TALLINNA TEHNIKAÜLIKOOL
Infotehnoloogia teaduskond

Ants Kristjan Rooma 175050IDDR

Privaatsust arvestavate kasutajapäringute mooduli arendus OÜ Net Group näitel

Diplomitöö

Juhendaja: Kristiina Hakk
PhD

Tallinn 2021

Autorideklaratsioon

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: Ants Kristjan Rooma

[17.05.2021]

Annotatsioon

Isikuandmete kaitse üldmäärus annab kodanikele muuhulgas õiguse igal ajahetkel paluda andmete haldajalt teavet andmete töötlemisest ning väljavõtet enda kohta kogutud andmetest. Lisaks on isikul õigus nõuda enda isikuandmete kustutamist. Tarkvaraarendusettevõte OÜ Net Group on paljude infosüsteemide arendaja ja haldaja, kus kogutakse ja töödeldakse isikuandmeid. Praegusel hetkel puudub OÜ Net Group poolt hallatavatel süsteemidel kiire ja lihtne viis isikuandmetest väljavõtete tegemiseks või andmete anonümiseerimiseks. Sellest probleemist on ajendatud ka käesoleva uurimustöö eesmärk.

OÜ Net Groupis läbiviidud küsitlusest selgus, et kasutajapäringute täitmine ei ole sage, kuid võtab olenevalt andmemudeli keerukusest aega. Probleemi lahendamiseks valmis prototüübina .NET Core raamistikul kirjutatud konsoolirakendus, mis võimaldab erinevatel projektidel koostada isikuandmete väljavõtteid ja teostada andmete anonümiseerimist.

Lõputöö on kirjutatud eesti keeles ning sisaldab teksti 26. leheküljel, kuus peatükki, üheksa joonist, ühte tabelit.

Abstract

Development of Privacy-Aware User Query Module: the Case of OÜ Net Group

Amongst other things the General Data Protection Regulation gives citizens the right to ask for information concerning data collection and the right to see a statement containing their personal information. In addition, a person has the right to ask for the removal of their personal data. Software company OÜ Net Group develops and manages a lot of information systems, which handle and store people's personal data. Currently, systems managed by OÜ Net Group do not have a fast and simple method for generating personal data statements or for personal data anonymization. This current situation is the catalyst for this thesis.

A survey conducted in OÜ Net Group revealed that user queries based on these two rights are not very populaar, but can take up a considerable amount of time based on the complexity of the data model. In order to alleviate these issues, a .NET Core console application was made. The console application gives different projects the ability to generate personal data statements and conduct personal data anonymization.

The thesis is in Estonian and contains 26 pages of text, six chapters, nine figures, one table.

Lühendite ja mõistete sõnastik

Camel case	Kirjastiil, kus sõnade vahel puuduvad tühikud ja iga sõna algab suure tähega
CSV	*Comma Separated Values,* koma eraldusega väärtuste fail
DOCX	Microsoft Wordi failiformaat
DOTX	Microsoft Wordi malliformaat
GDPR	*General Data Protection Regulation,* isikuandmete kaitse üldmäärus
HTML	*Hyper Text Markup Language,* hüpertekst-märgistuskeel
IP-aadress	*Internet Protocol* aadress, standardse võrgukihi protokolli aadress
JSON	*JavaScript Object Notation,* andmevahetusvorming
SQL	*Structured Query Language*, struktuurpäringukeel
XLSX	Microsoft Exceli failiformaat
XML	*Extensible Markup Language,* laiendatav märgistuskeel

Sisukord

Jooniste loetelu

Tabelite loetelu

1 Sissejuhatus

E-riigi toimimiseks ja arenguks on vaja saavutada kodanike usaldus erinevate e-teenuste, infosüsteemide ja teenusepakkujate vastu. Kodanik peab tundma end turvaliselt oma andmete jagamiseks ja nende kasutamise lubamiseks. Kindlustunne, et andmeid kogutakse, kasutatakse ja säilitatakse ausalt, turvaliselt ja õigesti on seega väga oluline. Eestis reguleerib kodanike andmete kasutamist isikuandmete kaitse seadus, mis põhineb Euroopa Parlamendi poolt vastu võetud isikuandmete kaitse üldmäärusel.

Isikuandmete kaitse üldmäärus annab kodanikele muuhulgas õiguse igal ajahetkel paluda andmete haldajalt teavet andmete töötlemisest ning väljavõtet enda kohta kogutud andmetest. Lisaks on isikul õigus nõuda enda isikuandmete kustutamist. Tarkvaraarendusettevõte OÜ Net Group on paljude infosüsteemide arendaja ja haldaja, kus kogutakse ja töödeldakse isikuandmeid. Töötades Net Groupis tarkvaraarendajana puutub autor kokku vajadusega teha andmebaasidest väljavõte konkreetse kasutaja andmetest. Ette tuleb ka isikute soovi andmeid kustutada, mis andmete terviklikkuse säilitamiseks tähendab enamasti andmete anonümiseerimist.

Kirjeldatud ülesanded on hetkel väga ajamahukad, sest nõuavad igal korral juhtumipõhist lähenemist ja analüüsi. Praegusel hetkel puudub OÜ Net Groupi poolt hallatavatel süsteemidel kiire ja lihtne viis isikuandmetest väljavõtete tegemiseks või andmete anonümiseerimiseks. Sellest probleemist on ajendatud ka käesoleva uurimustöö eesmärk. Diplomitöö eesmärk on automatiseerida, standardiseerida ja lihtsustada isikuandmetest väljavõtete genereerimist ja andmestiku anonümiseerimist. Sealjuures disainida ja pakkuda välja prototüüp, mida saab rakendada kirjeldatud probleemi lahendamiseks.

Esimeses peatükis antakse ülevaade isikuandmete kaitsest ja selle mõjust andmete töötlejatele. Järgmises peatükis analüüsitakse Net Groupi praeguseid protsesse ja viiakse läbi küsimustik välja selgitamaks probleeme kasutajapäringute teostamisel. Järgmisena kirjeldatakse loodavat kasutajapäringute täitmise moodulit, selle seadistamist, tööd ja väljundit. Viimases peatükis analüüsitakse tulemusi, arvutatakse mooduli arendamise tasuvus ja kirjeldatakse võimalike tulevikusuundasid.

2 Taust

Selles peatükis antakse ülevaade isikuandmete kaitse üldmäärusest ja füüsilise isiku õigustest nimetatud määruse raames. Detailsemalt käsitletakse käesoleva töö fookuses olevaid üldmääruse osi ning nende rakendamist infosüsteemides.

Lisaks tuuakse välja turul olemasolevad võimalused kasutajapäringute täitmiseks ja analüüsitakse nende sobivust OÜ Net Groupis rakendamiseks.

2.1 Isikuandmete kaitse

Eesti Vabariigis on iga kodaniku põhiseaduslikuks õiguseks õigus isikuandmete kaitsele. Kaitstakse inimese eraelu puutumatust ja vabadust. Isikuandmete kaitset reguleerib Eestis 2019. aastal vastu võetud isikuandmete kaitse seadus, mis põhineb 2016. aastal Euroopa Parlamendis vastu võetud Euroopa Liidu isikuandmete kaitse üldmäärusel [1]. Euroopa Liidu isikuandmete kaitse üldmäärus on laiemalt tuntud oma inglise keelse lühendi GDPR (*general data protection regulation*) järgi. Euroopa Parlamendi määrus asendab eelmise, 1995. aastal vastu võetud, määruse mis ei arvestanud uute tehnoloogiliste arengutega [2].

Üldmääruse eesmärk on kaitsta füüsiliste isikute õigusi nende kohta käivate andmete kogumisel, töötlemisel ja säilitamisel [3]. Isikuandmeid tohib töödelda üksnes õiguspärastel eesmärkidel ja vajalik on asjaomase isiku nõusolek [4]. Lisaks peab olema tagatud ligipääs inimese kohta kogutud isikuandmetele ning infole nende andmete töötlemise kohta. Viimase all peetakse silmas millisel määral, eesmärgil ja õiguslikul alusel andmeid on kasutatud [5].

Eestis teostab kontrolli isikuandmete kaitse üle Andmekaitse Inspektsioon. Andmekaitse Inspektsioon jaotab isikuandmed kolmeks liigiks [6]:

- Tavalised isikuandmed – andmed, mille kaudu saab füüsilist isikut tuvastada. Nendeks on näiteks nimi, isikukood, asukohateave, IP-aadress, samuti füüsilised, füsioloogilised, geneetilised, vaimsed, majanduslikud, kultuurilised ja muud tuvastamist võimaldavad tunnused ja nende kombinatsioonid.

- Tundlikud isikuandmed – andmed, mis valmistavad füüsilise isiku eraelule suuremat ohtu, kui näiteks andmete avaldamisega kaasneb oht elule, tervisele,

identiteedivargusele, võib kaasneda varaline ja mainekahju. Sellised andmed on näiteks sotsiaalabi saamine, kriminaal- ja väärteomenetluse andmed, krediitkaardi andmed, varanduslik seis, sideandmed, reaalajas asukohatuvastuse andmed, krediidireiting.

- Eriliiki isikuandmed – andmed, mida võib pidada delikaatseks. Sellised andmed on näiteks rassiline või etniline päritolu, poliitilised vaated, usulised või filosoofilised veendumused, ametiühingusse kuulumine, biomeetrilised andmed, terviseandmed, seksuaalse sättumus.

Isikuandmete kaitse reeglid ei kehti juriidilistele isikutele, andmete isiklikul otstarbel kasutamisele, manuaalselt andmete töötlemisele, kui andmetest ei looda andmekogumit ja anonüümsetele andmetele. Lisaks ei kehti isikuandmete kaitse reeglid ka riikliku julgeoleku tagamisel [7].

Isikutel on igal ajahetkel võimalik paluda teavet tema andmete töötlemise kohta. Sealjuures on isikul õigus teada kes ja milliseid andmeid töötleb, mis eesmärgil seda tehakse ja kellele andmeid edastatakse. Lisaks on isikul õigus teada andmete säilitamise aega, andmete allikat ja ka seda, et kas tema andmete põhjal tehakse automatiseerituid otsuseid [8]. Automatiseeritud otsuste tegemine tähendab infotehnoloogiliste vahenditega ja ilma inimese sekkumiseta andmete töötlemise tulemuste pealt otsuste langetamist. Automatiseeritud otsustamisega võib käia kaasas profiilianalüüs ehk isiku andmete hindamine erinevate prognooside tegemiseks [9].

Isik võib oma andmeid parandada, nende töötlemist piirata või paluda andmed kustutada [10] [11] [12]. Samuti on võimalik paluda andmete töötlejal enda kohta kogutud andmete ülekandmist teisele ettevõttele ehk teisele andmete töötlejale. Selleks peavad andmed olema struktureeritult masinloetaval kujul. Kusjuures andmete ülekandmine ei tähenda andmete kustutamist algse töötleja andmebaasist [13].

Kõik andmeid töötlevad ettevõtted peavad endale koostama andmekaitsetingimused, kus sisaldub kogu info eelnevalt tutvustatud füüsilise isiku õigustest ja nende rakendamise detailidest konkreetses ettevõttes. Andmekaitsetingimustest tuleb isikut teavitada koheselt, kui tema kohta andmeid koguma hakatakse, ning lisaks ka siis, kui ettevõtte andmekaitsetingimused ajas muutuvad [5].

Käesolev lõputöö keskendub edaspidi isikute õigusele igal ajahetkel paluda teavet tema kohta kogutud andmetest ja nende töötlemisest. Lisaks ka isiku õigusele nõuda enda kohta käivate andmete kustutamist.

2.2 Isikuandmetest väljavõtete tegemine

Isiku juurdepääsu tema kohta kogutud isikuandmetele ja andmete töötlemise teabe andmist reguleerivad artiklid 12-15 Euroopa Parlamendi poolt 2016. aastal vastu võetud isikuandmete kaitse üldmääruses. Artikkel 12 ütleb muuhulgas, et andmetöötlejal on teabe ja andmete väljastamiseks õigus nõuda isiku tuvastamist, et olla kindel isiku identiteedis. Lisaks on teabe esitamise kohta öeldud järgmist: „Vastutav töötleja võtab asjakohased meetmed, et esitada … teave ning teavitada teda … isikuandmete töötlemisest kokkuvõtlikult, selgelt, arusaadavalt ning lihtsasti kättesaadavas vormis, kasutades selget ja lihtsat keelt … Kõnealune teave esitatakse kirjalikult või muude vahendite abil, sealhulgas asjakohasel juhul elektrooniliselt.“ [14]

Artiklites 13 ja 14 on välja toodud nõutud teave andmete töötlemise kohta vastavalt sellele, kas andmed on kogutud isikult endalt või pärinevad andmed muust allikast. Esitatavaks teabeks on näiteks töötleja kontakt, töötlemise eesmärk, teave isiku õiguste kohta ja kaebuste esitamise kohta [14]. Kuna teabe hulk on suur ja tegemist on standardse infoga soovitab Andmekaitse Inspektsioon koostada töötlejal andmekaitsetingimuste dokumendi ja avalikustada see oma veebilehel. See võimaldab isikutel igal ajahetkel tutvuda organisatsiooni andmekaitsepoliitikaga ja saada üldist teavet andmete töötlemise kohta [15].

Kui isik soovib lisaks üldistele andmekaitsetingimustele tutvuda ka enda kohta kogutud andmetega, tuleb esitada sellekohane taotlus. Töötlejal on kohustus rahuldada taotlus ilma tasu küsimata ühe kuu jooksul [15]. Isiku kohta kogutud andmetega tutvumist reguleerib artikkel 15, kus on öeldud, et lisaks andmetele tuleb isikule väljastada [14]:

- andmete töötlemise eesmärk;
- andmete liik;
- vastuvõtjad, kellele andmeid on avalikustatud;

- andmete säilitamise ajavahemik;
- teave õiguse kohta taotleda enda andmete parandamist, kustutamist või töötlemise piiramist;
- teave kaebuse esitamise kohta;
- andmete allikas, kui andmeid ei ole saadud isikult endalt;
- teave automatiseeritud otsuste ja profiilianalüüsi tegemise kohta.

Üks isikuandmete väljavõtte tegemise erijuht on andmete ülekandmine. Seda reguleerib isikuandmete kaitse üldmääruse artikkel 20, kus on kirjas järgnev: „Andmesubjektil on õigus saada teda puudutavaid isikuandmeid, mida ta on vastutavale töötlejale esitanud, struktureeritud, üldkasutatavas vormingus ning masinloetaval kujul ning õigus edastada need andmed teisele vastutavale töötlejale, ilma et vastutav töötleja, kellele kõnealused isikuandmed on esitatud, seda takistaks, …“ [14] Andmekaitse Inspektsioon lisab, et üldkasutatav vorming on selline faili formaat, mis ei ole seotud mõne kitsalt kasutatava kommertstarkvaraga. Masinloetav vorming on näiteks XML, JSON või CSV [15].

2.3 Isikuandmete kustutamine

Isikuandmete kaitse üldmääruse artikkel 17 ütleb, et isikul on õigus nõuda oma isikuandmete kustutamist [14]. Andmete kustutamise õigus ei ole absoluutne ja sõltub olukorrast. Andmed tuleb kustutada, kui neid ei ole enam vaja või on neid kasutatud ebaseaduslikult. Andmete kustutamise taotlust ei pea rahuldama, kui andmeid säilitatakse mõne teise õiguse, näiteks sõnavabaduse või teadusuuringute kaitseks [16].

Pole üheselt selge mida mõeldakse andmete kustutamise all (inglise keeles *erasure*), ning on palju arutatud selle üle, et kas andmete täieliku kustutamise asemel piisab ka andmete anonümiseerimisest [17]. Andmekaitse Inspektsioon defineerib anonümiseerimist kui andmetest kõikide jälgede kaotamist, mis võiksid viia isiku tuvastamiseni [15]. Lisaks tuleb märkida, nagu ka peatükis 2.1 välja toodud, et isikuandmete kaitse reeglid ei kehti anonüümsetele andmetele.

Kinnituseks, et andmete kustutamise taotluse võib rahuldada ka läbi isikuandmete anonümiseerimise, võib toetuda 2018. aastal Austrias vastu võetud kohtuotsusele. Sealne

ettevõte otsustas enda kliendi andmete kustutamise taotluse rahuldada selliselt, et kustutas isikuandmed ainult osaliselt. Teine osa andmetest otsustati anonümiseerida, et neid andmed edaspidi andmeanalüütikas kasutada. Kuna alles jäänud andmete pealt ei olnud võimalik isikut tuvastada võeti vastu otsus mis leidis, et isikuandmete anonümiseerimine vastab isikuandmete kaitse reeglitele [17].

Seejuures peab anonümiseerimise protsess olema testitud ja korrektselt järgitud [17]. Anonümiseerimist ei tohi segamini ajada pseudonüümimisega, mida Andmekaitse Inspektsioon defineerib järgnevalt: „Pseudonüümine tähendab äratuntavate isikuandmete asendamist varjunimede, numbrikoodide ja muude tunnustega, mida asjassepuutumatud isikud ei oska ära arvata.“ [15] Anonümiseerimise ja pseudonüümimise vahe seisneb selle tegevuse tagasipööratavuses. Anonümiseerimine on tagasipööramatu ja lõplik andmete umbisikustamine [15].

2.4 Olemasolevad lahendused

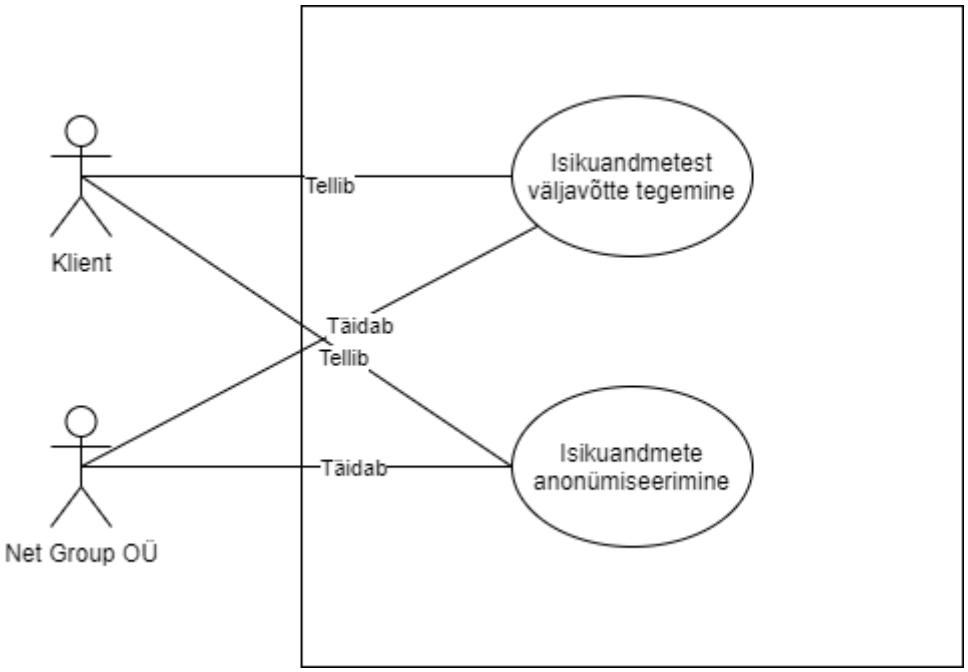
Universaalset toodet või teeki isikuandmete väljavõtte loomiseks või isikuandmete anonümiseerimiseks turult leida ei õnnestunud. Selle taga võib olla asjaolu, et sellise tööriista loomine mis kataks ära kõik ärilised vajadused, erinevad platvormid ja andmebaasid ning arhitektuurilised võimalused on väga keerukas ja seega ka kulukas. Võimalik, et praegusel hetkel on GDPR veel liiga uudne ning sellega kohanemine ja erinevate tööriistade loomine võtab aega.

Suured ettevõtted nagu Facebook või Google on arendanud sama probleemi lahendamiseks eraldi funktsionaalsuse oma olemasolevate toodete külge. Facebook võimaldab kasutajal alla laadida kogu info kasutaja tegevuse kohta. Samuti on võimalik andmed paluda kustutada. Google pakub sarnaseid võimalusi lubades alla laadida näiteks kogu kasutaja otsinguajaloo või asukoha info. Kogu tegevus käib läbi kasutajaliideste ja on automaatne ehk puudub vajadus suhelda kasutajatoega [18]. Eesti infosüsteemides nagu näiteks Pilet.ee, Osta.ee või Sais.ee automaatsed võimalused puuduvad. Oma andmetest väljavõtte saamiseks või andmete kustutamiseks tuleb pöörduda süsteemi haldaja poole kas läbi infotelefoni või saates e-kirja.

3 Kasutajapäringute täitmise hetkeolukord OÜ Net Groupis

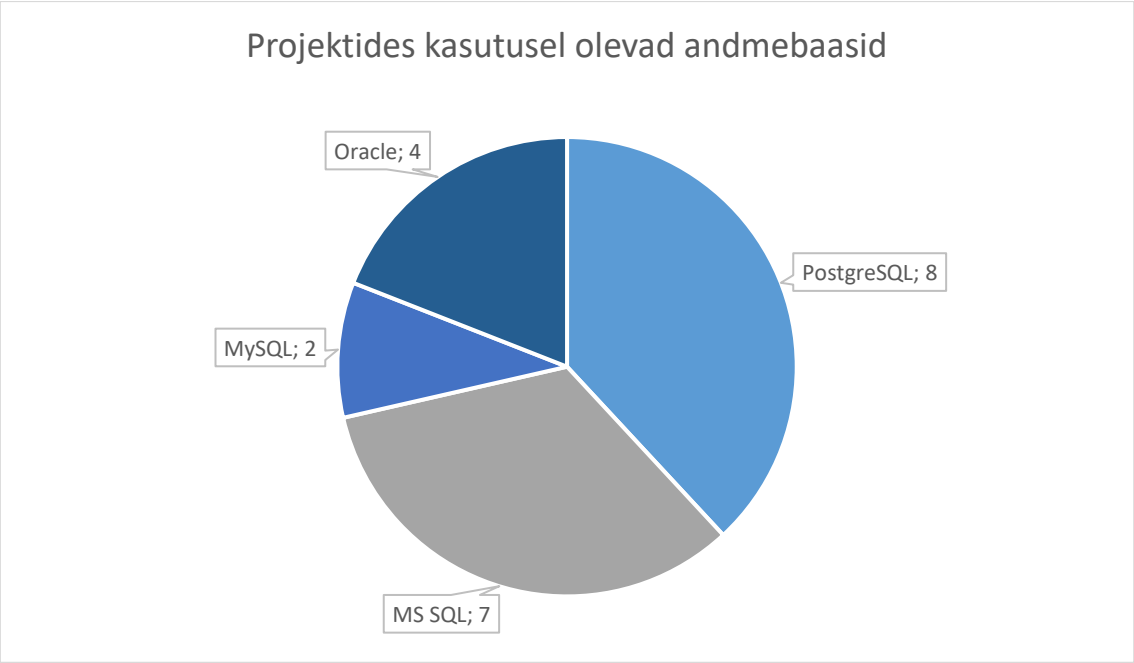
OÜ Net Group on tarkvaraarenduse ettevõte, kelle klientideks on erinevad organisatsioonid, kes soovivad oma toimimist läbi tarkvaralahenduste efektiivsemaks muuta. Lisaks tarkvara arendusele pakub Net Group ka olemasolevatele infosüsteemidele hooldusteenust, mis hõlmab tarkvara kasutamisel tekkivate probleemide lahendamist, parenduste tegemist, tarkvara testimist ja vajadusel ka andmete analüüsi ja korrastamist. Sellise teenuse pakkumisel tuleb ette ka kasutajapäringute täitmist, millest käesoleva töö fookuses on isikuandmetest väljavõtte tegemine ja isikuandmete kustutamine või anonümiseerimine. Isikuandmete kustutamise võimalus sõltub tarkvara andmemudelist. Standardse lahenduse loomiseks ja andmete terviklikkuse hoidmiseks käsitleme edaspidi kõiki isikuandmete kustutamise päringuid kui andmete anonümiseerimist.

OÜ Net Group on kõikidel juhtudel andmete volitatud töötleja, sest andmeid töödeldakse vastutava töötleja ehk OÜ Net Groupi kliendi nimel. Isikuandmete vastutav töötleja on OÜ Net Groupi klient, kellel on oma äriliste eesmärkide saavutamiseks vaja isikuandmeid töödelda [15]. Seega tulevad kõik kasutajapäringud Net Groupile läbi andmete vastutava töötleja ja Net Groupil endal puudub kohustus päringu esitajat identifitseerida või päringu täitmise üle otsustada. Fookuses olevad kasutuslood on toodud ka järgneval joonisel (Joonis 1).



Joonis 1. Kasutusjuhtude diagramm.

Järgnevates alapeatükkides käsitletakse detailselt mõlemat joonisel toodud kasutusjuhtu ja kirjeldatakse nende täitmise hetkeolukorda. Lisaks viidi Net Group'i siseselt läbi küsitlus, et välja selgitada kasutajapäringute tegemisele kuluv aeg ja päringute esinemise sagedus erinevate projektide lõikes. Küsitlus on toodud Lisas 2. Küsitlusele vastasid 21 projekti esindajad, mis teeb 42% kõikidest Net Groupi projektidest arvestades, et keskmiselt on Net Groupis arendamisel ja haldamisel 50 erinevat projekti korraga. Küsitlusest selgus ka erinevate andmebaaside kasutamise statistika. Kõige populaarsem on PostgreSQL andmebaas, millele ei jää palju alla ka MS SQL. Täpsem jaotus on toodud järgneval joonisel (Joonis 2).

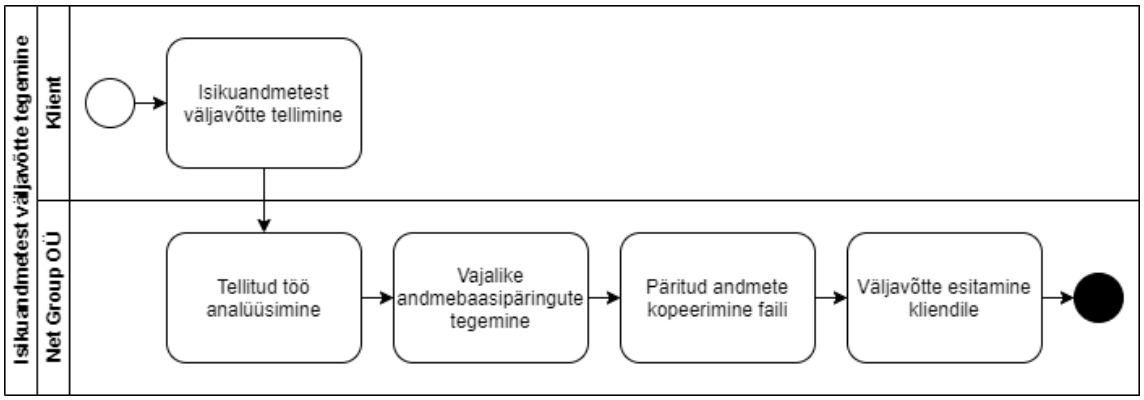


Joonis 2. Projektides kasutusel olevad andmebaasid.

3.1 Isikuandmetest väljavõtete tegemise hetkeolukord

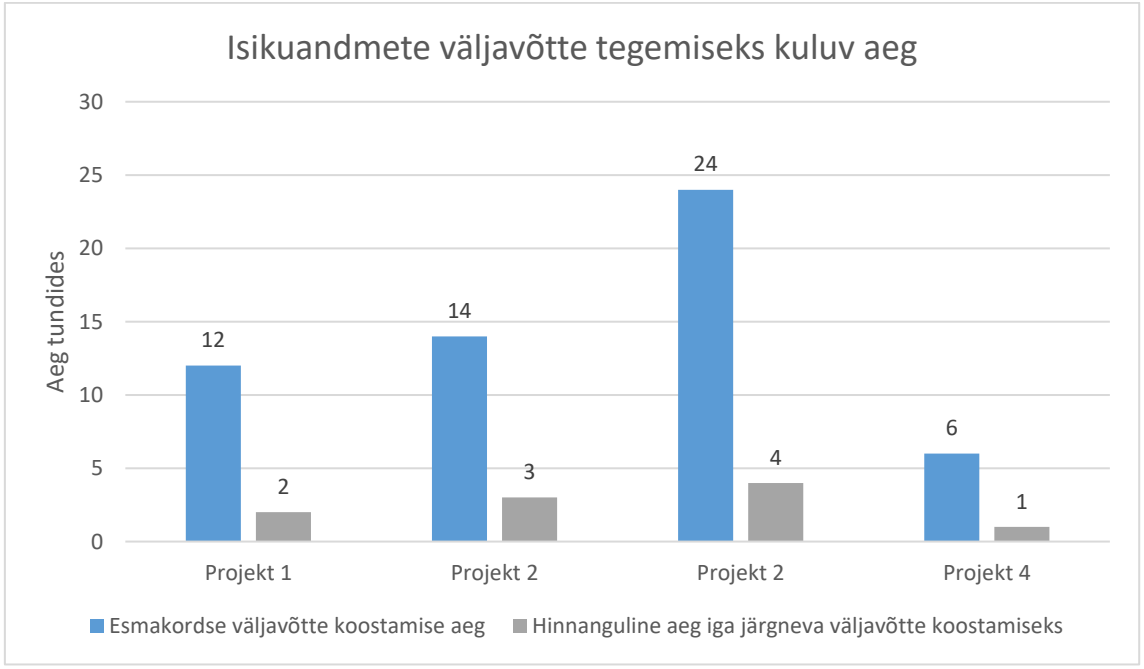
Isikuandmetest väljavõtte tegemise protsess algab sellega, kui Net Groupi klient otsustab Net Groupilt konkreetse isiku isikuandmetest väljavõtte tellida. Esmalt tuleb Net Groupi arendustiimil ülesandepüstitust analüüsida ja välja selgitada, millistes andmebaasitabelites info on salvestatud. Seejärel jooksutatakse vajalikud andmebaasipäringud. Esmakordselt uues projektis isikuandmetest väljavõtet tehes kulub analüüsimisele kindlasti rohkem aega, kui korduvatel juhtudel. Tavaliselt salvestab väljavõtet tegev tarkvaraarendaja väljavõtte tegemiseks vaja läinud päringud, et järgmistel kordadel oleks protsess kiirem ja oleks võimalik varem loodud päringuid taaskasutada.

Kui andmed on olemas, tuleb need kopeerida kliendile edastatavasse faili. Failiformaat on seejuures iga arendaja enda valida, kuid üldjuhul kasutatakse Microsoft Exceli failiformaati XLSX. Viimaks edastatakse isikuandmetest väljavõtet sisalduv fail kliendile ja sellega on Net Groupi jaoks väljavõtte tegemise protsess lõppenud. Klient peab veel arvestama, et isikuandmete kaitse üldmääruse kohaselt tuleb tal väljavõtet täiendada andmete töötlemise kohta käiva infoga nagu on tutvustatud peatükis 2.2. Kirjeldatud protsess on välja toodud ka järgneval joonisel (Joonis 3).



Joonis 3. Isikuandmetest väljavõtte tegemise hetkeolukord.

Läbiviidud küsitlusest selgus, et isikuandmetest väljavõtte tegemist on ette tulnud ühel korral neljas erinevas projektis. Kolm nendest projektidest kasutavad PostgreSQL andmebaasi ja ühes projektis on kasutusel MS SQL andmebaas. Keskmiselt kulus neil esmakordselt isikuandmetest väljavõtte tegemisele 14 tundi. Täpsed ajalised mahud on toodud järgmisel joonisel (Joonis 4). Joonisel on välja toodud ka vastajate ajaline hinnang iga järgneva väljavõtte koostamiseks.



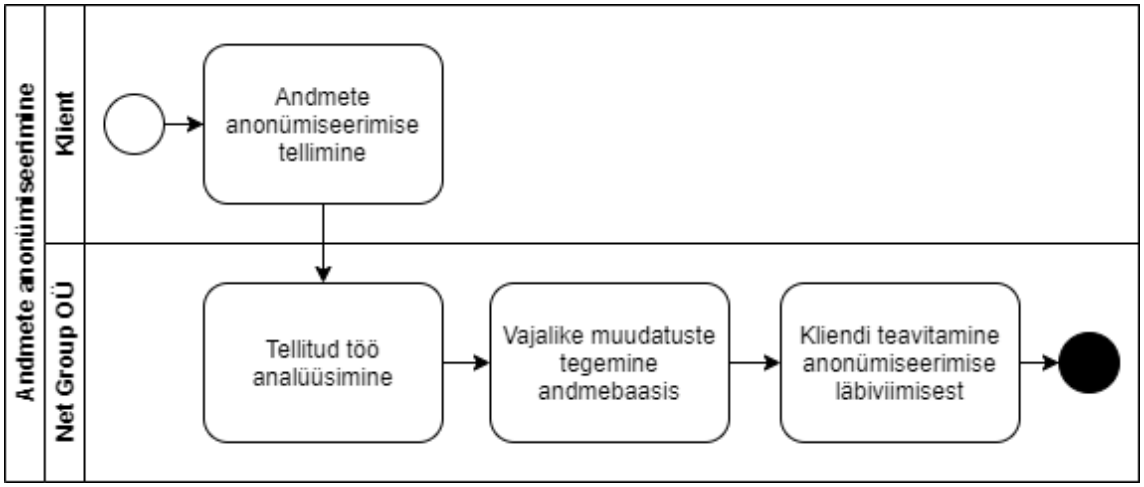
Joonis 4. Isikuandmete väljavõtte tegemiseks kuluv aeg.

Kõik ülejäänud vastajad, kes ei olnud varasemalt isikuandmetest väljavõtte koostamisega kokku puutunud hindasid ülesande täitmise ajaliseks mahuks 16-80 tundi.

3.2 Isikuandmete anonümiseerimise hetkeolukord

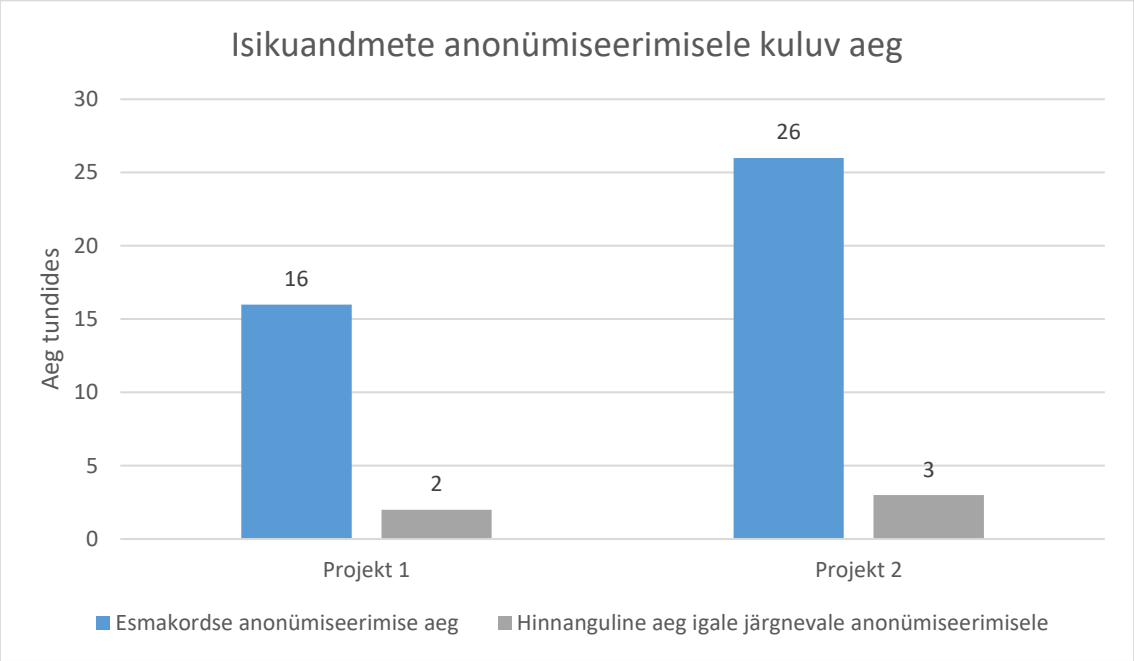
Isikuandmete anonümiseerimise protsess sarnaneb suuresti isikuandmete väljavõtte tegemise protsessile alates sellega, kui klient tellib Net Group'ilt mõne konkreetse isiku andmete anonümiseerimise. Seejärel tuleb Net Groupi tarkvaraarendajal välja selgitada millised on need isikuandmed, mis kuuluvad anonümiseerimisele. Jällegi on esimest korda anonümiseerimise teostamine palju ajamahukam, kui tegevuse hilisem kordamine. Andmete anonümiseerimisel tuleb olla kindel, et hiljem ei ole võimalik alles jäänud andmete pealt isikut tuvastada ning, et anonümiseerimine oleks tagasipööramatu.

Kui anonümiseerimine on läbi mõeldud, analüüsitud ja ka testitud tuleb teha vajalikud muudatused andmebaasis, millega eemaldatakse kõik võimalikud viited konkreetsele isikule. Viimase sammuna teavitab Net Group töö tellinud klienti andmete anonümiseerimise läbiviimisest. Kirjeldatud protsess on toodud järgneval joonisel (Joonis 5).



Joonis 5. Isikuandmete anonümiseerimise hetkeolukord.

Läbiviidud küsitlusest selgus, et ainult kahes projektis on ette tulnud isikuandmete anonümiseerimist või kustutamist. Mõlemas projektis on kasutusel PostgreSQL andmebaas. Küsitlusest ei selgunud andmete anonümiseerimise või kustutamise täpne lahenduskäik, kuid ajaline maht esmakordsel anonümiseerimisel oli keskmiselt 21 tundi. Täpsed ajalised mahud on toodud järgmisel joonisel (Joonis 6). Joonisel on välja toodud ka vastajate ajaline hinnang iga järgneva anonümiseerimise teostamiseks. Kõik ülejäänud vastajad, kes ei olnud varasemalt isikuandmete anonümiseerimisega kokku puutunud hindasid ülesande lahendamise hinnanguliseks ajaks 24-48 tundi.



Joonis 6. Isikuandmete anonümiseerimisele kuluv aeg.

4 Kasutajapäringute täitmise moodul

Kasutajapäringute täitmise lihtsustamiseks ja standardiseerimiseks on mõistlik luua kõikides Net Groupi projektides kasutatav ühtne päringute täitmise moodul. Moodulile seatud funktsionaalsed nõuded on järgnevad:

- Kasutaja peab saama genereerida konkreetse isiku isikuandmete väljavõtet.
- Kasutaja peab saama anonümiseerida konkreetse isiku andmeid. Anonümiseerimise all mõeldakse isikuandmete asendamist kas *null* väärtusega või mõne muu väärtusega, millega ei ole võimalik isikut tuvastada.

Seejuures kehtivad moodulile järgnevad mittefunktsionaalsed nõuded:

- Mooduli arendamisel on kasutatud läbivalt inglise keelt.
- Muutujad ja funktsioonid on nimetatud selgitavalt, kasutatakse *camel case* stiili.
- Funktsionaalsus on struktureeritud ja taaskasutatav.

Mooduli loomisel otsustas autor oma eelnevast kogemusest tulenevalt kasutada .NET tehnoloogiaid, et hoida kokku aega, mis kuluks mõne uue raamistikuga kohanemiseks. Samuti on pooled Net Groupi projektid suunatud .NET tehnoloogiatele ja olemas on kõrge kompetents. See lihtsustab abi küsimist või olukorda, kus keegi teine soovib ise moodulit täiendada. .NET on Microsofti poolt loodud platvorm erinevate rakenduste ehitamiseks. Kasutatavateks programmeerimisekeelteks on C#, F# või Visual Basic. Läbi aja on Microsoft loonud .NET erinevaid implementatsioone nagu .NET Framework või .NET Core [19].

.NET Framework on väga laialdaselt kasutusel olev ja põhjalikult dokumenteeritud raamistik, mida saab kasutada ehitamaks erinevaid rakendusi mis on mõeldud kasutamiseks Windowsi operatsioonisüsteemis. .NET Core on uuema generatsiooni implementatsioon .NET tehnoloogiatest, mis on avatud lähtekoodiga, suurema jõudlusega kui .NET Framework ja kasutatav erinevate operatsioonisüsteemidega [20].

Käesolevas töös võetakse kasutusele just .NET Core, sest lisaks eelnevalt välja toodud eelistele toetab .NET Core ka mikroteenuseid [20].

.NET Core võimaldab arendada rakenduse veebirakendusena, konsoolirakendusena või töölaua rakendusena kasutades Windows Forms, Windows Presentation Foundation või Windows Universal Platform teeki. Kolm viimasena nimetatud teeki on mõeldud ainult Windowsi töölaua rakenduste jaoks ja ei ole kasutatavad teistes operatsioonisüsteemides [21]. Seega taandub valik veebirakenduse ja konsoolirakenduse vahele. Autor otsustas esialgse prototüübi luua konsoolirakendusena, et keskenduda põhifunktsionaalsuse arendamisele. Hiljem on võimalik konsoolirakenduse funktsionaalsus üle kanda veebirakendusse, kuid prototüübi loomise faasis ei ole mõistlik veebirakenduse keskkondade või serverite seadistamise peale aega kulutada. Lisaks on konsoolirakendust lihtne testida ja ei pea muretsema rakenduse turvalisuse pärast, sest rakendus käivitatakse iga arendaja juures lokaalselt.

Enne mooduli käivitamist tuleb väärtustada järgmised rakenduse globaalsed muutujaid:

- *StatementTemplatePath* – viide väljavõtte mallile, väljavõtte malli on täpsemalt kirjeldatud peatükis 4.2.
- *StatementOutputDirectory* – genereeritud väljavõtte salvestamise asukoht.
- *StatementInputsPath* – viide isikuandmete väljavõtte sisendfailile.
- *StatementConfigurationPath* – viide isikuandmete väljavõtte konfiguratsiooni failile.
- *PersonalDataRemovalInputsPath* – viide isikuandmete anonümiseerimise sisendfailile.
- *PersonalDataRemovalConfigurationPath* – viide isikuandmete anonümiseerimise konfiguratsiooni failile.

Rakenduse käivitamisel kontrollitakse esimese asjana globaalsete muutujate olemasolu. Vajaliku väärtuse puudumisel katkestatakse kogu protsess ja teavitatakse sellest kasutajat. Järgnevates alapeatükkides on kirjeldatud mooduli isikuandmete väljavõtte

genereerimise ja isikuandmete anonümiseerimise funktsionaalsusi, selgitatud on mooduli käivitamisele eelnevat seadistamist, mooduli tööd ja tulemusi.

4.1 Isikuandmetest väljavõtte genereerimine

Konsoolirakenduse isikuandetest väljavõtte funktsionaalsuse kasutamiseks erinevates projektides tuleb koostada iga projekti kohta eraldi XML tüüpi konfiguratsiooni fail, kus juurelemendiks on *StatementConfiguration*. Näide sellisest failist on toodud järgmisel joonisel (Joonis 7). Selle elemendi alla luuakse massiiv *Databases*, kus on kohustuslik kirjeldada vähemalt üks alamelement *Database*. *Database* kirjeldab andmebaasi ja koosneb kahest elemendist – *ConnectionString* ja *Tables*. *ConnectionString* on kohustuslik väärtus, mida kasutatakse andmebaasiga ühenduse loomiseks. *Tables* on andmebaasitabelite massiiv, mille alla kuuluvad ühte tabelit kirjeldavad elemendid nimega *Table*. Kõikidest *Table* elementidest saadav informatsioon kuvatakse lõpuks isikuandmete väljavõttele. *Table* elemendi alamelemendid on:

- *NameInDatabase* – tabeli nimi andmebaasis.
- *DisplayName* – tabeli nimi väljavõttes kuvamiseks.
- *SelectColumns* – massiiv andmebaasitabeli veergudest.
    - *SelectColumn* – tabeli veerg väljavõttes kuvamiseks.
        - *NameInDatabase* – veeru nimi andmebaasis.
        - *DisplayName* – veeru nimi väljavõttes kuvamiseks.
- *Filter* – SQL keeles kirjutatud lause (WHERE osa) andmete filtreerimiseks. Lauses saab kasutada {N} muutujat, kus N on väljavõtte sisendfailis oleva väärtuse indeks.
- *RawSql* – mittekohustuslik element, SQL keeles kirjutatud terviklik lause. Lauses kasutatavad veerud peavad olema kirjeldatud samas järjekorras *SelectColumns* elemendis. Lauses saab kasutada {N} muutujat, kus N on väljavõtte sisendfailis oleva väärtuse indeks. Elemendi olemasolul eelistatakse seda *Filter* elemendile.

- *DisplayStyle* – viide väljavõtte formaadile. Võimalikud väärtused on *KeyValueDataTable* või *CascadingDataTable*, mis on kirjeldatud peatükis 4.2.

```
<?xml version="1.0" encoding="utf-8"?>
<StatementConfiguration xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:xsd="http://www.w3.org/2001/XMLSchema">
	<Databases>
		<Database>
			<ConnectionString></ConnectionString>
			<Tables>
				<Table>
					<NameInDatabase></NameInDatabase>
					<DisplayName></DisplayName>
					<SelectColumns>
						<SelectColumn>
							<NameInDatabase>
							</NameInDatabase>
							<DisplayName></DisplayName>
						</SelectColumn>
					</SelectColumns>
					<Filter></Filter>
					<DisplayStyle></DisplayStyle>
					<RawSql></RawSql>
				</Table>
			</Tables>
		</Database>
	</Databases>
</StatementConfiguration>
```

Joonis 7. Isikuandmete väljavõtte konfiguratsiooni fail.

Lisaks konfiguratsiooni failile kasutab konsoolirakendus ka sisendina XML faili, kus on määratud isiku andmed, kelle kohta päringuid jooksutatakse. Faili sisu on toodud järgneval joonisel (Joonis 8). Sisendfaili juurelemendiks on *ActivityInputs*, mille alla luuakse massiiv *Values*. Sisendiks antavad muutujad defineeritakse selle massiivi alamelementides nimega *Value*.

```
<?xml version="1.0" encoding="utf-8"?>
<ActivityInputs xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:xsd="http://www.w3.org/2001/XMLSchema">
	<Values>
		<Value></Value>
	</Values>
</ActivityInputs>
```

Joonis 8. Sisendfail.

Isikuandmetest väljavõtte genereerimiseks tuleb jooksutada käsku *GdprPersonalData create-statement*. Rakendus valideerib esmalt globaalsete muutujate olemasolu, seejärel isikuandmetest väljavõtte genereerimise konfiguratsiooni faili olemasolu ja ka faili sisu, et vajalikud väärtused oleksid olemas. Rakenduse töö katkestatakse juhul, kui kõik vajalikud elemendid ei ole täidetud. Kolmanda sammuna kontrollitakse sisendfaili olemasolu. Kui konfiguratsiooni faili väärtustes *Filter* või *RawSql* esineb {N} muutujaid, teostab rakendus nende muutujate asendamise vastavalt sisendfailis olevatele väärtustele.

Edasi saab rakendus kokku panna vajalikud SQL laused ja pärida andmebaasist andmed. Päritud andmed paigutatakse väljavõtte mallile, mida on kirjeldatud järgmises peatükis. Viimaks salvestatakse genereeritud väljavõte globaalsetes muutujates määratud asukohta. Kui mõne sammu käigus tekib ootamatu viga, peatatakse protsess ning teavitatakse sellest kasutajat.

4.2 Väljavõtte formaat

Tehniliselt kõige lihtsam oleks väljavõte koostada JSON, CSV või TXT failina. Tegemist on universaalsete ja masinate poolt loetavate failiformaatidega, kuid tavakasutaja jaoks ei ole selliste failide lugemine eriti mugav. Samuti puudub võimalus neid faile kujundada. Tavakasutaja jaoks on harjumuspärasem lugeda PDF või DOCX formaadis faili, mille plussiks on ka võimalus kujundada faili visuaali. Sedasi oleks võimalik genereerida lõplik väljavõte ja puuduks vajadus genereeritud väljavõtet täiendada.

Valides PDF või DOCX formaadis väljavõtte vahel osutus valituks DOCX formaat koos DOTX malliformaadiga. DOCX failiformaati on võimalik avada vabavaraliste programmidega nagu LibreOffice ja tasuta veebiteenustega nagu Google Docs ja Word Online. DOTX malliformaadis on võimalik ära kirjeldada isikuandmete väljavõtte struktuur ja kohustuslikud sisuelemendid. Erinevate projektide jaoks mallide loomine ja loodud mallide muutmine ei ole üleliia keerukas ja seda tegevust ei pea tegema arendaja rollis isik. PDF formaadi puhul nõuaks aga väljavõtte loomine ja kujundamine kasutajalt HTML ja CSS teadmisi ning erinevate projektide spetsiifilised nõuded võivad kaasa tuua lisaarendusi. DOTX malliformaadi puhul on üheks piiranguks Microsoft Wordi programmi olemasolu.

Kasutades DOCX formaadis väljavõtet on iga projekti isikuandmete väljavõtte DOTX mallis võimalik kasutada projektile omast kujundust, stiile ja logosid. Malli algusesse on võimalik iga projekti puhul ära kirjeldada staatiline info GDPR määruse täitmisest. Vajalik info on loetletud peatükis 2.2. Üldinfo mallil defineerimine kaotab ära ajalise kulu, mis hetkel kulus peale väljavõtte koostamist selle täiendamisele.

Kohustuslike sisuelementide loomiseks kasutatakse mallil Microsoft Word Content Control elemente. Kohustuslikud sisuelemendid on:

- *StatementCreationDateAndTime* – elemendi sisu täidetakse väljavõtte genereerimisel kuupäeva ja kellaajaga.
- *Content* – elemendi sisu täidetakse väljavõtte genereerimisel andmebaasist päritud andmetega.

Andmebaasist päritud andmete kuvamiseks kasutatakse eeldefineeritud stiile:

- *DataTableNameHeading* – stiil rakendatakse andmebaasi tabeli nimele.
- *DataTableRowHeading* – stiil rakendatakse andmebaasi tabeli kirjete eraldajatele.
- *KeyValueDataTable* – kahe veeruga tabeli stiil. Tabeli esimeses veerus on andmebaasi tabeli kirje kuvatavate atribuutide nimetused. Tabeli teises veerus on atribuudi väärtus. Stiili kasutatakse andmetabelite puhul, kus kirjete veergude väärtused ei ole pika sisuga.
- *CascadingDataTable* – ühe veeruga tabeli stiil. Tabelil on 2*n* rida, kus *n* tähistab andmebaasi tabelis kuvatavate atribuutide arvu. Tabeli ridade järjekorra numbrid algavad ühest. Paaritute ridade puhul kuvatakse välja andmebaasi tabeli atribuudi kuvatav nimetus. Paaris ridade puhul kuvatakse välja andmetabeli kirje väärtus, mis kuulub eelmisele tabeli real kuvatud atribuudile. Stiili kasutatakse andmetabelite puhul, kus kirjete veergude väärtused on pika sisuga.

4.3 Isikuandmete anonümiseerimine

Isikuandmete anonümiseerimise funktsionaalsuse kasutamiseks tuleb samuti kõikides projektides koostada XML konfiguratsiooni fail. Näidis sellisest failist on toodud

järgmisel joonisel (Joonis 9). Konfiguratsiooni faili juurelemendiks on *PersonalDataRemovalConfiguration*, mille alla luuakse sarnaselt isikuandmetest väljavõtte genereerimise konfiguratsiooni failile massiiv *Databases*. Selle elemendi alamelemendid on *Database*, kus on kohustuslik kirjeldada ära vähemalt üks andmebaas. *Database* koosneb kahest elemendist – *ConnectionString* ja *Tables*. *ConnectionString* on kohustuslik väärtus, mida kasutatakse andmebaasiga ühenduse loomiseks. *Tables* on andmebaasitabelite massiiv, mille alla kuuluvad ühte tabelit kirjeldavad elemendid nimega *Table*. Andmete anonümiseerimine toimub kõikides *Table* elementides kirjeldatud tabelites ning selleks on vaja kirjeldada järgnevad *Table* alamelemendid:

- *NameInDatabase* – tabeli nimi andmebaasis.
- *UpdateColumns* – massiiv andmebaasitabeli veergudest.
    - *UpdateColumn* – anonümiseeritav tabeli veerg.
        - *NameInDatabase* – veeru nimi andmebaasis.
        - *ReplacementAction* – viide anonümiseerimise tegevusele. Lubatud on neli erinevat väärtust: *SetNull*, *ReplaceString*, *ReplaceSubstring*, *ReplaceInteger*. *SetNull* asendab veeru väärtuse null väärtusega. *ReplaceString* asendab veeru väärtuse kasutaja poolt antud või rakenduse poolt genereeritud sõnega. *ReplaceSubstring* asendab ainult ette antud osa veeru väärtusest kasutaja poolt antud või rakenduse poolt genereeritud sõnega. *ReplaceInteger* asendab veeru väärtuse kasutaja poolt antud või rakenduse poolt genereeritud arvulise väärtusega.
        - *ReplaceWhat* – asendatav osa tabeli veerus olevast väärtusest. Kohustuslik element ainult juhul, kui *ReplacementAction* väärtus on *ReplaceSubstring*. Selle elemendi väärtusena võib kasutada muutujat {N}, kus N on sisendfailis oleva väärtuse indeks.
        - *ConstantValue* – veeru uus väärtus. Selle elemendi väärtusena võib kasutada muutujat {N}, kus N on sisendfailis oleva väärtuse indeks. Kohustuslik ainult siis, kui *ReplacementAction* väärtuseks

on *ReplaceString*, *ReplaceSubString* või *ReplaceInteger* ja element *RandomValueLength* on väärtustamata.

- *RandomValueLength* – veeru uue genereeritud väärtuse täisarvuline pikkus. Kohustuslik ainult siis, kui *ReplacementAction* väärtuseks on *ReplaceString*, *ReplaceSubstring* või *ReplaceInteger* ja element *ConstantValue* on väärtustamata. *ReplaceString* ja *ReplaceSubString* puhul genereeritakse vastava pikkusega sõne. *ReplaceInteger* puhul näitab selle elemendi väärtus, mis on täisarvu bittide kogus. Lubatud väärtused: 8, 16, 32, 64. Bittide arv näitab, mis on täisarvu minimaalne ja maksimaalne väärtus. Elemendi olemasolul eelistatakse seda *ConstantValue* elemendile.

- *Filter* - SQL keeles kirjutatud lause (WHERE osa) andmete filtreerimiseks. Lauses saab kasutada {N} muutujat, kus N on väljavõtte sisendfailis oleva väärtuse indeks.

- *RawSql* - mittekohustuslik element, SQL keeles kirjutatud terviklik lause. Lauses kasutatavad veerud peavad olema kirjeldatud samas järjekorras *SelectColumns* elemendis. Lauses saab kasutada {N} muutujat, kus N on väljavõtte sisendfailis oleva väärtuse indeks. Elemendi olemasolul eelistatakse seda *Filter* elemendile.

```
<?xml version="1.0" encoding="utf-8"?>
<PersonalDataRemovalConfiguration
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:xsd="http://www.w3.org/2001/XMLSchema">
	<Databases>
		<Database>
			<ConnectionString></ConnectionString>
			<Tables>
				<Table>
					<RawSql></RawSql>
					<NameInDatabase></NameInDatabase>
					<UpdateColumns>
						<UpdateColumn>
							<NameInDatabase>
							</NameInDatabase>
							<ReplacementAction>
							</ReplacementAction>
							<ReplaceWhat></ReplaceWhat>
							<ConstantValue></ConstantValue>
							<RandomValueLength>
							</RandomValueLength>
						</UpdateColumn>
					</UpdateColumns>
					<Filter></Filter>
				</Table>
			</Tables>
		</Database>
	</Databases>
</PersonalDataRemovalConfiguration>
```

Joonis 9. Isikuandmete anonümiseerimise konfiguratsiooni fail.

Lisaks konfiguratsiooni failile kasutab rakendus sisendfaili, mille sisu sama isikuandmetest väljavõtte genereerimise sisendfailile ja on toodud varasemalt (Joonis 8). Isikuandmete anonümiseerimiseks tuleb konsoolirakendus jooksutada käsuga *GdprPersonalData remove-data*. Rakendus valideerib esmalt globaalsete muutujate, konfiguratsiooni faili ja sisendfaili olemasolu ja failide sisu. Rakenduse töö katkestatakse, kui kõik vajalikud väärtused ei ole täidetud. Kui konfiguratsiooni faili väärtustes *Filter*, *RawSql*, *ReplaceWhat* või *ConstantValue* esineb {N} muutujaid, teostab rakendus nende muutujate asendamise vastavalt sisendfailis olevatele väärtustele. Järgmisena saab rakendus kokku panna SQL laused, mis käivitatakse vastavas andmebaasis ning teostatakse andmete muutmine. Kui mõne sammu käigus tekib ootamatu viga, peatatakse protsess ning teavitatakse sellest kasutajat.

5 Tulemused

Diplomitöö raames valmis kasutajapäringute täitmise mooduli prototüüp, milleks on .NET Core raamistikul baseeruv konsoolirakendus. Erinevad projektid saavad kasutada rakendust isikuandmete väljavõtete loomiseks ja isikuandmete anonümiseerimiseks. Tegevuste läbiviimiseks kasutatakse DOTX formaadis malli ja erinevaid konfiguratsiooni faile, mis seadistatakse projekti halduri poolt.

Prototüübi arendamisele kulus autoril 56 töötundi. Esmalt testiti loodud rakendust edukalt autori kohalikus arenduskeskkonnas, seejärel katsetati rakenduse toimimist ühe OÜ Net Group projekti testkeskkonnas, mis kasutab PostgreSQL andmebaasi. Isikuandmete väljavõtte konfiguratsiooni faili täitmisele kulus selles projektis 7 tundi, isikuandmete anonümiseerimise konfiguratsiooni faili täitmise peale 5 tundi. Rakenduse töökäiku demonstreeriti ka antud projekti kliendile, kes jäi tulemusega rahule. Kliendile meeldis, et ilma arendaja sekkumiseta on lihtne muuta isikuandmete väljavõtte malli visuaali. Klient saab loodud anonümiseerimise funktsionaalsust kasutada ka olukorras, kus anonümiseerimine peab toimuma perioodiliselt ja automaatselt. Selleks saab kasutada konsoolirakenduse ja Windows Task Scheduler komponendi kombinatsiooni. Lisaks pakuti välja, et tulevikus võiks juurde arendada võimaluse isikuandmete väljavõte automaatseks edastamiseks päringu esitanud isikule.

Rakenduse tasuvuse hindamiseks on koostatud järgnev tabel (Tabel 1).

Tabel 1. Loodud rakenduse tasuvus.

	Isikuandmete väljavõtte genereerimine	**Isikuandmete anonümiseerimine**	**Kokku**
Küsitlusest selgunud keskmine ajakulu kasutaja päringu esmaseks täitmiseks	14 tundi	21 tundi	**35 tundi**
Küsitlusest selgunud hinnanguline keskmine ajakulu iga järgneva kasutaja päringu täitmiseks	2,5 tundi	2,5 tundi	5 tundi
Rakenduse arendus	-	-	**56 tundi**
Loodud konsoolirakenduse esmakordsele seadistusele ja testimisele kulunud aeg testprojektis	7 tundi	5 tundi	**12 tundi**
Igale järgnevale päringu tegemisele kulunud aeg testprojektis	0,5 tundi	0,1 tund	0,6 tundi

Net Group peab loodud konsoolirakenduse kasutusele võtma vähemalt kolmes projektis, et rakenduse loomisele kulunud aeg ennast ära tasuks:

$$\frac{56}{35-12} \approx 2{,}43 \approx 3$$

Konsoolirakenduse juurutamine igas järgnevas projektis oleks täiendav ajaline ja rahaline kasum Net Groupile.

Loodud rakendust võiks tulevikus edasi arendada, et toetatud oleks suurem hulk erinevaid andmebaase. Lisaks pakuks klientidele lisaväärtust võimalus isikuandmete väljavõtet genereerida ka muudes failiformaatides, näiteks HTML. Konsoolirakenduse põhifunktsionaalsuse osast on võimalik luua ka NuGet teek, mis kataks väljavõtte faili koostamise funktsionaalsuse. Autor plaanib aga prototüübi põhjal edasi arendada .NET Core platvormil baseeruva veebirakenduse. Uue rakenduse versiooniga kaasneks graafiline kasutajaliides kasutaja päringute konfiguratsioonide loomiseks ja muutmiseks, võimekus automatiseerida kasutaja andmete väljavõtte edastamist, võimekus luua ja jälgida taustategevusi, mis anonümiseerivad kasutaja andmeid. Iga projekt saaks seadistada enda isikliku veebiteenuse. Kui Net Group otsustab kasutada ühtset veebirakendust kogu ettevõtte projektide kasutajapäringute täitmiseks, nõuaks veebirakendus ka turvalist andmekäsitlust ja autoriseerimise ning autentimise kihti.

6 Kokkuvõte

Isikuandmete kaitse üldmäärus annab kodanikele muuhulgas õiguse paluda andmete haldajalt teavet andmete töötlemisest ning väljavõtet enda kohta kogutud andmetest. Lisaks on isikul õigus nõuda enda isikuandmete kustutamist. Tarkvaraarendusettevõte OÜ Net Group arendab ja haldab paljusid infosüsteeme, kus kogutakse ja töödeldakse isikuandmeid. Praegusel hetkel puudub OÜ Net Group poolt hallatavatel süsteemidel kiire ja lihtne viis isikuandmetest väljavõtete tegemiseks või andmete anonümiseerimiseks. Probleemist tulenevalt oli töö eesmärk automatiseerida, standardiseerida ja lihtsustada isikuandmetest väljavõtete genereerimist ja andmestiku anonümiseerimist. Sealjuures disainida ja pakkuda välja prototüüp, mida saab rakendada kirjeldatud probleemi lahendamiseks.

Töö käigus analüüsiti esmalt isikuandmete kaitset Eestis ja anti ülevaade isiku õigustest pärida infot enda kohta kogutud andmetest ja nende töötlemisest. Keskenduti kahele päringule – väljavõte isikuandmetest ja andmete kustutamine või anonümiseerimine. Edasi kirjeldati nimetatud päringute täitmise protsessi Net Groupis ja probleemi täpsustamiseks viidi läbi küsitlus erinevates projektimeeskondades. Küsitlusest selgus, et kasutajapäringute täitmine ei ole sage, kuid võtab olenevalt andmemudeli keerukusest aega.

Probleemi lahendamiseks valmis prototüübina .NET Core raamistikul kirjutatud konsoolirakendus. Rakendus kasutab sisendina XML faili, kus on kirjeldatud päringu aluseks olevad muutujad. Lisaks on esmakordselt rakenduse kasutamisel vajalik kirjeldada konfiguratsiooni fail, mille alusel pannakse kokku andmebaasipäringud isikuandmete pärimiseks. Isikuandmete väljavõte puhul kasutatakse DOTX malli, mis andmetega täidetakse ja kasutajale DOCX formaadis väljastatakse.

Loodud prototüübi testimine ühes OÜ Net Groupi poolt arendatavas infosüsteemis sai positiivse tagasiside nii tiimiliikmetelt kui ka kliendilt. Võib öelda, et diplomitöö täitis oma eesmärgi ja loodud prototüüp lahendab püstitatud probleemi. Tulevikus on plaanis prototüübi põhjal arendada veebirakendus, et kasutajapäringute täitmise moodul oleks

lihtsamini hallatav ja integreeritav. Lisaks oleks võimalik luua prototüübi baasloogika NuGet teek, kui mõni projekt avaldab soovi ainult funktsionaalse koodi osas.

Kasutatud kirjandus

[1] Andmekaitse Inspektsioon, *Andmekaitse Reform,* 2019. [Online]. Loetud aadressil: https://www.aki.ee/et/teavitus-uudised/andmekaitse-reform Kasutatud: 12.03.2021.

[2] V. Ayala-Rivera ja L. Pasquale, „"The Grace Period Has Ended": An Approach to Operationalize GDPR Requirements," *2018 IEEE 26th International Requirements Engineering Conference*, Banff, Canada, 2018, lk 1. DOI: 10.1109/RE.2018.00023 Kasutatud: 10.03.2021.

[3] N. Gruschka, V. Mavroeidis, K. Vishi ja M. Jensen, „Privacy Issues and Data Protection in Big Data: A Case Study Analysis under GDPR," *2018 IEEE International Conference on Big Data*, Seattle, USA, 2018, lk 2. DOI: 10.1109/BigData.2018.8622621 Kasutatud 10.03.2021.

[4] Your Europe, *Isikuandmete kaitse üldmääruse kohane andmekaitse,* 2020. [Online]. Loetud aadressil: https://europa.eu/youreurope/business/dealing-with-customers/data-protection/data-protection-gdpr/index_et.htm Kasutatud: 05.03.2021.

[5] Andmekaitse Inspektsioon, *Teave isikuandmete kasutamise kohta*, 2020. [Online]. Loetud aadressil: https://www.aki.ee/et/eraelu-kaitse/teave-isikuandmete-kasutamise-kohta Kasutatud: 12.03.2021.

[6] Andmekaitse Inspektsioon, *Isikuandmed ja töötlemine*, 2019. [Online]. Loetud aadressil: https://www.aki.ee/et/eraelu-kaitse/isikuandmed-ja-tootlemine Kasutatud: 12.03.2021.

[7] Andmekaitse Inspektsioon, *Isikuandmete kaitse reeglid ei kehti*, 2019. [Online]. Loetud aadressil: https://www.aki.ee/et/eraelu-kaitse/isikuandmed-ja-tootlemine/isikuandmete-kaitse-reeglid-ei-kehti Kasutatud: 12.03.2021.

[8] Andmekaitse Inspektsioon, *Andmetega tutvumine*, 2019. [Online]. Loetud aadressil: https://www.aki.ee/et/eraelu-kaitse/andmetega-tutvumine Kasutatud: 12.03.2021.

[9] Andmekaitse Inspektsioon, *Automatiseeritud otsused ja profiilianalüüs*, 2019. [Online]. Loetud aadressil: https://www.aki.ee/et/eraelu-kaitse/automatiseeritud-otsused-ja-profiilianaluus Kasutatud: 12.03.2021.

[10] Andmekaitse Inspektsioon, *Andmete parandamine*, 2019. [Online]. Loetud aadressil: https://www.aki.ee/et/eraelu-kaitse/andmete-parandamine Kasutatud: 12.03.2021.

[11] Andmekaitse Inspektsioon, *Andmete töötlemise piiramine*, 2019. [Online]. Loetud aadressil: https://www.aki.ee/et/eraelu-kaitse/andmete-tootlemise-piiramine Kasutatud: 12.03.2021.

[12] Andmekaitse Inspektsioon, *Andmete kustutamine*, 2019. [Online]. Loetud aadressil: https://www.aki.ee/et/eraelu-kaitse/andmete-kustutamine Kasutatud: 12.03.2021.

[13] Andmekaitse Inspektsioon, *Andmete ülekandmine*, 2019. [Online]. Loetud aadressil: https://www.aki.ee/et/eraelu-kaitse/andmete-ulekandmine Kasutatud: 12.03.2021.

[14] *Isikuandmete kaitse üldmäärus,* Euroopa Parlament, 2016. [Online]. Loetud aadressil: https://eur-lex.europa.eu/legal-content/ET/TXT/HTML/?uri=CELEX:32016R0679&from=ET Kasutatud: 13.03.2021.

[15] *Isikuandmete töötleja üldjuhend*, Andmekaitse Inspektsioon, 2019. [Online]. Loetud aadressil: https://www.aki.ee/sites/default/files/dokumendid/isikuandmete_tootleja_uldjuhend.pdf Kasutatud: 14.03.2021.

[16] Euroopa Komisjon, *Kas ma võin paluda ettevõttel minu isikuandmed kustutada?*, *aasta puudu*. [Online]. Loetud aadressil: https://ec.europa.eu/info/law/law-topic/data-protection/reform/rights-citizens/my-rights/can-i-ask-company-delete-my-personal-data_et Kasutatud: 17.03.2021.

[17] B. Tolson, *Is the Anonymization of Personal Data the Same as Data Erasure?*, 2019. [Online]. Loetud aadressil: https://www.ironmountain.com/blogs/2019/is-the-anonymization-of-personal-data-the-same-as-data-erasure Kasutatud: 18.03.2021.

[18] D. Walker, *How to reclaim your data from Google, Facebook, Microsoft, Apple under GDPR*, 2021. [Online]. Loetud aadressil: https://www.itpro.co.uk/general-data-protection-regulation-gdpr/31330/how-to-reclaim-your-data-from-google-facebook Kasutatud: 16.04.2021.

[19] Microsoft, *What is .NET?*, 2021. [Online]. Loetud aadressil: https://dotnet.microsoft.com/learn/dotnet/what-is-dotnet Kasutatud: 12.04.2021.

[20] G. H. Siciliano, *.NET Core vs .NET Framework*, 2018. [Online]. Loetud aadressil: https://medium.com/wolox/net-core-vs-net-framework-a694f1fbdb26 Kasutatud: 12.04.2021.

[21] Microsoft, *.NET desktop apps*, 2021. [Online]. Loetud aadressil: https://dotnet.microsoft.com/apps/desktop. Kasutatud 12.04.2021.

Lisa 1 – Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks[1]

Mina, Ants Kristjan Rooma

1. Annan Tallinna Tehnikaülikoolile tasuta loa (lihtlitsentsi) enda loodud teose „Privaatsust arvestavate kasutaja päringute mooduli arendus OÜ Net Group näitel“, mille juhendaja on Kristiina Hakk
   1.1. reprodutseerimiseks lõputöö säilitamise ja elektroonse avaldamise eesmärgil, sh Tallinna Tehnikaülikooli raamatukogu digikogusse lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;
   1.2. üldsusele kättesaadavaks tegemiseks Tallinna Tehnikaülikooli veebikeskkonna kaudu, sealhulgas Tallinna Tehnikaülikooli raamatukogu digikogu kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.
2. Olen teadlik, et käesoleva lihtlitsentsi punktis 1 nimetatud õigused jäävad alles ka autorile.
3. Kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest ning muudest õigusaktidest tulenevaid õigusi.

[17.05.2021]

1 Lihtlitsents ei kehti juurdepääsupiirangu kehtivuse ajal vastavalt üliõpilase taotlusele lõputööle juurdepääsupiirangu kehtestamiseks, mis on allkirjastatud teaduskonna dekaani poolt, välja arvatud ülikooli õigus lõputööd reprodutseerida üksnes säilitamise eesmärgil. Kui lõputöö on loonud kaks või enam isikut oma ühise loomingulise tegevusega ning lõputöö kaas- või ühisautor(id) ei ole andnud lõputööd kaitsvale üliõpilasele kindlaksmääratud tähtajaks nõusolekut lõputöö reprodutseerimiseks ja avalikustamiseks vastavalt lihtlitsentsi punktidele 1.1. ja 1.2, siis lihtlitsents nimetatud tähtaja jooksul ei kehti.

Lisa 2 – Küsimustik hetkeolukorra täpsustamiseks

1. Milliseid isikuandmete päringuid on kasutajad esitanud?
    - Isikuandmete väljavõte
    - Isikuandmete eemaldamine/anonümiseerimine
    - Kasutajad ei ole selliseid päringuid esitanud
2. Kui palju aega kulus isikuandmete väljavõtte loomiseks?
3. Kui palju aega kuluks iga järgneva isikuandmete väljavõtte loomiseks?
4. Kui palju aega kulus isikuandmete eemaldamiseks/anonümiseerimiseks?
5. Kui palju aega kuluks igale järgnevale isikuandmete eemaldamisele/anonümiseerimisele?
6. Milliseid andmebaase te oma projektides kasutate?