

KAUNO TECHNOLOGIJOS UNIVERSITETAS
INFORMATIKOS FAKULTETAS
INFORMACIJOS SISTEMŲ KATEDRA

Evaldas Karazinas

**Semantinių web technologijų tyrimas ir taikymas
video paskaitų sistemoje VIPS**

*(Semantic Web Technology Research and Application
in Video Lecturing System VIPS)*

Magistro darbas

Darbo vadovas

prof. habil. dr. A. Targamadžė

KAUNAS, 2008

KAUNO TECHNOLOGIJOS UNIVERSITETAS
INFORMATIKOS FAKULTETAS
INFORMACIJOS SISTEMŲ KATEDRA

Evaldas Karazinas

**Semantinių web technologijų tyrimas ir taikymas
video paskaitų sistemoje VIPS**

*(Semantic Web Technology Research and Application
in Video Lecturing System VIPS)*

Magistro darbas

Recenzentas

Doc. Dr. S. Maciulevičius
2008.01.15

Vadovas

prof. habil. dr. A. Targamadžė
2008.01.15

Atliko

IFM-2/4 gr. stud.
Evaldas Karazinas
2008-01-15

Kaunas, 2008

TURINYS

1. Semantinės web technologijos	7
1.1. Tyrimo sritis ir sprendžiamos problemos	7
1.2. Tyrimo objektas	7
1.3. Sistemos vartotojai	8
1.3.1. Vartotojų aibė, tipai ir savybės	8
1.3.2. Vartotojų tikslai ir problemos	9
1.4. Problemos sprendimo metodų literatūros šaltiniuose analizė	9
1.4.1. paieškos organizavimo galimybės	9
1.4.2. Semantinis žiniatinklis ir paieška	9
1.4.3. Paieška po „gilųjį internetą“ (Deep web)	12
1.4.4. Paieška per vartotojų žymes	13
1.5. Panašių sistemų analizė	14
1.6. Darbo tikslas ir siekiami privalumai	16
2. Folksonomija grįstas paieškos modelis	17
2.1. Rekomenduojamų žymių išgavimas	18
2.2. Įrašų žymėjimas ir paieška	19
2.3. Apibendrinimas	20
3. Paieškos posistemis video paskaitų sistemoje VIPS	21
3.1. Reikalavimai	21
3.1.1. Siekiamos paieškos sistemos apibrėžimas	21
3.1.2. Kompiuterizuojamos funkcijos	21
3.1.3. Panaudojimo atvejų specifikacijos	22
3.1.4. Reikalavimai duomenims	23
3.1.5. Nefunkciniai reikalavimai ir apribojimai	24
3.1.5.1. Reikalavimai sąsajai	24
3.1.5.2. Reikalavimai kokybei	24
3.1.6. Rizikos veiksniai	24
3.1.7. Rezultato kokybės kriterijai	24
3.2. Projektas	25
3.2.1. Statinė struktūra	25
3.2.1.1. Loginė architektūra	25
3.2.1.2. Vartotojo paslaugų objektai	27
3.2.1.3. Veiklos logikos objektai	28
3.2.1.4. Duomenų bazės architektūra	29
3.3. Sistemos realizacija	30
3.3.1. Komponentų diagrama	30
3.3.2. Sistemos testavimo planas	31
4. Žodynėlis	32
5. Išvados	33
6. Literatūra	34
7. Priedai	36
7.1. Programos prototipo langai	36
7.2. Panaudojimo atvejai	39
7.3. Straipsnis	41

Lentelių sąrašas

1 lentelė.	PA „Įkelti skaidres ir žymes” specifikacija	22
2 lentelė.	PA „Žymėti” specifikacija	23
3 lentelė.	PA „Ieškoti” specifikacija	23
4 lentelė.	Reikalavimai vartotojo sąsajai	24
5 lentelė.	Reikalavimai kokybei	24
6 lentelė.	Projekto rizikos faktoriai	24
7 lentelė.	Programinės įrangos kokybės kriterijai	24
8 lentelė.	Nefunkcinių reikalavimų testavimo planas	31
9 lentelė.	Funkcinio testavimo planas	31

Paveikslų sąrašas

1.	pav. Paieškos veiklos proceso diagrama.....	8
2.	pav. Bakterijos hierarchinėje struktūroje	11
3.	pav. Verslo proceso modelį sudarančių elementų priklausomybės	17
4.	pav. Skaidrių sukėlimo ir žymių išgavimo veiklos modelis	18
5.	pav. Žymių suteikimo įrašams ir paieškos veiklos modelis	19
6.	pav. Kompiuterizuojamos funkcijos.....	22
7.	pav. Loginės architektūros modelis	25
8.	pav. Analitinis modelis.....	26
9.	pav. Vartotojo sąsajos klasių modelis.....	27
10.	pav. Veiklos paslaugų klasių modelis	28
11.	pav. Duomenų bazės schema	29
12.	pav. Komponentų diagrama.....	30
13.	pav. Paieškos langas.....	36
14.	pav. Žymėjimo langas	37
15.	pav. Metaduomenų redagavimo langas.....	38
16.	pav. Panaudojimo atvejis „Ieškoti“	39
17.	pav. Panaudojimo atvejis „Įkelti skaidres ir žymes“	40
18.	pav. Panaudojimo atvejis „tvarkyti savo žymes“	40
19.	pav. Panaudojimo atvejis „žymėti“	40

Semantic Web Technology Research and Application in Video Lecturing System VIPS

Summary

The aim of this paper is to present search organization opportunities in VIPS system archive in semantic manner. The Semantic Web is an evolving extension of the World Wide Web in which web content can be expressed not only in natural language, but also in a format that can be read and used by software agents, thus permitting them to find, share and integrate information more easily. Social tagging can also group information in semantic manner. Theoretical background of specified model is delivered in the beginning sections of this thesis. As experimental solution information system was introduced. Thesis is concluded and recommendations are suggested for future.

Įvadas

Informacinių technologijų bei kompiuterių tinklų sparti plėtra įgalino apdoroti, saugoti bei perduoti didžiulius informacijos kiekius, o tai paskatino dar spartesnę informacijos kūrimą. Dideliais tempais augant informacijos kiekiui ypač svarbus tampa jos organizavimas bei paieška. Nors šią problemą sprendžia galingos paieškos sistemos tokios kaip Google, didelė informacijos dalis joms nėra prieinama, nes ji yra sukuriama ir saugoma uždaroje sistemoje kurios reikalauja vartotojo autorizavimo. Tokios sistemos paprastai turi savo vidinius informacijos resursų kategorizavimo bei paieškos būdus, tačiau jie ne visada yra efektyvūs. Dažnai informacijos resursų, kurie būtų aktualūs vartotojui, radimas tampa sudėtingu, o neturint galimybės dideliame informacijos kiekyje greitai rasti tai ko reikia, net ir labai vertinga informacija lieka nepanaudota. Tokiu atveju, net ir turint didžiules virtualias bibliotekas su milijonais įrašų, negalima gauti maksimalios naudos, o pačios bibliotekos praranda savo reikšmę.

Galima teigti, kad paieškos sėkmė priklauso ne tik nuo paieškos sistemos algoritmų sudėtingumo, bet labiau – nuo informacijos resursų aprašymo. Pakankamai ir prasmingai aprašytus resursus būtų paprasta surasti ir panaudoti efektyviai, todėl yra kuriamos įvairios semantinės metodikos nuo laisvo tipo aprašymo – folksonomijos iki griežtos klasifikacijos pasitelkiant ontologijas.

Šiame darbe nagrinėjamos semantines web technologijos ir pabandoma jas praktiškai pritaikyti Video paskaitų sistemos VIPS įrašų ir kursų informacijos archyvo organizavimui bei paieškos jame realizavimui. KTU Distancinio mokymo centro serveryje įdiegtoje VIPS sistemoje jau yra sukauptas daugiau kaip 4 tūkst. įrašų archyvas iš įvairių renginių bei paskaitų. Šis archyvas gali būti naudojamas kaip skaitmeninių mokymosi resursų saugykla išplečiant jų panaudojimą tiek formaliame tiek ir neformaliame švietime. Norint efektyviai panaudoti VIPS sistemoje sukauptus mokymosi resursus reikalinga paieška, galinti pateikti tikslus bei aktualius rezultatus.

1. Semantinės web technologijos

Darbo tikslas yra semantinių web technologijų pagrindu sukurti paieškos posistemį Video paskaitų sistemos VIPS vaizdo įrašų bibliotekai. Tam reikia apžvelgti literatūros šaltiniuose siūlomus metodus bei su jomis siejamus PĮ sprendimus.

1.1. Tyrimo sritis ir sprendžiamos problemos

Tyrimo sritis – semantinės web technologijos.

Šiuo metu daug dirbama bandant sukurti paieškos sistemas ir jas taip organizuoti, kad programiniai agentai suvoktų internete esančių resursų prasmę ir pagal tai atrastų reikiamus vartotojui resursus. Be to, video paskaitų sistemoje VIPS saugomi tūkstančiai mokymosi objektų, tarp kurių surasti norimą objektą arba visus susijusius pagal prasmę objektus vartotojui yra sudėtinga, ko pasėkoje mokymosi objektai dažnai lieka mažai naudojami, bei neretai yra dubliuojami.

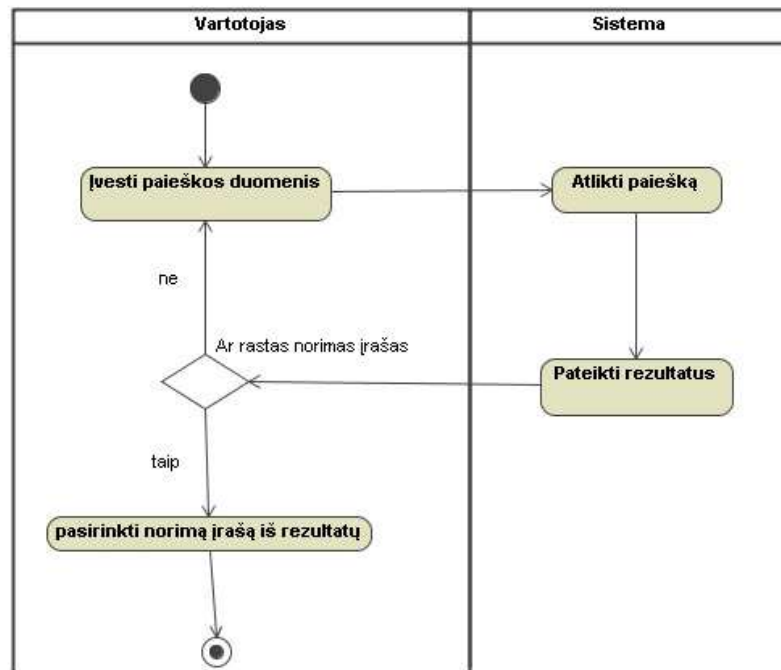
1.2. Tyrimo objektas

Paieškos efektyvumas priklauso tiek nuo turimos informacijos, aprašančios resursus, tiek nuo struktūros, kurioje ji saugoma, tiek nuo algoritmų, įgyvendinančių patį paieškos procesą.

VIPS sistemoje yra įgyvendinta pati paprasčiausia paieškos forma – proceso metu yra ieškomas paieškos žodis (-iai) tik resurso pavadinime, ir nors resurso pavadinimas gali gana tiksliai apibūdinti įrašo turinį, tačiau į patį turinį atsižvelgiama nėra.

Šio darbo tyrimo objektas - paieškos organizavimas virtualioje bibliotekoje pasitelkiant semantines web technologijas:

- duomenų apie resursus išgavimas ir saugojimas, aptarnavimas,
- paieškos proceso organizavimas.



1. pav. Paieškos veiklos proceso diagrama

1.3. Sistemos vartotojai

1.3.1. Vartotojų aibė, tipai ir savybės

- Vartotojų aibė video paskaitų sistemoje VIPS – tai asmenys, besimokantys nuotoliniu būdu arba teikiantys nuotolinio mokymosi paslaugas.
- Pagal tai skirstomi ir pagrindiniai vartotojų tipai:
- Dėstytojai – kuria kursus, susideda paskaitų skaidres, paveda paskaitas.
- studentai – dalyvauja paskaitoje.
- Kiti galimi sistemos vartotojai:
- Administratoriai – sistemos prižiūrėtojai, turintys visas teises.
- Asistentai – tai yra dėstytojų pagalbininkai, apribotas dėstytojo vartotojo variantas.
- Svečiai – svečiai yra labiau apribotų teisių studento tipo vartotojų variantas.

1.3.2. Vartotojų tikslai ir problemos

- Studento tikslas yra rasti reikiamą medžiagą ir ją įsisavinti.
- Dėstytojo tikslas yra pateikti medžiagą, praveisti paskaitą.
- Asistento tikslas yra pavaduoti dėstytoją techniniais klausimais.
- Administratoriaus pagrindinė veikla ir atsakomybė – stabilus sistemos veikimas, komunikacija su vartotojais ir pagalba jiems.

1.4. Problemos sprendimo metodų literatūros šaltiniuose analizė

1.4.1. Paieškos organizavimo galimybės

Atliekant analizę, buvo surasti keletas būdų kaip galima būtų organizuoti paiešką po virtualias mokymosi objektų bibliotekas:

- Organizuoti paiešką mokymosi objektus suskirsčius pagal semantinę prasmę.
- Leisti paiešką atlikti nepriklausomiems nuo bibliotekos paieškos varikliams.
- Organizuoti taikant folksonomiją.

1.4.2. Semantinis žiniatinklis ir paieška

Semantinis žiniatinklis yra mums įprasto žiniatinklio išplėtimas, kuriame tinklo turinys gali būti išreikštas ne tik tai gimtąja kalba, bet taip pat ir formatu, kuris gali būti perskaitytas ir panaudotas programinių agentų, tokiu būdu leisdamas jiems lengviau surasti bei organizuoti informaciją. Semantikos tinklas apima filosofiją, projektavimo principus, bendradarbiaujančias darbo grupes, ir įvairias įgalinančias technologijas. Kai kurie semantikos tinklo elementai yra išreikšti kaip numatomos ateities galimybės. Kiti semantikos tinklo elementai yra išreikšti oficialioje specifikacijoje, kai kurie iš jų: Išteklių Apibūdinimo Struktūrą (RDF), duomenų apskaitos formatų įvairovę (pavyzdžiui. RDF/XML, N3), ir žymėjimas ženklais toks kaip RDF Schema (RDFS) ir Tinklo Ontologijos Kalba (OWL), iš kurių visi yra numatyti, kad aprūpintų oficialų apibūdinimą sąvokų, terminų, ir santykių duotos žinių srities viduje.

Semantinio žiniatinklio tikslas yra padidinti interneto panaudojimo patogumą ir jo sujungtų išteklių naudingumą:

- Dokumentai, „pažymėti“ su semantikos informacija (iššesimas HTML KALBOS <Meta> žymės, panaudotos šiandieniuose tinklalapiuose, kad tiekėtų informaciją tinklo paieškos sistemoms, naudojančioms tinklo roplius). Tai galėtų būti mašinai

suprantama informacija apie žmogaus suprantamą dokumento turinį (toks kaip kūrėjas, pavadinimas, apibūdinimas, ir taip toliau) Semantikos žiniatinklis gali samprotauti apie gyvulius, žmones, vietas, mintis, ir taip toliau.) Semantikos žymėjimas yra dažnai kuriamas automatiškai, o ne rankomis.

- Bendri metaduomenų žodynai (ontologija) ir žemėlapiai tarp žodynų, kurie leidžia dokumento kūrėjams žinoti kaip žymėti dokumentus, kad programiniai agentai galėtų suprasti informaciją gautuose metaduomenyse (Pavyzdžiui, kad nemaišytų tinklalapio autoriaus su knygos, kuri pristatoma tinklalapyje, autoriumi)
- Automatizuoti agentai, kad įvykdytų užduotis semantikos tinklo vartotojams, naudodamas šiuos duomenis.
- Tinklu pagrįstos paslaugos (dažnai su savo programiniais agentais), kad tiekėtų informaciją specialiai agentams (pavyzdžiui, „Patikėta“ paslauga, kurios programinis agentas galėtų paklausti, jei kažkokia internetinė parduotuvė turi blogos paslaugos ar internetinio šiukšlinimo istoriją).

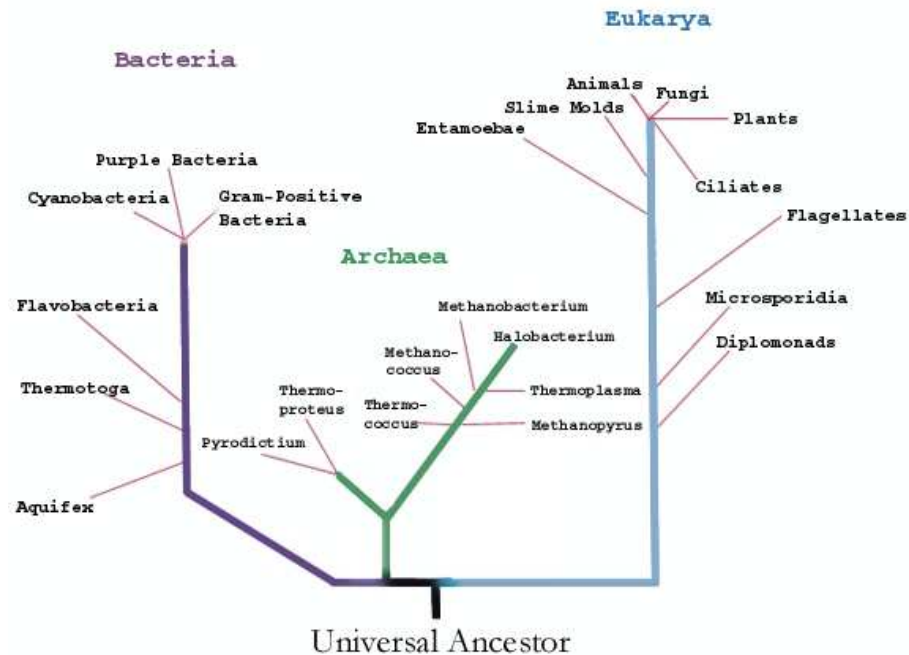
Paieška yra vienas iš populiariausių taikymų. Aiškios semantikos pridėjimas gali pagerinti paiešką. Semantinė paieška bando papildyti ir pagerinti tradicinius paieškos būdus.

Tradicinės informacijos paieškos technologija dažniausiai remiasi tikrai ieškomų žodžių atitikimo dokumentuose patikrinimu. Esant dideliame kiekiui suformuotos, kompiuteriui suprantamos informacijos apie plačią sritį objektų semantiniame tinkle leidžia išplėsti paieškos galimybes ir pagerinti tradicinę paiešką. [1]

Vienas iš galimų būdų spręsti paieškos problemas yra aprašyti objektus RDF schema arba OWL kalba. Tam reiktų sudaryti ontologiją, pagal kurias būtų galima suklasifikuoti objektus ir nustatyti ryšius tarp jų. Ontologija yra duomenų modelis, kuriame nurodomi konceptai tam tikrame žinių domene ir ryšiai tarp tų konceptų. Ontologiją sudaro:

- klasės: rinkiniai, kolekcijos ar objektų tipai,
- atributai, nustatymai, charakteristikos, parametrai, kuriuos turi objektai,
- ryšiai tarp objektų,
- įvykiai – atributų ar ryšių kitimas. [2]

Aprašius objektų savybes ir nurodžius ryšius pagal ontologiją, resursus galima suskirstyti hierarchiškai pagal kontekstinę prasmę. Pavyzdžiui, jei turėtume biblioteką apie bakterijas, tai hierarchiškai medžio tipo struktūroje jos ontologijos dokumentas galėtų atrodyti taip, kaip pavaizduota paveikslėlyje nr.2.[3]



2. pav. Bakterijos hierarchinėje struktūroje

Naudojantis semantiniiais metodais yra galimybės padaryti „protingą“ paiešką. Paprasčiausias būdas – vietoje to, kad atliktume paiešką ir pateiktume rezultatus paieščiui, rezultatai galėtų būti suklasifikuoti pagal prasmę. Taigi, jei ieškotume žodžio „Jaguaras“ galėtume gauti rezultatus, suklasifikuotus pagal tai ar jie apie automobilius, dideles kates, ar kitus objektus kartu. Ieškant informacijos pagal prasmę galima sujungti informaciją iš panašių dokumentų, panaikinti perteklių ir pateikti reziumė. Ryšiai tarp pagrindinių objektų dokumentuose gali būti pateikti vizualiai.[4]

VIPS sistemoje buvo įdiegta Dublin Core metaduomenų schema aprašyti video įrašams. Tai yra ontologija, pagal kurią galima aprašyti objektus, apibūdinant juos tokiais laukais:

- Pavadinimas (Title)
- Autorius (Creator)
- Pavadinimas (Subject)
- Apibūdinimas (Description)
- Publikuotojas (Publisher)
- Bendradarbiai(Contributor)
- Data (Date)
- Tipas (Type)
- Formatas (Format)
- Identifikatorius (Identifier)
- Šaltinis (Source)
- Kalba (Language)
- Ryšiai (Relation)

- Apimtis (Coverage)
- Teisės (Rights)

Programiniai agentai galėtų atsižvelgdami į šiuos metaduomenų laukus atlikti semantinę paiešką, taip pat, turint objektus, aprašytus šiais metaduomenimis, galima leisti paiešką atlikti išorinėms paieškos sistemoms. Tačiau bandymo metu gauti rezultatai parodė, kad šis sprendimas netenkina VIPS sistemos – nėra galimybių papildyti metaduomenis jau esančiai 4000 video įrašų bibliotekai.

1.4.3. Paieška po „gilųjį internetą“ (Deep web)

Virtualios bibliotekos priskiriamos „giliesiems interneto resursams“, nes paieškos sistemos tokios kaip Google arba Yahoo negali pasiekti jose slypinčios informacijos savo įprastais metodais, pasitelkę interneto vorus – programas kurios naršo po internetą pagal specialistų sukurtus algoritmus, automatiškai, aptikdami interneto nuorodas ir sekdami jomis toliau.[5] Šie interneto resursai nėra dalis pasaulinio interneto tinklo „paviršiaus“, kurį galėtų suindeksuoti naršyklės. Taip gali atsitikti, kai informacija yra pateikiama dinamiškai, atsakant į užklausą per formas (nėra tiesioginių nuorodų), reikalinga registracija ir prisijungimas, turinys pateikiamas per skriptus arba užkoduotas ne HTML kalba ar paprastu tekstu, o yra vaizdo įrašė arba paveikslėliuose.

Dauguma skaitmeninių bibliotekų turi paieškos sąsają, per kurią leidžia atlikti paiešką, taip pat turi specialius puslapius, arba svetainės žemėlapius (sitemap), pagal kuriuos paieškos varikliai gali atlikti paiešką. Skaitmeninės bibliotekos dažnai naudoja OAI-PMH[6] protokolą (Open Archives Initiative Protocol for Metadata Harvesting), kurio pagalba atskleidžia metaduomenis apie turimus resursus kitoms skaitmeninėms bibliotekoms ir paieškos naršyklėms kaip Google, Yahoo! Ir Scirus. Taigi paieškos problemą iš dalies galima spręsti perleidžiant paiešką atlikti interneto paieškos varikliams.

Tokios paieškos sistema labiau tinkama, kai norima atlikti paiešką per daugumą bibliotekų iš karto. Tuomet išorinis paieškos variklis apklausia bibliotekas ir gautus rezultatus gali sujungti, panaikinti perteklinę informaciją ar apibendrinti ir pateikti vartotojui.[7,8]

Darbo autorius dalyvaudamas ES FP6 projekte iCamp, VIPS sistemai realizavo SQI (Simple Query Interface) protokolo sąsają, per kuria galima gauti metaduomenis apie įrašus ir gauti nuorodas įrašams peržiūrėti.

Vartotojas yra suinteresuotas naudotis agreguojančiomis rezultatus paieškos sistemomis, nes viena ir ta pačia užklausa atliekama iš karto per visas paieškos sistemai prieinamas virtualias

bibliotekos bei rezultatai būna apjungti ir surikiuoti. Be to, tokia paieška išsaugotų laiką, kurį reiktų sugaišti kuriant paieškos variklį.

Norint, kad išorinės paieškos sistemos galėtų atlikti paiešką VIPS sistemoje reiktų įgyvendinti OAI-PMH protokolo sąsają, per kurią vyktų standartizuotas duomenų apsikeitimas. Šiuo metu išorinę paiešką su VIPS sistema galima atlikti per SQI (Simple Query Language) sąsają, kuria naudojasi to pačio iCamp projekto metu sukurtas rezultatų surinkimo variklis (mediatorius), lygiagrečiai atliekantis paiešką keletose skirtingų mokymosi objektų saugyklų kaupiamų įvairiuose Europos universitetuose ir iš karto pateikiantis agreguotus rezultatus. Pavyzdžiui, mokymosi objektų biblioteka A suformuotą sistemos viduje užklausą transformuoja ją pagal susitartą užklausų kalbą tarp saugyklų ir per SQI sąsają pasiunčia mokymosi objektų saugyklai B. Čia atėjusi užklausa transformuojama į sistemos viduje naudojamų užklausų formatą ir įvykdoma. Rezultatai gražinami pagal tą patį principą mokymosi objektų bibliotekai A.

Tačiau tokia paieška turi trūkumų, jei resursai nėra pilnai apibūdinti metaduomenimis. VIPS sistemoje yra įrašų, kuriuos galima apibūdinti tik sukurtimo data, pavadinimu ir autoriumi. Tačiau tai nėra pakankama, jeigu paieškos metu būtų ieškoma pagal apibūdinimą, turinį arba kitus parametrus. Be to, atliekant išorinę paiešką prarandami duomenys, pagal kuriuos būtų galima personalizuoti vartotojo aplinką VIPS sistemoje - paieškos rezultatai ir vartotojo pasirinkti peržiūrėti video įrašai. Šiuos duomenis taip pat būtų galima panaudoti ateityje tobulinant paieškos sistemą.

Atliekant bandymus su VIPS sistemoje esančia SQI sąsaja gauti rezultatai parodė, kad pirmiau reiktų spręsti metaduomenų trūkumo problemą.

1.4.4. Paieška per vartotojų žymes

Paiešką galima organizuoti per raktinius žodžius – vartotojų žymes, kuriomis vartotojai pažymi objektus. Toks procesas vadinamas Folksonomija (Folksonomy), dar žinomas kaip kolaboratyvus žymėjimas, socialinė klasifikacija, socialinė indeksacija. Folksonomijos rezultatas – daug vartotojų pažymi tą patį objektą ta pačia žyme, pagal kuria patys ir atlieka paiešką. Jei dauguma vartotojų pasirinko konkrečią žymę apibūdinti objektą, didelė tikimybė, kad kiti vartotojai, atlikę paiešką pagal tą žymę, gautų geriausią rezultatą.[9]

Pliusai – visa interneto resursų žymėmis paremta klasifikacija yra atliekama pagal vartotojus, kurie supranta resurso turinį, kas yra priešinga programiniams agentams, kurie algoritmiškai stengiasi suteikti resursui prasmę. Tai suteikia semantiškai suklasifikuotas žymes. Be to, skirtingai nuo tradicinio indeksavimo, metaduomenys yra ne tik sugeneruojami ekspertų, bet taip pat turinio

kūrėjų ir vartotojų. Dažniausiai naudojami laisvai pasirenkami raktiniai žodžiai vietoje kontroliuojančio žodyno.

Kita vertus vartotojai paprastai žymi objektus kai nori juos surasti antrą kartą ir vėliau, nes jiems šie objektai yra naudingi. Kuo daugiau vartotojų atranda, kad jiems naudingas konkretus objektas, tuo didesnis rangas tam objektui suteikiamas ir paieškos metu turi prioritetą eilėje su kitais surastais objektais. [10]

Socialinio žymėjimo didelis privalumas yra tas, kad folksonomija evoliucionuoja taip pat kaip ir žmonių kalba per naudojimą ir sąveiką, todėl negali „pasenti“.

Minusai:

- Nėra žymių standartinio rinkinio (kas taip pat yra vadinama kontroliuojančiu žodynu)
- Nėra žymių standartinės struktūros (ta pati žymė gali būti parašyta tiek didžiosiomis, tiek mažosiomis raidėmis, tiek daugiskaita, tiek vienaskaita, skirtingais linksniais ir pan.)
- Gali būti neteisingai pažymėtos padarant rašybos klaidas
- Žymės žodis gali turėti daugiau nei vieną reikšmę (sinonimai/antonimai),
- naudojamas ne pagal tradicinę reikšmę - personalizuotos žymės tam tikrai vartotojų grupei.
- Nėra mechanizmo pagal kurį vartotojai galėtų nurodyti hierarchinius ryšius tarp žymių (pvz resursas gali būti apibūdintas kaip „transporto priemonė“ arba kaip „automobilis“, nors „automobilis“ yra „transporto priemonės“ subšaka).[11,12]

Kiti minusai – vartotojai gali specialiai žymėti tam tikrus objektus daug kartų su populiariomis žymėmis, kad kiti vartotojai galvotų, jog tuos objektus verta peržiūrėti, kad paieškos varikliai būtų apgaunami. [13]

Folksonomija gali prisidėti prie semantinio žiniatinklio kūrimo. Tačiau yra sunku numatyti kaip didelė ir įvairi publika galėtų būti priversta pridėti metaduomenis, žymes interneto resursams nuosekliu, patikimu būdu, o autoriams patiems pridėti metaduomenis yra brangus ir daug laiko reikalaujantis dalykas. Palyginus su labiau formalizuota klasifikacija naudojant „kontroliuojančius žodynus“, folksonomija yra paskirstyta klasifikacija su mažomis palaikymo išlaidomis.[14]

1.5. Panašių sistemų analizė

Sculpteur5 (<http://www.sculpteurweb.org>) – naudoja semantinę technologiją multimedijos informacijos valdymui. Paslauga nukreipta į muziejus. Ontologijos su tam tikrais įrankiais buvo

sukurtos, pagal kurias yra apibrėžiami ir apibūdinami objektai, taip pat interneto voras ieško trūkstamos informacijos internete.

DELOS (<http://www.delos.info/>) – pagrindinis tikslas yra sukurti tolimesnės kartos skaitmeninių bibliotekų technologijas. DELOS projektas suskirstytas į septynis modulius, iš kurių vienas - „Žinių išgavimas ir semantinis bendradarbiavimas“. Pagrindinis šio modulio tikslas yra semantinis bendradarbiavimas tarp daugelio žodynų skaitmeninėje erdvėje.

DILIGENT (<http://diligentproject.org/>) – naudojamos Grid technologijos, kurios padeda susikurti virtualias dinamiškas skaitmenines bibliotekas virtualioms „eScience“ bendruomenėms.[15]

del.icio.us - interneto tinklalapis [16,17] yra socialinių interneto adresų žymėjimas interneto adresų saugojimui, dalinimuisi ir paieškai. Jo pagalba galima:

- išsisaugoti interneto nuorodas į kitus interneto resursus (straipsnius, blogus, muzikos, apžvalgų, naujienų ir pan. ir pasiekti šias nuorodas iš bet kurio kompiuterio, prijungto prie interneto),
- dalintis adresais į resursus su kitais interneto vartotojais,
- atrasti naujų dalykų, kuriuos jau pažymėjo kiti vartotojai.

Šioje sistemoje interneto adresų organizavimui ir saugojimui yra naudojamos žymės, kas yra daug lankstesnė sistema palyginus su katalogizavimu dedant nuorodas į iš anksto sugalvotus aplankus ar hierarchijas. Žymės interneto adresams vartotojai pasirenka laisvai. Tam tikras adresas, sužymėtas skirtingų vartotojų tuo metu turi jam priskirtus tagus, kurie tam adresui suteikia prasmę ir pagal tai galima rasti adresą paieškoje. Kuo daugiau vartotoju uždeda tam tikrą žymę vienam ir tam pačiam adresui, tuo labiau galima tikėtis, kad žymėjimas yra tikslus, taip pat tai parodo adreso populiarumą ir vartotojų domėjimosi kryptį.

Flickr yra nuotraukų pasidalinimo internetinis portalas. šį portalą interneto dienoraščių rašytojai (bloggeriai) naudoja kaip savo nuotraukų saugyklą. Populiarumą nulėmė inovatyvūs internetinės komunikacijos įrankiai kurie leidžia nuotraukas sužymėti ir atlikti paiešką (folksonomijos prasme). Flickr leidžia žmonėms, įkeliantiems nuotraukas, jas kategorizuoti naudojantis žymėmis. Flickr taip pat leidžia vartotojams kategorizuoti savo nuotraukas į „rinkinius“, grupes, kur nuotraukos yra po vienu pavadinimu. Tai yra daug lanksčiau nei tradicinis skirstymas į aplankus, nes viena ir ta pati nuotrauka gali priklausyti vienam nuotraukų rinkiniui, keliems, arba iš viso nei vienai (Konceptija yra analogiška Google kompanijos Gmail „etikečių“ principui). Flickr „rinkiniai“ reprezentuoja kategorinius metaduomenis labiau nei fizinę hierarchiją.

1.6. Darbo tikslas ir siekiami privalumai

Darbo tikslas – realizuoti paieškos posistemį ir palengvinti vartotojui įrašų radimo galimybes.

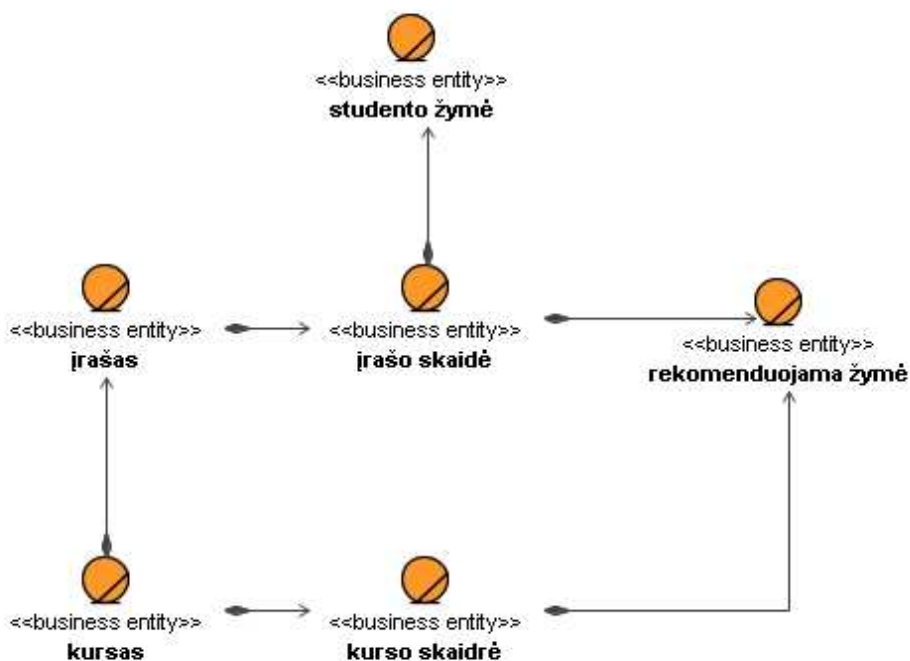
Siekiami privalumai:

- Vartotojas galėtų pasižymėti įrašą jam patinkančia žyme ir vėliau pagal tą žymę surasti tą įrašą.
- Atliekant paiešką, būtų galima ieškoti ne tik pagal pavadinimą, bet ir pagal turinį.
- Paieškos vienas matmenų – kiek vartotojų pažymėjo tą patį įrašą ta pačia žyme. Tai rodytų, jog dauguma vartotojų sutaria dėl konkrečios žymės konkrečiam kursui ir atliekant paiešką pagal tą žymę rezultatai būtų tikslesni ir pirmesni.
- Vartotojas turi savo žymių sąrašą, kas personalizuotų jo prisijungimo vartotoją, ir parodo jo domėjimosi sritį, bei pagal tai galima vartotojui pasiūlyti jį galimai dominančius kitus įrašus.

2. Folksonomija grįstas paieškos modelis

Socialinis žymėjimas galėtų išspręsti duomenų trūkumo problema, kadangi įrašų apibūdinimas nuo ekspertų būtų perkeltas vartotojams. Taip būtų išvengta reikalavimo samdyti ekspertus, o vartotojai galėtų patys žymėti video įrašus. Tai taip pat galėtų būti panaudota personalizuojant vartotojų aplinką.

VIPS sistemoje yra galimybė išvengti socialinio žymėjimo trūkumų – kontroliuojančio žodyno nebuvimo, nes vartotojams būtų leidžiama ne tik laisvai įvesti savo norimas žymes, bet būtų rekomenduojamos žymės, kurios būtų ištrauktos iš skaidrių. Kadangi VIPS projektas nėra komercinis, išlaidos jo palaikymui turi būti minimalios, todėl sprendimas leisti vartotojams paties žymėti įrašus būtų vienas iš geresnių sprendimų, o rekomenduojamos žymės neleistų nuklysti. Rekomenduojamas žymes dėstytojas ar administratorius galėtų redaguoti, pridėti savų. Taip būtų išspręsta senų įrašų apibūdinimo problema, nors senesnių įrašų žymės, nesant dideliame vartotojų populiarumui ir dėstytojo, pravedusių tų įrašų paskaitos susidomėjimui, gali būti ne visiškai tikslios.



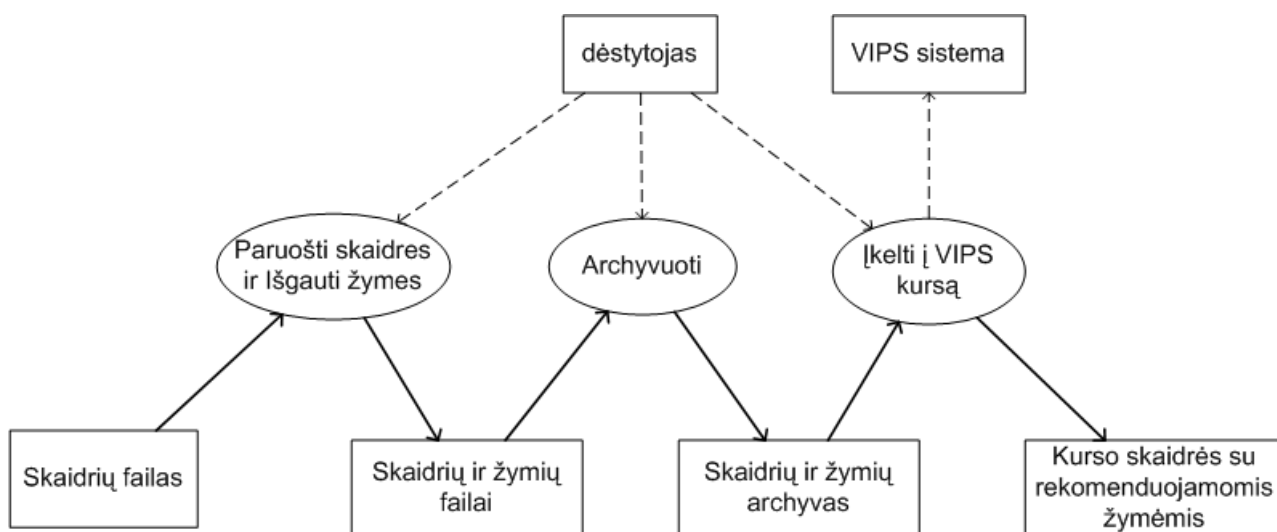
3. pav. Verslo proceso modelį sudarančių elementų priklausomybės

2.1. Rekomenduojamų žymių išgavimas

Rekomenduojamos žymės būtų gaunamos iš dėstytojo sukeltų skaidrių teksto. Šiuo metu nėra aiškių standartų aprašyti duomenų semantikai taip, kad galima būtų nestructūrizuotuose duomenyse programiškai aptikti reikalingą informaciją semantine prasme.[18]

Žemiau esančiame paveikslėlyje pagal prof. Remigijaus Gusto[19] metodiką nubraižytas skaidrių sukėlimo ir žymių išgavimo veiklos modelis:

- Paruoštas skaidrių failas (ppt. formatu) būtų apdorojamas specialia programa, kuri paruošia skaidres įkėlimui į VIPS sistemą. Tuo pat metu būtų sugeneruojamas XML failas, kuriame išsaugota skaidres aprašanti informacija – skaidrių pavadinimai ir skaidrių tekstas.
- Paruošti skaidrių failai ir aprašantysis failas būtų suarchyvuojami ir įkeliami į VIPS sistemą.
- VIPS sistema įkėlimo metu išsaugotų skaidres serveryje, susietų su pavadinimais ir išfiltruotų rekomenduojamas žymes iš skaidrių teksto.

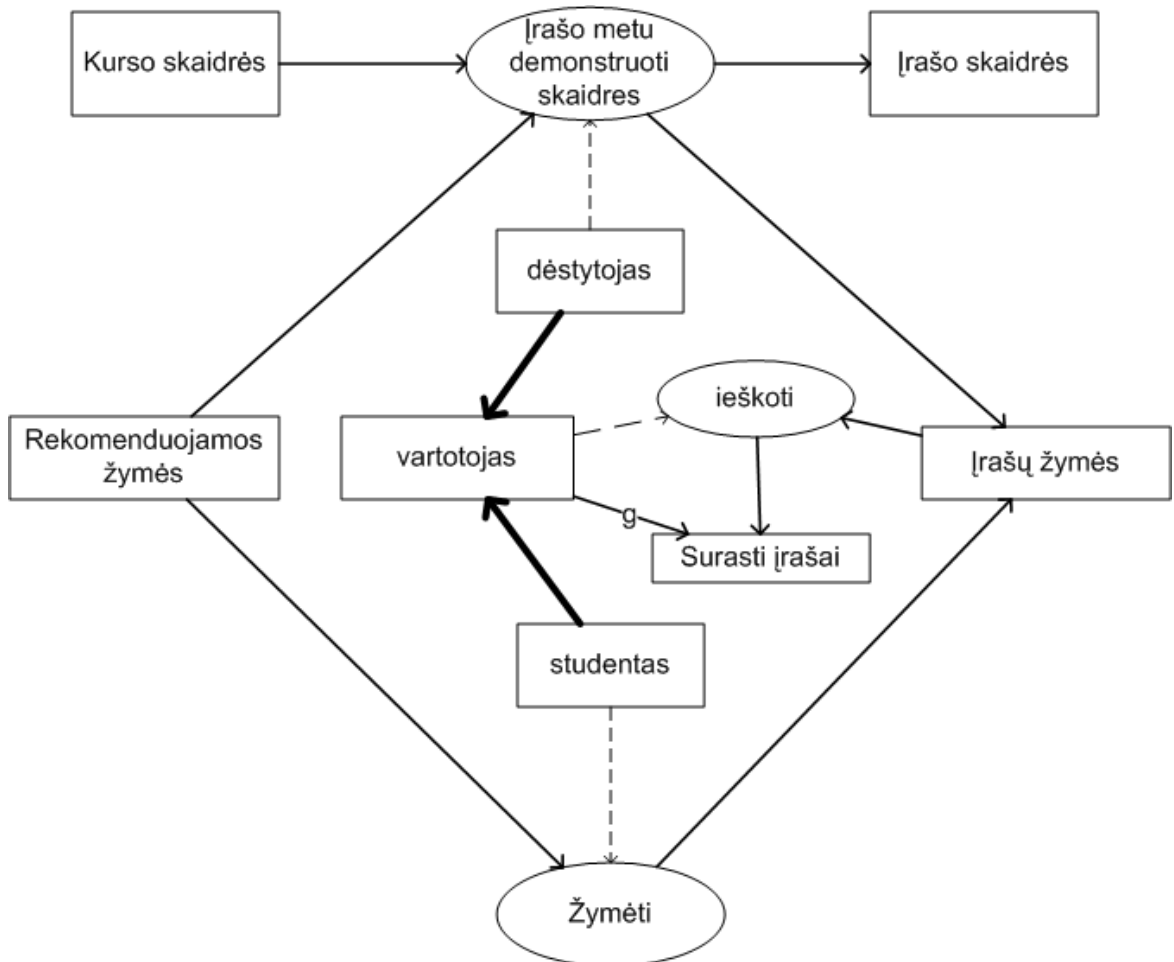


4. pav. Skaidrių sukėlimo ir žymių išgavimo veiklos modelis

2.2. Įrašų žymėjimas ir paieška

Įrašo metu demonstruojamos skaidrės būtų saugomos kaip įrašo skaidrės, o prie kurso rekomenduojamos žymės būtų saugomos kaip įrašo žymės. Tos pačios žymės būtų įkeliamos tik vieną kartą, kadangi įrašo peržiūros metu studentas turėtų turėti galimybę gauti rekomenduojamas žymes.

Žymių suteikimo įrašams veiklos modelyje nurodyta, kad rekomenduojamos žymės vieną kartą būtų įtraukiamos į įrašo žymes, todėl paieškos proceso metu būtų užtikrinta, kad paieška būtų vykdoma per video įrašų turinį. Paieška apimtų didesnę kiekį duomenų ir tikėtina, duotų geresnius paieškos rezultatus.



5. pav. Žymių suteikimo įrašams ir paieškos veiklos modelis

2.3. Apibendrinimas

Socialinio žymėjimo metodas papildo informaciją apie įrašus ir įgalina paiešką. Taip pat yra personalizuojama vartotojų prisijungimo erdvė. Rekomenduojamos žymės palengvina vartotojui pasirinkti žymę įrašui, todėl vartotojui turėtų būti didesnė motyvacija žymėti įrašus ir taip prisidėti prie paieškos sistemos, pagrįstos įrašų žymėjimu, tobulinimo.

3. Paieškos posistemis video paskaitų sistemoje VIPS

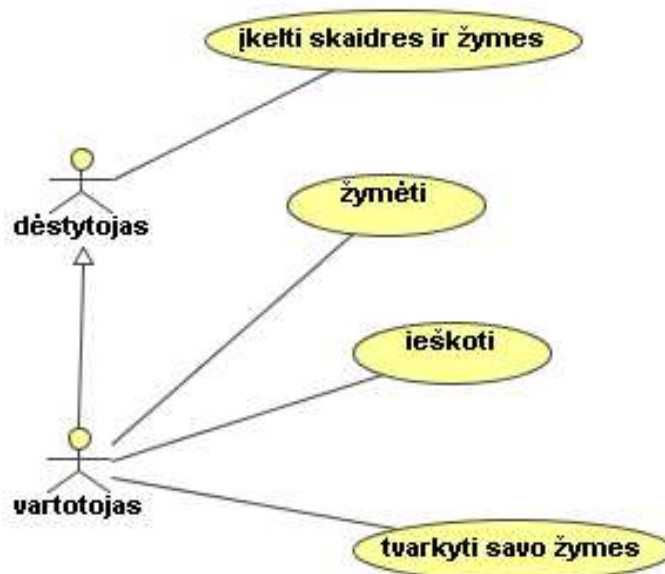
3.1. Reikalavimai

3.1.1. Siekiamos paieškos sistemos apibrėžimas

Siekama, kad dėstytojai įkeldami skaidres, kartu įkeltų ir žymių tekstą. Apdorojus tekstą būtų gautos rekomenduojamos žymės, kurias dėstytojas arba administratorius galėtų pageduoti, pridėti, išmesti. Vartotojai, žiūrėdami įrašą ar dalyvaudami paskaitoje galėtų pažymėti skaidrę rekomenduojama žyme arba pridėti savo. Sąraše jie matytų žymes, kuriomis atitinkama skaidrė buvo daugiausia kartų pažymėta. Vėliau atlikdami paiešką, jie galėtų vesti žymę, pagal kurią būtų ieškomas įrašas ir jiems būtų siūloma rinktis iš jau esančių. Taip pat jie turėtų savo personalizuotą aplinką, kur būtų matomi jų paskutiniai lankyti kursai, žymių sąrašas ir rekomenduojami jiems peržiūrėti įrašai, kurie galėtų juos dominti.

3.1.2. Kompiuterizuojamos funkcijos

Kompiuterizuotos VIPS funkcijos yra leisti vartotojams apibūdinti pasirinktą skaidrę savo žymėmis ir pasinaudoti paieška vėliau ieškant įrašų. Taip pat yra kompiuterizuojamas automatinis skaidrių žymėjimas pagal jose esantį tekstą.



6. pav. Kompiuterizuojamos funkcijos

3.1.3. Panaudojimo atvejų specifikacijos

Kiekvienas vartotojo panaudojimo atvejis specifikuotas pagal specifikavimo lentelės šabloną, kurioje nurodomi pagrindiniai panaudojimo atvejo parametrai: pavadinimas, sąlyga prieš, sąlyga po, susiję panaudojimo atvejai, pagrindinis įvykių srautas ir sistemos reakcija. Šių lentelių informacija papildo vartotojo PA modelį. Panaudojimo atvejų specifikacijos pateiktos lentelėse

1 lentelė. PA “Įkelti skaidres ir žymes” specifikacija

PA „Įkelti skaidres ir žymes“		
Prieš sąlyga		Vartotojas identifikuotas Prisijungta prie konkretaus kurso
Sužadinimo sąlyga		Vartotojas nori įkelti skaidres
Susiję panaudojimo atvejai	Išplečia PA	
	Apima PA	
	Specializuoja PA	
Pagrindinis įvykių srautas		Sistemos reakcija ir sprendimai
1. Vartotojas pateikia skaidres		Skaidrės yra apdorojamos, nuskaitomos žymės ir gražinamas skaidrių ir žymių aplankas
2. Vartotojas nurodo suarchyvuotą sistemos gražintą aplanką		Sistema įkelia skaidres ir žymes į duomenų bazę
Po sąlyga		Skaidrės sukeltos į serverį, žymės į duomenų bazę.

2 lentelė. PA „Žymėti” specifikacija

PA „Žymėti“	
Prieš sąlyga	Vartotojas identifiukuotas Prisijungta prie konkretaus įrašo
Sužadinimo sąlyga	Vartotojas nori žymėti įrašą
Susiję panaudojimo atvejai	Išplečia PA
	Apima PA
	Specializuoja PA
Pagrindinis įvykių srautas	Sistemos reakcija ir sprendimai
1. Vartotojas pasirenka žymę iš rekomenduojamų žymių	1. Sistema užregistruoja vartotojo pasirinktą žymę, skirtą konkrečiai skaidrei.
2. Vartotojas įveda savo žymę	2. Sistema užregistruoja vartotojo įvestą žymę, skirtą konkrečiai skaidrei.
3. Vartotojas redaguoja savo žymę	3. Sistema registruoja naują žymę ir ištrina seną
4. Vartotojas atsisako rekomenduojamos žymės	4. Sistema atšaukia vartotojo registruotą žymę
Po sąlyga	Įrašo skaidrė pažymėta vartotojo žyme.

3 lentelė. PA „Ieškoti” specifikacija

PA „Ieškoti“	
Prieš sąlyga	Vartotojas identifiukuotas
Sužadinimo sąlyga	Vartotojas nori surasti įrašą
Susiję panaudojimo atvejai	Išplečia PA
	Apima PA
	Specializuoja PA
Pagrindinis įvykių srautas	Sistemos reakcija ir sprendimai
	1. Išvedamos vartotojo žymės
1. Pasirinkti savo žymę arba įvesti	2. Atlikama paieška
Po sąlyga	Pateikiami rezultatus

3.1.4. Reikalavimai duomenims

Duomenų saugyklai ir joje esantiems duomenims keliami reikalavimai:

- Duomenų suderinamumo. Duomenų saugyklą realizuojančios duomenų bazės schema turės atitikti kuriamos programinės įrangos duomenų struktūrą.
- Duomenų saugumo. Saugykloje esantys duomenys negali būti iškraipomi pašalinės programinės įrangos ar vartotojų. Saugumas yra viena iš pagrindinių priemonių užtikrinti duomenų integralumą.

3.1.5. Nefunkciniai reikalavimai ir apribojimai

3.1.5.1. Reikalavimai sąsajai

Reikalavimai vartotojo sąsajai yra išdėstyti 4 lentelėje.

4 lentelė. Reikalavimai vartotojo sąsajai

Reikalavimas	Paaiškinimas
Sąsajos panaudojamumas	Sąsaja turi būti intuityvi.
Funkcinių reikalavimų realizavimas	Visi funkciniai reikalavimai turi būti realizuoti vartotojo sąsajoje. Negali būti ne grafinėje vartotojo sąsajoje realizuotų reikalavimų.

3.1.5.2. Reikalavimai kokybei

Reikalavimai kokybei yra išdėstyti 5 lentelėje.

5 lentelė. Reikalavimai kokybei

Reikalavimas	Paaiškinimas
Saugumas	Programa neturi atlikti neleistinių veiksmų operacinėje ir failų sistemose. Su sistema nesusijusios programinės įrangos darbai sistemos veikla negali turėti jokios įtakos, išskyrus racionalų dalinimąsi resursais.
Patikimumas	Programa turi dirbti stabiliai, operacijų rezultatai turi atitikti vartotojų laukiamus rezultatus.

3.1.6. Rizikos veiksniai

Rizikos veiksnių apžvalga yra pateikta 6 lentelėje.

6 lentelė. Projekto rizikos faktoriai

Rizikos veiksnys	Metodai rizikai sumažinti
Vartotojui nesuprantama ar nepriimtina sąsaja	Bendradarbiaujant su užsakovu sukurti vartotojo sąsajos prototipą.
Pasikeitę arba nauji reikalavimai	Prieš įvedant reikalavimų pakeitimus būtina įsitikinti pokyčių būtinumu. Jei pakeitimai yra motyvuoti, būtina juos dokumentuoti.

3.1.7. Rezultato kokybės kriterijai

Rezultato kokybės kriterijai yra išdėstyti 7 lentelėje.

7 lentelė. Programinės įrangos kokybės kriterijai

Kokybės kriterijus	Aprašas
Vartotojo sąsajos kokybė	Vartotojo sąsaja turi būti intuityvi ir suprantama visiems dalykinę sritį išmanantiems vartotojams.
Nefunkcinių reikalavimų realizacija	Programinė įranga turi atitikti iškeltus stabilumo, saugumo reikalavimus.
Funkcinių reikalavimų realizacija	Programinė įranga turi įgyvendinti visus projektavimo metu nustatytus funkcinius reikalavimus ir jų pakeitimus.

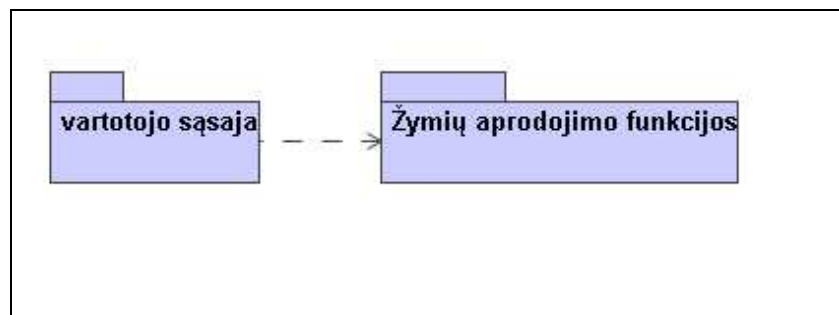
3.2. Projektas

3.2.1. Statinė struktūra

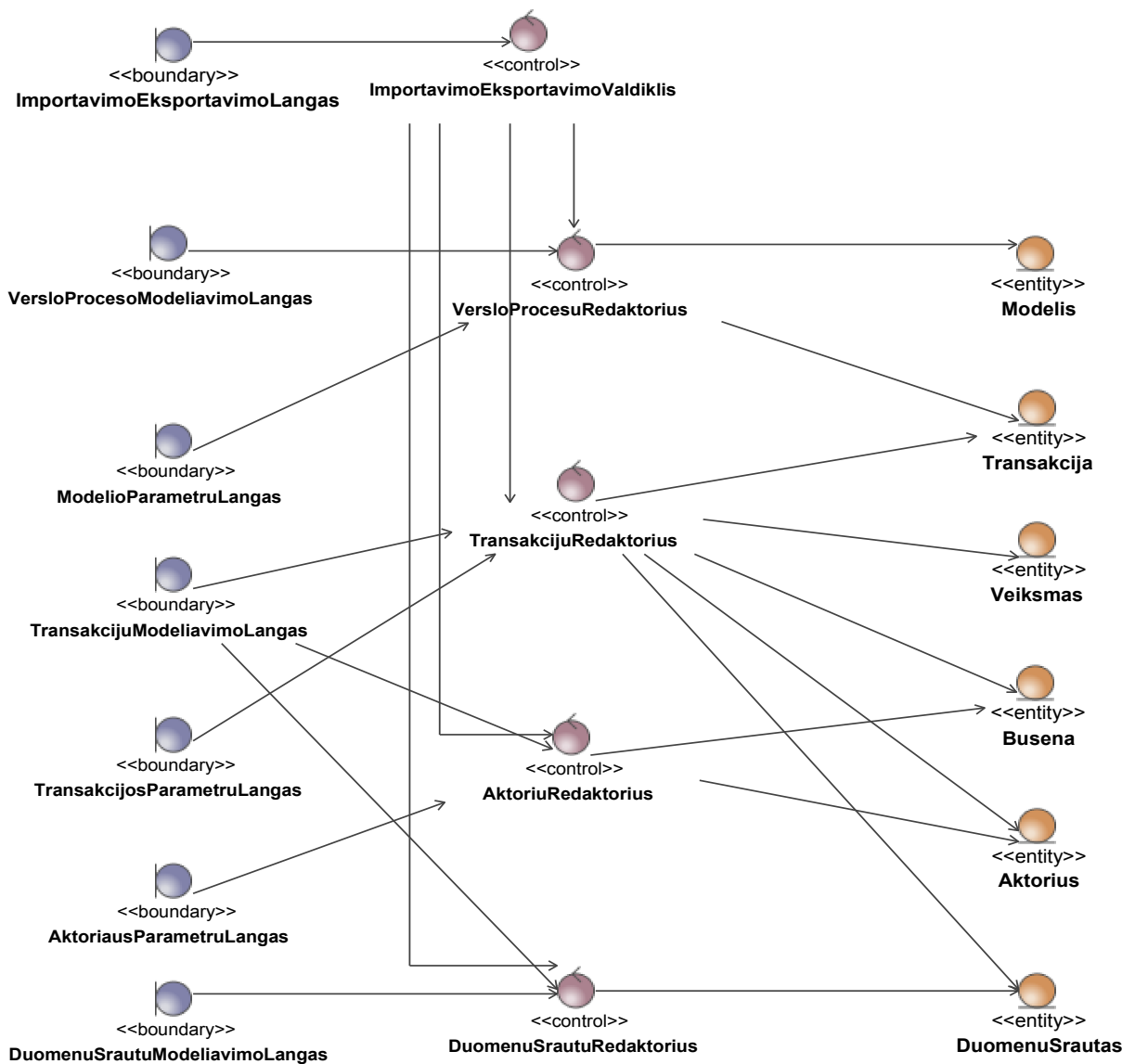
Šiame poskyryje pateikiama kuriamos programinės įrangos statinė struktūra: bendra loginė architektūra, vartotojo paslaugų klasių diagrama; klasių, teikiančių veiklos paslaugas, diagrama; klasių, realizuojančių veiklos esybes, diagrama.

3.2.1.1. Loginė architektūra

Loginės architektūros modelis (7 pav.) parodo, kaip sistemoje bus paskirstyta funkcijų realizacija į paketus ir kokios yra paketų tarpusavio priklausomybės.



7. pav. Loginės architektūros modelis



8. pav. Analitinis modelis

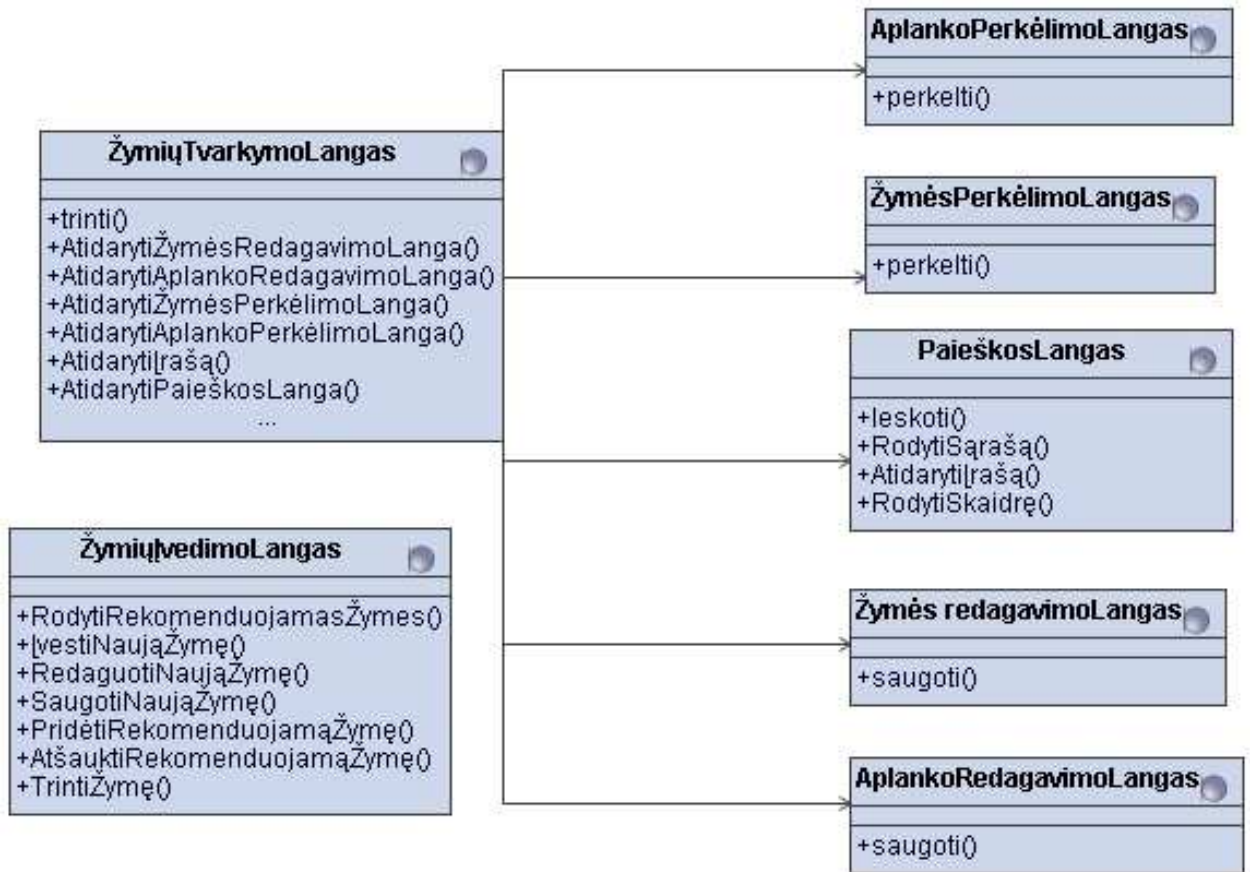
Analitiniame modelyje yra identifikuojami galimi sistemos objektai. Šie objektai yra skirstomi į grupes pagal atliekamas funkcijas ir atsakomybes.

Įprastai skiriamos šios klasių grupės:

1. Ribinės klasės. Naudojamos, modeliuojant sąveikas tarp sistemos išorės ir vidinių klasių. Ribinės klasės perduoda sistemai išorinius įvykius ir grąžina sistemos pranešimus
2. Valdikliai. Atlieka koordinuojantį vaidmenį ir valdo kitas klases. Valdikliai koordinuoja kitas klases ir perduoda joms įvykius iš ribinių klasių.
3. Esybės. Vaizduoja informaciją, kuri turi būti saugoma ir apdorojama. Esybių klasės vaizduoja dalykinės srities objektus.

3.2.1.2. Vartotojo paslaugų objektai

Klasių, teikiančių vartotojo paslaugas, diagrama yra pateikta 9 pav. Šia diagrama parodomos vartotojo sąsają realizuojančios klasės ir jų tarpusavio priklausomybės.

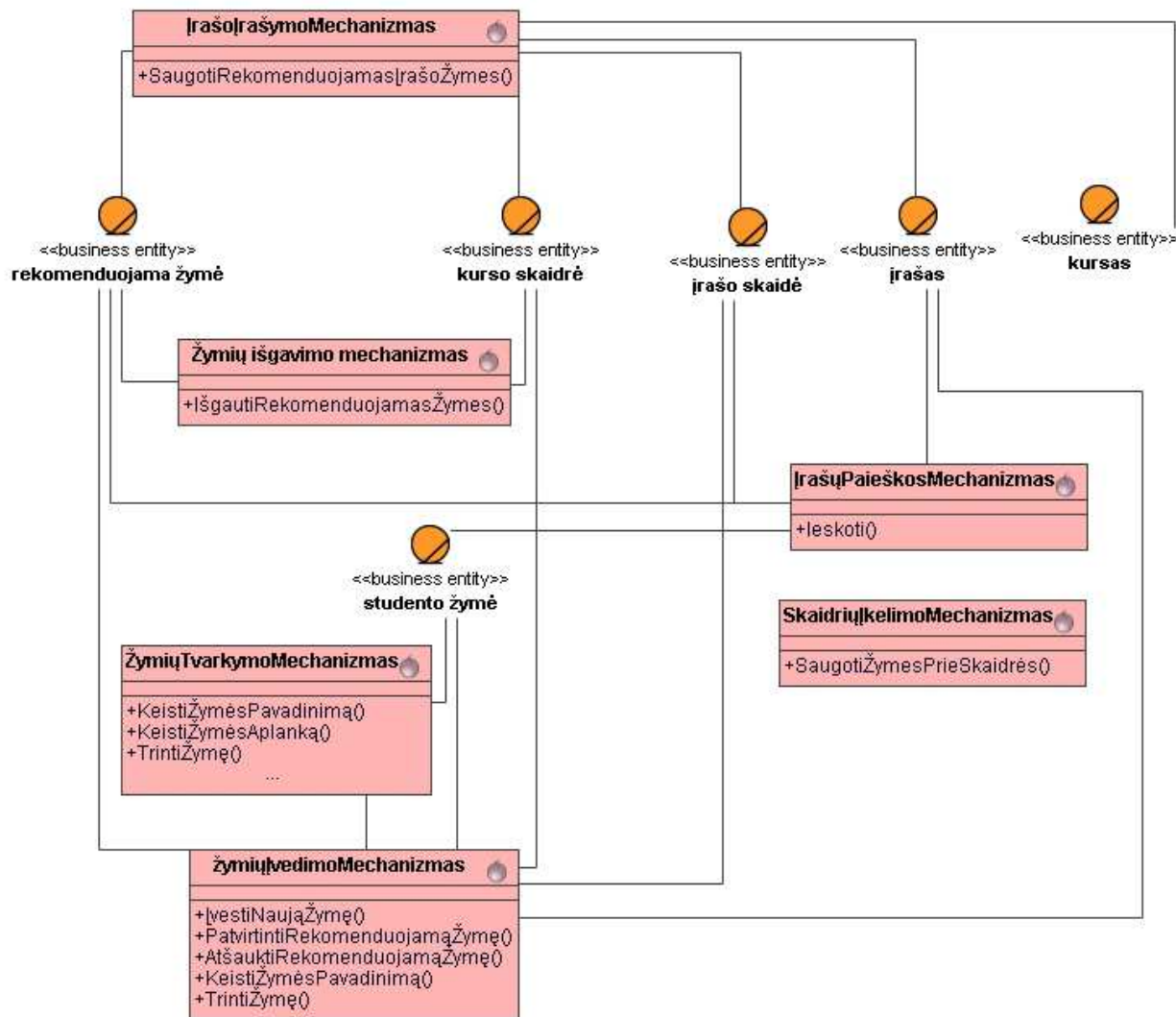


9. pav. Vartotojo sąsajos klasių modelis

Kadangi modelyje specifikuotomis klasėmis bus realizuoti vartotojo sąsajos langai, paveiksle esančioje diagramoje specifikuotus klasių ryšius galima traktuoti kaip galimus navigavimo po grafine vartotojo sąsają kelius.

3.2.1.3. Veiklos logikos objektai

Veiklos paslaugų klasių diagramoje (10 pav.) yra specifikuojamas redaktoriaus logikuos lygmuo.



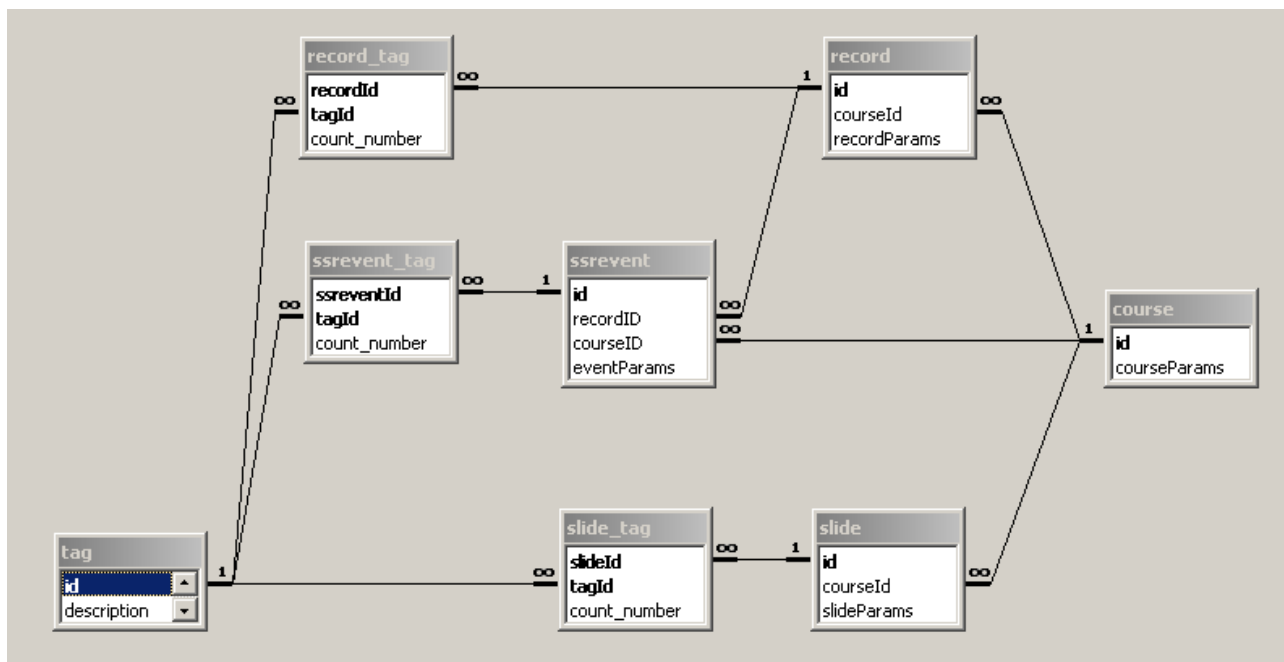
10. pav. Veiklos paslaugų klasių modelis

Redaktoriaus logikos lygmens klasės yra suskirstytos į dvi grupes:

1. valdiklių klasės;
2. duomenų klasės.

Toks klasių skirstymas atskiria vartotojo sąsają nuo duomenų klasių ir gerai atitinka MVC architektūrą. Šio sprendimo privalumas – galimybė modifikuoti vartotojo sąsają ir teikiamą funkcionalumą nekeičiant duomenų klasių. Analogiškai, pagrindinis trūkumas – minimalūs duomenų modelio pokyčiai gali reikalauti esminių pakeitimų vartotojo sąsajoje ar verslo logikos klasėse.

3.2.1.4. Duomenų bazės architektūra

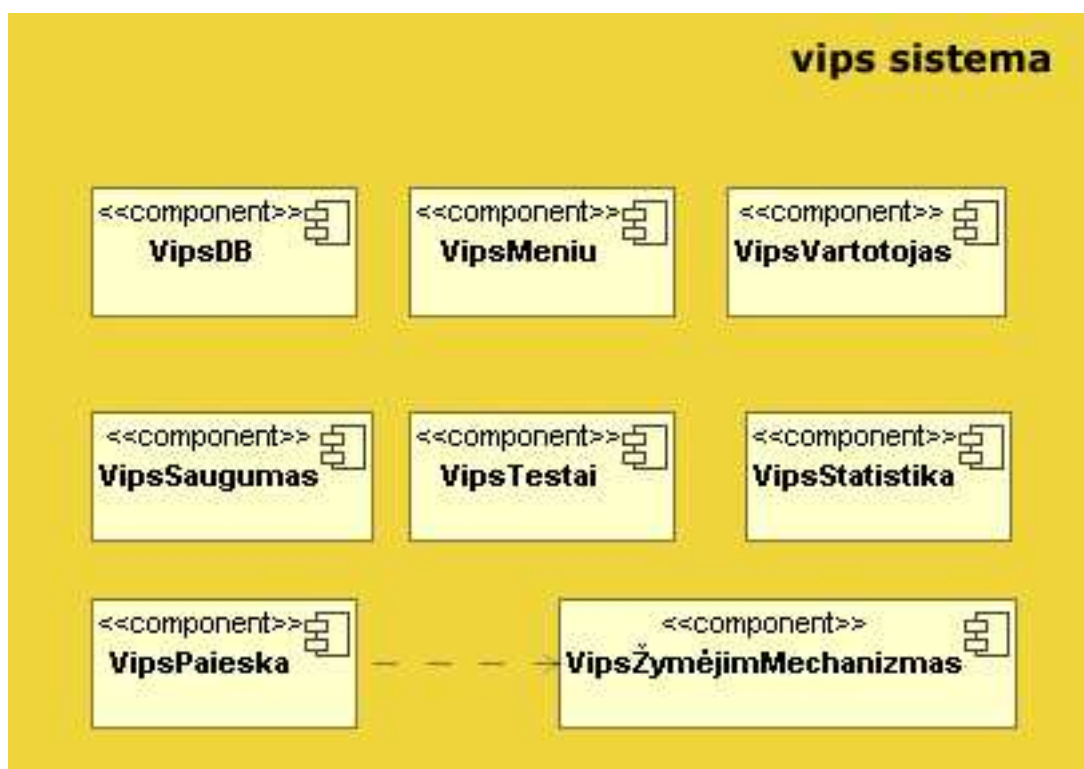


11. pav. Duomenų bazės schema

3.3. Sistemos realizacija

3.3.1. Komponentų diagrama

Komponentų diagramoje pavaizduoti kuriamo įrankio pagrindiniai komponentai ir ryšiai tarp jų. Išvedimas ir vykdomos operacijos yra paremtos Asinchronine Javascript technologija (AJAX), kad neerzintų vartotojo ir padidintų interaktyvumą bei naudojimosi paprastumą.[20,21]



12. pav. Komponentų diagrama

3.3.2. Sistemos testavimo planas

8 lentelė. Nefunkcinių reikalavimų testavimo planas

Testas	Veiksmų seka	Laukiamas rezultatas
Operacinės sistemos testas	Paleidžiamas Linux aplinkoje panaudojus Apache.	Sistema veikia be klaidų.
Vartotojo naršyklės testas	Prie sistemos prisijungiamas naudojant IE naršyklę. Tikrinama vartotojo sąsaja. Prie sistemos prisijungima naudojant FireFox naršyklę. Tikrinama vartotojo sąsaja.	Vartotojo sąsajoje klaidų neaptikta.
Apsauga nuo kodo injekcijos	Į sistemos kliento dalį patalpinamas žalingas kodas.	Žalingas kodas neįvygdytas.

Testavimo plano veiksmų seka ir laukiamas rezultatas funkciniam testavimui dar nesudaryti dėl prototipo realizacijos statuso.

9 lentelė. Funkcinio testavimo planas

Testas	Veiksmų seka	Laukiamas rezultatas
Skaidrių ir žymių įkėlimas		
Skaidrių pašalinimas		
Užklausa paieškai		
Skaidrės žymėjimas		
Žymių redagavimas		

4. Žodynėlis

Sąvoka	Paiškinimas
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting. Metaduomenų perdavimo protokolas
AJAX	(angl. Asynchronous Javascript And Xml) Web puslapių kūrimo technologija.
PI	Programinė įranga
URL	(angl. Unique resource identifier) Protokolas vardų specifikuojimui internete.
RDF	W3C specifikacija duomenų modeliui ir komponentams
RDFS	RDF schema yra plečiama žinių atvaizdavimo kalba, suteikianti pradinis elementus ontologijų apibūdinimui, kitaip vadinama RDF žodynais, skirtais struktūrizuoti RDF resursus
OWL	Web ontologijų kalba, skirta apibrėžti ontologijas
iCamp	Mokslo vystymo projektas, remiamas Europos Sąjungos pagal FP6 programą
SQI	(angl. Simple Query Language) Protokolas metaduomenų apsaugai
Deep web	„Gilusis internetas“ – pasaulinio interneto išteklių, nepasiekiami paieškos sistemoms standartiniais būdais
N3	(angl. Notation 3) RDF modelis, kurio sintaksė specialiai kurta, kad būtų patogesnė skaityti žmogui.
MVC	(angl. Model-view-controller) architektūrinis šablonas, naudojamas programinėje inžinerijoje.
W3C	Organizacijos sutrumpinimas, atsakingos už pasaulinio žiniatinklio standartus

5. Išvados

Semantinių web technologijų tyrimui, kaip objektas buvo pasirinkta KTU Distancinio mokymo centre naudojama Video paskaitų sistema VIPS kurioje per 6 jos eksploatacijos metus jau buvo sukauptas didelis vaizdo įrašų archyvas, kas leido tyrimą atlikti su realiais duomenimis. Sukurti papildomi moduliai VIPS sistemoje leido papildyti aprašymus struktūrizuotais meta-duomenimis panaudojant Dublin Core standarto schemą bei nestruktūrizuotais metaduomenimis – vartotojų pasirinktomis žymėmis. Apibendrinant darbą galima pateikti tokias išvadas:

1. Paieškos organizavimo būdų analizė parodė, kad VIPS sistemoje egzistuoja metaduomenų trūkumo problema, todėl buvo pasirinkta organizuoti paiešką pagal žymes.
2. Siekiant palengvinti žymėjimą ir sušvelninti socialinio žymėjimo minusus (kontroliuojančio žodyno nebuvimas), nuspręsta vartotojams pateikti rekomenduojamas žymes.
3. Prisijungę vartotojai prie VIPS sistemos įgytų galimybę žymėti objektus ir taip personalizuoti savo prisijungimo aplinką, tuo pačiu būtų galima gauti tikslesnę naudojimosi VIPS sistema statistiką.
4. Projektavimui buvo panaudota CASE priemonė MagicDraw UML 12.0, kuri palengvino darbą.

Ateityje tobulinant sistemą reikėtų koncentruotis į šiuos uždavinius:

1. Filtruoti žymes, išskiriant reikšminius žodžius bei pagal juos surasti susijusius mokymosi objektus.
2. Personalizuoti aplinką, pateikiant vartotojui aktualius įrašus pagal jo suvestas žymes arba artimus įrašus.
3. Panagrinėti galimybės išgauti metaduomenų informaciją iš garso arba video resursų.

6. Literatūra

1. R. Guha, Rob McCool, Eric Miller. Semantic Search. Prieiga per internetą: <http://www2003.org/cdrom/papers/refereed/p779/ess.html> [2007.03.21]
2. W3C. Web Ontology Language. Prieiga per internetą: <http://www.w3.org/TR/owl-features/> [2007.05.25]
3. Vikipedija, Laisvoji enciklopedija. Prieiga internete: http://en.wikipedia.org/wiki/Semantic_web [2007.08.14].
4. Heiner Stuckenschmidt, Frank van Harmelen. Information Sharing on the Semantic Web, Springer, 2003
5. Vikipedija, Laisvoji enciklopedija. Prieiga internete: http://en.wikipedia.org/wiki/Deep_web [2007.08.17].
6. Vikipedija, Laisvoji enciklopedija. Prieiga per internetą: <http://en.wikipedia.org/wiki/OAI-PMH> [2007.08.19]
7. Vikipedija, Laisvoji enciklopedija. Prieiga internete: http://en.wikipedia.org/wiki/Federated_Search [2007.02.19].
8. Vikipedija, Laisvoji enciklopedija. Prieiga internete: http://en.wikipedia.org/wiki/Web_crawler [2007.08.21].
9. Thomas Vander Wal. [Explaining and Showing Broad and Narrow Folksonomies](#). Prieiga per internetą : http://www.personalinfocloud.com/2005/02/explaining_and_.html [2007.08.25]
10. Vikipedija, Laisvoji enciklopedija. Prieiga per internetą: http://en.wikipedia.org/wiki/Social_bookmarking [2007.03.18]
11. Marieke Guy Ukoln, Emma TonkinUkoln. Folksonomies:Tidying up Tags? Prieiga per internetą: <http://www.dlib.org/dlib/january06/guy/01guy.html> [2007.10.03]
12. Tony Hammond, Timo Hannay, Ben Lund, Joanna Scott. Social Bookmarking Tools (I): A General Review. Prieiga per internetą: <http://www.dlib.org//dlib/april05/hammond/04hammond.html> [2007.10.05]

13. Vikipedija, Laisvoji enciklopedija. Prieiga per internetą: <http://en.wikipedia.org/wiki/Folksonomy> [2007.10.13]
14. Nigel Shadbolt and Wendy Hall, Tim Berners-Lee. The Semantic Web Revisited. Prieiga per internetą: http://eprints.ecs.soton.ac.uk/12614/1/Semantic_Web_Revisted.pdf [2007.10.13]
15. J. Davie's, R. Studer, P. Warren. Semantic Web Technologies, John Wiley & Sons Ltd, 2006, ISBN-10: 0-470-02596-4.
16. Socialinio žymėjimo svetainė del.icio.us. Prieiga per internetą: <http://del.icio.us> [2007.10.15]
17. Vikipedija, Laisvoji enciklopedija. Prieiga per internetą: <http://en.wikipedia.org/wiki/Del.icio.us> [2007.10.18]
18. Aurelija Kazlauskienė. Semantinių metaduomenų aprašymo ir efektyvaus panaudojimo problema. Informacinės technologijos'2006, 2006.
19. Gustas R. Pragmatinių ir dinaminių priklausomybių integravimo principai: Monografija. Kaunas: Technologija, 1997.
20. Vikipedija, Laisvoji enciklopedija. Prieiga per Internetą: <http://en.wikipedia.org/wiki/AJAX> [2007.11.06]
21. Bruce W. Perry Ajax Hacks, O'Reilly March 2006,438 p. ISBN 0-596-10169-4

7. Priedai

7.1. Programos prototipo langai

Viršuje (13 pav.) yra paieškos laukas, kuriame galima formuoti paieškos užklausą. Užklausa formuojama iš žymių, jei norima ieškoti pagal keleta žymių, tuomet žymės yra sujungiamos pliuso ženklu, kitais atvejais bus ieškoma žymė, sudaryta iš dviejų žodžių, atskitų tarpu.

Surasti įrašai yra demonstruojami ekrane, medžio tipo lentelėje, iš karto po įrašo rodant to įrašo skaidres, tokiu būdu iš karto galima pažiūrėti ar konkretus įrašas yra susijęs su vartotoją dominančia tema.

Dešinėje pusėje yra „asmeninių žymių“ sritis – tai vartotojo panaudotos žymės pažymint konkrečius įrašus.

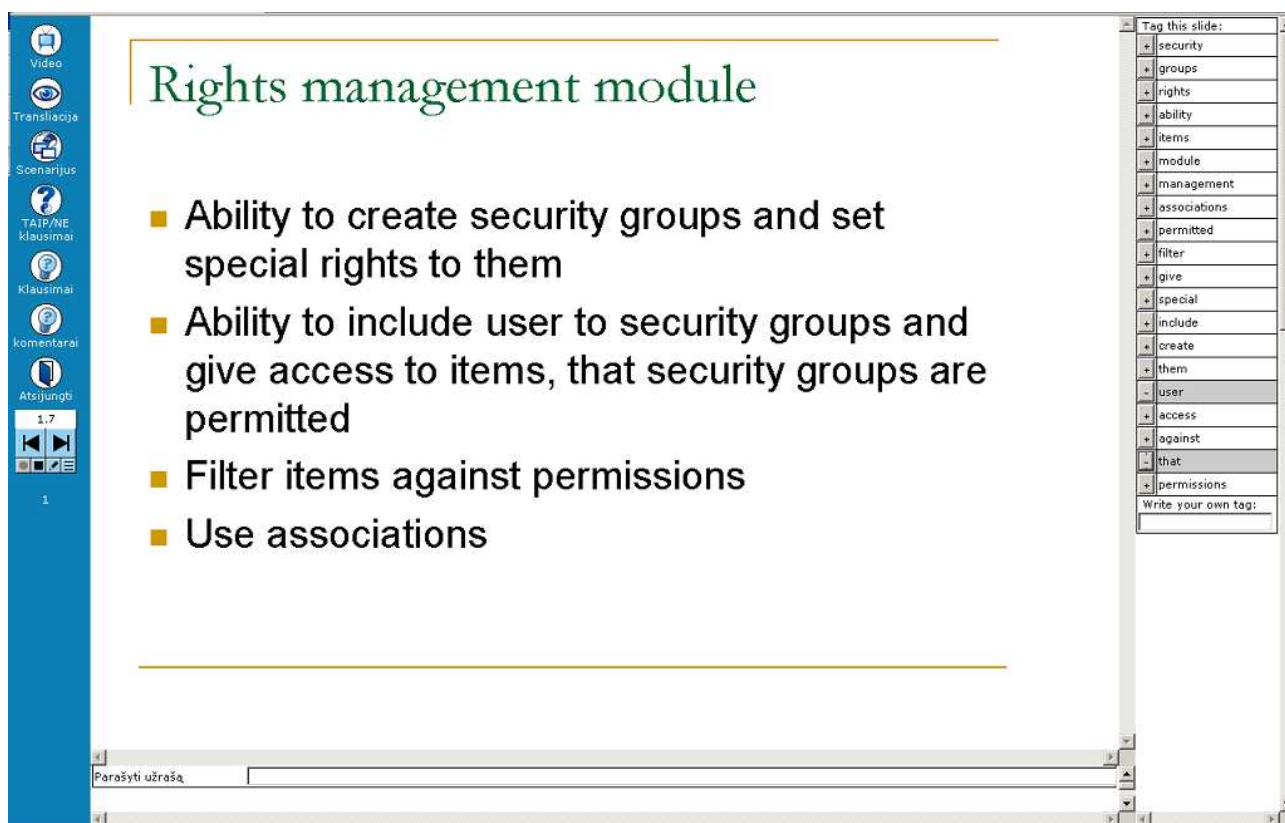
Ieškoti:

Įrašai	Tags Valdymas
<input type="checkbox"/> G.Cibulskis. Vaizdo ir garso apdorojimo ir perdavimo technologijos - paskaitų transliavimas VIPS (II dalis).	<input type="checkbox"/> <input type="checkbox"/>
<input type="checkbox"/> D. Rutkauskienė, G. Cibulskis. Šeštoji paskaita - Sinchroninio mokymo technologijos	<input type="checkbox"/> <input type="checkbox"/>
<input type="checkbox"/> G. Cibulskis. VK dalyvių atsakomybės	<input type="checkbox"/> <input type="checkbox"/>
<input type="checkbox"/> G.Cibulskis. VK metodikos pagrindai. 2. VK administraciniai aspektai - dalyvių atsakomybės	<input type="checkbox"/> <input type="checkbox"/>
<input type="checkbox"/> G. Cibulskis, A. Varneckas (KTU DMC). Sinchroninio mokymosi sistemų plėtros tendencijos	<input type="checkbox"/> <input type="checkbox"/>
VIPS integravimas su LieDM	<input type="checkbox"/>
Išplėstinė konfigūracija	<input type="checkbox"/>
Pradžią	<input type="checkbox"/>
VIPS konfigūracijos	<input type="checkbox"/>
Galimos VIPS konfigūracijos	<input type="checkbox"/>
Bazinė konfigūracija	<input type="checkbox"/>
Privalumai & trūkumai	<input type="checkbox"/>
Privalumai & trūkumai	<input type="checkbox"/>
Minimalios konfigūracijos privalumai & trūkumai	<input type="checkbox"/>
Rekomendacijos VIPS diegimui	<input type="checkbox"/>
Plėtros tendencijos	<input type="checkbox"/>
Technologinė plėtra	<input type="checkbox"/>
Prielaidos	<input type="checkbox"/>
Naujas projektas	<input type="checkbox"/>
Kontaktai	<input type="checkbox"/>
<input type="checkbox"/> G.Cibulskis_Sinchroninio_mokymo_tehnologijos	<input type="checkbox"/> <input type="checkbox"/>
<input type="checkbox"/> G.Cibulskis. Sinchroninio mokymo technologijos	<input type="checkbox"/> <input type="checkbox"/>
<input type="checkbox"/> Mokymosi technologijų integracija	<input type="checkbox"/> <input type="checkbox"/>
<input type="checkbox"/> 7. G.Cibulskis (KTU DMC). VIPS funkcinių galimybių plėtra	<input type="checkbox"/> <input type="checkbox"/>

Personal Tags:

13. pav. Paieškos langas

Žymėjimo lange (14 pav.) vaizduojama skaidrė apie kuria kalba dėstytojas. Dešinėje pusėje yra šios skaidrės rekomenduojamos žymės, iš kurių vartotojas gali pasirinkti arba įvesti savo apačioje esančiame laukelyje.



The image shows a presentation slide titled "Rights management module" in green text. The slide contains four bullet points:

- Ability to create security groups and set special rights to them
- Ability to include user to security groups and give access to items, that security groups are permitted
- Filter items against permissions
- Use associations

On the right side of the slide, there is a "Tag this slide:" panel with a list of tags: security, groups, rights, ability, items, module, management, associations, permitted, filter, give, special, include, create, them, user, access, against, that, permissions. Below the list is a text input field labeled "Write your own tag:". The slide is part of a presentation interface, with a blue sidebar on the left containing icons for Video, Transliacija, Scenarijus, TAIP/NE klausimai, Klausimai, komentarai, and Atsijungti. At the bottom of the slide, there is a search bar labeled "Porošyti užrašą" and a navigation bar with a "1" indicator.

14. pav. Žymėjimo langas

Metaduomenų redagavimo lange kursui pagal Dublin Core metaduomenų aprašymo standartą galima aprašyti metaduomenis: pavadinimą, apibūdinimą, temą ir t.t.

Edit Broadcast metadata

[LT] [EN]

Broadcast Metadata

[Title]	[Source]	[Creator]	[Rights]	[Provenance]
[Subject]	[Relation]	[Publisher]	[Date]	[RightsHolder]
[Description]	[Coverage]	[Contributor]	[Audience]	[Free Form]

Title

European Conference for Sustainable Water Management

? [\[Top\]](#)

Subject

Water Management

? [\[Top\]](#)

Description

Leonardo da Vinci project KNOWLEDGE BRIDGE is pleased to announce the European Conference for Sustainable Water Management. This supreme two-day event in Vilnius, Lithuania is dedicated to the potentials and the future of sustainable water management in Europe. Main topic is the challenges to implement the European Water Framework Directive and the solutions and benefits that online training and the European Water-Online-Network are offering for these tasks. You are invited to become a member of the growing European community Water-Online-Network (WON) of water experts for sustainable water management!

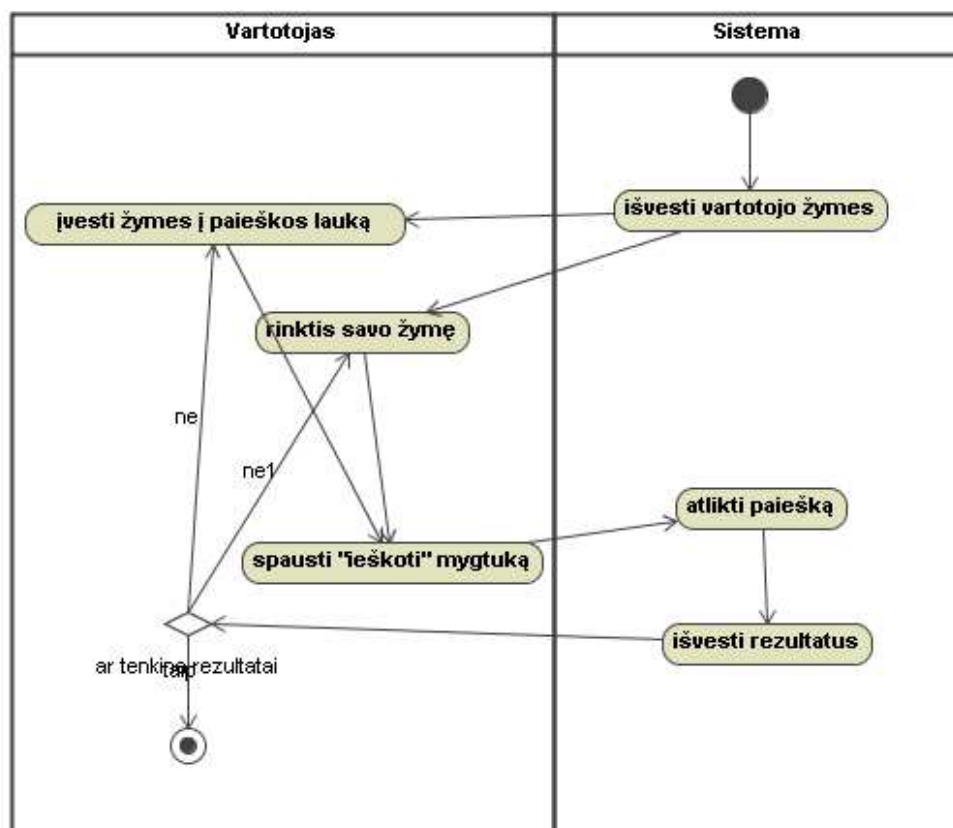
? [\[Top\]](#)

Source

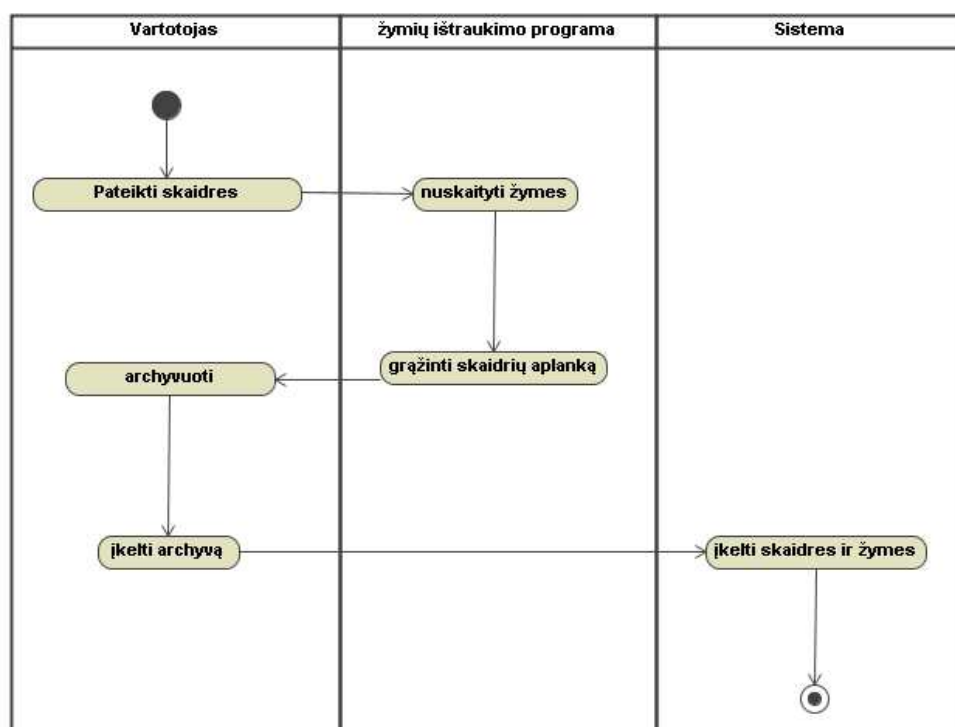
No entries in this table

15. pav. Metaduomenų redagavimo langas

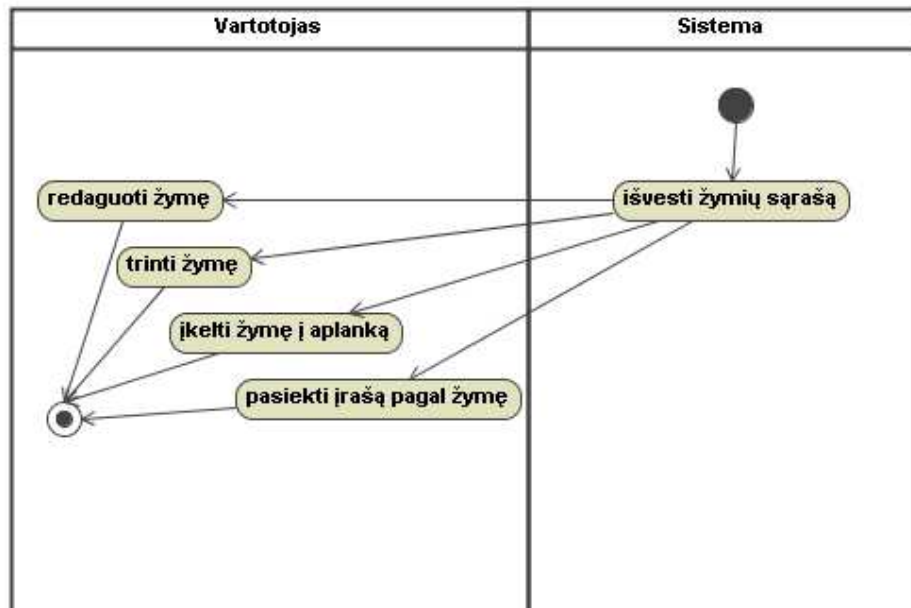
7.2. Panaudojimo atvejai



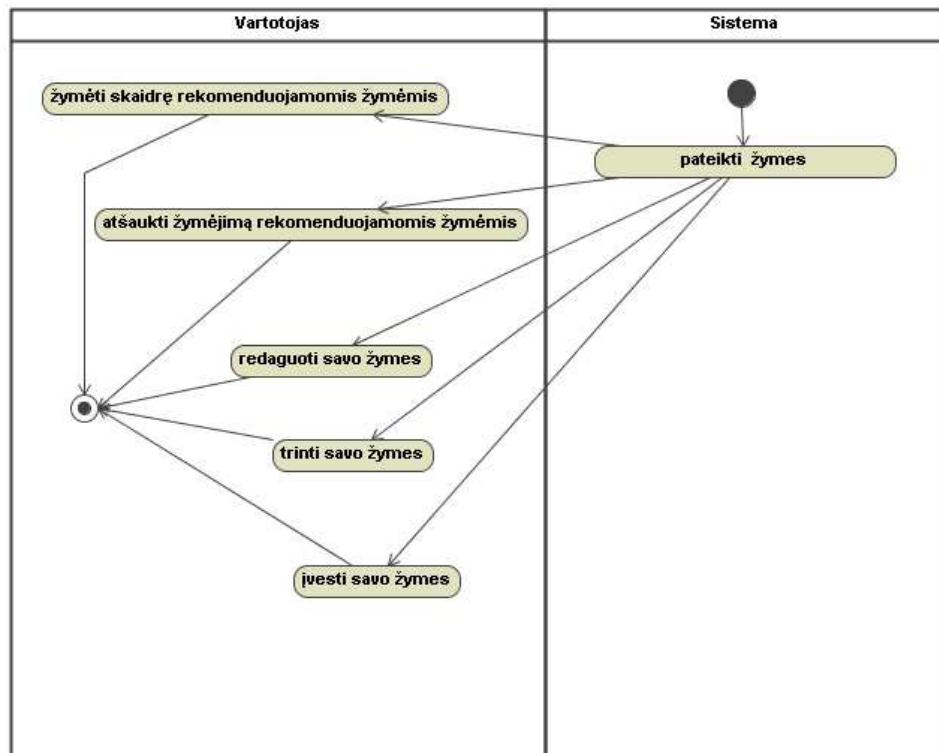
16. pav. Panaudojimo atvejis „Ieškoti“



17. pav. Panaudojimo atvejis „Įkelti skaidres ir žymes“



18. pav. Panaudojimo atvejis „tvarkyti savo žymes“ (Manage my notes)



19. pav. Panaudojimo atvejis „žymėti“ (Mark)

7.3. Straipsnis

Semantinių Web technologijų tyrimas ir taikymas Video paskaitų sistemoje VIPS

Evaldas Karazinas, Gytis Cibulskis

Kauno technologijos universitetas, Studentų g. 48a-312, Kaunas LT-5136

Pranešime pristatomos vaizdo paskaitų sistemos VIPS išplėtimo priemonės. Pagal Europos Sąjungos remiamą FP6 IST projektą iCamp atliekamas Europos universitetų skaitmeninių bibliotekų tinklo kūrimas, tuo siekiant išplėsti besimokančiųjų auditoriją bibliotekoms bei padidinti naudojimosi efektyvumą. Analizuojama naudojama technologija virtualios bibliotekos praplėtimui bei aptariami galimi paieškos virtualių bibliotekų tinkle sprendimai panaudojant semantines web technologijas.

20. Įvadas

Jau šeši metai Kauno technologijos universiteto Distantinio mokymo centre (DMC) pagal tarptautinę Eureka programą vykdomi moksliniai taikomieji tyrimai plėtojant interaktyvias mokymo priemones. Mokslo ir studijų fondo finansuotų Eureka projektų TESTVIL ir TESTVIL II metu buvo sukurta ir išplėta vaizdo paskaitų sistema VIPS, kurios pagalba galima internetu transliuoti interaktyvias paskaitas, konferencijas ar seminarus, realiu laiku vykdyti transliuojamų renginių įrašymą bei automatiškai kaupti įrašų archyvą. Sistema naudojasi ne tik Lietuvos nuotolinio mokymosi tinklo LieDM institucijos bet ir vienas didžiausių Lietuvos bankų, seniausias Europoje Bolonijos universitetas bei keletas kitų tarptautinių partnerių. KTU DMC serveryje įdiegtoje VIPS sistemoje jau yra sukauptas daugiau kaip 2,5 tūkst. įvairių renginių įrašų archyvas. Šis archyvas gali būti naudojamas kaip skaitmeninių mokymosi resursų saugykla išplečiant jų panaudojimą tiek formaliame tiek ir neformaliame švietime. Norint efektyviai panaudoti VIPS sistemoje sukauptus mokymosi resursus reikalinga patogi paieška bei galimybė gauti kuo daugiau informacijos apie saugykloje esančius įrašus. Tokiai paieškai realizuoti yra būtina išplėsti informaciją apie saugomus įrašus įdiegiant standartizuotą meta-duomenų schemą. Kadangi dalis įrašų yra apsaugoti slaptažodžiais viešai prieinamos paieškos rezultatuose galima vaizduoti tik užklausą atitinkančius meta-duomenis, o priėjimą prie pačio resurso suteikti tik autorizuotiems vartotojams. Vykdamas kitą Europos Sąjungos FP6 IST programos finansuojamą projektą "iCamp - Inovatyvi, viską apimanti, interaktyvi ir tarpkultūrinė mokymosi aplinka (Innovative, inclusive, interactive and intercultural learning campus)" yra numatoma sukurti Europos universitetų skaitmeninių mokymosi objektų saugyklų ir virtualių bibliotekų tinklą įgalinantį atlikti informacijos paiešką iš karto visose tinklui priklausančiose saugyklose. Tam reikia išspręsti technologines jungimo problemas bei sukurti tikslią paieškos sistemą, leidžiančią greitai dideliame duomenų kiekyje surasti reikiamą informaciją.

21. Skaitmeninių objektų saugyklų tinklo sukūrimas

21.1. iCamp projektas

Pagrindinis iCamp projekto tikslas yra sukurti atviras virtualias mokymosi aplinkas aukštajam mokslui Europoje sujungiant skirtingas mokymosi sistemas. iCamp nekuria naujos eMokymosi sistemos, bet palaiko esamų sistemų plėtimą ir bendradarbiavimą [2]. Tam buvo pasirinkta sukurti papildomus modulius virtualioms bibliotekoms bei skaitmeninių mokymosi objektų saugykloms, kurios tarpusavyje bendrautų panaudojant SQI specifikacija paremtus standartizuotus Web servisus.

21.2. SQI bendradarbiavimo sąsaja

SQI (Simple Query Interface) – taikomųjų programų sąsaja (API), skirta atlikti užklausoms į mokymosi objektų bibliotekas. Siekiant heterogeninių bibliotekų bendradarbiavimo, ši sąsaja tenkina tokius reikalavimus[3]:

- Neutralumo – rezultatų ir užklausų formatai gali būti pasirinkami. Bendradarbiaujančios sistemos turi susitarti dėl:
 - atributų rinkinio ir žodyno, kuris bus naudojamas užklausose,

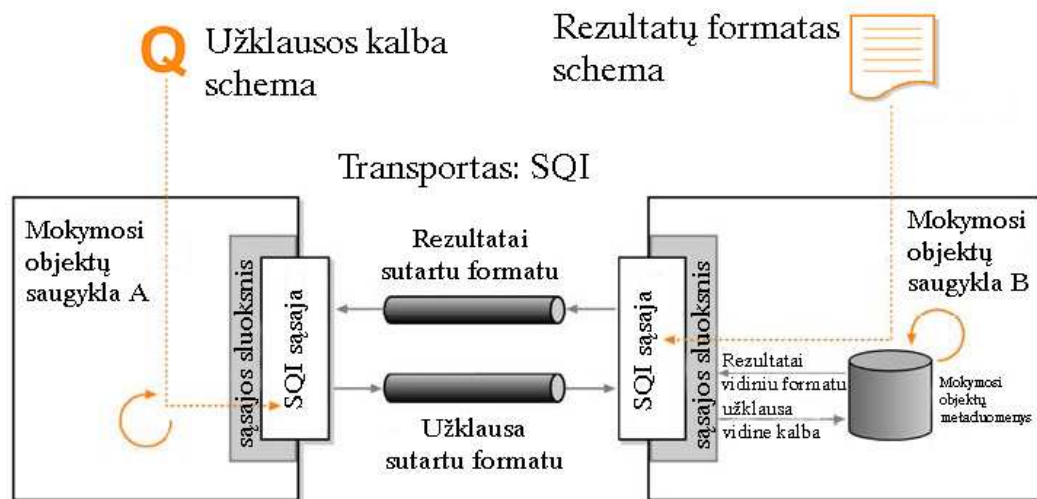
- užklauso kalbos,
- mokymosi objektų reprezentacijos
- meta-duomenų, skirtų mokymosi objektams, reprezentacijos

Kaip pavyzdys, toks susitarimas gali būti išreikštas XML arba RDF schemomis.

- Palaiko sinchronines ir asinchronines užklausas
- Gali dirbti išlaikant būseną arba be jos
- Paremta sesijų valdymo koncepcija tam kad būtų atskirtas autentifikavimas ir užklausių valdymas.
- Pats API suprojektuotas remiantis tokiais principais:
 - Komandos – užklauso atskyrimo principas
 - Paprastos komandos (su mažai formalijų parametru) ir išplečiamumas.

21.3. SQI komunikacija

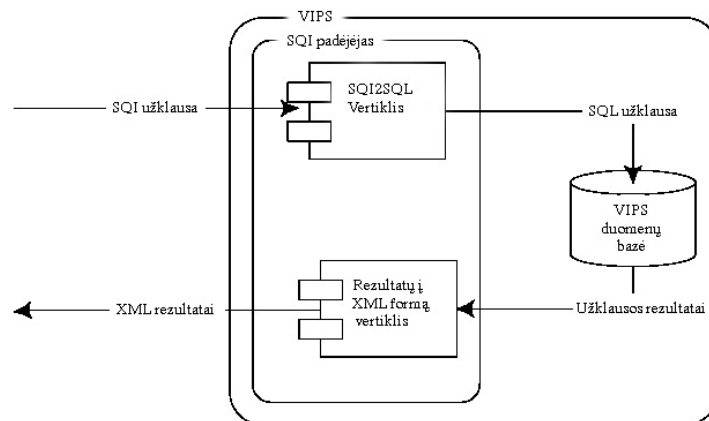
Mokymosi objektų biblioteka A suformuotą sistemos viduje užklausą transformuoja ją pagal susitartą užklausių kalbą tarp saugyklų ir per SQI sąsają pasiunčia mokymosi objektų saugykli B. Čia atėjusi užklausa transformuojama į sistemos viduje naudojamą užklausių formatą ir įvykdoma. Rezultatai gražinami pagal tą patį principą mokymosi objektų bibliotekai A. (žr. 1 paveikslėli)



1 pav. SQI komunikacijos procesas.

21.4. VIPS SQI sąsaja

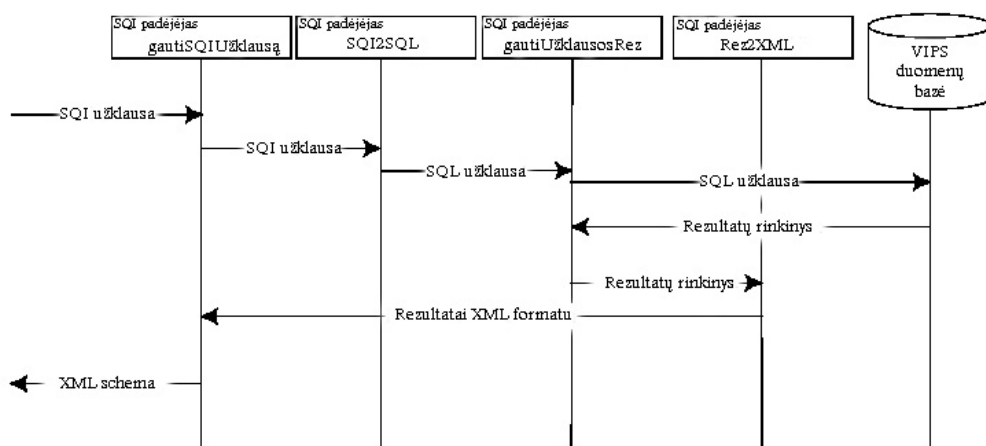
VIPS sistemoje SQI sąsaja realizuota sesijos ir užklausių valdymo Web servais. Pradinę architektūrą sudaro SQI padėjėjas su savo dviem moduliais (žr. 2 paveikslėli)



2 pav. SQI sąsaja VIPS sistemoje (pradinė architektūra).

SQI padėjėjas yra SQI sąsaja, įdėta kaip modulis į VIPS sistemą. SQI padėjėjas atsakingas už įeinančios užklauso priėmimą iš siuntėjo. Toliau žinutė yra siunčiama į SQI2SQL modulį, kuris transformuoja užklausą į SQL

(Structured Query Language) formatą. Po to SQI padėjėjas susisieikia su VIPS duomenų baze ir paima paieškos rezultatus. Rezultatai yra sudedami į XML formatą ir nusiunčiami gavėjui (žr. 3 paveikslėlį)



3 pav. Komunikacijos sekų diagrama VIPS sistemoje.

21.5. Paieškos servaisai

Pagal SQI specifikaciją, mokymosi objektų saugyklos gali būti pasiekiamos iš kitų sistemų apie prieinamą web servisą paskelbus nepriklausomame web servisų registre (UDDI). Tačiau, pagal iCamp projektą yra numatyta sukurti portalą, kuris leistų atlikti mokymosi medžiagos paiešką visose prieinamose saugyklose iš karto. Tuo tikslu portale turėtų būti paieškos sistema, sudaryta iš dviejų servisų:

- Mediatorius, turintis savyje paieškos mechanizmą ir bendraujantis su bibliotekomis (middleware logic)
- Portletas, papildantis mediatorių suteikiant paieškai sąsają su vartotoju (presentation logic)

22. Paieškos aspektai

Paieškos sistema priimtų iš vartotojo raktinius žodžius ir pagal juos nusiųstų paieškos užklausas saugykloms. Šios atliktų paiešką per savo metaduomenų įrašus ir gražintų rezultatus. Esant dideliems duomenų kiekiams, bei prisijungiant papildomoms virtualioms bibliotekoms ateityje tokia paieška gali tapti neefektyvi, surasti reikiamus mokymosi objektus gali tapti sudėtinga dėl[4]:

- Netikslios užklausų konstrukcijos. Paieškos žodžiai dokumentuose gali būti išreikšti kitais žodžiais, arba turėti kitokia prasmę kitame kontekste. Taip būtų gražinami nepilni arba netgi ne tokie rezultatai, kokių tikisi vartotojas. Nors pateikus daugiau raktinių žodžių šią problemą galima apeiti, tačiau iš bandymų rezultatų matyti, kad 99% vartotojų naudoja užklausoms iki dviejų raktinių žodžių.
- Sinonimai vietoje ieškomų žodžių. Paieškos sistema neduos rezultatų, jei dokumentuose bus ne tikslūs žodžiai, bet sinonimai. Net jeigu paieškos sistema atsirinktų ir sinonimus pagal kokį nors žodyną, rezultatų gali duoti per daug rezultatų ir vartotojas turės pats ieškoti reikiamos medžiagos per paieškos rezultatus.
- Konteksto trūkumas. Paieškos rezultatai galėtų būti tikslesni, jei atliekant paiešką būtų atsižvelgiama į vartotoją (duomenys apie specialybę, mokslo šaką, kurioje dirba vartotojas)
- Paprasto rezultatų vaizdavimo. Vartotojai paprastai nežiūri toliau dešimtos rezultato eilutės.

Galimas sprendimo būdas būtų naudoti semantines technologijas bibliotekose:

- Gauti turtingesnius metaduomenis. VIPS sistemoje vartotojai per SQI sąsają gali gauti aprašymus ir nuorodas į kursus ir įrašus. Abu šie mokymosi objektų tipai turi savo metaduomenis. Tačiau juos reikia suvesti žmogui. Tai labai sumažina galimybes susirasti norimus objektus, nes metaduomenys gali būti neužpildyti arba būti su klaidomis. Norint palengvinti procesą, galima būtų iš dalies automatizuoti metaduomenų išgavimą:
 - Pagal skaidres ir įvykius – teksto atpažinimo priemonėmis galima būtų iš skaidrių nuskaityti užrašus ir pagal tai atlikti santrauką arba išrinkti raktinius žodžius.
 - Pagal komentarus – kursai ir įrašai turi vartotojų komentarus. Esant didesniam jų kiekiui, galima išrinkti raktinius žodžius.
 - Pagal garsinę informaciją iš įrašo atsiradus tinkamoms priemonėms.
- Taikyti ontologijas. Ontologija – tai duomenų modelis, nusakantis sritį ir ryšius tarp objektų toje srityje. Naudojant ontologijas ir metaduomenis vartotojui galima leisti tiksliau išreikšti savo užklausą. Vartotojai

galėtų pasirinkti ontologinius konceptus užklausų patikslinimui arba pakartoti patikslintą užklausą pagal gražintus ontologinius konceptus.

Literatūros sąrašas

- [1] **Gytis Cibulskis, Giedrius Balbieris, Marius Siegas, Deividas Meškauskas.** Vaizdo paskaitų sistemos VIPS funkcinių galimybių plėtra. *Informacinės technologijos '2005*, 2005.
- [2] iCamp Consortium (2005): iCamp Technical Annex, p. 36
- [3] **B.Simon, D. Massart, Frans Van Assche, S. Ternier, E. Duval.** Simple Query Interface. Prieiga žiniatinklyje: <http://www.prolearn-project.org/lori> (2006 m. spalio 17 d.).
- [4] **J. Davie's, R. Studer, P. Warren.** Semantic Web Technologies, *John Wiley & Sons Ltd*, 2006, ISBN-10: 0-470-02596-4.
- [5] **Heiner Stuckenschmidt, Frank van Harmelen.** Information Sharing on the Semantic Web, *Springer*, 2003
- [6] iCamp (2006): Prototype of the Repository Network. Prieiga žiniatinklyje: http://www.icamp.eu/content/d3.2_icamp_final.pdf, (2006 m. spalio 17 d.).

Semantic Web Technology Research and Application in Video Lecturing System VIPS

This article is mainly covering extension possibilities of Video Lecturing System VIPS developed according Eureka project TESTVIL II. Development of digital learning objects repository based on VIPS recorded events archive and interconnection of it with European wide digital objects repository network is started in the framework of FP6 IST project iCamp - Innovative, inclusive, interactive and intercultural learning campus. Main aim is to expand the repository network and increase usability of digital repositories. Article provides analysis of technologies used for interconnecting repositories to the network and discusses solutions for searching of the information in this network.