



## **Geograafilise ekspertsüsteemi “Koolivõrk” loomine**

**Koostööleping nr 10-10/895**

Lõpparuanne 31. jaanuar 2005

Töö tellija: Haridus- ja  
Teadusministeerium

Töö täitja: AS Regio  
Tõnu Randaru  
Jüri Jagomägi  
Edgar Sepp  
Andres Ideon  
Mart Kelder  
Margus Tiru

Tartu 2005

# Sisukord

1.	Sissejuhatus .....	4
1.1.	Koostöölepingu objekt .....	5
1.2.	Planeeritud tööd ja tegevused.....	6
2.	Andmestikud .....	7
2.1.	Koolide teeninduspiirkonnad .....	7
2.2.	Koolide register .....	9
2.3.	Õpilaste- ja rahvastikuregister .....	11
2.4.	2000.a. rahvaloenduse andmebaas .....	15
2.5.	Pedagoogide register .....	16
2.6.	Projekti tulemusena loodud andmebaas .....	17
3.	Teemakaardid .....	20
3.1.	Koolivõrk ja õpilaste tihedus .....	20
3.2.	Koolide avamised ja sulgemised .....	21
3.3.	Õpilaste ränne ja tõmbekeskused .....	23
3.4.	Äärelise asendi ja/või hõreda asustustihedusega vallad .....	26
4.	Koolide veebirakendus .....	27
4.1.	Funktsionaalsus .....	27
4.2.	Kasutatud andmestikud .....	30
4.3.	Koolide andmestike uuendamine ja arendus .....	30
5.	Ekspertsüsteem "Koolivõrk" .....	32
5.1.	Andmestikud ja nende uuendamine .....	32
5.2.	Funktsionaalsus .....	33
5.3.	Kasutajaliides.....	36
5.4.	Tulemuste esitlus.....	40
6.	Tähelepanekud ja soovitused .....	41
6.1.	Algandmestikud ja nende süstematiseerimine .....	41
6.2.	Ekspertsüsteemi edasiarendamine ja visioonid .....	42
7.	Kokkuvõte.....	44

## JOONISED

<b>Joonis 1.</b> Algkoolide teeninduspiirkonnad (andmed seisuga 03.12.2004).....	7
<b>Joonis 2.</b> Põhikoolide teeninduspiirkonnad (andmed seisuga 03.12.2004).....	8
<b>Joonis 3.</b> Õpilaste paiknemine 2003. aastal (rahvastikuregistri elukoha järgi).....	11
<b>Joonis 4.</b> Rahvastiku tihedus 2000.a. rahvaloenduse andmetel.....	16
<b>Joonis 5.</b> Andmebaasi struktuur.....	18
<b>Joonis 6.</b> Õpilaste tihedus ja üldhariduskoolide tüübid 2003.a. seisuga.....	20
<b>Joonis 7.</b> Õpilaste tihedus ja paiknemine 2000.a rahvaloenduse andmetel.....	21
<b>Joonis 8.</b> Koolide avamised, sulgemised ja ühendamised 10 aasta jooksul (2. kooliaste).....	22
<b>Joonis 9.</b> Õpilaste ränne Harjumaal 2003. aasta andmete põhjal.....	23
<b>Joonis 10.</b> Õpilaste ränne Lääne-Virumaal 2003. aasta andmete põhjal.....	24
<b>Joonis 11.</b> Tõmbekeskused 2003.a. õpilaste rände järgi.....	25
<b>Joonis 12.</b> Äärelise asendi ja/või hõreda asustustihedusega vallad.....	26
<b>Joonis 13.</b> Koolide veebirakenduse avavaade.....	27
<b>Joonis 14.</b> Koolide veebirakenduse navigatsiooni nupud.....	27
<b>Joonis 15.</b> Koolide valiku kriteeriumid.....	28
<b>Joonis 16.</b> Kiirpäringu tulemused koolide kohta.....	29
<b>Joonis 17.</b> Detailse päringu tulemus kooli kohta.....	30
<b>Joonis 18.</b> Ekspertsüsteemi põhimõtteline skeem.....	32
<b>Joonis 19.</b> Ekspertsüsteemi hindamissüsteemi põhimõtteline skeem.....	35
<b>Joonis 20.</b> Teemakaartide valiku dialoogiaken.....	36
<b>Joonis 21.</b> Koolide valiku dialoogiaken.....	37
<b>Joonis 22.</b> Teemakaardi ja -info parameetrite muutmise dialoogiaken.....	37
<b>Joonis 23.</b> Õpilaste ränne dialoogiaken.....	38
<b>Joonis 24.</b> Kooli hindamiseks kasutatavate näitajate valimise dialoogiaken.....	38
<b>Joonis 25.</b> Näitajate valimise dialoogiaken koos andmete sisestuse nupuga.....	39
<b>Joonis 26.</b> Kooli andmete sisestamise dialoogiaken.....	39
<b>Joonis 27.</b> Valitud näitajate kaalude, vahemik ja hindepunktide seadmise dialoogiaken.....	40

## TABELID

<b>Tabel 1.</b> Koolide registri andmetabelite veergude nimetused.....	9
<b>Tabel 2.</b> Koolid, millel on muutunud kood.....	10
<b>Tabel 3.</b> Koolid, kus koolide registris oli koodi asemel registreerimisnumber.....	10
<b>Tabel 4.</b> Õpilaste- ja rahvastikuregistri andmetabelite veergude nimetused.....	12
<b>Tabel 5.</b> Koolide registrist puuduvad koolid, millele viidatakse õpilaste registrist.....	12
<b>Tabel 6.</b> Koolide registri ja õpilaste registri (2003 a. seis) vahelised ebakõlad.....	13
<b>Tabel 7.</b> Rahvastikuregistri ja õpilaste registri vahelised elukoha erinevused.....	13
<b>Tabel 8.</b> 2003. a. õpilaste nimekirjade geokodeerimise täpsus.....	14
<b>Tabel 9.</b> Õpilaste registri ja koolide registri vaheline õpilaste arvu võrdlus.....	14
<b>Tabel 10.</b> Suuremad erinevused õpilaste ja koolide registri õpilaste arvudes kooli kohta.....	14
<b>Tabel 11.</b> Pedagoogide registri andmetabelite veergude nimetused.....	16
<b>Tabel 12.</b> Näiteid erinevatest andmebaasis esinevatest aadressi kirjetest.....	17

# 1. Sissejuhatus

Infotehnoloogia suured võimalused ja üha kiirem ruumiandmete uuenemine on loonud olukorra, kus nende efektiivne kasutamine nõuab geoinformaatika valdkonna tippasjatundjailt järjest rohkem tähelepanu ja aega orienteerumiseks ruumiandmetes, nende tähenduses ja kvaliteedis. Lisaks kulub palju energiat ja aega tulemuste ilmestamiseks ja vormistamiseks. Enamlevinud tarkvarad on liiga universaalsed, piiratud funktsionaalsustega või keerulised kasutada ja kohmakad. Ilma eriväljaõppeta osatakse kasutada universaalsete tarkvarade võimalustest vaid mõni kümnendik. Kõik see viib asjatundja tähelepanu kõrvale tema põhiobjektist ja -ülesandest. Napib ka eksperte, kes lisaks väga heale valdkonna tundmisele, saaksid kasvõi rahuldavalt hakkama kõigi nende geoinformaatiliste tehniliste tugitegevuste valdkondadega (andmed, mudelid, GIS-tarkvara, vormistus).

Sarnaste probleemidega puutume kokku, kui on vaja kaasata laiem otsustajate ring või avalikkus. Lähteandmetele võrreldavusele lisaks hõlpsasti kasutatavad analüüsivahendid aitavad oluliselt lihtsamaks muuta erinevaid ruumis paiknemisvariantide, tekkida võivate olukordade või probleemide läbikaalumist.

Olukorda aitavad leevendada nn temaatilised "tööriistad", mis on kasutajasõbralikud vahendeid, aitamaks toime tulla ruumiandmete suurte mahtudega ja lahenduste kuvamisega. "Tööriistade" valmistamisel on vaja arvesse võtta ka erivaldkondade jaoks oluliste komponentide omavahelisi seosed ja koosmõjusid. Neid vahendeid saab käsitleda kui (territoriaalseid) **ekspertsüsteeme**, millelt oodatakse lihtsat ja operatiivset ümberhäälestamise ja -lülitamise võimlust vastavalt temaatika varieerimisele, lähtetingimuste, andmestike ja sihtgruppide muutmisele.

Koolivõrk so koolid oma astmete ning eri vormidega on ruumiga (territooriumiga) väga tihedalt seotud ja moodustab süsteemse käsitluse järgi suure ja väga keerulise ruumsüsteemi. Igakülgsest käsitlev analüüsi ja lahendusvariante genereeriv ning illustreeriv tarkvara ja andmekäsitlus ei saagi olla lihtne. Teistest riikidest sisseostmine on väheefektiivne, kuna selliste ekspertsüsteemide rakendamisel on antud maa (piirkonna) olude jaoks häälestamine ja tulemuste kultuurikontekstis arusaadavaks tegemine.

Käesolev projekt on ülevaade esimestest sammudest **koolivõrgu ekspertvahendi** loomisest. Töö kulges kahes suunas. Üks keskendas tähelepanu ekspertsüsteemile, mis on ümberhäälestatav erinevatele temaatilistele mudelitele ja tarkvaralahendustele. Teine suund pööras tähelepanu reaalsele andmetele ja nende interpreteerimisele, et valminud tarkvaralahendused annaksid koheselt koolivõrgu üle otsustamiseks kasutatavaid tulemusi.

Kogu tegevus toimus ja tulemus saavutati tihedas koostöös Haridus- ja Teadusministeeriumiga ja Poliitikauuringute Keskusega Praxis. Järgnevalt on kirjeldatud projekti tegevusi, vahendeid ja andmed, mis koos moodustavad ekspertsüsteemi, selle sügavamas tähenduses.

## 1.1. Koostöölepingu objekt

Projekti aluseks on koostööleping nr 10-10/895, mille ülesandeks on geograafilise ekspert-süsteemi "Koolivõrk" loomine algatamine koos reaalsetel andmetel põhinevate näidetega, mille abil saab luua erinevaid teemakaarte leidmaks lahendusvariante üldhariduse koolivõrgu kujundamiseks erinevate kriteeriumite (nt kooli tüüp, kooli teeninduspiirkond, klassi täituvus jm) korral.

Projekti põhitegevused valdkondade kaupa olid:

### Ettevalmistus:

- Tegevuskava koostamine.
- Metoodika koostamine – lähtuvalt Tellija esindajate ootustest, kasutada olevatest andmetest.
- Mudelite ja reeglite kirjeldamine – ekspertsüsteemi ülesande põhimõtete koostamine, soovitava lõpptulemuse kirjeldamine.

### Andmestikud:

- Algandmete süstematiseerimine – andmemudeli koostamine, tabelite vahel sidemete loomine, erinevate klassifikatsioonide välja selgitamine, Tellijalt vajalike andmete saamine.
- Algandmete korrastamine – eri formaatide ühildamine, nõuetele vastavalt struktureerimine.
- Edaspidise andmevajaduse kirjeldamine – milliseid andmeid ja kuidas võiks koguda/kasutada, ettepanekud erinevatele instantsidele.

### Ekspertsüsteem:

- Olemasoleva koolivõrgu ruumipõhine kirjeldamine.
- Eksperti töövahend – MapInfo keskkonnas vahend, kus saab teatud parameetreid muutes genereerida erinevaid teemakaarte.
- Veebirakendus projekti tulemuste (algandmete) presenteerimiseks.

### Aruandlus:

- Vahearuanne projekti käigust 10. september 2004.  
Seoses algandmestike osalise üleandmise viibimisega Tellija poolt Täitjale muudeti 29.11.2004 Lepingu lõpptähtaega.
- Lisandus vahearuanne projekti käigust 10. detsember 2004.
- Projekti lõpptähtaeg 1. veebruar 2005.

## 1.2. Planeeritud tööd ja tegevused

Projekti alustades koostati detailsem tegevuskava, kus toodi välja oodatav tulemus, kasutatavatele andmetele loetelu struktuuride täpsustamine ja tähtajad. Andmestike laekumisega seotud viivituste tõttu tuli ka tähtaegu muuta.

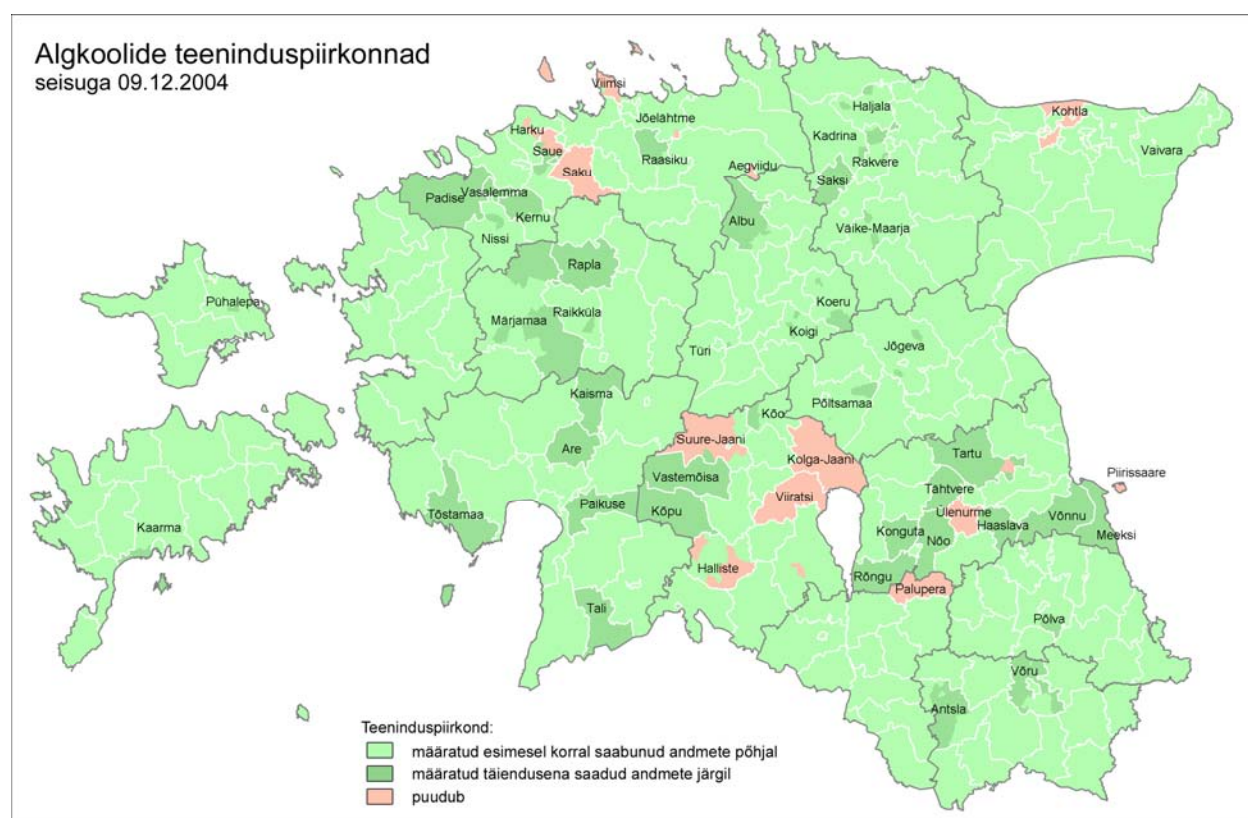
Tööde loetelu	Andmed	Tegevus	Tulemus
1. Andmete ettevalmistamine			
1.1. Andmebaaside struktuuri kirjeldamine		Dokumenteerimine	Teame milliseid andmeid vaja
1.2. Andmete hankimine	Rahvastiku register, rahvaloenduse andmed, õpilaste register, koolide andmed 1999-2003, õpilaste liikumise andmed, õpetajate register	Andmehaldajatega suhtlemine	Vajalikud andmestikud Regios
1.3. Andmete töötlemine		AB loomine	Andmed valmis analüüsiks
2. Olemasoleva olukorra analüüs			
2.1. Koolide ja koolipiirkondade võrgustik tüüpide kaupa	Koolide koolipiirkondade AB	Teemakaartide koostamine	Kaart Eesti/maakondade kaetusest koolivõrguga
2.2. Õpilaste mobiilsus ehk pendelränne	Õpilaste AB	Teemakaartide koostamine	Kaart õpilaste liikumisest elu- ja õppekoha vahel
2.3. Koolipiirkondade iseloomustus – õpilaste arv ja teised kvaliteedi ja kvantiteedi näitajad	Koolipiirkondade ja koolide AB	Teemakaartide koostamine	Kaardid koolipiirkondadest erinevate tunnuste alusel – nt õpilaste arv koolipiirkonnas, klasside mahutavus ja täituvus jms
3. Õpilaste arvu prognoos 2010	Rahvaloenduse AB	AB loomine, teemakaartide koostamine	Õpilaste arv valdades 2010 a, klasside kaupa
4. Ekspertsüsteemi kontseptuaalse mudeli väljatöötamine	Vajadused tulevad projekti ekspertidelt projekti käigus	Vajaduste analüüsimine ja sisendit/väljundite täpsustamine	Ideaalpilt sellest, milliseid analüüse kuidas soovitakse teha. Kirjeldatud andmed, andmetega tehtavad protsessid, tulemuste vormistamise viis jms
5. Ekspertsüsteemi osaline loomine			
5.1. Reeglite loomine	Reeglid, kriteeriumid tulevad projekti ekspertidelt projekti käigus	Ekspertarvamuste kogumine ja programmeerimine	Mudelid, algoritmid, kriteeriumid, millele vastavalt andmeid esitatakse
5.2. Kasutajaliidese loomine		Programmeerimine	Rakendus - eksperdi töövahend
5.3. Infosüsteemi programmeerimine		Programmeerimine	Valmis infosüsteem
5.4. Analüüsikaartide loomine ja vormistamine	Projekti käigus loodud koondandmebaas	Teemakaartide loominekoostamine	Analüüsikaardid
6. Veebirakenduse loomine koolivõrgu andmete presenteerimiseks		Programmeerimine	Veebirakendus kaartide ja tabelandmete kuvamiseks ja pärimiseks
7. Ekspertsüsteemi juurutamise ja edasiarendamise kava		Testimine, tegevusplaani kirjeldamine	Dokument, milles kirjeldatud ekspertsüsteemi kasutamise võimalused
8. Aruandlus			
8.1. VahearuanDED		Hetkeolukorra kirjeldamine koos süsteemi kirjeldusega sh näited teemakaartidest	HTM on aruande kätte saanud.
8.2. Lõpparuanne		Kirjalik kokkuvõte koos GES ja avaliku veebi-lahendusega	Üleandmisakt on HTM poolt allkirjastatud
8.3. Tööseminarid		Arutelud, terminoloogia täpsustamine	

## 2. Andmestikud

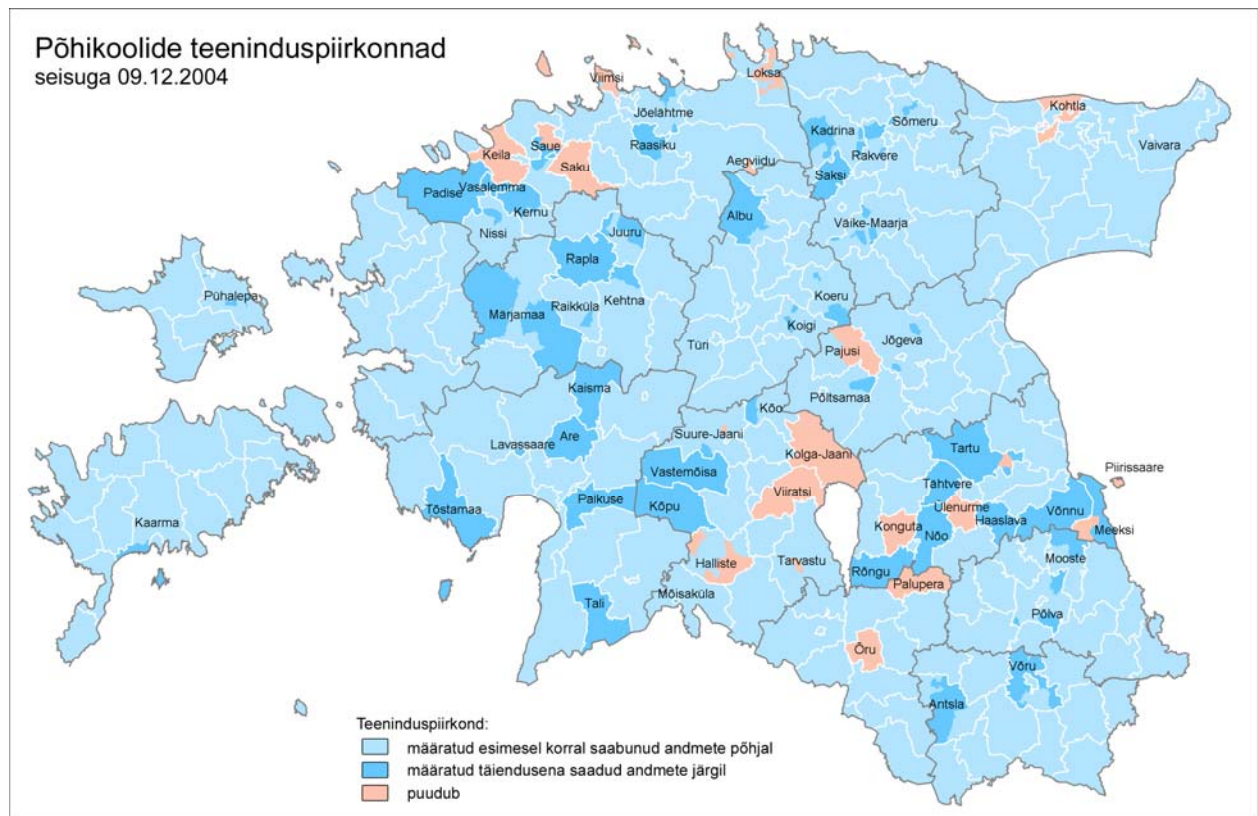
### 2.1. Koolide teeninduspiirkonnad

Teeninduspiirkonnad on esitatud 515 algkooli ja 436 põhikooli osas. Paberkoopiaste ja elektrooniliste dokumentide näol edastatud algmaterjal on digitaliseeritud ja ühendatud teiste andmestikega ühtseks andmebaasiks. Üle on vaadatud andmete sisestamisel tekkinud küsimused ja lõpetatud on ka Tallinna koolide teeninduspiirkondade vektoriseerimine. Viimane osutus päris tömahukaks, kuna andmed olid esitatud koolide kohta tänava ja majanumbrite loeteluna. Alljärgnevalt kahelt jooniselt on näha, millistel aladel on määratud algkooli (joonis 1) ja põhikooli (joonis 2) teeninduspiirkonnad.

Kuna andmete laekumine oli väga visa, siis on teeninduspiirkonnaga kaetud alad jagatud kaheks. Jagamise põhimõtteks on omavalitsusest saadud materjalid teeninduspiirkondade kohta peale esimest päringut ja tumedama värviga andmed, mis esitati peale teistkordset päringut ja puuduolevate alade loendi esitamist omavalitsustele. Vallanimed on kaardile kantud nende valdade kohta, kus valla territooriumi osas teeninduspiirkond osaliselt või teravikuna puudub või on andmed edastatud peale teist päringut.



Joonis 1. Algkoolide teeninduspiirkonnad (andmed seisuga 03.12.2004).



Joonis 2. Põhikoolide teeninduspiirkonnad (andmed seisuga 03.12.2004).

### 2.1.1. Tekkinud probleemid

Koolide teeninduspiirkondade andmeid ei ole esitatud täies mahus. Üheks põhjuseks oli see, et teeninduspiirkondi lihtsalt ei ole omavalitsuste poolt määratud. Oluline on eraldada saadud teeninduspiirkondade andmed selle järgi, kas need on saadud peale esimest või teist päringut. Põhjus on selles, et sageli hakati otsuseid teeninduspiirkondade kohta tegema alles siis, kui seda otse nõuti.

Soovisime koolide teeninduspiirkonnad eraldada kooliastmete kaupa. Siin tekkisid aga mitmed probleemid. Nimelt on enamus andmeid esitatud just algkoolide ja põhikoolide osas, kuna aga algkool võib olla nii 4, 5 või ka 6-klassiline, siis on esiteks väga aeganõudev ja tülikas otsida iga kooli puhul, mis astmeid see sisaldab. Teiseks kerkib küsimus, kuidas täita tabel siis, kui algkoolis on ainult osa teisest kooliastmest. Kolmandaks oleks 1. ja 2. kooliastme teeninduspiirkonnad suures osas kattunud. Neljanda kooliastme puhul ei ole teeninduspiirkondade määramine kohustuslik, mis tõttu laekus ka andmeid nende kohta suhteliselt vähe. Nii kujuneski kokkuleppel Praxis'ega töö käigus välja, et koolide teeninduspiirkonnad oli otstarbekas eraldatud vaid algkooli ja põhikooli tasemel.

Koolide teeninduspiirkondade sisestamine oli väga tömahukas. Enamasti saadi informatsioon paber kandjal või elektroonilises dokumendis lihtsa loeteluna. Teeninduspiirkond oli määratletud külade loeteluna, suuremate koolide puhul hõlmas teeninduspiirkond terve valda. Eriti tömahukas oli Tallinna teeninduspiirkondade vektoriseerimine, kus loetletud olid koolide kaupa tänavad ja majanumbrid, mis antud kooli teeninduspiirkonda kuulusid. Ainsa omavalitsusena õnnestus juba kaardifailina koolide teeninduspiirkonnad saada Tartu linnavalitsuselt.



## 2.2. Koolide register

Koolide 10 aasta andmed on koondatud ühte andmebaasi (Access) tabelisse. Kirjeid on kokku 787 kooli kohta, mida 10 aasta lõikes (1994-2003) kajastub 7007 kirjet. Detailsemad andmed klasside kohta on koondatud teise Accessi tabelisse, kus on kokku 60 445 kirjet. Algsetes lähteandmetes esineb segadusi koolide koodidega, mistõttu on nii koolide kui ka klasside tabelisse juurde tekitatud uus unikaalne identifikaator (kooli\_ID), mis seob omaval klassid vastavate koolidega kindlal aastal. Koolide geokodeerimiseks<sup>1</sup> leiti kõik erinevad koolid ja ühtlustati nende nimekuju, aadress ja kood viimase aasta seisuga. Edasi koolid geokodeeriti ja praeguseks on kõigi nende aadress ja asukoht ükshaaval maja täpsusega üle kontrollitud. Geokodeeriti ka Eestis olevad kutsekoolid 2003. aasta seisuga. Koolide avamiste, sulgemiste ja ühendamiste kindlaks tegemisega tegeles Praxis, kelle poolt kogutud andmed imporditi samuti andmebaasi. Lõpptulemusena on 10 aasta väljavõtte koolide registrist (10 erinevat Exceli faili) üle vaadatud ja vajalikul määral korrastatud ning ühendatud koos õpilaste, õpetajate, teeninduspiirkondade ja haldusjaotuse andmetega ühtsesse relatsiooni- lisse andmebaasi, võimaldades andmete ristkasutust. Andmebaas on omakorda seotud uni- kaalsete identifikaatorite abil MapInfo formaadis kaardipildiga, võimaldades viimasesse tösta andmebaasist erinevate päringute tulemusel saadud informatsiooni ning seda seal teemaki- hina kaardipildis visualiseerida.

### 2.2.1. Andmetega tehtud operatsioonid ja tekkinud probleemid

Esmalt olgu tabelis 1 ära toodud koolide registrist saadud andmetabelite veergude pealkir- jad. Sellisel kujul esitati andmed 10 väljavõttena erinevatest aastatest, mis kajastasid seisu ajavahemikus 1994-2003.

Tabel 1. Koolide registri andmetabelite veergude nimetused.

KOOD	Klassiruumide pindala	klassikomplektide arv (eesti õk)
nimi	SAAL	tüdrukud (vene õk)
ülalpidaja - kood	Saali(de) pindala	poisid (vene õk)
ülalpidaja nimi	UJULA	klassikomplektide arv (vene õk)
omandivorm -kood	internaatide/ õpilaskodude arv	õpilaste arv 1.vahetuses (eesti õk)
omandivorm - nimi	mahutavus	õpilaste arv 1.vahetuses (vene õk)
tüüp - kood	õppurite arv	õpilaste arv 2.vahetuses (eesti õk)
tüüp - nimi	Sööklakohtade arv	õpilaste arv 2.vahetuses (vene õk)
õpetajate kasutuses	Raamatute arv	õpilaste arv 3.vahetuses (eesti õk)
õpilaste kasutuses	Õpikute arv	õpilaste arv 3.vahetuses (vene õk)
üldkasutatavad (raamatukogus, internetipunktis jms)	Katseaia pindala	õpilaste arv kokku
kooli haldamiseks	Klass	jäeti klassikursust kordama
kokku	tüdrukud (eesti õk)	pikapäevarühmas
Klassiruumide arv	poisid (eesti õk)	

Palju segadust tekitasid koolide koodid. Unikaalsed identifikaatorid (edaspidi ID), kas siis tekstilised, numbrilised või kombineeritud, on relatsioonilise andmebaasi pidamise aluseks. Lisaks sellele, et ID oleks igal andmebaasi kirjel unikaalne ühes tabelis, on väga oluline ka selle identifikaatori järjepidevus läbi aja.

Käesoleval juhul kasutati koolide eristamiseks kooli koodi. Teatud juhtudel koolil kood muu- tus ja mis veel hullem, see oli antud teisele koolile. Andmebaasi seisukohalt on see põhi- mõtteline viga ja muudab relatsioonilise andmebaasi tõrgeteta funktsioneerimise võimatuks. Tabelis 2 ongi näha koolid, millel on muutunud kood, huvitaval kombel on see toimunud valdavalt ühel aastal. Olukorra muudab keerulisemaks asjaolu, et osa koodid on omistatud teistele koolidele. See sundis antud koodidesse suhtuma ettevaatusega ja nendega eraldi tegelema. Kokkuvõttes tuli klasside ja koolide tabeli sidumiseks luua uus unikaalne kood, milleks on kooli\_ID (vt täpsemalt peatükist 2.6.). Lisaks tehti ära ka suur töö kooli koodi ühtlustamiseks kõigi 10 aasta lõikes. Selleks võeti aluseks viimase (2003) aasta koodid, kontrolliti, et need ei korduks ja omistati siis need ka eelnevate aastate koolidele tagantjär- gi.

<sup>1</sup> Geokodeerimine on objektile aadressi alusel geograafiliste koordinaatide (st asukoha) määramine. Kasutades kooli aadressi määrati kooli koordinaadid ehk asukoht.

**Tabel 2.** Koolid, millel on muutunud kood.

Kooli nimi	kuni 1996. aastani	alates 1997. aastast
Hageri Lasteaed-Algkool	61002585	61018287
Harmi Põhikool	71005969	71063725
Johannese Kool Rosmal	54008423	54026585
Jõhvi Vene Gümnaasium	13005832	48020710
Kolga Keskkool	71005998	71005969
Kuri Külakool	73002062	80011029
Lehmja Põhikool	71047070	71047040
Lohusalu Lasteaed-Algkool	71047040	99999001
Nõo Põhikool	64000010	64043159
Paldiski Gümnaasium	71006222	71052934
Püüsi Põhikool	71047077	71064050
Vilisuu Erapõhikool	01100471	01386384

Selline kooli koodide ühtlustamine saab toimuda praktiliselt ükshaaval kontrollides, sest ainus kindel vahend andmete muutmiseks tabelis automaatsel teel ongi just seesama kood, mis nüüd ise parandamist nõudis. Seega tuli muuta esialgseid kooli koode 766 juhul. Sellest 742 oli koolide 1994. aasta koodid olid 5 kohalised, erinevalt teistest, mis olid 8 kohalised.

Andmebaasis säilitati ka koolide koodid nende esialgsel kujul, et oleks võimalik tuvastada tehtud muudatused.

Suurt segadust põhjustas ka teistes registrites (õpilaste ja õpetajate register) kooli identifitseerimiseks kasutusel olnud kooli registreerimisnumber. Esines juhtumeid isegi koolide registris kus oli kooli kood ja registreerimisnumber vahetuses. Mõlemad on täpselt ühepikkused (8 kohalised) ja seetõttu väga sarnased numbrilised koodid. Erinevuseks on see, et kui kooli registreerimisnumber algab 7-ga (va paar erandit), siis kooli kood võib alata ükskõik mis numbriga, kaasa arvatud 7. Seega võib need tunnused väga lihtsalt segamini ajada ja lõpuks ei teagi kumb on kumb. Selline viga tuli koolide registris välja kolme kooli osas (vaata tabel 3). Analoogsetest probleemidest üle saamiseks on andmebaasis ka tabel, mis arvatavasti ei ole täielik, kus on omavahel seotud kooli kood ja kooli registreerimisnumber.

**Tabel 3.** Koolid, kus koolide registris oli koodi asemel registreerimisnumber.

Kooli nimi	kood	registreerimisnumber
Haanja-Ruusmäe Põhikool	38001663	75013078
Valga Põhikool	33003841	75009786
Kunda Ühisgümnaasium	59005912	75027695

Probleeme tekitas kooli kood veel ka seetõttu, et tihti algas ta numbriga 0 (null). Nii oli see 787 koolist 108 puhul. Probleem on lihtsalt selles, et kui hoida koodi numbrilisel kujul, siis kaotavad teatud programmid selle 0 eest ära. Sama võib juhtuda ka andmete tõstmisel ühest programmist teise. Seetõttu tuleb sellist numbrilist koodi hoida hoopiski tekstilisel kujul. Tekstilisel kujul hoidmiseks on aga vaja tunduvalt rohkem mälu, mis suurte andmebaaside korral muutub väga keeruliseks faktoriks. Suureneb andmebaasi maht, see aeglustab aga töötamist ning erinevate otsingute tegemist.

Järgmine probleem oleks aga üldse olemata, kui igal koolil oleks olnud kasutusel oma ID ja andmed esitatud sellega seostatuna. Nimelt on vaja koolide sidumiseks asukohaga (geokodeerimiseks) nende aadresse. Aadressid esitati aga aastate kaupa eraldi Exceli töölehtedel. Faili sisuks oli kooli nimi, postii indeks ja aadress. Kahjuks ei olnud sinna juurde pandud ei kooli koodi ega ka registreerimisnumbrit, mis oleks lihtsustanud aadresside sidumist koolide andmetega. Seega tuli kasutada aadresside sidumiseks ka kooli nimesid. Olu korra tegi keeruliseks veel asjaolu, et peale selle, et kooli nimi ise on 10 aasta jooksul palju muutunud, esineb ka sama nimega koole ning ühe ja sama kooli nimi kirjutatud erinevalt. Kokkuvõttes tuli siin palju erandeid, mis tuli ükshaaval üle kontrollida.

Pärast aadresside ja koolidega sidumist selgus, et ka aadressid ei ole standardsed. Erinevatel aastatel on aadressi kirjalpilt antud erinevalt, kuigi tegelik asukoht ei ole muutunud. Nii

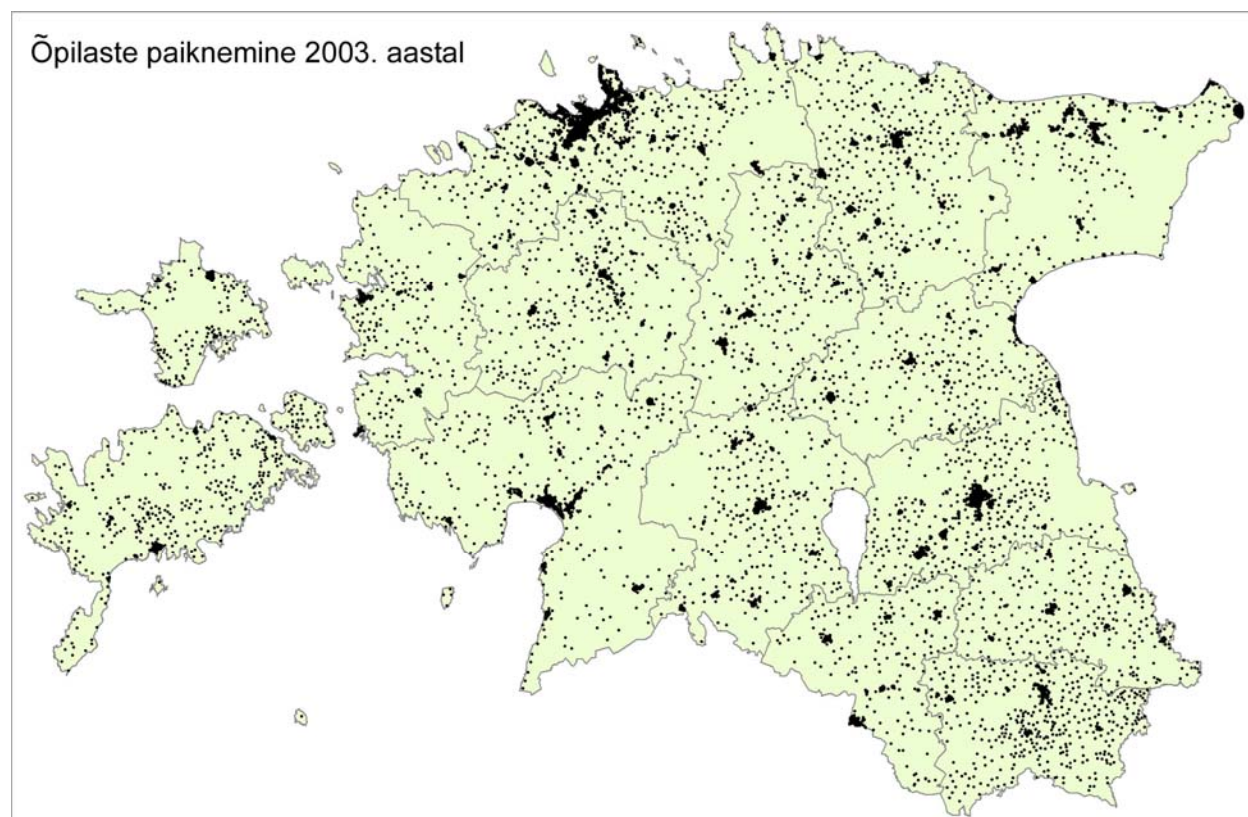
tuli ka aadressid enne geokodeerimist korrastada.

Kõiki aadresse geokodeerimiseks korrektsele kujule viia oleks olnud liialt töömahukas, seega leiti koolidele aadress viimase seisuga järgi, mis andmetabelites kajastus. Sama tehti ka kooli nimega. Peamiseks tunnuseks, mille alusel koolid läbi 10 aasta tuvastati, oli kooli kood, selle puhul pidi arvestama eelnevalt kirjeldatud probleeme. Siin tuleb olulise eelisena välja relatsioonilise andmebaasi eelis lihtsate tabelandmete ees, kus ei ole vaja kooli aadressi säilitada kõigi aastate kaupa, vaid piisab ühest aadressist, mis seotakse kooliga ID abil. Aadressi muutuse võib salvestada arhiiveeritud seisundina teises tabelis.

Pärast koolide geokodeerimist omistati enamusele koolidest asukohta näitav x- ja y-koordinaat juba Regios olnud andmete põhjal. Andmete kontrollimiseks ja uute aadressidega sidumiseks, kontrolliti kõikide koolide asukohad uuesti hoone täpsusega üle. Maapiirkondades ei ole sageli kasutusel tänavaid ja majanumbreid (tihti on täpsemaks viiteks küla nimi), siis kasutati asukoha täpsustamiseks erinevaid aluskaarte, vajadusel ka helistamist koolidesse.

### 2.3. Õpilaste- ja rahvastikuregister

2003. aasta õpilaste registri andmed on koondatud ühtsesse andmebaasi ja juurde on tekitatud igale õpilasele klassi\_ID, mis seob õpilased andmebaasis oleva klasside tabeliga ja seeläbi ka kooliga. Selle saavutamiseks asendati õpilaste registris olev kooli registreerimisnumber kooli koodiga. Õpilaste eristamiseks on tekitatud juurde ka teine unikaalne kood opil\_ID (õpilase ID). 2003. aasta õpilased on rahvastikuregistrist saadud aadressi järgi geokodeeritud ja MapInfos on tekitatud ka õpilaste kaardikiht. See kõik annab võimaluse siduda õpilased koolidega ja uurida õpilase ja kooli vahelisi ruumilisi seoseid, eelkõige õpilaste rännet.



Joonis 3. Õpilaste paiknemine 2003. aastal (rahvastikuregistri elukoha järgi).

### 2.3.1. Andmetega tehtud operatsioonid ja tekkinud probleemid

Tabelis 4 on ära toodud õpilaste registri ja rahvastikuregistri ühendamisel saadud andmetabeli veerud esialgsel kujul.

**Tabel 4.** Õpilaste- ja rahvastikuregistri andmetabelite veergude nimetused.

Aasta	synnikuu	synniaasta	emakeel	oppekeel	klass	oppeasutus	kaugus	elukoht
Maja	tänav	Asula/küla/linnaosa		Kylakd	Vald/linn	KD	Maakond	MK

Esimese operatsioonina tekitati igale õpilase ID, kuna see alguses tabelis puudus. Edasi tõsteti Exceli failis 3-l erineval töölehel olev õpilaste register koos rahvastikuregistri aadressiga Accessi tabelisse. Järgmise sammuna asendati klassi veerus olevad kirjed G1, G2 ja G3 vastavalt kirjetega 10, 11 ja 12. Seejärel otsiti koolide registreerimisnumbri juurde ka vastav kooli kood ning registreerimisnumbrid eemaldati, kuna viimane on olemas kooli koodi ja registreerimisnumbrit siduvas tabelis. Õpilaste sidumiseks klasside ja seeläbi ka koolidega, tekitati juurde klassi\_ID. Selleks liideti üheks stringiks järgmised tunnused: õppeaasta, kooli kood ja klassi number. Kooli koodide leidmist ei saanud teha automaatselt, kuna koodid võisid olla muutunud üle läinud teisele koolile. Seetõttu leiti kõigepealt igale õpilasele kooli kood kolme erineva tabeli järgi, kui need omavahel klappisid, siis omistati see ka vastavale registreerimisnumbrile koodiks. Sellisel moel jäi kood leidmata 56-le registreerimisnumbrile. Edasise päringuga tehti kindlaks, et 26 kooli puhul, millele õpilaste register viitas, oli registreerimisnumber segamini aetud koodiga. Ülejäänud 30 registreerimisnumbri puhul leiti vastav kood nende üksikhaaval üle kontrollimisel.

Kontrollides, kas kõikidele õpilastele on koolide registrist leitav ka kool (sidumiseks kasutatakse klassi\_ID'd ja kooli\_ID'd), olid tulemused järgmised. Õpilastele leiti kooli vaste **187 189** juhul, ehk ilma koolita jääb **5336** õpilast, mis on 2,8% kogu õpilaste arvust (õpilaste arv 2003. aasta õpilaste registris 192 525). 5336-st õpilasest 4226 õpib koolides, mida koolide registris toodud ei ole (vaata tabel 5). Tegemist on täiskasvanute- või erikoolidega, mis jäävad uurimusest välja ning seega endast probleemi ei kujuta.

**Tabel 5.** Koolide registrist puuduvad koolid, millele viidatakse õpilaste registrist.

Kooli kood	õpilasi	kooli nimi
01050000	984	Tallinna Täiskasvanute Gümnaasium
01053702	466	Tallinna Vanalinna Täiskasvanute Gümnaasium
13005855	252	Kohtla-Järve Õhtukeskkool
15008757	227	Narva Õhtukeskkool
18002407	380	Pärnu Täiskasvanute Gümnaasium
24007314	579	Tartu Täiskasvanute Gümnaasium
35006105	56	Viljandi Erikutsekool
38021393	230	Võru Täiskasvanute Gümnaasium
41023378	139	Jõgeva Täiskasvanute Keskkool
45003444	224	Kuressaare Õhtukeskkool
51000830	180	Paide Täiskasvanute Gümnaasium
61007447	219	Rapla Õhtukeskkool
69001852	149	Haapsalu Täiskasvanute Gümnaasium
71014900	141	Rummu Erikutsekool
<b>Kokku</b>	<b>4226</b>	<b>2,2% kõigist õpilastest</b>

Järele jääb veel 1110 õpilast, keda ei õnnestu kooliga siduda. Osalt seetõttu, et õpilastele on märgitud klass, mida koolide registri andmetel selles koolis 2003. aastal ei ole. Kolmel juhul ilmnes, et kool oli 2003. aastaks suletud (täpsemalt vaata tabel 6). Ilma koolide või õpilaste registrit täiendamata neid õpilasi koolidega siduda ei saa, nii et esialgu jäävadki need õpilased uurimusest välja. Täiendamine tähendaks, et osa koolide andmed tuleks koolide registris üle kontrollida, sest seal ei ole kirjas klasse, millele viidatakse õpilaste registris.

**Tabel 6.** Koolide registri ja õpilaste registri (2003 a. seis) vahelised ebakõlad.

Kooli kood	Õpilaste arv	Kooli nimi	Võrreldes õpilaste ja koolide registrit
01251508	450	Vana-Kalamaja Gümnaasium	Koolide registrist puudu klassid 10, 11 ja 12
35002461	192	Viljandi Täiskasvanute Gümnaasium	Koolide registrist puudu klassid 8, 9, 11 ja 12
01053688	146	Mustamäe Üldhariduskool	Koolide registrist puudu klassid 10, 11 ja 12
59004826	122	Rakvere Õhtukeskkool	Koolide registrist puudu klassid 10, 11 ja 12
33003686	84	Valga Kaugõppegümnaasium	Koolide registrist puudu klassid 10, 11 ja 12
01101884	22	Akadeemia Nord õigusgümnaasium	Suletud 2003ndal aastal
24007308	15	Tartu Kroonuaia Kool	Koolide registrist puudu 10nes klass
35012137	14	Saarepeedi Põhikool	Õpilaste registris märgitud 0 klass
01204667	9	Erakommertsgümnaasium	Koolide registrist puudu 11nes klass
56003708	9	Tõstamaa Keskkool	Õpilaste registris märgitud 13nes klass
59006363	9	Rakvere Vene Gümnaasium	Koolide registrist puudu 12nes klass
01345267	7	Tallinna Kurtide Kool	Koolide registrist puudu 12nes klass, aga on 13nes
35009836	7	Viljandi Vene Gümnaasium	Koolide registrist puudu 11nes klass
13005861	5	Ahtme Kool	Õpilaste registris märgitud 0 klass
18009999	5	Pärnu Toimetulekukool	Koolide registrist puudu 4jas klass
69003928	5	Haapsalu Ühisgümnaasium	Õpilaste registris märgitud 0 klass
01309923	3	Fouette Balletikool	Suletud 2003ndal aastal
33106357	2	Ritsu Lasteaed-Algkool	Koolide registrist puudu 5es klass
56003602	2	Massiaru Algkool	Koolide registrist puudu 6es klass
24007293	1	Tartu Puistee Kool	Suletud alates 2002st aastast
41001365	1	Tiheda Algkool	Koolide registrist puudu 1ne klass
<b>Kokku</b>	<b>1110</b>		

Järgmiseks probleemseks kohaks oli õpilaste elukoha sidumine kaardiga. Teatud informatsiooni andis rahvastikuregistri ja õpilaste registri elukohtade võrdlemine. Võrdluse teeb võimalikuks omavalitsuse kood, mis on nii õpilaste kui ka rahvastikuregistri andmetes olemas. Samas on nende andmete tekkelugu kindlasti erinev. Seega on järgnev statistika lihtsalt huvitav võrdlus. Nimelt erinevusi, kus rahvastikuregistri elukoht ei lange kokku õpilaste registris kajastuva elukohaga 12 304, mis on õpilaste arvust 6,4%. Järgnevas tabelis on suuremad erinevused ära toodud.

**Tabel 7.** Rahvastikuregistri ja õpilaste registri vahelised elukoha erinevused.

Vald/linn EHAK (rahvastikuregister)	Õpilasi	Vald / linn
Erinevad Tallinna linnaosade koodid (176, 298, 339, 387, 482, 524, 596, 614)	1474	Tallinn
795	1096	Tartu linn
625	319	Pärnu linn
727	307	Saue vald
728	211	Saue linn
Erinevad Kohtla-Järve linnaosade koodid (120, 265, 340, 553, 747, 893)	206	Kohtla-Järve
890	203	Viimsi vald
446	190	Maardu linn
198	185	Harku vald
653	160	Rae vald
296	157	Keila linn
897	157	Viljandi linn
...	...	...
<b>Kokku</b>	<b>12304</b>	

Tabeli 8 annab ülevaate õpilaste geokodeerimise täpsusest. Arvestatud on 2003. aasta õpilasi, keda registri järgi oli 192 525. Liites kokku koolide registri andmetes oleva õpilaste arvu, saame 43 õpilast rohkem (192 568 õpilast). Õpilaste registri ja koolide registri erinevus õpilaste arvu osas on toodud tabelis 9. Erinevused tulevad sellest, et osa koole, mis kajastuvad õpilaste registris ei ole kirjas koolide registrist saadud andmetes. Osaliselt on tege mist erikoolidega, mis antud uuringus kajastuma ei peagi. Samas näitab see fakt, et registre eesmärk ja pidamine ei ole päris ühesugune ning kooskõlas. Erinevused tulevad võib olla ka registre e erinevate haldajate ja täitjate tõttu. Geokodeerimise tulemusena saadud

õpilaste paiknemine on toodud joonisel 3 oleval kaardil.

**Tabel 8.** 2003. a. õpilaste nimekirjade geokodeerimise täpsus.

Õpilaste arv	Protsent	Geokodeerimise täpsusklass
137296	71,31%	Hoone
1125	0,58%	Lähim hoone
34	0,02%	Tänav
331	0,17%	Lähim tänav
50171	26,06%	Asula
3568	1,85%	Vigane – geokodeerimata
192525	100,00%	

**Tabel 9.** Õpilaste registri ja koolide registri vaheline õpilaste arvu võrdlus.

	2001	2002	2003
Õpilaste arv õpilaste registri järgi	191602	199122	192525
Õpilaste arv koolide registri järgi	207612	200478	192568
erinevus	<b>16010</b>	1356	43

Geokodeerimise täpsusklass, näitab et 71,3% õpilastest on geokodeeritud hoone täpsusega, suhteliselt väike osa lähima hoone ja tänava täpsusega. Natuke üle veerandi on geokodeeritud asula täpsusega, see tähendab, et õpilase elukohaks on võetud asula tsentroidi koordinaadid. Enamuse juhudest, hinnanguliselt üle 90%, tähendab see seda, et tegemist on küla (haldusareali) tsentroidiga. Teades, et koolide teeninduspiirkonnad ühelgi juhul ei poolita küla, võib seda lugeda täiesti piisavaks täpsuseks. Kaardipildis tähendab see, et suuremate külade puhul võib küla tsentroidi kohal olla ühe täpi asemel mitu kohakuti asetsevat täppi. Olulist rolli hakkab mängima õpilase ja küla tsentroidi võrdsustamine siis, kui on tarvis leida näiteks õpilaste kaugust koolist või kui tegemist on suhteliselt suure külaga, kus õpilase tegelik elukoht võib küla tsentroidist päris kaugel asuda. Täpsemat geokodeerimist praegusel juhul otstarbekas teha, kuigi maapiirkondades võiks õpilaste geokodeerimisel arvestada nende tegelikku elukohta, mitte kasutada selleks küla tsentroidi<sup>2</sup>.

Geokodeerimata jäänud 1,85% õpilastest (so 3568) on need, kelle kohta rahvastikuregistris puudusid elukoha andmed ja ka need, kelle kohta andmed olid esitatud ainult valla tasemel (186 õpilast). 119 õpilast polnud võimalik ei geokodeerida (kokku geokodeerimata 3568 õpilast) ega seostada ühegi kooliga (kokku seostamata 5336 õpilast).

Võrreldes omavahel õpilaste arvu koolide registri ja õpilaste registri järgi, siis 576st 2003.a. tegutsevast üldhariduskoolist on õpilaste arvus erinevus 320-nel koolil. 67 juhul on õpilaste arv suurem õpilaste registris ja ülejäänud 253 juhul on õpilaste arv suurem koolide registri järgi. 13-nel juhul on erinevus 100 või rohkem õpilast (vt tabel 10).

**Tabel 10.** Suuremad erinevused õpilaste ja koolide registri õpilaste arvudes kooli kohta.

Kooli kood	Kooli nimi	õpilaste register	koolide register	erinevus
15008835	Narva 6. Kool	45	735	690
45003415	Kuressaare Vanalinna Kool	1	503	502
01251508	Vana-Kalamaja Gümnaasium	648	149	-499
41001253	Mustvee Vene Gümnaasium	1	295	294
15008792	Narva Kesklinna Gümnaasium	899	1139	240
10306970	Audentese Erakool	824	651	-173
64003013	Elva Gümnaasium	1236	1067	-169
24007231	Tartu Kunstigümnaasium	832	996	164
71005892	Keila Gümnaasium	1072	928	-144
01043661	Maardu Gümnaasium	1233	1090	-143
01053688	Mustamäe Üldhariduskool	502	377	-125
01053547	Gustav Adolfi Gümnaasium	906	1006	100
01043572	Lasnamäe Gümnaasium	1211	1311	100

<sup>2</sup> Tsentroidi võiks sellisel juhul panna küla asustuse keskpunkti, mitte küla geomeetrilisse keskmesse

### 2.3.2. Rahvastikuprognosis

Projekti käigus sai selgeks, et rahvastikuprognosisi maakondade tasemel Eestis ei tehta. Probleemiks on maakondadevahelise rände arvestamine, selle mahu kindlakstegemine on keeruline ja tulemus oleks liialt ebatäpne. Seega puudub võimalus rahvastikuprognosisi kaasamiseks ekspertsüsteemi töös. Mõningat leevendust pakub rahvaloenduse andmetel baseeruv õpilaste arv aastaste vanusgruppide lõikes, mida saab interpoleerida tulevikku. Nendest andmetest kirjeldatakse lähemalt peatükis 2.4. 2000.a. rahvaloenduse andmebaas.

Nagu eelpool mainitud, on peamiseks probleemiks rahvastikuprognosiside tegemisel just maakondadevahelise rände arvestamine. Üheks põnevaks tulemuseks on käesoleva projekti raames õpilaste rännet kajastavad teemakaardid. Nendel kujutatud rändesuunad ja hulgad on saadud kolme erineva registri ühendamisel ja seal sisalduva ruumiinfo geokodeerimisel. Teisi sõnu rahvastikuregister õpilase elukoha andmetega, õpilaste register, mis viib kokku õpilase ja kooli ning koolide register koolide asukohaga. Tulemus on näha paari näitena teemakaartide osas (joonis 9 ja 10) ja ülejäänud kaardid leiab koolide veebirakendusest.

### 2.4. 2000.a. rahvaloenduse andmebaas

Algandmestike loetellu lisati töö käigus Statistikaameti poolt 2000.a. rahva ja eluruumide loenduse käigus kogutud rahvastiku vanuselise jaotuse andmestik.

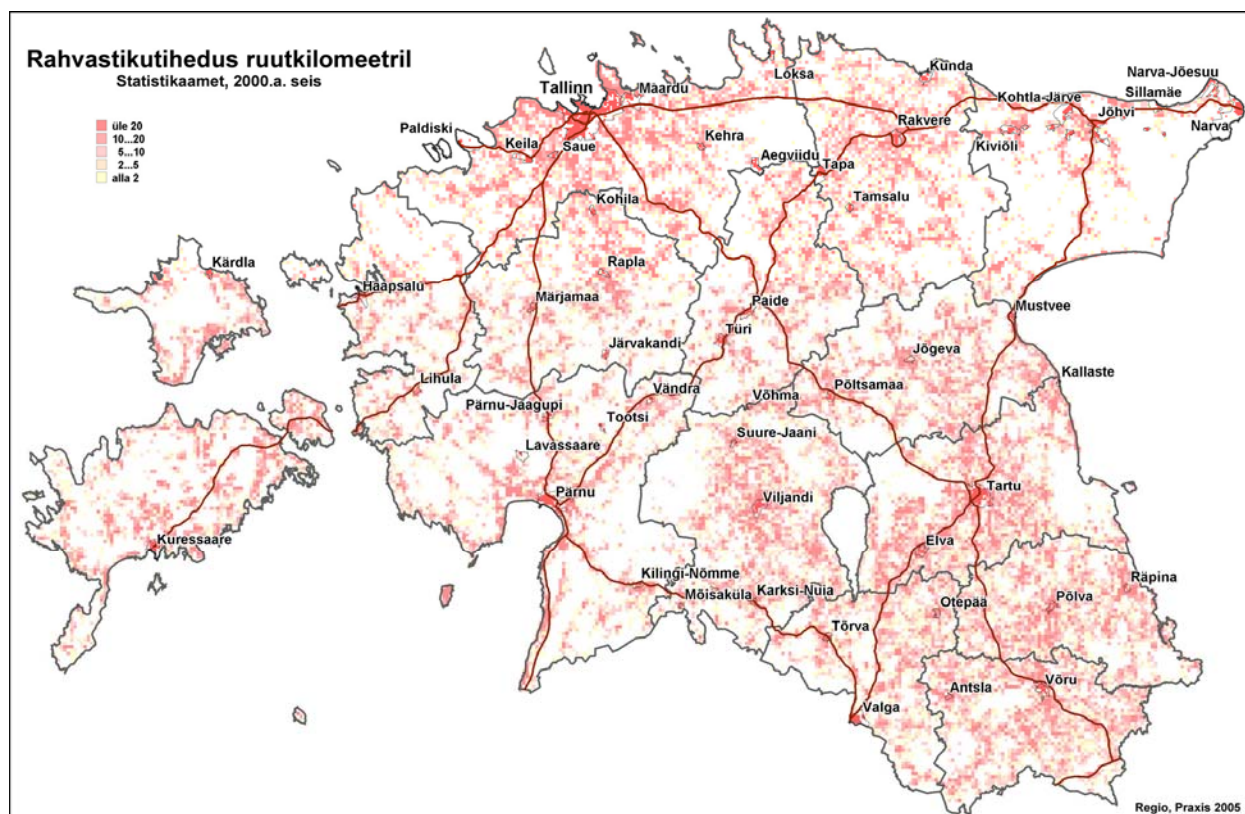
Andmed koostati ruudustikupõhiselt järgmise struktuuriga:

- Kogurahvastiku asustustihedus (elanikku km<sup>2</sup> kohta).
- Rahvastiku ühe-aastane vanuseline struktuur, vanuse vahemikus 0 kuni 24 aastat.
- Andmed vormistada kogu Eestimaa ulatuses 1x1 km ruutude kaupa, sh linnad ja alevid 0,5x0,5 km ruutude kaupa.
- Andmete formaat: ASCII-fail (comma separated - CSV).

NB! Statistikaameti poolt Andmebaasist RL2000 tehtud ruudustikupõhised andmed kajastavad tavaleibkondades ja tavaeluruumides elavaid isikuid.

Andmebaasi kohta täiendav info <http://www.stat.ee/129014> ja eraldi rahvastiku andmestikust <http://www.stat.ee/files/eva2003/RV200101.pdf>.

Rahvastiku andmeid kasutati taustakaardina koolivõrgu paiknemise analüüsil, mis annab täiendava infovoo koolide praegusele paiknemisele ning paiknemise võimalikele arengusuundadele tulevikus. Esialgses formaadis esitatud andmed seoti MapInfo kaardifailidega, mis võimaldavad kaardipildina kujutada näiteks teatud vanusgruppide kaupa rahvastiku tihedust ja näha nende diferentseeritust Eesti eri piirkondades. Joonisel 4 on näha rahvastiku tihedus 2000. aastal.



Joonis 4. Rahvastiku tihedus 2000. a. rahvaloenduse andmetel.

## 2.5. Pedagoogide register

Pedagoogide register ei olnud esialgse plaani kohaselt käesolevasse töösse kaasatud. Kuid mitmed õpetajatega seotud näitajad (mees ja naisõpetajate suhtarv, kõrgharidusega õpetajate osakaal, õpilaste arv õpetaja kohta jms) pakkusid huvi koolide hindamiskriteeriumite vallas, siis otsustati ka pedagoogide registrit kasutada. Saime 18. jaanuaril 2005 Praxise kaudu väljavõtte pedagoogide registrist, mis osutus poolikuks, sest paljud kirjed jäid mingil põhjusel edastamata. 20. jaanuaril saadud uus väljavõte sisaldas palju enam kirjeid, kokku 31 994. Pedagoogide registris sisaldunud andmeid kajastab järgnev tabel 10, kus on toodud veergude pealkirjad.

Tabel 11. Pedagoogide registri andmetabelite veergude nimetused.

kood	dbo_class_step_name	maakond	ametikoha suurus	sugu	tase_reg	õppeaine
erialane haridus	vald	õppeasutus	haridustase	pedagoogiline	asutuse kood	

Andmeväli "kood" sisaldas endas õpetajat identifitseerivat tunnust, mis võeti edaspidi õpetajate arvu leidmise aluseks. Eeldati, et iga kood tähendab erinevat õpetajat. Meid huvitasid pedagoogide registri puhul eelkõige järgmised andmed: õpetajate arv koolis, kõrgharidusega õpetajate osakaal, õpetajate töökoormus, mees- ja naisõpetajate suhtarv ning õpetajate keskmine vanus koolis.

Oluline on siinkohal märkida, et pedagoogide registri väljavõttest puudusid täielikult Tallinna koolide õpetajad. Tegelikult puudub praegusel juhul üldse võimalus kontrollida, kuivõrd on õpetajaid sellest väljavõttest puudu, sest mingit võrdlusmomenti teise andmestikuga ei ole.

Õpetajate arvu leidmiseks on kaks viisi: võib leida õpetajate arvu neile antud ID järgi, mis andis tulemuseks 11467 õpetajat 508 koolis. Teine võimalus leida õpetajate arv koolide kaupa, st summeerida 1 õpetaja kaks korda, kui ta töötab kahes koolis korraga. Sellisel moel saame õpetajate arvuks 11 927. Põhimõtteliselt on võimalik arvutada ka täiskoormu-



sega õpetajate arv ehk summeerida lihtsalt töökoormused, nii oleks tulemuseks 10 246,6 õpetajat.

Lisaks sellele, et pedagoogide registrist on puudu Tallinna õpetajate andmed, esineb seal muidki puudujääke. Näiteks ei ole määratud haridustaset märgitud 317 pedagoogi puhul (13 õpetaja puhul on selleks väärtuseks 0) ja pedagoogiline haridus 1673 õpetajal. Selgusetuks jääb, kas tegemist on lihtsalt puudulike andmetega või nendel õpetajatel vastav haridus puudub.

Pedagoogide registrist tehtud väljavõtte sidumisel koolide registriga tuli välja, et puudu on õpetajate andmed 125 2003. aastal töötanud kooli osas. Kui siit välja jätta vangla- ja erikoolid ning koolid tüüpidega 160, 162 ja 180<sup>3</sup> ja Tallinna koolid (85 kooli, mille kohta on andmete puudumine teada), siis jääb ilma õpetajateta veel 24 kooli.

## 2.6. Projekti tulemusena loodud andmebaas

Projekti käigus saadud erinevatest registritest (õpilaste register, rahvastikuregister, pedagoogide register, koolide register) saadud andmed koondati ühte relatsioonilisse andmebaasi võimaldades nimetatud registrite riskasutuse. Loodud andmebaasi struktuur on toodud joonisel 5.

Relatsioonilise andmebaasi puhul on üheks oluliseks aspektiks andmete liiasuse vältimine. Näiteks, kui koolide tabelis on 7007 kirjet, igaühe juures ka aadress, kuigi erinevaid koole on 747 ja aadresse umbes 800 ringis, siis on suur vahe, kas hoida andmebaasis 7007 aadressi kirjet või 800. Ühelt poolt on see oluline andmebaasi mahu vähendamiseks ja teiselt poolt erinevate vormiliste- ja kirjavigade ärahoidmiseks. Näiteks praeguses andmebaasis on kokku 2364 erineval kujul aadresside kirjet, mis väga paljudel juhtudel tähendavad ühte ja sedasama aadressi. Vaata näiteid tabelist 12.

**Tabel 12.** Näiteid erinevatest andmebaasis esinevatest aadressi kirjetest.

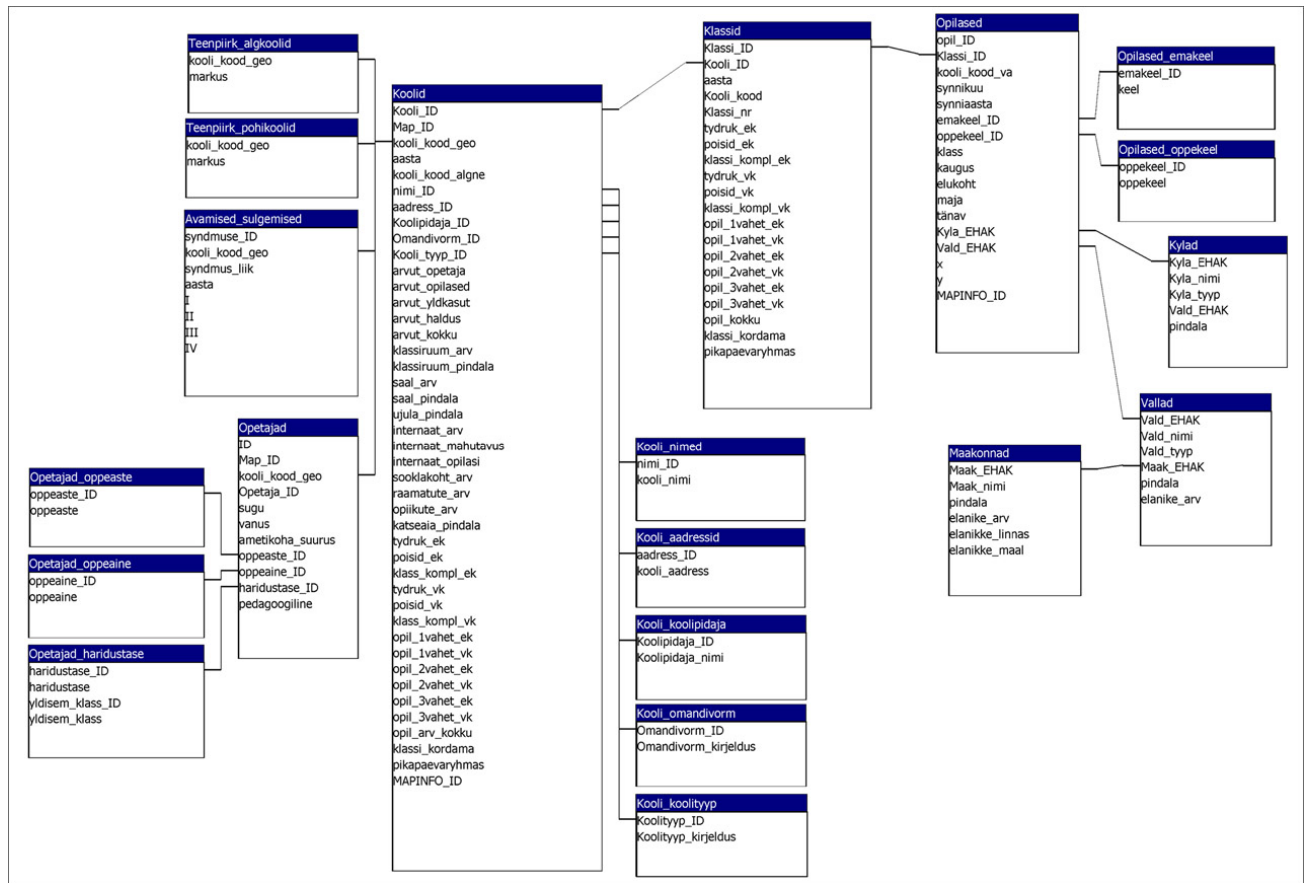
aadress	aadress
Ida-Viru maakond, Sillamäe, Tskalovi 25	Harju maakond, Saue vald, Ääsmäe, Kasesalu 16
Ida-Viru maakond, Sillamäe, Tskalovi 6/25	Harjumaa, Saue vald, Ääsmäe, Kasesalu 16
Ida-Viru maakond, Sillamäe, TSKALOVI TN. 25	Harjumaa, Saue vald, Ääsmäe, Kasesalu 10
Ida-Virumaa, Sillamäe, Tskalovi 25	Harjumaa, Saue vald, Ääsmäe, Kasesalu 10a
Ida-Virumaa, Sillamäe, TSKALOVI TN. 25	Harjumaa, Saue vald, Ääsmäe, Kasesalu 10a

Sama loogika kehtib ka kooli nimede puhul. Ei ole vaja hoida kooli nimesid andmebaasis 7007 korda, kui piisab ainult erinevate nimekujude säilitamisest, mis siis seotakse kooliga vastava ID abil.

Samadel põhjustel on koolide tabeliga ID abil seotud ka omandivormi, koolipidajat ja koolitüüpi näitavad tabelid. Õpilaste tabeliga on seotud emakeelt ja õppekeelt näitavad tabelid. Õpetajate tabeliga on ID abil ühendatud õpetajate haridustase, nende poolt õpetatav aine ja kooliaste.

Suhteliselt kõige rohkem tuli vaeva näha kooli ülalpidajate tabeliga, sest esines hästi palju samu ülalpidaja koode, kus ülalpidaja nimi oli erinev. Enamasti oli see tingitud igasugustest väikestest kirjapildi erinevustest (näit. Võru Maavalitsus ja võru Maakonnavalitsus või siis Kolga-Jaani vallavalitsus ja Kolga-Jaani vallavalitsus), mis sisuliselt tähendasid sama. Palju esines ka olukordi, kus üks ülalpidaja oli andmetabelis mitme erineva koodiga. Sellisel juhul jäeti tabelisse alles uuema seisuga kood.

<sup>3</sup> öhtukoolid, kaugõppekoolid ja koolid, mille juures puuduvad üldhariduskassid



Joonis 5. Andmebaasi struktuur.

Et andmebaasi süsteemselt ja analüüsiprogrammidega tööle rakendada, tekitati andmebaasis juurde mitmeid erinevaid identifikaatoreid (ID). Siinkohal on välja toodud olulisemad nendest ning lisatud ka mõne muu olulisema lihtsa andmevälja kirjeldus:

- **kooli\_kood\_algne** – esialgne kooli kood, mis saadi esialgsetest andmetest. Igaks juhuks on see säilitatud, et saaks võrrelda koole esialgsete andmetega;
- **kooli\_kood\_geo** – unikaalseks tehtud kood iga kooli osas. On võetud kooli kood viimase aasta seisuga ja juhul kui mingil varasemal aastal on olnud teistsugune kood, siis on see asendatud viimase aasta koodiga;
- **Map\_ID** – sisuliselt kopeerib eelnevat (**kooli\_kood\_geo**). On vajalik siis, kui tahetakse siduda kolme või enam tabelit, kuid siduvaks ID'ks on **kooli\_kood\_geo**, siis vajadusel saab selle asendada **Map\_ID**'ga;
- **Kooli\_ID** – unikaalne kooli kood, erinevalt **kooli\_kood\_geo**'st on see unikaalne ka igal aastal. See on saadud sel moel, et aastanumbrile on juurde lisatud kooli algne kood;
- **Klassi\_ID** – unikaalne kood iga klassi, kooli ja aasta lõikes. On saadud nii, et **kooli\_ID**'le on juurde lisatud klassi number;
- **Opil\_ID** – unikaalne kood iga õpilase kohta, on saadud lihtsalt õpilasi järjest numbrdades. Alates 1st kuni 192525-ni.

Sellise relatsioonilise andmebaasi eelised lihtsate tabelandmete ees on andmete optimeerimine. Välditakse samade andmete kordamist erinevates tabelites, mis hoiab kokku andmebaasi mahtu ja võimaldab vabaneda erinevatest vormilistest vigadest andmete hoidmisel. Tüüpilisem näide ongi erinevad aadressid ja nimed. Relatsioonilise andmebaasi korral säilitatakse aadresse ühekordse kirjena, kui kuskil aadressis ka esineb viga, siis piisab selle parandamisest ühes kohas.

Samas eeldab selline andmebaas väga korrektset andmete pidamist, sest kui aadress on

tabelis vale, siis on see vale kõigil kirjetel, mis sellele ID abil viitavad. Enam ei piirduks viga ühe kirjega nagu see oleks tavalise lameda andmetabeli korral. Täpselt peab olema paigas ka erinevate ID'de pidamine, vead ID's (nagu antud juhul oli kooli koodidega) ei võimalda sellise andmebaasi pidamist.

Eelnevas lõigus kirjeldatud aspekte ei saa aga kuidagi pidada relatsioonilise andmebaasi negatiivseteks külgedeks võrreldes lameda andmebaasiga. Otse vastupidi, see peaks olema iga registripidaja esmane huvi ja eesmärk, et tema andmed on korras. Relatsioonilises andmebaasis tulevad sellised vead palju lihtsamalt välja kui tavalistes andmetabelites.

Selline relatsiooniline andmebaas, kus on seotud erinevad registrid, on väga võimas tööriist. Antud juhul on siin juures ka ruumiinfo, mis geokodeerituna (ruumiga seotuna) on aluseks geoinfosüsteemi ja ruumilise (geograafilise) ekspertsüsteemi loomiseks. Praeguselt tasandilt, kus on läbi tehtud andmete kogumise ja korrastamise etapp, on suurepärane võimaluse minna edasi järgmisele tasandile – ekspertsüsteemi loomisele ja arendamisele. Praeguse seisuga ei ole saadud andmed päris täielikud ja registrites esineb mitmeid ebatäpsusi, kuid andmekogumist jätkates ja puudusi likvideerides võime jõuda väga hea lähtekohani ekspertsüsteemi tööks. Üheks lihtsaks näiteks antud andmebaasi poolt pakutavast lisandväärtusest tabelandmetele ongi näiteks maakondade tasemel õpilaste rändekaardid, mis annavad tabelandmetele ruumilise dimensiooni. Väga olulise mõõtmeks ajadimensiooni lihsamine. Muutuste jälgimine ajas annaks väga palju informatsiooni erinevate otsuste tegemiseks koolivõrgu korraldamisel. Jõuame jällegi välja ekspertsüsteemi pakutavate võimaluste juurde. Ekspertsüsteemi võimalusi ja võimalikke arengusuundi kirjeldab täpsemalt peatükk 5.

### **2.6.1. Andmebaasi uuendamine**

Loodud andmebaasi uuendamiseks tuleb väga selgelt paika panna edasised arengusuunad andmete kogumisel.

Milliseid andmestikke kogutakse ja kuidas neid kogutakse? Kui soovetakse koguda mingeid uusi andmeid, mida varem ei ole kogutud, siis tuleks mõelda, kuidas tahetakse nad andmebaasiga siduda. Üldiselt on selline andmebaas väga avatud ja andmeid saab ID kaudu siduda andmebaasiga. Peamiseks ID-ks jääb arvatavasti kooli kood. Seega tuleks kohe läbi mõelda, kas praegune kooli kood on edaspidiseks andmekogumiseks sobiv või mitte. Kui on ette näha, et koode on vaja muuta, tuleks need muutused korrektselt dokumenteerida ja paika panna ka see, kuidas saab vanad koodid uude süsteemi üle viia.

Kui muutusi plaanis ei ole ja andmekogumine jätkub samadel alustel on andmete uuendamine andmebaasis lihtne. Loomulikult tuleks täpselt jälgida, et kooli kood püsiks unikaalse tunnuseks ja seda ei omistataks vahepeal teistele koolidele. Käesoleva projekti raames kasutatud andmebaasi formaat (Access) võimaldab uusi andmeid importida paljudes erinevates andmeformaatides. Samas peab arvestama ka turvalisusega, andmete eeldatava kogumahu ning andmebaasi kasutamise kiirusega, mille osas Access teatud hetkel enam sobida ei pruugi.

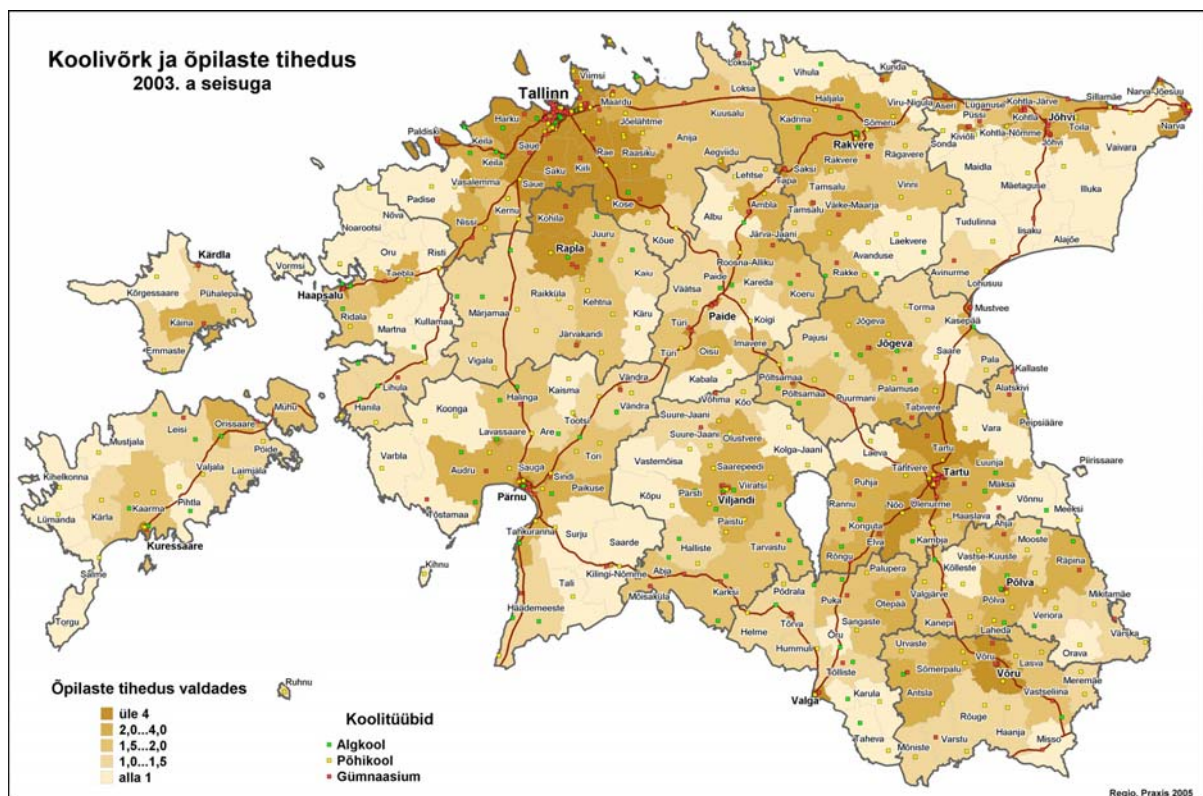
### 3. Teemakaardid

Projekti tulemusena koostati loodud andmebaasi (vt ptk. 2.6) abil suur hulk teemakaarte, millest väike osa on järgnevalt näidatena toodud. Lihtsamate teemakaartide tegemise võimaluse annab ekspertsüsteemi näidisvariant (õpilaste arv, klassikursust kordamajäänute arv, internaadi olemasolu jne), kuhu juurde saab tekitada uusi teemakaarte. Näiteks ei ole käesolevas projektis teemakaarte õpetajate kohta, sest vastav andmestik ei ole päris täielik ja jõudis esialgse plaani väliselt projekti viimastel päevadel. Lisaks esialgsetele ekspertsüsteemi vahenditele, mida saab edasi arendada, oleks edaspidi võimalik teemakaarte teha ka andmebaasist keerukate päringute põhjal, visualiseerida erinevaid näitajaid ja kriteeriume koolide jm kohta.

#### 3.1. Koolivõrk ja õpilaste tihedus

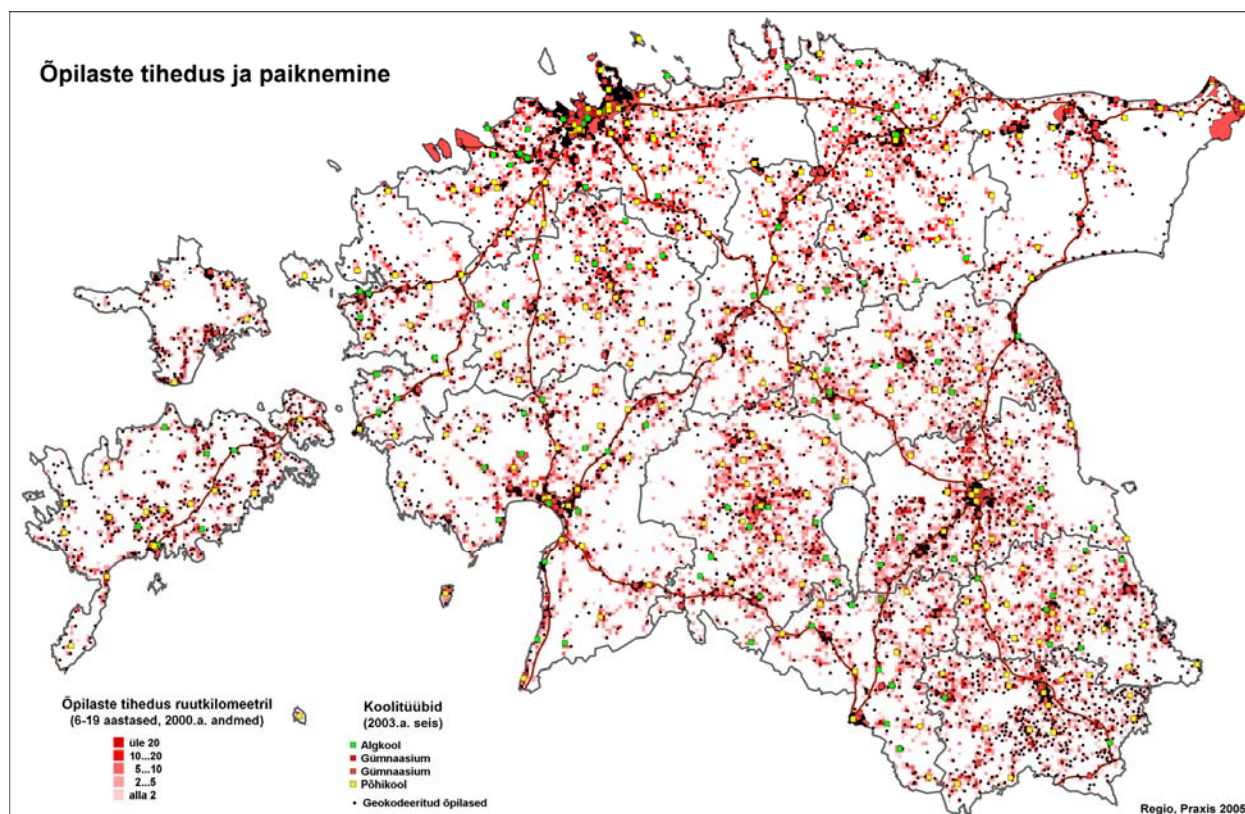
Koolide teeninduspiirkonnad on näha joonistel 1 ja 2. Joonisel 6 on Eesti koolivõrku kajastav kaart 2003.a. seisuga ja 576 kooli andmetega. Välja on jäetud kõik vangla- ja erikoolid ning kaardilt puuduvad öhtukoolid, kaugõppekoolid ja koolid, mille juures puuduvad üldhariduskassid (st koolid tüüpidega 160, 162 ja 180).

Õpilaste tihedus on saadud õpilaste registris olnud elukohta näitava valla koodi järgi. Iga valla õpilaste arv on summeeritud ja jagatud vastava valla pindalaga.



Joonis 6. Õpilaste tihedus ja üldhariduskoolide tüübid 2003. a. seisuga.

Õpilaste tihedust on võimalik näidata ka Statistikaametist saadud 2000. aasta rahvaloenduse andmete põhjal (vt joonis 7). Andmed on küll vanemad, kuid võimaldavad näidata õpilaste paiknemist ruumis palju diferentseeritumalt, kui valdade kaupa näitamine. Teemakaardil joonisel 7 on näidatud 6-19 aastaste õpilaste paiknemine 2000. aasta seisuga. Juurde on lisatud õpilaste geokodeerimisel (õpilaste elukohad 2003. aasta rahvastikuregistri järgi) saadud tulemus. Lisatud on veel ka koolid tüüpide kaupa.



Joonis 7. Õpilaste tihedus ja paiknemine 2000.a rahvaloenduse andmetel.

### 3.2. Koolide avamised ja sulgemised

Koolide avamiste, sulgemiste ja ühendamiste kohta saadi andmed Praxiselt. Teemakaardid kajastavad 10 aastast perioodi ja selle ajal toimunud muutusi koolivõrgus. Parema ülevaatlikkuse saamiseks koostati siin teemakaardid kooli astmete kaupa. Näitena on siin joonis 8 koolivõrgu muutused teise kooliastme kohta (ülejäänud kolm kaarti on näha internetis koolide veebirakenduses <http://kaardid.regio.ee/koolid/>). Nende kaartide puhul tuleb arvestada, et algselt koolide registris olnud 787-st koolist (10 aasta andmete põhjal) on kaardilt ära jäetud järgnevad koolid:

- kõik vangla- ja erikoolid;
- kõik õhtukoolid, kaugõppekoolid ja koolid, mille juures puuduvad üldharidusklassid (st koolitüüpidega 160, 162 ja 180);
- kõik Tallinnas asuvad koolid.

Kokku on kaardi koostamisel arvesse võetud jääb esialgselt 787-st koolist 629 kooli. Kooliastmete kaupa kaartide tegemisel on igale koolile leitud päringuga andmeväli näitamaks, kas kooliaste on olemas tervikuna, osaliselt või puudub täiesti, arvestades kõiki 10-nel aastal olnud klasse. Teisi sõnu kui mingil aastal koolis 3. kooliaste puudus, mõnel muul aastal aga oli, siis on selles andmeväljas tunnuseks „jah“, ja kool kaasatakse ka 3. kooliastet kajastavasse teemakaarti. Teemakaardi tegemisse on kaasatud ka need koolid, kui see kooliaste on olemas osaliselt.

1. kooliastet kajastaval kaardil on 621 kooli, 2. kooliastme kaardil 615, 3. kooliastme kaardil 444 ja 4. kooliastme kaardil 182 kooli.

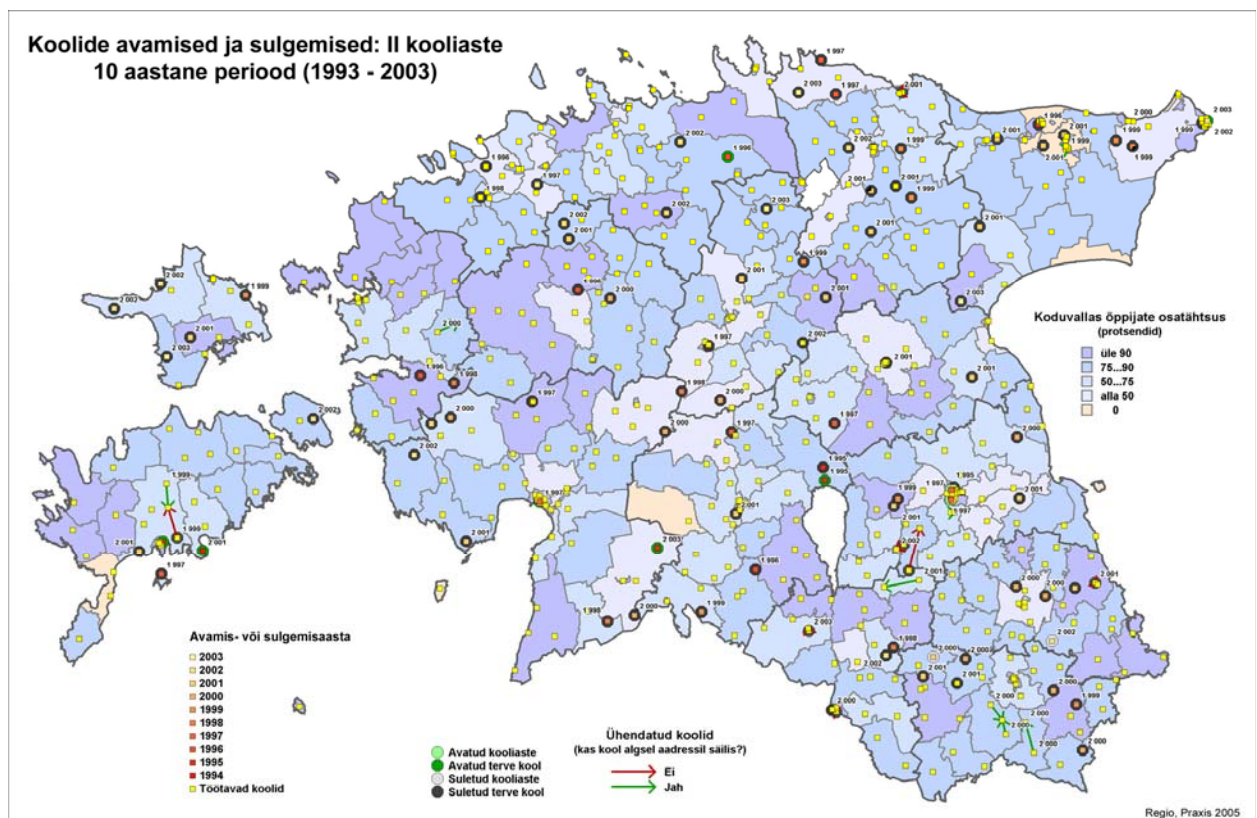
Taustaandmetena on nendel teemakaartidel kujutatud koduvallas õppijate osatähtsus 2003.a. seisuga. Vastav protsent (koduvallas õppijad) on saadud rahvastikuregistrist oleva õpilaste elukoha andmete ja kooli asukoha võrdlemisel. Leiti igas omavalitsuses elavate õpilaste arv ja seoti koolidega. Omavalitsuses elavate õpilaste ja omavalitsuse piires õppijate

suhtarv näitab koduvallas õppijate osatähtsust.

Kuna koduvallas õppijate osatähtsuse arvutamine sõltub mitmest andmetabelist ja andmete omavahelisest sidumisest, siis ei saanud arvutamisel kasutada kõiki õpilaste registris olevaid õpilasi, sest :

- osa õpilaste puhul ei suudetud leida rahvastikuregistrist elukoha andmeid;
- osa õpilasi ei olnud võimalik siduda kooliga kuna vastav kooli kood puudus õpilaste registrist;
- välja jäeti õpilased, kes käivad vangla- ja erikoolides;
- välja jäeti õpilased, kes käivad koolides tüübiga 160, 162 või 180.

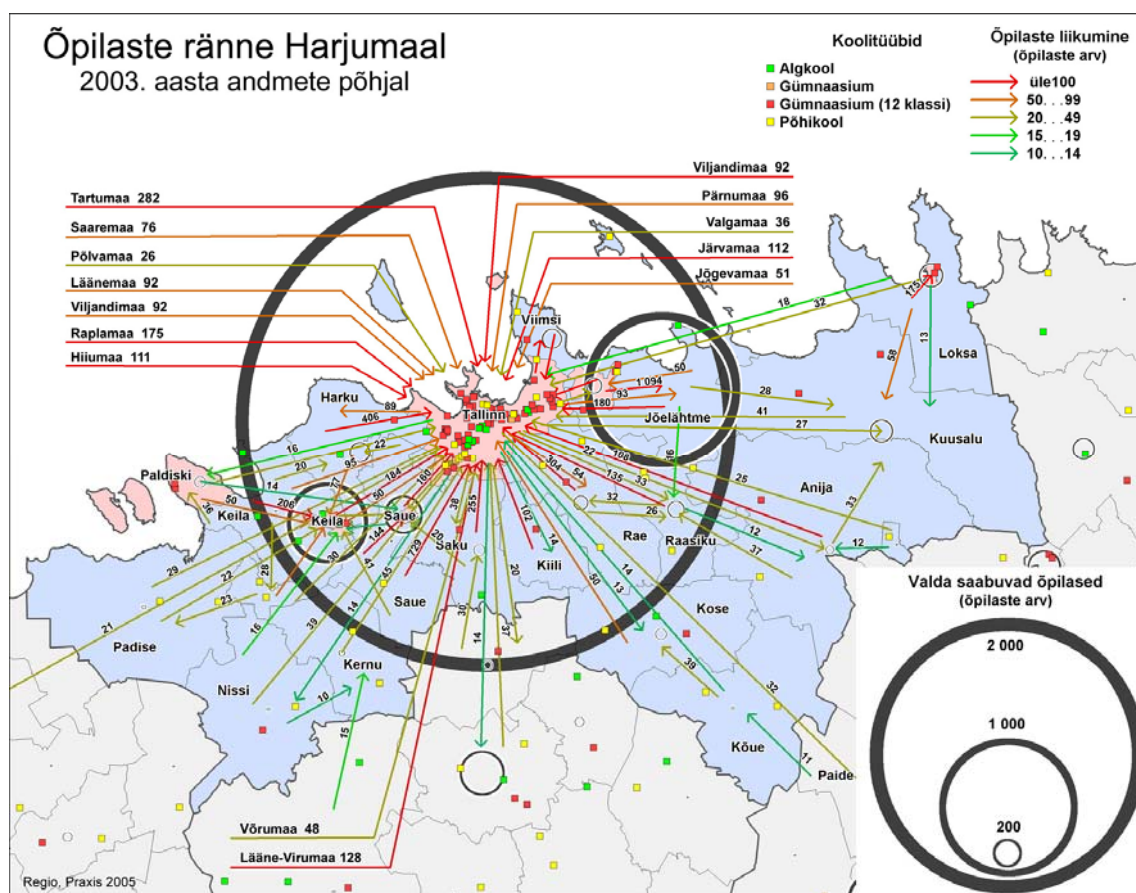
Kokku on kaardi koostamisel arvesse võetud 19 045 õpilase andmed (õpilaste registrisse on kantud 192 525 õpilast). Erinevate kooliastmete kaupa jagunevad õpilased järgnevalt: 1. kooliastmes – 38 260, 2. kooliastmes – 44 378, 3. kooliastmes – 62 729 ja 4. kooliastmes – 33 678 õpilast.



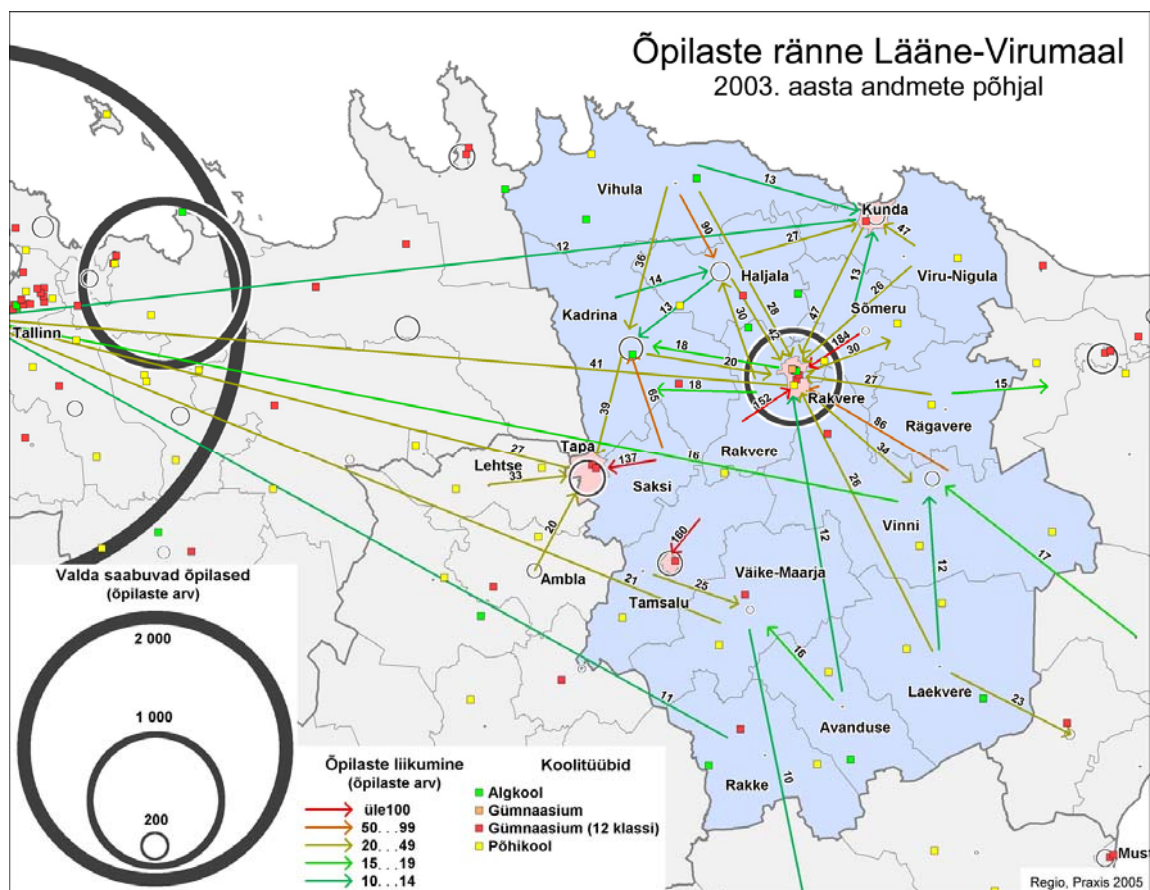
Joonis 8. Koolide avamised, sulgemised ja ühendamised 10 aasta jooksul (2. kooliaste).

### 3.3. Õpilaste ränne ja tõmbekeskused

Suurt huvi pakkus õpilaste ränne. Selle illustreerimiseks koostati iga maakonna kohta õpilaste rände kaart. Eelnevalt oli leitud koduvallas õppijate osatähtsus, st kui palju õpilasi käib koolis oma koduvallas. Tehti ka kindlaks, millised on suuremad tõmbekeskused ja milline on nende suhteline osakaal üksteisega võrreldes. Projekti käigus loodud andmebaas võimaldas visualiseerida erinevates registrites sisalduva informatsiooni. Nii koostati iga maakonna kohta õpilaste rändekaart. Joonisel 9 ja 10 olevad teemakaardid illustreerivadki kirjeldatud näiteid. Kuid samamoodi nagu eelmistest peatükkidest nähtub, ei õnnestu kõiki andmeid omavahel siduda, seega jääb ka õpilaste rände kaartidel kajastamata 13 480 õpilase info (192 525-st õpilasest).



Joonis 9. Õpilaste ränne Harjumaa 2003. aasta andmete põhjal.

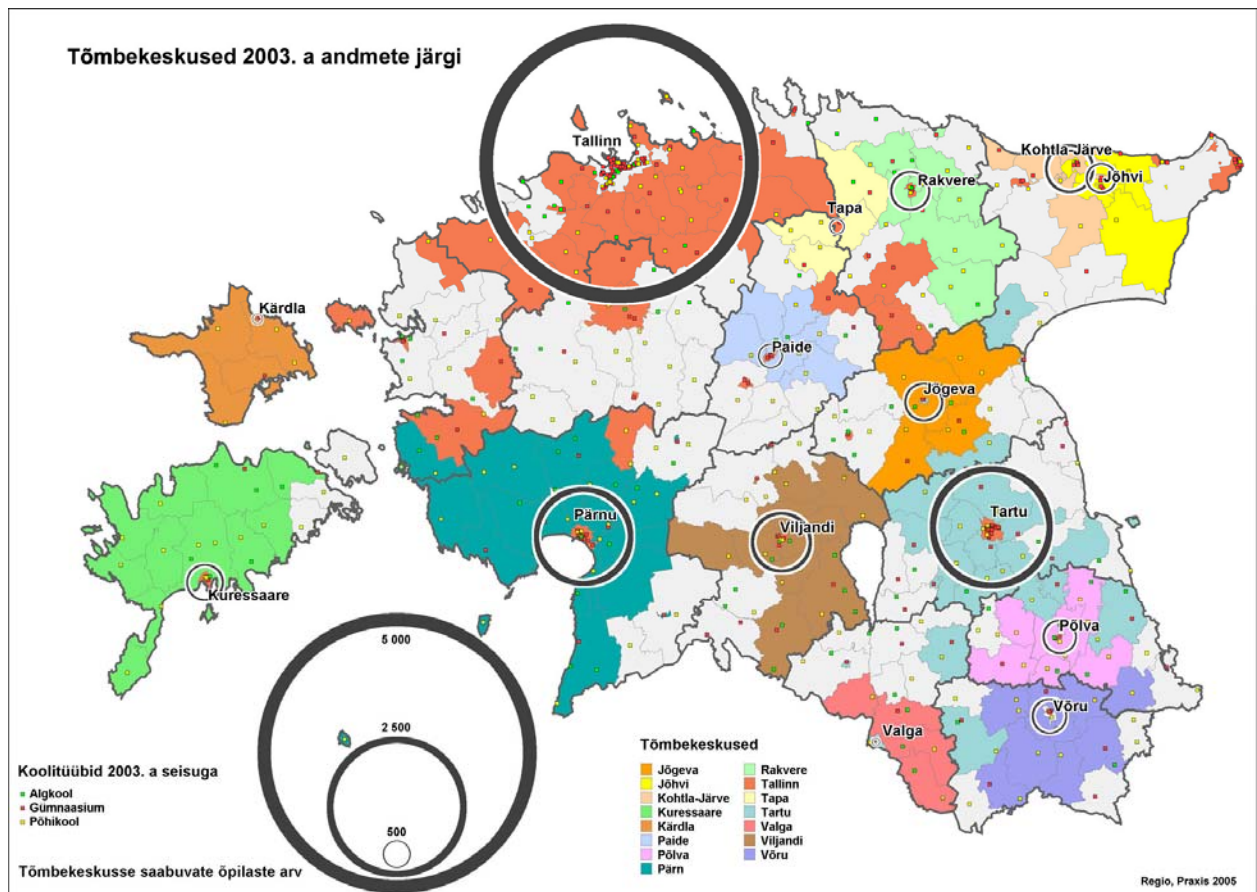


**Joonis 10.** Õpilaste ränne Lääne-Virumaal 2003. aasta andmete põhjal.

Praegusel juhul on rände kaardid koostatud rahvastikuregistrist saadud õpilase elukoha andmete alusel. Ka õpilaste registris oli õpilase elukoht valla koodi näol olemas. Peatükis 2.3. Õpilaste- ja rahvastikuregister on neid elukohaandmeid ka võrreldud (tabel 7) ja erinevus leiti 12 304 õpilase osas. Kui siia juurde liita veel ka need õpilased (3382 õpilast), kellele elukohta rahvastikuregistrist juurde ei leitud, siis saame kokku 15 686 õpilast, mis on 8,1% kogu õpilaste arvust. Võib arvata, et rände kaartide puhul tuleb arvestada analoogse veaprotsendiga.

Rände kaartidel kujutatud ringid näitavad õpilaste absoluutarvu, kes saabuvad teisest omavalitsusest. Terve Eesti kohta on välja toodud 15 tõmbekeskust (Joonis 11).





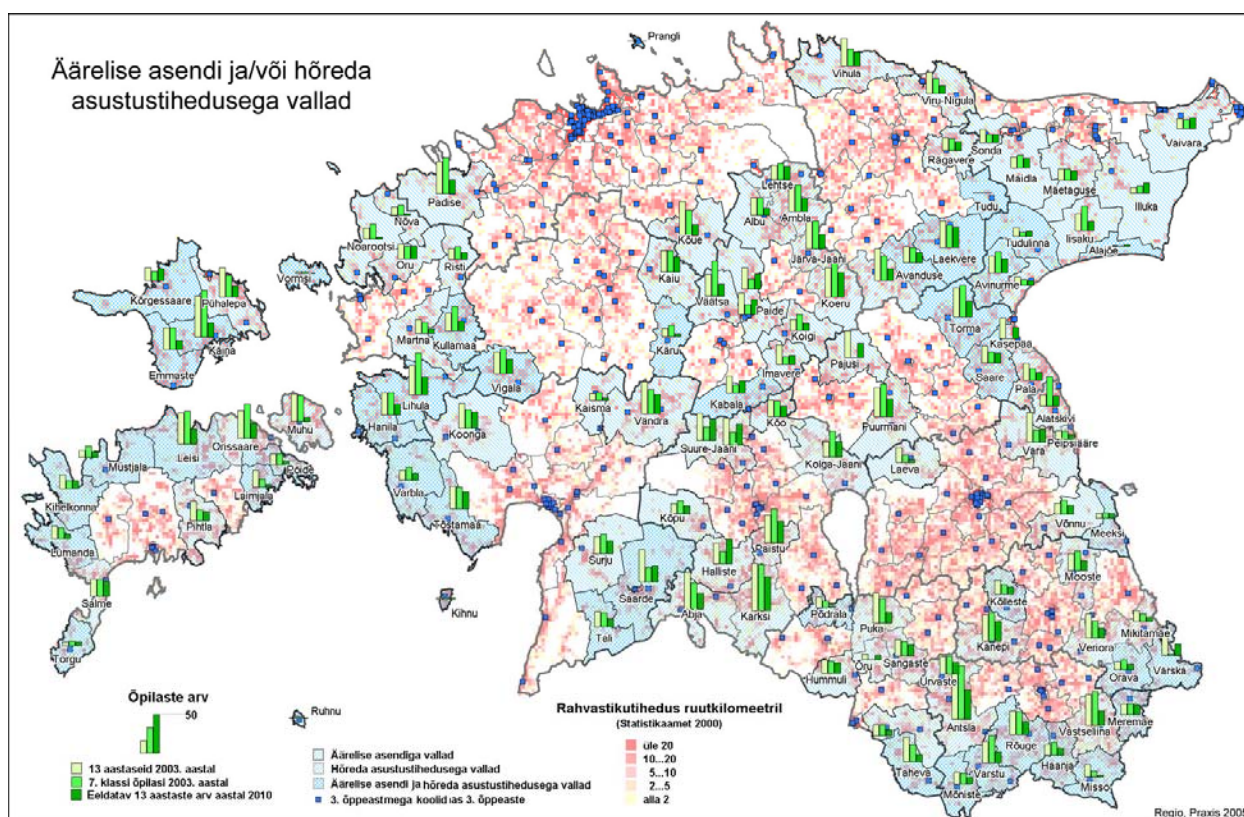
**Joonis 11.** Tõmbekeskused 2003.a. õpilaste rände järgi.

Tõmbekeskuste kaartide koostamise põhimõtted:

- Värvitud on omavalitsused ehk tõmbekeskus on määratud, ainult nendes valdades, kust lahkub üle 20% õpilastest;
- iga vald on toonitud selle tõmbekeskuse järgi, kuhu läheb kõige rohkem õpilasi;
- tõmbekeskust ei ole eristatud, kui see hõlmab endas vähem kui nelja omavalitsust;
- arvestatud on kõikide õpilaste rännet, hoolimata kooliastmetest.

### 3.4. Äärelise asendi ja/või hõreda asustustihedusega vallad

Äärelise asendiga ja/või hõreda asustustihedusega omavalitsused on välja toodud Jaak Kliimaski andmete põhjal. Temalt saadud tabelandmed sisaldasid lisaks valdade äärelisele asendile ja/või hõredale asustustihedusele ka vallas elavate 13 aastaste laste arvu, 7-ndas klassis õppivate laste arvu ja prognoositavat 13-aastaste laste arvu aastal 2010. Vastav info seoti kaardiga ja koostati teemakaardid neljas erinevas versioonis, millest ühte kokkuvõtvat teemakaarti näeb joonisel 12. Ülejäänud kahel kaardil on kujutatud ainult äärelise asendi ja hõreda asustusega omavalitsusi. Kolmas kaart kujutab omavalitsusi, mis on üheaegselt nii äärelise asendi kui ka hõreda asustusega. Ülejäänud kolm kaarti on kättesaadavad koolide veebirakenduses.



Joonis 12. Äärelise asendi ja/või hõreda asustustihedusega vallad.

## 4. Koolide veebirakendus

Koolide veebirakendus on atraktiivne, lihtsasti kasutatav ning ülevaatlik vahend, kust on võimalik saada informatsiooni:

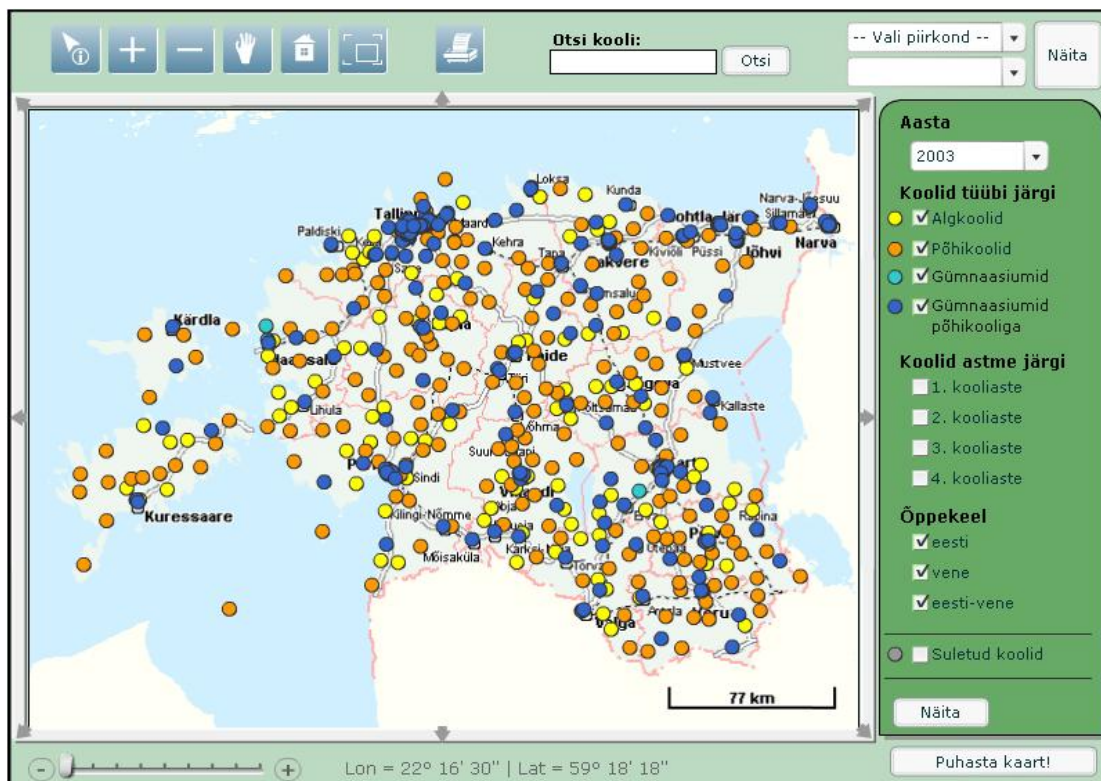
- koolide asukoha, kooli tüübi ja kooli astme ning õppekeele kohta,
- näha koolivõrgu arengut aastatel 1994-2003,
- olulisemaid andmeid koolide kohta aastatel 1994-2003.

### 4.1. Funktsionaalsus

Koolide veebirakendus asub aadressil <http://kaardid.regio.ee/koolid/>.

Rakendus on piiratud ligipääsuga ning nõuab kasutajatunnust ning salasõna.

Rakenduse kasutajatunnus ja salasõna on üle antud kaardiserveri kasutamislepingus.



Joonis 13. Koolide veebirakenduse avavaade.

### Rakenduse kasutamine

Kaardil navigeerimise tarvis on kaardirakenduse ülaserivas nuppude rida.



Joonis 14. Koolide veebirakenduse navigatsiooni nupud.

Funktsionaalsuse nupud (vasakult paremale):

- Info: koolide üldandmete küsimine
- Suurenda: vasaku hiireklahviga kaardil klikkides või vasakut hiireklahvi all hoides kaardil vedades suurendatakse kaarti
- Vähenda: vasaku hiireklahviga kaardil klikkides või vasakut hiireklahvi all hoides kaardil vedades vähendatakse kaarti
- Nihuta: vasakut hiireklahvi all hoides nihutatakse kaarti soovitud suunas
- Terve piirkond: avavaate taastamine
- Naviaken: navigatsiooniakna sisse- ja väljalülitamine
- Trükkimine: kaardiakna väljatrukkimine
- Otsi kooli: kooli nime osa sisestamise saab otsida koole, mis sisaldavad vastavat sõna või selle osa
- Näita maakond ja vald: rippmenüüst valides ning "Näita" nuppu vajutades liigutakse suurendatakse kaart valitud maakonna või valla tasemele

### Koolide valimine

Koolide kuvamiseks on rakenduse paremal küljel valikute tegemise võimalused

- Aasta: kuvatakse valitud aastal tegutsevad koolid
- Kooli tüübi järgi: näidatakse valitud tüübiga koole
- Kooli astme järgi: näidatakse koole, kus on vastav aste olemas
- Õppekeel: saab valida, millise õppekeelega koole kuvatakse
- Suletud koolid: näidatakse suletud koole vastavalt valitud kooli tüübile või astmele

NB! Pärast igakordset valiku tegemist peab vajutama nuppu "Näita", alles selle järel kuvatakse need kaardil.

**Aasta**  
2003

**Koolid tüübi järgi**

- Algkoolid
- Põhikoolid
- Gümnaasiumid
- Gümnaasiumid põhikooliga

**Koolid astme järgi**

- 1. kooliaste
- 2. kooliaste
- 3. kooliaste
- 4. kooliaste

**Õppekeel**

- eesti
- vene
- eesti-vene

Suletud koolid

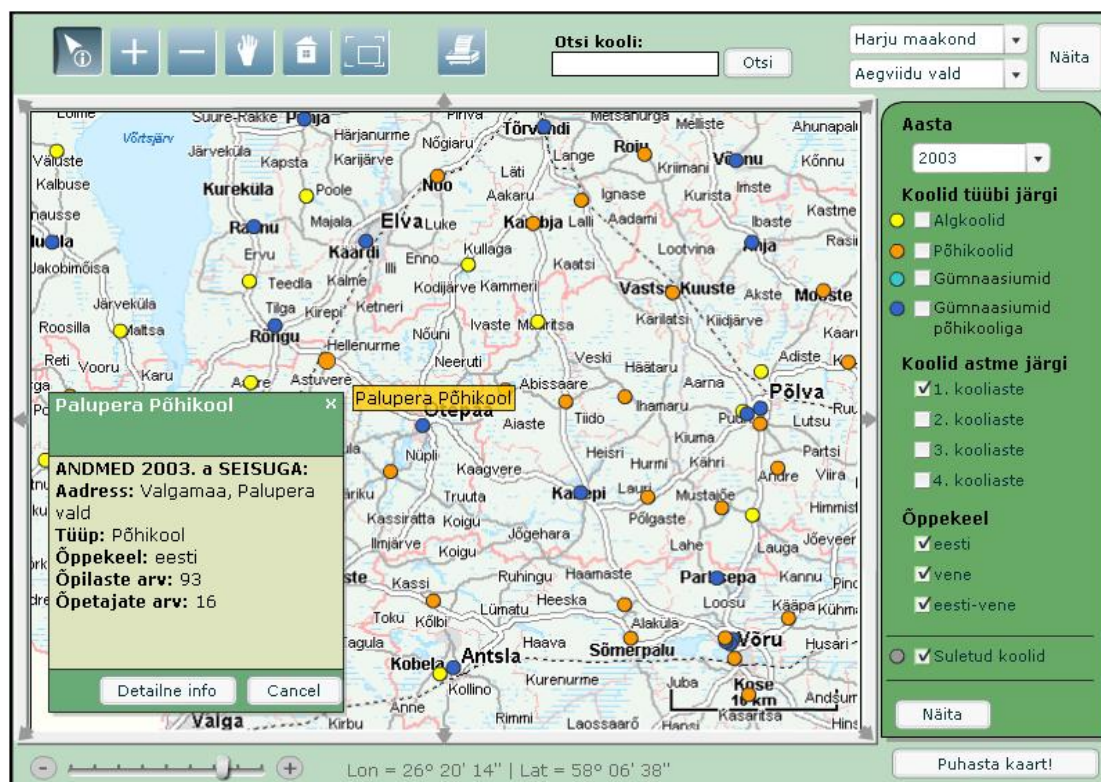
Näita

Puhasta kaart!

Joonis 15. Koolide valiku kriteeriumid.

## Kooli kohta info pärimine

Kooli kohta informatsiooni pärimiseks tuleb valida ülemisest nuppudereast "Info" nupp. Lii- kude kooli sümbolile kuvatakse kooli nimi. Sümbolil klikkides avaneb infoaken, milles on ku- vatud kooli üldandmed vastavalt valitud aastale.



Joonis 16. Kiirpäringu tulemused koolide kohta.

Kiirpäringus on toodud järgmised andmed:

- Address
- Tüüp
- Õppekeel
- Õpilaste arv
- Õpetajate arv

Valides kiirpäringust nupu "Detailne info" avaneb täiendav infoaken, milles on täpsem info õpilaste arvu kohta klasside kaupa, klassikomplektide arv jne. Võimalik on rippmenüüst valida aasta, mille kohta andmeid soovitatakse kuvada.

Kooli info: Palupera Põhikool - Microsoft Internet Explorer

**Palupera Põhikool**  
2001 a. seisuga

Vali muu aasta seis: 2001

2003
2002
2001
2000
1999
1998
1997
1996
1995
1994

<b>Aadress</b>	Valga maakond, Palupera v
<b>Koolipidaja</b>	Palupera Vallavalitsus
<b>Omandivorm</b>	Munitsipaalomand
<b>Koolitüüp</b>	Põhikool
<b>1. kooliaste (1-3 klass)</b>	jah
<b>2. kooliaste (4-6 klass)</b>	jah
<b>3. kooliaste (7-9 klass)</b>	jah
<b>4. kooliaste (10-12 klass)</b>	ei
<b>Õppekeel</b>	eesti

<b>Õpilaste arv</b>	94
<b>Tüdrukute arv (ek)</b>	47
<b>Poiste arv (ek)</b>	47
<b>Tüdrukute arv (vk)</b>	0
<b>Poiste arv (vk)</b>	0
<b>Klassikomplekte (ek)</b>	10
<b>Klassikomplekte (vk)</b>	0
<b>Klassiruumide arv</b>	11
<b>Õpilasi pikapäevarühmas</b>	26
<b>Klassikursust kordamajäänuid</b>	5

**Õpetajate arv (ainult 2003)**

**Avamisaasta**

**Sulgemisaasta**

Klass	Õpilaste arv klassis
1	8
2	15
3	9
4	13
5	14
6	11
7	6
8	10
9	8

Joonis 17. Detailse päringu tulemus kooli kohta.

## 4.2. Kasutatud andmestikud

Koolide veebirakenduses on kasutatud väljavõtte käesoleva projekti tulemusena loodud andmebaasist (vt ptk 2.6), mis on tõstetud Accessi andmebaasist veebipäringuid võimaldasse MySQL andmebaasi.

Rakendus baseerub Regio kaardiserveril, milles asuvat kaardiandmebaasi uuendatakse pidevalt. Kaardiserverist päritud kaardil kuvatakse vastavalt kasutaja poolt veebirakenduses tehtud valikutele MySQL andmebaasist võetav koolide info.

## 4.3. Koolide andmestike uuendamine ja arendus

Koolide andmestike edasine uuendamine ei ole käesoleva projekti raames planeeritud. Põhimõtteliselt on see edaspidise nõudluse tekkimisel võimalik. Koolide informatsiooni uuendamiseks on võimalik projekti tulemusena loodud andmebaasi igal aastal eelmise õppeaasta andmetega täiendada (kui hetkel on uusimad andmed seisuga 2003, siis sel aastal tuleks täiendada baasi 2004. a. andmetega jne). Andmestike uuendamise automatiseerimine on võimalik tulevikus, kuid praegu saaks vastavate andmete olemasolu korral uuendusi teha Regio.

Koolide andmeid veebirakenduses saab uuendada ka Regiost sõltumatult, seda juhul kui

MySQL andmebaas hakkaks tulevikus asuma Haridus- ja Teadusministeeriumi serveris. Sellisel juhul hakkab rakendus pärima kaardiandmeid Regio kaardiserverist ning koolide andmeid Haridus- ja Teadusministeeriumi serverist ja kasutaja arvutis asuv brauser paneb kaks andmestikku veebirakenduse abil kokku.

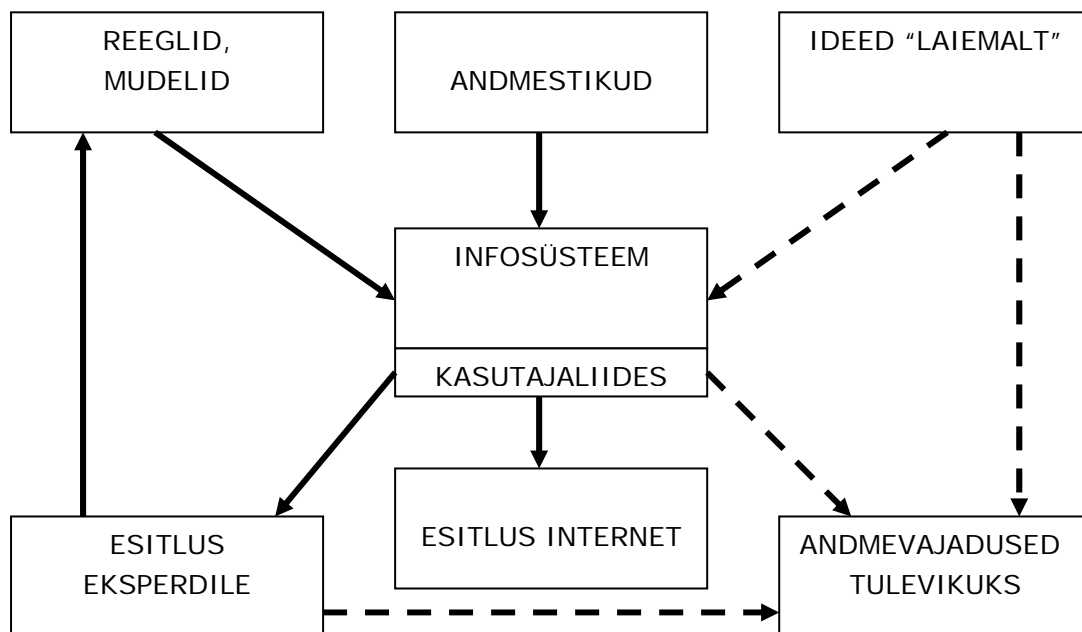
Võimalikud veebirakenduse edasiarendused:

- Uute ja/või keerulisemate päringute võimalus (nt koolide kuvamine õpilaste arvu alusel; millistest valdadest tuleb antud kooli õpilasi, millistesse koolidesse lähevad õppima valla õpilased jne).
- Kooli kohta olemasoleva informatsiooni laiendamine (nt arvutite arv, saali pindala, ujula olemasolu jne).

## 5. Ekspertsüsteem “Koolivõrk”

Käesolev projekt oli katse- ja uurimusprojekt, mille üks eesmärke oli selgitada välja ekspertsüsteemi loomise võimalikkust antud andmete, tarkvara ja rahalises situatsioonis, samuti täpsustada selle süsteemi ideestiku ja eesmärkide vastuvõetavust. Valminud töö põhjal saab selgitada huviliste ringi, rakendusvõimalused ja vajadused. Saadud kogemuste alusel saab asuda suuremate võimalustega ekspertsüsteemi projekteerimisele.

Ekspertsüsteemi ülesanne on toetada eksperte otsuste tegemisel, aidates otsuste tegeise reeglistiku ja parameetrite muutmisel saada uusi visualiseeritud kujutisi, mis omakorda aitavad teha järeldusi olukorra kohta. Näidissüsteem realiseeriti MapInfo keskkonnas. Joonisel 18 on kujutatud ekspertsüsteemi põhimõtteline skeem, mida järgnevatel alapeatükkides osade kaupa lahti seletatakse.



Joonis 18. Ekspertsüsteemi põhimõtteline skeem.

### 5.1. Andmestikud ja nende uuendamine

Ekspertsüsteem vajab oma tööks andmeid. Andmetest üksi jääb aga ekspertsüsteemi tööks väheseks. Suurte ja mahukate tabelandmetega pole midagi peale hakata. On vaja informatsiooni ja andmeid kasutajale arusaadaval kujul. Andmetest informatsiooni välja sõelumise viisid on kirjas ekspertsüsteemi reeglite ja mudelite osas. Kasutajaliidesega tihedalt seotud infosüsteem on poolautomaatne vahend, kus kasutaja saab valida andmetega tehtavad operatsioonid ja tulemusena edastatakse talle andmed töödeldud kujul, mille põhjal kasutaja võib teha teatud järeldusi, st ta saab nendest andmetest informatsiooni.

Andmete ja informatsiooni vahe ongi see, et informatsioon omab kasutaja seisukohalt mingit tähendust ja toob välja olulised aspektid. Olenevalt sellest, kuidas andmeid esitada, võib neist välja lugeda erinevaid aspekte. Seepärast ei tohiks olla ekspertsüsteem suletud süsteem. Kõik operatsioonid peaks olema kasutaja poolt kontrollitavad, läbi mängitavad ja läbi nähtavad.

Kõik saab alguse ikkagi andmetest, mistõttu on väga oluline, et nad sisaldaks võimalikult vähe müra (valesid või puudulikke andmeid) ja oleksid ajakohased. Käesolevas projektis kasutatud andmestike kvaliteet (täielikkus, usaldusväärsus, vormistus jms) oli suhteliselt väike. Erinevate andmestike kasutamisel tuli välja rida lünki (vt peatükk 2. Andmestikud.).



Ekspertsüsteem on realiseeritud MapInfo keskkonnas, lähteandmebaas aga MS Accessis. MapInfo oskab avada Accessi tabeleid, seega ei ole mingeid probleeme ka ekspertsüsteemil uute andmete kättesaamisega. Väike puudujääk seisneb selles, et MapInfo ei oska kasutada Accessis defineeritud relatsiooni andmetabelite vahel. See on aga väga kergesti lahendatav, kas eraldi väljavõtte tegemisega Accessist või siis ühendades erinevad tabelid MapInfos uuesti SQL päringu abil.

Kokkuvõttes tähendab, et kui andmebaas on uuendatud, siis on andmed kohe kättesaadavad ka ekspertsüsteemi jaoks. Andmebaasi uuendamine ei ole samuti probleem, kui edaspidi jälgida soovitusi andmete kogumise osas. Andmebaasi uuendamisest loe täpsemalt peatükist 2.6. Projekti tulemusena loodud andmebaas.

## 5.2. Funktsionaalsus

Ekspertsüsteem baseerub andmetel ja sellel, kuidas me neid andmeid tõlgendame ja sealt informatsioonini jõuame. See pole aga kõik, on vaja teada ekspertsüsteemi eesmärki ja lähitult sellest paika panna hindamiskriteeriumid (näitajad). Käesoleva ekspertsüsteemi eesmärgiks oleks ühelt poolt eksperdi töövahendi loomine, mis aitaks erinevaid otsuseid teha. Teiselt poolt võiks see olla ka lihtsalt tavakasutaja töövahend koolide hindamiseks, abivahend kooli valikuks või tagasiside saamiseks.

Andmetest vajaliku informatsiooni ülesleidmiseks on mudelid ja reeglid, mille abil on võimalik andmed otsuste tegemiseks sobivamale kujule ette valmistada. Üheks selliseks viisiks on teemakaartide tegemine, mis erinevate matemaatiliste ja või loogiliste tehete abil visualiseerivad erinevaid teemakihte. Näiteks on võimalik valida järgnevate teemakaartide vahel:

1. Kooli suurus õpilaste arvu järgi – näidatakse lihtsalt kooli suurusi erinevate klasside-na.
2. Õpilaste arv koolitüüpides (põhikool) – arvutatakse põhikoolis käivate õpilaste arv (summeeritakse 1.-9. klassini õpilaste arvud) ja võrreldakse piirnormiga (min 90 õpilast põhikoolis).
3. Õpilaste arv koolitüüpides (gümnaasium) – arvutatakse kõikide koolide, 4. kooliastet omavate, gümnaasiumiosa õpilaste arv ja võrreldakse piirnormiga 60 õpilast ja rohkem.
4. Klassiruumi ühe õpilase kohta – klassiruumide pindala jagatakse õpilaste arvuga koolis.
5. Arvuteid õpilase kohta.
6. Klassikursust kordamajäänud õpilaste osatähtsus.
7. Õpilasi õpetaja kohta.
8. Kõrgharidusega õpetajate osatähtsus.

On tehtud ka mudel, mis näitab õpilaste rännet st arvestatud kõigi õpilastega, kes lähevad oma koduvallast teise omavalitsusse. Õpilaste liikumised on ära kaardistatud ja üldistatud valla tasemele. Seega tabelandmetest, mis sisaldavad aadresse ja omavalitsuse koode, on jõutud süsteemi, mis lubab saada ülevaadet omavalitsuse kaupa, kust tuleb õpilasi juurde ja kuhu nad lahkuvad. Edasi on juba eksperdi ja tõlgendaja küsimus, miks õpilased selliselt liiguvad. Võib-olla on tegemist teatud valdkondades paremate koolidega? Sellistele küsimustele hakkab vastuseid saama koolide hindamise ekspertsüsteemi järgmisest osast: koolide hindamine.

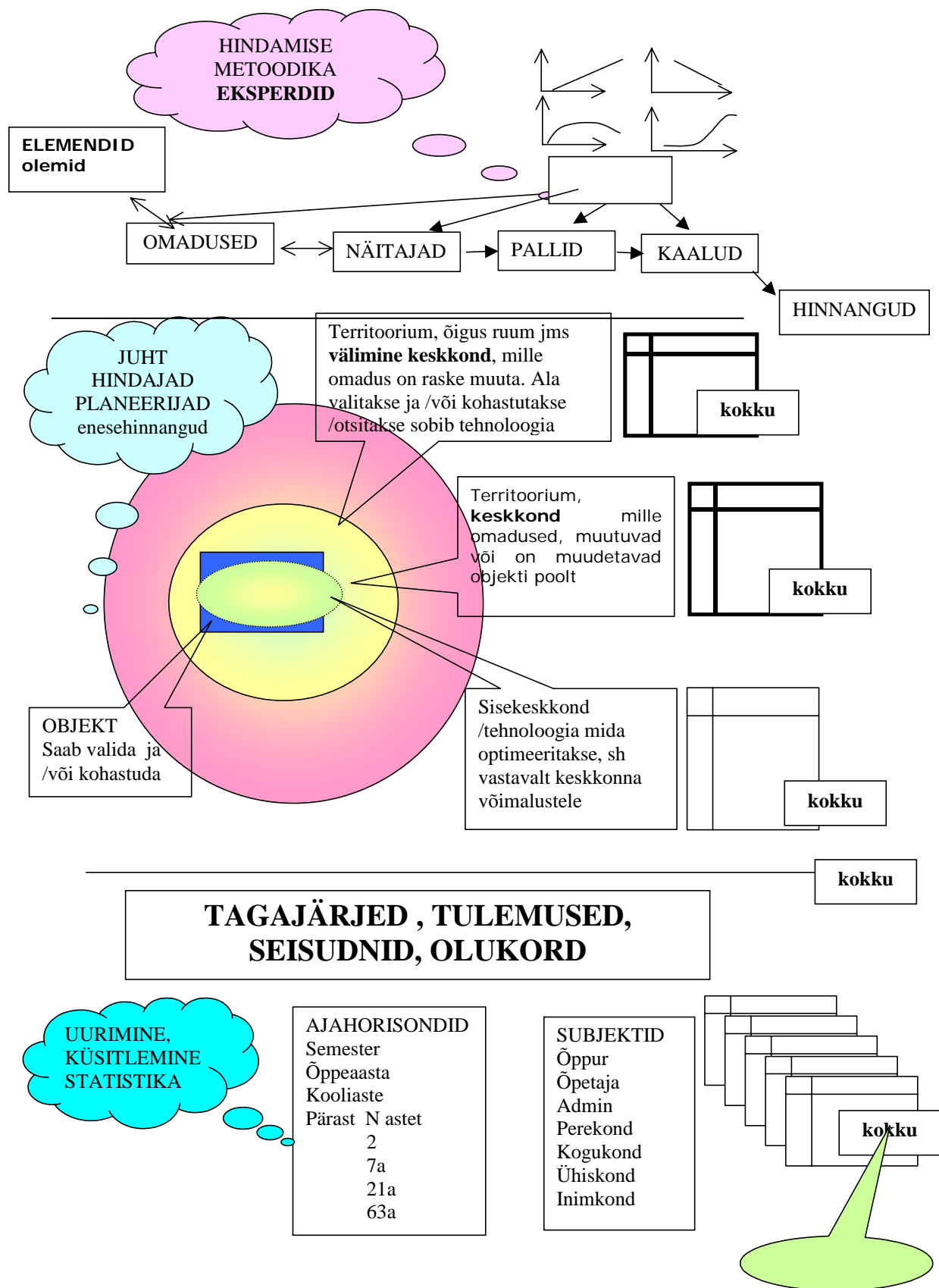
Koolide hindamine olekski tulevikus ekspertsüsteemi üheks peamiseks osaks. Hindamissüsteemi põhimõtteline skeem on toodud joonisel 19. Joonise ülemine kolmandik kujutab endast koolide hindamise üldskeemi. Elemendid (olemid) on antud juhul koolid. Koolidel on teatud omadused, mida on võimalik erinevate näitajatega mõõta. Mõõdetavad tulemused on omakorda ümber arvutatavad või klassifitseeritavad pallideks ehk hindepunktideks. Kõik see toimub läbi ekspertsüsteemi kasutajaliidese eksperdi kontrolli all ja viimase poolt valitud

meetoditel. Näidatud graafikud tähendavadki seda, et näitajalt üleminek pallideks toimub erinevate funktsioonide abil, mida ekspert saab valida sõltuvalt näitaja jaotusest. Järgmise sammuna antakse erinevatele näitajatele kaalud. Kaalud võimaldavad teatud näitajate olulisust tõsta ja teiste oma vähendada. Kaale arvestades ja palle summeerides saame kokkuvõtva hinnangu. Kaalud võimaldavad hindamissüsteemi häälestada ja leida olulised ning vähem olulised näitajad.

Joonise 19 keskmisel osal on vaadatud kooli omadusi erinevatest aspektidest. Kooli ümber on väline ruum ja keskkond, mida ei saa muuta, millega tuleb kohastuda (seadused, teedevõrk) ja see osa, mida on võimalik muuta – mõlemal on aga omadused, mida võib teatud näitajatega iseloomustada. Mõlema aspekti osas on võimalik valida näitajate abil ja teha läbi hindamisprotsessi ja saada kokkuvõttev hinnang. Analoogselt välise keskkonnaga on ka koolil sisekeskkond ja koolihoone mille kohta saab koostada läbi erinevate näitajate koondhinnangu. Hinnangute andmine võib toimuda erinevate huvigruppide seisukohalt vaadatuna, kes hindavad oluliseks erinevaid näitajaid.

Andes erinevatele hindamisaspektidele erinevad kaalud võime koostada hierarhilise süsteemi ja leida ühe koondhinde tervele koolile.

Joonise 19 alumine osa kajatab suundumist tulevikku. Lisades ekspertsüsteemile ajalise mõõtme, saame hakata koole hindama erinevatel ajahetkedel. Neid hindeid on võimalik säilitada, juurde teha küsitlusi ja statistikat. Lõppkokkuvõttes saame rääkida lisaks hetkeolukorrale koolivõrgus ka otsuste tulemustest ja tagajärgedest. Ekspertsüsteem võiks kujuneda ka erinevatele kasutajagruppidele (subjektidele) kättesaadavaks. Selleks on tarvis erinevaid kasutajasemeid. Näiteks võiksid õpilased täita rahulolu ankeeti oma kooli kohta läbi interneti. Direktor näeks selle tulemust ja ekspert saaks küsitluse tulemusi võrrelda teiste näitajatega ning arvestada tulemusi kooli hindamisel, omistades neile mingi kaalu.



Joonis 19. Ekspertsüsteemi hindamissüsteemi põhimõtteline skeem.

## 5.3. Kasutajaliides

Programmi käivitamine toimub MapInfo keskkonnas valides rippmenüüst 'Tools' käsu 'Run Mapbasic Program'. Seejärel tuleb leida üles kaust arvuti kõvakettalt, kuhu programm ja andmed salvestati ning avada sealt fail 'haridus3.mbx'. MapInfo rippmenüüde hulka tekib menüü 'HARIDUS'. Rippmenüüst käsku valides saab avada kasutajaliidese, mille abil saabki alustada tööd ekspertsüsteemiga.

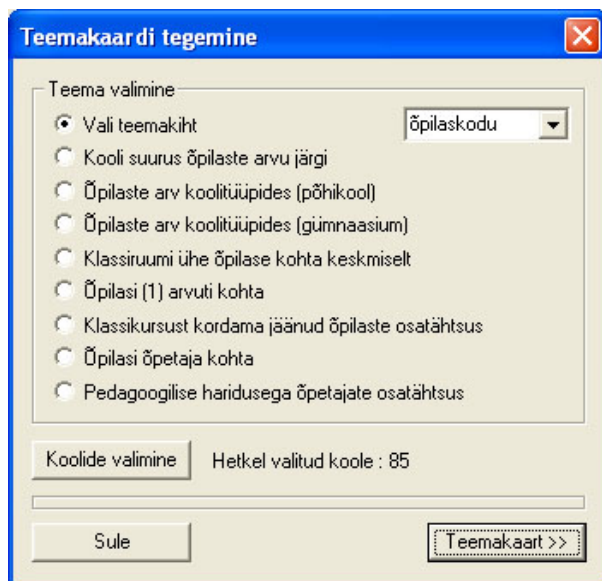
Kasutajaliides on vahend, millega suhelda andmetega ja muuta reegleid, mudelite abil tõlgendada andmeid teabeks ning kujutada neid tabelite ja kaartidena. Ekspertsüsteemi kasutajaliidese võiks jagada kolme ossa:

1. Koolide teemakaartide tegemine.
2. Õpilaste rände kaartide tegemine.
3. Koolide hindamine ja hinnangukaartide tegemine.

### 5.3.1. Koolide teemakaartide tegemine.

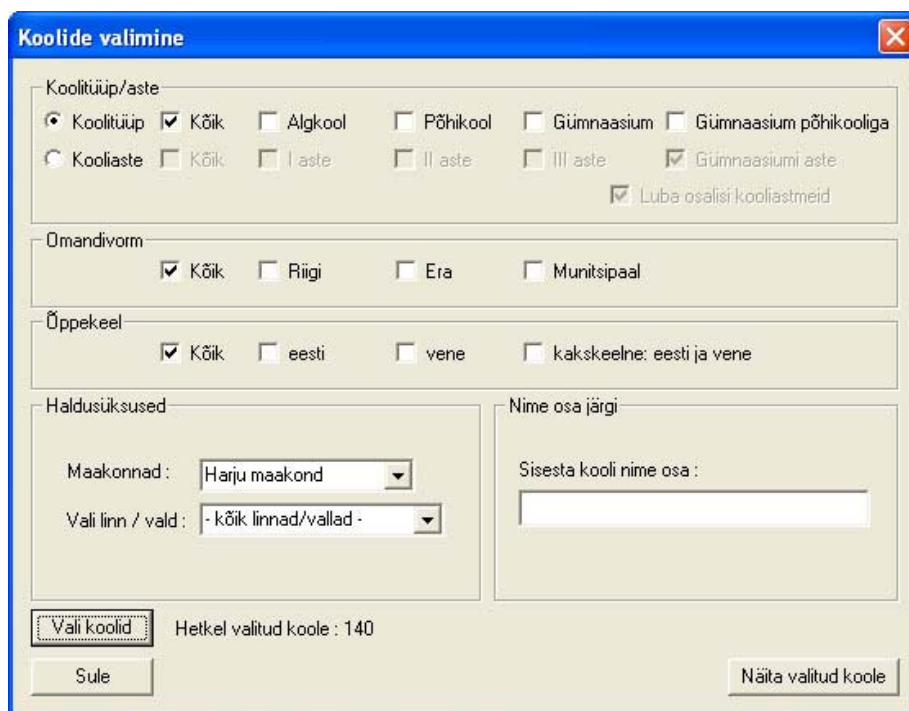
Koolide teemakaartide tegemine toimub kasutades joonisel Joonis 20 kirjeldatud dialoogiakent.

Teemakaarte saab koostada ühe teemakihi põhjal valides rippmenüüst – kooli tüüp, õppekeel, omandivorm, õpilaskodu olemasolu, ujula olemasolu. Samuti saab luua teemakaarte eelnevalt päringutega ettevalmistatud andmetest – vt joonis 1. Korraga saab valida vaid ühe teemakaardi tegemise.



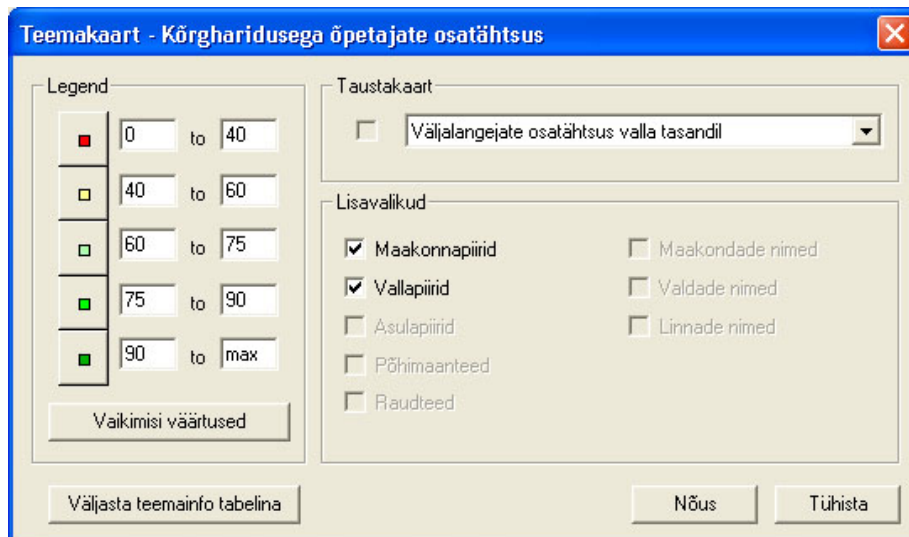
Joonis 20. Teemakaartide valiku dialoogiaken.

Teemakaardi valimise eel tuleks teostada koolide valimi moodustamine. Vajutades nupule 'Koolide valimine' avaneb dialoogiaken (Joonis 21), kus saab erinevate tunnuste abil moodustada valimit – koolitüüp/kooliaste, omandivorm, õppekeel, haldusüksus (maakond/omavalitsus), kooli nime osa. Dialoogiaknas on kuvatud valitud koolide arv. Samuti saab valikut vaadata kaardil – 'Näita valitud koole'. Koolide valiku järel tuleb aken sulgeda.



Joonis 21. Koolide valiku dialoogiaken.

Teemakaardi tegemise aknast edasi minnes, vajutades nupule 'Teemakaart' avaneb järgmine dialoogiaken (Joonis 22), kus on võimalik valida taustakaarti, kuvatavaid infokihte (nt maakonna või omavalitsuse piirid, teed, nimed jne), samuti muuta legendi. Dialoogiaknas saab valida ka andmete väljastamist tabelina.



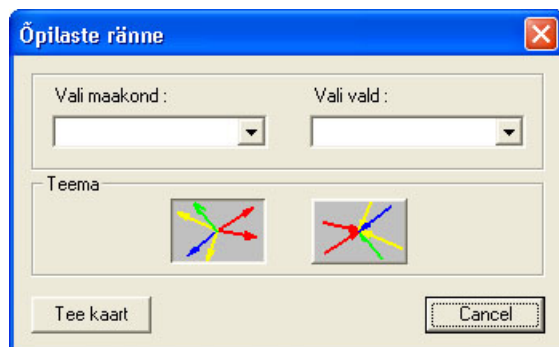
Joonis 22. Teemakaardi ja -info parameetrite muutmise dialoogiaken.

### 5.3.2. Õpilaste rände kaartide tegemine.

Geokodeeritud õpilaste elukohtade andmed lubavad visualiseerida õpilaste rännet. Selleks on rippmenüüs moodul 'Õpilaste ränne' (Joonis 23). Valides teemaks õpilaste väljarände näidatakse valitud omavalitusest õpilaste liikumist teiste omavalitsuste koolidesse. Valides teemaks õpilaste sisserände, näidatakse millistest valdadest liigutakse valitud omavalitusesse õppima. Kaardil iseloomustab rände intensiivsust joone värv, mis on toodud ka lisatud legendis.

Õpilaste elukohaandmed ja kooli asukoht (omavalitsuse tasemel) on saadud rahvastikure-

gistrist. Algandmete puudulikkuse tõttu ei ole võimalik kõiki õpilasi arvesse võtta, samuti pole täiesti usaldusväärsed inimeste registreeritud elukohtade andmed ja tulemused ei ole seetõttu absoluutselt õiged. Õpilaste rände puhul arvestatakse kõigi kooliastmete õpilaste liikumist.



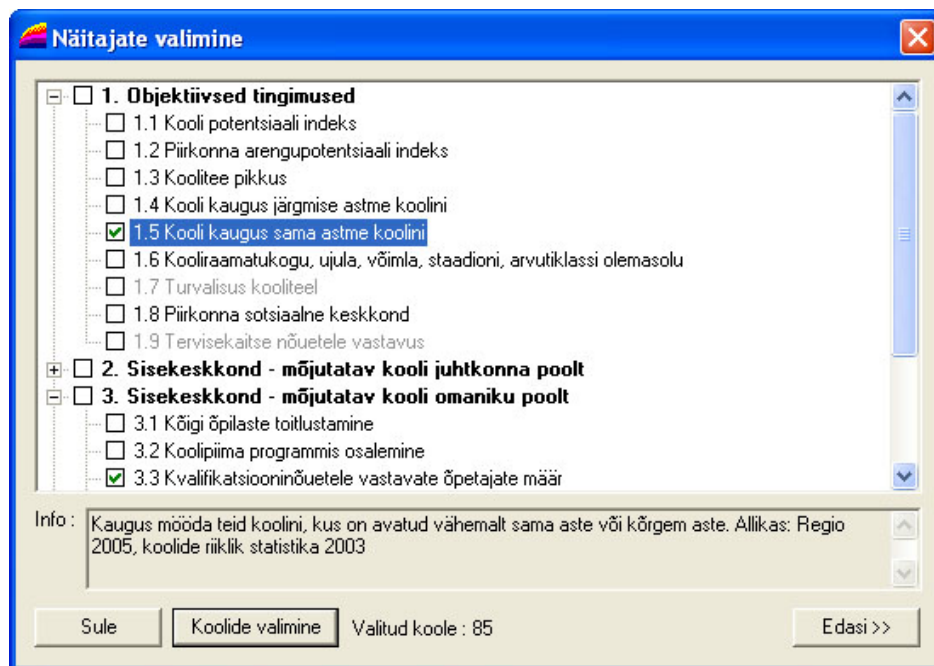
Joonis 23. Õpilaste ränne dialoogiaken.

### 5.3.3. Koolide hindamine ja hinnangukaartide tegemine

Valides rippmenüüst 'Koolide hindamine' avaneb dialoogiaken (Joonis 24). Selle mooduli abil on võimalik: a) võrrelda erinevaid koole näitajate alusel b) hinnata koole etteantud ideaalse näitaja väärtuste järgi.

Näitajad on jaotatud viide rühma. Igas rühmas on mõned näitajad aktiivsed ja mõned mitte. Mitteaktiivsete näitajate puhul pole esialgu olemas andmestikke.

Koole saab hinnata ühe või mitme näitaja alusel. Selleks tuleb linnukesed teha näitaja ees olevasse kasti. Näitajal klikkides on dialoogiakna alaosas antud lühike kirjeldus näitaja sisu ja andmete päritolu kohta.

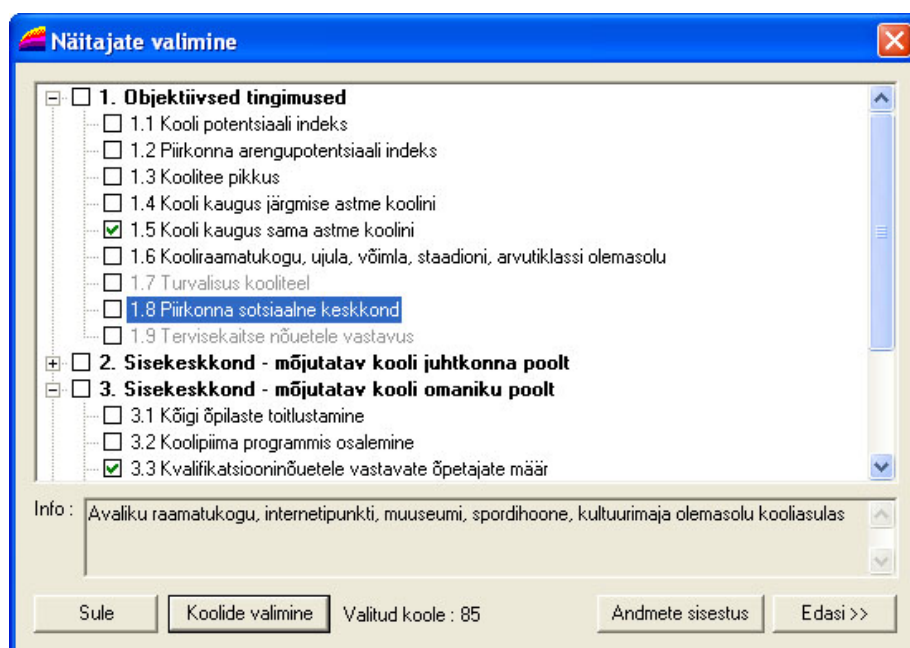


Joonis 24. Kooli hindamiseks kasutatavate näitajate valimise dialoogiaken.

Vajalik on teostada koolide valimine, mida soovitakse hinnata (Joonis 21).

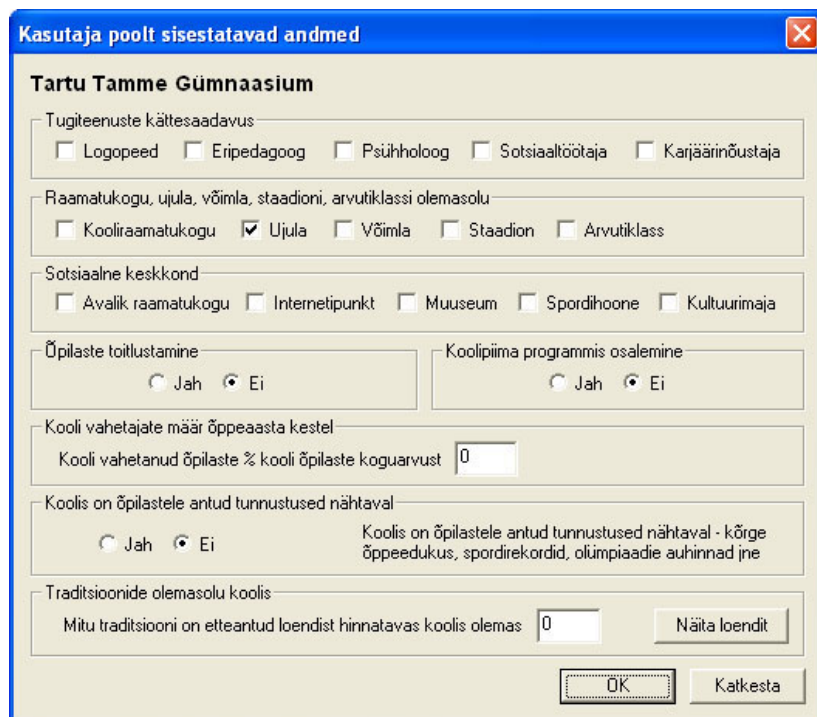
Enamike näitajate puhul pärinevad andmed riiklikust statistikast või on arvutatud kõigi koo-

lide kohta. Teatud kriteeriumite puhul tuleks eksperdil andmed sisestada. Liikudes sellisele kriteeriumile lisandub dialoogiakna allosas nupp 'Andmete sisestus'



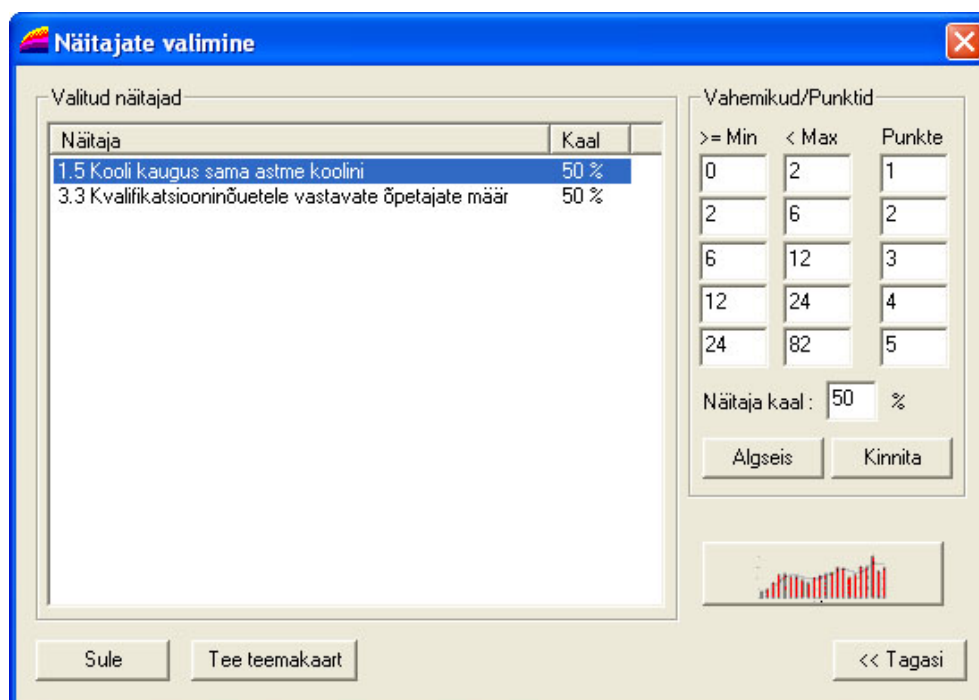
Joonis 25. Näitajate valimise dialoogiaken koos andmete sisestuse nupuga

Vajutades nupule andmete sisestus avaneb dialoogiaken (Joonis 26), kus eksperdil on võimalik oma teadmistele võimalik kooli kohta sisestada andmed, mis salvestatakse programmi juurde kuuluvasse andmebaasi. Edaspidi saab kooli kohta teha päringuid salvestatud andmeid kasutades. NB! Sisestada saab korraga ainult ühe kooli andmeid. Peale uuritava ala (vald, maakond) kõigi koolide andmete sisestamist ja salvestamist saab teha koolide hinnangukaarte tuginedes eksperdi poolt sisestatud andmetele.



Joonis 26. Kooli andmete sisestamise dialoogiaken

Vajutades dialoogiaknas 'Näitajate valimine' nupule 'Edasi' liigutakse dialoogiaknale (Joonis 27), kus on võimalik iga näitajat modifitseerida – muuta tema kaalu, muuta hinnete vahemikke ja hindeid. Graafiku sümbolile vajutades on näha, kuidas koolid jaotuvad antud tunnuse järgi.



Joonis 27. Valitud näitajate kaalude, vahemik ja hindepunktide seadmise dialoogiaken.

Valides nupu 'Tee teemakaart' avaneb juba varem nähtud teemakaardi kujundamise dialoogiaken (Joonis 22) mille abil saab kujundada hinnangukaardi tausta ja legendi. Sealt edasi liikudes arvutatakse välja näitajad ja neile antavad punktid ning tulemus kuvatakse kas teemakaardina või tabelina, kus kõik näitajad ja neile antud punktid on muudetud koondindeks.

## 5.4. Tulemuste esitlus

Käesoleva töö tulemusena on valminud esitlus eksperdile – teemakaardid, hindamise tulemus, mis on kaalude ja punktide poolest ümber mängitav ja tagasi mudeli juurde mindav. Tuleks luua ekspertsüsteem ka erinevatele kasutajatasemetel. Lihtsad kasutajad/vaatajad ja siis ekspert, kes saab esitluse/tulemuse abil luua uusi reegleid ja mudeleid.

Esitlus Internetti – lihtne rakendus esialgu väljavõtte andmebaasist, mida saab lihtsalt uuendada – ei ole otseselt ekspertsüsteemi osa.



## 6. Tähelepanekud ja soovitused

Projekti ettevalmistamise ja teostamise käigus kerkisid esile hulk küsimusi ning mille üle käesolevas peatükis püütakse osaliselt ka vastuseid pakkuda.

- Milliseid andmeid oleks mõttekas tulevikus koguda?
- Mida tasuks andmestike vormistamisel tähele panna?
- Kuidas saaks ministeerium enam abi geograafiliste andmete kasutamisest sh andmestike asukohapõhisel kasutamisel?
- Millisel määral tuleks ka õpetajate andmestik käsitleda?
- Millised oleks antud projekti loogilised jätkud?
- Millises vahekorras hinnata administreeritavust, juhitavust ja õpetajate, õpilaste rahulolu?
- Kuidas anda hinnanguid eri gruppide (subjektide seisukohalt)?
- Kooli tulemuslikkuse ja tagajärgede võrreldavus eri ajalõigetel?
- Millised on ja kuidas arvestada kooli "ökonoomika", perekonna, omavalitsuse, riigi (otsese ja kaudsed) kulutusi?
- Kuidas peaks kajastuma kooli ja selle ümbruskonna, sh kooli piirkonna, koostoime?
- Millises vahekorras summeerida eri näitajates tuletatud osa hindetud?

### 6.1. Algandmestikud ja nende süstematiseerimine

Töö käigus ilmnes rida puudujääke andmetes. **Kooli kood ja kooli registreerimisnumber** on liialt sarnased ja tihti andmebaasides segamini aetud. Andmebaasides põhjustab see palju probleeme. Võib olla peaks ühest neist loobuma. Üldiselt ei soovita erinevate koodide ja ID-de osas kasutada numbrilisi sümboleid, mille alguses on 0 (null). Selliseid koode ei saa hoida numbrilisel kujul, sest siis on tõenäoline, et tarkvarades number 0 eest ära. Seega tuleks hoida sellist koodi tekstilisel kujul, mis koormab rohkem mälu. Otsingud on tekstilise tunnuse puhul aeglasemad, mis annab tunda eriti suuremate andmebaaside puhul.

**Erinevused õpilaste registri ja koolide registri andmetes.** Õpilaste arv on registrites on erinev ning õpilaste registris on 4000 õpilast vähem. Õpilaste registris viidatakse aga koolidele, mida koolide registris pole. Enamjaolt oli tegemist täiskasvanute gümnaasiumide ja erikoolidega. Samas tähendab see ka seda, et õpilaste arvu erinevus suureneb koolide registri ja õpilaste registri andmete vahel rohkem kui neljale tuhandele õpilasele viimase kahjuks. Kas see tähendab, et ka õpilaste register on mingis osas puudulik? Küllap on viga mõlemas ja arvatavasti põhjus erinevates andmekogumisallikates, erinevates andmebaasi pidajates ja spetsifikatsioonis. Kas on selgelt paigas, mis on nende eesmärgid?

**Erinevate nimekujude kasutamine ja kirjavigade esinemine** raskendas oluliselt nende geokodeerimist. Probleemi aitaks lahendada näiteks andmebaasis etteantud haldusüksuste nimed ja aadress sisestamine sisestusvormist valikuid tehes. See tagaks andmete ühtse probleemideta kuju ja võrreldavuse.

Teine võimlaus on kasutada näiteks relatsioonilise andmebaasi eeliseid, kus aadresse hoitakse eraldi tabelis ja need seotakse kooliga ID abil. Aadress oleks sellisel juhul ainult ühes kohas ja piisaks selle muutmiseks ühes kohas, et see oleks kaasajastatud või õiges formaadis terves andmebaasis. Nii ei tekiks ka andmete liiasust ja andmebaasi ülekoormamist.

Relatsioonilise andmebaasi pidamisega ei ole otstarbekas ka liiale minna ja loobuda kõigist dubleerivatest andmekogumisviisidest. Mingeid andmeid on soovitatav koguda kontrolli mõt-

tes ka topelt. Nende võrdlemisel välja tulevad erinevused viitavad andmebaasi vigadele ja võimaldavad heita kriitilise pilgu andmekvaliteedile. Näiteks võrreldes õpilaste arvu koolide registri andmete ja õpilaste registri andmete järgi, näitas et 565-st üldhariduskoolist 2003.a. langes õpilaste arv kokku 245 koolis st 320 oli see erinev. Suurim viga küündis, vastavalt Narva 6. Kooli puhul 690 ja Kuressaare Vanalinna Kooli puhul 520 õpilaseni.

Väga oluline on kuidas teha kindlaks andmete puudumine või sisestamata jätmine. Eriti oluline on see numbriliste andmete kogumisel. Kas andmete puudumisel jätta lahter täitmata või kirjutada sinna 0? Kumbki viis pole sobiv, sest automaatsel statistika tegemisel jääksid esimesel juhul andmed statistikast välja ja teisel juhul võetaks näitaja väärtuseks 0. Tulemus on mõlemal juhul vale. Andmete puudumist peaks tähistama mingi väga selgelt eristatav kood, siis selline, mida tunnus ise kunagi omandada ei saa. Erinevate loendamisel saadud andmete puhul (õpilaste arv) kasutatakse näiteks koodi -1, -999 või midagi analoogset. Seeläbi peaks olema terve andmebaasi olema üheselt kasutusel.

Edaspidi võiks lisaks praegu kogutavatele andmetele hakata koguma ka "hea kooli" kriteeriume. Näiteks õpilaste ja õpetajate tervis. Oluline on selliste andmete kogumisel just ruumiline aspekt. Näiteks stressi all kannatavate õpilaste protsent, annab meile küll teatud informatsiooni, kuid on tarvis teada millistes koolides need õpilased käivad. Kui see protsent on kõrge terve Eesti osas, siis võiks ju järeldada, et üldises suunas või juhtimises on midagi korrast ära. Kui stressi all kannatavate õpilaste arv on teistest suhteliselt kõrgem paaris maakonnas, tuleks probleeme otsida nende maakondade tasandilt. Kui aga sellised õpilased jagunevad teatud üksikute koolide vahel üle Eesti, siis tuleks probleemi otsida omavalitsuse või kooli tasandilt. Ruumiga sidumata andmed jätavad meid ilma väga suurest osast informatsioonist.

Täpselt sama moodi annab väga palju hea kooli leidmisele ja probleemsete kohtade otsimisele juurde teiste hea kooli kriteeriumide kogumine ja ruumiline analüüsimine. Näiteks võiks leida ka piirkonnad/koolid, kust tulevad head maksumaksjad, milline on hilisem osalus iibes, kust astuvad ellu rahulolematud ja ka kriminaalsele teele sattuvad inimesed. Väga oluline on nende andmete juures just ruumiinfo, kas siis sidumine kooli, õpilase või õpetajaga, et kätte näidata, mis tasandil tuleb otsida probleemi põhjuseid.

Läbi tuleks mõelda kooli koodide pidamine, senine kord, kas jätkata samas vaimus või tuleks seda siiski muuta. Kui muuta, siis tuleks see rangelt ja korrektselt dokumenteerida, et oleksid võimalikud mõlemas suunas üleminekud (saamaks võrdlust ajaloolise seisuga või võrrelda uusi andmeid vanematega). Kindlasti soovitaks loobuda kahe paralleelse koodi kasutamisest, mis on praegu oma formaadilt liiga sarnased.

## 6.2. Ekspertsüsteemi edasiarendamine ja visioonid

Projekti edasiarendusena näeksime õpetajate kaasamist ekspertsüsteemi. Ekspertsüsteemi tuleks arendada rohkem võimalusi ja dünaamikat pakkuvaks, sest teemasid, mida ruumipõhiselt läbi mängida saab on palju rohkem kui projekti raames vormistatud paar põhimõtetlist näidet. Käesolev projekt osutuski pigem andmete koondamise, nende ristkasutuse võimaldamise ja erinevate teemakaartide koostamiseks, mille käigus ilmnisid probleemsete kohad erinevate teemade lõikes.

Edaspidi peaks lisaks praegu loodud ühele keerulisemale näitele ekspertsüsteemi töös, edasi arendada just erinevate parameetritega koolivõrgu läbimängimise võimalusi. Suurendades-vähendades ajalise kauguse piirväärtusi koolitee pikkuse osas ning luua võimalus olemasolevast koolivõrgust mõne tegutseva kooli elimineerimist, vaadates kuidas see omakorda mõjutaks olemasolevat situatsiooni. Selleks genereeritakse peale parameetrite muutmisi uued teeninduspiirkonnad ja vaadatakse kuidas siis õpilased jagunevad, kui kaugelt õpetajad peavad hakkama koolis käima jne.

Ekspertsüsteemi edasiarendamine toimub kolmes suunas:

### **I. Ekspertsüsteemi arendamine ja täiustamine**

- Ekspertsüsteemi enda ülesannete ja eesmärkide täpsustamine ning laiendamine
- Koolisüsteemi käsitlemise mudeli täiustamine
- Hindamisreeglite ja -eeskirjade täiustamine

### **II. Koondhinnangute ja kooli tulemuslikkuse seoste leidmine**

- Kasutajaskonna laiendamine ja väljaõpetamine koos kasutuse hõlbustamine
- Funktsionaalsuste lisamine ja interaktiivsuse suurendamine
- Kasutajasõbralikkuse tõstmine, et ei nõuaks eri ettevalmistust.

### **III. Andmehõive kvaliteedi ja operatiivse tõstmise**

- Andmekogudega operatiivsete suhete loomine
- Integreerumine Eesti Hariduse Infosüsteemiga ja teiste andmekogude
- Andmete auditeerimine ja eeltöötlemine

Materjalide analüüsil ja ekspertsüsteemi esimese versiooni väljatöötamisel saadud tähelepanekud on järgmised:

- Andmeid oli rohkem ja kättesaadavamad valdkondades, mis puudutasid kooli administreerimist, vähem neid, mis peegeldaksid õpilase ja kooli suhet.
- Koolide andmestik on oodatust vähem kvaliteetne ja kättesaadavus raske.
- Koolile ja seda ümbritsevale pöörati vähem tähelepanu, kui terviksüsteemile.
- Esialguses versioonis on hindamistulemused MapInfos temaatiliste kaartidena. Sel viisi saadud hinnangud on raskesti arvutiga integreeritavad ja statistiliselt töödeldavad, kuid ekspertidele üle vaatesaamiseks sobivad.

Soovitused, mida tuleb edaspidises arendamises arvestada on järgmised:

- Loodav süsteem peab võimaldama hinnangute saamist eri huvigruppide, subjektide seisukohalt eraldi. Nende hinnangute summeerimine ei olegi alati vajalik.
- Leida näitajad ja täiustada statistikat, mis kajastaksid koolide tulemuslikkust ja tagajärjekust eri ajalõigetel (nt semestrite, aasta, kooliastme lõpus ja 2, 6, 20 aasta pärast kooli lõpetamist).
- Ekspertsüsteemi lõplikuks välja töötamiseks on vajalikud regulaarsed rahulolu ja stressiuuringud kooliga seotud osaliste (õppurid, õpetajad, vanemad, administraatorid jt) hulgas.
- Ekspertsüsteemi on vaja integreerida õpilaste ja õpetajate tervise andmed.
- Kooli "ökonomika", perekonna, omavalitsuse, riigi (otsese ja kaudsed) kulutused peaksid olema haaratud ekspertsüsteemi.

## 7. Kokkuvõte

Käesolev töö keskendus praktilisemale suunale – süstematiseeritud ja korrigeeritud andmetele toetudes anti uut, integreeritud teavet Eesti koolivõrgu kohta.

Suurim osa projekti töömahust kulus algandmete kogumisele, nende viibimine põhjustas ka projekti tähtaegade osalise muutmise ja rõhkude ümberasetumise. Ilma algandmetest täielikku ülevaadet omamata ei olnud võimalik sisulist tööd andmebaasi loomisel alustada. Saadud andmete korrastamise ja analüüsi tulemuste põhjal sai tehtud haridusvaldkonna andmekogudele rida soovitusi ja ettepanekuid andmete sisu, struktuuri ja kogumise osas. Kooli, õpilasi ja õpetajaid kirjeldavatele andmetele lisas Regio oma andmekogudest logistika (teede ja nende kvaliteedi info), kooli asukoha ja asustusandmestiku ning kaartide taustainfo. Taustinformatsiooni kasutati teemakaartides, koolide veebirakenduses ja ekspertsüsteemi kasutajaliidese väljundites.

Töö tulemusena valminud andmebaasi põhjal valmistati üle 30 teemakaardi. Viimaste sisuline kvaliteet sõltub andmete kvaliteedist, kuid ka esitusviisist. Teemakaartidega visualiseeriti ja toodi esile olulisem osa väga suurest andmemahust, mida tabelandmeid vaadates on võimatu hoomata. Seetõttu on teemakaartide suur osatähtsus projekti töömahust põhjendatud. Ülevaatliku ja üldistatud pildi saamise kõrval saab kaartidelt üksikasjalikku infot – kaardid on koostatud detailsel alusel ja mitmeplaanilisena. See võimaldab leida detailid ja asjaolud, milles ühe või teise koolipiirkonna eripära seisneb. Analüütiliste teemakaartide komplekt moodustab eksperdi otsustele olulise aluse.

Koolivõrgu ekspertsüsteemi oluline osa on koolide veebirakendus koos valmistatud teemakaartide koguga. Veebirakenduses on koolide andmed üldistamata, leitavad iga kooli kohta eraldi ja vaadeldavad eri detailsusega kaartide taustal. Veebiakenduses on võimalik pärida „oma“ kooli asukohta saada täiendavat teavet koolist, mis on justkui kontrollmehhanism, sest „oma“ kooli kohta oskavad kõik kaasa rääkida. Veebirakendus on keskkond, kus on kättesaadavaks tehtud oluline koolivõrgu iseloomustav andmestik. Seda keskkonda saab täiendada soovitud suundades, näiteks uuendades koolide andmeid ja lisades uusi teemakaarte.

Ekspertsüsteemi olulisest osast, kasutajaliidese osast on valminud MapInfo kaardiprogrammile tuginev keskkond ja dialoogi akende süsteem. See võimaldab opereerida andmehulkadega kool, õpilane ja kooli tagamaa (koolikeskkond). Süsteem võimaldab mitmeid kooli ja koolipiirkonda iseloomustavaid andmeid kombineerida saades võrreldavaid hinnanguid (pallides). Kasutajaliidese abil saab tekitada olulisematest kooli iseloomustavatest andmetest erinevate reeglite ja mudelite abil uusi teadmisi ja teavet. Uus informatsioon võib olla tabeli, teemakaardi või teksti näol. Valminud ekspertsüsteemi kasutajaliides annab võimaluse erinevaid variante läbi mängida ja proovida. Ekspertsüsteemi oluliseks osaks on hinnangute andmise (analüüsi) eeskirjad ja tabelid ning koondhinnangute (sünteesi) eeskirjad, mis koostati koos Poliitikauuringute Keskusega Praxis ekspertidega.

Väljatöötatud lahendus ja korrastatud andmebaas annab võimaluse tulemuse kiireks levitamiseks/presenteerimiseks, sealhulgas veebipõhiste rakenduste või väljatrükkide näol. Projekti tulemusi rakendades saab tulevikus veelgi veelgi vähendada ekspertide rutiinset tegevust. Valminud ekspertsüsteemi komponendid võimaldavad arendada tulemuslikku dialoogi koolitemaatika ekspertidega, et saada ühistööna vahend motiveeritud ja põhjendatud ettepanekute tegemiseks otsustajate tarvis.