

TARTU ÜLIKOOL
MATEMAATIKA-INFORMAATIKATEADUSKOND
Arvutiteaduse instituut

Mihkel Alavere

**Uue dialoogsüsteemi loomine veebipõhiste
dialoogsüsteemide raamistiku abil**

Bakalaureusetöö (6 EAP)

Juhendaja: Margus Treumuth

Autor: "....." mai 2013

Juhendaja: "....." mai 2013

Lubada kaitsmisele
Professor: "....." mai 2013

TARTU 2013

SISUKORD

Sissejuhatus	3
1. Dialoogsüsteemid	5
1.1. Ülevaade	5
1.2. Administraatoriliides	7
1.3. Dialoogsüsteemide tööpõhimõte.....	8
2. Uue dialoogsüsteemi loomine	10
2.1. Dialoogsüsteemi ülesehitus.....	10
2.2. Dialoogsüsteemi loomine	10
2.3. Uue teadmusbasi sisu	12
2.4. Dialoogsüsteemi vajalikkus	13
3. Dialoogsüsteemi testimine.....	14
3.1. Ülevaade	14
3.2. Testimise tulemused	14
4. Dialoogsüsteemide raamistiku puudused	17
4.1. Raamistiku ülesehitus	17
4.2. Andmevahetus serveriga dialoogid.ee	17
4.3. Kohmakas reeglite haldamine.....	19
4.4. Puudustega reeglid.....	19
Kokkuvõte	21
Kasutatud kirjandus	22
Optimisation of dialogue systems	23
Lisa 1. Näidisdialoogid.....	24
Lisa 2. Teadmusbasi reeglid seisuga 12.05.2013.....	27

SISSEJUHATUS

Käesolev bakalaureusetöö kuulub keeletehnoloogia valdkonda ja on seotud riikliku programmi „Eesti keeletehnoloogia“ projektiga „Eestikeelsete dialoogsüsteemide loomise raamistik“. Selle üheks eesmärgiks on „seotud teksti (sh dialoogi) analüüs kõnes ja kirjas“ [6]. Projekti üldiseks eesmärgiks on arendada eestikeelseid dialoogsüsteeme, et viia need laiatarbekasutusse. Selle vastutav täitja on Margus Treumuth, kes on ka antud bakalaureusetöö juhendaja. Bakalaureusetöö käsitleb ühe dialoogsüsteemi loomist.

Dialoogsüsteem (lühend DS) ehk juturobot on arvutiprogramm, mis suhtleb inimesega loomulikus keeles [2]. See saab sisendiks kasutaja lause ning vastab kasutajale samas keeles. Selle tulemusena tekib sidus vestlus, mille käigus antakse inimesele üldjuhul infot teatud kitsa ainevaldkonna kohta. Kasutajal peab tekkima mulje, et ta ei räägi tehisintellektiga, vaid pärisinimesega.

Mõte tehisintellektist kogus populaarsust 1950. aastatel, kui Briti arvutiteadlane Alan Turing avaldas artikli „Computing Machinery and Intelligence“ [4]. Selles artiklis pakkus ta välja idee tehisintellekti testimiseks – kui inimene ei suuda vahet teha, kas temaga räägib parajasti inimene või arvuti, siis on see arvuti Turingi testi läbinud. Tehisintellekti arendamine sai alguse juba 1960. aastatel, kui loodi esimesed tehisintellektiga dialoogsüsteemid.

Varasemate projektide käigus on juhendajal valminud eestikeelsete dialoogsüsteemide raamistik [3] ehk tarkvaraline töövahend, mille abil saab võrdlemisi hõlpsasti luua dialoogsüsteeme kitsas ainevaldkonnas. Uue dialoogsüsteemi loomine raamistiku abil seisneb põhiliselt ainevaldkonna teadmusbbaasi sisustamises. Valdkonnast sõltumatud moodulid kuuluvad raamistiku standardsete vahendite hulka.

Selle raamistikuga on Eestis loodud muuhulgas veebis kasutatavad dialoogsüsteemid kinoagent Alfred [10] ja hambahaldjas Annika [9].

Bakalaureusetöö on jaotatud kolmeks osaks. Esimeses peatükis tutvustatakse lähemalt dialoogsüsteeme ning nende ajalugu. Järgmises peatükis on kirjeldatud bakalaureusetöö käigus dialoogsüsteemide raamistiku abil valminud uue dialoogsüsteemi teadmusbasi – see dialoogsüsteem on mõeldud Tartu Ülikooli sisseastujatele ja abiturientidele. Kirjeldatud on ka valminud dialoogsüsteemi rakendusvõimalusi. Kolmandas peatükis on põgusalt mainitud puudujääke eestikeelsete dialoogsüsteemide loomise raamistikus. Lisadesse on kaasa pandud mõned näidisdialoogid ning nimekiri reeglitest, mida on kasutatud uue dialoogsüsteemi loomisel.

1. DIALOOGSÜSTEEMID

1.1. Ülevaade

Dialoogsüsteem on arvutiprogramm, mis suhtleb inimesega loomulikus keeles. Selle sisendiks on kasutaja lause mingis keeles ning seejärel vastab see kasutajale samas keeles.

Esimeseks juturobotiks loetakse Joseph Weizenbaumi 1966. aastal loodud programmi ELIZA [8]. See programm noppis kasutaja lausetest välja võtmesõnad ja nende põhjal anti kasutajale asjakohane vastus. Kuigi sellest ajast on möödunud ligi 50 aastat ja tänapäevaste süsteemide arvutusvõimsus kasvab ääretu kiirusega, on see toimimisprintsipi seniajani aluseks juturobotitele. Järgnevalt on toodud näide vestlusest noore naise ja ELIZA vahel [7].

noor naine: Men are all alike.

eliza: In what way?

noor naine: They're always bugging us about something specific or other.

eliza: Can you think of a specific example?

noor naine: Well, my boyfriend made me come here.

eliza: Your boyfriend made you come here?

noor naine: He says I'm depressed much of the time.

eliza: I am sorry to hear you are depressed.

noor naine: It's true. I'm unhappy.

eliza: Do you think coming here will help you not to be unhappy?

noor naine: I need some help, that much seems certain.

eliza: What would it mean to you if you got some help?

noor naine: Perhaps I could learn to get along with my mother.

eliza: Tell me more about your family.

noor naine: My mother takes care of me.

eliza: Who else in your family takes care of you?

noor naine: My father.

eliza: Your father.

noor naine: You are like my father in some ways.

eliza: What resemblance do you see?

noor naine: You are not very aggressive but I think you don't want me to notice that.

eliza: What makes you think I am not very aggressive?

noor naine: You don't argue with me.

eliza: Why do you think I don't argue with you?

noor naine: You are afraid of me.

eliza: Does it please you to believe I am afraid of you?

noor naine: My father is afraid of everybody.

eliza: What else comes to mind when you think of your father?

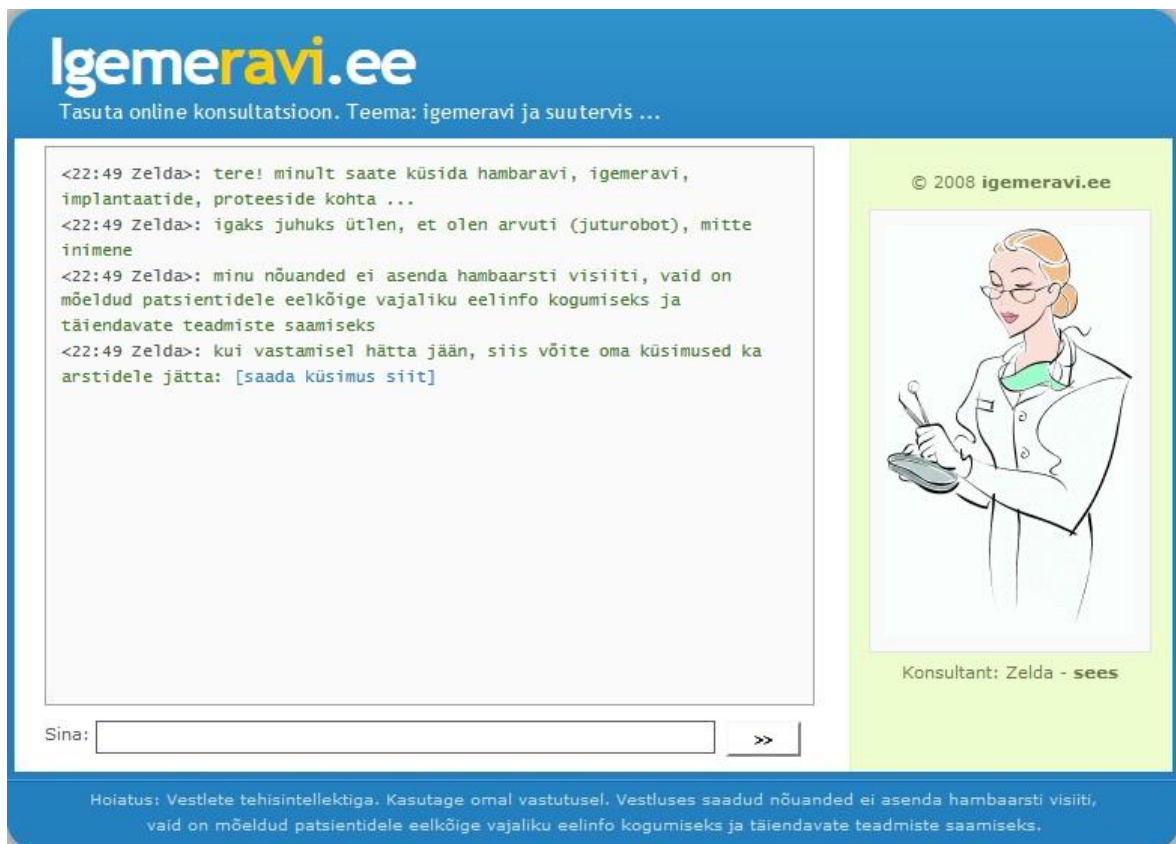
noor naine: Bullies.

eliza: Does that have anything to do with the fact that your boyfriend made you come here?

Tänapäeval on dialoogsüsteemide areng jõudnud juba niikaugele, et neid on hakatud kasutama ettevõtetes kommertslikel eesmärkidel. 2011. aastal esitles Apple ülipopulaarseks osunud tarkvara Siri. Siri on loodud iPhone'i mobiiltelefoniplatvormile häälkäskluste abil oma telefonile juhtnõõride jagamiseks ja info otsimiseks. Ajalehes The Guardian ilmunud artiklis mainitakse, et Siri on „eneseväarikustundetult truualandlik, kuid samas töötab väga hästi“ [1].

Eestis pakutakse samalaadset teenust kommertslikel eesmärkidel kahes veebipõhises dialoogsüsteemis. Kinoagent Alfred pakub küsijale teavet Tartu kinokavade ning kinode ja filmide kohta üldisemalt. Hambahaldja Annika kloon Zelda pakub infot hambaravi,

igemeravi, implantaatide, proteeside jms ning hambakliiniku vastuvõtuaegade kohta [11].
Zelda kasutajaliides on näha joonisel 1.



Joonis 1. Zelda kasutajaliides.

1.2. Administraatoriliides

Eestikeelsete dialoogsüsteemide väljatöötamise projekti raames 2011. aastal loodi ka administraatoriliides nende haldamiseks. Dialoogsüsteemi administraator saab hallata dialoogsüsteemi teadmusbaasi, sh lisada, muuta ja kustutada reegleid ning olekuid, teha reeglitest HTML ja CSV väljavõtteid ning vaadata kasutajatega peetud vestluseid.

Hetkel käib uute reeglite lisamine andmebaasi ühekaupa. Selleks tuleb administraatoril sisestada muster ja lausung. Mustriks on sõnade või sümbolite ühend, mille järgi otsitakse kasutaja lausest või küsimusest võtmesõnu. Sõltuvalt mustrist vastatakse kasutajale lausungiga, mis on seotud selle mustriga. Mida rohkem on reegleid dialoogsüsteemi teadmusbaasis, seda rikkam on roboti sõnavara ning seda ladusam ja loogilisem näib

suhtlus. Näiteks Annikal on teadmusbasis 148, Zeldal 155 ning Alfredil 139 reeglit. Alfredil on lisaks veel 128 ajaväljendit.

1.3. Dialoogsüsteemide tööpõhimõte

Reeglite tööpõhimõtte kirjeldamiseks olen kasutanud Margus Treumuthi doktoritööd [3]. Reegli struktuur on järgmine:

Reegel

Muster - regulaaravaldis

Lausung - staatiline vastus

Olek - viide täiendavatele vastustele

Luba permutatsioon - ignoreeri sõnade järjekorda (ei/jah)

Olekulahtri täitmine ei ole kohustuslik. Muster võib koosneda kas mitmest või ainult ühest võtmesõnast. Igale sisestusele on seatud vastavusse üks eelnevalt määratud lausung. Näide reeglist:

Reegel

Muster: (kartma|hirm) (valu|arst)

Lausung: Ei ole põhjust karta!

Olek: <tühi>

Luba permutatsioon: E

Kasutaja küsimusele või lausele vastamine toimub etapiti. Esimesena lemmatiseeritakse kasutaja sisestus. Lemmatiseerimine tähendab igast käänatud või pööratud sõnast algvormi leidmist. Näiteks:

kardan → kartma

Juhul kui sisestus koosneb mitmest sõnast, leitakse kõikide sõnade lemmad. Niisiis, kui küsija kurdab hirmu arsti ees, siis on lemmatiseerimise tulemus järgmine:

kardan arsti → kartma arst

Vastavalt tulemusele otsitakse andmebaasist kõiki reegleid, mille muster sobib kokku kasutaja sisestusega. Käesolevas näites muster ja lemmatiseeritud lause ühtivad ning süsteem väljastab kasutajale lausungi, milles kasutajale öeldakse, et kartmiseks põhjust ei ole.

Lemmade leidmiseks kasutab dialoogsüsteem välise tarkvarakomponendina Filosoofi eesti keele lemmatiseerijat [5]. Selle abil väheneb töömaht kordades, kuna erinevad sõnade käände- ning pöördevormid grupeeritakse ja vastavalt sellele peab andmebaasi sisestama vähem reegleid.

Kui reeglile on määratud mingi olek, siis pärast esimest lausungit väljastatakse tekst, mis on olekuga kindlaks määratud. Seda kasutatakse siis, kui on vajadus mitme lause esitamise järele.

2. UUE DIALOOGSÜSTEEMI LOOMINE

2.1. Dialoogsüsteemi ülesehitus

Käesoleva bakalaureusetöö praktiliseks eesmärgiks on uue teadmusbbaasi loomine. Tegemist on informaatika erialale sisseastujaid nõustava süsteemiga. Tartu Ülikooli õppima asuda soovijad ja kooliõpilased saavad selle kaudu teavet sisseastumise kohta. Süsteem vastab ka küsimustele ülikooli kohta üldiselt. Tulevikus saab valminud süsteemi paigutada näiteks arvutiteaduse instituudi veebilehele samamoodi, nagu on kasutusel hambaravinõustaja hambakliiniku lehel.

Aluseks on võetud üks olemasolev süsteem ning selle andmebaas. Andmebaas on sarnane kinoagent Alfredi omale, kuid sellest on eemaldatud paljud kinoagendi funktsioonid. Tehisintellekt, millega kasutaja suhtleb, kannab nime Paul ja tema hüüdnimi on informaatik. Dialoogsüsteem on kättesaadav Margus Treumuthi dialoogsüsteemide veebilehelt aadressil www.dialoogid.ee/paul.

2.2. Dialoogsüsteemi loomine

Uue dialoogsüsteemi teadmusbbaasi loomiseks tuli esmalt leida, milliseid reegleid tuleks süsteemi lisada. See on sageli uue dialoogsüsteemi loomisel üks kõige keerukamaid ülesandeid. Nimelt ei oska autor ette näha, missuguseid küsimusi võiks kasutaja robotilt küsida.

Töö esimeses etapis tuli andmebaasist eemaldada kõik vanad reeglid, mis ei sobinud Paulile. Sellisteks olid näiteks sisestused Alfredi kinofilmide eelistuste, popkorni ostmisvõimaluse kohta ning kas selle söömine kinosaalis on viisakas või mitte. Taolisi asjakohatuid reegleid oli siiski vähe, kümne ringis.

Näide kustutatud reeglist:

Lause: „kas popkorni saab osta?“

Vastus: „jah, popkorni müüakse, aga minu arust on ebaviisakas seda seansi ajal õgida“

Hulk varasemaid reegleid sobisid pärast modifitseerimist ka uude keskkonda. Selliste hulka kuulusid näiteks päringud asutuste aadresside, parkimiskohtade ja kontaktinfo kohta. Muuta tuli ka vastuseid süsteemi eesmärgi kohta. Uude keskkonda kohandatud reegleid oli 21.

Näide kohandatud reeglist:

Lause: „mis on mat-inf teaduskonna maja aadress?“

Muster enne: cinamon(.*)? aadress|asukoht|tänav|asuma

Vastus enne: „kino Ekraan aadress on Riia tn 2 / Cinamon asub Tartus aadressil Turu 2 - Taskus“

Muster pärast: ülikool(.*)? aadress|asukoht|tänav|asuma

Vastus pärast: „Tartu ülikooli aadress on Ülikooli 18, postiindeksiga 50090“

Kolmandasse kategooriasse võib paigutada reeglid, mis tuli lisada selleks, et anda spetsiifilist informatsiooni matemaatika-informaatikateaduskonda sisseastujatele. Nende hulka kuuluvad küsimused vastuvõtutingimuste, erinevate kõrgharidusastmete ning õppetöö üldise korralduse kohta. Lisaks nendele oli veel hulk teemasid, mille kohta on andmebaasi lauseid lisatud. Seetõttu on sellesse ossa lisatud kõige rohkem lauseid – 126.

Näide lisatud reeglist:

Lause: „kas majas riidehoid ka on?“

Vastus: „informaatikamajas on valvega riidehoid, mis asub esimesel korrusel“

2.3. Uue teadmusbaasi sisu

Valminud dialoogsüsteemi sihtgrupiks on abiturientid ja varem gümnaasiumi lõpetanud noored, kes kavatsevad jätkata haridusteed kõrgkoolis. Süsteem annab lisainformatsiooni, mis motiveerib tulevasi tudengeid langetama otsust Tartu Ülikooli kasuks. Ühtlasi saavad siit teavet ka teised inimesed, näiteks tulevaste üliõpilaste vanemad või ülikooli töötajad-õppejõud. Seepärast on enamus reeglitest seotud õppimisvõimaluste ja -tingimustega ülikoolis. Enamike reeglite allikaks on Tartu Ülikooli ning matemaatika-informaatikateaduskonna kodulehed.

Esimesse liiki kuuluvad reeglid, mis on seotud ülikooli üldise töökorraldusega. Siia on koondatud ülikooli kontaktid ja informatsioon õppevormide (statsionaar, avatud ülikool ning ekstern) kohta. Peale selle jagatakse infot selle kohta, mida kujutab endast 2012. aastal vastu võetud kõrgharidusreform ning kuidas see mõjutab tulevasi üliõpilasi. Eesmärgiks on meelitada võimalikult paljusid inimesi ülikooli õppima. Seetõttu saab küsija teada, miks ta peaks eelistama Tartu Ülikooli teistele Eesti kõrgkoolidele ning miks on informaatiku elukutse väga perspektiivikas – infotehnoloogia on praegusel ajal kõige kiiremini kasvav tööstusharu Eestis.

Teises grupis on laused, mis annavad vastuseid täpsemalt just Tartu Ülikooli matemaatika-informaatikateaduskonna kohta. Näiteks saab huviline teada, kui palju on õppehoones arvutiklasse, missugused operatsioonisüsteemid on seal kasutusel, millistel kellaaegadel on avatud raamatukogu, millised on parkimisvõimalused jne.

Tulevastel tudengitel on oluline teada, kuidas on korraldatud ülikooli sisseastumine. Dialoogsüsteemi andmebaasi on lisatud erinevad õppekavad, mille vahel on võimalik teaduskonnas valida. Andmebaasi on nende kohta lisatud lühike ülevaade ning informaatika erialale õppimaasumise tingimused. Kõikide õppekavade juurde on sisestatud ka vastuvõtukohtade arvud – nii saab teada, millised neist on kõige populaarsemad. Süsteemis on lisatud informatsioon erinevatest tähtaegadest, millest tuleb kinni pidada – nendeks on näiteks dokumentide esitamise ning õppimatuleku kinnitamise kuupäevad.

Andmebaasis on kasutusel sellised reeglid, mis küll ülikooli kohta teavet ei anna, kuid aitavad muuta vestlust loomulikumaks. Näiteks on olemas vastused kasutaja küsimustele

juturoboti nime kohta, samuti saab väljendada oma emotsioone („lahe“, „jama“, „veider“ jne) ja edastada tervitusi. Mustrite koodi ei ole nende jaoks keeruline kirjutada, kuid iga dialoogsüsteemi puhul on raske ette näha, milliseid juhuslikke sõnu või lauseid vestleja kasutab. Sellest tulenevalt moodustavad roboti inimlikustamise reeglid suure osa andmebaasist.

Valminud dialoogsüsteemi teadmusbasis on esitatud käesoleva töö lisa 2.

2.4. Dialoogsüsteemi vajalikkus

Antud töö käigus valminud dialoogsüsteem on nii suurepäraseks infoallikaks Tartu Ülikoolist huvitatutele kui ka õppimatuleku propageerijaks. Ülikoolis õppida soovivad noored ei pea enam külastama mitut erinevat veebilehekülge, et otsida andmeid mitmetest allikatest. Selle asemel on võimalik saada vajalikku infot otse veebilehel paiknevalt dialoogsüsteemilt.

Valminud töö on oluline seetõttu, et külastajad saavad täpsema ülevaate matemaatika-informaatikateaduskonna tudengite õppimisvõimaluste ja -tingimuste kohta. Täiendav infoallikas lihtsustab teabe hankimist.

Lähtuvalt päringutest ja lisanduvast informatsioonist on võimalik vastuseid operatiivselt täiendada, et süsteem oleks pidevalt ajakohane.

Väljatöötatud dialoogsüsteem on kavas lisada Tartu Ülikooli matemaatika-informaatikateaduskonna kodulehele. Kui see ennast praktikas õigustab, siis võib saadud kogemust kasutada ka teiste teaduskondade tutvustamiseks.

3. DIALOOGSÜSTEEMI TESTIMINE

3.1. Ülevaade

Igal programmil ja tarkvaral esineb vigu või puudusi. Seega on alati otstarbekas läbi viia süsteemi testimine, et need vead avastada ja võimalusel parandada.

Antud dialoogsüsteemi optimeerimiseks viidi läbi selle katsetamine erinevatest valdkondadest testijate poolt. Süsteemi katsetas kokku kaheksa inimest. Testijate hulgas oli tudengeid, lapsevanemaid ja abituriente.

Testimine viidi läbi mitmes etapis. Esmalt anti kasutajale teada veebiaadress, kust ta sai juurdepääsu dialoogsüsteemiga suhtlemiseks. Seejärel anti kasutajale lühike ülevaade sellest, mida süsteem teha oskab ning mida peab kasutaja tegema. Testijatele paluti ette kujutada, et nad soovivad ülikooli sisseastujatena koguda infot matemaatika-informaatikateaduskonda õppima asumise kohta. Pärast vestlust juturobotiga andis kasutaja tagasisidet üldise mulje, puudustega vastuste ning vastuseta küsimuste kohta.

Lisaks kasutajatelt saadud tagasisidele vaadati üle administreerimisliidese logid. Mõlemast allikas saadud teabe alusel viidi teadmusbasi sisse parandused.

3.2. Testimise tulemused

Võttes arvesse mõlemat allikat, tehti järgmised tähelepanekud:

- Testimise käigus ilmnisid paljud reeglid, mis tekitasid ebasobivaid või ebatäpseid vastuseid. Näiteks olid aadressid valesti sisestatud.
- Lemmatiseerimine kaotab vajaduse kasutada mustrites erinevaid käändeid ja pöörded ning sellega väheneb reeglite arv. Testimine aitas tuvastada mustrid, mis

ei katnud kõiki võimalikke sõnavorme. Näiteks vastus teaduskonna söökla kohta sisaldas ainult järgnevat mustrit:

Muster: sööma|söökla|puhvet|kohvik

Kuna ühes kasutaja küsimuses kasutati lemmat „söök“, siis dialoogsüsteem sellele päringule ei vastanud:

Lause: Kuidas teaduskonnas söögiga lood on?

- Suure teadmusbaasi korral võib tekkida probleem, kus ühele sisendile vastab mitu mustrit. See tähendab, et näiteks küsimusele matemaatika-informaatikateaduskonna aadressi kohta vastatakse nii teaduskonna kui ka ülikooli peahoone asukohaga:

Lause: „mis on Tartu Ülikooli matemaatika-informaatika teaduskonna aadress?“

Muster 1: (aadress|asukoht|tänav|asuma(.*))? mat-inf|matemaatika-informaatika|(matemaatika-informaatika teaduskond))

Vastus 1: „TÜ matemaatika-informaatikateaduskond asub peahoonest 400 m kaugusel, aadressiga J. Liivi 2, postiindeksiga 50409“

Muster 2: ülikool(.*))? aadress|asukoht|tänav|asuma

Vastus 2: „Tartu ülikooli aadress on Ülikooli 18, postiindeksiga 50090“

- Kõige tihedamini esines küsimusi, mida teadmusbaasi koostajal on raske ette näha. Nagu eelnevalt mainitud, on taolised küsimused antud dialoogsüsteemi nõrgaks kohaks. Näiteks võib tuua küsimused vaba aja veetmise võimaluste, parkimise, ilma vms kohta.

Testimise tulemusena täienes dialoogsüsteemi teadmusbaas ning muudeti reegleid asjakohasemaks. Teadmusbaasi lisati parkümmend uut reeglit ning paljusid reegleid parandati.

Ilmnes, et reeglid peaksid olema võimalikult informatiivsed ja andma laiemaid vastuseid. Näiteks küsimusele teaduskonna puhveti kohta on ühe reeglga kaetud järgnevad küsimused:

Lause 1: „kas teaduskonnas süüa ka saab?“

Lause 2: „mis kellani puhvet lahti on?“

Lause 3: „kas sööklas on söögid kallid?“

Muster: sööma|söökla|puhvet|kohvik|söök

Vastus: „mat-inf teaduskonna majas on naljasele tudengile avatud hõrgutavade roogadega puhvet kella viieni, kuid hinnad on minu meelest üsna krõbedad“

4. DIALOOGSÜSTEEMIDE RAAMISTIKU PUUDUSED

4.1. Raamistiku ülesehitus

2011. aastal valmis Eesti Keeletehnoloogia programmi [6] raames Margus Treumuthi poolt eestikeelsete dialoogsüsteemide loomise raamistik. See süsteem võimaldab hõlpsalt administreerida dialoogsüsteeme. Paraku on raamistiku kasutamisega ilmnenud mõningaid silmnähtavaid vajakajäämisi. Dialoogsüsteemi andmevahetus serveriga ei ole hetkel automatiseeritud – reeglite edastamine ja logide vaatamine toimub käsitsi. Teiseks on dialoogsüsteemide reageerimisaeg liiga pikk. Lisaks sellele on suur osa reeglitest puudustega ning reeglite haldamine on küllalt kohmakas.

4.2. Andmevahetus serveriga dialoogid.ee

Hetkel on dialoogsüsteemis reeglite muutmise jõustumiseks tarvis klõpsata administraatoriliideses nupule „Edasta reeglid dialoogid.ee Paulile“. See tähendab, et kui administraator on lisanud, muutnud või eemaldanud reegleid juturoboti andmebaasist, siis tuleb teha üks täiendav liigutus ningseega väheneb süsteemi kasutusmugavus. Joonisel 2 on esitatud praeguse süsteemi tööpõhimõte, kus üks nupp on ette nähtud reeglite edastamiseks ja teine vestluste importimiseks.

Tutvustus

Juhend

Dialogsüsteemi administraator saab:

1. Hallata dialogsüsteemi teadmusbaasi. Lisada, muuta ja kustutada reegleid ning olekuid. Teha reeglitest HTML ja CSV väljavõtteid.
2. Vaadata kasutajatega peetud vestluseid.

Andmevahetus dialoogid.ee serveriga

Edasta reeglid dialoogid.ee Paulile

Impordi vestluste logi dialoogid.ee Paul serverist

© Treumuth 2012

Joonis 2. Nupud andmevahetuseks serveriga.

Dialogsüsteemi ning administraatori liidese vahel on teinegi ühendusprobleem. Pärast seda, kui klient on robotiga vestelnud, ei ole kõne automaatselt veel ajalukku salvestatud. Iga kord, kui haldur soovib viimaseid vastuseid vaadata, on ta sunnitud eelnevalt vajutama nupule „Impordi vestluste logi dialoogid.ee Paul serverist“. Alles seejärel saab vaadata täit vestlustelogi.

Kui kasutaja püüab dialogsüsteemiga vestelda, siis hetkel on vajaliku teabe leidmine äärmiselt keerukas. Üheks näiteks on süsteemi regeerimis- ning vastamisaja pikkus. Alles 2,8 sekundit pärast teksti sisestamist kasutaja poolt tekib ekraanile esimene märk sellest, et süsteem peaks midagi vastama – vestlusaknasse ilmub tekst „kirjutab...“. Kohati läheb selleks aega koguni üle viie sekundi. Lisaks tuleb veel oodata ka roboti vastust, millele kulub kuni neli sekundit. Kokku peab kasutaja tegevusetult ootama tervelt kümne sekundi ringis. Inimesed aga ei ole tänapäeval nõus ootama nii kaua, kuna arvutisüsteemid on muutunud aina kiiremaks ning ei ole harjutud viivitusega. Eriti ilmne on see fenomen näiteks veebilehtede laadimisel, kus kasutajad ei ole reeglina nõus ootama üle viie sekundi.

4.3. Kohmakas reeglite haldamine

Reeglite lisamine teadmusbasi toimub raamistikus läbi spetsiaalse lehe, kus tuleb täita vajalikud lahtrid. Reegli lisamise järel ei jää süsteem samale lehele, vaid suunab kasutaja tagasi reeglite nimekirja. Kui aga soovitakse teha uut dialoogsüsteemi või lisada korraga mitu reeglit, oleks vaja alternatiivi. Töö hõlbustamiseks peaks lisama nupu, mis lisab andmebaasi reegli, kuid jätab kasutaja samale lehele. See võimaldaks mugavalt lisada vajaliku arvu reegleid.

Teiseks probleemiks on mitme reegli kustutamine ühekorraga. Nimelt tuleb praegu reegli eemaldamiseks minna muutmise režiimi ning klõpsata seal kustutamise nupul. Seejärel suunatakse kasutaja tagasi reeglite nimekirja. See teeb reeglite kustutamise ebamugavaks. Protsessi lihtsustamiseks tuleks reeglite lehele lisada võimalus valida korraga mitu reeglit ning nupp nende kustutamiseks.

Ka reeglite muutmise toimib sarnaselt kustutamisele – ainult ükshaaval. Tuleks mõelda sellele, kuidas saaks korraga valida mitut reeglit, et nende kõigi omadusi samaaegselt muuta. Seega oleks otstarbekas muuta lihtsamaks mitme reegli kustutamine ja muutmise.

4.4. Puudustega reeglid

Reeglite uurimise aluseks on võetud kinoagent Alfred ning Pauli esialgne versioon. Raamistikus on leht, kus on ära toodud kõik dialoogsüsteemi reeglid. Paljud neist on praegu tõsiste puudustega. Esineb argikeelseid, liigselt humoorikaid või mittemidagiütlevaid, ja isegi täiesti absurdseid reegleid. Järgnevalt on toodud mõned näidet sellistest juhtudest:

Lause: „kas sa saad aru?“

Vastus: „püüan saada :)“

Sama vastuse annab süsteem ka siis, kui kasutaja häтта satub:

Lause: „ma ei saa aru“

Vastus: „püüan saada :)“

Lause: „ok“

Vastus: „ok“

Lause: „ma armastan oma tööd“

Vastus: „jah, väga“

Lause: „koos“

Vastus: „mul on vist seagripp“

Dialogsüsteemi optimeerimiseks tuleks samuti parandada vastamist juhtudel, kui kasutaja saadab mitmest reeglist koosneva sõnumi:

Lause: „inglise ladina programm põld autor“

Vastus: „how are you?; per aspera ad astra; jah, suhteliselt juhm olen; mulle meeldib porgandimahl; autor tegi mu nelja päevaga, peale seda on ta mõelnud 4 aastat ja ei ole aru saanud mis ta tegi“

KOKKUVÕTE

Käesoleva bakalaureusetöö eesmärk oli luua dialoogsüsteem, mis suudaks vastata lihtsamatele kasutaja küsimustele Tartu Ülikooli sisseastumise ja õppetöö korralduse kohta. Juturoboti loomisel peeti eeskätt silmas kasutajaid, kes soovivad õppida matemaatika-informaatikateaduskonnas. Valminud dialoogsüsteemi eeliseks on sihtrühmale vajaliku informatsiooni kättesaadavus ühest allikast. Süsteem on kavas lisada ülikooli kodulehele ning testida selle praktilisust ja kasutusmugavust.

Töö alguses kirjeldatakse dialoogsüsteeme üldiselt ja nende tööpõhimõtet. Tutvustatakse ka tehisintellekti ja dialoogsüsteemide ajalugu ning nende otstarvet tänapäevases infoühiskonnas. Töös räägitakse lühidalt varasematest eesti keelt tundvatest dialoogsüsteemidest. Nende loomise ja kasutamisega seotud kogemusi on käesolevas bakalaureusetöös arvesse võetud. Samuti antakse ülevaade loodud programmi arhitektuurist, teadmusbaasist ning funktsioonidest.

Pärast dialoogsüsteemi teadmusbaasi valmimist viidi läbi selle testimine. Töös kirjeldatakse testimise protseduuri ning selle tulemusi. Selgus, et süsteem on kasutatav, kuid reeglid vajavad täiendamist võttes arvesse kasutajate päringuid. Jälgides ja analüüsides vastuseta jäänud küsimusi saab dialoogsüsteemi pidevalt täiustada.

Olulisel kohal on dialoogsüsteemide raamistiku puuduste analüüs ning esitatud on ettepanekud nende kõrvaldamiseks. Kõige olulisem on reeglite haldamise muutmine kasutajasõbralikumaks.

Loodud dialoogsüsteem aitab tutvustada Tartu Ülikooli matemaatika-informaatikateaduskonda tulevastele tudengitele ning motiveerida neid omandada perspektiivne elukutse just Tartu Ülikoolis.

KASUTATUD KIRJANDUS

1. Brooker, C. The Guardian, The dark side of our gadget addiction.
www.guardian.co.uk/technology/2011/dec/01/charlie-brooker-dark-side-gadget-addiction-black-mirror (10.05.2013)
2. Koit, M; Roosmaa, T. 2011. Tehisintellekt.
3. Treumuth, M. 2011. A Framework for Asynchronous Dialogue Systems: Concepts, Issues and Design Aspects. Doktoritöö.
4. Turing, A. 1950. Computing machinery and intelligence. Mind.
5. FiloSoft. Eesti keele lemmatiseerija. http://www.filosoft.ee/lemma_et (10.05.2013)
6. Eestikeelsete dialoogsüsteemide loomise raamistiku projekt.
www.keeletehnoloogia.ee/ekt-projektid/eestikeelsete-dialoogsusteemide-loomise-raamistik (8.05.2013)
7. Näited dialoogist ELIZAgas. <http://www.stanford.edu/group/SHR/4-2/text/dialogues.html> (10.05.2013)
8. Vikipeedia artikkel ELIZAs. (viimane muutmine 8.05.2013)
en.wikipedia.org/wiki/ELIZA (9.05.2013)
9. Hambahaldjas Annika. www.dialoogid.ee/hambahaldjas (10.05.2013)
10. Kinoagent Alfred. www.dialoogid.ee/kinoagent (10.05.2013)
11. Konsultant Zelda. igemeravi.ee/zelda (10.05.2013)

OPTIMISATION OF DIALOGUE SYSTEMS

Bachelor's thesis

Mihkel Alavere

Abstract

The purpose of this Bachelor's thesis was to create a new dialogue system for the University of Tartu. A dialogue system is a computer program intended to receive user input as a sentence or a question and correspond to the user in the same language. As a result, a interactive communication is formed on a specific subject.

Some general information and history is included about dialogue systems. The work gives examples of dialogue systems that are in use in Estonia as well as around the world. A brief description is also provided about the working principle of such systems.

A new dialogue system was built upon an existing framework for dialogue systems, which was created by Margus Treumuth for his Doctoral thesis in 2011. The system focuses on people applying for studying in the University of Tartu and attempts to answer their questions. With this system, it is no longer necessary to search for information from various webpages on the university homepage, because all the required information is gathered into this dialogue system.

After the system was developed, it underwent a series of tests and the results were taken into consideration when developing the system further. The existing rules, according to which an answer is generated, were improved.

A short summary of its applications is also provided in the thesis.

A complete list of all patterns is included in the appendix of the work, as well as examples of conversations.

LISA 1. NÄIDISDIALOGID

Dialoog 1

Kasutaja: Tere

Paul: tere!

Kasutaja: Mis sa informaatika kohta oskad rääkida?

Paul: informaatik on arvutipõhise infotöötlemisega tegelev teaduse ja tehnika haru, mis tegeleb nii riistvara kui tarkvaraga

Kasutaja: Olen kuulnud ÕISist, misasi see on?

Paul: õppeinfosüsteem ehk ÕIS on õppekorralduse infovahetuskeskkond, seal saab õppeainetele registreeruda, kirju vaadata ning taotleda õppetoetust, mis on tudengi jaoks kõige olulisem :)

Kasutaja: Räägi bakalaureuse õppekavadest

Paul: matemaatika-informaatikateaduskonda sisseastujatel on valida kolme õppekava vahel: informaatika (155 kohta), matemaatika (35 kohta) ja matemaatiline statistika (25 kohta)

Kasutaja: Aga doktorandi õppekavad?

Paul: kui soovid doktoriõppesse astuda, on valikuid kolm: informaatika (8 kohta), matemaatika 2 (kohta) ning matemaatiline statistika (2 kohta)

Kasutaja: Mida bakalaureuse informaatikas tehakse?

Paul: informaatika bakalaureuseõpe annab tudengitele põhjalikud teadmised arvutiteaduse teoreetilistest ja praktilistest alustest

Kasutaja: Selge, aitäh!

Paul: võtke heaks!

Dialoog 2

Kasutaja: tere

Paul: tere!

Kasutaja: kuidas läheb?

Paul: mul läheb väga hästi :)

Kasutaja: kas loengutes peab käima?

Paul: pooleteisttunnistes loengutes käimine ei ole reeglina kohustuslik, küll aga saab sealt kasulikku infot

Kasutaja: kes on kõige lähendam õppejõud?

Paul: kõik Tartu Ülikooli õppejõud on äärmiselt toredad inimesed, eriti mat-inf teaduskonnas

Kasutaja: kas tudengiüritusi on lähedaid?

Paul: õppuritele mõeldud üritustest soovitan soojalt tudengite kevad- ja sügispäevi

Kasutaja: mitu arvutiklassi on?

Paul: mat-inf majas on neli arvutiklassi, kus on peal Windowsi ning Linuxi operatsioonisüsteemid

Kasutaja: kas wifi on majas tasuta?

Paul: Tartu Ülikool pakub tasuta traadita interneti teenust enamikes õppehoonetes

Kasutaja: mitu aastat tuleb ülikoolis käia, et saada baka?

Paul: bakalaureuseõpe on kõrghariduse esimene aste, mille nominaalkestus on kolm aastat

Kasutaja: kas IT-sektoris on tööd kõvasti?

Paul: Tartu ülikooli IT-sektoris õppides omandab tudeng IT-hariduse ning selles vallas on valitsenud pidev tööjõupuudus. sellise haridusega inimesele pole töö leidmine probleem. lisaks on Tartu Ülikool on ainsa Eesti ülikoolina 3% maailma edukamate ülikoolide seas

Kasutaja: arusaadav, nägemist!

Paul: peatse jälleenägemiseni!

Dialog 3

Kasutaja: tervist

Paul: tere!

Kasutaja: kas tööandjatele läheb diplom korda?

Paul: pärast ülikooli bakalaureuse-, magistri- või doktorikava lõpetamist antakse üliõpilasele ilus diplom, mis aitab tal tulevikus tööle saada

Kasutaja: kas raamatukogu on lähedal?

Paul: matemaatika-informaatikateaduskonna raamatukogu on avatud E-R 8.00-17.00

Kasutaja: kas ühiselamuid tartus on?

Paul: Tartus on kaheksa üliõpilaselamut, hinnad varieeruvad 50 - 90 euro vahel

Kasutaja: kuhu ma pargin?

Paul: tasuta parkimiskohti õppehoone kõrval on, kuid hetkel jääb neid vajaka

Kasutaja: kas jalgrattaruumi on?

Paul: jalgratturid saavad jätta rattad maja taha, parkimiskohti on piisavalt

Kasutaja: kas ülikoolis kohvik on või tuleb võileib hommikul kaasa võtta?

Paul: mat-inf teaduskonna majas on näljasele tudengile avatud hõrgutavade roogadega puhvet kella viieni, kuid hinnad on minu meelest üsna krõbedad

Kasutaja: kas jalgrattaruumi on?

Paul: jalgratturid saavad jätta rattad maja taha, parkimiskohti on piisavalt

LISA 2. TEADMUSBAASI REEGLID SEISUGA 13.05.2013

Muster	Lausung	Olek
kus(.*)? (õpe õppimine õppima)	õpingud toimuvad reeglina Tartus, ent näiteks magistri tarkvaratehnika kursusel on võimalus õppida Tallinnas	KUSÕPE
õppekava(.*)? magister	magistrisõppes on järgmised õppekavad ja nende vastuvõtuarv: finants- ja kindlustusmatemaatika (12), finantsmatemaatika (10) Informaatika (38), matemaatika (10). matemaatika-informaatikaõpetaja (12), matemaatiline statistika (10), tarkvaratehnika (57)	-
õppekava(.*)? (doktor doktoriõpe)	kui soovid doktoriõppesse astuda, on valikuid kolm: informaatika (8 kohta), matemaatika 2 (kohta) ning matemaatiline statistika (2 kohta)	-
SAIS	SAISi kaudu saab lihtsastl esitada avaldusi sisseastumiseks - andmed pärinevad lõpetatud kooli kohta riiklikest registritest ning sisseastumisdokumente ei ole vaja isiklikult ülikoolile esitada	-
arvutiklass	mat-inf majas on neli arvutiklassi, kus on peal Windowsi ning Linuxi operatsioonisüsteemid	-
raamatukogu	matemaatika-informaatikateaduskonna raamatukogu on avatud E-R 8.00-17.00	-
tarkvaratehnika(.*)? (magistriõpe magister)(.*)? (kandideerimine kandideerima sisse astumine tingimus)	tarkvaratehnikasse astumisel arvestatakse eelmise õppeastme keskmist hinnet (75%) ning movitatsioonikirja (25%)	-
õppejõud	kõik Tartu Ülikooli õppejõud on äärmiselt toredad inimesed, eriti mat-inf teaduskonnas	-
akadeemiline puhkus	akadeemilisel puhkusel üliõpilane vabastatakse õppe- ja teadustöö kohustustest ning puhkuse ajaks peatub üliõpilase nominaalne õppeaeg. Seda võimaldatakse omal soovil, tervislikel põhjustel, kaitseväeteenistusse asumisel või seoses lapse hooldamisega	-
loeng	pooleteisttunnistes loengutes käimine ei ole reeglina kohustuslik, küll aga saab	-

	sealt kasulikku infot	
tulekahju	tulekahju korral tuleb maja otsekohe evakueerida!	-
(ilus mitu) tüdruk	informaatikat õpivad noormeeste kõrval ka paljud kaunid neiud, neid on veidi alla poole kõigist õppuritest	-
dekanaat dekaan tõnu kollo	dekanaat korraldab matemaatika-informaatikateaduskonna igapäevast tööd, selle juhatajaks on dekaan Tõnu Kollo	-
kuhu(.*)? tööle	pärast informaatikaõppe lõpetamist saab asuda tööle näiteks andmebaaside haldurina, programmeerijana, võrguadministraatorina või tarkvara testijana	-
puhkama puhkeruum	õppehoone neljandal korrusel on puhkeruum, kus saab rahulikult oma õpingutega tegeleda	PUHKERUUM
(tööl käia töötada)(.*)? (ja kui ka samal ajal)(.*)? (õppida õping loengutes käima ülikoolis käima)	samal ajal töötada ja õppida ei ole mõtslik minu meelest, kuid kui rahaline olukord muudmoodi ei võimalda, siis tuleb töötada samal ajal	-
kalle kaarli matemaatika(.*)? juhataja	Kalle Kaarli on matemaatika-informaatikateaduskonna matemaatika instituudi juhataja	-
kalev pärna statistika(.*)? juhataja	Kalev Pärna on matemaatika-informaatikateaduskonna matemaatilise statistika instituudi juhataja	-
matemaatiline statistika(.*)? (magister magistriõpe)(.*)? (kandideerimine kandideerima sisse astumine tingimus)	matemaatilise statistika magistriõppesse kandideerides arvestatakse bakalaureuseõppe õppeastme keskmist hinnet, kusjuures minimaalne punktisumma on 60	-
avatud ülikool kaugõpe	avatud ülikoolis (kaugõppes) õppimine eeldab peamiselt iseseisvat tööd ja veebipõhistel kursustel osalemist. Õppesessioonid toimuvad kord kuus nädalavahetustel	ÕPPEVORM
(õppimatulek tulen õppima)(.*)? (kinnitamine kinnitama tähtaeg)	õppimatuleku pead kinnitama luuli teiseks pooleks, täpne kuupäev varieerub aastast-aastasse	-
(doktor doktoriõpe)(.*)? (kandideerimine kandideerima sisse astumine tingimus)	doktoriõppesse sisse pääsemiseks tuleb sooritada eksam, mis toimub juunis	-
(mis millised kus kust)(.*)? õppekava	jagan infot bakalaureuse-, magistri- ja doktoriõppe õppekavade kohta	-
(EAP õppemaht)(.*)? õppetoetus	põhi- ja täiendava toetuse taotlemiseks ja saamiseks ei pea iga semester just 30 EAP sooritama. Õppetoetuse määramisel vaadatakse õppemahu täitmise protsenti kumulatiivselt, st õpitud semestrite peale kokku	-

õppetoetus(.*)? (saaja nimi)	õppetoetuse saajate nimesid ei avalikustata. Neile, kellele on õppetoetus määratud, saavad õppetoetuse määramise kohta ülikooli e-postile teate ning näevad õppetoetuse määramise perioodi ja summat õppeinfosüsteemis	-
õppetoetus(:*)? (millal kuna)	õppetoetus makstake välja hiljemalt jooksva kuu 20. kuupäevaks	-
tarkvaratehnika	tarkvaratehnika tegeleb tarkvara projekteerimise ning tarkvaraprojektide teostuse, testimise ja dokumenteerimisega – eesmärgiga muuta tarkvara loomise protsess efektiivsemaks. Õppe jooksul on võimalik spetsialiseeruda kahele valdkonnale: ettevõtte tarkvaratehnika ja reaalaja sardsüsteemide arendus	-
õppetoetus stipendium stipp	alates 2013/2014. õppeaastast immatrikuleeritud üliõpilased saavad taotleda vajaduspõhist õppetoetust suurusega 75, 135 või 220 eurot kuus	-
õppelaen	õppelaenu on õigus saada üliõpilastel, kes õpivad täiskoormusega, selle maksimaalne suurus on 1917,35 eurot	-
välisüliõpilane välistudeng välismaalane õppima	Tartu Ülikooli üliõpilastel on võimalik Erasmuse õppida vahetusüliõpilasena 1-2 semestrit välismaal	-
ratas jalgrattaparkla jalgratas parkima jalgratas parkimine jalgratas	jalgratturid saavad jätta rattad maja taha, parkimiskohti on piisavalt	-
praktikum	praktikum on selline tund, kus juhendaja ei räägi klassi ees üksinda, vaid tudengid võtavad aktiivselt õppest osa ja koos lahendatakse ülesandeid, umbes nagu keskkooli matemaatikatunnid	-
prodekaan varmo vene	matemaatika-informaatikateaduskonna prodekaaniks on Varmo Vene, kes annab ka loenguid mõndades ainetes	-
merle marandi dekanaadi juhataja	matemaatika-informaatikateaduskonna dekanaadi juhataja on Merle Marandi	-
Liivi Luik õppekorralduse spetsialist	Liivi Luik mat-inf teaduskonna õppekorralduse spetsialist, temale saab kirja saata, kui on probleeme või muresid ÕISiga, näiteks õppainetele registreerimisega	-
niitsoo	Niitsoo on noorim teadusdoktor, kaitstes oma doktoritööd kõigest 24 aasta vanusena	-
matemaatika(.*)? (bakalaureus baka bakalaureuseõpe)	matemaatika erialal saab põhjaliku ettevalmistuse matemaatika teoreetilistes valdkondades, tutvutakse hulgateooria, algebra, geomeetria ja matemaatilise analüüsi põhitõdedega	-
(ülikool kool)(.*)? infot	jagan infot Tartu Ülikooli kohta	-

(andma jagama)		
corp selts korporatsioon	Tartus on kümneid tudengiorganisatsioone, -seltse ja -ühinguid, millega saab elu lõpuni ühineda, näiteks suurepärase Korp! Vironia	-
programm TULE	TULE raames saavad õpinguid jätkata need, kellel haridustee on erinevatel põhjustel omal ajal pooleli jäänud	-
toimetulekutoetus	toimetulekutoetust on võimalik taotleda isikutel (perekondadel), kelle kuusissetulekud pärast eluasemekulude tasumist normpinna ulatuses jäävad alla toimetulekupiiri	TOIMETULE KUTOETUS
(lõputöö bakatöö bakalaureusetöö magistritöö dokoritöö)(.*)? (kõitmine kõitma)	lõputöid saab printida, paljundada ja kõita üliõpilasesinduses ning Tartu Ülikooli raamatukogus	-
erasmus (stipendium toetus)	Erasmuse raames saab elamistoetust reeglina 340-550 eurot kuus, lisaks ühekordne reisitoetus vastavalt sihtriigile (70 kuni 600 eurot)	-
isep	ISEPi raames saab minna õppima nendesse riikidesse üle maailma, kuhu Erasmusega ei saa	-
facebook	minul Facebooki kontot ei ole, aga Tartu Ülikoolil on see täitsa olemas	-
e-õpe	e-õppe all mõeldakse õpetamist osaliselt või täielikult arvuti ja interneti abil	-
moodle	Tartu Ülikooli Moodle'i keskkond on mõeldud kasutamiseks Tartu Ülikooli üliõpilastele ning selle kaudu saab õppida ning vaadata õppeainete materjale	-
vilistlane	vilistlane on ülikooli lõpetanu	-
seljakott kohver sülearvutikohver	sülearvutit on kergem tassida seljakotis, aga minu meelest näeb kohvriga tudeng välja palju elegantsem	-
cum laude	cum laude antakse üliõpilasele, kes kaitses lõputöö hindele A või sooritas lõpueksami hindele A ning kelle keskmine hinne on 4,60 või kõrgem	-
(mis pikk kaua mitu aastat)(.*)? (bakalaureus baka bakalaureusekraad bakalaureuseõpe)	bakalaureuseõpe on kõrghariduse esimene aste, mille nominaalkestus on kolm aastat	-
(mis pikk kaua)(.*)? (magister magistrikraad magistriõpe)	magistriõpe on kõrghariduse teine tase, mis kestab kaks aastat	-
kõrgharidusreform	2012. aastal vastu võetud kõrgharidusreformiga seoses on kõrgharidus Eestis tasuta, kui õpitakse täiskoormusega eestikeelses õppekavas. Teisest semestrist alates sõltub hüvitamine täidetud õppekava mahust	-

(avaldus dokument)(.*)? (esitama esitamine)	avaldusi ülikooli sisseastumiseks saab esitada veebipõhiselt SAISis või dokumentide esitamisel isiklikult	-
sellest(.*)? piisab	ahah, selge	-
teeme(.*)? nii jääb(.*)? nii (tulen astun) (sisse kooli ülikooli õppima)	väga hea	-
mul(.*)? üks(.*)? küsimus	küsimuste tekkimisel saan aidata mina, aga abi saab ka dekanaadist	-
(aadress asukoht tänav asuma)(.*)? (mat-inf matemaatika- informaatika matemaatika- informaatika teaduskond teaduskond)	TÜ matemaatika-informaatikateaduskond asub peahoonest 400 m kaugusel, aadressiga J. Liivi 2, postiindeksiga 50409	-
ei meeldi mulle	miks ei meeldi? palun selgita	-
informaatika(.*)? (bakalaureus bakalaureuseõpe)(.*)? (kandideerimine kandideerima sisse astumine tingimus)	informaatikasse kandideerimiseks on vaja saada 60,0 punkti - arvesse võetakse emakeele (25%), võõrkeele (25%) ning matemaatika riigieksam (50%)	-
sisseastumiseksam sisseastumiskats e	sisseastumiseksamid toimuvad reeglina juulis, nende põhjak otsustatakse, kes saab ülikooli sisse ning kes jääb välja	-
(mis mida)(.*)? (tead oskad annad)	annan infot Tartu ülikooli matemaatika-informaatikateaduskonda sisseastujatele	-
(misasi mis)(.*)? informaatika	informaatika on arvutipõhise infotöötajusega tegelev teaduse ja tehnika haru, mis tegeleb nii riistvara kui tarkvaraga	-
moodul(.*)? õppeinfosüsteem	õppekavas tehtud moodulite valikud tuleb õppeinfosüsteemis ära märkida, moodulite registreerimise juhend asub ÕISis info ja abi keskkonnas	-
(valima saama)(.*)? kõrvaleriala	matemaatika-informaatikateaduskonnas õppides saab valida kõrvaleriala, kuid õppekavades, kus on suurendatud peaeriala mahtu, pole see võimalik	-
riidehoid rõivistu garderoob riidehoi uruum riie	informaatikamajas on valvega riidehoid, mis asub esimesel korrusel	-
ühikas ühiselamu üliõpilaselamu	Tartus on kaheksa üliõpilaselamut, hinnad varieeruvad 50 - 90 euro vahel	-
wifi nett internet	Tartu Ülikool pakub tasuta traadita interneti teenust enamikes õppehoonetes	-
loodusteaduskond	loodusteaduskond on mõeldud rohkem arvuti riistara huvilistele, informaatika tegeleb pigem tarkvara arendamisega	-
vabaaine vabaainemoodul vabaainet emoodul	kõik bakalaureuseõppekavad sisaldavad 12 EAP mahus vabaaineid, mille täitmiseks võib valida aineid kõigest teaduskondadest	-
marlon gerardo dumas menjivar hond duras	Marlon on hästi pika nimega, Hondurasest pärit õppejõud Tartu Ülikoolis	-
(matemaatika informaatika)(.*)?	mat-inf teaduskonnal on kaks infotelefoni:	-

(telefon telefoninumber number heli stama)	737 5860 ning 737 5862	
parkla parkima parkimine auto sõiduk	tasuta parkimiskohti õppehoone kõrval on, kuid hetkel jääb neid vajaka	-
margus treumuth arhitekt	Margus Treumuth on keeleressursside analüütik ning ühtlasi selle dialoogsüsteemi arhitekt	-
nõme tobe igav	kõik ei peagi meeldima	-
(pane(.*)?(end ennast)(.*)? põlema) (sae(.*)? pekki) (tõmba(.*)? oksa)	see väljend on vana ja kulunud	-
majanduskriis majanduslangus	igale langusele järgneb tõus	-
mare koit tehisintellekt	kui soovid rohkem teada saada tehisintellekti kohta, siis soovitan küsida Mare Koidu käest	-
(sa sina ta tema)(.*)? (ilus kaunis kole inetu)	ilu on vaataja silmades	-
matemaatiline(.*)? statistika(.*)? (bakalaureus bakalaureuseõpe baka)	matemaatilise statistika õppekaval õpitakse lisaks süvendatud matemaatikale erinevate andmete kogumist, nende süviti analüüsimist ja järelduste tegemist	-
sport sportima sportimine	Tartu Ülikooli akadeemiline spordiklubi pakub suurepäraseid võimalusi sportimiseks, soovitan soojalt!	-
hello how are you	hello! This system is designed for people speaking estonian	-
rääkima(.*)? veel(.*)? (anekdoot nali)	rohkem nalju mul hetkel ei meenu	-
(mis kes)(.*)? olema(.*)? (parim parem halvem halvim)	mm ... aga mida sina arvad?	-
(mis mida)(.*)? (sa sina)(.*)? arvama	loeb see, mida sina arvad	ARVAMUS
(mis kuidas)(.*)? ilm	ilma kohta mul teave puudub	ILM
(sa sina)(.*)? (kena lahe armas ilus kaunis vahvalt ore nunnu intelligentne)	tänan	KOMPLIMENT
laula	eelistan kuulata teiste laulmist	LAULA
(rääkima tegema tee)(.*)? (anekdoot nali naljake)	räägin ühe anekdoodi	NALI_1
naljakas lõbus naljamees joke	nalja peab saama	NALI_2
turing.*	teen iga päev kümneid Turingi teste	TURING
mõte mõtlemine mõtlemine aju	mõtlen ainult kuidas andmebaasist uus vastus leida	-
tunne tundma	tundeid mul ei ole, ainult terve hunnik lauseid	-
eksam	ülikoolis tuleb ainete läbimiseks üldiselt läbida kursuse lõpus eksam	-
solvuma solvumine solvang kuri ülbitsema	anna andeks, kui sind solvasin	-
muusika laulma laulja staar	mulle meeldib Black Rockiti pala	-

	"itimees"	
(sa sina)(.*)? magama	arvutid ei maga kunagi	-
viin õlu konjak	alkohol aeglustab mind	-
lumi külmpakane talv suvi	mulle krõbedad talved sobivad, kuna arvuti ei tunne sooja ega külma	-
autor	minu autoriks on üks Tartu ülikooli bakalaureusetudeng	-
ladina	per aspera ad astra	-
inglise	how are you?	-
(facebook.* orkut.* rate.*)	see on mingi tutvumisportaal?	-
nõme hull	olen igavuse etalon	-
muu soovitama	soovitaks siis teha kaks kätekõverdust	-
sööma söök söökla puhvet kohvik	mat-inf teaduskonna majas on näljasele tudengile avatud hõrgutavade roogadega puhvet kella viieni, kuid hinnad on minu meelest üsna krõbedad	-
teater kino	kunagi olin teatrifriik, siis filmifriik, nüüd olen ülikooli sisseastuja nõustaja	-
juhe	juhe jooksis kokku ... see ei tohiks olla eriti üllatav, või mis? ...	-
raamat kirjandus kunst lugema	mulle meeldib vabal ajal lugeda tudengite lõputöid	-
reaalne	sina oled reaalne, mina olen illusioon	-
elu(.*)? mõte	elu mõte on kuulata tähti ja rääkida oma lugu ... vaikselt	-
hea(.*)? (aastavahetus uus(.*)? aasta vana(.*)? aasta)	sullegi head ja parimat jne	-
vaikus	vaikus on digitaalselt võimalik, digitaalne nullpunkt	-
(su sinu)(.*)? perekonna.*nimi	võid kutsuda mind eesnime pidi	-
(sa sina)(.*)? (robot masin arvuti)	igatahes inimene ma ei ole	ROBOT
õppekava(.*)? bakalaureus	matemaatika-informaatikateaduskonda sisseastujatel on valida kolme õppekava vahel: informaatika (155 kohta), matemaatika (35 kohta) ja matemaatiline statistika (25 kohta)	-
(armastad meeldib)(.*)? töö	jah, väga	TÖÖ
miks(.*)? (vaikima vait)	vaikin, sest ei oska midagi sobivat öelda	VAIKIMINE
ei usu valetama vale valelik	tahtlikult ma ei valeta	VALE
kui vana(.*)? (sa sina oled)	piisavalt vana	VANUS
(mis pikk kaua)(.*)? (doktor doktorikraad doktoriõpe)	doktoriõpe on kõrghariduse tipp, milles pärast nelja aastat õppimist saadakse doktoriks	DOKTORITÕÖ
ükskõik suva kama kaks savi	ahah	ÜKSKÕIK
kus(.*)? elama	mina elan internetis	-
palun aidake	püüan aidata	-
saa(.*)? aru	püüan saada :)	-
ei huvitama	sorry, ma mõtlesin, et huvitab ...	-

statsionaar statsionaarne päevane õpe	päevane õpe (mõnes kõrgkoolis statsionaarne õpe) eeldab igapäevast osalemist õppetööl	ÕPPEVORM
suits suitsetama suitsetamine tubaka s sigar sigarett	suitsetamine on väga halb harjumus, mina jätsin maha	-
super	super	-
mis(.*)?(su sinu)(.*)? nimi	minu nimi on Paul	-
kes sa oled	tehisintellekt olen, noh, arvuti põhimõtteliselt	-
tervist tere	tere!	-
rõõmus	tore	-
tsau	tsau	-
vait	vait jään siis, kui saab öeldud, mis öelda on :)	-
veider	veider jah :)	-
aitäh	võtke heaks!	-
väga hea	väga hea jah	-
narva pärna viljandi	Tartu Ülikoolil on kolledžid Narvas, Pärnus ning Viljandis, kuid seal informaatikat õppekava puudub	-
ülbe ülba	ülbusega ei ole siin midagi pistmist	-
pole paha	üsna hea jah	-
kuidas(.*)? läheb	mul läheb väga hästi :)	-
sina(.*)? inimene	ei, ma ei ole inimene, olen tehisintellekt	-
aeglane	ei saa kiiremini	-
miks(.*)? (tartu ülikool tartu ülikoolis) (itsektor it-sektor)	Tartu ülikooli IT-sektoris õppides omandab tudeng IT-hariduse ning selles vallas on valitsenud pidev tööjõupuudus	MIKSUT
imelik	imelik jah :)	-
pole viga	jah, pole viga	-
ei (saama saanud)(.*)? aru	jah, raskusi on :)	-
jama	jama? nojah, teie otsus	-
jääma(.*)? vait	jään vait siis, kui ise tahan	-
halvasti	kahju küll	-
kurb	kurb lugu jah	-
küsima	küsida võib kõike, iseasi kas oskan vastata	-
lahe	lahe jah :)	-
loll	pigem lollikindel	-
minu nimi	meeldiv tutvuda, mina olen Paul	-
normaalne	normaalne jah	-
hüvasti head aega nägemist nägemiseni	peatse jälleenägemiseni!	-
ei(.*)? (tahtma taha)	ok	-
ok okei	ok	-
oskama	oskan paljugi, kuigi üsna pea peaksid avastama, et olen samas üpris juhm	-

tänan	palun	-
sobib	tore, et sobib, kas saan veel kuidagi aidata?	-
ekstern	eksternina õppimine tähendab õppimist üliõpilasstaatust omamata	ÕPPEVORM
(peahoone ülikool)(.*)? (aadress asukoht tänav asuma)	Tartu ülikooli aadress on Ülikooli 18, postiindeksiga 50090	-
külastisüliõpilane	Tartu Ülikooli üliõpilased saavad tasuta kuulata loenguid ja osaleda seminarides EELK Usuteaduse Instituudis, Eesti Maaülikoolis, Eesti Muusika- ja Teatriakadeemias, Eesti Kunstiakadeemias, Eesti Lennuakadeemias, Kõrgemas Usuteaduslikus Seminaris, Tallinna Tehnikaülikoolis, Tallinna Ülikoolis, Tartu Kõrgemas Kunstikoolis ja Tartu Teoloogia Akadeemias	-
informaatika(.*)? (magister magistriõpe õppekava)	informaatika magistriõppes jätkatakse bakalaureuseõppes omandatud baasoskuste süvendamist, käsitletakse masinõpet, hajussüsteeme, formaalseid keeli ja krüptograafiat	-
sõjavägi sõjaväeteenistus armee kaitsevägi kaitseväeteenistus	sõjaväeteenistuse ajaks saab võtta ühe akadeemilise puhkuse lisaks ning üliõpilane vabastatakse õppetööst	-
vaba aeg vaba aja huvitav üritus kevadpäev sügispäev tudengipäev tudengiüritus	õppuritele mõeldud üritustest soovitan soojalt tudengite kevad- ja sügispäevi	-
diplom	pärast ülikooli bakalaureuse-, magistri- või doktorikava lõpetamist antakse üliõpilasele ilus diplom, mis aitab tal tulevikus tööle saada	-
instituut	matemaatika-informaatikateaduskonda kuulub kolm instituuti: arvutiteaduse instituut, matemaatilise statistika instituut ja matemaatika instituut	-
üliõpilane tudeng arv koguarv kui palju	üliõpilaste koguarv meie teaduskonnas on hetkel 1030	-
üliõpilaspilet ISIC	on olemas kaks erinevat üliõpilaspiletit: tasuta Eesti üliõpilaspilet üheks aastaks ja tasuline rahvusvahelise (ISIC) üliõpilaspilet, millega saab soodustusi toodetele ja teenustele nii Eestis kui välismaal	-
täiskoormus	täiskoormusega õppides täidab üliõpilane iga õppeaasta lõpuks vähemalt 75% (45 EAP)	-
osakoormus	osakoormusega õppides jääb tudengi õppemaht 50% ja 75% vahele (30 - 44 EAP)	-
eksmatt eksmatrikuleerimine eksma	eksmatrikuleerimine ehk eksmatt on	-

trikuleerima	üliõpilaste nimekirjast väljaarvamine	
immatrikuleerimine immatrikuleerima	immatrikuleerimine tähendab isiku taasarvamist üliõpilaste nimekirja	-
keskmine hinne	keskmise hinde arvutamisel arvestatakse tähed numbriteks järgmiselt: A = 5, B = 4, C = 3, D = 2, E = 1, F = 0	-
õppekava (vahetus vahetama)	õppekava vahetamiseks on kaks võimalust: kandideerida õpingute jooksul soovitud eriala vabale õppekohale või astuda uuesti ülikooli sisse	-
VÕTA	VÕTA raames saab varasemate õpingute ning töökogemuste eest kanda ülikooli üle ainepunkte	-
vaidlustamine vaidlustama	kui üliõpilane soovib vaidlustada õppekorraldusega seonduvat otsust, tuleb esmalt pöörduda otsuse teinud isiku poole ning vaidlustada otsus kas kirjalikult või suuliselt seitsme päeva jooksul otsuse teatavaks tegemisest alates	-
ravikindlustus	üliõpilased on ravikindlustusega kindlustatud, neil on õigus saada arstiabi kõikides haigekassaga lepingu sõlminud raviasutustes, kusjuures raviteenuse eest tasub haigekassa	-
tuutor	tuutor ehk üliõpilasest nõustaja on spetsiaalse koolituse läbinud üliõpilasest vabatahtlik, kes abistab oma eriala esmakursuslasi ja välisüliõpilasi õpingute alustamisel Tartu Ülikoolis	-
tasuta (operatsioonisüsteem windows) dreamspark msdn msdnaa	DreamSpark on Microsofti programm, millelga saab kasutada õppetöös ja mittetulunduslikus teadustöös suurt hulka Microsofti tarkvara, näiteks Windowsi operatsioonisüsteeme	-
(lõputöö dokoritöö bakatöö bakalauareusetöö magistritöö)(.*)? (vaatamine vaatama nägemine)	varasemate tudengite lõputöid saab vaadata lõputööde repositooriumis DSpace	-
telefon telefoninumber number helistama	Tartu ülikooli infotelefon on 737 5100. Lisaks on infotelefon ka matemaatika-informaatikateaduskonnal	-
ahsoo (ah soo) (või nii)	just	-
dokument esitamine esitama isiklikult	dokumentide isiklikuks esitamiseks on vaja esitada avaldus, diplom ja akadeemiline õiend ning ID-kaardi või passi isikukoodiga koopia Tartu Ülikooli peahoonesse	-
informaatika(.*)? (magister magistriõpe)(.*)? (kandideerimine kandideerima sisseastumine tingimus)	informaatika magistriastmesse kandideerimiseks vajalik minimaalne punktisumma on 60,0, arvesse võetakse eelmise õppeastme keskmine hinne	-
ÕIS	õppeinfosüsteem ehk ÕIS on	-

	õppekorralduse infovahetuskeskkond, seal saab õppeainetele registreeruda, kirju vaadata ning taotleda õppetoetust, mis on tudengi jaoks kõige olulisem :)	
informaatika(.*)? (bakalaureus bakalaureuseõpe baka õppekava)	informaatika bakalaureuseõpe annab tudengitele põhjalikud teadmised arvutiteaduse teoreetilistest ja praktilistest alustest. Saadud oskused on heaks eelduseks nii edasistele õpingutele kui ka edukale karjäärile näiteks tarkvaratestija, programmeerija, süstesteemianalüütikuna või ka paljudel teistel elualadel	-
peaeriala	enamik bakalaureuseõppe õppekavadest sisaldab ainult üht peaeriala ning on ära toodud õppekavas	-
õppetoetus(.*)? mitu	põhi- ja täiendava toetuse saajate arv arvutatakse haridus- ja teadusministri poolt kinnitatud koefitsiendi alusel	-
jaak vilo arvutiteaduse instituudi juhataja	doktor Jaak Vilo on TÜ Arvutiteaduse instituudi juhataja, ta on tore inimene	-
eduroam	eduroam võimaldab akadeemiliste võrkude kasutajatele vaba juurdepääsu internetile kõigis projektiga ühinenud asutustes, näiteks saab Amsterdams ligipääsu internetti oma tavalise kasutajanimega	-
valikaine valikainemoodul	valikainete moodul on 6 EAP mahus ainete komplekt, mida õppija võib valida kõigist õppekavadest	-
ajalugu	matemaatika-informaatikateaduskonna ajaloo tean vaid niipalju, et sellele pandi alus 1967. aastal	-
eap ap	õpingute mahtu mõõdetakse ainepunktides, ühele ainepunktile vastab 26 tundi üliõpilase tööd	-
ap	AP on vana õppemahtu hindav karakteristik, 1 AP võrdub 1,5 EAP.	-
õppematerjal	õppematerjalid on saadaval ÖISI kaudu ning raamatuid saab raamatukogust	-
tasub(.*)? magister magistriõpe	magistriõppesse tasub ikka minna, saab kaks aastat tudengimõnusi nautida	-
saama tuutor	tuutoriks on oodatud kõik üliõpilased, kes on abivalmid, ettevõtlikud ning suhtlemisaldid	TUUTOR
põhitoetus	põhitoetuse suuruseks on 55,93 eurot kuus, kuid seda pärast 2013. aasta sisseastunutele enam ei jagata	-
üliõpilasesindus TÜÜE	TÜ Üliõpilasesindus (TÜÜE) tegutseb Tartu Ülikooli enam kui 17 000 tudengi vaadete ja huvide esindamiseks nii ülikoolisiselt kui ka laiemalt ja sellesse saad kandideerida ka sina!	-
erasmus	Erasmuse raames saab minna välismaale	ERASMUS

	õppima, selle kestvus on reeglina 1 semester	
täiendusõpe	täiendusõppena pakub ülikool erineva pikkusega kursusi erialase või ametialase pädevuse tõstmiseks ning huvialakoolitusi, koolitusseminare, õppepäevi, arenguprogramme, e-õpet, suveülikooli ja talveülikooli kursusi	-
ülikool rektor	Tartu Ülikooli rektoriks on praegu Volli Kalm	-
muuseum ajalugu	esimesel korrusel on tilluke ülevaade matemaatika-informaatikateaduskonna ajaloost ning keldrikorrusel on spetsiaalne ruum, kus hoitakse vanu arvutivideoid, aga sinna minekuks peab luba küsima	-
sülearvuti	sülearvutit õppimiseks osta ei ole tingimata tarvis	SÜLEARVUT I
matemaatiline statistika	Tartu Ülikoolis on matstati õppekava nii baka-, magistri- kui ka doktoriõppes	-
matemaatiline (.*)? statistika (.*)? (magister magistriõpe)	matemaatiline statistika magistris tähendab tööd informatsiooniga: selle kogumist, töötlemist, analüüsimist ja lahtimõtestamist	-
matemaatika (.*)? (bakalaureus bakalaureuseõpe) (.*)? (kandideerimine kandideerima sisse astumine tingimus)	matemaatika bakalaureuseõppesse kandideerides arvestatakse emakeele riigieksamit (25%), võõrkeele riigieksamit (25%) ja matemaatika riigieksamit (50%)	-
matemaatiline statistika (.*)? (bakalaureus bakalaureuseõpe) (.*)? (kandideerimine kandideerima sisse astumine tingimus)	matemaatilise statistika bakalaureuseõppesse kandideerides arvestatakse emakeele riigieksamit (25%), võõrkeele riigieksamit (25%) ja matemaatika riigieksamit (50%)	-
matemaatika (.*)? (magister magistriõpe) (.*)? (kandideerimine kandideerima sisse astumine tingimus)	matemaatilika magistriõppesse kandideerides arvestatakse bakalaureuseõppe õppeastme keskmist hinnet, kusjuures minimaalne punktisumma on 60	-

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina, Mihkel Alavere

(sünnikuupäev: 04.08.1990)

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) enda loodud teose
Uue dialoogsüsteemi loomine veebipõhiste dialoogsüsteemide raamistiku abil,
mille juhendaja on Margus Treumuth,

- 1.1.reprodutseerimiseks säilitamise ja üldsusele kättesaadavaks tegemise eesmärgil, sealhulgas digitaalarhiivi DSpace-is lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;
 - 1.2.üldsusele kättesaadavaks tegemiseks Tartu Ülikooli veebikeskkonna kaudu, sealhulgas digitaalarhiivi DSpace'i kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.
2. olen teadlik, et punktis 1 nimetatud õigused jäävad alles ka autorile.
 3. kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest tulenevaid õigusi.

Tartus, **08.05.2013**