|  |
| --- |
| **UURINGU EETILISE HINDAMISE TAOTLUS EESTI BIOEETIKA JA INIMUURINGUTE NÕUKOGULE** |

Taotlusele 06.05.2024 lisatud täiendused on märgitud kollase värviga.

**1. UURIMISTÖÖ ANDMED**

Uurimistöö täielik nimetus: **TEHISINTELLEKTI KASUTAMISE VÕIMALUSED MEDITSIINIS**

Uurimistöö toimumiskoht: Narva mnt 18, Tartu (Tartu Ülikool, arvutiteaduse instituut)

**2. UURIMISTÖÖ TEOSTAJAD JA UURINGUKESKUSED**

2.1. Vastutav uurija

ees- ja perekonnanimi: **Jaak Vilo**  
teaduslik kraad: PhD  
amet: bioinformaatika professor  
töökoht: Tartu Ülikool, arvutiteaduse instituut  
töökoha aadress: Narva mnt 18, Tartu  
telefoninumber: +372 5049365

e-post: [vilo@ut.ee](mailto:vilo@ut.ee)   
allkiri: allkirjastatud digitaalselt

ees- ja perekonnanimi: **Sven Laur**  
teaduslik kraad: PhD  
amet: kaasprofessor

töökoht: Tartu Ülikool, arvutiteaduse instituut  
töökoha aadress: Narva mnt 18, Tartu  
telefoninumber: +372 5309 0584

e-post: [sven.laur@ut.ee](mailto:sven.laur@ut.ee)   
allkiri: allkirjastatud digitaalselt

2.2. Kaastöötajad

ees- ja perekonnanimi: **Sulev Reisberg**  
amet: terviseinformaatika teadur   
töökoht: Tartu Ülikool, arvutiteaduse instituut   
töökoha aadress: Narva mnt 18, Tartu  
telefoninumber:+372 5248123

e-post: [sulev.reisberg@ut.ee](mailto:sulev.reisberg@ut.ee)

ees- ja perekonnanimi: **Raivo Kolde**

amet: terviseinformaatika kaasprofessor  
töökoht: Tartu Ülikool, arvutiteaduse instituut  
töökoha aadress: Narva mnt 18, Tartu  
telefoninumber:+372 506 7961

e-post: [raivo.kolde@ut.ee](mailto:raivo.kolde@ut.ee)

ees- ja perekonnanimi: **Harry-Anton Talvik**  
amet: terviseinformaatika nooremteadur (juhendaja Sulev Reisberg)  
töökoht: Tartu Ülikool, arvutiteaduse instituut  
töökoha aadress: Narva mnt 18, Tartu

e-post: harry-anton.talvik@ut.ee

ees- ja perekonnanimi: **Marek Oja**  
amet: terviseinformaatika teadur  
töökoht: Tartu Ülikool, arvutiteaduse instituut  
töökoha aadress: Narva mnt 18, Tartu  
telefoninumber:+372 58002684

e-post: marek.oja@ut.ee

ees- ja perekonnanimi: **Sirli Tamm**  
amet: programmeerija (juhendaja Marek Oja)  
töökoht: Tartu Ülikool, arvutiteaduse instituut  
töökoha aadress: Narva mnt 18, Tartu

e-post: sirli.tamm@ut.ee

ees- ja perekonnanimi: **Kerli Mooses**   
amet: terviseinformaatika teadur  
töökoht: Tartu Ülikool, arvutiteaduse instituut  
töökoha aadress: Narva mnt 18, Tartu  
e-post: ​[kerli.mooses@ut.ee](mailto:kerli.mooses@ut.ee)

ees- ja perekonnanimi: **Valerija Jerina**amet: Tartu Ülikooli bakalaureusetudeng (juhendaja Sven Laur)  
e-post: ​[shionkaiko121@gmail.com](mailto:shionkaiko121@gmail.com)

ees- ja perekonnanimi: **Markus Haug**   
amet: terviseinformaatika nooremteadur (juhendaja Raivo Kolde)  
töökoht: Tartu Ülikool, arvutiteaduse instituut  
töökoha aadress: Narva mnt 18, Tartu  
e-post: [markus.haug@ut.ee](mailto:markus.haug@ut.ee)

ees- ja perekonnanimi: **Maarja Pajusalu**  
amet: terviseinformaatika nooremteadur (juhendaja Raivo Kolde)  
töökoht: Tartu Ülikool, arvutiteaduse instituut  
töökoha aadress: Narva mnt 18, Tartu  
e-post: [maarja.pajusalu@ut.ee](mailto:maarja.pajusalu@ut.ee)

ees- ja perekonnanimi: **Marianna Laht**  
amet: Tartu Ülikooli personaalmeditsiini spetsialist (juhendaja Raivo Kolde)  
töökoht: Tartu Ülikool, arvutiteaduse instituut  
töökoha aadress: Narva mnt 18, Tartu  
e-post: [marianna.laht@ut.ee](mailto:marianna.laht@ut.ee)

ees- ja perekonnanimi: **Ida Maria Orula**

amet: terviseinformaatika nooremteadur (juhendaja Sulev Reisberg)  
töökoht: Tartu Ülikool, arvutiteaduse instituut  
töökoha aadress: Narva mnt 18, Tartu

e-post: [ida.maria.orula@ut.ee](mailto:ida.maria.orula@ut.ee)

ees- ja perekonnanimi: **Elena Sügis**  
amet: terviseinformaatika teadur  
töökoht: Tartu Ülikool, arvutiteaduse instituut  
töökoha aadress: Narva mnt 18, Tartu   
e-post: [elena.sugis@ut.ee](mailto:elena.sugis@ut.ee)

ees- ja perekonnanimi: **Kaire Koljal**  
amet: programmeerija (juhendaja Sven Laur)  
töökoht: Tartu Ülikool, arvutiteaduse instituut  
töökoha aadress: Narva mnt 18, Tartu  
e-post: [kaire.koljal@ut.ee](mailto:kaire.koljal@ut.ee)

ees- ja perekonnanimi: **Hendrik Šuvalov**  
amet: terviseinformaatika nooremteadur (juhendaja Raivo Kolde)  
töökoht: Tartu Ülikool, arvutiteaduse instituut  
töökoha aadress: Narva mnt 18, Tartu  
e-post: [hendrik.suvalov@ut.ee](mailto:hendrik.suvalov@ut.ee)

ees- ja perekonnanimi: **Egert Vinogradov**  
amet: Tartu Ülikooli magistrant (juhendaja Marek Oja)  
e-post: [egert.vinogradov@ut.ee](mailto:egert.vinogradov@ut.ee)

ees- ja perekonnanimi: **Grete Mägi**  
amet: Tartu Ülikooli magistrant (juhendaja Marek Oja)  
e-post: [grete.magi@ut.ee](mailto:grete.magi@ut.ee)

ees- ja perekonnanimi: **Õie Renata Siimon**  
amet: terviseinformaatika nooremteadur (juhendaja Raivo Kolde)  
töökoht: Tartu Ülikool, arvutiteaduse instituut  
töökoha aadress: Narva mnt 18, Tartu  
e-post: [renata.siimon@ut.ee](mailto:renata.siimon@ut.ee)

ees- ja perekonnanimi: **Päivi Ojala**  
amet: Tartu Ülikooli magistrant (juhendaja Sulev Reisberg)

ees- ja perekonnanimi: **Taavi Tillmann**  
amet: Tartu Ülikooli rahvatervishoiu kaasprofessor

ees- ja perekonnanimi: **Laura Ilves**  
amet: Tartu Ülikooli magistrant (juhendaja Marek Oja)

ees- ja perekonnanimi: **Eve-Maarja Mänd**  
amet: Tartu Ülikooli magistrant (juhendaja Raivo Kolde)

ees- ja perekonnanimi: **Angela Kannukene**  
amet: Tartu Ülikooli nooremteadur (juhendaja Sulev Reisberg)

ees- ja perekonnanimi: **Kati Koido**  
amet: Tartu Ülikooli kaasprofessor (juhendaja Kerli Mooses)

ees- ja perekonnanimi: **Birgitt Majas**  
amet: Tartu Ülikooli magistrant (juhendaja Markus Haug)

ees- ja perekonnanimi: **Kunnar Kukk**  
amet: terviseinformaatika nooremteadur (juhendaja Sulev Reisberg)  
töökoht: Tartu Ülikool, arvutiteaduse instituut  
töökoha aadress: Narva mnt 18, Tartu  
e-post: [kunnar.kukk@ut.ee](mailto:kunnar.kukk@ut.ee)

ees- ja perekonnanimi: **Telver Objärtel**  
amet: Tartu Ülikooli magistrant (juhendaja Raivo Kolde)

ees- ja perekonnanimi: **Kermo Saarse**  
amet: Tartu Ülikooli magistrant (juhendaja Sulev Reisberg)

ees- ja perekonnanimi: **Oliver-Erik Suik**  
amet: Tartu Ülikooli magistrant (juhendaja Sulev Reisberg)

ees- ja perekonnanimi: **Belinda Lepmets**  
amet: Tartu Ülikooli bakalaureusetudeng (juhendaja Sulev Reisberg)  
  
ees- ja perekonnanimi: **Robin Piir**  
amet: Tartu Ülikooli bakalaureusetudeng (juhendaja Õie Renata Siimon)  
  
perekonnanimi: **Andres Püss**  
amet: Tartu Ülikooli bakalaureusetudeng (juhendaja Õie Renata Siimon)  
  
ees- ja perekonnanimi: **Kristiina Miller**  
amet: Tartu Ülikooli magistrant (juhendaja Marek Oja)

ees- ja perekonnanimi: **Fedor Stomakhin**  
amet: Tartu Ülikooli magistrant (juhendaja Sven Laur)

ees- ja perekonnanimi: **Grete Kelli Aava**amet: Tartu Ülikooli bakalaureusetudeng (juhendaja Markus Haug)

ees- ja perekonnanimi: **Helene Loorents**amet: Tartu Ülikooli bakalaureusetudeng (juhendaja Raivo Kolde)

ees- ja perekonnanimi: **Anett Sandberg**amet: Tartu Ülikooli magistrant (juhendaja Raivo Kolde)

ees- ja perekonnanimi: **Teele Kuri**amet: Tartu Ülikooli magistrant (juhendaja Raivo Kolde)

ees- ja perekonnanimi: **Margus Lehtme**amet: Tartu Ülikooli magistrant (juhendaja Raivo Kolde)

ees- ja perekonnanimi: **Mihhail Sokolov**amet: Tartu Ülikooli magistrant (juhendaja Raivo Kolde)

ees- ja perekonnanimi: **Hanna Keidong**  
amet: Tartu Ülikooli spetsialist (juhendaja Kerli Mooses)

ees- ja perekonnanimi: **Nikita Umov**  
amet: rahvatervishoiu nooremteadur (juhendaja Taavi Tillmann)  
töökoht: Tartu Ülikool, Tartu Ülikooli peremeditsiini ja rahvatervishoiu instituut  
e-post: [nikita.umov@ut.ee](mailto:nikita.umov@ut.ee)

ees- ja perekonnanimi: **Ave Põld**  
amet: rahvatervishoiu nooremteadur (juhendaja Mikk Jürisson)  
töökoht: Tartu Ülikool, Tartu Ülikooli peremeditsiini ja rahvatervishoiu instituut  
e-post: [ave.pold@ut.ee](mailto:ave.pold@ut.ee)

ees- ja perekonnanimi: **Ami Sild**  
amet: Tartu Ülikooli magistrant (juhendaja Raivo Kolde)

ees- ja perekonnanimi: **Margus Lehtme**  
amet: Tartu Ülikooli magistrant (juhendaja Raivo Kolde)

ees- ja perekonnanimi: **Andra Bluum**  
amet: Tartu Ülikooli bakalaureusetudeng (juhendaja Sven Laur)

ees- ja perekonnanimi: **Annika Talvet**  
amet: Tartu Ülikooli magistrant (juhendaja Sven Laur)

ees- ja perekonnanimi: **Kristel Agu**  
amet: Tartu Ülikooli magistrant (juhendaja Sulev Reisberg)

ees- ja perekonnanimi: **Kai Tiitsaar**  
amet: Tartu Ülikooli magistrant (juhendaja Marek Oja)

ees- ja perekonnanimi: **Stella Lelov**amet: Tartu Ülikooli rahvatervise magistrant (juhendaja Raivo Kolde)

ees- ja perekonnanimi: **Helen Reimand**amet: Tartu Ülikooli rahvatervise magistrant (juhendaja Raivo Kolde)

ees- ja perekonnanimi: **Neeme Ilves**amet: kliiniline spetsialist (juhendaja Raivo Kolde)

Lisaks on plaanis Tartu Ülikooli lõputööde kirjutamiseks kaasata meeskonda Tartu Ülikooli tudengeid (magistrandid ja doktorandid).

2.3.Vastutava uurija asutuse juhataja või tema kohusetäitja

on uurimistöö korraldamisega nõus:

ees- ja perekonnanimi: **Leho Ainsaar**

amet: Dekaan, TÜ Loodus- ja Täppisteaduste valdkond.

töökoht: Tartu Ülikool

töökoha aadress: Vanemuise 46, Tartu

allkiri: allkirjastatud digitaalselt

**3. UURIMISTÖÖ FINANTSEERIMINE**

Allikas: **RITA1/02-96 projekt „Tehisintellekti kasutamise võimalused e-valitsemises“ (teenusleping nr 7.2-2/19/11)**

Uurimistöö üldmaksumus, sh töötasude jaotus uuringu teostajatele (kellele ja millises ulatuses): **242160 eurot. Töötasude planeeritud jaotus:**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Tartu Ülikool | STACC | Cybernetica |
| Vastutavad uurijad | 13 300 | 36 000 | 15 000 |
| Kaastöötajad | 56 700 | 39 000 | 10 000 |
| Kaudsed kulud | 13 650 | 11 250 | 6 900 |
| Käibemaks | 16 730 | 17 250 | 6 380 |
| KOKKU | 100 380 | 103 500 | 38 280 |

**4. KAVANDATAVA UURIMISTÖÖ TEADUSLIK EESMÄRK JA PÕHJENDUS**

Eesti avalik sektor on keskmisest väiksem, kuid selle koormus on kõrge, sest riigi ja kohalike omavalitsuste pakutavad teenused peavad olema samaväärsed. Praegu pole teenuste pakkumine Eestis optimaalselt korraldatud. Teenuste kvaliteedi parandamiseks ametnike arvu suurendamata on üks võimalus rakendada tehisintellekti ehk krattide võimalusi.

Kratte juba kasutatud avalikus sektoris lihtsamate protsesside automatiseerimisel. See on aidanud tõsta teenuste kvaliteeti ja vähendada bürokraatiat. Taustal toimivaks riigiks on aga vaja muuta ka keerulisemaid protsesse. Järgmise sammuna nähaksegi vajadust põhjalikult uurida ja rakendada kratte ka keerulisemates otsustusprotsessides[[1]](#footnote-1).

Selleks, et viia e-riigi areng uuele tasemele – taustal toimivaks, sündmuspõhiseks ja teadmuspõhistest otsustest lähtuvaks, korraldas Eesti Teadusagentuur (ETAG) 2019. a kevadel välja konkursi uuringu „Tehisintellekti kasutamise võimalused e-valitsemises“ läbiviimiseks, mille lähteülesande pani kokku Majandus- ja Kommunikatsiooniministeerium ning Riigikantselei[[2]](#footnote-2). Uuringu eesmärk on analüüsida tehisintellekti rakendatavust avaliku sektori asutustes otsuste automatiseerimiseks. Vaatluse all on ka praegused kitsaskohad avalikus sektoris. Loodud tehisintellekti lahendused tuleb uuringu läbiviijal piloteerida ja valideerida.

Konkursi võitis konsortsium, millesse kuuluvad Tartu Ülikool, Tallinna Tehnikaülikool, Cybernetica AS ja STACC OÜ. Konsortsiumi eesmärk on luua avaliku sektori andmetel põhinevate teenuste prototüübid ja need rakendada, et hinnata kasu ning kaardistada piirangud. Konsortsium piloteerib masinõppe võimalusi neljas valdkonnas: tööturg, päästeamet, meditsiin, küberkaitse.

Käesolev uuring käsitleb tehisintellekti võimaluste rakendamist meditsiinis. Eestis on võrdlemisi heal tasemel riiklikud terviseandmeid sisaldavad andmekogud (raviarved, digiretsept, tervise infosüsteem, terviseregistrid jne), kuid nende omavaheline ühendamine, mis tehisintellekti rakendamiseks pakuks kõige efektiivsemaid võimalusi, on seni olnud komplitseeritud – seda tehniliselt, aga ka õiguslikust vaatepunktist. Esiteks on andmekogud erineva struktuuri ja ülesehitusega. Teiseks, mitmed elektroonsed tervisedokumendid (nt Digiloos) sisaldavad lisaks struktureeritud infole ka vabatekstilisi andmeid, mille analüüs nõuab suurt käsitööd. Seetõttu on patsiendi terviseseisundist tervikliku ülevaate saamine komplitseeritud ning nende andmete masinõppeline analüüs äärmiselt keeruline. Käesoleva uuringu raames **demonstreeritakse Eesti riigiasutustele** (majandus- ja kommunikatsiooniministeerium, sotsiaalministeerium, Eesti Haigekassa, Tervise- ja Heaolu Infosüsteemide Keskus), **milliseid masinõppelisi lahendusi meditsiiniliste andmekogude ühendatud uurimiseks on olemas ja millist kasu nende rakendamiseks võiks patsient ja Eesti riik tervikuna sellest saada.**

Täpsemalt, uuringus lingitakse (ühendatakse) eri meditsiinilistest andmekogudest pärit pseudonüümitud info ja viiakse ühtsele rahvusvaheliselt tunnustatud OHDSI OMOP andmekujule[[3]](#footnote-3), rakendatakse erinevaid andmeanalüüsi, s.h masinõppe, tehisintellekti ja privaatsust säilitavaid metoodikaid, ning demonstreeritakse ning analüüsitakse nende metoodikate rakendatavust ja rakendamisest saadavat kasu. Uuringu üheks osaks on ka õigusruumi analüüs – kas selliste lahenduste rakendamine igapäevatöösse nõuaks õigusruumi muudatusi. Täpsem uuringu metoodika on kirjeldatud ptk 7.

Uuringu meeskonnal on pikaaegne teadustöö kogemus selles vallas. Tartu Ülikool on osalenud/osaleb mitmes Euroopa Liidu projektis (IMI EMIF, IMI2 EHDEN), mille sisuks on OMOP andmekuju rakendamine Euroopa terviseandmetel. Tartu Ülikoolis õpetatakse andmeteadust, masinõppe- ja tehisintellektitehnoloogiaid. STACC OÜ omab pikaaegset terviseandmete analüütika kogemust – nad on Eestis analüüsinud teadaolevalt kõige suuremas mahus Digiloo andmeid (Reisberg 2013; Reisberg 2019), samuti Haigekassa raviarvete andmeid (Tamm; Liivlaid 2018; Liivlaid 2019). STACC OÜ on omandanud 2019. a detsembris ka EHDEN sertifikaadi selle kohta, et nad oskavad teostada andmete kvaliteetset teisendamist OMOP andmekujule[[4]](#footnote-4). Cybernetica AS omab pikaaegset kogemust privaatsust säilitavate metoodikate ja lahenduste arenduse ja publitseerimise vallas.

Kavandatava uuringu eesmärgid ühtivad **RITA1/02-96** ja **Eesti nutika spetsialiseerumise** projektieesmärkidega, samuti **Eesti teadus- ja arendustegevuse ning innovatsiooni strateegiaga 2014-2020 „Teadmistepõhine Eesti“[[5]](#footnote-5)** ning **Eesti digitaalse agendaga 2020[[6]](#footnote-6)**.

Viited:

*Reisberg, Sulev (2019). Developing computational solutions for personalized medicine. (Doktoritöö, Tartu Ülikool). Tartu: Tartu Ülikooli Kirjastus.*

*Reisberg, S; Sirel, R; Kalda, R; Merzin, M; Pruulmann, J; Vilo, J (2013). Elektrooniliste terviselugude analüüsimise võimalused Tartu perearstide infosüsteemi näitel. Eesti Arst, 92 (8), 452−459.*

*Liivlaid, Hedi; Eigo, Natalja; Reisberg, Sulev (2019). Eriarstiabi haigestumusstatistika võrdlus Tervise Arengu Instituudi ja Eesti Haigekassa andmetel. Eesti Arst, 1, 17−26.*

*Sirli Tamm, magistrikraad, 2019, (juh) Marek Oja; Sulev Reisberg, Eesti eelkooliealiste laste terviseseisund ja tervisekäitumine sünnist kuni 7aastaseks saamiseni 2010. aasta sünnikohordi põhjal Eesti Haigekassa raviarvete alusel, Tartu Ülikool, Loodus- ja täppisteaduste valdkond, arvutiteaduse instituut.*

*Hedi Liivlaid, magistrikraad, 2018, (juh) Sulev Reisberg; Natalja Jedomskihh-Eigo, Eriarstiabi haigestumusstatistika võrdlus Tervise Arengu Instituudi ja Eesti Haigekassa andmestikes, Tartu Ülikool, Loodus- ja täppisteaduste valdkond, arvutiteaduse instituut.*

**5. UURIMISTÖÖ TEOSTAMISE AEG**

*Veebruar 2020-juuni 2025*

**6. LOODAVA ANDMEBAASI STRUKTUURI KIRJELDUS**

Uuringus kasutatakse 10% juhuvalimi kohta (vt täpsemalt ptk 7) järgmisi andmeid:

1. Eesti Haigekassa raviarved:
   1. Raviarve number (id)
   2. Patsiendi pseudonüüm
   3. Patsiendi sünniaasta
   4. Patsiendi sugu
   5. TTO registrikood
   6. Haigla liigi tunnus (kood)
   7. Arve esitanud arsti eriala
   8. Arve alustatud (raviarve alguskuupäev)
   9. Arve lõpetatud (raviarve lõpukuupäev)
   10. Põhieriala kood
   11. Järgarve (jah/ei)
   12. Ravitüüp
   13. Raviarve lõpetamise põhjuse kood
   14. Arve diagnoosid:
       1. Diagnoosi kood ja nimetus RHK-10 järgi
       2. Diagnoosi tunnus
   15. Teenused:
       1. Tervishoiuteenuse kood ja nimetus
       2. Hulk
       3. Hulk2
       4. Kordi
       5. Kogus
       6. Teenuse osutamise kuupäev
       7. Material group ehk teenusgrupi kood ja nimetus
       8. Arve rea järjekorranumber
   16. Protseduurid:
       1. NCSP kood ja nimetus
       2. Protseduuri kuupäev
       3. Protseduuri rea järjekorranumber
2. Eesti Haigekassa digiretseptid:
   1. isiku pseudonüüm
   2. Retsepti üldandmed (number, koostamise aeg, kehtivuse aeg, liik)
   3. Patsient (pseudonüüm, sünniaasta)
   4. Arsti eriala kood
   5. TTO info
   6. Ravi info (kõik väljad)
   7. Väljastaja info (müümise aeg, apteegi kood)
3. TEHIK Digilugu CDA-dokumendid (epikriisid, saatekirjad, saatekirja vastused):
   1. Dokumendi üldandmed (haigusjuhtumi liik, dokumendi tüüp, number, versioon, kuupäev)
   2. Tervishoiuasutus
   3. Koostaja eriala
   4. Patsiendi pseudonüüm, sünniaasta, sugu
   5. Haigusjuhtum
   6. Surm
   7. Lõplik kliiniline diagnoos
   8. Kliiniline diagnoos
   9. Anamnees
   10. Ravimid
   11. Ravisoovitused
   12. Allergia
   13. Objektiivne leid
   14. Uuringud ja protseduurid, operatsioonid, analüüsid
   15. Immuniseerimine
   16. Kõrvalnähud
   17. Kokkuvõte patsiendi ravist
   18. Seisund väljakirjutamisel
   19. Vastuvõtule pöördumise aeg
   20. Dokumendi koostamise kuupäev
   21. Vastutaja asutuse nimi ja kood
4. Uuringu käigus teisendatakse ülalloetletud andmed OMOP andmekujule (kirjeldatud siin: <https://github.com/OHDSI/CommonDataModel/wiki/Standardized-Clinical-Data-Tables>), sama patsiendi andmed on patsiendi pseudonüümi kaudu lingitavad.

**7. METOODIKA (VALIMI KIRJELDUS)**

Uuring koosneb kahest osast:

1. Analüüsitavate terviseandmete teisendamine OMOP andmekujule ja selle protsessi automatiseerimise võimaluste uurimine. Selleks kasutatakse olemasolevaid teisendustööriistu White Rabbit ja Rabbit In a Hat (<https://github.com/OHDSI/WhiteRabbit>) ja OMOP sõnastikke, samuti arendatakse ise välja masinõppel põhinevaid tööriistu.
2. OMOP kujul andmete automaatsel analüüsil põhinevate ja privaatsust säilitavaid meetodeid kasutavate otsustustugede võimekuse uuring, käsitledes järgmisi patsiendi ravi kvaliteediga seotud küsimusi:
   1. Kuidas jälgitakse Haigekassa poolt ja Eesti Perearstide Seltsi poolt publitseeritud ravijuhiseid (<https://www.ravijuhend.ee/tervishoiuvarav/juhendid>)? Selleks võrreldakse ravijuhistes nimetatud tegevusi tegelike terviseandmetega.
   2. Kuidas jälgitakse ravimite kasutamisel ravimite infolehtedel toodud riski minimeerimise juhiseid? Ca 200 ravimil on infolehel kirjas, et ravimi tarvitamisel on vajalikud regulaarsed maksanäitajate kontrollid. Ravimite väljakirjutamine on kirjas Digiretsepti andmetes, aga raviarvete alusel saaks hinnata, kas ka vastavad kontrollid on tehtud. Et andmete kokkupanek on keerukas, ei ole seda seni olnud võimalik automaatselt jälgida ja seni on seda teadaolevalt uuritud üldse Eestis vaid ühes uuringus ühele ravimile (Kurvits) . Käesolevas uuringus analüüsitakse selliste analüüside automatiseerimisvõimalusi.
   3. Kas terviseandmetest on võimalik automaatselt tuvastada ravimite kõrvaltoimeid? Kõrvaltoimed on Eestis alaraporteeritud - ühelt poolt, kuna see on arstile tülikas, teiselt poolt aga võib olla seotud ka arstide hirmudega. Samas on kõrvaltoimete tuvastamine andmete teisesel analüüsil väga oluline ja on seotud ka ravimite ohutusega. Samas, infot kõrvalmõjude kohta võib leida tervisedokumentide diagnoosikoodidest, allergia blokist, samuti vabatekstilistest osadest ning neid võib kinnitada nt kaasuv ravimi vahetus. Seetõttu rakendatakse käesolevas uuringus andmetel erinevaid masinõppelisi mudeleid, et kõrvaltoimeid automaatselt tuvastada.
   4. OMOP kujul andmetel tehtud rahvusvaheliste uuringute replitseerimine. Universaalne andmekuju loob unikaalse võimaluse replitseerida mujal läbi viidud uuringuid, kasutades nendes uuringutes kasutatud (publitseeritud) arvutusalgoritme ning hinnata nende valiidsust/rakendatavust eesti populatsioonis. Käesolevas uuringus plaanitakse replitseerida hiljutist suuremahulist OMOP kujul andmetel tehtud uuringut, kus uuriti erinevate vererõhuravimite efektiivsust (Suhard).
   5. Võimalusel ka muud küsimused, mille tehisintellektiga lahendamine võiksid potentsiaalselt Eestile avalikule sektorile kasu tuua

Uurimiseks on vajalik järgmine andmestik:

* 10% juhuvalim (isikukoodi kontrollnumbri põhjal)
* Valimi kohta kõik peatükis 6 nimetatud andmeväljad ajavahemikust 2012-2019

Uurimistöö tulemused tehakse kättesaadavaks ptk 4 nimetatud asutustele, samuti on plaanitud neid avaldada Tartu Ülikooli lõputööde, teaduspublikatsioonidena. Kõik avaldatud tulemused on statistilist laadi (üksikindiviidi tasemel tulemusi ei avaldata). Uurimistöö tulemusena tekkivaid mudeleid (ei sisalda andmeid) on perspektiivis võimalik hiljem juurutada Eesti riigi tasemel automatiseeritud krattidena.

*Viited:*

*Kurvits K, Uusküla M, Laius O. Agomelatiinraviga seotud riskivähendamise meetmed praktikas. Eesti Arst. 2016 Nov 25;95(10):687-9.*

*Suchard MA, Schuemie MJ, Krumholz HM, You SC, Chen R, Pratt N, Reich CG, Duke J, Madigan D, Hripcsak G, Ryan PB. Comprehensive comparative effectiveness and safety of first-line antihypertensive drug classes: a systematic, multinational, large-scale analysis. The Lancet. 2019 Nov 16;394(10211):1816-26.*

**8. UURIMISTÖÖS KOGUTAVATE ISIKUANDMETE KAITSE KIRJELDUS**

**Andmete Haigekassast väljastamise protseduur**

* Andmete väljastamisel järgivad Haigekassa ja TEHIK kõiki seadusest tulenevaid nõudeid.
* Väljastuse eelduseks on TÜ eetikakomitee luba
* Väljastatakse vaid eetikakomitee taotluses nimetatud andmed
* Väljastatakse 10% juhuvalim, mis määratakse nõnda: valimisse kuuluvad isikud, kelle isikukoodi viimane number ehk kontrollnumber on 5.
* Isikuandmeid uurimismeeskonnale ei väljastata. Isikuandmed (isikukoodid, täpsed sünniajad) eemaldatakse ja asendatakse pseudonüümiga, et sama isikuga seotud andmeid oleks võimalik uuringugrupil omavahel kokku viia. Asendamine toimub Haigekassa ja TEHIK poolt enne andmete uurimismeeskonnale edastamist. Selleks lepivad Haigekassa ja TEHIK omavahel kokku isikukoodil põhineva salajase pseudonüümimisvõtme, mis on tagasipööramatu ja pärast pseudonüümimist hävitatakse. Nii puudub võimalus uurimistöös kasutatavaid andmeid uuesti isikustada (tagasi kodeerida).
* Haigekassa ja TEHIK pseudonüümivad ja edastavad uurimisgrupile pseudonüümitud andmed teineteisest sõltumatult
* Nii Haigekassa kui TEHIK koostavad väljastamisprotokolli, kuhu märgitakse väljastatud andmed ning üleandmise viis ja kuupäev

**Uurimistöö teostaja kohustused andmete töötlemisel**

* Uurimistöö läbiviija on kohustatud järgima kõiki seadusest tulenevaid ja väljastamise korralduses toodud andmekaitse põhimõtteid.
* Uurimistöö teostaja kasutab andmeid ainult uurimistöö eesmärkidel ja ulatuses
* Uurimistöö teostaja garanteerib andmete konfidentsiaalsuse.
* Uurimistöö teostaja hävitab andmed peale uurimistöö valmimist või väljastamise korralduses toodud andmete kasutamise tähtpäeva saabumisel.

Andmeid hoitakse TÜ Teaduspilves, ning ei ole selle raames kättesaadavad kõrvalistele isikutele. TÜ Teaduspilv on läbinud ISKE auditi 12.2017 ja vastab ISKE M taseme nõuetele (turvaklass T1K1S2). Eriti tundlike andmete puhul kasutatakse ka täiendavaid turvameetmeid - andmete krüpteerimist ja protsesside pidevat monitooringut. Kõik meeskonnaliikmed on allkirjastanud konfidentsiaalsuslepingu ja informeeritud andmetele kohanduvatest andmekaitse nõuetest.

Nagu eespool mainitud, tehakse igale uurijale kättesaadavaks andmed vastavalt minimaalsuse printsiibile ehk ainult sellele lõigule, mis konkreetselt tema tööks vajalik on. Nii välditakse olukordi, et kõik uurijad pääseks ligi kõikidele andmetele. Andmete täpse koosseisu otsustab projekti juht kaasates selleks asutuste vastutavaid esindajaid (Vilo, Reisberg, Laur) ning serverite administraatoreid TÜ teadusarvutuste keskusest.

**9. TEAVE SAMA PROJEKTI VARASEMATEST VÕI SAMAAEGSETEST HINDAMISTEST VÕI KOOSKÕLASTAMISTEST MUJAL**

**10. LISAD**

**10.1 Andmetabeli struktuur (vt ptk 6)**

**10.2** **Uuringu läbiviija(te) CV(d) (lisatud eraldi dokumentidena)**

1. <https://www.etag.ee/teadlased-asuvad-uurima-tehisintellekti-kasutamise-voimalusi-avalike-teenuste-pakkumisel/> [↑](#footnote-ref-1)
2. <https://www.etag.ee/uritus/tehisintellekti-kasutamise-voimalused-e-valitsemises-konkursi-lopptahtaeg/> [↑](#footnote-ref-2)
3. <https://github.com/OHDSI/CommonDataModel/wiki> [↑](#footnote-ref-3)
4. <https://www.ehden.eu/ehden-certifies-the-next-6-smes/> [↑](#footnote-ref-4)
5. <https://www.hm.ee/sites/default/files/59705_teadmistepohine_eesti_est.pdf> [↑](#footnote-ref-5)
6. <https://www.mkm.ee/sites/default/files/digitalagenda2020_final.pdf> [↑](#footnote-ref-6)