



Kauno technologijos universitetas

Informatikos fakultetas

Individų atpažinimo tikslumo parametru papildyta ontologija ir jos tyrimas

Baigiamasis magistro krypties studijų projektas

Dovilė Žygė

Projekto autorė

Doc. Dr. Rita Butkienė

Vadovė

Kaunas, 2019



Kauno technologijos universitetas

Informatikos fakultetas

Individų atpažinimo tikslumo parametru papildyta ontologija ir jos tyrimas

Baigiamasis magistro krypties studijų projektas

Informacinių sistemų inžinerijos studijų programa (kodas 621E15001)

Dovilė Žygė

Projekto autorė

doc. dr. Rita Butkienė

Vadovė

lekt. dr. Algirdas Šukys

Recenzentas

Kaunas, 2019



Kauno technologijos universitetas

Informatikos fakultetas

Dovilė Žygė

Individų atpažinimo tikslumo parametru papildyta ontologija ir jos tyrimas

Akademinio sąžiningumo deklaracija

Patvirtinu, kad mano, Dovilės Žygės, baigiamasis projektas tema „Individų atpažinimo tikslumo parametru papildyta ontologija ir jos tyrimas“ yra parašytas visiškai savarankiškai ir visi pateikti duomenys ar tyrimų rezultatai yra teisingi ir gauti sąžiningai. Šiame darbe nei viena dalis nėra plagijuota nuo jokių spausdintinių ar internetinių šaltinių, visos kitų šaltinių tiesioginės ir netiesioginės citatos nurodytos literatūros nuorodose. Įstatymų nenumatytų piniginių sumų už šį darbą niekam nesu mokėjęs.

Aš suprantu, kad išaiškėjus nesąžiningumo faktui, man bus taikomos nuobaudos, remiantis Kauno technologijos universitete galiojančia tvarka.

(vardą ir pavardę įrašyti ranka)

(parašas)

Žygė, Dovilė. Individų atpažinimo tikslumo parametru papildyta ontologija ir jos tyrimas. Magistro baigiamasis projektas / vadovas doc. dr. Rita Butkienė; Kauno technologijos universitetas, Informatikos fakultetas.

Mokslo kryptis ir sritis: Informatikos inžinerija, technologijos mokslai

Reikšminiai žodžiai: *semantika, ontologija, tikslumo parametras, individų atpažinimas.*

Kaunas, 2019. 123 p.

SANTRAUKA

Individų atpažinimo tikslumo parametru papildytos ontologijos tyrimo tikslas yra sudaryti galimybę semantinės paieškos vartotojams matyti atpažintų individų tikslumo įverčius ir jais remiantis atrinkti, rikiuoti užklausos rezultatus. Tyrimo metu yra vertinama kokią įtaką ontologijos papildymas tikslumo parametru turės paieškos greitaveikai.

Šiam tikslui sukurta eksperimentinė sistema, paruošti keli ontologijos variantai, apimantys individus su tikslumo įverčiu ir be jo, sudarytos įvairaus sudėtingumo SPRQL užklausos. Ontologijoms saugoti naudojama semantinių grafų duomenų bazė *GraphDB*. Eksperimentinė sistema leidžia kurti įvairaus sudėtingumo užklausas, vykdyti semantinę paiešką, peržiūrėti paieškos rezultatus ir užklausos vykdymo greitaveikos duomenis.

Eksperimento metu nustatyta, kad sprendimas pateikti paieškos vartotojams užklausos išvestų individų atpažinimo tikslumo įverčius turi akivaizdžią įtaką užklausos vykdymo greitaveikai. Įtaka greitaveikai priklauso nuo užklausos sudėtingumo, rezultato struktūros dydžio, išvedamų rezultatų kiekio ir pačios ontologijos dydžio.

Žygė, Dovilė. *Ontology supplemented with the individuals' recognition accuracy parameter and its investigation*. Master's thesis in Information Systems Engineering / supervisor assoc. doc. dr. Rita Butkienė. The Faculty of Informatics, Kaunas University of Technology.

Research area and field: Informatics Engineering, Technology Science

Key words: *semantic, ontology, precision measure, name entity recognize*

Kaunas, 2019. 123 p.

SUMMARY

The purpose of the ontology study, supplemented with an individual recognition accuracy parameter, is to enable the users of semantic search to see the accuracy estimates of identified individuals and to select, sort and query the results based on them. The impact of the ontology supplementation on the accuracy parameter will be assessed by the search speeds.

For this purpose, an experimental system has been developed, several ontology variants have been prepared that include individuals with and without accuracy estimates, and SPRQL requests of varying complexity. The semantic graph database, GraphDB, is used to store ontologies. The experimental system allows you to create queries of varying complexity, perform semantic search, view search results and query execution speed data.

The experiment found that the decision to provide search engine users with accuracy estimates of the queries generated by the query has a clear impact on query execution speeds. The impact on speeds depends on the complexity of the query, the size of the structure of the result, the amount of output and the size of the ontology itself.

TURINYS

Lentelių sąrašas	8
Paveikslų sąrašas	10
Terminų ir santrumpų žodynas	12
Įvadas.....	13
1. Ontologijos praturtintos individų atpažinimo tikslumo įverčiais modelio probleminės srities analizė.....	15
1.1. Analizės tikslas	15
1.2. Tyrimo objektas, sritis ir problema	15
1.3. Darbo tikslas, uždaviniai, planas ir siejami privalumai.....	16
1.4. Ontologijos praturtintos individų atpažinimo tikslumo įverčiais tyrimo objektų analizė ...	16
1.4.1. Duomenų ištraukimas interneto puslapiuose	16
1.4.2. Semantinės paieškos procesas.....	17
1.4.3. Teksto paruošimas anotavimui.....	17
1.4.4. Semantinės paieškos tikslumo vertinimo parametrai (matai)	27
1.5. Tyrimo objekto naudotojų analizė	28
1.6. Esamų problemos sprendimo metodų analizė (Lietuvos ir tarptautiniu mastu)	28
1.7. Siekiamo sprendimo apibrėžimas	32
1.8. Analizės išvados.....	32
2. Ontologijos praturtintos individų atpažinimo tikslumo įverčiais modelio sprendimo reikalavimų specifikacija ir projektas, formalus aprašas	33
2.1. Funkcinių reikalavimų specifikacija	33
2.2. Ontologijos praturtintos individų atpažinimo tikslumo įverčiais sprendimo panaudojimo atvejų modelis.....	33
2.2.1. Nefunkciniai reikalavimai.....	39
2.3. Dalykinės srities modelis.....	41
2.4. Naudotojų sąsajos modelis	44
2.5. Formalus teksto anotavimo tikslumo įverčio skaičiavimo aprašymas	44
2.6. Reikalavimų apibendrinimas	51
3. Ontologijos praturtintos individų atpažinimo tikslumo įverčiais modelio realizacijos projektas	53
3.1. Sistemos loginė architektūra.....	53
3.2. Sistemos elgsenos modelis	54
3.3. Duomenų bazės modelis.....	58
3.4. Detalus projektas.....	59

3.5. Realizacijos projekto apibendrinimas	60
4. Ontologijos praturtintos individų atpažinimo tikslumo įverčiais eksperimentinės sistemos realizacija ir testavimas	61
4.1. Sistemos realizacijos modelis ir veikimo aprašymas	61
4.2. Diegimo modelis	62
4.3. Naudojamos technologijos	63
4.4. Testavimo planas.....	63
4.4.1. Kontrolinis pavyzdys.....	64
4.5. Realizacijos apibendrinimas	66
5. Individų atpažinimo tikslumo įverčiais praturtintos ontologijos eksperimentinis tyrimas	67
5.1. Eksperimento planas	67
5.2. Eksperimento rezultatai.....	78
5.3. Sprendimo veikimo ir savybių analizė, kokybės kriterijų įvertinimas	86
5.4. Sprendimo taikymo rekomendacijos.....	87
6. Rezultatų apibendrinimas ir išvados	88
7. Turinys.....	89
8. Priedai.....	91
8.1. priedas. Lietuvių kalbos morfologiniame anotatoriuje naudojamos žymos	91
8.2. priedas. Teksto anotavimo pavyzdys	92
8.3. priedas. Ontologijos dokumentų įverčių failas -„Evaluation_egz.“	95
8.4. priedas. Užklausos.....	103
8.5. priedas. Fizinė duomenų bazė.....	120

LENTELIŲ SĄRAŠAS

1.1 lentelė. Kalbos dalių žymų anglų kalbai paaiškinimas. [10]	21
2.1 lentelė. Panaudojimo atvejis: „Įkelti dokumentą“	34
2.2 lentelė. Panaudojimo atvejis: „Valdyti užklausas“	35
2.3 lentelė. Panaudojimo atvejis: „Peržiūrėti užklausų ir statistikos rezultatus“	36
2.6 lentelė. Nefunkcinio reikalavimo NF-1 aprašas	39
2.7 lentelė. Nefunkcinio reikalavimo NF-2 aprašas	39
2.8 lentelė. Nefunkcinio reikalavimo NF-3 aprašas	40
2.9 lentelė. Nefunkcinio reikalavimo NF-4 aprašas	40
2.10 lentelė. Nefunkcinio reikalavimo NF-5 aprašas	40
2.11 lentelė. Nefunkcinio reikalavimo NF-6 aprašas	40
2.12 lentelė. Nefunkcinio reikalavimo NF-7 aprašas	40
2.13 lentelė. Nefunkcinio reikalavimo NF-8 aprašas	40
2.14 lentelė. Nefunkcinio reikalavimo NF-19 aprašas	41
2.15 lentelė. Dokumento tipo aprašas	43
2.16 lentelė. Įvykių tipų aprašas	43
2.17 lentelė. Objekto tipų aprašas	44
2.18 lentelė. Skaičiavimo rezultatai T1 tekstynui	46
2.19 lentelė. Skaičiavimo rezultatai T2 tekstynui	48
2.20 lentelė. Skaičiavimo rezultatai T3 tekstynui	50
4.1 lentelė. Valdyti paieškos sritis testavimo atvejis	63
4.2 lentelė. Valdyti užklausas testavimo atvejis	64
4.3 lentelė. Valdyti užklausų rezultatus testavimo atvejis	64
4.4 lentelė. Vykdyti semantinę paiešką testavimo atvejis	64
5.1 lentelė. Ontologijų schemų sąrašas	67
5.2 lentelė. Vieno objekto sudėtingumo užklausos pavyzdys	72
5.3 lentelė. Dviejų objektų sudėtingumo užklausos pavyzdys	73
5.4 lentelė. Du objektai+ Vokietija sudėtingumo užklausos pavyzdys	74
5.5 lentelė. Užklausos „Ką kalbėjo asmenys minėdami 'Vokietija?'“ vykdymo rezultatai	75
5.7 lentelė. Vienodo sunkumo užklausų rezultatai	78
5.8 lentelė. Vienodo sunkumo užklausų rezultatai naudojant paieškos limitą	80
5.9 lentelė. Skirtingo sudėtingumo užklausų greitaveikos rezultatai	82
5.10 lentelė. Skirtingo sudėtingumo užklausų greitaveikos rezultatai naudojant limitą	84
5.11 lentelė. Skirtingo išvedamų objektų kiekio greitaveikos rezultatai	85
8.1 lentelė. Morfologijos paslauga	92
8.2 lentelė. Morfologinio nagrinėjimo reikšmės	92
8.3 lentelė. Segmentavimo paslauga	93
8.4 lentelė. Įvardintų esybių paslauga	94
8.5 lentelė. Anotuoto teksto paslauga	95
8.6 lentelė. Užklausos „Visi asmenys?“ duomenys	103
8.7 lentelė. Užklausos „Visi asmenys?“ rezultatai	103
8.8 lentelė. Užklausos „Visos vietovės?“ duomenys	104
8.9 lentelė. Užklausos „Visos vietovės?“ rezultatai	104
8.10 lentelė. Užklausos „Visos organizacijos?“ duomenys	104
8.11 lentelė. Užklausos „Visos organizacijos?“ rezultatai	104

8.12 lentelė. Užklauso „Kokie asmenys yra minimi?“ duomenys	105
8.13 lentelė. Užklauso „Kokie asmenys yra minimi?“ rezultatai	105
8.14 lentelė. Užklauso „Kokios vietovės yra minimos?“ duomenys	105
8.15 lentelė. Užklauso „Kokios vietovės yra minimos?“ rezultatai	106
8.16 lentelė. Užklauso „Kokios organizacijos yra minimos?“ duomenys	106
8.17 lentelė. Užklauso „Kokios organizacijos yra minimos?“ rezultatai	106
8.18 lentelė. Užklauso „Ar yra koks nors kainų pokytis?“ duomenys	107
8.19 lentelė. Užklauso „Ar yra koks nors kainų pokytis?“ rezultatai	107
8.20 lentelė. Užklauso „Ar padidėjo kainos?“ duomenys	108
8.21 lentelė. Užklauso „Ar padidėjo kainos?“ rezultatai	108
8.22 lentelė. Užklauso „Ar sumažėjo kainos?“ duomenys	109
8.23 lentelė. Užklauso „Ar sumažėjo kainos?“ rezultatai	109
8.24 lentelė. Užklauso „Ką kalbėjo asmenys?“ duomenys	110
8.25 lentelė. Užklauso „Ką kalbėjo asmenys?“ rezultatai	110
8.26 lentelė. Užklauso „Kokios pareigos užimamos organizacijose?“ duomenys	111
8.27 lentelė. Užklauso „Kokios pareigos užimamos organizacijose?“ rezultatai	111
8.28 lentelė. Užklauso „Kokias pareigas užima asmenys?“ duomenys	112
8.29 lentelė. Užklauso „Kokias pareigas užima asmenys?“ rezultatai	112
8.30 lentelė. Užklauso „Ką kalbėjo asmenys minėdami „Vokietija“?“ duomenys	113
8.31 lentelė. Užklauso „Ką kalbėjo asmenys minėdami „Vokietija“?“ rezultatai	114
8.32 lentelė. Užklauso „Ką kalbėjo asmenys minėdami "Europa"“?“ duomenys	114
8.33 lentelė. Užklauso „Ką kalbėjo asmenys minėdami "Europa"“?“ rezultatai	115
8.34 lentelė. Užklauso „Ką kalbėjo asmenys minėdami "Lietuva"“?“ duomenys	115
8.35 lentelė. Užklauso „Ką kalbėjo asmenys minėdami "Lietuva"“?“ rezultatai	116
8.36 lentelė. Užklauso „Ką pareiškė asmenys?“ duomenys	116
8.37 lentelė. Užklauso „Ką pareiškė asmenys?“ rezultatai	117
8.38 lentelė. Užklauso „Ką samprotavo asmenys?“ duomenys	118
8.39 lentelė. Užklauso „Ką samprotavo asmenys?“ rezultatai	119
8.40 lentelė. Užklauso „Ką akcentavo asmenys?“ duomenys	119
8.41 lentelė. Užklauso „Ką akcentavo asmenys?“ rezultatai	120
8.42 lentelė. Duomenų bazės lentelių sąrašas	121
8.43 lentelė. Duomenų bazės lentelės „aspnetuserclaims“ aprašas	121
8.44 lentelė. Duomenų bazės lentelės „aspnetusers“ aprašas	121
8.45 lentelė. Duomenų bazės lentelės „queries“ aprašas	122
8.46 lentelė. Duomenų bazės lentelės „aspnetuserlogins“ aprašas	122
8.47 lentelė. Duomenų bazės lentelės „aspnetroles“ aprašas	122
8.48 lentelė. Duomenų bazės lentelės „aspnetuserroles“ aprašas	122
8.49 lentelė. Duomenų bazės lentelės „categories“ aprašas	122
8.50 lentelė. Duomenų bazės lentelės „requets“ aprašas	122
8.51 lentelė. Duomenų bazės lentelės „__migrationhistory“ aprašas	123
8.52 lentelė. Duomenų bazės lentelės „responses“ aprašas	123

PAVEIKSLŲ SĄRAŠAS

1 pav. Segmentinio anotavimo rezultato struktūros pavyzdys	20
2 pav. Žymų radimas pirmam žodžiui [9]	20
3 pav. Žymų išvardinimas antram žodžiui [9]	21
4 pav. Žymėjimo galutinis įvertinimas [9]	21
5 pav. Morfologinio anotavimo rezultato struktūros pavyzdys pirmas.....	23
6 pav. Morfologinio anotavimo rezultato struktūros pavyzdys antras	23
7 pav. Sintaksinio anotavimo rezultato pirmas pavyzdys	24
8 pav. Sintaksinio anotavimo rezultato antras pavyzdys	24
9 pav. Sintaksinio anotavimo rezultato trečias pavyzdys	24
10 pav. Įvardintų esybių anotavimo rezultato palyginamasis pavyzdys	25
11 pav. Google paieškos rezultatas	26
12 pav. „Ontotex „ semantinės paieškos rezultatas	27
13 pav. Semantinės paieškos LKSSAIS karkasas.....	29
14 pav. Semantinės paieškos modelis [20]	30
15 pav. „Ontotex“ kompanijos saityno anotuoto teksto ištrauka	30
16 pav. Anotuoto teksto rezultatas	31
17 pav. Anotuoto teksto rezultatas	31
18 pav. Semantinės paieškos eksperimentinės sistemos reikalavimai	33
19 pav. Semantinės paieškos eksperimentinės sistemos PA modelis	34
20 pav. PA „Įkelti dokumentą“ scenarijus.....	35
21 pav. PA „Valdyti užklausas“ scenarijus	36
22 pav. PA „Peržiūrėti užklausų ir statistikos rezultatus“ scenarijus.....	37
23 pav. PA „Pakartoti semantines paieškas“ scenarijus.....	38
24 pav. PA „Vykdyti semantinę paiešką“ scenarijus	39
25 pav. Dalykinės srities esybių ryšių modelis	41
26 pav. Dokumento esybių ryšių modelis	42
27 pav. Įvykio esybės kategorijos	43
28 pav. Objekto esybės kategorijos.....	43
29 pav. Semantinės paieškos eksperimentinės sistemos navigacijos modelis.....	44
30 pav. Konceptinis modelis	45
31 pav. Tikslumo įverčio skaičiavimo modelis T1 tekstynui	46
32 pav. Tikslumo įverčio skaičiavimo modelis T2 tekstynui	48
33 pav. Tikslumo įverčio skaičiavimo modelis T3 tekstynui	50
34 pav. Sistemos loginė architektūra.....	53
35 pav. Sistemos loginės architektūros detalizavimas	53
36 pav. PA1 „Įkelti dokumentą“ sistemos modelio sekų diagrama	54
37 pav. PA2 „Valdyti užklausas“ sistemos modelio sekų diagrama.....	54
38 pav. PA2 detalizuota alternatyva „Valdyti paieškos sritis“	55
39 pav. PA2 detalizuota alternatyva „Valdyti SPARQL užklausas“	56
40 pav. PA3 „Peržiūrėti užklausos ir statistikos rezultatus“ sistemos modelio sekų diagrama.....	56
41 pav. PA3 detalizuota alternatyva „Peržiūrėti rezultatus“	57
42 pav. PA3 detalizuota alternatyva „Valdyti statistikas“.....	57
43 pav. PA4 „Pakartoti semantines paieškas“ sistemos modelio sekų diagrama	58
44 pav. PA5 „Inicijuoti semantinę paiešką“ sistemos modelio sekų diagrama	58

45 pav. Loginė duomenų bazės schema	59
46 pav. Detalizuota eksperimentinės sistemos loginė architektūra.....	60
47 pav. Ontologijos praturtintos individų atpažinimo tikslumo įverčiais sistemos komponentų modelis.....	61
48 pav. Grafinės sąsajos komponento modelis	61
49 pav. Logikos valdymo komponento modelis	62
50 pav. Duomenų paslaugų komponento modelis	62
51 pav. Semantinės paieškos eksperimentinės sistemos diegimo modelis.....	63
52 pav. Eksperimentinės sistemos paieškos langas.....	65
53 pav. Eksperimentinės sistemos užpildytas paieškos langas.....	66
54 pav. Užklauso „Ką kalbėjo asmenys?“ paieškos rezultato fragmentas.....	66
55 pav. LKSSAIS ontologijų schemų ir „ <i>Evaluation_schema</i> “ importavimo struktūra.....	68
56 pav. Ontologijos schema „ <i>Evaluation_schema</i> “	69
57 pav. Papildyta ontologijos schema duomenų savybe F-matas.....	69
58 pav. Ontologijos dokumentų įverčių failo -„ <i>Evaluation_egz.</i> “ fragmentas.....	70
59 pav. Individas papildyta duomenų savybe F-matas.....	70
60 pav. Užklauso „Ką kalbėjo asmenys minėdami 'Vokietija?'“ nuo 1-20 testo rezultatų diagrama	76
61 pav. Užklauso „Ką kalbėjo asmenys minėdami 'Vokietija?'“ nuo 20-40 testo rezultatų diagrama	76
62 pav. Užklauso „Visos vietovės“ nuo 1-20 testo rezultatų diagrama.....	77
63 pav. Užklauso „Visos vietovės“ nuo 21-40 testo rezultatų diagrama.....	78
64 pav. Vienodo sudėtingumo užklausų greitaveikos diagrama.....	79
65 pav. Vienodo sudėtingumo užklausų greitaveikos diagrama naudojant limitą.....	81
66 pav. Užklauso "VISI ASMENYS" greitaveikos be limitu ir su limitu grafikas.....	81
67 pav. Skirtingo sudėtingumo užklausų greitaveikos diagrama su tikslumu dokumento atžvilgiu	83
68 pav. Skirtingo sudėtingumo užklausų greitaveikos diagrama su tikslumu individo atžvilgiu..	83
69 pav. Skirtingo sudėtingumo užklausų greitaveikos diagrama su tikslumu dokumento atžvilgiu naudojant limitą.....	84
70 pav. Skirtingo sudėtingumo užklausų greitaveikos diagrama su tikslumu individo atžvilgiu naudojant limitą.....	85
71 pav. Skirtingo išvedamų objektų kiekio greitaveikos diagrama	86
72 pav. Ontologijos praturtintos individų atpažinimo tikslumo įverčiais sistemos fizinės duomenų bazės schema.....	121

TERMINŲ IR SANTRUMPŲ ŽODYNAS

Ontologija (angl. *Ontology*). Tam tikros srities sąvokų visumos specifikacija; [1]

Protégé (angl. *Protégé*). Nemokamas, atviro kodo ontologijos kūrimo įrankis; [2]

OWL (angl. *Web Ontology Language*). Tai speciali pritaikyta ontologijų kalba skirta įvairioms informacinėms sistemoms, gebančiomis dirbti su internetu ir siekiančios informacijos jame. Ši kalba skirta visam ontologijų atpažinimo procesui, kartu naudojant RDF bei teikia papildomas klasių objektų savybių žinias, bei susieja objektų sąryšius; [3]

Auksinis standartas (angl. *gold standard*) – tai lingvistinis išteklius (pvz., tekstynas, lingvistinė duomenų bazė), kurį peržiūrėjo ir išanalizavo ekspertas (-ai) ir kuris naudojamas kompiuterinėms programoms mokyti bei vertinti.

LKSSAIS Lietuvių kalbos sintaksinės ir semantinės analizės informacinė sistema.

ĮVADAS

Darbas parengtas pagal Informacinių sistemų inžinerijos magistratūros studijų programos reikalavimus.

Darbo problematika ir aktualumas

Pasaulinis saitynas (angl. World Wide Web) – populiarus, viešai prieinamos informacijos paieškos šaltinis. Tačiau dėl informacijos gausos ir jos pateikimo formato paieška jame nėra tokia efektyvi kokia galėtų būti. Didžioji dali informacijos pasauliniame tinkle yra tekstas, kurį reikia struktūruoti, jei norima jame atlikti paiešką. Teksto segmentavimas į žodžius sudaro galimybę įgyvendinti raktažodinė paiešką. Tačiau raktažodinė paieška apriboja vartotoją pateikti abstraktesnius klausimus, nevertina užklausos prasmės. Dėl to vartotojui gali tekti dar ilgai ieškoti tinkamo atsakymo tarp pateiktų paieškos rezultatų gausos.

Semantinės technologijos suteikia galimybę įgyvendinti prasmingesnę ir efektyvesnę paiešką. Pagrindinis semantinės paieškos privalumas – gauti klausimo prasmę atitinkančius atsakymus ir tokiu būdu sutaupyti laiko analizuojant ir filtruojant pateiktus užklausos rezultatus. Tačiau, jei pagrindinis informacijos šaltinis yra nestruktūruotas tekstas, semantinės paieškos rezultato kokybė (tikslumas) labai priklauso nuo kalbinių technologijų, kurios buvo panaudotos analizuojant tekstą, veikimo tikslumo.

Kalbinių technologijos sparčiai tobulėja, tačiau jos dar neišsprendžia visų su rašytinės kalbos analize susijusių problemų. Pagrindinės iš jų būtų: žodžių daugiaprasmiškumas, teksto struktūros sudėtingumas ir paties teksto korektiškumas. Teksto analizę atliekantys programiniai komponentai ne visada sugeba tiksliai nustatyti žodžių ar jų junginių prasmę, t.y., tiksliai atlikti informacijos ištraukimą (angl. *Information Extraction*). O šios analizės pagrindu tam tikrų tipų individais ir ryšiais tarp jų užpildytos ontologijos, kuriose ir vykdoma semantinė paieška, negali vartotojui su vienodu ir pakankamu patikimumu pateikti atsakymų į klausimus.

Kadangi kalbinių technologijos iš principo negali išspręsti visų iš paties teksto kylančių sudėtingų analizės atvejų. Todėl, realizuojant semantinės paieškos sprendimus, svarbu į tai atsižvelgti. Vieni iš svarbesnių uždavinių būtų: tinkamų teksto analizės tikslumo vertinimo metodų ir priemonių sukūrimas, ir tikslumo įverčių panaudojimas semantinėje paieškoje

Darbo tikslas ir uždaviniai

Šio magistrinio darbo tikslas yra sudaryti ir ištirti galimybę semantinės paieškos vartotojams matyti atpažintų individų tikslumo įverčius ir jais remiantis atrinkti, rikiuoti užklausos rezultatus.

Tyrimo uždaviniai:

1. išanalizuoti tekstų semantinio anotavimo procesą;
2. išanalizuoti ontologijos kūrimo principus;
3. išanalizuoti ontologijų atvaizdavimo kalbą (OWL 2);
4. išanalizuoti veiksnus, kurie turi įtakos atpažintų esybių ir ryšių tarp jų tikslumui;
5. išanalizuoti individų atpažinimo tikslumo parametru papildytų ontologijų vertinimą;
6. sudaryti modelį ontologijoms praturtintoms individų atpažinimo tikslumo įverčiais tikslumui fiksuoti;
7. eksperimentiškai ištirti pasiūlytą metodiką ir modelį.

Darbo struktūra

Baigiamąjį darbą sudaro tokios pagrindinės dalys:

- įvadas,
- pirmame skyriuje pateikiama tikslumo įverčio panaudojimo analizė, ieškoma panašių, esamų sprendimų,
- antrame skyriuje modelio tyrimas, kur aprašyta: reikalavimų specifikacija, dalykinės srities modelis, naudotojų sąsajos modelis ir formalus sprendimo aprašas,
- trečiame skyriuje pateikiama modelio realizacijos projekto aprašymas,
- ketvirtame skyriuje pateikiama eksperimentinės sistemos realizacijos ir testavimo aprašymas,
- penktame skyriuje pateikiama sistemos eksperimentinis tyrimas, kurio metu bus vertinama sukurta sistema.

1. ONTOLOGIJOS PRATURTINTOS INDIVIDŲ ATPAŽIMIMO TIKSLUMO ĮVĖRČIAIS MODELIO PROBLEMINĖS SRITIES ANALIZĖ

Šiame skyriuje pateikiama su darbo problematika susijusios informacijos analizė, analizuojamas tyrimo objektas, esami problemos sprendimo būdai (technologijos, metodai), kitų autorių literatūra.

1.1. Analizės tikslas

Analizės tikslas – geriau susipažinti, kaip naudotojui pateikti kuo tikslesnę informaciją semantinės paieškos metu, išsiaiškinti ar jau yra panašių sprendimų, egzistuojančius sprendimus išnagrinėti ir įvertinti jų privalumus ir trūkumus.

Analizei atlikti naudojami šiais metodais:

- studijuojami literatūros šaltiniai;
- panašių sprendimų radimas bei jų analizė.

1.2. Tyrimo objektas, sritis ir problema

Magistrinio darbo tyrimo objektas – ontologijos, kurios užpildytos individualiais, sukurtais išanalizavus tekstą bei paieška tokiose ontologijose. Objektą taip pat apima ir tekstą analizuojantys programiniai komponentai, jų veikimas ir tarpusavio priklausomybės pateikiant ieškomą informaciją kaip tikslą, apdorojamų tekstų šaltiniai.

Šio magistrinio darbo tyrimo sritis – teksto analizės tikslumo vertinimo metodai, modeliai, tikslumo įvėrčio panaudojimo paieškoje sprendimai.

Naudojimasis pasauliniu saitynu darosi vis populiariesnis. Sudėtingėja informacijos pateikimas vartotojui. Vartotojui tampa sunku pasiekti ieškomą informaciją. Taip yra todėl, kad pats vartotojas nežino, kaip tai jam padaryti.

Paieškos sistema atitinkanti vartotojo poreikius yra vienas iš svarbiausių elementų, kuris gali padėti gauti tiksliai ieškomą informaciją. Nuolatos yra kuriamos paieškų sistemos pagal vartotojo poreikius, nes dabartinės paieškų sistemos duoda daug netikslių rezultatų. Aktualu tampa mokėjimas tinkamai ieškoti reikalingos informacijos.

Sparčiai tobulėjant informacinėms technologijoms vis daugiau technologiniu procesų yra siejama su kalbos apdorojimu. Vienas iš būdų paieškas padaryti efektyvesnes yra – semantinė paieška. Tačiau kiekvienas vartotojas nori kuo tikslesnės ir naujesnės ieškomos informacijos. Kai ko nors ieškome, į paieškos sistemų laukelį įvedame žodį ar frazę ir tikimės, kad paieškos sistema teisingai traktuoja mūsų informacijos poreikius ir formuojamą užklausą pagal parinktus raktinius žodžius. Deja, kartais gauname rezultatus, kuriuos turime gebėti įvertinti, atsirinkti iš daugiau nei kelių ar keliasdešimt tūkstančių galimų.

Atsiranda problema, kad analizuojami tekstai ne visada yra tikslūs, korektiškai parašyti, ar išsaugoti iš patikimų šaltinių, todėl yra sunku nustatyti automatiškai atpažintų esybių tikslumą. Tai atsiliepia tekstų vertinimui. Norint, spręsti šią problemą reikia ieškoti galimai tinkamų tikslumo vertinimo metodų ar vertinimo modelių

Numatomos problemos sprendimas – norit išspręsti šią problemą reikia sudaryti tikslumo vertinimo metodą, modelį, kurį naudojant vartotojas galėtų atskirti patikimą šaltinį, nuo mažiau patikimų, juos reitinguojant.

1.3. Darbo tikslas, uždaviniai, planas ir siekiami privalumai

Šis poskyris apibrėžia darbo tikslą ir numatomus darbo uždavinius, kurie bus sprendžiami, siekiant išspręsti kilusią problemą.

Darbo tikslas – sudaryti ir ištirti galimybę semantinės paieškos vartotojams atsirinkti tiksliausius užklauso rezultatus.

Magistro baigiamajame darbe suformuluoti uždaviniai:

1. išanalizuoti tekstų semantinio anotavimo procesą;
2. išanalizuoti darbo tyrimo objektus;
3. išanalizuoti esybių ir ryšių tarp jų automatinio atpažintų tikslumo vertinimą;
4. sudaryti modelį ontologijos praturtintos individų atpažinimo tikslumo įverčiais;
5. realizuoti ontologijos praturtintos individų atpažinimo tikslumo įverčiais eksperimentinę sistemą;
6. eksperimentiškai ištirti ir įvertinti pasiūlytą modelį.

1.4. Ontologijos praturtintos individų atpažinimo tikslumo įverčiais tyrimo objektų analizė

1.4.1. Duomenų ištraukimas interneto puslapiuose

Ištraukiant duomenis (angl. *Information extraction (IE)*) internetiniuose šaltiniuose pagrindinis uždavinys yra automatiškai išgauti struktūrizuotą informaciją iš nestruktūrizuotos ar pusiau struktūrizuotos [4]. Daugeliu atveju tai yra susiję su duomenų apdorojimu naudojant natūralios kalbos apdorojimo (angl. *NLP*) būdą. Struktūrizuoti duomenys yra tiksliai semantiškai apibrėžti iš pasirinkto teksto. Tokie duomenys gali būti tiesiogiai pateikiami galutiniam vartotojui arba gali naudoti kitos kompiuterinės sistemos, pvz., paieškos ir duomenų bazių valdymo sistemos, kad vartotojams būtų teikiamos geresnės paslaugos [5].

Dabartinė svarba

Informacijos ištraukimas labai glaudžiai yra susijęs su didėjančiu informacijos kiekiu, kuris yra gaunamas nestruktūrizuotoje formoje. Pasaulinio tinklo išradėjas *Tim Berners-Lee* priskiria pasaulinį tinklą dokumentų tinklui ir pasisako už tai, kad daugiau informacijos būtų galima pateikti kaip duomenų tinklą [4]. Kol tai nėra įgyvendinta, žiniatinklis daugiausia susideda iš nestruktūrizuotų dokumentų, kuriuose trūksta semantinių metaduomenų. Informacija esanti dokumentuose gali būti lengviau prieinama apdorojimui transformuojant į reliacinę formą arba pažymint XML žymas. Stebint naujienų informacijų srautus duomenų pateikimas reikalauja, kad duomenų ištraukimas pertvarkytų nestruktūrizuotus duomenis į tai kas būtų logiškai suprantama. Esami sprendimai ištraukiant duomenis yra skirti nuskaityti duomenų rinkinius, kurie yra parašyti natūralia kalba, ir jais užpildyti duomenų bazę.

Konkrečios informacijos, kurią reikia ištraukti, tipas ir struktūra priklauso nuo konkretaus pritaikymo poreikio. Toliau pateikiama keletas tokių taikymo pavyzdžių:

- Biomedicinos moksle dažnai reikia išanalizuoti mokslininkų publikuojamus straipsnius, kad atrinkti konkrečias biomedicinos esybes. Todėl paprasta paieška pagal raktinių žodžių atrinkimą gali būti netinkama, nes biomedicinos įstaigos dažnai turi sinonimus ir dviprasmiškus pavadinimus, todėl sunku tiksliai surasti tinkamas esybes.
- Finansų specialistams dažnai reikia ieškoti tam tikros informacijos iš straipsnių, kad padėtų jų kasdieniam sprendimų priėmimui. Pavyzdžiui, finansų įmonei gali tekti sužinoti visus bendrovės perėmimus, įvykusius per tam tikrą laikotarpį, ir informaciją apie kiekvieną naują įsigijimą. Norint automatiškai rasti tokią informaciją iš teksto, reikalingos standartinės informacijos ištraukimo technologijos, tokios kaip įvardintų esybių atpažinimas [4].

1.4.2. Semantinės paieškos procesas

Trumpiausias žodžių junginys, kuriuo būtų galima apibūdinti kas yra semantika būtų „Prasmės tyrimas“.

Lyginant semantinę paiešką ir „Google“ paiešką jos tarsi ir niekuom nesiskiria, tačiau semantinė paieška yra sudėtingas procesas, kuris savyje turi nematomas vartotojui technologijas. Vartotojas ieškodamas informacijos paieškos lauke įveda žodį ar frazę ir tikisi, kad paieškos sistema teisingai supranta jo poreikius. Dažnu atveju gražinta informacija būna gana plati ir tenka pačiam vartotojui ją atsirinkti iš kelių ar keliasdešimt tūkstančių galimybių.

Semantinė paieška siekia pagerinti ieškomos informacijos tikslumą, atsižvelgiant į ieškotojo ketinimus ir konteksto terminų reikšmę, neatsižvelgiant į atskirų žodžių prasmę paieškos užklausoje.

Semantinė paieška remiasi įvairiais aspektais, įskaitant paieškos kontekstą, lokaciją, tikslą, žodžių variacijas, sinonimus, apibendrintas ir specializuotas užklausas, atitinkančias sąvokas ir natūralios kalbos užklausas, pateikiant atitinkamus paieškos rezultatus.

Poreikis analizuoti gaunamą informaciją yra svarbus naudotojui, kuris nori gauti kuo tiksliau suprastos jo užklaustos atsakymą. Tikslumas nereiškia vien tik suradimas tokio pačio žodžio ar frazės. Tikslumas turi būti ir šaltinyje, kuris tą informaciją pateikia ir ar juo mes galime pasitikėti. Kuo daugiau nusistatysime tikslumo kriterijų komponentas, kuriuos naudojame anotavime, tuo rezultatas bus tikslesnis ir išvesdami atsakymą naudotojui galėsime patį tiksliausią atsakymą išvesti pirmą.

1.4.3. Teksto paruošimas anotavimui

Anotacija (žymėjimas) yra atributų, nuorodų ir t.t. pridėjimas prie dokumento arba prie pasirinktos teksto vietos. Anotuota informacija suteikia papildomą informaciją (metaduomenis) apie dokumento gabalus. Tipinės teksto anotacijos yra: segmentavimo, morfologinė, sintaksinė, įvardintų esybių atpažinimo ir semantinė.

Technologijos žengia greitai žingsniu į priekį, bet yra vietų kur žmogaus rankinis darbas dar yra neapkeičiamai svarbus. Tai būtų ir tekstų atrinkimas ir jų paruošimas prieš semantinę anotavimą. Automatizuotas darbas kompiuteriu ne visada supras taip gerai kaip žmogus, kuris tekstas yra rišlus ir turintis prasmę, ar tokių savybių neturintis ir t.t. Semantinio anotavimo procesas rankiniu būdu reikalauja turėti ir specifinių žinių ir didelių laiko sąnaudų. Norint šių problemų išvengti semantinio anotavimo procesas yra automatizuojamas.

Kompiuterio atliktu darbu semantiniame anotavime sunku pasitikėti, tad turime paruošti ir sukurti tekstynus, kurie bus rankiniu būdu ekspertų sužymėti. Tokius tekstynus galima vadinti auksiniu

standartu (angl. *gold standard*) arba etalonu. Rankiniu būdu semantiškai anotuotas tekstynas bus etalonu, jei rankinis anotavimas bus atliktas 100% tikslumu.

Naudojant automatizuotus semantinius anotatorius yra svarbu jų tikslumas, nes nuo to ir priklausys išvedamos užklauskos atsakymo tikslumas. Turėdami tokius tekstynus, jais galime patikrinti ir įvertinti atskirų semantinio anotavimo komponentų tikslumo įverčius, kas bus svarbu ir galutiniame rezultate.

Šiuo metu yra vertinamas galutinis tikslumas, kurį gauname po semantinio anotavimo, bet labai svarbu būtų vertinti kiekvieno naudojamo komponento tikslumą, taip tiksliau įvertinant kur yra didžiausi netikslumai ir tas spragas užpildant. Nes vieno komponento netinkamas paruošimas anotavimui gali lemti ir galutinį rezultatą.

Žinoma yra labai svarbu pastebėti, kad turint tekstyno standartą vienos srities, o paduodant tekstą kuriam mes visai neturime tekstyno standarto, tikėtis gero galutinio rezultato neverta.

Semantinio anotavimo paaiškinimas

Semantinis anotavimas - teksto žymėjimas remiantis pasirinkta metodika, kurios pagalba galima išspręsti natūralioje kalboje pasitaikančias dviprasmybes. Semantinio anotavimo tikslas tekste panaikinti dviprasmybes ir svarbiausia paskirtis yra pažymėti, kokią prasmę išreiškia teksto fragmentas. Kasdieninėje kalboje dviprasmybės, tai kai žodis tuo pačiu metu turi keletą skirtingų reikšmių, pvz., Rasa gali būti ir vardas ir gamtos reiškiny.

Semantinis anotavimo procesas susieja anotavimui skirtą teksto žodžius, frazes su ontologijoje esančios apibrėžtos srities informacija. Susieję žodžius ir frazes su ontologijos informacija, semantiškai anotuotas tekstas yra išsaugojamas faile kuris yra paruoštas naudojimui.

Semantinis tekstų anotavimas priskiriamas natūralios kalbos apdorojimo (angl. *natural language processing*) uždavinių grupei. Šio metodo esmė – gauti semantiškai struktūruotą tekstą iš nestruktūruoto teksto. Jeigu pažiūrėtume į šio uždavinio sprendimą kaip į juodą dėžę, įeigoje matytume nestruktūruotą (arba struktūruotą HTML žymų lygmenyje) tinklo dokumento tekstą, o išėjime – dalykinės srities ontologiją, praturtintą naujais, iš anotuoto teksto gautais klasių egzemplioriais.

Semantinis tekstų anotavimas yra sudėtingas uždavinys, reikalaujantis pilnos kalbinių komponentų grandinės įvykdymo, kurią dažniausiai sudaro šie lingvistiniai komponentai:

1. Leksinė analizė – teksto (sakinių) skaidymas į smulkesnius vienetus (angl. *tokens*), t.y. į žodžius, simbolius, skyrybos ženklus ir pan.
2. Morfologinė analizė – kalbos dalių priskyrimas leksinės analizės metu išskaidytiems žodžiams.
3. Žodžių lemavimas – pagrindinės žodžio formos nustatymas.
4. Kolokacijų atpažinimas – dažnai vartojamų žodžių sekų (junginių) nustatymas statistiniais metodais.
5. Įvardytų esybių atpažinimas – žodžių, junginių klasifikavimas į iš anksto apibrėžtas kategorijas, tokias kaip vardai, organizacijų pavadinimai, vietovardžiai ir pan.
6. Sintaksinė analizė – sakinio fragmentų skaidymas į sintaksės vienetus – kalbos dalis, frazes, nusakant jų tarpusavio sąryšius. Skiriami du pagrindiniai sąryšių tipai: hierarchinis (angl. *constituency-based*) ir priklausomybės (angl. *dependency-based*). Sintaksinės analizės rezultatas dažniausiai pateikiamas kaip sakinio sintaksinių vienetų sąryšių medis (angl. *parse tree*).

7. Semantinių vaidmenų žymėjimas – dar kitaip vadinamas paviršutine semantine analize. Šis metodas skirtas iš anksto veiksmožodžių leksikone apibrėžtų vaidmenų priskyrimui sintaksiniam sakinio vienetams. Automatinio semantinių vaidmenų žymėjimo metodai grindžiami mašininio mokymo taikymu naujiems tekstams anotuoti jau suanotuotų tekstynų (angl. *gold-standart*) atžvilgiu. Labiausiai paplitę tekstynai ir leksikonai semantinių vaidmenų žymėjimui yra *PropBank* ir *FrameNet*.
8. Anaforų, kataforų nustatymas – semantiškai lygiaverčių, bet skirtingomis formomis sakinyje išreikštų esybių nustatymas (angl. *coreference*).
9. Semantinio tekstų anotavimo rezultatas plačiai taikomas semantinėje paieškoje, kuomet standartiniai, raktažodžių pasikartojimu grindžiami paieškos metodai papildomi arba pakeičiami semantine paieška ontologijose ir tekstų anotacijose.

Ontologijos paaiškinimas

Ontologijos apibrėžimas būtų toks: "ontologija - tam tikros srities sąvokų visumos specifikuojimas išreikštu pavidalu" (*merriam-webster* žodynas) [6].

Ontologijos panaudojimas semantiniame anotavime

Ontologijos panaudojimas semantiniame teksto anotavime yra atpažinti kokios ontologijos klasę ar jos egzempliorių išraiška nagrinėjamas teksto žodis, su kokiais žodžiais jis gali būti susietas ir su kokiais negali.

1.4.3.1. Segmentavimo analizė

Segmentavimas yra gana sudėtinga sritis ir kelianti daugiausiai sunkumų. Nuo šios srities ir pačio pirmo anotavimo įrankio priklauso tolimesni rezultatai. Nuo jo tikslumo priklauso ar teisingai buvo atskirti simboliai, kurie perduoti tolimesniam apdorojimui. Dėl šios priežasties būtent segmentavimas kelia didelį susidomėjimą mokslininkams.

Atliekant teksto segmentavimą iškyla nemažai problemų priklausančių nuo teksto stiliaus. Jas galima padalinti į tokias dalis [7]:

- teksto eilutės išskyrimas iš teksto,
- žodžio atskyrimas iš eilutės,
- didžiųjų ir mažųjų raidžių skirtumo,
- sakinio pradžios ir pabaigos aptikimas.

Segmentiniame (seg) anotavimo lygmenyje saugomos žymos, nurodančios žodžio vietą tekste. Visos šio lygio žymos yra patalpintos tarp žymų. Į segmentinį failą, pagalba yra įterpiama antraštė. Toliau seka žyma, kurioje nurodomas indeksas "lex": { "seg" :}. Po žymos seka visos segmentiniam lygiui būdingos žymos.

Segmentinis failas aprašo suskirstymą į paragrafus ir į sakinius. Pasinaudojus saitynu Semantika.lt ir jame esančiu įrankiu, segmentavimą galima atvaizduoti pavyzdžiu (1 pav.). Plačiau žiūrėti prieduose 8.2 poskyryje.

```

Jonaitis atėjo į Kauno Technologijos Universitetą pirmą kartą 1985m rugsėjo 15d .
{"seg": [[0,1], [1,1], [3,8], [12,5], [18,1], [20,5], [26,13], [40,12], [53,5], [59,5], [65,5], [71,7], [79,3], [82,1]], "s": [[0,83]], "p": [[0,83]]}

Jonaitis atėjo į Kauno Technologijos Universitetą pirmą kartą 1985m rugsėjo 15d .
{"seg": [[0,1], [1,1], [2,1], [4,5], [10,1], [12,5], [18,13], [32,12], [45,5], [51,5], [57,5], [63,7], [71,3], [74,1]], "s": [[0,75]], "p": [[0,75]]}

```

1 pav. Segmentinio anotavimo rezultato struktūros pavyzdys

Leksinė analizė – teksto (sakinių) skaidymas į smulkesnius vienetus (angl. *tokens*), t.y. į žodžius, simbolius, skyrybos ženklus ir pan.

Leksinė analizė apima tik skaidymą į leksemas, o segmentavimas apima skaidymą į pastraipas, sakinius, leksemas.

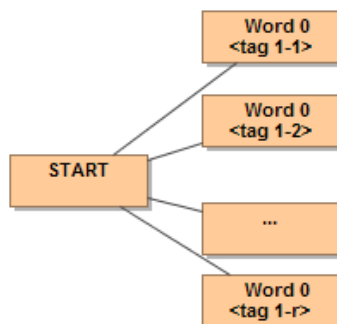
1.4.3.2. Morfologinė analizė

Morfologinė analizė – kalbos dalių priskyrimas leksinės analizės metu išskaidytiems žodžiams Šio anotatoriaus pagalba yra pažymimos įvairios sakinio dalys pagal tam tikras požymių grupes: kalbos dalis, linksniuotes, asmenuotes, gimines ir pan.

Pasak VDU Kompiuterinės lingvistikos centro anotatoriaus autorių [8]: „Automatiškai morfologiškai anotuojant tekstus susiduriama su dideliu morfologiniu daugiareikšmiškumu – reiškiniu, kai tam pačiam žodžiui ar žodžio formai pateikiamos kelios lemos (antraštiniai žodžiai) arba kelios gramatinės pažymos, pvz.: *galvos* – daiktavardžio *galva* vns. kilmininkas arba dgs. vardininkas, taip pat veiksmazodžio *galvoti* būsimosios laiko trečias asmuo; *pasakyti* – bendratis arba neveikiamosios rūšies dalyvio vyr. gim. dgs. vardininkas; *ir* – jungtukas, dalelytė,rieveiksmis. Nustatyta, kad beveik pusė lietuvių kalbos žodžių ir žodžių formų yra morfologiškai daugiareikšmiai“. Lietuvių kalbos morfologiniame anotatoriuje naudojamų pažymų lentelė yra priede (žr.8.1 priedas.).

Vienas iš labiausiai paplitusių statistinių metodų yra paslėptas *Markovo* modelis [9], kuris bando išmokti kalbos dalių žymėjimus. Paslėptas *Markovo* modelis kalbos dalių žymas traktuoja kaip nematytus paslėptus narius, o sakinių žodžius kaip stebėjimus. Modelis daro prielaidą, kad žymos formuoja *Markovo* seką tokią, kad kiekvienos žymos perėjimas prie kitos sakinio žymos su tikimybe priklausančia tik nuo prieš tai ėjusios žymos. Su paslėpto *Markovo* modelio kalbos dalių žymėjimu yra naudojamas *Viterbi* algoritmas, kad būtų sugeneruota optimali žymų seka duotam sakiniui. Paslėpto *Markovo* modelio paieška:

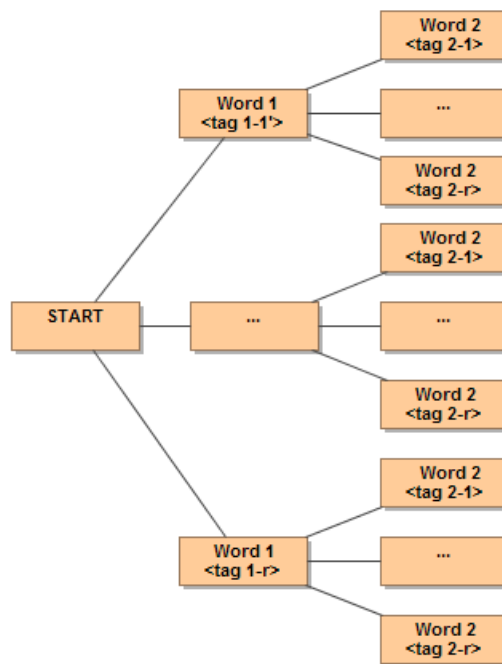
1 žingsnis. Išvardyti visas galimas žymas pirmam žodžiui.



2 pav. Žymų radimas pirmam žodžiui [9]

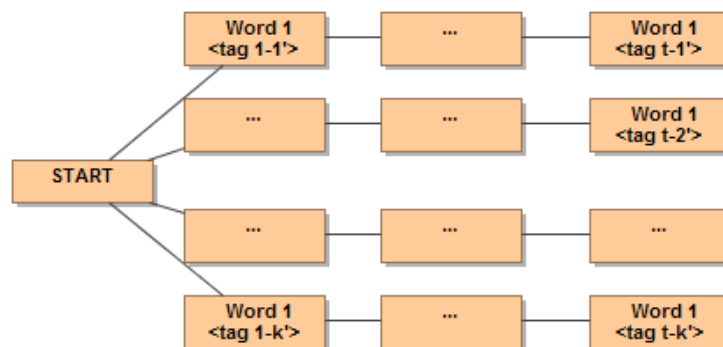
2 žingsnis. Įvertinti kiekvieną sužymėjimą naudojant apmokytą paslėptą *Markovo* modelį.

3 žingsnis. Kiekvienam k žymėjimui iš ankstesnio žingsnio, išvardyti visas galimas žymas antram žodžiui.



3 pav. Žymų išvardinimas antram žodžiui [9]

4 žingsnis. Įvertinti kiekvieną dviejų žodžių sakinio žymėjimą ir išmesti visą geriausią k.



4 pav. Žymėjimo galutinis įvertinimas [9]

Žemiau 1.1 lentelė pateikta kalbos dalių žymų anglų kalbai pavyzdys, tokios žymos naudojamos „Penn Treebank“.

1.1 lentelė. Kalbos dalių žymų anglų kalbai paaiškinimas. [10]

Number	Tag	Description
1.	CC	<i>Coordinating conjunction</i>
2.	CD	<i>Cardinal number</i>
3.	DT	<i>Determiner</i>
4.	EX	<i>Existential there</i>
5.	FW	<i>Foreign word</i>
6.	IN	<i>Preposition or subordinating conjunction</i>
7.	JJ	<i>Adjective</i>
8.	JJR	<i>Adjective, comparative</i>
9.	JJS	<i>Adjective, superlative</i>

10.	LS	List item marker
11.	MD	Modal
12.	NN	Noun, singular or mass
13.	NNS	Noun, plural
14.	NNP	Proper noun, singular
15.	NNPS	Proper noun, plural
16.	PDT	Predeterminer
17.	POS	Possessive ending
18.	PRP	Personal pronoun
19.	PRP\$	Possessive pronoun
20.	RB	Adverb
21.	RBR	Adverb, comparative
22.	RBS	Adverb, superlative
23.	RP	Particle
24.	SYM	Symbol
25.	TO	to
26.	UH	Interjection
27.	VB	Verb, base form
28.	VBD	Verb, past tense
29.	VBG	Verb, gerund or present participle
30.	VBN	Verb, past participle
31.	VBP	Verb, non-3rd person singular present
32.	VBZ	Verb, 3rd person singular present
33.	WDT	Wh-determiner
34.	WP	Wh-pronoun
35.	WP\$	Possessive wh-pronoun
36.	WRB	Wh-adverb

Morfologinio anotavimo lygmuo saugo metaduomenis apie patį žodį. Morfologinio anotavimo lygmenyje saugoma ši informacija apie žodį: visos žodžio prasmės, pagrindinė (antraštė) žodžio forma, kitaip dar vadinam lema, žodžio kalbos dalis ir įvairios morfologinės charakteristikos. Pasinaudojus saitynu Semantika.lt ir jame esančiu įrankiu, morfologinį anotavimą galima atvaizduoti pavyzdžiais (žr. 5 pav. ir 6 pav.).

Pirmame pavyzdyje asmens vardas ir pavardė nėra nurodomi (J.J.), todėl morfologinis anotatorius negali priskirti morfologinių charakteristikų, antrame pavyzdyje jau yra nurodoma asmens pavardė (J. Jonaitis) ir morfologinis anotatorius jau priskiria morfologines charakteristikas. Tai parodo kaip yra svarbu pateikti kuo tikslesnę informaciją, kad galėtume priskirti morfologines charakteristikas. Plačiau žiūrėti prieduose 8.2 priede.

```

      J      .      J
{"msd": [{"J", "X-"}, [{".", "Tp"}], [{"J", "X-"}],
      atėjo
[["ateiti", "Vgma3p---ni-"], ["ateiti", "Vgma3s---ni-"], ["ateiti", "Vgma3---ni-"], [{"i", "Sga"}],
      Kauno
[["Kaunas", "Npmsgng"], ["Kaunas", "Npmsgns"]],
      Technologijos      Universitetą .
[["technologija", "Ncfsgn-"], ["technologija", "Ncfpnn-"], ["technologija", "Ncfpvn-"], [{"universitetas", "Ncmsan-"}],
      pirmą
[["pirmas", "Momsaln"], ["pirmas", "Agpmsan"], ["pirmas", "Mofsaln"], ["pirmas", "Agpfsan"]],
      kartą
[["kartas", "Ncmsan-"], ["karta", "Ncfsan-"], ["karti", "Vgps-sfpnnan-p"], ["karti", "Vgps-smpnnan-p"]],
      1985m      rugsėjo      15d
[["1985m", "X-"], [{"rugsėjis", "Ncmsgn-"}], [{"15d", "X-"}], [{".", "Tp"}]]
      , "stem": ["J", ".", "J", "atėj", "i", "Kaun", "Technologij", "Universitet", "pirm", "kart", "1985m", "rugsėj", "15d", "."]}

```

5 pav. Morfologinio anotavimo rezultato struktūros pavyzdys pirmas

```

      J      .      Jonaitis
{"msd": [{"J", "X-"}, [{".", "Tp"}], [{"Jonaitis", "Npmsnns"}],
      atėjo
[["ateiti", "Vgma3s---ni-"], ["ateiti", "Vgma3s---ni-"], ["ateiti", "Vgma3---ni-"],
      i
[["i", "Sga"], [{"Kaunas", "Npmsgng"}, {"Kaunas", "Npmsgns"}],
      Kauno
[["technologija", "Ncfsgn-"], ["technologija", "Ncfpnn-"], ["technologija", "Ncfpvn-"], [{"universitetas", "Ncmsan-"}],
      Technologijos      Universitetą
      pirmą
[["pirmas", "Momsaln"], ["pirmas", "Agpmsan"], ["pirmas", "Mofsaln"], ["pirmas", "Agpfsan"]],
      kartą
[["kartas", "Ncmsan-"], ["karta", "Ncfsan-"], ["karti", "Vgps-sfpnnan-p"], ["karti", "Vgps-smpnnan-p"]],
      1985m      rugsėjo      15d
[["1985m", "X-"], [{"rugsėjis", "Ncmsgn-"}], [{"15d", "X-"}], [{".", "Tp"}]]
      , "stem": ["J", ".", "Jonait", "atėj", "i", "Kaun", "Technologij", "Universitet", "pirm", "kart", "1985m", "rugsėj", "15d", "."]}

```

6 pav. Morfologinio anotavimo rezultato struktūros pavyzdys antras

1.4.3.3. Sintaksinė analizė

Sintaksinė analizė – sakinio fragmentų skaidymas į sintaksės vienetus – kalbos dalis, frazes, nusakant jų tarpusavio sąryšius. Skiriami du pagrindiniai sąryšių tipai: hierarchinis (angl. *constituency-based*) ir priklausomybės (angl. *dependency-based*). Sintaksinės analizės rezultatas dažniausiai pateikiamas kaip sakinio sintaksinių vienetų sąryšių medis (angl. *parse tree*) [11].

Naudojant saityną Semantika.lt ir jame esantį įrankį, sintaksinį anotavimą galima atvaizduoti pavyzdžiais (žr. 7 pav., 8 pav., 9 pav.). Paveikslėliuose galime patyti, kad yra labai svarbu kaip yra pateikiama informaciją ir kokį skirtingą rezultatą gauname. Plačiau žiūrėti prieduose 8.2 priede.

Lietuviško teksto analizė ir taisymas

Paslaugos / Lietuviško teksto analizė ir taisymas 📖 Paslaugos aprašymas 🗨️ Pagalba

Analuojamas tekstas Rašybos klaidos (2) Gramatikos klaidos (0) Morfologija Įvardintos esybės (2) Žodžių junginiai Sintaksė

Tekstas:

J. J. atėjo į Kauno Technologijos Universitetą pirmą kartą 1985m rugsėjo 15d. studijos truko ketverius metus.

Žymėti pagal tipą:

Visi tipai ▼

Analizės suvestinė:

Analizuoti sakiniai	1
Neanalizuoti sakiniai	1

7 pav. Sintaksinio anotavimo rezultato pirmas pavyzdys

Lietuviško teksto analizė ir taisymas

Paslaugos / Lietuviško teksto analizė ir taisymas 📖 Paslaugos aprašymas 🗨️ Pagalba

Analuojamas tekstas Rašybos klaidos (2) Gramatikos klaidos (0) Morfologija Įvardintos esybės (2) Žodžių junginiai Sintaksė

Tekstas:

J. J. atėjo į Kauno Technologijos Universitetą pirmą kartą 1985m rugsėjo 15d. studijos truko ketverius metus.

Žymėti pagal tipą:

Visi tipai ▼

Analizės suvestinė:

Analizuoti sakiniai	1
Neanalizuoti sakiniai	2

8 pav. Sintaksinio anotavimo rezultato antras pavyzdys

Lietuviško teksto analizė ir taisymas

Paslaugos / Lietuviško teksto analizė ir taisymas 📖 Paslaugos aprašymas 🗨️ Pagalba

Analuojamas tekstas Rašybos klaidos (1) Gramatikos klaidos (0) Morfologija Įvardintos esybės (3) Žodžių junginiai Sintaksė

Tekstas:

J. Jonaitis atėjo į Kauno Technologijos Universitetą pirmą kartą 1985m rugsėjo 15d. Studijos truko ketverius metus.

Žymėti pagal tipą:

Visi tipai ▼

Analizės suvestinė:

Analizuoti sakiniai	2
Neanalizuoti sakiniai	0

9 pav. Sintaksinio anotavimo rezultato trečias pavyzdys

1.4.3.4. Įvardintų esybių atpažinimas tekste

Įvardinta esybė (angl. *named entity*) – tai įvardintos esybės kategorija, apibrėžianti tekste esančius žodžius ar frazes, pasižyminčius tomis pačiomis semantinėmis savybėmis. Pvz.:

- asmenvardžiu,
- organizacijos pavadinimu,
- vietovardžiu,
- data, laikotarpiu,
- ir t. t. ir pan.

Įvardintų esybių atpažinimas (angl. *Įvardintų esybių atpažinimas- Named Entity Recognition - NER R*) – tai paieškos procesas, kurio metu tekste surandamos įvardintos esybės.

Internetiniame „Semantika.lt“ puslapyje esantis įrankis, atpažinęs įvardintas esybes jas atvaizduoja kaip pateikta (10 pav.). Paveikslėlyje matome kokias įvardintas esybes pavyko atpažinti ir įvardinti skirtingais anotavimo atvejais. Plačiau žiūrėti prieduose 8.2 priede.

```
1 J. J. atėjo į Kauno Technologijos
2 Universitetą pirmą kartą 1985m rugsėjo 15d.
3 "ner": {
4   "ne": {
5     "money": [],
6     "dates": [
7       {
8         "text": "1985m rugsėjo 15d.",
9         "ref": [
10          11,
11          4
12        ]
13      }
14    ],
15    "products": [],
16    "organizations": [],
17    "locations": [
18      {
19        "text": "Kauno",
20        "ref": [
21          6,
22          1
23        ]
24      }
25    ],
26    "persons": []
27  },
28  "associations": {}
29
30
31
32
33
34
35
```

```
1 J. Jonaitis atėjo į Kauno Technologijos
2 Universitetą pirmą kartą 1985m rugsėjo 15d.
3 "ner": {
4   "ne": {
5     "money": [],
6     "dates": [
7       {
8         "text": "1985m rugsėjo 15d.",
9         "ref": [
10          10,
11          4
12        ]
13      }
14    ],
15    "products": [],
16    "organizations": [],
17    "locations": [
18      {
19        "text": "Kauno",
20        "ref": [
21          5,
22          1
23        ]
24      }
25    ],
26    "persons": [
27      {
28        "text": "J. Jonaitis",
29        "ref": [
30          0,
31          3
32        ]
33      }
34    ]
35  },
36  "associations": {}
37
```

10 pav. Įvardintų esybių anotavimo rezultato palyginamasis pavyzdys

Priklausomai nuo poreikio atpažįstant įvardintas esybes, galima rinktis tam tinkamą atpažinimo metodą. Žemiau pateikiama galimas įvardintų esybių atpažinimo metodų sąrašas.

Įvardintų esybių atpažinimo metodai:

- Taisyklėmis pagrįstas metodas (angl. *rule-base approach*);
- Statistinis metodas (angl. *statistical learning approach*) [12];
 - Paslėptas *Markovo* modelis (angl. *Hidden Markov Models*) [12];

- Maksimalios entropijos modelis (angl. *Maximum Entropy Markov Models*) [12];
- Sąlyginiai atsitiktiniai laukai (angl. *Conditional Random Fields*) [12];
- Ryšių ištraukimas (angl. *Relation Extraction*) [13];
 - Savybėmis grįstas klasifikavimas (angl. *Feature-based Classification*)
 - Karnel metodas (angl. *Kernel Methods*) [13];
 - Branduolių sekų pagrindu metodas (angl. *Sequence-based Kernels*) [13];
 - Medžiu grįstas branduolių metodas (angl. *Tree-based Kernels*) [13];
 - Prižiūrimas savaitinis mokymo metodas (angl. *Weakly Supervised Learning Methods*) [14];
- Neprižiūrimas mokymas (angl. *Unsupervised Information Extraction*) [15];
 - Ryšių radimas ir šablono indukcija (angl. *Relation Discovery and Template Induction*) [15];
 - Atviros informacijos ištraukimas (angl. *Open Information Extraction*) [15];

1.4.3.5. Semantinė analizė

Semantinių vaidmenų žymėjimas – dar kitaip vadinamas paviršutine semantine analize. Šis metodas skirtas iš anksto veiksmožodžių leksikone apibrėžtų vaidmenų priskyrimui sintaksiniam sakinio vienetams. Automatinio semantinių vaidmenų žymėjimo metodai grindžiami mašininio mokymo taikymu naujiems tekstams anotuoti jau suanotuotų tekstynų (angl. *gold-standard*) atžvilgiu. Labiausiai paplitę tekstynai ir leksikonai semantinių vaidmenų žymėjimui yra *PropBank* ir *FrameNet*.

Semantinė paieška pranašesnė už standartinę paiešką, surandančią informaciją pagal vartotojo įvestus raktinius žodžius, tuo, kad ji pateikia labiau vartotojo reikalavimus atitinkančius rezultatus. Pavaizduotose paveikslėliuose pasirinkta dvi skirtingos paieškos naršyklės t.y. „Google.com“ ir „Ontotex.com“.

The image shows a Google search results page for the query "donatas motiejūnas". The search bar at the top contains the text "donatas motiejūnas" and shows 633,000 results. Below the search bar, there are several search filters: "Viskas", "Vaizdai", "Naujienos", "Vaizdo įrašai", "Žemėlapiai", "Daugiau", "Nustatymai", and "Rankiai".

The main content area displays several search results:

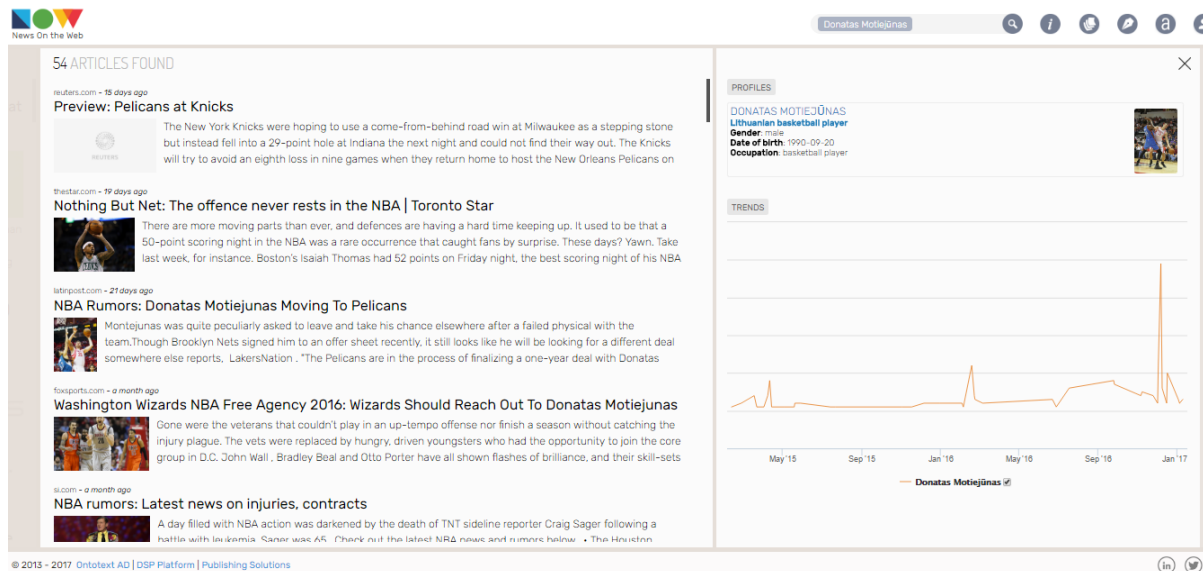
- A top result from ESPN: "Donatas Motiejūnas Stats, News, Videos, Highlights, Pictures, Bio ...". It includes a link to the player's profile and a snippet of text: "www.espn.com/nba/player/_id/6464/donatas-motiejunas".
- A result from NBA.com: "Donatas Motiejūnas stats, details, videos, and news. | NBA.com". It includes a link to the player's profile and a snippet of text: "www.nba.com/players/donatas/motiejunas/202700".
- A result from TV3.lt: "Donatas Motiejūnas prieš NBA čempionus sužaidė sezono mačą". It includes a link to the article and a snippet of text: "tv3.lt - Prieš 20 valandų".
- A result from Vakarų ekspresas: "Donatas Motiejūnas pelnė 14 taškų, o "Pelicans" nugalėjo NBA čempionus | ve.lt". It includes a link to the article and a snippet of text: "Vakarų ekspresas - Prieš 21 valandą".
- A result from DELFI.lt: "Donatas Motiejūnas - išsamiai DELFI.lt". It includes a link to the article and a snippet of text: "www.delfi.lt - Temos".

On the right side of the page, there is a detailed profile card for Donatas Motiejūnas, a Lithuanian basketball player. The card includes a photo of the player, his name "Donatas Motiejūnas", and his position "Krepšininkas". Below the name, there is a brief biography: "Donatas Motiejūnas – Lietuvos krepšininkas. Kauno krepšinio mokyklos „Aisčiai“ auklėtinis, rungtyniaujantis Houston „Rockets“ komandoje. Draftexpress.com reitinge tarp žaidėjų, gyvenančių ne JAV, jis užima antrąją vietą po ispano Ricky Rubio. Vikipedija". It also lists his birth date (1990 m. rugsėjo 20 d.), height (7 ft 0 in), and net worth (2.288 milijonų USD (2016 m.)). The card also mentions his parents (Vidas Motiejūnas) and his current team (Naujojo Orleono „Pelicans“). At the bottom of the card, there is a section titled "Žmonės taip pat ieškojo" (People also searched for) with a grid of player photos and names: Jonas Valančiūnas, Terrence Jones, Clint N'Dumba..., Mindaugas Kuzmickas, and Patrick Beverley.

11 pav. Google paieškos rezultatas

Abejose paieškos sistemose buvo parašyta užklausa „Donatas Motiejūnas“ -gauti rezultatai akivaizdžiai skirtingi. Google sistema dešinėje pusėje pateikia autobiografijos trumpus duomenis, o kairėje pusėje informacija kur šis žmogus paminėtas (žr. 11 pav.).

Ontotex kompanijos sukurtoje paieškoje yra galimybė pasirinkti sritį kur norima ieškoti duomenų apie šį žmogų: sportas, laisvalaikis, mokslas ir technologijos, verslas, tarptautinės naujienos. Akivaizdu, kad susiaurinus ir konkretizavus paieškos sritį gausiu tikslesnius duomenis apie vykdomą užklausą (žr. 12 pav.).



12 pav. „Ontotex „, semantinės paieškos rezultatas

1.4.4. Semantinės paieškos tikslumo vertinimo parametrai (matai)

Prasidedant 1987 metams buvo surengta keletas angliško teksto automatinio natūralios kalbos apdorojimo sistemų konferencijų MUC (angl. *Message Understanding Conference*). Šios konferencijos buvo skirtos įvertinti ir tarpusavyje palyginti teksto apdorojimo sistemas, kaip yra atliekamos tokios pačios užduotys. Užduočių metu buvo pateikta apdoroti tokie patys tekstai ir siekiama gauti tuos pačius rezultatus. Atlikus tyrimą priimta išvada, kad toks lyginimas negalėjo būti atliktas, nes skirtingi algoritmai vertinant skirtingo ilgio ar žanro tekstus gali kiekvieną kart apdoroti skirtingai. Kad tyrimas būtų teisingas reikėjo automatiniam sprendimams apdoroti pateikti tas pačias užduotis. Tuo tikslu buvo sukurtas tekstynas ir sudarytas tekstyno etalonas (angl. *gold standard*). Norint išvengti netikslumų dėl minėtų problemų, tekstas turėjo būti ilgas ir turintis daug skirtingų žodžių [16].

Surastas ir atpažintas esybės svarbu įvertinti ar jos yra tikslios. Semantinė paieška naudoja du vertinimo kriterijus, kaip įvertinti efektyvumą literatūroje. Tai yra tikslumo (angl. *precision*) ir jautrumas (angl. *recall*) parametrai [17]. Sėkmingai atlikus semantinę paiešką sistemoje parametrai būtų lygūs vienetui.

$$jautrumas = \frac{\text{rastų tinkamų dokumentų kiekis}}{\text{tinkamų dokumentų kiekis}}; \quad (1)$$

$$tikslumas = \frac{\text{rastų tinkamų dokumentų kiekis}}{\text{rastų dokumentų kiekis}}; \quad (2)$$

Norit padidinti rezultatų jautrumą reikia vartotojo užklausas formuluoti pakeičiant bendresnėmis ontologijoje numatytomis sąvokomis. Jautrumo rezultato padidinimas sumažina tikslumą, nes bendresnių sąvokų naudojimas paieškoje apima daugiau dokumentų išvedamame rezultate, kai didėjant jautrumui dokumentų skaičius didėja ir dokumentų, kurie neturi vartotojui reikalingos informacijos.

Šiuo atveju yra naudinga vartotojui leisti pasirinkti kokio tikslumo paiešką jis norėtų atlikti. Todėl vartotojas žinantis tiksliai ko ieško gaus tikslią informaciją rezultate, o vartotojas ieškantis bendresnės informacijos gaus platesnę gražinamą informaciją.

Siekiant įvertinti šių metodų išsamumą ir tikslumą, sukurtas bandomasis terminų identifikavimo tekstynas, kuriame neautomatiškai ekspertų sužymėti terminai yra laikomi auksiniu standartu (angl. *gold standard*) [17].

F_β - matas parodo kokią įtaką harmoniniam vidurkiui turi jautrumo ir tikslumo parametrai, F_1 reiškia, kad įtaka yra vienoda, jei beta yra daugiau už 1 didesnę įtaką turi jautrumas, jei beta mažiau nei vienetą didesnę įtaką turi tikslumas.

$$F_\beta = (1 + \beta^2) * \frac{\text{tikslumas} * \text{jautrumas}}{(\beta^2 * \text{tikslumas}) + \text{jautrumas}}; \quad (3)$$

Daugelio NLP (angl. *Natural Language Processing*) sistemų veikimo tikslumas yra vertinamas naudojant F-matą (angl. *f-measure*), kuris yra apibrėžiamas kaip tikslumo ir pateikimo metrikų harmoninis vidurkis, kuris apibūdina sistemos efektyvumą, taikant intervalą nuo nulio iki vieno [18].

$$F_{\text{matas}} = \frac{\text{jautrumas} * \text{tikslumas}}{\frac{\text{jautrumas} + \text{tikslumas}}{2}}; \quad (4)$$

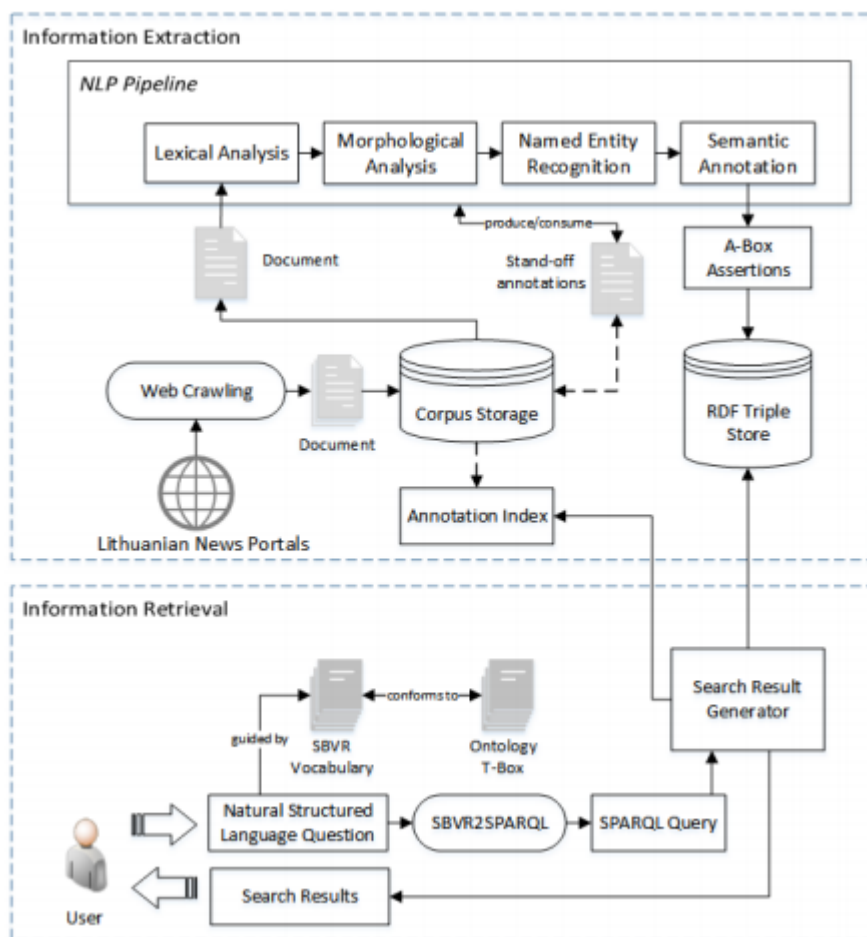
1.5. Tyrimo objekto naudotojų analizė

Analizuojamos ontologijos vartotojai yra pačios sistemos kūrėjai, kurie tvarko, prižiūri, ruošia informaciją. Sistemą kuriantys nori, kad naudotojus dominanti informacija būtų tiksli ir tikslumo kriterijus būtų atvaizduotas naudotojui matoma forma. Sistemos kūrėjai sprendžia, kaip bus vertinamas šaltinių tikslumas.

1.6. Esamų problemos sprendimo metodų analizė (Lietuvos ir tarptautiniu mastu)

LKSSAIS sprendimas

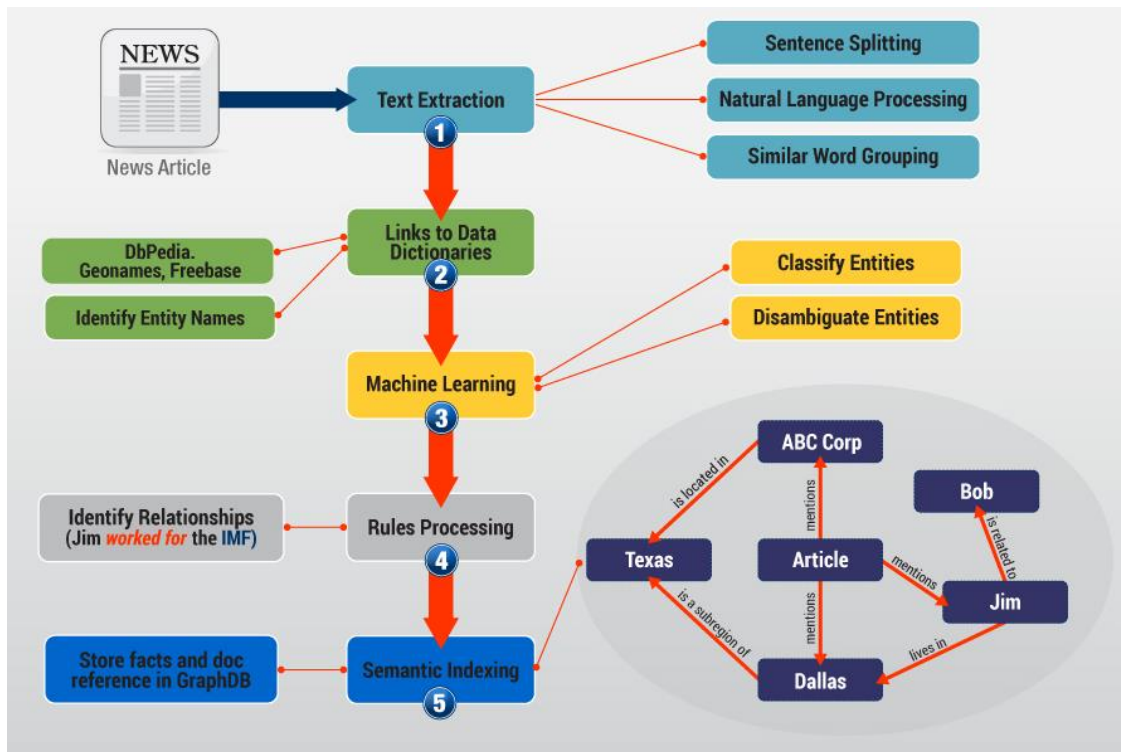
Yra sukurta lietuvių kalbos sintaksinės – semantinės analizės sistema. Semantinės paieškos LKSSAIS karkasas pavaizduotas 13 pav. Schemoje matoma ne tik natūralios kalbos anotavimo grandinė, bet ir pilnas procesas informacijos apdorojimo ir gražinimo naudotojui. Šios sistemos kūrėjai nėra išanalizavę ir pateikę konkrečių sprendimų, kaip įvertinti paieškos tikslumą kiekvienam natūralios kalbos anotavimo komponentui.



13 pav. Semantinės paieškos LKSSAIS karkasas

„Ontotext“ kompanijos sprendimas

Organizacija „Ontotext“, pateikia savo sprendimą semantinei paieškai (žr. 14 pav.). Teksto atpažinimo dalis vaizduojama pirmame blokelyje. Patys kūrėjai šį blokelį aprašo taip : „Tekstas yra išgaunamas iš straipsnių, dokumentų ar bet kurių nestruktūrotų duomenų formų. Sakiniai yra padalinti ir suskirstyti į kalbos dalis. Skirtingos formos su tais pačiais žodžiais yra sugrupuotos [19].“



14 pav. Semantinės paieškos modelis [20]

„Ontotext“ kompanijos vizija remiasi, kad technologijos ir verslas - tai teksto ir duomenų suvokimas. Didinant žinių grafą yra pagerinamas teksto analizės tikslumas. Tokiu būdu gerėja paieška, ištyrimas, klasifikacija ir rekomendacijos įvairiose paieškos srityse. „Ontotext“ kompanijos platforma ir platformos projektavimo pasirinkimas yra skirtas apdoroti didelius nestruktūruoto turinio kiekius, naudojant labai didelius žinių grafus, užtikrinant puikią anotacijos kokybę ir efektyviausią duomenų valdymą [21].

Semantinio anotavimo kokybei tiesioginę įtaką turi naudojamos programos. Kokybės užtikrinimui „Ontotex“ kompanija naudoja „Curation“ įrankį ir analizuoja šio įrankio įtaką gerinant anotavimo kokybę. Įrankio tikslas pagerinti komentarų kokybę, žodyno išsamumą ir teksto analizės tikslumą (angl. *precision*) ir gražinimą (angl. *recall*). Prie anotuoto teksto yra pateikiama kokybės parametrai (žr. 15 pav.) [21] :

- Parametras, kuris parodo su kokia tikimybe mašininio mokymo algoritmas, atpažino tam tikros klasės objektą (angl. *confidence*) [22];
- Procentinis termino dažnis tekste, įvertinat nuokrypį nuo teksto antrašte (angl. *relevance*) [22].

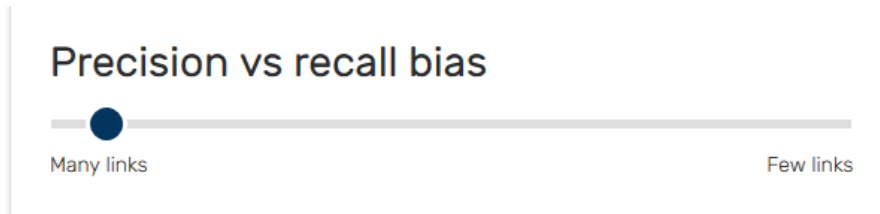
More than a decade after its first Central Asia strategy was unveiled, the European Union has released an update, i new opportunities for a stronger partnership." While the strategy is a hand interests with reg which serve as a glaring reminder that policies on paper can be far removed abilities in the reg Central Asia's importance to the EU is rooted, per the strategy, in its strat / resources, its r security. The strategy focuses on three priority areas: "Partnering for Resilience," "Partnering for Prosperity," and "Wc crowded a bevy of interrelated matters.



15 pav. „Ontotex“ kompanijos saityno anotuoto teksto ištrauka

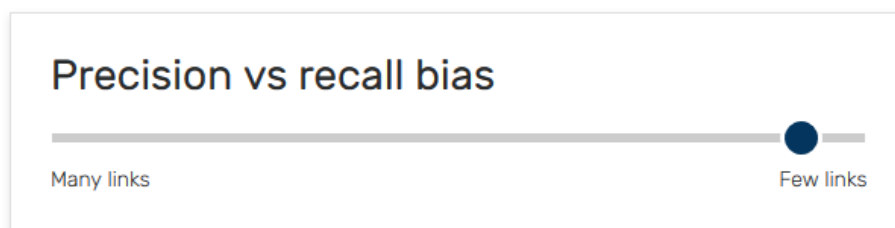
Anotacijas reikia tvarkyti jeigu pastebima dviprasmybių arba kai parametras (angl. *confidence*) taškų ribos yra per mažos. Yra svarbu, kad dviprasmiški atvejai būtų fiksuojami ir išaiškinami, kad tikslumas būtų kuo didesnis.

„Ontotex“ saityne galima filtruoti duomenų rezultatus pagal pageidaujamą tikslumą. Žemiau pateiktuose 16 pav. ir 17 pav. matosi išvedamų rezultatų kiekio pokytis.



Central Asia's importance to the EU is rooted, per the strategy, in its strategic location, its energy resources, its market and, security. The strategy focuses on three priority areas: "Partnering for Resilience," "Partnering for Prosperity," and "Working Better crowded a bevy of interrelated matters.

16 pav. Anotuoto teksto rezultatas



Central Asia's importance to the EU is rooted, per the strategy, in its strategic location, its energy resources, its market and, of course, security. The strategy focuses on three priority areas: "Partnering for Resilience," "Partnering for Prosperity," and "Working Better Together." crowded a bevy of interrelated matters.

17 pav. Anotuoto teksto rezultatas

Sprendimas medicinoje

Semantinė paieška gali būti naudinga ir medicinoje. Šiuolaikiniame pasaulyje ligos sparčiai progresuoja ir vis dažniau remiasi tik laikinomis sąvokomis. Suradus automatizavimo sprendimus skirtus laiko santykiui su medicinoje naudojamais laiko apibrėžimais, leistų turėti duomenis, kurie leistų atskleisti ligos simptomus ir paciento esamą situaciją. Mokslininkai iš „*Mayo Clinic*“ ir „*University of Colorado*“ užsibrėžė sau tikslus:

- 1) Įvertinti galimų technologijų, tokių kaip gryniesi analizatoriai (*off-the-shelf deep parsers*) ir semantinio tikslumo žymėjimai, siekiant nustatyti ar reikės papildomų kalbų tekstynų (angl. *Treebanking*) ir tokių tekstynų kūrimo (angl. *PropBanking*), kuriame yra anotuoti veiksmožodiniai teiginiai su jų argumentais [23];
- 2) Sukurti laiko anotacijos schemą klinikinėje srityje, pagrįstą žymų darbalaukio ir įvykių išraiškoms žymėti (angl. *TimeML*), ir pridėti anotacijos modulį *Knowtator* (<http://knowtator.sourceforge.net/>), kad atliktų anotacijas [23];

Norėdami pasiekti savo tikslus mokslininkai sukūrė duomenų rinkinį, kuris buvo sudarytas iš „*Mayo Clinic*“ pateiktų duomenų (klinikinių ir radiologinių užrašų). Medicininiai duomenys yra sudėtingi užrašai, daugiau kaip trečdalis yra sakinių frazės, kurios nėra tinkamai suprantamos *off-the-shelf* analizatoriaus. Tyrimo metu buvo aptikta tik 11% teisingų semantinių atitikmenų, nors 42% procentai turėjo teisingus semantinius atitikmenis. Didelis netikslumas gali būti dėl elementų, kurie nebuvo įtraukti į aprašytus duomenis arba atsirado ten per klaidą. Didelę dalį tokių problemų galima išspręsti skiriant pakankamai dėmesio duomenų paruošimui, ypač jei tiksliai pavadintos subjekto (NE) žymos gali būti naudojamos kaip pradiniai duomenys specializuotam žodynui. Mokslininko *Eugene Charniak Postagger* modelio rezultatai pablogėjo 5%, o F-matas pablogėjo 8%, bendrai nuo 88,25% iki 80,84%. [23] Bandant pagerinti analizatoriaus našumą mokslininkai mano, kad reikia atsižvelgti į tokius dalykus kaip leksikos ir sintaksės konstrukcijos neatitikimai. Jeigu sintaksinės konstrukcijos neatitikimai yra vidutiniai, analizatoriaus tikslumas gali būti gerokai padidintas sprendžiant leksikos problemas. Jeigu sintaksinės konstrukcijos neatitikimai yra dideli reikia papildomo kalbų tekstyno (angl. *treebanking*). Mokslininkas *Walter Daelemans* tai apibūdina kaip „paprastą bet veiksmingą sprendimo būdą (perkeliama problema): modulių perkvalifikavimas naudojant anotacijų duomenis iš naujo domeno“ [23]. Dažniausiai pasitaikančių klaidų dėl sakinių fragmentų dominavimas yra sintaksės konstrukcijos neatitikimo pavyzdys.

1.7. Siekiamo sprendimo apibrėžimas

Siekiamas sprendimas - individų atpažinimo tikslumo parametru papildytos ontologijos modelis.

1.8. Analizės išvados

Tiriamąjį darbo pirmos dalies išvados:

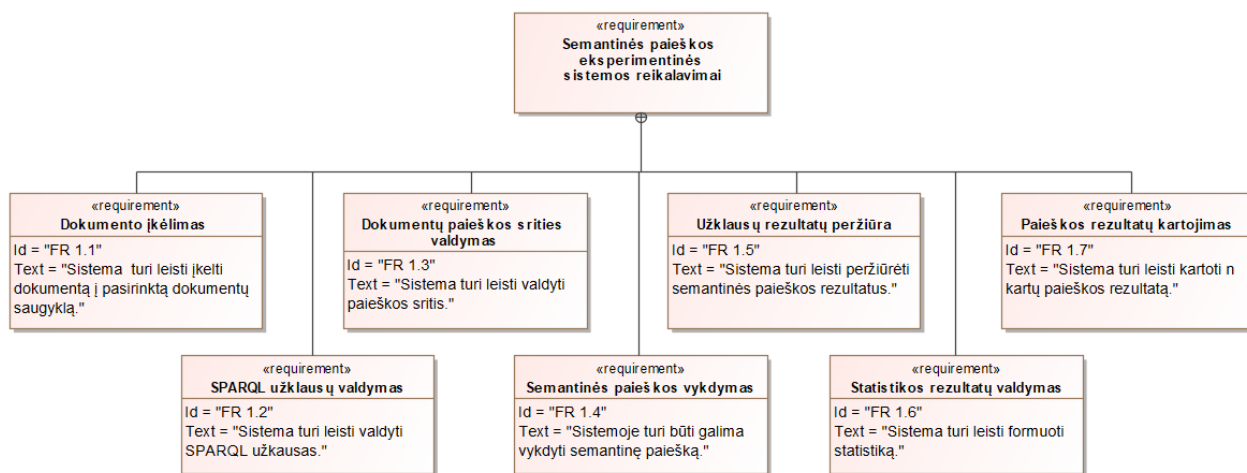
1. Išanalizavus tekstų analizės ir informacijos ištraukimo būdus, padaryta išvada, kad informacijos ištraukimo kokybė tiesiogiai priklauso nuo natūralios kalbos apdorojimą atliekančių komponentų veikimo kokybės. O tai tiesiogiai įtakoja ir paieškos kokybę.
2. Išanalizavus informacijos ištraukimo kokybės vertinimo būdus, nustatyta, kad teksto analizę atliekančių komponentų veikimo tikslumo įvertis yra kompleksinis, priklausantis nuo komponento realizuojamo algoritmo ir nuo tekstyno, kurio atžvilgiu vertinamas analizės rezultatas.
3. Semantinės paieškos LKSSAIS sprendimo karkasas suteikia semantinės paieškos galimybę, bet išvestas rezultatas yra atsitiktinis. Nėra galimybės rikiuoti ar filtruoti pateiktus rezultatus.
4. „Ontotext“ kompanijos sprendimas saugo semantinės analizės komponentų veikimo tikslumą nusakančių parametru įverčius. Tačiau nėra duomenų kiek tokio parametro saugojimas turi įtakos greitaveikai. Todėl lieka neaišku ar tai neapsunkina paieškos ir ar galima tikėtis efektyvios paieškos.
5. Įvertinus atliktą analizę, yra poreikis ištirti kiek tikslumo įverčių pridėjimas ontologijoje atsilieptų semantinės paieškos greitaveikai.

2. ONTOLOGIJOS PRATURTINTOS INDIVIDŲ ATPAŽINIMO TIKSLUMO ĮVERČIAIS MODELIO SPRENDIMO REIKALAVIMŲ SPECIFIKACIJA IR PROJEKTAS, FORMALUS APRAŠAS

Šio darbo tyrimo tikslas – sudaryti ir ištirti galimybę semantinės paieškos vartotojams matyti ontologijos praturtintos individų atpažinimo tikslo įverčiais tikslo parametras ir juo remiantis atrinkti, rikiuoti užklauso rezultatus. Reikalavimų specifikacijoje bus parodyti reikalavimai skirti eksperimentinei sistemai.

2.1. Funkcinių reikalavimų specifikacija

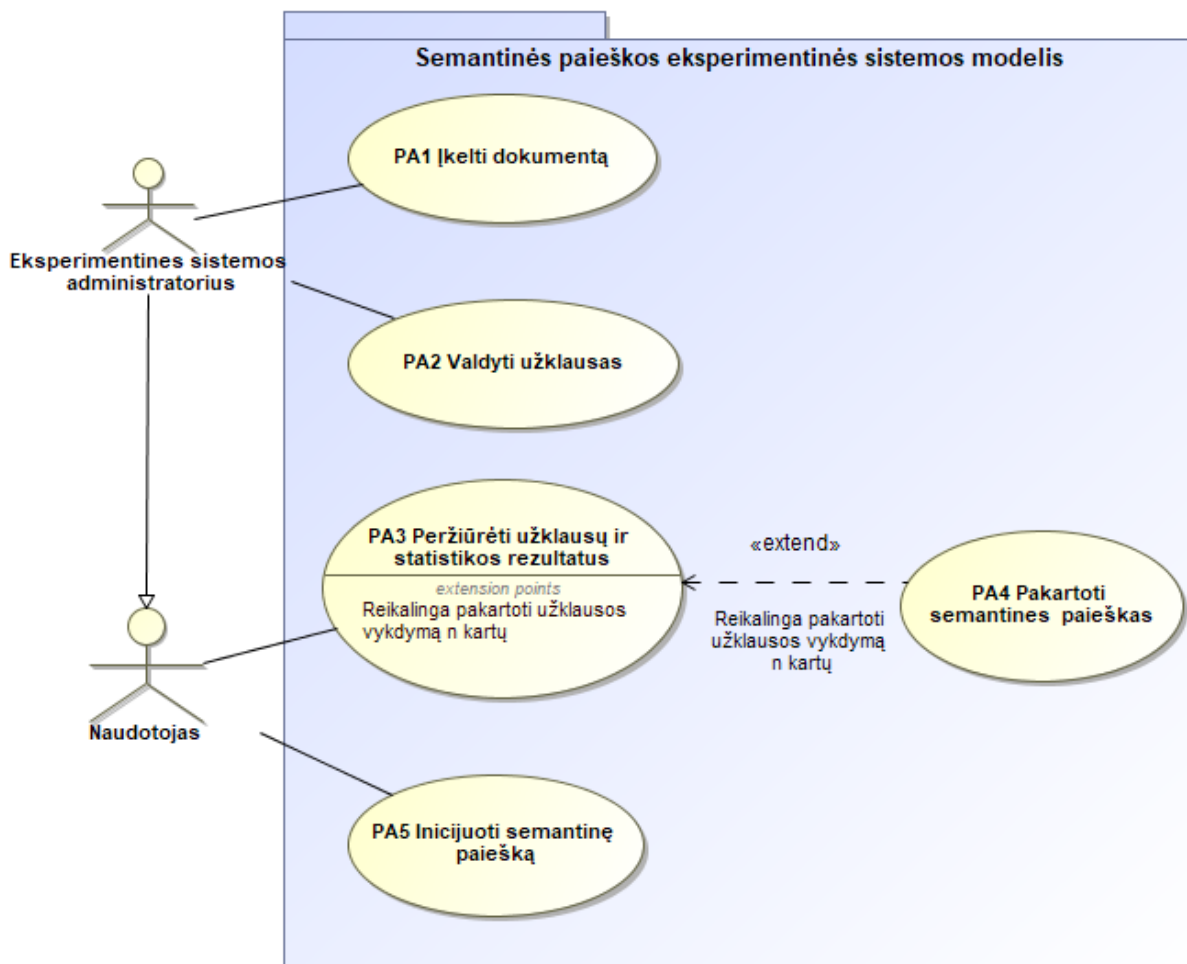
Funkciniai reikalavimai aprašo eksperimentinės sistemos funkcionalumą. Baigiamajame darbe yra suformuluoti septyni funkciniai reikalavimai eksperimentinei sistemai išpildyti. Reikalavimuose apibrėžiama ką sistema turi daryti ir kaip elgtis konkrečioje situacijoje. Funkciniai reikalavimai eksperimentinei sistemai pateikti 18 pav.



18 pav. Semantinės paieškos eksperimentinės sistemos reikalavimai

2.2. Ontologijos praturtintos individų atpažinimo tikslo įverčiais sprendimo panaudojimo atvejų modelis

19 pav. pateikti ontologijos praturtintos individų atpažinimo tikslo įverčiais modelio panaudojimo atvejai. Panaudojimo atvejų modelyje yra pateikiama visos naudotojų valdymo galimybės eksperimentinėje sistemoje. Modelyje yra išskirti du naudotojai – eksperimentinės sistemos administratorius ir naudotojas.



19 pav. Semantinės paieškos eksperimentinės sistemos PA modelis

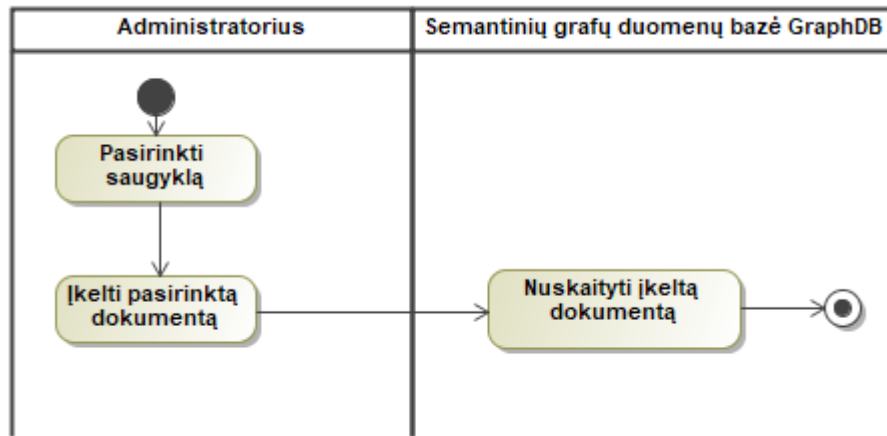
Detalus panaudojimo atvejų aprašas

Žemiau esančiose lentelėse pateikiama semantinės paieškos eksperimentinės sistemos panaudojimo atvejų modelio detalus aprašas.

2.1 lentelė. Panaudojimo atvejis: „Įkelti dokumentą“

1. PANAUDOJIMO ATVEJIS: Įkelti dokumentą		
Tikslas/uždavinys	Įkelti dokumentą iš saugyklos į sistemą.	
Aprašymas	Apima procesą, kurio metu eksperimentinės sistemos administratorius inicijuoja dokumento paėmimą ir perkėlimą į eksperimentinės sistemos saugyklą.	
Prieš-sąlyga	Eksperimentinės sistemos administratorius yra užklauso pateikimo lange.	
Aktorius:	Eksperimentinės sistemos administratorius.	
Sužadinimo sąlyga	Naudotojas nori inicijuoti dokumento paėmimą.	
Išpildomo reikalavimo numeris	FR 1.1	
Susiję panaudojimo atvejai	Išplečiantys PA	-
	Apimami PA	-
	Specializuoti PA	-
Veiklos taisyklės	- Eksperimentinės sistemos administratorius gali vienu metu įkelti ne daugiau nei vieną dokumentą.	
Pagrindinis scenarijus		
Pagrindinis įvykių srautas	Sistemos reakcija	

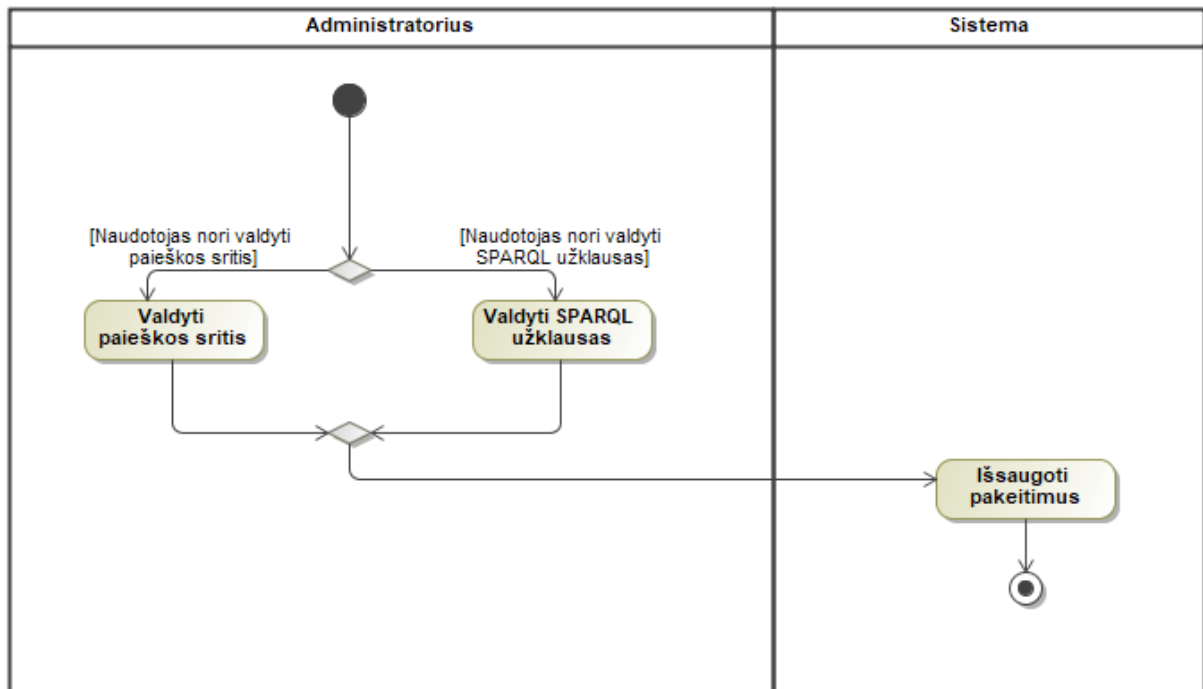
1. Eksperimentinės sistemos administratorius pasirenka dokumentą įkėlimui.	1.1 Sistema tikrina dokumento teisingumą
2. Eksperimentinės sistemos administratorius inicijuoja dokumento paėmimą.	2.1 Sistema įkelia dokumentą
Po-sąlyga	Išsaugomas įkeltas dokumentas
Alternatyvūs scenarijai	
Nėra	Nėra



20 pav. PA „Įkelti dokumentą“ scenarijus

2.2 lentelė. Panaudojimo atvejis: „Valdyti užklausas“

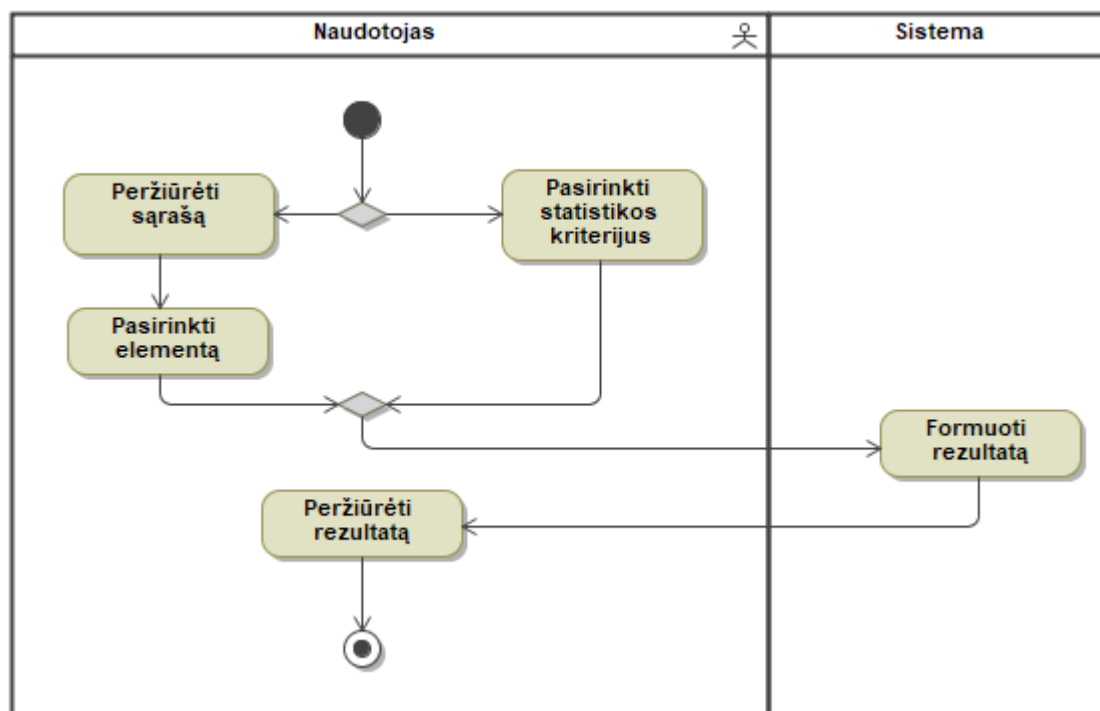
2. PANAUDOJIMO ATVEJIS: Valdyti užklausas		
Tikslas/uždavinys	Valdyti paieškos sritis ir SPARQL užklausas pagal poreikius.	
Aprašymas	Apima procesą, kurio metu naudotojas gali paieškos sritis ir užklausas kurti, redaguoti ir jas šalinti.	
Prieš-sąlyga	Eksperimentinės sistemos administratoriui reikalinga nauja paieškos sritis ar SPARQL užklausa.	
Aktorius:	Eksperimentinės sistemos administratorius.	
Sužadinimo sąlyga	Eksperimentinės sistemos administratorius nori sukurti naują paieškos sritį ar SPARQL užklausa	
Išpildomo reikalavimo numeris	FR 1.2	
Susiję panaudojimo atvejai	Išplečiantys PA	-
	Apimami PA	-
	Specializuoti PA	-
Veiklos taisyklės	✓	
Pagrindinis scenarijus		
Pagrindinis įvykių srautas	Sistemos reakcija	
1. Eksperimentinės sistemos naudotojas užklausų srities lange pasirenka mygtuką „Sukurti naują“	1.1. Atsidariusiame lange sistema pateikia detalesnį kūrimo langą.	
2. Naudotojas suveda naujus duomenis.	2.1.	
3. Naudotojas pasirenka mygtuką „Saugoti“	3.1. Sistema išsaugo naują įrašą.	
Po-sąlyga	Eksperimentinės sistemos administratorius sukuria naują užklausa	
Alternatyvūs scenarijai		
Nėra	Nėra	



21 pav. PA „Valdyti užklaudas“ scenarijus

2.3 lentelė. Panaudojimo atvejais: „Peržiūrėti užklauso ir statistikos rezultatus“

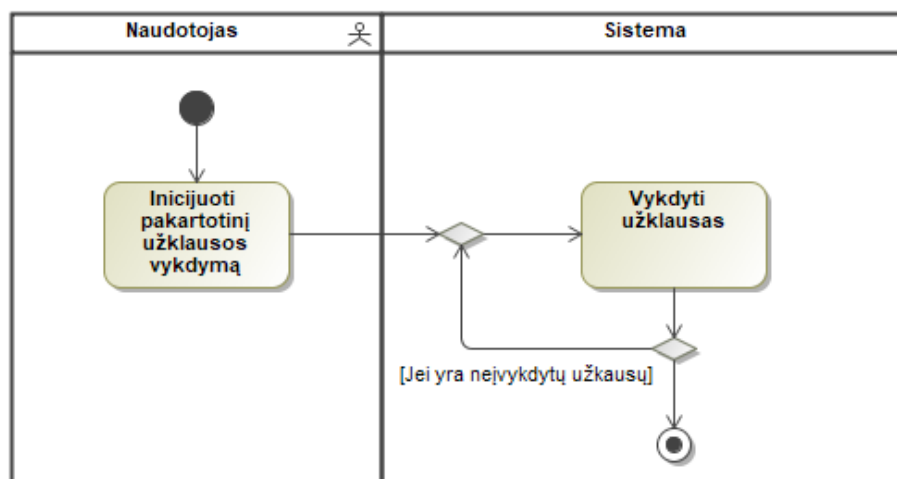
3. PANAUDOJIMO ATVEJIS: Peržiūrėti užklauso ir statistikos rezultatus		
Tikslas/uždavinys	Peržiūrėti užklauso rezultatus ir statistiką	
Aprašymas	Apima procesą, kurio metu naudotojas gali peržiūrėti semantinės paieškos rezultatus.	
Prieš-sąlyga	Įvykdytas PA.5 Inicijuoti semantinę paiešką.	
Aktorius:	Eksperimentinės sistemos administratorius /Naudotojas	
Sužadinimo sąlyga	Naudotojas nori peržiūrėti paieškos rezultatus.	
Išpildomo reikalavimo numeris	FR 1.5 FR 1.6	
Susiję panaudojimo atvejai	Išplečiantys PA	PA4 Pakartoti semantines paieškas
	Apimami PA	-
	Specializuoti PA	-
Veiklos taisyklės	✓	
Pagrindinis scenarijus		
Pagrindinis įvykių srautas		Sistemos reakcija
1. Naudotojas pasirenka paieškos rezultatą detalesnei peržiūrai.		1.1 Atsidariusiame lange sistema pateikia detalesnį paieškų rezultatą.
2. Naudotojas spaudžia mygtuką „Peržiūrėti“.		2.1 Sistema pasirinktą testą pateikia su detaliais paieškos rezultatais..
Po-sąlyga	Peržiūrimas dokumentas suanotuotas.	
Alternatyvūs scenarijai		
Nėra	Nėra	



22 pav. PA „Peržiūrėti užklausų ir statistikos rezultatus“ scenarijus

2.4 lentelė. Panaudojimo atvejais: „Pakartoti semantines paieškas“

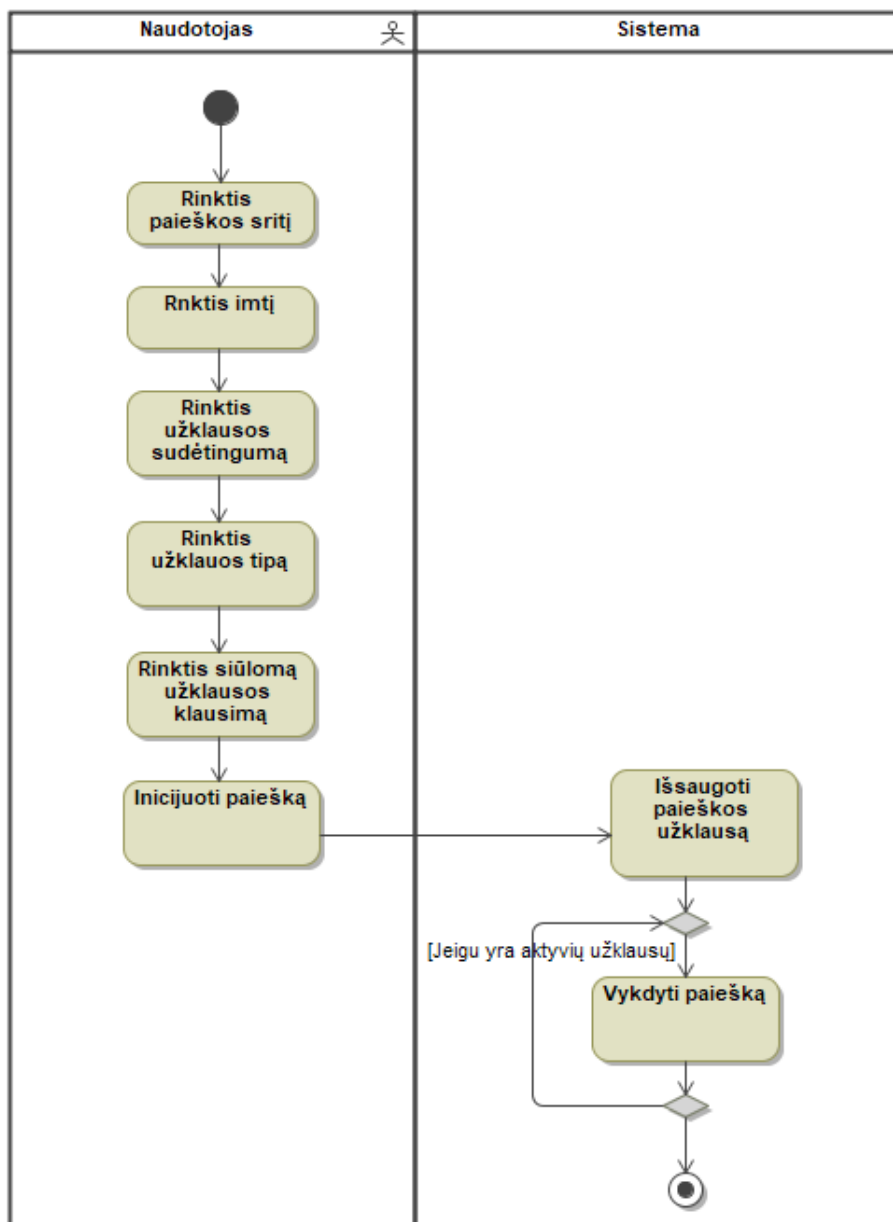
4. PANAUDOJIMO ATVEJIS: Pakartoti semantines paieškas		
Tikslas/uždavinys	Pakartoti vykdytą semantinę paiešką.	
Aprašymas	Apima procesą, kurio metu naudotojas nori pakartoti nustatytą n kartų jau vykdytą semantinę paiešką.	
Prieš-sąlyga	Naudotojas turi įvykdytą užklausą pagal nustatytus kriterijus.	
Aktorius:	Naudotojas	
Sužadinimo sąlyga	Naudotojas nori pakartoti vykdytą semantinę paiešką.	
Išpildomo reikalavimo numeris	FR 1.7	
Susiję panaudojimo atvejai	Išplečiantys PA	
	Apimami PA	-
	Specializuoti PA	-
Veiklos taisyklės	✓ Naudotojas gali pakartoti tik jau vykdytas semantines paieškas.	
Pagrindinis scenarijus		
Pagrindinis įvykių srautas	Sistemos reakcija	
1. Naudotojas pasirenka vykdytą paiešką.	1.1 Sistema atidaro rezultato langą.	
2. Naudotojas lauke „Kiek kartų vykdyti“ įveda skaičių.		
3. Naudotojas spaudžia „Vykdyti“.	3.1 Sistema įvykdo paiešką nustatytą n kartų.	
Po-sąlyga	Pasirinkta užklausa įvykdyta nustatytą n kartų.	
Alternatyvūs scenarijai		
Nėra	Nėra	



23 pav. PA „Pakartoti semantines paieškas“ scenarijus

2.5 lentelė. Panaudojimo atvejis: „Inicijuoti semantinę paiešką“

5. PANAUDOJIMO ATVEJIS: Vykdyti semantinę paiešką		
Tikslas/uždavinys	Vykdyti semantinę paiešką pagal užklausą	
Aprašymas	Apima procesą, kurio metu naudotojas vykdo semantinę paiešką pagal pasirinktus paieškos parametrus.	
Prieš-sąlyga	Naudotojas yra prisijungęs prie semantinės paieškos sistemos.	
Aktorius:	Naudotojas/ Eksperimentinės sistemos administratorius	
Sužadinimo sąlyga	Naudotojas nori vykdyti semantinę paiešką.	
Išpildomo reikalavimo numeris	FR 1.4	
Susiję panaudojimo atvejai	Išplečiantys PA	-
	Apimami PA	-
	Specializuoti PA	-
Veiklos taisyklės	✓ Naudotojas privalo pasirinkti paieškos parametrus.	
Pagrindinis scenarijus		
Pagrindinis įvykių srautas	Sistemos reakcija	
1. Paieškos srities vietoje naudotojas pasirenka paieškos sritį.	1.1 Sistema pateikia galimas paieškos sritis.	
2. Paieškos srities vietoje naudotojas pasirenka imtį.	2.1 Sistema pateikia galimas imtis.	
3. Paieškos srities vietoje naudotojas pasirenka užklauso sudėtingumą.	3.1 Sistema pateikia galimus užklauso sudėtingumus..	
4. Paieškos srities vietoje naudotojas pasirenka užklauso tipą.	4.1 Sistema pateikia galimus užklauso tipus	
5. Paieškos srities vietoje naudotojas pasirenka paieškos klausimą.	5.1 Sistema pateikia galimus paieškos klausimus.	
6. Naudotojas inicijuoja paiešką.	6.1 Sistema vykdo paiešką.	
Po-sąlyga	Įvykdyta semantinė paieška.	
Alternatyvūs scenarijai		
1a. Naudotojas nepasirenka nors vieno parametro.	1a. 1 Sistema nevykdo semantinės paieškos.	



24 pav. PA „Vykdėti semantinę paiešką“ scenarijus

2.2.1. Nefunkciniai reikalavimai

2.2.1.1. Reikalavimai sistemos išvaizdai

Šiame poskyryje pateikiamai eksperimentinės sistemos nefunkciniai reikalavimai sistemos išvaizdai.

2.6 lentelė. Nefunkcinio reikalavimo NF-1 aprašas

Reikalavimas #:	NF-1	Reikalavimo tipas:	1	Įvykis/panaudojimo atvejis #:	PA.2 ,3,4,5
Aprašymas:	Sistemos naudotojo grafinė sąsaja turi būti lengvai perprantama				
Pagrindimas:	Semantinės paieškos naudotojas turi lengvai susiorientuoti pateikiamuose naršyklės languose ir išmokti sistema naudotis per 1 valandą.				
Atitikimo kriterijus:	Lengvai suprantama vartotojo grafinė sąsaja per 1 valandą.				

2.7 lentelė. Nefunkcinio reikalavimo NF-2 aprašas

Reikalavimas #:	NF-2	Reikalavimo tipas:	2	Įvykis/panaudojimo atvejis #:	PA.2 ,3,4,5
Aprašymas:	Paieškos sistemos naudotojo sąsajos valdymo elementai turi būti standartinės naršyklės palaikomi valdymo elementai, navigacijos mygtukai ir pan.				

Pagrindimas:	Naudotojas dažniausiai geba naudotis interneto naršyklės programa.
Atitikimo kriterijus:	Naudotojo sąsajos valdymo elementai veikia naudojant naršyklės valdymo elementus.

2.8 lentelė. Nefunkcinio reikalavimo NF-3 aprašas

Reikalavimas #:	NF-3	Reikalavimo tipas:	3	Įvykis/panaudojimo atvejis #:	PA.2 ,3,4,5
Aprašymas:	Užvedus žymeklį ant naudojamų funkcijų turi atsirasti mažas iššokantis paaiškinimą atvaizduojantis informacinis langelis.				
Pagrindimas:	Tikslas naudotojui paaiškinti nesuprantamus meniu punktus.				
Atitikimo kriterijus:	Paaiškinti meniu punktai.				

2.9 lentelė. Nefunkcinio reikalavimo NF-4 aprašas

Reikalavimas #:	NF-4	Reikalavimo tipas:	4	Įvykis/panaudojimo atvejis #:	PA.2 ,3,4,5
Aprašymas:	Paieškos sistemos naudotojo aplinka turi būti neperkrauta papildoma informacija, esami elementai patogiai išdėstyti.				
Pagrindimas:	Semantinės paieškos naudotojas neturi būti blaškomas pašalinės informacijos bei grafikos pertekliumi.				
Atitikimo kriterijus:	Neperkrauta ir informatyvi naudotojo sąsaja – nėra papildomos reklamos.				

2.2.1.2. Reikalavimai panaudojimui

Šiame poskyryje pateikiamai eksperimentinės sistemos nefunkciniai reikalavimai sistemos panaudojimui.

2.10 lentelė. Nefunkcinio reikalavimo NF-5 aprašas

Reikalavimas #:	NF-5	Reikalavimo tipas:	5	Įvykis/panaudojimo atvejis #:	PA.2 ,3,4,5
Aprašymas:	Semantinės paieškos naudotojas turi galėti naudotis sistema be išankstinio apmokymo.				
Pagrindimas:	Daugelis moka naudotis internetu, o sistema pasiekama per interneto naršyklę.				
Atitikimo kriterijus:	90% darbo be klaidų dirbant su sistema pirmą kartą				

2.11 lentelė. Nefunkcinio reikalavimo NF-6 aprašas

Reikalavimas #:	NF-6	Reikalavimo tipas:	6	Įvykis/panaudojimo atvejis #:	PA.2 ,3,4,5
Aprašymas:	Paieškos duomenų įvedimas naudotojui turi būti aiškus ir paprastas.				
Pagrindimas:	Paieškos sistema naudotojas naudosis jei ji bus paprasta ir aiški.				
Atitikimo kriterijus:	95% darbo be klaidų dirbant su sistema įvedant duomenis.				

2.2.1.3. Reikalavimai veikimo sąlygoms

Šiame poskyryje pateikiamai eksperimentinės sistemos nefunkciniai reikalavimai sistemos veikimo sąlygoms.

2.12 lentelė. Nefunkcinio reikalavimo NF-7 aprašas

Reikalavimas #:	NF-8	Reikalavimo tipas:	8	Įvykis/panaudojimo atvejis #:	PA.5
Aprašymas:	Semantinės užklauskos turi būti vykdomos SPARQL kalba.				
Pagrindimas:					
Atitikimo kriterijus:	Naudotojas vykdo semantinę užklauską SPARQL kalba ir gauna atsakymą.				

2.13 lentelė. Nefunkcinio reikalavimo NF-8 aprašas

Reikalavimas #:	NF-9	Reikalavimo tipas:	9	Įvykis/panaudojimo atvejis #:	PA.2
Aprašymas:	Sistema turi leisti pasirinkti iš kelių iš anksto paruoštų paieškos užklauskų.				
Pagrindimas:	Vartotojas turi turėti galimybę matyti užklauskų pavyzdžius.				
Atitikimo kriterijus:	Sistemos naudotojas gali pasirinkti paruoštą užklauską.				

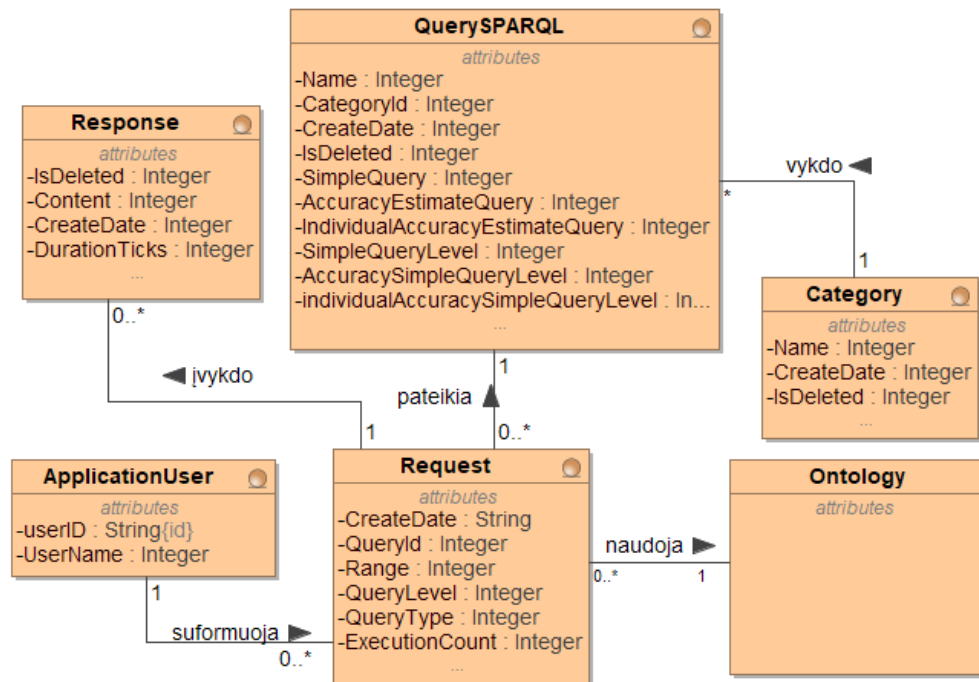
2.14 lentelė. Nefunkcinio reikalavimo NF-19 aprašas

Reikalavimas #:	NF-10	Reikalavimo tipas:	10	Įvykis/panaudojimo atvejis #:	PA.2,5
Aprašymas:	Sistema turi leisti matyti pačią SPARQL užklausą ir ją koreguoti.				
Pagrindimas:	Naudotojas geriau žino kokią SPARQL užklausą reikalinga.				
Atitikimo kriterijus:	Sistemos naudotojas koreguoja SPARQL užklausas pagal savo poreikį.				

2.3. Dalykinės srities modelis

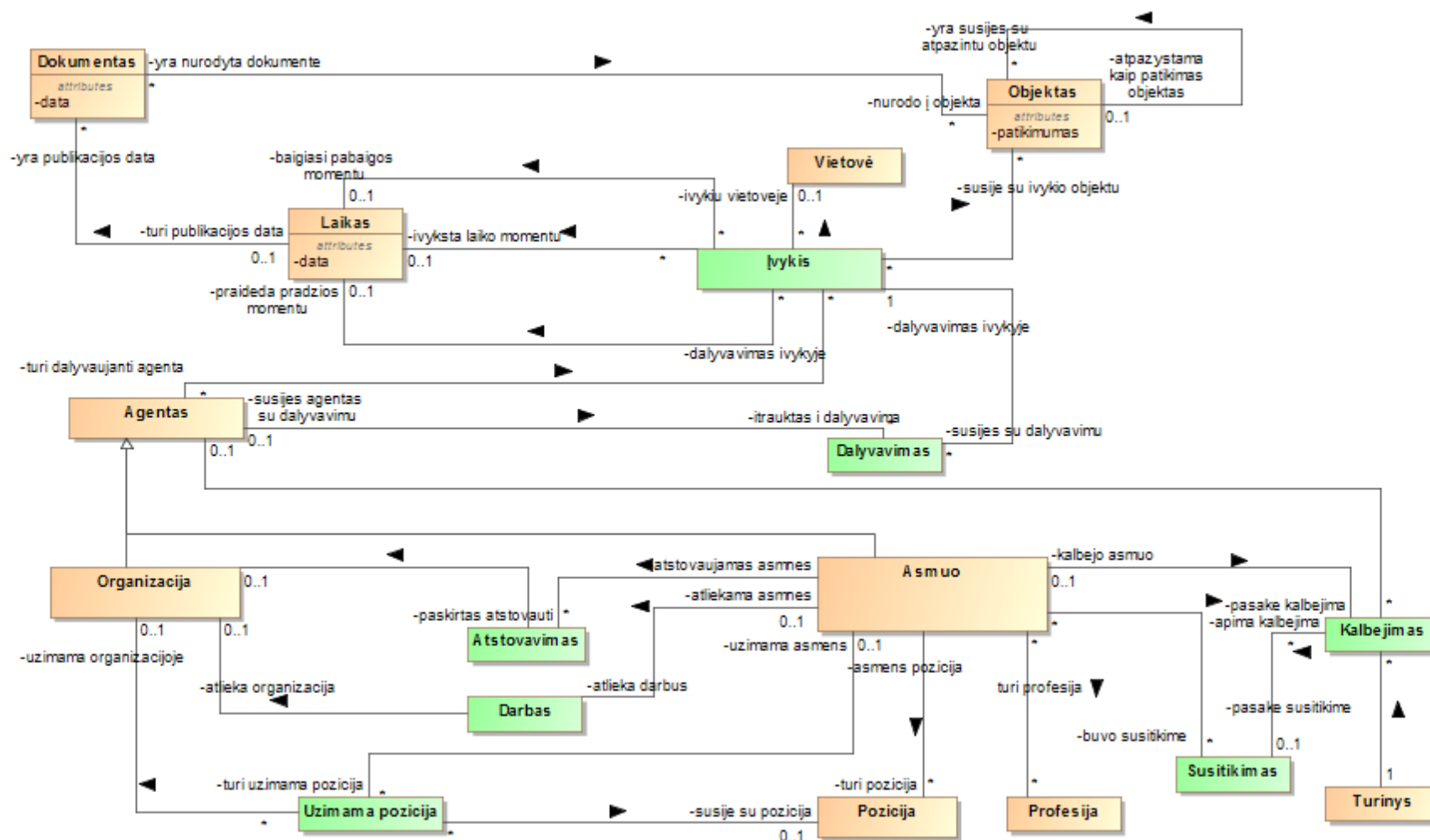
25 pav. yra pateikta ontologijos praturtintos individų atpažinimo tikslumo įverčiais modelio dalykinės srities esybių modelis. Jis sudarytas iš 6 esybių: rezultatų („Response“), paieškos srities („Category“), SPARQL užklausų („QuerySPARQL“), vartotojo („ApplicationUser“), užklauso („Request(test case)“).

Sistemos vartotojas suformuoja paieškai užklausą, priklausomai nuo poreikio tai gali būti daugiau nei viena užklausa. Įvykdžius užklausą gaunamas paieškos rezultatas. Formuojant paiešką joje gali būti viena paieškos sritis vienai pateiktai SPARQL užklausiai.



25 pav. Dalykinės srities esybių ryšių modelis

Baigiamajame darbe naudojama ontologija užpildyta individais, kurie sukurti automatiškai išanalizavus interneto portaluose publikuojamus straipsnius politikos ir ekonomikos tematika. Pagrindinis objektų tipas (klasė) yra dokumentas. Nagrinėjamas dokumentas gali būti apibudintas detalio esybių ryšių diagrama, pavaizduota 26 pav., matomi ryšiai ir kardinalumai tarp objektų. Diagramoje žalia spalva yra išskirta įvykių detalizuojančios esybės.

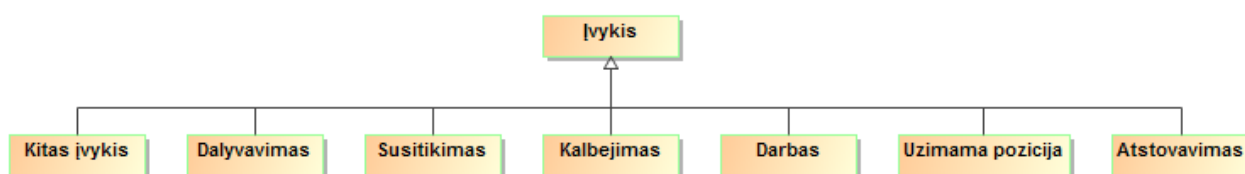


26 pav. Dokumento esybių ryšių modelis

2.15 lentelė. Dokumento tipo aprašas

Įvykio Nr.	Įvykio tipas	Pastaba
1.	Dokumentas	Dokumentas pagal sritį gali būti politinis ,ekonominis ir kt.;
2.	Objektas	Tai realiai egzistuojantis žmogus, pasaulis ir kt.;
3.	Turinys	Kalbėjimo metu išsakyta informacija;
4.	Organizacija	Žmonių kolektyvas suburtas tam tikrai veiklai atlikti;
5.	Asmuo	Subjektas , socialinis asmuo;

Kiekvienas įvykis gali būti kelių skirtingų tipų, kurie pateikiami 27 pav.



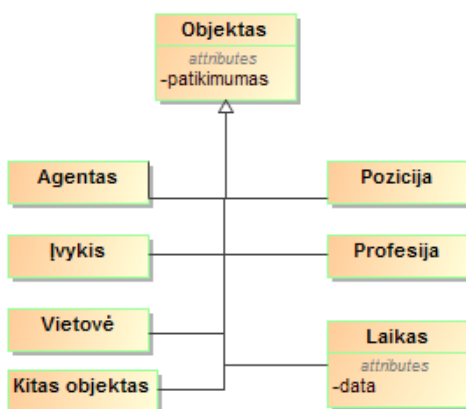
27 pav. Įvykio esybės kategorijos

Įvykis gali būti sudarytas iš sekančių tipų:

2.16 lentelė. Įvykių tipų aprašas

Įvykio Nr.	Įvykio tipas	Pastaba
1.	Atstovavimas	Asmuo atstovauja organizacijai, bendrijai, seimui ir kt.;
2.	Užimama pozicija	Asmuo užima poziciją susijusią su darbo veikla;
3.	Darbas	Asmuo dirba žurnalistu, seimo nariu, mechaniku ir kt.;
4.	Kalbėjimas	Asmuo kalbėjo susitikime, atliko viešą pranešimą renginyje ir kt.;
5.	Susitikimas	Asmuo dalyvavo susitikime pvz. NATO dalyviu susitikimas;
6.	Dalyvavimas	Asmuo dalyvavo susitikime, konferencijoje ar kt.;
7.	Kitas įvykis.	Asmuo buvo dalyviu kitame čia neįvardintame įvykyje;

Objektas taip pat gali būti skirstomas į keletą skirtingų tipų, kurie pateikti 28 pav.



28 pav. Objekto esybės kategorijos

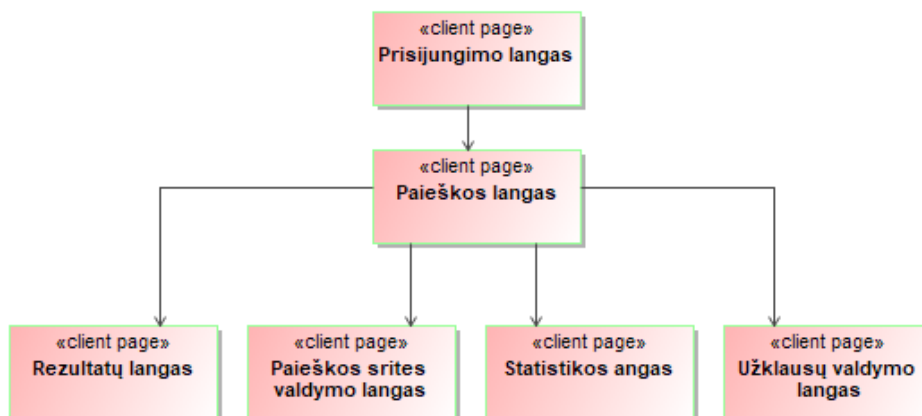
Objektas gali būti sudarytas iš tokių tipų:

2.17 lentelė. Objekto tipų aprašas

Objekto Nr.	Objekto tipas	Pastaba
1.	Agentas	Veikti galintis objektas;
2.	Įvykis	Objekto įvykiu gali būti kalbėjimas, darbas, susitikimas ir kt.;
3.	Vietovė	Vietovė yra šalis ,miestas , gatvė ir kt.;
4.	Pozicija	Objekto pozicija yra europarlamentaras, direktorius ir kt.;
5.	Profesija	Profesiją yra veikla, kaip inžinerija, medicina ir kt.;
6.	Laikas	Laikas yra konkrečiai įvardinta data, laiko tarpas ar kita laiką nusakanti išraiška.
7.	Kitas objektas	Objektas turi kitų čia neišvardintų objektų.

2.4. Naudotojų sąsajos modelis

Šiame skyriuje pateikiamas eksperimentinės sistemos navigacijos modelis (žr. 29 pav.). Sistemos pagrindinis vaizdas yra „Pagrindinis langas“ lange. Iš jo galima pasiekti paieškos vaizdą (vaizdą „Paieškos langas“), statistikos vaizdą (vaizdą „Statistikos langas“), paieškos srities valdymo vaizdą (vaizdą „Paieškos srities valdymo langą“), užklausų valdymo vaizdą (vaizdą „Užklausų valdymo langas“). Iš paieškos vaizdo galima pasiekti rezultatų vaizdą (vaizdą „Rezultatų langas“).

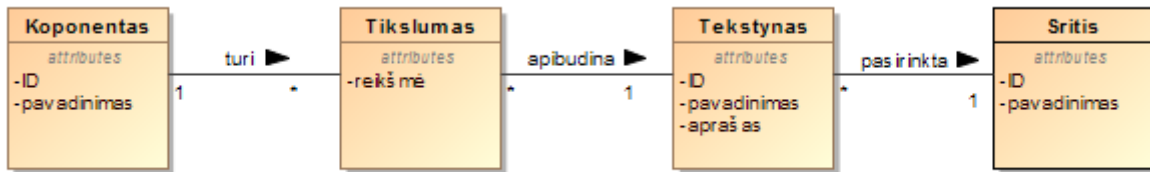


29 pav. Semantinės paieškos eksperimentinės sistemos navigacijos modelis

2.5. Formalus teksto anotavimo tikslumo įverčio skaičiavimo aprašymas

Atliekant semantinę paiešką aktualu žinoti kiek tikslus išvestas rezultatas. Rezultatas pirmiausiai priklauso nuo anotavimo grandinėės ir ją sudarančių komponentų tikslumo įverčių. Kiekvieno grandinėės komponento tikslumo įvertis priklauso nuo tekstyno, kuris buvo naudojamas suskaičiuoti tikslumo įvertį ir kitų anotavimo komponentų suformuotų rezultatų (pirmiau grandinėėje veikiančių anotavimo komponentų tikslumo). Jei sudarant konkrečią teksto anotavimo grandinėę būtų galimybė rinktis iš kelių to paties tipo anotavimo komponentų, tai atsirastų galimybė susikonstruoti tos srities tekstams anotuoti tiksliausiai veikiančią grandinėę. Bet tokiu atveju reikia žinoti individualų kiekvieno komponento veikimo tikslumą tam tikros srities tekstyno atžvilgiu. Individualus tikslumas reiškia, kad jis buvo nustatytas komponentui padavus tekstyno dokumentus su 100% tikslumo anotacijomis, kurios reikalingos vertinamo komponento veikimui. t.y. skaičiuojant individualų

komponento anotavimo tikslumą kitų komponentų veikimo tikslumui įtaka eliminuojama. Toks komponentų tikslumo paskaičiavimas leidžia pasirinkti tinkamus komponentus ir sudaryti anotavimo grandinę su kuo tikslesniu rezultatu galutiniame rezultate.



30 pav. Konceptinis modelis

31 pav. pateikiamas galimas tikslumo įverčio skaičiavimo modelis, kuris susideda iš trijų skirtingų komponentų. Kiekvienas komponentas turi savo tikslumo reikšmes, kurias galima pasirinkti, taip pat tikslumo reikšmę apibudina skirtingas tekstynas, kurį anotuojant buvo gauta tikslumo reikšmė. Tikslumo reikšmės ribos nuo 0 iki 1. Teksto anotavimo komponentų grandinės tikslumas anotuojant tam tikros srities tekstyną parodo, kokių tikslumų anotavimo rezultato galima tikėtis iš grandinės gale esančio komponento. Šį tikslumą galima suskaičiuoti pagal tokią formulę:

$$T_m = \prod_{k=1}^n T_{mk} , \quad (5)$$

čia T_{mk} – k-ojo komponento anotavimo grandinėje individualus tikslumas m-ojo tekstyno atžvilgiu,

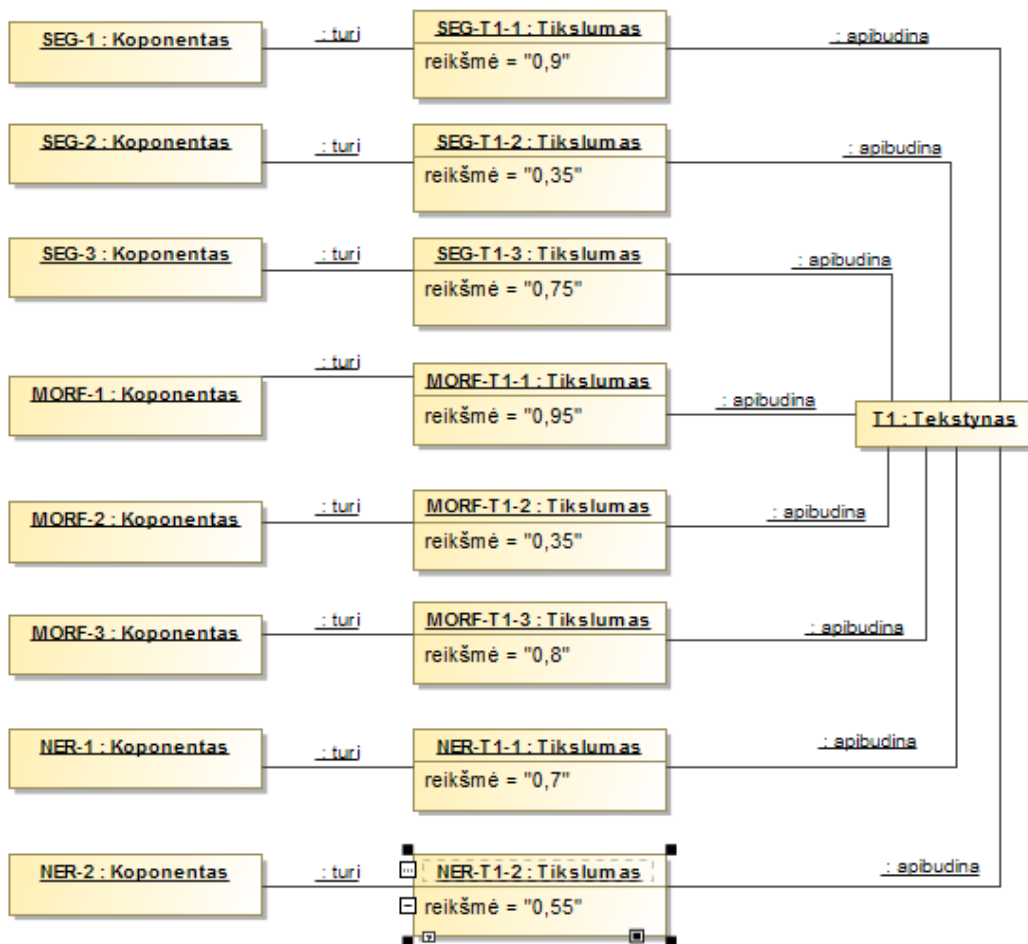
n- komponentų kiekis anotavimo grandinėje.

Pasirinkimas tekstynui T1: Jei norime sužinoti, kokio anotavimo tikslumo galime tikėtis iš anotavimo grandinės, sudarytos iš komponentų SEG1-T1-1, MORF1-T1-1 ir NER1-T1-2, anotodami tekstus, panašius į tekstyno T1 tekstus, tai išsirenkame tekstyno komponentų veikimo tikslumą, nustatytą anotuojant tekstyną T1 ir duotam pavyzdžiui gauname:

$$T1 = SEG1 - T1 - 1 * MORF1 - T1 - 1 * NER1 - T1 - 2,$$

$$T1 = 0,9 * 0,95 * 0,55,$$

$$T1 = 0.47025,$$



31 pav. Tikslumo įverčio skaičiavimo modelis T1 tekstynui

Žemiau esanti 2.18 lentelė parodo T1 tekstyno skaičiavimo rezultatus.

2.18 lentelė. Skaičiavimo rezultatai T1 tekstynui

Anotavimo grandinės variantas	Anotavimo grandinės komponentų tikslumo įverčiai	Anotavimo grandinės tikslumo įvertis
1.	SEG1-T1-1 = 0.9 MORF1-T1-1 = 0.95 NER1-T1-2 = 0.55	$T1 = 0.47025$
2.	SEG1-T1-1 = 0.9 MORF1-T1-1 = 0.95 NER1-T1-1 = 0.7	$T1 = 0.5985$
3.	SEG1-T1-1 = 0.9 MORF1-T1-2 = 0.35 NER1-T1-1 = 0.7	$T1 = 0.2205$
4.	SEG1-T1-1 = 0.9 MORF1-T1-3 = 0.95 NER1-T1-1 = 0.7	$T1 = 0.5985$
5.	SEG1-T1-1 = 0.9 MORF1-T1-2 = 0.35 NER1-T1-2 = 0.55	$T1 = 0.1732$
6.	SEG1-T1-1 = 0.9 MORF1-T1-3 = 0.8 NER1-T1-2 = 0.55	$T1 = 0.396$

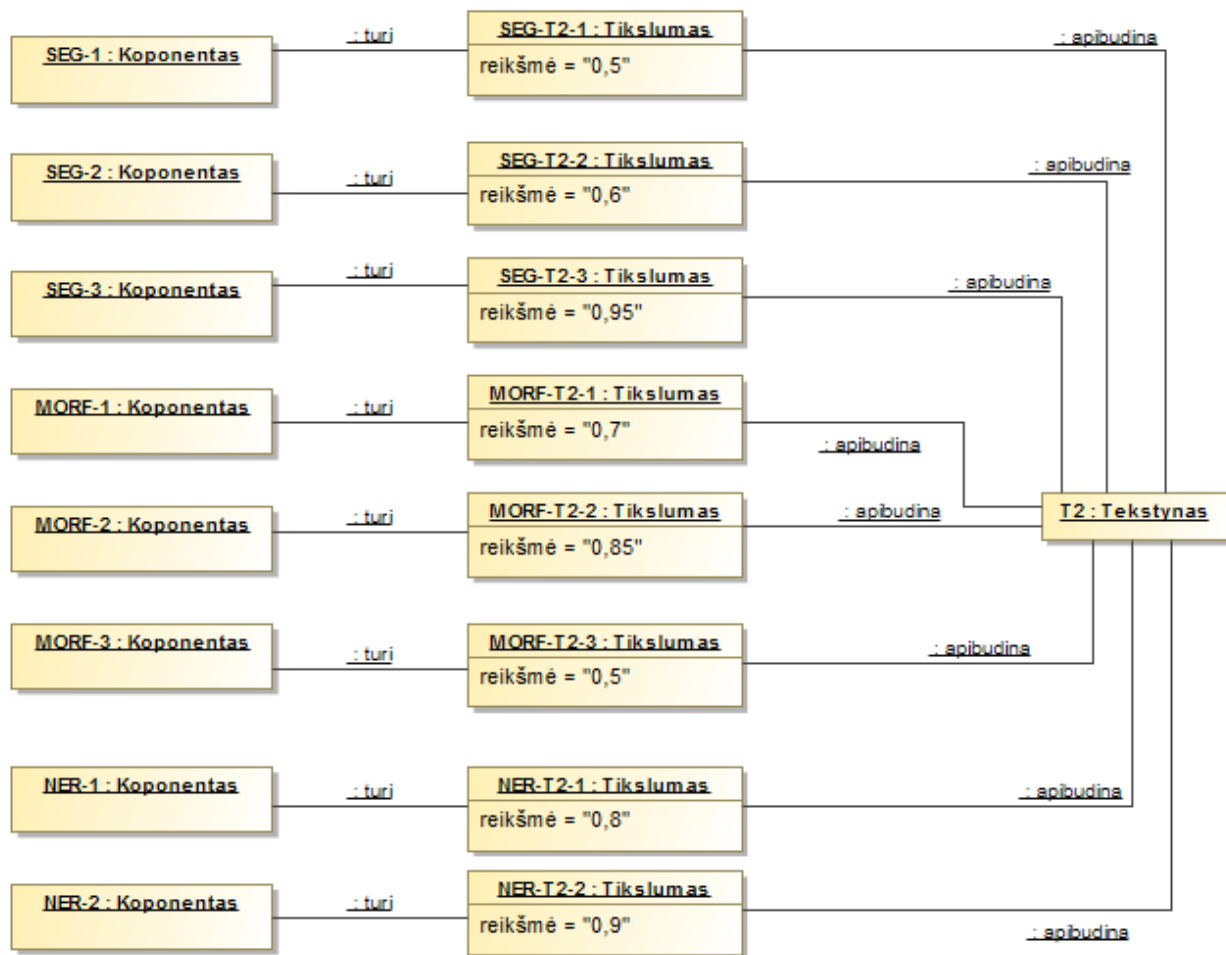
7.	SEG1-T1-2 = 0.35 MORF1-T1-1 = 0.95 NER1-T1-1 = 0.7	$T1 = 0.2327$
8.	SEG1-T1-2 = 0.35 MORF1-T1-2 = 0.35 NER1-T1-2 = 0.55	$T1 = 0.0673$
9.	SEG1-T1-2 = 0.35 MORF1-T1-3 = 0.8 NER1-T1-1 = 0.7	$T1 = 0.196$
10.	SEG1-T1-2 = 0.35 MORF1-T1-2 = 0.35 NER1-T1-2 = 0.55	$T1 = 0.0673$
11.	SEG1-T1-2 = 0.35 MORF1-T1-3 = 0.8 NER1-T1-2 = 0.55	$T1 = 0.154$
12.	SEG1-T1-2 = 0.35 MORF1-T1-1 = 0.95 NER1-T1-2 = 0.55	$T1 = 0.1828$
13.	SEG1-T1-3 = 0.75 MORF1-T1-1 = 0.95 NER1-T1-1 = 0.7	$T1 = 0.4987$
14.	SEG1-T1-3 = 0.75 MORF1-T1-2 = 0.35 NER1-T1-1 = 0.7	$T1 = 0.1837$
15.	SEG1-T1-3 = 0.75 MORF1-T1-3 = 0.8 NER1-T1-1 = 0.7	$T1 = 0.42$
16.	SEG1-T1-3 = 0.75 MORF1-T1-2 = 0.35 NER1-T1-2 = 0.55	$T1 = 0.1443$
17.	SEG1-T1-3 = 0.75 MORF1-T1-3 = 0.8 NER1-T1-2 = 0.55	$T1 = 0.33$
18.	SEG1-T1-3 = 0.75 MORF1-T1-1 = 0.95 NER1-T1-2 = 0.55	$T1 = 0.3918$

Pasirinkimas tekstynui T2: Jei norime sužinoti, kokio anotavimo tikslumo galime tikėtis iš anotavimo grandinės, sudarytos iš komponentų SEG1-T2-1, MORF1-T2-1 ir NER1-T2-2, anotodami tekstus, panašius į tekstyno T2 tekstus, tai išsirenkame tekstyno komponentų veikimo tikslumą, nustatytą anotuojant tekstyną T2 ir duotam pavyzdžiui gauname:

$$T2 = SEG1 - T2 - 1 * MORF1 - T2 - 1 * NER1 - T2 - 2,$$

$$T2 = 0,5 * 0,7 * 0,9,$$

$$T2 = 0.315,$$



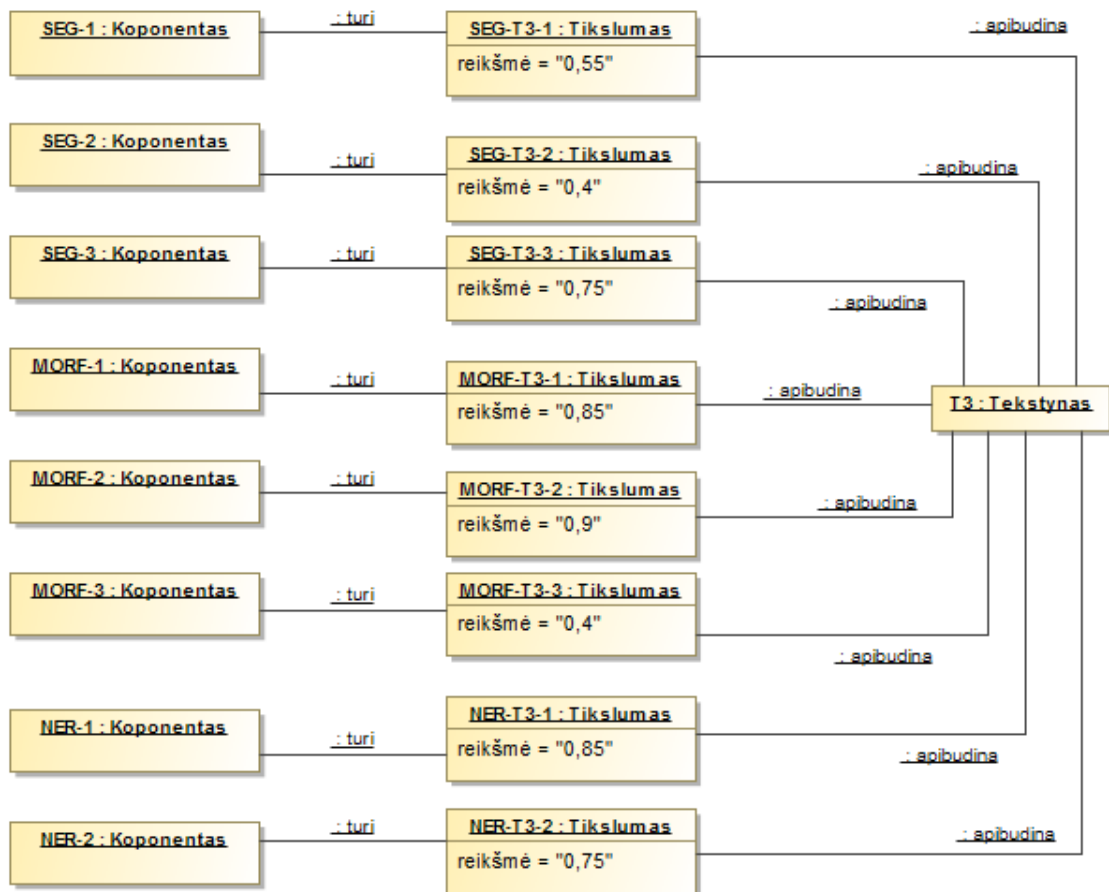
32 pav. Tikslumo įvertio skaičiavimo modelis T2 tekstynui

Žemiau esanti 2.19 lentelė, parodo T1 tekstinio skaičiavimo rezultatus.

2.19 lentelė. Skaičiavimo rezultatai T2 tekstynui

Anotavimo grandinės variantas	Anotavimo grandinės komponentų tikslumo įverčiai	Anotavimo grandinės tikslumo įvertis
1.	SEG1-T2-1 = 0.5 MORF1-T2-1 = 0.7 NER1-T2-2 = 0.9	$T2 = 0,315$
2.	SEG1-T2-1 = 0.5 MORF1-T2-1 = 0.7 NER1-T2-1 = 0.8	$T2 = 0.28$
3.	SEG1-T2-1 = 0.5 MORF1-T2-2 = 0.85 NER1-T2-1 = 0.8	$T2 = 0.34$
4.	SEG1-T2-1 = 0.5 MORF1-T1-3 = 0.5 NER1-T2-1 = 0.8	$T2 = 0.2$
5.	SEG1-T2-1 = 0.5 MORF1-T2-2 = 0.85 NER1-T2-2 = 0.9	$T2 = 0.3825$
6.	SEG1-T2-1 = 0.5	$T2 = 0.225$

	MORF1-T2-3 = 0.5 NER1-T2-2 = 0.9	
7.	SEG1-T2-2 = 0.6 MORF1-T2-1 = 0.7 NER1-T2-1 = 0.8	$T2 = 0.336$
8.	SEG1-T2-2 = 0.6 MORF1-T2-2 = 0.85 NER1-T2-2 = 0.9	$T2 = 0.459$
9.	SEG1-T2-2 = 0.6 MORF1-T2-3 = 0.5 NER1-T2-1 = 0.8	$T2 = 0.24$
10.	SEG1-T2-2 = 0.6 MORF1-T2-2 = 0.85 NER1-T2-2 = 0.9	$T2 = 0.459$
11.	SEG1-T1-2 = 0.6 MORF1-T2-3 = 0.5 NER1-T2-2 = 0.9	$T2 = 0.27$
12.	SEG1-T2-2 = 0.6 MORF1-T2-1 = 0.7 NER1-T2-2 = 0.9	$T2 = 0.378$
13.	SEG1-T2-3 = 0.95 MORF1-T2-1 = 0.7 NER1-T2-1 = 0.8	$T2 = 0.532$
14.	SEG1-T2-3 = 0.95 MORF1-T2-2 = 0.85 NER1-T2-1 = 0.8	$T2 = 0.646$
15.	SEG1-T2-3 = 0.95 MORF1-T2-3 = 0.5 NER1-T2-1 = 0.8	$T2 = 0.38$
16.	SEG1-T2-3 = 0.95 MORF1-T2-2 = 0.85 NER1-T2-2 = 0.9	$T2 = 0.7267$
17.	SEG1-T2-3 = 0.95 MORF1-T2-3 = 0.5 NER1-T2-2 = 0.9	$T2 = 0.4275$
18.	SEG1-T2-3 = 0.95 MORF1-T2-1 = 0.7 NER1-T2-2 = 0.9	$T2 = 0.5985$



33 pav. Tikslumo įverčio skaičiavimo modelis T3 tekstynui

Pasirinkimas tekstynui T3: Jei norime sužinoti, kokio anotavimo tikslumo galime tikėtis iš anotavimo grandinėls, sudarytos iš komponentų SEG1-T2-3, MORF1-T3-3 ir NER1-T3-1, anotuodami tekstus, panašius į tekstyno T3 tekstus, tai išsirenkame tekstyno komponentų veikimo tikslumą, nustatytą anotuojant tekstyną T3 ir duotam pavyzdžiui gauname:

$$T3 = SEG1 - T3 - 3 * MORF1 - T3 - 3 * NER1 - T3 - 1,$$

$$T3 = 0,75 * 0,4 * 0,85$$

$$T3 = 0.255$$

2.20 lentelė. Skaičiavimo rezultatai T3 tekstynui

Anotavimo grandinės variantas	Anotavimo grandinėls komponentų tikslumo įverčiai	Anotavimo grandinėls tikslumo įvertis
1.	SEG1-T3-1 = 0.55 MORF1-T3-1 = 0.85 NER1-T3-2 = 0.75	$T3 = 0.3506$
2.	SEG1-T3-1 = 0.55 MORF1-T3-1 = 0.85 NER1-T3-1 = 0.85	$T3 = 0.3973$
3.	SEG1-T3-1 = 0.55 MORF1-T3-2 = 0.9 NER1-T3-1 = 0.85	$T3 = 0.4207$
4.	SEG1-T3-1 = 0.55	$T3 = 0.187$

	MORF1-T3-3 = 0.4 NER1-T3-1 = 0.85	
5.	SEG1-T3-1 = 0.55 MORF1-T3-2 = 0.9 NER1-T3-2 = 0.75	$T3 = 0.3712$
6.	SEG1-T3-1 = 0.55 MORF1-T3-3 = 0.4 NER1-T3-2 = 0.75	$T3 = 0.165$
7.	SEG1-T3-2 = 0.4 MORF1-T3-1 = 0.85 NER1-T3-1 = 0.85	$T3 = 0.289$
8.	SEG1-T3-2 = 0.4 MORF1-T3-2 = 0.9 NER1-T3-2 = 0.75	$T3 = 0.27$
9.	SEG1-T3-2 = 0.4 MORF1-T3-3 = 0.4 NER1-T3-1 = 0.85	$T3 = 0.136$
10.	SEG1-T3-2 = 0.4 MORF1-T3-2 = 0.9 NER1-T3-2 = 0.75	$T3 = 0.27$
11.	SEG1-T3-2 = 0.4 MORF1-T3-3 = 0.4 NER1-T3-2 = 0.75	$T3 = 0.12$
12.	SEG1-T3-2 = 0.4 MORF1-T3-1 = 0.85 NER1-T3-2 = 0.75	$T3 = 0.255$
13.	SEG1-T3-3 = 0.75 MORF1-T3-1 = 0.85 NER1-T3-1 = 0.85	$T3 = 0.5418$
14.	SEG1-T3-3 = 0.75 MORF1-T3-2 = 0.9 NER1-T3-1 = 0.85	$T3 = 0.5737$
15.	SEG1-T3-3 = 0.75 MORF1-T3-3 = 0.4 NER1-T1-1 = 0.85	$T3 = 0.255$
16.	SEG1-T3-3 = 0.75 MORF1-T3-2 = 0.9 NER1-T3-2 = 0.75	$T3 = 0.5062$
17.	SEG1-T3-3 = 0.75 MORF1-T3-3 = 0.4 NER1-T3-2 = 0.75	$T3 = 0.225$
18.	SEG1-T3-3 = 0.75 MORF1-T3-1 = 0.85 NER1-T3-2 = 0.75	$T3 = 0.4781$

Apibendrinant skaičiavimų rezultatus matome, kad jeigu turėtume tokią meta informaciją apie komponentus, galėtume susikonstruoti tiksliausią vartotojo poreikius tenkinančią tekstų anotavimo grandinę.

2.6. Reikalavimų apibendrinimas

Atlikus reikalavimų analizę šio darbo tikslo ribose, bus kuriama „Semantinės paieškos IS“, kuri turi turėti galimybę parodyti paieškos rezultato įverčio tikslumą ir turėti galimybę pateikti vartotojui tikslumo įverčio reikšmę (19 pav.).

Dokumentas turintis tikslumo įvertį bus naudojamas paieškose ir naudotojas galės matyti paieškos dokumento tikslumą.

Naudojant eksperimentinį modelį padaryta išvada, kad jeigu turėtume informaciją apie grandinėle sudarančio kiekvieno komponento tikslumą, galėtume sudaryti tiksliausią vartotojo poreikius tenkinančią tekstų anotavimo grandinėle.

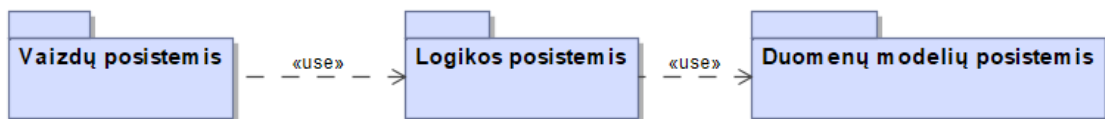
3. ONTOLOGIJOS PRATURTINTOS INDIVIDŲ ATPAŽINIMO TIKSLUMO ĮVERČIAIS MODELIO REALIZACIJOS PROJEKTAS

Ontologijos praturtintos individų atpažinimo tikslo įverčiais modelio realizacija susideda iš šių dalių:

- sistemos loginė architektūra;
- sistemos elgsenos modelis;
- duomenų bazės modelis;
- detalus projektas;
- diegimo modelis.

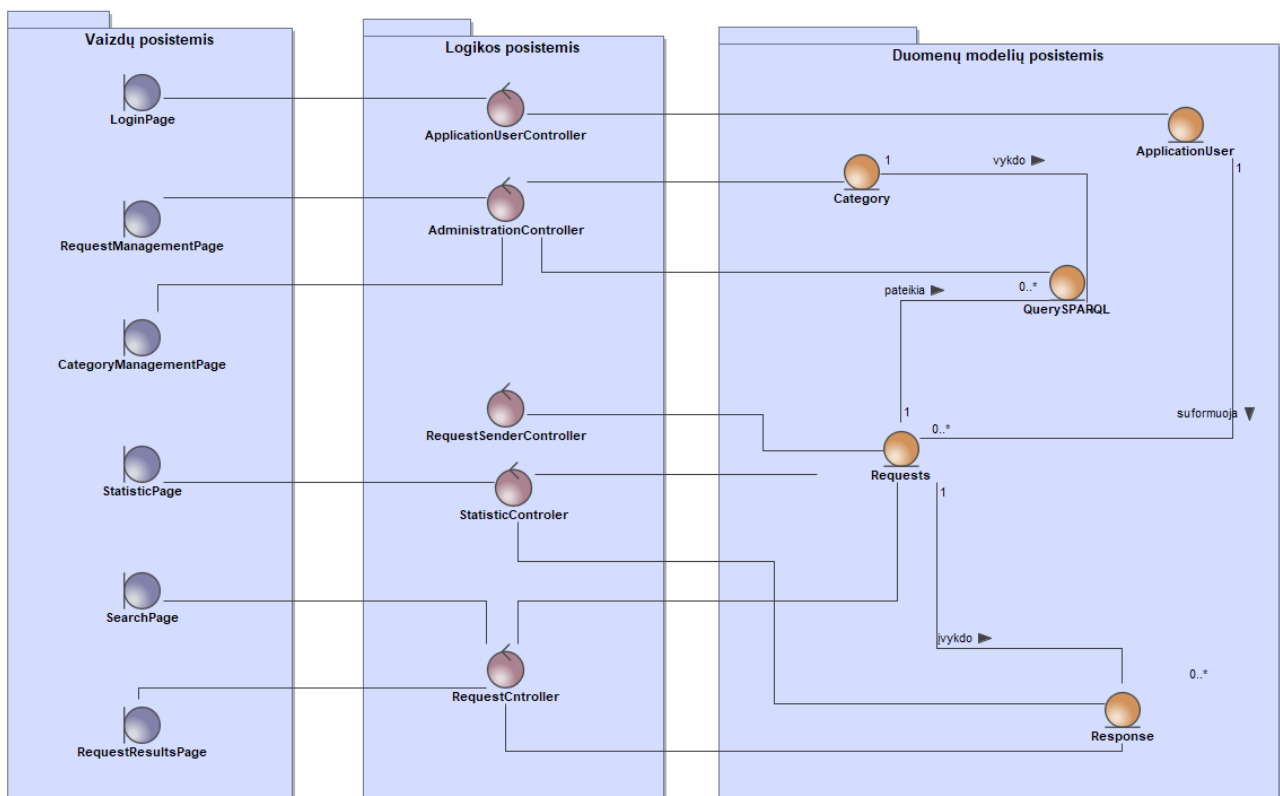
3.1. Sistemos loginė architektūra

Sistemos loginė architektūra vaizduoja tarpusavio ryšį tarp eksperimentinės sistemos posistemų (žr. 34 pav.). Vaizdų posistemyje yra vaizdų valdymo elementai. Logikos posistemyje yra logikos valdymo elementai. Duomenų modelių posistemyje yra esybės ir ryšiai siejantys konkrečias klases.



34 pav. Sistemos loginė architektūra

Žemiau esančiame paveikslėlyje 35 pav. pavaizduota posistemų sudėtis bei jų tarpusavio ryšiai tarp posistemų elementų.



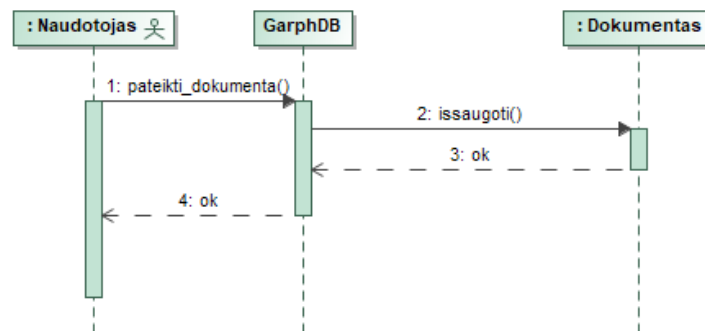
35 pav. Sistemos loginės architektūros detalizavimas

Vaizdų posistemis yra atsakingas už vaizdų elgseną. Logikos posistemio valdikliai yra atsakingi už loginius veiksmus. Logikos valdikliai komunikuoja su duomenų bazės valdymo sistema per duomenų modulių posistemį. Šis posistemis neturi tiesioginio ryšio su vaizdų elementais.

3.2. Sistemos elgsenos modelis

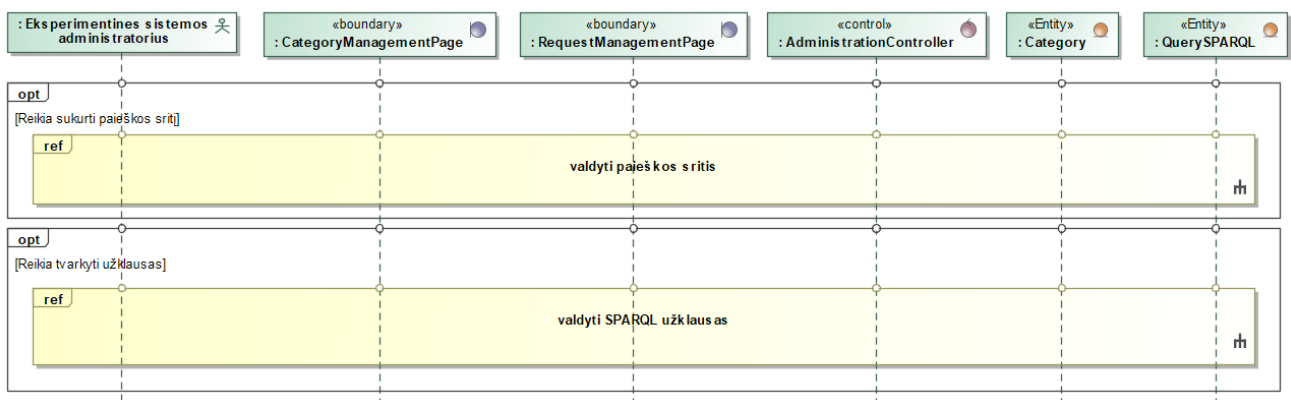
Šiame skyriuje panaudojimo atvejams pateikiamos projekto lygio sekų diagramos (žr. 36 pav. - 44 pav.)

36 pav. pavaizduota diagrama vaizduoja panaudojimo atvejį „Įkelti dokumentą“. Dokumento įkėlimas į *GarphDB* vyksta naudotojui pasirinkus norimą įkelti dokumentą, dokumentas yra pateikiamas sistemai ir išsaugomas.



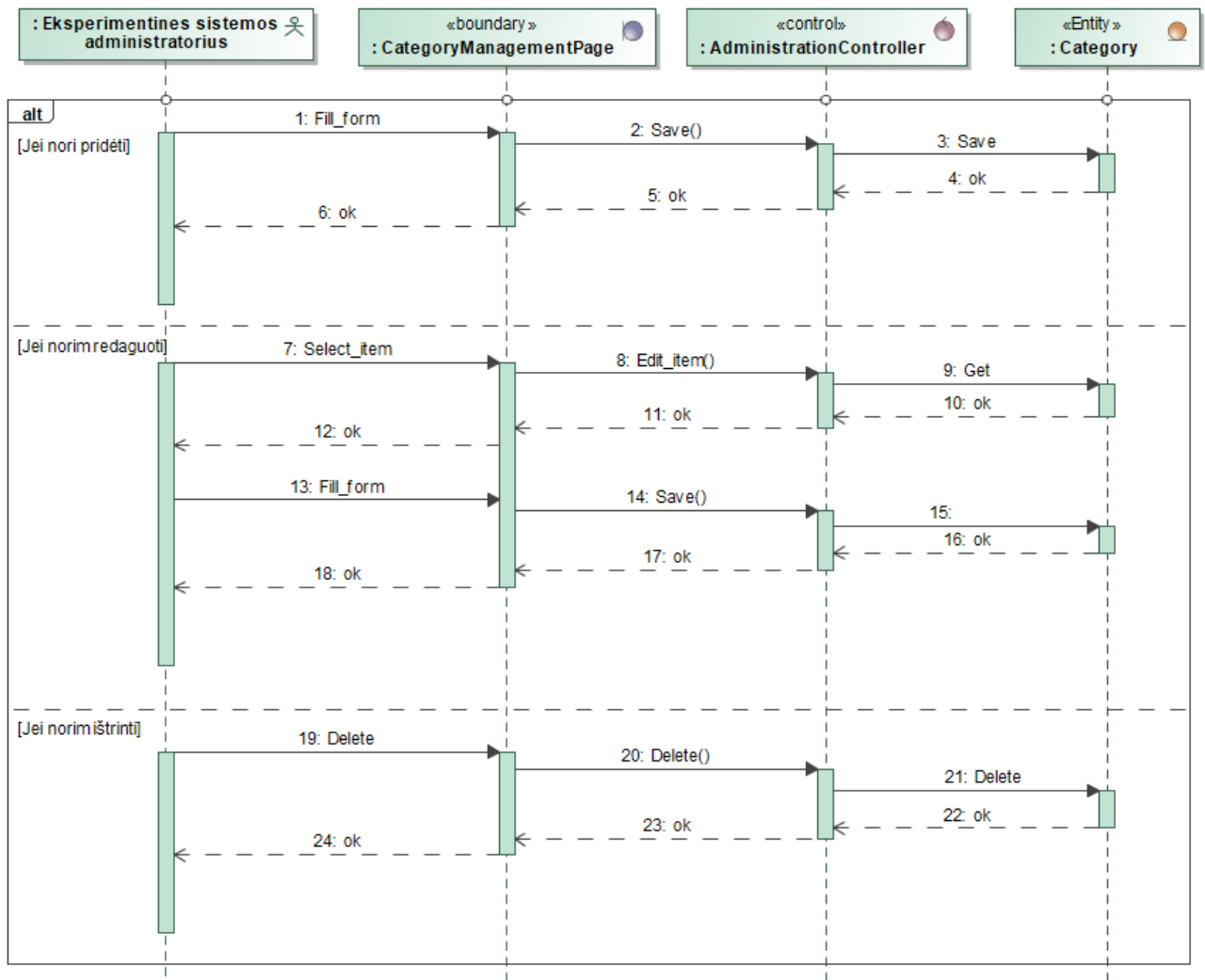
36 pav. PA1 „Įkelti dokumentą“ sistemos modelio sekų diagrama

37 pav. pavaizduota diagrama vaizduoja panaudojimo atvejį „Valdyti užklausas“. Sekų diagrama susideda iš dviejų pasirinkimų: „Reikia sukurti paieškos sritį“ ir „Reikia tvarkyti užklausas“. Detaliau pasirinkimo atvejai parašyti atitinkamai 38 pav. ir 39 pav.



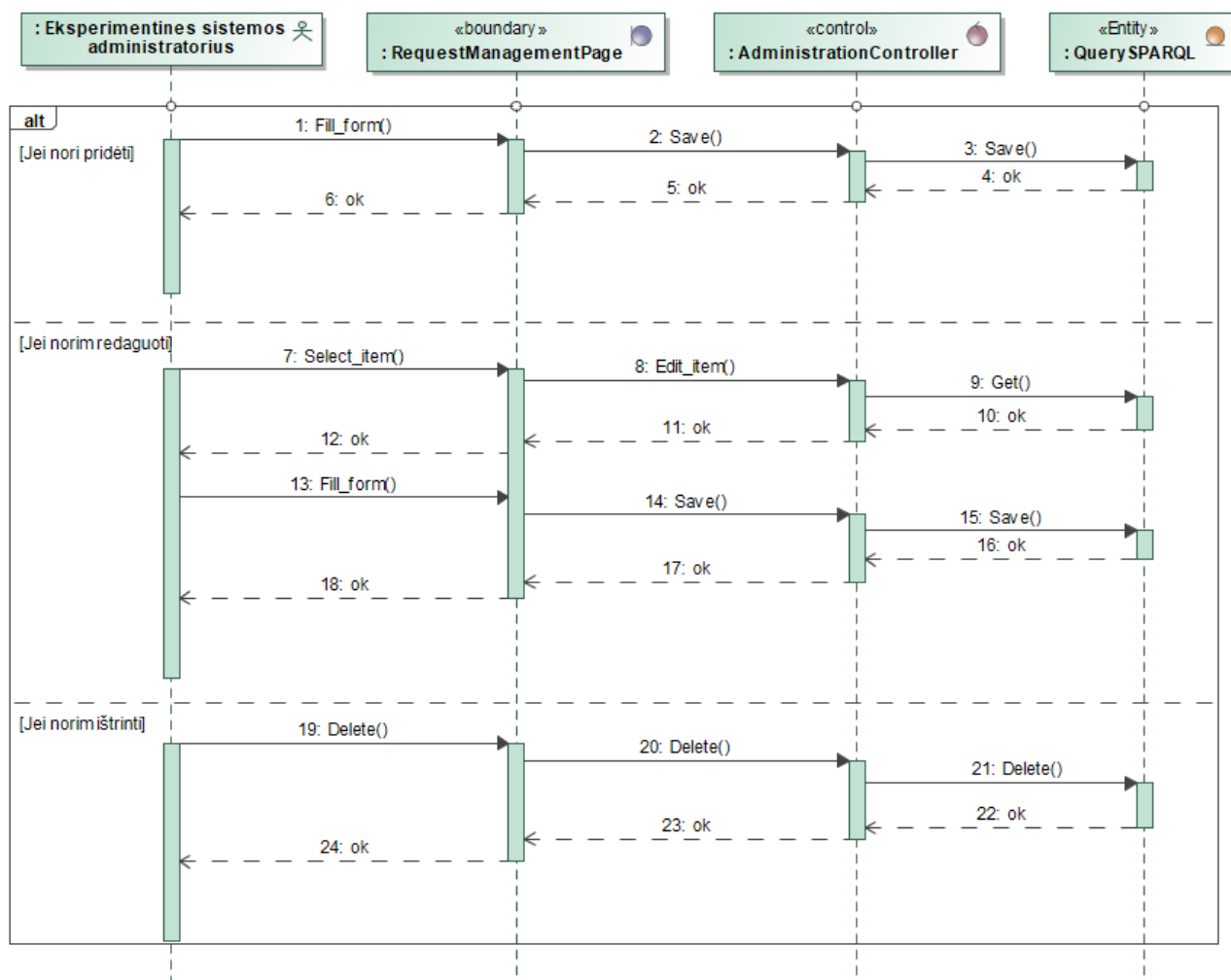
37 pav. PA2 „Valdyti užklausas“ sistemos modelio sekų diagrama

38 pav. pavaizduota diagrama vaizduoja panaudojimo atvejo „Valdyti užklausas“ detalizuotą alternatyvą „Valdyti paieškos sritis“. Diagramoje matome tris scenarijus: vartotojas gali pridėti, redaguoti ir ištrinti paieškos sritį.



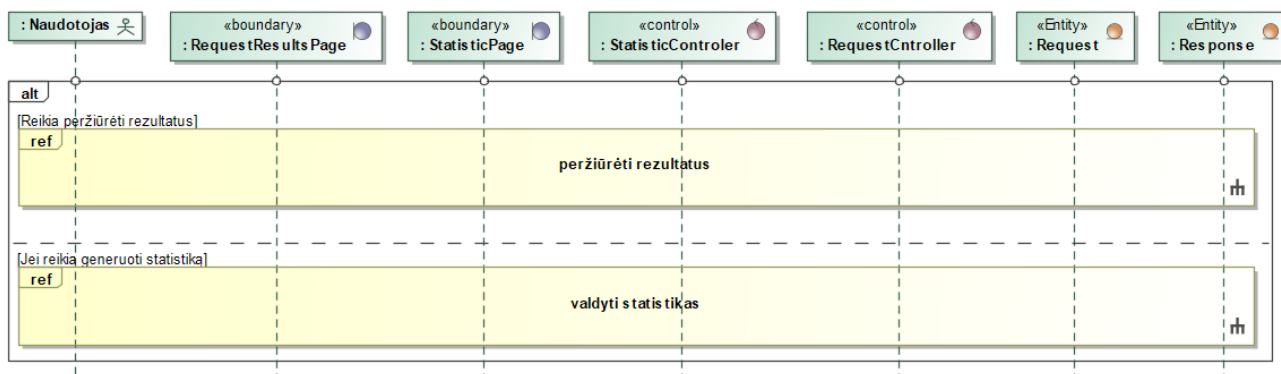
38 pav. PA2 detalizuota alternatyva „Valdyti apieškos sritis“

39 pav. pavaizduota diagrama vaizduoja panaudojimo atvejo „Valdyti užklausas“ detalizuotą alternatyva „Valdyti SPARQL užklausas“. Diagramoje matome tris scenarijus: vartotojas gali pridėti, redaguoti ir ištrinti SPARQL užklausa.



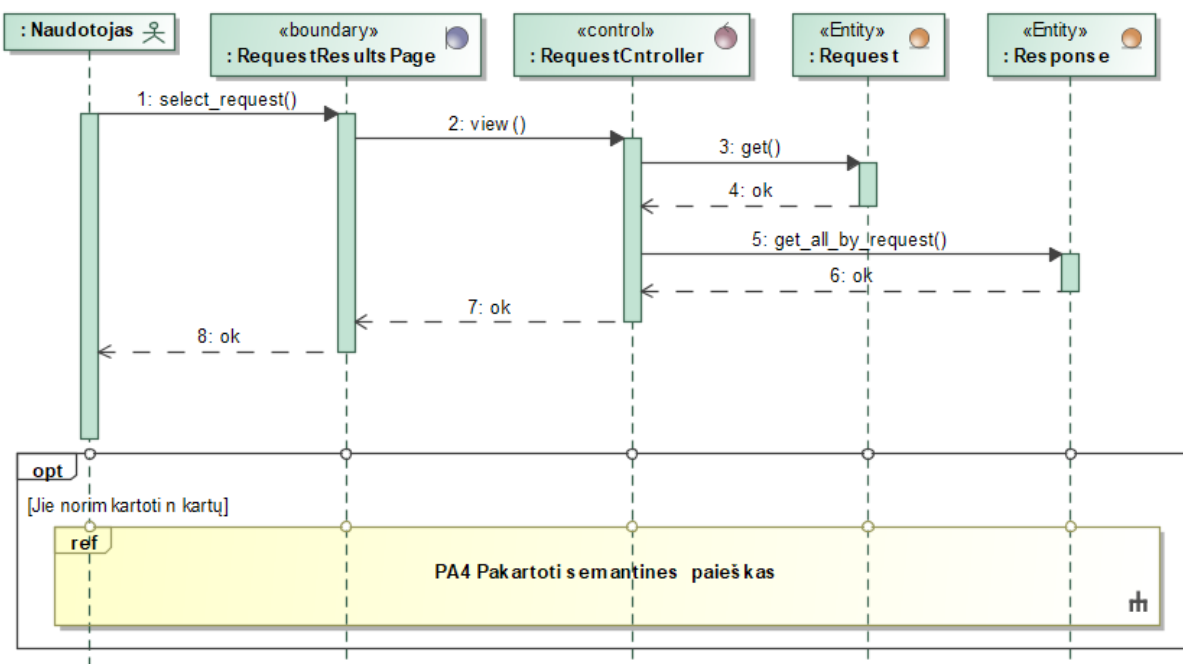
39 pav. PA2 detalizuota alternatyva „Valdyti SPARQL užklausas“

40 pav. pavaizduota diagrama vaizduoja panaudojimo atvejį „Peržiūrėti užklauso ir statistikos rezultatus“. Sekų diagrama susideda iš dviejų pasirinkimų: „Peržiūrėti rezultatus“ ir „Valdyti statistikas“. Detaliau pasirinkimo atvejai parašyti atitinkamai 41 pav. ir 42 pav.



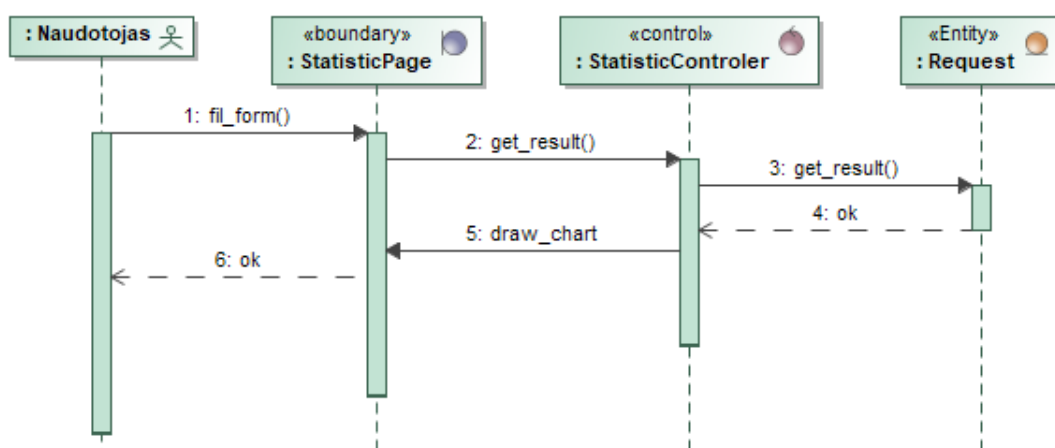
40 pav. PA3 „Peržiūrėti užklauso ir statistikos rezultatus“ sistemos modelio sekų diagrama

41 pav. pavaizduota diagrama vaizduoja panaudojimo atvejo „Peržiūrėti užklauso ir statistikos rezultatus“ detalizuotą alternatyvą „Peržiūrėti rezultatus“. Naudotojas peržiūrėjęs rezultatus juos gali pakartoti pasirinktą n kartų, detaliau panaudojimo atvejį „Pakartoti semantines paieškas“ žiūrėti 42 pav.



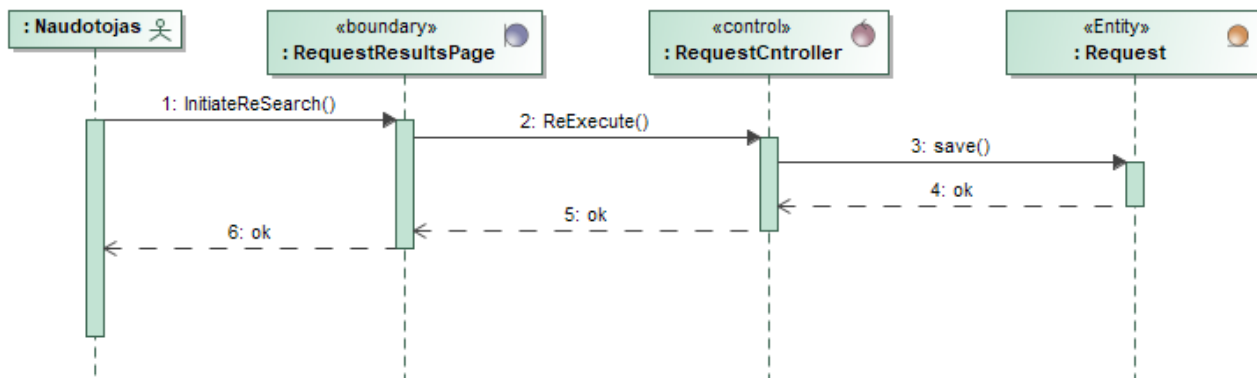
41 pav. PA3 detalizuota alternatyva „Peržiūrėti rezultatus“

42 pav. pavaizduota diagrama vaizduoja panaudojimo atvejo „Peržiūrėti užklauso ir statistikos rezultatus“ detalizuotą alternatyva „Valdyti statistikas“. Naudotojas statistikų lange pateikęs formą gali peržiūrėti vykdytų užklauso statistikos diagramas.



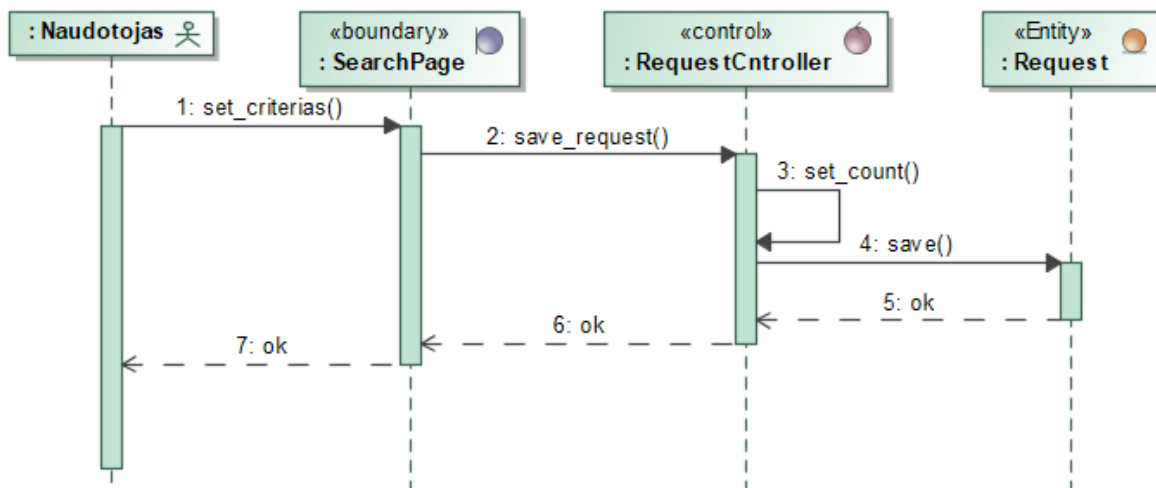
42 pav. PA3 detalizuota alternatyva „Valdyti statistikas“

43 pav. pavaizduota diagrama vaizduoja panaudojimo atvejį „Pakartoti semantinę paiešką“. Naudotojas turi poreikį kartoti jau vykdytą paiešką. Jis inicijuoja pakartotinę paiešką. Sistema išsaugo paieškos rezultatus.



43 pav. PA4 „Pakartoti semantines paieškas“ sistemos modelio sekų diagrama

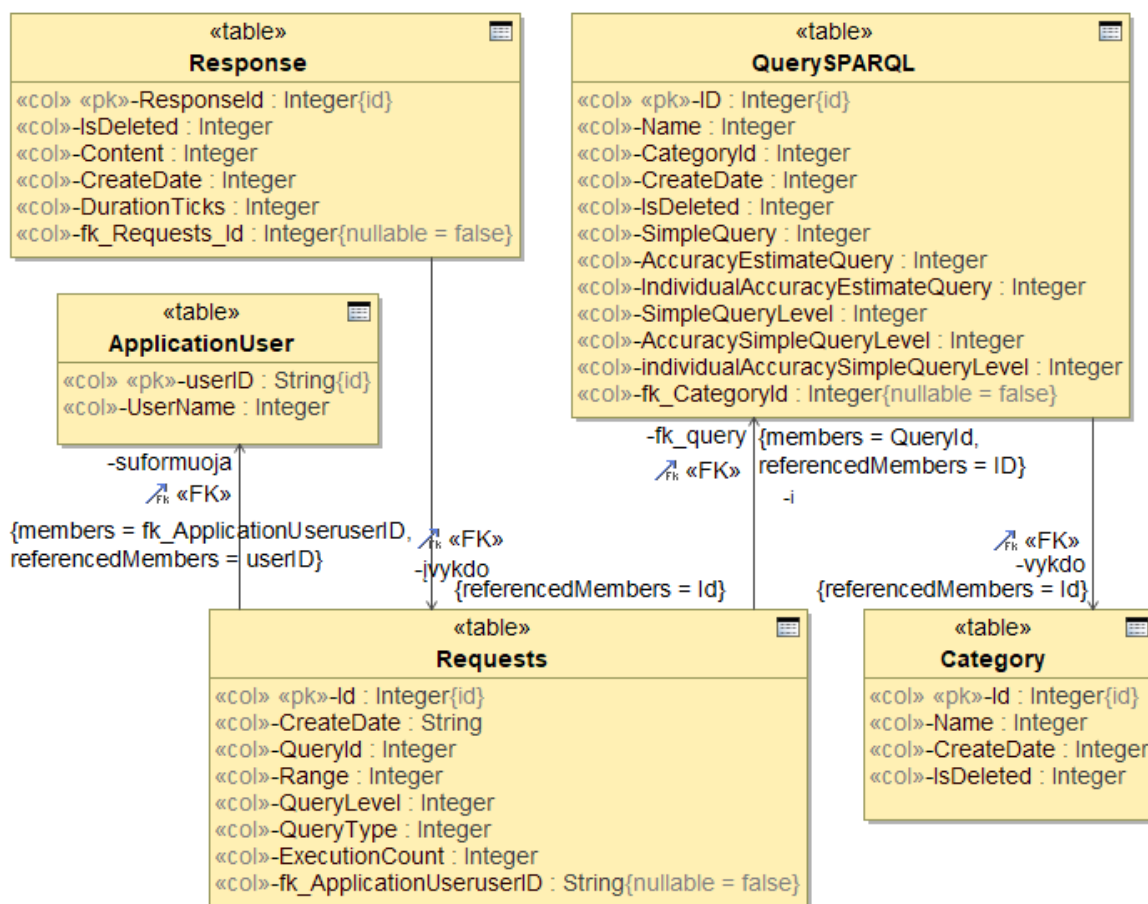
44 pav. pavaizduota diagrama vaizduoja panaudojimo atvejį „Inicijuoti semantinę paiešką“. Naudotojas paieškos lange užpildęs paieškos kriterijus gali inicijuoti užklausą.



44 pav. PA5 „Inicijuoti semantinę paiešką“ sistemos modelio sekų diagrama

3.3. Duomenų bazės modelis

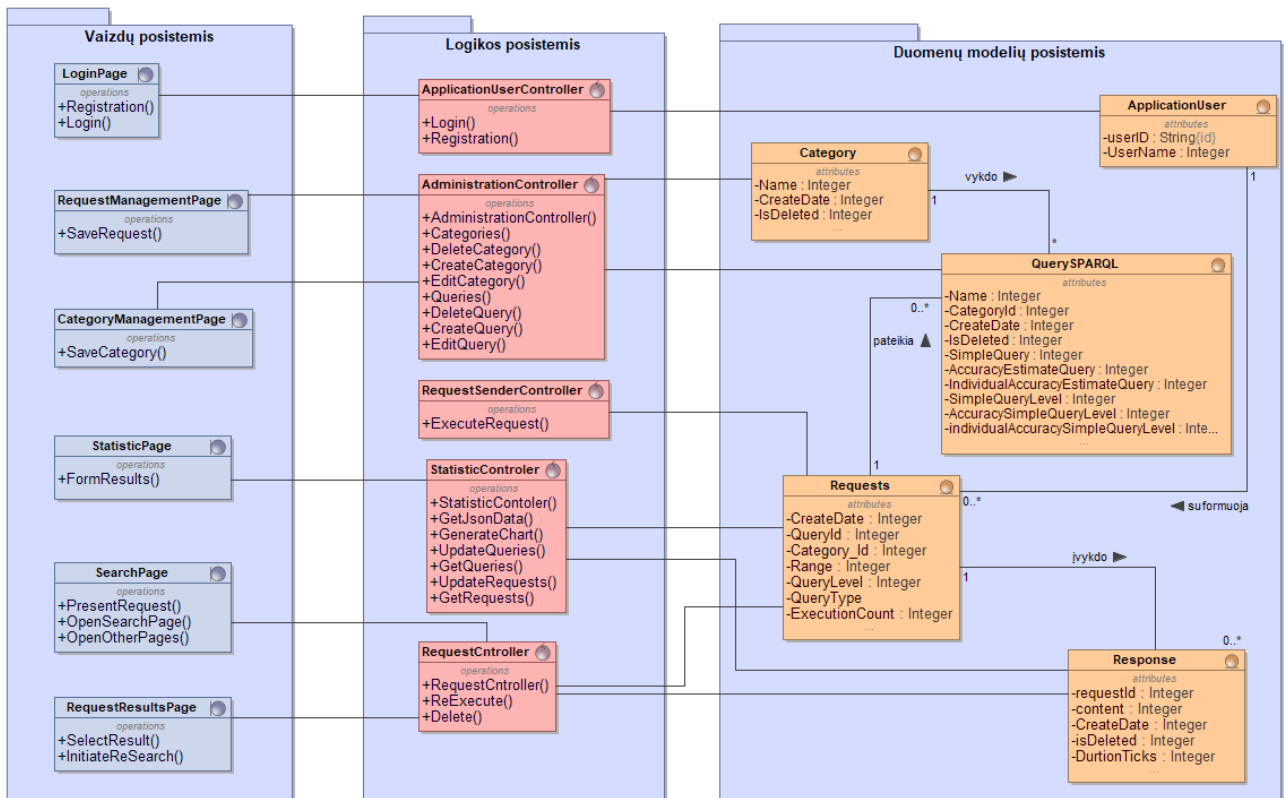
Šiame skyriuje pateikta loginė duomenų bazės schema (žr. 45 pav.). Schema yra sugeneruota pagal esybių klasių modelį (žr. 25 pav.) naudojant duomenų bazės transformacijas. Loginėje duomenų bazės schemoje pateikta reikalingos duomenų bazės lentelės su savo atributais ir ryšiais tarp duomenų bazės lentelių. Fizinė duomenų bazės schemą ir detalūs tos duomenų bazės schemos lentelių aprašymai pateikiami 8.5 priede. Fizinėje duomenų bazės schemoje pateiktos toje schemoje esančios lentelės su naudojamų laukų tipais ir kardinalumais tarp tų lentelių.



45 pav. Loginė duomenų bazės schema

3.4. Detalus projektas

Šiame skyriuje pateikta detalus eksperimentinės sistemos posistemių klasių modelis (46 pav.). Vaizdų posistemį sudaro 6 vaizdai, prisijungimo lango vaizdas (vaizdas „LoginPage“), užklausų valdymo lango vaizdas (vaizdas „RequestManagementPage“), paieškos sričių valdymo lango vaizdas (vaizdas „CategoryManagementPage“), statistikos lango vaizdas (vaizdas „StatisticPage“), paieškos lango vaizdas (vaizdas „SearchPage“), užklausų rezultatų lango vaizdas (vaizdas „RequestResultsPage“). Šie vaizdai kreipiasi/pasiekiami per logikos posistemį. Logikos posistemyje „ApplicationUserController“ valdiklis naudoja „ApplicationUser“ esybę. Valdiklis „AdministrationController“ naudoja „Category“ ir „QuerySPARQL“ esybes, o valdiklis „RequestSenderController“ naudoja „Request(test case)“ esybę. Valdiklis „StatisticContoller“ naudoja dvi esybes - „Request(test case)“ ir „Response“. „RequestControllesr“ valdiklis naudoja „Request(test case)“, „Response“ esybes.



46 pav. Detalizuota eksperimentinės sistemos loginė architektūra

3.5. Realizacijos projekto apibendrinimas

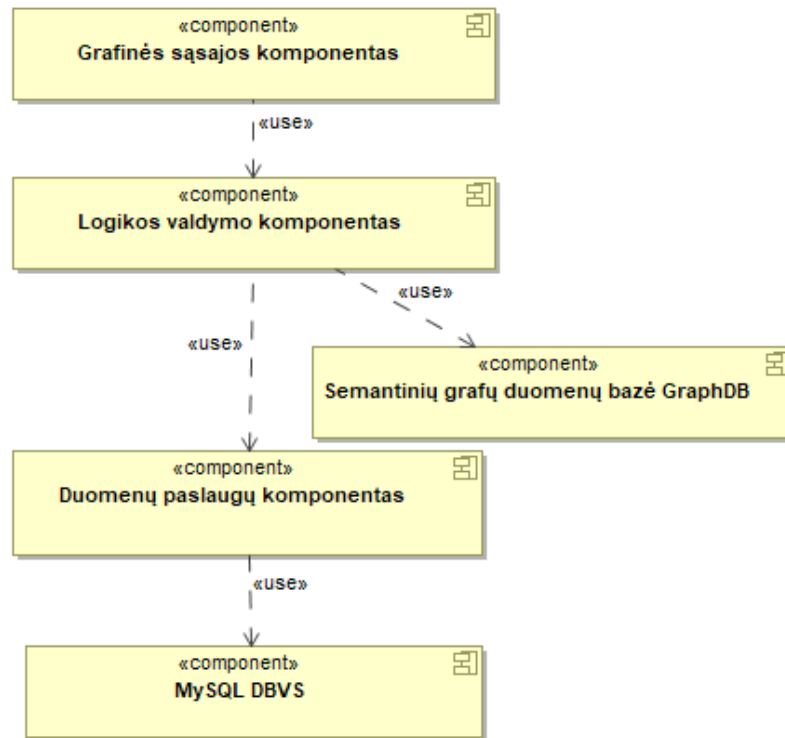
Išanalizavus kuriamai sistemai keliamus reikalavimus, buvo sukurta:

- sistemos loginė architektūra,
- eksperimentinės sistemos panaudojimo atvejai detalizuoti būsimos sistemos elgsenos modeliais,
- duomenų bazės modelis pagal esybių klasių modelį,
- detalus eksperimentinės sistemos posistemijų klasių modelis.

4. ONTOLOGIJOS PRATURTINTOS INDIVIDŲ ATPAŽINIMO TIKSLUMO ĮVERČIAIS EKSPERIMENTINĖS SISTEMOS REALIZACIJA IR TESTAVIMAS

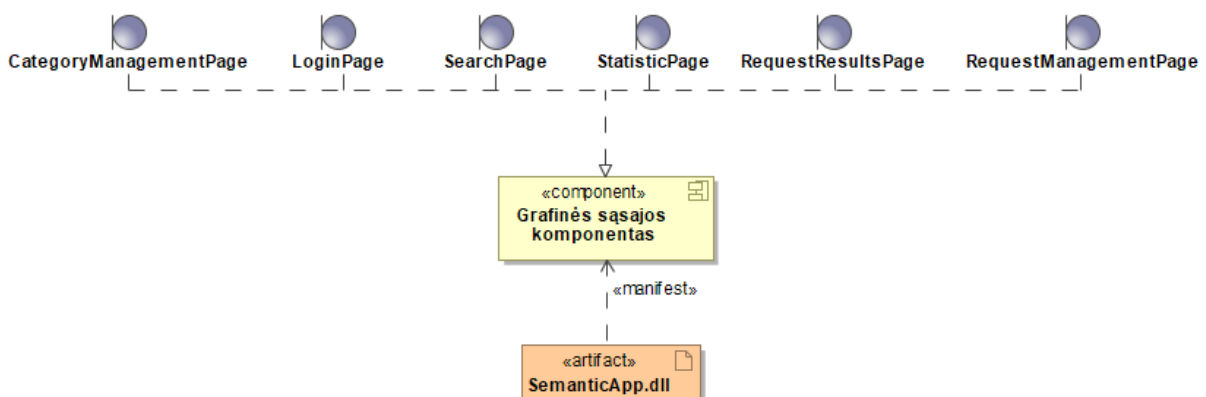
4.1. Sistemos realizacijos modelis ir veikimo aprašymas

Ontologijos praturtintos individų atpažinimo tikslo įverčiais komponentų modelyje (žr. 47 pav.) pateikiami esantys komponentai ir ryšiai tarp esamų komponentų.



47 pav. Ontologijos praturtintos individų atpažinimo tikslo įverčiais sistemos komponentų modelis

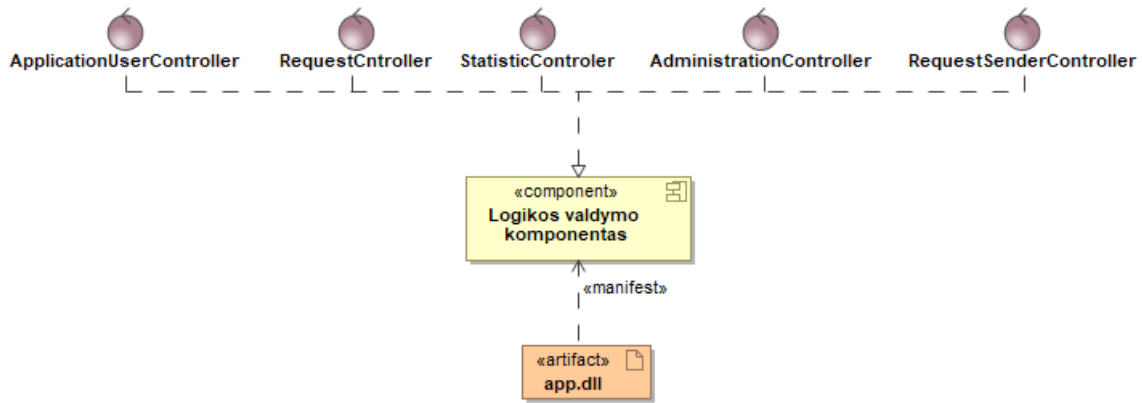
48 pav. pavaizduota grafinės sąsajos komponento modelis. Grafinės sąsajos komponento modelis sudarytas iš sistemos langų. Kiekvienas puslapis yra valdomas savo valdikliu, kuris informuoja apie pasikeitimus ir atitinkamai reaguoja į lange atliekamus naudotojo žingsnius. Grafinės sąsajos komponentas realizuojamas grafinės sąsajos atitinkamu artefaktu.



48 pav. Grafinės sąsajos komponento modelis

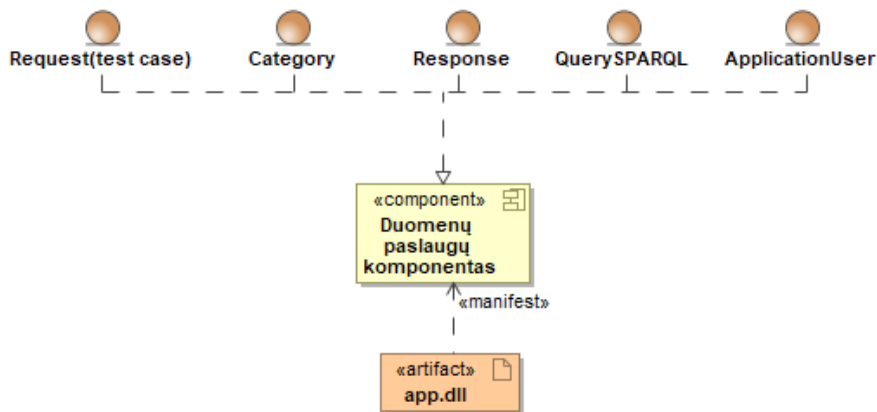
Logikos valdymo komponentas yra atsakingas už vykdomą logiką tarp grafinės sąsajos ir duomenų paslaugų komponento (žr. 49 pav.). Esantys valdikliai leidžia peržiūrėti ir redaguoti duomenų esybes,

kurti naujas, vykdyti semantinę paiešką pagal pateiktą užklausą, saugoti rezultatus ir formuoti rezultatų statistiką. Logikos vykdymo komponentas realizuotas logikos artefaktu.



49 pav. Logikos valdymo komponento modelis

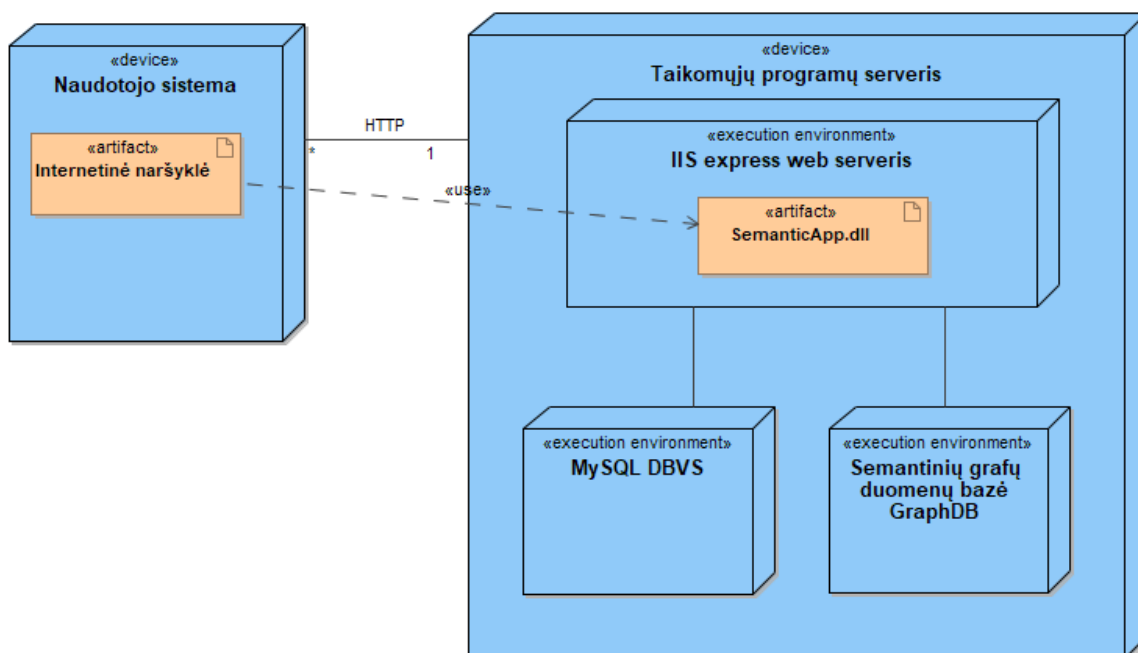
Duomenų paslaugų komponentas (žr. 50 pav.) suteikia galimybę manipuluoti duomenų esybių objektais, nesigilinant į naudojamas duomenų saugojimo technologijas. Komponentas skirtas komunikuoti su duomenų bazių valdymo sistema.



50 pav. Duomenų paslaugų komponento modelis

4.2. Diegimo modelis

Šiame skyriuje pateikiama eksperimentinės sistemos diegimo modelis (žr. 51 pav.). Naudotojo sistemą sudaro bet kuri pasirinkta internetinė naršyklė. Interneto naršyklė naudoja HTTP protokolą komunikavimui su taikomųjų programų serveriu.



51 pav. Semantinės paieškos eksperimentinės sistemos diegimo modelis

4.3. Naudojamos technologijos

Eksperimentinė sistema neturi reikalavimų kokiomis technologijomis turi būti realizuota eksperimentinė sistema. Ontologijų praturtintų individų atpažinimo tikslumo įverčiais eksperimentinė sistema yra realizuota naudojant .NET 4.6 programavimo kalbą, grafinė sąsaja realizuota su javascript. Sistemos duomenų saugojimui yra naudojama *MySQL* duomenų bazių valdymo sistema. Dokumentų saugojimui yra naudojama semantinių grafų duomenų bazė *GraphDB*.

4.4. Testavimo planas

Šiame poskyryje yra aprašoma realizuoto metodo testavimo planas. Aprašyti pagrindiniai žingsniai testuojant modelį, taip pat pateikiamas kontrolinis pavyzdys vaizduojantis testavimo plano vykdymo eigą.

Individų atpažinimo tikslumo įverčiais praturtintos ontologijos modelio ir jo tyrimo realizacijos testavimui sudaromas testavimo planas, siekiant pilnai išpildyti visus testavimo atvejus. Detalus testavimo planas pateiktas žemiau esančiose lentelėse.

4.1 lentelė. Valdyti paieškos sritis testavimo atvejais

Nr.	Testavimo atvejis	Valdyti paieškos sritis
Sužadinimo sąlyga(-os):		
	1. Paieškos sritis nėra sukurta;	
	2. Paieškos sritį reikia pašalinti;	
	3. Redaguoti paieškos sritį;	
Atliekami žingsniai:	Laukiamas sistemos atsakas:	Gautas sistemos rezultatas:
1) spausti mygtuką „Pridėti paieškos sritį“;	1) atvertas paieškos srities kūrimo langas;	1) sukurta paieškos sritis;
1) spausti mygtuką „Šalinti paieškos sritį“;	1) paieškos sritis šalinama;	1) pašalinta paieškos sritis;
1) spausti mygtuką „Redaguoti paieškos sritį“;	1) atvertas paieškos srities redagavimo langas;	1) poredaguota paieškos sritis;

Vykdyto tipas:	Rankinis
Rezultatas:	Įvykdytas

4.2 lentelė. Valdyti užklausas testavimo atvejis

Nr.	Testavimo atvejis Valdyti užklausas		
Sužadinimo sąlyga(-os): 1. Užklausa nėra sukurta; 2. Užklausa reikia pašalinti; 3. Redaguoti užklausa;			
Atliekami žingsniai:		Laukiamas sistemos atsakas:	Gautas sistemos atsakas:
1) spausti mygtuką „Pridėti užklausa“;		1) atvertas užklausos kūrimo langas;	1) sukurta užklausa;
1) spausti mygtuką „Šalinti užklausa“;		1) užklausa šalinama;	1) pašalinta užklausa;
1) spausti mygtuką „Redaguoti užklausa“;		1) atvertas užklausos redagavimo langas;	1) poredaguota užklausa;
Vykdyto tipas:		Rankinis	
Rezultatas:		Įvykdytas	

4.3 lentelė. Valdyti užklausų rezultatus testavimo atvejis

Nr.	Testavimo atvejis Valdyti užklausų rezultatus		
Sužadinimo sąlyga(-os): 1. Peržiūrėti užklausos rezultatus; 2. Šalinti užklausos rezultatus;			
Atliekami žingsniai:		Laukiamas sistemos atsakas:	Gautas sistemos atsakas:
1) spausti mygtuką „Peržiūrėti“;		1) atvertas užklausos rezultatų langas;	1) rezultatų peržiūra;
1) spausti mygtuką „Šalinti“;		1) rezultatas šalinamas;	1) pašalintas rezultatas;
Vykdyto tipas:		Rankinis	
Rezultatas:		Įvykdytas	

4.4 lentelė. Vykdyti semantinę paiešką testavimo atvejis

Nr.	Testavimo atvejis Sukurti semantines paieškos testavimo atvejį		
Sužadinimo sąlyga(-os): 1. Prisijungęs vartotojas nori sukurti paieškos testavimo atvejį.			
Atliekami žingsniai:		Laukiamas sistemos atsakas:	Gautas sistemos atsakas:
1) Pasirinkta paieškos sritis; 2) Pasirinkta imtis; 3) Pasirinkta užklausos sudėtingumas; 4) Pasirinkta užklausos tipas;		1) Sistema formuoja užklausų klausimus.	1) Suformuotas užklausų sąrašas;
5) Pasirinkta siūlomas užklausos klausimas; 6) Pasirinkta kreiptis;		1) Sistema formuoja užklausos atsakymą;	1) Suformuotas užklausos rezultatas;
Vykdyto tipas:		Rankinis	
Rezultatas:		Įvykdytas	

4.4.1. Kontrolinis pavyzdys

Kontrolinio pavyzdžio metu imituosime informacijos paiešką sukurtoje eksperimentinėje sistemoje. 52 pav. pavaizduotas eksperimentinės sistemos paieškos langas. Paieškos lange reikia užpildyti visus reikalaujamus laukus.

Užklauso

Paieškos sritis	Pasirinkite	▼
Imtis	Pasirinkite	▼
Užklauso sudėtingumas	Pasirinkite	▼
Užklauso tipas	Pasirinkite	▼
Siūlomų klausimų pasirinkimas	Pasirinkite	▼

52 pav. Eksperimentinės sistemos paieškos langas

Vartotojas nori atlikti paiešką pasirinkęs „Politikos“ paieškos sritį, 25 ontologijos straipsnių imtį, užklauso sudėtingumas „Dviejų tipų objektai“, užklauso tipas „Išvedamas tikslumo įvertis individo atžvilgiu“, iš siūlomų klausimų pasirenka „Ką kalbėjo asmenys?“ (žr. 53 pav.)

Paieškos sritis	Politika	▼
Imtis	25 dokumentų (tikslumas individo atžvilgiu)	▼
Užklauso sudėtingumas	Dviejų skirtingų tipų objektai	▼
Užklauso tipas	Išvedamas tikslumo įvertis individo atžvilgiu	▼
Siūlomų klausimų pasirinkimas	Ką kalbėjo asmenys?	▼
SPARQL užklausa	<pre>PREFIX ann: <http://semantika.lt/ns/Annotations#> PREFIX politics: <http://semantika.lt/ns/Politics#> PREFIX semLT: <http://semantika.lt/ns/SemLT#> PREFIX xsd: <http://www.w3.org/2001/XMLSchema#> PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#> PREFIX publicAdministrative: <http://semantika.lt/ns/PublicAdministrative#> PREFIX abstractObject: <http://semantika.lt/ns/AbstractObject#> PREFIX agents: <http://semantika.lt/ns/Agents#> PREFIX economyAndBusiness: <http://semantika.lt/ns/EconomyAndBusiness#> PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> PREFIX locations: <http://semantika.lt/ns/Locations#> PREFIX time: <http://semantika.lt/ns/Time#> PREFIX measurementUnits: <http://semantika.lt/ns/MeasurementUnits#> PREFIX events: <http://semantika.lt/ns/Events#> PREFIX evaluation: <http://semantika.lt/ns/Evaluation#></pre>	

```

SELECT DISTINCT ?asmuo_label ?turinys_label ?tikslumo_ivertis ?asmens_tikslumas ?turinio_tikslumas
WHERE
{
  ?dokumentas semLT:refers_to__object ?asmuo_i .
  ?dokumentas rdf:type semLT:document .
  OPTIONAL {?asmuo_i evaluation:fmeasure ?asmens_tikslumas .}
  OPTIONAL {?turinys_i evaluation:fmeasure ?turinio_tikslumas .}
  ?dokumentas semLT:refers_to__object ?kalbėjimas_i .
  ?asmuo_i events:talked__talking ?kalbėjimas_i .
  ?asmuo_i rdf:type agents:person .
  ?kalbėjimas_i rdf:type politics:talking .
  ?dokumentas semLT:refers_to__object ?turinys_i .
  BIND ( (?asmens_tikslumas*?turinio_tikslumas) AS ?tikslumo_ivertis )
  ?kalbėjimas_i events:talked__substance ?turinys_i .
  ?turinys_i rdf:type events:substance .
  ?asmuo_i rdfs:label ?asmuo_label .
  ?turinys_i rdfs:label ?turinys_label
}
order by DESC (?tikslumo_ivertis)

```

Kreiptis

53 pav. Eksperimentinės sistemos užpildytas paieškos langas

Įvykdžius užklausą „Ką kalbėjo asmenys?“ gauto rezultato fragmentas pavaizduotas 54 pav.

Rezultatai - 4/14/2019 10:04:41 PM

Trukmė 00:00:02.8306401 (2830.6401 ms)

Show 10 entries Search:

asmuo_label	turinys_label	tikslumo_ivertis	asmens_tikslumas	turinio_tikslumas
Stasys Kropas	„ Bankininkystės modeliai tokie , kad jie skaičiuoja bankroto tikimybę , ta tikimybė yra kliento nemokumo tikimybė . Modeliai įskaičiuoja krizės laikotarpio klientus , kurie turėjo tam tikrų mokumo problemų . Taip bus dar keletą metų į priekį , todėl maržų sumažėjimo galima tikėtis tik po kelių metų “	0.81	0.9	0.9
V. Grubliauskas	kad Uosto direkcijai pasiūlyta suformuoti darbo grupę , kuri galutinai ir atsakytų į klausimą , ar valčių prieplauką verta statyti senoje , ar siūlomoje naujoje vietoje .	0.81	0.9	0.9
Šarūnas Ruzgys	jog 9 proc. gyventojų planuoja rinktis pensiją ateityje gauti tik iš " Sodros " .	0.81	0.9	0.9
Radvilė Morkūnaitė	kad teisiniai susitarimai galėjo būti sukurti jau iki šiol .	0.81	0.9	0.9

54 pav. Užklausos „Ką kalbėjo asmenys?“ paieškos rezultato fragmentas

4.5. Realizacijos apibendrinimas

Atlikus realizaciją buvo realizuota ir ištestuota:

- ontologijos praturtintos individų atpažinimo tikslumo įverčiais sistemos komponentų modelis,
- eksperimentinės sistemos diegimo modelis,
- aprašytos sistemos realizavime naudojamos technologijos,

Ištestavus eksperimentinę sistemą buvo nustatyta, kad sistema tenkina iškeltus reikalavimus.

5. INDIVIDŲ ATPAŽINIMO TIKSLUMO ĮVERČIAIS PRATURTINTOS ONTOLOGIJOS EKSPERIMENTINIS TYRIMAS

Eksperimento metu bus tiriami galimi ontologijos praturtinimo individų atpažinimo tikslo įverčiai modelio realizavimo variantai. Tyrimo tikslas - įvertinti kokią įtaką semantinei paieškai (rezultatų turiniui ir greitaveikai) turi ontologijos papildymas individų (esybų ir ryšių tarp jų) atpažinimo tikslo įverčiais.

Toliau yra pateikiama eksperimentinio tyrimo aprašymo planas, pradiniai duomenys, tyrimo eiga, gautų eksperimento rezultatų analizė.

5.1. Eksperimento planas

Šiame skyriuje aprašyta kaip buvo organizuojamas eksperimentas.

Semantinių užklausų eksperimentinio tyrimo sistema.

Eksperimentui atlikti naudojama šiam tikslui sukurta semantinių užklausų eksperimentinio tyrimo sistema (žr. 4.1 skyrius). Pagrindinė sistemos funkcija – registruoti pateiktų sistemai SPARQL užklausų įvykdymo laiką. Sistemos naudotojas gali sukurti įvairaus sudėtingumo SPARQL užklausas ir jas vykdyti sistemos semantiniėje saugykloje. Užklausų vykdymo rezultatai yra susisteminti ir grafiškai pateikiami tolimesnei analizei.

Eksperimento duomenys.

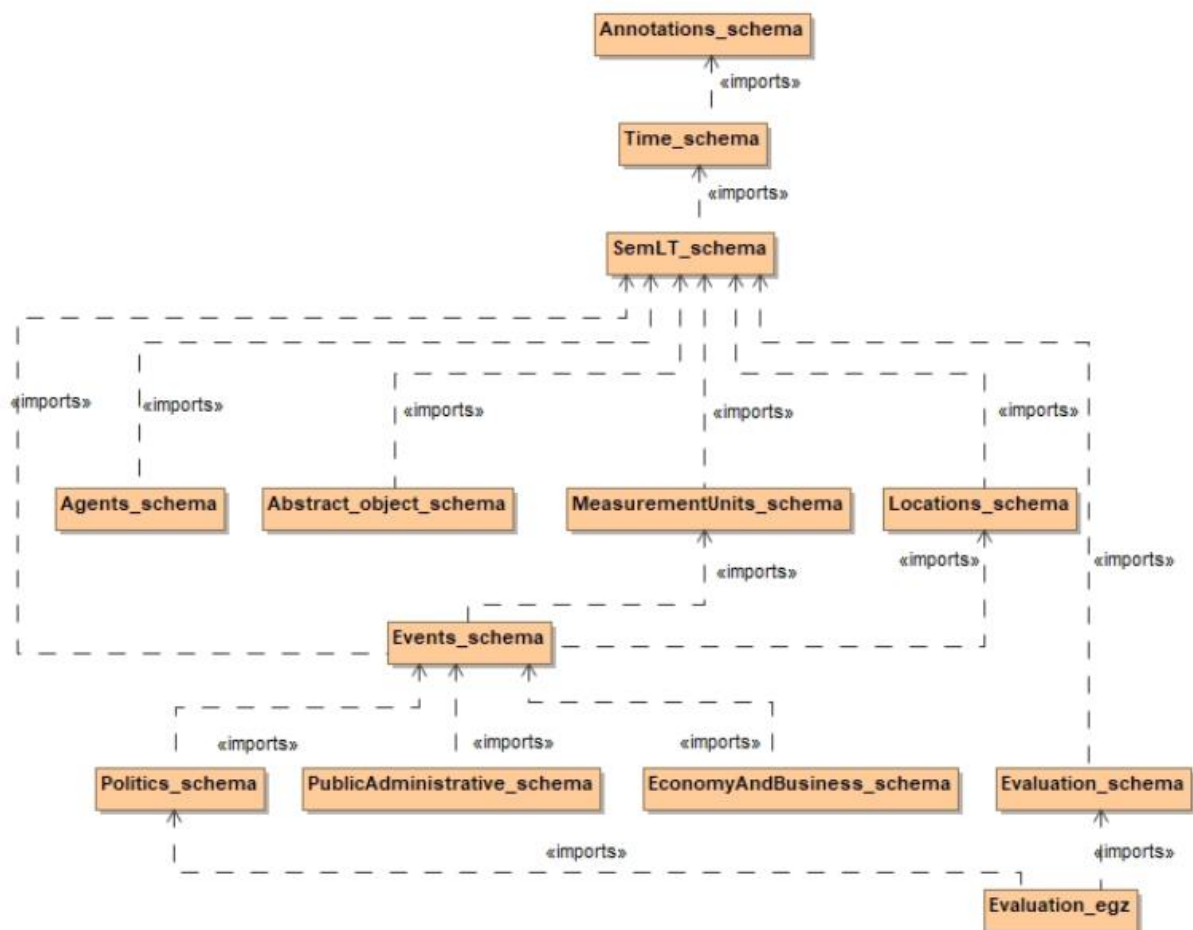
Eksperimentui vykdyti į eksperimentinės sistemos semantinių grafų duomenų bazę *GraphDB* buvo įkelta Lietuvių kalbos sintaksinės ir semantinės analizės informacinė sistemos (LKSSAIS) sukurta ontologija [24]. Šios ontologijos schema yra sukurta modulinio principu ir ją sudaro kelios integruotos schemas (žr. 5.1 lentelė). Ontologija užpildyta individais, kurie sukurti automatiškai išanalizavus interneto portaluose publikuojamus straipsnius politikos ir ekonomikos tematika.

5.1 lentelė. Ontologijų schemų sąrašas

Nr.	Ontologijos schemas pavadinimas	Paaiškinimas
1.	<i>SemLT_schema.rdf</i>	SemLT ontologija aprašo du pagrindinius konceptus: dokumentus ir objektus.
2.	<i>PublicAdministrative_schema.rdf</i>	Viešojo administravimo ontologija skirta nelaimių, nusikaltimų, paieškos ir pasiūlos įvykiams aprašyti.
3.	<i>Politics_schema.rdf</i>	Politikos ontologija skirta politikos įvykiams (sakymams, susitikimams) aprašyti ir sakymams kategorizuoti pagal pateiktą sakymų kategorizavimo schemą.
4.	<i>MeasurementUnits_schema.rdf</i>	Matavimo vienetų ontologija skirta aprašyti matavimo vienetams ir jų sąryšiams, kurie naudojami specialiųjų sričių ontologijų rinkiniuose.
5.	<i>Locations_schema.rdf</i>	Vietovių ontologija skirta aprašyti vietovėms (žemynams, regionams, šalims, miestams, vandens telkiniams ir kt.) ir jų sąryšiams, kurie naudojami specialiųjų sričių ontologijų rinkiniuose.
6.	<i>Events_schema.rdf</i>	Įvykių ontologija skirta aprašyti bendriniais įvykiams (pvz., dalyvavimui, darbui, susitikimui), susijusiems su objektų, agentų, vietovių ir laiko konceptais.

7.	<i>Agents_schema.rdf</i>	Agentų ontologija skirta aprašyti asmenų, organizacijų ir grupių conceptams ir jų sąryšiams, kurie naudojami specialiųjų sričių ontologijų rinkiniuose.
8.	<i>Abstract_object_schema.rdf</i>	Abstrakčių objektų ontologija skirta viešojo sektoriaus ir verslo poreikiams aktualiems abstrakčių objektų conceptams (turtui, nedarbui, skundams, lėšoms, ekonomikai ir kt.) aprašyti, kategorizuojant juos į grupes.
9.	<i>Time_schema.rdf</i>	Laiko ontologija skirta laiko klasėms ir jų sąryšiams aprašyti.
10.	<i>EconomyAndBusiness_schema.rdf</i>	Ekonomikos ir verslo ontologija skirta aprašyti šios srities įvykiams, pavyzdžiui, pokyčiams, pasiūlai ir paklausai.
11.	<i>Annotations_schema.rdf</i>	Anotacijų ontologija yra skirta aprašyti anotacijoms (metaduomenų savybėms), kurios naudojamos visuose kituose specialiųjų sričių ontologijų rinkiniuose.

Naudojamos LKSSAIS ontologijų schemas (žr. 5.1 lentelė.) papildytos dar viena schema „*Evaluation_schema*“, kuri aprašo F-mato parametą. F-matas yra naujas parametras, kuris nebuvo aprašytas ankstesnėse ontologijų schemose. Žemiau esančiame paveikslėlyje 55 pav. LKSSAIS ontologijų schemų ir „*Evaluation_schema*“ importavimo struktūra. Taip pat matoma ontologijos schemų priklausomybių struktūra, kurioje yra pridėta ir nauja „*Evaluation_schema*“.



55 pav. LKSSAIS ontologijų schemų ir „*Evaluation_schema*“ importavimo struktūra

Paveikslėlyje 56 pav., vaizduojama struktūra su joje aprašytu F-matas parametru, kuris yra pridedamas dokumentui arba individui.

```

<?xml version="1.0"?>
<rdf:RDF xmlns="http://semantika.lt/ns/Evaluation#"
  xml:base="http://semantika.lt/ns/Evaluation"
  xmlns:ann="http://semantika.lt/ns/Annotations#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:swrlb="http://www.w3.org/2003/11/swrlb#"
  xmlns:swrl="http://www.w3.org/2003/11/swrl#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:eval="http://semantika.lt/ns/Evaluation#"
  xmlns:s="http://semantika.lt/ns/SemLT#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:terms="http://purl.org/dc/terms/"
  xmlns:xml="http://www.w3.org/XML/1998/namespace"
  xmlns:vann="http://purl.org/vocab/vann/"
  xmlns:foaf="http://xmlns.com/foaf/0.1/"
  xmlns:dc="http://purl.org/dc/elements/1.1/">
  <owl:Ontology rdf:about="http://semantika.lt/ns/Evaluation#"
    <owl:imports rdf:resource="http://semantika.lt/ns/SemLT"/>
  </owl:Ontology>

  <!--
  ////////////////////////////////////////////////////////////////////
  //
  // Data properties
  //
  ////////////////////////////////////////////////////////////////////
  -->

  <!-- http://semantika.lt/ns/Evaluation#fmeasure -->

  <owl:DatatypeProperty rdf:about="http://semantika.lt/ns/Evaluation#fmeasure">
    <rdfs:domain rdf:resource="http://semantika.lt/ns/SemLT#document"/>
    <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#decimal"/>
    <ann:label_lt xml:lang="lt">F matas</ann:label_lt>
    <ann:label_sbvr xml:lang="lt">F_matas</ann:label_sbvr>
    <rdfs:comment xml:lang="lt">Čia saugomas anotavimo grandinės tikslumo įvertinimo matas F</rdfs:comment>
    <rdfs:label xml:lang="lt">F matas</rdfs:label>
  </owl:DatatypeProperty>
</rdf:RDF>

<!-- Generated by the OWL API (version 4.2.6.20160910-2108) https://github.com/owlcs/owlapi -->

```

56 pav. Ontologijos schema „Evaluation_schema“

Nagrinėjami dokumentai pasipildė papildomu parametru F-matas, lyginant nuo pradinės esybių ryšių diagramos pavaizduotos 26 pav., dabar tai vaizduoja paveikslėlis 57 pav.:



57 pav. Papildyta ontologijos schema duomenų savybe F-matas

Taip pat ontologijų schemos papildytos „Evaluation_egz“ failu, kuriame išdėstomi dokumentai su atitinkamu F-matas parametru, 58 pav. pateikiama failo fragmentas, pilnas failo aprašymas išdėstytas 8.3 priede :

```

<?xml version="1.0"?>
<rdf:RDF xmlns="http://www.w3.org/2002/07/owl#"
  xml:base="http://www.w3.org/2002/07/owl"
  xmlns:eval="http://semantika.lt/ns/Evaluation#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:terms="http://purl.org/dc/terms/"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:xml="http://www.w3.org/XML/1998/namespace"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:vann="http://purl.org/vocab/vann/"
  xmlns:dc="http://purl.org/dc/elements/1.1/" >
<owl:Ontology>
  <imports rdf:resource="http://semantika.lt/ns/Evaluation"/>
  <imports rdf:resource="http://semantika.lt/ns/Politics"/>
</owl:Ontology>

<!--
////////////////////////////////////
//
// Individuals
//
////////////////////////////////////
-->

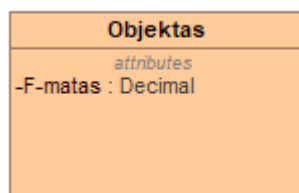
<!--http://semantika.lt/ns/SemLT#document~62c84fdd-27f9-44eb-95af-6c34aa6e55db~SemLT.document-1 -->

<NamedIndividual      rdf:about="http://semantika.lt/ns/SemLT#document~62c84fdd-27f9-44eb-95af-6c34aa6e55db~SemLT.document-1">
  <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.9</eval:fmeasure>
</NamedIndividual>
  <NamedIndividual      rdf:about="http://semantika.lt/ns/SemLT#document~38291e6f-a9ed-4b13-862f-8646888dffc4~SemLT.document-1">
  <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.7</eval:fmeasure>
</NamedIndividual>
  <NamedIndividual      rdf:about="http://semantika.lt/ns/SemLT#document~dce0a47c-86b4-4782-9c36-ad8a49bf7e73~SemLT.document-1">
  <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.8</eval:fmeasure>
</NamedIndividual>
  .....

```

58 pav. Ontologijos dokumentų įverčių failo „Evaluation_egz.“ fragmentas

Nagrinėjamų dokumentų individai buvo papildyti parametru F-matas, tai vaizduoja paveikslėlis 59 pav.



59 pav. Individas papildyta duomenų savybe F-matas

Iš turimos ontologijos atsitiktine tvarka yra išrinkta 100 dokumentų (semantiškai išanalizuotų žiniasklaidos straipsnių) ir su jais susijusi informacija (tai kas straipsniuose buvo atpažinta atlikus semantinę analizę). Eksperimente naudojamas ontologijos dydis yra matuojamas tripletų kiekiu. Eksperimentui atrinktą ontologiją sudaro:

- ontologijos dydis be schemos (*A-Box*) yra 266,595 tripletų, o su schema (*A-Box* ir *T-Box*) – 523,323 tripletų;
- vidutiniškai vieną straipsnį aprašo 4465 tripletų.

Viename šimte eksperimente naudojamų ontologijos straipsnių yra atpažinta tokių klasių individų:

- 1564 asmenų,
- 776 vietovių,
- 664 organizacijų,
- 676 pasisakymų.

Baigiamojo darbo eksperimento vykdymui naudojamus dokumentus reikia pasiruošti.

Dokumentų pasiruošimas papildant naujais faktais.

Tam, kad eksperimento metu įvertinti kokią įtaką rezultatų turiniui ir greitaveikai turi ontologijos papildymas individų atpažinimo tikslumo įverčiais, prieš eksperimentą straipsniai papildyti tikslumo įverčiais dviem būdais:

- dokumento atžvilgiu;
- individo atžvilgiu;

Ontologijų praplėtimas dokumento atžvilgiu, remiasi prielaida, kad eksperimente naudojamus straipsnius praleidus pro natūralios kalbos ir semantinės analizės ir anotavimo komponentų grandinėle, visi individai yra atpažinti vienodu tikslumu. Tokiu atveju tikslumo įvertį pakanka nurodyti tik straipsniui, o ne kiekvienam individui. Tokiu būdu yra gerokai mažiau padidinama ontologijos apimtis.

Ontologijų praplėtimas individo atžvilgiu, remiasi prielaida, kad eksperimente naudojamus straipsnius praleidus pro natūralios kalbos ir semantinės analizės ir anotavimo komponentų grandinėle, visi individai yra atpažinti su skirtingu tikslumu. Tokiu atveju tikslumo įvertį reikia nurodyti kiekvienam straipsnyje atpažintam individui.

Ontologijos praplėstos tikslumo įverčiu, kurio reikšmė atsitiktinai priskirta iš reikšmių intervalo nuo 0,9 iki 0,4. Kuo didesnė priskirta skaitinė reikšmė tuo jame minimos esybės ir įvairūs faktai yra laikomi tiksliau atpažinti.

Prieš papildant ontologiją tikslumo įverčiais ontologijos schema papildyta nauja schema *Evaluation_schema.rdf* (žr.56 pav.), kurioje specifikuotas tikslumo įverčio parametras. Praplėstų ontologijos straipsnių dokumento atžvilgiu tikslumo įverčiai išsaugoti atskirame *Evaluation_egz.rdf* faile (žr. 8.3 priedas). Ontologijos straipsniai praplėsti individų atžvilgiu tikslumo įverčius saugo konkrečiame straipsnyje, papildant konkretų individą.

Dokumentų suskirstymas į skirtingas imtis.

Kad eksperimento metu pamatyti kokią įtaką turi pridėtas tikslumo įverčio parametras greitaveikai, įtaką tiriama bandant užklausas ant skirtingos dydžio ontologijų. Eksperimente naudojamas vienas

šimtas ontologijos straipsnių yra suskirstytas į tris skirtingas imtis. Mažiausioje imtyje yra 25-i ontologijų straipsniai, vidutinėje imtyje yra 50-imt ontologijos straipsnių, didžiausioje imtyje yra 100-tas ontologijos straipsnių. Ontologijos straipsniai į imtis suskirstomi taip, kad išvedamų rezultatų kiekis kistų proporcingai.

Eksperimentinėje sistemoje ontologijai saugoti yra naudojama semantinių grafų duomenų bazė *GraphDB*. Šioje sistemoje galima sukurti kelias saugyklas. Suformuotos trys dokumentų imtys yra sukeltos į eksperimentinės sistemos duomenų devynias saugyklas, pagal suformuotas imtis.

SPARQL užklausos.

Eksperimentui iš viso yra sukurtos 54 užklausos. Jos buvo sudarytos pasirinkus 18 skirtingų klausimų politikos ir ekonomikos sritims.

Užklausos yra sugrupuotos į tris tipus, pagal tai ar užklausa reikalauja išvesti tikslumo įvertį atsakymui ar ne:

- neišvedamas tikslumo įvertis;
- išvedamas tikslumo įvertis dokumento atžvilgiu;
- išvedamas tikslumo įvertis individo atžvilgiu;

Taip pat užklausos suskirstytos pagal tai, kiek skirtingų tipų objektų (klasių individų) ieškoma ir pateikiama atsakyme:

- vienas tipo objektas;
- dviejų skirtingų tipų objektai;
- dviejų skirtingų tipų objektai ir konkretus objektas (pavyzdžiui, „Vokietija“, „Lietuva“ arba „Europa“);

Lentelėse (5.2 lentelė,

5.3 lentelė, 5.4 lentelė.) pateikiama po vieną pavyzdį skirtingo sudėtingumo SPARQL užklausų. Visas užklausas naudotas eksperimente žr. 8.3 priede.

5.2 lentelė. Vieno objekto sudėtingumo užklausos pavyzdys

Užklausos sudėtingumas	Užklausos tipas	Kokie asmenys yra minimi?
Vieno tipo objektas	neišvedamas tikslumo įvertis	<pre>SELECT DISTINCT ?asmuo_label WHERE { ?dokumentas semLT:refers_to__object ?asmuo_i . ?dokumentas rdf:type semLT:document. ?asmuo_i rdf:type agents:person . ?asmuo_i rdfs:label ?asmuo_label . }</pre>
	išvedamas tikslumo įvertis dokumento atžvilgiu	<pre>SELECT DISTINCT ?asmuo_label ?tikslumo_iverdis WHERE { ?dokumentas semLT:refers_to__object ?asmuo_i . ?dokumentas rdf:type semLT:document. OPTIONAL {?dokumentas evaluation:fmeasure ?tikslumo_iverdis . } ?asmuo_i rdf:type agents:person . ?asmuo_i rdfs:label ?asmuo_label . } order by DESC (?tikslumo_iverdis)</pre>

	išvedamas tikslumo įvertis individo atžvilgiu	<pre> SELECT DISTINCT ?asmuo_label ?tikslumo_ivertis WHERE { ?dokumentas semLT:refers_to__object ?asmuo_i . ?dokumentas rdf:type semLT:document. OPTIONAL {?asmuo_i evaluation:fmeasure ?tikslumo_ivertis. } ?asmuo_i rdf:type agents:person . ?asmuo_i rdfs:label ?asmuo_label . } order by DESC (?tikslumo_ivertis) </pre>
--	---	--

5.3 lentelė. Dviejų objektų sudėtingumo užklauso pavyzdys

Užklauso sudėtingumas	Užklauso tipas	Ar yra koks nors kainų pokytis
Dviejų skirtingų tipų objektai	neišvedamas tikslumo įvertis	<pre> SELECT DISTINCT ?kaina_label ?pokytis_label WHERE { ?dokumentas semLT:refers_to__object ?kaina_i . ?dokumentas rdf:type semLT:document. ?dokumentas semLT:refers_to__object ?pokytis_i . ?pokytis_i economyAndBusiness:has__abstract_object ?kaina_i. ?kaina_i rdf:type abstractObject:abstract_object . ?pokytis_i rdf:type economyAndBusiness:change . ?kaina_i rdfs:label ?kaina_label . ?pokytis_i rdfs:label ?pokytis_label . } </pre>
	išvedamas tikslumo įvertis dokumento atžvilgiu	<pre> SELECT DISTINCT ?kaina_label ?pokytis_label ?tikslumo_ivertis WHERE { ?dokumentas semLT:refers_to__object ?kaina_i . ?dokumentas rdf:type semLT:document. OPTIONAL {?dokumentas evaluation:fmeasure ?tikslumo_ivertis. } ?dokumentas semLT:refers_to__object ?pokytis_i . ?pokytis_i economyAndBusiness:has__abstract_object ?kaina_i. ?kaina_i rdf:type abstractObject:abstract_object . ?pokytis_i rdf:type economyAndBusiness:change . ?kaina_i rdfs:label ?kaina_label . ?pokytis_i rdfs:label ?pokytis_label . } order by DESC (?tikslumo_ivertis) </pre>
	išvedamas tikslumo įvertis individo atžvilgiu	<pre> SELECT DISTINCT ?kaina_label ?pokytis_label ?tikslumo_ivertis ?kainos_ivertis ?pokycio_ivertis WHERE { ?dokumentas semLT:refers_to__object ?kaina_i . ?dokumentas rdf:type semLT:document. OPTIONAL {?kaina_i evaluation:fmeasure ?kainos_ivertis. } OPTIONAL {?pokytis_i evaluation:fmeasure ?pokycio_ivertis. } ?dokumentas semLT:refers_to__object ?pokytis_i . ?pokytis_i economyAndBusiness:has__abstract_object ?kaina_i. ?kaina_i rdf:type abstractObject:abstract_object . BIND ((?kainos_ivertis*?pokycio_ivertis) AS ?tikslumo_ivertis) ?kaina_i rdfs:label ?kaina_label . ?pokytis_i rdfs:label ?pokytis_label . } order by DESC (?tikslumo_ivertis) </pre>

5.4 lentelė. Du objektai+ Vokietija sudėtingumo užklauso pavyzdys

Užklauso sudėtingumas	Užklauso tipas	Ar yra koks nors kainų pokytis
Dviejų skirtingų tipų objektai + konkretus objektas	neišvedamas tikslumo įvertis	<pre>SELECT DISTINCT ?asmuo_label ?turinys_label WHERE { ?dokumentas semLT:refers_to__object ?asmuo_i . ?dokumentas rdf:type semLT:document. ?dokumentas semLT:refers_to__object ?kalbėjimas_i . ?asmuo_i events:talked__talking ?kalbėjimas_i . ?asmuo_i rdf:type agents:person . ?kalbėjimas_i rdf:type politics:talking . ?dokumentas semLT:refers_to__object ?turinys_i . ?kalbėjimas_i events:talked__substance ?turinys_i . ?turinys_i rdf:type events:substance . ?asmuo_i rdfs:label ?asmuo_label . ?turinys_i rdfs:label ?turinys_label. ?salis ann:label_lemma "vokietija"@lt . ?salis rdf:type locations:location . ?kalbėjimas_i events:included__object ?salis. }</pre>
	išvedamas tikslumo įvertis dokumento atžvilgiu	<pre>SELECT DISTINCT ?asmuo_label ?turinys_label ?tikslumo_ivertis WHERE { ?dokumentas semLT:refers_to__object ?asmuo_i . ?dokumentas rdf:type semLT:document. OPTIONAL {?dokumentas evaluation:fmeasure ?tikslumo_ivertis . } ?dokumentas semLT:refers_to__object ?kalbėjimas_i . ?asmuo_i events:talked__talking ?kalbėjimas_i . ?asmuo_i rdf:type agents:person . ?kalbėjimas_i rdf:type politics:talking . ?dokumentas semLT:refers_to__object ?turinys_i . ?kalbėjimas_i events:talked__substance ?turinys_i . ?turinys_i rdf:type events:substance . ?asmuo_i rdfs:label ?asmuo_label . ?turinys_i rdfs:label ?turinys_label. ?salis ann:label_lemma "vokietija"@lt . ?salis rdf:type locations:location . ?kalbėjimas_i events:included__object ?salis. } order by DESC (?tikslumo_ivertis)</pre>
	išvedamas tikslumo įvertis individo atžvilgiu	<pre>SELECT DISTINCT ?asmuo_label ?turinys_label ?tikslumo_ivertis ?asmens_tikslumas ?kalbejimo_tikslumas ?turinio_tikslumas WHERE { ?dokumentas semLT:refers_to__object ?asmuo_i . ?dokumentas rdf:type semLT:document. OPTIONAL {?asmuo_i evaluation:fmeasure ?asmens_tikslumas . } OPTIONAL {?kalbėjimas_i evaluation:fmeasure ?kalbejimo_tikslumas . } OPTIONAL {?turinys_i evaluation:fmeasure ?turinio_tikslumas . } ?dokumentas semLT:refers_to__object ?kalbėjimas_i . ?asmuo_i events:talked__talking ?kalbėjimas_i . ?asmuo_i rdf:type agents:person . ?kalbėjimas_i rdf:type politics:talking . ?dokumentas semLT:refers_to__object ?turinys_i . BIND ((?asmens_tikslumas*?kalbejimo_tikslumas*?tutinio_tikslumas) AS ?tikslumo_ivertis) ?kalbėjimas_i events:talked__substance ?turinys_i .</pre>

		?turinys_i rdf:type events:substance . ?asmuo_i rdfs:label ?asmuo_label . ?turinys_i rdfs:label ?turinys_label. ?salis ann:label_lemma "vokietija"@lt . ?salis rdf:type locations:location . ?kalbėjimas_i events:included__object ?salis. } order by DESC (?tikslumo_ivertis)
--	--	--

Eksperimento vykdymas.

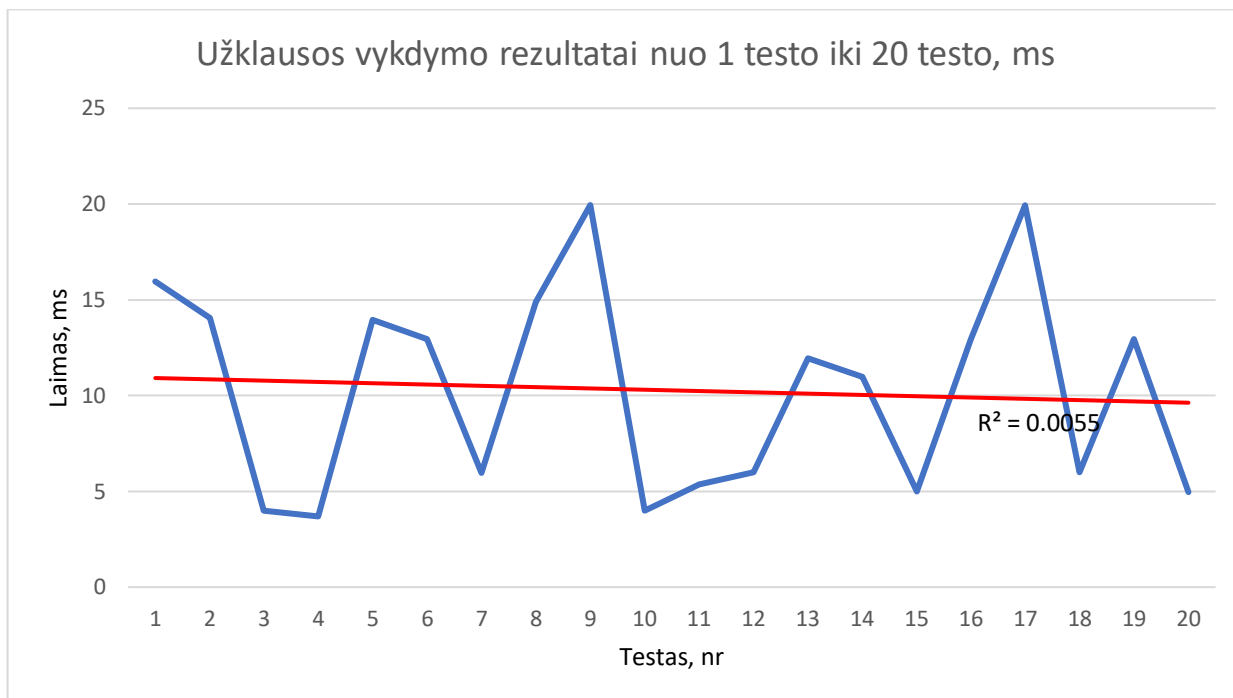
Eksperimentinėje sistemoje kiekviena suplanuota užklausa vykdyta 40 kartų. Išvedamų rezultatų greita veika fiksuota kiekvienu kartu išvedant duomenis. Rezultatų duomenys yra saugoti eksperimentinės sistemos duomenų bazėje. Iš 40 vienos užklauskos užfiksuotų vykdymo greita veikos rezultatų iš duomenų bazės paimti paskutiniai 20 rezultatų. Toks rezultatų kiekis naudojamas, nes po 20 pirmų užklauskos vykdymų fiksuojami greita veikos rezultatai tampa pastovesni, t.y. nebūna didelių trukmės šuolių. Trukmės šuoliai susidaro dėl *GraphDB* užklauskų variklio veikimo algoritmo, susijusio su atminties resursų valdymui.

5.5 lentelė. parodo „Ką kalbėjo asmenys minėdami 'Vokietija?'“ užklauskos 40 testų vykdymo rezultatus.

5.5 lentelė. Užklauskos „Ką kalbėjo asmenys minėdami 'Vokietija?'“ vykdymo rezultatai

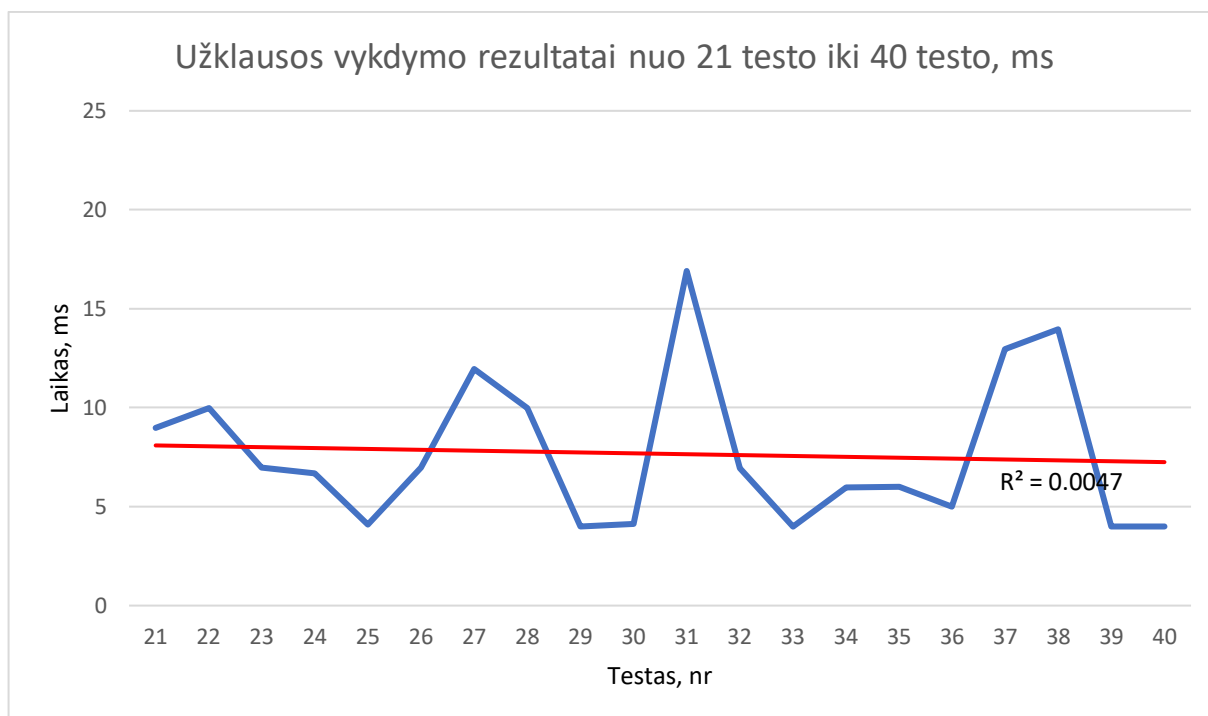
Užklauskos pavadinimas	Užklauskos sunkumas	Užklauskos tipas	Imtis	nuo 1 iki 20, ms	nuo 20 iki 40, ms
Ką kalbėjo asmenys minėdami 'Vokietija?'	Dviejų skirtingų tipų objektai ir konkretus objektas 'Vokietija'	be tikslumo įverčio	50 dokumentų (tikslumo nėra)	15.9577	8.9815
				14.0578	9.9685
				3.9971	6.9809
				3.6921	6.6737
				13.9612	4.0799
				12.9618	6.9813
				5.9507	11.9617
				14.9022	9.9725
				19.9475	3.9896
				3.9901	4.1225
				5.3638	16.9079
				5.9829	6.9396
				11.9583	3.989
				10.9604	5.9849
				4.9827	5.9867
				12.9151	4.9897
				19.946	12.963
				5.9829	13.9634
				12.9642	3.99
				4.9492	3.9893

Užklauskos „Ką kalbėjo asmenys minėdami 'Vokietija?'“ vykdymo nuo 1 iki 20 testo rezultatų diagrama pateikta 60 pav. Gauti rezultatai parodo, kad vykdant užklauską pirmuose 20 testų įvykdymo laikas turi laiko šuolius. Vykdymo laiko skalės ribos nuo 3.6921 ms iki 19.9475 ms., kur skirtumas yra lygus 16.2554 ms. Išskaičiuota tiesinė regresija yra lygi 0,0055.



60 pav. Užklauso „Ką kalbėjo asmenys minėdami 'Vokietija?'“ nuo 1-20 testo rezultatų diagrama

Užklauso „Ką kalbėjo asmenys minėdami 'Vokietija?'“ vykdymo nuo 21 iki 40 testo rezultatų diagrama pateikta 61 pav. Gauti rezultatai parodo, kad vykdant užklausą nuo 21 testo įvykdymo laikas turi mažesnius laiko šuolius lyginant su pirmais 20 vykdomų testu. Vykdyto laiko skalės ribos nuo 3.9893 ms iki 16.9079 ms, kur skirtumas lygus 12.9186 ms. Išskaičiuota tiesinė regresija yra lygi 0,0047.



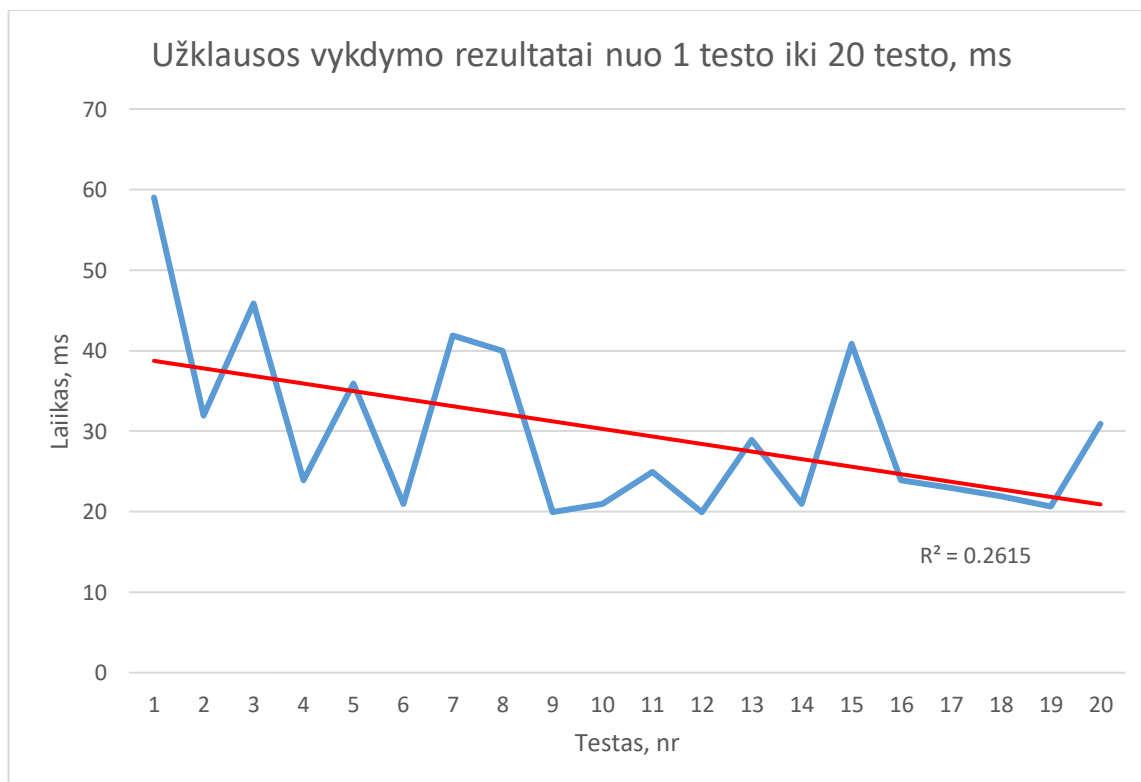
61 pav. Užklauso „Ką kalbėjo asmenys minėdami 'Vokietija?'“ nuo 20-40 testo rezultatų diagrama

5.6 lentelė parodo „Visos vietovės“ užklauso 40 testų vykdyto rezultatus. Šie testai yra atlikti su dvigubai didesniu duomenų skaičiumi.

5.6 lentelė Užklauso „Visos vietovės“ vykdymo rezultatai

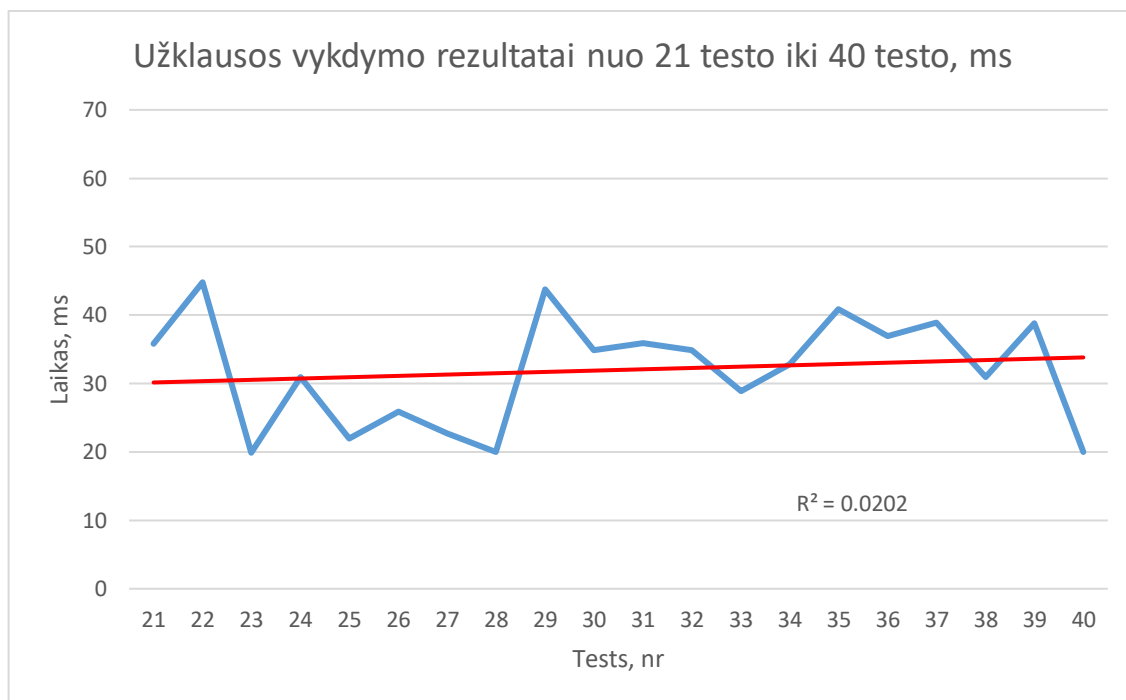
Užklauso pavadinimas	Užklauso sunkumas	Užklauso tipas	Imtis	nuo 1 iki 20, ms	nuo 20 iki 40, ms
Visos vietovės?	Vieno tipo objektas	be tikslumo įverčio	100 dokumentų (tikslumo nėra)	59.0005	35.8198
				31.9138	44.7966
				45.8745	19.885
				23.9353	30.9163
				35.9037	21.9378
				20.9463	25.9295
				41.8406	22.6877
				39.9302	19.9473
				19.9464	43.8071
				20.9811	34.9057
				24.9329	35.9414
				19.9492	34.9016
				28.9226	28.8787
				20.9471	32.8529
				40.8419	40.8872
				23.9366	36.8962
				22.9397	38.9364
				21.9385	30.9175
				20.6562	38.8393
				30.8725	19.948

Užklauso „Visos vietovės“ vykdymo nuo 1 iki 20 testo rezultatų diagrama pateikta 62 pav. Gauti rezultatai parodo, kad vykdant užklausą pirmuose 20 testų įvykdymo laikas turi laiko šuolius. Vykdomo laiko skalės ribos nuo 19.9464 ms iki 59.0005 ms., kur skirtumas yra lygus 39.0541 ms. Išskaičiuota tiesinė regresija yra lygi 0.2615.



62 pav. Užklauso „Visos vietovės“ nuo 1-20 testo rezultatų diagrama

Užklauso „Visos vietovės“ vykdymo nuo 21 iki 40 testo rezultatų diagrama pateikta 63 pav. Gauti rezultatai parodo, kad vykdant užklausą nuo 21 testo įvykdymo laikas turi mažesnius laiko šuolius lyginant su pirmais 20 vykdymų testu. Vykdyto laiko skalės ribos nuo 19.948 ms iki 44.7966 ms, kur skirtumas lygus 12.9186 ms. Išskaičiuota tiesinė regresija yra lygi 0,0202.



63 pav. Užklauso „Visos vietovės“ nuo 21-40 testo rezultatų diagrama

Lyginant dviejų pavyzdžių rezultatus galima pastebėti, kad laiko šuoliai išlieka nepriklausomai nuo to koks yra naudojamas duomenų kiekis.

Iš paskutinių 20 rezultatų yra apskaičiuojamas greitaiveikos vidurkis kiekvienai užklausiai. Taip pat yra suskaičiuojamas imties standartinis nuokrypis (STDEV.S), kad pamatyti kiek reikšmė skiriasi nuo vidurkio.

$$STDEV.S = \sqrt{\frac{\sum(x-\bar{x})^2}{(n-1)}}; \quad (6)$$

čia \bar{x} yra imties vidurkis $AVERAGE(\text{skaičius1}, \text{skaičius2}, \dots)$ ir n yra imties dydis.

5.2. Eksperimento rezultatai

Atlikus eksperimentą užklausų greitaiveika buvo suskaičiuota ant skirtingos imties duomenų. Gauti eksperimento rezultatai pateikti lentelėse ir pavaizduoti diagramomis. Lentelėse pateikiamas vykdytos užklauso pavadinimas, sudėtingumas, užklauso tipas, imtis, išvedamų įrašų kiekis, išvedamų įrašų greitaiveikos vidurkis ir standartinis nuokrypis nuo vidurkio

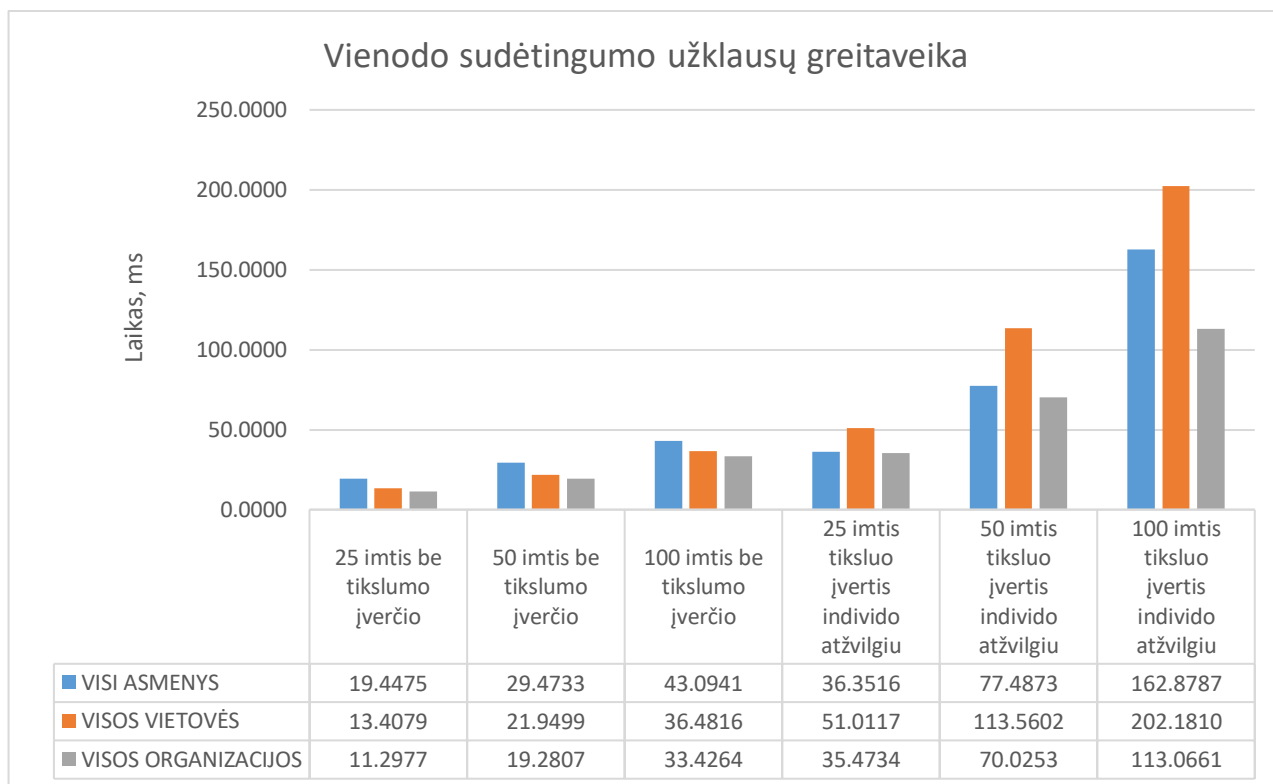
5.7 lentelė parodo vienodo sudėtingumo užklausų greitaiveikos rezultatus.

5.7 lentelė. Vienodo sunkumo užklausų rezultatai

Užklauso pavadinimas	Užklauso sudėtingumas	Užklauso tipas	Imtis	Išvedamų įrašų skaičius	Vidurkis, ms	Paklaida, ms
VISI ASMENYS		be tikslumo įverčio	25 imtis	294	19.4475	9.9721

	Vieno tipo objektas	tikslumo įvertis individo atžvilgiu	50 imtis	792	29.4733	14.3419
			100 imtis	1564	43.0941	12.8955
			25 imtis	294	36.3516	16.7687
			50 imtis	792	77.4873	28.0378
			100 imtis	1564	162.8787	77.6477
VISOS VIETOVĖS	Vieno tipo objektas	be tikslumo įverčio	25 imtis	291	13.4079	5.8457
			50 imtis	504	21.9499	10.8477
			100 imtis	776	36.4816	36.4816
		tikslumo įvertis individo atžvilgiu	25 imtis	291	51.0117	11.1467
			50 imtis	504	113.5602	39.7000
VISOS ORGANIZACIJOS	Vieno tipo objektas	be tikslumo įverčio	25 imtis	170	11.2977	6.4312
			50 imtis	406	19.2807	10.6170
			100 imtis	664	33.4264	12.2048
		tikslumo įvertis individo atžvilgiu	25 imtis	170	35.4734	12.5113
			50 imtis	406	70.0253	18.3308
			100 imtis	664	113.0661	22.5586

Vienodo sudėtingumo užklausų vykdymo greitaveikos diagrama pateikta 64 pav. Gauti rezultatai rodo kokią įtaką rezultatų išvedimo trukmei turi pridėtas papildomas tikslumo įverčio parametras ir naudojamas ontologijos dokumentų kiekis. Didėjant ontologijos dokumentų kiekiui didėja ir išvedamų rezultatų laiko trukmė. SPARQL užklausų, kur užklausų tipas „be tikslumo įverčio“, įvykdymo laiko skirtumas yra minimalus t.y. apie 6 ms. Naudojant SPARQL užklausias, kur užklausų tipas „tikslumo įvertis individo lygyje“, įvykdymo laikas beveik padvigubėja.



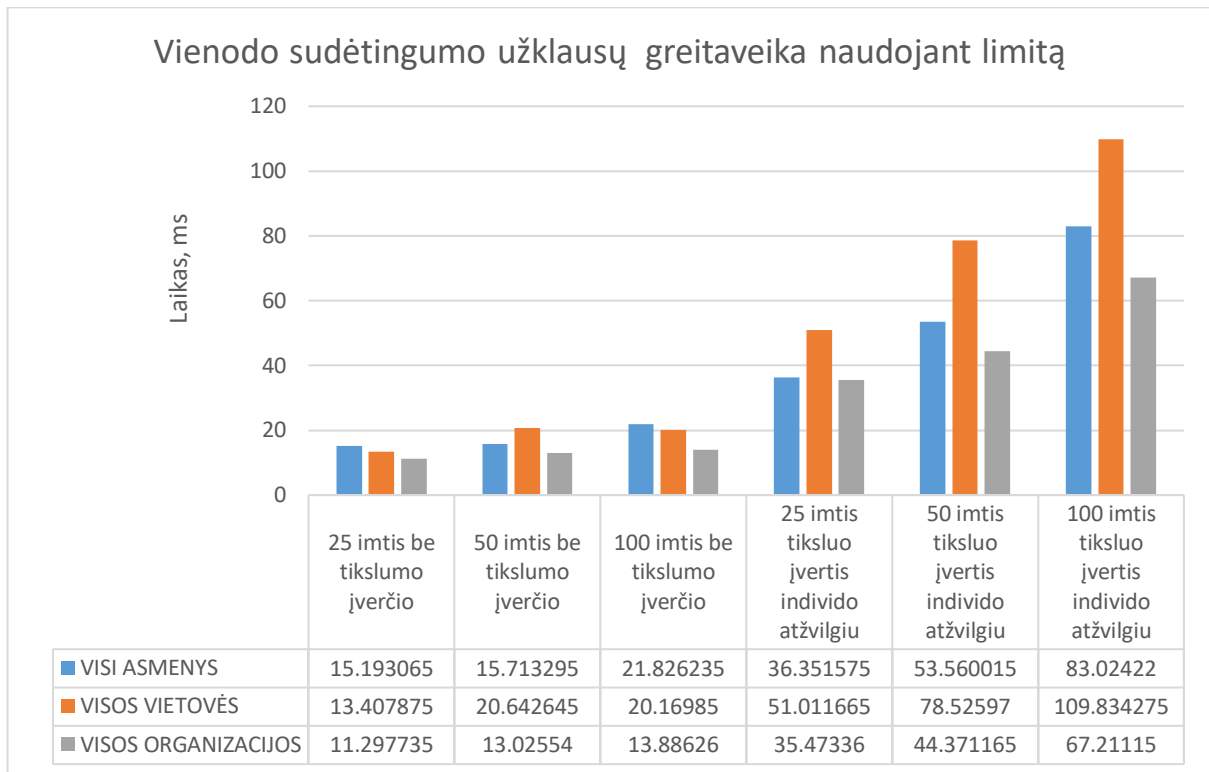
64 pav. Vienodo sudėtingumo užklausų greitaveikos diagrama

5.8 lentelė pateikia vienodo sudėtingumo užklausų rezultatus naudojant paieškos limitą. Paieškos limitas nustatytas pagal mažiausią dokumentų imtį ir joje išvedamų atsakymų skaičių.

5.8 lentelė. Vienodo sunkumo užklausų rezultatai naudojant paieškos limitą

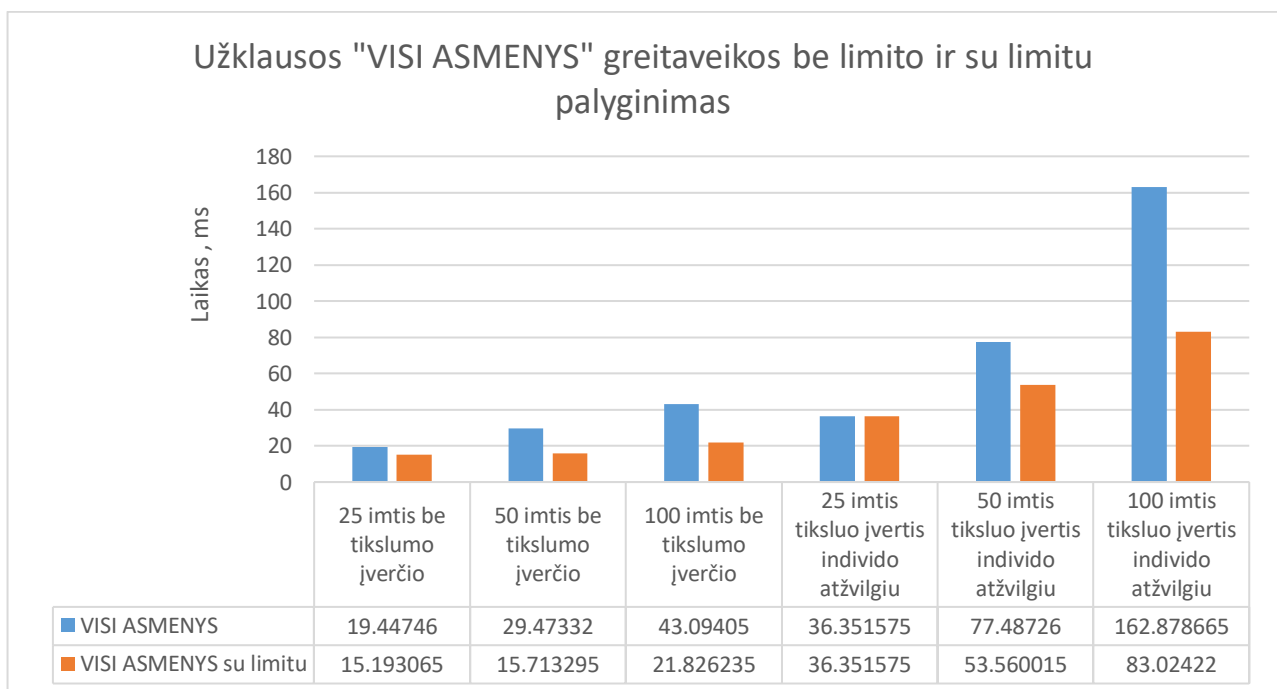
Užklausos pavadinimas	Užklausos sudėtingumas	Užklausos tipas	Imtys	Išvedamų įrašų skaičius	Vidurkis, ms	Paklaida, ms
VISI ASMENYS?	Vieno tipo objektas	be tikslumo įverčio	25 imtis	294	15.1931	8.0330
			50 imtis	294	15.7133	6.9111
			100 imtis	294	21.8262	11.9719
		tikslumo įvertis individo atžvilgiu	25 imtis	294	36.3516	16.7687
			50 imtis	294	53.5600	14.0825
			100 imtis	294	83.0242	22.0911
VISOS VIETOVĖS	Vieno tipo objektas	be tikslumo įverčio	25 imtis	291	13.4079	5.8457
			50 imtis	291	20.6426	9.5797
			100 imtis	291	20.1699	12.5238
		tikslumo įvertis individo atžvilgiu	25 imtis	291	51.0117	11.1467
			50 imtis	291	78.5260	19.3673
			100 imtis	291	109.8343	18.2725
VISOS ORGANIZACIJOS	Vieno tipo objektas	be tikslumo įverčio	25 imtis	170	11.2977	6.4312
			50 imtis	170	13.0255	6.8809
			100 imtis	170	13.8863	11.4655
		tikslumo įvertis individo atžvilgiu	25 imtis	170	35.4734	12.5113
			50 imtis	170	44.3712	10.8553
			100 imtis	170	67.2112	12.4053

Vienodo sudėtingumo užklausų greitaveikos diagrama naudojant limitą pateikta 65 pav. Gauti rezultatai rodo kokią įtaką rezultatų išvedimo trukmei turi pridėtas papildomas tikslumo įverčio parametras, naudojamas ontologijos dokumentų kiekis ir uždėtas apribojimas išvedamų rezultatų kiekiui.



65 pav. Vienodo sudėtingumo užklausų greitimeikos diagrama naudojant limitą

Užklausos „Visi asmenys“ greitimeikos palyginimas be limito ir su limitu pateiktas 66 pav.. Gauti rezultatai rodo kokią įtaką rezultatų išvedimo greičiui turi pridėtas limitas išvedamų duomenų kiekiui. Grafikas parodo, kad nepriklausomai nuo užklausos tipo išvedamų rezultatų greitimeika padidėja nuo 1 iki 2 kartų.



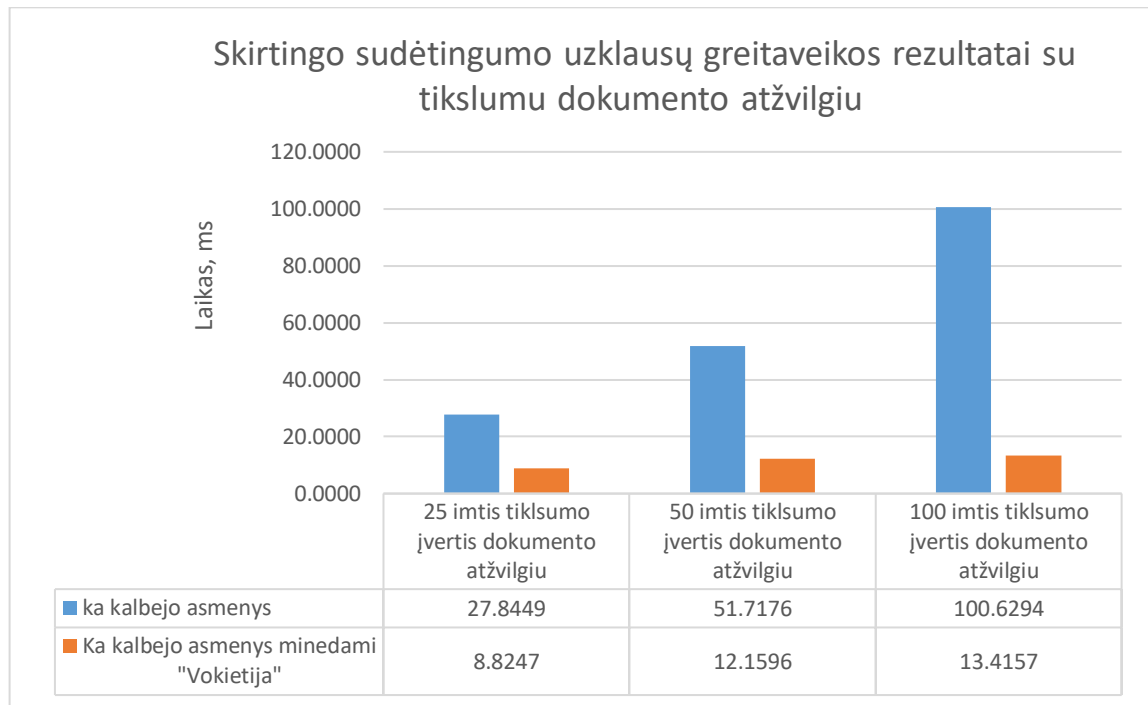
66 pav. Užklausos "VISI ASMENYS" greitimeikos be limito ir su limitu grafikas

5.9 lentelė pateikia skirtingo sudėtingumo užklausų rezultatus.

5.9 lentelė. Skirtingo sudėtingumo užklausų greitaveikos rezultatai

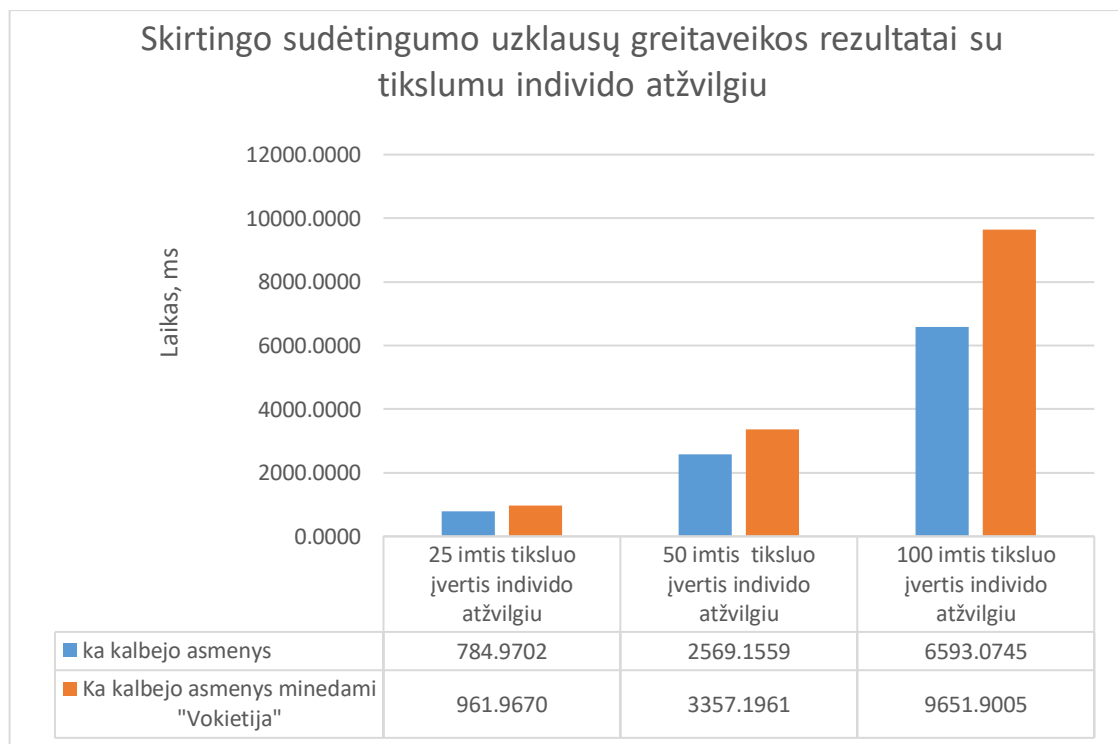
Užklausos pavadinimas	užklausos sunkumas	užklausos tipas	imtys	Išvedamų įrašų skaičius	Vidurkis, ms	Paklaida, ms
Ką kalbėjo asmenys	dviejų skirtingų tipų objektai	be tikslumo įverčio	25 imtis	69	18.8370	8.8633
			50 imtis	278	40.4862	16.1472
			100 imtis	684	74.6463	16.5418
		tikslumo įvertis dokumento atžvilgiu	25 imtis	69	27.8449	13.9733
			50 imtis	278	51.7176	15.2986
			100 imtis	684	100.6294	27.7176
		tikslumo įvertis individo atžvilgiu	25 imtis	69	784.9702	199.4744
			50 imtis	278	2569.1559	615.2406
			100 imtis	684	6593.0745	1285.9527
Ką kalbėjo asmenys minėdami "Vokietija"	dviejų skirtingų tipų objektai	be tikslumo įverčio	25 imtis	1	8.0777	4.1834
			50 imtis	4	7.6708	3.8298
			100 imtis	9	9.9269	4.5281
		tikslumo įvertis dokumento atžvilgiu	25 imtis	1	8.8247	5.1608
			50 imtis	4	12.1596	5.0935
			100 imtis	9	13.4157	11.0395
		tikslumo įvertis individo atžvilgiu	25 imtis	1	961.9670	86.1543
			50 imtis	4	3357.1961	90.2974
			100 imtis	9	9651.9005	522.9373

Skirtingo sudėtingumo užklausų greitaveikos diagrama su tikslumu dokumento lygyje pateikta 67 pav. Gauti rezultatai rodo kokią įtaką rezultatų išvedimo trukmei turi pridėtas papildomas tikslumo įverčio parametras dokumento lygyje, naudojamas ontologijos dokumentų imtis ir išvedamų įrašų kiekis. Užklausos „Ką kalbėjo asmenys minėdami "Vokietija"“ sąlygos dalis „WHERE“ yra sudaryta iš daugiau tripletų nei užklausos „Ką kalbėjo asmenys“, šiuo atveju įtaką rezultatų išvedimo greitaveikai turi išvedamų rezultatų kiekis. Su užklausa „Ką kalbėjo asmenys minėdami "Vokietija"“ yra išvedama įrašų 69/278/684 kiekis atitinkamai pagal imtis, o „Ką kalbėjo asmenys“ – 1/4/9 kiekis atitinkamai pagal imtis.



67 pav. Skirtingo sudėtingumo užklausų greitaveikos diagrama su tikslumu dokumento atžvilgiu

Skirtingo sudėtingumo užklausų greitaveikos diagrama su tikslumu individo lygyje pateikta 68 pav. Gauti rezultatai rodo kokią įtaką rezultatų išvedimo trukmei turi pridėtas papildomas tikslumo įverčio parametras individo atžvilgiu, naudojama ontologijos dokumentų imtis ir išvedamų įrašų kiekis. Šiuo atveju įtaką rezultatų išvedimo greitaveikai turi užklausos sudėtingumas, o ne išvedamų rezultatų kiekis, kaip kad 67 pav. atveju.



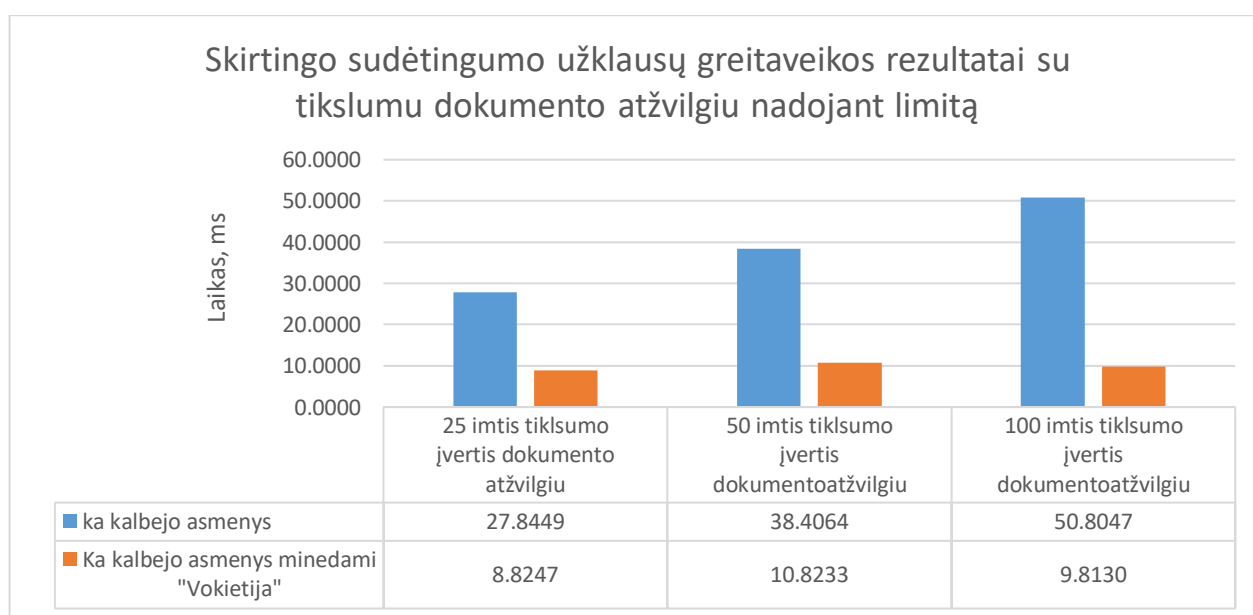
68 pav. Skirtingo sudėtingumo užklausų greitaveikos diagrama su tikslumu individo atžvilgiu

5.10 lentelė pateikia skirtingo sudėtingumo užklausų rezultatus naudojant išvedamų rezultatų limitą. Paieškos limitas nustatytas pagal mažiausią dokumentų imtį ir joje išvedamų atsakymų skaičių.

5.10 lentelė. Skirtingo sudėtingumo užklausų greitimeikos rezultatai naudojant limitą

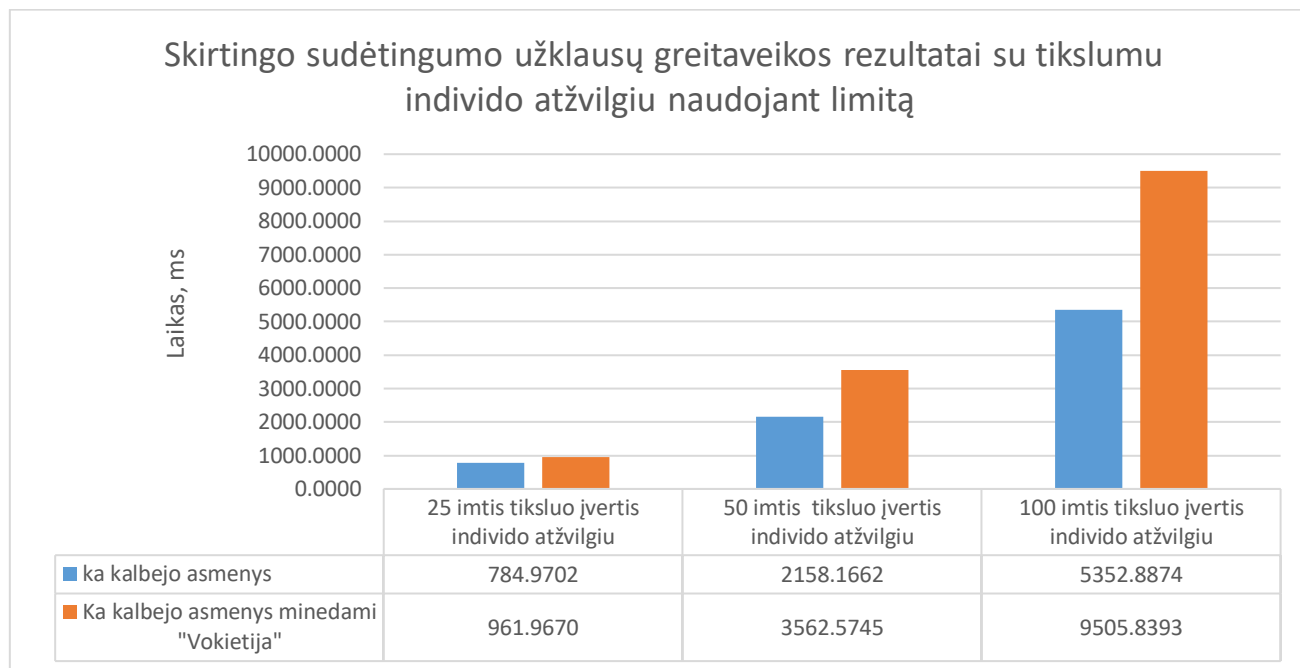
Užklausos pavadinimas	Užklausos sunkumas	Užklausos tipas	Imtys	Išvedamų įrašų skaičius	Vidurkis, ms	Paklaida, ms
Ką kalbėjo asmenys	dviejų skirtingų tipų objektai	be tikslumo įverčio	25 imtis	69	16.0771	9.6220
			50 imtis	69	17.4378	8.9166
			100 imtis	69	22.2695	8.1427
		tikslumo įvertis dokumento atžvilgiu	25 imtis	69	27.8449	13.9733
			50 imtis	69	38.4064	16.1787
			100 imtis	69	50.8047	15.2539
		tikslumo įvertis individo atžvilgiu	25 imtis	69	784.9702	199.4744
			50 imtis	69	2158.1662	119.6074
			100 imtis	69	5352.8874	62.3543
Ką kalbėjo asmenys minėdami "Vokietija"	dviejų skirtingų tipų objektai	be tikslumo įverčio	25 imtis	1	8.0777	4.1834
			50 imtis	1	11.1140	6.6747
			100 imtis	1	8.0046	5.0395
		tikslumo įvertis dokumento atžvilgiu	25 imtis	1	8.8247	5.1608
			50 imtis	1	10.8233	5.8316
			100 imtis	1	9.8130	4.3127
		tikslumo įvertis individo atžvilgiu	25 imtis	1	961.9670	86.1543
			50 imtis	1	3562.5745	570.8062
			100 imtis	1	9505.8393	952.9009

Skirtingo sudėtingumo užklausų greitimeikos diagrama su tikslumu dokumento atžvilgiu naudojant limitą pateikta 69 pav.. Gauti rezultatai rodo kokią įtaką rezultatų išvedimo trukmei turi pridėtas papildomas tikslumo įverčio parametras dokumento atžvilgiu, naudojama ontologijos dokumentų imtis ir išvedamų įrašų kiekis. Nurodytas išvedamų įrašų kiekis greitimeikai teigiamos įtakos neturėjo. Išvedimo greitis minimaliai, bet padidėjo lyginant su 67 pav. diagrama.



69 pav. Skirtingo sudėtingumo užklausų greitimeikos diagrama su tikslumu dokumento atžvilgiu naudojant limitą

Skirtingo sudėtingumo užklausų greitaveikos diagrama su tikslumu individo atžvilgiu naudojant limitą pateikta 70 pav.. Gauti rezultatai rodo kokią įtaką rezultatų išvedimo trukmei turi pridėtas papildomas tikslumo įverčio parametras individo atžvilgiu, naudojama ontologijos dokumentų imtis ir išvedamų įrašų kiekis. Nurodytas išvedamų įrašų kiekis greitaveikai teigiamos įtakos neturėjo. Išvedimo greitis minimaliai, bet padidėjo lyginant su 69 pav. diagrama.



70 pav. Skirtingo sudėtingumo užklausų greitaveikos diagrama su tikslumu individo atžvilgiu naudojant limitą

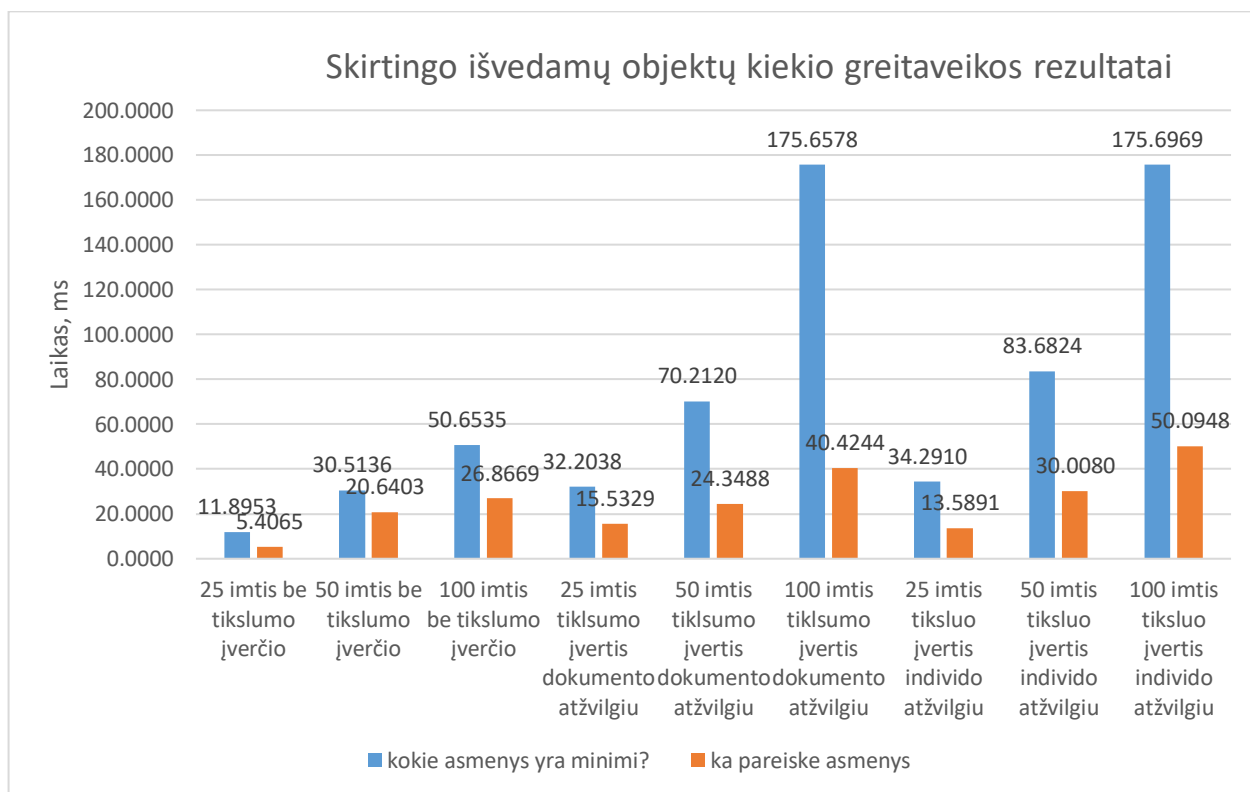
5.11 lentelė pateikia skirtingo išvedamų objektų kiekio rezultatus.

5.11 lentelė Skirtingo išvedamų objektų kiekio greitaveikos rezultatai

Užklausos pavadinimas	Užklausos sunkumas	Užklausos tipas	Imtys	Išvedamų įrašų skaičius	Vidurkis, ms	Paklaida, ms
Kokie asmenys yra minimi?	Vieno tipo objektas	be tikslumo įverčio	25 imtis	294	11.8953	4.7820
			50 imtis	792	30.5136	11.4743
			100 imtis	1564	50.6535	13.8651
		tikslumo įvertis dokumento atžvilgiu	25 imtis	294	32.2038	11.1928
			50 imtis	792	70.2120	14.6600
			100 imtis	1564	175.6578	73.8982
		tikslumo įvertis individo atžvilgiu	25 imtis	294	34.2910	14.3459
			50 imtis	792	83.6824	18.5298
			100 imtis	1564	175.6969	50.7378
Ką pareiškė asmenys?	dviejų skirtingų tipų objektai	be tikslumo įverčio	25 imtis	20	5.4065	15.9520
			50 imtis	66	20.6403	12.2778
			100 imtis	176	26.8669	9.4184
		tikslumo įvertis dokumento atžvilgiu	25 imtis	20	15.5329	6.5376
			50 imtis	66	24.3488	9.2255
			100 imtis	176	40.4244	12.6702
		tikslumo įvertis individo atžvilgiu	25 imtis	20	13.5891	5.8770
			50 imtis	66	30.0080	12.2363

			100 imtis	176	50.0948	14.3864
--	--	--	-----------	-----	---------	---------

Pateikta skirtingo išvedamų objektų kiekio greitaveikos diagrama pavaizduota 71 pav. Gauti rezultatai rodo kokią įtaką rezultatų išvedimo trukmei turi pridėtas papildomas tikslumo įverčio parametras dokumento arba individo atžvilgiu, naudojama ontologijos dokumentų imtis ir išvedamų įrašų kiekis. Diagramos rezultatai parodo, kad greitaveika priklauso nuo išvedamų rezultatų kiekio, nors išvedamas tik vienas objektas ir užklausa „Kokie asmenys yra minimi“ pagal sudėtingumą yra su mažesniu tripletų kiekiu nei „Ką pareiškė asmenys“.



71 pav. Skirtingo išvedamų objektų kiekio greitaveikos diagrama

5.3. Sprendimo veikimo ir savybių analizė, kokybės kriterijų įvertinimas

Atlikto eksperimento rezultatai rodo, kad papildomas parametras dokumento atžvilgiu ar individo atžvilgiu turi akivaizdžios įtakos paieškos greitaveikai. Įtaka priklauso nuo užklauso sudėtingumo, atsakymo struktūros, išvedamų rezultatų kiekio ir ontologijos dydžio. Paieškos atvejais kai naudojamos nesudėtingos užklauso, greitaveiką galima padidinti iki 2 kartų nustačius išvedamų rezultatų limitą, ko pasėkoje vartotojas matytų maksimaliai pačius tiksliausius rezultatus, t.y., rezultatus su aukščiausiai tikslumo įverčiais.

Sukurtose užklauso paieškos rezultatai rikiuojami pagal geriausią tikslumo įvertį. Netaikant tokio rikiavimo, išvedamų rezultatų tikslumas tampa neaiškus, reikalaujantis papildomos vartotojo analizės, pvz. rikiavimas pagal pavadinimą. Tačiau rikiuojant užklauso atsakymo rezultatus užklauso vykdymo trukmė išauga. Iš eksperimento duomenų matyti, kad kuo tikslesnio atsakymo norima tuo didesnė rezultato išvedimo trukmė. Ilgiausiai tukęs išvedimo laikas pateiktuose rezultatuose yra 9651.9005 ms, vykdant užklausą ant 100 ontologijos straipsnių, o trumpiausias išvedimo laikas 5.4065 ms, vykdant užklausą ant 25 ontologijos straipsnių, nenaudojant limito duomenų išvedimui.

5.4. Sprendimo taikymo rekomendacijos

Ontologijos praturtintos individų atpažinimo tikslumo įverčiais eksperimentinis tyrimas atskleidė, kad specifikuotas papildomas parametras greitaveiką įtakoja ją sulėtinant. Todėl norint taikyti pasiūlytą semantinės paieškos modelį, reikia ieškoti sprendimų kaip padidinti užklausų vykdymo greitaveiką neprarandant galimybės tikslesnius duomenis pateikti pirmiau.

Atlikto eksperimento metu vartotojui buvo pirmiausiai pateikti tiksliausiai atpažinti užklausos rezultatai. Tačiau natūralu, kad gali kilti poreikis gauti tam tikro pasirinkto tikslumo rezultatus. Todėl reiktų atlikti eksperimentą su užklausomis, kuriose būtų nurodyti tikslumo režiai, ir įvertinti tokių režijų įtraukimo įtaką užklausos vykdymo greitaveikai.

Sukurtą eksperimentinę sistemą būtų galima padaryti lankstesne saugomų užklausų atžvilgiu. Norint tirti kitus semantinių užklausų greitaveikos aspektus eksperimentinę sistemą reiktų patobulinti praplečiant galimybę saugoti įvairesnes užklausas.

6. REZULTATŲ APIBENDRINIMAS IR IŠVADOS

Šio darbo metu atlikus analizės, realizacijos ir eksperimento darbus buvo gautos tokios išvados:

1. Semantinės paieškos LKSSAIS sprendimo karkasas suteikia semantinės paieškos galimybę, bet išvestas rezultatas yra atsitiktinis. Atlikus analizę yra nustatyta, kad yra sprendimų tokių kaip „Ontotext“ kompanijos, kurie saugo semantinės analizės komponentų veikimo tikslumą nusakančių parametrų įverčius. Tačiau nėra duomenų kiek tokio parametro saugojimas turi įtakos semantinės paieškos greitaveikai.
2. Atlikus esamų sprendimų analizę nustatyta, kad visuotinai pripažintų semantinių užklausų greitaveikai tirti skirtų reikalavimų, sistemų, semantinių užklausų ir duomenų nėra sukurta, juos reikia sukurti.
3. Atlikus suplanuoto tyrimo reikalavimų analizę, buvo suprojektuota ir realizuota eksperimentinė aplinka, t.y. sistema, turinti galimybę užfiksuoti įvairaus sudėtingumo semantinių užklausų vykdymo greitaveiką, sukurtos įvairaus sudėtingumo semantinės užklauso bei paruošti duomenys, t.y. įvairių individų kiekį turinčios ontologijos, tiek turinčios, tiek neturinčias individų atpažinimo tikslumo įverčius.
4. Atlikus eksperimentą nustatyta, kad ontologijos praturtintos individų atpažinimo tikslumo įverčiais, užklauso vykdymo greitaveiką įtakoja ją sulėtinant. Lyginant užklauso vykdymo laiką ontologijoje be tikslumo įverčių su vykdymu ontologijoje su tikslumo įverčiais vykdymo laikas ilgėja nuo 4 kartų iki 89 kartų.
5. Atlikus eksperimentą galima teigti norit semantinę paiešką padaryti patrauklia ne tik pateikiant tikslumą, būtina ieškoti sprendimų kaip didinti greitaveiką.

7. TURINYS

- [1] B. Yildiz, „Ontology Evolution and Versioning,“ 2006. [Tinkle]. Available: <http://se-pubs.dbs.uni-leipzig.de/files/Yildiz2006OntologyEvolutionand.pdf>. [Kreiptasi 20 10 2016].
- [2] „Protege,“ [Tinkle]. Available: http://protege.stanford.edu/publications/ontology_development/ontology101.pdf. [Kreiptasi 20 11 2016]. [Kreiptasi 20 11 2016].
- [3] „W3C Semantic web,“ [Tinkle]. Available: <http://www.w3.org/2001/sw/wiki/OWL>. [Kreiptasi 20 11 2016].
- [4] D. FREITAG, „Machine Learning for Information Extraction in Informal Domains,“ 26 balandžio 2000. [Tinkle]. Available: <http://www.cs.bilkent.edu.tr/~guvenir/courses/CS550/Seminar/freitag2000-ml.pdf>. [Kreiptasi 11 gegužės 2019].
- [5] C. Z. Charu C. Aggarwal, „Mining Text Data,“ 2012. [Tinkle]. Available: https://ink.library.smu.edu.sg/cgi/viewcontent.cgi?article=2710&context=sis_research. [Kreiptasi 11 gegužės 2019].
- [6] „Merriam-Webster,“ [Tinkle]. Available: <https://www.merriam-webster.com/dictionary/ontology>. [Kreiptasi 25 11 2016].
- [7] M. Shridhar, G. F. Houle ir F. Kimura, „Recognition strategies for general handwritten text documents,“ t. 16, nr. 4, 2009.
- [8] „Kompiuterinės lingvistikos centras,“ [Tinkle]. Available: http://donelaitis.vdu.lt/main.php?id=4&nr=7_1. [Kreiptasi 15 10 2016].
- [9] „Part of Speech Tagging with Discriminatively Re-ranked Hidden Markov Models,“ [Tinkle]. Available: <http://nlp.stanford.edu/courses/cs224n/2011/reports/highfill.pdf>. [Kreiptasi 23 01 2017].
- [10] „Alphabetical list of part-of-speech tags used in the Penn Treebank,“ [Tinkle]. Available: http://www.ling.upenn.edu/courses/Fall_2003/ling001/penn_treebank_pos.html. [Kreiptasi 23 01 2017].
- [11] C. D. Manning, C. D. Manning ir H. Schütze, Foundations of Statistical Natural Language Processing, MIT Press. ISBN 978-0-262-13360-9., 1999.
- [12] V. Vapnik, „Statistical Learning Theory,“ [Tinkle]. Available: <http://read.pudn.com/downloads161/ebook/733192/Statistical-Learning-Theory.pdf>. [Kreiptasi 20 04 2019].
- [13] C. A. A. R. Dmitry Zelenko, „Kernel Methods for Relation Extraction,“ [Tinkle]. Available: <http://www.jmlr.org/papers/volume3/zelenko03a/zelenko03a.pdf>. [Kreiptasi 25 10 2018].
- [14] E. A. a. L. Gravano, „Eugene Agichtein and Luis Gravano. Snowball: Extracting relations from large plain-text collections,“ įtraukta *In Proceedings of the 5th ACM Conference on Digital Libraries*.
- [15] Y. S. a. S. Sekine., „Preemptive information extraction using unrestricted relation discovery,“ įtraukta *In Proceedings of the Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics*, 2006.
- [16] D. L. i. L. H. N. Chincor, „Evaluating Message Understanding Systems,“ įtraukta *An Analysis Of The Third Message Understanding Conference (MUC-3)*, 1993.
- [17] C. Mangold, „A survey and classification of semantic search approaches,“ *Semantics and Ontology*, t. 2, p. pp.23–34, 2007.
- [18] L. Derczynski, „Complementarity, F-score, and NLP Evaluation,“ LREC , 2016.

- [19] „Ontotext,“ [Tinkle]. Available: http://ontotext.com/documents/white_papers/The-Truth-About-Triplestores.pdf . [Kreiptasi 26 01 2017].
- [20] T. Agresta, „PRweb,“ Ontotext, 21 08 2014. [Tinkle]. Available: <http://www.prweb.com/releases/2014/08/prweb12085470.htm>. [Kreiptasi 25 01 2017].
- [21] J. Rayfield, „Ontotext,“ 18 01 2019. [Tinkle]. Available: <https://www.ontotext.com/ontotext-platform-annotation-quality-assurance-inter-annotator-agreement/>. [Kreiptasi 01 05 2019].
- [22] „Ontotext,“ [Tinkle]. Available: <https://tag.ontotext.com/documentation/>. [Kreiptasi 01 05 2019].
- [23] „The National Center for Biotechnology Information,“ 14 11 2009 . [Tinkle]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2815499/> . [Kreiptasi 23 10 2017].
- [24] „Semantika,“ [Tinkle]. Available: <http://semantika.lt/Resources/Ontologies>. [Kreiptasi 01 04 2018].

8. PRIEDAI

8.1. priedas. Lietuvių kalbos morfologiniame anotatoriuje naudojamos žymos

Nr.	Požymių grupės pavadinimas	Požymis	Pažyma
1	Kalbos dalis	daiktavardis	dk.
		būdvardis	bdv.
		skaitvardis	sktv.
		įvardis	įv.
		veiksmažodis	vksm.
		prieveiksmis	prv.
		jaustukas	jst.
		ištiktukas	išt.
		dalelytė	dll.
		prielinksnis	prl.
		jungtukas	jng.
		akronimas	akronim.
sutrumpinimas	sutr.		
2	Daiktavardžių rūšys	tikrinis	tikr.
		bendrasis	
3	Veiksmažodis ir jo formos	bendrasis	bendr.
		dalyvis	dlv.
		padalyvis	pad.
		pusdalyvis	pusd.
		būdinys	būdn.
		tiesioginė nuosaka	tiesiog. n.
		liepiamoji nuosaka	liep. n.
tariamoji nuosaka	tar. n.		
4	Skaitvardžiai	kiekinis	kiek.
		kelintinis	kelint.
		dauginis	daugin.
		kuopinis	kuopin.
5	Apibrėžtumas	įvardžiuotinis	įvardž.
		neįvardžiuotinis	neįvardž.
6	Sangražiškumas	sangražinis	sngr.
		nesangražinis	nesngr.
7	Rūšis	veikiamoji rūšis	veik. r.
		neveikiamoji rūšis	neveik. r.
		reikiamybės	reik.
8	Laikas	esamasis laikas	es. l.
		būtamasis kartinis laikas	būt. k. l.
		būtamasis dažninis laikas	būt. d. l.
		būsimasis laikas	būs. l.
		būtamasis laikas	būt. l.
9	Laipsnis	nelyginamasis laipsnis	nelygin. l.
		aukštesnysis laipsnis	aukšt. l.
		aukščiausiasis laipsnis	aukšč. l.
10	Giminė	moteriškoji giminė	mot. g.
		vyriškoji giminė	vyr. g.
		bevardė giminė	bev. g.
		bendroji giminė	bendr. g.
11	Skaičius	vienaskaita	vns.
		daugiskaita	dgs.
		dviskaita	dvisk.
12	Linksnis	vardininkas	V.
		kilmininkas	K.
		naudininkas	N.
		galininkas	G.
		įnagininkas	Įn.
		vietininkas	Vt.
		šauksmininkas	Š.
		iliatyvas	Il.
13	Asmuo	pirmasis asmuo	1 asm.
		antrasis asmuo	2 asm.
		trečiasis asmuo	3 asm.
14	Teigiamumas, neigiamumas	teigiamas	teig.
		neigiamas	neig.
15	Romėniški skaičiai		rom. sk.
16	Stabiliosios frazės		idprl. (pvz., <i>be to</i>), idjngt. (pvz., <i>ir kt.</i>), idPS (p. s., taip žymimas junginio <i>post scriptum</i> sutrumpinimas)

8.2. priedas. Teksto anotavimo pavyzdys

Anotuotino teksto pavyzdys:

„Balandžio 21 dieną žadame atidaryti Gedimino kalną. Į jį bus galima pakilti istoriniu taku nuo Vilnios upės pusės“, – trečiadienį BNS sakė Nacionalinio muziejaus atstovams ryšiams su žiniasklaida ir visuomene Arminas Šileikis. Jo teigimu, tuo pačiu metu ant kalno bus vykdomi ir archeologiniai kasinėjimai, ir šlaito tvirtinimo darbai, tačiau jie netrukdyt lankytojams. Kalno teritorija vasaros sezono metu bus atvira nuo 7 val. iki 23 val., Gedimino bokštas – nuo 10 val. iki 21 valandos.

8.1 lentelė. Morfologijos paslauga

Operacijos vardas		Morfologija	
Operacijos paaiškinimas		Paduoto teksto suanotavimas morfologiškai	
Įeigos parametrai			
Nr.	Vardas	Duomenų tipas	Paaiškinimas
1	Tekstas	Tekstas	Tekstas kurį reikia suanotuoti
Grąžinami parametrai			
Nr.	Vardas	Duomenų tipas	Paaiškinimas
1	Rezultatas	JSON	Struktūruotai pateikia teksto morfologines anotacijas.

Morfologinio anotavimo rezultato JSON struktūros pavyzdys:

8.2 lentelė. Morfologinio nagrinėjimo reikšmės

Balandžio:	Atidaryti :
<pre>[["balandis", "Ncmsgn-"], N – daiktavardis, c – bendrinis, m – vyr. giminės, s – vienaskaita, g – kilmininkas, n – nesangražinis,</pre>	<pre>[["atidaryti", "Vgi-----n--n--"], V – veiksmažodis, g – bendras, i – bendratis, n- neigiamas, n – nesangražinė,</pre>
<pre>[["Balandis", "Npmsgns"], N – daiktavardis, p - tikrinis, m – vyr. giminės, s – vienaskaita, g – kilmininkas, n – nesangražinis,</pre>	<pre>[["atidaryti", "Vgps-pmpnnn-p"], V – veiksmažodis, g – bendras, p – dalyvis, s – būtasis laikas, p – daugiskaita, m – vyriška giminė, p – neveikiamoji rūšis, n – neigiamas, n – neįvardžiuotinis, n- vardininkas n – nesangražinė, p – nelyginamasis laipsnis,</pre>
<pre>[["Balandis", "Npmsgng"]] N – daiktavardis, p – tikrinis, m – vyr. giminės, s – vienaskaita, g – kilmininkas, n – nesangražinis,</pre>	<pre>[["atidaryti", "Vgps-pmpnvnvnp"]] V – veiksmažodis, g – bendras, p – dalyvis, s – būtasis laikas, p – daugiskaita, m – vyriška giminė, p – neveikiamoji rūšis, n – neigiamas, n – neįvardžiuotinis, v- šauksmininkas n – nesangražinė, p – nelyginamasis laipsnis,</pre>

```

{"msd": [{"/","Tu"}], [{"balandis", "Ncmsgn"}, {"Balandis", "Npmsgns"}, {"Balandis", "Npmsgng"}], [{"21", "M----d-"}], [{"diena", "Ncfsan-"}, {"žadėti", "Vgmp1p--n--ni-"}, {"atidaryti", "Vgi-----n--n--"}, {"atidaryti", "Vggs-pmpnnnn-p"}, {"atidaryti", "Vggs-pmpnvnvnp"}], [{"Gediminas", "Npmsgng"}, {"Gediminas", "Npmsgns"}, {"Gediminas", "Npmsgnf"}], [{"k alnas", "Ncmsan"}, {".", "Tp"}, {"i", "Sga"}, {"jis", "Pgmsan"}, {"būti", "Vgmf3---n--ni-"}, {"būti", "Vgmf3s--n--ni-"}, {"būti", "Vgmf3p--n--ni-"}, {"busti", "Vgmf3s--n--ni-"}, {"busti", "Vgmf3p--n--ni-"}, {"busti", "Vgmf3---n--ni-"}, {"galėti", "Vgpp--npnn-n-p"}, {"galimas", "Agpfsnn"}, {"galimas", "Agpn--n"}, {"galėti", "Vgpp-sfpnnnn-p"}, {"galimas", "Agpfsin"}, {"galėti", "Vgpp-sfpnnin"}, {"galimas", "Agpfsvn"}, {"galėti", "Vgppsfpnvnvnp"}], [{"pakilti", "Vgi-----n--n--"}, {"pakilti", "Vggs-pmpnnnnp"}, {"pakilti", "Vggs-pmpnvnvnp"}], [{"istorinis", "Agpmsin"}, {"takas", "Ncmsin-"}, {"takus", "Agpn--n"}, {"nuo", "Sgg"}, {"Vilnia", "Npfsngng"}, {"upė", "Ncfsng-"}, {"upė", "Ncftpnn-"}, {"upė", "Ncftpvn-"}, {"pusė", "Ncfsng-"}, {"pusė", "Ncftpnn-"}, {"pusė", "Ncftpvn"}, {"", "Tu"}, {"", "Tc"}, {"", "Th"}, {"trečiadienis", "Ncmsan-"}, {"BNS", "Ya"}, {"sakė", "Ncfsnn-"}, {"sakyti", "Vgma3p--n--ni-"}, {"sakyti", "Vgma3s--n--ni-"}, {"sakyti", "Vgma3---n--ni-"}, {"nacionalinis", "Agpmsgn"}, {"muziejus", "Ncmsgn-"}, {"atstovas", "Ncmpdn-"}, {"ryšys", "Ncmpdn-"}, {"su", "Sgi"}, {"žiniasklaida", "Ncfsin-"}, {"žiniasklaida", "Ncfsnn-"}, {"žiniasklaida", "Ncfsvn-"}, {"ir", "Cg"}, {"ir", "Qg"}, {"visuomenė", "Ncfsin-"}, {"visuomenė", "Ncfsvn-"}, {"Arminas", "Npmsnnf"}, {"Arminas", "Npmsnns"}, {"šileikis", "Npmsnns"}, {".", "Tp"}, {"jis", "Pgmsgn"}, {"teigimas", "Ncmsin-"}, {"", "Tc"}, {"tuo pačiu", "Rgp"}, {"metas", "Ncmsin-"}, {"mesti", "Vgmpls--n--ni-"}, {"ant", "Sgg"}, {"ant", "Qg"}, {"kalnas", "Ncmsgn-"}, {"būti", "Vgmf3p--n--ni-"}, {"būti", "Vgmf3s--n--ni-"}, {"būti", "Vgmf3p--n--ni-"}, {"busti", "Vgmf3s--n--ni-"}, {"busti", "Vgmf3p--n--ni-"}, {"busti", "Vgmf3---n--ni-"}, {"vykdyti", "Vgpp-pmpnnnn-p"}, {"vykdyti", "Vgpp-pmpnvnvnp"}], [{"ir", "Cg"}, {"ir", "Qg"}, {"archeologinis", "Agmpnn"}, {"archeologinis", "Agmpvn"}, {"kasinėjimas", "Ncmpnn-"}, {"kasinėjimas", "Ncmpvn-"}, {"", "Tc"}, {"ir", "Cg"}, {"ir", "Qg"}, {"šlaitas", "Ncmsgn-"}, {"tvirtinimas", "Ncmsgn-"}, {"darbas", "Ncmpnn-"}, {"darbas", "Ncmpvn-"}, {"", "Tc"}, {"tačiau", "Cg"}, {"jis", "Pgmpnn"}, {"netrukdyti", "Vgmf3p--y--ni-"}, {"netrukdyti", "Vgmf3s--y--ni-"}, {"netrukdyti", "Vgmf3---y--ni-"}, {"lankytojas", "Ncmpdn-"}, {".", "Tp"}, {"kalnas", "Ncmsgn"}, {"Kalnas", "Npmsgng"}, {"teritorija", "Ncfsnn-"}, {"teritorija", "Ncfsin-"}, {"teritorija", "Ncfsvn-"}, {"vasara", "Ncfsng-"}, {"vasara", "Ncftpnn-"}, {"vasaroti", "Vgmf3s--n--ni-"}, {"vasaroti", "Vgmf3p--n--ni-"}, {"vasaroti", "Vgmf3---n--ni-"}, {"vasara", "Ncftpvn-"}, {"sezonas", "Ncmsgn-"}, {"metas", "Ncmsin-"}, {"mesti", "Vgmpls--n--ni-"}, {"būti", "Vgmf3s--n--ni-"}, {"būti", "Vgmf3p--n--ni-"}, {"būti", "Vgmf3---n--ni-"}, {"busti", "Vgmf3s--n--ni-"}, {"busti", "Vgmf3p--n--ni-"}, {"busti", "Vgmf3---n--ni-"}, {"atviras", "Agpfsnn"}, {"atviras", "Agpfsin"}, {"atviras", "Agpn--n"}, {"atviras", "Agpfsvn"}, {"nuo", "Sgg"}, {"7", "M----d-"}, {"val.", "Ys"}, {"iki", "Sgg"}, {"iki", "Cg"}, {"iki", "Rgp"}, {"23", "M----d-"}, {"val.", "Ys"}, {"", "Tc"}, {"Gediminas", "Npmsgng"}, {"Gediminas", "Npmsgns"}, {"Gediminas", "Npmsgnf"}, {"bokštas", "Ncmsnn-"}, {"-", "Th"}, {"nuo", "Sgg"}, {"10", "M----d-"}, {"val.", "Ys"}, {"iki", "Sgg"}, {"iki", "Cg"}, {"iki", "Rgp"}, {"21", "M----d-"}, {"valanda", "Ncfsng-"}, {"valanda", "Ncftpnn-"}, {"valanda", "Ncftpvn-"}, {"", "Tp"}, {"stem": [{"/","Baland", "21", "dien", "žad", "atidar", "Gedimin", "kaln", ".", "i", "j", "bus", "galim", "pakilt", "istorin", "tak", "nes", "Viln", "up", "pus", "", "", "", "-"}, {"trečiadien", "BNS", "sak", "Nacionalin", "muziej", "atstov", "ryš", "su", "žiniasklaid", "ir", "visuomen", "Armin", "šileik", ".", "Jo", "teigim", "", "tuopat", "met", "ant", "kaln", "bus", "vykdom", "ir", "archeologin", "kasinėjim", "", "ir", "šlait", "tvirtini m", "darb", "", "tač", "jie", "netruk", "lankyt", "."}, {"Kaln", "teritorij", "vasar", "sezon", "met", "bus", "atvir", "nes", "7", "val.", "ik", "23", "val.", " ", "Gedimin", "bokšt", "-", "nes", "10", "val.", "ik", "21", "valand", "."}]}

```

8.3 lentelė. Segmentavimo paslauga

Operacijos vardas		Segmentavimas	
Operacijos paaiškinimas		Paduoto teksto segmentavimas	
Įėjigos parametrai			
Nr.	Vardas	Duomenų tipas	Paaiškinimas
1	Tekstas	Tekstas	Tekstas kurį reikia segmentuoti
Grąžinami parametrai			

Nr.	Vardas	Duomenų tipas	Paiškinimas
1	Rezultatas	JSON	Struktūruotai pateikia teksto segmentavimą.

```

    „ Balandžio 21 dieną žadame atidaryti Gedimino kalną .
{"seg": [[0,1],[1,9],[11,2],[14,5],[20,6],[27,9],[37,8],[46,5],[51,1],
↓ jį bus galima pakilti istoriniu taku
[53,1],[55,2],[58,3],[62,6],[69,7],[77,9],[87,4],
nuo Vilnios upės pusės „
[92,3],[96,7],[104,4],[109,5],[114,1],[115,1],
- trečiadienį BNS sakė Nacionalinio muziejaus atstovams
[120,1],[122,11],[134,3],[138,4],[143,12],[156,9],[166,9],
ryšiams su žiniasklaida ir visuomene Arminas Šileikis .
[176,7],[184,2],[187,12],[200,2],[203,9],[213,7],[221,8],[229,1],
Jo teigimu , tuo pačiu metu ant kalno bus
[233,2],[236,7],[243,1],[245,9],[255,4],[260,3],[264,5],[270,3],
vykdomi ir archeologiniai kasinėjimai , ir šlaito
[274,7],[282,2],[285,14],[300,11],[311,1],[313,2],[316,6],
tvirtinimo darbai , tačiau jie netrukdydys lankytojams .
[323,10],[334,6],[340,1],[342,6],[349,3],[353,9],[363,11],[374,1],
Kalno teritorija vasaros sezono metu bus atvira nuo
[376,5],[382,10],[393,7],[401,6],[408,4],[413,3],[417,6],[424,3],
7 val. iki 23 val. ,
[428,1],[430,4],[435,3],[439,2],[442,4],[446,1],
Gedimino bokštas - nuo 10 val. iki
[448,8],[457,7],[465,1],[467,3],[471,2],[474,4],[479,3],
21 valandos .
[483,2],[486,8],[494,1]],
"s": [[0,52],[53,63],[120,110],[233,142],[376,119]],
"p": [[0,116],[120,110],[233,262]]]

```

8.4 lentelė. Įvardintų esybių paslauga

Operacijos vardas		Įvardintų esybių atpažinimas	
Operacijos paaiškinimas		Paduoto teksto įvardintų esybių atpažinimas	
Įeigos parametrai			
Nr.	Vardas	Duomenų tipas	Paiškinimas
1	Tekstas	Tekstas	Tekstas kuriame reikia atpažinti įvardintas esybes
Grąžinami parametrai			
Nr.	Vardas	Duomenų tipas	Paiškinimas
1	Rezultatas	JSON	Struktūruotai pateikia teksto įvardintas esybes.

```

"ner": { "ne": { "money": []},
"dates": [
  {"text": "Balandžio 21 dieną", "ref": [ 1, 3 ]},
  {"text": "trečiadienį", "ref": [23, 1]},
  {"text": "nuo 7 val. iki 23 val.", "ref": [ 67,6]},
  {"text": "nuo 10 val. iki 21 valandos", "ref": [77,6]}],
"products": [],
"organizations": [
  {"text": "BNS", "ref": [24, 1]},
  {"text": "Nacionalinio muziejaus", "ref": [26,2]}],
"locations": [
  {"text": "Vilnios upės", "ref": [17,2]}],
"persons": [
  {"text": "Gedimino", "ref": [6, 1 ]},
  {"text": "Arminas Šileikis", "ref": [ 34,2]},
  {"text": "Gedimino", "ref": [ 74,1 ]}],

```

"associations": {"persons": "Gedimino"}},

8.5 lentelė. Anotuito teksto paslauga

Operacijos vardas		Teksto suanotavimas	
Operacijos paaiškinimas		Paduoto teksto suanotavimas	
Įeigos parametrai			
Nr.	Vardas	Duomenų tipas	Paaiškinimas
1	Tekstas	Tekstas	Tekstas kurį reikia suanotuoti.
Grąžinami parametrai			
Nr.	Vardas	Duomenų tipas	Paaiškinimas
1	Rezultatas		Struktūruotai pateikia suanotuotą tekstą.

„Balandžio 21 dieną žadame atidaryti Gedimino kalną. Į jį bus galima pakilti istoriniu taku nuo Vilnios upės pusės“ – trečiadienį BNS sakė Nacionalinio muziejaus atstovams ryšiams su žiniasklaida ir visuomene Arminas Šileikis. Jo teigimu, tuo pačiu metu ant kalno bus vykdomi ir archeologiniai kasinėjimai, ir šlaito tvirtinimo darbai, tačiau jie netrukdyt lankytojams. Kalno teritorija vasaros sezono metu bus atvira nuo 7 val. iki 23 val., Gedimino bokštas – nuo 10 val. iki 21 valandos.

DATA: Balandžio 21

DATA: trečiadienį

TURINYS: Balandžio 21 dieną žadame atidaryti Gedimino kalną. Į jį bus galima pakilti istoriniu taku nuo Vilnios upės pusės“, – trečiadienį BNS sakė

TURINYS: tuo pačiu metu ant kalno bus vykdomi ir archeologiniai kasinėjimai, ir šlaito tvirtinimo darbai, tačiau jie netrukdyt lankytojams.

ASMUO: Arminas Šileikis;

ASMUO: Jo

UŽIMAMA POZICIJA: Nacionalinio muziejaus atstovams ryšiams su žiniasklaida ir visuomene

ORGANIZACIJA: BNS

VIETOVĖ: Gedimino kalną

VIETOVĖ: Vilnios upės

LAIKAS: nuo 7 val. iki 23 val.

LAIKAS: nuo 10 val. iki 21 valandos

LAIKAS: vasaros sezono

OBJEKTAS: Gedimino bokštas

KALBĖJIMAS: sakė

KALBĖJIMAS: teigimu

8.3. priedas. Ontologijos dokumentų įverčių failas - „Evaluation_egz.“

```
<?xml version="1.0"?>
<rdf:RDF xmlns="http://www.w3.org/2002/07/owl#"
  xml:base="http://www.w3.org/2002/07/owl"
  xmlns:eval="http://semantika.lt/ns/Evaluation#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:terms="http://purl.org/dc/terms/"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:xml="http://www.w3.org/XML/1998/namespace"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
```

```

xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns:vann="http://purl.org/vocab/vann/"
xmlns:dc="http://purl.org/dc/elements/1.1/" >
<owl:Ontology>
  <imports rdf:resource="http://semantika.lt/ns/Evaluation"/>
  <imports rdf:resource="http://semantika.lt/ns/Politics"/>
</owl:Ontology>

<!--
////////////////////////////////////
//
// Individuals
//
////////////////////////////////////
-->

<NamedIndividual rdf:about="http://semantika.lt/ns/SemLT#document~25edb85e-7d42-4093-bb68-
b4c566135a4b~SemLT.document-1">
  <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.9</eval:fmeasure>
</NamedIndividual>
<NamedIndividual rdf:about="http://semantika.lt/ns/SemLT#document~62c84 added-27f9-44eb-95af-
6c34aa6e55db~SemLT.document-1">
  <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.9</eval:fmeasure>
</NamedIndividual>
  <NamedIndividual rdf:about="http://semantika.lt/ns/SemLT#document~38291e6f-a9ed-4b13-862f-
8646888dff4~SemLT.document-1">
  <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.7</eval:fmeasure>
</NamedIndividual>
  <NamedIndividual rdf:about="http://semantika.lt/ns/SemLT#document~dce0a47c-86b4-4782-9c36-
ad8a49bf7e73~SemLT.document-1">
  <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.8</eval:fmeasure>
</NamedIndividual>
  <NamedIndividual rdf:about="http://semantika.lt/ns/SemLT#document~e764a3cc-ee76-47ff-b01b-
d5f4c0fd3830~SemLT.document-1">
  <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.8</eval:fmeasure>
</NamedIndividual>
  <NamedIndividual rdf:about="http://semantika.lt/ns/SemLT#document~t764a3cc-ee76-47ff-b01b-
d5f4c0fd3830~SemLT.document-1">
  <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.5</eval:fmeasure>
</NamedIndividual>
  <NamedIndividual rdf:about="http://semantika.lt/ns/SemLT#document~550d8ba6-f0ad-4c47-9e1e-
26a24dced86~SemLT.document-1">
  <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.6</eval:fmeasure>
</NamedIndividual>
  <NamedIndividual rdf:about="http://semantika.lt/ns/SemLT#document~4e9fee93-61c1-4985-aec8-
cdd96d0874c2~SemLT.document-1">
  <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.5</eval:fmeasure>
</NamedIndividual>
  <NamedIndividual rdf:about="http://semantika.lt/ns/SemLT#document~2ae68bb5-8771-4a1e-bdca-
f3767c9e9742~SemLT.document-1">
  <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.9</eval:fmeasure>
</NamedIndividual>
  <NamedIndividual rdf:about="http://semantika.lt/ns/SemLT#document~12f3a565-d884-4391-93fa-
d0af35a3a112~SemLT.document-1">
  <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.7</eval:fmeasure>
</NamedIndividual>
  <NamedIndividual rdf:about="http://semantika.lt/ns/SemLT#document~6b3215b2-91c7-44fd-9cec-
e797501a3527~SemLT.document-1">
  <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.4</eval:fmeasure>
</NamedIndividual>

```



```

      <NamedIndividual      rdf:about="http://semantika.lt/ns/SemLT#document~8d9a77ed-38ec-4249-8ebc-
8f72c7b87bc1~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.8</eval:fmeasure>
    </NamedIndividual>
      <NamedIndividual      rdf:about="http://semantika.lt/ns/SemLT#document~e8e8a44b-aeb4-4baa-9eaa-
25bb3c93c672~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.6</eval:fmeasure>
    </NamedIndividual>
      <NamedIndividual      rdf:about="http://semantika.lt/ns/SemLT#document~5d7a945e-190c-4ac5-816b-
381e0d2610bd~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.5</eval:fmeasure>
    </NamedIndividual>
      <NamedIndividual      rdf:about="http://semantika.lt/ns/SemLT#document~44bce81f-ea79-4b04-9966-
e4ff0b567f03~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.9</eval:fmeasure>
    </NamedIndividual>
      <NamedIndividual      rdf:about="http://semantika.lt/ns/SemLT#document~f2feee9d-d65f-42c2-94c0-
242f007e690b~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.7</eval:fmeasure>
    </NamedIndividual>
      <NamedIndividual      rdf:about="http://semantika.lt/ns/SemLT#document~c11428ca-1afb-48af-9df0-
9482b7682039~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.4</eval:fmeasure>
    </NamedIndividual>
      <NamedIndividual      rdf:about="http://semantika.lt/ns/SemLT#document~c0084d31-068f-41a1-973a-
2e1a7071e940~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.8</eval:fmeasure>
    </NamedIndividual>
      <NamedIndividual      rdf:about="http://semantika.lt/ns/SemLT#document~671b266e-9502-40bf-879f-
2ac5ae4542e1~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.6</eval:fmeasure>
    </NamedIndividual>
      <NamedIndividual      rdf:about="http://semantika.lt/ns/SemLT#document~a44bf17f-1967-48da-b6e9-
a52d171ae946~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.4</eval:fmeasure>
    </NamedIndividual>
      <NamedIndividual      rdf:about="http://semantika.lt/ns/SemLT#document~a1a99a8e-767d-4e55-9391-
7eb8255cc812~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.9</eval:fmeasure>
    </NamedIndividual>
      <NamedIndividual      rdf:about="http://semantika.lt/ns/SemLT#document~9bdf7f-43a3-49ab-a3b6-
7f5c3502376d~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.6</eval:fmeasure>
    </NamedIndividual>
      <NamedIndividual      rdf:about="http://semantika.lt/ns/SemLT#document~2036aaab-6cd3-43bb-900e-
cdd5f454b8ff~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.8</eval:fmeasure>
    </NamedIndividual>
      <NamedIndividual      rdf:about="http://semantika.lt/ns/SemLT#document~88d81d4e-cbbd-4b4c-b634-
426620fc2e72~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.9</eval:fmeasure>
    </NamedIndividual>
      <NamedIndividual      rdf:about="http://semantika.lt/ns/SemLT#document~a9980747-96c8-4fbe-bb74-
96ff46b0dc84~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.9</eval:fmeasure>
    </NamedIndividual>
      <NamedIndividual      rdf:about="http://semantika.lt/ns/SemLT#document~503e85b9-d911-42e1-90cf-
1ac2bc342827~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.8</eval:fmeasure>
    </NamedIndividual>

```

```

        <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~13d0d4e5-62a0-49fd-a1fa-62bf04d96e28~SemLT.document-1">
        <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.7</eval:fmeasure>
    </NamedIndividual>
        <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~2e2abec4-fcec-40d9-9470-898c67f87b95~SemLT.document-1">
        <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.5</eval:fmeasure>
    </NamedIndividual>
        <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~bac0cd8b-5965-451a-ba19-fc4f34d7dc65~SemLT.document-1">
        <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.8</eval:fmeasure>
    </NamedIndividual>
        <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~a4ad4aac-fc9d-4275-90db-7c834a56de7e~SemLT.document-1">
        <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.8</eval:fmeasure>
    </NamedIndividual>
        <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~5df15aad-993f-4560-904d-21ccd94779f3~SemLT.document-1">
        <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.9</eval:fmeasure>
    </NamedIndividual>
        <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~fb1907f0-7f3b-4996-b622-e8df4e9097a2~SemLT.document-1">
        <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.6</eval:fmeasure>
    </NamedIndividual>
        <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~ac543eaa-7039-485d-aecb-6436baeabf1a~SemLT.document-1">
        <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.8</eval:fmeasure>
    </NamedIndividual>
        <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~4500204a-7c5d-48fa-9f3e-5496de8d4bb5~SemLT.document-1">
        <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.4</eval:fmeasure>
    </NamedIndividual>
        <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~f4924464-e77c-4eb5-9268-df1f05f21bb3~SemLT.document-1">
        <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.7</eval:fmeasure>
    </NamedIndividual>
        <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~8e13af19-392c-4d04-a001-ccf3ff59fd6d~SemLT.document-1">
        <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.8</eval:fmeasure>
    </NamedIndividual>
        <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~b5958952-eeb7-4a34-9fe3-56a83ab70158~SemLT.document-1">
        <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.5</eval:fmeasure>
    </NamedIndividual>
        <NamedIndividual          rdf:about="http://semantika.lt/ns/SSemLT#document~ab044eaf-d311-4ebd-ba53-0f0d87019542~SemLT.document-1">
        <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.9</eval:fmeasure>
    </NamedIndividual>
        <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~5f75fefd-2473-4939-b7b8-cc0d424d954d~SemLT.document-1">
        <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.8</eval:fmeasure>
    </NamedIndividual>
        <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~5051bc17-7e93-492d-819e-88ab3a0a5b30~SemLT.document-1">
        <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.6</eval:fmeasure>
    </NamedIndividual>
        <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~b69ba536-c163-4a9b-aec4-ec217803d0a~SemLT.document-1">
        <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.8</eval:fmeasure>
    </NamedIndividual>

```

```

    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~315ba038-4248-425e-a96d-
a6b98dd8255c~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.8</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~847bf2ba-d112-40ab-acf9-
ee144ee4edc9~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.4</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~b99c67e4-4856-4e47-ad1b-
608012fbd7fc~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.6</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~60b1573f-4d6d-47c9-b071-
15cc87eb7322~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.9</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~dca8e2ea-7c92-4973-9624-
b116a5756ef4~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.8</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~301bb6ed-e19b-4ae5-902c-
078c42a0e5bc~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.5</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~b44a0d54-7e0b-423a-8b83-
16dca524638d~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.7</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~f969b375-85cf-44ae-9694-
745e2f776835~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.8</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~56407cc1-4f36-4482-b55a-
2f1efe41bf85~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.5</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~e469b4d9-b9d9-4884-8d02-
7b5c8ea27a05~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.9</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~9fb5336c-45e6-49af-a091-
22667a8a4749~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.8</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~7ce335a5-4169-424f-914f-
b4fe42670f9a~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.4</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~57f5f607-727f-4592-9e76-
57489af68039~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.5</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~7df9ea6c-1d75-470b-9b3d-
d3d5ae739c3d~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.6</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~4c9b9f1f-b49d-41b4-b2b8-
f4c0e65d9591~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.7</eval:fmeasure>
    </NamedIndividual>

```

```

    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~25615966-dd29-4a03-b95c-
fc08d00db6d3~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.8</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~343b4cad-b6e0-4179-b5f8-
9c7c481db0ef~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.9</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~c6922cd1-d041-4cb9-a258-
a9fbf2a13143~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.7</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~2bfe3c50-f1aa-4498-98d1-
0da036e2c666~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.8</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~4652750a-ebf9-47e5-ae40-
8865cfb6b2ec~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.7</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~d2d51af8-13b4-4cec-8a6e-
350010d823c0~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.8</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~1ee1d069-af29-4034-87df-
b6db1ce6b74f~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.7</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~f915ce62-3460-4198-bf79-
fd3cbddb1551~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.5</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~005ad3b3-c305-44f4-83a6-
d6d0d4d6ac16~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.7</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~b20c28dc-a49c-4701-890e-
4df8ff802268~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.9</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~d073f0c7-52c3-45d4-9f5b-
252f6e8cc816~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.7</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~25dd52d4-c8cc-46fe-ad4f-
fab583aa76ee~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.8</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~43057e81-1dfd-47e7-95df-
cadd2a24035~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.6</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~2908e806-e7ee-4e0b-a6b0-
660385731090~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.6</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~bd619959-acbf-4667-82f7-
f3f097c23409~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.9</eval:fmeasure>
    </NamedIndividual>

```

```

    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~40fbe217-d60c-42bc-83df-
0f163d38c733~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.7</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~1ff37cdf-7273-4746-a9e5-
35f164ccc7cd~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.8</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~7606f7e9-a580-4437-9cad-
5ee3f617993a~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.8</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~b36c0505-90db-4f0d-b1df-
842b901fda8e~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.4</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~c6f56f8e-2c77-479f-9652-
6cf99fe065f3~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.5</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~fb752e50-c822-4bf7-944a-
beed0d2c3359~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.6</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~055088eb-68a9-4a21-aa66-
58654200343b~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.7</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~e48eae92-4300-408e-96ab-
0c9af475839f~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.8</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~d359ac96-e592-4a8f-bc73-
e701ac3c9449~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.8</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~ecc27bed-5170-4dab-b106-
20efb9632e4f~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.9</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~b8f3eff3-fae5-41f8-aea0-
58caec121275~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.6</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~d9f1acf1-d878-434c-8ec2-
a161fc15a73e~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.7</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~e21ddec5-d21e-498d-be5f-
8d866e9cbcf2~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.9</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~666f9631-a815-49df-acf0-
7dcb3ed1889c~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.8</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~dc1d7f28-a326-4a78-b4d1-
71dd8304ecda~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.7</eval:fmeasure>
    </NamedIndividual>

```

```

    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~a917c231-0009-478f-a936-049fa4dd1dd0~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.6</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~ca808cbc-1381-4ec7-b3a2-69f7889845e4~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.5</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~4cb5673b-ed4b-4373-b856-6e7767f8d7eb~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.4</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~8c0cde7f-1053-42bc-a3a5-3f403399417e~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.7</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~98717883-0803-4c8e-a0fd-830b7a8b8239~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.6</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~fda29506-942c-4790-ae45-d9223e63dd01~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.9</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~e9f9509f-24cc-4dff-ab7c-b474c6b43626~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.9</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~d3e5672b-1f11-4ae9-9c2f-9fbe030fbe2e~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.4</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~663f2799-6159-4c04-ae71-d0fde0669671~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.5</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~42324a6e-83e8-491d-a3e1-6a36d48f983a~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.9</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~af0aea34-0f97-43f5-b937-262b726ddd88~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.9</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~3202d977-2807-4330-998b-1bef71f1580b~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.8</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~a7afba84-147d-467d-81fd-b3d10f4dfe0e~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.7</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~6cbaf478-d282-4145-af7f-eddfcbe4cc4~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.6</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~877cd23a-d1e3-4b96-b96d-38eaa65996b3~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.7</eval:fmeasure>
    </NamedIndividual>

```

```

    <NamedIndividual          rdf:about="http://semantika.lt/ns/SSemLT#document~90910f05-64a1-4717-90ec-06825b1cf6a8~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.9</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~ab044eaf-d311-4ebd-ba53-0f0d87019542~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.9</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~t1a99a8e-767d-4e55-9391-7eb8255cc812~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.6</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~25edb85e-7d42-4093-bb68-b4c566135a4b~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.7</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~90910f05-64a1-4717-90ec-06825b1cf6a8~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.6</eval:fmeasure>
    </NamedIndividual>
    <NamedIndividual          rdf:about="http://semantika.lt/ns/SemLT#document~ab044eaf-d311-4ebd-ba53-0f0d87019542~SemLT.document-1">
      <eval:fmeasure rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.8</eval:fmeasure>
    </NamedIndividual>
  </rdf:RDF>

```

8.4. priedas. Užklauso

8.6 lentelė. Užklauso „Visi asmenys?“ duomenys

Užklauso sudėtingumas	Užklauso tipas	Visi asmenys?
Vieno tipo objektas	neišvedamas tikslumo įvertis	<pre> SELECT DISTINCT ?asmuo_label WHERE { ?asmuo_i rdf:type agents:person . ?asmuo_i rdfs:label ?asmuo_label . } </pre>
	išvedamas tikslumo įvertis individo atžvilgiu	<pre> SELECT DISTINCT ?asmuo_label ?tikslumo_ivertis WHERE { ?asmuo_i rdf:type agents:person . OPTIONAL { ?asmuo_i evaluation:fmeasure ?tikslumo_ivertis . } ?asmuo_i rdfs:label ?asmuo_label . } order by DESC (?tikslumo_ivertis) </pre>

8.7 lentelė. Užklauso „Visi asmenys?“ rezultatai

Užklauso pavadinimas	Užklauso sudėtingumas	Užklauso tipas	Imtys	Išvedamų įrašų skaičius	Vidurkis, ms	Paklaida, ms
VISI ASMENYS	Vieno tipo objektas	be tikslumo įverčio	25 imtis	294	19.4475	9.9721
			50 imtis	792	29.4733	14.3419
			100 imtis	1564	43.0941	12.8955
		tikslumo įvertis individo atžvilgiu	25 imtis	294	36.3516	16.7687
			50 imtis	792	77.4873	28.0378
			100 imtis	1564	162.8787	77.6477

8.8 lentelē. Uzsklausos „Visos vietovēs?” duomenys

Uzsklausos sudėtingumas	Uzsklausos tipas	Visi vietovės?
Vieno tipo objektas	neišvedamas tikslumo įvertis	SELECT DISTINCT ?vietove_label WHERE { ?vietove_i rdf:type locations:location . ?vietove_i rdfs:label ?vietove_label . }
	išvedamas tikslumo įvertis individo atžvilgiu	SELECT DISTINCT ?vietove_label ?tikslumo_ivertis WHERE { ?vietove_i rdf:type locations:location . OPTIONAL {?vietove_i evaluation:fmeasure ?tikslumo_ivertis . } ?vietove_i rdfs:label ?vietove_label . } order by DESC (?tikslumo_ivertis)

8.9 lentelē. Uzsklausos „Visos vietovės?” rezultatai

Uzsklausos pavadinimas	Uzsklausos sudėtingumas	Uzsklausos tipas	Imtys	Išvedamų įrašų skaičius	Vidurkis, ms	Paklaida, ms
VISOS VIETOVĖS	Vieno tipo objektas	be tikslumo įverčio	25 imtis	291	13.4079	5.8457
			50 imtis	504	21.9499	10.8477
			100 imtis	776	36.4816	36.4816
		tikslumo įvertis individo atžvilgiu	25 imtis	291	51.0117	11.1467
			50 imtis	504	113.5602	39.7000
			100 imtis	776	202.1810	22.7396

8.10 lentelē. Uzsklausos „Visos organizacijos?” duomenys

Uzsklausos sudėtingumas	Uzsklausos tipas	Visi organizacijos?
Vieno tipo objektas	neišvedamas tikslumo įvertis	SELECT DISTINCT ?organizacija_label WHERE { ?organizacija_i rdf:type agents:organization . ?organizacija_i rdfs:label ?organizacija_label . }
	išvedamas tikslumo įvertis individo atžvilgiu	SELECT DISTINCT ?organizacija_label ?tikslumo_ivertis WHERE { ?organizacija_i rdf:type agents:organization . OPTIONAL {?organizacija_i evaluation:fmeasure ?tikslumo_ivertis . } ?organizacija_i rdfs:label ?organizacija_label . } order by DESC (?tikslumo_ivertis)

8.11 lentelē. Uzsklausos „Visos organizacijos?” rezultatai

Uzsklausos pavadinimas	Uzsklausos sudėtingumas	Uzsklausos tipas	Imtys	Išvedamų įrašų skaičius	Vidurkis, ms	Paklaida, ms
VISOS ORGANIZACIJOS	Vieno tipo objektas	be tikslumo įverčio	25 imtis	170	11.2977	6.4312
			50 imtis	406	19.2807	10.6170
			100 imtis	664	33.4264	12.2048
		tikslumo įvertis individo atžvilgiu	25 imtis	170	35.4734	12.5113
			50 imtis	406	70.0253	18.3308
			100 imtis	664	113.0661	22.5586

8.12 lentelė. Užklauso „Kokie asmenys yra minimi?“ duomenys

Užklauso sudėtingumas	Užklauso tipas	Kokie asmenys yra minimi?
Vieno tipo objektas	neišvedamas tikslumo įvertis	SELECT DISTINCT ?asmuo_label WHERE { ?dokumentas semLT:refers_to__object ?asmuo_i . ?dokumentas rdf:type semLT:document. ?asmuo_i rdf:type agents:person . ?asmuo_i rdfs:label ?asmuo_label . }
	išvedamas tikslumo įvertis dokumento l atžvilgiu	SELECT DISTINCT ?asmuo_label ?tikslumo_įvertis WHERE { ?dokumentas semLT:refers_to__object ?asmuo_i . ?dokumentas rdf:type semLT:document. OPTIONAL {?dokumentas evaluation:fmeasure ?tikslumo_įvertis . } ?asmuo_i rdf:type agents:person . ?asmuo_i rdfs:label ?asmuo_label . } order by DESC (?tikslumo_įvertis)
	išvedamas tikslumo įvertis individo atžvilgiu	SELECT DISTINCT ?asmuo_label ?tikslumo_įvertis WHERE { ?dokumentas semLT:refers_to__object ?asmuo_i . ?dokumentas rdf:type semLT:document. OPTIONAL {?asmuo_i evaluation:fmeasure ?tikslumo_įvertis . } ?asmuo_i rdf:type agents:person . ?asmuo_i rdfs:label ?asmuo_label . } order by DESC (?tikslumo_įvertis)

8.13 lentelė. Užklauso „Kokie asmenys yra minimi?“ rezultatai

Užklauso pavadinimas	Užklauso sunkumas	Užklauso tipas	Imtis	Išvedamų įrašų skaičius	Vidurkis, ms	Paklaida, ms	
Kokie asmenys yra minimi?	Vieno tipo objektas	be tikslumo įverčio	25 imtis	294	11.8953	4.7820	
			50 imtis	792	30.5136	11.4743	
			100 imtis	1564	50.6535	13.8651	
		tikslumo įvertis dokumento atžvilgiu		25 imtis	294	32.2038	11.1928
				50 imtis	792	70.2120	14.6600
				100 imtis	1564	175.6578	73.8982
		tikslumo įvertis individo atžvilgiu		25 imtis	294	34.2910	14.3459
				50 imtis	792	83.6824	18.5298
				100 imtis	1564	175.6969	50.7378

8.14 lentelė. Užklauso „Kokios vietovės yra minimos?“ duomenys

Užklauso sudėtingumas	Užklauso tipas	Kokie vietovės yra minimos?
Vieno tipo objektas	neišvedamas tikslumo įvertis	SELECT DISTINCT ?vietove_label WHERE { ?dokumentas semLT:refers_to__object ?vietove . ?dokumentas rdf:type semLT:document. ?vietove_i rdf:type locations:location . ?vietove_i rdfs:label ?vietove_label . }
	išvedamas tikslumo įvertis dokumento atžvilgiu	SELECT DISTINCT ?vietove_label ?tikslumo_įvertis WHERE { ?dokumentas semLT:refers_to__object ?vietove_i . ?dokumentas rdf:type semLT:document. OPTIONAL {?dokumentas evaluation:fmeasure ?tikslumo_įvertis . } ?vietove_i rdf:type locations:location . ?vietove_i rdfs:label ?vietove_label . } order by DESC (?tikslumo_įvertis)

	išvedamas tikslumo įvertis individo atžvilgiu	<pre> SELECT DISTINCT ?vietove_label ?tikslumo_ivertis WHERE { ?dokumentas semLT:refers_to__object ?vietove_i . ?dokumentas rdf:type semLT:document. OPTIONAL {?vietove_i evaluation:fmeasure ?tikslumo_ivertis . } ?vietove_i rdf:type locations:location . ?vietove_i rdfs:label ?vietove_label . } order by DESC (?tikslumo_ivertis) </pre>
--	---	--

8.15 lentelė. Užklauso „Kokios vietovės yra minimos?“ rezultatai

Užklauso pavadinimas	Užklauso sunkumas	Užklauso tipas	Imtis	Išvedamų įrašų skaičius	Vidurkis, ms	Paklaida, ms
Kokios vietovės yra minimos?“	Vieno tipo objektas	be tikslumo įverčio	25 imtis	545	4294.6299	360.70882
			50 imtis	1080	44267.5486	1490.4629
			100 imtis	1773	250328.79	2117.0435
		tikslumo įvertis dokumento atžvilgiu	545	545	50.4518	16.0553
			1080	1080	108.3164	108.3164
			1773	1773	212.4426	41.01255
			25 imtis	545	51.2521	15.7617
			50 imtis	1080	114.8147	32.74612
			100 imtis	1773	237.526	60.2502

8.16 lentelė. Užklauso „Kokios organizacijos yra minimos?“ duomenys

Užklauso sudėtingumas	Užklauso tipas	Kokios organizacijos yra minimos?
Vieno tipo objektas	neišvedamas tikslumo įvertis	<pre> SELECT DISTINCT ?organizacija_label WHERE { ?dokumentas semLT:refers_to__object ?organizacija . ?dokumentas rdf:type semLT:document. ?organizacija_i rdf:type agents:organization . ?organizacija_i rdfs:label ?organizacija_label . } </pre>
	išvedamas tikslumo įvertis dokumento atžvilgiu	<pre> SELECT DISTINCT ?organizacija_label ?tikslumo_ivertis WHERE { ?dokumentas semLT:refers_to__object ?organizacija_i . ?dokumentas rdf:type semLT:document. OPTIONAL {?dokumentas evaluation:fmeasure ?tikslumo_ivertis . } ?organizacija_i rdf:type agents:organization . ?organizacija_i rdfs:label ?organizacija_label . } order by DESC (?tikslumo_ivertis) </pre>
	išvedamas tikslumo įvertis individo atžvilgiu	<pre> SELECT DISTINCT ?organizacija_label ?tikslumo_ivertis WHERE { ?dokumentas semLT:refers_to__object ?organizacija_i . ?dokumentas rdf:type semLT:document. OPTIONAL {?organizacija_i evaluation:fmeasure ?tikslumo_ivertis . } ?organizacija_i rdf:type agents:organization . ?organizacija_i rdfs:label ?organizacija_label . } order by DESC (?tikslumo_ivertis) </pre>

8.17 lentelė. Užklauso „Kokios organizacijos yra minimos?“ rezultatai

Užklauso pavadinimas	Užklauso sunkumas	Užklauso tipas	Imtis	Išvedamų įrašų skaičius	Vidurkis, ms	Paklaida, ms
Kokios organizacijos yra minimos?“	Vieno tipo objektas	be tikslumo įverčio	25 imtis	170	2692.2056	436.19286
			50 imtis	405	27929.470	2727.5629
			100 imtis	662	129325.95	5139.1782
		tikslumo įvertis dokumento atžvilgiu	25 imtis	170	32.1108	10.7416
			50 imtis	405	70.7822	20.5214
			100 imtis	662	122.0153	36.62702
			25 imtis	170	35.2017	15.9054

		tikslumo įvertis	50 imtis	405	72.1958	17.8719
		individo atžvilgiu	100 imtis	662	117.8232	29.6473

8.18 lentelė. Užklausa „Ar yra koks nors kainų pokytis?“ duomenys

Užklauso sudėtingumas	Užklauso tipas	Ar yra koks nors kainų pokytis?
Dviejų skirtingų tipų objektai	neišvedamas tikslumo įvertis	<pre>SELECT DISTINCT ?kaina_label ?pokytis_label WHERE { ?dokumentas semLT:refers_to__object ?kaina_i . ?dokumentas rdf:type semLT:document. ?dokumentas semLT:refers_to__object ?pokytis_i . ?pokytis_i economyAndBusiness:has__abstract_object ?kaina_i . ?kaina_i rdf:type abstractObject:abstract_object . ?pokytis_i rdf:type economyAndBusiness:change . ?kaina_i rdfs:label ?kaina_label . ?pokytis_i rdfs:label ?pokytis_label . }</pre>
	išvedamas tikslumo įvertis dokumento atžvilgiu	<pre>SELECT DISTINCT ?kaina_label ?pokytis_label ?tikslumo_ivertis WHERE { ?dokumentas semLT:refers_to__object ?kaina_i . ?dokumentas rdf:type semLT:document. OPTIONAL {?dokumentas evaluation:fmeasure ?tikslumo_ivertis . } ?dokumentas semLT:refers_to__object ?pokytis_i . ?pokytis_i economyAndBusiness:has__abstract_object ?kaina_i . ?kaina_i rdf:type abstractObject:abstract_object . ?pokytis_i rdf:type economyAndBusiness:change . ?kaina_i rdfs:label ?kaina_label . ?pokytis_i rdfs:label ?pokytis_label . } order by DESC (?tikslumo_ivertis)</pre>
	išvedamas tikslumo įvertis individo atžvilgiu	<pre>SELECT DISTINCT ?kaina_label ?pokytis_label ?tikslumo_ivertis ?kainos_ivertis ?pokycio_ivertis WHERE { ?dokumentas semLT:refers_to__object ?kaina_i . ?dokumentas rdf:type semLT:document. OPTIONAL {?kaina_i evaluation:fmeasure ?kainos_ivertis . } OPTIONAL {?pokytis_i evaluation:fmeasure ?pokycio_ivertis . } ?dokumentas semLT:refers_to__object ?pokytis_i . ?pokytis_i economyAndBusiness:has__abstract_object ?kaina_i . ?kaina_i rdf:type abstractObject:abstract_object . BIND ((?kainos_ivertis*?pokycio_ivertis) AS ?tikslumo_ivertis) ?kaina_i rdfs:label ?kaina_label . ?pokytis_i rdfs:label ?pokytis_label . } order by DESC (?tikslumo_ivertis)</pre>

8.19 lentelė. Užklausa „Ar yra koks nors kainų pokytis?“ rezultatai

Užklauso pavadinimas	Užklauso sunkumas	Užklauso tipas	Imtys	Išvedamų įrašų skaičius	Vidurkis, ms	Paklaida, ms
Ar yra koks nors kainų pokytis?	Dviejų skirtingų tipų objektai	be tikslumo įverčio	25 imtis	23	11.2908	6.4971
			50 imtis	76	13.471	6.007
			100 imtis	148	27.6027	12.3196
		tikslumo įvertis dokumento atžvilgiu	25 imtis	23	17.4841	8.40187
			50 imtis	76	22.7329	8.33054
			100 imtis	148	27.8696	27.8696
		tikslumo įvertis individo atžvilgiu	25 imtis	23	20.7796	8.67675
			50 imtis	76	31.3011	11.3755
			100 imtis	148	42.1411	17.3402

8.20 lentelė. Užklauso „Ar padidėjo kainos?“ duomenys

Užklauso sudėtingumas	Užklauso tipas	Ar padidėjo kainos?
Dviejų skirtingų tipų objektai	neišvedamas tikslumo įvertis	<pre>SELECT DISTINCT ?kaina_label ?pokytis_label WHERE { ?dokumentas semLT:refers_to__object ?kaina_i . ?dokumentas rdf:type semLT:document. ?dokumentas semLT:refers_to__object ?pokytis_i . ?pokytis_i economyAndBusiness:has__abstract_object ?kaina_i. ?kaina_i rdf:type abstractObject:abstract_object . ?pokytis_i rdf:type economyAndBusiness:change . ?pokytis_i economyAndBusiness:has__change_type economyAndBusiness:increase_change_type. ?kaina_i rdfs:label ?kaina_label . ?pokytis_i rdfs:label ?pokytis_label . }</pre>
	išvedamas tikslumo įvertis dokumento atžvilgiu	<pre>SELECT DISTINCT ?kaina_label ?pokytis_label ?tikslumo_ivertis WHERE { ?dokumentas semLT:refers_to__object ?kaina_i . ?dokumentas rdf:type semLT:document. OPTIONAL {?dokumentas evaluation:fmeasure ?tikslumo_ivertis . } ?dokumentas semLT:refers_to__object ?pokytis_i . ?pokytis_i economyAndBusiness:has__abstract_object ?kaina_i. ?kaina_i rdf:type abstractObject:abstract_object . ?pokytis_i rdf:type economyAndBusiness:change . ?pokytis_i economyAndBusiness:has__change_type economyAndBusiness:increase_change_type. ?kaina_i rdfs:label ?kaina_label . ?pokytis_i rdfs:label ?pokytis_label . } order by DESC (?tikslumo_ivertis)</pre>
	išvedamas tikslumo įvertis individo atžvilgiu	<pre>SELECT DISTINCT ?kaina_label ?pokytis_label ?tikslumo_ivertis ?kainos_ivertis ?pokycio_ivertis WHERE { ?dokumentas semLT:refers_to__object ?kaina_i . ?dokumentas rdf:type semLT:document. OPTIONAL {?kaina_i evaluation:fmeasure ?k . } OPTIONAL {?pokytis_i evaluation:fmeasure ?c . } ?dokumentas semLT:refers_to__object ?pokytis_i . ?pokytis_i economyAndBusiness:has__abstract_object ?kaina_i. ?kaina_i rdf:type abstractObject:abstract_object . ?pokytis_i rdf:type economyAndBusiness:change . BIND ((?kainos_ivertis*?pokycio_ivertis) AS ?tikslumo_ivertis) ?pokytis_i economyAndBusiness:has__change_type economyAndBusiness:increase_change_type. ?kaina_i rdfs:label ?kaina_label . ?pokytis_i rdfs:label ?pokytis_label . } order by DESC (?tikslumo_ivertis)</pre>

8.21 lentelė. Užklauso „Ar padidėjo kainos?“ rezultatai

Užklauso pavadinimas	Užklauso sunkumas	Užklauso tipas	Imtis	Išvedamų įrašų skaičius	Vidurkis, ms	Paklaida, ms	
Ar padidėjo kainos?	Dviejų skirtingų tipų objektai	be tikslumo įverčio	25 imtis	19	14.701	15.5555	
			50 imtis	56	11.2301	6.74194	
			100 imtis	126	19.3923	10.7598	
		tikslumo įvertis dokumento atžvilgiu		25 imtis	19	11.8994	8.05411
				50 imtis	56	22.6351	9.71355
				100 imtis	126	37.6376	14.8016
		tikslumo įvertis individo atžvilgiu		25 imtis	19	16.4935	7.42563
				50 imtis	56	24.167	11.9737
				100 imtis	126	44.279	44.279

8.22 lentelė. Užklauso „Ar sumažėjo kainos?“ duomenys

Užklauso sudėtingumas	Užklauso tipas	Ar sumažėjo kainos?
Dviejų skirtingų tipų objektai	neišvedamas tikslumo įvertis	<pre>SELECT DISTINCT ?kaina_label ?pokytis_label WHERE { ?dokumentas semLT:refers_to__object ?kaina_i . ?dokumentas rdf:type semLT:document. ?dokumentas semLT:refers_to__object ?pokytis_i . ?pokytis_i economyAndBusiness:has__abstract_object ?kaina_i. ?kaina_i rdf:type abstractObject:abstract_object . ?pokytis_i rdf:type economyAndBusiness:change . ?pokytis_i economyAndBusiness:has__change_type economyAndBusiness:decrease_change_type. ?kaina_i rdfs:label ?kaina_label . ?pokytis_i rdfs:label ?pokytis_label . }</pre>
	išvedamas tikslumo įvertis dokumento atžvilgiu	<pre>SELECT DISTINCT ?kaina_label ?pokytis_label ?tikslumo_ivertis WHERE { ?dokumentas semLT:refers_to__object ?kaina_i . ?dokumentas rdf:type semLT:document. OPTIONAL {?dokumentas evaluation:fmeasure ?tikslumo_ivertis . } ?dokumentas semLT:refers_to__object ?pokytis_i . ?pokytis_i economyAndBusiness:has__abstract_object ?kaina_i. ?kaina_i rdf:type abstractObject:abstract_object . ?pokytis_i rdf:type economyAndBusiness:change . ?pokytis_i economyAndBusiness:has__change_type economyAndBusiness:decrease_change_type. ?kaina_i rdfs:label ?kaina_label . ?pokytis_i rdfs:label ?pokytis_label . } order by DESC (?tikslumo_ivertis)</pre>
	išvedamas tikslumo įvertis individo atžvilgiu	<pre>SELECT DISTINCT ?kaina_label ?pokytis_label ?tikslumo_ivertis ?kainos_ivertis ?pokycio_ivertis WHERE { ?dokumentas semLT:refers_to__object ?kaina_i . ?dokumentas rdf:type semLT:document. OPTIONAL {?kaina_i evaluation:fmeasure ?k . } OPTIONAL {?pokytis_i evaluation:fmeasure ?c . } ?dokumentas semLT:refers_to__object ?pokytis_i . ?pokytis_i economyAndBusiness:has__abstract_object ?kaina_i. ?kaina_i rdf:type abstractObject:abstract_object . ?pokytis_i rdf:type economyAndBusiness:change . BIND ((?kainos_ivertis*?pokycio_ivertis) AS ?tikslumo_ivertis) ?pokytis_i economyAndBusiness:has__change_type economyAndBusiness:decrease_change_type. ?kaina_i rdfs:label ?kaina_label . ?pokytis_i rdfs:label ?pokytis_label . } order by DESC (?tikslumo_ivertis)</pre>

8.23 lentelė. Užklauso „Ar sumažėjo kainos?“ rezultatai

Užklauso pavadinimas	Užklauso sunkumas	Užklauso tipas	Imtis	Išvedamų įrašų skaičius	Vidurkis, ms	Paklaida, ms	
Ar sumažėjo kainos?	Dviejų skirtingų tipų objektai	be tikslumo įverčio	25 imtis	16	9.6827	4.48077	
			50 imtis	51	12.2115	5.8861	
			100 imtis	105	13.6555	7.5193	
		tikslumo įvertis dokumento atžvilgiu		25 imtis	16	12.4795	6.11371
				50 imtis	51	15.6599	7.21993
				100 imtis	105	25.3708	12.0837
		tikslumo įvertis individo atžvilgiu		25 imtis	16	13.0633	7.48621
				50 imtis	51	21.4595	9.97848
				100 imtis	105	28.7440	9.81845

8.24 lentelė. Užklauso „Ką kalbėjo asmenys?“ duomenys

Užklauso sudėtingumas	Užklauso tipas	Ką kalbėjo asmenys?
Dviejų skirtingų tipų objektai	neišvedamas tikslumo įvertis	<pre>SELECT DISTINCT ?asmuo_label ?turinys_label WHERE { ?dokumentas semLT:refers_to__object ?asmuo_i . ?dokumentas rdf:type semLT:document. ?dokumentas semLT:refers_to__object ?kalbėjimas_i . ?asmuo_i events:talked__talking ?kalbėjimas_i . ?asmuo_i rdf:type agents:person . ?kalbėjimas_i rdf:type politics:talking . ?dokumentas semLT:refers_to__object ?turinys_i . ?kalbėjimas_i events:talked__substance ?turinys_i . ?turinys_i rdf:type events:substance . ?asmuo_i rdfs:label ?asmuo_label . ?turinys_i rdfs:label ?turinys_label }</pre>
	išvedamas tikslumo įvertis dokumento atžvilgiu	<pre>SELECT DISTINCT ?asmuo_label ?turinys_label ?tikslumo_ivertis WHERE { ?dokumentas semLT:refers_to__object ?asmuo_i . ?dokumentas rdf:type semLT:document. OPTIONAL {?dokumentas evaluation:fmeasure ?tikslumo_ivertis . } ?dokumentas semLT:refers_to__object ?kalbėjimas_i . ?asmuo_i events:talked__talking ?kalbėjimas_i . ?asmuo_i rdf:type agents:person . ?kalbėjimas_i rdf:type politics:talking . ?dokumentas semLT:refers_to__object ?turinys_i . ?kalbėjimas_i events:talked__substance ?turinys_i . ?turinys_i rdf:type events:substance . ?asmuo_i rdfs:label ?asmuo_label . ?turinys_i rdfs:label ?turinys_label } order by DESC (?tikslumo_ivertis)</pre>
	išvedamas tikslumo įvertis individo atžvilgiu	<pre>SELECT DISTINCT ?asmuo_label ?turinys_label ?tikslumo_ivertis ?asmens_tikslumas ?tutinio_tikslumas WHERE { ?dokumentas semLT:refers_to__object ?asmuo_i . ?dokumentas rdf:type semLT:document. OPTIONAL {?asmuo_i evaluation:fmeasure ?asmens_tikslumas . } OPTIONAL {?turinys_i evaluation:fmeasure ?tutinio_tikslumas . } ?dokumentas semLT:refers_to__object ?kalbėjimas_i . ?asmuo_i events:talked__talking ?kalbėjimas_i . ?asmuo_i rdf:type agents:person . ?kalbėjimas_i rdf:type politics:talking . ?dokumentas semLT:refers_to__object ?turinys_i . BIND ((?asmens_tikslumas*?tutinio_tikslumas) AS ?tikslumo_ivertis) ?kalbėjimas_i events:talked__substance ?turinys_i . ?turinys_i rdf:type events:substance . ?asmuo_i rdfs:label ?asmuo_label . ?turinys_i rdfs:label ?turinys_label } order by DESC (?tikslumo_ivertis)</pre>

8.25 lentelė. Užklauso „Ką kalbėjo asmenys?“ rezultatai

Užklauso pavadinimas	užklauso sunkumas	užklauso tipas	imtys	Išvedamų įrašų skaičius	Vidurkis, ms	Paklaida, ms
Ką kalbėjo asmenys	dviejų skirtingų tipų objektai	be tikslumo įverčio	25 imtis	69	18.8370	8.8633
			50 imtis	278	40.4862	16.1472
			100 imtis	684	74.6463	16.5418
		tikslumo įvertis dokumento atžvilgiu	25 imtis	69	27.8449	13.9733
			50 imtis	278	51.7176	15.2986
			100 imtis	684	100.6294	27.7176
tikslumo įvertis individo atžvilgiu	25 imtis	69	784.9702	199.4744		
	50 imtis	278	2569.1559	615.2406		

		100 imtis	684	6593.0745	1285.9527
--	--	-----------	-----	-----------	-----------

8.26 lentelė. Užklauso „Kokios pareigos užimamos organizacijose?“ duomenys

Užklauso sudėtingumas	Užklauso tipas	Kokios pareigos užimamos organizacijose?
Dviejų skirtingų tipų objektai	neišvedamas tikslumo įvertis	<pre> SELECT DISTINCT ?organizacija_label ?pozicija_label WHERE { ?dokumentas semLT:refers_to__object ?organizacija_i . ?dokumentas rdf:type semLT:document. ?dokumentas semLT:refers_to__object ?einamos_pareigos_i . ?organizacija_i events:provide__position_held ?einamos_pareigos_i . ?organizacija_i rdf:type agents:organization . ?einamos_pareigos_i rdf:type events:position_held. ?einamos_pareigos_i events:assign__position ?pozicija_i . ?pozicija_i rdf:type agents:position . ?pozicija_i rdfs:label ?pozicija_label . ?organizacija_i rdfs:label ?organizacija_label . } </pre>
	išvedamas tikslumo įvertis dokumento atžvilgiu	<pre> SELECT DISTINCT ?organizacija_label ?pozicija_label ?tikslumo_ivertis WHERE { ?dokumentas semLT:refers_to__object ?organizacija_i . ?dokumentas rdf:type semLT:document. OPTIONAL {?dokumentas evaluation:fmeasure ?tikslumo_ivertis . } ?dokumentas semLT:refers_to__object ?einamos_pareigos_i . ?organizacija_i events:provide__position_held ?einamos_pareigos_i . ?organizacija_i rdf:type agents:organization . ?einamos_pareigos_i rdf:type events:position_held. ?einamos_pareigos_i events:assign__position ?pozicija_i . ?pozicija_i rdf:type agents:position . ?pozicija_i rdfs:label ?pozicija_label . ?organizacija_i rdfs:label ?organizacija_label . } order by DESC (?tikslumo_ivertis) </pre>
	išvedamas tikslumo įvertis individo atžvilgiu	<pre> SELECT DISTINCT ?organizacija_label ?pozicija_label ?tikslumo_ivertis ?organizacijos_tikslumas ?pozicijos_tikslumas WHERE { ?dokumentas semLT:refers_to__object ?organizacija_i . ?dokumentas rdf:type semLT:document. OPTIONAL {?organizacija_i evaluation:fmeasure ?organizacijos_tikslumas . } OPTIONAL {?pozicija_i evaluation:fmeasure ?pozicijos_tikslumas . } ?dokumentas semLT:refers_to__object ?einamos_pareigos_i . ?organizacija_i events:provide__position_held ?einamos_pareigos_i . ?organizacija_i rdf:type agents:organization . ?einamos_pareigos_i rdf:type events:position_held. ?einamos_pareigos_i events:assign__position ?pozicija_i . BIND ((?organizacijos_tikslumas*?pozicijos_tikslumas) AS ?tikslumo_ivertis) ?pozicija_i rdf:type agents:position . ?pozicija_i rdfs:label ?pozicija_label . ?organizacija_i rdfs:label ?organizacija_label . } order by DESC (?tikslumo_ivertis) </pre>

8.27 lentelė. Užklauso „Kokios pareigos užimamos organizacijose?“ rezultatai

Užklauso pavadinimas	Užklauso sunkumas	Užklauso tipas	Imtys	Išvedamų rašų skaičius	Vidurkis, ms	Paklaida, ms
Kokios pareigos užimamos organizacijose?	Dviejų skirtingų tipų objektai	be tikslumo įverčio	25 imtis	7	6.7237	3.4166
			50 imtis	24	7.7336	4.1149
			100 imtis	50	11.0104	5.8014
		tikslumo įvertis dokumento atžvilgiu	25 imtis	7	7.569	4.0067
			50 imtis	24	9.7604	4.9906

		100 imtis	50	17.8518	8.5653
		25 imtis	7	11.8164	8.0835
	tikslumo įvertis individo atžvilgiu	50 imtis	24	13.8509	6.0264
		100 imtis	50	18.1399	5.0499

8.28 lentelė. Užklauso „Kokias pareigas užima asmenys?“ duomenys

Užklauso sudėtingumas	Užklauso tipas	Kokias pareigas užima asmenys?
Dviejų skirtingų tipų objektai	neišvedamas tikslumo įvertis	<pre>SELECT DISTINCT ?asmuo_label ?pozicija_label WHERE { ?dokumentas semLT:refers_to__object ?einamos_pareigos_i. ?dokumentas rdf:type semLT:document. ?dokumentas semLT:refers_to__object ?asmuo_i. ?einamos_pareigos_i events:is_hold_by__person ?asmuo_i . ?einamos_pareigos_i rdf:type events:position_held . ?asmuo_i rdf:type agents:person . ?dokumentas semLT:refers_to__object ?pozicija_i. ?einamos_pareigos_i events:assign__position ?pozicija_i . ?pozicija_i rdf:type agents:position. ?pozicija_i rdfs:label ?pozicija_label . ?asmuo_i rdfs:label ?asmuo_label . }</pre>
	išvedamas tikslumo įvertis dokumento atžvilgiu	<pre>SELECT DISTINCT ?asmuo_label ?pozicija_label ?tikslumo_ivertis WHERE { ?dokumentas semLT:refers_to__object ?einamos_pareigos_i. ?dokumentas rdf:type semLT:document. OPTIONAL {?dokumentas evaluation:fmeasure ?tikslumo_ivertis . } ?dokumentas semLT:refers_to__object ?asmuo_i. ?einamos_pareigos_i events:is_hold_by__person ?asmuo_i . ?einamos_pareigos_i rdf:type events:position_held . ?asmuo_i rdf:type agents:person . ?dokumentas semLT:refers_to__object ?pozicija_i. ?einamos_pareigos_i events:assign__position ?pozicija_i . ?pozicija_i rdf:type agents:position. ?pozicija_i rdfs:label ?pozicija_label . ?asmuo_i rdfs:label ?asmuo_label . } order by DESC (?tikslumo_ivertis)</pre>
	išvedamas tikslumo įvertis individo atžvilgiu	<pre>SELECT DISTINCT ?asmuo_label ?pozicija_label ?tikslumo_ivertis ?asmens_tikslumas ?pozicijos_tikslumas WHERE { ?dokumentas semLT:refers_to__object ?einamos_pareigos_i. ?dokumentas rdf:type semLT:document. OPTIONAL {?asmuo_i evaluation:fmeasure ?asmens_tikslumas . } OPTIONAL {?pozicija_i evaluation:fmeasure ?pozicijos_tikslumas . } ?dokumentas semLT:refers_to__object ?asmuo_i. ?einamos_pareigos_i events:is_hold_by__person ?asmuo_i . ?einamos_pareigos_i rdf:type events:position_held . ?asmuo_i rdf:type agents:person . ?dokumentas semLT:refers_to__object ?pozicija_i. ?einamos_pareigos_i events:assign__position ?pozicija_i . BIND ((?asmens_tikslumas*?pozicijos_tikslumas) AS ?tikslumo_ivertis) ?pozicija_i rdf:type agents:position. ?pozicija_i rdfs:label ?pozicija_label . ?asmuo_i rdfs:label ?asmuo_label . } order by DESC (?tikslumo_ivertis)</pre>

8.29 lentelė. Užklauso „Kokias pareigas užima asmenys?“ rezultatai

Užklauso pavadinimas	Užklauso sunkumas	Užklauso tipas	Imtys	Išvedamų įrašų skaičius	Vidurkis, ms	Paklaida, ms
Kokias pareigas užima asmenys?	Dviejų skirtingų tipų objektai	be tikslumo įverčio	25 imtis	95	12.5502	6.8364
			50 imtis	251	22.3299	11.9536
			100 imtis	467	32.633	10.5275

	tikslumo įvertis dokumento atžvilgiu	25 imtis	95	20.0869	9.6972
		50 imtis	251	36.9395	12.3834
		100 imtis	467	61.6308	14.8047
	tikslumo įvertis individo atžvilgiu	25 imtis	95	22.744	10.2284
		50 imtis	251	49.8156	10.3478
		100 imtis	467	112.2495	53.1721

8.30 lentelė. Užklauso „Ką kalbėjo asmenys minėdami „Vokietija“?“ duomenys

Užklauso sudėtingumas	Užklauso tipas	Ką kalbėjo asmenys minėdami "Vokietija"?
Dviejų skirtingų tipų objektai + konkretus objektas	neišvedamas tikslumo įvertis	<pre> SELECT DISTINCT ?asmuo_label ?turinys_label WHERE { ?dokumentas semLT:refers_to__object ?asmuo_i . ?dokumentas rdf:type semLT:document. ?dokumentas semLT:refers_to__object ?kalbėjimas_i . ?asmuo_i events:talked__talking ?kalbėjimas_i . ?asmuo_i rdf:type agents:person . ?kalbėjimas_i rdf:type politics:talking . ?dokumentas semLT:refers_to__object ?turinys_i . ?kalbėjimas_i events:talked__substance ?turinys_i . ?turinys_i rdf:type events:substance . ?asmuo_i rdfs:label ?asmuo_label . ?turinys_i rdfs:label ?turinys_label. ?salis ann:label_lemma "vokietija"@lt . ?salis rdf:type locations:location . ?kalbėjimas_i events:included__object ?salis. } </pre>
	išvedamas tikslumo įvertis dokumento atžvilgiu	<pre> SELECT DISTINCT ?asmuo_label ?turinys_label ?tikslumo_ivertis WHERE { ?dokumentas semLT:refers_to__object ?asmuo_i . ?dokumentas rdf:type semLT:document. OPTIONAL {?dokumentas evaluation:fmeasure ?tikslumo_ivertis . } ?dokumentas semLT:refers_to__object ?kalbėjimas_i . ?asmuo_i events:talked__talking ?kalbėjimas_i . ?asmuo_i rdf:type agents:person . ?kalbėjimas_i rdf:type politics:talking . ?dokumentas semLT:refers_to__object ?turinys_i . ?kalbėjimas_i events:talked__substance ?turinys_i . ?turinys_i rdf:type events:substance . ?asmuo_i rdfs:label ?asmuo_label . ?turinys_i rdfs:label ?turinys_label. ?salis ann:label_lemma "vokietija"@lt . ?salis rdf:type locations:location . ?kalbėjimas_i events:included__object ?salis. } order by DESC (?tikslumo_ivertis) </pre>
	išvedamas tikslumo įvertis individo atžvilgiu	<pre> SELECT DISTINCT ?asmuo_label ?turinys_label ?tikslumo_ivertis ?asmens_tikslumas ?kalbejimo_tikslumas ?turinio_tikslumas WHERE { ?dokumentas semLT:refers_to__object ?asmuo_i . ?dokumentas rdf:type semLT:document. OPTIONAL {?asmuo_i evaluation:fmeasure ?asmens_tikslumas . } OPTIONAL {?kalbėjimas_i evaluation:fmeasure ?kalbejimo_tikslumas . } OPTIONAL {?turinys_i evaluation:fmeasure ?turinio_tikslumas . } ?dokumentas semLT:refers_to__object ?kalbėjimas_i . ?asmuo_i events:talked__talking ?kalbėjimas_i . ?asmuo_i rdf:type agents:person . ?kalbėjimas_i rdf:type politics:talking . ?dokumentas semLT:refers_to__object ?turinys_i . BIND ((?asmens_tikslumas*?kalbejimo_tikslumas*?turinio_tikslumas) AS ?tikslumo_ivertis) ?kalbėjimas_i events:talked__substance ?turinys_i . ?turinys_i rdf:type events:substance . ?asmuo_i rdfs:label ?asmuo_label . </pre>

		?turinys_i rdfs:label ?turinys_label. ?salis ann:label_lemma "vokietija"@lt . ?salis rdf:type locations:location . ?kalbėjimas_i events:included__object ?salis. } order by DESC (?tikslumo_ivertis)
--	--	--

8.31 lentelė. Užklauso „Ką kalbėjo asmenys minėdami „Vokietija“?“ rezultatai

Užklauso pavadinimas	užklauso sunkumas	užklauso tipas	imtys	Išvedamų rašų skaičius	Vidurkis, ms	Paklaida, ms
Ką kalbėjo asmenys minėdami "Vokietija"	dviejų skirtingų tipų objektai	be tikslumo įverčio	25 imtis	1	8.0777	4.1834
			50 imtis	4	7.6708	3.8298
			100 imtis	9	9.9269	4.5281
		tikslumo įvertis dokumento atžvilgiu	25 imtis	1	8.8247	5.1608
			50 imtis	4	12.1596	5.0935
			100 imtis	9	13.4157	11.0395
		tikslumo įvertis individo atžvilgiu	25 imtis	1	961.9670	86.1543
			50 imtis	4	3357.1961	90.2974
			100 imtis	9	9651.9005	522.9373

8.32 lentelė. Užklauso „Ką kalbėjo asmenys minėdami "Europa"?“ duomenys

Užklauso sudėtingumas	Užklauso tipas	Ką kalbėjo asmenys minėdami "Europa"?
Dviejų skirtingų tipų objektai + konkretus objektas	neišvedamas tikslumo įvertis	<pre> SELECT DISTINCT ?asmuo_label ?turinys_label WHERE { ?dokumentas semLT:refers_to__object ?asmuo_i . ?dokumentas rdf:type semLT:document. ?dokumentas semLT:refers_to__object ?kalbėjimas_i . ?asmuo_i events:talked__talking ?kalbėjimas_i . ?asmuo_i rdf:type agents:person . ?kalbėjimas_i rdf:type politics:talking . ?dokumentas semLT:refers_to__object ?turinys_i . ?kalbėjimas_i events:talked__substance ?turinys_i . ?turinys_i rdf:type events:substance . ?asmuo_i rdfs:label ?asmuo_label . ?turinys_i rdfs:label ?turinys_label. ?salis ann:label_lemma "europa"@lt . ?salis rdf:type locations:location . ?kalbėjimas_i events:included__object ?salis. } </pre>
	išvedamas tikslumo įvertis dokumento atžvilgiu	<pre> SELECT DISTINCT ?asmuo_label ?turinys_label ?tikslumo_ivertis WHERE { ?dokumentas semLT:refers_to__object ?asmuo_i . ?dokumentas rdf:type semLT:document. OPTIONAL {?dokumentas evaluation:fmeasure ?tikslumo_ivertis . } ?dokumentas semLT:refers_to__object ?kalbėjimas_i . ?asmuo_i events:talked__talking ?kalbėjimas_i . ?asmuo_i rdf:type agents:person . ?kalbėjimas_i rdf:type politics:talking . ?dokumentas semLT:refers_to__object ?turinys_i . ?kalbėjimas_i events:talked__substance ?turinys_i . ?turinys_i rdf:type events:substance . ?asmuo_i rdfs:label ?asmuo_label . ?turinys_i rdfs:label ?turinys_label. ?salis ann:label_lemma "europa"@lt . ?salis rdf:type locations:location . ?kalbėjimas_i events:included__object ?salis. } order by DESC (?tikslumo_ivertis) </pre>

	išvedamas tikslumo įvertis individo atžvilgiu	<pre> SELECT DISTINCT ?asmuo_label ?turinys_label ?tikslumo_ivertis ?asmens_tikslumas ?kalbejimo_tikslumas ?turinio_tikslumas WHERE { ?dokumentas semLT:refers_to__object ?asmuo_i . ?dokumentas rdf:type semLT:document. OPTIONAL {?asmuo_i evaluation:fmeasure ?asmens_tikslumas . } OPTIONAL {?kalbejimas_i evaluation:fmeasure ?kalbejimo_tikslumas . } OPTIONAL {?turinys_i evaluation:fmeasure ?turinio_tikslumas . } ?dokumentas semLT:refers_to__object ?kalbejimas_i . ?asmuo_i events:talked__talking ?kalbejimas_i . ?asmuo_i rdf:type agents:person . ?kalbejimas_i rdf:type politics:talking . ?dokumentas semLT:refers_to__object ?turinys_i . BIND ((?asmens_tikslumas*?kalbejimo_tikslumas*?turinio_tikslumas) AS ?tikslumo_ivertis) ?kalbejimas_i events:talked__substance ?turinys_i . ?turinys_i rdf:type events:substance . ?asmuo_i rdfs:label ?asmuo_label . ?turinys_i rdfs:label ?turinys_label. ?salis ann:label_lemma "europa"@lt . ?salis rdf:type locations:location . ?kalbejimas_i events:included__object ?salis. } order by DESC (?tikslumo_ivertis) </pre>
--	---	--

8.33 lentelė. Užklauso „Ką kalbėjo asmenys minėdami "Europa"?“ rezultatai

Užklauso pavadinimas	Užklauso sunkumas	Užklauso tipas	Imtys	Išvedamų įrašų skaičius	Vidurkis, ms	Paklaida, ms
Ką kalbėjo asmenys minėdami "Europa"?	Dviejų skirtingų tipų objektai + konkretus objektas	be tikslumo įverčio	25 imtis	2	12.8607	10.3631
			50 imtis	10	11.0675	5.3911
			100 imtis	27	14.7028	6.5701
		tikslumo įvertis dokumento atžvilgiu	25 imtis	2	10.2663	5.8342
			50 imtis	10	12.045	8.4707
			100 imtis	27	12.8325	5.0862
		tikslumo įvertis individo atžvilgiu	25 imtis	2	990.1925	114.4206
			50 imtis	10	3422.466	90.0331
			100 imtis	27	11809.53	2537.28

8.34 lentelė. Užklauso „Ką kalbėjo asmenys minėdami "Lietuva"?“ duomenys

Užklauso sudėtingumas	Užklauso tipas	Ką kalbėjo asmenys minėdami "Lietuva"?
Dviejų skirtingų tipų objektai + konkretus objektas	neišvedamas tikslumo įvertis	<pre> SELECT DISTINCT ?asmuo_label ?turinys_label WHERE { ?dokumentas semLT:refers_to__object ?asmuo_i . ?dokumentas rdf:type semLT:document. ?dokumentas semLT:refers_to__object ?kalbejimas_i . ?asmuo_i events:talked__talking ?kalbejimas_i . ?asmuo_i rdf:type agents:person . ?kalbejimas_i rdf:type politics:talking . ?dokumentas semLT:refers_to__object ?turinys_i . ?kalbejimas_i events:talked__substance ?turinys_i . ?turinys_i rdf:type events:substance . ?asmuo_i rdfs:label ?asmuo_label . ?turinys_i rdfs:label ?turinys_label. ?salis ann:label_lemma "lietuva"@lt . ?salis rdf:type locations:location . ?kalbejimas_i events:included__object ?salis. } </pre>

išvedamas tikslumo įvertis dokumento atžvilgiu		<pre> SELECT DISTINCT ?asmuo_label ?turinys_label ?tikslumo_ivertis WHERE { ?dokumentas semLT:refers_to__object ?asmuo_i . ?dokumentas rdf:type semLT:document. OPTIONAL {?dokumentas evaluation:fmeasure ?tikslumo_ivertis .} ?dokumentas semLT:refers_to__object ?kalbėjimas_i . ?asmuo_i events:talked__talking ?kalbėjimas_i . ?asmuo_i rdf:type agents:person . ?kalbėjimas_i rdf:type politics:talking . ?dokumentas semLT:refers_to__object ?turinys_i . ?kalbėjimas_i events:talked__substance ?turinys_i . ?turinys_i rdf:type events:substance . ?asmuo_i rdfs:label ?asmuo_label . ?turinys_i rdfs:label ?turinys_label . ?salis ann:label_lemma "lietuva"@lt . ?salis rdf:type locations:location . ?kalbėjimas_i events:included__object ?salis . } order by DESC (?tikslumo_ivertis) </pre>
išvedamas tikslumo įvertis individo atžvilgiu		<pre> SELECT DISTINCT ?asmuo_label ?turinys_label ?tikslumo_ivertis ?asmens_tikslumas ?kalbejimo_tikslumas ?turinio_tikslumas WHERE { ?dokumentas semLT:refers_to__object ?asmuo_i . ?dokumentas rdf:type semLT:document. OPTIONAL {?asmuo_i evaluation:fmeasure ?asmens_tikslumas .} OPTIONAL {?kalbėjimas_i evaluation:fmeasure ?kalbejimo_tikslumas .} OPTIONAL {?turinys_i evaluation:fmeasure ?turinio_tikslumas .} ?dokumentas semLT:refers_to__object ?kalbėjimas_i . ?asmuo_i events:talked__talking ?kalbėjimas_i . ?asmuo_i rdf:type agents:person . ?kalbėjimas_i rdf:type politics:talking . ?dokumentas semLT:refers_to__object ?turinys_i . BIND ((?asmens_tikslumas*?kalbejimo_tikslumas*?turinio_tikslumas) AS ?tikslumo_ivertis) ?kalbėjimas_i events:talked__substance ?turinys_i . ?turinys_i rdf:type events:substance . ?asmuo_i rdfs:label ?asmuo_label . ?turinys_i rdfs:label ?turinys_label . ?salis ann:label_lemma "lietuva"@lt . ?salis rdf:type locations:location . ?kalbėjimas_i events:included__object ?salis . } order by DESC (?tikslumo_ivertis) </pre>

8.35 lentelė. Užklauso „Ką kalbėjo asmenys minėdami "Lietuva"?“ rezultatai

Užklauso pavadinimas	Užklauso sunkumas	Užklauso tipas	Imtys	Išvedamų įrašų skaičius	Vidurkis, ms	Paklaida, ms
Ką kalbėjo asmenys minėdami "Lietuva"?	Dviejų skirtingų tipų objektai + konkretus objektas	be tikslumo įverčio	25 imtis	9	9.5558	4.5357
			50 imtis	37	14.8023	6.6363
			100 imtis	97	242.6634	108.6552
		tikslumo įvertis dokumento atžvilgiu	25 imtis	9	10.0057	4.6672
			50 imtis	37	19.1632	9.4428
			100 imtis	97	318.2696	142.2825
		tikslumo įvertis individo atžvilgiu	25 imtis	9	1154.9726	229.9169
			50 imtis	37	5496.6701	687.7215
			100 imtis	97	23521.514	2891.1948

8.36 lentelė. Užklauso „Ką pareiškė asmenys?“ duomenys

Užklauso sudėtingumas	Užklauso tipas	Ką pareiškė asmenys?
-----------------------	----------------	----------------------

Dviejų skirtingų tipų objektai	neišvedamas tikslumo įvertis	<pre> SELECT DISTINCT ?asmuo_label ?turinys_label WHERE { ?dokumentas semLT:refers_to__object ?asmuo_i . ?dokumentas rdf:type semLT:document. ?dokumentas semLT:refers_to__object ?kalbėjimas_i . ?asmuo_i events:talked__talking ?kalbėjimas_i . ?asmuo_i rdf:type agents:person . ?kalbėjimas_i rdf:type politics:talking . ?dokumentas semLT:refers_to__object ?turinys_i . ?kalbėjimas_i events:talked__substance ?turinys_i . ?turinys_i rdf:type events:substance . ?kalbėjimas_i politics:has__talking_type politics:statement_talking_type. ?asmuo_i rdfs:label ?asmuo_label . ?turinys_i rdfs:label ?turinys_label } </pre>
	išvedamas tikslumo įvertis dokumento atžvilgiu	<pre> SELECT DISTINCT ?asmuo_label ?turinys_label ?tikslumo_ivertis WHERE { ?dokumentas semLT:refers_to__object ?asmuo_i . ?dokumentas rdf:type semLT:document. OPTIONAL {?dokumentas evaluation:fmeasure ?tikslumo_ivertis .} ?dokumentas semLT:refers_to__object ?kalbėjimas_i . ?asmuo_i events:talked__talking ?kalbėjimas_i . ?asmuo_i rdf:type agents:person . ?kalbėjimas_i rdf:type politics:talking . ?dokumentas semLT:refers_to__object ?turinys_i . ?kalbėjimas_i events:talked__substance ?turinys_i . ?turinys_i rdf:type events:substance . ?kalbėjimas_i politics:has__talking_type politics:statement_talking_type. ?asmuo_i rdfs:label ?asmuo_label . ?turinys_i rdfs:label ?turinys_label } order by DESC (?tikslumo_ivertis) </pre>
	išvedamas tikslumo įvertis individo atžvilgiu	<pre> SELECT DISTINCT ?asmuo_label ?turinys_label ?tikslumo_ivertis ?asmens_tikslumas ?turinio_tikslumas WHERE { ?dokumentas semLT:refers_to__object ?asmuo_i . ?dokumentas rdf:type semLT:document. OPTIONAL {?asmuo_i evaluation:fmeasure ?asmens_tikslumas .} OPTIONAL {?turinys_i evaluation:fmeasure ?turinio_tikslumas .} ?dokumentas semLT:refers_to__object ?kalbėjimas_i . ?asmuo_i events:talked__talking ?kalbėjimas_i . ?asmuo_i rdf:type agents:person . ?kalbėjimas_i rdf:type politics:talking . ?dokumentas semLT:refers_to__object ?turinys_i . ?kalbėjimas_i events:talked__substance ?turinys_i . ?turinys_i rdf:type events:substance . ?kalbėjimas_i politics:has__talking_type politics:statement_talking_type. BIND ((?asmens_tikslumas*?turinio_tikslumas) AS ?tikslumo_ivertis) ?asmuo_i rdfs:label ?asmuo_label . ?turinys_i rdfs:label ?turinys_label } order by DESC (?tikslumo_ivertis) </pre>

8.37 lentelė. Užklauso „Ką pareiškė asmenys?“ rezultatai

Užklauso pavadinimas	Užklauso sunkumas	Užklauso tipas	Imtys	Išvedamų įrašų skaičius	Vidurkis, ms	Paklaida, ms
Ką pareiškė asmenys?	dviejų skirtingų tipų objektai	be tikslumo įverčio	25 imtis	20	5.4065	15.9520
			50 imtis	66	20.6403	12.2778
			100 imtis	176	26.8669	9.4184
		tikslumo įvertis dokumento atžvilgiu	25 imtis	20	15.5329	6.5376
			50 imtis	66	24.3488	9.2255

		100 imtis	176	40.4244	12.6702
	tikslumo įvertis individo atžvilgiu	25 imtis	20	13.5891	5.8770
		50 imtis	66	30.0080	12.2363
		100 imtis	176	50.0948	14.3864

8.38 lentelė. Užklauso „Ką samprotavo asmenys?“ duomenys

Užklauso sudėtingumas	Užklauso tipas	Ką samprotavo asmenys?
Dviejų skirtingų tipų objektai	neišvedamas tikslumo įvertis	<pre>SELECT DISTINCT ?asmuo_label ?turinys_label WHERE { ?dokumentas semLT:refers_to__object ?asmuo_i . ?dokumentas rdf:type semLT:document. ?dokumentas semLT:refers_to__object ?kalbėjimas_i . ?asmuo_i events:talked__talking ?kalbėjimas_i . ?asmuo_i rdf:type agents:person . ?kalbėjimas_i rdf:type politics:talking . ?dokumentas semLT:refers_to__object ?turinys_i . ?kalbėjimas_i events:talked__substance ?turinys_i . ?turinys_i rdf:type events:substance . ?kalbėjimas_i politics:has__talking_type politics:reasoning_talking_type. ?asmuo_i rdfs:label ?asmuo_label . ?turinys_i rdfs:label ?turinys_label }</pre>
	išvedamas tikslumo įvertis dokumento atžvilgiu	<pre>SELECT DISTINCT ?asmuo_label ?turinys_label ?tikslumo_ivertis WHERE { ?dokumentas semLT:refers_to__object ?asmuo_i . ?dokumentas rdf:type semLT:document. OPTIONAL {?dokumentas evaluation:fmeasure ?tikslumo_ivertis . } ?dokumentas semLT:refers_to__object ?kalbėjimas_i . ?asmuo_i events:talked__talking ?kalbėjimas_i . ?asmuo_i rdf:type agents:person . ?kalbėjimas_i rdf:type politics:talking . ?dokumentas semLT:refers_to__object ?turinys_i . ?kalbėjimas_i events:talked__substance ?turinys_i . ?turinys_i rdf:type events:substance . ?kalbėjimas_i politics:has__talking_type politics:reasoning_talking_type. ?asmuo_i rdfs:label ?asmuo_label . ?turinys_i rdfs:label ?turinys_label } order by DESC (?tikslumo_ivertis)</pre>
	išvedamas tikslumo įvertis individo atžvilgiu	<pre>SELECT DISTINCT ?asmuo_label ?turinys_label ?tikslumo_ivertis ?asmens_tikslumas ?turinio_tikslumas WHERE { ?dokumentas semLT:refers_to__object ?asmuo_i . ?dokumentas rdf:type semLT:document. OPTIONAL {?asmuo_i evaluation:fmeasure ?asmens_tikslumas . } OPTIONAL {?turinys_i evaluation:fmeasure ?turinio_tikslumas . } ?dokumentas semLT:refers_to__object ?kalbėjimas_i . ?asmuo_i events:talked__talking ?kalbėjimas_i . ?asmuo_i rdf:type agents:person . ?kalbėjimas_i rdf:type politics:talking . ?dokumentas semLT:refers_to__object ?turinys_i . ?kalbėjimas_i events:talked__substance ?turinys_i . ?turinys_i rdf:type events:substance . ?kalbėjimas_i politics:has__talking_type politics:reasoning_talking_type. BIND ((?asmens_tikslumas*?turinio_tikslumas) AS ?tikslumo_ivertis) ?asmuo_i rdfs:label ?asmuo_label . ?turinys_i rdfs:label ?turinys_label } order by DESC (?tikslumo_ivertis)</pre>

8.39 lentelė. Užklauso „Ką samprotavo asmenys?“ rezultatai

Užklauso pavadinimas	Užklauso sunkumas	Užklauso tipas	Imtys	Išvedamų įrašų skaičius	Vidurkis, ms	Paklaida, ms
Ką samprotavo asmenys?	Dviejų skirtingų tipų objektai	be tikslumo įverčio	25 imtis	3	8.2394	3.4776
			50 imtis	15	11.3037	4.8314
			100 imtis	32	11.3475	6.0266
		tikslumo įvertis dokumento atžvilgiu	25 imtis	3	6.003	2.1761
			50 imtis	15	10.9243	6.5872
			100 imtis	32	18.8653	9.5048
		tikslumo įvertis individo atžvilgiu	25 imtis	3	10.2362	5.4556
			50 imtis	15	13.6846	7.3678
			100 imtis	32	17.4279	11.0847

8.40 lentelė. Užklauso „Ką akcentavo asmenys?“ duomenys

Užklauso sudėtingumas	Užklauso tipas	Ką akcentavo asmenys?
Dviejų skirtingų tipų objektai	neišvedamas tikslumo įvertis	<pre> SELECT DISTINCT ?asmuo_label ?turinys_label WHERE { ?dokumentas semLT:refers_to__object ?asmuo_i . ?dokumentas rdf:type semLT:document. ?dokumentas semLT:refers_to__object ?kalbėjimas_i . ?asmuo_i events:talked__talking ?kalbėjimas_i . ?asmuo_i rdf:type agents:person . ?kalbėjimas_i rdf:type politics:talking . ?dokumentas semLT:refers_to__object ?turinys_i . ?kalbėjimas_i events:talked__substance ?turinys_i . ?turinys_i rdf:type events:substance . ?kalbėjimas_i politics:has__talking_type politics:emphasis_talking_type. ?asmuo_i rdfs:label ?asmuo_label . ?turinys_i rdfs:label ?turinys_label } </pre>
	išvedamas tikslumo įvertis dokumento atžvilgiu	<pre> SELECT DISTINCT ?asmuo_label ?turinys_label ?tikslumo_ivertis WHERE { ?dokumentas semLT:refers_to__object ?asmuo_i . ?dokumentas rdf:type semLT:document. OPTIONAL {?dokumentas evaluation:fmeasure ?tikslumo_ivertis .} ?dokumentas semLT:refers_to__object ?kalbėjimas_i . ?asmuo_i events:talked__talking ?kalbėjimas_i . ?asmuo_i rdf:type agents:person . ?kalbėjimas_i rdf:type politics:talking . ?dokumentas semLT:refers_to__object ?turinys_i . ?kalbėjimas_i events:talked__substance ?turinys_i . ?turinys_i rdf:type events:substance . ?kalbėjimas_i politics:has__talking_type politics:emphasis_talking_type. ?asmuo_i rdfs:label ?asmuo_label . ?turinys_i rdfs:label ?turinys_label } } order by DESC (?tikslumo_ivertis) </pre>
	išvedamas tikslumo įvertis individo atžvilgiu	<pre> SELECT DISTINCT ?asmuo_label ?turinys_label ?tikslumo_ivertis ?asmens_tikslumas ?turinio_tikslumas WHERE { ?dokumentas semLT:refers_to__object ?asmuo_i . ?dokumentas rdf:type semLT:document. OPTIONAL {?asmuo_i evaluation:fmeasure ?asmens_tikslumas .} OPTIONAL {?turinys_i evaluation:fmeasure ?turinio_tikslumas .} ?dokumentas semLT:refers_to__object ?kalbėjimas_i . ?asmuo_i events:talked__talking ?kalbėjimas_i . ?asmuo_i rdf:type agents:person . ?kalbėjimas_i rdf:type politics:talking . ?dokumentas semLT:refers_to__object ?turinys_i . ?kalbėjimas_i events:talked__substance ?turinys_i . ?turinys_i rdf:type events:substance . } </pre>

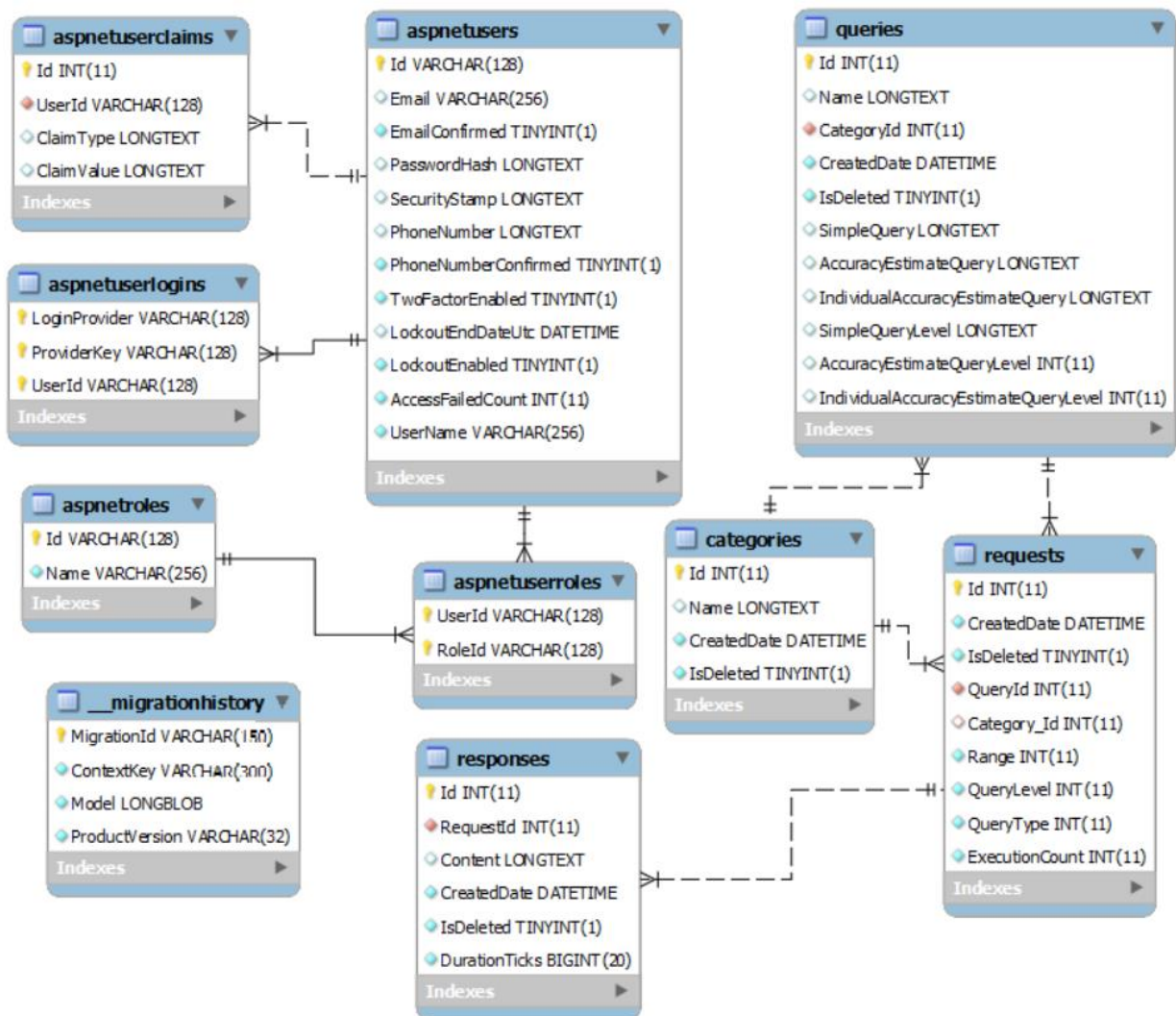
		?kalbėjimas_i politics:has_talking_type politics:emphasis_talking_type. BIND ((?asmens_tikslumas*?turinio_tikslumas) AS ?tikslumo_ivertis) ?asmuo_i rdfs:label ?asmuo_label . ?turinys_i rdfs:label ?turinys_label } lorder by DESC (?tikslumo_ivertis)
--	--	---

8.41 lentelė. Užklauso „Ką akcentavo asmenys?“ rezultatai

Užklauso pavadinimas	Užklauso sunkumas	Užklauso tipas	Imtys	Išvedamų įrašų skaičius	Vidurkis, ms	Paklaida, ms
Ką akcentavo asmenys?	Dviejų skirtingų tipų objektai	be tikslumo įverčio	25 imtis	7	9.2295	6.322
			50 imtis	34	13.6856	6.3887
			100 imtis	78	19.072	9.453
		tikslumo įvertis dokumento atžvilgiu	25 imtis	7	8.9316	3.7931
			50 imtis	34	12.4025	7.9057
			100 imtis	78	21.3327	10.9783
		tikslumo įvertis individo atžvilgiu	25 imtis	7	12.4334	6.5177
			50 imtis	34	15.5535	8.5705
			100 imtis	78	28.8211	14.5366

8.5. priedas. Fizinė duomenų bazė

Fizinė duomenų bazės schema (žr. 72 pav.) pateikiamos sukurtos duomenų bazės lentelės. Lentelių sąrašas (žr. 8.42 lentelė.) aprašoma kiekvienos lentelės paskirtis.



72 pav. Ontologijos praturtintos individų atpažinimo tikslumo įvertinimo sistemos fizinės duomenų bazės schema

8.42 lentelė. Duomenų bazės lentelių sąrašas

Pavadinimas	Aprašymas
aspnetuserclaims	Duomenų bazės lentelė skirta saugoti vartotojų prašymus (žr. 8.43 lentelė).
aspnetusers	Duomenų bazės lentelė skirta saugoti vartojus (žr. 8.44 lentelė).
queries	Duomenų bazės lentelė skirta saugoti naudojamus užklausas (žr. 8.45 lentelė).
aspnetuserlogins	Duomenų bazės lentelė skirta saugoti vartotojo prisijungimo duomenis (žr. 8.46 lentelė).
aspnetroles	Duomenų bazės lentelė skirta saugoti teises (žr. 8.47 lentelė).
aspnetuserroles	Duomenų bazės lentelė skirta saugoti vartotojo teises (žr. 8.48 lentelė).
categories	Duomenų bazės lentelė skirta saugoti paieškos sričių pavadinimus (žr. 8.49 lentelė).
requests	Duomenų bazės lentelė skirta saugoti suformuotas konkrečias užklausas (žr. 8.50 lentelė).
__migrationhistory	Duomenų bazės lentelė skirta saugoti duomenų bazės pakeitimus (žr. 8.51 lentelė).
responses	Duomenų bazės lentelė skirta saugoti užklausų rezultatus (žr. 8.52 lentelė).

8.43 lentelė. Duomenų bazės lentelės „aspnetuserclaims“ aprašas

Pavadinimas	Tipas	Privalomas	Paskirtis
Id	INTEGER	Taip	Identifikavimo numeris
RequestId	INTEGER	Taip	Suformuotas konkrečias užklausos identifikavimo numeris
Content	LONGTEXT	Taip	Turinys
CreatedDate	DATETIME	Taip	Sukūrimo data
IsDeleted	TINYINT(1)	Taip	Ar ištrinta
DurationTicks	BIGINT(20)	Taip	Sugaištas tikų kiekis

8.44 lentelė. Duomenų bazės lentelės „aspnetusers“ aprašas

Pavadinimas	Tipas	Privalomas	Paskirtis
-------------	-------	------------	-----------

Id	VARCHAR(128)	Taip	Identifikavimo numeris
Email	VARCHAR(256)	Taip	El. paštas
EmailConfirmed	TINYINT(1)	Taip	EL. paštas patvirtintas
PasswordHash	LONGTEXT	Taip	Slaptažodžio patvirtinimas
SecurityStamp	LONGTEXT	Taip	Saugumo atspaudas
PhoneNumber	LONGTEXT	Taip	Telefono numeris
PhoneNumberConfirmed	TINYINT(1)	Taip	Telefono numeris patvirtintas
UserName	VARCHAR(256)	Taip	Vartotojo vardas

8.45 lentelė. Duomenų bazės lentelės „queries“ aprašas

Pavadinimas	Tipas	Privalomas	Paskirtis
Id	INTEGER(11)	Taip	Identifikavimo numeris
Name	LONGTEXT	Taip	Užklauskos pavadinimas
CategoryId	INTEGER(11)	Taip	Paieškos srities identifikavimo numeris
CreatedDate	DATETIME	Taip	Sukūrimo data
IsDeleted	TINYINT(1)	Taip	Ar ištrinta
SimpleQuery	LONGTEXT	Taip	Neišvedamas tikslumo įvertis
AccuracyEstimateQuery	LONGTEXT	Taip	Išvedamas tikslumo įvertis dokumento atžvilgiu
IndividualAccuracyEstimateQuery	LONGTEXT	Taip	Išvedamas tikslumo įvertis individo atžvilgiu
SimpleQueryLevel	LONGTEXT	Taip	Vienas tipo objektas
AccuracyEstimateQueryLevel	INTEGER(11)	Taip	Dviejų skirtingų tipų objektai
IndividualAccuracyEstimateQueryLevel	INTEGER(11)	Taip	Dviejų skirtingų tipų objektai ir konkretus objektas

8.46 lentelė. Duomenų bazės lentelės „aspnetuserlogins“ aprašas

Pavadinimas	Tipas	Privalomas	Paskirtis
LoginProvider	VARCHAR(128)	Taip	Prisijungimo paslaugos teikėjas
ProviderKey	VARCHAR(128)	Taip	Teikėjo raktas
UserId	VARCHAR(128)	Taip	Vartotojo identifikavimo numeris

8.47 lentelė. Duomenų bazės lentelės „aspnetroles“ aprašas

Pavadinimas	Tipas	Privalomas	Paskirtis
Id	VARCHAR(128)	Taip	Identifikavimo numeris
Name	VARCHAR(256)	Taip	Pavadinimas

8.48 lentelė. Duomenų bazės lentelės „aspnetuserroles“ aprašas

Pavadinimas	Tipas	Privalomas	Paskirtis
UserId	VARCHAR(128)	Taip	Identifikavimo numeris
RoleId	VARCHAR(128)	Taip	Rolės identifikavimo numeris

8.49 lentelė. Duomenų bazės lentelės „categories“ aprašas

Pavadinimas	Tipas	Privalomas	Paskirtis
Id	INTEGER(11)	Taip	Paieškos srities identifikavimo numeris
Name	LONGTEXT	Taip	Paieškos srities pavadinimas
CreatedDate	DATETIME	Taip	Paieškos srities sukūrimo data
IsDeleted	TINYINT(1)	Taip	Ar ištrinta paieškos srities

8.50 lentelė. Duomenų bazės lentelės „requests“ aprašas

Pavadinimas	Tipas	Privalomas	Paskirtis
Id	INTEGER(11)	Taip	Užklauskos identifikavimo numeris
CreatedDate	DATETIME	Taip	Užklauskos sukūrimo data
IsDeleted	TINYINT(1)	Taip	Ar ištrinta užklausa
QueryId	INTEGER(11)	Taip	Užklauskos identifikavimo numeris
Category_Id	INTEGER(11)	Taip	Užklauskos paieškos sritis
Range	INTEGER(11)	Taip	Užklauskos imtis
QueryLevel	INTEGER(11)	Taip	Užklauskos sudėtingumas
QueryType	INTEGER(11)	Taip	Užklauskos tipas

ExecutionCount	INTEGER(11)	Taip	Užklauskos vykdymų skaičius
----------------	-------------	------	-----------------------------

8.51 lentelė. Duomenų bazės lentelės „__migrationhistory“ aprašas

Pavadinimas	Tipas	Privalomas	Paskirtis
MigrationId	VARCHAR(150)	Taip	Identifikavimo numeris
ContextKey	VARCHAR(300)	Taip	Nuoroda į turinio identifikatorių
Model	LONGBLOB	Taip	Modelis
ProductVersion	VARCHAR(32)	Taip	Versija

8.52 lentelė. Duomenų bazės lentelės „responses“ aprašas

Pavadinimas	Tipas	Privalomas	Paskirtis
Id	INTEGER	Taip	Identifikavimo numeris
RequestId	INTEGER	Taip	Suformuotas konkrečias užklauskos identifikavimo numeris
Content	LONGTEXT	Taip	Turinys
CreatedDate	DATETIME	Taip	Sukūrimo data
IsDeleted	TINYINT(1)	Taip	Ar ištrinta
DurationTicks	BIGINT(20)	Taip	Sugaištas tikų kiekis