

Upotreba semantičkog pretraživača nad registrom za akutni koronarni sindrom

Pavle Kostić, *MSc E.E.*, Suzana Blesić, *PhD*, Zorana Vasiljević, *MD, PhD* Nataša Mickovski-Katalina, *MD*, Jelica Grujić Milanović, *PhD*, Siniša Pavlović, *MD, PhD*, Slađan Milanović, *PhD*

Sadržaj — Ovaj rad predstavlja upotrebu semantičkog pretraživača (SP) nad registrom za akutni koronarni sindrom (AKS) u okviru sistema za upravljanje znanjem (SZUZ). SP je deo sistema za podršku kliničkom odlučivanju i koristi se kao pomoć u donošenju odluke u toku kliničkih procesa vezanih za zbrinjavanje i lečenje osoba obolelih od AKS.

Ključne reči — Semantički pretraživač, registar, akutni koronarni sindrom, upravljanje znanjem, kardiologija.

I. UVOD

SEMANTIČKO pretraživanje nastoji da poboljša tačnost pretrage razumevajući nameru pretraživača i kontekstualno značenje termina koje se pojavljuje u prostoru pretraživanja, nezavisno da li na Webu ili unutar zatvorenog sistema, generišući više relevantnih rezultata [1]. Polazeći od definicije semantičkog pretraživanja, a imajući u vidu potrebe lekara u koronarnim jedincama (KJ) da im se omogući lakše pronalaženje sličnih slučajeva iz Registra za akutni koronarni sindrom (REAKS) u ovom radu je opisan semantički pretraživač (SP) kao deo sistema za podršku kliničkom odlučivanju (SZPKO). U nastavku je dat pregled osnovnih pojmova i definicija

Ovaj rad je nastao kao rezultata rada na projektu, iz oblasti Tehnološkog razvoja koje finansira Ministarstvo za nauku i tehnološki razvoj Republike Srbije, broj TR13010-A, pod nazivom „Sistem za upravljanje znanjem - primena u kardiologiji“.

Pavle Kostić, Univerzitet u Beogradu, Institut za medicinska istraživanja, Laboratorija za biomedicinski inženjering i telemedicinu, PO BOX 102, 11129 Beograd, Srbija (telefon: 381-11-2685788, e-mail: pkostic@imi.bg.ac.rs)

Suzana Blesić, Univerzitet u Beogradu, Institut za medicinska istraživanja, Laboratorija za biomedicinski inženjering i telemedicinu, PO BOX 102, 11129 Beograd, Srbija (telefon: 381-11-2685788, e-mail: suzana@imi.bg.ac.rs)

Zorana Vasiljević, Univerzitet u Beogradu, Medicinski fakultet, Katedra za internu medicinu, Dr Subotića 8, 11000 Beograd, Srbija (telefon: 381-11-3615552, e-mail: zoranav@eunet.rs)

Nataša Mickovski-Katalina, Institut za javno zdravlje Srbije „Dr Milan Jovanović - Batut“, Dr Subotića 5, 11000 Beograd, Srbija (telefon: 381-11-2684 566, e-mail: natasa_m_katalina@batut.org.rs)

Jelica Grujić Milanović, Univerzitet u Beogradu, Institut za medicinska istraživanja, Laboratorija za kardiovaskularnu fiziologiju, PO BOX 102, 11129 Beograd, Srbija (telefon: 381-11-2685788, e-mail: jeca@imi.bg.ac.rs)

Siniša Pavlović, Univerzitet u Beogradu, Medicinski fakultet, Katedra za internu medicinu, Dr Subotića 8, 11000 Beograd, Srbija (telefon: 381-11-3615617, e-mail: pavlosini@yahoo.com)

Slađan Milanović, Univerzitet u Beogradu, Institut za medicinska istraživanja, Laboratorija za biomedicinski inženjering i telemedicinu, PO BOX 102, 11129 Beograd, Srbija (telefon: 381-11-2685788 e-mail: sladjan.milanovic@imi.bg.ac.rs)

vezanih za AKS i Registar za AKS. Opisana je i veza vodiča dobre kliničke prakse i Sistema za podršku kliničkom odlučivanju, kao i veza sa SZUS u okviru koga je realizovan SP. Opisana je platforma na kojoj je realizovan SP, funkcionalni zahtevi SP, kao način realizacije aplikacije SP. Izvedeni su i zaključci o mogućnosti upotrebe SP.

II. AKUTNI KORONARNI SINDROM

Dosta često oboljenje koje obuhvata oko 50% svih kardiovaskularnih bolesti je AKS. AKS obuhvata spektar kliničkih prezentacija, koje, prema kliničkom nalazu, elektrokardiografskoj slici ST segmenta i prema nalazu kardiospecifičnih enzima može da se posmatra kao akutni infarkt miokarda (AIM) sa ST elevacijom (STAIM), AIM bez ST elevacije (NSTAIM) i kao nestabilna angina pektoris (NAP) [2]-[3]. Blagovremeno postavljanje dijagnoze i lečenje savremenim metodama su značajno povećali prognozu bolesnika [4].

III. VODIČI DOBRE KLINIČKE PRAKSE I SISTEMI ZA PODRŠKU KLINIČKOM ODLUČIVANJU

U oblasti medicine definisani su vodiči dobre kliničke prakse (*Good Clinical Practice* - GCP) uglavnom kao „sistematski razvijane instrukcije za pomoć lekaru i pacijentu pri odlučivanju o odgovarajućoj zdravstvenoj zaštiti za specifične kliničke okolnosti“ [5]. Dok se kompjuterizovani GCP smatraju kritičnim komponentama medicine zasnovane na dokazima [6]-[8] njihov pravi potencijal tek treba da bude realizovan [9].

Glavni izazov je integracija modela vodiča u kliničke poslovne procese za generisanje pacijent-centričnih personalizovanih preporuka i akcija.

Sistemi za podršku kliničkom odlučivanju bazirani na vodičima se pojavljuju kao paradigma za pomoć u smanjenju grešaka, smanjenju troškova i poboljšanju kvaliteta u medicini zasnovanoj na dokazima.

Funkcije sistema za podršku kliničkom odlučivanju mogu se kategorizovati na oblast konteksta, znanja i izvora podataka, sisteme za podršku odlučivanju, isporučivanju informacija i radne procese [10]. Za uspeh implementacije u kompleksno okruženje zdravstvenih IS, kompjuterizovani GCP i druge glavne tehnologije za podršku kliničkom odlučivanju treba da budu osnova platforme sistema za upravljanje znanjem koja uključuje standarde za terminologiju (*Systematic Nomenclature of Medicine – Clinical Terms (SNOMED-CT)*; *Logical Observations*,

Identifiers, Names, and Codes (LOINC); Unified Medical Language System (UMLS); General Architecture for Languages, Encyclopedias, and Nomenclatures in Medicine (GALEN)), razmenu podataka (Health Level Seven (HL7) za generalne zdravstvene informacije; Digital Imaging and Communications in Medicine (DICOM) za medicinske slike; ISO/IEEE 11073 za podatke sa medicinskih uređaja) i EHR.

IV. SISTEM ZA UPRAVLJANJE ZNANJEM

SZUS je realizovan kao (1) Registar (baza podataka sa pratećom aplikacijom) za Akutni Koronarni sindrom (REAKS), (2) Modulom za poslovno izveštavanje i analitiku (Business Intelligence - BI) i (3) Modul za semantičko pretraživanje. Sve delovi su eksponirani preko Portala koji nalazi na adresi <http://www.kardionet.org/> [17]. U ovom radu je detaljnije opisan SP. SP je razvijen nad *on-line* verzijom REAKS kojoj se pristupa preko [17] u čiju su bazu importovani podaci za period 2002.-2008. godina.

V. CILJ

Cilj razvoja SP bio je da se za potrebe lekara u KJ napravi aplikacija koja će im omogućiti lakše pronalaženje sličnih slučajeva iz REAKS-a.

Funkcionalni zahtevi SP su bili:

- Razvijanje semantičkog meta-modela za preprezentaciju znanja koji se nalazi u registru
- Razvijanje portleta za menadžment meta-modela znanja i semantičku pretragu.

VI. SEMANTIČKI PRETRAŽIVAČ

Za realizaciju SP je odabrana D2RQ platforma i komponente iz kojih se sastoji. U nastavku je opisan način realizacije SP.

A. Jena framework

Jena je Javin okvir (framework) za kreiranje aplikacija Semantičkog Web-a. Obezbeđuje programsko okruženje za RDF, RDFS i OWL kao i za SPARQL (upitni jezik za RDF) i uključuje mehanizam zaključivanja zasnovan na pravilima. Jena je skup Javinih klasa i paketa, programa i alatki za parsiranje, kreiranje i pretragu RDF grafova i modela. Uključuje:

- RDF API
- Čitanje i pisanje RDF u RDF/XML, N3 i N-Trojka
- OWL API
- Memorijsko i trajno skladištenje
- SPARQL upitni mehanizam (query engine)

Aktuelna verzija je 2.6.2, koju smo i mi koristili, dostupna je za preuzimanje na adresi [12]

B. SPARQL

SPARQL je upitni jezik i protokol koji služi za rad sa RDF koga je dizajnirala W3C RDF Data Access Working Group. Kao upitni jezik SPARQL je orjentisan na podatke (data-oriented) tj. može da izvršava upite samo nad podacima koji se nalaze u modelu. U samom upitnom

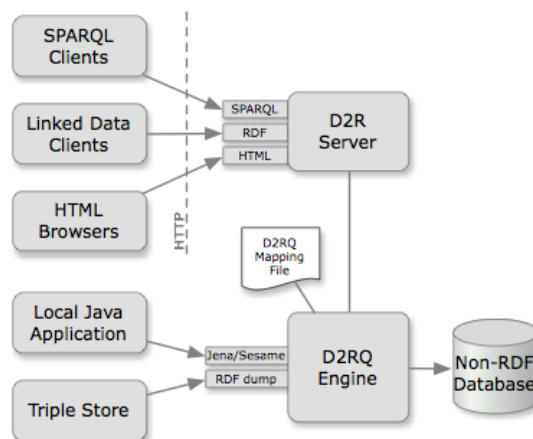
jeziku ne postoji mehanizam zaključivanja (inference). Naravno, Jena model može biti „pametan“ tako što će stvoriti neke trojke na zahtev (on-demand) koristeći OWL mehanizam zaključivanja (OWL reasoning). SPARQL ne radi ništa drugo sem da izvršava ono što je u aplikaciji definisano u formi upita i vraća skup informacija u obliku RDF grafa.

C. D2RQ platforma

D2RQ platforma se sastoji iz:

- D2RQ Mapping Language – deklarativni jezik za mapiranje koji služi za opisivanje relacija između ontologija i relacionih modela podataka.
- D2RQ Engine – plugin za Jenu i Sesame Semantic Web toolkits, koji koristi mapiranje da bi prepravio Jena i Sesame API SQL upite nad bazom i da bi prosledio rezultate do viših nivoa okvira (framework).
- D2R Server - HTTP server koji se može iskoristiti za pružanje Linked Data pogleda, HTML pogleda za debugovanje kao i SPARQL endpoint nad bazom podataka.

D2RQ platforma je prikazana na Slici 1.



Sl.1 Arhitektura D2RQ platforme

D2RQ mehanizam (D2RQ Engine) je implementiran kao Jena graf, osnovni objekat za predstavljenje informacija u okviru Jena okvira. D2RQ graf “obmotava” lokalnu bazu podataka virtuelnim RDF grafom, koji služi samo za čitanje (read-only). D2RQ mehanizam prepravlja pozive od strane Jena okvira ili Sesame API, metodu find() i SPARQL upite, tako da odgovaraju modelu date aplikacije. Rezultati ovih upita su transformisani u RDF trojke ili u SPARQL rezultate (result sets) koji se prosleđuju višim slojevima okvira.

D2RQ server je alat za prezentovanje relacione baze podataka na Semantičkom web-u. Omogućava RDF i HTML pretraživačima da pretražuju sadržaj baze podataka, i omogućava aplikacijama da izvršavaju upite nad bazom koristeći SPARQL upitni jezik. D2RQ server je napravljen nad D2RQ mehanizmom (D2RQ Engine).

D. Realizacija SP

Za realizaciju semantičkog pretraživanja koristili smo prethodno opisane tehnologije, Jena framework nam je

poslužio kao osnova za predstavljanje znanja koje se nalazi u bazi. Da bi mogli da koristimo podatke iz postojeće baze a da ne moramo da pravimo novu bazu podataka koja bi bila podobna za korišćenje u semantičkim aplikacijama koristili smo D2RQ platformu. Ona nam je omogućila da napravimo jedan sloj iznad postojeće relacione baze koji nam omogućava da pretražujemo bazu kao da je u pitanju graf i to graf koji možemo samo čitati, što je nama bilo dovoljno da bi smo ispunili projektni zadatak. Da bi smo mogli da koristimo taj sloj morali smo prvo da kreiramo fajl za mapiranje, u okviru D2RQ postoji alat koji olakšava taj, izuzetno obiman posao mapiranja tabela iz baze i njihovih veza u klase u semantičkom modelu, taj alat je generate-mapping (korišćenjem D2RQ Mapping Language).

Kao rezultat procesa mapiranja dobijen je fajl asc-mapping.n3. Kao ilustracija, u nastavku, je prikazan deo fajla za mapiranje koji se odnosi na tabelu EPISODEOF CARE.

```
# Table ACS.EPISODEOF CARE
<jdbc:oracle:thin:@imiserver:1521:imldb/vocab#ACS
.EPISODEOF CARE> a d2rq:ClassMap;
  d2rq:dataStorage map:database;
  d2rq:uriPattern
"jdbc:oracle:thin:@imiserver:1521:imldb/vocab#ACS
.EPISODEOF CARE/@@ACS.EPISODEOF CARE.ID@";
  d2rq:class
<jdbc:oracle:thin:@imiserver:1521:imldb/vocab#ACS
.EPISODEOF CARE>;
  d2rq:classDefinitionLabel "ACS.EPISODEOF CARE";
.
<jdbc:oracle:thin:@imiserver:1521:imldb/vocab#ACS
.EPISODEOF CARE_label> a d2rq:PropertyBridge;
  d2rq:belongsToClassMap
<jdbc:oracle:thin:@imiserver:1521:imldb/vocab#ACS
.EPISODEOF CARE>;
  d2rq:property rdfs:label;
  d2rq:pattern "ACS.EPISODEOF CARE
@@@ACS.EPISODEOF CARE.ID@";
.
map:EPISODEOF CARE_ID a d2rq:PropertyBridge;
  d2rq:belongsToClassMap
<jdbc:oracle:thin:@imiserver:1521:imldb/vocab#ACS
.EPISODEOF CARE>;
  d2rq:property vocab:EPISODEOF CARE_ID;
  d2rq:propertyDefinitionLabel "EPISODEOF CARE
ID";
  d2rq:column "ACS.EPISODEOF CARE.ID";
  d2rq:datatype xsd:decimal;
```

Fajl za mapiranje je osnovni element D2RQ platforme. Na osnovu ovog fajla aplikacija je u stanju da posmatra i pretražuje relacionu bazu podataka kao da je u pitanju jedan read-only (samo za čitanje) graf.

Nakon mapiranja relacione baze dobili smo jedan virtuelni graf koji smo u stanju da pretražujemo korišćenjem SPARQL upitnog jezika. Na slici 2 prikazan je graf koji se odnosi na epizodu posete.

Na osnovu specifikacije SPARQL upitnog jezika napravili smo upite koji nam kao rezultat vraćaju pacijente koji ispunjavaju zadate uslove.

VII. REZULTATI

A. Aplikacija za semantičko pretraživanje

Aplikacija za semantičku pretragu se sastoji iz:

- dela za unos parametara pretrage i
- dela za prikaz rezultata pretrage.

B. Parametri po kojima se vrši pretraživanje

Parametri su podeljeni u nekoliko grupa:

1. demografski podaci (datum rođenja, pol)
2. faktori rizika (bolesti srca i krvnih sudova kod oca, majke, braće, sestara, dece, pušenje u poslednjih mesec dana, hiperlipoproteinemije, hipertenzija, diabetes melitus)
3. dijagnoza akutnog koronarnog sindroma
4. terapije tokom hospitalizacije
5. ishod bolesti.



Slika 2. Deo grafa koji se odnosi na epizodu posete

U zavisnosti od toga koji su parametri izabrani, u delu za prikaz rezultata pretrage prikazuju se podaci koji zadovoljavaju zadate uslove (slika 3).

VIII. ZAKLJUČAK

U ovom radu je pokazano kako se SP nad registrom za akutni koronarni sindrom (AKS) može upotrebiti kao deo sistema za podršku kliničkom odlučivanju i koristi se kao pomoć u donošenju odluke u toku kliničkih poslovnih procesa vezanih za zbrinjavanje i lečenje osoba obolelih od AKS u okviru sveobuhvatnog sistema za upravljanje znanjem. Ovo je jedan od prvih pokušaja u našoj zemlji da se uprkos nespостоjanju elektronske zdravstvene dokumentacije izgradi sistem zasnovan na upravljanju znanjem. Potencijal ovog pristupa će biti istraživani u budućnosti posebno poredeći ga sa već postojećim sistemima koji su razvijani u zemljama razvijenog sveta [10], [15]-[17].

ZAHVALNICA

Posebnu zahvalnost autori duguju članovima Ekspertskog tima za Akutni koronarni sindrom Ministarstva zdravlja Republike Srbije.

Registar za akutni koronarni sindrom - REAKS

Semantičko pretraživanje

Datum rođenja od: 1.1.1960. Datum rođenja do: 1.1.1970. Pol: Muški. Dijagnoza: infarkt miokarda sa ST-elevacijom. Bolesti srca i krvnih sudova kod oca, majke, brace, sestara, dece: < izaberi >. Pušenje u poslednjih mesec dana: < izaberi >. Fibrinoliza: Ne. Urgentna PTCA: Da. Hipertenzija: Da. Diabetes melitus: Ne. Ishod: Oporavljen.

Epizoda posete	Pacijent	Pol	Datum rođenja	Otpust	Dijagnoza	Fibrinoliza	Urgentna PTCA	Bolesti srca i krvnih sudova kod oca, majke, brace, sestara, dece	Pušenje u poslednjih mesec dana	Hipertenzija	Diabetes melitus
30636	30636	Muški	24.03.1963.	Oporavljen	Infarkt miokarda sa ST-elevacijom	Ne	Da	Ne		Da	Ne
44176	44176	Muški	30.06.1960.	Oporavljen	Infarkt miokarda sa ST-elevacijom	Ne	Da	Ne		Da	Ne
25567	25567	Muški	09.02.1962.	Oporavljen	Infarkt miokarda sa ST-elevacijom	Ne	Da	Da		Da	Ne
27190	27190	Muški	12.07.1960.	Oporavljen	Infarkt miokarda sa ST-elevacijom	Ne	Da	Da		Da	Ne
28696	28696	Muški	02.12.1961.	Oporavljen	Infarkt miokarda sa ST-elevacijom	Ne	Da	Da	20	Da	Ne

Sl 3. Izgled aplikacije za semantičko pretraživanje. Prikazane su epizode posete sa dodatnim detaljima.

LITERATURA

- [1] http://en.wikipedia.org/wiki/Semantic_search/ (pristupano oktobra 2010. godine)
- [2] Van de Werf F, Ardissino D, Betriu A, et al. Management of acute myocardial infarction in patients presenting with ST-segment elevation. *Eur Heart J* 2003; 24:28-66
- [3] Bassand J, Hamm C, Ardissino D, et al. Guidelines for the diagnosis and treatment of non-ST-segment elevation acute coronary syndrom. Task Force for the Diagnosis and Treatment of non-ST-Segment Elevation Acute Coronary Syndromes of the European Society of Cardiology. *Eur Heart J* 2007; 14:1-63.
- [4] Vasiljević Z. Akutni koronarni sindrom: patofiziološki mehanizam, klasifikacija i klinički oblici. *Acta Clinica* 2006;6(1): 29-36.
- [5] Field MJ, Lohr KN, editors. *Guidelines for Clinical Practice: From Development to Use*. Washington, DC: National Academy Press; 1992
- [6] Peleg M, Tu SW, Bury J, Ciccarese P, Fox J, Greenes RA, et al. Comparing Computer-Interpretable, Guideline Models: A Case-Study Approach. *J Am, Med Inform Assoc* 2003;10(1):52-68.
- [7] Wang D, Peleg M, Tu S, Boxwala A, Greenes R, Patel V. Representation primitives, process models and patient data in computer-interpretable clinical practice guidelines: A literature review of guideline representation methods. *Int J Med Inform* 2002;68(1-3):59-70
- [8] de Clercq PA, Blomb JA, Korsten HHM, Hasman A.. Approaches for creating computer-interpretable guidelines that facilitate decision support. *Artif Intell Med* 2004;31:1-27
- [9] Sonnenberg FA, Hagerty CG. Computer-interpretable clinical practice guidelines: Where are we and where are we going? In: Kulikowski C, Haux R, editors. *IMIA Yearbook of Medical Informatics 2006*. *Methods Inf Med* 2006;45 Suppl 1: 145-58
- [10] Berlin A, Sorani M, Sim I. A taxonomic description of computer-based clinical decision support systems. *J Biomed Inform* 2006;39(6):656-67
- [11] <http://www.kardionet.org/> (pristupano oktobra 2010. godine)
- [12] Jena - A Semantic Web Framework for Java <http://jena.sourceforge.net/> (pristupano oktobra 2010. godine)
- [13] SPARQL Query Language for RDF - <http://www.w3.org/TR/rdf-sparql-query/> (pristupano oktobra 2010. godine)
- [14] The D2RQ Platform - Treating Non-RDF Databases as Virtual RDF Graphs <http://sourceforge.net/projects/d2rq-map/> (pristupano oktobra 2010. godine)

- [15] Chen CC, Chen K, Hsu CY, Li YC. Developing guideline-based decision support systems using protégé and jess. *Comput Methods Programs Biomed*. 2010 Jun 29. [Epub ahead of print]
- [16] Ash JS, Sittig DF, Dykstra R, Wright A, McMullen C, Richardson J, Middleton B. Identifying best practices for clinical decision support and knowledge management in the field. *Stud Health Technol Inform*. 2010;160(Pt 2):806-10.
- [17] Hsieh NC, Chang CY, Lee KC, Chen JC, Chan CH. Technological Innovations in the Development of Cardiovascular Clinical Information Systems. *J Med Syst*. 2010 Jul 23. [Epub ahead of print]

ABSTRACT

This paper presents the use of semantic search engine (SP) of the register for acute coronary syndrome (ACS) within the knowledge management systems. SP is part of clinical decision support system and is used as an aid in decision making in clinical processes related to the care and treatment of patients with ACS.

USING SEMANTIC SEARCH ENGINE OF THE REGISTER FOR ACUTE CORONARY SYNDROME

Pavle Kostić, Suzana Blesić, Zorana Vasiljević, Nataša Mickovski-Katalina, Jelica Grujuć Milanović, Siniša Pavlović, Slađan Milanović