

TALLINNA TEHNIKAÜLIKOOL
Infotehnoloogia teaduskond
Informaatikainstituut

IDX70LT

Rauno Kulla 143686IAPM

**MASSANDMETE PÄRINGUTE
TÖÖTLEMISE AUTOMATISEERIMINE**

Magistritöö

Juhendaja: Jaak Henno

PhD

Dotsent

Tallinn 2016

Autori deklaratsioon

Kinnitan, et olen koostanud antud lõputöö iseseisvalt ning seda ei ole kellegi teise poolt varem kaitsmisele esitatud. Kõik töö koostamisel kasutatud teiste autorite tööd, olulised seisukohad, kirjandusallikatest ja mujalt pärinevad andmed on töös viidatud.

Autor:

Kuupäev:

Annotatsioon

Massandmete päringu töötlemise automatiseerimine

Käesolevas magistritöös analüüsiti ja seejärel realiseeriti veebilahendusena alternatiivne meetod massandmete päringu töötlemisele, mis oli seni Äriregistris toimunud ainult emailide vahendusel. Analüüsimisel kaardistati eelnevalt sissetulnud kirjalikud massandmete päringud, millest leiti tüüp-andmeskoobid ja –kriteeriumid ning nende tugi lisati veebirakendusse. Selleks, et vältida võimalikke tekkivaid probleeme käideldavusega, tehti massandmete päringu käivitamiseks sobiva hetke leidmine süsteemi ülesandeks.

Lahendus loodi dünaamiliseks, et lihtsustada tulevikus täiendavate võimaluste loomist. Mitmeid muudatusi süsteemi loogikas on võimalik teostada rakenduse lähtekoodi muutmata. Realiseeritud funktsionaalsuse testimisel selgus, et lahendus on efektiivne ka ekstreemolukorras, mil päringus töödeldavate andmete maht on väga suur. Arvestades, et Äriregistri infosüsteemid on arendatud Python'i programmeerimiskeeles, realiseeriti ka massandmete päringu koostamise rakendus Python'is.

Lõputöö on kirjutatud eesti keeles ning sisaldab teksti 72 leheküljel, 7 peatükki, 23 joonist, 4 tabelit.

Abstract

Automating the Process of Mass Data Querying

The necessity of this work was raised from the need to find an optimal solution for processing mass data queries in the Estonian Commercial Register. So far the process had solely been handled by business correspondences. Criteria and data scopes that could be interpreted as typical were detected by analysing queries that had been submitted by clients via previous method. Thereupon, the functionality for these criteria and data scopes was designed and developed as a web application. The created solution was added into the Central Commercial Register as new functionality.

In the developed software mass data queries were designed to be asynchronous. In order to avoid potential problems with system availability the task of finding suitable moment for starting queries was given to the system. Additionally, that decision resulted in the scenario of aborted queries becoming impossible.

The solution was developed to be as dynamic as possible to simplify additions that will be made in the future. Many adjustments with the system's logic can be done without any further developments. Instead, it is suffice just to add or change data records in database tables that were created in course of this thesis.

While testing the developed software, it became apparent that the solution is capable of delivering results even in extreme circumstances when the amount of data that being processed is huge. Additionally it was concluded, that it is not going to be an issue for the system to deliver results while staying in the boundaries of the established requirements even in the future when the amount of data increases.

Because all of the Commercial Register's software systems including the Central Commercial Register have been developed in Python programming language, the solution for querying mass data was also developed in Python.

The thesis is in Estonian and contains 72 pages of text, 7 chapters, 23 figures, 4 tables.

Lühendite ja mõistete sõnastik

Ajax	<p><i>Ajax</i></p> <p>„Asünkroonne JavaScript ja XML Lühend AJAX võeti kasutusele 2005.a. ja sellega tähistatakse selliste interaktiivsete veebirakenduste loomise meetodit, kus toimub jooksev kulissidetagune andmevahetus brauseri ja veebiserveri vahel, nii et kasutaja iga liigutuse peale pole vaja kogu veebilehte uuesti alla laadida.“ [1]</p>
Andmeskoop	<p><i>Scope of data</i></p> <p>Andmeskoobiks nimetakse käesoleva töö raames komplekti ühes ja samas andmebaasi tabelis paiknevatest andmeveergudest, mis andmeskoobi valimise puhul kaasatakse tulemusse. Tulemusse mittekaasatavad veerud, mis asuvad samas tabelis, ei kuulu andmeskoobi koosseisu. Töö üleselt kasutatakse andmeskoobi kohta ka andmestiku mõistet.</p>
Kriteerium	<p><i>Criterion</i></p> <p>Kriteeriumiks on kitsenduse täpsustatud väärtus, millele peavad tulemusse lisatavad andmed vastama. Töö üleselt on kasutatud kriteeriumi kohta ka kitsenduse mõistet.</p>
Massandmete päring	<p><i>Query of mass data</i></p> <p>Massandmete päringuks nimetakse toimingut, millega saadakse andmebaasist andmed, mis vastavad päringus seatud üldistavatele tüüp-kriteeriumitele. Üldistavate tüüp-kriteeriumite all mõistetakse tingimusi, millele vastavad mitu ettevõtet. Kuna päringu vastuses sisalduvate andmete hulk on enamasti suur nimetataksegi selliseid päringuid tulemuste kvantiteedi tõttu massandmete päringuteks.</p>

Proksi	<i>Proxy server</i> „Välisliiklust vahendav tulemüüri komponent. Kui kasutaja veebilehitseja pöördub tulemüüri kaitstud veebiserveri poole, siis proksi asub veebilehitseja ja selle veebiserveri vahel. Ta püüab kinni veebiserverile saadetud päringud ja kontrollib, kas ta saab neile päringuile ise vastata. Kui mitte, siis edastab proksi päringu veebiserverile.“ [1]
Tasuta kasutaja	<i>Free user</i> Äriregistri teabesüsteemi kasutaja tüüp, kellel on õigus Äriregistri andmeid tasuta pärida. Tasuta kasutajateks on riigiasutuste, kohalike omavalituste ning RIKi töötajad.
Tasuline kasutaja	<i>Chargeable user</i> Äriregistri teabesüsteemi kasutaja tüüp, kelle teostatavad päringud kuuluvad tasustamisele. Sellised kasutajad on eraisikud ning erasektori ettevõtete, kes on RIK'iga partnerlus lepingu sõlminud, töötajad.
Triger	<i>Trigger</i> „Tegevus, mis põhjustab protseduuri automaatse käivitumise, näit. viitetervikluse säilitamiseks. Triger käivitub, kui kasutaja üritab andmebaasis sisalduvaid andmeid muuta lisamise (<i>insert</i>) , kustutamise (<i>delete</i>) või värskendamise (<i>update</i>) käsu abil. Triger võib anda süsteemile käsu sooritada mistahes arv tegevusi, kui üritatakse teha teatud kindlat muudatust. Vältides andmete ebakorrektsust, autoriseerimata või süsteemitut muutmist aitavad trigerid hoida andmebaasi terviklust“ [1]
Äriregistri teabesüsteem	<i>Central Commercial Register</i> Äriregistri teabesüsteem on keskkond, mille kaudu infosüsteemi kasutaja pääseb ligi äriregistri andmebaasis olevatele andmetele. Ligi pääsemise all peetakse silmas, et kasutajal on võimalik ettevõtete andmeid näha ning osta. Ettevõtete andmeid modifitseerida ei ole teabesüsteemi kaudu võimalik.

Jooniste loetelu

Joonis 1. Teabesüsteemi arhitektuuri mudel.	17
Joonis 2. Kasutusjuhtude diagramm.	34
Joonis 3. Andmebaasi kontseptuaalset struktuuri kirjeldav olemi-suhte diagramm	41
Joonis 4. Päringu seisundidiagramm	44
Joonis 5. Loogilise disaini andmebaasi diagramm (UML notatsioonis).	45
Joonis 6. Päringu töötluse algoritmi tegevusdiagramm.	48
Joonis 7. Teabesüsteemi autentimise vaade.	51
Joonis 8. Massandmete päringu asukoht.	52
Joonis 9. Massandmete päringu koostamise vaade	53
Joonis 10. Massandmete päringu koostamise info modaalaken	54
Joonis 11. Päringu kinnitamisel emaili sisestamine	55
Joonis 12. Õigusliku vormi kriteeriumina täpsustamise vaade.	56
Joonis 13. Aadressi kriteeriumina täpsustamise vaade.	57
Joonis 14. Tegevusala kriteeriumina täpsustamise vaade.	58
Joonis 15. Selgitav pilt kuidas jäetakse vahele tasemed, millel puuduvad valikud.	58
Joonis 16. Esmakande aja kriteeriumina täpsustamise vaade.	59
Joonis 17. Registrikoodide nimekirja kriteeriumina lisamise vaade.	60
Joonis 18. Ettevõtte oleku kriteeriumina täpsustamise vaade.	61
Joonis 19. Tellitud päringute ülevaate asukoht.	61
Joonis 20. Kasutajaga seotud massandmete päringu nimekiri.	62
Joonis 21. Ekstreemumolukorra kriteeriumi valikud.	64
Joonis 22. Ekstreemumolukorra testimise koostatud päringu vaade.	65
Joonis 23. Ekstreemumolukorra testi tulemuste faili suurus.	66

Tabelite loetelu

Tabel 1. Cartesiuse korrutise tulemuse näide.	21
Tabel 2. Olemitüüpide definitsioonide kirjeldused.	42
Tabel 3. Andmebaasi lisatavate tabelite kirjeldused.	46
Tabel 4. Ekstreemumolukorra testimise tulemuse metaandmed.	65

Sisukord

Sissejuhatus	11
1. Hetkeolukorra kirjeldus	13
1.1 Kehtiva töökorralduse head küljed	13
1.2 Kehtiva töökorralduse probleemid	14
1.3 Kehtiva töökorralduse parandamise võimalused	15
1.4 Alternatiivsed lahendused.....	16
2. Teoreetiline taust	17
2.1 Arhitektuurilised piirangud.....	17
2.2 Päringute hinnastamine ja tasumine	18
2.3 Päringute teostamine.....	19
2.4 Cartesiuse korrutise vältimine	21
2.5 Tulemuste salvestamine.....	22
3. Süsteemi kavandamine ja realiseerimine.....	23
3.1 Ülesande püstitus	23
3.2 Kriteeriumite ja andmeskoopide valiku analüüs	24
3.2.1 Kriteeriumide kaardistamine	24
3.2.2 Andmeskoopide kaardistamine	26
3.3 Funktsionaalsed nõuded	27
3.4 Kasutusjuhtude mudel	33
3.5 Mittefunktsionaalsed nõuded.....	39
3.6 Kontseptuaalne andmemudel.....	40
3.6.1 Olemi suhte diagramm	40
3.6.2 Olemitüüpide definitsioonid	42
3.7 Päringu seisundidiagramm	44
3.8 Andmebaasi lisanduvad tabelid	44
3.8.1 Tabelite kirjeldus	45
3.9 Kontroll päringu loogika sisestuse üle.....	47
3.10 Päringu töötlemise algoritm.....	47
4. Kasutajaliides	50

4.1 Massandmete päringu funktsionaalsuse asukoht.....	50
4.2 Massandmete päringu koostamine.....	52
4.3 Päringu kriteeriumite täpsustamiste vaated	55
4.3.1 Õigusliku vormi kriteeriumina täpsustamine	55
4.3.2 Aadressi kriteeriumina täpsustamine.....	56
4.3.3 Tegevusala kriteeriumina täpsustamine	57
4.3.4 Esmakande kuupäeva kriteeriumina täpsustamine	58
4.3.5 Registrikoodide kriteeriumina täpsustamine	59
4.3.6 Ettevõtte oleku kriteeriumina täpsustamine	60
4.4 Ülevaade tellitud päringutest	61
5. Ekstreemumolukorra testimine.....	63
5.1 Ekstreemumolukorra testimiseks sobiva päringu koostamine.....	63
5.2 Ekstreemumolukorra testimise tulemused.....	65
5.3 Ekstreemolukorra testimise järeldused	66
6. Planeeritud edasiarendused.....	67
7. Kokkuvõte	68
Summary.....	70
Kasutatud kirjandus	72
Lisa 1	73
Lisa 2	79
Lisa 3	89

Sissejuhatus

Eesti Äriregister (edaspidi Äriregister) on alates 2007. aastast pakkunud võimalust pärida Äriregistri andmebaasis sisalduvaid andmeid Internetis. Siiani on olnud võimalik näha andmeid ainult ettevõtte põhisel. Paralleelselt teabesüsteemi käigus hoidmisele on Äriregister pakkunud taustal klientidele võimalust pärida andmeid ka massiliselt vastavalt kliendi vajadustele.

Massiliselt andmete pärimise (edaspidi massandmete päringu) all mõistetakse olukorda, milles klient omab soovi kindlasse andmeskoopi või andmeskoopidesse kuuluvate andmete järele, mille tulemused peavad lisaks vastama ka tema poolt seatud tingimustele. Sellistele tingimustele, ehk kriteeriumitele vastab enamasti suur hulk andmeid, mistõttu nimetataksegi Äriregistris taolisi päringuid massandmete päringuteks.

Tänapäevani on massandmete päringute teenindamise tööprotsess käinud eranditult emailide saatmise vahendusel. Klient pöördub Registrate ja Infosüsteemide Keskuse (edaspidi RIK) klienditeeninduse poole emaili teel ning esitab oma soovi andmete saamiseks. Seejärel edastatakse soov Äriregistri talituse halduri kätte, kes teostab päringu alusel andmebaasis SQL päringu ning edastab emailiga tulemused kliendile. Tänapäevase meetodiga kulub massandmete päringute töötlemisele 25% kogu Äriregistri talituse halduskoormusest.

Käesoleva diplomitöö ülesandeks on realiseerida veebilahendusena alternatiiv hetkel kasutatavale aegunud teguviisile, millega raisatakse Äriregistri talituse halduskoormust. Uue lahenduse loomiseks analüüsitakse tänapäevase meetodiga sissetulnud massandmete päringuid, millest kaardistatakse tüüp-kriteeriumid ning -andmeskoobid. Leitud tüüp-kriteeriumite ning -andmeskoopide järgi massandmete päringu teostamise võimekus lisatakse töö tulemusena Äriregistri teabesüsteemi.

Käesoleva magistratöö esimeses peatükis tutvustatakse hetkeolukorda massandmete päringute täitmisel ning mooduseid, kuidas on võimalik protsessi parandada. Samuti tehakse ülevaade alternatiivsetest lahendustest, millest järeldub, et käesoleva töö

kirjutamise ajal tarbijal enamasti puuduvad veebilahenduse vormis alternatiivid emaili teel teostatavatele ettevõtete massandmete päringutele. Teises peatükis kirjeldatakse teoreetilisi aluseid, millega loodud tarkvara arendamise käigus arvestati. Kolmanda peatüki algul loetletakse esmalt eesmärgid, mida töö tulemusena soovitakse saavutada ning seejärel kirjeldatakse realiseeritud lahendust erinevate mudelite abil. Neljandas peatükis esitatakse kasutajaliidese funktsionaalne prototüüp ning kirjeldatakse kogu loodud kasutajaliidese funktsionaalsust. Viiendas peatükis testitakse realiseeritud tarkvara ekstreemumolukorras ning tehakse sellest järeldused. Kuuendas tutvustatakse täiendavaid arendusi, mida süsteemiga plaanitakse teostada ning seitsmendas peatükis tehakse loodud tarkvara kohta kokkuvõtlikud järeldused.

1. Hetkeolukorra kirjeldus

Järgnevas peatükis selgitatakse Äriregistrist massandmete pärimise hetkeolukorda, selle probleeme ning mooduseid nende parandamiseks. Lisaks selgitatakse olemasolevaid võimalusi Eesti ettevõtete massandmete pärimiseks veebilahenduste kaudu.

Äriregistri teabesüsteemi kaudu on kõigil soovijatel võimalik tutvuda Äriregistrisse kantud juriidiliste isikute kehtivate ning ajalooliste andmetega. Teabesüsteemi kaudu on kasutajal võimalik andmeid mugavalt vaadata ettevõtete põhiselt, valides esmalt ettevõtte ning seejärel konkreetsed andmed mida näha soovitakse. Massiliste andmete päringut ei ole Äriregistri teabesüsteem seni toetanud.

Äriregistris talletatavaid andmeid soovivad tellida erinevad isikud ning organisatsioonid. Soovides massiliselt andmeid, pöörduakse kirjalikult emailiga Registrate ja Infosüsteemide Keskuse, edaspidi RIK, poole ning kirjeldatakse milliste tingimustele peavad andmed vastama. Vajaminevaid andmeid kirjeldab klient massandmete päringu esitamiseks loodud malli kasutades. [2]

Sissetulevate kirjalike päringutega tegelemine võtab Äriregistri talituses ühel halduril kogu töömahu. Töönädalas tuleb keskmiselt sisse 50 päringut. Hinnanguliselt on 80% kirjalike päringute puhul on tegemist tüüppäringutega, millistele ei ole vajalik koostada SQL päring algusest lõpuni, vaid piisab juba eelnevalt valmis tehtud päringutes tingimuste väärtuste muutmisest.

1.1 Kehtiva töökorralduse head küljed

Hetkel kasutatava töökorralduse kohaselt on masspäringute juures alati kliendi ning süsteemi vahel Äriregistri töötaja, kes vajadusel aitab kliendil jõuda sihini täpsustades kriteeriume või selgitades võimalike tulemuste skoopi. Olukorras, kus klient ise ei oska sõnastada oma vajadusi on võimalik pöörduda telefoni või emaili teel Äriregistri talituse poole saamaks konsultatsiooni võimaluste kohta. Samuti jätab Äriregistri halduri poolne

isiklik tagasiside kliendile riigiga suhtlemisest meeldivama mulje kui infosüsteemis orienteerumine.

Äriregistri andmed on talletatud keeruka struktuuriga andmebaasidesse. Sellest tulenevalt luues veebirakenduse, mille kaudu oleks võimalik täita kõik tellimused mida klient teostada soovib, on selle kasutamine klientide jaoks hoomamatu. Kuna erinevaid kriteeriumite ning andmeskoopide valikuid oleks kasutaja jaoks liiga palju, et oleks võimalik nende erinevustest selgelt aru saada. Kui praegu tuleb sisse päring, mille sarnast varem ei ole olnud, siis haldur peab lihtsalt hindama, kas soovi on võimalik täita ning seejärel päringu valmis kirjutama.

Massandmete päringud on enamasti andmebaasile kurnavad. Sellised päringud võtavad palju aega ning täitmise ajal mälu. Kuid kuna päringuid käivitab Äriregistri töötaja, siis on tagatud kontroll selle üle, et taolisi päringuid andmebaasi samaaegselt suurel hulgal ei jõua. Vastav olukord võiks lõppeda andmebaasi mälu täitumisega ning see tähendaks omakorda, et kõikide rakenduste, mis sama andmebaasi kasutavad, töö hanguks. Sel juhul oleks Äriregistri talituses tegu lühiajalise kriisiga, seni kuni baasiserver saab taas käivitatud. Seega on hetkel kasutusel oleva töökorraldusega kirjeldatud potentsiaalse ohtliku sündmuse toimumise tõenäosus väike.

1.2 Kehtiva töökorralduse probleemid

Meetod, kuidas Äriregistris täidetakse sissetulevaid kirjalikke tellimusi, on aegunud ning sellega kaasnevad mitmed negatiivsed aspektid nii kliendile kui ka teenusepakkuja – RIK'i seisukohast. Järgnevalt on välja toodud praeguse töökorraldusega kaasnevad mured kliendi jaoks ning kohad, mis sisaldavad arenguruumi Äriregistri vaatenurgast:

- Klient soovib andmeid koheselt või vähemalt samal päeval. Tavaliselt on halduril laual mitu päringut, mida ta teostab järjekorras, seetõttu esineb tihti situatsioon, kus kliendi päring saab vastuse alles viiendal päeval. Erasektoris on sagedane olukord, kus viiendaks päevaks on juba vajadus andmeid saada möödunud, mille tõttu ei ole klient enam andmete tellimisest huvitatud ning seega jääb RIK ilma tulust, mida oleks saanud päringu kiirema töötlemise eest.

- Sageli esineb kirjalike päringute puhul mitmekordne edasi-tagasi kirjavahetus, sest klient esimene kord andmeid soovides ei tea täpselt kuidas küsida või jätab mõne vajaliku kriteeriumi edastamata. Seetõttu peab Äriregistri haldur küsima kliendilt täpsustusi enne kui päringu saab käivitada. Kirjavahetusega kaasneb mõlema osapoole jaoks tarbetu ajakulu.
- Kasutades seesugust aegunud teguviisi massandmete päringute täitmiseks kulutab Äriregistri talitus igas kuus ühe halduri terve töökoormuse. Arvestades, et Äriregistri talituses on neli halduri kohta moodustab masspäringute pakkumine Äriregistri halduskoormusest tervelt veerandi.

1.3 Kehtiva töökorralduse parandamise võimalused

Eelnevalt klientide päringuid analüüsid kindlaks tehtud tüüp-kriteeriumide järgi massiliselt andmete pärimise võimaluse lisamine teabesüsteemile, parandab protsessi nii RIK'i kui ka klientide jaoks mitmel moel:

- Kaob suur osa hetkel kirjalikest sissetulevatest päringutest halduri töölaualt, sest need kliendid saaks suunata kasutama rakenduse võimalusi. Järelejäänud klientide, kes soovivad väga spetsiifilisi andmeid, mida ei ole võimalik ka loodava rakenduse kaudu tellida, hulk on nii väike, et nende päringute vastamisele ei teki mitme päevast viidet. Seega oleks minimeeritud kliendi jaoks soovitud tulemuse saavutamise aeg.
- Vabaneb hulga halduri töömahtu. Eelnevalt on välja toodud, et Äriregistri talituses nelja halduri töömahust kulub hetkeolukorras kirjalike päringute vastamisele terve ühe halduri töömaht ning hinnanguliselt moodustab tüüppäringute osakaal kõikidest sisse tulevatest kirjalikest massandmete päringutest 80%, siis kogu talituse haldusmaht väheneb 20% võrra. Seda vabanevat ressursi saab hakata kasutama teiste igapäevaste Äriregistri talituse halduri tööülesannete täitmise peale.
- Teabesüsteemi kasutab käesoleval hetkel (märts 2016) päevas keskmiselt 20860 kasutajat. Lisades massandmete päringu funktsionaalsuse teabesüsteemi, suureneb selle võimaluse olemasolust teadlikkus Äriregistri andmete kasutajate

seas. Suurem teadlikkus toob kaasa mõningase kasutatavuse tõusu ning see omakorda RIK-i omatulu kasvu.

1.4 Alternatiivsed lahendused

Ajal, mil diplomitöö kirjutati, ei paku Äriregister ise võimalust massiliselt andmeid pärida veebirakenduse kaudu, seega kirjeldatakse järgnevalt millised on hetkel klientide võimalused veebi kaudu Eesti ettevõtete andmeid massiliselt tellida.

Ainus Eesti ettevõtete kohta masspäringuid võimaldav teadaolev süsteem on Krediidiinfo AS poolt loodud otseturunduse andmebaas: „Baltic Target“. See on veebirakendus mille kaudu on kasutajatel võimalik osta ette määratud kriteeriumite alusel ettevõtete kontaktandmeid. Lisaks Eestis paiknevatele, on võimalik tellida ka teistes Baltikumi riikides asuvate ettevõtete kontakte. [3]

Baltic Target'i headeks külgedeks on kindlasti kriteeriumite küllus, mille alusel saab valimit paika panna. Negatiivseteks külgedeks võrreldes planeeritava lahendusega on tarbija jaoks asjaolu, et süsteemi kaudu on võimalik pärida ainult ettevõtete kontaktandmeid, planeeritav süsteem seevastu hakkab pakkuma tüüp andmeskoopidena käsitletavaid andmeid. Teisalt on probleemiks asjaolu, et Eesti ettevõtete puhul on võimalus, et vastuses sisaldub aegunud andmeid, kuna Krediidiinfo saab oma andmed Äriregistrilt 15 päevases nihkes.

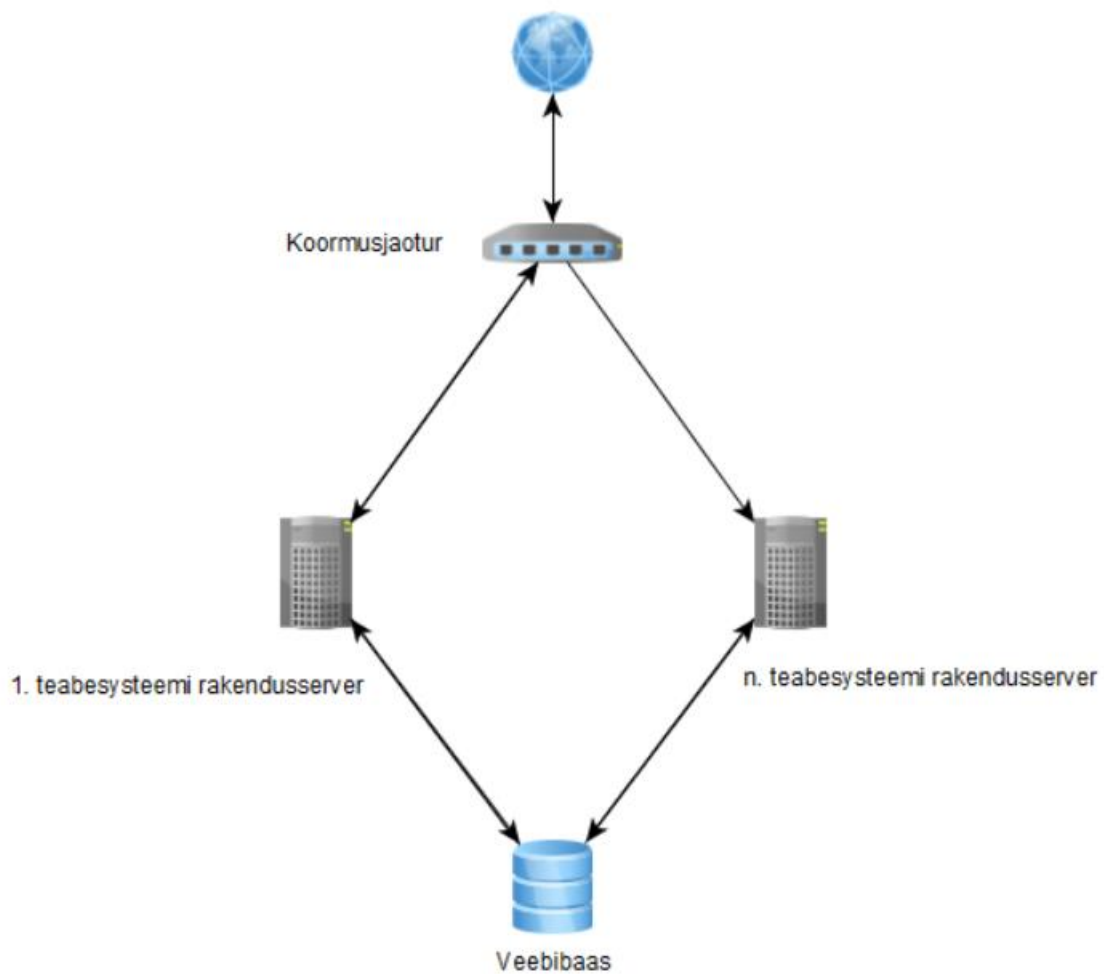
Teisi süsteeme, mis Eesti ettevõtete andmeid massiliselt pakuks, ei ole. Seega käesoleval hetkel tarbijal, kes soovib seesuguseid andmeid, on võimatu oma soovi otse veebikeskkonnas täita, vaid ta on sunnitud pöörduma kirjalikult mõne asutuse või ettevõtte poole.

2. Teoreetiline taust

Järgnevas peatükis käsitletakse teoreetilisi aluseid, mida käesoleva magistritöö raames loodud tarkvara realiseerimisel kasutati ning millega arvestati.

2.1 Arhitektuurilised piirangud

Äriregistri teabesüsteemist on koormuse ühtlasema jaotamise ja käideldavuse parandamiseks alati aktiivsed mitu rakendusserverit. Seda, millise rakendusserveriga ühendub iga kasutaja otsustab koormusjaotur, nagu näha jooniselt 1. [4]



Joonis 1. Teabesüsteemi arhitektuuri mudel.

Joonisel on näha, et kliendi päring suundub esmalt koormusjaoturile, mis suunab päringu edasi ühte aktiivsesse teabesüsteemi rakendusserverisse. Seejärel server vajadusel ühendub andmebaasiga ning pärib vajaminevat. Koormusjaotur saates päringu serverisse, jääb ootama vastust, mida kliendile tagasi saata. Vastuse ootamisele on seatud ette maksimaalne ajaline piir, ehk *timeout*, milleks on üks minut. Seda ületades jaotur teavitab klienti vastuse tulemata jäämisest. Ajalise limiidi eesmärgiks on vähendada ohtu, et koormusjaoturisse jääksid püsima samaaegselt pikad päringud, mis võivad lõppeda mälu täitumisel kogu teabesüsteemi hangumisega.

Kuna koormusjaoturi *timeout* on ette seatud aeg, mida pikendada ei ole mõistlik, on välistatud võimalus, et läbi teabesüsteemi saaks kasutaja masspäringuid otse andmebaasi saata. Suurte andmemahtude tõttu saaksid enamus päringuid koormusjaoturit läbides katkestatud.

2.2 Päringute hinnastamine ja tasumine

Hinnastamise alla kuuluvad kõik valitavad andmeskoobid. Hind võib jaguneda kolme erineva liigi vahel:

1. Hinnastatud kirjepõhiselt.
2. Hinnastatud objektipõhiselt.
3. Kuulub hinnastamiselt teise andmeskoobi objekti koosseisu.

Kirjepõhine hinnastamine tähendab seda, et kõik andmeobjekti kirjed eraldi on tasulised. Näiteks kui ühel ettevõttel on kaks seotud isikut, siis hinnastamise alla kuuluvad mõlemad seotud isikud eraldi seotud isikute andmeskoobi hinnaga.

Objektipõhine hinnastamine tähendab seda, et kõik andmeobjekti kirjed on hinnastatud korraga ühekordselt. Eelnevas näites, kui ühel ettevõttel on kaks seotud isikut, siis hinnastatakse mõlemad isikud ühekordselt seotud isikute andmeskoobi hinnaga.

Kolmanda variandina – kuulub hinnastamiselt teise andmeskoobi koosseisu – tähendab seda, et eelnevas näites ei hinnastata kumbagi tulemust, vaid nad kuuluksid hinnastamiselt ettevõtte andmete hinna sisse.

Päringute hinnastamine sõltub kasutajast, kes päringu tellib. Kasutajaks võib olla mõne riigiasutuse töötaja, siis on tegu tasuta kasutajaga ning temale on päringu teostamine tasuta. Tasuliste kasutajate puhul, kelleks on teabesüsteemi tasuline kasutaja ning eraisikust kasutaja, on tarvis hind aktsepteerida enne päringu tulemustele ligi pääsemist. Teabesüsteemi tasuliste kasutajate puhul saab kasutaja tasuda päringute eest kuulõpus tema lepingupartnerile esitatava arve alusel. Eraisikust kasutajal on vajalik tasuda päringu eest pangalingi kaudu enne tulemustele ligipääsemist.

Veebi kaudu massandmete päringu teostamise võimaldamine on Registrite ja Infosüsteemide Keskuse jaoks uus äriteenus, mille hinnastamine on RIK-is üldosakonna ülesandeks. Ajal, mil diplomitöö kirjutati, on märgitud teenuse hinnastamise paikapanek üldosakonna äriteenuse halduri tööplaani ning planeeritav valmimise aeg on 2016. aasta sügis. Seega, ei ole hetkel teada hinnastamise reeglid, mistõttu jääb selle realiseerimine diplomitöö skoobist välja, kuid siiski on plaanis ettenägelikult luua hiljem hinnastamise implementatsiooni lisamiseks andmebaasi poolne tugi hinnastamiseks vajalike andmete salvestamiseks.

2.3 Päringute teostamine

Nagu kirjeldatud arhitektuuriliste piirangute peatükis, ei ole võimalik lubada päringuid teostada veebirakendusest otse andmebaasi. Seega on tarvis magistritöö raames realiseerida asünkroonne päringute teostamise lahendus, kus kasutaja valib esmalt omale vajaminevad andmed ning paneb paika kriteeriumid, seejärel kinnitab päringu, mis tegelikult koheselt ei käivitu, vaid käivitatakse süsteemi poolt hiljem.

Päringuid on võimalik käivitada taustaprotsessi kaudu, kasutades Äriregistri serverites kasutusel olevas Linuxi operatsiooni süsteemis sisalduvat vahendit – *cron*. Kasutades *cron*'i säilib kontroll - samaaegselt käivituvate päringute üle - süsteemil. Automaatselt käivitavas programmis on võimalik paika seada mitu massandmete päringut paralleelselt käivitub ning seda väärtust on võimalik hiljem muuta vastavalt sellele, kas päringud saavad jooksvalt täidetud või mitte. Lastes aga päringuid käivitada kasutajatel otse rakendusest, sõltuks samaaegselt käivitavate päringute arv hoopis kasutajatest, mis ei ole soovitud tulemus.

Cron'i käivitab *crontab* fail, mis määrab paika pandud ajakava järgi käivitatavad ülesanded või programmid. Olemasolevasse teabesüsteemi *crontab* faili on vajalik lisada rida, mis hakkab kasutaja soovitud päringuid käivitama. Lisaks on vajalik lisada *crontab* faili teine rida, tulemuste faili kustutamise jaoks. Lisatavas reas tuleb määrata käivitamise ajad ning käivitatava programmi asukoht.

Käivitamise aja määramine käib järgneva formaadi kohaselt:

MI HH DD MM DW asukoht

Käesolevas töös realiseeritava projekti raames on vaja käivitada taustaprotsess võimalikult tihti, eesmärgiga anda kasutajale vastus ilma mõttetute ajaliste viideteta. Teisalt on taustaprotsessi poolt teostatavad päringud mahukad, enamasti läheb taoliste päringute vastamisele mitmeid minuteid. Vältides mõttetut arvuti ressursi kasutust on hinnanguliselt sobivaim vahemik, millal päringuid käivitada 3 minutit, ehk taustaprotsess peaks uuesti käima minema iga päev, kolm minutit pärast eelneva taustaprotsessi algust. Sellistel aegadel käivitab taustaprotsessi järgnev lause:

```
1,4,7,10,13,16,19,22,25,28,31,34,37,40,43,46,49,52,55,58 * * * * *  
/data/teabesysteem/cronid/mass/teosta_paring.py
```

Tulemuste faili kustutamiseks kasutatav taustaprotsess peab käivituma kord päevas. Selle ülesande teostuseks on valitud ajaks 02:30. Sellisel ajal käivitab tulemuste kustutamise taustaprotsessi järgnev lause:

```
20 2 * * * /data/teabesysteem/cronid/mass/kustuta_tulemused.py
```

Ülal kirjeldatud käskudes tähendab tärn, et seda aega ei täpsustata ehk igal võimalikul ajal. [5]

Võttes arvesse, et *cron*'iga sarnane funktsionaalsus on olemas ka Windows Server operatsiooni süsteemis, nagu näiteks *CRONw* [6] või *pycron* [7], on võimalik käesolevas töös realiseeritavat masspäringute teostamise lahendust kasutada automaatselt käivitatava taustaprotsessina ka Windows Serveri operatsiooni süsteemiga.

2.4 Cartesiuse korrutise vältimine

Lastes kasutajal erinevaid andmeskoope valides koostada tema soovitud tulemust, tuleb süsteemi poolt jälgida, et valikusse ei lisataks sellist andmeskoopide kombinatsiooni, mis annaks tulemuseks Cartesiuse korrutise. Cartesiuse korrutis tähendab igat võimalikku kombinatsiooni elementidest, milles on kaks või rohkem rida. Viidatud tulemus tekib olukorras kui teostatakse päring, milles on mitu andmeskoopi, mis omavad ettevõtte andmetega seost üks-mitmele.

Näitena on ühe ettevõttega seotud isikud Rauno Kulla ning Mari Maasikas ning selle sama ettevõttega seotud emaili aadressid info@rik.ee, mari@rik.ee ning rauno@rik.ee, valides tulemuste sekka nii seotud isikute kui ka emaili aadresside andmeskoobid, saame tulemuseks kuus rida sellisel kujul nagu kuvatud tabelis 1. [8]

Tabel 1. Cartesiuse korrutise tulemuse näide.

Ettevõtte nimi	Seotud isik	Email
Näite ettevõtte OÜ	Rauno Kulla	info@rik.ee
Näite ettevõtte OÜ	Rauno Kulla	mari@rik.ee
Näite ettevõtte OÜ	Rauno Kulla	rauno@rik.ee
Näite ettevõtte OÜ	Mari Maasikas	info@rik.ee
Näite ettevõtte OÜ	Mari Maasikas	mari@rik.ee
Näite ettevõtte OÜ	Mari Maasikas	rauno@rik.ee

Kuna tulemuste hinna arvestus käib eelkõige tulemuse ridade arvu pealt, siis Cartesiuse korrutis rikub hinna arvestuse loogika. Teisalt ei ole ka kasutajal soovi saada sellist tulemust, mis koosneb erinevatest andmeskoopide tulemuste kombinatsioonidest ning millega on vajalik teha lisatoiminguid, et tulemus kasulik oleks. Seega on vajalik kehtestada rakenduses selline nõue, et oleks võimalik valida igas päringus ainult ühte sellist andmeskoopi, mis oleks ettevõtte andmetega seoses üks-mitmele.

Samas peab olema päringu koostamise algoritmile arusaadav, et vajadusel teeks andmeskoobi tegelikust üks-mitu seosest ettevõtte andmetega - üks-ühele seose, sidudes tabeli ainult kõige viimase sobiva kirjega. Lisaks peab algoritmile olema eristatav olukord, kui mitme tabeli tulemused ühendada omavahel funktsiooni kaudu. Äriregistris andmebaasis paiknevad kandeisikud ning kandevälised isikud erinevates tabelites ning kõigi seotud isikute vastusesse saamisel Cartesiuse korrutise vältimiseks, on vajalik

ühendada ettevõttega seotud kirjed mõlemast tabelist funktsiooni abil ning seejärel tulemus siduda teiste andmeskoopide valikutega.

2.5 Tulemuste salvestamine

Päringutest saadud vastused pannakse *csv* formaadis faili ning kompressitakse *zip* failiks. Tulemuseks olev fail omakorda salvestatakse koos mõningate metaandmetega päringu kohta andmebaasi. Faili jaoks võetakse kasutusele Äriregistris kasutusel oleva andmebaasihalduri *PostgreSQL* andmetüüp *bytea*.

Bytea on suuteline salvestama kuni 1 GB suuruseid faile. Suuremate failide salvestatakse päringu tulemused mitmesse faili, sedasi, et iga faili suurus jääb alla 1 GB. Alternatiivina kaalutud andmetüüpi *OID* ei võeta kasutusse, kuna sellega kaasneb rohkem negatiivseid külgi kui positiivseid. Negatiivseteks külgedeks on, et andmeobjekt ei kustu lihtsalt rea kustutamise, samuti ei oleks fail tervikuna kättesaadav, vaid seda tuleks käsitleda jadana. *OID* eeliseks *bytea* ees on selle võimekus talletada kuni 2GB sisu, kuna faili suuruse probleemi lahendus on juba käsitletud, ei anna *OID* andmetüübi kasutamine siinkohal eelist. [9]

Arvestades, et massilistelt andmete pärimisel on vastusteks suured failid ning kui süsteemi kasutades neid faile pidevalt lisandub, toob see endaga kaasa andmebaasi suuruse kasvu. Baasi suuruse kasvuga kaasnevad erinevad miinused. Neist kõige olulisemaks pean andmebaasi varunduse teostamise ajakulu märgatavat kasvu. Seetõttu on vajalik pärast päringu tulemuste kasutamist või mõistliku aja, mis on kasutajatega kokkulepitud, möödumist tulemuse fail andmebaasist kustutada igal ööl käivituva taustaprotsessi abil. Seda tööd hakkab igaõiselt teostama peatükis 2.3 kirjeldatud tulemuste faile andmebaasist kustutav *cron*.

3. Süsteemi kavandamine ja realiseerimine

Järgnevas peatükis püstitatakse esmalt diplomitöö eesmärgid. Seejärel määratakse tüüpkrriteeriumid ning tüüpandmestikud analüüsides sissetulnud päringuid. Järgnevalt kirjeldatakse süsteemi loomiseks vajalikud mudelid ja diagrammid.

3.1 Ülesande püstitus

Järgnevalt on välja toodud eesmärgid mida käesoleva töö tulemusena soovitakse saavutada.

Töö peamiseks ning lõppeesmärgiks on massandmete päringute teostamise võimaluse lisamine olemasolevale Äriregistri andmemüügi ning -pärimise platvormile, milleks on Äriregistri teabesüsteem. Käesoleva töö kirjutamise ajal on teabesüsteem koht, kust kasutajad vastavalt oma kasutajatüübile saavad pärida või osta ettevõtete andmeid iga ettevõtte kohta eraldi. Massandmete päringuid teenindab Äriregistri haldur ning suhtlemine toimub kirjalikult emaili teel. Peamise eesmärgi tulemusena soovitakse enim tellitavad massandmed teha kättesaadavaks Internetis teabesüsteemi kaudu.

Sekundaarseks eesmärgiks on kaardistada kriteeriumid, mille alusel on vajalik lasta teabesüsteemi kasutajatel panna paika just neile vajaminevad andmed. Kaardistamine toimub analüüsides eelnevalt sissetulnud massandmete päringuid. Analüüsimisel selgitatakse välja kriteeriumid mille järgi on kliendid kõige rohkem andmete valimit piiranud ning iga tingimuse kohta tuvastatakse kas seda on võimalik veebirakenduses realiseerida.

Eelneva eesmärgi osaks on ka kaardistamine, milliseid andmeid kõige enam massiliselt tellitakse. Teostus toimub samal meetodil. Saadav tulemus on juba iseenesest Äriregistrile vajalik, kuna annab ülevaatliku pildi, mis andmeid ning mille järgi tellitakse, hetkel on vastav selgus ainult halduril, kes massandmete päringutega tegeleb. Alameesmärgina võib siinkohal välja tuua, et andmeskoobid, mida realiseeritav süsteem

võimaldab pärida, peab olema laialdasem, kui hetkel olemasolevas konkureerivas süsteemis – Baltic Target.

Kolmandaks eesmärgiks on anda kontroll massandmete päringute käivitamise üle süsteemile. Teisisõnu ei tohiks anda teabesüsteemi kasutajale võimalust selliseid päringuid baasis alustada. Andes päringu käivitamise õiguse süsteemile, saab reeglite alusel määrata millal on optimaalne aeg päringut teostada. Tulemusena minimeeritakse võimalus massandmete päringuga süsteem hanguma lasta, seega ei teki probleeme käideldavusega.

Arvestades, et Äriregistri teabesüsteem on loodud Python'i programmeerimiskeeles, on vajalik loodav lahendus samuti realiseerida Python'is.

3.2 Kriteeriumite ja andmeskoopide valiku analüüs

Kriteeriumeid, ehk milliseid tingimusi peab olema võimalik täpsustada ning andmeskoobid ehk mis andmeid peab olema võimalik tulemusele lisada, pandi paika analüüsides alates 2016. aasta algusest sissetulnud klientide kirjalikke massandmete päringuid. Analüüsimisele kulus aega kokku 40 tundi.

3.2.1 Kriteeriumide kaardistamine

Kriteeriumide kaardistamisel analüüsiti sissetulnud päringuid eesmärgiga teha kindlaks, milliste tingimuste alusel on kliendid siiani andmeid soovinud. Päringuid analüüsides leiti tüüp-kriteeriumid, mida kõige enim kliendid soovivad. Kuna klientide soovid võivad varieeruda suurel mahul kattes sisuliselt kogu Äriregistri andmebaasi võimalused, siis esines ka soove, mida ei ole võimalik tüüp-kriteeriumina käsitleda ning automaatselt töödelda.

Kaardistamisest järeldus, et luues veebipõhise lahenduse, mis toetab leitud seitset tüüp-kriteeriumi on võimalik 82% sissetulevaid massandmete päringuid läbi selle töödelda. Ülejäänud osa jääks endiselt Äriregistri halduri ülesandeks.

Tüüp-kriteeriumiteks kujunesid järgmised tingimused:

- Nimekiri registrikoodidest: Klient andis oma nimekirja csv failis temale huvipakkuvatest registrikoodidest ning soovis sellele lisaks täiendavaid

andmeid. Kõikidest analüüsitud massandmete päringutest esitati 10% juhtudel enda registrikoodide nimekiri. Kuna kriteerium leidis kasutust ning sellise kriteeriumi alusel päringut teostada ei muuda päringut aeglaseks, siis on otsustatud rakendusele kriteeriumi tugi luua.

- Ettevõtte olek: Klient täpsustas mis olekus ettevõtteid vastusesse soovis. Kõikidest analüüsitud massandmete päringutest sooviti 4% juhtudel täpsustada ettevõtte olekut. Kuigi küsijate hulk oli väike, ainult 4%, ei muuda kriteeriumi lisamine päringut aeglaseks, mistõttu otsustati kriteeriumi tugi luua.
- Õiguslikud vormid ning õigusliku vormi alaliigid: Klient täpsustas millise õigusliku vormi või õigusliku vormi alaliigiga ettevõtete andmeid soovis. Õigusliku vormi ning selle alaliiki on mõistlik käsitleda ühe kriteeriumina. Kokku täpsustati õigusliku vormi või alaliigi vastet 40% kõikidest analüüsitud massandmete päringutest, ehk tegu oli väga sageli esineva kriteeriumiga, seega on vajalik rakendusele tugi luua.
- Tegevusala: Klient täpsustas, millise tegevusalaga ettevõtete andmeid soovis. Tegevusala oli täpsustatud 24% kõikidest analüüsitud massandmete päringutest, mis on piisav alus rakendusele tegevusala täpsustamise tugi kriteeriumina luua.
- Majandusaasta aruanne: Klient täpsustas, et soovis andmeid ettevõtete kohta, kellel on majandusaasta aruanne esitamata kindla aasta kohta. Kuna seda kriteeriumit oli täpsustatud ainult 2% kõikidest analüüsitud päringutest ning toe lisamine muudaks päringut suurusjärg aeglasemaks, sest vastuseid tuleb otsida ühest kõige suuremast tabelist Äriregistri andmebaasis, siis kriteeriumi tuge rakendusele ei looda.
- Aadress: Klient täpsustas, millises haldusüksuses paiknevate ettevõtete andmeid soovis. Tegu oli kõige enim küsitud kriteeriumiga, mida täpsustati 52% kõikidest analüüsitud päringutest. Seega on vajalik rakendusele aadressi määramise tugi kriteeriumina luua.
- Esmakande kuupäev: Klient täpsustas, mis ajast alates sisse kantud ettevõtteid tulemusse soovis. Esmakande kuupäeva täpsustati 14% kõikidest analüüsitud

massandmete päringutest, mis on piisav maht, et otsustati rakendusse kriteeriumi tugi luua.

Ühe kriteeriumi - majandusaasta aruanne - toe lisamine jääb käesoleva magistritöö skoobist välja, seega on võimalik töödelda 80% sissetulevaid massandmete päringuid kasutades töö skoopi jäänud kriteeriumeid. Kriteeriumid järjestatakse rakenduses vastavalt nende täpsustamise sagedusele analüüsitud massandmete päringute hulgas.

3.2.2 Andmeskoopide kaardistamine

Andmeskoopide kaardistamisel analüüsiti samu sissetulnud päringuid nagu kriteeriumite kaardistamiselgi. Seekord oli eesmärgiks teha kindlaks mis andmeskoope, ehk andmestike on kliendid siiani soovinud. Erinevalt kriteeriumide kaardistamisest, analüüsides andmestikke leidus, et kõik kliendid küsisid selliseid andmestike, mida võib käsitleda tüüpandmestikuna.

Tüüp-andmestikeks kujunesid järgmised andmeskoobid:

- Ettevõtte andmed: Klient soovis tulemusse saada andmeid ettevõtte kohta. Ettevõtte andmete koosseisu kuuluvad: ettevõtte nimetus, registrikood, õiguslik vorm, ettevõtte olek, aadress, postikood ja kapital. Ettevõtte andmete andmestikku kuuluvaid andmeid küsiti kokku 90% analüüsitud massandmete päringutes. Tegu on kõige enim küsitava andmestikuga, ehk kindlasti tuleb rakendusele lisada andmestiku valimise võimalus.
- Käibemaksukohuslase number: Klient soovis tulemusse saada lisaveergu ettevõtte käibemaksukohuslase numbri kohta. Kuigi andmestiku sooviti ainult 4% analüüsitud massandmete päringutes, ei muutu päring selle andmestiku toe lisamisel aeglasemaks, mistõttu otsustati rakendusele andmestiku valimise võimalus luua.
- Seotud isikute andmed: Klient soovis tulemusse saada andmeid ettevõttega seotud isikute kohta. Seotud isikute andmete koosseisu kuuluvad: isiku eesnimi, isiku nimi, isiku kood ja roll ettevõttes. Seotud isikute andmestiku koosseisu kuuluvaid andmeid sooviti kokku 16% analüüsitud massandmete päringutes, mis on piisav alus rakendusele andmestiku valimise toe lisamiseks.

- Aastakäive ja töötajate arv: Klient soovis tulemusse andmeid majandusaasta aruande näitajate kohta. Aastakäivet ning töötajate arvu käsitletakse koos, kuna andmed kuuluvad mõlemad majandusaasta aruande näitajate koosseisu ning hinnastatakse ühekordselt. Majandusaasta aruande näitajate koosseisu kuuluvad: töötajate arv, käive, kasum, majandusaasta algus ja majandusaasta lõpp. Andmestikku kuuluvaid andmeid küsiti kokku 34% analüüsitud massandmete päringutes, mis on piisav alus rakendusele andmestiku valimise toe lisamiseks.
- Kontaktandmed kõik: Klient soovis tulemusse kõikide sidevahendi liikide andmeid. Sidevahendite andmete koosseisu kuuluvad: sidevahendi liik, sidevahend. Andmestiku kuuluvaid andmeid küsiti 22% analüüsitud massandmete päringutes, mis on piisav alus rakendusele andmestiku valimise toe lisamiseks.
- Kontaktandmed - email: Klient soovis tulemusse sidevahenditest ainult emaili tüüpi andmeid. Emaili sidevahendite andmete koosseisu kuuluvad: sidevahendi liik, sidevahend. Andmestiku küsiti 14% analüüsitud massandmete päringutes, mis on piisav alus eraldi emaili valiku toe lisamiseks.
- Tegevusala: Klient soovis tulemusse saada lisaveergudena andmeid ettevõtte tegevusala kohta. Tegevusala andmete koosseisu kuuluvad: tegevusala kood, tegevusala nimetus. Andmestiku küsiti 18% analüüsitud massandmete päringutes, mis on piisav alus tegevusala andmestiku valimise toe lisamiseks.

Valitavad andmeskoobid järjestatakse rakenduses vastavalt nende soovimise sagedusele analüüsitud massandmete päringute hulgas.

3.3 Funktsionaalsed nõuded

1. **Massandmete päringu teostamise võimalus konfigureeritavaks** – Administraatoril peab olema võimalik rakenduse konfiguratsioonifailist massandmete päringu funktsionaalsust sisse ja välja lülitada. Välja lülitades on vajalik, et rakenduses ei oleks funktsionaalsust võimalik kasutada.
2. **Samaaegselt käivitavate päringute arv konfigureeritavaks** – Kuna on tegu uue süsteemiga, mille kasutatavust ajas ei saa kindlalt ette näha on vajalik anda

administraatorile võimalus muuta samaaegselt käivitatavate päringute arvu konfiguratsioonifailist. Võimalus lahendab et ootel päringud kuhjuksid, kui kasutajad teostavad rohkem päringuid kui iga kolme minuti tagant käivitatav taustaprotsess suudab läbi käia.

3. **Massandmete päringu võimalus ainult autenditud kasutajatele** – Päringuid ei teostata koheselt, vaid siis kui süsteemile sobib, seega peab olema täielikult tuvastatav milline kasutaja päringu teostas, et oleks võimalik talle päringu valmimisel tulemused kättesaadavaks teha.
4. **Massandmete päringu käivitamine** – Päringuid peab käivitama süsteem automaatse taustaprotsessi vahendusel. Päringut käivitav taustaprotsess peab alustama tööd iga kolme minuti tagant. Päringuid tuleb käivitada FIFO põhimõttel sissetuleku järjekorras.
5. **Valitud skoobi salvestamine** – Kasutaja poolt paika pandud tulemuste skoop peab baasis säilima, et vajadusel oleks teoreetiliselt võimalik teostada sama päring uuesti.
6. **Valitud kriteeriumite salvestamine** – Kasutaja paika pandud päringu kriteeriumid peavad andmebaasis säilima, et vajadusel oleks võimalik teostada sama päring uuesti.
7. **Päringu tulemuse salvestamine** – Päringu tulemused tuleb panna *csv* faili, mis omakorda pakitakse kokku *zip* failiks. Saadud fail tuleb salvestada andmebaasi.
8. **Päringu tulemuse maksimaalne ridade arv failis** – Tulemuse failis võib ridu olla kuni 1 000 000. Juhul kui ridu on rohkem kui etteantud maksimum tuleb tulemused salvestada mitmes failis ning teha failid kasutajale eraldi kättesaadavaks.
9. **Päringu tulemuse kustutamine** – Päringu tulemusi on vajalik hoida alles kuni kasutaja tulemused on kätte saanud või kuni viis päeva. Rakendatakse tingimust, kumb enne täitub. Tulemuste kustutamist teostab iga ööl automaatselt käivitatav taustaprotsess, mille ainsaks ülesandeks on massandmete päringu tulemuste kustutamine.

10. **Päringu valmimise teavitamine** – Päringuid teostav taustaprotsess tulemuste salvestamise järgselt teavitab päringu tellijat töö valmimisest emaili teel.
11. **Päringu tulemuse allalaadimine** – Päringu valmides ilmub kasutajale päringute ülevaate vaatesse, mis asub saki „Minu toimingud“ valikus „Massandmete päringud“ vastava päringu real kinnitus tulemuse olemasolu kohta ning võimalus tulemus failina alla laadida.
12. **Kriteeriumite kombineerimine** – Kriteeriumeid peab olema võimalik päringule seada rohkem kui üks.
13. **Kokku sobimatud kriteeriumid** – Kriteeriumile peab olema võimalik halduril lisada viit teistele kriteeriumitele, mille päringus täpsustamine koos antud kriteeriumiga ei ole lubatud.
14. **Ainuvõimalik kriteeriumi liik** – Kriteeriumit peab olema halduril võimalik märkida liigilt selliseks, et tema valimisel ei ole enam teisi kriteeriumeid võimalik valida. Sellisel juhul ei kehti nõue nr 12.
15. **Kriteeriumina peab olema võimalik sisestada registrikoodide nimekirja** – Massandmete päringu süsteemi kaudu peab olema kasutajal võimalik laadida üles manus, mis sisaldab nimekirja registrikoodidest, mille kohta andmeid soovitakse.
16. **Kriteeriumina registrikoodide nimekirja lisamisel laetava faili lubatud formaadid** – Kasutajal on lubatud ainsa faili formaadina laadida üles *csv* formaadis faile.
17. **Registrikoodide nimekiri olgu liigilt ainuvõimalik kriteerium** – Kui kasutaja sisestab manusena faili registrikoodidest, mille kohta andmeid soovib, peavad teised kriteeriumid muutuma mitteaktiivseks, ehk teisi kriteeriume ei tohi saada täpsustada.
18. **Registrikoodide nimekirjas aktsepteeritavad kirjed** – Manus olevas *csv* failis võetakse väärtus kriteeriumina arvesse, kui see vastab Eestis kehtestatud ettevõtte registrikoodi formaadile, milleks on kaheksa järjestikust numbrit ning

esimene number peab olema üks, kaheksa või üheksa. Ebasobivad kirjed jäetakse välja.

19. **Registrikoodide nimekirjas kirjete arvu võimalik vahemik** – Massandmete päringu kriteeriumina laadides üles nimekirja, peab see sisaldama vahemikus 10 kuni 5000 aktsepteeritavat kirjet. Alampiir on lisatud, kuna väiksemale hulgale registrikoodidele andmeid pärides on otstarbekam kasutajal kasutada teabesüsteemi lihtpäringu võimalust. Aktsepteeritavad kirjeid tuleb kontrollida lähtudes tingimustest, mis on toodud nõudes 18. Jäädes piiridest välja, teavitatakse kasutajat mitesobivast sisendist.
20. **Registrikoodide nimekirjas veergude arv** – Massandmete päringu kriteeriumina laadides üles faili registrikoodide nimekirjaga, peavad andmed paiknema ühes veerus.
21. **Registrikoodide nimekirja edukalt töödeldud andmete tagasiside** – Kasutaja poolse nimekirja sisestamise järgselt peab süsteem kasutajale kuvama arvu, kui mitu aktsepteeritavat kirjet sisse laeti. Aktsepteeritavad kirjeid tuleb kontrollida lähtudes tingimustest, mis on toodud nõudes 18.
22. **Kriteeriumina peab olema võimalik täpsustada ettevõtte olekut** – Kasutajal peab olema võimalik massandmete päringus valida talle vajalikku ettevõtte oleku aktiivsete olekute hulgast. Aktiivseteks olekuteks loetakse kolme olekut: registrisse kantud, pankrotis, likvideerimisel.
23. **Ettevõtte oleku valiku reeglid** – Kasutaja peab saama valida rohkem kui ühte olekut. Jättes valiku tegemata ei ole võimalik valikut kinnitada, tagasilikumine on sel juhul võimalik. Kõiki olekuid valida ei ole lubatud kuna vastasel juhul ei oleks tegu kitsendusega.
24. **Kriteeriumina peab olema võimalik täpsustada ettevõtte õiguslikku vormi ning alaliiki** – Kasutajal peab olema võimalik massandmete päringus valida talle vajalikku ettevõtte õigusliku vormi ning selle alaliiki.
25. **Õigusliku vormi valiku reeglid** – Kasutaja peab saama valida rohkem kui ühte õiguslikku vormi või selle alaliiki. Jättes valiku tegemata ei ole võimalik valikut

kinnitada, tagasiliikumine on sel juhul võimalik. Kõiki vorme valida ei ole lubatud kuna vastasel juhul ei oleks tegu kitsendusega.

26. **Õigusliku vormi alaliigi valikute ilmumine** – Valides õiguslikuks vormiks „MTÜ – mittetulundusühing“ või „TÜ – tulundusühistu“, siis peab ilmuma lisavalikuna võimalus täpsustada alaliiki. Teiste õigusliku vormide puhul alaliigi valikut ei teki.
27. **Kriteeriumina peab olema võimalik täpsustada ettevõtte peamist tegevusala** – Kasutajal peab olema võimalik massandmete päringus valida talle vajaliku ettevõtte tegevusala.
28. **Tegevusala valikute paigutus** – Kasutajale tuleb kuvada tegevusala valikud järjestatult esmalt vastavalt tegevusala tasemele ning seejärel tähestikulises järjekorras.
29. **Tegevusala valikute reeglid** – Kasutaja peab saama valida rohkem kui ühte tegevusala. Jättes valiku tegemata ei ole võimalik valikut kinnitada, tagasiliikumine on sel juhul võimalik. Kõiki valikuid valida ei ole lubatud kuna vastasel juhul ei oleks tegu kitsendusega.
30. **Kriteeriumina peab olema võimalik täpsustada ettevõtete aadressi** – Kasutajal peab olema võimalik massandmete päringus valida talle vajalikku haldusüksust, kus paiknevad ettevõtted teda huvitavad.
31. **Aadressi valiku reeglid** – Kasutajal peab tekkima aadressi sisendväljale *autocomplete* teenusega antud valikud. Peab saama valida üks kuni kümme erinevat haldusüksust. Jättes valiku tegemata ei ole võimalik valikut kinnitada, tagasiliikumine on sel juhul võimalik. Kõiki valikuid valida ei ole lubatud kuna vastasel juhul ei oleks tegu kitsendusega.
32. **Kriteeriumina peab olema võimalik täpsustada esmakande kuupäeva** – Kasutajal peab olema võimalik massandmete päringus valida talle olulist kuupäeva, mis ajast alates peab olema ettevõtte registrisse kantud.
33. **Esmakande kuupäeva valiku reeglid** – Kasutajal peab tekkima kuupäeva sisendväljale kalender, kust on kasutajal võimalik kuupäeva valida. Võimalik on

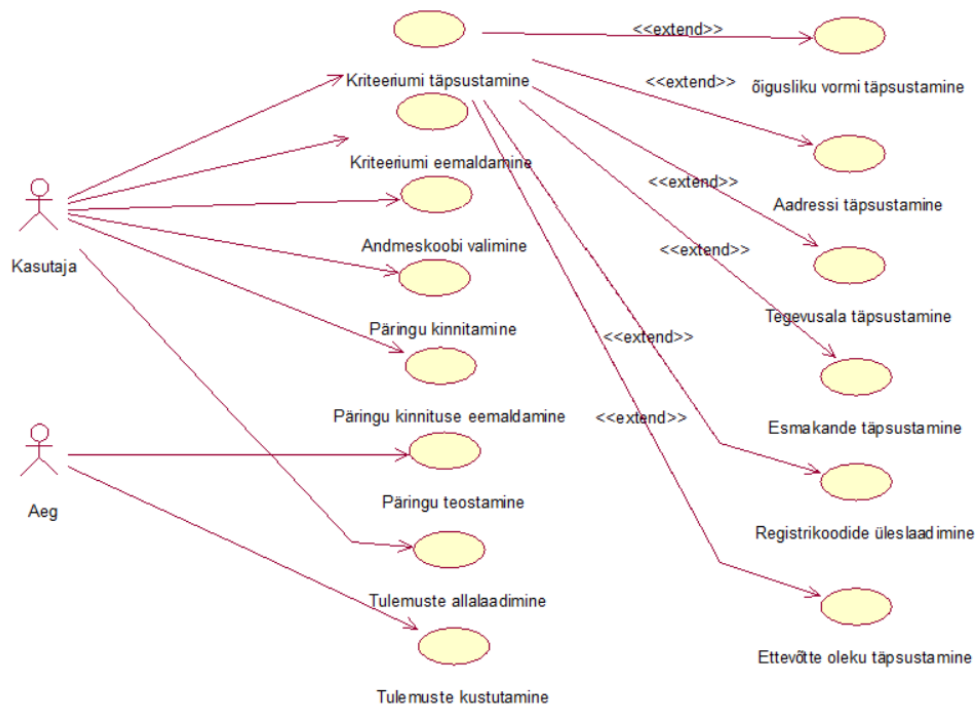
valida päringu kohta ühe kuupäeva. Jättes valiku tegemata ei ole võimalik valikut kinnitada, tagasilikumine on sel juhul võimalik.

34. **Andmeskoopide kombineerimine** – Andmeskoope peab olema võimalik päringule lisada rohkem kui üks.
35. **Kokkusobimatud andmeskoobid** – Andmeskoobile peab olema võimalik halduril lisada viit teistele andmeskoopidele, mille päringusse lisamine koos antud andmeskoobiga ei ole lubatud.
36. **Ainuvõimalik andmeskoobi liik** – Andmeskoopi peab olema võimalik halduril märkida liigilt selliseks, et tema valimisel ei ole enam teisi andmeskoope võimalik valida. Sellisel juhul ei kehti nõue nr 34.
37. **Andmeskoobina peab olema võimalik tulemustesse kaasata ettevõtte andmeid** – Kasutajal peab olema võimalik kaasata tulemusse ettevõtte andmestikku, kuhu peavad kuuluma järgmised andmed: ettevõtte nimetus, registrikood, õiguslik vorm, ettevõtte olek, aadress, postikood ja kapital.
38. **Andmeskoobina peab olema võimalik tulemustesse kaasata kõik kontaktandmed** – Kasutajal peab olema võimalik kaasata tulemusse kõikide kontaktandmete andmestikku, kuhu peavad kuuluma järgmised andmed: sidevahendi liik, sidevahend. Kõikide kontaktandmete valikuga kokku sobimatud andmeskoobid on emaili aadressid ning seotud isikud.
39. **Andmeskoobina peab olema võimalik tulemustesse kaasata emaili aadressid** – Kasutajal peab olema võimalik kaasata tulemusse sidevahenditest ainult emailide andmestikku, kuhu peavad kuuluma järgmised andmed: sidevahendi liik, sidevahend. Emaili aadresside valikuga kokku sobimatud andmeskoobid on kõik kontaktandmed ning seotud isikud.
40. **Andmeskoobina peab olema võimalik tulemustesse kaasata majandusaasta aruande näitajad** – Kasutajal peab olema võimalik kaasata tulemusse majandusaasta aruannete näitajate andmestikku, kuhu peavad kuuluma järgmised andmed: töötajate arv, käive, kasum, majandusaasta algus ja majandusaasta lõpp.

41. **Majandusaasta aruande näitajate andmestiku seos ettevõtte andmetega** – Majandusaasta aruande näitajate leidmisel tuleb seos luua ainult viimase sobiva kirjega. Ehk tuleb luua üks-mitu seosest üks-ühele seos, lisades vastusesse andmed ainult viimase aasta kohta.
42. **Andmeskoobina peab olema võimalik tulemustesse kaasata seotud isikud** – Kasutajal peab olema võimalik kaasata tulemusse ettevõttega seotud isikute andmestikku, kuhu peavad kuuluma järgmised andmed: isiku eesnimi, isiku nimi, isiku kood ja roll ettevõttes. Seotud isikute valikuga kokku sobimatud andmeskoobid on kõik kontaktandmed ning emaili aadressid.
43. **Seotud isikute andmestiku seos ettevõtte andmetega** – Seotud isikute andmete leidmisel tuleb seos ettevõtte andmetega luua funktsiooni abil, kuna seotud isikute andmeid paiknevad kahes erinevas tabelis, mille asjakohased andmed tuleb funktsiooniga ühendada.
44. **Andmeskoobina peab olema võimalik tulemustesse kaasata ettevõtte peamise tegevusala andmeid** – Kasutajal peab olema võimalik kaasata tulemusse ettevõtte tegevusala andmestikku, kuhu peavad kuuluma järgmised andmed: tegevusala kood, tegevusala nimetus.
45. **Andmeskoobina peab olema võimalik tulemustesse kaasata ettevõtte käibemaksukohuslase numbrit** – Kasutajal peab olema võimalik kaasata tulemusse ettevõtte käibemaksukohuslase numbriga andmestik, kuhu peavad kuuluma järgmised andmed: käibemaksukohuslase number.
46. **Kohustuslikud sisendid** – Kasutaja peab päringus täpsustama vähemalt ühte kriteeriumit ning kaasama päringu tulemusse vähemalt ühe andmeskoobi. Lisaks peab kasutaja päringu kinnitamisel sisestama emaili, kuhu saabub päringu valmimisel vastav teavitus.

3.4 Kasutusjuhtude mudel

Järgnevalt on kirjeldatud massandmete päringu funktsionaalsuse kasutusjuhte. Selgitav diagramm on esitatud joonisel 2.



Joonis 2. Kasutusjuhtude diagramm.

Kasutusjuht: Kriteeriumi täpsustamine

Primaarne tegutseja: Kasutaja

Osapooled ja nende huvid:

- Kasutaja soovib saada tema vajadustele võimalikult sobivaid andmeid.

Käivitav sündmus: Kasutaja soovib täpsustada temale vajalikku tingimust.

Eeltingimused: Kasutaja on ennast teabesüsteemi autentunud kas eraisikuna või teabesüsteemi kasutajana.

Järeltingimused: Päringuga on seotud täpsustatud kriteeriumi andmed.

Stsenaarium (tüüpiline sündmuste käik):

1. Kasutaja suundub massandmete päringu vaatesse.
2. Kasutaja valib kriteeriumina õigusliku vormi, mida soovib täpsustada.
3. Süsteem avab õigusliku vormi täpsustamise vaate.
4. Kasutaja valib nimekirjast õiguslikud vormid, mille vasteid päringu tulemusena soovitakse.
5. Kasutaja kinnitab valikud.
6. Süsteem salvestab valikud.
7. Süsteem suunab kasutaja tagasi massandmete päringu täitmise vaatesse.

Laiendused (või alternatiivne sündmuste käik):

- 2a. Kasutaja valib kriteeriumina aadressi, mida soovib täpsustada.
1. Süsteem avab aadressi täpsustamise vaate.
 2. Kasutaja sisestab haldusüksused, kus paiknevaid ettevõtteid päringu tulemusena soovitakse.
- 2b. Kasutaja valib kriteeriumina tegevusala, mida soovib täpsustada.
1. Süsteem avab tegevusala täpsustamise vaate.
 2. Kasutaja valib tegevusalad, milles tegutsevaid ettevõtteid päringu tulemusena soovitakse.
- 2c. Kasutaja valib kriteeriumina esmakande aja, mida soovib täpsustada.
1. Süsteem avab esmakande aja täpsustamise vaate.
 2. Kasutaja valib talle pakutavast kalendrist kuupäeva, millest alates sisse kantud ettevõtteid tulemusena soovitakse.
- 2d. Kasutaja valib kriteeriumina registrikoodid, mida soovib sisendada.
1. Süsteem avab registrikoodide lisamise vaate.
 2. Kasutaja laeb üles kuni 5000 realise csv formaadis faili temale huvipakkuvatest registrikoodidest.
- 2e. Kasutaja valib kriteeriumina ettevõtte olekud, mida soovib täpsustada.
1. Süsteem avab ettevõtte olekute täpsustamise vaate.
 2. Kasutaja valib nimekirjast ettevõtte olekud, millele vastavaid ettevõtteid päringu tulemusena soovitakse.
- 5a. Kasutaja ei soovi valikuid salvestada ning liigub tagasi massandmete päringu koostamise lehele
1. Süsteem ei salvesta valikuid.

Kasutusjuht: Kriteeriumi eemaldamine

Primaarne tegutseja: Kasutaja

Osapooled ja nende huvid:

- Kasutaja: soovib tühistada eelnevalt tehtud kriteeriumi täpsustamise valikut.

Käivitav sündmus: Kasutaja on ekslikult täpsustanud tingimust.

Eeltingimused: Päringuga on seotud kriteerium, mille valikuid on salvestatud andmebaasi

Järelingimused: Seos kriteeriumi ning päringu ja kriteeriumi vahel on eemaldatud ning kriteeriumi valikud kustutatud.

Stsenaarium (tüüpiline sündmuse käik):

1. Süsteem kuvab kasutajale kõik täpsustamiseks võimalikud kriteeriumid nimekirjas.
2. Täpsustatud kriteeriumil kuvab süsteem valiku eemaldamise võimalust.
3. Kasutaja valib sellise eelnevalt täpsustatud kriteeriumi, mille valikut soovib tühistada.
4. Süsteem tühistab kriteeriumi ning päringu seose.
5. Süsteem kustutab andmebaasist kõik tehtud valikud kriteeriumi kohta.

Kasutusjuht: Andmeskoobi valimine

Primaarne tegutseja: Kasutaja

Osapooled ja nende huvid:

- Kasutaja: soovib, määrata tulemusse jõudvate andmete koosseisu.

Käivitatav sündmus: Kasutaja määrab milliseid andmeid tulemusse kaasata.

Eeltingimused: Kasutaja on päringus täpsustanud tema jaoks vajalikud kriteeriumid.

Järelingimused: Päringuga on seotud valitud andmeskoobid.

Stsenaarium (tüüpiline sündmuste käik):

1. Subjekt soovib valida andmeskoobid päringu tulemusteks.
2. Süsteem kuvab valitavate andmeskoopide nimekirja.
3. Subjekt valib nimekirjast andmeskoobi, mida soovib tulemusena saada.
4. Süsteem seob valitud andmeskoobi päringuga.

Laiendused (või alternatiivne sündmuste käik):

- 5a. Subjekt soovib eemaldada andmeskoobi päringust.
 1. Subjekt eemaldab andmeskoobi päringust.
 2. Süsteem kustutab andmeskoobi ja päringu seose.

Kasutusjuht: Päringu kinnitamine

Primaarne tegutseja: Kasutaja

Osapooled ja nende huvid:

- Kasutaja: Soovib tema märgitud kriteeriumite ning valitud andmeskoopidele vastavaid andmeid.

Käivitatav sündmus: Kasutaja kinnitab oma päringu teostamise valikud.

Eeltingimused: Päringuga on seotud vähemalt üks kriteerium ning vähemalt üks andmeskoop.

Järelingimused: Päringu olek on muudetud kinnitatuks ja päringule on lisatud kinnitamise aeg ning email.

Stsenaarium (tüüpiline sündmuste käik):

1. Kasutaja kinnitab päringu.

2. Süsteem avab modaalakna, milles palub kasutajal sisestada emaili, kuhu saadetakse hiljem teavitust päringu valmimisest. Andmeväljal on vaikimisi väärtuseks kasutaja email, kui seda on võimalik välja võtta kas eelnevate päringute küljest või lepingulise kasutaja lepingu partneri andmetest.

3. Kasutaja sisestab emaili ning kinnitab päringu

4. Süsteem salvestab andmed ja muudab päringu staatuse kinnitatuks.

Kasutusjuht: Päringult kinnituse eemaldamine

Primaarne tegutseja: Kasutaja

Osapooled ja nende huvid:

- Kasutaja: Soovib muuta eelnevalt kinnitatud päringus tehtud valikuid

Käivitav sündmus: Kasutaja eemaldab päringu kinnituse.

Eeltingimused: Päring on olekus „kinnitatud“.

Järeltingimused: Päringu olek on muudetud „alustatud“ ning päringult on eemaldatud kinnitamise aeg.

Stsenaarium (tüüpiline sündmuste käik):

1. Kasutaja avab vaate, milles kuvatakse ülevaade tellitud päringutest.
2. Süsteem kuvab kasutajale kinnitamise ajast kahanevalt nimekirjas kasutajaga seotud päringuid ning nende olekuid. Kinnitatud olekus päringute puhul annab süsteem kasutajale võimaluse kinnituse eemaldada.
3. Kasutaja valib päringu, milliselt ta soovib kinnituse eemaldada ning vajutab nuppu „Eemalda kinnitus“.
4. Süsteem muudab päringu olekuks „alustatud“ ning eemaldab päringult kinnituseaja ja suunab kasutaja päringu teostamise vaatesse, milles on täidetud kõik väljad, mis kasutajal olid eelnevalt salvestatud.

Kasutusjuht: Päringu teostamine

Primaarne tegutseja: Aeg

Osapooled ja nende huvid:

- Aeg: Ettemääratud ajal käivitatakse päringu teostamise taustaprotsess.

Käivitav sündmus: Käivitub *python*'i skript, mis koostab SQL päringu ning käivitab selle ja teeb tulemused kasutajale kättesaadavaks.

Eeltingimused: Taustaprotsess on määratud käivituma kindlatel aegadel, lisaks on olemas vähemalt üks päring, mille staatus on kinnitatud.

Järelingimused: Päring on teostatud, tulemus on lisatud. Päringu staatus muudetud „valmis“.

Stsenaarium (tüüpiline sündmuste käik):

1. Süsteem valib massandmete päringu id, mida hakkab töötleva.
2. Süsteem muudab päringu staatuseks „tões“.
3. Süsteem koostab vastavalt valitud andmeskoopidele SQL päringu *SELECT* jaotise.
4. Süsteem koostab vastavalt valitud kriteeriumitele ning andmeskoopidele SQL päringu *FROM* jaotise.
5. Süsteem koostab vastavalt valitud kriteeriumitele ning andmeskoopidele SQL päringu *WHERE* jaotise.
6. Süsteem käivitab loodud SQL lause.
7. Süsteem salvestab tulemused *csv* faili.
8. Süsteem pakendab *csv* faili *zip* formaati.
9. Süsteem salvestab seostab tekitatud faili päringuga.
10. Kui tellijaks on tasuta kasutaja, muudab süsteem päringu staatuseks „makstud“ ning märgib lõppajaks praeguse kuupäeva ning kellaaja.
11. Süsteem teavitab kasutajat päringu valmimisest sisestatud emailile.

Laiendused (või alternatiivne sündmuste käik):

10a. Kui tellijaks on tasuline kasutaja, muudab süsteem päringu staatuseks „Valmis“ ning märgib lõpp ajaks praeguse kuupäeva ning kellaaja.

Kasutusjuht: Tulemuste allalaadimine

Primaarne tegutseja: Kasutaja

Osapooled ja nende huvid:

- Aeg: Ettemääratud ajal käivitatakse päringu tulemuse faili kustutamise taustaprotsess

Käivitav sündmus: Kasutaja laeb päringu tulemuse enda arvutisse.

Eeltingimused: Päring on süsteemi poolt töödeldud ning päringu staatuseks on „Makstud“.

Järelingimused: Päring on kasutaja poolt alla laetud ning päringu staatuseks on „Alla laetud“.

Stsenaarium (tüüpiline sündmuste käik):

1. Kasutaja valib teabesüsteemist „Minu toimingud“ sakist alajaotuse „Massandmete päringud“.

2. Süsteem avab tellitud päringute ülevaate vaate.
3. Kasutaja valib päringu, mille staatuseks on „Makstud“, real olevale nupule „Lae tulemus“.
4. Süsteem serveerib kasutajale päringu tulemuse faili.
5. Süsteem muudab päringu staatuseks „Alla laetud“.

Kasutusjuht: Tulemuste kustutamine

Primaarne tegutseja: Aeg

Osapooled ja nende huvid:

- Äriregistri talitus: soovib et andmebaas oleks võimalikult kompaktne.

Käivitav sündmus: Käivitus *python*'i skript, mis kustutab päringu tulemuse faili andmebaasist.

Eeltingimused: Päring on teostatud ja /või kasutaja on tulemus alla laadinud.

Järeltingimused: Päring tulemuse juurest on kustutatud fail andmetega.

Stsenaarium (tüüpiline sündmuste käik):

1. Süsteem kontrollib päringute olemasolu, mis on „alla laetud“ staatuses ning mille tulemus ei ole null.
2. Leides selliseid kirjeid, kustutab süsteem tulemuse kirjelt faili väärtuse.
3. Süsteem kontrollib päringute olemasolu, mis on „valmis“ staatuses ning mille lõpp ajaks on märgitud aeg, millest on möödas viis päeva.
3. Leides selliseid kirjeid, kustutab süsteem tulemuse kirjelt faili väärtuse.

3.5 Mittefunktsionaalsed nõuded

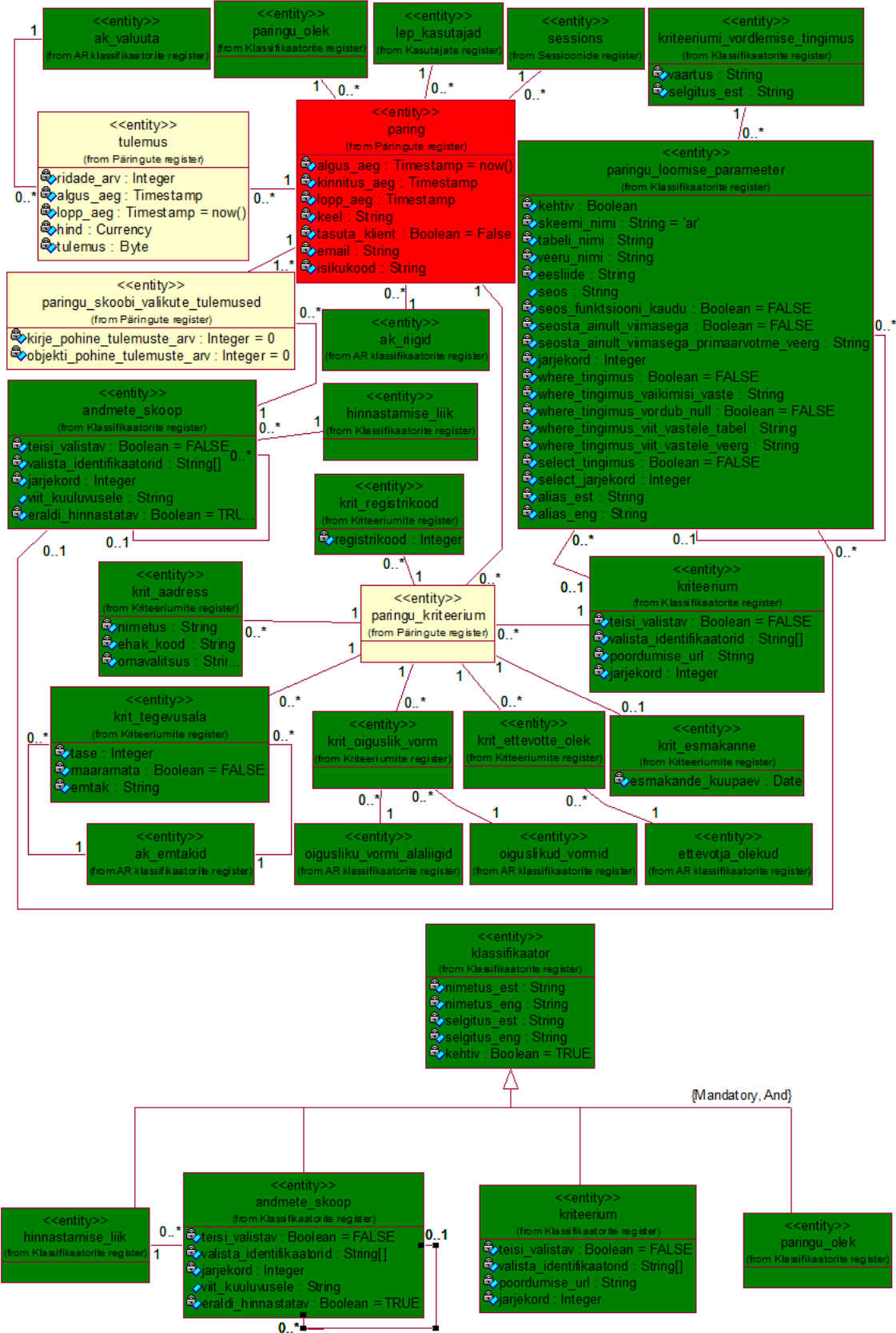
- **Veebilehitsejate tugi** – Rakendus peab töötama tõrgeteta Äriregistri poolt toetatavates veebilehitsejates, milleks on 2016. aasta mai seisuga IE11, Chrome 49 ja 50, Mozilla Firefox 44 ja 45 ning Safari 9 ja 9.1.
- **Jõudlus** – Massandmete päringu töötlemise taustaprotsess peab jõudma töödelda iga päringu poole tunni jooksul alates päringu töötlemise alustamise hetkest.
- **Teisaldatavus** – Massandmete päringut teostav lahendus peab olema võimalikult sõltumatu Äriregistri infosüsteemi spetsiifilistest komponentidest, et seda oleks vajadusel võimalik integreerida teiste süsteemidega.

- **Dokumenteeritus** – Massandmete päringu süsteem peab olema dokumenteeritud, et Äriregistri töötajatel oleks süsteemist parem arusaam ning võimalike edasiarenduste jaoks oleks arendajatel võimalik saada ülevaade süsteemist.
- **Kasutajaliides** – Peab jälgima Äriregistri teabesüsteemi disaini.
- **Turvalisus** – Kuna rakenduse taustaprotsess koostab valitud kriteeriumite ning andmeskoopide alusel dünaamiliselt SQL lauseid, siis ei tohi anda ebavajalikke ligipääse SQL lausete koostamise aluseks olevatele tabelitele ja rakenduse kood tuleb koostada selliselt, et SQL süstimine ei oleks võimalik. [10]

3.6 Kontseptuaalne andmemudel

3.6.1 Olemi suhte diagramm

Järgneval joonisel 3 on tähistatud punasega päringute registri põhiobjekt. Kollasega on tähistatud päringute registrisse kuuluvad mitte-põhiobjektid. Rohelisega on tähistatud teistesse registrisse kuuluvad objektid, mida on vaja massandmete päringute töötlemise automatiseerimise toimimise tagamiseks.



Joonis 3. Andmebaasi kontseptuaalset struktuuri kirjeldav olemi-suhte diagramm

3.6.2 Olemitüüpide definitsioonid

Loodavad tabelid jaotatakse alamosadeks ehk registriteks ning tabelis 2 on muuhulgas näha olemitüüpide jaotus erinevate registrite vahel. Joonisel 3 on kujutatud juba Äriregistri andmebaasis olemasolevad tabelid ilma atribuutide jaotiseta.

Tabel 2. Olemitüüpide definitsioonide kirjeldused.

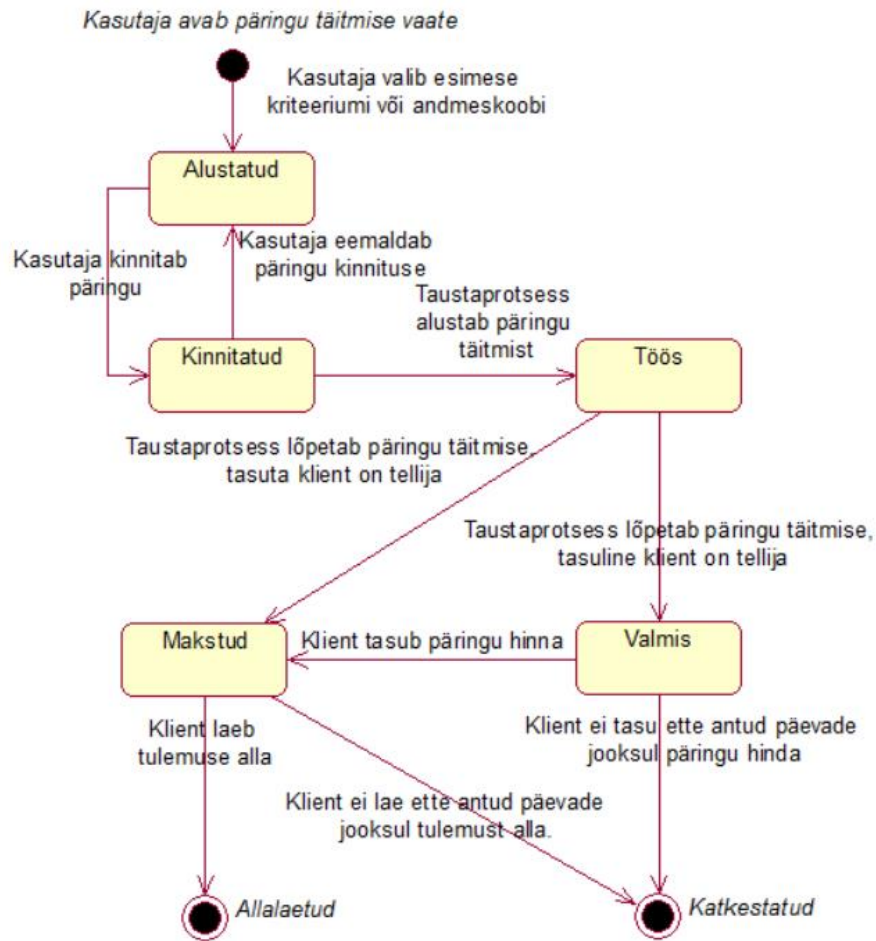
Olemitüübi nimi (aliased)	Kuuluvus registrisse	Definitsioon
Paring	Päringute register	Iga kord kui teabesüsteemi kasutaja asub teostama massandmete tellimise protsessi, salvestatakse siia esialgsed andmed, millega hakatakse kogu tellimise protsessi vältel salvestatavaid andmeid seostama.
Tulemus	Päringute register	Kirjeldab päringu tulemusena saadud vastuseid.
Paringu_kriteerium	Päringute register	Kirjeldab päringu ja kriteeriumi liigi omavahelist sõltuvust. Seda sõltuvust kasutatakse kriteeriumi valikute seostamisel päringu ning kriteeriumi liigiga.
Paringu_skoobi_valikute_tulemused	Päringute register	Kirjeldab päringu ja valitud andmestiku omavahelist sõltuvust. Lisaks peab arvestust iga valitud skoobi saadud tulemuste arvu üle, mida kasutatakse tulemuse koguhinna määramisel.
Andmete_skoop	Klassifikaatorite register	Kirjeldab erinevaid andmestike, mida on võimalik tellida. Andmestikud koosnevad semantiliselt sarnastest veergudest.
Kriteerium	Klassifikaatorite register	Kirjeldab erinevaid tingimusi, mille alusel on võimalik teostatavat päringut kitsendada.
Paringu_loomise_parameeter	Klassifikaatorite register	Kirjeldab reegleid, mille alusel päringu kinnitades taustaprotsessil on võimalik koostada valitud kriteeriumite ning andmeskoopide järgi SQL päring.
Krit_aadress	Kriteeriumite register	Kirjeldab kriteeriumina täpsustatud haldusüksuse valikuid.

Olemitüübi nimi (aliased)	Kuuluvus registrisse	Definitsioon
Krit_registrikood	Kriteeriumite register	Kirjeldab kriteeriumina üles laetud csv failis olevaid ettevõtete registrikoodide väärtuseid, mis on läbinud vastava sobivus-algoritmi kontrolli.
Krit_esmakanne	Kriteeriumite register	Kirjeldab kriteeriumina täpsustatud kuupäeva, millest alates sisse kantud ettevõtjaid soovitakse.
Krit_ettevotte_olek	Kriteeriumite register	Kirjeldab kriteeriumina valitud ettevõtte olekuid
Krit_oiguslik_vorm	Kriteeriumite register	Kirjeldab kriteeriumina valitud ettevõtte õigusliku vormi väärtuseid.
Krit_tegevusala	Kriteeriumite register	Kirjeldab kriteeriumina valitud tegevusala väärtuseid.
Hinnastamise_liik	Klassifikaatorite register	Kirjeldab, mis viisil on iga andmeskoop hinnastatav.
Kriteeriumi_vordlemise_tingimus	Klassifikaatorite register	Kirjeldab, mis viisil päringut koostades võrreldakse nii kasutaja valitud kui ka vaikimisi väärtuseid.
Paringu_olek	Klassifikaatorite register	Kirjeldab olekut, milles iga päring hetkel paikneb.
Lep_kasutajad	Kasutajate register	Kirjeldab teabesüsteemis registreeritud kasutajaid.
Sessions	Sessioonide register	Kirjeldab kasutaja sessiooni teabesüsteemis.
Ak_valuuta	AR klassifikaatorite register	Kirjeldab valuutasid, mis on erinevates riikides ametlikult käibel.
Ak_emptak	AR klassifikaatorite register	Kirjeldab tegevusalasid, mida on võimalik ettevõtetel enda tegevusaladeks valida
Oigusliku_vormi_alaliigid	AR klassifikaatorite register	Kirjeldab ettevõtte õigusliku vormi alaliiki
Oiguslikud_vormid	AR klassifikaatorite register	Kirjeldab ettevõtte õigusliku vormi väärtust.
Ettevõtja_olekud	AR klassifikaatorite register	Kirjeldab millises olekus ettevõtja hetkel on.

Olemitüüpides sisalduvate atribuutide kirjeldus on leitav lisa 1 olevas tabelis 5.

3.7 Päringu seisundidiagramm

Järgnevalt on joonisel 4 välja toodud loodava süsteemi põhiolemitüübi, milleks on päring, seisundidiagramm.



Joonis 4. Päringu seisundidiagramm

3.8 Andmebaasi lisanduvad tabelid

Järgnevalt on joonisel 5 välja toodud loodava süsteemi andmebaasi diagramm. Andmebaasi tabelite loomisele, klassifikaatorite täitmisele ning andmebaasi õiguste paika panekule kulus kokku 65 tundi ning tulemuseks oli kokku ca 500 rida SQL koodi.

Massandmete päringu süsteemi tabelid saavad asuma samas skeemis, kus asuvad kõik Äriregistri teabesüsteemi tabelid. Juba Äriregistri andmebaasis eksisteerivate tabelite puhul piirduakse seose kuvamisega loogilise andmebaasi diagrammis. Eraldi nende detailset kirjeldust välja ei tooda. Järgnevas tabelis 3 on lisatud andmebaasi lisatavate tabelite kirjeldused. Tabelite detailne kirjeldus on välja toodud lisa 2 olevas tabelis 6.

Tabel 3. Andmebaasi lisatavate tabelite kirjeldused.

Tabeli nimi	Millise olemitüübi, atribuudi või seosetüübi põhjal on loodud?
Mass_paring	Olemitüüp: paring
mass_tulemus	Olemitüüp: tulemus
Mass_sys_paringu_loomise_parameeter	Olemitüüp: paringu_loomise_parameeter
Mass_sys_andmete_skoop	Olemitüüp: andmete_skoop
Mass_sys_kriteerium	Olemitüüp: kriteerium
Mass_sys_kriteeriumi_vordlemise_tingimus	Olemitüüp: kriteeriumi_vordlemise_tingimus
Mass_krit_registrikood	Olemitüüp: krit_registrikood
Mass_krit_aadress	Olemitüüp: krit_aadress
Mass_krit_tegevusala	Olemitüüp: krit_tegevusala
Mass_krit_ettevotte_olek	Olemitüüp: krit_ettevotte_olek
Mass_krit_oiguslik_vorm	Olemitüüp: krit_oiguslik_vorm
Mass_krit_esmakanne	Olemitüüp: krit_esmakanne
Mass_paringu_kriteerium	Olemitüüp: paringu_kriteerium
Mass_paringu_skoobi_valikute_tulemused	Olemitüüp: paringu_skoobi_valikute_tulemused
Mass_kl_hinnastamise_liik	Olemitüüp: hinnastamise_liik
Mass_kl_paringu_olek	Olemitüüp: paringu_olek
Kl_ettevotte_olekud	Olemitüüp: ettevotja_olekud
Kl_oigusliku_vormi_alaliigid	Olemitüüp: oigusliku_vormi_alaliigid
Kl_oiguslikud_vormid	Olemitüüp: oiguslikud_vormid
Ak_emtak	Olemitüüp: ak_emtak
Ak_valuuta	Olemitüüp: ak_valuuta
Ak_riigid	Olemitüüp: ak_riigid
Sessions	Olemitüüp: sessions
Lep_kasutajad	Olemitüüp: lep_kasutajad

3.9 Kontroll päringu loogika sisestuse üle

Kogu koostatava SQL päringu ülesehituse loogika paikneb tabelis *mass_sys_paringu_loomise_parameeter*. Eesmärgiga vältida süsteemi esialgse loogika paika panekul ning hilisemal täiendamisel vigaseid sisestusi, mis võivad päringu koostamise või käivitamise rikkuda, on loodud andmebaasi triger, mis kontrollib enne *mass_sys_paringu_loomise_parameeter* tabelisse rea lisamist selle reeglitele vastavust. Kontrollitavateks tingimusteks on:

1. Sisestatava veeru eksisteerimine sisestatavas skeemis olevas sisestatavas tabelis.
2. Kui tegu on funktsiooniga, kas funktsioon eksisteerib andmebaasis.
3. Kõik kirjed sama tabeli kohta peavad kasutama ühte ja sama eesliidet.
4. Kõik kirjed sama tabeli kohta peavad kasutama sama järjekorra numbrit.
5. Kirje ei tohi olla samal ajal *select* tingimus ja *where* tingimus.
6. Viimasega seostamisel peab olema täpsustatud ka seostatava tabeli primaarvõtme veerg.
7. *Where* tingimuse vaste viida olemasolu andmebaasis.

Triger SQL lause kujul on näha lisa 3 oleval joonisel 24.

3.10 Päringu töötlemise algoritm

Päringu töötlemise algoritmi loomisele kulus 30 tundi ning moodustab kokku ca 400 rida koodireast.

Joonisel 6 esitatakse tegevusdiagrammil päringu töötlemise algoritm, milles viidatakse konkreetsetele andmebaasi tabelitele, mille poole algoritm pöördub. Algoritm alustab tööd iga kolme minuti möödudes.

väiksem, kui konfiguratsioonifailis seatud maksimaalne arv, siis valib algoritm kõige varem kinnitatud päringu ning asub seda täitma.

Esmalt muudab algoritm päringu oleku „Töös“. Seejärel pärib tabelist *mass_paringu_loomise_parameeter* kõik päringus valitud andmeskoopidega seotud andmed, mille *where_tingimus* on väär ning *select_tingimus* on tõene. Neid andmeid kasutades loob algoritm hiljem käivitatava päringu *SELECT* jaotise. Seejärel koostatakse SQL päringu *FROM* jaotis. Selle jaoks pärib algoritm *mass_paringu_loomise_parameeter* tabelist kõik päringus valitud andmeskoopide ning kriteeriumitega seotud andmed, mille *where_tingimus* on väär ning *select_tingimus* on väär. Viimasena koostab taustaprotsess SQL päringu *WHERE* jaotise. Selleks pärib algoritm *mass_paringu_loomise_parameeter* tabelist kõik päringus valitud andmeskoopide ning kriteeriumitega seotud andmed, mille *where_tingimus* on tõene ning *select_tingimus* on väär ning saadud andmeid kasutades loob *WHERE* jaotisesse kuuluvad tingimused.

Päring valmis, käivitab taustaprotsess selle ning ootab vastust. Vastuse saabudes lisatakse 1 000 000 rea kaupa tulemused eraldi *csv* failidesse, mis omakorda pakendatakse *zip* failideks ning lisatakse *mass_tulemus* tabelisse. Seejärel sõltuvalt päringu tellinud klienditüübist muudab algoritm päringu kas „Makstud“ või „Valmis“ olekusse ning saadab päringu tellinud kasutajale e-maili teavituse päringu valmimisest ja lõpetab töö.

4. Kasutajaliides

Peatükis kirjeldatakse töö eesmärgina loodud kasutajaliidese funktsionaalsust. Kasutajaliidese toimimise jaoks vajaliku serveripoolse funktsionaalsuse loomine võttis aega 170 tundi ning lahendus koosneb ca 1900 koodireast. *Front end* funktsionaalsuse loomine võttis aega 220 tundi ning lahendus koosneb ca 2700 koodireast.

Kasutajaliides on lisatud töö autori Äriregistri teabesüsteemi arenduskeskkonna külge ning seda on võimalik testida spetsiaalselt selle projekti jaoks loodud proksi vahendusel aadressil <https://demo-ariregister-massandmed.rik.ee>. Proksi on loodud IP piiranguga, kuid on ligipääsetav TTÜ arvutipargi arvutitele. Arenduskeskkond on seotud Äriregistri pööratud andmebaasiga, milles on sassi ajatud ettevõtete registrikoodid ning mitmed muud andmestikud asendatud võltsväärtustega. Seega reaalseid andmeid selle rakenduse kaudu ei saa. Peatüki kirjutamise hetkeks (2016 aprill) on realiseeritud töötaval kujul kogu plaanitud kasutajaliidese funktsionaalsus, kuid kasutajaliidese disain ei kuulu töös püstitatud eesmärkide hulka ning ei ole sellisel kujul toodangu keskkonna kõlbulik.

Järgnevalt on kirjeldatud loodud kasutajaliidese funktsionaalsust selle vaadete piltide ning selgituste abil.

4.1 Massandmete päringu funktsionaalsuse asukoht

Funktsionaalsuse kasutamine eeldab eelnevat kasutaja tuvastamist infosüsteemi poolt. Selle vajaduse tingib soov päringu valmides kasutajale tulemused kättesaadavaks teha. Seetõttu on esmalt vajalik kasutajal ennast autentida kasutades ühte teabesüsteemi autentimiseks olemasolevatest võimalustest, sel põhjusel kirjeldatakse esmalt käesoleva magistritöö osaks mitteolevat funktsionaalsust, kuidas Äriregistri teabesüsteemi autentida. Teabesüsteemi sisselogimise asukoht on liikudes avalehelt „Sisene“ nupule vajutades autentimise vaatesse, mis on väljatoodud joonisel 7.

Esimesed kolm autentimislahendust: ID-kaardiga, mobiil-ID'ga ning pangalingi kaudu autentimine – pärinevad kõik RIK'i ühtsest autentimislahendust pakkuvast keskkonnast

– Autentimisportaalist, mis üle *proxy* lähenedes ei pruugi funktsioneerida. Sel juhul saab kasutada massandmete päringu funktsionaalsusele ligi pääsemiseks nimekirjas viimast autentimislahendust, milleks on parooliga sisenemine lepingulisele kliendile. Kasutada on võimalik testimise otstarbeks loodud kasutajat: „massandmed“, mille parooliks on „Massandmed“.

TEST
e-äriregister
Registrite ja Infosüsteemide Keskus

ÄRIREGISTRI TEABESÜSTEEM

0 SISENE

Avaleht Päringud Nimekirjad Abiinfo

Sisenemine

Aega jäänud: 12:42

Kasuta ID-kaardiga autentimist

ID-KAART

Kasuta Mobiil-ID'ga autentimist

Isikukood:

Mobiilinumbr:

MOBIIL-ID

Sisene pangalingi kaudu

SEB Swedbank Danske Bank Nordea LHV

Parooliga sisenemine lepingulisele kliendile

Kasutajanimi:

Parool:

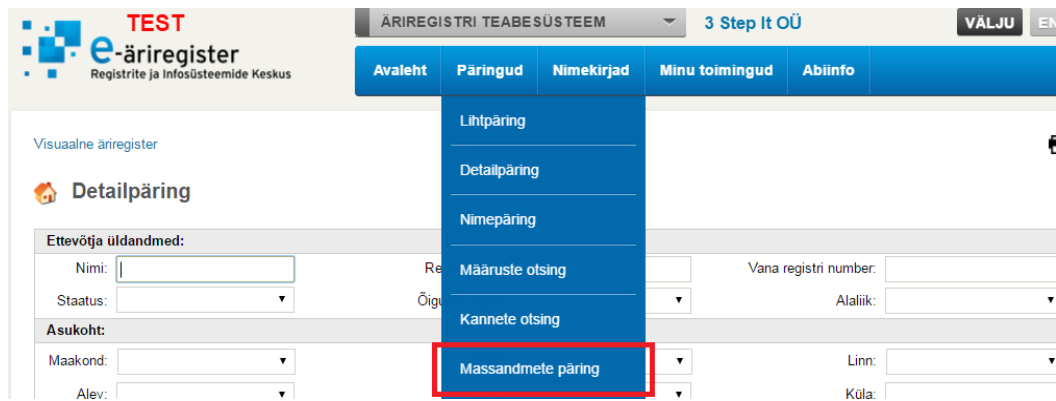
SISENE

Soovides saada meie lepinguliseks kliendiks, palun täita registreerimisvormi.

Registreeru

Joonis 7. Teabesüsteemi autentimise vaade.

Olles süsteemis autenditud kasutaja kuvatakse eelnevale lisaks sakis „Päringud“ valikut „Massandmete päring“, mis ongi diplomitöö objektiks oleva funktsionaalsuse asukohaks. Valik on näha joonisel 8.



Joonis 8. Massandmete päringu asukoht.

4.2 Massandmete päringu koostamine

Valides eelnevas punktis väljatoodud „Massandmete päringu“ valiku, avaneb joonisel 9 kuvatud vaade, kus kasutaja saab hakata päringule kriteeriumeid ning andmestikke lisama.

Kriteeriumite täpsustamine

Õiguslik vorm ?	+ Vali	
Aadress ?	+ Vali	
Tegevusala ?	+ Vali	
Esmakande aeg ?	✍ Muuda	✖ Kustuta
Registrikoodide nimekiri ?	+ Vali	
Ettevõtte olek ?	✍ Muuda	✖ Kustuta

Andmestike valimine

Valik ettevõtte andmetest ?	<input type="checkbox"/>
Kõik kontaktandmed ?	<input checked="" type="checkbox"/>
E-maili aadressid ?	<input type="checkbox"/>
Majandusaasta aruande näitajad ?	<input type="checkbox"/>
Seotud isikud ?	<input type="checkbox"/>
Tegevusala ?	<input type="checkbox"/>
Käibemaksukohuslase number ?	<input checked="" type="checkbox"/>

Vastus koosneb järgmistest veergudest:

```

ettevotte_registrikood;ettevotte_nimetus;sidevahendi_liik;sidevahend;KMKR_number;
  
```

Joonis 9. Massandmete päringu koostamise vaade

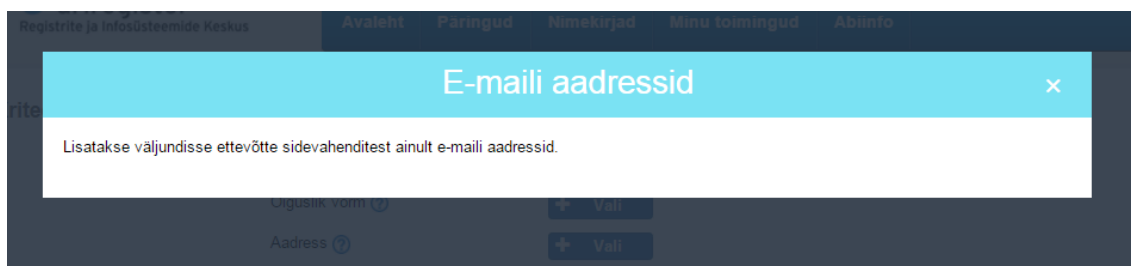
Esmalt on loetletud kriteeriumid, mille järgi on võimalik päringut kitsendada. Kriteeriumeid saab lisada igas päringus üks kuni mitu. Joonisel 9 näidatud päringus on kasutaja täpsustanud esmakande aega ning ettevõtte olekut. Täpsustatud valikuid on soovi korral võimalik muuta või kustutada. Samuti on jooniselt näha, et kriteerium „Registrikoodide nimekiri“ on mitteaktiivses olekus, mis on tingitud kriteeriumi omadusest olla ainuke võimalik valik. Kuna näites on juba valitud teised kriteeriumid, siis registrikoodide nimekirja enam täpsustada ei saa.

Järgmisena on loetletud andmestikud, mille valikute alusel pannakse kokku päringu tulemus. Joonisel 9 on näha, et kasutaja on valinud tulemusse lisatavateks andmestikkudeks „Kõik kontaktandmed“ ja „Käibemaksukohuslase number“. Samuti on märgatavalt erinevus emaili aadresside ning seotud isikute andmestiku valikus võrreldes

teiste mittevalitud andmestikega. Tumedam lülitusnupp näitab, et seda valikut ei ole võimalik juba tehtud valikutega koos lisada. Kui kasutaja nüüd ikkagi valib lisaks emaili aadresside andmestiku siis kuvatakse talle veeteade, mis selgitab tagamaid, miks ei ole valiku tegemine lubatud.

Iga andmestiku kohta tehtud valikuga muutub veergude nimekiri, mis paikneb „Vastus kuvatakse järgmistest veergudest“ alapealkirja all. Nimekiri on lisatud andmaks kasutajale koheselt tagasisidet, mis andmeid veergudena tema tehtud andmestike valikute alusel tulemuse faili lisatakse. Joonisel 9 toodud näites sisaldaks tulemus ettevõtte registrikoodi, ettevõtte nimetuse, sidevahendi liigi, sidevahendi ning käibemaksukohuslase numbri andmeväljasid.

Vajutades kriteeriumi või andmestiku valiku nimetuse järel oleva küsimärgi ikooni peale avaneb objekti selgitust sisaldav modaalaken, nagu väljatoodud joonisel 10.



Joonis 10. Massandmete päringu koostamise info modaalaken

Päringu katkestades, päring kustutatakse ning kasutaja suunatakse teabesüsteemi avalehele. Päringu kinnitamine õnnestub vaid juhul, kui kasutaja on päringule lisanud vähemalt ühe andmestiku ning täpsustanud vähemalt ühte kriteeriumi, vastasel juhul kuvatakse kasutajale veeteade. Kinnituse reeglite täitmisel avaneb modaalaken nagu näidatud joonisel 11, kus kasutaja peab sisestama kahekordselt oma emaili, millele taustaprotsess, mis päringu teostab, saadab päringu valmides vastava teavituse.

E-meili aadressi kinnitus

Palume sisestada emaili aadress, kuhu saadetakse teavitust päringu valmides.

Emaili aadress rauno.kulla@just.ee

Korda emaili aadressi rauno.kulla@just.ee

Loobu Kinnita

Joonis 11. Päringu kinnitamisel emaili sisestamine

Emaili sisestamisel kontrollitakse kahe sisendi vastavust emaili formaadile ning sisendite omavahelist kokkulangemist. Nende tingimuste täitmisel päring kinnitatakse ning kasutaja suunatakse massandmete päringu nimekirja vaatesse, mis on kirjeldatud punktis 4.4.

4.3 Päringu kriteeriumite täpsustamiste vaated

Joonisel 9 kujutatavas massandmete päringu koostamise vaates on iga kriteeriumi järel võimalus täpsustada selle kriteeriumi väärtust. Vajutades kriteeriumi real nupule „Vali“ olenevalt, kui kriteerium on eelnevalt päringule lisamata või „Muuda“ juhul kui kriteeriumi valikud on juba salvestatud, suunatakse kasutaja kriteeriumi andmete sisestamise vaadetesse.

Kõikidel kriteeriumitel välja arvatud esmakande aeg on lubatud sisestada mitu vastet, kuid ei ole lubatud salvestada valikut, mis sisaldab kõiki valiku variante, kuna siis ei oleks tegu päringu kitsendusega ning sellega kaasneks koostatavale SQL päringule ebavajalike ridade lisandumine, mis teeks päringut hoopis aeglasemaks. Samuti ei ole võimalik kriteeriumit salvestada tegemata ühtki valikut kriteeriumi piiramiseks.

4.3.1 Õigusliku vormi kriteeriumina täpsustamine

Soovides täpsustada massandmete päringu koostamise vaates, kujutatud joonisel 9, ettevõtete õiguslikku vormi, liigub kasutaja õigusliku vormi määramise vaatesse, kus on tal võimalik valida kõikide õiguslike vormide seast just sellised mille vasteid tulemusse soovib. Vaade on kujutatud joonisel 12. Võimalikud valikud on järjestatud kasutades javascript jQuery raamistiku akordioni tööriista, milles avanevad lahti sellised

õiguslikud vormid, millel on alaliike. Alaliikidega õiguslikud vormid on tulundusühistu ning mittetulundusühing.

The screenshot shows the 'Valik õigusliku vormi järgi' (Selection by legal form) page on the Estonian Business Register website. The page has a blue header with the logo and navigation tabs: 'Avaleht', 'Päringud', 'Nimekirjad', 'Minu toimingud', and 'Abiinfo'. Below the header, the title 'Valik õigusliku vormi järgi' is displayed. The main content area contains a list of legal forms with checkboxes for selection. The 'Osühing' (Partnership) form is selected. Under the 'Tulundusühistu' (Business Entity) category, the following sub-forms are selected: 'liikmete isikliku vastutusega tulundusühistu' (Business entity with limited liability), 'lisavastutusega tulundusühistu' (Business entity with unlimited liability), and 'piiratud vastutusega tulundusühistu' (Business entity with limited liability). Other forms listed include 'Aksiaselts', 'Täisühing', 'Füüsilisest isikust ettevõtja', 'Usaldusühing', 'Mittetulundusühing', 'Sihtasutus', 'Välismaa äriühingu filiaal', 'Euroopa majandushuviühing', 'Euroopa äriühing (Societas Europaea)', 'Euroopa ühistu', and 'Euroopa territoriaalse koostöö rühmitus'. At the bottom of the form, there are two buttons: '← Tagasi' (Back) and 'Salvesta →' (Save).

Joonis 12. Õigusliku vormi kriteeriumina täpsustamise vaade.

4.3.2 Aadressi kriteeriumina täpsustamine

Soovides täpsustada massandmete päringu koostamise vaates, kujutatud joonisel 9, haldusüksust, kus piires ettevõtte paikneb, liigub kasutaja aadressi määramise vaatesse, kus on tal võimalik sisestada haldusüksused, mille piires peavad tulemusse kaasatavad ettevõtted paikema. Vaade on kujutatud joonisel 13.

TEST
e-äriregister
Registrite ja Infosüsteemide Keskus

ÄRIREGISTRI TEABESÜSTEEM Registrite ja Infosüst.. VAL

Avaleht Päringud Nimekirjad Minu toimingud Abiinfo

Valik ettevõtte paiknemise järgi

Haldusüksus (linn, alev, alevik, küla) Viiratsi alevik, Viljandi vald, Viljandi maakond + Salvesta ja lisa järgmine

Päringusse kaasatavad haldusüksused

Olustvere alevik, Suure-Jaani vald, Viljandi maakond	Kustuta
Suure-Jaani linn, Suure-Jaani vald, Viljandi maakond	Kustuta
Võhma linn, Viljandi maakond	Kustuta
Viiratsi alevik, Viljandi vald, Viljandi maakond	Kustuta

← Tagasi Salvesta →

Joonis 13. Aadressi kriteeriumina täpsustamise vaade.

Haldusüksuse valiku sisendväljal on kasutusel *javascript*'i *jQuery* raamistiku *autocomplete* tööriist, mis pakub kasutaja sisendile sobivaid vasteid haldusüksuste näol alates kahe tähe pikkusest sisendist. Lahendus on võetud üle Äriregistri ettevõtjaportaali aadressi valikust ning kohandatud vastavaks. Kui ettevõtjaportaali aadressivalikus töötab *autocomplete* kahetasemeliselt, kus teiseks tasemeks on tänava ning maja ja korteri valik, siis siinse lahenduse jaoks on pandud teenus tööle ainult haldusüksuse taseme jaoks.

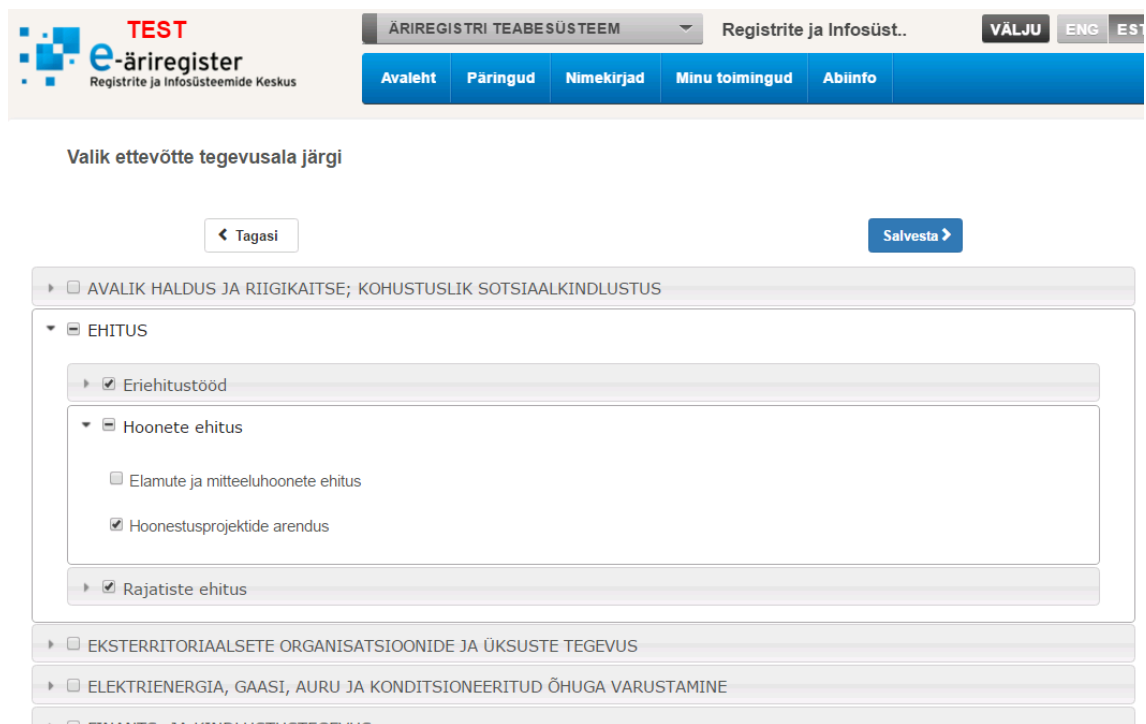
Ühtlasi on kasutajal võimalik lisada päringu kriteeriumiks päringu kohta korraga kuni 10 haldusüksust. Selle jaoks tuleb *autocomplete* vaste valiku järgselt vajutada nupule „Salvesta ja lisa järgmine“, mis tühjendab sisendvälja uue sisendi jaoks. Iga tehtud valikut on võimalik kustuta nupu vajutusega päringu kriteeriumi osana eemaldada.

4.3.3 Tegevusala kriteeriumina täpsustamine

Soovides täpsustada massandmete päringu koostamise vaates, kujutatud joonisel 9, ettevõtete tegevusala, liigub kasutaja tegevusala määramise vaatesse, kus tal on võimalik valida kõikide tegevusalade seast sellised, milles tegutsevaid ettevõtteid tulemusse soovib. Vaade on kujutatud joonisel 14. Võimalikud valikud on järjestatud kasutades sarnaselt õigusliku vormi vaatele *javascript*'i *jQuery* akordioni tööriista, milles avanevad kliki peale lahti sellised valikud, mis sisaldavad alamvalikuid.

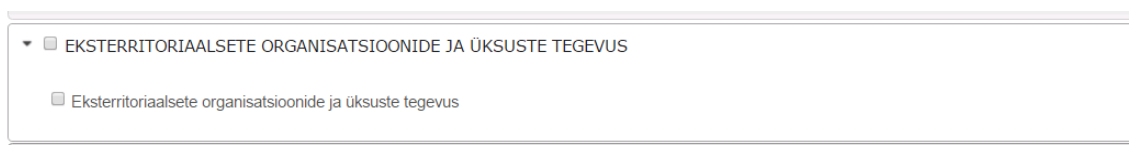
Kuna tegevusala kõikvõimalike valikuid on sadu, siis annab rakendusele kiiruse eelise lahendus, milles laetakse esmalt esimese taseme tegevusalad korraga ning seejärel iga

valik teostab *jQuery ajax* päringu saades vastuseks selle tegevusala järgmise taseme valikud.



Joonis 14. Tegevusala kriteeriumina täpsustamise vaade.

Võrreldes eelnevalt Äriregistri süsteemides kasutusel olnud tegevusala valikutele on parandatud kasutusmugavust pannes kõik tegevusala valikud tähestiku järjekorda. Samuti on vähendatud kasutaja poolt tehtavate klikkide arvu tehes kõige alumise, viienda taseme, tegevusala valiku kättesaadavaks kohe, kui temast tase kõrgemal eksisteerib ainult üks tegevusala, mis tähendab, et sel tasemel valikud puuduvad. Siinkohal on toodud näiteks joonisel 15 kuvatav esimese taseme tegevusala „Eksterritoriaalsete organisatsioonide ja üksuste tegevus“, millel on ainult üks viienda taseme tegevusala ja igal vahetasemel on samuti ainult üks tegevusala, seega kasutajal pole vajadust vahepealseid tasemeid läbi klikkida, et jõuda viienda taseme tegevusalani.



Joonis 15. Selgitav pilt kuidas jäetakse vahele tasemed, millel puuduvad valikud.

4.3.4 Esmakande kuupäeva kriteeriumina täpsustamine

Soovides täpsustada massandmete päringu koostamise vaates, kujutatud joonisel 9, kuupäeva, alates millest registrisse kantud ettevõtteid tulemusse soovitakse, liigub

kasutaja esmakande aja määramise vaatesse, kus tal on võimalik valida soovitud kuupäev. Vaade on kujutatud joonisel 16. Esmakande teostamise kuupäeva sisendväljal klikk avab *javascript*'i *jQuery* raamistiku kuupäeva valiku tööriista, mis aitab kasutajal visuaalselt soovitud kuupäev valida.



The screenshot shows the top navigation bar of the 'ÄRIREGISTRI TEABESÜSTEEM' (Business Register Information System) in 'TEST' mode. The main menu includes 'Avaleht', 'Päringud', 'Nimekirjad', 'Minu toimingud', and 'Abiinfo'. Below the navigation bar, the page title is 'Valik ettevõtte esmakande aja järgi'. The main content area features a form with the label 'Esmakande teostamise kuupäev:' and a text input field containing '01.03.2016'. A date picker calendar is open, showing the month of 'Märts 2016'. The calendar grid has columns for days of the week (E, T, K, N, R, L, P) and rows for dates from 1 to 31. Navigation buttons 'Tagasi' and 'Salvesta' are visible at the bottom of the calendar.

Joonis 16. Esmakande aja kriteeriumina täpsustamise vaade.

4.3.5 Registrikoodide kriteeriumina täpsustamine

Soovides saada oma registrikoodide nimekirjale lisaks teisi andmestike on kasutajal võimalik laadida süsteemi üles oma *csv* formaadis nimekiri registrikoodidest, millele lisavälju valitavate andmestike näol soovib. Selleks valib kasutaja massandmete päringu koostamise vaates, mis on kujutatud joonisel 9, registrikoodide nimekirja kriteeriumi ning liigub üleslaetava registrikoodide nimekirja vaatesse, mis on kujutatud joonisel 17.

Valik üleslaetavate registrikoodide järgi

Teil on võimalik laadida üles enda nimekiri registrikoodidest, mis on salvestatud CSV formaadis. Nimekiri peab olema ühes veerus ning ridade arv peab jääma vahemikku 10 - 5000

Registrikoodide fail (CSV):

[← Tagasi](#)

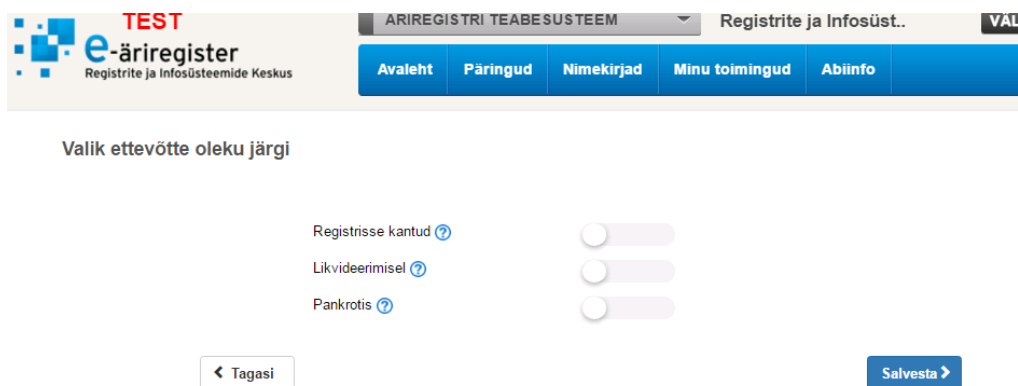
[Salvesta →](#)

Joonis 17. Registrikoodide nimekirja kriteeriumina lisamise vaade.

Kasutajal on lubatud sisestada *csv* formaadis fail, milles on üks veerg ning ridade vahemik peab jääma 10 ja 5000 vahele, vastasel juhul saab kasutaja serveri poolt veateate. Registrikoodide nimekirja üleslaadimisel ei ole kasutajal lubatud lisada päringule täiendavaid kriteeriume kuna tehakse eeldus, et kasutaja soovib kõikidele oma nimekirjas olevatele registrikoodidele lisaandmeid.

4.3.6 Ettevõtte oleku kriteeriumina täpsustamine

Soovides täpsustada massandmete päringu koostamise vaates, kujutatud joonisel 9, ettevõtete oleku liiki, liigub kasutaja ettevõtte oleku määramise vaatesse, kus on tal võimalik valida ettevõtte olekute seast sellised mille vasteid tulemusse soovib. Vaade on kujutatud joonisel 18.



Joonis 18. Ettevõtte oleku kriteeriumina täpsustamise vaade.

4.4 Ülevaade tellitud päringutest

Sisseloginud Äreregistri teabesüsteemi kasutaja saab näha tema poolt tellitud päringute nimekirja valides peamenüüst saki „Minu toimingud“ alt valiku „Massandmete päringud“. Valik on toodud välja joonisel 19.



Joonis 19. Tellitud päringute ülevaate asukoht.

Avanevas nimekirjas kuvatakse tellimuse aja järgi kahanevalt kasutajaga seotud massandmete päringud, millega on võimalik teostada iga päringu olekust sõltuvat tegevust. Nimekiri kasutajaga seotud massandmete päringutest on toodud välja joonisel 20. Iga tellimuse rea kohta kuvatakse kasutajale tunnust, kas tulemus eksisteerib, ehk päring on töödeldud.

Tellimuse aeg	Tellimuse staatus	Tulemus olemas	
	Alustatud	-	Täitma
01.05.2016 0258	Kinnitatud	-	Eemalda kinnitus
01.05.2016 0258	Töös	-	
30.04.2016 1620	Makstud	+	Lae tulemus
30.04.2016 1618	Makstud	+	Lae tulemus
29.04.2016 2323	Makstud	+	Lae tulemus
29.04.2016 2320	Makstud	+	Lae tulemus
29.04.2016 2311	Makstud	+	Lae tulemus
29.04.2016 2259	Makstud	+	Lae tulemus
29.04.2016 2251	Makstud	+	Lae tulemus

[Vanemad ->](#)

Joonis 20. Kasutajaga seotud massandmete päringu nimekiri.

Nimekirjas kuvatakse kokku kuni kümme kasutajaga seotud massandmete päringut. Juhul kui kasutajal on temaga seotud massandmete päringuid rohkem kui 10, siis antakse kasutajale võimalus liikuda eelmiste päringute juurde nupult „vanemad“.

Päringu puhul, mis on olekus „alustatud“ kuvatakse kasutajale võimalus liikuda päringu täitmise vaatesse. Olekus „kinnitatud“ päringu puhul antakse kasutajale võimalus päringult kinnitus eemaldada. Kinnituse eemaldades muutub päringu staatus tagasi „alustatud“ olekusse ning kasutaja suunatakse massandmete päringu teostamise vaatesse, milles on salvestatud kõik väljad, mis olid kasutaja poolt eelnevalt lisatud. Olek „töös“ tähendab, et päring on hetkel süsteemi poolt töötluses, mille tõttu kasutajale võimalust päringuga lisategevust teha ei võimaldata. Kasutajal tuleb oodata, kuni süsteem lõpetab päringu töötlemise.

Kuna hinnastamine ning sellega koos päringu eest tasumine ei kuulu käesoleva töö skoopi, on realiseeritud lahendus, mille järgselt süsteem päringuga omapoolse töö lõpetamisel märgib päringu staatuseks „makstud“. Makstud olek tähendab, et päringuga on seotud tulemus ning kasutajale kuvatakse võimalus tulemus allalaadida. Valides allalaadimise serveritakse kasutajale tulemuse andmed pakendatult *zip* faili ning muudetakse päringu staatus „allalaetud“.

Päringu tulemuse allalaadimise võimalust pakutakse kasutajale kuni päringu tulemuste faili andmebaasist kustutav taustaprotsess faili eemaldab.

5. Ekstreemumolukorra testimine

Järgnevas peatükis teostatakse loodud rakendusele ekstreemumolukorra test, mille eesmärgiks on katsetada rakenduse toimimist juhul, kui kasutaja teostab väga suuri andmebahte käsitleva päringu. Kuigi arenduskeskkond on ühendatud Äriregistri pööratud andmebaasiga, on selles andmebaasis kirjete arv suurusjärgult võrreldav toodangu andmebaasi kirjete arvuga. Seega on ekstreemumolukorra test asjakohane. Ekstreemolukorra testimise ning ettevalmistuse peale kulus ajaliselt 5 tundi.

5.1 Ekstreemumolukorra testimiseks sobiva päringu koostamine

Kuna rakenduse kaudu on kasutajal nõutud lisada päringule vähemalt üks tingimus, siis sellise olukorra imiteerimiseks lisatakse päringule kriteerium, mis kitsendab tulemusse kaasatavate ettevõtete arvu minimaalselt. Seesuguseks tingimuseks võetakse ettevõtte õiguslik vorm, milles valitakse kõik valikud va euroopa territoriaalse koostöö rühmitus, mida esineb õiguslikkudest vormidest kõige vähem. Kriteeriumi valikud on toodud välja joonisel 21.

Valik õigusliku vormi järgi

<input checked="" type="checkbox"/> Aktsiaselts
<input checked="" type="checkbox"/> Täisühing
<input checked="" type="checkbox"/> Füüsilisest isikust ettevõtja
<input checked="" type="checkbox"/> Usaldusühing
<input checked="" type="checkbox"/> Osaühing
<input checked="" type="checkbox"/> Tulundusühistu
<input checked="" type="checkbox"/> Mittetulundusühing
<input checked="" type="checkbox"/> Sihtasutus
<input checked="" type="checkbox"/> Välismaa äriühingu filiaal
<input checked="" type="checkbox"/> Euroopa majandushuviühing
<input checked="" type="checkbox"/> Euroopa äriühing (Societas Europaea)
<input checked="" type="checkbox"/> Euroopa ühistu
<input type="checkbox"/> Euroopa territoriaalse koostöö rühmitus

[← Tagasi](#)

[Salvesta >](#)

Joonis 21. Ekstreemumolukorra kriteeriumi valikud.

Eesmärgiga saada andmemahtu maksimaalseks kaasame testimisse andmeskoope nii, et tulemuse ridu tekiks maksimaalselt. Selleks on testimiseks kaasatud järgmised andmestikud: valik ettevõtte andmetest, majandusaasta aruande näitajad, seotud isikud, tegevusala ning käibemaksukohuslase numbri. Seotud isikud kaasame, kuna kontaktandmeid on ettevõtete kohta koguselt vähem kui seotud isikuid. Koostatud päringu vaade on toodud välja joonisel 22.

Kriteeriumite täpsustamine

Õiguslik vorm ?	Muuda Kustuta
Aadress ?	+ Vali
Tegevusala ?	+ Vali
Esmakande aeg ?	+ Vali
Registrikoodide nimekiri ?	+ Vali
Ettevõtte olek ?	+ Vali

Andmestike valimine

Valik ettevõtte andmetest ?	<input checked="" type="checkbox"/>
Kõik kontaktandmed ?	<input type="checkbox"/>
E-maili aadressid ?	<input type="checkbox"/>
Majandusaasta aruande näitajad ?	<input checked="" type="checkbox"/>
Seotud isikud ?	<input checked="" type="checkbox"/>
Tegevusala ?	<input checked="" type="checkbox"/>
Käibemaksukohuslase number ?	<input checked="" type="checkbox"/>

Vastus koosneb järgmistest veergudest:

```
ettevotte_registrikood;ettevotte_nimetus;oiguslik_vorm;oigusliku_vormi_alaliik;ettevotte_olek;haldusykskus;aadress;
```

* Katkesta
✓ Kinnita

Joonis 22. Ekstreemumolukorra testimise koostatud päringu vaade.

Päringu kinnitamise järgselt käivitatakse taustaprotsess ning jälgitakse, kas see tegutseb seatud nõuete kohaselt. Nõuete alusel peab päring vastama iga päringu vähem kui tunni ajaga ning salvestama ühte tulemuse faili maksimaalselt 1 000 000 rida.

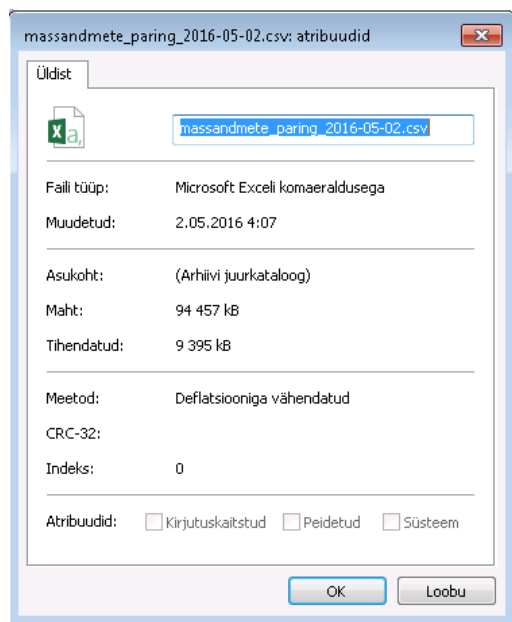
5.2 Ekstreemumolukorra testimise tulemused

Tulemuse metaandmeid kuvatakse tabelis 4. Tabelist on näha, et päring suutis ettenähtud aja piires tulemuse töödelda. Nõudena on püstitatud, et iga päring tuleb töödelda vähem kui tunni ajaga. Algus ja lõppaja vahemik on 1 minut ning 43 sekundit. Seega päringu töötlemise ajalise kestvuse nõue täidetakse probleemivabalt ka ekstreemumolukorra puhul.

Tabel 4. Ekstreemumolukorra testimise tulemuse metaandmed.

Id	Mass_paring_id	Ridade_arv	Algus_aeg	Lopp_aeg
17	25	431338	2016-05-02 04:05:57.312706+03	2016-05-02 04:07:40.241831+03

Rakendusest faili alla laadides on näha, et kokkupakitud faili suuruseks on 9.17Mb, kujutatud ka joonisel 23. Suurus mahub *bytea* andmetüübi maksimaalse suuruse, milleks on 1GB piiresse.



Joonis 23. Ekstreemumolukorra testi tulemuste faili suurus.

5.3 Ekstreemolukorra testimise järeldused

Testimisel osutus, et algselt arvatud probleem, et vastuse failid tekivad liiga suured ning neid tuleb salvestada mitme eraldi failina, tegelikult süsteemi töös murekoht ei ole, ning võib tõsta maksimaalset kirjete arvu, mida ühte faili salvestatakse. Levinud *csv* faili töötlemise programm Windows Excel ei võimalda korraga laadida ühte faili rohkem kui 1 048 576 rida. Seetõttu jäetakse esialgu tulemuste faili piiriks 1 000 000 rida. [12]

Teisalt ilmnes, et massandmete päringu süsteem töötab kiiremini, kui esialgu hinnatud. Kuna ekstreemumolukorras andmemaht, mida taustaprotsess pidi töötleva oli väga suur, sai süsteem päringuga hakkama vähem kui kahe minutiga, siis on mõistlik kasutatavuse kasvamisel toodangu keskkonnas vähendada taustaprotsessi käivitamise vahelist aega. Esialgu sai ajavahemikuks seatud kolm minutit, kuid see oleks otstarbekas muuta väiksemaks. Muudatuse tulemusena väheneks aeg, kui pikalt klient peaks tulemuste järgi ootama.

6. Planeeritud edasiarendused

Realiseeritud süsteem on loodud võimalikult dünaamiliselt lihtsustamaks tulevasi arendusi ning täiendusi. Vajaduse tekkimisel on võimalik lahendusse lisada valikutena nii kriteeriume kui ka andmeskoope. Andmeskoope on võimalik lisada ilma, et oleks vajalik teostada arendusi. Piisab loogikat sisaldavasse tabelisse *mass_sys_paringu_loomise_parameeter* ning andmeskoobi klassifikaatorisse uute ridade lisamisest. Kriteeriumi lisamisel tekib vajadus ka arenduste järele, kuna iga kriteerium omab spetsiifilist kasutajaliidese vaadet ning andmebaasi tabelit, kus logitakse kasutaja määratud kriteeriumi vasted.

Esimeseks planeeritud arenduseks on luua süsteemile tugi massandmete päringu hinnastamise kalkuleerimiseks. Käesoleva töö raames hinnastamist ei realiseeritud, seetõttu ei arvuta süsteem päringu hinda ning seetõttu ei ole võimalik küsida Äriregistri teabesüsteemi tasuliselt kasutajatüübilt päringu eest nõutavat tasu. Seega täiendus muudaks lahenduse kasutatavaks ka tasulistel kasutajatel. Arenduse eelduseks saab olema hinnastamise reeglite kinnitamine RIK'i üldosakonna poolt. Teiseks planeeritud edasiarenduseks on kasutajaliidese disaini parandamine muutes rakenduse oluliselt kasutajasõbralikumaks, seeläbi parandades ka kasutajakogemust.

7. Kokkuvõte

Käesoleva magistritöö raames loodi massandmete päringu täitmise jaoks alternatiiv veebirakenduse näol, mis oli Äriregistris seni käinud emailide vahendamise teel. Loodud lahendus lisandus Äriregistri teabesüsteemi uueks funktsionaalsuseks. Esmalt analüüsiti eelnevalt klientidelt sissetulnud massandmete päringuid ning leiti millised on tüüp-kriteeriumid ja tüüp-andmestikud, mille järgi kliendid enim oma soove täpsustavad. Analüüsis otsustati kriteeriumi ning andmestiku päringus esinemise sageduse alusel, kas realiseerida rakendusse nende jaoks tugi. Välja jäi ainult üks kasutaja poolt täpsustatud tingimus, mida oleks võinud tüüp-kriteeriumina käsitleda. Täiendavalt selgus, et 80% päringute puhul on võimalik päring teostada tüüp-kriteeriumeid ja tüüp-andmestikke kasutades, mille pealt tehti järeldus, et veebirakenduse loomisel väheneks kogu Äriregistri talituse halduskoormus tervelt 20%.

Massandmete päringute käivitamine tehti taustaprotsessi tööülesandeks, seeläbi minimeerides võimaluse, et teabesüsteemi töö hangub andmebaasi mälu täitumise tõttu nagu oleks võinud juhtuda olukorras, kus päringute käivitamise õigus jäänuks kasutajatele. Võttes arvesse, et tegu on uue funktsionaalsusega, ei ole täpselt teada kui suur hakkab olema selle kasutatavus, mistõttu tehti päringute käivitamise maht konfiguratsioonifailist juhitavaks, sellega vähendades ootel olevate päringute kuhjumise ohtu.

Hinnastamise reeglid ei ole nende ridade kirjutamise ajaks RIK'i üldosakonna poolt veel paika pandud. Selle tõttu jäi magistritöö skoobist välja hinnastamise toe põhjalik analüüs ning realiseerimine, kuigi süsteemi kavandamisel lisati ettenägelikult andmebaasi suur osa hinnastamiseks vajalikke andmevälju. Hetkeseisuga on süsteemi funktsionaalsus kasutatav ainult teabesüsteemi tasuta kasutajate jaoks, sest tasulised kasutajad peavad oma päringute eest maksma, aga hinnastamise puuduoleku tõttu tasumisele kuuluvat summat ei arvutata.

Lahenduse testimise käigus jõuti järeldusele, et süsteemil ei ole probleemiks püstitatud nõuete piires püsimine ka ekstreemolukordades, kus peab toimetama suurte

andmemahutudega. Lisaks tehti soovitus muuta taustaprotsessi käivituse sagedust tihedamaks, kuna esialgu seatud väärtuse puhul jääb süsteemil päringute töötamise vahel vaba aega, mis väljendub kliendi poolse tulemuse järgi ootamises.

Lahenduse loomisel peeti oluliseks, et süsteemi oleks tulevikus võimalikult lihtne edasi arendada. Esimeseks lisanduseks loodud tarkvarale planeeritakse realiseerida tugi hinnastamisele. Seeläbi oleks ka teabesüsteemi tasulisel kasutajal võimalik täies mahus funktsionaalsusest kasu lõigata. Teisalt on vajalik kasutajaliidese disain muuta kasutajasõbralikumaks. Uue andmeskoobi loomiseks ei ole vajalik süsteemi arendusi teha, piisab õigetesse tabelitesse ridade lisamisest. Uue kriteeriumi loomiseks vastavatesse tabelitesse ridade lisamisest ei piisa. Vajalik on ka luua uue kriteeriumi jaoks kasutajaliidese vaade ning uus tabel, kus salvestada kasutaja valikud.

Summary

In this work an alternative for the process of querying mass data was developed in the form of a web application. In the Estonian Commercial Register this process had previously been handled by exchanging emails. Developed solution was added into Central Commercial Register as a new functionality. Firstly, mass data queries that had been submitted by Commercial Register's clients were analyzed and out of this analysis, criterions and data scopes that could be regarded as typical were detected. The decision, whether to develop these typical criterions and data scopes were made taking into account how frequently they had been included in analyzed queries. Additionally, it appeared that in case of 80% of the queries they could be performed using only the criterions and data scopes that were generalized as typical. Therefrom, it was deduced that with developing a web application follows a 20% deduction of entire Commercial Register unit's administration load.

The task of starting mass data queries was designed as a background process, thereby minimizing the possibilities of stalling the Central Commercial Register due to database memory problems. Latter could have been an issue if the option of starting mass data queries would have been left to users. The volume of starting mass data queries was made adjustable from configuration file considering that the functionality is new and its usability is uncertain at this time. Therefore reducing the change of accumulating queries that are waiting to be started.

The guidelines of pricing queries have yet to be established by the General Department of the Center of Registers and Information Systems. Because of this, thorough analysis and development of query pricing was left out of the thesis scope. Even though many essential data fields for pricing were foresightedly included while designing the system. Currently the developed functionality is useful only for the free users of Central Commercial Register. Chargeable users would have to pay for their queries before they are granted access to results but because the pricing has not been implemented the transaction amount is not being calculated.

While testing the application it became apparent that staying in the boundaries of set requirements is not going to be an issue even in extreme circumstances when the

amount of data being processed is huge. Moreover, a recommendation was made that the frequency of starting background process should be increased. Whereas using the initially set frequency would cause unnecessary delays between processing. The processing delay would translate to customer side waiting for results.

While designing and developing the solution it was taken into account that future developments would be made as easily as possible. First planned addition to created application is to supplement implementation for pricing the queries. Such addition would make the system useful for all of the users of Central Commercial Register. Secondly, it is planned to make user interface design more user friendly. In order to add new data scope to the system any further developments are not necessary. It can be done simply by adding rows to correct tables. On the other hand, while adding criterion it is additionally required to develop a new view for web application and also to create a table where to store user selections for that criterion.

Kasutatud kirjandus

1. E-teatmik [WWW] <http://www.vallaste.ee> (02.05.2016)
2. Infopäringu mall [WWW] http://www.rik.ee/sites/www.rik.ee/files/elfinder/article_files/infoparing_1.rtf (02.05.2016)
3. Baltic Target [WWW] <http://www.krediidiinfo.ee/index.php?m=26> (02.05.2016)
4. Durkin, N. The Benefits of Load Balancing [WWW] <http://blog.wsol.com/the-benefit-of-load-balancing> (02.05.2015)
5. Indiana University, What are cron and crontab, and how do I use them? [WWW] <https://kb.iu.edu/d/afiz> (02.05.2016)
6. CRONw – CRON for Windows [WWW] <http://cronw.sourceforge.net/> (02.05.2016)
7. Pycron [WWW] <http://www.kalab.com/freeware/pycron/pycron.htm> (02.05.2016)
8. Cartesian Join [WWW] <http://c2.com/cgi/wiki?CartesianJoin> (02.05.2016)
9. BinaryFilesInDB [WWW] <https://wiki.postgresql.org/wiki/BinaryFilesInDB> (02.05.2016)
10. Roman, E. SQL Injection Defense in Python [WWW] <http://www.slideshare.net/openpbs/sql-injection-defense-in-python> (02.05.2016)
11. Andmekaitse inspektsiooni isikukoodi kasutamise juhend [WWW] http://www.aki.ee/sites/www.aki.ee/files/elfinder/article_files/Isikukoodi%20kasutamise%20juhend_2.rtf (02.05.2016)
12. George Zhao, What is the Maximum file size for CSV that Excel can open? [WWW] <https://social.technet.microsoft.com/Forums/office/en-US/54ca6d3a-4b1c-4318-aa8c-7ac803d3af69/what-is-the-maximum-file-size-for-csv-that-excel-can-open-excel-2013-64bit?forum=excel> (02.05.2016)

Lisa 1

Tabel 5. Atribuutide definitsioonide kirjeldused.

Olemitüübi nimi	Atribuudi nimi	Atribuudi definitsioon	Näiteväärtus
Paring	Algus_aeg	Kuupäev ja kellaaeg koos ajatsooniga, millal kasutaja alustas päringu tellimuse täpsustamist.	'2016-04-15 14:02:21'
Paring	Kinnitus_aeg	Kuupäev ja kellaaeg koos ajatsooniga, millal kasutaja kinnitas päringu tellimuse.	'2016-04-15 14:12:21'
Paring	Lopp_aeg	Kuupäev ja kellaaeg koos ajatsooniga, millal süsteem lõpetas päringu töötlemise ning salvestas tulemused kasutajale kättesaadavaks.	'2016-04-15 14:30:21'
Paring	Keel	Päringu kinnitamisel salvestatakse kasutaja rakenduses valitud keel andmebaasi. Seda arvestades koostab taustaprotsess tulemuste päringu, milles on veergude nimetused salvestatud keeles	'est'
Paring	Tasuta_klient	Tõeväärtus veerg, kas päringu tellinud kasutaja puhul on tegemist teabesüsteemi tasuta kasutajaga. Kliendi tunnusest sõltub, millisesse olekusse muudab süsteem päringu pärast tulemuste salvestamist.	TRUE
Paring	Email	Kasutaja, kes päringu tellis, poolt sisestatud emaili aadress, kuhu süsteem saadab teavituse päringu tulemuste valmimise kohta	'Rauno.Kulla@just.ee'
Paring	Isikukood	Kasutaja isikukood on talle riigi, mille kodakondne isik on, poolt antud. "Oluline on teada, et isikukoodid ei kuulu delikaatsete isikuandmete alla. Isikukoodid on tavalised isikuandmed ja nende kasutamisele ei ole seatud rohkem piiranguid kui näiteks inimese nime või sünniaja kasutamisele." [11]	'39105206010'
Tulemus	Ridade_arv	Vastuses sisalduvate ridade arv, mida võetakse arvestusse päringu eest hinna	201

Olemitüübi nimi	Atribuudi nimi	Atribuudi definitsioon	Näiteväärtus
		arvutamisel.	
Tulemus	Algus_aeg	Aeg, millal taustaprotsess alustas päringu täitmist.	'2016-05-01 20:30:21'
Tulemus	Lopp_aeg	Aeg, millal taustaprotsess oli salvestanud päringu tulemused andmebaasi. Ehk algus ning lõppaja vahel on ajavahemik, mis kulus andmebaasist andmete pärimise peale ning lisaks ajavahemik mis kulus süsteemil andmete tulemuse failina salvestamiseks.	'2016-05-01 20:32:33'
Tulemus	Hind	Päringu tulemuste eest kokku arvutatud summa, mis tuleb tasulisel kliendil maksta, enne kui ta pääseb vastusele ligi	14.03
Tulemus	Tulemus	Päringu tulemuseks olevad andmed, mis on salvestatud csv formaadis ning csv fail on omakorda pakendatud <i>zip</i> failiks.	PK...
Päringu_skoobi_valikute_tulemused	Kirje_pohine_tulemuste_arv	Iga päringus valitud skoobi kohta arvutatud kokku ridade arv, kui mitu tulemust valitud skoop andis, ehk ühe ettevõtte kohta käivat kolm kirjet maksustatakse kõik eraldi. Arvestatakse kui skoopi hinnastatakse kirje põhiselt,	121
Päringu_skoobi_valikute_tulemused	Objekti_pohine_tulemuste_arv	Iga päringus valitud skoobi kohta kokku arvutatud ridade arv, kui mitu tulemust valitud skoop andis. Tulemused on koondatud ettevõtte põhiselt, ehk ühe ettevõtte kohta kolm kirjet maksustatakse ühe kordselt. Arvestatakse kui skoopi hinnastatakse objekti põhiselt.	60
Andmete_skoop	Teisi_valistav	Tunnus, mis kirjeldab andmeskoobi kirje kohta, kas tema valimisel, on võimalik valida ka mõnda muud andmeskoopi või mitte.	FALSE
Andmete_skoop	Valista_identifikaatorid	Nimekiri teistest andmeskoopide primaarvõtmetest, mida juhul kui andmeskoop ise on valitud, ei ole võimalik enam valida.	{'1', '2'}
Andmete_skoop	Järjekord	Järjekorra unikaalsed täisarvu tüüpi väärtused panevad paika, mis järjekorras andmeskoobi objekte kasutajale kuvatakse. Järjekord on esialgselt paika pandud analüüsides klientide sissetulevaid	1

Olemistüübi nimi	Atribuudi nimi	Atribuudi definitsioon	Näiteväärtus
		päringuid ning esimesena kuvatakse need andme-skoobid mida küsitakse kõige rohkem.	
Andmete_skoop	Viit_kuuluvusele	Viit kuuluvusele viitab sama tabeli primaarvõtmele, ning iseloomustab kirje tunnust, kuuluda hinnastamiselt viidatud objekti alla.	'3'
Andmete_skoop	Eraldi_hinnastata v	Tunnus, mis kirjeldab kas andmeskoopi hinnastatake eraldi või kuulub ta hinnastamiselt mõne teise andmeskoobi koosseisu	TRUE
Kriteerium	Teisi_valistav	Tunnus, mis kirjeldab kriteeriumi kirje kohta, kas tema valimisel, on võimalik täpsustada ka mõnda muud kriteeriumi või mitte.	FALSE
Kriteerium	Valista_identifikaatorid	Nimekiri teistest kriteeriumite primaarvõtmetest, mida juhul kui kriteerium ise on valitud, ei ole võimalik enam valida.	{'2'}
Kriteerium	Poordumise_url	kirjeldab rakenduse vaadet, kuhu kriteeriumi täpsustamisel rakendus kasutaja suunab.	'oiguslik_vorm'
Paringu_loomise_parameeter	Kehtiv	Tunnus, kas kirje on aktiivne ning seda võetakse süsteemi poolt päringu loomisel arvesse või mitte.	TRUE
Paringu_loomise_parameeter	Skeemi_nimi	Kirjeldab, millise andmebaasiskeemi kohta kirje käib. Sisendi puhul kontrollitakse skeemi olemasolu Äriregistri andmebaasis.	'ar'
Paringu_loomise_parameeter	Tabeli_nimi	Kirjeldab, millise andmebaasitabeli või funktsiooni kohta kirje käib. Sisendi puhul kontrollitakse tabeli või funktsiooni olemasolu skeemis	'ettevotjad'
Paringu_loomise_parameeter	Veeru_nimi	Kirjeldab, millise andmebaasitabeli veeru kohta kirje käib. Sisendi puhul kontrollitakse veeru olemasolu kirjes märgitud tabelis	'nimi'
Paringu_loomise_parameeter	Eesliide	Kirjeldab päringus kasutatavat eesliidet kirjes märgitud andmebaasitabeli kohta.	'et'
Paringu_loomise_parameeter	Seos	Viit seosele viitab sama tabeli primaarvõtmele, ning iseloomustab kirje	'40'

Olemistüübi nimi	Atribuudi nimi	Atribuudi definitsioon	Näiteväärtus
		seost väärtusena märgitud kirjega. Seost kasutatakse SQL päringu FROM jaotise koostamisel tabelite seoste koostamisel.	
Paringu_loomise_parameeter	Seos_funktsiooni_kaudu	Tunnus, kas tabeli asemel on tegu funktsiooniga. Funktsiooni seost tuleb kasutada juhul kui soovitud andmed asuvad erinevates tabelites ning neid on vaja väljastada samades veergudes. Seega on vaja funktsioonis UNION operaatoriga tabelid ühendada ning kriteerium või skoop seostada loodud funktsiooniga.	TRUE
Paringu_loomise_parameeter	Seosta_ainult_vii_masega	Tunnus, kas kirje seos tuleb luua ainult viimase sobiva kirjega tabelis. Kasutatakse juhul kui on soov teha 1-mitu seosest 1-1 seos kahe tabeli vahel.	TRUE
Paringu_loomise_parameeter	Seosta_ainult_vii_masega_primaarvotme_veerg	Juhul kui soovitakse seostada tabelit ainult viimase kirjega, siis on vajalik anda kaasa seostatava tabeli primaarvõtme veerg, millest leitakse maksimum.	'id'
Paringu_loomise_parameeter	Jarjekord	Järjekorra unikaalsed täisarvu tüüpi väärtused panevad paika, mis järjekorras SQL päringu FROM jaotis luuakse.	2
Paringu_loomise_parameeter	Where_tingimus	Tunnus, kas kirje puhul on tegu SQL päringu WHERE jaotise objektiga. Sel juhul kasutatakse kirje andmeid ainult WHERE jaotise kirje loomisel.	TRUE
Paringu_loomise_parameeter	Where_tingimus_vaikimisi_vaste	Vaikimisi väärtus, millele WHERE tingimus peab vastama	'(,L“, „R“, „N“)
Paringu_loomise_parameeter	Where_tingimus_vordub_null	Tunnus, kas WHERE tingimus peab võrduma NULL'iga.	TRUE
Paringu_loomise_parameeter	Where_tingimus_viiit_vastele_tabel	Viit tabelile, kust kirjega seotud kriteeriumi vasteid tuleb otsida	'ettevotjad'
Paringu_loomise_parameeter	Where_tingimus_viiit_vastele_veerg	Viit tabeli veerule, kust kirjega seotud kriteeriumi vasteid tuleb otsida	'ettevotja_oleku_d_kl'
Paringu_loomise_parameeter	Select_tingimus	Tunnus, kas kirje puhul on tegu SQL päringu SELECT jaotise objektiga. Sel juhul kasutatakse kirje andmeid ainult SELECT jaotise kirje loomisel.	FALSE
Paringu_loomise_parameeter	Select_jarjekord	Select_jarjekord väärtuste alusel sõltub mis järjekorras koostab süsteem SQL lause	4

Olemitüübi nimi	Atribuudi nimi	Atribuudi definitsioon	Näiteväärtus
		SELECT jaotise.	
Paringu_loomise_parameeter	Alias_est	Eestikeelne vaste, millega SELECT päringu objekti, ehk veeru väärtus asendatakse.	'nimetus'
Paringu_loomise_parameeter	Alias_eng	Inglisekeelne vaste, millega SELECT päringu objekti, ehk veeru väärtus asendatakse.	'name'
Krit_aadress	Nimetus	Kasutaja poolt valitud haldusüksuse, kus ettevõtja paikneb, nimetus.	'Tallinn, Harju maakond'
Krit_aadress	Ehak_kood	Kasutaja poolt valitud haldusüksuse, kus ettevõtja paikneb, Eesti haldus- ja asutusjaotuse klassifikaatori kood.	'0541'
Krit_aadress	Omavalitsus	Kasutaja poolt valitud haldusüksuse, kus ettevõtja paikneb, omavalitsuse kood.	'6967'
Krit_registrikood	Registrikood	Kasutaja poolt üles laetud ettevõtte registrikood, mis on läbinud registrikoodi algoritmi kontrolli.	10104309
Krit_esmakanne	Esmakande_kuupaev	Kasutaja poolt valitud esmakande kuupäev, alates millisest ajast registrisse kantud ettevõtjaid soovitakse	'2016-01-01'
Krit_tegevusala	Tase	Kasutaja poolt valitud tegevusala tase tegevusalade nimekirjas.	3
Krit_tegevusala	Maaramata	Tunnus, kas kasutaja poolt valitud tegevusalal on mõni alama taseme tegevusala valitud ning mõni valimata, ehk valitud tegevusala ise on olekus määramata.	TRUE
Krit_tegevusala	Emtak	Kasutaja poolt valitud tegevusala kood, millele vastavaid ettevõtjaid otsitakse.	71101
Kriteeriumi_vordlemise_tingimus	Vaartus	Väärtus, milline võrdlustehe SQL lause WHERE jaotises objektiga teostatakse.	'='
Kriteeriumi_vordlemise_tingimus	Selgitus_est	Selgitab võrdlemise tingimust Äriregistri haldurile.	'võrdub täpselt.'
Klassifikaator	Nimetus_est	Klassifikaatori nimetus eesti keelselt.	'Esmakande kuupäev'
Klassifikaator	Nimetus_eng	Klassifikaatori nimetus inglise keelselt	'The date of the first entry'
Klassifikaator	Selgitus_est	Klassifikaatori selgitus eesti keelselt	'Kuupäev,

Olemitüübi nimi	Atribuudi nimi	Atribuudi definitsioon	Näiteväärtus
			millal ettevõtte kanti registrisse'
Klassifikaator	Selgitus_eng	Klassifikaatori selgitus inglise keelselt	'Date when the company was founded'
Klassifikaator	Kehtiv	Tunnus, kas klassifikaator on hetkel aktiivne või mitte.	TRUE

Lisa 2

Tabel 6. Andmebaasi tabelite detailed kirjeldused.

Mass_paring						
Veeru nimi	Tüüp/ domeen	Tüübi/ domeeni nimi	Pikkus	Võimalikud väärtused	Vaike- väärtus	Kohus- tuslik
id	T	suur täisarv		Unikaalsed väärtused. Automaatselt genereeritavad väärtused. Samm=1		jah
algus_aeg	T	kuupäev/ kellaaeg/ ajavöönd			praegune aeg	jah
kinnitus_aeg	T	Kuupäev/ kellaaeg/ ajavöönd				ei
lopp_aeg	T	kuupäev/ kellaaeg/ ajavöönd				ei
keel	T	tekst	3			ei
tasuta_klient	T	tõeväärtus			FALSE	jah
email	T	tekst				ei
isikukood	T	tekst	20			ei
isikukood_riik	T	tekst	3			ei
kasutaja_id	T	suur täisarv				ei
kl_paringu_olek	T	tekst			'1'	jah
sessions_id	T	tekst	64			jah
Primary Key (id) Foreign Key (isikukood_riik) REFERENCES ar.ak_riigid (id) ON UPDATE CASCADE Foreign Key (kasutaja_id) REFERENCES teabesysteem.lep_kasutajad (kasutaja_id) ON UPDATE CASCADE						

Foreign Key (kl_paringu_olek) REFERENCES teabesysteem.mass_kl_paringu_olek (id) ON UPDATE CASCADE

Foreign Key (sessions_id) REFERENCES teabesysteem.sessions (sessions_id)

Mass_tulemus

Veeru nimi	Tüüp/ domeen	Tüübi/ domeeni nimi	Pikkus	Võimalikud väärtused	Vaike- väärtus	Kohus- tuslik
id	T	suur täisarv		Unikaalsed väärtused. Automaatselt genereeritavad väärtused. Samm=1		jah
ridade_arv	T	täisarv			0	jah
algus_aeg	T	kuupäev/ kellaeg/ ajavöönd				jah
lopp_aeg	T	kuupäev/ kellaeg/ ajavöönd			Praegune aeg	jah
tulemus		bait				ei
hind		valuuta				ei
valuuta	T	tekst	3			ei
mass_paring_id	T	suur täisarv				jah

Primary Key (id)

Foreign Key (valuuta) REFERENCES ar.ak_valuuta (id) ON UPDATE CASCADE

Foreign Key (mass_paring_id) REFERENCES teabesysteem.mass_paring (id) ON UPDATE CASCADE ON DELETE CASCADE

Mass_sys_paringu_loomise_parameeter

Veeru nimi	Tüüp/ domeen	Tüübi/ domeeni nimi	Pikkus	Võimalikud väärtused	Vaike- väärtus	Kohus- tuslik
id	T	tekst		Unikaalsed väärtused.		jah
kehtiv	T	tõeväärtus			TRUE	jah
skeemi_nimi	T	tekst			'ar'	jah
tabeli_nimi	T	tekst				ei
veeru_nimi	T	tekst				ei

eesliide	T	tekst				ei
seosta_funktsiooni_kaudu	T	tõeväärtus			FALSE	jah
seosta_ainult_viimasega	T	tõeväärtus			FALSE	jah
seosta_ainult_viimasega_primaarv otme_veerg	T	tekst				ei
jarjekord	T	täisarv				jah
where_tingimus	T	tõeväärtus			FALSE	jah
where_tingimus_vaikimisi_vaste	T	tekst				ei
where_tingimus_vordub_null	T	tõeväärtus			FALSE	jah
where_tingimus_viit_vastele_tabel	T	tekst				ei
where_tingimus_viit_vastele_veer g	T	tekst				ei
select_tingimus	T	tõeväärtus			FALSE	jah
select_jarjekord	T	täisarv				ei
alias_est	T	tekst				ei
alias_eng	T	tekst				ei
seos	T	tekst				ei
where_tingimus_liik	T	tekst				ei
kriteerium_id	T	tekst				ei
andme_skoop_id	T	tekst				ei
<p>Primary Key (id)</p> <p>Alternate Key (andme_skoop_id, select_jarjekord)</p> <p>Alternate Key (andme_skoop_id, kehtiv, skeemi_nimi, tabeli_nimi, veeru_nimi, eesliide, seos)</p> <p>Alternate Key (kriteerium_id, kehtiv, skeemi_nimi, tabeli_nimi, veeru_nimi, eesliide, seos)</p> <p>Foreign Key (seos) REFERENCES teabesysteem.mass_sys_paringu_loomise_parameeter (id) ON UPDATE CASCADE</p> <p>Foreign Key (where_tingimus_liik) REFERENCES teabesysteem.mass_sys_kriteeriumi_vordlemise_tingimus (id) ON UPDATE CASCADE</p> <p>Foreign Key (kriteerium_id) REFERENCES teabesysteem.mass_sys_kriteerium (id) ON UPDATE CASCADE</p> <p>Foreign Key (andme_skoop_id) REFERENCES teabesysteem.mass_sys_andmete_skoop (id) ON UPDATE CASCADE</p>						
Mass_sys_andmete_skoop						
Veeru nimi	Tüüp/ domeen	Tüübi/ domeeni nimi	Pikkus	Võimalikud väärtused	Vaike- väärtus	Kohus- tuslik

id	T	tekst		Unikaalsed väärtused.		jah
nimetus_est	T	tekst				jah
nimetus_eng	T	tekst				jah
selgitus_est	T	tekst				ei
selgitus_eng	T	tekst				ei
kehtiv	T	tõeväärtus			TRUE	jah
teisi_valistav	T	tõeväärtus			FALSE	jah
valista_identifikaatorid	T	tekst[]				ei
eraldi_hinnastatav	T	tõeväärtus			TRUE	jah
jarjekord	T	täisarv		Unikaalsed väärtused		jah
hinnastamise_liik	T	tekst				jah
viit_kuuluvusele	T	tekst				ei

Primary Key (id)

Alternate Key (jarjekord)

Foreign Key (hinnastamise_liik) REFERENCES teabesysteem.mass_kl_hinnastamise_liik (id) ON UPDATE CASCADE

Foreign Key (viit_kuuluvusele) REFERENCES teabesysteem.mass_sys_andmete_skoop (id) ON UPDATE CASCADE

Mass_sys_kriteerium

Veeru nimi	Tüüp/ domeen	Tüübi/ domeeni nimi	Pikkus	Võimalikud väärtused	Vaike- väärtus	Kohus- tuslik
id	T	tekst		Unikaalsed väärtused.		jah
nimetus_est	T	tekst				jah
nimetus_eng	T	tekst				jah
selgitus_est	T	tekst				ei
selgitus_eng	T	tekst				ei
kehtiv	T	tõeväärtus			TRUE	jah
teisi_valistav	T	tõeväärtus			FALSE	jah
valista_identifikaatorid	T	tekst[]				ei
poordumise_url	T	tekst				jah
jarjekord	T	täisarv		Unikaalsed		jah

				väärtused		
Primary Key (id)						
Alternate Key (jarjekord)						
Mass_sys_kriteeriumi_vordlemise_tingimus						
Veeru nimi	Tüüp/ domeen	Tüübi/ domeeni nimi	Pikkus	Võimalikud väärtused	Vaike- väärtus	Kohus- tuslik
id	T	tekst		Unikaalsed väärtused.		jah
vaartus	T	tekst		Unikaalsed väärtused.		jah
selgitus_est	T	tekst				ei
Primary Key (id)						
Mass_krit_registrikood						
Veeru nimi	Tüüp/ domeen	Tüübi/ domeeni nimi	Pikkus	Võimalikud väärtused	Vaike- väärtus	Kohus- tuslik
id	T	suur täisarv		Unikaalsed väärtused. Automaatselt genereeritavad väärtused. Samm=1		jah
registrikood	T	täisarv				jah
mass_paringu_kriteerium_id	T	suur täisarv				jah
Primary Key (id)						
Alternate Key (registrikood, mass_paringu_kriteerium_id)						
Foreign Key (mass_paringu_kriteerium_id) REFERENCES teabesysteem.mass_paringu_kriteerium (id) ON UPDATE CASCADE ON DELETE CASCADE						
Mass_krit_aadress						
Veeru nimi	Tüüp/ domeen	Tüübi/ domeeni nimi	Pikkus	Võimalikud väärtused	Vaike- väärtus	Kohus- tuslik
id	T	suur täisarv		Unikaalsed väärtused. Automaatselt genereeritavad väärtused.		jah

				Samm=1		
nimetus	T	tekst				jah
ehak_kood	T	tekst	4			jah
omavalitsus	T	tekst	4			jah
mass_paringu_kriteerium_id	T	suur täisarv				jah
<p>Primary Key (id)</p> <p>Alternate Key (nimetus, ehak_kood, omavalitsus, mass_paringu_kriteerium_id)</p> <p>Foreign Key (mass_paringu_kriteerium_id) REFERENCES teabesysteem.mass_paringu_kriteerium (id) ON UPDATE CASCADE ON DELETE CASCADE</p>						
Mass_krit_tegevusala						
Veeru nimi	Tüüp/domeen	Tüübi/domeeni nimi	Pikkus	Võimalikud väärtused	Vaikeväärtus	Kohustuslik
id	T	suur täisarv		Unikaalsed väärtused. Automaatselt genereeritavad väärtused. Samm=1		jah
tase	T	täisarv				jah
maaramata	T	tõeväärtus			FALSE	jah
emtak	T	tekst				jah
mass_paringu_kriteerium_id	T	suur täisarv				jah
emtak_id	T	tekst				jah
vanem_emtak_id	T	tekst				ei
<p>Primary Key (id)</p> <p>Alternate Key (emtak_id, mass_paringu_kriteerium_id)</p> <p>Foreign Key (mass_paringu_kriteerium_id) REFERENCES teabesysteem.mass_paringu_kriteerium (id) ON UPDATE CASCADE ON DELETE CASCADE</p> <p>Foreign Key (emtak_id) REFERENCES ar.ak_emtak (id) ON UPDATE CASCADE</p> <p>Foreign Key (vanem_emtak_id) REFERENCES ar.ak_emtak (id) ON UPDATE CASCADE</p>						
Mass_krit_ettevotte_olek						

Veeru nimi	Tüüp/ domeen	Tüübi/ domeeni nimi	Pikkus	Võimalikud väärtused	Vaike- väärtus	Kohus- tuslik
id	T	suur täisarv		Unikaalsed väärtused. Automaatselt genereeritavad väärtused. Samm=1		jah
ettevotte_olek	T	tekst				jah
mass_paringu_kriteerium_id	T	suur täisarv				jah
Primary Key (id) Alternate Key (ettevotte_olek, mass_paringu_kriteerium_id) Foreign Key (mass_paringu_kriteerium_id) REFERENCES teabesysteem.mass_paringu_kriteerium (id) ON UPDATE CASCADE ON DELETE CASCADE Foreign Key (ettevotte_olek) REFERENCES ar.kl_ettevotja_olekud (id) ON UPDATE CASCADE						
Mass_krit_oiguslik_vorm						
Veeru nimi	Tüüp/ domeen	Tüübi/ domeeni nimi	Pikkus	Võimalikud väärtused	Vaike- väärtus	Kohus- tuslik
id	T	suur täisarv		Unikaalsed väärtused. Automaatselt genereeritavad väärtused. Samm=1		jah
oiguslik_vorm	T	tekst				jah
oigusliku_vormi_alaliik	T	tekst				ei
mass_paringu_kriteerium_id	T	suur täisarv				jah
Primary Key (id) Alternate Key (oiguslik_vorm, oigusliku_vormi_alaliik, mass_paringu_kriteerium_id) Foreign Key (mass_paringu_kriteerium_id) REFERENCES teabesysteem.mass_paringu_kriteerium (id) ON UPDATE CASCADE ON DELETE CASCADE Foreign Key (oiguslik_vorm) REFERENCES ar.kl_oiguslikud_vormid (id) ON UPDATE CASCADE Foreign Key (oigusliku_vormi_alaliik) REFERENCES ar.kl_oiguslikud_vormi_alaliigid (id) ON UPDATE CASCADE						
Mass_krit_esmakanne						

Veeru nimi	Tüüp/ domeen	Tüübi/ domeeni nimi	Pikkus	Võimalikud väärtused	Vaike- väärtus	Kohus- tuslik
id	T	suur täisarv		Unikaalsed väärtused. Automaatselt genereeritavad väärtused. Samm=1		jah
esmakande_kuupaev	T	kuupaev				jah
mass_paringu_kriteerium_id	T	suur täisarv		Unikaalsed väärtused		jah
Primary Key (id) Alternate Key (mass_paringu_kriteerium_id) Foreign Key (mass_paringu_kriteerium_id) REFERENCES teabesysteem.mass_paringu_kriteerium (id) ON UPDATE CASCADE ON DELETE CASCADE						
Mass_paringu_kriteerium						
Veeru nimi	Tüüp/ domeen	Tüübi/ domeeni nimi	Pikkus	Võimalikud väärtused	Vaike- väärtus	Kohus- tuslik
id	T	suur täisarv		Unikaalsed väärtused. Automaatselt genereeritavad väärtused. Samm=1		jah
mass_paring_id	T	suur täisarv				jah
kriteerium_id	T	tekst				jah
Primary Key (id) Alternate Key (mass_paring_id, kriteerium_id) Foreign Key (mass_paring_id) REFERENCES teabesysteem.mass_paring (id) ON UPDATE CASCADE ON DELETE CASCADE Foreign Key (kriteerium_id) REFERENCES teabesysteem.mass_sys_kriteerium (id) ON UPDATE CASCADE						
Mass_paringu_skoobi_valikute_tulemused						
Veeru nimi	Tüüp/ domeen	Tüübi/ domeeni nimi	Pikkus	Võimalikud väärtused	Vaike- väärtus	Kohus- tuslik

id	T	suur täisarv		Unikaalsed väärtused. Automaatselt genereeritavad väärtused. Samm=1		jah
kirje_pohine_tulemuste_arv	T	täisarv			0	jah
objekti_pohine_tulemuste_arv	T	täisarv			0	jah
mass_paring_id	T	suur täisarv				jah
andme_skoop_id	T	tekst				jah
<p>Primary Key (id)</p> <p>Alternate Key (mass_paring_id, andme_skoop_id)</p> <p>Foreign Key (mass_paring_id) REFERENCES teabesysteem.mass_paring (id) ON UPDATE CASCADE ON DELETE CASCADE</p> <p>Foreign Key (andme_skoop_id) REFERENCES teabesysteem.mass_sys_andmete_skoop (id) ON UPDATE CASCADE</p>						
Mass_kl_hinnastamise_liik						
Veeru nimi	Tüüp/ domeen	Tüübi/ domeeni nimi	Pikkus	Võimalikud väärtused	Vaike- väärtus	Kohus- tuslik
id	T	tekst		Unikaalsed väärtused.		jah
nimetus_est	T	tekst				jah
nimetus_eng	T	tekst				jah
selgitus_est	T	tekst				ei
selgitus_eng	T	tekst				ei
kehtiv	T	tõeväärtus			TRUE	jah
<p>Primary Key (id)</p>						
Mass_kl_paringu_olek						
Veeru nimi	Tüüp/ domeen	Tüübi/ domeeni nimi	Pikkus	Võimalikud väärtused	Vaike- väärtus	Kohus- tuslik

id	T	tekst		Unikaalsed väärtused.		jah
nimetus_est	T	tekst				jah
nimetus_eng	T	tekst				jah
selgitus_est	T	tekst				ei
selgitus_eng	T	tekst				ei
kehtiv	T	tõeväärtus			TRUE	jah
Primary Key (id)						

Lisa 3

```
--plpgsql trigger andmebaasi lisatava
--päringu loogika sisestuse üle

CREATE OR REPLACE FUNCTION
teabesysteem.kontrolli_sys_paringu_loomise_parameetri_s
obivust() RETURNS trigger AS
$BODY$
DECLARE
    status boolean = FALSE;
BEGIN
    SELECT EXISTS (SELECT 1 FROM
information_schema.columns WHERE
table_schema=new.skeemi_nimi AND
table_name=new.tabeli_nimi AND
column_name=new.veeru_nimi) INTO status;
    IF status = FALSE THEN
        SELECT EXISTS (SELECT 1 FROM pg_proc WHERE
proname = new.tabeli_nimi) INTO status;
        IF status = FALSE THEN
            SELECT EXISTS (SELECT 1 FROM
information_schema.columns WHERE
table_schema=new.skeemi_nimi AND
table_name=new.tabeli_nimi) INTO status;
            IF status = FALSE THEN
                RAISE EXCEPTION '%.%.%' - Sellist tabeli
veergu andmebaasis ei leidu, kontrolli sisendit.',
new.skeemi_nimi, new.tabeli_nimi, new.veeru_nimi;
            ELSE
                IF new.veeru_nimi IS NOT NULL THEN
                    RAISE EXCEPTION '%.%.%' - Sellist tabeli
veergu andmebaasis ei leidu, kontrolli sisendit.',
new.skeemi_nimi, new.tabeli_nimi, new.veeru_nimi;
                END IF;
            END IF;
        END IF;
    ELSE
        SELECT EXISTS (SELECT 1 FROM
teabesysteem.mass_sys_paringu_loomise_parameeter a
WHERE a.kehtiv = new.kehtiv AND
a.skeemi_nimi=new.skeemi_nimi AND
a.tabeli_nimi=new.tabeli_nimi AND a.eesliide <>
new.eesliide) INTO status;
        IF status = TRUE THEN
            RAISE EXCEPTION '%.% %' -
mass_sys_paringu_loomise_parameetrisse on juba selline
tabel salvestatud, aga teise eesliitega, kontrolli
sisendit', new.skeemi_nimi, new.tabeli_nimi,
```

```

new.eesliide;
    ELSE
        SELECT EXISTS (SELECT 1 FROM
teabesysteem.mass_sys_paringu_loomise_parameeter a
WHERE a.kehtiv = new.kehtiv AND a.jarjekord <>
new.jarjekord AND a.skeemi_nimi=new.skeemi_nimi AND
a.tabeli_nimi=new.tabeli_nimi AND (a.kriteerium_id =
new.kriteerium_id OR a.andme_skoop_id =
new.andme_skoop_id)) INTO status;
        IF status = TRUE THEN
            IF new.kriteerium_id IS NOT NULL THEN
                RAISE EXCEPTION '"%.%" -
mass_sys_paringu_loomise_parameetrisse on juba selline
tabel salvestatud selle kriteeriumi kohta, aga teise
jarjekorra numbriga, kontrolli sisendit',
new.skeemi_nimi, new.tabeli_nimi, new.jarjekord;
            END IF;
            IF new.andme_skoop_id IS NOT NULL THEN
                RAISE EXCEPTION '"%.%" -
mass_sys_paringu_loomise_parameetrisse on juba selline
tabel salvestatud selle andme_skoobi kohta, aga teise
jarjekorra numbriga, kontrolli sisendit',
new.skeemi_nimi, new.tabeli_nimi, new.jarjekord;
            END IF;
        END IF;
    END IF;
    END IF;
    END IF;
    IF new.select_tingimus AND new.where_tingimus THEN
        RAISE EXCEPTION '"%.%" - kirje ei tohi olal
samalajal select tingimus ja where tingimus.',
new.skeemi_nimi, new.tabeli_nimi, new.veeru_nimi;
    END IF;
    IF new.seosta_ainult_viimasega AND
new.seosta_ainult_viimasega_primaarvotme_veerg IS NULL
THEN
        RAISE EXCEPTION '"%.%" - viimasega seostates
tuleb täpsustada ka seostatava tabeli primaarvotme
veerg.', new.skeemi_nimi, new.tabeli_nimi,
new.veeru_nimi;
    END IF;
    IF new.where_tingimus AND
new.where_tingimus_viit_vastele_tabel IS NOT NULL THEN
        SELECT EXISTS (SELECT 1 FROM
information_schema.columns WHERE
table_name=new.where_tingimus_viit_vastele_tabel AND
column_name=new.where_tingimus_viit_vastele_veerg) INTO
status;
        IF status = FALSE THEN
            RAISE EXCEPTION '"%.%" - WHERE tingimuse viit
veergu ei eksisteeri.',
new.where_tingimus_viit_vastele_tabel,
new.where_tingimus_viit_vastele_veerg;
        END IF;
    END IF;

```

```
        RETURN NEW;  
    END;  
    $BODY$  
LANGUAGE 'plpgsql';
```

Joonis 24. Andmebaasi trigger, mis kontrollib päringu loogika sisestust.