

---

# KULTURNA BAŠTINA NA SEMANTIČKOM VEBU

---

Tamara Butigan-Vučaj  
Narodna biblioteka Srbije, Beograd

UDK 004.822:02(4)  
930.25:004.822(4)

---

*„Sanjam o vebu koji može da analizira sve podatke na mreži – sadržaje, njihove međusobne veze, razmenu podataka između ljudi i računara. Taj semantički veb još ne postoji, ali kada bude ostvaren svakodnevni postupci u trgovini, administraciji i organizaciji života obavljajuće mašine u komunikaciji sa drugim mašinama. Konačno će biti materijalizovani inteligentni agenti koje tako dugo najavljujemo.”*  
(Tim Berners-Lee: *Weaving the Web*, 1999)<sup>1</sup>

Svest o mogućnostima semantičkog veba, od kojih zastaje dah, daje našim životima još jednu porciju uzbuđenosti i iščekivanja. Isto toliko uzbuđenja semantički veb unosi i u sve potencijalne prostore primene. Tvorac semantičkog veba, Tim Berners-Li, možda nije bio svestan da će semantički veb

naći primenu u kulturnoj baštini, pre nego u našim svakodnevnim životima. On i njegovi saradnici vrlo rano su uočili da će se evolucija veba bazirati na interoperabilnosti sadržaja. Šta semantički veb može doneti sektoru kulturne baštine? U prvom redu mogućnost da njegovi sadržaji budu pristupačniji, i to na maksimalno svrsishodan način, ne samo stručnoj javnosti već i korisnicima. Kreatori kulturnih sadržaja mogli bi znatno da olakšaju svakodnevni posao pomoću inteligentnih agenata za organizaciju izložbi, tribina, pa zašto ne i automatskih agenata za nabavku ili katalogizaciju u bibliotekama. Za korisnike kulturnih sadržaja će, na primer, biti moguće zadati inteligentnom agentu da isplanira posetu Viminacijumu, prikupivši prethodno informacije o mogućim načinima za dolazak tamo, redovima vožnje i radnom vremenu kustosa na lokalitetu. Potrebno je samo da računari imaju dovoljno informacija koje su semantički povezane.

Sve ovo deo je semantičkog veba, koji se i dalje često određuje kao vizija postojećeg veba, uprkos činjenici da je ideja stara gotovo čitavu deceniju.

---

<sup>1</sup> Prevela Ljiljana Kovačević

Sektor kulturne baštine, tradicionalno tromiji u prihvatanju i primeni novih tehnologija, pokušava da prati ovu viziju, o čemu svedoči rastući broj međunarodnih konferencija i radionica o kulturnoj baštini na semantičkom vebu kao, i projekata, sa prvim konkretnim rezultatima. Prva međunarodna radionica o kulturnoj baštini na semantičkom vebu održana je u novembru 2007. godine, <http://www.cs.vu.nl/%7EElaroyo/CH-SW.html>, uz 6. Međunarodnu konferenciju o semantičkom vebu<sup>2</sup> u Koreji. U junu 2008. godine održana je prva radionica o semantičkoj interoperabilnosti u Evropskoj digitalnoj biblioteci,<sup>3</sup> <http://multimedia.semanticweb.org/siedl/>, u okviru 5. Evropske konferencije o semantičkom vebu.<sup>4</sup> Ustanovljena je i nagrada *Izazov semantičkog veba*,<sup>5</sup> koja se svake godine dodeljuje tokom međunarodne konferencije, a nekoliko puta je dodeljena za primene semantičkog veba u sektoru kulturne baštine.

Veliki broj evropskih projekata koji su se bavili kulturnom baštinom u periodu nakon 2000, bavili su se i semantičkim vebom. Projekat DigiCULT,<sup>6</sup> koji je realizovan između 2002. i 2004. godine, jedan je od prvih koji se bavio semantičkim vebom za resurse baštine. Kako je to lepo rekao Guntram Gezer,<sup>7</sup> saradnik ovog projekta: „Semantički veb podseća na lavirint za koji trenutno nemamo mapu niti Arijadninu nit na raspolaganju“ [1]. U cilju skidanja vela misterije sa semantičkog veba, u okviru ovog projekta dat je primer jednog objekta iz sektora baštine sa razrađenom semantičkom inter-

operabilnošću metapodataka, uključujući i XML, RDF i ontologije, sve ono što tehnološki čini semantički veb. Semantička interoperabilnost mogla bi da se definiše kao mogućnost da dva ili više računarskih sistema razmenjuju informacije i interpretiraju značenje tih informacija na način kako to očekuju krajnji korisnici ovih sistema. XML je skraćenica za *Extensible Markup Language*, metajezik za označavanje dokumenata i jedan je od W3C standarda. RDF je skraćenica za *Resource Description Framework*, grupu W3C specifikacija, zamišljenih kao model metapodataka za opis veb izvora. Format opisa je triplet: subjekat – predikat – objekat. Za biblioteke su posebno zanimljive ontologije kako ih je definisao profesor Džejms Hendler,<sup>8</sup> kao «skupove termina, uključujući i rečnike, semantičke veze, i neka osnovna pravila izvođenja i logike». U vezi sa ontologijama postavlja se čitav niz pitanja, počev od ekonomskih, kao što su cene formiranja ontologija preko pitanja interoperabilnosti, korišćenja ontologija iz drugih oblasti u sektoru baštine, do toga koliko OWL<sup>9</sup> može da zadovolji potrebe ontologija u sektoru baštine. Prema rečima Nikole Guarina,<sup>10</sup> ontologije, budući kreirane za određene primene, uglavnom nisu interoperabilne. Isti termini imaju različita značenja u različitim primenama. On ističe da su zato *dobro zasnovane*<sup>11</sup> generičke ontologije apsolutni imperativ i preduslov semantičkog veba. Dobar primer klasifikacije u kulturnoj baštini predstavlja holandski *Iconclass* <http://www.iconclass.nl/>. Nastao je sistematizacijom predmeta, tema i motiva zastupljenih u umetničkim delima, a sadrži 28.000 pojmova svrstanih u 10 grupa,

<sup>2</sup> ISWC – International Semantic Web Conference, <http://iswc.semanticweb.org/>

<sup>3</sup> SIEDL – Semantic Interoperability for European Digital Library

<sup>4</sup> ESWC 2008 – 5<sup>th</sup> European Semantic Web Conference, <http://www.eswc2008.org/>

<sup>5</sup> Semantic web Challenge, <http://challenge.semanticweb.org/>

<sup>6</sup> <http://www.digicult.info>

<sup>7</sup> Guntram Gezer, Salzburg Research, Austria, <http://www.salzburgresearch.at/>

<sup>8</sup> James Handler, Rensselaer Polytechnic Institute (RPI), USA, <http://www.rpi.edu/>

<sup>9</sup> OWL – Web Ontology Language, grupa jezika za kreiranje ontologija, razvijena u W3C

<sup>10</sup> Nicola Guarino, Laboratory for Applied Ontology, <http://www.loa-cnr.it/>

<sup>11</sup> Well-founded, termin koji se koristi u matematici za relacije

14.000 ključnih reči i 40.000 bibliografskih referenci. Onlajn verzija ovog sistema podrazumeva i prikaz objekata (slika) koje su opisane pojmovima. Kraljevska biblioteka Holandije učestvuje u ovom projektu zbirkom srednjovekovnih iluminiranih rukopisa.

Holandija je jedna od prvih zemalja koja je počela da se bavi semantičkim vebom u sektoru baštine, tačnije Holandska asocijacija za digitalnu baštinu,<sup>12</sup> kliringhaus za informaciono-komunikacione tehnologije i kulturnu baštinu, koja je osnovana 1999. godine, a okuplja vodeće institucije baštine. Ova asocijacija je kreirala kulturni portal Cultuurwijzer,<sup>13</sup> koji omogućava pristup bazama podataka svih organizacija članica, između ostalog i preko polja predmeta u Dablinskom jezgru. Još jedan projekat ove asocijacije, mali korak bliže semantičkom vebu, bavio se razvojem jedinstvene sheme za opis građe na nivou zbirke, kako bi se omogućio uniformniji pristup baštinskim zbirkama. Koren ovog projekta jeste u holandskom projektu MUSIP,<sup>14</sup> koji se bavio opisom zbirke muzejske građe.

Sektor kulturne baštine vremenom je postao značajniji prostor za primenu semantičkog veba. Tokom 2006. godine održano je više konferencija koje su se bavile dodirnim tačkama semantičkog veba i kulturne baštine. U Parizu je održan seminar Digitalni semantički sadržaj kroz kulturu,<sup>15</sup> na kome su razmenjena francuska i finska iskustva u ovoj oblasti.

Finska je 2003. godine pokrenula nacionalni projekat za razvoj ontologija FinnONTO, baziran na nacionalnom tezaurusu od 23.000 pojmova, uz

predviđeni niz primena u različitim oblastima, od kojih su najuspešnije bile one u oblasti eKulture: Finska kultura na semantičkom vebu, <http://www.seco.tkk.fi/applications/kulttuurisampo/> i druga vezana za muzeje, semantički portal: Finski muzeji na semantičkom vebu, <http://www.museosuomi.fi>. Portal je nagrađen nagradom *Izazov semantičkog veba* 2004. godine. Rezultat istog projekta su i portal Orava koji sadrži video sadržaje i materijale za učenje, <http://demo.seco.tkk.fi/orava/> i semantički Suomi.fi portal, <http://demo.seco.tkk.fi/suomifi/>, razvijen za eUpravu.

Projekat SCULPTEUR, <http://www.sculpteurweb.org/> je od 2002. do 2005. godine udružio nekoliko muzejskih institucija iz Francuske i Velike Britanije i nekoliko akademskih institucija za tehničku podršku, razvivši *konceptualni brauser* heterogenih muzejskih baza podataka, polazeći od zajedničke ontologije: CIDOC CRM.<sup>16</sup> Pokušaji da se CIDOC CRM primeni i na bibliografske baze podataka vodili su preko FRBR-a,<sup>17</sup> koji nije ontologija ali je konceptualni model bibliografskih podataka. Probni model FRBR ontologije, kao proširenja CIDOC CRM, razvijen je 2006. godine u saradnji sa IFLA-om. Između 2004. i 2007. godine, na bazi CIDOC CRM ontologije, kreiran je portal zavičajnih zbirke jedne oblasti u Finskoj, <http://www.museo24.fi/>.

U Nemačkoj nacionalnoj biblioteci pokrenuto je nekoliko projekata koji istražuju kako bi se maksimalno mogla iskoristiti semantika bibliografskih metapodataka pomoću alata semantičkog veba. Prvi je saradnja sa nemačkom Vikepedijom od 2005. godine, naime svi biografski članci u Vikipediji dobili su strukturalne metapodatke kojima se ostvaruje veza

<sup>12</sup> Vereniging DEN, <http://www.den.nl>

<sup>13</sup> Culture pointer – Kulturni pokazatelj, <http://www.cultuurwijzer.nl/>

<sup>14</sup> MUSIP – Museum Inventarisation Project, <http://www.musip.nl/>

<sup>15</sup> Symposium on Digital Semantic Content across Cultures, <http://www.seco.tkk.fi/events/2006/2006-05-04-websemanitique/>

<sup>16</sup> CIDOC CRM - The International Committee for Museum Documentation Conceptual Reference Model, <http://cidoc.ics.forth.gr/>

<sup>17</sup> FRBR, Functional Requirements for Bibliographic Records – Funkcionalni zahtevi za bibliografske zapise, razvijeni u IFLA-i 1998.

sa onlajn katalogom Nemačke nacionalne biblioteke, bilo da je lično ime autorska ili predmetna odrednica. Ovaj projekat je ukazao na nedostatak jedinstvenih identifikatora izvora (URI) u normativnoj datoteci ličnih imena i odgovarajućeg mehanizma za njihovo razrešavanje, kao preduslov za realizaciju semantičkog vebe. Drugi projekat *CrissCross* bavi se kreiranjem višejezičnog, korisnicima prilagođenog onlajn rečnika, a baziran je na sistemu predmetnih odrednica Nemačke nacionalne biblioteke i njihovoj vezi sa DDC odrednicama, kao i drugim sistemima predmetnih odrednica, kao što je onaj Kongresne biblioteke.

Jedna od najznačajnijih evropskih inicijativa u domenu kulturne baštine jeste izgradnja Evropske digitalne biblioteke (EDL), koja za cilj ima da omogući jednostavan pristup evropskoj kulturnoj i naučnoj baštini, uključujući i sve vrste multimedijalne građe iz svih institucija čuvara baštine: arhiva, biblioteka i muzeja. Ovaj jedinstveni portal pratiće multikulturalno i višejezičko okruženje, a već 2010. godine na njemu će biti dostupno preko 6 miliona digitalnih dokumenata iz čitave Evrope. Evropska digitalna biblioteka direktno proizilazi iz agende i2010 Evropske komisije za digitalne biblioteke, a projekat *EDLnet tematska mreža*,<sup>18</sup> koji se realizuje u okviru programa *eContentPlus* Evropske komisije, ima za cilj da pripremi teren za izgradnju buduće Evropske digitalne biblioteke – *Europeane*. U ovom projektu jedan od ključnih ciljeva jeste postizanje interoperabilnosti, u prvom redu semantičke interoperabilnosti, kao automatskog tumačenja značenja razmenjenih informacija od strane mašina, kod pristupa kulturnim sadržajima. Jedan radni paket u projektu *EDLnet* upravo se bavi interoperabilnošću, jer je ona samo srce tehničke vizije *Europeane* [4]. Zbirke kulturne i naučne baštine indeksirane su u skladu sa različitim standardima metapodataka: MARC za bibliografski

opis, EAD za arhivsku građu, AMICO za muzejsku građu, Dablinsko jezgro za digitalne zbirke. Prvi korak ka međurazumevanju ovih opisa jeste formalni reprezentativni jezik za razmenu formalne semantike na nivou informacija, a sledeći je izgradnja sistema koji će razumeti ovu semantiku, kao što su alati za (automatsko) rezonovanje i semantičke mašine za pretraživanje. Njihova je uloga da omoguće veb servise za pretraživanje i pronalaženje informacija. U interesu je Evropske digitalne biblioteke da regrutuje sve postojeće inicijative i istraživanja tehnologija semantičkog vebe, kako bi semantički povezala evropsku kulturnu baštinu u jedinstveni portal.

Projekat *eCulture*, <http://e-culture.multimedial.nl/> pokrenulo je nekoliko holandskih istraživačkih i akademskih institucija, sa ciljem da izgrade jedinstvenu semantičku mašinu za pretraživanje sadržaja zbirki više holandskih repozitorijuma. Radi se uglavnom o muzejskoj građi, a projektom je semantičkom vebu prilagođeno desetak različitih skupova metapodataka. U projektu je razvijena metodologija za postavljanje sadržaja umetničkih zbirki na semantički veb kroz RDF platformu. Krenulo se od podataka kojima raspolažu baštinske institucije: metapodaci koji opisuju objekte, fotografije (skenovi) objekata i lokalni kontrolisani rečnici koji se koriste za neke od metapodataka. Različite institucije koristile su različite skupove metapodataka, različite rečnike i različito organizovane opise, od baza podataka, preko XML dokumenata do običnih tekstualnih podataka. Prvi korak bila je transformacija sheme metapodataka i postojećih rečnika u jednu standardizovanu RDF shemu, u kojoj su korišćeni standardi: Dablinsko jezgro za metapodatke i SKOS<sup>19</sup> za rečnike. A drugi je korak bio pronalaženje odgovarajućih termina iz lokalnih tezaurusa u nekim od standardnih rečnika;

<sup>18</sup> EDLnet Thematic Network,

<sup>19</sup> SKOS, Simple Knowledge Organization System, <http://www.w3.org/2004/02/skos/>

u ovom projektu korišćeni su Getty rečnici<sup>20</sup>: Tezaurus za umetnost i arhitekturu sa preko 130.000 termina, Jedinstvena lista imena umetnika sa 300.000 imena i Tezaurus geografskih naziva koji opisuje preko milion geografskih lokacija. Korišćeno je i nekoliko holandskih rečnika i tezaurusa. Na osnovu definisane metodologije razvijena je aplikacija za automatsku konverziju izvornih podataka u standardizovanu shemu i povezivanje termina sa onim u standardnim rečnicima. Ovo programsko rešenje, pod imenom *Annocultor*, je onlajn dostupno. Virtuelnim korisnicima portala projekta na raspolaganju je demo verzija pretraživanja, kojese ubrzo pretvara u jedno uzbudljivo krstarenje kroz vreme, prostor i umetnost. Ova aplikacija nagrađena je na Međunarodnoj konferenciji o semantičkom vebu 2006. godine nagradom *Izazov semantičkog veba*. Demo verzija Europeane<sup>21</sup> upravo podseća na ovo nagrađeno rešenje.

U domenu arhiva zanimljiv je projekat *Video Active*, <http://videoactive.wordpress.com/>, pokrenut u okviru programa *eContent Plus* Evropske komisije 2006. godine, koji za cilj ima da preko jedinstvenog portala omogući pristup bogatoj televizijskoj baštini iz 11 evropskih arhiva audiovizuelne građe na 10 različitih jezika. Pored arhiva, partneri u konzorcijumu projekta su i akademske institucije. Svaki arhiv napravio je izbor od pet sati audio i video materijala od 1890. godine do danas. Za postizanje semantičke interoperabilnosti korišćeni su postojeći metapodaci iz arhiva, prilagođeni skupu podataka iz Dablinskog jezgra, uz nekoliko dodatnih elemenata koji su specifični za ovu vrstu građe. Ovi opisi su zatim prevedeni u sheme RDF-a. Kao semantičko skladište metapodataka u ovom projektu se koristi Java platforma *Sesame*. Korisnicima je omogućeno

pretraživanje i pregledanje sadržaja. Projekat *Video Active* je partnerski projekat *EDLnet-a* i rezultati koji budu postignuti do isteka projekta biće korišćeni u Evropskoj digitalnoj biblioteci za audiovizuelnu građu.

Nemački muzeji pokrenuli su inicijativu koja ih vodi bliže semantičkom vebu kroz izradu jedinstvenog formata metapodataka za muzejsku građu pod nazivom *museumdat*, <http://www.museumdat.org/>, kao i za izradu jedinstvenog tezaurusa, pod nazivom *museumvoc*, <http://www.museumsvokabular.de>.

U institucijama baštine u Srbiji, bibliotekama, muzejima, arhivima i zavodima za zaštitu spomenika, najkraći put do ontologija, imoguće semantičkog veba, treba da pređu biblioteke, zahvaljujući jednom broju tezaurusa i masovnijem korišćenju baza podataka za opis dokumenata. Projekat izrade makrotezaurusa u Narodnoj biblioteci Srbije nažalost nije završen. Uspeh semantičkog veba u sektoru baštine zavisi i od spremnosti navedenih institucija da usvoje XML pristup u opisivanju zbirki i objekata. Kratak osvrt na prilike u Srbiji, vezan za korišćenje XML-a u sektoru kulturne baštine, krajnje je razočaravajući. Samo nekoliko institucija koristi ovaj jezik, a interesovanje za radionice XML-a je poražavajuće malo. Kao i za proces digitalizacije tako i za primenu semantičkog veba, institucijama kulture u Srbiji potrebna je jedna krovna asocijacija, organizacija ili institucija koja bi se, između ostalog, bavila postizanjem semantičke interoperabilnosti i razvojem rečnika i ontologija za nacionalnu kulturnu baštinu, kako bi je sa postojećeg veba polako migrirala na viši, semantički veb.

#### Literatura:

1. Towards a Semantic Web for Heritage Resources, May 2003, [http://www.digicult.info/downloads/thematic\\_issue\\_3\\_low.pdf](http://www.digicult.info/downloads/thematic_issue_3_low.pdf) [31.8.2008]

<sup>20</sup> [http://www.getty.edu/research/conducting\\_research/vocabularies/](http://www.getty.edu/research/conducting_research/vocabularies/)

<sup>21</sup> <http://www.europeana.eu>, ali i <http://www.europeana.rs>, <http://www.europeana.fr>, ...

Tamara Butigan-Vučaj

2. SIEDL Workshop Proceedings, <http://image.ntua.gr/swamm2006/SIEDLproceedings.pdf> [31.8.2008]
3. First CHSW workshop proceedings, <http://www.cs.vu.nl/%7Elaroyo/CH-SW/ISWC-wp9-proceedings.pdf> [31.8.2008]
4. One more step towards the European digital library, International Conference, <http://www.edlproject.eu/conference/> [31.8.2008]

Tamara Butigan-Vučaj

*Kulturna baština na semantičkom vebu*

Rezime

Digitalizacija, anotiranje, arhiviranje, pronalaženje i prezentacija kulturne baštine jesu preduslovi, ali tek postizanje semantičke interoperabilnosti pravi je ključ za ostvarivanje vizije semantičkog veba u ovom sektoru. Institucijama koje se bave kulturnom baštinom ova vrsta interoperabilnosti nije jedan od primarnih ciljeva, i veliki izazov za njih predstavlja usvajanje tehnologija, standarda i načina

razmišljanja neophodnih za primenu semantičkog veba. Semantički veb nudi rešenje problema informacionog preopterećenja koji proizvodi veb, a od sektora kulturne baštine očekuje se da prepozna priliku i iskoristi je.

Rad sadrži pregled napora, aktivnosti i projekata primene semantičkog veba na široki prostor evropske kulturne baštine.

Tamara Butigan-Vučaj

*Cultural Heritage on the Semantic Web*

Résumé

Digitization, annotating, archiving, discovering and presentation of cultural heritage are the preconditions, but the true key to the realization of the vision of the semantic web in this sector is the achieving of semantic interoperability. The institutions dealing with cultural heritage do not consider this kind of interoperability as one of the primary aims, and it is a great challenge for them to adapt technologies, standards and ways of thinking necessary for the application of the semantic web. The semantic web offers a solution to the problem regarding the information overburdening produced by the web, and it is upon the cultural heritage sector to recognize the opportunity and take advantage of it.

The work contains a survey of efforts, activities and projects concerning the applying of the semantic web to the wide area of the European cultural heritage.