

Arvamus Andmekaitse Inspektsioonile

Kuupäev: 07.09.26

Uuring: „Leibkonna abivajaduse andmebaaside-põhise hindamise võimalikkus toimetulekutoetuse ja tervishoiu omaosaluskoormuse näidete alusel“

1. Konsultatsiooni küsimise asjaolud

Andmekaitse Inspektsioon pöördus ETAGi teaduseetika komitee poole isikuandmete kaitse seaduse § 6 lg 5¹ alusel, et saada arvamus uuringu eetiliste aspektide kohta. Teaduseetika valdkonna juht Marten Juurik edastas 4. augustil 2026 konsultatsioonipalve ja uuringu materjalid komitee liikmetele. Vastutavaks määrati terviseandmete ja andmeteaduse alamkomitee. Esialgne arvamus koostati kirjalikus menetluses komitee liikmete tähelepanekute alusel. 3. septembril toimus kohtumine AKI esindajatega, kus arutati esmaseid teaduseetika komitee tähelepanekuid. Kohtumisel antud selgitusi võeti arvamuse lõplikul vormistamisel arvesse.

2. Üldhinnang

Teaduseetika komitee hinnangul ei ole esitatud materjalide põhjal võimalik pidada kogu kavandatud andmetöötlust kirjeldatud ulatuses piisavalt põhjendatuks ja proportsionaalseks. Enne töötlemise lubamist tuleb eristada uurimisülesanded, põhjendada iga ülesande andmevajadust ning vähendada andmekoosseisu ulatuseni, mis on vastava eesmärgi saavutamiseks vältimatult vajalik.

Minimaalselt tuleb eristada kolme uurimisülesannet: 1) registripõhiste leibkondade konstrueerimine; 2) leibkondade majandusliku olukorra ja võimaliku abivajaduse hindamine; 3) tervishoiu omaosaluse koormuse analüüs. Leibkondade konstrueerimine võib eeldada kogu elanikkonda hõlmavaid lähteandmeid, kuid mitte kõiki taotletud tunnuseid. Abivajaduse ja tervishoiu omaosaluse analüüsi puhul tuleb eraldi põhjendada nii üldkogumi ulatust kui ka isikutasandi andmete detailsust ning hinnata valimi või agregeeritud andmete kasutamise võimalust.

3. Uuringu asjaolud

Uuringu tellija, AKI-le taotluse esitaja ja vastutav töötaja on Sotsiaalministeerium. Volitatud töötlejad on Sihtasutus Mõttekoda Praxis, Proud Engineers OÜ, Kasecampers OÜ ja Statistikaamet. Uuringu eesmärk on hinnata toimetulekutoetuse ja tervishoiu omaosaluse digitaliseerimise, automatiseerimise ja personaliseerimise võimalusi.

Taotluses lähtutakse probleemist, et toimetulekutoetuse saamiseks peab inimene praegu ise abi otsima ja esitama kohalikule omavalitsusele ulatuslikult andmeid, millest osa võib riigil juba olemas olla. Tervishoiu omaosalus ei sõltu praegu leibkonna sissetulekust ega koosseisust. Uuringuga otsitakse võimalusi kasutada abivajaduse tuvastamiseks rohkem registriandmeid ja pakkuda teenuseid proaktiivsemalt.

Uuringu raames toimub:

- 1) **Registripõhine leibkondade konstrueerimine.** Elukoha-, lisa-aadressi-, viibimiskoha-, kinnisvara-, suhte-, hooldusõiguse-, eestkoste- ja toetuste andmete põhjal püütakse tuvastada, millised inimesed elasid aastatel 2024–2025 koos ja moodustasid ühe leibkonna. Andmepõhiseid leibkondi võrreldakse küsitlusuuringutes ja toimetulekutoetuse menetlustes kirjeldatud leibkondadega.
- 2) **Leibkondade tulude, kulude ja abivajaduse hindamine.** Uuritakse, millises ulatuses on registriandmete põhjal võimalik arvestada leibkonna tulusid, vara ja eluasemekulusid. Selle põhjal hinnatakse toimetulekutoetuse taotlemise ja määramise digitaliseerimise, automatiseerimise ning proaktiivsemaks muutmise võimalusi. Taotleja seob selle vajadusega vähendada inimeselt juba riigi valduses olevate andmete korduvat küsimist ning leida võimalikud toetuse saajad, kes praegu toetust ei taotle. Uuring peab muu hulgas näitama, kas andmestikud ja loodav leibkonnamudel võimaldavad neid eesmärke piisava täpsusega täita.
- 3) **Tervishoiu omaosaluse koormuse analüüs.** Uuringus ühendatakse tervishoiuteenuste kasutamise ja omaosaluse andmed leibkonna koosseisu, sissetuleku ja muude sotsiaal-majanduslike andmetega. Selle põhjal hinnatakse, kuidas tervishoiuga seotud kulud mõjutavad eri leibkondade majanduslikku olukorda. Selle põhjal omakorda analüüsitakse, kas ja kuivõrd oleks võimalik arvestada tervishoiu omaosaluse piirmäära leibkonna põhiselt.
- 4) **Riskide kaardistus ja tulevaste lahenduste kavandamine.** Uuringu tulemusena tehakse ettepanekuid võimaliku süsteemi arenduseks, mis arvestaks leibkonna majanduslikku seisu olemasolevate andmete põhiselt ja lihtsustaks toetuse taotlemist.
- 5) **Ekspertintervjuud ja KOVide küsitlus.** Need ei ole osa AKI taotlusest, kuid lisainfo ja ekspertide sisendi saamiseks kogutakse täiendavalt andmeid.

AKI-le esitatud taotlus puudutab olemasolevate isikuandmete kasutamist ja ühendamist andmepõhiste leibkondade konstrueerimiseks ning nende majandusliku olukorra, võimaliku abivajaduse ja tervishoiu omaosaluse koormuse modelleerimiseks.

Taotluse järgi ei ole uuringu kogu andmekoosseis veel lõplik. Pärast kirjanduse analüüsi ja esimese uurimisülesande täitmist on kavandatud esitada AKI-le lisataotlus

leibkondade majandusliku toimetuleku analüüsimiseks vajalike andmete kasutamiseks. Käesolev arvamus puudutab seetõttu ainult praegu esitatud andmepäringut.

Taotluses on ette nähtud pseudonümiseerimine, Statistikaameti turvaline teadlaste keskkond, üldistatud väljundid ning leibkonnamääratluse valideerimine eri andmestike võrdlemisega. Samuti on käsitletud EUROMODi ja Tervisekassa andmete kasutuseesmärgi ning õiguslike riskide, läbipaistvuse ja vaiete küsimusi.

4. Andmekoosseis

Taotluses on kirjeldatud üheksat nummerdatud andmestikku, millest EUROMOD ja ESU sisaldavad omakorda eri allikatest ühendatud andmeid. Taotluses on nimetatud: 1) EUROMODi maksude ja toetuste mudeli sisendandmestik (komitee loenduse põhjal kokku 171 tunnust); 2 ja 3) sotsiaalteenuste ja -toetuste andmeregister STAR (2 andmestikku, kokku 25 tunnust); 4) Töötukassa andmekogu (5 tunnust); 5) rahvastikuregister (18 tunnust); 6) kinnistusraamat (5 tunnust); 7) Tervisekassa andmekogu ja Retseptikeskuse andmed, mida taotluses käsitletakse ühe andmestikuna (kokku 54 tunnust); 8) Eesti Sotsiaaluuring (ESU); 9) Eesti leibkondade finantskäitumise ja tarbimisharjumuste uuring (HFCS).

- 1) **Eesti Sotsiaaluuringu andmestik.** Lisaks inimestelt küsitlusega kogutud vastustele sisaldab ESU andmestik haldusandmetest saadud tunnuseid. Taotluse lisas on nende päritolu tähistatud valdavalt asutuste lühenditega MTA, SKA, ETK ja HK. Lisa ei nimeta nende tunnuste täpseid algandmekogusid ega kirjelda, millal, millisel alusel ja millises menetluses on need küsitlusandmetega ühendatud.
 - a) **Maksu- ja Tolliameti haldusandmed (MTA):** palgatulu, koondamisega seotud tulud, pensionisamba maksed ja väljamaksed ning muud maksuarvestusega seotud tulud.
 - b) **Sotsiaalkindlustusameti haldusandmed (SKA):** perehüvitis, elatisabi, pensionid, vanemahüvitis, puudetoetus, toetuste maksmise kestus ning muud sotsiaalkaitsega seotud rahalised hüvitisid.
 - c) **Eesti Töötukassa haldusandmed (ETK):** registreeritud töötus, töötuskindlustushüvitis, töötutoetus, koondamis- ja maksejõuetushüvitis, õppimise ja ettevõtluse alustamise toetused, töövõimetoetus ning muud Töötukassa makstavad toetused.
 - d) **Tervisekassa haldusandmed (HK):** sünnitushüvitis ning haigus- ja hooldushüvitis.
 - e) **Puude ja töövõime andmed:** puude raskusaste, töövõimetus protsent ja määratud töövõime ulatus; täpne algandmekogu ei ole lisas nimetatud.

Lisaks sisaldab ESU andmestik küsitluse vastustest ja haldusandmetest arvutatud või tuletatud tunnuseid, näiteks leibkonna neto- ja ekvivalentsissetulek, tuludetsiilid ja -kvintiilid, suhtelise ja absoluutse vaesuse tunnused, ilmajaetus, tööintensiivsus ning

leibkonnatüübid. Failist ei selgu iga tuletatud tunnuse puhul täpselt, millised algandmed on saadud küsitlusega ja millised haldusallikatest.

2) **EUROMOD** on maksude ja toetuste mikrosimulatsioonimudel. Mudel ise ei ole andmekogu, vaid kasutab sisendandmestikku, millele rakendatakse maksude ja toetuste arvutamise reegleid. Käesolevas uuringus taotletakse ligipääsu Statistikaameti registripõhisele EUROMODi sisendandmestikule, mis hõlmab taotluse järgi kõiki Eesti residentide 2024. ja 2025. aastal. Kõik tunnused on aastase täpsusega.

Taotluse lisas märgitakse, et kõiki tunnuseid ei ole saadud ühest andmekogust. Osa tunnuseid on kokku pandud mitme registri põhjal ning osa on agregeeritud teistest EUROMODi tunnustest. Tunnuste allikana on korduvalt nimetatud ka Statregi, kuid taotlusest ei selgu, mida Statreg selles kontekstis täpselt sisaldab ega millistest algandmetest selle tunnused pärinevad.

- a. **Rahvastikuregister (RR):** isikute, leibkondade, vanemate ja partnerite pseudonüümitud identifikaatorid ning osa demograafilistest, perekondlikest, haridus- ja elukohaandmetest.
- b. **Eesti hariduse infosüsteem (EHIS):** omandatav ja omandatud haridustase, hariduse detailsem liigitus, hariduse omandamise aeg ning õppimisega seotud hüvitised.
- c. **Maksudkohustuslaste register ja muud Maksu- ja Tolliameti haldusandmed (MKR):** töötasu, ettevõtlustulu, muud maksustatavad tulud, maksud, pensionisamba maksed ja väljamaksed ning osa hüvitiste andmeid.
- d. **Sotsiaalkaitse infosüsteem (SKAIS):** pensionid, perehüvitised, vanemahüvitis, puuetega inimeste sotsiaaltoetused, üksi elava pensionäri toetus ning muud sotsiaalkaitseandmed.
- e. **Sotsiaalteenuste ja -toetuste andmeregister (STAR):** toimetulekutoetus, energiahüvitis, osa eluasemega seotud toetustest ning teatavad hariduse ja leibkonna kohta kasutatavad tunnused.
- f. **Töötukassa andmekogud EMPIS ja TKIS:** töötuna arvel olemine, tööturuteenused, töötutoetus, töötuskindlustushüvitis, töötusega seotud hüvitiste kestus ja muud tööturutoetused.
- g. **Töövõime hindamise ja töövõimetoetuse andmekogu TETRIS:** puuduva või osalise töövõime andmed.
- h. **Tervisekassa andmekogu KIRST:** haigus-, hooldus-, rasedus- ja sünnitushüvitised ning muud terviseseisundiga seotud rahalised hüvitised.

- i. **Kohalike omavalitsuste andmed (KOV):** eelkõige kohalike toetuste ja energiahüvitisega seotud andmed.
- j. **KredExi andmed:** KredExi kaudu makstud toetused.
- k. **Statistikaameti statistiline register ja muud Statistikaameti andmestikud (Statreg):** demograafilised, haridus-, tööturu-, leibkonna-, puude-, toetuste ja muud tunnused. Statregi kasutatakse nii iseseisva allikana kui ka teistest registritest pärinevate tunnuste arvutamisel ja täiendamisel.

5. Hinnang andmekoosseisu ulatusele

Komitee peab andmekoosseisu erakordselt ulatuslikuks. Olemasolevale EUROMODi ühendandmestikule lisatakse detailsemad andmed elukohtade, suhete, kinnisvara, toimetulekutoetuse menetluste, retseptide, diagnooside, meditsiiniseadmete, hambaravi ja tervishoiuteenuste kohta. Nii tekib kogu elanikkonda hõlmav ühendandmestik, mis võimaldab kujundada detailse tervikpildi nii üksikisikust kui ka tema oletatavast leibkonnast, sidudes omavahel inimese elukoha- ja peresuhted, sissetuleku, vara, toetused, tööturuseisundi, puude, töövõime ning tervishoiuteenuste kasutamise.

Kõigi tunnuste päritolu ei ole võimalik lõpuni jälgida. Alati ei selgu, millised andmed pärinevad konkreetsest registrist, millised on Statistikaamet arvutanud ning millised on koondatud mitmest allikast. See läbipaistmatus raskendab andmete korduvuse ja iga algandme vajalikkuse hindamist.

Taotluses põhjendatakse kogu elanikkonna hõlmamine registripõhiste leibkondade määratlemise ning tervishoiu omaosaluse jaotuse hindamisega, kuid ei võrrelda seda vähem ulatuslike alternatiividega. Soovitav on uurimisülesande kaupa välja tuua, miks vajavad need analüüsid kõigi Eesti residentide detailseid isikutasandi andmeid ning miks ei piisa esinduslikust valimist, vähem detailsest andmestikust või eri analüüside tegemisest eraldi.

Läbipaistvuse probleem ei piirdu andmeallikate nimetamisega. Menetleval asutusel ja avalikkusel peab olema võimalik mõista, **milline uus teadmine inimeste kohta andmete ühendamisel tekib**. Praegu on uuringu andmekoosseis jagatud taotluse, andmepäringu kirjelduse, EUROMODi tunnuste, ESU Exceli faili ja HFCS ankeedi vahel. Ühegi dokumendi põhjal eraldi ei teki tervikpilti sellest, et uuring võimaldab kogu elanikkonna tasandil siduda leibkondlikud suhted, elukoha, kinnisvara, sissetuleku, vara, toetused, tööturuseisundi, töövõime ja terviseandmed.

Komitee hinnangul ei piisa sellise uuringu läbipaistvuse tagamiseks uuringu pealkirja, üldeesmärgi ja andmekogude nimetuste avaldamisest. Taotluses on kirjeldatud andmete liikumist, pseudonümiseerimist ja ühendamist Statistikaameti keskkonnas, kuid ei ole ühtset tunnusepõhist ülevaadet kasutatavate andmete algallikast, seosest

konkreetses uurimisülesandes ja analüütikule kättesaadava lõpliku andmestiku koosseisust. Üldine soovitus on liikuda andmepõhiste uuringute osas masinloetavate andmekataloogide suunas, mis võimaldaks taotluses välja tuua kõik kasutatavad andmeväljad koos metaandmetega.

Teaduseetika komitee soovib töötada välja andmetöötamise ulatuse ja mõju klassifikatsioon, mis võimaldaks eristada piiratud registriuuringuid kogu elanikkonda ja mitut eluvaldkonda hõlmavatest ühendandmestikest ning määrata viimastele põhjalikumad läbipaistvuse ja põhjendamise nõuded.

6. Andmepõhiste leibkondade konstrueerimine

Uuringus tuletatakse registriandmete põhjal, millised inimesed elavad koos või moodustavad ühise leibkonna. Nii tekib inimese elukorralduse kohta uus teadmine, mida ta ei ole ise esitanud ega saa uuringu käigus kontrollida.

Ühine aadress või kinnisvaraseos ei tõenda kooselu, ühist majandamist ega vastastikust vastutust. Seos võib tuleneda aegunud registreeringust, ajutisest elukorraldusest, omandist, hooldus-, pere- või üürisuhetest või jagatud eluasemest. Samuti ei pruugi tegelik kooselu registrites kajastuda. Ekslik tuletus võib tulevikus mõjutada inimesele omistatavat sissetulekut, vara, abivajadust või toetuseõigust. Põhimõtteline küsimus on ka selles, millise õigustusega võib riik järeldada inimeste kooselu ja majanduslikku seotust andmetest, mis on kogutud muudel eesmärkidel, eriti kui inimene ise sellist seost ei kinnitaks.

Komitee soovib AKI-l paluda taotlejal kirjeldada:

1. milliste andmete ja reeglite põhjal inimesed samasse leibkonda liigitatakse;
2. kuidas lahendatakse eri registrite vastuolud ning hinnatakse tuletatud seoste ebakindlust ja võimalikke vigu;
3. kuidas eristatakse ühist aadressi tegelikust kooselust ja ühisest majandamisest.

Andmepõhist leibkonda tuleb käsitada ebakindla analüütilise konstruktsioonina, mitte tõendina tegeliku kooselu, ühise majandamise või vastastikuse vastutuse kohta. Sellise mudeli võimalik halduslik kasutamine eeldaks eraldi põhjendust ning inimesele arusaadavat võimalust tema kohta tuletatud leibkonnaseost näha, parandada ja vaidlustada.

7. Tervishoiu omaosaluse koormuse andmed

Taotlus põhjendab Tervisekassa ja Retseptikeskuse andmete üldist vajadust tegeliku omaosaluse ja selle võimalike piirmäärade hindamisega, kuid mitte kõigi detailtunnuste kasutamist. Andmeid küsitakse retsepti, ostu, raviarve ja ravijuhu tasandil ning need sisaldavad peale tasutud summade diagnoosi, ravisutust, teenusekoode ja -kuupäevi, ravile saabumise liiki ja raviarve lõpetamise põhjust. Omaosaluse rahalise koormuse

analüüsist ei tulene iseenesest vajadust kirjeldada inimese tervises seisundit ja raviteekonda sellise detailsusega.

Tervishoiu omaosaluse analüüs on eraldiseisev uurimisülesanne, mille andmevajadus ja proportsionaalsus tuleb põhjendada eraldi. Komitee soovib näidata tunnuste kaupa, millised terviseandmed on vältimatult vajalikud ning kas piisaks koondatud omaosaluse summast, laiematest teenuserühmadest või muul viisil agregeeritud andmetest.

8. Automatiseeritud otsustamine ja uuringu eesmärgi õiguslik teostatavus

Taotlus eristab digiteeritud, proaktiivset ja personaliseeritud teenust ning mõnab, et kõiki toimetulekutoetuse otsuseid ei saa automatiseerida. Samuti nähakse ette õiguslike riskide, vaiete ja vajalike õigusmuudatuste analüüs. Uuringu kaugem eesmärk on siiski kasutada registripõhist leibkonda ja majandusandmeid toetuse määramise automatiseerimiseks, abivajaduse proaktiivseks tuvastamiseks ning sissetulekust sõltuva tervishoiu omaosaluse rakendamiseks. Seetõttu tuleb kavandatava lahenduse õiguslikku teostatavust hinnata juba uuringu kavandamisel.

IKÜM-i artikkel 22 piirab üksnes automatiseeritud töötlusel põhinevaid otsuseid, mis tekitavad inimesele õiguslike tagajärgi või mõjutavad teda muul viisil märkimisväärselt. Erandid eeldavad konkreetset õiguslikku alust ja inimese õigusi kaitsvaid meetmeid; eriliiki isikuandmetel põhinevatele otsustele kehtivad täiendavad piirangud.

Käesolevas uuringus ei tehta veel inimeste toetuseõiguse või tervishoiu omaosaluse kohta siduvaid otsuseid. Uuringu eesmärk on aga ette valmistada andmestik, meetodid ja võimalikud süsteemilahendused, mille abil saaks selliseid menetlusi tulevikus senisest automaatsemalt korraldada. Seetõttu peab taotlus selgelt eristama vähemalt kolme võimalikku kasutust:

1. mudeliga analüüsitakse üksnes poliitikastsenaariume ja üldist mõju;
2. mudel tuvastab võimaliku abivajaduse või käivitab inimesele teenuse pakkumise, kuid ei määra tema õigusi;
3. mudeli väljund mõjutab toetuse, omaosaluse või muu inimese õigusi puudutava otsuse sisu.

Nende kasutusviiside õiguslik tähendus ja inimesele avalduv mõju on erinevad.

Üksnes ametniku formaalne osalemine ei välista artikli 22 kohaldamist, kui mudel määrab otsuse sisuliselt ette. Taotleja peab selgitama, mida saab kehtiva õiguse alusel automatiseerida, millised otsused eeldavad sisulist inimese sekkumist ning kuidas saab inimene näha, parandada ja vaidlustada nii lähteandmeid, tuletatud leibkonda kui ka selle põhjal tehtud otsust.

Tulevase lahenduse õiguslik teostatavus mõjutab ka praeguse andmetöötluste põhjendatust. Kogu elanikkonda hõlmavat andmetöötlust ei saa õigustada lahendusega, mille keskne kasutus võib olla lubamatu või eeldada alles eraldi õigusliku aluse loomist. Mudeli sobivusest poliitikaanalüüsiks ei tulene selle sobivus inimese leibkonna, abivajaduse, toetuseõiguse või omaosaluse määramiseks.

9. Isiku informatsioonilise enesemääramise vähenemine

Uuring tõstatab põhimõttelise küsimuse inimese võimalusest kontrollida tema kohta käiva teabe kasutamist. Registriandmete ühendamise ei kasutata üksnes olemasolevaid haldusfakte, vaid tuletatakse uut teavet inimese elukorralduse, leibkondlike suhete, majandusliku seotuse ja võimaliku abivajaduse kohta. Eriti ettevaatlik tuleb olla olukorras, kus andmete ühendamise kaudu võidakse järeldada asjaolusid, mida inimene ise ei sooviks küsitluses või muus otsesuhtluses avaldada.

See vähendab inimese tegelikku kontrolli tema kohta käiva teabe üle. Küsimus ei ole ainult selles, et töötlemine toimub nõusolekuta. Registripõhise tuletamise korral ei tea inimene üldjuhul, milline tervikpilt tema kohta on loodud, milliseid seoseid on tema andmetest järeldatud ega seda, kas need järeldused vastavad tema tegelikule olukorrale. Uuringus ei ole inimesel võimalik tema kohta tuletatud teavet kinnitada, parandada ega selle loomisele vastu vaielda.

Asjaolu, et riigil on õiguslik alus hoida üksikuid andmeid maksustamise, toetuste maksmise, tervishoiu, statistika või muul eesmärgil, ei põhjenda iseenesest nende ühendamist uue ja detailse isikupildi loomiseks. Hinnata tuleb, millist uut teavet ühendamise saadakse, kui eraeluline see teave on ning miks ei ole uuringu eesmärki võimalik saavutada viisil, mis jätaks inimesele suurema kontrolli tema kohta käiva teabe üle. Seda hinnangut toetab IKÜM artikli 5 lõike 1 punktis b sätestatud eesmärgi piirangu põhimõte: asjaolu, et andmed on algsel eesmärgil õiguspäraselt kogutud, ei muuda neid riigi jaoks vabalt kasutatavaks ressursiks. Iga hilisema töötlemise puhul tuleb eraldi hinnata töötlemise eesmärki, õiguslikku alust, vajalikkust ja proportsionaalsust.

10. Riigi informatsiooniline võim ja usaldus

Registrite ühendamine suurendab riigi teadmisevõimet inimese suhtes. Riigiasutustel, andmevaldajatel ja uurijatel on järjest täpsem arusaam sellest, milliseid andmeid saab omavahel ühendada ning milliseid uusi järeldusi nende põhjal teha. Inimesel endal ei ole üldjuhul võrreldavat ülevaadet sellest, milline tervikpilt tema kohta eri andmeallikate ühendamisel tekib. Selline teadmiste ebavõrdsus suurendab avalike institutsioonide informatsioonilist võimu.

Üksikute andmekogude eesmärgid võivad inimestele olla üldjoontes arusaadavad: maksuandmeid kasutatakse maksustamisel, terviseandmeid tervishoiu ning sotsiaalandmeid toetuste ja teenuste korraldamisel. Sellest ei järeldu, et inimesed

oskaksid eeldada nende andmete ühendamist selleks, et tuletada nende leibkonna koosseisu, majanduslikku sõltuvust, haavatavust või võimalikku abivajadust. Eri eesmärkidel kogutud andmete ühendamine loob riigile kvalitatiivselt uue teadmise ja tegutsemisvõime.

Sellise arengu ühiskondlikku legitiimsust ei saa eeldada üksnes töötlemise õiguslikust alusest või tehnilisest teostatavusest. Käesolev uuring on osa laiemast liikumisest registripõhise, proaktiivse ja automatiseeritud avaliku halduse poole. See muudab riigi ja inimese suhet ning tõstatab küsimuse, millist teavet võib riik tuletada, milleks seda kasutada ja millised peaksid olema kasutamise piirid.

Läbipaistvus on siin iseseisev eetiline kaitsemeede. Avalikkusel peab olema lihtsas ja arusaadavas vormis võimalik teada, milliseid andmestikke riik ühendab, milliseid järeldusi nende põhjal tehakse ning kuidas neid järeldusi võidakse kasutada. Üksnes andmekogude nimetuste, õiguslike aluste või uuringu üldeesmärgi avaldamine ei anna tegelikku ülevaadet riigi andmetöötamise ja tuletusvõime ulatusest. Läbipaistvus on ühtlasi õiglase kohtlemise eeltingimus. Kui inimene ei saa aru, milliste andmete, eelduste ja seoste põhjal tema olukorda hinnatakse, ei saa ta kontrollida järelduste õigsust, tuvastada enda jaoks ebasoodsat või ebavõrdset kohtlemist ega oma õigusi tõhusalt kaitsta.

Taotluses käsitletakse tulevase süsteemi läbipaistvust, vaiete võimalust ja tulemuste kommunikatsiooni, kuid selliste lahenduste kujundamine ei tohiks toimuda ainult haldus- ja eksperdikogukonnas. Komitee soovib käsitada avalikku arutelu ja selgitamist süsteemi arendamise vältimatu osana. Enne uute võimsate infosüsteemide kasutuselevõttu peab olema ühiskondlikult arusaadav, millist uut võimekust riik loob ning kuidas inimese õigused ja autonoomia selles muutuses säilivad.

Ulatuslik ja inimese jaoks raskesti hoomatav registriandmete ühendamine võib kahjustada usaldust riigi ja avalike institutsioonide vastu. Usaldusrisk tekib eriti siis, kui inimesed ei saa mõistlikult ette näha, et ühe avaliku ülesande täitmiseks antud andmeid kasutatakse hiljem teise eesmärgiga ühendandmestiku loomiseks ning nende eluolu, suhete või võimaliku abivajaduse kohta uute järelduste tegemiseks. Seetõttu ei ole läbipaistvus üksnes andmekaitseõude täitmise vahend, vaid see aitab hoida usaldust avaliku võimu suhtes.

11. Tehnilised kaitsemeetmed ja nende piirid

Komitee peab asjakohaseks, et andmed pseudonümiseeritakse Statistikaametis, uurijad kasutavad neid piiratud ligipääsuga teadlaste keskkonnas ning kõik keskkonnast väljastatavad tulemid läbivad Statistikaameti konfidentsiaalsuskontrolli. Samuti taaskasutatakse võimalusel olemasolevaid andmeid, et vältida uusi päringuid. Need meetmed vähendavad isikute otsese tuvastamise, andmete lekkimise ja uuringukeskkonnast väljaviimise riski.

Turvaline töötlemiskeskond ei lahenda siiski uuringu keskseid eetilisi küsimusi. See ei vähenda ühendandmestiku sisulist ulatust ega vasta küsimusele, kas kogu elanikkonna kohta nii paljude eluvaldkondade andmete ühendamine on vajalik ja proportsionaalne. Samuti ei lahenda pseudonümiseerimine andmepõhiste leibkondade tuletamise ebakindlust, inimeste kohta uue eraelulise teabe loomise põhjendatust ega sellise teadmise tulevase kasutamise legitiimsust.

12. Kokkuvõtte võimalikest eetilistest riskidest

Järgnevalt hinnatakse uuringu eetilisi riske bioetikas laialt kasutatava nelja põhimõtte alusel: austus inimese autonoomia vastu, mittekahjustamine, heategemine ja õiglus¹. Samadest põhimõtetest lähtutakse „Inimuuringute hea tava“ kavandis², kus esimene põhimõte on sõnastatud laiemalt austusena inimeste vastu, hõlmates nende autonoomia ja privaatsuse austamist ning lugupidavat kohtlemist.

Autonoomia: taotluses kirjeldatud andmetöötlus riivab oluliselt inimeste informatsioonilist enesemääramisõigust, mille järgi peaks inimesel endal olema võimalikult suur kontrolliõigus enda kohta käivate andmete üle. Antud juhul ühendatakse inimeste kohta kokku ulatuslik ja detailne andmestik, ilma et inimesele oleks tagatud lihtne ja arusaadav võimalus andmetöötlust kontrollida, vaidlustada või parandada. Inimese vaba enesemääramist piirab ka kaugem eesmärk määratleda leibkond andmepõhiselt, mitte lasta inimesel ise määratleda oma leibkond.

Ilma inimese teadva nõusolekuta toimuv andmetöötlus on autonoomia põhimõttega olemuslikus vastuolus, sest otsus inimese kohta käivate andmete kasutamise üle tehakse ilma tema osaluseta. See ei tähenda, et nõusolekuta andmetöötlus oleks alati eetiliselt lubamatu, kuid sellist autonoomia piirangut peavad õigustama muud kaalukad hüved. Mida ulatuslikum ja eraelulisem on andmetöötlus ning mida väiksem on inimese võimalus seda mõista, kontrollida ja vaidlustada, seda tugevam peab olema töötlemise vajalikkuse ja proportsionaalsuse põhjendus.

Võimalikud kahjud (mittekahjustamine): plaanitav andmetöötlus on olulise privaatsuse riivega, kuna töödeldakse kõigi Eesti elanike andmeid väga suures ulatuses. Otseste tuvastamise, loata ligipääsu ja andmete lekkimise riske maandatakse tehniliste ja korralduslike turvameetmetega. Need meetmed ei vähenda siiski andmetöötluse sisulist ulatust ega lahenda eesmärgipärasuse, vajalikkuse, proportsionaalsuse ja inimeste kohta uue teabe tuletamise eetilisi küsimusi. Oluline privaatsuse riive on ka see, kui riik tuletab inimese kohta uut teadmist, mida inimene ise ei oleks nõus enda kohta avaldama. Isegi kui tegemist on inimese enda initsieeritud protsessiga – nt varasemate toetuste taotlemiste käigus saadud andmeid – on inimene andnud andmed eesmärgiga saada toetust, mida ta toimetulekuks vajab, mitte selleks, et riik saaks

¹ Beauchamp, T. L., & Childress, J. F. (2026). Principles of biomedical ethics (9th ed.). Oxford University Press.

² Eesti Teadusagentuur. (2025). Inimuuringute hea tava (kavand, viimati täiendatud 1. aprillil 2025).

kogutud andmeid igavesti säilitada ja kõikvõimalike muude andmetega siduda. See on kooskõlas ka isikuandmete töötlemise eesmärgi piirangu põhimõttega (vt arvamuse punkti 9).

Privaatsuse kahjustamise argument on tihedalt seotud eelpool käsitletud informatsioonilise enesemääramise piiratusena. Suurem läbipaistvus ja kontrolliõigus aitaks privaatsuse kahjustamist tasakaalustada, kuid antud uuringus ei ole seda piisavalt tehtud.

Võimalikud kasud (heategemine): plaanitaval andmetöötlusel ei ole vahetut kasu inimestele, kelle andmeid töödeldakse. Samas on võimalik uuringu raames loodud mudelite edukaks osutumise korral ümber korraldada senine toetuste andmise süsteem ning tervishoiu omaosaluse arvestamise alused, et see arvestaks inimeste tegeliku majandusliku olukorraga. Taotluses mainiti ka võimalust, et andmepõhine leibkonna mudel aitab tuvastada inimesi, kellel on õigus saada riigilt toetust, kuid kes ei ole seda ise küsinud.

Üldistatult võivad uuringu tulemused aidata kaasa järgnevale: 1) riigil tekib võimekus leida üles inimesed, kes päriselt abi vajaksid ja pakkuda neile toetusi; 2) toetuse taotlemine muutub inimesele lihtsamaks ja kiiremaks; 3) tervishoiuteenused saab muuta vaestele leibkondadele odavamaks.

Õiglus: riigi soov pakkuda teenuseid proaktiivselt tõstatab üldisema küsimuse riigi ja kodaniku suhtest. Proaktiivsus eeldab paratamatult, et riik ise teab inimeste kohta praegusega võrreldes rohkem – kes vajab millist abi ja kui palju – ning oskab seejärel ise inimesele toetuse välja pakkuda. Kuigi see aitab kindlasti jõuda nende haavatavate isikuteni, kes ise võib-olla ei teaks või oskaks riigilt toetust küsida, on selle hinnaks detailne teadmiste kogum kogu elanikkonna kohta, kellest enamus ei pruugi ulatuslikust privaatsuse riivist mingit kasu saada. Ühtlasi oleks see liikumine paternalistlikuma elukorralduse suunas: riik ei piirdu enam inimese esitatud taotluse lahendamisega, vaid tuvastab ise tema võimaliku abivajaduse, konstrueerib tema leibkonna ning algatab selle põhjal teenuse pakkumise. See võib suurendada abi kättesaadavust, kuid vähendab samal ajal inimese võimalust ise määratleda, millal ja millises ulatuses ta oma eraelulisi asjaolusid riigiga jagab.

Üldisem õigluse küsimus on ka see, kuidas ühiskond sellisesse muutusesse suhtub, kuivõrd on võimalikke alternatiive avalikult debateeritud ning kes ja millise legitiimsusega on teinud kogu ühiskonna nimel otsuse, et just sellise tuleviku suunas liigume.

13. Kokkuvõte ja soovitus

Uuring käsitleb olulisi sotsiaal- ja tervisepoliitika küsimusi, kuid eesmärgi olulisus ei põhjenda iseenesest kogu elanikkonda hõlmava detailse ühendandmestiku kasutamist. Esitatud materjalide põhjal ei ole võimalik veenduda, et kõik andmed on iga

uurimisülesande jaoks vajalikud ja proportsionaalsed. Peamised probleemid on andmekoosseisu erakordne ulatus ja osaline läbipaistmatus, leibkonna ja majandusliku seotuse tuletamine, detailsete terviseandmete küsitav vajalikkus ning inimese kontrolli vähenemine tema kohta loodava teabe üle.

Komitee soovib AKI-l paluda taotlejal enne kavandatud andmetöötluse alustamist:

1. eristada uuringu empiirilised eesmärgid ja nende andmevajadused, näidates eraldi, milliseid andmeid kasutatakse leibkondade konstrueerimiseks, toimetulekutoetuse modelleerimiseks ja tervishoiu omaosaluse analüüsiks;
2. esitada terviklik andmestiku kirjeldus, mis näitab algandmekogusid, varem loodud ühendandmestikke, käesolevas uuringus lisatavaid andmeid, andmete seostamise viisi ja uurijale kättesaadava lõpliku andmestiku koosseisu;
3. kaaluda uuringu läbiviimiseks vähem andmeid kasutavaid alternatiive, sealhulgas eri uurimisülesannete etapiti ja eraldi täitmist, valimipõhist analüüsi, väiksemat tunnuste koosseisu ning agregeeritud andmete kasutamist;
4. kirjeldada andmepõhiste leibkondade moodustamise reegleid ja ebakindlust, sealhulgas seda, kuidas eristatakse ühist aadressi tegelikust kooselust ja ühisest majandamisest ning kuidas hinnatakse ekslike leibkonnaseoste mõju;
5. põhjendada eraldi Tervisekassa ja Retseptikeskuse andmete detailsust, selgitades, miks on omaosaluse rahalise koormuse hindamiseks vaja diagnoosi-, retsepti-, raviarve- ja ravijuhutasandi andmeid, millistest detailandmetest on võimalik loobuda ning kuivõrd oleks võimalik uuringu eesmärgi saavutada agregeeritud andmetega;
6. selgitada kavandatava lahenduse õiguslikult realistlikku kasutusviisi, eristades poliitikaanalüüsi, abivajaduse võimaliku tuvastamise, teenuse proaktiivse pakkumise ja inimese õigusi mõjutava automatiseeritud otsustamise;
7. kirjeldada uuringut avalikkusele lihtsas ja terviklikus vormis, tuues välja, kelle andmeid kasutatakse, milliseid eluvaldkondi ühendandmestik hõlmab, milliseid uusi seoseid andmete põhjal tuletatakse ning milliste tulevaste lahenduste väljatöötamiseks uuringut tehakse.

Komitee hinnangul ei ole esitatud materjalide põhjal praegu piisavalt põhjendatud kogu taotletud andmekoosseisu kasutamine. Enne loa andmist tuleks eespool nimetatud küsimused lahendada ning näidata iga uurimisülesande puhul eraldi kavandatud andmetöötluse vajalikkust ja proportsionaalsust.

Käesolev arvamus puudutab uuringuks kavandatud andmetöötlust. See ei laiene registripõhise leibkonnamudeli, automatiseeritud abivajaduse hindamise ega personaliseeritud või proaktiivse teenuse hilisemale kasutuselevõtule. Selline kasutus

oleks eraldiseisev ja oluliselt suurema mõjuga tegevus, mis vajaks omaette õiguslikku alust ning eetilise ja ühiskondliku mõju hindamist.

Lugupidamisega

(allkirjastatud digitaalselt)

Kadi Krinal

Teaduseetika komitee terviseandmete ja andmeteaduse alamkomitee esimees

(allkirjastatud digitaalselt)

Marten Juurik

teaduseetika valdkonna juht

Eesti Teadusagentuur