

Asia: VN/24821/2021

Lausuntopyyntö luonnoksesta hallituksen esitykseksi eduskunnalle laiksi Genomikeskuksesta ja terveyteen liittyvän geneettisen analyysin suorittamisen edellytyksistä

Lausunnonantajan lausunto

Voitte kirjoittaa lausuntonne alla olevaan tekstikenttään

Lausunto luonnoksesta hallituksen esitykseksi laiksi Genomikeskuksesta ja terveyteen liittyvän geneettisen analyysin suorittamisen edellytyksistä

Antti Honkela, TkT

Apulaisprofessori, Helsingin yliopisto, Tietojenkäsittelytiede

Haluan käyttää mahdollisuutta antaa lausunto luonnoksesta tietojenkäsittelytieteen ja tietosuojan tutkimuksen näkökulmasta. Johdan Suomen tekoälykeskus FCAI:ssa Privacy-preserving and Secure AI -tutkimusohjelmaa ja toimin varajohtajana vastuullista sote-tiedon käyttöä tutkivassa poikkitieteellisessä Datalit-projektissa. Olen lisäksi pysyvänä asiantuntijana STM:n nimittämässä sote-tiedon toisiokäytön ohjausryhmässä. Omassa tutkimuksessani olen työskennellyt paljon erilaisen genomitiedon analysoinnin ja hyödyntämisen parissa kehittäen tähän uusia laskennallisia menetelmiä.

Ehdotetun lain keskeisenä sisältönä ehdotetaan uuden viranomaisen, Genomikeskuksen, perustamista. Tämä vaikuttaa erittäin huonosti perustellulta seuraavista syistä:

* Genomikeskukselle ei esitetä tehtäviä, jotka perustelisivat uuden viranomaisen perustamisen. Genomikeskuksen perustaminen voisi olla perusteltavissa genomitietorekisterin ylläpitäjäksi, mutta sitä ei tässä esitetä. Toisaalta genomitietorekisterin perustamisen hyöty on hyvin kyseenalainen (tästä lisää alla).

* Luonnoksessa ei ole arviota esityksen vaikutuksista julkiseen talouteen, mutta perusteluissa viitataan tarpeeseen merkittävästä resursseista. Uuden merkittäviä resursseja vaativan viranomaisen perustaminen tilanteessa, jossa esimerkiksi Sosiaali- ja terveysalan tietolupaviranomainen on pahasti aliresursoitu vaikuttaisi vastuuttomalta.

* Nopeasti kehittyvällä alalla kuten genomilääketieteessä parhaan asiantuntemuksen ylläpito edellyttää asiantuntijoiden osallistumista alan tutkimukseen. Tästä syystä asiantuntijatehtäviä toteuttavan viranomaisen tehtäviin pitäisi sisältyä tieteellinen tutkimus. Maallikon näkökulmasta jää epäselväksi, mikseivät olemassaolevat tahot kuten THL selviytyisi Genomikeskukselle tässä laissa esitetyistä tehtävistä jopa paremmin kuin suunniteltu Genomikeskus.

Genomikeskuksen perustamista on hyvin vaikea perustella muuten kuin genomitietorekisterin kehittämistä ja ylläpitoa varten.

Kansallinen genomitietorekisteri muodostaisi merkittävän riskin kansalaisten tietosuojalle ja yksityisyydelle. Lisäksi informaatioteknologian kehitys tulee erittäin suurella todennäköisyydellä tekemään tuollaisen rekisterin turhaksi melko pian.

Genomitieto eroaa luonteeltaan muista potilastiedoista, koska se on äärimmäisen identifioivaa. Kun viranomaisten pitää tunnistaa vainaja tai rikoksen tekijä, paras työkalu on usein DNA eli genomi. Identtisiä kaksosia lukuun ottamatta jokaisen genomi on yksilöllinen. Usein jo pieni osa genomista riittää tunnistustiedoksi. Pieni määrä yleisiä variantteja ei riitä kenenkään tunnistamiseen, mutta harvinaisten varianttien kohdalla jo yksi voi joskus riittää. Koska ihminen saa puolet genomistaan kummaltakin vanhemmalta, genomi kertoo yksilön lisäksi tietoa tämän lähisukulaisista ja jopa vielä syntymättömistä lapsista.

Kansallinen genomitietorekisteri olisi valtava yksityisyysriski, koska se mahdollistaisi käytännössä kaikkien kansalaisten tunnistamisen suhteellisen helposti arkielämässä jälkeen jäävistä biologisista näytteistä. Kerran luotu rekisteri säilyisi käytännössä ikuisesti. Yksittäisen ihmisen genomien selvittäminen on periaatteessa helppoa: hankit vain hiuksen tai muun biologisen näytteen kohteesta, minkä perusteella voit sekvensoida genomien korkeintaan tuhannella eurolla. Erona yksittäisen ihmisen ja kansallisen rekisterin välillä on, että yksittäisiä ihmisiä seuraamalla ei ole mahdollista tehdä tunnistushakuja koko väestöstä.

Mahdollisesti perustettavan kansallisen genomitietorekisterin tietoturva pyrittäisiin toki turvaamaan mahdollisimman vahvasti. Valitettavasti täydellistä turvaa ei ole, joten kerätyn rekisteritiedon paljastumisriskiä on mahdoton eliminoida muuten kuin jättämällä rekisteri keräämättä. Genomitietorekisterin luonteen vuoksi tietojen paljastumisen aiheuttama riski olisi ainutlaatuinen ja aiheuttaisi mittamatonta pitkäaikaista haittaa huomattavalle osalle Suomen kansasta. Toisin kuin muut tiedot, genomitieto ei vanhene koskaan, eikä paljastunutta genomia ole mahdollista muuttaa tai vaihtaa. Genomitutkimus on suhteellisen nuorta, emmekä vielä tiedä, millaista vahinkoa genomitiedolla olisi mahdollista aiheuttaa.

Viimeaikainen tekoäly- ja data-analytiikkamenetelmien kehitys on liikkunut nopeasti kohti hajautettua mallia, jossa data pysyy mahdollisimman lähellä sen kohteita. Uudet salausmenetelmät (esim. homomorfinen salakirjoitus) mahdollistavat analyysien tekemisen salakirjoitetulla datalla.

Tämän kehityksen perusteella vaikuttaa todennäköiseltä, että viimeistään 10 vuoden aikajänteellä olisi mahdollista toteuttaa hajautettu järjestelmä, jossa kansalaisilla säilyisi kontrolli heidän genomitietoihinsa, mutta joka samalla mahdollistaisi niiden käytön tutkimuksessa suostumuksen pohjalta. Tarvittavat informaatioteknologian perusteknologiat ovat jo olemassa, enää tarvitaan soveltavaa tutkimusta ja tuotekehitystä niiden skaalaamiseksi tällaisiin sovelluksiin.

Yhteenvetona esitys tulisi palauttaa uuteen valmisteluun, jossa selvitetäisiin genomitiedon tallennukseen vaihtoehtoja, jotka eivät edellytä kansallisen genomitietorekisterin perustamista ja varmistavat kansalaisten mahdollisuudet kontrolloida omien genomitietojensa käyttöä ja säilytystä.

Esimerkiksi European Health Data Space -valmistelun yhteydessä kehitetään ratkaisuja hajautettujen terveystietojen hyödyntämiseen. Tämän työn tuloksena syntyy todennäköisesti ratkaisuja, jotka tekevät genomitietorekisterin tarpeettomaksi, minkä jälkeen myöskään Genomikeskukselle ei jäisi riittävästi tehtäviä sen perustamisen perustelemiseksi.

Antti Honkela