

**UNIVERZA V LJUBLJANI
FAKULTETA ZA RAČUNALNIŠTVO IN
INFORMATIKO**

Dejan Čokara

**SPLETNI ISKALNIK DNEVNIH MALIC
IN KOSIL V LJUBLJANSKIH
RESTAVRACIJAH Z MOŽNOSTJO
PRIKAZA POTI**

**DIPLOMSKO DELO NA VISOKOŠOLSKEM STROKOVNEM
ŠTUDIJU**

MENTOR: viš. pred. dr. Aleksander Sadikov

Ljubljana, 2009



Št. naloge: 00406/2008

Datum: 15.10.2008

Univerza v Ljubljani, Fakulteta za računalništvo in informatiko izdaja naslednjo nalogo:

Kandidat: **DEJAN ČOKARA**

Naslov: **SPLETNI ISKALNIK DNEVNIH MALIC IN KOSIL V LJUBLJANSKIH
RESTAVRACIJAH Z MOŽNOSTJO PRIKAZA POTI**
**SEARCHING THROUGH LJUBLJANA'S DAILY OFFER OF LUNCHES
WITH THE CAPABILITY OF DISPLAYING THE ROUTE**

Vrsta naloge: Diplomsko delo visokošolskega strokovnega študija

Tematika naloge:

Kandidat naj zasnuje in implementira spletno aplikacijo za iskanje po dnevni ponudbi ljubljanskih restavracij. Aplikacija naj vsebuje vsakodnevno avtomatsko ažuriranje ponudbe s spletnih strani restavracij, ki so podane v različnih formatih in hranjenje le-te v podatkovni bazi za hitrejše iskanje. Poleg samega iskanja po ponudbi naj aplikacija omogoča uporabniku še prikaz poti do izbrane restavracije.

Mentor:


viš. pred. dr. Aleksander Sadikov



Dekan:


prof. dr. Franc Solina

IZJAVA O AVTORSTVU

diplomskega dela

Spodaj podpisani/-a DEJAN ČOKARA,

z vpisno številko 63040202,

sem avtor/-ica diplomskega dela z naslovom:

SPLETNI ISKALNIK DNEVNIH MALIC IN KOSIL V LJUBLJANSKIH
RESTAVRACIJAH Z MOŽNOSTJO PRIKAZAPOTI

S svojim podpisom zagotavljam, da:

- sem diplomsko delo izdelal/-a samostojno pod mentorstvom (naziv, ime in priimek)

viš. pred. dr. Aleksander Sadikov

in somentorstvom (naziv, ime in priimek)

- so elektronska oblika diplomskega dela, naslov (slov., angl.), povzetek (slov., angl.) ter ključne besede (slov., angl.) identični s tiskano obliko diplomskega dela
- soglašam z javno objavo elektronske oblike diplomskega dela v zbirki »Dela FRI«.

V Ljubljani, dne _____

Podpis avtorja/-ice:

Zahvala

Zahvaljujem se viš. pred. dr. Aleksandru Sadikovu za prevzeto mentorstvo, za nasvete in predvsem za potrpežljivost, ki jo je imel pri nastajanju tega diplomskega dela.

Posebna zahvala velja mojim staršem, sestri, njenemu možu, (nečakinji in nečaku) ter puncu, ki so me moralno podpirali skozi študij, me spravljali v dobro voljo med pisanjem diplome ter z mano trepetali pred izpiti.

Zahvala gre tudi sošolcu in sodelavcu Mihi Kollerju, ki mi je ob kavah pomagal s svojimi nasveti pri izdelavi diplomske naloge ter sošolcu Dušanu Šmitranu, ki mi je prav tako pomagal z nasveti.

Kazalo vsebine

Povzetek	1
Abstract.....	2
1 Uvod.....	3
2 Opis uporabljenih orodij in tehnologij	4
2.1 Razporejevalnik opravil	4
2.2 Podatkovna baza MySQL	4
2.3 Skriptni jezik Python	5
2.4 Skriptni jezik PHP	5
2.5 Dreamweaver	6
2.6 Spletni strežnik Apache	7
2.7 XAMPP.....	7
3 Prikaz komunikacije posameznih komponent.....	9
4 Podroben opis in funkcionalnost posameznih komponent ter njihova problematika.....	10
4.1 Zagotavljanje podatkov podatkovni bazi MySQL.....	10
4.1.1 Dnevno zaganjanje skripte v Python-u.....	10
4.1.2 Razčlenjevanje spletnih strani izbranih restavracij	10
4.1.3 Podatkovna baza MySQL s podatki o restavracijah.....	18
4.1.4 Obveščanje skrbnika sistema.....	19
4.2 Spletni iskalnik v PHP-ju.....	20
4.2.1 Vnos zelene jedi in poizvedovanje po podatkovni bazi	20
4.2.2 Prikaz zadetkov.....	23
4.3 Prikazovalnik poti	25
5 Prikaz delovanja spletnega iskalnika.....	28
6 Sklep.....	31
7 Viri	32

Kazalo slik

Slika 1 Xampp	8
Slika 2 Komunikacija komponent	9
Slika 3 Primer HTML kode po uporabi Beautiful Soup-a.....	12
Slika 4 Primer strani iz katere izvlečem ponudbo kosil in malic	13
Slika 5 Primer urejene značke	14
Slika 6 Primer neurejene značke	15
Slika 7 Primer prevedene spletne strani iz PDF datoteke v tekstovno datoteko	16
Slika 8 Primer neuspešne pretvorbe spletne strani v Flash-u v tekstovno datoteko [12]	17
Slika 9 Tabela s podatki o naslovih restavracije	18
Slika 10 Tabela s podatki o vsebini malic in kosil	18
Slika 11 Primer rezultatov razčlenjevanja	20
Slika 12 Prikaz napačnega izpisa rezultata poizvedbe	24
Slika 13 Prikaz zelenega izpisa rezultata poizvedbe	24
Slika 14 Primer vnosa naslova za vsako prikazano restavracijo	26
Slika 15 Prikaz začetne spletne strani iskalnika	28
Slika 16 Prikaz restavracij, ki ponujajo uporabnikovo zeleno jed	29
Slika 17 Slika Podroben prikaz ponujenih današnjih malic in kosil	29
Slika 18 Slika Prikaz poti od uporabnikovega naslova do naslova restavracije.....	30

Povzetek

Za diplomsko nalogo sem izdelal spletni iskalnik malic in kosil, ki vsebuje nekaj ljubljanskih restavracij ter njihovo ponudbo. Namen izdelave spletnega iskalnika je uporabniku skrajšati čas iskanja zelene jedi po posameznih restavracijah. Ko uporabnik najde zeleno jed v iskalniku, mu ta ponudi prikaz poti do restavracije, kjer je iskana jed na jedilniku.

Pred začetkom izdelave sem se moral odločiti katera razvojna orodja in tehnologije bom uporabljal pri izdelavi diplomske naloge. Seveda vseh nisem mogel predvideti, tako je bil seznam uporabljenih orodij daljši od načrtovanega. Prvotna zamisel spletnega iskalnika malic in kosil je bila, da se bo le-ta izvajal na spletnem strežniku. Zaradi raznih prilagajanj tehnologij in podpiranju le-teh na spletnem strežniku, sem se odločil izvajati spletni iskalnik na lokalnem strežniku.

Orodja in tehnologije s katerimi sem si pomagal pri izdelavi diplomske naloge so opisana v prvem delu. Po opisu orodij sledi prikaz komunikacije posameznih komponent in njihove soodvisnosti. V nadaljevanju sledi podroben opis izdelave posameznih komponent in njihova problematika. Opis izdelave posameznih komponent zajema, zagotavljanje podatkov podatkovni bazi MySQL z dnevnim zaganjanjem skripte v Pythonu, ki razčlenjuje spletne strani restavracij in obvešča skrbnika preko elektronske pošte o uspešnosti razčlenjevanja. Ena od komponent je spletni iskalnik, ki uporabniku zeleno jed poišče v podatkovni bazi ter prikaže restavracije v katerih se ponuja, zadnja od komponent pa je spletni prikaz poti, ki uporabniku na njegovo zahtevo prikaže pot od vnesenega naslova uporabnika do naslova restavracije. Pri posameznih komponentah so predstavljeni problemi, ki so se pojavljali pri izdelavi diplomske naloge, podani so tudi njihovi primeri ter način rešitve.

Uspešno rešeni predstavljeni problemi so prikazani na primeru uporabe spletnega iskalnika malic in kosil v zadnjem delu.

Ključne besede: restavracija, podatkovna baza, spletni iskalnik, prikaz poti, razčlenjevanje spletnih strani

Abstract

For thesis, I created a web search engine for brunches and lunches, which includes several restaurants in Ljubljana and their offer. The purpose of making a web search engine is to short the time that each user spends on searching dishes by individual restaurants. When a user finds a desired dish in the web search, it offers him to display the route to the restaurant, where the desired dish is served.

Before the beginning of making my thesis, I had to decide which development tools and technologies I will be using in creating the thesis. Of course, I could not foresee all of them, so the list of used tools was longer than planned. The primary idea of the web search engine for brunches and lunches was to be implemented on the web server. Because of the various adjustments of technologies and supports for them on their web server, I decided to implement a web search on the local server.

Tools and technologies, with which I have helped myself in the manufacture of thesis, are described in the first part. After the description of tools there follow the demonstration of communication for individual components and their co-dependence. The following is a detailed description of manufacture of individual components and their problems. The description of manufacture the individual components contents, the assuring of data to MySQL database with daily running scripts in Python, which is parsing web pages of restaurants and informs an administrator via e-mail on the performance of parsing. One of the components is a web search engine that searches for user-desired dish in the database, and displays the restaurants in which the dish is offered. The last of the components is a web display of the route, which shows the user the way from the address he entered to the restaurant he has chosen. In individual components, problems that have occurred in creating the thesis are presented, and ways of solution and examples for those problems are given.

Successfully solved expressed problems are shown in the case of use of web search engine for brunches and lunches in the last part.

Keywords: restaurant, database, web search engine, routing, website parse

1 Uvod

Včasih, ko si zaželim nekaj pojesti, moramo pregledati celoten splet. Dostikrat nas po iskanju in pregledovanju posameznih spletnih strani restavracij mine lakota ali celo obupamo nad iskanjem. Prav zaradi tega je prišla zamisel o spletnem iskalniku malic in kosil, ki potratno delo opravi namesto nas.

Diplomska naloga predstavlja zamisel o spletnem iskalniku malic in kosil, po različnih restavracijah, ki bi namesto nas opravil zamudno delo brskanja po spletu. Googlov iskalnik je služil kot osnova za prikaz in delovanje predstavljenega spletnega iskalnika. Uporabnik v spletni iskalnik vpiše želeno jed in kot rezultat se prikaže: ime restavracije, delni zapis jedi, ki jo je uporabnik vpisal, povezavo do spletne strani restavracije na kateri je podrobnejši opis malic in kosil ter možnost prikaza poti do restavracije od naslova, ki ga uporabnik vpiše.

V prvem delu diplomske naloge sem predstavil orodja in tehnologije, ki sem jih uporabljal pri izdelavi diplomske naloge. Z nekaterimi sem se na novo seznanil, pri nekaterih pa sem okrepil svoje znanje. V drugem delu je prikazana medsebojna komunikacija posameznih komponent, odvisnost med njimi in zaporedje njihovega delovanja.

V nadaljevanju sledi opis razvoja posameznih komponent. Podrobneje sem predstavil izbrane probleme in rešitve, s katerimi sem se spopadal pri izdelavi diplomske naloge. Primeri uspešno odpravljenih problemov so prikazani na spletnem iskalniku malic in kosil v zadnjem delu.

2 Opis uporabljenih orodij in tehnologij

V tem poglavju predstavljam orodja, s katerimi sem si pomagal pri izdelavi diplomske naloge.

2.1 Razporejevalnik opravil

Z Razporejevalnikom opravil [1] lahko določimo nekemu opravilu (programu, orodju, paketni datoteki, ...), da se zažene ob nastavljenem času. Zagon posameznega opravila je lahko enkratno ali periodičen, na primer dnevno, tedenski, mesečni. V Razporejevalniku opravil nastavimo časovni zagon tistim opravilom, na katera ne smemo pozabiti in tistim, ki se izvajajo v času naše odsotnosti.

2.2 Podatkovna baza MySQL

Podatkovna baza MySQL [2] je najbolj razširjena odprtokodna relacijska podatkovna baza in ima več kot 100 milijonov namestitev. SQL v imenu MySQL je kratica za »Structured Query Language« in je standardiziran jezik, ki se uporablja za dostop do podatkovnih baz. Določen je z ANSI/ISO SQL standardom. MySQL se najpogosteje uporablja v navezi s skriptnimi jeziki PHP, Perl in Python. PHP in MySQL sta osnovni komponenti pri večini sistemov za upravljanje spletnih vsebin. Izvorna koda projekta MySQL je dostopna pod pogoji licence GNU, do nje pa lahko dostopamo tudi preko različnih komercialnih licenčnih različic.

Podatkovno bazo MySQL je mogoče uporabljati na različnih operacijskih sistemih. Nekateri od teh (FreeBSD, Linux, MacOS X, NetBSD, Solaris, Windows 95, 98, ME, NT, 2000, XP in Windows Vista) so znani večini uporabnikov.

Za administracijo podatkovne baze MySQL prek spleta se dostikrat uporablja orodje imenovano phpMyAdmin, napisano v skriptnem jeziku PHP. Za njegovo delovanje potrebujemo predhodno nameščene programe Apache, PHP in MySQL.

2.3 Skriptni jezik Python

Python [3] je dinamično objektno usmerjen programski jezik, ki se lahko uporablja za številne vrste razvoja programske opreme. Je odprtokodni skriptni jezik, kar pomeni, da je prost za uporabo kot tudi za tržne proizvode. Nudi močno podporo za povezovanje z drugimi jeziki in orodji. Zaradi obsežnih standardnih knjižnic se ga je možno naučiti v nekaj dneh.

Nekaj Python-ovih ključnih značilnosti:

- zelo jasna in čitljiva sintaksa,
- je prijazen uporabnikom in enostaven za učenje,
- močna samo povpraševalna zmogljivost,
- je predmetno usmerjen,
- naravni izrazi procedurene kode,
- popolna modularnost,
- visoka stopnja dinamičnih podatkovnih vrst,
- obsežne standardne knjižnice,
- razširitveni moduli so lahko napisani v C, C++, v Javi za Jython ali v .NET jeziku za IronPython, ...

Python dopušča hitro pisanje kode, ki jo potrebujemo. Za to se lahko zahvalimo visoko optimiziranemu Bajtovem prevajalniku in podpornim knjižnicam, ki Python-ovo kodo izvajajo dovolj hitro za večino aplikacij. Python deluje na Windows, Linux/Unix, Mac OS X, OS/2, Amiga, Palm dlančnikih in mobilnih telefonih Nokia.

2.4 Skriptni jezik PHP

PHP [4] se kot odprtokodni skriptni programski jezik uporablja za izdelavo dinamičnih spletnih strani. Izvaja se lahko na strani strežnika ali pa se zažene kot skripta v ukaznem načinu.

Pri izvajanju na spletnem strežniku jemlje izvorno kodo za vhod in generira spletno stran (HTML kodo) kot izhod. Trenutno sta v uporabi dve večji različici 5.2.x in 4.4.x, razvija pa se tudi različica 6.0, ki bo odpravila precej znanih težav prejšnjih verzij programskega jezika.

Glavne prednosti skriptnega jezika PHP so:

- nizki stroški
 - o PHP je na voljo zastonj pod pogoji PHP licence,
- enostavna uporaba
 - o veliko primerov je objavljeno na spletu, tako da ga lahko prilagodimo svojim potrebam,
- popularnost
 - o skripta PHP se izvaja na več kot 24% vseh spletnih strežnikov,
- ima veliko podporo spletne skupnosti,
- je hiter,
- je stabilen, ...

Podobni plačljivi produkti so Microsoft Active Server Pages, Macromedia ColdFusion in SUNov Java Server Pages.

2.5 Dreamweaver

Dreamweaver [5] je program za ustvarjanje spletnih strani, ki omogoča:

- vizualno načrtovanje in urejanje, ter neposredno pisanje HTML kode,
- ustvarjanje dinamičnih vsebin spletnih strani, ki uporabljajo strežnik v jezikih kot so ASP, ASP.NET, Java, PHP, Cold Fuzion, Markup Language (CFML),
- da se pri njegovi uporabi čas izdelave spletnih strani znatno skrajša zaradi naprednih tehnik enostavne uporabe.

Dreamweaver je zelo prilagodljiv program, ki združuje veliko možnosti oblikovanja in načrtovanja programov. Prav zaradi tega je zelo uporabno orodje tako za začetnike in tiste z malo manj izkušnjami, kot za izkušene uporabnike.

Dreamweaver na prvi pogled daje vtis prenatrpanosti z odvečnimi podrobnostmi, vendar kakor hitro začnemo z resnim izdelovanjem spletne strani, ugotovimo, da postavitve gumba na določenem mestu ni slučajnost.

Dreamweaver ponuja vzporedno pisanje HTML kode ter spreminjanje postavitve strani. Poleg tega nudi osebno oblikovanje HTML in CSS stilov. Velika prednost tega programa je tudi v

tem, da je združljiv s številnimi proizvajalci drugih programov. V brskalniku omogoča predogled spletne strani, navadno v Internet Explorer-ju.

2.6 Spletni strežnik Apache

Pri Apache [6,7] HTTP Server Project so si prizadevali razviti in vzdrževati odprto kodo HTTP strežnika za sodobne operacijske sisteme, vključno z UNIX in Windows NT. Cilj projekta je bil zagotoviti varen, učinkovit in nadgradljiv strežnik, ki podpira HTTP ter omogoča storitve v sinhronizaciji s sedanjimi HTTP standardi. Od aprila 1996, je strežnik Apache najpopularnejši spletni strežnik na svetu.

Nekaj značilnosti strežnika Apache HTTP:

- je močan, prilagodljiv in skladen s HTTP/1.1,
- izvaja najnovejše protokole, vključno HTTP/1.1 (RFC2616),
- je konfigurabilen in ima možnost razširitve z drugimi moduli,
- lahko ga prilagodimo s pisanjem 'Modulov' z Apache modul API,
- omogoča dostop do celotne izvorne kode vključno z neomejeno licenco,
- izvaja se na Windows NT/9x, Netware 5.x in novejši, OS/2, na večini različic Unix, kot tudi na številnih drugih operacijskih sistemih,
- neprestano se razvija,
- spodbuja povratne informacije od uporabnikov skozi nove ideje, poročila programskih napak in popravkov.

Ime "Apache" je bilo izbrano iz spoštovanja do ameriškega indijskega plemena Apache, dobro znano po svojih odličnih spretnostih v strategiji vojskovanja in njihovi neizčrpni vzdržljivosti.

Sekundarno in bolj popularno ime (čeprav napačno) je "A PAtCHy server", ki temelji že na obstoječi kodi in vrsti popravkov.

2.7 XAMPP

XAMPP [8] ima vse funkcije AMPP (Apache MySQL, PHP, Perl) paketa, ki je ena od redkih nekomercialnih AMPP nakopičenih vmesnih oprem, na voljo v Linux-u. S tesno povezavo,

XAMPP omogoča od delovanja funkcij na osebnih domačih straneh, do delovanja vseh funkcij produkcijske strani, čeprav samo za namene razvoja. Zaradi varnostnih vprašanj XAMPP ni mišljen za uporabo kot produkcijski strežnik.

XAMPP ima naslednje prednosti:

- je enostaven za namestitev in nastavitve,
- vsebuje številne uporabne pakete, ki poenostavijo stvari, kot na primer izdelavo prometnih poročil in pospešitev PHP vsebin,
- bil je temeljito preizkušen na operacijskih sistemih SUSE, Red Hat, Mandrake in Debian Linux distribucijah, kot tudi na Windows ® in Solaris.

V osnovni paket je vključen sistem, programiranje in strežniška programska oprema:

- Apache, najbolj znan spletni strežnik,
- MySQL, odlična, brezplačna in odprtokodna podatkovna baza,
- PHP programski jezik,
- Perl programski jezik,
- ProFTPD, FTP server,
- OpenSSL, za Secure Sockets Layer podporo.



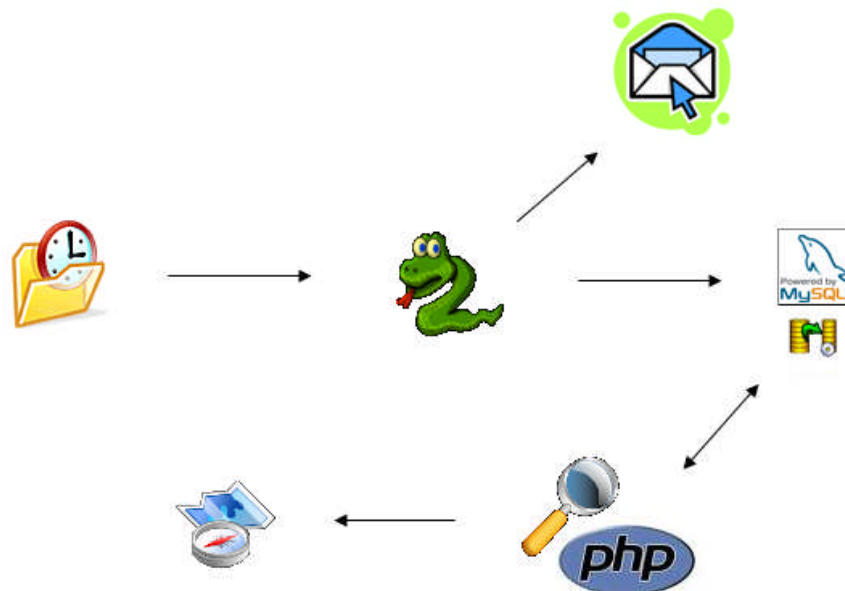
Slika 1 Xampp

3 Prikaz komunikacije posameznih komponent

V tem poglavju sledi prikaz medsebojne komunikacije posameznih komponent sistema.

Vsako jutro ob 9. uri zjutraj Razporejevalnik opravi najprej zažene skripto v Python-u. Nato skripta iz spletnih strani vnaprej izbranih restavracij, izvleče vsebino današnjega menija malic in kosil ter vsebino zapiše v podatkovno bazo MySQL. Po končanem zapisovanju v podatkovno bazo MySQL, pošlje skripta v Python-u elektronsko pošto skrbniku spletnega iskalnika. Elektronska pošta vsebuje priponko v kateri so rezultati uspešnega ali neuspešnega zapisovanja vsebine v podatkovno bazo. Na podlagi neuspešnega zapisovanja vsebine v podatkovno bazo, lahko skrbnik spletnega iskalnika razbere, na katerih spletnih straneh izbranih restavracij je prišlo do vsebinskih sprememb.

Uporabnik zažene spletni iskalnik malic in kosil in vpiše zeleno jed. Spletni iskalnik, napisan v skriptnem jeziku PHP, pošlje vprašanje SQL podatkovni bazi MySQL. Nato se na spletnem iskalniku prikaže seznam restavracij, ki ga vrne vprašanje SQL. Nadalje ima uporabnik tudi možnost prikaza poti do restavracije od naslova, ki ga uporabnik vnese.



Slika 2 Komunikacija komponent

4 Podroben opis in funkcionalnost posameznih komponent ter njihova problematika

V nadaljevanju bom podrobno opisal komponente in njihovo funkcionalnost ter probleme, s katerimi sem se spopadal pri izdelovanju diplomske naloge in način odprave le-teh.

4.1 Zagotavljanje podatkov podatkovni bazi MySQL

Podatkovni bazi MySQL podatke zagotavlja skripta v Python-u. Za zagotovljenost podatkov pa, skrbi Razporejevalnik opravil, ki omenjeno skripto zažene vsak dan.

4.1.1 Dnevno zaganjanje skripte v Python-u

V Razporejevalniku opravil sem kreiral opravilo, ki ima nalogo zaganjanja skripte v Python-u vsakodnevno ob 9. uri zjutraj. Za vsakodnevno zaganjanje skripte sem se odločil, saj nekatere restavracije na spletnih straneh nimajo objavljene tedenske ponudbe malic in kosil. V podatkovni bazi se tako nahajajo samo dnevne ponudbe malic in kosil.

Za Razporejevalnik opravil sem se odločil, ker se spletni iskalnik in njegove ostale komponente izvajajo na lokalnem strežniku.

4.1.2 Razčlenjevanje spletnih strani izbranih restavracij

Za polnjenje podatkovne baze MySQL sem potreboval ustrezne podatke. Te podatke sem dobil tako, da je skripta v Python-u razčlenila spletne strani izbranih restavracij. Za izbor le nekaterih restavracij sem se odločil, ker je bilo spletne strani nekaterih restavracij zaradi njihove sestave, na primer v Flash-u, nesmiselno razčlenjevati. Z razčlenitvijo sem dobil skupek podatkov, ki so potrebni za podatkovno bazo.

Izbor skriptnega jezika Python za razčlenitev spletnih strani se je izkazal kot izziv. Ker sem se s tem jezikom prvič srečal, je bilo programiranje v njem še toliko bolj težavno, a zanimivo, ker si rad zadam nove in nekoliko težje cilje. Izvrsten je za hitro preverjanje spremenjene

kode, saj lahko v Python Interpreter-ju obdelujemo neko spremenljivko dokler ne dosežemo želene učinkovitosti, pri tem pa nam ni potrebno vedno znova zaganjati celotne kode.

Razčlenjevanje spletnih strani v HTML-ju

Razčlenjevanje nekaterih spletnih strani je bilo dokaj enostavno, nekaterih malo težje, nekaterih pa skoraj nemogoče. Probleme z razčlenjevanjem bom opisal v nadaljevanju.

Pri razčlenjevanju sem si pomagal z uporabo razčlenjevalnika BeautifulSoup in z regularnimi izrazi. Z uporabo le-teh sem si olajšal kar precej dela.

- BeautifulSoup [9] je Python-ov HTML/XML razčlenjevalnik, zasnovan za hitro spremenjene projekte, kot so na primer zajem zaslona.

Glavne značilnosti, zaradi katerih je tako koristen:

- o BeautifulSoup se ne zaduši s podatki, če smo slabo podali zahteve za razčlenitev. Poda nam drevesno strukturo razčlenjenih podatkov, ki spominja na prvotni dokument. To je navadno dovolj za zbiranje podatkov, ki jih potrebujemo.
- o BeautifulSoup ponuja nekaj preprostih načinov in Python-ovih slogov za krmiljenje, iskanje in prilagajanje razčlenjenega drevesa. Služi kot orodje za razčlenjevanje dokumentov in pridobivanje tistega kar potrebujemo. Zato nam ni potrebno ustvarjati lastnega razčlenjevalnika.
- o BeautifulSoup samodejno pretvarja vhodne dokumente v Unicode in izhodne dokumente v UTF-8. Zato nam ni potrebno razmišljati o kodiranju, razen če dokument ne opredeljuje natančnega kodiranja, potem BeautifulSoup ne more zaznati načina kodiranja in moramo navesti izvirno kodiranje.

BeautifulSoup razčleni karkoli mu ponudimo in nam kot rezultat vrne drevesno strukturo. Kot argumente mu lahko navedemo »Poišči vse spletne povezave«, »Poišči vse spletne povezave razreda externalLink«, »Poišči vse spletne povezave, kateri spletni naslovi se ujemajo s foo.com« ali »Poišči naslove tabel, ki ima krepko besedilo, nato pa mi vrni to besedilo«. Dragoceni podatki, ki so bili nekoč zaklenjeni

v slabo zasnovanih spletnih straneh, so sedaj dosegljivi. Projekti za katere bi drugače porabili nekaj ur, z Beautiful Soup potrebujemo le nekaj minut.

- Regularno izražanje [10] (včasih skrajšan v "Regex") je način s katerim računalniški programer ali uporabnik izrazi, kako naj računalniški program poišče določen vzorec v besedilu in kaj naj računalniški program naredi z vsakim najdenim vzorcem.

Primer regularnega izražanja bi bil argument programu naj poišče vse vrstice v besedilu, ki vsebujejo besedo "Windows 95" in nato izpiše vsako vrstico, v kateri je ugotovljeno ujemanje ali nadomeščanje z drugim besedilom (na primer, samo "Windows"), kjerkoli se pojavi podobnost.

Najbolj znano orodje za določanje in ravnanje incidence regularnih izrazov je grep. Njegova uporabnost v Unix-u temelji na lastnem operacijskem sistemu, za Windows in druge operacijske sisteme pa je potreben samostojen program.

Spletna stran katero sem se najprej lotil razčlenjevati je bila <http://www.jurman-sp.si/>. Vsako spletno stran razčlenjuje lastna skripta. Razlog temu je želja po čim bolj preciznem spletnem iskalniku malic in kosil. Skripta deluje tako, da na začetku z uporabo knjižnice **urllib** dobimo vsebino celotne spletne strani. Za lažji pregled in razčlenjevanje dobljeno vsebino preoblikujemo v drevesno strukturo z uporabo Beautiful Soup. Primer preoblikovanja je prikazan na spodnji sliki.

```
<td>\r\n\t\t\t\t\t<a href="default.asp" class="headerLink">Domov</a>\r\n\t\t\t\t\t<span class="headerLinkSeperator"></span>\r\n\t\t\t\t\t<a href="Sitemap.asp" class="headerLink">Na\xce8rt strani</a>\r\n\t\t\t\t\t<span class="headerLinkSeperator"></span>\r\n\t\t\t\t\t<a href="Obrazec.asp?form=pisiteNam" class="headerLink">Kontakt</a>\r\n\t\t\t\t\t<span class="headerLinkSeperator"></span>\r\n\t\t\t\t\t<a href="Obrazec.asp?form=pisiteNam" class="headerLink">Delovni \xe8as</a>\r\n\t\t\t\t\t<span class="headerLinkSeperator"></span>\r\n\t\t\t\t\t<a href="TellAFriend.asp" class="headerLink">Priporo\xe8i nas prijatelju</a>\r\n\t\t\t\t\t<span class="headerLinkSeperator"></span>\r\n\t\t\t\t\t<a href="Guestbook.asp" class="headerLink">Knjiga gostov</a>\r\n\t\t\t\t\t</td>

<td>
<a href="default.asp" class="headerLink">Domov</a>
<span class="headerLinkSeperator"></span>
<a href="Sitemap.asp" class="headerLink">NaÄrt strani</a>
<span class="headerLinkSeperator"></span>
<a href="Obrazec.asp?form=pisiteNam" class="headerLink">Kontakt</a>
<span class="headerLinkSeperator"></span>
<a href="Obrazec.asp?form=pisiteNam" class="headerLink">Delovni Äas</a>
<span class="headerLinkSeperator"></span>
<a href="TellAFriend.asp" class="headerLink">PriporoÄi nas prijatelju</a>
<span class="headerLinkSeperator"></span>
<a href="Guestbook.asp" class="headerLink">Knjiga gostov</a>
</td>
```

Slika 3 Primer HTML kode po uporabi Beautiful Soup-a

V zgornjem delu slike (Slika 3) je HTML koda dobljena z uporabo knjižice urllib, spodnji del slike pa prikazuje HTML kodo po uporabi BeautifulSoup-a. Poleg dobljene drevesne strukture so bili odstranjeni tudi nepotrebni znaki za odstavke, novo vrstico, predelčnik itd.. Odstranitev teh znakov je bila koristna pri izpisovanju rezultatov najdenih jedi na spletnem iskalniku.

Z uporabo BeautifulSoup in regularnih izrazov na uvodni spletni strani dobim ime restavracije in nadaljujem z iskanjem spletnega naslova do spletne strani Jedilnik. Že opisani postopek razčlenjevanja vsebine in iskanja spletnega naslova ponavljam dokler ne najdem spletnega naslova do spletne strani Dnevne malice.

Pri razčlenjevanju mi je delo olajšalo tudi to, da ima restavracija Jurman posamezno spletno stran za vsak dan. S pomočjo knjižice **time** in uporabo funkcije **strptime** sem dobil sistemski čas. Na podlagi sistema časa, sem poiskal spletno stran na kateri je vsebina današnjih malic in kosil.

Domov | Načrt strani | Kontakt | Delovni čas | Priporoči nas prijatelju | Knjiga gostov

JURMAN

Če res pravi si gurman, obišči gostilno in pizzerijo Jurman !!

Jedilnik Ambient Dogodki O gostilni

Jedilnik

- Jedilnik
- Otroški menu
- Malice
- **Dnevne malice**
- Ponedeljek
- Torek
- Sreda
- Četrtek
- **Petek**
- Ponudba za vegetarijance

Domov - Jedilnik - Dnevne malice - **Petek**

Petek

1. Mesno zelenjavna enolončnica	4,00 EUR
2. Makaronovo meso (Rdeča pesa v solati)	5,20 EUR
3. Ocvrti file osliča ali soma (Francoska solata)	5,50 EUR
4. Pizza s suho salamo (OD 11:00 DO 17:00 URE)	4,90 EUR

Slika 4 Primer strani iz katere izvlečem ponudbo kosil in malic

Po končani razčlenitvi spletne strani, shranim podatke o trenutnem naslovu spletne strani, o današnji ponudbi malic in kosil ter o imenu restavracije. Po končanem zbiranju podatkov je bilo potrebno podatke ustrezno pripraviti za vpis v podatkovno bazo MySQL.

Zaradi dobre strukture spletne strani restavracije Jurman, razčlenjevanje ni bilo težavno. Dobra struktura strani mi pomeni, da je spletna stran enostavna, berljiva in lepo razporejena z značkami. Značke pa imajo tudi enkratno uporabljena imena, kar nam je v veliko pomoč pri razčlenjevanju, saj moramo pri razčlenitvi natančno navesti lastnosti značk znotraj katerih naj poišče iskano vsebino. Če navedba ni natančna je lahko rezultat vsebine razčlenjevanja napačen oziroma neuporaben.

```
<div id="vsebina_dnevnih_malic_in_kosil">
  <ol>
    <li>Pašta s šinkom</li>
    <li>Kuhani njoki s tunino omako</li>
    <li>Ocvrti piščanec</li>
  </ol>
</div>
<div id="pijača">
  <ol>
    <li>Ledeni čaj</li>
    <li>Gosti sok</li>
    <li>Jabolčni sok</li>
  </ol>
</div>
```

Slika 5 Primer urejene značke

Kot je prikazano na sliki (Slika 5) je identifikacija značke div »vsebina_dnevnih_malic_in_kosil«. Tako sem podal argument razčlenjevalniku, da shrani vse značke li znotraj značke div, z identifikacijo »vsebina_dnevnih_malic_in_kosil«. Tako so bili v podatkovno bazo vneseni samo podatki o kosilih.

Nekatere spletne strani so bile za razčlenjevanje veliko bolj zahtevne. Primer ene takih je <http://www.puttanesca.si>. Ponudba malic in kosil je prikazana na eni spletni strani, poleg tega ima tudi slabo strukturo strani. Nima enkratnih lastnosti značk oziroma jih sploh nima (Slika 6), tako da je zelo težko razčleniti samo tisto vsebino, ki je iskana.

```

<div>
  <ol>
    <li>Pašta s šinkom</li>
    <li>Kuhani njoki s tunino omako</li>
    <li>Ocvrti piščanec</li>
  </ol>
</div>
<div>
  <ol>
    <li>Ledeni čaj</li>
    <li>Gosti sok</li>
    <li>Jabolčni sok</li>
  </ol>
</div>

```

Slika 6 Primer neurejene značke

V tem primeru sem razčlenjevalniku podal argument naj izpiše vse značke li, ki se nahajajo znotraj značk div. Kot rezultat sem dobil vsa kosila in pijače. Zaradi predhodne ureditve vsebine v drevesno strukturo, sem moral izvesti še en korak, da sem dobil iskane podatke.

Razčlenjevanje spletnih strani z vsebino v formatu PDF

Ena od restavracij z vsebino malic in kosil v formatu PDF je restavracija Interspar. Razčlenjevanje PDF vsebine je bilo dokaj težavno. Za uspešno razčlenjevanje sem uporabil pretvorbo PDF datoteke v tekstovno datoteko.

Na spletu je kar pestra ponudba pretvornikov. Mnogi so plačljivi ali neuporabni, saj ne zajemajo celotne vsebine. Eden najboljših pregledovalnikov je Xpdf.

- Xpdf [11] je odprtokodni pregledovalnik za Portable Document Format (PDF) vsebino. Te so včasih imenovane tudi "Acrobat" datoteke iz imena Adobe PDF's software. Xpdf Projekt vključuje tudi luščenje besedil PDF, PDF-to-PostScript pretvornik in različne druge javne storitve.

Xpdf poteka v skladu z X Windows sistemi na UNIX-u, VMS in OS/2. Ne X komponente (pdftops, pdftotext itd.) se prav tako izvajajo na Win32 sistemih in na skoraj vseh sistemih z dostojnim C++ prevajalnikom. Xpdf je zasnovan tako, da je čim

manjši in tem bolj učinkovit. Za pretvorbo uporablja Type 1, TrueType ali standard X pisave.

Xpdf se lahko izvaja na skoraj vsakem sistemu, ki izvaja X11 in Unix podobnim (POSIX) knjižnicam. Vse kar potrebujete je ANSI C++ in C prevajalnike za njihovo prevajanje.

S knjižico pdftotext je bila pretvorba zelo učinkovita in podobna originalu (Slika 7).

Edini problem so bili šumniki, saj jih knjižica pdftotext ne zna pretvoriti. Lahko pa jih pretvori v sičnike in tako ne uniči originalnosti dokumenta.

Dobljeno tekstovno datoteko je bilo potrebno še nadalje razčleniti. Ker se nizi v Python-u vedejo podobno kot seznamih jih lahko indeksiram in režem. Tako sem z nekaj zankami in pogoji prišel do želenih podatkov.

Ponudba malic in kosil je razdeljena v dva stolpca. Temu primerno je potekalo tudi razčlenjevanje. Najprej je bilo potrebno pri vsaki prebrani vrstici pregledati ali vsebuje niz današnjega dne. Za uspešno najden dan sem indeksno mesto v prebrani vrstici shranil. Če se je indeksno mesto nahajalo na levi polovici sredinskega indeksa, je bilo potrebno naslednje vrstice brati samo do polovice in obratno.

<div style="display: flex; justify-content: space-between; align-items: center;"> <div style="text-align: center;"> Restoracija </div> </div>	
TEDENSKA SEZONSKA JED PEČENO PIŠCANCJE BEDRO, GRAH, SMETANOVA OMAKA, KROMPIR S ŠAMPINJONI IN JAJČEVCI € 3,30	
PONEDELJEK	TOREK
KOSILO: - Dnevna juha, sesekljana puranja pečenka, krompirjeva pire, 200 g solate MENU 1: - Poltravnovi štruki, zelenjavna omaka MENU 2: - Ražnja, ajvar, polenta ENOLONČNICA: - Pšička s klobaso VITAL MENU (1749k/422kcal): - Morski pes na žaru, ohrovt s krompirjem	KOSILO: - Dnevna juha, file pange v pivskem testu, krompirjev pire, 200 g solate MENU 1: - Poltravnovi štruki, zelenjavna omaka MENU 2: - Ražnja, ajvar, pomfrit ENOLONČNICA: - Štampa na žaru VITAL MENU (1749k/422kcal): - Svinjski zrezek na žaru, ohrovt s korenčkom
SREDA	ČETRTEK
KOSILO: - Dnevna juha, pretlakano goveje pleče, praženi krompir, 200 g solate MENU 1: - Špečetni golaž, slan krompir MENU 2: - Svinjski zrezek v ciganski omaki, kuhane rezine ENOLONČNICA: - Štampa na žaru VITAL MENU (1441k/349kcal): - Šampinjoni na žaru s peteršiljem	KOSILO: - Dnevna juha, polentni njočki z govejim naparjem, 200 g solate MENU 1: - Špečetni golaž, slan krompir MENU 2: - Svinjski zrezek v ciganski omaki, kruhove rezine ENOLONČNICA: - Jota s klobaso VITAL MENU (1441k/349kcal): - Šampinjoni na žaru s peteršiljem
PETEK	SOBOTA
KOSILO: - Dnevna juha, čebulna brzola, široki rezanci, 200 g solate MENU 1: - Skutine palačinke, sadna solata MENU 2: - Svinjski zrezek parma, stroji fižol s slanim ENOLONČNICA: - Pšička s klobaso VITAL MENU (1339k/323kcal): - Postev na žaru, kuhana ovsenka	KOSILO: - Dnevna juha, svinjska zarezanka po mehiško, pražen krompir, 200 g solate MENU 1: - Ocvrta sardelice, krompirjeva solata MENU 2: - Pšička s klobaso VITAL MENU (1339k/323kcal): - Postev na žaru, kuhana ovsenka
NEDELJA 22.02.2009	
KOSILO 1: Dnevna juha, puranji zrezek po pariško, pečen krompir (steakhouse), 200g solate, menu sladica KOSILO 2: Dnevna juha, mrična pečurka, gobov riž, 200g solate, menu sladica KOSILO 3: Dnevna juha, sarajevski čevapčiči, prebrance, 200g solate, menu sladica	KOSILO: - Dnevna juha, file pange v pivskem testu, pečeni krompir, 200g solate MENU 1: - Zelenjavni narastek, francoska solata MENU 2: - Goveji golaž, polenta ENOLONČNICA: - Vampir po trzasko VITAL MENU (2749k/422kcal): - Svinjski zrezek na žaru, ohrovt s korenčkom CETRTEK KOSILO: - Dnevna juha, polentni njočki z govejim naparjem, 200 g solate MENU 1: - Zelenjavni kroznik z nacho svinjo omako MENU 2: - Pšička s klobaso ENOLONČNICA: - Štampa na žaru VITAL MENU (1441k/349kcal): - Šampinjoni na žaru s peteršiljem PETEK KOSILO: - Dnevna juha, čebulna brzola, široki rezanci, 200 g solate MENU 1: - Skutine palačinke, sadna solata MENU 2: - Svinjski zrezek parma, stroji fižol s slanim ENOLONČNICA: - Pšička s klobaso VITAL MENU (1339k/323kcal): - Postev na žaru, kuhana ovsenka NEDELJA 22.02.2009 KOSILO 1: Dnevna juha, puranji zrezek po pariško, pečen krompir (steakhouse), 200g solate, menu sladica KOSILO 2: Dnevna juha, mrična pečurka, gobov riž, 200g solate, menu sladica KOSILO 3: Dnevna juha, sarajevski čevapčiči, prebrance, 200g solate, menu sladica
TEDENSKA PONUDBA VELJA OD 16.02.2009 DO 22.02.2009	

Slika 7 Primer prevedene spletne strani iz PDF datoteke v tekstovno datoteko

Nadalje je bilo potrebno samo še pripraviti podatke za vnos v podatkovno bazo. Nekatere nepotrebne podatke, dobljene z razčlenjevanjem, sem izbrisal s pomočjo regularnih izrazov, saj pri izpisovanju na spletnem iskalniku ne bi imeli pomena.

Razčlenjevanje spletnih strani v Flash-u

Pri restavracijah s spletno stranjo v Flash-u, razčlenjevanje vsebine ni smiselno, vsaj za zdaj ne. Primer take restavracije je Portal. Poskus pretvorbe Flash-ove datoteke v tekstovno datoteko ni uspela. Težava se pojavi pri šumnikih in pri slabem zaznavanju celotne vsebine spletne strani. Primer pretvorbe je prikazan na sliki (Slika 8), ker pa je večina pretvornikov plačljivih, je slika simbolna. Posledica tega je, da je kar nekaj restavracij izpadlo iz izbora, ker pretvorjeni podatki niso bili koristni.

firstgard
 have seven (7) senator; 7) the Seventh Senatorial District
 term
 and shall have one (1) senator;
 the mleng
 effecti
 ve d'atp
 of this
 8) the Eighth Senatorial District shall be
 consist
 composed of Angaur, Sonsorol and Tobi and shall
 an ik and
 have one (1) senator.
 enashall
 ng detve
 designo

Slika 8 Primer neuspešne pretvorbe spletne strani v Flash-u v tekstovno datoteko [12]

Kot je razvidno iz opisanih primerov različnih spletnih strani, bi bilo zelo težko narediti enotni avtomatski razčlenjevalnik spletnih strani.

4.1.3 Podatkovna baza MySQL s podatki o restavracijah

Podatkovna baza MySQL je sestavljena iz dveh tabel. Prva tabela vsebuje podatke o naslovu restavracij, druga tabela pa podatke o vsebini malic in kosil. Kot prvo bom opisal tabelo s podatki o naslovih izbranih restavracij.

Tabela (Slika 9) vsebuje podatke o imenu restavracije (kar je tudi primarni ključ), imenu ulice, hišni številki, mestu in državi. Podatki o naslovu restavracije so vneseni ročno. Razlog za to je, da imajo nekatere restavracije naslov podan v sliki ali Flash-u in bi se zopet zataknilo pri pretvorbi v tekstovno datoteko.

		ime_restavracije	ulica	st	mesto	drzava
<input type="checkbox"/>			Gostilna in pizzerija Jurman	Zaloška	151	Ljubljana Slovenija
<input type="checkbox"/>			Gostilna A estica	Slovenska cesta	40	Ljubljana Slovenija
<input type="checkbox"/>			Špageterija in pizzerija Favola	Dunajska cesta	129	Ljubljana Slovenija
<input type="checkbox"/>			Interspar Restavracija	Šmartinska	152	Ljubljana Slovenija
<input type="checkbox"/>			Puttanesca — pasta & amp; pizza & amp; salad	Bravnišarjeva	13	Ljubljana Slovenija
<input type="checkbox"/>			Gostišče Julij	Zelena pot	10	Ljubljana Slovenija
<input type="checkbox"/>			Restavracija in Klub Via Bona	Tbilisjska	59	Ljubljana Slovenija
<input type="checkbox"/>			Gostilna Guliver	Vilharjeva cesta	43	Ljubljana Slovenija

↑ Označi vse / Odznači vse Z označenim:

Slika 9 Tabela s podatki o naslovih restavracije

V naslednji tabeli (Slika 10) so zajeti podatki o vsebini malic in kosil po razčlenjevanju. Primarni ključ te tabele je zaporedna številka restavracije. Poleg tega je vpisano še ime restavracije, povezava do spletne strani z vsebino malic in kosil ter posamezne jedi. V zadnjem stolpcu omenjene tabele so posamezne jedi zapisane z majhnimi črkami, saj to poenostavi iskanje in primerjanje nizov med seboj.

		ID	ime_restavracije	link	kosilo	
<input type="checkbox"/>			1	Gostilna in pizzerija Jurman	http://www.jurman-sp.si/katalog.asp?CatalogueGroupID=112	jota s prekajanim mesom testenine z jurkovo omako puranji medaljoni pizza kraska
<input type="checkbox"/>			2	Interspar Restavracija	http://spar.si.spar.at/rsync_spar_si/Restavracija/Restavracijajedilnik2.pdf	dnevna juha njoki s skampi in rdecim radicem parmezan 200 g solate skutini struklji kremna spinaca svinjski zrezek s porom in sampinjoni kroketi pasta fizol s klobaso piscancji zrezek z zelenjavo kapo kuhan riz
<input type="checkbox"/>			3	Špageterija in pizzerija Favola	http://www.favola.si/kosila.php	golazeva juha puran polnjen s suhimi slivami v rahli smetanovi omaki široki rezanci solata telecja obara z žličniki zelenjavna rizota s parmezanom domaca torta z borovnicevo kremo in belo cokolado
<input type="checkbox"/>			4	Puttanesca — pasta & amp; pizza & amp; salad	http://www.puttanesca.si/malice.php	malica mesna lasagna sestavljena solata kosilo dnevna juha piscanec v vrtnarski omaki priloga njoki z maslom sestavljena solata sladica malica pizza (polnozrnato testo) mala pizza vesuvio (pelati sir pikantna salama in omaka pekoci feferoni)
<input type="checkbox"/>			5	Gostišče Julij	http://julci.si/index.php?option=com_content&task=view&id=7&Itemid=	pasulj z mesom vrtnarsko meso solata vampi v omaki dnevna juha kraski zrezek rezanci s spinaco solata sladica
<input type="checkbox"/>			6	Gostilna A estica	http://www.sestica.si/dnevna-kosila/	golazeva ali zelenjavna juha ramsteak in zelenjava na zaru pecen krompir solata sladica ali kava golazeva ali zelenjavna juha pecenica kislo zelje in prazen krompir sladica ali kava golazeva ali zelenjavna juha domaci ajdovi zepki z buckami in smetano solata sladica ali kava
<input type="checkbox"/>			7	Restavracija in Klub Via Bona	http://www.via-bona.com/kosila.asp	dnevna juha goveja ali jesprenckova juha z zelenjavo široki rezanci s porom sampinjoni in smetano mesna lasagna piscancji file v sirovi omaki ocvtvi svaljki mesana zelenjava pljučne rezinice s sampinjoni in parmezanom na rukoli zapecen krompir solata sladica
<input type="checkbox"/>			8	Gostilna Guliver	http://www.guliver.net/kosila.html	dnevna juha goveja ali zelenjava s pisanim fizolom puranji zrezek s sampinjoni v smetanovi omaki kruhov cmok široki rezanci z belusi in zelenjavo mesna lasagna solata - solatni bar sladica

↑ Označi vse / Odznači vse Z označenim:

Slika 10 Tabela s podatki o vsebini malic in kosil

Tabele se med seboj povezujejo z relacijo ena proti nič oziroma neskončno. Toda v konkretnem primeru bi lahko bile povezane tudi z relacijo ena proti nič oziroma ena. Za relacijo ena proti nič oziroma neskončno sem se odločili zaradi možnosti kasnejše razširitve spletnega iskalnika. Z dodanimi novimi atributi bi lahko vodil statistiko malic in kosil skozi celoten mesec oziroma vse leto. Tako bi lahko spremljal spreminjanje vsebine malic in kosil skozi mesec oziroma vse leto, lahko bi tudi predpostavil ponudbo restavracij v prihodnjih mesecih.

4.1.4 Obveščanje skrbnika sistema

Želja po čim bolj preciznem spletnem iskalniku malic in kosil ter natančnem prikazu le-teh, je s seboj prinesla tudi nekaj slabosti. Največja slabost je občutljiv razčlenjevalnik. Zaradi neprestanega razvoja tehnologij ali želje po čim boljši promociji restavracije prek spleta, se temu primerno spreminjajo tudi njihove spletne strani.

Razčlenjevalniku je potrebno natančno navesti lastnosti značk, znotraj katerih se črpajo potrebni podatki za vnos v podatkovno bazo. S spremembo spletnih strani se spremenijo tudi lastnosti značk, ki so prvotno navedene razčlenjevalniku. Kot posledica tega se pojavi napaka pri razčlenjevanju oziroma prazen vnos v bazo. Da bi se izognili napakam zaradi sprememb spletnih strani, se vsakodnevno obvesti skrbnika spletnega iskalnika o uspešnosti razčlenjevanja.

Skrbnik spletnega iskalnika prejme vsak dan po elektronski pošti tekstovni dokument, v katerem so rezultati razčlenjevanja. Dokument vsebuje podatke o uspešnosti praznjenja tabele v podatkovni bazi, kjer je zapisana vsebina malic in kosil ter uspešnost razčlenjevanja po posameznih restavracijah. V primeru neuspešnega razčlenjevanja se poleg imena restavracije zapiše tudi napaka, ki je privedla do neuspešnega razčlenjevanja. Primer tekstovnega dokumenta, ki vsebuje informacije o razčlenjevanju je prikazan na sliki (Slika 11).

```

Brisanje vsebine kosila: OK
Jurman: napaka pri branju ('NoneType' object has no attribute 'string')
Spar: OK
Favola: OK
Puttanesca: OK
Julci: OK
Sestica: OK
Via Bona: OK
Guliver: OK

```

Slika 11 Primer rezultatov razčlenjevanja

4.2 Spletni iskalnik v PHP-ju

Kot sem že omenil, je Googlov spletni iskalnik služil kot osnova za prikaz in delovanje spletnega iskalnika malic in kosil. Za Googlov spletni iskalnik sem se odločil zato, ker je uporabniku prijazen in hkrati zelo učinkovit.

Spletni iskalnik malic in kosil je napisan v skriptnem jeziku PHP. Za uporabo skriptnega jezika PHP sem se odločil zaradi predhodnih izkušenj in njegove enostavne uporabe. Dodaten razlog izbire skriptnega jezika PHP je razširjenost uporabe pri izdelovalcih spletnih strani kar pomeni, da je veliko primerov uporabe določenih funkcij in problemov pri programiranju objavljenih na spletu. Zato sem si pri reševanju problemov pomagal iz podobnih primerov objavljenih na spletu.

4.2.1 Vnos zelene jedi in poizvedovanje po podatkovni bazi

Po uporabnikovi povezavi s spletnim naslovom se prikaže spletna stran iskalnika malic in kosil. V zgornjem delu spletne strani se nahaja vnosno polje, v katerega uporabnik vpiše zeleno jed. Vnosno polje kontrolira dolžino vpisanega niza. Če je dolžina vpisanega niza manjša ali enaka dvema znakoma, se v spodnjem delu spletne strani prikaže opozorilo namenjeno uporabniku, kjer piše »Niz mora vsebovati najmanj tri znake.«.

Nizu ustrezne dolžine se najprej odstranijo vsi šumniki, ki se nadomestijo s sičniki. Razlog temu je različen zapis v podatkovni bazi. Kot sem že prej omenil imajo nekatere restavracije na svojih spletnih straneh vsebine malic in kosil objavljene v PDF datoteki. Tako je vsebina malic in kosil, ki so objavljena v PDF datoteki v podatkovno bazo zapisana s sičniki, vsebina malic in kosil objavljena na drugih spletnih straneh pa s šumniki. Zaradi različnih vnosov v podatkovno bazo se je pojavljal problem pri iskanju zelene jedi.

Celotna dnevna ponudba restavracije se v podatkovno bazo zapiše v eno polje tabele kot en sam niz. Zapisani niz je sestavljen iz več podnizov. Za iskanje po bazi sem uporabil pogoj LIKE. Ker sem hotel primerjati nize enega z drugim sem moral uporabiti tudi znak »%«, katerega sem dodal na začetek in konec vsakega niza, ki ga je uporabnik vpisal. Znak »%« služi za nadomestne znake.

- Primer uporabe znaka »%«:
 - o V podatkovni bazi obstaja niz »Pečen piščanec s krompirjem in solato«. Če uporabnikovem nizu »piščanec« ne nastavim znaka »%« na začetku in na koncu niza, se pri primerjanju obeh nizov ne bo prikazal zapis iz podatkovne baze. Pogoj LIKE primerja celoten niz iz podatkovne baze z nizom, ki ga je uporabnik vpisal.
 - o Če nizu »piščanec«, ki ga je vpisal uporabnik na začetku in na koncu nastavim znak »%« (»%piščanec%«), pogoj LIKE pri iskanju podniza dovoljuje obstoj podnizov pred in za podnizom »piščanec« v podatkovni bazi in tako se kot rezultat vprašanja SQL prikaže zapis iz podatkovne baze.

Kot primer si predstavljajmo, da uporabnik vpiše v vnosno polje spletnega iskalnika zeleno jed »piščanec«. V podatkovni bazi pa obstajata dva zapisa. V prvem zapisu je niz piščanec zapisan s šumniki, v drugem pa s sičniki. Po izvedenem vprašanju SQL se kot rezultat prikaže le prvi zapis iz podatkovne baze, ker v podatkovni bazi pri primerjanju niza »piscanec« z nizom »piščanec«, ki ga je uporabnik vpisal v vnosno polje, niz ni enak, se le-ta ne prikaže in obratno. Tak način iskanja jedi ni bil zadovoljiv.

Zato sem se odločil spremeniti način iskanja po podatkovni bazi. Primeren način za rešitev problema je bila uporaba znaka »_«. Tako sem vse šumnike v nizu, ki ga je vpisal uporabnik,

zamenjal z znakom »_«. Namesto tega znaka se lahko nahaja katerakoli črka. Pri zamenjavi šumnikov ali sičnikov z znakom »_« je pogoj LIKE primerjal niz »pi__ane_« z zapisoma v podatkovni bazi in kot rezultat prikazal oba zapisa. Problem se je zdel rešen dokler nisem ugotovil, da v primeru, ko uporabnik vpiše niz »riž« se kot rezultat lahko prikaže tudi zapis, ki vsebuje niz »rib«. Tak, delno precizen iskalnik ni ponujal prave rešitve.

Problem sem poskusil rešiti še na naslednji način. Po vpisanem uporabnikovem nizu sem najprej primerjal niz s šumniki z obema zapisoma v podatkovni bazi in na podlagi tega dobil kot rezultat prvi zapis. Nato sem vse šumnike zamenjal s sičniki in dobil drugi zapis iz baze. Tako sem zopet dobil oba zapisa iz baze. Ker šumnike ni bilo težavno spremeniti v sičnike, je pa nastala težava pri obratnem spreminjanju. Na primer kako vedeti, če bi uporabnik vpisal niz »piscanec« s sičniki, katere od teh je potrebno spremeniti v šumnike, da bi kot rezultat vprašanje SQL dobil oba zapisa iz podatkovne baze. Tudi ta težava bi se lahko rešila tako, da bi zapise v podatkovni bazi primerjal z vsemi možnimi kombinacijami zamenjave sičnikov v šumnike in obratno v nizu, ki ga je vpisal uporabnik. Prekomerno primerjanje bi zahtevalo veliko časa. Tako tudi ta rešitev ni zadovoljila želje po preciznem spletnem iskalniku malic in kosil.

Alternativa zgoraj opisanemu problemu je bila, da pred vnosom v podatkovno bazo odstranim vse šumnike in jih nadomestim z sičniki. Ta rešitev je bila zadovoljiva, zato sem jo uporabil. Najprej sem moral prilagoditi skripto Python. Z nekaj ukazi sem problem poizvedovanja rešil z zamenjavo šumnikov v sičnike pred vnosom v podatkovno bazo. Tako sem dobil vse zapise o vsebini malic in kosil s sičniki. Pri vnosu uporabnikove zelene jedi s šumniki, te zamenjam s sičniki in tako dobim vse zapise iz podatkovne baze v katerih se nahaja uporabnikova zelena jed. Pri vpisu uporabnikove zelene jedi s sičnike je pa vse skupaj še toliko lažje.

Z rešenim problemom šumnikov in sičnikov poizvedovanje še ni bilo zaključeno. Problem, ki se je pojavljal pri poizvedovanju je zaradi sklanjatev jedi. Tako sem moral vsak uporabnikov vpis zelene jedi skrajšati za nekaj zadnjih znakov. Na primer, če uporabnik vpiše v vnosno polje »obara« kot zeleno jed in tega skrajšam za zadnja dva znaka, po podatkovni bazi poizvedujemo z nizom »oba«. Zaradi uporabe že prej omenjenega znaka »%«, ki sem ga dodal na začetek in na konec niza, sem kot rezultat poizvedbe dobil zapis iz podatkovne baze, ki je vseboval jed »obara« in »klobasa«. Oba zapisa pa sem dobil zato, ker oba niza vsebujeta podniz »oba«. Z uporabo regularnih izrazov je bil tudi ta problem rešljiv. Z regularnim

izrazom lahko poizvedovanje omejim na začetek vsakega podniza znotraj niza. Tako da pri vnosu uporabnikove zelene jedi »obara«, kot rezultat poizvedovanja dobim le zapis, v katerem se nahaja niz »obara«.

Dodatna funkcionalnost pri iskanju zelene jedi in s tem tudi dodaten problem, je bila razširitev iskanja na več jedi hkrati. Uporabniku tako ni potrebno iskati vsake jedi posebej, ampak vpiše celoten seznam jedi, ločenih s presledkom ali vejico. Kot rezultat poizvedovanja dobi vse restavracije, ki ponujajo vsaj eno od zelenih jedi uporabnika. Največji problem pri omogočanju iskanja več jedi hkrati, je bil prvotno v pomanjkanju idej za rešitev.

Težava je bila tudi v tem, da nisem znal sestaviti vprašanja SQL. Na koncu sem z nekaj raziskovanja po spletu in z večkratnim poskušanjem prišel do rešitve. Problem sem rešil tako, da sem zelene jedi, ki jih je uporabnik vpisal spravil v seznam. Vprašanje SQL je sestavljeno v odvisnosti od velikosti seznama in sicer tako, da kot rezultat poizvedbe prikaže vse tiste restavracije, ki vsebujejo vsaj eno od vpisanih uporabnikovih zelenih jedi.

4.2.2 Prikaz zadetkov

Po predstavitvi uspešno izvedene poizvedbe bom predstavil prikaz zadetkov. Rezultate poizvedbe sem želel prikazati na podoben način kot Googlov spletni iskalnik. Najprej prikaže ime restavracije, ki je napisano z nekoliko večjimi črkami kot ostali podatki in modro barvo z namenom, da jo hitreje opazimo. Pod imenom restavracije se z nekoliko manjšimi črkami izpiše delni zapis jedi, ki jo je uporabnik vpisal. Najdene posamezne jedi sem pri tem zapisu hotel zapisati krepko z razlogom, da uporabnik hitreje ugotovi, katere jedi so bile najdene v posameznih restavracijah.

Pri tem izpisovanju se je pojavilo kar nekaj manjših problemov. Prvi problem, ki se je pojavil je bil ta, kako vedeti katero jed je vprašanje SQL našlo. Težavo sem rešil tako, da sem zelene jedi, ki jih je uporabnik vpisal, spravil v seznam. Pred izpisovanjem posameznih najdenih restavracij, sem vsebino malic in kosil tudi spravil v seznam. Tako sem lahko vsako vpisano jed uporabnika, primerjal z jedmi iz podatkovne baze. Ob ustreznem primerjanju sem si jed zapomnil in ji nastavil značko za krepek izpis. S krepkim zapisom prikazane zelene jedi še vedno ni zadovoljivo. Pri takemu prikazu uporabnik pri iskanju juhe ne bi mogel vedeti,

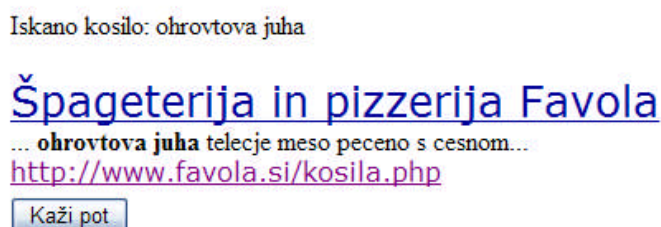
katero juho ponujajo v restavraciji. Zato sem se tudi pri tem izpisu zgledoval po Googlovem spletnem iskalniku, saj ta prikaže nekaj podnizov pred najdenim ustreznim podnizom, in nekaj podnizov za njem.

Ker sem vsebino malic in kosil iz podatkovne baze predhodno spravil v seznam, sem tako z zanko izpisal tri nize pred iskano jedjo in tri nize za iskano jedjo. S takim izpisom ima uporabnik vsaj opisano jed, ki jo je iskal. Po odpravljenem izpisu krepko zapisanih jedi se je pojavil nov problem. Ta je prikazan na sliki (Slika 12).



Slika 12 Prikaz napačnega izpisa rezultata poizvedbe

Kot je razvidno na sliki (Slika 12) je uporabnikova zelena jed »ohrovtova juha«. Spletni iskalnik je najprej s primerjanjem našel podniz »ohrovtova«. Ker je v podatkovni bazi v vsebini malic in kosil to prvi podniz, se zato pred njem ne izpiše nič, za njem pa naslednji trije podnizi. V drugem primerjanju je našel jed »juha«. Pri izpisu sta se tako izpisala dva podobna podniza, kar seveda ni bilo prav. Uporabniku ne preveč prijazen izpis sem poskusil spremeniti, tako da sem v prvem nizu »... ohrovtova juha telecje meso...« poskusil s primerjanjem najti še drug podniz zelene jedi (**juha**). Želja je bila tudi, da če se v tem nizu nahaja še kakšna jed, se niz podaljša pri izpisu še za nekaj podnizov od zadnje najdene jedi. Primer uspešno rešenega in zelenega izpisa se nahaja na sliki (Slika 13).



Slika 13 Prikaz zelenega izpisa rezultata poizvedbe

Težavo pri izpisu sem rešil tako, da sem se najprej sprehodil čez celoten seznam vsebin malic in kosil. Vsakemu ustrezno najdenemu podnizu sem dodal značko za krepek izpis na spletni strani. S tem sem dobil niz s krepkim zapisom ustrezno najdenih jedi v njem. Po dodajanju značk za krepek izpis, sem se zopet sprehodil skozi seznam. V vsakem podnizu iz seznama sem preveril, če vsebuje značko za krepek izpis. Pri uspešno najdenem podnizu sem si zapomnil mesto nahajanja v seznamu in če se v naslednjih treh podnizih nahaja še kakšna značka za krepek izpis, se izpis podaljša še za tri podnize. V primeru, da se naslednja značka za krepek izpis nahaja dlje od treh podnizov, se za to značko izpiše posamezen podniz, ki se pripne predhodnemu podnizu. Izpis je tako bolj uporabniku prijazen.

4.3 Prikazovalnik poti

Za prikaz poti od uporabnikovega naslova do restavracije sem uporabil zemljevid Google Maps. Zemljevid Google Maps s podporo JavaScript omogoča pridružitve zemljevida k lastnim spletnim stranem in prilagajanje našim zahtevam. S pridružitvijo prikaza poti k spletnemu iskalniku malic in kosil bi bila spletna stran predolga in nepregledna. Namesto pridružitve prikaza poti sem se odločil prikazati pot v novem oknu brskalnika. Tako ima uporabnik možnost prikaza več poti naenkrat. Takšen način prikaza je veliko bolj pregleden in uporabniku prijazen. Za način prikazovanja poti sem izbral prikaz poti na zemljevidu s tekstovnim opisom poti in razdalj ter z možnostjo povečanja ali zmanjšanja zemljevida.

Za omogočanje prikaza poti sem moral posredovati podatke, o vpisanem naslovu uporabnika in naslovu restavracije, iz spletnega iskalnika malic in kosil v spletno stran prikaza poti. Pri prenosu podatkov o naslovu uporabnika in restavracije sem si pomagal z značko form, ki vsebuje vnosna polja in gumbe. Gumb ima v znački metodo za shranjevanje podatkov v spletni naslov, ki so bila vpisana v vnosna polja. Za vsako prikazano restavracijo sem kreiral svojo značko. Za zbiranje podatkov o naslovu uporabnika in restavracije iz spletnega naslova sem uporabil značko form z metodo POST. Spremenljivka `$_POST` je niz spremenljivk imen in vrednosti, ki ga je poslala HTTP POST metoda. Informacije, poslana iz značke form z metodo POST, niso vidne drugim in pri pošiljanju ni omejitev glede količine le-teh. Metoda se lahko uporablja pri spletnih straneh napisanih v skriptnem jeziku PHP. Shranjene podatke v spletnem naslovu iskalnika lahko spletna stran prikaza poti te podatke izlušči in uporabi.

Kot sem že prej omenil, se je za vsak izpis restavracije kreirala svoja značka form. Uporabnikov naslov sem shranjeval v spletni naslov na podlagi vpisa v vnosna polja. Naslov restavracije pa sem shranjeval tako, da se je pri izpisu podatkov o restavraciji na spletnem iskalniku, v uporabniku nevidno vnosno polje vpisal naslov restavracije iz podatkovne baze. Po vpisu uporabnikovega naslova in pritisku na gumb »Kaži pot« sta se tako oba naslova shranila v spletni povezavi iskalnika.

Pri začetnem pošiljanju podatkov je moral uporabnik vpisati lasten naslov za vsako prikazano restavracijo. Omenjeno pošiljanje podatkov je prikazano na sliki (Slika 14).

[Špageterija in pizzerija Favola](#)
 ... brokoli kremna juha spageti s peceno...
<http://www.favola.si/kosila.php>

Ulica Hišna št. Mesto

[Gostišče Julči](#)
 ... stajerska kisla juha makaronovo meso solata... meso solata prebranec dnevna juha pariski zrezek prazen krompir solata...
http://julci.si/index.php?option=com_content&task=view&id=7&Itemid=

Ulica Hišna št. Mesto

[Gostilna Šestica](#)
 ... goveja ali zdrobova juha prekajena kuhana svinjska... prekajena kuhana svinjska kraca kisla repa zelje... goveja ali zdrobova juha piscancje nabadalo z... goveja ali zdrobova juha domaci njoki z...
<http://www.sestica.si/dnevna-kosila/>

Ulica Hišna št. Mesto

Slika 14 Primer vnosa naslova za vsako prikazano restavracijo

Zgoraj prikazan primer pošiljanja in vnašanja naslova uporabnika ni primeren, saj bi se uporabnik ob prikazu večjega števila restavracij, naveličal vpisovanja znova istega naslova.

Alternativa zgoraj omenjenem vpisu naslova uporabnika je bila, da bi uporabnik vnesel naslov samo enkrat in ta bi se shranil dokler ga ne bi uporabnik spremenil ali zapustil spletni iskalnik. Osnovni princip shranjevanja in pošiljanja podatkov preko spletnega naslova je isti. Pri kreiranju značk form se podatki o naslovih pri vsakem izpisu posamezne restavracije še nahajajo v vnosnih poljih, le da uporabnikove podatke o naslovu posredujem iz neke druge značke form. Težava se je pojavila pri posredovanju podatkov iz ene značke form v drugo.

Najprej sem poskusil težave odpraviti s pomočjo funkcije v JavaScript. Funkcija ob vsaki spremembi v vnosnem polju, na podlagi identifikacije značke, posreduje podatke znački form, ki se nahaja pri izpisu restavracije. Rešitev je bila zadovoljiva le pri izpisu ene restavracije, pri izpisu več restavracij pa se podatki o naslovu uporabnika niso posredovali zaradi večkratne pojavitve istih identifikacij značk.

Način posredovanja podatkov sem spremenil, tako da sem podatke o naslovu uporabnika posredoval znački form prek spletnega naslova iskalnika. Vnosna polja v katera uporabnik vpiše naslov se nahajajo v isti znački form kot vpis želene jedi. Za tak način rešitve problema sem se odločil zato, ker so se ob vsakem novem vpisu jedi in osveževanju spletnega iskalnika podatki o naslovu uporabnika izbrisali in obratno. V skupni znački form sta dva gumba, prvi je namenjen iskanju želene jedi, drugi pa za shranjevanje naslova. Ob pritisku katerega koli gumba se iz vseh vnosnih polj, ki se nahajajo v skupni znački form, shranijo vpisani podatki v spletni naslov. Podatke posredujem vsem izpisanim značkam form pri posameznih restavracijah s pomočjo metode POST. Zaradi shranjevanja podatkov v spletni naslov se ob vpisu novih zelenih jedi ali naslova uporabnika prikažejo na osveženem spletnem iskalniku tako novi kot predhodno vpisani podatki.

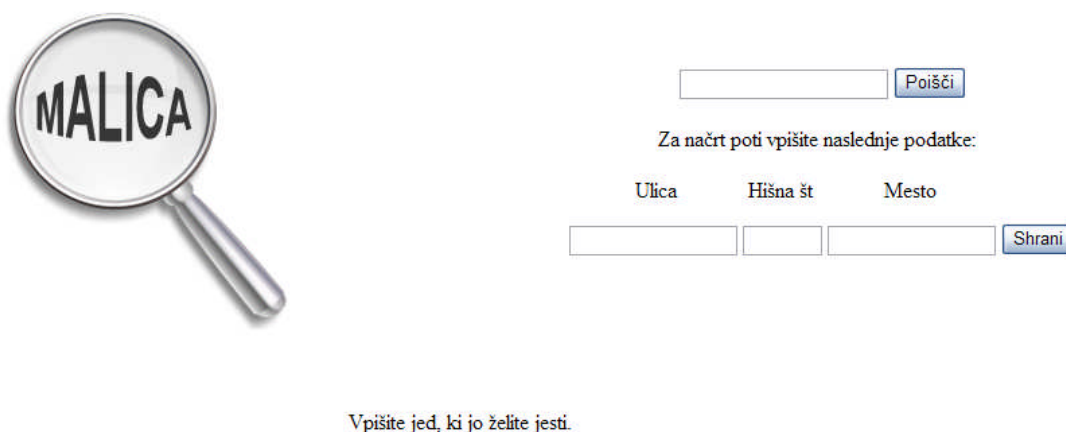
Novo okno v brskalniku, ki prikaže pot od uporabnikovega naslova do naslova restavracije, dobi podatke iz spletnega naslova iskalnika z uporabo metode POST in spremenljivke `$_POST`. Dobljene podatke shranim v spremenljivke skriptnega jezika JavaScript. Te spremenljivke uporabim pri Google Maps metodi `directions.load()`. Omenjena metoda skrbi za prikaz poti na zemljevidu od navedenega začetnega naslova uporabnika do navedenega končnega naslova restavracije. Pri prikazu poti na zemljevidu sta uporabljeni še metodi:

- `GDirections`, ki skrbi za tekstovni prikaz poti in
- `GLargeMapControl`, ki skrbi za povečanje ali pomanjšanje zemljevida.

Natančnost prikaza poti je odvisna od natančnosti zemljevidov Google Maps. Pri nekaterih krajih, ki so v zemljevidu natančno opisani, nam pot prikaže do hišne številke natančno. Pri slabo opisanih krajih pa je natančnost nekoliko slabša, zato prikaže pot le do vpisane ulice in ne do podane hišne številke.

5 Prikaz delovanja spletnega iskalnika


V nadaljevanju je na primeru opisano in prikazano delovanje spletnega iskalnika malic in kosil. Po uporabnikovi zahtevi po spletnem iskalniku malic in kosil se prikaže spletna stran prikazana na sliki (Slika 15).



The image shows a web search interface. On the left, there is a magnifying glass icon with the word 'MALICA' inside it. To the right, there is a search form. At the top, there is a text input field followed by a blue button labeled 'Poišči'. Below this, the text 'Za načrt poti vpišite naslednje podatke:' is displayed. Underneath, there are three labels: 'Ulica', 'Hišna št', and 'Mesto', each followed by a text input field. To the right of the 'Mesto' field is a blue button labeled 'Shrani'. Below the input fields, the text 'Vpišite jed, ki jo želite jesti.' is displayed.

Slika 15 Prikaz začetne spletne strani iskalnika

Po prikazu strani spletnega iskalnika malic in kosil uporabnik vpiše želeno jed v vnosno polje. Uporabnikova želja iskane jedi je »dnevna juha« in »krača«. Po vpisu želene jedi v vnosno polje, uporabnik pritisne na gumb »Poišči«. Spletni iskalnik malic in kosil prikaže uporabniku restavracije (Slika 16), ki ponujajo vpisane zelene jedi.



Za načrt poti vpišite naslednje podatke:

Ulica Hišna št. Mesto

Iskano kosilo: dnevna juha, krača

[Špageterija in pizzerija Favola](#)
 ... brokoli kremna juha spageti s peceno...
<http://www.favola.si/kosila.php>

[Gostišče Julči](#)
 ... stajerska kisa juha makaronovo meso solata... meso solata prebranec dnevna juha pariski zrezek prazen krompir solata...
http://julci.si/index.php?option=com_content&task=view&id=7&Itemid=

[Gostilna Šestica](#)
 ... goveja ali zdrobova juha prekajena kuhana svinjska... prekajena kuhana svinjska kraca kisa repa zelje... goveja ali zdrobova juha piscancje nabodalo z... goveja ali zdrobova juha domaci njoki z...
<http://www.sestica.si/dnevna-kosila/>

Slika 16 Prikaz restavracij, ki ponujajo uporabnikovo želeno jed

Poleg imena restavracij se prikaže tudi delni opis uporabnikove želene jedi, ki jo ponuja prikazana restavracija. Za zelo podroben opis (Slika 17) ali ogled preostalih jedi uporabnik pritisne na spletno povezavo, ki se nahaja pod delnim opisom uporabnikove želene jedi.



FAVOLA

[Drobnjica](#)
[Dnevna Kosila](#)
[Dnevna hrana](#)
[Vinska karta](#)
[Zadovoljstvo](#)

• danes
 • jutri
 • pojutrišnjem
 • 28.02.2009
 • 01.03.2009
 • 02.03.2009

Kosila
25.02.2009

Ohrovtova juha
Cena: 2,50 €

Telečje meso pečeno s česnom, korenjem, slanino, figami in jabolki, njoki, solata
Cena: 7,00 €

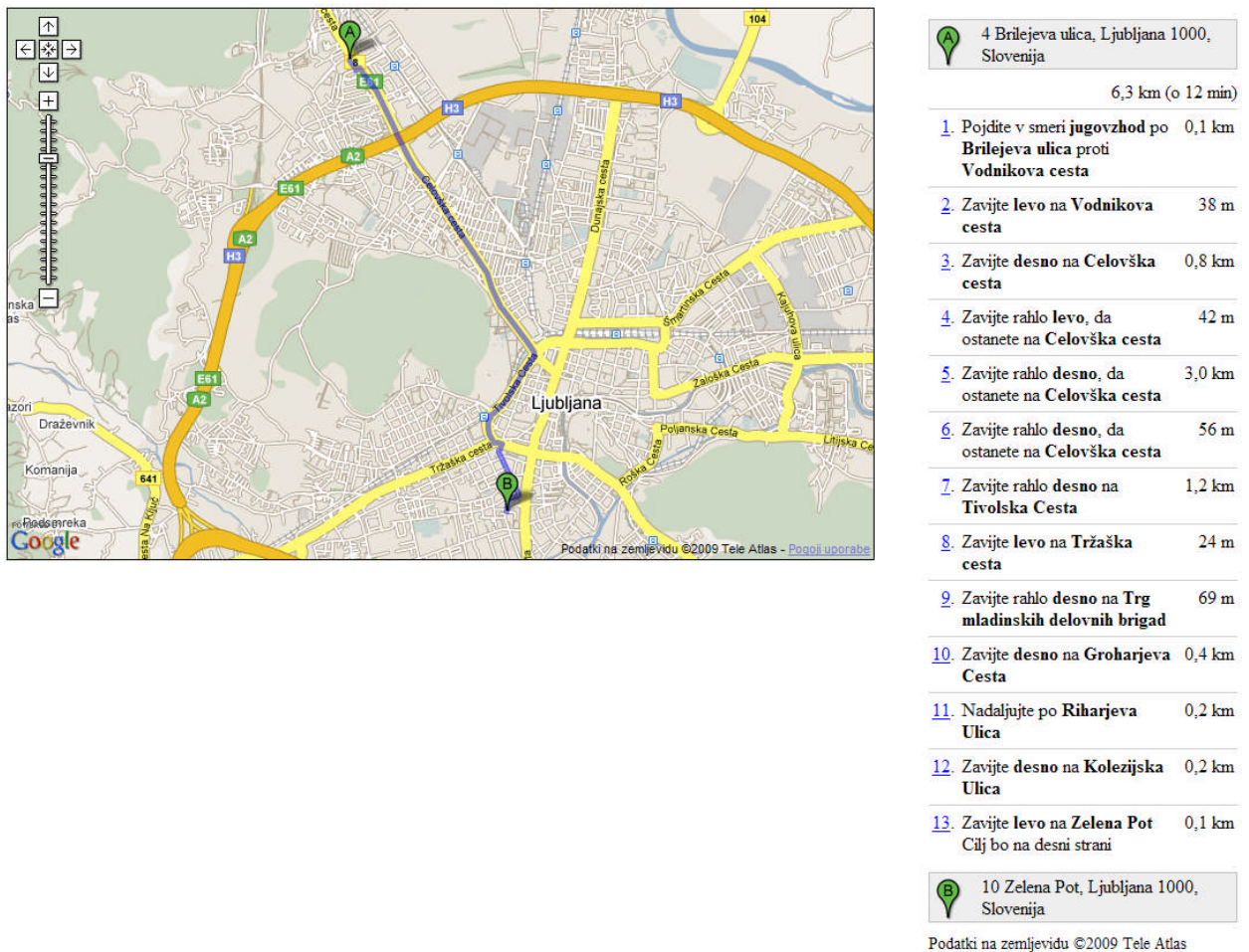
Bučke polnjene z ajdovo kašo in lososom
Cena: 6,70 €

Segedin golaž
Cena: 6,00 €

Krho testo z jogurtovo kremo in gozdnimi sadeži
Cena: 2,70 €

Slika 17 Slika Podroben prikaz ponujenih današnjih malic in kosil

Zaradi nepoznavanja vseh restavracij in tako tudi njenih naslovov, ima uporabnik na razpolago prikaz poti od svojega vpisanega naslova do naslova restavracije. Uporabnik vpiše želen naslov tako, kot zahtevajo vnosna polja. Najprej vpiše ime ulice, nato številko ulice in ime mesta na katerem se nahaja oziroma iz katerega želi na pot proti restavraciji. Po vpisanem naslovu uporabnik pritisne na gumb »Shrani« in tako se njegov naslov shrani na spletni strani vse dokler je ne spremeni oziroma ne zapusti spletnega iskalnika. Za prikaz poti do restavracije pa uporabnik pritisne na gumb »Kaži pot«. Poleg prikazane poti (Slika 18) so prikazana tudi navodila kako naj uporabnik pride do iskane restavracije, oddaljenost v kilometrih in približen čas potovanja.



Slika 18 Slika Prikaz poti od uporabnikovega naslova do naslova restavracije

6 Sklep

Z izdelavo spletnega iskalnika malic in kosil v ljubljanskih restavracijah z možnostjo prikaza poti do le-teh, sem uporabniku omogočil hitrejšo, enostavnejšo in bolj preciznejšo iskanje zelenih jedi. S preciznostjo je mišljeno, da iskalnik prikaže le restavracije, ki imajo na današnji ponudbi malic in kosil zeleno jed.

Težave, s katerimi sem se spopadal, so bile zaradi različno sestavljenih spletnih strani restavracij, šumnikov v besedilu in pomanjkanja znanja na področjih, s katerimi sem se srečeval pri izdelavi diplomskega dela. Probleme in lastne rešitve sem se trudil čim bolj natančno razložiti in prikazati na tistih primerih, ki bi jih bralec lahko čim bolj razumel.

Možnosti za izboljševanje spletnega iskalnika je veliko. Ena od teh je gotovo univerzalni razčlenjevalnik. S tem bi namesto podrobnega razčlenjevanja le izbranih spletnih strani restavracij lahko razčlenjevali bolj obširno, izgubili pa bi preciznost pri prikazu besedila malic in kosil. Z izdelavo univerzalnega razčlenjevalnika, ki bi znal razčlenjevati večino spletnih strani restavracij, bi lahko dnevno spremljal objave novih restavracij na spletu. Spletne strani restavracij, tako ne bi bilo potrebno ročno iskati in vpisovati v seznam restavracij za razčlenjevanje ampak, bi se seznam kreiral avtomatsko. Tako bi bila tudi ponudba malic in kosil novo objavljenih spletnih strani restavracij dnevno osvežena in dostopna uporabniku pri iskanju zelene jedi. Izboljššan bi bil lahko tudi izbor izpisanih restavracij na podlagi zaznavanja uporabnikovega naslova. To pa pomeni, da bi se prikazale le restavracije, ki so v njegovem okolišu.

Že pri sami izdelavi sem nekatere rešitve sproti izboljševal, pri nekaterih pa je bilo enostavno premalo časa za izboljšavo.

7 Viri

- [1] (2009)Opis Razporejevalnika opravil. Dostopno na:
http://oraport.fov.uni-mb.si/pls/portal/PORTAL.wvw_user_utilities.exec_procedure?p_session_id=&p_module_id=1389253212&p_version=1&p_arg_names=SIFRA&p_arg_values=1682
- [2] (2009>About MySQL. Dostopna na:
<http://www.mysql.com/about/>
- [3] (2009)Python. Dostopna na:
<http://www.python.org>
- [4] (2009)What is PHP. Dostopna na:
<http://www.amazon.com/PHP5-MySQL-Bible-Tim-Converse/dp/0764557467#reader>
- [5] (2009)O Programu Dreamweaver CS4. Dostopna na:
<http://www.unze.ba/am/pzi/2008/EmirSabic/programi/wd-dreamweaver.html>
- [6] (2009>About Apache. Dostopno na:
<http://httpd.apache.org/>
- [7] (2009)Apache Server Frequently Asked Questions. Dostopna na:
<http://httpd.apache.org/docs/1.3/misc/FAQ.html#what>
- [8] (2009)Introducing XAMPP. Dostopna na:
<http://www.ibm.com/developerworks/linux/library/l-xampp/>
- [9] (2009>About Beautiful Soup. Dostopna na:
<http://www.crummy.com/software/BeautifulSoup/>
- [10] (2009)Definition of regular expression. Dostopna na:
http://searchsoftwarequality.techtarget.com/sDefinition/0,,sid92_gci517740,00.html
- [11] (2009)What is Xpdf? Dostopno na:
<http://www.foolabs.com/xpdf/about.html>
- [12] (2009)Flash Paper Convert Service. Dostopna na:
<http://www.designs-services.com/Flashpaper-Service/Flashpaper-Convert.htm>