TARTU ÜLIKOOL

Majandusteaduskond

Robert Pirk

ARENEVATE TEHNOLOOGIATE KASUTUSELEVÕTT – INNOVATSIOON EESTI AUDIITORETTEVÕTJATES

Magistritöö

Juhendajad: Ulvi Sloog (MBA)

professor Urmas Varblane

Tartu 2025

Olen koostanud töö iseseisvalt. Kõik töö koostamisel kasutatud teiste autorite tööd, põhimõttelised seisukohad, kirjandusallikatest ja mujalt pärinevad andmed on viidatud.

Sisukord

Lühikokkuvõte 4

1. Sissejuhatus 4

2. Kirjanduse ülevaade 8

2.1 Innovatsiooni uurimine, innovatsiooni mõiste 8

2.2 Varasemad uuringud audiitorteenuste valdkonnas 11

2.3 Protsessiteooria täiustatud andmeanalüüsi tehnoloogiate kasutuselevõtu kohta audiitorettevõtjates 15

3. Uurimisobjekt, metoodika ja andmed 19

4. Tulemused ja analüüs 29

4.1 Ülevaade vastustest 29

4.2 Eesti audiitorettevõtjate innovaatilisuse põhinäitajad 30

4.3 Arenevate tehnoloogiate kasutuselevõtt 33

5. Arutlus, järeldused 41

Kokkuvõte 47

Viidatud allikad 50

Lisad 54

LISA A 54

LISA B 63

LISA C 68

LISA D 69

LISA E 70

LISA F 71

Summary 131

Lühikokkuvõte

Käesoleva uurimuse eesmärgiks on anda hinnang innovatsioonile Eesti audiitorettevõtjates läbi arenevate tehnoloogiate kasutuselevõtu ja anda soovitusi selle protsessi edendamiseks. Täpsemalt antakse ülevaade Eesti audiitorettevõtjate (AEV) innovaatilisuse põhinäitajatest ning analüüsitakse arenevate tehnoloogiate (AT) kui efektiivsust tõstvate lahenduste kasutuselevõtu hetkeseisu, tuleviku väljavaateid, potentsiaalseid kasutegureid ning takistavaid tegureid protsessiteooria kontekstis. Selle saavutamiseks on läbi viidud e-maili teel valikvastustega küsitlus Eesti AEV-de seas (kokku 116), kust saadi 34 vastust, mida analüüsiti tarkvaras R ja RStudio. Uuringu tulemused näitavad, et peamiselt tegelevad AEV-d protsessiinnovatsioonidega, innovatsioonide arvu piiravad peamiselt vajaduse ja ressursside puudus. AT-de kasutuseni on jõudnud vähem kui pool AEV-dest, ilmneb mahajäämus väiksemate AEV-de poolt. Suuremate ja väiksemate AEV-de vahel esineb seoses AT-de kasutuselevõtu protsessiga veel mitmeid erinevusi. Kujundatud on soovitused AT-de kasutuselevõtu protsessi edendamiseks Eesti AEV-des, millest olulisemad on seotud AT-de alaste teadmiste ning omavahelise ja IKT sektoriga koostöö edendamisega eriti väiksemate AEV-de jaoks eesmärgiga vähendada mahajäämust suurematest.

Märksõnad: arenevad tehnoloogiad, audit, innovatsioon, arenevate tehnoloogiate kasutuselevõtu protsess, protsessiinnovatsioon

CERCS: S190, S192

1. Sissejuhatus

Tehnoloogia areng tänapäeva maailmas on väga kiire ning see avaldab mõju paljudele ettevõtetele erinevates valdkondades. Auditeerimise valdkond on viimase poolesaja aasta jooksul läbi elanud tehnoloogia arenguga seoses drastilise muutuse – töötajate arv sektoris on seoses arvuti kasutuselevõtuga ja digitaalse raamatupidamise arenemisega kordades langenud. Viimastel aastatel on seoses koroona pandeemiaga astutud kiireid samme distantsilt auditeerimise arenguteel (Seidenstein jt, 2024). Eesotsas tehisintellektiga (AI[1]) on praegu arengufaasis mitmeid arenevaid tehnoloogiaid (AT), millel on potentsiaali oluliselt mõjutada audiitorite tööd lühemas ja pikemas perspektiivis (IAASB, 2020). Globaalses vaates võib AT-de kasutuselevõtt ohtu seada ka paljud tänapäevased töökohad finantsteenuste valdkonnas (Palekhova & Kramarenko, 2020). Audiitorteenuste puhul nähakse AT-de

[1] Magistritöös kasutatud lühendid on välja kirjutatud esmasel mainimisel, kokkuvõtlikult on lühendid välja toodud lisas LISA E.

kasutuselevõtus võimalust efektiivsuse tõstmiseks (Krieger jt, 2021; Ruhnke, 2023; Zemankova, 2019; Vitali & Giuliani, 2024), kuid selles, kas see mõjutab audiitorettevõtjate (AEV) töötajate arvu või struktuuri, ollakse erinevatel arvamustel (Vitali & Giuliani, 2024).

Eesti audiitorteenuste turul valitseb praegu audiitorite defitsiit, kuna madalate piirmäärade tõttu on audiitorkontrolli kohustusega ettevõtete arv suurenenud (Õun, 2024). Vandeaudiitorite keskmine vanus on olnud tõusutrendis, pidevalt kehtestatakse uusi regulatsioone, millega kaasas käimine on koormav ning soodustab vanemate vandeaudiitorite kutsetegevusest lahkumist (Õun, 2024). Seega võib Eesti kontekstis AT-de kasutuselevõtust saadav efektiivsuse suurenemine võimaldada turuolukorra normaliseerumist, sest käesoleva uurimuse lõpetamise ajaks vastu võetud uued audiitorkontrolli kohustuse piirmäärad, ei pruugi leevendust siiski tuua (Audiitortegevuse seadus, 2025; Õun, 2024). Sama seadusemuudatusega jõustub alates 17.01.2025 ka kestlikkusaruannete audiitorkontrolli kohustus, mis võib järgnevate aastate jooksul osa vandeaudiitoritest finantsauditite valdkonnast kõrvale juhtida (Audiitortegevuse seadus, 2025; Õun, 2024).

Teadmistemahukate teenuste sektoris, kuhu audiitortegevus ka kuulub, on leitud, et kõrgem innovatsioonivõimekus on seotud klientidele pakutava suurema lisandväärtusega, mis omakorda tingib ka finantsiliselt edukamad äritulemused (Santos-Vijande jt, 2013). Samuti on leitud, et sellel sektoril on majanduses oluline roll ka innovatsiooni vahendaja ja edasikandjana – teenuse osutamise protsessi raames võivad uued tehnoloogiad levida ka teenuseosutaja klientidele (Miles, 2005). Üldiselt on poliitikategijatel suur huvi selle vastu, kuidas edeneb AT-de kasutuselevõtt ettevõtetes (OECD/Eurostat, 2018). Varasemates artiklites on soovitatud uurida digitaalse arengu erinevusi suurte ja väikeste AEV-de vahel (Barr-Pulliam jt, 2022); kuidas väiksemad AEV-d võtavad kasutusse väliste osapoolte poolt välja töötatud tööriistu (Barr-Pulliam jt, 2022); tehisintellekti kasutuselevõttu *Big 4* võrgustikku mittekuuluvates AEV-des (Fedyk jt, 2022). Eestis tegutseb 116 AEV-d, kellest 20 suurema käive moodustab 80% kogu audiitorteenuste turu käibest (AJN, 2024), mistõttu on tegemist hea populatsiooniga suurte ja väikeste erinevuste analüüsimiseks. AT-de kasutuselevõttu Eesti AEV-des on uuritud vaid põgusalt – Mikhaylova (2024) on oma magistritöös leidnud, et tehisintellektil põhinevatel tööriistadel on Eesti AEV-de silmis mitmeid potentsiaalseid kasutusvõimalusi.

Arvestades üldist huvi AT-de kasutuselevõtuga seoses, turuolukorda, mis nõuab AEV-delt efektiivsuse suurendamist ning võimalust täiendada olemasolevat teaduskirjandust suurte ja väikeste AEV-de erinevusi käsitleva analüüsiga on seatud järgmine uurimiseesmärk:

Anda hinnang innovatsioonile Eesti audiitorettevõtjates läbi arenevate tehnoloogiate kasutuselevõtu ja anda soovitusi selle protsessi edendamiseks.

Eesmärgi täitmiseks on püstitatud järgmised uurimisülesanded:

1. määratleda uurimisobjekt ning selgitada, millistes valdkondades on AEV-des võimalik AT-sid kasutusele võtta;
2. anda ülevaade teaduskirjandusest, mis käsitleb innovatsiooni ja AT-de kasutelevõtu protsessi ning seda mõjutavaid tegureid AEV-des;
3. kujundada uurimuse metoodika ning koostada küsimustik;
4. viia läbi küsitlus ning tulemuste analüüs;
5. arutleda tulemuste üle ja kujundada soovitused AT-de kasutuselevõtu protsessi edendamiseks Eesti AEV-des.

Uurimus tugineb peamiselt kahele teoreetilisele raamistikule – AEV-de innovaatilisuse põhinäitajate analüüs tugineb Oslo käsiraamatus (OECD/Eurostat, 2018) välja toodud teoreetilistele kontseptsioonidele. AT-de kasutuselevõttu käsitlev osa uurimusest tugineb Krieger jt (2021) poolt välja töötatud protsessiteooriale täiustatud andmeanalüüsi tehnoloogiate (ADA) kasutuselevõtu kohta AEV-des. See on spetsiifiliselt audiitortegevuse valdkonnas konkreetse tehnoloogia kasutuselevõttu kirjeldav edasiarendus E. M. Rogersi (2003) poolt välja töötatud innovatsiooni levimise teooriast ning kirjeldab ADA kasutuselevõtu etappe ning seda protsessi mõjutavaid AEV siseseid ning väliseid tegureid (Krieger jt, 2021). Kriegeri jt teooriaga on mitmeid sarnaseid jooni ka Goto (2023) poolt välja töötatud protsessiteoorias, mis keskendub samuti auditi valdkonnale.

Käesoleva uurimuse empiiriline osa on läbi viidud Oslo käsiraamatus (OECD/Eurostat, 2018) esitatud soovitustest lähtuvalt – koostatud on valikvastustega küsimustik, mis võimaldab kvantitatiivset andmeanalüüsi ja millega kogutakse andmeid nii AEV-de innovaatilisuse põhinäitajate kui ka AT-de kasutuselevõtu kui spetsiifilise fookusteema kohta. Fookusteema küsimused ning vastusevariandid on välja töötatud varasema kirjanduse ja Rahvusvaheline Auditeerimise ja Kindlustandvate Töövõttude Standardite Komitee (IAASB) tehnoloogia töörühma poolt koostatud materjalide põhjal (täpsemalt peatükis 2). Kvantitatiivne lähenemine võimaldab anda üldistatud hinnangu AT-de kasutuselevõtu ning sellega seotud tegurite kohta Eesti AEV-de seas ning audiitorteenuste turul. Andmed on kogutud perioodil 29.11.2024 kuni 07.12.2024 Eesti AEV-delt. Kvantitatiivne andmeanalüüs on läbi viidud tarkvaras R ja RStudio.

Uuringu tulemustest selgub, et Eesti AEV-d ei erine innovaatiliste ettevõtete osakaalult Eesti ettevõtete innovatsiooniuuringu keskmistest tulemustest, kuid on näha, et

AEV-de hulgas on oluliselt suurem osakaal üksnes protsessiinnovatsiooniga ettevõtetel. Protsessiinnovatsioone on läbi viinud enamik AEV-sid ning peamiselt on need uuendused välja töötanud AEV ise. Olulisemad põhjused, miks ei ole rohkem innovatsioone ellu viidud, on vajaduse ja ressursside puudus. AT-de kasutuselevõtt ei ole AEV-de hulgas veel laialt levinud, kuid suuremate AEV-de kiirema arengu mõjul on vähemalt üks AT rakendatud juba rohkem kui poolel audiitorteenuste turust. AT-de kasutuselevõtu protsessiga seotud küsimustes, sealhulgas nendest saadava kasu ja kasutuselevõtu takistuste osas ilmneb mitmeid erinevusi suurte ja väikeste AEV-de vahel – üldiselt on suuremad AT-de osas optimistlikumad. Ligikaudu pooltel AEV-del ei ole ühtegi töötajat, kellel oleks konkreetne tööülesanne tegeleda innovatsiooniga, mis võib osutuda oluliseks takistuseks ka AT-de kasutuselevõtu protsessis. Tulemustest lähtuvalt on autor andnud mitmeid soovitusi AEV-dele ja selle turuga seotud organisatsioonidele, mis võiksid kiirendada AT-de kasutuselevõtu protsessi AEV-des ning vähendada ka väikeste AEV-de mahajäämust suurematest. Soovituste põhirõhk on AT-de alaste teadmiste tõstmine AEV-des ja koostöö soosimine AEV-de vahel ja teiste asutustega vähendamaks ressursside ja IKT alaste teadmiste puudusest tingitud takistusi.

Uuringu läbiviimist ja uuringu disaini mõjutasid töö autoriga seotud piirangud. Autor on ise praktiseeriv vandeaudiitor ning ühe keskmise suurusega AEV osanik ning juhatuse liige. Ühelt poolt tuleb see taustateadmiste ning valdkonna tundmise osas kasuks, kuid samuti tähendab see, et uurimisobjektid on samas ka autori konkurendid. Selleks, et AEV-d oleksid nõus vastuseid esitama, otsustati, et küsimustik on anonüümne ning ei sisalda infot, mille alusel oleks autoril võimalik tuvastada, kes on vastanud. Seetõttu on küsimustikus ka töötajate arv ning käive küsitud vahemikuna, mitte konkreetse numbrina, mis oleks muidu statistilise analüüsi puhul olnud lihtsamini kasutatav variant. See on üks oluline põhjus, miks on otsustatud kvantitatiivse, valikvastustega uuringu kasuks – kvalitatiivse uuringu puhul, näiteks intervjuude formaadis, oleks arutelu pidanud minema AT-de teemasse nii süvitsi, et vastajad ei oleks pruukinud soovida sellist infot konkurendile anda. Edasised uuringud võiksid neid lünki täita.

Edasine töö on üles ehitatud järgnevalt. Teises peatükis antakse ülevaade teaduskirjandusest, mis käsitleb innovatsiooni uurimist ja mõistet, AT-de kasutuselevõttu audiitorteenuste ja teiste teadmistemahukate äriteenuste valdkonnas ning Kriegeri jt (2021) loodud protsessiteooriat ADA kasutuselevõtu kohta AEV-des. Kolmandas peatükis kirjeldatakse käesoleva uuringu uurimisobjekti ehk Eesti AEV-sid, metoodikat ning andmeid. Neljandas peatükis analüüsitakse uuringu tulemusi. Viiendas peatükis arutletakse uuringu

tulemuste üle, tuuakse välja töö tugevused ja nõrkused ning antakse soovitusi edasiseks uurimiseks. Töö lõppeb kokkuvõttega.

2. Kirjanduse ülevaade

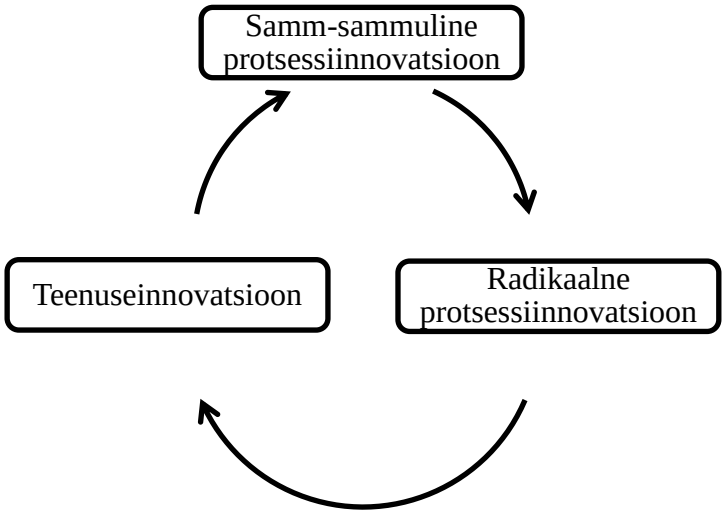
2.1 Innovatsiooni uurimine, innovatsiooni mõiste

Innovatsiooni uurimise alguseks võib pidada Schumpeteri töid 20. sajandi esimesel poolel, milles ta uuris innovatsiooni kui majandusarengu võtit. Läbi aja on mõistet „innovatsioon“ defineeritud väga erinevalt, sealhulgas paljude alaliikide kohta, kuid enamik neist sisaldab keskse põhimõttena uue idee või uuendusliku idee edukat rakendamist (Varadarajan, 2024). Käesolevas uurimistöös kasutatakse OECD (2018) poolt välja töötatud definitsiooni, mis on innovatsiooniuuringutes laialt levinud. Selle mõistmiseks on oluline teada, et mõiste „toode“ hõlmab nii kaupu kui teenuseid:

> „Innovatsioon on uus või täiustatud toode või protsess (või nende kombinatsioon), mis erineb oluliselt üksuse poolt varasemalt pakutud toodetest või üksuse poolt varasemalt kasutatud protsessist, ja mille üksus on kättesaadavaks teinud potentsiaalsetele kasutajatele (toode) või üksuse poolt rakendatud (protsess).“ (OECD/Eurostat, 2018)

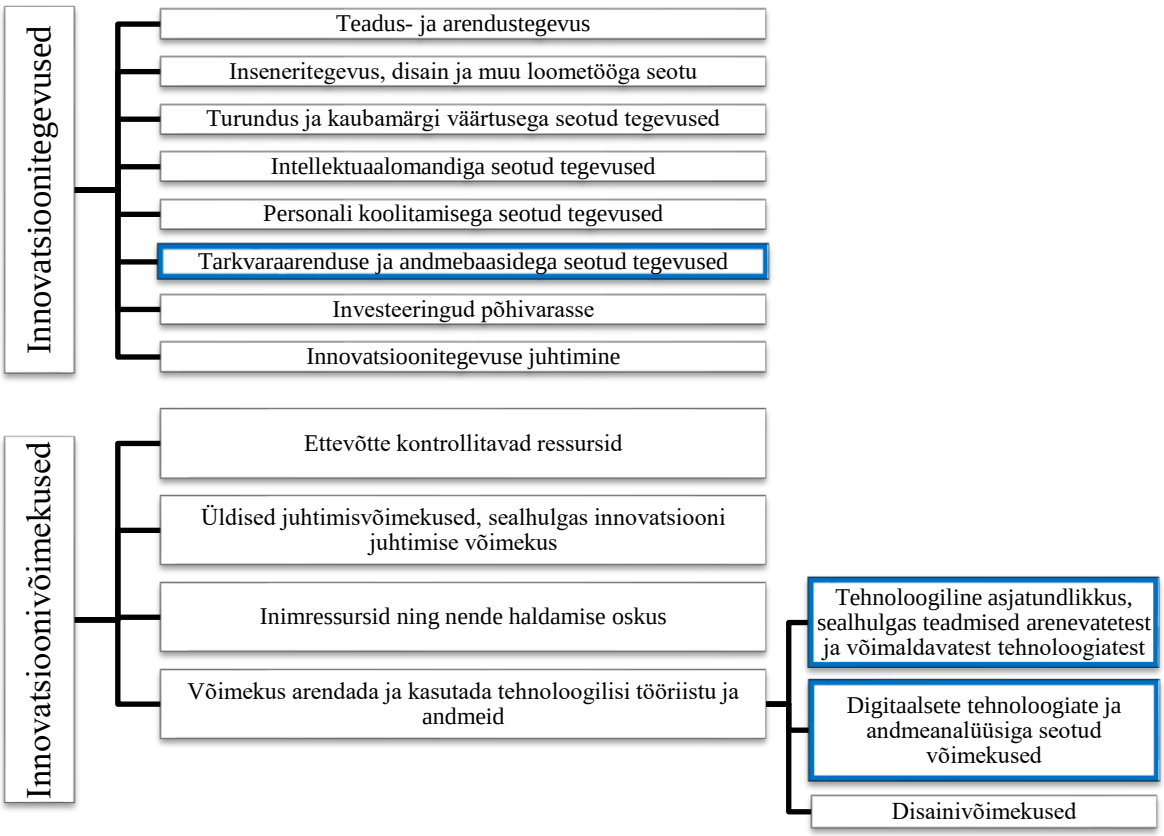
Definitsioonist on loogiliselt tuletatav ka Oslo käsiraamatus (2018) toodud innovatsiooni jaotus toote- ja protsessiinnovatsiooniks olenevalt sellest, kas uudne on toode või protsess. AEV-de kontekstis, kes pakuvad üksnes teenuseid, on käesolevas uurimuses kasutatud enamasti tooteinnovatsiooni asemel mõistet „teenuseinnovatsioon“ vältimaks põhimõistega kaasnevat segadust, mis võib tekkida intuitiivsest ootusest, et „toode“ on midagi füüsilist, kuigi käsiraamatu kohaselt peetakse tooteks ka teenuseid.

Teenuste sektoris on toote- ja protsessiinnovatsioonid omavahel seotud ning seda on käsitlenud R. Barras (1986) oma pööratud toote elutsükli teoorias. Selle teooria kohaselt teenuste innovatsiooni edasi info- ja kommunikatsioonitehnoloogia (IKT) kasutus: esimeses etapis toimub samm-sammuline protsessiinnovatsioon teenuste osutamise protsesside efektiivsuse tõstmiseks; teises etapis toimub radikaalne protsessiinnovatsioon, mille käigus muudetakse põhjalikult teenuse osutamise viise ning pakutakse välja uusi lahendusi; kolmandas etapis kulmineerub tsükkel teenuseinnovatsiooniga – eelneva põhjal ja tehnoloogia abil jõutakse turule uute teenuste pakkumiseni (Barras, 1986) (Joonis 1).



Joonis 1 Innovatsiooni etapid pööratud toote elutsükli teooria kohaselt
Allikas: Autori koostatud Barras (1986) põhjal

Tavakasutuses kasutatakse sõna „innovatsioon“ nii tegevuste kui ka tulemuste kohta. Innovatsiooni mõõtmiseks ettevõtetes on oluline neid selgelt eristada. Lähtuvalt definitsioonist peetakse käesolevas uuringud innovatsiooniks tulemust – toodet või protsessi, mis on reaalselt rakendust leidunud. Innovatsiooni elluviimiseks peetakse innovatsiooni esmakordset rakendamist, sealjuures on oluline esmakordne rakendamine ettevõttes – tegemist ei pea olema maailma, riigi või turu kui terviku jaoks esmakordne kasutuselevõtt. Hiljem võivad ettevõtted teha veel muudatusi ja arendusi, millest võib, aga ei pruugi kujuneda eraldi innovatsioon. Innovatsioonitegevused on ettevõtte poolt läbi viidavad tegevused, mille sihiks on ellu viia uusi innovatsioone ning need jagunevad kaheksasse laia kategooriasse, mis on välja toodud joonisel 2. Käesoleva uurimustöö fookuses on arenevate tehnoloogiate kasutuselevõtt, mis on seotud tarkvaraarenduse ja andmebaasidega seotud tegevustega. Viimaste all on mõeldud nii ettevõtte siseselt arendatud kui väljast ostetud tarkvara, sealhulgas standardseis tarkvarapakette, ettevõtte jaoks sobitatud tarkvara ning seadmetega kaasa tulevat tarkvara; digitaalseid andmeid ja andmebaase. Need tegevused on innovatsioonitegevused siis kui nende eesmärk on uute või oluliselt edasi arendatud toodete või äriprotsesside arendamine. (OECD/Eurostat, 2018)



Joonis 2 Ettevõtete innovatsioonivõimekuste ja innovatsioonitegevuste jaotus ning käesoleva uurimuse fookusteemad
Märkus: Käesoleva uuringu fookusteemaga seotud innovatsioonitegevused ja -võimekused on joonisel tähistatud sinise raamiga.
Allikas: Autori koostatud Oslo käsiraamatu põhjal (OECD/Eurostat, 2018)

Ettevõtete innovatsioonitegevuste tulemuslikkust mõjutavad järgmised võimekused (Joonis 2): ettevõtte kontrollitavad ressursid; üldised juhtimisvõimekused, sealhulgas innovatsiooni juhtimise võimekus; inimkapital ning selle haldamise oskus; ning võimekus arendada ja kasutada tehnoloogilisi tööriistu ja andmeid. Viimati mainitud võimekuse hulka kuuluvad digitaalsete tehnoloogiate, sealhulgas paljude arenevate ja võimaldavate tehnoloogiate, ning andmeanalüüsiga seotud teadmised ja oskused. Toote- ja protsessiinnovatsioonid, mis sisaldavad info- ja kommunikatsioonitehnoloogiaid (IKT) või sõltuvad olulisel määral IKT-st moodustavad olulise osa innovatsioonidest kõikides sektorites. Innovatsioonivõimekuste kohta info kogumine võimaldab analüüsida, milline on innovatsiooni mõju ettevõtte ärilisele edule ja miks mõned ettevõtted on innovaatilised ja mõned mitte. (OECD/Eurostat, 2018)

Teadmistemahukate teenuste sektoris on leitud, et kõrgem innovatsioonivõimekus on seotud klientidele pakutava suurema lisandväärtusega, mis omakorda tingib ka finantsiliselt edukamad äritulemused (Santos-Vijande jt, 2013). Teadmistemahukate äriteenuste osakaal

majanduses on IKT kiire arengu toel oluliselt kasvanud ja on leitud, et sellel valdkonnal on majanduses oluline roll ka innovatsiooni vahendaja ja edasikandjana – teenuse osutamise protsessi raames võivad uued tehnoloogiad levida ka teenuseosutaja klientidele (Miles, 2005). Tüüpilised poliitikameetmed innovatsiooni soodustamiseks, mis peaksid olema sektorineutraalsed, on siiski disainitud pigem tootmisettevõtete profiili silmas pidades, mistõttu ei pruugi need arvesse võtta teadmistemahukate äriteenuste eripärasid, näiteks

- uurimis- ja arendustegevus ei ole innovatsiooni allikana niivõrd olulisel kohal ja
- sageli ei ole selles valdkonnas eraldi osakondi ja juhte, vaid innovatsiooniga tegelevad arendustiimid;
- patente kasutatakse vähem intellektuaalomandi kaitseks ja muu (Miles, 2005).

2.2 Varasemad uuringud audiitorteenuste valdkonnas

Eestis on audiitorettevõtjates toimuva innovatsiooniga seoses varem avaldatud vähe uurimusi. Ka üleriigiline ettevõtete innovatsiooniuuring, mida viib läbi Statistikaamet, ei ole kaasanud audiitorettevõtjaid hõlmavat tegevusala (*Statistikaamet*, 2024a). Kui aastal 2017 oli võimalik tõdeda, et Eesti audiitorteenuste turul valitses hinnasurve, mis tõi kaasa vajaduse suurendada auditite efektiivsust, muuhulgas läbi tehnoloogia kasutuselevõtu (Budarov, 2017), siis oma 2024. aastal kaitstud magistritöös tõdeb H. Õun (2024), et turul valitseb audiitorite defitsiit, mida tõenäoliselt süvendab seadusemuudatus, mis toob kaasa kestlikkusaruannete kohustusliku audiitorkontrolli. Nii hinnasurve kui ka defitsiidi korral on oluline suurendada AEV-de efektiivsust, mida varasema kirjanduse põhjal arenevate tehnoloogiate kasutuselevõtt võimaldab (Ruhnke, 2023; Zemankova, 2019; Vitali & Giuliani, 2024).

Tehisintellekti kasutamise väljavaateid Eesti AEV-des on käsitlenud oma magistritöös V. Mikhaylova (2024), kelle tööst selgub, et

- suhtumine tehisintellektil põhinevate tööriistade rakendamisse on avatud ja positiivne;
- enamik uuringu käigus küsimustikule vastanud vandeaudiitoreid peab tõenäoliseks, et nende AEV-s võetakse sellised tööriistad kasutusele;
- usutakse, et sellised tööriistad võimaldavad efektiivsemat ajakasutamist ning tõsta teenuste kvaliteeti;
- rohkem kasutusvõimalusi nähakse nii töövõtu aktsepteerimise, planeerimise ja läbiviimise faasides näiteks esmase riskihinnangu läbiviimisel, kliendist arusaamise omandamisel, finantsandmete analüüsimisel ja ebatavaliste tehingute tuvastamisel, kuid

- on protseduure ja tegevusi, milles on audiitoril vaja teha kutsealaseid otsustusi ja kasutada kutsealast skeptitsismi, milles tehisintellekt inimest asendada ei saa;
- tehisintellekti kasutuselevõtt on väiksematele AEV-dele keeruline, kuna nõuavad suuri investeeringuid.

Uuring põhines vandeaudiitorite seas läbi viidud küsitlusel ning intervjuudel ning ei süvenenud muude AT-de kui tehisintellekti kasutamisele. (Mikhaylova, 2024)

Peamiselt Saksamaa vastajatel põhinevas uuringus leiti, et AEV siseses juhtimises on digitaalsete tehnoloogiate kasutuselevõtt juba kõrgelt arenenud, kuid audiitorteenuse osutamise raames protseduuride läbiviimisel on digitaalsete tööriistade kasutuselevõtt veel küllaltki piiratud (Krieger jt, 2021). Norras läbi viidud uuringus on leitud, et väikeste AEV-de seas on protseduuride automatiseerimisega tegelenud vaid üksikud ja kasutuselevõtt on pigem olematu (Wiklund & Fallan, 2024). Saksamaa audiitorid ei näinud 2019. aastal läbi viidud uuringus, et arenevad tehnoloogiad tooksid kaasa olulisi muutusi auditi valdkonnas (Tiberius & Hirth, 2019). Vastupidiselt on ka leitud, et mõju valdkonnale on juba olemas – tehisintellekti kasutuselevõtu järgselt on täheldatud madalama taseme töötajate arvu langust, samas kui kesk- ja kõrgema taseme töötajate arv ei ole muutunud (Fedyk jt, 2022), prognoositud on kogu finantsteenuste sektoris töökohtade arvu vähenemist (Palekhova & Kramarenko, 2020). Itaalia audiitorite põhjal läbi viidud uuringus leiti, et osad audiitorid arvavad, et AT-de kasutuselevõtt võib vähendada väiksemate oskustega töötajate osakaalu töötajate struktuuris, kuna nende tööd on võimalik automatiseerida, samas kui on ka neid, kes leiavad, et muutuvad üksnes protseduurid ja meetodid, aga töötajate struktuur püsib sarnasena olemasolevale (Vitali & Giuliani, 2024). Samas uurimuses leitakse, et osad audiitorid arvavad, et AT-de kasutuselevõtuga kaasnev suur investeerimisvajadus võib Big-4 ettevõtete konkurentsieelist suurendada, aga on ka neid, kes leiavad, et keskmises perspektiivis muutuvad AT-d odavamaks ning see võimaldab väiksematel AEV-del hoopis vähendada vahet *Big 4* ettevõtetega (Vitali & Giuliani, 2024). Anica-Popa jt (2024) on leidnud, et seoses tehisintellekti kasutuselevõtuga muutuvad tulevikus teadmised ja oskused, mida audiitoritel on vaja, mis muudab ka AEV-de värbamispraktikat. Tabelis 1 on välja toodud IAASB tehnoloogia töörüma poolt välja töötatud materjalide ning autori poolt läbi töötatud kirjanduse põhjal tuvastatud AT-d, millel on (potentsiaalne) mõju audiitorteenuste osutamisele, nende kirjeldused ja näited kasutusvõimalustest auditi kontekstis.

Tabel 1

Arenevad tehnoloogia, nende kirjeldused ja kasutusvõimalused audiitortegevuse valdkonnas

Arenev tehnoloogia	**Tehnoloogia kirjeldus**	**Näiteid kasutusvõimalustest audiitortegevuse valdkonnas**
1. Tehisintellekt *(Artificial Intelligence – AI)*	Arvutiteaduste haru, milles arendatakse arvutisüsteeme, mis suudavad lahendada ülesandeid, mis tavapäraselt nõuavad inimese intelligentsust, nagu nägemine, kõnetuvastus, otsuste tegemine ja tõlkimine. Hõlmab endas mitmeid tehnoloogiaid, millest osa on edaspidi ka eraldi välja toodud: masinõpe, planeerimine ja optimeerimine, robootika, ekspertsüsteemid, loomuliku keele töötlemine, arvutinägemine, kõnest arusaamine. (IAASB, 2022b)	Kasutusvõimalusi auditis on palju, näiteks: * planeerimisel – automatiseeritud töörühma loomine, aktsepteerimise protseduuridel andmete analüüsimine, sh mittetraditsioonilistest allikatest; * kliendist arusaamise omandamisel – erinevatest allikatest pärit info analüüs, sisekontrollidest arusaamine vastava dokumentatsiooni põhjal, ebaharilike tegevuste tuvastamine, tüüpriskide tuvastamine; * substantiivsete protseduuride juures – erisuste tuvastamine andmetes, dokumentide töötlemine ja kokkuvõtete tegemine OCR kaudu, inventuuri läbiviimine droonide ja arvutinägemise abil; * kokkuvõtete tegemisel – aruannetes avalikustatava info nõuetele vastavuse kontroll, ristviidetega auditidokumentatsiooni koostamine ja kontrollküsimustike täitmine (IAASB, 2022b; Seidenstein jt, 2024). * Võimaldab audiitoritel omandada teadmisi, õppida ja teha informeeritumaid otsustusi auditite käigus (Anica-Popa jt, 2024)
2. Loodusliku keele töötlemine *(Natural Language Processing - NLP)*	AI haru, millega antakse arvutitele võimekus aru saada ja kasutada inimkeeli nii kõnes kui kirjalikult. Selle saavutamiseks rakendatakse arvutuslingvistika, statistilise modelleerimise ja masinõppe tehnoloogiaid. (IAASB, 2022d)	* Audiitorile nõu andmine auditi protsessi jooksul häälnõustamise ja abistava info tekstirobotite kasutamisega; * riskihindamise protseduuride ajal audiitorile info andmine kliendi kohta, mittestruktureeritud andmete analüüsimine erinevatest allikatest; * audiitori toetamine kliendi sisekontrollidest arusaamisel; * dokumentatsiooni koostamine koos audiitoriga, näiteks analüüsi kommentaaride sõnastamisel. (IAASB, 2022d; Seidenstein jt, 2024) * Dokumentide analüüsimine ja nendest kokkuvõtete tegemine
3. Robootiline protsesside automatiseerimine *(Robotic Process Automation - RPA)*	Tarkvara robotite tehnoloogia, mille käigus automatiseeritakse rutiinsed, korduvad muidu manuaalselt läbi viidavad ülesanded (IAASB, 2023b)	Kasutusvõimalusi on nii audiitorteenuste osutamisel kui ka AEV sisemistes protsessides: * andmete väljavõtmise, transformeerimise ja väljastamise protsessi automatiseerimine andmete töötlemisel, mis võib olla sisendiks ka edasiste AT-de rakendamisel; * andmetabelite ühendamine; * automatiseeritud sisemiste ja väliste andmete kokku viimine, sh raamatupidamise andmete kokku viimine dokumentidega; * kontrollküsimustike, mallide jt dokumendipõhjade ettevalmistamine. (IAASB, 2023b; Seidenstein jt, 2024)
4. Asjade internet *(Internet of Things – IoT)*	Omavahel digitaalses ühenduses olevate seadmete süsteem, mille igal osalisel on unikaalne tunnus ja võime edastada andmeid ilma inimese sekkumiseta. (IAASB, 2023a)	* IoT tehnoloogiad võivad olla osa kliendi sisekontrollisüsteemist, mistõttu tuleb neid arvesse võtta riskihindamisel ja arusaamise omandamisel; * IoT süsteemide poolt genereeritud andmeid võib olla võimalik kasutada tõendusmaterjalina auditi käigus (IAASB, 2023a).
5. Homomorfne krüpteerimine *(HKrüpt)*	Tehnoloogia, mis võimaldab arvutuste tegemist krüpteeritud andmetega ilma nende lahti krüpteerimiseta – seega saab andmeid analüüsida tagades konfidentsiaalsuse (IAASB, 2022c).	* Võimaldab erinevate osapoolte andmeid kombineerida masinõppeks ilma seejuures andmete konfidentsiaalsust rikkumata; * kolmandate osapoolte andmete kasutamine andmekaitse reegleid rikkumata;

		* riikidevaheliste auditite puhul andmete ühine analüüsimine rikkumata eri jurisdiktsioonide andmekaitsereegleid; * võrdlusuuringute tegemise lihtsustamine (IAASB, 2022c; Seidenstein jt, 2024).
6. Rakendusliidesed *(Application Programming Interfaces – API)*	Kindlaksmääratud reeglite kogum, mis määrab, kuidas rakendused või arvutid üksteisega suhtlevad. Võimaldab ettevõtetel jagada andmeid väliste osapooltega – standardiseeritud päringutele antakse standardseeritud väljundid. (IAASB, 2022a)	Ligipääs kliendi andmetele, päringud andmetest; ligipääs kliendispetsiifilistele kolmandate osapoolte andmetele, näiteks panga väljavõtetele; ligipääs auditi jaoks asjakohasele välisele infole. See võimaldab andmete kasutamist muuta efektiivsemaks ja kiiremaks, muuhulgas ka standardiseeritud andmete kaudu. Saab kombineerida teiste AT-dega, näiteks RPA-ga (IAASB, 2022a; Seidenstein jt, 2024)
7. Andmete standardiseerimine *(Standard.)*	Protsess, mille käigus muundatakse andmed ühtsesse formaati, mis võimaldab hiljem paremat andmeanalüüsi läbi viia ja andmeid efektiivsemalt kasutada. Samuti avaneb võimalus kasutada laiaulatusille analüütilisi meetodeid. (IAASB, 2021)	Standardiseeritud andmete analüüs on kiirem, kuna on võimalik kasutada ühte keskset andmeanalüüsi mudelit erinevatest andmebaasidest pärit andmetele, mistõttu vabaneb aega olulisemate protseduuride jaoks. Võimaldab ADA rakendamist. Üldiselt analüütiliste protseduuride läbiviimine erinevates etappides, sh kogu populatsiooni analüüs. (IAASB, 2021; Seidenstein jt, 2024)
8. Täiustatud andmeanalüütika *(Advanced Data Anlytics – ADA)*	Algoritmid ja statistika, mis võimaldab tuvastada erandeid ja teha ennustusi; statistiline modelleerimine; andmete visualiseerimine; andmete hoidmine „pilves“; sisaldab protsessikaevandamise, tekstikaevandamise tehnoloogiaid, tõenäosusmudeleid, uskumuste võrgustikke, regressiooni, kirjeldavat statistikat, hüpoteeside testimist jm (Appelbaum jt, 2018; Ruhnke, 2023; Seidenstein jt, 2024).	* Andmete kvaliteedi hindamine, mustrite tuvastamine, vigade ja pettuse tuvastamine (Seidenstein jt, 2024); * suurandmete analüüs, sh erinevatest allikatest pärit andmete kombineerimine, potentsiaalselt substantiivsete protseduuride hulga vähendamine auditi protsessis (Ruhnke, 2023). * ADA kasutusvõimalusi on paljudes auditi etappides, sh aktsepteerimisel, planeerimisel, testimisel, ülevaatamisel, arvamuse kujundamisel, pideval auditeerimisel.
9. Plokiahela tehnoloogia *(Blockchain)*	Krüptograafiat rakendav hajutatud pearaamat, mis moodustab digitaalse, andmebaasi ahelas toimunud tehingute kohta, mida on võimalik iga inimesel ise üle kinnitada. (IAASB, 2023d; Seidenstein jt, 2024)	Hajus pearaamat digitaalsete tehingute kohta, mille „sisse kandmist“ peavad süsteemi osalised kinnitama. Iga tehing saab unikaalse koodi ja on tagantjärele kontrollitav. Potentsiaalselt vähendab pettust ning võimaldab pidevat auditeerimist. Võimaldab luua tarkasid lepinguid. Leitakse, et see tehnoloogia võib auditi protsessi tulevikus muuta fundamentaalselt. (Seidenstein jt, 2024)
10. Protsessikaevandamine *(Process Mining - PMining)*	Protsessi käigus salvestatud digitaalsete andmete, näiteks logide põhjal protsessi kirjeldamine (Seidenstein jt, 2024).	* Võimaldab tuvastada tegelikult toimunud sündmuste põhjal konstrueerida, milline on protsess. Seda analüüsi on võimalik võrrelda kliendi sise-eeskirjadega, mis võimaldab tuvastada sisekontrollide puudusi. (Seidenstein jt, 2024)
11. Arvutinägemine *(Computer Vision - CVision)*, sh optiline märgituvastus *(Optical Character Recognition - OCR)*	Tehnoloogia, mis võimaldab arvutil aru saada märkidest, näiteks kirjakeelest, käekirjast, numbritest; piltidel olevast infost; näotuvastus (IAASB, 2022b).	* Dokumentidest auditi jaoks vajaliku teksti väljavõtmine; * võimaldab kasutada NLP-d; * dokumentide AI abil analüüsimise võimaldamine (Seidenstein jt, 2024).

Märkus: Arenevate tehnoloogiate nimede järele on lisatud inglisekeelsed nimetused ja lühendid

Allikas: Autori koostatud tabelis viidatud allikate põhjal

Varasemates artiklites on soovitatud uurida digitaalse arengu erinevusi suurte ja väikeste AEV-de vahel (Barr-Pulliam jt, 2022); kuidas väiksemad AEV-d võtavad kasutusse

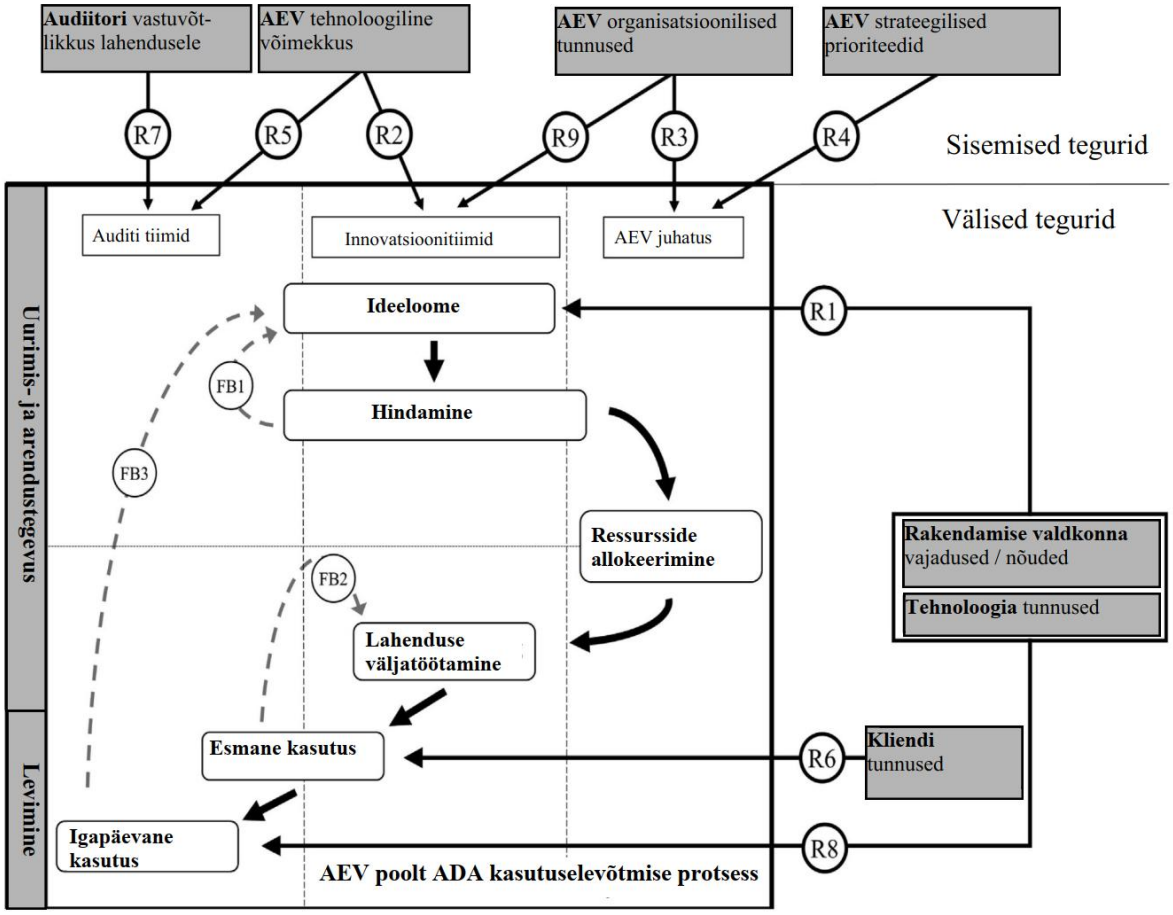
väliste osapoolte poolt välja töötatud tööriistu (Barr-Pulliam jt, 2022); tehisintellekti kasutuselevõttu *Big 4* võrgustikku mittekuuluvates AEV-des (Fedyk jt, 2022); kas AT-de kasutuselevõtt suurendab audiitorteenuste turu kontsentreerumist suurte AEV-de kätte või võimaldab väiksematel vahet vähendada (Ruhnke, 2023; Vitali & Giuliani, 2024). Käesolev uurimus üritab täita rahvusvahelises teaduskirjanduses tuvastatud uurimislünkasid ning täiendab Mikhaylova (2024) poolt tehtut laiendades uuritavate AT-de ringi, rakendades küsitluse puhul põhimõtet, et ühe AEV kohta antakse üks vastus, mis võimaldab tulemusi laiendada AEV-de kogumile, ning kaardistab lisaks tulevikuväljavaadetele ka hetkeolukorda AT-de kasutuselevõtmisel. Eesti on Euroopa Liidu nende riikide hulgas, kus digitaalsete tehnoloogiate kasutuselevõtt on laiemalt levinud, mis viitab suuremale innovatsioonivõimekusele ja potentsiaalselt kõrgemale konkurentsivõimele (Vărzaru & Bocean, 2024). Seega on edasistes uuringutes ehk võimalik läbi viia ka riikidevahelisi võrdlusi.

2.3 Protsessiteooria täiustatud andmeanalüüsi tehnoloogiate kasutuselevõtu kohta audiitorettevõtjates

Krieger jt, (2021) on loonud teoreetilise raamistiku täiustatud andmeanalüüsi (ADA) tehnoloogiate kasutuselevõtu kohta audiitorettevõtetes, mis selgitab

1. kuidas AEV-d ADA tehnoloogiaid kasutusele võtavad;
2. kuidas mõjutavad kasutuselevõtu protsessi nii AEV endaga, selle keskkonnaga ja kasutusse võetava tehnoloogiaga seotud tegurid;
3. kuidas eelnev mõjutab kasutuselevõtu protsessi tulemust (Joonis 3).

See mudel on audiitortegevuse valdkonnas ühe konkreetse AT kasutuselevõtu spetsiifikast lähtuv edasiarendus E. M. Rogersi innovatsiooni difusiooni teooriast (Krieger jt, 2021). Rogersi teooria selgitab, kuidas uuendused levivad ühiskonnas viiesammulise protsessi kaudu, mis hõlmab teadlikkuse teket, huvi ülesnäitamist, hindamist, katsetamist ja omaksvõttu, rõhutades, et leviku tempo ja ulatus sõltuvad innovatsiooni tajutud kasulikkusest, sotsiaalsetest mõjudest ning varajaste kasutajate rollist üldise aktsepteerimise protsessis (Rogers, 2003). Kriegeri jt (2021) teooriaga sarnase raamistiku teadus- ja arendustegevuse kohta AT-de kasutuselevõtul on välja töötanud ka Goto (2023), kes toob välja, et uurimis- ja arendustegevus on teenuste sektoris tihti tagaplaanile jäänud, kuigi sellel on oluline roll, eriti kui on vaja kasutusele võtta uusi tehnoloogiaid, mille kohta ei ole praktiseerijatel endil põhjalikke teadmisi (Goto, 2023). Käesoleva uurimuse peamise teoreetilise raamistikuna käsitletakse Kriegeri jt (2021) teooriat.



Joonis 3 Protsessiteooria täiustatud andmeanalüüsi tehnoloogiate kasutuselevõtu kohta audiitorettevõtjates
Märkus: R1-R9 – tehnoloogiate kasutuselevõttu mõjutavad tegurid, FB1-FB3 – tagasiside-ahelad, AEV – audiitorettevõtja, ADA – täiustatud andmeanalüüsi tehnoloogiad
Allikas: Krieger jt, (2021), autori tõlgitud

Joonisel 3 kujutatud Kriegeri jt teooria kohaselt osalevad AEV siseselt protsessis innovatsioonitiimid, juhatus, ja audititiimid. Protsess koosneb kuuest järjestikusest sammust, milles võib esineda ka tagasisideahelaid: ideeloome, hindamine, ressursside allokeerimine, lahenduste väljatöötamine, esmane kasutus ja igapäevane kasutus. Esimesed neli sammu moodustavad uurimis- ja arendustegevuse etapi ja esmane ning igapäevane kasutus moodustavad levimise etapi. Kahe etapi vahel peab toimuma otsus uus lahendus kasutusele võtta ning selle levikuks igapäevasesse kasutusse on vaja, et auditiimid selle aktsepteeriksid. Kõiki samme mõjutavad nii AEV sisesed kui väliskeskkonnast tulenevad tegurid, mis võivad saada määravaks protsessi lõpptulemusele, ehk sellele, kas AT võetakse kasutusse või mitte. (Krieger jt, 2021)

Ideeloome. Protsessi esimene samm on ideeloome, mille käigus tuvastatakse, kuidas ja milleks AT-d potentsiaalselt kasutada saaks. Seda viivad läbi peamiselt innovatsioonitiimid, kellel on AT-st kõige paremad teadmised, kuid aidata saavad ka

auditititiimid, kellel on paremad teadmised audiitorteenustest. Kasutusviiside tuvastamiseks võib olla vaja erinevate teadmistega töötajate koostööd ning protsessi hõlbustab see kui innovatsioonitiimides on audiitortegevuse alaseid teadmisi ning auditi tiimidel on üldine arusaamine AT-dest. (Krieger jt, 2021) Innovatsioonitöötajate ja auditítöötajate vahelise tasakaalu leidmist ideeloome faasis on välja toonud ka Goto (2023). Valdkonnaüleste teadmistega töötajad on väärtuslikud nii ideeloome kui igapäevasesse kasutamisse juurutamisel ja AT aktsepteerimisel audiitorite poolt (Goto, 2023; Krieger jt, 2021). Väliste teguritena mõjutavad ideeloome faasi valdkonna nõuded, sealjuures AEV-de puhul kehtivad regulatsioonid, ja AT tunnused, eriti selle keerukus ning tingimused, mis on vajalikud kasutamiseks (Krieger jt, 2021).

Hindamine. Kui on välja segitatud potentsiaalne kasutusviis, hinnatakse idee vastavust AEV eesmärkidega – kulutõhusus, efektiivsus ja asjakohasus. Sammu viivad läbi enamasti innovatsioonitiimid, aga oma panuse võivad anda ka auditititiimid ja juhatus. Mida paremini on AT kasutusviis kooskõlas eesmärkidega, seda tõenäolisem on kasutuselevõtu protsessi jätkumine. Kui kasutusviis ei vasta ootustele, võib protsess pöörduda uuesti ideeloome faasi. (Krieger jt, 2021)

Audiitortegevuse valdkonnale keskendunud uuringute põhjal avaldub AT-de kasutuselevõtmise kasu, mida võib pidada ka AEV eesmärkideks, mitmel moel:

- tõuseb audiitorite töö kvaliteet (Ruhnke, 2023);
- tõuseb töö efektiivsus (Krieger jt, 2021; Ruhnke, 2023; Zemankova, 2019; Vitali & Giuliani, 2024);
- väheneb vigade hulk töövõttudes (Zemankova, 2019);
- töötajatel avaneb võimalus panustada rohkem aega keerulisemate ja kõrgema lisandväärtusega tegevustele (Zemankova, 2019);
- tõuseb ettevõtte konkurentsivõime (Ruhnke, 2023).

Ressursside allokeerimine. AT kasutuselevõtt eelmiste sammudega tuvastatud viisil nõuab nii finantsilisi kui inimressursse ning kui neid ei võimaldata, jääb arendus pooleli või minimaalsele tasemele. Tuleb ka otsustada, kas arendus viiakse läbi ise, ostetakse sisse või viiakse läbi koostöös kellegi teisega. Otsused sõltuvad muu hulgas ka ettevõtte suurusest – suurematel ettevõtetel on rohkem finantsvahendeid ning tihtipeale on neil ka lihtsam talente tööle meelitada. Investeerimisotsust mõjutavad ka nii ettevõtte tulumudel, omanike struktuur, AT-de kasutuselevõtuga seotud strateegilised otsused kui ka klientide ootused ja vajadused. (Krieger jt, 2021)

AEV-de innovatsioonitegevusi, sealhulgas AT-de kasutuselevõttu, mõjutab nii AEV juhtkonna (Zemankova, 2019) kui ka kliendi juhtkonna suhtumine AT-desse ja innovatsiooni (Barr-Pulliam jt, 2022). Mida optimistlikum on juhtkond arenevatest tehnoloogiatest saadava kasu osas, seda suurem on tõenäosus, et arenevaid tehnoloogiaid kasutusele võetakse (Zemankova, 2019). AT-de kasutuselevõttu takistavate teguritena nähakse

- regulatiivseid takistusi, sealhulgas spetsiifiliste kutsetegevuse standardite puudumist (Krieger jt, 2021);
- töötajate ajalise ressursi puudust (Krieger jt, 2021; Seidenstein jt, 2024);
- finantsressursside puudust, sh nii ise arendamiseks kui teistelt ostmiseks (Krieger jt, 2021; Seidenstein jt, 2024);
- võimetust värvata piisavate oskuste ja teadmistega tööjõudu (Krieger jt, 2021);
- väiksemate AEV-de puhul ka kliendipoolse nõudluse ja vajaduse puudumist (Wiklund & Fallan, 2024);
- audiitorettevõtte juhtkonna hinnangut, et investeeringud tehnoloogiasse ei ole tasuvad (Wiklund & Fallan, 2024);
- ning väiksematest ettevõtetest koosneva kliendiportfelli puhul ka võimetust saada kliendilt algandmeid kujul, mis on sobilik innovaatiliste tööriistade kasutamiseks (Vitali & Giuliani, 2024).

Kuigi rahvusvahelisel tasandil on mitmeid suuri AEV-de võrgustikke, mis on juba aastaid AT-del põhinevaid lahendusi kasutanud, on paljudel väiksematel võimatu ise arendada, kuna selleks ei ole piisavalt finants- ega inimressursse (Seidenstein jt, 2024). Kõik audiitorid ei ole tehnoloogia vallas pädevad, mis takistab AT-de kasutuselevõttu ning tekitab AEV-dele täiendava vajaduse investeerida ka töötajate koolitamisse (Seidenstein jt, 2024).

Lahenduste väljatöötamise sammu käigus vormitakse esialgsest ideest tegelikult toimiv lahendus. See protsess võib nõuda ideeloomega võrreldes teistsuguseid teadmisi ja oskusi, sealjuures tarkvaraarendusest, infosüsteemidest ja äriprotsessidest. Mõju avaldab ka AEV organisatsioonistruktuur – arendus võib toimuda tsentraliseeritult või detsentraliseeritult. **Esmase kasutuselevõtu** sammu käigus testitakse lahendust, liidetakse see audititiimide tööprotsessi ja suunatakse tegelikku kasutusse. Võib selguda, et lahendus on veel poolik – siis liigutakse tagasi väljatöötamise etappi. **Igapäevane kasutus** on AT kasutuselevõtu protsessi viimane samm. AEV-d ei pruugi muuta lahenduse kasutamist kohustuslikuks, mistõttu peab iga audiitor ise lahenduse aktsepteerima. See sõltub tehnoloogia keerukusest, lahenduse poolt pakutavast kasutegurist kui ka välistest teguritest,

näiteks regulatiivsest keskkonnast. Sealjuures võivad regulatsioonid olla nii kasutamist soodustava (näiteks on auditi käigus kohustuslik testida päevaraamatu kandeid) kui ka takistava (näiteks täiendava dokumentatsiooni koostamise kohustus) mõjuga. (Krieger jt, 2021a)

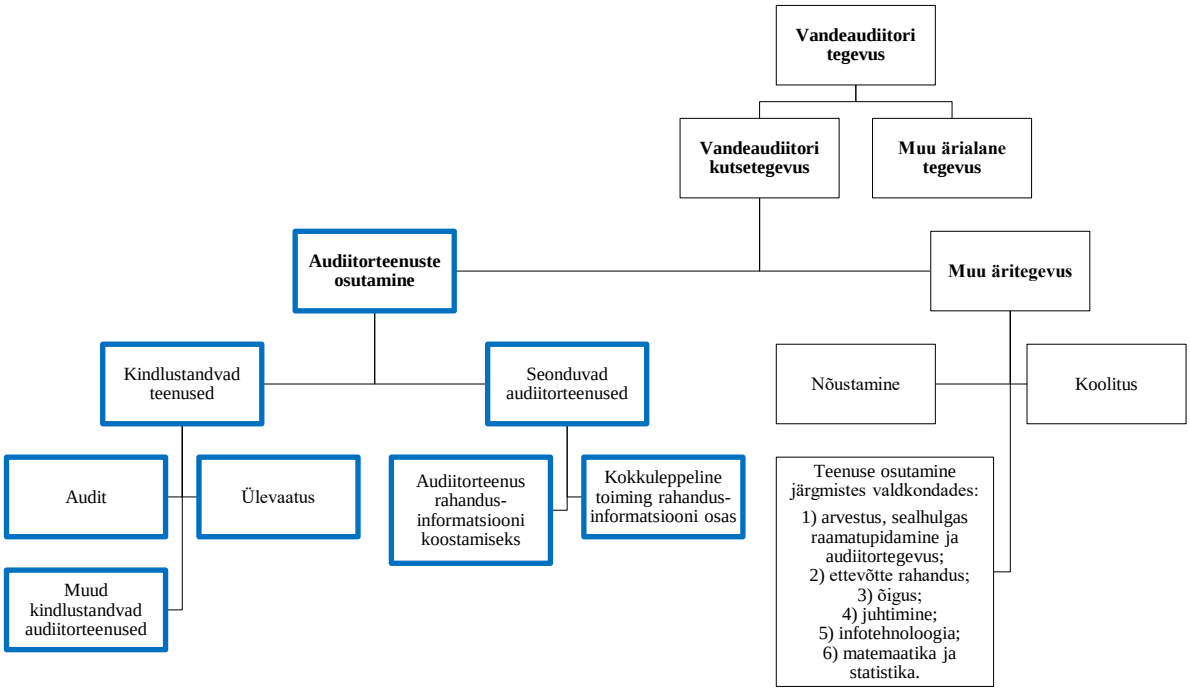
Käesoleva uurimuse empiirilises osas viiakse läbi Oslo käsiraamatus kirjeldatud põhimõtete kohaselt uuring innovatsioonide kohta Eesti AEV-des, keskendudes AT-de kasutuselevõtule (OECD/Eurostat, 2018). Küsimustikuga tuvastatakse ettevõtete innovatsiooniuuringuga (Statistikaamet, 2023) sarnasel viisil AEV-de innovaatilisuse põhinäitajad, mis annavad üldise konteksti ning kaudse võrdlusbaasi teiste valdkondadega Eestis. AT-de osas uuritakse nende kasutuselevõttu, kasutuselevõtu tõenäosust nelja aasta jooksul, AT-dest saadavat kasu nii AT-de kui kasutegurite lõikes, kasutuselevõttu takistavaid tegureid, innovatsioonitöötajate olemasolu AEV-de struktuuris ja seda, kes on AT-de väljatöötaja. Kõik uuritavad küsimused AT-de kohta mõjutavad Krieger jt (2021) poolt välja töötatud AT-de kasutuselevõtmise protsessi mudeli ühte või mitut osa, mis annab võimaluse arutluse osas siduda uuringu tulemused teoreetilise raamistikuga ning anda soovitusi AT-de kasutuselevõtu soodustamiseks.

3. Uurimisobjekt, metoodika ja andmed

Uurimisobjekt. Käesoleva uurimuse uurimisobjektiks on Eesti AEV-d kui juriidilised isikud, mitte vandeaudiitorid indiviidina. Oslo käsiraamatu (OECD/Eurostat, 2018) kohaselt on innovatsiooniuuringute puhul ettevõtte tasandi info kogumine tavapärane. Eestis oli käesoleva uuringu andmete kogumise alguse seisuga 116 AEV-d ning viimase viie aasta jooksul on see arv langenud ligikaudu 12 % võrra. 30.06.2024 lõppenud aastal töötas AEV-des audiitorteenuste valdkonnas kokku täistööajale taandatuna 593 töötajat ning kehtiv kutsetunnistus oli sel kuupäeval 337 vandeaudiitoril. Audiitorteenuste käive oli 53,5 mln eurot, mis moodustas 67% AEV-de kogukäibest ning oli jooksevhindades 24% rohkem kui sellele eelnenud aastal ning 98% rohkem kui viis aastat tagasi perioodil 01.07.2019-01.07.2020. Sama viia aasta jooksul on tarbijahinna indeks tõusnud 42%. Müügitulu ühe töötaja kohta on selle aja jooksul suurenenud 67%. (AJN, 2024; Audiitorkogu, 2024a; Statistikaamet, 2024b)

Perioodil 01.07.2023-30.06.2024 moodustas suurima 20 AEV käive kogu audiitorteenuste käibest 80%, sealjuures *Big 4* võrgustikku kuuluvate ettevõtete käive moodustas kogukäibest 50% (AJN, 2024). Järelikult, umbes poole audiitorteenuste turust katavad iseseisvad ja väiksematesse võrgustikesse kuuluvad AEV-d, mille osas on AT-de

kasutuselevõttu rahvusvaheliselt vähem uuritud. Samuti on võimalus selgitada väiksemate ja suuremate AEV-de erinevusi.



Joonis 4 Vandeaudiitori tegevuste jaotus ja uurimisobjekti piiritlus
Märkus: Sinise raamiga on välja toodud audiitorteenused, mis on käesoleva uurimistöö fookuseks AT-de kasutuselevõtu valdkonnas.
Allikas: Autori koostatud audiitortegevuse seaduse §43-44 alusel (Audiitortegevuse seadus, 2023)

Vandeaudiitorite tegevused jaotuvad kutsetegevuseks (audiitorteenuste osutamine ja muu äritegevus) ning muuks ärialaseks tegevuseks (Joonis 4; Audiitortegevuse seadus, 2023). Audiitorteenuste osutamine on rangelt reguleeritud valdkond – neid teenuseid peavad vandeaudiitorid osutama audiitorettevõtja kaudu, kellel peab olema tegevusluba; teenustele kehtivad rahvusvahelised standardid (rahvusvahelised auditeerimise standardid, rahvusvahelised ülevaatamise teenuse standardid, rahvusvahelised seonduvate teenuste standardid); ja AEV-de üle teostab järelevalvet Audiitortegevuse Järelevalve Nõukogu (AJN).

Kõige tavapärasemad audiitorteenused on raamatupidamise aastaaruannete audit ja ülevaatus, mille läbiviimise kohustus on sätestatud audiitortegevuse seadusega. Kui ettevõttel on sellist teenust vaja, on tal võimalik seda tellida endale sobivalt AEV-lt. Audit on kõrgemat kindlustunnet andev töövõtt, mille eesmärgiks on avaldada arvamust selle kohta, kas finantsaruanded on koostatud kõikides olulistes osades kooskõlas rakendatava

finantsaruandluse raamistikuga (IAASB, 2022e). Auditi käigus tuleb lisaks järelepärimistele, mis üksinda ei ole piisav tõendusmaterjal, viia läbi auditiprotseduure, mis hõlmavad inspekteerimist, vaatlust, kinnitamist, uuesti arvutamist, taasläbiviimist ja analüütilisi protseduure või nende meetodite kombinatsioone (IAASB, 2022f). Ülevaatus on madalamat kindlustunnet andev töövõtt eesmärgiga suurendada finantsaruannete usaldusväärsust, mille käigus viib audiitor auditiga võrreldes läbi vähem ulatuslikke protseduure, millest peamine osa on analüütilised protseduurid ning järelepärimised (IAASB, 2023c).

Käesolev uurimistöö keskendub arenevate tehnoloogiate kasutuselevõtule just audiitorteenuste osutamise kontekstis, sest just see osa audiitorite tegevusest on kõige rohkem reguleeritud ning see on peamine, mis fundamentaalselt eristab AEV-de tegevust enamikust teistest teenuseid pakkuvatest ettevõtetest.

Andmete kogumine ja töötlemine. Uuringu läbiviimisel on aluseks võetud OECD ja Eurostati poolt koostatud Oslo käsiraamat (2018), mis annab suuniseid innovatsiooniga seotud uuringute läbiviimise kohta alustades andmete kogumisest kuni andmete analüüsini. Kuna Statistikaameti poolt läbi viidud Eesti ettevõtete innovatsiooniuuringusse (edaspidi ka CIS) ei ole, siis on küsimustiku kahte osasse lisatud ettevõtete innovaatilisuse põhinäitajaid käsitlevad küsimused, mis võimaldavad hankida andmeid sarnasel kujul nagu neid on hangitud ka Eesti ja Euroopa tasandil teiste tegevusalade kohta (*Statistikaamet*, 2024a). Eraldi alajaotuses on esitatud küsimused AT-de kasutuselevõtu kui uuringu fookusteema kohta. Uurimisstrateegiat saab pidada subjektipõhise (AEV-d) ja objektipõhise (AT-de kasutuselevõtt) lähenemisviisi kombinatsiooniks (OECD/Eurostat, 2018).

Küsimustiku koostamine. CIS-ga võrreldavate andmete kogumiseks innovaatilisuse põhinäitajate kohta, on küsimustikku kaasatatud sellised küsimused innovatsioonide elluviimise kohta, et nende põhjal oleks võimalik tuvastada innovaatiliste AEV-de osakaal ning innovatsioonide jaotus teenuse- ja protsessiinnovatsioonide lõikes. Selles osas on nii küsimuste kui vastusevariantide sõnastus analoogne CIS-i küsimustikuga. Defineeriti ka järgmised mõisted:

> **Innovaatiline audiitorettevõtja** on AEV, kes on perioodil 01.07.2022-30.06.2024 raporteerinud ühe või mitu innovatsiooni, sealjuures hõlmab see nii enda välja töötatud kui ka kellegagi koostöös läbi viidud nii teenuse- kui protsessiinnovatsioone (Statistikaamet, 2023).
>
> **Mitteinnovaatiline audiitorettevõtja** on AEV, kes ei ole perioodil 01.07.2022-30.06.2024 raporteerinud ühtegi innovatsiooni (Statistikaamet, 2023).

Teenuseinnovatsioone käsitlevad küsimused nr 8-13, protsessiinnovatsioone küsimused 14, 15 ja 17 (LISA A). Eesti AEV-d on kohustatud esitama tegevusaruande iga aasta augustis perioodide kohta, mis algavad 1. juulil ning lõpevad 30. juunil. Seetõttu on käesolevas uuringus enamik andmeid esitatud ning kogutud ka sarnaste, AEV-de jaoks harjumuspäraste raporteerimisperioodide kohta. Küsitakse 30.06.2024 seisu kohta ning kaheaastase perioodi 01.07.2022-30.06.2024 kohta. Tulemused ei ole otseselt võrreldavad viimatise CIS-ga, kuna see käsitleb kolmeaastast perioodi 2020-2022 (Statistikaamet, 2023). Edasises analüüsis on konteksti loomiseks siiski kõige laiemal tasemel võrdlus CIS-i tulemustega tehtud selles osas, kui suur osa ettevõtetest on innovaatilised.

Lähtuvalt Krieger jt (2021) poolt koostatud AT-de kasutuselevõtu protsessi teooriast ning eesmärgist anda soovitusi AT-de kasutuselevõtu soodustamiseks, on esitatud küsimusi, millega kogutud andmete põhjal on võimalik hinnata AT-de kasutamist, tulevikus kasutusele võtmise potentsiaali, AT-de kasutuselevõtust saadava kasu suurust ja kasutegureid, takistavaid tegureid, AT-de väljatöötajaid ja innovatsioonitöötajate osa AEV-de struktuuris. Võimaldamaks kvantitatiivset analüüsi ja üldistamist kogumile on kõik küsimused valikvastustega. Küsimuste sõnastamisel on lähtutud Oslo käsiraamatu (OECD/Eurostat, 2018) juhistest, sealjuures on lisatud selgitusi, et küsimused oleksid vastajatele üheselt arusaadavad.

Oslo käsiraamatu (OECD/Eurostat, 2018) kohaselt on AT-de kohta küsimiseks kaks varianti:

1. esitada avatud vastustega küsimused, milles saavad vastajad avaldada, millised AT-d on neil kasutuses;
2. esitada valikvastustega küsimused, milles on antud ammendav loetelu asjakohastest AT-dest.

Käesolevas uurimuses on rakendatud teist varianti. Selle lähenemise keerukuseks on ammendava loetelu koostamine võttes arvesse sektorit, millele on küsimustik suunatud (OECD/Eurostat, 2018). Sarnastel põhimõtetel on koostatud vastusevariandid ka AT-de kasutuselevõtu kasutegurite ja takistavate tegurite kohta.

Kuna käesolev uurimus käsitleb AEV-sid, siis esmaseks allikaks asjakohaste AT-de tuvastamisel oli Rahvusvaheline Auditeerimise ja Kindlustandvate Töövõttude Standardite Komitee (IAASB) tehnoloogia töörühma poolt 2020. aasta novembris koostatud ülevaade arenevatest tehnoloogiatest auditi ja muude kindlustandvate töövõttude valdkonnas ning sealt alates koostatud täpsustavad ülevaated nendest arenevatest tehnoloogiatest (IAASB, 2020, 2021, 2022c, 2022a, 2022d, 2022b, 2023b, 2023a, 2023d). Valdkondliku akadeemilise

kirjanduse läbitöötamisel tuvastati täiendavad AT-d, mida on audiitortegevusega seostatud, kuid mida ei ole IAASB ülevaates käsitletud. Teemakohaseid artikleid otsiti EBSCO Discovery ja Google Scholar andmebaasidest. Esialgu olid fookuses auditi, professionaalsete teenuste ja raamatupidamisteenustega seotud eelretsenseeritud artiklid, mis käsitlesid innovatsiooni, AT-de kasutuselevõttu, tehisintellekti, täiustatud andmeanalüüsi. Eriti asjakohaste artiklite puhul vaadati üle, millised uuemad artiklid on Scopus andmebaasi põhjal loetud artiklile viidanud ning kui tuvastati uuem teemakohane artikkel, töötati läbi ka see. Sellise metoodikaga on tuvastatud asjakohased AT-d, kasutuselevõtu kasutegurid ja takistavad tegurid, mis on välja toodud peatükis 2 „Kirjanduse ülevaade“ ja mille põhjal on küsimustikus esitatud küsimused 1-7 (LISA A).

Krieger jt (2021) teooria kohaselt on AT-de kasutuselevõtul oluline innovatsioonitiimide roll, mistõttu esitati küsimus selle kohta (küsimus nr 16). Arvestades, et Eesti AEV-de hulgas on palju väikseid ja keskmisi, siis lisas autor selle küsimuse puhul vastusevariandid, mis võimaldavad tuvastada, kas ettevõtetes, kus ei ole innovatsioonile spetsialiseerunud tiimi, on siiski üks või mitu töötajat, kes tegelevad innovatsioonidega põhitöö kõrvalt. Kui innovatsioonidega tegeletakse põhitöö kõrvalt andis autor võimaluse eristada olukorda, kus innovatsioonidega tegelemine põhitöö kõrvalt on juhtkonna poolt suunatud või tegutsetakse omal algatusel ilma juhtkonna poolse suunamiseta. Need vastusevariandid on koostatud autori valdkondliku kogemuse põhjal.

Küsimustikus on neli küsimust, mis nõuavad vastamist intervallskaalal antava hinnanguna. Selleks, et vastamine oleks AEV-de jaoks lihtne, on rakendatud kitsaid, nelja ja viie palliseid skaalasid. Hindamaks kui suurt potentsiaalset kasu näeb AEV juhtkond oma ettevõttes AT-de kasutuselevõtul AT-de lõikes (LISA A - küsimus 1) ja kasutegurite lõikes (LISA A - küsimus 2), on rakendatud skaalat „kasu puudub“ (0), „väike kasu“ (1), „mõõdukas kasu“ (2) ja „suur kasu“ (3). Kasutuselevõtmist takistavate tegurite mõju hindamiseks on kasutatud skaalat „ei ole takistuseks“ (0), „väike mõju“ (1), „mõõdukas mõju“ (2), „suur mõju“ (3) (LISA A - küsimus 3). Hindamaks AT-de kasutuselevõtu tõenäosust järgmise nelja aasta jooksul (LISA A - küsimus 6) ja tõenäolist tulevikus kasutusele võetava AT väljatöötajat (LISA A - küsimus 7), on kasutatud skaalat „väga ebatõenäoline“ (1), „ebatõenäoline“ (2), „ei oska öelda“ (3), „tõenäoline“ (4), „väga tõenäoline“ (5). Kõik ülejäänud teemaküsimused on „tõene / väär“, „üks õige variant“ ja „mitu õiget mitmest“ tüüpi küsimused. Ka metainfo küsimused on valikvastustega.

Küsimustiku vormistamine ja vastuste kogumine. Küsimustiku koostamiseks on kasutatud Google Forms ja Pretty Froms Designer tarkvarasid. Küsimustik on anonüümne, st

küsimustikule vastamiseks ei ole vastajal vaja esitada audiitorettevõtet identifitseerivaid tunnuseid ning pole vaja esitada ka kontaktandmeid. Andmete kvaliteedi tagamiseks on kõikidele küsimustele, välja arvatud viimane küsimus, vastamine muudetud kohustuslikuks. Küsimustes, kus on küsitud hinnanguid mitme tehnoloogia või mõjuteguri kohta, on vastamine iga rea kohta kohustuslik. Enne küsimustiku saatmist vastajatele, piloteeriti küsimustik ühe vastaja peal, kellelt saadud tagasiside põhjal muudeti küsimuste järjestust, vormingut ning täiustati küsimustikus esitatud selgitavat materjali.

AEV-de nimekiri koos kontaktandmetega on võetud 28.11.2024 seisuga Audiitorkogu kodulehelt (Audiitorkogu, 2024b) ja kontrollitud üle audiitortegevuse portaalist (Rahandusministeerium, 2024). Küsimustik saadeti kõigile 116-le sel hetkel registris olnud AEV-le e-maili teel 28.11.2024 ning vastuseid koguti kuni 07.12.2024 hommikuni. Selleks, et vastamisaktiivsust tõsta, moodustati kõikidest AEV-dest juhuslik neljakümne ettevõttega valim, kellele helistati 04.12.2024, et tuletada küsimustikku meelde ja paluda vastata. Selleks, et vastamisaktiivsust tõsta ka suuremate AEV-de seas, helistati täiendavalt ka kõikidele Audiitortegevuse Järelevalve Nõukogu aastaraamatus välja toodud suurimale kahekümnele AEV-le, kes ei olnud sattunud eelnevalt mainitud valimisse. Kokku helistati 04.12.2024 viiekümne viiele ettevõttele. 06.12.2024 hommikul saadeti kõikidele AEV-dele viimane meeldetuletuskiri vastamiseks, milles toodi välja kui palju oli selle hetke seisuga vastajaid igast suurusklassist.

Google Forms kaudu kogutud andmete korrastamine viidi läbi tarkvaras Excel. Töötati välja küsimuste ja vastuste koodid ning kodeeriti andmed, määrati igale andmeväljale muutuja tüüp – enamik andmeid viidi binaarsele või kategoorilisele, sealjuures järjestiktunnuse kujule ja numbrilisse formaati. Osade küsimuste puhul viidi mitme valikuga vastused binaarsele kujule ning loodi abimuutujad iga vastusevariandi jaoks. Vastusevariandid ning koodid on välja toodud lisas B. Küsimustiku reliaabluse hindamiseks viiakse tarkvaras R läbi Cronbachi alfa test. Juhul kui testi tulemus viitab ebapiisavale reliaablusele (<0,8), analüüsitakse, millised küsimused mõjutavad reliaablust negatiivselt ning kaalutakse nende vastuste eemaldamist analüüsist.

Kaalumine tulemuste üldistamiseks. Tabelis 2 on välja toodud ülevaatlikud näitajad Eesti audiitorteenuste turuosaliste kohta perioodi 01.07.2023-30.06.2024 andmete põhjal ning nende jaotus suurusklasside kaupa. Suurusklasside moodustamisel on lähtutud Audiitortegevuse Järelevalve Nõukogu (AJN) aastaraamatus kasutatud jaotusest kui turuosalistele tuttavast jaotusest (AJN, 2024). Selline jaotus ei ole küll rahvusvaheliselt võrreldav, kuid arvestades, et Eesti AEV-d on rahvusvaheliste piirmäärade alusel valdavalt

väike- ja mikroettevõtted, üksikute eranditega, on selle turu kontekstis selline jaotus asjakohane. On näha, et Eesti audiitorteenuste turg on koondunud suurte AEV-de kätte – nad moodustavad 64% audiitorteenuste turu käibest, aga vaid 7% ettevõtete arvust (Tabel 2).

Tabel 2
Ülevaade Eesti audiitorteenuste turuosalistest ning nende jaotus suurusklassidesse

Suurusklass	**Audiitorteenuste käibe vahemik suurusklassis**	**Audiitorettevõtjate arv ja osakaal**	**Audiitorteenuste käive ja osakaal**
Suur	>1,0 mln EUR	8 (7%)	34 288 678 (64%)
Keskmine	0,1-1,0 mln EUR	57 (49%)	17 211 215 (32%)
Väike	<0,1 mln EUR	51 (44%)	2 015 497 (4%)
Kokku		116 (100%)	53 515 390 (100%)

Märkus: Audiitorteenuste käive on perioodi 01.07.2023-30.06.2024 kohta ning liigitamine on sama perioodi andmete põhjal
Allikas: Autori koostatud Audiitorkogult saadud andmete põhjal, kes on andmed koondanud audiitorettevõtjate tegevusaruannetest

Selleks, et teha andmete põhjal üldistusi audiitorteenuste turu ja Eesti AEV-de kohta kogumina, rakendatakse andmeanalüüsis valimi esinduslikkuse parandamiseks kaale (Sauga, 2020). Kaalude arvutamise võrrandid on tuletatud A. Sauga (2020) statistikaõpiku põhjal. Analüüsis kasutatakse kahte erinevat kaalu, millest esimene kohandab tulemusi lähtuvalt mittevastajate osakaalust suurusklasside lõikes ning selliselt kaalutud andmeid võib pidada üldistuseks audiitorettevõtjate kohta olenemata sellest, et audiitorettevõtjad on väga erineva suurusega (Võrrand 1). Teine kaalumise viis, mida analüüsis kasutatakse, on üldistus audiitorteenuste turule, mis iseloomustab tõenäoliselt paremini seda, kuidas tajuvad turgu audiitorteenuste tarbijad. Audiitorteenuste turu mahtu arvestatakse käesolevas uurimistöös AEV audiitorteenuste käibe alusel perioodil 01.07.2023-30.06.2024. Kaalumine käibe alusel on ka Oslo käsiraamatu kohaselt tavapärane (OECD/Eurostat, 2018). Uuringu disainist ja autoriga seotud piirangutest lähtuvalt ei ole uuritavatelt küsitud täpset käivet, vaid käibe vahemik. Turumahule kaalumiseks on kasutatud suurusklassi kuuluvate AEV-de keskmise käibe ning suurusklassi käibe osakaalu kogu audiitorteenuste käibest korrutist (Võrrand 2). Audiitorteenuste turu üldistuse tegemiseks rakendatakse kogutud andmetele mõlema kaalu korrutist (Võrrand 1, Võrrand 2).

Võrrand 1

$$Kaal_vastajad_i = \frac{N_i}{n_i}$$

kus n_i – vastajate arv suurusklassis
N_i – audiitorettevõtjate arv suurusklassis
$Kaal_vastajad_i$ – suurusklassi *i* mittevastajate korrigeerimise kaal

Võrrand 2

$$Kaal_turg_i = \frac{R_i}{N_i} \times \frac{N}{R}$$

kus N – audiitorettevõtjate arv kokku
N_i – audiitorettevõtjate arv suurusklassis
R – audiitorteenuste käive kokku
Ri – audiitorteenuste käive suurusklassis
$Kaal_turg_i$ – suurusklassi *i* kaal

Tabel 3
Andmete põhjal arvutatud kaalud

Suurusklass	**Kaal_vastajad (**Võrrand 1**)**	**Kaal_turg (**Võrrand 2**)**
Suur	1,33	9,29
Keskmine	2,71	0,65
Väike	7,29	0,09

Märkus: Kaal_vastajad näitab mitut ettevõtet oma suurusklassis esindab iga vastaja, Kaal_turg näitab, mitmega tuleb korrutada suurusklassi kuuluva ettevõtte keskmine käive selleks, et saada tulemuseks suurusklassi kogukäive.

Andmete põhjal arvutatud kaalud on toodud tabelis 3. Mittevastajate osa korrigeeriva kaalu tõlgendus on, mitut AEV-d oma suurusklassis iga vastaja esindab. Kuna väikeste AEV-de hulgas oli vastamisaktiivsus kõige madalam, on kaal ka kõige kõrgem. Kaalu kasutamine tagab, et suurusklasside lõikes esinev erinevus vastamisaktiivsuses saab tasandatud. Audiitorteenuste turu kaal näitab kui palju kordi on ühe suurusklassi kuuluva AEV keskmine käive suurem kõikide AEV-de keskmisest käibest. See kaal tagab, et turule tehtud üldistustes on igal suure AEV vastusel oluliselt suurem mõju keskmiste arvutamisel kui keskmiste ja väikeste AEV-de vastustel. Selliselt kaalumine võimaldab anda hinnangu, millisena tunnetavad audiitorteenuste turgu teenuste tellijad.

Veapiiride arvutamine. Binaarsete tunnuste esinemise tõenäosuse veapiiride arvutamisel on kasutatud Wilsoni meetodit (Wilson, 1927), mida on kohandatud kaalutud andmete jaoks (Võrrand 3). Wilsoni meetod on sobilik meetod küsitluse andmete binaarsete andmete proportsiooni usaldusvahemike määramisel, mis tagab, et usalduspiirde alumised otsad ei lähe negatiivseks ning on sobilik ka väikeste valimite puhul (Franco jt, 2019).

Võrrand 3

$$Alumine\ piir = \frac{(2n_w\hat{p} + z^2) - z * \sqrt{z^2 - \frac{1}{n_w} + 4n_w\hat{p}(1-\hat{p})}}{2(n_w + z^2)}$$

$$Ülemine\ piir = \frac{(2n_w\hat{p} + z^2) + z * \sqrt{z^2 - \frac{1}{n_w} + 4n_w\hat{p}(1-\hat{p})}}{2(n_w + z^2)}$$

kus n_w – kaalutud vastajate arv
$\hat{p}$ – positiivsete vastuste kaalutud osakaal
z – Z-väärtus vastavalt usaldusnivoole (0,95 puhul z = 1,96)

Intervallskaalal antud hinnangute kaalutud keskmiste väärtuste veapiiride arvutamisel on kasutatud *Bootstrap* meetodit 10 000 simulatsiooniga. Iga simuleeritud valimi põhjal on leitud kaalutud keskmine hinnang ning veapiiridena rakendatakse kõikide simuleeritud hinnangute protsentiile. Usaldusnivoo 0,95 juures on alumise piirina kasutatud 2,5% protsentiili ja ülemise piirina 97,5% protsentiili, mis määravad jaotuse keskosa, kuhu jääb 95% andmetest. *Bootstrap* meetodit peetakse kasulikuks just väikeste valimitega töötamisel. (Davidson & Hinkley, 1997)

Andmeanalüüs. Pärast esmast andmete töötlemist Excelis on edasine andmeanalüüs läbi viidud ja joonised koostatud tarkvaras R ja RStudio. Koodi kirjutamisel on suheldud ka ChatGPT 4o tekstirobotiga lahenduse ideede saamiseks, koodi kirjutamiseks ja kommenteerimiseks, tekkinud vigade kontrollimiseks ja parandamiseks (OpenAI, 2025). Autor on koodi kirjutanud ka ise, seda muutnud ja kohandanud, kõik tekstiroboti pakutud koodijupid on autori poolt kontrollitud veendumaks analüüsi õigsuses. Programmi R kood on toodud lisas LISA F.

Uurimistulemuste tõlgendamiseks AEV-de innovaatilisuse põhinäitajate raames ning võrdlemiseks Statistikaameti poolt läbi viidud innovatsiooniuuringuga moodustati küsimuste põhjal fiktiivne muutuja, mis tähistab, kas tegemist on innovaatilise AEV-ga vaatlusalusel perioodil. Selleks, et luua konteksti uuringu raames antavale hinnangule Eesti AEV-de innovaatilisuse kohta, on analüüsitud vastuseid küsimustele, mis on sarnased ettevõtete innovatsiooniuuringuga. Selleks on välja arvutatud kaalutud keskmine innovaatiliste AEV-de osakaal, üksnes protsessiinnovatsiooniga, üksnes teenuseinnovatsiooniga ja mõlemat tüüpi innovatsiooniga ettevõtete osakaalud. Lisaks sellele analüüsitakse ka seda, kes on protsessiinnovatsioonide elluviija, millistes valdkondades protsessiinnovatsioone on ellu viidud ja mis on peamine põhjus, miks ei ole ellu viidud rohkem innovatsioone. Rakendatud

on mittevastajate kohandust (Võrrand 1) ning rohkemate innovatsioonide puudumise peamise põhjuse puhul ka turu kohandust (Võrrand 1, Võrrand 2). Hinnangute veapiirid on arvutatud Wilsoni meetodil.

Arenevate tehnoloogiate kasutuselevõtu analüüsimisel on tulemused laiendatud nii AEV-de kogumile kui ka audiitorteenuste turule. Erinevalt kaalutud tulemused võimaldavad teha järeldusi erinevas kontekstis ja kaudselt tuvastada erinevusi suuremate ja väiksemate AEV-de vahel, mida on soovitatud uurida ka varasemas kirjanduses (Barr-Pulliam jt, 2022; Ruhnke, 2023; Vitali & Giuliani, 2024). Binaarsetel andmetel põhinevate kaalutud osakaalude leidmiseks on väärtusele 1 kodeeritud „jah“ vastused korrutatud läbi kaalu(de)ga ning seejärel jagatud kogu AEV-de populatsiooniga (N=116). Osakaalude illustreerimiseks on koostatud tulpdiagrammid, millel on kujutatud ka veapiirid usaldusnivoo 0,95 juures Wilsoni meetodil. AEV-de kogumile ja audiitorteenuste turule laiendatud tulemusi on võrreldud. Sellise lähenemisega on analüüsitud järgmist:

- 30.06.2024 seisuga audiitorteenuste osutamisel AT kasutusele võtmise osakaal iga AT kohta ja vähemalt ühe AT kohta;
- AT-de väljatöötajate lõikes (AEV ise, AEV koostöös võrgustiku teise liikmega, AEV koostöös muu ettevõttega ja teine ettevõte) osakaal;
- innovatsioonitöötajate kaasatuse alusel osakaalude jaotus nii tüüpide lõikes kui binaarsena – kas on innovatsioonitöötaja roll täidetud või mitte.

Tuvastamaks seost AT kasutuselevõtmises nähtava kasu ja kasutuselevõtmise tõenäosuse vahel, on läbi viidud läbi paariviisiline korrelatsioonanalüüs, mille käigus on korrelatsioonikordajad välja arvutatud iga AT (kokku 11) kohta järgmiste paaride viisi: 1) Crameri V korrelatsioonikordajad AT kasutuselevõtust oodatav kasu ja AT 30.06.2024 seisuga kasutamise fakti vahel; ja 2) Spearmani korrelatsioonikordajad AT kasutuselevõtust oodatava kasu ja AT nelja aasta jooksul kasutusele võtmise tõenäosuse vahel (LISA D). Crameri V on valitud järjestiktunnuse ja binaarse tunnuse vahelise korrelatsiooni arvutamiseks ning Spearmani kordaja on valitud kahe järjestiktunnuse vahelise korrelatsiooni arvutamiseks. Esimeses analüüsis on tulemuse statistilist olulistust testitud Fischeri testiga, teises analüüsis Feldti ja Duhacheki meetoditega. Korrelatsioonikordaja näitab kahe muutuja vahelist seost skaalal -1 kuni +1. Positiivne seos viitab sellele, et kui ühe muutuja väärtus on kõrgem, siis on ka teise muutuja väärtus keskmiselt kõrgem. Korrelatsioon ei väljenda seoses põhjuslikkust. Tavapäraselt peetakse seost nõrgaks kui korrelatsioonikordaja absoluutväärtus on ≤0,3; mõõdukaks kui see on vahemikus 0,3-0,7 ja tugevaks kui see on ≥0,7 (Sauga, 2020).

Intervallskaalal antud hinnangute osas on leitud kaalutud keskmised hinnangud koos *Bootstrap* meetodil leitud veapiiridega. Kaalutud keskmise hinnangu arvutamiseks on korrutatud vastuste väärtused esmalt kaalu(de)ga ning seejärel jagatud AEV-de koguarvuga (N=116). Tulemuste visualiseerimiseks on loodud graafikud, kus on välja toodud ka veapiirid. Võrreldud on kaalutud keskmisi AEV-de kogumile ning audiitorteenuste turule laiendatuna. Hindamaks, kas keskväärtusi mõjutavad skaala äärmustesse jäävad vastused või pigem keskmised, on arvutatud kõikide vastusevariantide osakaalud ning võrreldud neid suurusklasside lõikes. Analüüsiks on tulemused visualiseeritud virnastatud tulpdiagrammidena. Seda metoodikat on rakendatud järgmiste teemade puhul:

- kui tõenäoliseks peetakse AT-de kasutuselevõtmist audiitorteenuste osutamisel järgmise nelja aasta jooksul ja sama küsimus võrdlemaks vastuseid nende vahel, kelle on ja kellel ei ole innovatsioonitöötaja;
- kui mõjusaks peetakse seitset erinevat AT-de kasutuselevõtu kasutegurit;
- kui mõjusaks peetakse üheksat erinevat AT-de kasutuselevõttu takistavat tegurit.

4. Tulemused ja analüüs

4.1 Ülevaade vastustest

Küsimustikule vastas kokku 34 AEV-d, mis moodustab 29% kõikidest AEV-dest, sealjuures on vastanud 75% suurtest, 37% keskmistest ja 14% väikestest AEV-dest (Tabel 4).

Tabel 4
Vastamisaktiivsus suurusklasside lõikes

Suurusklass	**Audiitorettevõtjate arv klassis**	**Vastajate arv**	**Vastajate osakaal**
Suur	8	6	75%
Keskmine	57	21	37%
Väike	51	7	14%
Kokku	116	34	29%

Märkus: Suurusklasside jaotus on välja toodud eespool tabelis 2

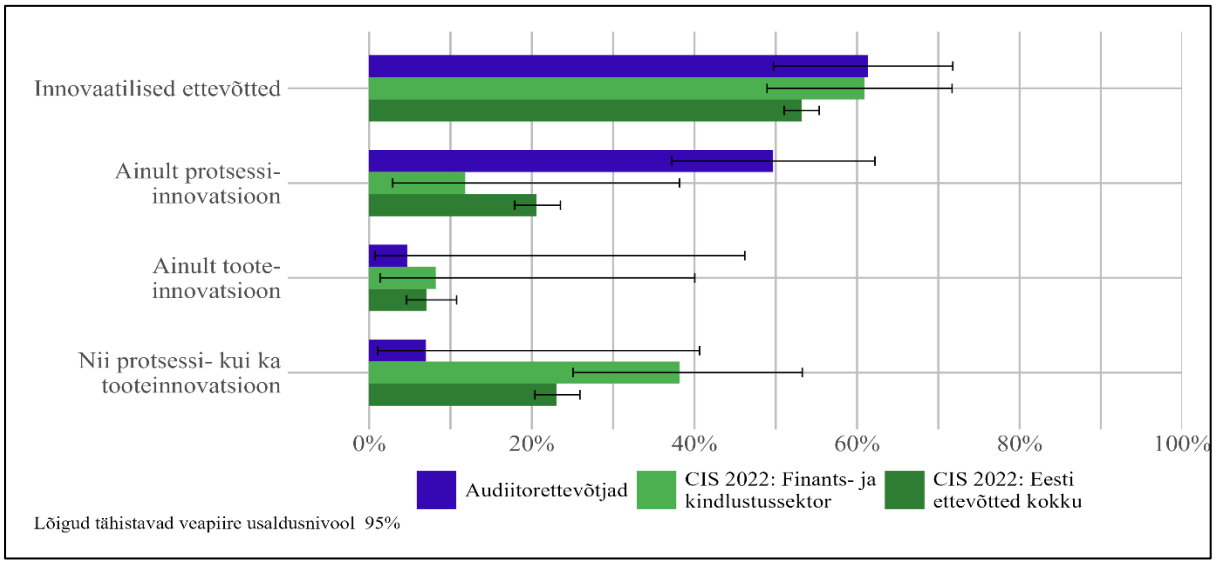
Küsimustiku reliaabluse hindamiseks arvutati välja Cronbachi alfa. Pärast selliste küsimuste eemaldamist, mille hulgas oli palju „ei tea“ vastuseid, oli Cronbachi alfa 0,94, mis on statistiliselt oluliselt kõrgem kui 0,8, mis viitab küsimustiku heale reliaablusele – andmed on sobivad edasise analüüsi läbiviimiseks (LISA D).

Vastamisaktiivsus on sellist tüüpi uurimuse jaoks olnud hea. Arvestades kogu populatsiooni suurust ületas saadud vastuste arv autori ootusi. Edasistes tulemustes on aga näha, et statistilise analüüsi jaoks on andmepunktide arv siiski väike, mis toob kaasa suured

veapiirid enamiku tulemuste puhul. Seega tuleb edasisi tulemusi tõlgendada ettevaatlikult. Autori hinnangul on vastuste hulk piisav uurimuse eesmärkide täitmiseks.

4.2 Eesti audiitorettevõtjate innovaatilisuse põhinäitajad

Kogutud andmete põhjal oli 57% Eesti AEV-dest perioodil 01.07.2022-30.06.2024 innovaatilised ehk raporteerisid vähemalt ühe teenuse- või protsessiinnovatsiooni. Statistikaamet on leidnud, et perioodil 2020-2022 oli 53% Eesti ettevõtetest innovaatilised, finants- ja kindlustussektoris oli vastav näitaja 61% (Statistikaamet, 2024). Innovaatiliste ettevõtete osakaal ei ole AEV-de hulgas statistiliselt oluliselt erinev Eesti ettevõtete keskmisest ega ka finants- ja kindlustussektori ettevõtete keskmisest.

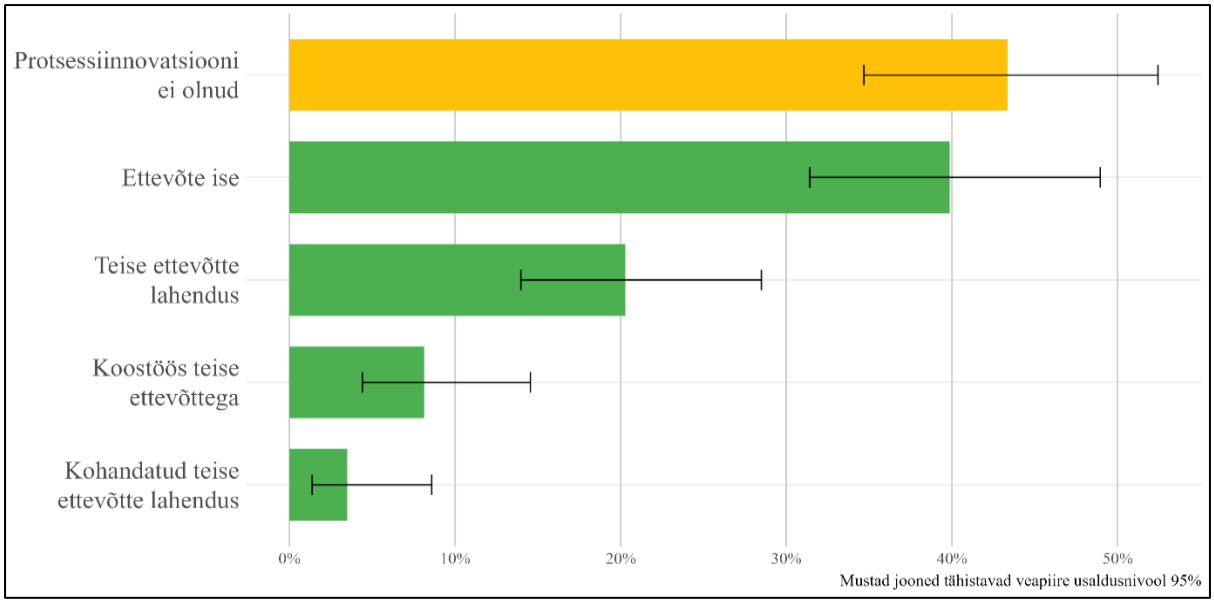


Joonis 5 Innovaatiliste ettevõtete osakaal innovatsiooni tüüpide lõikes – audiitorettevõtjad võrdluses Eesti ettevõtete innovatsiooniuuringu (2022) tulemustega finants- ja kindlustustegevuse sektori ning innovatsiooni põhitegevusaladega
Märkus: Eesti ettevõtete innovatsiooniuuringu andmed (CIS 2022) pärinevad Statistikaametist ning on perioodi 2020-2022 kohta, käesolev uuring on perioodi 01.07.2023-30.06.2024 kohta (Statistikaamet, 2024). Veapiirid on autori poolt arvutatud Wilsoni meetodil.
Allikas: Autori koostatud käesoleva uuringu ja Statistikaameti andmete põhjal

Joonisel 5 on näidatud andmete põhjal leitud osakaalud ning vastuste hulga põhjal arvutatud veapiirid usaldusnivool 0,95. Tulemuste tõlgendamisel saab arvestada, et statistiliselt oluline erinevus esinev siis kui võrreldavate gruppide usaldusvahemikud üheski osas ei kattu. Statistiliselt olulised erinevused selguvad kahes aspektis: esiteks, AEV-de seas on mõlema grupiga võrreldes üksnes protsessiinnovatsiooniga ettevõtete osakaal suurem, teiseks, AEV-de seas on mõlema grupiga võrreldes nii protsessi- kui tooteinnovatsiooniga

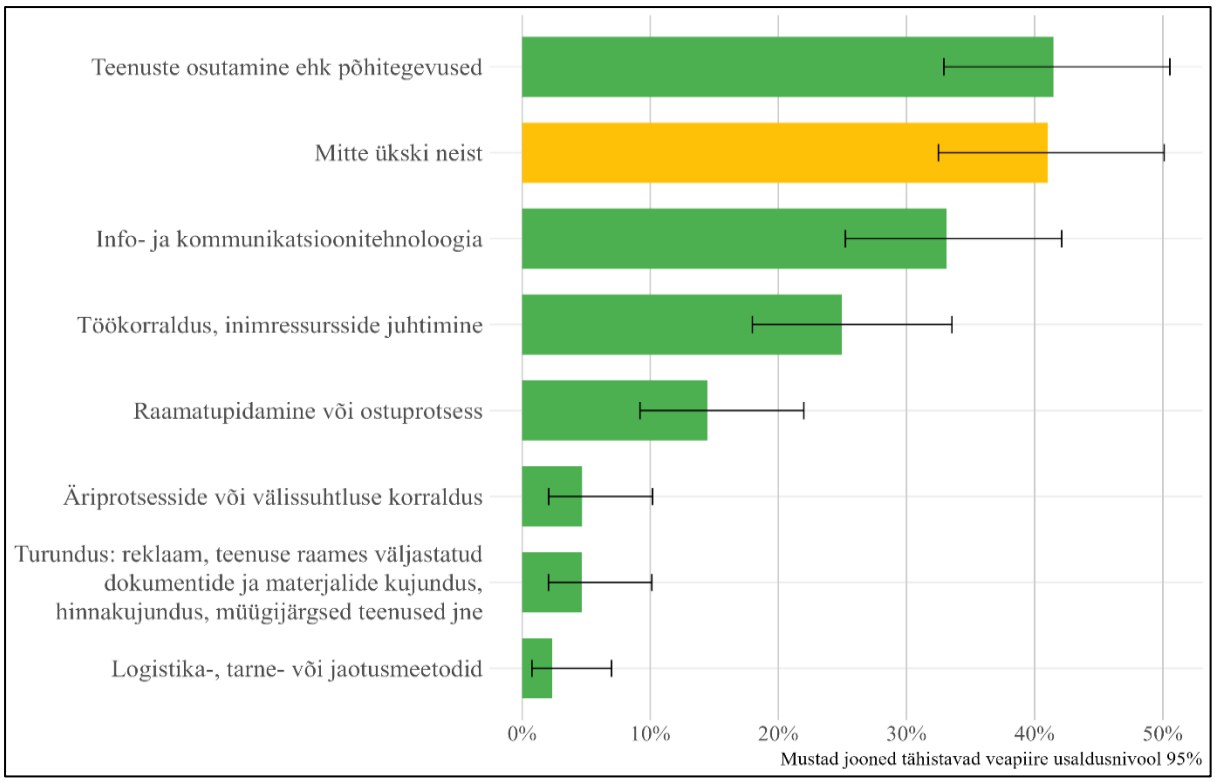
ettevõtete osakaal väiksem. Tulemused näitavad, et tooteinnovatsiooniga AEV-de osakaal on väga väike, mistõttu ei ole tooteinnovatsiooni detailsemalt analüüsitud.

Valdav osa AEV-dest, kes on innovatsiooni läbi viinud, on seda teinud ainult protsessiinnovatsioonina. Joonisel 6 on näha, et ilma protsessiinnovatsioonita oli 43% AEV-dest, mis näitab, et rohkem kui pooled on oma protsessides uuendusi läbi viinud. 40% AEV-dest on ise olnud vähemalt osade protsessiinnovatsioonide väljatöötaja ning 20% on rakendanud teiste ettevõtte poolt välja töötatud lahendusi. Koostöö ja teiste ettevõtete protsesside kohandamine on protsessiinnovatsioonide puhul AEV-de hulgas vähe levinud. Vastamisel valiti kõik asjakohased variandid, ehk kui oli mitu protsessiinnovatsiooni erinevate väljatöötajatega, siis valiti mitu vastust.

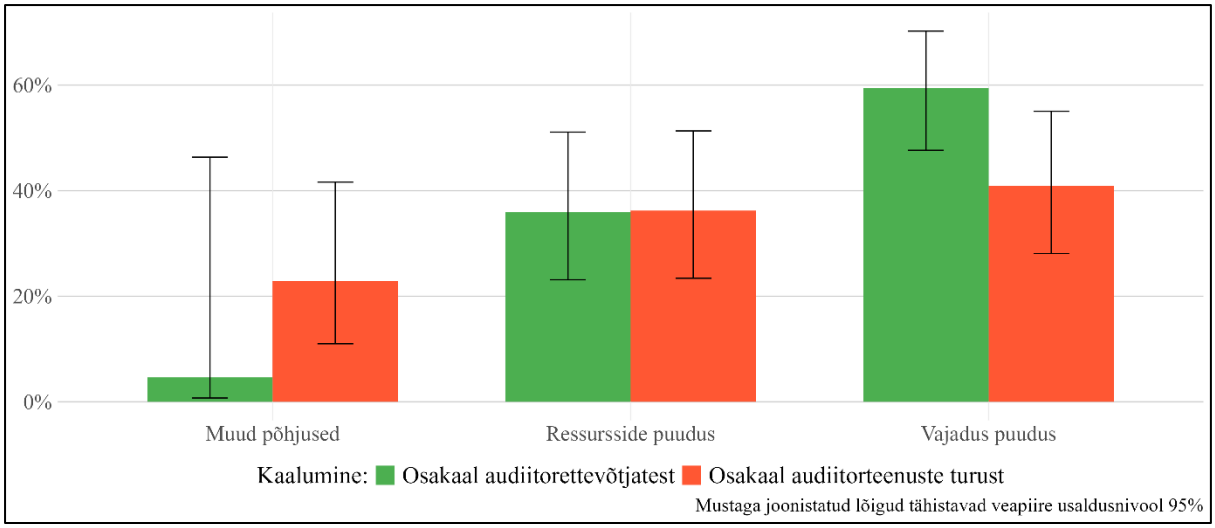


Joonis 6 Perioodil 01.07.2022-30.06.2024 AEV-des toimunud protsessiinnovatsioonide väljatöötajad osakaaluna audiitorettevõtjatest

Märkus: Kollasega on tähistatud selliste AEV-de osakaal, kellel ei ole protsessiinnovatsiooni olnud, rohelisega on tähistatud osakaal AEV-dest kelle puhul on protsessiinnovatsiooni väljatöötaja vastav isik. Vastates tuli määrata kõik õiged, seega võib samal AEV-l olla erinevate väljatöötajatega protsessiinnovatsioone. Veapiirid on arvutatud Wilsoni meetodiga.



Joonis 7 Valdkondade lõikes perioodil 01.07.2022-30.06.2024 protsessiinnovatsiooni läbi viinud AEV-de osakaal kõikidest AEV-dest
Märkus: Valida oli võimalik mitu vastusevarianti, sealjuures võib ka sama protsessiinnovatsioon avaldada mõju mitmes valdkonnas. Kollasega on tähistatud selliste AEV-de osakaal, kellel ei ole küsitud valdkondades protsessiinnovatsiooni olnud, ülejäänud on tähistatud rohelisega. Veapiirid arvutatud Wilsoni meetodiga.



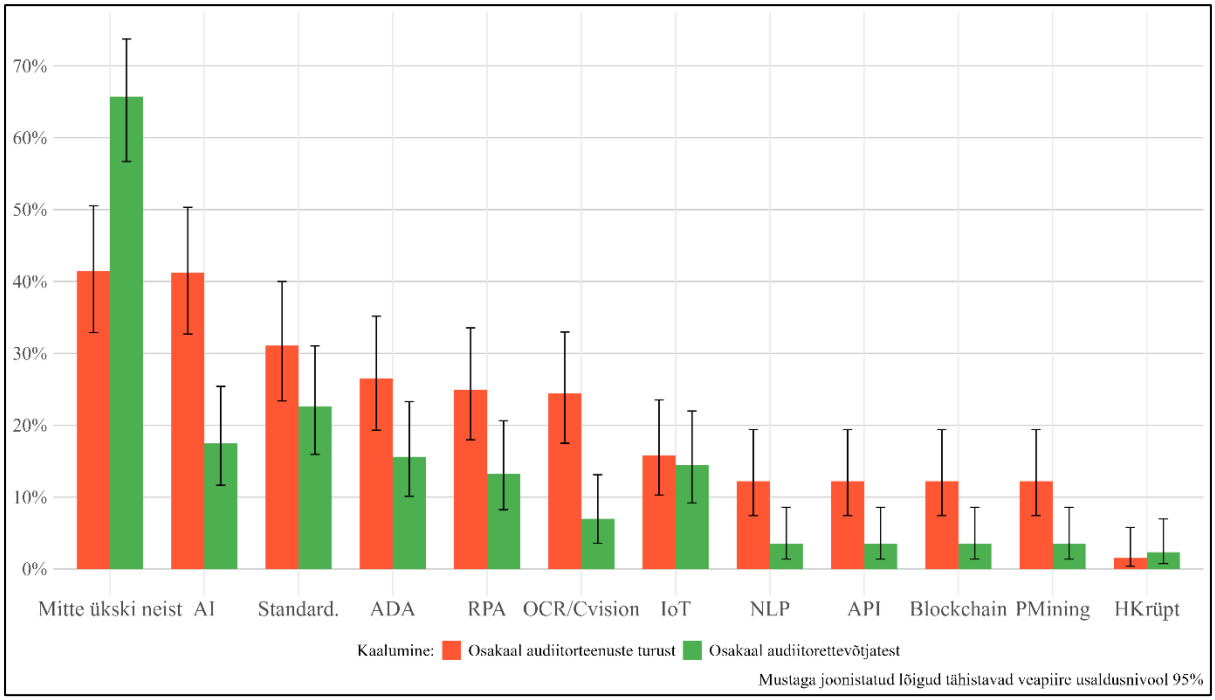
Joonis 8 Peamine põhjus, miks ei ole perioodil 01.07.2022-30.06.2022 ellu viidud rohkem innovatsioone
Märkus: Veapiirid on arvutatud Wilsoni meetodiga

Joonisel 7 on näidatud, millistes valdkondades AEV-d protsessiinnovatsioone läbi viinud. On näha, et 41% AEV-dest on viinud ellu uuendusi oma põhitegevustes ehk teenuste

osutamisel, 33% IKT kasutamises ning 25% töökorralduses ja inimressursside juhtimises. Ülejäänud valdkondades, sealhulgas turunduses, on innovatsioone ellu viinud vähesed AEV-d. Joonisel 8 on välja toodud peamine põhjus, miks AEV-d ei ole läbi viinud rohkem innovatsioone. Ressursside puudus on peamine takistus 36% AEV-de jaoks (turust samuti 36%), kõige suurema osa AEV-de jaoks (59%) oli peamiseks põhjuseks aga vajaduse puudus – siin on aga turule kaalutud tulemuste põhjal osakaal vaid 41%, mis viitab sellele, et väiksemate AEV-de puhul on vajadus rohkemate innovatsioonide järele väiksem kui suurtes AEV-des.

4.3 Arenevate tehnoloogiate kasutuselevõtt

Uuringu tulemustest selgub, et enamik Eesti AEV-dest ei ole 30.06.2024 seisuga kasutusele võtnud ühtegi uurimuse objektiks olevat arenevat tehnoloogiat (Joonis 9). Turule laiendatud tulemuse põhjal on 59% turust juba vähemalt ühe AT kasutusele võtnud. See näitab, et suuremad AEV-d on rohkem jõudnud AT-de kasutamiseni kui väiksemad AEV-d. Joonisel 9 on välja toodud, kui suur osa AEV-dest ja audiitorteenuste turust on 30.06.2024 seisuga erinevaid AT-sid kasutusele võtnud.

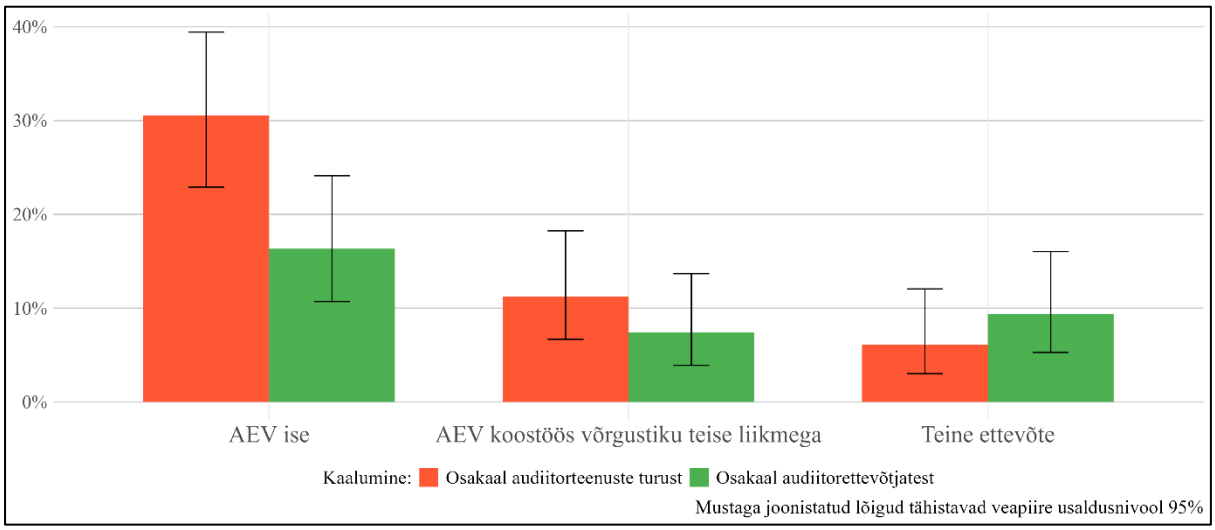


Joonis 9 Arenevate tehnoloogiate kasutuselevõtt 30.06.2024 seisuga osakaaluna Eesti audiitorettevõtjatest ning audiitorteenuste turust
Märkus: Veapiirid on arvutatud Wilsoni meetodiga.

Kõige rohkem on juba kasutusele võetud tehisintellekti tehnoloogiat, mis on kasutusel ligikaudu 41% turust, järgnevad andmete standardiseerimine (31%), täiustatud andmeanalüüs

(26%), robootiline protsesside automatiseerimine (29%) ja OCR / arvutinägemine (24%) (Joonis 9). Kõige vähem on seni kasutusele võetud homomorfse krüpteerimise tehnoloogiat. AT-de kasutuselevõttu soodustab see kui optimistlik või entusiastlik on selles osas ettevõtte juhtkond (Zemankova, 2019). Analüüsimaks, kas Eesti AEV-de puhul on näha, et AT-de suhtes optimistlikuma juhtkonnaga AEV-d on võtnud AT kasutusse tõenäolisemalt, on analüüsitud, kas AT-st nähtava kasu ja kasutuselevõtu vastused on omavahel seotud. Läbi viidud Crameri V korrelatsioonianalüüs näitas, et statistiliselt olulisi seoseid ühegi AT puhul ei selgunud ning samuti ei olnud ükski seos tugev, kuigi kõik seosed olid positiivse märgiga (LISA C).

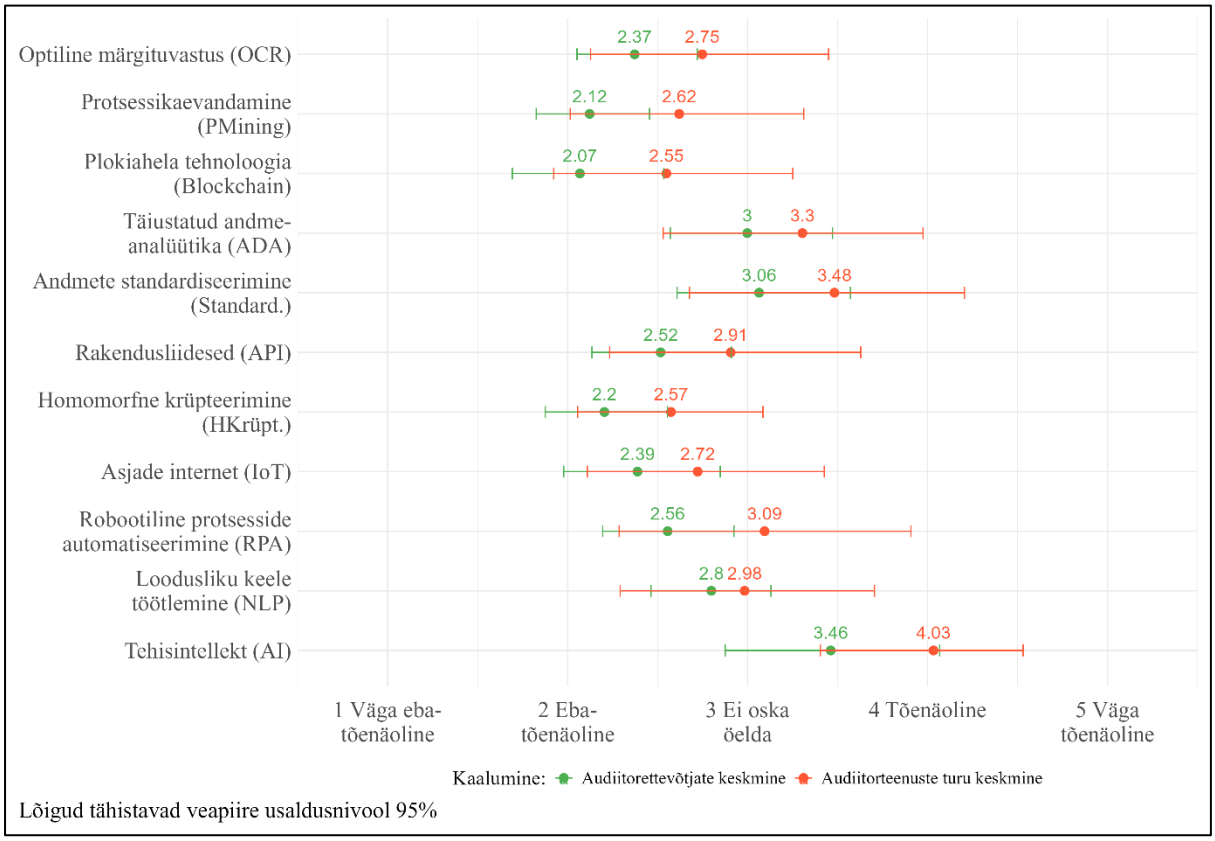
Analüüsides, kes on olnud AT väljatöötaja, on joonisel 10 näha, et kõige rohkem on väljatöötajaks olnud AEV ise, nagu ka protsessiinnovatsioonide puhul (Joonis 6). Küsimustikule vastajate seas ei olnud ühtegi AEV-d, kes oleks AT välja töötanud koostöös teise, samasse võrgustikku mittekuuluva ettevõttega.



Joonis 10 Kasutusele võetud AT-de väljatöötajate jaotus osakaaluna Eesti audiitorettevõtjatest ning audiitorteenuste turust
Märkus: AEV – audiitorettevõtja. Andmed on esitatud osakaaluna kogu turust ja kõikidest audiitorettevõtjatest, jooniselt on välja jäetud need, kellel ei ole AT kasutusele võetud. Sama ettevõte võis valida mitu vastusevarianti, seega ei anna tulpade summa tervikut. Veapiirid on arvutatud Wilsoni meetodiga.

Analüüsides seda, kui tõenäoliseks peavad AEV-d arenevate tehnoloogiate kasutuselevõttu audiitorteenuste osutamisel järgmise nelja aasta jooksul, on üldiselt joonisel 11 näha, et pigem ei peeta seda tõenäoliseks – kaalutud keskmised hinnangud jäävad alla intervallskaala keskpunkti. Vastused peegeldavad üksnes kasutuselevõtmise tõenäosust audiitorteenuste osutamisel, seega võib muu äritegevuse käigus AT-de kasutamise tõenäosus

olla erinev. Enamiku tehnoloogiate puhul jääb hinnang veidi alla keskmise ehk kaldub pigem ebatõenäolise poole. Kõige tõenäolisemaks peetakse tehisintellekti kasutuselevõttu, millele järgneb andmete standardiseerimine ja täiustatud andmeanalüütika. Ka nende vastuste puhul on märgata, et AEV-dele kaalutud tulemused on madalamad kui audiitorteenuste turule kaalutud tulemused, mis näitab, et suurema käibega AEV-d peavad AT-de kasutuselevõttu tõenäolisemaks kui madalama käibega AEV-d.



Joonis 11 AEV-de kaalutud keskmised hinnangud AT-de audiitorteenuste osutamisel kasutusele võtmise tõenäosuse kohta järgmise nelja aasta jooksul skaalal 1 – „väga ebatõenäoline“ kuni 5 – „väga tõenäoline“
Märkus: Veapiirid on arvutatud *Bootstrap* meetodiga

Mida optimistlikum on juhtkond AT-de osas, seda tõenäolisem on AT kasutuselevõtt (Zemankova, 2019). Kriegeri jt (2021) poolt väljatöötatu protsessis tähendaks see, et juhtkond on tõenäolisemalt valmis ressursse allokeerima. Käesolevad uurimuses on AEV hinnangut AT potentsiaalse kasu kohta kasutatud lähendina juhtkonna optimismi väljendamiseks. Seega võib oodata, et mida suurem on AT-s nähtav kasu, seda tõenäolisemaks peetakse selle AT kasutuselevõttu. Tabelis 5 on toodud iga AT kohta arvutatud Spearmani korrelatsioonid AT-st oodatava kasu ning selle tehnoloogia kasutuselevõtmise tõenäosuse vahel. Paariviisilise korrelatsioonianalüüsi tulemused näitavad,

et kõikide AT-de puhul on seos positiivne. See tähendab, et kui nähakse suuremat kasu, siis on ka järgmise nelja aasta jooksul kasutuselevõtmise tõenäosus suurem. Tabelis 5 on toodud nii kaalumata kui ka AEV-dele ja audiitorteenuste turule laiendatud tulemused. Kõik korrelatsioonikordajad, välja arvatud üks on statistiliselt olulised. Seose tugevus on valdavalt mõõdukas, tugev seos on täheldatav üksnes tehisintellekti ja loomuliku keele töötluse puhul.

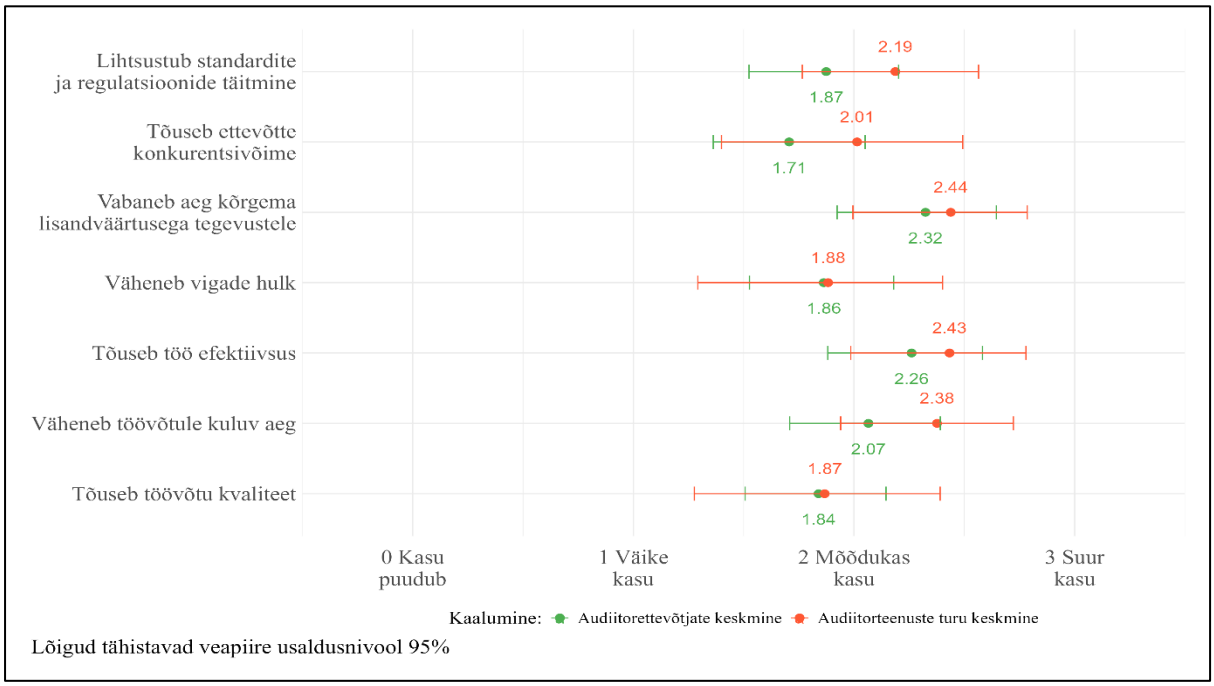
Tabel 5
Spearmani (kaalutud) korrelatsioonikordajad AT kasutuselevõtust nähtava kasu ja AT nelja aasta jooksul kasutusele võtmise tõenäosuse vahel

Paar – kasu vs kasutuselevõtu tõenäosus	Kaalumata		Mittevastajate kaaluga		Mõlema kaaluga	
	r	**p**	**r**	**p**	**r**	**p**
AI	0,67	0,000	0,74	0,000	0,55	0,001
NLP	0,66	0,000	0,62	0,000	0,75	0,000
RPA	0,46	0,006	0,42	0,014	0,66	0,000
IoT	0,40	0,020	0,45	0,007	0,44	0,009
HKrüpt	0,58	0,000	0,65	0,000	0,44	0,009
API	0,43	0,011	0,50	0,003	0,35	0,042
Standard.	0,42	0,014	0,40	0,019	0,42	0,014
ADA	0,39	0,022	0,47	0,005	0,32	0,068
Blockchain	0,41	0,015	0,43	0,010	0,42	0,014
PMining	0,41	0,015	0,45	0,008	0,56	0,001
CVision	0,53	0,001	0,55	0,001	0,44	0,009

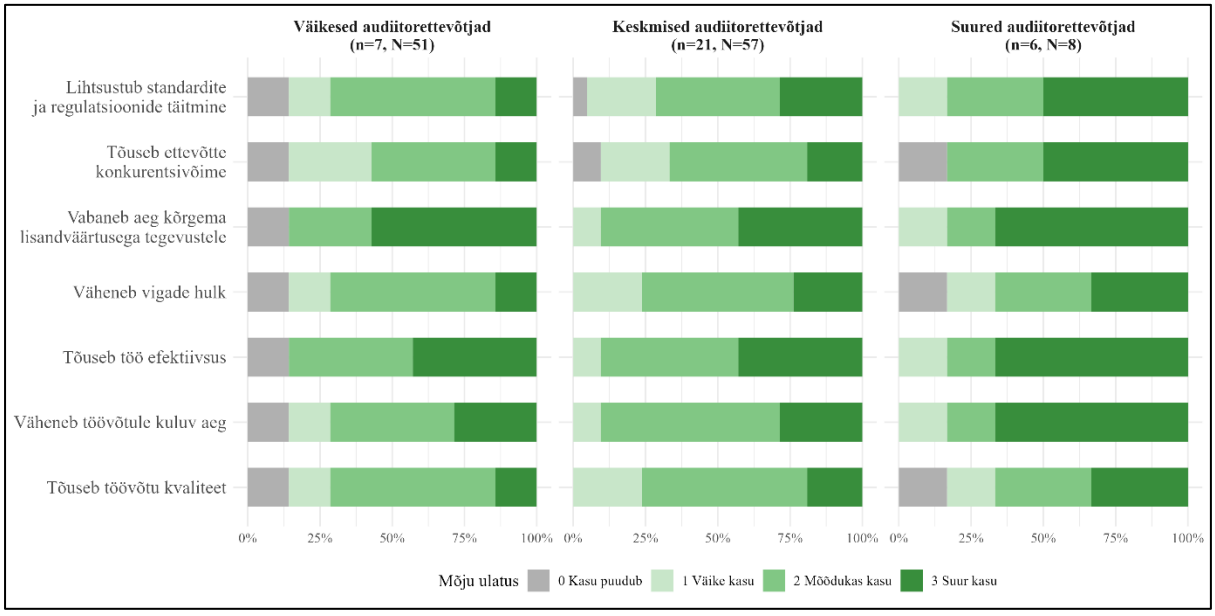
Märkus: r – Spearmani (kaalutud) korrelatsioonikordaja, p – olulisuse tõenäosus, mittevastajate kaal - Võrrand 1, mõlema kaaluga - Võrrand 1 ja Võrrand 2 korrutis. Tumesinisega on tähistatud tugevad seosed, sinisega mõõdukad seoses, rohelise taustaga on tähistatud statistiliselt olulised olulisuse nivoo 0,95 juures. AT-de lühendite selgitused on välja toodud lisasLISA E.

Üldistades võib öelda, et kõik kirjanduse põhjal tuvastatud variandid, kuidas võib avalduda AT-de kasutuselevõtust saadav kasu, on potentsiaalsed kasutegurid ka Eesti AEV-de hinnangul, sealjuures kõige olulisem mõju on selles, et töötajatel vabaneb rohkem aega kõrgema lisandväärtusega tegevusetele, tõuseb töö efektiivsus ning väheneb ka töövõtule kuluv aeg (Ruhnke, 2023; Zemankova, 2019; Vitali & Giuliani, 2024;Joonis 12). Mõõdukat kasu nähakse ka mujal – lihtsustub standardite ja regulatsioonide täitmine, tõuseb AEV konkurentsivõime, väheneb vigade hulk tehtavas töös ning tõused töövõtu kvaliteet (Ruhnke, 2023; Zemankova, 2019; Vitali & Giuliani, 2024). Nende vastuste puhul on näha, et audiitorteenuste turu keskmine hinnang on enamasti kõrgem kui AEV-de keskmine hinnang, mis viitab, et suuremad AEV-d on keskmiselt optimistlikumad saavutatava kasu osas. Vaadates aga vastusevariantide jaotust suurusklasside lõikes, on näha, et vähemalt vastajate hulgas on suured AEV-d olnud pigem optimistlikumad kui väiksemad AEV-d – nende puhul on üldiselt suurt kasu nägevate vastajate osakaal suurem kui teistes kategooriates (Joonis 13).

Samas saab aga välja tuua, et mõõdukat või suurt kasu näevad kõikides suurusklassides kõikide kasutegurite puhul rohkem kui pooled vastajad.

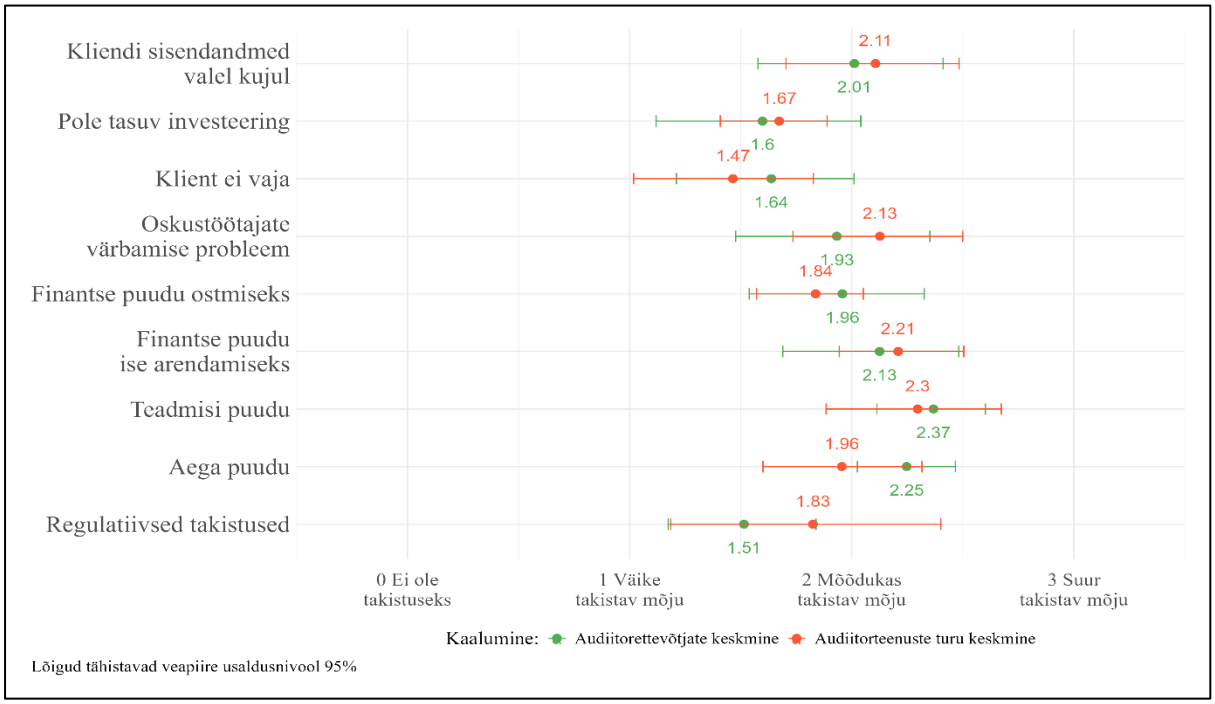


Joonis 12 AEV-de kaalutud keskmised hinnangud AT-de kasutuselevõtu kasutegurite mõju ulatusele skaalal 0 – „kasu puudub“ kuni 3 – „suur kasu“
Märkus: Veapiirid on arvutatud *Bootstrap* meetodil.



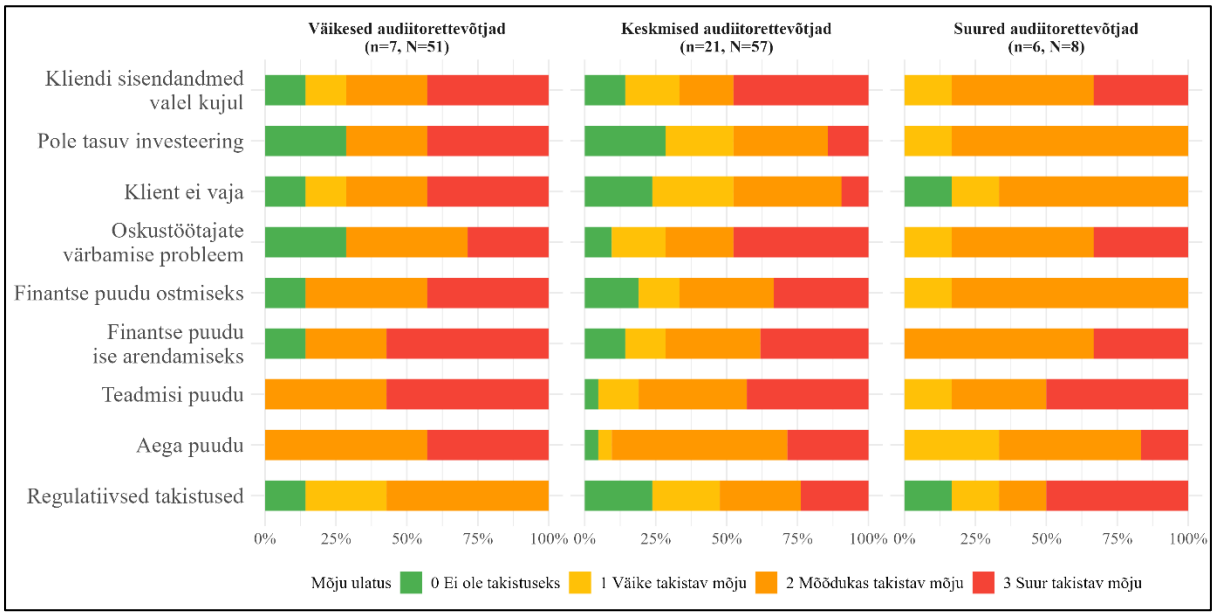
Joonis 13 AT-de kasutuselevõtust saadava kasu hinnangute osakaalud vastajate seas kasutegurite ja suurusklasside lõikes
Märkus: Osakaalud näitavad, mitu protsenti suurusklassi vastajatest valis konkreetse vastuse. N – audiitorettevõtjate arv suurusklassis, n – küsitlusele vastajate arv suurusklassis.

Ka Eestis on kõik üheksa kirjanduse põhjal tuvastatud AT-de kasutuselevõttu takistavat tegurit takistuseks – Eestis ei ole seega olukord oluliselt erinev (Krieger jt, 2021; Vitali & Giuliani, 2024; Wiklund & Fallan, 2024). Joonisel 14 on näha, et takistava mõju ulatus on skaalal 0-3 erinevate tegurite puhul varieeruv, aga jääb vahemikku 1,47-2,3 turu keskmise hinnangu alusel ning 1,51-2,37 AEV-de keskmise hinnangu alusel (Joonis 14). Ei esine tegurit, mille mõju keskmine hinnang läheneks nullile, mis viitaks, et tegemist ei ole takistusega Eesti AEV-de jaoks. Audiitorteenuste turule laiendatud keskmise alusel, mis rõhutab suuremate AEV-de hinnanguid, on suurematekstakistusteks (keskmine hinnang >2) takistuseks teadmiste puudus, finantside puudus ise AT-de arendamiseks, oskustöötajate värbamise probleem ning klientidelt saadavate sisendandmete vale vorming. AEV-de keskmise puhul, milles on suurem osakaal väikeste ja keskmise suurusega AEV-de hinnangutel, on suurimateks takistusteks (keskmine hinnang >2) teadmiste puudus, ajapuudus, finantside puudus ise arendamiseks ja klientide sisendandmete vale vorming. Samal joonisel on ka näha, et selle küsimuse osas ei jää turu keskmised ettevõtete keskmisest alati samale poole – järelikult on takistava mõju hinnangud suurusklasside lõikes eri tegurite osas erineva mõjuga ning seega on oluline analüüsida, millised tegurid on suurusklasside lõikes olulisema mõjuga (Joonis 14).



Joonis 14 Audiitorettevõtjate kaalutud keskmised hinnangud AT-de kasutuselevõttu takistavate tegurite mõju ulatusele skaalal 0 – „ei ole takistuseks“ kuni 3 – „suur takistav mõju“
Märkus: Veapiirid on arvutatud *Bootstrap* meetodil.

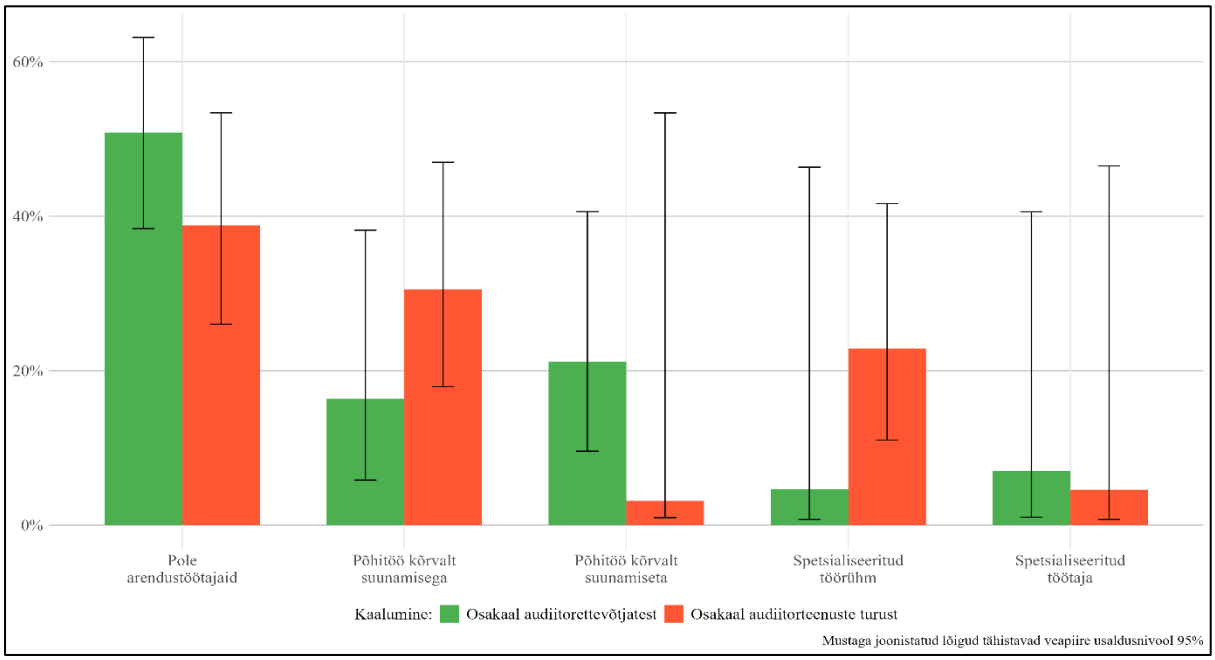
Joonisel 15 on välja toodud vastajate hinnangud takistavate tegurite mõjule suurusklasside lõikes. On näha, et üldiselt on suurte AEV-de hinnangute hulgas suure mõjuga hinnangute osakaal väiksem kui teistes suurusklassides. Näiteks ei ole vastajate hulgas ükski suur AEV pidanud suureks takistuseks investeeringu tasuvuse probleemi, kliendi vajaduse puudust ega AT ostmiseks vajalike finantside puudust. Suurtest AEV-dest 50% vastajate hinnangul regulatiivsed takistused suured, samas kui keskmiste puhul on vastav osakaal 24% ja väikeste AEV-de puhul ei ole ükski vastaja pidanud seda suureks takistuseks (Joonis 15). Ka on näha, et kuigi hinnangut „suur“ on suurte AEV-de puhul vähem kui teistes rühmades, on nende hulgas vähe ettevõtjaid, kes peavad takistusi väikseks või puuduvaks. Vaadeldes hinnanguid, mis on väiksemad kui mõõdukad (roheline ja kollane), on keskmise suurusega AEV-de hulgas suurim osakaal neid, kes on takistuste osas optimistlikumad.



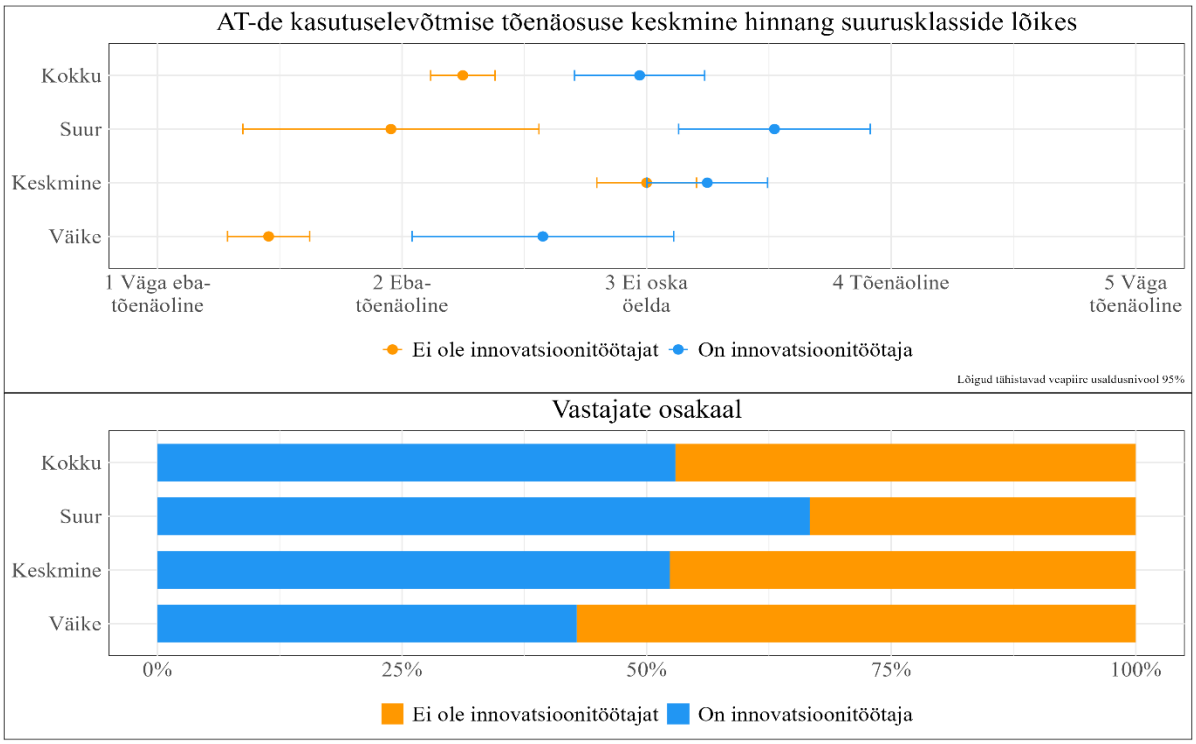
Joonis 15 AT-de kasutuselevõttu takistavate tegurite mõju hinnangud suurusklasside lõikes
Märkus: Osakaalud näitavad, mitu protsenti suurusklassi vastajatest valis konkreetse vastuse. N – audiitorettevõtjate arv suurusklassis, n – küsitlusele vastajate arv suurusklassis.

Joonisel 16 on välja toodud AEV-de jaotus innovatsioonitöötajate kaasatuse alusel. On näha, et ligikaudu pooltel (51%) AEV-del ei ole töötajaid, kes tegeleksid innovatsiooniga. Turu kontekstis on selliste osakaal 45% - järelikult on innovatsioonitöötajateta rohkem väikseid AEV-sid, kuid ka keskmiste ja suuremate seas on selliseid omajagu. Ülejäänud andmetulbad joonisel 16 näitavad, et on vaid üksikud AEV-d, kellel on tööl inimesed või inimene, kes tegelevad innovatsiooniga põhitööna. Ülejäänud AEV-des on vähemalt üks töötaja, kes põhitöö kõrvalt tegeleb innovatsiooniga – sealjuures on turule ja AEV-dele taandatud hinnangute põhjal näha, et juhtkonna poolse suunamisega on vastajate seas rohkem

suuremaid AEV-sid kui väikseid ja juhtkonna poolse suunamiseta innovatsioonitöötajaid kasutatakse peamiselt väiksemates AEV-des.



Joonis 16 AEV-de ja audiitorteenuste turu jaotus innovatsioonitöötajate kaasatuse alusel
Märkus: Veapiirid on arvestatud Wilsoni meetodil.



Joonis 17 AT-de kasutuselevõtmise tõenäosuse keskmine hinnang suurusklasside lõikes ja innovatsioonitöötaja olemasolu osakaal vastajate ja suurusklasside lõikes
Märkus: Kategooria „On innovatsioonitöötaja“ hõlmab kõiki variante alates ühest põhitöö kõrvalt tegelevast töötajast kuni spetsialiseerunud innovatsioonitiimini. AT-de kasutuselevõtu

tõenäosusena on kasutatud joonisel 11 näidatud üheteistkümne erineva tehnoloogia hinnangute aritmeetilist keskmist. Usalduspiirid on arvutatud Wilsoni meetodiga.

Joonise 17 alumises osas on näha, et vastajate hulgas on ilma innovatsioonitöötajata AEV-de osakaal seda kõrgem, mida väiksemasse suurusklassi vastaja kuulub, kuid selgelt on selliseid AEV-sid igas suurusklassis. Sama joonise ülemises osas on näha, et AEV-de keskmised hinnangud AT-de kasutuselevõtu tõenäosuse kohta on igas suurusklassis kõrgemad siis kui AEV-s on mingilgi kujul töötaja, kelle ülesanne on innovatsiooniga tegeleda. Olenemata suurtest veapiiridest on näha, et see erinevus on ka statistiliselt oluline.

5. Arutlus, järeldused

Uuringu tulemusena saab tõdeda, et Eesti AEV-d ei erine innovaatilisuse põhinäitajate poolest oluliselt Eesti ettevõttetest keskmiselt ega ka finants- ja kindlustussektorist, välja arvatud selles, et innovatsiooni AEV-des ilmestab suurem keskendumine protsessi-innovatsioonidele – neid on läbi viinud üle poole AEV-dest. Arvestades, et valdkonnas on teenuste sisu suuresti määratud rahvusvaheliste standarditega, on tulemus ootuspärane – uute audiitorteenustega tullakse turule harva ja pigem uutest regulatsioonidest ajendatuna. Lähiaastate suurem muutus selles osas on seotud kestlikkusaruannete audiitorkontrolli lisandumisega audiitorteenuste loetelusse (Audiitortegevuse seadus, 2025). Barrase (1986) pööratud toote elutsükli teooria raamistikus saab autori hinnangul paigutada Eesti audiitortegevuse valdkonna esimesse, samm-sammulise protsessiinnovatsiooni etappi, kus tegeletakse olemasolevate protsesside uuendamise ja parandamisega. AT-de kasutuselevõtu kiirendamine, mille üheks eelduseks on ka paindlikumad regulatsioonid, võib viia radikaalse protsessiinnovatsioonini, kus muutub oluliselt see, kuidas audiitorteenuseid osutatakse, näiteks liikumine pideva auditeerimise, analüütikal põhineva auditi suunas (Seidenstein jt, 2024).

Protsessiinnovatsioone lähemalt vaadates selgub, et kõige sagedamini on uuenduse välja töötanud AEV ise, kuid 20% AEV-dest on rakendanud ka teiste poolt välja töötatud lahendusi. Järelikult näeb suur osa AEV-dest võimalusi oma töö paremini organiseerimiseks ning on sellega ka aktiivselt tegelenud. 41% AEV-dest on protsesse uuendanud ka põhitegevuse valdkonnas. Arvestades, et enamiku AEV-de teenustest moodustavad audit ja ülevaatus, siis on põhjust arvata, et ka nendes on uuendusi tehtud. Autori hinnangul on teenuste osutamise arendamine audiitorteenuste valdkonnas vajalik, et tagada teenuste areng ning kvaliteet. Auditite kvaliteedi tõstmine on ka oluline motivaator, miks IAASB on otsustanud AT-de kasutuselevõttu toetada ning analüüsib aktiivselt ka vajadust muuta

kehtivaid standardeid (Seidenstein jt, 2024). AEV-de protsesside uuendamine võib olla vajalik ka uute regulatsioonidega vastavuses olemiseks, mistõttu tuleks tähelepanu pöörata AEV-dele, kes ei ole oma protsesse uuendanud. Uuritavasse ajavahemikku jäi näiteks uue Rahvusvahelise Kvaliteedijuhtimise Standardi (Eesti) 1 (Täiendatud) „Kvaliteedijuhtimine ettevõtete puhul, mis viivad läbi finantsaruannete auditeid või ülevaatusi või muid kindlustandvaid või seonduvate teenuste töövõtte“ jõustumine (alates 15.12.2022), mis oleks autori hinnangul enamiku AEV-de puhul pidanud nõudma ka protsesside uuendamist. Seega võib protsessiinnovatsioonide puudumine tekitada küsimusi ka AEV-de vastamises kehtivatele regulatsioonidele.

Seda enam on huvitav, et peamine põhjus, miks AEV-d ei ole viinud läbi rohkem innovatsioone, on vajaduse puudumine (Joonis 8 osas 4.2). Sellest võib järeldada, et 59% AEV-dest on kas suutnud ellu viia kõik innovatsioonid, mida nad on pidanud vajalikuks või ei pea nad innovatsioonide elluviimist üldse vajalikuks või ei ole nad suutnud tuvastada kohti oma tegevustes, mida oleks vajalik või võimalik uuendada. Käesolev uurimus ei anna vastuseid, millised on sellise tulemuse põhjused, kuid seda oleks huvitav edaspidi uurida. Kui põhjus on suutmatus tuvastada vajadust, võib see viidata täiendava koolitamise vajadusele. Tulemused näitavad, et vajaduse puudus on peamiseks põhjuseks pigem väiksematel AEV-del kui suurematel, mis muudab tulemuse tõlgendamise ja potentsiaalse edaspidise uurimise veelgi huvitavamaks. Peamine rohkemate innovatsioonide puudumise põhjus on 42% AEV-de (ja ka audiitorteenuste turu) puhul ressursside puudus, mis näitab, et on suur hulk AEV-sid, kes sooviksid rohkem innovatsioone läbi viia – seega potentsiaali valdkonna arenguks on. Tuues viimatise tulemuse aga AT-de kasutuselevõtu protsessi konteksti, võib ressursside puudus viidata sellele, et AT-de kasutuselevõtu protsess jääb ressursside allokeerimise sammul pooleli (Joonis 3 osas 2.3).

Enamik Eesti AEV-dest ei olnud 30.06.2024 seisuga võtnud audiitorteenuste osutamisel kasutusele veel ühtegi AT-d, kuid suuremates AEV-des on jõutud kaugemale ja kokkuvõttes on enamik (59%) audiitorteenuste turust siiski vähemalt ühe AT juba kasutusele võtnud (Joonis 9 osas 4.3). Kõige enam on kasutusele võetud tehisintellekti tehnoloogiaid, mis võib ühelt poolt tuleneda nende kasvavast populaarsusest inimeste igapäevaelus, aga ka sellest, et tehisintellekti rakendatakse väga paljudes erinevates rakendustes ning sellel on ka palju erinevaid potentsiaalseid kasutusvõimalusi auditi raames (IAASB, 2022b; Mikhaylova, 2024; Seidenstein jt, 2024).

Sarnaselt protsessiinnovatsioonidele on kasutusele võetud AT-dest kõige suurema osa välja töötanud AEV-d ise. Kooskõlas varasema kirjandusega (Seidenstein jt, 2024; Vitali &

Giuliani, 2024) on ka uuringu tulemustest näha, et ise väljatöötamise võimekus on olnud suuremates AEV-des kõrgem ja nad on ressursside allokeerimise sammul on otsustatud ise arendamise kasuks. See võimaldab paindlikumate ja kasutajaspetsiifilisemate lahenduste loomist kui valmis lahenduse ostmine turult (Krieger jt, 2021). Märkimisväärne on, et ükski küsimustikule vastanud AEV ei ole AT kasutuselevõttu läbi viinud koostöös teise, samasse võrgustikku mittekuuluva ettevõttega. Ühelt poolt näitab see süsteemi suletust – koostööd väliste osapooltega ei tehta. Samas on see potentsiaalne arenguhüppe koht – ühelt poolt võimaldaks koostöö erinevate väiksemate AEV-de vahel ressursse ühendada, mis muudaks investeeringu kulutõhusamaks, teiselt poolt, kui kaasata koostöösse ka IKT või andmeteaduse sektori osapooli, on võimalik ületada ka IKT alaste teadmiste puudujääk (Seidenstein jt, 2024). Koostöö teiste osapooltega võib AT-de kasutuselevõtu protsessi hõlbustada kõikides sammudes kuni esmase kasutuselevõtuni. Homomorfse krüpteerimise tehnoloogia kasutuselevõttu peetakse pigem ebatõenäoliseks (Joonis 11 osas 4.3), kuid just selle tehnoloogia rakendamine võimaldaks potentsiaalselt AI lahenduste väljatöötamisel masinõppe kaudu kasutada erinevate audiitorettevõtjate andmeid mudelite treenimiseks, sealjuures klientide andmete konfidentsiaalsust kaotamata (Seidenstein jt, 2024).

Suuremad AEV-d on hetkel optimistlikumad AT-de lähiajal kasutusele võtmise tõenäosuse osas (Joonis 11 osas 4.3). Tulemus on igati ootuspärane ning kooskõlas ka rahvusvahelise kirjandusega (Seidenstein jt, 2024), kuid väljendab seda, et väiksemad AEV-d on tuleviku perspektiivis kehvemas seisus ning võivad seetõttu kaotada konkurentsivõimes (Vitali & Giuliani, 2024). Juhtkonna optimism mõjutab ka kasutuselevõtmise tõenäosust (Zemankova, 2019), seega, et vältida väiksemate AEV-de mahajäämust suurematest, on vaja tõsta väiksemate AEV-de teadlikkust AT-dest saadava kasu osas ning näidata, kuidas on kasutuselevõtt võimalik. Käesoleva uuringu andmetel ei ole väljavaated AT-de kasutuselevõtuks head – enamiku AT-de puhul jäävad AEV-de ja ka turu keskmised hinnangud kasutuselevõtmise tõenäosuse kohta alla intervallskaala keskpunkti, mis viitab, et kasutuselevõtt on pigem ebatõenäoline. Kõige tõenäolisemaks peetakse tehisintellekti kasutuselevõttu, millele järgneb andmete standardiseerimine ja täiustatud andmeanalüütika. Ka siinkohal võib põhjus olla tehnoloogia rakendamise mitmekülgsus (IAASB, 2022b; Mikhaylova, 2024; Ruhnke, 2023; Seidenstein jt, 2024).

Kooskõlas varasama kirjandusega (Zemankova, 2019) ilmneb tulemustest, et AT-de kasutuselevõtust oodatava kasu ja kasutuselevõtmise tõenäosuse vahel esineb enamasti mõõduka tugevusega statistiliselt oluline positiivne korrelatsioon – AEV-d, kes hindavad kasu kõrgemalt peavad ka kasutuselevõttu tõenäolisemaks (Tabel 5 osas 4.3). Kõiki

rahvusvahelise kirjanduse põhjal tuvastatud AT-de kasutuselevõtu kasutegurid kehtivad ka Eesti AEV-de puhul (Joonis 12 osas 4.3). Kõige rohkem oodatakse, et AT-de kasutuselevõtt annab töötajatele võimaluse keskenduda kõrgema lisandväärtusega tööle ja et töö efektiivsus tõuseb. Taas on näha erinevust suuremate ja väiksemate AEV-de vahel – suuremad on kasu osas optimistlikumad kui väiksed, sealjuures on nende hulgas rohkem vastajaid, kes näevad AT-de kasutuselevõtust suurt kasu, mis oli uuringus maksimaalne hinnang. Üldiselt saab järeldada, et Eesti AEV-d on AT-de kasutuselevõtust saadava kasu osas optimistlikult meelestatud, sealhulgas suuremad AEV-d rohkem kui väiksemad. Erinevus võib olla seotud väikeste AEV-de poolt tugevamini tunnetatava piiranguga, et klientidel puudub AT-de järele vajadus või et klientide andmed ei ole kujul, mis võimaldaks AT-de kasutuselevõttu. Autori hinnangul võib põhjuseks olla ka vähene teadlikkus sellest, milliseid rakendusvõimalusi AT-del on. Huvitav on, et võrdlemisi kõrgete potentsiaalse kasu hinnangute juures on AT-de kasutuselevõtu tõenäosuse hinnangud siiski võrdlemisi madalad. Järelikult nähakse, et kasutuselevõtu protsessi mõnes sammus esineb nii suuri takistusi, et ka oodatav kasu ei kaalu neid üles – investeeringut, mis ei pruugi olla alati otseselt rahaline, ei peeta tasuvaks (Wiklund & Fallan, 2024).

Tulemuste põhjal saab järeldada, et AT-de kasutuselevõttu takistavad tegurid on Eestis sarnased rahvusvahelise kirjanduse põhjal tuvastatuga (Krieger jt, 2021; Vitali & Giuliani, 2024; Wiklund & Fallan, 2024) – ei olnud ühtegi tegurit, mille keskmised hinnangud oleksid nii madalad, et saaks öelda, et Eesti kontekstis ei ole see takistuseks (Joonis 14 osas 4.3). Kõige suuremaks takistuseks peetakse teadmiste puudust, mistõttu on oluline tegeleda audiitorite koolitamisega, mis tähendab AEV-de jaoks ka investeeringuid töötajate koolitamisse (Seidenstein jt, 2024). Seidenstein jt (2024) kohaselt on üks tõenäolisi koolitamist vajavaid valdkondi teadmised IKT-st. See võib pidurdada AT kasutuselevõtu protsessi nii ideeloome, hindamise kui ka lahenduse väljatöötamise sammus, kuid võib olla takistuseks ka igapäevases kasutamises. Hea lahendus on hea üksnes siis kui see reaalselt kasutusse võetakse. Takistavate tegurite puhul on huvitav, et erinevate tegurite puhul on suurusklasside lõikes hinnangud erinevad (Joonis 15 osas 4.3). Võib öelda, et suured AEV-d ei pea finants- ega ajaliste ressursside puudust nii suureks probleemiks kui väiksed, samuti on suured optimistlikumad investeeringu tasuvuse osas. Koos juhtkonna optimismiga, mida väljendavad hinnangud AT-de potentsiaalsele kasule võib eeldada, et suuremates AEV-des jõuab AT kasutuselevõtmise protsess tõenäolisemalt ressursside allokeerimise etapist edasi (Zemankova, 2019). Samas on näha, et regulatiivseid takistusi on suurte silmis rohkem kui

väikeste puhul. Takistuste osas kõige optimistlikumad hinnangud (takistus puudub) on väikeste ja keskmiste AEV-de juures.

Regulatiivsete takistuste puhul on näha, et mida suuremas suurusklassis on vastajad, seda suurem osakaal vastajatest peab seda oluliseks takistavaks teguriks. Ühelt poolt võib see näidata, et suured AEV-d on oma hinnangutes realistlikumad, võimalik, et väiksemad AEV-d ei ole ka regulatiivsete takistustega niivõrd hästi kursis. Samas võib see aga näidata, et suuremates AEV-des on protsessid regulatsioonide piires juba kaugemale arendatud, nii et edasine arendus võib juba kombata regulatsioonide piire. Väiksemates, kus AT-sid on vähem kasutusele võetud, võib olla veel „vaba vett“ arendamiseks seni kuni regulatiivsed takistused akuutseks muutuvad. Regulatiivsed takistused mõjutavad AT-de kasutuselevõtmist juba esimesel, ideeloome sammul – see tähendab, et AEV-d ei suuda tuvastada AT-le sobilikku kasutusviisi, kuna teatud regulatiivsetel põhjustel paistab AT kasutamine juba eos võimatu.

Tulemustest selgub, et ligikaudu pooltel Eesti AEV-dest, kes moodustavad ligikaudu 40% audiitorteenuste turu käibest, ei ole ühtegi töötajat, kelle tööülesandeks, olgu siis põhitööna või kõrvaltegevusena, oleks innovatsioonide väljatöötamine (Joonis 16 osas 4.3). See võib olla ka üheks põhjuseks, miks AT-de kasutuselevõttu peetakse turul pigem ebatõenäoliseks ja miks kõige suuremaks takistuseks AT-de kasutuselevõtul peetakse teadmiste puudust. Kriegeri jt (2021) protsessiteooria kohaselt on innovatsioonitöötajate roll AT-de kasutuselevõtul oluline ideeloome, hindamise ja lahenduste väljatöötamise etappides (Joonis 3 osas 2.3). AEV-des, kus selline roll on jaotatud, võib olla AT-de kasutuselevõtt toimuda sujuvamalt, eeldusel, et innovatsioonitöötajad on kursis AT-de kasutusvõimalustega. Samuti võivad keskkonnast tulenevad takistused olla siis olla lihtsamini ületatavad. Uuringu tulemused näitavad, et AT-de kasutuselevõtu tõenäosust peetakse suuremaks ettevõtetes, kus on mingilgi moel innovatsioonitöötaja funktsioon täidetud (Joonis 17 osas 4.3). See on ootuspärane, sest teades infovoogu, millega audiitorid peavad igapäevaselt tegelema, on isegi AT-de kohta käiva infoga enda kursis hoidmine keeruline kui selleks eraldi aega ei võeta. Autori hinnangul on seega innovatsioonitöötaja funktsiooni loomine või selleks aja eraldamine töötajate graafikus üks võimalusi AT-de kasutuselevõtu protsessi sujuvamaks ja kiiremaks muuta.

Eelneva põhjal teeb autor AEV-dele, katuseorganisatsioonile ja muudele valdkonnaga seotud osapooltele järgmised soovitused AT-de kasutuselevõtu soodustamiseks Eesti AEV-de seas.

- AT-de kasutuselevõttu tuleks Audiitorkogu ja regulaatorite poolt soosida, sest kasutuselevõtul on mitmeid positiivseid aspekte ja potentsiaali pakkuda leevendust Eesti turul valitsevale audiitorite defitsiidile.
- Teadmiste puudus võib AT-de kasutuselevõtmise protsessi takistada peaaegu igal sammul, mistõttu peaks AT-de alaseid teadmisi kaasama nii audiitorite täiendkoolituse programmidesse (näiteks Audiitorkogu koolitused) kui ka finantsteadmistele suunatud õppekavadesse ülikoolides.
- Teadmisi saaks arendada ka näiteks info lisamisega Audiitorkogu uudiskirjadesse. Näiteks oleks Audiitorkogul võimalik audiitoreid suunata IAASB kodulehele https://www.iaasb.org/focus-areas/technology, kus tuuakse välja ja koondatakse ülevaateid, juhtumiuuringuid ja muid materjale AT-de valdkonna trendide kohta just audiitortegevuse vaatevinklist.
- Audiitorkogu oma uudiskirjades võiks välja tuua ka reaalseid valmis lahendusi, mida AEV-d saaks kohe kasutusele võtta. Soetamine või rentimine on enamasti vähemalt esmainvesteeringu osas odavam kui ise arendamine. Kui selline info oleks kättesaadav uudiskirjades või muus lihtsasti kättesaadavas formaadis, vähendaks see ka ajakulu võimalike lahenduste otsimisele – väiksemates AEV-des on ajapuudus oluline takistus.
- Väiksematel audiitorbüroodel on finantsiliselt keeruline AT-de kasutuselevõtuga tegeleda ning väiksemate AEV-de mahajäämuse ja konkurentsivõime languse vältimiseks on vaja astuda samme. Uuringust paistab, et AEV-d ei ole innovatsiooni valdkonnas teinud koostööd teiste ettevõtetega väljaspool oma võrgustikku. Suunamine koostööle (väiksemad AEV-d omavahel ja / või AEV-d teiste sektorite ettevõtetega) võiks olla üks variant väiksemate AEV-de järele aitamiseks, mis pakub ka leevendust finantsilistele takistustele AEV-de kasutuselevõtul. Näiteks võiks mitmed väiksemad AEV-d ühiselt mõne AT hankida / arendada või muul moel kasutusele võtta. Ühiste tööriistade kasutamisel võib andmete konfidentsiaalsuse säilitamiseks rakendada homomorfse krüpteerimise tehnoloogiat. Koostööd võiks arendada ka IKT sektori osalistega ja miks mitte ka ülikoolide ja innovatsiooni ja AT-sid käsitlevate teadusasutustega. Audiitorkogu võiks võtta ka selles vallas kontaktide võrgustiku loomiseks ühendust asjakohaste erialaliitude ja teiste (katus)organisatsioonidega.
- AT-de kasutuselevõtu kiirendamiseks on oluline, et AEV-s oleks keegi, kelle ülesanne on innovatsioonidega tegeleda. Nii on võimalik suunata innovatsioonitegevusi ning ka AT-de kasutuselevõtu tõenäosus suureneb. Väiksemate AT-de puhul on eraldi inimest

keeruline palgata, aga ka see võib olla abiks, kui see ülesanne on ühel või mõnel töötajal põhitöö kõrvalt või toimiva koostöö puhul ka mitme AEV peale ühiselt.

- Kõikidele AEV-dele tuleks kasuks analüüsida AT-de kasutuselevõtu protsessi mudelit (Joonis 3 osas 2.3) ning tuvastada, milline on peamine samm, kus protsess pooleli jääb. Sealt edasi on võimalik hinnata, milliseid samme astuda, et protsess lõpuni jõuaks. Kui protsess ei algagi, sest puuduvad teadmised AT-de ja nende potentsiaalsete kasutusvaldkondade kohta, siis on selle tõkke ületamine kõige olulisem, sest seejärel võib avaneda mitmeid uusi võimalusi arenemiseks.

Kokkuvõte

Arenevate tehnoloogiate (AT) nagu tehisintellekt, robootiline protsesside automatiseerimine, täiustatud andmeanalüütika kasutus ja muu rakendamine audiitortegevuse valdkonnas on viimasel ajal kiiresti arenenud ning potentsiaalselt muudab valdkonda tulevikus märkimisväärselt (Seidenstein jt, 2024). Võimalikud on muutused audiitorettevõtjates (AEV) vajaliku tööjõu struktuuri, teadmiste ja oskuste ning audiitorteenuste osutamise protsessi osas (Palekhova & Kramarenko, 2020; Seidenstein jt, 2024; Vitali & Giuliani, 2024). Audiitorteenuste turg väga reguleeritud valdkonnana erineb üldisest teenuste turust, mis mõjutab ka AT-de kasutuselevõtu protsessi (Krieger jt, 2021) ja muudab spetsiifiliselt selle valdkonna uurimise huvitavaks. Eestis valitseb hetkel audiitorite defitsiit ning see püsib tõenäoliselt ka järgmistel aastatel (Õun, 2024), mistõttu on asjakohane uurida võimalusi olemasolevate AEV-de efektiivsuse tõstmiseks – AT-de kasutuselevõtt pakub selleks võimaluse (Krieger jt, 2021; Ruhnke, 2023; Zemankova, 2019; Vitali & Giuliani, 2024). AT-de kasutuselevõttu Eesti AEV-des ei ole varem palju uuritud, peamine on Mikhaylova (2024) magistritöö, mis käsitles tehisintellekti kasutuselevõtu väljavaateid.

Käesoleva uurimuse eesmärgiks on anda hinnang innovatsioonile Eesti audiitorettevõtjates läbi arenevate tehnoloogiate kasutuselevõtu ja anda soovitusi selle protsessi edendamiseks. Selle saavutamiseks on täidetud järgmised uurimisülesanded. Uurimisobjektina on määratletud AEV-d, on selgitatud nende ettevõtete eripärasid ja tegevusi ning tuvastatud valdkonnad, milles on AEV-des võimalik AT-sid kasutusele võtta. Koostatud on ülevaade teaduskirjandusest, mis käsitleb innovatsiooni ja AT-de kasutelevõtu protsessi ning seda mõjutavaid tegureid AEV-des. Oslo käsiraamatu (OECD/Eurostat, 2018) juhiste kohaselt ning kirjanduse ülevaatest saadud info põhjal on välja töötatud uurimuse metoodika ning koostatud küsimustik. Uurimuse teoreetilise raamistikuna käsitletakse Oslo käsiraamatus (OECD/Eurostat, 2018) toodud põhimõtteid ettevõtete innovaatilisuse põhinäitajate kohta ning Krieger jt (2021) poolt kujundatud protsessiteooriat tehnoloogiate

kasutuselevõtuks AEV-des. Viidi läbi valikvastustega küsitlus Eesti AEV-de seas (kokku 116), millele saadi 34 vastust ning analüüsiti tulemusi peamiselt kvantitatiivse kirjeldava statistika abil tarkvaras R ja RStudio. Andmete analüüsimise metoodika hõlmab kaalude kasutamist, et tulemusi oleks võimalik üldistada Eesti AEV-de kogumile ja eraldi ka audiitorteenuste turule. Viimane üldistus annab aimu sellest, kuidas tunnetab AT-de kasutuselevõttu Eesti audiitorteenuste klientide kogum. Tulemuste usaldusväärsuse hindamiseks on arvutatud veapiirid usaldusnivool 0,95 ning tulemuste väljendamiseks on koostatud joonised. Lõpuks on arutletud tulemuste üle ja kujundatud soovitused AT-de kasutuselevõtu protsessi edendamiseks Eesti AEV-des.

Uuringu tulemused näitavad, et enamikust Eesti ettevõtetest erinevad AEV-de innovaatilisuse põhinäitajad oluliselt suurema üksnes protsessiinnovatsiooniga ettevõtete osakaalu poolest, mis on tõenäoliselt seotud valdkonna kõrge reguleeritusega. Protsessiinnovatsioone on läbi viinud enamik AEV-sid ning peamiselt on need uuendused välja töötanud AEV ise. Innovatsioone ei ole rohkem ellu viidud peamiselt seoses vajaduse või ressursside puudumisega. AT-de kasutuselevõtt ei ole AEV-de hulgas veel laialt levinud, eriti väikeste AEV-de seas. Üldiselt näitavad tulemused AT-de kasutuselevõtu protsessi mõjutavate tegurite osas mitmeid erinevusi suuremate ja väiksemate AEV-de vahel. Üldiselt on suuremad AT-de osas optimistlikumalt meelestatud. Peamine takistus AT-de kasutuselevõtule on teadmiste puudus, mida võib võimendada asjaolu, et ligikaudu pooltel AEV-del ei ole ühtegi töötajat, kellel oleks konkreetne tööülesanne tegeleda innovatsiooniga.

Tulemustest lähtuvalt on uurimuse lõpuosas antud mitmeid soovitusi Eesti audiitorteenuste turu osalistele AT-de kasutuselevõtu protsessi edendamiseks. Peamised soovitused on seotud AT-de alaste teadmiste tõstmisega ning turuosaliste omavahelise ja väliste osapooltega koostöö edendamisega, eriti väiksemate AEV-de puhul, et vähendada ressursside ja IKT alaste teadmiste puudusest tingitud takistusi. Väikeste AEV-de järele aitamine on oluline, et vähendada nende mahajäämust AT-de kasutuselevõtul, kuna see võib tulevikus vähendada nende konkurentsivõimet.

Uuringu disainist lähtuvalt esineb piiranguid, mis tekitavad edasisi uurimisvõimalusi. Kvantitatiivne lähenemine ei võimalda tuvastada uusi tehnoloogiaid, mida varasemas kirjanduses ja allikates ei ole käsitletud – seda oleks võimalik täiendada intervjuudel või avatud vastustega küsimustikul põhinevate uuringutega. Vastuseid oli 34, mis on hea tulemus arvestades Eesti AEV-de koguarvu (116), kuid mis on paljude statistiliste analüüside jaoks siiski ebapiisav, mistõttu tuli uurimuses piirduda peamiselt kirjeldava statistikaga ning leppida suurte veapiiridega. Uuringu ajakava tõttu jäi ka vastamise periood väga lühikeseks –

pikem aeg oleks võinud anda mõnevõrra rohkem vastuseid. Selle piirangu mõju minimeerimiseks võttis autod uuritavatega korduvalt ühendust, sealhulgas telefoni teel. Tulevikus võiks sarnast uuringut läbi viia suuremas regioonis, sealhulgas mitmeid riike hõlmava uuringuna. Sellisel juhul oleks võimalik saada ka piisavalt vastuseid statistiliste mudelite koostamiseks ning avaneks võimalus uurida, kas AEV-de innovaatilisuse põhinäitajates ja AT-de kasutuselevõtu protsessis esineb ka regionaalseid erinevusi. Sarnase metoodikaga uuringud võimaldaks tekitada ka aegridade analüüsi ajas toimuvate muutuste kohta arenevate tehnoloogiate kasutuselevõtu valdkonnas. Asjaolu, et autor on ise vandeaudiitor ning uurimisobjektide konkurent, piiras küsimuste detailsusastet ja oli ka oluline põhjus kvantitatiivse lähenemise kasutamiseks. Edaspidised uuringud võiksid keskenduda ka sellele, millised konkreetsed AT-del põhinevad lahendused on Eesti AEV-des kasutusel.

Viidatud allikad

1. Anica-Popa, I. F., Vrîncianu, M., Anica-Popa, L. E., Cișmașu, I. D., & Tudor, C. G. (2024). Framework for Integrating Generative AI in Developing Competencies for Accounting and Audit Professionals. *Electronics (Switzerland)*, *13*(13). https://doi.org/10.3390/electronics13132621
2. Appelbaum, D. A., Kogan, A., & Vasarhelyi, M. A. (2018). Analytical procedures in external auditing: A comprehensive literature survey and framework for external audit analytics. *Journal of Accounting Literature*, *40*, 83–101. https://doi.org/10.1016/j.acclit.2018.01.001
3. Audiitorkogu. (2024a). Audiitorkogu üldkoosolek. M.-M. Arengu (Toim), *Audiitorkogu üldkoosolek*.
4. Audiitorkogu. (2024b, november 28). *Audiitorettevõtjate nimekiri*. https://www.audiitorkogu.ee/est/audiitorettevotjad
5. AJN. (2024). *AJN aastaraamat 2024*. https://media.voog.com/0000/0041/6835/files/AJN%20aastaraamat%202024.pdf
6. Audiitortegevuse seadus (2023). https://www.riigiteataja.ee/akt/117032023006
7. Audiitortegevuse seadus (2025). https://www.riigiteataja.ee/akt/107012025013
8. Barras, R. (1986). Towards a theory of innovation in services. *Research Policy*, *15*(4), 161–173. https://doi.org/10.1016/0048-7333(86)90012-0
9. Barr-Pulliam, D., Brown-Liburd, H. L., & Munoko, I. (2022). The effects of person-specific, task, and environmental factors on digital transformation and innovation in auditing: A review of the literature. *Journal of International Financial Management and Accounting*, *33*(2), 337–374. https://doi.org/10.1111/jifm.12148
10. Budarov, R. (2017). *Auditi efektiivsuse suurendamine muutuval audiitorteenuste turul*.
11. Davidson, A. C., & Hinkley, D. V. (1997). *Bootstrap Methods and Their Application* (10. tr). Cambridge University Press.
12. Statistikaamet. (2023). *Ettevõtete innovatsiooniuuring. Aastad 2020-2022. Käsiraamat*. https://stat.ee/sites/default/files/2023-04/Innovatsiooniuuring_kasiraamat_1293_2023.pdf
13. Fedyk, A., Hodson, J., Khimich, N., & Fedyk, T. (2022). Is artificial intelligence improving the audit process? *Review of Accounting Studies*, *27*(3), 938–985. https://doi.org/10.1007/s11142-022-09697-x

14. Franco, C., Little, R. J. A., Louis, T. A., & Slud, E. V. (2019). Comparative Study of Confidence Intervals for Proportions in Complex Sample Surveys. *Journal of Survey Statistics and Methodology*, *7*(3), 334–364. https://doi.org/10.1093/jssam/smy019
15. Goto, M. (2023). Anticipatory innovation of professional services: The case of auditing and artificial intelligence. *Research Policy*, *52*(8). https://doi.org/10.1016/j.respol.2023.104828
16. IAASB. (2020). *Technology Disruption in Audit and Assurance Agenda Item 1-A*.
17. IAASB. (2021, oktoober 27). *IAASB Digital Technology Market Scan: Data Standardization*. https://www.iaasb.org/news-events/2021-10/iaasb-digital-technology-market-scan-data-standardization
18. IAASB. (2022a). *IAASB Digital Technology Market Scan: API Access*. https://www.iaasb.org/news-events/2022-01/iaasb-digital-technology-market-scan-api-access
19. IAASB. (2022b). *IAASB Digital Technology Market Scan: Artificial Intelligence - A Primer*. https://www.iaasb.org/news-events/2022-03/iaasb-digital-technology-market-scan-artificial-intelligence-primer
20. IAASB. (2022c). *IAASB Digital Technology Market Scan: Homomorphic Encryption*. https://www.iaasb.org/news-events/2022-10/iaasb-digital-technology-market-scan-homomorphic-encryption
21. IAASB. (2022d). *IAASB Digital Technology Market Scan: Natural Language Processing*. https://www.iaasb.org/news-events/2022-06/iaasb-digital-technology-market-scan-natural-language-processing
22. IAASB. (2022e). *ISA (EE) 200 rahvusvaheline auditeerimise standard (Eesti) 200: Sõltumatu audiitori üldised eesmärgid ja auditi läbiviimine kooskõlas rahvusvaheliste auditeerimise standarditega (Eesti)*.
23. IAASB. (2022f). *Rahvusvaheline auditeerimise standard (Eesti) 500: Auditi tõendusmaterjal*.
24. IAASB. (2023a). *IAASB Digital Technology Market Scan: Internet of Things Technologies*. https://www.iaasb.org/news-events/2023-05/iaasb-digital-technology-market-scan-internet-things-technologies
25. IAASB. (2023b). *IAASB Digital Technology Market Scan: Robotic Process Automation*. https://www.iaasb.org/news-events/2023-01/iaasb-digital-technology-market-scan-robotic-process-automation
26. IAASB. (2023c). *Rahvusvaheline ülevaatamise teenuse standard*.

27. IAASB. (2023d, märts 6). *IAASB Digital Technology Market Scan: Digital Assets*. https://www.iaasb.org/news-events/2023-03/iaasb-digital-technology-market-scan-digital-assets
28. Krieger, F., Drews, P., & Velte, P. (2021). Explaining the (non-) adoption of advanced data analytics in auditing: A process theory. *International Journal of Accounting Information Systems, 41*. https://doi.org/10.1016/j.accinf.2021.100511
29. Mikhaylova, V. (2024). *Tehisintellekti kasutamise väljavaated Eesti audiitorettevõtetes*. Tallinna Tehnikaülikool.
30. Miles, I. (2005). Knowledge intensive business services: Prospects and policies. *Foresight, 7*(6), 39–63. https://doi.org/10.1108/14636680510630939
31. OECD/Eurostat. (2018). *Oslo Manual 2018: Guidelines for Collecting, Reporting and Using Data on Innovation, 4th Edition* (4th Edition). The Measurement of Scientific, Technological and Innovation Activities, OECD Publishing, Paris/Eurostat, Luxembourg. https://doi.org/10.1787/9789264304604-en
32. OpenAI. (2025, jaanuar 13). *ChatGPT 4o*. ChatGPT 4o. https://chatgpt.com/
33. Palekhova, V., & Kramarenko, O. (2020). The impact of technological innovations on employment in the financial sector. *Technology audit and production reserves, 6*(4(56)), 45–49. https://doi.org/10.15587/2706-5448.2020.220290
34. Rahandusministeerium. (2024, november 28). *Audiitorettevõtjate nimekiri*. https://www.audiitortegevus.ee/atr/web/register/audiitorettevotja-otsing?
35. Rogers, E. M. . (2003). *Diffusion of innovations, 5th edition* (5., illustreeritud tr, Kd 576). Simon and Schuster.
36. Ruhnke, K. (2023). Empirical research frameworks in a changing world: The case of audit data analytics. *Journal of International Accounting, Auditing and Taxation*, *51*. https://doi.org/10.1016/j.intaccaudtax.2023.100545
37. Santos-Vijande, M. L., González-Mieres, C., & López-Sánchez, J. Á. (2013). An assessment of innovativeness in KIBS: Implications on KIBS' co-creation culture, innovation capability, and performance. *Journal of Business and Industrial Marketing, 28*(2), 86–102. https://doi.org/10.1108/08858621311295236
38. Sauga, A. (2020). *Statistika õpik majanduseriala üliõpilastele* (2. väljaanne). Taltech Kirjastus.
39. Seidenstein, T., Marten, K. U., Donaldson, G., Föhr, T. L., Reichelt, V., & Jakoby, L. B. (2024). Innovation in Audit and Assurance: A Global Study of Disruptive Technologies.

Journal of Emerging Technologies in Accounting, *21*(1), 129–146. https://doi.org/10.2308/JETA-2022-026

40. Statistikaamet. (2024, detsember 8). *Innovaatilised ettevõtted 2022 Statistikaamet*. https://andmed.stat.ee/et/stat/majandus__teadus-tehnoloogia-innovatsioon__innovatiivne-tegevus/TDI1601/table/tableViewLayout2
41. *Statistikaamet*. (2024a, november 11). https://www.stat.ee/et/avasta-statistikat/metoodika-ja-kvaliteet/esms-metaandmed/21702#3-Statistika-esitus-2
42. Statistikaamet. (2024b, detsember 6). *Statistikaamet IA02: TARBIJAHINNAINDEKS*. https://andmed.stat.ee/et/stat/majandus__hinnad/IA02/table/tableViewLayout2
43. Zemankova, A. (2019). Artificial Intelligence in Audit and Accounting: Development, Current Trends, Opportunities and Threats-Literature Review. *Proceedings - 2019 3rd International Conference on Control, Artificial Intelligence, Robotics and Optimization, ICCAIRO 2019*, 148–154. https://doi.org/10.1109/ICCAIRO47923.2019.00031
44. Tiberius, V., & Hirth, S. (2019). Impacts of digitization on auditing: A Delphi study for Germany. *Journal of International Accounting, Auditing and Taxation*, *37*. https://doi.org/10.1016/j.intaccaudtax.2019.100288
45. Varadarajan, R. (2024). Mitigating the definitional quagmire in innovation research: An inclusive definition of innovation as a template for defining various types of innovations uniformly. *Journal of Product Innovation Management*, *41*(4), 703–723. https://doi.org/10.1111/jpim.12729
46. Vărzaru, A. A., & Bocean, C. G. (2024). Digital Transformation and Innovation: The Influence of Digital Technologies on Turnover from Innovation Activities and Types of Innovation. *Systems*, *12*(9), 359. https://doi.org/10.3390/systems12090359
47. Vitali, S., & Giuliani, M. (2024). Emerging digital technologies and auditing firms: Opportunities and challenges. *International Journal of Accounting Information Systems*, *53*. https://doi.org/10.1016/j.accinf.2024.100676
48. Wiklund, T. E., & Fallan, E. (2024). Adoption of RPA Technology in Micro and Small Audit Entities. *Journal of Emerging Technologies in Accounting*, 1–21. https://doi.org/10.2308/jeta-2023-036
49. Wilson, E. B. (1927). Probable Inference, the Law of Succession, and Statistical Inference. *Source: Journal of the American Statistical Association* (Kd 22, Number 158).
50. Õun, H. (2024). *Audiitortegevuse minevik, olevik ja tulevik*. Tallinna Tehnikaülikool.

LISA A
Küsimustik

Sissejuhatus

Lugupeetud audiitorettevõtja esindaja,

Küsimustikuga kogutakse andmeid Robert Pirk'i magistritöö jaoks, mille eesmärk on uurida Eesti audiitorettevõtjate innovaatilisust perioodil 01.07.2022-30.06.2024 ja arenevate tehnoloogiate kasutuselevõttu ja nende tehnoloogiate kasutuselevõtu väljavaateid Eesti audiitorettevõtjate seas.

Küsimustik on mõeldud täitmiseks eelkõige ettevõtte juhile või juhtkonna liikmele, kes tegutseb audiitorettevõtja arendamise valdkonnas ja on kursis ettevõttes läbi viidud ja planeeritavate uuendustega. Küsimustik on anonüümne ning vastuseid kasutatakse uurimistöös üksnes üldistatuna kõikide vastajate peale. Küsimustik on suunatud audiitorettevõtjale, mistõttu palun kindlaks teha, et Teie ettevõtte kohta esitatakse üks vastus.

Küsimustikule vastamine võtab hinnanguliselt 10-20 minutit.

Küsimustiku esimene osa keskendub arenevate tehnoloogiate kasutuselevõtule audiitorettevõtjates. Teine ja kolmas osa on üldisemad ning uurivad ettevõtte innovaatilisust. Selle osa küsimused põhinevad Eurostati läbiviidaval Community Innovation Survey (CIS) küsimustel. Need küsimused on välja töötatud selliselt, et küsimustikule saaks vastata kõikide valdkondade ettevõtted. Seetõttu võib esineda vastusevariante, mis ei ole audiitorettevõtja seisukohast relevantsed.

Palun küsimustikule vastata hiljemalt 06.12.2024.

Lisateavet küsimustiku täitmise ja esitamise kohta saab telefonil +372 5692 7469 või e-posti aadressil ro_bert_pirk@hotmail.com.

Disclaimer: Robert Pirk on praktiseeriv vandeaudiitor.

I osa - Arenevate tehnoloogiate kasutuselevõtt audiitortegevuses

Globaalsel tasandil on audiitortegevuses toimumas muutused seoses erinevate arenevate tehnoloogiate kasutuselevõtuga, millel võib olla oluline mõju kogu valdkonnale. Käesoleva uuringu käigus on eesmärgiks uurida, millises seisus on arenevate tehnoloogiate kasutuselevõtt Eestis tegutsevates audiitorettevõtjates.

1. Kuivõrd suurt potentsiaalset kasu näete oma ettevõttes järgmiste arenevate tehnoloogiate kasutuselevõtul? Vastused intervallskaalal
   0 – kasu puudub
   1 – väike kasu
   2 – mõõdukas kasu
   3 – suur kasu
   a. Tehisintellekt (Artificial Intelligence - AI): *AI hõlmab erinevaid tehnoloogiaid, sh masinõpe ja süvaõpe, mida kasutatakse auditites näiteks andmeanalüüsi ja otsuste tegemise parandamiseks.*
   b. Loodusliku keele töötlemine (Natural Language Processing - NLP): *AI alamvaldkond, mis võimaldab arvutitel mõista ja tõlgendada inimkeelt, hõlbustades näiteks struktureerimata andmete, nagu e-kirjad ja dokumendid, analüüsi auditites.*
   c. Robootiline protsesside automatiseerimine (Robotic Process Automation - RPA): *RPA automatiseerib korduvaid ülesandeid.*
   d. Asjade internet (Internet of Things - IoT): *IoT hõlmab omavahel ühendatud seadmeid, mis koguvad ja vahetavad andmeid ning mida saab näiteks kasutada auditites reaalajas jälgimiseks ja äriprotsesside ning kontrollide hindamiseks.*
   e. Homomorfne krüpteerimine: *See krüpteerimismeetod võimaldab sooritada arvutusi krüpteeritud andmetel neid dekrüpteerimata, suurendades andmete privaatsust ja turvalisust auditite ajal.*
   f. Rakendusliidesed (Application Programming Interfaces – API)
   g. Andmete standardiseerimise tarkvara
   h. Täiustatud andmeanalüütika (Advanced Data Anlytics – ADA): Kirjeldav analüütika, diagnostiline analüütika, ennustav analüütika, soovituslik analüütika, sh suurandmete analüütika, kogu populatsiooni analüüsimine jm.
   i. Plokiahela tehnoloogia (Blockchain)
   j. Protsessikaevandamine (Process Mining) - andmeanalüüsi meetod, mille abil uuritakse ja analüüsitakse organisatsiooni protsesse, kasutades jälgi, mis on salvestatud IT-süsteemides. Protsessikaevanduse eesmärk on mõista, kuidas protsessid tegelikult toimivad, võrreldes neid plaanitud protsessidega, ning tuvastada võimalusi nende parendamiseks.
   k. Optiline märgituvastus (Optical Character Recognition) ja / või arvutinägemine (Computer Vision) – tehnoloogiad, mis võimaldavad tuvastada teksti. Võimalik dokumente digiteerida.

   Vastused

a. Tehisintellekt (Artificial Intelligence – AI)
b. Loodusliku keele töötlemine (Natural Language Processing - NLP)
c. Robootiline protsesside automatiseerimine (Robotic Process Automation - RPA)
d. Asjade internet (Internet of Things – IoT)
e. Homomorfne krüpteerimine.
f. Rakendusliidesed (Application Programming Interfaces – API)
g. Andmete standardiseerimise tarkvara
h. Täiustatud andmeanalüütika (Advanced Data Anlytics – ADA)
i. Plokiahela tehnoloogia (Blockchain)
j. Protsessikaevandamine (Process Mining)
k. Optiline märgituvastus (Optical Character Recognition) ja / või arvutinägemine (Computer Vision)

2. Millisel moel järgnevatest ja mil määral avaldub või tõenäoliselt avalduks ühe või mitme areneva tehnoloogia kasutuselevõtust saadav kasu? *Vastused intervallskaalal*

0 – kasu puudub

1 – väike kasu

2 – mõõdukas kasu

3 – suur kasu

a. Tõuseb töövõtu kvaliteet (Ruhnke, 2023)
b. Väheneb töövõtule kuluv aeg
c. Tõuseb töö efektiivsus (Zemankova, 2019), (Ruhnke, 2023), (Vitali & Giuliani, 2024)
d. Väheneb vigade hulk töövõttudes (Zemankova, 2019)
e. Võimaldab töötajatel panustada rohkem aega keerulisematele ja kõrgema lisandväärtusega tegevustele (Zemankova, 2019)
f. Tõuseb meie ettevõtte konkurentsivõime audiitorteenuste turul (Ruhnke, 2023)
g. Võimaldab paremini täita töövõttudega seotud standardite ja muude regulatsioonide nõudeid

3. Kui suur mõju on Teie hinnangul Teie ettevõttes arenevate tehnoloogiate kasutuselevõtu kontekstis järgmistel takistavatel teguritel? *Vastused intervallskaalal*

0 – ei ole takistuseks

1 – väike takistav mõju

2 – mõõdukas takistav mõju

3 – suur takistav mõju

a. **Regulatiivsed takistused**, sh selliste tehnoloogiate kasutamist reguleerivate auditeerimise standardite puudumine ning teadmatus, kas järelevalveorganid peavad uudseid lahendusi võrdväärseks traditsiooniliste protseduuridega (Krieger jt, 2021)
b. **Ei ole piisavalt ajalist ressurssi**, et töötada välja arenevatel tehnoloogiatel põhinevaid lahendusi, mida saaks kasutada audiitorteenuste osutamisel, või neil

pole aega, et leida välisest allikast pärit lahendustele spetsiifilist rakendust audiitorteenuste kontekstis (Krieger jt, 2021)

c. **Ei ole piisavaid teadmisi**, et leida olemasolevatele tehnoloogiatele asjakohane kasutusvõimalus audiitorteenuste osutamise käigus, ehk ei suudeta tehnoloogiatele leida konkreetseid kasutusvõimalusi. Tehniliste teadmistega töötajatel ei pruugi olla piisavat arusaamist protseduuridest ja sellest, millised eesmärgid töövõtul on, samas kutsealaste teadmistega töötajatel ei pruugi olla piisavat arusaamist, kuidas uus tehnoloogia toimib, et leida sellele rakendust protseduuride läbiviimisel. (Krieger jt, 2021)
d. **Ei ole piisavalt finantsilisi ressursse**, et ise arendada arenevatel tehnoloogiatel põhinevaid lahendusi. (Krieger jt, 2021)
e. **Ei ole piisavalt finantsilisi ressursse**, et võtta kasutusele mujal välja töötatud arenevatel tehnoloogiatel põhinevaid lahendusi (Krieger jt, 2021)
f. **Ei õnnestu värvata piisavate teadmiste ja oskustega tööjõudu** (Krieger jt, 2021)
g. **Puudub kliendist tulenev vajadus tehnoloogiate kasutuselevõtuks** (näiteks: üksnes väiksed kliendid, kelle puhul arenevate tehnoloogiate kasutamine ei oleks tasuv) (Wiklund & Fallan, 2024)
h. **Investeeringud arenevatesse tehnoloogiatesse ei ole tasuvad** (Wiklund & Fallan, 2024)
i. **Kliendid ei suuda anda andmeid sellisel kujul, mis on vajalik arenevatel tehnoloogiatel põhinevate tööriistade kasutamiseks** (Vitali & Giuliani, 2024)

Vastused

a. **Regulatiivsed takistused**
b. **Ei ole piisavalt ajalist ressurssi**
c. **Ei ole piisavaid teadmisi**
d. **Ei ole piisavalt finantsilisi ressursse**, et ise arendada arenevatel tehnoloogiatel põhinevaid lahendusi.
e. **Ei ole piisavalt finantsilisi ressursse**, et võtta kasutusele mujal välja töötatud arenevatel tehnoloogiatel põhinevaid lahendusi
f. **Ei õnnestu värvata piisavate teadmiste ja oskustega tööjõudu**
g. **Puudub kliendist tulenev vajadus tehnoloogiate kasutuselevõtuks**
h. **Investeeringud arenevatesse tehnoloogiatesse ei ole tasuvad**
i. **Kliendid ei suuda anda andmeid sellisel kujul, mis on vajalik arenevatel tehnoloogiatel põhinevate tööriistade kasutamiseks**

4. Millised arenevad tehnoloogiad olid Teie ettevõttes audiitorteenuste osutamisel 30.06.2024 seisuga kasutusele võetud? Valige kõik sobivad vastusevariandid
    a. Tehisintellekt (Artificial Intelligence – AI)
    b. Loodusliku keele töötlemine (Natural Language Processing - NLP)
    c. Robootiline protsesside automatiseerimine (Robotic Process Automation - RPA)
    d. Asjade internet (Internet of Things – IoT)
    e. Homomorfne krüpteerimine.
    f. Rakendusliidesed (Application Programming Interfaces – API)
    g. Andmete standardiseerimine
    h. Täiustatud andmeanalüütika (Advanced Data Anlytics – ADA)

i. Plokiahela tehnoloogia (Blockchain)
j. Protsessikaevandamine (Process Mining)
k. Optiline märgituvastus (Optical Character Recognition) ja / või arvutinägemine (Computer Vision)
l. Mitte ükski neist

5. Kes on 30.06.2024 seisuga kasutusele võetud arenevat tehnoloogiat rakendava(te) tööriista(de) väljatöötaja? Valige kõik sobivad.
   a. Ettevõte ise
   b. Samasse võrgustikku kuuluv teine ettevõte või ettevõte koostöös võrgustikku kuuluva teise ettevõttega
   c. Ettevõte koostöös teise ettevõtte või asutusega, kes ei kuulu samasse võrgustikku
   d. Teine ettevõte või asutus
   e. Ettevõttes ei ole 30.06.2024 seisuga nimetatud arenevaid tehnoloogiaid kasutusele võetud
6. Kui tõenäoliseks peate, et Teie ettevõte võtab järgmise nelja aasta jooksul kasutusele nimetatud arenevaid tehnoloogiaid audiitorteenuste osutamisel? *Vastused intervallskaalal*

   1 – väga ebatõenäoline

   2 – ebatõenäoline

   3 – ei oska öelda

   4 – tõenäoline

   5 – väga tõenäoline

   a. Tehisintellekt (Artificial Intelligence – AI)
   b. Loodusliku keele töötlemine (Natural Language Processing - NLP)
   c. Robootiline protsesside automatiseerimine (Robotic Process Automation - RPA)
   d. Asjade internet (Internet of Things – IoT)
   e. Homomorfne krüpteerimine.
   f. Rakendusliidesed (Application Programming Interfaces – API)
   g. Andmete standardiseerimise tarkvara
   h. Täiustatud andmeanalüütika (Advanced Data Anlytics – ADA)
   i. Plokiahela tehnoloogia (Blockchain)
   j. Protsessikaevandamine (Process Mining)
   k. Optiline märgituvastus (Optical Character Recognition) ja / või arvutinägemine (Computer Vision)
7. Kui tõenäoliselt pärinevad Teie ettevõttes tulevikus kasutusse võetavad arenevad tehnoloogiad järgmistest allikatest? *Vastused intervallskaalal*

   1 – väga ebatõenäoline

   2 – ebatõenäoline

   3 – ei oska öelda

   4 – tõenäoline

   5 – väga tõenäoline

a. Ettevõte ise
b. Samasse võrgustikku kuuluv teine ettevõte või Ettevõte koostöös võrgustikku kuuluva teise ettevõttega
c. Ettevõte koostöös teise ettevõtte või asutusega, kes ei kuulu samasse võrgustikku
d. Teine ettevõte või asutus

II osa - Teenuseuuendused perioodil 01.07.2022-30.06.2024

Kõikidele selles osas esitatud küsimustele vastates pidage silmas kaheaastast perioodi 01.07.2022-30.06.2024

Teenuseuuendus on uus või täiustatud teenus, mis erineb oluliselt ettevõte poolt varasemalt osutatud teenustest ja mida on turule tutvustatud.

Näited:

- Turu jaoks uue teenuse pakkumine (näiteks ESG aruannete audiitorkontrolli,)
- Ettevõtte jaoks uue kutseteenuse pakkumine (näiteks pakendiaudit, kui ettevõte pole varem pakendiauditi teenust osutanud)
- Varem pakutud kutseteenuse oluline muutus

Küsimused baseeruvad CISil ja on välja töötatud kõikides sektorites töötavate ettevõtete jaoks.

8. Kas Teie ettevõte tõi perioodil 01.07.2022–30.06.2024 turule või esitles sihtrühmale **turu jaoks** uusi teenuseid?
    a. Jah
    b. Ei
    c. Ei oska öelda
9. Kas Teie ettevõte tõi perioodil 01.07.2022–30.06.2024 turule või esitles sihtrühmale **ettevõtte jaoks** uusi teenuseid?
    a. Jah
    b. Ei
10. Kas Teie ettevõte tõi perioodil 01.07.2022–30.06.2024 turule või esitles sihtrühmale **oluliselt täiustatud** teenuseid?
    a. Jah
    b. Ei
11. Kas mõni uus või täiustatud teenus perioodil 01.07.2022–30.06.2024 oli selline, mida enne teid ei pakkunud Eestis ükski Teie konkurent?
    a. Jah
    b. Ei
    c. Ei oska öelda
12. Kas mõni uus või täiustatud teenus perioodil 01.07.2022–30.06.2024 oli identne või väga sarnane teenustega, mida Teie konkurendid Eestis juba pakkusid

a. Jah
b. Ei
c. Ei oska öelda

13. Kes oli nende teenuseuuenduste väljatöötaja? Samasse võrgustikku kuuluvad teised ettevõtted liigatada ka „teisteks ettevõteteks või asutusteks“. Valige kõik sobivad vastusevariandid.
    a. ettevõte ise
    b. ettevõte koostöös teise ettevõtte või asutusega
    c. ettevõte kohandas teises ettevõttes või asutuses väljatöötatud teenuseid enda või kliendi soovidele
    d. teine ettevõte või asutus
    e. ettevõttel puudusid teenuseuuendused perioodil 01.07.2022-30.06.2024

III osa - Protsessiuuendused perioodil 01.07.2022-30.06.2024

Kõikidele selles osas esitatud küsimustele vastates pidage silmas kaheaastast perioodi 01.07.2022-30.06.2024

Protsessiuuendus on uus või täiustatud äriprotsess ühe või mitme ärifunktsiooni jaoks, mis erineb oluliselt ettevõtte poolt varem rakendatud äriprotsessidest ja mis on ettevõtja poolt kasutusse võetud .

Näited:

- Olulised muutused kutseteenuse osutamise protsessides, näiteks uue tarkvara kasutuselevõtt, olulised arendused tööpaberites, protseduuride automatiseerimine.
- Olulised muutused kutseteenuste turunduse ja müügiga seotud protsessides.
- Uute info- ja kommunikatsioonitehnoloogia süsteemide, tarkvarade kasutuselevõtt
- Olulised muudatused kutseteenuse juhtimise ja haldamise protsessides

Küsimused baseeruvad CISil ja on välja töötatud kõikides sektorites töötavate ettevõtete jaoks. Kõik vastusevariandid ei pruugi seega paista audiitorettevõtja seisukohas asjakohased.

14. Millistes järgmistes valdkondades võttis Teie ettevõte perioodil 01.07.2022–30.06.2024 kasutusele uusi protsesse või arendas protsesse edasi nii, et need erinesid oluliselt Teie varasematest protsessidest? Valida kõik asjakohased vastusevariandid.
    a. teenuste osutamine (s.t põhiprotsess, sh arendus) (ettevõtte poolt kliendile teenuse osutamine)
    b. logistika-, tarne- või jaotusmeetodid (seotud kaupade liikumisega)
    c. info- ja kommunikatsioonitehnoloogia (arvutid, tarkvara, andmebaasid jm)
    d. raamatupidamine või ostuprotsess (st audiitorettevõtja enda protsess, mitte klientide protsessid)

e. äriprotsesside korraldus või välissuhtluse korraldus (external relations) (äriprotsesside ümberkorraldamine jms)
f. töökorraldus, inimressursside juhtimine (juhtimisstruktuuri muutused, kohustuste muutumine, uued allüksused, meeskonnatöö jms)
g. turundus: reklaam, teenuse raames väljastatud dokumentide ja materjalide kujundus, hinnakujundus, müügijärgsed teenused, jne
h. mitte üheski neist

15. Kes oli nende protsessiuuenduste väljatöötaja? Samasse võrgustikku kuuluvad teised ettevõtted liigatada ka „teisteks ettevõteteks või asutusteks“. Vali kõik sobivad vastusevariandid.
    a. ettevõte ise
    b. ettevõte koostöös teise ettevõtte või asutusega
    c. ettevõte kohandas teises ettevõttes või asutuses väljatöötatud teenuseid enda või kliendi soovidele
    d. teine ettevõte või asutus
    e. ettevõttel puudusid protsessiuuendused perioodil 01.07.2022-30.06.2024
16. Milline järgnevatest väidetest iseloomustab Teie ettevõtet kõige paremini? (Goto, 2023; Krieger jt, 2021)
    a. ettevõttes on uuenduste väljatöötamisele, leidmisele ja / või kasutuselevõtmisele spetsialiseeritud töörühm
    b. ettevõttes on uuenduste väljatöötamisele, leidmisele ja / või kasutuselevõtmisele spetsialiseerunud üks töötaja
    c. ettevõttes on üks või mitu töötajat, kes põhitöö kõrvalt tegelevad uuenduste väljatöötamise, leidmise ja / või kasutuselevõtmisega ja nende selline tegevus on juhtkonna poolt suunatud.
    d. ettevõttes on üks või mitu töötajat, kes põhitöö kõrvalt tegelevad uuenduste väljatöötamise, leidmise ja / või kasutuselevõtmisega omal algatusel ilma juhtkonna poolsete suunamiseta.
    e. ettevõttes ei ole töötajaid, kes tegeleksid uuenduste väljatöötamise, leidmise ja / või kasutuselevõtmisega.
17. Miks ei olnud Teie ettevõttel perioodil 01.07.2022-30.06.2024 rohkem uuendusi (teenuse- ja protsessiuuendused kokku)? Valida peamine põhjus.
    a. ressursside puudus
    b. vajadus puudus
    c. muud põhjused

IV osa – Üldinfo

Üldinfo vastava ettevõtte kohta

18. Audiitorettevõtja keskmine töötajate arv täistööajale taandatuna perioodil 01.07.2023-30.06.2024
    a. 1-5
    b. 6-10
    c. 11-20
    d. 20-50
    e. Üle 50

19. Audiitorettevõtja kutseteenuste käive eurodes perioodil 01.07.2023-30.06.2024
    a. > 1 000 000
    b. 100 000 – 1 000 000
    c. < 100 000
20. Kas Teie ettevõte kuulub rahvusvahelisse võrgustikku?
    a. Jah
    b. Ei
21. Kui soovite midagi lisada arenevate tehnoloogiate rakendamise kohta audiitortegevuses või käesoleva küsimustiku kohta, siis on seda võimalik teha siin:
    Avatud vastusega lahter

LISA B

Vastusevariandid ja koodid

Nr	Lühikirjeldus	Tüüp	Variant 1	Variant 2	Variant 3	Variant 4	Variant 5	Kood 1	Kood 2	Kood 3	Kood 4	Kood 5
1a	Kasu - AI	järjestus	0 – kasu puudub	1 – väike kasu	2 – mõõdukas kasu	3 – suur kasu		0	1	2	3	
1b	Kasu - NLP	järjestus	0 – kasu puudub	1 – väike kasu	2 – mõõdukas kasu	3 – suur kasu		0	1	2	3	
1c	Kasu - RPA	järjestus	0 – kasu puudub	1 – väike kasu	2 – mõõdukas kasu	3 – suur kasu		0	1	2	3	
1d	Kasu - IoT	järjestus	0 – kasu puudub	1 – väike kasu	2 – mõõdukas kasu	3 – suur kasu		0	1	2	3	
1e	Kasu - HKrüpt	järjestus	0 – kasu puudub	1 – väike kasu	2 – mõõdukas kasu	3 – suur kasu		0	1	2	3	
1f	Kasu - API	järjestus	0 – kasu puudub	1 – väike kasu	2 – mõõdukas kasu	3 – suur kasu		0	1	2	3	
1g	Kasu - Standard.	järjestus	0 – kasu puudub	1 – väike kasu	2 – mõõdukas kasu	3 – suur kasu		0	1	2	3	
1h	Kasu - ADA	järjestus	0 – kasu puudub	1 – väike kasu	2 – mõõdukas kasu	3 – suur kasu		0	1	2	3	
1i	Kasu - Blockchain	järjestus	0 – kasu puudub	1 – väike kasu	2 – mõõdukas kasu	3 – suur kasu		0	1	2	3	
1j	Kasu - PMining	järjestus	0 – kasu puudub	1 – väike kasu	2 – mõõdukas kasu	3 – suur kasu		0	1	2	3	
1k	Kasu - OCR / CVision	järjestus	0 – kasu puudub	1 – väike kasu	2 – mõõdukas kasu	3 – suur kasu		0	1	2	3	
2a	Kasu - Kvaliteedi tõus	järjestus	0 – kasu puudub	1 – väike kasu	2 – mõõdukas kasu	3 – suur kasu		0	1	2	3	
2b	Kasu - Ajavõit	järjestus	0 – kasu puudub	1 – väike kasu	2 – mõõdukas kasu	3 – suur kasu		0	1	2	3	
2c	Kasu - Efektiivsuse tõus	järjestus	0 – kasu puudub	1 – väike kasu	2 – mõõdukas kasu	3 – suur kasu		0	1	2	3	
2d	Kasu - Vigade vähenemine	järjestus	0 – kasu puudub	1 – väike kasu	2 – mõõdukas kasu	3 – suur kasu		0	1	2	3	

2e	Kasu - Rohkem aega keerulisema jaoks	järjestus	0 – kasu puudub	1 – väike kasu	2 – mõõdukas kasu	3 – suur kasu		0	1	2	3	
2f	Kasu - Konkurentsivõime tõus	järjestus	0 – kasu puudub	1 – väike kasu	2 – mõõdukas kasu	3 – suur kasu		0	1	2	3	
2g	Kasu - Standardite täitmise paranemine	järjestus	0 – kasu puudub	1 – väike kasu	2 – mõõdukas kasu	3 – suur kasu		0	1	2	3	
3a	Takistus - Regulatiivne	järjestus	0 – ei ole takistuseks	1 – väike takistav mõju	2 – mõõdukas takistav mõju	3 – suur takistav mõju		0	1	2	3	
3b	Takistus - Aega puudu	järjestus	0 – ei ole takistuseks	1 – väike takistav mõju	2 – mõõdukas takistav mõju	3 – suur takistav mõju		0	1	2	3	
3c	Takistus - Teadmisi puudu	järjestus	0 – ei ole takistuseks	1 – väike takistav mõju	2 – mõõdukas takistav mõju	3 – suur takistav mõju		0	1	2	3	
3d	Takistus - Finantse puudu ise arendamiseks	järjestus	0 – ei ole takistuseks	1 – väike takistav mõju	2 – mõõdukas takistav mõju	3 – suur takistav mõju		0	1	2	3	
3e	Takistus - Finantse puudu ostmiseks	järjestus	0 – ei ole takistuseks	1 – väike takistav mõju	2 – mõõdukas takistav mõju	3 – suur takistav mõju		0	1	2	3	
3f	Takistus - Oskustöötajate värbamise probleem	järjestus	0 – ei ole takistuseks	1 – väike takistav mõju	2 – mõõdukas takistav mõju	3 – suur takistav mõju		0	1	2	3	
3g	Takistus - Klient ei vaja	järjestus	0 – ei ole takistuseks	1 – väike takistav mõju	2 – mõõdukas takistav mõju	3 – suur takistav mõju		0	1	2	3	
3h	Takistus - Pole tasuv investeering	järjestus	0 – ei ole takistuseks	1 – väike takistav mõju	2 – mõõdukas takistav mõju	3 – suur takistav mõju		0	1	2	3	
3i	Takistus - Kliendi sisendandmed valel kujul	järjestus	0 – ei ole takistuseks	1 – väike takistav mõju	2 – mõõdukas takistav mõju	3 – suur takistav mõju		0	1	2	3	
4a	Kasutusel - AI	binaarne	valitud	pole valitud				1	0			
4b	Kasutusel - NLP	binaarne	valitud	pole valitud				1	0			
4c	Kasutusel - RPA	binaarne	valitud	pole valitud				1	0			
4d	Kasutusel - IoT	binaarne	valitud	pole valitud				1	0			
4e	Kasutusel - HKrüpt	binaarne	valitud	pole valitud				1	0			
4f	Kasutusel - API	binaarne	valitud	pole valitud				1	0			
4g	Kasutusel - Standard.	binaarne	valitud	pole valitud				1	0			
4h	Kasutusel - ADA	binaarne	valitud	pole valitud				1	0			
4i	Kasutusel - Blockchain	binaarne	valitud	pole valitud				1	0			

4j	Kasutusel - PMining	binaarne	valitud	pole valitud				1	0			
4k	Kasutusel - OCR / CVision	binaarne	valitud	pole valitud				1	0			
4l	Kasutusel - ükski pole	binaarne	valitud	pole valitud				1	0			
5a	Kasutusel väljatöötaja - Ise	binaarne	valitud	pole valitud				1	0			
5b	Kasutusel väljatöötaja - Koostöö võrgustikuga	binaarne	valitud	pole valitud				1	0			
5c	Kasutusel väljatöötaja - Koostöö teistega	binaarne	valitud	pole valitud				1	0			
5d	Kasutusel väljatöötaja - Teine	binaarne	valitud	pole valitud				1	0			
5e	Kasutusel väljatöötaja - Ei ole kasutusel	binaarne	valitud	pole valitud				1	0			
6a	Kasutusele tõen. - AI	järjestus	1 – väga ebatõenäoline	2 – ebatõenäoline	3 – ei oska öelda	4 – tõenäoline	5 – väga tõenäoline	1	2	3	4	5
6b	Kasutusele tõen. - NLP	järjestus	1 – väga ebatõenäoline	2 – ebatõenäoline	3 – ei oska öelda	4 – tõenäoline	5 – väga tõenäoline	1	2	3	4	5
6c	Kasutusele tõen. - RPA	järjestus	1 – väga ebatõenäoline	2 – ebatõenäoline	3 – ei oska öelda	4 – tõenäoline	5 – väga tõenäoline	1	2	3	4	5
6d	Kasutusele tõen. - IoT	järjestus	1 – väga ebatõenäoline	2 – ebatõenäoline	3 – ei oska öelda	4 – tõenäoline	5 – väga tõenäoline	1	2	3	4	5
6e	Kasutusele tõen. - HKrüpt	järjestus	1 – väga ebatõenäoline	2 – ebatõenäoline	3 – ei oska öelda	4 – tõenäoline	5 – väga tõenäoline	1	2	3	4	5
6f	Kasutusele tõen. - API	järjestus	1 – väga ebatõenäoline	2 – ebatõenäoline	3 – ei oska öelda	4 – tõenäoline	5 – väga tõenäoline	1	2	3	4	5
6g	Kasutusele tõen. - Standard.	järjestus	1 – väga ebatõenäoline	2 – ebatõenäoline	3 – ei oska öelda	4 – tõenäoline	5 – väga tõenäoline	1	2	3	4	5
6h	Kasutusele tõen. - ADA	järjestus	1 – väga ebatõenäoline	2 – ebatõenäoline	3 – ei oska öelda	4 – tõenäoline	5 – väga tõenäoline	1	2	3	4	5
6i	Kasutusele tõen. - Blockchain	järjestus	1 – väga ebatõenäoline	2 – ebatõenäoline	3 – ei oska öelda	4 – tõenäoline	5 – väga tõenäoline	1	2	3	4	5
6j	Kasutusele tõen. - PMining	järjestus	1 – väga ebatõenäoline	2 – ebatõenäoline	3 – ei oska öelda	4 – tõenäoline	5 – väga tõenäoline	1	2	3	4	5
6k	Kasutusele tõen. - OCR / CVision	järjestus	1 – väga ebatõenäoline	2 – ebatõenäoline	3 – ei oska öelda	4 – tõenäoline	5 – väga tõenäoline	1	2	3	4	5
7a	Tõen. tul. väljatöötaja - Ise	järjestus	1 – väga ebatõenäoline	2 – ebatõenäoline	3 – ei oska öelda	4 – tõenäoline	5 – väga tõenäoline	1	2	3	4	5

7b	Tõen. tul. väljatöötaja - Koostöö võrgustikuga	järjestus	1 – väga ebatõenäoline	2 – ebatõenäoline	3 – ei oska öelda	4 – tõenäoline	5 – väga tõenäoline	1	2	3	4	5
7c	Tõen. tul. väljatöötaja - Koostöö teistega	järjestus	1 – väga ebatõenäoline	2 – ebatõenäoline	3 – ei oska öelda	4 – tõenäoline	5 – väga tõenäoline	1	2	3	4	5
7d	Tõen. tul. väljatöötaja - Teine	järjestus	1 – väga ebatõenäoline	2 – ebatõenäoline	3 – ei oska öelda	4 – tõenäoline	5 – väga tõenäoline	1	2	3	4	5
8	TI - turule uus	binaarne	jah	Ei	Ei oska öelda			1	0	99		
9	TI - ettevõttele uus	binaarne	jah	Ei				1	0			
10	TI - täiustatud	binaarne	jah	Ei				1	0			
11	TI - Eestis uus	binaarne	jah	Ei	Ei oska öelda			1	0	99		
12	TI - Eestis varem olemas	binaarne	jah	Ei	Ei oska öelda			1	0	99		
13												
13a	TI - väljatöötaja ise	binaarne	valitud	pole valitud				1	0			
13b	TI - väljatöötaja koostöö	binaarne	valitud	pole valitud				1	0			
13c	TI - kohandatud teenus	binaarne	valitud	pole valitud				1	0			
13d	TI - väljatöötaja keegi teine	binaarne	valitud	pole valitud				1	0			
13e	TI puudus	binaarne	valitud	pole valitud				1	0			
14a	PI - teenuste osutamine	binaarne	valitud	pole valitud				1	0			
14b	PI - logistika	binaarne	valitud	pole valitud				1	0			
14c	PI - IKT	binaarne	valitud	pole valitud				1	0			
14d	PI - raamatupidamine jms	binaarne	valitud	pole valitud				1	0			
14e	PI - äriprotsesside korraldus	binaarne	valitud	pole valitud				1	0			
14f	PI - töökorraldus, inimressursid	binaarne	valitud	pole valitud				1	0			
14g	PI - turundus	binaarne	valitud	pole valitud				1	0			
14h	PI - mitte ükski neist	binaarne	valitud	pole valitud				1	0			
15a	PI - väljatöötaja ise	binaarne	valitud	pole valitud				1	0			
15b	PI - väljatöötaja koostöö	binaarne	valitud	pole valitud				1	0			

15c	PI - kohandatud Protsess	binaarne	valitud	pole valitud				1	0			
15d	PI - väljatöötaja keegi teine	binaarne	valitud	pole valitud				1	0			
15e	PI puudus	binaarne	valitud	pole valitud				1	0			
16	Innovatsioonitöötajad	nominaal	a.	b.	c.	d.	e.	1	2	3	3	5
17	Rohkemate innovatsioonide puudumise põhjus	nominaal	a. ressursside puudus	b. vajadus puudus	c. muud põhjused			1	2	3		
18	Töötajate arv	järjestus	a. 1-5	b. 6-10	c. 11-20	d. 20-50	e. Üle 50	1	2	3	4	5
19	Suurus käibe alusel	järjestus	a. > 1 000 000	b. 100 000 – 1 000 000	c. < 100 000			3	2	1		
20	Kuuluvus RV võrgustikku	binaarne	a. Jah	b. Ei				1	0			
21	Lisainfo											

LISA C

Korrelatsioonanalüüs

Crameri V korrelatsioonianalüüsi tulemused AT kasu hinnangu ja 30.06.2024 seisuga kasutamise fakti vahel		
Paar	**Crameri V**	**p-väärtus**
X1a..Kasu...AI vs X4a..Kasutusel...AI	0,28	0,665
X1b..Kasu...NLP vs X4b..Kasutusel...NLP	0,28	0,439
X1c..Kasu...RPA vs X4c..Kasutusel...RPA	0,37	0,231
X1d..Kasu...IoT vs X4d..Kasutusel...IoT	0,37	0,196
X1e..Kasu...HKrüpt vs X4e..Kasutusel...HKrüpt	0,34	0,206
X1f..Kasu...API vs X4f..Kasutusel...API	0,42	0,152
X1g..Kasu...Standard. vs X4g..Kasutusel...Standard.	0,31	0,446
X1h..Kasu...ADA vs X4h..Kasutusel...ADA	0,26	0,769
X1i..Kasu...Blockchain vs X4i..Kasutusel...Blockchain	0,39	0,276
X1j..Kasu...PMining vs X4j..Kasutusel...PMining	0,30	0,504
X1k..Kasu...OCR...CVision vs X4k..Kasutusel...OCR...CVision	0,27	0,547

Siin on arvutatud tehnoloogiast oodatava kasu korrelatsioonid (Crameri V) sellega, kas tehnoloogia on juba kasutusel iga üheteistkümne tehnoloogia kohta. Tulemused näitavad, et kuigi korrelatsioonid on positiivsed, ei ole need ühegi tehnoloogia puhul statistiliselt olulised.

LISA D

Cronbachi alfa

Cronbachi alfa kõikide küsimuste ning korrigeeritud andmete põhjal

Kõik küsimused				
Toores alfa	95% usalduspiirid	Alumine usalduspiir	Cronbachi alfa	Ülemine usalduspiir
0,41	Feldti meetodil	0,09	0,41	0,66
	Duhacheki meetodil	0,23	0,41	0,59
Pärast küsimuste 8-13 eemaldamist				
Toores alfa	95% usalduspiirid	Alumine usalduspiir	Cronbachi alfa	Ülemine usalduspiir
0,94	Feldti meetodil	0,9	0,94	0,96
	Duhacheki meetodil	0,91	0,94	0,96

Märkus: Küsimustiku usaldusväärsus on hea kui Cronbachi alfa on suurem kui 0,8

LISA E

Tekstis kasutatud lühendite tähendused

ADA – täiustatud andmeanalüüsi tehnoloogiad *(Advanced Data Analytics)*

AEV – audiitorettevõtja

AI – tehisintellekt *(Artificial Intelligence)*

AJN - Audiitortegevuse Järelevalve Nõukogu

API – rakendusliides *(Application Programming Interfaces)*

AT – arenev tehnoloogia

Blockchain – plokiahela tehnoloogia

CIS - ettevõtete innovatsiooniuuring (osa Euroopaülesest uuringust *Community Innovation Survey*)

CVision – arvutinägemine *(Computer Vision)*

HKrüpt – homomorfne krüpteerimine

IAASB - Rahvusvaheline Auditeerimise ja Kindlustandvate Töövõttude Standardite Komitee

IKT – info- ja kommunikatsioonitehnoloogia

IoT – asjade internet *(Internet of Things)*

NLP – loodusliku keele töötlemine *(Natural Language Processing)*

OCR – optiline märgituvastus *(Optical Character Recognition)*

PMining – protsessikaevandamine *(ProcessMining)*

RPA – robootiline protsesside automatiseerimine *(Robotic Process Automation)*

Standard. – andmete standardiseerimine

TI – teenuseinnovatsioon

LISA F

Programmi R koodid

```
####################
# Keskkonna ettevalmistus #
####################

# Laadime vajalikud paketid
if (!require(pacman, quietly = TRUE)) install.packages(pacman)

pacman::p_load(
  dplyr,         # Andmete vormindamine
  readr,         # Andmete lugemine failidest
  tidyr          # Tabelite teisendused
)


#############
# Andmete lugemine #
#############

responses_path <- "responses.csv"
responses <- readr::read_csv(responses_path, show_col_types = FALSE)

respondent_key_path <- "respondent_key.csv"
respondent_key <- readr::read_csv(respondent_key_path, show_col_types = FALSE)

weights_path <- "weights.csv"
weights <- readr::read_csv(weights_path, show_col_types = FALSE)

estonian_innovation_survey_data_path <- "estonian_innovation_survey_data_fin.csv"
estonian_innovation_survey_data <- readr::read_csv(estonian_innovation_survey_data_path, show_col_types = FALSE)

estonian_innovation_survey_data_overall_path <- "estonian_innovation_survey_data_overall.csv"
estonian_innovation_survey_data_overall <- readr::read_csv(estonian_innovation_survey_data_overall_path, show_col_types = FALSE)

# innovative_companies_path <- "innovative_companies_wilsonCI.csv"
# innovative_companies <- readr::read_csv(innovative_companies_path, show_col_types = FALSE) |>
#   dplyr::mutate(across(everything(), ~ gsub("", '', .)))  # Eemaldame jutumärgid kõikidest veergudest


############################
# Innovatsiooni parameetrite arvutamine #
############################

confidence_interval_confidence_level <- 0.95
z_value <- qnorm(1 - (1 - confidence_interval_confidence_level) / 2)

process_innovation_question_ids <- 82
service_innovation_question_ids <- 60:62



# Innovatsiooni parameetrid ettevõtte kaupa
innovators_by_respondent <- responses |>
  dplyr::filter(
    question_id %in% c(process_innovation_question_ids, service_innovation_question_ids)) |>
  dplyr::mutate(
    is_selected = case_when(
      code == 0 ~ FALSE,
      code == 1 ~ TRUE,
      # Kood 99 ("ei tea") muudetakse väärtuseks 0 ("ei")
      # code == 99 ~ FALSE
      )) |>
  dplyr::group_by(respondent_id) |>
  dplyr::summarise(
    has_process_innovation = any(question_id %in% process_innovation_question_ids & !is_selected)  ,
    has_service_innovation = any(question_id %in% service_innovation_question_ids & is_selected)) |>
  dplyr::mutate(
    has_only_process_innovation = has_process_innovation & !has_service_innovation,
    has_only_service_innovation = has_service_innovation & !has_process_innovation,
    has_process_and_service_innovation = has_process_innovation & has_service_innovation,
    # Ettevõte on innovaatiline, kui vähemalt ühes kategoorias on innovatsiooni tegevusi
    has_process_or_service_innovation = has_process_innovation | has_service_innovation)
```

```
# print(innovators_by_respondent)
# summary(innovators_by_respondent)

# Innovatsiooni parameetrid ettevõtte suuruse kaupa
innovators_by_size <- innovators_by_respondent |>
  dplyr::left_join(
    respondent_key,
    join_by(respondent_id)) |>
  dplyr::group_by(respondent_size_id) |>
  dplyr::summarise(
    process_innovation_proportion = mean(has_only_process_innovation),
    service_innovation_proportion = mean(has_only_service_innovation),
    process_and_service_innovation_proportion = mean(has_process_and_service_innovation),
    process_or_service_innovation_proportion = mean(has_process_or_service_innovation),
    n_respondents = n()) |>
  tidyr::pivot_longer(
    cols = ends_with("_innovation_proportion"),
    names_to = "innovation_type",
    values_to = "proportion_raw",
    names_transform = \(string) stringr::str_remove(string, "_innovation_proportion"))|>
  # Add n_true column by multiplying proportion_raw by n_respondents
  dplyr::mutate(
    n_true = proportion_raw * n_respondents
  )


print(innovators_by_size)
# summary(innovators_by_size)

# Wilsoni meetodi rakendamine ettevõtte suuruse kaupa
innovation_proportions <- innovators_by_size |>
  dplyr::left_join(
    weights,
    join_by(respondent_size_id)) |>
  dplyr::group_by(innovation_type) |>
  dplyr::mutate(
  #   weight_multiplier = weight_post_stratification / sum(weight_post_stratification)) |>
  #   # weight_multiplier = 1/3)
  weight_multiplier = weight_post_stratification) |>
  dplyr::summarise(
    # proportion_weighted = sum(proportion_raw * weight_multiplier),
    # n_weighted = sum(weight_multiplier * n_respondents),
    n_weighted = sum(n_true * weight_multiplier),
    proportion_weighted = n_weighted / 116,
    lower_bound = (
      (2 * n_weighted * proportion_weighted + z_value^2) -
        z_value * sqrt(z_value^2 - (1 / n_weighted) + 4 * n_weighted * proportion_weighted * (1 - proportion_weighted))) /
      (2 * (n_weighted + z_value^2)),
    upper_bound = (
      (2 * n_weighted * proportion_weighted + z_value^2) +
        z_value * sqrt(z_value^2 - (1 / n_weighted) + 4 * n_weighted * proportion_weighted * (1 - proportion_weighted))) /
      (2 * (n_weighted + z_value^2))
  )

print(innovation_proportions)

# Kokkuvõtvad innovatsiooni parameetrid turu kohta (CIS 2022: Finants- ja kindlustussektor)
estonian_innovation_survey_proportions <- estonian_innovation_survey_data |>
  # Filter out rows with NA innovation_type
  dplyr::filter(!is.na(innovation_type)) |>
  # Calculate total number of valid companies (excluding NA rows)
  dplyr::mutate(total_companies = 110) |>
  # Compute proportions based on the total number of valid companies
  dplyr::mutate(
    source = "estonian_innovation_survey",
    proportion_weighted = n_companies / total_companies,
    # Variance calculation for proportions
    variance = proportion_weighted * (1 - proportion_weighted) / (n_companies - 1),
    # Calculate Wilson confidence interval lower and upper bounds
    lower_bound = (
      (2 * n_companies * proportion_weighted + z_value^2) -
        z_value * sqrt(z_value^2 - (1 / n_companies) + 4 * n_companies * proportion_weighted * (1 - proportion_weighted))) /
      (2 * (n_companies + z_value^2)),
    upper_bound = (
```

```
    (2 * n_companies * proportion_weighted + z_value^2) +
      z_value * sqrt(z_value^2 - (1 / n_companies) + 4 * n_companies * proportion_weighted * (1 - proportion_weighted))) /
    (2 * (n_companies + z_value^2))
  )

# print(estonian_innovation_survey_proportions)


# Kokkuvõtvad innovatsiooni parameetrid turu kohta (CIS 2022: Kõik ettevõtted)
estonian_innovation_survey_overall_proportions <- estonian_innovation_survey_data_overall |>
  # Filter out rows with NA innovation_type
  dplyr::filter(!is.na(innovation_type)) |>
  # Calculate total number of valid companies (excluding NA rows)
  dplyr::mutate(total_companies = 3859) |>
  # Compute proportions based on the total number of valid companies
  dplyr::mutate(
    source = "estonian_innovation_survey_overall",
    proportion_weighted = n_companies / total_companies,
    # Variance calculation for proportions
    variance = proportion_weighted * (1 - proportion_weighted) / (n_companies - 1),
    # Calculate Wilson confidence interval lower and upper bounds
    lower_bound = (
      (2 * n_companies * proportion_weighted + z_value^2) -
        z_value * sqrt(z_value^2 - (1 / n_companies) + 4 * n_companies * proportion_weighted * (1 - proportion_weighted))) /
      (2 * (n_companies + z_value^2)),
    upper_bound = (
      (2 * n_companies * proportion_weighted + z_value^2) +
        z_value * sqrt(z_value^2 - (1 / n_companies) + 4 * n_companies * proportion_weighted * (1 - proportion_weighted))) /
      (2 * (n_companies + z_value^2))
  )

# print(estonian_innovation_survey_overall_proportions)



# Lõplik tabel kõigi andmetega
innovative_companies <- innovation_proportions |>
  dplyr::mutate(
    source = "current_study",
    confidence_level = confidence_interval_confidence_level  # Lisa usaldusnivoo
  ) |>
  dplyr::bind_rows(
    estonian_innovation_survey_proportions |>
      dplyr::mutate(confidence_level = confidence_interval_confidence_level)  # Lisa survey data
  ) |>
  dplyr::bind_rows(  # Lisa kõikide ettevõtete survey data
    estonian_innovation_survey_overall_proportions |>
      dplyr::mutate(confidence_level = confidence_interval_confidence_level)
  ) |>
  dplyr::select(
    innovation_type,
    source,
    proportion_weighted,
    confidence_level,
    lower_bound,
    upper_bound
  )



#############
# Andmete salvestamine #
#############

write.csv(innovative_companies, "innovative_companies_wilsonCI.csv", row.names = FALSE)

print(innovative_companies)

####################
# Keskkonna ettevalmistus #
####################

# Laadime vajalikud paketid
if (!require(pacman, quietly = TRUE)) install.packages(pacman)
```

```
pacman::p_load(
  dplyr,         # Andmete vormindamine
  readr,         # Andmete lugemine failidest
  tidyr          # Tabelite teisendused
)


#############
# Andmete lugemine #
#############

responses_path <- "responses.csv"
responses <- readr::read_csv(responses_path, show_col_types = FALSE)

respondent_key_path <- "respondent_key.csv"
respondent_key <- readr::read_csv(respondent_key_path, show_col_types = FALSE)

weights_path <- "weights.csv"
weights <- readr::read_csv(weights_path, show_col_types = FALSE)

estonian_innovation_survey_data_path <- "estonian_innovation_survey_data_fin.csv"
estonian_innovation_survey_data <- readr::read_csv(estonian_innovation_survey_data_path, show_col_types = FALSE)

estonian_innovation_survey_data_overall_path <- "estonian_innovation_survey_data_overall.csv"
estonian_innovation_survey_data_overall <- readr::read_csv(estonian_innovation_survey_data_overall_path, show_col_types = FALSE)

# innovative_companies_path <- "innovative_companies_wilsonCI.csv"
# innovative_companies <- readr::read_csv(innovative_companies_path, show_col_types = FALSE) |>
#   dplyr::mutate(across(everything(), ~ gsub("", '', .)))  # Eemaldame jutumärgid kõikidest veergudest


############################
# Innovatsiooni parameetrite arvutamine #
############################

confidence_interval_confidence_level <- 0.95
z_value <- qnorm(1 - (1 - confidence_interval_confidence_level) / 2)

process_innovation_question_ids <- 82
service_innovation_question_ids <- 60:62



# Innovatsiooni parameetrid ettevõtte kaupa
innovators_by_respondent <- responses |>
  dplyr::filter(
    question_id %in% c(process_innovation_question_ids, service_innovation_question_ids)) |>
  dplyr::mutate(
    is_selected = case_when(
      code == 0 ~ FALSE,
      code == 1 ~ TRUE,
      # Kood 99 ("ei tea") muudetakse väärtuseks 0 ("ei")
      # code == 99 ~ FALSE
      )) |>
  dplyr::group_by(respondent_id) |>
  dplyr::summarise(
    has_process_innovation = any(question_id %in% process_innovation_question_ids & !is_selected)  ,
    has_service_innovation = any(question_id %in% service_innovation_question_ids & is_selected)) |>
  dplyr::mutate(
    has_only_process_innovation = has_process_innovation & !has_service_innovation,
    has_only_service_innovation = has_service_innovation & !has_process_innovation,
    has_process_and_service_innovation = has_process_innovation & has_service_innovation,
    # Ettevõte on innovaatiline, kui vähemalt ühes kategoorias on innovatsiooni tegevusi
    has_process_or_service_innovation = has_process_innovation | has_service_innovation)

# print(innovators_by_respondent)
# summary(innovators_by_respondent)

# Innovatsiooni parameetrid ettevõtte suuruse kaupa
innovators_by_size <- innovators_by_respondent |>
  dplyr::left_join(
    respondent_key,
    join_by(respondent_id)) |>
  dplyr::group_by(respondent_size_id) |>
```

```
  dplyr::summarise(
    process_innovation_proportion = mean(has_only_process_innovation),
    service_innovation_proportion = mean(has_only_service_innovation),
    process_and_service_innovation_proportion = mean(has_process_and_service_innovation),
    process_or_service_innovation_proportion = mean(has_process_or_service_innovation),
    n_respondents = n()) |>
  tidyr::pivot_longer(
    cols = ends_with("_innovation_proportion"),
    names_to = "innovation_type",
    values_to = "proportion_raw",
    names_transform = \(string) stringr::str_remove(string, "_innovation_proportion"))|>
  # Add n_true column by multiplying proportion_raw by n_respondents
  dplyr::mutate(
    n_true = proportion_raw * n_respondents
  )


print(innovators_by_size)
# summary(innovators_by_size)

# Wilsoni meetodi rakendamine ettevõtte suuruse kaupa
innovation_proportions <- innovators_by_size |>
  dplyr::left_join(
    weights,
    join_by(respondent_size_id)) |>
  dplyr::group_by(innovation_type) |>
  dplyr::mutate(
  #   weight_multiplier = weight_post_stratification / sum(weight_post_stratification)) |>
  #   # weight_multiplier = 1/3)
  weight_multiplier = weight_post_stratification) |>
  dplyr::summarise(
    # proportion_weighted = sum(proportion_raw * weight_multiplier),
    # n_weighted = sum(weight_multiplier * n_respondents),
    n_weighted = sum(n_true * weight_multiplier),
    proportion_weighted = n_weighted / 116,
    lower_bound = (
      (2 * n_weighted * proportion_weighted + z_value^2) -
        z_value * sqrt(z_value^2 - (1 / n_weighted) + 4 * n_weighted * proportion_weighted * (1 - proportion_weighted))) /
      (2 * (n_weighted + z_value^2)),
    upper_bound = (
      (2 * n_weighted * proportion_weighted + z_value^2) +
        z_value * sqrt(z_value^2 - (1 / n_weighted) + 4 * n_weighted * proportion_weighted * (1 - proportion_weighted))) /
      (2 * (n_weighted + z_value^2))
  )

print(innovation_proportions)

# Kokkuvõtvad innovatsiooni parameetrid turu kohta (CIS 2022: Finants- ja kindlustussektor)
estonian_innovation_survey_proportions <- estonian_innovation_survey_data |>
  # Filter out rows with NA innovation_type
  dplyr::filter(!is.na(innovation_type)) |>
  # Calculate total number of valid companies (excluding NA rows)
  dplyr::mutate(total_companies = 110) |>
  # Compute proportions based on the total number of valid companies
  dplyr::mutate(
    source = "estonian_innovation_survey",
    proportion_weighted = n_companies / total_companies,
    # Variance calculation for proportions
    variance = proportion_weighted * (1 - proportion_weighted) / (n_companies - 1),
    # Calculate Wilson confidence interval lower and upper bounds
    lower_bound = (
      (2 * n_companies * proportion_weighted + z_value^2) -
        z_value * sqrt(z_value^2 - (1 / n_companies) + 4 * n_companies * proportion_weighted * (1 - proportion_weighted))) /
      (2 * (n_companies + z_value^2)),
    upper_bound = (
      (2 * n_companies * proportion_weighted + z_value^2) +
        z_value * sqrt(z_value^2 - (1 / n_companies) + 4 * n_companies * proportion_weighted * (1 - proportion_weighted))) /
      (2 * (n_companies + z_value^2))
  )

# print(estonian_innovation_survey_proportions)


# Kokkuvõtvad innovatsiooni parameetrid turu kohta (CIS 2022: Kõik ettevõtted)
estonian_innovation_survey_overall_proportions <- estonian_innovation_survey_data_overall |>
```

```
  # Filter out rows with NA innovation_type
  dplyr::filter(!is.na(innovation_type)) |>
  # Calculate total number of valid companies (excluding NA rows)
  dplyr::mutate(total_companies = 3859) |>
  # Compute proportions based on the total number of valid companies
  dplyr::mutate(
    source = "estonian_innovation_survey_overall",
    proportion_weighted = n_companies / total_companies,
    # Variance calculation for proportions
    variance = proportion_weighted * (1 - proportion_weighted) / (n_companies - 1),
    # Calculate Wilson confidence interval lower and upper bounds
    lower_bound = (
      (2 * n_companies * proportion_weighted + z_value^2) -
        z_value * sqrt(z_value^2 - (1 / n_companies) + 4 * n_companies * proportion_weighted * (1 - proportion_weighted))) /
      (2 * (n_companies + z_value^2)),
    upper_bound = (
      (2 * n_companies * proportion_weighted + z_value^2) +
        z_value * sqrt(z_value^2 - (1 / n_companies) + 4 * n_companies * proportion_weighted * (1 - proportion_weighted))) /
      (2 * (n_companies + z_value^2))
  )

# print(estonian_innovation_survey_overall_proportions)



# Lõplik tabel kõigi andmetega
innovative_companies <- innovation_proportions |>
  dplyr::mutate(
    source = "current_study",
    confidence_level = confidence_interval_confidence_level  # Lisa usaldusnivoo
  ) |>
  dplyr::bind_rows(
    estonian_innovation_survey_proportions |>
      dplyr::mutate(confidence_level = confidence_interval_confidence_level)  # Lisa survey data
  ) |>
  dplyr::bind_rows(  # Lisa kõikide ettevõtete survey data
    estonian_innovation_survey_overall_proportions |>
      dplyr::mutate(confidence_level = confidence_interval_confidence_level)
  ) |>
  dplyr::select(
    innovation_type,
    source,
    proportion_weighted,
    confidence_level,
    lower_bound,
    upper_bound
  )



#############
# Andmete salvestamine #
#############

write.csv(innovative_companies, "innovative_companies_wilsonCI.csv", row.names = FALSE)

print(innovative_companies)

####################
# Keskkonna ettevalmistus #
####################

# Laadime vajalikud paketid
if (!require(pacman, quietly = TRUE)) install.packages(pacman)
# Laadime vajalikud paketid
if (!require(pacman)) install.packages("pacman")
pacman::p_load(openxlsx, dplyr)

pacman::p_load(
  dplyr,         # Andmete vormindamine
  readr,         # Andmete lugemine failidest
  tidyr          # Tabelite teisendused
)
```

```
#############
# Andmete lugemine #
#############

vastused_kodeeritud <- read_excel("vastused_kodeeritud.xlsx")

## Loon muutujad, mis sisaldavad vastavate andmete veeru numberid "vastused_kodeeritud" failis
process_innov_source_ids <- 23:27
weight_ids <- c(132, 136)
suurusklass_id <- 130
confidence_interval_confidence_level <- 0.95

## Kasutades eelnevalt loodud ID-de vektoreid, toon andmed uutesse muutujatesse
data_process_innov_source <- vastused_kodeeritud[, process_innov_source_ids]
data_weights <- vastused_kodeeritud[, weight_ids]
data_suurusklass <- vastused_kodeeritud[, suurusklass_id]

# # Arvutan ka raha kaalu (kaal_turg on juba kahe kaalu korrutis andmetes)
# data_weights <- data_weights %>% dplyr::mutate(
#   kaal_raha = Kaal_turg / Kaal_innov
# )

## Teen koondandmestiku kõikide vajalike muutujatega
data <- cbind(
  data_process_innov_source,
  data_weights,
  data_suurusklass
)

print(data)

# Add labels to the column names
column_labels <- c(
  "15a: PI - väljatöötaja ise" = "Ettevõte ise",
  "15b: PI - väljatöötaja koostöö" = "Koostöös teise\nettevõttega",
  "15c: PI - kohandatud Protsess" = "Kohandatud teise\nettevõtte lahendus",
  "15d: PI - väljatöötaja keegi teine" = "Teise ettevõtte\nlahendus",
  "15e: PI puudus" = "Protsessiinnovatsiooni\nei olnud"
)

# Apply labels to the first 5 columns
colnames(data)[1:5] <- unname(column_labels[colnames(data)[1:5]])

# Veendu, et 16. veerg (suurusklass) on korrektselt tuvastatud
suurusklass_column <- colnames(data)[8]

# Arvuta esimese 5 veeru summa, grupeerituna suurusklassi järgi
grouped_sums <- data %>%
  dplyr::select(1:5, !!sym(suurusklass_column)) %>%  # Vali esimesed 8 veergu ja suurusklassi veerg
  dplyr::group_by(!!sym(suurusklass_column)) %>%     # Grupeeri andmed suurusklassi järgi
  dplyr::summarise(across(1:5, sum, na.rm = TRUE))   # Arvuta igas grupis iga veeru summa, ignoreerides NA väärtusi

# Kuvab gruppide summad
print(grouped_sums)

weight_NR_column <- colnames(data)[7]
print(weight_NR_column)

# Arvuta esimese 8 veeru summa korrutades non-response kaaludega ja grupeerituna suurusklassi järgi
NR_weighted_grouped_sums <- data %>%
  dplyr::mutate(across(1:5, ~ . * !!sym(weight_NR_column))) %>%  # Korruta vastused kaaludega
  dplyr::select(1:5, !!sym(suurusklass_column)) %>%          # Vali muudetud veerud ja suurusklass
  dplyr::group_by(!!sym(suurusklass_column)) %>%             # Grupeeri suurusklassi järgi
  dplyr::summarise(across(1:5, sum, na.rm = TRUE))           # Arvuta igas grupis iga veeru summa

# Kuvab kaalutud summad gruppide lõikes
print(NR_weighted_grouped_sums)

# Calculate NR weighted averages and Wilson confidence intervals for columns 1 to 8
process_innov_source_NR <- data %>%
  dplyr::summarise(across(1:5, ~ {
    # Calculate the weighted sum and n_weighted (sum of weights)
    weighted_sum <- sum(.x * Kaal_innov, na.rm = TRUE)  # Weighted sum
    n_weighted <- sum(Kaal_innov, na.rm = TRUE)         # Effective sample size (sum of weights)
    weighted_proportion <- weighted_sum / 116           # Weighted proportion
```

```
    z_value <- qnorm(1 - (1 - confidence_interval_confidence_level) / 2)

    # Calculate Wilson confidence interval using n_weighted
    lower_bound <- (
      (2 * n_weighted * weighted_proportion + z_value^2) -
        z_value * sqrt(max(0, z_value^2 - (1 / n_weighted) + 4 * n_weighted * weighted_proportion * (1 - weighted_proportion)))) /
      (2 * (n_weighted + z_value^2))
    upper_bound <- (
      (2 * n_weighted * weighted_proportion + z_value^2) +
        z_value * sqrt(max(0, z_value^2 - (1 / n_weighted) + 4 * n_weighted * weighted_proportion * (1 - weighted_proportion)))) /
      (2 * (n_weighted + z_value^2))

    # Return results as a list
    c(proportion_weighted = weighted_proportion, lower_bound = lower_bound, upper_bound = upper_bound)
  }))

# Print the results
print(process_innov_source_NR)

# Salvesta andmestik Exceli formaadis
write.xlsx(process_innov_source_NR, "process_innov_source_NR.xlsx", rowNames = FALSE)

#####################
# Keskkonna ettevalmistus #
#####################

# Laadime vajalikud R-paketid
if (!require(pacman, quietly = TRUE)) install.packages(pacman)

pacman::p_load(
  ggplot2,       # Graafika genereerimine
  readr,         # Andmete lugemine failidest
  tidyr,         # Tabelite teisendused (pikk vs lai formaat)
  dplyr,         # Andmete töötlemine
  scales,        # Telgede siltide vormindamine (nt protsendid)
  stringr        # Tekstidega töötlemine
)
library(readxl)

#############
# Andmete lugemine #
#############

# # Määrame andmefailide asukohad ja loeme need sisse
# market_weighted_columns_summary_path <- "market_weighted_columns_summary.csv"  # Turu kaalutud andmed
# market_weighted_data <- readr::read_csv(market_weighted_columns_summary_path, show_col_types = FALSE)

NR_weighted_columns_summary_path <- "process_innov_source_NR.xslx"  # andmed protsessiinnovatsiooni valdkondade kohta
NR_weighted_data <-process_innov_source_NR <- read_excel("process_innov_source_NR.xlsx")

# Lisame veeru "Source" (allikas), mis näitab, kummast andmestikust rida pärineb
# market_weighted_data <- market_weighted_data %>%
#   mutate(Source = "Osakaal audiitorteenuste turust")  # Turu kaalutud osakaal

NR_weighted_data <- NR_weighted_data %>%
  mutate(Source = "Osakaal audiitorettevõtjatest")  # Vastamisest kõrvale jäänud osakaal

# Kombineerime mõlemad andmestikud üheks suureks andmetabeliks
combined_data <- bind_rows(NR_weighted_data)
print(combined_data)



# Teisendame andmed pikaks formaadiks
plot_data <- combined_data %>%
  pivot_longer(
    cols = -Source,  # Kõik veerud peale "Source"
    names_to = "Category",
    values_to = "Proportion"
  ) %>%

  # Lisame veeru "Metric_Type", et eristada kaalutud keskmisi, alumisi ja ülemisi veapiire
  mutate(Metric_Type = rep(c("Weighted_Mean", "Lower_Bound", "Upper_Bound"), each = 5, times = 1)) %>%
  # Teisendame veerud uuesti laia formaati
  pivot_wider(names_from = Metric_Type, values_from = Proportion)
```

```
  # Kohandame kategooriate sildid (lisame reavahetused)



print(plot_data)

# Loome horisontaalse tulpdiagrammi koos veapiiridega
process_innov_source <- ggplot(plot_data, aes(
  y = reorder(Category, Weighted_Mean),  # Nimetused Y-teljel
  x = Weighted_Mean,                                    # Kaalutud osakaal X-teljel
  fill = Category                                       # Värvid vastavalt allikale
)) +
  geom_bar(
    stat = "identity",                  # Kasutame staatilist identiteeti tulbadiagrammi jaoks
    position = position_dodge(width = 0.7),  # Nihutame tulbad allika järgi
    width = 0.7                         # Tulpade laius
  ) +
  geom_errorbar(
    aes(xmin = Lower_Bound, xmax = Upper_Bound),  # Alumine ja ülemine veapiir
    position = position_dodge(width = 0.7),      # Nihutame veapiirid vastavalt tulpadele
    width = 0.2,                        # Veapiiride horisontaaljoonte laius
    color = "black"                       # Veapiirid musta värvi
  ) +
  labs(
    x = NULL,                           # Eemaldame X-telje pealkirja
    y = NULL,                           # Eemaldame Y-telje pealkirja
    fill = "Kaalumine:",                  # Legendi pealkiri
    caption = "Mustad jooned tähistavad veapiire usaldusnivool 95%"  # Selgitus graafiku all
  ) +
  scale_fill_manual(
    values = c(
      "Protsessiinnovatsiooni\r\nei olnud" = "#FFC107",  # Kollane
      "Ettevõte ise" = "#4CAF50",                  # Roheline
      "Teise ettevõtte\r\nlahendus" = "#4CAF50",        # Roheline
      "Koostöös teise\r\nettevõttega" = "#4CAF50",      # Roheline
      "Kohandatud teise\r\nettevõtte lahendus" = "#4CAF50"  # Roheline
    )
  ) +
  scale_x_continuous(
    labels = percent_format(accuracy = 1)      # X-telg protsentidena
  ) +
  theme_minimal() +
  theme(
    text = element_text(family = "Times New Roman"),  # Kirjastiil
    panel.background = element_rect(fill = "white", color = "white"),  # Valge taust
    plot.background = element_rect(fill = "white", color = "white"),   # Valge taust
    axis.text.x = element_text(size = 12),                          # X-telje tekstide suurus
    axis.text.y = element_text(size = 18, lineheight = 1.),          # Y-telje tekstide suurus ja reavahe
    plot.title = element_text(size = 16, face = "bold", hjust = 0.5),  # Pealkiri
    plot.caption = element_text(size = 12),                          # Selgituste suurus
    # legend.position = "bottom",                                    # Legendi asukoht all
    legend.text = element_blank(),   #_text(size = 14),                          # Legendi tekstide suurus
    legend.title = element_blank(),     #text(size = 14),                        # Legendi pealkirja suurus
    legend.key = element_blank(),
    panel.grid.major.x = element_line(color = "gray80", size = 0.5),   # X-telje gridlines
    panel.grid.minor.x = element_blank()                            # Eemaldame väiksed gridlines
  )

# # Kuvame graafiku
print(process_innov_source)



# Salvestame graafiku
ggsave(
  filename = "07_process_innov_source.png",  # Faili nimi
  plot = process_innov_source,
  width = 12,
  height = 6,
  units = "in",
  dpi = 300
)
#####################
# Keskkonna ettevalmistus #
#####################
```

```
# Laadime vajalikud paketid
if (!require(pacman, quietly = TRUE)) install.packages(pacman)

pacman::p_load(
  dplyr,         # Andmete vormindamine
  readr,         # Andmete lugemine failidest
  tidyr          # Tabelite teisendused
)

#############
# Andmete lugemine #
#############

vastused_kodeeritud <- read_excel("vastused_kodeeritud.xlsx")

## Loon muutujad, mis sisaldavad vastavate andmete veeru numberid "vastused_kodeeritud" failis
AT_in_use_ids <- 67:78
weight_ids <- c(132, 136)
suurusklass_id <- 130
confidence_interval_confidence_level <- 0.95

## Kasutades eelnevalt loodud ID-de vektoreid, toon andmed uutesse muutujatesse
data_AT_in_use <- vastused_kodeeritud[, AT_in_use_ids]
data_weights <- vastused_kodeeritud[, weight_ids]
data_suurusklass <- vastused_kodeeritud[, suurusklass_id]

# Arvutan ka raha kaalu (kaal_turg on juba kahe kaalu korrutis andmetes)
data_weights <- data_weights %>% dplyr::mutate(
  kaal_raha = Kaal_turg / Kaal_innov
)

## Teen koondandmestiku kõikide vajalike muutujatega
data <- cbind(
  data_AT_in_use,
  data_weights,
  data_suurusklass
)

# Add labels to the column names
column_labels <- c(
  "4a: Kasutusel - AI" = "AI",
  "4b: Kasutusel - NLP" = "NLP",
  "4c: Kasutusel - RPA" = "RPA",
  "4d: Kasutusel - IoT" = "IoT",
  "4e: Kasutusel - HKrüpt" = "HKrüpt",
  "4f: Kasutusel - API" = "API",
  "4g: Kasutusel - Standard." = "Standard.",
  "4h: Kasutusel - ADA" = "ADA",
  "4i: Kasutusel - PBlockchain" = "Blockchain",
  "4j: Kasutusel - PMining" = "PMining",
  "4k: Kasutusel - OCR / CVision" = "OCR/Cvision",
  "4l: Kasutusel - ükski pole" = "Mitte ükski neist"
)

# Apply labels to the first 12 columns
colnames(data)[1:12] <- unname(column_labels[colnames(data)[1:12]])

# Veendu, et 16. veerg (suurusklass) on korrektselt tuvastatud
suurusklass_column <- colnames(data)[16]

# Arvuta esimese 12 veeru summa, grupeerituna suurusklassi järgi
grouped_sums <- data %>%
  dplyr::select(1:12, !!sym(suurusklass_column)) %>%  # Vali esimesed 12 veergu ja suurusklassi veerg
  dplyr::group_by(!!sym(suurusklass_column)) %>%      # Grupeeri andmed suurusklassi järgi
  dplyr::summarise(across(1:12, sum, na.rm = TRUE))   # Arvuta igas grupis iga veeru summa, ignoreerides NA väärtusi

# Kuvab gruppide summad
print(grouped_sums)

weight_NR_column <- colnames(data)[14]
print(weight_NR_column)

# Arvuta esimese 12 veeru summa korrutades non-response kaaludega ja grupeerituna suurusklassi järgi
NR_weighted_grouped_sums <- data %>%
```

```
  dplyr::mutate(across(1:12, ~ . * !!sym(weight_NR_column))) %>%  # Korruta vastused kaaludega
  dplyr::select(1:12, !!sym(suurusklass_column)) %>%          # Vali muudetud veerud ja suurusklass
  dplyr::group_by(!!sym(suurusklass_column)) %>%              # Grupeeri suurusklassi järgi
  dplyr::summarise(across(1:12, sum, na.rm = TRUE))           # Arvuta igas grupis iga veeru summa

# Kuvab kaalutud summad gruppide lõikes
print(NR_weighted_grouped_sums)

# Calculate NR weighted averages and Wilson confidence intervals for columns 1 to 12
NR_weighted_columns_summary <- data %>%
  dplyr::summarise(across(1:12, ~ {
    # Calculate the weighted sum and n_weighted (sum of weights)
    weighted_sum <- sum(.x * Kaal_innov, na.rm = TRUE)  # Weighted sum
    n_weighted <- sum(Kaal_innov, na.rm = TRUE)         # Effective sample size (sum of weights)
    weighted_proportion <- weighted_sum / 116           # Weighted proportion
    z_value <- qnorm(1 - (1 - confidence_interval_confidence_level) / 2)

    # Calculate Wilson confidence interval using n_weighted
    lower_bound <- (
      (2 * n_weighted * weighted_proportion + z_value^2) -
        z_value * sqrt(max(0, z_value^2 - (1 / n_weighted) + 4 * n_weighted * weighted_proportion * (1 - weighted_proportion)))) /
      (2 * (n_weighted + z_value^2))
    upper_bound <- (
      (2 * n_weighted * weighted_proportion + z_value^2) +
        z_value * sqrt(max(0, z_value^2 - (1 / n_weighted) + 4 * n_weighted * weighted_proportion * (1 - weighted_proportion)))) /
      (2 * (n_weighted + z_value^2))

    # Return results as a list
    c(proportion_weighted = weighted_proportion, lower_bound = lower_bound, upper_bound = upper_bound)
  }))

# Print the results
print(NR_weighted_columns_summary)

# Save to a CSV file (optional)
write.csv(NR_weighted_columns_summary, "NR_weighted_columns_summary.csv", row.names = FALSE)

######################

# Arvuta esimese 12 veeru summa korrutades turu kaaludega ja grupeerituna suurusklassi järgi
weight_market_column <- colnames(data)[13]
print(weight_market_column)

market_weighted_grouped_sums <- data %>%
  dplyr::mutate(across(1:12, ~ . * !!sym(weight_market_column))) %>%  # Korruta vastused kaaludega
  dplyr::select(1:12, !!sym(suurusklass_column)) %>%          # Vali muudetud veerud ja suurusklass
  dplyr::group_by(!!sym(suurusklass_column)) %>%              # Grupeeri suurusklassi järgi
  dplyr::summarise(across(1:12, sum, na.rm = TRUE))           # Arvuta igas grupis iga veeru summa

# Kuvab kaalutud summad gruppide lõikes
print(market_weighted_grouped_sums)

# Calculate market weighted averages and Wilson confidence intervals for columns 1 to 12
market_weighted_columns_summary <- data %>%
  dplyr::summarise(across(1:12, ~ {
    # Calculate the weighted sum and n_weighted (sum of weights)
    weighted_sum <- sum(.x * Kaal_turg, na.rm = TRUE)  # Weighted sum
    n_weighted <- sum(Kaal_turg, na.rm = TRUE)         # Effective sample size (sum of weights)
    weighted_proportion <- weighted_sum / 116          # Weighted proportion
    z_value <- qnorm(1 - (1 - confidence_interval_confidence_level) / 2)

    # Calculate Wilson confidence interval using n_weighted
    lower_bound <- (
      (2 * n_weighted * weighted_proportion + z_value^2) -
        z_value * sqrt(max(0, z_value^2 - (1 / n_weighted) + 4 * n_weighted * weighted_proportion * (1 - weighted_proportion)))) /
      (2 * (n_weighted + z_value^2))
    upper_bound <- (
      (2 * n_weighted * weighted_proportion + z_value^2) +
        z_value * sqrt(max(0, z_value^2 - (1 / n_weighted) + 4 * n_weighted * weighted_proportion * (1 - weighted_proportion)))) /
      (2 * (n_weighted + z_value^2))

    # Return results as a list
    c(proportion_weighted = weighted_proportion, lower_bound = lower_bound, upper_bound = upper_bound)
  }))
```

```
# Print the results
print(market_weighted_columns_summary)

# Save to a CSV file (optional)
write.csv(market_weighted_columns_summary, "market_weighted_columns_summary.csv", row.names = FALSE)

####################
# Keskkonna ettevalmistus #
####################

# Laadime vajalikud R-paketid
if (!require(pacman, quietly = TRUE)) install.packages(pacman)

pacman::p_load(
  ggplot2,       # Graafika genereerimine
  readr,         # Andmete lugemine failidest
  tidyr,         # Tabelite teisendused (pikk vs lai formaat)
  dplyr,         # Andmete töötlemine
  scales,        # Telgede siltide vormindamine (nt protsendid)
  stringr        # Tekstidega töötlemine
)

#############
# Andmete lugemine #
#############

# Määrame andmefailide asukohad ja loeme need sisse
market_weighted_columns_summary_path <- "market_weighted_columns_summary.csv"  # Turu kaalutud andmed
market_weighted_data <- readr::read_csv(market_weighted_columns_summary_path, show_col_types = FALSE)

NR_weighted_columns_summary_path <- "NR_weighted_columns_summary.csv"  # Vastamisest kõrvale jäänud ettevõtete kaalutud
andmed
NR_weighted_data <- readr::read_csv(NR_weighted_columns_summary_path, show_col_types = FALSE)

# Lisame veeru "Source" (allikas), mis näitab, kummast andmestikust rida pärineb
market_weighted_data <- market_weighted_data %>%
  mutate(Source = "Osakaal audiitorteenuste turust")  # Turu kaalutud osakaal

NR_weighted_data <- NR_weighted_data %>%
  mutate(Source = "Osakaal audiitorettevõtjatest")  # Vastamisest kõrvale jäänud osakaal

# Kombineerime mõlemad andmestikud üheks suureks andmetabeliks
combined_data <- bind_rows(market_weighted_data, NR_weighted_data)

# Teisendame andmed laia formaadist pikkvormi
combined_data_long <- combined_data %>%
  pivot_longer(
    cols = -Source,          # Kõik veerud peale "Source" teisendatakse ridadeks
    names_to = "Technology", # Tehnoloogia nimed tulevad veerunimedest
    values_to = "Value"      # Vastavad väärtused tulevad tabeli sisust
  ) %>%
  # Lisame veeru "Metric_Type", et eristada kaalutud keskmisi, alumisi ja ülemisi veapiire
  mutate(Metric_Type = rep(c("Weighted_Mean", "Lower_Bound", "Upper_Bound"), each = 12, times = 2)) %>%
  # Teisendame veerud uuesti laia formaati
  pivot_wider(names_from = Metric_Type, values_from = Value)

# Järjestame tehnoloogiad turu kaalutud osakaalu alusel kahanevalt
sorted_technologies <- combined_data_long %>%
  filter(Source == "Osakaal audiitorteenuste turust") %>%  # Filtreerime turu andmed
  arrange(desc(Weighted_Mean)) %>%                         # Järjestame kahanevalt
  pull(Technology)                                         # Saame nimekirja tehnoloogiatest

# Rakendame sorteeritud järjekorra tehnoloogia veerule
combined_data_long <- combined_data_long %>%
  mutate(Technology = factor(Technology, levels = sorted_technologies)) %>%
  mutate(Source = factor(Source, levels = c("Osakaal audiitorteenuste turust", "Osakaal audiitorettevõtjatest")))  # Turu osakaal vasakule

##################
# Graafiku loomine #
##################

plot <- ggplot(combined_data_long, aes(x = Technology, y = Weighted_Mean, fill = Source)) +
  geom_bar(
    stat = "identity",                    # Kasutame staatilist identiteeti tulbadiagrammi jaoks
    position = position_dodge(width = 0.7),  # Nihutame tulbad üksteise kõrvale
```

```
  width = 0.7                          # Tulpade laius
 ) +
 geom_errorbar(                        # Lisame veapiirid
  aes(ymin = Lower_Bound, ymax = Upper_Bound),  # Alumine ja ülemine piir
  position = position_dodge(width = 0.7),      # Nihutame veapiirid vastavalt tulpadele
  width = 0.2,                         # Veapiiride horisontaaljoonte laius
  color = "black"                      # Veapiirid musta värvi
 ) +
 labs(
  x = "Arenev tehnoloogia",            # X-telje silt
  y = "Tehnoloogia kasutajate kaalutud osakaal (%)",  # Y-telje silt
  fill = "Kaalumine:",                 # Legendi pealkiri
  caption = "Mustaga joonistatud lõigud tähistavad veapiire usaldusnivool 95%"  # Pildi all olev selgitus
 ) +
 scale_fill_manual(                    # Kohandame tulpade värve
  values = c("Osakaal audiitorteenuste turust" = "#FF5733",  # Punane
             "Osakaal audiitorettevõtjatest" = "#4CAF50")    # Roheline
 ) +
 scale_y_continuous(labels = scales::percent_format(accuracy = 1)) +  # Y-telg protsentidena
 theme_minimal() +                     # Minimaalne graafiku stiil
 theme(
  text = element_text(family = "Times New Roman"),  # Kirjastiil
  panel.background = element_rect(fill = "white", color = "white"),  # Valge taust
  plot.background = element_rect(fill = "white", color = "white"),   # Valge taust
  axis.text.x = element_text(angle = 0, hjust = 0.5, size = 18),     # X-telje sildid keskele
  # axis.title.x = element_text(size = 16),                          # X-telje pealkirja suurus
  # axis.title.y = element_text(size = 16),                          # Y-telje pealkirja suurus
  axis.title.x = element_blank(),                                    # X-telje pealkirja ei näita
  axis.title.y = element_blank(),                                    # Y-telje pealkirja ei näite
  axis.text.y = element_text(size = 16),                             # Y-telje siltide suurus
  plot.title = element_text(size = 16, hjust = 0.5),                 # Graafiku pealkiri keskele
  plot.caption = element_text(size = 14),                            # Selgituste suurus
  legend.position = "bottom",                                        # Legendi asukoht all
  legend.text = element_text(size = 14),                             # Legendi tekstide suurus
  legend.title = element_text(size = 14),                            # Legendi pealkirja suurus
  panel.grid.major.y = element_line(color = "gray80", size = 0.5),   # Lisame y-telje gridlines
  panel.grid.minor.y = element_blank()                               # Eemaldame väikesed gridlines
 ) +
 scale_y_continuous(
  breaks = seq(0, 1, by = 0.1),                                      # Iga 10% juures y-telje sildid
  labels = scales::percent_format(accuracy = 1)                      # Sildid protsentidena
 )

# Salvestame graafiku pildifailina
ggsave(
 filename = "09_AT_in_use_plot.png",  # Faili nimi
 plot = plot,                         # Graafik
 device = "png",                      # Faili tüüp
 width = 14,                          # Laius tollides
 height = 8,                          # Kõrgus tollides
 units = "in",                        # Mõõtühik
 dpi = 300                            # Resolutsioon
)

# Kuvame graafiku
print(plot)
######################
# Keskkonna ettevalmistus #
######################

# Laadime vajalikud paketid
if (!require(pacman, quietly = TRUE)) install.packages(pacman)

pacman::p_load(
 dplyr,        # Andmete vormindamine
 readr,        # Andmete lugemine failidest
 tidyr         # Tabelite teisendused
)

#############
# Andmete lugemine #
#############

vastused_kodeeritud <- read_excel("vastused_kodeeritud.xlsx")
```

```
## Loon muutujad, mis sisaldavad vastavate andmete veeru numberid "vastused_kodeeritud" failis
AT_in_use_source_ids <- 82:85
weight_ids <- c(132, 136)
suurusklass_id <- 130
confidence_interval_confidence_level <- 0.95

## Kasutades eelnevalt loodud ID-de vektoreid, toon andmed uutesse muutujatesse
data_AT_in_use_source <- vastused_kodeeritud[, AT_in_use_source_ids]
data_weights <- vastused_kodeeritud[, weight_ids]
data_suurusklass <- vastused_kodeeritud[, suurusklass_id]

# Arvutan ka raha kaalu (kaal_turg on juba kahe kaalu korrutis andmetes)
data_weights <- data_weights %>% dplyr::mutate(
  kaal_raha = Kaal_turg / Kaal_innov
)

## Teen koondandmestiku kõikide vajalike muutujatega
data <- cbind(
  data_AT_in_use_source,
  data_weights,
  data_suurusklass
)

# Add labels to the column names
column_labels <- c(
  "5a: Kasutusel väljatöötaja - Ise" = "AEV ise",
  "5b: Kasutusel väljatöötaja - Koostöö võrgustikuga" = "AEV koostöös võrgustiku teise liikmega",
  "5c: Kasutusel väljatöötaja - Koostöö teistega" = "AEV koostöös teise ettevõttega (mitte võrgustikus)",
  "5d: Kasutusel väljatöötaja - Teine" = "Teine ettevõte"
)

# Apply labels to the first 4 columns
colnames(data)[1:4] <- unname(column_labels[colnames(data)[1:4]])

# Veendu, et 8. veerg (suurusklass) on korrektselt tuvastatud
suurusklass_column <- colnames(data)[8]

# Arvuta esimese 4 veeru summa, grupeerituna suurusklassi järgi
grouped_sums <- data %>%
  dplyr::select(1:4, !!sym(suurusklass_column)) %>%  # Vali esimesed 4 veergu ja suurusklassi veerg
  dplyr::group_by(!!sym(suurusklass_column)) %>%      # Grupeeri andmed suurusklassi järgi
  dplyr::summarise(across(1:4, sum, na.rm = TRUE))   # Arvuta igas grupis iga veeru summa, ignoreerides NA väärtusi

# Kuvab gruppide summad
print(grouped_sums)

weight_NR_column <- colnames(data)[6]
print(weight_NR_column)

# Arvuta esimese 4 veeru summa korrutades non-response kaaludega ja grupeerituna suurusklassi järgi
NR_weighted_grouped_sums <- data %>%
  dplyr::mutate(across(1:4, ~ . * !!sym(weight_NR_column))) %>%  # Korruta vastused kaaludega
  dplyr::select(1:4, !!sym(suurusklass_column)) %>%          # Vali muudetud veerud ja suurusklass
  dplyr::group_by(!!sym(suurusklass_column)) %>%             # Grupeeri suurusklassi järgi
  dplyr::summarise(across(1:4, sum, na.rm = TRUE))           # Arvuta igas grupis iga veeru summa

# Kuvab kaalutud summad gruppide lõikes
print(NR_weighted_grouped_sums)

# Calculate NR weighted averages and Wilson confidence intervals for columns 1 to 4
AT_in_use_source_NR_weighted <- data %>%
  dplyr::summarise(across(1:4, ~ {
    # Calculate the weighted sum and n_weighted (sum of weights)
    weighted_sum <- sum(.x * Kaal_innov, na.rm = TRUE)  # Weighted sum
    n_weighted <- sum(Kaal_innov, na.rm = TRUE)         # Effective sample size (sum of weights)
    weighted_proportion <- weighted_sum / 116           # Weighted proportion
    z_value <- qnorm(1 - (1 - confidence_interval_confidence_level) / 2)

    # Calculate Wilson confidence interval using n_weighted
    lower_bound <- (
      (2 * n_weighted * weighted_proportion + z_value^2) -
        z_value * sqrt(max(0, z_value^2 - (1 / n_weighted) + 4 * n_weighted * weighted_proportion * (1 - weighted_proportion)))) /
      (2 * (n_weighted + z_value^2))
    upper_bound <- (
      (2 * n_weighted * weighted_proportion + z_value^2) +
```

```
      z_value * sqrt(max(0, z_value^2 - (1 / n_weighted) + 4 * n_weighted * weighted_proportion * (1 - weighted_proportion)))) /
    (2 * (n_weighted + z_value^2))

  # Return results as a list
  c(proportion_weighted = weighted_proportion, lower_bound = lower_bound, upper_bound = upper_bound)
 }))

# Print the results
print(AT_in_use_source_NR_weighted)

# Save to a CSV file (optional)
write.csv(AT_in_use_source_NR_weighted, "AT_in_use_source_NR_weighted.csv", row.names = FALSE)

#####################

# Arvuta esimese 4 veeru summa korrutades turu kaaludega ja grupeerituna suurusklassi järgi
weight_market_column <- colnames(data)[5]
print(weight_market_column)

market_weighted_grouped_sums <- data %>%
 dplyr::mutate(across(1:4, ~ . * !!sym(weight_market_column))) %>%  # Korruta vastused kaaludega
 dplyr::select(1:4, !!sym(suurusklass_column)) %>%          # Vali muudetud veerud ja suurusklass
 dplyr::group_by(!!sym(suurusklass_column)) %>%             # Grupeeri suurusklassi järgi
 dplyr::summarise(across(1:4, sum, na.rm = TRUE))           # Arvuta igas grupis iga veeru summa

# Kuvab kaalutud summad gruppide lõikes
print(market_weighted_grouped_sums)

# Calculate market weighted averages and Wilson confidence intervals for columns 1 to 4
AT_in_use_source_market_weighted <- data %>%
 dplyr::summarise(across(1:4, ~ {
  # Calculate the weighted sum and n_weighted (sum of weights)
  weighted_sum <- sum(.x * Kaal_turg, na.rm = TRUE)  # Weighted sum
  n_weighted <- sum(Kaal_turg, na.rm = TRUE)         # Effective sample size (sum of weights)
  weighted_proportion <- weighted_sum / 116          # Weighted proportion
  z_value <- qnorm(1 - (1 - confidence_interval_confidence_level) / 2)

  # Calculate Wilson confidence interval using n_weighted
  lower_bound <- (
   (2 * n_weighted * weighted_proportion + z_value^2) -
    z_value * sqrt(max(0, z_value^2 - (1 / n_weighted) + 4 * n_weighted * weighted_proportion * (1 - weighted_proportion)))) /
   (2 * (n_weighted + z_value^2))
  upper_bound <- (
   (2 * n_weighted * weighted_proportion + z_value^2) +
    z_value * sqrt(max(0, z_value^2 - (1 / n_weighted) + 4 * n_weighted * weighted_proportion * (1 - weighted_proportion)))) /
   (2 * (n_weighted + z_value^2))

  # Return results as a list
  c(proportion_weighted = weighted_proportion, lower_bound = lower_bound, upper_bound = upper_bound)
 }))

# Print the results
print(AT_in_use_source_market_weighted)

# Save to a CSV file (optional)
write.csv(AT_in_use_source_market_weighted, "AT_in_use_source_market_weighted.csv", row.names = FALSE)
#####################
# Keskkonna ettevalmistus #
#####################

# Laadime vajalikud R-paketid
if (!require(pacman, quietly = TRUE)) install.packages(pacman)

pacman::p_load(
 ggplot2,       # Graafika genereerimine
 readr,         # Andmete lugemine failidest
 tidyr,         # Tabelite teisendused (pikk vs lai formaat)
 dplyr,         # Andmete töötlemine
 scales,        # Telgede siltide vormindamine (nt protsendid)
 stringr,       # Tekstidega töötlemine
 pacman
)

#############
# Andmete lugemine #
```

```
#############

# Määrame andmefailide asukohad ja loeme need sisse
AT_in_use_source_market_weighted_path <- "AT_in_use_source_market_weighted.csv"  # Turu kaalutud andmed
market_weighted_data <- readr::read_csv(AT_in_use_source_market_weighted_path, show_col_types = FALSE)

AT_in_use_source_NR_weighted_path <- "AT_in_use_source_NR_weighted.csv"  # Vastamisest kõrvale jäänud ettevõtete kaalutud
andmed
NR_weighted_data <- readr::read_csv(AT_in_use_source_NR_weighted_path, show_col_types = FALSE)

# Lisame veeru "Source" (allikas), mis näitab, kummast andmestikust rida pärineb
market_weighted_data <- market_weighted_data %>%
  mutate(Source = "Osakaal audiitorteenuste turust")  # Turu kaalutud osakaal

NR_weighted_data <- NR_weighted_data %>%
  mutate(Source = "Osakaal audiitorettevõtjatest")  # Vastamisest kõrvale jäänud osakaal

# Kombineerime mõlemad andmestikud üheks suureks andmetabeliks
combined_data <- bind_rows(market_weighted_data, NR_weighted_data)

# Teisendame andmed laia formaadist pikkvormi
combined_data_long <- combined_data %>%
  pivot_longer(
    cols = -Source,          # Kõik veerud peale "Source" teisendatakse ridadeks
    names_to = "AT_valjatootaja", # Tehnoloogia nimed tulevad veerunimedest
    values_to = "Value"      # Vastavad väärtused tulevad tabeli sisust
  ) %>%
  # Lisame veeru "Metric_Type", et eristada kaalutud keskmisi, alumisi ja ülemisi veapiire
  mutate(Metric_Type = rep(c("Weighted_Mean", "Lower_Bound", "Upper_Bound"), each = 4, times = 2)) %>%
  # Teisendame veerud uuesti laia formaati
  pivot_wider(names_from = Metric_Type, values_from = Value)

# Järjestame AT väljatöötajad turu kaalutud osakaalu alusel kahanevalt
sorted_AT_valjatootajad <- combined_data_long %>%
  filter(Source == "Osakaal audiitorteenuste turust") %>%  # Filtreerime turu andmed
  arrange(desc(Weighted_Mean)) %>%                    # Järjestame kahanevalt
  pull(AT_valjatootaja)                               # Saame nimekirja tehnoloogiatest

# Rakendame sorteeritud järjekorra tehnoloogia veerule
combined_data_long <- combined_data_long %>%
  mutate(AT_valjatootaja = factor(AT_valjatootaja, levels = sorted_AT_valjatootajad)) %>%
  mutate(Source = factor(Source, levels = c("Osakaal audiitorteenuste turust", "Osakaal audiitorettevõtjatest")))  # Turu osakaal vasakule

##################
# Graafiku loomine #
##################

# Filter out rows where Weighted_Mean is 0
filtered_data <- combined_data_long %>%
  filter(Weighted_Mean > 0)  # Exclude rows with 0-valued bars

# Create the plot
plot <- ggplot(filtered_data, aes(x = AT_valjatootaja, y = Weighted_Mean, fill = Source)) +
  geom_bar(
    stat = "identity",                    # Kasutame staatilist identiteeti tulbadiagrammi jaoks
    position = position_dodge(width = 0.7),  # Nihutame tulbad üksteise kõrvale
    width = 0.7                           # Tulpade laius
  ) +
  geom_errorbar(                          # Lisame veapiirid
    aes(ymin = Lower_Bound, ymax = Upper_Bound),  # Alumine ja ülemine piir
    position = position_dodge(width = 0.7),      # Nihutame veapiirid vastavalt tulpadele
    width = 0.2,                          # Veapiiride horisontaaljoonte laius
    color = "black"                         # Veapiirid musta värvi
  ) +
  labs(
    x = "Areneva tehnoloogia väljatöötaja",              # X-telje silt
    y = "Tehnoloogia väljatöötaja kaalutud osakaal (%)",  # Y-telje silt
    fill = "Kaalumine:",                  # Legendi pealkiri
    caption = "Mustaga joonistatud lõigud tähistavad veapiire usaldusnivool 95%"  # Pildi all olev selgitus
  ) +
  scale_fill_manual(                      # Kohandame tulpade värve
    values = c("Osakaal audiitorteenuste turust" = "#FF5733",  # Punane
               "Osakaal audiitorettevõtjatest" = "#4CAF50")    # Roheline
  ) +
  scale_y_continuous(labels = scales::percent_format(accuracy = 1)) +  # Y-telg protsentidena
```

```
  theme_minimal() +                    # Minimaalne graafiku stiil
  theme(
    text = element_text(family = "Times New Roman"),  # Kirjastiil
    panel.background = element_rect(fill = "white", color = "white"),  # Valge taust
    plot.background = element_rect(fill = "white", color = "white"),   # Valge taust
    axis.text.x = element_text(angle = 0, hjust = 0.5, size = 20),     # X-telje sildid keskele
    # axis.title.x = element_text(size = 18),                          # X-telje pealkirja suurus
    axis.title.x = element_blank(),
    axis.title.y = element_blank(),
    # axis.title.y = element_text(size = 18),                          # Y-telje pealkirja suurus
    axis.text.y = element_text(size = 16),                            # Y-telje siltide suurus
    plot.title = element_text(size = 18, hjust = 0.5),                # Graafiku pealkiri keskele
    plot.caption = element_text(size = 16),                           # Selgituste suurus
    legend.position = "bottom",                                       # Legendi asukoht all
    legend.text = element_text(size = 16),                            # Legendi tekstide suurus
    legend.title = element_text(size = 16),                           # Legendi pealkirja suurus
    panel.grid.major.y = element_line(color = "gray80", size = 0.5),   # Lisame y-telje gridlines
    panel.grid.minor.y = element_blank()                              # Eemaldame väikesed gridlines
  ) +
  scale_y_continuous(
    breaks = seq(0, 1, by = 0.1),                                     # Iga 10% juures y-telje sildid
    labels = scales::percent_format(accuracy = 1)                     # Sildid protsentidena
  )

# Salvestame graafiku pildifailina
ggsave(
  filename = "11_AT_in_use_source_plot.png",  # Faili nimi
  plot = plot,                                # Graafik
  device = "png",                             # Faili tüüp
  width = 14,                                 # Laius tollides
  height = 6,                                 # Kõrgus tollides
  units = "in",                               # Mõõtühik
  dpi = 300                                   # Resolutsioon
)

# Kuvame graafiku
print(plot)
######################
# Keskkonna ettevalmistus #
######################

# Laadime vajalikud paketid
if (!require(pacman, quietly = TRUE)) install.packages(pacman)

pacman::p_load(
  dplyr,         # Andmete vormindamine ja töötlemine
  purrr,         # Iteratsioonide ja listide töötlemine
  readr,         # Andmete lugemine CSV-failidest
  tidyr          # Tabelite teisendused (lai ja pikk formaat)
)

#############
# Andmete lugemine #
#############

# Defineerime failide teed ja loeme vastavad andmed CSV-failidest
responses_path <- "responses.csv"  # Vastuste fail
responses <- readr::read_csv(responses_path, show_col_types = FALSE)

respondent_key_path <- "respondent_key.csv"  # Vastajate võtmefail (nt suurusklassid)
respondent_key <- readr::read_csv(respondent_key_path, show_col_types = FALSE)

weights_path <- "weights.csv"  # Kaalude fail
weights <- readr::read_csv(weights_path, show_col_types = FALSE)

question_key_path <- "question_key.csv"  # Küsimuste võtmefail
question_key <- readr::read_csv(question_key_path, show_col_types = FALSE)

#############
# Funktsioonid #
#############

# Funktsioon bootstrap'i proovide loomiseks
get_bootstrap_samples <- function(codes, n_respondents, n_samples) {
  # Loob n_samples bootstrap-proovi, kus igas proovis on n_respondents vastust
```

```
  bootstrap_samples <- purrr::map(
    .x = 1:n_samples,  # Iteratsioonide arv
    .f = \(x) sample(codes, n_respondents, replace = TRUE))  # Proovide loomine asendamisega

  return(bootstrap_samples)  # Tagastab bootstrap-proovid
}

############################
# Filtreeri olulised küsimused #
############################

# Määrame innovatsiooni tõenäosuste küsimuste ID-d
innovation_probabilities_question_ids <- 45:55

# Filtreerime vastused, mis vastavad küsimuste ID-dele
innovation_probabilities_responses <- responses |>
  dplyr::filter(question_id %in% innovation_probabilities_question_ids)

########################################
#NON-RESPONSE KAALUGA
########################################
# Bootstrap'i veapiiride arvutamine #
########################################

# Määrame usaldusnivoo ja bootstrap'i simulatsioonide arvu
confidence_interval_confidence_level <- 0.95
n_bootstrap_simulations <- 10^4  # 10 000 simulatsiooni

# Bootstrap'i sisendandmete ettevalmistamine
bootstrap_input <- innovation_probabilities_responses |>
  dplyr::left_join(
    respondent_key,  # Lisame vastaja võtme
    join_by(respondent_id)) |>
  dplyr::group_by(respondent_size_id, question_id) |>
  dplyr::summarise(
    codes = list(code),  # Vastuste koodid listina
    n_respondents = n(),  # Vastajate arv igas grupis
    .groups = "drop") |>
  dplyr::left_join(
    weights,  # Lisame kaalud
    join_by(respondent_size_id)) |>
  dplyr::group_by(question_id) |>
  dplyr::mutate(
    # weight_multiplier = weight_post_stratification / sum(weight_post_stratification)
    weight_multiplier = weight_post_stratification
    ) |>
  dplyr::arrange(question_id, respondent_size_id) |>
  dplyr::select(
    question_id,
    respondent_size_id,
    codes,
    n_respondents,
    weight_multiplier)

# Lisame iga rea jaoks bootstrap-proovid
bootstrap_samples <- bootstrap_input |>
  dplyr::rowwise() |>
  dplyr::mutate(
    bootstrap_samples = list(
      get_bootstrap_samples(
        codes,  # Vastuste koodid
        n_respondents,  # Vastajate arv
        n_bootstrap_simulations)),  # Simulatsioonide arv
    mean_raw = list(
      purrr::map(
        bootstrap_samples,
        \(x) mean(x))))  # Arvutame iga proovi keskmise


# Loome tühja tabeli bootstrap'i keskmiste jaoks
bootstrap_mean_weighted <- tibble::tibble()

# Läbime iga küsimuse ID
for (question_id in unique(bootstrap_samples$question_id)) {
```

```
  bootstrap_values_by_question_id <- bootstrap_samples |>
    dplyr::filter(question_id == !!question_id)

  # Raw (toore) keskmised bootstrap-proovidest
  mean_raw <- bootstrap_values_by_question_id |>
    dplyr::pull(mean_raw) |>
    purrr::map(\(listed_values) unlist(listed_values))

  weight_multipliers <- bootstrap_values_by_question_id$weight_multiplier
  n_respondents <- bootstrap_values_by_question_id$n_respondents

  # Läbime kõik simulatsioonid ja arvutame kaalutud keskmised
  mean_weighted_by_simulation <- c()
  for (i_simulation in seq.int(n_bootstrap_simulations)) {
    mean_raw_by_simulation <- purrr::map_dbl(mean_raw, \(x) x[i_simulation])
    mean_weighted_by_simulation <- c(mean_weighted_by_simulation, sum(weight_multipliers *
    mean_raw_by_simulation * n_respondents)/116)
  }
  print(mean_weighted_by_simulation)

  # Lisame kaalutud keskmised iga küsimuse kohta
  bootstrap_mean_weighted <- bootstrap_mean_weighted |>
    dplyr::bind_rows(
      tibble::tibble(
        question_id = !!question_id,
        mean_weighted = list(mean_weighted_by_simulation)))
}
print(bootstrap_mean_weighted)
# Bootstrap'i veapiiride arvutamine
bootstrap_confidence_intervals <- bootstrap_mean_weighted |>
  dplyr::rowwise() |>
  dplyr::mutate(
    confidence_level = confidence_interval_confidence_level,  # Usaldusnivoo
    confidence_interval_start = quantile(
      unlist(mean_weighted),
      probs = (1 - confidence_level) / 2, names = FALSE),  # Alumine piir
    confidence_interval_end = quantile(
      unlist(mean_weighted),
      probs = 1 - ((1 - confidence_level) / 2), names = FALSE)) |>  # Ülemine piir
  dplyr::select(-mean_weighted)

################################
# Innovatsiooni tõenäosuste arvutamine #
################################

# Arvutame toored keskmised suurusegrupi ja küsimuse alusel
innovation_probabilities_by_size <- innovation_probabilities_responses |>
  dplyr::left_join(
    respondent_key,  # Lisame vastaja võtme
    join_by(respondent_id)) |>
  dplyr::group_by(respondent_size_id, question_id) |>
  dplyr::summarise(
    mean_raw = mean(code),  # Toorne keskmine
    n_respondents = n(),  # Vastajate arv
    .groups = "drop")

# Lisame kaalud ja arvutame tegelikud kaalutud keskmised
innovation_probabilities <- innovation_probabilities_by_size |>
  dplyr::left_join(
    weights,
    join_by(respondent_size_id)) |>
  dplyr::group_by(question_id) |>
  dplyr::mutate(
    # weight_multiplier = weight_post_stratification / sum(weight_post_stratification)
    weight_multiplier = weight_post_stratification
    ) |>
  dplyr::summarise(
    mean_weighted = sum(mean_raw * weight_multiplier * n_respondents)/116) |>
  dplyr::left_join(
    bootstrap_confidence_intervals,  # Lisame bootstrap'i veapiirid
    join_by(question_id)) |>
  dplyr::left_join(
    question_key,  # Lisame küsimuste nimed
    join_by(question_id)) |>
  dplyr::select(
```

```
    question_id,
    question_name = name,  # Küsimuse nimi
    mean = mean_weighted,  # Kaalutud keskmine
    confidence_level,
    confidence_interval_start,
    confidence_interval_end)

#############
# Andmete salvestamine #
#############

# Salvestame lõpliku andmestiku CSV-failina
write.csv(innovation_probabilities, "innovation_probabilities_NR.csv", row.names = FALSE)



##########################################
#TURU KAALUGA
########################################
# Bootstrap'i veapiiride arvutamine #
########################################

# Bootstrap'i sisendandmete ettevalmistamine
bootstrap_input_market <- innovation_probabilities_responses |>
  dplyr::left_join(
    respondent_key,  # Lisame vastaja võtme
    join_by(respondent_id)) |>
  dplyr::group_by(respondent_size_id, question_id) |>
  dplyr::summarise(
    codes = list(code),  # Vastuste koodid listina
    n_respondents = n(),  # Vastajate arv igas grupis
    .groups = "drop") |>
  dplyr::left_join(
    weights,  # Lisame kaalud
    join_by(respondent_size_id)) |>
  dplyr::group_by(question_id) |>
  dplyr::mutate(
    # weight_multiplier = weight_post_stratification / sum(weight_post_stratification)
    weight_multiplier = weight_post_stratification * weight_revenue
  ) |>
  dplyr::arrange(question_id, respondent_size_id) |>
  dplyr::select(
    question_id,
    respondent_size_id,
    codes,
    n_respondents,
    weight_multiplier)

# Lisame iga rea jaoks bootstrap-proovid
bootstrap_samples_market <- bootstrap_input_market |>
  dplyr::rowwise() |>
  dplyr::mutate(
    bootstrap_samples_market = list(
      get_bootstrap_samples(
        codes,  # Vastuste koodid
        n_respondents,  # Vastajate arv
        n_bootstrap_simulations)),  # Simulatsioonide arv
    mean_raw = list(
      purrr::map(
        bootstrap_samples_market,
        \(x) mean(x))))  # Arvutame iga proovi keskmise


# Loome tühja tabeli bootstrap'i keskmiste jaoks - TURG
bootstrap_mean_weighted_market <- tibble::tibble()

# Läbime iga küsimuse ID
for (question_id in unique(bootstrap_samples_market$question_id)) {
  bootstrap_values_by_question_id_market <- bootstrap_samples_market |>
    dplyr::filter(question_id == !!question_id)

  # Raw (toore) keskmised bootstrap-proovidest
  mean_raw <- bootstrap_values_by_question_id_market |>
    dplyr::pull(mean_raw) |>
    purrr::map(\(listed_values) unlist(listed_values))
```

```
  weight_multipliers_market <- bootstrap_values_by_question_id_market$weight_multiplier
  n_respondents <- bootstrap_values_by_question_id_market$n_respondents

  # Läbime kõik simulatsioonid ja arvutame kaalutud keskmised
  mean_weighted_by_simulation_market <- c()
  for (i_simulation in seq.int(n_bootstrap_simulations)) {
    mean_raw_by_simulation_market <- purrr::map_dbl(mean_raw, \(x) x[i_simulation])
    mean_weighted_by_simulation_market <- c(mean_weighted_by_simulation_market, sum(weight_multipliers_market *
                                  mean_raw_by_simulation_market * n_respondents)/116)
  }
  print(mean_weighted_by_simulation_market)

  # Lisame kaalutud keskmised iga küsimuse kohta
  bootstrap_mean_weighted_market <- bootstrap_mean_weighted_market |>
    dplyr::bind_rows(
      tibble::tibble(
        question_id = !!question_id,
        mean_weighted_market = list(mean_weighted_by_simulation_market)))
}
# print(bootstrap_mean_weighted_market)
# Bootstrap'i veapiiride arvutamine
bootstrap_confidence_intervals_market <- bootstrap_mean_weighted_market |>
  dplyr::rowwise() |>
  dplyr::mutate(
    confidence_level = confidence_interval_confidence_level,  # Usaldusnivoo
    confidence_interval_start = quantile(
      unlist(mean_weighted_market),
      probs = (1 - confidence_level) / 2, names = FALSE),  # Alumine piir
    confidence_interval_end = quantile(
      unlist(mean_weighted_market),
      probs = 1 - ((1 - confidence_level) / 2), names = FALSE)) |>  # Ülemine piir
  dplyr::select(-mean_weighted_market)

################################
# Innovatsiooni tõenäosuste arvutamine #
################################

# Lisame kaalud ja arvutame tegelikud kaalutud keskmised
innovation_probabilities_market <- innovation_probabilities_by_size |>
  dplyr::left_join(
    weights,
    join_by(respondent_size_id)) |>
  dplyr::group_by(question_id) |>
  dplyr::mutate(
    # weight_multiplier = weight_post_stratification / sum(weight_post_stratification)
    weight_multiplier = weight_post_stratification * weight_revenue
  ) |>
  dplyr::summarise(
    mean_weighted = sum(mean_raw * weight_multiplier * n_respondents)/116) |>
  dplyr::left_join(
    bootstrap_confidence_intervals_market,  # Lisame bootstrap'i veapiirid
    join_by(question_id)) |>
  dplyr::left_join(
    question_key,  # Lisame küsimuste nimed
    join_by(question_id)) |>
  dplyr::select(
    question_id,
    question_name = name,  # Küsimuse nimi
    mean = mean_weighted,  # Kaalutud keskmine
    confidence_level,
    confidence_interval_start,
    confidence_interval_end)

#############
# Andmete salvestamine #
#############

# Salvestame lõpliku andmestiku CSV-failina
write.csv(innovation_probabilities_market, "innovation_probabilities_market.csv", row.names = FALSE)

####################
# Keskkonna ettevalmistus #
####################
```

```
# Laadime vajalikud R-paketid
if (!require(pacman, quietly = TRUE)) install.packages(pacman)

pacman::p_load(
  dplyr,        # Andmete töötlemine ja vormindamine
  ggplot2,      # Graafika loomine
  readr,        # Andmete lugemine CSV-failidest
  scales,       # Telgede siltide vormindamine
  stringr       # Teksti töötlemise operatsioonid
)


#############
# Andmete lugemine #
#############

# Loeme kaks failid: NR kaalutud ja turu kaalutud andmed
innovation_probabilities_NR_path <- "innovation_probabilities_NR.csv"  # NR kaalutud fail
innovation_probabilities_market_path <- "innovation_probabilities_market.csv"  # Turu kaalutud fail

innovation_probabilities_NR <- readr::read_csv(innovation_probabilities_NR_path, show_col_types = FALSE)
innovation_probabilities_market <- readr::read_csv(innovation_probabilities_market_path, show_col_types = FALSE)


###############
# Andmete ühendamine #
###############

# Lisame veeru, mis näitab, millisest allikast andmed pärinevad
innovation_probabilities_NR <- innovation_probabilities_NR |>
  dplyr::mutate(source = "Audiitorettevõtjate keskmine")
innovation_probabilities_market <- innovation_probabilities_market |>
  dplyr::mutate(source = "Audiitorteenuste turu keskmine")

# Ühendame kaks andmestikku üheks suureks tabeliks
innovation_probabilities <- dplyr::bind_rows(
  innovation_probabilities_NR,
  innovation_probabilities_market
)


###############
# Graafiku andmed #
###############

plot_data <- innovation_probabilities

# Defineerime tehnoloogiate nimed
technology_labels <- c(
  "tõenäosus_ai" = "Tehisintellekt (AI)",
  "tõenäosus_nlp" = "Loodusliku keele\ntöötlemine (NLP)",
  "tõenäosus_rpa" = "Robootiline protsesside\nautomatiseerimine (RPA)",
  "tõenäosus_iot" = "Asjade internet (IoT)",
  "tõenäosus_hkrüpt" = "Homomorfne krüpteerimine\n(HKrüpt.)",
  "tõenäosus_api" = "Rakendusliidesed (API)",
  "tõenäosus_standard" = "Andmete standardiseerimine\n(Standard.)",
  "tõenäosus_ada" = "Täiustatud andme-\nanalüütika (ADA)",
  "tõenäosus_plokiahel" = "Plokiahela tehnoloogia\n(Blockchain)",
  "tõenäosus_protsessikaeve" = "Protsessikaevandamine\n(PMining)",
  "tõenäosus_arvutinägemine" = "Optiline märgituvastus (OCR)"
)

# Defineerime tõenäosuse sildid
probability_labels <- c(
  "1 Väga eba-\ntõenäoline",
  "2 Eba-\ntõenäoline",
  "3 Ei oska\nöelda",
  "4 Tõenäoline",
  "5 Väga\ntõenäoline"
)


###############
# Graafiku loomine #
###############
```

```
plot <- plot_data |>
  # Loome graafiku aluse
  ggplot2::ggplot(
    aes(
      x = `mean`,  # Tõenäosuse keskmine väärtus
      y = reorder(question_name, question_id),  # Tehnoloogia nimi
      color = source)) +  # Kasutame allikat värvidena

  # Lisame punktid tõenäosuste jaoks
  geom_point(
    size = 3) +

  # Lisame horisontaalsed veapiirid
  geom_errorbarh(
    aes(xmin = confidence_interval_start, xmax = confidence_interval_end),
    height = 0.2) +

  # Lisame tõenäosuste väärtused punktide kohale
  geom_text(
    aes(label = round(mean, 2)),
    nudge_y = 0.3,
    size = 5) +

  # X-telje skaala
  scale_x_continuous(
    limits = c(0.5, 5.5),  # Tõenäosuse ulatus (1-5)
    expand = c(0, 0),
    labels = probability_labels,  # Tõenäosuse tasemed
    breaks = c(1, 2, 3, 4, 5)) +

  # Y-telje skaala
  scale_y_discrete(labels = technology_labels) +

  # Kohandame värvide legendi
  scale_color_manual(
    name = "Kaalumine:",
    values = c("Audiitorettevõtjate keskmine" = "#4CAF50",  # Roheline värv NR jaoks
               "Audiitorteenuste turu keskmine" = "#FF5733")) +  # Oranž värv turu jaoks

  # Lisame graafiku sildid ja selgituse
  labs(
    x = "Kasutusele võtmise tõenäosus järgmise nelja aasta jooksul (1-5)",
    y = "Arenev tehnoloogia",
    caption = stringr::str_glue("Lõigud tähistavad veapiire usaldusnivool {unique(plot_data$confidence_level) * 100}%")) +

  # Graafiku kujundus
  theme_minimal() +

  theme(
    text = element_text(family = "Times New Roman"),  # Kirjastiil
    plot.margin = margin(10, 10, 10, 10),  # Graafiku ääred
    # axis.title.x = element_text(size = 24, margin = margin(t = 30)),
    # axis.title.y = element_text(size = 24, margin = margin(r = 30)),
    axis.title.x = element_blank(),
    axis.title.y = element_blank(),
    axis.text = element_text(size = 18),
    axis.text.x = element_text(margin = margin(t = 10)),
    plot.caption.position = "plot",
    plot.caption = element_text(size = 18, hjust = 0),
    legend.position = "bottom",  # Legendi asukoht
    legend.text = element_text(size = 14),
    legend.title = element_text(size = 16),
    # panel.grid = element_line(colour = "black")
    legend.spacing = unit(0.0, "points")
    )


###############
# Graafiku salvestamine #
###############

# Salvestame graafiku pildifailina
ggplot2::ggsave(
  filename = "13_innovation_probabilities_comparison.png",
```

```
  plot = plot,
  device = "png",
  width = 4000,
  height = 2750,
  units = "px",
  bg = "white")

#####################
# Keskkonna ettevalmistus #
#####################

# Laadime vajalikud paketid
if (!require(pacman, quietly = TRUE)) install.packages(pacman)

pacman::p_load(
  dplyr,         # Andmete vormindamine ja töötlemine
  purrr,         # Iteratsioonide ja listide töötlemine
  readr,         # Andmete lugemine CSV-failidest
  tidyr          # Tabelite teisendused (lai ja pikk formaat)
)

#############
# Andmete lugemine #
#############

# Defineerime failide teed ja loeme vastavad andmed CSV-failidest
responses_path <- "responses.xlsx"
responses <- read_excel(responses_path)

respondent_key_path <- "respondent_key.csv"
respondent_key <- readr::read_csv(respondent_key_path, show_col_types = FALSE)

weights_path <- "weights.csv"
weights <- readr::read_csv(weights_path, show_col_types = FALSE)

question_key_path <- "question_key.xlsx"
question_key <- read_excel(question_key_path)

#############
# Funktsioonid #
#############

# Funktsioon bootstrap'i proovide loomiseks
get_bootstrap_samples <- function(codes, n_respondents, n_samples) {
  # Loob n_samples bootstrap-proovi, kus igas proovis on n_respondents vastust

  bootstrap_samples <- purrr::map(
    .x = 1:n_samples,  # Iteratsioonide arv
    .f = \(x) sample(codes, n_respondents, replace = TRUE))  # Proovide loomine asendamisega

  return(bootstrap_samples)  # Tagastab bootstrap-proovid
}

############################
# Filtreeri olulised küsimused #
############################

# Määrame innovatsiooni tõenäosuste küsimuste ID-d
AT_benefit_question_ids <- 12:18

# Filtreerime vastused, mis vastavad küsimuste ID-dele
AT_benefit_responses <- responses |>
  dplyr::filter(question_id %in% AT_benefit_question_ids)

########################################
#NON-RESPONSE KAALUGA
########################################
# Bootstrap'i veapiiride arvutamine #
########################################

# Määrame usaldusnivoo ja bootstrap'i simulatsioonide arvu
confidence_interval_confidence_level <- 0.95
n_bootstrap_simulations <- 10^4  # 10 000 simulatsiooni

# Bootstrap'i sisendandmete ettevalmistamine
```

```
bootstrap_input <- AT_benefit_responses |>
  dplyr::left_join(
    respondent_key,  # Lisame vastaja võtme
    join_by(respondent_id)) |>
  dplyr::group_by(respondent_size_id, question_id) |>
  dplyr::summarise(
    codes = list(code),  # Vastuste koodid listina
    n_respondents = n(),  # Vastajate arv igas grupis
    .groups = "drop") |>
  dplyr::left_join(
    weights,  # Lisame kaalud
    join_by(respondent_size_id)) |>
  dplyr::group_by(question_id) |>
  dplyr::mutate(
    # weight_multiplier = weight_post_stratification / sum(weight_post_stratification)
    weight_multiplier = weight_post_stratification
  ) |>
  dplyr::arrange(question_id, respondent_size_id) |>
  dplyr::select(
    question_id,
    respondent_size_id,
    codes,
    n_respondents,
    weight_multiplier)

# Lisame iga rea jaoks bootstrap-proovid
bootstrap_samples <- bootstrap_input |>
  dplyr::rowwise() |>
  dplyr::mutate(
    bootstrap_samples = list(
      get_bootstrap_samples(
        codes,  # Vastuste koodid
        n_respondents,  # Vastajate arv
        n_bootstrap_simulations)),  # Simulatsioonide arv
    mean_raw = list(
      purrr::map(
        bootstrap_samples,
        \(x) mean(x))))  # Arvutame iga proovi keskmise


# Loome tühja tabeli bootstrap'i keskmiste jaoks
bootstrap_mean_weighted_NR <- tibble::tibble()

# Läbime iga küsimuse ID
for (question_id in unique(bootstrap_samples$question_id)) {
  bootstrap_values_by_question_id <- bootstrap_samples |>
    dplyr::filter(question_id == !!question_id)

  # Raw (toore) keskmised bootstrap-proovidest
  mean_raw <- bootstrap_values_by_question_id |>
    dplyr::pull(mean_raw) |>
    purrr::map(\(listed_values) unlist(listed_values))

  weight_multipliers <- bootstrap_values_by_question_id$weight_multiplier
  n_respondents <- bootstrap_values_by_question_id$n_respondents

  # Läbime kõik simulatsioonid ja arvutame kaalutud keskmised
  mean_weighted_by_simulation <- c()
  for (i_simulation in seq.int(n_bootstrap_simulations)) {
    mean_raw_by_simulation <- purrr::map_dbl(mean_raw, \(x) x[i_simulation])
    mean_weighted_by_simulation <- c(mean_weighted_by_simulation, sum(weight_multipliers *
                                                mean_raw_by_simulation * n_respondents)/116)
  }
  print(mean_weighted_by_simulation)

  # Lisame kaalutud keskmised iga küsimuse kohta
  bootstrap_mean_weighted_NR <- bootstrap_mean_weighted_NR |>
    dplyr::bind_rows(
      tibble::tibble(
        question_id = !!question_id,
        mean_weighted = list(mean_weighted_by_simulation)))
}
print(bootstrap_mean_weighted_NR)
# Bootstrap'i veapiiride arvutamine
bootstrap_confidence_intervals_NR <- bootstrap_mean_weighted_NR |>
```

```
  dplyr::rowwise() |>
  dplyr::mutate(
    confidence_level = confidence_interval_confidence_level,  # Usaldusnivoo
    confidence_interval_start = quantile(
      unlist(mean_weighted),
      probs = (1 - confidence_level) / 2, names = FALSE),  # Alumine piir
    confidence_interval_end = quantile(
      unlist(mean_weighted),
      probs = 1 - ((1 - confidence_level) / 2), names = FALSE)) |>  # Ülemine piir
  dplyr::select(-mean_weighted)


################################
# AT-st nähtava kasu arvutamine #
################################

# Arvutame toored keskmised suurusegrupi ja küsimuse alusel
AT_benefit_by_size <- AT_benefit_responses |>
  dplyr::left_join(
    respondent_key,  # Lisame vastaja võtme
    join_by(respondent_id)) |>
  dplyr::group_by(respondent_size_id, question_id) |>
  dplyr::summarise(
    mean_raw = mean(code),  # Toorne keskmine
    n_respondents = n(),  # Vastajate arv
    .groups = "drop")

# Lisame kaalud ja arvutame tegelikud kaalutud keskmised
AT_benefits <- AT_benefit_by_size |>
  dplyr::left_join(
    weights,
    join_by(respondent_size_id)) |>
  dplyr::group_by(question_id) |>
  dplyr::mutate(
    # weight_multiplier = weight_post_stratification / sum(weight_post_stratification)
    weight_multiplier = weight_post_stratification
  ) |>
  dplyr::summarise(
    mean_weighted = sum(mean_raw * weight_multiplier * n_respondents)/116) |>
  dplyr::left_join(
    bootstrap_confidence_intervals_NR,  # Lisame bootstrap'i veapiirid
    join_by(question_id)) |>
  dplyr::left_join(
    question_key,  # Lisame küsimuste nimed
    join_by(question_id)) |>
  dplyr::select(
    question_id,
    question_name = name,  # Küsimuse nimi
    mean = mean_weighted,  # Kaalutud keskmine
    confidence_level,
    confidence_interval_start,
    confidence_interval_end)

#############
# Andmete salvestamine #
#############

# Salvestame lõpliku andmestiku CSV-failina
write.csv(AT_benefits, "AT_benefit_NR.csv", row.names = FALSE)



############################################
#TURU KAALUGA
########################################
# Bootstrap'i veapiiride arvutamine #
########################################

# Bootstrap'i sisendandmete ettevalmistamine
bootstrap_input_market <- AT_benefit_responses |>
  dplyr::left_join(
    respondent_key,  # Lisame vastaja võtme
    join_by(respondent_id)) |>
  dplyr::group_by(respondent_size_id, question_id) |>
  dplyr::summarise(
```

```
    codes = list(code),  # Vastuste koodid listina
    n_respondents = n(),  # Vastajate arv igas grupis
    .groups = "drop") |>
  dplyr::left_join(
    weights,  # Lisame kaalud
    join_by(respondent_size_id)) |>
  dplyr::group_by(question_id) |>
  dplyr::mutate(
    # weight_multiplier = weight_post_stratification / sum(weight_post_stratification)
    weight_multiplier = weight_post_stratification * weight_revenue
  ) |>
  dplyr::arrange(question_id, respondent_size_id) |>
  dplyr::select(
    question_id,
    respondent_size_id,
    codes,
    n_respondents,
    weight_multiplier)

# Lisame iga rea jaoks bootstrap-proovid
bootstrap_samples_market <- bootstrap_input_market |>
  dplyr::rowwise() |>
  dplyr::mutate(
    bootstrap_samples_market = list(
      get_bootstrap_samples(
        codes,  # Vastuste koodid
        n_respondents,  # Vastajate arv
        n_bootstrap_simulations)),  # Simulatsioonide arv
    mean_raw = list(
      purrr::map(
        bootstrap_samples_market,
        \(x) mean(x))))  # Arvutame iga proovi keskmise


# Loome tühja tabeli bootstrap'i keskmiste jaoks - TURG
bootstrap_mean_weighted_market <- tibble::tibble()

# Läbime iga küsimuse ID
for (question_id in unique(bootstrap_samples_market$question_id)) {
  bootstrap_values_by_question_id_market <- bootstrap_samples_market |>
    dplyr::filter(question_id == !!question_id)

  # Raw (toore) keskmised bootstrap-proovidest
  mean_raw <- bootstrap_values_by_question_id_market |>
    dplyr::pull(mean_raw) |>
    purrr::map(\(listed_values) unlist(listed_values))

  weight_multipliers_market <- bootstrap_values_by_question_id_market$weight_multiplier
  n_respondents <- bootstrap_values_by_question_id_market$n_respondents

  # Läbime kõik simulatsioonid ja arvutame kaalutud keskmised
  mean_weighted_by_simulation_market <- c()
  for (i_simulation in seq.int(n_bootstrap_simulations)) {
    mean_raw_by_simulation_market <- purrr::map_dbl(mean_raw, \(x) x[i_simulation])
    mean_weighted_by_simulation_market <- c(mean_weighted_by_simulation_market, sum(weight_multipliers_market *
                                                                                      mean_raw_by_simulation_market * n_respondents)/116)
  }
  print(mean_weighted_by_simulation_market)

  # Lisame kaalutud keskmised iga küsimuse kohta
  bootstrap_mean_weighted_market <- bootstrap_mean_weighted_market |>
    dplyr::bind_rows(
      tibble::tibble(
        question_id = !!question_id,
        mean_weighted_market = list(mean_weighted_by_simulation_market)))
}
# print(bootstrap_mean_weighted_market)
# Bootstrap'i veapiiride arvutamine
bootstrap_confidence_intervals_market <- bootstrap_mean_weighted_market |>
  dplyr::rowwise() |>
  dplyr::mutate(
    confidence_level = confidence_interval_confidence_level,  # Usaldusnivoo
    confidence_interval_start = quantile(
      unlist(mean_weighted_market),
      probs = (1 - confidence_level) / 2, names = FALSE),  # Alumine piir
```

```
    confidence_interval_end = quantile(
      unlist(mean_weighted_market),
      probs = 1 - ((1 - confidence_level) / 2), names = FALSE)) |>  # Ülemine piir
  dplyr::select(-mean_weighted_market)

################################
# AT-st nähtava kasu arvutamine - TURULE KAALUTUD #
################################

# Lisame kaalud ja arvutame tegelikud kaalutud keskmised
AT_benefits_market <- AT_benefit_by_size |>
  dplyr::left_join(
    weights,
    join_by(respondent_size_id)) |>
  dplyr::group_by(question_id) |>
  dplyr::mutate(
    # weight_multiplier = weight_post_stratification / sum(weight_post_stratification)
    weight_multiplier = weight_post_stratification * weight_revenue
  ) |>
  dplyr::summarise(
    mean_weighted = sum(mean_raw * weight_multiplier * n_respondents)/116) |>
  dplyr::left_join(
    bootstrap_confidence_intervals_market,  # Lisame bootstrap'i veapiirid
    join_by(question_id)) |>
  dplyr::left_join(
    question_key,  # Lisame küsimuste nimed
    join_by(question_id)) |>
  dplyr::select(
    question_id,
    question_name = name,  # Küsimuse nimi
    mean = mean_weighted,  # Kaalutud keskmine
    confidence_level,
    confidence_interval_start,
    confidence_interval_end)

#############
# Andmete salvestamine #
#############

# Salvestame lõpliku andmestiku CSV-failina
write.csv(AT_benefits_market, "AT_benefit_market.csv", row.names = FALSE)




################################################
# Vastuste proportsioonid suuruserühmade lõikes iga küsimuse kohta
################################################

################################
# Kaalutud proportsioonid suurusgruppide kaupa #
################################

# Suurusklasside lõikes roportsioonide arvutamine
AT_benefit_weighted_proportions <- AT_benefit_responses |>
  dplyr::left_join(
    respondent_key,  # Liidame vastajate suurusklassi andmed
    join_by(respondent_id)) |>
  dplyr::left_join(
    weights,  # Liidame kaalude andmed
    join_by(respondent_size_id)) |>
  dplyr::group_by(question_id, respondent_size_id, code) |>
  dplyr::summarise(
    n_responses = n(),  # Vastuste arv igas grupis (küsimus, suurusklass, vastuse kood)
    weighted_responses_NR = weight_post_stratification * n_responses,  # NR Kaalutud vastuste koguarv
    weighted_responses_market = weight_post_stratification * weight_revenue * n_responses,  # turule Kaalutud vastuste koguarv
    .groups = "drop"  # Tühistame grupeerimise
  ) |>
  dplyr::distinct(question_id, respondent_size_id, code, .keep_all = TRUE) |>  # Eemaldame duplikaadid
  dplyr::group_by(question_id, respondent_size_id) |>
  dplyr::mutate(
```

```
    total_weighted_in_group_NR = sum(weighted_responses_NR),  # Vastuste NR kaalutud koguarv grupis (suurusklassis)
    total_weighted_in_group_market = sum(weighted_responses_market),  # Vastuste turule kaalutud koguarv grupis (suurusklassis)
    total_responses_in_group = sum(n_responses),    #Vastuste kaalumata koguarv suurusklassis
    kogumist_kaalutud_osakaal_NR = weighted_responses_NR / 116,  # NR Kaalutud proportsioon igale vastuse koodile kogumis 116
    kogumist_kaalutud_osakaal_market = weighted_responses_market / 116,  # NR Kaalutud proportsioon igale vastuse koodile kogumis 116
    suurusklassist_kaalumata_osakaal = n_responses / total_responses_in_group,   # kaalumata proportsioon suurusklassis
  ) |>

  # # KONTROLLVEERUD
  #   dplyr::group_by(question_id) |>
  # dplyr::mutate(
  #   Kontroll_NR = sum(kogumist_kaalutud_osakaal_NR),    # Iga küsimuse NR kaalutud osakaalud peavad kokku andma 1
  #   Kontroll_market = sum(kogumist_kaalutud_osakaal_market)    # Iga küsimuse market kaalutud osakaalud peavad kokku andma 1
  # ) |>
  # dplyr::group_by(question_id, respondent_size_id) |>
  # dplyr::mutate(
  #   Kontroll_kaalumata_osakaal = sum(suurusklassist_kaalumata_osakaal)   #iga küsimuse iga suurusklassi kaalud peavad kokku andma 1
  #     ) |>
  dplyr::ungroup()  # Tühistame kõik grupeerimised

# print(AT_benefit_weighted_proportions)


# Loome kokkuvõtliku tabeli
AT_benefit_proportions_table <- AT_benefit_weighted_proportions |>
  dplyr::left_join(
    question_key,  # Lisame küsimuste nimed
    join_by(question_id)) |>
  dplyr::select(
    question_id,
    question_name = name,  # Küsimuse nimi
    respondent_size_id,
    code,
    n_responses,
    weighted_responses_NR,
    weighted_responses_market,
    total_weighted_in_group_NR,
    total_weighted_in_group_market,
    total_responses_in_group,
    kogumist_kaalutud_osakaal_NR,
    kogumist_kaalutud_osakaal_market,
    suurusklassist_kaalumata_osakaal

  ) |>
  dplyr::arrange(question_id, respondent_size_id, code)  # Sortimine küsimuse, suurusklassi ja vastuse koodi järgi

# Salvestame tulemused CSV-faili
write.csv(AT_benefit_proportions_table, "AT_benefit_proportions.csv", row.names = FALSE)

################################
# Andmed valmis graafiku jaoks #
################################
library(ggplot2)
library(dplyr)
library(extrafont)

# Laeme tabeli, mis sisaldab proportsioonide andmeid
AT_benefit_proportions_table <- read.csv("AT_benefit_proportions.csv")

# Kohandame suurusklasside ja vastuse koodide silte
size_group_labels <- c(
  "1" = "Väikesed audiitorettevõtjad (n=7, N=51)",
  "2" = "Keskmised audiitorettevõtjad (n=21, N=57)",
  "3" = "Suured audiitorettevõtjad (n=6, N=8)"
)

# Define labels for benefits and response levels
benefit_labels <- c(
  "kasu_kvaliteet" = "Tõuseb töövõtu kvaliteet",
  "kasu_aeg" = "Väheneb töövõtule kuluv aeg",
  "kasu_efektiivsus" = "Tõuseb töö efektiivsus",
  "kasu_vĆ¤hem_vigu" = "Väheneb vigade hulk",
  "kasu_rohkem_aega_keerulisemaks" = "Vabaneb aeg kõrgema\nlisandväärtusega tegevustele",
  "kasu_konkurentsivĆµime" = "Tõuseb ettevõtte\nkonkurentsivõime",
  "kasu_standardite_tĆ¤itmine" = "Lihtsustub standardite\nja regulatsioonide täitmine"
```

```
)

magnitude_labels <- c(
  "0 Kasu puudub",
  "1 Väike kasu",
  "2 Mõõdukas kasu",
  "3 Suur kasu"
)



# Loome kihilise tulpgraafiku, kus iga takistuse kohta on üks rida
horizontal_grouped_bar_plot <- ggplot(AT_benefit_proportions_table, aes(
  x = suurusklassist_kaalumata_osakaal,  # X-teljel osakaal
  y = factor(question_name, levels = unique(question_name)),  # Y-teljel küsimused
  fill = factor(code, levels = c("0", "1", "2", "3"))  # Värvid vastuse koodide järgi
)) +
  geom_bar(stat = "identity", position = position_stack(reverse = TRUE), width = 0.6) +  # Kihilised horisontaalsed tulbad
  facet_grid(~ factor(respondent_size_id, labels = size_group_labels_hor),  # Jaotame suurusklasside järgi kõrvuti
             scales = "free_x", space = "free") +
  scale_fill_manual(
    name = "Mõju ulatus",  # Legendi pealkiri
    values = c(
      "3" = "#388E3C",  # tumeroheleni - suur kasu
      "2" = "#81C784",  # roheline - mõõdukas kasu
      "1" = "#C8E6C9",  # heleroheline - väike kasu
      "0" = "#B0B0B0"   # hall - kasu puudub
    ),
    labels = magnitude_labels  # Legendi sildid
  ) +
  scale_x_continuous(
    labels = scales::percent_format(accuracy = 1)  # Vormindame X-telge protsentidena
  ) +
  scale_y_discrete(labels = benefit_labels) +  # Y-telje sildid
  labs(
    x = "Osakaal",  # X-telje pealkiri
    y = "AT kasutuselevõtust saadav kasu",  # Y-telje pealkiri
    # title = "Takistuste jaotus suurusklasside kaupa",  # Graafiku pealkiri
    # subtitle = "Proportsioonid vastuste ja suurusklasside järgi",  # Alapealkiri
    # caption = "Andmed kajastavad suurusklasside kaalumata proportsioone"  # Selgitus
  ) +
  theme_minimal() +  # Minimalistlik kujundus
  theme(
    text = element_text(family = "Times New Roman"),  # Teksti stiil
    axis.text.x = element_text(size = 10),  # X-telje teksti suurus
    axis.text.y = element_text(size = 13),  # Y-telje teksti suurus
    # axis.title.x = element_text(size = 14, margin = margin(t = 10)),  # X-telje pealkirja suurus
    # axis.title.y = element_text(size = 14, margin = margin(r = 10)),  # Y-telje pealkirja suurus
    axis.title.x = element_blank(),
    axis.title.y = element_blank(),
    legend.position = "bottom",  # Legendi asukoht
    legend.title = element_text(size = 12),  # Legendi pealkirja suurus
    legend.text = element_text(size = 10),  # Legendi teksti suurus
    strip.text = element_text(size = 12, face = "bold")  # Suurusklasside pealkirjad (facet_grid)
  )

# Salvestame graafiku pildifailina
ggsave(
  filename = "15.2_AT_benefit_size_comparison_horizontal.png",
  plot = horizontal_grouped_bar_plot,
  device = "png",
  width = 12,
  height = 6,
  units = "in",
  bg = "white"
)

# Kuvame graafiku
# print(horizontal_grouped_bar_plot)
####################
# Keskkonna ettevalmistus #
####################

# Laadime vajalikud R-paketid
if (!require(pacman, quietly = TRUE)) install.packages(pacman)
```

```
pacman::p_load(
  dplyr,          # Andmete töötlemine ja vormindamine
  ggplot2,        # Graafika loomine
  readr,          # Andmete lugemine CSV-failidest
  scales,         # Telgede siltide vormindamine
  stringr         # Teksti töötlemise operatsioonid
)


#############
# Andmete lugemine #
#############

# Loeme kaks failid: NR kaalutud ja turu kaalutud andmed
AT_benefit_NR_path <- "AT_benefit_NR.csv"  # NR kaalutud fail
AT_benefit_market_path <- "AT_benefit_market.csv"  # Turu kaalutud fail

AT_benefit_NR <- readr::read_csv(AT_benefit_NR_path, show_col_types = FALSE)
AT_benefit_market <- readr::read_csv(AT_benefit_market_path, show_col_types = FALSE)


###############
# Andmete ühendamine #
###############

# Lisame veeru, mis näitab, millisest allikast andmed pärinevad
AT_benefit_NR <- AT_benefit_NR |>
  dplyr::mutate(source = "Audiitorettevõtjate keskmine")
AT_benefit_market <- AT_benefit_market |>
  dplyr::mutate(source = "Audiitorteenuste turu keskmine")

# Ühendame kaks andmestikku üheks suureks tabeliks
AT_benefit <- dplyr::bind_rows(
  AT_benefit_NR,
  AT_benefit_market
)


###############
# Graafiku andmed #
###############

plot_data <- AT_benefit

# Defineerime kasu saamise viiside nimed
benefit_labels <- c(
  "kasu_kvaliteet" = "Tõuseb töövõtu kvaliteet",
  "kasu_aeg" = "Väheneb töövõtule kuluv aeg",
  "kasu_efektiivsus" = "Tõuseb töö efektiivsus",
  "kasu_vĆ¤hem_vigu" = "Väheneb vigade hulk",
  "kasu_rohkem_aega_keerulisemaks" = "Vabaneb aeg kõrgema\nlisandväärtusega tegevustele",
  "kasu_konkurentsivĆµime" = "Tõuseb ettevõtte\nkonkurentsivõime",
  "kasu_standardite_tĆ¤itmine" = "Lihtsustub standardite\nja regulatsioonide täitmine"
)

# Defineerime kasu suuruse sildid
magnitude_labels <- c(
  "0 Kasu\npuudub",
  "1 Väike\nkasu",
  "2 Mõõdukas\nkasu",
  "3 Suur\nkasu"
)

###############
# Graafiku loomine #
###############

plot <- plot_data |>
  # Loome graafiku aluse
  ggplot2::ggplot(
    aes(
      x = `mean`,  # Kasu hinnangu keskmine väärtus
      y = reorder(question_name, question_id),  # Kasu saamise viis
      color = source)) +  # Kasutame allikat värvidena
```

```
 # Lisame punktid tõenäosuste jaoks
 geom_point(
  size = 3) +

 # Lisame horisontaalsed veapiirid
 geom_errorbarh(
  aes(xmin = confidence_interval_start, xmax = confidence_interval_end),
  height = 0.2) +

 # Lisame tõenäosuste väärtused punktide kohale
 # Lisame tõenäosuste väärtused punktide kohale või alla sõltuvalt allikast
 geom_text(
  aes(
   label = round(mean, 2),
   y = ifelse(source == "Audiitorteenuste turu keskmine",
              as.numeric(reorder(question_name, question_id)) + 0.2,
              as.numeric(reorder(question_name, question_id)) - 0.2)),
  size = 5,
  vjust = ifelse(plot_data$source == "Audiitorteenuste turu keskmine", -0.5, 1.5)) +

 # X-telje skaala
 scale_x_continuous(
  limits = c(-0.5, 3.5),  # Kasu suurus (0-3)
  expand = c(0, 0),
  labels = magnitude_labels,  # Kasu tasemed
  breaks = c(0, 1, 2, 3)) +

 # Y-telje skaala
 scale_y_discrete(labels = benefit_labels) +

 # Kohandame värvide legendi
 scale_color_manual(
  name = "Kaalumine:",
  values = c("Audiitorettevõtjate keskmine" = "#4CAF50",  # Roheline värv NR jaoks
             "Audiitorteenuste turu keskmine" = "#FF5733")) +  # Oranž värv turu jaoks

 # Lisame graafiku sildid ja selgituse
 labs(
  x = "Kasu määr (0-3)",
  y = "Kasutuselevõtust saadav kasu",
  caption = stringr::str_glue("Lõigud tähistavad veapiire usaldusnivool {unique(plot_data$confidence_level) * 100}%")) +

 # Graafiku kujundus
 theme_minimal() +

 theme(
  text = element_text(family = "Times New Roman"),  # Kirjastiil
  plot.margin = margin(20, 20, 20, 20),  # Graafiku ääred
  # axis.title.x = element_text(size = 24, margin = margin(t = 30)),
  # axis.title.y = element_text(size = 24, margin = margin(r = 30)),
  axis.title.x = element_blank(),
  axis.title.y = element_blank(),
  axis.text = element_text(size = 18),
  axis.text.x = element_text(margin = margin(t = 10)),
  plot.caption.position = "plot",
  plot.caption = element_text(size = 18, hjust = 0),
  legend.position = "bottom",  # Legendi asukoht
  legend.text = element_text(size = 14),
  legend.title = element_text(size = 16),
  # panel.grid = element_line(colour = "black")
  legend.spacing = unit(0.0, "points")
 )


###############
# Graafiku salvestamine #
###############

# Salvestame graafiku pildifailina
ggplot2::ggsave(
 filename = "15_AT_benefit_type.png",
 plot = plot,
 device = "png",
 width = 4000,
```

```
  height = 2500,
  units = "px",
  bg = "white")
####################
# Setup environment #
####################

# Load libraries
if (!require(pacman, quietly = TRUE)) install.packages(pacman)

pacman::p_load(
  dplyr,        # data formatting
  purrr,        # list operations
  readr,        # read data from files
  tidyr         # table transformations
)


#############
# Read data #
#############

responses_path <- "responses.xlsx"
responses <- read_excel(responses_path)

respondent_key_path <- "respondent_key.csv"
respondent_key <- readr::read_csv(respondent_key_path, show_col_types = FALSE)

weights_path <- "weights.csv"
weights <- readr::read_csv(weights_path, show_col_types = FALSE)

question_key_path <- "question_key.xlsx"
question_key <- read_excel(question_key_path)


#############
# Functions #
#############


get_bootstrap_samples <- function(codes, n_respondents, n_samples) {
  # Helped function to sample n_samples of size n_respondents
  # from the input vector of codes

  bootstrap_samples <- purrr::map(
    .x = 1:n_samples,
    .f = \(x) sample(codes, n_respondents, replace = TRUE))

  return(bootstrap_samples)
}


############################
# Filter relevant_questions #
############################

# Takistavate tegurite küsimuste ID-d
AT_barrier_question_ids <- 83:91

# Filtreerime vastused, mis vastavad küsimuste ID-dele
AT_barrier_responses <- responses |>
  dplyr::filter(question_id %in% AT_barrier_question_ids)



# Defineerime tehnoloogiate nimed
# column_labels <- c(
#   "takistus_Regulatiivsne" = "Regulatiivne",
#   "takistus_Aega_puudu" = "Aega puudu",
#   "takistus_Teadmisi_puudu" = "Teadmisi puudu",
#   "takistus_Finantse_puudu_ise_arendamiseks" = "Finantse puudu ise arendamiseks",
#   "takistus_Finantse_puudu_ostmiseks" = "Finantse puudu ostmiseks",
#   "takistus_OskustĆ¶Ć¶tajate_vĆ¤rbamise_probleem" = "Oskustöötajate värbamise probleem",
#   "takistus_Klient_ei_vaja" = "Klient ei vaja",
#   "takistus_Pole_tasuv_investeering" = "Pole tasuv investeering",
```

```
#   "takistus_Kliendi_sisendandmed_valel_kujul" = "Kliendi sisendandmed valel kujul"
# )


#############
# Funktsioonid #
#############

# Funktsioon bootstrap'i proovide loomiseks
get_bootstrap_samples <- function(codes, n_respondents, n_samples) {
  # Loob n_samples bootstrap-proovi, kus igas proovis on n_respondents vastust

  bootstrap_samples <- purrr::map(
    .x = 1:n_samples,  # Iteratsioonide arv
    .f = \(x) sample(codes, n_respondents, replace = TRUE))  # Proovide loomine asendamisega

  return(bootstrap_samples)  # Tagastab bootstrap-proovid
}

########################################
#NON-RESPONSE KAALUGA
########################################
# Bootstrap'i veapiiride arvutamine #
########################################

# Määrame usaldusnivoo ja bootstrap'i simulatsioonide arvu
confidence_interval_confidence_level <- 0.95
n_bootstrap_simulations <- 10^4  # 10 000 simulatsiooni

# Bootstrap'i sisendandmete ettevalmistamine
bootstrap_input <- AT_barrier_responses |>
  dplyr::left_join(
    respondent_key,  # Lisame vastaja võtme
    join_by(respondent_id)) |>
  dplyr::group_by(respondent_size_id, question_id) |>
  dplyr::summarise(
    codes = list(code),  # Vastuste koodid listina
    n_respondents = n(),  # Vastajate arv igas grupis
    .groups = "drop") |>
  dplyr::left_join(
    weights,  # Lisame kaalud
    join_by(respondent_size_id)) |>
  dplyr::group_by(question_id) |>
  dplyr::mutate(
    # weight_multiplier = weight_post_stratification / sum(weight_post_stratification)
    weight_multiplier = weight_post_stratification
  ) |>
  dplyr::arrange(question_id, respondent_size_id) |>
  dplyr::select(
    question_id,
    respondent_size_id,
    codes,
    n_respondents,
    weight_multiplier)

# Lisame iga rea jaoks bootstrap-proovid
bootstrap_samples <- bootstrap_input |>
  dplyr::rowwise() |>
  dplyr::mutate(
    bootstrap_samples = list(
      get_bootstrap_samples(
        codes,  # Vastuste koodid
        n_respondents,  # Vastajate arv
        n_bootstrap_simulations)),  # Simulatsioonide arv
    mean_raw = list(
      purrr::map(
        bootstrap_samples,
        \(x) mean(x))))  # Arvutame iga proovi keskmise


# Loome tühja tabeli bootstrap'i keskmiste jaoks
bootstrap_mean_weighted <- tibble::tibble()

# Läbime iga küsimuse ID
for (question_id in unique(bootstrap_samples$question_id)) {
```

```
  bootstrap_values_by_question_id <- bootstrap_samples |>
    dplyr::filter(question_id == !!question_id)

  # Raw (toore) keskmised bootstrap-proovidest
  mean_raw <- bootstrap_values_by_question_id |>
    dplyr::pull(mean_raw) |>
    purrr::map(\(listed_values) unlist(listed_values))

  weight_multipliers <- bootstrap_values_by_question_id$weight_multiplier
  n_respondents <- bootstrap_values_by_question_id$n_respondents

  # Läbime kõik simulatsioonid ja arvutame kaalutud keskmised
  mean_weighted_by_simulation <- c()
  for (i_simulation in seq.int(n_bootstrap_simulations)) {
    mean_raw_by_simulation <- purrr::map_dbl(mean_raw, \(x) x[i_simulation])
    mean_weighted_by_simulation <- c(mean_weighted_by_simulation, sum(weight_multipliers *
                                                mean_raw_by_simulation * n_respondents)/116)
  }
  print(mean_weighted_by_simulation)

  # Lisame kaalutud keskmised iga küsimuse kohta
  bootstrap_mean_weighted <- bootstrap_mean_weighted |>
    dplyr::bind_rows(
      tibble::tibble(
        question_id = !!question_id,
        mean_weighted = list(mean_weighted_by_simulation)))
}
print(bootstrap_mean_weighted)
# Bootstrap'i veapiiride arvutamine
bootstrap_confidence_intervals <- bootstrap_mean_weighted |>
  dplyr::rowwise() |>
  dplyr::mutate(
    confidence_level = confidence_interval_confidence_level,  # Usaldusnivoo
    confidence_interval_start = quantile(
      unlist(mean_weighted),
      probs = (1 - confidence_level) / 2, names = FALSE),  # Alumine piir
    confidence_interval_end = quantile(
      unlist(mean_weighted),
      probs = 1 - ((1 - confidence_level) / 2), names = FALSE)) |>  # Ülemine piir
  dplyr::select(-mean_weighted)

################################
# Takistuse keskmise arvutamine #
################################

# Arvutame toored keskmised suurusegrupi ja küsimuse alusel
AT_barrier_by_size <- AT_barrier_responses |>
  dplyr::left_join(
    respondent_key,  # Lisame vastaja võtme
    join_by(respondent_id)) |>
  dplyr::group_by(respondent_size_id, question_id) |>
  dplyr::summarise(
    mean_raw = mean(code),  # Toorne keskmine
    n_respondents = n(),  # Vastajate arv
    .groups = "drop")

# Lisame kaalud ja arvutame tegelikud kaalutud keskmised
AT_barrier <- AT_barrier_by_size |>
  dplyr::left_join(
    weights,
    join_by(respondent_size_id)) |>
  dplyr::group_by(question_id) |>
  dplyr::mutate(
    # weight_multiplier = weight_post_stratification / sum(weight_post_stratification)
    weight_multiplier = weight_post_stratification
  ) |>
  dplyr::summarise(
    mean_weighted = sum(mean_raw * weight_multiplier * n_respondents)/116) |>
  dplyr::left_join(
    bootstrap_confidence_intervals,  # Lisame bootstrap'i veapiirid
    join_by(question_id)) |>
  dplyr::left_join(
    question_key,  # Lisame küsimuste nimed
    join_by(question_id)) |>
  dplyr::select(
```

```
    question_id,
    question_name = name,  # Küsimuse nimi
    mean = mean_weighted,  # Kaalutud keskmine
    confidence_level,
    confidence_interval_start,
    confidence_interval_end)

#############
# Andmete salvestamine #
#############

# Salvestame lõpliku andmestiku CSV-failina
write.csv(AT_barrier, "AT_barrier_NR.csv", row.names = FALSE)



##########################################
#TURU KAALUGA
########################################
# Bootstrap'i veapiiride arvutamine #
########################################

# Bootstrap'i sisendandmete ettevalmistamine
bootstrap_input_market <- AT_barrier_responses |>
  dplyr::left_join(
    respondent_key,  # Lisame vastaja võtme
    join_by(respondent_id)) |>
  dplyr::group_by(respondent_size_id, question_id) |>
  dplyr::summarise(
    codes = list(code),  # Vastuste koodid listina
    n_respondents = n(),  # Vastajate arv igas grupis
    .groups = "drop") |>
  dplyr::left_join(
    weights,  # Lisame kaalud
    join_by(respondent_size_id)) |>
  dplyr::group_by(question_id) |>
  dplyr::mutate(
    # weight_multiplier = weight_post_stratification / sum(weight_post_stratification)
    weight_multiplier = weight_post_stratification * weight_revenue
  ) |>
  dplyr::arrange(question_id, respondent_size_id) |>
  dplyr::select(
    question_id,
    respondent_size_id,
    codes,
    n_respondents,
    weight_multiplier)

# Lisame iga rea jaoks bootstrap-proovid
bootstrap_samples_market <- bootstrap_input_market |>
  dplyr::rowwise() |>
  dplyr::mutate(
    bootstrap_samples_market = list(
      get_bootstrap_samples(
        codes,  # Vastuste koodid
        n_respondents,  # Vastajate arv
        n_bootstrap_simulations)),  # Simulatsioonide arv
    mean_raw = list(
      purrr::map(
        bootstrap_samples_market,
        \(x) mean(x))))  # Arvutame iga proovi keskmise


# Loome tühja tabeli bootstrap'i keskmiste jaoks - TURG
bootstrap_mean_weighted_market <- tibble::tibble()

# Läbime iga küsimuse ID
for (question_id in unique(bootstrap_samples_market$question_id)) {
  bootstrap_values_by_question_id_market <- bootstrap_samples_market |>
    dplyr::filter(question_id == !!question_id)

  # Raw (toore) keskmised bootstrap-proovidest
  mean_raw <- bootstrap_values_by_question_id_market |>
    dplyr::pull(mean_raw) |>
    purrr::map(\(listed_values) unlist(listed_values))
```

```
  weight_multipliers_market <- bootstrap_values_by_question_id_market$weight_multiplier
  n_respondents <- bootstrap_values_by_question_id_market$n_respondents

  # Läbime kõik simulatsioonid ja arvutame kaalutud keskmised
  mean_weighted_by_simulation_market <- c()
  for (i_simulation in seq.int(n_bootstrap_simulations)) {
    mean_raw_by_simulation_market <- purrr::map_dbl(mean_raw, \(x) x[i_simulation])
    mean_weighted_by_simulation_market <- c(mean_weighted_by_simulation_market, sum(weight_multipliers_market *
                                                                                      mean_raw_by_simulation_market * n_respondents)/116)
  }
  print(mean_weighted_by_simulation_market)

  # Lisame kaalutud keskmised iga küsimuse kohta
  bootstrap_mean_weighted_market <- bootstrap_mean_weighted_market |>
    dplyr::bind_rows(
      tibble::tibble(
        question_id = !!question_id,
        mean_weighted_market = list(mean_weighted_by_simulation_market)))
}
# print(bootstrap_mean_weighted_market)
# Bootstrap'i veapiiride arvutamine
bootstrap_confidence_intervals_market <- bootstrap_mean_weighted_market |>
  dplyr::rowwise() |>
  dplyr::mutate(
    confidence_level = confidence_interval_confidence_level,  # Usaldusnivoo
    confidence_interval_start = quantile(
      unlist(mean_weighted_market),
      probs = (1 - confidence_level) / 2, names = FALSE),  # Alumine piir
    confidence_interval_end = quantile(
      unlist(mean_weighted_market),
      probs = 1 - ((1 - confidence_level) / 2), names = FALSE)) |>  # Ülemine piir
  dplyr::select(-mean_weighted_market)

################################
# Innovatsiooni tõenäosuste arvutamine #
################################

# Lisame kaalud ja arvutame tegelikud kaalutud keskmised
AT_barrier_market <- AT_barrier_by_size |>
  dplyr::left_join(
    weights,
    join_by(respondent_size_id)) |>
  dplyr::group_by(question_id) |>
  dplyr::mutate(
    # weight_multiplier = weight_post_stratification / sum(weight_post_stratification)
    weight_multiplier = weight_post_stratification * weight_revenue
  ) |>
  dplyr::summarise(
    mean_weighted = sum(mean_raw * weight_multiplier * n_respondents)/116) |>
  dplyr::left_join(
    bootstrap_confidence_intervals_market,  # Lisame bootstrap'i veapiirid
    join_by(question_id)) |>
  dplyr::left_join(
    question_key,  # Lisame küsimuste nimed
    join_by(question_id)) |>
  dplyr::select(
    question_id,
    question_name = name,  # Küsimuse nimi
    mean = mean_weighted,  # Kaalutud keskmine
    confidence_level,
    confidence_interval_start,
    confidence_interval_end)

#############
# Andmete salvestamine #
#############

# Salvestame lõpliku andmestiku CSV-failina
write.csv(AT_barrier_market, "AT_barrier_market.csv", row.names = FALSE)
```

```
################################################
# Vastuste proportsioonid suuruserühmade lõikes iga küsimuse kohta
################################################

################################
# Kaalutud proportsioonid suurusgruppide kaupa #
################################

# Suurusklasside lõikes roportsioonide arvutamine
AT_barrier_weighted_proportions <- AT_barrier_responses |>
  dplyr::left_join(
    respondent_key,  # Liidame vastajate suurusklassi andmed
    join_by(respondent_id)) |>
  dplyr::left_join(
    weights,  # Liidame kaalude andmed
    join_by(respondent_size_id)) |>
    dplyr::group_by(question_id, respondent_size_id, code) |>
  dplyr::summarise(
    n_responses = n(),  # Vastuste arv igas grupis (küsimus, suurusklass, vastuse kood)
    weighted_responses_NR = weight_post_stratification * n_responses,  # NR Kaalutud vastuste koguarv
    weighted_responses_market = weight_post_stratification * weight_revenue * n_responses,  # turule Kaalutud vastuste koguarv
    .groups = "drop"  # Tühistame grupeerimise
  ) |>
  dplyr::distinct(question_id, respondent_size_id, code, .keep_all = TRUE) |>  # Eemaldame duplikaadid
  dplyr::group_by(question_id, respondent_size_id) |>
  dplyr::mutate(
    total_weighted_in_group_NR = sum(weighted_responses_NR),  # Vastuste NR kaalutud koguarv grupis (suurusklassis)
    total_weighted_in_group_market = sum(weighted_responses_market),  # Vastuste turule kaalutud koguarv grupis (suurusklassis)
    total_responses_in_group = sum(n_responses),    #Vastuste kaalumata koguarv suurusklassis
    kogumist_kaalutud_osakaal_NR = weighted_responses_NR / 116,  # NR Kaalutud proportsioon igale vastuse koodile kogumis 116
    kogumist_kaalutud_osakaal_market = weighted_responses_market / 116,  # NR Kaalutud proportsioon igale vastuse koodile kogumis 116
    suurusklassist_kaalumata_osakaal = n_responses / total_responses_in_group,   # kaalumata proportsioon suurusklassis
  ) |>

  # KONTROLLVEERUD
  #   dplyr::group_by(question_id) |>
  # dplyr::mutate(
  #   Kontroll_NR = sum(kogumist_kaalutud_osakaal_NR),    # Iga küsimuse NR kaalutud osakaalud peavad kokku andma 1
  #   Kontroll_market = sum(kogumist_kaalutud_osakaal_market)    # Iga küsimuse market kaalutud osakaalud peavad kokku andma 1
  # ) |>
  # dplyr::group_by(question_id, respondent_size_id) |>
  # dplyr::mutate(
  #   Kontroll_kaalumata_osakaal = sum(suurusklassist_kaalumata_osakaal)   #iga küsimuse iga suurusklassi kaalud peavad kokku andma 1
  #     ) |>
  dplyr::ungroup()  # Tühistame kõik grupeerimised

# print(AT_barrier_weighted_proportions)


# Loome kokkuvõtliku tabeli
AT_barrier_proportions_table <- AT_barrier_weighted_proportions |>
  dplyr::left_join(
    question_key,  # Lisame küsimuste nimed
    join_by(question_id)) |>
  dplyr::select(
    question_id,
    question_name = name,  # Küsimuse nimi
    respondent_size_id,
    code,
    n_responses,
    weighted_responses_NR,
    weighted_responses_market,
    total_weighted_in_group_NR,
    total_weighted_in_group_market,
    total_responses_in_group,
    kogumist_kaalutud_osakaal_NR,
    kogumist_kaalutud_osakaal_market,
    suurusklassist_kaalumata_osakaal

  ) |>
  dplyr::arrange(question_id, respondent_size_id, code)  # Sortimine küsimuse, suurusklassi ja vastuse koodi järgi

# Salvestame tulemused CSV-faili
write.csv(AT_barrier_proportions_table, "AT_barrier_proportions.csv", row.names = FALSE)
```

```
##############################
# Andmed valmis graafiku jaoks #
##############################
####################
# Keskkonna ettevalmistus #
####################

# Laadime vajalikud R-paketid
if (!require(pacman, quietly = TRUE)) install.packages(pacman)
library(extrafont)

pacman::p_load(
  dplyr,        # Andmete töötlemine ja vormindamine
  ggplot2,      # Graafika loomine
  readr,        # Andmete lugemine CSV-failidest
  scales,       # Telgede siltide vormindamine
  stringr       # Teksti töötlemise operatsioonid
)


#############
# Andmete lugemine #
#############

# Loeme kaks failid: NR kaalutud ja turu kaalutud andmed
AT_barrier_NR_path <- "AT_barrier_NR.csv"  # NR kaalutud fail
AT_barrier_market_path <- "AT_barrier_market.csv"  # Turu kaalutud fail

AT_barrier_NR <- readr::read_csv(AT_barrier_NR_path, show_col_types = FALSE)
AT_barrier_market <- readr::read_csv(AT_barrier_market_path, show_col_types = FALSE)


###############
# Andmete ühendamine #
###############

# Lisame veeru, mis näitab, millisest allikast andmed pärinevad
AT_barrier_NR <- AT_barrier_NR |>
  dplyr::mutate(source = "Audiitorettevõtjate keskmine")
AT_barrier_market <- AT_barrier_market |>
  dplyr::mutate(source = "Audiitorteenuste turu keskmine")

# Ühendame kaks andmestikku üheks suureks tabeliks
AT_barrier <- dplyr::bind_rows(
  AT_barrier_NR,
  AT_barrier_market
)


###############
# Graafiku andmed #
###############

plot_data <- AT_barrier


# Defineerime takistuste nimed
barrier_labels <- c(
  "takistus_Regulatiivne" = "Regulatiivsed takistused",
  "takistus_Aega_puudu" = "Aega puudu",
  "takistus_Teadmisi_puudu" = "Teadmisi puudu",
  "takistus_Finantse_puudu_ise_arendamiseks" = "Finantse puudu\nise arendamiseks",
  "takistus_Finantse_puudu_ostmiseks" = "Finantse puudu ostmiseks",
  "takistus_OskustĆ¶Ć¶tajate_vĆ¤rbamise_probleem" = "Oskustöötajate\nvärbamise probleem",
  "takistus_Klient_ei_vaja" = "Klient ei vaja",
  "takistus_Pole_tasuv_investeering" = "Pole tasuv investeering",
  "takistus_Kliendi_sisendandmed_valel_kujul" = "Kliendi sisendandmed\nvalel kujul"
)

# Defineerime kasu suuruse sildid \n
magnitude_labels <- c(
  "0 Ei ole\ntakistuseks",
  "1 Väike\ntakistav mõju",
  "2 Mõõdukas\ntakistav mõju",
  "3 Suur\ntakistav mõju"
```

```
)

###############
# Graafiku loomine #
###############

plot <- plot_data |>
  # Loome graafiku aluse
  ggplot2::ggplot(
    aes(
      x = `mean`,  # Kasu hinnangu keskmine väärtus
      y = reorder(question_name, question_id),  # Kasu saamise viis
      color = source)) +  # Kasutame allikat värvidena

  # Lisame punktid tõenäosuste jaoks
  geom_point(
    size = 3) +

  # Lisame horisontaalsed veapiirid
  geom_errorbarh(
    aes(xmin = confidence_interval_start, xmax = confidence_interval_end),
    height = 0.2) +

  # Lisame tõenäosuste väärtused punktide kohale
  # Lisame tõenäosuste väärtused punktide kohale või alla sõltuvalt allikast
  geom_text(
    aes(
      label = round(mean, 2),
      y = ifelse(source == "Audiitorteenuste turu keskmine",
                 as.numeric(reorder(question_name, question_id)) + 0.2,
                 as.numeric(reorder(question_name, question_id)) - 0.2)),
    size = 5,
    vjust = ifelse(plot_data$source == "Audiitorteenuste turu keskmine", -0.5, 1.5)) +

  # X-telje skaala
  scale_x_continuous(
    limits = c(-0.5, 3.5),  # Takistava mõju suurus (0-3)
    expand = c(0, 0),
    labels = magnitude_labels,  # Takistava mõju tasemed
    breaks = c(0, 1, 2, 3)) +

  # Y-telje skaala
  scale_y_discrete(labels = barrier_labels) +

  # Kohandame värvide legendi
  scale_color_manual(
    name = "Kaalumine:",
    values = c("Audiitorettevõtjate keskmine" = "#4CAF50",  # Roheline värv NR jaoks
               "Audiitorteenuste turu keskmine" = "#FF5733")) +  # Oranž värv turu jaoks

  # Lisame graafiku sildid ja selgituse
  labs(
    x = "Takistava mõju ulatus (0-3)",
    y = "AT kasutselevõttu takistav tegur",
    caption = stringr::str_glue("Lõigud tähistavad veapiire usaldusnivool {unique(plot_data$confidence_level) * 100}%")) +

  # Graafiku kujundus
  theme_minimal() +

  theme(
    text = element_text(family = "Times New Roman"),  # Kirjastiil
    plot.margin = margin(20, 20, 20, 20),  # Graafiku ääred
    # axis.title.x = element_text(size = 24, margin = margin(t = 30)),
    # axis.title.y = element_text(size = 24, margin = margin(r = 30)),
    axis.title.x = element_blank(),
    axis.title.y = element_blank(),
    axis.text.x.bottom = element_text(size = 16),
    axis.text.y = element_text(size = 20),
    axis.text.x = element_text(margin = margin(t = 10)),
    plot.caption.position = "plot",
    plot.caption = element_text(size = 14, hjust = 0),
    legend.position = "bottom",  # Legendi asukoht
    legend.text = element_text(size = 14),
    legend.title = element_text(size = 16),
    # panel.grid = element_line(colour = "black")
```

```
   legend.spacing = unit(0.0, "points")
  )


###############
# Graafiku salvestamine #
###############

# Salvestame graafiku pildifailina
ggplot2::ggsave(
  filename = "17_AT_barriers.png",
  plot = plot,
  device = "png",
  width = 4000,
  height = 2500,
  units = "px",
  bg = "white")
################################
# Stacked Bar Chart (Kihiline tulpgraafik) #
################################

library(ggplot2)
if (!require(extrafont)) install.packages("extrafont")
library(extrafont)

# # Import and register fonts (this step can take time)
# font_import(prompt = FALSE)

# # Load fonts for Windows
# loadfonts(device = "win")


# Laeme tabeli, mis sisaldab proportsioonide andmeid
AT_barrier_proportions_table <- read.csv("AT_barrier_proportions.csv")

# Kohandame suurusklasside ja vastuse koodide silte
size_group_labels <- c(
  "1" = "Väikesed audiitorettevõtjad (n=7, N=51)",
  "2" = "Keskmised audiitorettevõtjad (n=21, N=57)",
  "3" = "Suured audiitorettevõtjad (n=6, N=8)"
)

# Defineerime takistuste nimed
barrier_labels <- c(
  "takistus_Regulatiivne" = "Regulatiivsed takistused",
  "takistus_Aega_puudu" = "Aega puudu",
  "takistus_Teadmisi_puudu" = "Teadmisi puudu",
  "takistus_Finantse_puudu_ise_arendamiseks" = "Finantse puudu\nise arendamiseks",
  "takistus_Finantse_puudu_ostmiseks" = "Finantse puudu ostmiseks",
  "takistus_OskustĆ¶Ć¶tajate_vĆ¤rbamise_probleem" = "Oskustöötajate\nvärbamise probleem",
  "takistus_Klient_ei_vaja" = "Klient ei vaja",
  "takistus_Pole_tasuv_investeering" = "Pole tasuv investeering",
  "takistus_Kliendi_sisendandmed_valel_kujul" = "Kliendi sisendandmed\nvalel kujul"
)

response_code_labels <- c(
  "0" = "0 Ei ole takistuseks",
  "1" = "1 Väike takistav mõju",
  "2" = "2 Mõõdukas takistav mõju",
  "3" = "3 Suur takistav mõju"
)

# # Loome kihilise tulpgraafiku
# stacked_bar_plot <- ggplot(AT_barrier_proportions_table, aes(
#   x = factor(question_name, levels = unique(question_name)),  # Küsimused X-teljel
#   y = suurusklassist_kaalumata_osakaal,  # Y-teljel kaalutud proportsioon
#   fill = factor(code, levels = c("0", "1", "2", "3"))  # Värvid vastuse koodide järgi
# )) +
#   geom_bar(stat = "identity", position = "stack", width = 0.7) +  # Kihiline tulpgraafik
#   facet_wrap(~ factor(respondent_size_id, labels = size_group_labels), ncol = 1) +  # Jaotame graafiku suurusklasside järgi
#   scale_fill_manual(
#     name = "Mõju ulatus",  # Legendi pealkiri
#     values = c(
#       "0" = "#4CAF50",  # Roheline "Takistus puudub"
#       "1" = "#FFC107",  # Kollane "Väike takistus"
```

```
#       "2" = "#FF9800",  # Oranž "Mõõdukas takistus"
#       "3" = "#F44336"   # Punane "Suur takistus"
#     ),
#     labels = response_code_labels  # Legendi sildid
#   ) +
#   # X-telje skaala
#   scale_x_discrete(labels = barrier_labels) +
#   # Y-telje skaala (protsentides)
#   scale_y_continuous(
#     labels = scales::percent_format(accuracy = 1)  # Vormindame väärtused protsentides täpsusega 1%
#   ) +
#
#   labs(
#     x = "AT kasutuselevõttu takistavad tegurid",
#     y = "Osakaal",
#     # title = "Takistuste jaotus suurusklasside kaupa",
#     # subtitle = "Proportsioonid vastuste ja suurusklasside järgi",
#     # caption = "Andmed kajastavad kaalutud proportsioone iga suurusklassi sees"
#   ) +
#   theme_minimal() +  # Minimalistlik kujundus
#   theme(
#     text = element_text(family = "Times New Roman"),  # Teksti stiil
#     axis.text.x = element_text(angle = 45, hjust = 1, size = 10),  # X-telg diagonaalselt
#     axis.text.y = element_text(size = 12),  # Y-telje teksti suurus
#     # axis.title.x = element_text(size = 14, margin = margin(t = 10)),  # X-telje pealkiri
#     # axis.title.y = element_text(size = 14, margin = margin(r = 10)),  # Y-telje pealkiri
#     axis.title.x = element_blank(),  # X-telje pealkiri
#     axis.title.y = element_blank(),  # Y-telje pealkiri
#     legend.position = "bottom",  # Legendi asukoht
#
#     legend.title = element_text(size = 12),  # Legendi pealkirja suurus
#     legend.text = element_text(size = 10),  # Legendi teksti suurus
#     strip.text = element_text(size = 12, face = "bold")  # Gruppide pealkirjad (facet_wrap)
#   )
#
# # Salvestame graafiku pildifailina
# ggsave(
#   filename = "18_AT_barrier_size_comparison.png",
#   plot = stacked_bar_plot,
#   device = "png",
#   width = 10,
#   height = 8,
#   units = "in",
#   bg = "white"
# )
#
# # Kuvame graafiku
# print(stacked_bar_plot)
#
#
#
```

```
##############################HORISONTAALNE VARIANT
################################
# Kõrvuti horisontaalsed kihilised tulbad #
################################

library(ggplot2)


# Kohandame suurusklasside ja vastuse koodide silte
size_group_labels_hor <- c(
  "1" = "Väikesed audiitorettevõtjad\n(n=7, N=51)",
  "2" = "Keskmised audiitorettevõtjad\n(n=21, N=57)",
  "3" = "Suured audiitorettevõtjad\n(n=6, N=8)"
)

# Loome kihilise tulpgraafiku, kus iga takistuse kohta on üks rida
horizontal_grouped_bar_plot <- ggplot(AT_barrier_proportions_table, aes(
```

```
  x = suurusklassist_kaalumata_osakaal,  # X-teljel osakaal
  y = factor(question_name, levels = unique(question_name)),  # Y-teljel küsimused
  fill = factor(code, levels = c("0", "1", "2", "3"))  # Värvid vastuse koodide järgi
)) +
  geom_bar(stat = "identity", position = position_stack(reverse = TRUE), width = 0.6) +  # Kihilised horisontaalsed tulbad
  facet_grid(~ factor(respondent_size_id, labels = size_group_labels_hor),  # Jaotame suurusklasside järgi kõrvuti
             scales = "free_x", space = "free") +
  scale_fill_manual(
    name = "Mõju ulatus",  # Legendi pealkiri
    values = c(
      "0" = "#4CAF50",  # Roheline "Takistus puudub"
      "1" = "#FFC107",  # Kollane "Väike takistus"
      "2" = "#FF9800",  # Oranž "Mõõdukas takistus"
      "3" = "#F44336"   # Punane "Suur takistus"
    ),
    labels = response_code_labels  # Legendi sildid
  ) +
  scale_x_continuous(
    labels = scales::percent_format(accuracy = 1)  # Vormindame X-telge protsentidena
  ) +
  scale_y_discrete(labels = barrier_labels) +  # Y-telje sildid
  labs(
    x = "Osakaal",  # X-telje pealkiri
    y = "AT kasutuselevõttu takistavad tegurid",  # Y-telje pealkiri
    # title = "Takistuste jaotus suurusklasside kaupa",  # Graafiku pealkiri
    # subtitle = "Proportsioonid vastuste ja suurusklasside järgi",  # Alapealkiri
    # caption = "Andmed kajastavad suurusklasside kaalumata proportsioone"  # Selgitus
  ) +
  theme_minimal() +  # Minimalistlik kujundus
  theme(
    text = element_text(family = "Times New Roman"),  # Teksti stiil
    axis.text.x = element_text(size = 12),  # X-telje teksti suurus
    axis.text.y = element_text(size = 16),  # Y-telje teksti suurus
    # axis.title.x = element_text(size = 14, margin = margin(t = 10)),  # X-telje pealkirja suurus
    # axis.title.y = element_text(size = 14, margin = margin(r = 10)),  # Y-telje pealkirja suurus
    axis.title.x = element_blank(),  # X-telje pealkiri
    axis.title.y = element_blank(),  # Y-telje pealkiri
    legend.position = "bottom",  # Legendi asukoht
    legend.title = element_text(size = 12),  # Legendi pealkirja suurus
    legend.text = element_text(size = 12),  # Legendi teksti suurus
    strip.text = element_text(size = 12, face = "bold")  # Suurusklasside pealkirjad (facet_grid)
  )

# Salvestame graafiku pildifailina
ggsave(
  filename = "18_AT_barrier_size_comparison_horizontal.png",
  plot = horizontal_grouped_bar_plot,
  device = "png",
  width = 12,
  height = 6,
  units = "in",
  bg = "white"
)

# Kuvame graafiku
# print(horizontal_grouped_bar_plot)
library(ggplot2)
library(dplyr)
library(extrafont)

# Read data
data <- read.csv("AT_barrier_proportions.csv")

# Define labels for barriers and response levels
barrier_labels <- c(
  "takistus_Regulatiivne" = "Regulatiivsed takistused",
  "takistus_Aega_puudu" = "Aega puudu",
  "takistus_Teadmisi_puudu" = "Teadmisi puudu",
  "takistus_Finantse_puudu_ise_arendamiseks" = "Finantse puudu ise arendamiseks",
  "takistus_Finantse_puudu_ostmiseks" = "Finantse puudu ostmiseks",
  "takistus_Oskustöötajate_värbamise_probleem" = "Oskustöötajate värbamise probleem",
  "takistus_Klient_ei_vaja" = "Klient ei vaja",
  "takistus_Pole_tasuv_investeering" = "Pole tasuv investeering",
  "takistus_Kliendi_sisendandmed_valel_kujul" = "Kliendi sisendandmed valel kujul"
)
```

```
response_code_labels <- c(
  "0" = "Ei ole takistuseks",
  "1" = "Väike takistus",
  "2" = "Mõõdukas takistus",
  "3" = "Suur takistus"
)

# Prepare data for plotting
plot_data <- data %>%
  pivot_longer(cols = c("kogumist_kaalutud_osakaal_NR", "kogumist_kaalutud_osakaal_market"),
               names_to = "weight_type",
               values_to = "proportion") %>%
  mutate(weight_type = recode(weight_type,
                              "kogumist_kaalutud_osakaal_NR" = "NR kaalutud",
                              "kogumist_kaalutud_osakaal_market" = "Turu kaalutud"))

# Create horizontal grouped stacked bar plot
horizontal_grouped_stacked_bar_plot <- ggplot(plot_data, aes(
  y = factor(question_name, levels = rev(unique(question_name))),  # Questions on Y-axis
  x = proportion,  # Proportion on X-axis
  fill = factor(code, levels = c("0", "1", "2", "3"))  # Color by response levels
)) +
  geom_bar(stat = "identity", position = "stack", width = 0.4) +  # Stacked bar for each weight type
  facet_grid(~ weight_type, scales = "free_x") +  # Place NR and Market side-by-side
  scale_fill_manual(
    name = "Vastuse tase",
    values = c(
      "0" = "#4CAF50",  # Green for "No Barrier"
      "1" = "#FFC107",  # Yellow for "Minor Barrier"
      "2" = "#FF9800",  # Orange for "Moderate Barrier"
      "3" = "#F44336"   # Red for "Major Barrier"
    ),
    labels = response_code_labels
  ) +
  scale_y_discrete(labels = barrier_labels) +
  scale_x_continuous(labels = scales::percent_format()) +
  labs(
    y = "Takistavad tegurid",
    x = "Osakaal (%)",
    title = "NR ja Turu kaalutud takistuste osakaalude võrdlus"
  ) +
  theme_minimal() +
  theme(
    text = element_text(family = "Times New Roman"),
    axis.text.x = element_text(size = 12),
    axis.text.y = element_text(size = 10),
    axis.title.x = element_text(size = 14, margin = margin(t = 10)),
    axis.title.y = element_text(size = 14, margin = margin(r = 10)),
    legend.position = "bottom",
    legend.title = element_text(size = 12),
    legend.text = element_text(size = 10),
    strip.text = element_text(size = 12, face = "bold")
  )

# Save the plot
ggsave(
  filename = "AT_barrier_comparison_NR_vs_Market_horizontal.png",
  plot = horizontal_grouped_stacked_bar_plot,
  device = "png",
  width = 12,
  height = 8,
  units = "in",
  bg = "white"
)

# Print the plot
print(horizontal_grouped_stacked_bar_plot)
#####################
# Setup environment #
#####################

# Load libraries
if (!require(pacman, quietly = TRUE)) install.packages(pacman)
```

```
pacman::p_load(
   dplyr,        # data formatting
   purrr,        # list operations
   readr,        # read data from files
   stringr,      # text transformations
   tidyr         # table transformations
)


#############
# Read data #
#############

responses_path <- "responses.csv"
responses <- readr::read_csv(responses_path, show_col_types = FALSE)

respondent_key_path <- "respondent_key.csv"
respondent_key <- readr::read_csv(respondent_key_path, show_col_types = FALSE)

respondent_size_key_path <- "respondent_size_key.csv"
respondent_size_key <- readr::read_csv(respondent_size_key_path, show_col_types = FALSE)

question_key_path <- "question_key.csv"
question_key <- readr::read_csv(question_key_path, show_col_types = FALSE)


#############
# Functions #
#############

get_wilson_confidence_interval <- function(p, n, confidence_level) {
   if (p %in% c(0, 1)) return(c(0, 0))

   z_value <- qnorm(1 - (1 - confidence_interval_confidence_level) / 2)

   pseudo_mean <- (2 * n * p + z_value^2) / (2 * (n + z_value^2))
   span <- z_value * sqrt(z_value^2 - (1 / n) + 4 * n * p * (1 - p)) / (2 * (n + z_value^2))

   confidence_interval <- c(
      pseudo_mean - span,
      pseudo_mean + span
   )
   return(confidence_interval)
}

get_normal_confidence_interval <- function(mean, variance, confidence_level) {
   p <- (1 - confidence_level) / 2
   confidence_interval <- c(
      qnorm(p, mean, sqrt(variance)),
      qnorm(p, mean, sqrt(variance), lower.tail = FALSE)
   )
   return(confidence_interval)
}


#################################
# Get innovation specialist data #
#################################

innovation_specialist_question_id <- 83

data_innovation_specialist <- responses |>
   dplyr::filter(question_id == innovation_specialist_question_id) |>
   # Company has innovation specialists if the answer is something other than "no innovation specialist" (code 5)
   dplyr::mutate(has_innovation_specialist = code != 5) |>
   dplyr::select(
      respondent_id,
      has_innovation_specialist)


################################
# Specialist vs technology use #
################################

no_techology_use_question_id <- 39
```

```
data_technology_use <- responses |>
   dplyr::filter(
      question_id == no_techology_use_question_id) |>
   dplyr::mutate(
      # Company uses some technology if "no technology in use" option is not marked
      uses_technology = code == 0) |>
   dplyr::left_join(
      data_innovation_specialist,
      join_by(respondent_id)) |>
   dplyr::select(
      respondent_id,
      uses_technology,
      has_innovation_specialist)

technology_use_by_size <- data_technology_use |>
   dplyr::left_join(
      respondent_key,
      join_by(respondent_id)) |>
   dplyr::group_by(
      has_innovation_specialist,
      respondent_size_id) |>
   dplyr::summarise(
      n = n(),
      n_uses_technology = sum(uses_technology),
      proportion = n_uses_technology / n,
      .groups = "drop")

technology_use_total <- technology_use_by_size |>
   dplyr::left_join(
      respondent_size_key,
      join_by(respondent_size_id)) |>
   dplyr::group_by(has_innovation_specialist) |>
   dplyr::mutate(
      weight_multiplier = n_companies / sum(n_companies)) |>
   dplyr::summarise(
      n = sum(n),
      n_uses_technology = sum(n_uses_technology),
      proportion = sum(proportion * weight_multiplier))

confidence_interval_confidence_level <- 0.95

technology_use <- dplyr::bind_rows(
   technology_use_by_size,
   technology_use_total) |>
   dplyr::rename(technology_use_proportion = proportion) |>
   dplyr::rowwise() |>
   dplyr::mutate(
      confidence_interval = list(
         get_wilson_confidence_interval(
            technology_use_proportion,
            n,
            confidence_interval_confidence_level)),
      technology_use_proportion_confidence_interval_start = confidence_interval[[1]],
      technology_use_proportion_confidence_interval_end = confidence_interval[[2]],
      respondent_size_id = case_when(
         is.na(respondent_size_id) ~ 99,
         TRUE ~ respondent_size_id)) |>
   ungroup() |>
   dplyr::select(
      respondent_size_id,
      has_innovation_specialist,
      n,
      technology_use_proportion,
      technology_use_proportion_confidence_interval_start,
      technology_use_proportion_confidence_interval_end) |>
   dplyr::arrange(respondent_size_id, has_innovation_specialist)


########################################
# Specialist vs innovation probability #
########################################

innovation_probability_question_ids <- 45:55
```

```
innovation_probability_by_size <- responses |>
  dplyr::filter(
    question_id %in% innovation_probability_question_ids) |>
  dplyr::left_join(
    data_innovation_specialist,
    join_by(respondent_id)) |>
  dplyr::left_join(
    respondent_key,
    join_by(respondent_id)) |>
  dplyr::group_by(
    has_innovation_specialist,
    respondent_size_id,
    question_id) |>
  dplyr::summarise(
    n = length(unique(respondent_id)),
    mean_by_question = mean(code),
    variance_by_question = var(code)) |>
  dplyr::summarise(
    n = unique(n),
    n_questions = length(unique(question_id)),
    mean_by_size_group = mean(mean_by_question),
    # Law of total variance
    variance_by_size_group = mean(variance_by_question) + var(mean_by_question),
    variance_of_mean_by_size_group = variance_by_size_group / n_questions / n) |>
  dplyr::rowwise() |>
  dplyr::mutate(
    confidence_interval = list(
      get_normal_confidence_interval(
        mean_by_size_group,
        variance_of_mean_by_size_group,
        confidence_interval_confidence_level)),
      confidence_interval_start = confidence_interval[[1]],
      confidence_interval_end = confidence_interval[[2]]) |>
  ungroup()

innovation_probability_total <- innovation_probability_by_size |>
  dplyr::left_join(
    respondent_size_key,
    join_by(respondent_size_id)) |>
  dplyr::group_by(has_innovation_specialist) |>
  dplyr::mutate(
    weight_multiplier = n_companies / sum(n_companies)) |>
  dplyr::summarise(
    n = sum(n),
    mean_by_innovation_expert_status = sum(mean_by_size_group * weight_multiplier),
    variance_of_mean_by_innovation_expert_status = sum(variance_of_mean_by_size_group * weight_multiplier^2)) |>
  dplyr::rowwise() |>
  dplyr::mutate(
    confidence_interval = list(
      get_normal_confidence_interval(
        mean_by_innovation_expert_status,
        variance_of_mean_by_innovation_expert_status,
        confidence_interval_confidence_level)),
      confidence_interval_start = confidence_interval[[1]],
      confidence_interval_end = confidence_interval[[2]]) |>
  ungroup()

innovation_probability <- innovation_probability_by_size |>
  dplyr::rename(
    innovation_probability = mean_by_size_group,
    innovation_probability_confidence_interval_start = confidence_interval_start,
    innovation_probability_confidence_interval_end = confidence_interval_end) |>
  dplyr::bind_rows(
    dplyr::rename(
      innovation_probability_total,
      innovation_probability = mean_by_innovation_expert_status,
      innovation_probability_confidence_interval_start = confidence_interval_start,
      innovation_probability_confidence_interval_end = confidence_interval_end)) |>
  dplyr::mutate(
    respondent_size_id = case_when(
      is.na(respondent_size_id) ~ 99,
      TRUE ~ respondent_size_id)) |>
  dplyr::select(
    respondent_size_id,
    has_innovation_specialist,
```

```
      n,
      innovation_probability,
      innovation_probability_confidence_interval_start,
      innovation_probability_confidence_interval_end) |>
   dplyr::arrange(respondent_size_id, has_innovation_specialist)


#############
# Save data #
#############

specialist_vs_innovation <- technology_use |>
   dplyr::left_join(
      innovation_probability,
      join_by(respondent_size_id, has_innovation_specialist, n))

write.csv(specialist_vs_innovation, "specialist_vs_innovation.csv", row.names = FALSE)
writexl::write_xlsx(specialist_vs_innovation, "specialist_vs_innovation.xlsx", col_names = TRUE)
####################
# Keskkonna ettevalmistus #
####################

# Laadime vajalikud paketid
if (!require(pacman, quietly = TRUE)) install.packages(pacman)

pacman::p_load(
  dplyr,         # Andmete vormindamine
  readr,         # Andmete lugemine failidest
  tidyr          # Tabelite teisendused
)

#############
# Andmete lugemine #
#############

vastused_kodeeritud <- read_excel("vastused_kodeeritud.xlsx")

## Loon muutujad, mis sisaldavad vastavate andmete veeru numberid "vastused_kodeeritud" failis
AT_in_use_source_ids <- 28
weight_ids <- c(132, 136)
suurusklass_id <- 130
confidence_interval_confidence_level <- 0.95

## Kasutades eelnevalt loodud ID-de vektoreid, toon andmed uutesse muutujatesse
data_AT_in_use_source <- vastused_kodeeritud[, AT_in_use_source_ids]
data_weights <- vastused_kodeeritud[, weight_ids]
data_suurusklass <- vastused_kodeeritud[, suurusklass_id]

# # Arvutan ka raha kaalu (kaal_turg on juba kahe kaalu korrutis andmetes)
# data_weights <- data_weights %>% dplyr::mutate(
#   kaal_raha = Kaal_turg / Kaal_innov
# )

## Teen koondandmestiku kõikide vajalike muutujatega
data <- cbind(
  data_AT_in_use_source,
  data_weights,
  data_suurusklass
)

print(data)

#
# # Add labels to the column names
# column_labels <- c(
#   "5a: Kasutusel väljatöötaja - Ise" = "AEV ise",
#   "5b: Kasutusel väljatöötaja - Koostöö võrgustikuga" = "AEV koostöös võrgustiku teise liikmega",
#   "5c: Kasutusel väljatöötaja - Koostöö teistega" = "AEV koostöös teise ettevõttega (mitte võrgustikus)",
#   "5d: Kasutusel väljatöötaja - Teine" = "Teine ettevõte"
# )
#
# # Apply labels to the first 4 columns
# colnames(data)[1:4] <- unname(column_labels[colnames(data)[1:4]])

# Veendu, et 8. veerg (suurusklass) on korrektselt tuvastatud
```

```
# suurusklass_column <- colnames(data)[3]


# Samm 1: Loenda vastuste arv suurusklasside lõikes
vastuste_arv <- data %>%
  group_by(`16: Innovatsioonitöötajad`, `19: Suurus käibe alusel`) %>%
  summarise(vastuste_arv = n())


print (vastuste_arv)


######################### NR kaaluga

# Samm 2: Arvutame kaalutud osakaalud ja Wilsoni usalduspiirid
kaalutud_tulemused_NR <- data %>%
  group_by(`16: Innovatsioonitöötajad`) %>%
  summarise(
    weighted_sum = sum(Kaal_innov, na.rm = TRUE),  # Kaalutud summa
    n_weighted = sum(Kaal_innov, na.rm = TRUE),   # Kaalutud vastuste arv
    .groups = "drop"
  )  %>%
  mutate(
    weighted_proportion = weighted_sum / 116,  # Kaalutud osakaal populatsioonis
    z_value = qnorm(1 - (1 - confidence_interval_confidence_level) / 2),  # z väärtus usaldusnivool
    sqrt_term = sqrt(pmax(0, z_value^2 - (1 / n_weighted) + 4 * n_weighted * weighted_proportion * (1 - weighted_proportion))),
    lower_bound = (
      (2 * n_weighted * weighted_proportion + z_value^2 - z_value * sqrt_term) /
        (2 * (n_weighted + z_value^2))
    ),
    upper_bound = (
      (2 * n_weighted * weighted_proportion + z_value^2 + z_value * sqrt_term) /
        (2 * (n_weighted + z_value^2))
    )
  ) %>%
  mutate(
    lower_bound = pmax(0, lower_bound),  # Piirame alumise piiri 0-ga
    upper_bound = pmin(1, upper_bound)   # Piirame ülemise piiri 1-ga
  )

# Tulemuste kuvamine
print(kaalutud_tulemused_NR)



# Tulemus salvestatakse vajadusel CSV-failina
write.csv(kaalutud_tulemused_NR, "innovation_specialist_NR.csv", row.names = FALSE)




########################### MARKET WEIGHT


# Samm 2: Arvutame kaalutud osakaalud ja Wilsoni usalduspiirid
kaalutud_tulemused_market <- data %>%
  group_by(`16: Innovatsioonitöötajad`) %>%
  summarise(
    weighted_sum = sum(Kaal_turg, na.rm = TRUE),  # Kaalutud summa
    n_weighted = sum(Kaal_turg, na.rm = TRUE),   # Kaalutud vastuste arv
    .groups = "drop"
  ) %>%
  mutate(
    weighted_proportion = weighted_sum / 116,  # Kaalutud osakaal populatsioonis
    z_value = qnorm(1 - (1 - confidence_interval_confidence_level) / 2),  # z väärtus usaldusnivool
    sqrt_term = sqrt(pmax(0, z_value^2 - (1 / n_weighted) + 4 * n_weighted * weighted_proportion * (1 - weighted_proportion),1)),
    lower_bound = (
      (2 * n_weighted * weighted_proportion + z_value^2 - z_value * sqrt_term) /
        (2 * (n_weighted + z_value^2))
    ),
    upper_bound = (
      (2 * n_weighted * weighted_proportion + z_value^2 + z_value * sqrt_term) /
        (2 * (n_weighted + z_value^2))
    )
```

```
  ) %>%
  mutate(
    lower_bound = pmax(0, lower_bound),  # Piirame alumise piiri 0-ga
    upper_bound = pmin(1, upper_bound)   # Piirame ülemise piiri 1-ga
  )

 # Tulemuste kuvamine
 print(kaalutud_tulemused_market)

 sum(kaalutud_tulemused_market$n_weighted)
 colSums(kaalutud_tulemused_NR[2:8])

 # Tulemus salvestatakse vajadusel CSV-failina
 write.csv(kaalutud_tulemused_market, "innovation_specialist_market.csv", row.names = FALSE)
library(ggplot2)
library(dplyr)
library(extrafont)

# Laeme tabeli, mis sisaldab proportsioonide andmeid
AT_benefit_proportions_table <- read.csv("AT_benefit_proportions.csv")

# Kohandame suurusklasside ja vastuse koodide silte
size_group_labels <- c(
  "1" = "Väikesed audiitorettevõtjad (n=7, N=51)",
  "2" = "Keskmised audiitorettevõtjad (n=21, N=57)",
  "3" = "Suured audiitorettevõtjad (n=6, N=8)"
)

# Define labels for benefits and response levels
benefit_labels <- c(
  "kasu_kvaliteet" = "Tõuseb töövõtu kvaliteet",
  "kasu_aeg" = "Väheneb töövõtule kuluv aeg",
  "kasu_efektiivsus" = "Tõuseb töö efektiivsus",
  "kasu_vĆ¤hem_vigu" = "Väheneb vigade hulk",
  "kasu_rohkem_aega_keerulisemaks" = "Vabaneb aeg kõrgema\nlisandväärtusega tegevustele",
  "kasu_konkurentsivĆµime" = "Tõuseb ettevõtte\nkonkurentsivõime",
  "kasu_standardite_tĆ¤itmine" = "Lihtsustub standardite\nja regulatsioonide täitmine"
)

magnitude_labels <- c(
  "0 Kasu puudub",
  "1 Väike kasu",
  "2 Mõõdukas kasu",
  "3 Suur kasu"
)



# Loome kihilise tulpgraafiku, kus iga takistuse kohta on üks rida
horizontal_grouped_bar_plot <- ggplot(AT_benefit_proportions_table, aes(
  x = suurusklassist_kaalumata_osakaal,  # X-teljel osakaal
  y = factor(question_name, levels = unique(question_name)),  # Y-teljel küsimused
  fill = factor(code, levels = c("0", "1", "2", "3"))  # Värvid vastuse koodide järgi
)) +
  geom_bar(stat = "identity", position = position_stack(reverse = TRUE), width = 0.6) +  # Kihilised horisontaalsed tulbad
  facet_grid(~ factor(respondent_size_id, labels = size_group_labels_hor),  # Jaotame suurusklasside järgi kõrvuti
             scales = "free_x", space = "free") +
  scale_fill_manual(
    name = "Mõju ulatus",  # Legendi pealkiri
    values = c(
      "3" = "#388E3C",  # tumeroheleni - suur kasu
      "2" = "#81C784",  # roheline - mõõdukas kasu
      "1" = "#C8E6C9",  # heleroheline - väike kasu
      "0" = "#B0B0B0"   # hall - kasu puudub
    ),
    labels = magnitude_labels  # Legendi sildid
  ) +
  scale_x_continuous(
    labels = scales::percent_format(accuracy = 1)  # Vormindame X-telge protsentidena
  ) +
  scale_y_discrete(labels = benefit_labels) +  # Y-telje sildid
  labs(
    x = "Osakaal",  # X-telje pealkiri
    y = "AT kasutuselevõtust saadav kasu",  # Y-telje pealkiri
    # title = "Takistuste jaotus suurusklasside kaupa",  # Graafiku pealkiri
```

```
    # subtitle = "Proportsioonid vastuste ja suurusklasside järgi",  # Alapealkiri
    # caption = "Andmed kajastavad suurusklasside kaalumata proportsioone"  # Selgitus
  ) +
  theme_minimal() +  # Minimalistlik kujundus
  theme(
    text = element_text(family = "Times New Roman"),  # Teksti stiil
    axis.text.x = element_text(size = 10),  # X-telje teksti suurus
    axis.text.y = element_text(size = 13),  # Y-telje teksti suurus
    # axis.title.x = element_text(size = 14, margin = margin(t = 10)),  # X-telje pealkirja suurus
    # axis.title.y = element_text(size = 14, margin = margin(r = 10)),  # Y-telje pealkirja suurus
    axis.title.x = element_blank(),
    axis.title.y = element_blank(),
    legend.position = "bottom",  # Legendi asukoht
    legend.title = element_text(size = 12),  # Legendi pealkirja suurus
    legend.text = element_text(size = 10),  # Legendi teksti suurus
    strip.text = element_text(size = 12, face = "bold")  # Suurusklasside pealkirjad (facet_grid)
  )

# Salvestame graafiku pildifailina
ggsave(
  filename = "15.2_AT_benefit_size_comparison_horizontal.png",
  plot = horizontal_grouped_bar_plot,
  device = "png",
  width = 12,
  height = 6,
  units = "in",
  bg = "white"
)

# Kuvame graafiku
# print(horizontal_grouped_bar_plot)
######################
# Keskkonna ettevalmistus #
######################

# Laadime vajalikud R-paketid
if (!require(pacman, quietly = TRUE)) install.packages(pacman)

pacman::p_load(
  ggplot2,       # Graafika genereerimine
  readr,         # Andmete lugemine failidest
  tidyr,         # Tabelite teisendused (pikk vs lai formaat)
  dplyr,         # Andmete töötlemine
  scales,        # Telgede siltide vormindamine (nt protsendid)
  stringr        # Tekstidega töötlemine
)

#############
# Andmete lugemine #
#############

# Määrame andmefailide asukohad ja loeme need sisse
market_weighted_columns_summary_path <- "innovation_specialist_market.csv"  # Turu kaalutud andmed
market_weighted_data <- readr::read_csv(market_weighted_columns_summary_path, show_col_types = FALSE)

NR_weighted_columns_summary_path <- "innovation_specialist_NR.csv"  # Vastamisest kõrvale jäänud ettevõtete kaalutud andmed
NR_weighted_data <- readr::read_csv(NR_weighted_columns_summary_path, show_col_types = FALSE)

# Lisame veeru "Source" (allikas), mis näitab, kummast andmestikust rida pärineb
market_weighted_data <- market_weighted_data %>%
  mutate(Source = "Osakaal audiitorteenuste turust")  # Turu kaalutud osakaal

NR_weighted_data <- NR_weighted_data %>%
  mutate(Source = "Osakaal audiitorettevõtjatest")  # Vastamisest kõrvale jäänud osakaal

# Kombineerime mõlemad andmestikud üheks suureks andmetabeliks
combined_data <- bind_rows(market_weighted_data, NR_weighted_data)

print(combined_data)

# Graafiku loomine
plot_innovation_specialist <- ggplot(combined_data, aes(
  x = `16: Innovatsioonitöötajad`,
  y = weighted_proportion,
  fill = Source
```

```
)) +
  geom_bar(
    stat = "identity",                    # Kasutame staatilist identiteeti tulbadiagrammi jaoks
    position = position_dodge(width = 0.7),  # Nihutame tulbad üksteise kõrvale
    width = 0.7                           # Tulpade laius
  ) +
  geom_errorbar(
    aes(ymin = lower_bound, ymax = upper_bound),  # Alumine ja ülemine veapiir
    position = position_dodge(width = 0.7),      # Nihutame veapiirid vastavalt tulpadele
    width = 0.2,                          # Veapiiride horisontaaljoonte laius
    color = "black"                       # Veapiirid musta värvi
  ) +
  labs(
    x = NULL,                             # Eemaldame X-telje pealkirja
    y = NULL,                             # Eemaldame Y-telje pealkirja
    fill = "Kaalumine:",                  # Legendi pealkiri
    caption = "Mustaga joonistatud lõigud tähistavad veapiire usaldusnivool 95%"  # Selgitus graafiku all
  ) +
  scale_fill_manual(                      # Kohandame tulpade värve
    values = c(
      "Osakaal audiitorteenuste turust" = "#FF5733",  # Punane
      "Osakaal audiitorettevõtjatest" = "#4CAF50"     # Roheline
    )
  ) +
  scale_y_continuous(
    labels = percent_format(accuracy = 1)         # Y-telg protsentidena
  ) +
  scale_x_discrete(labels = function(x) stringr::str_wrap(x, width = 20)) +  # Teksti mähkimine kahe rea peale
  theme_minimal() +
  theme(
    text = element_text(family = "Times New Roman"),  # Kirjastiil
    panel.background = element_rect(fill = "white", color = "white"),  # Valge taust
    plot.background = element_rect(fill = "white", color = "white"),   # Valge taust
    axis.text.x = element_text(angle = 0, hjust = 0.5, size = 14),     # X-telje sildid sirged
    axis.text.y = element_text(size = 14),                             # Y-telje siltide suurus
    plot.title = element_text(size = 16, hjust = 0.5),                 # Graafiku pealkiri keskele
    plot.caption = element_text(size = 12),                            # Selgituste suurus
    legend.position = "bottom",                                        # Legendi asukoht all
    legend.text = element_text(size = 14),                             # Legendi tekstide suurus
    legend.title = element_text(size = 14),                            # Legendi pealkirja suurus
    panel.grid.major.y = element_line(color = "gray80", size = 0.5),   # Y-telje gridlines
    panel.grid.minor.y = element_blank()                               # Eemaldame väikesed gridlines
  )

# Salvestame graafiku pildifailina
ggsave(
  filename = "24_innovation_specialist.png",  # Faili nimi
  plot = plot_innovation_specialist,          # Graafik
  device = "png",                             # Faili tüüp
  width = 14,                                 # Laius tollides
  height = 8,                                 # Kõrgus tollides
  units = "in",                               # Mõõtühik
  dpi = 300                                   # Resolutsioon
)
######################
# Keskkonna ettevalmistus #
######################

# Laadime vajalikud paketid
if (!require(pacman, quietly = TRUE)) install.packages(pacman)
# Laadime vajalikud paketid
if (!require(pacman)) install.packages("pacman")
pacman::p_load(openxlsx, dplyr)

pacman::p_load(
  dplyr,         # Andmete vormindamine
  readr,         # Andmete lugemine failidest
  tidyr          # Tabelite teisendused
)

#############
# Andmete lugemine #
#############

vastused_kodeeritud <- read_excel("vastused_kodeeritud.xlsx")
```

```
## Loon muutujad, mis sisaldavad vastavate andmete veeru numberid "vastused_kodeeritud" failis
process_innov_fields_ids <- 13:20
weight_ids <- c(132, 136)
suurusklass_id <- 130
confidence_interval_confidence_level <- 0.95

## Kasutades eelnevalt loodud ID-de vektoreid, toon andmed uutesse muutujatesse
data_process_innov_fields <- vastused_kodeeritud[, process_innov_fields_ids]
data_weights <- vastused_kodeeritud[, weight_ids]
data_suurusklass <- vastused_kodeeritud[, suurusklass_id]

# # Arvutan ka raha kaalu (kaal_turg on juba kahe kaalu korrutis andmetes)
# data_weights <- data_weights %>% dplyr::mutate(
#   kaal_raha = Kaal_turg / Kaal_innov
# )

## Teen koondandmestiku kõikide vajalike muutujatega
data <- cbind(
  data_process_innov_fields,
  data_weights,
  data_suurusklass
)

print(data)

# Add labels to the column names
column_labels <- c(
  "14a: PI - teenuste osutamine" = "Teenuste osutamine ehk põhitegevused",
  "14b: PI - logistika" = "Logistika-, tarne- või jaotusmeetodid",
  "14c: PI - IKT" = "Info- ja kommunikatsioonitehnoloogia",
  "14d: PI - raamatupidamine jms" = "Raamatupidamine või ostuprotsess",
  "14e: PI - äriprotsesside korraldus" = "Äriprotsesside või välissuhtluse korraldus",
  "14f: PI - töökorraldus, inimressursid" = "Töökorraldus, inimressursside juhtimine",
  "14g: PI - turundus" = "Turundus: reklaam, teenuse raames väljastatud dokumentide ja materjalide kujundus, hinnakujundus,
müügijärgsed teenused, jne",
  "14h: PI - mitte ükski neist" = "Mitte ükski neist"
)

# Apply labels to the first 12 columns
colnames(data)[1:8] <- unname(column_labels[colnames(data)[1:8]])

# Veendu, et 16. veerg (suurusklass) on korrektselt tuvastatud
suurusklass_column <- colnames(data)[11]

# Arvuta esimese 8 veeru summa, grupeerituna suurusklassi järgi
grouped_sums <- data %>%
  dplyr::select(1:8, !!sym(suurusklass_column)) %>%  # Vali esimesed 8 veergu ja suurusklassi veerg
  dplyr::group_by(!!sym(suurusklass_column)) %>%      # Grupeeri andmed suurusklassi järgi
  dplyr::summarise(across(1:8, sum, na.rm = TRUE))   # Arvuta igas grupis iga veeru summa, ignoreerides NA väärtusi

# Kuvab gruppide summad
print(grouped_sums)

weight_NR_column <- colnames(data)[10]
print(weight_NR_column)

# Arvuta esimese 8 veeru summa korrutades non-response kaaludega ja grupeerituna suurusklassi järgi
NR_weighted_grouped_sums <- data %>%
  dplyr::mutate(across(1:8, ~ . * !!sym(weight_NR_column))) %>%  # Korruta vastused kaaludega
  dplyr::select(1:8, !!sym(suurusklass_column)) %>%          # Vali muudetud veerud ja suurusklass
  dplyr::group_by(!!sym(suurusklass_column)) %>%             # Grupeeri suurusklassi järgi
  dplyr::summarise(across(1:8, sum, na.rm = TRUE))           # Arvuta igas grupis iga veeru summa

# Kuvab kaalutud summad gruppide lõikes
print(NR_weighted_grouped_sums)

# Calculate NR weighted averages and Wilson confidence intervals for columns 1 to 8
process_innov_fields_NR <- data %>%
  dplyr::summarise(across(1:8, ~ {
    # Calculate the weighted sum and n_weighted (sum of weights)
    weighted_sum <- sum(.x * Kaal_innov, na.rm = TRUE)  # Weighted sum
    n_weighted <- sum(Kaal_innov, na.rm = TRUE)         # Effective sample size (sum of weights)
    weighted_proportion <- weighted_sum / 116           # Weighted proportion
    z_value <- qnorm(1 - (1 - confidence_interval_confidence_level) / 2)
```

```
    # Calculate Wilson confidence interval using n_weighted
    lower_bound <- (
      (2 * n_weighted * weighted_proportion + z_value^2) -
        z_value * sqrt(max(0, z_value^2 - (1 / n_weighted) + 4 * n_weighted * weighted_proportion * (1 - weighted_proportion)))) /
      (2 * (n_weighted + z_value^2))
    upper_bound <- (
      (2 * n_weighted * weighted_proportion + z_value^2) +
        z_value * sqrt(max(0, z_value^2 - (1 / n_weighted) + 4 * n_weighted * weighted_proportion * (1 - weighted_proportion)))) /
      (2 * (n_weighted + z_value^2))

    # Return results as a list
    c(proportion_weighted = weighted_proportion, lower_bound = lower_bound, upper_bound = upper_bound)
  }))

# Print the results
print(process_innov_fields_NR)

# Salvesta andmestik Exceli formaadis
write.xlsx(process_innov_fields_NR, "process_innov_fields_NR.xlsx", rowNames = FALSE)
####################
# Keskkonna ettevalmistus #
####################

# Laadime vajalikud R-paketid
if (!require(pacman, quietly = TRUE)) install.packages(pacman)

pacman::p_load(
  ggplot2,       # Graafika genereerimine
  readr,         # Andmete lugemine failidest
  tidyr,         # Tabelite teisendused (pikk vs lai formaat)
  dplyr,         # Andmete töötlemine
  scales,        # Telgede siltide vormindamine (nt protsendid)
  stringr        # Tekstidega töötlemine
)
library(readxl)

#############
# Andmete lugemine #
#############

# # Määrame andmefailide asukohad ja loeme need sisse
# market_weighted_columns_summary_path <- "market_weighted_columns_summary.csv"  # Turu kaalutud andmed
# market_weighted_data <- readr::read_csv(market_weighted_columns_summary_path, show_col_types = FALSE)

NR_weighted_columns_summary_path <- "process_innov_fields_NR.xslx"  # andmed protsessiinnovatsiooni valdkondade kohta
NR_weighted_data <-process_innov_fields_NR <- read_excel("process_innov_fields_NR.xlsx")

# Lisame veeru "Source" (allikas), mis näitab, kummast andmestikust rida pärineb
# market_weighted_data <- market_weighted_data %>%
#   mutate(Source = "Osakaal audiitorteenuste turust")  # Turu kaalutud osakaal

NR_weighted_data <- NR_weighted_data %>%
  mutate(Source = "Osakaal audiitorettevõtjatest")  # Vastamisest kõrvale jäänud osakaal

# Kombineerime mõlemad andmestikud üheks suureks andmetabeliks
combined_data <- bind_rows(NR_weighted_data)
print(combined_data)



# Teisendame andmed pikaks formaadiks
plot_data <- combined_data %>%
  pivot_longer(
    cols = -Source,  # Kõik veerud peale "Source"
    names_to = "Category",
    values_to = "Proportion"
  ) %>%

  # Lisame veeru "Metric_Type", et eristada kaalutud keskmisi, alumisi ja ülemisi veapiire
  mutate(Metric_Type = rep(c("Weighted_Mean", "Lower_Bound", "Upper_Bound"), each = 8, times = 1)) %>%
  # Teisendame veerud uuesti laia formaati
  pivot_wider(names_from = Metric_Type, values_from = Proportion)  %>%
  # Kohandame kategooriate sildid (lisame reavahetused)
  mutate(Category = case_when(
```

```
    str_detect(Category, "Turundus") ~ "Turundus: reklaam, teenuse raames väljastatud\ndokumentide ja materjalide
kujundus,\nhinnakujundus, müügijärgsed teenused jne",
    TRUE ~ Category
  ))

print(plot_data)

# Loome horisontaalse tulpdiagrammi koos veapiiridega
process_innov_fields <- ggplot(plot_data, aes(
  y = reorder(Category, Weighted_Mean),  # Nimetused Y-teljel
  x = Weighted_Mean,                                    # Kaalutud osakaal X-teljel
  fill = Category                                       # Värvid vastavalt allikale
)) +
  geom_bar(
    stat = "identity",                  # Kasutame staatilist identiteeti tulbadiagrammi jaoks
    position = position_dodge(width = 0.7),  # Nihutame tulbad allika järgi
    width = 0.7                         # Tulpade laius
  ) +
  geom_errorbar(
    aes(xmin = Lower_Bound, xmax = Upper_Bound),  # Alumine ja ülemine veapiir
    position = position_dodge(width = 0.7),      # Nihutame veapiirid vastavalt tulpadele
    width = 0.2,                        # Veapiiride horisontaaljoonte laius
    color = "black"                       # Veapiirid musta värvi
  ) +
  labs(
    x = NULL,                           # Eemaldame X-telje pealkirja
    y = NULL,                           # Eemaldame Y-telje pealkirja
    fill = "Kaalumine:",                  # Legendi pealkiri
    caption = "Mustad jooned tähistavad veapiire usaldusnivool 95%"  # Selgitus graafiku all
  ) +
  scale_fill_manual(                      # Kohandame tulpade värve
    values = c(
      "Mitte ükski neist" = "#FFC107",  # Kollane
      "Teenuste osutamine ehk põhitegevused" = "#4CAF50",                  # Roheline
      "Info- ja kommunikatsioonitehnoloogia" = "#4CAF50",    # Roheline
      "Töökorraldus, inimressursside juhtimine" = "#4CAF50",                  # Roheline
      "Raamatupidamine või ostuprotsess" = "#4CAF50",
      "Äriprotsesside või välissuhtluse korraldus" = "#4CAF50",                  # Roheline
      "Turundus: reklaam, teenuse raames väljastatud\ndokumentide ja materjalide kujundus,\nhinnakujundus, müügijärgsed teenused jne" =
"#4CAF50",      # Roheline
      "Logistika-, tarne- või jaotusmeetodid" = "#4CAF50"  # Roheline

    )
  ) +
  scale_x_continuous(
    labels = percent_format(accuracy = 1)      # X-telg protsentidena
  ) +
  theme_minimal() +
  theme(
    text = element_text(family = "Times New Roman"),  # Kirjastiil
    panel.background = element_rect(fill = "white", color = "white"),  # Valge taust
    plot.background = element_rect(fill = "white", color = "white"),   # Valge taust
    axis.text.x = element_text(size = 16),                          # X-telje tekstide suurus
    axis.text.y = element_text(size = 18, lineheight = 1.),          # Y-telje tekstide suurus ja reavahe
    plot.title = element_text(size = 16, face = "bold", hjust = 0.5),  # Pealkiri
    plot.caption = element_text(size = 14),                          # Selgituste suurus
    legend.position = "none",                                  # Legendi asukoht all
    # legend.text = element_text(size = 14),                          # Legendi tekstide suurus
    # legend.title = element_text(size = 14),                          # Legendi pealkirja suurus
    panel.grid.major.x = element_line(color = "gray80", size = 0.5),   # X-telje gridlines
    panel.grid.minor.x = element_blank()                          # Eemaldame väiksed gridlines
  )

# # Kuvame graafiku
print(process_innov_fields)



# Salvestame graafiku
ggsave(
  filename = "26_process_innov_fields.png",  # Faili nimi
  plot = process_innov_fields,
  width = 12,
  height = 8,
  units = "in",
```

```
  dpi = 300
)
library(ggplot2)
library(dplyr)
library(patchwork)

# Andmete loomine
andmed <- read_excel("specialist_vs_innovation.xlsx")

#   data.frame(
#   respondent_size_id = c(1, 1, 2, 2, 3, 3, 99, 99),  # Vastajate suurusklassid
#   has_innovation_specialist = c(FALSE, TRUE, FALSE, TRUE, FALSE, TRUE, FALSE, TRUE),  # Kas innovatsioonispetsialist on
olemas
#   n = c(4, 3, 10, 11, 2, 4, 16, 18),  # Vastajate arv
#   innovation_probability = c(1.45, 2.58, 3.00, 3.25, 1.95, 3.52, 2.25, 2.97),  # Innovatsioonitõenäosus
#   innovation_probability_confidence_interval_start = c(1.29, 2.04, 2.80, 3.00, 1.35, 3.13, 2.12, 2.71),  # Usaldusvahemiku algus
#   innovation_probability_confidence_interval_end = c(1.62, 3.11, 3.20, 3.49, 2.56, 3.91, 2.38, 3.24)  # Usaldusvahemiku lõpp
# )

# Lisame selgitavad sildid
andmed <- andmed %>%
  mutate(
    specialist_label = ifelse(has_innovation_specialist, "On innovatsioonitöötaja", "Ei ole innovatsioonitöötajat"),
    respondent_size_label = factor(
      respondent_size_id,
      levels = c(1, 2, 3, 99),
      labels = c("Väike", "Keskmine", "Suur", "Kokku")
    )
  )

# Innovatsioonitõenäosuse võrdlus (täielik skaala)
innovatsioonitõenäosus_graafik <- ggplot(andmed, aes(
  x = innovation_probability,  # Innovatsioonitõenäosus X-teljel
  y = respondent_size_label,  # Suurusklassid Y-teljel
  color = specialist_label  # Värvid vastavalt sellele, kas spetsialist on olemas
)) +
  geom_point(size = 3) +  # Lisame punktid
  geom_errorbarh(aes(xmin = innovation_probability_confidence_interval_start, xmax = innovation_probability_confidence_interval_end),
height = 0.2) +  # Veapiirid
  labs(
    x = NULL,  # Eemaldame X-telje pealkirja
    y = "Suurusklass",  # Y-telje pealkiri
    color = "Innovatsioonispetsialist:",  # Legendi pealkiri
    title = "AT-de kasutuselevõtmise tõenäosuse keskmine hinnang suurusklasside lõikes",  # Pealkiri
    caption = "Lõigud tähistavad veapiire usaldusnivool 95%"  # Selgitav tekst
  ) +
  scale_color_manual(
    values = c(
      "On innovatsioonitöötaja" = "#2196F3",  # Sinine
      "Ei ole innovatsioonitöötajat" = "#FF9800"  # oranž
    )
  ) +
  scale_x_continuous(
    limits = c(1, 5),  # Näitame kogu skaalat (1-5)
    breaks = c(1, 2, 3, 4, 5),  # X-telje väärtused
    labels = c("1 Väga eba-\ntõenäoline", "2 Eba-\ntõenäoline", "3 Ei oska\nöelda", "4 Tõenäoline", "5 Väga\ntõenäoline")
  ) +
  theme_minimal() +
  theme(
    text = element_text(family = "Times New Roman"),
    axis.text.x = element_text(size = 16),
    axis.text.y = element_text(size = 16),
    axis.title.y = element_blank(),
    plot.title = element_text(size = 20, hjust = 0.5),
    legend.position = "bottom",
    legend.text = element_text(size = 16),
    legend.title = element_blank(),
    panel.background = element_rect(fill = "white"),
    plot.background = element_rect(fill = "white")
  )

# Proportsioonide graafik suurusklasside lõikes (horisontaalne, "Kokku" tagasi)
proportsioon_graafik <- ggplot(andmed, aes(
  y = respondent_size_label,  # Suurusklassid Y-teljel
  x = n,  # Vastajate arv X-teljel
```

```
  fill = specialist_label  # Värvid vastavalt sellele, kas spetsialist on olemas
)) +
  geom_bar(stat = "identity", position = "fill", width = 0.7) +  # Lisame täidetud tulbad
  scale_x_continuous(labels = scales::percent_format()) +  # X-telg protsentidena
  labs(
    x = "Osakaal vastajatest (%)",  # X-telje pealkiri
    y = "Suurusklass",  # Y-telje pealkiri
    fill = "Innovatsioonispetsialist:",  # Legendi pealkiri
    title = "Vastajate osakaal"  # Pealkiri
  ) +
  scale_fill_manual(
    values = c(
      "On innovatsioonitöötaja" = "#2196F3",  # Sinine
      "Ei ole innovatsioonitöötajat" = "#FF9800"  # oranž
    )
  ) +
  theme_minimal() +
  theme(
    text = element_text(family = "Times New Roman"),
    axis.text = element_text(size = 16),
    axis.title.y = element_blank(),
    axis.title.x = element_blank(),
    plot.title = element_text(size = 20, hjust = 0.5),
    legend.position = "bottom",
    legend.text = element_text(size = 16),
    legend.title = element_blank(),
    panel.background = element_rect(fill = "white"),
    plot.background = element_rect(fill = "white")
  )

# Kombineerime graafikud
kombineeritud_graafik <- innovatsioonitõenäosus_graafik / proportsioon_graafik +
  plot_annotation(
    title = element_blank(), # "Innovatsioonispetsialisti mõju innovatsioonitõenäosusele",
    theme = theme(
      plot.title = element_text(size = 18, hjust = 0.5)
    )
  )

# Salvestame graafiku
ggsave(
  filename = "27_specialist_vs_innovation_prob.png",
  plot = kombineeritud_graafik,
  width = 12,
  height = 8,
  units = "in",
  dpi = 300
)

# Kuvame graafiku
print(kombineeritud_graafik)
######################
# Keskkonna ettevalmistus #
######################

# Laadime vajalikud paketid
if (!require(pacman, quietly = TRUE)) install.packages(pacman)

pacman::p_load(
  dplyr,         # Andmete vormindamine
  readr,         # Andmete lugemine failidest
  tidyr          # Tabelite teisendused
)

#############
# Andmete lugemine #
#############

vastused_kodeeritud <- read_excel("vastused_kodeeritud.xlsx")

## Loon muutujad, mis sisaldavad vastavate andmete veeru numberid "vastused_kodeeritud" failis
main_reason_source_ids <- 29
weight_ids <- c(132, 136)
suurusklass_id <- 130
confidence_interval_confidence_level <- 0.95
```

```
## Kasutades eelnevalt loodud ID-de vektoreid, toon andmed uutesse muutujatesse
data_main_reason_source <- vastused_kodeeritud[, main_reason_source_ids]
data_weights <- vastused_kodeeritud[, weight_ids]
data_suurusklass <- vastused_kodeeritud[, suurusklass_id]

# # Arvutan ka raha kaalu (kaal_turg on juba kahe kaalu korrutis andmetes)
# data_weights <- data_weights %>% dplyr::mutate(
#   kaal_raha = Kaal_turg / Kaal_innov
# )

## Teen koondandmestiku kõikide vajalike muutujatega
data <- cbind(
  data_main_reason_source,
  data_weights,
  data_suurusklass
)

print(data)


# Samm 1: Loenda vastuste arv suurusklasside lõikes
vastuste_arv <- data %>%
  group_by(`17: Rohkemate innnovatsioonide puudumise põhjus`, `19: Suurus käibe alusel`) %>%
  summarise(vastuste_arv = n())


print (vastuste_arv)


########################## NR kaaluga

# Samm 2: Arvutame kaalutud osakaalud ja Wilsoni usalduspiirid
kaalutud_tulemused_NR <- data %>%
  group_by(`17: Rohkemate innnovatsioonide puudumise põhjus`) %>%
  summarise(
    weighted_sum = sum(Kaal_innov, na.rm = TRUE),  # Kaalutud summa
    n_weighted = sum(Kaal_innov, na.rm = TRUE),   # Kaalutud vastuste arv
    .groups = "drop"
  )  %>%
  mutate(
    weighted_proportion = weighted_sum / 116,  # Kaalutud osakaal populatsioonis
    z_value = qnorm(1 - (1 - confidence_interval_confidence_level) / 2),  # z väärtus usaldusnivool
    sqrt_term = sqrt(pmax(0, z_value^2 - (1 / n_weighted) + 4 * n_weighted * weighted_proportion * (1 - weighted_proportion))),
    lower_bound = (
      (2 * n_weighted * weighted_proportion + z_value^2 - z_value * sqrt_term) /
        (2 * (n_weighted + z_value^2))
    ),
    upper_bound = (
      (2 * n_weighted * weighted_proportion + z_value^2 + z_value * sqrt_term) /
        (2 * (n_weighted + z_value^2))
    )
  ) %>%
  mutate(
    lower_bound = pmax(0, lower_bound),  # Piirame alumise piiri 0-ga
    upper_bound = pmin(1, upper_bound)   # Piirame ülemise piiri 1-ga
  )

# Tulemuste kuvamine
print(kaalutud_tulemused_NR)



# Tulemus salvestatakse vajadusel CSV-failina
write.csv(kaalutud_tulemused_NR, "main_reason_for_no_more_innovations_NR.csv", row.names = FALSE)




########################## MARKET WEIGHT


# Samm 2: Arvutame kaalutud osakaalud ja Wilsoni usalduspiirid
kaalutud_tulemused_market <- data %>%
```

```
  group_by(`17: Rohkemate innnovatsioonide puudumise põhjus`) %>%
  summarise(
    weighted_sum = sum(Kaal_turg, na.rm = TRUE),  # Kaalutud summa
    n_weighted = sum(Kaal_turg, na.rm = TRUE),   # Kaalutud vastuste arv
    .groups = "drop"
  ) %>%
  mutate(
    weighted_proportion = weighted_sum / 116,  # Kaalutud osakaal populatsioonis
    z_value = qnorm(1 - (1 - confidence_interval_confidence_level) / 2),  # z väärtus usaldusnivool
    sqrt_term = sqrt(pmax(0, z_value^2 - (1 / n_weighted) + 4 * n_weighted * weighted_proportion * (1 - weighted_proportion),1)),
    lower_bound = (
      (2 * n_weighted * weighted_proportion + z_value^2 - z_value * sqrt_term) /
        (2 * (n_weighted + z_value^2))
    ),
    upper_bound = (
      (2 * n_weighted * weighted_proportion + z_value^2 + z_value * sqrt_term) /
        (2 * (n_weighted + z_value^2))
    )
  ) %>%
  mutate(
    lower_bound = pmax(0, lower_bound),  # Piirame alumise piiri 0-ga
    upper_bound = pmin(1, upper_bound)   # Piirame ülemise piiri 1-ga
  )

# Tulemuste kuvamine
print(kaalutud_tulemused_market)

sum(kaalutud_tulemused_market$n_weighted)
colSums(kaalutud_tulemused_NR[2:8])

# Tulemus salvestatakse vajadusel CSV-failina
write.csv(kaalutud_tulemused_market, "main_reason_for_no_more_innovations_market.csv", row.names = FALSE)
#####################
# Keskkonna ettevalmistus #
#####################

# Laadime vajalikud R-paketid
if (!require(pacman, quietly = TRUE)) install.packages(pacman)

pacman::p_load(
  ggplot2,       # Graafika genereerimine
  readr,         # Andmete lugemine failidest
  tidyr,         # Tabelite teisendused (pikk vs lai formaat)
  dplyr,         # Andmete töötlemine
  scales,        # Telgede siltide vormindamine (nt protsendid)
  stringr        # Tekstidega töötlemine
)

#############
# Andmete lugemine #
#############

# Määrame andmefailide asukohad ja loeme need sisse
market_weighted_columns_summary_path <- "main_reason_for_no_more_innovations_market.csv"  # Turu kaalutud andmed
market_weighted_data <- readr::read_csv(market_weighted_columns_summary_path, show_col_types = FALSE)

NR_weighted_columns_summary_path <- "main_reason_for_no_more_innovations_NR.csv"  # Vastamisest kõrvale jäänud ettevõtete
kaalutud andmed
NR_weighted_data <- readr::read_csv(NR_weighted_columns_summary_path, show_col_types = FALSE)

# Lisame veeru "Source" (allikas), mis näitab, kummast andmestikust rida pärineb
market_weighted_data <- market_weighted_data %>%
  mutate(Source = "Osakaal audiitorteenuste turust")  # Turu kaalutud osakaal

NR_weighted_data <- NR_weighted_data %>%
  mutate(Source = "Osakaal audiitorettevõtjatest")  # Vastamisest kõrvale jäänud osakaal

# Kombineerime mõlemad andmestikud üheks suureks andmetabeliks
combined_data <- bind_rows(market_weighted_data, NR_weighted_data)

print(combined_data)

# Graafiku loomine
plot_main_reason_for_no_more_innovations <- ggplot(combined_data, aes(
  x = `17: Rohkemate innnovatsioonide puudumise põhjus`,
```

```
  y = weighted_proportion,
  fill = Source
)) +
  geom_bar(
    stat = "identity",                    # Kasutame staatilist identiteeti tulbadiagrammi jaoks
    position = position_dodge(width = 0.7),  # Nihutame tulbad üksteise kõrvale
    width = 0.7                           # Tulpade laius
  ) +
  geom_errorbar(
    aes(ymin = lower_bound, ymax = upper_bound),  # Alumine ja ülemine veapiir
    position = position_dodge(width = 0.7),      # Nihutame veapiirid vastavalt tulpadele
    width = 0.2,                          # Veapiiride horisontaaljoonte laius
    color = "black"                       # Veapiirid musta värvi
  ) +
  labs(
    x = NULL,                             # Eemaldame X-telje pealkirja
    y = NULL,                             # Eemaldame Y-telje pealkirja
    fill = "Kaalumine:",                  # Legendi pealkiri
    caption = "Mustaga joonistatud lõigud tähistavad veapiire usaldusnivool 95%"  # Selgitus graafiku all
  ) +
  scale_fill_manual(                      # Kohandame tulpade värve
    values = c(
      "Osakaal audiitorteenuste turust" = "#FF5733",  # Punane
      "Osakaal audiitorettevõtjatest" = "#4CAF50"     # Roheline
    )
  ) +
  scale_y_continuous(
    labels = percent_format(accuracy = 1)         # Y-telg protsentidena
  ) +
  scale_x_discrete(labels = function(x) stringr::str_wrap(x, width = 20)) +  # Teksti mähkimine kahe rea peale
  theme_minimal() +
  theme(
    text = element_text(family = "Times New Roman"),  # Kirjastiil
    panel.background = element_rect(fill = "white", color = "white"),  # Valge taust
    plot.background = element_rect(fill = "white", color = "white"),   # Valge taust
    axis.text.x = element_text(angle = 0, hjust = 0.5, size = 18),     # X-telje sildid sirged
    axis.text.y = element_text(size = 18),                             # Y-telje siltide suurus
    plot.title = element_text(size = 20, hjust = 0.5),                 # Graafiku pealkiri keskele
    plot.caption = element_text(size = 16),                            # Selgituste suurus
    legend.position = "bottom",                                        # Legendi asukoht all
    legend.text = element_text(size = 20),                             # Legendi tekstide suurus
    legend.title = element_text(size = 20),                            # Legendi pealkirja suurus
    panel.grid.major.y = element_line(color = "gray80", size = 0.5),   # Y-telje gridlines
    panel.grid.minor.y = element_blank()                               # Eemaldame väikesed gridlines
  )

# Salvestame graafiku pildifailina
ggsave(
  filename = "29_main_reason_for_no_more_innovations.png",  # Faili nimi
  plot = plot_main_reason_for_no_more_innovations,          # Graafik
  device = "png",                         # Faili tüüp
  width = 14,                             # Laius tollides
  height = 6,                             # Kõrgus tollides
  units = "in",                           # Mõõtühik
  dpi = 300                               # Resolutsioon
)
```

Summary

ADOPTION OF EMERGING TECHNOLOGIES – INNOVATION IN ESTONIAN AUDIT FIRMS

Robert Pirk

The use of emerging technologies (AT) such as artificial intelligence, robotic process automation, advanced data analytics, and other applications in the audit industry has recently developed rapidly and has the potential to significantly change the industry in the future (Seidenstein jt, 2024). Changes are possible in the workforce structure, knowledge and skills required by audit firms (AEV), and the process of providing audit services (Palekhova & Kramarenko, 2020; Seidenstein jt, 2024; Vitali & Giuliani, 2024). The audit services market, as a highly regulated sector, differs from the general services, which also affects the process of adopting ATs (Krieger jt, 2021) and makes this area of specific interest for research. There is currently a shortage of auditors in Estonia and this is likely to persist in the coming years (Õun, 2024), which is why it is appropriate to explore opportunities to increase the efficiency of existing AEVs - the introduction of ATs offers an opportunity for this (Krieger jt, 2021; Ruhnke, 2023; Zemankova, 2019; Vitali & Giuliani, 2024). The introduction of ATs in Estonian AEVs has not been studied much before, the main one being Mikhaylova's (2024) master's thesis, which dealt with the possibilities of introducing artificial intelligence.

The aim of this study is to assess innovation in Estonian audit firms through the introduction of emerging technologies and to provide recommendations for promoting this process. To achieve this, the following research tasks have been completed. AEVs have been defined as the object of study, the specific features and activities of these companies have been explained, and the areas in which ATs can be introduced in AEVs have been identified. An overview of the scientific literature on the process of innovation and adoption of ATs and the factors influencing it in AEVs has been prepared. In accordance with the guidelines of the Oslo Manual (OECD/Eurostat, 2018) and based on the information obtained from the literature review, the research methodology and questionnaire have been developed. The theoretical framework of the study is the principles set out in the Oslo Manual (OECD/Eurostat, 2018) on the key indicators of innovativeness of enterprises and the process theory for the adoption of technologies in AEVs developed by Krieger jt (2021). A multiple-choice survey was conducted among Estonian AEVs (116 in total), to which 34 responses were received, and the results were analysed mainly using quantitative descriptive statistics in

the software R and RStudio. The data analysis methodology includes the use of weights so that the results can be generalized to the population of Estonian AEVs and separately to the audit services market. The latter generalization provides an idea of how the adoption of ATs is perceived by the population of clients of Estonian audit services. To assess the reliability of the results, error limits have been calculated at a confidence level of 0.95 and figures have been prepared to express the results. Finally, the results have been discussed, and recommendations have been formulated for promoting the process of introducing ATs in Estonian AEVs.

The results of the study show that the main indicators of innovativeness of AEVs differ from most Estonian companies in a significantly higher proportion of companies with only process innovation, which is probably related to the high level of regulation in the field. Process innovations have been carried out by most AEVs, and these innovations have mainly been developed by the AEVs themselves. Innovations have not been implemented more mainly due to the need or lack of resources. The introduction of ATs is not yet widespread among AEVs, especially among small AEVs. In general, the results show several differences between larger and smaller AEVs in terms of factors influencing the process of introducing ATs. In general, larger AEVs are more optimistic about ATs. The main obstacle to the introduction of ATs is the lack of knowledge, which may be amplified by the fact that approximately half of the AEVs do not have any employees who would have a specific job task to deal with innovation.

Based on the results, the final part of the study provides several recommendations for the participants of the Estonian audit services market to promote the process of introducing ATs. The main recommendations are related to increasing knowledge about ATs and promoting cooperation between market participants themselves and with external parties, especially in the case of smaller AEVs, to reduce obstacles caused by the lack of resources and ICT knowledge. Helping small AEVs catch up is important to reduce their lag in introducing ATs, as this may reduce their competitiveness in the future.

There are limitations to the design of the study that open further research opportunities. The quantitative approach does not allow for the identification of new technologies that have not been discussed in previous literature and sources – this could be supplemented with studies based on interviews or open-ended questionnaires. There were 34 responses, which is a good result considering the total number of Estonian AEVs (116), but which is still insufficient for many statistical analyses, which is why the study had to be limited to mainly descriptive statistics and accept large margins of error. Due to the study

schedule, the response period was also very short – a longer period could have yielded somewhat more responses. To minimize the impact of this limitation, the respondents were contacted repeatedly, including by telephone. In the future, a similar study could be conducted in a larger region, including as a study covering several countries. In this case, it would also be possible to obtain enough responses to compile statistical models and would open the opportunity to investigate whether there are regional differences in the main indicators of AEVs' innovativeness and the process of introducing ATs. Studies with a similar methodology would also allow for time series analysis of changes over time in the adoption of emerging technologies. The fact that the author is a certified auditor himself, and a competitor of the research subjects limited the level of detail of the questions and was also an important reason for using a quantitative approach. Future studies could also focus on which specific AT-based solutions are in use in Estonian AEVs.

Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks

Mina, Robert Pirk,

annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose „Arenevate tehnoloogiate kasutuselevõtt - innovatsioon Eesti audiitorettevõtjates“, mille juhendajad on Ulvi Sloog ja Urmas Varblane, reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada digitaalarhiivi DSpace kuni autoriõiguse kehtivuse lõppemiseni.

Annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi DSpace kaudu Creative Commonsi litsentsiga CC BY NC ND 3.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, kuni autoriõiguse kehtivuse lõppemiseni.

Olen teadlik, et punktides 1 ja 2 nimetatud õigused jäävad alles ka autorile.

Kinnitan, et lihtlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

Robert Pirk
16.01.2025