

5.7.2019

STM/071:00/2018 ja STM/2688/2018



## CSC - Tieteen tietotekniikan keskus Oy

### Lausunto hallituksen esitysluonnoksesta koskien Genomikeskusta ja genomitietojen käsittelyn edellytyksiä

CSC - Tieteen tietotekniikan keskus Oy (CSC) kiittää mahdollisuudesta osallistua sosiaali- ja terveysministeriön kuulemiseen koskien hallituksen esitysluonnosta Genomikeskuksesta ja genomitietojen käsittelyn edellytyksistä. 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. CSC on yli 40 vuoden ajan tehnyt ICT-ratkaisuja edistyneen ja haastavan tutkimuksen tarpeisiin ja yli 20 vuoden kokemuksella tuottanut genomitiedon käsittelyn ja hallinnan palveluja tutkimukselle. CSC on biologisen tiedon hallintaa kehittävän kansainvälisen ELIXIR-infrastruktuurin Suomen osakeskus.

CSC näkee Genomikeskuksen perustamiseen tähtäävän ja genomitietojen käsittelyn edellytykset määrittävän lainsäädännön luomisen erittäin tärkeänä askeleena erityisesti terveystutkimuksen kehityksen näkökulmasta. Lakiehdotus on syventynyt työstön aikana, ja CSC kiittää valmistelijoita erittäin kattavasta lainsäädännöllisen ympäristön ja nykytilan kuvauksesta. Lakiehdotuksessa huomioidaan varsin hyvin terveydenhuollon käytön mahdollistaminen ja Kelan rooli siinä. Nykyisellään CSC näkee lakiehdotuksen kuitenkin painottuvan vahvasti genomitiedon terveydenhuollon käyttöön, jolloin yhteys toisiokäyttöön jää valitettavan heikoksi. Lisäksi lakiehdotus sisältää riskejä tutkimustoiminnan onnistumisen ja siten myös terveydenhoidon kehityksen kannalta. Lain jatkokäytössä onkin huomioitava, että tieteellisten läpimurtojen saavuttamiseksi terveydenhuollon genomidata on saatava tutkimuskäyttöön, osana SOTE-toisioissa määritettyä lupaprosessia, josta tulokset voidaan taas palauttaa takaisin terveydenhuollon pariin.

#### Toisiokäytön vaatima tekninen osaaminen

Toisiokäytön tekniset vaatimukset eroavat merkittävästi terveydenhuollon käytön vaatimuksista. Genomidatan käyttöpalveluiden toteuttamisessa on olennaista, että sitä toteuttavilla organisaatioilla on riittävästi genomidatan hallintaan liittyvää ICT-asiantuntemusta sekä yhteyksiä vastaavia keskuksia muihin maihin rakentaviin toimijoihin. Jatkokäytössä on syytä pohtia, tulisiko lainsäädännössä eritellä tarkemmin terveydenhuollon käyttöön ja toisiokäyttöön liittyvät vaatimukset sekä toteuttajat, samalla huomioiden tietoturva ja käytännön sanelemat tekniset vaatimukset. CSC näkee, että yksilötason genomitiedon hallinta kannattaa perustaa valtiosektorin datanhallinnan palveluiden varaan, kunnes kokemuksen perusteella tiedetään tarkemmin, miten suuria määriä genomitietoa voidaan soveltaa. Lain ei kuitenkaan pitäisi radikaalisti rajoittaa anonyymien genomitiedon prosessointia, jonka varassa moni terveydenhuollon kehityksen kannalta tärkeä tutkimus toimii.

CSC:llä on runsaasti asiantuntemusta sensitiivisen datan hallinnasta ja ICT-ratkaisuista, joiden avulla genomitiedon toisiokäyttöpalvelut voidaan mahdollistaa. CSC on esimerkiksi mukana kehittämässä kansainvälisten yhteistyökumppaneiden kanssa tutkimuksen käyttöön genomiaineistojen vastaanottojärjestelmää, joka perustuu Euroopan genomi- ja fenomiarkiston<sup>1</sup> teknologiaan mahdollistaen mm. kansainvälisten referenssiaineistojen

---

<sup>1</sup> European Genome-phenome Archive, EGA



käytön osana kotimaista genomitutkimusta. Lisäksi CSC:n, THL:n ja HUS:n välisessä yhteistutkimuksessa on luotu työnkulku geneettisen tiedon tallentamiselle ja yhdistämiselle potilasasiakirjoista tuotettuun tietoon. Aineistoja ei välttämättä tarvitse keskittää CSC:n infrastruktuuriin, vaan tärkeintä on varmistaa aineistojen kytkeminen tietoturvallisesti osaksi federoitua ja integroitua palvelua, jonka CSC tarjoaa yhdessä muiden toimijoiden kanssa. Geneettisen tiedon tietoturallinen jakaminen terveydenhuollon ja tutkimuksen käyttöön voidaan mahdollistaa etätyöpöydän avulla, mikä varmistaa aineistojen säilymisen järjestelmän sisällä ja takaa skaalautuvat resurssit tutkimuksen käyttötapauksille. Nämä käyttötapaukset eroavat olennaisesti terveydenhuollon käytön vaatimuksista. CSC:llä on vankka kokemus käyttötapauksien muutosten hallinnasta kansallisessa tutkimuskentässä.

### Genomitiedon käyttö tutkimuksessa

Lakiehdotuksen 10 § määrittää, miten variaatiodataa voidaan luovuttaa tutkimuskäyttöön, mikä jättäisi genomitiedon (esim. raakasekvenssitiedon) käytön ulos tutkimuksesta. Tutkimusluvan rajaaminen lakitekstissä variaatitietoihin on ongelmallista, sillä variaatitietojen muodostaminen geenitiedosta ei ole yksiselitteisesti määriteltävissä oleva prosessi, vaan sen yksityiskohdat riippuvat käyttötarkoituksesta. Lisäksi tämä prosessi on osa jatkuvaa ja välttämätöntä tutkimusprosessia, jonka kehitys vaarantuu, mikäli genomikeskuksen genomidata rajataan tutkimuksen ulkopuolelle. Käytännössä tutkija harvoin käyttää tutkimuksessaan ainoastaan yhtä dataa vaan yhdistelee useaa, mikä ei ole ongelmatonta tällä lain määrittelyllä. Jatkotyöstössä tulee varmistaa, että genomitiedon toisiokäyttö mahdollistetaan tutkimukselle.

### Kansainvälinen tutkimusyhteistyö

Kansainvälinen yhteistyö on välttämätöntä genomitutkimuksessa jo suomalaisen datan määrän rajallisuuden vuoksi. Suomalaisen tutkimustoiminnan kilpailukyvyyn ja houkuttelevuuden kannalta on tärkeää, että kansainvälisen yhteistyön toimintamahdollisuuksia ei kavenneta. Kansainvälinen yhteistyö ja kansainvälisen referenssidatan käyttö genomitietoa hyödyntävissä suomalaisissa tutkimuksissa on välttämätöntä esimerkiksi lääketieteellisten läpimurtojen saavuttamiseksi. Toisiokäyttö ei rajoitu ainoastaan suomalaisten tekemään tutkimukseen, vaan toisiolupakäytännön pitää mahdollistaa kansainvälisten yhteistyökumppaneiden pääsy tutkimusaineistoihin. Tämä pitää huomioida käyttäjien tunnistautumisessa, joka on vaikeampaa järjestää kansainvälisille käyttäjille.

### CSC:llä on lisäksi seuraavia muita huomioita lakiluonnoksesta:

Genomitiedon saatavuuteen ja luvitukseen liittyen CSC näkee, että aineiston löytäminen ja käyttöluvan hakeminen tulee tehdä hakijalle mahdollisimman helpoksi. Koska keskitetyt ratkaisut takaavat paremman kokonaiskuvan olemassa olevasta aineistosta, olisi sekä genomitiedon hyödyntäjän että luvittajien näkökulmasta tarkoituksenmukaista, että genomitiedon saatavuus- ja luvituspalvelut olisivat riippumattomia itse luvanantajasta.

Lakiehdotuksen 2 § 4. kohdassa määritellään termi *viitetieto*. Tässä yhteydessä yhtenä genomiaineiston osana *viitetieto* terminä on harhaanjohtava. Yleiskielessä ja tieteellisessä kontekstissa viitetieto-termillä tarkoitetaan lähdeviitteitä eli bibliografisia tietoja. Englanniksi käytetään termiä ”allele frequency”, joka tarkoittaa tunnistetun genomivariantin suhteellista esiintymistä populaatiossa. Tässä yhteydessä voisi käyttää esimerkiksi termiä genomivariaation ”yleisyystieto”.

Lakiehdotuksen 7 § 3. momentin mukaan genomitiedot on tallennettava genomitietorekisteriin vuoden kuluessa siitä, kun geneettinen analyysi on valmistunut. Lakiehdotuksen perusteluista s. 215 käy ilmi, että tässä viitataan genomitiedon raakadataan, mikä ei sisällä sen analysointia informatiikan keinoin. Perustelun teksti ei tee aivan selväksi, onko aineiston oltava aina avoimena toisiokäyttöön vuoden sisällä, sillä kappaleen ensimmäisessä lauseessa



viitataan vain tallentamiseen, eikä saatavuuteen. Lisäksi ehdotuksessa seuraava lause hämärtää merkitystä edelleen käsittelemällä määräaikaisten rajoitusten kirjaamista. Näin lain sananmuoto korostaa tallennusta eikä yhdistä sitä toisilain mukaiseen saatavuuteen.

Lakiehdotuksen 5 § 4. kohdassa määritetään Genomikeskuksen tehtäväksi ”luoda vastaanottamastaan tai tuottamastaan genomitiedosta variaatio- ja viitetietorekisteri”. Genomikeskus ei kuitenkaan lakiehdotuksen mukaan tuota genomirekisteriin tallennettua tietoa, vaan sen tuottavat terveydenhuolto tai biopankit, jotka myös hallinnoivat sitä Genomikeskuksen tietokannassa. Ilmeisesti edellä mainitussa lakiehdotuksen kohdassa termi genomitieto ymmärretään siis laajemmin kuin ”genomirekisteriin tallennettuna tietona”. Muutoin viittaus rekisteriin tallennetun tiedon tuottamiseen tulee poistaa laista.

Lakiehdotuksen 9 § määrittää, että tiedot tulee hävittää tai arkistoida viimeistään 50 vuodessa rekisteröidyn kuolemasta tai 120 vuodessa henkilön syntymästä, jos kuolinaika ei ole tiedossa. Lakiehdotuksen s. 241 todetaan kliinisen genetiikan potilastietoaaineistojen kriittisten säilytysaikojen olevan 120 vuotta potilaan kuolemasta. On vaikea olla rinnastamatta genomitietoa kriittisiin potilastietoaaineistoihin. Genominlaajuinen tieto on vielä harvinaista ja vanhimpien kansalaisten kohdalla lakiehdotuksen vaatimus voisi johtaa tiedon poistamiseen aktiivisesta käytöstä jo 50 vuoden kuluttua. Useamman sukupolven kattava geno- ja fenotyyppiaineisto on vielä niin harvinaisen, että sen vaikutusta terveydenhuoltoon on vaikea arvioida, mutta sitä kuitenkin ei tule aliarvioida.

Lakiehdotuksen 11 § 3. momentissa säädetään genomitiedon luovuttamisesta terveydenhuoltoon. Lakiehdotus korostaa genomitiedon erityisluonnetta, joka seuraa yhteisen genomiaineiston jakautumisesta sukulaisten parissa. Perusteluissa todetaan, että henkilön lupaa ei tarvita genomitiedon käyttämiseksi eikä terveydenhuollon tule kertoa potilaalle kenen genomitietoa hoidossa on käytetty. Lakiehdotuksessa ei kuitenkaan oteta kantaa, miten sukulaissuhteita selvitetään ja kuka niitä selvittää. Genomikeskuksella tulee olla valtuudet ja helpot keinot selvittää viranomaistietoihin kirjattuja sukulaissuhteita. Genomikeskuksen täytyy siis käytännössä päästä näihin tietoihin henkilötietorekisterissä.

Lakiehdotuksen s. 210 todetaan, että rekisteri sisältäisi suurimmaksi osaksi vain yleisiä suomalaisessa väestössä esiintyviä variaatiotietoja. Kuitenkin s. 13 viitataan EU:n perusoikeusasiakirjan syrjintää koskevaan artiklaan sekä s. 244 todetaan, että ehdotetun sääntelyn tavoitteena on edistää väestön terveyttä yhdenvertaisesti. Genomikeskus ei voi asettaa ihmisiä eriarvoiseen asemaan jättämällä huomiotta alkuperältään valtion nykyisten rajojen ulkopuolelta tulleiden Suomen terveydenhuollon piiriin kuuluvien ihmisten genomitietoa. Näin määritelty etnisen taustan tuntemus on Genomikeskuksen toiminnan tavoitteille tärkeää, sillä tämä tieto mahdollistaa geneettisen variaation yhdistämisen oikeaan kansainväliseen kontekstiin, mikä voi olla oleellista kliinisen merkityksen selvittämisessä.

Lakiehdotuksessa s. 240 todetaan, että genomitieto ei lähtökohtaisesti ole muuttuvaa muuten kuin syöpien kohdalla. Genomitieto voi yksilön kehityksen aikana muuttua merkittävästi. Eri kudostyyppisiin päätyneet mutaatiot voivat poiketa toisistaan merkittävästi. Genomin määrittäminen tehdään yleensä vain yhdestä solutyypistä. Genetiikan kirjallisuudesta löytyy lukuisia esimerkkejä, joissa somaattinen muuntelu on johtanut tautiin, joka ei ole ollut nähtävissä toisesta somaattisesta kudoksesta tai sukusolulinjasta tehdyn määrittelyn perusteella. Viite: *The human body is a mosaic of different genomes*, <https://www.nature.com/articles/d41586-019-01780-9>.

Lakiehdotuksen perusteluissa s. 213 rajataan Genomikeskuksessa säilytettävästä aineistosta pois genomitieto, joka ei täytä tarkemmin määrittelemättömiä terveydenhuollon laatukriteereitä. Toisaalta s. 166 todetaan kaupallisten suoraan yksityisille tarjottujen geneettisten palvelujen (Direct to Consumer, DTC) olevan lähtökohtaisesti Genomikeskuksen tietokantoihin talletettavia. On tärkeää huomioida, että geneettistä tietoa arvioitaessa laadun ja



merkitsevyyden välille ei voi vetää yhtäläisyyttä. Geneettisessä riskiarvioinnissa lähisukulaisen huonompilaatuinenkin aineisto voi olla merkittävä jo siksi, että se on jo olemassa. Käytännössä Genomikeskuksen on otettava vastaan kaikki merkittävä genomitieto ja kehitettävä luokittelujärjestelmä, joka ohjaa sen käyttöä.

Espoossa, 5.7.2019

CSC – Tieteen tietotekniikan keskus Oy

Kimmo Koski

Toimitusjohtaja

Tommi Nyrönen

Ohjelmajohtaja