Univerzitet Mediteran

Fakultet za informacione tehnologije



**Zavrsni rad iz predmeta**

**Inteligentni sistemi**

**Tema: Ekstrakcija strukturnih podataka o eventima**

**Student: Predmetni profesor:**

Misurovic Milos 28/09 Prof. dr Jelena Jovanovic

**Predmetni asistent:**

Doc. dr Ivana Ognjanovic

Podgorica , 2013

Sadržaj

[Opis aplikacije 1](#_Toc358369001)

[Domenski model 1](#_Toc358369002)

[Rešenje 2](#_Toc358369003)

[Crawler 2](#_Toc358369004)

[Kreiranje RDF Model 2](#_Toc358369005)

[Pretrazivanje eventa 2](#_Toc358369004)

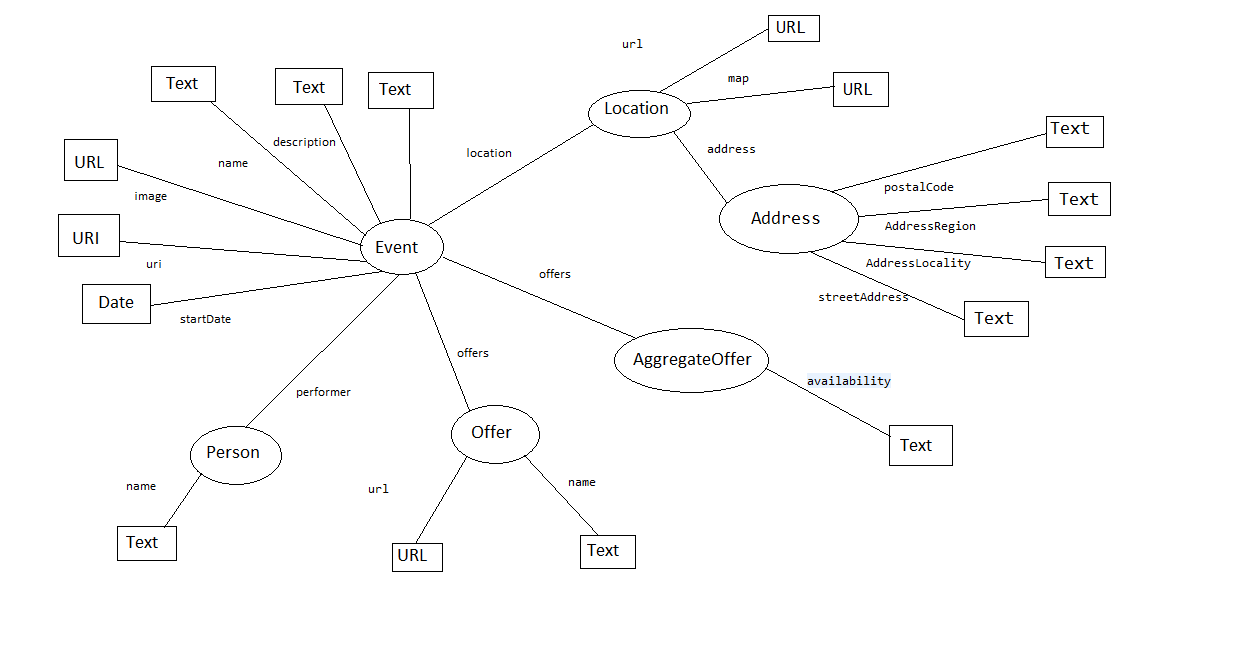
[Tehnička realizacija aplikacije 3](#_Toc358369006)

**Opis aplikacije**

Aplikacija Eventful omogucava prikupljanje i ekstrakciju podataka o dogadjajima ( event-ima) sa Web sajta “eventful.com” i cuva podatke u fajlu. Takodje omogucava i prebacanje podataka iz RDF/XML fajla u objekto orjentisane klase koje smo napravili u skladu sa Schema.org vokabularom. Model koji smo napravili cuvamo u fajlu in ad njim mozemo izvrsavat sparql upite kako bi pretrazivali evente po nekim zadatim atributima.

**Domenski model**

Domenski model opisan je preko grafa (Slika 1). Model je zasnovan na modelu za opisa eventa definisanom u okviru Schema.org vokabulara.



Slika 1 - domenski model

**Resenje Eventful aplikacije**

Aplikacija Eventful posjeduje tri bitne funkcionalnosti prestavljne u nastavku teksta.

**Crawler**

Jedna od bitni stavki u ovom program je Crawler koji nam sluzi da na zadatu Web stranicu prikupi sve Url koje su povezane sa tom stranicom kako bi mi mogili da izvucemo sto vise podataka sa web stranice. Za Crawler sam koristio crawler4j , otvorenog koda web crawler za javu.

**Kreiranje RDF Modela**

Druga bitna stavka u ovom programu je kreiranje RDF Modela definisane u okviru Schema.org vokabularu. Punjenje objektno orjentisanih klasa kojim smo predstavili Objekte koje cemo sacuvati u modelu smo izvrsili tako sto smo anizirali liniju po linij RDF/XML fajla i punili objekte koje smo kasnije sacuvali u modelu. Klase koje smo napravili su da bi kreirali RDF Model su :

* Thing
* Event
* Location
* Address
* Person
* Offer
* AggregateOffer

**Pretrazivanje dogadjaja (eventa)**

Treca i takodje bitna stavka u ovom programu je pretraga i analiza dogadja pomocu odredjenih upita. Metode kojim mozemo pretrazivati evente su pretrazivanje eventa po nazivu ulice , drugi je pretrazivanje po datumu kad se event dogadja a treci je po mejstu gdje se event odrzava;

**Tehnicka realizacija aplikacije**

Aplikacija je napisana u objektnom orjentisano jeziku Java. Za pretrazivanje URL sa odredjene web stranice sam koristi crawler4j kojeg sam skinu sa <https://code.google.com/p/crawler4j/> .Za cuvanje podatak sam koristi obican fajl u kojem cuvam Rdf scheme , a za pretragu i analizu podataka iz rdf grapha sam koristi sparql upite .