# Tietosuoja tutkimuksessa

*Arja Kuula-Luumi (Tietoarkisto)*

*Tutkimusaineistojen anonymisointi -seminaari*

*5.4.2017 Tampereen yliopisto*

TIETOARKISTO

TIETOARKISTO
FINNISH SOCIAL SCIENCE DATA ARCHIVE

# Sisältö

- Henkilötieto
- Henkilötietojen käsittelyn laillisuusperusteet tutkimuksessa
- Henkilötietolain ja EU:n tietosuoja-asetuksen mukaiset suojatoimet
- Määritelmät
  - Tunnisteellinen tieto
  - Pseudonyymi aineisto
  - Anonyymi aineisto
- Tietoarkiston suojatoimet

TIETOARKISTO
FINNISH SOCIAL SCIENCE DATA ARCHIVE

# Mitä on henkilötieto?

## Henkilötietolaki

- *henkilötiedolla* tarkoitetaan kaikenlaisia luonnollista henkilöä taikka hänen ominaisuuksiaan tai elinolosuhteitaan kuvaavia merkintöjä, jotka voidaan tunnistaa häntä tai hänen perhettään tai hänen kanssaan yhteisessä taloudessa eläviä koskeviksi;

## EU:n tietosuoja-asetus

- 'henkilötiedoilla' tarkoitetaan kaikkia tunnistettuun tai tunnistettavissa olevaan luonnolliseen henkilöön, jäljempänä 'rekisteröity', liittyviä tietoja; tunnistettavissa olevana pidetään luonnollista henkilöä, joka voidaan suoraan tai epäsuorasti tunnistaa erityisesti tunnistetietojen, kuten nimen, henkilötunnuksen, sijaintitiedon, verkkotunnistetietojen taikka yhden tai useamman hänelle tunnusomaisen fyysisen, fysiologisen, geneettisen, psyykkisen, taloudellisen, kulttuurillisen tai sosiaalisen tekijän perusteella.

TIETOARKISTO
FINNISH SOCIAL SCIENCE DATA ARCHIVE

# Henkilötietojen käsittelyn laillisuusperusteet tutkimuksessa

- Tutkittavan suostumuksella kerätyt tutkimusaineistot
  - Käsittelyperusteena henkilötietolain mukainen suostumus yksilöityyn käyttötarkoitukseen

- Viranomaisen asiakirja- tai rekisteriaineistot, jotka saatu viranomaisen luvalla
  - Käsittelyperusteena henkilötietolain 14 § ja siihen sisältyvät suojatoimet

TIETOARKISTO
FINNISH SOCIAL SCIENCE DATA ARCHIVE

# Erityistapaukset

## Big Data

- *” Tietojen massamuotoinen hyödyntäminen tavalla tai toisella on mahdollista ainoastaan, jos tietosuojaa koskevat periaatteet otetaan asianmukaisesti huomioon.”* Lausunto liikenne- ja viestintäministeriön massadataa (big data) koskevasta selvityksestä (tietosuojavaltuutettu 18.12.2017 )
- Paul Ohm (2010). Broken Promises of Privacy: Responding to the Surprising Failure of Anonymisation. 57 Ucla Law Review 1701-1777.

## Sosiaalisen median aineistot

- Tietosuojan ohella sovellettavaksi tulee usein myös tekijänoikeudet
- Sopimusoikeudelliset seikat (esim. FB, Twitter jne.)
- http://www.fsd.uta.fi/fi/ajankohtaista/tapahtumat/humanistinen_tutkimus_ja_lainsaadanto_2016/

TIETOARKISTO
FINNISH SOCIAL SCIENCE DATA ARCHIVE

# Henkilötietolain 14 § mukaiset suojatoimet, jos tutkimusta ei voi suorittaa ilman tunnisteita

- Tutkimussuunnitelma
- Aineistoa käytetään vain suostumuksessa yksilöityyn tarkoitukseen
- Aineisto tulee suojata ulkopuolisilta
- Käyttötarkoituksen päätyttyä
  - aineisto hävitetään tai
  - arkistoidaan tunnisteineen Kansallisarkiston luvalla (HetiL 35 §) tai
  - **muutetaan sellaiseen muotoon, etteivät tutkittavat ole enää tunnistettavissa**

TIETOARKISTO
FINNISH SOCIAL SCIENCE DATA ARCHIVE

# EU:n tietosuoja-asetuksen mukaiset suojatoimet (1/2)

- Tekniset ja organisatoriset suojatoimet

  - Esim. koulutus, ohjeet, salassapitositoumukset, käytönvalvonta, tietoturva, tietojen salaus, etäkäyttöyhteydet, pseudonymisointi, **anonymisointi**

TIETOARKISTO
FINNISH SOCIAL SCIENCE DATA ARCHIVE

# EU:n tietosuoja-asetuksen mukaiset suojatoimet (2/2)

- Seloste henkilötietojen käsittelytoimista (ent. rekisteriseloste)
- Tietosuojan vaikutusten arviointi erityisryhmiin kuuluvien (ent. arkaluonteisten) tietojen käsittelylle

  - Tarkasteltava erityisesti suunniteltuja toimenpiteitä sekä suojatoimia ja mekanismeja, joiden avulla voidaan lievittää käsittelyyn kohdistuvaa riskiä ja varmistaa henkilötietojen suoja sekä asetuksen vaatimusten toteutuminen

TIETOARKISTO
FINNISH SOCIAL SCIENCE DATA ARCHIVE

# Lisätietoa EU:n tietosuoja-asetuksesta

- Tietosuojavaltuutetun toimiston sivusto sisältää tilannetietoa ja oppaita suomeksi
- www.tietosuoja.fi/fi/index/euntietosuojauudistus.html

- Article 29 Data Protection Working Party
- http://ec.europa.eu/newsroom/just/item-detail.cfm?item_id=50083

TIETOARKISTO
FINNISH SOCIAL SCIENCE DATA ARCHIVE

# Milloin ihmistä koskevat tiedot ovat henkilötietoja?

KAIKKI LUONNOLLISEEN HENKILÖÖN LIITTYVÄT TIEDOT OVAT HENKILÖTIETOA

- Lausunnot, toteamukset
- Väitteet, mielipiteet, arvoarvostelmat
  - Voi olla sekä objektiivista että subjektiivista
  - Ei edellytä totuutta tai todennettavuutta
- Yksityiselämä, perhe-elämä
- Ammatillinen toiminta
- Taloudellinen ja sosiaalinen käyttäytyminen

**HENKILÖÖN LIITTYVÄT TIEDOT TULEE OLLA RIITTÄVÄN YKSILÖIVIÄ YKSIN TAI YHDESSÄ, JOTTA HENKILÖN VOI TUNNISTAA**

- Nimi, osoite, henkilötunnus, auton rekisteri, sormenjälki, kasvokuva, ääni jne.
- Syntymäaika tai ikä
- Sukupuoli
- Etninen tausta
- Ammatti
- Koulutus
- Asuinpaikka
- Työpaikka, opiskelupaikka, koulu
- Kotitalouden koostumus
- Siviilisääty jne.
- Fyysiset piirteet
- Ainutkertaiset elämäntapahtumat
- Jne.

TIETOARKISTO
FINNISH SOCIAL SCIENCE DATA ARCHIVE

# Pseudonyymi aineisto

- Pseudonymisoinnilla tarkoitetaan tietueen yhden (tavallisesti ainutkertaisen) attribuutin korvaamista toisella.
  - Oikean nimen tilalle toinen nimi
  - Havaintotunnuksen koodaaminen toiseksi
- Jos tunnisteellinen aineisto edelleen olemassa, yhdistäminen on mahdollista = pseudonyymi aineisto on henkilötietoa
- Pseudonymisointi on suojatoimi, ei anonymisointia

TIETOARKISTO
FINNISH SOCIAL SCIENCE DATA ARCHIVE

# Anonyymi aineisto

- Aineisto on anonyymi vasta kun siitä ei voi millään *kohtuullisesti toteutettavissa olevilla* keinoilla tunnistaa yksittäisiä tutkittavia.
  - Otettava huomioon tunnistamisesta aiheutuvat kulut, tunnistamiseen tarvittava aika, käytettävissä oleva teknologia.

- Anonyymistä aineistosta ei voi tunnistaa yksittäisiä tutkittavia esimerkiksi epäsuorien tunnisteiden avulla tai yhdistämällä aineistoon muualta saatavia tietoja.

- Anonyymiin aineistoon ei voi myöskään yhdistää samoja tutkittavia koskevia uusia tietoja.

- Anonymisoinnin tulee olla peruuttamaton, jotta voidaan puhua anonyymistä aineistosta.

TIETOARKISTO
FINNISH SOCIAL SCIENCE DATA ARCHIVE

Latanya Sweeney, Akua Abu, Julia Winn (2013). Identifying Participants in the Personal Genome Project by Name.
Report number: Harvard University, Data Privacy Lab 1021-1
Cite as: arXiv:1304.7605 [cs.CY]
Figure 1. in page 3

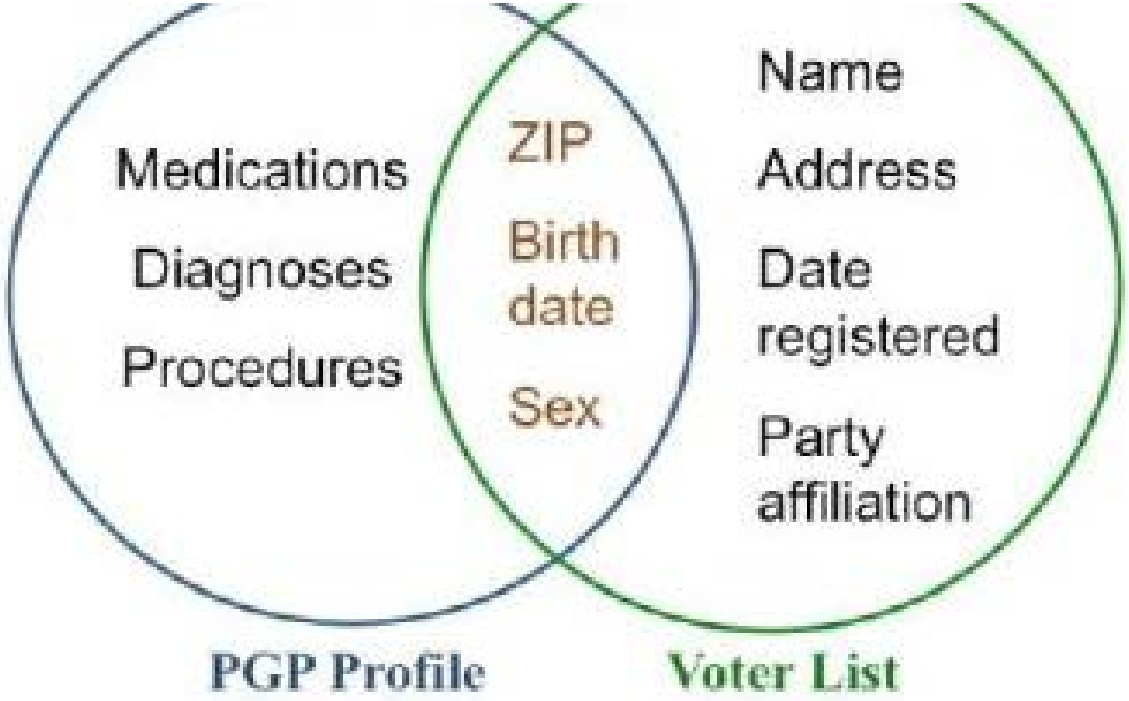


**Figure 1.** Linking PGP profile to a voter list using demographics to put names to the medical and genomic

TIETOARKISTO
FINNISH SOCIAL SCIENCE DATA ARCHIVE

# Henkilötietojen käsittelyperusteet Tietoarkistossa

1. Aineiston tulee olla arkistoitaessa anonyymi
   - Aineiston tarkistus arkistointisopimuksen nojalla: ”Tietoarkisto voi muokata vastaanottamaansa aineistoa voimassa olevien tietosuoja- ja tietoturvallisuusnormien ja pitkäaikaissäilytyksen vaatimusten mukaisesti.”
2. Anonymisointi toimeksiannosta (HetiL 8 §:n 1. mom. 7 kohta)
3. Harvoissa tapauksissa arkistoitavaan aineistoon voidaan soveltaa henkilötietolain poikkeussäännöstä, joka koskee toimituksellisia, taiteellisia ja kirjallisia tarkoituksia (HetiL 2 § 5 mom.).
   - Vain kuuden aineiston arkistointiin käytetty poikkeusperustetta (aineistoja Tietoarkistossa 1358)

TIETOARKISTO
FINNISH SOCIAL SCIENCE DATA ARCHIVE

# Tietoarkiston suojatoimet

- Sisäänoton tietoturva
- Säilytyksen tietoturva
- Anonymisointi
  - Henkilöstökoulutukset
  - Laadunvarmistus
  - Jäännösriskin arviointi
- Käyttäjätunnistus
- Käytön valvonta

TIETOARKISTO
FINNISH SOCIAL SCIENCE
DATA ARCHIVE

# Kiitos!

asiakaspalvelu.fsd @ uta.fi