



Tilastokeskus 

Tietotilinpäätös 2020

Kansikuva: Shutterstock
Kannen suunnittelu: Riikka Turunen
Taitto: Riikka Turunen

Tietoja lainattaessa lähteenä on mainittava Tilastokeskus.

© 2021 Tilastokeskus

Tilastokeskuksen pääjohtaja
Markus Sovala.



Pääjohtajan alkusanat

Tiedon määrä yhteiskunnassa on kasvanut, ja tietoa sekä tuottaa että hyödyntää yhä suurempi joukko toimijoita. Vaikka toimijoiden moninaisuus aiheuttaa variaatiota tiedon laatuun, on virallisten tilastojen rooli luotettavan tiedon lähteenä säilyttänyt asemansa.

Tilastotoimi on organisoitu riippumattomaksi toiminnoksi, ja sitä ohjaavat tilastolain lisäksi tilastoalan ammattieettiset periaatteet ja käytäntesäännöt. Alan ammattietiikka on toiminnan keskeinen perusta, ja tilastoviranomainen näyttää toiminnallaan esimerkkiä myös muille tiedontuottajille tilastotuotannon ja tietojen käytön eettisten periaatteiden toteuttamisessa.

Tilastotoimen yleisviranomaisena Tilastokeskus hyödyntää tilastoja kootessaan laajasti yhteiskunnan muiden toimijoiden toiminnan yhteydessä syntyviä tietoaineistoja. Niiden avulla tilastoja voidaan tuottaa

tehokkaasti ja laadukkaasti. Tilastokeskus tuottaa ajantasaista tietoa yhteiskunnan hyödyksi ja edistää tietoaaineistojensa käyttöä tilastointi- ja tutkimustarkoituksiin yhteiskunnassa.

Tilastoviranomaisen toimintaa leimaa pyrkimys avoimuuteen. Julkistamme tilastoinnissa käytetyt menetelmät sekä kuvaamme tilastotuotannossa käytetyn aineiston ja analyysimenetelmät tilastolukujen yhteydessä. Tämä luo pohjan luottamukselle tilastoja ja Tilastokeskusta kohtaan.

Tietosuojalainsäädännön uudistumisen myötä henkilötietojen läpinäkyvä ja luotettava käsittely on muodostunut yhä tärkeämmäksi. Tilastokeskuksen tietosuojaperiaatteet varmistavat sen, että yksittäisen tiedonantajan tiedot ovat luottamuksellisia. Tietotilinpäätöksen avulla raportoimme henkilötietojen käsittelystä vuosittain.

Tietotilinpäätökseen on koottu vuoden 2020 keskeisimmät tietosuojan täytäntöönpanoon liittyvät asiat. Poimin tähän kolme tärkeää huomiota raportin sisällöstä:

- Pandemia aiheutti työn tekemisessä ison muutoksen, kun maaliskuussa siirryttiin pääosin etätyöhön. Käytössä olevat viestintävälineet ja toimintatavat tukivat hyvin tietosuojan ja -turvan toteutumista myös etätyössä.
- Syksyllä käyttöön otettu uusi asianhallintajärjestelmä VirastoVahva ottaa aiempaa paremmin huomioon asianhallinnan ja tietosuojan vaatimukset.
- Vuoden aikana valmistui esitys tilastotuotannon tietoaaineistojen säilytysajoista ja tietoaaineistojen hävittämisestä. Sen täytäntöönpano alkaa vuonna 2021.

Sisällys

Pääjohtajan alkusanat	3
1 Tiedonhallintalain toimeenpano Tilastokeskuksessa.....	6
2 Kokonaisarkkitehtuuri	8
3 Tietosuojan vaikutustenarviointi	9
4 Koulutus ja ohjeet	9
5 Tietoturva	10
6 Sisäinen tarkastus	11
7 Tiedonkeruu poikkeusoloissa	12
8 Suomi.fi-tunnistautumisen käyttöönotto alkoi	13
9 Väestölaskentavuosi 2020.....	13
10 VirastoVahvan käyttöönotto	15
11 Arkisto	15
12 Mittarit vuonna 2020	16
12.1 Rekisteritietojen tarkastuspyynnöt	16
12.2 Tilastot	16
12.3 Tilastotuotannossa käytettävien aineistojen vastaanotto	16
12.4 Suorat tiedonkeruut	17
12.5 Tilaustutkimukset	20
12.6 Asianhallinta ja arkistointi.....	21
12.7 Tutkimusaineistot	22

1 Tiedonhallintalain toimeenpano Tilastokeskuksessa

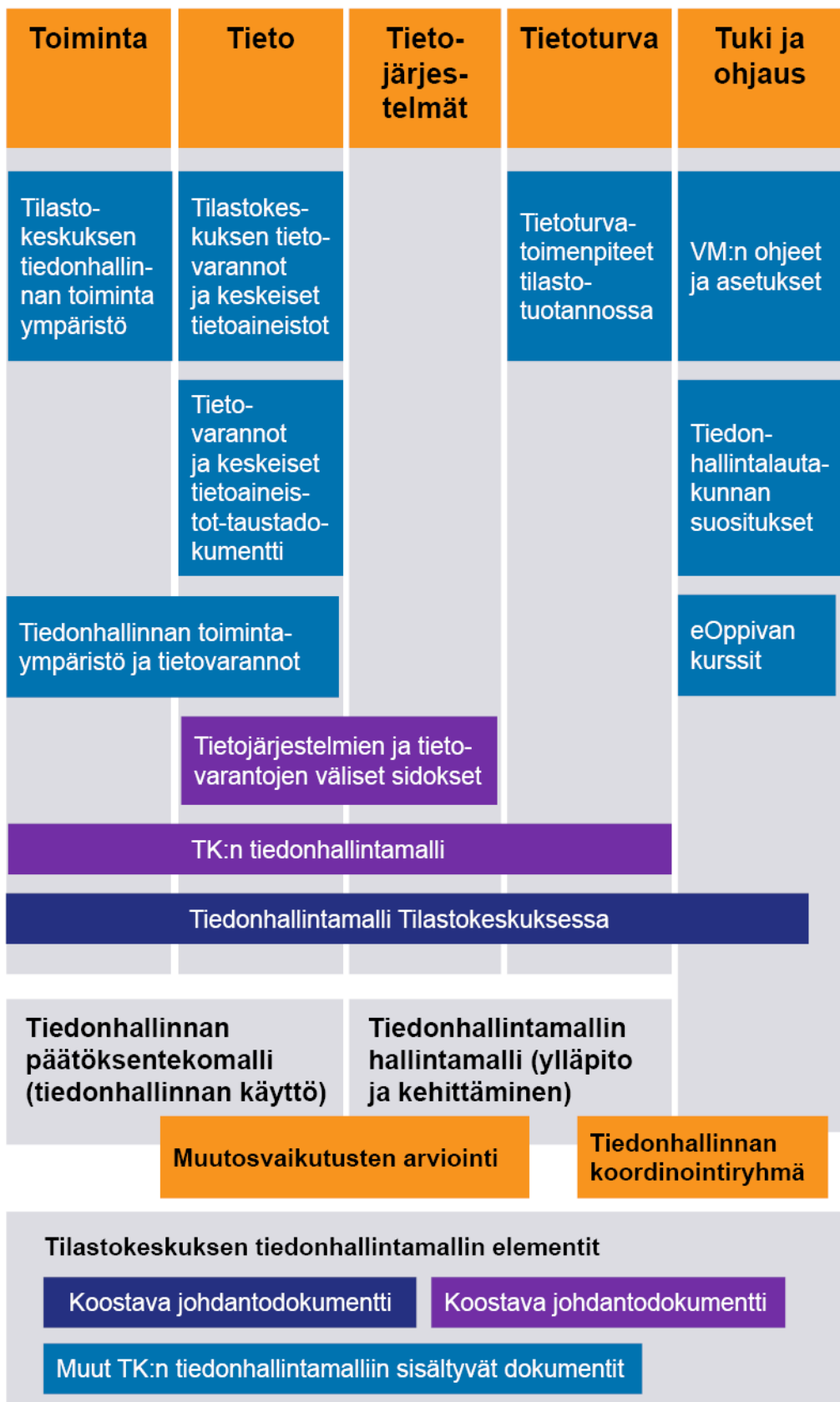
Vuoden 2020 alussa voimaan tullut laki julkisen hallinnon tiedonhallinnasta (tiedonhallintalaki) luo yleislain tasolla perustan yhtenäisille menettelytavoille, joilla julkinen hallinto käsittelee hallussaan olevia tietoja koko niiden elinkaaren ajan.

Tiedonhallintalain täytäntöönpanemiseksi Tilastokeskuksessa perustettiin kaksi työryhmää: Tiedonhallintalaki tilastotuotantoprosessiin sekä Tiedonhallintamalli. Tiedonhallintalain täytäntöönpanon koordinoimisesta vastasi koordinaatioryhmä, joka selvitti lain vaatimuksia Tilastokeskuksessa. Työryhmien työ valmistui joulukuussa 2020.

Tiedonhallintalaki tilastotuotantoprosessiin -työryhmän tavoitteena oli selvittää tilastotuotantoprosessin tietojärjestelmiin ja tietoaineistoihin kohdistuvien tiedonhallintalain vaatimusten toteutumisen nykytilaa, laatia aikataulutettu suunnitelma näiden vaatimusten toteuttamiseksi sekä laatia vaatimusten toteuttamiseksi tarvittavat ohjeet. Työryhmä laati muun muassa raportin lokitietojen keräämisen nykytilasta ja ehdotuksen tiedonhallintalain lokitietojen keräämistä koskevien vaatimusten täyttämiseksi, muistion tietojärjestelmien ohjeista ja niiden päivittämisestä, esityksen tilastotuotannon tietoaineistojen säilytysajoista, päätösluonnoksen tietoaineistojen hävittämisestä sekä ajoitetun suunnitelman tiedonhallintalain vaatimusten toteuttamisesta tilastotuotantoprosessissa.

Tiedonhallintalaki edellyttää tiedonhallintayksiköiltä toimintaympäristönsä tiedonhallintaa määrittelevää ja kuvaavaa tiedonhallintamallia. Tilastokeskuksen tiedonhallintamallin laatimisesta vastasi Tiedonhallintamalli-työryhmä. Tiedonhallintamallin ensimmäinen versio toteutettiin syksyn 2020 aikana (kuva 1).

Kuva 1. Tilastokeskuksen tiedonhallintamallin rakenne ja sen ensimmäiseen versioon sisältyvät dokumentit ja muu materiaali.



Ensimmäisessä versiossa pitäydettiin kuvaamisessa suhteellisen ylätasolla, jolloin tietovarannot, tietojärjestelmät ja prosessit saatiin kuvattua vaatimusten mukaisesti rinnakkain, yhteen toimivasti ja

koko organisaation kattavasti. Mallin työstämisessä lähtöperiaatteena oli hyödyntää jo olemassa olevia KA-kuvauksia ja dokumentteja mahdollisimman pitkälle. Toteutetulla mallilla onkin selkeä linkki Tilastokeskuksen kokonaisarkkitehtuurityöhön. Malli on Tilastokeskuksen johdon hyväksymä.

2 Kokonaisarkkitehtuuri

Kokonaisarkkitehtuurityön organisoinnissa on tunnistettu arkkitehtuurin kerroksellinen rakenne. Alkuvuonna 2020 koko Tilastokeskuksen organisaatiotasaisen arkkitehtuurin alle rakentui tilasto- ja tietotuotannon yhteinen arkkitehtuuri. Lisäksi joissakin kohdealuearkkitehtuureissa keskityttiin tietyn arkkitehtuurin osa-alueen yhteiseen kehittämiseen. Organisaatiomuutoksen myötä kokonaisarkkitehtuurityö on sijoitettuna Kehittäminen ja digitalisaatio -palvelualueelle, jossa on käynnistynyt Tilastokeskuksen arkkitehtuurin hallintamallin päivitystyö ja sen sovittaminen Tilastokeskuksen uuteen kehittämismalliin.

Arkkitehtuuritoiminta ja -kyvykkyys arvioidaan vuosittain. Vuosiarviointien pohjalta laaditaan ehdotus kehittämishankkeista ja -toimenpiteistä ja päivitetään Tilastokeskuksen KA-toiminnan kehittämispolku. Lisäksi tavoitteena on yhdistää kokonaisarkkitehtuuri osaksi Tilastokeskuksen kehittämistoimintaa. Tavoitteena on hahmottaa projektin kehittämiskohde osana Tilastokeskuksen koko arkkitehtuurikokonaisuutta, sekä toisaalta arkkitehtuurilinjausten, -periaatteiden, valmiiden ratkaisujen ja yhteisten palvelujen hyödyntäminen kehittämisprojekteissa. Jo projektin suunnitteluvaiheessa tulee arvioida, miten olemassa olevia arkkitehtuuriperiaatteita ja -linjauksia sovelletaan kyseisessä projektissa sekä miten projekti asemoituu Tilastokeskuksen arkkitehtuurissa.

Toistaiseksi Tilastokeskuksessa on pääosin keskitytty organisaatiokohtaiseen kokonaisarkkitehtuurityöhön, mutta yhteisten palvelukokonaisuuksien arkkitehtuurit, tiedonhallintalaki ym. tuovat kokonaisarkkitehtuuriryhmille eteen myös tilanteita, joissa täytyy toimia julkisen hallinnon yhteisen arkkitehtuuritoiminnan sekä myös yksittäisten toimijoiden suuntaan. Arkkitehtuuriryhmät seuraavat myös julkisen hallinnon yhteisen kokonaisarkkitehtuurin sekä toimialojen arkkitehtuurityötä ja työn etenemistä. Tilastokeskus seuraa myös kansainvälistä arkkitehtuurityötä ja osallistuu kansainvälisten viitekehikoiden kehittämiseen.

3 Tietosuojan vaikutustenarviointi

Vuosina 2018–2019 tietosuojan vaikutustenarviointit on Tilastokeskuksessa toteutettu itsearviointina tilastoprosesseittain. Vuonna 2020 päätettiin toteuttaa tietosuojan vaikutustenarviointi suunnitteilla olevaan Stiina-henkilötietovarantoon, johon on lähivuosien aikana vaiheittain tarkoitus siirtää kaikki Tilastokeskuksen henkilötietoja sisältävät tietoaaineistot.

Stiina-hankkeen tavoitteena on rakentaa palvelupohjaisen tietoaarkkitehtuurin mukaisesti yhteiskäyttöiset, jatkuvapäivitteiset ja metatieto-ohjautuvat tuotantojärjestelmät. Tavoitteena on, että mitään henkilötiedon käsittelyä ei toisteta Tilastokeskuksen tuotantoprosesseissa, vaan kaikki tieto on käsitelty ja tallennettu kertaalleen henkilötietovarantoon, josta tietoa hallinnoidaan ja käytetään keskitetysti. Uudistuksessa huomioidaan tärkeänä osana henkilötietojen käytön tietosuoja. Hankkeen myötä otetaan käyttöön henkilön suorien tunnistetietojen pseudonymisointi kaikissa henkilö pohjaisissa tuotantoprosesseissa.

Tietosuojan vaikutustenarviointi hankittiin järjestelmän merkittävyyden ja laajuuden vuoksi ostopalveluna. Raportti saatiin loppuvuodesta 2020. Henkilötietojen käsittelyä Stiina-tietovarannossa kehitetään edelleen arvioinnin pohjalta.

Loppuvuonna 2020 valmisteltiin tietosuojan vaikutustenarvioinnin hankintaa Tilastokeskuksen tutkijapalvelun Fiona-etäkäyttöjärjestelmään. Tietosuojan vaikutustenarviointi on tarkoitus toteuttaa alkuvuonna 2021.

4 Koulutus ja ohjeet

Tilastokeskus on vuoden aikana päivittänyt ja uusinnut tietosuojaan liittyvää ohjeistustaan. Vuonna 2020 ohjeistusta tarkistettiin ja täsmennettiin erityisesti Tilastokeskuksen organisaatiouudistuksen ja sähköisen asianhallinnan järjestelmän VirastoVahvan käyttöönoton johdosta. Ohjeita peilattiin myös tietosuojalainsäädännön soveltamiskäytännön valossa.

Vuoden 2020 aikana Tilastokeskuksessa päivitettiin tietosuojapolitiikka vuosille 2021–2024, Tilastokeskuksen tietosuoja- ja tietoturvaohje, tilastolain soveltamisohje ja tietoaineistojen tutkimuskäyttöön luovuttamista koskeva käyttölupaohje.

Koronatilanteesta johtuen loppukeväälle 2020 suunnitelluja koulutustilaisuuksia siirrettiin syksylle. Poikkeuksellisen tilanteen jatkuessa myös syksyllä ohjattiin koulutusta erityisesti eOppivassa olevaan verkkokoulutukseen, johon sisältyy kaikille käyttäjille suunnatun sisällön lisäksi myös Tilastokeskuksen itse tuottamaa koulutusta mm. tiedonhallinnasta ja tilastolainsäädännöstä. Lisäksi vuoden 2020 aikana pidettiin joitain pienempimuotoisia, tiettyyn teemaan keskittyviä keskustelutilaisuuksia.

5 Tietoturva

Maaliskuun puolivälin jälkeen Tilastokeskuksen työntekijöistä valtaosa, yli 85 prosenttia, työskenteli etänä. Etätöitä oli tehty jo tätä aiemminkin, mutta ei tässä laajuudessa.

Suurin tekninen haaste kohdistui tietoliikenneyhteyksien toimivuuteen. Vaikka ongelmiakin oli, etenkin ensimmäisinä kuukausina, kokonaisuutena arvioiden kaikki sujui paremmin kuin osattiin odottaa.

Sujuvan etätöiden tekivät mahdolliseksi jo aiemmin käyttöönotetut uudet viestintävälineet sekä Valtorin toimittama tilastotuotannon tarpeisiin rakennettu etäkäyttöympäristö.

Tietoturvatapahtumista suurin osa koostui erilaisista huijausyrityksistä. Teknisen tuen huijauspuheluita tuli runsaasti isolle osaa viraston työntekijöistä. Aiempien vuosien tapaan myös ns. toimitusjohtajahuijauksia sekä laskuiksi naamioituja haittaohjelmia esiintyi runsaasti.

Marraskuussa Tilastokeskus osallistui Digi- ja väestötietoviraston järjestämään julkishallinnolle suunnattuun tietoturva- ja tietosuojaloukkauksien hallinnan harjoitukseen.

6 Sisäinen tarkastus

Vuoden aikana valmistuivat käyttövaluushallintaa ja digitaalisten tilastoaineistojen tarkatusta koskevat tarkastukset.

Käyttövaluushallintaa koskeva sisäinen tarkastus tehtiin Avain-palvelusta. Avain-palvelu on Valtorin tuottama identiteetinhallintapalvelu ja käyttövaluushallinnan työkalu, jonka Tilastokeskus otti käyttöön keväällä 2019. Tarkastuksen tavoitteena oli antaa Tilastokeskuksen johdolle arvio Avain-palvelun käytettävyydestä, tietoturvallisuudesta ja tietosuojasta. Lisäksi arvioitiin sisäisen valvonnan toimivuutta ja riittävyttä. Tarkastuksessa annettujen suositusten toimeenpanoa edistettiin vuoden aikana Käyttövaluushallinnan uudelleen organisointi -projektissa.

Toinen tarkastus kohdistui digitaalisten tilastoaineistojen arkistointiin. Kansallisarkiston päätöksen mukaan valmiit tilastoaineistot säilytetään digitaalisessa muodossa pysyvästi. Tarkastuksessa käytiin läpi sähköisessä muodossa olevan tilastoaineiston arkistoinnin tilannetta, prosessia ja ohjeistusta. Tarkastuksen suositusten toimeenpano on annettu Tieto- ja tilastoprosessit (Tipa)-palvelualueen tehtäväksi.

Joulukuussa aloitettiin tarkastus Tilastokeskuksen varmistusjärjestelmästä. Tarkastuksen kohteena ovat Tilastokeskuksen tilastotuotannon tiedonkeruuaineiston, käsitellyn aineiston, tilastoaineiston ja arkistoidun aineiston varmistusjärjestelmien toimivuus ja luotettavuus sekä varmistusjärjestelmissä olevan tiedon luottamuksellisuus, eheys ja saatavuus. Tarkastusraportti valmistui tammikuussa 2021.

Ennalta sovittujen tarkastusten lisäksi Tilastokeskuksessa tehtiin lokitietojen käsittelyyn liittyvä selvitys. Lainsäädäntö määrittää yleisvaatimukset lokitietojen käsittelylle. Lisäksi lokitietojen keräämisestä ja käsittelystä (ns. lokienhallinnasta) on saatavilla valtiosektorille osoitettuja ohjeita.

Sisäinen tarkastus selvitti tilastotuotannon tietojärjestelmien ja tiedonhallinnan lokitietojen käsittelyä (keruu, säilytys, katselu, analysointi, raportointi, luovutus ja tuhoaminen). Tiedonhallintalain mukaan tietojärjestelmien käytöstä ja niistä tehtävistä tietojen luovutuksista kerätään tarpeelliset lokitiedot, jos tietojärjestelmän käyttö edellyttää tunnistautumista tai muuta kirjautumista. Tilastokeskuksen

tilastotuotannon järjestelmät vaativat tunnistautumista, joten lokitietojen kerääminen koskettaa virastossa laajaa joukkoa järjestelmiä.

Raportti suosittaa lokitietojen käsittelyn edelleen kehittämistä tiedonhallintalain vaatimusten mukaisesti.

7 Tiedonkeruu poikkeusoloissa

Tilastokeskus sai viranomaisten ja yksityisen sektorin tietoaaineistot pääosin sovitulla tavalla tilastotuotannon käyttöön poikkeusvuonna 2020. Koronasta johtuvat lisääjat saattoivat hidastaa aineistojen muodostumista. Suora tiedonkeruu jatkui poikkeusoloista huolimatta.

Yritysten siirtäessä työntekijänsä etätöihin Tilastokeskus julkaisi tiedonkeruusivuillaan yrityksille ohjeet sujuvaan ja tietoturvalliseen vastaamiseen haastavassa etätötilanteessa.

Tilastokeskus päätti huhtikuussa, että epidemian aikana kaikki Tilastokeskuksen tiedonkeruissa suunnitellut haastattelut tehdään puhelimitse. Käyntihaastattelujen tekeminen ja hyllyhintojen kerääminen myymälöistä lopetettiin turvallisuussyistä. Ahvenanmaalla hintojen keruu jatkui, koska keruuta ei voitu tehdä netissä.

Tilastokeskus oli varautunut tekemään vuonna 2020 kaksi hyvin laajaa ja yhteiskunnallisesti merkittävää kansainvälistä tutkimusta, jotka olisivat pohjautuneet täysin käyntihaastatteluihin. Nämä tutkimukset on siirretty alkamaan myöhempänä ajankohtana.

Verkkovastaamiseen siirtyminen onnistui, koska siihen oli ehditty aiemmin panostaa riittävästi sekä henkilö- että yrityskeruissa. Lisäksi onnistumiseen vaikutti se, että Tilastokeskus saattoi tarjota vastaajille julkishallinnon välineitä vahvaan tunnistautumiseen verkossa.

Tilanne synnytti tiedonhankinnassa voimakkaan kehittämispaineen. Puuttuvia vastauksia korvattiin puhelinhaastatteluilla ja -karhunnalla. Testattiin tietojen hakemista, esimerkiksi hintatietoja haettiin verkosta henkilötyönä ja verkkoharavoinnilla. Hintatilastoissa otettiin käyttöön merkittävä kassapäätteaineisto korvaamaan puuttuvia hintahavaintoja.

Joihinkin tiedonkeruisiin lisättiin kysymyksiä koronaan liittyen. Kysymykset koskivat mm. etätyön lisääntymistä, etäkokouksia ja muutoksia verkkokaupassa.

Tilastokeskus on raportoinut aineistotoimitusten ja suoran tiedonkeruun haasteita ja onnistumisia tarkalla tasolla vuoden 2020 tilinpäätöksessä.

Tilastokeskus pystyi uudessa tilanteessa mukauttamaan tiedonhankintaansa joustavasti, ja tiedontuotannon laatu ja toimintavarmuus saatiin pysymään hyvällä tasolla.

8 Suomi.fi-tunnistautumisen käyttöönotto alkoi

Suomi.fi on julkishallinnon asiointipalveluiden yhteinen tunnistuspalvelu, joka mahdollistaa turvallisen sähköisen tunnistamisen. Suomi.fi-tunnistuksesta vastaa Digi- ja väestötietovirasto (DVV).

Tilastokeskus otti vuonna 2020 käyttöön Suomi.fi-tunnistuksen ja -valtuudet yritys- ja yhteisökeruissa. Käyttöönotto jatkuu vuonna 2021.

Tilastokeskus laajensi Suomi.fi-tunnistautumisen koskemaan myös henkilö- ja kotitalouspohjaisia tiedonkeruita, joissa verkkolomakkeella oli esitätetty vastaajan tai hänen kotitaloutensa tietoja.

9 Väestölaskentavuosi 2020

Väestölaskenta tehdään maailman eri maissa noin joka kymmenes vuosi. YK:n suositusten mukaan kaikkien maailman maiden tulisi tehdä väestölaskenta kymmenlukuun päättyvinä vuosina, mutta käytännössä eri maissa laskennan ajankohta voi vaihdella. EU:n asetuksen (EU) 2017/712 mukaan väestölaskennan ajankohdaksi on määritelty vuosi 2021. Suomen osalta tämä tarkoittaa käytännössä vuodenvaihteen 2020/21 tilannetta, jossa tiedot kuvaavat vuoden 2020 viimeistä päivää tai viikkoa.

Suomessa väestölaskenta on nykyisin kymmenvuosittain tehtävä kooste erillisistä tilastoista, joita tuotetaan vuosittain. Tärkeä osa väestölaskentaa

on juuri kansainvälinen näkyvyys ja tietojen tuottaminen kansainvälisiin tarkoituksiin, esimerkiksi EU:ta varten. Väestölaskennassa jäsenmaat tuottavat Eurostatille laajan taulukkopaketin, jonka sisältö on tarkkaan määritelty eri asetuksissa.

Vuonna 2008 tuli voimaan Euroopan parlamentin ja neuvoston asetus väestö- ja asuntolaskennoista (EY) N:o 763/2008, joka määrittelee entistä sitovammin mm. väestölaskennan yhteydessä tuotettavan tietosisällön. Tätä puiteasetusta on lisäksi täydennetty komission säädöksillä ja toimeenpanoasetuksilla, joilla pyritään lisäämään tietojen laatua, yhtenäisyyttä ja vertailtavuutta maiden välillä. Väestölaskentojen ensisijaisena tarkoituksena on kuvata maassa asuvaa väestöä tietyssä ajanhetkenä: väestön määrää, väestörakennetta ja alueellista jakaumaa. Lisäksi laskennoissa kerätään tietoa väestön elinoloista, koulutusrakenteesta, perheistä ja asutokunnista, lasten lukumäärästä, muuttaneista, taloudellisesta toiminnasta, ammateista ja toimialasta. Väestölaskentojen yhteydessä kerätään usein myös tietoja rakennuksista ja asunnoista.

Keväällä 2020 maailmalle levinnyt covid-19 -pandemia on vaikuttanut maailmalla myös väestölaskentoihin. Yhteensä 16 maata 57:stä UNECE:n piirissä on raportoinut joutuneensa lykkäämään väestölaskennan toteuttamista johtuen covid-19 aiheuttamasta pandemiasta. Suomen väestölaskennan toteuttamiseen pandemialla ei ole vaikutusta, sillä tiedot tuotetaan hyödyntäen useita eri hallinnollisia rekistereitä.

Eurostatille tuotetaan laaja tietopaketti (Census-Hub -palvelu). Sen tietojen suojaamisessa otetaan käyttöön uudenlainen tietosuojausmenetelmä (cell key -menetelmä). EU:n väestölaskenta-asetuksen vaatimuksesta suojausta ei voida toteuttaa peittämällä pieniä solufrekvenssejä, kuten kansallisessa taulukkotuotannossa yleensä on tehty.

10 VirastoVahvan käyttöönotto

Tilastokeskuksessa otettiin lokakuussa 2020 käyttöön uusi sähköinen asianhallintajärjestelmä VirastoVahva. Järjestelmä on Valtorin toimittama TietoEvryn tuote Public 360. Käyttöönotto oli osa Tilastokeskuksen asian- ja dokumentinhallinnan kehittämishanketta, ja uudella järjestelmällä korvattiin aikaisemmin käytössä ollut Dynasty-asianhallintajärjestelmä. Dynastyssa olleet aineistot migroitiin eArchiveen, joka on sähköinen arkisto.

VirastoVahva mahdollistaa asiakirjojen sähköisen käsittelyn ja mm. sähköisen allekirjoituksen Tilastokeskuksen sisällä. Perustoiminnallisuuksien lisäksi VirastoVahvassa otettiin käyttöön sopimushallinnan moduuli.

Kaikilla tilastokeskuslaisilla on järjestelmään vähintään selaajaoikeus. Rekisteröityjen asioiden ja asiakirjojen näkyvyyttä rajoitetaan käyttöoikeuksin. VirastoVahvassa pystytään aikaisempaa paremmin ottamaan huomioon asianhallinnan ja tietosuojan vaatimukset.

11 Arkisto

Valtionhallinnon pysyvästi säilytettävät paperiasiakirjat muutetaan jatkossa digitaaliseen muotoon. Tilastokeskus on osallistunut Kansallisarkiston massadigitoitihankkeeseen vuonna 2020 valmistelemalla 280 hyllymetriä asiakirjoja digitointiin. Digitoitujen asiakirjojen tietosuoja varmistetaan aineistokohtaisin metatiedoin sekä lakiperusteisin käyttörajoituksin. Digitoidut asiakirjat ovat yleisön käytettävissä Astia-verkkopalvelun kautta käyttörajoitukset huomioiden. Aineiston muuttamisesta analogisesta digitaaliseen muotoon on tehty yhdessä Kansallisarkiston kanssa tietosuoja- ja tietoturvalähtöinen riskikartoitus.

Paperiasiakirjojen vuosittaista hävittämistä on tehty asiakirjojen säilytysaikojen mukaisesti. Analogisten tietoaineistojen kokonaismäärä oli vuoden 2020 lopussa noin 2,1 hyllykilometriä, Kansallisarkiston päätöksellä pysyvästi säilytettäviä tietoaineistoja on tästä noin 1,2 hyllykilometriä. Arkiston tilat on suojattu rajatuin kulkuoikeuksin.

12 Mittarit vuonna 2020

12.1 Rekisteritietojen tarkastuspyynnöt

Tilastokeskukselta on tiedusteltu EU:n yleisen tietosuoja-asetuksen (EU) 2016/679 artikla 15 mukaisesti tietoa siitä, mitä henkilöä itseään koskevia tietoja on tallennettuna Tilastokeskuksen henkilörekistereihin seuraavasti:

Vuosi	Tarkastuspyyntöjä, kpl
2018	5
2019	4
2020	3

12.2 Tilastot

Tilastokeskuksella oli vuonna 2020 tuotannossa 158 tilastoa, joista 144 julkaisi tietoa. Vuonna 2020 perustettiin yksi uusi tilasto, ilmaliikenne. Veronalaiset tulot -tilasto merkittiin lakkautetuksi. Lisäksi seuraavien viiden liikennetilaston tuotanto siirtyi Traficomista Tilastokeskukseen: tietilasto, kotimaan vesiliikenne, ulkomaan meriliikenne, kauppalaivasto ja rautatietilasto.

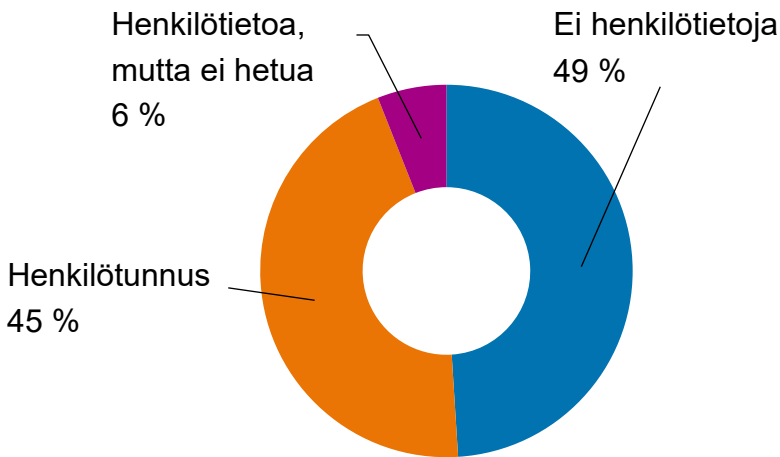
12.3 Tilastotuotannossa käytettävien aineistojen vastaanotto

Tilastokeskuksessa toimii keskitetty aineistojen vastaanottopalvelu (Ava). Seuraavat tiedot perustuvat Tilastokeskuksen tiedonkeruurekisteriin kuvattujen aineistokuvausten tietoihin.

Vuonna 2020 Tilastokeskukseen toimitti aineistoja 98 eri tiedontuottajaa. Tiedon siirtää Tilastokeskukseen joko sen tuottaja tai jokin tuottajan valtuuttama taho (esim. Tieto Oy, CGI).

Vuonna 2020 Tilastokeskukseen toimitettiin aineistoja 307 kappaletta. Näistä aineistoista 76 prosenttia toimitettiin linjasiirtona ja 24 prosenttia muulla tavoin. Tilastokeskukseen saapuu keskitetyn vastaanoton ohi jonkin verran lähinnä taulukkomuotoisia aineistoja.

12.3.1 Henkilötiedot aineistojen vastaanotossa (N=307)



Aineistoista (N=307) 45 prosenttia sisälsi henkilötunnuksia. Kuusi prosenttia sisälsi henkilötietoja, mutta ei henkilötunnuksia. Loput 49 prosenttia saaduista aineistoista ei sisältänyt henkilötietoja.

12.4 Suorat tiedonkeruut

Tilastokeskus teki vuonna 2020 suoria tiedonkeruita 93 kappaletta. Näistä keruista henkilöön tai kotitalouteen kohdistui 11 kpl eli 11 prosenttia.

Vuoden 2020 aikana tehtyjen henkilö- ja kotitalouspohjaisten suorien tiedonkeruiden (N=11) otoksissa oli yhteensä 248 830 kohdetta.

Kun suoran keruun kohteena ovat muut kuin yksityishenkilöt tai kotitaloudet, keruukohteet luokitellaan yhteisöiksi. Esimerkkeinä yhteisöistä ovat elinkeinonharjoittajat, osakeyhtiöt, avoimet yhtiöt, rekisteröidyt yhdistykset, kunnat, kuntayhtymät, rahoituslaitokset, vakuutusyhtiöt, osuuskunnat, seurakunnat, yliopistot, säätiöt ja oppilaitokset.

Kolmessa vuoden 2020 yritystiedustelussa otokseen kuului yhteisöjen lisäksi myös yksityishenkilöitä esimerkiksi rakennuksen omistajana tai ajoneuvon haltijana. Tämän vuoksi nämä tiedonkeruut luokitellaan tässä tarkastelussa mukaan henkilöpohjaisiin tiedonkeruisiin. Keruut on merkitty tähdellä (*) seuraavassa luettelossa:

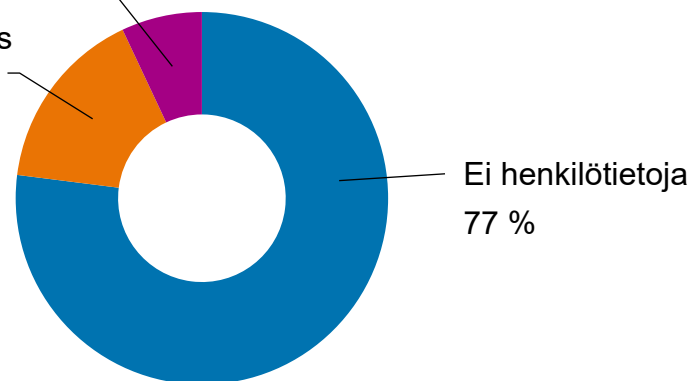
12.4.1 Henkilö- ja kotitalouspohjaiset tiedonkeruut otoskoon mukaan

1. Työvoimatutkimus	140 000
2. Suomalaisten matkailukysely	28 200
3. Kuluttajien luottamustutkimus	26 400
4. Tulo- ja elinolotutkimus	13 000
5. Tieliikenteen tavarankuljetukset*	10 000
6. Ajankäyttötutkimus	8 840
7. Tulo- ja elinolotutkimuksen yhdistelmäpilottitiedonkeruu	6 690
8. Väestön tieto- ja viestintätekniikan käyttö -tutkimus	6 000
9. Asuntojen ja omakotitalojen korjausrakentaminen*	5 000
10. Kulutustutkimus pilottitiedonkeruu	4 000
11. Teollisuus- ja varastorakennusten korjausrakentaminen*	700

12.4.2 Henkilötiedot suorassa tiedonkeruussa (N=93)

Henkilötietoa,
mutta ei hetua
7 %

Henkilötunnus
16 %



Kun tarkastellaan kaikkia vuonna 2020 tehtyjä suoria tiedonkeruita (N=93), niissä 16 prosentissa esiintyi henkilötunnuksia ja seitsemässä prosentissa puolestaan käsiteltiin henkilötietoja ilman henkilötunnusta. Lopuissa 77 prosentissa suorista keruista ei käsitelty henkilötietoja.

Henkilötiedoksi luettavaa tietoa kerätään myös yrityskeruissa, kuten esimerkiksi palkkatiedusteluissa. Tilastokeskus tarvitsee tiedonkeruussa henkilötunnuksen, jotta työnantajan toimittamia tietoja voidaan yhdistää muihin aineistoihin, esimerkiksi tutkintorekisterin tutkintotietoihin.

12.4.3 Erityisiin tietoryhmiin kuuluvat tiedot suorassa tiedonkeruussa ja aineistoissa

Erityisiin tietoryhmiin kuuluvat tiedot ovat EU:n yleisen tietosuoja-asetuksen 9 artiklassa määritellyt rotu tai etninen alkuperä, poliittinen mielipide, uskonnollinen tai filosofinen vakaumus tai ammattiliiton jäsenyys sekä geneettiset tai biometriset tiedot, terveyttä koskevat tiedot taikka luonnollisen henkilön seksuaalista käyttäytymistä ja suuntautumista koskevat tiedot. Lisäksi ryhmään kuuluvat EU:n yleisen tietosuoja-asetuksen 10 artiklassa määritellyt rikostuomioihin ja rikkomuksiin liittyvät tiedot.

Tilastokeskuksen käyttöön toimitetuista aineistoista (N=307) 18 prosenttia sisälsi joko erityisiin tietoryhmiin kuuluvia tietoja tai rikostuomioihin tai rikkomuksiin liittyviä tietoja tai molempia. Jäljelle jäävissä 82 prosentissa saaduista aineistoista ei esiintynyt erityisiin tietoryhmiin kuuluvia tietoja eikä rikostuomioihin tai rikkomuksiin liittyviä tietoja.

Vuoden 2020 suorista keruista (N=93) viisi prosenttia sisälsi joko erityisiin tietoryhmiin kuuluvia tietoja tai rikostuomioihin tai rikkomuksiin liittyviä tietoja tai molempia. 95 prosentissa tehdyistä suorista tiedonkeruista ei esiintynyt erityisiin tietoryhmiin kuuluvia tietoja eikä rikostuomioihin tai rikkomuksiin liittyviä tietoja.

12.4.4 Muutoksia suorissa tiedonkeruissa vuonna 2020

Uusiin aineistolähteisiin liittyen Tilastokeskus luopui vuonna 2020 suorista tiedonkeruista oppilaitosmuotoisen ammatillisen koulutuksen opiskelijat ja tutkinnon suorittaneet kalenterivuodelta sekä oppisopimuskoulutuksen opiskelijat ja tutkinnon suorittaneet kalenterivuodelta. Nämä tiedot saadaan jatkossa pääosin Opetushallituksen ylläpitämästä Koski-palvelusta, eli Kansallisten opiskeluoikeuksien ja suoritusten keskitetystä integraatiopalvelusta. Koski-palvelusta ei saada peruskoulujen erityisopetusta koskevia tietoja. Koski-palvelusta saatavia tietoja täydennetään edelleen keräämällä palvelun ulkopuolelle jäävää tilastoitavaa koulutusta mm. Ahvenanmaalta.

Vuonna 2020 ei enää toteutettu myöskään järjestäjä- ja oppilaitostason oppilaat ja opiskelijat -keruuta. Täydentäväksi keruuksi tuli opiskelijakysely, jonka tietoja käytetään oppilaitosrekisterin opiskelijamäärätietoihin. Kaikista Tilastokeskuksen suorista tiedonkeruista löytyy lisätietoa verkkosivulta <https://www.tilastokeskus.fi/keruu>.

12.5 Tilaustutkimukset

Toimeksiantoina asiakkaille tehtävien tiedonkeruiden aiheet ja kokoluokka vaihtelevat paljon. Toimeksiantoina tehdään ainoastaan yhteiskunnallisesti tai tieteellisesti merkittäviä tutkimuksia. Osa keruista on säännöllisesti toistuvia, osa kertaluonteisia. Tutkimuksia tilaavat pääasiassa julkishallinto ja tutkimuslaitokset. Anonymisoitu aineisto voidaan toimittaa tapauskohtaisesti toimeksiantajan lisäksi Yhteiskuntatieteelliseen tietoaarkistoon (FSD).

Toimeksiantoina tehtäviä tutkimuksia koskevat samat salassapitovelvoitteet kuin Tilastokeskuksen omiakin tiedonkeruita. Henkilötietoja käsitellään aina luottamuksellisesti, eikä Tilastokeskus missään olosuhteissa luovuta eteenpäin henkilön tunnistamisen mahdollistavia tietoja.

12.5.1 Toimeksiantojen lukumäärä

Vuosi	Toimeksiannot, kpl
2010	23
2011	15
2012	11
2013	10
2014	7
2015	14
2016	11
2017	12
2018	12
2019	17
2020	12

12.5.2 Henkilöpohjaiset tilaustutkimukset 2020 otoskoon mukaan

Tilaustutkimuksen nimi	Otoskoko
Kansalaispulsssi, Valtioneuvosto	40 000
Suhteellinen tulotieto, Hanken	22 500
Tutkimus kuluttajien inflaatio-odotuksista, Suomen Pankki	20 000
Työvoimapalvelut osana Avointen työpaikkojen tutkimusta, TEM	10 000
Tasa-arvo työpaikoilla 2020, STM	8 000
Luonnon virkistyskäytön valtakunnallinen inventointi, Luke	4 000
Luottamus julkiseen hallintoon, Valtioneuvosto ja OECD	2 200
Työolobarometri osana Työvoimatutkimusta, TEM	2 200
Yksityisen sektorin palkkatiedustelu: Palvelulaitosten (AVAINTA ry) tiedot	<700

12.6 Asianhallinta ja arkistointi

12.6.1 Digitaalisten tilastoaineistojen arkistointi

Digitaalisten tilastoaineistojen arkistointi on osa tilastotuotantoprosessia. Arkistoinnin tarkoituksena on varmistaa tutkimus- ja tilastointitarkoituksen kannalta arvokkaan tiedon pysyvä säilyminen hyödynnettävässä muodossa. Tilastoaineistojen huolellinen kuvaaminen eli metatietojen tallentaminen on oleellista aineistojen ymmärtämiseksi.

Tilastokeskus arkistoi pysyvästi Kansallisarkiston päätöksen mukaisesti jokaisesta tilastosta julkaisun tuottamiseen kerätyn tarkimman tason, yleensä havaintoyksikkötason (esim. yritys, henkilö, kotitalous) tilastoaineiston. Jos havaintoyksikkökohtaista tilastoaineistoa (=mikroaineisto) ei kerätä, arkistoidaan summattu eli aggregoitu aineisto. Arkistoitava tilastoaineisto sisältää datan ja sen sisällön ja rakenteen kuvaavat metatiedot, mm. muuttujakuvaukset, käsitteet ja luokitukset. Tilastoaineistojen arkistointien vuosittainen toteuma on yksi mittari tulosityksiköiden tuloskorteissa.

Tilastoaineistojen pysyvä arkistointi on osa tietoaineistojen tiedonhallinnan kokonaisuutta. Arkistolakiin perustuen Tilastokeskus arkistoi pysyvästi tilastoaineiston elinkaaren loppuvaiheen aineiston. Kaikkien muiden, määräajan säilytettävien, tilastoaineistojen säilytysajoista päättää Tilastokeskus. Tiedon säilyttämisen laatuksiterit tulee huomioida jo tiedon elinkaaren alkupäässä.

12.6.2 Arkistoidut digitaaliset tilastoaineistot 2011–2020

Vuosi	Arkistoituja aineistoja, kpl
2011	106
2012	182
2013	122
2014	83
2015	20
2016	38
2017	11
2018	211
2019	155
2020	114

12.7 Tutkimusaineistot

Tilastolain mukaan Tilastokeskus voi myöntää käyttöoikeuden yksikkötasoisiin tutkimusaineistoihin käytettäväksi tieteellisissä tutkimuksissa ja yhteiskuntaoloja koskevissa tilastollisissa selvityksissä. Tutkimusaineistojen pohjana ovat Tilastokeskuksen tilastointia varten saadut aineistot, joihin usein yhdistetään muiden tahojen tai tutkijan omia aineistoja.

Tilastokeskuksen tutkijapalveluissa muodostetaan sekä tutkimushankekohtaisia tutkimusaineistoja että etukäteen valmistettuja tutkimusaineistoja. Ns. valmisaineistojen kuvaukset löytyvät tutkimusaineistokatalogista Taikasta (stat.taika.fi).

Sosiaali- ja terveysalan tietolupaviranomainen Findata voi vuodesta 2020 alkaen myöntää käyttölupia Tilastokeskuksen kuolemansyyaineistoihin.

12.7.1 Tutkimusaineistojen käyttölujarahakemukset 2017–2020

	2017	2018	2019	2020
Kuolemansyyaineistot	82	66	77	70
Käyttölupien jatkot ja laajennukset	166	162	226	195
Mikrosimulointi	20	12	18	17
Valmisaineistojen etäkäyttö	50	77	79	94
Räätälöityjen tutkimus-aineistojen etäkäyttö	87	95	93	57
Räätälöityjen tutkimus-aineistojen luovutus	35	20	14	5
Tunnisteellisena luovutet-tavat tutkimusaineistot	33	34	42	33
Toisen viranomaisen lupa	5	3	3	3
Palveluaineistot	-	-	-	22
Ulkopuolisen aineiston lisääminen-	-	-	-	13
Yhteensä	478	469	552	509

12.7.2 Kansallinen etäkäyttöpalvelu

Tilastokeskus luovuttaa ensisijaisesti tutkimusaineistot tutkijoiden käyttöön tietoturvalisessa etäkäyttöjärjestelmässä. Fiona-etäkäyttöjärjestelmässä työskentely vastaa omalla koneella työskentelyä, mutta järjestelmästä ulos siirrettävät tiedostot ja tulokset tarkistetaan tietosuojaan osalta sekä tutkijan että Tilastokeskuksen toimesta. Tilastokeskus tarjoaa Fiona-etäkäyttöjärjestelmää tieturvalisena ratkaisuna myös muiden viranomaisten tutkimusaineistoluovutuksiin.

12.7.3 Fiona-etäkäyttöpalvelu 2017–2020

	2017	2018	2019	2020
Etäkäytön organisaatio-sopimukset	71	88	91	96
Etäkäyttöhankeet	163	196	219	197
Henkilöitä ¹	363	466	584	659
Käyttäjätunnuksia ²	501	601	782	983
Räätälöityjen tutkimusaineistojen etäkäyttö	87	95	93	57
Räätälöityjen tutkimusaineistojen luovutus	35	20	14	5

1 Henkilö laskettu vain kerran välittämättä siitä, vaikka olisi useassa hankkeessa mukana.

2 Henkilö huomioitu lukumäärään jokaisella käyttäjätunnuksellaan.

Tilastokeskus 

www.tilastokeskus.fi

www.facebook.com/Tilastotohtori

www.twitter.com/tilastokeskus