

TALLINNA TEHNIKAÜLIKOOL

Infotehnoloogia teaduskond

Maili Kark 1829511AAM

**VALIMISANDMETE VEEBIARHIIV.
ANDMEMUDEL JA AVAANDMETE
STRUKTUURID RIIGI VALIMISTEENISTUSE
VALIMISINFO PÕHJAL**

Magistritöö

Juhendaja: Priit Rospel
MA

Tallinn 2020

Autori deklaratsioon

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

Autor: Maili Kark

[17.05.2020]

Annotatsioon

Käesoleva magistritöö “Valimisandmete veebiarhiiv. Andmemudel ja avaandmete struktuurid Riigi valimisteestuse valimisinfo põhjal” eesmärgiks on leida ühtne struktuur 1992-2019. aastal toimunud üleriigiliste valimiste andmestikule, et parandada andmete kättesaadavust. Töö käigus viiakse läbi andmete detailanalüüs, mille tulemina koostatakse andmemudel tulevase valimisandmete arhiivi andmebaasi loomiseks. Analüüsis esitatakse uuringu tulemusel järeldused, milliseid andmeid on võimalik esitada ning kuidas neid struktureerida, täpsustades tulevasele andmebaasile seni esitatud ärioloogilisi nõudeid. Detailanalüüs hõlmab seadusest tulenevaid nõudeid andmetele, analüüsi nende muutumisest lähtuvalt ning klassimudeleid leitud andmehulkade ja seostega. Andmemudel esitatakse töö kavandi osas. Andmete kättesaadavuse teise aspektina on töö käigus leitud valimisliigiti ühtsed XML-struktuurid valimisandmestiku avaldamiseks avaandmetena.

Lõputöö on kirjutatud eesti keeles ning sisaldab teksti 63 leheküljel, 4 peatükki, 13 joonist, 2 tabelit.

Abstract

Election Data Web Archive. Data Model and Open Data Structures on Election Data from The State Electoral Office

This Master's thesis is concerned with the usability of the data from previous elections on the State Electoral Office's webpage. At the moment, the data can be accessed through HTML-webpages, which differ in their design and structure. Open data is only available about election results from 2014-2019. These circumstances make it difficult for researchers or scientists to analyse or study data concerning elections, and therefore cause extra work for the State Electoral Office to provide information for incoming requests. Overall statistics about elections that is produced by the State Electoral Office are available, but it has proved to be too general for specific interests. The thesis only includes data from nationwide elections from 1992-2019, including Parliament elections, European Parliament elections, local government councils elections and referendums.

The aim of the thesis is to improve the usability of the nationwide elections data by creating election data web archive in the future. According to previous business analyse on this matter, the data on these webpages was analysed in detail to find the patterns, gaps and structures in this information, forming class models to present necessary attributes and relations. This enables to make the requirements for the future database more specific and to structure the data in substantive and logical manner, setting data in definite, recurrent patterns considering different types of elections.

Resulting from the analysis, a entity relationship diagram was created to accommodate all the recurrent election data. The migration of data from the webpages to a new database, based on this data model, is depicted on the component diagram, involving a data scraping tool. The structures for open data on all the published election information are also designed and presented with XML-file examples.

The thesis is in Estonian and contains 63 pages of text, 4 chapters, 13 figures, 2 tables.

Lühendite ja mõistete sõnastik

Mõisted

Hääletaja - hääle andnud valija

RVT veebileht - valimised.ee leht. Vastavatesse jaotistesse on lingitud toimunud valimiste veebilehed, üldstatistika ja analüüs ning avaandmete allalaadimise leht.

Kolme täрни аваандмед - Tim Berners-Lee loodud viie täрни süsteemi kohaselt tähendab see veebis kättesaadavat, masinloetavat, struktureeritud ning avatud formaadiga andmestikku.

Valija - hääletamisõiguslik isik (reguleeritud RK, EP, KOV ja EH valimisseadustega)

Valimissündmus - määratud aastal toimuv valimine. Siin: üleriigiliselt korraldatud RK, EP, KOV valimine või RH

Valimiste kood - Valimisliigi kokkulepitud lühendist ja toimumise aastaarvust koosnev kood tekstis. Nt RK2015 - 2015. aastal toimunud Riigikogu valimine

Valimiste veebileht - Igaks valimissündmuseks luuakse valimissündmuse koodi sisaldav veebileht. Näiteks: <https://kov2017.valimised.ee/> . Vanemate valimissündmuste korral on valimiste veebilehe URL ka teises formaadis, nt kujul rh03 - 2003. aasta rahvahääletus.

Välisääл - alaliselt välismaal elava valija antud ääl. Välismaal elavate valijate ääли kogutakse ainult RK, EP ja RH valimisliigi korral.

Lühendid

EHS - E-hääletussüsteem

EP - Euroopa Parlament

KOV - Kohalik omavalitsus

RK - Riigikogu

VARH - planeeritav valimisandmete veebiarhiiv, andmeid kuvav kasutajaliides ning sellega ühenduses olev andmebaas

VIS - Valimiste infosüsteem

VP - Valimispäev

Sisukord

Autori deklaratsioon	1
Jooniste loetelu	7
Sissejuhatus	8
1. Ülesandepüstitus ja taust	11
1.1. Probleemi kirjeldus ja töö eesmärk	11
1.2. Ülevaade valdkonnast	13
1.2.1. Valimiste korraldamine	14
1.3. Andmete koosseis ja as-is protsess	17
1.4. Varasemad uuringud ja nõuded valimisandmete veebiarhiivile	22
1.5. Olemasolevad juhendid avaandmetele	25
1.6. Uurimismeetod	26
2. Andmehulkade analüüs	29
2.1. Valimisjaoskonnad ja -komisjonid	32
2.2. Kandidaadid	37
2.3. Hääletamisest osavõtu statistika	45
2.4. Hääletamis- ja valimistulemus	49
2.4.1. Hääletamis- ja valimistulemus erakonniti	50
2.4.2. Valituks osutunud, asendusliikmed	55
2.4.3. Mandaatide omandamise viis	55
2.5. Detailne hääletamistulemus	57
2.5.1. Häälte jagunemine	57
2.6. Üldised nõuded analüüsist lähtuvalt	63
3. Kavand	64
3.1. Andmemudel olemi-suhte diagrammina	64
3.2. Andmete migreerimine HTML-lehtedelt andmebaasi	66
3.3. Andmestruktuurid avaandmetele	68
3.3.1. Jaoskonnad	69
3.3.2. Kandidaadid	69
3.3.3. Hääletamisest osavõtt	70
3.3.4. Hääletamis- ja valimistulemus	70
3.3.5. Detailne hääletamistulemus	71
4. Kokkuvõte	73
Kasutatud kirjandus	75
Lisa 1 – Olemi-suhte diagrammi olemite kirjeldused	79

Lisa 2 – XML faili struktuur haldusüksuste ja jaoskondadega RK valimissündmusel	91
Lisa 3 – XML faili struktuur ühe valimissündmuse kandidaatidega KOV valimissündmusel	93
Lisa 4 – XML faili struktuur kõikide kandidaatidega	95
Lisa 5 – XML faili struktuur detailse hääletamisest osavõtu infoga RK valimissündmusel	97
Lisa 6 – XML faili struktuur hääletamis- ja valimistulemusega - RK valimissündmus	98
Lisa 7 – XML faili struktuur hääletamis- ja valimistulemusega RH valimissündmusel	100
Lisa 8 – XML faili stuktuur detailse hääletamistulemusega RK valimissündmusel	102

Jooniste loetelu

Joonis 1. Näited veebilehtedest: 1992., 2003., 2011. ja 2019. aasta RK valimistulemustest erakondade kaupa

Joonis 2. Valimisperioodi ajaline graafik 2019. aasta Euroopa Parlamendi valimiste näitel

Joonis 3. Andmete tekkimise põhiprotsess viimaste valimiste näitel

Joonis 4. Andmehulkadeks struktureeritud valimisinfo

Joonis 5. Ringkonna asukoht haldusüksuste struktuuris eri valimisliikide korral

Joonis 6. Valimisüksuste, ringkondade ja jaoskondade andmetega klassimudel

Joonis 7. Näited RK2003 ja KOV2005 kandidaadi andmetega

Joonis 8. Klassimudel nimekirjade, kandidaatide andmete ja seostega

Joonis 9. Klassimudel osavõtuandmete kuvamiseks vajalike tabelite ja seostega

Joonis 10. Klassimudel valimistulemuse ja mandaatide jaotuse andmehulkadega

Joonis 11. Klassimudel detailse hääletamisandmestiku esitamiseks

Joonis 12. Komponentmudel andmete “kraapimise” rakendusega

Joonis 13. Valimisandmete olemi-suhte diagramm

Tabelite loetelu

Tabel 1. Valimisandmete kasutajagrupid ja infovajadus

Tabel 2: Toimunud üleriigilised valimised toimumisaja ja liikide järgi

Sissejuhatus

Käesoleva magistritöö eesmärgiks on parandada Riigi valimisteenistuse veebilehel kuvatud 1992-2019. aastate valimissündmuste andmete kättesaadavust. Praegu asuvad vaadeldava perioodi andmed omaaegsetel HTML-veebilehtedel. Valimiste puhul on avalikkuse infovajadus eriti terav valimispäevale eelneva ja järgneva ajaperioodi jooksul, milleks enamasti on paar kuud ning ka olemasolevad veebilehed on loodud laiale kasutajate ringile ühe valimisündmuse raames optimaalse, järk-järgult lisanduva info andmiseks. Selle ajaperioodi - 17 aasta jooksul - on muutunud nii veebilehtede disain kui andmehulkade kuvamise loogika. Uurijatele, kes soovivad juba möödunud valimissündmuste andmeid võrrelda liigiti või kindla uurimisfookusega, tähendab see lisatööd andmete leidmisel ja töötlemisel ning Riigi valimisteenistusele ajaressurssi teabenõuetele vastamisel. Riigi valimisteenistuse lehel on olemas valimissündmuste üldiste andmete võrdlus aastate lõikes, kuid nagu võib järeldada saabunud teabenõuetest, on paljude uurijate huvi märksa detailsem või esitatust erineva fookusega. Avaliku sektori andmete avaldamine selget ja ühtlaselt kuvatuna ning avaldamine avaandmetena on järjest enam liitujaid koguv protsess, mille tähtsaima eesmärgina võib näha demokraatlike protsesside läbipaistvuse suurendamist.

Töö tulemuseks on tulevikus loodava valimiste arhiivandmete (VARH) lehel kuvatavate andmete struktureerimine, mis tugineb vaadeldava perioodi üleriigiliste valimissündmuste andmete detailanalüüsile. Analüüsi järelduste põhjal valmib andmemudel, mille abil saab luua vajalikke andmeid mahutava andmebaasi, kuhu saab edaspidi sisestada ka tulevaste valimissündmuste andmed. Andmete migreerimine veebilehtedelt andmebaasi on esitatud komponentmudelina. Kuigi ka praegu on võimalik alla laadida 2014-2019. aastate valimissündmuste tulemusi, luuakse töö käigus avaandmete struktuurid seni avaldamata andmete kohta (haldusüksuste struktuur ja jaoskonnad, kandidaadid) ning ühtlustatakse olemasolevaid.

Esimeses osas tutvustatakse probleemi tausta ning seni tehtud uuringuid, info hulka ja sisu määravat seadusandlust. Tutvustatakse andmete tekkimise praegust (*as-is*) protsessi

ning andmete vastutavaks töötlejaks olevaid institutsioone üldisemalt. Eesti avalikule sektorile andmemudelite loomiseks ning avaandmete väljastamiseks olevatest standarditest antakse ülevaade.

Teises osas võetakse ette valimisanandmete detailanalüüs. Analüüsimeetodiks on info uurimine kolmes aspektis: ajalises, valimisliigilises ning andmete sisulises plaanis ning leitud seosed esitatakse graafiliselt klassimudelitena. Varem esitatud nõuded täpsustatakse uute tehniliste võimaluste kontekstis ning määratletakse konkreetseid andmehulgad, mida on vajalik või võimalik avaldada.

Kolmandas osas esitatakse analüüsile ning seal leitud nõuetele tuginev kavand - andmemudel olemi-suhte diagrammina koos olemite kirjeldusega. Antakse, ülevaade võimalustest andmed veebilehtedelt andmebaasi liigutada ning komponentmudel andmete migreerimiseks koos rakendusega andmete kontrolliks. Luuakse andmete failistruktuurid avaandmete jaoks, mis esitavad andmeid analüüsis leitud seostes ning hulkades, täpsustatakse erinevaid andmehulki eri valimisliikide korral.

Magistritöö tulemuseks olev analüüs ja kavand on aluseks loodavale VARH (Valimiste arhiiv) andmebaasile, mille kasutajaliides kuvab valimisliigiti andmeid ühtlaselt ning võrreldaval moel. Iga andmehulka, mis analüüsi käigus on leitud ja struktureeritud, saab masinloetaval kujul alla laadida. See loob uurijatele võimaluse esiteks tutvuda kõigi valimissündmuse aluseks või tulemuseks olnud andmetega ning teiseks teha nende pinnalt uusi järeldusi.

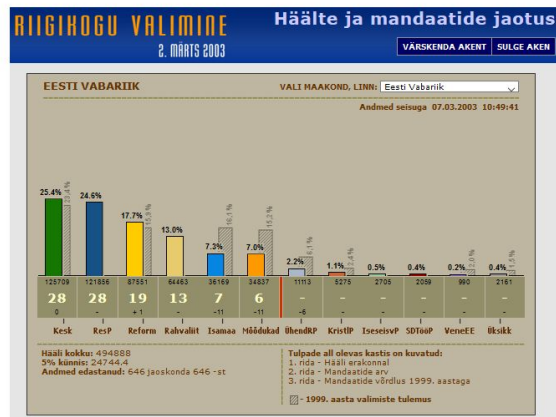
1. Ülesandepüstitus ja taust

Selles peatükis avatakse magistritöö aluseks oleva probleemi olemust ning valimiste korraldamist üldisemalt. Eraldi tuuakse esile valimiste korraldamisega seotud institutsioonide tegevused ja kohustused valimisandmete tekkimise juures ning antakse ülevaade seadustest ja määrustest, mis valimisandmete koosseisu määravad.

1.1. Probleemi kirjeldus ja töö eesmärk

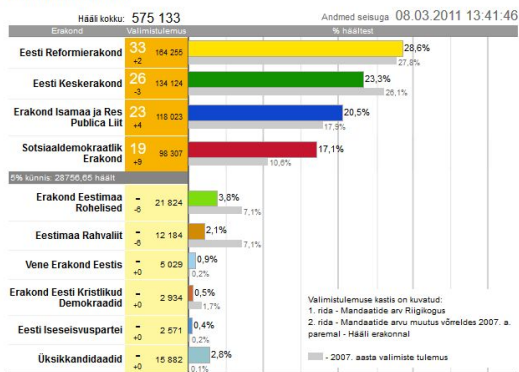
RVT veebilehe arhiivi jaotises [1] on avaldatud valimistulemuste andmestik 1992-2019 toimunud üleriigilistest valimistest. Praegusel kujul koosneb andmestik iga sündmuse kohta tolleaegsetest HTML lehtede kogumist, mis hõlmab ligi 20 erineva kujundusega veebilehte (joonis 1). Erandiks on 1992. aasta rahvahääletus, mille tulemuste kohta eraldi valimiste veebileht puudub, kuid mille andmed on RVT veebilehel esitatud eraldi [2].

Praegusel kujul olemasoleva valimisandmete arhiivi probleemiks on andmete leidmise keerukus. Andmeid ei ole võimalik hõlpsasti võrrelda ega analüüsida. Masinloetaval kujul on hetkel kättesaadavad ainult 2014-2019 aasta valimissündmuste tulemused [3]. See teeb keerukaks valimisi analüüsivate teadlaste, üliõpilaste, uurijate ja meedia töö, kes peavad andmete võrdlemiseks neid HTML lehtedelt otsima. Puudub võimalus kõiki esitatud andmeid töödeldaval kujul alla laadida. RVT veebilehel on olemas RVT koostatud ülevaatlid ning graafiliselt esitatud statistika, mis hõlmab näiteks osavõtu statistikat, kandidaatide arve ning nende keskmist vanust [4], e-hääletanute statistikat [5], samas on pdf-failina kättesaadav ja allalaetav ka põhjalik statistikakogumik "Valimised Eestis 1992-2015" [6], mis sisaldab andmeid ja ülevaateid selle perioodi valimistest. Mitmetel uurijatel on andmete osas aga oma tööst lähtuvalt spetsiifilisemad huvid, mida need ülevaatlidud statistikaandmed ei rahulda.



Valimistulemus

Eesti Vabariik kokku



Hääletamis- ja valimistulemus

Eesti Vabariik

Andmed seitsuga 8.03.2019 15:59



Joonis 1. Näited veebilehtedest: 1992., 2003., 2011. ja 2019. aasta RK valimistulemustest erakondade kaupa

2015. aastal on Eneli Pung koostanud RVT tellimisel lähteülesande eespool käsitletud probleemide kaardistamiseks ning võimaliku lahenduse leidmiseks [7]. Lähteülesandes on eesmärgina kirjeldatud valimisandmete ühtsele struktuurile viimist ning selleks digitaalse, veebis avaldatud andmebaasi loomine ning andmete avalikustamist - kokkuvõtlikult VARHi (valimisandmete arhiivi) loomist. Teel VARHini on lähteülesande kohaselt esimesteks ülesanneteks valimisandmete detailanalüüs ning selle põhjal andmemudeli loomine, mis oleks tulevase andmebaasi aluseks [7]. Käesolevas töös võetakse ette andmete detailanalüüs ning luuakse tehniline andmemudel. Valimiste veebilehtedel olevate andmete analüüsimise kõrval võetakse arvesse seni laekunud andmepäringuid [8], et võimalusel arvestada juba andmestruktuuride loomisel erinevate

kasutajagruppide huvidega. Andmete avalikustamise teemal analüüsitakse juba kättesaadavate - 2014-2019 aastate valimistulemuste - failistruktuure, et vaadata, kas vanemad andmed sobivad samuti neisse olemasolevatesse struktuuridesse. Vajadusel luuakse andmete avaldamiseks uued failistruktuurid. Lähteülesandes kirjeldatud valimisliikide (Riigikogu valimised, kohaliku omavalitsuse volikogu valimised, Euroopa Parlamendi valimised, Vabariigi Presidendi valimised, Riigikogu juhatuse valimised, Rahvahääletused) vaadeldakse siin ainult RK, KOV, EP ja RH valimisi, kuna sel puhul on tegemist üleriigiliste valimistega, kus andmehulgad ja koosseisud on omavahel võrreldavad. Vabariigi Presidendi ning Riigikogu juhatuse valimised on kajastatud ühtsel kujul ka juba RVT veebilehel [9][10].

Skoobist jäävad välja andmebaasi uute andmete sisestamise konkreetse vahendi määramine ning avaandmete kasutajaliidese loomine. Viimase osas on valimiste veebilehel juba välja kujunenud praktika võimaldada allalaadimist vastavalt veebilehelt zip-laiendiga failidena, selleks eraldi kasutajaliidese loomiseks puudub hetkel vajadus. Töös ei käsitleta ka andmete tekkimise ega liikumise turvalisuse küsimusi ega mittefunktsionaalseid nõudeid. Kuna töö teemaks on juba avaldatud valimisandmete struktureerimine, siis märgitakse valimistulemuse arvutamisse puutuv ära ainult seal, kus see on andmete organiseerimise puhul vajalik. Valimistulemuse arvutamisest annab ülevaate RVT veebileht [11].

Kokkuvõttes on käesoleva uurimistöö eesmärgiks VARHi loomiseks vajalike esmaste tööde läbiviimine, mis edaspidi võib olla sisendiks andmemudelil põhineva andmebaasi ning kasutajaliidese loomisele.

1.2. Ülevaade valdkonnast

Selles peatükis antakse lühike ülevaade vaatlusaluste valimisandmete loomise ja edastamisega tegelevast Riigi valimisteenistusest ning Vabariigi Valimiskomisjonist, nende struktuurist, ülesannetest ning kohustustest. Edasi kirjeldatakse seda, kuidas saadakse andmed vaatlusaluste valimissündmuste HTML-failidele. See protsess on olnud 1992. aastast alates vastavuses tolaegsete tehnoloogiliste võimalustega. Alates

1999. aastast on valimisandmete allikaks olnud valimiste infosüsteem VIS, mille ülesehitusest ja funktsioonidest, samuti valimisandmete tekkimisest antakse põgus ülevaade.

1.2.1. Valimiste korraldamine

Valimiste korraldamises on eristatavad riiklik ning kohalik tase. Riikliku taseme korraldajateks on Vabariigi valimiskomisjon ning Riigi valimisteenistus. Kohalikul tasemel korraldavad valimisi valla/linna (KOV) komisjonid ning neile alluvad jaoskonnakomisjonid.

Valimiste korraldamine toimub vastavuses valimisseadustega. Vaadeldava perioodi jooksul on neid seadusi pidevalt muudetud ja täpsustatud. Käsitletavas töös märgitakse ära ainult need muudatused, mis andmete esitamist otseselt puudutavad.

1.2.1.1. Riigi valimisteenistus

Riigi valimisteenistus on Riigikogu Kantselei struktuuri kuuluv üksus, mille ülesandeks on Riigikogu, Euroopa Parlamendi, kohaliku omavalitsuse volikogu valimiste ning rahvahääletuste seadusekohane korraldamine. Valimisteenistus abistab ka Vabariigi Presidendi ning Riigikogu esimehe ja aseesimeeste valimistel [12].

Riigi valimisteenistuse ülesanded eri valimisliikide korral on kirjeldatud vastavates valimisseadustest [13a]-[16a]. Põhiülesandeks on valimiste seadusekohase korraldamise tagamine - hääletamise ettevalmistamine Eestis ja välisriikides, hääletamisvahendite hankimine ning valimiste korraldajate koolitamine [12]. RVT avaldab enne valimisi kandideerimisdokumentide näidised ning võtab vastu kandideerimisdokumendid RK ja EP valimistel [13][15]. Valimisteenistus valvab piirkondlike valimiste korraldajate tegevuse üle hääletamise korraldamisel ning registreerib esitatud avaldused leitud puuduste kohta [17].

Hääletustulemuste kogumisel kasutab RVT kahte süsteemi: üheks on elektroonilise hääletamise süsteem (EHS) ning teiseks valimiste infosüsteem (VIS), mille mõlema

haldamine kuulub RVT pädevusse. Enne valimisi valmistab RVT e-hääletussüsteemi ning valimiste infosüsteemi valimisteks ette ning teeb hiljem kindlaks e-hääletamise tulemuse [17] ning nii e- kui paberhääli ühendava hääletamis- ja valimistulemuse.

1.2.1.2. Vabariigi Valimiskomisjon

Järelevalvet Riigi valimisteenistuse seadusest tulenevate ülesannete kohta korraldab Vabariigi Valimiskomisjon, kuhu kuuluvad Riigikogu valimise seaduse alusel määratud liikmed. Valimiskomisjoni ülesanneteks on valimiste aluspõhimõtete (vabaduse, üldisuse, ühetaolisuse ja otsesuse, hääletamise salajasuse) täitmise jälgimine nii paber- kui e-hääletamise korral, kandidaatide registreerimine ning pärast valimisi hääletamis- ja valimistulemuste kindlakstegemine. Valimiskomisjon korraldab Vabariigi Presidendi valimise. Samuti kuuluvad Vabariigi Valimiskomisjoni ülesannete hulka valimisteemaliste kaebuste läbivaatamine ning vajadusel hääletustulemuste tühistamine [18]. Andmete tekkimise seisukohast on oluline, et Vabariigi Valimiste infosüsteemi põhimääruse kohaselt on Vabariigi Valimiskomisjon infosüsteemi vastutav töötleja [19]. Valimiskomisjoni liikmed on VISis määratud riikliku tasemega valimiste korraldajateks, kes võivad valimisperioodil kõiki sisestatud andmeid kontrollida.

1.2.1.3. Valimiste infosüsteem - VIS

VIS on elektrooniline töökeskkond valimiste korraldajatele, mille esimene versioon loodi 1999. aasta Riigikogu valimisteks. Enne VISi kasutuselevõttu edastati piirkondlikud hääletamisprotokollid RVTle erinevaid, varem kokku lepitud kanaleid pidi - kasutati faksi ning andmete toomist disketil. Seejärel sisestati protokollide andmed 1990. aastatel sisestajate poolt vastavasse programmi, mis summeeris protokollid ning arvutas valimistulemused. Valimiste infosüsteemi esimese versiooni (sel ajal VALIS-projekti) peamised töö-protseduurid olid valimiskomisjonide, erakondade ja kandidaatide andmete kogumine ja esitamine, valimisprotokollide täitmise võimaldamine ning hääletustulemuste veebilehekülgede genereerimine, vahetulemuste edastamine meediale, valimistulemuse väljaselgitamine ja esitamine ning valimiste statistiliste koondite koostamine [20].

Sellest ajast on VIS läbinud mitmeid muudatusi ja uuendusi, kuid põhifunktsioonid on jäänud samaks. Praegu on töös VISi kolmanda versiooni arendused, mis peavad võimaldama 2018. aastal vastu võetud ning 01.01.2021 jõustuvaid valimisseaduse muudatusi. Uuendatud seadused sätestavad näiteks elektroonilise valijate nimekirja kasutamise, hääletusperioodi muudatused ja kandideerimisdokumentide digitaalse esitamise võimaluse [13b]-[16b] Käesoleva töö skooopi jäävad aga üksnes 1992-2019 aastate valimisandmed.

Enne igat valimist on VISis tehtud väiksemaid või suuremaid muudatusi, mille on tinginud vajadus kohandada infosüsteem vastavaks valimiste vahelisel ajal vastu võetud seadusemuudatustele. Haldusreformi käigus 2017. aastal liideti valdu, enne 2019. aasta valimisi kaotati maakondade valimiskomisjonid ning loodi KOV komisjonid. Muudatusi on tehtud ka kasutajate vajadustest ning soovidest lähtuvalt mõne töö hõlbustamiseks või huviväärse info pakkumiseks. 2005. aastast, mil Eestis sai võimalikuks e-valimine, on VISil funktsionaalsus e-häälte arvestamiseks ning e-hääletajate arvestamiseks - viimane on vajalik selleks, et tühistada nende e-hääletajate hääled, kes on andnud ka paber-hääle.

2013. aastal vastu võetud Valimiste infosüsteemi põhimäärus sätestab infosüsteemi ülesanded järgmiselt:

1. vahendada andmeid valimiskomisjonide vahel;
2. töödelda kandideerimisandmeid ja hääletamis- ning valimistulemusi;
3. tagada ülevaade valimiste käigust;
4. võimaldada valimisandmete elektroonilist esitamist;
5. valmistada valimiskomisjonide, kandideerimise ja hääletamis- ning valimistulemuste andmed ette avalikustamiseks [19, §2]

VISi kantud andmetel on määruse kohaselt informatiivne tähendus [19, §4]. See tähendab, et õiguslikku jõudu on siiani omanud vastava komisjoni poolt allkirjastatud paber-protokollid. Sellega on tasalülitatud risk, kui VISi töös peaks valimiste perioodil esinema tõrkeid või vigu. Tegevuskäik selleks puhuks antakse valimiste korraldajatele

teada kirjalikes juhendites ning RVT läbiviidavatel koolitustel. VIS infosüsteemina on aga andnud kiiret ning süstematiseeritud infot valimiste kulgemise ning tulemuste kohta nii valimiste korraldajatele kui avalikkusele juba valimispäeva õhtuks.

Määrus sätestab ka infosüsteemi turvaklassid ja turbeastmed, andmete kaitse, andmete väljastamise korra ning muud üksikasjad VISi toimimise kohta [19].

VISi kasutusõiguse võib määruse kohaselt anda valimiskomisjoni liikmele, valimiskomisjoni otsusel komisjoni abistavale isikule, Vabariigi Valimiskomisjoni sekretariaadi liikmele, välisesinduses hääletamist korraldavale isikule. Kasutajad registreeritakse ning nad saavad VISis oma õigustele ja kohustustele vastava rolli. VISi pääseb reeglina sisse ainult isiku digitaalset tuvastamist võimaldava sertifikaadiga [19].

1.3. Andmete koosseis ja *as-is* protsess

Kõik valimiste veebilehtedel olevad andmed tulevad VISist ning selles peatükis antakse andmete tekkimisest ja liikumisest ülevaade. 2013. aastal vastu võetud määrus sätestab, mis andmed VISi sisestatakse. Andmed sisestatakse järgmiste üksuste kohta:

1. Valimiskomisjon
2. Valimiskomisjoni liige
3. Valimisjaoskond
4. Erakond või valimisliit
5. Erakonna või valimisliidu esindaja
6. Valimisringkond
7. Kandidaat
8. Välisesindus
9. Välisesinduses hääletamist korraldav isik
10. Elektroonilise hääletamise korraldamiseks valija info
11. Hääletamisest osavõtu info
12. Hääletamistulemused vastavalt Vabariigi Valimiskomisjoni määrusega kehtestatud protokollile vormile

13. Valimistulemused vastavalt Vabariigi Valimiskomisjoni määrusega kehtestatud protokollile vormile

Mis andmed täpsemalt iga punkti kohta sisestatakse, on loetletud määruses [19, §6]. Neist punktidest 1.-9. sisestatakse enne valimisi ning see info on seotud valimiste korraldusliku poolega. Veebilehtedele on enne valimisi avalikkuse teavitamiseks kuvatud info ringkondade, jaoskondade (ning jaoskondade esimeeste) ning valimistel osalevate erakondade, valimisliitude (KOV valimiste korral) ning kandidaatide kohta. Seda, mis andmekogumid täpsemalt valimiste eel või järgselt veebilehtedele kuvatakse, on aja jooksul muudetud. 2017. aastal kasutati näiteks jaoskondade info kuvamiseks Rahvastikuregistri kaardirakendust, mis praeguseks on suletud.

Valimisnädalal lisanduvad eelnevale infole statistilised andmed hääletajate kohta. 2013. aastal võeti kasutusele e-hääletanute lugeja, mis valimiste veebilehel kuvab e-hääletajate arvu reaalselt. Pabersedeliga hääletajate arvu kohta annavad jaoskonnad RVTle VISi kaudu infot valimisnädala neil päevadel, mil hääletamine toimub, kell 20:00, valimispäeval aga kolm korda päevas - kell 12:00, 16:00 ja 20:00 (joonis 2). Jaoskondades sisestatakse andmed ka väljaspool elukohta hääletanute kohta. Kuna siiani on kasutusel olnud paberil valijate nimekirjad, siis imporditakse e-valimisperioodi lõppedes VISi e-hääletanute nimekiri jaoskondade kaupa sorteeritult (joonis 3). Jaoskondades märgitakse e-hääletanute nimekirjas need isikud, kes on andnud ka paber-hääle ning nende isikute e-hääled tühistatakse.

EUROOPA PARLAMENDI VALIMISED 2019

N	R	L	P	E	T	K	N	R	L	P
16.05	17.05	18.05	19.05	20.05	21.05	22.05	23.05	24.05	25.05	26.05.2019
EELHÄÄLETAMINE							Hääletamist ei toimu			VALIMISPÄEV
Maakonnakeskuses hääletamine 12.00 - 20.00			Jaoskonnas hääletamine 12.00 - 20.00		Jaoskonnas hääletamine 9.00 - 20.00					
E-hääletamine 9.00 - - 18.00					Kodus hääletamine

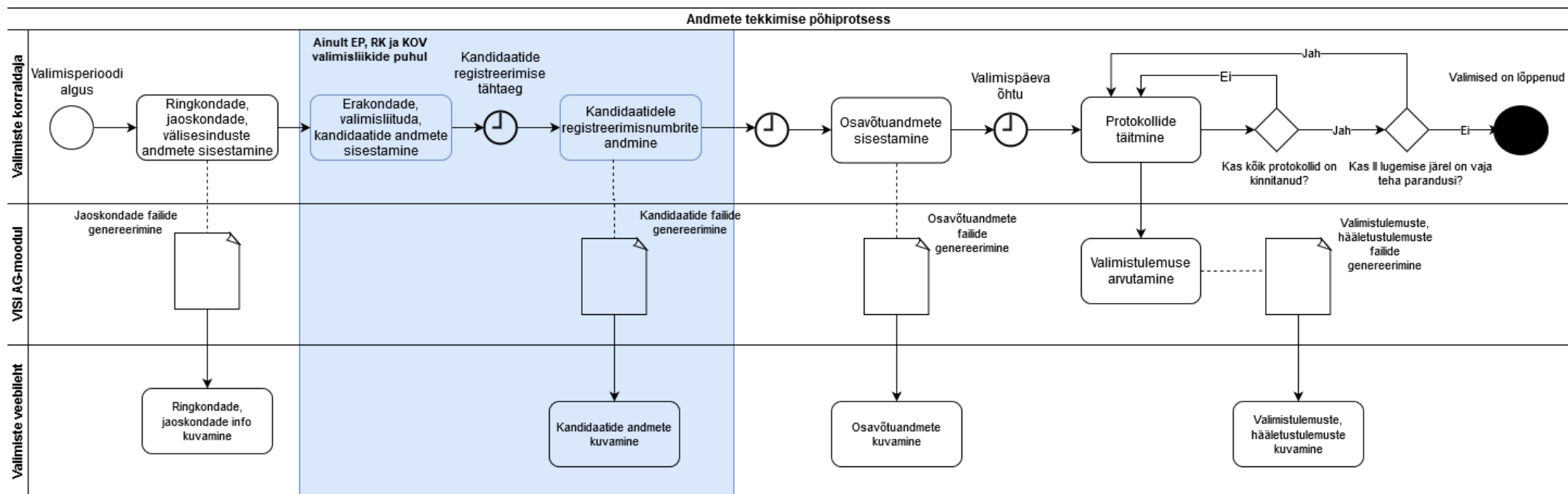
Joonis 2 . Valimisperioodi ajaline graafik 2019. aasta Euroopa Parlamendi valimiste näitel [21]

Valimispäeva õhtul algab pärast jaoskondade sulgemist hääle lugemine ning jaoskonnaprotokollide täitmine. Lisaks iga kandidaadi saadud häälele märgitakse sinna ka info hääletussedelite kasutamise kohta (nt rikutud, kasutamata sedelid). E-hääled (nende hulgast on eemaldatud tühistatud e-hääled) imporditakse VISi samuti valimispäeva õhtul. Kui protokoll on VISis jaoskonnakomisjoni poolt kinnitatud, siis summeeritakse sisestatud andmed edasi kõrgema taseme protokoll (näiteks RK valimiste korral ringkonnaprotokoll). Kõiki andmeid hoitakse VISi andmebaasis. Alates 2009. aastast on VISile lisatud nn aruandegeneraatori moodul, mis genereerib vastavalt VISi andmetele XML ja HTML faile. HTML failid kuvatakse valimiste veebilehel, XML failid on kättesaadavad meediaasutustele, kes on selles eelnevalt RVTga kokku leppinud. Andmeid oleks võimalik veebilehtedele kuvada ka otse VISi andmebaasist, kuid kuna veebilehtede koormus on valimisperiodil, eriti aga valimispäeva õhtul väga suur, on valitud turvalisem variant - kuvada veebilehel staatilisi HTML-faile. Valimisinfo failid on sisaldanud järgmisi andmehulki:

- Jaoskonnad (viimastel valimistel asendatud üldkasutaja jaoks avaliku kaardirakendusega, jaoskonnaandmeid sisaldavad failid on genereeritud ainult asutusesiseseks kasutamiseks)

- Kandidaadid (fail genereeritakse enne valimisi siis, kui kandidaatidele on antud registreerimisnumbrid, andmed kuvatud samanimelisel veebilehel, nt: <https://rk2019.valimised.ee/et/index.html>)
- Hääletamisest osavõtu statistika (fail genereeritakse eelhääletamise ajal 6., 5., ja 4. päeval enne valimispäeva ning valimispäeval kell 12:00, 16:00 ja 20:00. Failid genereeritakse siis, kui kõik jaoskonnad on info edastanud. Andmed on kättesaadavad vastaval veebilehel, nt: <https://rk2019.valimised.ee/et/participation/participation.html>)
- Valimistulemus (Genereeritakse valimispäeva õhtul. Info kuvatud veebilehel nimega “Hääletamis- ja valimistulemus”, vastavalt valimisliigile valitud organi liikmed, nt: <https://rk2019.valimised.ee/et/election-result/election-result.html>)
- Detailne hääletamistulemus ringkonnas, maakonnas (Genereeritakse valimispäeva õhtul, nt: <https://rk2019.valimised.ee/et/voting-result/district-1-voting-result.html>)
- Hääletamistulemus valdades ja linnades (Genereeritakse valimispäeva õhtul siis, kui vastavas vallas/linnas asuvad jaoskonnad on protokollid kinnitanud, nt: <https://rk2019.valimised.ee/et/voting-result/local-municipality-0205-voting-result.html>)

Joonis 3 esitab andmete sisestamise ning failide tekkimise vahelised seosed BPMN (*Business Process Modelling Notation*) põhiprotsesside mudelina.



Joonis 3. Andmete tekkimise *as-is* põhiprotsess viimaste valimiste näitel. Valimiste korraldaja all on olenevalt valimisliigist mõeldud RVTd või KOV valimiskomisjoni.

Valimiste infosüsteemi põhimäärus sätestab infosüsteemi kantavate andmete säilitamise tähtajatult [19, §15], seda, millised andmed peavad olema aga avalikustatud, ei ole iga andmehulga kohta täpselt määratud. Praeguse seisuga on kokkuvõtlikult valimiste veebilehtedel läbivalt avaldatud valimis- ja hääletamistulemused ning osavõtustatistika, kandidaatide nimekirjad ning enamasti ka jaoskonnad. Veebilehtedelt puuduv info on olemas RVT asutusesiseseks kasutamiseks mõeldud digitaalsetel ja paber-dokumentidel, mille abil on veebiandmestikku võimalik täiendada.

1.4. Varasemad uuringud ja nõuded valimisandmete veebiarhiivile

Vajadus süstematiseerida toimunud valimiste andmeid on olnud juba varasemalt ning esimene kirjalik sissevaade probleemi ning võimalikesse lahendustesse on eespool mainitud 2015. aasta lähteülesanne [7]. VARHi, kui laiale kasutajaskonnale mõeldud rakenduse loomise juures on põhiküsimuseks, milliseid andmeid, millises koosseisus ning millise detailsusega erinevad kasutajagrupid vajavad. VARHi kasutajad ning nende vajadused on vastavalt lähteülesandele kaardistatud järgnevas tabelis.

Tabel 1. Valimisandmete kasutajagrupid ja infovajadus [7]

		Grupp	Kirjeldus
Eesti kasutajad	Vahetult seotud kasutajad	Erakonnad	Erakond analüüsib kandidaatide edukust, nii erakonnasiseselt kui ka üle erakondade. Tulevikus hakkab suure tõenäosusega huvi pakkuma tulemus hääletamisviiside kaupa.
		Kandidaadid	Kandidaate huvitab eelkõige enda tulemus oma ringkonnas ja jaoskonniti ning erakonna kontekstis, tulevikus ka hääletamisviiside kaupa.
		Valimiste korraldajad	Korraldajad kasutavad valimisandmeid laialdaselt üldiseks analüüsiks, seadusloomeks, valimiste protsessi kvaliteedi tagamiseks ning ka oma töötulemuste analüüsiks.
		Erasektor	Erasektor võib huvi tunda

			valimisandmete vastu eesmärgil luua neile lisaväärtus, mida hakatakse edasi müüma toote või rakendusena.
		Tavakodanikud	Valija rollis tavakodanik tunneb huvi, kuidas läks tema hääle saanud kandidaadil. Laiemas mõttes tavakodanikud võivad huvituda pea kõigist valimisandmetest isikliku statistilise analüüsi eesmärgil.
		Meedia	Meediat huvitavad eelkõige üksikisikute andmed, nii konkreetse valimissündmuse kontekstis kui ka läbi aegade. Suurim huvi on vahetult pärast valimisi.
		Teadlased	Sotsiaalteadlased soovivad näha valimisandmestikus tuvastatavaid mustreid ja tendentse. Neid huvitavad võimalikult täpsed aegread ja omavahel võrreldavad andmed aastate lõikes. Teadlasi huvitavad eelkõige konsolideeritud isikustamata andmestikud, mis iseloomustavad tervikut.
		Tudengid	Tudengid soovivad eelkõige massandmeid, mille hulgast huvipakkuvaid välja filtreerida.
Välisriigi kasutajad		Välisriigi teadlased ja ametnikud	Euroopa Komisjon analüüsib eelkõige Euroopa Parlamendi valimiste andmeid. OSCE ja muud organisatsioonid soovivad teha üldanalüüse oma eesmärkide täitmiseks.
		Välisriigi meedia	Välisriigi meedia kasutab valimisandmeid uudiste koostamiseks, sh eelkõige elektroonilise hääletamise andmeid.

Tabelis märgitud kasutajagruppidele võib lisada veel välisriikide tavakodanikud, kelle huvi võib sarnaselt Eesti tavakodanikule puudutada kas konkreetse kandidaadi tulemust või valimistulemusi üldisemalt.

Kui VARH suudab pakkuda kasutajatele vajaliku täpsusega infot, väheneb RVT halduskoormus teabenõuetele vastamise osas. Valimisandmete andmebaas annaks

kasutajatele ülevaate, milliseid andmeid on aastate jooksul kogutud ning avaneb ülevaade andmetest, mida varem pole ehk osatud küsida. Andmete kogukondlik tõlgendamine ja kontroll suurendavad valimisprotsessi läbipaistvust [7].

Eeltoodust tulenevalt esitatakse lähteülesandes nõuded valimisandmete veebiarhiivile kahes aspektis. Esimene puudutab andmete importimist (VISist valimisandmete andmebaasi) ja haldamist ning seda vaadeldakse lähemalt käesoleva töö 3. peatükis. Teine aspekt kirjeldab nõudeid andmete vaatamisele ja eksportimisele. Kuna nõuded on lähteülesandes toodud loogiliselt ja struktuurselt seotuna, on need ka siin toodud koos, kuigi osa neist puudutab avaandmetele esitatud nõudeid.

1. Andmete vaatamine ja eksportimine
 - a. Kasutaja peab saama valida arhiivi keelt (eesti/inglise)
 - b. Kasutaja peab saama sirvida valimisandmeid:
 1. Koondandmeid ühe valimissündmuse liigi lõikes
 2. Ühe valimissündmuse kõiki andmeid
 - c. Kasutaja peab saama kandidaate otsida
 - d. Kandidaatide kandideerimise andmebaasi peab saama valimissündmuse ja otsingu kaupa alla laadida
 - e. Valimisandmeid peab saama tervikuna alla laadida
 - f. Iga tabeli kujul infohulka peab saama eksportida XML formaati
 - g. Üldtabeleid peab saama eksportida lisaks CSV formaati
 - h. Iga tabelit ja diagrammi peab saama välja trükkida ja läbi brauseri välja trükkimise vahendi salvestada ja PDFina [7]

Neist nõuetest täpsustatakse käesolevas töös nõudeid 1.b, 1.c, 1.d., 1.e. 1.f , 1.g. ehk tuuakse täpselt esile, millised on valimise koondandmeid, millised andmed sõltuvad valimisliigist ning millised andmehulgad on ajas muutunud niipalju, et nende võrdlemine pole asjakohane.

Nõude 1.a. osas saab VARHi puhul kasutada ka juba praegu hilisemate valimiste veebilehtede puhul kasutatavat lahendust, kus tõlgitud märksõnad võimaldavad ka

välisuurijatel andmetest ülevaate saada. Allalaaditavate andmete osas on ka praegu kättesaadavatel andmetel XML-atribuudid loodud inglise keeles ning selliselt, et andmehulga sisu on neis arusaadav. Kasutajaliidese keele valik lahendatakse veebilehe poolal. Nõue 1.h. puudutab VARH veebilehte, kuhu tuleb vastav funktsionaalsus lisada, kuid see tuleb lahendada VARHi kasutajaliidese juures ega kuulu siinse töö skoopi.

Nõue 1.c on praegu täidetud RVT veebilehel kandidaadi otsinguga täidetud [22], kuid otsingu tulemust pole praeguse seisuga võimalik alla laadida. Töös esitatakse sellele nõudele täpsustus, mis käsitleb kandidaadi otsingut juhul, kui kandidaadi nimi on valimissündmuste vahel muutunud.

Nõue 1.f. on täidetud valimistulemuste osas juba praegu 2014-2019 valimisandmete juures, kus kõik tulemuste tabelid on avaandmetena kättesaadavad. Selles töös tehtud analüüsi põhjal saab lisada jaoskondade ja kandidaatide andmed ning lisaks varasemate valimissündmuste kõiki andmed. Nõue 1.g. saab täita kasutajale juhendmaterjali esitamisega, kuidas XML-faili on võimalik CSV-failina avada, vastavad andmehulgad, mis analüüsi põhjal vastavad määratlusele “üldtabel” määratakse analüüsi käigus.

1.5. Olemasolevad juhendid avaandmetele

Käesolevas peatükis analüüsitakse Eesti kontekstis olemasolevaid juhiseid avaandmete loomiseks. Peamiseks allikaks on seejuures Riigi Infosüsteemide Ameti poolt koostatud “Avaandmete loomise ja avaldamise juhend” [23].

Avaandmed on Andmekaitse Inpektsiooni definitsiooni kohaselt avaliku või erasektori teave, mille kasutamist ei ole seadusega või seadusega kehtestatud korras piiratud. Avaandmetel on kolm tunnust: need on tehtud tasuta kättesaadavaks kõigile isikutele mistahes kasutuseesmärgil; andmed on digitaalsed, masinloetavad ja riskasutatavad teiste andmetega; andmete levitamiseks ja taaskasutamiseks ei ole seatud piiranguid. Taaskasutamiseks ei loeta teabevaldajate vahel teabe vahetamist [24]. “Avaandmete loomise ja avaldamise juhend” soovib avaandmete planeerimisel, projekteerimisel,

loomisel ja avaldamisel lähtuda mitmetest kriteeriumitest, millest lühike ülevaade on toodud allpool.

Juhendi kohaselt tuleb tagada andmete **eraldatus** - avaldatud andmed peavad olema eraldi nn. tööandmetest. Andmed peavad olema lihtsasti **kättesaadavad** ja **juurdepääsetavad** kõigile ning sõltumata kasutaja residentsusest või juriidilisest staatusest. Samasisulisi andmeid tuleb esitada võrreldavate ajavahemike kaupa ehk **perioodiliselt**, ning selliselt, et võrreldavus varasemate andmetega ei kaoks ning kogu ajatelge katvalt **järjepidevana**. Kokku kuuluvad andmed tuleb omavahel siduda, millega tagatakse andmete **terviklikkus**. Tagada tuleb andmete **terviklus** –avaandmete füüsiliselt erinevad kogumid, mis kõigi tunnuste järgi on tuvastatavad samade andmete kogumina, ei tohi olla variantsed ja ajas muutuvad. Avaandmete **esitusviis** peab olema üldtunnustatud, vaba litsensiga kaetud formaadiga formaliseeritud kujul. Andmete **mõistetavuse** huvides tuleb kättesaadavaks teha andmeelementide semantilised kirjeldused. **Litsenseerida** tuleb andmed selliselt, et on võimaldatud nende andmete töötlemine sh. nende andmete rikastamine teistest mistahes andmekogumitest pärit andmetega. Alustavalt peab andmete valdaja andmeid avaldama oma initsiatiivist lähtuvalt ja enda poolt kehtestatud korra alusel ja seaduses kehtestatud mahus ehk **proaktiivselt**. Avaandmetena esitatavad andmed ei tohi üldjuhul olla seostatud/seostatavad isikuga ehk peavad olema anonümiseeritud ega tohi sisaldada **konfidentsiaalseid** andmeid. Andmete sisu ja **õigsuse** eest vastutab nende esitaja so. andmete omanik [23].

Eespool toodud nõudeid võetakse läbivalt arvesse andmete klassimudelite loomisel ning andmestruktuuride koostamisel.

1.6. Uurimismeetod

Töö üheks eesmärgiks oleva detailanalüüsi käigus uuritakse valimisandmete veebilehtedel olevat infot, et korrastada ja struktureerida korduvad või eristuvad andmehulgad. Iga korduvat andmehulka uuritakse eraldi peatükis. Kirjeldatakse

seaduses esitatud nõudeid andmetele, mida kuvatakse veebilehele ning millist infot on nõuete kohaselt vaja migreerida tuleviku andmebaasi. Andmeid analüüsitakse kolmest aspektist:

- **valimisliikide mõistes** - nelja selles töös käsitletava valimisliigi korral tingivad valimisseadused erinevate andmestruktuuride tekkimise
- **andmehulkade mõistes** - andmehulga sisu muudatused valimisliigi siseselt
- **ajalises mõistes** - ajas võib muutuda esitatud andmete hulk või sisu

Analüüsi tulemusel leitakse andmete vahelised seosed, mille puhul arvestatakse leitud aspektide ja eripäradega. Andmehulkade omadused ning omavahelised seosed esitatakse klassimudelina. Analüüsist leitu ning klassimudelite ühildamisel on kasutatud Chi-Lun Liu väljapakutud meetodikat, mis tugineb konstruktivismi ning ontoloogia teooriatele. Selle järgi koosneb protsess neljast osast:

- 1) eelnev teadmiste kogumine
- 2) muudatuste spetsifitseerimine
- 3) muudatuste analüüs
- 4) muudatuste juurutamine

Siit kaks esimest punkti tegelevad konstruktivistliku teooria kohaselt varasemate ontoloogiate ning seoste assimileerimise ning kirjeldamisega, kaks viimast leiavad puuduolevad või vajadustele mittevastavaid seosed ning adapteerivad uued vajadused [25]. Valimisandmete kontekstis tähendab see olemasolevate, HTML-lehtedel asuvate andmestruktuuride võrdlust VISi aluseks oleva andmemudeliga, selles ühisosa leidmist ning uute seoste leidmist vastavalt andmete loogilistele seostele. Ontoloogiliselt on valimiste spetsiifilised mõisted (nt. kandidaat, erakond, ringkond) juba VISi loomisel leitud ning detailiseeritud ja käesolev töö taaskasutab neid mõisteid. RIA avaldatud "Ontoloogiate loomise meetodika" [26, lk 5] valdkonnaontoloogiate loomise protsessi järgi on selle töö analüüsi osas ette võetud seoste ja atribuutide lisamine vastavalt analüüsis leitud nõuetele.

Iga andmehulga juures on klassimudelites lisatud niipalju seoseid, kui selle hulga struktuuri selgitamiseks vajalik. Andmehulkade analüüsi iga peatüki lõpus tuuakse esile

eelnevast tulenevad täpsustused 2015. aastal esitatud nõuetele. Nõuete kirjeldamisel võetakse arvesse RVT infosüsteemide valdkonna peaspetsialistiga tehtud küsitlusi, kel on ülevaade laekuvatest teabenõuetest [8]. Klassimudelitega struktureeritud andmetest luuakse töö 3. peatükis relatsiooniline andmemudel. Olemite kirjeldused on toodud töö lisades. Andmemudeli loomise ja kirjelduse juures on juhiseks võetud RIA relatsiooniliste andmemudelite koostamise juhendist [27] ning Bill Inmoni andmebaaside arhitektuuri ja kavandamise soovitusel [28].

Avaandmete struktuurid, mis arvestavad analüüsis tehtud järeldustele, on koos selgitustega 3. peatükis.

Piirangutena tuleb käesoleva töö puhul arvestada valimisseadustes ning VISi määruuses esitatud nõuetega, mis tuuakse välja iga andmehulga analüüsi juures. Kui seaduses mõne andmehulga kohta piiranguid ei ole, siis arvestatakse valimiste veebilehtedel väljakujunenud tavasid andmete esitamise detailsuse osas. Selles, milliseid andmeid esitada, on lähtutud Avaliku teabe seadusest [29], isikuandmetesse puutuvast RVT andmekaitsetingimustest [30] ja Isikuandmete kaitse seadusest [31]. Andmete esitamisel tuleb käsitleda piiranguna ka olukorda, kus mingid andmed on RVTs olemas, kuid mitte valimiste veebilehel avaldatud, kuna vaadeldaval ajajärgul polnud see tavaks. Hilisemate valimissündmustega võrreldes tekitab see lünga andmete võrreldavusse. Andmed tuleb sel juhul RVT arhiivdokumentidest VARH andmebaasi sisestada, kuid tuleb arvestada võimalusega, et see töö on ajamahukas. Veebilehtedelt puuduvate (kuid RVT arhiivis olemasolevate) andmete korral antakse analüüsis juhiseid olukorra kommunikeerimiseks.

2. Andmehulkade analüüs

Lähteülesanded esitatud nõuded käsitlevad valimisandmeid ühtse tervikuna. Lähteülesanne nõue 1.d.i sätestab näiteks, et valimisandmeid peab saama tervikuna alla laadida. Praegusel kujul sisaldavad valimisandmetega veebilehed aga ka infot, mis ilmselt ei kujuta endast senise teadmise ning laekunud päringute põhjal tuleviku uurijate jaoks väärtust, vaid on tähtsust omanud üksnes valimisperioodil. Siia alla kuuluvad näiteks valimisjaoskondade kontaktandmed, valimiskomisjonide tööajad ja aadressid jms. Juhul, kui peaks siiski mingil põhjusel tekkima vajadus selle ajaloolise info leidmiseks, on see leitav ka RVT või KOVide arhiividest. Järgnevas analüüsis määratletakse täpsemalt andmehulgad, mis kujutavad endast kasutajagruppidele huviobjekti ning kaasatakse valimisandmete tervik-komplekti.

Kõiki valimisandmete hulki analüüsitakse nende tervikluse ning asjakohasuse suhtes. Andmehulkade näidisstruktuuri juures juures võetakse eeskujuks kõige viimasema samaliigilise valimise andmehulki, kuna need on kujunenud nõuete pideva uuendamise ja täiendamise teel ning arvestavad tänaseks teada olevate infovajadustega. Näiteks e-hääled olid esialgu veebis eristatavad vaid ringkondade kaupa (RK), kuid 2012. aastal vastu võetud valimisseaduste muudatuste järel hakati e-hääletustulemus kindlaks tegema ka jaoskonniti [6, lk 5].

Valimissündmused saavad olla järgmistes liikides:

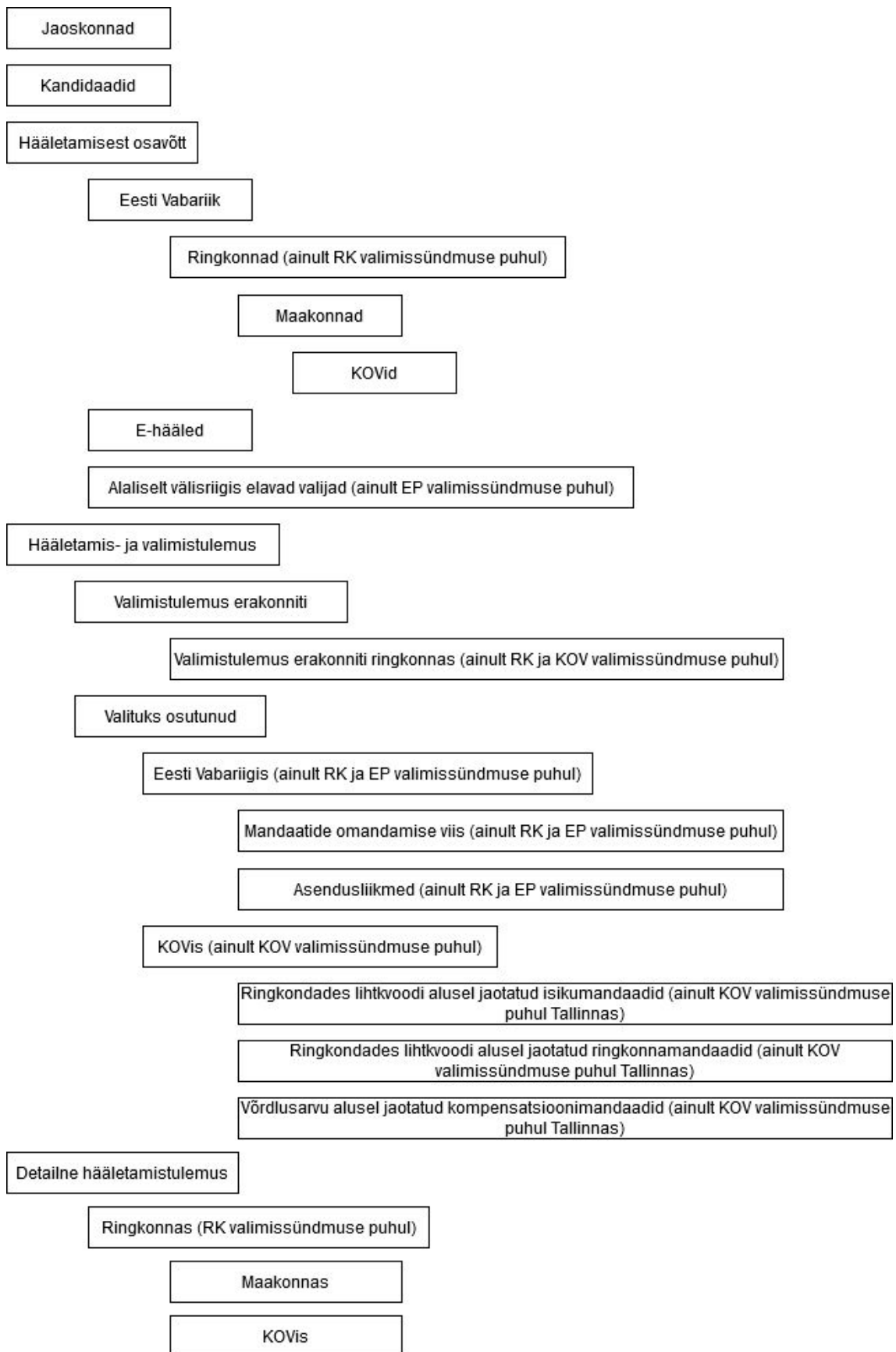
- Riigikogu valimised
- Kohaliku omavalitsuse volikogu valimised
- Euroopa parlamendi valimised
- Vabariigi Presidendi valimised
- Riigikogu juhatuse valimised
- Rahvahääletused

Selles töös on vaatluse all on neli üleriigilise korraldusega valimisliiki, nende toimumisajad liigiti on märgitud tabelis 2.

Tabel 2: Toimunud üleriigilised valimised toimumisaja ja liikide järgi

Valimisliik	RK	EP	KOV	RH
1.	20.09.1992	13.06.2004	17.10.1993	28.06.1992
2.	05.03.1995	07.06.2009	20.10.1996	14.09.2003
3.	07.03.1999	25.05.2014	17.10.1999	
4.	02.03.2003	26.05.2019	20.10.2002	
5.	04.03.2007		16.10.2005	
6.	06.03.2011		18.10.2009	
7.	01.03.2015		20.10.2013	
8.	03.03.2019		15.10.2017	

Veebilehtedel esitatud andmehulkadest annab ülevaate joonis 4. Sisuliselt võib andmed jagada viite suuremasse gruppi: jaoskonnad, kandidaadid, hääletamisest osavõtt, hääletamis- ja valimistulemus ning detailne hääletamistulemus. Kolm viimast hulka jagunevad omakorda alaliikideks, mis mahukat valimisandmestikku eri valimisliikidele vastavalt esitavad ja reastavad. Andmehulgad on siin joonisel esitatud käsitletavate valimisliikide suhtes hirerahiliselt, vastavalt Bill Inmoni soovitusel andmestruktuuride loomiseks [28].



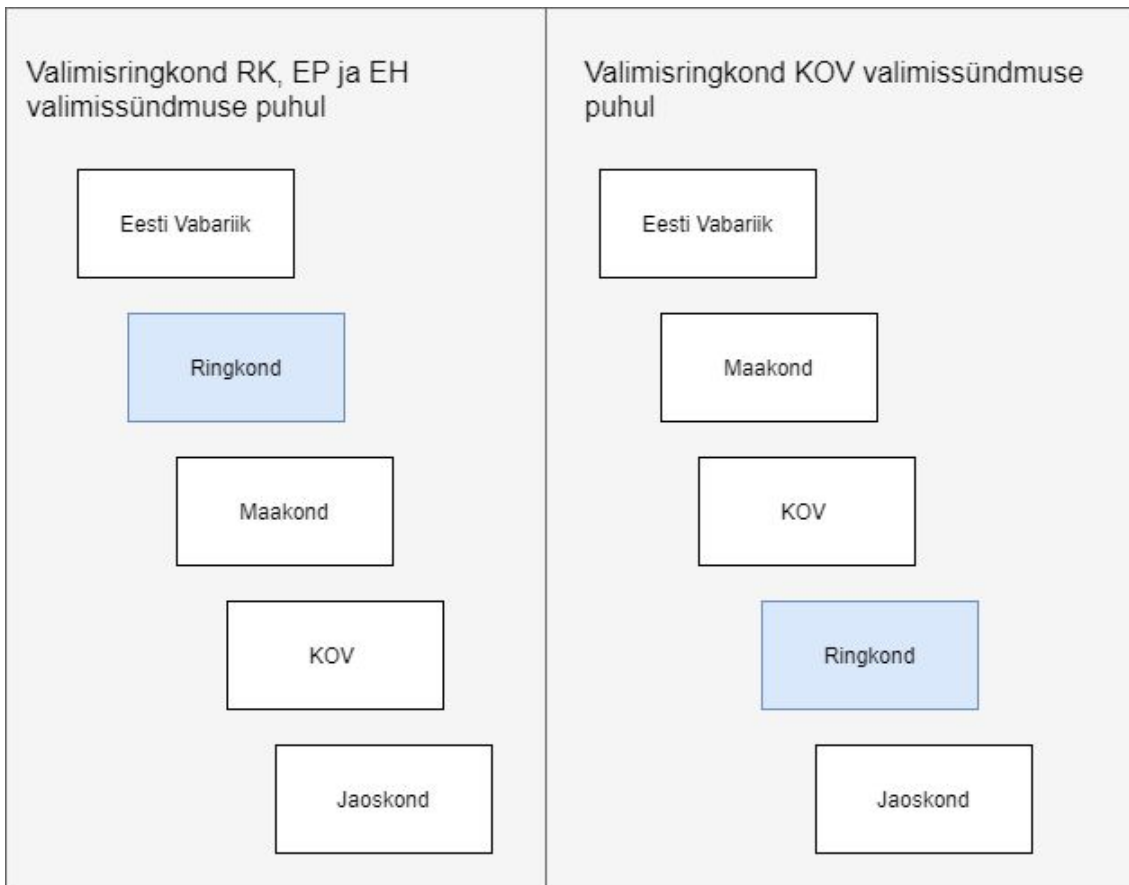
Joonis 4. Andmehulkadeks struktureeritud valimisinfo

Järgmistes peatükkides esitatakse iga andmehulga analüüsi lõpuks klassimudel andmehulkade mõistmiseks vajalike tabelite, atribuutide ning seostega. Vastava andmehulga all on järeldustena toodud täpsustused lähteülesandes esitatud nõuetele.

2.1. Valimisjaoskonnad ja -komisjonid

Valimiskomisjonide ja -jaoskondade näol on tegemist valimiste kõige nähtavama osaga häälte kogumisel. Valimisseaduste järgi moodustatakse iga valla ja linna ning Tallinna linnaosa territooriumil vähemalt üks valimisjaoskond, kus korraldab hääletamist jaoskonnakomisjon [13]-[16]. Hierarhiliselt on valimiskomisjonide puhul eristatavad kolm taset. Kõigi valimisliikide korral asub jaoskondade hierarhias kõige kõrgemal kohal Vabariigi Valimiskomisjon. RK, EP ja RH valimissündmuste korral allusid sellele kuni 2015. aastani maakonna valimiskomisjonid, mille alla omakorda kuulusid jaoskonnakomisjonid. KOV valimiste korral kuulusid maakonnakomisjonide alla omakorda KOV taseme komisjonid, millele allusid jaoskonnakomisjonid. Alates 2019. aasta RK valimistest ei moodustata enam üheski valimisliigis maakonnakomisjone, vaid nende rolli asendavad KOV taseme komisjonid. Erandlikud on kõigi valimisliikide korral Tallinn ja Tartu linn, mis, olles sisuliselt KOVid, asuvad kõigis struktuurides maakondadega samal tasemel, kuna on teistest KOVidest märgatavalt suurema elanike arvuga. Tallinnal ja Tartul on ka RK, EP ja RH valimissündmuse korral oma valimiskomisjon.

VISis valimissündmust luues on üheks esimeseks tegevuseks valimisüksuste nõ. puud sisaldava JSON faili importimine. Fails on kirjeldatud vastava valimissündmuse korral tekkiv ringkonnastruktuur, mida kujutatakse joonisel 5. EP ja RH korral on kogu Eesti üks ringkond, RK valimiste korral on ringkondi 12. KOV valimistel asuvad ringkonnad KOVi sees. Tallinnas moodustatakse ringkonnad linnaosade kaupa.



Joonis 5. Ringkonna asukoht haldusüksuste struktuuris eri valimisliikide korral

VISi laetavas JSON failis on maakonna ja KOV taseme üksused kirjeldatud väljade “code” - üksuse EHAK kood, “name” - üksuse nimi, “idLevel” ja “type” abil. Nii EHAK koodid kui maakondade-KOVide nimed on aja jooksul muutunud ning 2017. aasta haldusreformiga ühinesid paljud vallad. Kuna haldusüksuste struktuur on pidevas muutumises ja muutub veelgi, on andmebaasi loomise mõttes parem arvestada, et iga valimissündmuse korral tekib eraldi haldusstruktuur, mida ei taaskasutata. Lähteülesande nõue 1.b.1. sätestab, et valimisandmete korral peab saama vaadata koondandmeid ühe valimissündmuse liigi lõikes ning valimissündmuse kõiki andmeid. Koondandmete all võib mõista üle-eestilisi andmeid jaoskondade koguarvu, valimisaktiivsuse vms suhtes, kuid võrrelda näiteks Häädemeeste valla valimisandmeid 1996 ja 2019 aasta lõikes on keerukas, kuna valla piirid on muutunud. Tulevases andmestruktuuris tuleb iga üksuse kohta salvestada tema omaegne EHAK kood, mis võimaldab Statistikaameti haldus- ja asustusjaotuse klassifikaatori abil leida haldusüksuste struktuuri muutused.

VISi sisestatakse enne valimisi VISi põhimääruses nõutav info iga valimiskomisjoni, valimiskomisjoni liikme, valimisjaoskonna ja välisesinduse kohta [19]. Kui seaduses ning määruses eristatakse eri andmehulki - valimiskomisjon, valimiskomisjoni liige, valimisjaoskond, siis veebilehtedel on põhifookuses olnud valija teavitamine valimisjaoskondade asukohtadest, lahtiolekuaegadest ning kontaktandmetest, kuhu probleemide korral pöörduda ning info on sellest lähtuvalt ka esitatud, sisaldades eri andmehulkade elemente ühe, kasutaja jaoks olulise tervikuna. Veebilehtedel ära märgitud ka maakonna- ja KOV valimiskomisjonid, kuigi need otseselt ei tegele hääletajate vastuvõtmisega. Vabariigi Valimiskomisjoni, maakonna valimiskomisjonide ning Tartu linna ja Tallinna linna valimiskomisjoni kohta on neis andmetes, mis veebilehelt kättesaadavad, järgmine andmestik*:

1. Valimiskomisjoni nimi
2. Aadress
3. Sidevahendid
4. Tööaeg
5. Esimees
6. Aseesimees (ainult 1999)
7. Sekretär (ainult 1999)
8. Liikmed (ainult 1999)
9. Hääletamise aeg valimispäeval (ainult 2013)

**1992-1996, 2011, 2019 RK andmed puuduvad, KOV2009 andmed maakondade, KOVide kohta puuduvad; 2017 oli jaoskondade osas kasutuses Rahvastikuregistri kaardirakendus, mis on tänaseks suletud; 2019 EP korral on kättesaadav exeli-formaadis andmefail jaoskonnainfoga.*

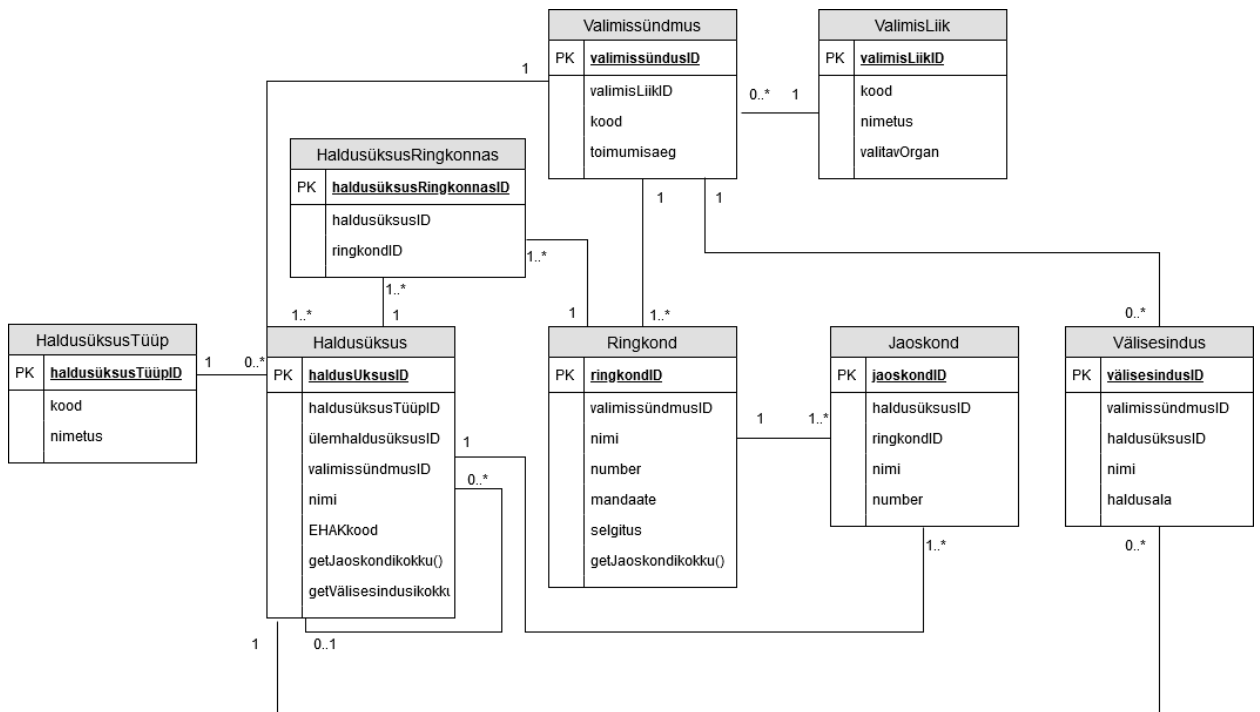
Jaoskonnana toimivad sisuliselt ka välisesindused, mis korraldavad hääletamist välisriikides. Välishääli loetakse ainult RK, EP ja RH valimiste korral ning nendeks valimisteks luuakse ka välisesindused. Komisjonide asemel on välisriikides hääletamist korraldavad isikud.

Välisesinduste kohta on leitav järgmine info (1992, 1993, 1995, 2003, 2004, 2007, 2011, 2014, 2015, 2019 andmed puuduvad, olemas on ainult 1999., 2009. aasta andmestik):

1. Riik (2009. aasta andmetes ka see, mis piirkondi see katab)
2. Aadress
3. Sidevahendid
4. Hääletamist korraldav isik
5. Kirja teel hääletamiseks taotluse esitamise tähtaeg
6. Kirja teel hääletamise tähtaeg
7. Välisesinduses hääletamise tähtaeg

Jaoskondade ning välisesinduste puhul on uurijate jaoks oluliseks ning huviväärseks infoks nende arv mingis üksuses ning selle muutused ajas. Alates 2014. aastast on jaoskondade arv veebilehtedel ka eraldi üksuste kaupa välja toodud. Jaoskondadega on seotud Eesti administratiivstruktuur, mille esitamine valimisaastale vastaval kujul tuleb avaandmetes ning andmebaasis lahendada.

Joonisel 6 esitatud klassimudel kujutab struktuuri, mis hõlmab valimissündmust, haldusüksusi, ringkondi, jaoskondi, välisesindusi ning nendevahelisi seoseid. Haldusüksuste all on mõistetud Eesti Vabariiki, maakondi ning KOVe (määratud tabeli HaldusüksusTüüp) abil. Nende omavaheline seos on väljendatud rekursiivse seose abil. RK, EP ja RH valimissündmuste korral asuvad ringkonnad Eesti Vabariigis, KOV valimiste korral KOVis ning see seos lahendatakse tabeli “HaldusüksusRingkonnas” abi. Välisesindused ei ole seotud ühegi ringkonnaga, vaid alluvad Eesti Vabariigile. Jaoskondade ja välisesinduste arv ei ole veebis enamuse välisesinduste puhul eraldi kuvatud kuid nende arvu saab küsida iga üksuse kohta *get*-päringute abil. Valimiste liik (RK, EP, KOV, RH) määratakse vastava tabeli abil. Enne valimisi määratakse igas ringkonnas mandaadid - st mitu kohta kandidaatide vahel jagatakse. Mandaatide arv sõltub vastava ringkonna elanike arvust.



Joonis 6. Valimisüksuste, ringkondade ja jaoskondade andmetega klassimudel

Täpsustused nõudele 1.b.2:

Valimisandmed, mida on jaoskondade kohta võimalik veebilehel ühe valimissündmuse kohta näha ning alla laadida on järgmised:

- Allalaaditav fail peab sisaldama Eesti haldusüksuste struktuuri, kus ringkond asub vastavalt valimisliigile määratud tasemel. Struktuuri osad on: Eesti Vabariik, maakond, KOV, jaoskond. Allalaaditavas failis peab iga struktuuriüksuse kohta olema märgitud:
 - Nimi
 - EHAKkood
- Kui valimisliigist tulenevalt on loodud välisesindused, siis nende kohta:
 - Nimi
 - Teeninduspiirkond
- Iga ringkonna kohta järgmine info:
 - Nimi
 - Number
 - Mandaatide arv

- Iga KOVi kohta jaoskonnad järgmise infoga
 - Jaoskonna nimi
 - Jaoskonna number

Loetletud andmed arvatakse l.e. nõude kohaselt valimisandmete koosseisu. Need andmed ei moodusta tabelit ning seetõttu ei rakendu siin nõuded l.f. ja l.g.

2.2. Kandidaadid

RK, EP või KOV valimistel osaleda soovivad kandidaadid peavad seaduse kohaselt täitma kandideerimisavalduse, kus väljendab soovi kandideerida erakonna (KOV valimiste korral või kuni 1995. aastani RK valimiste korral ka valimisliidu) nimekirjas või üksikkandidaadina ning kinnitab, et vastab kandidaadile esitatavatele nõuetele [13]-[15]. Avalduses esitab kandidaat ka enda kohta nõutavad andmed. Seda, millised andmed tuleb esitada ning milliseid võib esitada, on ajas muutunud. 1999. aastal kohalike omavalitsuste valimisteks välja antud määruse kohaselt **peab** kandidaat kandideerimisankeedis esitama oma ees- ja perekonnanime, isikukoodi, erakondliku kuuluvuse, postiaadressi ja sidevahendite numbrid. Esitada **võib** andmed sünnikoha, hariduse ja eriala koos õppeasutuse nime ja lõpetamise ajaga, teaduskraadi, töökohta ja ameti kohta [13, §28][14,§32’][15,§27]. Sama seis on 2002. aasta Riigikogu valimisteks välja antud määruses [33]. 2013. aasta määruse järgi kandidaat aga esitab järgmised andmed: ees- ja perekonnanimi, isikukood, erakondlik kuuluvus, elukoha aadress, sidevahendite numbrid, kontaktandmed, andmed hariduse kohta, töökoht ja amet [34]. Nõuetele vastavad kandidaadid registreeritakse ning neile antakse registreerimisnumber vastavalt nende erakonnale/valimisliidule antud liisunumbri järgi. Liisku hakati heitma alles 2004. aastast, enne seda määras erakondade või valimisliitude järjestuse kandideerimisdokumentide esitamise järjekord [6, lk 31]. Kandidaadid avalikustatakse valimiste veebilehel. Milline komplekt kandidaadi avalduses esitatud andmetest avalikustatakse, on valimisseadustega reguleeritud. Veebis olevate andmete analüüsimisel on kandidaadi kohta avaldatud järgmised andmed:

- Registreerimisnumber
- Nimi

- Sünniaeg
- Sünnikoht (kuni 2014, edaspidi ainult EP valimistel EL kodaniku puhul)
- Perekonnaseis (ainult 1992-1999)
- Amet (kuni 2007)
- Töökoht (kuni 2011)
- Elukutse (ainult 2011)
- Töökoht ja amet (alates 2013)
- Haridus (alates 2015 haridustase)
- Õppeasutus (kuni 2007)
- Eriala (kuni 2007)
- Teaduslik kraad (1992-2014)
- Haridus, õppeasutus, eriala (ainult 2009)
- Erakond / erakondlik kuuluvus (1999-2019)
- Nimekiri või üksikkandidaat (2013-2019)
- Kontaktandmed (2015-2019)

Ka enne valimisi, valimiste ajal ning vahetult pärast valimisi kandidaatide kohta veebilehel kättesaadav info võib olla hiljem kandidaadi nõudel osaliselt kustutatud (näiteks e-postiaadress, telefoni number). Erinev on olnud ka lähenemine kandidaadi hariduse ja töökoha küsimusele - osade valimissündmuste korral on need praeguseks kustutatud kõigil kandidaatidel, kuid kustutatud on ka nt üksikul kandidaadil, kes selleks on soovi avaldanud. Kandideerimisavaldust on siiani saanud täita nii paberil kui vastavalt RVT antud vormile digitaalselt ning väljade täitmise automaatset kontrolli ei toimu. Nii on hariduse ja töökoha väljadele sattunud väga erisuguse kvaliteediga infoti. Alates 2015. aastast tuleb hariduse asjus märkida nii paber- kui digitaalset avaldust täites üks järgmiste valikute vahel: põhihariduseta, põhiharidus, keskharidus (sh keskeriharidus), kõrgharidus. Hariduse ja töökoha info on olnud siiani ütluspõhine, st seda andmestikku ei pärita ühestki registrist vaid tugineb eeldusele, et kandidaat kannab avaldusele tõese info. Andmestiku ebakorrapärasuse tõttu ei ole võimalik seda kuidagi filtreerida ega sorteerida ning seetõttu pole mõistlik seda infot veebilehel säilitada.

Uurijatele on kandidaatide puhul huvi pakkunud kandidaatide ning valituks osutunute vanus või sugu. Kandidaatide keskmine vanus on valimisandmete veebilehel aastate kaupa avaldatud, ning vastavat statistikat on valimissündmuse kaupa tehtud ka 1992-2009. aastatel [4]. Valimiste veebilehe statistikaosas on iga aasta kohta toodud ka kandidaatide ning valituks osutunute sooline võrdlus. Maakondade ja valdade kaupa seda võrdlust nii ülevaatlikul kujul pole (v.a. KOV2002 ja KOV2005 valimissündmus, kus on vastavas statistikaosas leitud ka sooline jaotus maakonniti). Uurijatele võib aga huvi pakkuda näiteks hoopis alla 25-aastaste kandidaatide osakaal valimistel või valituks osutunute hulgas. Kandidaadi sünniaeg on reeglina veebis kuvatud, kuid masinloetav andmekuju teeks sellele küsimusele vastuse leidmise filtreerimise teel kiireks ja lihtsaks. Arvestades, et ka kandidaatide ja valituks osutunute teemal sel on RVTsse tulnud mitmeid päringuid, tuleb andmemudelisse luua väli ka soolise kuuluvuse määramiseks.

RVTI on juba olemas kaks kandidaatide infot sisaldavat XML-faili struktuuri, mis väljastavad VISi sisestatud kandidaadi andmed. Üks neist on mõeldud veebilehel kuvamiseks ning meediale edastamiseks. Selles failis sisaldub ka iga kandidaadi sooline kuuluvus. Veebis soolist kuuluvust ei kuvata, kuna seal see kandidaadi andmete juures enamasti nimest järeldatav, failis aga võimaldab kiiret sorteerimist. Teine kandidaatide fail on mõeldud asutusesiseseks kasutamiseks ning sisaldab kogu infot, mis kandidaat oma ankeedis avaldab, sh isikukood. 1999. aastast alates on kõik kandidaatide andmed - sh isikukood, mille abil sooline kuuluvus leida - sisestatud VISi, kuid varasematest aegadest tuleks selle info leidmiseks uurida paberkandjal kandideerimisavaldusi. Valmiv andmebaas peaks arvestama võimalusega, et see aeganõudev töö võetakse ette millalgi tulevikus, kuid esialgu jääks need andmed täitmata. Isikukood kuulub Andmekaitse inspeksiooni isikuandmete liigituse kohaselt tavaliste isikuandmete hulka [35], kuid veebis ega avalikustamiseks mõeldud failist jääb see mõistagi välja. Isikukoodi VARH andmebaasi salvestamiseks annab põhjuse lähteülesande nõue 1.c., mille järgi peab olema võimalik kandidaatide andmebaasi otsingu ja valimissündmuse kaupa alla laadida. Otsingu kaupa alla laadimise all on mõeldud praeguse kandidaatide otsingu lehel ilmutat: <http://vvk.ee/arhiiv/kandidaadi-otsing/>, kus tuleb otsinguks sisestada vähemalt nelja tähemärgiline otsisõna. Seal kuvatakse otsitud kandidaadi nimi,

valimissündmus, vald/linn (KOV), ringkonna number, kandidaadi number, nimekiri, saadud häälte arv ning see, kas kandidaat osutus valituks. Lähteülesandes esitatud nõude kohaselt peaks ilmuvat tabelit saama alla laadida, kuid praeguseks on RVTsse saabunud teabenõuete põhjal selge, et vajadus on kõigi valimiste kõiki kandidaate sisaldava andmefaili järele. Praegu toimub otsing kandidaadi nime järgi, kuid kui kandidaadi nimi on vahepeal muutunud, tuleb teha uus otsing uue nimega, et leida sama kandidaadi kõik valimised. Võimalus nime vahetanud kandidaatide kohta siiski kogu kandideerimisajalugu saada peaks toimuma siiski isikukoodi kaudu isikut ja valimissündmust ühendades ning seetõttu on vajalik isikukoodi salvestamine andmebaasi.

Aastate lõikes on erisusi ka selles, kas tühjaks jäänud väljad (andmed, mida kandidaat ei ole esitanud) kuvataksegi veebis tühjadena, või siis ei kuvata neid välju üldse (joonis 7). Iseenesest on tegemist infoga, millised andmed on kandidaat otsustanud jagamata jätta. Seda informatsiooni ei ole plaanis aga tulevasse andmebaasi kaasata.

KANDIDEERIMINE	KANDIDAADID
<p>EESTI REFORMIERAKOND</p> <p>Registreerimiseks esitamise kuupäev: 28.04.2004 Registreerimise kuupäev: Address: Tõnismägi 9, Tallinn, 10119 Sidevahendid: Esindajad: Kristen Michal Arto Aas</p> <hr/> <p>Kandidaat nr. 170 - SIGNE KIVI</p> <p>Erakond: Eesti Reformierakond Sünniaeg: 24.02.1957 Sünnikoht: Tallinn Amet: Riigikogu liige Töökoht: Riigikogu Haridus: Kõrgem Õppeasutus: ERKI, 1980 Eriala: Tekstiilikonstnik Teaduslik kraad:</p>	<p>Kandidaat nr 110 - TIIU SAATLA</p> <p>Sünniaeg: 24.01.1942 Amet: pensionär</p>

Joonis 7. Näited RK2003 ja KOV2005 kandidaadi andmetega

Viimasteks, 2019. aasta RK ja EP valimistel avalikustati vastavalt valimisseadustele [13][15] veebilehel kandidaadi kohta*:

- Registreerimisnumber
- Ees- ja perekonnanimi

- Sünniaeg**
- Kodakondsus (ainult EP ja KOV valimistel EL liikmesriigi kodaniku puhul)
- Erakondlik kuuluvus
- Kontaktandmed
- Haridus
- Töökoht ja amet

**KOV2017 aasta andmetest on kandidaadi kohta muu andmestik peale registreerimisnumbri, nime ja erakondliku kuuluvuse kustutatud.*

***Sünniaja avaldamist ei nõua seadus, kuid see võimaldab eristada samanimelisi kandidaate*

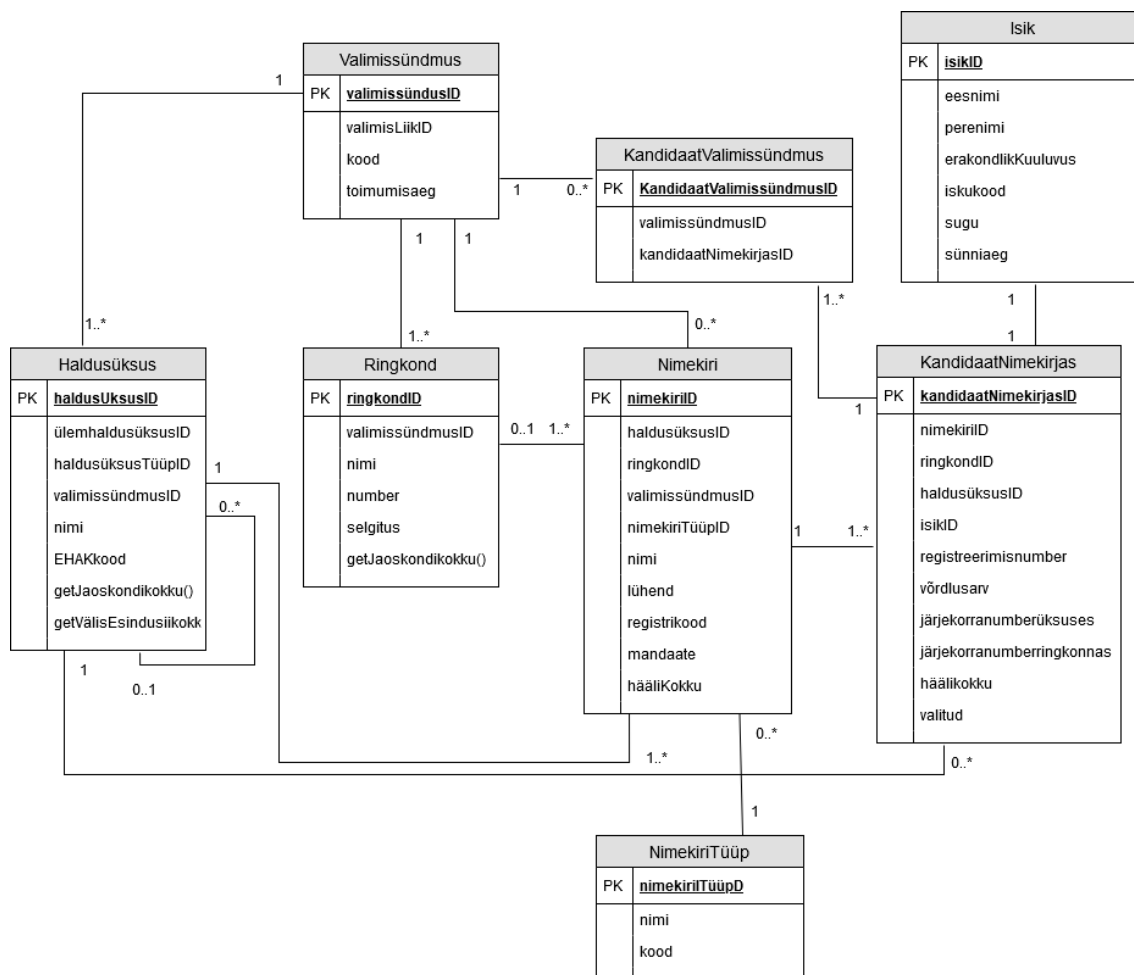
RVTs on vastu võetud otsus jätta eelnev kandidaadi andmestik tulevikus valimiste veebilehele kuueks kuuks pärast valimisi, hiljem aga kuvatakse ainult valimistega seotud andmed:

- Registreerimisnumber
- Ees- ja perekonnanimi
- Sünniaeg
- Erakondlik kuuluvus (võib sõltuvalt valimisliigist olla erinev nimekirjast, milles kandidaat kandideerib või ka puududa)

Eelnev andmestik on selles vaates, mis avaneb veebilehel kandidaadi nimel klikkides. Veebilehtedel on aga valimisüksuste struktuuripuud lahti klikkides arusaadav ka ringkond ja nimekiri, kus kandidaat kandideerib. Avaandmete kontekstis tuleb see info samavõrra arusaadavalt esitada. Kuna VARH andmebaas esitab valimiste järgset infot, saab iga kandidaadi juures lisaks kuvada ka talle antud häälte arvu, võrdlusarvu (kui see on arvatud) ning seda, kas kandidaat osutus valituks.

Kandidaadid on erakonna sees jaotatud nimekirjadesse ning kuna positsioon selles võib valituks osutumise küsimuses mängida suurt rolli, pakub see huvi ka uurijatele. Kui veebilehel on see struktuuripuus jälgitav, siis andmefaili tuleb see juurde lisada. EP

valimissündmuse korral on kandidaat ainult ühes, üleriigilises nimekirjas, kuna EP valimiste korral on kogu Eesti üks ringkond (joonis 5), RK valimissündmuse korral aga ringkonnanimikirjas ning üleriigilises nimekirjas. KOV valimissündmuse korral on kandidaat ringkonnanimikirjas, Tallinna puhul on olemas ka erakondade/valimisliitude ülelinnalised nimekirjad. Klassimudel joonisel 8 esitab eespool leitud vajaduste kohaselt nimekirjade ja kandidaatide andmestiku. Kandidaadid on ainult RK, EP ja KOV valimissündmusel. Kandidaat saab kandideerida valimistel ainult ühes koosseisus - EP valimistel erakonnas, RK valimistel erakonnas (kuni 1995. aastani said RK valimistel osaleda ka valimisliidud) ning KOV valimistel erakonnas või valimisliidus, või siis üksikkandidaadina. Neid kooslusi on otstarbekas hoida tabelis nimega “Nimekiri”, kus koosseisu liik - erakond, valimisliit, üksikkandidaadid - on päritav teisest tabelist. Üksikkandidaatide puhul on nimekirjaks “Üksikkandidaadid”, mis moodustatakse igas ringkonnas, kus üksikkandidaate on. Võimaldamaks samanimelise erakonna valimistel osalemise ajaloo leidmist, tuleb andmebaasi salvestada selle registrikood, mis jääb samaks ka siis, kui erakonna nimi peaks muutuma. Registreerimiskoodid on Äriregistris avalikud [36] ning need saab lisada ka avaandmete failis nimekirja (erakonna, valimisliidu) koosseisu. Üksikkandidaatide puhul jääb see väli tühjaks. Tabelis “Kandidaat” täidetakse väli haldusüksusID KOV valimiste korral selle KOVi infoga, kus kandidaat kandideerib, EP ja RK valimiste korral on selleks aga Eesti Vabariigile vastav haldusüksuseID.



Joonis 8. Klassimudel nimekirjade, kandidaatide andmete ja seostega

Täpsustused nõudele 1.b.2:

Valimisandmed, mida on kandidaatide kohta võimalik veebilehel ühe valimissündmuse kohta näha ning alla laadida on järgmised:

- Kandidaadi nimi
- Registreerimisnumber
- Sugu*
- Sünniaeg*
- Erakondlik kuuluvus (juhul, kui vastav väli kandidaatide juures puudub, sisestatakse “Andmed puuduvad”, kui väli on täitmata - “Puudub”)
- Nimekiri, kus kandidaati kandideerib

- Ringkond
- KOV (KOV valimissündmuse korral)
- Järjekorranumber ringkonnas
- Järjekorranumber üksuses
- Saadud häälte arv
- Võrdlusarv (juhul, kui see on vastavalt arvutatud, vastasel korral jääb väli tühjaks)
- Valituks osutamine

Need andmed arvatakse 1.e. nõude järgi valimisandmete koosseisu.

Täpsustused nõudele 1.c. ja 1.d.:

Kasutajal on võimalik alla laadida kõigi 1992-... valimissündmustel osalenud kandidaatide faili, milles on iga kandidaadi kohta:

- Nimi (kui see on muutunud, siis nimi vastavalt valimissündmusele)
- Sugu*
- Sünniaeg*
- Valimissündmus
 - Registreerimisnumber
 - Ringkond (nimi ja number)
 - KOVi nimi (KOV valimissündmuse korral)
 - Nimekiri
 - Häälte arv
 - Valituks osutamine

** Leitakse isikukoodi järgi, mis 1999. aastast varasemate valimissündmuste korral leitakse RVT arhiivdokumentidest. Kuni vastav sisestus ei ole tehtud, kuvatakse väli tühjana.*

Samasugust andmekomplekti peab kasutaja saama vaadata ja alla laadida kandidaadi otsingus sisestatud ja leitud kandidaadi kohta. Andmetest ei moodustu üldtabelit ning sellele ei rakendu nõue 1.g.

2.3. Hääletamisest osavõtu statistika

Hääletamisest osavõtu statistika tabel on tänaseks kujunenud vahendiks, mis annab juba valimisperioodil infot, kui paljud valijad on maakonna, Tallinna ja Tartu vaates valimas käinud ning kui palju on e-hääletajaid. Siiani kasutusel olnud *as-is* protsess on esitatud joonisel 3, kuid seoses 2021. aastast kasutusele võetava elektroonilise valijate nimekirjaga on ka statistika kuvamise osas oodata ümberkorraldusi [13b]-[16b]. Elektrooniline valijate nimekiri võimaldab igal ajahetkel teada täpselt hääletanute ja e-hääletanute arvu, kuid kokkuvõtete, meedia ning andmete võrdluse huvides tuleb siin leppida kokku ajahetk, mil salvestatakse päeva jooksul hääletanute hulk.

Hääletamisest osavõtu statistika kõnekaim number - osavõtu protsent - leitakse valijate arvu (hääleõiguslike isikute) ja hääletanute (hääle andnud isikute) arvu võrdlusest.

Osavõtuandmete tabeleid on esitatud alates uuritava perioodi algusest 1992. aasta RK valimistest. Kuni 2003. aastani esitati RK valimiste korral andmeid ühtse tabelina ringkonniti, KOV valimiste korral maakonniti. Tabelis esitati valijate ja hääletajate arv ning osavõtu protsent. RK korral on maakondade info leitav eraldi, ringkondade kaupa jaotatud lehtedelt. Edasi tuli Eestis juurde aga kaks seni olemata valimisliiki - rahvahääletus 2003. aastal Euroopa Liitu astumise üle ning sellele järgnes Euroopa Parlamendi valimine 2004. aastal. Mõlema valimisliigi korral on kogu Eesti üks ringkond ning ka sel puhul on osavõtuandmed esitatud maakondade kaupa ning edaspidi juba kõigi valimisliikide korral ainult protsendiliselt. Kuna neljast valimisliigist kolme puhul esitati nüüd osavõtuandmed ühe ülevaatliku tabelina maakonniti, on ka RK valimiste korral kuvatud edaspidi osalust maakonniti ning ringkondade valijate ja hääletajate arv on liikunud detailsete andmete alla, nagu see on ka KOV valimiste korral.

Alates 2003. aasta rahvahääletusest uuendatakse valimiste veebilehel osavõtustatistikat kokku kuuel ajahetkel: eelhääletuse päeval pärast jaoskondade sulgemist (joonis 2) ning valimispäeval kolm korda päevas. Osalusprotsenti on kuvatud varem nii ühe kui kahe kohaga peale koma. Meediale väljastatav XML-fail sisaldab lisaks ka hääletanute

arvu numbrilisel kujul. Edaspidi peab osavõtustatistika tabel pakkuma andmeid nii arvu, kui protsendi kujul, et uurijad saaksid kätte just endale vajaliku täpsusega andmestiku.

Alates 2005. aastast on andmehulgaga liitunud e-hääletajad. KOV2005 valimistel on need osavõtu andmetes liidetud iga maakonna hääletate arvule, kuid juba järgmistest, RK2007 valimistest kuvatakse need kogu Eesti vaates eraldi andmehulgana. E-hääled antakse ringkonniti, kuid 2013. aastal lisati lahendus, kus e-hääli võib näha ka jaoskondade kaupa (sellest lähemalt peatükis 2.5.). Jaoskondade e-hääletajate summadest on võimalik summeerida vastava KOVi või maakonna e-hääletajate arv.

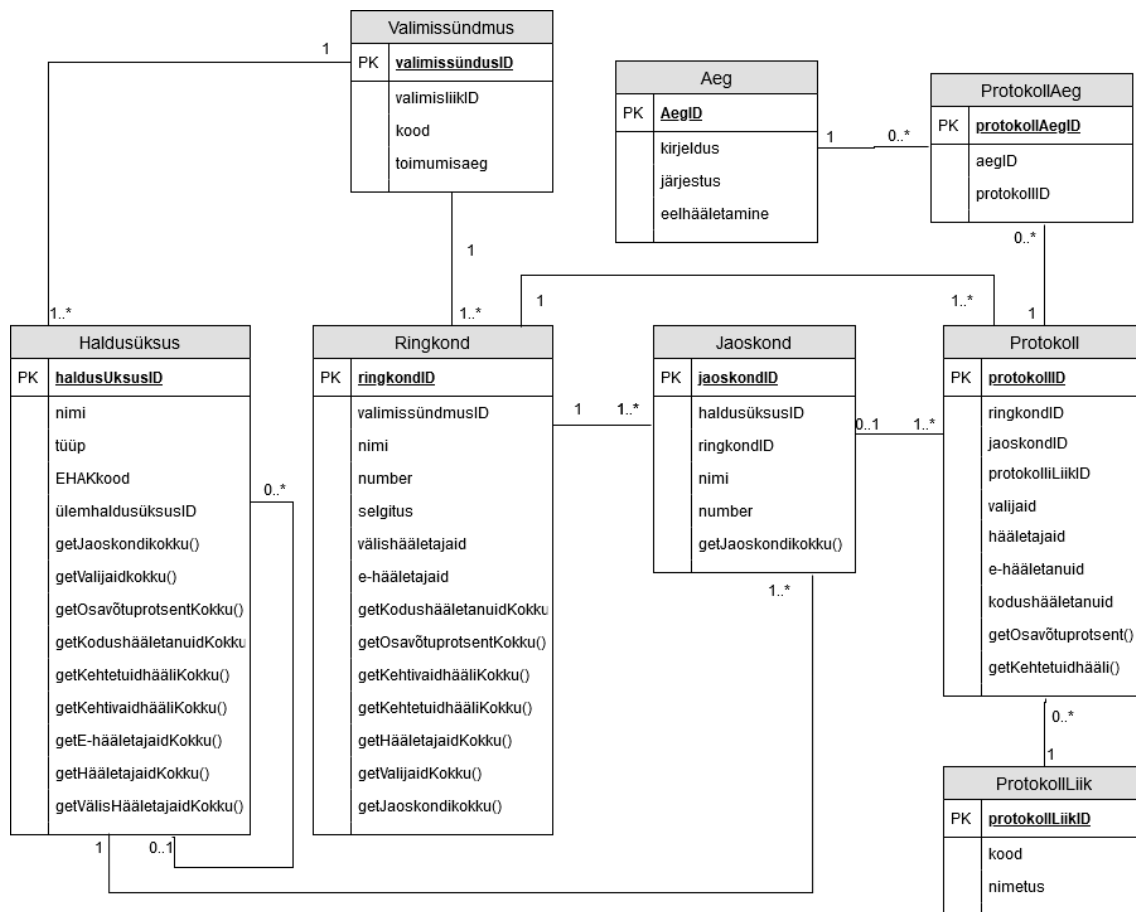
Sarnaselt e-häälele arvestatakse ringkonniti ka välishääli (ainult RK, EP ja RH valimisliigi korral). RK valimissündmuse korral lisatakse välisriigis pabersedeliga antud hääled igas ringkonnas ühe kokkulepitud maakonna arvestusse, EP valimissündmuse korral lisatakse need Tallinna häälele.

Hääletanute arv jaguneb veel omakorda kodushääletajateks, eelhääletajateks, välishääletajateks jne. See, kas need numbrid on teada, sõltub ajahetkest. Need arvud leitakse reeglina valimispäeva õhtuks ning on siiani kuvatud ainult EP, RK ja RH üleriigiliselt ning kuvatav andmekomplekt on valimiseti olnud erinev. Viimasteks valimisteks on see kujunenud järgmiseks:

- Valijaid
- Hääletanuid
- Kehtivaid sedeleid
- Kehtetuid sedeleid
- Eelhääletanuid
- Kodus hääletanuid
- Elektrooniliselt hääletanuid
- Välisriikides hääletanuid

VARH andmebaas peab arvestama võimalusega, et kui neid andmeid ei leita veebilehtedelt, siis sisestatakse need hiljem protokollide andmetest ning tulevikus ka

KOV valimiste kohta. Seni jäävad täitmata andmeväljad tühjaks ning selgitav info kuvatakse veebilehel. Klassimudel joonisel 9 näitab, et mõned protokollid - välishäälte protokoll RK, EP ja RH valimissündmuste korral, e-häälte protokoll - ei ole seotud kindla jaoskonnaga vaid ainult ringkonnaga ning nende andmed sisestatakse selleks määratud tasemel. Jaoskonnad saavad oma jaoskonnaga seotud e-hääletajate arvu teada aga pärast eelhääletamise lõppu. Üleriigilised osavõtuarvud leitakse Haldusüksuse tabelist, kui haldusüksuseks on Eesti Vabariik. Osavõtuprotsendid arvutatakse valijate arvu ja hääletanute arvu järgi, et saada ühtlane tulemus: varasemate valimiste korral on seda arvutatud kahe komakoha täpsusega, hilisematel ühe komakohaga. RK, EP ja RH korral on välishäälte protokollil puhul tegemist protokollil eriliigiga, kus täidetakse ainult hääletajate osa.



Joonis 9. Klassimudel osavõtuandmete kuvamiseks vajalike tabelite ja seostega

Täpsustused nõudele 1.b.2:

Valimisandmed, mida on hääletamisest osavõtu kohta võimalik veebilehel ühe valimissündmuse kohta näha ning alla laadida on järgmised:

- Osalusstatistika eelhääletamise kohta
 - Eelhääletanute arv ja protsent maakonniti kokku kuuel ajahetkel (alates 2003. aastast)
 - Eelhääletanute arv ja protsent e-hääletajaid (alates 2005. aastast)
 - Eelhääletanute arv ja protsent Eestis kokku (alates 1993. aastast)
- Detailne osalusstatistika
 - Valijate arv
 - Hääletajate arv
 - Kehtivaid sedeleid
 - Kehtetuid sedeleid
 - Eelhääletanuid (alates 1993. aastast)
 - Kodus hääletanuid (alates 1999. aastast)
 - Elektrooniliselt hääletanuid (alates 2005. aastast)
 - Välisriikides hääletanuid (RK, EP ja RH valimissündmuste korral)
 - Iga maakonna kohta:
 - Valijate arv
 - Hääletajate arv
 - Kehtivate häälte arv
 - E-häälte arv (alates 2005. aastast)
 - Osalusprotsent
 - Välisriikides alaliselt elavate hääletajate kohta (ainult EP valimiste korral):
 - Hääletajate arv
 - Kehtivate häälte arv
 - E-häälte arv (alates 2005. aastast)
 - Iga ringkonna kohta (ainult RK valimistel)
 - Iga ringkonnas asuva haldusüksuse kohta

- Valijate arv
- Hääletajate arv
- Osalusprotsent
- Ringkonna e-häälte kohta
 - Hääletajate arv
- Ringkonna välishäälte kohta
 - Valijate arv
 - Hääletajate arv
- Iga KOV kohta
 - Valijate arv
 - Hääletajate arv
 - Kehtivate häälte arv
 - E-häälte arv (alates 2005. aastast)

Puuduvad andmehulgad - e-hääletajad enne 2005 aastat jätta allalaaditavas andmefailis tühjaks. Põhjus tuleb avaldada kasutajale infoteates. Kõik loetletud andmehulgad esitatakse veebis tabelitena ning neile rakendub nõue 1.f. ja 1. g. Üldtabelid moodustuvad loetletud üksuste kaupa.

2.4. Hääletamis- ja valimistulemus

Hääletamis- ja valimistulemuse all on siin peatükis mõistetud mõistetud mitut andmehulka. Tegemist on RK ja EP valimiste korral esiteks ülevaatega valimistulemustest, st sellest, mis nimekirjad on saanud mandaate, teiseks sellest, millised kandidaadid või üksikkandidaadid on osutunud valituks. KOV valimiste korral on see jaotus nime all “Hääletamistulemus erakonniti”. Neid andmehulki on eri aegadel liigitatud veebilehel eri pealkirjade või alamlehtede alla ning ka visuaalselt on kujundus liikunud tabelite ning graafikute vahel. Andmehulgad on valimisliigiti erinevad, kuid neisse hulkadesse jagamine on aja jooksul olnud võrdlemisi ühtlane: iga valimisliigi korral esitatakse samasugune andmestik. Teistest valimisliikidest oluliselt erinev on hääletamis- ja valimistulemus RH valimisliigis, mil esitatakse jah- ja ei-häälte arv ja protsent nii KOVides, maakondades kui üleriigiliselt. Vaadeldava perioodi

rahvahääletused toimusid 1992. ja 2003. aastal, mil e-hääletamine polnud veel juurutatud. Tulevaste samasuguste sündmuste jaoks tuleb aga andmemudelisse vastav lahendus lisada.

Alates 2015. aastast on valimis- ja hääletamistulemuste juures ka mäрге: “Andmed seisuga ...”, mis markeerib ajahetke, mil valimistulemus on arvutatud. See andmehulk omab aga tähtsust üksnes valimispäeva õhtul ning pärast häälte teist lugemist, sest hiljem andmetes enam muutusi ei toimu. Sama tähendusrikas on valimistulemuste väljaselgitamise ajal samas esitatud andmestik mitu jaoskonda on tulemused edastanud ning jaoskondade arv kokku, mis võimaldab vaatajal oletada, kas graafikus nähtav tulemus võib veel palju muutuda. Eraldi kuvatakse ka, kas e-hääled on edastatud või mitte. Valimiste järgselt, kui kõik jaoskonnad on andmed edastanud ja e-hääled arvestatud, ei oma see andmestik enam tähtsust. Järgnevalt analüüsitakse hääletamis- ja valimistulemuste andmeid iga andmehulga kaupa.

2.4.1. Hääletamis- ja valimistulemus erakonniti

Tegemist on kõige ülevaatlilikuma andmestikuga, mis näitab (olenevalt valimissündmuse lehe disainist graafiku või tabelina), kui palju said erakonnad hääli ning mandaate. RK valimiste korral jaotatakse 101 mandaati erakondade või üksikkandidaatide vahel vastavalt valimistulemusele, EP valimiste korral jaotatakse 6 (2019. aastani) või 7 (alates 2019. aasta valimistest) mandaati. KOV valimiste korral leitakse mandaadid KOVi siseselt.

Eesti Vabariigi tasemel esitatakse iga nimekirja kohta saadud häälte arv, protsent kõigist kehtivatest häältest, saadud mandaatide arv ning alates 2003. aastast võrdlusena sama nimekirja häälte ja mandaatide arv sama valimisliigi eelmise valimissündmuse korral, lisaks kehtivate häälte arv kokku. Samal ajal on viidud kokku varem eraldi lehtedel asunud andmehulgad - “Valimispäev” nimekirjadele antud häältega ning “Valimistulemus” ainult mandaadi saanud nimekirjade andmetega, kuna sisuliselt oli tegemist dubleerimisega. Selles andmehulgas on sellest ajast hakatud kuvama ka neid erakondi, kes mandaate ei saanud (EP valimiste korral vastavalt 2004. aastast). Eesti

Vabariigi tasemel asuvad üksikkandidaadid RK valimissündmuse korral ühes andmehulgas koos, EP valimiste korral nimeliselt ning eraldi.

KOV valimiste korral, kus nimekirjadega osalevad nii erakonnad üleriigiliselt kui valimisliidud KOVi siseselt, on Eesti Vabariigi tasemel vastavas tabelis alates 2005. aastast esitatud ainult üleriigiliselt või mitmetes KOVides osalenud erakonnad.

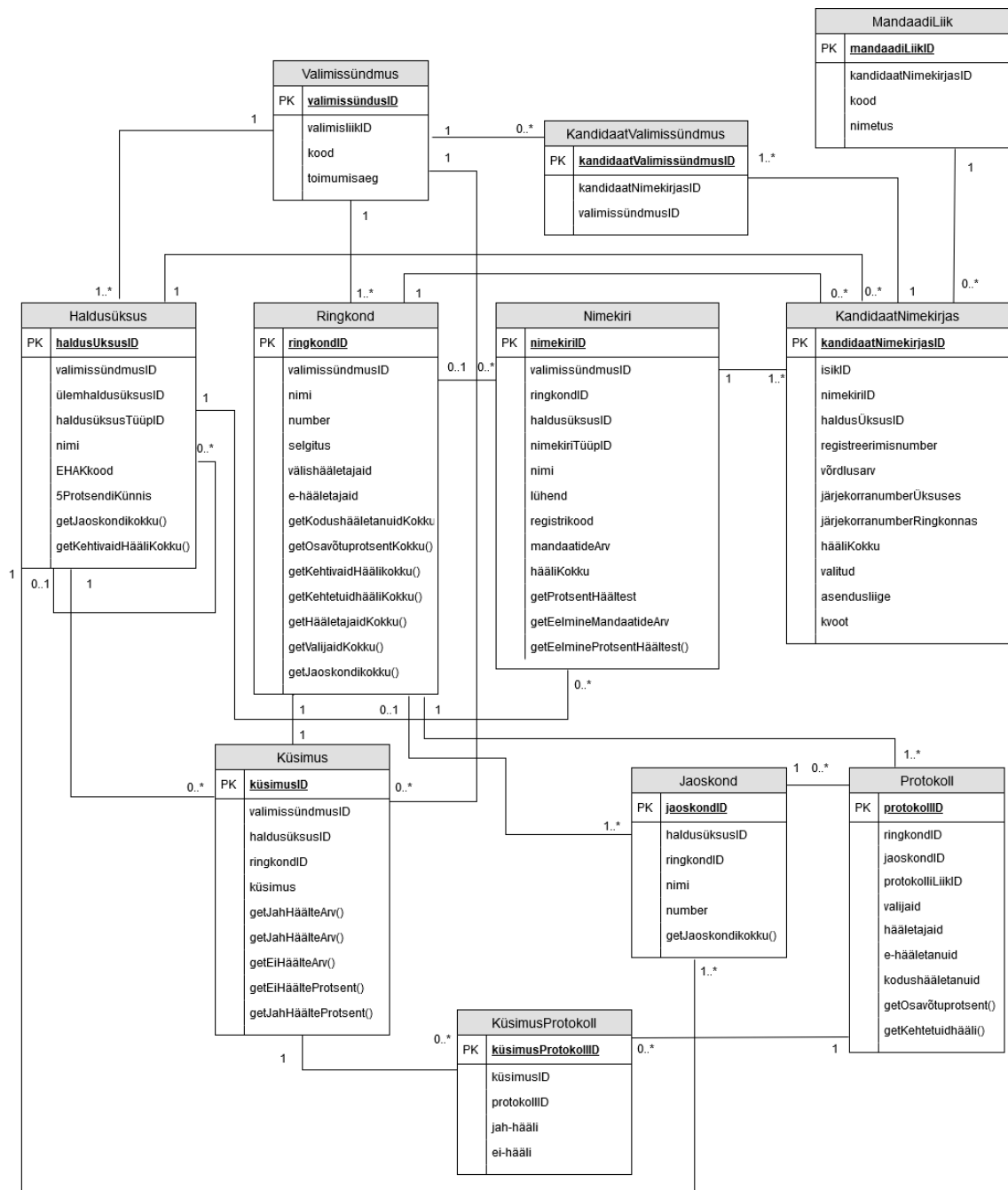
Eesti Vabariigi tasemest allpool on RK valimiste korral andmestikku esitatud väiksemate haldusüksuste - üldjuhul maakondade - kaupa, 2011. aastast aga ringkondade kaupa. Kuna häälte arv mingile nimekirjale summeeritakse jaoskonnas antud pabersedelitest, välisriikides antud häältest alates 2005. aastast e-häältest, siis on võimalik seda andmestikku kokku võtta nii maakonna kui ringkonna tasemel. Maakonnatase on ühest küljest üldkasutajale arusaadavam, kuid kuna RK valimiste korral asub Tallinnas kolm ringkonda, on tekkiv nimekiri kujul "Harjumaa, Hiiumaa, Tallinn 1, Tallinn 2, Tallinn 3, ...". Ringkondade kaupa kuvamisel tekib ringkondade nimekiri, kus on lisaks ringkonna numbrile olemas ka ringkonna kirjeldus, näiteks "Valimisringkond number 4 - Harjumaa, Raplammaa". Erakondade kaupa valimistulemusi esitades tuleb selle selguse huvides eelistada ringkondade kaupa kuvamist. Andmete hulk jääb seejuures samaks.

EP ja KOV valimiste korral kuvatakse üleriigiliselt osalenud nimekirjade andmed maakonniti ning RK valimistega samas hulgas. KOV valimiste korral on lisaks veel üks andmehulk "Valimistulemus suuremates keskustes", kus esitatakse sama andmestik nimekirjade kohta linnades ning suuremates valdades. Tegemist on laiale kasutajaskonnale mõeldud andmestikuga, kus valimistulemuste hulgast on linnad esile toodud. Avaandmete vastavas failis on need võimalik välja filtreerida ning eraldi suuremate keskuste andmeid sisaldava faili jaoks vajadus puudub.

Mandaatide arvu võrdluse saab esile tuua ainult RK ja EP valimisliikide korral, kuna KOV valimistel erineb aastati osalevate erakondade ja valimisliitude komplekt tunduvalt. RH valimiste korral kuvatakse Jah- ja Ei-häälte arv ja protsent Eesti Vabariigis, maakonniti ja KOVi kaupa. RK ja EP korral täidetakse tabelis Nimekiri

mandaatideArv ja getEelmineMandaatideArv() (sama nimekirja mandaatide arv eelmise samaliigilise valimissündmuse korral - veebis kuvatud aastast 2003), kuna mandaadid leitakse üleriigiliselt. Alates 2011 aastast on RK ja KOV valimisliikide korral kuvatud alati ka 5% künnis. RH valimisliigi korral ei täideta nimekirjade ega kandidaatidega seotud andmestikku. Üle-eestiliselt ja maakonniti kuvatakse jah- ja ei-häälte arv ja protsent. KOV valimiste korral kuvatakse tulemus nimekirjade kaupa KOVis lisaks ka detailandmete koosseisus, kuid kuna seal lisandub ainult üks andmeväli - nimekirja protsent kõigist ringkonna mandaatidest, siis on see lisatud joonisel 10.

Seni toimunud RH sündmustel aastatel 1992 ja 2003 ei olnud veel e-hääletamist. E-hääled ja välisääled lisatakse RH korral ringkonnale, milleks RH valimiliigis on kogu Eesti ning avaandmete failis on see lahendus sellisena loodud.



Joonis 10. Klassimudel valimistulemuse ja mandaatide jaotuse andmehulkadega

Täpsustused nõudele 1.b.2:

Valimisandmed, mida on hääletamistulemuste kohta võimalik veebilehel ühe valimissündmuse kohta näha ning alla laadida on Eesti Vabariigi kohta tervikuna kui RK korral iga valimisringkonna ning EP ja KOV valimiste korral iga maakonna, KOV korral lisaks ka suuremate keskuste kohta järgmine:

- Nimekirja nimi (erakond või valimisliit, RK korral lisaks andmehulk “Üksikkandidaadid”, EP korral iga üksikkandidaat eraldi)
 - Nimekirjale (EP korral üksikkandidaadile) antud häälte arv
 - Protsent kõigist häältest
 - Saadud mandaatide arv (ainult RK ja EP valimiste korral)
 - Võrdlusena sama nimekirja sama valimisliigi eelmise valimissündmuse protsent häältest
 - Võrdlusena sama nimekirja sama valimisliigi eelmise valimissündmuse mandaatide võrrelduna käesoleval sündmusel saadud mandaatide arvuga (+märgiga, kui mandaate saadi juurde, -märgiga, kui kaotati, ainult EP ja RK korral)
- 5% künnis (s.o. 5% üleriigiliselt kehtivatest häältest. Alla selle hääli saanud erakonnad / üksikkandidaadid ei osale mandaatide jagamisel. 5% künnis arvutatakse RK korral kõigist kehtivatest häältest, KOV valimiste korral vastava KOVi kehtivatest häältest)

RH valimisliigi korral

- Eesti Vabariigis, maakonnas, KOVis, eraldi pabersedeliga-hääletajate hulgas, e-häälte hulgas (alates 2005. a.), välishäälte hulgas
 - Jah-häälte hulk
 - Jah-häälte protsent
 - Ei-häälte hulk
 - Ei-häälte protsent

Andmed, mida vastava alajaotuse veebilehtedelt korjata ei saa, summeeritakse detailsetest hääletamisandmetest, kasutades sama arvutusloogikat, mis nende valimissündmuste korral, kus vastav andmestik on olemas. Need andmed on:

- Valimistulemus nimekirjade kaupa ringkonniti RK valimiste korral enne 2003. aastat
- Üleriigiliselt osalenud nimekirjade vastav andmestik KOV valimiste korral enne 2005. aastat

Loetletud andmed saab esitada tabelite kujul ning neile rakendub nõue 1.f. ja 1. g. Üldtabelid moodustuvad loetletud üksuste kaupa.

2.4.2. Valituks osutunud, asendusliikmed

Valituks osutunute andmehulk esitab vastavalt valimissündmusele mandaadi saanud kandidaate. RK valimiste korral kuvatakse 101 valituks osutunud või asendusliikmeks jäänud kandidaadid oma erakonna kaupa sorteeritult. Alates 2011. aastast on sealsamas kuvatud ka erakonna mandaatide arv kokku, et seda ei peaks otsima teiselt lehelt. Samamoodi erakondade kaupa kuvatakse asendusliikmed. KOV valimistel valitakse kohaliku omavalitsuse volikogu liikmed ning asendusliikmeid valimiste veebilehele ei kuvata. EP valimistel valiti enne 2019. aastat, mil Suurbritannia otsustas Euroopa Liidust lahkuda, Eestist 6 esindajat, 2019. aasta valimistel aga 7, kusjuures üks valitu jäi nõ. ootele kuni Suurbritannia lõpliku lahkumisotsuseni. Kuigi valituks osutunute arv on väike, esitatakse hääletustulemuste andmeid mitmete eri väljadega, võrdlusarvu või häälte arvu järgi reastatult.

Valimistulemus - sh valitud liikmed, asendusliikmed - arvutatakse VISis ning neid andmeid võetakse konstantidena. Kandidaatidele arvutatud võrdlusarvud ning kvoodid tulevad samuti kas impordi teel VISist või HTMLidelt migreerituna. Klassimudel esitatakse järgmise peatüki koosseisus, kuna eelnevale lisatavate andmete hulk on väike.

2.4.3. Mandaatide omandamise viis

RK valimiste korral saab mandaadi omandada kolmel viisil: isikumandaadina, ringkonnamandaadina või kompensatsioonimandaadina ning nende omandamise juures on olulised eri arvud - kvoot või võrdlusarv [37]. KOV valimiste korral leitakse Tallinnas erandina samuti isiku-, ringkonna- ja kompensatsioonimandaadid, teiste KOVide korral saab mandaat olla isiku- või nimekirjamandaat [38]. Nimekirjamandaat saadakse võrdlusarvu järgi. KOV valimistel ei kuvata aga välja, mis liigilise mandaadi on kandidaat saanud, vaid see on järeldatav kvootide ja võrdlusarvude kaudu. Mandaadi liik leitakse VISis.

Täpsustused nõudele 1.b.2:

Valimisandmed, mida on valitud liikmete, asendusliikmete ja kohta võimalik veebilehel ühe valimissündmuse kohta näha ning alla laadida on Eesti Vabariigi kohta tervikuna kui RK korral iga valimisringkonna ning EP ja KOV valimiste korral iga maakonna, KOV korral lisaks ka suuremate keskuste kohta järgmine:

Isikumandaadi ja ringkonnamandaadi saanute hulgas kuvatakse:

- Järjekorranumber
- Kvoot
- Ringkond
- Nimekiri (erakond)
- Hääli
- Registreerimisnumber
- Kandidaadi nimi

Kompensatsioonimandaadi (RK või KOV Tallinna erandi korral) ja nimekirjamandaadi (KOV valimistel KOVides) saanute hulgas kuvatakse:

- Järjekorranumber
- Võrdlusarv
- Nimekiri
- Hääli
- Registreerimisnumber
- Kandidaadi nimi

Valituks osutunud kandidaadid eristatakse atribuudi “valitud” või “asendusliige” alusel. Mandaatide omandamise viis leitakse tabeli MandaadiLiik abil (joonis 10).

Loetletud andmed saab esitada tabelite kujul ning neile rakendub nõue 1.f. ja 1. g. Üldtabelid moodustuvad mandaadi liikide kaupa.

2.5. Detailne hääletamistulemus

Detailsete hääletustulemuste all esitatakse andmed kõige detailsemal kujul - igale kandidaadile antud häälte arv jaoskondade kaupa KOV tasemel, EP ja RK valimissündmuste korral ka maakonna tasemel ja üleriigiliselt. EP valimissündmuse korral on eraldi andmehulgaks ka kandidaatidele antud välishääled, RK korral lisatakse need igale ringkonnale ning kuvatakse eraldi. Eesti Vabariigi kohta on esitatud viimastel aastatel üldstatistika andmestik ning hääletamisest osavõtt ringkondade kaupa. Ringkondade kaupa esitatakse samuti statistilised andmed (üksuse nimi, valijate arv, häälte arv ning häälte protsent) ringkonda kuuluvate haldusüksuste (üldjuhul maakondade, Tallinnas linnaosade) kaupa, kuid neid statistilisi andmeid on siin töös käsitletud osavõtustatistika peatükis. Näiteks kui ringkonda nr 9 kuuluvad Tartumaa ning Jõgevamaa, siis esitatakse need andmed maakondade kohta, lisaks veel eraldi e-hääled ning välisriikides antud hääled EP, RK ja RH valimissündmuse korral.

2.5.1. Häälte jagunemine

Hääletamistulemused esitatakse vastavalt valimisliigile järgmiselt: RK valimissündmuse korral ringkonniti ka KOVi kaupa (Tallinnas linnaosade kaupa) ning jaoskonniti, EP korral üle-eestiliselt, maakonniti, KOVi kaupa (Tallinnas lisaks linnaosad) ja jaoskondade kaupa ning KOV valimistel KOVi ja jaoskondade kaupa. Alljärgnevalt uuritakse neid hulki eraldi.

Tegemist on kõige detailsema andmestikuga valimisandmete hulgas, mis esitab igale kandidaadile antud häälte arvu jaoskondade kaupa. Kui e-valimised toimuma hakkasid, esitati e-hääled esialgu ainult ringkonna tasemel. 2013. aasta KOV valimistest saadik kuvatakse e-hääli ka jaoskonna tasemel, mis võimaldab neid summeerida ka KOVi ja maakonna kaupa.

Eesti Vabariik (ainult EP)

Eesti Vabariigi tasemel saab siin andmehulgas kandidaatide kaupa hääletamistulemust esitada ainult EP valimiste korral, kuna siis osalevad kõik kandidaadid üleriigiliselt, kuna Eesti on üks ringkond. Kuna kandidaat saab kandideerida ainult ühes ringkonnas, siis teiste valimisliikide korral selline vaade võimalik ei ole. Kandidaadid on erakondade kaupa ning reastatud registreerimisnumbrite järgi. Nimekirjad kuvatakse selles järjekorras, nagu on enne kandidaatidele registreerimisnumbrite andmist toimunud liisuheitmisel leitud, mistõttu nimekirjal peab andmebaasis olema olema ka väli selle numbri jaoks. Liisuheitmine nimekirjade järjestuseks juurutati 2004. aastal - samal ajal, kui Eestis hakkasid toimuma EP valimised ning enne seda olenes nimekirjade kuvamise järjestus kandideerimisdokumentide esitamise järjekorrast. Seetõttu on ka vastava atribuudi nimeks klassimudelil "järjekorranumber". Üksikkandidaadid asuvad eraldi viimases andmehulgas. Alloleva klassimudeli järgi saab kandidaatidele antud häälte eri liigid (jaoskonna hääl, e-hääl, välishääl) eristade protokollide liikide kaudu.

Ringkonnas (RK)

Ringkondade kaupa, kus igasse ringkonda kuulub üks või mitu maakonda või linnaosa (Tallinnas) esitatakse kandidaatidele antud häáli RK valimiste korral. Sel juhul kuvatakse erakonnad samuti liisunumbrite või järjekorranumbri järgi reastatult ning kandidaadid häälte arvu järgi reastatult. Alates 2011. aastast on andmestik parema ülevaate saamiseks kahe tabelina - esimeses on nimekirjad ning neile antud hääled ringkonna üksuste kaupa, teises kandidaatide kaupa. Need andmed on leitavad juba eelmise klassimudeli põhjal, eristades tabeli KandidaatProtokoll antud hääled ringkonnaID ja haldusüksusID järgi, e-hääled ja välishääled protokollide liigi järgi.

Maakond (ainult EP ja RK)

Maakonniti eristuvaid hääletustulemusi kandidaatide kaupa saab esitada EP ja RK valimistel. Mõlemal puhul tuuakse sarnaselt eelmistele vaadetele reastatud

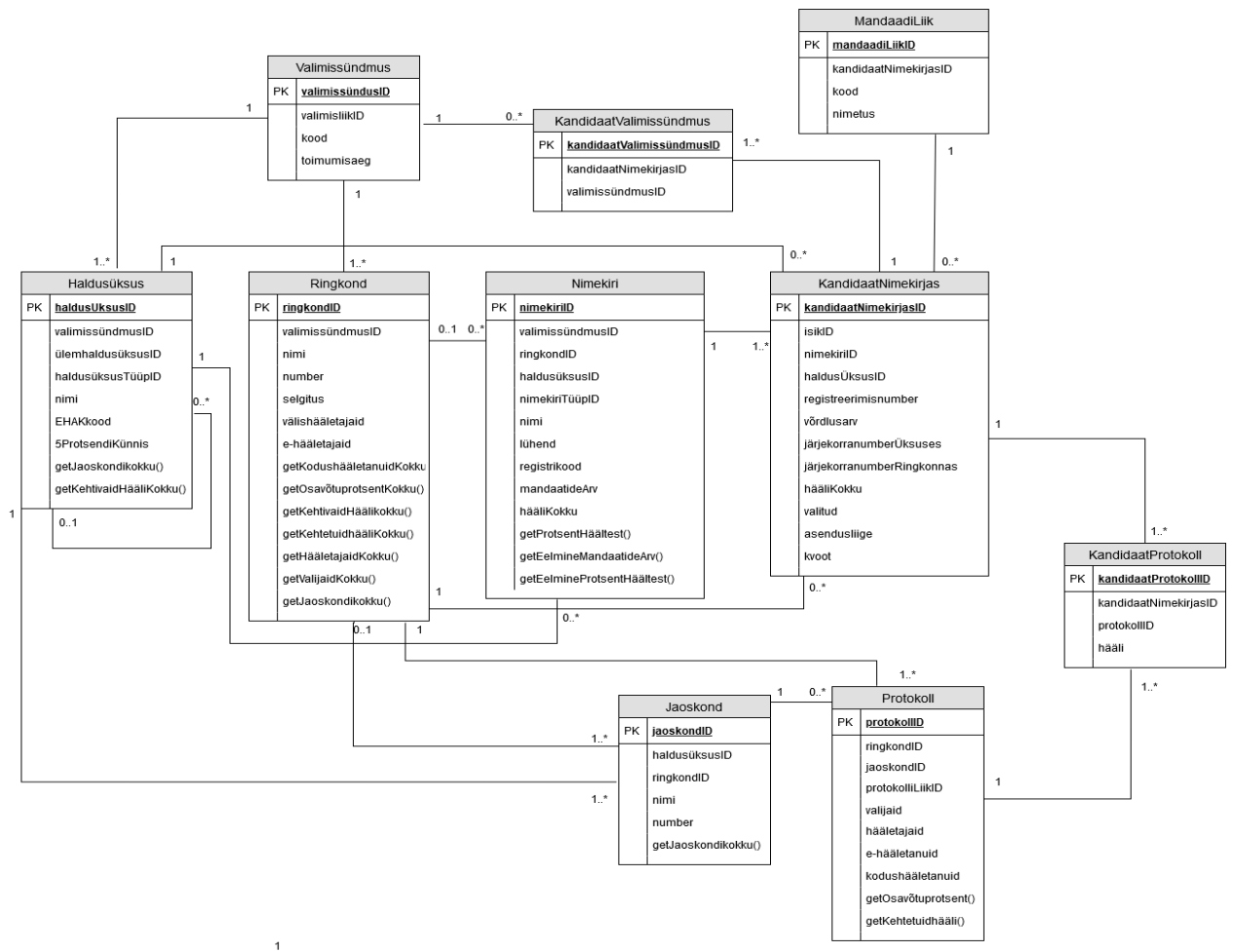
nimekirjadele ja kandidaatidele välja kehtivad (s.o. kandidaadile antud) hääled kokku, EP valimiste korral ka osalejate hulk (kehtivad + kehtetud hääled) ning osavõtuprotsent vastavas maakonnas või KOVis. Osavõtustatistika on siin töös juba kaasatud vastavasse peatükki.

EP valimiste korral lisandub maakondade tasemele veel alaliselt välisriigis elavate valijate hulk. Sel puhul eristatakse nimekirja kaupa ja kandidaaditi pabersedeliga antud hääled (jaoskonna hääled) ning e-hääled.

KOV (EP, RK ja KOV valimisliigid)

See andmehulk on valimisliigiti ühtlane, st esitatakse sama andmestik - igale kandidaadile antud hääled jaoskonna kaupa vastavas KOVis. Jaoskonnad on tähistatud vastavalt nummerdusele nt "J1, J2, ...". Nimekirjadele antud hääled summeeritakse tabelis kokku, ning kokku on arvatud ka kogu KOVi kehtivad hääled. KOV ja RK valimiste korral on samas tabelis esitatud veel ka vastava KOVi valijate arv ning osalusprotsent. Kui KOVi tasemel osalusstatistikat on käsitletud juba hääletamisest osavõtu peatükis, siis jaoskonna tasemel osalusprotsent jääb ainult detailse hääletamise andmestiku osaks.

Siiani on valija olnud seotud ühe, elukohajärgse jaoskonnaga. Kui valija annab e-hääle, loetakse see tema nn "kodujaoskonna" e-hääleks. Alates 2013. aastast, mil e-hääled jaoskonniti lisati detailse hääletamistulemuse koosseisu, kuvatakse need e-hääled vastavalt "J1E, J2E, ...". Kõik eespool nõutavad andmed on võimalik leida joonisel 11 kujutatud klassimudelil olevate tabelite kaudu.



Joonis 11. Klassimudel detailse hääletamisandmestiku esitamiseks

Täpsustused nõudele 1.b.2:

Valimisandmed, mida on detailsete hääletamistulemuste kohta võimalik veebilehel ühe valimissündmuse kohta näha ning alla laadida on järgmine:

Eesti Vabariigi kohta EP valimisel

- Nimekiri
 - Kandidaat
 - Registreerimisnumber
 - Jaoskonna hääli kokku
 - E-hääli kokku (alates 2005. a.)
 - Välishääli kokku

- Nimekiri kokku
 - Jaoskonna hääli kokku
 - E-hääli kokku
 - Välisahäali kokku
- Kehtivaid hääli kokku

Ringkondade kaupa RK valimistel

Toodud on häälte arv kokku, ringkonda kuuluvates üksustes eraldi, välisahälte arv ja e-häälte arv:

- Nimekiri
 - Kandidaat
 - Registreerimisnumber
 - Hääli kokku
 - Hääli ringkonda kuuluvates üksustes (KOVides, Tallinnas linnaosades)
 - E-hääli
 - Välisahäali
- Nimekiri kokku
 - Kvooti
 - Hääli kokku
 - Hääli ringkonda kuuluvates üksustes (KOVides, Tallinnas linnaosades)
 - E-hääli
 - Välisahäali

Maakondade kaupa RK ja EP valimistel

- Nimekiri
 - Kandidaat
 - Registreerimisnumber
 - Hääli kokku maakonnas (EP valimissündmuse korral Tallinas)
 - Hääli kokku maakonda kuuluvates KOVides (EP valimissündmuse ja Tallinnas linnaosades)

- Nimekiri kokku
 - Maakonnas
 - KOVides
- Hääli kokku
 - Maakonnas
 - KOVides

KOVis (EP, RK ja KOV valimisliikide korral)

- Nimekiri
 - Kandidaat
 - Registreerimisnumber
 - VA (KOV valimisliigi korral)
 - valitud (KOV valimisliigi korral)
 - kokku
 - kokku ringkonnas (KOV valimisliigi korral)
 - P kokku
 - E kokku (alates 2013. aastast)
 - J1+n P
 - J1+n E (alates 2013. aastast)
- Nimekiri kokku
- Kehtivaid hääli kokku
- Osalejaid kokku
- Valijaid kokku

Puuduvate andmehulkade kohta tuleb esitada põhjendus infoteates. Loetletud andmed saab esitada tabelite kujul ning neile rakendub nõue 1.f. ja 1. g. Üldtabelid moodustuvad loetletud üksuste kaupa.

2.6. Üldised nõuded analüüsist lähtuvalt

Analüüsis on iga peatüki lõpuks esile optimaalne andmehulk, mida on vastava andmestiku puhul otstarbekas kuvada, moodustades täpsustused olemasolevatele lähteülesande nõuetele. Üldine reegel RVT veebilehtedel on, et andmefailis peab sisaldama kogu info, mis on valimiste veebilehel kuvatud ja vastupidi. Erandid, kus see selliselt ei ole, on vastava andmehulga juures välja toodud.

Nagu analüüsist võis järeldada, jaguneb puuduv andmestik kaheks osaks:

- andmed, mida ei ole vastavalt täpsustatud nõuetele leitud (kokku summeeritud)
- andmed, mida ei ole vastavalt täpsustatud nõuetele kogutud

Esimeses hulgas on see andmestik võimalik olemasolev põhjal leida. Teises on osa andmeid, mida polegi võimalik leida ning see tuleb uurijale infotekstis selgitada. Siin võib näitena tuua 2009. aasta EP valimised, kus kinniste nimekirjade tõttu ei ole võimalik esitada kandidaatidele antud häáli. Osa andmeid on aga RVT arhiivist võimalik VARH andmebaasi sisestada. Käeolev töö, mis nõutavad andmehulgad välja toob, annab selleks tööks sisendi. Analüüsist lähtuvad uued nõuded on seega järgmised:

1. Kasutaja peab saama infot, millistes andmehulkades on aastate või valimisliikide lõikes lünki.
2. Lünkade esinemise põhjused tuleb infoteatena esitada vastava valimissündmuse juures.
3. Avaandmed peavad sisaldama puuduva andmehulga atribuute, kuid need jätta tühjaks, et tagada andmete võrreldavus aastate lõikes.
4. Avaandmed peavad sisaldama sama andmehulka, mis on kuvatud valimiste veebilehel.
5. Avaandmete failinimi peab sisaldama valimiste koodi (kujul näiteks RK2015).

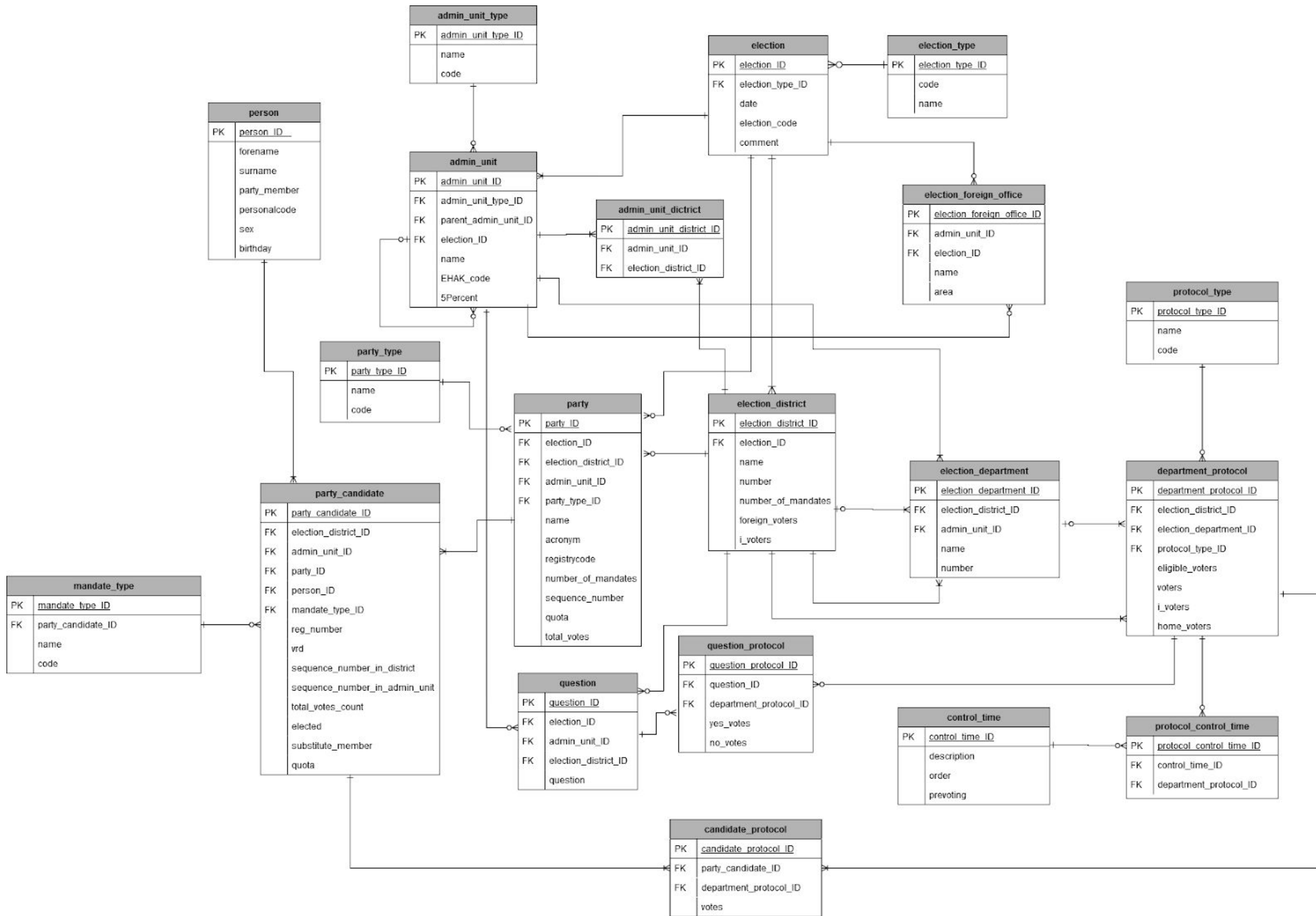
3. Kavand

Kavandi osas esitatakse töö üheks eemärgiks olnud valimisandmete andmemudel, mis lähtub analüüsi peatükis leitud vajadustest andmete hoidmiseks ning leidmiseks. Sissevaade lähteülesandes toodud nõuetele andmete migreerimise osas on kõrvutatud tänaste uute võimalustega andmeid HTML-lehtedelt eraldada. Järgneb komponentmudel rakendustega, mis võimaldavad andmete migreerimise loodavasse andmebaasi. Töö teiseks eemärgiks olnud andmefailide struktuuride erinevused valimiliikide lõikes on toodud vastavas peatükis koos XML-formaadis failistruktuuride näidetega.

3.1. Andmemudel olemi-suhte diagrammina

Andmemudeli eesmärgiks on tulevikus loodava VARH andmebaasi struktuuri loomine. Mudel täidab 2015. aasta lähteülesande nõuded ning analüüsi käigus neile nõuetele lisatud täpsustused. Andmemudeli otseseks kasutajaks on andmebaasi loomise ja andmete migreerimisega tegelevad isikud RVTst ning kaudselt kõik valimisandmetest huvituvad isikud, kes otsivad infot valimiste veebilehtedelt, kuna veebilehtedele kuvatakse edaspidi andmebaasist tulev info.

Andmebaasi loomiseks on koostatud relatsiooniline andmemudel, mis sisaldab kõiki analüüsi klassimudelites loodud tabeleid. Kasutatud on “varesejala” notatsiooni (*crows foot notation*). Mudeli loomisel on lähtutud RIA koostatud “Relatsiooniliste andmemudelite koostamise juhendist” lähtuvalt [27], mis kirjeldab reegleid riigi infosüsteemide relatsiooniliste andmemudelite ja mudelite alusel relatsiooniliste andmebaaside loomisel ning andmemudeli puhul on nii tabelite nimed kui atribuudid inglisekeelsed, et saavutada ühtsus olemasoleva VIS andmebaasiga. Tabelite ja atribuutide nimetustes on seal, kus võimalik, taotletud samade andmehulkade juures samu pealkirju olemasoleva VISi andmemudeliga. Andmemudel on kujutatud joonisel 13.



Joonis 13. Valimisandmete olemi-suhte diagramm

Iga olemi kohta on esitatud kirjeldus lisa 1. Kirjelduses on toodud ka inglisekeelse olemi-suhte diagrammi ning eestikeelsete klassimudelite vastavused.

Andmed on normaliseeritud vastavalt eksisteerimiskriteeriumile, st esitatud selle olemi koosseisus, millega see otseselt seotud on [28]. Normaliseerimise juures on arvesse võetud nii olemasolevat VIS andmemudelit kui analüüsi käigus täpsustatud nõudeid andmetele, st milliseid andmeid on üldse vajadus säilitada.

3.2. Andmete migreerimine HTML-lehtedelt andmebaasi

Uuemate valimissündmuste korral on võimalik migreerida andmed VISi andmebaasist otse VARH andmebaasi, küsides vajalikud andmeväljad samade päringutega, mis täidavad HTML-lehti. Vanemate sündmuste korral aga tuleb leida muu võimalus andmed HTML-lehtedelt andmebaasi liigutada. Nagu märgitud töö esimeses osas, sätestas lähteülesanne funktsionaalsete nõuete osas ka nõuded andmete migreerimisele. See nägi ette veebi valimisandmetest kokkulepitud struktuuriga CSV faili koostamise ning selle importimise VARH andmebaasi. Uuemate valimissündmuste korral imporditaks VIS andmebaasist loodud XML failid samuti VARHi andmebaasi. Selleks saab praegusel ajal kasutada mitmesuguseid andmete “kraapimise” tööriistu, mis vastavate seadistuste kohaselt koguvad HTML-i info ning sisestavad selle andmebaasi.

Lähteülesanne esitab migreerimise kohta järgmised nõuded:

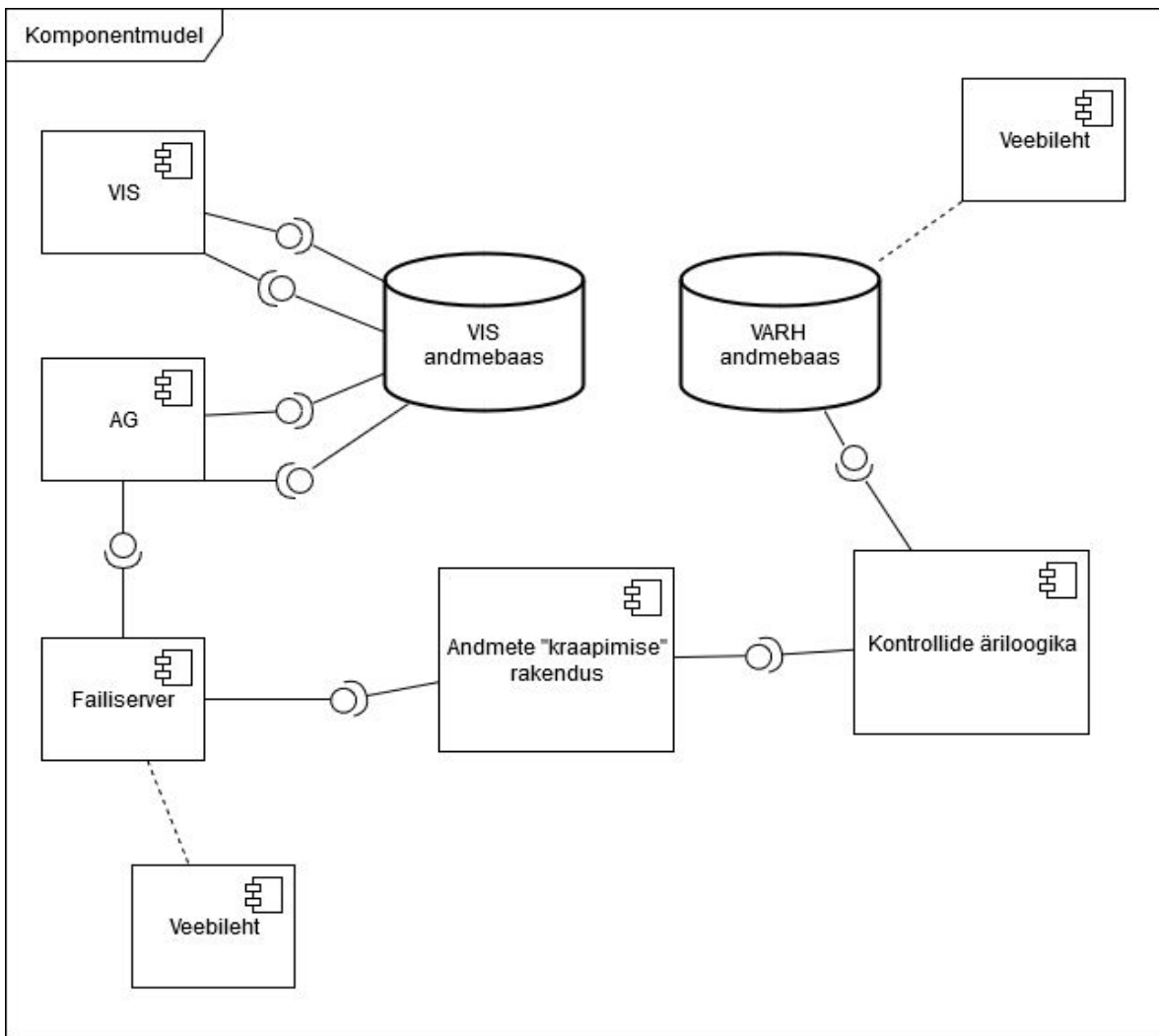
1. Süsteemiadministraator peab saama lisada ajaloolised valimisandmed otse andmebaasi.
2. Süsteemiadministraator peab saama importida kokkulepitud struktuuriga CSV faile, kust loetakse andmed andmebaasi.
3. Veebiadministraator peab saama importida VIS põhimoodulist eksporditud kokkulepitud formaadis valimisandmete XML faile valimisandmete arhiivi testkeskkonda ja töökeskkonda.
4. Süsteemiadministraator peab saama kergelt tagasi pöörata andmebaasi eelmist versiooni.

5. Veebiadministraator peab saama statistikabelitele ja veebilehtedele lisada selgitavaid infotekste eesti ja inglise keeles [7].

Käesoleva töö analüüsisist ning praegustest võimalusest lähtuvalt saab ajaamahuka CSV failide koostamise kõrval nõuetele lisada veel järgmist:

6. Süsteemiadministraator peab saama valimiste veebilehtedelt sobiva andmete “kraapimise” tööriistaga eraldada vajalikke andmeid.
7. Andmete “kraapimise” tööriist peab võimaldama andmete esmast süstematiseerimist.

Komponentmudel joonise 12 esitab andmete migreerimise vanemate valimissündmuste korral, mil andmestik tuleb koguda HTML-lehtedelt.



Joonis 12. Komponentmudel andmete migreerimiseks veebilehelt VARH andmebaasi

Andmete kogumise järgselt tuleb läbi viia nende kvalitatiivne analüüs, et leida võimalikud migreerimisel tekkinud vead. Olenevalt tööriistast saab seda teha juba andmete “kraapimise” juures või hiljem vastavate päringutega andmebaasis. Andmebaas peab võimaldama kasutajal aru saada, kas mõni andmehulk on mõne valimissündmuse juurest üldse puudu (näiteks e-hääled enne 2005. aastat) või ootavad need alles eraldi käsitsi sisestamist. Esimesel juhul ei peaks puuduvat andmehulka üldse veebilehtedele välja kuvama, teiste puhul tuleb andmeväli jätta tühjaks ning selgitada seda infotekstiga.

3.3. Andmestruktuurid avaandmetele

Siiani kättesaadavad andmed arvestavad suuresti valimispäeva õhtu infovajadust ning on selliselt ka struktureeritud. VARH, mis on loodud tabelis 1 loetletud kasutajagruppidele, on suunitlusega teadustööks ja analüüsiks. Vastavalt peatükis 1.5 toodud RIA loetletud nõuetele on koostatud järgnevad andmestruktuurid. Esitatavad andmed on tööandmetest **eraldatud** - tegemist on tööandmetest lahutatud, avaliku andmestikuga. Avaldatuna RVT veebilehel on need lihtsasti **kättesaadavad** ja **juurdepääsetavad** kõigile. Esitamise **periodilisus** on vastavuses valimiste toimumise tsükliga. **Järjepidevuse** ja **tervikluse** küsimust on käsitletud analüüsis ning andmete kasutajale edastatakse info võimalike lünkade kohta. **Terviklikkuse** huvides on seniseid struktuure muudetud. **Esitusviisiks** on laialt kasutatav XML-formaat. **Mõistetavuse** huvides on muudetud atribuutide nimetused selliseks, et need oleksid maksimaalselt “iseseletavad” ning kuna sama andmestik esitatakse VARH veebilehel, on seal olemas võti andmete mõistmiseks. Andmed on **litsenseeritud** Creative Commons litsensiga CC BY 4.0 [39]. Andmed avaldatakse **proaktiivselt** - seni on uurijate infovajadus rahuldatud teabepäringutega. Avaldatav andmestik isikute kohta on reguleeritud valimisseadusega, ega sisalda **konfidentsiaalseid** andmeid. Andmete sisu ja **õigsuse** tagab RVT.

Järgnevalt esitatakse struktuurid viie analüüsis käsitletud andmehulga kohta, sõltuvalt valimisliigist.

3.3.1. Jaoskonnad

Jaoskonnad asuvad haldusüksuste “puus” sõltuvalt valimisliigist erineval tasemel: RK, EP ja RH valimissündmuste korral KOVis ning KOV valimiste korral ringkonnas. Haldusüksuste “puu” moodustub haldusüksuste tüüpide kaupa. Välisesindused kuuluvad failis haldusüksuse mõttes Eesti Vabariigi alla. EP ja RH korral on ringkondi ainult üks. Näitena on esitatud haldusüksuste struktuur ja jaoskonnad RK valimisündmuse korral (lisa 2), teiste valimisliikide puhul on faili moodustamise loogika sama vastavalt haldusüksuste ja ringkondade kujunemisele.

3.3.2. Kandidaadid

Kandidaatide infot esitatakse kahe eri huljana - kandidaadid ühel valimissündmusel ning kõik kandidaadid kõigi valimissündmuste korral. Esimesel puhul jagatakse kandidaadid Eesti Vabariigis ringkondadesse, nagu need vastavalt valimisliigile kujunevad. Iga kandidaadi kohta on märgitud tema järjekorranumber üksuses (RK valimiste korral tähendab see järjekorda üleriigilises nimekirjas, KOV valimiste korral ainult Tallinna ülelinnalises nimekirjas) ning ringkonnas. EP valimissündmuse korral, mil üksus (Eesti Vabariik) ja ringkond kattuvad, puudub vajadus ringkonna ja üksuse kaupa järjekorda eristada. Failistruktuuri ühtse mõistetavuse huvides võib ka see andmestik sarnaneda teistega ning ringkonna ja üksuse järjekorranumber on lihtsalt alati sama. Näitena on toodud kandidaadid KOV valimiste puhul (lisa 3).

Teine kandidaatide andmehulk - kandidaadid kõigi valimissündmuste korral - sisaldab 1992-... aastate kandidaate, kus iga kandidaadi info sisaldab osaletud valimissündmuse, registreerimisnumbrit, ringkonda, valla/linna nime (juhul kui tegemist oli KOV valimistega) ja nimekirja, milles kandideeriti, häälte arvu ning valituks osutumist. Senises kandidaatide otsingus lehel: <http://vvk.ee/arhiiv/kandidaadi-otsing/> ei kuvata RK ega EP valimiste korral kandidaadi ringkonna nime (EP korral oleks see siis alati Eesti Vabariik), kuid ühtsuse huvides võiks selle samuti sinna lisada. Otsingus ei ole võimalik näidata 2009. aasta Euroopa Parlamendi valimistel kinniste nimekirjadega süsteemis erakonna nimekirjas kandideerinud kandidaatide kohta hääli - sel puhul tuleb

häälte arvu märgendite vahele kirjutada “andmed puuduvad” ning selgitus lisada andmete allalaadimise lehele. Erinevate valimiste vahel nime muutnud kandidaadid viiakse kokku isikukoodi abil, kuid isikukoodi ennast kusagil ei avaldata.

Toodud näites (lisa 4) on näidisenä kandidaat, kes on osalenud nii EP, RK kui KOV valimistel erakonna või valimisliidu nimekirjas ja ka üksikkandidaadina.

3.3.3. Hääletamisest osavõtt

Analüüsis leitakse, et hääletamisest osavõtt jaguneb kahte ossa - eelhääletamise statistika maakonniti ja nn lõpliku seisuga, detailne statistika. Hääletamisest osavõtu puhul puuduvad enne 1999. aastat eelhääletamise andmed veebist. Eelhääletamise osa sisaldava osavõtuandmestiku saab seetõttu väljastada ainult 1999 -2019. aasta andmete kohta. Vastav failistruktuur on RVTs juba olemas ning kasutusel, mistõttu seda siin ei esitata. Uus on detailsete osavõtuandmete fail (lisa 5), kus osaliselt taaskasutatakse juba olemasoleva detailse hääletamistulemuse faili elemente, kuid selliselt, et andmed oleksid kergesti mõistetavad ka välisriigi uurijale.

3.3.4. Hääletamis- ja valimistulemus

Hääletus- ja valimistulemuste fail sisaldab kõiki eespool analüüsis vastavas peatükis loetletud andmehulki - valimistulemuste erakonniti, valituks osutunud liikmeid ning asendusliikmeid, samuti seda, millise mandaadi liigi on valituks osutunu saanud. Sellist andmestikku sisaldav fail on RVTs avaandmetena olemas, kuid selle fookus on olnud valimispäeva õhtul info edastamine, mistõttu sisaldab juba toimunud valimise uurijale ebavajalikke elemente. Nagu eespool märgitud, peab avaandmete fail sisaldama infot, mis veebis andmete vaatajale on lehekülje liigenduse järgi arusaadav ning ka vastupidi - peatükis 2.4. nõutavatest väljadest ei ole kandidaadi juures nimekirja kuvamine oluline, kuna XML faili vastav pesastus muudab selle arusaadavaks. Failidesse on lisatud elemente, mis valimiste lehtedel andmete pideva muutumise tõttu olid lahendatud dünaamiliselt veebilehe poolel, kuid nüüd vaadeldavad staatilisena, näiteks 5% künnis.

Näidisenä on toodud RK vastav fail (lisa 6), kus üks kandidaat on saanud isikumandaadi ning teine jäänud asendusliikmeks. Iga kandidaadi kohta kuvatakse andmete sorteerimist võimalda komplekt (nt võrdlusarv, kvoot), mis sel juhul, kui kandidaadile neid ei arvatud, jäävad tühjaks. Vastav selgitus antakse uurijale veebilehel.

Uue lahendusena on näide RH valimisliigi vastavast failist (lisa 7), milles on toodud Eesti vabariigi tasemel atribuudid, mis korduvad iga maakonna ning KOVi valimissündmuse korral. Ainsana ei kordu välishäälte osa, kuna see on lisatud ainult üleriigilisel tasemel.

3.3.5. Detailne hääletamistulemus

Detailset hääletamistulemust on siiani esitatud RK ja EP valimissündmuste korral kahe osana - üks fail maakondade infoga ning teine grupp, kus iga fail sisaldab ühe KOVi detailseid hääletamistulemusi - siin tekib nii palju faile, kui on Eestis KOVe. Nagu vastava hulga analüüsist nähtus, sisaldab see andmehulk infot, mis sisuliselt on liigitatavad hääletamisest osavõtu andmete alla ning nii on see ka osalusandmete faili kaasatud. RH korral ei ole seni hääletamisinfot jaoskondade kaupa kuvatud ning kuna 2018. aastate seadusemuudatustega ei ole valijad enam kindla jaoskonnaga seotud, vaid ainult valimisringkonnaga, siis ei ole jaoskondade tasemel lahendust siin ka välja toodud [13b]-[16b].

Siin esitatud näide (lisa 8) on seni kasutatud failidest hääletamise osavõtu eristanud ning sisaldavad ainult hääletamistulemuse osa. Varem on XMLe kasutatud HTMLi loomiseks, kuid edaspidi, kui XMLi otstarbeks on andmestiku arusaadaval moel kuvamine nii Eesti kui välisuurijale, on seni kasutusel olnud failistruktuure muudetud. RK ning EP korral tekivad ringkondade erisuse tõttu erinevad struktuurid - RK valimisliigi korral saab kandidaatide andmeid näidata ringkonniti ja maakonniti ning igas ringkonnas on eri kandidaadid, EP korral samuti ringkonnas ja maakonnas, kuid samu kandidaate. Andmestik on loodud vastavalt analüüsis välja toodud nõuetele. Kuna olulisi muudatusi sisaldab siin ainult suuremate üksuste - maakonde ja ringkondade fail

- siis on siin töös näitena toodud ainult seda liiki fail. KOV taseme andmefaili, mis praegu kasutusel on, muudatusi ei planeerita ning seda siinkohal ei esitata.

4. Kokkuvõte

Valimiste eelselt ja järgselt avalikustatud ametlik andmestik on olnud vaadeldava perioodi - 1992-2019 - jooksul pidevas muutuses. Kui osa andmestikku on sel perioodil olnud alati esitatud - nt osavõtuprotsent, valimistulemused - siis andmete detailsus ning loogilistesse hulkadesse jagamine on reageerinud avalikkuse ootustele ning infovajadustele. Valimisandmed asuvad praegu omaaegsetel veebilehtedel. Käesolevaks ajaks on tekkinud vajadus andmete masinloetavuseks, mis võimaldaks vastavalt uurija hüpoteesile esitada uusi ja huvitavaid analüüse. Selle töö eesmärgiks on alustada protsessiga, mille lõpptulemuseks on valimisandmete arhiiv, mille kõiki andmeid saab alla laadida ning võrrelda.

Andmete detailanalüüs, mis selles töös on läbi viidud, on kätte juhatanud lüngad andmestikus, mis mõnel puhul on vastava algoritmiga lihtsasti lahendatavad, mõnel juhul vajavad aga täiendavat arhiivitööd seni avaldamata allikatest. Suures enamuses on valimisandmed aga olemas ning see töö jagab need sisuliselt loogilistesse hulkadesse, mis on kasutaja jaoks vastavalt XML failidesse jagatud. Failstruktuuride ühtlustamine tagab andmete võrreldavuse sama valimisliigi erinevate valimissündmuste ning üldjoontes ka eri valimisliikide vahel.

Detailanalüüsi tulemusena valimis töös andmemudel, mis on sisendiks tulevase VARH andmebaasi loomiseks. Kaasaegsed võimalused annavad võimaluse andmeid veebilehtedelt “kraapida”, mis võimaldab kogu andmete migreerimise veebilehtedelt andmebaasi teha kiiremini.

Töös on loodud uued andmestruktuurid andmetele, mida seni pole avaandmete koosseisu arvatud ning muudetud olemasolevaid avaandmete failistruktuure selliseks, et sisaldaksid sisu poolest loogiliselt seotud andmehulki ning need oleksid kergemini mõistetavad ka välisriigi uurijatele.

Kokkuvõtteks lõi täitis käesolev töö ülesandepüstituses võetud eesmärgid - loodi andmestruktuurid seniste andmete mahutamiseks, kuvamiseks ja avaldamiseks.

Kasutatud kirjandus

- [1] Riigi valimisteenistus, “Toimunud valimised”, [Võrgumaterjal].
<https://www.valimised.ee/et/toimunud-valimiste-arhiiv>. [Kasutatud 28.01.2020]
- [2] Riigi valimisteenistus, “Rahvahääletus Eesti Vabariigi põhiseaduse eelnõu, põhiseaduse rakendamise seaduse eelnõu ja nendega kaasneva lisaküsimuse osas 28. juunil 1992. aastal”, [Võrgumaterjal].
<https://www.valimised.ee/et/valimiste-arhiiv/rahvah%C3%A4%C3%A4letus-est-i-vaba-riigi-p%C3%B5hiseaduse-eeln%C3%B5u-p%C3%B5hiseaduse-rakendamise-seaduse>.
[Kasutatud 01.03.2020]
- [3] Riigi valimisteenistus, “Valimiste avaandmed”, [Võrgumaterjal].
<https://www.valimised.ee/et/valimiste-arhiiv/valimiste-avaandmed>. [Kasutatud 29.01.2020]
- [4] Riigi valimisteenistus, “Statistika ja analüüs”, [Võrgumaterjal].
<https://www.valimised.ee/et/valimiste-arhiiv/statistika-ja-anal%C3%BC%C3%BCs>.
[Kasutatud 29.01.2020]
- [5] Riigi valimisteenistus, “E-hääletamise statistika”, [Võrgumaterjal].
<https://www.valimised.ee/et/valimiste-arhiiv/elektroonilise-h%C3%A4%C3%A4letamis-e-statistika>. [Kasutatud 02.02.2020]
- [6] Riigi valimisteenistus, “Valimised Eestis 1992-2015”, [Võrgumaterjal].
https://www.valimised.ee/sites/default/files/uploads/misc/Valimised_Eestis_1992-2015.pdf. [Kasutatud 01.02.2020]
- [7] Pung, E. Valimiste arhiivi lähteülesanne. Ärianalüüs. Hansson, Leego ja Partner. AK. 2015. RVT arhiivis
- [8] Intervjuud RVT infosüsteemide valdkonna peaspetsialisti Leino Mandrega 02.03. ja 05.03.2020
- [9] Riigi valimisteenistus, “Toimunud Vabariigi Presidendi valimised”, [Võrgumaterjal].
<https://www.valimised.ee/et/valimiste-arhiiv/toimunud-vabariigi-presidendi-valimised>.
[Kasutatud 15.03.2020]

- [10] Riigi valimisteenistus, "Toimunud Riigikogu juhatuse valimised", [Võrgumaterjal].
<https://www.valimised.ee/et/valimiste-arhiiv/toimunud-riigikogu-juhatuse-valimised>.
[Kasutatud 03.02.2020]
- [11] Riigi valimisteenistus, "Hääletamis- ja valimistulemuste kindlakstegemine",
[Võrgumaterjal].
<https://www.valimised.ee/et/valimiste-meelespea/h%C3%A4%C3%A4letamis-ja-valimistulemuste-kindlakstegemine>. [Kasutatud 06.02.2020]
- [12] Riigikogu, "Riigi valimisteenistuse põhimäärus", [Võrguteavik].
<https://www.riigikogu.ee/wpcms/wp-content/uploads/2014/11/Riigi-valimisteenistuse-p%C3%B5him%C3%A4%C3%A4rus-1.pdf>. [Kasutatud 05.03.2020]
- [13] a) Riigikogu valimise seadus. *Riigi Teataja*, nr I, 03.01.2020, 12, [Võrguteavik].
<https://www.riigiteataja.ee/akt/103012020012>. [Kasutatud 11.03.2020]
- b) Riigikogu valimise seadus. *Riigi Teataja*, nr I, 09.07.2018, 5, [Võrguteavik].
<https://www.riigiteataja.ee/akt/109072018005>. [Kasutatud 12.03.2020]
- [14] a) Kohaliku omavalitsuse volikogu valimise seadus. *Riigi Teataja*, nr I, 03.01.2020, 10, [Võrguteavik]. <https://www.riigiteataja.ee/akt/103012020010>. [Kasutatud 11.03.2020]
- b) Kohaliku omavalitsuse volikogu valimise seadus. *Riigi Teataja*, nr I, 09.07.2018, 3, [Võrguteavik]. <https://www.riigiteataja.ee/akt/109072018003>. [Kasutatud 11.03.2020]
- [15] a) Euroopa Parlamendi valimise seadus. *Riigi Teataja*, nr I, 03.01.2020, 8, [Võrguteavik]. <https://www.riigiteataja.ee/akt/103012020008>. [Kasutatud 11.03.2020]
- b) Euroopa Parlamendi valimise seadus. *Riigi Teataja*, nr I, 22.12.2018, 21, [Võrguteavik]. <https://www.riigiteataja.ee/akt/122122018021>. [Kasutatud 15.03.2020]
- [16] a) Rahvahääletuse seadus. *Riigi Teataja*, nr I, 03.01.2020, 14, [Võrguteavik].
<https://www.riigiteataja.ee/akt/103012020014>. [Kasutatud 11.03.2020]
- b) Rahvahääletuse seadus. *Riigi Teataja*, nr I, 09.07.2018, 4, [Võrguteavik].
<https://www.riigiteataja.ee/akt/109072018004>. [Kasutatud 15.03.2020]
- [17] Riigi valimisteenistus, "Riigi valimisteenistus", [Võrguteavik].
<https://www.valimised.ee/et/valimiste-korraldajad/riigi-valimisteenistus>. [Kasutatud 28.02.2020]
- [18] Riigi valimisteenistus, "Vabariigi Valimiskomisjoni koosseis, pädevus ja ülesanded", [Võrguteavik].

- [19] Valimiste infosüsteemi põhimäärus. Riigi Teataja, I, 02.07.2013, 68, [Võrguteavik]. <https://www.riigiteataja.ee/akt/102072013068>. [Kasutatud 02.03.2020]
- [20] Aruanne 1999. a Riigikogu valimiste infotehnoloogilise tagamise kohta. IC Süsteemide AS ja Microlinki Süsteemide AS. 17.03.99, lk. 2. RVT arhiivis
- [21] Riigi valimisteenistus, “Hääletamisviisid”, [Võrguteavik]. <https://www.valimised.ee/et/euroopa-parlament-2019/h%C3%A4%C3%A4letamisviisid>. [Kasutatud 14.03.2020]
- [22] Riigi valimisteenistus, “Toimunud valimiste kandidaatide otsing”, [Võrguteavik]. <http://vvk.ee/arhiiv/kandidaadi-otsing/>. [Kasutatud 28.02.2020]
- [23] Riigi Infosüsteemi Amet. “Avaandmete loomise ja avaldamise juhend”. Versioon 1.1. Tallinn 2015, uuendatud 2016, [Võrguteavik]. https://www.ria.ee/sites/default/files/content-editors/publikatsioonid/avaandmete_loomise_juhend.pdf. [kasutatud 23.03.2020]
- [24] Andmekaitseinspeksioon, “Avaandmetest”, [Võrguteavik]. <https://www.aki.ee/et/teabe-avalikkus/avaandmetest>. [Kasutatud 23.03.2020]
- [25] Liu, Chi-Lun, “Knowledge Representation and Inconsistency Reasoning of Class Diagram Maintenance in Big Data,” *World Academy of Science, Engineering and Technology International Journal of Computer and Information Engineering*. Vol:12, No:7, 2018, lk 541-545, [Võrguteavik]. <https://zenodo.org/record/1317424>. [Kasutatud 15.04.2020]
- [26] Haav, Hele-Mai. Ontoloogiatega loomise meetoodika. Versioon Ver 4. 21.10.2011. Küberneetika Instituut, [Võrguteavik]. https://moodle.ria.ee/pluginfile.php/375/mod_resource/content/2/ontoloogiatega_loomise_meetoodika_v4.PDF. [Kasutatud 15.04.2020]
- [27] Riigi Infosüsteemi Amet. “Relatsiooniliste andmemudelite koostamise juhend”. Ver 1.0. Tallinn 2015, [Võrguteavik]. https://www.ria.ee/sites/default/files/content-editors/publikatsioonid/relatsiooniliste_andmemudelite_koostamise_juhend_ver_1.0.pdf. [Kasutatud 10.03.2020]
- [28] Inmon, Bill. “Database Architecture and Design”, *Technics Publications*, 2018, [Võrguteavik]. https://learning.oreilly.com/videos/database-architecture-and/9781634623308/9781634623308-9781634623308_2. [Kasutatud 12.03.2020]
- [29] Avaliku teabe seadus. *Riigi Teataja*, I, 15.03.2019, 11, [Võrguteavik]. <https://www.riigiteataja.ee/akt/122032011010?leiaKehtiv>. [Kasutatud 12.04.2020]

- [30]] Riigi valimisteenistus. “Riigi valimisteenistuse andmekaitsetingimused”, [Võrguteavik]. <https://www.valimised.ee/et/andmekaitsetingimused>. [Kasutatud 27.02.2020]
- [31] Isikuandmete kaitse seadus. *Riigi Teataja*, I, 04.01.2019, 14, [Võrguteavik]. <https://www.riigiteataja.ee/akt/106012016010?leiaKehtiv>. [Kasutatud 27.02.2020]
- [32] Eesti Statistikaamet. “Klassifikaatorite nimistu. Eesti haldus- ja asustusjaotuse klassifikaator. Varasemad versioonid”, [Võrguteavik]. <http://metaweb.stat.ee/>. [Kasutatud 25.02.2020]
- [33] Kandidaatide registreerimiseks esitamise ja registreerimise kord Riigikogu valimistel. *Riigi Teataja lisa*, 2002, 127, 1851, [Võrguteavik]. <https://www.riigiteataja.ee/akt/249486>. [Kasutatud 21.03.2020]
- [34] Kandidaatide registreerimise avalduse vormi, kandidaatide ringkonnanimikirja vormi, kandidaatide ülevallalise või -linnalise nimekirja vormi, kandideerimisavalduse vormi ja kandidaadi ankeedi vormi kehtestamine kohaliku omavalitsuse volikogu valimisteks. *Riigi Teataja*, I, 16.01.2013, 1, [Võrguteavik]. <https://www.riigiteataja.ee/akt/116012013001>. [Kasutatud 23.03.2020]
- [35] Andmekaitse inspeksioon, “Isikuandmete liigitus”, [Võrguteavik]. <https://www.aki.ee/et/eraelu-kaitse/isikuandmed-ja-tootlemine/isikuandmete-liigitus>. [Kasutatud 17.03.2020]
- [36] E-äriregister, “Erakonnad”, [Võrguteavik]. <https://ariregister.rik.ee/erakonnad?lang=est>. [Kasutatud 04.04.2020]
- [37] Riigi valimisteenistus, “Valimistulemuste kindlakstegemine Riigikogu valimistel”, [Võrguteavik]. <https://www.valimised.ee/et/valimiste-meelespea/valimistulemuste-kindlakstegemine-riigikogu-valimistel>. [Kasutatud 04.04.2020]
- [38] Riigi valimisteenistus, “Valimistulemuste kindlakstegemine kohaliku omavalitsuse volikogu valimistel”, [Võrguteavik]. <https://www.valimised.ee/et/valimiste-meelespea/valimistulemuste-kindlakstegemine-kohaliku-omavalitsuse-volikogu-valimistel>. [Kasutatud 04.04.2020]
- [39] Creative Commons, “Autorile viitamine 4.0 Rahvusvaheline (CC BY 4.0)”, [Võrguteavik]. <https://creativecommons.org/licenses/by/4.0/deed.et>. [Kasutatud 29.03.2020]

Lisa 1 – Olemi-suhte diagrammi olemite kirjeldused

1. Tabel election (Valimissündmus)*

Tabelis hoitakse kõiki toimunud valimissündmusi vaadeldavast perioodist. Iga valimissündmuse kohta moodustub uus kirje.

*Sulgudes märgitud tabeli nimi klassimudelites

Atribuut	Andmetüüp	Klassimudel	Selgitus
election_ID	serial not null	ValimissündmusID	automaatselt genereeritud, unikaalne ID
election_type_ID	int4 not null	ValimisLiikID	Valimisliigi tüüp vastavast tabelist
code	int4 not null	kood	Valimissündmuse kood, nt RK2015
date	date	toimumisaeg	Valimissündmuse valimispäeva kuupäev, näiteks 01.03.2015
comment	varchar(400)	-	Võib sisaldada olulist infot valimissündmuse, erisuste vms kohta

2. Tabel election_type (ValimisLiik)

Tabelis hoitakse valimisliike üleriigiliste valimissündmuste hulgast.

Atribuut	Andmetüüp	Klassimudel	Selgitus
election_type_ID	serial not null	ValimisLiikID	automaatselt genereeritud, unikaalne ID
code	varchar(20)	kood	Valimisliigi kood 1- Riigikogu valimised 2- Euroopa

			parlamendi valimised 3- KOV valimised 4- Rahvahääletus
name	varchar(80) not null	nimi	Valimisliigi nimi

3. Tabel election_district (Ringkond)

Tabel valimisringkondadega, mis iga valimissündmuse puhul luuakse.

Atribuut	Andmetüüp	Klassimudel	Selgitus
election_district_ID	serial not null	RingkondID	automaatselt genereeritud, unikaalne ID
election_ID	int4 not null	ValimissündmusID	Viide valimissündmusele
name	varchar(255)	nimi	Ringkonna nimi, näiteks: "Harju- ja Raplamaa"
number	int4	number	Ringkonna number
number_of_mandates	int4 not null	mandaate	Ringkonnas jagatavate mandaatide arv
foreign_voters	int4	välisääletajaid	Välisääletajate arv ringkonnas
i-voters	int4	e-hääletajaid	E-hääletajate arv ringkonnas

4. Tabel admin_unit (HaldusÜksus)

Tabel haldusüksutega, mis vastavad valimissündmuse korraldamise aja haldusüksuste struktuurile

Atribuut	Andmetüüp	Klassimudel	Selgitus
admin_unit_ID	serial not null	HaldusÜksusID	automaatselt genereeritud, unikaalne ID
admin_unit_type_ID	int4	HaldusüksusTüüpID	Viide haldusüksuse

D		D	tüübile
parent_admin_unit_ID	int4	ÜlemhaldusüksusID	Rekursiivne seos ühe haldusüksuse alluvussuhte väljendamiseks teise haldusüksusega
election_ID	int4	ValimissündmusID	Viide valimissündmusele
name	varchar(255)	nimi	Haldusüksuse nimi, näiteks "Harjumaa"
EHAK_code	varchar(6) not null	EHAKkood	Haldusüksuse kood Statistikaameti klassifikaatorist, mis on aluseks vastava valimissündmuse haldusstruktuurile, näiteks 0037
5PercentThreshold	numeric(10,5)	5ProtsendiKünnis	5% kehtivatest häältest

5. Tabel admin_unit_type (HaldusüksusTüüp)

Haldusüksuse jaotus riigi, maakonna, KOV, linnaosa taseme järgi

Atribuut	Andmetüüp	Klassimudel	Selgitus
admin_unit_type_ID	serial not null	HaldusüksusTüüpID	automaatselt genereeritud, unikaalne ID
name	varchar(100)	nimi	Tüübi kood
code	int4	kood	Tüübi nimi: 1- Eesti Vabariik 2- maakond 3-KOV 4-linnaosa

6. Tabel admin_unit_district (HaldusüksusRingkonnas)

Tabel ringkondade ja haldusüksuste vahelise alluvussuhte väljendamiseks vastavalt valimisliigile

Atribuut	Andmetüüp	Klassimudel	Selgitus
admin_unit_district_ID	serial not null	HaldusüksusRingkonnasID	automaatselt genereeritud, unikaalne ID
admin_unit_ID	int4	HaldusüksusID	Viide haldusüksusele
district_ID	int4	RingkondID	Viide ringkonnale

7. Tabel party (Nimekiri)

Valimisel osalev koosseis - erakond, valimisliit, mis esitab oma kandidaatide nimekirja

Atribuut	Andmetüüp	Klassimudel	Selgitus
party_ID	serial not null	nimekiriID	automaatselt genereeritud, unikaalne ID
election_ID	int4	valimissündmusID	Viide valimissündmusele
election_district_ID	int4	ringkondID	Viide ringkonnale, kus nimekiri kandideerib
admin_unit_ID	int4	haldusüksusID	Viide haldusüksusele, kus nimekiri kandideerib
party_type_ID	int4	nimekiriTüüpID	Viide nimekirja esitanud koosseisu tüübile
name	varchar(250)	nimi	Erakonna, valimisliidu nimi, üksikkandidaatide puhul "Üksikkandidaadid"
acronym	varchar(40)	lühend	Lühend koosseisu nimest

registrycode	int4 not null	registrikood	Registrikood Äriregistris
number_of_mandates	int4	mandaatideArv	Valimistel saadud mandaatide arv
sequence_number	int4	järjekorranumber	Liisuheitmise number
quota	numeric(10,5)	kvoot	Nimekirja kvoot ringkonnas
total_votes	int4	hääliKokku	Nimekirjale antud häälte arv kokku

8. Tabel party_type (NimekiriTüüp)

Valimistel osaleva koosseisu tüüp

Atribuut	Andmetüüp	Klassimudel	Selgitus
party_type_ID	serial not null	nimekiriTüüpID	automaatselt genereeritud, unikaalne ID
name	varchar(40)	nimi	Valimistel osaleva koosseisu tüübi nimi 1-erakond 2-valimisliit 3-üksikkandidaadid
code	int4	kood	Koosseisu tüübi number

9. Tabel person (Isik)

Tabelis hoitakse valimistel kandidaatidena osalevaid isikuid

Atribuut	Andmetüüp	Klassimudel	Selgitus
person_ID	serial not null	isikID	automaatselt genereeritud, unikaalne ID
forename	varchar(40) not null	eesnimi	Eesnimi
surname	varchar(40) not	perekonnanimi	Perekonnanimi

	null		
party_member	varchar(200)	erakondlikKuuluvus	Isiku erakondlik kuulumus
personalcode	varchar(11)	Isikukood	Isikukood
birthday	date	Sünniaeg	Isiku sünniaeg, näiteks 01.05.1980
sex	char(1)	Sugu	Isiku sugu (M/N)

10. Tabel party_candidate (KandidaatNimekirjas)

Tabelis hoitakse isikute ja nimekirjade seoseid.

Atribuut	Andmetüüp	Klassimudel	Selgitus
party_candidate_ID	serial not null	kandidaatNimekirjasID	automaatselt genereeritud, unikaalne ID
election_district_ID	int4	ringkondID	Viide valimisringkonnale
admin_unit_ID	int4	haldusüksusID	Viide haldusüksusele
party_ID	int4	nimekiriID	Viide nimekirjale
person_ID	int4	isikID	Viide isikule
mandate_type_ID	int4	mandaadiTüüp	Viide mandaadi tüübile, kui kandidaat osutus valituks
reg_number	int4	registreerimisnumber	Kandidaadi registreerimisnumber
vrđ	numeric(10,4)	Võrdlusarv	Kandidaadile arvutatud võrdlusarv
sequence_number_in_district	int4	järjekorranumberRingkonnas	Kandidaadi järjekorranumber ringkonnas
sequence_number_i	int4	järjekorranumberÜ	Kandidaadi

n_admin_unit		ksuses	järjekorranumber RK valimistel üleriigilises nimekirjas, KOV valimiste ja Talinna puhul ülelinnalises nimekirjas
total_votes_count	int4	hääliKokku	Saadud häälte arv kokku
elected	boolean	valitud	Valituks osutumine (true/false)
substitute_member	boolean	asendusliige	Asendusliikmeks osutumine (true/false)
quota	numeric(10,5)	kvoot	Kandidaadi kvoot

11. Tabel mandate_type (MandaadiTüüp)

Mandaatide tüübid, mida valimisliigiti on võimalik kandidaadil saada

Atribuut	Andmetüüp	Klassimudel	Selgitus
mandate_type_ID	serial not null	mandaadiTüüpID	automaatselt genereeritud, unikaalne ID
party_candidate_ID	int4	kandidaatNimekirjasID	Viide kandidaadile
name	varchar(20)	nimetus	Mandaadi tüübi nimetus: 1-isikumandaat 2-ringkonnamandaat 3-kompensatsiooni mandaat 4-nimekirjamandaat
code	int4	kood	Mandaadi tüübi kood

12. Tabel election_department (Jaoskond)

Valimisteks moodustatud jaoskonnad.

Atribuut	Andmetüüp	Klassimudel	Selgitus
election_department_ID	serial not null	jaoskondID	automaatselt genereeritud, unikaalne ID
admin_unit_ID	int4	haldusüksusID	Viide KOVile või linnaosale, milles jaoskond asub
election_district_ID	int4	ringkondID	Viide ringkonnale, kus jaoskond asub
name	varchar(100)	nimi	Jaoskonna nimi, näiteks "Abja valla valimisjaoskond"
number	int4 not null	number	Jaoskonna number

13. Tabel election_foreign_office (Välisesindus)

Vastavatel valimissündmustel moodustatud välisesindused, kus korraldatakse hääletamist.

Atribuut	Andmetüüp	Klassimudel	Selgitus
election_foreign_office_ID	serial not null	välisesindusID	automaatselt genereeritud, unikaalne ID
admin_unit_ID	int4	haldusüksusID	Haldusüksus, millele välisesindus allub
election_ID	int4	valimissündmusID	Viide valimissündmusele
name	varchar(250)	nimi	Välisesinduse nimi, näiteks "EV Suursaatkond Athens"
area	varchar(1000)	teeninduspiirkond	Loetelu piirkondadest, mida välisesindus teenindab, näiteks "Küpros,

			Albaania”
--	--	--	-----------

14. Tabel department_protocol (Protokoll)

Hääletamistulemusi sisaldav protokoll.

Atribuut	Andmetüüp	Klassimudel	Selgitus
department_protocol_ID	serial not null	protokollID	automaatselt genereeritud, unikaalne ID
election_district_ID	int4	ringkondID	Viide ringkonnale, kus protokoll loodi
election_department_ID	int4	jaoskondID	Viide jaoskonnale, kus protokoll loodi
protocol_type_ID	int4	protokolliLiikID	Viide protokollitüübile
eligible_voters	int4	valijaid	Valijate arv üksuses, kus protokoll koostati
voters	int4	hääletajaid	Hääletajate arv üksuses, kus protokoll koostati
i_voters	int4	e-hääletajaid	E-hääletajate arv protokollis
home_voters	int4	kodushääletanuid	Kodus hääletajate arv protokollis

15. Tabel protocol_type (ProtokolliLiik)

Tabel erinevate protokollitüüpidega.

Atribuut	Andmetüüp	Klassimudel	Selgitus
protocol_type_ID	serial not null	protokolliLiikID	automaatselt genereeritud, unikaalne ID
name	varchar(40)	nimetus	Protokolli tüübi nimetus: 1-jaoskonnas

			pabersedeliga antud hääled 2-e-hääled jaoskonnas 3-välisahääled ringkonnas 4-e-hääled ringkonnas
code	int4	kood	Protokolli tüübi kood

16. Tabel control_time (Aeg)

Andmete esitamise hetk eelhääletamise perioodil ja valimispäeval

Atribuut	Andmetüüp	Klassimudel	Selgitus
control_time_ID	serial not null	aegID	automaatselt genereeritud, unikaalne ID
description	varchar(40)	kirjeldus	Aja määratlus
order	int4	järjestus	Aegade järjestus: 1 - 6. päev enne VP 2- 5. päev enne VP 3 - 4. päev enne VP 4 - VP kell 12:00 5 - VP kell 16:00 6 - VP kell 20:00
prevoting	boolean	eeelhääletamine	Märge, kas tegemist on eeelhääletusperioodi ga (true/false)

17. Tabel protocol_control_time (ProtokollAeg)

Tabel protokolliandmete ja kontrollaeegade seostamiseks

Atribuut	Andmetüüp	Klassimudel	Selgitus
protocol_control_time		ProtokollAeg	
protocol_control_time_ID	serial not null	protokollAegID	automaatselt genereeritud,

			unikaalne ID
control_time_ID	int4	aegID	Viide kontrollajale
department_protocol_ID	int4	protokollID	Viide hääletusprotokollile

18. Tabel question (Küsimus)

Küsimus, mis on pandud rahvahääletusele.

Atribuut	Andmetüüp	Klassimudel	Selgitus
question_ID	serial not null	küsimusID	Automaatselt genereeritud, unikaalne ID
election_ID	int4	valimissündmusID	Viide valimissündmusele
admin_unit_ID	int4	haldusÜksusID	Viide haldusüksusele
election_district_ID	int4	ringkondID	Viide ringkonnale
question	varchar(250)	küsimus	Rahvahääletuse küsimus

19. Tabel question_protocol (KüsimusProtokoll)

Tabel RH küsimuse ja hääletusprotokollide ühendamiseks

Atribuut	Andmetüüp	Klassimudel	Selgitus
question_protocol_ID	serial not null	küsimusProtokollID	Automaatselt genereeritud, unikaalne ID
question_ID	int4	küsimusID	Viide RH küsimusele
department_protocol_ID	int4	protokollID	Viide protokollile
yes-votes	int4	jah-hääli	Poolthäälte arv
no-votes	int4	ei-hääli	Vastuhäälte arv

20. Tabel candidate_protocol (KandidaatProtokoll)

Tabel hoiab kandidaadi häältearvu ühes protokollis

Atribuut	Andmetüüp	Klassimudel	Selgitus
candidate_protocol_ID	serial not null	kandidaatProtokoll ID	Automaatselt genereeritud, unikaalne ID
party_candidate_ID	int4	kandidaatNimekirjasID	Viide kandidaadile
department_protocol_ID	int4	protokollID	Viide protokollile
votes	int4	hääli	Kandidaadile antud häälte arv protokollis

Lisa 2 – XML faili struktuur haldusüksuste ja jaoskondadega RK valimissündmusel

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<electionDepartmentsOutput>
  <electionCode>KOV2013</electionCode>
  <electionTime>20.10.2013</electionTime>
  <reportType>ELECTION_DEPARTMENTS</reportType>
  <generated>2013-10-28T11:27:18.656+02:00</generated>
  <data>
    <adminUnit>
      <adminUnitName>Eesti Vabariik</adminUnitName>
      <ehakCode>0000</ehakCode>
      <numberOfDepartments>400</numberOfDepartments>
      <counties>
        <county>
          <countyName>Harju maakond</countyName>
          <ehakCode>0037</ehakCode>
          <numberOfDepartments>47</numberOfDepartments>
          <parishes>
            <parish>
              <parishName>Anija vald</parishName>
              <ehakCode>0141</ehakCode>
              <numberOfDepartments>3</numberOfDepartments>
              <districts>
                <district>
                  <districtName>Anija valla ringkond nr 1</districtName>
                  <districtNumber>4</districtNumber>
                  <numberOfDepartments>3</numberOfDepartments>
                  <departments>
                    <department>
                      <departmentID>54678</departmentID>
                      <departmentName>Anija valla jaoskond</departmentName>
```

```
        <departmentNumber>1</departmentNumber>
        <departmentArea></departmentArea>
    </department>
</departments>
</district>
</districts>
</parish>
</parishes>
</county>
</counties>
</adminUnit>
</data>
</electionDepartmentsOutput>
```

Lisa 3 – XML faili struktuur ühe valimissündmuse kandidaatidega KOV valimissündmusel

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<electionCandidatesOutput>
  <electionCode>KOV2013</electionCode>
  <electionTime>20.10.2013</electionTime>
  <reportType>ELECTION_CANDIDATES</reportType>
  <generated>2013-10-28T11:27:18.656+02:00</generated>
  <data>
    <adminUnit>
      <adminUnitName>Eesti Vabariik</adminUnitName>
      <ehakCode>0000</ehakCode>
      <numberOfCandidates>10000</numberOfCandidates>
      <counties>
        <county>
          <countyName>Harju maakond</countyName>
          <ehakCode>0037</ehakCode>
          <numberOfCandidates>47</numberOfCandidates>
          <parishes>
            <parish>
              <parishName>Anija vald</parishName>
              <ehakCode>0141</ehakCode>
              <numberOfCandidates>3</numberOfCandidates>
              <districts>
                <district>
                  <districtName>Anija valla ringkond nr 1</districtName>
                  <districtNumber>4</districtNumber>
                  <numberOfCandidates>30</numberOfCandidates>
                  <parties>
                    <party>
                      <partyID>54678</partyID>
                      <partyName>Anija valimisliit</partyName>
```

```
<partyCode>AV</partyCode>
  <candidates>
    <candidate>
      <candidateID>83920</candidateID>
      <forename>Mari-Liis</forename>
      <surname>Männik</surname>
      <candidateRegNumber>101</candidateRegNumber>
<sequenceNumberInAdminUnit>1</sequenceNumberInAdminUnit>
      <politicalParty>Eesti Testerakond</politicalParty>
      <birthday>11.01.2000</birthday>
      <sex>N</sex>
      <votes>100</votes>
      <elected>true</elected>
    </candidate>
  </candidates>
</party>
</parties>
</district>
</districts>
</parish>
</parishes>
</county>
</counties>
</adminUnit>
</data>
</electionCandidatesOutput>
```

Lisa 4 – XML faili struktuur kõikide kandidaatidega

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<allCandidatesOutput>
  <reportType>ALL_CANDIDATES</reportType>
  <generated>2019-10-28T11:27:18.656+02:00</generated>
  <data>
    <candidates>
      <candidate>
        <candidateID>83920</candidateID>
        <elections>
          <election>
            <electionCode>RK2015</electionCode>
            <electionTime>01.03.2015</electionTime>
            <forename>Mari-Liis</forename>
            <surname>Männik</surname>
            <candidateRegNumber>333</candidateRegNumber>
            <adminUnit>Tartu linn</adminUnit>
            <districtNumber>10</districtNumber>
            <party>Üksikkandidaat</party>
            <votes>4000</votes>
            <elected>true</elected>
          </election>
          <election>
            <electionTime>25.05.2014</electionTime>
            <electionCode>EP2014</electionCode>
            <forename>Mari</forename>
            <surname>Männik</surname>
            <candidateRegNumber>101</candidateRegNumber>
            <adminUnit>Eesti Vabariik</adminUnit>
            <districtNumber>1</districtNumber>
            <party>Erakond Edasi</party>
            <votes>3500</votes>
            <elected>>false</elected>
          </election>
        </elections>
      </candidate>
    </candidates>
  </data>
</allCandidatesOutput>
```



```
</election>
<election>
  <electionTime>20.10.2013</electionTime>
  <electionCode>KOV2013</electionCode>
  <forename>Mari</forename>
  <surname>Männik-Kuusik</surname>
  <candidateRegNumber>222</candidateRegNumber>
  <adminUnit>Tallinn, kesklinna linnaosa</adminUnit>
  <districtNumber>2</districtNumber>
  <party>Valimisliit Tallinna kodanik</party>
  <votes>5500</votes>
  <elected>true</elected>
</election>
</elections>
</candidate>
</candidates>
</data>
</allCandidatesOutput>
```

Lisa 5 – XML faili struktuur detailse hääletamisest osavõtu infoga RK valimissündmusel

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<electionParticipationOutput>
  <electionCode>RK2019</electionCode>
  <electionTime>03.03.2019</electionTime>
  <reportType>ELECTION_PARTICIPATION</reportType>
  <generated>2019-03-08T11:27:18.656+00:00</generated>
  <data>
    <electionParticipation>
      <adminUnitName>Eesti Vabariik</adminUnitName>
      <ehakCode>0000</ehakCode>
      <overallStatistics>
        <eligibleVoters>887420</eligibleVoters>
        <voters>565045</voters>
        <validBallots>561141</validBallots>
        <invalidBallots>3904</invalidBallots>
        <advanceVoters>99177</advanceVoters>
        <homeVoters>5127</homeVoters>
        <iVoters>247232</iVoters>
        <foreignVoters>2107</foreignVoters>
      </overallStatistics>
      <districts>
        <district>
          <districtNumber>7</districtNumber>
          <districtName>Ida-Virumaa</districtName>
          <votersInList>63380</votersInList>
          <voted>34863</voted>
          <percentage>55.0</percentage>
          <districtSimpleQuota>4926.00</districtSimpleQuota>
          <districtNumberOfMandates>7</districtNumberOfMandates>
          <participation>
            <participationRow>
              <name>Ida-Viru maakond</name>
              <votersInList>63359</votersInList>
              <votedCount>29396</votedCount>
              <percentage>46.4</percentage>
            </participationRow>
          </participation>
        </district>
      </districts>
    </electionParticipation>
  </data>
</electionParticipationOutput>
```

```
</participationRow>
<participationRow>
  <name>Välisriigid</name>
  <votersInList>21</votersInList>
  <votedCount>21</votedCount>
</participationRow>
<participationRow>
  <name>E-hääled</name>
  <votedCount>5446</votedCount>
</participationRow>
<adminUnit>
  <adminUnitName>Alutaguse vald</adminUnitName>
  <eligibleVoters>3298</eligibleVoters>
  <votedCount>1870</votedCount>
  <percentage>56.7</percentage>
</adminUnit>
</participation>
</district>
</districts>
</electionParticipation>
</data>
</electionParticipationOutput>
```

Lisa 6 – XML faili struktuur hääletamis- ja valimistulemusega - RK valimissündmus

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<electionResultOutput>
  <electionCode>RK2015</electionCode>
  <electionTime>01.03.2015</electionTime>
  <reportType>ELECTION_RESULT</reportType>
  <generated>2015-03-01T11:27:18.656+02:00</generated>
  <data>
    <adminUnitName>Eesti Vabariik</adminUnitName>
    <ehakCode>0000</ehakCode>
    <fivePercentThreshold>28057,1</fivePercentThreshold>
    <votes>561141</votes>
    <parties>
      <party>
        <partyName>Testerakond 1</partyName>
        <acronym>TE</acronym>
        <partyNumberOfMandates>33</partyNumberOfMandates>
        <totalPartyVotes>162363</totalPartyVotes>
        <percentOfVotes>28,9</percentOfVotes>
        <lastElectionTotalPartyVotes>20.0</lastElectionTotalPartyVotes>
        <lastElectionPartyNumberOfMandates>30</lastElectionPartyNumberOfMandates>
        <candidates>
          <candidate>
            <forename>Mari-Liis</forename>
            <surname>Männik</surname>
            <candidateRegNumber>101</candidateRegNumber>
            <candidateID>10101</candidateID>
            <sequenceNumber>1</sequenceNumber>
            <districtNumber>2</districtNumber>
            <candidateVotes>15000</candidateVotes>
            <elected>true</elected>
            <reserved>false</reserved>
            <quota></quota>
            <vrd>3,3177</vrd>
            <mandateType>Isikumandaat</mandateType>
          </candidate>
          <candidate>
            <forename>Toomas</forename>
            <surname>Kaasik</surname>
            <candidateRegNumber>140</candidateRegNumber>
            <candidateID>10403</candidateID>
```

```
<sequenceNumber>36</sequenceNumber>
<districtNumber>5</districtNumber>
<candidateVotes>10021</candidateVotes>
<elected>>false</elected>
<reserved>>true</reserved>
<quota></quota>
<vrd></vrd>
<mandateType></mandateType>
</candidate>
</candidates>
</party>
</parties>
</data>
</electionResultOutput>
```

Lisa 7 – XML faili struktuur hääletamis- ja valimistulemusega RH valimissündmusesel

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<electionResultOutput>
  <electionCode>RH2030</electionCode>
  <electionTime>01.05.2030</electionTime>
  <reportType>ELECTION_RESULT</reportType>
  <generated>2030-05-01T11:27:18.656+02:00</generated>
  <data>
    <adminUnit>
      <adminUnitName>Eesti Vabariik</adminUnitName>
      <ehakCode>0000</ehakCode>
      <totalVotes>55311</totalVotes>
      <yesVotesCount>369657</yesVotesCount>
      <yesVotePrecentage>66.83</yesVotePrecentage>
      <noVotesCount>183454</noVotesCount>
      <noVotesPercentage>33.17</noVotesPercentage>
      <ballotVotes>
        <yesVotesCount>184828</yesVotesCount>
        <yesVotePrecentage>61.5</yesVotePrecentage>
        <noVotesCount>101000</noVotesCount>
        <noVotesPercentage>38.1</noVotesPercentage>
      </ballotVotes>
      <iVotes>
        <yesVotesCount>160128</yesVotesCount>
        <yesVotePrecentage>70.2</yesVotePrecentage>
        <noVotesCount>90028</noVotesCount>
        <noVotesPercentage>29.8</noVotesPercentage>
      </iVotes>
      <foreignVotes>
        <yesVotesCount>986</yesVotesCount>
        <yesVotePrecentage>71.0</yesVotePrecentage>
        <noVotesCount>132</noVotesCount>
        <noVotesPercentage>29.0</noVotesPercentage>
      </foreignVotes>
    </adminUnit>
    <counties>
      <county>
        <countyName>Harjumaa</countyName>
        <ehakCode>0034</ehakCode>
        <totalVotes>55311</totalVotes>
        <yesVotesCount>37670</yesVotesCount>
        <yesVotePrecentage> 69.86</yesVotePrecentage>
      </county>
    </counties>
  </data>
</electionResultOutput>
```

```
<noVotesCount>16251</noVotesCount>  
<noVotesPercentage>30.14</noVotesPercentage>  
</county>  
</counties>  
</adminUnit>  
</data>  
</electionResultOutput>
```

Lisa 8 – XML faili struktuur detailse hääletamistulemusega RK valimissündmusel

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<electionVotingResultInCountiesOutput>
  <electionCode>RK2025</electionCode>
  <electionTime>01.03.2025</electionTime>
  <reportType>ELECTION_VOTING_RESULT_IN_COUNTIES</reportType>
  <generated>2025-03-01T11:27:18.656+02:00</generated>
  <data>
    <districts>
      <district>
        <districtName>Harju- ja Raplamaa</districtName>
        <districtNumber>4</districtNumber>
        <districtQuota>6080.60</districtQuota>
        <districtNumberOfMandates>15</districtNumberOfMandates>
        <votesDistributionByParties>Testerakond 2
          </votesDistributionByParties>
        <acronym>TE</acronym>
        <totalVotesInDistrict>1078
          </totalVotesInDistrict>
        <partyQuotaInDistrict>0.1782</partyQuotaInDistrict>
        <iVotesInDistrict>102</iVotesInDistrict>
        <foreignVotesInDistrict>100</foreignVotesInDistrict>
        <votesInAdminUnit>
          <name>Harjumaa</name>
          <value>100</value>
        </votesInAdminUnit>
        <votesInAdminUnit>
          <name>Raplamaa</name>
          <value>100</value>
        </votesInAdminUnit>
        <candidates>
          <candidate>
            <forename>Mari-Liis</forename>
            <surname>Männik</surname>
            <candidateRegNumber>101</candidateRegNumber>
            <candidateID>10120</candidateID>
            <votesInUnits>
              <totalVotes>200</totalVotes>
              <iVotes>100</iVotes>
              <foreignVotes>50</foreignVotes>
              <votesInAdminUnit>
```



```

        <name>Harjumaa</name>
        <value>30</value>
    </votesInAdminUnit>
    <votesInAdminUnit>
        <name>Raplamaa</name>
        <value>20</value>
    </votesInAdminUnit>
</votesInUnits>
</candidate>
</candidates>
<counties>
    <county>
        <countyName>Harjumaa</countyName>
        <ehakCode>0034</ehakCode>
        <totalVotesInCounty>14643</totalVotesInCounty>
        <parties>
            <party>
                <partyName>Testerakond 1</partyName>
                <totalVotesInCounty>392</totalVotesInCounty>
                <candidates>
                    <candidate>
                        <forename>Mari-Liis</forename>
                        <surname>Männik</surname>
                        <candidateRegNumber>101 </candidateRegNumber>
                        <candidateID>10120</candidateID>
                        <votesInAdminUnit>
                            <name>Anija vald</name>
                            <value>10</value>
                        </votesInAdminUnit>
                    </candidate>
                </candidates>
            </party>
        </parties>
    </county>
</counties>
</district>
</districts>
</data>
</electionVotingResultInCountiesOutput>

```