



Virve Lans
Andmekaitse Inspeksioon
info@aki.ee

Teie 15.01.2024 nr 2.2.-1/24/29-2

Meie 28.01.2024 nr 24.8-4/S/5235-3

Parandatud taotlus isikuandmete töötlemiseks teadusuuringus ilma isiku nõusolekuta

Austatud Virve Lans

Andmekaitse Inspeksiooni menetluses on Majandus- ja Kommunikatsiooniministeeriumi taotlus uuringus „Meeste väikse osakaalu põhjused EHW valdkonna õppurite ja töötajate hulgas“ isikuandmete töötlemiseks andmesubjekti nõusolekuta.

Andmekaitse Inspeksioon palus oma 15.01.2024 kirjas kõrvaldada taotluses esinevad puudused ja vastata taotlusega seotud küsimustele hiljemalt 30.01.2024.

Parandatud taotlus isikuandmete töötlemiseks teadusuuringus on lisatud käesolevale kirjale kui Lisa 1. Siinkohal toome ära meie vastused taotlusega seotud küsimustele:

- 1. Taotluse punktis A (õppimise kategooria all) on öeldud: õpingute alustamise, katkestamise ja lõpetamise kirjeldamine lisaks soole teiste taustatunnuste lõikes, et /.../ Palun selgitage, millised on teised taustatunnused.*

Taustainfona tahame juhtida tähelepanu sellele, mida rikkalikum on taustatunnuste loetelu, seda sisukamalt on võimalik nähtust hinnata. Kuna tegemist on teadusuuringuga, siis ei ole meil ette teada, millised taustatunnused on nähtusega seotud. Uuringu osaks ongi tuvastada ja otsida selliseid tunnuseid ja seoseid. Teisalt piiravad analüüsi registrite tunnuste loetelu, mida üldse saab taustatunnusena kasutada. Lähtudes neist kaalutlustest määratleme andmete koosseisus teised taustatunnused järgmisena:

- vanus;
- kodakondsus;
- emakeel;
- elukoha maakond;
- haridustase.

- 2. Taotluse punktis A (töötamise kategooria all) on öeldud: Töötasu EHW ametialadel: meeste ja naiste kuine brutotasu 2014–2023 eriala, vanuse, **rahvuse** ja tööandja piirkonna lõikes. Juhin tähelepanu, et rahvus kuulub eriliiki isikuandmete hulka ning sellisel juhul tuleb pöörduda ka eetikakomiteesse (isikuandmete kaitse seaduse § 6 lg 4). Samas ei ole rahvust nimetatud taotluse punktis 5 kirjeldatud andmekoosseisudes. Kas te soovite rahvust pärida?*

Kui jah, siis millisest andmekogust ning kas olete konsulteerinudeetikakomiteega kooskõlastuse taotlemise vajadusest? Kui ei soovi pärida, siis palun taotlusest rahvus välja jätta.

Taotluse punkt A viitab riigihanke ülesandepüstitusele, kus on rahvus võimaliku tasutatunnusena märgitud. Arvestades uuringu ülesehitust ja seda, et rahvus kuulub eriliiki isikuandmete alla, samuti ajakulu, mis sellise andmeühiku lisamine uuringule kaasa toob eetikakomitee menetluse lisandumisega, otsustasime selle analüüsist välja jätta ning seepärast ei kajastu see ka andmete koosseisu päringus. Kustutasime selle taotluse tekstis ülesandepüstitusest.

- 3. Taotluse punktis C on öeldud, et teistes uuringu osades töödeldakse andmeid isiku nõusolekul, mistõttu ei puuduta taotluse nende andmete töötlemist. Palun siiski kirjeldada ka seda osa uuringust, mis tugineb nõusoleku alusele, sealhulgas, kuidas planeerite jõuda sihtrühmani. Palun edastada nõusoleku vorm või selle kavand. Eeldan, et nõusoleku alusel toimuva uuringuosa puhul küsitletakse sihtgruppi kuuluvaid isikuid. Palun kirjeldada, mil moel küsitlus läbi viiakse (intervjuu, küsimustik vms), kes viib, kuidas küsitlusandmeid töödeldakse (sh kas kasutatakse pilveteenust ja küsitlusplatvormi vms ning millises riigis asub selle teenusepakkuja server jms).*

Uuringu käigus tehakse intervjuud sihtrühmaga. Sihtrühmauuringus osalejateni jõudmiseks võetakse ühendust EHW valdkonna erialade esindajate ja ülikoolilistide ning erialaliitudega, samuti kasutatakse n-ö lumepalli meetodit, kogudes võimalike intervjueeritavate kohta soovitusi ja jagades üleskutset nii tiimiliikmete võrgustike kui ka intervjueeritavate kaudu. Võimalikel intervjueeritavatel lastakse täita ära lühike veebivorm (screener) Google Forms platvormil. Veebivormis küsitakse järgnevaid isikuandmeid: nimi, meiliaadress, vanus, elukoht, seotus EHW valdkonnaga. Vastuste põhjal tehakse valik, kes intervjuudesse kutsutakse. 60 päeva jooksul pärast valiku tegemist veebivormi kaudu kogutud andmed kustutatakse.

Sihtrühma esindajatega tehakse individuaalsed intervjuud kas Zoomi platvormil või ruumikohtumise vormis. Intervjuud viivad läbi projektimeeskonna liikmed. Kõik intervjuud salvestatakse edasiseks analüüsiks, heli- ja videosalvestised jäävad uuringu lõpuni intervjuu teinud tiimiliikme arvuti kõvakettale ning neid ei jagata teiste tiimiliikmete, kolmandate osapooltega ega hoiustata pilveteenust kasutades. Intervjuu teinud uurija kuulab intervjuu läbi ja teeb nendest mittetäielikud transkriptsioonid, mida jagab vaid teiste uuringutiimi liikmetega ja kus kasutatakse varjunimesid. Intervjuude helisalvestised kustutatakse pärast projekti lõppemist, IV kvartalis 2024. Uuringu lõppraportis võime kasutada tsitaate vestlustest, kuid teeme seda kasutades tunnuseid, mis ei võimalda isikut tuvastada. Intervjuude informeerimise ja nõusolekuvorm on lisatud kirjale. Intervjuuereimise korraldab uuringus Rakendusliku Antropoloogia Keskus, kes ei tee ühendatud registriandmete analüüsi ning kes ei ole ka volitatud töötleja AKIle esitatud taotluse andmete töötlemise mõttes. Seetõttu me ei lisanud nimetatud infot täiendatud taotluse teksti.

- 4. Taotluse kohaselt teeb Statistikaamet uuringu läbiviijale andmed kättesaadavaks pseudonüümitud kujul. Kas Statistikaametis säilib koodivõti? Kui jah, siis palun selgitage miks ning kui kaua koodivõti säilib.*

Jah, statistikaametis säilib koodivõti. Statistikaamet kasutab andmete pseudonüümimiseks oma koodivõtit, mida kasutatakse ka teistes statistikatöodes. See on vajalik, et erinevatest andmeallikatest saadud andmed oleks pseudonüümitud samal viisil ja oleks omavahel lingitavad nii täna, minevikus kui tulevikus statistikaametisse saabuvate andmetega. Kuna tegemist on regulaarses kasutuses oleva koodivõtmega, siis seda ei kustutata. Lisasime vastava info taotluse punkti 7.

- 5. Taotluse punktis 5 andmekoosseisus on välja toodud, et registritest päritakse pseudokood, kas siiski päritakse isikukood ning selle abil seotakse andmed ja määratakse pseudokood või kasutatakse Statistikaametis juba inimese kohta olevaid pseudokoode?*

Uuringu käigus registritest lisapäringuid ei tehta. Kasutatakse andmeid, mis on statistikaameti kasutuses statistikatööde tegemiseks. Seega kasutatakse juba pseudonüümitud andmeid, millel on pseudokoodid (nii inimese isikukoodi kui ettevõtte registrikoodi tasemel).

- 6. Uuringu eesmärgiks on välja selgitada meeste väikse osakaalu põhjuseid EHW valdkonnas, kuid taotluse kohaselt päritakse kõikide selle valdkonna inimeste andmeid, sh naiste. Palun põhjendada, miks on vajalik uuringu eesmärgi täitmiseks pärida ka naiste andmeid?*

Meeste väike osakaal ilmneb ainult võrdluses tervikuga, mille moodustavad mehed ja naised koos. Ühelt poolt on uurimisobjektiks karjääriteekonnad, teiselt poolt aga meeste ja naiste võrdlus, kas ja mille poolest on valdkonnas mehed ja naised erinevad ning mille poolest sarnased.

- 7. Taotluse punktis 5 on öeldud: II EHW valdkonna tööturu olukorra võrdluseks kõik perioodil 2014–2023 vähemalt ühel kuul töötanud inimesed, kelle kohta on vajalik andmete koosseis väiksem. Täpne isikute arv pole teada, kuid hõivatute arv on aastas ca 650000. Palun põhjendada, miks on vajalik pärida võrdluseks nii suure hulga inimeste andmeid? Miks ei piisa võrdluseks vähemate inimeste andmed? Palun kaaluda väiksema hulga inimeste andmete kaasamist.*

Võrdlusandmete päringut on vaja selleks, et vastata taotluse punktis A toodud uurimisülesandele: Töötasu võrdlus teiste ametialadega (kuna ametialade arv ISCO klassifikatsiooni järgi on suur, tehakse valik ametialadest, kelle töötasudega EHW töötajate tasu võrreldakse, või grupeeritakse võrdlusametialasid). Töötasu võrdluse sisuks on hinnata, kuidas erinevad samasugust tööd tegevate inimeste töötasud EHW valdkonnas ja teistes valdkondades. Kuna töötasu võrdlus peab arvestama rikkaliku tasutaandmestikuga, sest tahame võrrelda võimalikult sarnaseid inimesi, siis on vaja võrdluses arvestada haridust, töökoormust, majandustegevusala ja ametiala jm tunnuseid, mida võimalik. Võrdluse tegemiseks peame selekteerima võrdlusandmetest välja taustatunnuste poolest sobivad võrdlusisikud, selleks kasutame ökonomeetrisi meetodeid nagu täppis- või tõenäosuspõhine sobitamine. Enne selle analüüsi tegemist ei ole võimalik teada, millised grupid ja isikud võrdluseks sobivad. Samuti otsustame, millist võrdlusgrupi selekteerimise meetodit kasutada analüüsi väljundite põhjal, sobitatud ja kontrollgrupi statistikute võrdlemine annab sisendi, milline meetod sobib võrdlusgrupi koostamiseks. Kuna korrektse ja parima kontrollgrupi leidmine on üks oluline analüüsi osa, siis pole võimalik enne analüüsi tegemist piiritleda võrdlusandmetena kasutatavat andmestikku.

Andmete koosseisu kirjelduses täpsustasime registrite nimetusi taotluses.

- 8. Palun Teil uuesti hinnata andmekoosseisu vajalikkust ning põhjendada registritest päritavate andmekoosseisude vajalikkust uuringu eesmärgi täitmiseks (iga andmekoosseisu kuuluv kirje eraldi). Hetkel tekitab näiteks küsimusi, miks on EHISest vajalik pärida õppekavagrupperi ja -rühma, miks ei piisa õppekava nimetusest; TÖRist päritakse tööandja pseudokoodi, tööandja tegevusala EMTAK koodi ja lisaks väljamakse tegija pseudokoodi, miks ei piisa tööandja pseudokoodist. TÖRist soovitakse ka töösuhte peatamise põhjust, miks see on vajalik ja kas need andmed on TÖRis? TÖRi andmekoosseisus on ka kirje algversioon, lõppversioon, versiooni kehtivuse algus ja lõpp, mida need kirjed sisaldavad ning kuidas need uuringu eesmärgi täitmiseks vajalikud on.*

Õppekavade klassifikatsioon on ehitatud üles nii, et iga õppekava nimetus kuulub õppekavarühma ja õppekavarühm omakorda õppekavagrupperi. (Näiteks kuuluvad õppekavagrupperi „Õpetajakoolitus ja kasvatusteadus“ järgmised õppekavarühmad: „Kasvatusteadus“, „Koolieelikute õpetajate koolitus“, „Klassiõpetaja koolitus“, „Aineõpetaja koolitus“, „Haridus, mujal liigitamata“ ja „Hariduse interdistsiplinaarne õppekavarühm“). Seega annab õppekavagrupperi ja õppekavarühma üldistuse võimaluse ehk võimaluse üldisemal tasandil analüüsi teha. Juhul, kui me päriksime ainult õppekava nimetuse, oleks vaja õppekavasid koondada käsitsi õppekavagrupperidesse ja õppekavarühmadesse.

Seega ei anna nende tunnuste küsimine isiku kohta detailsemat infot, vaid aitab õppekavade analüüsi lihtsamal moel üldisemal tasandil läbi viia.

TÖRi päringu kirjete põhjused on järgmised:

- Tööandja pseudokood on tehniline näitaja, millega saab erinevaid andmestikke omavahel siduda ja sellel puudub sisuline tähendus analüüsi mõistes.
- Tööandja tegevusala EMTAK kood tähistab Eesti Majandustegevusalade Klassifikaatorit, mis näitab millisel tegevusalal tööandja tegutseb ja on oluliseks analüüsitavaks tunnuseks, kuna tegevusalati erinevad töötasud, tööjõu nõudlus ja pakkumine, töötajate oskuste ja teadmiste spetsiifika, siis on vaja seda arvestada tööturu analüüsid.
- Väljamakse tegija pseudokood on vajalik, kuna selle abil saab siduda töötamise registri ja tulu- ja sotsiaalmaksudeklaratsiooni andmed avaliku sektori asutustel, kus tööandjat kitsamalt (nt munitsipaalkool) ja töötasude maksja (kohalik omavalitsus) ei lange kokku. Kuna analüüs keskendub EHW valdkonna karjääriteedele, siis peame siduma ka töötasude ja kitsamalt tööandja andmed.
- Töösuhte peatamise põhjuse välja panime analüüsi, kuna üheks peatamise põhjuseks on lapsehoolduspuhkus, mida võtavad endiselt rohkem naised kui mehed ja selleks et see ei viiks kallutatud järeldustele tööturu teekonna hindamisel, kavatsesime seda arvesse võtta. Samas nõuab selline analüüs, kus arvestame muuhulgas töösuhte peatamise põhjuseid rohkem aega, kui meil on kasutada praeguse analüüsi tegemisel, siis loobume sellest ja jätame töösuhte peatamise põhjuse välja.
- Kirje algversioon, lõppversioon, versiooni kehtivuse algus ja lõpp on andmed, mis võimaldavad eristada TÖRi registris tehtud muudatusi ja üle erinevate versioonide jätkuvaid kirjeid omavahel ühtlustada.

9. *Taotluse punkti 8 kohaselt ei pea Te andmesubjekti teavitamist vajalikuks. Selgitan, et isikuandmete kaitse üldmääruse (IKÜM) üheks oluliseks põhimõtteks on läbipaistvuse põhimõte. Isikuandmete vastutaval töötlejal on kohustus esitada läbipaistvalt teavet isikuandmete töötlemise kohta. IKÜM artiklid 12–14 sätestavad, milline teave tuleb andmesubjektile esitada. Isikuandmete vastutav töötleja peab võtma aktiivseid meetmeid kõnealuse teabe esitamiseks andmesubjektile või suunama andmesubjekti aktiivselt selle asukohta. IKÜM artikkel 14 lõike 5 punkti b kohaselt on tõesti võimalik tugineda teavitamata jätmisele, kuid sama punkt ütleb ka, et sellistel juhtudel võtab vastutav töötleja asjakohased meetmed andmesubjekti õiguste, vabaduste ja õigustatud huvide kaitsmiseks, sealhulgas teeb teabe avalikult kättesaadavaks. Antud uuringu puhul planeerite töödelda suure hulga inimeste andmeid ning ei nähtu, et neil oleks võimalik sellisest andmetöötlustest teada saada, seega leian, et tuleb teha vastav teavitus enne andmetöötluste algust. Kuivõrd töödeldakse suure hulga inimeste andmeid ning kontaktandmeid registritest ei pärita, siis teavitus on võimalik teha ka näiteks ministeeriumi kodulehel, meedias vms.*

Majandus- ja Kommunikatsiooniministeerium avaldab enne andmetöötluste algust oma kodulehel uuringu info vastavalt IKÜM artiklites 12-14 äratoodule.

Lugupidamisega

(allkirjastatud digitaalselt)

Tiit Riisalo
majandus- ja infotehnoloogiainminister

Lisa. 1) Parandatud taotlus isikuandmete töötlemiseks teadusuuringus

Grete Kaju
Grete.Kaju@mk.m.ee