

TALLINNA TEHNIKAÜLIKOOL
Infotehnoloogia teaduskond

Kalle Johannes Reinmets
185454IAAB

Hoonete geomeetria automaatne uuendamine OpenStreetMapi näitel

Bakalaureusetöö

Juhendaja: Siim Vene

MSc

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: Kalle Johannes Reinmets

17.05.2021

Annotatsioon

Töö põhiprobleem lähtub OpenStreetMap kaardil 2008. aastal Tallinnasse ning Tallinna lähivaldadesse lisatud hoonete impordist. 2021. aastaks on hooned kaasaegsete arusaamade järgi madala kvaliteediga ja vananenud ning väärivad asendamist uute kvaliteetsemate andmetega Maa-ameti andmekogust. Töö põhieesmärk oli luua lahendus, mis suudaks peale vajalike kooskõlastuste saamist vähemalt osad vanad hoonekujud tuvastada ning automaatselt uuendada. Kooskõlastuste saamine on pikk välistest teguritest sõltuv protsess ning see ei kuulunud töö skoopi.

Algselt töö kõrvaleesmärk, kuid hiljem kõige mahukam osa, oli 2008. aasta andmete päritolu ja juriidilise puhtuse analüüs. Peale pikki kirjavahetusi paljude osapooltega jõuti järelduseni, et vana import põhines suure tõenäosusega Maa-ameti 1996–2007 põhikaardil, mille kasutamine oli siis ja on ka praegugi igati lubatud ning puudub põhjus karta tulevikus autoriõiguste rikkumiste süüdistuste pärast.

Lõputöö on kirjutatud eesti keeles ning sisaldab teksti 35 leheküljel, 3 peatükki, 5 joonist, 1 tabelit.

Abstract

Automatic Updating of Building Geometry on Basis of OpenStreetMap

Main problem this thesis originates from 2008 building import added to OpenStreetMap (OSM) online map in and near Tallinn, Estonia. Imported buildings' shapes are considered to have very low quality by modern standards and are long overdue to be replaced by better quality data from Estonian Land Board. Main purpose of this thesis was to create technical solution that would be capable to replace old building geometry with more modern one while preserving editing history of old data. Actual replacement at public instance of OSM server was not within scope of thesis due to long and complicated process needed for automated editing approval.

Initially mere side-objective, but later major part of thesis was to identify source and legitimacy of data used for import in 2008 as failure to do so may have left buildings subject to removal even after they were updated. Investigation started from contacting Garmin because of their signature-tags featured in initial import, but later focused on trying to identify user who made the import. Finally, after comparing similarity of building shapes, it was concluded that data had most likely originated from Estonian Land Board's 1996–2007 Basic map, which was source allowed for OSM mapping back then and is still today.

In order to update building footprints, Python 3 script was created. Upload script was tested multiple times in local OSM server set up on Ubuntu 20.04 virtual machine and it managed to safely update 27 thousand buildings (of 45 thousand remaining from original import) taking 10 hours per try.

Script is available at Github [1].

The thesis is in Estonian and contains 35 pages of text, 3 chapters, 5 figures, 1 table.

Lühendite ja mõistete sõnastik

Digiteerimine	Kaardielementide ülesmärkimine teise kaardi või aerofoto järgi joonistades.
DWG	<i>Data Working Group</i> (Andmete töögrupp) – OSMi moderaatori-tiim, mis tegeleb vandalismi, plagiaadi jms.
ETAK	Eesti topograafia andmekogu – avaandmed, mille põhjal koostab Maa-amet Eesti kaarti.
Muudatuskogum	Kaardistaja poolt tervikuna OpenStreetMap serverisse laaditud mingi piirkonna muudatused.
ODbL	<i>Open Database License</i> (avatud andmebaasi litsents) – Litsents, mille alusel OSM andmeid levitatakse.
OSM	<i>OpenStreetMap</i> – Avalik kaardiandmete andmebaas.
OSMF	<i>OpenStreetMap Foundation</i> – OSMi esindav juriidiline isik.
Ruumiandmed	Andmed, mis otseselt või kaudselt osutavad konkreetsele asukohale või geograafilisele alale [2].

Sisukord

1 Sissejuhatus	10
2 Taust	12
2.1 Geinfosüsteemid	12
2.2 Vabatahtlik geograafiline teave	12
2.3 OpenStreetMap	14
2.3.1 OpenStreetMap kui projekt	14
2.3.2 OpenStreetMap kui kaart	15
2.3.3 OpenStreetMap kui infosüsteem	16
2.4 OpenStreetMap ja Eesti	18
3 Analüüs	20
3.1 Üldise lähenemise valimine	20
3.2 Vananenud hoonekujude leidmine OSM andmebaasist	22
3.3 Impordi lähteandmete tausta uurimine	24
3.4 Arenduskeskkonna valimine	30
3.4.1 Arenduskeskkonna virtuaalmasina põhinäitajad	30
4 Praktiline osa	32
4.1 Keskkonna seadistamine	32
4.2 Hoonete uuendamise skript	34

4.2.1 Algoritmi selgitused	34
4.2.2 Skripti tulemused.....	35
4.3 Edasiarendamise suunad tulevikus	40
4.3.1 Massmuudatuse ettevalmistamine	40
4.3.2 Edasiarendused kaugemas tulevikus	42
5 Kokkuvõte	43
Kasutatud kirjandus	45
Lisa 1 – Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks	60
Lisa 2 – Lõpliku Overpassi päringu näide.....	61
Lisa 3 – Kuvatõmmised kolmest erinevast muudatuskogumite jaotusest	62

Jooniste loetelu

Joonis 1. Kuvatõmmis kasutaja „Verbatiumi“ hooneimpordi muudatuskogumite tabelist.	25
Joonis 2. Maa-ameti 1996–2007 põhikaardi (vasakul) ja OpenStreetMapi andmete (paremal) võrdlus.....	28
Joonis 3. Näide vigaselt lisatud hoonekujust.....	36
Joonis 4. Skripti käigus tagastatud tulemuste jaotus	39
Joonis 5. Näiteid skripti töö tulemusest.....	40

Tabelite loetelu

Tabel 1. Skripti käigus tagastatud tulemuste jaotus	38
---	----

1 Sissejuhatus

Käesolev töö läheneb Eesti-kesksele probleemile, kus OpenStreetMap (OSM) kaardil on enamik Tallinna ja Tallinna lähivaldade hoonete kujudest imporditud 2008. aastal ebakvaliteetsest tundmatu päritoluga andmekogust. Hoolimata sellest, et enamike atribuutandmete (valdavalt aadressiinfo) uuendamine on tänaseks automatiseeritud [3], püsib probleem geomeetriaga, mis on kaasaegsete arusaamade kohaselt kehva kvaliteediga ning vajab kaasajastamist. Tulenevalt Eesti OSM kaasautorite tagasihoidlikust aktiivsusest [4], ei ole mõistlik kaasajastamist ilma automatiseerimiseta läbi viia. Seda enam, et erinevalt 13 aasta tagusest ajast on tänapäeval avaandmetena saadaval Maa-ameti Eesti topograafilise andmekogu (ETAK) hoonekujud.

Hoolimata probleemi vanusest on teema uudne, sest OSMi Eesti alal ei ole varem automatiseeritult geomeetriat veel uuendatud [5]. Skriptid on Maa-ameti avaandmeid seni kasutatud vaid uute objektide lisamiseks (näiteks metsamaade ning hoonete lisamine vastavalt 2007. ja 2013. aastal) või olemasolevate elementide atribuutandmete uuendamiseks (2019. aadressiandmete uuendamine).

Oluliseks teguriks, miks seni pole keegi tahtnud probleemiga tegeleda, on 2008. a impordi teadmata allikas, mistõttu kahtlustati võimalikku autoriõiguste rikkumist [6]. Töö analüütiline osa keskendub suuresti andmete allika otsimisele erinevate kaardistustevõtete vahendusel. Algne hüpotees oletas, et andmed võisid pärineda Garmin kaardilt, kuid Garmin esindaja lükkas hoonekujusid nähes väite ümber.

Praktilises osas valmis kaks skripti. Neist põhiline, mis oli ka töö eesmärgiks, on Pythoni skript, mis omab võimekust vajadusel uuendada piirkonna hoonekujusid võttes aluseks Maa-ameti Eesti topograafilise andmekogu (ETAK) avalikult kättesaadavad SHP-failid. Andmete uuendamine ei ole iseenesest töö eesmärk, sest andmete korrektne massmuutmine hõlmab pikka protsessi, mille heakskiit oleneb välistest faktoritest. Teine on suhteliselt lihtne Bash skript, mis on mõeldud Ubuntu virtuaalmasinas testimiskeskonna seadistamiseks.

Uuritava probleemi lahendus on piisavalt universaalne, et sarnast lahendust saaks peale mõningate muudatuste sisseviimist kasutada tulevikus üle maailma mitmesuguste andmekomplektide sünkroniseerimiseks OSM andmekoguga. Minimaalselt peaks lahendus olema rakendatav Eesti 2013. aasta hooneimpordi värskendamiseks. Peale skripti valmimist tundsid lahenduse vastu huvi Poola kasutajad, kes on hiljuti saanud juurdepääsu nende kohalikule hoonete andmekogule ning otsivad lahendust vanade hoonekujude asendamiseks [7].

2 Taust

Teoreetiline osa tutvustab geoinfosüsteeme ning vabatahtlikku geograafilist teavet (VGI) ning seejärel annab ülevaate OpenStreetMapist (OSM) käesoleva probleemi kontekstis.

2.1 Geoinfosüsteemid

Geoinfosüsteemid (GIS) on ülejäänud käesolevas töös käsitlevate valdkondadega võrreldes suhteliselt vana ja lai valdkond, mille juured ulatuvad 1970-aastate ülikoolide arvutuskusteni [8]. Geoinfosüsteemi saab määratleda erinevalt, kuid käesolevas töös lähuti GISist kui automatiseeritud süsteemi ruumiandmetega seotud toimingute tegemiseks [9]. Võimalikud toimingud saavad hõlmata tervet andmete elukaart ning nende hulka kuuluvad näiteks andmete kogumine, töötlemine, analüüs ning esitamine [9].

GIS andmed saab üldiselt jagada kaheks: vektorandmed ning rasterandmed. Vektorandmestik koosneb koordinaatidega määratud tippudest ja neid ühendavatest lõikudest. Vektorkujul sobivad esitamiseks geomeetrilised kujundid, millel saab tuua esile selge piiri. Rasterandmed koosnevad pikslitest ning neid kasutatakse olukordades, kus soovitakse kirjeldada ühtset pinda katvaid andmeid ning nende jagamine diskreetseteks objektideks ei ole mõistlik (näiteks kui soovitakse esitada eri värvi murulaike). [8]

Ruumiandmed on Eesti „Ruumiandmete seaduse“ tähenduses „andmed, mis otseselt või kaudselt osutavad konkreetsele asukohale või geograafilisele alale, sealhulgas /../ andmed, mis kirjeldavad ruumiobjektide asukohta, omadusi ja kuju geograafilises ruumis“ [2]. Eelnimetatud seadus reguleerib valdavalt avaliku sektori toodetud ruumiandmeid, muuhulgas Maa-ameti hallatavat Eesti topograafia andmekogu (ETAK). ETAKi andmeid kasutatakse käesolevas töös lähteandmetena, mida kanti üle OSM kaardile.

2.2 Vabatahtlik geograafiline teave

Koos Veeb 2.0 ehk interaktiivse veebi esile kerkimisega 10–15 aastat tagasi hakati rääkima kasutajapoolsest sisuloomest [10] [11]. Veeb 2.0-l puudub täpne definitsioon, kuid

üks konkreetsemaid selgitusi väidab, et varasema ühepoolse suhtluse asemel (veebilehe omanikult kasutajatele) on nüüd sisuloomesse kaasatud ka lõppkasutaja (kasutajalt kasutajale) [12]. Kõige tüüpilisema Veeb 2.0 näitena võib tuua mistahes sotsiaalmeedia platvormi. Tänu samal perioodil tarbijaelektronikaseadmetesse jõudnud GPS-toele kujunes uus trend ka geoinfosüsteemide valdkonnas – vabatahtlik geograafiline teave (*Volunteered Geographic Information, VGI*) [13].

VGI kui nähtuse formuleeris 2007. aastal uurija Goodchild [14], kes kirjeldas VGId kui kasutajate loodud sisu alaliiki [15]. Kuni käesoleva sajandi algusaastateni oli geoinformatsiooni tootmine koondunud enamasti valitsusasutuste (tavaliselt kohaliku omavalitsuse) monopoliks ning andmekogude kasutusõiguse soetamine oli jõukohane tavaliselt ainult suurematele ettevõtetele [9]. VGI puhul käib andmete kogumine tavaliselt detsentraliseeritult, mis tähendab, et info pärineb paljudest erinevatest allikatest, tihtilugu on esmaseks allikaks erinevad üksikisikud [16]. VGI olulise erisusena tavapärasest GISist ei ole enamik kaasautoreid erialaspetsialistid, vaid vabatahtlikud internetist [16].

VGI raames lisatav info võib varieeruda väga lihtsast tegevusest (näiteks sotsiaalmeedias asukohainfo jagamisest) keerulisamate tegevusteni, näiteks avaliku kaardi põhjaliku muutmiseni [17]. Mõned autorid on VGId võrrelnud kasutajate ühisteadvusega (*collective intelligence*), kus isikute grupi koostööna loodud süsteem on targem (detailsem) kui kõige targem indiviid [18]. Seeläbi on oletatud, et tänu paljude üksikisikute erinevatele motiividele ja vaadetele on võimalik luua palju mitmekesisem infosüsteem, kui ainult ekspertide (kartograafide) rühma kasutades, samal ajal olulist täpsust kaotamata [18].

OpenStreetMap on hetkel küll suurim ja tuntuim VGI [19], aga varem on eksisteerinud ning on ka praegugi olemas alternatiivseid platvorme, näiteks Platial [16], Tagzania [19] ja Frapper [20]. Esmakordselt kasutati VGId humanitaarkatastroofide lahendamiseks 2005. aastal (orkaan Katarina ja Google Maps) [21].

VGI-l leidub ka vastuoluline varjukülg. Mõned autorid on vabatahtlike kaasamist käsitlenud nende ärakasutamisena. Mõned ettevõtted on motiveeritud kasutama avaandmeid, sest nende kasutamine on odavam – ettevõttel pole vaja maksta ei andmete litsentsitasusid ega palka majasisestele kartograafidele [16]. Teisest küljest on VGI raames loodud andmed kasulikud laiemale kogukonnale kui ainult ühele organisatsioonile [16]. Eelneva argumentatsiooni puhul viidati algallikana Nielsen 2006. aasta uuringule, mis tuvastas, et

keskmises sotsiaalvõrgustikus kehtis „90-9-1-reegel“ – 90% kasutajast ei sekku sisuloomesse, 9% on sekkuvad harva ning vaid 1% kasutajatest on aktiivsed sisuloojad [22].

2.3 OpenStreetMap

Käesolev alapeatükk tutvustab lugejale OpenStreetMapi, liikudes järgemööda üldisest projekti tutvustusest tehnilisemaks.

2.3.1 OpenStreetMap kui projekt

OpenStreetMapi (OSM) nime järgi oletatakse tihti, et tegemist on kaardiga, teinekord räägitakse OSMist ka kui VGId koondavast andmebaasist. Tegelikuses on mõiste tähendus veidi laiemgi – OSM on projekt, mille eesmärk on luua avaandmetest koosnev geoinfo kogumik [23]. Mõningaste mõõndustega saab projekti lühidalt kokku võtta kui Vikipeedia ristamine Google Mapsiga [24] [25]. Projekt sai Vikipeediast inspireerituna alguse 2004. aastal Londonist [13], ajendiks tollal tavainimesele kättesaamatud kvaliteetsed ruumiantmed [25]. Kuigi OSMi juured pärinevad Suurbritanniast, leidis liikumine suurimat kõlapinda Saksamaal, Prantsusmaal ja Austrias [19] [26].

Projekti huvisid esindab juriidilise kehana Suurbritannias registreeritud mittetulundusühing OpenStreetMap Foundation (OSMF, OpenStreetMapi sihtasutus), mis loodi 2006. aastal eesmärgiga hallata projektiga seotud taristut ning annetusi [23]. OSMF mõjutab OSMi läbi erinevate töögruppide, mis on mõeldud kindlate ülesannete täitmiseks. Töögrupe on kokku 8, käesoleva töö seisukohast on olulisim andmete töögrupp (*Data Working Group*, DWG), millest tuleb juttu allpool. Samuti väärivad esiletõstmist operatsioonide töögrupp (OWG), mis vastutab serverite (sh OSM-API) hoolduse ja arenduse eest, ning litsentsimise töögrupp (*Licensing Working Group*, LWG), mis tegeleb juriidiliste küsimustega, olulisimana 2012. aastal toimunud litsentsivahetusega. [27]

OSMi eripära paljudest teistest VGI platvormidest on ODbL litsentsipoliitika, mis lubab erinevalt nt Google Mapsist kasutada lähteandmestikku ka kolmandatel osapooltel [19]. OSMi andmeid levitati algul CC-BY-SA litsentsi alusel, kuid 2012. aasta oktoobris mindi üle OSMi tarbeks loodud avatud andmebaasi litsentsile (ODbL) [23] [28]. Vana litsentsi suurimateks puudusteks olid samadel tingimustel jagamise (SA) ning algallikale viitamise (BY) nõude hägusad piirid [23]. ODbL-le ülemineku käigus oli toonastel kasutajatel

võimalik valida, kas nad soovivad uue litsentsi tingimustega nõustuda. Kaardimuudatused, mille autorid uue litsentsiga ei nõustunud, eemaldati juulis 2012 [29].

OSMi kasvus on ajalooliselt suurimad hüpped leidnud aset aastatel 2007–2008 ning 2015–2016 [30]. Viimaste aastate trendina kaarti hakanud täiendama mitmed suurfirmad [31]. Tänapäevaks on OSM kasvanud suurimaks vabalt kättesaadavaks geoinfo platvormiks [32] ning OSMi kantud geoinfot kasutavad paljud erinevad veebiteenused üle maailma [33]. Kõige tüüpilisem näide on navigaatoriäpp mobiiltelefonis, aga samast andmekogust on loodud kümneid temaatilisi kaarte kitsamatele huvigruppidele, näiteks kaardid jalgrattaristust või ühistranspordist [34]. Leidub ka ambitsioonikamaid projekte, kus OpenStreetMapi taristu vahenditega on loodud uusi iseseisvaid andmekogusid [35] [36].

2.3.2 OpenStreetMap kui kaart

OSM kaardile jõuab uus lisatav teave erinevaid teid pidi. Andmete lisamine kaardile tugineb reeglina kasutajate *ad-hoc* kaardistamistegevusel, mis pole ei automatiseeritud ega põhjalikult koordineeritud. Traditsiooniliselt toimus kaardistamine GPS-jälgede abil [13], kuid kaasajal on levinuimaks meetodiks digiteerimine, ehk objektide ülesjoonistamine litsentsitud aerofotode või kaartide põhjal [37]. Vabatahtlike poolt andmete lisamiseks tehtavat rohket käsitööd peetakse samaaegselt projekti tugevuseks ja nõrkuseks [38]. Kui tugevuseks on kaardi kattuvus arengumaades, kuhu kommertsteenused ei laiene, siis nõrkusena võib juba põhjalikult kaardistatud piirkondade ajakohasena hoidmine käia inimestele üle jõu. Kolmandat kaardi täiendamise meetodit ehk importimist tutvustatakse allpool.

Järelevalvega kaardiandmete õigsuse üle ning väiksemate tülide lahendamise tegeleb tavaliselt kohalik kogukond [13]. Suuremahuliste intsidentide lahendamise tegeleb üks kuuest töögrupist, DWG (*Data Working Group*, andmete töögrupp). Sisuliselt on tegemist kaardiprojekti moderaatoritiimiga, kellega tuleks ühendust võtta tõsisete reeglite rikkumiste või vandalismi korral [39]. Lisaks vandalismiga võitlemisele on DWG-l võimekus vajadusel tagasiulatuvalt kustutada (*redact*) andmestikust autoriõiguseid rikkuvat sisu [40]. Käesoleva töö kontekstis on oluline just viimane vastutusala. DWG ja LWG peamise erinevusena andmete töögrupp tegeleb andmetega, mis lisatakse OSMi, litsentsimise töögrupp OSMi andmete väljavõtetega [27].

Eelmainitud kolmas üldine suund kaardi muutmiseks on andmete importimine mõnest vabalt kättesaadavast andmekogust. Mõne üksiku kasutaja töö tulemusel saavutatakse tulemus, mis käsitsi kaardistamisel võib võtta tervel kogukonnal nädalaid või isegi kuid [41]. Importimise eeliseks on ühtlane kvaliteet, kiirus ja mugavus, aga puudustena tuleb tagada ühilduvus kaardil juba olemasolevate andmetega ning hilisem andmete ajakohasena hoidmine on keerulisem kui nende lisamine [41]. Kuigi impordid tagavad kaardi kvaliteedile miinimumtaseme, on nii Eestis kui ka mujal maailmas esinenud juhtumeid, kus uutele kasutajatele paistab kaart valmis olevat ning nad ei soovi midagi juurde lisada [17] [6]. Käesolev töö keskendub varem imporditud teabe uuendamisele.

Importid liigituvad OSMi ökosüsteemis automatiseeritud muudatuste alla ning nende läbiviimine on rangelt reguleeritud vastavasisulise käitumisjuhenditega [17]. Eeskirjade kohaselt tuleb esiteks saada kohalike kasutajate nõusolek, mis Eestis tähendab kohalikku meililisti talk-ee kirjutamist. Järgnevad lisatava teabe omanikult nõusoleku küsimine (käesoleva probleemi puhul pole samm vajalik, sest Maa-ameti ruumiandmeid on varem juba impordides kasutatud), tegevusplaani avaldamine vikis ning kooskõlastamine rahvusvahelises importide meililistis [42].

Kui eelnevalt kirjeldati importimise kooskõlastamist, siis automaatmuudatuste läbiviimine on eelmisega sarnane, kuid on kohati leebem ning põhjalikumalt lahti kirjutatud. Näiteks ei ole teise juhendi väitel vaja saata teadet mitmesse meililisti, vaid piisab ainult ühest (kasutaja valikult üldine, kohalik või importide meililist). Samuti selgitatakse detailsemalt, mida peab vikis avaldatav impordi dokumentatsioon sisaldama: olulisimana nõutakse muudatuste läbiviija kontaktteavet ning skripti tööpõhimõtte tutvustamist [43]. Segaduse vältimiseks olgu mainitud, et OSMis on olemas sarnase nime, kuid teistsuguse sisuga juhend nimega Organiseeritud Muudatuste Eeskiri (*Organised Editing Guidelines*), mis reguleerib organisatsioonide töötajate tegevust OSMis, aga mitte käesoleva töös käsitletavat automatiseeritud tegevust [44].

2.3.3 OpenStreetMap kui infosüsteem

Järgnevalt liigutakse tehnilisemaks ning vaadeldakse OpenStreetMapi kui konkreetseid rakendusi. Kaart, mida kasutaja OSMi veebilehel näeb, ning samuti ka andmebaas ise on vaid jäämäe tipp nende taha jäävast taristust omaette teenuste, protsesside ning rakendustega. OSMi peamine dokumenteerimise viis on nende viki [45].

OSM andmemudel hoiustab kolme tüüpi elemente: sõlmi (*node*), mis on matemaatikas võrreldav punktiga, jooni (*way*), mille lähim matemaatiline vaste on murdjoon, ning relatsioone (*relation*), mida kasutatakse mitme joone ja/või sõlme sidumisel üheks loogiliseks elemendiks [46]. Kõik elemendid saavad omada atribuutandmete silte (*tag*), kuid geograafilised koordinaadid talletatakse ainult sõlmedele [10]. Sildid on võti-väärtus paarid (*key, value*), mida kirjeldatakse kujul *key=value*. Kõrgema taseme elemendid (jooned ja relatsioonid) saavad oma asukohainfo läbi viidete madalama taseme elementidele [17]. Siit järeldub üks analüüsi lähtepunkte, et leidmaks hooneid, mille kuju ei ole aastate jooksul muudetud, tuleb vaadelda just sõlmede muutmise ajalugu.

Kõige levinum kaardi muutmise viis on teha seda vastava rakendusega, mida võib üldjoontes jagada kolme kategooriasse: veebi-, mobiili- ja töölauarakendused. 2020. aastal olid üldarvestuses ning iga kategooria populaarseimad vastavalt iD (80%), Maps.Me (12%) ja Josm (7%) [47]. Protsent näitab, kui suur osa kõigist kasutajatest nimetatud rakendust kasutasid. Lisaks neile kolmele väärub mainimist populaarsuselt järgmine, ommoodi mänguline telefonirakendus StreetComplete (4%), sest viimane on Eesti kasutajate seas populaarsem kui Maps.me [4] [48] [49].

Kõik muudatused, mida kasutajad teevad, grupeeritakse alates 2009. aastast, peale API versioon 0.6 kasutuselevõttu, muudatuskogumitesse (*changeset*) [50]. Muudatuskogum on komplekt muudatusi, mida üks kasutaja tegi ühe kaardistussessiooni käigus, ning nad laaditakse OpenStreetMap serverisse ühtse tervikuna [51]. Muudatuskogumite kasutuselevõtu eesmärk oli muuta muudatuste grupeerimine ning seeläbi rämpssisuga võitlemine lihtsamaks – kui varem pidi DWG pöörama tagasi terve piirkonna muudatused, siis uues API versioonis said nad keskenduda ainult konkreetse kasutaja tegevusele [50]. 2009. aasta versioonivahetuse käigus jagati kõik varasemad muudatused tagantjärele automaatselt muudatuskogumitesse [50].

Andmete uuendamist OSMi andmemudelis on võrreldud pinu-tüüpi andmestruktuuriga, kus muudetud element salvestatakse uue versioonina [10]. Andmebaasi struktuur on disainitud selliselt, et midagi püsivalt andmebaasist ära ei kustutata [52]. Kõik, mis kasutajad kaardiredaktoris kustutavad, jäävad tegelikult baasi alles, ning teised kasutajad saavad neid spetsiaalsete tööriistadega lugeda ja vajadusel taastadagi. Kuigi DWG-l on võimekus ajaloost lubamatut sisu peita, siis andmebaasi tasemel ei ole isegi neil võimalik elemente täielikult kustutada [53].

Kui vaadata OSMi IT-lahendusi süsteemiadministraatori vaatepunktist, siis võib näha moodulitest koosnevat infosüsteemi [54]. Vabavarale omaselt leidub populaarseimatele moodulitele alternatiivseid lähenemisi ning piisavate oskuste korral suudaks soovija enamikku komponente enda serveriski käivitada. Täisväärtuslik OSMi veebileht tugineb erinevatele moodulitele, mis kõik töötlevad omaenda versiooni igaminutiliselt uuendatavast OSMi andmestikust [13]. Näiteks Mapnik joonistab kaardiaknas kuvatava pildi (kaardi-paanid), otsingukasti taga töötab Nominatimi otsingumootor ning otsingutulemuseni juhatavad sisse ehitatud navigatsioonimoodulid OSRM ja Graphhopper [55]. Käesoleva töö seisukohast on olulised vaid saidi põhifunktsionaalsuse tagav OSM-API ning Overpass API.

OpenStreetMapi põhiline rakendusliides ehk OSM API on mõeldud toorandmete vahetuks andmebaasiga, peamiselt kaardi muutmiseks [56]. API töötab koos põhilise veebisaidiga tarkvarakomplektis, mida kutsutakse Rails-pordiks. Nimi Rails-port tuleneb rakenduse jaoks kasutatud Ruby-on-Rails raamistikust. Ühte Rails rakendusse on kokku koondatud erinevad liidesed, mis on vajalikud toimiva OSM veebilehe põhiliseks funktsionaalsuseks: OSM API, PostgreSQL andmebaasi konfigureerimine ja veebilehe kasutajaliides koos iD kaardiredaktoriga [57]. Kaardi lugemiseks ja ruumipäringute tarbeks kasutatakse andmete lugemisele optimeeritud lahendusi, näiteks Overpass API [55].

Kui tavakasutajale piisab otsitava koha leidmiseks lihtsast veebiotsingust, siis kogenumad kasutajad, kes töötavad suuremate andmemahtudega, saavad lähtuvalt otsitavate andmete olemusest valida kolme lähenemise vahel [58]. Täielik tervet maailma kattev väljavõte nimega, kohalikud väljavõtted ning temaatilised väljavõtted läbi ruumipäringute. Kõige levinum ruumipäringute tegemise viis OSMi ökosüsteemis on eelmainitud Overpass API päringud [10]. Andmete küsimiseks spetsiaalsetest serveritest kasutatakse Overpass QL (*Query language*, päringukeel) või XML-vormingus päringuid [59] [60]. Overpass on ise suhteliselt robustne API, mistõttu kasutavad paljud OSMi kasutajad teenust läbi Overpass-Turbo veebiliidese, mis lubab hõlpsalt uusi päringuid koostada ja testida [10] [61].

2.4 OpenStreetMap ja Eesti

Vanimad siiani Eesti alal paiknevad OSM kaardi elemendid pärinevad juulist 2006 ning kujutavad endast põhimaanteid [62]. Eestis on OSMi formaalseks kohalikuks ühinguks

MTÜ Avatud Maakaardi Selts ning peamine suhtluskanal on ametlik Eesti meililist talk-ee [63]. Vaadeldes Eesti meilisti ja viki lehe ajalugu, või julgelt väita, et Eestis jäi kasutajate rohkuse ja aktiivsuse kõrgaeg aastatesse 2009–2011 ja 2013–2015 [64] [4].

Uute allikate kaasamine OSMi on keeruline ja rangelt reguleeritud valdkond, sest paljude internetis kättesaadavate andmestike kasutamine algsest eesmärgist muul põhjusel võib olla keelatud [65]. Eestis on olukord veidi leebem, sest siin tugineb OSM andmestik suuressti Maa-ameti ortofotodele ja rasterkaartidele, mis on olnud aastaid avaandmetena vabalt kättesaadavad [66] [67]. Eesti topograafia andmekogu (ETAK) vektorandmestik muudeti vabalt kättesaadavaks 2018. aasta suvel [68]. Lugejal võib siinkohal tekkida küsimus, miks ei võiks inimesed lihtsalt Maa-ameti ortofotot vaadata, miks on vaja vahele OSMi vahendajaks? Olenevalt kasutaja eesmärgist võibki ortofoto täita soovitud eesmärgi paremini kui OSM kaart, aga samas võib OpenStreetMapi vaadelda kui globaalset ning suhteliselt standardiseeritud liidest, läbi mille saaks siduda paljusid avalikke andmekogusid üle maailma [69]. Eestlasele võib olla loomulik kasutada Maa-ameti kaarti, aga OSM aitab olukorras, kui välismaine kasutaja ei ole Maa-ameti olemasolustki teadlik.

Käesoleva töö kontekstis võib Harjumaal paiknevate hoonete päritolu jagada laias laastus kolmeks – kasutaja „Verbatiumi“ import septembris 2008, kasutaja „Svimiku“ impordid aastatel 2013 ja 2019–2020 ning kasutajate poolt käsitsi kaardile kantud hooned. Viimase kahe puhul on hooned lisatud reeglina Maa-ameti andmestiku põhjal. „Verbatiumi“ andmete päritolu kohta puudus seni igasugune informatsioon, samuti võib oletada, et kohalikel kasutajatel puudus tahe probleemiga tegeleda, sest 13 aasta eest kaardile kantud andmestik on tihtilugu muutmata kujul tänaseni säilinud. Üldjoontes Eestiga analoogne, kuid palju tuntum probleem OSM andmete kehva kvaliteedi osas on „Verbatiumiga“ samal perioodil USAs toimunud TIGER tänavate import [17].

3 Analüüs

Analüüsi osa tutvustatakse probleemsete kaardiosade automaatse tuvastamise meetodikat, antakse ülevaate probleemsete andmete päritolu uurimisest ning praktilise osa testimiskeskonna seadistamisest.

3.1 Üldise lähenemise valimine

Uuritava probleemi puhul ei olnud suurimaks küsimuseks mitte uute andmete kopeerimine, vaid mida teha vanade hoonekujudega. Varsti peale käesoleva tööga alustamist selgus sõnumivahetusest kasutajaga „SviMik“, et tagamaks loodavale lahendusele mingitki praktilist väärtust, tuleb uurida „Verbatiumi“ hooneimpordi lähteandmete juriidilist puhust, mille päritolu oli kuni käesoleva ajani tundmatu. Võimalikud lähenemised saab suuressti jagada kaheks: kas vanad andmed kustutatakse alt ära ning uus andmestik lisatakse tühjale kohale kui tavaline hooneimport või lisatakse uued hoonekujud vanade uue versioonina. Mõlemal variandil on kaks alavarianti, mis puudutavad kustutamise meetodit.

Esiteks variant, kus vanad andmed eemaldatakse enne uute hoonete lisamist. Lähenemise eelisena alustatakse „puhtalt lehelt“, ehk kaotatakse viited varasematele versioonidele ning nende võimalikele intellektuaalomandi rikkumistele. Puudub vajadus märkimisväärseks lisapingutuseks, et tagada ühilduvust olemasolevate andmetega. Samuti on OSMi kohalikul kui ka rahvusvahelisel kogukonnal varasem kogemus sarnaste olukordadega, vastavalt kasutaja „SviMik“ impordid ning 2012. litsentsivahetus [70] [29].

Suurim puudus on sama, mis eelis, ehk kaotatakse varasem hoone muutmiste ajalugu ning vähe kogunud kasutajatel ei ole võimalik tutvuda hoonete kujunemislooga. Samuti tähendab igasugune masskustutamine, et muudatust ei ole võimalik teha hetkeliselt, vaid muutmine võib võtta tunde, kui mitte päevi või nädalaid. 2019. aastal toimus seoses Teneti filmivõttega vestlus, kus arutati ajutiste, aga oluliste muudatuste mõju OSMis. Jaak Laineste argumenti võrguta režiimis töötavatest kaardirakendustest kõlbab kasutada ka siin [71]. Leidub kasutajaid, kes uuendavad oma kaarte OSM andmestiku põhjal väga harva, kuid kui mõne suure ja olulise navigatsiooniteenuse kaardi uuendamine langeb täpselt

vanade hoonete kustutamise ja uute impordi vahele, siis võib Eesti suurima linna puuduvad aadressid mõjutada pika perioodi vältel tuhandeid kasutajaid. [71]

Esimese lähenemise kaks alavarianti lahknevad andmete kustutamise protsessis. 2019. aastal toimunud vestluses toetati ideed hoonete kustutamiseks [70]. Eeldatavasti mõeldi kustutamise all nn „pehmet kustutamist“, mida saavad teha OSMi tavakasutajad [52]. Teisalt oleks juriidiliselt korrektsem alternatiiv oleks paluda DWG-1 „Verbatiumi“ lisatud hooned jäädavalt andmebaasist peita.

Teine, töömahukam lähenemisviis on vanade hoonekujude uuendamine, lisades uued andmed järgmise versioonina. Puuduseks on suurem töö maht, hõlmates keerulisemat asendamise loogikat ning uute ja vanade elementide omavahelist sidumist. Eelised on suuresti samad, mis eelneva puudused – muudatuse tegemise ajal kaart ei tühjene ning säilib muudatuste ajalugu. Kustutamine tugineb DWG moderaatoriõigustele, mis tähendab, et Eesti kasutajad üksi ei suuda vanu versioone vajadusel eemaldada. Samas lükkab lähenemine vanade versioonide eemaldamise vajadust edasi. Käesoleva lähenemise alavariandid ongi, kas peale muutmist tuleks vanad versioonid OSMi ajaloost peita või mitte.

Kui eelmised kaks lahendust keskenduvad automaatsele muudatusele, siis 2019. aasta arutelus pakuti välja ka Osmose tööriista abil kaardistustalgute korraldamist, kus vanad hooned kustutatakse ning asendatakse uutega käsitsi [70]. Käesolevas analüüsis ei peetud täielikult manuaalset tegevust realistlikuks. Kui lähtuda kasutajate aktiivsusest perioodil märts–aprill, kuluks isegi siis, kui kõik 115 kasutajat ei muudaks oma OSMi kasutusmustrit, aga pühendaksid oma tegevuse probleemi lahendamisele, kuluks olukorra lahendamiseks üle pooleteise kuu [4].

Arvutamisel kasutati valemit, kus võeti kõik kasutajate tehtud muudatused (*Map changes*), lahutati lisatud sõlmede arv (*Created nodes*) ning vahe jagati 59-ga (päevade arv veebruaris–märtsis). Saadud vahetulemus 840 näitab veidi liialdatud kujul, kui palju hooned suudaksid kasutajad keskmiselt päevas töödelda. Reaalne tulemus oleks kindlasti väiksem, sest muudatuste hulka on arvestatud ka elementide kustutamised. Lisatud sõlmede hulk oli ainus viidatud veebilehel olnud näitaja, mis käsitles sõlmi, mitte jooni [4]. Kui arvestada, et 40 903 hoonet [70] töödeldakse sagedusega 840 hoonet päevas, saab lihtsa arvutustehtega leida ajakuluks 48 päeva.

3.2 Vananenud hoonekujude leidmine OSM andmebaasist

Leidmaks hooneid, mille kuju on tõenäoliselt aegunud, kasutati Overpass API päringut, tehes seda läbi avaliku Kumi Systems serveri. Algul kasutati tuntuimat Overpassi serverit Overpass-api.de, kuid hiljem valiti Kumi server, sest nende server on mõeldud suuremate päringute jaoks [59]. Päringud koostati ja siluti Overpass-turbo veebikeskkonnas ning kopeeriti skripti sisse alles hiljem. Lõplik skript kasutas API kaudu päringute tegemiseks overPy tarkvarateeki.

Overpy eelisteks teiste sarnaste Pythoni teekide ees oli orienteeritus ühele eesmärgile (overpassi päringute tegemine) ning võimekus võtta sisendpäring vastu peaaegu samasuguse tekstina nagu oleks kasutaja selle Overpass-turbo leheküljel koostanud. Alternatiivsed variandid kas üritasid võtta päringu koostamise vaeva enda kanda (Overpassify) või sisaldasid liigseid käesoleva eesmärgi saavutamiseks ebavajalikke lisamooduleid, näiteks Nominatimi otsingumootori liidest (OSMPythonTools) [59]. Esimene oleks tähendanud päringu kirjutamisel topelttöö tegemist, sest peale soovitud päringu koostamist oleks päring tulnud tõlkida teegile arusaadavale kujule, et teek päringu taas Overpass QL kujule viiks [72].

Overpass QL päringukeelel, mida kasutati päringu kirjutamisel, ilmnisid omapärased puudujäägid. Näiteks selgus, et Overpassiga ei saa otsida etteantud ajatemplist vanemaid objekte, ainult uuemaid, mistõttu tuli alternatiivina leida hulk kõigi elementidega ning lahutada neist uuemate elementide hulk [73]. Päringu kirjutamisel otsustati mitte kasutada kasutajanime-põhist filtrit, vaid 10-aastast vanusepiirangut, et muuta skript tulevikus korduvkasutatavamaks. Näiteks võiks 2023. aastal kasutada skripti järgmise, 2013. a hoonet importi kaasajastamiseks. Valminud päringu näide koos ingliskeelsete kommentaaridega on lisas 3. Kommentaarid on inglise keeles, sest päringut jagati peatükis 3.3 kirjeldatud uuringute käigus teiste, s.h ingliskeelsete osapooltega. Lõplikus Pythoni skriptis asendati väljad nagu kuupäev ja otsinguala automaatselt skriptis määratud muutujatega.

Päringu põhiloogikaks on leida kõik (Harjumaa) hooned, mille geomeetriat (s.t mitte üh-tegi sõlme) ei ole viimase 10 aasta jooksul muudetud (sõlmi pole liigutatud). Overpass QL-i eripärade tõttu tuli probleemile läheneda vastanduvate loogikatehetega. Näiteks eelmainitud päring kujunes lõpuks „leia kõigi (Harjumaa) hoonete hulgast need, mille ükskõik millist sõlme on viimase 10 aasta jooksul muudetud ning eemalda nad ülejäänute

seast“. Vältimaks tulevikus skripti töö käigus soovimatuid kõrvaliste elementide liigutamisi, lisati täiendava filtrina tingimus, et ükski hoone ei tohi jagada ühtegi sõlme mittehoonega. Päringule ei suudetud lisada tingimust, mis välistaks ühiseid sõlmi jagavate hoonete (nt Tallina vanalinna majade) leidmise. Sellise tingimuse kontroll lahendati kohalikus skriptis üleslaadimise funktsiooni sees.

1. märtsi 2021 seisuga leiti Harjumaalt 44302 tingimustele vastavat hoonet. Kokku on Harjumaal ligikaudu 188 000 hoonet [74]. Kuigi kasutatud Kumi Overpassi serveris ei olnud koormuspiirajat, otsustati päringu tegemisele kuluva paari minuti säästmiseks tulemus puhverdada. Puhvri kasuks räägib uuritavate andmete harv muutmine ning tagastatava faili suhteliselt suur maht (60 MB). Puhvril on kasu ka olukorras, kui mõni kasutaja soovib tulevikus kasutada koormuspiirajaga serverit, mis keelab peale teist päringut olenevalt päringu keerukusest 5–6 minutiks uute päringute vastuvõtmise [75],

OverPy teek ei osanud iseseisvalt päringuid puhverdada. Esialgse lahendusena täiendati algul päringu puhverdamiseks OverPy lähtekoodi, asendades kindlatele tingimustele vastava päringu korral veebilehelt andmete lugemise failist lugemisega. Lahendus ei olnud pikas perspektiivis kuigi otstarbekas, sest esiteks toetas selline lahendus korraga vaid ühe tulemuse puhverdamist ning puhvri uuendamine nõudnuks puhvrifaili käsitsi asendamist. Samuti oleks tulnud skripti levitamisel lisada põhiskriptile ka muudetud lähtekoodiga OverPy teek. Lahendusena lisati skripti ja OverPy teegi vahele abstraktsioonikiht, mis haldas erinevate päringute puhverdamist kasutades selleks päringu krüptoräsi.

Skriptide kirjutamise keelteks valiti testimiskeskonna loomise tarbeks Bash, sest tegu on Linuxi käsurea üldtuntud skriptimiskeelega, ning andmetöötluks Python, sest autor omas kõige rohkem varasemalt kogemust just Pythoni keelega.

SHP-faili töötlemiseks kaaluti mitut alternatiivset teeki. Varakult jäeti kõrvale Fiona ja ArcPy. Mõlema kõrvalejätmise põhjuseks oli liigne käesoleva projekti raames ebavajalik funktsionaalsus, Fiona lisapuuduseks oli suhteliselt keeruline paigaldusprotsess [76] ning ArcPy puhul selgus, et tegemist oli üldsegi jaosvaraga, mis nõuab tasulist ArcGis programmi [77]. Esialgse valiku järel eristusid PyShp [78] ja dbfRead [79] teistest teekidest ühest küljest spetsiifilise, kuid samas lihtsa suunitluse poolest. Peale PyShp ja dpfRead teegi proovimist ning langetati lõplik valik PyShp kasuks, sest esiteks dpfRead ei olnud võimeline iseseisvalt geometriainfot avama, mistõttu pyShp kasutamine olnuks niikuinii

vajalik, ning teiseks ilmnes skripti testimisel, et pyShp oli faili avamisel 5–10 korda kiirem kui dbfRead (3–5 vs 25–30 sek), sest dbfRead paistis üritavat faili avamisel tervet faili korraga mällu laadida.

3.3 Impordi lähteandmete tausta uurimine

Suurim osa käesoleva tööle pühendatud ajast kulus 2008. aasta septembris „Verbatiumi“-nimelise kasutaja poolt läbi viidud hooneimpordi lähteandmete leidmisele. Juhul kui andmete päritolu ei oleks saanud kontrollida, oleks alternatiivne lahendus olnud kõigi imporditud hoonekujude kustutamine ning uue impordi läbiviimine.

Kuna „Verbatiumi“ OSM konto viimasest elumärgist möödub tänavu 10 aastat [80], peeti kasutajale sõnumi saatmist tupikuks ning keskendus algul andmete allika leidmisele läbi tuntud kaardistustevõtetega kontakteerumise. Uurimise lähtepunktiks kujunes „Verbatiumi“ impordi käigus kõigile hoonetele lisatud Type=0x13 andmesilt, mis oletatavasti pärines Garmin kaardilt [81]. Type=0x13 tähistab Garmin kaardivormingus hoonet [82].

Garmini spetsialist teatas peale infoga tutvumist, et ta ei suutnud imporditud andmete ja Garmin kaardiarhiivi andmete sarnasust tuvastada. Type-sildid võisid spetsialisti meelest olla lisatud impordi käigus, et tagada OSMi kaardi kasutatavaus Garmin seadmetes. Garmin jurist viitas oma vastuses ühe kaardiandmete võimaliku allikana Here'ile ning soovitas samuti pöörduda kohalike ettevõtete poole.

Teisel katsel õnnestus Here'iga kontakteeruda ning saada vastus, et kiri on kätte saadud, kuid peale seda saabus vaikus ning töö esitamise seisuga sisulist infot ei ole lisandunud. Kohalikest ettevõtetest suheldi kahega, mis on teinud põhikaardistust Maa-ametile [83], Regio ja EOMapiga, kuid mõlemad eitasid sarnaselt Garminiga imporditud andmete sarnasust ettevõtte enda arhiiviga

Here'ilt vastust oodates uuriti „Verbatiumi“ muudatuste ajalugu ning selgus, et kasutaja ei teinud kunagi muudatusi mujal maailmas kui ainult Eestis [80]. Muudatuskogumite atribuutide analüüsimisel selgus, et „Verbatium“ oli venekeele kasutaja [84], kes tegutses peamiselt Väike-Õismäel ja Tiskres [85]. Muudatuste analüüsi muutis raskemaks 2009. aastal kasutusele võetud API versioon 0.6, mille käigus grupeeriti varasemad muudatused automaatselt muudatuskogumitesse [50].

Järgnevalt on Joonis 1 näha kuvatõmmis tabelist, mis saadeti kõigile kaardistusettevõtetele illustreerimaks uuritavate muudatuskogumite omavahelist ajalist seotust. Tabel, mille põhjal kuvatõmmis loodi, ei ole säilinud, aga see valmis „Verbatiumi“ ajaloo põhjal HTML-faili regulaaravaldistega töödeldes [80]. Tabelist selgub, et alates 2008. aasta 6. septembri hilisõhtust kuni 8. septembri varahommikuni (UTC aja järgi) toimus andmete üleslaadimine katkematu voona. Kuni muudatuskogumini 561094 (k.a) sisaldasid üles laaditud andmete hulgas ainult sõlmed. Kogumisse 569055 sattus mõlemad: 9248 sõlme ning 40753 hoonet. 569277 lisas ainult hooneid.

Chset	Changes	Comment	Start	End	Duration	Time from prev.
545638	2	(kommentaare pole)	05.09.2008 07:51	05.09.2008 08:51	01:00.00	18:44, 0 days
546348	91	(kommentaare pole)	06.09.2008 01:02	06.09.2008 02:02	01:00.00	16:11, 0 days
549217	11	(kommentaare pole)	06.09.2008 02:23	06.09.2008 03:23	01:00.00	00:21, 0 days
550438	309	(kommentaare pole)	06.09.2008 18:41	06.09.2008 20:58	02:17.00	15:18, 0 days
556899	50000	(kommentaare pole)	06.09.2008 22:21	07.09.2008 01:02	02:41.00	01:23, 0 days
557193	50000	(kommentaare pole)	07.09.2008 01:02	07.09.2008 04:08	03:06.00	00:00, 0 days
557358	50000	(kommentaare pole)	07.09.2008 04:08	07.09.2008 06:18	02:10.00	00:00, 0 days
557568	50000	(kommentaare pole)	07.09.2008 06:18	07.09.2008 08:34	02:16.00	00:00, 0 days
558056	50000	(kommentaare pole)	07.09.2008 08:34	07.09.2008 11:18	02:44.00	00:00, 0 days
558636	50000	(kommentaare pole)	07.09.2008 11:18	07.09.2008 14:04	02:46.00	00:00, 0 days
559707	50000	(kommentaare pole)	07.09.2008 14:04	07.09.2008 16:54	02:50.00	00:00, 0 days
561094	50000	(kommentaare pole)	07.09.2008 16:54	07.09.2008 19:55	03:01.00	00:00, 0 days
569055	50000	(kommentaare pole)	07.09.2008 19:55	08.09.2008 00:10	04:15.00	00:00, 0 days
569277	9346	(kommentaare pole)	08.09.2008 00:10	08.09.2008 02:56	02:46.00	00:00, 0 days
572107	24160	(kommentaare pole)	08.09.2008 05:51	08.09.2008 13:41	07:50.00	02:55, 0 days
579407	4220	(kommentaare pole)	08.09.2008 19:38	08.09.2008 22:57	03:19.00	05:57, 0 days
591093	3377	(kommentaare pole)	09.09.2008 20:55	09.09.2008 22:32	01:37.00	21:58, 0 days
617599	155	(kommentaare pole)	12.09.2008 10:19	12.09.2008 12:15	01:56.00	11:47, 2 days

Joonis 1. Kuvatõmmis kasutaja „Verbatiumi“ hooneimpordi muudatuskogumite tabelist.

572107 on arvatavasti impordijärgne pool-käsitsi korrastamine, kus näiteks parandati ilmselged kirjavead nagu building ja Name (õige on vastavalt building ja name, mõlemad väikese algustähega) [86]. Samuti leidis muudatusi, mis oli tehtud käsitsi erinevate programmidega, nt Merkaartori ja siis paar tundi hiljem Potlatchiga. Võib oletada, et API 0.6-le üleminekul pandi ühte muudatuskogumisse kokku viimased järelejäänud hooned impordist ning hiljem käsitsi tehtud muudatused.

Talk-ee meililisti arhiivist „Verbatiumi“ otsimisel tagastati ühe vastena neljast kiri, mille oli saatnud isik nimega Valeri Kuzmin [87], kuid mitte kuskil nimetatud lõimes ei mainitud kordagi „Verbatiumi“ nime. Algselt jäeti kirjavahetus kõrvale kui oletatav valepositiivne otsingutulemus ning jätkati süvitsi Bartosz Fabianowskit tsiteeriva kirjaga [88],

millest selgus, et Fabianowski paistis teadaolevalt ainus inimene, kes on kunagi „Verbatiumiga“ suhelnud. Kirjavahetus Fabianowskiga oli edukas – selgus, et eelmainitud Valeri oligi „Verbatiumi“ kasutaja. Samuti oletas Bartoz internetiotsingut kasutades, et Valeri Kuzmini nimeline isik on kas tarkvaraarendaja Tallinnast [89] või päästja Sillamäelt [90].

Eelpool mainitud kirja saatmisel meililisti kasutas Kuzmin meiliaadressi kartograff@hotmail.ee [87], mis tänapäeval enam küll ei tööta, aga lubab oletada, et OSM wiki kasutaja „Kartograff“ võis olla „Verbatiumi“ teine konto [91]. Samuti mainiti kõnealuses lõimes eksportimist „XGisist“ [87], mis lubab oletada, et ka hoonekujud võivad pärineda Maa-ameti X-GIS kaardirakendusest [92]. „Kartograffi“ tegevuste logi uurides leiti viiteid Maa-ameti 1996–2007 põhikaardi kasutamisele [91].

Päästeametnikku peeti tõenäolisemaks vasteks kui tarkvaraarendajat. Peatükis 4.3.1 mainitud kasutajate aktiivsuse analüüsi läbi viies tehti Osmiumi käsus kirjaviga, kirjutades tekstifiltrisse 2017. aasta asemel 2007. Ekslikult tehtud päring näitas, et samal perioodil tegid suuremahulisi muudatusi teisedki kasutajad. Nii mõnedki kasutajad lisasid enda muudetud elementidele @rescue.ee-lõpulise signatuuri või kasutasid OSMi oletatavasti oma pärisnime alt, kusjuures nime sisestamisel otsimootoris selgus, et tegu oli (endiste) päästjatega. Aktiivseim nende seas oli kasutaja „Andres Kaaber“ [93].

Järgnevalt püstitati hüpotees, et kuna Kaaber ja Kuzmin on mõlemad OSM kaardistajad ning mõlemal paistab olevat seos vastavalt Lääne- ja Ida-Virumaaga, siis võivad nad olla tuttavad ning kasutasid tõenäoliselt üht ja sama Päästeametile kasutuseks mõeldud andmekomplekti, mis Fabianowski arvates pärines Maa-ametist. Maa-ameti põhikaardistuse ajaloo tutvustusest selgus, et kõige tõenäolisemad põhikaardi kaardistajad olid Eesti Kaardikeskus (liidetud hiljem Maa-ametiga) ning EOMap Lõuna [83]. Kuna senini ei olnud ükski kolmest varem küsitatud ettevõttest sidunud hoonekujusid enda kaardistamisega ning niivõrd ebakorrapärased hoonekujud tundusid kahtlased, eeldati edaspidi, et „Verbatium“ võis jälgede segamiseks omandatud vektorkihti enne üleslaadimist muuta. Järgnevate kirjade puhul ei palutud asutustel enam OSM andmeid nende endi kaardiarhiiviga võrrelda, vaid paluti hinnata autoriõigustega kaitstud kaardiandmete OSMi jõudmise ja tagajärgede lahendamise võimalusi.

Päästeametile, Maa-ametile, EOMapile ja Andres Kaaberile saadetud kirjadele vastasid esimesena EOMap ja Maa-amet. EOMap soovitas küsimusega pöörduda Maa-ameti poole, kuid pidasid leket võimalikuks. Maa-amet teatas, et isegi kui kasutati autoriõigustega kaitstud vektorandmeid, siis kuna tegemist on ammuse juhtumiga ning andmed on tänapäeval vabalt kättesaadavad, ei pea nad „juhtumi täiendavat menetlemist oma prioriteediks“. Siiski kinnitati, et WMSi kasutamine oli OSMis lubatud ka enne kui Maa-amet 2009. aastal [66] sellekohase konkreetse kirjaliku nõusoleku andis.

Päästeameti praeguse arendusosakonna nõuniku edastatud info läks vastuollu endise arendusjuhi Kaaberi vastusega [93]. Nõuniku sõnul oli Päästeametil ligipääs Maa-ameti kaardiandmetele, aga neid kasutati ainult paaril töökohal paberkaartide trükkimiseks. Andres Kaaberi väitel Päästeametis ei olnud üldse Maa-ameti (vektor-) kaarte. Samuti ei osanud ta kommenteerida midagi „Verbatiumi“ või tema andmete päritolu kohta, kui ainult nii palju, et „Verbatium“ tegutses pigem iseseisvalt. Kaaber soovitas kirjutada Jaak Lainestele ja Margus Värtsonile, sest mõlemad kuuluvad kohaliku OSM organisatsiooni juhatusse [94].

Kahele Andres Kaaberi poolt mainitud inimesele saadeti kiri. Värtson algul ei vastanud ning Laineste vastas möödaminnes üsna lühidalt, kuid tema kirjast selgus, et tegelikult on andmete päritolu üritatud varemgi uurida. Teadaolevatest suurfirmade andmetest (s.h Googlest) sarnasusi ei ole leitud, kuid Jaak pidas võimalikuks varianti, et andmed võisid lekkida SHP-failina mõnest kohalikust (tänapäevaks tegevuse lõpetanud) väikefirmast. Värtson vastas hiljem, et tema Maa-ameti hüpoteesi õigeks osutumisel andmete juriidilise korrektsuse pärast ei muretseks, kuid saatis vastuse kogemata ka avalikku meililisti, kus kirjavahetusega liitus Anti Kasuk. Viimane teadis kommenteerida, et impordi toimumise ajal ringles Eesti OSM kasutajate seas andmekogu, mis väidetavalt olevat pärinenud Tallinna andmekogudest. Kuigi varasemalt oldi tutvutud Tallinna 2006. aasta aluskaardiga ning järeldatud, et „Verbatiumi“ andmed on kehvema kvaliteedi ning laiema ulatusega (Sauelt Maarduni vs ainult Tallinn), saadeti linnaplaneerimisameti geoinfosüsteemide osakonnale siiski nende andmekogu päritolu ja ulatust täpsustav küsimus. Töö esitamise ajaks ei ole linn kirjale vastanud.

Kuni uurimise meililisti jõudmiseni lähtuti uurimisel eeldusest, et impordiga lisati kaardile juba olemasolevad vektorandmed, sest 2013. aasta Maa-ameti rasterkaardi vektoreerimisel saadi parema kvaliteediga tulemused. Omades ligipääsu 2006. aasta Tallinna

[95] ning 2008. aasta Regio [96] paberatlastele, kontrolliti nimetatud kaartidel hoonekujusid ning leiti, et nad on märgatavalt parema kvaliteediga kui uuritava andmekogul. Lisaks oletati, et vektoriseerimise tarkvara loomine ning reaalajas WMS-teenusest kaardi-paanide laadimine, vektoriseerimine ning üleslaadimine OSM APIsse sagedusega 4,5 sõlme sekundis oli 2008. aasta riistvara kasutades ebaotstarbekalt keeruline. Pidades nõu kogenuma kasutaja Mateusz Koniecznyga [97] saadi teada, et vastava tarkvara ühendamine OSM APIga on suhteliselt lihtne ning isegi kui arvuti poleks suutnud kaarti reaalajas analüüsida, oleks üleslaadimise saanud läbi viia peale analüüsi lõppu. Konieczny oletust peeti realistlikuks, sest „Verbatium“ laadis esmalt üles kõik sõlmed ning alles siis lisas juurde sõlmi ühendavad jooned. Võib järeldada, et üleslaadimisel tuli hoonete jooned vahepeal salvestada, sest kui vahemälu poleks kasutatud, oleks joonte ja sõlmede üleslaadimine pidanud kestma ligikaudu sama kaua. Eelnev oletus eeldab, et pildianalüüs on hoo-neimpordi protsessi kõige aeglasem osa.

Rasterkaardi hüpoteesi toetab teinegi tähelepanek. Praktilises osas loodud skripti testimisel tagastati veateate seoses joonega 26855405 (koordinaatidel 59.48357, 24.88379). Antud juhul pole küsimus mitte hoones endas, vaid selle kõrval asuvas seletamatus kolm-nurgas, mis on kaardile kantud mõne meetri suuruse hoonena (joone nr 26905970). Joonis 2 näitab kõrvuti kuvatõmmiseid – vasakul Maa-ameti ajalooliste kaartide veebirakenduse 1996–2007. a põhikaardist ning paremal OSMi iD kaardiredaktorist, kus on aluskaardiks valitud 2021. a põhikaart ning OSM hooned joonistatud jämeda punase joonega. Kolm-nurga ajalugu uurides selgub, et tegu oli ühega sadadest „Verbatiumi“ impordiga lisatud väikestest moodustistest, millest enamik kustutati 2016. aastal [98]. See konkreetne kolm-nurk jäeti alles, sest aerofotol asus väidetav hoone puuvõra varjus ning hoone eksistentsi ei saanud ilma kohapealse vaatluseta kontrollida [98].



Joonis 2. Maa-ameti 1996–2007 põhikaardi (vasakul) ja OpenStreetMapi andmete (paremal) võrdlus

Kui algul Fabianowskile kirjutati, soovitas viimane teavitada ka OSMi ametlikku mode-
raatoritiimi DWG, kuid nende teavitamisega viivitati kõheldes, et DWG võib olukorrale
läheneda nii, nagu hoiatati Garmini eest – reageeritakse läbi andmete kustutamise ilma,
et algsele kirjale vastataks. DWG-le saadetud kirjas tugineti kirjutamise hetkel kõige tõe-
näolisemana tundunud Maa-ameti päritolule. DWG-l võttis juhtumi osas arvamuse ku-
jundamine kauem kui tavaliselt, kuid nädal hiljem küsimust korrates anti vastus kõigest
mõne tunniga. Ametlik seisukoht oli, et DWG ei näe probleemi seni, kuni nad pole amet-
likku kaebust saanud, ning esitatud info põhjal on kaebuse saamine ebatõenäoline. „Kui
te just ise ei soovi kaebust esitada, mispuhul peaksime vajadusel kõik (algsed hooned
ning nende hilisemad versioonid) kustutama.“ DWG vastus saabus siis, kui muu kirjava-
hetus oli juba meililisti jõudnud. Värtson pidas DWG seisukohta mõistlikuks [99].

Lisamärkus kaardiandmete sarnasuse kontrollimise teemal. Ei ole teada, millist metoodi-
kat kasutasid käesolevas töös mainitud teised ja kolmandad osapooled kaartide sarnasuse
määramiseks. Siiski varustati neid veidi muudetud Overpassi päringuga (vt Lisa 2) ja ta-
beliga „Verbatiumi“ anomaalselt suurtest muudatuskogumitest (vt Joonis 1). Käesolevas
peatükis kasutati kahe kaardi võrdlemiseks visuaalset vaatlust. Kontrollpunktideks valiti
üksikute hoonetena Estonia pst 4 (Estonia teater) ja Tornimäe tn 3 (nn Tornimäe kaksik-
tornid), sest mõlemad hooned olid „Verbatiumi“ versioonis joonistatud kahe eraldi hoo-
nena. Piirkondade valimisel eelistati kohti, kus ruumipäring (vt Lisa 2) leidis tavapärasest
tihedamalt tingimustele vastavaid hooned. Sõelale jäi Künnapuu tänav Maarjamäel,
Kesk tee Haabneemes ning Nurme tänav Nõmmel. Kõigi kolme piirkonna puhul tegu
vanemate eramute piirkonnaga, kus võib majadelt leida ümmargusi juurdeehitusi, mida
„Verbatiumi“ andmestikus ja vanal Maa-ameti põhikaardil kajastati 2–3 sõlmega, mood-
sal Maa-ameti kaardil 6–9 ning 2006. aasta Tallinna aluskaardil veel enama sõlmega.
Andmete kvaliteeti hinnati lisaks eelmainitud ümaratele osadele ka täisnurkade esinemi-
ses. „Verbatiumi“ andmetes olid päriselus riskikülükujulised hooned harva korrektse ku-
juga. Hoonekuju käsitleti korrektseks kui sellel olid tavaoludes paralleelsed vastasküljed,
valdavalt täisnurksed nurgad ning kaarjad seinad olid joonistatud ühtlaste vahedega sõl-
medega.

3.4 Arenduskeskkonna valimine

Töö ülesandepüstituse üks lähtetingimusi oli tagada põhiskripti toimimine testkeskkonnas, sest niivõrd mahuka muudatuse läbiviimine avalikus serveris eeldaks põhjalikku eeltööd, mis võib jääda toppama väliste tegurite taha. OSMi avalikku põhiserverit [100] ei ole lubatud kasutada arendustegevuseks. Algul prooviti küll kasutada avalikku arendusserverit [101], aga seal oli kitsaskohaks andmete esmane serverisse üle kandmine, sest sellisel juhul oleks tulnud kasutada märkimisväärselt aeglasemat üldkasutatavat APIt, mitte PostgreSQL'i funktsionaalsust nagu kohaliku serveri korral. Peale 20 hoone üleslaadimist avalikku testkeskkonda selgus, et hoonetele anti testkeskkonnas uued IDd, mis ei olnud kuidagi sarnased OSM põhiserveris vastavate hoonete IDga. Identifikaatorite erinevus oleks tähendanud märkimisväärselt lisatööd skripti arendamisel, kuid nende lahknevust ei peetud otsuse langetamisel oluliseks näitajaks. Kohaliku virtuaalmasina kasuks otsustati just andmete ülekandmiskiiruse ja meetodi mugavuse pärast. Avastus, et Osmo-sise tööriist oskas andmete baasi migreerimisel nende algseid ID-sid säilitada, oli ootamatu, aga see-eest meeldiv üllatus. Kolmas võimalik valik skripti testimiskeskonnaks oleks olnud paluda konto avamist arendusserveris, aga selleks ei nähtud vajadust, sest kohaliku seadme jõudlusest piisas [102].

3.4.1 Arenduskeskkonna virtuaalmasina põhinäitajad.

Testkeskkonna virtuaalse riistvara võtmenäitajateks valiti 4 gigabaidi (GB) muutmälu, 4-tuumaline protsessor ning 40 GB salvestusruumi SSD salvestusseadmel. 40 GB salvestusruum on hüperviisori vaikesoovitus Ubuntu 20.04 operatsioonisüsteemile. Arvestades, et loodavasse keskkonda plaaniti laadida ainult Eesti kaart, puudus vajadus salvestusruumi suurendamiseks. Testkeskkonna jaoks kasutati SSD-d, sest kasutatud tööjaamas ei olnud tavalisel kõvekettal (HDD) lihtsalt piisavalt ruumi, SSD eelistamisel ei peetud kiirust oluliseks teguriks, sest eeldatavasti oleks pudelikaelaks kujunenud Pythoni skript koos CPUga või hiljem *live*-andmete korral võrguühendus. 4 protsessorituuma kasutamine oli samuti host-süsteemi piirangutest tulenenud vaikesoovitus. Mäluruumi vajaduse leidmine oli veidi keerukam.

Tervet maailma kattev avalik API-server kasutab 504 GB muutmälu [103]. 2021. aasta alguse seisuga oli ülemaailmse OSM kaardi maht pakitud kujul (osm.bz2-laiendiga) veidi üle 100 GB [104], samal ajal kui Eesti ala kaart jäi samas vormingus alla 150 MB [105].

Lihtsast arvutusest järeldub, et ainult Eesti alaga peaks server vajama umbkaudu 800 MB mälu (RAM). Samas tuleb sellele juurde liita lisavaru, mille reserveerib endale operatsioonisüsteem koos graafilise kasutajaliidesega. Ubuntu dokumentatsiooni andmetel vajab serveri versioon 1 GB [106] ning töölauakeskkond 3 gibibaiti (GiB) [107], ehk 3,2 GB [108]. Liites OSM serveri ja Ubuntu kasutajaliidese mäluvajadused, saadi kokku 4 GB. Oletuse õigsusele viitab ka kaardiserveri ülesseadmist juhendav veebileht Switch2OSM, mis soovitab arvestada suurele linnale 4 GB mäluruumi [109].

Virtuaalmasina operatsioonisüsteemiks valiti Ubuntu 20.04 LTS töölauaversioon. Ubuntu, sest tegu on levinud Linuxi distributsiooniga ning enamik OSMiga seonduvaid juhendmaterjale on loodud või vähemalt kontrollitud Ubuntu 18.04 või 20.04 LTS versiooniga [109]. 20.04 valiti, sest see oli uuem ning sama versiooni kasutatakse ka avaliku teenuse serveris [110]. Olulise otsusena eelistati töölauaversiooni serverivariandile, peamiselt graafilise kasutajaliidese mugavuse pärast. Nimelt võimaldab viimane teha erinevaid failihalduse ja silumise toiminguid kiiremini ja mugavamalt kui ainult käsuriid kasutades. Kolm kõige olulisemat eelist olid käsurea lõikepuhvri tugi, lihtsam failide liigutamine ning sisseehitatud veebibrauser. Viimane on eriti oluline, sest eelmainitud juhendmaterjalid ei katnud veebiteenuse võrguseadistust, vaid vaikimisi lubati teenusele ainult kohaliku hosti (s.t VMi 127.0.0.1) võrguliidese kuulamise.

4 Praktiline osa

Käesolev peatükk katab skriptide valmimise ja keskkonna ülesseadmisesse puutuva. Tutvustatakse testimiskeskkonna virtuaalmasina loomist, hoonekujude uuendamise skripti põhimõtteid, skripti töö tulemusi ning lõpuks antakse ülevaade võimalikest tuleviku edasiarendustest, sealhulgas soovitusel, mida tuleks jälgida hoonete uuendamisel OSM serveris.

4.1 Keskkonna seadistamine

Testkeskkonna virtuaalmasin seati tööle Vmware Workstation hüperviisoris Windows 10 operatsioonisüsteemiga tööjaamas, sest eelnimetatud tarkvara oli varasema ülikooli õpetöö tõttu arvutisse paigaldatud. Virtuaalmasinat ümbritsev füüsiline host-süsteem ei tohiks omada mõju loodavale testkeskkonnale ning soovijal peaks olema võimalik testkeskkonda korrata ka testsuguste hüperviisoritega. Link keskkonna seadistamise skripti juurde on saadaval kokkuvõtte lõpus.

Juhuks kui lugeja soovib keskkonna loomise skripti oma testseadmes kontrollida olgu igaks juhuks mainitud, et käesoleva skripti esimeses testkeskkonnas kasutati värskest installitud virtuaalmasinat, mille kasutajanimeks oli user. Lisaks kontrolliti skripti teises virtuaalmasinas teistsuguse kasutajanime ja parooliga, bashi skript töötas ootuspäraselt hästi. Turvakaalutustel tasub teada, et konkreetse virtuaalmasina võrgukaart oli lähtuvalt Workstationi vaikesätetest seadistatud töötama NAT-režiimis, ehk välises võrgus paiknevad seadmed ei oleks saanud VMiga ühendada [111]. Samuti paiknes host-seade tulemüüri kaitstud võrgus.

Testkeskkonna seadistamise skript on üldjoontes identne Rails-pordi Githubi repositooriumis avaldatud juhendiga [110], kuid suurima erisusena asendati vaid kohad, kus eeldati teise, interaktiivsesse konsooli, sisse logimist (nt `sudo -u postgres -i`), iseseisvale skriptile sobivamate käskudega. Üks selline olukord oli veebilehele administraatori kasutaja loomine. Juhendmaterjali järgi oleks tulnud kasutaja luua veebibrauseris ning seejärel kasutajakonto Ruby konsoolis aktiveerida [112]. Eraldi aktiveerimist on vaja sellepärast,

et server ei ole vaikesätetes võimeline iseseisvalt konto aktiveerimise e-kirju saatma. Kasutaja loomist prooviti algul automatiseerida wget-iga veebibrauserit imiteerides, kuid autentimine ei õnnestunud sessiooni küpsiste tõttu. Läbi Ruby konsooli kasutajate lisamiseks ametlikku dokumentatsiooni ei leitud, aga see eest avastati Githubi repositooriumis mitteaktiivne parandusettepanek, mille ühe osana initsialiseeriti server mõningate eelmääratud kasutajatega [113]. Leitud Ruby skriptist sai kergesti tuletada Bashi käsud, mis sarnased korraldused Ruby konsoolile edastaksid [114].

Ka andmete ülekandmine Osmosisega ei tahtnud alguses töötada nii, nagu juhendis oli ette nähtud, sest paketiindeksis sisalduv Osmosis versioon kasutas vananenud andmebaasi struktuuri [115]. Lahendus oli leida Githubi lehelt uuem versioon ning see käsitsi arvutisse paigaldada. Seejuures selgus, et kuna Osmosis tugineb oma töös paljudele spetsiifilistele Java teekidele, on mõistlik lasta Apt-l esmalt vanem versioon koos sõltuvustega installida ning siis uuem versioon teise kausta alla laadida. Osmosis on tööriist, mis on mõeldud OSM andmete töötlemiseks ning erinevate vormingute/rakenduste vahel liigutamiseks [116]. Käesolevas töös kasutati Osmosist kaardi kopeerimiseks andmebaasi.

Praktilise osa läbiviimisel märtsi lõpus oli serveri ülesseadmisel suurim takistus MimeMagic teegi litsentsivahetus MIT litsentsilt GPL-le [117]. Ruby-on Rails vajab teegi MimeMagic versiooni 0.3.x, aga litsentsivahetuse tagajärjel eemaldati kõik varasemad versioonid enne 0.4.0 [118]. Railsi arendajad lahendasid olukorra nädal hiljem asendades MimeMagicu Marceliga [117]. Vahepealsel perioodil asendati ajutise lahendusena OSM veebiserveri paketi sõltuvuste failis senine 0.3.x nõue 0.4.0-ga.

Vaikimisi kuulab server ainult kohaliku hosti (VMi 127.0.0.1) liidest. Ligipääsu lubamine füüsilise hostseadme IP-lt oli ootamatult lihtne. Kuigi eelnevalt uuriti põhjalikult rails-pordi lähtekoodi selgus, et sätet muudeti käsurealt ning tegelikkuses piisas vaid `-b 0.0.0.0` lisamisest serveri käivitamise käsule [119]. Tulenevalt NATi kasutamisest lubati selle käsuga ligipääs ainult hostseadmele, aga mitte teistele kohtvõrgus olevatele arvutitele. Käesolevas töös kirjeldatud keskkonna loomise skripti kasutamine avalikult kättesaadava API serveri seadistamiseks on tungivalt mittesoovitav, vaid lähtuda tuleks sellisel juhul spetsiaalselt vastavaks otstarbeks loodud juhistest [112].

4.2 Hoonete uuendamise skript

Töö praktilise osa põhiväärtus on Pythoni skript, mis üritab seada OSMi ja ETAK hooned omavahel vastavusse ning seejärel OSMis üle API andmeid uuendada. Skripti koostamisel üritati lähtuda põhimõttest kasutada minimaalselt lisateeke, mis ei ole Python 3.7 sissehitatud.

4.2.1 Algoritmi selgitused

Skripti võib üldjoontes jagada 4 ajamahukaks osaks: Overpassi päringu tegemine, ETAK-andmete ruumiindeksi loomine, kahe allika andmete ühildamine ning lõpuks OpenStreetMapi API kaudu hoonekujude uuendamine. Esimesed kolm osa sisaldavad tulemuse puhverdamist püsivalt, et muuta samade lähteandmetega skripti käivitamine tulevikus kiiremaks. Puhver salvestatakse enamasti stringiks konverteeritud Pythoni objektina, et tulemus oleks vajadusel inimloetav ning samuti loetav erinevate programmide poolt.

Kuna niivõrd väikesel maa-alal nagu seda on Eesti puudub vajadus arvestada maa kumerusega, siis kasutab skript geomeetria töötlemiseks ETAK-andmekogu vaikimisi ristkoordinaatide süsteemi L-EST97, mille puhul koordinaatide sammuvahe on mõlemas suunas üks meeter [120] [121]. Võrdluseks näiteks tavapärase geograafiliste koordinaatide puhul muutub meridiaanide vahekaugus võrdeliselt laiuskraadi koossinusega [122], mistõttu oleks kaugusarvutuste tegemine geograafiliste koordinaatide põhjal arvuti jaoks palju aeglasem ja keerulisem. Skripti loomisel lähtuti Maa-ameti Geo-Lest konverteerimise PHP-funktsioonidest [123], aga optimeerimiseks eemaldati korduvad ja etteennustatavad operatsioonid, näiteks konstantide omavahel korrutamine ning neli korda järjest samast muutujast siinuse leidmine.

Kahe andmekogu hoonete ligikaudse samasuse määramiseks kasutati ruumiindeksi, mis kujutas endast lihtsakoelist 100-meetriste vahedega ruudustikku ning hoonekujude külgnevuskastide kattuvuse võrdlemist. Ruudustik, mis andmestruktuurina kujutab endast paisktabelit, peab järge ühes 100x100 m ruudus paiknevatest Maa-ameti hoonetest. Ruudustiku tarbeks valiti paisksalvestus, sest sellisel juhul on otsitava ruudu mälust leidmine konstantse keerukusega [124]. Samuti on paisksalvestus lisatud Pythonisse sissehitatud andmestruktuurina – sõnastikuna (*dictionary*) [125]. Hoonet kindlate ruudustiku lahtriga siduv funktsioon sisaldab eripära, kus hoone kantakse ruudustikku kaks korda – esimese

ning keskmise sõlme asukohta. Eeldades, et enamik hooneid on ristkülikud, jääb seega tavaliselt esimene ja keskmine sõlm kujundi vastastippudesse. Seeläbi saavutatakse indekseerimisruudustikus parem tõrkekindlus olukordade vastu, kus OSMi ja Maa-ameti hoonekuju kattuvad vaid väga väikeses ulatuses. Külgnevuskastide kattuvuse võrdlemine tähendab, et kuna hulknurkade täpset kattuvuse arvutamist peeti liiga keeruliseks, joonistati hoonete ümber koordinaatteljestikuga paralleelsed ristkülikud, mille ülekatet omavahel võrreldi. Skript jättis koheselt kõrvale külgnevuskastid, mis kattusid vähem kui 15% ulatuses.

Skriptis uuendatakse ainult need hooned, mille kuju muutmine ei oma otsest mõju naaberhoonetele (ehk kahel naaberhoonel puuduvad ühised sõlmed). Seda, et hoone ei jagaks sõlmi mitte-hoonega, kontrolliti juba Overpassi päringut tehes. Kuna API vastav päring ei tee vahet, mis objektid sõlme jagavad, siis tegelikult puudub seega oluline vajadus päringu tegemisel kõnealust tingimust kontrollida. Samas vähendatakse seeläbi hooneid, mida skript peab hiljem niikuinii kontrollima, andes sellega mõningase ajasäästu. Lisaks jäetakse kõrvale hooned, mille ükskõik milline sõlm omab mistahes silte, sest skripti töö käigus selgus, et enamasti on sellisel juhul tegu hoone uksega. Kuna ETAK-andmekogus ukseava asukohta ei salvestata, võib skript ukse tahtmatult majanurka liigutada.

Lisaks olgu mainitud, et skript ei ole mõeldud hindama Maa-ameti spetsialistide kaardistuse õigsust, vaid eeldab, et elukutselised spetsialistid oskavad oma tööd piisavalt hästi. Käesoleva töö teema tutvustamisel on mõnel kuulajal tekkinud küsimus, mis saaks siis, kui kellegi hoovis seisaks lauavirn, mis on kantud Maa-ameti kaardile eraldiseiseva hoonena? Vastus: juhul, kui rohkem kui 10 aastat tagasi on samasse kohta (ekslikult) joonistatud hoone, ning viimase 10 aasta jooksul ei ole keegi kõnealuse hoone kuju muutnud, siis skript käsitleb joont nii nagu teisigi hooneid ning uuendab lauavirna kuju uuemale infole vastavaks.

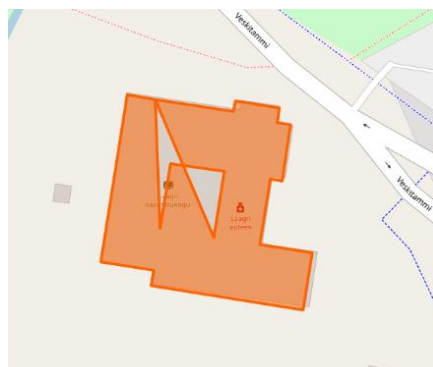
4.2.2 Skripti tulemused

Skripti esimene täispikkuses käivitamine toimus õhtusel ajal. Umbkaudu 44 000 hoonekuju töötlemine võib võtta üle 10 tunni, avaliku API korral oletatavasti isegi kauem. Esimese päeva jooksul jõuti töödelda ~8000 hoonet, neist 1600 andsid veateate, sh 600 jagatud sõlme ning 400 hoonet ei leidnud vastet. Sarnased proportsioonid püsisid ka täiskomplekti töötlemisel (allpool).

Skripti töötlemise järjestuse aluseks oli algul hoonete OSMi identifikaator (ID), mis tähendab, et kujundite muutmisel võisid muudatuskogumite külgnevuskastid katta väga suuri alasid. Mõistlikum oleks olnud hoonete grupeerimine umbes ruutkilomeetri suurus-tesse kastidesse, et külgnevuskastid oleksid väiksemad ning annaksid parema ülevaate skripti tööst. Vastav funktsionaalsus oleks teretulnud ka tulevikus, kui skriptile plaani-takse luua juurde veebiliides senimuutmata hoonete haldamiseks. Näiteks 2019. aastal, kui „SviMik“ taas-tutvustas oma uut importi, anti talle tagasisidet, et kasutajad oleks oodanud võimalust eristada oma kodukohas asuvaid hooneid [126]. Hoonete asukohapõ-hine sorteerimine lisatigi esimese ja teise täispikkuses impordikatse vahel (muudatuse tagajärgi saab näha Lisas 4).

Täismähus hoonekujude massuuendamise viidi töö raames testkeskkonnas läbi kolm korda. Uuendamiseks valitud hoonete puhul ei piiratud ainult „Verbatiumi“ hoonetega, vaid Overpassi päringu tingimuseks seati sarnaselt Lisa 2 – Lõpliku Overpassi päringu näide toodud päringuga Harjumaal paiknevad hooned, mille ühtegi sõlme pole viimase 10 aasta jooksul uuendatud ning ükski sõlm pole jagatud mitte-hoonega. Kohalikus test-keskkonnas käitatud skript talletas logisid töödeldud või töötlemata hoonete kohta, pida-des meeles, mis põhjusel hoonet ei õnnestunud uuendada.

Kuigi päringu koostamisel ning skripti kirjutamisel hoiduti vigade ennetamiseks muut-mast hooneid, mille muutmine võib ka midagi muud mõjutada, esines ikkagi vigu. Kui OSMis ilma sisehoovita hoone oli Maa-ameti kaardil koos sisehooviga, nt Laagri raama-tukogu (Joonis 3, joon 26877570), ei osanud skript midagi kahtlustada ning kandis ühe joonega kaardile nii hoone sisemise kui ka välise piirjoone. OSM ökosüsteemis on olemas validaatorid, näiteks Osmose [127], mis suudavad sellised konfliktid kergesti märgata-vaks teha.



Joonis 3. Näide vigaselt lisatud hoonekujust

Kolme käivituse suurimateks erinevusteks olid erinevate muudatuskogumitesse jaotamise viisid. Esimesel juhul tuli skript toime 28 muudatuskogumiga, millest igatükk kattis ligemale poole maakonnast. Teisel käivitamise eesmärk oli testida piiratud suurusega ühikruutude korrapärasest töötlemist, mistõttu iga muudatuskogum töötles ainult ühe ühikruuduga seotud hooneid. Tol korral edastati serverisse üle 500 muudatuste komplekti, enamik 1x1 km suurused. Kolmas katse kombineeris kaks eelmist, kus andmete üleslaadimine toimus kuni 10 km laiuste kogumikena, üritades seejuures mahutada võimalikult palju muudatusi ühte kogumisse. Käesolevat lõiku illustreerivad joonised on koondatud Lissasse 3.

Kümnekilomeetrised suurused saavutati ühe ühikruudu küljepikkuse (kilomeetri) kümnega korrutamisel, kordaja 10 valiti, et vastav koodirida oleks hiljem lugejale paremini mõistetav. Sarnane põhjus kehtis ka kilomeetrise ühikruudu suurusega, kuid seal arvestati lisatingimustega, et kuna ühikruudu tunnus on sisuliselt saadud ristkoordinaatide täisarvulisel jäägita jagamisel, peaks tõrkeotsingut teostaval kasutajal olema võimalik peast arvutades tuletada ühikruudu tunnusest ruudu asukoht. 1000-ga korrutamist peeti piisavalt kergeks tehteks, et kasutaja oleks suuteline avaldist peast lahendada.

Andmete uuendamine testkeskkonnas võttis kõigil kordadel aega veidi alla 11 tunni ning selle aja jooksul muudeti 32364 hoonekuju. Testkäivituste käigus üritas programm muuta lisaks eelmainitud 32364 veel 11561 hoonet, kuid nende puhul tagastas andmete serveris uuendamise funktsioon veateate. Alljärgnev Tabel 1 annab ülevaate unikaalsetest väljunditest, mis hoonete töötlemine tagastas. Näiteks kui hoonel oli neli nurka ja iga nurk oli jagatud mõne teise hoonega, siis tagastas skript küll neli korda „Hoone sõlm on jagatud teise hoonega“, aga käesolevas tabelis loeti selline olukord ühe korrana. Mitme põhjuse alla on koondatud juhud, kui skript andis 2 erinevat veateadet. 3 või enam erinevat veateadet ei esinenud. Kõige sagedamalt esines juhtum olukorras, kus viisnurkne hoone jagas ühist külge naabermajaga, aga Maa-ameti kaardile oli hoone kantud nelinurgana. Versioonikonflikt tähistab paari üksikut juhtu, kus teadmata asjaoludel üritas skript uuendada sama sõlme kaks korda. Reeglina tekkis viga kohe esimese sõlme juures ning andmete terviklus ei kahjustunud. Kuigi eelnevalt kirjeldati nõuet, et ükski sõlm ei tohi omada atribuutandmeid, siis vastavale tingimusele mittevastamist skript ei tuvastanud, sest arvatavasti eemaldati sellised elemendid juba Overpassi päringu tegemisel. Erandlik väljund

Stop tähistab vastavalt kirjeldusele kas skripti katkestamist või töö lõppemist. Paksus kirjas Fail, Stop ja Success tähistavad väljundi kategooriat üldisema filtreerimise jaoks.

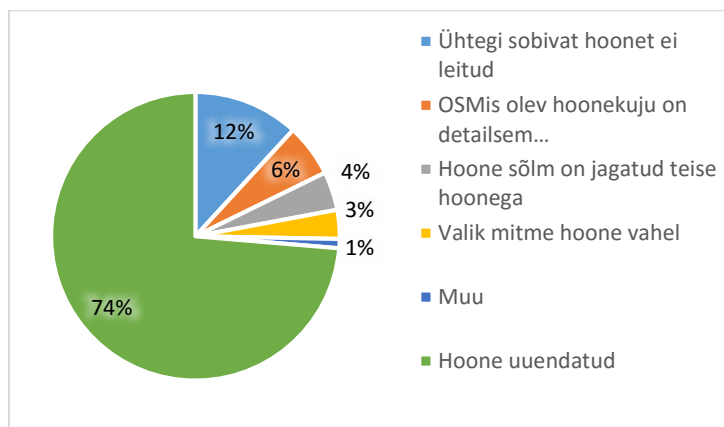
Tabel 1. Skripti käigus tagastatud tulemuste jaotus

Funktsiooni väljund	Hoonekujude arv
Fail	11561
Ühtegi sobivat hoonet ei leitud	5198
OSMis olev hoonekuju on detailsem kui Maa-ameti kontuur	2594
Hoone sõlm on jagatud teise hoonega	1897
Valik mitme hoone vahel	1424
Mitu põhjust (2 tk)	442
Versioonikonflikt	6
Stop	4
Skript katkestatud	3
Skript lõpetas edukalt	1
Success	32364
Hoone uuendatud	32364
Kokku	43929

Tabelis 1 ja peatükis 3.2 toodud hoonete koguarv ei lange täpselt kokku, sest kahe päringu vahele jäänud perioodil (märtsi algus vs aprilli lõpp) võisid teised kasutajad kaarti muuta ning seega skripti alguses tehtud päringu ajaks jäi alles paarsada hoonet vähem.

Alljärgneval joonisel kuvatakse parema visualiseerimise huvides eelmises tabelis toodud andmed sektordiagrammina. Marginaalse tähtsusega väljundid, ehk olukorrad, mida esines vähem kui 1000 korda koondati ruumi kokkuhoiu huvides ühise nimetaja „Muu“ alla. Eelnenud tabelist 1 ning järgnevast Joonis 4 selgub, et skript tuli ilma märkimisväärsete probleemideta toime peaaegu kolmveerandil juhtudest. Minimaalse kasutajapoolse sisendiga oleks arvatavasti lahendatav 10% seni katmata juhtudest. Esiteks juhtumid, kus skript tuvastas uuritava hoone külgnevuskasti enam kui 15% kattuvuse vähemalt kahe hoonega („Valik mitme hoone vahel“), ning teiseks olukorrad, kus varasem vana hoonekuju on detailsem kui Maa-ameti kaardistatu. Viimast võisid „Verbatiumi“ andmestikus põhjustada näiteks kõrvalhoonete kompleksid, mis olid varem kaardistatud ühe suure hoonena, kuid tänapäevasel kaardil on iga ehitis kaardistatud eraldiseisvana. Olukorrad,

kus hoone sõlm on jagatud teise hoonega, on samuti automaatselt lahendatavad, kuid võivad tekitada tarbetu sõlmede dubleerimise. Lisaks enesestmõistetavale tähendusele võib sobiva hoone puudumist tõlgendada ka nii, et sobiv hoone on läheduses küll olemas, kuid kujundite suhteline nihe on nii suur, et nad ei kattu piisaval määral, et ETAK hoonekuju külgnõlvuskast kattuks OSMi omaga vähemalt 15% ulatuses.



Joonis 4. Skripti käigus tagastatud tulemuste jaotus

Testkeskkonnas kasutati 20. märtsi seisuga saadaval olnud versioone nii OSMist kui ka ETAK andmekogust. Kui muudatus viia läbi tegelike, töö avaldamise ajal OSM serveris kasutusel olnud andmetega, siis peaks tulemus olema suuresti sama, äärmuslikul juhul on vaid paarsada hoonet vahepeal kustutatud või muudetud.

Peale mitut skripti kontrollkäivitust selgus, et mõnikord uuendati andmeid isegi siis, kui seda polnud vaja, sest sama hoone oli varem juba parandatud. Lihtne üldtuntud lahendus oli võtta üle sarnane lähenemine, mida kasutatakse ka populaarsetes kaardiredaktorites, ehk uus versioon saadetakse serverisse vaid siis, kui seda tegelikkuses ka on muudetud.

Peatüki lõpetamiseks sobib skripti töö tulemust illustreerima järgnev joonis, mis näitlikustab hooneid, kus skripti oli kõige suurem mõju varasemale hoonekujule. Lisaks eelmainitud tabelis 1 kajastatud logile salvestas skript ka edukalt muudetud hoonete statistikat. See oli vaid lihtne tabel, kus oli kirjas hoone ID, sõlmede arv ning mitu sõlme skripti lisas.

Järgneval Joonis 5 on näidatud hooned, mille kuju muutus vastava logi põhjal kõige enam detailsemaks, ehk teisisõnu kuvatakse hooned, millele lisati kõige rohkem uusi sõlmi või oli neil palju sõlmi juba olemas. Pruun hulknurk on töö kirjutamise ajal kasutusel olnud vana hoonekuju kuvatuna avaliku OSM kaardi peal, sinine täiteta murdjoon on testserveri

andmekiht, ehk hoone kuju peale töötlemist. Iga hoone juures on punane arv, mis näitab selle hoone ID-d OSM andmekogus. Näiteks saab aadressil <https://osm.org/way/26905329> avada veebikaardil ülemise vasakpoolse hoone ning vaadata selle täpset asukohta.



Joonis 5. Näiteid skripti töö tulemusest

4.3 Edasiarendamise suunad tulevikus

Käesoleva töö skooopi ei kuulunud andmete reaalne massuendamine OpenStreetMapi põhisaadil, sest ei ole võimalik prognoosida võimalikku ajakulu, kui muudatuse kooskõlastamisel oleks tekkinud pikem arutelu. Käesolevas peatükis antakse esmalt ülevaade tegevustest, mida võiks lisaks peatükis 2.3.2 kirjeldatud kohustuslikele tegevustele veel teha, ning seejärel esitatakse näiteid edasiarendamise võimalustest kaugemaks tulevikuks peale 2008. aasta impordi probleemi lahendamist.

4.3.1 Massmuudatuse ettevalmistamine

Osana tavapärasest automaatmuudatuste tegemise protseduurist, mida kirjeldati OSMi tutvustavas peatükis, tuleb tegevust arutada kohalikus ametlikus meililistis [42]. Sarnaselt üldise kaasautorite vähesusega on ka Eesti talk-ee meililistis inimeste osalus väga tagasihoidlik ning meiliarhiivi [128] ja kaardimuudatuste ajalugu [105] võrreldes pandi tähele, et parimal juhul vaid pooled meililisti osalised on ise viimase aasta jooksul Eestis mida-

gigi teinud. Siit tulenevalt arvati, et saamaks meililistist Eesti kasutajate arvamust objektiivset kajastav arutelu, tuleks kutsuda aktiivsemad kohalikud (käesoleval juhul Harjumaa) kasutajad arutelus osalema. Analüüsi on hea teha keskkonna seadistamise skripti osana paigaldatud Osmiumiga [129].

Järgnev Bashi käsk leiab Osmiumi tööriista abiga kõige aktiivsemad kasutajad aastatel 2019–2021. Esimene käsk konverteerib binaarfaili tekstikujule; teine eraldab regulaarvaldisega read, kus ajatempel algab kujul t2019, t2020 või t2021. Seejärel eraldatakse kasutajanimed, nimed sorteeritakse, loendatakse iga nime esinemissagedus ning lõpuks tagastatakse nimekiri 20 kõige aktiivsema kasutajaga. Skripti tulemust ei tohi käesolevas dokumendis avaldada, sest sisendina kasutatud Eesti kaardi muudatuste ajalugu pärineb Geofabriku piiratud ligipääsuga serverist ning nende kasutustingimused keelavad privaat-suskaalutustel tuletatud andmete levitamise OSMis registreerimata kasutajatele [105].

```
osmium cat -f opl estonia-internal.osh.pbf | grep -P ' t20(19|20|21)-' | grep -Po '(?<= u)[^ ]+(?= )' | sort | uniq -c | sort -nr | head -20
```

Maa-ametil oli töö kirjutamise ajal väljatöötamisel uus avaaandmete komplekt, kus ETAK-andmekihile lisatakse info 3D-mudelite kohta [130]. Arvestades, et OSM andmemudel lubab talletada teavet ka hoonete või nende osade kõrguse, värvi ja katuse kuju kohta ning leidub erinevaid rakendusi [131], mis oskavad nimetatud infot ka kasutajale kuvada, võib eeldada, et uued kõrgusandmed peaksid olema imporditavad ka OSM kaardile. Maa-amet pakkus aerolaserskaneerimisel saadud kõrgusandmeid juba varem, kuid uus strateegia peaks lubama täpsemat ning paremini töödeldavat infot [130]. Käesolevas töös loodud skript lisab võimaluse korral hoonele kui tervikule automaatselt kõrgusinfo, aga uus andmekogu lubab oletatavasti kasutada kõrgusinfot juba detailsemalt.

Käesolev suur planeeritav import on võimalus viia läbi muid iseenesest kasulikke, kuid potentsiaalselt vastuolulisi muudatusi. Esiteks uuendada Maa-ameti ETAK andmekogule viitavad maaamet:ETAK atribuudid, mida lisab „SviMik“ oma aadressiimportimistel hoonetele. Aprilli lõpu seisuga on silt lisatud peaaegu veerandile Harjumaa hoonetele [132]. Järgmiseks võiks kaaluda vananenud ning võimalike spämm-siltide eemaldamiskorrastamist. Vastuolulist Type=0x13 OSMis enam ei leidu, kuid järgmises alapeatükis mainitavaid võtmeid CityIdx ning created_by on tänapäeval lubatud automaatselt eemaldada [133]. Kaht silti esineb kokku pisut üle 100 korral, valdavalt Väike-Õismäel. Muudatusi võidakse pidada vastuoluliseks, sest esiteks andmeid kustutataks automaatselt

tegevuse käigus ning samuti võib väita, et kuna eemaldatavatel andmetel puudub märkimisväärne sisuline väärtus, puudub väärtus ka nende kustutamisel.

4.3.2 Edasiarendused kaugemas tulevikus

Juhuks kui tulevikus soovitakse valminud skripte kellegi poolt edasi arendada, siis tuleb kindlasti lisada paremad infoturbemehhanismid, sest praegune lahendus sisaldab halbu praktikaid. Kõige karjuvam näide on Pythoni skripti poolt puhverdatud andmete mällu laadimine läbi faili käivitamise [134], samas kui turvalisem alternatiiv on Jsoni või Pickle teegi kasutamine [135]. Pickle on küll kiire [135], aga inimloetava vahemälu tingimusest tulenevalt jääb arvatavasti valikuks Pythoni Jsoni-teen. Käesolevaga tasub korrata üle, et juhul kui lugeja soovib testkeskkonna skripti kasutada avaliku OSM API alternatiivi siseseadmiseks peab ta kindlasti võtma ette turvalisusealaseid lisasamme, alustades näiteks tulemüüri seadistamisest.

Testkäivituse tulemusi jagati rahvusvahelises OSM-teemalises foorumis ning Poola kasutajad tundsid huvi võimaluse vastu kasutada skripti nende sarnase riikliku andmestiku importimiseks [7]. See tähendab, et esimesed sammud peale eelmises alapeatükis kirjeldatud tegevuste sooritamist on tõlkida Pythoni skripti kommentaarid ja muutujanimed täies mahus inglise keelde, vähendada sõltuvust Eesti EPSG:3301 koordinaatsüsteemist ning paranda skripti dokumentatsiooni.

Üks täiendav suund, mida ei olnud arvutiriistvarast tingitud piirangute tõttu võimalik uurida, oli „Verbatiumi“ vanade hoonete atribuutandmed. Muudatuskogumite uurimiseks tuleb varuda palju aega ja kannatust, sest impordi muudatuskogumid on väga suured, 5 korda suuremad kui kaasajal OSM APIs lubatud [136]. Veidi uuriti impordi viimast muudatuskogumit 572107, kus esines tundmatu võti RoadID ning paarkümmend hoonet sisaldasid silti `cityId=213` [86]. Oletatavasti pärineb silt andmekogust, mida impordi käigus kasutati.

Käesoleva töö kirjutamise käigus loodi Overpassi päring varem kui Pythoni skript. Päringu koostamisel ei arvestatud tulevase skripti eripäradega, vaid üritati ühe päringuga saavutada võimalikult ülevaatlik tulemus, mis sobiks teiste osapooltega jagamiseks. Töö koostamise lõpufaasis selgus, et Overpass toetab väljundina ka külgnevuskasti terve hoone asemel [137]. Hetkeseisuga ei vaja skript andmeid terve hoone kohta, vaid esimene sammuna peale hoonekuju saamist arvutatakse viimasest külgnevuskast.

5 Kokkuvõte

Käesolev töö keskendus probleemile, kus OpenStreetMap (OSM) kaardile on Tallinna ja lähivaldade hooned lisatud 2008. aastal tundmatust allikast pärineva impordiga ning tänaseks hoonekujud kaasaegsete arusaamade kohaselt ebakvaliteetsed. Töö põhieesmärk oli luua skript, mis suudaks uuendada hooneid, mille kuju ei ole peale 2008. aasta impordi muudetud. Eesmärk ei olnud hoonete tegelik uuendamine avalikus veebiteenuses, sest niivõrd suure ja mahuka toimingu tegemine oleks nõudnud keerulist kooskõlastusprotsessi, mis ei oleks pruukinud töö tähtajaks lõppeda.

Konkreetsest impordiga lisati suurusjärgus 50 kuni 60 tuhat hoonet, neist 40 211 sisaldavad vähemalt üht nurka, mida pole tänini liigutatud [138]. Selleks, et minimeerida automaatmuudatuste käigus tehtavat tahtmatut kahju, jäeti kohe esialgse päringuga kõrvale hooned, mille mistahes sõlme oli vahepeal liigutatud või oli hoone ühendatud mitte-hoonega. Muutmaks päringut universaalsemaks, loobuti konkreetse kasutaja tingimusest otsiti selle asemel sõlmi, mille viimasest muutmisest oli möödunud üle 10 aasta.

Varsti peale tööga alustamist selgus, et töö praktiline osa omaks mingisugustki väärtust, tuleb otsida üles 2008. aasta impordi allikas, sest tegu võis olla autoriõiguste rikkumisega. Andmete päritolu leidmisest sai töö kõrvaleesmärk ning hüpoteesina oletati, et andmed võisid pärineda Garmin kaardilt, sest impordis varustati kõik hooned atribuudiga `Type=0x13`, mida kasutatakse Garmin kaardil. Garmin esindaja lükkas hüpoteesi kiirelt ümber viidates, et nimetatud atribuutandmeid kasutatakse kolmandate osapoolte kaartide kuvamiseks Garmin sedametes, mitte vastupidi.

Uurimine tuvastas, et impordi läbiviija oli Eestis elav OSMi kasutaja. Kasutaja endaga ei õnnestunud küll kontakteeruda, kuid uuriti avalikult kättesaadavat sisu, mida oli isik oma kontodelt OSMiga seonduvatesse teenustesse laadinud ning vesteldi mitmete kaardistusalaste organisatsioonidega. Kõrvutades imporditud andmeid teiste uuritava perioodist kättesaadavate kaartidega avastati märgatav sarnasus Maa-ameti 1996–2007 raster-põhi-kaardiga. Peale konsulteerimist Maa-ameti, OSMi moderaatoritiimi ja kohalike pikaajsete kasutajatega ning jõuti järeldusele, et kuna võimalikust rikkumisest on möödunud

niivõrd palju aega, Maa-amet ei pidanud menetluse alustamist vajalikuks ning seni ei ole keegi ametlikku kaebust esitanud; ei pidanud moderaatorid andmete eemaldamist vajalikuks. Töö käigus ei õnnestunud ammendavalt tõestada 1996–2007 põhikaardi kasutamist, kuid kui nimetatud kaarti siiski kasutati, siis Maa-amet kinnitas, et tegemist oli avaandmetega, mida oli lubatud kasutada isegi varem, kui amet vastavasisulisel loa 2009. aasta sügisel andis.

Praktilises osas koostati esmalt lihtne Bash'i skript, mis valmistaks kohalikuks testimiseks ette Ubuntu virtuaalmasina OSM serverina koos Eesti kaardi andmetega, ning seejärel loodi Pythoni programm. Pythoni skript suutis Maa-ameti ETAK ja OSMi andmekogu alusel kohalikus testserveris hoonekujusid uuendada. Skripti käivitati täismahus kolmel korral, kõigil juhtudel kulus 37 tuhande hoone töötlemiseks pisut üle 10 tunni. Edukalt õnnestus uuendada üle 27 tuhande hoonekuju ehk 74% kõigist esialgsest filtrist läbi pääsenud hoonetest. Skripti tulemust peetakse piisavalt heaks, sest ülejäänud 26% muutmine oleks võinud kahjustada teiste andmete kvaliteeti.

Varsti peale käesoleva töö valmimist, võib-olla isegi enne kaitsmist, plaanitakse tähtsaimana reaalse andmete uuendamist avalikus teenuses. Võimalikud edasiarenduste suunad tulevikuks hõlmavad 2008. a impordiga lisatud atribuutandmete süvitsi uurimist ning skripti korrastamist ja dokumenteerimist viisil, mis sobib kasutamiseks ka välismaalastele, mitte ainult eestikeelsetele kasutajatele L-EST97 projektsioonis andmete jaoks.

Valminud skript on kättesaadav Githubis [1].

Kasutatud kirjandus

- [1] K. J. Reinmets, „Link käesolevas töös valminud skriptidele,“ [Võrgumaterjal]. Kättesaadav: https://github.com/kallejre/geo/tree/main/building_geometry_importer.
- [2] „Ruumiandmete seadus,“ 1.12.2020. [Võrgumaterjal]. Kättesaadav: <https://www.riigiteataja.ee/akt/109102020004?leiaKehtiv#para3lg1>. [Kasutatud 27.04.2021].
- [3] „Maa-amet, Maintaining imported data,“ 5.10.2019. [Võrgumaterjal]. Kättesaadav: https://wiki.openstreetmap.org/wiki/Maa-amet#Maintaining_imported_data. [Kasutatud 19.02.2021].
- [4] P. Neis, „Eesti ala redigeerimise statistika,“ 20.02.2021. [Võrgumaterjal]. Kättesaadav: <https://osmstats.neis-one.org/?item=countries&country=Estonia&date=21-2-2021>. [Kasutatud 21.02.2021].
- [5] OpenStreetMap Wiki kaasautorid, „Import/Catalogue - One-Time Imports,“ 4.02.2021. [Võrgumaterjal]. Kättesaadav: https://wiki.openstreetmap.org/wiki/Import/Catalogue#One-Time_Imports. [Kasutatud 18.02.2021].
- [6] M. Värtson, J. Laineste ja "SviMik", „Talk-ee: Huvitavad tööriistad Eesti jaoks,“ 2.04.2013. [Võrgumaterjal]. Kättesaadav: <https://groups.google.com/g/openstreetmap-ee/c/SnTG1oYOOdE>. [Kasutatud 15.02.2021].
- [7] "Zen [PL]", „Vestlus kasutajaga "Zen [PL]",“ 25.04.2021. [Võrgumaterjal]. Kättesaadav: <https://discordapp.com/channels/413070382636072960/599196002699837461/835594948211900426>. [Kasutatud 25.04.2021].
- [8] T. Sutton, O. Dassau and M. Sutton, A Gentle Introduction to GIS, Eastern Cape: Department of Land Affairs, 2009.

- [9] T. Jagomägi, *Geoinformaatika praktikule*, Tartu: Regio, 1999.
- [10] P. Mooney ja M. Minghini, „A Review of OpenStreetMap Data,“ *Mapping and the Citizen Sensor*, London, Ubiquity Press, 2017, pp. 37-60.
- [11] J.-F. Girres ja G. Touya, „Quality Assessment of the French OpenStreetMap Dataset,“ *T. GIS*, kd. 14, pp. 435-459, august 2010.
- [12] G. Blank ja B. C. Reisdorf, „The Participatory Web,“ *Information, Communication & Society*, kd. 15, nr 4, pp. 537-554, 2012.
- [13] P. Neis, M. Goetz ja A. Zipf, „Towards Automatic Vandalism Detection in OpenStreetMap,“ *ISPRS International Journal of Geo-Information*, kd. 1, pp. 315-332, 22 november 2012.
- [14] M. F. Goodchild, „Citizens as Voluntary Sensors: Spatial Data Infrastructure in the World of Web 2.0,“ *International Journal of Spatial Data Infrastructures Research*, kd. 2, pp. 24-32, 2007.
- [15] C. Barron, P. Neis ja A. Zipf, „A Comprehensive Framework for Intrinsic OpenStreetMap Quality Analysis,“ *Transactions in GIS*, kd. 18, nr 6, pp. 877-895, 2013.
- [16] M. Haklay, „How Good is Volunteered Geographical Information? A Comparative Study of OpenStreetMap and Ordnance Survey Datasets,“ *Environment and Planning B: Planning and Design*, kd. 37, nr 4, pp. 682-703, 2010.
- [17] D. Zielstra, H. H. Hochmair ja P. Neis, „Assessing the Effect of Data Imports on the Completeness of OpenStreetMap – A United States Case Study,“ *Transactions in GIS*, kd. 17, nr 3, pp. 315-334, 28.05.2013.
- [18] S. E. Spielman, „Spatial Collective Intelligence? Credibility, accuracy, and Volunteered Geographic Information,“ *Cartography and geographic information science*, kd. 41, nr 2, pp. 115-124, 2014.

- [19] P. Neis ja A. Zipf, „Analyzing the Contributor Activity of a Volunteered Geographic Information Project — The Case of OpenStreetMap,“ *ISPRS International Journal of Geo-Information*, kd. 1, nr 2, pp. 146-165, 2012.
- [20] M. Arrington, „Social Mapping: Game Not Over Yet,“ 18.10.2007. [Võrgumaterjal]. Kättesaadav: <https://techcrunch.com/2007/10/18/social-mapping-game-not-over-yet>. [Kasutatud 6.05.2021].
- [21] P. A. Johnson ja R. Sieber, „Situating the Adoption of VGI by Government,“ %1 *Crowdsourcing geographic knowledge*, 2013, pp. 65-81.
- [22] J. Nielsen, „The 90-9-1 Rule for Participation Inequality in Social Media and Online Communities,“ 8.10.2006. [Võrgumaterjal]. Kättesaadav: <https://www.nngroup.com/articles/participation-inequality>. [Kasutatud 6.05.2021].
- [23] J. Bennett, *OpenStreetMap: Be Your Own Cartographer*, Birmingham: Packt Publishing, 2010.
- [24] G. Boeing, „Exploring Urban Form through OpenStreetMap Data: A Visual Introduction,“ Routledge, Abingdon, 2020.
- [25] J. Jokar Arsanjani, A. Zipf, P. Mooney ja M. Helbich, „An Introduction to OpenStreetMap in Geographic Information Science: Experiences, Research, and Applications,“ %1 *OpenStreetMap in GIScience*, Springer, 2015, pp. 1-15.
- [26] S. Poole, „Contributors per capita,“ 22.11.2015. [Võrgumaterjal]. Kättesaadav: https://wiki.openstreetmap.org/wiki/File:Contributors_per_capita.png. [Kasutatud 10.05.2021].
- [27] OpenStreetMap Foundation, „Working Groups,“ [Võrgumaterjal]. Kättesaadav: https://wiki.osmfoundation.org/wiki/Working_Groups. [Kasutatud 8.05.2021].
- [28] OpenStreetMap Foundation, „Licence and Legal FAQ,“ 29.10.2020. [Võrgumaterjal]. Kättesaadav: https://wiki.osmfoundation.org/wiki/Licence/Licence_and_Legal_FAQ. [Kasutatud 8.05.2021].

- [29] H. Wood, „Automated redactions complete,“ 26.07.2012. [Võrgumaterjal]. Kättesaadav: <https://blog.openstreetmap.org/2012/07/26/automated-redactions-complete/>. [Kasutatud 28.04.2021].
- [30] OpenStreetMap wiki kaasautorid, „Stats,“ 22.04.2021. [Võrgumaterjal]. Kättesaadav: <https://wiki.openstreetmap.org/wiki/Stats>. [Kasutatud 29.04.2021].
- [31] J. Anderson, D. Sarkar ja L. Palen, „Corporate Editors in the Evolving Landscape of OpenStreetMap,“ *ISPRS International Journal of Geo-Information*, kd. 8, nr 5, pp. 1-18, 2019.
- [32] M. Moscholaki, „Comparison of remotely sensed and OpenStreetMap registered water reservoirs,“ Delft University of Technology, Delft , 2020.
- [33] A. Schutzberg, „Ten Things You Need to Know About OpenStreetMap,“ 19.04.2012. [Võrgumaterjal]. Kättesaadav: <https://www.directionsmag.com/article/1823>. [Kasutatud 23.04.2021].
- [34] F.-B. Mocnik, A. Mobasher ja A. Zipf, „Open source data mining infrastructure for exploring and analysing OpenStreetMap,“ Springer, Heidelberg, 2018.
- [35] „OpenHistoricalMap: About,“ 2021. [Võrgumaterjal]. Kättesaadav: <https://openhistoricalmap.org/about>. [Kasutatud 26.04.2021].
- [36] „OGF: About OpenGeofiction,“ 30.08.2020. [Võrgumaterjal]. Kättesaadav: https://wiki.opengeofiction.net/wiki/index.php/OGF:About_OpenGeofiction. [Kasutatud 25.04.2021].
- [37] K. T. Jacobs, „Quality Assessment of Volunteered Geographic Information,“ Carleton University, Ottawa, 2018.
- [38] „OSM SWOT,“ 7.09.2020. [Võrgumaterjal]. Kättesaadav: https://wiki.openstreetmap.org/wiki/OSM_SWOT. [Kasutatud 6.03.2021].

- [39] OpenStreetMap wiki kaasautorid, „Data working group,“ 22.11.2019. [Võrgumaterjal]. Kättesaadav: https://wiki.openstreetmap.org/wiki/Data_working_group. [Kasutatud 11.04.2021].
- [40] „Kasutaja blokeeringute loetelu,“ 26.04.2021. [Võrgumaterjal]. Kättesaadav: https://www.openstreetmap.org/user_blocks. [Kasutatud 26.04.2021].
- [41] M. A. Brovelli ja G. Zamboni, „A New Method for the Assessment of Spatial Accuracy and Completeness of OpenStreetMap Building Footprints,“ 24.07.2018. [Võrgumaterjal]. Kättesaadav: https://res.mdpi.com/d_attachment/ijgi/ijgi-07-00289/article_deploy/ijgi-07-00289.pdf. [Kasutatud 25.04.2021].
- [42] OpenStreetMap wiki kaasautorid, „Import/Guidelines,“ 21.02.2021. [Võrgumaterjal]. Kättesaadav: <https://wiki.openstreetmap.org/wiki/Import/Guidelines>. [Kasutatud 29.04.2021].
- [43] OpenStreetMap wiki kaasautorid, „Automated Edits code of conduct,“ 21.02.2021. [Võrgumaterjal]. Kättesaadav: https://wiki.openstreetmap.org/wiki/Automated_Edits_code_of_conduct. [Kasutatud 30.04.2021].
- [44] OpenStreetMap wiki kaasautorid, „Organised Editing Guidelines,“ 11.09.2019. [Võrgumaterjal]. Kättesaadav: https://wiki.openstreetmap.org/wiki/Organised_Editing_Guidelines. [Kasutatud 30.04.2021].
- [45] OpenStreetMap Wiki kaasautorid, „Main Page,“ 19.07.2020. [Võrgumaterjal]. Kättesaadav: https://wiki.openstreetmap.org/wiki/Main_Page. [Kasutatud 30.04.2021].
- [46] OpenStreetMap wiki kaasautorid, „Elements,“ 3.11.2020. [Võrgumaterjal]. Kättesaadav: <https://wiki.openstreetmap.org/wiki/Elements>. [Kasutatud 20.03.2021].
- [47] OpenStreetMap wiki kaasautorid, „Editor usage stats,“ 9.01.2021. [Võrgumaterjal]. Kättesaadav: https://wiki.openstreetmap.org/wiki/Editor_usage_stats. [Kasutatud 11.05.2021].

- [48] R. Westerholt, H. Lorei ja B. Höfle, „Behavioural Effects of Spatially Structured Scoring Systems in Location-Based Serious Games—A Case Study in the Context of OpenStreetMap,“ *ISPRS International Journal of Geo-Information*, kd. 9, nr 2, 2020.
- [49] "SviMik", „Active users,“ 5.05.2021. [Võrgumaterjal]. Kättesaadav: <http://osm.svimik.com/users.php>. [Kasutatud 5.05.2021].
- [50] OpenStreetMap wiki kaasautorid, „API v0.6 - Changesets,“ 8.04.2021. [Võrgumaterjal]. Kättesaadav: https://wiki.openstreetmap.org/wiki/API_v0.6#Changesets_2. [Kasutatud 17.04.2021].
- [51] OpenStreetMap Wiki kaasautorid, „Changeset,“ 5.02.2021. [Võrgumaterjal]. Kättesaadav: <https://wiki.openstreetmap.org/wiki/Changeset>. [Kasutatud 11.05.2021].
- [52] M. Haklay ja P. Weber, „OpenStreetMap: User-Generated Street Maps,“ *IEEE Pervasive Computing*, kd. 7, nr 4, pp. 12-18, 2008.
- [53] OpenStreetMap wiki kaasautorid, „Database,“ 23.03.2020. [Võrgumaterjal]. Kättesaadav: <https://wiki.openstreetmap.org/wiki/Database>. [Kasutatud 10.05.2021].
- [54] OpenStreetMapi kaasautorid, „Component overview,“ 3.11.2018. [Võrgumaterjal]. Kättesaadav: https://wiki.openstreetmap.org/wiki/Component_overview. [Kasutatud 10.04.2021].
- [55] L. Adenaw, J. Kreibich, M. Wittmann, L. Merkle, A. Waclaw ja M. Lienkamp, „MAGIS – A Geographic Information System for Mobility Data Analysis,“ %1 *IEEE Intelligent Transportation Systems Conference (ITSC)*, 2019.
- [56] OpenStreetMap wiki kaasautorid, „API,“ 27.07.2020. [Võrgumaterjal]. Kättesaadav: <https://wiki.openstreetmap.org/wiki/API>. [Kasutatud 15.04.2021].

- [57] „Openstreetmap website,“ 5.01.2021. [Võrgumaterjal]. Kättesaadav: <https://github.com/openstreetmap/openstreetmap-website/blob/master/README.md>. [Kasutatud 24.04.2021].
- [58] C. Fonte, M. Minghini ja V. Antoniou, „Automated methodology for converting OSM data into a Land Use/Cover,“ %1 *6th International Conference on Cartography & GIS*, 2016.
- [59] OpenStreetMap wiki kaasautorid, „Overpass API,“ 30.12.2020. [Võrgumaterjal]. Kättesaadav: https://wiki.openstreetmap.org/wiki/Overpass_API. [Kasutatud 5.04.2021].
- [60] „OpenStreetMap and the Overpass API,“ [Võrgumaterjal]. Kättesaadav: <https://dev.overpass-api.de/overpass-doc/en/preface/preface.html>. [Kasutatud 14.04.2021].
- [61] „Overpass Turbo,“ [Võrgumaterjal]. Kättesaadav: <https://dev.overpass-api.de/overpass-doc/en/targets/turbo.html>. [Kasutatud 17.04.2021].
- [62] „Overpassi päring eesti vanimate sõlmede leidmiseks,“ [Võrgumaterjal]. Kättesaadav: <https://overpass-turbo.eu/s/17kI>.
- [63] Avatud Maakaardi Selts, „Seltsist,“ [Võrgumaterjal]. Kättesaadav: <https://www.openstreetmap.ee/seltsist>. [Kasutatud 12.05.2021].
- [64] OpenStreetMap wiki kaasautorid, „WikiProject Estonia,“ 23.01.2020. [Võrgumaterjal]. Kättesaadav: https://wiki.openstreetmap.org/wiki/WikiProject_Estonia. [Kasutatud 10.05.2021].
- [65] V. Antoniou, L. See, G. Foody ja C. C. Fonte, „The Future of VGI,“ %1 *Mapping and the Citizen Sensor*, London, Ubiquity Press, 2017, pp. 377-390.
- [66] Maa-amet, „Vastus taotlusele,“ 17.09.2009. [Võrgumaterjal]. Kättesaadav: https://svimik.com/Maa-amet_vastus_OSM.pdf. [Kasutatud 15.02.2021].

- [67] Maa-amet, „Maa-ameti avatud ruumiandmete litsents,“ 10.07.2018. [Võrgumaterjal]. Kättesaadav: https://geoportaal.maaamet.ee/docs/Avaandmed/ETAK_ruumiandmete_litsentsileping.pdf. [Kasutatud 12.05.2021].
- [68] K. Kaukver, M. Aren ja T. Hurt, „Eesti topograafia andmekogu andmed, katastripiirid ja kitsenduste mõjualad muutusid avaandmeteks,“ Maa-amet, 1.07.2018. [Võrgumaterjal]. Kättesaadav: <https://www.maaamet.ee/et/uudised/eesti-topograafia-andmekogu-andmed-katastripiirid-ja-kitsenduste-mojualad-muutusid>. [Kasutatud 10.05.2021].
- [69] M. Hentschel ja B. Wagner, „Autonomous robot navigation based on OpenStreetMap geodata,“ *13th International IEEE Conference on Intelligent Transportation Systems*, 2010.
- [70] "SviMik" ja J. Laineste, „Talk-ee: Verbatium buiding import discussion,“ 14.09.2019. [Võrgumaterjal]. Kättesaadav: <https://groups.google.com/g/openstreetmap-ee/c/4tXttib2vFg>. [Kasutatud 22.04.2021].
- [71] J. Laineste, „Talk-ee: Laagna Tee,“ 7.07.2019. [Võrgumaterjal]. Kättesaadav: <https://groups.google.com/g/openstreetmap-ee/c/6SKdNPzy0fg/m/wbVZGBjcBgAJ>. [Kasutatud 12.05.2021].
- [72] G. Appleton, „Overpassify,“ 7.07.2018. [Võrgumaterjal]. Kättesaadav: <https://github.com/gappleto97/overpassify/>. [Kasutatud 29.03.2021].
- [73] DaveF, „Overpass to return entities created > 2 years ago,“ 9.12.2016. [Võrgumaterjal]. Kättesaadav: <https://help.openstreetmap.org/questions/53391>. [Kasutatud 14.02.2021].
- [74] „Overpassi päring kõigi Harjumaa hoonete leidmiseks,“ [Võrgumaterjal]. Kättesaadav: <https://overpass-turbo.eu/s/16Wk>. [Kasutatud 3.05.2021].
- [75] Overpass-API, „Overpass-API olekuleht,“ [Võrgumaterjal]. Kättesaadav: <https://overpass-api.de/api/status>.

- [76] S. Gillies, „Fiona - Installation,“ 2011. [Võrgumaterjal]. Kättesaadav: <https://fiona.readthedocs.io/en/latest/README.html#installation>. [Kasutatud 4.03.2021].
- [77] Esri, „Install ArcPy,“ 22.11.2020. [Võrgumaterjal]. Kättesaadav: <https://pro.arcgis.com/en/pro-app/latest/arcpy/get-started/installing-arcpy.htm>. [Kasutatud 5.03.2021].
- [78] GeospatialPython, „PyShp,“ 14.01.2021. [Võrgumaterjal]. Kättesaadav: <https://github.com/GeospatialPython/pyshp>. [Kasutatud 10.03.2021].
- [79] O. M. Bjørndalen, „dbfRead,“ 22.12.2020. [Võrgumaterjal]. Kättesaadav: <https://github.com/olemb/dbfread>. [Kasutatud 3.03.2021].
- [80] "Verbatium", „Kasutaja "verbatium" muudatuskogumid,“ 9.08.2011. [Võrgumaterjal]. Kättesaadav: <https://www.openstreetmap.org/user/verbatium/history>. [Kasutatud 15.02.2021].
- [81] S. Mikhailov, „Verbatium buiding import discussion,“ 14.09.2019. [Võrgumaterjal]. Kättesaadav: <https://groups.google.com/g/openstreetmap-ee/c/4tXttib2vFg/m/eTqhwQRGAQAJ>. [Kasutatud 15.02.2021].
- [82] Berndw, „All in one Garmin Map/Garmin Types,“ 28.01.2013. [Võrgumaterjal]. Kättesaadav: https://wiki.openstreetmap.org/wiki/DE:All_in_one_Garmin_Map/Garmin_Types. [Kasutatud 10.02.2021].
- [83] Maa-amet, „Põhikaardistuse ajalugu,“ 27.11.2019. [Võrgumaterjal]. Kättesaadav: <https://geoportaal.maaamet.ee/est/Ruumiandmed/Topokaardid-ja-aluskaardid/Eesti-pohikaart-1-10000/Pohikaardistuse-ajalugu-p113.html>. [Kasutatud 21.03.2021].
- [84] "Verbatium", „Verbatiumi muudatuskogumite nimekiri XML-vormingus,“ 8.08.2011. [Võrgumaterjal]. Kättesaadav: <https://www.openstreetmap.org/api/0.6/changesets/?user=63077>. [Kasutatud 14.04.2021].

- [85] P. Neis, „How did you contribute to OpenStreetMap? - verbatim,“ 8.08.2011. [Võrgumaterjal]. Kättesaadav: <https://hdyc.neis-one.org/?verbatim>. [Kasutatud 13.04.2021].
- [86] "Verbatim", „Muudatuskogum: 572107 osmChange XML,“ 8.09.2008. [Võrgumaterjal]. Kättesaadav: <https://www.openstreetmap.org/api/0.6/changeset/572107>. [Kasutatud 25.04.2021].
- [87] V. Kuzmin, „<http://gis.tallinn.ee/kasiinod/>,“ 8.01.2010. [Võrgumaterjal]. Kättesaadav: https://groups.google.com/g/openstreetmap-ee/c/KCx7p_R3PYA. [Kasutatud 23.03.2021].
- [88] J. Laineste, „OSM: Buss/tramm/troll Tallinnas,“ 3.10.2011. [Võrgumaterjal]. Kättesaadav: https://groups.google.com/g/openstreetmap-ee/c/_fv9IhSwdhU. [Kasutatud 23.03.2021].
- [89] V. Kuzmin, „Kuzmini LinkedIni leht,“ [Võrgumaterjal]. Kättesaadav: <https://ee.linkedin.com/in/valeri-kuzmin-1b454949>. [Kasutatud 24.03.2021].
- [90] Päästeamet, „Sillamäe PK 3 kontaktid,“ [Võrgumaterjal]. Kättesaadav: https://www.rescue.ee/et/kontaktid/ida_paastekeskus/ida-virumaa_paastepiirkond/sillamae_pk_3/ametnik/valeri.kuzmin. [Kasutatud 24.03.2021].
- [91] Kartograff, „Kasutaja Kartograff kaastöö,“ 12.12.2008. [Võrgumaterjal]. Kättesaadav: <https://wiki.openstreetmap.org/wiki/Special:Contributions/Kartograff>. [Kasutatud 24.03.2021].
- [92] Maa-amet, „Maainfo kaardirakenduse kirjeldus,“ 16.02.2021. [Võrgumaterjal]. Kättesaadav: <https://geoportaal.maaamet.ee/est/Kaardirakendused/Maainfo-X-GIS-2/Maainfo-kaardirakenduse-kirjeldus-p106.html>. [Kasutatud 15.04.2021].
- [93] A. Kaaber, 01.12.2020. [Võrgumaterjal]. Kättesaadav: <https://ee.linkedin.com/in/andres-kaaber>. [Kasutatud 24.03.2021].

- [94] Äripäev, „Avatud Maakaardi Selts MTÜ - Juhatuse liikmed,“ Äripäeva infopank, [Võrgumaterjal]. Kättesaadav: <https://infopank.ee/ettevote/351764/avatud-maakaardi-selts-mtu>. [Kasutatud 21.04.2021].
- [95] Eesti Kaardikeskus, Tallinna atlas, Tallinn: Eesti Kaardikeskus, 2005.
- [96] Regio, Eesti teed, Tartu: Regio, 2008.
- [97] M. Konieczny, 14.04.2021. [Võrgumaterjal]. Kättesaadav: <https://discordapp.com/channels/413070382636072960/413070502580453387/831827930660995112>. [Kasutatud 14.04.2021].
- [98] M. Rämmel, „Muudatuskogum: 13479101,“ 13.10.2012. [Võrgumaterjal]. Kättesaadav: <https://www.openstreetmap.org/changeset/13479101>. [Kasutatud 22.04.2021].
- [99] M. Värtson, „Talk-ee: Küsimus Verbatiumi hooneimpordi teemal,“ 15.04.2021. [Võrgumaterjal]. Kättesaadav: https://groups.google.com/g/openstreetmap-ee/c/n4CjEd_sf8k/m/cPIeWbSICwAJ. [Kasutatud 15.04.2021].
- [100] OSMi kaasautorid, „OpenStreetMapi põhiline veebileht,“ 4.05.2021. [Võrgumaterjal]. Kättesaadav: <https://www.openstreetmap.org>. [Kasutatud 4.05.2021].
- [101] OSMi kaasautorid, „OpenStreetMapi arendusserveri veebileht,“ 4.05.2021. [Võrgumaterjal]. Kättesaadav: <https://api06.dev.openstreetmap.org>. [Kasutatud 4.05.2021].
- [102] OpenStreetMapi kaasautorid, „Using the dev server,“ 22.06.2016. [Võrgumaterjal]. Kättesaadav: https://wiki.openstreetmap.org/wiki/Using_the_dev_server. [Kasutatud 22.04.2021].
- [103] OpenStreetMap Foundation, „OSMF Server Info: snap-01.openstreetmap.org,“ 27.02.2021. [Võrgumaterjal]. Kättesaadav: <https://hardware.openstreetmap.org/servers/snap-01.openstreetmap.org/>. [Kasutatud 27.02.2021].

- [104] „Index of /planet/2021,“ 25.03.2021. [Võrgumaterjal]. Kättesaadav: <https://planet.osm.org/planet/2021>. [Kasutatud 25.03.2021].
- [105] Geofabrik, „Geofabrik internal download server: Estonia,“ 2.04.2021. [Võrgumaterjal]. Kättesaadav: <https://osm-internal.download.geofabrik.de/europe/estonia.html>. [Kasutatud 3.04.2021].
- [106] Canonical, „Ubuntu Server: Basic installation,“ 07.2020. [Võrgumaterjal]. Kättesaadav: <https://ubuntu.com/server/docs/installation>. [Kasutatud 1.05.2021].
- [107] Ubuntu Documentation Team, „Ubuntu Desktop Edition: Recommended Minimum System Requirements,“ 28.05.2020. [Võrgumaterjal]. Kättesaadav: <https://help.ubuntu.com/community/Installation/SystemRequirements>. [Kasutatud 2.05.2021].
- [108] A. J. Thor, „Prefixes for binary multiples,“ Metrologia, kd. 37, nr 1, p. 81, 1.02.2000.
- [109] „Switch2OSM: Serving Tiles,“ 2020. [Võrgumaterjal]. Kättesaadav: <https://switch2osm.org/serving-tiles>. [Kasutatud 24.02.2021].
- [110] T. Hughes, A. Allan ja M. Amos, „Openstreetmap website: INSTALL.md,“ 3.02.2021. [Võrgumaterjal]. Kättesaadav: <https://github.com/openstreetmap/openstreetmap-website/blob/master/INSTALL.md>. [Kasutatud 24.02.2021].
- [111] Vmware, „Configuring Network Address Translation,“ 31.05.2019. [Võrgumaterjal]. Kättesaadav: <https://docs.vmware.com/en/VMware-Workstation-Pro/16.0/com.vmware.ws.using.doc/GUID-89311E3D-CCA9-4ECC-AF5C-C52BE6A89A95.html>. [Kasutatud 21.04.2021].
- [112] T. Hughes, A. Allan ja M. Amos, „Openstreetmap website: CONFIGURE.md,“ 5.01.2021. [Võrgumaterjal]. Kättesaadav: <https://github.com/openstreetmap/openstreetmap-website/blob/master/CONFIGURE.md>. [Kasutatud 24.02.2021].
- [113] T. Jordans, „Openstreetmap website: Add seed data for users, diary entries and iD,“ 2.01.2021. [Võrgumaterjal]. Kättesaadav: <https://github.com/openstreetmap/openstreetmap-website/pull/3030/files>. [Kasutatud 24.02.2021].

- [114] „The Rails Command Line,“ [Võrgumaterjal]. Kättesaadav: https://guides.rubyonrails.org/command_line.html. [Kasutatud 24.02.2021].
- [115] M. Migurski, „Openstreetmap website: ERROR: column "nearby" of relation "users",“ 15.03.2020. [Võrgumaterjal]. Kättesaadav: <https://github.com/openstreetmap/openstreetmap-website/issues/2449>. [Kasutatud 24.02.2021].
- [116] B. Henderson ja M. Migurski, „Osmosis,“ 27.05.2020. [Võrgumaterjal]. Kättesaadav: <https://github.com/openstreetmap/osmosis>. [Kasutatud 3.04.2021].
- [117] T. Claburn, „Ruby off the Rails: Code library yanked over license blunder, sparks chaos for half a million projects,“ 29.03.2021. [Võrgumaterjal]. Kättesaadav: https://www.theregister.com/2021/03/29/rails_mime_fix. [Kasutatud 31.03.2021].
- [118] Q. Daley, „Dependency on mimemagic 0.3.x no longer valid,“ 24.03.2021. [Võrgumaterjal]. Kättesaadav: <https://github.com/rails/rails/issues/41750>. [Kasutatud 26.03.2021].
- [119] N. Blaschke, „3.8.0: rails s -b 0.0.0.0 will set port to default puma port,“ 27.03.2017. [Võrgumaterjal]. Kättesaadav: <https://github.com/puma/puma/issues/1255>. [Kasutatud 2.04.2021].
- [120] Maa-amet, „L-EST97 koordinaatsüsteem,“ 27.11.2019. [Võrgumaterjal]. Kättesaadav: <https://geoportaal.maaamet.ee/est/Ruumiandmed/Geodeetilised-andmed/Eesti-geodeetiline-susteem/L-EST97-koordinaatsusteem-p173.html>. [Kasutatud 1.03.2021].
- [121] MapTiler, „Estonian Coordinate System of 1997,“ 27.05.2005. [Võrgumaterjal]. Kättesaadav: <https://epsg.io/3301>. [Kasutatud 1.03.2021].
- [122] OpenStreetMap wiki kaasautorid, „Zoom levels,“ 27.08.2019. [Võrgumaterjal]. Kättesaadav: https://wiki.openstreetmap.org/wiki/Zoom_levels. [Kasutatud 2.03.2021].

- [123] H. Jürgenson ja T. Riips, „Lamberti konformse koonilise projektsiooni koordinaatide süteem ja arvutusvalemid,“ 1995. [Võrgumaterjal]. Kättesaadav: <https://www.maaamet.ee/rr/geo-lest/>. [Kasutatud 23.02.2021].
- [124] J. Kiho, Algoritmid ja andmestruktuurid, Tartu: Tartu Ülikooli kirjastus, 2003.
- [125] Tutorials Point, „Python - Hash Table,“ 2021. [Võrgumaterjal]. Kättesaadav: https://www.tutorialspoint.com/python_data_structure/python_hash_table.htm. [Kasutatud 12.05.2021].
- [126] M. Oviir, „SviMik_import is back to life,“ 8.08.2019. [Võrgumaterjal]. Kättesaadav: <https://groups.google.com/g/openstreetmap-ee/c/HnD85p3Wg9s/m/EQ4O89f1CgAJ>. [Kasutatud 20.04.2021].
- [127] J. Jaubert ja F. Rodrigo, „Osmose-QA validaator,“ 2021. [Võrgumaterjal]. Kättesaadav: <http://osmose.openstreetmap.fr/>. [Kasutatud 27.04.2021].
- [128] „Meililisti talk-ee arhiiv,“ 7.02.2021. [Võrgumaterjal]. Kättesaadav: <https://groups.google.com/g/openstreetmap-ee>. [Kasutatud 3.04.2021].
- [129] M. J. A. Eugster ja T. Schlesinger, „Osmar: OpenStreetMap and R,“ *R-Journal*, 2012.
- [130] A. Ots, „Maa-amet valmistub tootma Eesti kohta rohkem kolmemõõtmelisi ruumiandmeid,“ 8.03.2021. [Võrgumaterjal]. Kättesaadav: <https://www.maaamet.ee/et/uudised/maa-amet-valmistub-tootma-eesti-kohta-rohkem-kolmemootmelisi-ruumiandmeid>. [Kasutatud 4.04.2021].
- [131] F4; OSMi kaasautorid, „F4map,“ 30.04.2021. [Võrgumaterjal]. Kättesaadav: <https://demo.f4map.com>. [Kasutatud 30.04.2021].
- [132] „Overpassi päring maaamet:ETAK võtmete kasutusest,“ [Võrgumaterjal]. Kättesaadav: <https://overpass-turbo.eu/s/16Wd>.
- [133] OpenStreetMap wiki kaasautorid, „Key:created_by,“ 31.03.2021. [Võrgumaterjal]. Kättesaadav: https://wiki.openstreetmap.org/wiki/Key:created_by. [Kasutatud 22.04.2021].

- [134] V. Li, „Hacking Python Applications,“ 15.11.2019. [Võrgumaterjal].
Kättesaadav: <https://medium.com/swlh/hacking-python-applications-5d4cd541b3f1>. [Kasutatud 12.05.2021].
- [135] M. Pilgrim, „Serializing Python Objects,“ %1 *Dive Into Python 3*, 2009, pp. 314-338.
- [136] F. Ramm ja P. Norman, „Reduce maximum changeset size to 10k changes,“ 1.08.2016. [Võrgumaterjal]. Kättesaadav: <https://github.com/openstreetmap/openstreetmap-website/pull/1259>. [Kasutatud 14.04.2021].
- [137] OpenStreetMap wiki kaasautorid, „Overpass QL,“ 1.09.2020. [Võrgumaterjal].
Kättesaadav: https://wiki.openstreetmap.org/wiki/Overpass_API/Overpass_QL.
[Kasutatud 25.04.2021].
- [138] „Overpassi päring impordiga lisatud hoonete loendamiseks,“ 4.05.2021.
[Võrgumaterjal]. Kättesaadav: <https://overpass-turbo.eu/s/16Wp>. [Kasutatud 4.05.2021].

Lisa 1 – Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks¹

Mina, Kalle Johannes Reinmets

1. Annan Tallinna Tehnikaülikoolile tasuta loa (lihtlitsentsi) enda loodud teose „Hoonete geomeetria automaatne uuendamine OpenStreetMapi näitel“, mille juhendaja on Siim Vene
 - 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 tähtjaks nõusolekut lõputöö reprodutseerimiseks ja avalikustamiseks vastavalt lihtlitsentsi punktidele 1.1. ja 1.2, siis lihtlitsents nimetatud tähtaja jooksul ei kehti.

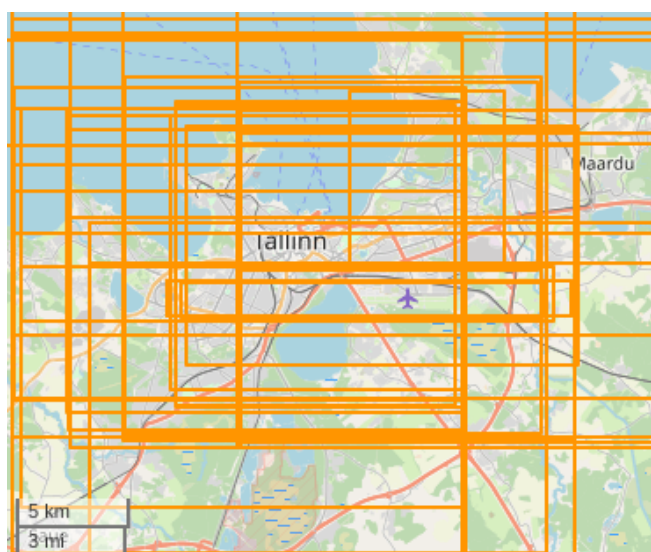
Lisa 2 – Lõpliku Overpassi päringu näide

Näide töö käigus valminud päringust, mille saab kirjutada nt `overpass-turbo.eu` veebilehele, et leida Tallinna kesklinnast hooneid, mille ühtegi sõlme ei ole peale 24. veebruari 2011 kordagi muudetud. Tegelikus skriptis kasutatakse küll sarnast loogikat, aga muutujate automaatseks asendamiseks valiti kehvemini jälgitav vormistus. Päringus kasutatakse ingliskeelseid kommentaare, sest päringu link edastati peatükis 3.1 kirjeldatud vestluste käigus kaardistusettevõtetele. Kõige olulisem Overpass QL süntaksi eripära järgevas päringus < ja > sümbolid. Neid kasutatakse, et leida vastavalt kõiki jooni, mis läbivad mingis hulgas olevaid sõlmi, ning kõiki sõlmi, mida etteantud joon sisaldab [137].

```
[out:json][timeout:25];
{{geocodeArea:Kesklinn}}->.searchArea;
// Find all buildings in area
((way["building"](area.searchArea);
// and save them and their nodes into set .all
);>)->.all;
// Find nodes within .all, that are newer than 10yrs and
(node.all(newer:"2011-02-24T12:00:00.000Z"));
// ...save WAYS containing them into set .new.
<)->.new;
// Subtract new set .new from .all and get nodes from that set
(((way.all; - way.new;);>);>);
// Then save ways going through any of these nodes into set test1
<)->.test1;
// Extract non-buildings from test1 ...
way.test1[!"building"];
// ...along all other ways sharing any common node with them...
>;<
// And save them into set test2
->.test2;
// Subtract ways of test2 from ways of test1
(way.test1; - way.test2;);
// Query all nodes of these ways
(._;>);
// Return output json with metadata.
out meta;
```

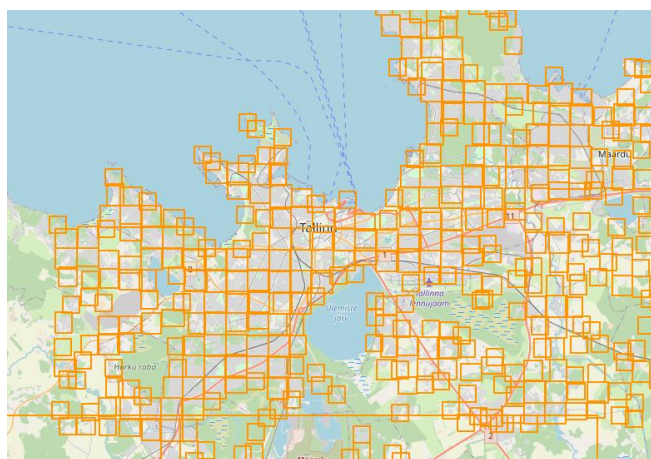
Lisa 3 – Kuvatõmmised kolmest erinevast muudatuskogumite jaotusest

Käesolevas lisa on toodud kolm kuvatõmmist, mis illustreerivad erinevaid peatükis 4.2.2 kirjeldatud muudatuskogumite jaotusvariante. Esimene (Joonis 3-1) näitab olukorda enne ühikruutude kasutuselevõttu, kui hooneid töödeldi lähtuvalt nende OSMi identifikaatorite järjestusest, ehk hooneid muudeti algupärasest kaardile kandmise järjekorras.



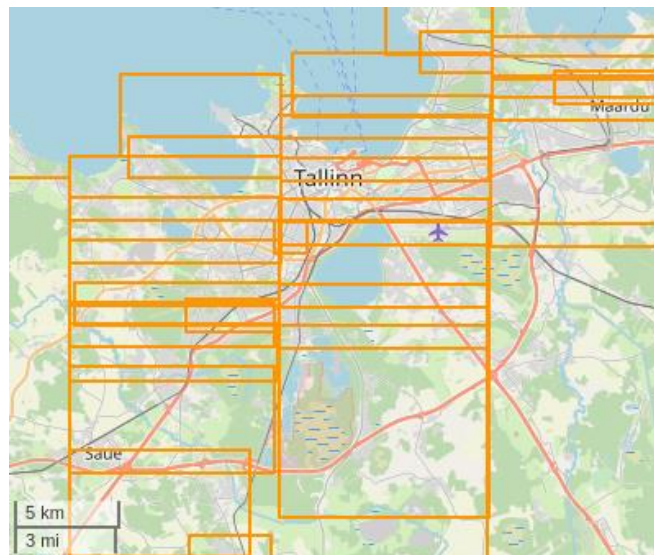
Joonis 3-1. Olukord enne ühikruutude kasutuselevõttu.

Teise kuvatõmmise (Joonis 3-2) saamiseks piirati iga 1x1 km ühikruudu kõik muudatused eraldiseisvasse muudatuskogumisse. Teine kuvatõmmis on teistest veidi erinev, sest kasutatakse suuremat skaalat, et anda paremini edasi muudatuskogumite külgnevuskastide ühtlast suurust.



Joonis 3-2. Muudatused jaotatuna ühikruutudesse

Kolmas ehk lõplik variant (Joonis 3-3) kombineerib esimese kahe parimad omadused – saavutati mõistlikult väikesed muudatuskogumid hoides samal ajal tehtud muudatuste hulga maksimaalsena. Ühe muudatuskogumi laius oli maksimaalselt 10 km.



Joonis 3-3. Lõplik muudatuskogumite jaotumine