

VILNIAUS UNIVERSITETAS  
MATEMATIKOS IR INFORMATIKOS FAKULTETAS  
PROGRAMŲ SISTEMŲ KATEDRA

# **Didelės apimties svetainės optimizavimas taikant SEO principus**

**Large-scale website optimization applying SEO principles**

Bakalauro darbas

Atliko: Modestas Kažinauskas (parašas)

Darbo vadovas: prof. dr. Saulius Minkevičius (parašas)

Darbo recenzentas: doc. dr. Valdas Undzėnas (parašas)

Vilnius – 2013

## **SANTRAUKA**

Šio bakalaurinio darbo „Didelės apimties svetainės optimizavimas taikant SEO principus“ autorius yra Modestas Kažinauskas. Dažnai didelės apimties svetainės yra prastai optimizuotos, todėl nesulaukia didelio lankytojų kiekio iš paieškos variklių, nenaudojant mokamos reklamos. Darbo tikslas yra pristatyti pagrindinius SEO principus ir pritaikyti juos optimizuojant didelės apimties muzikos svetainę.

Pirmoje darbo dalyje bus aptartas „Google“ paieškos sistemos veikimo principas. Antrojoje – autorius aptars pagrindinius optimizavimo analizei reikalingus įrankius ir jų naudojimą. Trečiojoje dalyje bus aptarti pagrindiniai vidiniai ir išoriniai svetainės optimizavimo principai, jų specifika ir svarba. Ketvirtojoje dalyje autorius peržvelgs optimizuojamos didelės apimties muzikos svetainės trūkumus ir jų ištaisymą. Penktojoje dalyje autorius ištirs svetainėje rastų optimizavimo klaidų ištaisymo naudą ir pateiks tyrimo rezultatus.

Darbo rezultatas: optimizuota didelės apimties svetainė „Google“ paieškos varikliui ir optimizavimu pasiekti rezultatai.

Raktiniai žodžiai: SEO, optimizavimas „Google“ paieškos varikliui, optimizavimo įrankiai, didelės apimties svetainės optimizavimas, SEO principai.

## **SUMMARY**

The author of this bachelor paper called “Large-scale website optimization applying SEO principles” author is Modestas Kažinauskas. Large-scale websites have low level of optimization in general, to achieve large amount of visitors from search engines, without paid advertisements. The goal of this paper is to present main SEO principles and apply it in optimising large-scale music website.

In the first chapter of this paper the principle of the “Google” search engine will be described. In the second part the author will describe the main tools required for optimization analysis and the use of it. In the third part the main inside and outside optimization principles, the specification and the importance will be described. In the fourth part the author will review the large-scale music website optimization leaks and the possibility to repair it. In the fifth part the author will investigate the benefit of the repaired optimization faults of the website.

The results are optimized large-scale website for “Google” search engine and the results of optimization.

Keywords: SEO, the optimization for “Google” search engine, optimization tools, large-scale website optimization, SEO principles.

# TURINYS

SUMMARY .....	3
ĮVADAS.....	6
1. „GOOGLE“ PAIEŠKOS SISTEMOS VEIKIMAS .....	8
1.1. Paieškos sistemos robotai .....	8
1.2. Svetainės turinio indeksavimas.....	8
1.3. Paieškos vykdymas .....	9
2. SVETAINĖS OPTIMIZAVIMO PAIEŠKOS SISTEMOMS ANALIZAVIMAS.....	10
2.1. Įrankiai .....	10
2.2. Raktažodžių analizė ir parinkimas .....	12
3. SVETAINĖS OPTIMIZAVIMAS „GOOGLE“ PAIEŠKOS SISTEMAI.....	13
3.1. Svetainės vidinis optimizavimas.....	13
3.1.1. Puslapio pavadinimas .....	13
3.1.2. Meta aprašymas ir raktažodžiai .....	14
3.1.3. Svetainės adresų struktūra .....	15
3.1.4. Svetainės žemėlapiai .....	17
3.1.5. „robots.txt“ instrukcijos.....	20
3.1.6. Svetainės klaidų tikrinimas ir šalinimas .....	22
3.2. Svetainės išorinis optimizavimas.....	23
4. MUZIKOS PAIEŠKOS SVETAINĖS OPTIMIZAVIMAS.....	25
4.1. Naudojamos priemonės.....	25
4.2. Optimizavimo principas ir eiga.....	25
4.3. Paieškos užklausų saugojimas .....	26
4.4. Svetainės žemėlapių generavimas iš paieškos užklausų.....	29
4.5. Svetainės žemėlapių pateikimas „Google“ sistemai .....	32
5. MUZIKOS PAIEŠKOS SVETAINĖS OPTIMIZAVIMO REZULTATAI.....	33
5.1. Svetainės puslapių skaičius „Google“ paieškoje .....	33
5.2. Svetainės puslapių matomumo paieškoje ir apsilankymo rodikliai.....	35

IŠVADOS .....	37
SĄVOKŲ APIBRĖŽIMAI IR SANTRUMPŲ SĄRAŠAS.....	39
LITERATŪROS SĄRAŠAS .....	40
PRIEDAI.....	42
1 PRIEDAS. APKLAUSA, KAIP SKIRTINGŲ AMŽIAUS GRUPIŲ ŽMONĖS APLANKO PUSLAPIUS INTERNETE.....	42
2 PRIEDAS. REKLAMOS RŪŠIŲ EFEKTYVUMAS .....	43
3 PRIEDAS. PAIEŠKOS VARIKLIŲ PASISKIRSTYMAS RINKOJE.....	44

## IVADAS

Kiekvienai interneto svetainei yra kritiškai svarbus lankytojų srautas, nes būtent dėl lankytojų jos vienaip ar kitaip generuoja pelną. Vienos svetainės parduodamos prekes ar paslaugas, kitos – rodydamos reklaminius skydelius.

Didžiausias lankytojų srautas vis dar yra gaunamas iš paieškos svetainių, kaip matome iš priede (žr. 1–ą priedą) pateiktos analizės. Todėl, svetainės savininkui norint sulaukti kuo didesnio lankytojų skaičiaus, reikia sutelkti dėmesį būtent kaip gauti kuo daugiau lankytojų iš paieškos svetainių.

Tai padaryti yra žinomi du būdai:

1. reklamuotis paieškos svetainėse (paieškos rezultatų pirkimas pagal paieškos žodžius);
2. optimizuoti svetainę paieškos varikliams.

Iš anksčiau išvardytų būdų pritraukti lankytojams, buvo pasirinktas optimizavimas paieškos varikliams (angl. Search engine optimization). Šis būdas turi didesnę efektyvumą, nei reklamavimasis paieškos svetainėse bei yra plačiau naudojamas (žr. 2–ą priedą). Jis yra skirtas internetinių puslapių lankomumui padidinti, gaunant lankytojų srautą iš paieškos svetainių, tokių kaip „Google“, „Yahoo“ ar „Bing“, optimizuojant esamos svetainės turinį.

2012–2013 metų duomenimis (žr. 3–ą priedą), yra nustatyta, kad populiariausia interneto paieškos sistema pasaulyje yra „Google“. Ja naudojasi 79,74% žmonių, „Bing“ naudojasi 9,94% žmonių, o „Yahoo“ – 7,28% žmonių. Pasaulyje yra dar ir kitų populiarių paieškos sistemų, tačiau jos daugiausiai naudojamos tik specifiniuose regionuose, tokiuose kaip Kinija („Baidu“ paieškos sistema), todėl šiame darbe tokios paieškos sistemos nėra svarbios. Kadangi „Google“ užimama rinkos dalis yra didžiausia, todėl yra svarbu sutelkti dėmesį, kaip užsitikrinti lankytojų srautą, būtent iš šios paieškos svetainės.

Iš pirmos žvilgsnio gali pasirodyti neaišku, kaip galima pritraukti daugiau žmonių į interneto svetainę, optimizavus ją paieškos varikliams. Naudotojui ieškant informacijos paieškos svetainėje, ji grąžina rezultatus, kuriuose yra daug puslapių, kurie tenkina paieškos raktažodžius. Naudotojas geriausiai pastebi ir renkasi rezultatus, kurie yra rodomi aukščiausiose rezultatų puslapių pozicijose. Kurioje pozicijoje atsiduria tinklapis, nustato paieškos sistema, analizuodama jo turinį. Tam ji naudoja sudėtingus algoritmus, kurie yra nuolat tobulinami paieškos variklių kūrėjų. Taigi, pagrindinis svetainės savininko tikslas yra pasiūlyti paieškos svetainėms savo puslapio turinį, kuris būtų kuo labiau atitinkantis paieškos sistemai priimtus turinio standartus. Tai leidžia svetainei atsidurti aukštesnėse paieškos rezultatų pozicijose. Šio

dalyko įgyvendinimui, yra naudojama svetainės optimizacija, kuri ir yra viena iš šio darbo nagrinėjamų temų.

Šio darbo tikslas yra išsiaiškinti ir pristatyti „Google“ paieškos variklio veikimo principus. Taip pat pritaikyti juos ieškant didelės apimties muzikos svetainės optimizacijos trūkumų bei ištaisyti juos, bei ištirti, kokią naudą padarė pakeitimai.

Darbo uždaviniai:

1. susipažinti ir aprašyti „Google“ paieškos sistemos veikimo principus;
2. susipažinti ir aprašyti vidinį ir išorinį svetainės optimizavimą „Google“ paieškos varikliui;
3. surasti muzikos svetainės optimizacijos klaidas ir jas ištaisyti;
4. ištirti ir aprašyti optimizacijos patobulinimo naudą svetainei.

Šiame darbe bus galima susipažinti su populiariausiais „Google“ kompanijos sukurtais svetainių administravimo įrankiais: „Keyword Tool“, „Analytics“ ir „Webmaster Tools“. Bus atliktas praktinis darbas patobulinant esamos muzikos paieškos svetainės optimizaciją paieškos varikliams. Optimizacijos metu, darbai bus atliekami PHP programavimo kalba, panaudojant „CodeIgniter 2“ karkasą.

Darbo reikšmė – aptarti „Google“ paieškos variklio ypatumus, susipažinti su pagrindiniais svetainės optimizavimo principais ir jos teikiama nauda, geriau suprasti didelių svetainių optimizavimo paieškos varikliams metodus.

# **1. „GOOGLE“ PAIEŠKOS SISTEMOS VEIKIMAS**

Kiekvieną dieną pasaulyje yra sukuriama tūkstančiai naujų internetinių puslapių, kuriuose talpinamas didelis kiekis informacijos. Ši informacija yra bevertė, jei negalima jos panaudoti. Tam, kad būtų galima greitai ieškoti tarp didelio kiekio informacijos, yra naudojamos paieškos sistemos.

Jos yra sudėtingas programų rinkinys, skirtas greitam pateiktos frazės ieškojimui tarp daugybės duomenų, kurie yra surinkti ir išsaugoti paieškos sistemos duomenų bazėje. Paieškos sistemose esanti informacija yra atnaujinama ir papildoma kiekvieną dieną, naudojant automatinius informacijos paieškos robotus.

Paieškos sistema pateikia rezultatus, kurie yra susiję su paieškos fraze. Jie gal būti sudaryti iš internetinių puslapių, paveikslėlių ar kitokio tipo failų.

## **1.1. Paieškos sistemos robotai**

Internetė esantys tinklapiai yra indeksuojami paieškos sistemų robotų pagalba. Tai yra automatinės programos naudojančios tam tikrus algoritmus. Jų pagalba yra atrandami nauji bei pastoviai, kas tam tikrą laiko tarpą, peržiūrimi jau esantys paieškos sistemos indekse puslapiai.

Paieškos sistemų robotų darbas prasideda nuo sąrašo tinklapių adresų, kurie bus peržiūrimi. Šie adresai yra gaunami iš prieš tai buvusio indeksavimo ir papildomi su adresais esančiais svetainės, kuri bus peržiūrima, žemėlapyje.

„Google“ robotas, kiekvieną kartą aplankydamas svetainę, aptinka nuorodas. Tai yra žymų atributuose „SRC“ ir „HREF“ esančias reikšmes. Jomis papildo jau esamą adresų sąrašą, kuris toliau bus peržiūrimas. Tai, kiekvieną kartą yra aplankomas vis didesnis kiekis svetainės puslapių skaičius.

Visa gauta informacija, tokia kaip: nauji puslapiai ar svetainės, nebeveikiančios nuorodos, yra panaudojama, atnaujinant paieškos svetainės indeksą [GWC12].

## **1.2. Svetainės turinio indeksavimas**

Paieškos robotui aplankius svetainę, jos turinys yra nusiunčiamas indeksavimo programoms. Jos išfiltruoja gautą turinį. Tai yra, pašalinamos nereikalingos žymos ir paliekamas tik tekstas, kuriame yra galima atlikti paiešką. Šis turinys toliau yra saugomas „Google“ indekso duomenų bazėje.



Indeksas yra surikiuotas abėcėlės tvarka, pagal paieškos frazę. Su kiekvienu indekso įrašu yra saugomas dokumentų sąrašas, kuriuose yra frazė ir vieta, kur tekste ji pasitaikė. Tokia duomenų struktūra leidžia labai greitai pasiekti paieškos frazes.

Paieškos vykdymo paspartinimui, „Google“ nenaudoja dažnai pasitaikančių anglišku žodžių (pvz.: „the“, „is“, „of“), pasikartojančių tarpo ar kitų simbolių arba atskirų skaičių arba raidžių. Šie žodžiai yra tokie dažni, kad gali per daug susiaurinti paiešką, todėl jų galima saugiai atsisakyti. Taip pat visos raidės yra verčiamos mažosiomis raidėmis [BP98].

### **1.3. Paieškos vykdymas**

Paieškos užklausų rezultatai yra matomi naudotojams. Kai jis įveda užklausą, paieškos variklis grąžina nuorodas su aprašymais, kurie atitinka tiksliausiai. Tai yra nustatoma naudojant per du šimtus įvairiausių veiksmų, kai kurių kompanija „Google“ neatskleidžia, nes tai yra komercinė paslaptis.

Vienas iš atskleidžiamų veiksmų yra puslapio reitingas (angl. „page rank“). Jis naudojamas išmatuoti puslapio svarbumą, pagal išorinių svetainių nuorodas. Kuo didesnis puslapio reitingas, vadinasi tuo daugiau svetainių turi nuorodas į tą puslapį ir jis yra patikimesnis. Žinoma, yra ir netinkamų nuorodų, kurių pagalba bandoma dirbtinai sukelti šį reitingą, tačiau „Google“ nuolat kuria naujus būdus, kaip aptikti tokias netinkamas nuorodas.

Tai pat „Google“ paieškos varikliui yra svarbesni puslapiai, kurie turi paieškos frazes šalia viena kitos ir išdėlioti ta pačia tvarka, kaip paieškos užklausoje. Variklis taip pat gali rasti daugiažodes frazes arba sakinius.

Paieškos sistemos naudotojams yra suteikiama galimybė uždrausti paieškai kreipti dėmesį į paieškos žodžius esančius puslapio „title“, „body“ ar kitų žymų viduje [GWC13][BP07].

Visas paieškos procesas susideda iš kelių dalių:

1. naudotojas pateikia užklausą „Google“ paieškos sistemai;
2. užklausa keliauja į tinklo serverį;
3. tada užklausą siunčiama į indekso serverius. Juose yra nustatomi puslapiai, kuriuose yra pateiktos užklausos žodžiai;
4. užklausa keliauja į dokumentų serverį. Jame pagal indekso serverio rezultatus yra grąžinami dokumentai (nuorodos į juos) ir sugeneruojami maži gabaliukai teksto, kurie apibūdina kiekvieną paieškos rezultatą;
5. rezultatai yra grąžinami per akimirksnį nuo naudotojo užklausos pateikimo.

## 2. SVETAINĖS OPTIMIZAVIMO PAIEŠKOS SISTEMOMS ANALIZAVIMAS

Kokybiškas svetainės optimizavimas yra sunkus darbas. Norint, kad svetainė sulauktų didelio skaičiaus lankytojų, yra būtina žinoti ne mažai duomenų apie svetainę. Šiems darbams atlikti „Google“ sukūrė puikius įrankius. Jų pagalba puslapio savininkas gali daug lengviau atlikti svetainės optimizavimo darbus.

### 2.1. Įrankiai

Svetainių optimizavimui yra naudojami tam tikri įrankiai. Jų pasirinkimas yra didelis, tačiau profesionalūs įrankiai yra mokami. Toliau yra pateikta keletas populiariausių įrankių, kurie yra nors ir nemokami, tačiau itin kokybiški ir populiarūs:

**„Google Webmaster Tools“** – šio įrankio pagalba puslapio savininkas gali:

- peržiūrėti sąrašą raktažodžių, dėl kurių svetainė buvo parodyta „Google“ paieškos rezultatų puslapyje;
- pateikti svetainės žemėlapi, kurį „Google“ naudoja puslapio indeksavimui;
- peržiūrėti puslapio, suindeksuoto „Google“ paieškoje, statistiką;
- pamatyti puslapio turinio sutvarkymo pasiūlymus;
- pamatyti sąrašą išorinių ir vidinių nuorodų, kurios rodo į svetainę;
- sugeneruoti „robots.txt“ failą.

Šis įrankis yra nuolat tobulinamas, todėl jame yra sukuriama vis naujų ir naudingų funkcijų svetainės savininkams. „Google Webmaster Tools“ yra pasiekiamas adresu: <https://www.google.com/webmasters/tools>.

**„Google Analytics“** – šio įrankio pagalba, puslapio administratorius gauna išsamią statistiką apie jo interneto svetainės naudojimą. Duomenys pateikiami tiek realiaame laike, tiek ataskaita iš duomenų bazės pagal pasirinktą laikotarpį, kuris gali būti nurodomas net valandų tikslumu. Pasinaudojęs šiuo įrankiu, svetainės savininkas realiaame laike gali pamatyti:

- kiek svetainėje yra lankytojų;
- kurie svetainės puslapiai yra šiuo metu lankomi;
- iš kokio šaltinio atėjo lankytojai. Jei iš „Google“, tai dar yra parodoma kokių raktažodžių pagalba naudotojas rado puslapį paieškos rezultatų lange;
- kokias naršykles naudoja lankytojai;
- iš kokių miestų ir šalių yra lankytojai.

Puslapio administratorius pasinaudojęs ataskaitą pagal pasirinktą laikotarpį gali:

- naudoti tas pačias funkcijas kaip ir realaus laiko ataskaitos;
- kiek lankytojų apsilankė svetainėje ar viename iš jos puslapių;
- kiek laiko užtruko puslapyje;
- į kokias svetaines išėjo puslapio lankytojai;
- pamatyti kiek lankytojų sugrįžo į svetainę ir kiek kartų apsilankė joje;
- kokia naršykle naudojosi lankytojai.

Be paminėtų funkcijų, puslapio savininkas gali stebėti, kokios nuorodos puslapyje susilaukė didžiausio paspaudimų kiekio. Tam yra sukurta funkcija, kuri prideda papildomą sluoksnį puslapyje, kuriame yra rodomi paspaudimai ant nuorodų.

Svarbu paminėti, kad šis įrankis yra nuolat tobulinamas, todėl pastoviai papildomas naujomis funkcijomis.

Įsidiesti šį įrankį į svetainę yra labai lengva. Tiesiog užtenka įterpti „JavaScript“ kodą į norimus stebėti svetainės puslapius ir jis jau galės vykdyti stebėjimus. Šis įrankis yra pritaikytas net tik svetainėms, kurios turi „JavaScript“ palaikymą, bet ir turi aplikacijų programavimo sąsają pritaikytą ir kitokių technologijų svetainėms.

„Google Analytics“ pasiekiamas adresu: <http://www.google.com/analytics/>.

„**Google keyword tool**“ – visiškai nemokamas ir naudojamas nustatant pagal kokius raktažodžius daugiausiai vykdomos paieškos. Šio įrankio pagalba naudotojas gali:

- gauti raktinių žodžių, bei jų junginių su kitais raktažodžiais pasiūlymus. Todėl lengviau gali išsirinkti naudosimą frazę;
- pamatyti raktažodžių konkurencingumą. Jei raktažodis yra naudojamas daugelio svetainių, tai jo konkurencija aukšta, todėl į pirmąsias pozicijas patekti yra gana sunku, todėl reiktų rinktis raktažodžius su mažiausiu konkurencingumu;
- gauti paieškų skaičių, kuris yra pateikiamas prie kiekvieno raktažodžio ar jo derinio, kuris padeda nuspręsti, kurie yra populiariesni. Atsižvelgiant į tai, palengvėja tinkamų raktažodžių parinkimas;
- gauti raktažodžius, pagal kuriuos yra reitinguojama tam tikra svetainė.

Tinkamiausi raktažodžiai yra tie, kurių konkurencingumo rodiklis yra kuo žemesnis, o paieškų skaičius kuo didesnis. Kadangi tokias frazes rasti yra gana sunku, nes jų yra mažai, todėl paieškai reikia sugaišti nemažai laiko. Tačiau tai padarius, rezultatai bus tikrai geri.

„Google keyword tool“ pasiekiamas adresu: <https://adwords.google.com/o/KeywordTool>.

## 2.2. Raktažodžių analizė ir parinkimas

Optimizavimo pradžioje yra svarbu ištirti rinką, kurioje bus konkuruojama. Tam reikia susirinkti raktažodžius, kurie bus naudojami puslapių turinyje. Žinoma, nustatant, kokie raktažodžiai bus naudojami, yra svarbu atsižvelgti į jų konkurencijos lygį. Jei raktažodžiai yra labai konkurencingi ir nesusieti su turiniu, tai naudojant juos puslapio optimizacija neduos gerų rezultatų. Jeigu tai bus atlikta teisingai, tai padės išvengti atvejų, kai lankytojai ateina į svetaines, kurios naudoja klaidinantį turinį ir yra bevertės. Tokiose svetainėse lankytojai nepraleidžia daug laiko, todėl yra mažesnė tikimybė, kad nuoroda į svetainę pasidalins internete [Hol13].

Svetainės puslapių raktažodžių parinkimui yra puikiai tinkantis įrankis – „Google Keyword tool“, apie kurį buvo aprašyta anksčiau. Naudojant jį, raktažodžius reikia parinkti pagal:

- lankytojų srautą. Prie kiekvieno raktažodžio ar jų junginio yra rodomas paieškų skaičius per mėnesį. Jeigu jis yra labai mažas, tai yra tikėtina, kad optimizuoti puslapiai sulauks mažo skaičiaus lankytojų. Kita vertus, pasirinkus labai populiarius raktažodžius yra tikimybė, jog jų konkurencija taip pat bus labai aukšta.
- Raktažodžių konkurencingumą. Kadangi labai yra sunku konkuruoti su rinkos lyderiais, kurių svetainės yra sukurtos tikrai seniai ir turi daug nuorodų iš kitų šaltinių, todėl reikia bandyti surasti mažiau konkurencingus raktažodžius. Labiausiai verti dėmesio yra mažo ar vidutinio konkurencingumo lygio raktažodžiai.

Dažnai pasitaiko atvejų, kad yra norima konkuruoti su tam tikra svetaine. Tada yra naudinga sužinoti dėl kokių raktažodžių svetainė turi tokius aukštus reitingus. Čia taip pat yra naudojamas „Google Keyword tool“. Tereikia nurodyti svetainės adresą, apie kurios raktažodžius norima sužinoti. Rezultatai rodomi taip pat, kaip ir paprastoje raktažodžių paieškoje.

Tikslingų raktažodžių parinkimas specialiai tam pritaikytais įrankiais atneša didelę naudą, nes sutaupoma laiko ir geriau išnaudojami potencialūs raktiniai žodžiai.

### 3. SVETAINĖS OPTIMIZAVIMAS „GOOGLE“ PAIEŠKOS SISTEMAI

#### 3.1. Svetainės vidinis optimizavimas

Šiame svetainės optimizavimo proceso dalyje yra naudojami veiksmai, kurie yra atliekami svetainės viduje. Šio etapo tikslas yra padaryti svetainę kuo lengviau skaitomą paieškos sistemos robotams. Tai atlikus, jau yra galima tikėtis, kad paieškos sistemos robotai lengvai suindeksuos svetainės turinį ir svetainė atsidurs aukščiau rezultatų sąrašo viršaus, nei buvo prieš tai.

##### 3.1.1. Puslapio pavadinimas

Vienas iš svarbiausių faktorių, lemiančių paieškos rezultatus, yra puslapio pavadinimas. Puslapio pavadinimas yra reikmė, priskirta „<title>“ žymos. Ji dažniausiai yra randama „HTML/XHTML“ dokumento viršuje tarp „<head>“ žymos .

```
<!DOCTYPE html>
<html xmlns="http://www.w3.org/1999/xhtml" lang="lt">
  <head>
    <title>Mano puslapio pavadinimas</title>
  </head>
```

Puslapio pavadinimas yra dažniausiai vaizduojamas naršyklės lango viršuje arba naršyklės kortelės viršuje.

„Google“ paieškos sistemoje kiekvienas rezultatas yra atvaizduojamas su puslapio pavadinimu (žr.: pav. 1). Jis pateikiamas kaip nuoroda į puslapį.

Google

matematikos ir informatikos fakultetas

Žiniatinklis Vaizdai Daugiau Paieškos įrankiai

Apie 55.100 rezult. (0,19 sek.)

**Mieli studentai, — Matematikos ir Informatikos fakultetas**  
[www.mif.vu.lt/](http://www.mif.vu.lt/) - „Google“ kopija  
Žinios apie **fakultetą** ir jos struktūrą, studijas, informacija stojantiems ir studentams.

**Paštas**  
Vilniaus universiteto Žiniatinklio paštas. Naujokų registracija ...

**Studijų skyriaus informacija**  
Jūs esate čia: Pradžia › Studijos › Studijų skyriaus informacija ...

**Fakulteto naujienos**  
Šiame skyriuje rasite naujienas, aktualias fakulteto ...  
[Daugiau rezultatų iš vu.lt »](#)

**Katedros**  
CERN projektai · STSC. Jūs esate čia: Pradžia › Struktūra ...

**Tvarkaraščiai**  
Bakalauro studijos. Programų sistemos. PS, 1 k., 1 gr. PS, 1 k ...

**Stojantiejiems**  
Čia pateikiamos reikalingos nuorodos ir informacija apie ...

Pav. 1: „Google“ paieškos sistemos pateiktas rezultatas.

Paieškos varikliai naudoja „<title>“ žymą, kad parodytų paieškos rezultatus. Jeigu puslapis neturi pavadinimo arba jis yra bereikšmis (T.y. neatitinkantis puslapyje esančio turinio), tai pasunkina puslapio indeksavimą. Puslapio pavadinimas apibūdina robotui apie ką yra puslapis.

Puslapio pavadinimas negali būti per daug ilgas, nes „Google“ paieškos sistema gali atvaizduoti tik 70 simbolių savo paieškos rezultatų puslapyje esančiuose pavadinimuose.

Žodžiams, kurie vaizduojami pavadinime pirmieji, yra suteikiama didesnė svarba, nei žodžiams kurie vaizduojami toliau. Pavyzdžiui, jeigu puslapyje yra kalbama apie tai, kaip sutaisyti kompiuterio Asus K53 sugedusį kietąjį diską, pagrindiniai raktažodžiai yra „sutaisyti“, „Asus“, „K53“, tai puslapio pavadinimas toks kaip:

```
<title>Modesto pagalbos naudotojui tinklapis: Kaip sutaisyti sugedusį kietąjį diską Asus K53 nešiojamam kompiuteriui</title>
```

Gali būti pakeistas į:

```
<title>Kaip sutaisyti Asus K53 nešiojamojo kompiuterio kietąjį diską</title>
```

Palyginus pradinį variantą su pakeistuoju, galime pastebėti, kad puslapio pavadinimas ne tik sutrumpėjo, bet ir tapo lengviau skaitomas bei informatyvesnis. Taip pat pastebėjome, jog atsisakėme ir tinklapio pavadinimo, tai svarbiausiems raktažodžiams atsidurti pirmiau, kas pagerina poziciją paieškos rezultatuose.

Kai svetainė yra sudaryta iš daug puslapių, yra svarbu atkreipti dėmesį, jog puslapių pavadinimai nesikartotų, nes tai sudaro sunkumų, kai rezultatai turi būti atvaizduojami rezultatų puslapyje ir taip yra sunkiau atskirti kuris puslapis yra kuris, be to kiekvienas puslapis turi būti unikalus [Man10].

### 3.1.2. Meta aprašymas ir raktažodžiai

Meta žymų pagalba, tinklapių savininkai pateikia informaciją apie savo puslapius paieškos varikliams [GWC12b]. Meta žymos yra rašomos tarp „<head>“, žymos. Žemiau yra pateiktas jos pavyzdys:

```
<!DOCTYPE HTML>
<html>
  <head>
    <META NAME="description" CONTENT="Puslapio aprašymas">
    <META NAME="keywords" CONTENT="raktažodis1, raktažodis2">
```

```
<title>Puslapio pavadinimas</title>
<META NAME="robots" CONTENT="noindex,nofollow">
</head>
</html>
```

Kiekvieno paieškos variklio robotas, skaitydamas puslapį, supranta tik tam tikras meta žymas ir tik į jas kreipia dėmesį.

Ankščiau buvo populiari naudoti „keywords“ meta žymos atributą, todėl daugelyje puslapių dar iki šiol galime jį rasti. Tačiau „Google“ robotas šios žymos atributo net nesupranta, todėl indeksavimui jis nedaro jokios įtakos. Jo buvo atsisakyta ne atsitiktinai, nes buvo pastebėtas per didelis piknaudžiavimas iš puslapių kūrėjų pusės. Labai dažnai raktažodžiai buvo parenkami ne pagal puslapio turinį, o tiesiog pagal konkurencingumą.

Tačiau, „Google“ robotas supranta meta žymą su „description“ atributu. Jos pagalba puslapio savininkas gali pateikti trumpą informaciją apie ką yra puslapis. Paieškos varikliai aptys nusprendžia, ar ši pateikta informacija yra tinkama ar ne. Jei taip, tai naudojama paieškos rezultatų puslapyje, kaip trumpas aprašymas po pavadinimu su nuoroda.

### 3.1.3. Svetainės adresų struktūra

Adresų struktūra svetainėje turi būti kuo paprastesnė, kad kuo lengviau būtų suprantama puslapio lankytojams [GWC10]. Puslapio kūrėjas, turi atsižvelgti į tai, kaip yra kuriami svetainės puslapių adresai. „Google“ pataria, kad jie turi būti konstruojami naudojant lengvai įskaitomus žodžius ir nenaudojant sesijos kintamųjų nuorodose, nes tai pasunkina suprantamumą lankytojams. Jeigu svetainėje esantis puslapis yra apie muziką, tai adresas <http://www.svetaine.lt/muzika> būtų kur kas daug logiškesnis ir aiškesnis, negu <http://www.svetaine.lt/index.php?id=15&item=a155s>. Pateikus tokią nuorodą lankytojui yra visiškai neaišku, kas bus puslapyje. „Google“ paieškos sistema tokiems adresams teikia daug mažesnį prioritetą. Dėl adresuose esančių sesijos kintamųjų, puslapių tikrinimo robotas taip pat gali sunaudoti didesnę duomenų srautą, negu yra reikalingas svetainės turiniui indeksuoti, kadangi gali būti tikrinami puslapiai su vienodu turiniu.

Svetainėje esantys puslapiai gali turėti per didelį nuorodų kiekį dėl [GWC12f]:

- puslapiuose, kuriose pateikiamos dainos, gali būti rūšiuojamas pagal autorių ar pavadinimą. Dėl to gali atsirasti nereikalingas nuorodų skaičiaus padidėjimas, kuris tiesiog nurodo į identišką turinį turinčius puslapius.

- Dinaminio turinio generavimo, dėl kurio galimi nedideli pakeitimai pagal skaitiklius, laiko formatus ar reklamas, nors kelių puslapių pagrindinis turinys yra identiškas.
- Netinkamų santykinių nuorodų, dėl kurių gali įvykti begalinės sekos. Tai dažniausiai įvyksta dėl pasikartojančių adreso kelio elementų – <http://www.svetaine.lt/kategorija/muzika/kategorija/mp3>.

Puslapių administratoriai siekdami išvengi panašių klaidų, gali:

- naudoti „robots.txt“ bylą, probleminių nuorodų blokavimui. Paprastai yra blokuojamas dinamiškas turinys arba galima begalinė seka pakeitus, kokų nors parametą.
- Trumpinti nuorodas su nereikalingais parametrais.
- Sesijoms naudoti „sausainius“ (angl. cookies).

„Google“ taip pat rekomenduoja skirti skyrybos ženklais žodžius esančius adresuose. <http://www.svetaine.lt/kategorija/pasaulinesmuzikostop40> yra ne tokia naudinga/suprantama svetainės naudotojams kaip <http://www.svetaine.lt/kategorija/pasaulines-muzikos-top-40>. Pastarojoje, žodžiai yra atskirti brūkšnelio ženklu „-“, tai daro adresą aiškesniu. Taip pat šis ženklas yra labiau rekomenduotinas, nei „\_“.

Pasaulyje yra labai daug „PHP“ kalba parašytų svetainių, kurios naudoja dinaminį turinį. Taip yra dėl to, kad svetainėje esantį turinį yra daug patogiau saugoti duomenų bazėje, nes taip yra lengviau jį prižiūrėti/administruoti. Tam yra naudojami ir dinaminiai adresai, dėl kurių parametų keičiasi puslapiuose esantis turinys. Tačiau dinaminiai adresai nėra gera praktika optimizuojant puslapius, nes „Google“ paieškos sistemoje jie yra vertinami prastesniu reitingu, nei puslapiai su statiniais adresais. Taigi, jeigu svetainė naudoja dinamines nuorodas yra galimybė naudoti „htaccess“ bylą, kurioje yra naudojamos tam tikros taisyklės. Viena iš jų yra „RewriteRule“.

Jeigu svetainė naudoja dinaminį adresą <http://www.svetaine.lt/search.php?type=song&page=1&search=britney>, tai ją galima pakeisti į tokią kaip ši: <http://www.svetaine.lt/dainu-paieska/britney/1.html> statinę nuorodą, kuri paieškos sistemoje yra daug vertingesnė. Norint tai padaryti, reikia „htaccess“ byloje įrašyti taisyklių rinkinį:

RewriteEngine on



```
RewriteRule ^dainu-paieska/([a-zA-Z0-9-]+)/(\d+).html?$
search.php?type=song&page=$2&search=$1[L]
```

Pirmoje šio failo eilutėje yra nurodoma, jog bus naudojamos taisyklės peradresavimui. O antroje eilutėje nurodoma, kuri nuorodos dalis turėtų būti peradresuota. „^“ simbolis reiškia domeno pradžią, „([a-zA-Z0-9-])“ parodoma reguliariosios išraiškos pagalba, kad gali būti pateikiami didžiosios, mažosios raidės, skaičiai ir brūkšniai. „(\d+)“ – parodo, kad už pasvirojo brūkšnio yra iš skaičių sudaryta adreso dalis. Visi adresai, kurie atitinka šią taisyklę yra peradresuojami į dinaminę „search.php“ nuorodą, kurioje bus sudedamos atitinkamos reikšmės ir įvykdoma užklausa, tačiau tikrasis dinaminės nuorodos adresas nebus matomas naudotojui.

Taigi, norit, kad svetainės adresas būtų suindeksuotas tinkamai, reikia naudoti adreso perrašymą raktažodžiais ir vengti nepageidaujamų simbolių.

### 3.1.4. Svetainės žemėlapiai

Interneto svetainės susideda ne iš vieno puslapio. Kartais jos turi kelis šimtus ar net kelis tūkstančius. Kadangi svetainės puslapiai ne visada yra susieti nuorodomis, naudoja dinaminį turinį, didelė dalis turinio užkraunama naudojant „AJAX“ technologiją ar turi didelį nesusietų nuorodų archyvą, todėl robotai indeksuodami turinį visų svetainės puslapių gali ir neaptikti [GWC12d]. Tam, kad būtų suindeksuoti tokie svetainės puslapiai, yra naudojami svetainių žemėlapiai, kurie yra pateikiami paieškos svetainėms (šiuo atveju „Google“) ir juos naudodamos indeksuoja svetaines. Tačiau „Google“ negarantuoja, kad bus peržiūrėti visi svetainės žemėlapyje esantys puslapiai, tačiau tai vis tiek pagerins bendrą svetainės puslapių aptikimo rodiklį.

Norint, kad svetainės žemėlapi „Google“ galėtų tinkamai nuskaityti, jį reikia sudarinėti atsižvelgiant į specialius reikalavimus [GWC12e]:

- turi naudoti „UTF-8“ koduotę.
- Turi būti sudarytas naudojant „XML“ kalbą.
- Svetainės žemėlapis turi prasidėti „XML“ schema su atributu *xmlns=http://www.sitemaps.org/schemas/sitemap/0.9* iš <http://www.sitemaps.org/>.
- Svetainių adresai esantys žemėlapyje turi būti unikalūs.
- Turi naudoti „UTF-8“ koduotę.
- Vienoje svetainės žemėlapyje neturi būti daugiau nei 50 000 adresų.
- Turint daugiau nei vieną žemėlapi, galima naudoti svetainių žemėlapių rodyklę, kad būtų pateikti visi svetainės žemėlapiai.

## Svetainės žemėlapis pavyzdys „XML“ kalba:

```
<?xml version="1.0" encoding="UTF-8"?>
  <sitemapindex xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">
    <url>
      <loc>http://www.svetaine.lt/</loc>
      <priority>0.8</priority>
    </url>
    <url>
      <loc>http://www.svetaine.lt/puslapis-1.html</loc>
      <changefreq>weekly</changefreq>
      <priority>0.8</priority>
    </url>
    <url>
      <loc>http://www.svetaine.lt/puslapis-2.html</loc>
      <lastmod>2013-02-02</lastmod>
    </url>
  </sitemapindex>
```

Svetainės žemėlapis yra sudarytas iš „XML“ žymų:

- „<loc>“ – skirta nurodyti vieną svetainėje esantį adresą.
- „<changefreq>“ – skirta nurodyti kaip dažnai keičiasi puslapyje esantis turinys. Ši žyma nėra privaloma. Šio elemento pagalba paieškos sistemos robotui nurodoma kaip dažnai bus peržiūrima svetainė. Galimos šio elemento reikšmės [GWC12e]:
  - „always“ – visada (puslapis pasikeičia kiekvieną kartą, kai jis buvo atidarytas);
  - „hourly“ – valandomis;
  - „daily“ – dienomis;
  - „weekly“ – savaitėmis;
  - „monthly“ – mėnesiais;
  - „yearly“ – metais;
  - „never“ – niekada (dažniausiai naudojamas archyvo adresams).
- „<priority>“ – nusako puslapio svarbumą. Reikšmės švytuoja nuo „1.0“ – labai svarbu iki „0.1“ – visiškai nesvarbu. Ši žyma taip pat nėra būtina. Ši žyma nedaro jokios įtakos „Google“ paieškos rezultatuose esančiai svetainės pozicijai, tačiau „Google“ robotui padės nuspręsti apie svetainės struktūrą.

- „<lastmod>“ – šioje žymoje esanti reikšmė nurodo, kada paskutinį kartą buvo modifikuotas turinys esantis konkrečiame puslapyje.

Kai svetainės turi labai didelį kiekį puslapių, yra sunku svetainės žemėlapij pateikti viename faile, kadangi dėl labai didelio jo dydžio pailgėja failo siuntimo laikas, todėl yra patariama svetainės žemėlapius skaidyti į mažesnius, kurie būtų atsiųsti optimaliai ir be didelio užlaikymo.

Norint kelis ar daugiau žemėlapių pateikti „Google“ paieškos sistemai, yra du galimi variantai:

- Naudojantis „Google Webmaster Tools“ įrankiu įkelti visus žemėlapius po vieną rankiniu būdu.
- Pateikti svetainės žemėlapių rodyklę, kuri jau turi visą sąrašą suskaidytų svetainės žemėlapių.

Svetainės žemėlapių rodyklė yra sudaryta XML failo pavidalu. Žemiau yra pateiktas jo pavyzdys, kuriame yra sujungiami du žemėlapiai. Sudarius tokią rodyklę, užtenka pateikti ją, kad būtų suindeksuoti atskiri žemėlapiai esantys joje.

```
<?xml version="1.0" encoding="UTF-8"?>
  <sitemapindex xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">
    <sitemap>
      <loc>http://www.svetaine.lt/zemelapis-1.xml</loc>
      <lastmod>2013-04-10</lastmod>
    </sitemap>
    <sitemap>
      <loc>http://www.svetaine.lt/zemelapus-2.xml</loc>
      <lastmod>2005-01-01</lastmod>
    </sitemap>
  </sitemapindex>
```

Svetainės žemėlapių rodyklės struktūra yra sudaryta iš XML žymų:

- „<sitemapindex>“ – su šia žyma yra apgaubiamas rodyklinis svetainės žemėlapis. Joje taip pat galima nurodyti atributą su svetainės žemėlapiu protokolu.
- „<sitemap>“ – šios žymos pagalba atskiriami svetainės žemėlapių parametrai.
- „<loc>“ – šioje žymoje yra saugomas svetainės žemėlapiu failo adresas.
- „<lastmod>“ – ši žyma skirta nurodyti paskutinę žemėlapiu modifikavimo datą.

Šis failas yra labai patogus, jei norima pateikti daug svetainės žemėlapių, tačiau jame esantis žemėlapių skaičius negali būti didesnis nei 500 [GWC11].

### 3.1.5. „robots.txt“ instrukcijos

Labai yra svarbu, kad norimas svetainės turinys būtų suindeksuotas, nereikalingas – uždraustas indeksuoti arba, jei reikia, pašalintas iš paieškos, o kai kurie robotai būtų uždrausti indeksuoti svetainę. Tam svetainėje yra patalpinamas specialus failas „robots.txt“, kurį naudoja paieškos sistemų robotai apsilankę svetainėje. Taigi kiekvieną kartą svetainėje <http://www.manosvetainė.lt/> robotas prieš pradėdamas darbą apsilankys adresu <http://www.manosvetainė.lt/robots.txt> [GWC12c]. Tik jeigu šiuo adresu neranda kitokių instrukcijų jie pradeda analizuoti ir indeksuoti svetainę.

Žinoma, ne visų paieškos svetainių robotai indeksuoja puslapius taikydami taisykles aprašytas „robots.txt“ faile, todėl jei svetainė vis gi turi turinio, kuris neturėtų būti įkeltas į paiešką, turėtų būti papildomai apsaugotas, jeigu tokių puslapių nėra, arba šis failas yra tuščias, nėra jokios būtinybės jį talpinti svetainėje.

Verta paminėti, jog jeigu „robots.txt“ faile yra užblokuotas puslapis nuo indeksavimo, jį vis tiek gali suindeksuoti [GWC12c]. Tai įvyktų tada, jei robotui pavyktų rasti nuorodų iš kitų svetainių, būtent į užblokuotą puslapį. Tokiu atveju, jis vis tiek bus suindeksuotas ir atsiras paieškos rezultatuose. Tačiau, jeigu norima visiškai užblokuoti puslapį nuo patekimo į rezultatų puslapį reikia naudoti „noindex“ meta žymą norimo užblokuoti puslapio antraštėje.

„Google“ robotas supranta šias taisykles:

- „User-agent“ – jos pagalba aprašomi robotai, kuriems bus taikomos taisyklės.
- „Disallow“ – ja nurodoma blokuojamas puslapis ar aplankas.
- „Sitemap“ – jos pagalba nurodoma svetainės žemėlapių vieta.

Paprastas „robots.txt“ failo pavyzdys:

```
User-agent: googlebot
Disallow: /naujienos/slapta/
Disallow: /archyvas/slapta/
Sitemap: http://pavyzdys.lt/zemelapis.xml
User-agent: *
Disallow: /domenys/neindeksuojami/
Disallow: /neindeksuojamas.html
Disallow: /nuotraukos/slapta*.png
```

Pirmoje pavyzdžio grupėje matome „googlebot“, tai reiškia, jog taisyklės taikomos tik „Google“ paieškos robotui. Taigi taip yra užblokuojami du aplankai nuo „Google“ roboto.

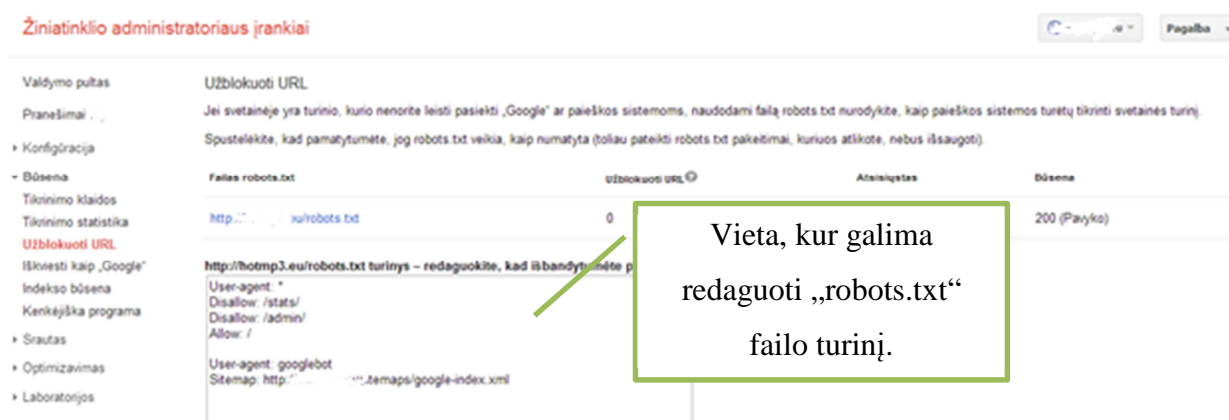
Taip pat „Sitemap“ taisyklės pagalba yra nurodomas svetainės žemėlapis. Būtina pastebėti, jog jis turi būti nurodomas su pilnu adresu. Jo pagalba robotas ieškos svetainės žemėlapio nurodytoje vietoje ir nebus reikalinga jį atskirai pateikti „Google Webmaster Tools įrankiui“.

Toliau matome „\*“ . Šis simbolis reiškia, jog ši taisyklė yra taikoma visiems robotams, ne tik „Google“. Kiekvienas skyrius prasidedantis „User-agent“ yra naujas ir jis nesiremia prieš tai buvusių skyriumi. Ši grupė pritaikoma vienai direktorijai, vienam „.html“ failui. Tai pat „Google“ robotai palaiko reguliarių išraiškų simbolius todėl ir visiems PNG tipo paveikslėliams prasidedantiems žodžiu „slapta“ ir turintiems galūnę „.png“ galioja ta pati draudimo indeksuoti taisyklė.

Aprašant taisykles robotams, kartais yra sunku pastebėti klaidas. Todėl jų tikrinimui galima naudoti išorinius įrankius, arba naudoti „Google Webmaster Tools“ esantį šio failo tikrinimo funkcionalumą.

Kartais tobulinant „robots.txt“ pasitaiko įvairių klaidų, todėl, kad „robots.txt“ būtų patikrintas, tereikia prisijungti į „Google Webmaster Tools“. Jame yra sukurtas įrankis, kurio pagalba įmanoma patikrinti ar teisingai sukurtos taisyklės robotams. Norint tai padaryti reikia:

- paspausti ant svetainės, kurios „robots.txt“ failas yra norimas patikrinti.
- Kairėje pusėje esančiame meniu paspausti „Būsena“.
- Išsiskleidusiam meniu praplėtime paspausti „Užblokuoti URL“.



Pav. 2: „robots.txt“ failo redagavimas ir klaidų tikrinimas naudojant „Google Webmaster Tools“ įrankį.

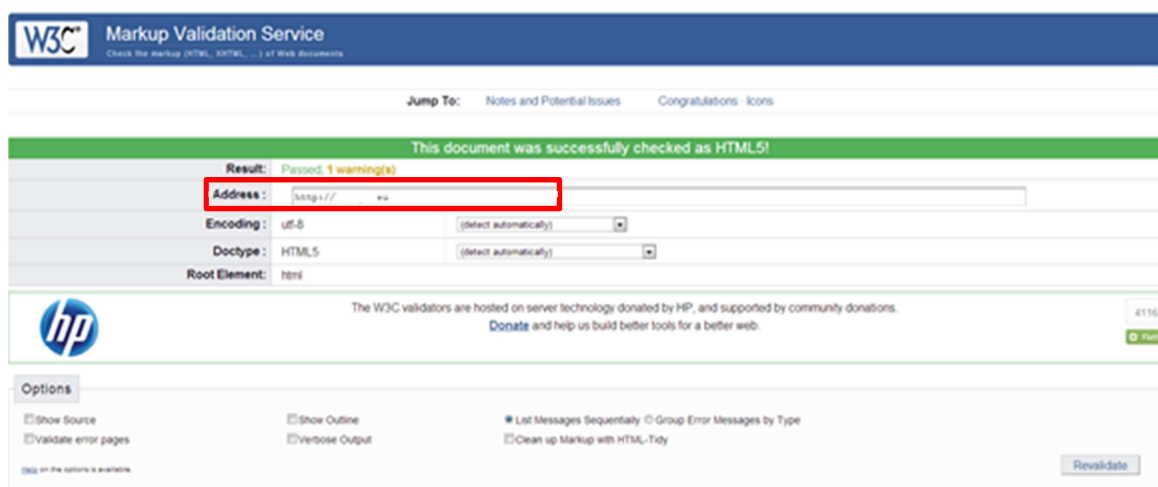
Nurodytoje vietoje galima įdėti redaguotą „robots.txt“ failo turinį ir paspaudus mygtuką „Išbandyti“, patikrinti ar failas yra suformuotas be klaidų. Jeigu jame randamos klaidos, tai pateikiami bandymo rezultatai su nurodytomis klaidomis ir jų aprašymais.

Ištaisius klaidas reikia nepamiršti „robots.txt“ failo turinį esantį serveryje pakeisti į ištaisytą, nes jis nėra išsaugomas automatiškai.

### 3.1.6. Svetainės klaidų tikrinimas ir šalinimas

Optimizuojant svetaines paieškos varikliams yra svarbu atkreipti ir į „HTML“ kodo taisyklumą. Jeigu „HTML“ kodas neatitinka „W3C“ kodavimo standartų, tai paieškos sistemų robotai gali nebaigti skaityti puslapio struktūros bei neteisingai suindeksuoti puslapius ir tada tokių puslapių optimizavimas tampa bevertis [Lei13]. Taip pat netaisyklingai parašius „HTML“, „XML“ ar „CSS“ kodą ar taisykles, svetainės vaizdas gali būti nevienodai atvaizduojamas skirtingose naršyklėse ir tai sukeltų daug nepatogumų svetainės naudotojams.

Norint patikrinti ar svetainės puslapis atitinka W3C standartus, galima patikrinti puslapyje <http://validator.w3.org/>, jame esančiame adreso (angl. „address“) laukelyje įvedus norimo patikrinti puslapio adresą (žr.: pav. 3).



Pav. 3: Svetainės klaidų tikrinimas pagal W3C standartus.

Jeigu svetainė atitinka kriterijus, tai parodo žinutė puslapio viršuje, o jei ne, tai yra pateikiamas visas klaidų ir perspėjimų sąrašas, su rastomis klaidomis ir jų vietomis. Kai svetainė yra sudaryta iš labai daug puslapių, yra sunku užtikrinti, kad visi puslapiai būtų taisyklingi, nes joje turinys yra dinamiškas ir jeigu jis dar būna keičiamas nepatyrusių naudotojų, tai labai lengva padaryti klaidų, dėl kurių atsiranda klaidų ir jos pasunkina robotų indeksavimo procesą.

Patikrinus svetainę su šiuo įrankiu bei ištaisius rodomas klaidas, yra daug didesnė tikimybė, jog svetainė bus vienodai atvaizduota visose interneto naršyklėse ir sumažės tikimybė netinkamai robotų suindeksuotų puslapių.

### **3.2. Svetainės išorinis optimizavimas**

Atlikus svetainės vidinį optimizavimą dar nereiškia, kad bus pasiekiami norimi rezultatai „Google“ paieškoje, kadangi rezultatuose pirmesni puslapiai yra rodomi su aukštesniu reitingu. Šis reitingas kyla (ar krinta) dėl kiekvieno naudotojo parašytos nuomonės su nuoroda į svetainę. Jei svetainė su geru atsiliepimu, kurioje yra įdėta nuoroda, turi gana aukštą reitingą, tai ši svetainės nuoroda padės pakelti reitingą, nes kiekviena tokia nuoroda parodo, kad turinys yra populiarus ir susilaukiantis dėmesio.

Svetainės turinys turi būti naudingas naudotojui, jei taip yra tai jis gali juo pasidalinti su kitais, tiek per socialines svetaines, forumus ar dienoraščius. Taip svetainė tampa vis populiareesnė. Sulaukia ne tik naujų lankytojų, bet ir yra padidinamas naudingų nuorodų skaičius, kuris padeda pasiekti geresnius optimizavimo rezultatus.

Nuorodų skaičiui padidinti yra naudojamas ir nuorodų ar nuorodų dedamų į straipsnius pirkimas bei nemokamų paslaugų/daiktų teikimas populiarioms svetainėms, kurių puslapių reitingai yra labai aukšti, tačiau tai pažeidžia „Google“ taisykles. Jeigu yra pastebima, kad svetainė pasinaudojo tokiomis paslaugomis, ji yra baudžiama sumažinant puslapio reitingą [GWC12g]. Taip pat stengiantis padidinti puslapių reitingą yra draudžiama:

- mainytis nuorodomis;
- talpinti nuorodas į nesusijusius ar pavojingus puslapius;
- mainytis nuorodomis su partneriais, kad patiltų puslapio reitingas;
- naudoti automatines programas, kurios talpina nuorodas svetainėse, kad pakeltų puslapio reitingą;
- naudoti tekstinę reklamą, kad būtų pakeltas puslapio reitingas;
- talpinti nuorodas į žemos kokybės nuorodų „fermas“;
- talpinti nuorodas puslapio apačioje į kitas svetaines;
- naudoti forumo komentaruose parašus, kuriuose yra nuorodos į puslapį.

Jeigu puslapyje yra komentarų ar kitokia skiltis, kurioje naudotojai patys redaguoja turinį ir talpina nuorodas į puslapius, kad nebūtų pažeistos taisyklės reikia naudoti rel="nofollow" atributą nuorodos žymoje „<a>“ į su svetainę nesusijusį puslapį.

```
<a href="http://xxx.eu/download/song/660617/indie-blog-heaven-the-strokes-hawaii-aloha.mp3" rel="nofollow" target="_blank" title="Indie blog heaven the - Strokes hawaii aloha mp3 download">Atsisiųsti dainą</a>
```

Dėl šio atributo jei nuoroda rodys ir į blogo turinio svetainę, puslapio reitingui tai nepakenks.



## 4. MUZIKOS PAIEŠKOS SVETAINĖS OPTIMIZAVIMAS

Optimizuojama muzikos paieškos svetainė yra <http://xxx.eu>. Ji skirta visiems interneto naudotojams. Svetinės duomenų bazėje yra saugoma per 2 milijonus nuorodų į „mp3“ tipo failus esančius visame pasaulyje, todėl ją sudaro labai didelis kiekis dainų paieškai skirtų puslapių. Pačių dainų serveryje nėra saugoma. Svetainė sukurta naudojant „PHP“ programavimo kalbą ir „CodeIgniter 2“ karkasą, kuris yra vienas populiariausių iš nemokamų karkasų. Duomenys saugomi „MySQL“ duomenų bazėje, o greitą paiešką svetainėje užtikrina „Sphinx“ duomenų bazės indeksavimo ir paieškos serveris. Pagrindinė svetainės optimizacijos paieškos varikliams rasta problema yra ta, kad paieškos rezultatų puslapiai beveik nėra susieti tarpusavyje, todėl paieškos robotui yra labai sunku surasti kitus svetainės puslapius. Dėl šio priežasties „Google“ paieškos rezultatų puslapyje yra labai mažas suindeksuotų puslapių kiekis. Tai patikrinti galima „Google“ paieškoje įvedus žodį „site:“ su domeno pavadinimu. Šiuo atveju – „site:xxx.eu“. Kadangi tokios užklausoje rezultatuose pateikiamas suindeksuotų puslapių skaičius yra labai nedidelis, todėl buvo pasirinkta sukurti svetainės žemėlapi.

### 4.1. Naudojamos priemonės

Svetainės vidiniam optimizavimui yra naudojama „PHP“ kalba ir jos karkasas „CodeIgniter 2“, kadangi svetainė jau yra sukurta naudojant juos. Programavimo įrankis buvo pasirinktas „Eclipse IDE“ pagrindu sukurtas „Aptana Studio“. Patogiam „MySQL“ duomenų bazės valdymui per interneto naršyklę naudojamas „phpMyAdmin“ įrankis. Svetainė yra talpinama jau veikiančiame „dedikuoti.lt“ dedikuotame serveryje. Optimizavimo rezultatams tikrinti yra naudojamas anksčiau aprašytas tinklapio administratoriui skirtas įrankis „Google Webmaster Tools“.

### 4.2. Optimizavimo principas ir eiga

Optimizuojamos svetainės veikimo principas yra dainų paieška duomenų bazėje. Lankytojas atėjęs į svetainę vykdo paiešką ir gauna tam tikrus rezultatus. Svetainė rezultatų gali pateikti labai didelį kiekį. Registruojant kiekvieno lankytojo paieškas galima sužinoti, kokių rezultatų lankytojas tikisi svetainėje ir koks turinys yra populiariausias. Tokių duomenų ši svetainė neregistruoja, todėl vienas pirmųjų darbų yra surinkti visas užklausoje, pagal kurias naudotojai ieško dainų ir išsaugoti jas duomenų bazėje. Tam, kad būtų išvengta nereikalingų duomenų kaupimo, užklausoje, pagal kurias nebuvo gauta jokių rezultatų, nėra saugomos duomenų bazėje.

Duomenų bazėje kuriamai lentelei reikia sudaryti 4 stulpelius. Vienas stulpelis skirtas „ID“, antrasis – „hash“, skirtas greičiau ieškoti užklausoje pagal sugeneruotus frazių „hash“ kodus,

trečiasis - skirtas saugoti paieškos užklausai, o ketvirtasis - užklausų skaičiui, pagal kurį bus nustatomas frazės populiarumas.

Duomenys išsaugoti duomenų bazėje bus naudojami sudarant svetainės žemėlapi „XML“ formatu ir pateikti „Google“ paieškos sistemai „Google Webmaster Tool“ pagalba. Pagal gautus duomenis ji suindeksuos svetainę, taip atsiras daugiau galimų rezultatų „Google“ paieškoje, dėl to ir padidės svetainės lankomumas.

Tokiu būdu svetainės naudotojai ieškodami dainų automatiškai generuos svetainės žemėlapi.

### 4.3. Paieškos užklausų saugojimas

Visų pirma darbas prasideda nuo naujos lentelės sukūrimo duomenų bazėje, kurioje ir bus saugomos sėkmingos mp3 dainų užklausos ir šių užklausų skaičius.

Darbo autorius duomenų bazės valdymui naudojo „phpMyAdmin“ įrankį. Šioje aplikacijoje esančiame užklausų lange buvo įvykdyta „SQL“ užklausa:

```
CREATE TABLE IF NOT EXISTS `searched_phrases_any` (  
  `id` int(20) NOT NULL AUTO_INCREMENT,  
  `hash` varchar(32) NOT NULL,  
  `phrase` varchar(500) NOT NULL,  
  `times` int(20) NOT NULL DEFAULT '1',  
  PRIMARY KEY (`id`),  
  UNIQUE KEY `hash` (`hash`)  
) ENGINE=MyISAM DEFAULT CHARSET=utf8 AUTO_INCREMENT=1;
```

„SQL“ funkcijos „CREATE TABLE“ pagalba yra sukuriama nauja lentelė. Šalia jos nurodoma komanda „IF NOT EXISTS“ ir naujos lentelės pavadinimas „searched\_phrases\_any“. Pastaroji komanda naudojama apsidraudimui, kad neįvyktų klaida duomenų bazėje, jei joje yra dar viena tokia pati lentelė. Toliau yra nurodomi lentelės stulpeliai ir jų parametrai.

Šioje duomenų bazės lentelėje „id“ yra lentelės identifikatorius. „Int(20)“ komanda nurodoma, kad tai bus integer tipo įrašai, užklausų lentelei visiškai užtenka tokio tipo lauko, kadangi maksimali šio lauko reikšmė yra 4294967295. „NOT NULL“ nurodymu, laukelis niekada nebus tuščias. Taip pat čia yra panaudojamas „AUTO\_INCREMENT“ atributas, kuris nurodo, jog šio laukelio reikšmė visada bus nustatoma automatiškai. Tai yra - didžiausią „id“ stulpelio reikšmę padidinant vienetu.

„hash“ stulpelyje yra saugomos naudotojų pateiktų užklausų maišos kodas, kuris yra 32 simbolių dydžio. Jame saugomos „MD5“ funkcija sugeneruotos frazės kodas.

„phrase“ stulpelyje saugoma dainų paieškos laukelyje įvestos naudotojų frazės. Saugoma informacija turi būti „VARCHAR“ tipo ir neilgesnė nei 500 simbolių. Įrašant frazę ilgesnę nei 500 simbolių, būtų nukerpama dalis iš dešinės į kairę. Dėl šio duomenų tipo yra sutaupoma duomenų bazės dydis.

„times“ stulpelyje yra saugomas naudotojų užklausų kiekis paieškoje. Jis yra tokio pat duomenų tipo kaip ir „id“ stulpelis.

„PRIMARY KEY (id)“ – nurodo, jog visi identifikatoriai bus nesidubliuojantys. „UNIQUE KEY“ – nurodo, jog „hash“ stulpelyje saugomos reikšmės bus unikalios ir nesidubliuos, taip pat padės greitai atlikti paiešką pagal šį stulpelį.

Kadangi reikiama lentelė jau yra sukurta, tai reikia ją užpildyti reikiamais duomenimis. Tam autorius parašė funkciją naudodamas „CodeIgniter 2“ karkaso teikiamomis funkcijomis. Aplikacijos modelių aplanke buvo sukurta klasė „searched\_phrases\_any\_model.php“ ir patalpinta į „/home/xxx.eu/public\_html/app/application/models/database“. Joje pridėta nauja funkcija:

```
<?php
class Searched_phrases_any_model extends CI_Model {
    private $table_name = 'searched_phrases_any';
    function __construct() {
        parent::__construct();
        $this->load->database();
    }
    public function add_phrase($phrase) {
        $encoded_phrase = md5($phrase);
        $sql = "INSERT INTO `".$this->table_name."` (`hash`, `phrase`)
VALUES (?, ?) ON DUPLICATE KEY UPDATE times = times + 1;";
        $this->db->query($sql, array($encoded_phrase, $phrase));
    }
    public function get_paged($limit, $page) {
        $this->db->select('phrase')->
        from($this->table_name)->order_by('times', 'desc')->
        limit($limit, $limit * ($page - 1));
        $query = $this->db->get();
        $phrases = array();
    }
}
```

```

        foreach ($query->result() as $row)
            $phrases [] .= $row->phrase;
        return $phrases;
    }
    public function get_size() {
        return $this->db->count_all_results($this->table_name);
    }
}
?>

```

Iš pradžių funkcijai perduota užklausa yra „\$phrase“ yra užkoduojama „MD5“ algoritmu, kad būtų naudojama kaip maišos kodas. Toliau sudaromas masyvas su duomenimis: maišos kodu ir paieškos fraze. Toliau naudojant „CodeIgniter 2“ karkase esančią biblioteką „db“ yra įdedami duomenys į duomenų bazės lentelę. Kadangi numatytasis „times“ stulpelio dydis yra „1“, tai pirmą kartą įdedant naują eilutę papildomai jo nurodyti nereikia.

Jeigu pasirodo, jog tokia eilutė lentelėje jau yra, tai „\$this -> db -> affected\_rows()“ gražinamas rezultatas yra „-1“. Tokiu atveju pagal maišos kodą surastose eilutėse yra padidinama 1 „times“ laukelio reikšmė. Taip yra skaičiuojamas frazės paieškų kiekis. Autorius nori pastebėti, jog nuo „SQL“ injekcijų automatiškai apsaugo „CodeIgniter 2“ karkase esantis mechanizmas, todėl papildomai to daryti nereikia.

Taip pat yra sukuriama funkcija „get\_paged“ skirta gauti išsaugotas paieškos frazes, pritaikant limitus ir puslapio numerį. Visi rezultatai yra rikiuojami pagal „times“ stulpelį pradedant nuo daugiausiai kartų ieškotos frazės, kadangi populiariausios frazės yra svarbiausios ir jos turi turėti pirmenybę. Jeigu būtų viršytas didžiausias leidžiamas svetainės žemėlapių skaičius, tai mažiausiai kartų ieškotos frazės būtų neberodomas, kas turėtų mažiausią įtaką, nei neberodymas puslapių su populiariausiomis frazėmis.

Ne ką mažiau svarbu gauti ir duomenų bazėje išsaugotų paieškos frazių skaičių, kadangi pagal jį bus nustatomas svetainės žemėlapių rodyklėje esančių žemėlapių skaičius. Tam bus naudojama funkcija „get\_size()“, kuri ir gražins visų frazių, esančių duomenų bazės lentelėje skaičių.

Kai darbas yra atliktas su paieškos frazių įrašymu į duomenų bazę, reikia paduoti duomenis įrašymui tik tada, kai duomenų bazėje buvo rastas rezultatas. Kadangi visos paieškos daromos per „Sphinx“ paieškos serverį, tai paieškos frazės įrašymas turi būti įvykdomas tada, kai šis serveris grąžina nors vieną rezultatą. Tai galima patikrinti

„sphinx\_song\_search.php“. Paieškos algoritmo apačioje pridodant papildomą tikrinimą:

```
if ($result["total"] > 0) {  
    $this -> load -> model("database/searched_phrases_exact_model", "phrases");  
    $this -> phrases -> add_phrase($query);  
}
```

„\$result["total"]“ - yra saugomas frazės paieškos rezultatų skaičius, jei jis yra didesnis už 0, tai vadinasi, kad paieškos serveris grąžino rezultatų ir paieškos frazę galima įrašyti į lentelę.

#### 4.4. Svetainės žemėlapių generavimas iš paieškos užklausų

„Google“ paieškai pateikiami puslapiai turi būti suformuoti iš svetainės naudotojų atliktų paieškų. Tam darbui atlikti autorius sukūrė „/home/xxx.eu/public\_html/app/application/controllers“ direktorijoje klasę „sitemaps.php“. Kadangi duomenų bazėje nuorodų į puslapius su dainomis yra per 2 milijonus, tai ir paieškų skaičius bus taip pat labai didelis. Todėl reikia sudaryti ne tik svetainės žemėlapius, bet ir svetainės žemėlapių rodyklę, kaip aprašyta 3.1.4 skyriuje.

```
class Sitemaps extends CI_Controller {  
    private $limit = 50000;  
    public function __construct() {  
        parent::__construct();  
        $this -> load -> model("database/searched_phrases_any_model",  
            "searched_phrases");  
    }  
    public function google_index() {  
        $total_phrases = $this -> searched_phrases -> get_size();  
        $total_sitemaps = ceil($total_phrases / $this -> limit);  
        $data = array(  
            'total_sitemaps' => $total_sitemaps,  
            'site_name' => $this -> config -> item("base_url")  
        );  
        $this -> load -> view('sitemaps/sitemap_index_view.php', $data);  
    }  
    public function google($page) {  
        $phrases = $this -> searched_phrases -> get_paged($this -> limit, $page);  
        $data = array(  
            'phrases' => $phrases,  
            'site_name' => $this -> config -> item("base_url")  
        );  
    }  
}
```

```

    );
    $this->load->view('sitemaps/sitemap_view.php', $data);
}
}

```

Kadangi svetainės žemėlapyje negali būti daugiau nei 50000 adresų, tai todėl nustatomas limitas, pagal kurį bus skaičiuojamas svetainės žemėlapių skaičius.

Toliau kontroleryje aprašoma funkcija „google\_index“, kurios pagalba yra rodoma svetainės žemėlapių rodyklė. Kiek rodyklėje bus rodoma žemėlapių yra nustatoma padalinant sėkmingų paieškos frazių lentelėje esančių eilučių skaičių iš žemėlapių adresų limitu. Šis dalmuo apvalinamas į viršų. Jei gautas dalmuo bus 2,1 tai galutinis rezultatas – 3. Apskaičiuotas žemėlapių skaičius žemėlapių rodyklėje ir svetainės pavadinimas yra perduodamas atvaizdavimui. Specialiai tam skirtas failas „sitemap\_index\_view.php“ yra patalpintas „/home/xxx.eu/public\_html/app/application/views/sitemaps“ kataloge.

```

<?php
header('Content-Type: text/xml');
echo "<?xml version = \"1.0\" encoding = \"UTF-8\" ?>\n\r";
?>
<sitemapindex xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">
    <?php
    for ($i = 0; $i < $total_sitemaps; $i++) {
        echo "<sitemap>\n\r";
        echo "<loc>" , $site_name , "/sitemaps/google-sitemap-" ,
        $i + 1 , ".xml</loc>\n\r";
        echo "<lastmod>" , date("Y-m-d") , "</lastmod>\n\r";
        echo "</sitemap>\n\r";
    }
    ?>
</sitemapindex>

```

Šis failas skirtas atvaizduoti gautus rezultatus. Jis suformuojamas, pagal pateiktas reikšmes iš kontrolerio. „header('Content-Type: text/xml')“ pagalba yra nurodoma, jog serveryje suformuoto failo tipas bus „XML“. Toliau naudojamos žymos, kurios buvo aprašytos 3.1.4 skyriuje. Toliau yra sukamas ciklas ir formuojami žemėlapių adresai tokios konstrukcijos: „http://xxx.eu/sitemaps/google-sitemap-1.xml“. Šioje konstrukcijoje esančio žemėlapių failo prievardis yra „.xml“, kad veiktų tokia konstrukcija reikia papildomai apsirašyti kelias

„CodeIgniter“ konfigūraciniame faile „routes.php“ esančiame „/home/xxx.eu/public\_html/app/application/config“ aplanke. Tai yra padaroma pridėdant eilutę:

```
$route['sitemaps/google-sitemap-(:num).xml'] = 'sitemaps/google/$1';
```

Ji nurodo, jog visi „http://xxx.eu/sitemaps/google-sitemap-\*.xml“ tipo adresai bus nukreipiami į tam tikrą kontrolerį. Tokiu būdu yra imituojamas statinis „XML“ tipo failas.

Taip pat yra patartina pridėti dar vieną eilutę, kurios pagalba svetainės žemėlapiu rodyklės adresas „http://xxx.eu/sitemaps/google-index.xml“ būtų nukreipiamas į tam tikrą kontrolerį.

```
$route['sitemaps/google-index.xml'] = 'sitemaps/google_index';
```

Šie du atvejai yra analogiški „htaaccess“ byloje esančių komandų naudojimui aprašytame 3.1.3 skyriuje. Tik visą darbą atlieka „Codeigniter“ karkasas, kas padaro darbą daug paprastesnį.

Taigi toliau pagrindiniame žemėlapių kontroleryje esanti funkcija „google(\$page)“ yra atsakinga už svetainės žemėlapių sukūrimą. Šiai funkcijai yra paduodamas puslapio numerio parametras, pagal kurį nustatomas, kuris žemėlapiu puslapis bus rodomas. Po to duomenys yra paimami iš duomenų bazės ir perduodami „sitemap\_view.php“ failui atsakingam už rezultatų atvaizdavimą esantį „/home/xxx.eu/public\_html/app/application/views/sitemaps“.

```
<?php
header('Content-Type: text/xml');
echo "<?xml version = \"1.0\" encoding = \"UTF-8\" ?>\n\r";
?>
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">
    <?php
        foreach ($phrases as $phrase) {
            echo '<url>';
            echo '<loc>' , $site_name , "/search/mp3/1/" , str_replace(" ", "-",
                $phrase) , ".html</loc>";
            echo '<changefreq>weekly</changefreq>';
            echo '<lastmod>' , date("Y-m-d") , '</lastmod>';
            echo '<priority>0.6</priority>';
            echo "</url>";
        }
    ?>
</urlset>
```

Šio failo pagalba yra atvaizduojami adresai sudaryti iš paieškos frazių išsaugotų duomenų bazėje.

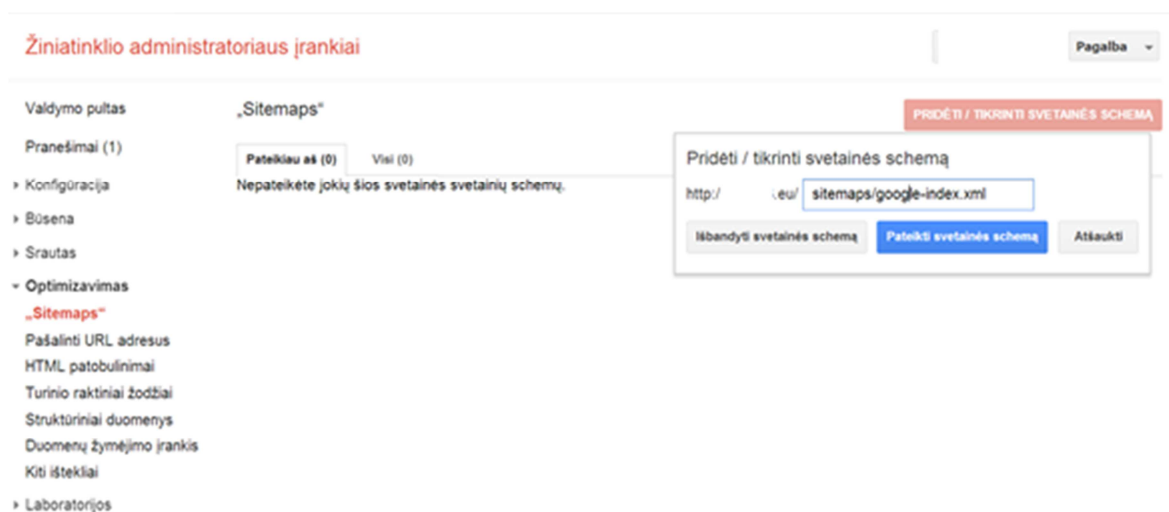
#### 4.5. Svetainės žemėlapių pateikimas „Google“ sistemai

Svetainės žemėlapiai gali būti pateikiami kelias būdais. Vienas iš jų yra pateikti žemėlapių adresą per „robots.txt“ bylą. Joje reikia pridėti tokias eilutes:

```
User-agent: googlebot  
Sitemap: http://xxx.eu/sitemaps/google-index.xml
```

Šis būdas nėra blogas, tačiau „Google“ robotas gali apsilankyti ir perskaityti šią bylą tik už vienos ar dviejų savaitių. O tai gana ilgas laiko tarpas.

Svetainės žemėlapių galima pateikti ir kitu būdu, kuriuo svetainės žemėlapis bus perskaitytas greičiau. Tai galima padaryti naudojantis „Google Webmaster Tools“ įrankio pagalba. Šio įrankio galimybės yra aprašytos 2.1 skyriuje. Prisijungus šiame įrankyje kairėje pusėje yra optimizavimo skiltis ir joje „Sitemaps“ meniu punktas (žr.: 4 pav.).



Pav. 4: Svetainės „xxx.eu“ žemėlapių pateikimas „Google“ sistemai

Dešinėje pusėje yra svetainės schemas pridėjimo mygtukas. Jį paspaudus galima nurodyti svetainės žemėlapių ar žemėlapių rodyklės adresą. Šiai svetainei yra naudojamas žemėlapių rodyklės adresas <http://xxx.eu/sitemaps/google-index.xml>. Tokiu būdu pateikus vieną failą, bus pateikiami ir visi svetainės žemėlapių, taip nereiks kiekvieno pateikinti atskirai.

Pateikti puslapiai yra indeksuojami. Progresą galima pamatyti tame pačiame puslapyje.

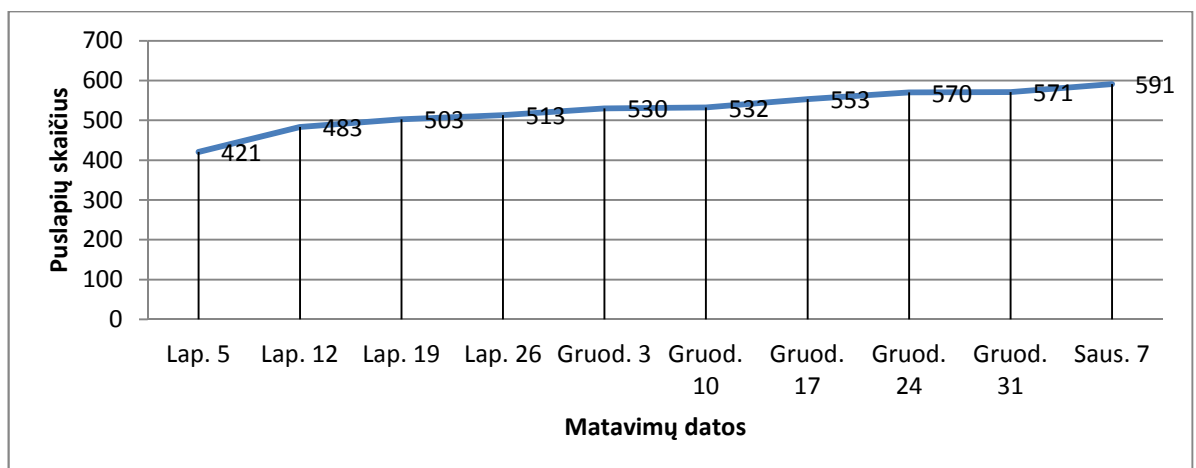


## 5. MUZIKOS PAIEŠKOS SVETAINĖS OPTIMIZAVIMO REZULTATAI

Optimizavus muzikos paieškos svetainę sukuriant puslapių žemėlapi buvo stebimi optimizavimo rezultatai „Google“ paieškoje. Svetainė buvo stebėta 2 mėnesius prieš optimizaciją ir 3 mėnesius po. Stebėjimams buvo naudojami „Google Webmaster Tools“ ir „Google Analytics“ įrankiai. Gauti rezultatai užfiksuoti ataskaitoje.

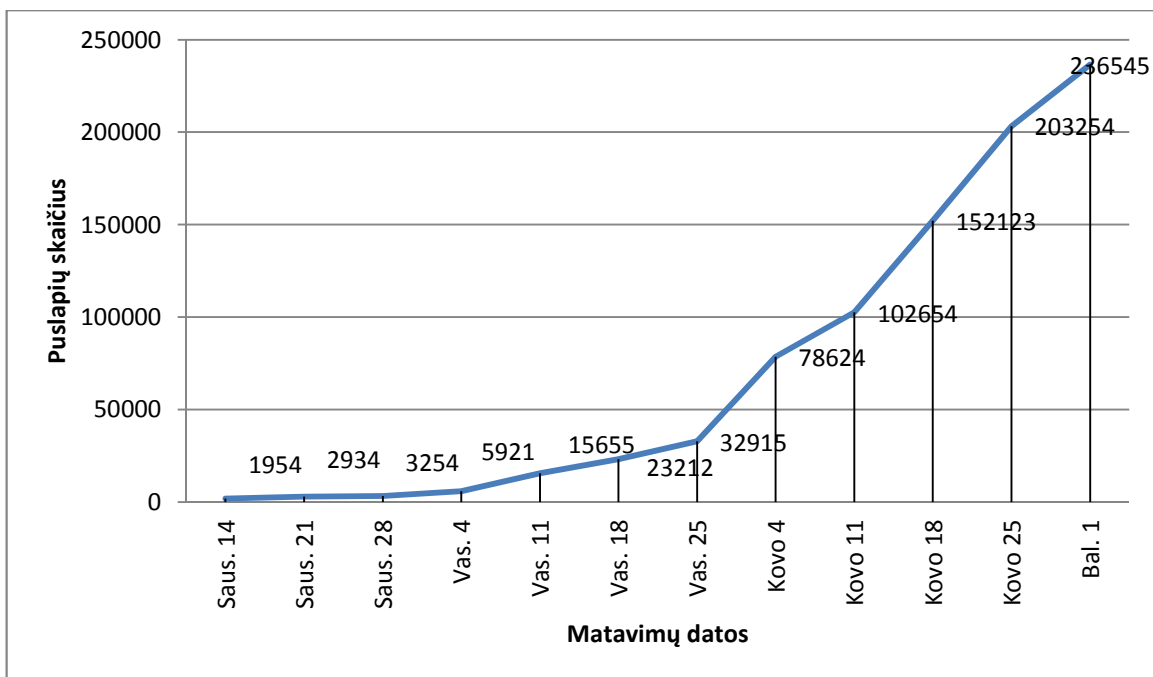
### 5.1. Svetainės puslapių skaičius „Google“ paieškoje

Darbo pradžioje muzikos paieškos svetainė neturėjo savo puslapių žemėlapi. „Google“ paieškos sistemoje suindeksuotų puslapių kiekis buvo labai mažas ir per 2 stebėjimo mėnesius keitėsi nežymiai. Tyrimas buvo vykdomas nuo 2012m. Lapkričio 4 dienos iki 2013m. Sausio 7 dienos. Šiems stebėjimams buvo naudojama komanda „Google“ paieškos laukelyje „site:xxx.eu“, kuri grąžindavo suindeksuotų puslapių skaičių. Kadangi svetainė buvo sukurta jau prieš pusę metų, tai „Google“ sistema jau buvo suindeksavusi kai kuriuos puslapius. Stebėjimai buvo vykdomi kas savaitę (žr.: 5 pav.).



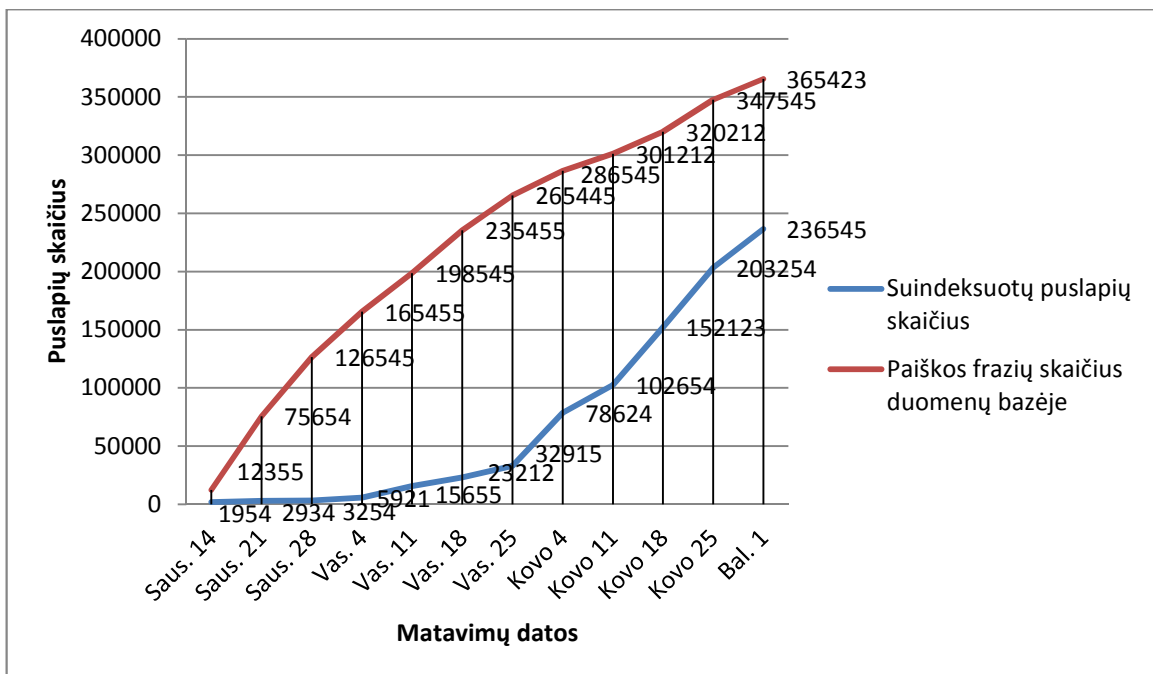
*Pav. 5 Svetainės „xxx.eu“ suindeksuotų puslapių skaičius prieš svetainės žemėlapio sukūrimą.*

Autoriui sukūrus svetainės žemėlapi generavimą iš paieškos frazių, suindeksuotų puslapių stebėjimai buvo vykdomi nuo 2013m. Sausio 13d. iki 2013m. Balandžio 1d. Stebėjimai taip pat buvo vykdomi kas savaitę.



Pav. 6 Svetainės „xxx.eu“ suindeksuotų puslapių skaičius po svetainės žemėlapio sukūrimo.

Palyginus dviejų grafikų duomenis buvo pastebėta, jog žymiai padidėjo svetainės per savaitę suindeksuojamų puslapių skaičius „Google“ sistemoje, kai buvo patalpintas svetainės žemėlapis.



Pav. 7 Svetainės „xxx.eu“ suindeksuotų puslapių skaičiaus priklausomybė nuo paieškos frazių skaičiaus duomenų bazėje.

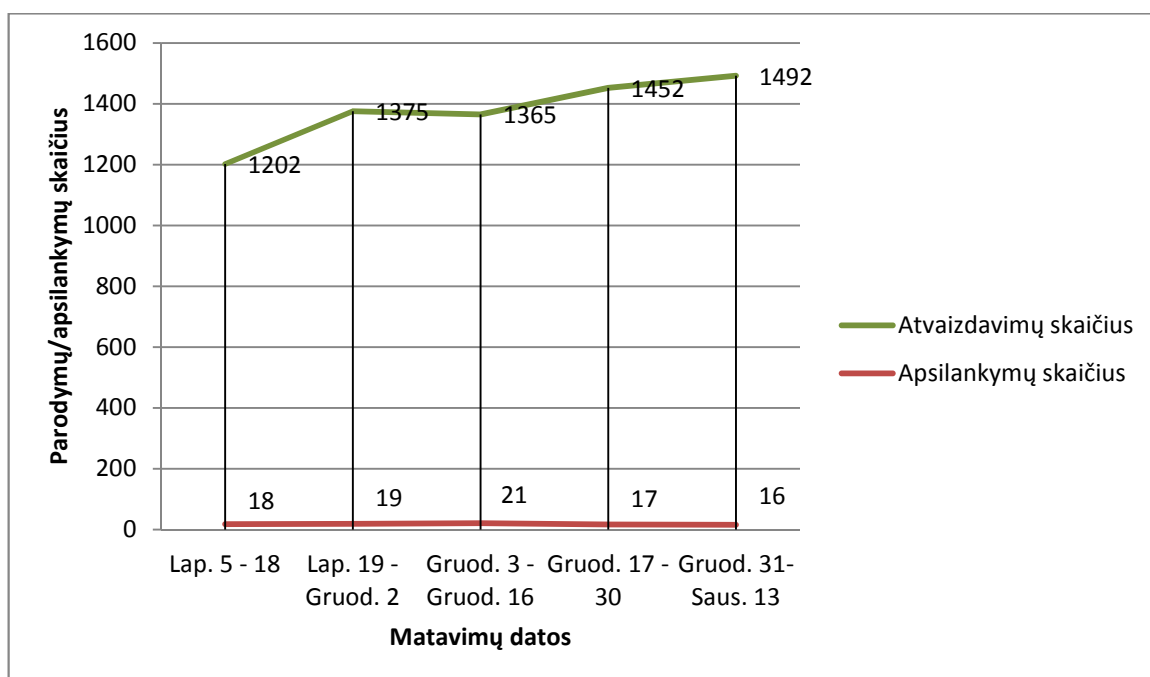
Autorius nori pastebėti, jog muzikos paieškos svetainės paieškos frazių skaičius duomenų bazėje labai sparčiai augo nuo svetainės žemėlapio sukūrimo pradžios. Taip yra dėl to, kad ši

muzikos svetainė turi sukurtą „API“, kuria naudoja kitos svetainės, muzikos paieškai. O kadangi frazių lentelė yra papildoma kiekvieną kartą su sėkminga paieškos fraze (tiek naudojantis „API“, tiek ir per pačios svetainės paiešką), tai paieškos frazių kiekis duomenų bazėje augo daug greičiau.

Palyginus muzikos paieškos svetainės suindeksuotų puslapių skaičių „Google“ paieškos sistemoje su paieškos frazių skaičiumi duomenų bazėje, galima matyti priklausomybę. Todėl galima daryti prielaidą, jog svetainės žemėlapiu pagalbą pateikiamų adresų kiekis daro įtaką „Google“ suindeksuotų puslapių skaičiui. Suindeksuotų svetainės puslapių skaičius yra daug didesnis, nei svetainės, kuri neturėjo žemėlapiu. Tai parodo, jog svetainės žemėlapiu nauda yra akivaizdi.

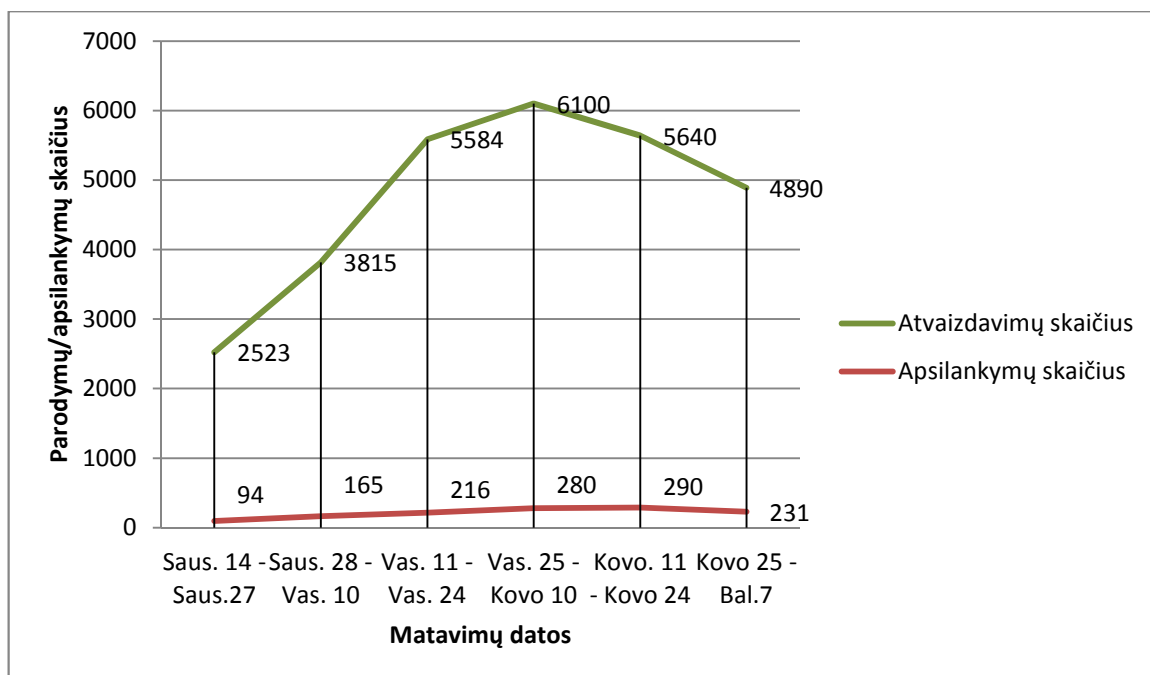
## 5.2. Svetainės puslapių matomumo paieškoje ir apsilankymo rodikliai

Svetainės puslapių parodymai „Google“ paieškos sistemoje ir apsilankymų skaičius taip pat yra labai svarbūs rodikliai. Naudojantis „Google Webmaster Tools“ įrankiu galima nustatyti kiek kartų svetainės puslapiai buvo parodyti „Google“ paieškos rezultatuose. Naudojant „Google Analytics“ įrankį galima sužinoti puslapio lankytojų skaičių iš „Google“ paieškos. 8 pav. esančioje diagramoje, žalia linija yra pavaizduota kiek kartų svetainė buvo atvaizduota paieškos rezultatuose, o raudona – kiek kartų naudotojai apsilankė iš „Google“ paieškos. Parodymų ir lankytojų skaičius du kart per mėnesį, dar nesukūrus svetainės žemėlapiu.



Pav. 8 Svetainės „xxx.eu“ puslapių parodymų skaičius „Google“ paieškoje prieš svetainės žemėlapiu sukūrimą.

Duomenys, surinkti po svetainės žemėlapio sukūrimo, yra atvaizduoti 9 pav. esančioje diagramoje.



Pav. 9 Svetainės „xxx.eu“ puslapių parodymų skaičius „Google“ paieškoje po svetainės žemėlapio sukūrimą.

Autorius nori pastebėti, jog svetainės duomenų bazėje yra laikoma nuorodų ir į nelegalius autorinius kūrinius, kadangi nėra galimybės išsiaiškinti visų surenkamų nuorodų legalumo, todėl puslapiai su nuorodomis, kurie yra aukštose paieškos rezultatų pozicijose yra užblokuojami „Google“ sistemos, gavus įspėjimą apie autoriaus teises pažeidžiantį turinį. Dėl kelių tokių užblokuotų puslapių esančiame 9 pav. grafike, Kovo 11 – Balandžio 25 dienomis yra sumažėjęs atvaizdavimų skaičius.

Palyginus šiuos grafikus galima pastebėti, jog po svetainės žemėlapio sukūrimo, puslapių atvaizdavimų skaičius žymiai padidėjo. Taip pat padidėjo ir apsilankymų skaičius iš „Google“ paieškos sistemos.

## IŠVADOS

Susipažinus su „Google“ paieškos sistema, buvo aptartos taisyklės ir metodai, kuriais vadovaujantis yra siūloma kurti ar tobulinti svetainę, kad ji būtų kuo aukštesnėse paieškos sistemos rezultatų pozicijose. Kadangi ši kompanija teigia, jog paieškos rezultatų rikiavimui naudoja per du šimtus įvairių kriterijų, kurių visų ji neatskleidžia (kompanijos komercinė paslaptis), tai yra svarbu taikyti kuo daugiau metodų ir jų variacijų bei dalintis kitų svetainių kūrėjų patirtimi, norint išsiaiškinti metodus, kurie yra svarbūs, reitingo augimui paieškos sistemoje.

Praktinio darbo metu autorius patobulino sukurtą didelės apimties svetainę, panaudojant išanalizuotas paieškos variklių optimizavimo metodikas. Peržvelgus jos trūkumus, buvo sukurtas bei integruotas paieškos frazių registravimas duomenų bazėje ir svetainės žemėlapių generavimas, panaudojant suregistruotus duomenis. Kadangi svetainėje yra didelės apimties (ją sudaro labai daug puslapių su skirtingu turiniu), tai tokiu būdu buvo siekiama pateikti kuo daugiau svetainės puslapių „Google“ paieškos robotui, taip padedant greičiau suindeksuoti juos. Tokiu optimizavimo metodu buvo padidintas puslapių parodymų skaičius „Google“ paieškos rezultatuose, taip pat ir lankytojų skaičius svetainėje.

Duomenų rinkimui panaudojus „Google Webmaster Tools“ ir „Google Analytics“ įrankius, buvo surinkti duomenys apie paieškos rezultatuose parodytus ir peržiūrėtus puslapius. Palyginus gautus duomenis, buvo nustatyta, jog kuo paieškoje dažniau rodomas puslapis, tuo daugiau potencialių lankytojų jis sulaukia. Žinoma, nereikia pamiršti, jog dar svarbiau yra pozicija paieškos rezultatuose, negu puslapių parodymų skaičius. Kuo aukštesnėje pozicijoje puslapis yra, tuo didesnis yra gaunamas lankytojų srautas, o dėl žemos vietos rezultatų puslapyje dažniausiai yra didelis svetainės puslapių parodymų skaičius, bet ne apsilankymų puslapyje skaičius.

Patobulinus muzikos paieškos svetainės optimizavimą „Google“ paieškos varikliui, tyrimo metu buvo pastebėtas naudotojų skaičiaus augimas, o tai ir buvo vienas iš optimizacijos siekiamų tikslų. Rezultatai parodo, jog svetainės optimizacija turi būti nuolat tikrinama ir tobulinama taikant naujausius puslapių optimizavimo principus paieškos sistemoms.

Optimizavimas paieškos varikliams yra mažai lėšų reikalaujantis procesas palyginus su reklamos pirkimu paieškos rezultatų puslapio viršuje. Kadangi paieškos sistemos yra labai plačiai naudojamos, tai naudotojai, ieškantys informacijos, renkasi puslapius, kurie yra paieškos rezultatų viršuje, o jeigu puslapiai yra gerai optimizuoti paieškos varikliams, tai jie ir yra rodomi viršuje, kur naudotojai geriausiai juos pastebi. Tačiau, norint pasiekti kuo geresnių rezultatų

svetainė turi būti nuolat tobulinama, ieškant spragų svetainėje ar jos optimizacijoje, tam naudojant specialius įrankius.

## SĄVOKŲ APIBRĖŽIMAI IR SANTRUMPŲ SĄRAŠAS

1. **API** – (angl. „Application Programming Interface“) tai sąsaja, kurią suteikia kompiuterinė sistema, biblioteka ar programa tam, kad programuotojas per kitą programą galėtų pasiekti jos funkcionalumą ar apsikeistų su ja duomenimis.
2. **HTML** – hipertekstų ženklinimo kalba, dažniausiai vartojama tinklalapiams rašyti.
3. **MySQL** – reliacinių duomenų bazių valdymo sistemų, palaikanti daugelį naudotojų, dirbanti SQL kalbos pagrindu.
4. **PHP** – (angl. HyperText Processor) skriptų kalba, kurios pagalba kuriami dinaminiai interneto puslapiai.
5. **SEO** – (angl. „Search Engines Optimization“) tinklalapių derinimas, siekiant aukštesnės vietos paieškos sistemų rezultatuose.
6. **W3C** – (angl. „World Wide Web Consortium“) yra konsorciumas leidžiantis programinės įrangos standartus/rekomendacijas žiniatinkliui.
7. **XML** – (angl. „Extensible Markup Language“) universali dokumentų ženklinimo kalba, skirta dokumentų struktūrai aprašyti.

## LITERATŪROS SĄRAŠAS

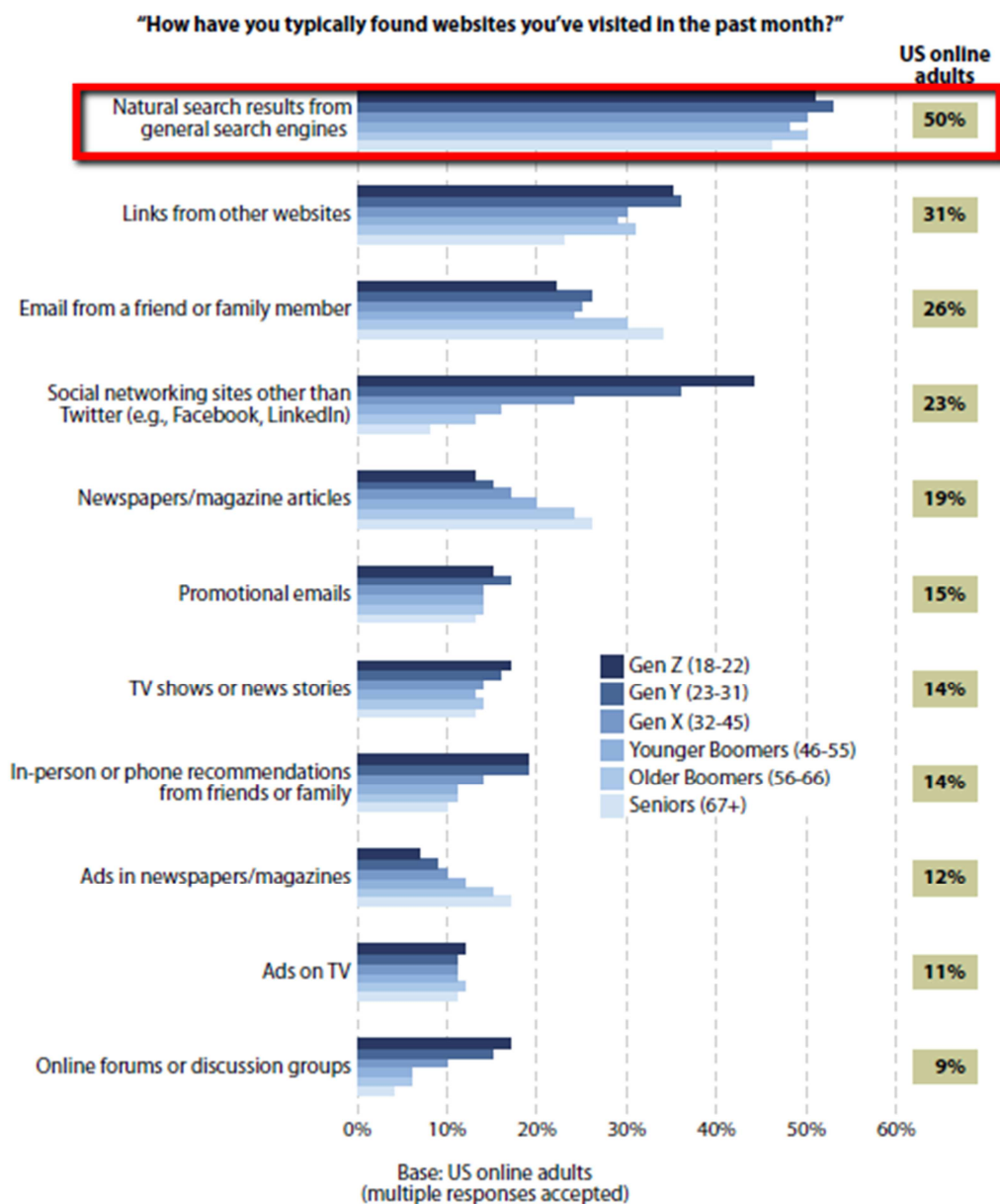
- [BP07] Nancy Blachman, Jerry Peek. How Google Works. Googleguide, 2007.  
[Žiūrėta 2013–03–07]. Prieiga per internetą:  
<[http://www.googleguide.com/google\\_works.html](http://www.googleguide.com/google_works.html)>
- [BP98] Sergey Brin, Lawrence Page. The Anatomy of a Large–Scale Hypertextual Web Search Engine, 1998.  
[Žiūrėta 2013–02–10]. Prieiga per internetą:  
<[http://static.googleusercontent.com/external\\_content/untrusted\\_dlcp/research.google.com/lt//pubs/archive/334.pdf](http://static.googleusercontent.com/external_content/untrusted_dlcp/research.google.com/lt//pubs/archive/334.pdf)>
- [GWC10] Google Webmaster Central. Google’s Search Engine Optimization Starter Guide.  
[Žiūrėta 2013–04–11]. Prieiga per internetą:  
<[http://static.googleusercontent.com/external\\_content/untrusted\\_dlcp/www.google.com/lt//webmasters/docs/search-engine-optimization-starter-guide.pdf](http://static.googleusercontent.com/external_content/untrusted_dlcp/www.google.com/lt//webmasters/docs/search-engine-optimization-starter-guide.pdf)>
- [GWC11] Google Webmaster Tools. Sitemap index file, 2011  
[Žiūrėta 2013–03–29]. Prieiga per internetą:  
<<http://support.google.com/webmasters/bin/answer.py?hl=en&answer=71453>>
- [GWC12] Google Webmaster Central. Googlebot, 2012.  
[Žiūrėta 2013–03–17]. Prieiga per internetą:  
<<http://support.google.com/webmasters/bin/answer.py?hl=en&answer=182072>>
- [GWC12b] Google Webmaster Tools. Meta tags, 2012.  
[Žiūrėta 2013–03–27]. Prieiga per internetą:  
<<http://support.google.com/webmasters/bin/answer.py?hl=en&answer=79812>>
- [GWC12c] Google Webmaster Tools. About the robots.txt file, 2012  
[Žiūrėta 2013–03–27]. Prieiga per internetą:  
<<http://support.google.com/webmasters/bin/answer.py?hl=en&answer=35303>>
- [GWC12d] Google Webmaster Central. About Sitemaps, 2012.  
[Žiūrėta 2013–04–10]. Prieiga per internetą:  
<<http://support.google.com/webmasters/bin/answer.py?hl=en&answer=156184>>



- [GWC12e] Google Webmaster Central. About Sitemaps, 2012.  
[Žiūrėta 2013–04–10]. Prieiga per internetą:  
<<http://support.google.com/webmasters/bin/answer.py?hl=en&answer=183668>>
- [GWC12f] Google Webmaster Tools. URL structure. 2012.  
[Žiūrėta 2013–04–12]. Prieiga per internetą:  
<<http://support.google.com/webmasters/bin/answer.py?hl=en&answer=76329>>
- [GWC12g] Google Webmaster Tools. Link schemes. 2012.  
[Žiūrėta 2013–04–14]. Prieiga per internetą:  
<<http://support.google.com/webmasters/bin/answer.py?hl=en&answer=66356>>
- [GWC13] Google Webmaster Central. Learn how Google discovers, crawls, and serves web pages, 2013.  
[Žiūrėta 2013–03–29]. Prieiga per internetą:  
<<http://support.google.com/webmasters/bin/answer.py?hl=en&answer=70897>>
- [Hol13] Laura Holloway. Choosing the Best Keywords. Vertical Measures. 2013.  
[Žiūrėta 2013–04–06]. Prieiga per internetą:  
<<http://www.verticalmeasures.com/resources/seo-tutorial-videos/choosing-the-best-keywords/>>
- [Lei13] Brandon Leibowitz. W3C Validation not part of Google Search Engine Ranking Factor.  
[Žiūrėta 2013-02-27]. Prieiga per internetą:  
<<http://www.pandia.com/sew/2120-w3c-validation-not-part-of-google-search-engine-ranking-factor.html>>
- [Man10] Modi Mann. 5 Common SEO Mistakes with Web Page Titles, 2010.  
[Žiūrėta 2013–03–20]. Prieiga per internetą:  
<<http://sixrevisions.com/content-strategy/5-common-seo-mistakes-with-web-page-titles/>>

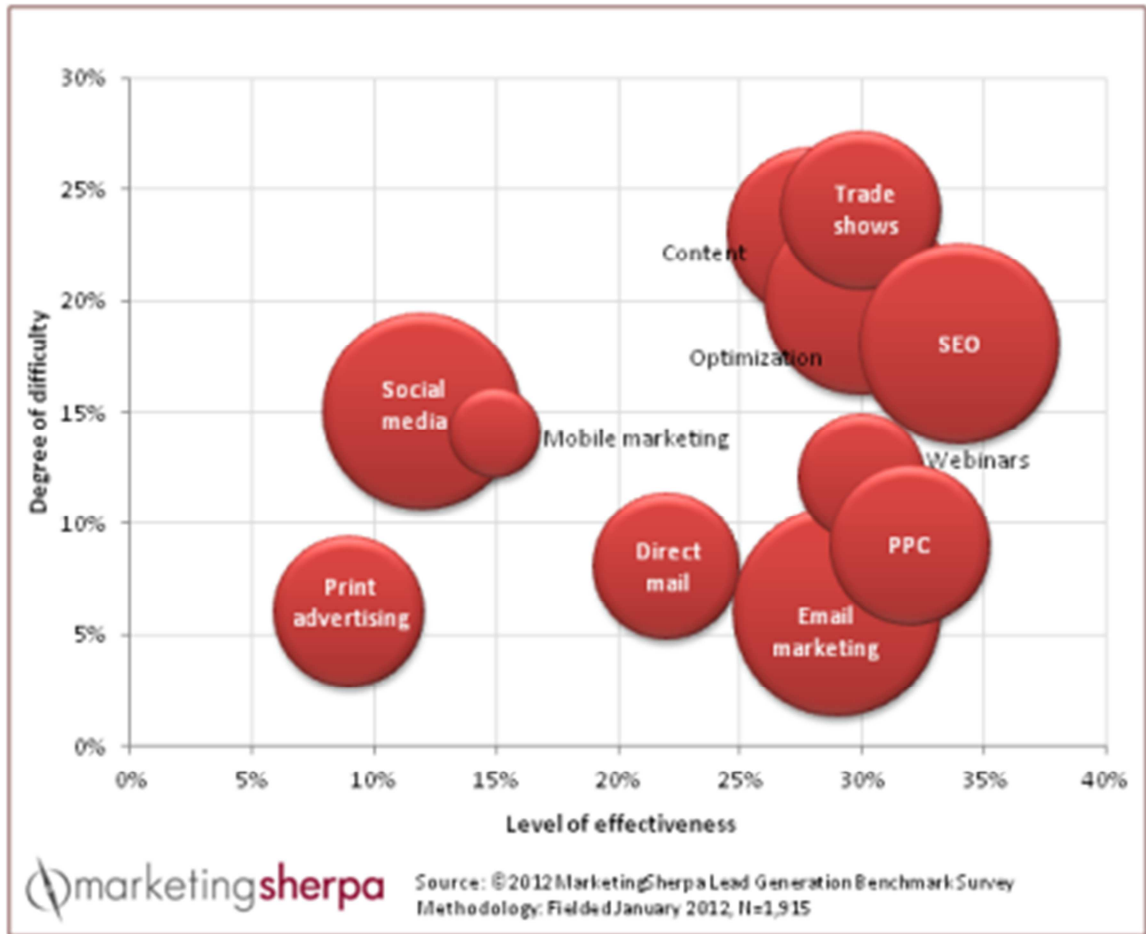
## PRIEDAI

### 1 PRIEDAS. APKLAUSA, KAIP SKIRTINGŲ AMŽIAUS GRUPIŲ ŽMONĖS APLANKO PUSLAPIUS INTERNETE



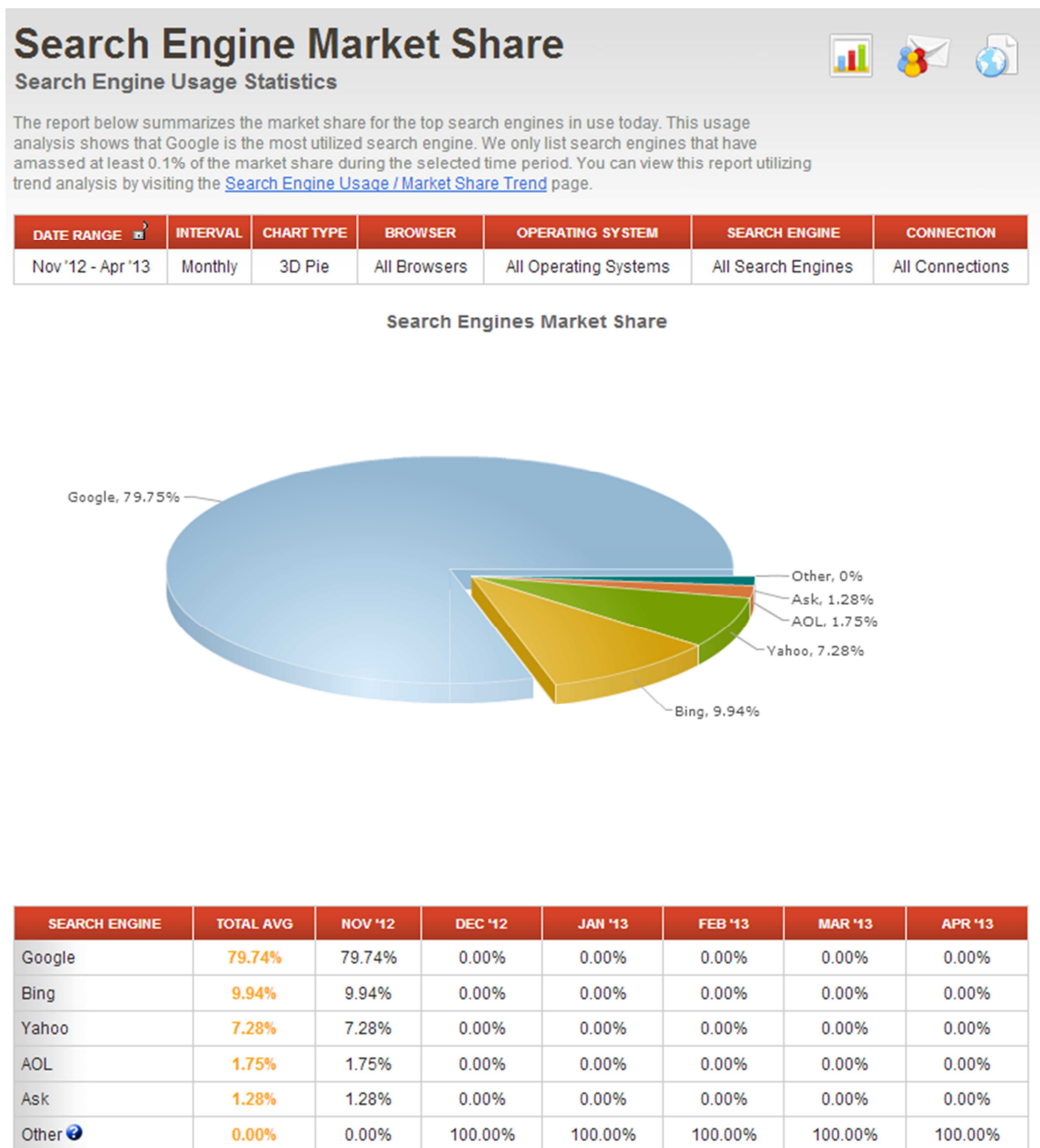
Pav. 10: Kompanijos „Forrester Research“ surinktos statistikos rezultatai. Šioje diagramoje yra matoma, jog Jungtinėse Amerikos Valstijose žmonės svetaines randa naudodamiesi paieškos varikliais.

## 2 PRIEDAS. REKLAMOS RŪŠIŲ EFEKTYVUMAS



Pav. 11: Kompanijos „MarketingSherpa“ 2012m. atlikto reklamos rūšių efektyvumo išvados. Y ašyje – reklamos rūšies įgyvendinimo sudėtingumas, X ašyje – reklamos efektyvumas. Diagramoje apskritimo dydžiai yra proporcingi reklamos rūšies populiarumui.

### 3 PRIEDAS. PAIEŠKOS VARIKLIŲ PASISKIRSTYMAS RINKOJE



*Pav. 12: Kompanijos „Stat Owl“ surinktos 6 mėn. statistikos rezultatai. Ši diagrama rodo, jog daugiausiai rinkos užima kompanijos „Google“ paieška, po jos seka „Bing“, o dar mažiau rinkos užima „Yahoo“ paieškos sistema.*