



VILNIAUS GEDIMINO TECHNIKOS UNIVERSITETAS

FUNDAMENTINIŲ MOKSLŲ FAKULTETAS

INFORMACINIŲ TECHNOLOGIJŲ KATEDRA

Jurgis Mikučionis

SOCIALINIŲ ŽINIATINKLIŲ ANALIZĖS PRIEMONIŲ TYRIMAS
SOCIAL-NETWORKING ANALYSIS TOOL RESEARCH

Baigiamasis magistro darbas

Informacinių technologijų studijų programa, valstybinis kodas 62407T104

Duomenų gavybos technologijų specializacija

Informatikos inžinerijos mokslo kryptis

Vilnius, 2009

VILNIAUS GEDIMINO TECHNIKOS UNIVERSITETAS
FUNDAMENTINIŲ MOKSLŲ FAKULTETAS
INFORMACINIŲ TECHNOLOGIJŲ KATEDRA

TVIRTINU
Katedros vedėjas

(Parašas)

(Vardas, pavardė)

(Data)

Jurgis Mikučionis

SOCIALINIŲ ŽINIATINKLIŲ ANALIZĖS PRIEMONIŲ TYRIMAS
SOCIAL-NETWORKING ANALYSIS TOOL RESEARCH

Baigiamasis magistro darbas

Informacinių technologijų studijų programa, valstybinis kodas 62407T104

Duomenų gavybos technologijų specializacija

Informatikos inžinerijos mokslo kryptis

Vadovas doc. dr. Regina Kulvietienė

(Moksl. laipsnis, vardas, pavardė)

(Parašas)

(Data)

Konsultantas doc. dr. Angelė Kaulakienė

(Moksl. laipsnis, vardas, pavardė)

(Parašas)

(Data)

Konsultantas

(Moksl. laipsnis, vardas, pavardė)

(Parašas)

(Data)

Vilnius, 2009

Vilniaus Gedimino technikos universitetas
Fundamentinių mokslų fakultetas
Informacinių technologijų katedra

ISBN	ISSN
Egz.sk.	2
Data	2009 06 01

Inžinerinės informatikos studijų programos baigiamasis magistro darbas

Pavadinimas: Socialinių žiniatinklų analizės priemonių tyrimas

Autorius: Jurgis Mikučionis

Vadovas: doc. dr. Regina Kulvietienė

Kalba

☒ Lietuvių

☐ Užsienio (Anglų)

Anotacija

Baigiamajame darbe išnagrinėtas internetinių socialinių tinklų tyrimo procesas bei tam skirti įrankiai. Darbo pagrindą sudaro trys dalys, kuriose: apžvelgti populiariausi internetiniai socialiniai tinklai, jų realizavimo technologijos; išanalizuotas internetinių socialinių tinklų tyrimo procesas, pasiūlytas trijų etapų tyrimo modelis, apžvelgti kiekviename etape naudotini įrankiai; taip pat pateikiama pasiūlyto modelio realizacija „PERL“ programavimo kalba, analizuojanti socialinio portalo „Twitter“ duomenis ir sugebanti pateikti trijų tipų ataskaitas.

Darbą sudaro 6 dalys: įvadas; internetiniai socialiniai tinklai; internetinių socialinių tinklų tyrimo proceso modelis, tyrimo įrankiai; internetinių socialinių tinklų tyrimo modelio realizavimas; išvados; literatūros sąrašas.

Darbo apimtis – 61 p. teksto be priedų, 28 iliustr., 2 lent., 18 bibliografiniai šaltiniai.

Atskirai pridedami darbo priedai.

Prasminiai žodžiai: socialiniai žiniatinkliai, internetiniai socialiniai tinklai, internetinių socialinių tinklų tyrimo procesas, internetinių socialinių tinklų tyrimo įrankiai.

Vilnius Gediminas Technical University
Faculty of Fundamental sciences
Information technologies department

ISBN ISSN
No of copies 2
Date 2009 06 01

Engineering Informatics study programme master thesis

Title: Social-Networking Analysis Tool Research

Author: Jurgis Mikučionis

Adviser: doc. dr. Regina Kulvietienė

Thesis language

☒ Lithuanian

☐ Foreign (English)

Summary

The research theme of this master's final work is social-networking analysis process and tools relevant to it. The framework of this thesis consists of three parts: an overview of the most popular online social networks and their implementation technologies; a research on social-networking analysis process with a three-stage model proposed, all stages and tools relevant to them analyzed; also, implementation of the proposed model, written in "PERL" programming language, able to analyze a social portal named "Twitter", also able to provide three types of reports.

Structure: introduction; social-networking; social-networking analysis process model, analysis tools; social-networking analysis process realization; conclusions, references.

Thesis consists of: 61 p. text without appendixes, 28 pictures, 2 tables, 18 bibliographical entries.

Appendixes included.

Basic words: social-networking, social-networking analysis process, social-networking analysis tools.

TURINYS

Turinys	5
Paveikslų sąrašas	6
Lentelių sąrašas	6
Santrumpos	7
Įvadas	9
1. Internetiniai socialiniai tinklai	11
1.1. Skyriaus įvadas	11
1.2. Internetinių socialinių tinklų fenomenas	12
1.2.1. „MySpace“	14
1.2.2. „Facebook“	14
1.3. Internetinių socialinių tinklų technologijos	15
1.3.1. „AJAX“	16
1.3.2. „CSS“	18
1.3.3. „XHTML“	18
1.3.4. „REST“ / „SOAP“	19
1.3.5. „RSS“ / „ATOM“	20
1.3.6. Tinklo paslaugos	21
1.4. Skyriaus išvados	24
2. Internetinių socialinių tinklų tyrimo proceso modelis, tyrimo įrankiai	25
2.1. Skyriaus įvadas	25
2.2. Duomenų išrinkimas	26
2.3. Duomenų analizė	33
2.4. Duomenų vaizdavimas	36
2.5. Skyriaus išvados	40
3. Internetinių Socialinių tinklų tyrimo modelio realizavimas	41
3.1. Skyriaus įvadas	41
3.2. Duomenų išrinkimas	42
3.2.1. Duomenų išrinkimo realizavimas	43
3.3. Duomenų analizė	46
3.3.1. Duomenų analizės realizavimas	46
3.4. Duomenų vaizdavimas	50
3.4.1. Duomenų vaizdavimo realizavimas	50
3.5. Skyriaus išvados	57
Išvados	58
Literatūra	59
Priedai	61
A priedas.	61

PAVEIKSLŲ SĄRAŠAS

1 pav. Socialinių tinklų susikūrimo chronologinė seka [4]	13
2 pav. „MySpace“ („Alexa“ 2009 m. balandis 24 d.)	14
3 pav. „Facebook“ („Alexa“ 2009 m. balandis 24 d.)	15
4 pav. „AJAX“ veikimo scenarijus [6]	18
5 pav. Socialinių tinklų analizės proceso etapai	25
6 pav. Duomenų išrinkimas	26
7 pav. Informacijos srautai įprasto prisijungimo atveju	27
8 pav. Pirmas duomenų išrinkimo būdas	27
9 pav. Antras duomenų išrinkimo būdas	28
10 pav. Trečias duomenų išrinkimo būdas	29
11 pav. „Twitter“ „XML“ kalba pateikti duomenys (2009 m. balandis 19 d.)	31
12 pav. Socialinio portalo „Facebook“ testavimo aplinka (2009 m. balandis 19 d.)	32
13 pav. Duomenų analizavimas	33
14 pav. „Yahoo“ siūlomos internetinės paslaugos „Geocoding“ atsakymo pavyzdys (2009 m. balandis 19 d.)	34
15 pav. Valiutų kursų sąrašas (2009 m. balandis 4d.)	35
16 pav. Duomenų vaizdavimas	36
17 pav. Duomenų struktūros, aprašytos kodu pavyzdys (http://www.graphviz.org/Gallery/undirected/philogv.txt)	36
18 pav. „Graphviz“ duomenų struktūros pavyzdys (http://www.graphviz.org/Gallery/undirected/philohml)	37
19 pav. „Walrus“ duomenų struktūros pavyzdys (http://www.caida.org/tools/visualization/walrus/gallery1/)	37
20 pav. „Google Chart API“ pavyzdys (2009 m. gegužė 3 d.)	38
21 pav. Duomenų pateikimas „Excel“ šablonu „NodeXL“ (http://www.codeplex.com/NodeXL)	39
22 pav. "cURL"	42
23 pav. Užklauso rezultatas naršyklėje	44
24 pav. Paprogramei paruoštas duomenų srautas	51
25 pav. „ItakaDuo.fdp“	53
26 pav. „Itaka.png“	54
27 pav. „Zodziai.png“	55
28 pav. „Issidestymas.png“	56

LENTELIŲ SĄRAŠAS

1 lentelė. „REST“ ir „SOAP“ palyginimas [9]	20
2 lentelė. Interneto paslaugų sandara [13]	22

SANTRUMPOS

1. Web 2.0 – terminas, apibūdinantis besikeičiančias žiniatinklio technologijos ir dizaino naudojimo tendencijas, kai yra siekiama padidinti kūrybiškumą, keitimąsi informacija bei bendradarbiavimą tarp naudotojų.
2. RSS – tai XML failų formatų šeima internetiniam duomenų rinkimui iš naujienų tinklų.
3. XML (angl. eXtensible Markup Language) yra W3C (angl. World Wide Web Consortium) rekomenduojama bendros paskirties duomenų struktūrų bei jų turinio aprašomoji kalba.
4. JavaScript – objektiškai orientuota scenarijų programavimo kalba, besiremianti prototipų principu.
5. HTML (angl. Hyper text Markup Language) – tai kompiuterinė žymėjimo kalba, vartojama turiniui internete pateikti. Kalbą standartizuoja W3C.
6. XMLHttpRequest (XHR) - tai užklauskos objekto sąsaja naudojama naršyklės scenarijų kalboje („JavaScript“).
7. DHTML (dinaminis HTML) – technologija, sujungianti kelių kitų technologijų kombinaciją („HTML“, „XML“, „CSS“ ir kt.).
8. LAMP (angl. Linux, Apache, MySQL, PHP) – akronimas, vartojamas kalbant apie atvirojo kodo programinę įrangą, skirtą internetinio serverio ir duomenų bazės valdymo funkcijoms atlikti.
9. ActiveX - tai programinės įrangos komponentų apibrėžimo taisyklių visuma.
10. XSLT (angl. eXtensible Stylesheet Language (XSL) transformations) – kalba, aprašanti XML dokumento transformaciją į HTML dokumentą arba į kitokios struktūros XML dokumentą.
11. MathML (angl. Mathematical Markup Language) – matematinių išraiškų bei struktūros formatus „XML“ žymėjimo kalbos pagrindu.
12. SVG (angl. Scalable Vector Graphics) – dvimačių vektorinių vaizdų saugojimo formatus „XML“ žymėjimo kalbos pagrindu.
13. MIME – (angl. Multi-Purpose Internet Mail Extensions). Tai protokolo TCP/IP papildymas, naudojamas siųsti internetu ne atvirą tekstą.
14. SMTP (angl. Simple Mail Transfer Protocol) — de facto standartas el.pašto laiškų perdavimui internete. Naudojamas elektroniniams laiškam pristatyti į gavėjo el.pašto dėžutę.
15. HTTPS (angl. Hypertext Transfer Protocol Secure) - tai HTTP ir kodavimo (SSL (Secure Socket Layer) arba TLS (Transport Layer Security)) protokolų kombinacija.

16. JMS (angl. Java Message Service) - standartizuota „Java“ „API“ (taikomųjų programų programavimo sąsaja (angl. „Application Programming Interface“)), skirta komunikacijai pranešimais realizuoti tarp dviejų ar daugiau programų sistemų ar jos dalių.
17. IIOP (angl. internet inter ORB protocol) – GIOP (ang. „general inter ORB protocol“ – dviejų kokių nors ryšio kanalų (paprastai kompiuterių tinklų) susijusių programų bendravimo protokolas) versija, papildyta standartais, reikalingais dirbti internete.
18. WSDL (angl. Web Services Definition) - speciali kalba „XML“ pagrindu skirta apibūdinti interneto paslaugoms (angl. Web Services).
19. UDDI (angl. Universal Description, Discovery, and Integration) – technologija interneto paslaugoms aptikti ir naudoti.
20. SOA (angl. Service Oriented Architecture) - tai toks verslo taikomųjų programų kūrimo būdas, kai programiniai procesai suskaidomi į paslaugas, kurios yra pateikiamos ir įmanomos surasti tinkle.
21. JSON (angl. JavaScript Object Notation) – paprastas, tekstu paremtas, žmogui įskaitomas duomenų perdavimo formatas.
22. PHP – plačiai paplitusi dinaminė interpretuojama programavimo kalba (angl. Hypertext Preprocessor), sukurta 1997 m. ir specialiai pritaikyta interneto svetainių kūrimui.
23. Venn diagrama - diagrama, naudojama užkloties sritims tarp elementų rodyti.
24. QR kodas - nedidelis baltų ir juodų taškų tinklelis, kuriame gali būti užkoduota net iki 3 KB duomenų.
25. URL (angl. Universal Resource Locators) - konkretaus informacijos šaltinio (tinklapių, vieno tinklalapio puslapio) adresui nurodyti internete sukurtas standartas.
26. PERL - pagal idėją tai ne santrumpa, o tiesiog pavadinimas, tačiau vėliau buvo pasiūlyti ir santrumpą atitinkantys pavadinimai, tokie kaip: „Practical Extraction and Report Language“ ar „Pathologically Eclectic Rubbish Lister“.
27. UNIX – grupė operacinių sistemų, kilusių iš 1969 m. - 1970 m. sukurtos UNICS (apie 1970 pervadintos į UNIX) sistemos.
28. PNG (angl. Portable Network Graphics) – bitų masyvo formatas, kuris suglaudinas „be nuostolių“. Sukurtas kaip GIF formato pakaitalas – formatas nereikalaujantis patento licencijos.
29. Maišos lentelė (angl. hash table) – duomenų struktūros, kuriose duomenys yra saugomi, priskiriant jiems unikalų raktą.
30. ISO 3166 - trijų dalių geografinis kodavimo standartas koduojantis valstybių ir priklausomų teritorijų pavadinimus.

IVADAS

Socialinio tinklo sąvoka neturi tikslaus apibrėžimo, plačiąja prasme – tai socialinė struktūra, kurią sudaro tarpusavyje tam tikrais saitais susiję mazgai (individaai ar organizacijos). Realiam gyvenime – tai žmonių grupė, kurią jungia bendri interesai (bendros vertybės, idėjos, draugystė, giminytė, darbas, seksualiniai santykiai).

Socialinis tinklas internete - tai „WEB 2.0“¹ portalas, leidžiantis savo naudotojams susikurti individualius profilius, jungtis į grupes, keistis informacija.

Darbo tikslas: Išanalizuoti internetinių socialinių tinklų tyrimo procesą ir tam skirtus įrankius.

Uždaviniai:

- Apžvelgti populiariausius socialinių tinklų portalus bei jų realizavimo technologijas.
- Suprojektuoti internetinių socialinių tinklų tyrimo proceso modelį, išanalizuoti visus jo etapus, kiekvienam etapui tinkamus įrankius.
- Realizuoti internetinių socialinių tinklų tyrimo proceso modelį.

Temos aktualumas: Tinklalapio „PewResearchCenter“ duomenimis, per pastaruosius 4 metus internetų (asmenų besinaudojančių internetu) turinčių profilį internetiniame socialiniame portale skaičius išaugo 4 kartus (nuo 8% iki 35% visų internetautų). Populiariausi socialiniai tinklai tarp jaunų žmonių, 18 – 25 metų amžiaus grupėje net 65% visų internetautų turi profilį kokiame nors socialiniame žiniatinklyje [1].

„Alexa“ duomenimis, net du socialinių tinklų portalai puikuoja populiariausių pasaulio interneto svetainių dešimtuose. Populiariausią socialinį portalą „Facebook“ kiekvieną dieną vidutiniškai aplanko 17% visų pasaulio internetautų.

Dėl milžiniško naudotojų skaičiaus ir potencialių galimybių socialiniai portalai šiandien vertinami milijonais, netgi kai kuriais atvejais milijardais dolerių. Prieš keletą metų socialinį portalą „MySpace“ už 580 mln. JAV dolerių nupirko žiniasklaidos grupė „News Corporation“, „Google“ į tą patį portalą investavo 900 milijonų JAV dolerių, kurie minėtoje svetainėje suteikia jai išskirtines reklamos teises [2].

Mokslinis naujumas: Socialinių tinklų analizės sąvoka nėra nauja. Socialiniai tinklai, kaip socialinės žmonių grupės, analizuojami ir tiriama jau daugybę metų. Socialiniai tinklai internete - kur kas naujesnė sąvoka ženkliai išplečianti socialinių tinklų tyrimo kaip mokslo ribas. Internetiniai socialiniai tinklai, nors iš esmės vienodi, tačiau realizuoti labai įvairiai, todėl jų tyrimas sudėtingas ir nuo konkretaus produkto priklausantis procesas.

Socialinių tinklų internete tyrimo sąvoka šiandien apsiriboja jų atvaizdavimu. Nėra sukurto jokio konkretaus tyrimo modelio ar įrankių, kurie apibrėžtų visą socialinių tinklų internete tyrimo

¹ Web 2.0 – terminas, apibūdinantis besikeičiančias žiniatinklio technologijas ir dizaino naudojimo tendencijas, kai yra siekiama padidinti kūrybiškumą, keitimąsi informacija bei bendradarbiavimą tarp naudotojų.

procesą. Socialiniams tinklams internete eksponentiškai augant, poreikis juos analizuoti didėja, todėl būtina šio proceso ir su juo susijusių problemų analizė.

Praktinė vertė: Atlikta analizė bei sukurtas internetinių socialinių tinklų tyrimo proceso modelis nusako, kaip turėtų būti analizuojama socialinių portalų informacija nuo jos gavybos iki rezultatų pateikimo šiuolaikinių technologijų terpėje. Tokia informacija skirta įmonės veiklai optimizuoti ar tam tikrai vartotojų auditorijai aptikti (jei kalbama apie reklamą).

Tyrimo metodika: Lyginamoji literatūros šaltinių bei ekspertų patirties analizė.

1. INTERNETINIAI SOCIALINIAI TINKLAI

1.1. Skyriaus įvadas

Šiuo metu pasauliniame internete ryškiausia tendencija – sparčiai augantis socialinių tinklų populiarumas. Sparčiausiai kyla portalo „Facebook“ naudotojų skaičius. JAV per pastaruosius metus jis išaugo 149 %, tuo tarpu Europoje – 314 %, kaip skelbia bendrovės „Hitwise“ ir „Comscore“. Socialinių tinklų populiarumas pastaruoju metu didėja dėl 35-54 metų amžiaus grupės naudotojų.

Lietuvoje socialinių tinklų augimas nelieka nepastebėtas internetinio verslo žaidėjų. Praėjusių metų pabaigoje didžiausias lietuviškas klasiokus ir grupiokus jungęs socialinis tinklas „Klase.lt“ pakeitė koncepciją ir pritaikė savo aplinką ne tik bendramokslių, bet taip pat savo draugų ir bendraminčių tinklui kurti. Remiantis interneto tyrimų bendrovės „Gemius“ duomenimis, po šio atsinaujinimo per 3 mėnesius „Klase.lt“ lankytojų skaičius išaugo 75 %.

Interneto vartojimo ir internete praleisto laiko rodikliai kasmet didėja, tačiau žmonių laikas ribotas. Tad socialinių tinklų augimas vyksta mažėjančio pramoginių interneto portalų populiarumo sąskaita. „NetPop Research“ duomenimis, pramoginėse svetainėse interneto vartotojai 2008 m. praleido 30 % mažiau laiko negu 2006 m. Video vaizdelių ir žaidimų portalai šiai tendencijai atsparūs, tačiau muzikinių svetainių ir interneto pokalbių kambarių lankomumas mažėja.

Socialinius tinklus įvairioms reklamos kampanijoms ir prekės ženklo stiprinimui vis dažniau ima naudoti įmonės bei reklamos agentūros. Maiklo Stelznerio atlikto tyrimo duomenimis, 77 % mažų ir vidutinių JAV įmonių bendrauja su savo vartotojais kuriame nors socialiniame tinkle [3].

Socialinius tinklus pamažu atranda ir Lietuvos įmonės. Pavyzdžiui, kavinių tinklas „Coffe Inn“ „Facebook“ anketoje aktyviai bendrauja su savo lankytojais. Mobilusis operatorius „Tele2“, į socialinį tinklą „Klase.lt“ integravo „Pildyk“ velykinį žaidimą. Įvairios šokių mokyklos per savo puslapius socialiniuose tinkluose renka naujas grupes bei informuoja apie būsimus renginius [3].

1.2. Internetinių socialinių tinklų fenomenas

Pasaulyje pirmosios tokio tipo internetinės svetainės buvo „Classmates“ (skirta bendraklasių paieškai) ir „SixDegrees“. Pastaroji buvo labai panaši į dabartinius socialinius portalus. Joje buvo galima susikurti profilį, draugų ratą ir nuo 1998 m. naršyti po draugų profilius. Tuo metu susikurti profilį jau buvo įmanoma ir daugelyje kitų pažinčių ar formalių bendruomenių svetainių. „Classmates“ iš pradžių leido tik prisijungti prie savo universiteto ar kolegijos tinklo, galimybės susikurti savo profilį ir draugų ratą nebuvo. Visas šias savybes pirmoji sujungė „SixDegrees“. Ši svetainė save reklamavo kaip priemonę susisiekti ir bendrauti. Dėl menko verslo plėtojimo 2000 m. „SixDegrees“ nutraukė savo veiklą (tuo metu ji jau buvo pritraukusi milijonus naudotojų). Vėliau 2007 m. tinklalapio įkūrėjas A. Weinreichas pareiškė, kad ši paslauga internete atsirado tiesiog per anksti, besibūriuojantys interneto naudotojai tuo metu dar neturėjo pakankamai daug draugų, kurie galėjo naudotis internetu, o pažintys su nepažįstamaisiais jų dar nedomino.

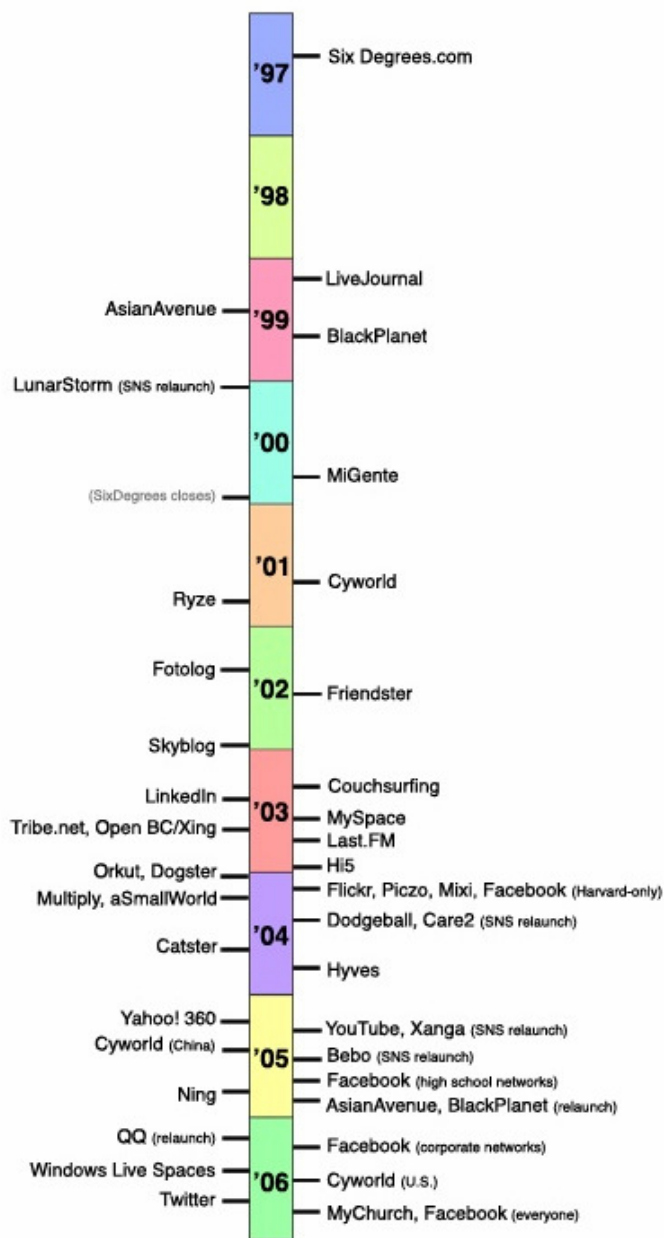
Nuo 1997 m. iki 2001 m. pradėjo veikti aibė tinklalapių, kurie siūlė įvairių įrankių profiliams kurti bei draugams skirstyti. „AsianAvenue“, „BlackPlanet“ bei „MiGente“ savo naudotojams leido vienu metu turėti kelis profilius (asmeninius, profesinius ar skirtus pasimatymams), taip suskirstyti tinklai įgalino portalų savininkus atskleisti daugiau informacijos apie savo tinklų narius, supaprastėjo naudotojų paieškos sistema, jie pradėjo aktyviau bendrauti ir „jungtis“. Po kurio laiko panašią sistemą pradėjo naudoti ir tinklaraščių bendruomenė „LiveJournal“, portalas taip pat pristatė naują funkciją, primenančią į šių dienų „RSS“² sistemą, naujoji funkcija suteikė galimybę sekti mėgstamų autorių (sistemoje draugų) veikalus.

Kita socialinių tinklų „banga“ prasidėjo 2001 m., pradėjus veikti portalui „Ryze“. „Ryze“ įkūrėjas pirmuosius su šiuo portalu supažindino savo draugus, įvairius verslininkus bei investuotojus, kurie priklausė San Francisko verslo ir technologijų bendruomenei. Įkūrėjas jiems pristatė projektą, kurį sudarė visa socialinių tinklų grupė. Grupės nariais turėjo tapti keli socialiniai portalai, orientuoti į skirtingas realaus pasaulio sritis. „Ryze“ turėjo tapti verslo portalu, „LinkedIn“ - socialiniu tinklu suvedančiu darbdavius ir darbuotojus, „Tribe“ - įvairių bendruomenių bei interesų grupių tinklu, o „Friendster“ – naujų draugų paieškos portalu. Galiausiai pripažinimo susilaukė tik „LinkedIn“ ir „Tribe“, „Ryze“, nors gyvuoja iki šiol, masiškai populiariu netapo, o „Friendster“ buvo įvardytas kaip vienas didžiausių nusivylimų interneto istorijoje. „Friendster“ nesėkmės priežastimi tapo per staigus ir per greit išaugęs naudotojų susidomėjimas, dėl kurio pradėjo strigti techninė portalo įranga. Be techninių kliūčių, socialinio tinklo kūrėjai susidūrė ir su privatumo problemomis. „Friendster“ veikė „draugo draugas – mano draugas“ principu, portalui masiškai išpopuliarėjus, visa asmeninė naudotojų informacija, skirta artimiausiems draugams, per ilgas „draugo draugo draugo...“ grandines tapo prieinama visiems sistemos nariams [4].

² RSS – tai XML failų formatų grupė internetiniam duomenų rinkimui iš naujienų tinklų.

Nuo 2003 m. po netikėtos ankstyvos portalų „Friendster“ sėkmės, internete pradėjo kurtis daugybė kitų socialinių tinklų, orientuotų į įvairias naudotojų grupes. Portalas „Dogster“ buvo skirtas susipažinti su nepažįstamaisiais pagal bendrus interesus, „MyChurch“ – sujungti krikščioniškosios bažnyčios narius, „CouchSurfing“ – sujungti keliautojus, ieškančius nakvynės.

Detalesnė internetinių socialinių tinklų susikūrimo seka pavaizduota 1 paveikslėlyje.



1 pav. Socialinių tinklų susikūrimo chronologinė seka [4]

2004 m. Dale'as Dougherty ir Craig'as Cline „O'Reilly Media Web 2.0“ konferencijoje oficialiai pristatė „Web 2.0“ sąvoką. Tuo metu daugelis socialinių portalų jau siūlė savo naudotojams įrankių vaizdinei ar garsinei informacijai keistis ar publikuoti. Keletas įrankių ypač išpopuliarėjo ir tapo naujų socialinių portalų paskirties dalimi. Portalas „Youtube“ tapo vaizdo

informacijos publikavimo ar keitimosi socialiniu tinklu, „Last.FM“ muzikos mėgėjų socialiniu tinklu, „Flickr“ – nuotraukų portalu [5].

1.2.1. „MySpace“

„Myspace“ pradėjo veikti nuo 2003 m., kartu su „Friendster“, „Xanga“, bei „AsianAvenue“, tačiau išpopuliarėjo tik nuo 2005 m. Tuo metu, „Alexa“ duomenimis (2 paveikslėlyje pavaizduoti portalo populiarumo rodikliai 2009 m. balandžio 22. d.), svetainė buvo populiareesnė net už „Google“. Pradinis „Myspace“ tikslas buvo pritraukti atsiskyrusius „Friendster“ naudotojus, kurie, pasklidus gandams, jog žadama portalą apmokestinti, ėmė ieškoti kitų nemokamų socialinių tinklų. Prie naudotojų skaičiaus augimo prisidėjo ir populiarių muzikinių grupių, kurios buvo pašalintos iš „Friendster“ dėl profilių taisyklių nesilaikymo, jungimasis į „Myspace“. „Myspace“ tapo puiki platforma muzikinių grupių reklamai bei populiarinimui. Masinis paauglių bendruomenių jungimasis prasidėjo 2004 m. Dauguma jų apie „Friendster“ nebuvo net girdėję, o tiesiog norėjo tapti mylimos grupės gerbėjų klubo nariais. Dar vienas svarbus aspektas, nulėmęs portalo populiarumą, buvo „Myspace“ prieigos neribojimas nei paaugliams, nei nepilnamečiams [4].



2 pav. „MySpace“ („Alexa“ 2009 m. balandis 24 d.)

1.2.2. „Facebook“

„Facebook“ sukūrė Markas Zuckerbergas 2004 m. vasario mėn. Pradžioje „Facebook“ buvo skirtas tik Harvardo universiteto studentams. Vėliau - visų JAV universitetų studentams, dar vėliau - įmonių darbuotojams, organizacijų nariams, 2006 m. gale – visiems. 2008 m. pavasarį platforma išversta į vokiečių, ispanų, prancūzų kalbas. Šiuo metu „Facebook“ jau išverstas į 37 įvairias pasaulio kalbas.

Kaip jau minėta, „Facebook“ - pats populiariausias socialinis portalas internete (3 paveikslėlyje pavaizduoti portalo populiarumo rodikliai 2009 m. balandžio 22. d.). Kodėl jis tapo toks populiarus, pasakyti sunku, dažniausiai įvardijamos priežastys: galimybė sukurti savo programinį modulį bei paprastas, bet kartu ir galingas draugų paieškos mechanizmas [4].



3 pav. „Facebook“ („Alexa“ 2009 m. balandis 24 d.)

1.3. Internetinių socialinių tinklų technologijos

Socialinio tinklo internete sąvoka remiasi „Web 2.0“ ideologija. „Web 2.0“ – terminas, apibūdinantis besikeičiančias žiniatinklio technologijos ir dizaino naudojimo tendencijas, kai yra siekiama padidinti kūrybiškumą, keitimąsi informacija bei bendradarbiavimą tarp naudotojų.

„Web 2.0“ sąvoka pirmą kartą buvo paminėta 2003 m. ir išpopuliarėjo po pirmosios „O’Reilly Media Web 2.0“ konferencijos 2004 m. Nors iš pažiūros terminas nusako naują interneto versiją, jis nereiškia naujos tinklo specifikacijos. Iš tikrųjų, jis apibūdina/nusako naujus būdus, kaip programuotojai ir galutiniai naudotojai turėtų naudotis interneto galimybėmis.

„Web 2.0“ technologijos:

- „AJAX“ (tam tikrų taisyklių visuma naudojama „Web 2.0“ tinklalapiams kurti).
- „CSS“ (kalba, skirta nusakyti kita struktūrine kalba aprašyto dokumento vaizdavimą. Dažniausiai „CSS“ aprašomas „HTML“ dokumentų pateikimas, tačiau ją galima taikyti ir įvairiems kitiems „XML“³ dokumentams).
- „XHTML“ (angl. „eXtensible HyperText Markup Language“, žymėjimo kalba, kuri turi panašias į „HTML“ žymėjimo taisykles, tik jos sintaksė yra griežtesnė. „HTML“ yra

³ XML (angl. eXtensible Markup Language) yra W3C (angl. World Wide Web Consortium) rekomenduojama bendros paskirties duomenų struktūrų bei jų turinio aprašomoji kalba.

pagrįsta „SGML“ (speciali aprašomoji (ne programavimo) kalba, skirta „HTML“ ir kitų giminingų kalbų struktūrai aprašyti.), taigi yra mažiau suvaržyta (galima praleisti kai kuriuos žymenis, pavyzdžiui, neužbaigti paragrafo), o „XHTML“ remiasi „XML“, todėl „XHTML“ dokumentas turi būti struktūriškai tvarkingas (angl. well-formed)).

- „REST“ (angl. „Representational State Transfer“ – supaprastinta sąsaja, skirta taikomųjų programų specifiniams duomenims perduoti per tinklą nenaudojant papildomų pranešimų lygių) / „SOAP“ (angl. „Simple Object Access Protocol“ – protokolas, skirtas keistis „XML“ pranešimais pasitelkus tinklą).
- „RSS“ (tai „XML“ failų formatų grupė internetiniam duomenų rinkimui iš naujienų tinklų) / „ATOM“ (sindikacijos standartas, pateikiamas „XML“ kalba ir naudojamas žiniatinklio srautuose (angl. „web feeds“)).
- Tinklo paslaugos („Web Services“).

„Web 2.0“ ideologija paremti sprendimai:

- Tinklaraščiai (angl. „Blogs“).
- „Cloud Computing“ (techninės bei programinės įrangos išteklių naudojimas kaip paslaugos per internetą).
- „Mesh Ups“ (taikomosios tinklo programos sujungiančios kitų taikomųjų programų duomenis ir/arba funkcionalumą).
- „RSS“ (duomenų išrinkimui skirta „XML“ formatų grupė).
- „Podcasts“ (prenumeruojamos transliacijos).
- Internetiniai socialiniai tinklai.
- „Wiki“ programinė įranga.

1.3.1. „AJAX“

„AJAX“ arba asinchroninis „JavaScript“⁴ ir „XML“ programavimas – terminas, apibrėžiantis svetainių programavimo technologiją, naudojančią šias priemones maksimaliam interaktyvumui pasiekti:

- „HTML“⁵ (ar „XHTML“) bei „CSS“ informacijos vaizdavimui;
- dokumento objektinį modelį (DOM) bei „JavaScript“ kalbą dinamiškam vaizdavimui bei interaktyvumui;
- „XML“, „XSLT“⁶ ir „XMLHttpRequest“⁷ objektą asinchroniniams duomenų mainams su serveriu („AJAX“ tam galėtų naudoti ir kitas technologijas).

⁴ JavaScript – objektiškai orientuota scenarijų programavimo kalba, besiremianti prototipų principu.

⁵ HTML (angl. Hyper text Markup Language) – tai kompiuterinė žymėjimo kalba, vartojama turiniui internete pateikti. Kalbą standartizuoja W3C.

⁶ XSLT (angl. eXtensible Stylesheet Language (XSL) transformations) – kalba, aprašanti XML dokumento transformaciją į HTML dokumentą arba į kitokios struktūros XML dokumentą.

Kaip ir „DHTML“⁸ ar „LAMP“⁹, „AJAX“ nėra technologija, bet tik abstraktus terminas, apimantis kelių technologijų kombinaciją.

Tradicinių svetainių interaktyvumas kuriamas kaskart kreipiantis į serverį. Pavyzdžiui, užpildžius anketą (ar formą), kreipiamasi į serverį ir įkraunamas naujas puslapis. Tokiu būdu bereikalingai iššvaistomi ištekliai, nes didelė dalis informacijos nepakinta ir siunčiama kaskart iš naujo. Taip pat tokiu būdu neįmanoma pasiekti tokio interaktyvumo, kokį gali suteikti neinternetinės taikomosios programos.

„AJAX“ svetainės gali siųsti užklausas serveriui, gauti atsakymą tam tikra apsibrėžta forma („SOAP“ ar kita „XML“ paremta) bei naudojant „JavaScript“ programavimo kalbą atnaujinti tik reikiamą puslapio dalį. Tokiu būdu sutaupomi tinklo ištekliai, nes žymiai sumažinami duomenų srautai. Taip pat taupomi ir serverio procesoriaus ištekliai, nes dalis logikos perkeliama klientui (6 paveikslėlis).

Kaip ir „DHTML“ svetainėms, „AJAX“ svetainėms reikia žymiai daugiau testavimo, nes kiekviena naršyklė gali elgtis šiek tiek skirtingai. Tačiau naudojant šią technologiją paspartinamas interaktyvumas. Naudojant „AJAX“, padaugėja mikroužklausų serveriui, toks neapgalvotas technologijos naudojimas gali sulėtinti bendravimą tarp kliento ir serverio.

Kartais sunku ar net neįmanoma „AJAX“ svetainių pritaikyti nepalaikomoms naršyklėms, taip pat sudėtinga pritaikyti svetaines standartams, kurie atsižvelgia į neįgalius žmones.

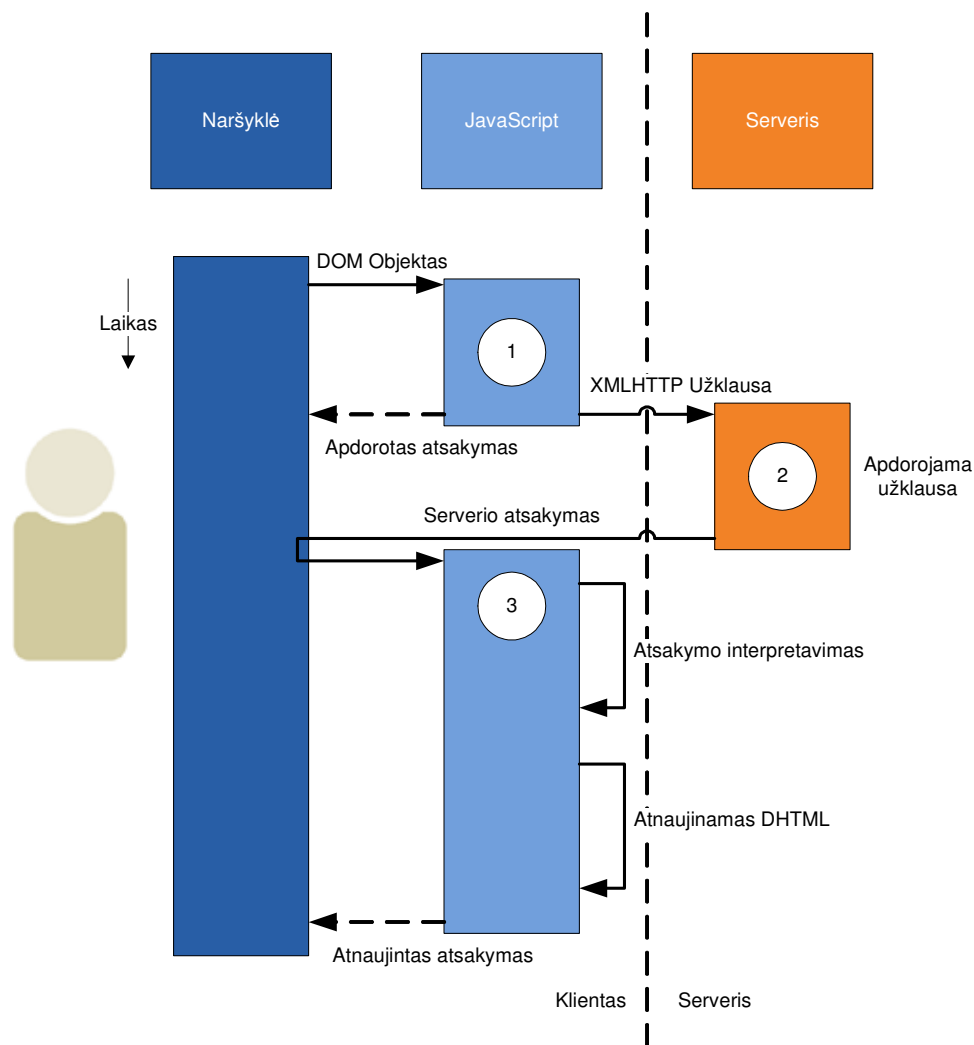
„AJAX“ tampa būtent „AJAX“ kai tinklalapio kūrėjas panaudoja „XMLHttpRequest“ „JavaScript“ objektą. „XMLHttpRequest“ pirmą kartą buvo panaudotas kompanijoje „Microsoft“ kaip „ActiveX“¹⁰ objektas „IE 5.0“ naršyklėje (1999 m. kovo mėn.). „Mozilla“ pridėjo „XMLHttpRequest“ objektą „Mozilla 1.0“ ir „Netscape 7“ naršyklėms (2001 m. kovo mėn.). Po to ėjo „Apple“ kompanijos naršyklė „Safari 1.2“ (2004 m. vasario mėn.), o „Opera“ „XMLHttpRequest“ objektą pradėjo taikyti nuo versijo 8.5 (2005 m. rugsėjo mėn.) [6].

⁷ XMLHttpRequest (XHR) - tai užklausos objekto sąsaja naudojama naršyklės scenarijų kalboje („JavaScript“).

⁸ DHTML (dinaminis HTML) – technologija, sujungianti kelių kitų technologijų kombinaciją („HTML“, „XML“, „CSS“ ir kt.).

⁹ LAMP (angl. Linux, Apache, MySQL, PHP) – akronimas, vartojamas kalbant apie atvirojo kodo programinę įrangą, skirtą internetinio serverio ir duomenų bazės valdymo funkcijoms atlikti.

¹⁰ ActiveX - tai programinės įrangos komponentų apsibrėžimo taisyklių visuma.



4 pav. „AJAX“ veikimo scenarijus [6]

1.3.2. „CSS“

„CSS“ arba pakopinės stilių lentelės - tai specialus „HTML“ kalbos priedas, kuriuo galima pagerinti naršyklėje pateikiamos informacijos išvaizdą. Su „CSS“ galima kiekvienam dokumento elementui nurodyti šriftą, spalvą, dydį ir pan.

„CSS“ atsiradimas buvo didžiulis šuolis kuriant internetinį dizainą, nes šia technologija galima apibrėžti stilių bet kokiam „HTML“ elementui ir jį priskirti prie tiek puslapių, kiek norima. Norint pakeisti viso puslapio struktūrą užtenka pakoreguoti stiliaus failą, visa kita atsinaujina automatiškai, taip sutaupoma nemažai laiko [7].

1.3.3. „XHTML“

„XHTML“ (angl. „eXtensible HyperText Markup Language“) yra žymėjimo kalba, kuri turi panašias į „HTML“ žymėjimo taisykles, tik jos sintaksė yra griežtesnė. „HTML“ yra pagrįsta „SGML“ (speciali aprašomoji (ne programavimo) kalba, skirta „HTML“ ir kitų giminingų kalbų struktūrai aprašyti.), taigi yra mažiau suvaržyta (galima praleisti kai kuriuos žymenis, pavyzdžiui,

neužbaigti paragrafo), o „XHTML“ remiasi „XML“, todėl „XHTML“ dokumentas turi būti struktūriškai tvarkingas (angl. well-formed).

„XHTML“ buvo sukurtas „W3C“ ir yra tapęs šio konsorciumo oficialia rekomendacija.

Labiau sintaksiškai suvaržytos „HTML“ versijos poreikis atsirado padidėjus naudojamų mobiliųjų įrenginių skaičiui. Šie įrenginiai yra maži, todėl neturi daug išteklių, kuriuos būtų galima panaudoti struktūriškai sudėtingam „HTML“ kodui.

Dauguma naujesnių žiniatinklio naršyklių gali dirbti su „XHTML“ kodu, bet senesnės naršyklės „XHTML“ kodą supranta kaip „HTML“. Taip pat beveik visos naršyklės, suderinamos su „XHTML“, gali naudoti „HTML“ kodą. Yra nuomonių, kad šis faktas stabdo „XHTML“ plėtrą. Nepaisant to, kad „Microsoft“ yra visateisė „W3C“ narė, šios kompanijos naršyklė „Internet Explorer“ nėra suderinama su kai kuriomis „XHTML“ rekomendacijomis.

Ypač naudinga „XHTML“ savybė yra ta, kad šia žymėjimo kalba parašytame dokumente galima vartoti kitų iš „XML“ kilusių kalbų (pvz., „MathML“¹¹, „SVG“¹²) elementus. Vis dėlto šią savybę galima panaudoti tik tada, kai „XHTML“ dokumentas yra siunčiamas kaip „XML“ (nurodant „application/xhtml“ bei „xml“ „MIME“¹³ tipą) [8].

1.3.4. „REST“ / „SOAP“

Tinklo komunikacijos protokolai taip pat laikomi „Web 2.0“ dalimi. Žinomiausi iš jų yra „REST“ ir „SOAP“. Abu jie skirti duomenims manipuliuoti nuotoliniu būdu (detalus abiejų protokolų palyginimas pateiktas 1 lentelėje). Kad paslauga būtų galima pasinaudoti, turi būti apibrėžti metodai, kaip tai padaryti. Tai priklauso nuo paslaugų tiekėjų.

„REST“ - gana paprastas protokolas, tačiau nelengvai įgyvendinamas, kai duomenų manipuliacija yra sudėtinga. Jis visiškai pagrįstas „HTTP“ protokolu, ir visos naudojamos funkcijos yra „HTTP“ protokolo „GET“, „POST“, „PUT“ ir „DELETE“ komandos.

„SOAP“ - kur kas sudėtingesnis ir lankstesnis protokolas. Faktiškai tai vienas iš tinklo tarnybų paslaugų („Web Services“) protokolų. „SOAP“ gali būti aptrauktas ant tokių protokolų, kaip „HTTP“ arba „SMTP“¹⁴. Labiausiai paplitusi „SOAP“ modifikacija veikia „HTTP“ ir „HTTPS“¹⁵ protokolų pagrindu ir su „POST“ metodu siunčia „XML“ užklausas serveriui.

Pirmoji „SOAP“ versija buvo išimtinai paremta „XML-RPC“ technologija, kuri išliko kaip pagrindinė iki šių dienų. „XML-RPC“ technologija yra specifikacija, kuri suteikia galimybę

¹¹ MathML (angl. Mathematical Markup Language) – matematinių išraiškų bei struktūros formatus „XML“ žymėjimo kalbos pagrindu.

¹² SVG (angl. Scalable Vector Graphics) – dvimačių vektorinių vaizdų saugojimo formatus „XML“ žymėjimo kalbos pagrindu.

¹³ MIME – (angl. Multi-Purpose Internet Mail Extensions). Tai protokolo TCP/IP papildymas, naudojamas siųsti internetu ne atvirą tekstą.

¹⁴ SMTP (angl. Simple Mail Transfer Protocol) — de facto standartas el.pašto laiškų perdavimui internete. Naudojamas elektroniniams laiškamsi pristatyti į gavėjo el.pašto dėžutę.

¹⁵ HTTPS (angl. Hypertext Transfer Protocol Secure) - tai HTTP ir kodavimo (SSL (Secure Socket Layer) arba TLS (Transport Layer Security)) protokolų kombinacija.

programinės įrangos sistemoms, veikiančioms skirtingose aplinkose, vykdyti nuotolines užklausas. Šios užklausos persiunčiamos kaip „XML“ pranešimai, panaudojant „HTTP“ protokolą transportavimui.

„XML-RPC“ technologija yra taikoma ir kaip atskiras komunikavimo standartas. Jo panaudojimas internetinėms paslaugoms realizuoti yra efektyvesnis sprendimas nei „SOAP“, nes pastarasis yra sudėtingesnis ir persiunčia kur kas daugiau metaduomenų [9] [10].

1 lentelė. „REST“ ir „SOAP“ palyginimas [9]

Kriterijus	REST	SOAP
Nėra persiunčiama daug perteklinių duomenų	+	-
Paprasta realizacija	+	-
Nepriklausomybė nuo protokolo	-	+
Duomenų tipų tikrinimas	-	+
Yra serviso aprašymo kalba	+	+
Egzistuoja pagalbinės kūrimo priemonės	-	+
Vieninga prieigos sąsaja	+	-
Galimybė naudoti spartinančiąją atmintį	+	-
Žmogui suprantamos užklausos ir jų rezultatai	+	-

1.3.5. „RSS“ / „ATOM“

„Web feed“ (žiniatinklio srautas) - tai toks duomenų formatas, kuris naudojamas dažnai atnaujinamam turiniui pateikti interneto naudotojams. Dažniausiai „Web feed“ vadinamas tiesiog „RSS“, bet tai nėra tas pats.

„Web feed“ - dokumentas, kuriame yra sutrumpintos (kartais ir visos) naujienų versijos su nuorodomis į originalius straipsnius. Dažniausiai tai naudojama naujienų svetainėse ir tinklaraščiuose (angl „Blogs“). „Web feed“ turi du pagrindinius formatus: „RSS“ ir „ATOM“. Dokumentas standartiškai pateikiamas „XML“, taigi jis labiau skirtas skaityti programoms nei žmonėms.

Kartais ši technologija naudojama perduoti informacijai iš vienos svetainės į kitą ir tam nereikia žmogaus įsikišimo.

„RSS“ – „XML“ duomenų formatas, dar kitaip vadinamas žiniatinklio srautu (angl. „Web Feeds“).

„RSS“ santrumpa turi tris reikšmes. Pagal populiarumą:

- tikrai paprastas surinkimas (angl. „Really Simple Syndication“);

- išsamus puslapio aprašymas (angl. „Rich Site Summary“) - (RSS 0.91);
- RDF (angl. „Resource Description Framework Site Summary“) - (RSS 0.9, 1.0 ir 1.1).

„ATOM“ - sindikacijos standartas, pateikiamas „XML“ kalba ir naudojamas žiniatinklio srautuose (angl. „Web Feeds“).

„ATOM“ buvo sukurtas dėl daug egzistuojančių neišbaigtų „RSS“ sindikacijos formatų. Pavyzdžiui dėl suderinamumo su „XML-RPC“ problemų.

„RSS 2.0“ įrašai gali būti sudaryti iš paprasto teksto (angl. „Plain text“) arba ištaisyto „HTML“ (angl. „Escaped HTML“), galimybės identifikuoti, kuris iš jų yra naudojamas nėra. „RSS 2.0“ nėra tiksliai suderintas darbui su „HTML“ kodu, pavyzdžiui, „&“ pakeičiamas į „&“ arba „amp“. Įrašo pavadinime (angl. „Title“) esantys laužtiniai skliaustai sukelia klaidą, kadangi jie paprasčiausiai nepalaikomi. Taip pat „RSS 2.0“ turinio modelis nepalaiko „XML“ žymių kalbos, todėl susiduriama su turinio panaudojimo (angl. „Usability“) problemomis.

„ATOM“ turi daug geriau sutvarkytą elementų palaikymo sistemą. Turiniui privalo būti paskirtas vienas iš šių tipų:

- paprastas tekstas, be jokių žymių (numatytasis tipas);
- ištaisytas (angl. „Escaped“) „HTML“;
- taisyklinga (angl. „Well formed“) „XHTML“ žymių kalba;
- kitos „XML“ žymės;
- „base64-encoded binary“ turinys.

„RSS 2.0“ turi aprašymo elementą (angl. „description“), kuriame galima įdėti santrauką arba visą straipsnį. „ATOM“ šį elementą realizavo patogiau – išskirstė į du skirtingus elementus: „Summary“ ir „Content“.

Kalboms (vartojamoms įrašams identifikuoti) „RSS 2.0“ naudoja savo žymę „<language>“, „ATOM“ naudoja „XML“ „xml:lang“ atributą.

„RSS 2.0“ suteikia galimybę nurodyti el. paštą žiniatinklio srauto autoriui/savininkui (angl. „Feed author“). Tam naudojamos žymės „<managingEditor>“ ir „<webMaster>“, o įrašų autoriams nurodyti naudojama „<author>“. Kai kurie autoriai nenori dalytis savo el. paštu ir todėl naudoja „<dc:creator>“.

„ATOM 1.0“ suteikia „<author>“ ir „<contributor>“ elementus abejiems lygiams - ir srauto aprašymo, ir įrašams. Abu elementai privalo turėti nurodytą autoriaus vardą, taip pat pasirinktai galima papildomai nurodyti el. paštą arba tinklalapio adresą [11].

1.3.6. Tinklo paslaugos

Tinklo paslaugos (angl. Web service) – tai protokolų ir standartų rinkinys, naudojamas duomenų mainams tarp taikomųjų programų ir sistemų. Įvairiomis programavimo kalbomis

parašytos taikomosios programos, veikiančios skirtingose operacinėse sistemose, naudoja šias paslaugas duomenų mainams kompiuteriniuose tinkluose [12].

Technologijas, naudojamas įgyvendinant tinklo paslaugas, galima suskirstyti į keletą lygmenų (2 lentelė.). Tai perdavimo (angl. „Transportation“), pranešimų (angl. „Messaging“), apibrėžimo (angl. „Definition“), atradimo (angl. „Discovery“), kokybės (angl. „Quality of Service“) ir sujungimo lygmenys (angl. „Service Aggregation“) [13].

2 lentelė. Interneto paslaugų sandara [13]

„UDDI“	Atradimo lygmuo
„WSDL“	Apibrėžimo lygmuo
„SOAP“	Pranešimų lygmuo
„XML“	Perdavimo lygmuo
„HTTP“, „JMS“, „SMTP“, „IIOP“	

Žemiausiame, perdavimo lygmenyje, gali būti panaudoti keli protokolai, tačiau dažniausiai naudojamas „HTTP“ bei jo atmainos. Kiti galimi perdavimo protokolai yra „JMS“¹⁶, „SMTP“ ir „IIOP“¹⁷.

Pranešimų lygmenyje pagrindinės naudojamos technologijos yra „XML“ ir „SOAP“. „XML“ taip pat pagrįstos visos kitos „Web service“ technologijos bei standartai. „SOAP“ protokolas, nors ir yra plačiai naudojamas, tačiau jį galima pakeisti kitais protokolais.

Kitas yra apibrėžimo lygmuo. Tinklo paslaugos privalo turėti aprašytą sąsają. Tai reikalinga, norint užtikrinti jų suderinamumą su ateities technologijomis bei „silpną susiejimą“ (angl. „lose coupling“). „Web service“ sąsajos apibrėžimą galima suprasti kaip susitarimą tarp paslaugos ir jos naudotojo, užtikrinantį, kad paslauga atliks tam tikrus veiksmus, gaudama tam tikrus parametrus. Tai sąsajos (angl. „Inteface“) analogas objektinio programavimo kalbose [12]. Tokiam paslaugos apibrėžimui yra sukurta „WSDL“¹⁸. „WSDL“ aprašams yra numatyti du panaudojimo atvejai. Vienu atveju aprašymas naudojamas paslaugos kvietimui, kitu – standartinei paslaugos dokumentacijai. „WSDL“ kalba paremta „XML“, tačiau „XML“ schemos apibrėžimo kalbą galima pasirinkti laisvai. Taip pat „WSDL“ numatyta galimybė naudoti ne tik „SOAP“ protokolą.

Siekiant suteikti galimybę išplėsti kalbą, į „WSDL“ buvo įtraukti tik būtiniausi „Web service“ aspektai. Tai yra informacinių pranešimų aprašymai, paslaugos atliekamos operacijos bei paslaugos adresas ir palaikomas protokolas. Reikėtų paminėti tai, jog pirmi du elementai vadinami abstrakčia

¹⁶ JMS (angl. Java Message Service) – standartizuota „Java“ „API“ (taikomųjų programų programavimo sąsaja (angl. „Application Programming Interface“)), skirta komunikacijai pranešimais realizuoti tarp dviejų ar daugiau programų sistemų ar jos dalių.

¹⁷ IIOP (angl. internet inter ORB protocol) – GIOP (ang. „general inter ORB protocol“ – dviejų kokių nors ryšio kanalų (paprastai kompiuterių tinklu) susijusių programų bendravimo protokolas) versija, papildyta standartais, reikalingais dirbti internete.

¹⁸ WSDL (angl. Web Services Definition) – speciali kalba „XML“ pagrindu skirta apibūdinti interneto paslaugoms (angl. Web Services).

aprašo dalimi, o kiti du – konkrečiai. Abstrakti dalis turėtų būti naudojama standartinės paslaugos dokumentacijai. Taip pat reiktų atkreipti dėmesį į tai, jog „WSDL“ neatspindi „Web service“ semantikos. Tai yra iš aprašo negalima tiksliai pasakyti nei ką daro „Web service“ metodai, nei kokia tvarka juos reiktų kviešti [13].

Atradimo lygmenyje, naudojama „UDDI“¹⁹ technologija. „UDDI“ specifikacija aprašo registrą, skirtą informacijai apie paslaugas bei jas teikiančias organizacijas kaupti. Toks registras yra būtinas, norint pasinaudoti vienu iš „SOA“²⁰ privalumų – daugkartiniu komponentų panaudojimu. Programinės įrangos kūrėjams reikėjo vietos, kur būtų saugomos nuorodos į paslaugas bei informacija, kaip su jais elgtis [12]. Tokia informacija suteikia galimybę dinamiškai surasti ir iškviešti reikiamą paslaugą tiek sistemos kūrimo, tiek jos veikimo metu. Pradžioje buvo planuojama, kad internete bus globalus viešas registras, kuriame visi galės dėti informaciją apie save ir savo teikiamas paslaugas. Globalaus registro idėją palaikė tokios kompanijos kaip „IBM“, „Microsoft“, „SAP“ ir kitos. Tačiau ši idėja nepasiteisino. Vis dėlto privatūs registrai, veikiantys kompanijų vidiniuose tinkluose, paplito gana plačiai. „UDDI“ registre gali būti saugoma informacija apie viską, ne vien tik apie verslo organizacijas. „UDDI“ įrašo struktūrą sudaro trys dalys. Tai baltieji, geltonieji ir žalieji puslapiai. Baltieji puslapiai – tai bendro pobūdžio informacija apie organizaciją. Geltonieji puslapiai skirti konkretesniam verslo aprašymui. Tai gali būti teikiamų paslaugų sąrašas. Taip pat čia saugomi verslo klasifikatoriai. Klasifikatorius identifikuoja verslą pagal tam tikrą klasifikavimo sistemą. Tai gali būti geografinis indeksas ar kita standartinė sistema. Klasifikavimo sistemą taip pat gali aprašyti ir pats naudotojas. Žaliuosiuose puslapiuose saugoma techninė informacija apie „Web service“. Tai nuoroda į pačią paslaugą, jos „WSDL“ aprašą ir pan.

¹⁹ UDDI (angl. Universal Description, Discovery, and Integration) – technologija interneto paslaugoms aptikti ir naudoti.

²⁰ SOA (angl. Service Oriented Architecture) - tai toks verslo taikomųjų programų kūrimo būdas, kai programiniai procesai suskaidomi į paslaugas, kurios yra pateikiamos ir įmanomos surasti tinkle.

1.4. Skyriaus išvados

- Bendrovės „Nielsen“ atliktas tyrimas parodė, kad socialinių tinklų portalus, tokius kaip „MySpace“, „Facebook“ ir kt., žmonės ėmė tikrinti dažniau nei savo pačių asmeninius el. paštus.
- Anot tyrėjų („Nielsen“), socialinių tinklų portalai dabar yra ketvirti pagal populiarumą, o juos lanko 67 % visų internautų. Pagal lankomumo rodiklius jie nusileidžia tik paieškos tinklalapiams, žinių portalams bei programinės įrangos tinklalapiams.
- Socialinio tinklo internete sąvoka remiasi „Web 2.0“ ideologija. „Web 2.0“ – terminas, apibūdinantis besikeičiančias žiniatinklio technologijos ir dizaino naudojimo tendencijas, kai yra siekiama padidinti kūrybiškumą, keitimąsi informacija bei bendradarbiavimą tarp naudotojų.
- Pagrindinė „Web 2.0“ idėja – atskirti pateikimo dizainą nuo pateikiamo turinio.
- Kažkada lygiagrečiai kurta serverių bei klientų programinė įranga dabar koncentruojasi serverio pusėje.
- Internetas, kaip žinių šaltinis, pamažu „virsta“ centralizuota platforma.

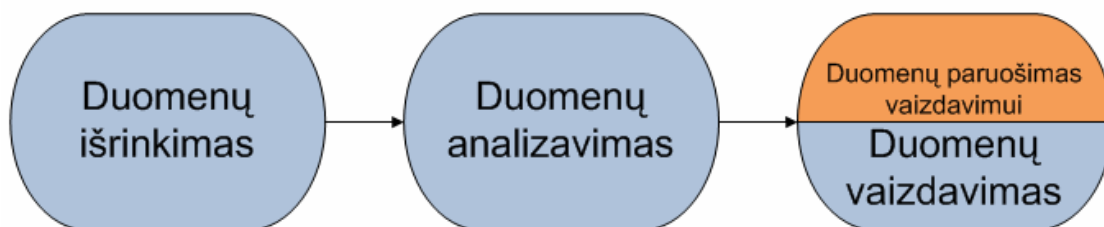
2. INTERNETINIŲ SOCIALINIŲ TINKLŲ TYRIMO PROCESO MODELIS, TYRIMO ĮRANKIAI

2.1. Skyriaus įvadas

Socialiniams tinklams eksponentiškai didėjant ir populiarėjant, kartu didėja ir poreikis juos analizuoti. Analizavimo galimybėmis internetiniai socialiniai tinklai kur kas pranašesni už realų pasaulį. Daugeliu atvejų analizuojama informacija jau yra tinkle, didžiausia problema – jos išrinkimas. Ką galima sužinoti išanalizavus vieną ar kitą socialinį tinklą – tik išmonės ir vaizduotės klausimas, kurio ribų čia nenagrinėta.

Rinkoje šiuo metu nėra jokių oficialių socialinių tinklų standartų. Nors iš esmės visi socialiniai tinklai yra vienodi, jų realizavimo ypatybės priklauso nuo jų kūrėjų išmonės. Tai apsunkina juos analizuojančių įrankių kūrimą, nes kiekvieną įrankį reikia adaptuoti specialiam tinklui (jeigu norima procesą automatizuoti). Geroji programavimo praktika reikalauja atskirti bet kokio proceso loginius etapus, taip išskaidyta programa būna kur kas universalesnė, atskiromis jos dalimis gali rūpintis atitinkamos srities specialistai. Todėl šiame darbe į socialinių tinklų analizę žiūrima kaip į procesą, susidedantį iš 3 etapų (5 paveikslėlis):

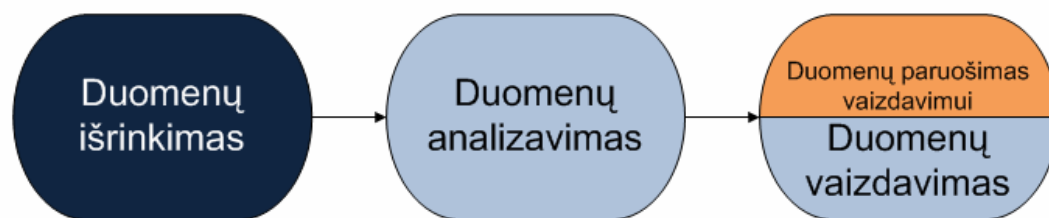
- duomenų išrinkimas;
- duomenų analizavimas;
- duomenų vaizdavimas.



5 pav. Socialinių tinklų analizės proceso etapai

Šio skyriaus tikslas – apžvelgti kiekviename socialinės tinklų analizės (kaip proceso) etape siūlomus įrankius, problemas ir praktikas.

2.2. Duomenų išrinkimas



6 pav. Duomenų išrinkimas

Yra du būdai, kaip galima išanalizuoti socialinį tinklą.

Pirmuoju atveju analizuojamas visas socialinis tinklas (dažniausiai įmonė arba kelių įmonių tinklas). Tokios analizės tikslai dažniausiai susiję su analizuojamos įmonės veiklos optimizavimu. Ko įmonės nenori matyti – tai pernelyg centralizuoto tinklo, kuris, kaip praktika parodė, susiduria su dideliais sunkumais vienam iš centrinių mazgų pasitraukus iš sistemos. Kita vartus, įmonės nenori ir per daug išplėsto tinklo, nes kuo mažesnis socialinio tinklo rangas (mažiausias atstumas tarp 2 tolimiausių mazgų), tuo greičiau nukeliauja duomenys iš vieno tinklo galo į kitą. Kaip ir daugelyje kitų dalykų, optimalus išsidėstymas slypi kažkur tarp dviejų kraštutinių.

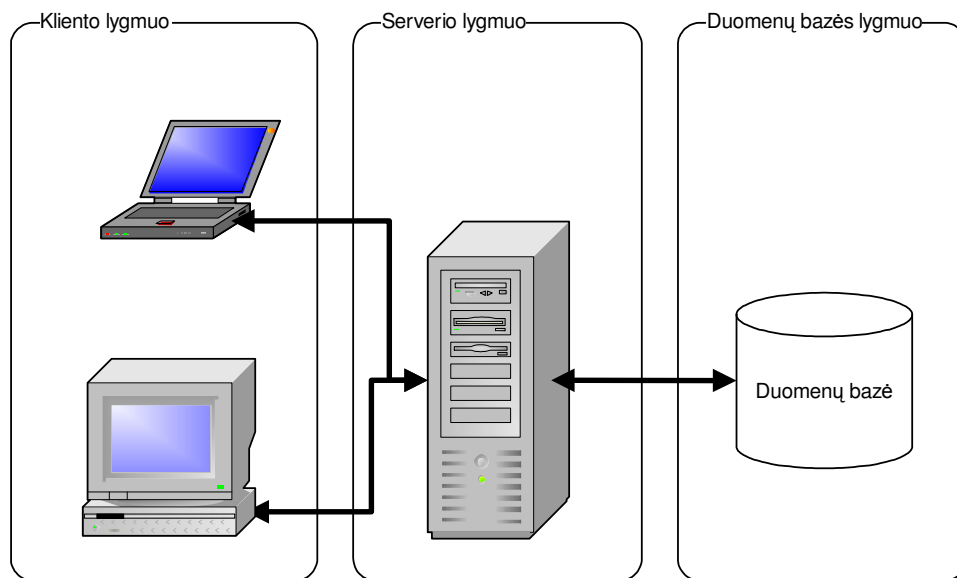
Socialinio tinklo mazgais gali būti ne tik individai, neretai jo dalimis tampa ir organizaciniai vienetai, taip tinklas tampa labiau centralizuotu be nepageidaujamo šalutinio poveikio (mazgų pasitraukimo iš sistemos).

Tiriant įmonių grupę, išryškėja jas siejantys mazgai, kurių įmonės nori matyti kuo daugiau.

Antruoju atveju analizuojama tik dalis socialinio tinklo. Neretai visas socialinis tinklas būna per didelis objektyviai analizei, jį sudaro pernelyg skirtingi mazgai bei grupės. Tokios analizės tikslai gali būti labai įvairūs, nuo naudotojo daromos įtakos savo „draugams“ iki tam tikrų interesų grupių išsidėstymo pasaulyje.

Kokį atvejį beanalizuotume, šio etapo tikslas – išrinkta tiksli reikalingų duomenų struktūra. Pagrindinės problemos: kaip ištraukti reikalingus duomenis ir kaip juos pateikti kitiems etapams. Tai, nepriklausomai nuo analizuojamos sistemos, galima įgyvendinti keliais būdais.

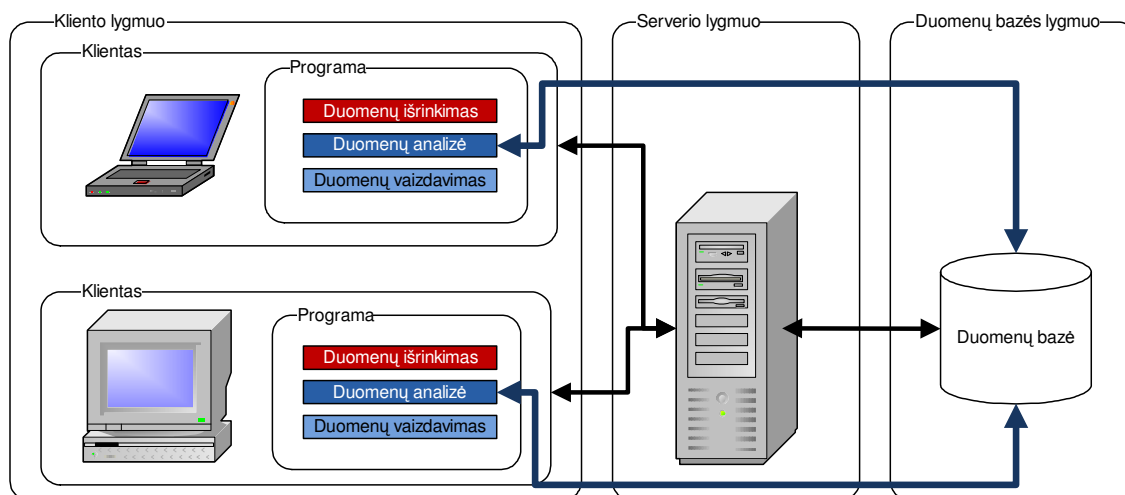
Įprastu atveju naudotojui prisijungus prie sistemos, duomenys apie jo padėtį socialiniame tinkle pateikiami tik vizualiai - „HTML“ kalba. Visą darbą atlieka serveris, kuris ne tik rūpinasi, kaip pateikiama informacija, bet ir pačios informacijos turiniu, kurį esant reikalui (jei informacijos neturi savo virtualioje atmintyje) parsiunčia iš duomenų bazės. Tokios sistemos architektūros reikalauja „Web 2.0“. Taip įgyvendintos sistemos labai apsunkina serverio darbą, tačiau naudotojui leidžia prisijungti paprasčiausia naršykle (7 paveikslėlis).



7 pav. Informacijos šaltiniai įprasto prisijungimo atveju

Išskirti reikalingos duomenų struktūros iš „HTML“ kodo praktiškai neįmanoma. Net ir atradus tam tikrą dėsningumą ar pritaikius keletą informaciją išrenkančių programinių modulių, susiduriama su daugybe problemų (pvz., tokius modulius reikėtų atnaujinti kiekvieną, kartą kai būtų atnaujintas puslapio dizainas).

Vienu iš galimų variantų galėtų būti tiesioginis kreipimasis į informacijos duomenų bazę (8 paveikslėlis).



8 pav. Pirmasis duomenų išrinkimo būdas

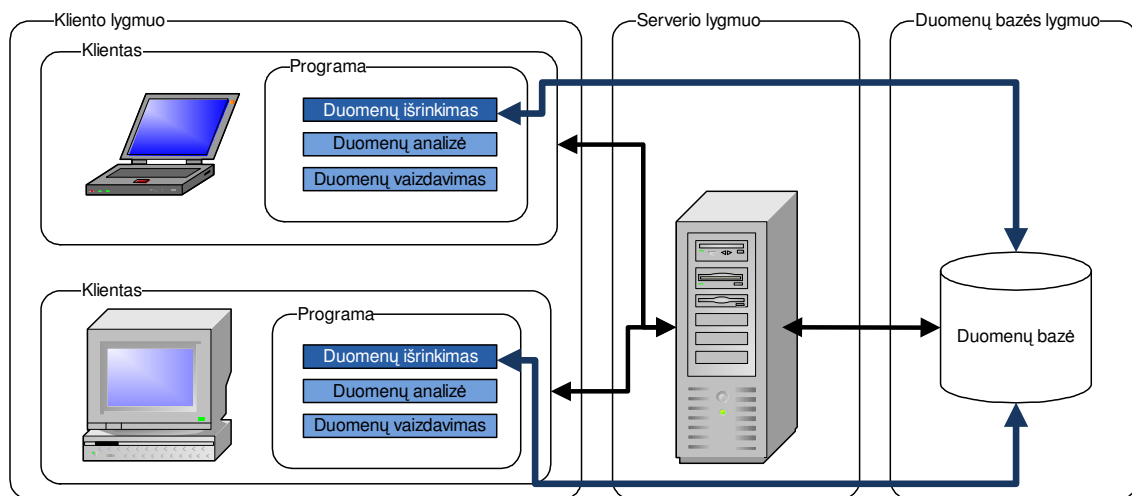
Sprendimo privalumai:

- Perteklinių duomenų nebuvimas (sistema galėtų atsisiųsti tik konkrečioms skaičiavimams reikalingus duomenis skaičiavimo metu (tačiau tam reikėtų sujungti du proceso etapus)).
- Sutaupyta laikas (duomenis naudotojas galėtų pasiekti tiesiogiai, ne per serverį).

Sprendimo trūkumai:

- Sudėtinga programinė kliento įranga.
- Dirbti būtų galima tik prisijungus prie duomenų bazės.
- Du proceso loginiai etapai būtų sujungti į vieną (prieštarauja gerai programavimo praktikai).
- Tinklo tyrėjas beveik visada turėtų turėti duomenų bazės kliento programinę įrangą (taip pat mokėti ja naudotis).
- Tinklo tyrėjas turėtų gerai žinoti duomenų struktūrą bazėje (o programa būtų labai jautri bet kokiems duomenų struktūros pokyčiams).
- Papildomai apkrauta duomenų bazė (Duomenų bazių ištekliai dideliuose socialiniuose tinkluose labai brangūs).

Kitu sprendimu galėtų būti tiesioginis kreipimasis į informacijos duomenų bazę ir visos galimai reikalingos informacijos parsisiuntimas (9 paveikslėlis).



9 pav. Antrasis duomenų išrinkimo būdas

Sprendimo privalumai:

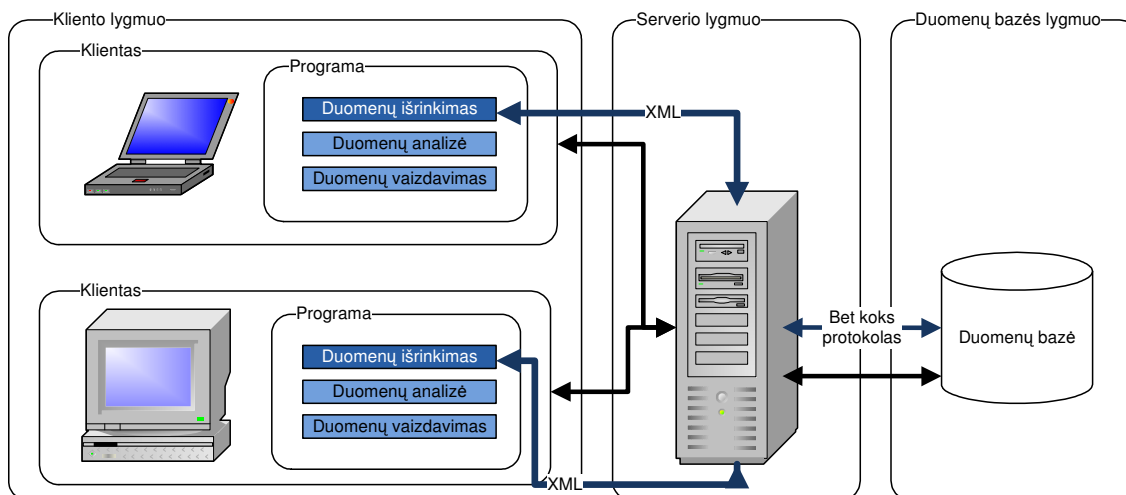
- Parsisiuntus duomenis, galima su jais dirbti be ryšio su duomenų baze.
- Sutaupyta laikas (duomenis naudotojas galėtų pasiekti tiesiogiai, ne per serverį).

Sprendimo trūkumai:

- Sudėtinga programinė kliento įranga.
- Atsisiųsti tektų viską (tiksliai nežinodami ką tirsim, privalėtume parsisiųsti visus duomenis, o naudoti tik reikiamus).
- Tinklo tyrėjas beveik visada turėtų turėti duomenų bazės programinę kliento įrangą (taip pat mokėti ja naudotis).

- Tinklo tyrėjas turėtų gerai žinoti duomenų struktūrą bazėje (programa būtų labai jautri bet kokiems duomenų struktūros pokyčiams).
- Papildomai apkrauta duomenų bazė (Duomenų bazių ištekliai dideliuose socialiniuose tinkluose labai brangūs).

Idealiu atveju duomenų pateikimu galėtų pasirūpinti socialinio tinklo kūrėjai, išplėsdami savo serverių funkcionalumą (10 paveikslėlis).



10 pav. Trečiasis duomenų išrinkimo būdas

Sprendimo privalumai:

- Paprasta programinė kliento įranga.
- Parsisiuntus duomenis, galima su jais dirbti be ryšio su duomenų baze.
- Neužimta duomenų bazė.
- Tinklo tyrėjui būtų nebūtina išmanyti duomenų struktūros duomenų bazėje.
- Tinklo tyrėjui nereikėtų jokios papildomos programinės įrangos.

Sprendimo trūkumai:

- Duomenis turėtų pateikti socialinių tinklų savininkai.
- Duomenų išrinkimas užtruktų ilgiau.
- Atsisiųsti tektų viską (tiksliai nežinodami ką tirsim, privalėtume parsisiųsti visus duomenis, o naudoti tik reikiamus).

Apžvelgus visus privalumus ir trūkumus, nesunku nustatyti, kad socialinių tinklų tyrėjams parankiausia būtų, jeigu duomenų pateikimu pasirūpintų portalų savininkai. Taip būtų apsunkinamas socialinių tinklų darbas, nes jie, be įprastų funkcijų, dar turėtų siūlyti ir duomenų struktūrų perdavimo paslaugas, tačiau alternatyvūs būdai nedžiugina nei sistemos tyrėjų, nei pačių savininkų.

Dėl milžiniškos duomenų apimtys populiariausių socialinių portalų duomenų bazės būna specifiskai optimizuotos, kaip jose laikomi duomenys ir kaip vyksta duomenų mainai tarp duomenų

bazės ir serverio – komercinės paslaptys, kuriomis portalų savininkai nelinkę dalytis. Socialinių tinklų kūrėjai gali rinktis, arba išvis neteikti tokios informacijos („one.lt“, „klase.lt“) arba išplėsti savo serverių funkcionalumą ir tokią informaciją teigti per juos („facebook“, „myspace“, „twitter“). Kodėl socialinių portalų savininkai turėtų teigti tokią informaciją? Tikrai ne dėl jų analizės. Vėl iškyla „Web 2.0“ idėja, kuri teigia, kad socialinių tinklų turinį kuria sistemos naudotojai. Viskas susiveda į portalo siūlomus turinio kūrimo įrankius. Vieni portalai leidžia savo naudotojams talpinti nuotraukas, nuomones, video ar audio informaciją, kiti savo funkcionalumą dar labiau išplečia suteikdami naudotojams galimybę sistemą išnaudoti kaip platformą ir kurti savo įrankius ar programėles.

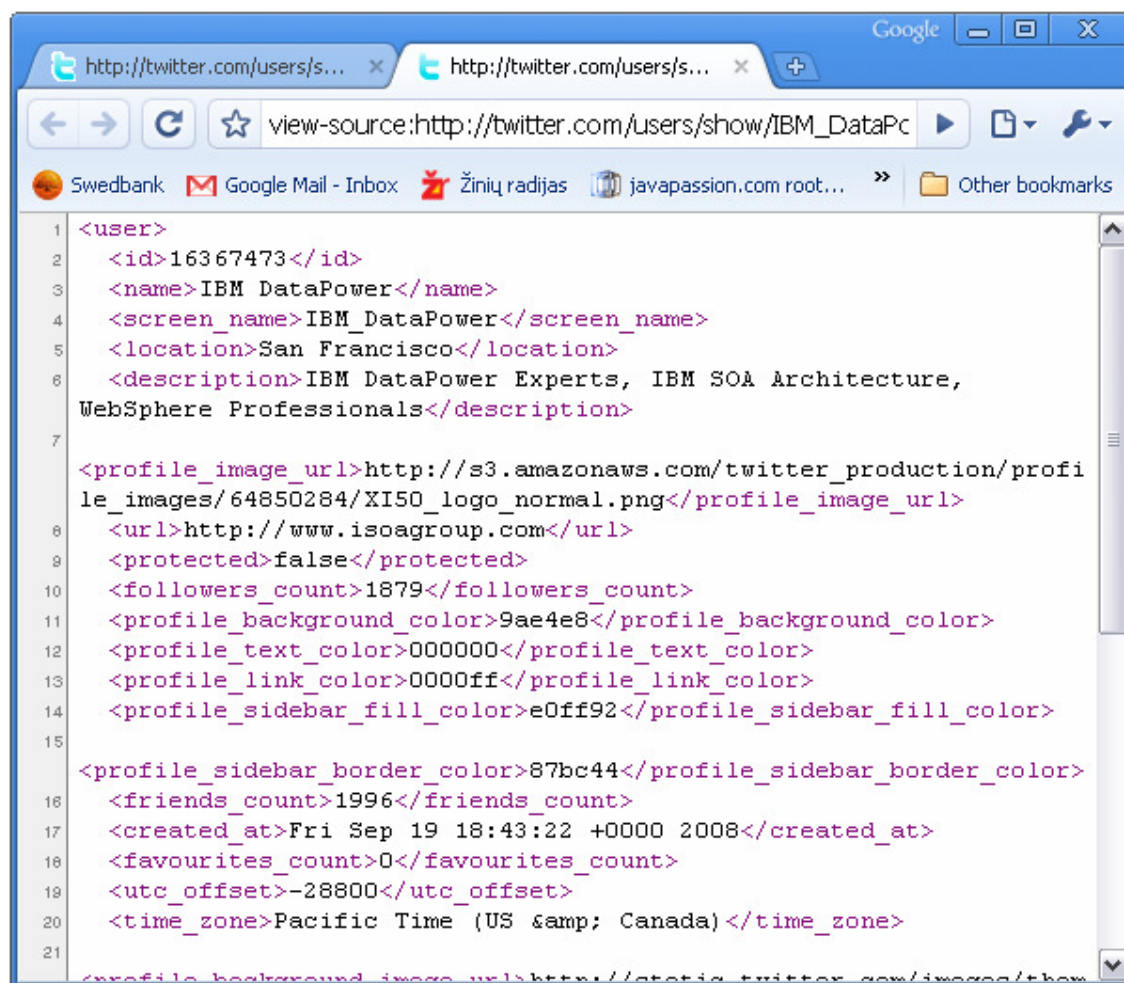
Akivaizdu, kad naudotojų kuriamoms programoms reikalinga kitokia (labiau standartizuota) sąsaja su sistema.

Dar vienas klausimas su kuriuos susiduria socialinių tinklų savininkai - kaip jie turėtų pateikti informaciją apie savo tinklų struktūrą naudotojų kuriamoms programoms. Vienas iš sprendimų – duomenis perduoti kaip tekstą. Nors ir paprasta išeitis, tačiau ne labai „grakšti“, vargu ar atsirastų daug naudotojų norinčių gilintis į tokį specifišką sprendimą. Toks būdas neabejotinai sumažintų ir sistemos universalumą, nes vargu ar taip pateikiamai informacijai būtų galima sugalvoti daugiau panaudojimo būdų.

Kitas sprendimo būdas - perduoti tokią informaciją standartizuotai – duomenų struktūrų bei jų turinio perdavimo kalba – „XML“. „XML“ kalba labai patogi neapibrėžto dydžio duomenų kiekiams perduoti, nes visi ja perduodami elementai turi vardus (visi „XML“ kalbos elementai randami per vardus arba vardų kelius). Taip perduodamų duomenų struktūras galima vėliau nesunkiai išplėsti papildomais laukais, nesibaiminant, kad visos iki tol naudotojų parašytos programos nustos veikti. „XML“ kaip duomenų perdavimo standartą „W3C“ (angl. „World Wide Web Consortium“) rekomenduoja jau 11 metų.

Socialinis portalas „Twitter“ siūlo net 32 metodus įvairiems duomenims gauti ir manipuluoti. Šiais 32 metodais įgyvendinamas visas portalo funkcionalumas, o tai reiškia, kad sistemos naudotojai „Twitter“ portalo klientą (programą su grafine sąsaja, sugebančią perduoti informaciją iš serverio naudotojui) gali susikurti bet kokioje „REST“ protokole ir „XML“ kalbą suprantančioje platformoje (Tarp kurių ne tik visos pagrindinės operacinės sistemos, bet ir mobilieji telefonai, įvairūs nešiojami įrenginiai) [14].

11 paveikslėlyje pavaizduotas vieno iš „Twitter“ metodų sugeneruotas atsakymas „XML“ kalba.

A screenshot of a web browser window displaying the source code of a Twitter user profile. The browser's address bar shows the URL 'http://twitter.com/users/show/IBM_DataPc'. The source code is an XML document representing a user profile. The code is color-coded with syntax highlighting. The user's name is 'IBM DataPower', location is 'San Francisco', and description is 'IBM DataPower Experts, IBM SOA Architecture, WebSphere Professionals'. The profile image URL is 'http://s3.amazonaws.com/twitter_production/profile_images/64850284/XI50_logo_normal.png'. The user has 1879 followers and 1996 friends. The profile was created on 'Fri Sep 19 18:43:22 +0000 2008'. The time zone is 'Pacific Time (US & Canada)'.

```
1 <user>
2   <id>16367473</id>
3   <name>IBM DataPower</name>
4   <screen_name>IBM_DataPower</screen_name>
5   <location>San Francisco</location>
6   <description>IBM DataPower Experts, IBM SOA Architecture,
   WebSphere Professionals</description>
7
8   <profile_image_url>http://s3.amazonaws.com/twitter_production/profi
   le_images/64850284/XI50_logo_normal.png</profile_image_url>
9   <url>http://www.isoagroup.com</url>
10  <protected>>false</protected>
11  <followers_count>1879</followers_count>
12  <profile_background_color>9ae4e8</profile_background_color>
13  <profile_text_color>000000</profile_text_color>
14  <profile_link_color>0000ff</profile_link_color>
15  <profile_sidebar_fill_color>e0ff92</profile_sidebar_fill_color>
16  <profile_sidebar_border_color>87bc44</profile_sidebar_border_color>
17  <friends_count>1996</friends_count>
18  <created_at>Fri Sep 19 18:43:22 +0000 2008</created_at>
19  <favourites_count>0</favourites_count>
20  <utc_offset>-28800</utc_offset>
21  <time_zone>Pacific Time (US & Canada)</time_zone>
```

11 pav. „Twitter“ „XML“ kalba pateikti duomenys (2009 m. balandis 19 d.)

Portalas „Facebook“ siūlo net 78 metodus duomenims gauti ir manipuluoti, kurie programoms pateikia net paprastiems naudotojams nepasiekiamų duomenų. Portalas reikalingus duomenis gali pateikti net 3 būdais („XML“, „JSON“²¹, tinklalapių kūrimo kalbos „PHP“²² kintamųjų tipais). Naudotojams siūloma ir testavimo aplinka (12 paveikslėlis) bei turtinga kiekvieno metodo dokumentacija [15].

²¹ JSON (angl. JavaScript Object Notation) – paprastas, tekstu paremtas, žmogui įskaitomas duomenų perdavimo formatas.

²² PHP – plačiai paplitusi dinaminė interpretuojama programavimo kalba (angl. Hypertext Preprocessor), sukurta 1997 m. ir specialiai pritaikyta interneto svetainių kūrimui.

facebook
DEVELOPERS
Documentation
Community
Resources
Tools
News

Tools

API Test Console
FBML Test Console
Feed Template Console
Registered Templates Console

You can experiment with functions and responses, and see what content Facebook Platform makes available. Select the method you wish want to call and the format of the return values.

User ID
745489727

Response Format
XML

Callback

Method (Documentation)
friends.get

Call Method

\$facebook->api_client->friends_get("");

<?xml version="1.0" encoding="UTF-8"?>
<friends_get_response xmlns="http://api.facebook.com/1.0/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-i
<uid>550757489</uid>
<uid>588704385</uid>
<uid>624795251</uid>
<uid>670434499</uid>
<uid>731203853</uid>
<uid>776984082</uid>
<uid>1221241082</uid>
<uid>1265200682</uid>
<uid>1314185842</uid>
<uid>1477238460</uid>
<uid>1549789111</uid>
</friends_get_response>

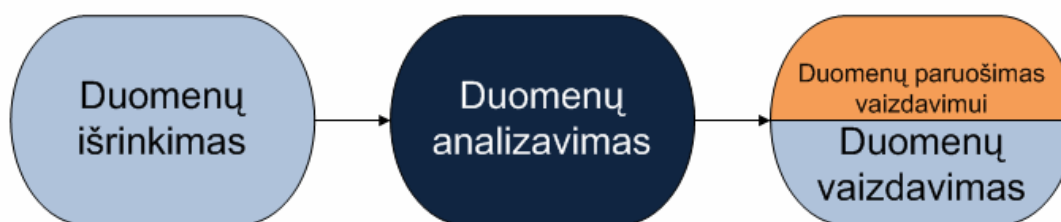
Facebook © 2009
About
Terms of Service
Privacy Policy

12 pav. Socialinio portalo „Facebook“ testavimo aplinka (2009 m. balandis 19 d)

Rinkoje yra ir daugiau socialinių tinklų siūlančių duomenų struktūrų perdavimo paslaugas, dauguma duomenis pateikia „XML“ kalba, tačiau elementus vadina savaip. Dėl pastarosios priežasties neįmanoma sukurti bendro duomenų išrinkimo įrankio visiems socialiniams tinklams, juos reikia adaptuoti konkretiems produktams.

Apibendrinamas norėčiau dar kartą atsakyti į šio skyriaus klausimus. Socialinių tinklų tyrimui reikalingus duomenis turėtų pateikti jų pačių savininkai. Pagrindiniai tyrėjo įrankiai: „REST“ arba „SOAP“ protokolai, kuriais pateikiami „XML“ kalba aprašyti duomenys. Kitiems etapams duomenys turėtų būti perduodami ta pačia „XML“ kalba. Viskas kuo turėtų pasirūpinti tinklo tyrėjas – tai teisingas elementų įvardinimas.

2.3. Duomenų analizė



13 pav. Duomenų analizavimas

Duomenų analizė – antras ir kai kuriais atvejais paskutinis proceso etapas (jeigu duomenų nereikia atvaizduoti grafiškai). Kompiuterio galimybė atlikti daug operacijų per minimalų laiką tiesiog tobulai tinka tokiam darbui.

Praeitame etape priėjome išvados, kad duomenys analizei turėtų būti perduodami „XML“ kalba, todėl šiame etape analizuoti būtent taip aprašyti duomenys.

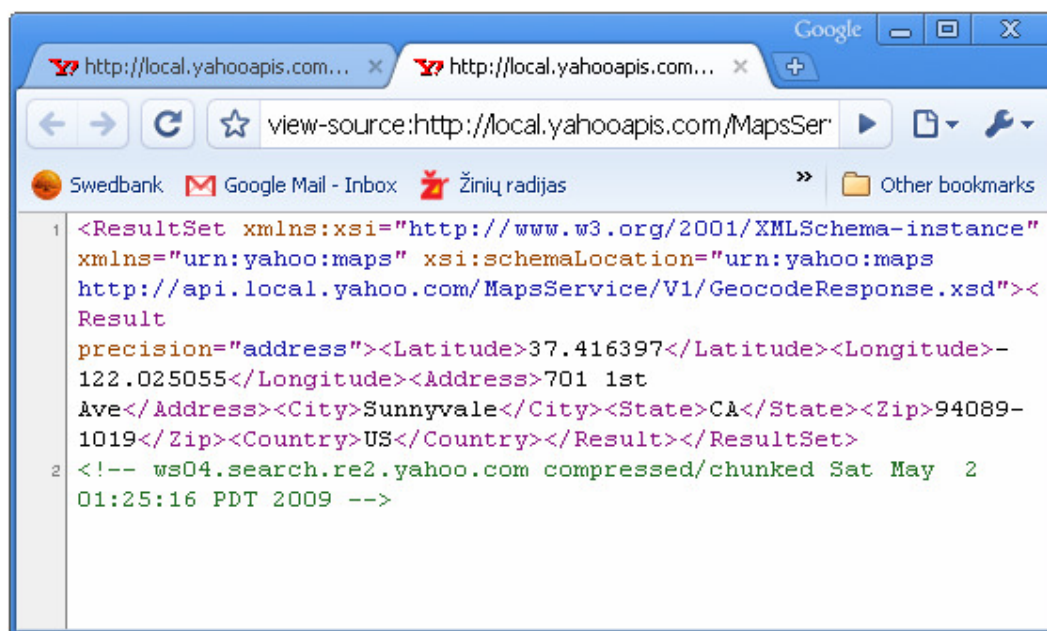
Šio etapo tikslas – apžvelgti kaip atvaizduojamos realios struktūros „XML“ kalba ir kaip tam tikri jos duomenys gali būti interpretuojami.

Svarbu suvokti, kad „XML“ duomenų modelis yra medis, sudarytas iš elementų, atributų ir kitokio tipo mazgų. Tekstinė sintaksė – tai tik „XML“ duomenų modelio serializacija. Norima struktūra išgaunama naudojantis elementais, kurie tekstone „XML“ sintakse atvaizduojami kaip žymės (angl. „tags“). „XML“ standartas nustato bendras jų sudarymo taisykles, bet neapibrėžia konkretaus elementų rinkinio ar jų reikšmės. Tai apibrėžia kiti „XML“ šeimos standartai („XHTML“, „XSLT“ ir t.t.). „XML“ galima pavadinti „kalba kitoms kalboms aprašyti“.[17]

Darbe jau minėta, kad iš esmės visi socialiniai tinklai vienodi, jų pagrindą sudaro tarpusavyje tam tikrais ryšiais susiję mazgai. „XML“ kalba, kiekvienas mazgas dažniausiai aprašomas kaip atskiras elementas, o jo ryšiai atvaizduojami per vieną iš jo laukų – „draugų“ sąrašą.

Viena iš problemų su kuria susiduria socialiniu tinklų analizuotojai – duomenų nepilnumas. Naudotojai, kurdami savo profilius, informaciją apie save įveda patys. Tokios informacijos praktiškai neįmanoma patikrinti, todėl dažniausiai ji išsaugoma taip, kaip buvo pateikta. Tikslumo ir išbaigtumo reikalaujantys duomenys tokie kaip naudotojo adresas neretai įvedami vienoje eilutėje. Nors ir daugumai kitų naudotojų tokie duomenys būna suprantami, sistemoms - toli gražu ne. Tokių duomenų įvertinimui reikalingi papildomi įrankiai.

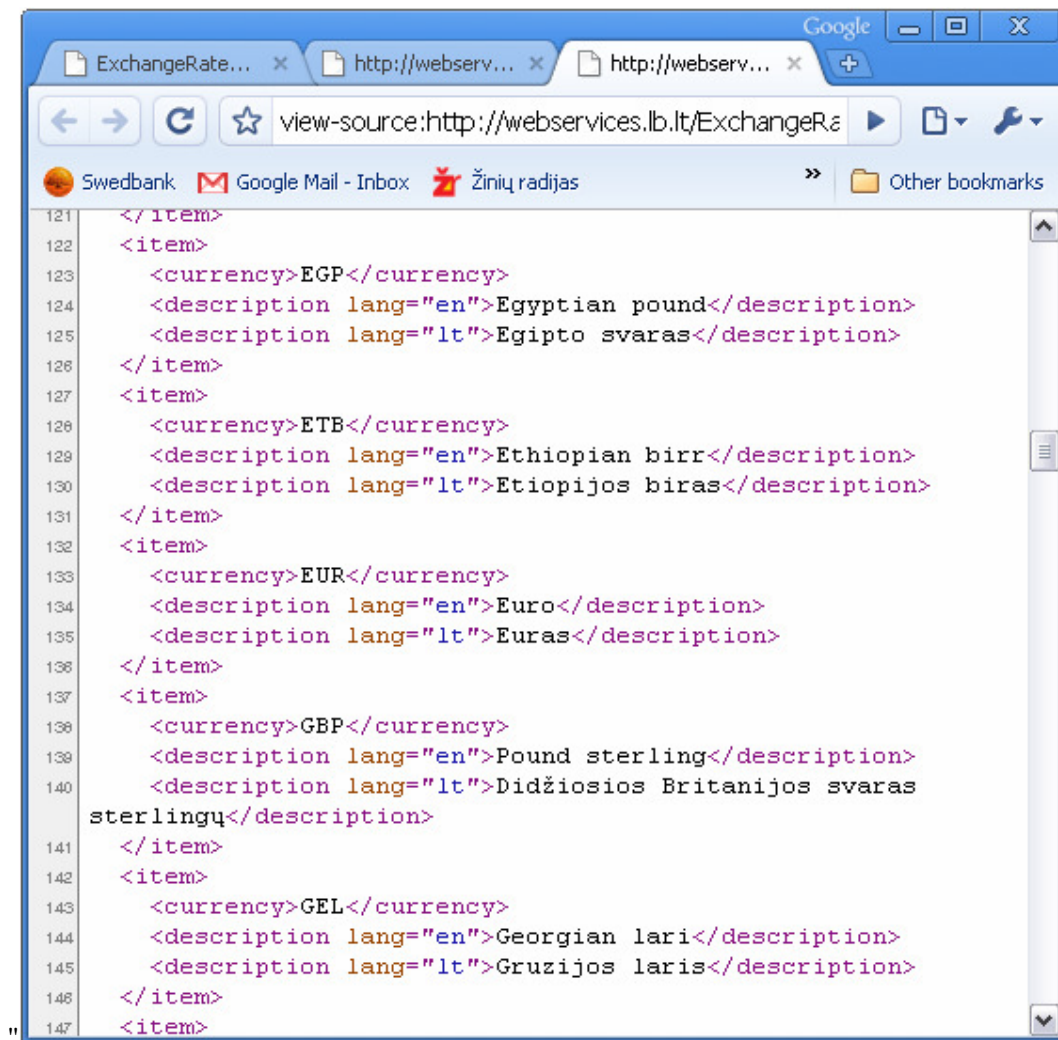
Vieną iš tokių įrankių siūlo kompanija „Yahoo“. Tai internetinė paslauga sugebanti iš minimalios informacijos nustatyti adresą pašto indeksu tikslumu bet kurioje pasaulio dalyje. Tinklo paslaugos užklausos atsakymą sudaro net 9 laukai, į kuriuos be valstybės, miesto ir pašto indekso įtraukiamos ir geografinė platuma bei ilguma, adreso tikslumą apibūdinantis laukas (14 paveikslėlis).



14 pav. „Yahoo“ siūlomos internetinės paslaugos „Geocoding“ atsakymo pavyzdys (2009 m. balandis 19 d)

Kita problema, su kuria susiduria socialinių tinklų analizuotojai – įvairūs formatai, būdingi skirtingoms šalims. Paprasti elementai, tokie kaip datos ar laikas, dažniausiai pateikiami pasauliniais standartais, bet yra ir tokių, kurių vienareikšmiškai išversti neįmanoma. Vienas iš tokių elementų – valiuta. Akivaizdu, kad prieš bet kokius skaičiavimus ar lyginimus, valiuta turėtų būti konvertuota.

Lietuvoje valiutų kursų lito atžvilgiu skelbia Lietuvos Bankas. Jis taip pat rūpinasi ir šių duomenų pateikimu interneto paslaugomis (15 paveikslėlis).



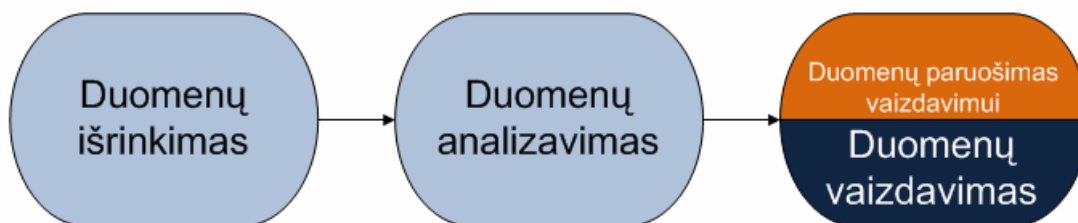
```
121 </item>
122 <item>
123   <currency>EGP</currency>
124   <description lang="en">Egyptian pound</description>
125   <description lang="lt">Egipto svaras</description>
126 </item>
127 <item>
128   <currency>ETB</currency>
129   <description lang="en">Ethiopian birr</description>
130   <description lang="lt">Etioipijos biras</description>
131 </item>
132 <item>
133   <currency>EUR</currency>
134   <description lang="en">Euro</description>
135   <description lang="lt">Euras</description>
136 </item>
137 <item>
138   <currency>GBP</currency>
139   <description lang="en">Pound sterling</description>
140   <description lang="lt">Didžiosios Britanijos svaras
sterlingų</description>
141 </item>
142 <item>
143   <currency>GEL</currency>
144   <description lang="en">Georgian lari</description>
145   <description lang="lt">Gruzijos laris</description>
146 </item>
147 <item>
```

15 pav. Valiutų kursų sąrašas (2009 m. balandis 4d.)

Kaip atliekami įvairūs skaičiavimai, priklauso nuo konkretaus įrankio. Beveik visais atvejais, tai kokia nors programavimo kalba parašytas kodas, atliekantis kelis ciklus ar palyginimus. Beveik visos šiuolaikinės programavimo kalbos siūlo daugybę pagalbinių įrankių skirtų „XML“ kalba aprašytiems duomenims manipuluoti ar interneto paslaugoms iškviesti.

Kaip skaičiavimų rezultatai turėtų būti perduoti sekančiam etapui – programuotojų išmonės reikalas. Tačiau kaip ir kitais atvejais, standartiškai perduodami duomenys („XML“) suteiktų daugiau laisvės naujų vaizdavimo (kitas etapas) įrankių integracijai.

2.4. Duomenų vaizdavimas



16 pav. Duomenų vaizdavimas

Duomenų vaizdavimas - paskutinis socialinių tinklų tyrimo proceso etapas. Priklausomai nuo konkretaus uždavinio, duomenis vaizdavimui gali tekti specifiškai apdoroti. Šį etapą išskaidžiau į dvi dalis: duomenų paruošimą ir duomenų vaizdavimą, nes net ir specifiškai paruošti duomenys gali būti suprantami keliems vaizdavimo įrankiams (pavyzdžiui „XML“ kalba aprašytus duomenis atvaizduoti gali daug programų).

Šio etapo tikslas – apžvelgti šiuo metu rinkoje siūlomus duomenų vaizdavimo įrankius.

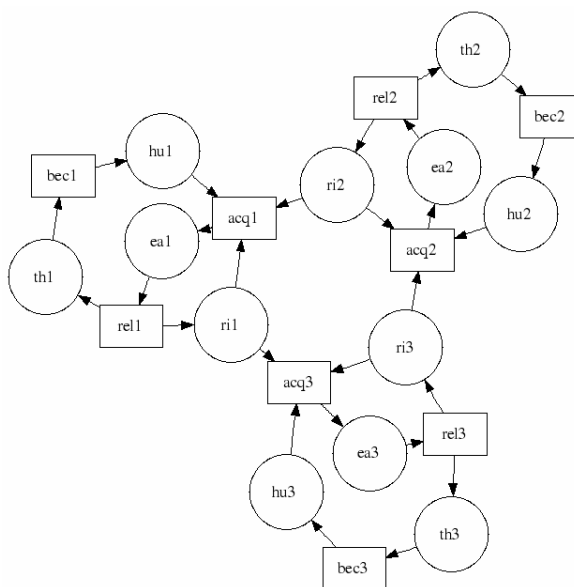
Pirmas įrankis kurį paminėsiu, kompanijos „AT&T“ darbuotojų atviro kodo projektas „Graphviz“. Aktualiausias įrankio privalumas, kuris ir daro jį tokiu patogi socialinių tinklų atvaizdavimui – galimybė duomenų struktūrą apsirašyti kodu. 17 paveikslėlyje pateiktas kodu aprašytų duomenų pavyzdys, 18 – pagal tą patį kodą atvaizduota struktūra. Tai įgalina socialinio tinklo tyrėją keisti vaizduojamos informacijos dizainą priklausomai nuo jos turinio.

```
ItakaDuom.fdp - Notepad
File Edit Format View Help
## "It encodes the so-called philosophers dilemma. Neato pretty much approximates the way how
## humans would layout the graph." Contributed by Manfred Jeusfeld.
## Command to generate the layout: "neato -Tpng thisfile > thisfile.png"

digraph PhiloDilemma {
  node [shape=box]; bec3; rel3; bec2; rel2; acq2; acq3; bec1; rel1; acq1;
  node [shape=circle, fixedsize=true, width=0.9]; hu3; th3; ri3; ea3; hu2; th2; ri2; ea2; hu1; th1; ri1; ea1;
  ri3->acq2;
  hu3->acq3;
  bec3->hu3;
  th3->bec3;
  rel3->th3;
  rel3->ri3;
  ea3->rel3;
  acq3->ea3;
  ri2->acq1;
  ri2->acq2;
  hu2->acq2;
  bec2->hu2;
  th2->bec2;
  rel2->th2;
  rel2->ri2;
  ea2->rel2;
  acq2->ea2;
  ri1->acq3;
  ri1->acq1;
  hu1->acq1;
  bec1->hu1;
  th1->bec1;
  rel1->th1;
  rel1->ri1;
  ea1->rel1;
  acq1->ea1;

  overlap=false
  label="PetriNet Model PhiloDilemma\nExtracted from ConceptBase and layed out by Graphviz "
  fontsize=12;
}
```

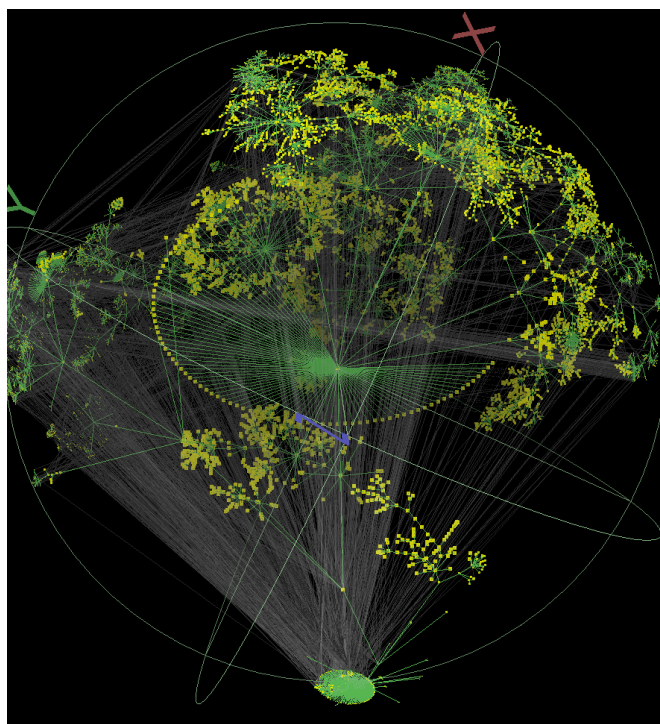
17 pav. Duomenų struktūros, aprašytos kodu pavyzdys
(<http://www.graphviz.org/Gallery/undirected/philogv.txt>)



18 pav. „Graphviz” duomenų struktūros pavyzdys (<http://www.graphviz.org/Gallery/undirected/philo.html>)

„Graphviz“ siūlo net keliolika būdų įvairioms duomenų struktūroms atvaizduoti, visas jas galima iš pradžių apsirašyti kodu. Šis įrankis ypatingai išryškina 2 atskiras etapo dalis (duomenų paruošimą ir duomenų atvaizdavimą), nes taip aprašytus duomenis sugeba atvaizduoti ir daugiau įrankių.

Vienas iš tokių įrankių, sugebančių atkurti duomenų struktūras iš kodo (paruošto įrankiui „Graphviz“) – kompanijos „Caida“ produktas „Walrus“ (19 paveikslėlis). „Walrus“ orientuotas į ypatingai dideles duomenų struktūras (iki milijono mazgų).



19 pav. „Walrus” duomenų struktūros pavyzdys (<http://www.caida.org/tools/visualization/walrus/gallery1/>)

Duomenis atvaizduoti galima ir interneto paslaugų pagalba. Tokių paslaugų lyderė – kompanija „Google“. „Google“ duomenis siūlo atvaizduoti net 9 tipų įvairaus dizaino grafikais:

- linijiniais grafikais;
- stulpeliniais grafikais;
- skritulinėmis diagramomis;
- venn diagramomis²³;
- sklaidos grafikais;
- spindulinėmis diagramomis;
- žemėlapiams;
- skalėmis;
- „QR“ kodais²⁴.

Duomenys tokiai paslaugai siunčiami kaip „URL“²⁵ parametrai (oficialiai „URL“ ilgis neribojamas). Akivaizdu, kad šiuo atveju duomenų paruošimo atskirti nuo jų vaizdavimo neįmanoma.

„Google Chart API“ užklausos pavyzdys (rezultatas pavaizduotas 20 paveikslėlyje):

<http://chart.apis.google.com/chart?>

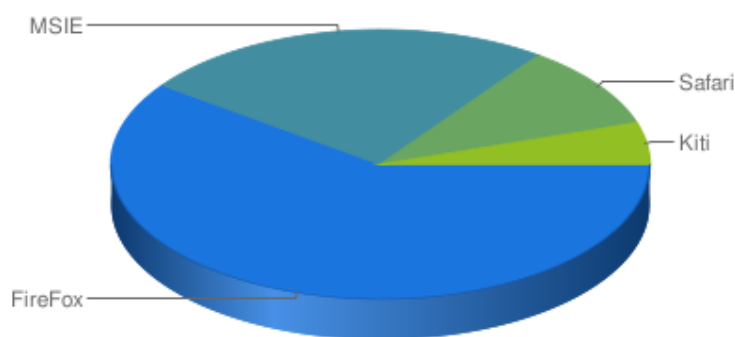
[cht=p3](#)

[&chd=s:8ZKF](#)

[&chs=400x200](#)

[&chl=FireFox|MSIE|Safari|Kiti](#)

[&chco=1B75DF,92BF23](#)



20 pav. „Google Chart API“ pavyzdys (2009 m. gegužė 3 d)

²³ Venn diagrama - diagrama, naudojama užkloties sritims tarp elementų rodyti.

²⁴ QR kodas - nedidelis baltų ir juodų taškų tinklėlis, kuriame gali būti užkoduota net iki 3 KB duomenų.

²⁵ URL (angl. Universal Resource Locators) - konkretaus informacijos šaltinio (tinklalapio, vieno tinklalapio puslapio) adresui nurodyti internete sukurtas standartas.

Dar vieną įrankį per vieną iš savo atviro kodo projektų („CodePlex“) siūlo kompanija „Microsoft“. Įrankis, pavadintas „NodeXL“ (21 paveikslėlis), iš tiesų yra vienas iš „Microsoft“ kompanijos produkto „Excel“ šablonų. Duomenys šiam įrankiui turi būti pateikti lentele.

21 pav. Duomenų pateikimas „Excel“ šablonu „NodeXL“ (<http://www.codeplex.com/NodeXL>)

2.5. Skyriaus išvados

- Socialinių tinklų tyrimui reikalingus duomenis turėtų pateikti jų pačių savininkai. Pagrindiniai tyrėjo įrankiai: „REST“ arba „SOAP“ protokolai, kuriais pateikiami „XML“ kalba aprašyti duomenys. Kitiems etapams duomenys turėtų būti perduodami „XML“ kalba. Viskas kuo turėtų pasirūpinti tinklo tyrėjas – teisingas elementų įvardinimas.
- Kaip atliekami įvairūs skaičiavimai priklauso nuo konkretaus įrankio. Beveik visais atvejais, tai kokia nors programavimo kalba parašytas kodas atliekantis kelis ciklus ar palyginimus. Beveik visos šiuolaikinės programavimo kalbos siūlo daugybę pagalbinių įrankių skirtų „XML“ kalba aprašytiems duomenims manipuluoti ar interneto paslaugoms iškviešti.
- Socialinių tinklų vaizdavimą (kaip proceso etapą) galima išskaidyti į dvi dalis: duomenų paruošimą ir duomenų vaizdavimą, nes net ir specifiškai paruošti duomenys gali būti suprantami keliems vaizdavimo įrankiams (pavyzdžiui „XML“ kalba aprašytus duomenis atvaizduoti gali daug programų).
- Kokiu įrankiu turėtų atvaizduoti socialinių tinklų tyrimo rezultatus priklauso tik nuo tyrimo specifikos.

3. INTERNETINIŲ SOCIALINIŲ TINKLŲ TYRIMO MODELIO REALIZAVIMAS

3.1. Skyriaus įvadas

Norint parodyti praktinę tyrimo svarbą, buvo įgyvendintas, praeitame skyriuje aptartas, 3 etapų socialinių tinklų tyrimo proceso modelis. Darbe remtasi Nathan'o Harrington'o (Programuotojas dirbantis kompanijoje „IBM“) straipsniu: „Social-networking open source visualization aids“ [18].

Įrankio kūrimui buvo pasirinkta „PERL“²⁶ programavimo kalba. „PERL“ tai scenarijų programavimo kalba dažniausiai naudojama „UNIX“²⁷ operacinėse sistemose, tačiau prie jos „neprižišta“. „PERL“ pasirinkta dėl jos paprastumo, universalumo bei didelės vartotojų sukurtų programinių modulių duomenų bazės. Kurtos programos kode panaudoti „XML::Simple“ bei „Geo::Coder::Yahoo“ kodo moduliai.

„PERL“ nėra geriausias pasirinkimas tokio tipo programai, tačiau darbo tikslas nebuvo sukurti naudojimui paruoštą įrankį, o sukurti tokio tipo programos modelį.

Be programavimo kalbos buvo naudota daug papildomų įrankių, daugiau apie juos bus paminėta tyrimo eigoje.

Programoje įgyvendinti visi 3 proceso etapai.

Programa sugeba išanalizuoti vieno socialinio tinklo duomenis ir apskaičiavus pateikti trijų tipų ataskaitas.

Kurta programa pritaikyta socialiniam tinklui „Twitter“. Būtent šis produktas buvo pasirinktas dėl vienos labai paprastos priežasties – „Twitter“ sistemos naudotojų informacija bei ryšiai su kitais sistemos naudotojais matomi visiems. Daugelis kitų sistemų tokią informaciją slepia, ir tik pats naudotojas savo noru gali ją atskleisti.

²⁶ PERL - pagal idėją tai ne santrumpa, o tiesiog pavadinimas, tačiau vėliau buvo pasiūlyti ir santrumpą atitinkantys pavadinimai, tokie kaip: „Practical Extraction and Report Language“ ar „Pathologically Eclectic Rubbish Lister“.

²⁷ UNIX – grupė operacinių sistemų, kilusių iš 1969 m. - 1970 m. sukurtos UNICS (apie 1970 pervadintos į UNIX) sistemos.

3.2. Duomenų išrinkimas

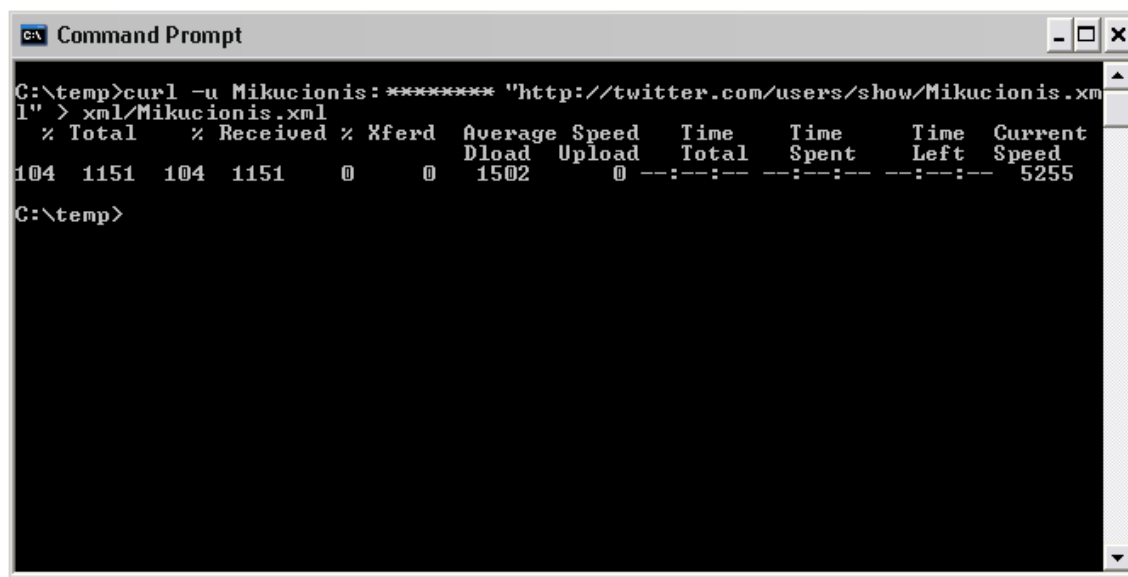
Analizuotas socialinis portalas „Twitter“ turi apie 2 milijonus naudotojų. Viso tinklo analizei reikėtų milžiniško kiekio skaičiavimo išteklių ir išskirtinių teisių sistemoje. Neturint nei vienos iš paminėtų privilegijų buvo nuspręsta analizuoti tik dalį tinklo.

„Twitter“ labai elementarus socialinis portalas. Jo esmė – suteikti galimybę konkrečiam draugų ratui nuolat sekti vienas kito veiksmus. Kiekvienas sistemos naudotojas turi savo „lentą“, kurioje gali nuolat skelbti ką veikia (vienas įrašas negali viršyti 140 simbolių (orientuota į mobiliųjų telefonų rinką)). Kiti sistemos naudotojai gali pasirinkti kieno veiksmus jie nori sekti, tai jie padaro įtraukdami konkretų naudotoją į savo draugų sąrašą.

Etape naudoti įrankiai:

- „curl“;
- „XML::Simple“ („PERL“ programinis modulis).

„curl“ labai plačiai naudojama atviro kodo programa, dėl kurios egzistavimo reikėtų dėkoti daugybei žmonių. Darbe naudotą versiją parašė Doug'as Kaufman'as (versija:7.10.5). „curl“ – veikimo principas labai paprastas, programa paprasčiausiai parsisiunčia serverio atsakymą, kurį gauna kreipdamasi pagal pateiktą „URL“ adresą (22 paveikslėlis). Pasirinktas buvo būtent šis įrankis, nes yra populiariausias, internete galima rasti daugybę šios programos alternatyvų, tačiau dėl paprasto jų funkcionalumo į alternatyvas nesigilinau.



```
C:\temp>curl -u Mikucionis:***** "http://twitter.com/users/show/Mikucionis.xml"
104% Total    104% Received    104% Xferd    Average Speed   Time    Time     Time  Current
           Dload    Upload   Total   Spent    Left   Speed
104 1151 104 1151    0      0  1502    0 --:--:-- --:--:-- --:--:-- 5255
C:\temp>
```

22 pav. "curl"

„XML::Simple“ – programavimo kalbos „PERL“ biblioteka, supaprastinanti „XML“ duomenų struktūrų naudojimą.

3.2.1. Duomenų išrinkimo realizavimas

Programavimo kalba „PERL“ stipriai „surišta“ su operacine sistema. Kurtoje programoje, daugelis komandų „konstruojamos“ „PERL“ aplinkoje, tačiau vykdyti atiduodamos operacinei sistemai.

Portalo „Twitter“ duomenys prieinami tik sistemos naudotojams, todėl prieš sistemos analizę, reikėjo tapti tinklo nariu. Mane atspindintis profilis buvo naudojamas kaip atskaitinis sistemos taškas.

„HTTP“ protokolas pats savaime nepalaiko sesijos tarp atskirų užklausų („HTTP“ protokolu veikia „cURL“), „cURL“ taip pat, todėl naudotojo vardas bei slaptažodis programoje išsaugojami kintamuosiuose (naudotojo vardą bei slaptažodį reikia pateikti su kiekviena „cURL“ užklausa, o jų programa turi labai daug).

Kintamasis „searchUser“ šioje įrankio versijoje nenaudojamas, bet yra įtrauktas dėl galimybės išplėsti programos funkcionalumą (jeigu sistemos atskaitos tašku pasirinktumėme kitą profilį, tačiau vis tiek norėtume prisijungti savo vardu).

Kintamasis „mode“ skirtas norimos ataskaitos raktiniam vardui.

„@ARGV“ komandinės eilutės argumentai (angl. „command line arguments“) pateikiami programos kvietimo metu.

```
use strict;
use XML::Simple;
die "specify searchUser, username, password, mode " unless @ARGV == 4;
my($search, $user, $pass, $mode) = @ARGV;
```

„\$cmd“ – tekstinis kintamasis, kuriame konstruojamos operacinei sistemai skirtos komandos.

Norint išryškinti socialinių tinklų analizės proceso etapus, parsiusi duomenis išsaugojami kietajame diske – kataloguose „xml“ bei „img“ (paprastesniame modelyje informacija būtų atiduota kintamajame).

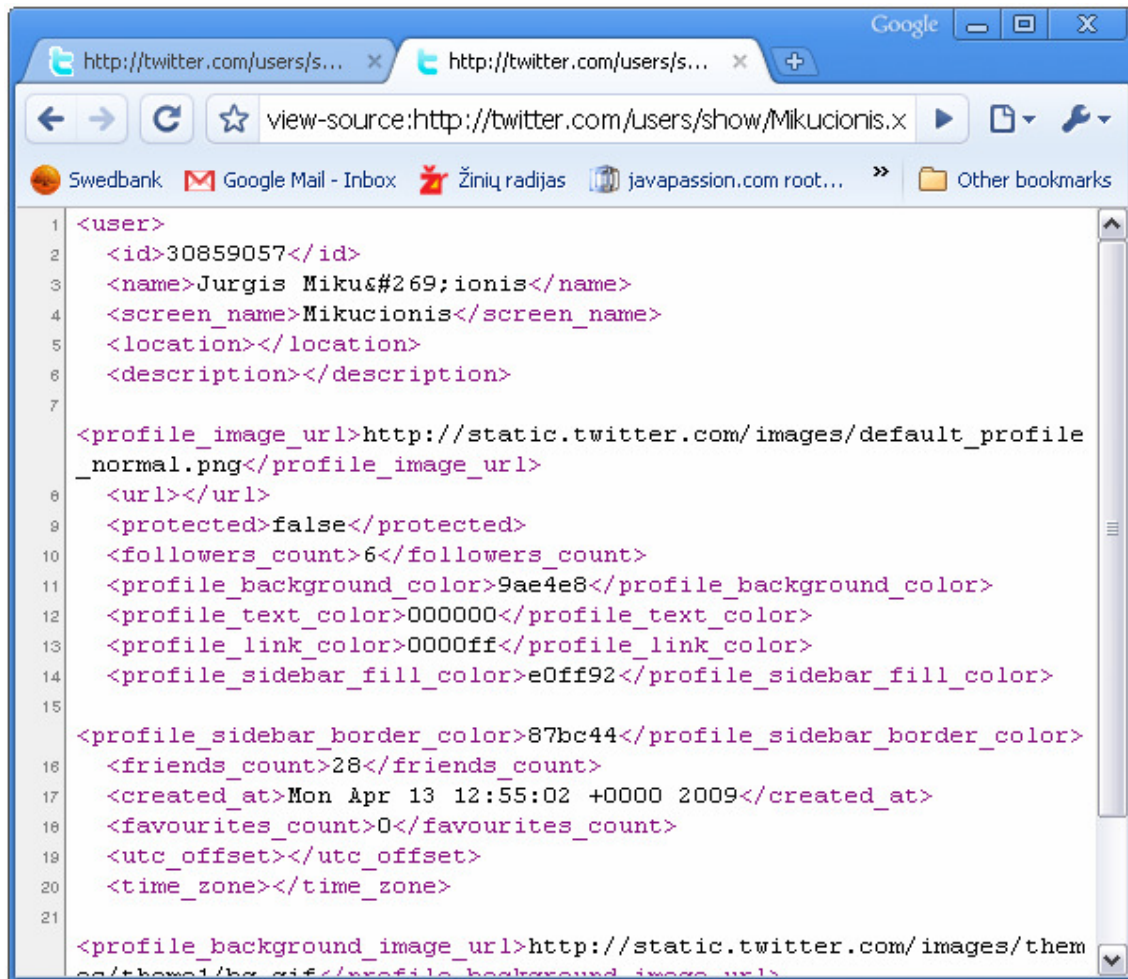
Prieš kurdama katalogus ar siųsdama reikiamą informaciją, programa visada patikrina ar jos neturi.

```
# Sukuriame programai reikalingus katalogus
my $cmd = "mkdir xml img";
system($cmd) unless(-d "xml" && -d "img");
```

Praeitame skyriuje minėta, kad „Twitter“ per specialią prieigą duomenis apie savo naudotojus pateikia „XML“ formatu (23 paveikslėlis).

```
# Issaugojame vartotojo duomenis
```

```
$cmd = qq{ curl -u $user:$pass "http://twitter.com/users/show/$user.xml" > xml/$user.xml};  
system($cmd) unless(-e "xml/$user.xml");
```



23 pav. Užklauso rezultatas naršyklėje

Avatarų portalas nesaugo savo duomenų bazėje, tačiau pateikia jų „URL“ adresus viename iš „XML“ kalba parsiųstų naudotojo atributų.

```
# Issaugojame vartotojo avatara
```

```
my $xmlImg = XMLin( "xml/$user.xml" );  
my $imgUrl = $xmlImg->{profile_image_url};  
$cmd = qq{ curl "$imgUrl" > img/$user.png};  
system($cmd) unless(-e "img/$user.png");
```

Naudotojo draugų sąrašui gauti naudojama kita užklausa. Programoms skirta „Twitter“ sąsaja informaciją apie visus naudotojus atiduoda bloke „user“, todėl naudotojo „draugai“ išsaugomi

kitoje byloje (viską išsaugojus vienoje byloje, atskirti kur informacija apie patį naudotoją, o kur informacija apie jo draugus, būtų neįmanoma).

```
# Issaugome vartotojo draugu sarasa
$cmd = qq{curl -u $user:$pass "http://twitter.com/statuses/friends.xml" >
xml/$user.friends.xml};
system($cmd) unless(-e "xml/$user.friends.xml");
```

Cikle analogiškai išrenkama informacija apie naudotojo „draugus“, jų „draugus“.

```
# Issaugome informacija apie visus vartotojo draugus
my $xmlFriend = XMLin("xml/$user.friends.xml");
for my $name(keys %{$xmlFriend->{user}})
{
    my $userFr = $xmlFriend->{user}->{$name}->{screen_name};
    # Issaugojame vartotojo draugo draugu sarasa
    $cmd = qq{curl -u $user:$pass "http://twitter.com/statuses/friends/$userFr.xml?page=1" >
xml/$userFr.friends.xml};
    system($cmd) unless(-e "xml/$userFr.friends.xml");
}
```

Tame pačiame cikle išsaugoma po 200 naujausių kiekvieno „draugo“ pasisakymų.

```
# Issaugojame vartotojo draugo 200 naujausių pasisakymų

$cmd = qq{curl -u $user:$pass
"http://twitter.com/statuses/user_timeline/$userFr.xml?count=200" >
xml/$userFr.user_timeline.xml};

system($cmd) unless(-e "xml/$userFr.user_timeline.xml");
```

Taip pat visų draugų avatarai.

```
# Issaugome vartotojo draugo avatara
my $imgUrl = $xmlFriend->{user}->{$name}->{profile_image_url};
$cmd = qq{curl "$imgUrl" > img/$userFr.png};
system($cmd) unless(-e "img/$userFr.png");
} # Cikla kartojame kiekvienam draugui
```

Šios programos dalies rezultatas – 2 nauji katalogai. Viena (pavadintame „xml“) išsaugota informacija apie naudotoją, jo „draugus“, jo „draugų draugus“, kita (pavadintame „img“) - „PNG“²⁸ formatu išsaugoti avatarai

²⁸ PNG (angl. Portable Network Graphics) – bitų masyvo formatas, kuris suglaudinas „be nuostolių“. Sukurtas kaip GIF formato pakaitalas – formatas nereikalaujantis patento licencijos.

3.3. Duomenų analizė

Jau minėta, kad programa sugeba pateikti 3 tipų ataskaitas:

- galima naudotojo įtaka savo „draugams“;
- naudotojo draugų rato pokalbių temos;
- naudotojo draugų išsidėstymas pasaulyje.

Šiame skyrelyje apžvelgta reali siūlomų ataskaitų reikšmė, skaičiavimo algoritmai. Visi skaičiavimai atlikti programavimo kalba „PERL“.

Naudoti įrankiai:

- „Geo::Coder::Yahoo“ („PERL“ programinis modulis);
- „XML::Simple“ („PERL“ programinis modulis).

Jau minėta „XML::Simple“ - programavimo kalbos „PERL“ biblioteka, supaprastinanti „XML“ duomenų struktūrų naudojimą.

„Geo::Coder::Yahoo“ - tai „PERL“ modulis, apibrėžiantis supaprastintą sąsają tarp kuriamos programos ir „Yahoo“ siūlomos internetinės paslaugos „GeoCoder“.

3.3.1. Duomenų analizės realizavimas

Sukurta programa turi tik vieną pradžios tašką. Sugeneruotos ataskaitos tipą apsprendžia naudotojo pateiktas kintamasis programos kvietimo metu (kintamasis „mode“). Parsiuntus visus duomenis, programos vykdymas atiduodamas vienam iš paprogramių.

```
# Programos vykdymas atiduodamas vienam is paprogramiu
visualizeInfluence() if($mode eq "influence");
visualizeKeywords() if($mode eq "keywords");
visualizeLocations() if($mode eq "locations");
```

Sociologai teigia, kad kiekvienas socialinės grupės narys daro įtaką aplink jį esantiems individams. Naudotojo daroma įtaka yra atvirkščiai proporcinga jo įtakojamo individo jungčių skaičiui. Todėl, norint paskaičiuoti (arba atvaizduoti) galimą naudotojo įtaką jo „draugams“, reikia suskaičiuoti, kiek kiekvienas naudotojo „draugas“ turi savo „draugų“. Kaip tik tai ir daro vienas iš kurtos programos paprogramių.

Įrankis realizuotas labai paprastai. Naudotojo draugų sąrašas išsaugojamas maišos lentelėje²⁹, kuri atiduodama specialiam ciklui, pritaikančiam eilę komandų kiekvienam maišos lentelės elementui.

²⁹ Maišos lentelė (angl. hash table) – duomenų struktūros, kuriose duomenys yra saugomi, priskiriant jiems unikalų raktą.

```

sub visualizeInfluence
{
    # Suskaiciuojame kiek kiekvienas vartotojo draugas turi draugu
    my %frHash = ();
    my $xmlFriend = XMLin("xml/$user.friends.xml");
    for my $name(keys %{$xmlFriend->{user}})
    {

```

Tokiu pačiu principu apskaičiuojamas kiekvieno „draugo“ „draugų“ skaičius. Rezultatas išsaugojamas maišos lentelėje, kurios raktai –naudotojo „draugų“ vardai („Twitter“ sistemoje visi naudotojų vardai unikalūs).

```

    my $userFr = $xmlFriend->{user}->{$name}->{screen_name};
    my $xmlSec = XMLin("xml/$userFr.friends.xml");
    $frHash{$userFr} = 0;
    for my $linkUser(keys %{$xmlSec->{user}}) {$frHash{$userFr}++}
}

```

Antras programos įrankis sąlyginai parodo apie ką kalba pasirinkto naudotojo draugų ratas. Skaičiavimai paremti tų pačių žodžių pasikartojimu sistemos naudotojo „draugų“ žinutėse.

Algoritmo pradžioje suskaiciuojama, kiek kartų buvo paminėtas kiekvienas žodis. „Ciklo cikle“ principu suskaiciuojami visi kiekvieno „draugo“ kiekvienoje žinutėje pavartoti žodžiai.

```

sub visualizeKeywords
{
    my %wordHash = ();
    my $xmlFriend = XMLin("xml/$user.friends.xml");
    # Po viena suskaiciuojame visu vartotojo draugu panaudotus zodzius
    for my $name (keys %{$xmlFriend->{user}})
    {
        my $userFr = $xmlFriend->{user}->{$name}->{screen_name};
        my $xmlSec = XMLin("xml/$userFr.user_timeline.xml");
        # Kiekvienas vartotojo draugas turi iki 200 zinuciu
        for my $linkUser(keys %{$xmlSec->{status}})
        {
            my $msgText = $xmlSec->{status}->{$linkUser}->{text};
            for my $key(split " ", lc($msgText)){$wordHash{$key}++}
        }
    }
}

```

Maišos lentelėje išsaugoti vartoti žodžiai surūšiuojami pagal jų pavartojimo skaičių ir išrenkami populiariausi (išrenkami tik iš 7 ar daugiau simbolių sudaryti žodžiai).

```
my $tStr = "";
my $chlStr = "";
# Gauta zodziu ir panaudojimo kartu lentele surusiuojame pagal kartus
for my $word (sort{ $wordHash{ $b } <=> $wordHash{ $a } }keys %wordHash)
{
# I galutini grafika itraukiame tik ilgus zodzius, t.y ilgesnius nei 7 simboliai
next unless(length($word) > 7);
$tStr .= "$wordHash{ $word },";
$chlStr .= "$wordl";
# Iseiname is ciklo kai turime 10 populiariausiu zodziu
last if(($tStr =~ s/,/,/g) == 10);
}
# Istrinam paskutinius kablelius
chop($tStr);
chop($chlStr);
```

Trečia programos ataskaita parodo naudotojo draugų išsidėstymą pasaulyje (geografiškai). Portalas „Twitter“ informacijai apie savo naudotojų geografinę padėtį yra išskyręs tik vieną lauką. Tai labai apsunkina įrankio darbą, nes į vienintelį lauką kiekvienas naudotojas gali įvesti ką tik nori, koku tik nori formatu (portalas net nesiūlo įvesti valstybės ar valstijos).

Valstybės nustatymui programa naudoja jau minėtą „Yahoo“ siūlomą įrankį – „Geocoder“ (įrankis naudojamas per specialią „PERL“ kalbai parašytą sąsają). „Geocoder“ visada grąžina 1 rezultatą (tiksliausią įmanomą).

Valstybės nustatymui programa naudoja jau minėtą „Yahoo“ siūlomą įrankį – „Geocoder“ (įrankis naudojamas per specialią „PERL“ kalbai parašytą sąsają). „Geocoder“ visada grąžina 1 rezultatą (tiksliausią įmanomą).

sub visualizeLocations

```
{
    # Sis modulis pagal adreso dali nustato valstybe, kuriai tas adresas priklauso
    use Geo::Coder::Yahoo;
    my $geocoder = Geo::Coder::Yahoo->new(appid => 'my_app');
    my %locData = ();
    my $total;
    my $xmlFriend = XMLin("xml/$user.friends.xml");
    # Po viena issiaiskiname visu vartotojo draugu valstybes
    for my $name (keys %{$xmlFriend->{user}})
    {
        my $userLoc = $xmlFriend->{user}->{$name}->{location};
        my $location = $geocoder->geocode(location => "$userLoc");
```

Interneto paslaugos atsakymą sudaro „XML“ kalba aprašytas objektas turintis 9 atributus.

Programai reikalingas tik vienas iš jų – valstybė (pateikiama pagal „ISO 3166“³⁰ standartą).

```
# Is grazintos duomenu lenteles issitraukiame tik viena reikalinga
```

```
elementa - valstybe
```

```
for my $param(@{$location})
```

```
{
```

```
    my %hashRef = %{$param};
```

```
    # Visas valstybes susirasome i lentele
```

```
    # Suskaiciuojame kiek tose valstybese vartotojas turi
```

```
draugu
```

```
    $locData{$hashRef{country}}++;
```

```
    $total++;
```

```
}
```

```
}
```

```
my $chld;
```

```
my $chd;
```

Visų trijų įrankių galutinis analizės rezultatas pateikiamas maišos lentelėse.

³⁰ ISO 3166 - trijų dalių geografinis kodavimo standartas koduojantis valstybių ir priklausomų teritorijų pavadinimus.

Pagal sudarytą 3 etapų socialinių tinklų tyrimo proceso modelį, informacijos mainams tarp atskirų etapų turėtų būti naudojama „XML“ kalba. Ši praktika programoje nepritaikyta, nes darbo tikslas nėra sukurti naudojimui paruoštą įrankį, kurio funkcionalumą vėliau ketinčiau išplėsti.

3.4. Duomenų vaizdavimas

Jau minėta, kad paskutinis socialinių tinklų tyrimo proceso etapas – duomenų vaizdavimas, susideda iš dviejų dalių.

Šiame skyrelyje apžvelgta, kaip realizuotas duomenų paruošimas bei vaizdavimas sukurtoje programoje.

Naudoti įrankiai:

- „Graphviz“;
- „Google Charts“.

Abu naudoti įrankiai jau aptarti praeitame skyriuje. „Google Charts“ ir „Graphviz“ įvairių duomenų atvaizdavimui skirtos programos.

3.4.1. Duomenų vaizdavimo realizavimas

Ataskaitos „Galima naudotojo įtaka savo draugams“ vaizdavimui pasirinkta programa „Graphviz“. Praitame skyriuje minėta, kad vaizduojamą informaciją programai galima pateikti kodu, kuriame aprašomi vaizduojami elementai ir jų tarpusavio ryšiai (elementus programa išdėsto automatiškai).

Duomenų analizės pateiktas rezultatas – maišos lentelė, kurios raktai naudotojo „draugų“ vardai, o reikšmės –naudotojo „draugų“ „draugų“ kiekiai.

Pirmas algoritmo etapas, surūšiuoti lentelę pagal galimą naudotojo įtaką („draugų“ „draugų“ skaičių). Tai daroma, nes programoje įgyvendinta galimybė riboti vaizduojamų „draugų“ skaičių (jeigu jų labai daug). Nerūšiuojant lentelės, vaizdavimui būtų atiduodami absoliučiai atsitiktiniai „draugai“, norimas rezultatas – „draugai“, kuriems naudotojas daro didžiausia įtaką.

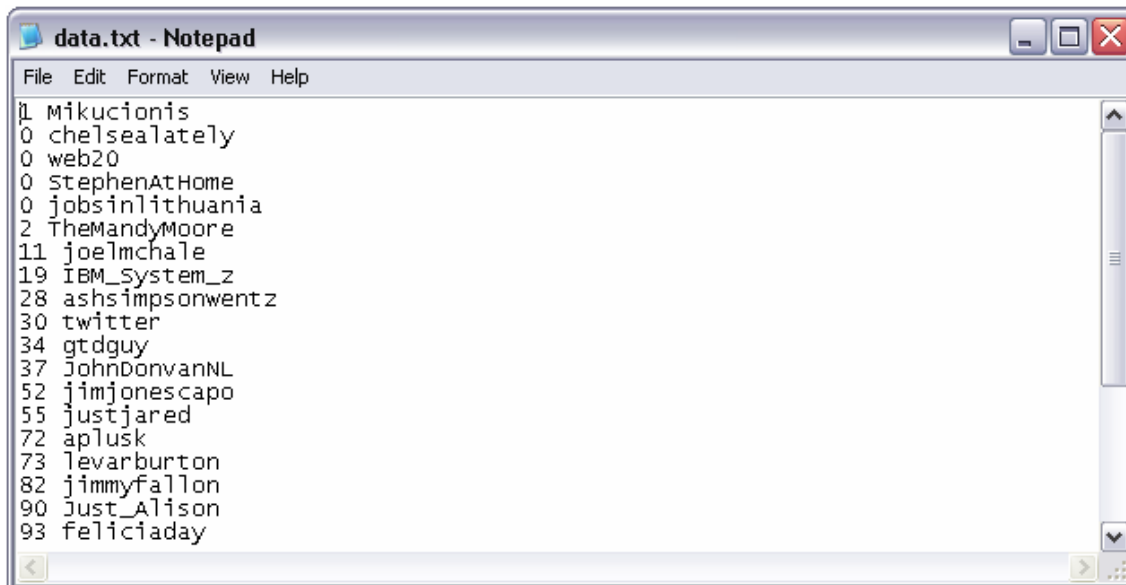
```
# Suformuojame duomenų srautą, kuri atiduosime paprogramiui
my $infList = "1 $user\n";
for my $name(sort{$frHash{$a} <=> $frHash{$b}} keys %frHash)
{
    $infList .= "$frHash{$name} $name\n";
    # Atkomentuokite eilutę žemiau jei norite riboti atvaizduotų draugų skaičių
    # last if(($infList =~ s/\n\n/g) == 10);
}
# Istringiname paskutini kableli
chop($infList);
```

Kodą įrankiui „Graphviz“ generuoja sukurtos programos paprogramė. Vienintelis būdas per operacinę sistemą iškviesti paprogramei atiduoti iš kelių eilučių sudarytą maišos lentelę, yra pateikti ją per iš anksto byloje išsaugotą duomenų srautą (24 paveikslėlis). Naudota operacinė sistema „MS Windows XP“ neturi komandos sugebančios pateikti kintamojo sudaryto iš kelių eilučių (taip kintamasis galėtų būti perduotas kaip komandinės eilutės parametras). Komanda „ECHO“ išveda visus duomenis iki simbolio „nauja eilutė“ („Enter“). Duomenis paprogramei perdaviau komandos „TYPE“ pagalba. „TYPE“ nuskaito visą nurodytos bylos turinį. Iš bylos nuskaitytą duomenų srautą paprogramei atdaviau per vieną iš numatytųjų (angl. „default“) kintamųjų - „STDIN“ (operacinė sistema kiekvienam savo procesui sukuria 4 numatytuosius srautus).

```
# Issaugome duomenų srautą byloje
open MYDATA, ">data.txt";
print MYDATA $infList;
close MYDATA;

# Paruosiamė duomenis

$cmd = qq{TYPE data.txt | Perl.exe twitdot.pl $user img > ItakaDuom.fdp};
system($cmd);
```



24 pav. Paprogramei paruoštas duomenų srautas

Bylos, į kurią išvedamas paruoštas kodas, vardas nurodomas pagrindinėje programoje. Be duomenų srauto, paprogramei kvietimo metu atiduodami dar du parametrai: centrinis mazgas, kurio įtaka paskaičiuota, bei avatarų katalogo vardas.

```

use strict;
die "usage: buildDot.pl <home_node> <thumbnail_dir>" unless @ARGV == 2;
my( $homeNode, $thumbsDir ) = @ARGV;
# GraphViz duomenų struktūros elementas
print "graph g {\n";
# Iprasta STDIN pakeiciame kviesdami si paprogrami vartotojo draugu, bei ju draugu skaiciu
lentele

```

Kiekvienai duomenų srauto eilutei įvykdoma viena ciklo iteracija. Kiekvieną srauto eilutę sudaro naudotojo „draugo“ vardas bei naudotojo „draugo“ „draugu“ skaičius. Pagal naudotojo „draugo“ „draugu“ skaičių apskaičiuojamas juos jungiančios rodyklės storis bei spalva (didelė įtaka išreiškiama storomis ir šviesiomis rodyklėmis).

```

while(<STDIN>)
{
    # Atskiriam vartotojo draugo varda nuo jo turimu jungciu skaiciaus
    my($lineWidth,$nodeName) = split;

    # Net jei vartotojo draugas neturi draugu jungciai vistiek priskiriamas minimalis svoris
    # Kitiems vartotojo draugams, draugu skaicius priskiriamas kaip jungties storis
    $lineWidth = 1 if($lineWidth == 0);

    # Suformuojamos vartotojo draugu sritys
    print qq(subgraph "cluster_$nodeName" );
    print qq({label="$nodeName"; labelloc="t"; "${nodeName}_icon"};\n);
    print qq("${nodeName}_icon" [label="",shape=box,style=invis,);
    print qq(shapefile="$thumbsDir/$nodeName.png"];\n);

    # Pagal vartotojo draugu draugu skaiciu priskiriame atitinkama jungties spalva
    my $lineColor = "#000000";

    # Kuo daugiau draugu, tuo sviesesne jungtis
    if($lineWidth > 20){ $lineColor = "#ff9900" }
    elsif($lineWidth > 15){ $lineColor = "#ca7836" }
    elsif($lineWidth > 10){ $lineColor = "#cabb36" }
    elsif($lineWidth > 5){ $lineColor = "#93ca36" }

```

```

# Kadangi storis matuojamas milimetrais, sumaziname visus svorius 10 kartu
$lineWidth = $lineWidth / 10;

# Nurodome rodykles tipa
print qq(edge[color="$LineColor",arrowtail="normal",arrowsize="3",]);
print qq(style="setlinewidth($lineWidth)" );\n);

# Nutraukiame iteracija, jei jungtis turi jungti vartotoja su juo paciu
if($nodeName =~ /$homeNode/){print qq(\n); next}

# Aprasome kas su kuo bus sujungta
# Musu atveju, vartotojas su vienu is savo draugu
print qq("cluster_$nodeName" -- "cluster_$homeNode";\n\n);
}
# GraphViz duomenų struktūros elementas
print qq(}\n);

```

Galutinis paprogramės rezultatas - kodu aprašyta, vaizdavimui paruošta duomenų struktūra išsaugota byloje „ItakaDuom.fdp“ (25 paveikslėlis).

```

graph TD
    subgraph cluster_Mikucionis
        Mikucionis[Mikucionis]
    end
    subgraph cluster_chelsealately
        chelsealately[chelsealately]
    end
    subgraph cluster_web20
        web20[web20]
    end
    subgraph cluster_StephenAtHome
        StephenAtHome[StephenAtHome]
    end
    subgraph cluster_jobsinlithuania
        jobsinlithuania[jobsinlithuania]
    end
    subgraph cluster_TheMandyMoore
        TheMandyMoore[TheMandyMoore]
    end
    subgraph cluster_joelmchale
        joelmchale[joelmchale]
    end

    Mikucionis --> chelsealately
    Mikucionis --> web20
    Mikucionis --> StephenAtHome
    Mikucionis --> jobsinlithuania
    Mikucionis --> TheMandyMoore
    Mikucionis --> joelmchale

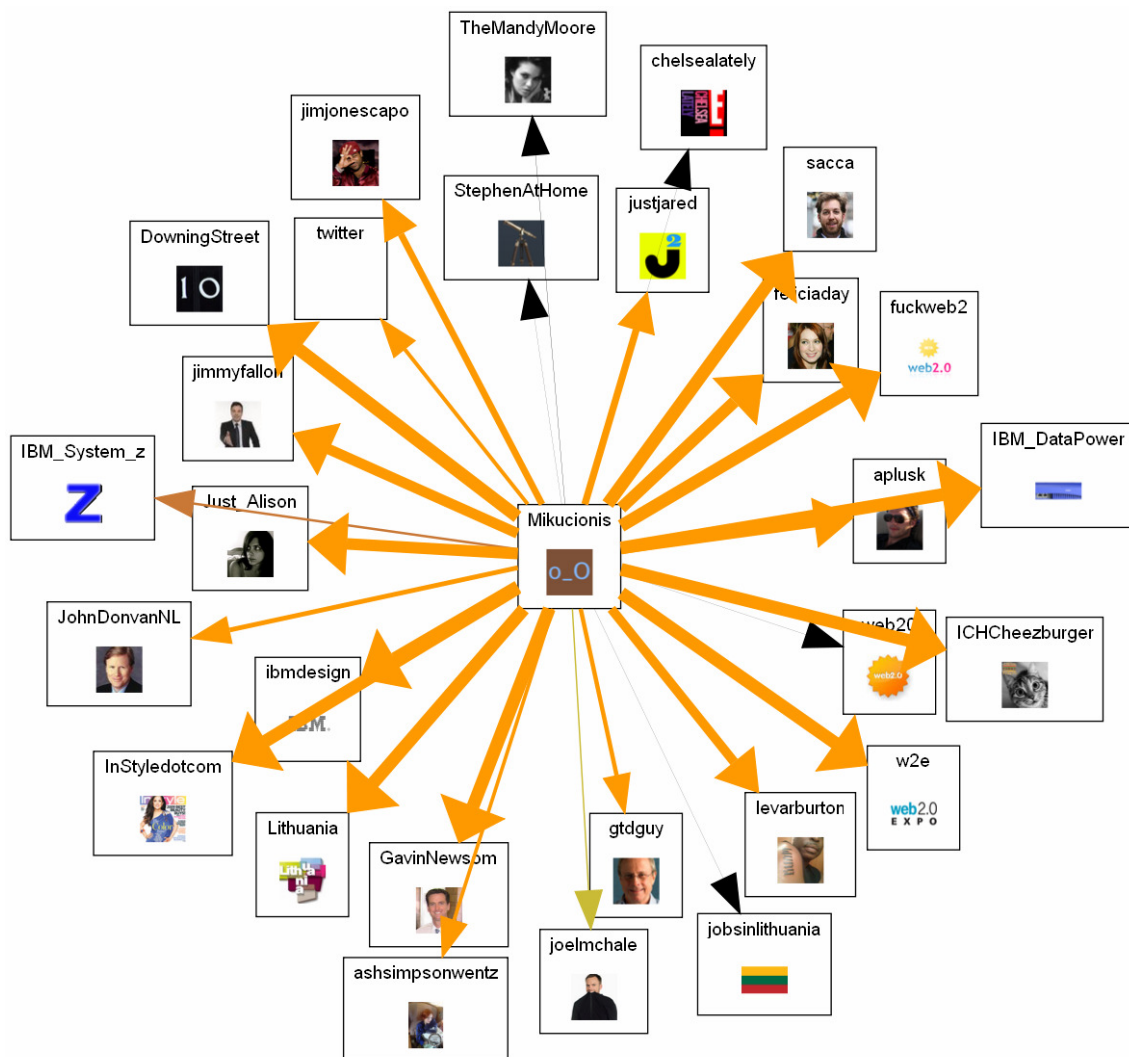
```

25 pav. „ItakaDuom.fdp“

Paskutinis proceso etapas – duomenų atidavimas „Graphviz“ programa.

```
# Atiduodame duomenis grafiniam apdorojimui
# Naudojama programa Graphviz
$cmd = qq{fdp ItakaDuom.fdp -Tpng -o Itaka.png};
system($cmd);
}
```

Ataskaitos rezultatas išsaugomas byloje „Itaka.png“ (26 paveikslėlis).



26 pav. „Itaka.png“

Kitų dviejų ataskaitų skaičiavimai pateikiami įrankio „Google Charts“ pagalba. Google Charts – tai kompanijos „Google“ siūloma internetinė paslauga skirta nesudėtingiems duomenims atvaizduoti. Duomenys paslaugai pateikiami kaip „URL“ užklauskos parametrai.

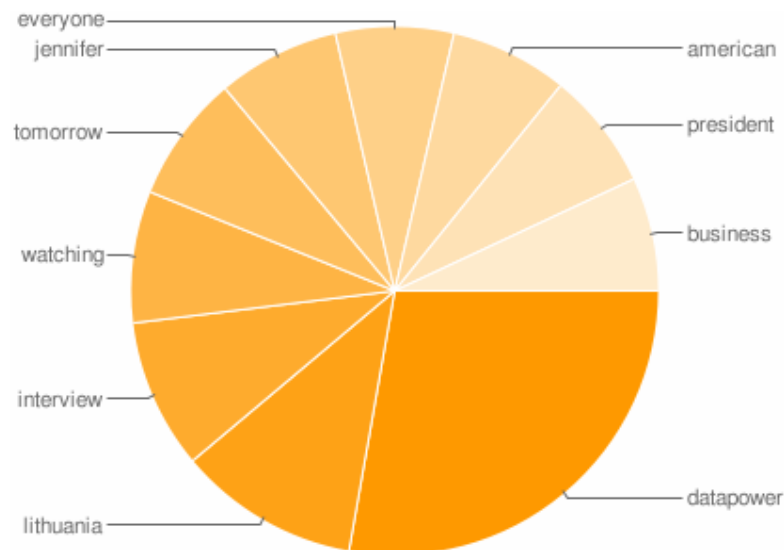
Naudotojo draugų rato pokalbių temoms atvaizduoti pasirinkta skritulinė diagrama. Tokio tipo diagrama reikalauja 4 kintamųjų:

- cht=3 – diagramos tipas;

- chs=1000x300 – diagramos dydis;
- chd=s:\$tStr – duomenų vertės (atskiriamos kableliais)(procentais reikšmių įvertinti nebūtina, paslauga tai padaro automatiškai);
- chl= \$chlStr – diagramos sričių vardai (turi būti pateikti tokia pačia tvarka kaip ir vertės);
- chco – spalvų spektro kintamasis (nebūtinai).

```
# Atiduodame duomenis grafiniam apdorojimui
# Naudojama paslauga Google Chart API, duomenys itraukiami kaip uzklauso
parametrai
$cmd = qq{curl "http://chart.apis.google.com/chart?cht=p
&chd=t:$tStr&chs=1000x300&chl=$chlStr" > Zodziai.png };
system($cmd);
}
```

Ataskaitos rezultatas išsaugomas byloje „Zodziai.png“ (27 paveikslėlis).



27 pav. „Zodziai.png“

Naudotojo draugų išsidėstymui pasaulyje pavaizduoti naudojama ta pati „Google Charts“ paslauga.

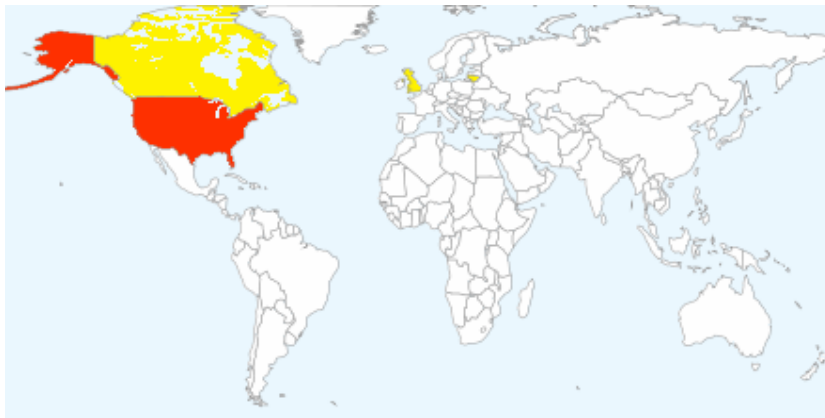
Tokio tipo diagrama reikalauja 7 kintamųjų:

- cht=t – diagramos tipas;
- chtm=world – diagramos potipis;
- chs=440x220 – diagramos dydis;
- chco=FFFFFF,FFFF00,FF0000 – spalvų spektro kintamieji (pirmasis – neįvardintoms kintamajame „chld“ šalims, kiti du – spalvų spektras „chld“ reikšmėms);

- chld=\$chld – „ISO 3166“ standartu išvardinti šalių kodai (kadangi visos šalys pagal šį standartą žymimos 2 raidėmis, kintamieji neišskiriami);
- chd=t:\$chd – duomenų vertės;
- chf=bg,s,EAF7FE" – vandenynų spalvos parametras.

```
# Atiduodame duomenis grafiniam apdorojimui
# Naudojama paslauga Google Chart API, duomenys itraukiami
  kaip uzklauso parametrai
$cmd = qq{curl "http://chart.apis.google.com/chart?cht=t&chtm=world
&chs=440x220&chco=FFFFFF,FFFF00,FF0000&chld=$chld
&chd=t:$chd&chf=bg,s,EAF7FE" > Issidestymas.png};
system($cmd);
}
```

Ataskaitos rezultatas išsaugomas byloje „Issidestymas.png“ (28 paveikslėlis).



28 pav. „Issidestymas.png“

3.5. Skyriaus išvados

- Sukurta programa skirta socialiniam tinklui „Twitter“ analizuoti. Programoje įgyvendinti visi trys socialinių tinklų tyrimo proceso etapai.
- Programa analizuoja tik nedidelę socialinio tinklo dalį – vieną naudotoją ir jo „draugų“ ratą (tokia ataskaitų specifi­ka), tačiau esant reikalui jos akiratį galima nesunkiai išplėsti.
- Mane atspindintis sistemos profilis naudojamas kaip atskaitinis, tačiau programa būtų kur kas naudingesnė jei atskaitos tašku būtų pasirinktas įmonė arba įmonės skyrių atspindintis profilis.

IŠVADOS

- Internetinių socialinių tinklų bei jų realizavimo technologijų apžvalga parodė, kad:
 - populiariausi internetiniai socialiniai tinklai: „Facebook“, „Youtube“, „MySpace“ bei „Twitter“;
 - beveik visi internetiniai socialiniai portalai remiasi „WEB 2.0“ idėjomis, o jų realizavimo technologijos iš esmės yra tokios pačios;
 - socialinių tinklų tyrimo įrankiai turi būti adaptuojami priklausomai nuo konkretaus portalo.
- Darbe buvo suprojektuotas 3-jų etapų internetinių socialinių tinklų tyrimo proceso modelis (duomenų išrinkimas, analizė, pateikimas), kurio siūloma laikytis kuriant internetinių socialinių tinklų analizei skirtus įrankius.
- Išanalizuoti visi pasiūlyto proceso etapai, apžvelgtos jų realizavimo technologijos, kiekviename etape naudoti įrankiai.
- Programavimo kalba „PERL“ įgyvendintas pasiūlytas internetinių socialinių tinklų tyrimo modelis. Sukurta programa analizuoja socialinio portalo „Twitter“ duomenis, sugeba pateikti 3-jų tipų ataskaitas:
 - „galima naudotojo įtaka savo draugams“ parodo naudotojo daromą įtaką aplink jį esantiems individams;
 - „naudotojo draugų rato pokalbių temos“ sąlyginai parodo, apie ką kalba pasirinkto naudotojo draugų ratas;
 - „naudotojo draugų išsidėstymas pasaulyje“ parodo naudotojo draugų išsidėstymą pasaulyje (geografiškai).
- Programa analizuoja tik nedidelę socialinio tinklo dalį – vieną naudotoją ir jo „draugų“ ratą (tokia ataskaitų specifi­ka), tačiau nesunkiai susidorotų ir su didesne struktūra (naudotojo, naudotojo „draugų“, naudotojo „draugų draugų“ ir t.t.).
- Atskaitos tašku, portalo „Twitter“ duomenų sistemoje, buvo pasirinktas mane atspindintis profilis, tačiau, programa būtų kur kas naudingesnė, jei atskaitos tašku būtų pasirinktas įmonė arba įmonės skyrių atspindintis profilis.

LITERATŪRA

1. Amanda Lenhart, „PewResearchCenter“, „Pew Internet & American Life“ projektas. Social Networks Grow: Friending Mom and Dad. [Žiūrėta 2009 m. balandžio mėn. 22 d.]. Prieiga per internetą: <http://pewresearch.org/pubs/1079/social-networks-grow>
2. „Thomson Reuters“. MySpace co-founder DeWolfe to step down. [Žiūrėta 2009 m. balandžio mėn. 22 d.]. Prieiga per internetą: <http://www.reuters.com/article/businessNews/idUSTRE53L6YL20090423>
3. „Balsas.lt“. Socialiniai tinklai plečiasi kosminiu greičiu. [Žiūrėta 2009 m. balandžio mėn. 24 d.]. Prieiga per internetą: <http://www.balsas.lt/naujiena/251291/socialiniai-tinklai-pleciasi-kosminiu-greiciu/rubrika:naujienos-verslas-reklamairziniasklaida>
4. Nicole B. Ellison, Mičigano Universitetas, Telekomunikacijų, Informacijos ir Daugelypių Terpių departamentas. Social Network Sites: Definition, History, and Scholarship. [Žiūrėta 2009 m. balandžio mėn. 24 d.]. Prieiga per internetą: <http://jcmc.indiana.edu/vol13/issue1/boyd.ellison.html>
5. Žydrūnas Sadauskas, Interneto Marketingo Laboratorija. Socialinių tinklapių dabartis ir perspektyvos. [Žiūrėta 2009 m. balandžio mėn. 24 d.]. Prieiga per internetą: <http://ffdesign-impl.blogspot.com/2008/10/socialini-tinklapi-dabartis-ir.html#links>
6. „CodeRS“. Ajax. [Žiūrėta 2009 m. balandžio mėn. 26 d.]. Prieiga per internetą: <http://www.coders.lt/ajax/start>
7. Bert Bos, „W3C“. W3C Core Styles. [Žiūrėta 2009 m. balandžio mėn. 26 d.]. Prieiga per internetą: <http://www.w3.org/StyleSheets/Core/>
8. „W3C“. XHTML2 Working Group Home Page. [Žiūrėta 2009 m. balandžio mėn. 26 d.]. Prieiga per internetą: <http://www.w3.org/MarkUp/>
9. Andrius Blažinskas, Kauno Technologijos Universitetas, Informatikos Fakultetas, Programų Inžinerijos Katedra. REST architektūros panaudojimo paskirstytos sistemos projektavime tyrimas. Magistro darbas, 2008 m., Kaunas

10. S. Dissanaïke, P. Wijkman, M. Wijkman, Stokholmo Universitetas, Karališkasis Technologijų Institutas. Utilizing XML-RPC or SOAP on an Embedded System. [Žiūrėta 2009 m. balandžio mėn. 26 d.]. Prieiga per internetą:
<http://ieeexplore.ieee.org/iel5/9027/28651/01284068.pdf?tp=&isnumber=28651&arnumber=1284068>
11. Aurelijus Valeiša, „Pixel.lt“. Web feeds (RSS, Atom). [Žiūrėta 2009 m. balandžio mėn. 26 d.]. Prieiga per internetą: <http://pixel.lt/web-feeds-rss-atom-2-dalis.html>
12. Booth D., Hass H., McCabe F. Web Services Architecture. [Žiūrėta 2009 m. gegužės mėn. 2 d.]. Prieiga per internetą: <http://www.w3.org/TR/ws-arch/>
13. S. Weerawarana, F. Curbera, F. Leymann, „Prentice Hall“. Web Services Platform Architecture: SOAP, WSDL, WS-Policy, WS-Addressing, WSBPEL, WS-Reliable Messaging, and More. JAV, New Jersey Valstija, 2005 m., p. 27–113.
14. Doug Williams. Twitter API Documentation. [Žiūrėta 2009 m. gegužės mėn. 2 d.]. Prieiga per internetą: <http://apiwiki.twitter.com/Twitter-API-Documentation>
15. „API“. [Žiūrėta 2009 m. gegužės mėn. 2 d.]. Prieiga per internetą:
http://wiki.developers.facebook.com/index.php/API#API_Methods
16. Aurelijus Valeiša, "Pixel.lt". Trumpas XML pradžiamokslis. [Žiūrėta 2009 m. gegužės mėn. 2 d.]. Prieiga per internetą: <http://pixel.lt/xml.html>
17. Xneox.com. Web 2.0 – kas tai?. [Žiūrėta 2009 m. balandžio 25 d.]. Prieiga per internetą:
<http://xneox.com/spauda/web-20--kas-tai.php>
18. Nathan Harrington, „IBM“. Social-networking open source visualization aids. [Žiūrėta 2009 gegužės mėn. 4 d.]. Prieiga per internetą:
<http://www.ibm.com/developerworks/opensource/library/os-socialtools/>

PRIEDAI

A priedas.

INFORMATIKA

11-osios Lietuvos jaunųjų mokslininkų konferencijos „Mokslas – Lietuvos ateitis“, įvykusios Vilniuje 2008 m. balandžio. 09 d., straipsnių rinkinys

VEIKLOS PROCESŲ VALDYMAS

Julija Škatulaitė, Turalas Rasulovas, Jurgis Mikučionis

Vilniaus Gedimino technikos universitetas, juskada@yahoo.com

Anotacija. Nagrinėjama veiklos procesų valdymo (angl. *Business Process Management, BPM*) prasmė, kilmė ir aktualumas mūsų laikais, procesų visumos valdymo galimybės. Aptariami veiklos procesų valdymo komplektai (angl. *Business Process Management Suite, BPMS*), architektūra, naudojami standartai. Apžvelgiami keli BPMS gamintojai, jų siūlomi programiniai sprendimai.

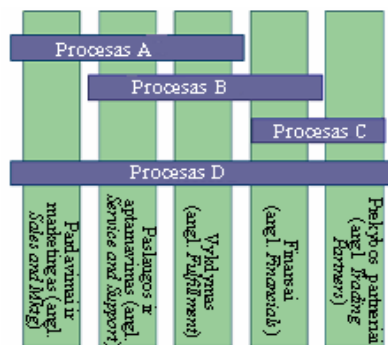
Įvadas

Efektyvus, produktyvus, inovatyvus verslo ar veiklos procesas galimas tik tinkamai naudojant informacines technologijas. XXI amžiuje, norint viską atlikti greitai ir kokybiškai, mažesnėmis sąnaudomis ir didesne apimtimi, žvilgsnis nukreiptas į: klientui aktualių (angl. *on-demand*) poreikių atitikimą; proceso visumos (angl. *end-to-end*) valdymą; galimybę atsižvelgti į geriausią elgseną konkrečiu atveju; inovatyvų atsaką pokyčiams; galimybę darant mažiau daryti daugiau. Šios aktualijos verčia tobulinti *BPM, BPMS* sprendimus. Pats *BPM*, kaip valdymo disciplina, atsirado prieš 20 metų kaip naujas požiūris į veiklą (kaip planuoti, analizuoti, vertinti ne atskiras dalis, bet procesus visame veiklos kontekste) (Silver 2007). Straipsnyje aptarsime *BPM* metodologiją ir įgyvendinimo akcentus.

Problematika

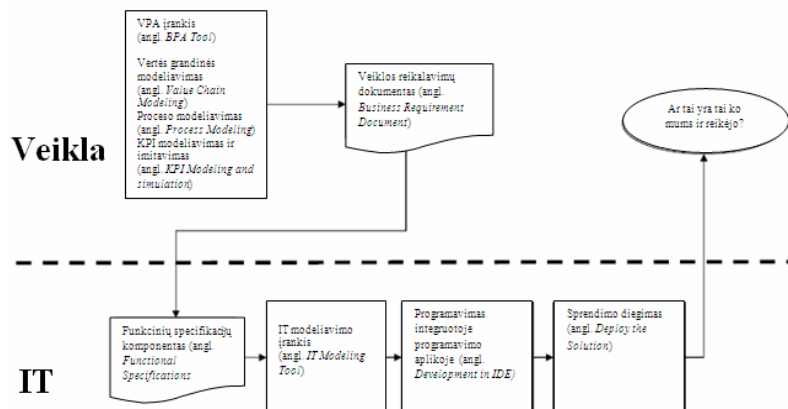
BPM – plati sąvoka, susijusi ir su veiklos modeliavimu, rodiklių vertinimu, ir su technologija, ir su verslo ir IT bendradarbiavimu, t. y. ir veiklos analizė, ir paskirstymas, ir integracija, ir *KPI* (angl. *Key Performance Indicators*), ir *Workflow* automatizavimas, ir *SOA* (angl. *Service-oriented architecture*), ir t. t.

BPM esmė – naujas požiūris į veiklą, kai atskirai analizuojama ne atskiros veiklos, duodančios veiklos rezultata, bet veikiantys keliuose veiklose procesai, kurie leidžia susikonsoliduoti ties procesu nuo jo pradžios iki pabaigos (angl. *end-to-end*). Daug privalumų jau žinoma naudojant tokias *enterprise* aplikacijas, kaip pvz., *ERP, CRM*, kurios individualių funkcijų ar sistemų lygmenyje automatizuoja veiklą, daro ją efektyvesne, suteikia galimybę duomenų integracijai ir kontrolei. Be to, tai ištakos besikuriant *end-to-end* procesų valdymui, kuriam vykstant dėmesys nukreiptas į galutinį tikslą (angl. *mission-critical*) ir procesai „laužo“ ribas tarp atskirų funkcijų (1 pav.).

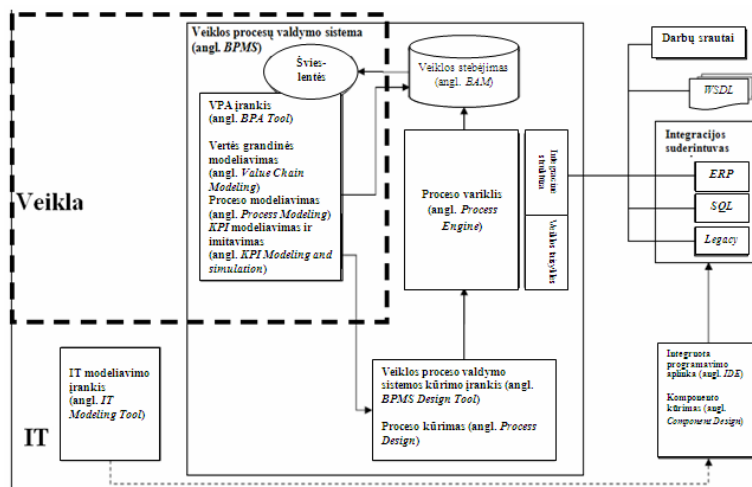


1 pav. Kompanijos ir *enterprise* aplikacijos organizuojamos pagal atskiras veiklos funkcijas, kai dėmesys sutelktas į tikslą, procesai „laužo“ ribas (Silver 2006)

BPM disciplina reikalinga ne tik dėl kitokios veiklos analizės galimybių. Ši naujo požiūrio metodika leidžia, akcentuojant modeliavimą ir analizę, paversti modelį vykdomu, matuojamu ir galimu optimizuoti. Tam tikslui naudojamas veiklos procesų valdymo komplektas – *BPMS*, kuris „laužo“ barjerus tarp atskirų funkcijų ir padidina efektyvumą (automatizuojant), paslankumą pokyčiams, suderinamumą (logika paremta taisyklėmis) ir aiškų proceso visumos, t. y. viso funkcijų derinio, matomumą. Taip pat *BPMS* visą veiklos procesą padaro vientisą, paremtą veiklos specialistų ir IT specialistų bendradarbiavimu (2 pav., 3 pav.).



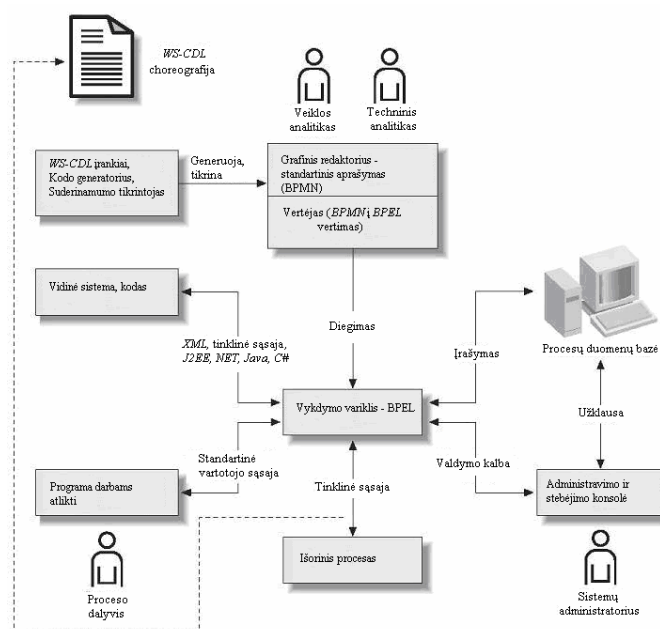
2 pav. Seno požiūrio *BPM* įgyvendinimas (Silver 2006)



3 pav. Naujo požiūrio *BPM* įgyvendinimas – *Clean Handoff* stilius (Silver 2006)

BPM standartai

BPM turi daug žinomų standartų, iš kurių renkamos geriausios kombinacijos. Dažnai veiklos procesų modeliavimo standartų grupės analizė pradedama veiklos procesų modeliavimo notacija *BPMN* (angl. *Business Process Modeling Notation*), kurios pagrindinis siekis yra suteikti ženklų sistemą, kurią lengvai suprastų visi procese dalyvaujantys žmonės: ir verslo analitikai, kuriantys juodraštinis veiklos procesų modelius, ir techniniai darbuotojai, kurie atsakingi už veiklos procesų vykdymą, reikiamų technologijų diegimą, ir galiausiai tie atstovai, kurie valdys ir prižiūrės šiuos procesus. Veiklos procesų diagramų braižymas yra loginė veiklos procesų valdymo įgyvendinimo pradžia. *BPMN* vaizduoja nuoseklų žingsnių ir veiksmų srautą veiklos procese (4 pav.).



4 pav. BPM architektūra (Havey 2005)

BPMN yra susieta su veiklos procesų vykdymo kalba *BPEL4WS* arba tiesiog *BPEL* (angl. *Business Process Execution Language*) ir yra vykdoma *BPEL* variklio. *BPEL*, kaip procesų vykdymo kalba, suformuoja esminį veiklos procesų valdymo sluoksnį. Egzistuoja ir kitos panašaus pobūdžio vykdomos kalbos, pavyzdžiui, *XPDL*, *BPML* ir kt., tačiau šiuo metu labiausiai dėmesys atkreiptas į didžiųjų gamintojų naudojamą *BPEL*, kuri, prireikus, gali būti modifikuojama, priderinant prie kitų kalbų ar perimant kai kurias reikiamas jų savybes.

Tinklinių paslaugų choreografija (angl. *choreography*) paremta *W3C* (angl. *World Wide Web Consortium*) konsorciumo *WS-CDL* (angl. *Web Services Choreography Description Language*) kalba. Choreografija apibūdina procesus iš globalaus požiūrio taško. Choreografijos globalus požiūris yra kontrastas *BPEL* orkestravimui (angl. *orchestration*). *BPEL* procesas yra vieno dalyvio procesas, o choreografija yra dalyvių grupės sąveikos modelis.

WSDL (angl. *Web Services Description Language*) arba tinklinių sąsajų apibūdinimo kalbą pristatė *Microsoft* kartu su *IBM* 2000-ųjų metų spalį. Ji buvo paremta *IBM* tinklu pasiekiamų paslaugų aprašymo kalba ir *Microsoft* paprastų objektų pasiekimo protokolų (*SOAP*) sutartine kalba. *WSDL* apibūdina verslo funkcijas nuo diegimo nepriklausomu būdu, kuris leidžia kitoms taikomosioms programoms jas surasti ir panaudoti. Taigi, *WSDL* yra kalba, pagal kurią nusakomos tinklinės sąsajos ir aprašomas jų naudojimas. Programa, naudojanti tinklines sąsajas, naudoja *WSDL* tam, kad surastų paslaugų vietą ir taip būtų galima prisijungti prie jų.

SOAP pristatė *Microsoft* korporacija 1999-aisiais metais, dalyvaujant *IBM* tam, kad būtų praplėstas pasiūlymo mastas. 2000-ųjų metų kovą abi kompanijos pateikė 1.1 versiją konsorciui – *W3C*. Šis jungtinis pateikimas labai padidino pasitikėjimą pramonės šakose, nes suteikė galimybių dirbti kartu su dviem šių organizacijų labiausiai paplitusiais kompiuteriniais modeliais, tokiais kaip *J2EE* ir *NET*.

BPEL specifikacija yra laikoma jau esamų tinklinių sąsajų standartų veiklos procesų išplėtimu. Anksčiau tinklinės sąsajos buvo apribotos iki neturinčių būklės sąveikų. *BPEL*, kiti procesai ir choreografijos kalbos rodo, kaip sukurti turinčius būklę, sąveikaujančius veiklos procesus tinklinėms sąsajoms. *BPEL* yra griežta kalba, kuri sukuria ir išplečia web paslaugas sąveikaujantiems procesams.

BPMN ryšys su *BPEL* – *BPMN* pateikia mechanizmą, kaip iš veiklos lygio notacijos sukurti vykdomą veiklos procesą. *BPMN* diagramos nebūtų tiek naudingos, jei jų nebūtų galima vykdyti. *BPEL* sąsaja su *BPMN* leidžia iš *BPMN* diagramų gauti *BPEL* kodą, o tai savo ruožtu suteikia galimybę vykdyti *BPMN* *BPEL* sistemose.

Gamintojai ir *BPMS* sprendimai

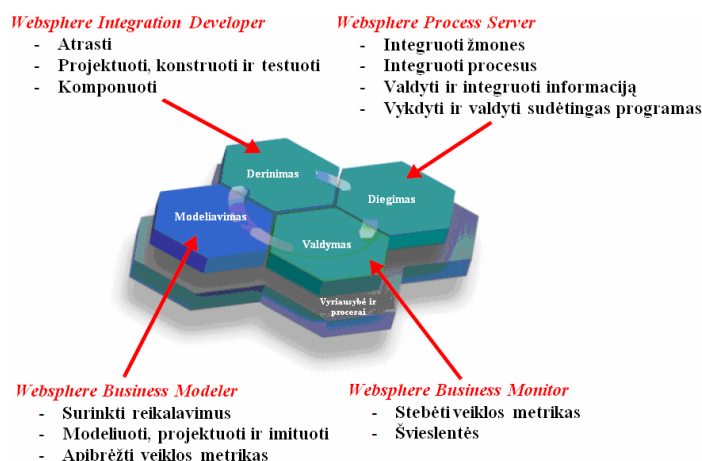
Rinkoje veiklos procesų valdymo sistemos jau siūlomos ne vieno gamintojo. Dauguma programinės įrangos kūrėjų savo *BPMS* sprendimus pateikia sudarytus iš kelių, jau praktiškai standartiniais tapusių komponentų.

Pagrindiniai veiklos procesų valdymo komponentai:

- *BPEL* redaktorius, arba programa, leidžianti grafiškai atvaizduoti įmonėje vykstančius procesus, jų tarpusavio sąveiką ir ryšius;
- veiklos logikos variklis – dažniausiai serveris, kuriame saugomi ir, reikalui esant, pritaikomi tam tikri sprendimai, išreikšti kaip taisyklių rinkiniai. Tokia struktūra leidžia sprendimus vykdyti automatizuotai, t. y. daug greičiau ir pigiau;
- veiklos proceso stebėjimo įrankiai. Vadovai ir sprendimų priėmėjai nuolat stebėdami vykstančius procesus gali kur kas greičiau identifikuoti bei eliminuoti iškilusias problemas;
- žmogiškųjų ir sisteminių resursų integratoriai į sistemą, palaikantys visus resursus sistemoje, taip pat užtikrinantys ryšį tarp atskirų komponentų.

Apžvelkime kelių tiekėjų konkrečius produktus.

Didelę *BPMS* rinkos dalį užima *IBM* – pirmasis programinės įrangos infrastruktūros tiekėjas, siūlantis nuoseklų ir visą (angl. *end-to-end*) gyvavimo ciklo palaikymą (Silver 2006). *IBM WebSphere Business Suite* – tai 4 atskirų *IBM* produktų rinkinys (5 pav.).

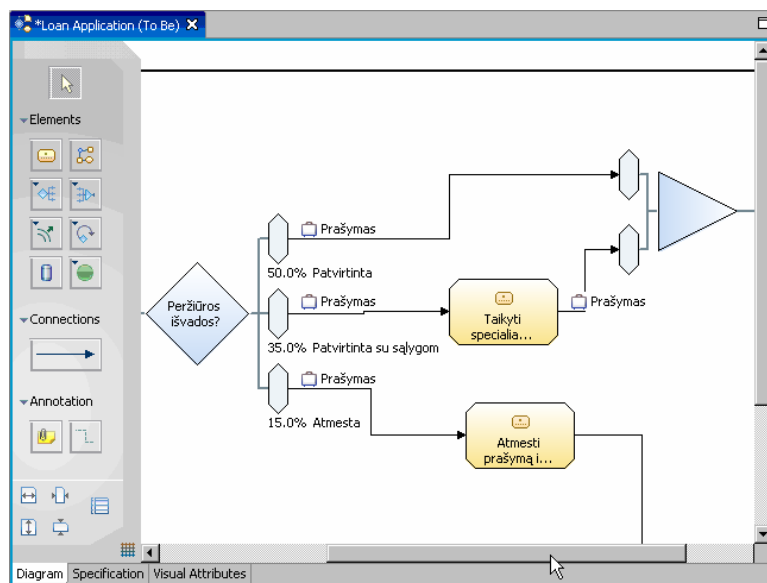


5 pav. *IBM WebSphere Business Suite* komponentai (Wahli, et. al. 2006)

- *WebSphere Business Modeler* – atlieka *BPEL* redaktoriaus vaidmenį, jame simuliuojami procesai;
- *WebSphere Integration Developer* – atlieka integratoriaus vaidmenį;
- *WebSphere Process Server* – veiklos logikos serveris;
- *WebSphere Business Monitor* – veiklos proceso stebėjimo įrankiai.

Šis produktas remiasi *BPEL* ir į paslaugų teikimą orientuota architektūra – *SOA*. Nors *IBM WebSphere Business Suite* parduodamas kaip atskiras produktas, iš tiesų tai sprendimas, kurį sudaro kiti 4 produktai. Palaikomas visas (angl. *end-to-end*) gyvavimo ciklas, komponentai suskirstyti į atskirus etapus. Procesas

pradedamas naudojant *IBM Business Modeler* komponentą, kuriame jau analizės metu (modeliuojant veiklos logiką, simuliuojant procesus) kuriamas vykdomasis modelis (6 pav.)



6 pav. *WebSphere Business Modeler* grafinė vartotojo sąsaja, skirta procesams modeliuoti ir veiklai simuliuoti

Microsoft taip pat siūlo produktų sąrašą, tačiau jie įvardino šį rinkinį kaip sprendimą, o ne kaip atskirą produktą.

- *Windows SharePoint* ir *Office SharePoint Server* – sprendimo integratoriai.
- *BizTalk Server's BRE* - sprendimo veiklos logikos variklis.
- *BizTalk Server's BAM* kartu su *Microsoft Excel* bei *Office PerformancePoint BAM* informacijai atvaizduoti – veiklos proceso stebėjimo įrankis.
- *Microsoft Visio* – sprendimo redaktorius (Chappell 2007).

Oracle BPEL Process Manager – tai *BPEL* standartu paremtas produktas, kurį rinkai siūlo *Oracle*. Šis produktas, be abejo, yra lengvai suderinamas su kitais *Oracle* gaminiais, taip pat *Oracle* teigia, kad jų produktas lengvai suderinamas ir su *Microsoft*, *IBM* bei kitų gamintojų programine įranga (*Oracle...2008*).

Active Endpoints rinkai siūlo produktą *ActiveBPEL*. Šis produktas įdomus tuo, kad jo kodas – atviras visuomenei, o pats produktas nemokamas. Sudarytas jis iš dviejų dalių: serverio ir redaktoriaus, kuris, beje, sukurtas *Eclipse* bazėje (*The ActiveBPEL...2008*). *ActiveBPEL* gerokai kuklesnis už savo giminaitį *ActiveVOX*, už kurį naujasis savininkas jau prašo mokėti.

Išvados

Veiklos procesų valdymas – aktuali šių dienų metodologija, lėmusi naują požiūrį į veiklos procesą, „laužant“ ribas tarp sprendimų, galimų tik atliekant atskiras funkcijas.

BPMS pagrindinis akcentas – vykdomieji modeliai (t. y. iš karto dvigubas modeliavimas – analitinis ir vykdomasis), sukuriama grafiškai ne programuojant kaip struktūrinė schema ir gali būti diegiami į procesą, kad automatizuotų, integruotų ir stebėtų visą (angl. *end-to-end*) procesą.

Kalbant apie *BPM* svarbūs yra standartai. Naudojami jau žinomi standartai, pavyzdžiui, *UML* (angl. *Unified Modeling Language*) buvo sukurtas visai kitam tikslui, bet naudojamas ir *BPM*. Be to, naujesnių technologijų naudojimas, tokių kaip tinklinių sąsajų (angl. *web services*) naudojimas *BPM* technologijoms, lėmė senesnių standartų tobulinimą ir naujų standartų kūrimą. *BPM* yra keli standartai, vieni yra geresni nei kiti. Geriausiai yra išsirinkti stipriausias jų dalis. Patogi adaptacija yra *BPEL* ir *BPMN*.

Visam procesui naudojami *BPMS* programiniai sprendimai turi valdymo organizavimo skirtumų priklausomai nuo gamintojo.

Literatūra

- The ActiveBPEL Community Edition Engine* [interaktyvus]. 2008 [žiūrėta 2008 m. balandžio 4 d.].
Prieiga per internetą: <<http://www.activevos.com/community-open-source.php>>.
- Chappell D. 2007. *Microsoft and BPM: A Technology Overview*. Microsoft Corporation. USA.
- Havey M. 2005. *Essential Business Process Modeling*. O'Reilly. USA.
- Oracle Business Process Management* [interaktyvus]. 2008 [žiūrėta 2008 m. balandžio 3 d.].
Prieiga per internetą: <<http://www.oracle.com/technologies/bpm/index.html>>.
- Silver B. 2007. *New Directions in BPM Tools and Technology*. Bruce Silver Associates. USA.
- Silver B. 2005. *Standards and the Process Lifecycle*. [interaktyvus]. Bruce Silver Associates
[žiūrėta 2008 m. balandžio 5 d.]. Prieiga per internetą:
<<http://www.bpminstitute.org/articles/article/article/standards-and-the-process-lifecycle>>.
- Silver B. 2007. *The BPMS ValueProposition*. BPMInstitute.org. USA.
- Silver B. 2006. *Understanding and Evaluating BPM Suites*. BPMInstitute.org. USA.
- Wahli U., et. al. 2006. *WebSphere Process Integration V6: Business Process Management Modeling through Monitoring*. IBM Press. USA.

BUSINESS PROCESS MANAGEMENT

J.Škatulaitė, T.Rasulovas, J.Mikučionis

Summary

The study deals with Business Process Management (BPM) as a new management discipline which enables management of end-to-end processes that extends across organizational and system boundaries. We have discussed its relevance to today's business process. We have also discussed Business Process Management Suite (BPMS) and used standarts. In addition to this, we have overviewed some BPMS vendors and products.