



Statistikanõukogu
20.06.2018

Ülevaade REGREL 2021 ettevalmistustest

Statistikaamet, REGREL projektijuht, Diana Beltadze

20.06.2018



Sisukord

- Ajakava
- REGREL andmekogude tööd
- Metoodika arendus
- Suurandmete projektid
- II prooviloenduse ettevalmistustööd

REGREL ettevalmistus 2010–2020

2021–2022
REGREL

2018 (31.12.2018 loendusmoment II
prooviloenduseks)

1. REGREL II prooviloenduse ettevalmistamine (programm, meetod, aeg, reeglid jn)
2. Metoodika arendus
3. Loenduse tunnuste andmekvaliteedi parendustööd
4. Andmekogude arendustööd

2019
II prooviloendus

1. REGREL metoodika täiendamine ja testimine
2. Andmekogudes andmekvaliteedi parendamine
3. ADS-süsteemi juurutamine
4. Andmehõive üle x-tee

REGREL projekti olulisemad tööd

LEKU CAPI ja CAWI(2)
järgude avamine ja
uuringu läbiviimine

15.10.2018

II prooviloenduse
programmi, momendi
ja meetodi kinnitamine

18.09.2018

TÖR andmekogumise
käivitamine ja
koolitused

30.november 2018

Prooviloenduseks
ettevalmistustööd on tehtud

15.12.2018

REGREL andmekogud

REGREL projekti andmekogude suuremad alamprojektid

1. TÖR registri rakenduse täiendavad arendused

Tähtaeg: 18.06.2018

2. Maa-ameti aadressi-korrastamise projekt

Tähtaeg: november 2018

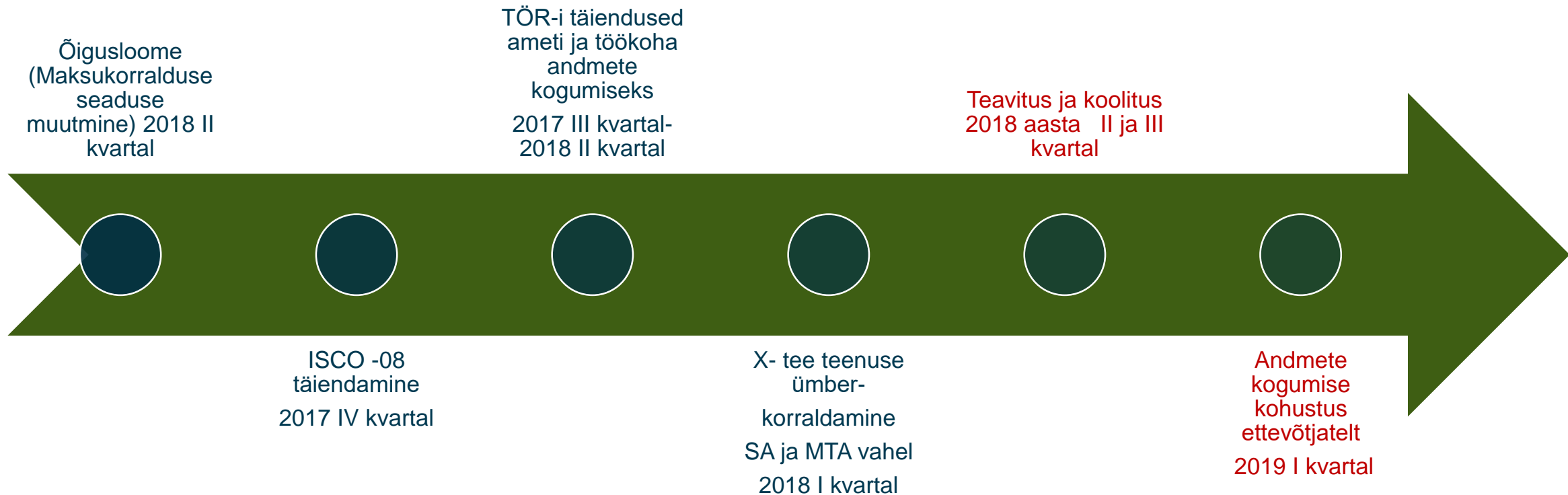
3. Ehitisregistri andmete korrastamine tulenevalt seadusest

Tähtaeg: detsember 2020

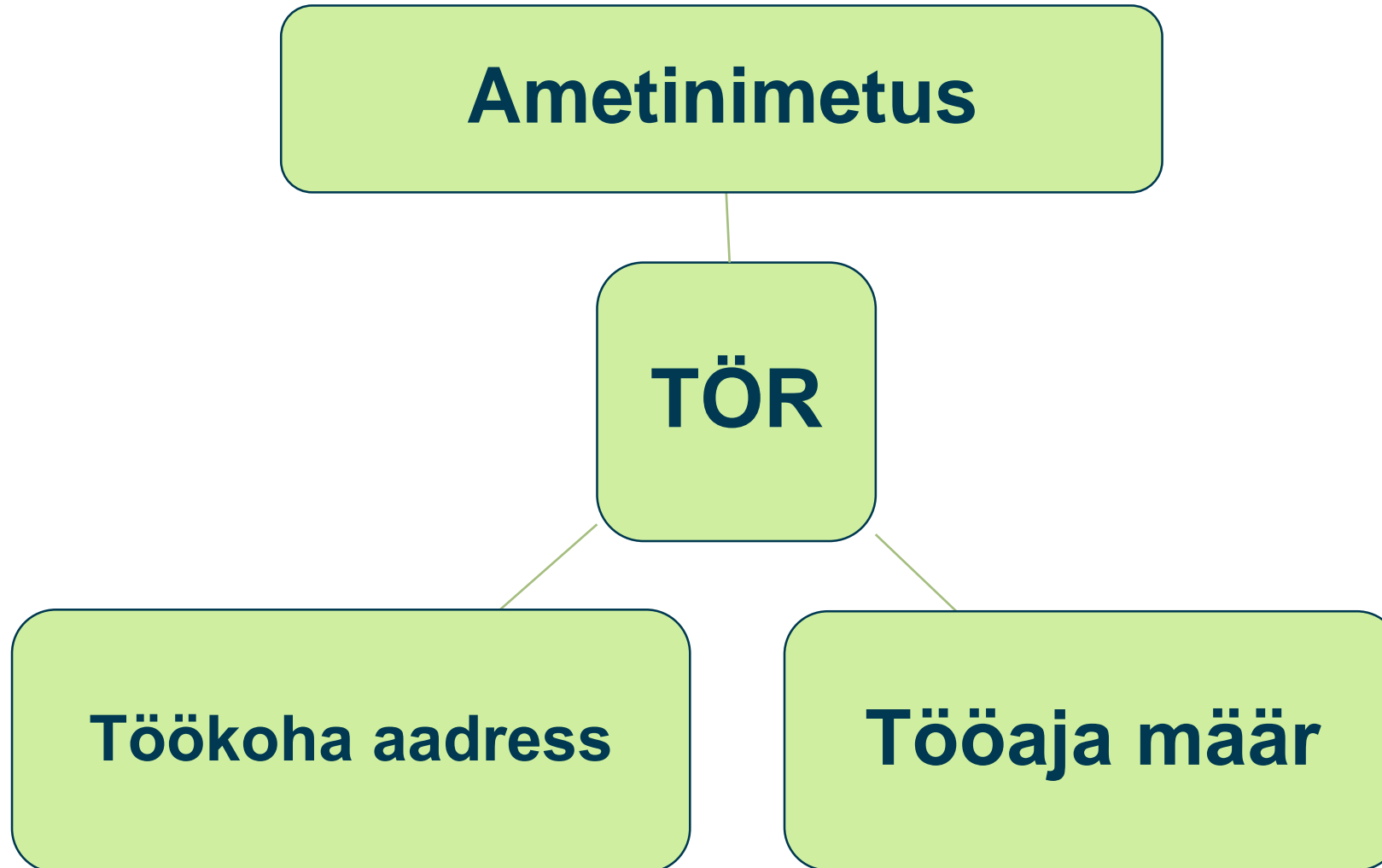
4. Rahvastiku-registriandmete kvaliteedi tööd

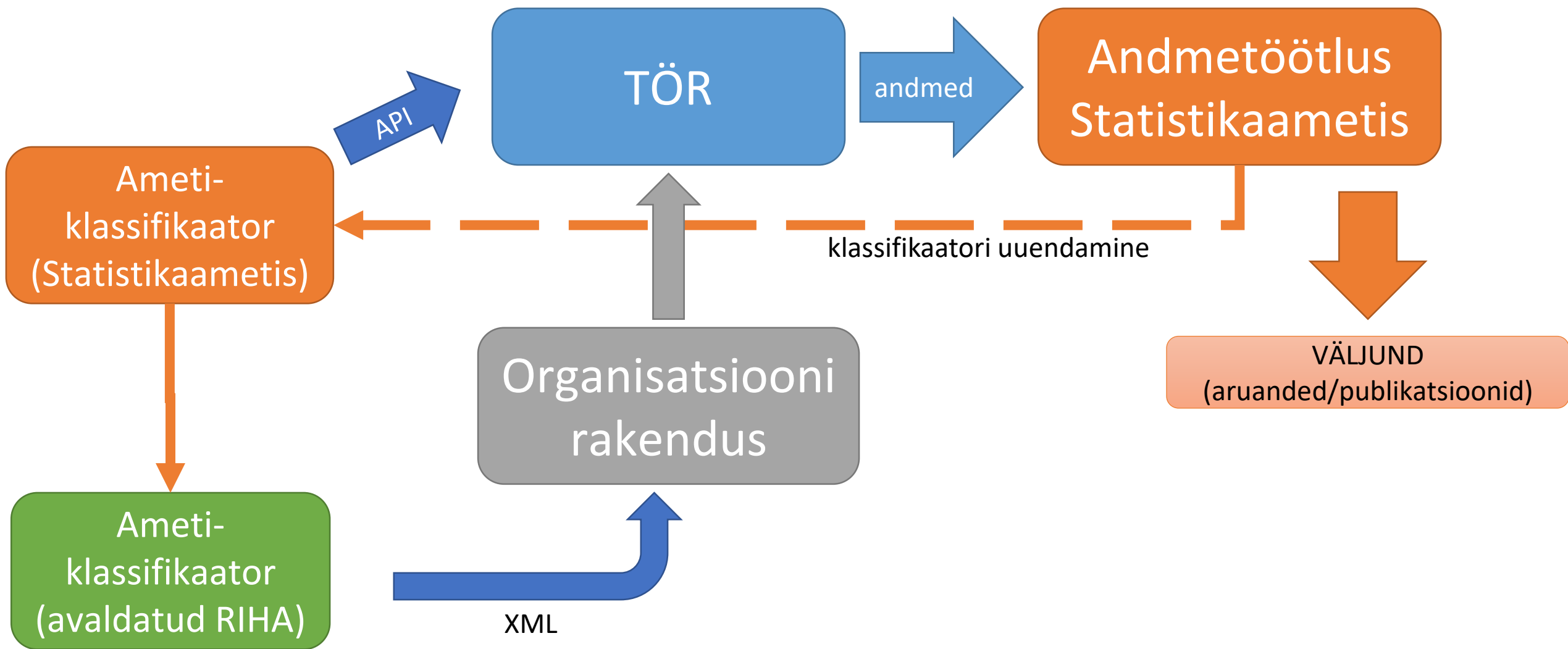
Tähtaeg: detsember 2020

TÖR-i teekaart ameti ja töökoha andmete kogumiseks



Uued tunnused, mida hakatakse TÖR-i koguma

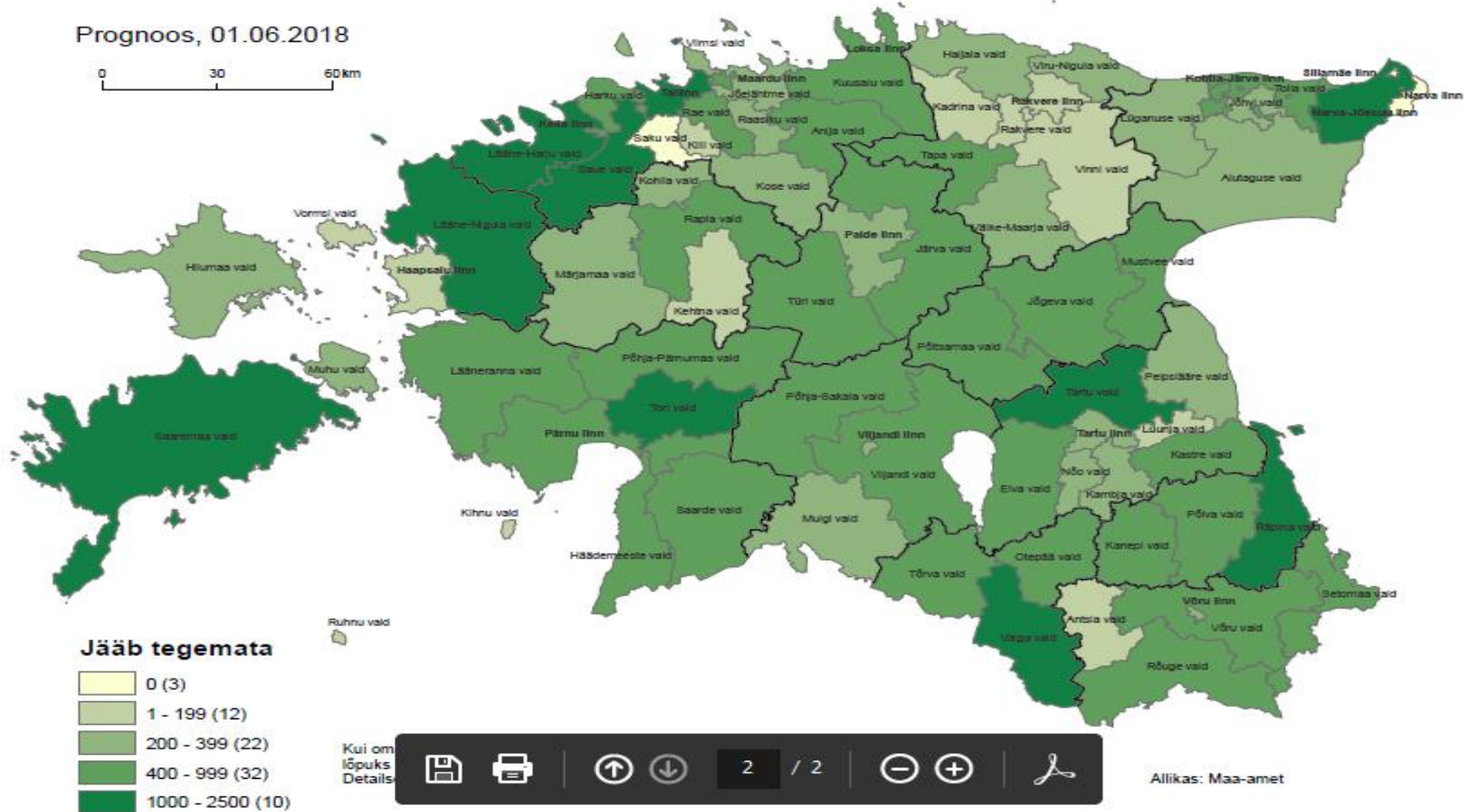




Aadressiandmete korrastamise projekt

- 1.juuni seisuga on Maa-amet raporteerinud:
tehtud **37,3%** töödest ja kulunud on **77%** tööde tegemiseks
plaanitud ajast. Prognoos on **1%** võrra paranenud: **50,6%**
töödest saab tehtud novembriks.

Tegemata tööde prognoos KOV-de lõikes



Lahendused

- Maa-amet esitas taotluse jätkuprojektiks
- Suurendatakse tariifide tasusid ca 7%, et kompenseerida inflatsiooni mõju ja komponendid, mida tänased tööde tariifid ei sisalda(nt tänava sildid)
- Maa-amet ja RM arutavad täiendavate toetusmeetmete lisamise võimalusi KOV-le.
- Maa-amet tegi 12.06 ELVL volikogu koosolekul ettekande aadressandmete korrastamise teemal.

Ehitisregistri parendustööd

- MKM tellis EHR-i ortofotode arenduse, millega tuuakse ADS-st EHR-i automaatselt objektid koos EHR-koodi, aadressi, ruumikuju ja Maa-ameti poolt tuvastatud pindala ja kõrgusega.
- Selle tööga saab valmis EHR-is objektid KOV-idele valideerimiseks ja on täiendavate puuduvate tehniliste andmete lisamiseks valmisolek tagatud.
- Samas on ortofotode funktsionaalsuse kasutuselevõtt viibinud, kuna oodati EHR-ADS x-tee V6 teenuste kasutuselevõtmise taga, et lahendust saaks testida.
- Eeltoodud EHR-i funktsionaalsus on KOV-idele kasutatav kui lõpeb selle tänaseks pikale veninud testimine.

Andmekogude x-teenus olemas 10 andmekogu pidajal

Olemas x-tee teenus Statistikaametiga

Rahvastikuregister ver 6

Aadressiandmete süsteem (ADS) ver 5

Kohustusliku kogumispensioni register (KOPIS) ver 5

EHIS ver 5

RIK andmekogud (3) ver 6

Töötukassa ver 6

EHR ver 6

KVKR ver 5

X-tee teenus loomisel

Andmekogu andmevahetuskanal hetkel FTB	Tähtaeg
Riiklik elamis- ja töölubade register (ETR/KMAIS/RKAS)	2018 IV kv
Maksudohustuslaste register (MKR)	2018 III kv
Töötamise register (TÖR)	2018 III kv
Sotsiaalteenuste ja -toetuste andmeregister (STAR)	2018 III kv
Ravikindlustuse andmekogu (KIRST)	2018 IV kv
Retseptikeskus (RETS)	2018 IV kv
Riiklik pensionikindlustuse register (SKAIS)	2018 IV kv
Vangide ja kriminaalhooldusaluste register (VANGIS/KIR)	2018 IV kv
Liiklusregister	2018 IV kv
E-toimiku süsteem (e-toimik)	Ei loodagi
Surma põhjuste register (SPR)	2018 IV kv
Meditsiiniline sünniregister (EMSR)	2018 IV kv

REGREL metoodika arendus

Lõppenud tööd

- Partnerlusindeksi arendustöö
- Mobiilpositsioneerimise andmete pilootprojekt
- Suurandmete pilootprojekt eluruumide loendamiseks

Indeksite arendus

- Töö indeksitega jätkub
- Arvestada tuleb andmekogudes toimuvate muutustega
 - Nt uus infosüsteem (võimalik uus struktuur, uued klassifikaatorid jmt)
- Arvestada tuleb nii seaduste muudatuste kui ka uute andmekogude/allikatega
 - Nt töövõime (TETRIS alates 2016)

Partnerluse indeks

- Registritest on tarvis leida partnereid siduvaid nn partnerluse märke.
- Nende põhjal tuleb defineerida partnereid siduv mudel – partnerluse indeks, mis omistatakse isikute paarile ja mis näitab tõenäosust, et need isikud on partnerid (moodustavad perekonnatuuma).
- Indeksi toimimiseks on tarvis hinnata mudeli parameetreid, mudelit testida, anda hinnang otsustus-vigade tõenäosusele.

Partnerluse märgid

- Abielu
- Poolabielu (abielu registreerimisdokumentides on vastuolu)
- Lahutus
- Poollahutus (lahutuse registreerimisdokumentides on vastuolu)
- Ühiste alaealiste (<19 a) laste olemasolu
- Ühine sissetulekute deklaratsioon Maksuametile
- Elatise nõue
- Ühine eluruum
- Ühine laen
- Ühisomand
- Eluruum_omand (isiku eluruumiks on partneri omand)

Paiknemise indeks

- Kui partnerid on leitud, moodustatakse perekond, lisades partnerite juurde nende kõik alaealised lapsed
- Perekonnale on tarvis määrata ühine eluruum. See võib olla kummagi partneri eluruum või ka hoopis erinev ruum
- Leibkonna sidumiseks eluruumiga on tarvis paiknemise märke

Töös

- Paiknemise indeksi arendus
- Indeksitel põhinev rände ja hargmaisuse metoodika arendus
- Ametinimetuse imputeerimise metoodika välja töötamine tulenevalt seaduse muudatustest (VÕS-lepingulised)

Leibkonna ja elukoha võrdlusuuring(LEKU)

- Palju keeldumisi CAP-s
- Palju kirju tuleb tagasi SA-sse (peamine põhjus: puudub postkast, vigane aadress, tühi maja vm)

Suurandmete projektid

Mobiiliandmete projekt

- Projekti eesmärk oli mobiilioperaatorilt koguda andmeid inimeste nõusolekul, et saaks nende reaalsel elukohta ja mobiiliandmeid kokku viia ja uurida, kui täpselt mobiilandmed ja elukoht kokku lähevad.
- Viidi läbi töötlus ja siis arvutati mobiilandmetest iga inimese jaoks ankurpunktid. Nendeks olid elukoht, tööajakoht ja muud regulaarsed kohad.
- Tehti analüüs, mis pidi selgitama välja täpsust mõjutavad faktorid, ning kaardistada veaallikad.
- Lõpuks pakuti välja metoodikat, kuidas mobiilandmeid valideerimiseks kasutada.

Meetod

- Inimeste liikumiste kaardistamisel lähtuti tegevuspõhisest (*activity-based*) tegevusruumide ja sotsiaalsete võrgustike analüüsi meetoditest, kasutades selleks passiivse MPD andmeid.
- Passiivne MPD on andmeanalüüsi tehnoloogia, kus iga mobiiltelefoni kasutaja (isiku) kohta leitakse kvantitatiivse analüüsi abil karakteristikud (tegevusajad, ankurpunktid tegevusruumis, liikumisvektorid, teekonnad, liikumisruum, jt) mobiilioperaatori poolt kogutud kliendi toimingute logis (näiteks sisenevad ja väljuvad kõned, SMS-id, MMS-id, andmekasutus) sisalduvate aegade ja asukohtade põhjal (mobiilsidemasti leviala või interpoleerimise meetodiga parandatud täpsusega).
- Eelkõige kasutatakse anonümiseeritud passiivseid MPD-d andmeid, kuid telefoni kasutaja nõusolekul on võimalik siduda konkreetne telefoni kasutaja konkreetse isikuga.

Tulemused(VVLK 18.04.2018)

- 93% respondentide puhul (242/261-st) oli võimalik MPD-d kasutades välja arvutada koduankurpunkt või kodu-töö ankurpunkt. Mõnede isikute puhul ei ole MPD meetodil võimalik ankurpunkte arvutada ja seega MPD-d elukoha valideerimiseks kasutada, sest nende telefonikasutus ei ole piisav või kaasnevad andmete töötlemise ja tõlgendamisega, sh ankurpunktide arvutamise ja mingisugused tehnilised probleemid.
- 82% puhul asus elukoha aadress koduankurpunkti või kodu-töö ankurpunkti levialas.
- 97% puhul asus ankru leviala tsentroid samas maakonnas.
- 87% puhul samas omavalitsusüksuses ja 67% puhul samas asustusüksuses. Asustusüksuse ja ankurpunkti leviala tsentroidi võrdlemise miinuseks on, et harilikult katavad mobiilimastide levialad mitu asustusüksust (eriti hõreasustusega piirkondades) ning see ei pruugi tähendada, et ankurpunkt on tingimata valesti arvutatud.

- Kauguste suhtes on elukohti võimalik valideerida kõige lihtsamini, kuid see meetod ei võimalda hinnata valideerimise (taseme) usaldusväärsust.
- Kauguspõhise valideerimise metoodikas ei võeta arvesse mobiilside levialade näitajaid. Seetõttu töötab kauguspõhine valideerimine eeskätt lihtsates olukordades (üks rahvastikuregistri aadress, üks statistiliselt tugev koduankurpunkt selle läheduses), aga ei suuda anda usaldusväärseid vastuseid keerulisemates olukordades. Ühtlasi ei võimalda kauguspõhine lähenemine hinnata valideerimise usaldusväärsust, kuna valideerimise usaldusväärsus ei sõltu otseselt kaugusest. Samas on see lähenemine metoodiliselt lihtsam ning väiksema arendustööga rakendatav.
- Levialapõhisel valideerimisel võetakse arvesse registriaadressi paiknemist levialade suhtes. See asjaolu võimaldab hinnata iga valideerimise usaldusväärsust ja anda madalama tasemega valideerituse ka keerulisematele ja piirijuhtudele. Levialapõhise valideerimise metoodika rakendamise teeb mõnevõrra töömahukamaks keerulisema otsustuspuu väljatöötamine ja selle abil leitud tulemuste kontrolli vajadus.
- Ankurpunktide määramise metoodika sisaldab vigu, eriti mittekorrapärase ajalis-ruumilise käitumisega inimeste ankurpunktide määramisel.

Suurandmete kasutamine eluruumide statistikas

Eluruumide üldkogum koosneb asustatud (76%) ja asustamata eluruumidest (24%).

(REL2011 asustamata eluruume on 14.4% eluruumide koguarvust).

Üldkogumi asustatud osa moodustamise alusel on Rahvastiku registris registreeritud elukoha aadress. On teada et 20%-25% püsielanikest elavad mitte seal, kus nad on ennast registreerinud.

Eesmärk 1. Kasutades elektriandmeid täpsustada eluruumi asustatust.

Eesmärk 2. Pärast partnerlusindeksi rakendamist tekib vajadus isikute elukohtade ümber paigutamisest, elektritarbimise andmed peaksid siin ka aitama.

Andmeallika miinused

- Andmete kirjeldus puudulik
- Korteriühistute korral ei saa eristada üksikute leibkondade tarbimist
- Aadressid vabatekstina, üks-ühene linkimine teiste allikatega õnnestub ainult osaliselt
- Rentnikud (eluruumid ja äripinnad) – ei ole elektrilepingu omanikud, nende tarbimine omistatakse omanikule, mitte neile endile

Järeldused

- Kõikidele eluruumidele ei õnnestunud elektritarbimise infot juurde linkida, kuid see info (kuigi osaline) on kindlasti kasutatav eluruumi asustatuse ja leibkonna/partnerite elukoha täpsustamisel.

Edasised sammud

- Eleringi 2016,2017, 2018 andmed. Algoritmide ja meetodite täiendamine.
- Elektriandmete info kasutamine
 - eluruumi üldkogumi moodustamisel;
 - partnerlusindeksis.

II prooviloenduse ettevalmistustööde olulisemad verstapostid

1. REGREL II prooviloenduse metoodika kirjeldus on koostatud 31.05.2018
2. Loendustunnuste moodustamine 30.10.2018
3. Prooviloenduse meetodi ja momendi kinnitamine 18.09.2018
4. Prooviloenduse väljundi kinnitamine 10.09.2018
5. Prooviloenduse programmi koostamine 25.10.2018
4. Prooviloenduse kvaliteedi kriteeriumite välja töötamine 10.05.2018

Kokkuvõte

- REGREL ettevalmistuses on fookuses metoodika arendustööd
- Käivitatud on uuring LEKU indeksite testimiseks
- On koostatud metoodika ülevaade II prooviloenduseks
- Kõik EL kohustuslikud loendustunnused on registritest kogutavad 20.06.2018 seisuga

