

HE 74/2021 Hallituksen esitys eduskunnalle avoimen datan direktiivin täytäntöönpanoa koskevaksi lainsäädännöksi

Helsingin kaupunki suhtautuu positiivisesti datan avaamiseen ja sen sääntelyyn lailla.

Helsingin keskeiset kommentit liittyvät seuraaviin asioihin:

- Komissio ei ole määritellyt arvokkaita tietoaaineistoja ennen hallituksen esityksen antamista, mikä aiheuttaa esitykseen monin paikoin epämääräisyyttä
- Suhde henkilötietolainsäädäntöön ja sen aiheuttamat vaikutukset ja muutostarpeet
- Säädösten lisäksi on tarvetta yleisiin käytännön ohjeisiin ja linjauksiin sekä suositukseen avoimen datan lisenssistä
- Hyödyt tulevat yhteentoimivuuden kautta
- Tutkimusaineistojen määritelmällinen epäselvyys

Niin direktiivissä kuin direktiiviä toimeenpaneavassa lakiesityksessä ensisijainen tavoite on julkisen tiedon uudelleenkäytön lisääminen. Direktiivin nimi taas viittaa avoimeen dataan, mikä on ristiriidassa sen kanssa, että yleisesti ymmärretyn määritelmän mukaan avoin data on täysin maksutonta, kun taas direktiivi mahdollistaa datan osittaisen maksullisuuden. Mielestämme tulevien säädösten tavoitteissa, termistössä ja nimissä olisi hyvä yhtenäisesti kirkastaa ja selventää avoimen datan ja uudelleenkäytön edistämisen näkökulmat.

Hallituksen esityksessä ei ole määritelty siirtymäaikoja sille, milloin direktiivin määrittelemät julkiset aineistot olisi oltava saatavilla direktiivin määrittelemissä muodoissa. Ainoa esityksessä mainittu siirtymäaika koskee arvokkaiden tietoaaineistojen maksuttomuutta; ei siis sitä, koska arvokkaat tietoaaineistot tulisi olla saatavilla koneluettavassa muodossa reaaliaikaisten rajapintojen kautta tai massalatauksena. Helsingin kaupunki toivoo selkeitä siirtymäaikoja.

Arvokkaat tietoaaineistot

On ongelmallista, että direktiivin kansallista lainsäädäntöä laaditaan, vaikka tietty keskeinen lainsäädännön kohde, arvokkaat tietoaaineistot, on määrittelemättä. Helsingin kaupunki ymmärtää syyt tilanteeseen, mutta toivoo, että arvokkaiden tietoaaineistojen listan tarkennuttua asiasta olisi vielä mahdollista käydä keskustelua.

High value dataset -termin suomennos "arvokas tietoaaineisto" on vaikeasti ymmärrettävä termi, kun kyse on yleistä hyötyä tuottavasta datasta, eikä rahallisesta arvosta. Olisiko "yleishyödyllinen tietoaaineisto" kuvaavampi termi?

Huomattava on myös, että arvo syntyy toimijoiden yhtenäisestä tavasta tuottaa ja avata dataa, ei välttämättä yhden toimijan aktiivisuudesta avata dataansa. Täten "arvokkaaksi tietoaaineistoksi" voi muuttua data, jonka yksittäinen toimija ei ymmärrä olevan arvokas ennen kuin kaikki muutkin samankaltaiset toimijat avaavat datansa samalla tavalla.

Henkilötietolainsäädäntö

Henkilötietoja sisältävät aineistot on rajattu esityksen ulkopuolelle. Helsingin kaupunki haluaa tuoda esiin, että tietosuoja-asetuksen mukainen henkilötiedon määritelmä on huomattavan laaja. On todennäköistä, että kaupungin hallussa oleviin arvokkaisiin tietoaaineistoihin sekä tiheästi päivittyviin tietoihin sisältyy henkilötietoja. Henkilötietojen poistaminen aineistosta vaatii hallinnollista työtä ja saattaa aiheuttaa odottamattomia kustannusvaikutuksia.

Tulisi arvioida, kuinka paljon esitykseen kuuluva aineisto todellisuudessa sisältää henkilötietoja. Henkilötietojen rajaaminen kokonaan soveltamisalan ulkopuolelle voi nimittäin johtaa epätarkoituksenmukaisiin lopputuloksiin.

Tarve yleisiin käytännön ohjeisiin ja suositukseen lisenssistä

Direktiivin positiiviset vaikutukset yksittäisille organisaatioille tulevat myös sen kautta, että avaamista tehdään tietosisällöltään yhteneväisellä mallilla ja velvoittavuus ulotetaan laajalti samankaltaisiin toimijoihin. Yhteentoimivan avoimen datan tuottaminen tuo etuja kaikille datan avaajille ja hyödyntäjille mm. palvelujen skaalattavuuden ja datan käytettävyyden kautta.

Tarvetta on yhteisille kansallisille ohjeille ja linjauksille siitä, millä tarkkuudella dataa avataan niin, ettei tietosuoja vaarannu. Tällä hetkellä esimerkiksi toiminimiä ja yritysten nimiin sisältyviä henkilötietoja (tavallisimmin nimi) koskien on kyseenalaista, voidaanko niitä koskevia tietoja sisällyttää sellaisenaan avoimena datana julkaistuihin aineistoihin (esim. voiko ostolaskudataan sisällyttää toiminimiä).

Koko sääntelyn kannalta oleellista olisi ohjata datan avaamista määrittelemällä tiedon yleistystasoa, tilastointiyksikköjen kokoa ja tiedon aggregoinnilla säädeltävissä olevaa tietosuojaa.

Lisäksi on tarve suositukselle käytettävistä lisensseistä (vrt. vanha JHS 189 Avoimen tietoaaineiston käyttö lupa) sekä muille datan avaamisen ehdoille ja käytännöille. Hallituksen esitys ei ota tähän kantaa.

Hyödyt yhteentoimivuuden kautta

Avoimen datan tuottamat hyödyt tulevat erityisesti yhteentoimivuuden (niin datan sisällön ja rakenteen kuin kuvailu- eli metatietojenkin) kautta, ei välttämättä yksittäisten organisaation avaamien datojen kautta. Myös yhteneväisiin metatietoihin kuin myös datan löydettävyyteen olisi hyvä tehdä kansallisia linjauksia.

Helsingin kaupunki kannattaa avoimien standardoitujen rajapintojen käyttöä. Kun organisaatio käyttää itsekin omia avoimia rajapintojaan, voidaan varmistaa paremmin datan ajantasaisuus ja laatu sekä rajapinnan toimivuus ja käytettävyys.

Tutkimusaineistot

Yleisesti ottaen on hyvä tavoite saattaa julkisesti tuotetut tutkimusaineistot avoimeen käyttöön eli saada aineistoista mahdollisimman laaja ja monipuolinen hyöty. Hieman epäselväksi kuitenkin jää, mikä kaikki on tutkimusaineistoa. HE:ssä todetaan: *"Tutkimustiedolla tarkoitetaan sähköisessä muodossa olevia asiakirjoja, jotka eivät ole tieteellisiä julkaisuja tai artikkeleita, mutta jotka on koottu tai tuotettu tieteellisten tutkimustoimien yhteydessä ja joita käytetään näyttönä tutkimusprosessissa, tai joita pidetään tutkimusyhteisössä yleisesti tarpeellisina tutkimustulosten tarkistamiseksi. Selkeyden vuoksi direktiivissä tarkoitettua tutkimustietoa kutsutaan tässä esityksessä tutkimusaineistoksi."*

Jos siis kyse on vain **tieteellistä** tutkimusta varten kootuista tutkimusaineistoista, ei kyseinen kirjaus koske kaupungin omissa suunnitteluprosesseissa syntyneitä dataa, jota ei ole kerätty tieteellistä tutkimusta varten – mutta olisi sellaiseen mahdollista käyttää. Tämä tulkinta jää siis osin epäselväksi.

Onko HE:n tarkoittamaa tutkimusaineistoa myös data, joka kertyy esimerkiksi erilaisista kyselyistä, joita kaupungin toimialat suorittavat palveluidensa parantamiseksi? Kyselyiden tarkoituksena ei useinkaan ole sinänsä tutkimus, mutta tutkijat voivat niistä kiinnostua. Esimerkkinä karttakyselytyökalu, jonka kautta saadaan kommentteja alueellisista suunnitteluhankkeista. Entä onko HE:n tarkoittamaan tutkimusaineistoa myös data, jota kertyy kaupungin ostamien konsulttiselvitysten, -raporttien ja -tutkimusten yhteydessä?

Hieman epäselväksi jää myös se, mihin tämä lakiehdotus käytännössä velvoittaa kuntia tutkimusaineistojen osalta. HE:ssä todetaan: *"Esityksessä ehdotetaan säädettäväksi uusi laki julkisin varoin tuotettujen"*

tutkimusaineistojen uudelleenkäytöstä. Sääntelyllä saatettaisiin kansallisesti voimaan ne velvoitteet, jotka direktiivissä kohdistuvat julkisin varoin tuotettuihin tutkimusaineistoihin. Lakia sovellettaisiin tutkimusaineistoihin, jotka niiden julkaisija on jo asettanut julkisesti saataville digitaalisessa muodossa."

Yllä oleva kirjaus tarkoittaa, että laki koskee vain sellaisia aineistoja, jotka on jo valmiiksi jaettu esim. pääkaupunkiseudun avoimen datan palvelun Helsinki Region Infosharen (hri.fi) kautta. Tähän Helsingin kaupungilla on jo valmiiksi toimivat käytännöt. Toisiko tämä lakiesitys niihin jotain muutosta?

Velvoittavuudessa huomioitavaa on myös se, että osa kaupungilla käytettävistä tutkimusaineistoista tulee järjestelmistä, jotka eivät ole kaupungin omistajuudessa. Tämä korostuu esim. 1.3.2021 alkaen kuntakokeiluissa.

Lisätietoja

Hami Kekkonen

Projektipäällikkö

Helsingin kaupunginkanslia

puh: 040-3344 780

hami.kekkonen@hel.fi

Kaisa Voipio

Suunnittelija

Helsingin kaupunginkanslia

puh: 040-3580 647

kaisa.voipio@hel.fi