# Poimintakuvaus ja arvioidut kustannukset

Tämä lomake on tarkoitettu hakijoiden, rekisterinpitäjien ja Findatan käyttöön.

* Hakija voi pyytää rekisterinpitäjiltä neuvontaa suunnitellun **poiminnan toteutettavuudesta** ja poimintaan liittyvistä **kustannuksista**, ennen kuin jättää tietolupahakemuksen tai tietopyynnön Findatalle. Findata vahvistaa arvioidut poimintakustannukset rekisterinpitäjiltä vielä hakemuskäsittelyn aikana.
* Rekisterinpitäjä voi antaa hakijoille toisiolain mukaista **neuvontaa rekisterien tietosisällöistä** ja **soveltuvuudesta** toisiolain mukaiseen käyttötarkoitukseen.
* Findata pyytää toisiolain nojalla rekisterinpitäjiltä maksutonta arviota poiminnan toteuttamisesta ja enimmäiskustannuksista.

Lomakkeen käyttäminen **vähentää** jälkikäteen tehtäviä **poimintojen täydennyksiä ja korjauksia**. Ennakkotiedot poimintaan liittyvistä kustannuksista helpottavat hankkeiden rahoituksen suunnittelua.

## **Ohje hakijalle**

**Täytä lomakkeen kohdat 1–3**. Tunnistat kohdat myös sinisestä taustaväristä.

Ennen kuin lähetät hakemuksen Findatalle:

1. Täytä yksi lomake per rekisterinpitäjä.
2. Lähetä täytetty lomake sille rekisterinpitäjille, jonka rekistereihin poiminta kohdistuu.
3. Rekisterinpitäjä antaa tarvittaessa neuvontaa poiminnan sisältöön ja valittuihin muuttujiin liittyen, ja palauttaa lomakkeen täydennettynä hakijalle.
4. Liitä täydennetty lomake Findatalle jätettävään hakemukseen Word-tiedostomuodossa.

*Vinkki: jos lähetät lomakkeen usealle eri rekisterinpitäjlle, täytä ensin lomakkeen kohdat 1 ja 2. Tee tämän jälkeen lomakkeesta oma kopio kullekin rekisterinpitäjälle ja täytä niihin kohta 3.*

## **Ohje rekisterinpitäjälle**

**Täytä lomakkeen kohdat 4–6** ja tarvittaessa kohtaan 1 diaarinumero tai muu tunnistenumero. Tunnistat kohdat myös oranssista taustaväristä.

Kun saat hakijan tai Findatan täyttämän lomakkeen:

1. Ota kantaa siihen, onko tällä lomakkeella kuvattu poiminta mahdollista toteuttaa. Mikäli ei, neuvo hakijaa muokkaamaan poimintaa toteuttamiskelpoiseksi.
2. Arvioi, kuinka monena tiedostona aineisto toimitetaan ja mikä tiedostojen arvioitu koko on.
3. Arvioi, paljonko tällä lomakkeella kuvattu poiminta enimmillään maksaisi.
4. Palauta täydennetty lomake Word-tiedostomuodossa.
   * Jos olet saanut lomakkeen hakijalta, lähetä lomake hakijalle.
   * Jos olet saanut lomakkeen Findatalta, lähetä lomake Findatalle.

## **1 Hankkeen lyhyt kuvaus** *[hakija täyttää]*

Tämän kohdan tarkoituksena on antaa rekisterinpitäjälle tietoa hankkeen ja poimintojen kokonaisuudesta.

Alla olevan **taulukon tulee olla aina samansisältöinen** riippumatta siitä, mille rekisterinpitäjälle lomake lähetetään.

|  |  |
| --- | --- |
| Tarvittava tieto | Täytä tiedot tähän sarakkeeseen |
| Hankkeen nimi |  |
| Lyhyt kuvaus hankkeesta | Rekisterinpitäjät tarvitsevat taustatietoa hankkeesta poimintojen toteutettavuuden arvioimiseksi. Kuvaile muutamalla lauseella (enintään n. 70 sanaa / 500 merkkiä). |
| Hakija / hakijaorganisaatio | Luvanhakijalla tarkoitetaan tietolupahakemuksessa tai tietopyynnössä nimettävää organisaatiota tai henkilöä, jolle tietolupaa haetaan tai kenelle tietopyyntöpäätöksellä luovutettava aineisto toimitetaan. |
| Hakijan yhteyshenkilö | Kirjoita yhteyshenkilön nimi, sähköposti ja puhelin. Yhteyshenkilöllä tarkoitetaan henkilöä, joka vastaa hakemusta ja poimintaa koskeviin tiedusteluihin. |
| Laskun maksaja | Kirjaa tähän seuraavat tiedot: Laskun maksajan nimi, ALV-tunnus ja kotimaa.  Laskun lopullinen maksaja vaikuttaa siihen, lisätäänkö poimintakustannuksiin arvonlisävero vai ei. Findata toimii vain laskun välittäjänä ja käsittelee laskua kirjanpidossaan läpikulkueränä. Siten rekisterinpitäjä määrittää arvonlisäverovelvollisuuden laskun lopullisen maksajan perusteella. |
| Toisiolain mukainen käyttötarkoitus | Valitse seuraavista: **Tietoluvat:** tieteellinen tutkimus; tilastointi; opetus; viranomaisen suunnittelu- ja selvitystehtävä **Tietopyynnöt:** edellä olevien lisäksi kehittämis- ja innovaatiotoiminta; tietojohtaminen; sosiaali- ja terveydenhuollon viranomaisohjaus ja -valvonta |
| Poiminnan toistuvuus | Poimitaanko tiedot vain yhden kerran, vai esimerkiksi vuosittain tiettyyn vuoteen asti? |
| Aiemmat tähän poimintaan liittyvät luvat | Jos hankkeeseen tai poimintoihin liittyy aiempia/vanhoja lupia, listaa tähän alla pyydetyt tiedot luvista. Näin rekisterinpitäjät tietävät, mihin poiminta liittyy.   * Luvan myöntäjä, pvm, diaarinumero * Luvan myöntäjä, pvm, diaarinumero |
| Rekisterinpitäjän diaarinumero |  |

## **2 Poimintajärjestys** *[hakija täyttää]*

Kuvaa alla olevaan taulukkoon aineiston muodostuksen vaiheet, jos aineiston muodostus sisältää useita poimintoja (esim. kohdejoukko, verrokit ja aineisto poimitaan eri rekisterinpitäjiltä).

Täydennä rekisterikohtaiset poiminnat omiin taulukoihinsa vasta kohdassa 3.

**Poimintajärjestyksen tulee olla aina samansisältöinen** riippumatta siitä, mille rekisterinpitäjälle lomake lähetetään.

|  |  |  |
| --- | --- | --- |
| Poiminnan vaihe nro | Vaiheen kuvaus | Poimintavaiheen enimmäisaika (arkipäiviä) |
| 1 | Tällä poiminnalla tunnistetaan kohdejoukkoon kuuluvat henkilöt. Kuvaa, minkä rekisterinpitäjän rekisteristä kohdejoukko tunnistetaan ja mikä kohdejoukon arvioitu koko tai suuruusluokka on.  Jos kohdejoukkoa ei tarvitse erikseen poimia, vaan sinulla on se jo olemassa, kuvaa tähän mistä se on peräisin.  Esim. Findata lähettää kohdejoukkoa koskevan aineistopyynnön THL:lle ja THL toimittaa kohdejoukon Findatalle. | 30 |
|  | Findata esikäsittelee kohdejoukon, ja poistaa tietojensa käsittelyn Findatassa kieltäneet henkilöt joukosta. | 5 |
| 2 | Tässä vaiheessa poimitaan mahdolliset verrokit ja/tai sukulaiset.  Esim. Findata lähettää kohdejoukon henkilötunnukset DVV:lle verrokkipoimintaa varten. DVV toimittaa verrokit Findatalle. | 30 |
|  | Findata esikäsittelee verrokit ja/tai sukulaiset, ja poistaa tietojensa käsittelyn Findatassa kieltäneet henkilöt joukosta. | 5 |
| 3 | Tässä vaiheessa poimitaan lopullinen aineisto. Kuvaa, mistä rekistereistä kohdejoukolle poimitaan tietoja.  Esim. Findata lähettää kohdejoukon ja verrokkien henkilötunnukset THL:lle ja Kelalle aineistopoimintaa varten. Kela ja THL toimittavat aineistot Findatalle. | 30 |
| 4 | Findata esikäsittelee ja pseudonymisoi aineiston, ja luovuttaa pseudonymisoidun aineiston päätöksen saajalle etäkäyttöympäristöön. | 30 |

## **3 Poimittavat muuttujat** *[hakija täyttää]*

Täytä alla olevaan taulukkoon poimittavat muuttujat mahdollisimman tarkalla tasolla.

Jos tarvitset samalta rekisterinpitäjältä muuttujia useasta eri rekisteristä tai datasetistä, kopioi ja täytä kullekin oma taulukko.

Huom! Jos olet ennen hakemuksen jättämistä saanut rekisterinpitäjältä alustavan kustannusarvion ja poimintakuvausta on tämän jälkeen muutettu, korosta taulukkoon selkeästi siihen tehdyt muutokset.

|  |  |
| --- | --- |
| **Poiminnan vaihe (kts. poimintajärjestys)** | Esim. Vaihe 1, kohdejoukon poiminta |
| **Rekisterinpitäjä** | Esim. Terveyden ja hyvinvoinnin laitos |
| **Rekisteri ja datasetti (osatietojoukko)** | Esim. Terveydenhuollon hoitoilmoitusrekisteri |
| **Rajaukset** | **Kohdejoukko:**  Tällä poiminnalla tunnistetaan kohdejoukon henkilöitä.  TAI  Tällä poiminnalla poimitaan aineistoa jo tunnistetulle kohdejoukolle. Findata toimittaa kohdejoukon henkilötunnukset rekisterinpitäjälle.  Kohdejoukolla tarkoitetaan sitä henkilöjoukkoa, joiden henkilötunnuksilla on tarkoitus poimia tietoja joko tästä kyseisestä rekisteristä tai muista rekistereistä.  Kirjaa tähän tieto siitä, jos tämän taulukon mukaiset tiedot poimitaan kohdejoukon lisäksi myös esim. verrokeille tai sukulaisille, tai jos kohdejoukkoja on useita. |
| **Kohdejoukon /kohdejoukkojen koko:** Merkitse tähän, kuinka iso kohdejoukko arviolta on. Arvioi mieluummin ylä- kuin alakanttiin. Jos tietoja poimitaan myös sukulaisille/verrokeille, laske heidät myös mukaan tähän summaan. |
| **Kohdejoukon koon arvioinut taho (jos toteutettu):** |
| **Aikaväli, jolta tiedot poimitaan:**  Esim. 1.1.2000–31.12.2010 |
| **Aikavälin rajausmuuttuja:**  Kirjaa tähän, minkä muuttujan mukaan aikaväli rajataan. Esim. syntymäpäivä tai lääkkeen määräämispäivä |
| **Poiminnan muut rajaukset:**  Esim. vain yli 18-vuotiaat |
| **Poimittavat muuttujat** | Kuvaa tähän poimittava aineisto muuttujatasolla.  Esim.  Henkilötunnus HETU Asiakkaan sukupuoli SUKUP  Asiakkaan ikä IKA  Päädiagnoosi PDGO, PDGE |
| **Rakenteettomat muuttujat** | Listaa tähän vielä uudelleen ne muuttujat, jotka ajattelet poimittavan rakenteettomassa muodossa.   * Perustele rakenteettomien muuttujien tarve lyhyesti * Määrittele, saako rekisterinpitäjä luoda rakenteettomista muuttujista mahdollisuuksien mukaan suoraan rakenteellista dataa   Esim. Potilas-asiakirjojen tekstit tai sanahaulla poimittavat tekstikatkelmat. Kyseistä tietoa ei ole mahdollista saada muulla tavalla. Tekstit ja tekstikatkelmat tulee toimittaa sellaisinaan. |
| **Tarkistuspiste (hakijan ehdottama)** | Ehdotan, että tästä poiminnasta pidetään tarkistuspiste poiminnan oikeellisuuden varmistamiseksi aineistopoimintojen yhteydessä.  Rastita ruutu, mikäli haluat ehdottaa rekisterinpitäjälle tarkistuspistettä poiminnan oikeellisuuden varmistamiseksi.  Tarkistuspiste on tarkoitettu lähtökohtaisesti kohdejoukon muodostuksen oikeellisuuden varmistamiseksi ketjuuntuvissa poiminnoissa, mutta sitä voidaan hyödyntää myös muiden monimutkaisten ei-ketjuuntuvien poimintojen yhteydessä.  Tarkistuspiste voidaan toteuttaa eri keinoin, esim. hakijalle toimitetaan laajennettu aineistokuvaus (kohdejoukon koko ja käytetyt koodit) tai hakija ja rekisterinpitäjät pitävät Teams-palaverin tai keskustelevat sähköpostitse.  Findata ei osallistu tarkistuspisteen käytännön järjestelyihin, vaan niistä sovitaan hakijan ja rekisterinpitäjän välillä. Findata ei selvitä kohdejoukkoja käsitellessään, onko tarkistus tehty. Hakija tai rekisterinpitäjä informoi Findataa, jos tarkistuspisteen seurauksena ilmenee epäselvyyksiä, jotka vaativat Findatan toimia. |
| **Muuta huomioitavaa** |  |

## **4** **Rekisterinpitäjän kommentit esitetystä poiminnasta** *[rekisterinpitäjä täyttää]*

|  |
| --- |
| **Rekisterinpitäjä**:  Täytä tähän laatikkoon mahdolliset kommentit poiminnasta.  **Tarkistuspiste (rekisterinpitäjän ehdottama):**  Ehdotan, että yllä olevista poiminnoista pidetään tarkistuspiste hakijan kanssa poiminnan oikeellisuuden varmistamiseksi aineistopoimintojen yhteydessä.  Rastita ruutu, mikäli haluaisit ehdottaa hakijalle tarkistuspistettä poiminnan oikeellisuuden varmistamiseksi.  Täytä tähän myös kommentit tarkistuspisteen tarpeellisuudesta, mikäli hakija on rastittanut toiveen tarkistuspisteestä poimintataulukossa.  Tarkistuspiste on tarkoitettu lähtökohtaisesti kohdejoukon muodostuksen oikeellisuuden varmistamiseksi ketjuuntuvissa poiminnoissa, mutta sitä voidaan hyödyntää myös muiden monimutkaisten ei-ketjuuntuvien poimintojen yhteydessä.  Tarkistuspiste voidaan toteuttaa eri keinoin, esim. hakijalle toimitetaan laajennettu aineistokuvaus (kohdejoukon koko ja käytetyt koodit) tai hakija ja rekisterinpitäjät pitävät Teams-palaverin tai keskustelevat sähköpostitse.  Findata ei osallistu tarkistuspisteen käytännön järjestelyihin, vaan niistä sovitaan hakijan ja rekisterinpitäjän välillä. Findata ei selvitä kohdejoukkoja käsitellessään, onko tarkistus tehty. Hakija tai rekisterinpitäjä informoi Findataa, jos tarkistuspisteen seurauksena ilmenee epäselvyyksiä, jotka vaativat Findatan toimia. |

## **5 Poiminnan määräarviot rekisterinpitäjältä** *[rekisterinpitäjä täyttää]*

Täytä oheinen taulukko haettavien tiedostojen määräarvioista. Määräarvioiden avulla voimme Findatassa arvioida aineistonkäsittelyn kustannukset hakijalle.

|  |  |
| --- | --- |
| Tarvittava tieto | Täytä tiedot tähän sarakkeeseen |
| Arvio toimitettavan aineiston sisältämästä henkilömäärästä | N: |
| Arvio rakenteellisena toimitettavien aineistojen sisältämistä rivimääristä ja tiedostomääristä | Alle 5 milj. rivin tiedostoja: XX kpl  5–20 milj. rivin tiedostoja: XX kpl  Yli 20 milj. rivin tiedostoja: XX kpl |
| Tunnistetietoja sisältävät muuttujat  Listaa tähän  - kaikki tunnistetietoja mahdollisesti sisältävät muuttujat, sekä  - millainen tunnistetieto on kyseessä. |  |
| Sisältääkö toimitettava aineisto ei-rakenteista tietoa? | Kyllä / Ei |
| Onko rekisterinpitäjän mahdollista jalostaa avointa tekstiä rakenteiseen muotoon avoimen tekstin toimittamisen sijaan?  Kyllä: mitä muuttujia tämä koskee?  Esim: tupakointitiedot voidaan tuottaa muodossa tupakoi/lopettanut tupakoinnin/ei ole koskaan tupakoinut | Kyllä / Ei |
| Rakenteettoman tiedon mahdollisimman tarkka kuvailu | 1. Sisältääkö rakenteeton tieto henkilöiden tunnistetietoja (esim. nimiä, osoitteita, puhelinnumeroita): 2. Missä tiedostomuodossa rakenteeton tieto on: 3. Arvio, kuinka monessa tiedostossa on rakenteetonta tietoa: 4. Arvio rakenteettoman tekstimuotoisen tiedon rivimääristä per tiedosto: 5. Kuvailu rakenteettomasta tiedosta (esim. röntgenkuvia, tekstikatkelmia 100 merkin pituudella tai kolme lausetta hakusanan ympäriltä): |
| Poimittavat rekisteritiedot toimitetaan csv-tiedostona  Jos poimittavaa aineistoa ei voi toimittaa csv-tiedostoina – esimerkiksi kuvantamisaineistot – ilmoita missä tiedostomuodoissa aineisto toimitetaan. | Kyllä |

## **6 Poiminnan enimmäiskustannusarvio** *[rekisterinpitäjä täyttää]*

Täytä alla olevaan kustannusarviotaulukkoon

* arvio edellä kuvatun poiminnan kokonaiskustannusten enimmäismäärästä
* kustannusarvion erittely eli tieto siitä, mistä kokonaiskustannusten arvio koostuu
* tieto arvonlisäverosta.

Voit muokata taulukon tietosisältöä tapauskohtaisesti vastaamaan poimintanne kustannuserittelyä. Taulukosta tulisi ilmetä alvillinen ja alviton hinta.

Jos pyydettyjen tietojen kokoaminen edellyttää useita poimintoja, kopioi alla oleva taulukko ja täytä kunkin poimintakerran enimmäiskustannusarviot erillisiin taulukoihin.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Tarvittava tieto | Suorite  Esim. poiminta rekisteristä X | Suoritteen yksikkö  Esim. Työtunti | Yksikkömäärä Esim. 35 kpl | Yksikköhinta Esim. 115 euroa/kpl | Suoritteen kokonaishinta Esim. 4025 euroa |
| Kustannusarvion erittely |  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Suoritteiden kokonaishinta ilman arvonlisäveroa | XX euroa  Huom! Luvanhakijan pyytämä hinta-arvio voidaan antaa suuntaa-antavana. Findata varmistaa hakemuskäsittelyn aikana poimintakustannukset rekisterinpitäjiltä. | | | | |
| Arvonlisäveron prosentti (mikäli arvonlisävero laskutetaan) | XX %  Huom! Findata toimii vain laskun välittäjänä ja käsittelee laskua kirjanpidossaan läpikulkueränä. Siten rekisterinpitäjä määrittää arvonlisäverovelvollisuuden laskun lopullisen maksajan perusteella. Laskun maksaja on ilmoitettu tämän lomakkeen kohdassa 1 Hankkeen lyhyt kuvaus.  Lisää Findatalle osoitettuihin laskuihin viittaus luvansaajasta tai luvansaajan ilmoittamasta muusta laskun lopullisesta maksajasta. Näin luvansaaja tai muu laskun lopullinen maksaja voi vähentää laskuun liittyvän arvonlisäveron omassa verotuksessaan.  Käytä seuraavaa tekstiä ja täydennä siihen maksajan tiedot: Laskun lopullinen maksaja on \*maksajan nimi (alv-tunnus)\*, minkä mukaisesti laskun arvonlisäverokäsittely määräytyy. Findata ei ole laskun osalta arvonlisäverovelvollinen, vaan käsittelee sen läpilaskutuksena. (Verohallinnon ohjaus)”  The final payer of the invoice is \*maksajan nimi (VAT-ID: )\*, according to which the VAT treatment of the invoice is determined. Findata treats the invoice as a pass-through invoicing and is therefore not liable for VAT on the invoice. (Guidance from Finnish Tax Administration) | | | | |
| Enimmäiskustannusten arvio arvonlisäveron kanssa yhteensä | XX euroa | | | | |
| Kustannusarvion antajan yhteystiedot | Nimi ja sähköposti | | | | |
| Kustannusarvion antamisen päivämäärä ja voimassaolo | pp.kk.vvvv | | | | |