16.6.2020

Antti Honkela
Datatieteen apulaisprofessori
Tietojenkäsittelytieteen osasto
Helsingin yliopisto

Eduskunnan sosiaali- ja terveysvaliokunnalle

Asia: Asiantuntijapyyntönne HE 101/2020 vp

Lausunto Hallituksen esityksestä eduskunnalle laiksi tartuntatautilain väliaikaisesta muuttamisesta

Kiitän sosiaali- ja terveysvaliokuntaa mahdollisuudesta esittää lausunto covid-19-tartuntaketjujen katkaisua tehostavan mobiilisovellukseen perustuvan tietojärjestelmän edellyttämästä lainmuutoksesta. Oma asiantuntemukseni liittyy erityisesti tekniseen ja tilastolliseen tietosuojaan, joten tarkastelen esitystä ensisijaisesti tästä näkökulmasta.

Esityksessä hahmotellun kaltainen tietojärjestelmä voi tietyissä olosuhteissa auttaa ratkaisevasti covid-19-epidemian hillitsemisessä ja täten säästää uusilta keväällä 2020 tarvittujen kohdentamattomien rajoitustoimien aiheuttamilta merkittäviltä kustannuksilta ja muilta haitoilta.

Odotettavien hyötyjen toteutumisen kannalta on välttämätöntä, että riittävän suuri osa väestöstä ottaa sovelluksen käyttöön erityisesti tiheään asutuilla alueilla, joissa kontakteja tuntemattomien kanssa on paljon. Koska käyttöönotto perustuu asianmukaisesti vapaaehtoisuuteen, on laajan käyttöönoton varmistamiseksi ensiarvoisen tärkeää varmistaa, että jokainen kansalainen kokee järjestelmän hyödyt itselleen haittoja suuremmiksi, ja varmistaa järjestelmän vahva tietoturva ja tietosuoja.

Esityksessä hahmoteltu järjestelmä on perusteiltaan vaaditun vahvan tietoturvan ja tietosuojan mukainen hajautettuun arkkitehtuuriin perustuva järjestelmä, jossa ei kerätä käyttäjien sijaintitietoa. Käyttäjän kannalta positiivista on, että pelkän sovelluksen havaitseman altistumisen perusteella ei voida määrätä karanteeniin (43 e §, 3. momentti). *Heikkoutena ehdotetussa laissa ei oteta selvää kantaa tiettyihin järjestelmän tietosuojan kannalta kriittisiin toteutuksen yksityiskohtiin*. Koska nämä ovat järjestelmän käyttäjien tietosuojan kannalta kriittisiä, niistä tulisi säätää suoraan laissa.

Esitys on pääpiirteissään hyvä. *Keskityn sen vuoksi alla muutamiin ehdotuksen heikkouksiin, joista erityisesti ensimmäinen olisi kriittisen tärkeää korjata ennen ehdotuksen hyväksymistä*. Muilta osin kannatan esityksen hyväksymistä.

Käytettyjen pseudonyymisten tunnisteiden vaihtuminen ja altistuneiden informoiminen vahvistetun tartunnan jälkeen

Ehdotetun järjestelmän tietosuojan kannalta on keskeistä, että mobiilisovelluksen lähettämät pseudonyymiset tunnisteet vaihtuvat riittävän usein. Muuten niistä muodostuisi kiinteä tunniste, joka mahdollistaisi yksittäisen käyttäjän seurannan.

Ongelma on tiedostettu esityksen kohdan 4.2 alakohdassa Tietosuojavaikutukset (s. 20), jossa todetaan: "... Jokainen käyttäjä saisi yksilöllisen pseudonyymisen tunnisteen, joka muuttuisi säännöllisesti. ... Tietosuojaan kohdistuvia riskejä madallettaisiin kuitenkin merkittävästi sillä, että suorien tunnistetietojen sijaan käsitellään tehokkaasti pseudonyymisiä tietoja, niin sanottuja pseudonyymisiä tunnisteita, jotka muuttuvat riittävän usein."

Toisaalta säännöskohtaisissa perusteluissa (43 c §, s. 25) todetaan ainoastaan: "Sovelluksen käyttöön ottaessaan käyttäjä ei antaisi sovellukselle suoria tunnistetietojaan, vaan jokaiselle käyttäjälle osoitettaisiin sattumanvarainen pseudonyyminen tunniste. Nämä pseudonyymiset tunnisteet voisivat muuttua esimerkiksi tietyn ajan jälkeen."

Tämän epämääräisyyden poistamiseksi ja asian tärkeyden takia vaatimus tunnisteiden muuttumisesta pitäisi olla suoraan lakitekstissä.

Toinen merkittävä riski liittyy järjestelmän väärinkäyttöön, jos käyttäjä voi tehdä muiden kiusaksi ilmoituksen tartunnasta, vaikka hänellä ei todellisuudessa sitä ole. Esitykseen sisältyvässä järjestelmän kuvauksessa tämä on estetty vaatimalla tartuntailmoituksen tekemiseen koodi, jonka saa vain terveysviranomaisilta tartunnan vahvistamisen jälkeen. Koska ominaisuus on järjestelmän tietoturvan ja tietosuojan kannalta keskeinen, siitä tulisi säätää itse lakitekstissä.

Ehdotus:

Lisätään pykälään 43 a § uusi momentti:
Terveyden ja hyvinvoinnin laitos vastaa, että järjestelmän mobiilisovellus turvaa käyttäjien tietosuojan varmistamalla, että sovelluksen käyttämät pseudonyymiset tunnisteet vaihtuvat riittävän usein käyttäjän pitkäaikaisen seurannan estämiseksi ja että 43 c § 2 momentin 3 kohdan mukaisen ilmoituksen tartunnasta voi tehdä ainoastaan vahvistetun tartunnan saanut sovelluksen käyttäjä.

Tietojärjestelmän tietoturvallisuuden arvioiminen

Esityksessä ehdotetaan Liikenne- ja viestintävirastolle vastuuta tietojärjestelmän tietoturvallisuuden arvioinnista. Tämä on hyvä alku, mutta ei riittävä. Koska järjestelmän toteutukseen sisältyy huomattavia käyttäjien yksityisyyteen liittyviä riskejä, olisi tärkeää mahdollistaa myös riippumaton erityisesti mobiilisovellukseen kohdistuva arviointi julkaisemalla sen lähdekoodi.

Lähdekoodin julkisuutta pidetään yleisesti erittäin hyödyllisenä tietojärjestelmien tietoturvallisuuden arvioinnin kannalta. Alalla pidetään yleisesti hyväksyttynä, että järjestelmän turvallisuus ei saa perustua sen toiminnan salaamiseen vaan sen pitää perustua julkisuuteen ja avoimeen arviointiin.

Lähdekoodin julkisuus todennäköisesti vahvistaisi erityisesti tässä tapauksessa kansalaisten luottamusta sovelluksen tietoturvaan ja voisi sikäli edistää sovelluksen laajaa käyttöönottoa, mikä olisi välttämätöntä esityksen hyötyjen toteutumiseksi.

Ehdotus:

Lisätään pykälään 43 g § maininta:
Lisäksi Terveyden ja hyvinvoinnin laitos velvoitetaan antamaan muille tahoille mahdollisuus arvioida tietojärjestelmän turvallisuutta julkaisemalla 43 a § 1 momentissa tarkoitetun mobiilisovelluksen toteutuksen lähdekoodi ennen järjestelmän käyttöönottoa.

Tilastotiedon kerääminen kontaktien lukumäärästä epidemian seurantaa ja mallinnusta varten

Esityksen mukainen järjestelmä tarjoaisi erinomaisen tilaisuuden kerätä tosiaikaista tilastotietoa epidemian leviämisen taustalla olevista kontakteista epidemian seurantaa ja mallinnusta varten. Tällainen tiedonkeruu olisi mahdollista toteuttaa käyttäjien tietosuojaa vaarantamatta[1]. Koska tilastotiedon keruu onnistuu varsinaista altistumisten jäljitystä pienemmillä käyttäjämäärillä, se lisäisi esitetyn järjestelmän toteutuksesta saatavia hyötyjä erityisesti tilanteessa, jossa sovelluksen käyttäjämäärä jää toivottua pienemmäksi.

Kontaktien lukumääristä kerätyt tilastotiedot ovat anonyymejä, joten ne eivät ole henkilötietoja. Nämä tilastotiedot kuitenkin perustuvat henkilötietoihin, joita pitää käsitellä anonyymin tilaston muodostamiseksi. Tämän mahdollistamiseksi on selvyyden vuoksi syytä lisätä järjestelmän henkilötietojen käsittelyperusteisiin kohta, joka sallii kontakteista kerättyjen henkilötietojen käsittelyn niiden anonymisoimiseksi tilastollista seurantaa varten.

Ehdotus:

Lisätään pykälän 43 c § 3 momenttiin uusi kohta:
2 momentin 2 kohdassa tarkoitettuja tietoja saa lisäksi käsitellä niiden anonymisoimiseksi epidemian tilastollista seurantaa varten.

[1] Vieraskynäkirjoitus A. Honkela, ”Koronavirusepidemian seuranta kaipaa lisädataa”, HS 1.6.2020