit rajapintapalveluiden saatavuuden, kapasiteetin ja/tai suorituskyvyn osalta, kuten INSPIRE-direktiivissä.

Yleiset kommentit

Yleiset kommentit hallituksen esityksen luonnoksesta ja ehdotetun sääntelyn vaikutuksista kokonaisuudessaan.

Erilaisten tietoaaineistojen saattaminen koneluettavaan muotoon on myös SYKEN tutkijoiden ja asiantuntijoiden etu. Kuitenkin koneluettavuuden vaatimus herättää kysymyksiä, tullaanko rajapintojen osalta antamaan tarkat tekniset ja saatavuutta koskevat määrittelyt (kuten INSPIRE-direktiivissä) vai jääkö organisaatioille vapaudet rajapintojen toteuttamiseen.

SYKE katsoo, että metatietojen rooli on keskeinen, kun tietoja asetetaan saataville. SYKEillä on oma, kaikkien SYKEN aineistojen, tietovarantojen ja palveluiden kuvaukset kokoava metatietopalvelu, josta metatietoja voidaan siirtää rajapintojen kautta erilaisiin kansallisiin/kansainvälisiin metatietoportaaleihin. Tavoitteena on, että yhdestä aineistosta/tietovarannosta/palvelusta tehdään vain yksi metatietokuvaus, joka toteuttaa keskeisimmät vaatimukset metatiedoilta. Esimerkiksi INSPIRE-direktiivin vaatimukset metadatalle, tutkimusaineistoja koskevat metadatavaatimukset ja

mahdolliset arvokkaita tietoaaineistoja koskevat metadatatvaatimukset tulisivat olla yhteensovitettavissa.

Nordqvist Lena
Suomen ympäristökeskus SYKE