



CSC - Tieteen tietotekniikan keskus Oy

Lausunto selvityksestä muiden kuin henkilötietojen vapaan liikkuvuuden esteistä Suomessa

CSC - Tieteen tietotekniikan keskus Oy (CSC) kiittää mahdollisuudesta lausua selvityksestä muiden kuin henkilötietojen vapaan liikkuvuuden esteistä Suomessa. CSC on suomalainen ICT-osaamiskeskus, joka ylläpitää opetus- ja kulttuuriministeriön toimeksiannosta valtakunnallista keskitettyä tietotekniikkainfrastruktuuria ja tarjoaa sen avulla kansallisia tietotekniikkapalveluita tutkimuksen, koulutuksen, kulttuurin ja julkishallinnon tarpeisiin.

Data ja sen vapaa liikkuvuus ovat keskeisiä tulevaisuuden kilpailukykytekijöitä ja innovaatioiden edistäjiä. Siksi on erittäin tärkeää pyrkiä luomaan suotuisa toimintaympäristö ja EU:n laajuinen yhtenäinen ja tarkoituksenmukainen lainsäädännöllinen pohja datan käytölle ja liikkuvuudelle. CSC pitää olemassa olevan sääntelyn kriittistä tarkastelua ja tarvittaessa purkamista aiheellisena ja tukee täten asetuksen vaatimusta siitä, että jäsenvaltioiden on poistettava perusteettomat sijaintia koskevat vaatimukset sähköisessä muodossa olevalle muulle kuin henkilötiedolle.

Liikenne- ja viestintäministeriön selvityksen perusteella Suomessa on voimassa vain kolme lakia, jotka rajoittavat muun kuin henkilötietojen vapaata liikkuvuutta. Kaikki ministeriöt eivät tosin ole osallistuneet selvityksen laatimiseen, joten on tärkeää, että kaikki muut mahdolliset lainsäädännölliset esteet eri sektoreilla selvitetään.

Tilanne näyttää joka tapauksessa olevan Suomessa selvästi parempi kuin monissa muissa EU:n jäsenvaltioissa, sillä esimerkiksi Saksassa vastaavia rajoituksia on komission vaikutusarvioinnin mukaan jopa 11. Jotta data liikkuisi vapaasti ja asetuksen tavoitteet täyttyisivät, on pidettävä huoli, että jäsenvaltiot todella poistavat perusteettomat sijaintia koskevat vaatimukset. Suomen tulee ottaa tässä proaktiivinen rooli ja vaatia komissiota valvomaan asetuksen toteutumista, mikäli näin ei ole muuten tapahtumassa.

CSC:n mielestä on aiheellista kiinnittää huomiota erityisesti seuraaviin seikkoihin:

Joissain tapauksissa, esimerkiksi genomitietoihin liittyen, on syytä selvittää missä menee henkilötiedon ja muun kuin henkilötiedon raja. Puhdas genomitieto on selvästi henkilötietoa, mutta osa sen johdelmista voi olla syytä tulkita muuksi kuin henkilötiedoksi. Anonymisoitu genomitieto ei välttämättä ole enää henkilötietoa, etenkin siinä vaiheessa, kun se on aggregoitu tilastotiedoksi. Suomalaisten genomitiedon pitäminen Suomen rajojen sisällä on perusteltua, mutta tutkimustulokset ja tilastolliset analyysit ovat muuta kuin henkilötietoa, eikä niiden sijaintia ole siksi syytä rajoittaa. Tällaisen tiedon vapaa liikkuvuus maiden ja alueiden välillä on välttämätöntä, jotta eri maissa olevat tutkijat voivat tehdä työtään ja kehittää entistä parempia ratkaisuja erilaisten sairauksien hoitoon ja ennaltaehkäisyyn.

Genomitiedon lisäksi myös osana terveydenhuoltoa otettavien kuvien kanssa tulee olla tarkkana. Magneettikuva tai mikroskoopilla näytteestä otettu kuva voivat olla henkilötietoa, jos kuvan mukana välitetään riittävästi tunnistetietoa, kuten nimi tai henkilöturvatus. THL:lla ja biopankeilla on kuitenkin keinot erottaa henkilötieto kuvantamistiedoista, jolloin kuvat voidaan välittää anonyyminä kuvamateriaalina esimerkiksi koneoppimisen koulutusta varten. Kuvan mukana välitetään täten vain ennustettava asia, vaikkapa mikroskooppikuvassa tieto, onko kyseessä ihosyöpä vai ei. Tämän tyyppisten tietojen siirtoa ei myöskään pidä rajoittaa, sillä kyse ei ole henkilötiedoista. Lisäksi tietojen liikkuvuus on välttämätöntä uusien hoito- ja diagnosointimuotojen kehittämiseksi.

On ensiarvoisen tärkeää ymmärtää, että pelkkä datan sijaintirajoitusten purkaminen ei ole riittävää datavetoiselle taloudelle suotuisien olosuhteiden rakentamiseksi. Ennen kaikkea tulee keskittyä sen varmistamiseen, että dataa todella voidaan käyttää ja hyödyntää eri sektoreiden (kaupallinen, tutkimus, julkinen) välillä. Hyvänä ohjenuorana toimivat eurooppalaiset yhteentoimivuusperiaatteet (EIF)¹.

CSC pitää erityisen tärkeänä datan semanttista yhteentoimivuutta – sitä, että dataa voidaan vastaanottaa ja käsitellä tavalla, jossa tietojen merkitys, käytettävyys ja ymmärrettävyys säilyvät. Yhtä lailla datanhallintaan liittyvät ohjeistukset ja datapolitiikat, metadatan laatu ja datanhallintaa tukevat palvelut vaikuttavat datan löydettävyyteen ja uudelleenkäyttöön. Tarvitaan yhteisiä tietomalleja, käytänteitä ja standardeja sekä ennen kaikkea osaamista, jotta data liikkuisi jäsenvaltioiden, toimijoiden ja sektoreiden välillä sujuvasti. Loppukäyttäjien näkökulman tulee olla keskiössä kaikessa suunnittelussa, jotta lopputuloksena on tehokas ja luotettava tiedonhallinta.

Datan tulee olla helposti saatavilla, jotta tutkijat pääsevät hyödyntämään dataa. Pääsynhallinnan tulee olla mahdollisimman sujuvaa ja luotettavaa, ja tutkijat on pystyttävä tunnistamaan koneellisesti. Myös kansainvälisten tutkijoiden tulee pystyä tunnistautumaan suomalaisissa palveluissa, jotta he pystyvät hyödyntämään suomalaisia tutkimusaineistoja. Sama pätee tietysti muissa EU-maissa sijaitseviin aineistoihin. EU:n tulee kehittää unioninlaajuisia prosesseja, joilla varmistetaan, että tutkimuksessa käytettävä data on tutkijoille mahdollisimman avointa, löydettävää ja helposti hyödynnettävää (FAIR-periaatteet²). Euroopan avoimen tieteen pilvipalvelu EOSC³ on tästä hyvä esimerkki. Suomen tulee jatkaa hankkeen tukemista ja pitää huoli, että EOSC toteutuu tarkoituksenmukaisella tavalla.

¹ https://ec.europa.eu/isa2/eif_en

² <https://www.fairdata.fi/miksi-fairdata/fair-periaatteet/>

³ <https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud>



EU:n on myös jatkossa syytä varmistaa, että datatalouteen liittyvä lainsäädäntö on yhtenäistä eikä eri hallinnonaloilla vedetä eri suuntaan. Rinnakkaisten aloitteiden, kuten tekijänoikeuksien modernisoinnin, tulee tukea datavetoisen talouden kehittämistä, eikä päinvastoin, kuten nyt on tapahtumassa. Tekijänoikeusdirektiivissä oleva mahdollisuus rajoittaa tietyille aineistoille tehtävä tekstin- ja datanlouhinta vain ei-kaupalliseen tutkimuskäyttöön toimii varoittavana esimerkkinä. Poikkeuksen laajentaminen koskemaan myös kaupallista käyttöä on ensiarvoisen tärkeää datan täysimääräisen hyödyntämisen ja eurooppalaisen datatalouden ja tekoälyn kehittämisen kannalta. Suomen täytyy puuttua tekijänoikeusdirektiivin, ja jatkossa tarvittaessa myös muun datatalouden kannalta ongelmallisen lainsäädännön, ongelmakohtiin viimeistään kansallisessa toimeenpanossa.

Yhteenvetona CSC haluaa korostaa, että datan sijaintirajoitusten poistaminen ei itsessään riitä takaamaan datan vapaata liikkuvuutta, koska digiaikana datan sijainnin sijaan yhä merkityksellisemmäksi nousevat muut tekijät. Datan liikkuvuus kulminoituu viime kädessä yhteentoimivuuteen, jonka tulee toteutua kaikilla tasoilla: paitsi teknisellä, myös organisatorisella sekä erityisesti semanttisella ja lainsäädännöllisellä tasolla. Jos halutaan saavuttaa toimiva datatalous, tätä aihetta tulee tarkastella kokonaisvaltaisemmin kuin pelkästään sijainnin näkökulmasta.

Espoossa, 16.10.2019

CSC – Tieteen tietotekniikan keskus Oy

Kimmo Koski

Irina Kupiainen

Toimitusjohtaja

Ohjelmajohtaja