

Semantička anotacija obrazovnih materijala u sistemu e-obrazovanja

Jovana Dadić, Marijana Despotović-Zrakić, Zorica Bogdanović, Miloš Milutinović

Fakultet organizacionih nauka

Univerzitet u Beogradu

Beograd, Srbija

jovana@elab.rs, maja@elab.rs, zorica@elab.rs, milosm@elab.rs

Sadržaj— Informacione tehnologije se u sve većoj meri koriste u različitim procesima obrazovanja. Jedan od ključnih procesa obrazovanja zasnovanog na savremenim informaciono-komunikacionim tehnologijama, jeste proces pronalazjenja obrazovnih materijala. U ovom radu biće predložen model za semantičku anotaciju obrazovnih materijala, u cilju poboljšanja procesa pronalazjenja istih. Model semantičke anotacije zasnovan je na konceptima semantičkog veba i karakteriše ga visok nivo interoperabilnosti i adaptabilnosti. (Abstract)

Ključne reči - e-obrazovanje; semantičko opisivanje; semantički veb; (key words)

I. UVOD

Visok nivo razvoja i dostupnosti informaciono-komunikacionih tehnologija na globalnom nivou doveo je do značajnih promena u obrazovnom procesu i doprineo razvoju i primeni e-obrazovanja. Sve veći broj obrazovnih institucija organizuje nastavu koristeći neki od koncepata e-obrazovanja. Jedna od osnovnih prednosti primene navedenih koncepata jeste mogućnost implementacije različitih oblika komunikacije među učesnicima u procesu obrazovanja. Prema Despotović-Zrakić, komunikacija putem Interneta u obrazovnom procesu javlja se u nekoliko pojavnih oblika: pisani načini komunikacije, verbalni i vizuelni načini komuniciranja i pronalazjenje obrazovnih materijala [1]. Svaki od navedenih oblika komunikacije nosi određene rizike i izazove u digitalnom okruženju. Ovaj rad će se pretežno baviti procesom pronalazjenja obrazovnih materijala, kao i metodama za unapređenje ovog procesa korišćenjem tehnologija semantičkog veba.

II. TEHNOLOGIJE SEMANTIČKOG VEBa

Veb je trenutno sastavljen od velikog broja dokumenata napisanih prirodnim jezikom i namenjenih ljudima. Ljudi na žalost mogu razumeti i procesuirati samo vrlo malu količinu informacija raspoloživih na vebu. Iz tog razloga, prepoznata je potreba za mehanizmima koji bi mašinama omogućili da razumeju informacije na vebu i na taj način pomognu ljudima u procesuiranju i analiziranju podataka koji su im potrebni. Današnji računari ne razumeju prirodni jezik i ne mogu pomoći ljudima u kompleksnim zadacima kao što je analiza podataka. Semantički veb predstavlja tehnologiju koja bi ovo omogućila. Osnovni koncept semantičkog veba je sledeći: dodavanjem

specijalnih podataka – metapodataka, veb stranama tj resursima, računari će biti u mogućnosti da razumeju značenje podataka na veb stranama, i uz pomoć toga pomognu ljudima u kompleksnim zadacima koje obavljaju na vebu.

Standardima i tehnologijama semantičkog veba se primarno bavi međunarodna organizacija za standardizaciju veba W3C. W3C semantički veb definiše kao veb podataka, tj veb koji obezbeđuje jedinstveni okvir za deljenje podataka između aplikacija, preduzeća i zajednice [2]. Mnogi autori predlažu različite definicije semantičkog veba. Prema Berners-Lee, semantički veb može se definisati kao princip koji će omogućiti mašinama da razumeju semantičke dokumente i podatke [3]. Gomez Perez i Corcho definišu semantički veb kao skup tehnologija koje podrazumevaju eksplicitno definisanje znanja ugrađenog u aplikacije na vebu, integraciju informacija na inteligentan način, obezbeđivanje pristupa internetu zasnovanog na semantici, kao i ekstrakciju informacija iz teksta [4]. McIlraith semantički veb definiše kao veb servisa čije su osobine, mogućnosti, interfejsi i rezultati šifrovani u nedvosmislenom obliku koji je razumljiv mašinama [5]. Neke od novijih teorija semantički veb smatraju vebom informacija na osnovu kojih se može delati, tj vebom informacija izvedenih iz podataka putem semantičke teorije za interpretaciju simbola [6].

A. Implementacija aplikacija zasnovanih na semantičkom vebu

Sama ideja semantičkog veba može se realizovati u 4 koraka:

1. Kreiranje ontologija domena,
2. Serijalizacija kreiranih ontologija korišćenjem tehnologija semantičkog veba,
3. Kreiranje semantičkog markup-a na resursima,
4. Kreiranje servisa semantičkog veba.

Izgradnja ontologija predstavlja prvi korak u realizaciji rešenja zasnovanog na semantičkom vebu. Reč ontologija potiče iz grčkog jezika, od reči *ontos* koja znači *biti* i reči *logos*, koja znači *svet*. U filozofiji, ontologija predstavlja istraživanje samog postojanja stvari. U kontekstu računarskih nauka, ontologija predstavlja skup koncepata koji opisuju određenu oblast [7]. Jepsen detaljnije definiše ontologije, i to

kao metod predstavljanja delova znanja (ideja, činjenica, stvari) na način koji definiše odnose i strukturu koncepata u određenom domenu znanja [8]. U procesu kreiranja domenske ontologije poželjno je koristiti neku od postojećih metodologija za razvoj ontologija i logičko projektovanje. Nakon kreiranja ontologija domena, može se pristupiti i serijalizaciji kreiranih ontologija, i to korišćenjem tehnologija semantičkog veba, kao što su: RDF, RDFS i OWL. Navedene tehnologije semantičkog veba zasnovane su na XML-u, i omogućavaju zapisivanje logičkih koncepata i relacija koje postoje u ontologijama. RDF (*Resource Description Framework*) je standardni model za razmenu podataka na vebu, čiju specifikaciju definišu W3C preporuke iz 2004. godine [9]. RDF model je sličan klasičnim konceptualnim modelima kao što su model objekti-veze (MOV) i dijagram klasa. Kao i pomenuti modeli, RDF koristi iskaze u obliku subjekat-predikat-objekat, poznatije pod nazivom *triple*. RDFS se odnosi na pojam RDF *Schema-e*, tj jezika koji obezbeđuje prikaz RDF rečnika, zasnovan na XML formatu [10]. RDF Schema je korisna pre svega za definisanje klasa kojima pripadaju resursi definisani u rdf rečniku, kao i hijerarhije tih klasa u RDF. U praksi se koriste spojeni RDF i RDFS rečnici. OWL (Web Ontology Language) predstavlja grupu jezika za reprezentaciju znanja koji za razliku od prethodno navedenih RDF i RDFS podržava rad semantičkog mehanizma zaključivanja (*semantic reasoner*) [11]. Semantički mehanizam zaključivanja je softver koji može doći do zaključaka na osnovu skupa navedenih činjenica ili aksioma. U aplikacijama semantičkog veba, semantički mehanizam zaključivanja uglavnom izvodi zaključke na osnovu pravila definisanih u ontologiji. Mehanizmi zaključivanja se oslanjaju na deskriptivnu logiku (*Description Logic*), a postoji više različitih implementacija. Model ontologije sa pridruženim mehanizmom zaključivanja omogućava generisanje izvedenog grafa, tj grafa instanci sa odnosima i zaključcima među instancama.

U sledećoj fazi, neophodno je kreirati semantičku anotaciju (*semantic markup*) resursa, kako bi resursi bili čitljivi i razumljivi mašinama, tj računarima. Semantička anotacija zasnovana na ontologijama predstavlja uređenu četvorku (*a,b,c,d*) gde je *a* predmet anotacije, *c* je sama anotacija (tekst anotacije), *b* je veza kojom su povezani predmet anotacije i sama anotacija, i *d* je kontekst u kome se vrši anotacija, pri čemu *a,b,c* i *d* predstavljaju ontološke klase ili osobine, a prikazani su pomoću URI [12]. URI (*Uniform Resource Identifier*) je jedinstveni identifikator pojmova koji pokazuje na adresu na kojoj se nalazi opis pojma.

Postoje tri nivoa automatizacije procesa anotacije resursa:

1. Manuelno anotiranje
2. Polu-automatsko anotiranje
3. Automatsko anotiranje

Ontološka anotacija resursa povezuje kreirane ontologije sa delovima ili celokupnim resursom, i omogućavaju zaključivanje o sadržaju resursa. Na osnovu serijalizovanih ontologija, moguće je takođe kreiranje i semantički anotiranih veb servisa. Semantički anotirani veb servisi, ili semantički veb servisi predstavljaju mašinama razumljive veb servise, pogodne za automatizovanu kompoziciju veb servisa [5].

B. Primena tehnologija semantičkog veba u e-obrazovanju

Prema Moore & Kearsley elektronsko obrazovanje, predstavlja kompleksan sistem koji uključuje sledeće elemente [13]:

- Učenje na daljinu (*distance learning*) i predavanja na daljinu (*distance teaching*), pri čemu su procesi učenja i predavanja odvojeni vremenski i prostorno,
- Jedinice učenja u raznim oblicima (štampani materijali, audio vizuelni i sl.),
- Proces učenja koji može da bude individualni i grupni,
- Tutorski rad kombinacijom raznovrsnih formi "face-to-face" komunikacija korišćenjem medija,
- Interaktivnost i postizanje sinergijskog efekta u radu sa grupom studenata.

Jedinice učenja, kao jedan od bitnih elemenata elektronskog obrazovanja, predmet je istraživanja. Postoji veći broj definicija koje se odnose na pojam jedinica učenja. Navedene su neke od značajnijih definicija:

- Jedinica učenja se može posmatrati digitalni resurs za učenje koji je namenjen jednom cilju učenja i koji se može iskoristi u različitim kontekstima [14].
- Jedinica učenja je svaki digitalni resurs koji može biti korišćen za podršku učenju [15].

Nekoliko organizacija predlaže standarde kreiranja jedinica učenja za korišćenje u e-obrazovanju. Dve najčešće korišćene grupe standarda koje se primenjuju u e-obrazovanju su SCORM i LOM. SCORM (*Sharable Content Object Reference Model*) je skup standarda i specifikacija za e-obrazovanje na vebu čiji je cilj povećanje interoperabilnosti aplikacija e-obrazovanja [16]. LOM (*Learning Object Metadata*) standard predlaže sintaksu i semantiku metapodataka kojima se jedinice učenja opisuju [17].

Prema Aroyo & Dicheva, može se tvrditi da savremeni sistemi e-obrazovanja moraju obezbediti 3 ključna svojstva, i to [18]:

- Interoperabilnost (*interoperability*)
- Mogućnost ponovnog korišćenja (*reusability*)
- Deljivost (*shareability*)

Raznovrsnost i distribuiranost jedinica učenja izazvali su probleme interoperabilnosti između sistema, probleme u pronalazenju odgovarajućih materijala i nemogućnosti višestrukog korišćenja nastavnih materijala.

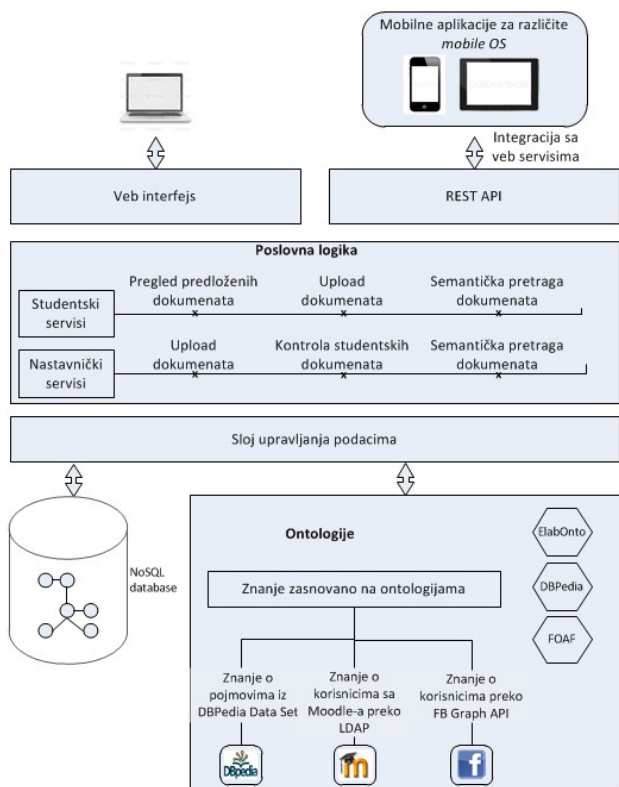
Mogućnost ponovnog korišćenja je vrlo važan aspekt jedinice učenja, koji se sve češće u literaturi spominje kao ključna osobina jedinice učenja. Zbog toga je vrlo važno obezbediti odgovarajuće metapodatke koji se pripisuju jedinici učenja, a koji opisuju samu jedinicu učenja i na taj način omogućavanju njeno iskorišćavanje u različitim kontekstima i različitim scenarijima učenja.

Primena tehnologija semantičkog veba u procesima e-obrazovanja pruža dobru osnovu za ostvarenje prethodno

navedenih svojstava savremenih sistema e-obrazovanja. Devežić tvrdi da su tehnologije semantičkog veća jedan od ključnih faktora u razvoju obrazovnih zajednica, kao i da su značajne oblasti istraživanja u e-obrazovanju vezane za primenu semantičkih tehnologija [19].

III. MODEL SEMANTIČKE ANOTACIJE OBRAZOVNIH MATERIJALA

Na Sl. 1 je dat predlog strukture sistema za rad sa semantički anotiranim dokumentima e-obrazovanja, prilagođen za Laboratoriju za elektronsko poslovanje, na Fakultetu organizacionih nauka.



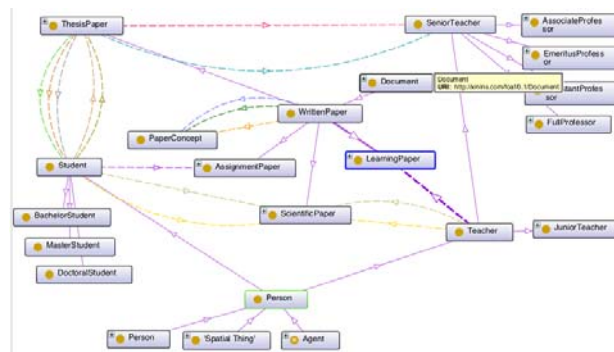
Slika 1. Predlog strukture sistema za rad sa semantički anotiranim dokumentima e-obrazovanja

Na sloju podataka, za skladištenje podataka se predlaže korišćenje No-SQL baze podataka, karakteristične po svojoj prilagodljivoj strukturi i efikasnom sistemu pretrage. No-SQL baze podataka u većini slučajeva brže procesuiraju podatke, pre svega zato što su šeme podataka jednostavnije [20]. Imajući u vidu da je *join* operacija u relacionim bazama veoma vremenski zahtevna, naročito ako se radi o *join*-u po spoljnjem ključu, No-SQL baze su izuzetno pogodne za nestruktuirane podatke koji su povezani velikim brojem veza [21]. U predloženom modelu, podaci iz No-SQL baze će biti struktuirani uz pomoć ontologija, i to:

- Domenska ontologija - ontologija koja se odnosi na konkretnu situaciju i oblast znanja u kojoj će se sistem primenjivati.

- DBPedia ontologije - najveći izvor struktuiranih informacija sa Wikipedia-e i.
- FOAF ontologije (*Friend of a friend*)[22] – ontologija za opis osoba i veza među osobama.

Kao domenska ontologija za model sa Sl. 1 razvijena je ontologija elabOnto prilagođena radu Laboratorije za elektronsko poslovanje, na Fakultetu organizacionih nauka. Pri kreiranju elabOnto korišćena je metodologija za razvoj ontologija sa univerziteta Stanford [23]. Ključni koncepti koji se javljaju u elabOnto ontologiji su: Student, Teacher, Written Paper i Paper Concept. Neke od klasa elabOnto ontologije nasleđuju klase iz ontologija višeg nivoa FOAF i DC (*Dublin Core*)[24], što doprinosi interoperabilnosti sistema. Uprošeni prikaz ontologije elabOnto prikazan je na Sl. 2.



Slika 2. Uprošeni prikaz elabOnto ontologije

Sami podaci, koji će biti struktuirani uz pomoć navedenih ontologija, i smešteni u No-SQL bazu, preuzimaju se iz 4 izvora:

1. Sa aplikativnog nivoa od korisnika,
2. Sa društvene mreže Facebook, koristeći Facebook Graph API – podaci o korisnicima,
3. Sa LMS (*Learning Management System*) platforme Moodle LMS, koja je već integrisana u proces nastave u Laboratoriji za elektronsko poslovanje, koristeći LDAP protokol,
4. Iz DBPedia data set [25] – podaci o pojmovima iz dokumenata.

Sloj upravljanja podacima u sistemu treba da obezbedi mapiranje podataka iz baze podataka u instance navedenih ontologija. Na sloju poslovne logike u sistemu nalazi se implementacija procesa studentskih i nastavničkih servisa. Pristup svim navedenim servisima omogućen je preko veb interfejsa, pomoću koga će i studenti i nastavnici imati pristup sistemu. Osim veb interfejsa, dizajniran je i REST API na osnovu koga je omogućena integracija sa veb servisima. REST (*Representational State Transfer*) je tip softverske arhitekture za distribuirane sisteme koji pokazuje izuzetne performanse po pitanjima skalabilnosti, generalizacije interfejsa i integriteta komponenti sistema [26]. Korišćenje veb servisa je izuzetno pogodno u sistemima e-obrazovanja, pre svega zato što omogućava jednostavno korišćenje servisa sistema u aplikacijama razvijenim u različitim okruženjima. Na ovaj

način, servisi sistema se mogu obezbediti korisnicima u različitim mobilnim aplikacijama kao što su: Android aplikacija, Windows Phone aplikacija i iOS aplikacija, bez potrebe za ponovnom implementacijom servisa.

IV. ZAKLJUČAK

Savremene Internet tehnologije, poput tehnologija semantičkog veba, u značajnoj meri doprinose unapređenju procesa komunikacije između učesnika e-obrazovanja. U radu je dat predlog modela za semantičku anotaciju dokumenata u sistemu e-obrazovanja, kao i predlog opšte strukture sistema za rad sa semantički anotiranim dokumentima e-obrazovanja. Takođe je kreirana domenska ontologija za opis dokumenata u sistemu e-obrazovanja Laboratorije za elektronsko poslovanje na Fakultetu organizacionih nauka. Predloženi sistem je u skladu sa troslojnom arhitekturom, a oslanja se na savremene koncepte struktuiranja podataka uz pomoć ontologija.

Budući fokus istraživanja biće pre svega na kreiranju pilot sistema kako bi bilo moguće izvršiti analizu efikasnosti predloženog modela. Biće izvršena detaljna analiza postojećih dokumenata u sistemu, kako bi domenska ontologija elabOnto bila dodatno prilagođena specifičnostima dokumenata. Takođe, biće ispitana mogućnost primene dinamičkih taksonomija u klasifikaciji jedinica učenja u sistemu, kao i mogućnost primene metoda procesiranja prirodnog jezika (NLP – *Natural Language Processing*) u procesu anotacije jedinica učenja.

LITERATURA

- [1] M. Despotovic-Zrakić, "Razvoj Metoda Poslediplomskog Obrazovanja na Daljinu Zasnovanog na Internet Tehnologijama," Univerzitet u Beogradu, 2006.
- [2] W3C, "W3C Semantic Web," 2011. [Online]. Available: <http://www.w3.org/standards/semanticweb/>.
- [3] T. Berners-lee, J. Hendler, and O. Lassila, "The Semantic Web will enable machines to," *Scientific American*, no. May, pp. 1–4, 2001.
- [4] A. Gómez-pérez, O. Corcho, and U. P. De Madrid, "Ontology Languages for the Semantic Web," *IEEE INTELLIGENT SYSTEMS*, vol. 4, no. 1, pp. 54–60, 2002.
- [5] S. a. McIlraith, T. C. Son, and H. Zeng, "Semantic Web services," *IEEE Intelligent Systems*, vol. 16, no. 2, pp. 46–53, Mar. 2001.
- [6] N. Shadbolt, W. Hall, and T. Berners-lee, "The Semantic Web Revisited," *IEEE Intelligent Systems*, vol. 21, no. 3, pp. 96–101, May 2006.
- [7] M. Uschold and M. Gruninger, "Ontologies: Principles, Methods and Applications," *Knowledge Engineering Review*, vol. 11, no. 2, pp. 93–136, 1996.
- [8] T. C. Jepsen, "Just What Is an Ontology, Anyway?," *IT Professional*, vol. 11, no. October, pp. 22–27, 2009.
- [9] W3C, "RDF," 2004. [Online]. Available: www.w3c.org/RDF.

- [10] W3C, "RDF Schema," 2004. [Online]. Available: <http://www.w3.org/2001/sw/wiki/RDFS>.
- [11] W3C, "OWL Web Ontology Language," 2004. [Online]. Available: <http://www.w3.org/TR/owl-features/>.
- [12] E. Oren, K. H. Moller, S. Scerri, S. Handschuh, and M. Sintek, "What are Semantic Annotations?"
- [13] M. G. Moore and G. Kearsley, *Distance Education: A Systems View*. Wadsworth Publishing Company, 1996.
- [14] P. Mohan, S. Augustine, and C. Brooks, "Learning Objects on the Semantic Web The University of the West Indies," in *3rd IEEE International Conference on Advanced Learning Technologies*, 2003.
- [15] D. A. Wiley, *Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy*, vol. 2830, no. 435. 2002, pp. 3–23.
- [16] Rustici, "SCORM," 1996. [Online]. Available: <http://scorm.com>.
- [17] IEEE, "LOM," 2002. [Online]. Available: <http://ltsc.ieee.org/wg12/index.html>.
- [18] L. Aroyo and D. Dicheva, "The New Challenges for E-learning: The Educational Semantic Web Interoperability of WBES," *Educational Technology & Society*, vol. 7, no. 4, pp. 59–69, 2004.
- [19] V. Devedžić, *Semantic web and education*. New York: Springer, 2006.
- [20] N. Leavitt, "Will NoSQL Databases Live Up to Their Promise?," *Computer*, vol. 43, no. 2, pp. 12–14, Feb. 2010.
- [21] L.-Y. Ho, J.-J. Wu, and P. Liu, "Distributed Graph Database for Large-Scale Social Computing," in *2012 IEEE Fifth International Conference on Cloud Computing*, 2012, pp. 455–462.
- [22] FOAF-Project, "FOAF," 2000. [Online]. Available: <http://www.foaf-project.org/>.
- [23] N. F. Noy and D. L. McGuinness, "Ontology Development 101: A Guide to Creating Your First Ontology," CA, 2000.
- [24] Dublin Core Metadata Initiative, "DC," 2000. [Online]. Available: <http://dublincore.org/>.
- [25] "DBPedia Data Sets," 2007. [Online]. Available: <http://wiki.dbpedia.org/Datasets>.
- [26] R. T. Fielding, "Architectural Styles and the Design of Network-based Software Architectures," University of California, 2000.

ABSTRACT

Information technologies are increasingly being used in various educational processes. Learning materials discovery is one of the key educational ICT based processes. A model for semantic annotation of learning materials, that will enhance the materials discovery process, will be proposed. The model for semantic annotation is based on semantic web concepts and is characterised by high interoperability and adaptability level.

SEMANTIC ANNOTATION OF LEARNING MATERIALS IN E-EDUCATION SYSTEMS

Jovana Dadić, Marijana Despotović-Zrakić, Zorica Bogdanović, Miloš Milutinović