

# Mogućnost primene ontologije u razvoju Web sajta za prodaju automobila

Emir Ugljanin, Amel Kolašinac, Dženan Avdić

**Sadržaj** — U ovom radu su izloženi osnovni pojmovi iz oblasti ontologije i semantičkog Web-a, kao i mogućnosti primene u poboljšanju usluga sajtova za prodaju polovnih automobila. Zbog ekspanzije broja sajtova koji pružaju sličan sadržaj i veliki broj podataka, primena novih tehnologija i softverskih rešenja treba da omogući bolje pozicioniranje preduzeća koja svoje proizvode i usluge pružaju preko interneta. U radu su izloženi osnovni principi koji su postavljeni u realizaciji projekta kao primer podrške poslovanju malim i srednjim preduzećima.

**Ključne reči** — Ontologija, semantički web, ontološke biblioteke, Web servisi.

## I. UVOD

OSNOVNI cilj ovog rada je da prikaže implementaciju ontologija na unapređivanje funkcionalnosti internet sajtova i olakšavanje pronalaska traženih informacija. Želi se prevazići trenutna ograničenost u pretraživanju web sajtova na funkcije i ponuđene mogućnosti koje je odredio programer ili administrator sajta i omogućiti korisnicima da uz ontologije dođu do odgovarajućeg traženog proizvoda ili usluge bez potrebe da budu detaljno informisani o svim njenim detaljima. Ovako definisan sajt bi omogućio svim korisnicima da bez obzira na obrazovanje i informisanost o proizvodu koji traže dođu do njega ili do nekog koji je njemu ekvivalentan po performansama i karakteristikama.

Za organizovanje i izradu ovog rešenja su korišćena najnovina rešenja iz oblasti semantičkog weba, ontologija, web servisa, ontološkog modeliranja web servisa kao i programskih jezika za pisanje aplikacija na internetu.

U radu su predstavljene osnove semantičkog weba i njegov nastanak, kao i pregled tehnologija koje se koriste u semantičkom webu, osnovne definicije ontologija i ostalih tehnologija potrebnih za njeno funkcionisanje na internetu, hronološki razvoj ontologija od prvih filozofskih shvatanja ontologija do današnjih dana kao bitnog alata za predstavljanje i deljenje znanja. Dalje su predstavljeni najpoznatiji jezici za pisanje ontologija, konstruisanje i

korišćenje ontoloških biblioteka kao i sama prednost upotrebe ontologija. Na kraju je opisana praktična primena ontologije na sajtu za prodaju automobila.

## II. SEMANTIČKI WEB

Govoreći o pretraživanju interneta i pronalaženju odgovarajućih informacija 1996, Berners-Lee [1] je posebno istakao činjenicu kako je vrlo teško pronaći određeni podatak ukoliko on nije prikazan u u obliku konkretnog semantičkog navoda. U narednim godinama on je na konferencijama WWW 7 (Brisbane, 1997) i WWW 8 (Toronto, 1998) izložio ideju globalnog Web-a sa samostalnom unutrašnjom logikom i službeno najavio sledeću fazu razvoja nazvanu semantički Web. Ključni deo čitave ideje sadržan je u semantici predmetnih metapodataka (*subject metadata*) i njihovom prikazu kroz kontekstualizovanu i mašinski "razumljivu" terminologiju. To očigledno zahteva i semantičku i sintaktičku interoperabilnost predmetnog rečnika, budući da je dobro poznato da su za temeljan i sveobuhvatan opis predmeta potrebni ne samo izolovani pojmovi već i propoziciona logika (Veltman, 2001). U tom kontekstu, postojeće mašine za organizaciju znanja (*knowledge organization systems* skr. KOS) kao što su, na primer, klasifikacioni sistemi, prepoznati su kao važni izvori strukturiranih i formalizovanih rečnika koji mogu biti izuzetno korisni u razvoju semantičkog Web-a.

Dve su razvojne tendencije na području sistema za organizaciju znanja, posebno unutar konteksta klasifikacionih struktura znanja, bitne za razvoj ideje semantičkog Web-a i vrlo će verovatno uticati na buduću upotrebu sistema za organizaciju znanja:

- 1.) standardi i programska rešenja za terminološku razmenu;
- 2.) ontološki rečnici (u obliku ontologija s područja veštačke inteligencije).

Uključujući ova dva razvojna područja, proces standardizacije se primarno usredsređuje na tehnološke okvire mreže i na daljnji napredak u razvoju mrežnih jezika za reprezentaciju sadržaja (uz korišćenje XML i XML/RDF sistema kodiranja).

Da bi se obezbedilo postojanje Semantičkog Web-a potrebno je obezbediti širok spektar standarda na kojima se Semantički Web bazira. Na slici Sl.1 je prikazana "torta" Semantičkog Web-a (*engl. Semantic Web cake*).

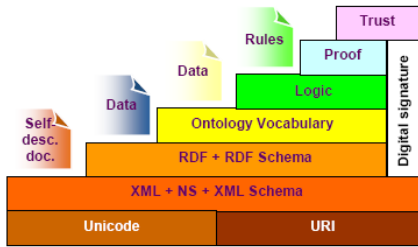
Pošto je predmet istraživanja u ovom radu prvenstveno usmeren na ontologije i ontološko inženjerstvo nećemo se zadržavati na objašnjavanju slojeva "torte" već ćemo detaljnije razmatrati ontologije.

Ovaj rad delimično je finansiralo Ministarstvo nauke Republike Srbije, Projekat tehnološkog razvoja TR 13012.

Emir Ugljanin, Fakultet Tehničkih Nauka, Državni Univerzitet u Novom Pazaru, Srbija (e-mail: emirugljanin@gmail.com).

A. Kolašinac, Fakultet Tehničkih Nauka, Državni Univerzitet u Novom Pazaru, Srbija; (telefon: 381-64-3683552; e-mail: akolisinac@np.ac.yu).

Dž. Avdić, Fakultet Tehničkih Nauka, Državni Univerzitet u Novom Pazaru, Srbija (e-mail: dzavdic@np.ac.yu).



Sl. 1

### III. ONTOLOGIJE

Filozofska definicija ontologije je učenje o biću kao bivstvujućem, kao i o njegovim opštim, fundamentalnim i konstitutivnim određenjima [2]. Ontologije se često vide kao osnovni građevni blokovi Semantičkog Web-a. One sadrže ponovo iskoristive delove znanja o specifičnim domenima. Međutim, ti delovi informacija nisu često statički, već evoluiraju tokom vremena. Promene na domenima, priladavanje različitim poslovima ili promene u konceptualizaciji zahtevaju modifikacije ontologija. Iz toga zaključujemo da je potrebna podrška za održavanje tih promena. To je jako važno u decentralizovanom i nekontrolisanom okruženju kao što je Web, gde se promene događaju bez obaveštenja.

Ontologije mogu povećati funkcionalnost Web-a na mnogo načina. Mogu poboljšati preciznost u pretraživanju Web-a – pretraživački programi mogu tražiti samo one stranice koje se odnose na određeni koncept umesto da pronalaze stranice koje koriste jednu od neodređenih ključnih riječi. Naprednije aplikacije će koristiti ontologije za povezivanje informacija na stranicama sa strukturama znanja i pravilima zaključivanja.

Unutar ideje semantičkog Web-a velika važnost pridaje se unapređenju mašinski obradivih struktura znanja koje su usko vezane uz područje inženjeringa znanja (*knowledge engineering*) i veštačke inteligencije (*artificial intelligence*).

Pojam “ontologija” postupno počinje obuhvatati čitav niz značenja i uključuje sve od taksonomija, rečnika koji se koriste u metapodacima, popisa proizvoda ili klasifikacija usluga, do rečnika baza podataka i njihovih odnosa. Ontologija u smislu formalne strukture podataka koja se koristi u izgradnji baze znanja je čak i unutar ovog područja relativno nov predmet istraživanja i datira iz 1990-ih.

Ontologija je usko povezana sa sistemima baza znanja (*knowledge based systems – KBS*) tj. ekspertnim sistemima koji su programirani da se “ponašaju inteligentno” i na taj način ili pomognu stručnjacima da izvrše svoje zadatke brže i ekonomičnije, ili u potpunosti zamene ljude u izvođenju opasnih i skupih rutinskih procedura. Takav računarski sistem potrebno je opremiti znanjem iz određenog područja tj. baze znanja (*knowledge base*) i programirati ga da vrši stalne procedure pomoću programa za zaključivanje (*inference engine*).

Kako bi se to postiglo, bazu znanja treba izgraditi na principu formalnih mašinski obradivih struktura podataka.

Sam pojam baze znanja zapravo je neformalan naziv za skup informacija u kojem je ontologija jedan od sastavnih delova. Međutim, taj skup takođe može sadržati i informacije izražene u obliku deklarativnog jezika kao što su logika (*logic*) ili pravila ekspertnog sistema (*expert system rules*), kao i informacije koje nisu formalizovane, već su izražene prirodnim jezikom ili proceduralnim kodom. Ontologija se izgrađuje tokom temeljne logičke procedure, a rezultat toga je određena klasifikaciona struktura sa jasno određenim kategorijama i pojmovnim odnosima koje se može prikazati putem “pojmovnih grafova” (*conceptual graphs*) i formatizirati na mašinski obradiv način [3].

Prikaz znanja (*knowledge representation*), kako se podrazumeva unutar područja veštačke inteligencije, bavi se širokim spektrom računarskog znanja tj. znanja izraženog putem strogih pravila logike.

Unutar konteksta semantičkog Web-a pojam ontologije može imati vrlo široko značenje koje se obično izvodi iz klasifikacione strukture i sistema upravljanja pojmovnim rečnikom koji su sastavni dijelovi svake ontologije. McGuinness (2002) nudi kategorizaciju u smislu praktične primene ontologija koja ilustrira sličnost s, np. bibliografskom klasifikacijom, a koja deli ontologije na **jednostavne i strukturirane**.

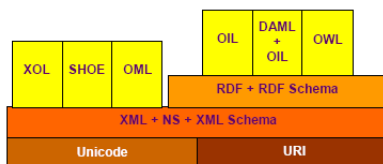
Govoreći o **jednostavnim ontologijama** McGuinness daje primere taksonomija ili jednostavnih hijerarhijskih rečnika i definiše njihovu svrhu na način koje naglašava njihovu sličnost s bibliotečkim klasifikacijama.

**Strukturirane ontologije**, s druge strane, osim mašinski čitljivih, kodiranih hijerarhijskih odnosa, sadrže informacije o pojedinim svojstvima i njihovim vrednosnim ograničenjima čime se određeni pojam stalno povezuje sa konkretnim primerom na kojeg se odnosi.

### IV. JEZICI ZA PISANJE ONTOLOGIJA

Cela infrastruktura Semantičkog Web-a je zasnovana na korišćenju standarda: za tekst (Unicode – čime se obezbeđuje standardno korišćenje predstavljanja znakova) i Uniform Resource Identifier (URI). Koncept URI ima šire značenje nego što je značenje Uniform Resource Locator (URL). Naime, URI može da predstavlja stvar kojoj se može pristupiti preko računarske mreže (npr. HTML dokument), stvar kojoj se ne može pristupiti preko računarske mreže (npr. čovek) i apstraktne koncepte koji fizički ne postoje (npr. pojam kreator Web stranice). Na sledećem sloju se nalazi definicija XML-a kao ključnog jezika za postizanje interoperabilnosti na Web-u. Korišćenjem XML-a definisani su ostali jezici Semantičkog Web-a. Primeri takvih jezika su: XML Ontology Language (XOL), Ontology Markup Language (OML), Simple HTML Ontology Extensions (SHOE). Međutim, korišćenjem XML-a definisani su Resource Description Framework (RDF) i RDF Schema (RDFS) – W3C standardi za opisivanje metapodataka i opisivanje rečnika konceptata na Web-u, a koji su zasnovani na XML-u. Zbog toga druga grupa ontoloških jezika počiva na RDF i RDFS kao i na XML-u. Primeri jezika koji su zasnovani na XML/RDF(S) su: Ontology Inference Layer, DARPA Agent Markup Language + OIL (DAML + OIL) i Web Ontology Language (OWL). Na slici Sl.2. su prikazane

relacije pomenutih jezika u odnosu na slojeve Semantičkog Web-a. Poslednji napori W3C-a koji se odnose na standardizaciju ontoloških jezika se prvenstveno odnose na one jezike koji su zasnovani na RDF(S). Zbog toga se u nastavku ovog odeljka razmatraju: XML, RDF i RDFS, kao i SHOE, OIL, DAML+OIL i OWL.



Sl. 2

#### A. Ontology Inference Layer (OIL)

Ontology Inference Layer ili Ontology Interchange Language (OIL) [4] je predlog za Web zasnovanu reprezentaciju ontologija i sloj za zaključivanje pomoću ontologija. U velikoj meri OIL se zasniva na, i uzima kao polaznu tačku RDF šeme i potpuno je zasnovan na XML sintaksi, dok je sa RDF-om u određenoj meri i isprepletan. Ovo se postiže proširivanjem definicije šeme RDFS-a dodatnim jezičkim primitivama. Treba napomenuti da pored XML-a i RDF-a, OIL poseduje još i korene u opisnoj logici (*Description Logic*) i sistemima zasnovanim na okvirima (*Frame-based systems*)

#### B. DAML Ontology library system

DAML (*DARPA Agent Markup Language*) [5] program je zvanično krenuo sa radom avgusta 2000. godine. Cilj truda DAML-a je razvitak jezika i alata koji bi omogućili stvaranje koncepta «Semantički Web». Cilj ove grupe, koja radi pod pokroviteljstvom *Defence Advanced Research Project Agency*-a je razvoj jezika i alata koji bi podržali koncept semantičkog Web-a. DAML je takođe i naziv jezika, sada već grupe jezika, razvijenih u okviru ovog programa. Kao i pominjani OIL, i DAML je jezik zasnovan na ontologijama razvijen kao proširenje XML-a i RDF-a. Da je reč o pravcima istraživanja koji neće nadalje ići odvojeno svedoči i poslednji iz grupe DAML jezika – DAML+OIL. Ovakav pristup, naravno, DAML uvrštava u najozbiljnijeg kandidata za buduće standarde. DAML je konceptualno vrlo sličan OIL-u. Razlike su mahom u nazivima komponenti. Tako npr. *slot-def* postaje *Property*, *slot-constraint* postaje *Restriction* i sl.

Ontološki bibliotečki sistem sadrži katalog ontologija razvijenih DAML-om. Taj katalog DAML ontologija je dostupan u XML, HTML i DAML formatu. Ljudi mogu poslati nove ontologije putem javnog DAML ontološkog bibliotečkog sistema.

#### C. DARPA Agent Markup Language + OIL (DAML+OIL)

DARPA Agent Markup Language + OIL (DAML+OIL) ontološki jezik je nastao spajanjem DAML-ONT (ranijeg DAML ontološkog jezika) i OIL jezika. Cilj ovog spajanja je bio da se integrišu formalni aspekti deskriptivne logike u DAML i da se izmeni DAML-ONT prema sugestijama

korisnika. Uopšte posmatrano, DAML+OIL u odnosu na DAML-ONT više naglašava preciznu semantiku korišćenjem revidirane semantike logike prvog reda (*engl. First-order logic – FOL*) i model-teorijske semantike. DAML+OIL je dosta sličan OIL-u u mnogim aspektima, ali je čvršće integrisan sa RDFS, koji samo omogućava specifikaciju jezika i njegovu serijalizaciju.

Iako zasnovanost na RDFS ima nekih prednosti u smislu ponovnog korišćenja postojeće RDFS infrastrukture i prenosivosti DAML+OIL ontologija, korišćenje RDFS za potpuno definisanje strukture DAML+OIL je dosta teško, jer za razliku od XML-a RDFS nije projektovan za definisanje precizne sintakse strukture. Rešenje ovog problema kod DAML+OIL je definisanje semantike jezika na takav način da se dá značenje svakoj ontologiji (ili delu ontologije) prema RDFS specifikaciji uključujući i “nepoznate” konstrukcije kao što su ograničenja slotova na više klasa. Kao i OIL i DAML+OIL vodi poreklo iz sistema zasnovanih na okvirima i deskriptivne logike. DAML+OIL pored modelteorijske definicije (pomoću RDFS) ima i preciznu aksiomatsku definiciju korišćenjem KIF-a.

Neke od razlika između OIL-a i DAML+OIL su zato što različito posmatraju individue, tj. DAML+OIL posmatra individue kao pojedinačne elemente u domenu od interesa, dok OIL ih posmatra kao primitivne koncepte. Ovo znači da je u DAML+OIL moguće naglasiti da su dve individue iste ili različite. DAML+OIL uključuje korišćenje tipova podataka iz XML Schema, a takođe se mogu definisati novi korisnički tipovi podataka pomoću XML Schema mehanizma za definisanje jednostavnih tipova podataka. Pored toga, DAML+OIL pravi razliku između instanci klasa (definisanih u ontologiji) i instanci tipova podataka (definisanih korišćenjem sistema tipova podataka od XML Schema). Ovo povlači i postojanje dve vrste osobina (*engl. properties*) u DAML+OIL: objektne osobine (*engl. object properties*) – kojima je opseg (*engl. range*) klasa i osobine tipa podataka (*engl. datatype properties*) – kojima je opseg neki tip podataka.

#### D. Web Ontology Language (OWL)

Web Ontology Language (OWL) [6] je semantički jezik za objavljivanje i deljenje ontologija na Web-u. OWL je razvijen kao proširenje rečnika RDF-a i izveden iz DAML+OIL ontološkog jezika. Pošto je OWL zasnovan na RDF(S) to znači da OWL ontologija predstavlja RDF graf, koji se dalje može predstaviti u obliku RDF trojki. Zbog toga je moguće za OWL koristiti slične sintaksne forme kao i za RDF. Trenutno postoji specifikacija OWL sintakse koja je zasnovana na korišćenju XML/RDF sintakse.

U razvoju OWL-a je uzeto u obzir da mora da podrži pretpostavke o otvorenom svetu i da dozvoli uvoženje i kombinovanje različitih ontologija. Neke od tih ontologija mogu čak da budu i kontradiktorne, ali nove informacije ne treba nikada da opozivaju postojeće, već smo mogu da se dodaju. Da bi se obezbedila takva kompatibilnost i u isto vreme podrška za izračunavanje i rezonovanje u konačnom vremenu sa alatima koji bi se pravili na postojećim ili budućim tehnologijama, OWL uvodi tri podjezika za različite namene: OWL Full, OWL DL i OWL Lite.

## V. PRIMER KREIRANJA ONTOLOGIJE

Postoji veliki broj sajtova koji pružaju usluge oglašavanja automobila. Google na primer nalazi 1.380.000 pogodaka za upit „auto oglaš“. Jedan od prvih na listi je sajt „autopijac.com“ [7] koji trenutno nudi oglašavanje vozila kao i pretragu prema unapred definisanim kriterijumima. Postoji veliki broj sajtova na kojima se potencijalni kupci mogu informisati o tipovima automobila, njihovoj potrošnji, kvalitetu, komforosti, održavanju itd. Svesni smo činjenice da veliki broj vozača ili potencijalnih vozača nema velikog znanja o automobilima i svim ovim detaljima od onoga što je čuo od prijatelja, kao i činjenice da mnogi potencijalni kupci nemaju vremena da prelistavaju razne forume da bi izučili sve detalje o automobilima koje hoće da kupe. Ovo rezultira čestim greškama pri kupovini vozila jer vozilo koje je kupljeno ne ispunjava uslove koji su potrebni kupcu, pa dolazimo do ponovne prodaje i kupovine vozila sve dok kupac ne dođe do vozila koje mu odgovara.

Na sajtu autopijac.com je omogućeno korisnicima da pretražuju vozila prema godištu, tipu motora, proizvođaču, modelu, dodacima, ceni, tipu vozila i još nekim vidljivim karakteristikama. Ovo je situacija i sa drugim sajtovima ovog tipa.

Da bi prikazali trenutni način funkcionisanja sajtova za prodaju automobila bez ontologija tražićemo Opel Omega sa cenom između 5000 i 6000 evra na sajtu autopijac.com preko forme za pretragu na početnoj strani kao što se vidi na slici Sl.3.

Sl. 3. Forma za pretragu sajta

Rezultat pretrage u ovom trenutku je 6 automobila marke Opel Omega čija je cena između 5000 i 6000 evra. Rezultat se može videti na slici Sl.4.

Ovakav način pretrage je odgovarajući samo za iskusne kupce koji tačno znaju šta traže. Problem nastaje ukoliko neiskusni kupac odluči da kupi Opel Omega, a ne nađe odgovarajući automobil u rezultatima. Ontologije bi omogućile da se kao rezultati prikažu i automobili koji nisu marke Opel već imaju iste ili slične karakteristike u komforosti, bezbednosti, potrošnji itd.

Ovo bi omogućilo da kompjuter sam napravi alternative za kupovinu vozila ukoliko nema vozila koje mi tražimo, a takođe bi kompjuter bio u mogućnosti da odabere vozilo prema našim potrebama umesto nas, tako što bi smo zadali kriterijume koje je potrebno da zadovoljava automobil.

Sl. 4. Rezultati pretrage sajta bez korišćenja Ontologija

Na primer, ukoliko tražimo automobil sa cenom između 5.000 i 6.000 evra, koji troši 5-6 l dizel goriva, koji ima servise u Novom Pazaru i okolini, i da taj automobil bude komforan karavan sa jakom karoserijom i kvalitetnim turbo motorom koji može da pređe stotine hiljada kilometara, kompjuter nam može predložiti VW passata b4 karavana koji ispunjava sve uslove koje tražimo.

Ontologija koja treba da unapredi rad ovog sajta i sličnih sajtova biće besplatna za korišćenje i treba da sadrži hijerarhije i opise kategorija proizvođača automobila sa svim detaljima u vezi svakog automobila. Ovi detalji bi sadržali potrošnju automobila, njegovu komforost, prosečan životni vek motora i limarije, cenu održavanja vozila kao i servisne centre za održavanje tog vozila itd.

Realizacija ove ontologije je u toku, te stoga nismo u mogućnosti da odmah realizujemo ovaj način pretrage.

## LITERATURA

- [1] <http://sh.wikipedia.org/wiki/Ontologija>
- [2] Berners-Lee, T., Fielding, R., i Masinter, L. (1998) RFC 2396: Uniform Resource Identifiers (URI): Generic syntax. Status: DRAFT STANDARD
- [3] <http://www.jfsowa.com/ontology/ontometa.htm>
- [4] Fensel, D., Van Harmelen, F., Horrocks, I., McGuinness, D. L. and Patel-Schneider, P. F., "OIL: An Ontology Infrastructure for the Semantic Web", IEEE Intelligent Systems, Vol. 16, Num. 2, March/April 2001.
- [5] "DAML – The DARPA Agent Markup Language Homepage", <http://www.daml.org>
- [6] <http://www.w3.org/TR/owl-guide/>
- [7] Besplatni auto oglasi <http://www.autopijac.com>

## ABSTRACT

This study examines basic concepts of ontology and semantic web, as well as possibility of improving car sale website services. Due to a rising number of websites that offer similar services and a great deal of data, using new technologies and software solutions is supposed to enable better positioning of companies which are offering their services via internet. This study examines basic principles we use as an example of support to small and developing companies.

## POSSIBILITY OF USING ONTOLOGY IN DEVELOPMENT OF WEBSITE FOR CAR SALES

Dženan Avdić, Amel Kolašinac, Emir Ugljanin