

## Summatasaisen tiedonkeruun selvitystarve

### Taustaa

Summatasaisen tiedonkeruun tarpeet tunnistettiin alun perin TOR-hankkeessa, jonka lopputuloksena syntyi nykyinen Koski-tietovaranto. Hankkeen keskeinen tavoite oli hankkia opetuksesta ja koulutuksesta tiedot henkilötasolla, kun ennen vuotta 2020 tiedot kerättiin Tilastokeskuksessa erillistiedonkeruulla. Vuonna 2020 tiedonkeruun oli tarkoitus muuttua Koski-pohjaiseksi eli tiedot poimittaisiin tilastoajankohdalta 20.9. Koskesta.

Koskeen tallennettavista tiedoista säädetään laissa valtakunnallisista opinto- ja tutkintorekistereistä (844/2017). Koski-tietovarannon käyttöönoton kanssa samoihin aikoihin astui voimaan EU:n yleinen tietosuojasetus, joka muutti henkilötietojen säädöserustaa merkittävästi. Tästä johtuen myös Koski-tietovarannon tiedonkeruun osalta Opetushallitus toteutti vaikutustenarviointia, jonka tuloksena 4.6.2020 Opetushallituksen juristitiimi linjasi, ettei erityisiin henkilötietoryhmiin liittyviä tietoja voida enää kerätä KOSKI-tietovarantoon, ellei ko. tietojen keräämiselle ole opinto- ja tutkintorekisterilain 4 §:n 1–2 momentin mukaista käyttötarkoitusta.

Tukimuotojen tietoa ei käytetä rahoituksenlaskennan perusteena vaan sen pääasiallinen tarvitsija on Tilastokeskus. Opinto- ja tutkintorekisterilaki ei sisällä perustetta tallentaa rekisteriin henkilötietoja vain tilastointitarkoitusta varten. Tämän tulkinnan johdosta KOSKI-tietovarannon tietomallista poistettiin osioita, joihin tallennettu tieto katsotaan erityisiksi henkilötiedoiksi. Opetushallitus tiedotti opetuksen järjestäjiä, että edellä mainittujen tietomallista poistettujen tietojen siirto tulee lopettaa.

Kosken tietomallista löytyvät kaikki erityisopetustilaston muodostamiseen tarvittavat tiedot, mutta muita kuin rahoituslaskennan perusteena käytettäviä tietoa ei voida kerätä.

- Esi-, perus- ja lisäopetuksen tehostetun tuen oppilasmäärät 20.9.
  - Tiedot ilmoitetaan ryhmissä esiopetus, vuosiluokat 1-6, vuosiluokat 7-9 ja lisäopetus ja jako sukupuolen mukaan
- Osa-aikaista erityisopetusta lukuvuoden aikana saaneet esi-, perus- ja lisäopetuksen oppilaat (yhteensä ja tytöt), jotka olivat peruskoulunne kirjoilla 20.9.
- Peruskoulujen esi-, perus- ja lisäopetuksen oppilaat, joilla on erityisen tuen päätös 20.9., toteutuspaikan mukaan
  - Tiedot ilmoitetaan ryhmissä esiopetus, vuosiluokat 1-6, vuosiluokat 7-9 ja lisäopetus
  - Toteutuspaikkatieto ilmoitetaan seuraavissa ryhmissä:
    - Opetus on kokonaan erityisryhmissä tai -luokassa
    - Opetuksesta 1-19 % on yleisopetuksen ryhmissä
    - Opetuksesta 20-49 % on yleisopetuksen ryhmissä
    - Opetuksesta 50-79 % on yleisopetuksen ryhmissä
    - Opetuksesta 80-100 % on yleisopetuksen ryhmissä

- Peruskoulujen esi-, perus- ja lisäopetuksen oppilaat, joilla on tehostetun tuen oppimissuunnitelma tai erityisen tuen päätös 20.9., tukimuotojen mukaan (oppilaan oppimissuunnitelmaan tai HOJKS:iin tai näihin liittyviin asiakirjoihin kirjatut tukimuodot). Oppilaalla voi olla useampia tukimuotoja.
  - Tiedot kerätään Kosken koodiston mukaisesti:
    - Osa-aikainen erityisopetus
    - Tukiopetus
    - Avustajapalvelut
    - Tulkitsemispalvelut
    - Erityiset apuvälineet
    - Sisältää
- Perusopetuksen lykkääntymisen tiedot Koulun oppivelvollisten määrä, joiden perusopetuksen aloittamista on tänä syksynä siirretty

### **Vuosien 2020 ja 2021 erillistiedonkeruu**

Opetus- ja kulttuuriministeriö lähti valmistelemaan Kosken tiedonkeruun ulkopuolelle jäävien tietojen osalta erillistä tiedonkeruuta. Opetus- ja kulttuuriministeriö teki päätöksen tilata puuttuvien erityisopetuksen tilastointiin tarvittavien tietojen erillistiedonkeruun Tilastokeskukselta. Tiedot tarvitaan hallitusohjelman toimeenpanon tueksi ja jotta erityisopetuksen tiedonkeruun aikasarjojen jatkuvuus voidaan taata. Opetus- ja kulttuuriministeriö ja Tilastokeskus lähettivät erillistiedonkeruusta tiedotteen opetuksen järjestäjille 18.9.2020.

Erillistiedonkeruussa kerätyt puuttuvat tiedot yhdistetään Kosken kerättäviin tuen tietoihin, jotka voidaan tallentaa Koskeen, sillä niitä käytetään valtionosuusrahoituksen laskentaperusteena. Vuosien 2020 ja 2021 erillistiedonkeruut sisällytettiin Tilastokeskuksen Tietopalvelusopimukseen. Opetus- ja kulttuuriministeriö sopii erillistiedonkeruun toteutuksesta vuosittain ja alustavasti tästä on sovittu vuoden 2023 osalta.

### **Asian edistämisen vaiheet**

Opetus- ja kulttuuriministeriö on viestittänyt Tilastokeskukselle selvittävänsä erillistiedonkeruulle vaihtoehtoisia toteutustapoja. Vuoden 2022 OKM-OPH-TK –johtajatapaamisessa Tilastokeskus ilmaisi toiveensa lopettaa summatasoiset erillistiedonkeruut ja pyysi opetus- ja kulttuuriministeriötä selvittämään vaihtoehtoisia toteuttamistapoja.

Opetus- ja kulttuuriministeriö on käynyt Opetushallituksen kanssa keskusteluja vuodesta 2021 saakka. Alun perin selvitettiin mahdollisuutta toteuttaa Koski-palvelun osana summatasoisien tiedonkeruun toteuttamistapoja. Vuoden 2022 aikana Opetushallituksen KOSKI-tiimi linjasi, että tietoarkkitehtuurin näkökulmasta organisaation toimintaprofiilia koskevan summatiedon paikka ei ole KOSKI-tietovarannossa, vaan tiedonkeruut summatasoisien tiedon osalta tulee toteuttaa erillisellä ratkaisulla. Osana tulossopimusta vuodelle 2023 sovittiin summatasoisien tiedonkeruun toteuttamisen selvityshanke, jossa nyt tunnistetuille summatasoisien tiedon keruulle etsitään toteuttamistapa ja suunnitellaan sen hankkeistaminen kustannusarvioineen.

On selvää, että jatkossakaan opetushallinnon tietotarpeita ei pystytä toteuttamaan pelkästään henkilö pohjaisilla tietovarannoilla vaan summatasoisia tietoja tarvitaan varhaiskasvatuksen- ja koulutuksen järjestäjiltä. Näitä varten tarvitaan pysyvä ja kestävä ratkaisu, joka helpottaa viranomaisten tiedonhankintaa ja vähentää järjestäjiin kohdistuvia eri viranomaisten tietopyyntöjä.

### **Muita tunnistettuja summatasoisien tiedonkeruun tarpeita**

Osana summatasaisen tiedonkeruun kehittämisen valmistelua on tunnistettu myös muita tietoja, joita tarvitaan ennakointi-, arviointi- ja selvitystyön toteutukseen. Haettavalla ratkaisulla tultaisiin toteuttamaan myös näitä tiedonkeruita:

- Tietosuojalainsäädännön muutoksen vuoksi lakkautettiin katsomusaineiden oppimäärätietojen keruu vuonna 2021. Tällä hetkellä Koski-tietovarannosta saa tiedon oppilaiden oppiainetiedoista (esim. uskonto/elämäkatsomustieto) mutta ei tietoa minkä katsomusaineen (ev-lut, islam, jne) opetukseen osallistuu. Katsomusaineiden oppimäärätietoja ei myöskään kerätä tilapäisellä tiedonkeruulla vaan tietoaineistot ovat tällä hetkellä puutteellisia siltä osin ja tilastoja tietyn katsomusaineen oppilasmääristä ei ole mahdollista muodostaa.
- Sairaalakoulujen hoitopäivien määrää ei tällä hetkellä ole saatavilla kootusti. Hoitopäivien määrää tarvitaan valtiosuuslaskennassa, jonka lisäksi tietosisältöä tarvitaan yleiseen toimialan tietopohjan vahvistamiseen.

### **Mahdollisesti jatkossa tarpeelliset tietotarpeet**

Marinin hallituskaudella osana oppivelvollisuuslain toimeenpanoa opetuksen järjestäjille muodostui velvollisuus antaa oppilaalle tehostettua henkilökohtaista oppilaanohjausta. Tehostetussa henkilökohtaisessa oppilaanohjauksessa oppilaalle annetaan henkilökohtaista ja muuta oppilaanohjausta sekä laaditaan henkilökohtainen jatko-opintosuunnitelma. Tehostettua henkilökohtaista oppilaanohjausta annetaan 8. ja 9. vuosiluokalla. Sivistysvaliokunta kiinnitti lausunnossaan (SiVM 15/2020) oppivelvollisuuslain mukaisen uudistuksen seurannan ja arvioinnin toteuttamiseen. Tehostetun henkilökohtaisen oppilaanohjauksen seuraamiseksi olisi perusteltua seurata oppilasmääriä. Lausunnossaan sivistysvaliokunta lisäksi toteaa, perustuslakivaliokunnasta poiketen, ettei valtakunnalliseen opetuksen ja koulutuksen tietovarantoon tulisi tallentaa arkaluonteisia tietoja.

Perusopetuslaissa (26 §) opetuksen järjestäjälle on asetettu veloitteeksi seurata perusopetukseen osallistuvan oppilaan poissaoloja ja ilmoittaa luvattomista poissaoloista oppilaan huoltajalle tai muulle lailliselle edustajalle. Kuitenkin, selvitysten mukaan arviolta vain 40 % opetuksen järjestäjistä seuraa poissaoloja. Koulukunnossa –hankkeessa on kehitetty hallituskauden aikana poissaolotietojen seuraamista ja kehitetty poissaolotietojen yhtenäisen kirjaamisen käytäntöjä. Hankkeessa 16 opetuksen järjestäjää Varsinais-Suomessa pilotoi yhtenäisen tiedon keruun toteutusta vuosien 2022 – 2023 aikana.

### **Summatasaisen tiedonkeruun säädöspohja**

Valmistelun yhteydessä on selvitetty summatasaisen tiedonkeruun säädöspohjaa. Summatason tiedonkeruun perusteena voisi tietyin ehdoin toimia perusopetuslain 41.2 §:

Opetuksen järjestäjän tulee pyynnöstä toimittaa valtion opetushallintoviranomaisille niiden määräämät koulutuksen arvioinnin, kehittämisen, tilastoinnin ja seurannan edellyttämät tiedot. Kuntien ja kuntayhtymien taloustietojen toimittamisesta säädetään lisäksi kuntalain (410/2015) 121 a §:ssä.

Opetus- ja kulttuuriministeriö sekä Opetushallitus ovat valtion opetushallintoviranomaisia ja pykälän perusteella voivat pyytää opetuksen järjestäjiä toimittamaan tarvitsemiaan tietoja käyttöönsä. Tietojen käyttötarkoitus olisi tältä osin tilastojen muodostaminen tutkimus-, arviointi-, ennakointi- ja suunnittelutarkoituksiin. Tietojen keruun toteuttaminen edellyttäisi Opetushallitukselta määräyksen antamista, jossa määritellään järjestäjiltä tarvittavat tiedot sekä ohjeistetaan tietojen oikeamuotoiseen luovutustapaan ja tiedon sisältöön.

Kyseessä olevan pykälän nojalla ei voida kerätä henkilötietoa, jolloin niin tiedonkeruun ohjeistuksessa kuin sen toteutuksessa tulee poistaa välillisen tunnistamisen riski. Tämä tulee toteuttaa erityistä huolellisuutta noudattaen silloin kun kyseessä on erityisiin henkilötietoryhmiin kuuluvat tai salassapidettävät tiedot. Esimerkiksi katsomusopetuksen tiedot kuuluvat erityisiin henkilötietoryhmiin ja ovat myös salassapidettäviä (JulkL 24.1 § 32 k, henkilökohtaiset olot).

Mikäli summatasoiseen tietoon muodostuu välillisen tunnistamisen riski, sitä käsitellään henkilötietona, joka puolestaan edellyttäisi säännöksiä henkilötietojen käsittelystä ja mahdollisesti uudesta rekisteristä säätämistä. Välillisen tunnistamisen riskin minimoimiseen on jo olemassa käytäntöjä, joissa niin tiedonkeruuta kuin sen raportointia ei toteuteta esimerkiksi jos tietyn tietojoukon koko on alle 5 henkilöä järjestäjää kohden. Noudattamalla näitä käytäntöjä sekä tekemällä kerättävien tietojoukkojen osalta riskiarvioita, on välillisen tunnistamisen riski suurella varmuudella vältettävissä