

Specifikacija projekta za ispit iz predmeta Skript jezici - školska 2017/2018. godina

projekat važi u januarskom, februarskom, septembarskom i oktobarskom roku

Naslov projekta: Veb aplikacija za pretraživanje knjiga

Kratak opis: Implementirati veb aplikaciju za pretraživanje bibliotečkog fonda jedna biblioteke. Prvi deo je import podataka o knjigama iz fajla i priprema prefiksa za pretraživanje. Podaci o knjigama su dati u fajlu po UNIMARC formatu (evropski standard za opis bibliotečke građe - knjiga, časopisa,...). Pretraživanje se vrši po prefiksima i potrebno je učitati iz posebnog fajla mapiranje prefiksa na elemente UNIMARC formata, kao i nazive prefiksa. Veb aplikacija za pretraživanje treba da podrži napredne funkcionalnosti pretraživanja po više kriterijuma koji se spajaju logičkim operatorima, dodavanje operatora, čuvanje istorije upita u sesiji, učitavanje upita iz sesije, korišćenje džoker znakova u upitu. Prikaz rezultata pretrage treba da podrži opcije za prikaz svih podataka i skraćeni prikaz.

Sa sledećeg linka možete preuzeti potrebne fajlove za import:

https://drive.google.com/a/raf.rs/file/d/1qukAy_Z14G7i3qP5ZZ3Fkoq05-ELbEVL/view?usp=sharing

U ovaj paket uključeni su sledeći fajlovi:

knjige.txt - podaci o knjigama, jedan red sadrži UNIMARC zapis jedne knjige

prefiksi.txt - mapiranje prefiksa na potpolja

PrefixNames_sr.properties - nazivi prefiksa na srpskom

Projekat se sastoji iz tri dela čiji opisi su dati u nastavku.

Prvi deo - bash skripta za pregled fajlova (5 poena)

Pre importovanja podataka u bazu treba pokrenuti Bash skriptu koja proverava da li su svi fajlovi za import pripremljeni i da li su fajlovi u dobrom formatu. Skripta treba da proveri sadržaj foldera (koji ste skinuli sa linka), da ispiše sve fajlove koji se u njemu nalaze, zatim broj linija u svakom fajlu i da proveri sadržaje fajlova na sledeći način.

U fajlu sa knjigama proveriti da li svaki zapis (svaka linija) sadrži polje 200 i u njemu potpolje a. Taj podatak predstavlja naslov knjige i ukoliko nema tog podatka, ta knjiga se ne može importovati (pogledati opis strukture fajla u opisu drugog dela projekta). Zatim proveriti fajl za mapiranje prefiksa - prefiksi.txt i iz njega izlistati sve linije koje nisu u dobrom formatu, a to je dva velika slova, crtica, tri cifre i jedno malo slovo - pogledati fajl. Na kraju proveriti fajl sa nazivima prefiksa - PrefixNames_sr.properties koji treba u svakoj liniji da ima dva velika slova, znak '=' i iza njega proizvoljan broj karaktera. Za sve tri provere ispisati redne brojeve linija, same linije i ukupan broj nepravilnih linija.

Pripremiti test primer u kom su bar po dve linija iz svakog fajla u pogrešnom formatu.

Drugi deo - import podataka o knjigama i priprema prefiksa za pretraživanje (15 poena) - Python

Napisati program u Python-u koji podatke iz fajlova importuje u relacionu bazu (ko želi može da koristi i neku NoSQL bazu). Osmisliti šemu baze (ili strukturu za NoSQL) kojom mogu da se predstavljaju podaci o jednoj knjizi, podaci o mapiranju prefiksa i nazivi prefiksa.

Podaci o knjizi su u fajlu knjige.txt dati tako da svaki red sadrži podatke o jednoj knjizi u UNIMARC formatu koji je objašnjen u Prilogu 1. Polje je označeno sa tri cifre a ispred svakog polja (osim prvog) stoji znak pod brojem 30 iz ASCII tabele (u Notepad++ se prikazuje kao RS). Posle oznake polja rezervisana su dva mesta za indikatore, mogu biti prazna ili sa nekom vrednošću indikatora i sledi lista potpolja koja pripadaju polju. Oznaka potpolja počinje karakterom 31 iz ASCII tabele (u Notepad++ se prikazuje kao US), zatim ide jedan karakter koji je naziv potpolja i iza njega sadržaj potpolja.

Fajl za mapiranje sadrži naziv prefiksa, crticu i oznaku polja i potpolja na koje se mapira, a značenje je sledeće, na primer mapiranje TI-200a znači da će string unet u prefiks TI tražiti u potpolju a polja 200.

Fajl sa nazivima prefiksa sadrži za svaki prefiks labelu na srpskom, ako za neki prefiks nema labelu, koristiti originalni naziv prefiksa iz fajla za mapiranje.

Treći deo - veb aplikacija (25 poena) - PHP i JavaScript

Implementirati veb aplikaciju za pretraživanje knjiga korišćenjem jezika PHP i JavaScript, a dozvoljeno je korišćenje bilo kog frejmworka za ova dva jezika.

Pretraživanje se vrši preko forme koja je slična formi datoj u prilogu 2, ali je potrebno napraviti moderniji izgled komponenti. U ovakvoj formi se kreira upit koji je niz prefiksa spojen logičkim operatorima and i or. Na početku forma treba da ima dva operanda (dva prefiksa) i napraviti dugme kojim se dodaju novi operandi. Prefiksi se biraju iz kombo boksa koji sadrži sve prefikse iz učitanih fajlova i to nazive prefiksa ili originalnu oznaku prefiksa, ako nije dat naziv u fajlu.

Korisnik aplikacije unosi kriterijume za pretragu, a to su stringovi za prefiks i bira logičke operatore. Ne mora se uneti string za svaki prefiks.

Kada se pretraži baza prikazuju se sve knjige koje zadovoljavaju upit, obezbediti dve vrste prikaza, jedan skraćeni u kom se prikazuju samo naslov, autori, izdavač, godina izdanja, mesto izdanja i puni prikaz u kom se prikazuje pun UNIMARC zapis (kao u prilogu 1).

Obezbediti funkcionalnost pregleda svih upita iz tekuće sesije (osmisliti format za predstavljanje upita) i ponovnog učitavanja već poslatog upita u formu za pretragu.

Bilo bi dobro da arhitektura aplikacije bude osmišljena tako da jednostavno može da se zameni serverska strana, znači da se upiti skladište u neki json string koji se šalje nekom serveru za pretragu.

Prilog 1 - Objašnjenje UNIMARC standarda

UNIMARC format propisuje pravila za čuvanje podataka o knjizi (ili časopisu, članaku, karti, rukopisu i ostalim vrstama građe). Podaci o knjizi se smestaju u strukturani oblik koji se sastoji od polja i potpolja. Polja se označavaju sa tri cifre, a potpolja sa jednim karakterom. Polja sadrže potpolja. Svaki podatak o knjizi se unosi u određeno potpolje prema pravilima koji propisuje standard. Polja imaju i dva indikatora koji se pišu pored naziva polja, oznaka # znači da indikator nije definisan.

Na primer naslov knjige se unosi u 200a (ova oznaka znači da se naslov unosi u potpolje a polja 200), autor knjige se unosi po pravilu 700a sadrži prezime, a 700b ime. U 205a unosi se izdanje, i 210c izdavač, u 100c i 210d godina izdanja, itd.

U nastavku je dat primer UNIMARC zapisa jedne knjige. Naslov je "Concepts of Programming Languages" (200a). Autor ove knjige je Robert W. Sebesta (700b i 700a), a međunarodni standardni broj za knjigu (ISBN broj) je 0-8053-7133-8 (010a). Jezik publikacije je engleski (101a) a zemlja izdavanja SAD (102a).

```
001 ## [7]ba[a]i[b]a[c]m[d]0
010 ## [a]0-8053-7133-8
100 ## [b]d[c]1996[h]scr
101 0# [a]eng
102 ## [a]usa
105 ## [a]a
200 0# [a]Concepts of Programming Languages[f]Robert W.
Sebesta
205 ## [a]3. izd.
210 ## [a]Reading [etc.][c]Addison-Wesley Publishing
Company[d]1996
215 ## [a]xv, 634 str.[c]ilustr.[d]24 cm
700 #1 [4]070[a]Sebesta[b]Robert W.
992 ## [b]crsale9701-old
```

Prilog 2 - Skica forme za pretraživanje

Autor ▼		OR ▼
Naslov ▼		AND ▼
Ključne reči ▼		OR ▼
Izdavač, distributer itd. ▼		AND ▼
Godina izdavanja ▼		
Pošalji		
