



Lausunto

23.11.2020

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?**

Lakiesityksessä säädetään mm., että arvokkaiden tietoaaineistojen (joita tässä vaiheessa vain sivutaan teemoittain) pitää olla saatavilla maksutta koneluettavasti ja rajapinnan kautta. Vaikutukset toimijoille tiedetään kuitenkin paremmin siinä vaiheessa, kun komissio tarkentaa arvokkaan tietoaaineiston määritelmää täytäntöönpanosäädöksissä. Lain soveltamisala koskee ainakin Helsingin seudun liikennettä (HSL), Helsingin energiaa (Helen) ja Helsingin seudun ympäristöpalveluja (HSY). Koska toimijat ovat jo avanneet tietoaan, muutosten vaikutusten arvioidaan olevan pieniä. Täyttä varmuutta tästä ei kuitenkaan voi olla ennen komission tarkennuksia.

Luonnoksessa hallituksen esitykseksi todetaan, että julkisista asiakirjoista voi joissakin tapauksissa olla mahdollista niitä yhdistelemällä luoda tai saada selville tietoja, joiden sisältöä viranomaisen ei alkuperäisiä asiakirjoja luovuttaessaan ole voinut tietää tai arvioida. Tarvetta on saada yhteisiä ohjeita ja linjauksia siitä, millä tarkkuudella dataa avataan niin, ettei tietosuojaa vaaraudu.

Ehdotetussa tiedonhallintalain 24 a §:ssä säädettäisiin tiheästi päivittyvän tiedon saatavuudesta. Esityksen mukaan tiedon on lähtökohtaisesti pyynnöstä oltava saatavilla teknisten rajapintojen avulla heti tiedon keräämisen jälkeen ja tarvittaessa useana kerralla ladattavana tiedostona. Esityksestä ei kuitenkaan käy ilmi, onko tiheästi päivittyvän tiedon suhdetta julkisuuslain mukaiseen sääntelyyn viranomaisen asiakirjan julkiseksi tulemisesta arvioitu. Esityksestä ei käy ilmi, onko tiheästi päivittävissä tiedossa kyse aina sellaisista asiakirjoista, jotka tulevat julkiseksi heti, kun ne on kerätty. Vai onko mahdollista, että asiakirja tulee vasta myöhemmin julkiseksi, jolloin sen on oltava saatavilla teknisten rajapintojen avulla heti sen tultua julkiseksi. Tällöin laissa tulisi todeta, että tiheästi päivittyvän tiedon on pyynnöstä oltava saatavilla teknisten rajapintojen avulla heti tiedon keräämisen jälkeen tai heti sen tultua viranomaisten toiminnan julkisuudesta annetun lain mukaan julkiseksi.

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

Elinkeinotoimintaan liittyen, tietoriskejä olisi pitänyt käsitellä erityisesti kaupparekisterin osalta. Kaupparekisterin tietojen avaaminen maksutta avoimen rajapinnan kautta voi lisätä uudenlaisia yrityspalveluja, mutta myös yrityksiin kohdistuvia kansainvälisiä tietoriskejä ja petoksia.

**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?**

Kannatamme avoimien rajapintojen julkaisua ja käyttöä aina, kun niiden kautta saatavissa oleva data on julkista. Kun käytetään itsekin omia avoimia rajapintoja, voidaan varmistaa paremmin datan ajantasaisuus ja laatu sekä rajapinnan toimivuus ja käytettävyys. Helsingin kaupunki käyttää jo useampia kaikille avaamiaan rajapintoja myös omassa käytössään.

Mikäli avoimella rajapinnalla tässä yhteydessä tarkoitetaan pääsyä aina luotettavan ja ajantasaisen datan äärelle, datan dokumentoimista ja avoimuuden ehdoin saatavilla olemista, on väite puollettavissa. Kaupungilla on kokemusta, että avointen, mieluiten standardoitujen rajapintojen ylläpito yhtäältä helpottaa datan tuottajien työtä ja toisaalta auttaa datan löytymisessä ja hyödyntämisessä. Mikäli tietoisuus avoimista rajapinnoista lisääntyisi entisestään ja mikäli rajapintoja avattaisiin järjestelmällisesti em. ehdoin, yleistyisi niiden käyttö todennäköisesti myös viranomaisten keskuudessa. Tämä puolestaan lisäisi yhteistyön edellytyksiä, loisi mahdollisuuksia kehittää datan yhteentoimivuutta ja standardointia ja parempia yhteisiä julkisen hallinnon digitaalisia palveluja.

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

1. Liikennepalvelulailla henkilökuljetuspalveluihin liittyviä nk. olennaisia tietoja (reitti-, pysäkki-, aikataulu-, hinta- ja saatavuustiedot sekä esteettömyystiedot) on velvoitettu avaamaan. Liikennepalvelulaissa ei kuitenkaan ole säädetty tietojen maksullisuudesta tai maksuttomuudesta, jota tässä myös käsitellään. Mahdollisesti pieniä muutoksia jouduttaisiin tekemään lakiin tämän vuoksi.

2. Lisäys lakiehdotukseen viranomaisten toiminnan julkisuudesta annetun lain 16 §:n 3 momenttiin

Ehdotus: Viranomaisen henkilörekisteristä saa antaa henkilötietoja sisältävän kopion tai tulosteen tai sen tiedot sähköisessä muodossa, jollei laissa ole toisin erikseen säädetty, jos luovutuksensaajalla on henkilötietojen suoja koskevien säännösten mukaan oikeus tallettaa ja käyttää sellaisia

henkilötietoja. Henkilötietoja saa kuitenkin luovuttaa suoramarkkinointia ja mielipide- tai markkinatutkimusta varten vain, jos niin erikseen säädetään tai jos rekisteröity on antanut siihen suostumuksensa. Viranomaisen henkilörekisteriin ja viranomaisen asiakirjaan sisältyvät täysin julkiset henkilötiedot voidaan pitää julkisesti saatavilla tietoverkossa, jos ne on pseudonymisoitu siten, että suora yhdistettävyyden henkilöön on poistettu.

Perustelut: Luonnoksessa hallituksen esitykseksi on todettu, että esitys ei koske asiakirjojen salassapitoa eikä se vaikuta viranomaisten henkilötietojen käsittelyyn. Käytännössä on kuitenkin tunnistettu, että useimmissa tietoa-aineistoissa on mukana henkilötietoja. Tällaisten aineistojen tilastoksi muuntaminen ja aggregointi sellaiselle tasolle, että henkilöt eivät varmasti ole tunnistettavissa vaatii tapauskohtaista arviointia ja on työlästä. Erityisesti tapauksissa, joissa kysymys on täysin julkisista tiedoista, on henkilön tunnistaminen usein helppoa. Tämä johtaa siihen, että aineistoja ei julkaista lainkaan tai tietoja voidaan julkaista vain muokattuina. Silloin kun tiedot ovat julkisina tietoina kenen tahansa saatavissa, pitäisi ne voida julkaista suoraan pseudonymisoidussa muodossa siten, että tieto voidaan julkaista vaikka otanta koskee vain muutaman henkilön tietoja.

### 3. Ehdotus tietosuojalain 4 §:n 1 momentin 1 kohdan muuttamisesta:

Henkilötietoja saa käsitellä tietosuoja-asetuksen 6 artiklan 1 kohdan e alakohdan mukaisesti, jos:

1) kysymys on henkilön asemaa, tehtäviä sekä niiden hoitoa julkisyhteisössä, elinkeinoelämässä, järjestötoiminnassa tai muussa vastaavassa toiminnassa kuvaavista tiedoista siltä osin kuin käsittelyn tavoite on yleisen edun mukainen ja käsittely on oikeasuhtaista sillä tavoiteltuun oikeutettuun päämäärään nähden. Tällaiset henkilötiedot saadaan julkaista yleisessä tietoverkossa, jos on ilmeistä, että se ei vaaranna perustuslain 10 §:ssä turvattua yksityiselämän suojaa.

Perustelut: Vaikka luonnoksessa hallituksen esitykseksi on rajattu henkilötietojen suoja esityksen ulkopuolelle, koska se ei ole direktiivin soveltamisalassa olisi samassa yhteydessä kuitenkin tarpeen käyttää tietosuoja-asetuksen 86 artiklan mukaista kansallista liikkumavaraa. Tietosuojalain 4 §:n 1 momentin 1 kohta on osoittautunut haastavaksi, koska tietosuojalain esitöissä on todettu, ettei se oikeuta ko. henkilötietojen julkisesti saatavilla pitämiseen yleisessä tietoverkossa. Tämä tiukensi tulkintaa henkilötietolain aikaisesta saman sisältöisestä säännöksestä. Tämä on asettanut rajoitteita esim. sellaisten asiakirjojen ja aineistojen pitämiseksi yleisesti saatavina tietoverkoissa, joihin sisältyvät henkilötiedot ovat täysin julkisia ja niihin ei sisälly yksityisyyden suojan intressiä. Erityisesti silloin, jos tiedot sisältyvät toisen viranomaisen julkiseen henkilörekisteriin, joka on julkaistu yleisessä tietoverkossa, kuten ytj:ssä, ei suojalle ole tarvetta. Tällä hetkellä esimerkiksi toiminimiä ja yritysten nimiin sisältyviä henkilötietoja (tavallisimmin nimi) koskien on kyseenalaista voidaanko niitä koskevia tietoja sisällyttää sellaisenaan avoimena datana julkaistuihin aineistoihin.

## Laki eräiden julkisten yritysten tietojen uudelleenkäytöstä

## **Kommentit laista eräiden julkisten yritysten tietojen uudelleenkäytöstä kokonaisuutena.**

Lakiesitys olisi pitänyt laatia vasta siinä vaiheessa, kun komissio olisi tarkentanut arvokkaan tietoaineiston määritelmää.

### **Millaisiin yrityksiin toimialallanne tai alueellanne ehdotettu laki tulisi sovellettavaksi?**

Lain soveltamisala koskee ainakin Helsingin seudun liikennettä (HSL), Helsingin energiaa (Helen) ja Helsingin seudun ympäristöpalveluja (HSY) ja mahdollisesti myös Helsingin satamaa.

### **Tunnistattekko julkisia yrityksiä koskevaa sääntelyä, jota olisi ehdotetun lain johdosta muutettava tai joka on sen kanssa ristiriidassa?**

Helsingin kaupungilla ei ole tiedossa tällaista sääntelyä.

### **Onko tiedossanne sääntelyä, joka velvoittaa ehdotetun lain soveltamisalassa olevia julkisia yrityksiä antamaan niiden hallussa olevia asiakirjoja uudelleenkäytettäväksi?**

Kaupungilla on velvollisuus tuottaa liikennejärjestelmätietoja rakenteisessa muodossa valtakunnalliseen tietovarantoon (Digiroad). Lakiesityksessä on jo todettu nykyisen liikennepalvelulain velvoitteet.

Tähän liittyen, onko Traffic Management Finland direktiivin soveltamisalassa erityistehtävänä? Liikennepalvelulain muutoksella (parhaillaan eduskunnan käsittelyssä) sille on ehdotettu siirrettäväksi liikkumispalveluihin liittyviä olennaisia tietoja (National Access Point). Jos komissio määrittäisi tiedot arvokkaiksi tietoaineistoiksi, pitäisi tiedot olla saatavissa maksutta. Tästäkin syystä on tärkeää, että tietojen maksuttomuudesta pidetään kiinni jatkossakin.

### **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?**

Ks. edellinen vastaus. Liikennepalvelulain tavoitteena on myös synnyttää kaupallisia palveluja, joilla edistetään esim. matkaketjujen tehostumista.

### **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?**

Tätä on osin vaikea arvioida, kun arvokkaita tietoaineistoja ei ole vielä täsmennetty. Ainakin turvallisuusnäkökulmaa pitäisi käsitellä enemmän, jotta tietojen väärinkäyttöä pystytään estämään. Tarvitaan esim. yhteisiä ohjeita ja linjauksia siitä, millä tarkkuudella dataa avataan niin, ettei tietosuoja vaaraudu.

### **Onko asiakirjojen uudelleenkäytön salliminen ominaista tai olennaista jonkin ehdotetun lain soveltamisalassa olevan julkisen yrityksen liiketoiminnalle?**

Helsingissä avoimen datan periaatetta ja datan avaamista on pyritty edistämään suunnitelmallisesti. Myös mm. Helsingin seudun liikenne ja Helsingin seudun ympäristöpalvelut ovat jo avanneet paljon julkista dataansa, pääosin rajapintojen kautta.

## Laki julkisin varoin tuotettujen tutkimusaineistojen uudelleenkäytöstä

### **Kommentit laista julkisin varoin tuotettujen tutkimusaineistojen uudelleenkäytöstä kokonaisuutena.**

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 syntynyttä dataa, jota ei ole kerätty tieteellistä tutkimusta varten – mutta olisi sellaiseen mahdollista käyttää. Tämä tulkinta jää siis osin epäselväksi.

Onko 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.

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?

Lakiesitys ei siis ilmeisesti koske kaikkia aineistoja, sellaisia, joita ei ole julkisesti jaettu. Tulisiko tämän lakiesityksen myötä lisävelvoitetta tutkimusaineistojen (esim. kyselyt) jakamiseen jatkossa entistä useammin? Jos ei, vaikutukset ovat nähdäkseni varsin pieniä.

Velvoittavuudessa huomioitavaa on myös se, että osa kaupungilla käytettävistä tutkimusaineistoista tulee järjestelmistä, jotka eivät ole kaupungin omistajuudessa. Tämä korostuu erityisesti 1.1. alkaen kuntakoikeiluissa.

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

Jos on kyse vain sellaisista aineistoista, jotka on jo valmiiksi jaettu julkisesti, ei tällä ehdotuksella nähdäksemme ole kovin suuria vaikutuksia. Julkisesti jaetut tutkimusaineistot ovat pääsääntöisesti jo valmiiksi maksuttomia.

Julkisesti saataville asetettujen aineistojen mahdollinen muokkaus ja metatietojen täydentäminen yms. voi toki aiheuttaa lisätyötä. HE:ssä todetaan: "Tutkimusaineisto olisi asetettava saataville kaikissa jo olemassa olevissa tiedostomuodoissa ja kaikkina jo olemassa olevina kielitoisintoina. Tutkimusaineisto olisi lisäksi asetettava saataville sähköisessä ja yleisesti käytettävässä koneluettavassa muodossa kuvailutietoineen, mikäli tämä on tarkoituksenmukaista ja mahdollista ilman kohtuutonta vaivaa."

Kysymyksenä herää, miten tämä laki vaikuttaisi Tilastokeskuksen ja THL:n toteuttamiin ja keräämiin tutkimusaineistoihin, vai vaikuttaisiko lainkaan.

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

Tutkimusaineistojen, kuten esimerkiksi kyselyaineistojen ja haastatteluaineistojen, julkiseksi ja avoimeksi asettamisessa keskeinen kysymys on henkilötietojen suojaaminen. Tätä säädellään jo tarkasti. Lakiesitys ei ilmeisesti koske sellaisia aineistoja, joiden julkiseksi asettamista rajoittaa henkilötietojen suojaamiseksi annetut säädökset.

Kun kyselyaineistoja tai haastatteluaineistoja kerätään, vastaajalle on kerrottava, mihin aineistoa käytetään. Voiko kyselyihin vastaamisen ja haastatteluihin osallistumisen into vähetä entisestään, jos saatekirjeessä olisi jatkossa todettava, että aineisto laitetaan julkisesti saataville (vaikka anonymisoitunakin) ja että se on tämän seurauksena käytettävissä myös kaupallisiin tarkoituksiin?

Mikäli lain edellyttämät vaatimukset julkisesti jaetun tutkimustiedon osalta ovat liian "raskaita" toteutettavaksi, voi tämä vähentää intoa jakaa aineistoja julkisesti. Lakiesityksessä kuitenkin todetaan, että vaaditut asiat tulisi toteuttaa mikäli "mahdollista ilman kohtuutonta vaivaa". Mitä tämä konkreettisesti tarkoittaa?

Osa kaupungin omaan tarkoitukseen keräämistä tilasto- ja tutkimusaineistoista on luonteeltaan hyvin sensitiivisiä. Tällaisia aineistoja ovat esim. vieraskielisten työttömyystilastot tai

sosiaalipalvelujen käyttötilastot. Anonymisoitujenkin sensitiivisten aineistojen avaamisessa tulee olla hyvin varovainen.

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

Tutkimusaineistojen avointa ja maksutonta saatavuutta on edistetty jo monin tavoin. Esimerkiksi Suomen akatemian verkkosivuilla todetaan: "Akademia edellyttää rahoittamiltaan hankkeilta sitoutumista tutkimustuloksistaan kertovien tiedejulkaisujen avoimeen saatavuuteen sekä hankkeen tuottamien tutkimusaineistojen avaamiseen. Aineistoilla voi kuitenkin perustelluista syistä olla erilaisia avoimuuden asteita vaihdellen kaikille täysin avoimesta salassa pidettävään."

Myös Tampereen yliopiston ylläpitämään tietovarastoon on jo pitkään kerätty erityisesti yhteiskuntatieteellisten tutkimusten tutkimusaineistoja avoimeen käyttöön.

Kuten HE:ssä todetaan: "Avoimuus on yksi tieteen peruseriaatteista. Suomessa Tieteellisten seurain valtuuskunta koordinoi tieteen avoimuutta ja koordinaatiota ohjaa Avoimen tieteen ja tutkimuksen kansallinen ohjausryhmä, joka valmistelee parhaillaan muun muassa tutkimusaineistojen ja –menetelmien avoimen saatavuuden linjausta sekä tutkimusdatan avoimen saatavuuden osalinjausta. Ehdotettu laki tukee jo käynnissä olevia toimia tieteen avoimuuden edistämiseksi, mutta esityksen vaikutuksia ei voida pitää erityisen merkittävänä, koska avoimuuden periaatteet ovat jo yleisesti omaksuttuja suomalaisessa tutkimusyhteisössä ja ehdotettu sääntely koskee ainoastaan sitä, millä tavoin jo julkaistu aineisto olisi asetettava saataville. "

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

Kuten jo aiemmin nostimme esiin, on tärkeä selkiyttää, miten tämä laki suhtautuu tietosuojalakiin ja tietosuoja-asetukseen ja siinä määriteltyihin henkilötietoja koskeviin säädöksiin.

## **Arvokkaat tietoaineistot**

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

Helsingin kaupungilla on paljon mm. tilastoja ja paikkatietoja sekä ympäristöön ja liikkumiseen liittyvää dataa avoimena datana avoimien rajapintojen kautta sekä tiedostoina (HRI.fi).

"Arvokas tietoaineisto" on hieman vaikeasti ymmärrettävä termi, kun kyse on yleistä hyötyä tuottavasta datasta, eikä rahallisesta arvosta. Olisiko "yleishyödyllinen tietoaineisto" kuvaavampi termi? Huomattava on myös, että arvo syntyy toimijoiden yhtenäisestä tavasta tuottaa ja avata dataa, ei välttämättä yhden toimijan aktiivisuudesta avata dataansa. Tämä on huomattu mm. pääkaupunkiseudun datojen avaamisessa yhtenäisellä tavalla. Täten "arvokkaaksi tietoaineistoksi" voi muuttua data, jonka yksittäinen toimija ei ymmärrä olevan arvokas ennen kuin kaikki muutkin samankaltaiset toimijat avaavat datansa samalla tavalla.

Arvokkaiksi tietoaaineistoiksi voidaan kaupungissa tunnistaa direktiivin mainitsemista paikkatietoaaineistoista kaikki kunnan/kaupungin tuottamat perustiedot, kuten mm. osoitteiden ja paikannimien maantieteelliset sijainnit, aluejakojen rajat, maaston peruskartoitustiedot, rakennetun ympäristön keskeiset elementit (kiinteistöt, rakennukset, liikenteen infratiedot jne), joita kaupungin alueelta ei viranomaistoimintana tuota mikään muu taho. Jokaisen julkishallinnon toimijan toimipiste- ja palvelutiedot saavutettavuus- ja aukioloaikatietoineen tulisi olla avointa dataa.

Liikenteenohjausjärjestelmän tiedot ovat tyypillisiä "arvokkaita tietoaaineistoja". Mm. liikennemerkkien sijainnit, väylätiedot ja poikkeustilanteen liikenteenohjaus ovat tyypillisiä kaupungin yleiseksi hyödyksi tuottamia tietoja, joita tarvittaisiin avoimina, koneluettavina tietoaaineistoina (tästä on jo erityislaki).

Samoin erilaiset tilastot ovat yleisesti tunnistettuja "arvokkaita tietoaaineistoja". Niiden avaamisesta olisi helpointa aloittaa, koska niitä tuotetaan jo runsaasti. Nämä tietoaaineistot ovat arvokkaita koneluettavina ja rakenteellisina avoimina aineistoina, kun niistä saa muodostettua valtakunnallisia yhdistelmiä.

Reaaliaikaisista, jatkuvasti päivittyvistä aineistoista kaupungissa on tunnistettu esim. liikenteen tilannetiedot (ruuhkat, onnettomuudet, katutyöt jne.), joukkoliikenteen reaaliaikaiset tiedot sekä energian-, veden-, sähkön- ym. kulutustiedot. Myös ilman- ja vedenlaatu- sekä jätemäärätiedot on tunnistettu arvokkaiksi ja yleisesti kiinnostaviksi mm. kestävän kehityksen ohjaamiseen vaadituiksi tiedoiksi.

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

Helsingin kaupungilla on valmius lisätä datan avaamista ja tehdä sitä myös pääkaupunkiseudun laajuisesti Helsinki Region Infoshare -toimintamallin ja -palvelun avulla. Sikäli kuin arvokkaiksi tietoaaineistoiksi määritellään joitakin sellaisia dataa, joita ei vielä tuoteta, avaamiseen on vain varauduttu HRI-palvelun lisäksi yleisellä tasolla rakentamalla kaupunkiyhteisiä tietoallasratkaisuja (esim. sensori- ja IoT-datojen keräämiseen) ja määrittelemällä muun muassa kaupunkiyhteiset API-linjaukset. Lisäksi kaupunki on pystyttämässä kaupunkiyhteistä API-hallintaa, joka auttaa hallinnoimaan reaaliaikaisen datan ulkopuolista käyttöä.

Datan avaamiseen tulisi varautua ja Helsingissä varaudutaankin jo tietojärjestelmien hankintavaiheessa (sisällyttämällä avaamisen ja avointen rajapintojen vaatimus sopimukseen) ja laitteistojen kustannuksista sopimisella jo varhaisessa vaiheessa (esim. GPS-laitteistot joukkoliikennevälineissä vaativat ylläpitoa). Samoin, datan avaamiseen tulisi jatkossa varautua jo sen keräämisen yhteydessä (esim. kyselyt, tutkimukset ja palautteet).



Valtakunnallisten toimijoiden datantuotanto nojaa usein kuntien dataan, muttei täysimääräisesti hyödytä kuntia toivotulla tavalla. Esimerkiksi tilastojen tuotannossa kunnilla on rooli datan tuottajana, mutta valtakunnallisesti yhtenäisten datojen käyttöehdoista ja avaamisesta ei ole selkeyttä.

Olisi tärkeää tehdä tämä arviointi siinä vaiheessa, kun käsite "arvokkaat tietoaaineistot" täsmentyy. Euro- tai lukumääräarvioita on vaikea antaa, kun arvokkaita tietoaaineistoja ei ole vielä täsmennetty.

## Yleiset kommentit

### **Yleiset kommentit hallituksen esityksen luonnoksesta ja ehdotetun sääntelyn vaikutuksista kokonaisuudessaan.**

Helsingin kaupunki suhtautuu positiivisesti datan avaamiseen ja sen sääntelyyn lailla.

Helsingin keskeiset kommentit liittyvät seuraaviin asioihin:

- Arvokkaita tietoaaineistoja ei ole määritelty ennen lausuntopyyntöä, mikä vaikeutti monissa kohden lausunnon antamista
- Suhde henkilötietolainsäädäntöön ja sen aiheuttamat vaikutukset ja muutostarpeet
- Säädösten lisäksi on tarvetta yleisiin käytännön ohjeisiin (esim. aggregointi ja pseudonymisointi) ja linjauksiin sekä suositukseen avoimen datan lisenssistä
- Hyödyt tulevat yhteentoimivuuden (niin datan sisällön ja rakenteen kuin metatietojenkin) kautta, ei välttämättä yksittäisten organisaation avaamien datojen kautta
- Tilasto- ja tutkimusaineistojen määritelmällinen epäselvyys

On ongelmallista, että direktiivin kansallista lainsäädäntöä laaditaan ja siitä pyydetään lausuntopyyntöjä, vaikka tietty keskeinen lainsäädännön kohde, tässä tapauksessa arvokkaat tietoaaineistot, on määrittelmättä. Lisäksi komissio voi delegointivaltuuksien mukaisesti lisätä tähän listaukseen uusia teemakokonaisuuksia ilman, että selvitetään mitä kustannusvaikutuksia tästä voi koitua jäsenvaltioille. Lakiluonnokseen on kirjattu vaatimus, että julkisen, ei-henkilötietoja sisältävät arvokkaan tietoaaineiston on oltava maksutta saatavilla koneluettavassa muodossa ohjelmointirajapinnan kautta. Tällä hetkellä tiedossa olevan arvokkaan tietoaaineiston luettelo (paikkatiedot, maan havainnointi ja ympäristö, säätiedot, yritys- ja yritysten omistustiedot sekä liikkuvuustiedot) ei ole riittävän täsmällinen. Toivomme, että esitystä olisi mahdollista kommentoida vielä uudemman kerran, kun arvokkaiden tietoaaineistojen lista on täsmennytyt.

Henkilötietoja sisältävät aineistot on rajattu esityksen ulkopuolelle. Helsingin kaupunki haluaa tuoda esiin, että tietosuoja-asetuksen mukainen henkilötiedon määritelmä on huomattavan laaja. On erittäin todennäköistä, että kaupungin hallussa olevaan arvokkaaseen tietoaaineistoon sekä tiheästi päivittyviin tietoihin sisältyy henkilötietoja. Henkilötietojen poistaminen aineistosta vaatii

hallinnollista työtä ja saattaa aiheuttaa odottamattomia kustannusvaikutuksia. Jatkovalmistelussa tulisi arvioida, kuinka paljon esitykseen kuuluva aineisto todellisuudessa sisältää henkilötietoja. Henkilötietojen rajaaminen kokonaan soveltamisalan ulkopuolelle voi nimittäin johtaa epätarkoituksenmukaisiin lopputuloksiin.

Tarvetta on yhteisille ohjeille ja linjauksille siitä, millä tarkkuudella dataa avataan niin, ettei tietosuoja vaarannu. Asettaako avoimen datan direktiivi rajoituksia tai suosituksia käytettävälle lisenssille (vrt. JHS 189 Avoimen tietoaineiston käyttöluja) tai muille datan avaamisen ehdoille ja käytännöille? Tai ollaanko ohjeistuksia laatimassa Suomessa?

Koko sääntelyn kannalta oleellista olisi ohjata datan avaamista määrittelemällä tiedon yleistystasoa, tilastointiyksikköjen kokoa ja tiedon aggregoinnilla säädeltävissä olevaa tietosuojaa, mieluummin kuin puhumalla yleisesti "arvokkaista tietoaineistoista", joihin oletettavasti kohdistuu suojaamista siksi, ettei yleistämistä, aggregointia ja anonymisointia ole ohjattu tekemään yhtenäisellä tavalla.

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, termistöissä ja nimissä olisi hyvä yhtenäisesti kirkastaa ja selventää avoimen datan ja uudelleenkäytön edistämisen näkökulmat.

Helsingin elinkeinoelämään lakiesityksellä on pääsääntöisesti myönteisiä vaikutuksia, sillä esityksellä edistetään julkisen tietoaineiston uudelleenkäyttöä, mikä voi synnyttää uusia innovaatioita ja uutta liiketoimintaa. 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 avaajille, mutta myös sen hyödyntäjille mm. palvelujen skaalattavuuden ja datan käytettävyyden kautta.

Karhu Kimmo  
Helsingin kaupunki - Strategiaosasto