

YK:N EUROOPAN TALOUSKOMISSIO
EUROOPAN TILASTOKONFERENSSI

(Epävirallinen käännös)

Tilastotietojen luottamuksellisuus & yksikötason tietojen käyttö

HYVIÄ KÄYTÄNTÖJÄ KOSKEVAT PERIAATTEET JA SUUNTAVIIVAT

YHDISTYNEET KANSAKUNNAT

New York ja Geneve, 2007

Huom.

Tässä julkaisussa käytetyt nimitykset tai tapa, jolla tiedot on esitetty, eivät kuvasta millään tavoin Yhdistyneiden kansakuntien sihteeristön mielipidettä minkään maan, alueen, kaupungin tai muun paikallisyhteisön, tai niiden viranomaisten, juridisesta asemasta, eivätkä niiden maantieteellisten rajojen sijainnista.

YHDISTYNEIDEN KANSAKUNTIEN JULKAISU

Sales No. E.07.II.E.7

ISBN 13: 987-92-1-116959-1

ISSN: 069-8458

Copyright © United Nations, 2007

All rights reserved

ii

Lisähuomautus

Nämä suuntaviivat on laatinut Australian tilastoviraston pääjohtajana toimineen Dennis Trewinin johtama task force -ryhmä Euroopan tilastokonferenssin (Conference of European Statisticians, CES) pyynnöstä.

Luottamuksellisuutta ja yksikötason tietojen käyttöä koskevat suuntaviivat ja pääperiaatteet hyväksyttiin Euroopan tilastokonferenssin täysistunnossa kesäkuussa 2006 ja hallituksessa lokakuussa 2006.

Suuntaviivat ovat käyttäjilleen varsin dynaaminen työkalu, sillä niitä päivitetään aika ajoin. Asiakirjan liitteiksi on odotettavissa erityisesti lisää malliesimerkkejä.

Suuntaviivojen englanninkielinen versio on saatavissa elektronisessa muodossa YK:n Euroopan talouskomission (United Nations Economic Commission for Europe, UNECE) tilastotoimiston verkkosivuilta osoitteesta <http://www.unece.org/stats/publ.htm>. Suomenkieliseen versioon ei ole käännetty liitteessä 1 olevia malliesimerkkejä (toimittajan huomautus).

Otamme mielellämme vastaan asiakirjaa koskevaa palautetta, jota voi lähettää osoitteeseen confidentiality@unece.org.

Kiitokset

Nämä periaatteet ja suuntaviivat ovat Euroopan tilastokonferenssin perustaman task force -ryhmän työn tulosta. Task force -ryhmän jäseniä olivat Ivan Fellegi (Kanada), Otto Andersen (Tanska), Teimuraz Beridze (Georgia), Luigi Biggeri (Italia) ja Tadeusz Toczynski (Puola). Task force -ryhmän puheenjohtajan Dennis Trewinin (Australia) panos on ollut merkittävä, sillä hän on laatinut asiakirjan alkuperäisen englanninkielisen tekstin ja liittänyt siihen monivaiheisen yhteistyön aikana eri valtioita ja kansainvälisiä järjestöjä koskevat osuudet.

Tiina Luige ja Gauri Khanna YK:n Euroopan talouskomission tilastotoimistosta olivat task force -ryhmälle suureksi avuksi.

Task force -ryhmä haluaa kiittää myös Svante Öbergiä (Ruotsi) ja Heinrich Brünggeriä (YK:n talouskomissio) aktiivisesta osallistumisesta työhön.

Useat valtiot ovat lisäksi tukeneet suuntaviivojen laadintaa toimittamalla asiakirjan liitteiksi arvokkaita malliesimerkkejä.

Lopuksi haluamme kiittää Euroopan tilastokonferenssin hallitusta koko projektin ajan saamastamme rakentavasta ohjauksesta.

SISÄLLYS

I. JOHDANTO.....	5
II. MIKSI KANSALLISTEN TILASTOVIRASTOJEN PITÄISI TUKEA TUTKIJAYHTEISÖÄ? ...	7
III. PÄÄPERIAATTEET	10
IV. ASIAA KOSKEVA LAINSÄÄDÄNTÖ	12
V. MENETELMÄT TUTKIJAYHTEISÖN TUKEMISEKSI.....	13
VI. KANSALLISTEN TILASTOVIRASTOJEN JA TUTKIJOIDEN VÄLISTEN JÄNNITTEIDEN HALLINTA	20
VII. YKSIKÖTASON AINEISTOJEN LUOVUTTAMISEEN LIITTYVIÄ HALLINNOLLISIA KYSYMYKSIÄ	24
VIII. MUITA ERITYISIÄ NÄKÖKOHTIA	26
LIITE 2. TERMINOLOGIAA.....	29

I. JOHDANTO

1. Tietosuojasta on perinteisesti vastattu lähinnä kansallisella tasolla. Internetin välityksellä tapahtuvan tiedon levityksen lisääntyttyä siitä on nyt kuitenkin tulossa myös kansainvälinen huolenaihe. Tutkijayhteisön jäsenet tekevät runsaasti kansainvälistä yhteistyötä, ja tutkijat voivat suhtautua hyvin kriittisesti maiden välisiin eroihin tietojen käyttöoikeuksissa. Usein tutkijoiden pääsy muiden maiden yksikkötason tietoihin on jopa estetty, koska tietosuojan pitävyyteen ei luoteta. Eri maiden väliset vertailut voivat kuitenkin olla oleellinen osa tutkimusprojektia. Asia kiinnostaa yliopistoissa toimivien tutkijoiden ohella myös muun muassa kansainvälisiä virastoja ja laitoksia, jotka käyttävät yksikkötason tietoja tutkimustarkoituksiin, erityisesti eri maiden välillä tehtäviin vertailuihin. Tällaiset tutkimukset ovat yleensä myös niissä mukana oleville maille erittäin kiinnostavia ja oleellisia.

2. Olisiko siis mahdollista sopia kansainvälisellä tasolla yhteisistä yksikkötason tietojen levittämistä koskevista periaatteista? Kysymystä tulee tarkastella vuoden 2003 Euroopan tilastokonferenssin valossa, jossa todettiin, että tutkimuksen tukeminen on tärkeä osa kansallisten tilastovirastojen toimintaa ja että tilastovirastoilla on näiden tarpeiden tyydyttämisessä yleisesti parantamisen varaa. Parannusta kaivataan esimerkiksi yksikkötason tietojen saatavuuteen, joka on myös näiden periaatteiden ja suuntaviivojen päätavoite. (Vaikka näissä suuntaviivoissa viitataan kansallisiin tilastovirastoihin, etenkin monissa sellaisissa maissa, joissa on hajautetut järjestelmät, tilastoja tuottavat useat tahot. Tilastovirastot tuleekin tässä ymmärtää laajasti kaikkina virallisia tilastoja tuottavina tahoina.)

3. Näillä suuntaviivoilla on kaksi päätavoitetta:

(i) yhdenmukaisempien käytäntöjen edistäminen eri maissa ja asianmukaisiin tarkoituksiin käytettävän yksikkötason tietojen saatavuuden parantaminen tutkijayhteisössä; ja

(ii) maiden tukeminen niiden pyrkimyksissä parantaa yksikkötason tietojen saatavuutta tarjoamalla niiden käyttöön nämä suuntaviivat ja niitä täydentävät malliesimerkit.

4. Asiakirjassa käytetään kautta linjan termiä ”yksikkötason tiedot”. Sillä voidaan viitata yksittäistä henkilöä, kotitaloutta, liikeyritystä tai muuta juridista henkilöä koskeviin tietoihin. Kyseessä voi olla tilastoviraston itsensä keräämä tieto tai muusta, esimerkiksi hallinnollisesta, lähteestä saatu tieto.

5. Suuntaviivoja laadittaessa on tunnustettu se tosiasia, että pääsy yksikkötason tietoihin toteutetaan eri maissa hieman erilaisin järjestelyin. Järjestelyihin vaikuttavat useat tekijät, kuten lainsäädäntö, yleinen mielipide ja tutkijayhteisön tukemiseen käytettävissä olevat resurssit. Edistyksellisessä tilastovirastossa nämä järjestelyt voivat olla aivan toisenlaiset kuin kehittymättömämmässä virastoissa. Joissakin maissa saatetaan myös ajatella, ettei tietosuojan varmistamiseksi tarvittavia järjestelmiä tai resursseja ole käytettävissä. Ei pidä odottaa, että joka maassa päädyttäisiin täsmälleen samanlaisiin järjestelyihin, vaikka suuntaviivojen toivotaankin johtavan nykyistä yhdenmukaisempaan käytäntöön.

6. On myös otettava huomioon, ettei kaikilla maille ole samanlaisia lähtökohtia. Erityisesti Itä-Euroopassa on maita, joissa ei ole perinteisesti ollut juurikaan tietosuojaa koskevaa lainsäädäntöä. Vaikka tilanne on muuttumassa, useissa tapauksissa lakimuutosten edellyttämä kulttuurin muutos voi kestää vielä pitkään.

7. Monissa maissa asiaa koskevaa lainsäädäntöä on jo olemassa. Myös Euroopan unionissa on tietosuojaa koskevaa lainsäädäntöä, joka sisältää lukuisia periaatteita ja sääntöjä. Niitä sovelletaan jo

monissa Euroopan talouskomission maissa, erityisesti EU-maissa. Tunnustamme, ettei voimassa olevaa lainsäädäntöä ole helppo muuttaa, ja että muutokset olemassa oleviin suuntaviivoihin edellyttävät yhteistyötä lukuisten sidosryhmien kanssa. Tilaisuuksia kuitenkin tarjoutuu aika ajoin, ja silloin näistä suuntaviivoista ja niihin liittyvistä malliesimerkeistä voi olla hyötyä tarvittavien muutosten määrittelyssä. Joissakin maissa suuntaviivat voivat toimia virikkeenä muutoksia koskeville keskusteluille ja sopimuksille.

8. Näitä suuntaviivoja koskevat kysymykset voi lähettää sähköpostitse Euroopan talouskomission tilastotoimistoon osoitteella confidentiality@unece.org.

II. MIKSI KANSALLISTEN TILASTOVIRASTOJEN PITÄISI TUKEA TUTKIJAYHTEISÖÄ?

9. Useimmissa maissa virallisia tilastoja kerätään paitsi valtionhallinnon tarpeisiin myös muun yhteiskunnan käyttöön. Tällainen on tilanne erityisesti demokratioissa, joissa virallisia tilastoja voidaan käyttää hallitusten toteuttaman politiikan ja ohjelmien tehokkuuden arviointiin, ts. ikään kuin yhteiskunnan peilinä.

10. Lainauksia Yhdistyneen kuningaskunnan avointa hallintoa koskevasta Valkoisesta kirjasta vuodelta 1993:

”Hallitus kerää virallisia tilastoja jakaakseen tietoja keskustelujen pohjaksi, päätöksentekoon ja tutkimustyöhön niin valtionhallinnossa kuin yhteisössä laajemminkin.”

”Ne tarjoavat objektiivisen näkökulman kansallisella tasolla tapahtuviin muutoksiin ja mahdollisuuden vertailuihin eri ajanjaksojen ja maantieteellisten alueiden välillä.”

”Vapaa pääsy virallisiin tilastoihin antaa kansalaiselle enemmän kuin pelkän kuvan yhteiskunnasta. Se avaa ikkunan maan hallituksen työhön ja aikaansaannoksiin osoittaen hallituksen toiminnan laajuuden yhteiskuntapolitiikan kaikilla eri osa-alueilla ja antaen mahdollisuuden yhteiskuntapoliittisten ohjelmien ja toimenpiteiden vaikutusten arviointiin.”

11. Tutkijayhteisöllä on erityisen tärkeä rooli politiikan analysoinnin ja julkisen keskustelun aktivoijana sekä hallituksen ohjelmien tehokkuuden arvioijana. Käyttökelpoiset analyysit edellyttävät kuitenkin laadukkaiden tilastotietojen saatavuutta. Jollei tutkijoilla ole käyttömahdollisuutta relevantteihin virallisen tilaston tietoihin, he pyrkivät usein itse keräämään tarvitsemansa tiedot. Tämä aiheuttaa sekä tietojen kerääjälle että tiedonantajalle ylimääräisiä kustannuksia, minkä lisäksi kerätyt tiedot ovat usein heikkolaatuisia.

12. Mahdollistamalla tutkijoiden pääsy yksikötason tietoihin voidaan myös hankkia lisäkatetta virallisten tilastojen keruusta aiheutuneille kustannuksille sekä arvokkaita näkemyksiä tietojen laadusta ja siitä, miten tilastotutkimusta voitaisiin parantaa tai laajentaa.

13. Mitä tutkijayhteisöllä tarkoitetaan? Yhteisöön luetaan tietenkin erilaisissa yliopistollisissa laitoksissa työskentelevät tutkijat. Lisäksi siihen kuuluvat kansalaisjärjestöjen ja kansainvälisten laitosten tutkijat. Osa yksikötason tietoja tarvitsevista tutkijoista työskentelee valtion rahoittamissa virastoissa ja laitoksissa. Näiden suuntaviivojen yhteydessä kaikki edellä mainitut tutkijat luetaan osaksi tutkijayhteisöä. Kuten suuntaviivoista jäljempänä ilmenee, eri näkökohtien painotukset saattavat kuitenkin vaihdella eri puolilla tutkijayhteisöä.

14. Seuraavissa kohdissa pyritään yhdistämään kansallisten tilastovirastojen ja tutkijayhteisön näkökulmat ja löytämään järjestelyt, jotka vastaisivat pääosin molempien osapuolten tarpeisiin. Tätä aihetta käsitellään yksityiskohtaisemmin luvussa 6.

Kansallisten tilastovirastojen näkökulma

15. Jotta tiedonantajat jatkossakin osallistuisivat tiedonkeruusiin, kansallisten tilastovirastojen on kyettävä säilyttämään heidän luottamuksensa. Tietosuojan turvaaminen on tämän luottamuksen kivijalka. Jos tiedonantajat luulevat tai havaitsevat, ettei tilastovirasto huolehdi heidän antamiensa tietojen tietosuojasta, heidän halukkuutensa yhteistyöhön tai (oikeiden) tietojen antamiseen todennäköisesti heikkenee. Yksi ainoa tapaus, etenkin jos se saa osakseen paljon median huomiota, voi vaikuttaa merkittävästi tiedonantajien yhteistyöhalukkuuteen ja siten virallisten tilastojen laatuun.

16. Tämä on kansallisten tilastovirastojen kannalta oleellisin asia, joskaan ei ainoa huolenaihe. Tärkeää on myös, että tilastovirastolla on riittävästi valtuuksia tukea tutkijoiden pääsyä yksikötason tietoihin, joko lain nojalla tai muulla tavoin saadulla valtuutuksella.

17. Joidenkin maiden tilastovirastot ovat huolissaan siitä, ovatko heidän keräämänsä tiedot riittävän laadukkaita laajempaan tiedonlevitykseen. Laatu saattaa olla riittävä tilastoihin mutta riittämätön kovin yksityiskohtaiseen analyysiin. Joissakin tapauksissa tilastoihin tehdään viimeistelyvaiheessa muutoksia, joita ei päivitetä perusaineistoihin. Seurauksena voi olla epä johdonmukaisuuksia yksikötason tietojen ja julkaistuihin aggregaattitietoihin perustuvien tutkimustulosten välillä.

18. Kansalliset tilastovirastot voivat olla huolissaan myös kustannuksista. Niitä aiheuttavat yksikötason tietoja sisältävien tiedostojen luomisen ja dokumentoinnin ohella käyttö- ja tietoturvajärjestelmien kehittäminen sekä tutkijayhteisöltä saapuvien tietopyyntöjen käsittely ja käyttö lupien jakaminen. Lisäksi uusia käyttäjiä on opastettava ymmärtämään monimutkaisia tiedostorakenteita ja muuttujien määritelmiä. Vaikka vastuu näistä kustannuksista on tilastovirastoilla, niille ei yleensä ole myönnetty lisäbudjettia näiden ylimääräisten töiden tekemiseen. Kaiken kaikkiaan on vielä todettava, ettei tutkijoidenkaan rahoitus riitä osallistumaan merkittävästi näihin kustannuksiin.

19. Kansalliset tilastovirastot ovat kuitenkin alkaneet tiedostaa yhä enemmän tutkijayhteisön tukemisen merkityksen sekä lisäarvon niiden omalle tiedonkeruulle ja -käsittelylle, kun kerättyjä tietoja hyödynnetään tehokkaasti tutkimuksessa. On erityisesti huomattava, että tilastotietoihin perustuvien havaintojen tuominen päättäjien ja suuren yleisön saataville on yleisen edun mukaista. Lisäksi, tilastotutkimusten tietojen entistä laajempi hyödyntäminen tällä tavoin voi tarjota lisäsuojaa näihin tilasto-ohjelmiin kohdistuvilta budjettileikkauksilta.

Tutkijayhteisön näkökulma

20. Tutkijayhteisön näkökulmasta yksikötason tietoihin perustuvan tutkimuksen tukemisen tulisi olla tärkeä osa mitä tahansa virallista tilastointijärjestelmää. Tästä koituisi muun muassa seuraavanlaisia etuja:

(i) yksikötason tiedot antavat päättäjille mahdollisuuden monitahoisten kysymysten asetteluun ja analysointiin. Esimerkiksi taloustieteissä aggregoitujen tilastotietojen analysointi ei anna riittävän tarkkaa kuvaa talouden toiminnasta tuottavuuden kasvuun vaikuttavien tekijöiden analysoimiseksi;

(ii) pääsy yksikötason tietoihin antaa analyytikoille mahdollisuuden laskea reunajakaumiin perustuvia vaikutuksia keskimääräisten vaikutusten lisäksi. Yksikötason tietojen avulla analyytikko voi esimerkiksi laatia monimuuttuja-regressioanalyysin, jossa tiettyjen muuttujien reunajakaumien vaikutuksia voidaan tarkastella erikseen;

(iii) yleisesti ottaen laajasti saatavilla olevat yksikötason tiedot antavat mahdollisuuden tärkeiden tutkimusten toistamiseen;

(iv) yksikötason tietojen saatavuus tutkimustarkoituksiin ja siitä saatava palaute voivat auttaa parantamaan tiedon laatua. Esimerkiksi Yhdysvalloissa tilastoinnista vastaava Bureau of the Census

on laatinut tutkijoiden käyttöön viralliset dokumentointimallit parantaakseen omien tutkimustensa laatua;

(v) yksikkötason tietojen saatavuus lisää kerätyistä tilastotiedoista saatavaa tuotosta ja siten tietojen keruusta kokonaisuudessaan saatavaa taloudellista arvoa.

21. Jos pääsy yksikkötason tietoihin on estetty, tutkijat saattavat kehittää ja panna toimeen omia tilastollisia tiedonkeruitaan, mikä lisää yhteiskunnalle aiheutuvaa vastaustaakkaa. Tästä (niin tietojen kerääjälle kuin tiedonantajillekin) aiheutuvien kustannusten lisäksi, kerätyt tiedot ovat usein laadultaan heikompia ja otokset ovat pienempiä kuin virallisissa otantatutkimuksissa. Tämä johtaa myös tutkimustulosten heikompaan laatuun. Hyväksytystä, virallisesta ja ensiluokkaisesta tietolähteestä on selkeää etua kaikelle tutkimustyölle, verrattuna vaihtoehtoon, jossa tutkijat käyttävät eri tietolähteitä eri aiheiden analysointiin. Kansallisilla tilastoviranomaisilla voi tässä suhteessa olla hyvin hyödyllinen rooli.

22. Tutkijat huomauttavat, etteivät he ole kiinnostuneita tunnistamaan aineistosta yksittäisiä henkilöitä, ja käytäntö on osoittanut tämän pitävän myös paikkansa. Tutkijat ovatkin sitä mieltä, että kansalliset tilastovirastot ovat olleet yleisesti liian konservatiivisia suhtautumisessaan yksikkötason tietojen käyttöoikeuksiin.

23. Yhdysvaltain tiedesäätiön (National Science Foundation) isännöimässä tutkimuksen luottamuksellisuutta koskevassa seminaarissa vuonna 2003 Peter Madsen viittasi ilmiöön nimeltä yksityisyyden paradoksi. Hän väittää, että "kova kiire varmistaa täydellinen yksityisyys tutkimusyhteyksissä johtaa paradoksiin, yhteiskunnallisen edun menetykseen, sen sijaan että edut lisääntyisivät". Hänen mukaansa käyttämällä käsitettä "yleishyödyllinen" voitaisiin saada aikaan toisenlaisia tuloksia:

"Ehkäpä tämä yleishyödyllisyyden käsite voisi saada ihmiset havaitsemaan, että vaikka heillä todellakin on oikeus yksityisyyteen, he ymmärtäisivät, että heillä on myös velvollisuus jakaa tietoa, jos sillä tavoin voidaan edistää yleistä etua."

Jotkut käyttävät ilmaisua "yksityisyyden vaje" (privacy deficit) kuvaamaan yksikkötason tietojen luovuttamiseen liittyviä yksityisyyttä koskevia ongelmia. Keskustelua voidaan sitten kohdistaa siihen, ovatko tutkimushankkeen tuomat edut tärkeämpiä kuin mahdollinen yksityisyyden vaje.

24. Tutkijayhteisö pitää tärkeänä myös tutkimusta nykyistä parempien tietosuojamenetelmien kehittämiseksi, sillä se lisäisi tietojen hyödynnettävyyttä. Kansalliset tilastovirastot lienevät tämän tutkimuksen merkityksestä samaa mieltä. Tämä tutkimus antaisi todennäköisesti kuitenkin vain osittaisen vastauksen kysymykseen tutkimustarkoituksiin käytettävien yksikkötason tietojen nykyistä paremmasta saatavuudesta, ja tutkijoita tuskin ilahduttaisi, jos tilastovirastot luottaisivat tietosuojassa pelkästään parantuneisiin tilastollisiin menetelmiin.

III. PÄÄPERIAATTEET

25. Yhdistyneiden kansakuntien virallisia tilastoja koskevassa kuudennessa peruseriaatteessa tilastojen luottamuksellisuus määritellään hyvin selvästi.

”Tilastoviranomaisten tilastojen laadintaa varten keräämää yksikkötasoista tietoa, riippumatta siitä, koskeeko tieto luonnollista tai juridista henkilöä, on käsiteltävä erittäin luottamuksellisesti ja käytettävä ainoastaan tilastotarkoituksiin.”

Kaikkien yksikkötason tietojen saatavuutta koskevien periaatteiden tulee olla yhdenmukaisia tämän peruseriaatteen kanssa.

26. Yksikkötason tietojen luottamuksellisuuden varmistamiseksi tulisi noudattaa seuraavia periaatteita. Kutakin periaatetta käsitellään tarkemmin seuraavissa kappaleissa.

Periaate 1: Virallisiin tilastotarkoituksiin kerättyjä yksikkötason tietoja on asianmukaista käyttää tutkimustyötä tukevaan tilastolliseen analyysiin, mikäli turvataan tietojen luottamuksellisuus.

Periaate 2: Pääsy yksikkötason tietoihin tulisi sallia ainoastaan silloin, kun niitä käytetään tilastotarkoituksiin.

Periaate 3: yksikkötason tietojen käyttöoikeuksien on oltava linjassa asianomaisen lainsäädännön ja muiden säännösten kanssa, jotta luovutettavien tietojen luottamuksellisuus voidaan turvata.

Periaate 4: Tutkijoiden käyttö lupien myöntämisessä käytettävistä toimintatavoista sekä yksikkötason tietojen käyttötarkoituksista ja käyttäjistä tulisi tiedottaa avoimesti ja julkisesti.

27. Yksikkötason tietojen käyttöoikeuden myöntäminen tutkimukseen ei ole ristiriidassa YK:n kuudennessa peruseriaatteen kanssa, niin kauan kun ei ole mahdollista tunnistaa tiedon viittaavan johonkin tiettyyn yksilöön. Periaate 1 ei kuitenkaan velvoita sallimaan pääsyä yksikkötason tietoihin. Päätökset tietojen käyttöoikeuksista tekee kansallinen tilastovirasto. Voi olla myös muita syitä (esim. laatu), joiden vuoksi pääsyä yksikkötason tietoihin ei voida sallia. Voi myös olla tiettyjä henkilöitä tai laitoksia, joille ei olisi asianmukaista antaa pääsyä näihin tietoihin.

28. Periaatetta 2 koskien on syytä selventää toisaalta käsitteitä tilastokäyttö tai analyysikäyttö ja toisaalta käsitettä hallinnollinen käyttö. Tilasto- tai analyysikäytössä tavoitteena on laatia tiettyyn (luonnollisista tai juridisista henkilöistä koostuvaan) ryhmään viittaavia tilastoja. Hallinnollisessa käytössä tavoitteena on sen sijaan saada tietoa jostakin tietystä luonnollisesta tai juridisesta henkilöstä, jotta pystytään tekemään päätös, joka voi aiheuttaa kohdeyksikölleen joko etua tai haittaa. Jotkut tietopyynnöt voivat esimerkiksi olla sinänsä laillisia (tuomioistuimen päätös) mutta vastoin tätä periaatetta. Yleisen luottamuksen säilyttämiseksi viralliseen tilastointijärjestelmään tällaisiin pyyntöihin on vastattava kielteisesti. Jos yksikkötason tietoja ei käytetä tilasto- tai analyysitarkoituksiin, pääsyä tietoihin ei tulisi sallia. Eettiset lautakunnat tai vastaavat organisaatiot voivat antaa ohjeita tilanteisiin, joissa vallitsee epävarmuutta siitä, pitäisikö pääsy tietoihin sallia vai ei.

29. Tutkijat tarvitsevat yksikkötason tietoja tutkimustarkoituksiin. Tutkimus voi edellyttää erilaisten kokonaistilastojen laadintaa, tilastollisten jakaumien esittämistä, tilastollisten mallien kehittämistä tai osajoukkojen välisten tilastollisten erojen analysointia. Nämä käyttötarkoitukset vastaavat tilastotarkoituksia. Siinä määrin kun edellä mainitut ovat yksikkötason tietojen tosiasiallisia käyttötapoja, niiden voidaan sanoa vastaavan myös tutkimustarkoituksia.

30. Periaatetta 3 koskien todettakoon, että luottamuksellisuuden suojaamiseksi tarvittavan juridisen perustan tulisi olla kunnossa ennen yksikkötason tietojen luovuttamista. Juridista perustaa on lisäksi täydennettävä hallinnollisilla ja teknisillä toimenpiteillä, jotta voidaan säädellä tietoihin pääsyä ja varmistaa yksilöä/yksikköä koskevan tiedon pysyminen salassa. Näiden järjestelyjen (laki tai sitä täydentävä asetus, määräys tms.) olemassaolo ja näkyvyys ovat tarpeen yleisen luottamuksen lisäämiseksi siihen, että tietoja käytetään asianmukaisesti. Lainsäädännön tarjoama perusta on selkeästi suositeltavin, mutta joissakin maissa se ei ole mahdollinen, jolloin tulisi ottaa käyttöön jokin muu hallinnollinen järjestely. Juridisesta (tai muusta) järjestelystä tulisi myös sopia kansallisten yksityisyyden suojasta vastaavien viranomaisten kanssa niissä maissa, joissa sellaisia on, ennen kuin säädetään asiaa koskeva laki. Jos tällaisia viranomaisia ei ole, yksityisyyden suojaa koskevia asioita saattaa valvoa epävirallisemmin jokin maassa toimiva kansalaisjärjestö. On viisasta hankkia juridiselle tai muulle järjestelylle kansalaisjärjestön tuki, tai ainakin selvittää järjestölle suurinta huolta aiheuttavat asiat.

31. Joissakin maissa yksikkötason tietojen luovuttamista koskevaa lainsäädäntöä ei ole. Sellaisessa tapauksessa yksikkötason tietojen luovuttamiselle tulisi hankkia vähintään jokin muu valtuutus. Lainsäädäntö on kuitenkin suositeltavin vaihtoehto.

32. Periaate 4 on tärkeä yleisen luottamuksen lisäämiseksi siihen, että yksikkötason tietoja käytetään asianmukaisesti ja että päätökset niiden luovuttamisesta tehdään objektiivisin perustein. On kansallisten tilastovirastojen harkinnassa, myönnetäänkö käyttöoikeus ja kenelle ja miten yksikkötason tietoja luovutetaan. Päätösten tulisi kuitenkin olla läpinäkyviä. Tilastoviraston omat verkkosivut on tehokas tapa varmistaa asianmukainen käytäntö ja myös jakaa tietoa siitä, miten luovutettujen tietojen pohjalta laadittuihin tutkimusraportteihin pääsee käsiksi.

IV. ASIAA KOSKEVA LAINSÄÄDÄNTÖ

33. Yksikötason tietojen luovuttamista koskevan lainsäädännön olemassaolo on hyvin tärkeää, kuten periaate 3 osoittaa (ks. luku 3). Tähän on useita syitä:

(i) yleisen luottamuksen varmistaminen käytäntöä kohtaan, ts. siihen, että on olemassa lainsäädännön asettamat rajat sille, mitä voidaan ja mitä ei voida tehdä;

(ii) kansallisten tilastovirastojen ja tutkijoiden välisen yhteisymmärryksen aikaansaaminen noudatettavasta käytännöstä;

(iii) yhdenmukaisuuden aikaansaaminen käyttölujarahakemusten käsittelyssä; ja

(iv) perustan luominen rikkomusten käsittelyyn.

34. Lainsäädännön ei tarvitse olla primäärioikeutta. Yksityiskohdat voivat sopia paremmin esimerkiksi asetuksiin tai määräyksiin, joilla on kuitenkin tiettyä lainsäädännöllistä voimaa. Ellei lainsäädäntöä ole, on olennaisen tärkeää hankkia jokin muu valtuutus. Kansallisen tilastoviraston maine vaarantuu, ellei yksikötason tietojen luovuttamiselle edes anonymisoituina saada jonkinlaista valtuutusta.

35. On tärkeää, että lainsäädäntö (tai valtuutus) kattaa seuraavat näkökohdat:

(i) mitä voidaan ja mitä ei voida tehdä ja mihin tarkoitukseen;

(ii) luovuttamisen ehdot; ja

(iii) seuraamukset näiden ehtojen rikkomisesta.

36. Malliesimerkissä 1 esitellään pääkohtia Australian ja malliesimerkissä 2 Suomen yksikötason tietojen luovuttamista koskevasta lainsäädännöstä.

V. MENETELMÄT TUTKIJAYHTEISÖN TUKEMISEKSI

37. Kansalliset tilastovirastot voivat tukea tutkimusta monin eri tavoin. Eri menetelmät esitetään ensin tiivistetysti alla olevissa taulukoissa. Menetelmiä käsitellään laajemmin taulukoiden jälkeen tulevissa tekstiosuuksissa. Lisävalaistusta eri menetelmiin saa asiakirjan liitteinä olevista malliesimerkeistä.

38. On syytä mainita eräs seikka, jonka merkitystä ei aina täysin ymmärretä. Yksikkötason tietoja sisältävät tiedostot voidaan anonymisoida poistamalla niistä nimet ja osoitteet sekä muilla toimenpiteillä (esim. karkeistamalla maantieteelliset yksityiskohtat), jotta varmistetaan, että yksilöiden tunnistaminen on erittäin epätodennäköistä, kun tiedostoja katsotaan erikseen. Tätä voidaan kutsua spontaanin tunnistamisen eliminoinniksi. Julkisella ja yksityisellä sektorilla on kuitenkin myös muita yksikkötason tiedostoja, joista yksilöt voivat joskus olla tunnistettavissa. Tutkimusten mukaan yhdistämällä tilastolliseen malliin perustuen tilastovirastojen yksikkötason tiedostoja tällaisiin tiedostoihin on mahdollista tunnistaa aineistossa esiintyviä harvinaisia tapauksia. Jos tilastoviraston yksikkötason tiedostot sisältävät paljon yksityiskohtaista tietoa, tunnistettavissa olevien tietojen määrä voi olla melko huomattava. Suhteellisesti mitattuna harvinaisia tapauksia esiintyy enemmän pienissä maissa. Näitä riskejä ei tilastovirastoissa aina täysin ymmärretä. Riskejä voidaan kuitenkin pienentää käsittelemällä tilastovirastojen yksikkötason tiedostoja erilaisin tekniikoin, joita ovat esimerkiksi tietojen muokkaaminen tai arvojen vaihtaminen yksiköiden välillä.

(i) Kansallisten tilastovirastojen ulkopuolella käytettävät tilastotuotteet:

Tiedonlevityskanava	Huomautuksia
Tilastotaulukot ja tietokuutiot	Nämä voivat sisältää sekä vakiotaulukoita että nimenomaan tutkijan tarpeisiin tuotettuja taulukoita (tai analyyseja). Jotkut virastot luovuttavat hyvin yksityiskohtaisia matriiseja, joita nimitetään tietokuutioiksi ja joita tutkijat voivat muokata omien tarpeidensa mukaan. Hyvin yksityiskohtaisten taulukoiden tai tietokuutioiden tietosuojariskit voivat olla samaa luokkaa kuin <u>yksikkötason tiedoilla</u> .
Anonymisoidut yksikkötason aineistot - Julkiseen käyttöön tarkoitetut tiedostot (PUFs)	Nämä ovat suuren yleisen käyttöön kansallisten tilastovirastojen ulkopuolella tarkoitettuja yksikkötason tiedostoja. Tiedostot on anonymisoitu ja niiden levityksessä käytetään usein jotakin välinettä, esim. CD-ROM-levyä, ja joskus ne kulkevat tietoarkiston kautta. (Huom. Anonymisoinnilla viitataan nimien ja osoitteiden poistamiseen lisäksi myös muihin toimenpiteisiin (esim. maantieteelliseen sijaintiin viittaavien yksityiskohtien karkeistamiseen), joilla varmistetaan, että yksilöiden tunnistaminen on erittäin epätodennäköistä. Julkiseen käyttöön tarkoitettujen tiedostojen tietosuojan tulisi olla sellaisella tasolla, että se estää myös tietojen välillisen tunnistamisen tiedostoa muihin tiedostoihin yhdistettäessä. Julkiseen käyttöön tarkoitetut tiedostot ovat joissakin maissa eräs tapa levittää tietoa tutkijoille.

<p>Anonymisoidut yksikkötason aineistot - Käyttölupaa edellyttävät tiedostot</p>	<p>Käyttölupaa edellyttävät tiedostot on myös anonymisoitu, mutta toisin kuin julkiseen käyttöön tarkoitetut tiedostot, ainoastaan hyväksytyillä tutkijoilla on niihin käyttöoikeus. Ennen tiedostojen luovuttamista tutkijoiden on allekirjoitettava sitoumus tai sopimus. Vaikka tiedostojen ilmoitettaisiin olevan yleisesti saatavilla, niitä ei luovuteta ennen kuin tutkija on toimittanut vaaditun sitoumuksen tai sopimuksen. Vaikka tiedostot anonymisoidaan ja muillakin toimenpiteillä varmistetaan, että yksilöiden tunnistaminen on erittäin epätodennäköistä, kun tiedostoja katsotaan erikseen, tiedostot saattavat sisältää mahdollisesti tunnistettavaa tietoa, mikäli tiedostot yhdistetään toisiin tiedostoihin. Tämä on eräs syy siihen, että vastaanottajalta vaaditaan etukäteen sitoumus tai sopimus. Kansalliset tilastovirastot voivat asettaa tutkijoille myös muita tiedostojen käyttöä koskevia ehtoja.</p>
--	---

(ii) Palveluikkuna, jonka kautta tutkijat voivat jättää tietopyyntöjään

Palvelu

Huomautuksia

<p>Etäyhteyksikäyttö</p>	<p>Monissa maissa tutkijoille järjestetään parhaillaan mahdollisuuksia tuottaa yksikkötason tiedoista tilastollisia tutkimustuloksia verkkoyhteyden välityksellä, ilman että tutkijat varsinaisesti ”näkevät” käyttämiään yksikkötason tietoja. Etäkäyttöön liittyvien lisävarmistusten ansiosta sekä sen vuoksi, etteivät käytettävät yksikkötason tiedot tosiasiallisesti siirry pois tilastovirastosta, voidaan tutkijoiden käyttöön tarjota entistä yksityiskohtaisempia tietoja.</p>
--------------------------	---

(iii) Järjestelyt, joilla tutkijoille annetaan mahdollisuus työskennellä kansallisen tilastoviraston tiloissa

Palvelu

Huomautuksia

<p>Datalaboratoriot</p>	<p>Datalaboratorioissa tutkijoille tarjotaan pääsy helpommin tunnistettaviin yksikkötason tietoihin kansallisen tilastoviraston tiloissa. Yleensä vaaditaan aukotonta kirjausketjua aineiston käytöstä ja tutkimus tehdään tilastoviraston tiukassa valvonnassa. Pääsy yksityiskohtaisempien tietojen ääreen aiheuttaa tutkijalle jonkin verran epämukavuutta, koska hänen on työskenneltävä tilastoviraston tiloissa tai erillisyyksikössä.</p>
-------------------------	--

Tilastotaulukot ja tietokuutiot

39. Tilastotaulukot ovat yhä taloudellisin tapa tyydyttää monien tutkijoiden tietotarpeet, joten niiden merkitystä ei pidä aliarvioida. Tietokuutioiden (hyvin yksityiskohtaisten moniulotteisten taulukoiden) käyttöönotto on lisännyt tilastotaulukoiden käyttökelpoisuutta tutkimuksessa, koska tutkijat voivat muokata tietokuutioita omien tarpeidensa mukaisiksi.

40. Alankomaiden tilastovirasto oli ensimmäisiä organisaatioita, joissa tietokuutiot otettiin käyttöön. Malliesimerkissä 3 kuvataan tietokuutioiden käyttöä keskeisenä osana viraston tiedonlevitysstrategiaa.

41. Tietosuojaongelmat koskevat myös tilastotaulukoita ja tietokuutioita. Esimerkiksi tilastointia koskeva lainsäädäntö edellyttää yleensä, ettei julkaistava taulukkoaineisto saa sisältää tunnistettavissa olevaa tietoa. Salassa pidettävien tietojen suojaamisesta huolehditaan jo ennen tietojen luovuttamista. Tilastotaulukoiden tietojen salaamiseen on olemassa tietokoneohjelmia, ja uusia tietosuojamenetelmiä kehitetään kaiken aikaa.

Anonymisoidut yksikkötason tiedostot – julkiseen käyttöön tarkoitetut tiedostot

42. Tutkijat pitävät näitä erittäin arvokkaana palvelumuotona. Tietojen yhdistelymahdollisuuksien lisääntyttä kehitys näyttäisi kuitenkin kulkevan siihen suuntaan, että julkiseen käyttöön tarkoitettujen tiedostojen osuus tiedon levittämisessä vähenee, ja käyttö lupaa edellyttävien anonymisoitujen tiedostojen, etäkäytön ja datalaboratorioiden merkitys tutkimustyössä kasvaa. Luvanvaraisten tiedostojen käytössä luotetaan tunnistamisen estämiseksi tehtävien toimenpiteiden ohella tutkijoiden antamiin sitoumuksiin tai sopimuksiin siitä, etteivät he yritä tunnistaa yksikkötason tiedoista yksilöitä tai yksiköitä. Nämä sitoumukset ovat usein keskeisessä asemassa käyttö lupaa edellyttäviä tiedostoja luovutettaessa (ks. seuraava tekstiosio).

43. Vaikka kansalliset tilastovirastot yleisesti tarjoavat kaikille yhtäläiset oikeudet tilastojen käyttöön, pääsy yksikkötason tietoihin muodostaa poikkeuksen. Sellaisiin käyttäjiin, joiden tutkimuksella ei ole vankkoja perusteita, tai joilla on pääsy tietokantoihin, joiden tietoihin anonymisoituja yksikkötason tietoja olisi helppo yhdistää, voidaan suhtautua eri tavalla kuin muihin.

44. Tämä ei kuitenkaan koske julkiseen käyttöön tarkoitettuja tiedostoja, jotka pyritään pitämään saatavilla mahdollisimman laajasti. Tutkijat ovat painottaneet yleisessä käytössä olevien tiedostojen merkitystä. Näitä tiedostoja arvostetaan ja käytetään laajasti niissä maissa, joissa niitä on. Paljastamista yrittävälle henkilölle tuskin kuitenkaan tuottaisi hankaluuksia tunnistaa tiedostoista yksilöitä yhdistämällä tiedot tilastolliseen malliin perustuen muihin tietokantoihin, erityisesti jos on kyse väkiluvultaan pienistä maista, joissa on väestörekisterit. Ennen julkiseen käyttöön tarkoitettujen tiedostojen luovuttamista tulisi tutkia huolellisesti luovutuksen ehdot, jotta voidaan paremmin hallita tietosuojaan liittyvät riskit. Esimerkiksi sopimus voisi olla yksi luovutuksen ehdoista. On mahdollista kehittää menettely, jossa tietojen luovutuksen ehtona on etukäteen allekirjoitettava sopimus, myös silloin, kun julkiseen käyttöön tarkoitettu tiedosto yleiskäyttötiedosto on saatavilla Internetin välityksellä. Yleisesti ottaen tietosuojariski on paljon suurempi väkiluvultaan pienissä maissa. Tästä johtuen tutkijoiden ei pidä odottaa kaikkien maiden luovuttavan julkiseen käyttöön tarkoitettuja tiedostoja.

45. Tunnistamisen riskiä voi pienentää käyttämällä erilaisia tekniikoita, kuten arvojen vaihtoa ja tietojen muokkaamista. Näitä menetelmiä käytetään yleisesti esimerkiksi Yhdysvalloissa. Haittapuolena on, että menetelmät voivat heikentää tiedoston sisältämien tietojen käytettävyyttä.

46. Malliesimerkissä 4 kuvataan julkiseen käyttöön tarkoitettujen tiedostojen yleiskäyttötiedostojen luovuttamista Yhdysvalloissa. Mielenkiintoinen yksityiskohta on yhteiskuntatieteellisen tietoarkiston rooli tutkijoille myönnettävien tiedostojen käyttöoikeuksien hallinnassa.

47. Yksikkötason aineistojen anonymisoinnissa käytettävistä menetelmistä on saatavilla runsaasti kirjallisuutta. Hyvä yhteenveto on esitetty julkaisussa Willenburg, L. & de Waal, T. (2001), Elements of Statistical Disclosure Control. Ohjelmistopaketti μ -ARGUS sisältää yksikkötason tietojen tietosuojaohjelman lukuisine menetelmineen.

Anonymisoidut yksikkötason aineistot – käyttö lupaa edellyttävät tiedostot

48. Tässä luovutusmenettelyssä tietyille käyttäjille myönnetään lupa käyttää anonymisoituja yksikötason aineistoja sen jälkeen, kun he ovat allekirjoittaneet asiaa koskevan sitoumuksen tai sopimuksen. Vaikka tiedostot on anonymisoitu, eikä tiedoista voi tunnistaa yksilöitä, kun tiedostoja katsotaan erikseen, tunnistaminen saattaa olla mahdollista yhdistämällä tiedot tilastolliseen malliin perustuen muihin tiedostoihin. Juuri tämän vuoksi käyttäjiltä vaaditaan käyttö lupaa. Käyttö lupaan liittyy erilaisia ehtoja, jotka voidaan eritellä tutkijan tai hänen tutkimuslaitoksensa allekirjoittamassa sitoumuksessa tai sopimuksessa. Ehdot voivat vaihdella maittain tai jopa tutkijoittain kulloinkin kyseessä olevan tutkimushankkeen tai tutkijan taustan mukaan.

49. Ehdot voivat sisältää joitakin tai kaikki seuraavista asioista:

- tutkijan suostumus tietojen luovutusta koskevien ehtojen noudattamiseen;
- tutkijan vakuutus, ettei hän yritä paljastaa aineistosta yksilöitä tai organisaatioita;
- tietoja käytetään ainoastaan tilastointi tai tutkimustarkoituksiin;
- yksikötason tietoja ei luovuteta edelleen muille henkilöille;
- yksikötason tiedot palautetaan kansalliseen tilastovirastoon tutkimushankkeen päätyttyä; ja
- tutkijan vakuutus, ettei hän yritä ilman erillistä lupaa yhdistää tiedostoja tilastolliseen malliin perustuen muihin tietokantoihin.

50. On hyvien käytäntöjen mukaista huolehtia, että tällaisella sitoumuksella olisi jonkin asteinen juridinen asema, esimerkiksi säätämällä siitä lailla. Se sallisi ryhtymisen oikeustoimiin sitoumuksen ehtoja rikottaessa. Tämä ei sulje pois muita toimenpiteitä, joihin voidaan ryhtyä ehtoja rikottaessa, kuten tietopalvelun lopettamista kyseiselle tutkijalle ja/tai hänen edustamalleen tutkimuslaitokselle. Tätä aihetta käsitellään lisää luvussa 7.

51. Käyttö lupaa edellyttävien tiedostojen välityksellä pitäisi olla mahdollista luovuttaa enemmän tietoja kuin julkiseen käyttöön tarkoitettujen tiedostojen välityksellä, mikäli menettelyyn voidaan sitoumuksin saada lisää luotettavuutta ja tietojen salassa pysyminen voidaan siten turvata. Tämä koskee nimenomaan niitä tapauksia, joissa tiedot saattavat olla muihin tiedostoihin yhdistettäessä välillisesti tunnistettavissa.

52. Käyttö lupaa edellyttävien tiedostojen luovuttamista kuvataan malliesimerkeissä 5 (Australia), 6 (Alankomaat) ja 7 (Ruotsi).

Mahdollisesti tunnistettavissa olevat yksikötason tiedot

53. Joissakin maissa luovutetaan tilasto- tai tutkimustarkoituksessa ulkopuoliseen käyttöön yksikötasonaineistoja, jotka sisältävät mahdollisesti tunnistettavissa olevaa tietoa. Tosin luovutus edellyttää ankaria ehtoja sisältävää käyttö lupaa. Käyttö luvan tulisi sisältää tietojen käyttöä koskevat ehdot, ja menettely tulisi olla yksityiskohtaisesti määriteltynä laissa. Tiedonantajien ja suuren yleisön luottamuksen säilyttäminen edellyttää menettelyltä ehdottoman tiukkaa sääntelyä. Etäyhteyskäyttö ja datalaboratoriot ovat muita tämän tyyppisiin tilanteisiin soveltuvia vaihtoehtoja.

Etäyhteyskäyttö

54. Etäkäyttöjärjestelmien merkitys tiedonlevityksessä kasvaa kaiken aikaa. Maiden välillä on kuitenkin huomattavia eroja etäkäytön toteutuksessa. Kaikille menetelmille on tyypillistä, että tutkijoille ei anneta pääsyä varsinaisiin yksikötason tietoihin vaan tietojen käyttöön perustuva toimeksianto voidaan lähettää etäyhteyden avulla Internetin välityksellä. Usein kansallisen tilastoviraston ja tutkijan tai hänen edustamansa tutkimuslaitoksen välillä on tehty asiaa koskeva sopimus.

55. Esimerkiksi Kanadan tilastovirasto tarjoaa tutkijoiden käyttöön synteettisiä aineistoja ja antaa heidän suorittaa ajoja koko aineistosta tietoverkon välityksellä. Kanadan tilastovirasto suorittaa toimeksiantojen ajot "offline" ja lähettää tulokset takaisin verkon välityksellä varmistettuaan tietosuojan. Vaikka myös Australian tilastovirastossa käytetään vastaavaa järjestelmää, maiden välillä on joitakin tärkeitä eroja. Yksikkötason tiedostojen tiedot salataan spontaanin tunnistamisen estämiseksi ennen kuin ne annetaan käyttäjän saataville etäyhteyden kautta. Etäkäyttötiedostojen koekäyttö on kuitenkin sallittu ja pieni määrä anonymisoituja yksikkötason tietoja on sallittua ladata verkosta poikkeavien havaintojen yms. tarkastelua varten. Lopputulos tarkastetaan ennen sen lähettämistä tutkijalle. Tällä hetkellä järjestelmä toimii eräajoperiaatteella mutta interaktiiviseen versioon siirtymistä harkitaan. Tanskan tilastoviraston käyttämät menetelmät eroavat edellisistä. Käytössä on online-järjestelmä, jossa tutkijat voivat tehdä analyysejä koko yksikkötason tietoja sisältävästä tiedostosta. Varsinaisia yksikkötason tietoja he eivät pysty lataamaan. Riskejä pyritään hallitsemaan lisäksi tutkimuslaitosten kanssa tehtävien sopimusten avulla sekä sopimusrikkomusten yhteydessä käytettävien sanktioiden (erityisesti kieltämällä tietoihin pääsy vastaisuudessa).

56. Käytössä on kahdenlaisia etäkäyttöjärjestelmiä.

(a) Etäyhteyden välityksellä lähetettävät toimeksiannot, missä tutkija lähettää tilastovirastoon ohjelmakoodin, jolla tuotettu aineisto toimitetaan hänelle myöhemmin Internetin välityksellä.

(b) Varsinainen etäyhteyksikäyttö, jossa tutkija tekee analyysin itse ja näkee tuloksen välittömästi tietokoneensa näyttöruudulta.

Monissa maissa on vaihtoehdon (a) tyyppisiä etäkäyttöjärjestelmiä ja, Tanskaa lukuun ottamatta, vaihtoehdon (b) tyyppiset järjestelmät ovat vielä kehitteillä. Eri järjestelmien hyväksyttävyyys todennäköisesti vaihtelee maasta toiseen.

57. Vaikka etäkäyttöjärjestelmiä on käytössä vasta muutamissa maissa ja vaikka mallit ja lähestymistavat vaihtelevat, kuten edellä kuvattiin, etäkäytöstä tähän mennessä saadut kokemukset ovat olleet yleisesti myönteisiä.

58. Kustannusnäkökulmasta katsottuna etäyhteyksikäyttö on suositeltavampaa kuin datalaboratoriot (ks. jäljempänä), sillä etäkäytön valvonta ei ole niin työvoimavaltaista kuin datalaboratorioissa vaadittava tietojen käytön valvonta.

59. Jos etäkäyttöjärjestelmissä ei pystytä täysin poistamaan tunnistamisen riskiä, tulisi tutkijoiden kanssa lisäksi tehdä sopimus, jotta varmistetaan, että he ovat täysin tietoisia velvollisuuksistaan. On hyvien käytäntöjen mukaista jakaa käyttöoikeuksia vain niille tutkijoille, jotka ovat allekirjoittaneet jonkin tyyppisen sopimuksen, jossa käytön ehdot on määritelty. On myös tärkeää huolehtia etäkäyttäjien kouluttamisesta sekä järjestelmän käytön säännöllisestä seurannasta ja tarkastamisesta.

60. Käytössä olevia etäkäyttöjärjestelmiä kuvataan malliesimerkeissä 8 (Kanada), 9 (Australia) ja 10 (Tanska).

Datalaboratoriot

61. Datalaboratorioita on käytetty joissakin tilastovirastoissa jo useita vuosia. Niiden etuja ovat tehokas tunnistamisen riskin hallinta ja tutkijoiden pääsy aineistoihin, jotka sisältävät yksikkötason aineistoja, joiden luovuttaminen muulla tavoin ei ole mahdollista. Riittävä tietosuoja edellyttää lisäksi käytölle asetettavia ehtoja. Datalaboratorioiden suurimpana haittapuolena on pidetty tutkijoille aiheutuvaa hankaluutta, joka voi johtua esimerkiksi sitä, että tutkija joutuu käyttämään tietojen analysoinnissa hänelle ennestään tuntematonta tietokoneohjelmaa. Muihin vaihtoehtoihin verrattuna laboratoriot ovat myös kansallisille tilastovirastoille kalliita ylläpitää.

62. Jotkut tilastovirastot (esim. Kanadan tilastovirasto) ovat perustaneet datalaboratorioille omia erillisiä toimitiloja, jotka sijaitsevat tutkijoiden kannalta edullisemmissä paikoissa. Nämä

tutkimuskeskukset (joita kutsutaan joskus nimellä Research Data Centres) voivat myös olla varsin kallis vaihtoehto, ellei tilastovirastolle ole myönnetty niitä varten erityistä rahoitusta.

63. Mitä ehtoja tietojen saatavuudelle datalaboratoriossa tulisi sitten asettaa? Tärkeimpiä ehtoja voisivat olla seuraavat: (a) tutkimuksen tarjoaman yleisen hyödyn kirjaaminen, (b) selvitys siitä, miten tulokset saatetaan yleisesti saataville, (c) todisteet sille, että tutkija toimii vilpittömässä mielessä, (d) laillisesti sitova sopimus tai sitoumus ja (e) kansallisen tilastoviraston valvonnan edellytykset.

64. Datalaboratoriotoimintaa eri maissa kuvataan malliesimerkeissä 11 (Kanada), 12 (Yhdysvallat), 13 (Alankomaat), 14 (Uusi-Seelanti), 15 (Italia) ja 16 (Brasilia).

Tutkijan palkkaaminen kansallisen tilastoviraston tilapäiseksi työntekijäksi

65. Yksi tapa tarjota tutkijoille pääsy yksikkötason aineistoihin on heidän palkkaamisensa kansallisen tilastoviraston tilapäisiksi työntekijöiksi, jolloin heitä koskevat samat salassapitovelvoitteet kuin tilastoviraston muuta henkilökuntaa. Tämä järjestely sopii ainoastaan tilanteisiin, joissa tutkija osallistuu tilastoviraston hankkeeseen – muussa tapauksessa järjestely saattaisi vaikuttaa epäilyttävältä. Jos tällaiseen epäilyttävään järjestelyyn olisi kuitenkin ryhdytty ja asia tulisi julkisuuteen, luottamus tilastovirastoon heikkenisi.

66. Tutkijan palkkaamiseen voidaan ryhtyä kansallisen tilastoviraston aloitteesta, jos tutkijan katsotaan pystyvän tuomaan tilastoviraston hankkeeseen erityisosaamista ja edistämään tiedon hyödyntämistä. Aloite voi tulla myös tutkijalta itseltään. Tutkijan osallistuminen kansallisen tilastoviraston hankkeeseen on helpompi osoittaa, jos yhteisestä työstä laaditaan julkaistava raportti (joka voidaan lisäksi ulkoasultaan erottaa kansallisen tilastoviraston tavanomaisista julkaisuista). Tällaisista järjestelyistä koituu tutkijoille luonnollisesti etuja. Voidaan myös sopia, että tutkijat julkaisevat tutkimustyönsä tulokset itse jollakin muulla tavalla (tilastoviraston kanssa asiasta neuvoteltuaan).

Yrityksiä koskeva tieto

67. Yrityksiin, joihin luetaan myös maatalousyrittäjät, liittyy tiettyjä erityisnäkökohtia. Yritykset, nimenomaan suuryritykset, ovat kotitalouksia tai yksilöitä helpommin tunnistettavissa, erityisesti spontaanin tunnistuksen kautta. Tämä johtuu siitä, että niiden ominaisuuksien jakauma on vinompi. Lisäksi moniin yrityksiä koskeviin tutkimuksiin suurimmat yritykset valitaan mukaan hyvin suurella todennäköisyydellä. Joissakin maissa yritystietoja sisältävät tietokannat ovat helpommin saatavilla, jolloin tietojen yhdistäminen on mahdollista. Monet yliopistojen tutkijat toimivat myös yritysten konsultteina, jolloin yrityksiä koskevien tietojen tutkiminen vilpittömässä mielessä ei välttämättä sovi yhteen tällaisen konsultin roolin kanssa (koska tutkijoita ei voida "aivopestä" tiedoista, joita he ovat saaneet tutkimustyössään). Myös taloudellinen kilpailu tai jopa turvallisuusseikat voivat rajoittaa joidenkin maiden mahdollisuuksia jakaa tunnistettavaa yrityksiä koskevaa tietoa muiden maiden tutkijoille.

68. Tutkijoille annettavien käyttöoikeuksien näkökulmasta suurin ero kotitalouksia tai yksilöitä koskevan tiedon ja yrityksiä koskevan tiedon välillä on, että parhaimman suojan tarjoavat tiedonlevityskanavat ovat yrityksiä koskevien tietojen luovutuksessa erityisen tärkeitä.

69. Tiedonlevityskanaviin liittyen:

- tilastotaulukot ovat yhä käyttökelpoisia, tosin suuremman tunnistusriskin vuoksi yrityksistä ei yleensä ole saatavilla kovin yksityiskohtaista tietoa sisältäviä taulukoita;
- anonymisoidut yksikkötason aineistot soveltuvat ehkä vain pienimpiin yrityksiin. Joissakin tutkimuksissa pienet yritykset voivat olla tutkijoiden erityisen kiinnostuksen kohteina. Silloinkin joitakin tietoja (esim. talouteen liittyviä lukuja) on tarpeen "muunnella" jotta estetään yhdistäminen muihin tietokantoihin (esim. verotustietoihin). Vaihtoehtona on lukujen esittäminen vaihteluvälinä. Anonymisoiduista yksikkötason tiedostoista lienee siten vain rajallista hyötyä;
- samantyyppisistä syistä myös etäyhteyskäyttö soveltuu ehkä vain pienimpiä yrityksiä koskeviin yksikkötason tiedostoihin. Etäkäyttöjärjestelmät antavat kansallisille tilastovirastoille kuitenkin mahdollisuuden yhdistämisen riskin hallintaan, joten tietojen "muunteleminen" tietosuojan takaamiseksi ei välttämättä ole tarpeen. Suuryritysten ollessa kyseessä tietojen salaaminen voi olla vaikeaa, vaikka tutkijat eivät pääsekään suoraan käsiksi yksikkötason tietoihin.
- datalaboratoriot soveltunevat parhaiten yrityksiä koskevaa yksikkötason tietoa sisältävien tiedostojen luovuttamiseen. Laboratorioita käytetään esimerkiksi Alankomaiden tilastovirastossa (ks. malliesimerkki 13).

70. Joitakin tutkimushankkeita on mahdollista tukea kyseisten yritysten suostumuksella.

VI. KANSALLISTEN TILASTOVIRASTOJEN JA TUTKIJOIDEN VÄLISTEN JÄNNITTEIDEN HALLINTA

Miten ratkaista kansallisten tilastovirastojen ja tutkijoiden näkökulmien väliset jännitteet?

71. Tehokkainta on kansallisten tilastovirastojen siirtyminen riskien välttämisenstrategiasta riskienhallintastrategiaan. Tapoja tämän toteuttamiseksi pohditaan yksityiskohtaisemmin seuraavissa kappaleissa.

72. On selvää, että riskejä on ja että ne on hallittava. Tunnistettavia yksilöitä koskevaa tietoa sisältävien tietokantojen nopea laajentuminen merkitsee, että on käytännössä mahdotonta täysin välttää lukuisten yksilöiden tunnistaminen tietoja yhdistämällä, huolimatta nimien ja osoitteiden poistamisesta, erityisesti jos tiedosto sisältää tiedon kotitalouden rakenteesta. Monet näistä tietokannoista ovat yksityisen sektorin ylläpitämiä, ja siellä niiden käyttöä ei yleisesti valvota yhtä tiukasti kuin julkisella sektorilla. Lisäksi teknologian kehitys on helpottanut tietojen yhdistämistä, joko suoraan tai tilastolliseen malliin perustuvaa yhdistämistä käyttäen (joka voi johtaa harvinaisten tapausten kohdalla tarkkaan vastaavuuteen). Pohjimmiltaan riskien hallinnassa on kyse siitä, ettei tunnistettavissa olevaa yksikkötason tietoa sallita luovutettavan kansallisen tilastoviraston ulkopuolelle ilman arvojen vaihtamista yksiköiden välillä tai muokkaamisen kaltaisia toimenpiteitä. (Huomaa, että riskit vaihtelevat muun muassa maan koon mukaan. Pienissä maissa riski on suhteellisesti tarkasteltuna suurempi, koska niissä esiintyy suhteellisesti enemmän harvinaisia tapauksia.)

73. Kansallisten tilastovirastojen myöntämät oikeudet yksikkötason tietojen käyttöön eivät kuitenkaan näytä antaneen aihetta julkiseen kiistelyyn. Nykyinen käytäntö vaikuttaisikin saaneen suurelta yleisöltä melko laajan, hiljaisen hyväksynnän, vaikka tiedossamme ei olekaan maita, joissa asia olisi herättänyt laajempaa julkista keskustelua. Yleinen huoli yksityisyyden suojasta antaa kuitenkin viitteitä siitä, että jossakin kulkee raja sen suhteen, mitä suuri yleisö on valmis hyväksymään. Yksi epäonnistunut tapaus voi riittää käynnistämään julkisen keskustelun (yli kansallisten rajojen). Sillä voi puolestaan olla vaikutuksia tilastotietojen keruussa tärkeisiin yhteistyösuhteisiin.

74. Läpinäkyvyys on tärkeää, jotta vältytään salailusyytöksiltä. Onkin hyvien käytäntöjen mukaista, että kansalliset tilastovirastot ilmaisevat avoimesti antavansa keräämiään anonymisoituja yksikkötason tietoja osin myös tutkijoiden käyttöön ja että käyttöoikeuksia myönnetään tietyillä ehdoilla perusteltuihin tarkoituksiin. Tämä asia on hoidettava hyvin, niin etteivät yksityisyyden suojaa ajavat tahot pääse ohjaamaan yleistä mielipidettä. Arvostettujen ja vaikutusvaltaisten henkilöiden tuki asialle on elintärkeä.

Miten kansalliset tilastovirastot voivat hallita yksikkötason tietojen luovuttamista koskevia riskejä?

75. Seuraavassa muutamia ehdotuksia:

(i) sovitaan tietyistä periaatteista, joita tulee noudattaa yksikkötason tietoja koskevia käyttöoikeuksia myönnettäessä (luvun 3 periaatteiden tapaan);

(ii) varmistetaan riittävä juridinen ja eettinen perusta (sekä teknologia ja menetelmät) tietosuojan turvaamiseksi. Juridinen ja eettinen perusta edellyttävät tasapainoista arviota, jossa toisessa vaakakupissa ovat tietosuojan ja toisessa tutkimuksen tuomat yleiset hyödyt. Päätös käyttöluvan myöntämisestä voi riippua kyseisen tutkimushankkeen ansioista ja tutkijan uskottavuudesta, joten

juridiseen puoleen tulisi jättää sitä varten jonkin verran liikkumavaraa. Käyttöluvan myöntämistä ei tule pitää itsestäänselvytenä;

(iii) kehitetään ns. arm's length -periaatteen mukainen prosessi näiden kahden yleisen hyödyn tasapainottamiseen. On hyvien käytäntöjen mukaista perustaa sisäinen lautakunta käymään näitä asioita koskevia keskusteluja ja tekemään suosituksia kansallisen tilastoviraston johdolle. Erilaisia arvioita tarvittaessa eettiset lautakunnat tai vastaavat elimet voivat olla avuksi, ja niiltä voi pyytää neuvoja esimerkiksi harkinnanvaraisissa käyttöoikeutta koskevissa päätöksissä. Yleisen hyödyn perusteleva on paljon vakuuttavampaa, jos tutkimuksen tulokset saatetaan yleisesti saataville;

(iv) ollaan täysin läpinäkyviä yksikötason tietojen täsmällisistä käyttötarkoituksista väärinkäyttöepäilyjen välttämiseksi;

(v) ollaan valmiita lisäämään tietojen saatavuutta etäkäyttöjärjestelmien kautta ja datalaboratorioissa, jos täysin tunnistamattomassa muodossa olevia tietoja ei voida luovuttaa yleiseen käyttöön ilman, että tietoja joudutaan merkittävästi "muuntelemaan". Tutkitaan muita keinoja hyödyntää uutta teknologiaa tietojen saatavuuden parantamiseksi siten, että turvataan riittävä tietosuojaa;

(vi) siirretään osa vastuusta tutkijayhteisölle. Varmistetaan, että tutkijat ymmärtävät, miksi kansalliset tilastovirastot haluavat niin aktiivisesti turvata tietojen luottamuksellisuuden. Varmistetaan, että tutkijat tietävät, mitä seurauksia heille ja heidän edustamilleen tutkimuslaitoksille aiheutuu sopimusten tai sitoumusten rikkomisesta. Huolehditaan seurannasta ja toimeenpannaan rikkomusten yhteydessä asianmukaisia sanktioita.

76. Viimeksi mainittu kohta kaipaa hieman lisäselvitystä. Tutkijayhteisön kulttuuri ja arvomaailma poikkeavat hyvin paljon kansallisissa tilastovirastoissa vallitsevista ajattelutavoista. Tutkijat pitävät yksikötason tietojen käyttöön luonnollisesti liittyviä valvontatoimia usein tarpeettomana byrokratiana. Vaikka tiedossa ei ole tapauksia, joissa tutkijat olisivat käyttäneet oikeuttaan yksikötason tietojen käyttöön tarkoituksella yksilöiden tunnistamiseen, tunnetaan kuitenkin tapauksia, joissa yksinomaan tutkijan omaan käyttöön luovutettuja tietoja on annettu ilman lupaa eteenpäin muille tutkijoille tai joissa tietoja on ilman lupaa yhdistetty tilastolliseen malliin perustuen muihin tietoihin kattavampien tiedostojen luomiseksi. Kyseiset tutkijat eivät ehkä omasta mielestään ole tehneet mitään väärää, koska eivät ole yrittäneet tunnistaa aineistosta tiedon kohteista. Tämän tyyppiset tapaukset voivat kuitenkin tullessaan julki horjuttaa suuren yleisön luottamusta, joten niihin tulisi suhtautua vakavasti. Kansalliset tilastovirastot ja tutkijat toimivat erilaisissa kulttuureissa, ja niiden näkemykset edellä mainitun kaltaisten tapausten riskeistä ovat erilaiset. Tämä on otettava huomioon määriteltäessä yksikötason tietojen luovuttamisessa noudatettavia käytäntöjä.

77. Miten kansalliset tilastovirastot voisivat siirtää osan riskeistä tutkijoille? Seuraavassa joitakin toimenpide-ehdotuksia:

(i) pyydetään heitä todistamaan vilpittömyytensä tutkijoina sekä osoittamaan tutkimushankkeesta koitua yleinen hyöty ja yksikötason tietojen tarpeellisuus tutkimukselle;

(ii) edellytetään heiltä laillisesti sitovan sitoumuksen allekirjoittamista, jolloin heitä koskevat samat salassapitovelvoitteiden rikkomisesta seuraavat sanktiot kuin tilastoviraston muuta henkilökuntaa;

(iii) selostetaan heille syyt kansallisten tilastovirastojen varovaisuuteen. Varmistetaan asianmukaisella koulutuksella, että tutkijat ovat täysin perillä velvoitteistaan. Huolehditaan seurannasta käyttämällä tehokkaita tarkastus- ja valvontamenettelyjä. Yhteisten käytäntöjen laatimisesta tutkijayhteisön kanssa saattaisi olla hyötyä;

(iv) jos rikkomuksia esiintyy, keskeytetään sen hetkiset ja tulevat tietopalvelut tutkijalle ja mahdollisesti myös hänen edustamalleen tutkimuslaitokselle määräajaksi (esim. kunnes tutkimuslaitos on ryhtynyt asianmukaisiin kurinpitotoimiin sitoumuksen rikkojaa vastaan). Yritetään saada heidät ymmärtämään, että kaikkien tutkijoiden pääsy yksikötason tietoihin voi vaarantua

tulevaisuudessa, jos asiasta syntyy paljon kriittistä keskustelua julkisuudessa. Ryhdytään tarvittaessa oikeustoimiin.

78. Tietojen luvattoman paljastumisen aiheuttamia ongelmia ei voida aliarvioida, etenkin jos tietojen paljastaminen on tahallista. Tällaisiin tilanteisiin on suhtauduttava erittäin vakavasti.

79. Todellisuudessa yleisen luottamuksen säilyttäminen järjestelmään edellyttää niin juridisia, hallinnollisia kuin teknisiäkin toimenpiteitä. Tutkijayhteisön on myös hyväksyttävä, ettei tietojen saatavuus ole itsestäänselvyys. Kansallisilla tilastovirastoilla voi olla oikeus käyttölupien myöntämiseen, mutta tutkijoita koskevien lupien tulee olla tilastoviraston harkintavallassa. Lisäksi käyttölupaan liittyy tiettyjä velvoitteita. Erityisesti tutkijoiden tulisi hyväksyä omalta osaltaan vastuu siitä, että mahdollisuudet tietojen saatavuuteen pysyvät yhtä hyvinä myös tulevaisuudessa. Rajoitukset ja turvatoimenpiteet saattavat olla ankarammat kuin muiden heidän käyttämiensä tietojen kohdalla, mutta ne ovat perusteltuja.

Muita näkökohtia

Suostumus

80. Joskus esitetään väite, jonka mukaan asiaan pitäisi saada tiedonantajan suostumus ennen yksikötason tietojen luovuttamista kansallisen tilastoviraston ulkopuolelle. Väitettä perustellaan sillä, että tiedonantajalla on oikeus päättää, miten hänen antamiaan tietoja käytetään, riippumatta siitä, onko tieto tunnistettavissa. Tällaiseen ei pitäisi ryhtyä, sillä

(i) suostumusten pyytämiseen ja käsittelyyn liittyy merkittäviä käytännön näkökohtia;

(ii) luovutettavat tiedot eivät ole tunnistettavissa, ja niitä käytetään vain tilastotarkoituksiin, tiedonkeruun tarkoituksen kanssa yhdenmukaisesti;

(iii) on hyvin vaikeaa antaa tiedonantajalle kaikkea sitä tietoa, jota todella harkitun päätöksen tekeminen vaatii, mikä johtaa siihen, että monet vastaisivat varmuuden vuoksi kieltävästi. Otos ei olisi enää edustava, jos se kattaisi vain niiden tiedonantajien tiedot, jotka ovat antaneet suostumuksensa.

Kuten toisaalla näissä suuntaviivoissa todetaan, tilastovirastoilta edellytetään tiedonantajien suuntaan avoimuutta tietojen luovuttamisessa noudatettavasta käytännöstä. Näin toimien voidaan väittää, että tiedonantajilta on saatu asiaan passiivinen suostumus.

(Huom. Jos laki sallii, suostumuksen pyytäminen voisi olla paikallaan tilanteessa, jossa käyttäjällä on julkaistavan pienehkön aggregaattaineiston perusteella mahdollisuus päätellä aineistoon kuuluvan yksikön (henkilön tai yrityksen) tilanne. Tällainen tapaus voi tulla vastaan lähinnä yrityksiä koskevissa tilastoissa.)

Hallinnollisista lähteistä saadut tiedot

81. Suostumukseen liittyy eräs toinenkin näkökohta. Kansallisten tilastovirastojen tietokannat voivat sisältää sekä viraston itse keräämää tietoa että hallintoviranomaisten keräämää ja tilastovirastoille toimittamaa tietoa. Ellei lainsäädännössä ole nimenomaista säännöstä tai menettelysääntöä, jossa määrättäisiin toisin, kansallisen tilastoviraston ei tulisi luovuttaa eteenpäin hallinnollisista lähteistä peräisin olevia yksikötason tietoja ilman hallintoviranomaisen suostumusta (joka ei sitä kenties voi antaa johtuen omille tiedonantajilleen antamistaan lupauksista). Silloinkin kun hallinnollisista lähteistä olevat tiedot ovat jo julkisesti saatavilla, olisi huomaavaista ilmoittaa asiasta hallintoviranomaisille, jotta he saavat tilaisuuden ottaa kantaa. Muussa tapauksessa hallinnollisten tietojen saanti voi hankaloitua. Hallintovirastojen on myös huolehdittava omalta osaltaan yksityisyyden suojaa ja tietosuojaa koskevista asioista.

Valmiussuunnitelma

82. Kansallisten tilastovirastojen on tärkeää laatia suunnitelma sen varalta, että yksikkötason tietojen saatavuus nousee julkisen keskustelun aiheeksi. Ei pidä olettaa, ettei tällaista tilannetta tule koskaan vastaan. Mitkä ovat tärkeimmät argumentit?

(i) Kansalliset tilastovirastot voivat viitata siihen, kuinka huolellisesti ne varmistavat tietosuojan erilaisin menetelmin, kuten anonymisoimalla tiedot, huolehtimalla korkeatasoisista fyysisistä tietoturvajärjestelyistä sekä kehittämällä prosessin toisaalta tietosuojasta ja toisaalta tutkimuksesta suurelle yleisölle koituvien, usein keskenään ristiriitaisten hyötyjen ja niiden keskinäisen tasapainon arviointiin.

(ii) Jos rikkomus on tapahtunut ja kansalliselle tilastovirastolle esitetään asiaa koskevia kysymyksiä, viraston tulisi suhtautua rikkomuksiin ja esitettyihin rangaistusvaatimuksiin avoimesti; ja viraston tulisi tehdä selväksi, että rikkomuksesta vastaa tutkija ja että tilastovirasto ryhtyy omalta osaltaan asianmukaisiin toimiin rikkomuksen johdosta.

(iii) Kansallisten tilastovirastojen tulisi viitata yksikkötason tietojen luovuttamisesta suurelle yleisölle koituihin kokonaisyötyihin, erityisesti koskien sitä tutkimushanketta, jonka yhteydessä rikkomus on sattunut, mutta myös muiden esimerkkien valossa.

(iv) Yhteistyöstä tunnettujen, arvostettujen henkilöiden kanssa, jotka ovat valmiita julkisesti tukemaan tilastoviraston käytäntöjä, tulisi sopia jo ennalta. Erityisen tärkeätä olisi saada mukaan yksityisyyden suojasta vastaavia ylempiä virkamiehiä.

VII. YKSIKÖTASON AINEISTOJEN LUOVUTTAMISEEN LIITTYVIÄ HALLINNOLLISIA KYSYMYKSIÄ

Tietosuoja koskevan päätöksenteon hallinta

83. Tunnistamisen mahdollisuutta ei voida koskaan täysin sulkea pois, tosin se voi joskus olla hyvin pieni. Nykyään on olemassa tietokoneohjelmia, joiden avulla voidaan arvioida harvinaisten – ja siksi helpommin tunnistettavien – tapausten osuus aineistosta.

84. Tilastoviraston johtaja tai hänen sijaisensa tekee käytännössä päätökset yksikötason aineistojen luovuttamisesta joko anonymisoidun tiedoston muodossa (yleiseen käyttöön tarkoitettuna tai käyttö lupaa edellyttävänä tiedostona), etäkäyttöjärjestelmän välityksellä tai datalaboratoriossa. Päätöksentekonsa tueksi johtaja tarvitsee tietoa muun muassa siitä,

- onko tunnistamisen riski riittävän pieni;
- ovatko käytetyt tietosuojakeinot vahingoittaneet yksikötason aineiston sisältöä siinä määrin, ettei sitä voida käyttää tutkimustarkoituksiin; ja
- ovatko muutettiin tehdyt muutokset sekä tutkijan tarpeiden että tunnistamisriskin kannalta asianmukaiset.

85. Esimerkkinä viimeksi mainitusta kohdasta todettakoon, että voidaan valita, kuinka yksityiskohtaista tietoa annetaan maantieteeseen, iän ja kotitalouden rakenteen kaltaisista tunnistamista helpottavista muuttujista.

86. Tällaisten tietojen säännöllisestä toimittamisesta johtajalle tulisi huolehtia asianmukaisin järjestelyin. Hänen päätöksentekonsa tueksi virastossa tulisi olla tutkimustoimintoja, jotka voisivat sijaita viraston metodologiayksikössä. Erilaisia käytäntöjä kuvataan malliesimerkeissä 17 (Slovenia) ja 18 (Australia).

Metatietojen hallinta

87. Yksikötason tietojen tehokas hyödyntäminen edellyttää, että käyttäjille annetaan myös asianmukaiset metatiedot. Ne voivat sisältää

(i) kuvauksen tilastoeselvityksestä sekä tietoa sen laadusta;

(ii) luettelon aineiston sisältämistä tiedoista ja käytetyistä luokitteluista (joskus käytetään nimitystä ”tietohakemisto”); ja

(iii) tietoyksiköiden määrittelyt.

Kohdan (i) mukaisten tietojen antaminen auttaa varmistamaan, ettei yksikötason tietoja käytetä, elleivät ne todella sovellu aiottuun tarkoitukseen.

88. Silloin kun yksikötason tiedot luovutetaan sähköisesti, myös niihin liittyvät metatiedot on annettava käyttäjän saataville. Mikäli mahdollista, metatiedot tulisi luovuttaa heti yksikötason tietojen yhteydessä. Paperituloste on siihen yhä varsin käypä keino, tosin kansallisten tilastovirastojen omien verkkosivujen käyttö tähän tarkoitukseen on kasvussa.

Tutkijoiden tekemien rikkomusten hallinta

89. Kuten edellisessä luvussa tuli esiin, rikkomusten todennäköisyyttä tulisi pyrkiä vähentämään. Kaikesta huolimatta rikkomuksia kuitenkin esiintyy, joten niiden varalle tulisi kehittää toimintamalli.

90. Rikkomuksiin tulee suhtautua vakavasti. Jos näin ei tehdä, suuren yleisön luottamus järjestelmään kärsii. Rikkomuksia myös esiintyy useammin, jos niihin ei puututa ankaralla kädellä.

91. Rikkomuksia voidaan käsitellä lukuisin eri tavoin. Jos on esimerkiksi rikottu lakia, tulisi harkita oikeustoimiin ryhtymistä. Vaikka se tulee kalliiksi, vakavampien rikkomusten yhteydessä on olennaisen tärkeää osoittaa, miten tärkeä asia tietojen luottamuksellisuus on tilastovirastolle. Oikeustoimet on myös tehokas keino vähentää rikkomusten todennäköisyyttä tulevaisuudessa.

92. Lisäksi tulisi estää tutkijan pääsy yksikötason tietoihin vastaisuudessa. Vähintään tämä kurinpidollinen toimenpide tulisi toteuttaa.

93. Tulisi myös harkita tietojen luovuttamisen keskeyttämistä tutkijan edustamalle tutkimuslaitokselle, ainakin kunnes

(i) tutkimuslaitos on ryhtynyt asianmukaisiin toimenpiteisiin tutkijan tekemän rikkomuksen johdosta; ja

(ii) kansallinen tilastovirasto on varma siitä, että kyseisessä tutkimuslaitoksessa on otettu käyttöön asianmukaiset menettelyt tulevien rikkomusten ehkäisemiseksi.

94. Tutkijayhteisö kannattaa yleensä ankariin toimenpiteisiin ryhtymistä niitä suhteellisen harvoja rikkomusten tekijöitä kohtaan, jotka voivat pilata koko tutkijayhteisön maineen. Se on pitkällä aikavälillä tutkijayhteisön oman edun mukaista.

95. Pienempien rikkomusten yhteydessä varoitus riittänee.

VIII. MUITA ERITYISIÄ NÄKÖKOHTIA

Kansainvälinen saatavuus

96. Maiden väliset vertailut ovat tärkeitä eri maissa toteutettujen politiikkojen ja ohjelmien tehokkuuden selvittämiseksi. Erityisesti maiden hallitukset pitävät vertailua hyvänä keinona politiikan arviointiin. Anonymisoitujen yksikötason tietojen käytön sallimisesta maiden välisiä vertailuja tekeville tutkijoille ja kansainvälisissä virastoissa työskenteleville on selkeitä etuja, mutta asiassa on myös riskinsä (esim. tietojen luovuttaminen eteenpäin ilman lupaa). Eurostatia lukuun ottamatta kansainvälisissä virastoissa toimivaa henkilöstöä ei koske mikään kansallinen tai kansainvälinen lainsäädäntö vaan ainoastaan kyseisen organisaation omaa henkilökuntaa koskevat säännöt. On siis syytä toimia varoen. Suurin ongelma on se, että kansainvälisten virastojen henkilöstön tai ulkomailla toimivien tutkijoiden tekemien rikkomusten yhteydessä on käytettävissä vain rajallinen määrä sanktioita. Toisaalta tunnistamisriski on paljon pienempi (niin kauan kun tutkija ei välitä yksikötason tietoja edelleen tietoja luovuttaneessa maassa olevalle kolmannelle osapuolelle). Ongelmia saattaa aiheuttaa myös se, ettei monilla mailla ole lainmukaisia valtuuksia luovuttaa tietoja kansainvälisille virastoille tai tutkijoille oman maansa ulkopuolella.

97. Näissä suuntaviivoissa on kehoitettu siirtymään yksikötason tietojen luovuttamisessa riskien välttämistä pikemminkin riskien hallinnan suuntaan. Riskit ovat pienempiä tai ne vaikuttavat pienemmiltä, jos vastaanottavalla virastolla on uskottavuutta ja luotettavan maine. Riskit ovat myös helpommin perusteltavissa, jos luovutettavien yksikötason tietojen käyttötarkoitus on tietoja tarjoavan maan omissa intresseissä. Kyse saattaa olla esimerkiksi osallistumisesta kansainväliseen tutkimukseen, jonka suorittaa kansainvälinen virasto tai arvostettu kansainvälinen tutkimusyhteistyöorganisaatio (esim. Luxembourg Income Study).

98. Globalisaatio on lisännyt tällaisten kansainvälisten tutkimusten merkitystä. Kansallisten tilastovirastojen tulisi olla oikeutettuja tukemaan tämäläisiä tutkimuksia tarjoamalla niihin tarvittavia tietoja. Kyse pitäisi kuitenkin olla nimenomaan oikeudesta tietojen luovuttamiseen (ts. kansallisella tilastovirastolla tulisi olla mahdollisuus päättää, luovuttaako tietoja vai ei), johon liittyy myös asianmukaisten turvatoimien ja luovutusehtojen määrittely. Tilastovirastojen tulisi olla valmiita luovuttamaan anonymisoituja yksikötason tietoja erityisesti silloin, kun riskit ovat pienet ja hyödyt suuret.

99. Mitä vaihtoehtoja tutkijoilla on käytettävissä, kun he haluavat käyttää muiden maiden tiedostoja? Miten kansainväliset laitokset voivat saada yksikötason tietoja tilasto- ja tutkimustarkoituksiin? Vaihtoehtoja ovat muun muassa seuraavat:

(i) kansainvälinen laitos(tai tutkija) kerää tiedot itse tai toimeksiantoa (esim. asiantuntevaa tutkimuslaitosta) käyttäen ja ilmoittaa yksikötason tietojen luovuttamisesta tutkimuskäyttöön heti tiedonkeruun yhteydessä;

(ii) julkiseen käyttöön tarkoitetut tiedostot, jos niitä on olemassa;

(iii) anonymisoidut käyttöluvat edellyttävät yksikötason aineistot niissä maissa, joissa menetelmä on käytössä;

(iv) asianmukaisin turvatoimenpitein varustetut etäkäyttöjärjestelmät;

(v) yhteistyö sellaisen kansallisessa tilastovirastossa tai kansallisen tilastoviraston kanssa samassa maassa työskentelevän tutkijan kanssa, jolla on käyttöluvat yksikötason aineistoihin.

100. Kansainvälisille tutkijoille suositeltavin tapa päästä yksikötason tietoihin on vaihtoehto (i). PISA-tutkimus on siitä hyvä esimerkki. Kansalliset tilastolait eivät kuitenkaan yleensä koske tällaisten tutkimusten tiedonkeruuta. Joissakin tutkimuksissa tietojen laatu, ja erityisesti vastausaktiivisuus, voivat kärsiä. Nämä riippuvat tutkimuksen tyypistä sekä tutkimuksen suorittavan organisaation maineesta. Lisäksi asiaan vaikuttaa kussakin maassa tiedonkeruuta suorittavien sitoutuneisuus. Nämä tekijät kansainvälisten tutkijoiden on syytä ottaa huomioon ennen päätöstään käyttää vaihtoehtoa (i). Tietojen saatavuutta ja laatua saattaa joutua punnitsemaan keskenään.

101. Joissakin tapauksissa voi olla tutkimuksen kannalta riittävää, että analysoitavaksi saadaan muuta hyvin yksityiskohtaista tietoa, joka ei kuitenkaan ole yksikötason tietoa. Tätä menetelmää käytettiin International Comparison Programme (ICP) -hankkeessa ja OECD:n/Eurostatin ostovoimapariteettitutkimuksessa.

102. Vaihtoehdossa (ii) mainittuja yleiskäyttötiedostoja on saatavissa vain joistakin maista. Käyttölupaa edellyttäviä anonymisoituja tiedostoja (vaihtoehto (iii)) voi käyttää, elleivät kansallisen tilastoviraston säännökset aseta rajoituksia. Jos kansalliset tilastovirastot voivat luovuttaa yksikötason tietoja näissä muodoissa, seuraavat seikat olisi hyvä ottaa huomioon:

- luottamus tutkijaan ja hänen edustamaansa tutkimuslaitokseen;
- tutkimuksen merkitys kyseiselle maalle; ja
- onko tietojen luovutus linjassa tiedonantajille tietoja kerätessä annettujen lupauksen kanssa.

103. On todennäköistä, että monet maat luovuttavat tietoja mieluummin tietyille kansainvälisille laitoksille tai tiettyihin tutkimusprojekteihin kuin yleisesti kansainvälisen tutkijayhteisön saataville. Lisäksi eri tutkijoille voidaan asettaa erilaisia ehtoja. Jotkut maat suosivat tietojen luovutusta tutkijan kotimaan kansallisen tilastoviraston kautta. Mutta oli käytetty luovutustapa sitten mikä tahansa, on aina hyvien käytäntöjen mukaista luovuttaa tietoja ainoastaan tietyillä, erillisessä sitoumuksessa tai muistiossa määritellyillä ehdoilla. Ehdot voivat sisältää seuraavanlaisia asioita:

- rajoitetaan tietojen käyttöoikeus vain kansainvälisten laitosten tiettyihin osastoihin ja kielletään tietojen välittäminen muille;
- rajoitetaan käyttötarkoituksia, joihin yksikötason tietoja voidaan käyttää ilman erillistä lupaa
- sovitaan yksikötason tietojen palauttamisesta pyydettyä (ne voivat esim. sisältää virheitä);
- varataan tilastovirastolle mahdollisuus ottaa kantaa julkaistuun, yksikötason tietoihin pohjautuvaan materiaaliin; ja
- tehdään selväksi, mitä luovutusehtojen rikkomisesta seuraa.

104. Tehokkain keino luovutusehtojen rikkomusten käsittelyyn on tietojen luovuttamisen keskeyttäminen toistaiseksi. Asia voidaan myös ottaa esille tutkimuslaitoksen ylemmän johdon kanssa. Kansainvälisten virastojen ollessa kyseessä vakavat rikkomukset voidaan käsitellä diplomaattivirkamiesten kanssa. Keskeistä on kuitenkin, ettei rikkomusten sallita tapahtuvan ilman, että niihin reagoidaan. Muussa tapauksessa ne yksinkertaisesti toistuvat.

105. Monissa maissa etäkäyttöjärjestelmät (vaihtoehto (iv)) voivat olla suosituin menetelmä luovuttaa tietoja tutkijoille kansainvälisesti. Näiden järjestelmien yhteydessä on useita varmistuksia ja kansallisten tilastovirastojen asema yksikötason tietojen tarjoajana on ongelmatilanteiden kannalta suojatumpi. Järjestelmien käytettävyydessä kansainväliseen tiedonjakoon on kuitenkin vielä parantamisen varaa. Kokeilut ovat tärkeitä.

106. Yksi kansainvälisille tutkijoille ja myös kansainvälisille virastoille sopiva toimintamalli voisi olla kansallisten tutkijoiden muodostaman verkoston hyödyntäminen tutkimustyössä (eli vaihtoehto (v)), esimerkiksi yksikötason tietojen erilaisissa analyyseissä. Kansalliset tutkijat voisivat toimia kansallisissa tilastovirastoissa erityisen tärkeiden kansainvälisten tutkimusten parissa.

107. Kansallisten tilastovirastojen on päätettävä, luovuttavatko ne tietoja kansainvälisille tutkijoille vai eivät, ottaen huomioon tässä luvussa käsitellyt asiat. Olisi pidettävä mielessä, että suositeltavinta on pyrkiä toiminnassa riskien hallintaan. Joissakin tutkimuksissa saavutettavat hyödyt voivat olla niin suuret, että tietyt riskit ovat perusteltuja, kunhan huolehditaan toiminnan laillisuudesta. Riskitaso voi vaihdella tutkimuslaitoksesta toiseen. Kansallisten tilastovirastojen on myös valittava sopivin menetelmä tietojen luovuttamiseen kansainvälisille tutkijoille. Päätöksenteon yhtenäistämistä helpottaisivat eri maiden välisenä yhteistyönä kehitetyt tietojen luovuttamista kansainvälisille virastoille ja tutkijoille koskevat suuntaviivat, joiden tulisi olla linjassa kansallisen lainsäädännön kanssa. Maat voivat myös päättää muuttaa omaa lainsäädäntöään sallimaan tietojen luovuttamisen perustelluissa tapauksissa.

108. Malliesimerkissä 19 kuvaillaan OECD:n PISA-tutkimuksen toteutusta ja malliesimerkissä 20 yksikötason tietojen luovuttamista kansainvälisiin tutkimuksiin Australiassa.

Tietojen yhdistäminen

109. Tietojen yhdistäminen, joko suoraan tai tilastolliseen malliin perustuvalla yhdistämisellä, voi tuoda aineistolle huomattavasti lisäarvoa. Sillä voidaan lisätä huomattavasti erilaisia analysointimahdollisuuksia. Esimerkiksi terveyteen liittyvissä tutkimuksissa yhdistetyt tiedostot voivat olla erityisen arvokkaita. Kansallisten tilastovirastojen on asianmukaista seurata tilastotarkoituksessa tehtävää tiedostojen yhdistämistä.

110. Tutkijat pyrkivät yhä enemmän hyödyntämään tietoja, jotka on yhdistetty kansallisen tilastoviraston tai vastaavan elimen tietokantoihin (esimerkiksi joissakin maissa väestölaskentatietoihin). Tällaisten yhdistettyjen tiedostojen hallinnointi tulee olla tilastoviraston vastuulla. On myös mahdollista, että tilastoviraston hallinnoitavaksi uskotaan myös tilastovirastojen ulkopuolelta peräisin olevia yhdistettyjä tiedostoja, koska virastoilla on jo tarvittavat tietoturvajärjestelyt ja suuren yleisön luottamus.

111. Vaikka aineistojen yhdistelyyn liittyy selkeitä etuja, siinä on myös riskinsä, etenkin jos yhdistetyn tiedoston hallinnoijalla ei ole vastaavanlaista tietosuojajärjestelmää kuin kansallisissa tilastovirastoissa. Tunnistamisriskit kasvavat yhdistetyissä tiedostoissa. Myös asiaa koskevilla käsityksillä on merkitystä. Monissa maissa tehdyt selvitykset ovat osoittaneet, että tietokantojen yhdistämisestä ollaan yleisesti huolissaan. Tiedostoja yhdistettäessä onkin erityisen tärkeä noudattaa luvussa 3 mainittua neljää periaatetta.

112. Niissä maissa, joissa toimii yksityisyyden suojasta vastaavia elimiä, tietojen yhdistämisessä noudatettavalle käytännölle tulisi saada näiden organisaatioiden tuki.

113. Tietojen yhdistämisen käytäntöjä kuvaillaan malliesimerkisä 21 (Kanada) ja 22 (Ruotsi).

LIITE 2. TERMINOLOGIAA

Tätä terminologiaa tulisi lukea rinnakkain UNECE:n ja Eurostatin yhteistyönä laaditun tilastotietojen luottamuksellisuutta koskevan Glossary on Statistical Disclosure Control –sanaston kanssa. (Sanasto on saatavissa englanninkielisenä osoitteesta:
www.unece.org/stats/documents/ece/ces/ge.46/2005/wp.45.e.pdf.)

Kansallinen tilastovirasto

Vaikka ilmaisu käytettäisiin yksikössä, sen on tarkoitus viitata kaikkiin tilastovirastoihin tai valtionhallinnon tilastointiyksiköihin, jotka tuottavat virallisia tilastoja ja tarjoavat pääsyn tilasto- tai tutkimustarkoituksiin käytettäviin yksikkötason tietoihin.

Tutkijayhteisö

Vaikka termi viittaa lähinnä tutkimuslaitoksissa, kuten yliopistoissa, työskenteleviin henkilöihin, se sisältää myös tutkijat, jotka toimivat valtion virastoissa, kansalaisjärjestöissä, kansainvälisissä virastoissa sekä yksityissektorilla. Joissakin maissa termi määritellään kapeammin, jolloin se saattaa kattaa ainoastaan tutkimuslaitoksissa toimivat tutkijat.

Tilastotarkoitukset

On erityisen tärkeää erottaa toisistaan tietojen tilastollinen ja hallinnollinen käyttö. Tilastollisessa käytössä yksikkökohtaisia tietoja käytetään panoksena tilastojen tuotannossa, jolloin tuotoksena saatavat tilastot viittaavat ryhmään luonnollisia tai juridisia henkilöitä. Tilastotarkoitus voi sisältää myös tukea kansallisen tilastoviraston muulle toiminnalle (esim. otoksen valinnassa yritysrekisteristä). Hallinnollisessa käytössä kyse on jotakin tiettyä luonnollista tai juridista henkilöä koskevasta päätöksestä, joka voi aiheuttaa kyseiselle yksilölle tai yksikölle joko etua tai haittaa.

Edellä mainitut tilastot voivat olla tilastollisia aggregaatteja, tilastollisia jakaumia, mallien parametrejä tai muita tilastoanalyysin muotoja, jotka viittaavat yksilöistä tai organisaatiosta koostuviin ryhmiin identifioimatta niitä.

Yksikkötason tietojen käyttö tutkimuksessa vastaa tilastotarkoituksia, jos tietoja käytetään edellä mainitun kaltaisten tilastojen tuottamiseen.

Anonymisoidut yksikkötason aineistot– yleiskäyttötiedostot Nämä ovat suuren yleisön käyttöön tarkoitettuja yksikkötason tiedostoja. Tiedostot on anonymisoitu ja niiden levityksessä käytetään usein jotakin välinettä, esim. CD-ROM-levyä, ja joskus ne kulkevat tietoarkiston kautta. Anonymisoinnilla viitataan nimien ja osoitteiden poistamisen lisäksi myös muihin toimenpiteisiin, joilla varmistetaan, että yksilöiden tunnistaminen on erittäin epätodennäköistä.

Anonymisoidut yksikkötason aineistot – käyttö lupaa edellyttävät tiedostot

Anonymisoinnilla viitataan nimien ja osoitteiden poistamisen lisäksi myös muihin toimenpiteisiin, joilla varmistetaan, että yksilöiden tunnistaminen on erittäin epätodennäköistä.

Toisin kuin yleiskäyttötiedostot, käyttö lupaa edellyttäviä tiedostoja voivat käyttää ainoastaan hyväksytyt tutkijat hyväksytyihin tarkoituksiin. Ennen tiedostojen luovuttamista tutkijoiden on allekirjoitettava lainvoimainen sitoumus.

Etäyhteyskäyttö, etäkäyttöjärjestelmä

Nämä järjestelmät tarjoavat tutkijoille mahdollisuuden tuottaa yksikkötason tiedoista tilastollisia tutkimustuloksia verkkoyhteyden välityksellä, ilman että tutkijat varsinaisesti ”näkevät” käyttämiään yksikkötason tietoja. Varsinaiset yksikkötason tiedot eivät siirry lainkaan kansallisen tilastoviraston ulkopuolelle. Käytössä on kahdenlaisia etäkäyttöjärjestelmiä.

(a) Etäyhteyden välityksellä lähetettävät toimeksiannot, joka tarkoittaa, että tutkija lähettää tilastovirastoon ohjelman ja saa sillä tuotetun aineiston jälkeenpäin sähköpostitse.

(b) Varsinainen etäyhteyskäyttö, jossa tutkija tekee analyysin itse ja näkee tuloksen välittömästi tietokoneensa näyttöruudulta.

Datalaboratoriot

Tässä vaihtoehdossa pääsy yksikkötason tietoihin edellyttää työskentelyä kansallisen tilastoviraston tiloissa tai sen erillisyyksikössä. Tiedot voivat olla saatavilla joko välittömästi tai välillisesti kansallisen tilastoviraston henkilökunnan valvonnassa. Jos tiedot ovat välittömästi tutkijan saatavilla, se merkitsee, että tutkijaa kohdellaan tilastoviraston tilapäisenä työntekijänä, jolloin häntä koskevat samat velvoitteet kuin muutakin henkilöstöä.

Arvojen vaihtaminen yksiköiden välillä

Yksikkötason tietojen tietosuojamenetelmä, jossa arvoja vaihdetaan yksiköiden välillä. Menetelmä säilyttää tilastolliset tunnusluvut, kuten keskiarvot, varianssit ja yksimuuttujajakaumat mutta voi vaikuttaa monimuuttujajakaumiin.

Tietojen muokkaaminen

Yksikkötason tietoja luovutettaessa käytettävä menetelmä, jossa tietoja muokataan ennen luovutusta siten, että tietojen tunnistamisriski pienenee mutta informaatio sisältö pysyy mahdollisimman pitkälti ennallaan. Muokkaavat menetelmät ”vääristävät” tietoja tekemällä niihin tietosuojasyistä tarkoituksella pieniä virheitä. Mahdollisia muokkaavia menetelmiä ovat:

- pyöristys
- kohinan lisääminen.

Riskien välttäminen

Tässä toimintamallissa pyritään kaikkien riskien eliminointiin. Yksikkötason tietojen tietosuojan edellytetään olevan absoluuttinen, ei pelkästään silloin kun tiedostoja katsotaan erikseen vaan myös yhdistettäessä niitä muihin tiedostoihin.

Riskienhallinta

Tämä tarkoittaa lainsäädännön suomissa rajoissa suoritettavaa riskien tunnistamista ja hallintaa erilaisten riskien suuruuden (vaikutus) ja todennäköisyyden mukaan. Eniten resursseja käytetään vaikutukseltaan ja todennäköisyydeltään suurimpien riskien hallintaan. Yksikkötason tietojen tietosuoja ei ole aina absoluuttinen silloin, kun tietoja tarkastellaan yhdistettyinä muihin tietoihin. Tietosuoja voitaisiinkin tarkastella osana riskien pienentämiseksi käytettävien menetelmien muodostamaa kokonaisuutta

Tietojen yhdistäminen

Tietoja voidaan yhdistää suoraan (esim. käyttämällä tunnistetietoja kuten nimeä ja osoitetta tai henkilötunnusta) tai tilastolliseen malliin perustuen (käyttämällä todennäköisiä vastaavuuksia). Kyse voi olla pelkästään kansallisen tilastoviraston aineistosta, tilastoviraston omasta ja hallinnollisista lähteistä peräisin olevasta aineistosta tai pelkästään hallinnollisten lähteiden aineistoista. Kunkin aineiston tiedot voidaan yhdistää pitkittäin. Kaikkia näitä mahdollisuuksia voidaan käyttää tietojen yhdistämisessä.

* * * * *