



**Riikliku programmi „Eesti keele keeletehnoloogiline tugi (2006–2010)“  
sihtevalveerimine**

**Evalveerimisraport**

**Tartu 2011**

# Riikliku programmi „Eesti keele keeletehnoloogiline tugi (2006–2010)“ sihtevalveerimine

## Evalveerimisraport

### Sisukord

1. Sissejuhatus.....	3
2. Evalveerimise protsess.....	4
3. Kokkuvõte .....	5
4. Programm tervikuna .....	7
5. Programmi raames rahastatud projektid.....	10
5.1. Projektide taotlemine .....	10
5.2. Projektide kestvus ja rahastamine .....	10
5.3. Projektide kulud olulisemate kuluartiklite lõikes.....	12
5.4. Projektide täitmise kaasatud personal .....	12
5.5. Projektide tulemuslikkus.....	13
6. Programmi juhtkomitee.....	16
7. Järeldused, soovitusel ja ettepanekud .....	17
8. Lisad .....	18

# 1. Sissejuhatus

Riikliku programmi “Eesti keele keeletehnoloogiline tugi (2006–2010)” (edaspidi RP EKKTT) (programmi teksti vt lisa 1) peaesmärgiks oli eesti keele keeletehnoloogilise toe arendamine tasemele, mis võimaldab eesti keelel edukalt toimida tänapäeva infotehnoloogilises keskkonnas. Keeletehnoloogia on infotehnoloogiat ja keeleteadust ühendav interdistsiplinaarne valdkond, mis tegeleb vahendite (meetodite, algoritmide, programmide) väljatöötamisega nii kirjutatud kui ka suulise keele arvutitöötuseks. Kuna eesti keele kõnelejate arv on liiga väike selleks, et keeletehnoloogiat arendavatel firmadel tekiks huvi eesti keelele sobivate vahendite väljatöötamise vastu, on keeletehnoloogia arendamiseks paratamatult vaja riigi toetust.

Keeletehnoloogia kuulub Eesti teadus- ja arendustegevuse strateegia “Teadmistepõhine Eesti” võtmevaldkonna “Kasutajasõbralikud infotehnoloogiad ja infoühiskonna areng” alla. Keeletehnoloogia oli ka Vabariigi Valitsuse heakskiidu leidnud “Eesti keele arendamise strateegia (2004–2010)” võtmevaldkond. Keeletehnoloogiline tugi koosneb keeleressurssidest ja keeletarkvarast ning viimase rakendustest. Keeleressursid on elektroonilised andmekogud, mida kasutatakse ka keeletarkvara väljatöötamiseks. Keeletarkvara koondab endas meetodeid, algoritme ja programme keelematerjali töötlemiseks ja rakendussüsteemide väljatöötamiseks. Keeletehnoloogia ülesandeks on luua ressursid ja vahendid, mis tagavad eesti keelele võrdsed võimalused infotehnoloogilises keskkonnas suhtlemiseks.

RP EKKTT on välja kasvanud riiklikust programmist “Eesti keel ja rahvuslik mälu (2004–2008)”, mis sisaldas keeletehnoloogia moodulit. EKKTT rahastas keeletehnoloogiaalast teadus- ja arendustegevust alates ressurside loomisest kuni keeletehnoloogiliste rakenduste (prototüüpide) loomiseni. Programmi eesmärgiks on, et selle tulemused jõuaksid e-ühiskonnas kasutusse. E-riik ja e-ühiskond on programmi tulemi kasutajad kõige laiemas mõttes. Programmi tulemused on mõeldud kasutamiseks e-valitsemises, e-hariduses, e-tervishoius, e-kaubanduses, e-panganduses, e-teeninduses, e-kirjastamises, e-(mobiil)sides, e-suhtluses, e-infopäringutes jne. Sellise e-ühiskonna loomiseks on vaja abivahendeid tekstitöötusel ning kõnetehnoloogia viimist uuele tasemele.

Keeletehnoloogiline tugi on vajalik infoühiskonna arendamiseks. See annab oma panuse meid ümbritseva tehisintellekti e-keskkonna loomisesse aastaks 2020. Tänu keeletehnoloogilisele toele toimib see keskkond eestikeelsena. Arvuti ja väliskeskkonna vahetu interaktsioon ja ühtesulamine toimub eesti keele ja masintõlge vahendusel. Meie visioon eeldab, et mitmesuunaline masintõlge toimib nii suuliste kui ka kirjalike tekstide puhul. Tänu kõneleja automaattuvastusele saab kõiki e-toiminguid teha suulist kõnet kasutades. Tõlkeprogrammid võimaldavad meil kasutada elektroonilisi ressursse paljudes keeltes.

Paljud maailma riigid on mõistnud, et oma keelilise ja kultuurilise identiteedi säilitamiseks infoühiskonnas on vaja välja töötada rahvuskeelte keeletehnoloogiline tugi ning seda arendada. Keeletehnoloogia on üks Euroopa Liidu prioriteete, mida toetatakse mitme programmi kaudu. Eesti jaoks on samasugune riiklik programm vajalik nüüd, mil kuulumine Euroopa Liitu nõuab meilt nüüdisaegsete keeletehnoloogiliste vahendite olemasolu samal tasemel teiste riikidega.

Eestis muutus keele- ja kõnetehnoloogia areng intensiivseks 1990. aastate keskpaigast alates, mil käivitus Euroopa Liidu COPERNICUS-programm, samuti on Eesti osalenud nii 4. kui ka 5. raamprogrammi keeletehnoloogia projektides. Vajalike spetsialistide ettevalmistuse tagab Tartu Ülikooli arvutilingvistika ja keeletehnoloogia alane kraadiõpe ning Tartu Ülikooli poolt koos Eesti Keele Instituudi ja Tallinna Tehnikaülikooli Küberneetika Instituudiga käivitatud doktorikool “Keeleteadus ja -tehnoloogia”. Spetsialistide ettevalmistusele annab uue mõõtme liitumine Põhjamaade eespool nimetatud keeletehnoloogia doktorikoolide võrgustikuga.

## 2. Evalveerimise protsess

EKKTT sihtevalveerimise viis läbi SA Archimedese juures tegutsev Teaduskompetentsi Nõukogu büroo. Evalveerimine toimus 2011. aasta kevadel ja see koosnes kolmest etapist:

- 1) programmiga seonduva materjali kogumine ja analüüsimine;
- 2) intervjuud programmi juhtkomitee liikmetega;
- 3) evalveerimisraporti koostamine.

Sihtevalveerimise käigus töötati läbi järgmised programmiga seotud materjalid:

- programmi kirjeldus;
- programmi eel- ja jätkuprogrammide kirjeldused;
- programmi raames rahastatud projektide taotlused ja lõpparuanded;
- programmi juhtkomitee koosolekute protokollid ning tegevusaruanded;
- programmi korraldusliku poolega seotud abimaterjalid;
- programmi kokkuvõtavad materjalid.

Peale selle intervjueriti programmi juhtkomitee liikmeid, et saada rohkem tagasisidet. Programmi juhtkomitee, sh intervjueritud liikmete nimekiri on toodud lisas 4.

EKKTT sihtevalveerimise metoodika oli sarnane 2010.a läbi viidud riiklike programmide „Eesti keel ja rahvuslik mälu (2004-2008)“ ning „Humanitaar- ja looduteaduslikud kogud (2004-2008)“ sihtevalveerimisel kasutatud metoodikaga. Et kõigil sihtevalveeritud programmidel oli ka sarnane ülesehitus, siis ka osad EKKTT sihtevalveerimise järeldused, soovitusel ja ettepanekud langevad kokku varasemate sihtevalveerimise käigus tehtutega.

### 3. Kokkuvõte

RP EKKTT käivitati Haridus- ja Teadusministeeriumi poolt 2006. aastal ja see kestis 2010. aastani (programmi teksti vt lisast 1). EKKTT kasvas välja riiklikust programmist „Eesti keel ja rahvuslik mälu (2004-2008)“ (EKRM), mis sisaldas keeletehnoloogia moodulit alajaotustega „Keeleressursid“ ja „Tarkvara rakendusala“. Et osa EKKTT raames rahastatud projekte olid töös juba EKRM-i ajal aastatel 2004-2005, siis need projektid, mis esitati samal kujul ka EKKTT programmi, said EKKTT raames rahastuse EKRM programmis määratud ajaks, st valdavalt 2008.a lõpuni. Nende projektide puhul, mille fookus muutus või tegevused laienesid, algas programmiperiood otsast peale ehk nende nn. uuenenud projektide pikkus oli valdavalt EKKTT lõpuni, st kuni 7 aastat.

EKKTT peaesmärgiks oli eesti keele keeletehnoloogilise toe arendamine tasemele, mis võimaldaks eesti keelel edukalt toimida tänapäeva infotehnoloogilises keskkonnas. Programmi tulemuslikkuse kriteeriumid ja oodatavad tulemused olid programmi tekstis algselt detailselt kirjeldatud – kokku määratleti programmdokumendis 30 alaeesmärki ja oodatavat tulemust. Et kokku rahastati programmi raames 33 projekti, siis sisuliselt olid kõik alaeesmärgid projektidega kaetud. Projektide tulemuslikkust hinnanud juhtkomitee hinnangul täideti 84% projektidest vähemalt rahuldavalt. Sihtevalveerimise läbiviija hinnangul täideti 79% projektide puhul püstitatud eesmärgid vähemalt osaliselt.

**Siit tulenevalt võib väita, et 4/5 projektide tulemustega võib üldjoontes rahule jääda.**

Ühe eesmärgina oli välja toodud programmi raames rahastatud projektide tulemuste avalik kättesaadavus. Kõik EKKTT raames rahastatud projektid (kirjeldus, metoodika, tulemused) on koondatud programmi veebilehele internetis ([www.keeletehnoloogia.ee](http://www.keeletehnoloogia.ee)) ning **3/4 projektide tulemustest on sealtkaudu ka avalikult kättesaadavad.**

Eesti keele arendamise strateegia 2009.a valminud vaheseire aruande (vt lisa 9) põhjal on töö keeletehnoloogilise toe loomisel olnud edukas ja EKKTT tegevused kulgenud plaanipäraselt ja tulemuslikult.

Juhtkomitee intervjueritud liikmete sõnul oli programm vajalik ja vastas seatud eesmärkidele. Olulisemate positiivsete saavutustena toodi välja:

- valdkonna stabiilse finantseerimise tagamine;
- programmist saadud lisaraha efektiivne kasutus (ka Euroopa kontekstis);
- kõigis planeeritud alavaldkondades toimus areng;
- programm võimaldas uute valdkonnaga seotud inimeste väljaõpet;
- programmi tulemite rist- ja avalik kasutus.

Puudustena toodi välja eelkõige juriidilised, sh andmekaitse ja intellektuaalse omandi kuuluvusega seotud küsimused, tulemuste mittepiisav rakendatavus lõppkasutaja seisukohast ning asutustevaheliste koostööprojektide võimaluste puudumine.

Programmi kui terviku seisukohalt võib välja tuua kaks suurt olulist lahendamist vajavat probleemide ringi:

**1. Programmi disainiga seonduvad küsimused:**

- a. Programmide raames rahastatavad tegevused peavad toetama programmi edukaks realiseerimiseks vajalikke kriitilisi protsesse. Ehkki EKKTT põhieesmärgiks oli programmi tulemuste jõudmine kasutusse e-ühiskonnas, rahastati selle raames teadus- ja arendustööd kuni prototüüpide loomiseni. Siia oli sisse programmeeritud vastuolu programmi eesmärkide ja rahastatavate tegevuste vahel, kuna katmata oli äärmiselt oluline lõik esialgsest prototüübist kuni tootmis- ja lõppkasutajale kasutuskõlbliku tooteni.
- b. Programmi juhtkomitee peaks tegelema programmi sisulise juhtimisega (sh rahastamismudeli rakendamisega) ning vastutama programmi eesmärkide saavutamise eest, projektitoetusi peaks hindama ekspertkomisjon. Kuigi programmi juhtkomitee ülesanded, sh analüüs ja vastutus, olid programmi tekstis (p 1.3) ka välja toodud, piirdus juhtkomitee põhiliselt hindamiskomisjoni rolliga.

**2. Programmivälised tegurid, mis pärsivad programmi tulemuste saavutamist, sh:**

- a. Intellektuaalomandiga seonduvad küsimused, mis puudutavad erinevate autorikaitse objektideks olevate andmebaaside kasutamist arendustöös, vabavarana kasutusel olevate programmiplokkide kasutamist kommertsrakendustes ning intellektuaalse omandi kuuluvuse küsimusi, kui üheks huvitatud osapooleks (realiseerijaks) on äriettevõtte.
- b. Programmi põhitulemuste saavutamist toetavad tegevused, mida ei rahastata programmi raames: loodud on tarkvaraprototüüp, mis vajab turule jõudmiseks ja seal püsima jäämiseks ning konkurentsis püsimiseks edasist jätkuvat arendustegevust ja täiendavat toetust. Seda tüüpi tegevused olid EKKTT puhul katmata.

Konkreetsed järeldused, soovitused ja ettepanekud on toodud iga peatüki järel, punktis 7 on toodud üldistatud, süstematiseeritud ning olulisemad soovitused ja ettepanekud.

Aruande koostasid SA Archimedes teadusevalvatsiooni peaspetsialist Viktor Muuli ja teaduskompetentsi nõukogu büroo juhataja Madis Saluveer. Eesti teadusinfosüsteemi päringud koostas Eesti teadusinfosüsteemi ETIS konsultant Priit Tuvike. Sihtevalveerimise lähteülesanne on toodud lisan 2.

## 4. Programm tervikuna

Keeletehnoloogia on infotehnoloogiat ja keeleteadust ühendav interdistsiplinaarne valdkond, mis tegeleb vahendite (meetodite, algoritmide, programmide) väljatöötamisega nii kirjutatud kui suulise keele arvutitöötamiseks. RP EKKTT käivitati Haridus- ja Teadusministeeriumi poolt 2006. aastal ja see kestis 2010. aastani. EKKTT kasvas välja riiklikust programmist „Eesti keel ja rahvuslik mälu (2004-2008)“, mis sisaldas keeletehnoloogia moodulit alajaotustega „Keeleressursid“ ja „Tarkvara rakendusala“. EKKTT sisaldas kahte alamoodulit:

- Keeletarkvara, mille all mõisteti meetodeid, algoritme ja arvutiprogramme keelematerjali töötlemiseks, mis EKKTT raames oli kavandatud eesti keele keeletehnoloogilise toe loomiseks ja vajalike rakendussüsteemide väljatöötamiseks. Selleks planeeriti kohandada eesti keelele mitmeid keelest sõltumatuid keeletöötlusmeetmeid ja –algoritme ning luua uusi ainult eesti keele spetsiifikast lähtuvaid lahendusi.
- Keeletehnoloogilised ressursid, mille all mõisteti elektroonilisi andmekogusid, mida kasutatakse keeletarkvara väljatöötamiseks: korpused (kõnesignaali ja tekstide kogumid), elektroonilised sõnastikud ja andmebaasid, formaalsed keelekirjeldused (grammatikad). Keeleressursid on keelespetsiifilised ja neid ei saa teistest keeltest üle võtta.

EKKTT peaeesmärgiks oli eesti keele keeletehnoloogilise toe arendamine tasemele, mis võimaldaks eesti keelel edukalt toimida tänapäeva infotehnoloogilises keskkonnas. Ehk: kuigi programm rahastas otseselt teadus- ja arendustööd kuni prototüüpide loomiseni, oli selle eesmärgiks siiski programmi tulemuste jõudmine kasutusse e-ühiskonnas. Siia oli sisse programmeeritud vastuolu programmi eesmärkide ja rahastatavate tegevuste vahel, kuna katmata oli äärmiselt oluline lõik esialgsest prototüübist kuni tootmis- ja lõppkasutajale kasutuskõlbliku tooteni.

Programmi valmistas ette haridus- ja teadusministeeriumi töörühm, programmi põhitäitjateks olid Eesti Keele Instituut, Tallinna Tehnikaülikool ning Tartu Ülikool.

Programmi alaeesmärgid on programmi tekstis detailselt lahti kirjutatud – kokku oodati konkreetseid tulemusi ligi 30 tegevusvaldkonnas. Selliselt põhjalikult lahtikirjutatud programmi alaeesmärkide ja oodatavate tulemuste kogum võimaldab pärast programmi lõppemist püstitatud eesmärkide saavutamist suhteliselt hõlpsasti hinnata. Samas võib olukorda, mille puhul ühele püstitatud alaeesmärgile vastab sisuliselt üks projekt (kokku rahastati programmi raames 33 projekti), tõlgendada kui eesmärkide püstitamist Eestis antud valdkonnas tegutsevate uurimisgruppide põhiselt ning programmi koostamist lähtuvalt ka nende huvidest ja kompetentsist, mitte niivõrd riigi- ja ühiskonnapoolsetest reaalistest vajadustest. Samuti tähendab see suurte eesmärkide lahustumist paljudes väikestes alaeesmärkides.

**Kokkuvõtteks: kui tööd keeletehnoloogiliste ressurssidega võib programmi raames lugeda kordaläinuks, siis lõppkasutajale suunatud keelerakvara loomine oli selgelt vähem tulemuslik.**

EKKTT-ga kattuvaid programme teiste ministeeriumite haldusalas ei olnud. Et programm rahastas teadus- ja arendustööd kuni prototüüpide loomiseni, sh ka alusuuringuid, siis võis täheldada läbiviidavate projektide temaatika kattumist samal ajal töös olnud sihtfinantseeritavate teadusteemade ja SA Eesti Teadusfondi grantide temaatikaga. Juhtkomitees olid enamuses teadlaste esindajad (vt 6. punkt), mistõttu (tõenäoliselt tahtmatult) sai suurema osakaalu projektide teaduslik pool, rakenduslikule küljele ja selle piisavale arvestamisele hindamiskriteeriumina pöörati vastavalt vähem tähelepanu.

RP EKKTT jätkuprogramm on riiklik programm „Eesti keeletehnoloogia (2011-2017)“, mille raames lisaks tarkvaraprototüüpide ja keeleressursside arendamisele pööratakse suurt tähelepanu keeletehnoloogia rakenduste loomisele ja olemasolevate ning loodavate ressursside ja tarkvara kättesaadavaks tegemisele.

Programmi raames käivitatud Eesti keeleressursside keskus on Eesti teaduse infrastruktuuri teekaardi (vt lisa 7) objektiks. Eesti keeleressursside keskus on infrastruktuur, mis võimaldab kõigile uurijatele keeleressursside ja tehnoloogiate (sh kõnetehnoloogia andmebaasid, emotsionaalse kõne korpus, eesti-X-keele sõnastiku andmebaas, leksikograafi töökeskkond, emotsionaalse keele korpus, kõnetuvastaja, grammatikakontrollija, sisukokkuvõtete koostaja, eesti-inglise masintõlge jm) kättesaadavuse. Selleks ühendatakse eksisteerivad digitaalsed arhiivid ja pakutakse keeletehnoloogia vahendeid kui veebiteenust, mis kasutab arhiveeritud andmeid. Teekaart on pikaajaline (10–20 aasta perspektiiviga) planeerimisvahend, mis sisaldab loetelu uutest või ajakohastamist vajavatest riiklikult olulistest teaduse infrastruktuuriüksustest. Infrastruktuuri lülitamine teekaardi koosseisu ei tähenda rahastamisotsust ega pane objekte tähtsuse järgi järjekorda, kuid teekaart on sisend ettevalmistamisel olevatele rahastamisotsustele.

Samuti kuulub Eesti keeleressursside keskus ESFRI infrastruktuuriobjekti CLARIN (Ühine keeleressursside ja -tehnoloogia infrastruktuur) koosseisu (vt lisa 7). CLARIN on suureulatuslik üle-Euroopaline koostööinitsiatiiv, millega luuakse, koordineeritakse ja tehakse keeleressursid ja -tehnoloogia kättesaadavaks ning täiel määral kasutatavaks. CLARIN pakub humanitaar- ja sotsiaalvaldkonna teadlastele arvutivõrgustik-tüüpi infrastruktuuri abil vahendeid, mis võimaldavad kasutada arvutipõhiseid keeletehnoloogiaid, et uurida keele erinevaid rolle ühiskonnas (nt kultuuri ja teadmiste kandjana, suhtlemisel, identiteediosana jt). CLARIN baseerub hulgal võtmetehnoloogiatel, mis on peamisteks alusteks e-Teaduse paradigmas: võrgustikutehnoloogiad, mille abil ühendatakse andmebaase ja repositooriume ning veebiteenuseid, mida erinevad keskused pakuvad; digitaalraamatukogud ja reaalaja-arhiivid; semantilised veebitehnoloogiad, millega ületatakse struktuurilised ja semantilised kodeerimisprobleemid; arenenud multilingvistilised keeletehnoloogiad, mis toetavad kultuurilist ja keelelist integratsiooni. Vt lisa: [www.clarin.eu](http://www.clarin.eu)

Positiivsena võib välja tuua, et programmi raames rahastatud projektide loetelu, kirjeldus ja tulemused on koondatud veebilehele [www.keeletehnoloogia.ee](http://www.keeletehnoloogia.ee). Samas paikneb info programmi kohta tervikuna ning kokkuvõtte programmi kolmest aruandekonverentsist, mis toimusid 2007., 2009. ja 2010.a aastal.

#### Järeldused ja soovitused.

- EKKTT fookus ja alaeesmärgid olid tervikuna laialivalguvad ning programmi raames rahastatavad tegevused kattusid osaliselt teiste Haridus- ja Teadusministeeriumi rahastamisinstrumentide raames toetatavate tegevustega. Programmid ja nende raames rahastatavad projektid peaksid olema selgelt fookuseeritud ning üldeesmärgid ja tulemuslikkuse indikaatorid peaksid olema välja toodud. Tulevaste programmide koostamisel peaks keskenduma senisest enam üldistele, ent olulisematele eesmärkidele ja vältima tegevuste topeltrahastamise võimalikkust.
- Programmide raames rahastatavad tegevused peavad toetama programmi edukaks realiseerimiseks vajalikke kriitilisi protsesse. Ehkki EKKTT põhieesmärgiks oli programmi tulemuste jõudmine kasutusse e-ühiskonnas, rahastati selle raames teadus- ja arendustööd kuni prototüüpide loomiseni. Siia oli sisse programmeeritud vastuolu programmi eesmärkide ja rahastatavate tegevuste vahel, kuna katmata oli äärmiselt oluline lõik esialgselt prototüübist kuni tootmis- ja lõppkasutajale kasutuskõlbliku tooteni.



- Lahendamist vajavad programmide eesmärkide saavutamist toetavad programmivälised tegevused, sh intellektuaalse omandiga seonduvad küsimused.
- Ülaltoodust lähtuvalt tuleks edaspidi senisest suuremat tähelepanu pöörata programmide väljatöötamisele, kaasates programme ettevalmistavate töögruppide koosseisu rohkem erinevate huvigruppide esindajaid.

## 5. Programmi raames rahastatud projektid

### 5.1. Projektide taotlemine

Projekte oli võimalik taotleda programmi toimumise ajal (2006–2010) igal aastal. Projektikonkursid kuulutas välja Haridus- ja Teadusministeerium, projektitaotluste hindamise viis läbi programmi juhtkomitee ning konkursside tulemused kinnitati juhtkomitee ettepanekul haridus- ja teadusministri käskkirjaga kas märtsis või aprillis (üks kord ka juunis ja septembris).

Projektide sellise menetlustsükli juures võib välja tuua kaks probleemi:

- Projektitaotluste ja -aruannete koostamine ning nende menetlemine igal aastal on liiga ressursimahukas.
- Projektide rahastamistulemuste kinnitamine alles rahastatava aasta märtsis-aprillis (või isegi veel hiljem) jätab asutused vähemalt esimeseks kvartaliks teadmatusse, kas ja mis mahus nad saavad oma plaane alanud aastal realiseerida.

Mõlema probleemi lahenduseks paremini sobiva taotluskorralduse loomine on põhimõtteliselt Haridus- ja Teadusministeeriumi pädevuses.

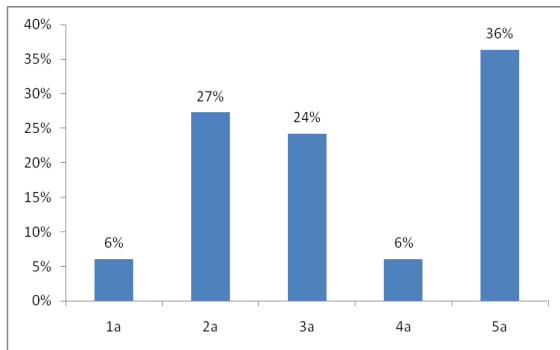
Kuigi projektid, ka erinevate asutuste poolt läbi viidavad, olid omavahel sageli loogiliselt, sisuliselt ja järjestikuliselt seotud, siis projektide taotlemine toimus asutusepõhiselt, formaalselt isegi puudus võimalus taotleda asutustevahelisi koostööprojekte.

#### Järeldused ja soovitused.

- Tuua projektide taotlemistsükkel rahastamisele eelneva aasta lõppu ning nende rahastamis- ja aruandlusperioodi pikendada kahe-kolme aastani.
- Vajalik on luua tingimused ja võimalused asutustevaheliste koostööprojektide taotlemiseks ja elluviimiseks.

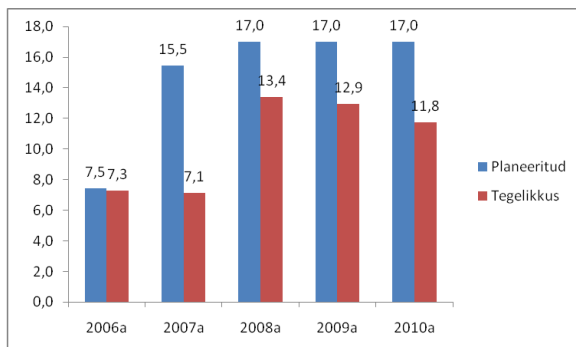
### 5.2. Projektide kestvus ja rahastamine

Programmi raames rahastati kokku 33 projekti. Projektide keskmine pikkus oli 3,4 aastat, suurima osa projektidest moodustasid kogu programmiperioodi, st 5 a kestvad projektid. 2 ja 3 aastat kestvate projektide suhteliselt suurt ja sarnast osakaalu saab seletada EKRM programmist ületulevate projektidega, mida rahastati EKKTT raames reeglina kolm aastat ja programmi raames oligi võimalik nende lõppedes 2008.a realiseerida veel vaid 2 a kestvaid projekte. Lühikese programmiperioodiga saab seletada ka 4 ja 1 a kestvate projektide väikest osakaalu projektide koguarvust (vt joonis 1) ja mille tekkimiseks puudusid sisuliselt võimalused.

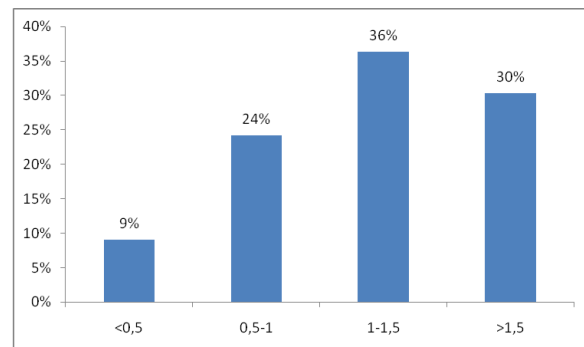


**Joonis 1.** EKKTT raames rahastatavate projektide kestvus.

Programmi raames rahastati projekte kogumahas 52,5 miljonit krooni. Programmi käivitamisel planeeriti programmi rahastamist kogumahas 73,9 miljonit krooni, sh infrastruktuuri ajakohastamist mahus 20 miljonit krooni (sh 2007.a – 8 milj kr, 2008.a – 6 milj kr, 2009.a – 4 milj kr ja 2010.a – 4 milj kr). Infrastruktuuri kaasajastamisplaaniga on seletatav ka planeeritud rahastamise suur tõus 2007.a. Programmi rahastamist planeeriti riigieelarvest, infrastruktuuri kaasajastamist planeeriti taotleja tõekefondidest. Programmi raames elluviidavate projektide tegelikuks rahastamismahuks kujunes 52,5 miljonit krooni, sh infrastruktuuri kaasajastamiseks 2,1 miljonit krooni. Rahastamise vähenemise peamiseks põhjuseks oli tõekefondide kasutamise reeglistiku muutused, mis ei võimaldanud programmi raames kasutada tõekefondide vahendeid, riigipoolne rahastamisplaan täideti aga isegi üldise majanduslanguse tingimustes (vt joonis 2).



**Joonis 2.** EKKTT planeeritud ja tegelik rahastamine aastate lõikes (mln kr).



**Joonis 3.** EKKTT raames rahastatud projektide osakaal rahastamismahu mahu alusel protsentuaalselt (summad toodud mln kr).

Programmi raames rahastatud projekti keskmine maht oli ca 1,6 miljonit krooni, alla 0,5 miljoni kroonise mahuga projektide osakaal oli 9%, üle 1 miljoni kroonise mahuga projekte oli 66% projektide üldarvust (vt joonis 3).

Programmi raames rahastatud projektide nimekiri koos detailsemate andmetega on toodud lisa 3.

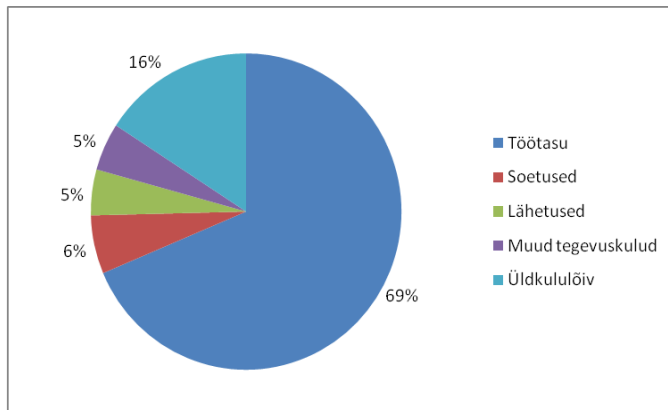
#### Järeldused ja soovitused.

- Programmide rahastamisdünaamika oli hea, selle üldmahtu suudeti hoida programmi lõpuni enam-vähem stabiilsena vaatamata üldisele majanduslangusele. Rahastamise stabiilsuse tagamine on kindlasti ka üks olulisi eeldusi tulevaste programmide puhul.
- Positiivne on rahastatud projektide ajalise kestvuse omavaheline suhe - 2/3 rahastatud projektidest kestsid vähemalt 3 aastat.

- Positiivne on ka rahastatud projektide mahu omavaheline suhe - 2/3 rahastatud projektide maht oli vähemalt 1 miljon krooni.

### 5.3. Projektide kulud olulisemate kuluartiklite lõikes

EKKTT raames rahastatud projektide kuluartiklid olid taotlustes ja aruandluses jaotatud järgmiselt: töötasu, soetused, lähetused, muud tegevuskulud, asutuse üldkulud. Samade kuluartiklite lõikes analüüsiti projektide kulusid ka sihtevalveerimise raames (vt lisa 8).



**Joonis 4.** EKKTT programmi vahendite keskmine jaotus projektide kuluartiklite lõikes.

Suurimaks kuluartikliks oli töötasu, mis moodustas 69% kogu kuludest (82% efektiivsetest kuludest, st kuludest ilma üldkululõivuta), sellele järgnes üldkululõiv 16%-ga. Ülejäänud kuluartiklid (soetused, lähetused ja muud tegevuskulud) olid oluliselt väiksemad ja moodustasid igaüks 5-6%. Välja võib tuua, et 2008. ja 2009.a eraldi soetusteks määratud summadel ei olnud asutustel lubatud üldkululõivu võtta. Lisaks kaasrahastasid asutused ETIS-e andmetel projekte kogumahus ca 1,4 miljonit krooni, mis teeb programmipoolseks keskmiseks rahastuse määraks 97%. Programmipoolne finantseerimine töötasude puhul moodustas kogufinantseerimisest 97%, soetuste puhul 99%, lähetuste puhul 97%, muude tegevuskulude puhul 99% ja üldkulude puhul 98%.

Lähtudes nii projektide põhitäitjate kui tulemuste kattuvusest erinevate rahastamisinstrumentide lõikes, võib tõstatada küsimuse üldkululõivumäära rakendamise otstarbekusest nii suures määras, kuna see vähendab oluliselt spetsiifilisi eesmärke teenivate programmide efektiivse rahastamise mahtu.

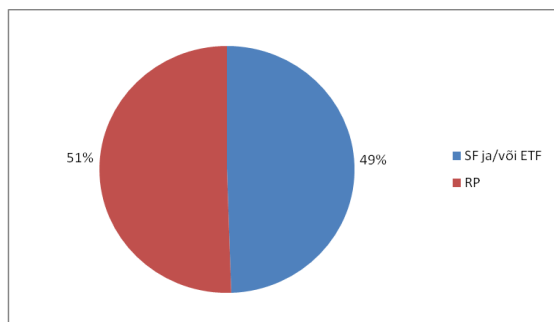
#### Järeldused ja soovitused.

- Abikõlblikud kuluartiklid ning nende määrad ja omavahelised proportsioonid olid piiritlemata, lõviosa programmi kuludest moodustasid tööjõukulud (69% programmi kogumaksumusest, 82% programmi efektiivsest maksumusest, st ilma asutuste üldkululõivuta). Raporti koostajates tekitas küsimusi üldkulude suur osakaal programmides, kuna suures osas kasutati teiste rahastamisinstrumentidega samu ressursse, mille eest on üldkululõivu juba võetud.
- Kaaluda valdavalt samade inim- ja infrastruktuuriressursside abil tehtava teadus- ja arendustöö osas üldkululõivu määra vähendamist spetsiifiliste programmide raames.

### 5.4. Projektide täitmisse kaasatud personal

EKKTT puhul osales projektide täitmisel lõpparuannete põhjal kokku 160 põhitäitjat, mis annab ühe projekti põhitäitjate arvuks keskmiselt 4,9. Põhitäitjate summaarne täistööaja ekvivalent (inglisekeelne

termin *full-time equivalent* (FTE)) oli 49,8, st keskmiselt 1,5 projekti kohta. Abitööjõuna kaasati 172 inimest, mis teeb keskmiselt 5,2 inimest (1,8 FTE) summaarse keskmise tööajaga 47,5 kuud projekti kohta.



**Joonis 5.** EKKTT raames rahastatud projektide põhitäitjate kattuvus sihtfinantseeritavate teadusteemade ja ETF-i grantide põhitäitjatega protsentides 2010.a ETISe andmetel. Joonisel: RP – riiklike programmide põhitäitjate osakaal, kes ei osale kas sihtfinantseeritavates teadusteemades või ETF-i grantides põhitäitjatena; SF ja/või ETF – riiklike programmide põhitäitjate osakaal, kes osalevad kas sihtfinantseeritavates teadusteemades või ETF-i grantides põhitäitjatena.

Projektidega seotud personal oli tihedalt seotud asutuses tehtava muu valdkondliku teadus- ja arendustegevusega – praktiliselt pooled põhitäitjatest olid seotud kas sihtfinantseeritava teadusteemaga või ETF-i grandiga samuti põhitäitja tasandil, 16% riiklike programmide projektide põhitäitjatest oli 2010. aastal tegevad nii SF-i kui ka ETF-i põhitäitjana. Detailne ülevaade põhitäitjate seotusest teiste projektidega 2010. aastal on toodud lisas 5.

#### Järeldused ja soovitused.

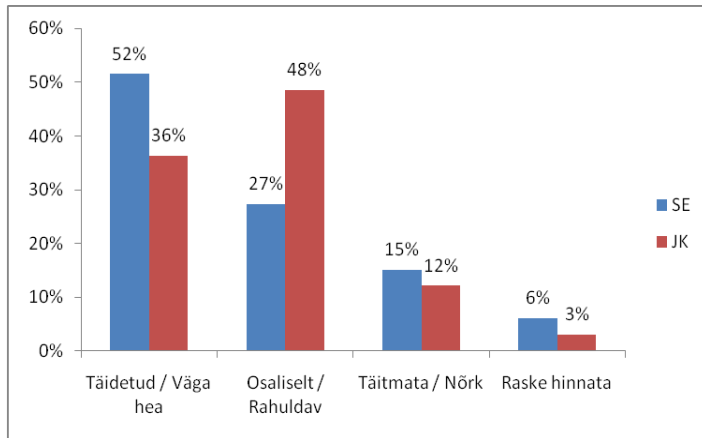
- Erinevate teadus- ja arendustegevuse projektide täitmise kaasatud personal kattub paljuski erinevate rahastamisinstrumentide lõikes. Rahastamisinstrumentide paljusus Eesti-suguses väikeriigis piiratud inimressurssi tingimustes killustab vahendeid, loob täiendavat bürokraatiat ning muudab konkreetsete projektide jälgimise ja hindamise keeruliseks.

### **5.5. Projektide tulemuslikkus**

Programmide sihtvalveerimise käigus analüüsiti lõpparuannetes kajastatud info ja selle piisavuse vastavust aruandlusvormile Eesti teaduse infosüsteemis ETIS ning võrreldi projektide raames tehtud lõpparuannetes sisalduva info alusel taotlustes planeeritud tegevustega.

Lõpparuannetes kajastatud info ja selle piisavuse analüüsi tulemusena tuli välja, et kõikide aruandekohustuslike projektide aruanded vastasid kehtestatud nõuetele.

Konkreetsete projektide tulemuslikkuse hindamiseks võrreldi projektitaotlustes püstitatud eesmärke lõpparuannetes kirjeldatud tulemustega. Juhtkomitee pidas 36% projektide täitmist heaks/väga heaks, 48% rahuldavaks/osaliselt täidetuks, 3% puhul polnud võimalik projekti edukust hinnata, 12% projektide tulemuslikkust hindas juhtkomitee nõrgaks (vt joonis 6).



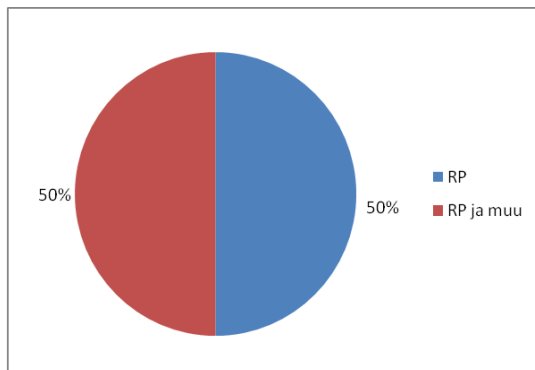
**Joonis 6.** Juhtkomitee (JK) ja evalveerimise läbiviija (SE) ühtlustatud hinnangud projektide tulemuslikkusele.

Evalveerimise läbiviija poolt antud hinnangu alusel täideti taotluses püstitatud eesmärgid täielikult või valdavalt 52% projektide puhul, vähemalt osaliselt 27% projektide puhul, 15% projektide puhul olid eesmärgid valdavalt saavutamata ja 6% projektide puhul oli projektide tulemuslikkusele lõpparuannete põhjal äärmiselt raske hinnangut anda, kuna eesmärgid ja tulemused olid kirjeldatud äärmiselt erinevalt (vt joonis 6).

Olulisemad tegurid, mis põhjustasid projektitaotluses toodud eesmärkide ja aruandluses kajastatud tulemuste erinevusi, olid järgmised:

- Projektide ebamääraselt sõnastatud eesmärgid ja/või tulemused, mistõttu eesmäärke ja tulemusi oli kohati raske hinnata.
- Projekti taotluses ja lõpparuandes toodud eesmärkide ja tulemuste kirjeldus ei kattunud, mis raskendas projekti eesmärkide saavutamise hindamist.
- Projektitaotlustes küsitud summasid vähendati kohati üsna oluliselt (taotluste rahastamisprotsent oli vahemikus 21-94%), millele ei järgnenud ei projektide tegevuste dokumenteeritud korrigeerimist ega põhjendamist, miks need tegemata jäeti. See aga ei võimaldanud projektide tulemuslikkust alati objektiivselt hinnata.

EKKTT raames ellu viidud projektide lõpparuannetes oli tulemustes välja toodud 12 kaitstud doktoritööd, neist 50% (6 doktoritööd) olid kajastatud lisaks teiste projektide ja rahastamisinstrumentide, sh sihtfinantseeritavate teadusteemade, ETF-i grantide ja teiste riiklike programmide tulemustena (vt joonis 7). Kaitstud doktoritööde kattuvust konkreetsete projektide lõikes vt lisa 6.



**Joonis 7.** EKKTT raames rahastatud projektide lõpparuannetes kajastatud doktoritööde protsentuaalne kattuvus teiste rahastamisvahendite aruandluses kajastatud doktoritöödega. Joonisel: RP – ainult RP aruandluses kajastatud doktoritööd; RP + muu – doktoritööd, mis on lisaks RP aruandlusele kajastatud ka teiste rahastamisinstrumentide, sh sihtfinantseeritavate teadusteemade, ETF-i grantide ja teiste riiklike programmide aruandluses.

Võrreldes projektide tulemusena kaitstud doktoritööde kattuvust teiste rahastamisinstrumentide aruandluses kajastatuga, võib väita, et ETIS-es kirjeldatud teadustöö väljundit kajastatakse korduvalt eri rahastamisvahendite aruandluses.

Sihtvalveerimise läbiviijas tekitas küsimusi olukord, mil rahastuse said projektid, mis olid jätkuprojektideks ekspertidelt nõrga lõpphinnangu saanud projektidele. Kui praegu on asutuste/töögruppide taset näitava mõõdikuna kasutusel projektide arv ja maht, siis täiendava mõõdikuna peaks neile lisanduma ka projektide läbiviimise kvaliteet, mis peaks olema kättesaadav eri rahastamisinstrumentide hindajatele.

Alamoodulite lõikes võib tuua välja, et keeletehnoloogiliste ressurssidega tegelevad projektid olid üldiselt tulemuslikumad kui lõppkasutajale suunatud keeletarkvaraga tegelevad projektid (vt lisa 3). Seda võib seletada ühelt poolt vastuoluga programmi eesmärgi (programmi tulemuste jõudmine kasutusse e-ühiskonnas) ja rahastatavate tegevuste (teadus- ja arendustegevus kuni prototüüpide loomiseni) vahel, kuna rahaliselt oli katmata eesmärgi saavutamiseks äärmiselt oluline lõik esialgselt prototüübist kuni tootmis- ja lõppkasutajale kasutuskõlbuliku tooteni. Teiseks võib välja tuua lõppkasutajale suunatud rakenduste puhul intellektuaalse omandiga seotud probleemid, mistõttu infotehnoloogiaga tegelevad ettevõtted polnud huvitatud potentsiaalselt tootearendusest. Ka ei tohi alahinnata massturu puudumist mõningate lõppkasutajale suunatud rakenduste osas, mis vähendab nii ettevõtete kui teadus- ja arendusasutuste huvi temaatikaga tegelda.

#### Järeldused ja soovitused.

- Vähenenud rahastamisest tingitud muudatused heakskiidetud projektide tegevustes tuleb läbi arutada, kooskõlastada ja dokumenteerida hindamiskomisjoni ja projekti elluviija vahel, et tagada rahaliste vahendite optimaalseim kasutamine.
- Asutuse töögrupi taset näitava mõõdikuna tuleks kasutusele võtta projektide elluviimise kvaliteet. Projektide vahe- ja lõpparuannete hinnangud tuleks teha kättesaadavaks ka teiste rahastamisvahendite ekspertidele.
- Teha projektide lõpparuannete lühiannotatsioonid ETISes avalikult kättesaadavaks.
- Erinevate rahastamisinstrumentide kaudu rahastatud projektid olid kohati nende temaatika ja tulemuste poolest kattuvad. Tulevikus võiks kaaluda kas rahastamisinstrumentide vähendamist või erinevate rahastamisinstrumentide senisest suuremat fokuseeritust.

## 6. Programmi juhtkomitee

Programmi juhtkomitee kinnitati haridus- ja teadusministri 04.09.2006. a käskkirjaga nr 759. Juhtkomitee põhiülesanne oli riikliku programmi raames laekunud projektitaotluste läbivaatamine ja ministrile finantseerimissetpaneku tegemine.

Üheksast juhtkomitee liikmest olid kolm riigi/rahastaja/huvigruppide esindajatena Haridus- ja teadusministeeriumi ning Majandus- ja kommunikatsiooniministeeriumi esindajad ning kuus teaduslike ekspertidena vastavas valdkonnas tegutsevate teadus- ja arendusasutuste esindajad. Juhtkomitee liikmetega läbi viidud intervjuude põhjal võib väita, et kõigil juhtkomitee liikmetel oli piisavalt võimalusi oma arvamuste ja ettepanekute esitamiseks, seega olid ka esindatud ministeeriumi esindajatel võimalik regulaarselt jälgida programmi ja konkreetsete projektide kulgu. Et tegelikkuses puudub Eestis riigi väiksuse ja eesti keelt kõnelevate inimeste arvu vähesuse tõttu nõudlik masskasutaja, siis oleks juhtkomitee võinud olla poliitiliselt tugevam, et tagada riigi ja lõppkasutaja seisukohast oluliste eesmärkide järgimist asutustepõhises projektide taotlemises ja elluviimises. Ja et mitmed lõppkasutajatena oluliste huvigruppidega (sh nt eakad või puuetega inimesed) tegelemine on riigi poolt antud sotsiaalministeeriumi haldusalasse, siis oleks võinud juhtkomiteesse kuuluda ka nende esindaja.

Juhtkomitee liikmed lugesid ja hindasid ise kõiki projektitaotlusi, kaasamata väliseid eksperte. Hindamiseks kasutati kokkulepitud hindamiskriteeriume, mille alusel anti taotlustele punkte. Summeeritud hindepunktide alusel moodustati taotluste pingerida, mis arutati veelkord läbi juhtkomitee koosolekul. Projektide rahastamine sõltus ka taotluse kvaliteedist ja pingereas eespool olevad taotlused said ka protsentuaalselt suurema toetuse. Otsustusprotsess oli demokraatlik ja selle aluseks oli hindamisel moodustunud pingerida, üldjuhul jõuti projektide rahastamises konsensusele. Suuri pingeid otsustusprotsess juhtkomiteede liikmete vahel esile ei toonud, aeg-ajalt oli probleemiks juhtkomitee liikmete roll asutuste esindajatena, samuti võib probleemina välja tuua mitmeid nõrku hinnanguid saanud töögrupi jätkuvat rahastamist.

Juhtkomiteed abistas tehniliselt programmi palgaline koordinaator, kelle olemasolu aitas olulisel määral kaasa juhtkomitee tegevuse sujuvamaks muutmisele.

### Järeldused ja soovitused.

- Programmi oluliste eesmärkide sõnastamisel on äärmiselt oluline roll programme ettevalmistaval töögrupil, kes täpsustab nii juhtkomitee ülesandeid kui sõnastab programmi strateegilised eesmärgid ja formuleerib võimalikud lahendused eesmärkide saavutamiseks.
- Erinevatest huvigruppidest ja lõppkasutajate esindajatest koosneva juhtkomitee roll on eelkõige programmi tervikuna jälgiv ja juhtiv.
- Projektitaotluste taset ja kvaliteeti peaks hindama valdkondlikest ekspertidest koosnev hindamiskomisjon. Taotluste hindamisprotsessi tuleks kindlasti kaasata väliseksperte.
- Juhtkomitee ülesanded peavad olema piisava selgusega määratletud ja lahti kirjutatud, liikmete tegevus võiks olla tasustatud. EKKTT juhtkomitee ülesanded, sh analüüs ja vastutus, olid programmi tekstis (p 1.3) küll välja toodud, ent juhtkomitee piirdus põhiliselt hindamiskomisjoni rolliga. Juhtkomitee peab võtma täieliku vastutuse programmi eduka täitmise eest.
- Keeruliste ja spetsiifiliste ülesannete lahendamiseks ei pruugi avatud konkursid olla alati parimaks lahenduseks. Juhtkomiteel peab olema õigus täiendavalt valida võimalusi tööde teostajate leidmiseks.
- Äärmiselt oluline on programmi koordinaatori toetav ja tehniline roll, sellist praktikat tuleks kindlasti jätkata.



## 7. Olulisemad järeldused, soovitused ja ettepanekud

- Erinevate huvigruppide ja lõppkasutajate esindajate (sh erinevate ministeeriumite) osakaal programme ettevalmistavates töögruppides peab olema suurem, et tagada programmi tulemuste lõppkasutajale suunatud huvide parem esindatus programmidokumentis.
- Programmid ja nende raames rahastatavad tegevused peaksid olema selgelt fokuseeritud ning üldeesmärgid ja tulemuslikkuse indikaatorid peaksid olema välja toodud, senisest enam peaks keskendumise üldistele, ent olulisematele eesmärkidele, samuti vältima tegevuste topeltrahastamise võimalusi erinevate rahastamisinstrumentide raames.
- Programmide raames toetatavad tegevused peavad toetama kõiki programmi edukaks realiseerimiseks vajalikke kriitilisi protsesse.
- Lisaks programmi raames toetatavatele tegevustele tuleb leida lahendused ka programmivälistele pärssivatele tingimustele, et saavutada programmide eesmärkide saavutamist toetav väliskeskond (nt intellektuaalomandiga seonduvad küsimused).
- Programmide juhtkomiteed peavad tegelema programmi kui terviku struktuurse jälgimise, analüüsi ja juhtimisega ning vastutama programmi eesmärkide saavutamise eest.
- Projektitaotluste taset ja kvaliteeti peab hindama valdkondlikest ekspertidest koosnev hindamiskomisjon, mille koosseisu tuleks kindlasti kaasata väliseksperte.
- Vähenenud rahastamisest tingitud muudatused heakskiidetud projektitaotluste tegevustes tuleb enne taotluste lõplikku heakskiitmist dokumenteeritult läbi arutada ja kooskõlastada hindamiskomisjoni ja projekti elluviija vahel.

## 8. Lisad

1. Riikliku programmi „Eesti keele keeletehnoloogiline tugi (2006–2010)“ tekst (\*)  
(otseviit <http://www.hm.ee/index.php?popup=download&id=4964>)
2. EKKTT sihtvalveerimise lähteülesanne
3. Koondinfo programmi raames rahastatud projektide kestuse, rahastamise, aruandluse ja kaasatud personali kohta (allikas: ETIS + aruanded paberil)
4. Juhtkomitee liikmete nimekiri, sh intervjueeritud liikmed
5. Projektide põhitäitjate kattuvus SF-i ja ETF-i põhitäitjatega 2010. aastal (allikas: ETIS)
6. Projektide tulemusena kajastatud kaitstud doktoritööde kattuvus teiste rahastamisinstrumentide kaudu rahastatud projektide aruandluses kajastatuga (allikas: ETIS)
7. Eesti teaduse infrastruktuuri teekaart 2010 (\*)  
(otseviide <http://www2.archimedes.ee/teadus/File/Teekaart.pdf>)
8. EKKTT projektide kulud kuluartiklite lõikes (allikas: ETIS)
9. Eesti keele arendamise strateegia (2004-2010) vaheseire 2009.a. (\*)  
(otseviide <http://ekn.hm.ee/node/39>)

### **Lisamaterjal:**

- RP EKKTT materjalid internetis: [www.keeletehnoloogia.ee](http://www.keeletehnoloogia.ee)
- EKKTT 2007, 2009 ja 2010.a toimunud aruandekonverentside trükised.

## **Lisa 2: Riikliku programmi „Eesti keele keeletehnoloogiline tugi (2006-2010)“ sihtevalveerimise lähteülesanne**

### **Taust**

Riiklik programm (*edaspidi RP*) „Eesti keele keeletehnoloogiline tugi (2006-2010)“ (*edaspidi EKKTT*) oli käigus aastatel 2006-2010.

Programmi peaesmärgiks oli eesti keele keeletehnoloogilise toe arendamine tasemele, mis võimaldaks eesti keelel edukalt toimida tänapäeva infotehnoloogilises keskkonnas, sh programmi tulemuste rakendamine e-ühiskonnas.

Programm oli üles ehitatud kolmes alajaotuses:

- Keeletehnoloogia valdkonnad;
- Keeletarkvara;
- Keeletehnoloogilised ressursid.

Programmi rahastati riigieelarvest Haridus- ja Teadusministeeriumi vahendusel, rahastati 31 projekti ning selle kogumaht oli ligi 53 milj krooni.

### **Eesmärgid**

Planeeritava sihtevalveerimise (*edaspidi SE*) eesmärgid on:

- hinnata ja analüüsida EKKTT kui rahastamisinstrumendi ettevalmistust, läbiviimist ja juhtimist, sh võrdlevalt EKRM-ga;
- hinnata ja analüüsida EKKTT tulemuslikkust ja väljundeid võrreldes programmi käivitamisel püstitatud eesmärkidega;
- hinnata ja analüüsida EKKTT raames rahastatud projektide tulemuste vastavust programmi ja projektide eesmärkidele ning nende jätkusuutlikkust;
- hinnata ja analüüsida EKKTT ja selle raames rahastatud projektide seotust/kattuvust teiste RP-de ja rahastamisvahenditega;
- anda soovitusi ja teha ettepanekuid edaspidiseks.

SE käigus saadud informatsiooni kasutatakse riiklike programmide mõjude ning rakendamise hindamiseks ja analüüsimiseks ning uute meetmete ettevalmistamisel.

### **SE tulemuste sihtgrupp**

SE tulemuste sihtgrupp on HTM, teaduspoliitika komisjon, teised ministeeriumid jt teaduspoliitikate, sh riiklike programmide väljatöötajad ning RP kaudu rahastatud asutused.

SEI osalevad RP kaudu rahastatud asutused:

- Eesti Keele Instituut
- Eesti Kirjandusmuuseum
- FiloSoft OÜ
- Tallinna Tehnikaülikool (sh Küberneetika Instituut)
- Tartu Ülikool
- ELIKO TAK

### **Metoodika/info allikad**

EKKTT üld- ja taustainfo saadakse olemasolevast kirjalikust dokumentatsioonist (teaduspoliitilised materjalid, vastavad riigikogu, VV ja ministrite õigusaktid) ning vajadusel intervjuudest EKKTT ettevalmistava töögrupi juhtide/liikmetega ning HTM toonaste ametnikega. Info juhtgrupi tegevuste kohta saadakse juhtgrupi koosolekute protokollidest ja vajadusel intervjuudest juhtkomiteede juhtide/liikmetega. Üld- ja taustainfo ning juhtkomitee materjalide kättesaadavuse tagab HTM.

Projektitaotluste, projektide ja -aruannete info saadakse taotlustest ja aruannetest ETISes ning vajadusel intervjuudest projektijuhtidega.

Info EKKTT teaduspoliitilise rolli, mõju ja jätkusuutlikkuse kohta saadakse eelnimetatud materjalidest ning vajadusel intervjuudest juhtkomitee liikmetega, HTM ametnikega ning toetatud asutuste juhtidega/projektijuhtidega.

### **Evalveerimise läbiviimine**

**Juhtgruppi** evalveerimise läbiviimise ettevalmistamiseks ja koordineerimiseks ei moodustata, SE läbiviimise ja tulemuslikkuse tagab vastavalt lähteülesandele SE läbiviija kooskõlastatult SE korraldajaga.

### **Väljundid**

SE väljundiks on evalveerimisraport, mis sisaldab hinnanguid lõppenud RP (tulemuslikkuse) kohta ja soovitusi tulevaste programmide väljatöötamiseks.

Raporti struktuur:

- Lühikokkuvõtte (taust, eesmärgid, tulemused, soovitus)
- Põhiaruanne
  - Üldeesmärk
  - Raporti sihtgrupp
  - Konkreetsed eesmärgid
  - Metoodika (infoallikad, andmeanalüüs, piirangud)
  - SE läbiviimine
  - Programmi lühikirjeldused (eesmärgid, vajalikkus, sisu, kasusaajad, tulemused)
  - SE tulemused. RP-de:
    - EKKTT tulemuslikkus ja mõju;
    - EKKTT ettevalmistus, läbiviimine ja juhtimine;
    - EKKTT raames rahastatud projektide tulemuslikkus ja jätkusuutlikkus;
    - EKKTT ja selle raames rahastatud projektide kattuvus teiste rahastamisvahenditega;
    - Muu
  - Järeldused ja soovitus

Raporti lisad:

- Lähteülesanne
- Ajakava
- Läbitöötatud dokumentatsioon
- Intervjueeritute nimekiri
- Intervjuude sisulised juhendid
- RPde programmdokumendid
- Muud raporti käigus valminud materjalid

### **SE ajakava**

<b>Tegevus</b>	<b>Ajakava (2011)</b>
Koostöölepingu sõlmimine Haridus- ja teadusministeeriumi ja SA Archimedes vahel, lähteülesande väljatöötamine ja SE sisuline kavandamine	30. märts
Olemasoleva dokumentatsiooni kogumine ja süstematiseerimine, vajadusel kommunikatsioon RP täitmisel osalenud asutuste ja juhtgrupi liikmetega	18. aprill
Andmeanalüüs, vajadusel kommunikatsioon RP täitmisel osalenud asutuste ja juhtgrupi liikmetega	31. mai
Tulemuste esmane analüüs ja raporti ettevalmistamine	15. juuni
Lõppraporti valmimine	30. juuni

### **Maksumus**

Evalveerimise maksumus, sh töötasu, transpordi- ja materjalikulu kaetakse SA Archimedese poolt.

### **Tehniline tugi**

Tehnilise toe (dokumentatsiooni kogumine, säilitamine, ajakava ja sellest kinnipidamine, transport ja majutus, kontoripind ja –tarbed) tagab SA Archimedes.

### Lisa 3. Koondinfo programmi raames rahastatud projektide kestuse, rahastamise, aruandluse ja kaasatud personali kohta (allikas: ETIS + aruanded paberil)

Kood	Pealkiri	Algus	Lõpp	Kestvus	Asutus	Kokku (tuh kr)	Eesmärgid taotluses	Tulemused lõpparuandes	Põhitajad isikud	Põhitajad FTE	Tööjõud isikud	Tööjõud FTE	Töökuud
EKKT06-1	Eesti emotsionaalse kõne korpus	2006	2010	5	Eesti Keele Instituut	3 541	Eesmärk on luua korpus, mis on semantiliselt koostöös suulise kõne korpusena (Karl Pajusalu projekt, TÜ) ja korpuselise tekst-kõne sünteesiga (Meelis Mihkla projekt, EKI). Selgelt tajutavate emotsioonidega ja neutraalseks tunnistatud laused võetakse korpusesse. Korpus koosneb kõnefaalidest (kokku u 32 000 sekundit), relatsioonilisest andmebaasist ja veebileidest. Kõnefaile hoitakse pakkimata WAV-failidena. Andmebaas hoitakse metaandmeid kõnefaalide kohta, ajasükkrooneid andmeid (erinevaid märgendusi (ortograafiline, foneetiline jm), põhitooni, intensiivsust jne), tajukatsete kirjeldusi ja tulemusi. Veebileides võimaldab: metaandmete sisestamise; tajukatsete läbiviimise; sünteesi ja analüüsi päringud.	Korpus on niisugusena, nagu ta teoreetiliselt kavandati valmis. Emotsionaalse kõne korpus koosneb andmebaasist ja kuulamis- ja lugemistestide läbiviimise ning tulemuste statistilise hindamise vahenditest. Korpus on teostatud veebipõhise rakendusena. Erandlik on Eesti emotsionaalse kõne korpus oma avaliku juurdepääsu poolest.	3	2,5	5	0,0	19
EKKT06-10	Mitmesõnaliste verbide ja nende kokku-lahku kirjutamise vigade äratundmine eestikeelsetes tekstides	2006	2008	3	Tartu Ülikool	1 120	Projekti lõpp-eesmärgiks on leida seaduspärasused, mis võimaldaksid automaatselt ära tunda eestikeelses tekstis esinevaid mitmesõnalisi verbe - ühend- ja väljendverbe ning tugiverbi-ühendeid. Seejuures lahendamist nõudev teoreetiline probleem (verbikeske) püsiühendi piiridest jääb väljapoole antud projekti raame. Väljendite äratundmine võimaldab asuda lahendama kokku-lahku kirjutamise lubatavuse probleemi konkreetses kontekstis.	On loodud kolm omavahel seotud keeleressurssi: 1. Verbikeskete püsiühendite andmebaas 2. Kolmest erinevast tekstiklassist (ilukirjanduse, ajakirjanduse ja horisondi tekstid, igaihes 100 000 sõna) koosnev morfoloogiliselt ühestatud tekstikorpus, milles verbikesked väljendid on märgendatud. 3. Programm, mis ühaldab andmebaasi ja morfoloogiliselt ühestatud korpuselise märgendab verbikeskete väljendeid. Verbikeskete püsiühendite märgendaja sisendiks on morfoloogiliselt ühestatud korpus.	3	0,3	0	0,0	0
EKKT06-11	Lihtlause semantiline analüüs	2006	2008	3	Tartu Ülikool	1 490	Eesmärgiks on luua automaatse semantilise analüüsi programm, mis antud projekti puhul analüüsib ainult lihtlauseid, s.o lauseid, kus pole alistavaid sidesõnu ega alistusseoses olevaid osalauseid. Esialgu on mõttekas piirata semantilist analüüsi teatud kitsama ontoloogilise valdkonnaga. Meie projekti puhul on selleks valitud liikumisega (k.a liigutamine, asetamine jne) seotud situatsioonid.	Projekti põhitulemusteks on: 1) semantiliselt analüüsiks vajalikud ressursid, milleks on a) freimileksikon ja b) semantiliselt märgendatud korpus; 2) semantilise analüüsi tarkvara, mille olulisimaks esindajaks praeguse seisuga on semantilisi järeldusi tegev programm.	7	1,3	7	1,5	60
EKKT06-12	Elektrooniliste teatmeteoste kasutajasõbralikud päringusüsteemid	2006	2008	3	Tartu Ülikool	1 060	1. sõnastike ja teiste teatmeteoste struktuurne märgendamine ja ettevalmistamine elektroonilise andmebaasi ning info-otsingute jaoks. 2. lihtsate ja mugavate päringusüsteemide välja töötamine ja arendamine. 3. elektrooniliste sõnastike kasutuse populariseerimine ennekõike läbi juhendmaterjalide ja koolitusprogrammide välja töötamise, koostamise ja koolituste läbi viimise.	1. Sõnastike keskkonna arendamise projekt on üle kandunud Eesti Keele Instituuti, kus on Ülle Viksi juhitud projekt. 2. Ligikaudse otsimise põhilised tööriistad: Üldistatud teiseduskauge katsetamise päringumootor: ( <a href="https://biit-dev.cs.ut.ee/~orasmaa/gen_ed_test/">https://biit-dev.cs.ut.ee/~orasmaa/gen_ed_test/</a> ); Inglise sõnastikust ligikaudse otsimise päringumootor: ( <a href="https://biit-dev.cs.ut.ee/~orasmaa/ing_ligikaudne/">https://biit-dev.cs.ut.ee/~orasmaa/ing_ligikaudne/</a> ) 3. Populariseerimise kohta aruandluses ülevaade puudub.	2	0,1	7	5,0	65
EKKT06-13	Eesti vanema kirjakeele elektroonilised kogud	2006	2010	5	Tartu Ülikool	759	Põhieesmärgiks on luua elektrooniliselt kasutatav tekstide väliskorpus, mis võimaldaks ligipääsu vanema eesti kirjakeele olulisematele tekstidele. Projekti teine olulisem eesmärk on luua 17. sajandi sõnaraamatute koondandmebaas, mis võimaldaks uurijatel ligipääsu sellele haruldasele keelevarale, mis seni on arhiivides kättesaadav olnud vaid üksikutele huvilistele.	Projekti käigus on loodud originaalne leksikograafi abivahend "Vakker". Vana kirjakeele korpusel on võimalik päringuid teha nii lemmatiseeritud kui ka lemmatiseerimata tekstidest (sellisel juhul on vaja aimata tekstisõna kuju). Projekti käigus on välja töötatud ka vanade sõnastike elektroonilise töötlemise meetodika. Projekti kolme esimese aasta jooksul (2006-2008) täiendati vana kirjakeele korpusel 18. ja 19. sajandi tekstidega (kokku 800 000 tekstisõna mahus). Aastatel 2006-2010 lemmatiseeriti 17. sajandi tekste, põhiliselt H. Stahli tekste, 80 000 tekstisõna mahus. Praegu on eesti kirjakeele sõnade esmaesinemused tuvastatavad kuni 17. sajandi keskpaigani.	5	1,0	18	4,6	193

EKKT06-14	Eesti keele koondkorpus	2006	2009	4	Tartu Ülikool	1 913	Käesoleva projekti eesmärgiks on eesti keele koondkorpuse arendamine 200 miljoni sõnani ja koondkorpuse automaatne morfoloogiline märgendamine. Käesolev projekt jätkab riikliku programmi „Eesti keel ja rahvuslik mälu“ projekti „Eesti keele koondkorpus“ raames tehtud tööd. Võrreldes programmi „Eesti keel ja rahvuslik mälu“ raames esitatud projektiga, on käesoleva projekti ülesannetesse lisandunud koondkorpuse morfoloogiline märgendamine ja sellega seoses pikeneb projekt ühe aasta võrra. Eesti keele koondkorpuse suurus on hetkel ligi sada miljonit sõna.	1) Koondkorpuse maht on nüüd 245 miljonit sõna, kasutajaliides (www.cl.ut.ee/essursid/kasutajaliides). 2) Koondkorpuse automaatne morfoloogiline märgendamine - selle töö tegi ära projekt "Korpusepäring Keeleveebis"	4	1,45	8	3,1	82
EKKT06-15	Eestikeelne infodialoog arvutiga	2006	2008	3	Tartu Ülikool	1 250	Projekti eesmärgiks on tarkvara väljatöötamine, mis võimaldaks eestikeelset küsimus-vastusdialoogi arvutiga inimestevahelise suhtluse reeglite kohaselt (sh täpsustavate küsimuste esitamine, tagasiside, kultuurispetsiifilised rituaalid jms). Sellise tarkvara loomiseks vajalik keeleressurs on märgendatud dialoogikorpus (praegune maht on u 100 000 dialoogiaktide tasemel märgendatud tekstisõna). Korpuse mahtu on vaja suurendada vähemalt 200 000 tekstisõnani Selle tulemusel valmib programm – eestikeelse telefoniteenuste süsteemi prototüüp.	1. Suuliste dialoogide kogumaht on 1148 litereeritud teksti ehk 221037 tekstisõna. Projekti käigus on maht kasvanud u 300 teksti e 70000 sõna võrra). 2. Projekti tulemusel on valminud dialoogsüsteem, kus on kasutatud varem loodud küsimus-vastussüsteemide Reisiagent ja Teatriagent arendamisel saadud kogemusi. See on veebis kasutatav "vestlusrobot", mis suudab eesti keeles anda infot Tartu kinokavade kohta. Kõnetuvastust ei ole lisatud, s.t süsteemile saab esitada üksnes kirjalikke küsimusi päringuaknas ja saada vastuseid eesti keeles, nii kirjutatud tekstina kui (soovi korral) ka tehiskõnes.	7	2,2	11	1,1	47
EKKT06-16	Eesti keele spontaanse kõne foneetiline korpus	2006	2010	5	Tartu Ülikool	3 078	Luu teiste eestikeelse kõne korpustega ühilduv spontaanse kõne foneetiliselt märgendatud korpus, mida saab kasutada eesti keele häälduse põhiparameetrite analüüsimisel ning eesti keele kõnesünteesi ja kõnetuvastuse ülesannete täitmisel. (Praeguseks on märgendatud 0,5 tundi kõnet, u 20 000 märgendusühikut). Salvatakse 40 keelejuhi (20 naist + 20 meest) kolme laadi kõnetekste. Salvastatud kõne tuleb foneetiliselt märgendada erinevatel märgenduskihtidel. Andmebaasi mahuks esimesel etapil on vähemalt 20 tunni (30 minutit 40 keelejuhilt) kõnematerjali märgendamine.	Projekti töö tulemusel salvestati 35 keelejuhi kõnet kogukestusega 28 tundi. Lingvistilistest kihtidest on lisaks märgendatud koguluatases automaatselt häälikustruktuuride ja lausungitasand. Poolautomaatselt on märgendatud silbitasand ning selle käsitsi kontrollimise käigus on ühtlasi kontrollitud ja ühtlustatud kahe esimese tasandi märgendamist (kontrollitud on kokku 17 tundi ja 30 minutit). Sõna- ja häälikutasandil on korpuses kokku 709 574 segmenti (sõnatasandil 180 008 ja häälikutasandil 529 566 segmenti) ja muudel tasanditel 291 391 segmenti. Segmentide koguarv on korpuses 1 000 965.	6	1,8	9	2,9	185
EKKT06-17	Eesti kõnekeele korpuse kogumine ja translitereerimine	2004	2008	5	Tartu Ülikool	1 045	Projekti eesmärk on koguda 5 aasta jooksul 2 000 000-tekstisõnaline kõnekeele korpus. 1. koostada üldkorpuse ja ka teiste projektide jaoks vajalike alamkorpuste (infodialoogid, eriti telefonikõned) struktuuri ja mahtude jaotus. a) tekstide lindistamine b) tekstide litereerimine ja märgendamine ning taustakirjalduste koostamine. c) varasema materjali digitaliseerimine.	Lindistatud on enam kui 2,1 miljoni sõnaline korpus, neist on täppislitereeritud (kõik suhtlusnähtused, täpsed pausipikkused) 679 439 sõna. Digitaliseeriti 379 linti (1563 kõnet), umbes 190 tundi.	5	1,3	0	0,0	0
EKKT06-18	Süntaksanalüüsil põhinev keeletarkvara ning selle arendamiseks vajalikud keeleressursid	2004	2008	5	Tartu Ülikool	1 240	Luuakse järgmiste keeletarkvarasüsteemide prototüübid: a) grammatikakorrekto b) süntaksanalüüsil põhinev automaatsete sisukokkuvõtete tegija c) süntaksanalüüsil põhinev infootsisüsteem	Loodi järgmised keeletarkvarasüsteemide prototüübid: 1. Grammatikakorrekto. 2. Eestikeelsete tekstide sisukokkuvõtja.  Loodi järgmised keeleressursid: 1. Automaatseks analüüsiks vajalikud grammatikad, vt (http://www.ut.ee/~kailli/grammatika); 2. Korpused - Pindsüntaktiliselt märgendatud korpus: http://math.ut.ee/~kailli/Korpus/pindmine/, mis on leitav ka Tartu ülikooli korpuste pealehelt (http://www.cl.ut.ee/korpused/). - Puudepank: http://math.ut.ee/~kailli/Korpus/puud/ 3. Vigadekorpus: http://lepo.it.da.ut.ee/~heli_u/vigadekorpus/	5	1,7	0	0,00	0

EKKTT06-2	Eestikeelne korpuspõhine kõnesüntees	2006	2010	5	Eesti Keele Instituut	448	<p>* Eestikeelne kõnekorpused on planeeritud professionaalsete diktorite ette loetud tekstide salvestistena (mees- ja naishääli).</p> <p>Kõnekorpused koostatakse foneetiliselt tasakaalustatud kõnematerjal (60-100 minutit ühe diktori kohta).</p> <p>* Sünteeskõne loomulik rütm ja kõla modelleeritakse prosoodia generaatoriga.</p> <p>* Eestikeelne korpuspõhine süntesaator seotakse SAPI liidese abil Windows-põhiste rakendusprogrammidega, sest eelkõige vajavad süntesaatorit nägemispuudega inimesed arvutiga suhtlemisel ekraanilugeja vahendusel.</p>	<p>1. Korpuspõhine kõnesüntees. Korpuspõhise kõnesünteesi jaoks on koostatud kaks tekstikorpust: ca 600 lausest ja ca 3000 lausest koosnevad. Kokku salvestasime 20,8 tundi kõnet 6 keelejuhi esituses (3meest ja 3 naist). Kõneüksuste valiku meetodil genereeriti viis uut sünteeshäält.</p> <p>2. Kõneprosoodia modelleerimine. Projekti käigus valminud doktoritöös (Mihkila 2007) töötati välja metodoloogia kõnekorpuste baasil eesti keele prosoodia modelleerimiseks ja uurimiseks erinevatel statistilistel tehnikatel.</p> <p>3. Rakendused. Üheks meie kõnesünteesi ja sünteeshääle rakenduseks on projekti käigus välja töötatud elektrooniliste tekstide ettelugemisüsteem ELTE <a href="http://elte.eki.ee/">http://elte.eki.ee/</a></p> <p>4. Kõnesünteesi keeleressursid ja tarkvara prototüübid</p> <p>A. Projekti käigus salvestatud ja märgendatud kõnekorpused hoitakse ja esitatakse serveris <a href="http://heli.eki.ee/">http://heli.eki.ee/</a>. Kõigi keelejuhtidega on koostatud kirjalik tõend, milles nad lubavad salvestatud kõnematerjali kasutada eesti keele tekst-kõne sünteesi tarbeks ja ka eesti keele uurimiseks. Kõnematerjali kasutamine ja tingimuste fikseerimine toimub kontaktisiku tonis@eki.ee vahendusel.</p> <p>B. Loodud kõnesünteesi tarkvara kasutus on vaba. Sünteeshääle näiteid on võimalik kuulata <a href="http://heli.eki.ee/syntees/">http://heli.eki.ee/syntees/</a>. Teksti- ja kõnekorpused ja sünteeshääle tarkvara on allalaaditav <a href="http://heli.eki.ee/">http://heli.eki.ee/</a>.</p>	4	2,3	1	0,2	2
EKKTT06-3	Leksikograafi töökeskkond	2006	2010	5	Eesti Keele Instituut	6433	<p>1. Luua leksikograafidele sobiv interaktiivne töökeskkond, st töövahendid, mis ühilduvad kehtiva rahvusvahelise märgistusstandardiga (XML) ja rakendavad nii universaalseid kui ka eesti keele põhiseid keeletehnoloogia vahendeid. Uues töökeskkonnas on võimalik luua uusi elektroonilisi sõnastikke mitmeks eri otstarbeks: nii tavalisele sõnastikusajatajale kui ka keeletehnoloogia sisetarbeks.</p> <p>2. Koostada eesti lähtekeele andmebaas kakskeelsete sõnaraamatute jaoks ehk Eesti-X-keele sõnastik (= EXS).</p>	<p>1. Valminud on leksikograafi töökeskkonna EELEX põhimoodul: sõnastike veebipõhine haldussüsteem. EELEXi on seni kasutatud või kasutatakse mitut tüüpi sõnastike koostamisel ja toimetamisel. Praeguse seisuga on EELEXi keskkonnas 25 sõnastikku. Sõnastike nimekirja on kodulehel: <a href="http://eelex.eki.ee/">http://eelex.eki.ee/</a>.</p> <p>2. EELEXi keskne rakendus on Eesti-X sõnastiku andmebaas (EXS). EXS andmebaas on eesti pool standardkujul olemas; on tehtud esialgne sõnavalik, esitatud grammatilised andmed, näidatud sõna tähendusliigendus koos vajalike seletustega, lisatud sõna kasutusinfo, olulised liitsõnad, sõnaühendid jne. Sihtkeele andmed puuduvad. EXS-i lähtematerjal on pärit mahukast (u 80000 märksõna) Eesti-vene sõnaraamatust (1997–2009), mida on täiendatud teiste sõnastike andmetega.</p> <p>3. Sõnastike avalikud veebiversioonid (vt <a href="http://portaal.eki.ee/">http://portaal.eki.ee/</a>) on suunatud eelkõige tavakasutajale.</p>	11	5,0	6	?	14
EKKTT06-4	Korpusepäring keeleveebis	2006	2010	5	Filosoft OÜ	1247	<p>1. Võimaldada eesti keele uurijatel ja eesti keelt teise keelena õppijatel tekstikorpuse mugavalt üle interneti kasutada.</p> <p>2. Suurendada veebis kättesaadavate erialasõnastike arvu, teisendades olemasolevaid standardsele XML-põhisele kujule TBX.</p> <p>3. Siduda sõnastiku- ja korpusepäringud omavahel, nii et lisaks sõnastiku-vastele saab kasutaja ka näiteid otse korpusest.</p> <p>Käesolevas projektis ühendatakse olemasolevad eesti keele ressursid - tekstikorpused, morfoloogiline analüsaator koos ühestajaga ning sõnastikud – uueks tervikuks, mis annab seni puudunud võimalused nii uurijatele, eesti keele õppijatele kui ka muidu huvilistele.</p> <p>Liidetakse järgmised ressursid, mille kasutamiseks antud projekti raames ei ole autoriõiguslikke takistusi: TÜ eesti keele koondkorpus (<a href="http://www.cl.ut.ee/">http://www.cl.ut.ee/</a>), Filosofti morfoloogiline analüsaator (<a href="http://www.filosoft.ee/">http://www.filosoft.ee/</a>) ja ühestaja ning mitmete eri autorite sõnastikud (nt. <a href="http://keeleveeb.edu.ee/">http://keeleveeb.edu.ee/</a>).</p> <p>Projekt realiseeritakse leheküljel <a href="http://keeleveeb.edu.ee/">http://keeleveeb.edu.ee/</a>; leping vajaliku serveriruumi kasutamiseks on olemas.</p>	<p>TÜ koondkorpuse tekstid mahuga 250 miljonit sõna on töödeldud. Päringud on tehtavad aadressil <a href="http://www.keeleveeb.ee">www.keeleveeb.ee</a></p> <p>Saab esitada ühispäringu kokku 60 sõnastikule.</p> <p>Kasutaja saab enne päringu tegemist valida nii sõnastikud kui korpused, sel moel saab kasutaja ühepäringuga nii sõnastikuvasted kui näited korpustest.</p>	3	0,9	4	3,4	22



EKKT06-5	Kõne analüüs ja variatiivsuse mudelid	2006	2010	5	Tallinna Tehnikaülikool	2573	<p>1. uurida ja arendada kõnesignaali analüüsi, müratöötuse ja akustilise keskkonna modelleerimise meetodeid – see on vajalik kõnetuvastussüsteemide rakendamiseks erinevates akustilistes tingimustes (kodu, büroo, telefonikanal, jne) ja erinevate kõne kodeerimise algoritmide korral;</p> <p>2. uurida kõneparametreid varieeruvust erinevate kõnelejade vahel ja töötada välja vastavad normaliseerimismeetodid kõnetuvastuse jaoks ning leida sobivad tunnused/meetodid kõneleja identifitseerimiseks tema häälenäidete alusel;</p> <p>3. luua eesti keelele omaste kvantiteedinähtuste mudelid kõnetuvastuse jaoks, samuti leida meetodid akustilise informatsiooni kasutamiseks liitsõnapiiride tuvastamiseks;</p> <p>4. uurida aktsendiga kõne akustilisi omadusi, luua akustilised mudelid aktsendiga kõne tuvastuseks</p> <p>5. arendada signaalitöötlusmeetodeid aktsendi automaatseks hindamiseks, luua e-õppe tarkvara eesti keele häälduse paremaks omandamiseks.</p>	Analüüsi ja uuriti. Kavandatud uurimisteedest ei jõutud käsitleda akustilise informatsiooni kasutamist liitsõnapiiride tuvastamiseks ja rakendada uurimistulemusi tarkvara prototüüpide arendamisel, nt e-õppe tarkvara eesti keele häälduse paremaks omandamiseks, akustilised mudelid aktsendiga kõne tuvastuseks, meetodid kõneleja identifitseerimiseks.	4	1,2	3	1,20	10
EKKT06-6	Kõnekeele ressursid ja kõnetehnoloogia andmebaasid	2006	2010	5	Tallinna Tehnikaülikool	2903	<p>1. kõnetuvastuse uuringuteks, tuvastussüsteemide treenimiseks ja testimiseks vajalike erinevate kõnekorpuste (spontaanne kõne, dialoogid fikseeritud valdkondades, uudiste lugemine, jms) salvestamine;</p> <p>2. aktsendikorpuse loomine eri emakeelega isikute eesti keele hääldusnäidetest. Salvestatakse kõnenäiteid vene, saksa, prantsuse, rootsi, soome, Ameerika ja Suurbritannia inglise ja muu emakeelega eesti keele kõnelejatelt;</p> <p>3. tehnoloogilise keskkonna loomine erinevate kõneandmebaaside haldamiseks ja efektiivseks kasutamiseks uurimistöös. Luuakse tehnoloogilised vahendid eri formaatides salvestatud kõnekorpuste ühildamiseks ja liides päringute esitamiseks;</p> <p>4. kõnesignaali segmenteerimise ja märgendamise ühtsete printsiipide ja reeglite väljatöötamine; vastava juhendi koostamine.</p>	<p>1. Kõnekorpuste loomine: Uudistekorpust sisaldab ca 300 tundi 10 erineva raadiodiktoriga loetud Eesti Raadio lühiaudiseid. Märgendatud korpuse maht on ca 30 tundi. Loengukõne korpust sisaldab ca 350 tundi eri ainevaldkondade akadeemiliste loengute salvestust. Vestlussaadete korpust sisaldab 34 saate salvestust kogumahuga ca 20 tundi.</p> <p>2. Aktsendikorpust sisaldab 160 keelejuhi salvestusi, võrdlusmaterjalina on salvestatud 20 eesti emakeelega keelejuhi kõnenäited.</p> <p>3. Korpuste haldamiseks ja kättesaadavaks tegemiseks on installeeritud ja kohandatud korpuste haldussüsteem LAMUS.</p> <p>4. Erinevate kõnekorpuste segmenteerimiseks ja märgendamiseks on koostatud reeglistik, mis on vabalt kättesaadav <a href="http://www.phon.ioc.ee/dokuwiki/doku.php?id=projects:ressursid:ressursid.et">http://www.phon.ioc.ee/dokuwiki/doku.php?id=projects:ressursid:ressursid.et</a></p>	2	0,9	8	2,6	90
EKKT06-7	Eestikeelse kõnetuvastuse meetodite uurimine ja arendamine	2006	2010	5	Tallinna Tehnikaülikool	3001	<p>1. uurida, millised on optimaalsed leksikonis salvestatavad tuvastusühikud (difoonid, silbid, pseudo-morfeemid, jms);</p> <p>2. luua valitud tuvastusüksustel baseeruv morfo-süntaktiline keelemudel; uurida semantiliste seoste modelleerimist statistilises keelemudelis;</p> <p>3. rakendada algoritme tuvastussüsteemi kiireks adapteerimiseks kõneleja häälduse eripäraga;</p> <p>4. uurida kõnetuvastussüsteemide rakendusprobleeme mobiilside kanalis;</p> <p>5. luua optimaalsed tehnoloogilised lahendused piiratud sõnavaraga (kuni 10 000 sõna) tuvastussüsteemide arendamiseks;</p> <p>6. uurida ja arendada tehnoloogilisi lahendusi piiramatult sõnavaraga tuvastussüsteemide loomiseks;</p> <p>7. välja töötada erinevaid tuvastussüsteemide prototüüpe.</p>	<p>1. leiti, et 4-gramm keelemudelite kasutamisel on sobilikumad ühikud liitsõnad ja "tükeldatud" liitsõnad. Liitsõnad rekonstrueeritakse tuvastatud liitsõnadest järeltöötuse käigus.</p> <p>2. semantiliste seoste modelleerimise meetodid ja tulemused avaldati mitmes konverentsipublikatsioonis.</p> <p>3. kõneleja adapteerimine on implementeeritud tuvastustehnoloogias.</p> <p>4. Osaliselt täidetud: seda probleemi intensiivselt ei uuritud, kuid implementeeritud tuvastustehnoloogia toetab kitsa sagedusribaga kõne tuvastamist, selleks treeniti spetsiaalsed akustilised mudelid.</p> <p>5. Osaliselt täidetud: loodi avatud lähtekoodiga näiterakendus piiratud sõnavaraga kõnetuvastuseks.</p> <p>6. uurida ja arendada tehnoloogilisi lahendusi piiramatult sõnavaraga tuvastussüsteemide loomiseks - Täidetud.</p> <p>7. välja töötada erinevaid tuvastussüsteemide prototüüpe - Täidetud.</p>	2	2,0	1	0,5	8
EKKT06-8	Veebipõhine interaktiivne keeleõpe ja selleks vajalikud ressursid	2006	2010	5	Tartu Ülikool	936	<p>1. Koostada eesti keele õppijate ja kasutajate tüüpiliste korpust, 2010 aastaks 300 000 sõna.</p> <p>2. Arendada kõigile ligipääsetavat ja hõlpsalt kasutatavat veebipõhist avatud õppekeskkonda, mis sisaldaks projekti käigus valmivaid interaktiivseid eesti keele õppematerjale toetamaks keeleõppijate tema õppeprotsessis; meetodilisi materjale eesti keele kui teise keele õpetamiseks; õppetooli poolt korraldatava konverentsi "Emakeel ja teised keeled" materjale ja ettekannete kogumikke; viidakogu eesti keele ja kultuuriga seotud veebilehtedest.</p>	<p>1. Õppijakeele korpust koosneb kahest elektroonilisest andmebaasist:</p> <p>a) paralleelkorpust 2006-2007. 2008. aasta veebruaris seisuga on paralleelkorpustes vigaseid lauseid 9000, neis kokku 128 000 sõna; parandusi lausetena on 9100, sõnadena 129 000.</p> <p>b) tekstikorpust 2008-2010. Tekstide sõnaarv on kokku u 210 000 sõna. Õppijakeele vigade märgendamiseks on välja töötatud detailne veamärgendussüsteem.</p> <p>2. Veebipõhist avatud õppekeskkonda arendatud ei ole.</p>	10	0,5	12	1,0	125
EKKT06-9	Masintõlge I	2006	2008	3	Tartu Ülikool	1280	<p>1. Parandada olemasoleva inglise-eesti korpuse paralleelstimise kvaliteeti, suurendada tema mahtu ja mitmekesisust tekstide valikut.</p> <p>2. Märgendada korpust lingvistiliselt nii põhjalikult kui see on automaatselt võimalik. Seda tuleb teha nii eesti kui inglise keele osas.</p> <p>3. Teha seeria masintõlke eksperimente nii olemasolevate vabavaraaluste kui TÜs välja töötatud tarkvarapakettidega, kasutades momeendi olemasolevaid korpust.</p>	Parandatud, märgendatud, eksperimenteeritud.	5	0,3	2	1,3	14

EKKTT07-21	TÜ eesti keele tesauruse (eesti wordneti) täiendamine	2007	2010	4	Tartu Ülikool	1 564	Leksikaal-semantilise andmebaasi laiendamine ja täiendamine. Projekti lõpuks (2010.a.) prognoosime tesaurusesse vähemalt 25 000 uut mõistet (arvestusega, et täiskohaga töötaja sisestab aastas u 1000-2000 uut mõistet koos selle semantiliste seostega). 2007.a sisaldas andmebaas 15000 mõistet.	Oleme tesaurusesse lisanud uusi sünohulki (mõisteid) 27676, milles on erinevaid sõnu 36608. Arvestusse läksid ka (pool)automaatselt lisatud sünohulgad (neid on üle 9000.	8	2,90	6	0,8	29
EKKTT07-23	Reeglipõhine keeletarkvara	2007	2008	2	Tartu Ülikool	440	1. arendada tarkvara selliselt, et viia teineteisest täielikult lahku reeglite kogum ja reeglite rakendamise mootor. 2. arendada edasi reeglite kogumeid, mida saaks kasutada morfoloogilise analüüsi ja sünteesi eesmärkidel. 3. arendada edasi algoritme, mis lubaks reeglite kogumeid kiiremini evalveerida. 4. luua vabavaraline eesti keele speller, mida saaks kasutada erinevatel platvormidel ja mis liidestuks standardsete spelleriiliidestega (ispell, aspell, myspell, hunspell).	1-3: tõenäoliselt on midagi tehtud. 4: realiseerimata. Tulemuste kirjeldus ei lange kokku eesmärkide kirjeldusega.	2	1,0	2	1,5	18
EKKTT08-26	VAKO - eesti vahekeele korpuse keeletarkvara ja keeletehnoloogilise ressursi arendamine	2008	2010	3	Tallinna Ülikool	1 244	1. Olemasoleva keeletarkvara alusel ning seda arendades luua EVKK automaatseks töötlemiseks sobivad tarkvararakendused. 2. Olemasoleva keeletehnoloogilise ressursi laiendamine, sh suurendamine EVKks 5 miljoni sõneni ning funktsionaalsuste laiendamine (õppijakeele elektroonilise sõnastiku loomine ja sagedussõnastiku koostamine).	1. Sõnajärje vealeidja prototüübi veebirakendus on alles valmimisejärgus. 2. Raske hinnata.	7	4,0	9	3,7	168
EKKTT08-28	Eesti keeleressursside keskus	2008	2010	3	Tartu Ülikool	1 354	Luua infrastruktuur, mis teeb eestikeelsed keeleressurssid ja keeletehnoloogilise tarkvara kõigile huvilistele kättesaadavaks. Lisaks keeleressursside kogumisele tuleb käivitada süsteem olemasolevate keeleressursside tutvustamiseks, teavitamiseks ja potentsiaalsete kasutajate koolitamiseks.	Tehti ära ettevalmistav töö Eesti keeleressursside keskuse käivitamiseks.	2	0,7	9	2,1	27
EKKTT08-29	Eesti fraseologismide elektroonilise alussõnastiku loomine	2008	2010	3	Eesti Kirjandusmuuseum	528	Toota Eesti kõnekäändude ja fraseologismide andmebaasi materjali põhjal eesti fraseologismide mõisteline alussõnastik elektroonsel kujul. Internetis üldkasutatav alussõnastik tagaks üld- ja erisõnastikega võrreldes hõlpsama ligipääsu fraseologismidele, võimaldaks kajastada fraseologismide kõige üldisemaid suhteid ja vahakordi.	Valminud on Eesti fraseologismide elektrooniline alussõnastik. Sõnastiku põhiüksuseks on 1893 mõisteartiklit, kokku on sõnastikus 26813 tähendusüksust.	3	0,4	2	1,2	48
EKKTT09-57	Intelligentne kasutajaliides andmebaasidele	2009	2010	2	Tartu Ülikool	978	Luuakse kasutajaliides, mis võimaldaks hõlpsat adapteerumist erinevatele ainevaldkondadele ja seostamist erinevate andmebaasidega. Liidest saab minimaalsete täienduste tegemise teel häälestada uutele ainevaldkondadele ja siduda andmebaasidega, andes seega kasutajale võimaluse pöörduda andmebaaside poole eesti keeles ning saada vastuseks adekvaatset, tõest infot. Kasutaja sisestab oma päringu eesti keeles ja saab intelligentse kasutajaliidese vahendusel vastuse samuti eesti keeles, tekstina või soovi korral tehiskõnes.	Loodi: Intelligentne veebiliides kinoinfole: <a href="http://www.dialoogid.ee/kinoagent/">http://www.dialoogid.ee/kinoagent/</a> Veebiliides võlur Ozi dialoogide läbiviimiseks <a href="http://www.dialoogid.ee/aivo/">http://www.dialoogid.ee/aivo/</a> Dialoogikorpuse tööpink <a href="http://www.dialoogid.ee/dialoogid/">http://www.dialoogid.ee/dialoogid/</a>	4	1,4	5	1,8	85

EKKT09-61	Tartu Ülikooli eesti kõnekeele audio- ja videokorpuse kogumine ja otsingitarkvara loomine	2009	2010	2	Tartu Ülikool	842	<p>Projekti üldeesmärgid:</p> <p>a) koguda kõnekeele tekste, translitereerida ja varustada taustakirjeldusega keelekasutust mõjutavate keeleväliste nähtuste kohta</p> <p>b) arendada arvutitarkvara, mis võimaldab otsida korpusest erinevaid keelelisi nähtusi ning neid analüüsida.</p> <p>Eelmist projekti jätkavad tegevused:</p> <p>-Tartu ülikooli suulise keele korpuse lindistamine, transkribeerimine ja taustakirjeldustega varustamine</p> <p>-selle sees inimese-arvuti suhtluse modelleerimiseks kasutatava institutsionaalseid infodialooge sisaldava korpuse (Dialogikorpuse) lindistamine, transkribeerimine ja taustakirjeldustega varustamine</p> <p>Professor Mare Koidu juhitud projektide tarvis</p> <p>-Enne elmist projekti, 1997-2004 kogutud analoogformaadis korpuseosa digitaliseerimine ja transkriptsioonide täpsustamine, viimaks need vastavusse transkriptsiooni praeguse seisuga</p> <p>-kogu korpuse taustakirjelduste korratamine, mis on eelduseks selle maksimaalselt arvutipõhisele kasutusele</p> <p>-täppistranskriptsiooni (maksimaalse põhjalikkusega tehtud transkriptsiooni) valdavate transkribeerijate koolitamine</p> <p>-korpuse kogumise ja kasutamisega seotud juriidiliste probleemide lahendamine vastavuses Eesti ja Euroopa Liidu areneva seadusandlusega</p> <p>Käesoleva projekti uued tegevused jagunevad kahte rühma:</p> <p>- Projekti üheks uueks eesmärgiks on videokorpuse kogumine.</p> <p>- Projekti teiseks uueks eesmärgiks on tarkvara arendamine, mis võimaldaks otsida ja analüüsida korpusest automaatselt erinevat materjali.</p>	Jätkati eelmise projekti jätkutegevusi lubatust väiksemas mahus. Tulemuseks on ca 1,5 miljoni litereeritud sõna suurune korpus.	8	2,5	0	0,0	0
EKKT09-62	Eesti keele semantika ressursid ja vahendid	2009	2010	2	Tartu Ülikool	576	<p>Korrastada semantiliselt ühestatud ja märgendatud (ühestatud sõnatähtendustega) korpus ja suurendada seda 500 000 sõnani (sisaldab praegu 100 000 sõna).</p> <p>Luaa sõnatähtenduste käsitsi ühestamist hõlbustav tarkvara, mis võimaldaks (1) valida ühestatavat faili (2) valida kasutatavat wordneti versiooni (3) leida ühestatavad sõnad failist (4) omistada ühestatavale sõnale wordnetis leiduv tähendus.</p> <p>Luaa leksikaal-semantilise andmebaasi loomist ja haldamist hõlbustav tarkvara.</p> <p>Üldontoloogia ja valdkonnaontoloogia-põhise mitmekeelse infootsingu raamistiku, ressursside ja rakenduste loomine, selleks (1) valitakse sobiv ontoloogiakeel, (2) luuakse ontoloogia ressursid (üldontoloogia ja mõned valdkonna-ontoloogiad) (3) luuakse ontoloogia kasutatav näidisrakendus.</p>	Semantiliselt ühestatud ja märgendatud (ühestatud sõnatähtendustega) korpus sisaldab 277 000 sõna. Loodi sõnatähtenduste käsitsi ühestamist hõlbustav tarkvara KYPAP, mis ei ole kättesaadav. On alustatud leksikaal-semantilise andmebaasi loomist ja kasutamist võimaldava tarkvara loomist: moodul eurown.py	4	1,4	1	0,5	24
EKKT09-63	Lihtlause semantiline analüüs 2	2009	2010	2	Tartu Ülikool	1 122	<p>1. Lause prediaadist sõltuvate argumentide semantiliste rollide automaatne määratlemine.</p> <p>2. Lausest järelduste tegemise programmi täiustamine.</p>	<p>1. Selgus, et ei ole mõttekas üritada isegi mitte lihtlause osas kogu eesti keelt semantiliselt katvat lähenemist.</p> <p>2. "Lausest järelduste tegemise programmi täiustamine" on jäänud osaliselt täitmata.</p>	8	1,6	2	0,7	30
EKKT09-64	Masintõlge II	2009	2010	2	Tartu Ülikool	1 501	<p>Projekti eesmärgiks on parandada olemasoleva masintõlke kvaliteeti. Selleks on kavas kolm suunda:</p> <p>1. Kasutada keelespetsiifilist tarkvara</p> <p>2. Kasutada tagasisidet ja alternatiivide võrdlemist</p> <p>3. Kombineerida erineval moel treenitud programmiversioone ja valida väljund mitme variandi hulgest.</p> <p>Masintõlke parandamine tagasiside kaudu toimuks tagasiside vormi kaudu. Projekti tulemusel loodav tõlkija on kõigile interneti kaudu tasuta kasutatav. Projekti käigus loodavad lingvistiliselt töödeldud korpused on samuti kõigile vabalt kasutatavad.</p>	<p>Projekti lõpptulemused</p> <p>On loodud järgmised veebirakendused:</p> <p>1. Eesti-inglise masintõlke süsteemi demoversioon (pluss võrdlus Mosese ja Google'iga): <a href="http://masintolge.ut.ee/">http://masintolge.ut.ee/</a></p> <p>2. Statistilise masintõlke printsibi demo <a href="http://masintolge.ut.ee/mt-for-kids">http://masintolge.ut.ee/mt-for-kids</a></p> <p>3. Masintõlge vanase eesti keelde (printsibi demo): <a href="http://masintolge.ut.ee/oest">http://masintolge.ut.ee/oest</a></p> <p>4. Tõlkeabisüsteem: <a href="http://kde.teataja.ee/">http://kde.teataja.ee/</a></p>	3	1,5	4	2,2	48
EKKT09-65	Automaatne parafrasaaside leidmine ning sõnade ja lühifraaside tõlkimine paralleelkorpuste abil	2009	2010	2	Tartu Ülikool	348	<p>Käeoleva projekti raames on plaanis luua veebileidesega tööriist, mis võimaldab kasutajal sisestada sõna või fraasi ning päringule vastuseks saada kas tõlked valitud võõrkeeles või parafrasaadid lähtekeeles. Sarnane tööriist on olemas inglise keele ja mõnede teiste enamräägitavate keelte jaoks. Nimetatud tööriist on kasutatav veebiaadressil <a href="http://linearb.co.uk">http://linearb.co.uk</a>. Meie tahaksime luua sama funktsionaalsusega tööriista, mis hõlmaks ka eesti keelt.</p>	<p>Parafrasaaside-tõlgete leidmise veebileides paikneb praegu aadressil <a href="http://ats.cs.ut.ee/parafrasaad">http://ats.cs.ut.ee/parafrasaad</a> ning tarkvara ja korpused asuvad TÜ Arvutiteaduse instituudi serveris (<a href="http://ats.cs.ut.ee">ats.cs.ut.ee</a>)</p>	6	1,6	1	0,1	24

EKKT09-66	Nutika süvaveebi- ja veebiresursside kombineeriva infootsüsteemi prototüüp	2009	2010	2	Tartu Ülikool	177	Edendada tehnoloogiat süntaksi- ja semantikapõhise infootsüsteemi toetamiseks nii veebis kui desktop- rakenduses. Luuakse või korduvkasutatakse (taotluses EKKT 09-62, "Eesti keele semantika ressursid ja vahendid" planeeritavat) esimest versiooni eesti üldontoloogiast ning seatakse see spetsiifilisemate konkreetsete rakendusvaldkonna ontoloogiate ning teiste mõistete süsteemide komponentidega lähtudes nutika otsimise rakendusest.	Loodi lõppkasutajale orienteeritud süntaksi- ja semantikapõhise infootsüsteemi prototüüp, kus seatakse eesti ja inglise keele NER tehnoloogiad; infootsüsteem rakendab TMI ja andmeteenuste semantilisi kirjeldusi infopäringute toetamiseks. Loodud prototüüp jäi juurutamata lõppkasutajale suunatud infokeskkonnas (a'la eesti.ee, delfi.ee, neti.ee).	6	0,9	11	7,4	67
EKKT09-67	Eesti keele sõltuvusgrammatika arendamine ja osaliselt mittekorrekse eestikeelse teksti morfoloogiline ühestamine ja süntaktiline analüüs	2009	2010	2	Tartu Ülikool	872	1. Grammatikakorrektori tööversioon. 2. Suulise keele süntaksianalüsaatori arendamine. 3. Murdetekstide süntaktiline analüüs. 4. Interneti keele (uue meedia keele) süntaktiline analüüs. 5. Õppijakeele süntaktiline analüüs. 6. Sügavamate sõltuvusseoste tuvastamine.	1. Koostati grammativigade korpus, mis on kasutatav ainult tööriistades või väga rangete juurdepääsupiirangutega. Koostati 210 grammatikavigade tuvastamise reeglit, mis konteksti kontrollides märgendavad vigasesse grammatilisse konstruktsiooni kuuluva sõna. Peamiselt on keskendunud komavigade tuvastamisele. Süsteem leiab 95% vigadest korrektselt, 5% on valealame ja 7% vigu jääb leidmata. Link süsteemi veebidemole <a href="http://www.keeletehnoloogia.ee/projektid/eesti-keele-soltuvusgrammatika">http://www.keeletehnoloogia.ee/projektid/eesti-keele-soltuvusgrammatika</a> 2. Analüsaator on võimeline analüüsima kõnetuvastaja väljundit, kuid täpseid arvutusi korrektsuse kohta tehtud ei ole. 3. Kohandati reegleid erinevate murrete analüüsiks. Märkimisväärset tõusu analüüsi korrektsuses (mitmesus 10%, vigu 3-4%) ei tekkinud, kuid nüüd on korpus, millel reeglistik baseerub, oluliselt suurem (kokku ca 35000 sõna). Murdekorpuse analüüsiks kasutatav reeglistik on kättesaadav projekti kodulehelt. 4. Interneti keele (uue meedia keele) süntaktiline analüüs: Praegune reeglistik ja näited analüüsitud korpusest on kättesaadav projekti kodulehelt. 5. Õppijakeele süntaktiline analüüs: Õppijakeele vigu tuvastavad reeglid on hetkel grammatikakorrektori üldreeglibaasis. 6. Sügavamate sõltuvusseoste tuvastamine: On loodud grammatika sõltuvusstruktuuri tuvastamiseks. Kui korrektselt morfoloogiliselt ühestatud lihtlause analüüsi edukus on pea 100%, siis lisatõde nõuab keerulisema struktuuriga põimlause analüüsi. Enamasti suudetakse küll tuvastada sõltuvusstruktuur lause sees, kuid mitte osalasete vahel.	2	0,6	13	4,30	44
EKKT10-74	Eesti keele koondkorpuse esituse ja kasutusvõimaluste arendamine	2010	2010	1	Tartu Ülikool	550	Üleminek UTF-8-le ja XML-keelele. Koondkorpuse kasutajaliidesele kollokatsioonide leidja esialgse versiooni lisamine. Blogide allkorpuse loomise alustamine. Uue meedia leksika, ortograafia ja morfoloogia analüüs automaatse morfoloogilise analüüsi vajadustest lähtuvalt.	1) Koondkorpus on teisendatud ASCII -> UTF-8; SGML -> XML ja TEI P3 (st Text Encoding Initiative versioon 3) -> TEI P4 -> TEI P5. 2) Kollokatsioonid on sellised sõnapaarid, mille liikmed esinevad üksteise naabruses sagedamini kui võiks eeldada nende üksikult esinemise sageduste põhjal. Tüüpilised kollokatsioonid on püsühendid - näiteks ühendverbid või idiomatiliselt väljendid. Meie kollokatsioonide leidja esialgne versioon võimaldab otsida nii lemmade kui ka sõnavormide kollokatsioone, kuid kollokaadid st päritud lemma/sõnavormiga olulisel määral koos esinevad sõnad antakse alati sõnavormidena e tekstisõnadena. Kollokatsioonide leidja esialgne versioon asub <a href="http://www.rabauti.ee/clc">http://www.rabauti.ee/clc</a> ; sinna läheb viit cl.ut.ee esilehelt. 3) Analüüsi tulemusena on koostöös EKKT projektiga "Korpusepäring Keeleveebis" valminud: - kasutajasõnastik morfoloogiaanalüsaatorile etmrf - meetod ja programm kasutajasõnastiku automaatseks täiendamiseks - meetod ja programm tekstide eeltöötuseks	6	2,05	2	1,0	3
EKKT10-75	Piiratud loomuliku eesti keele mooduli prototüüp teadmushaldusplatvormidele	2010	2010	1	ELIKO Tehnoloogia Arenduskeskus OÜ	110	Projekti peamine eesmärk on luua masintöödeldava piiratud eesti keele (MTEK) raamistik, sh järgmised raamistiku komponendid: disainiprintsiibid, lihtsustatud grammatika ja (esimest järku loogikal põhinev) esitusviis, piiratud sõnavara eesti keele üldontoloogia baasil (korduvkasutatakse ja arendatakse edasi EKKT 2009 aasta projekti <a href="http://ats.cs.ut.ee/semantika/wiki/index.php/%C3%9Cldontoloogia">http://ats.cs.ut.ee/semantika/wiki/index.php/%C3%9Cldontoloogia</a> tulemeid), semantilise vikimootori prototüüp.	1. kirjeldatud on 4 raaltöödeldava eesti keele (RTEK) kasutuslugu OpenUP meetodikast lähtuvalt. <a href="http://www.keeletehnoloogia.ee/projektid/piiratud-loomuliku-eesti-keele-mooduli-prototuuup/kasutuslugu">http://www.keeletehnoloogia.ee/projektid/piiratud-loomuliku-eesti-keele-mooduli-prototuuup/kasutuslugu</a> 2. raaltöödeldava eesti keele omadused on määratletud ca 60-s piiratud loomulike keelte ontoloogia dimensioonis (vt <a href="http://www.keeletehnoloogia.ee/projektid/piiratud-loomuliku-eesti-keele-mooduli-prototuuup/piiratud-loomuliku-eesti-keele-omaduste-1">http://www.keeletehnoloogia.ee/projektid/piiratud-loomuliku-eesti-keele-mooduli-prototuuup/piiratud-loomuliku-eesti-keele-omaduste-1</a> ). 3. loetelu projektidest, teadusartiklitest, tarkvaraplattformidest, keeleressurssidest ja piiratud loomulike keelte esitlusviisidest koos hinnanguga nende kasutatavusest RTEKI loomiseks (vt <a href="http://www.keeletehnoloogia.ee/projektid/piiratud-loomuliku-eesti-keele-mooduli-prototuuup/rtek-loomisel-korduvkasutatavad-ressursid">http://www.keeletehnoloogia.ee/projektid/piiratud-loomuliku-eesti-keele-mooduli-prototuuup/rtek-loomisel-korduvkasutatavad-ressursid</a> ). 4. projekti käiku ja tulemeid kajastab avalik veebisait <a href="http://www.keeletehnoloogia.ee/projektid/piiratud-loomuliku-eesti-keele-mooduli-prototuuup">http://www.keeletehnoloogia.ee/projektid/piiratud-loomuliku-eesti-keele-mooduli-prototuuup</a>	3	0,8	3	1,9	17

#### Lisa 4. EKKTT juhtkomitee

- Jaak Vilo - komisjoni esimees, Tartu Ülikool, matemaatika-informaatikateaduskond, Arvutiteaduse instituut, erakorraline vanemteadur (intervjueeriti SE käigus);
- Heiki-Jaan Kaalep - Tartu Ülikool, filosoofiateaduskond, eesti ja soome-ugri keeleteaduse osakond, vanemteadur (intervjueeriti SE käigus);
- Einar Meister - Tallinna Tehnikaülikool, Küberneetika Instituut, foneetika ja kõnetehnoloogia labori juhataja (intervjueeriti SE käigus);
- Kaili Müürisep - Tartu Ülikool, matemaatika-informaatikateaduskond, Arvutiteaduse instituut, keeletehnoloogia teadur;
- Karl Pajusalu - Tartu Ülikool, filosoofiateaduskond, eesti keele ajaloo ja murrete professor, eesti ja soome-ugri keeleteaduse osakonna juhataja;
- Indrek Reimand - Haridus- ja Teadusministeerium, teadusosakonna juhataja;
- Urmas Sutrop - Eesti Keele Instituut, direktor (intervjueeriti SE käigus);
- Uuno Vallner - Majandus- ja Kommunikatsiooniministeerium, riigi infosüsteemide osakond, infotehnoloogia arhitektuuri talituse juhataja (intervjueeriti SE käigus);
- Kadri Vider - Haridus- ja Teadusministeerium, teadusosakonna peaekspert.

**Lisa 5. Projektide põhitäitjate kattuvus SF-i ja ETF-i põhitäitjatega 2010. aastal (allikas: ETIS).  
 Numbritega on tähistatud vastavate projektide arv, kus EKKTT projekti põhitäitja osaleb  
 põhitäitjana kas SF või ETF teema/projekti täitmisel.**

Eesnimi	Perekonnanimi	SF osalus	ETF osalus
Rene	Altrov		
Tanel	Alumäe	1	1
Madis	Arukask	1	1
Andrei	Dementjev		
Lauri	Eesmaa		
Pille	Eslon		2
Mark	Fišel		1
Mark	Fišel		1
Olga	Gerassimenko		1
Küllli	Habicht	1	1
Indrek	Hein		
Tiit	Hennoste	1	
Ingmar	Jaska		
Heiki- Jaan	Kaalep	1	
Heiki- Jaan	Kaalep	1	
Neeme	Kahusk	1	
Neeme	Kahusk	1	
Neeme	Kahusk	1	
Riina	Kasterpalu		1
Kadri	Kerner		
Kadri	Kerner		
Indrek	Kiissel		
Valve- Liivi	Kingisepp		
Mare	Koit	1	1
Kristina	Koppel		
Jüri	Kuusik		
Peep	Küngas		
Krista	Liin		
Krista	Liin		
Pärtel	Lippus		1
Andres	Loopmann		
Martin	Luts		
Martin	Luts		
Erkki	Luuk	1	
Erika	Matsak		
Einar	Meister	1	
Einar	Meister	1	
Lya	Meister		
Lya	Meister		
Krista	Mihkels		2
Meelis	Mihkla	1	1

Kadri	Muischnek	1	
Anni	Muru		
Kaili	Müürisep	1	
Helen	Nigol		1
Tõnis	Nurk		
Anni	Oja		1
Siim	Orasmaa		
Siim	Orasmaa		
Heili	Orav	1	1
Heili	Orav	1	1
Hille	Pajupuu	1	
Sirli	Parm		1
Pille	Penjam	1	1
Eesnimi	Perenimi		
Liisi	Piits		1
Liisi	Pool		
Liisi	Pool		
Kristiina	Praakli		
Esta	Prangel		
Küllli	Prillop		
Rene	Prillop		
Siiri	Pärkson		1
Siiri	Pärkson		1
Kristel	Ress		
Tiit	Roosmaa	1	1
Tiit	Roosmaa	1	1
Andriela	Rääbis		2
Monika	Saarmann		
Kati	Sein		
Hanna	Sinijärv		
Raul	Sirel		
Kairit	Sirts		
Oleg	Šelajev		
Kairi	Tamuri		1
Piia	Taremaa		
Ain	Teesalu		
Mark	Tehver		
Pire	Teras	1	1
Daniel	Tikkerbär		
Karol	Toompalu		
Maarika	Traat	1	
Maarika	Traat	1	
Margus	Treumuth		1
Katrin	Tsepelina		
Tuuli	Tuisk		2
Kristel	Uiboaed		1
Elle	Vaimann		
Kaarel	Veskis		

Kaarel	Veskis		
Ülle	Viks		
Jaak	Vilo	1	1
Asta	Õim	1	1
Haldur	Õim	1	1
Katre	Õim	1	1



**Lisa 6. Projektide tulemusena kajastatud kaitstud doktoritööde kattuvus teiste rahastamisinstrumentide kaudu rahastatud projektide aruandluses kajastatuga (allikas: ETIS).**

Projekti number	Doktorant		Dokoritöö pealkiri	Paralleelsed projektid			
	Eesnimi	Perenimi		I	II	III	IV
EKKTT06-11	Heili	Orav	Isiksuseomaduste sõnavara semantika eesti keeles	EKKTT06-11	EKKTT09-63	ETF6744	SF0182541s03
EKKTT06-8	Raili	Pool	Eesti keele teise keelena omandamise seaduspärasusi täis- ja osasihitise näitel	EKKM09-115	EKKTT06-8		
EKKTT09-57	Andriela	Rääbis	Eesti telefonivestluse sissejuhatus: struktuur ja suhtlusfunktsioonid	EKKTT09-57			
EKKTT09-63	Ann	Veisman	Eesti keele kaas- ja mäarsõnade semantika võimalusi	EKKTT09-63	ETF7552		
EKKTT09-66	Riina	Maigre	Hajusrakenduste automaatne kompositsioon	EKKTT09-66			
EKKTT06-2	Meelis	Mihkla	Kõne ajalise struktuuri modelleerimine eestikeelsele tekst-kõne sünteesile	EKKTT06-2			
EKKTT06-13	Pille	Penjam	Eesti kirjakeele da- ja ma-infinitiiviga konstruktsioonid	EKKTT06-13	ETF6510	EKKM09-88	
EKKTT06-10	Kadri	Muischnik	Verbi ja noomeni püsiühendid eesti keeles	EKKTT06-10	SF0182541s03	SF0182571s03	ETF5787

EKKTT0 6-3	Margit	Langemets	Nimisõna süstemaatiline polüseemia eesti keeles ja selle esitus eesti keelevaras	EKKTT0 6-3			
EKKTT0 6-18	Kaarel	Kaljurand	Attempto Controlled English as a Semantic Web Language	EKKTT0 6-18			
EKKTT0 7-21	Kadri	Kerner	Valdkonnaspetsiifilise ühetähenduslikkuse tõhusus sõnatähenduste ühestamisel	EKKTT0 7-21	EKKTT09-62	EKT2	EKT12
EKKTT0 9-63	Erkki	Luuk	The noun/verb and predicate/argument structures	EKKTT0 9-63			

Lisa 8: EKKTT projektide kulud kuluartiklite lõikes.

Projekt / kuluartikkel		Töötasu ja maksud	Soetused	Lähetused	Muud tegevuskulud	Üldkululõiv	Kokku
EKKTT06-1	RP	2 289 820	230 451	187 279	268 660	565 268	3 541 478
EKKTT06-1	MUU	0	0	0	0	0	0
EKKTT06-10	RP	809 692	0	0	94 308	216 000	1 120 000
EKKTT06-10	MUU	125 641	0	0	0	0	125 641
EKKTT06-11	RP	941 533	54 045	142 532	63 890	288 000	1 490 000
EKKTT06-11	MUU	23 000		20 174	11 923		55 097
EKKTT06-12	RP	692 401	49 849	0	109 350	208 400	1 060 000
EKKTT06-12	MUU	26 312	0	0	0	0	26 312
EKKTT06-13	RP	513 544	0	47 111	54 579	143 816	759 050
EKKTT06-13	MUU	770 000	20 000	14 500	16 500	16 000	837 000
EKKTT06-14	RP	1 490 455	114 817	34 355	28 400	339 747	2 007 774
EKKTT06-14	MUU	94 480	0	0	0	0	94 480
EKKTT06-15	RP	591 771	45 000	73 860	298 369	241 000	1 250 000
EKKTT06-15	MUU	0	0	0	0	0	0
EKKTT06-16	RP	2 285 873	172 560	32 609	53 163	533 431	3 077 635
EKKTT06-16	MUU	0	0	0	0	0	0
EKKTT06-17	RP	645 493	45 159	40 448	148 013	165 887	1 045 000
EKKTT06-17	MUU					66 000	66 000
EKKTT06-18	RP	682 155	69 839	233 059	28 920	226 027	1 240 000
EKKTT06-18	MUU					74 000	74 000
EKKTT06-2	RP	3 044 061	312 862	147 026	176 460	767 331	4 447 740
EKKTT06-2	MUU	0	0	45 000	0	0	45 000
EKKTT06-3	RP	4 757 838	264 100	144 462	218 433	1 048 347	6 433 180
EKKTT06-3	MUU	0	0	0	0	0	0

Projekt / kuluartikkel		Töötasu ja maksud	Soetused	Lähetused	Muud tegevuskulud	Üldkululõiv	Kokku
EKKTT06-4	RP	1 247 300	0	0	0	0	1 247 300
EKKTT06-4	MUU	0	0	0	0	0	0
EKKTT06-5	RP	1 574 991	238 094	212 127	103 962	443 376	2 572 550
EKKTT06-5	MUU	0	0	0	0	0	0
EKKTT06-6	RP	1 585 381	688 234	72 316	168 000	388 899	2 902 830
EKKTT06-6	MUU	0	0	0	0	0	0
EKKTT06-7	RP	1 726 895	529 104	192 850	49 322	502 904	3 001 075
EKKTT06-7	MUU	0	0	0	0	0	0
EKKTT06-8	RP	655 090	45 846	28 655	36 169	170 360	936 120
EKKTT06-8	MUU	0	0	0	0	0	0
EKKTT06-9	RP	814 887	61 340	118 278	41 495	244 000	1 280 000
EKKTT06-9	MUU	105 721	0	0	0	0	105 721
EKKTT07-21	RP	1 163 253	46 339	76 816	3 846	273 416	1 563 671
EKKTT07-21	MUU	0	0	0	0	0	0
EKKTT07-23	RP	349 077	0	2 923		88 000	440 000
EKKTT07-23	MUU						0
EKKTT08-26	RP	1 082 798	36 966	91 972	32 289	0	1 244 025
EKKTT08-26	MUU	0	0	0	0	0	0
EKKTT08-28	RP	746 837	19 557	199 677	149 411	237 829	1 353 311
EKKTT08-28	MUU	0	0	0	0	0	0
EKKTT08-29	RP	465 110	45 195	3 215	416	0	513 935
EKKTT08-29	MUU	0	0	0	0	0	0
EKKTT09-57	RP	764 482	0	20 460	30 175	163 023	978 140
EKKTT09-57	MUU	0	0	0	0	0	0
EKKTT09-61	RP	634 178	0	1 243	70 390	136 458	842 269
EKKTT09-61	MUU	0	0	0	0	0	0

Projekt / kuluartikkel		Töötasu ja maksud	Soetused	Lähetused	Muud tegevuskulud	Üldkululõiv	Kokku
EKKTT09- 62	RP	388 627	22 495	61 632	10 995	92 050	575 800
EKKTT09- 62	MUU	0	0	0	0	0	0
EKKTT09- 63	RP	758 881	24 967	115 227	42 155	180 767	1 121 997
EKKTT09- 63	MUU	0	0	0	0	0	0
EKKTT09- 64	RP	1 143 356	36 890	41 523	32 210	241 838	1 495 818
EKKTT09- 64	MUU	0	0	0	6 890	0	6 890
EKKTT09- 65	MUU	0	0	0	0	0	0
EKKTT09- 66	RP	847 282	0	0	133 752	196 206	1 177 240
EKKTT09- 66	MUU	2 856	0	0	0	0	2 856
EKKTT09- 67	RP	622 228	0	84 588	19 584	145 280	871 680
EKKTT09- 67	MUU	0	0	0	0	0	0
EKKTT10- 74	RP	359 271	30 719	45 845	22 498	91 667	550 000
EKKTT10- 74	MUU	0	0	0	0	0	0
EKKTT10- 75	RP	0	0	0	110 000	0	110 000
EKKTT10- 75	MUU	0	0	0	0	0	0