

# Tilastokeskuksen tietotilinpäätös 2017, raportti henkilötietojen käsittelystä

26.6.2018

## Sisällys

1	Tiivistelmä.....	4
2	Tilastokeskuksen strategia.....	5
	2.1.1 Kuva: Strategiset päämäärät.....	5
	2.1.2 Kuva: Kriittiset menestystekijät.....	6
3	Tilastokeskuksen rooli Suomen tilastotoimessa.....	7
4	Lainsäädäntö, tietosuoja ja tietoturva.....	8
	4.1 Tilastokeskuksen toimintaa ohjaava lainsäädäntö.....	8
	4.1.1 Laki Tilastokeskuksesta.....	8
	4.1.2 Tilastolainsäädäntö.....	8
	4.1.3 Henkilötieto ja henkilötietolaki.....	8
	4.1.4 Laki viranomaisten toiminnan julkisuudesta.....	9
	4.2 Muuttuvan tietosuojalainsäädännön vaikutus Tilastokeskuksen toimintaan.....	9
	4.3 Tietoturva.....	9
	4.3.1 Tietoturvan hallinta.....	9
	4.3.2 Käyttövaltuuksien hallinta.....	9
	4.3.3 Tekninen tietoturva.....	9
	4.3.4 Tietoturvapoikkeamien raportointi.....	10
	4.3.5 Jatkuvuus ja riskienhallinta.....	10
	4.3.6 Kuva. Riskienhallintaprosessi.....	10
	4.4 Tilastojen ja taulukoiden tietosuoja.....	11
	4.5 Yksikköaineisto ja taulukkoaineisto.....	11
	4.6 Tilastolliset tietosuojamenetelmät.....	11
	4.7 Pseudonymisointi eli tunnusten korvaaminen satunnaisilla tunnisteilla.....	12
	4.8 Sisäinen tarkastus.....	13
	4.9 Henkilöstön osaamisen varmistaminen.....	14
5	Tilastojen laatu, ammattietiikka ja käytännesäännöt (CoP).....	15
	5.1 Tilastojen laatu.....	15
	5.2 Ammattietiikka ja käytännesäännöt.....	15
	5.2.1 Kuva: Euroopan tilastoja koskevat käytännesäännöt.....	16
	5.3 Vertaisarviointit.....	16
6	Tiedonhankinta.....	17
	6.1 Tiedonhankinnan keskittäminen.....	17
	6.2 Tiedonkeruun yleiset periaatteet.....	17
	6.2.1 Tiedonantajien informointi.....	17
	6.2.2 Tiedonantorasitteen minimointi.....	18
	6.2.3 Tiedonkeruumenetelmät.....	18

6.3	Aineistojen käyttö ja aineistojen ajantasaisuus .....	19
6.3.1	Suurimmat aineistojen toimittajat .....	19
6.4	Tilastokeskuksen henkilöihin ja kotitalouksiin kohdistuvat tiedonkeruut .....	20
<b>7</b>	<b>Tilastotuotanto .....</b>	<b>21</b>
7.1	Tuotettavat tilastot.....	21
7.2	Henkilötietojen elinkaari .....	21
7.2.1	Kuva: tiedonhankinnan tietovarannot.....	22
7.3	Tilastotuotannon prosessimalli.....	23
7.3.1	Kuva: tilastotuotannon prosessimalli.....	23
7.4	Tilastotietovarannot.....	23
7.5	Nykyinen henkilötietovaranto .....	24
7.6	Nykyinen palkkatietovaranto .....	25
7.7	Tieliikenneonnettomuudet.....	26
7.7.1	Kuva: Tietojärjestelmät.....	27
<b>8</b>	<b>Tilastotiedon jakelu .....</b>	<b>28</b>
8.1	Tilastojen julkistamisperiaatteet Tilastokeskuksessa .....	28
8.2	Avoin data .....	28
<b>9</b>	<b>Tietojen luovuttaminen.....</b>	<b>28</b>
9.1	Tietojen luovutus tilastotarkoituksiin.....	28
9.2	Tietojen luovutus tutkimuskäyttöön .....	29
9.2.1	Henkilöaineistojen käsittely edellyttää Tilastokeskuksen käyttö lupaa .....	29
9.2.2	Taulukko: Vuonna 2017 myönnettyjen käyttö lupien ja aineistoluovutusten määrät tuoteluokittain .....	30
9.2.3	Sisu-mikrosimulointimalli.....	30
9.2.4	Etäkäyttöjärjestelmä.....	30
9.3	Yritys- ja toimipaikkarekisteri.....	31
9.4	Yritystiedonantajarekisteri.....	32
<b>10</b>	<b>Toimeksiantoina tehtävä henkilötietojen käsittely .....</b>	<b>33</b>
10.1	Asiakasrahoitteiset haastattelu- ja tiedonkeruupalvelut.....	33
10.1.1	Taulukko. Asiakasrahoitteiset tiedonkeruut.....	33
10.2	Asiakaskohtaiset tietopalvelusopimukset .....	34
10.3	Tilastointipalvelut.....	34
10.4	Kansallinen etäkäyttöpalvelu.....	34
<b>11</b>	<b>Asiakas- ja sidosryhmätiedot.....</b>	<b>35</b>

12	Asianhallinta ja arkistointi.....	36
12.1	Asianhallinta ja analoginen asiakirja-aineisto.....	36
12.2	Tilastoaineiston arkistointi.....	36
13	Kehittämistoimenpiteet vuodelle 2018.....	37

## 1 Tiivistelmä

Tämä on Tilastokeskuksen ensimmäinen tietotilinpäätös. Tietotilinpäätöksessä kuvataan yleisellä tasolla Tilastokeskuksen tilastojen tuottamisen kokonaisprosessi, toimeksiantona tehtävä tietojen keruu ja tutkimus, tietopalvelu tilastollisesta perusrekisteristä sekä tietojen luovuttaminen tutkimuskäyttöön, jotta lukija pystyisi hahmottamaan henkilötietojen käsittelyn lähtökohdat ja niihin liittyvät toimintatavat sekä tietojen käsittelyn laajuuden Tilastokeskuksessa. Raportin tiedot kuvaavat vuotta 2017.

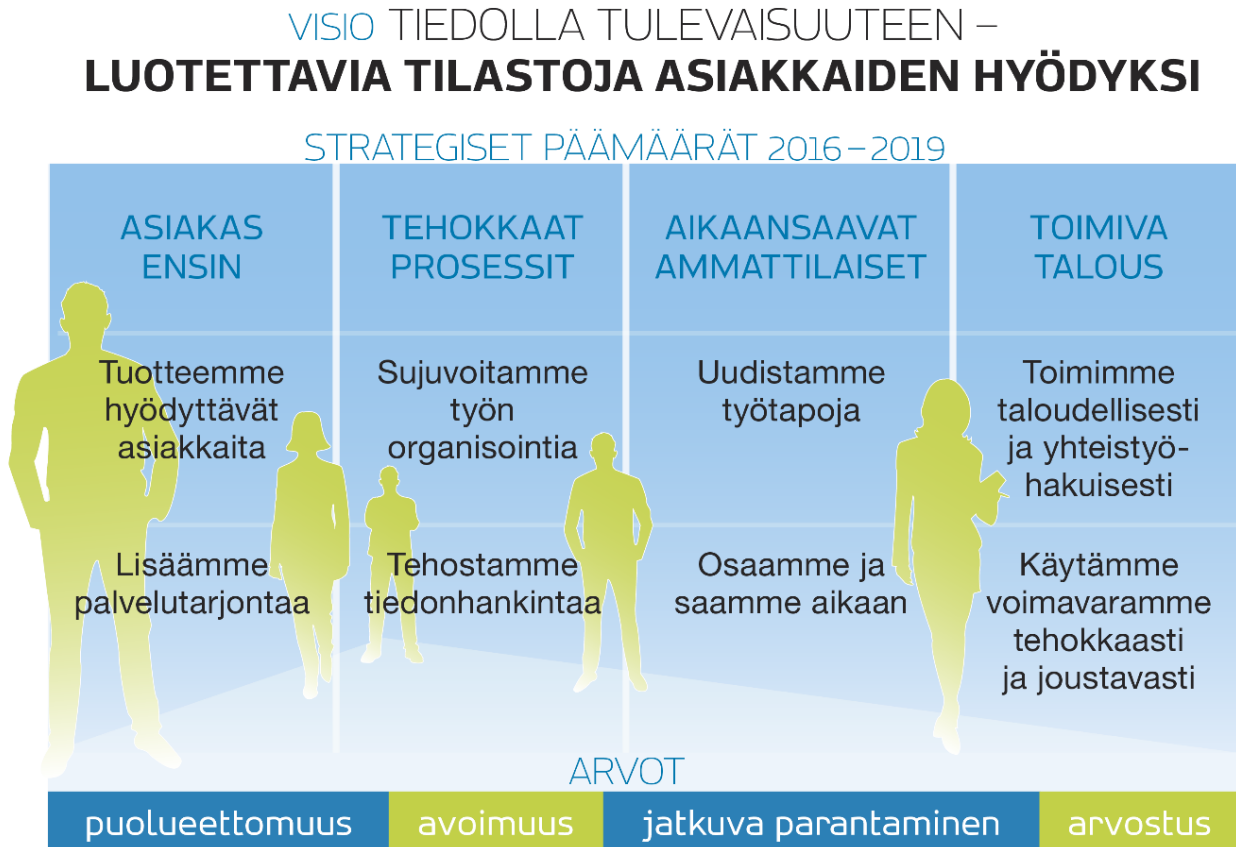
Tilastokeskus tuottaa lähes 160 eri tilastoa, joista suurin osa kuuluu Suomen virallisen tilaston kokonaisuuteen. Kaikista tilastojen tietoaineistoista valtaosa saadaan julkishallinnon rekistereistä tai muista hallinnollisista lähteistä. Tilastokeskus kerää tietoja myös kyselyillä ja haastatteluilla silloin, jos tarvittavia tietoja ei voida muuten saada. Yksityishenkilöiltä tai kotitalouksilta Tilastokeskus tiedustelee suoraan tietoja vuosittain noin kymmenessä suorassa tiedonkeruussa. Suorista henkilö- tai kotitalouspohjaista tiedonkeruista saatavia tietoja Tilastokeskus hyödyntää kahdessatoista eri tilastossa.

Lisäksi Tilastokeskus käsittelee henkilötietoja tutkimusaineistoja muodostettaessa, erilaisissa toimeksiantoissa sekä asiakas- ja sidosryhmäsuhteiden hoitamisessa.

Avainsanat: tietotilinpäätös, tietosuoja, tietovirrat, aineistotyypit

## 2 Tilastokeskuksen strategia

### 2.1.1 Kuva: Strategiset päämäärät



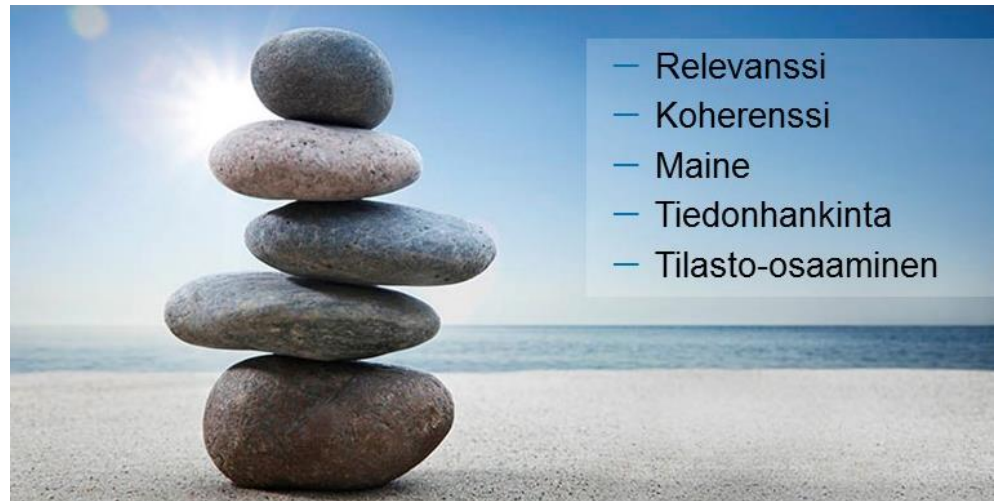
Tilastokeskuksen toiminta-ajatuksen lähtökohtana ovat lainsäädännössä määritellyt tehtävät: *Tilastokeskus tukee demokraattista ja tietoon perustuvaa päätöksentekoa sekä tutkimusta tuottamalla yhteiskuntaoloja kuvaavia luotettavia tilastoja, selvityksiä ja aineistoja. Tilastokeskus kehittää valtion tilastointia yhteistyössä muiden valtion viranomaisten kanssa.*

Strategiakaudella 2016–2019 Tilastokeskuksen tavoitteena on nostaa asiakkaat toiminnan keskiöön. Tilastokeskuksen visio on *Tiedolla tulevaisuuteen – luotettavia tilastoja asiakkaiden hyödyksi*. Strategian kolme keskeistä teemaa ovat asiakkaat, digitalisaatio ja toimintatavat. Strategian mukaan jokapäiväisessä työssä tavoitteena tulee olla hyödyn tuottaminen asiakkaille. Asiakkaita ovat paitsi tiedonkäyttäjät, myös tiedonantajat eli kansalaiset, yritykset ja yhteisöt.

Tiedonantajien ja tiedonkäyttäjien näkemykset tietojen salassapidosta voivat olla hyvinkin erilaiset. Jotkut tiedonkäyttäjryhmät voivat haluta hyvinkin yksityiskohtaista tietoa. Tiedonantajien kannalta on tarpeen, etteivät yksittäisten tiedonantajien tiedot paljastu. Tilastokeskus arvioi jatkuvasti yksilön tietosuojaa ja tiedonkäyttäjien tarvetta yksityiskohtaiseen tietoon.

Tilastokeskuksen toiminnassa on tunnistettu viisi kriittistä menestystekijää. Näistä maine ja maineen säilyttämistä uhkaavat riskit liittyvät läheisesti tilastotuotannon ja -tuotteiden laatuun ja luotettavuuteen sekä tietosuojaan ja tietoturvaan.

### 2.1.2 Kuva: Kriittiset menestystekijät



### 3 Tilastokeskuksen rooli Suomen tilastotoimessa

Valtion tilastotoimi on tilastoalan viranomaisten verkosto, joka tuottaa Suomen virallisia tilastoja (SVT), Euroopan tilastojärjestelmän (ESS) tilastoja sekä muita tilastoja. Sen tehtäviin kuuluu sellaisten yhteiskuntaoloja ja niiden kehitystä kuvaavien tilastojen tuottaminen, joita yhteiskunnassa yleisesti tarvitaan. Valtion tilastotoimeen kuuluu 14 virastoa ja laitosta, joista neljä toimii tilastoviranomaisina, kahdeksan muuna tilastoja laativana viranomaisena ja kaksi muuna tilastoja laativana laitoksena.<sup>1</sup>

Tilastokeskus on tilastolaissa määritelty tilastotoimen yleisviranomaiseksi. Yleisviranomaisena Tilastokeskuksen tehtävänä on yhteistyössä muiden tilastontuottajien kanssa huolehtia siitä, että kansallisiin ja kansainvälisiin velvoitteisiin perustuvat tilastointitehtävät hoidetaan.

Tilastokeskuksen tehtävänä kansallisessa koordinoinnissa on edistää ja tukea laadukkaiden virallisten tilastojen tuotantoa. Suomen virallisen tilaston neuvottelukunta (SVT-neuvottelukunta) perustettiin vuonna 2002 Tilastokeskuksen koordinoitiroolin tehostamiseksi ja yhteistyön tiivistämiseksi tilastoja tuottavien viranomaisten välillä.<sup>2</sup> Neuvottelukuntaan kuuluvat kaikki virallisia tilastoja laativat viranomaiset ja laitokset. Sen tehtävänä on tilastotoimen kehittämislinjausten ja Euroopan tilastojärjestelmän kansallisen tason koordinoinnin tavoitteiden edistäminen.

Vuosille 2016–2017 asetetun neuvottelukunnan tehtävänä oli säännöllisesti käsitellä tilastojen julkaisemiseen, laatuun ja tuotannon kehittämiseen liittyviä kysymyksiä, edistää kansallisen tilastotoimen koordinointia ja yhteistyötä Euroopan tilastojärjestelmän tavoitteiden mukaisesti, edistää virallisen tilaston näkyvyyttä ja tunnettuutta kotimaassa ja kansainvälisesti, laatia Suomen virallisen tilastotoimen kehittämislinjaukset ja seurata niiden täytäntöönpanoa sekä seurata kansallisen ja EU:n tilastolain muutoksia ja yhteistyössä sopeuttaa tilastotoimea muutosten mukaisesti.

---

<sup>1</sup> <http://tilastokeskus.fi/org/tilastotoimi/index.html>

<sup>2</sup> <http://www.stat.fi/meta/svt/svt-ntk.html>



## 4 Lainsäädäntö, tietosuoja ja tietoturva

### 4.1 Tilastokeskuksen toimintaa ohjaava lainsäädäntö

#### 4.1.1 Laki Tilastokeskuksesta

Tilastokeskuksesta annetun lain mukaan Tilastokeskuksen tehtävänä on laatia yhteiskuntaoloja koskevia tilastoja ja selvityksiä sekä huolehtia valtion tilastotoimen yleisestä kehittämisestä yhteistyössä muiden valtion viranomaisten kanssa. Tilastokeskus on valtion tilastotoimen yleisviranomainen.

#### 4.1.2 Tilastolainsäädäntö

Suomen tilastotoimea velvoittavat EU:n antamat tilastoasetukset. EU:n toimivalta Euroopan tilastojen tuottamisessa perustuu Euroopan unionin toiminnasta tehdyn sopimuksen (SEUT) artiklaan 338. EU:n tilastosäädökset ovat yleisimmin neuvoston ja parlamentin asetuksia, joita täydennetään komission antamalla täytäntöönpanosäädöksillä. Keskeisin asetuksista on Euroopan tilastoista annettu Euroopan parlamentin ja neuvoston asetus (EY) No 223/2009 eli EU:n tilastolaki, joka luo yleiset puitteet tilastojen laadinnalle Euroopan unionissa. Tilastokohtaiset asetukset sääntelevät EU-maiden tilastojen yhdenmukaistamista, kuten luokituksia ja muita tilastostandardeja sekä komissiolle toimitettavia tietoja ja noudatettavia aikatauluja.

Tilastolaki on koko valtion tilastotoimea koskeva yleislaki. Sen tarkoituksena on varmistaa yhteiskunnallista päätöksentekoa ja suunnittelua varten yleisesti tarvittavan luotettavan tiedon saanti. Tavoitteena on myös kansainväliseen tilastoyhteistyöhön liittyvien velvoitteiden toteuttaminen, tilastotoimen yhtenäistäminen ja tehostaminen sekä hyvän tilastotavan toteutuminen. Lain tavoitteena on selkeyttää menettelytapoja suhteessa tietojen antajiin ja käyttäjiin sekä varmistaa tiedonantajien ja tiedon kohteena olevien oikeuksien toteutuminen. Lain tarkoituksena on lisäksi edistää tilastotarkoituksiin kerättyjen tietojen käyttöä tieteellisessä tutkimuksessa.

Tilastolaki ohjaa Tilastokeskuksen ja muidenkin valtion viranomaisten tilastojen laadintaa. Laissa säädetään Tilastokeskuksen oikeudesta kerätä tietoja tiedonantovelvollisuuden perusteella. Tilastolain mukaan tilastoinnissa tulee ensi sijassa käyttää muussa yhteydessä kerättyjä tietoja. Suurin osa tiedoista saadaankin erilaisista rekistereistä. Tiedonantajilta kerätään vain ne välttämättömät tiedot, joita ei saada muualta. Valtion viranomaisilla on yleinen tiedonantovelvollisuus hallussaan olevista tiedoista. Yrityksillä, kunnallisilla organisaatioilla ja voittoa tavoittelemattomilla yhteisöillä on velvollisuus antaa tietoa laissa erikseen määritellyissä asioissa. Yksittäisellä kansalaisella ei ole velvollisuutta antaa tietoja Tilastokeskukselle, vaan tietojen antaminen on vapaaehtoista.

#### 4.1.3 Henkilötieto ja henkilötietolaki

Henkilötiedolla tarkoitetaan kaikenlaisia luonnollista henkilöä taikka hänen ominaisuuksiaan tai elinolosuhteitaan kuvaavia merkintöjä, jotka voidaan tunnistaa häntä tai hänen perhettään tai hänen kanssaan yhteisessä taloudessa eläviä koskeviksi. Esimerkiksi nimi ja osoite muodostavat henkilötiedon samoin kuin myös esimerkiksi tarkka ammatti, kunta ja ikä. Henkilötiedoilla tarkoitetaan tunnistettavia tietoja kaikista luonnollisista henkilöistä myös heidän toimiessaan ammatinharjoittajana.

Henkilötietojen käsittelyyn tilastojen laadinnassa sovelletaan tilastolain lisäksi henkilötietolakia. Henkilötietolaki edellyttää tilastojen laadinnassa

tietojenkäsittelyprosessin huolellista suunnittelua ennen henkilökisterien perustamista. Tietojen tarve selvitetään huolellisesti ja tietojen antajilta pyydetään vain tilaston laatimisen kannalta välttämättömät tiedot. Vastaavanlainen suunnitteluvaikeus tulee myös tilastolaista.

#### 4.1.4 Laki viranomaisten toiminnan julkisuudesta

Tilastotarkoituksiin kerättyjen tietojen salassapidosta säädetään laissa viranomaisten toiminnan julkisuudesta (621/1999, julkisuuslaki), jonka mukaan tilastoviranomaiselle tilastojen laatimista varten annetut asiakirjat ja tilastojen perusaineistot ovat salassapidettavia. Poikkeuksena salassapidosta ovat valtion ja kunnallisten viranomaisten toimintaa ja taloutta kuvaavat julkiset tiedot sekä eräät yrityksiä koskevat tiedot, joista on säädetty erikseen.

### 4.2 Muuttuvan tietosuojalainsäädännön vaikutus Tilastokeskuksen toimintaan

EU:n yleisen tietosuoja-asetuksen voimaantulon jälkeen Tilastokeskuksessa asetettiin projekti, jonka tehtävänä oli selvittää henkilötietojen käsittelyn nykytila Tilastokeskuksessa. Projekti kartoitti vuoden 2016 syksyllä tuotantokäytössä olevat tilastoaineistot, niiden käsittely- ja tallennusympäristöt, henkilökistereiden rekisteriselosteet, arkistoinnin tilanteen, voimassa olevat palvelujen ostoihin liittyvät sopimukset, tietojen luovutukseen liittyvät sopimukset, erilaiset lähiaikoina tehdyt auditoinnit ja tarkastukset, tietosuojaan ja tietoturvaan liittyvän ohjeistuksen sekä henkilöstön tietosuojaan ja tietoturvaan liittyvän perehdytyksen ja koulutuksen.

Projektin tekemän selvitystyön yhteydessä todettiin lisäksi, että Tilastokeskuksessa on käynnissä useita projekteja, jotka edesauttavat ja ovat osin myös edellytys EU:n yleisen tietosuoja-asetuksen vaatimusten täyttymiselle. Nämä projektit liittyvät henkilötunnuksen pseudonymisointiin eli tunnusten korvaamiseen satunnaisilla tunnisteilla, käyttövaltuushallinnan kehittämiseen, dokumentin- ja asianhallintaan sekä lokipolitiikan täytäntöönpanoon.

Tilannekartoituksen jälkeen keväällä 2017 käynnistettiin toteutusprojekti, jonka tehtävänä on mm. varmistaa tietosuoja-asetuksen edellyttämä rekisteröityjen oikeuksien toteutuminen sekä sopimusten asetuksenmukaisuus, ohjeiden päivittäminen, koulutuksen ja perehdytyksen uudistaminen sekä asetuksen edellyttämän osoitusvelvollisuuden varmistaminen. Projektin määräaika päättyy kesäkuussa 2018.

### 4.3 Tietoturva

#### 4.3.1 Tietoturvan hallinta

Tilastokeskuksen tietoturvaperiaatteet toteuttavat viraston toimintaan liittyvää lainsäädäntöä. Tietoturvaperiaatteet ovat johdon hyväksymät ja henkilökunta soveltaa niitä työssään. Tilastokeskus hyödyntää toiminnassaan myös julkisen hallinnon digitaalisen turvallisuuden johtoryhmän (VAHTI) laatimaa ohjeistusta.

#### 4.3.2 Käyttövaltuuksien hallinta

Tilastokeskuksessa käyttövaltuudet perustuvat työtehtäviin. Valtuuksien myöntämisessä noudatetaan pienimmän käyttövaltuuden periaatetta, eli oikeudet annetaan vain niihin tietoihin, joita työtehtävien tekeminen edellyttää.

#### 4.3.3 Tekninen tietoturva

Tilastokeskuksen järjestelmien teknisestä ylläpidosta vastaa Valtion tieto ja viestintätekniikkakeskus Valtori yhdessä Tilastokeskuksen tietohallinnon kanssa.

Tilastokeskuksen tietohallinto ylläpitää ja kehittää viraston omia tilastointitarkoituksiin olevia järjestelmiä ja sovelluksia. Näiden tietoturvasta huolehditaan mm. tekemällä säännöllisiä auditointeja. Auditoinnit suorittaa aina ulkopuolinen toimija.

Tilastokeskuksen alihankkijat tai palveluntuottajat, joiden henkilöstö liikkuu Tilastokeskuksen tiloissa tai joiden henkilökunta käsittelee salassa pidettäviä tietoja Tilastokeskuksen toimeksiannosta, allekirjoittavat salassapitositoumuksen. Henkilötietojen käsittelyyn liittyvät velvoitteet kuten salassapitositoumuksen allekirjoittaminen on kirjattu myös tietojen käsittelijöiden kanssa tehtyihin sopimuksiin.

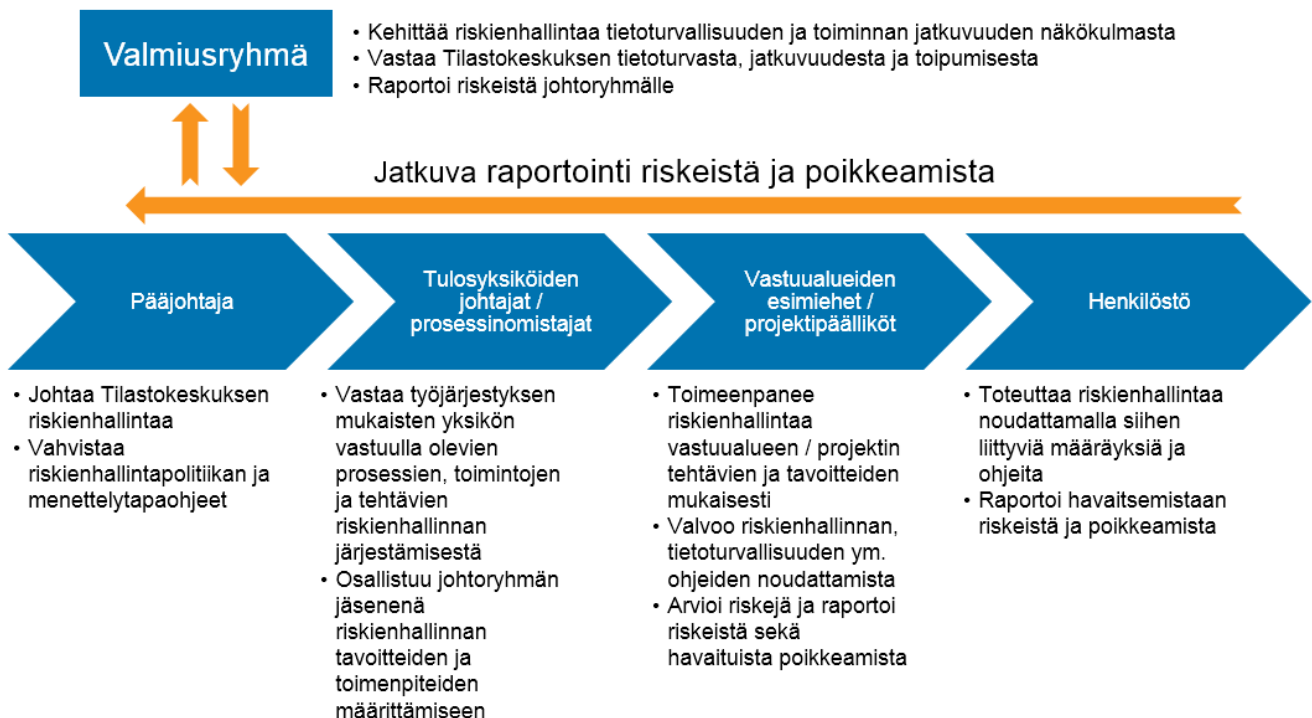
#### 4.3.4 Tietoturvapoikkeamien raportointi

Tilastokeskuksen järjestelmiin ja toimintaympäristöön kohdistuneet tietoturvapoikkeamat raportoidaan ja kirjataan keskitettyyn poikkeamienhallintajärjestelmään. Poikkeamien arvioinnista vastaa viraston valmiusryhmä. Tietoturvapoikkeamia käsitellään viraston johtoryhmässä vähintään kaksi kertaa vuodessa.

#### 4.3.5 Jatkuvuus ja riskienhallinta

Toiminnan jatkuvuudesta ja riskienhallinnasta vastaa viraston valmiusryhmä. Ryhmä kartoittaa riskejä osana riskienhallintaprosessia yhdessä viraston johdon kanssa. Riskit arvioidaan ja niille nimetään vastuutahot. Tietojärjestelmiin kohdistuvia riskejä seurataan viraston tietohallinnossa.

#### 4.3.6 Kuva. Riskienhallintaprosessi



#### 4.4 Tilastojen ja taulukoiden tietosuoja

Tilastokeskuksen tilastotarkoituksiin keräämät tiedot ovat eräitä tilastolaissa säädettyjä poikkeuksia lukuun ottamatta salassapidettäviä. Tietosuojan tavoitteena on turvata tiedonantajien yksityisyys sekä liike- ja ammattisalaisuus. Tietosuoja varmistetaan lakien, määräysten, ohjeiden, menetelmien ja näiden soveltamisen avulla. Tietosuoja otetaan huomioon jo tiedonkeruuvaiheessa ja sen suunnittelussa huolehtimalla, ettei tilastojen laatimisen kannalta tarpeettomia tietoja kerätä. Tietoja käsitellään tietosuojaperiaatteita noudattaen eli salassapidettävien tietojen paljastuminen ulkopuolisille estetään tietojen käsittelyn kaikissa vaiheissa.

Tietosuojaperiaatteiden tarkoituksena on ylläpitää luottamuksellisia suhteita tiedonantajiin. Näin varmistetaan kattavien ja luotettavien tietojen saaminen tilastojen laadintaa varten. Tietosuojaperiaatteet rajaavat sen, mitä tietoja voidaan kerätä, miten tietoja voidaan käsitellä sekä missä muodossa ne voidaan julkaista. Tietosuoja vaikuttaa siten tilastojen laatimisessa sovellettaviin menettelytapoihin ja julkaistavien tilastojen sisältöön.

Tilastokeskuksen julkaisemat tai luovuttamat tiedot ovat sellaisessa muodossa, ettei tietojen kohde tai kohdetta koskevia salassapidettäviä tietoja paljastu. Julkaistavat taulukkoaineistot suojataan käyttämällä soveltuvia tilastollisia tietosuojamenetelmiä. Tietosuojamenetelmät voivat vaikuttaa aineistojen sisältöön, joten suojauksessa kiinnitetään huomiota tietojen laatuun, arkaluonteisuuteen sekä aineiston käyttötarkoitukseen sopivan menetelmän valintaan. Salassapidettäviä tietoja voidaan julkaista suojaamattomina tietojen kohteen kirjallisella suostumuksella. Tätä mahdollisuutta on käytetty pääasiassa vain suuryritysten tietoihin, ei koskaan henkilötietoihin.

#### 4.5 Yksikköaineisto ja taulukkoaineisto

Aineistoja, jotka sisältävät yksikkökohtaisia tietoja, esimerkiksi henkilöistä tai yrityksistä, kutsutaan yksikköaineistoiksi. Yksikköaineistoissa tiedot kohdistuvat suoraan tiettyyn yksikköön ja siksi suojaamattomasta aineistosta on yleensä helppo päätellä tiedon kohteita, vaikka suorat tunnistet olisivatkin poistettu.

Taulukkoaineistosta (taulukoista) puhutaan, kun viitataan taulukkomuotoon järjestettyyn aggregoituun aineistoon.

#### 4.6 Tilastolliset tietosuojamenetelmät

Salassapidettävien tietojen suojaamiseen on käytettävissä erilaisia tilastollisia tietosuojamenetelmiä. Kaikkien tietosuojamenetelmien tavoitteena on saada aineisto muokattua muotoon, josta yksikön tai sen ominaisuuksien paljastaminen on lähes mahdotonta. Tietoja ei pidetä tunnistettavina, kun tunnistaminen vaatisi kohtuuttomasti aikaa, kustannuksia tai työvoimaa. Taulukkoaineistoille ja yksikköaineistoille on olemassa omat, toisistaan eroavat tietosuojamenetelmänsä.

Sekä taulukko- että yksikköaineistot voivat poiketa toisistaan niin sisällön, rakenteen, julkaisukonseptin, käyttäjien kuin käyttötarkoituksenkin perusteella. Eri tietosuojamenetelmien vaikutukset suojattavaan aineistoon voivat poiketa suurestikin toisistaan. Näiden syiden takia yksityiskohtaisia, kaikkiin taulukko- tai yksikköaineistoihin sopivia suojausohjeita on vaikea antaa. Käytettävät suojausmenetelmät tulisi valita aina tilannekohtaisesti.

Suojausmenetelmän valinnassa tärkeimmät kriteerit ovat vaaditun suojaustason saavuttaminen ja aineiston tärkeiden ominaisuuksien säilyttäminen. Mitä paremmin kunkin aineiston erityispiirteet on otettu huomioon, sitä paremmin aineisto voidaan suojata säilyttäen samalla sen hyödyllisyys. Usein menetelmän valintaan vaikuttaa eniten käytettävissä olevat resurssit, kuten aika ja saatavilla oleva ohjelmisto, jolla suojaus toteutetaan. Suojausmenetelmän tulisi olla myös läpinäkyvä siten, että suojatun aineiston käyttäjä ymmärtää käytetyn suojausmenetelmän pääpiirteet ja näin osaa ottaa huomioon suojauksen mahdollisesti aiheuttamat muutokset aineistossa.

Taulukoiden suojauskäytäntöihin on tehty yleisiä suojausperiaatteita. Yksikköaineistot sisältävät yleensä taulukkoaineistoa enemmän tietoa, joten niihin kohdistuu taulukoita suurempi paljastumisriski. Näin ollen yksikköaineistojen suojaus suunnitellaan ja toteutetaan aina aineistokohtaisesti. Mitä enemmän ja arkaluonteisempaa tietoa ollaan käsittelemässä, sitä tiukemmat tietosuojamenetelmät ovat käytössä.

#### 4.7 Pseudonymisointi eli tunnusten korvaaminen satunnaisilla tunnisteilla

Tilastokeskuksessa on toteutettu palvelu, jonka mahdollistaa tilastotuotannossa käsiteltävien henkilötietoja sisältävien tietojen pseudonymisoinnin yhtenäisellä tavalla. Henkilötietojen pseudonymisointi on keino estää henkilöiden suora tunnistaminen. Pseudonymisointi tarkoittaa suorien tunnistetietojen poistamista tai niiden korvaamista aineistoissa muilla tunnisteilla. Poistetut tai korvatut tiedot säilytetään hyvin suojattuna erillään muusta aineistosta.

Kaikki Tilastokeskukseen saapuva henkilötieto pseudonymisoidaan jatkossa heti aineistojen vastaanoton ensimmäisessä vaiheessa riippumatta siitä, onko kyseessä viranomaisten tai yksityisten tahojen toimittamat tiedot vai Tilastokeskuksen suoraan henkilöiltä tai yrityksiltä keräämä tieto.

Pseudonymisointipalvelu on osa uutta aineistojen vastaanottoprosessia, joka otetaan käyttöön vaiheittain vuoden 2018 syksystä alkaen. Tiedot ohjataan pseudonymisointiin vastaanottoprosessin ensimmäisessä vaiheessa. Pseudonymisoinnissa henkilötunnukset korvataan henkilöön kiinnitetyllä satunnaisella tunnukseksi. Tunnus ei sisällä henkilötunnuksen tapaan informaatiota henkilön ominaisuuksista, kuten esimerkiksi sukupuolta tai syntymäaikaa. Henkilötunnuksen korvaamisen lisäksi henkilöiden nimitiedot ja sähköpostiosoite tyhjennetään käsiteltävästä tiedostosta. Pseudonymisoinnin jälkeen aineisto jatkaa vastaanottoprosessin seuraaviin vaiheisiin ja siitä edelleen tilastotuotannon käsittelyvaiheisiin.

Vastaanottoprosessi on automaattinen ja sitä ohjataan metatietojen avulla. Käyttövaltuudet vastaanottoprosessin ensimmäiseen vaiheeseen tietoihin, jossa aineistoissa on vielä henkilöiden suorat tunnistetiedot mukana, ovat hyvin rajatut ja perustuvat prosessissa tapahtuvien mahdollisten virhetilanteiden selvittämiseen.

Pseudonymisoinnissa korvatut ja poistetut henkilötunnus- ja nimitiedot tallennetaan erilliseen suojattuun tietovarantoon. Käyttövaltuudet tietojen palauttamiseen voidaan myöntää vain henkilöille, jotka työtehtäviensä suorittamiseksi tarvitsevat käyttöönsä henkilöiden suoria tunnistetietoja. Tällaisia työtehtäviä ovat esimerkiksi yhteydenotot tiedonantajiin.

Kaikki pseudonymisointipalvelussa tapahtuvat tietojen käsittelyt kirjataan keskitettyyn lokijärjestelmään, josta tapahtumien kulku ja syy käsittelylle voidaan selvittää.

#### 4.8 Sisäinen tarkastus

Sisäinen tarkastus on riippumatonta ja objektiivista arviointi-, varmistus- ja konsultointitoimintaa, jonka tarkoituksena on tukea Tilastokeskusta sen tavoitteiden saavuttamisessa.

Sisäisen tarkastuksen tehtävänä on selvittää johdolle sisäisen valvonnan ja riskienhallinnan asianmukaisuutta ja riittävyyttä sekä suorittaa pääjohtajan vahvistamat tarkastustehtävät. Tarkastuksen kohteena voivat olla Tilastokeskuksen kaikki toiminnot ja tulosyksiköt.

Tilastokeskuksen sisäinen tarkastus tekee vuosittain noin 3-4 tarkastusta sekä muutamia selvityksiä. Aikaisempina vuosina tarkastusten painopiste on ollut hallinnon tarkastamisessa, mutta viime vuosina tarkastuksen painopistettä on siirretty Tilastokeskuksen ydintoimintaan – tilastotuotantoon.

Vuonna 2015 tarkastettiin tilastotuotannon dokumentaatiota. Vuonna 2016 tarkastettiin tilastotuotannon laadunvarmistusta, tilastojen perustamis- ja lakkauttamisprosessien toimivuutta, salassapitositoumuksia sekä tiedonkeruupäätösten ajantasaisuutta. Vuonna 2017 painopiste tarkastustoiminnassa oli EU:n tietosuojasetuksen vaatimuksissa. Tarkastuksen kohteena oli henkilötietojen käsittely tarkastukseen valituissa 5 tilastossa, jotka sisältävät henkilötietoja. Lisäksi tehtiin selvitys ”Henkilötietojen käsittely haastattelutyössä”, jossa tarkastettiin haastattelutyönä tehtävien kotitalous- ja henkilötiedonkeruiden aineistojen keruuta ja käsittelyä.

Oman sisäisen tarkastuksen lisäksi Tilastokeskus hankkii sisäisten tarkastusten asiantuntijapalveluja ostopalveluna. Vuonna 2016 ostopalveluna tarkastettiin yritystilastojen integroidusta tietojärjestelmästä. Tarkastus painottui hallinnollisiin kontroleihin ja tavoitteena oli arvioida tuotantojärjestelmän kontrollien ja tietoturvamekanismien tarkoituksenmukaisuutta ja toimivuutta. Vuonna 2017 ulkoistettuna tarkastuksena tehtiin kehitteillä olevan sosiaalitulastojen integroidun tietojärjestelmän suunnitelmien tietosuojaan tarkastus. Järjestelmän tarkastuksessa tavoitteena oli selvittää tietosuojausta suhteessa tietosuojaasetuksen vaatimuksiin. Tilastokeskuksen henkilötietoaineistojen käsittely ja julkistaminen on tarkoitus siirtää uuteen järjestelmään vaiheittain vuodesta 2018 alkaen.

Kaikissa vuonna 2017 valmistuneissa tarkastuksissa Tilastokeskuksen toimintaa, prosesseja, tietojärjestelmiä ja ohjeistusta on arvioitu suhteessa EU:n tietosuojaasetuksen vaatimuksiin ja on nostettu esille kehittämissuosituksia. Toimintaa on parannettu saatujen kehittämis ehdotusten pohjalta.

Loppuvuodesta 2017 aloitettiin tarkastus ”Käyttöoikeuksien hakeminen ja lakkauttaminen henkilön vaihtaessa tehtäviä”, jossa tavoitteena on varmistua siitä, että henkilöstön käyttöoikeudet ovat ajan tasalla eikä liian laajoja oikeuksia ole käytössä.

#### 4.9 Henkilöstön osaamisen varmistaminen

Tilastokeskuksessa otettiin käyttöön Euroopan tilastoja koskevien käytännesääntöjen mukaiset salassapitositoumukset vuonna 2005. Jokainen uusi työntekijä allekirjoittaa salassapitositoumuksen ensimmäisenä työpäivänään. Allekirjoittamisen yhteydessä uudelle työntekijälle kerrotaan tietosuojasta ja tilastotarkoituksiin kerättyjen tietojen salassapidosta. Salassapitovelvoitteesta on laadittu lyhyt muistio, joka annetaan uudelle työntekijälle. Perehdyttäjät käyvät uuden työntekijän kanssa läpi työtä koskevat tietosuoja- ja -turvamääräykset ja mitä ne käytännössä tarkoittavat.

Tilastokeskuksessa tietosuojaan liittyvää perehdytystä ja koulutusta järjestetään sekä yksikkö- että yleisperehdytyksenä. Yleisperehdytykseen kuuluu Tilastokeskuksessa suunniteltu ja toteutettu tilasto-osaamisen koulutusohjelma. Ohjelmassa käydään läpi ammattieettiset periaatteet ja tilastoalan lainsäädäntö, jossa tietosuojalla on keskinen merkitys. Koko organisaatiolle tarkoitettua koulutusta on järjestetty lainsäännön muutosten yhteydessä

Vuonna 2017 Tilastokeskuksessa järjestettiin uudistuvaan tietosuojalainsäädäntöön liittyvää koulutusta kohdennetusti eri henkilöstöryhmille. Tietosuojalainsäädäntö oli aiheena ylimmän johdon riskienhallintatyöpajassa, esimiesten yhteistilaisuudessa ja koko henkilöstölle eri tulosityksiköiden kokouksissa. Tilastokeskus on osallistunut Vahti-Juhta yhteishankkeen työpajoihin. Lisäksi on osallistuttu erilaisiin uudistuvaa tietosuojalainsäädäntöä käsitteleviin tilaisuuksiin ja seminaareihin.

## 5 Tilastojen laatu, ammattietiikka ja käytäntesäännöt (CoP)

### 5.1 Tilastojen laatu

Tilastoprosessin kaikissa vaiheissa noudatetaan yhteisesti sovittuja laatukriteereitä. Kriteeristöjä on useita ja ne koskevat tilastotuotannon toimintoja tai tilastotuotteita tai molempia. Tilastokeskuksessa käytössä olevat kriteeristöt ovat keskenään yhteensopivia.

Tilastokeskuksessa laadunhallinnan kokonaiskehikkona käytetään Euroopan laatupalkinnon periaatteita (EFQM-periaatteita). Tilastoalan oma laadunhallinnan kehikko on Euroopan tilastojen käytäntesäännöt (CoP). Kehikot täydentävät toisiaan. Myös Suomen virallisen tilaston (SVT) laatukriteerit ovat yhteensopivat Euroopan tilastojen käytäntesääntöjen kanssa. Laadusta raportoidaan käyttäjille jokaisen tilaston julkistuksen osalta erikseen laatuselosteilla. Laatuselosteet ovat julkisesti saatavilla tilastojen internetsivuilla.

Laadun toteutumista seurataan Tilastokeskuksessa virastotasolla puolivuositain. Mittaristossa seurataan sekä tuloksia että laadunhallintaa. Tilastotuotteiden määrän ja tuottavuuden lisäksi seurataan mm. julkistusten oikea-aikaisuutta, virheettömien julkistusten osuutta, julkaisuviivettä ja vastausosuuksia.

Tilastokeskuksen strategiassa asiakkaat, asiakastarpeet ja vuorovaikutus asiakkaiden kanssa ovat keskeisessä osassa ja tätä seurataan säännöllisesti tutkimuksin ja kyselyin. Näiden kohteena ovat niin yrityskuva, asiakkaiden näkemykset tiedon käytön muutoksista, verkkopalvelut kuin yksittäiset tuotteet ja palvelutkin. Tämän lisäksi seurataan tilastokeskus.fi -palvelun palautelaatikkoon tulleita palautteita, kurssiarviointeja, tiedonantajilta saatua palautetta, maksullisten toimeksiantojen asiakkaiden tyytyväisyyttä ja verkkopalvelujen käyttöä.

### 5.2 Ammattietiikka ja käytäntesäännöt

Tilastoalalla ammattietiikkaa pidetään tärkeänä. Tilastoeettisten periaatteiden noudattaminen on tilastoviranomaisen toiminnan edellytys. Luotettavien tilastojen perustana ovat tilastotoimen perinteiset arvot eli luotettavuus, puolueettomuus, riippumattomuus ja objektiivisuus.

Tilastoalan työntekijöillä on velvollisuuksia yhteiskuntaa, asiakkaita ja rahoittajia kohtaan sekä niitä yksilöitä ja yhteisöjä kohtaan, joita tilastot kuvaavat. Tilastoja laadittaessa on otettava huomioon tilastoihin kohdistuvat käyttötarpeet, tietosuoja ja tiedonantajille aiheutuva tiedonantotaakka.

Euroopan tilastojen käytäntesäännöt (European Statistics Code of Practice, CoP) muodostavat toimintaohjeiston, johon Euroopan tilastojen tuotannon tulee perustua. Käytäntesääntöjen laatimisen lähtökohtana oli tarve laatia minimistandardi, jolla taataan Euroopan tilastojärjestelmien riippumattomuus, rehellisyys ja vastuullisuus. Käytäntesäännöissä määritellään myös Euroopan tilastojärjestelmän toimijoiden roolit. Käytäntesääntöjen asema on sittemmin määritelty Euroopan tilastoja koskevassa Euroopan parlamentin ja neuvoston asetuksessa ja periaatteiden toteutumista seurataan säännöllisesti.

Käytäntesäännöt koskevat tilastoviranomaisten riippumattomuutta ja vastuuvollisuutta sekä prosessien ja julkaistavan tiedon laatua. Periaatteet ovat yhteensopivat YK:n tilastokomission hyväksymien virallisen tilaston periaatteiden



kanssa ja täydentävät niitä. Periaatteet ovat myös yhteensopivat Euroopan laatupalkintoperiaatteiden (EFQM) kanssa.

Euroopan tilastojen käytännesäännöt muodostuvat 15 periaatteesta. Periaatteiden toteutumista arvioidaan indikaattoreiden avulla:

### 5.2.1 Kuva: Euroopan tilastoja koskevat käytännesäännöt

Institutionaalinen ympäristö	Tilastoprosessit	Tuotetut tilastotiedot
1. Ammatillinen riippumattomuus	7. Pätevät menetelmät	11. Relevanssi
2. Tiedonkeruuvaltuudet	8. Asianmukaiset tilastolliset menettelytavat	12. Tarkkuus ja luotettavuus
3. Riittävät voimavarat	9. Kohtuullinen vastausrasite	13. Oikea-aikaisuus ja täsmällisyys
4. Laatuun sitoutuminen	10. Kustannustehokkuus	14. Johdonmukaisuus ja vertailukelpoisuus
5. Tilastosalaisuus (tilastollinen tietosuojaja)		15. Saatavuus ja selkeys
6. Puolueettomuus ja objektiivisyys		

### 5.3 Vertaisarvioinnit

Euroopan tilastoja koskevien käytännesääntöjen toteutumista varmistetaan vertaisarvioinneilla. EU:n tilastotoimessa on tehty kaksi vertaisarviointikierrosta, joista ensimmäinen tehtiin vuosina 2006–2008. Tämä tehtiin tilastovirastoista kootuilla arviointitiimeillä. Vuosina 2013–2015 käytiin toinen arviointikierron, joka toteutettiin käyttäen ulkopuolisia arvioijia. Toisella kierroksella arvioitiin myös EU:n tilastoviraston Eurostatin toimintaa. Raportit on julkaistu Eurostatin kotisivulla. Vertaisarvioinneissa todettuja kehittämistoimia seurataan vuosittain.

## 6 Tiedonhankinta

### 6.1 Tiedonhankinnan keskittäminen

Tilastokeskus keskitti organisaatiomuutoksella suuren osan tiedonhankintaansa erilliseen yksikköön vuonna 2013. Keskittämisellä on pyritty hallinnollisten ja muiden massa-aineistojen käytön tehostamiseen, aineistojen vastaanoton automatisointiin, aineistojen laajaan hyödyntämiseen, uusien aineistojen aktiiviseen etsimiseen sekä kustannussäästöihin. Toimintatapojen muutoksella pyritään tiiviimpään sidosryhmäyhteistyöhön ja parempaan riskien hallintaan sekä vastausrasitteen vähentämiseen.

Aineistojen laadun kannalta oleellista on aineistojen kattavuus ja pysyvyys sekä riskienhallinnan kannalta muutostilanteissa riittävän aikainen tiedonsaanti sisällöllisistä tai kattavuusmuutoksista. Tämän vuoksi Tilastokeskus osallistuu aktiivisesti uusien hallinnollisten rekistereiden kuten tulorekisterin suunnitteluun.

### 6.2 Tiedonkeruun yleiset periaatteet

Tilastokeskuksessa tietojen keräämistä ohjaavat ennen kaikkea tilastolaki ja henkilötietolaki 25.5.2018 saakka (ks. tarkemmin luku 4).

Tilastokeskus selvittää tilastoinnin tarpeet ja mahdolliset tietolähteet huolellisesti ennen tilastoinnin aloittamista ja kerää tietoja suoraan yksittäisiltä tiedonantajilta vain jos se on välttämätöntä. Tilastoja laadittaessa käytetään ensisijaisesti hyväksi muihin tarkoituksiin kerättyjä tietoja kuten viranomaisten hallintotehtäviä hoitaessa kertyneitä tietoja. Pääosa Tilastokeskuksen tilastointia varten saamasta ja käsittelemästä henkilötiedosta on peräisin muiden viranomaisten omassa toiminnassa syntyneistä hallinnollisista aineistoista. Tietoja kerätään yksittäisiltä tiedonantajilta suoraan vain, jos muihin tarkoituksiin kerättyjä tietoja ei ole käytettävissä. Tiedonantajilta pyydetään vain tilastotuotannon kannalta välttämättömät tiedot.

Tietojen antaminen tilastointia varten on vapaaehtoista, jollei tiedonantovelvollisuutta ole laissa säädetty. Yksittäisellä kansalaisella ei ole säädetty tiedonantovelvollisuutta. Sen sijaan yrityksillä ja muilla organisaatioilla kuten hallinnollisten aineistojen toimittajille on säädetty tiedonantovelvollisuus tilastolaissa (14 ja 15 §).

Tilastokeskus pyrkii tietojen keruussa hyvään yhteistyöhön tiedonantajien ja heidän edustajiensa kanssa. Tilastokeskus neuvottelee pääsääntöisesti tiedonantajien edustajien kanssa tarkoitusta varten perustetuissa yhteistyöryhmissä. Suoria neuvotteluja käydään hallinnollisia aineistoja toimittavien viranomaisen sekä muiden aineistoja toimittavien yhteistyötahojen kanssa.

Tilastokeskus voi oman tiedonkeruunsa yhteydessä kerätä tietoja samanaikaisesti myös muille viranomaisille ja yhteisöille näiden toimeksiannosta. Tällaisia toimeksiantoja voidaan ottaa vastaan, jos toimeksiannosta ei aiheudu haittaa Tilastokeskuksen omalle toiminnalle eli silloin, jos kerättävien tietojen käyttötarkoitus on rajattu tilastonlaaitimis- ja tutkimuskäyttöön. Tilastokeskuksella voi myös olla yhteistiedonkeruita muiden viranomaisten kanssa edellyttäen, että tilastotarkoituksiin kerättävien tietojen tietosuojavaatimukset toteutuvat.

#### 6.2.1 Tiedonantajien informointi

Tietojenkeruun käynnistämisestä ja kerättävistä tiedoista ilmoitetaan hyvissä ajoin tiedonantajille ja heidän edustajilleen, jotta tiedonantajat voivat asianmukaisesti varautua antamaan tietoja.

Tiedonkeruun yhteydessä tiedonantajille kerrotaan keskeisimmät vastaamiseen liittyvät asiat kuten: tietojen käyttötarkoitus, tietojen suojaamisjärjestelyt, keskeiset tilaston laatimiseen ja tiedon käsittelemiseen liittyvät menettelytavat, tilaston laadinnassa mahdollisesti käytettävät muut tietolähteet ja muualta kerättävät tiedot, tietojen säilytysajat, tiedonantajien oikeudet, perustuuko tietojen kerääminen tiedonantovelvollisuuteen vai vapaaehtoisuuteen ja muut tiedonantajan kannalta tärkeät tiedot, esimerkiksi tietojen edelleen luovuttaminen. Tiedonantajien informointikäytännöt muuttuvat hieman tarkemmiksi uuden tietosuojaa-asetuksen määrittellessä osin uudelleen tiedonantajien informointia.

Tilastokeskuksen tiedonantajille tarjoama informaatio on kerroksellista. Tämä tarkoittaa, että vastaamisen kannalta oleelliset asiat on kerrottu tiedonantajalle lähetetyssä materiaalissa. Lähetetyssä materiaalissa kerrotaan, että tarkempia lisätietoja voi saada Tilastokeskuksen internetsivuilta tai ottamalla yhteyttä tiedonkeruuta hoitaviin henkilöihin.

## 6.2.2 Tiedonantorasiitteen minimointi

Tilastokeskus pyrkii minimoimaan tietojen keräämisestä sekä tietojen antajalle että Tilastokeskukselle aiheutuvia kustannuksia ja yleisesti vastausrasitetta.

Tilastokeskus seuraa yrityksille kohdistuvaa vastausrasitetta ja pyrkii eri keinoin minimoimaan tiedonantajille aiheutuvaa rasitetta.

Tämä tapahtuu tiedonkeruun huolellisella suunnittelulla, käyttämällä olemassa olevia tietolähteitä, hyödyntämällä tilastollisia ja muita menetelmiä tiedon antamisesta aiheutuvien kustannusten ja työn vähentämiseksi aina silloin, kun se tilastojen luotettavuuden kannalta on mahdollista. Esimerkiksi ns. otannan käyttäminen valikoitaessa henkilöitä tai yrityksiä tiedonkeruuseen pienentää tilastoinnin aiheuttamaan rasitusta tilastoinnin kohdejoukossa. Erityisesti yksittäisen kansalaisen todennäköisyys päästä otokseen on hyvin pieni.

Vastaamista pyritään helpottamaan lomakesuunnittelulla ja tiedonkeruuvälineitä kehittämällä. Lomaketestauksella varmistetaan toteutettavan tiedonkeruun laatua. Lomaketestauksessa joukko tiedonantajia pääsee testaamaan aitoa tiedusteluun vastaamistilannetta ja vaikuttamaan lomakkeiden ymmärrettävyyteen ja tiedonkeruuvälineiden käytettävyyteen.

## 6.2.3 Tiedonkeruumenetelmät

Tiedonkeruuta suunniteltaessa huomioidaan tiedonkeruun yleiset periaatteet arvioitaessa, mikä keruumenetelmä on tilaston tai tutkimuksen tavoitteiden, laadun ja kustannusten sekä kohdejoukon kannalta paras vaihtoehto.

Seuraavassa on lueteltu keskeisimmät tiedonkeruumenetelmät:

**Hallinnolliset aineistot ja rekisterit** ovat Suomessa kattavia. Niiden käyttö tilastojen pohjana vähentää kansalaisten tai yritysten tiedonantorasiitusta. Haasteena on hallinnolliseen tarkoitukseen kerätyn tiedon ajoittain hankala muuntaminen tilastokäyttöön sekä se, että lainsäädännön ja hallintokäytäntöjen muutokset saattavat vähentää tietojen vertailukelpoisuutta. Hallinnollisten aineistojen tietoja täydennetään muilla keruumenetelmillä kerätyillä tiedoilla.

**Puhelinhaastattelussa** kysymykset ovat sellaisia, että ne ovat puhelimesta helposti ymmärrettäviä. Vastaaja voi valita hänelle sopivan ja mieluisan paikan ja ajankohdan haastattelulle. Menetelmä on yleisin henkilöitä ja kotitalouksia koskevien tietojen keräämisessä. Haastattelut ovat kestoltaan suhteellisen lyhyitä. Yrityksille puhelinhaastatteluja tehdään vähän.

**Käyntihaastattelut** eli kasvokkain tehtävät haastattelut tehdään useimmiten vastaajan kotona tai muussa ennalta sovitussa paikassa. Käyntihaastattelut helpottavat sisällöllisesti haastavien tiedon keräämistä. Ne ovat kestoaltaan puhelinhaastatteluja pitempiä. Tätä menetelmää käytetään pääasiassa henkilöihin suunnatussa tiedonkeruussa. Käyntihaastatteluun rinnastuva tiedonkeruun muoto on ns. suora havainnointi, käytännössä tällä tarkoitetaan hyllyhintojen keruuta liikkeistä.

**Postikyselyssä** vastaajille lähetetään täytettäväksi paperilomake. Menetelmää käytetään sekä henkilö- että yritystiedonkeruussa, mutta verkossa tapahtuvan tiedonkeruun lisääntyminen on vähentänyt tämän menetelmän käyttöä. Erityismuotona postikyselystä on päiväkirjamenetelmä, jossa vastaaja pitää kirjaa tutkittavasta ilmiöstä, esimerkiksi ajankäytöstä tai kulutuksesta.

**Vastaaminen verkossa** on muita tiedonkeruumenetelmiä nopeampi ja edullisempi. Se mahdollistaa monipuoliset kysymysasettelut erilaisille vastaajaryhmille. Yritykset vastaavat pääosin vain verkossa olevalla lomakkeella. Myös joissakin henkilötiedonkeruissa voi vastata verkossa.

Tilastokeskus pyrkii kehittämään ja ottamaan käyttöön uusia tiedonkeruun muotoja. Tällaisia ovat mm. automatisoitu tiedonsiirto suoraan vastaajan tietojärjestelmästä ja automatisoitu web-tiedonkeruu (nk. web-scaperit).

### 6.3 Aineistojen käyttö ja aineistojen ajantasaisuus

Hallinnollisten aineistojen ja rekisterien käyttö tilastotuotannossa pienentää sekä tilastoinnin kustannuksia että tiedonantajien vastausrasitetta, mikä on yksi keskeisistä tilastolain tavoitteista

Nopeimmat hallinnolliset aineistot ja rekisterit ovat käytettävissä muutaman viikon tai parin kuukauden kuluttua tilastoitavasta ajankohdasta (ajoneuvorekisteri, työnhakijarekisteri, väestötietojärjestelmä ja verottajan maksuvalvonta-aineisto), kun taas hitaimpien valmistumiseen kuluu jopa vuosi (esimerkiksi verotustietokanta). Ajantasaisuuteen vaikuttaa se, kuinka kauan kuluu aikaa ennen kuin tapahtuma rekisteröidään ja se, kuinka usein tietoja päivitetään rekisteriin. Väestötietoja päivitetään joka päivä, tulotietoja päivitetään tällä hetkellä pääosin kerran vuodessa, mutta verohallinnon tulorekisterin käyttöönoton myötä palkka-, työsuhde- ja etuustietoja saadaan päivittäin. Tilastokeskus ottaa tulorekisterin tietoja vastaan vuoden 2020 alusta alkaen.

Ajantasaisuus vaikuttaa myös tietojen kattavuuteen. Esimerkiksi maksuvalvonta-aineistojen saapuessa ensimmäistä kertaa tilastoviranomaiselle kattavuus on noin 80 prosenttia yritysten liikevaihdon ja 75 prosenttia säännöllisten työnantajien maksaman palkkasumman osalta lopullisesta määrästä. Yksittäisen yrityksen kohdalla on ongelmallista tietää, puuttuuko ilmoitus tiedonantoviipeen vai yrityksen toiminnan loppumisen vuoksi. Tämän vuoksi annetun kuukauden tietoja pyydetään tilastotoimen käyttöön toistuvasti kuuden kuukauden ajan, kunnes ne ovat kattavia. Tämä ei kuitenkaan vastaa ajantasaisuuden haasteeseen, jota pyritään pienentämään tilastollisin menetelmin.

#### 6.3.1 Suurimmat aineistojen toimittajat

Tilastokeskus kerää tiedonantovelvoitteen perusteella suuren määrän tietoja eri viranomaisilta niiden hallinnollisessa toiminnassa kertyneistä rekistereistä. Merkittäviä aineistotoimittajia ovat mm. Verohallinto, Väestörekisterikeskus, Luonnonvarakeskus, Eläketurvakeskus, Finanssivalvonta, Posti, Kansaneläkelaitos, Keva, oikeusministeriö, Oikeusrekisterikeskus, Opetushallitus, Patentti- ja rekisterihallitus, Poliisihallitus, Suomen Pankki, Terveiden ja

hyvinvoinnin laitos, Tulli, työ- ja elinkeinoministeriö, Valtiokonttori ja Viestintävirasto

#### 6.4 Tilastokeskuksen henkilöihin ja kotitalouksiin kohdistuvat tiedonkeruut

Tilastokeskus kerää tietoja suoraan yksityishenkilöiltä ja kotitalouksilta. Suurin tiedusteluista on työvoimatutkimus, jossa tiedonkeruu kohdistuu kuukausittain noin 12 500 henkilöön. Vuoden kaikki kalenteriviikot ovat työvoimatutkimuksen tutkimusviikkoja. Samaa vastaajaa haastatellaan puhelimitse kaikkiaan viisi kertaa puolentoista vuoden aikana. Haastattelu kestää keskimäärin 4–8 minuuttia.

Kuluttajabarometrissa tutkimuksen kohteena on kuukausittain kokonaan vaihtuva 2 350 henkilön satunnaisotos. Samaa otosta käytetään myös suomalaisten matkailu -tutkimuksen tiedonkeruussa.

Vuosittain tietoja kerätään asuntojen ja omakotitalojen korjausrakentamisen tiedustelussa, tulo- ja elinolotutkimuksessa sekä väestön tieto- ja viestintätekniikan käyttö -tutkimuksessa. Lisäksi tehdään epäsäännöllisesti tai harvemmin tehtäviä tutkimuksia, joita ovat mm. aikuiskoulutustutkimus, ajankäyttötutkimus, kulutustutkimus, työolotutkimus ja vapaa-aikatutkimus. Lisäksi toteutetaan kertaluonteisia tiedonkeruista yhteiskunnallisesti tärkeistä ja ajankohtaisista aiheista.

Vuonna 2017 yhdeksässä suorassa tiedonkeruussa kyselyiden kohteena oli noin 250 000 yksityishenkilöä tai kotitaloutta.

## 7 Tilastotuotanto

### 7.1 Tuotettavat tilastot

Tilastokeskus tuottaa tällä hetkellä noin 160 tilastoa. Tilastot julkaistaan Tilastokeskuksen verkkosivuilla. Tilastoista tehdään vuosittain noin 560 uutta julkistusta.

Vuotuinen julkistamiskalenteri julkaistaan edellisen vuoden lopulla Internetissä. Se sisältää kaikki vuoden tilastojulkistukset aikataulutettuna.

Raporttimuotoinen vuosiohjelma ”Tilastokeskuksen vuosiohjelma – tilastot, tuotteet ja palvelut” julkaistaan keväisin verkkosivulla: <http://www.stat.fi/org/tilastokeskus/vuosiohjelma.html>

- Tilastojen yksityiskohtaiset kuvaukset löytyvät osoitteesta: <http://www.tilastokeskus.fi/meta/til/index.html>
- Tilastokeskuksen tiedonkeruista kerrotaan osoitteessa: <http://www.tilastokeskus.fi/keruu/index.html>
- Lisätietoja tilastoissa tapahtuvista muutoksista julkaistaan ajantasaisesti: <http://tilastokeskus.fi/til/muutokset.html>

### 7.2 Henkilötietojen elinkaari

Tilastotarkoituksiin kerättävä tieto käsitellään ennalta suunnitellussa ja määritellyssä tilastotuotantoprosessissa (kuva 7.3.1). Tilastotarkoituksiin kerättävän tiedon elinkaari virastossa määritellään tiedonkeruukohtaisissa tiedonkeruupäätöksissä.

Tilastot ja niiden tuotantoprosessit luokitellaan tavallisesti lähdeaineiston perusteella. Tilastotyyppejä ovat:

- **otostilastot**, joissa perusjoukon ominaisuuksia arvioidaan tilastollisia otanta- ja estimointimenetelmiä käyttäen ja tiedot hankitaan yleensä suoralla tiedonkeruulla
- **rekisteritilastot** ja muut epäsuorasti kerättyä mikroaineistoa hyödyntävät tilastot, joissa lähdeaineiston muuttujat ja kohdejoukko on määritelty muuta kuin tilastointitarkoitusta varten
- **kokoomatilastot**, joissa hyödynnetään valmiita laskelmia tai indikaattoreita kuvaamaan usein hankalasti suoraan mitattavaa ilmiötä
- **edellisten yhdistelmät**. Kasvava määrä tilastoja perustuu erityyppisten tietolähteiden yhdistelyyn.

Tilastokeskus käsittelee henkilötietoja useassa tilastoprosessin vaiheessa ja eri käyttötilanteissa. Henkilötietoon pohjautuva tilastointi edellyttää aina Väestörekisterikeskuksen tietojen hyödyntämistä. Väestörekisterikeskuksesta saatuja väestötietoja käytetään sekä otospohjaisten henkilötiedonkeruiden otospoimintaan että kokonaistilastojen tuotantoon. Henkilöpohjaisissa kokonaistilastoissa käsitellään koko Suomen väestöä koskevaa tietoa. Henkilöpohjaisia otostutkimuksia varten puolestaan poimitaan väestötiedot kerran puolessa vuodessa. Keruun käynnistyessä otosta rikastetaan päivitystiedoilla, joita haetaan Tilastokeskukseen Väestörekisterikeskuksesta viikoittain.

Tilastokeskuksen kerätessä henkilötietoja suoraan yksityishenkilöiltä ja kotitalouksilta otokseen poimittujen henkilöiden ja kotitalouksien henkilötietoja käsitellään myös henkilöiden tavoittelun yhteydessä. Kirjeen lähettämistä varten ja haastattelun järjestämiseksi kohdehenkilöille haetaan osoitetietojen lisäksi

joissain tapauksissa puhelinnumerot tai sähköpostiosoitteet julkisista tietolähteistä.

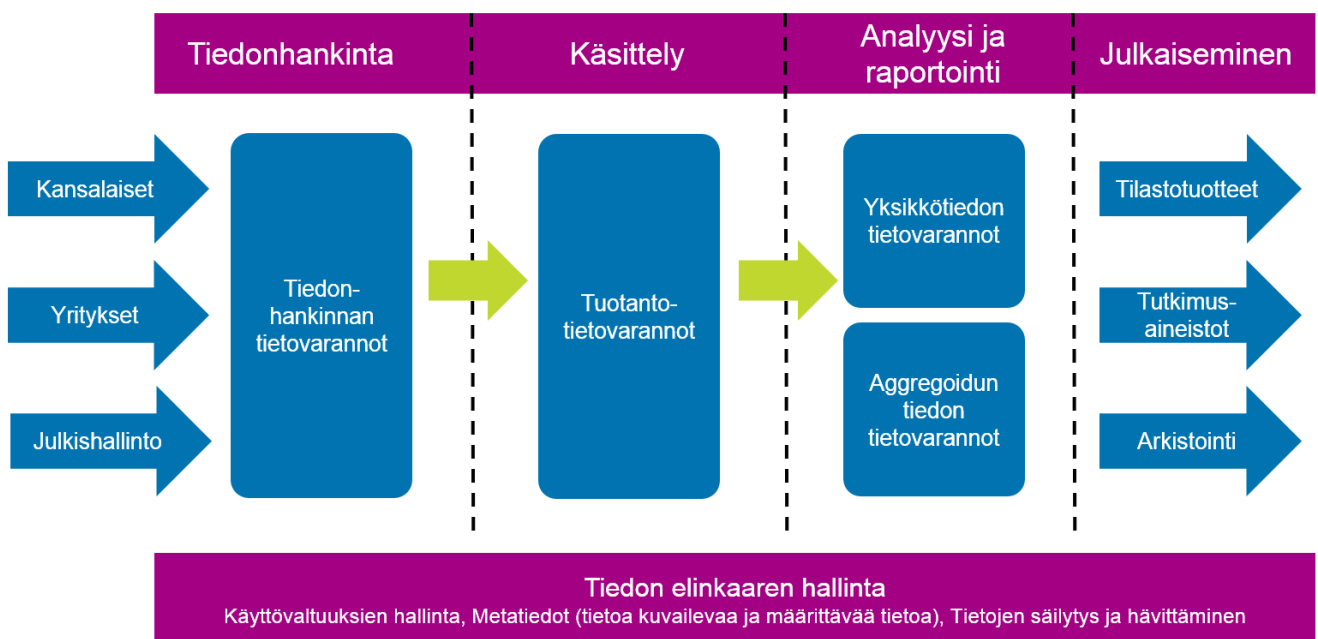
Yritystiedusteluissa Tilastokeskus kerää yrityksiltä henkilötietoja koskien niiden työntekijöitä – tyypillisesti palkkaan ja työsuhteisiin liittyvissä tiedusteluissa.

Tiedon elinkaari Tilastokeskuksessa alkaa, kun Tilastokeskus saa toisen viranomaisen tai muun tiedontuottajan tietoaineiston hallintaansa. Kansalaisilta tai yrityksiltä suoraan kerättävien henkilötietojen elinkaari alkaa sillä hetkellä, kun tiedonantaja vastaa Tilastokeskuksen tiedusteluun. Tiedontoimittajilta ja tiedonantajilta suoraan kerätyt tiedot talletetaan tiedonkeruukohtaisiin tietovarantoihin, joissa tieto on sellaisena kuin se on Tilastokeskukseen saatu. Tietojen käsittely tapahtuu erillisellä tietojen käsittelyalueella, jossa eri lähteistä kerätyt yksittäistä henkilöä tai organisaatiota koskevia tietoja yhdistetään ja käsitellään tilastotiedon tuottamisen vaatimusten mukaisesti. Käsittely tallentuu ns. tuotantotietovarantoihin. Tuotantotietovarantoja ja aiemmin mainittuja raakadatavarastoja säilytetään tilastoprosessien vaatimuksista riippuen eripituisia aikoja.

Tuotantotietovarannon perusteella muodostetaan kunkin ilmiön kuvaamisessa tarvittavat yksikkötason (esimerkiksi henkilö tai yritys) tietovarannot, jotka säilytetään pysyvästi. Pysyvä säilyttäminen tarkoittaa sitä, että tiedot ja tietoa koskevat kuvaukset (aineiston ja sen muuttujien kuvaukset) arkistoidaan siten, että niitä on mahdollista käsitellä uudelleen ilmoitetuissa käyttötarkoituksissa vuosikymmentenkin jälkeen.

Yksikkötietovarannosta muodostetaan myös aggregoituja tietoja sisältävät tilastotietovarannot, joista tuotetaan erilaisia tilastotuotteita kuten julkistuksia ja taulukoita. Tässä vaiheessa käsiteltävät tiedot ovat tilastotietoa, jota voidaan jakaa erilaisina tietotuotteina yhteiskunnan käyttöön. Näihin tilastoinnin tuloksiin voi jokainen tutustua mm. Tilastokeskuksen internetsivuilla.

### 7.2.1 Kuva: tiedonhankinnan tietovarannot



## 7.3 Tilastotuotannon prosessimalli

### 7.3.1 Kuva: tilastotuotannon prosessimalli

Metatiedon hallinta ja laadunhallinta							
1 Tietotarpeen määrittely	2 Tilastoinnin suunnittelu	3 Tuotantojärjestelmän rakentaminen	4 Tiedonhankinta	5 Käsittely	6 Analysointi ja raportointi	7 Julkaiseminen	8 Prosessin arviointi
1.1 Tietotarpeen tunnistaminen	2.1 Tietosisällön suunnittelu	3.1 Keruuvälineiden toteuttaminen	4.1 Otanta	5.1 Lähdeaineistojen yhdistäminen	6.1 Julkistettavien tietojen muodostaminen	7.1 Jakelutietokannan päivittäminen	8.1 Seurantatietojen kokoaminen
1.2 Tietotarpeen varmistaminen	2.2 Muuttujien kuvaaminen	3.2 Käsittely- ja analyysiratkaisujen toteuttaminen	4.2 Tiedonkeruun valmistelu	5.2 Luokittelu ja koodaaminen	6.2 Julkistettavien tietojen tarkistaminen	7.2 Julkaisun valmistelu	8.2 Arviointi
1.3 Tietosisällöstä sopiminen	2.3 Tiedonhankinnan suunnittelu	3.3 Jakeluratkaisujen toteuttaminen	4.3 Tiedonkeruun suorittaminen	5.3 Virheiden tunnistaminen	6.3 Tulosten tulkinta ja raportointi	7.3 Julkaiseminen	8.3 Toimenpiteistä sopiminen
1.4 Käsitteiden tunnistaminen	2.4 Otannan suunnittelu	3.4 Työnkulkujen järjestäminen	4.4 Tiedonkeruun viimeistely	5.4 Editointi ja imputointi	6.4 Tietosuojan varmistaminen	7.4 Tiedon käytön edistäminen	
1.5 Tietolähteiden selvittäminen	2.5 Käsittely- ja analyysimenetelmien suunnittelu	3.5 Tuotantojärjestelmän testaaminen		5.5 Johdettujen tietojen muodostaminen	6.5 Julkistettavien tietojen viimeistely	7.5 Tiedon käytön tukeminen	
1.6 Toteutus-suunnitelman valmistelu	2.6 Työnkulun ja välineiden suunnittelu	3.6 Tilastotuotannon testaaminen		5.6 Painojen laskenta			
		3.7 Tuotantojärjestelmän viimeistely		5.7 Estimointi			
				5.8 Tietojen viimeistely			

Yleisen tilastotuotannon prosessimallin käytön tavoitteena on yhtenäistää tilastojen tuotantoprosesseja. Tilastotuotannon prosessimalli (Generic Statistic Business Process Model, GSBPM) on YK:n Euroopan talouskomission UNECE:n ja eurooppalaisten tilastovirastojen ylläpitämä tilastojen tuotantoprosessin (tilastointiprosessin) kuvaus. Tilastokeskuksessa malli otettiin käyttöön vuonna 2015, minkä jälkeen mallin mukaista jäsenystä on käytetty mm. uusissa työohjeissa, prosessikuvauksissa, uusien tietojärjestelmien käyttöliittymissä, työhakemistorakenteen soveltamisessa ja muissa soveltuviissa kohteissa.

## 7.4 Tilastotietovarannot

Tilastotuotanto perustuu lähes kaikilla aihealueilla tarkasteltavaa ilmiötä eri näkökulmista kuvaavien **aineistojen** yhdistelemiseen. Yhdisteltävän tiedon määrän kasvaessa tiedon hallittavuus, johdonmukaisuus ja ristiriidattomuus on yhä tärkeämpää.



Suurimmat aineistokokonaisuudet on koottu **tietovarannoksi**. Kokoamalla tiedot yhteen estetään tilastotietojen epäjohdonmukaisuutta ja aineistojen päällekkäistä käsittelyä. Pääsääntönä on, että yksi tieto tallennetaan yhteen paikkaan.

Tilastokeskuksen suurimmat tilastotietovarannot ovat:

1. Yritystietovaranto
2. Väestöä ja elinoloja kuvaavien rekisteriaineistojen tietovaranto
3. Tulonjaon tietovaranto
4. Palkkatietovaranto
5. Kansantalouden tilinpidon ja maksutaseen tietovaranto
6. Kuntataloustietovaranto
7. Ympäristö- ja energiatilastojen tietovaranto
8. Hint- ja kustannusindeksien tietovaranto
9. Uudisrakentamisen tietovaranto.

## 7.5 Nykyinen henkilötietovaranto

Nykyiseen henkilötietovarantoon on kerätty henkilöihin liittyvät tilastoinnissa käytettävät rekisteripohjaiset kokonaisaineisot. Vanhimmat tiedot ovat vuoden 1950 väestölaskennasta.

Tietovarasto sisältää mm. seuraavia tietoja:

- väestölaskentatiedot vuodesta 1970 lähtien (lisäksi 10 prosentin otos vuoden 1950 väestölaskennasta)
- väestörakennetiedot, perhe- ja asuntokuntatiedot
- rakennus- ja asuntotiedot
- väestönmuutokset: syntyneet, kuolleet, muutot jne.
- aluetiedot (osa-alueet, taajamat, ruudut ja postinumerot)
- tutkintorekisteri, oppilaitosrekisteri, opiskelija-aineisto
- työssäkäyntitiedot
- kuolemansyytiedot.

Tietovarannon tiedot on tuotettu yhdistelemällä, luokittelemalla ja uudelleen koodaamalla useita hallinnollisia rekistereitä. Lähdeaineistojen käsittelyt tehdään erillisissä tuotantokannoissa ja tiedot siirretään lopullisten aineistojen valmistuttua henkilötietovarantoon. Tietovarannossa henkilön nimet ja henkilötunnukset on korvattu henkilönumerolla (ns. surrogaateilla).

Tietovarannon tiedot ovat pääasiassa poikkileikkaustietoja jostain ajankohdasta, yleensä vuoden lopusta. Tietokantatauluihin on kerätty lisäksi jaksotietoja esimerkiksi työsuhteista, työttömyydestä ja avioliitoista.

Henkilötietovarannosta tuotetaan useita tilastoja mm. työssäkäyntitilastot, väestöä koskevat rakenne- ja muutostilastot sekä asunto- ja asuinoloja kuvaavat tilastot.

Suurin osa tietovarannossa käytetyistä luokituksista perustuu EU-direktiiveillä vahvistettuihin kansainvälisiin standardeihin (mm. toimialaluokitus ja koulutusluokitus). Luokitussuosituksia käyttämällä edistetään eri aikoina ja eri alueilta tuotettujen tilastotietojen vertailtavuutta.

Yksittäisistä tilastoista kertyy dokumentoitavaa ja säilytettävää metatietoa läpi tuotantoprosessin. Metatiedolla tarkoitetaan esimerkiksi käytettyjä määritelmiä,

luokituksia, standardeja, otantakehikkoa, käytettyä lisäinformaatiota, tehtyjä editointeja ja muuttujamuunnoksia. Henkilötietovarannon taulujen aineisto- ja muuttujakuvaukset on tallennettu metatietovarastoon.

Jokaisesta julkaistusta tilastosta on laadittu laatuseloste, jossa on kuvattu tarkemmin kyseisessä tilastossa käytetyt tutkimusmenetelmät eli tilaston perusjoukko, käytetyt perusaineistot, tutkimusasetelma, tiedonkeruutapa, tärkeimmät luokitukset ja estimointimenetelmät. Osalle tilastoista on tehty myös erillinen yksityiskohtainen menetelmäseloste.

Tilastokeskuksessa käytetään tilastojen kokonaislaatua kuvaavina keskeisinä laadunarviointimenettelyinä tilastojen läpivalaisua ja tilastotuotantoon räätälöityä itsearviointia.

Henkilötietovarannon tietoja hyödynnetään kokonaisaineistoista julkaistavien henkilötilastojen lisäksi useissa otoksiin perustuvissa tilastoissa kuten mm. työvoimatutkimuksessa, tulonjakotilastossa, kulutustutkimuksessa ja asumisen energiankulutustilastossa. Lisäksi henkilötietovarannosta tehdään suoraan tutkijoille tutkimusaineistoja ja muita asiakaskohtaisia erityisselvityksiä.

## 7.6 Nykyinen palkkatietovaranto

Nykyinen palkkatietovaranto sisältää tilastoinnissa ja tutkimuskäytössä käytettävät rekisteripohjaiset sekä Tilastokeskuksen omilla tiedonkeruilla kerätyt henkilön palkkaa ja työsuhdetta koskevat tiedot. Vanhimmat tiedot ovat 1970-luvulta ja uusimmat tiedot ovat tällä hetkellä vuodelta 2017. Tietovarannon tiedot kuvaavat korotettuna 1,6 miljoonan palkansaajan ansiotietoja.

Tietovaranto sisältää:

- Yksityisen sektorin kuukausi- ja tuntipalkkatiedot
- Kuntasektorin kuukausi- ja tuntipalkkatiedot
- Valtion palkkatiedot
- Palkkarakennetilaston tiedot

Tietovarannon tiedot on tuotettu yhdistelemällä, luokittelemalla ja uudelleen koodaamalla käyttäen apuna tilastoalueen omia ja Tilastokeskuksen yleisiä luokituksia sekä yhteisten tietovarantojen tietoja mm. YTY-tietovarannosta.

Lähdeaineistojen käsittelyt tehdään erillisessä tuotantokannassa ja tiedot siirretään aineistojen valmistuttua tietovarastoon.

Tietovarannon tiedot ovat poikkileikkaustietoja vuoden viimeisen neljänneksen ajankohdalta. Yksittäisistä tilastoprosesseista kertyy metatietoa läpi tuotantoprosessin mm. käytettyjä määritelmiä, luokituksia, käytettyä lisäinformaatiota, tehtyjä editointeja. Palkkatietovaraston aineistokuvaukset sekä tietopalvelutietokantojen kuvaukset on tallennettu metatietovarastoon.

Palkkatietovarastosta tuotetaan seuraavat tilastot:

- yksityisen sektorin kuukausipalkkatilasto
- yksityisen sektorin tuntipalkkatilasto
- kuntasektorin kuukausi- ja tuntipalkkatilasto (yhteinen)

- valtion palkkatilasto
- palkkarakennetilasto (EU-tilasto)

Jokaisesta julkaistusta tilastosta on laadittu laatuseloste, jossa on kuvattu tarkemmin kyseisessä tilastossa käytetyt tutkimusmenetelmät eli tilaston perusjoukko, käytetyt perusaineistot, tutkimusasetelma, tiedonkeruutapa, tärkeimmät luokitukset ja estimointimenetelmät.

Palkkatietovarannon tietoja hyödynnetään julkaistavien palkka- ja palkkarakennetilastojen lisäksi ansiotasoindeksin tuotannossa. Palkkatiedoista tehdään vakiotietopalveluja ja erillisselvityksiä. Isoimpia palkkatietojen hyödyntäjiä ovat työmarkkinajärjestöt, viranomaiset, yritykset ja media. Lisäksi palkkarakennetiedot toimitetaan Tilastokeskuksen tutkijapalveluihin, jossa tiedoista muodostetaan yhtenäinen harmonisoitu palkkarakennetilasto tutkijoiden käyttöön (aikasarja vuodesta 1995).

## 7.7 Tieliikenneonnettomuudet

Tieliikenneonnettomuustilastot sisältävät tietoa henkilövahinkoon johtaneista tieliikenneonnettomuuksista ja niiden osallisista. Tietosisältö kattaa poliisin tietoon tulleet ja Tilastokeskukselle ilmoitetut henkilövahinkoon johtaneet tieliikenteen onnettomuudet. Tiedot kerätään poliisin tietojärjestelmään kirjatuista ja Tilastokeskukselle ilmoittamista henkilövahinkoon johtaneista tieliikenteen onnettomuuksista. Poliisin tietoja täydennetään Liikenneviraston ja Liikenteen turvallisuusviraston tiedoilla sekä Tilastokeskuksen kuolemansyytilaston tiedoilla.

Aineisto sisältää kaikki kuolemaan johtaneet onnettomuudet ja suuren osan loukkaantumisonnettomuuksista. Tietoaineisto sisältää onnettomuuksiin osallisiin henkilöihin liittyviä tietoja kuten tienkäyttäjärühmä, ikä, sukupuoli, kansalaisuus, puhalluskokeen tulos, päihteiden käyttö ja tiedot ajokortista. Lisäksi aineisto sisältää onnettomuutta koskevia tietoja (vakavuus, onnettomuustyyppi, ajankohtaan liittyvät tiedot, osallisten määrät), onnettomuuspaikkaan liittyviä tietoja (maakunta, kunta, paikka tieverkolla, nopeusrajoitus, risteys) ja onnettomuuden olosuhteisiin liittyviä tietoja (keli, sää, valoisuus).

## 7.7.1 Kuva: Tietojärjestelmät



Tilastokeskuksen tietojärjestelmät jakaantuvat kolmeen pääryhmään. Tilastotuotantojärjestelmät käsittävät tuotannon neljä eri päävaihetta. Jako noudattaa yleisen tilastotuotannon prosessimallia (GSBPM).

Tilastotuotantoa tukevat järjestelmät puolestaan vaikuttavat läpi koko tilastotuotantoprosessin. Nämä ovat hyvin toimialasidonnaisia ja näistä järjestelmistä suurin osa on Tilastokeskuksen itsensä tuottamia.

## 8 Tilastotiedon jakelu

### 8.1 Tilastojen julkistamisperiaatteet Tilastokeskuksessa

Tilastojen julkaisuajankohdat vahvistetaan syksyisin toiminnan suunnittelun yhteydessä. Seuraavan vuoden julkistamiskalenteri julkaistaan käyttäjille joulukuussa.

Tilastokeskuksen tilastojulkistukset, tilastotietokantapäivitykset sekä tiedotteet julkistetaan kaikille samanaikaisesti arkipäivisin kello 9.00 internetissä. Tiedot ovat julkisia sen jälkeen, kun ne ovat päivittyneet internetsivuille. Poikkeuksen muodostavat mm. tiedotustilaisuuksissa julkistettavat tilastot ja tiedotteet, joiden julkistamisajankohdasta päätetään erikseen. Ennen julkistamista tilaston tietoja saa Tilastokeskuksessa käsitellä vain henkilö, joka osallistuu ko. tilaston laadintaan.

- [Julkistamiskalenteri](#)

### 8.2 Avoin data

Kansainvälisesti vertailtuna Suomen tilastotoimi tuottaa laadukkaasti avointa dataa. [ODIN \(Open data inventory\) -mittari](#), joka erityisesti painottaa eri maiden tilastovirastojen verkkosivuilta avoimena datana saatavissa olevaa aineistoa, luokittelee Suomen tilastotoimen maailmanlaajuisesti parhaimmiston joukkoon.

Asiakkaiden avoimeen dataan liittyvät tarpeet ja odotukset ovat kasvaneet. Avoimen datan palautetta on saatu myös avoimen datan yhteisöjen kautta. Eri asiakastutkimusten ja palautteiden mukaan Tilastokeskuksen avoimeen dataan liittyy edelleen kehittämistarpeita muun muassa metatiedon laatuun ja datan käytettävyyteen sekä datan löydettävyyteen ja näkyvyyteen.

Tilastokeskus on panostanut avoimen datan saatavuuden lisäämiseen ([www.stat.fi/avoindata](http://www.stat.fi/avoindata)). Asiakkaiden tietotarpeiden kartoitus maakunta- ja SOTE-uudistuksen näkökulmasta aloitetaan. Kartoitus on osa laajempaa hankekokonaisuutta, jossa tavoitteena on tulevaisuudessa lisätä alueellisen tiedon tarjontaa avoimena datana.

Tiedon avaamisen lisäksi Tilastokeskus on osallistunut avoimia aineistoja hyödyntäviin hackathon-tapahtumiin ja järjestänyt avoimen datan asiakaskoulutuksia.

## 9 Tietojen luovuttaminen

### 9.1 Tietojen luovutus tilastotarkoituksiin

Tilastotarkoituksiin kerättyjen tietojen luovuttamisesta ja luovuttamiseen liittyvistä rajoitteista säädetään Euroopan tilastoista annetussa Euroopan parlamentin ja neuvoston asetuksessa ja tilastolaissa. Tilastotarkoituksiin kerättyjä salassapidettäviä tietoja voidaan luovuttaa tilastointia varten toiselle tilastoviranomaiselle, Euroopan unionin tietotarpeisiin tilastoja tuottavalle viranomaiselle ja Suomen Pankille.

Luovutusta ja tietojen vaihtoa koskevia sopimuksia on tehty Luonnonvarakeskuksen, Terveiden ja hyvinvoinnin laitoksen, Ahvenanmaan tilastotoimiston ÅSUB:n ja Suomen Pankin kanssa.

Lisäksi Tilastokeskus luovuttaa Euroopan unionin tilastovirastolle Eurostatille tilastointia koskevien Euroopan parlamentin ja neuvoston asetusten edellyttämiä anonymisoituja tai aggregoituja tietoja. Tilastomuotoista aineistoa luovutetaan kansainvälisten sopimusten velvoittamana mm. YK:lle ja OECD:lle.

## 9.2 Tietojen luovutus tutkimuskäyttöön

Tilastotarkoituksiin kerättyihin tietoihin voidaan antaa käyttöoikeus eli niitä voidaan käsitellä valvotussa etäkäyttöympäristössä, jos aineistoista voidaan välillisesti tunnistaa yksittäinen henkilö käsittelyn aikana. Valmiissa taulukoissa tunnistaminen ei enää tarkastusprosessien jälkeen ole mahdollista. Anonymisoitu tutkimusaineisto, jossa välillinenkin tunnistaminen ei ole mahdollista, voidaan luovuttaa tutkijalle.

Tilastokeskuksessa toimii pääjohtajan asettama Tilastoeettinen lautakunta. Sen tehtävänä on antaa lausuntoja mikroaineistojen luovuttamisesta tutkimuskäyttöön sekä muista tietosuoja- ja tilastoeettisistä kysymyksistä, jotka liittyvät esimerkiksi tiedonkeruuseen, käsittelyyn tai julkaisemiseen. Tilastoeettisen lautakunnan käsittelyn tavoitteena on varmistaa Tilastokeskuksessa yhtenäinen menettely periaatteellisissa lainsäädännön ja tilastoeettisten periaatteiden soveltamiskysymyksissä kokoamalla tilastollinen ja juridinen asiantuntemus päätöksenteon tueksi.

### 9.2.1 Henkilöaineistojen käsittely edellyttää Tilastokeskuksen käyttö lupaa

Tilastokeskuksen yksikkötason aineistojen käyttö edellyttää käyttö lupaa. Tietoja saa käyttää ainoastaan käyttö luvan saanut henkilö ja vain päätöksessä hyväksytyyn tarkoitukseen. Aineistoja käsittelevät henkilöt allekirjoittavat salassapitositoumuksen.

Tilastoinnin kohteiden ts. tilastoaineistoissa esiintyvien havaintojen kannalta tietosuojasta huolehtiminen on tärkeää. Yksittäistä henkilöä, kotitaloutta, yritystä tai muuta tilastoyksikköä koskevien tietojen paljastuminen ulkopuolisille on estettävä Tilastokeskuksen tutkimuskäyttöön antamien aineistojen pohjalta julkaistuista tiedoista. Käyttö lupien ehtoihin kuuluu, ettei tietojen kohteita saa yrittää tunnistaa aineistosta. Aineistojen käyttäjiä on laadittu ohjeita aineiston käyttöön liittyvästä tietosuojasta. Tutkijoita varten on myös tehty Tutkimusaineistot etäkäytössä -verkkokurssi.

Käyttö lupa myönnetään määräajaksi. Tilastokeskus lähettää tutkimushankkeille muistutuksen käyttö luvan päättymisestä. Käyttäjälle luovutettu aineisto on palautettava tai hävitettävä Tilastokeskukselle käyttö lupapäätöksessä määrättyssä ajassa hankkeen päättyessä. Samoin on hävitettävä aineistosta mahdollisesti otetut kopiot ja hävittämistä on ilmoitettava Tilastokeskukselle.

Etäkäyttöjärjestelmässä tutkijoiden kirjautumistunnukset eivät enää toimi käyttöajan päätettyä. Käyttö luvan voimassaololle voi hakea jatkoa.

Käyttö luvan ehdoksi voidaan asettaa, että aineisto on käytettävissä vain Tilastokeskuksen tiloissa tai etäkäyttöjärjestelmässä (esim. kokonaisaineistot). Käyttö luvan saajalta voidaan myös edellyttää selvityksen tai tutkimuksen esittämistä Tilastokeskukselle tietosuojatarkistusta varten ennen sen julkistamista. Julkaistut lopputuotteet tai niiden viitetiedot tulee myöskin toimittaa Tilastokeskukselle.

Tilastokeskus poistaa kaikista tutkimuskäyttöön luovutettavista aineistoista henkilötunnukset ja korvaa ne pseudotunnisteilla. Poikkeuksena Tilastokeskus voi luovuttaa tutkimustarkoituksiin tietoja iästä, sukupuolesta, ammatista, sosioekonomisesta asemasta ja koulutuksesta myös tunnistetietoineen, mikäli tutkijalla on oikeus käsitellä henkilötietoja. Erityislainsäädännön perusteella myös kuolinsyytietoja voidaan luovuttaa tunnistetietoineen tutkimustarkoituksiin.

Aineistojen yhdistäminen tutkijan omiin tai muiden organisaatioiden luovuttamiin aineistoihin on myös mahdollista. Tällöin tunnisteita (henkilötunnus, nimi) sisältävät aineistot toimitetaan Tilastokeskuksen tutkijapalveluihin, jossa suorat tunnistet poistetaan ja korvataan hankekohtaisilla pseudotunnuksilla ennen aineistojen siirtämistä etäkäyttöjärjestelmään.

#### 9.2.2 Taulukko: Vuonna 2017 myönnettyjen käyttöluvien ja aineistoluovutusten määrät tuoteluokittain

Tuoteluokka	Myönnetyt käyttöluvat	Aineisto-toimitus
Kuolemansyyaineistot	82	84
Jatkoluvat, lupien laajennukset	141	
Mikrosimulointi	21	14
Valmisaineistot/etäkäyttö	49	40
Räätälöidyt / etäkäyttö	84	76
Räätälöidyt / luovutus	29	19
Tunnisteelliset	26	21
Muut	5	6
Yhtensä	437	260

#### 9.2.3 Sisu-mikrosimulointimalli

Tilastokeskus kehittää ja ylläpitää Suomen henkilövero- ja sosiaaliturvajärjestelmää mallintavaa SISU (Simuloitu Suomi) -mikrosimulointimallia. SISU-malli on henkilöverotuksen ja sosiaaliturvalainsäädännön suunnitteluun, seurantaan ja arvioimiseen tarkoitettu laskentaväline. SISU-mallissa käytettävää laajaa ja väestöä hyvin edustavaa rekisteripohjaista mikroaineistoa (n. 800 000 henkilön otosaineisto) voidaan käyttää vain etäympäristön kautta.

Vuonna 2017 SISU-mallilla oli 75 käyttäjää. Käyttäjät ovat pääsääntöisesti ministeriöistä, tutkimuslaitoksista sekä etujärjestöistä.

#### 9.2.4 Etäkäyttöjärjestelmä

Tilastokeskuksella on tietojen etäkäyttöjärjestelmä (FIONA–Finnish Online Access). Etäkäyttöpalvelun kautta tutkija pääsee käyttämään mikroaineistoja omalta työpisteeltään valvotussa ja tietoturvallisessa ympäristössä. Etäkäyttöjärjestelmä käyttää vahvaa tunnistautumista. Tutkija kirjautuu etäkäyttöympäristöön joko toimeksiantojen tai mikrosimuloinnin palvelimeen.

Etätyöpöydällä työskentely vastaa omalla koneella työskentelyä, mutta tutkimusaineisto siirtyy käyttäjälle ainoastaan työpöytänäkymänä. Tutkimusaineistoissa yksikötason suorat tunnistet, kuten henkilötunnukset tai yritystunnukset, on etäkäyttöpalvelussa korvattu pseudotunnisteilla. Pseudotunniste on suoran yksikötunnisteen korvaava keinotekoinen tunniste, jolla yksikköä (tai yksilöä) voidaan seurata aineistossa ja yksikön tietoja linkittää

eri aineistojen välillä. Kun tutkija haluaa siirtää järjestelmästä ulos tiedostoja ja tuloksia, Tilastokeskus tarkistaa niiden tietosuojan ennen siirtoa.

Etäkäyttöpalvelun teknisestä ylläpidosta vastaa Opetus- ja kulttuuriministeriön hallinnoima Tieteen tietotekniikan keskus – CSC -yhdessä Valtorin kanssa. Järjestelmän tietoturvasta ja CSC:n palvelimella säilytettävien tutkimusaineistojen tietosuojasta vastaavat Tilastokeskus ja CSC organisaatioiden välillä tehtyjen sopimusten mukaisesti.

Tutkijan käytössä on analyysiohjelmiä (SPSS, STATA ym.) ja kullakin hankkeella on oma työskentelytila tutkimusaineistojen, analyysitulosten ja koodien säilyttämiseksi. Mikrosimuloinnissa tutkijalla on käytössä SISU-mikrosimulointimallin käyttöön tarkoitettu SAS EG -ohjelma. Järjestelmän käyttö on maksullista ja siitä kerätään lokitietoja laskutusta, järjestelmän ylläpitoa ja valvontaa varten.

Etäkäyttö edellyttää, että tutkijan oma organisaatio on laatinut sopimuksen etäkäyttöyhteyden avaamisesta ja ylläpidosta ja on toimittanut Tilastokeskukselle selvityksen etäyhteyteen liittyvistä tietosuoja- ja tietoturvakäytännöistään. Sopimuksessa määritellään käytön ehdot sekä organisaation hallinnollinen ja tekninen yhteyshenkilö.

Etäkäyttäjien määrä kasvaa koko ajan. Tilastokeskus kehittää CSC:n kanssa mittareita, joilla käyttöä voi seurata. Vuoden 2017 lopussa 71 organisaatiolla oli sopimus etäkäytöstä Tilastokeskuksen kanssa. Hankkeita oli yhteensä 163 ja käyttäjiä 363.

### 9.3 Yritys- ja toimipaikkarekisteri

Tilastokeskus ylläpitää yritys- ja toimipaikkarekisteriä, joka sisältää kaikki yritykset, yhteisöt ja yksityiset elinkeinonharjoittajat, jotka ovat liiketoiminnasta arvonlisäverovelvollisia tai toimivat työnantajina. Rekisterissä oli vuoden 2017 lopussa noin 735 000 luonnollista henkilöä, joista edelleen toimintaa harjoittaa noin 287 000. Luonnollinen henkilö harjoittaa yritystoimintaa omalla nimellään tai rekisteröidyllä toiminimellä. Tähän ryhmään kuuluvat ammatinharjoittajat, useimmat maatalouseläinten harjoittajat sekä liikkeenharjoittajat.

Yritysrekisterin pääasiallinen käyttötarkoitus on toimia tilastollisena perusrekisterinä, jonka tietoja käytetään tiedonkeruissa ja tilastojen laatimisessa perustietojen lähteenä, otoskehikkona sekä yritystiedonantajarekisterin perustana. Lisäksi rekisteriä käytetään tietopalvelurekisterinä ja sen tilastolain 18 §:ssä määriteltyjä julkisia tietoja toimitetaan asiakkaille. Yritysrekisterin tietoja ovat mm. yritysten osoitteet, toimialat, henkilöstön ja liikevaihdon suuruusluokat, aloittamisaika sekä tuoja/viejä-tieto. Tietoja saa niin yritys- kuin toimipaikkatasolla.

Yritysrekisterin tietolähteitä ovat useat eri hallinnolliset aineistot sekä omat tiedustelut suoraan yrityksille. Tärkeimmät aineistot ovat Verohallinnon asiakastietokanta ja muut Verohallinnon tietoaineistot. Lisäksi hyödynnetään Patentti- ja rekisterihallituksen ja Verohallinnon yhteistä yritystietojärjestelmää (YTJ), Tullin ulkomaankauppa-aineistoa, Patentti- ja rekisterihallituksen kaupparekisteriä, Väestörekisterikeskuksen rakennus- ja huoneistorekisteriä, Posti Oy:n osoiteaineistoa sekä Valtiokonttorin ja Kevan työsuhteaineistoja.



Tilastokeskuksen omissa tiedusteluissa selvitetään mm. yritysten toimiala, henkilöstön määrä ja toimipaikkarakenteeseen liittyvät tiedot. Tiedusteluja lähetetään vuosittain valtaosalle monitoimipaikkaisista ja yli 20 hengen yksitoimipaikkaisista yrityksistä. Muille yrityksille lähetetään tiedusteluja harvemmin rotaatioon perustuen. Uudet yritykset tiedustellaan melko pian toiminnan aloittamisen jälkeen.

#### 9.4 Yritystiedonantajarekisteri

Yritystiedonantajarekisteriin kerätään tiedot siitä, mihin eri Tilastokeskuksen tiedonkeruisiin yritystiedonantaja kuuluu ja mitkä ovat tiedonkeruiden ajankohdat. Yritystiedonantajarekisterissä ovat nähtävissä kaikki suorat yritystiedonkeruut ja kaikki yksittäiseen yritykseen kohdistuvat tiedonkeruut. Tietoja rekisteriin on kerätty vuodesta 2006 alkaen.

Yritystiedonantajarekisteri sisältää tiedonantajista seuraavat tiedot: y-tunnus, tiedonantajatunnus, tiedonantajan nimi, yhteyshenkilön nimi, vastautustieto ja vastaustapa. Tietoja käytetään mm tiedonantajien informoisissa (tiedonkeruupalvelu yrityksille) ja Tilastokeskuksen sisällä esim. selvitettäessä keitä ovat suurtiedonantajat ja miten tiedonantorasite kohdentuu yrityksille.

Yritystiedonantajarekisteristä luovutetaan tietoja Tilastokeskus.fi/tiedonkeruut - sivuilla sijaitsevan Tiedonkeruupalvelu yrityksille -sovelluksen kautta ainoastaan yritystiedonantajalle itselle. Tietoja voidaan luovuttaa myös konsernin emoyhtiön prokuristille, mikäli emoyhtiö haluaa nähdä tytäryhtiöidensä tiedonkeruutietoja. Yritystiedonantajalle on tämän palvelun käyttöä varten erikseen toimitettava käyttäjätunnukset.

## 10 Toimeksiantoina tehtävä henkilötietojen käsittely

### 10.1 Asiakasrahoitteiset haastattelu- ja tiedonkeruupalvelut

Tilastokeskus tarjoaa maksullisena toimintana haastattelu- ja tiedonkeruupalveluita. Ulkoiset asiakkaat ovat esimerkiksi eri tutkimuslaitoksia, ministeriöitä, yliopistoja ja tutkijaryhmiä. Tiedonkeruun kohde voi keruusta riippuen olla henkilö, kotitalous, yrityksen toimipiste tai kunta. Tietoja kerätään keruusta riippuen verkossa, puhelimitse, käyntihaastatteluna tai postitse, mutta yleisimmin edellä mainittujen tapojen erilaisilla yhdistelmillä.

Keruiden aiheet ja kokoluokka vaihtelevat paljon. Tiedonkeruut vaihtelevat esimerkiksi kymmenille tuhansille vastaajille lähetettävästä verkko-postikyselystä 3 000 henkilön käyntihaastattelututkimukseen ja 5 000 henkilön verkko-puhelinkeruusta kaikkia Suomen kuntia ja kuntayhtymiä koskevaan verkkokyselyyn.

Yksittäisen tutkimuksen toteuttamiseen liittyy monia työvaiheita keruun suunnittelusta aineiston käsittelyyn, toimittamiseen ja sopimuksesta riippuen myös raportointiin. Keruun aiheesta riippumatta tehtäviin kuuluu useimmiten kysymysten suunnittelu yhdessä asiakkaan kanssa, kyselylomakkeiden tuottaminen eri keruuvälineisiin, vastaajamateriaalien (saate- ja muistutusviestit, esitteet ja muut tavoittelumateriaalit sekä tiedonkeruusivujen sisältö) suunnittelu ja tekeminen, vastausten kertymisen seuranta ja ohjaaminen, aineiston siistiminen ja perusanalysointi sekä tutkimusselosteen tekeminen. Seloste, peruseräraportti ja aineisto toimitetaan keruun päätteeksi asiakkaalle.

Asiakasrahoitteisten tutkimusten arkistointi hoidetaan niin, että kaikesta materiaalista säilytetään vain jälkikäsitellyn kannalta välttämättömät henkilötiedot sekä lopullinen saatemateriaali, lomake ja data.

Arkistointiin vaikuttaa se, mitä sopimukseen on asiakkaan kanssa kirjattu. Joissain tapauksissa voidaan sopia aineiston tai muun materiaalin laajemmasta tai pitkäaikaisemmasta säilyttämisestä esim. saman tiedonkeruun seuraavaa kierrosta tai paneeliasetelmaa varten.

#### 10.1.1 Taulukko. Asiakasrahoitteiset tiedonkeruut

Vuosi	Toimeksiantojen lukumäärä
2010	23
2011	15
2012	11
2013	10
2014	7
2015	14
2016	11
2017	12

## 10.2 Asiakaskohtaiset tietopalvelusopimukset

Tilastokeskus tekee maksullisena palvelutoimintana erityisselvityksiä asiakkaan tarpeiden mukaan. Erityisselvityksissä käytetään Tilastokeskuksen tilastotarkoituksiin keräämiä aineistoja kuten esim. väestö-, työssäkäynti-, asuntokunta-, asunto-, ja rakennusaineistoja. Joskus aineistoihin yhdistetään asiakkaan aineistoja tai tietoja toimitetaan asiakkaan toimittamaan aluejakoon perustuen. Tiedot raportoidaan yksittäisenä taulukkona, taulukkopaketina, tilastollisena selvityksenä, graafisina kuvioina tai teemakarttoina

## 10.3 Tilastointipalvelut

Tilastokeskus tuottaa tilastoja tilauksesta myös muille organisaatioille. Tilastointipalvelu voi sisältää koko tilaston tuottamiseen liittyvän prosessin tiedonkeruusta jakeluun tai vain jonkin osatehtävän prosessista. Palveluun sisältyvä tiedonkeruu voidaan toteuttaa joko haastattelemalla, sähköisenä kyselynä tai yhdistelemällä valmiita rekisteriaineistoja. Tilastointipalveluita tehtäessä noudatetaan aina samoja tilastoeettisiä työtapoja ja tietosuojakäytänteitä kuin muussakin tilastotuotannossa.

Tilastointipalvelu suunnitellaan aina tiiviissä yhteistyössä asiakkaan kanssa ja palvelun toteutuksesta sovitaan yksityiskohtaisesti sopimuksella.

Tilastointipalvelun osana voidaan julkaista myös valmiita kuvioita, taulukoita sekä analyysitekstejä. Tarvittaessa asiakkaalle toimitetaan tilastoaineisto myös tunnistamattomassa muodossa tutkimuskäyttöön.

Tällä hetkellä Tilastokeskuksessa tehdään noin 20 erilaista tilastointipalvelua. Asiakkaina on mm. Verohallinto ja Liikenteen turvallisuusvirasto Trafi. Palvelutarjonta on laajentunut melko voimakkaasti viime vuosina. Laajentuminen perustuu pitkälti hyviin asiakassuhteisiin ja hyvään julkisuuskuvaan.

## 10.4 Kansallinen etäkäyttöpalvelu

Tilastokeskus on sopinut neljän viranomaisen ja 23 korkeakoulun kanssa siitä, että ne käyttävät FIONA-etäkäyttöjärjestelmää tutkimusaineistojen luovutuskanavana. Tällöin rekisteripitäjän muodostama ns. valmisaineisto on toimitettu Tilastokeskuksen tutkijapalveluihin, jossa aineistoon on viety pseudotunniste ja siirretty FIONA-etäkäyttöjärjestelmään. Tilastokeskuksen tutkijapalvelut avaa rekisteripitäjän myöntämien käyttöluupien mukaisesti etäkäyttöyhteyden valmisaineistomoduuleihin käyttöluuvissa nimetyille tutkijoille.

## 11 Asiakas- ja sidosryhmätiedot

Tilastokeskuksen lopputuotteita käyttäviä asiakkaita ja sidosryhmiä ovat esimerkiksi valtion virastot ja laitokset, kunnat, yritykset ja yhteisöt, yliopistot ja oppilaitokset, tutkijat sekä media. Asiakkuus perustuu aina sopimukseen, suostumukseen tai asemaan ja tiedot kerätään pääsääntöisesti asiakkaalta tai yhteyshenkilöltä itseltään. Asiakas- ja sidosryhmätiedoille on määritelty säilytysajat arkistonmuodostussuunnitelmassa. Säilytysajat vaihtelevat tilanteen mukaan. Esimerkiksi Tilastokeskuksen uutisviestejä tilanneiden tiedot poistetaan kun asiakas lopettaa tilauksen. Sopimukseen liittyviä tietoja puolestaan säilytetään kymmenen vuotta sopimuksen voimassaolon jälkeen, jonka jälkeen tiedot poistetaan.

Asiakas- ja sidosryhmätietojen hallintaa ja käyttöä Tilastokeskuksessa on kehitetty viime vuosina. Vuoden 2016 lopussa otettiin käyttöön uusi asiakas- ja sidosryhmähallintajärjestelmä, asiakas- ja sidosryhmätiedot päivitettiin perin pohjin ja vanhentuneet tiedot poistettiin asianmukaisesti. Uuden järjestelmän avulla asiakas- ja sidosryhmätietojen oikeellisuuden tarkistaminen on helpottunut ja on nyt systemaattinen osa asiakas- ja sidosryhmyötä.

Vuonna 2017 asiakas- ja sidosryhmähallintajärjestelmään integroitiin asiakas- ja sidosryhmäviestinnän uusia välineitä sekä sähköpostiviestinnän että sosiaalisen median kanavien tueksi. Meneillään on myös Tilastokeskuksen tilastojulkistusten ja erilaisten uutisviestien tilausjärjestelmän kehittäminen mm. siten, että tilaajien mahdollisuus hallita omia tilauksiaan ja -tietojaan paranee.

Viime vuosien kehittämistyön ansiosta Tilastokeskus tuntee keskeisimmät asiakkaansa ja sidosryhmänsä aiempaa paremmin ja samalla pystyy varmistamaan asiakas- ja sidosryhmätietojen ajantasaisuuden ja hallinnon EU:n tietosuojasetuksen mukaisesti.

## 12 Asianhallinta ja arkistointi

### 12.1 Asianhallinta ja analoginen asiakirja-aineisto

Tilastokeskuksen hallinnollisten asioiden arkistonmuodostussuunnitelma on päivitetty ja vahvistettu yleis-, talous- ja henkilöstöhallinnon osalta vuonna 2012. Vuonna 2017 aloitettiin AMS, TOS ja prosessikuvaukset -projektissa sekä jo vahvistettujen osioiden päivitys nykytilanteen tasolle sekä muiden tehtäväryhmien (tilastotuotanto, tiedonhallinta, viestintä, markkinointi ja tietopalvelu sekä kansainvälinen yhteistyö) päivitys. Säilytysaikaesitys lähetettiin Kansallisarkistolle keväällä 2018.

Lokakuussa 2017 käynnistyi Tilastokeskuksessa asian- ja dokumentinhallinnan kehittämishanke (ADOHA). Hankkeen tavoitteena on luoda asian- ja dokumentinhallinnalle selkeät prosessit ja toimintatavat sekä ottaa käyttöön helppokäyttöiset asian- ja dokumentinhallinnan sähköiset järjestelmät. Asianhallinnan osalta hanke alkoi vuoden 2017 lokakuun lopussa projektilla, jossa laaditaan tiedonohjaussuunnitelma keskeisimmiltä osin. Jatkoprosjekteissa tullaan hankkimaan ja ottamaan käyttöön sähköinen asianhallinta.

Valtioneuvoston periaatepäätöksen 21.6.2017 mukaisesti viranomaisten hallussa oleva analoginen asiakirja-aineisto digitoidaan vuoteen 2030 mennessä. Päätöksen mukaan asiakirjatieto arkistoidaan vain digitaalisessa muodossa, jolloin suurin osa analogisessa muodossa olevista asiakirjoista hävitetään digitoinnin jälkeen. Digitointi tapahtuu Kansallisarkiston koordinoimassa massadigitoitihankkeessa. Tilastokeskus kartoitti vuonna 2017 sen hallussa olevan manuaalisen asiakirja-aineiston määrää ja laatua osana Kansallisarkiston esiselvityshanketta.

### 12.2 Tilastoaineiston arkistointi

Tilastoaineistojen pysyvä arkistointi varmistaa yhteiskuntatutkimuksen kannalta arvokkaan tiedon säilymisen ja hyödyntämisen myös tulevaisuudessa. Vuonna 2017 uudistetut sähköisen tilastoaineiston arkistointiohje ja arkistoinnin välineet sujuvoittavat arkistointiprosessia.

Arkistoitavan tilastoaineiston kuvaaminen on olennaista hyödyntämisen kannalta, jotta muuttujien sisällöt avautuvat oikeina ja välttytään virhetulkinnoilta. Aineiston ja muuttujien metatietojen laatu tarkoittaa mm. luokitusten liittämistä muuttujille ja muuttujien sisältöjen kuvaamista kattavasti ja ymmärrettävästi. Arkistopaketti sisältää varsinaisen aineiston, muuttujien kuvaukset ja SÄHKE2-normin mukaiset arkistometatiedot, esimerkiksi tiedon siitä, sisältääkö aineisto yksikkötason tietoja.

Jokainen tilasto arkistoi kerran vuodessa tilaston julkaisua varten kerätyn, muokatun ja valmiin havaintoyksikkötason tilastoaineiston. Jos yhdestä laajasta aineistosta tuotetaan monta tilastoa, aineisto arkistoidaan vain kerran. Tilasto itse määrittää, milloin aineisto on valmis arkistoitavaksi ja kirjaa arkistointinsa tavoitepäivämäärän arkistointisuunnitelmaan.

Arkistointisuunnitelmat ja niiden toteumat ovat seurattavissa ajantasaisesti raportointijärjestelmässä Portissa. Arkistoinnin toteuma on vuonna 2018 ensimmäistä kertaa mittarina tulosityksiköiden tuloskorteissa. Tilastokeskus lähetti 2017 Kansallisarkiston päätettäväksi säilytysaikaesityksen pysyvästi säilytettävistä sähköisessä muodossa arkistoitavista tilastoaineistoista.

## 13 Kehittämistoimenpiteet vuodelle 2018

Kehittämistyö jatkuu. Vuonna 2018 Tilastokeskuksessa tietosuojaan liittyviä kehittämiskohteita ovat:

- Kokonaisarkkitehtuurin kehittäminen
- Tietoarkkitehtuuri ja loogiset tietovarannot
- Pseudonymisoinnin käyttöönotto
- Käyttövaltuushallinnan kehittäminen
- Asian- ja dokumentinhallinnan kehittäminen
- Arkistointi
- Lokipolitiikan toimeenpano ja lokien valvonnan kehittäminen
- Tietosuojan vaikuttavuusarviointien tekeminen henkilö pohjaisten tilastojen osalta

Kehittämistoimenpiteiden etenemisestä raportoidaan vuotta 2018 koskevassa tietotilinpäätöksessä.