



Tilastokeskus 

# TIETOTILINPÄÄTÖS 2019

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

Tietoja lainattaessa lähteenä on mainittava Tilastokeskus.

© 2020 Tilastokeskus

# Tiivistelmä

Tilastokeskuksen uusi strategia valmisteltiin syksyn ja talven 2018–2019 aikana. Vuoden 2019 aikana pohdittiin, miten Tilastokeskuksen tulisi muuttaa toimintatapojaan ja järjestää toimintojaan, jotta strategian tavoitteet saavutettaisiin. Tilastokeskuksen strategian ytimessä on tieto ja sen käyttö. Eri lähteistä tulevaa tietoa halutaan yhdistellä toisiinsa nykyistä paremmin ja tehdä se eettisesti ja läpinäkyvästi siten, että tietojen käyttö on luotettavaa ja avointa. Toiminnan tulee olla sitä kautta yhteiskunnassa hyväksyttävää. Jokainen tietoja toiminnassaan käsittelevä organisaatio on vastuussa siitä, että luottamus säilyy eri toimijoiden välillä.

Tilastolain muutos valmisteltiin vuonna 2019. Muutoksessa huomioitiin erityisesti muuttunut tietosuojalainsäädäntö. Lakimuutos tuli voimaan 1.2.2020.

Tilastokeskuksen tietotilinpäätökset vuosilta 2017 ja 2018 löytyvät Tilastokeskuksen kotisivuilta<sup>1</sup>. Tämä Tilastokeskuksen kolmas tietotilinpäätös täydentää kahta aikaisemmin laadittua ja tässä keskitytään erityisesti Tilastokeskuksen tiedonhankinnan kehittämiseen ja siihen, miten sillä pyritään parantamaan tietosuojaa.

Avainsanat: tietotilinpäätös, tietosuoja, tiedonhankinta, henkilötiedot, aineistot

---

1 [www.tilastokeskus.fi](http://www.tilastokeskus.fi)



Tilastokeskuksen pääjohtaja Marjo Bruun.

# Sisälllys

<b>Tiivistelmä</b> .....	<b>3</b>
<b>1 Riskienhallinta ja tietoturva</b> .....	<b>6</b>
1.1 Riskienhallinta- ja valmiusorganisaatioiden uudistus .....	6
1.2 Riskienhallintapolitiikka ja -ohje .....	6
1.3 Käyttövaltuushallinnan Avain-järjestelmän käyttöönotto .....	7
1.4 ESS IT Security Framework -auditointi .....	7
1.5 Pilvipalvelut .....	8
<b>2 Henkilöstön osaamisen varmistaminen</b> .....	<b>8</b>
<b>3 Tiedonhankinnan kehittäminen</b> .....	<b>9</b>
3.1 Yleistä .....	9
3.2 Tiedonantajien informoinnin kehittäminen .....	10
3.3 Uusia tietoaineistoja käyttöön .....	10
3.4 Yksityisen sektorin hallussa olevat aineistot .....	11
3.5 Yksityisen sektorin datat ja kotitaloustilastot .....	12
3.6 Verkkoharavoinnin periaatteet Tilastokeskuksessa .....	13
3.7 Koski-palvelun käyttöönotto .....	15
3.8 Verohallinnon tulorekisterin käyttöönottoon valmistautuminen ...	16
3.9 Henkilötietovarannon uudistamishanke (STIINA) .....	17
3.10 Sijaintitietojärjestelmä .....	19
<b>4 Tietoarkkitehtuuri ja loogiset tietovarannot</b> .....	<b>19</b>
<b>5 Asianhallinta ja arkistointi</b> .....	<b>22</b>
<b>6 Mittarit</b> .....	<b>23</b>
6.1 Henkilöstön osallistuminen tietosuojakoulutukseen vuonna 2019 ..	23
6.2 Rekisteritietojen tarkastuspyynnöt .....	23
6.3 Tilastot .....	24
6.4 Tilastokeskuksen tilastotuotannossaan käyttämien aineistojen vastaanotto .....	24
Kuvio 1: Aineistojen vastaanoton tavat vuonna 2019 (N=253) .....	24
Kuvio 2: Saapuvien aineistojen toimitustiheydet 2019 (N=253) .....	25
Kuvio 3: Tilastokeskuksen käyttöön saadut aineistot vuonna 2019 (N=253) .....	25
Kuvio 4: Erityisiin tietoryhmiin* kuuluvia henkilötietoja tai rikostuomioihin tai rikkomuksiin liittyviä tietoja sisältävien aineistojen osuus saaduista aineistoista vuonna 2019 (N=253) .....	26

6.5	Suorat tiedonkeruut . . . . .	27
	Kuvio 5: Henkilö- ja kotitalouspohjaiset suorat tiedonkeruut kaikista vuoden 2019 suorista tiedonkeruista (N=89) . . . . .	27
	Luettelo 1: Vuonna 2019 toteutetut suorat henkilö- ja kotitalouspohjaiset tiedonkeruut. . . . .	29
	Kuvio 6: Henkilö- ja kotitalouspohjaiset suorat tiedonkeruut vuonna 2019 otoskoon mukaan (N=15). . . . .	30
	Taulukko 1: Henkilö- ja kotitalouspohjaiset suorat tiedonkeruut vuonna 2019 otoskoon mukaan (N=15). . . . .	31
	Kuvio 7: Henkilötieto ja suoralla tiedonkeruulla kerättävät tiedot vuonna 2019 (N=89) . . . . .	31
	Kuvio 8: Arkaluonteinen/erityisiin tietoryhmiin kuuluva tieto suorassa tiedonkeruussa vuonna 2019 (N=89) . . . . .	32
	Taulukko 2: Haastattelujen määrät vuosina 2017–2019. . . . .	32
6.6	Tilaustutkimukset. . . . .	33
	Kuvio 9: Toimeksiantojen lukumäärä vuosittain: asiakasrahoitteiset tilaustutkimukset vuosina, 2010–2019, kpl. . . . .	34
	Taulukko 3. Toimeksiantojen lukumäärä 2010–2019, kpl . . . . .	34
	Kuvio 10: Tilaustutkimukset vuonna 2019 otoskoon mukaan . . . . .	35
6.7	Asianhallinta ja arkistointi . . . . .	37
	Kuvio 11: Arkistoitujen tilastoaineistojen lukumäärät vuosina 2001–2019. . . . .	38
	Taulukko 5: Arkistoitujen tilastoaineistojen lukumäärät vuosina 2001–2019. . . . .	38
6.8	Tutkimusaineistot. . . . .	39
	Taulukko 6: Tutkimusaineistojen käyttöluvut 2017–2019 . . . . .	39
6.9	Kansallinen etäkäyttöpalvelu . . . . .	40
	Taulukko 7: Kansallinen etäkäyttöpalvelu 2017–2019 . . . . .	40

# 1 Riskienhallinta ja tietoturva

## 1.1 Riskienhallinta- ja valmiusorganisaatioiden uudistus

Tilastokeskuksen riskienhallinta- ja valmiusryhmä asetettiin helmikuussa 2019. Vuodesta 2000 asti toimineen ryhmän tehtäväkenttää laajennettiin ja täsmennettiin kattamaan

- tietoturvallisuuden, tietosuojan, jatkuvuuden ja toipumiskyvyn ylläpitäminen, kehittäminen ja koordinointi
- riskienhallinnan koordinointi ja kehittäminen
- varautuminen poikkeusoloihin sekä normaaliajan häiriöihin
- tehtäviin liittyvä sidosryhmäyhteistyö.

Ryhmään kuuluvat ylijohtaja, IT-kehittämispäällikkö, hallintojohtaja, tietohallintojohtaja ja tietoturvapäällikkö. Ryhmän kokoonpanoa täsmennettiin siten, että joka toinen kuukausi riskienhallinta- ja valmiusryhmän kokouksiin osallistuvat tietosuojavastaava, riskienhallinnan koordinoija, viestintäpäällikkö ja toimitiloista vastaava toimistoesimies.

## 1.2 Riskienhallintapolitiikka ja -ohje

Tilastokeskus laati jo vuonna 2013 riskienhallinnan toimeenpanoa varten ”Tilastokeskuksen riskienhallintapolitiikka” ja ”Riskien arviointi Tilastokeskuksessa” -ohjeet. Tämän jälkeen on yhteiskunnassa tapahtunut aihetta sivuavia muutoksia. VAHTI-ohjeistusta riskienhallinnasta on päivitetty ja tietosuojalainsäädäntö ja tiedonhallintalaki ovat tuoneet uudet vaatimukset riskienhallinnalle. Valtiontalouden tarkastusvirasto antoi vuonna 2018 suosituksen riskien- ja jatkuvuuden hallinnan liittämistä osaksi virastojen arjen johtamista ja toimintaa. Julkisen sektorin riskienhallinnan kehittäminen nostettiin hallituskauden toimenpiteenä esille pääministeri Rinteen hallitusohjelmaan.

Tarve riskienhallintaohjeistuksen päivittämiselle oli ilmeinen ja päivitys tehtiin vuoden 2019 aikana. Riskienhallintapolitiikkaa täsmennettiin ja riskien arviointiohjetta laajennettiin kattamaan koko riskienhallinnan prosessi. Riskienhallinnan kehittämistarpeista tehtiin loppuvuodesta kysely, jonka pohjalta laadittiin suunnitelmat riskienhallinnan edelleen kehittämiseksi ja toimeenpanemiseksi vuoden 2020 aikana. Riskienhallintapolitiikka julkaistaan Tilastokeskuksen verkkosivuilla yhdessä loppuvuodesta vahvistetun tietosuojapolitiikan kanssa.

### 1.3 Käyttövaltuushallinnan Avainjärjestelmän käyttöönotto

Tilastokeskus otti tuotantokäyttöön Valtorin tuottaman käyttövaltuuksien hallintajärjestelmä Avaimen maaliskuussa 2019.

Käyttöönoton myötä joukko aiemmin paperilla tehtyjä käyttövaltuusprosesseja sähköistettiin. Tällä muutoksella valtuuksien myöntämisprosessia nopeutettiin merkittävästi.

### 1.4 ESS IT Security Framework -auditointi

Euroopan unionin tilastovirasto Eurostat teetätti tietoturva-auditoinnin jäsenmaille vuosien 2017 ja 2019 välisenä aikana. Suomessa auditointi koski Tullia ja Tilastokeskusta.

Auditoinnin suoritti PricewaterhouseCoopersin Hollannin yksikkö.

Auditoinnissa käytettiin Eurostatin ESS IT Security -asiantuntijaryhmän laatimaa viitekehikkoa, joka perustuu ISO 27001:2013 -standardiin. Tilastokeskus täytti hyväksytysti viitekehiksen vaatimukset syksyllä 2019.



Auditoinnin seurauksena joitain viraston sisäisiä prosesseja muutettiin. Esimerkiksi sovelluskehityksessä tilastosovelluksen julkaisu kehityksestä tuotantoon tapahtuu jatkossa tuoteomistajan kirjallisen hyväksynnän perusteella.

## 1.5 Pilvipalvelut

Tilastokeskuksen pääjohtaja vahvisti keväällä 2019 pilvipalvelulinjaukset. Datan vieminen pilvipalveluihin on aloitettu Tilastokeskuksen Intranetista, viestintävälineistä, sekä asiakas- ja sidosryhmätiedoista.

Seuraavassa vaiheessa on luontevaa ottaa käsittelyyn varsinaiset tilastoaineistot. Tilastoaineistojen pilvikelpoisuus arvioidaan hyväksytyyn kriteeristön pohjalta. Pilvipalveluihin liittyviä toimijoita ja riskejä arvioidaan tätä varten toteutetun riskianalyysityökalun avulla.

Tilastokeskuksen riskienhallinta- ja valmiusryhmä käsittelee uusien pilvipalveluiden riskiarvioinnit. Tiedon tallentamisesta pilvipalveluun päättää kunkin tiedon omistava johtaja laaditun riskiarvioinnin pohjalta. Pilvipalveluiden käyttöä ja vaatimustenmukaisuutta valvoo tätä tarkoitusta varten perustettu tiimi.

## 2 Henkilöstön osaamisen varmistaminen

Tilastokeskus on vuoden aikana päivittänyt ja uusinnut tietosuojaan liittyvää ohjeistustaan. Alkuvuodesta 2019 ohjeistusta tarkistettiin ja täsmennettiin tietosuojalain voimaantulon myötä.



Vuoden aikana Tilastokeskuksessa valmisteltiin ja hyväksyttiin

- Tilastokeskuksen pilvipalvelulinjaukset
- riskienhallintapolitiikka ja -ohje
- verkkoharavoinnin toimintaperiaatteet
- tietosuojapolitiikka vuosille 2020–2023
- tietosuoja- ja tietoturvaohjeet.

Henkilöstön tietosuojaosaamista pyrittiin lisäämään Tietosuojan ABC julkishallinnon henkilöstölle -verkkokurssilla, jonka lähes koko henkilöstö suoritti. Syksyllä henkilöstölle järjestettiin infotilaisuus, jossa asialistalla oli myös ajankohtaiset tietosuoja-asiat.

Tämän lisäksi uuden henkilöstön perehdyttämistä tietosuoja-asioihin lisättiin. Uusille työntekijöille, joista merkittävä osa oli tilastohaastattelijoita, järjestettiin perehdyttämistilaisuuksia. Joitain pienempimuotoisia, tiettyyn teemaan keskittyviä keskustelutilaisuuksia pidettiin vuoden aikana.

Tietosuojavastaavan tukena toimii tietosuojaverkosto.

## 3 Tiedonhankinnan kehittäminen

### 3.1 Yleistä

Tilastolaki velvoittaa tilastoviranomaisia käyttämään tilastoja laadittaessa ensi sijassa hyväksi julkishallinnon tehtävien hoitamisessa kertyneitä sekä elinkeinon- ja ammatinharjoittajien, yhteisöjen ja säätiöiden tavanomaisen toiminnan seurauksena syntyneitä tietoja. Laki myös velvoittaa keräämään tiedot taloudellisesti ja tiedonantajille aiheutuva haitta ja kustannukset minimoiden. Tilastoja laativan viranomaisen on huolehdittava siitä, että tiedonantajilta pyydetään vain tilastojen laatimisen kannalta välttämättömät tiedot.

EU:n yleisen tietosuoja-asetuksen 5 artiklassa on säädetty henkilötietojen käsittelyn yleisistä periaatteista, jotka on tilastoinnissakin huomioitava.

Nämä periaatteet on huomioitu Tilastokeskuksen tiedonhankinnan kehittämisessä.

## 3.2 Tiedonantajien informoinnin kehittäminen

Tilastokeskus pyrkii kehittämään tiedonantajien informointia sekä verkossa että tiedonantajille toimitettavien saatekirjeiden osalta<sup>2</sup>. Samalla on ryhdytty hyödyntämään uusia viestintätapoja, kuten videoita. Vuonna 2019 Metropolian AMK:n opiskelijat tekivät harjoitustyönä Tilastokeskukselle tiedonkeruusta ja tietosuoja-asioista animaatiot ”Sinun vastauksesi ovat tärkeitä” ja ”Tietosuojasi on meille tärkeää”.

Ne ovat nähtävissä YouTubessa sekä Tilastokeskuksen tiedonkeruusivuilla:

- Sinun vastauksesi ovat tärkeitä: <http://www.tilastokeskus.fi/keruu> ja
- Tietosuojasi on meille tärkeää: <http://www.tilastokeskus.fi/keruu/luottamuksellisuus.html>

Yksittäisissä henkilö- ja kotitalouspohjaisissa tiedonkeruissa videot ovat myös yleistyneet tiedonantajien informoinnissa ja motivoinnissa.

## 3.3 Uusia tietoaineistoja käyttöön

Lähtökohdat uusien aineistojen käyttöönotolle tilastotoimessa ovat hyvät, koska Suomessa on yleisesti ottaen laadukkaat hallinnolliset aineistot, joita kehitetään jatkuvasti. Tämä kehittäminen näkyy olemassa olevien rekis-

---

<sup>2</sup> Tilastokeskuksen omissa ja asiakasrahoitteisissa suorissa tiedonkeruissa tiedonantajille toimitettavia informointikirjeitä on päivitetty tarkemmin vastaamaan tietosuoja-asetusta.

tereiden laadun parannuksina ja uusien tietojen kokoamisena uusiksi rekistereiksi tai palveluiksi. Uusimpia merkittäviä tietovarantoja ovat vuonna 2019 käyttöön otetut tulorekisteri ja Koski opintosuoritus- ja tutkinto-tietovaranto sekä lähivuosina käyttöön otettava sähkön vähittäismarkkinoiden keskitetty tiedonvaihtojärjestelmä, Fingrid Datahub. Sähköinen osakehuoneistorekisteri (ASREK) puolestaan yhdistää asunto-osakkeita koskevat tiedot ja omistajamerkinnot sähköiseksi palveluksi. Uudet taloyhtiöt on rekisteröity järjestelmään vuodesta 2019, vanhat taloyhtiöt ovat voineet hyödyntää palvelua toukokuusta 2019.

### 3.4 Yksityisen sektorin hallussa olevat aineistot

Julkisen sektorin ulkopuolella syntyy yhä enemmän digitaalisessa muodossa olevaa tietoa. Uutta digitaalista tietoa generoivat yritysten tiedon- ja asiakashallintaan liittyvät järjestelmät, tekniset sensorit ja mittarit, www-asiointialustat ja sosiaalisen median alustat sekä myös digitalisoituvat julkishallinto. Tällaisia tietolähteitä voidaan käyttää parantamaan nykyisiä tilastoja ja niiden avulla voidaan tuottaa kokonaan uusia tilastoja ja muita tietoon perustuvia julkisia palveluita. Haasteena yksityisen sektorin datan käyttämiselle on usein toiminnan sopimusperusteisuus sekä myöskin datan luonne. Yksityisen sektorin mielenkiintoisin data koskee yleensä yritysten asiakkaita ja henkilöstöä.

Pääpaino uusien aineistojen hyödyntämisessä on olemassa olevien tilastojen ja tietotuotteiden kehittäminen, mutta Tilastokeskus pyrkii myös julkaisemaan uusia kokeellisia tilastoja.

Sisäisissä tietojenkäsittelyprosesseissa Tilastokeskuksen tavoitteena on keskittää tietoja ns. loogisiin tietovarastoihin ja yhtenäistää tietojen käsittelyä. Henkilötiedon käsittelyn osalta tämä tarkoittaa, että **tieto käsitellään vain kerran ja tieto on tallennettu kertaalleen** henkilötietovarantoon, josta tietoa hallinnoidaan ja käytetään keskitetysti.

## 3.5 Yksityisen sektorin datat ja kotitaloustilastot

Yksityisen sektorin rekisteriaineistojen sisältöjä ja käyttömahdollisuuksia otospohjaisissa kotitaloustilastoissa selvitettiin keväällä 2019.

Tilastokeskuksen kotitaloustilastoissa hyödynnetään kattavasti julkisen sektorin hallinnollisia aineistoja. Sen sijaan yksityisen sektorin aineistoihin pääsy on ollut hyvin rajattua, eikä aineistojen käyttökelpoisuudesta tilastointikäyttöön ole ollut selkeää kuvaa.

Selvityksen tavoite oli kartoittaa mm. seuraavien aineistojen soveltuvuutta tilastointiin: kauppojen kanta-asiakas- ja maksupääteaineistot, vakuutusyhtiöiden asiakasaineistot sekä pankkialan tietovarannot koskien opintolainoja, talletuksia ja maksutapahtumia.

Selvityksessä keskityttiin aineistojen sisältöön ja soveltuvuuteen, siihen katsomatta onko Tilastokeskuksella nykytilanteessa mahdollisuus todella saada näitä aineistoja käyttöönsä.

Työssä tunnistettiin kaksi uutta tietolähdettä: Opetushallituksen varhaiskasvatuksen tietovaranto (VARDA) ja edellä mainittu sähkön vähittäismarkkinoiden tiedonvaihtojärjestelmä Fingrid Datahub. Mainitut tietolähteet ovat vielä rakenteilla ja ne pyritään integroimaan tilastotuotantoon vuodesta 2022 alkaen.

Sen sijaan selvitys ei päätynyt suosittamaan kauppaketjujen kanta-asiakasaineistojen käytön edistämistä kotitalouksien kulutuksen tilastoinnissa. Aineistojen vinous saattaisi aiheuttaa tilanteen, jossa aineiston paikkaaminen ja korjaaminen olisi jopa työlämpi prosessi kuin nykyinen malli.

Selvityksen ja Norjan tilastoviranomaisten tekemien kokeilujen perusteella arvioitiin, että kehittämisponnistukset kannattaa suunnata laajemmin kauppojen maksupääteaineistojen ja pankkien transaktioaineistojen hyödyntämiseen.

## 3.6 Verkkoharavoinnin periaatteet Tilastokeskuksessa

Tilastokeskuksessa valmisteltiin vuoden 2019 aikana ”Verkkoharavoinnin periaatteet Tilastokeskuksessa” -ohjeistus, jossa kerrotaan millaisia periaatteita Tilastokeskus noudattaa kerätessään tietoja verkosta automatisoidusti web scraping- eli verkkoharavointimenetelmin. Periaatteisiin voi tutustua Tilastokeskuksen kotisivuilla [www.tilastokeskus.fi](http://www.tilastokeskus.fi).

Tilastokeskuksen on tilastolain mukaan kerättävä tilastoihin tarvittavat tiedot mahdollisimman tehokkaasti ja tiedonantajien vastausrasite minimoiden. Muun muassa tästä syystä Tilastokeskus kehittää ja ottaa käyttöön erilaisia uusia tiedonkeruumenetelmiä kuten verkkoharavointi, jonka avulla voidaan kerätä internetistä jo valmiiksi löytyviä tietoja. Lisäksi verkkoharavoinnilla voidaan tehostaa tilastotuotantoa, kun kalliita suoria tiedonkeruita voidaan korvata automatisoiduilla ratkaisuilla.

Tilastokeskus huomioi verkkoharavointia suunnitellessaan juridiset ja eettiset kysymykset ennen varsinaiseen verkkoharavointiin ryhtymistä.

Tilastokeskus toteuttaa verkkoharavointia kuten muutakin toimintaansa; läpinäkyvästi ja avoimesti, jolloin sitä voidaan tarvittaessa arvioida. Tämän toteuttamiseksi Tilastokeskus tekee verkkoharavoinnista ja sillä hankittavan aineiston käytöstä aina tiedonkeruupäätöksen. Aineisto kuvataan muiden hankittavien aineistojen kaltaisesti Tilastokeskuksen tiedonkeruurekisteriin.

Eettisen kestävyuden varmistamiseksi kaikessa Tilastokeskuksen harjoittamassa verkkoharavoinnissa noudatetaan seuraavia periaatteita:

**Lainmukaisuus:** Lainsäädäntö, tietosuoja mukaan lukien, huomioidaan jo haravoinnin suunnitteluvaiheessa, ja sitä noudatetaan täysimääräisesti. Mahdollisia muutoksia oikeustilassa (lainsäädäntö, oikeuskäytäntö, vaikiintuneet tulkinnat) seurataan.

**Läpinäkyvyys:** Verkkoharavoinneista ilmoitetaan julkisesti Tilastokeskuksen kotisivuilla. Samassa yhteydessä annetaan kustakin verkkoharavoinnista seuraavat tiedot: haravoinnin tarkoitus, haravoinnin kohteena olevat tietotyypit sekä yhteystiedot, joihin verkkosivuston ylläpitäjä voi ottaa yhteyttä lisätietojen pyytämiseksi tai haravoinnin rajoittamiseksi. Jos kyse on tiedonantovelvollisuuteen perustuvasta tiedon keräämisestä, tietoja kerätessä huomioidaan lisäksi tilastolain mukainen tiedottamisvelvollisuus. Henkilötietojen käsittelyä koskevaa informointivelvollisuutta noudatetaan.

**Vähimmän haitan periaate:** Verkkoharavointi toteutetaan kaikin puolin siten, että haravoinnista aiheutuu verkkosivuston toiminnalle ja sen omistajille mahdollisimman vähän haittaa ja kustannuksia.

**Oikeus kieltää:** Verkkosivuston ylläpitäjille annetaan oikeus kieltää haravointi (opt-out) ottamalla yhteyttä Tilastokeskukseen. Kieltopyyntöjä kunnioitetaan, ja ne ilmoitetaan yhteiseen listaan (ns. black list).

**Tilastointiperiaatteiden noudattaminen:** Tilastojen laatimiseen sovellettavia tiedonkeruuta sekä tilastojen suunnittelua ja laadintaa koskevia menettelytapoja ja periaatteita noudatetaan myös verkkoharavoinnissa. Sama koskee Tilastokeskuksen toiminnan perustana olevia tilasto- ja tutkimusalan ammattieettisiä periaatteita.

**Käyttöehtojen tarkistaminen:** Verkkoharavointi kohdistetaan toistaiseksi ainoastaan sellaisille verkkosivustoille, joiden käyttöehdot on tarkistettu. Haravoinnin katsotaan olevan sallittua, jos sitä ei ole nimenomaisesti kielletty tai kieltö on selkeästi rajattu koskemaan pelkästään kaupallista toimintaa. Epäselvissä tilanteissa voidaan ottaa yhteyttä sivuston ylläpitäjään. Jos vastausta ei saada kohtuullisen ajan kuluessa, sivustoa voidaan haravoida.

## 3.7 Koski-palvelun käyttöönotto

Opetushallituksen ylläpitämän Koski-palvelun (Kansallisten opiskeluoikeuksien ja suoritusten keskitetty integraatiopalvelu) tietojen käyttöönotto tilastoinnissa alkoi vuonna 2019 lukiokoulutuksesta, kun sopimus tietojen luovuttamisesta Tilastokeskukselle allekirjoitettiin kesäkuussa. Koski-palvelun tietojen käyttöön siirrytään vaiheittain; vuonna 2020 vuorossa ovat ammatillinen koulutus sekä esi- ja perusopetus.

Ensimmäiset tiedot Koskesta saatiin vuoden 2019 syksyllä tiedostotoimituksena. Vuoden 2020 alusta tietojen haku tapahtuu rajapinnan kautta. Koskesta saatavia tietoja täydennetään suoralla tiedonkeruulla koulutuksen järjestäjille, koska palvelun ulkopuolelle jää tilastoitavaa koulutusta (mm. Ahvenanmaa).

Uusien aineistojen myötä Tilastokeskus luopui lukiokoulutuksen päättäneiden ainevalintoja koskevasta tiedonkeruusta ja lukiokoulutuksen opiskelijat – tiedonkeruusta. Uuden palvelun myötä Tilastokeskus ei enää kysy tarvittavia tietoja erikseen vaan koulutuksen järjestäjät raportoivat tietonsa yhteen palveluun, josta tietojen käyttöön oikeutetut tahot noutavat tarvitsemansa tiedot. Tilastojen näkökulmasta järjestely on taloudellisesti tehokas ja tietoturvallinen tapa kerätä ja liikutella suuria tietomääriä.

Koski-palvelu on kohdannut alkuvaiheessa tyypillisiä laatuhaasteita. Näitä haasteita pyritään korjaamaan yhdessä palveluun tietoja tuottavien, palvelua ylläpitävien ja tilastojien yhteistyöllä. Tällaiset haasteet eivät ole ainutlaatuisia vaan yleensä uusien tietoaineistojen käyttöönottoon liittyy eri toimijoista johtuvaa alkukankeutta. Uuden palvelun kattavuuden ja tietosisällön laatuhaasteina joudutaan alkuvaiheessa selvittelemään ja korjaamaan.

Uusilla keskitetyillä palveluilla on tyypillisesti paljon hyödyntäjiä, joten tietojen laatuun kiinnitetään paljon huomiota ja tarve saattaa toiminta vaatimusten mukaiseksi on suuri. Tämä tapahtuu, kun tietojen tuottajat



tottuvat uuteen raportointiin, tiedon siirto heiltä järjestelmään toimii ja järjestelmän laadun valvontaa tehostetaan.

Koski-palvelun laajentaminen kattamaan tutkintoon johtamaton koulutus sekä opetushallinnon toimialaan kuulumattomat koulutukset antaisivat mahdollisuuden luopua täydentävistä tiedonkeruista, joita varten Tilastokeskuksessa joudutaan ylläpitämään tiedonkeruuvalmiuksia.

Koski-palvelu myös parantaa tilastoinnin mahdollisuuksia ja laatua, koska vuonna 2020 ensimmäistä kertaa esi- ja perusopetuksen oppilaiden tiedot tullaan saamaan henkilötasoisina.

### 3.8 Verohallinnon tulorekisterin käyttöönottoon valmistautuminen

#### Tulorekisterin haasteet

Tilastokeskus tuli mukaan tulorekisterin käyttäjäksi 1.1.2020 alkaen. Tulorekisteri on monipuolinen kokonaisuus, joka aiheuttaa nykyiselle tilastotuotannolle haasteita, mutta antaa myös uusia mahdollisuuksia uudelleen tilastoinnille ja tietojen jakelulle.

Tulorekisteri sisältää ensimmäisessä vaiheessa (2019 ja 2020) palkka- ja työsuhdetietoja. Etuus- ja eläketiedot tallennetaan tulorekisteriin vuoden 2021 alusta lähtien. Tulorekisteri tulee ensimmäisessä vaiheessa vaikuttamaan palkka- ja työssäkäyntitilastossa sekä yritystilastoissa. Tulorekisteriin on mahdollista tallentaa myös vapaaehtoisia tietoja. Tilastokeskuksella on lain mukaan mahdollisuus hyödyntää näitä tietoja. Jos palkanmaksajat, tietojen ilmoittajat ovat antaneet myös vapaaehtoisesti ilmoitettavat tiedot tulorekisteriin täydellisesti, on sillä vaikutusta siihen, mitä tietoja Tilastokeskus erillisillä kyselyillä tiedustelee. Tulorekisterin käytön alkuvaiheessa vaikuttaa siltä, että yleensä tietojen ilmoittajat ovat antaneet vain pakolliset tiedot. Oletettavaa on, että

tulorekisterin vakiintuessa myös vapaaehtoisia tietoja ilmoitetaan yhä useammin. Mikäli näin käy, se vähentää erityisesti yritysten vastausrasitetta.

### Tulorekisterin edut

Tulorekisterin etuna voidaan pitää sitä, että tiedot saadaan keskitetyksi yhteen paikkaan ja nopealla aikataululla. Tilastokeskus saa lähes koko tulorekisterin tiedot käyttöönsä tilastointia ja sen kehittämistä varten. Näin on mahdollista luopua useista suorista tiedonkeruista, mikä parantaa tietosuojaa vähentämällä henkilötietojen siirtoa ja käsittelyä. Tulorekisteritietoja on mahdollista rikastaa myös Tilastokeskuksen muilla tiedoilla. Tämä tarjoaa uusia mahdollisuuksia uuden, ajantasaisen tiedon käyttöön niin tilastoinnissa kuin julkiseen keskusteluunkin.

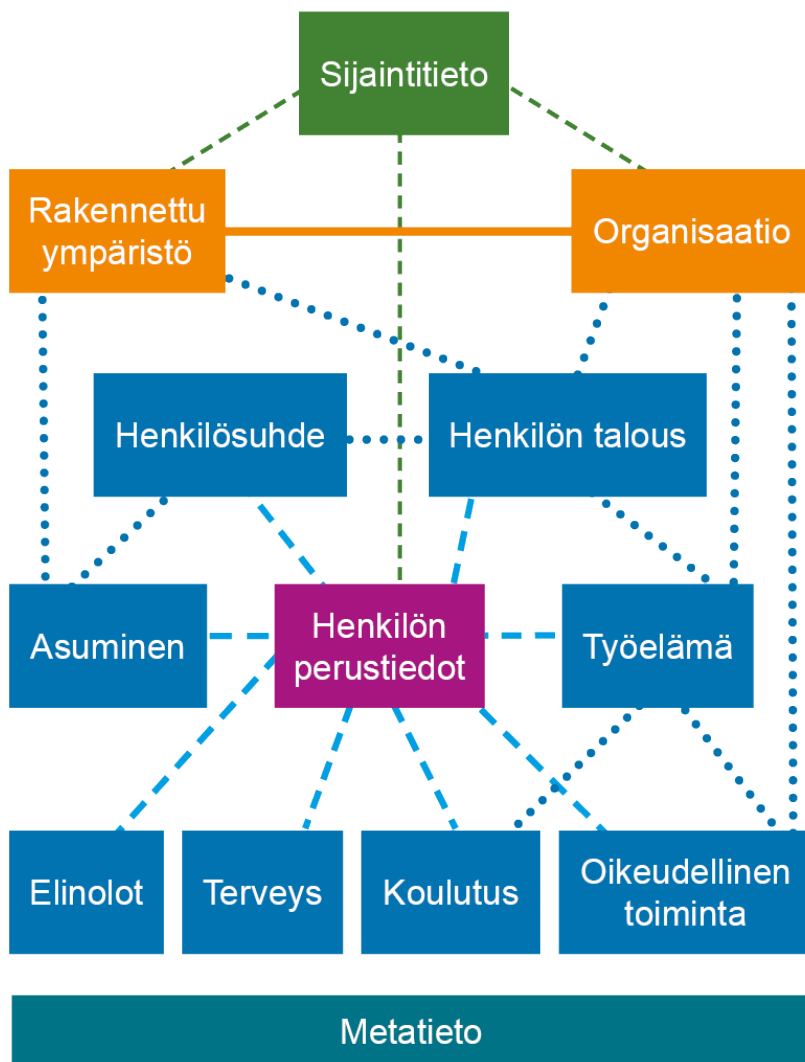
## 3.9 Henkilötietovarannon uudistamishanke (STIINA)

Tilastokeskus on uudistamassa henkilötietovarantoaan henkilötietokokonaisuuden uudistushankkeessa. Hankkeessa toteutetaan ensimmäiset osat uudesta henkilötietovarannosta ja siihen liittyvistä tilastotuotantoprosesseista vuosien 2016–2020 aikana. Hankkeen tavoitteena on rakentaa palvelupohjaisen tietoarkkitehtuurin mukaisesti yhteiskäyttöiset, jatkuvapäivitteiset ja metatieto-ohjautuvat tuotantojärjestelmät. Hankkeen taustalla on henkilökokonaisuuteen kuuluvien tilastojen tietojärjestelmien uudistustarpeen lisäksi toimintaympäristön suuret muutokset. Näistä merkittävimpiä ovat tulorekisterin ja opetushallinnon Koski-palvelun käyttöönotto.

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 tietosuojaa. Hankkeen myötä otetaan käyttöön henkilön suorien

tunnistetietojen pseudonymisointi kaikissa henkilö pohjaisissa tuotantoprosesseissa. Kun uusi tietojärjestelmä saadaan käyttöön, ei tilastoaineistoissa tarvitse enää käyttää henkilön suoria tunnistetietoja keruuvaiheen jälkeen.

Kuva 1: Henkilötietokokonaisuuden loogiset tietovarannot



## 3.10 Sijaintitietojärjestelmä

Sijaintitiedot ovat yksi Tilastokeskuksen tietoarkkitehtuurin tunnistamasta kuudesta käsittelyvaiheen loogisesta tietovarannosta (luku 4). Sijainti on tiedon ominaisuus. Sijaintitieto voi olla täsmällistä koordinaattitietoa tai viitteellistä osoitteiden tai tunnusten avulla kohteita paikantavaa tietoa. Paikkatieto on tietoa, jolla on sen kohteita yhdistävä ominaisuus: sijainti.

Tilastokeskus julkaisi kesäkuussa 2018 sijaintitiedon viitearkkitehtuurin. Siinä esitettiin, miten Tilastokeskuksessa saavutetaan keskitetty toimintamalli sijaintitietojen tuottamiseen, miten sijaintitiedot on määritetty keskitettyyn sijaintitietovarantoon ja kuinka sijaintitietoja käytetään yhdenmukaisesti. Viitearkkitehtuurin ehdottamat sijaintitietojärjestelmän keskeiset osat otettiin käyttöön kesäkuussa 2019. Sijaintitietojärjestelmä koostuu sijaintitietovarannosta sekä sen päälle sijaintitietojen toistuviin käsittelytarpeisiin toteutetuista teknologiariippumattomista palveluista. Sijaintitietovaranto kattaa sijaintikohteiden koordinaatti- ja osoitetiedot, tietoprosesseissa käytettävät alueraja-aineistot sekä sijaintikohteiden linkit kyseisiin alueisiin. Sijaintitietojärjestelmä integroidaan tulevana vuosina Tilastokeskuksen tietoarkkitehtuurin mukaisiin tietovarantoihin siten, että niissä ei toisteta sijaintitietoja, vaan sijaintitiedot poimitaan tunnisteiden avulla sijaintitietovarannosta. Sijaintitietovaranto integroidaan henkilötietovarantoon STIINA-hankkeen tuloksena vuoden 2020 aikana.

## 4 Tietoarkkitehtuuri ja loogiset tietovarannot

Tietoarkkitehtuurihanke edisti vuosina 2017–2018 tilastotuotannon tietoarkkitehtuurin kehittymistä ja tuotti ylätasoa tietoarkkitehtuurimäärittämiä. Tilastotuotannon tietoarkkitehtuuri -hankkeessa tuotettujen tietoarkkitehtuurikuvausten keskiössä on tieto. Suhteessa nykytilaan tietolähtöisyyden voidaan katsoa olevan ajattelutavan muutos perinteiseen ti-

lastolähtöiseen kehittämistoimintaan verrattuna. Keskittymällä tietoihin voidaan tavoitella määritellä siten, että tiedot ovat yhteiskäyttöisiä sekä laadukkaita, ja toiminnassa voidaan välttää tarpeetonta kopiointia ja tietojen päällekkäisyyttä tai epäyhtenäisyyttä.

Tietoarkkitehtuurihanke tuotti tilastoprosessille loogisten tietovarantojen ylätasoa tavoitetilan ja vei eteenpäin Tilastokeskuksen tilastotuotannon Generic Statistical Information Model (GSIM) -pohjaista käsittemallia. Hanke myös suunnitteli Tilastokeskuksen tieto-objektien tunnistamisen hallinnan tulevaisuuden ja koordinoi tunnistepalvelun määrittelyä ja toteuttamista. Tämän lisäksi hanke hahmotteli käsitteiden hallinnan tulevaisuutta ja yhteyksiä kansalliseen yhteentoimivuuteen ja sen välineisiin.

Tilastotuotannon tavoitetilan ylätasoa loogiset tietovarannot ilmenevät kuvasta kaksi. Tilastoja tuotetaan käsittelemällä ja yhdistelemällä tietoja yhdestä tai useammasta tietovarannosta. Tietovaranto voi sisältää useiden tietojärjestelmien tietokantoja tai rekistereitä. Vastaavasti sama looginen tietovaranto voi sisältää useiden eri tahojen hallinnoimia tietoja, vaikka tiedot sijaitsisivatkin samassa fyysisessä tietokannassa.

Kuva 2: Tilastotuotannon loogiset tietovarannot

<b>Tiedon- hankinta</b>	<b>Tiedonhankinnan tietovarannot</b>	
<b>Käsittely</b>	<b>Henkilö- tietovaranto</b>	<b>Hyödyke- tietovaranto</b>
	<b>Organisaatio- tietovaranto</b>	<b>Päästö- ja energia- tietovaranto</b>
	<b>Rakennetun ympäristön tietovaranto</b>	<b>Sijainti- tietovaranto</b>
<b>Analyysi ja raportointi</b>	<b>Mikro- tietovaranto</b>	<b>Makro- tietovaranto</b>
<b>Julkaiseminen</b>	<b>Julkaisemisen tietovarannot</b>	
<b>Metatietovaranto</b>		

Yksittäisten tietovarantojen täsmällinen sisältö ja rakenne tulee määrittellä erikseen. Henkilöpohjaisten tilastojen STIINA-kehittämishankkeessa toteutettiin tarkemman tason loogisten tietovarantokokonaisuuksien kuvaus henkilötietokokonaisuuden osalta. Myös sijaintitietovaranto, rakennetun ympäristön tietovaranto sekä päästöjen ja energian tietovaranto on määri-

telty tarkemmalla tasolla. Seuraava vaihe kohti tavoitetilaa etenemisessä on tarvittavien käsitteellisen ja loogisen tason kuvausten tuottaminen kaikille tarkemmille osakokonaisuuksille.

## 5 Asianhallinta ja arkistointi

Tilastokeskuksen hallinnollisten asioiden arkistonmuodostussuunnitelma on päivitetty ja vahvistettu yleis-, talous- ja henkilöstöhallinnon osalta vuonna 2012. Päivitettyyn ja täydennettyyn AMSiin saatiin Kansallisarkiston päätös pysyvästi säilytettävien asiakirjojen osalta tammikuussa 2019. Tämän päätöksen pohjalta Tilastokeskus vahvisti määräaikaisesti säilytettävien asiakirjojen säilytysajat.

Tilastokeskuksessa lokakuussa 2017 käynnistynyt asian- ja dokumentinhallinnan kehittämishanke (ADOHA) jatkui. Hankkeen alaprojektissa täydennettiin tiedonohjaussuunnitelmaluonnosta vahvistetuilla säilytysajoilla salassa pidettävien asiakirjojen metatiedoilla. Projekti suositteli etenemään sähköisen asianhallintajärjestelmän hankinnassa siten, että virasto valitsisi Valtorin tarjoaman VirastoVAHVA-asianhallintajärjestelmän, jonka käyttöönotto olisi vuonna 2020.

Valtionhallinnon pysyvästi säilytettävät paperiset asiakirja-aineistot tullaan muuttamaan digitaaliseen muotoon lähivuosina Kansallisarkiston koordinoiman massadigitointihankkeen muodossa. Digitoinnin jälkeen aineistot ovat julkisesti helposti saatavilla käyttörajoituksista riippuen. Tilastokeskus valmistautui vuonna 2020 alkavaan digitointiin aineiston inventoinnilla sekä laatimalla digitoitavasta aineistosta seulontaesitykset Kansallisarkistolle. Seulontapäätökset saatiin vuoden 2019 aikana. Määräajan säilytettäviä, säilytysajaltaan umpeutuneita aineistoja hävitettiin. Aineiston luettelointia arkistonhallintajärjestelmään metatietoineen jatkettiin.



Analogisia asiakirja-aineistoja eli tietovarantoja oli vuoden 2019 lopussa noin 2400 hyllymetriä, joista pysyvästi säilytettävää on noin 1300 hyllymetriä. Aineistot pitävät sisällään sekä Tilastokeskuksen hallinnon asiakirjoja että tilastoaineistoja.

## 6 Mittarit

### 6.1 Henkilöstön osallistuminen tietosuojakoulutukseen vuonna 2019

- Tietosuojan ABC -koulutuksen ja siihen liittyvän testin suoritti 694 työntekijää.
- Uusien haastattelijoiden tietosuojaperehdytykseen osallistui 30 henkilöä.
- Uusien työntekijöiden tietosuojainfoon osallistui 25 henkilöä.

### 6.2 Rekisteristietojen tarkastuspyynnöt

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

Rekisteritietojen tarkastuspyynnöt

#### **Vuosi Pyyntöjä**

2018	5
2019	4

## 6.3 Tilastot

Vuonna 2019 Tilastokeskuksella oli tuotannossa 155 tilastoa, joista 140 julkaisi tietoja. Tilastokeskuksen tuottamat otospohjainen tulonjakotilasto ja rekisteripohjainen tulonjaon kokonaistilasto yhdistyivät yhdeksi tulonjakotilastoksi.

## 6.4 Tilastokeskuksen tilastotuotannossaan käyttämien aineistojen vastaanotto

Tilastokeskuksessa toimii keskitetty aineistojen vastaanottopalvelu (Ava). Vuonna 2018 saapui 26 000 (korjattu tieto) tiedostoa ja vuonna 2019 yhteensä 24 000 tiedostoa.

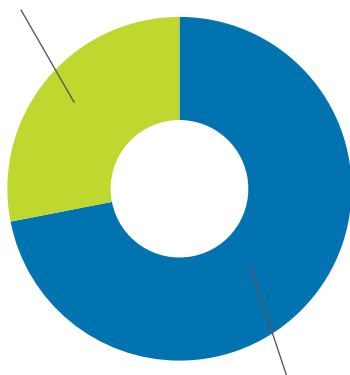
Edellä mainittujen aineistojen lisäksi Tilastokeskukseen saapuu keskitetyn vastaanoton ohii jonkin verran lähinnä taulukkomuotoisia aineistoja.

Seuraavat kuviot perustuvat Tilastokeskuksen tiedonkeruurekisteriin kuvattujen aineistokuvausten tietoihin.

Kuvio 1: Aineistojen vastaanoton tavat vuonna 2019 (N=253)

Muulla tavalla saapuvat aineistotoimitukset

28 %

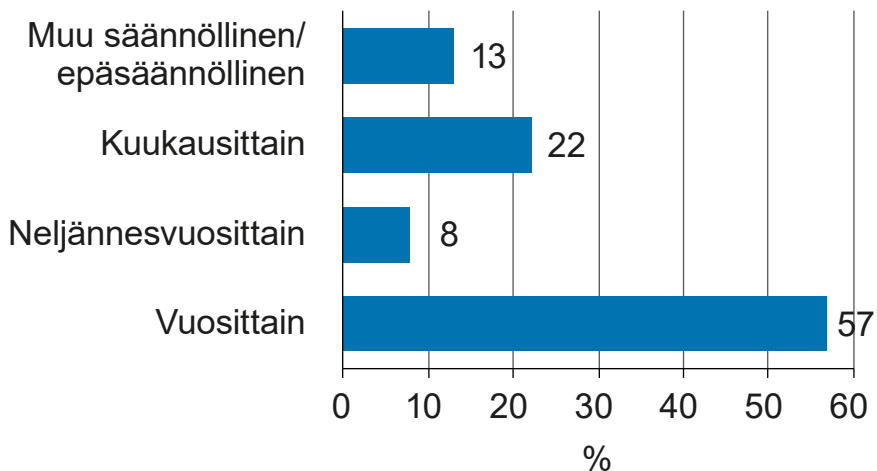


Linjasiirtona saapuvat aineistotoimitukset

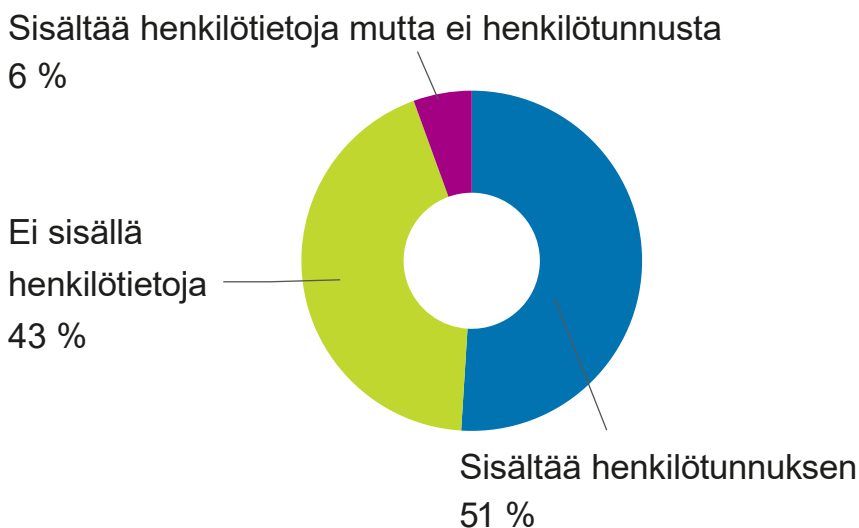
72 %

Keskitetyn aineistojen vastaanoton kautta tietoja toimitti 68 eri tiedon-  
tuottajaa. Tiedon siirtää Tilastokeskukseen joko sen tuottaja tai jokin hei-  
dän valtuuttamansa taho (esim. Tieto Oy, CGI).

Kuvio 2: Saapuvien aineistojen toimitustiheydet 2019 (N=253)



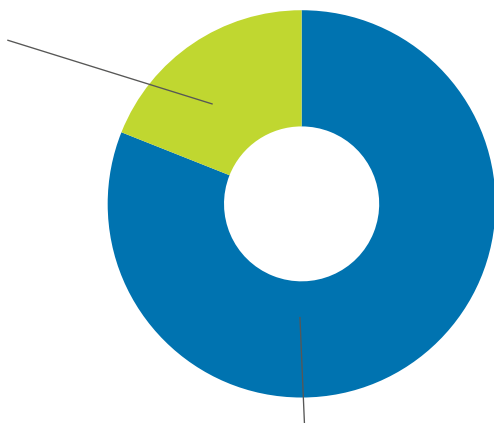
Kuvio 3: Tilastokeskuksen käyttöön saadut aineistot vuonna 2019 (N=253)



Kuvio 4: Erityisiin tietoryhmiin\* kuuluvia henkilötietoja tai rikostuomioihin tai rikkomuksiin liittyviä tietoja sisältävien aineistojen osuus saaduista aineistoista vuonna 2019 (N=253)

Sisältää erityisiin tietoryhmiin kuuluvia henkilötietoja tai rikostuomioihin tai rikkomuksiin liittyviä tietoja

19 %



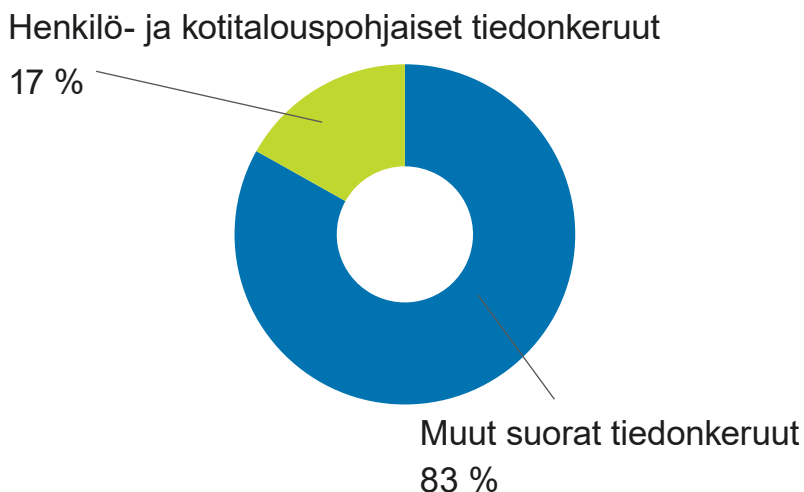
Ei sisällä erityisiin tietoryhmiin kuuluvia henkilötietoja eikä rikostuomioihin tai rikkomuksiin liittyviä tietoja

81 %

\*Erityisiin tietoryhmiin kuuluvat tiedot ovat EU:n yleisen tietosuojasetuksen 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) ja EU:n yleisen tietosuojasetuksen 10 artiklassa määritellyt rikostuomioihin ja rikkomuksiin liittyvät tiedot.

## 6.5 Suorat tiedonkeruut

Kuvio 5: Henkilö- ja kotitalouspohjaiset suorat tiedonkeruut kaikista vuoden 2019 suorista tiedonkeruista (N=89)



Kolmessa korjausrakentamisen tiedustelussa otokseen kuuluu yhteisöjen lisäksi myös yksityishenkilöitä, rakennusten omistajia. Nämä kolme tiedonkeruuta luokitellaan tässä tarkastelussa mukaan henkilöpohjaisiin tiedonkeruisiin.

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

Poimintoja vuoden 2019 muutoksista suorissa tiedonkeruissa Vuonna 2019 tehtiin uudet henkilö- ja kotitalouspohjaiset Jakamistalous-tiedustelu, sekä laaja Ulkomailla eli muualla kuin Suomessa suoritettut koulutukset -tiedonkeruu. Niistä löytyy lisätietoa verkkosivulta [www.tilastokeskus.fi/keruu](http://www.tilastokeskus.fi/keruu).

Opetushallituksen ylläpitämästä Koski-palvelusta<sup>3</sup> saatavien tietojen myötä Tilastokeskus luopui lukiokoulutuksen päättäneiden ainevalintoja koskevasta tiedonkeruusta ja lukiokoulutuksen opiskelijat –tiedonkeruusta. Koski-palvelusta saatavia tietoja täydennetään edelleen keräämällä palvelun ulkopuolelle jäävää tilastoitavaa koulutusta mm. Ahvenanmaalta.

## Suomalaisten matkailukysely

Suomalaisten matkailu -tiedonkeruussa tiedot kerättiin puhelinhaastatteluna edellisen kuukauden aikana päättyneistä matkoista vuodesta 2012 huhtikuuhun 2019.

Suomalaisten matkailu -tilaston tiedonkeruumalli muuttui vuoden 2019 alusta siten, että tiedot kerätään omana erillisenä kyselynään, jossa vastaaja voi ilmoittaa matkansa joko verkossa lomakkeella tai puhelinhaastattelussa. Mallin muutoksen yhteydessä tiedonkeruun otoskokoa pienennettiin 1 333 vastaajaan kuukaudessa (aiemmin 2 350 henkilöä) ja tutkimusjaksoa pidennettiin kolmeen kuukauteen. Uusi keruu on nimeltään Suomalaisten matkailukysely.

## Kuluttajien luottamustutkimus

Kuluttajien luottamustutkimus oli huhtikuuhun 2019 asti nimeltään kuluttajabarometri. Se tehtiin puhelinhaastatteluilla yhdessä suomalaisten matkailututkimuksen kanssa.

Kuluttajabarometrin tiedot kerättiin puhelinhaastatteluilla vuodesta 1987 lähtien. Toukokuusta 2019 alkaen kuluttajien luottamustutkimus toteutettiin puhelinhaastattelujen lisäksi itse täytettävällä verkkolomakkeella. Kumpaakin keruutapaa varten on oma kyselylomakkeensa.

Tutkimuksen uuden perusjoukon muodostavat Suomessa asuvat 18–74-vuotiaat henkilöt (kuluttajabarometrissa 15–84-vuotiaat). Otokoko pieni

---

3 Kansallisten opiskeluoikeuksien ja suoritusten keskitetty integraatio-palvelu.

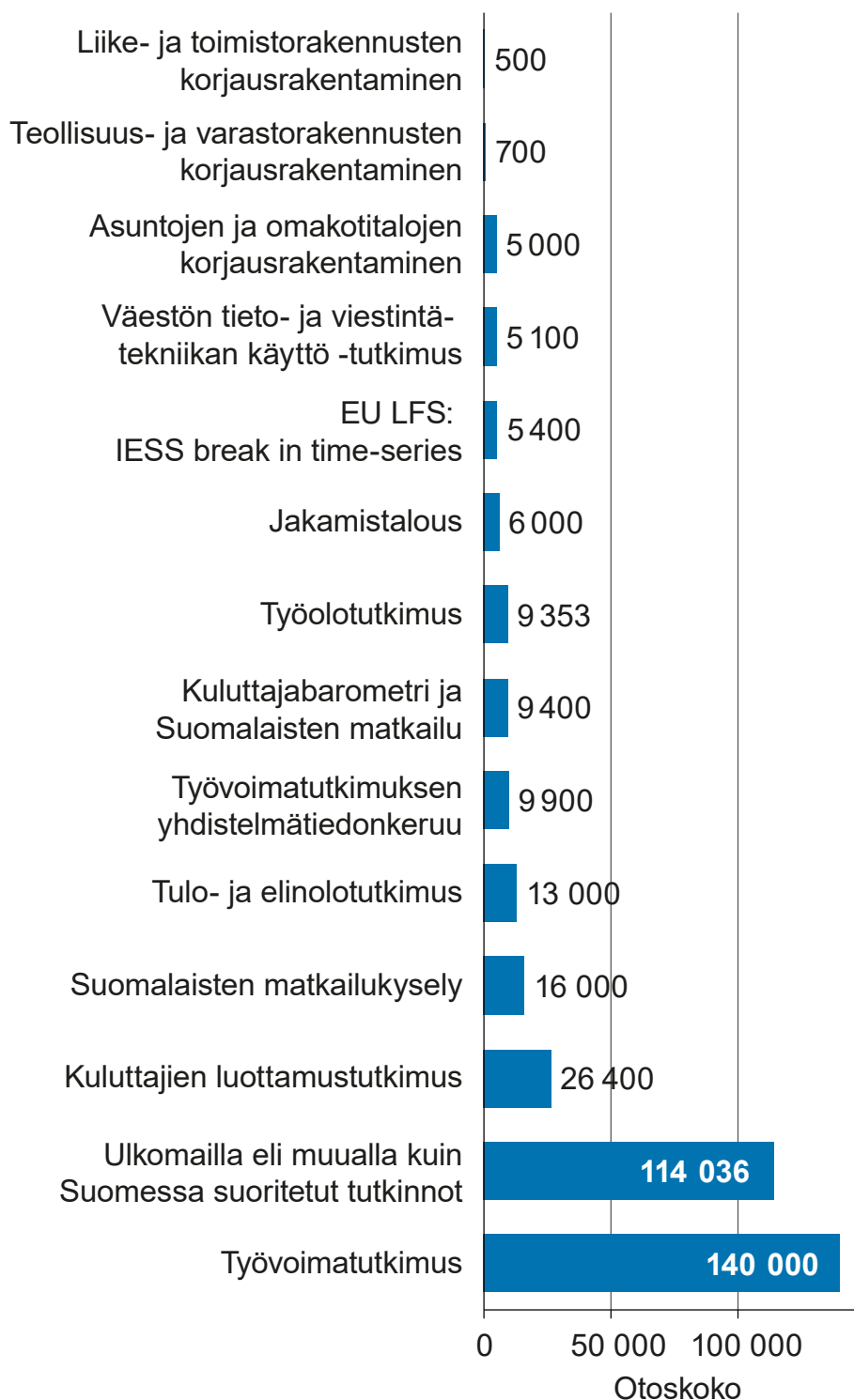
hieman ja on kuukausittain 2 200 henkilöä (kuluttajabarometrissa 2 350 henkilöä). Lisäksi jatkossa on käytössä paneeliasetelma: jokainen henkilö osallistuu tutkimukseen kaksi kertaa noin puolen vuoden sisällä, ja joka kuukausi puolet eli 1 100 on uusia otoshenkilöitä (kuluttajabarometrissa kokonaan uusi otos kuukausittain). Uudet menetelmät kuvataan tarkemmin kuluttajien luottamus -tilaston [laatuselosteessa](#).

Luettelo 1: Vuonna 2019 toteutetut suorat henkilö- ja kotitalouspohjaiset tiedonkeruut

1. Asuntojen ja omakotitalojen korjausrakentaminen
2. EU LFS: IESS break in time-series
3. Jakamistalous
4. Kuluttajabarometri (lakkautettu 2019)
5. Kuluttajien luottamustutkimus
6. Liike- ja toimistorakennusten korjausrakentaminen
7. Suomalaisten matkailu (lakkautettu 2019)
8. Suomalaisten matkailukysely
9. Teollisuus- ja varastorakennusten korjausrakentaminen
10. Tulo- ja elinolotutkimus
11. Työolotutkimus
12. Työvoimatutkimuksen yhdistelmätiedonkeruu
13. Työvoimatutkimus
14. Ulkomailla eli muualla kuin Suomessa suoritetut koulutukset
15. Väestön tieto- ja viestintätekniikan käyttö -tutkimus



Kuvio 6: Henkilö- ja kotitalouspohjaiset suorat tiedonkeruut vuonna 2019 otoskoon mukaan (N=15)

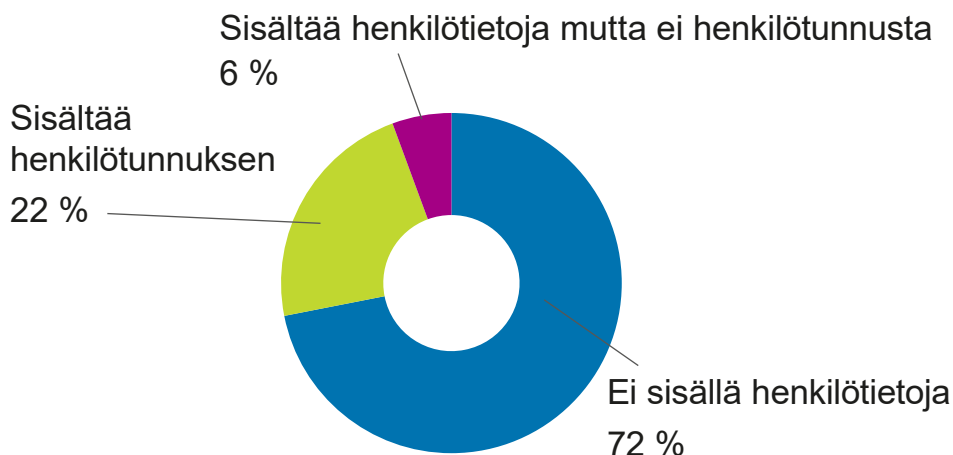


Vuoden 2019 aikana tehtyjen henkilö- ja kotitalouspohjaisten suorien tiedonkeruiden otoksissa oli yhteensä 360789 kohdetta.

Taulukko 1: Henkilö- ja kotitalouspohjaiset suorat tiedonkeruut vuonna 2019 otoskoon mukaan (N=15)

<b>Keruun nimi</b>	
Työvoimatutkimus	140 000
Ulkomailla eli muualla kuin Suomessa suoritettut tutkinnot	114 036
Kuluttajien luottamustutkimus	26 400
Suomalaisten matkailukysely	16 000
Tulo- ja elinolotutkimus	13 000
Työvoimatutkimuksen yhdistelmätiedonkeruu	9 900
Kuluttajabarometri ja Suomalaisten matkailu	9 400
Työolotutkimus	9 353
Jakamistalous	6 000
EU LFS: IESS break in time-series	5 400
Väestön tieto- ja viestintätekniikan käyttö -tutkimus	5 100
Asuntojen ja omakotitalojen korjausrakentaminen	5 000
Teollisuus- ja varastorakennusten korjausrakentaminen	700
Liike- ja toimistorakennusten korjausrakentaminen	500
<b>Yhteensä</b>	<b>360 789</b>

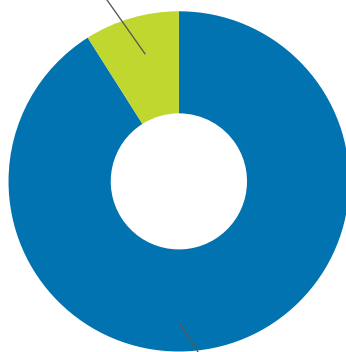
Kuvio 7: Henkilötieto ja suoralla tiedonkeruulla kerättävät tiedot vuonna 2019 (N=89)



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

Kuvio 8: Arkaluonteinen/erityisiin tietoryhmiin kuuluva tieto suorassa tiedonkeruussa vuonna 2019 (N=89)

Sisältää erityisiin tietoryhmiin kuuluvia henkilötietoja tai rikostuomioihin tai rikkomuksiin liittyviä tietoja  
9 %



Ei sisällä erityisiin tietoryhmiin kuuluvia henkilötietoja eikä rikostuomioihin tai rikkomuksiin liittyviä tietoja  
91 %

Taulukko 2: Haastattelujen määrät vuosina 2017–2019

	2017	%	2018	%	2019	%
Käyntihaastattelu	6 412	4	3 905	3	10 201	7
Puhelinhaastattelu*	130 535	96	108 840	97	130 017	93
<b>Yhteensä</b>	<b>151 331</b>	<b>100</b>	<b>126 364</b>	<b>100</b>	<b>144 903</b>	<b>100</b>

Tilastokeskus on suoran tiedonkeruun (N=89) yhteydessä tehtävien haastattelujen osalta rekisterinpitäjänä.

\*2017–2019 Kuluttajabarometrilla ja Suomalaisten matkailulla yhteinen otos

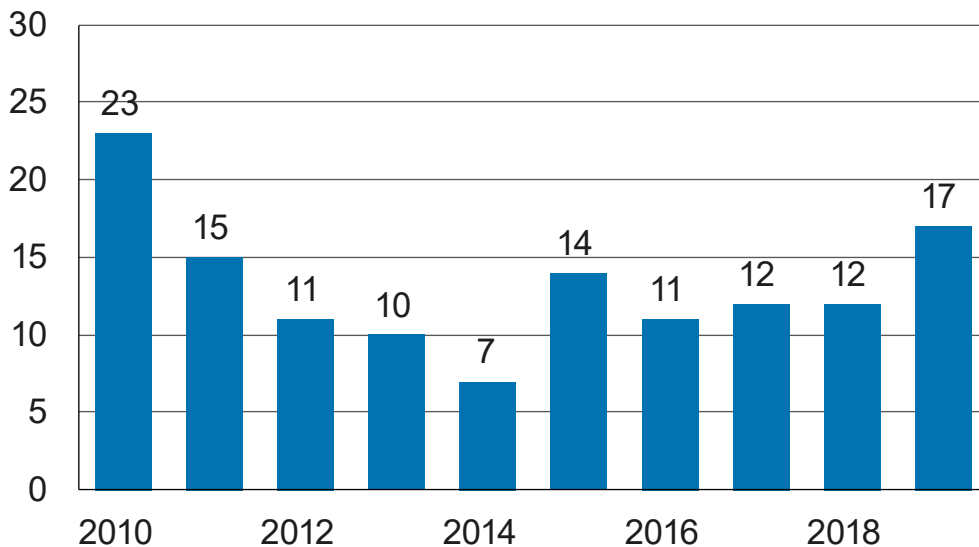
## 6.6 Tilaustutkimukset

Toimeksiantoina asiakkaille tehtävien tiedonkeruiden aiheet ja kokoluokka vaihtelevat paljon. Toimeksiantoina tehdään ainoastaan yhteiskunnallisesti tai tieteellisesti merkittäviä tutkimuksia. Osa keruista on säännöllisesti toistuvia, osa kertaluonteisia.

Tiedonkeruita tilaavat pääasiassa julkishallinto ja tutkimuslaitokset –tyypillisimmin ministeriöt, yliopistot sekä tutkimusorganisaatiot. Aineiston tietosisällöstä, tiedonkeruusopimuksesta ja otoshenkilöiden informoinnista riippuen anonymisoitu aineisto voidaan toimittaa toimeksiantajan lisäksi Yhteiskuntatieteelliseen tietoarkistoon (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.

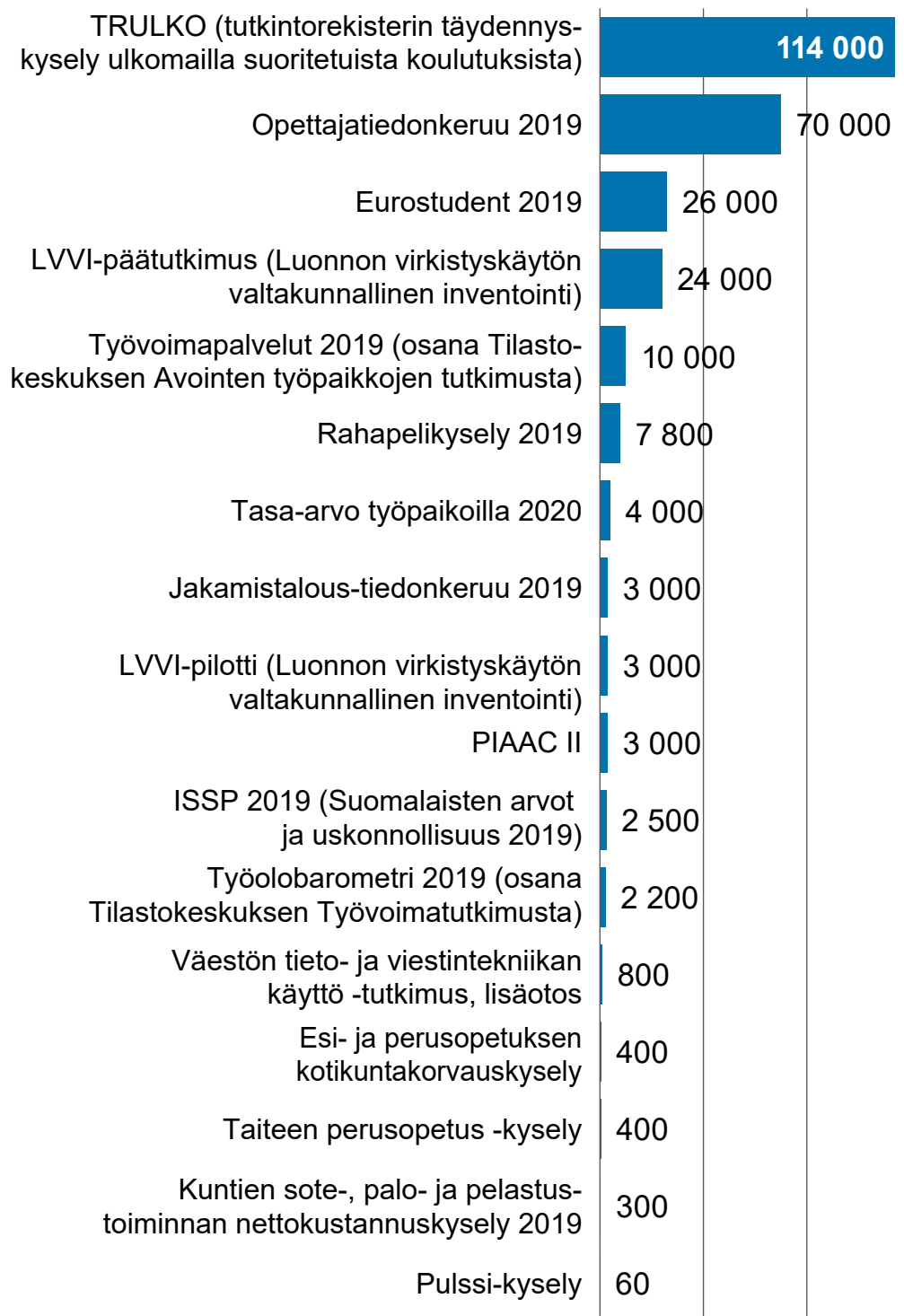
Kuvio 9: Toimeksiantojen lukumäärä vuosittain: asiakasrahoitteiset tilaustutkimukset vuosina, 2010–2019, kpl



Taulukko 3. Toimeksiantojen lukumäärä 2010–2019, kpl

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

Kuvio 10: Tilaustutkimukset vuonna 2019 otoskoon mukaan



Tilastokeskuksen tekemien asiakasrahoitteisten tilaustutkimusten otokset olivat vuonna 2019 yhteensä 271 460 yksikköä.

Taulukko 4. Tilaustutkimukset vuonna 2019 otoskoon mukaan, kpl

<b>Tilaustutkimus</b>	<b>Otoskoko, kpl</b>
Ulkomaille eli muualla kuin Suomessa suoritettavat koulutukset (suora tiedonkeruu)	114 000
Opettajatiedonkeruu	70 000
Eurostudent	26 000
Luonnon virkistyskäytön valtakunnallinen inventointi - päätutkimus	24 000
Työvoimapaalvelut osana Avointen työpaikkojen tutkimusta	10 000
Rahapelikysely	7 800
Tasa-arvo työpaikoilla	4 000
Jakamistalous (suora tiedonkeruu)	3 000
Luonnon virkistyskäytön valtakunnallinen inventointi - pilotti	3 000
PIAAC, Kansainvälinen aikuisten taitotutkimus -pätutkimus	3 000
Suomalaisten arvot ja uskonnollisuus	2 500
Työolobarometri (osana Työvoimatutkimusta)	2 200
Väestön tieto- ja viestintätekniiikan käyttö, lisäotos	800
Esi- ja perusopetuksen kotikuntakorvauskysely (suora tiedonkeruu)	400
Taiteen perusopetus -kysely	400
Kuntien sote-, palo- ja pelastustoiminnan nettokustannuskysely	300
Pulssi-kysely	60



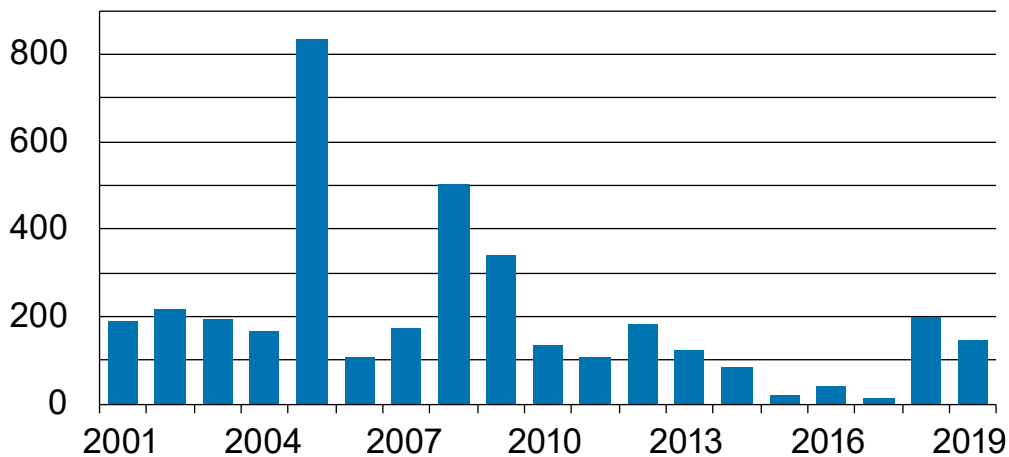
## 6.7 Asianhallinta ja arkistointi

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. muutujakuvaukset, käsitteet ja luokitukset. Tilastoaineistojen arkistointien vuosittainen toteuma on yksi mittari tulosityksiköiden tuloskorteissa.

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

Kuvio 11: Arkistoitujen tilastoaineistojen lukumäärät vuosina 2001–2019



Taulukko 5: Arkistoitujen tilastoaineistojen lukumäärät vuosina 2001–2019.

Vuosi	Toimeksiantojen lukumäärä, kpl
2001	188
2002	216
2003	195
2004	165
2005	835
2006	108
2007	175
2008	500
2009	340
2010	134
2011	106
2012	182
2013	122
2014	83
2015	20
2016	38
2017	11
2018	199
2019	146

## 6.8 Tutkimusaineistot

Tilastokeskuksessa tutkimusaineistojen pohjana ovat Tilastokeskuksen tilastointia varten saadut aineistot, joihin usein yhdistetään muiden tahojen tai tutkijan omia aineistoja.

Taulukko 6: Tutkimusaineistojen käyttöluvut 2017–2019

Tuoteluokka	Myönnetyt käyttöluvut 2017	Myönnetyt käyttöluvut 2018	Myönnetyt käyttöluvut 2019
Kuolemansyyaineistot	80	70	73
Voimassaolevien käyttö lupien jatkot ja laajennukset	144	135	222
Mikrosimulointi	21	12	14
Valmisaineistojen etäkäyttö	49	62	65
Räätälöityjen tutkimusaineistojen etäkäyttö	90	74	75
Räätälöityjen tutkimusaineistojen luovutus	29	26	9
Tunnisteellisena luovutettavat tutkimusaineistot	26	23	50
Muut (ainoastaan toisten viranomaisten aineistot)	5	4	1
<b>Yhteensä</b>	<b>444</b>	<b>406</b>	<b>509</b>

## 6.9 Kansallinen etäkäyttöpalvelu

Tilastokeskus voi tilastolain nojalla antaa käyttöoikeuden hallussaan oleviin, tilastointia varten kerättyihin salassa pidettäviin tietoihin tieteellistä tutkimusta ja yhteiskuntaoloja koskevaa tilastollista selvitystä varten.

Etätyöpöydällä työskentely vastaa omalla koneella työskentelyä, mutta järjestelmästä ulos siirrettävät tiedostot ja tulokset tarkistetaan tietosuojan osalta sekä tutkijan että Tilastokeskuksen toimesta.

Taulukko 7: Kansallinen etäkäyttöpalvelu 2017–2019

<b>Kansallinen etäkäyttöpalvelu</b>	<b>2017</b>	<b>2018</b>	<b>2019</b>
Etäkäytön organisaatiosopimukset	71	88	91
Etäkäyttöhankeet	163	196	219
Henkilöitä (mahdollisesti mukana useassa hankkeessa)	363	466	584
Käyttäjätunnuksia (hanke- ja henkilökohtaisia tunnuksia)	501	601	782

Tilastokeskus 

[www.tilastokeskus.fi](http://www.tilastokeskus.fi)  
[www.facebook.com/Tilastotohtori](https://www.facebook.com/Tilastotohtori)  
[www.twitter.com/tilastokeskus](https://www.twitter.com/tilastokeskus)