

## Tilastokeskuksen ja tutkijoiden välinen yhteistyö - Arvio tietotoimitusten aikatauluihin vaikuttavista tekijöistä ja nykytilanteen parantamisesta

### Tiivistelmä

Tilastokeskus luovuttaa empiirisiä tutkimusaineistoja tutkijoille. Näiden aineistojen avulla tehdään yhteiskuntapoliittista tutkimus- ja selvitystyötä sekä akateemisia opinnäytteitä. Tilastokeskuksen tutkimusaineistojen keskeisiä käyttäjiä ovat erilaiset yliopistot, yliopistolliset sairaalat ja sairaanhoitopiirit sekä tutkimuslaitokset, kuten VATT, EK, PT ja Pellervo.

Tilastokeskuksella on ollut vaikeuksia toimittaa tutkijoille heidän tarvitsemiaan tutkimusaineistoja kohtuullisessa määräajassa. Ongelman taustalla on nähtävissä kolme tekijää:

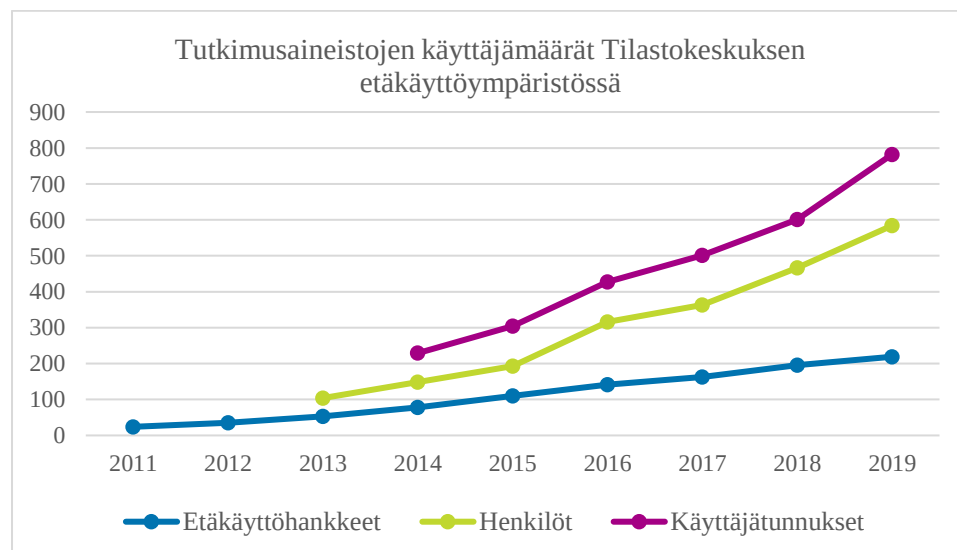
1. Tutkimusaineistojen kysyntä ja käyttäjämäärät ovat kasvaneet merkittävästi. Vuonna 2019 Tilastokeskuksen aineistoja käytti liki 600 tutkijaa, kun vuonna 2016 heitä oli reilut 300.
2. Tutkimusaineistojen käyttöön liittyy tietoturva-vaatimuksia, jotka tekevät aineiston käyttölupapyyntöä käsittelystä, tarvittavan tutkimusaineiston muodostamisesta ja tutkimuksen tulosten tarkistamisesta työlästä.
3. Tutkijoiden aineistopalveluun osoitetut resurssit eivät ole juuri muuttuneet edellisten vuosien aikana. Vuonna 2019 aineistopalveluun käytettiin likimain yhtä paljon aikaa kuin vuonna 2016.

Näihin haasteisiin on vastattu käynnistämällä Tilastokeskuksen sisäisiä kehittämistoimia, perustamalla keskeisiä käyttäjiä sisältävä työryhmä hahmottelemaan ehdotuksia toiminnan parantamiseksi sekä lisäämällä palvelun resursointia.

### Johdanto

Empiiristen tutkimusaineistojen tuottaminen tutkijoille on Tilastokeskuksen merkittävä ydintoiminto, joka on keskitetty yhdelle vastuualueelle vuodesta 2010 alkaen. Osa toiminnasta on asiakas- ja osa budjettirahoitteista.

Tilastokeskus vastaanottaa vuosittain noin 500 aineiston käyttölupahakemusta ja valmistaa noin 200-250 aineistotoimitusta vuodessa. Aineistojen kysyntä ja käyttäjämäärät ovat kasvaneet nopeasti.



Tilastokeskuksen aineistopalveluun käytetyt resurssit ovat puolestaan pysyneet miltei samansuuruisina. Tämän on aiheuttanut haasteita vastata kasvaneeseen kysyntään.

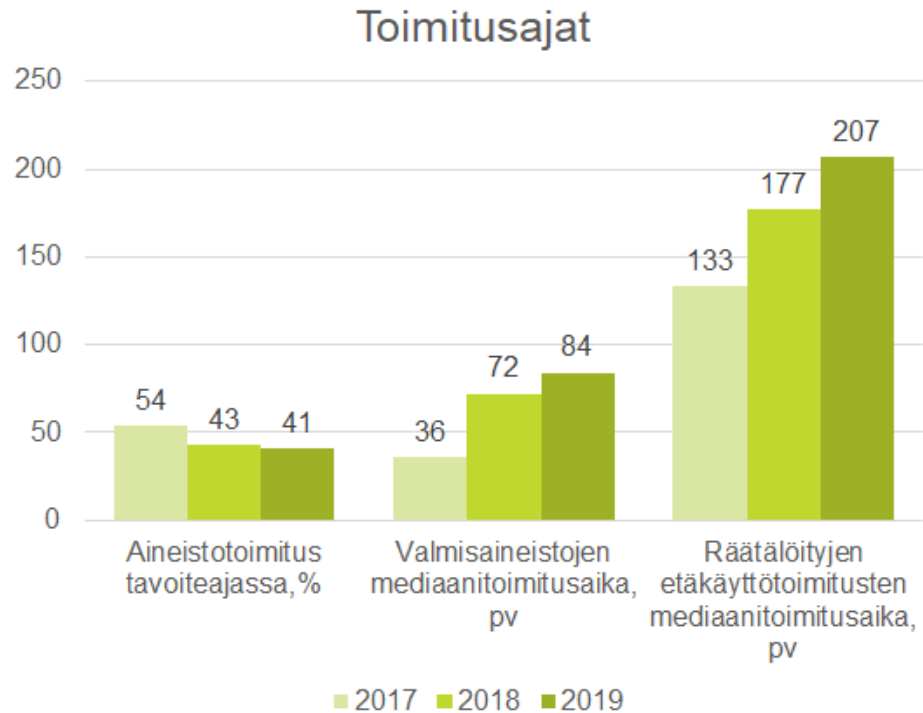
Tietotoimituksiin liittyvä tietoturva vaatimukset tekevät prosessista nykyisellään työlää. Tilastokeskuksessa arvioidaan kukin tutkimussuunnitelma ja tutkimushankkeen tarvitseman aineiston kattavuus erikseen. Tyypillisesti tutkijat tekevät laskentansa etäkäyttöympäristössä ilman, että tarvittava tutkimusaineisto poistuu Tilastokeskuksen palvelimelta. Näin saadut tulokset tarkistetaan Tilastokeskuksessa ja varmistetaan, ettei tuloksissa ole yksilöitävissä olevia henkilötietoja. Nykyisellään tämä tarkistusprosessi on manuaalinen ja työläs, minkä vuoksi siihen liittyviä kehittämistoimia on ryhdytty panemaan toimeen.



Tutkijoille toimitetuista aineistoista vain 41% on saatu valmiiksi tavoiteajassa. Aineistojen kysynnän kasvu on heijastunut suoraan toimitusaikojen kasvuna.

Tietotoimitukset voidaan jakaa karkeasti kahteen ryhmään. Valmisaineistoja on laadittu suosituimmista tutkimusaiheista. Nämä aineistot ovat nopeasti käytettävissä ja ovat sisällöltään monipuolisia. Yksi valmisaineisto voi palvella useampaa tutkimuskysymystä. Rääätälöidyt aineistot, toisin kun valmisaineistot, saadaan yhdistelemällä monista eri lähteistä koostuvista tiedoista, joista osa voi olla tutkijan itsensä keräämiä. Kukin rääätälöity aineisto on tutkimushankekohtainen ja laaditaan tutkijan toiveiden mukaan, minkä vuoksi niiden toimitusajat ovat pidemmät.

3.3.2020



## Tilastokeskuksen sisäinen kehittäminen

Tilastokeskuksessa on syksyllä 2019 laadittu kehittämissuunnitelman, jossa listatuilla toimenpiteillä saadaan tehostettua tietotoimituksia tutkijoille. Taustalla on Tilastokeskuksen ja sen sidosryhmien kanssa käyty keskustelu aineistotoimitusten viivästymisestä.

Kehittämissuunnitelman keskeisiä elementtejä ovat:

- Aineistotoimitusta koskevien prosessien automatisointi. Joidenkin aineistotyyppien osalta on luotu poimintakoodi, jolla voidaan vähentää näiden aineistojen muodostamiseen vaatimaa työaika.
- Tutkijat suorittavat laskentansa tietoturvalisessä etäkäyttöympäristössä. Tilastokeskuksen asiantuntijat tarkistavat täällä laskettuja tuloksia mahdollisten tietoturvapoikkeamien varalta. Tämän manuaalisen tarkistamisen korvaaminen vähemmän työintensiivisellä, koneellisella tarkastamisella, on pitkän aikavälin tavoite.
- Tilastokeskus on laatinut suosituimmista tutkimusaiheista niin sanottuja valmisaineistoja, joita päivitetään säännöllisesti tuoreilla tiedoilla. Näiden valmisaineistojen luovuttaminen tutkijoiden käyttöön on verrattain suoraviivaista. Tavoitteena on kehittää, ylläpitää ja lisätä näiden valmisaineistojen määrää.
- Tavoitteena on uudistaa tutkimusaineistoja koskevaa hakulomaketta. Pitkän aikavälin tavoitteena siirtyminen sähköiseen käyttöluupahakemukseen. Aineistojen käyttöluupahakemusten tuki perustuisi jatkossa enemmän kirjallisiin ohjeisiin ja vähemmän henkilökohtaiseen neuvontaan.
- Käyttöluupapäätösten käsittelyn suoraviivaistaminen siirtymällä mm. sähköisiin allekirjoituksiin.
- Jatkossa asiakkaita velvoitettaisiin vastaamaan lisätietopyyntöihin määritellyssä ajassa, jotta heidän hakemuksensa pysyy aktiivisena. Lisäksi asiakasta velvoitettaisiin tarkastamaan heille toimitettu aineisto

reklamointiajan sisällä. Mahdollisia korjauksia ei suoriteta jonon ohi ilmaiseksi tämän reklamaatioajan päätyttyä.

Osa tässä listatuista kehittämistoimista edellyttää taloudellisia ja ajankäyttöllisiä investointeja.

## Julkishallinnon aineistojen hyödyntämistä pohtiva ryhmä

Valtionvarainministeriö (VM) ja opetus- ja kulttuuriministeriö (OKM) ovat asettaneet työryhmän, jonka tehtävänä on laatia suunnitelma julkishallinnon aineistojen hyödyntämisen parantamiseksi. Ryhmän toimikausi on 1.3.-31.5.2020 ja siihen kuuluu edustajia Valtionvarainministeriöstä, Valtion taloudellisesta tutkimuskeskuksesta (VATT), Helsinki Graduate School of Economicsista (HGSE) ja Tilastokeskuksesta.

Työryhmä hahmottelee järjestelmää, jossa rekistereihin kerääntyvät tiedot saadaan tehokkaasti politiikan selvitys- ja tutkimuskäyttöön.

Työryhmän tehtävänä on luoda kokonaiskuva nykytilanteesta sekä määrittää ja aikatauluttaa toimenpiteet, jotka toteuttamalla päästään tilanteeseen, jossa aineistot ovat tehokkaaseen analyysiin soveltuvassa muodossa, lupaprosessit yksinkertaisia ja aineistojen käyttöpalveluita tarjoavien organisaatioiden resursointi ja palveluiden hinnoitteluperiaatteet kunnossa. Ryhmän toimeksiannossa todetaan, että järjestelmän tulee olla sellainen, että siihen voidaan joustavasti liittää uusia aineistotoimittajia. Ryhmä ottaa tietosuojan huomioon toimenpiteiden suunnittelussa.

## Tutkijapalveluiden resurssit

Tutkijoille osoitettujen tietotoimitusten nopeuttamiseksi tutkijapalveluihin on alkuvuodesta 2020 palkattu määräaikaista työvoimaa purkamaan tietotoimituksiin syntynyttä ruuhkaa.

Keskustelut tutkijapalvelun pysyvistä resurssilisäyksistä on käynnistetty. Osa Tilastokeskuksen sisäisistä kehittämistoimista edellyttää lisärahoitusta ja -henkilöstöä.

## Johtopäätökset

Tutkijapalvelun perustana on tilastolain (23.4.2004/280) 3§, jossa veloitetaan Tilastokeskusta seuraavasti: ”Lain tarkoituksena on myös edistää tilastotarkoituksia varten kerättyjen tietojen käyttöä tieteellisissä tutkimuksissa ja yhteiskuntaoloja koskevissa tilastollisissa selvityksissä.”

Tietoon perustuva päätöksenteko edellyttää nopeaan reagointiin kykenevää tutkimustoimintaa ja nopeasti saatavilla olevia tutkimusaineistoja. Yhteiskuntapoliittiset tutkimus- ja selvitystehtävät ovat muodostuneet vakiintuneeksi osaksi hallitusohjelmia, ja tutkimusrahoitusta on suunnattu entistä enemmän strategiaan, nopeaa yhteiskuntapoliittista reagointia vaativiin aiheisiin.

Tutkimusaineistojen käyttäjien ja kysynnän kasvu nähdään Tilastokeskuksessa hyvänä asia. Nyt käynnissä olevat kehittämistoimenpiteet, (määräaikaiset) resurssilisäykset sekä VM:n asettaman työryhmän ehdotukset vievät toimintaa parempaan suuntaan.