

UNIMARC za autorizirane podatke u RDF-u



UNIMARC for authority data in RDF

Leonardo Jelenković

UNIBIS, Zagreb

leonardo.jelenkovic@zg.t-com.hr

Mirna Willer

Odjel za informacijske znanosti, Sveučilište u Zadru
Department of Information Sciences, University of Zadar

mwiller@unizd.hr

UDK / UDC 025.31:004

Istraživanje / Research

Primljeno / Received on: 26. 5. 2014.

Prihvaćeno / Accepted on: 26. 6. 2014.

Sažetak

Rad je nastavak istraživanja koje se bavilo problematikom objavljivanja bibliografskih podataka UNIMARC kao povezanih otvorenih podataka u RDF-u, standardu semantičkog weba. Kako u vrijeme kad je rađeno to istraživanje (2012.-2013.) format UNIMARC nije imao objavljene svoje imenske prostore koji bi se mogli koristiti pri označivanju podataka, izrađena je pretvorba formata UNIMARC u druge rječnike koji su imali objavljene imenske prostore, kao što su ISBD, Dublin Core, bibo itd. U ovom radu autori se bave autoriziranim podacima. Metodologija pristupa problemu je ista, s tom razlikom što se sada koriste i imenski prostori formata UNIMARC za autorizirane podatke koji su konstruirani u skladu s metodom opisanom za imenske prostore formata UNIMARC objavljenom tijekom 2013. Prikazat će se problemi i postignuta rješenja objavljivanja autoriziranih podataka kao povezanih otvorenih podataka.

Ključne riječi: format za autorizirane podatke UNIMARC, RDF: Resource Description Framework, povezani otvoreni podaci (LOD), semantički web, metode objavljivanja LOD-a, CROLISTKIS

Summary

The paper describes the second phase of the research dealing with publishing UNIMARC bibliographic data as linked open data in RDF, the standard of the Semantic Web. As at the time of the first research (2012-2013) UNIMARC formats did not have published their namespaces to be used in publishing bibliographic data, the mapping was made between UNIMARC and some of those schemas that had their namespaces published, such as ISBD, Dublin Core, bibo, etc. In this paper the issues of publishing authority data are dealt with. The approach to the problem is the same as in the first phase of the research, while the difference being that UNIMARC namespaces are being additionally used being constructed following the method published by the end of 2013. Problems and solutions of publishing authority data as linked open data are being described.

Key words: UNIMARC/Authorities format, RDF: Resource Description Framework, LOD: Linked Open Data, Semantic Web, LOD publishing methods, CROLIST LIS

Uvod

Rad je nastavak istraživanja koje su autori objavili u zborniku radova 16. seminara *Arhivi, knjižnice, muzeji*, a koje se bavilo problematikom objavljivanja bibliografskih podataka kao povezanih otvorenih podataka u RDF-u, standardu semantičkog weba.¹ Istraživanje prikazano u tom radu, kao i u ovome, rađeno je na UNIMARC-zapisima iz baze podataka skupnog kataloga CROLIST, ali je naglasak stavljen na format UNIMARC, ne na njegovu implementaciju. Kako u vrijeme rada na istraživanju format UNIMARC nije imao objavljene svoje imenske prostore koji bi se mogli koristiti pri označavanju podataka, prikazan je način pretvorbe (*mapiranja*) formata UNIMARC u druge rječnike ili standarde koji su već imali objavljene imenske prostore kao što su ISBD, Dublin Core, bibo itd. Rad je posebno naglasio potrebu objavljivanja podataka usporedo u više rječnika. Razlog tomu bilo je razmišljanje da korisnici povezanih otvorenih podataka mogu preuzeti podatke u onom rječniku koji im je prikladniji, čime se povećava mogućnost ponovnog korištenja tih podataka i time povezivanje s drugim podacima, što je i temeljna svrha koncepta povezanih podataka, dok se pretvorba izvornog formata UNIMARC u druge rječnike nalazi pod nadzorom autora knjižnično-informacijskog sustava, tj. CROLIST-a.

U ovom se radu autori bave autoriziranim podacima te slijede isti pristup rješavanju problema. Metodološki, pristup problemu razlikuje se jedino u tomu da se sada koriste uz ostale i imenski prostori formata UNIMARC za autorizirane podatke koji su konstruirani u skladu s metodom koja opisuje postupak izrade imenskih prostora za te formate, a koja je u međuvremenu objavljena u okviru projekta IFLA-ina Strateškog programa

1 Jelenković, Leonardo; Mirna Willer. Kako do povezanih podataka iz UNIMARC zapisa? // 16. seminar *Arhivi, knjižnice, muzeji: mogućnosti suradnje u okruženju globalne informacijske infrastrukture: zbornik radova / uredile Nives Tomašević i Ivona Despot*. Zagreb: Hrvatsko knjižničarsko društvo, 2013. Str. 30-56.

UNIMARC *UNIMARC in RDF*.² Time se, a u skladu sa zaključkom prethodnog istraživanja, ovaj rad temelji na iskazanoj potrebi da se podaci objavljuju u rječniku u kojem su izvorno nastali kako bi se „zadržao potpuni kontekst zapisa, tj. sačuvala izvorna informacijska vrijednost zapisa“,³ dok se istodobno usporedo objavljuju podaci i u nekim drugim rječnicima. Ova metoda objavljivanja povezanih podataka u izvornim formatima, odnosno rječnicima, tj. metoda „odozdo prema gore“, suprotstavlja se metodi „odozgo prema dolje“, tj. metodi koja polazi od općeg prema posebnom, odnosno objavljivanju podataka u imenskim prostorima rječnika (npr. Dublin Core) koji udovoljavaju globalnom, poopćenom pristupu nasuprot lokalnom ili specifičnom.⁴

Potrebno je uputiti čitatelja i na još jedan zaključak prethodnog rada, a koji se odnosi na početnu pretpostavku „da povezani podaci objavljeni na webu – *webu podataka* – ne zamjenjuju postojeću infrastrukturu informacijskih sustava, već su komplementarni toj strukturi, tj. da je to još jedan od načina objavljivanja bibliografskih podataka“.⁵ Istraživanje je međutim ukazalo na suprotno:

„Premda neki autori smatraju da je potonja mogućnost [raspoloživi sustavi] realna, tj. da je dovoljna nadgradnja u smislu izrade softvera za konverziju podataka u LOD [*linked open data*], smatramo da takvo rješenje može biti samo privremeno. Naime baveći se problematikom

2 Dunsire, Gordon; Mirna Willer. UNIMARC and Linked Data. // IFLA Journal 37, 4 (December 2011), 314-326 [citirano: 2014-05-23]. Dostupno na: http://www.ifla.org/files/hq/publications/ifla-journal/ifla-journal-37-4_2011.pdf; Willer, Mirna; Gordon Dunsire; Predrag Perožić. The UNIMARC in RDF project: name spaces and linked data. // IFLA World Library and Information Congress, 79th IFLA General Conference and Assembly, 17-23 August 2013, Singapore. Session 222 — Expanding MARC metadata services with linked open data — UNIMARC Core Activity (UCA) [citirano: 2014-05-23]. Dostupno na: <http://library.ifla.org/156/1/222-willer-en.pdf>; Dunsire, Gordon; Mirna Willer; Predrag Perožić. Representation of the UNIMARC bibliographic data format in Resource Description Framework. // DC-2013, Lisbon, Portugal, [International Conference on Dublin Core and Metadata Applications]: Linking to the Future, Lisbon, 2 - 6 September, 2013 [citirano: 2014-05-23]. Dostupno na: <http://dcevents.dublincore.org/IntConf/dc-2013/paper/view/173/91>; IFLA UNIMARC Strategic Programme [citirano: 2014-05-23]. Dostupno na: <http://www.ifla.org/unimarc>

3 Jelenković, L.; M. Willer. Nav. dj., str. 31.

4 Hillmann, Dianne; Gordon Dunsire; Jon Phipps. Maps and gaps: strategies for vocabulary design and development. // Proc. Int'l Conf. on Dublin Core and Metadata Applications 2013, 82-89 [citirano: 2014-05-23]. Dostupno na: <http://dcevents.dublincore.org/IntConf/dc-2013/paper/view/185/80>

5 Jelenković, L.; M. Willer. Nav. dj., str. 31. Citirani autori su: Heath, Tom; Christian Bizer. *Linked Data: Evolving the Web into a Global Space*. [San Rafael, Calif.]: Morgan & Claypool, 2011.

objavljivanja bibliografskih podataka kao LOD, tijekom istraživanja došli smo do zaključka da je potrebno razmotriti, a time i planirati razvoj nove generacije softvera koja će omogućiti da se tijekom obrade građe osim strojne stvaraju ili uspostavljaju veze prema onim LOD-ima koji su intelektualno obrađeni, tj. koji su prepoznati kao vjerodostojni i autentični, a time i prihvatljivi za povezivanje s našim, lokalnim podacima.“⁶

Taj je zaključak posebno značajan za autorizirane podatke, na čije daljnje istraživanje i upućuje navedeni rad.

Istraživanje

Kao i u istraživanju koje se bavilo bibliografskim podacima opisanom u prethodnom radu, prvo se pristupilo analizi korištenja imenskim prostorima u većim nacionalnim knjižnicama i skupnim katalozima koji objavljuju svoje autorizirane podatke. Nakon prikaza te analize iznose se rješenja do kojih se došlo u objavljivanju autoriziranih podataka kao povezanih otvorenih podataka (LOD) iz sustava CROLIST.

Istraživanje (1): Upotreba ontoloških rječnika za autorizirane podatke u nacionalnim knjižnicama i skupnim katalozima

U Tablici 1 prikazan je popis rječnika koje koriste neke nacionalne knjižnice i skupni katalozi za objavljivanje autoriziranih podataka kao LOD. Tabela ne sadrži prikaz sintaktičkih standarda kao što su RDFS (Resource Description Framework Schema), SKOS (Simple Knowledge Organization System) i OWL (Web Ontology Language). Analizirani rječnici navedeni u prvom redu tablice su bio – Bio: A Vocabulary of Biographical Information,⁷

⁶ Isto, str. 33.

⁷ Bio: A vocabulary of biographical information [citirano: 2014-05-23]. Dostupno na: <http://vocab.org/bio/0.1/html>

foaf – FOAF: Friend of a Friend,⁸ dct: DC terms,⁹ rda: preliminarna verzija imenskih prostora RDA¹⁰ i org: Anorganization ontology.¹¹

Tablica 1. Usporedba korištenja imenskih prostora autoriziranih podataka i rječnika u nekim nacionalnim knjižnicama i skupnim katalozima

Ustanova/ Schema	bio	foaf	dct	rda	org	drugo
British Library, UK	x	x	X	x	x	
BnF, F			X	x		Inter/UNIMARC à RDF
DNB, DE						GND Ontology
LC, SAD		x				MADS
VIAF		x				
CERL Thesaurus		x		x		wgs84_pos, gnd, gn, ct, rel, edm

Britanska knjižnica koristi sve navedene rječnike u prikazu autoriziranih podataka,¹² dok Nacionalna knjižnica Francuske, osim Dublin Corea, koristi neke elemente kataložnih pravila *RDA: Resource Description & Access* te format INTERMARC, odnosno UNIMARC za pojedine elemente. Njemačka nacionalna knjižnica razvila je vlastiti imenski prostor: GND Ontology,¹³ a Kongresna knjižnica koristi MADS, verziju formata MARC21 za autorizirane podatke¹⁴ uz rječnik FOAF. Skupni katalog autoriziranih podataka VIAF: Virtual International Authority File¹⁵ koristi samo rječnik foaf, dok skupni katalog stare građe CERL Thesaurus (HPB: European Printed Heritage in

8 FOAF: Friend of a Friend [citirano: 2014-05-23]. Dostupno na: <http://www.foaf-project.org/>

9 DCMI Metadata Terms [citirano: 2014-05-23]. Dostupno na: <http://dublincore.org/documents/dcmi-terms/>

10 Preliminary version of RDA [citirano: 2014-05-23]. Dostupno na: <http://metadataregistry.org/schema/show/id/15.html>

11 Anorganization ontology [citirano: 2014-05-23]. Dostupno na: <http://www.epimorphics.com/public/vocabulary/org.html>

12 British Library Data Model – Book [citirano: 2014-05-23]. Dostupno na: <http://www.bl.uk/bibliographic/pdfs/bl-datamodelbook.pdf>; British Library Metadata Services, Free Data Services [citirano: 2014-05-23]. Dostupno na: <http://www.bl.uk/bibliographic/datafree.html>

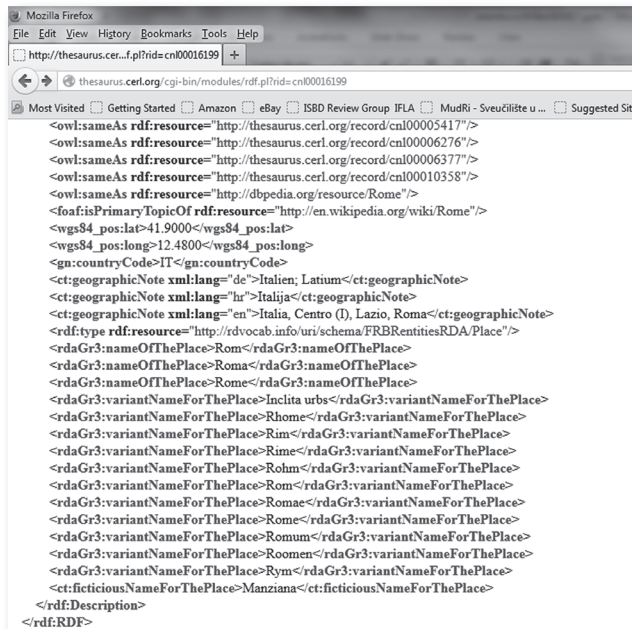
13 GND Ontology [citirano: 2014-05-23]. Dostupno na: <http://d-nb.info/standards/elementset/gnd>

14 MADS/RDF Primjer [citirano: 2014-05-23]. Dostupno na: <http://www.loc.gov/standards/mads/rdf/#t11>

15 VIAF: Virtual International Authority File [citirano: 2014-05-23]. Dostupno na: <http://viaf.org/>

the Hand-press Period, up to c. 1830),¹⁶ osim rječnika foaf i imenskih prostora RDA, koristi više drugih, kao što su gnd, edm (Europeana Data Model),¹⁷ rječnik za geopozicioniranje wgs84-pos¹⁸ i geonames – gn.¹⁹

Može se zaključiti da samo tri organizacije među promatranima koriste dosljedno jedan rječnik, dok ostale koriste metodu *mix & match*, tj. kombinaciju različitih rječnika za objavljivanje autoriziranih podataka. Ta metoda kombiniranja različitih rječnika uočena je i kod objavljivanja bibliografskih podataka u navedenim ustanovama. Na Slici 1. prikazan je autorizirani RDF-zapis za naziv grada Rima iz CERL-ova tezaurusa u kojem se koriste navedeni rječnici, a na Slici 2. taj isti zemljopisni naziv iz VIAF-a koji koristi samo rječnik foaf.



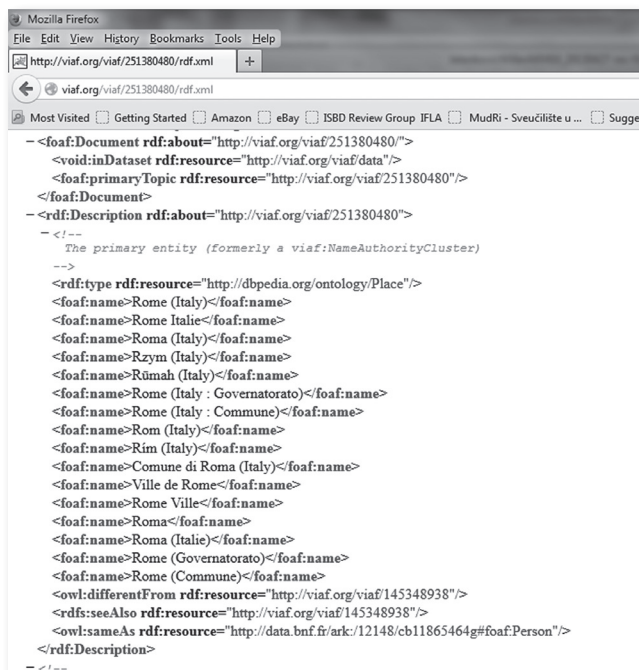
Slika 1. RDF-zapis iz CERL-ova tezaurusa skupnog kataloga HPB za mjesto tiskanja Rim

16 CERL Thesaurus [citirano: 2014-05-23]. Dostupno na: http://www.cerl.org/resources/cerl_thesaurus/main

17 Europeana Data Model (EDM) documentation [citirano: 2014-05-23]. Dostupno na: <http://pro.europeana.eu/edm-documentation>

18 W3C Semantic Web Interest Group, Basic Geo (WGS84 lat/lon) Vocabulary [citirano: 2014-05-23]. Dostupno na: <http://www.w3.org/2003/01/geo/>

19 Geo Names Ontology [citirano: 2014-05-23]. Dostupno na: <http://www.geonames.org/ontology/documentation.html>



Slika 2. RDF-zapis iz međunarodnog skupnog kataloga autoriziranih podataka VIAF za mjesto tiskanja Rim

Promatrajući navedeno, na pitanje koji ontološki rječnik izabрати odgovor se trebao pronaći između sljedećega: (1) izabрати rječnik koji se koristi u najvišeslučajeva, odnosno koji se koristi za najveći broj zapisa, npr. foaf kataloga VIAF ili rječnik lokalnog kataloga, npr. UNIMARC kataloga CROLIST ili (2) izabрати samo jedan rječnik ili više njih za objavljivanje podataka autoriziranih zapisa.

U prvom slučaju radi se o tzv. situaciji *Poor and popular vs. rich and lonely*,²⁰ tj. odabрати rječnik za objavljivanje podataka koje svatko može čitati jer su prikazani u jednostavnom, „siromašnom“ i, time, općem „jeziku“ koji se koristi u velikom broju slučajeva, pa će i pronalaženje, ponovno korištenje, a time i povezivanje s „lokalnim“ podacima biti uspješnije. U tom slučaju,

20 Vidjeti informaciju o radionici istog naslova održanoj na konferenciji DCMI International Conference on Dublin Core and Metadata Applications, Lisbon, September, 2013 [citirano: 2014-05-23]. Dostupno na: <http://dcevents.dublincore.org/IntConf/index/pages/view/rich-poor>

povezani podaci koje objavljuju knjižnice bit će dostupni široj skupini potencijalnih korisnika i izvan te zajednice, omogućit će interoperabilnost podataka s onima tih zajednicama, što je i cilj objavljivanja povezanih podataka kao otvorenih povezanih podataka. Nasuprot tomu, posljedica odabira korištenja „vlastitog“, „lokalnog“ formata, koji je u upotrebi u određenoj knjižničnoj zajednici,²¹ ali koji u potpunosti prikazuje „bogatu“ informacijsku vrijednost autoriziranih podataka, bit će manja dostupnost i, time, smanjena mogućnost ponovnog korištenja tih podataka, što pretpostavlja i smanjene mogućnosti uključivanja u mrežu povezanih otvorenih podataka.

U drugom slučaju, stavlja se u opreku odabir jednog rječnika nasuprot odabiru njih više. U prvom slučaju, jedan rječnik koriste ustanove koje su razvile vlastite rječnike, npr. MADS i gnd, nad kojima imaju potpuni nadzor u odnosu na upotrebu i razvoj. U drugom slučaju, odlučeno je da se koristi tzv. tehnika *mix & match* u nedostatku jednog jedinstvenog rječnika, npr. Britanska knjižnica i CERL Tezaurus, kako bi se objavili izvorni podaci u svom kontekstu jer nijedan raspoloživi rječnik ne udovoljava potrebi prikaza potpune informacijske vrijednosti lokalnog zapisa.

Možemo dakle zaključiti da je u promatranim slučajevima, bez obzira na odluku o strategiji, odnosno metodi objavljivanja podataka, ipak naglasak na objavljivanju „bogatih“ otvorenih povezanih podataka koji zadržavaju izvornu informativnu vrijednost zapisa izrađenih u katalozima prethodne tehnologije.

Istraživanje (2): Objavljivanje UNIMARC-zapisa autoriziranih podataka kao povezanih otvorenih podataka

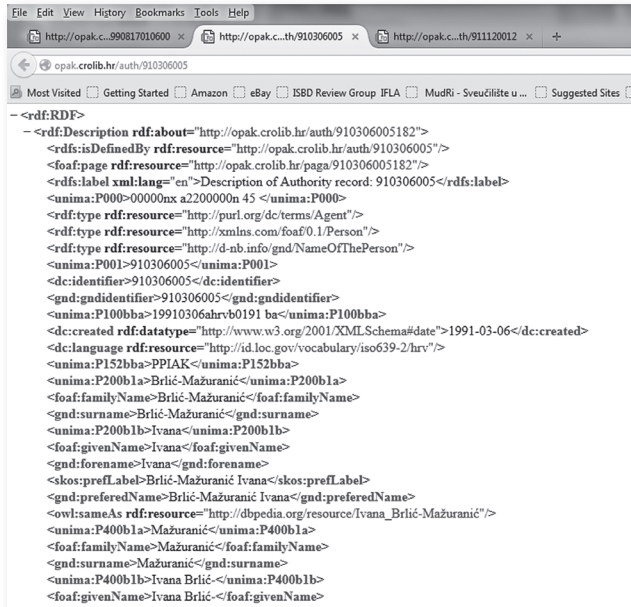
U projektu objavljivanja UNIMARC-zapisa autoriziranih podataka skupnog kataloga CROLIST kao povezanih otvorenih podataka slijedila se ista odluka donesena za bibliografske podatke. Naime potrebno je zadržati preciznost katalogiziranih zapisa/podataka, ali istovremeno objavljivanjem

21 Za primjer odluke odabira lokalnog rječnika za objavljivanje bibliografskih podataka vidjeti: Hamilton, Gill. Nacionalne knjižnice i otvoreni i povezani podaci: nove usluge pronalaženja i dostupnosti. // 16. seminar Arhivi, knjižnice, muzeji: mogućnosti suradnje u okruženju globalne informacijske infrastrukture: zbornik radova / uredile Nives Tomašević i Ivona Despot. Zagreb: Hrvatsko knjižničarsko društvo, 2013. Str. 57-65.

podataka manje granularnosti, tj. korištenjem općeg jezika, osigurati njihovo ponovno korištenje od strane sustavakoji nisu upoznati s izvornim rječnikom te time povećati mogućnosti povezivanja s vanjskim izvorima u oblaku podataka. Za prikaz izvornih podataka u formatu UNIMARC za autorizirane podatke korišten je iskonstruirani imenski prostor u skladu s metodom objavljenom u radovima Dunsire, Willer i Perožić,²² dok je za prikaz tih istih podataka usporednim korištenjem drugih imenskih prostora izrađena tablica podudarnosti elemenata formata UNIMARC/A s najčešće korištenim rječnicima BIO, gnd, FOAF, dct i elementi RDA Group 2 (osobe, korporativna tijela i obitelji) (Prilog 1.). Od navedenih rječnika u ovoj fazi projekta korišteni su samo imenski prostori rječnika unima, dc, gnd i foaf te sintaktičkih standarda²³ RDFS za identifikaciju izvornog zapisa, SKOS, tj. skos:prefLabel za ime autora ujedinstvenomodnosno autoriziranomobliku, tj. polja bloka 2-- , skos:altLabel za varijantni i srodni oblik imena, tj. polja bloka 4-- i 5-- i skos: editorial Note za polja bloka 8-- (815,810,820,825,830,835,836,886) te owl:sameAs za povezivanje s vanjskim izvorima, npr. dbpedia. Na Slici 3. prikazan je RDF-zapis za osobno ime usporednim korištenjem imenskih prostora više rječnika. Potpuni RDF-zapis vidjeti u prilogu 2. Potrebno je napomenuti da kodirani podaci formata UNIMARC nisu uzeti u obzir jer je ovdje prikazani rad eksperimentalan, ali i stoga što kodovi i njihove vrijednosti još nisu objavljeni kao SKOS-rječnici vrijednosti (*value vocabularies*).

22 Dunsire, G.; M. Willer; P. Perožić. Nav. dj.

23 Za sintaktičke standarde RDFS (Reource Description Framework Schema), SKOS (Simple Knowledge Organization System) i OWL (Web Ontology Language) vidjeti npr., w3schools.com. Introduction to RDF [citirano: 2014-05-23]. Dostupno na: http://www.w3schools.com/webservices/ws_rdf_intro.asp i W3C Semantic Web. SKOS: Simple Knowledge Organization System [citirano: 2014-05-23]. Dostupno na: <http://www.w3.org/2004/02/skos/>



Slika 3. RDF-zapis za osobno ime korištenjem imenskih prostora za format UNIMARC za autorizirane podatke (unima) i usporedo rječnika gnd, foaf i dc (dio)²⁴

Zaključno

Univerzalna bibliografska kontrola pretpostavlja da su nacionalne bibliografske službe obvezne objavljivati bibliografske i autorizirane podatke nacionalne produkcije. Nova tehnološka infrastruktura i tehnologije semantičkog weba za objavljivanje podataka kao povezanih otvorenih podataka ne mijenja tu obvezu, dapače. Mnoge su nacionalne knjižnice aktivne u istraživanju u tome području, kao i u objavljivanju svojih bibliografskih podataka. Za očekivati je dakle da će Nacionalna i sveučilišna knjižnica u Zagrebu također započeti slična istraživanja kao dio svojih obveza i odgovornosti.

U međuvremenu, pristupilo se istraživanju prikazivanja formata UNIMARC u RDF-u za zapise iz skupnog kataloga CROLIST. Projekt se bavio načelnim pitanjima upotrebe formata UNIMARC, ne bitno vezanim

24 RDF-zapis za <http://opak.crolib.hr/auth/910306005> iz autorizirane baze podataka CROLIST.

uz knjižnični sustav CROLIST, čime se željelo otvoriti novo područje istraživanja u Hrvatskoj, ali i stvoriti poligon za istraživanje u okviru međunarodnog projekta *UNIMARC in RDF*.

Kao i u slučaju prve faze projekta koji se bavio bibliografskim podacima, i u ovoj, drugoj, fazi istraživanja s autoriziranim podacima postavljena su načelno ista pitanja. Prvo je pitanje znamo li komu su potrebni/korisni takvi podaci. Gledano s aspekta koncepta Univerzalne bibliografske kontrole, i na toj tehnološkoj platformi potrebno je osigurati identifikaciju i okupljanje bibliografskih podataka, a za korisnika i kretanje katalogom, odnosno bibliografskim univerzumom. Koncept korisnika sada je potrebno proširiti na one znanstvenike/istraživače koji će sami moći pristupiti i koristiti povezanе otvorene podatke u skladu sa svojim partikularnim potrebama, ali i na pretraživače i agregatore povezanih otvorenih podataka koji su prvenstveno strojevi.

U odnosu na pitanja vezana uz objavljivanje imenskih prostora formata UNIMARC za autorizirane podatke, što uključuje i bibliografske podatke, uočava se potreba definiranja potpoljâ unutar određenog polja kao skupova elemenata, tzv. BAG, SEQence i ALT-skupine, a ujedno vezano uz ponovljivost i redoslijed potpoljâ unutar polja. Naime svaki podatkovni element formata UNIMARC objavljuje se kao zasebna RDF-rečenica – *triplet*, što olakšava pristup pretraživačima, ali predstavlja problem pri dizajnu usluga za prikaz podataka krajnjem korisniku, bez obzira u kojem formatu ili obliku.

Analiza postojećih rješenja pokazuje da su neke nacionalne knjižnice i skupni katalogi pribjegli praktičnim rješenjima i koriste samo jedan, općeprihvaćen rječnik u zajednici semantičkog weba, kao naprimjer skupni katalog VIAF koji koristi foaf. Druge su knjižnice razvile složene rječnike koji odgovaraju specifičnostima vlastite upotrebe autoriziranih podataka, kao što su Nacionalna knjižnica Njemačke koja je razvila rječnik gnd i Kongresna knjižnica koja koristi imenski prostor MADS razvijen na temelju formata MARC21 za autorizirane podatke. Treća vrsta ustanova koristi usporedo više rječnika s namjerom da se informacijska vrijednost izvornih podataka ne izgubi tijekom pretvorbe lokalnog formata u imenske prostore vanjskih rječnika. Odluka donesena u ovom projektu slijedi potrebu da se osigura prenošenje potpune vrijednosti izvornih podataka u novo okruženje, ali i potrebu da ti podaci budu identificirani, odnosno pronađeni i iznova korišteni u drugim sustavima i time povezani.

Projekt ukazuje na potrebu za daljnjim istraživanjima, posebno dopune tablica pretvorbi imenskih prostora različitih rječnika, uključivanje jedinstvene identifikacije entiteta²⁵ i identifikaciju ciljnih autoriziranih i vjerodostojnih izvora za automatsko povezivanje vlastitih podataka. U tom smislu treba istaknuti i u prethodnom radu prepoznatu potrebu za nadgradnjom infrastrukture koja bi uključila i tehničko rješenje koje bi osiguralo vanjskim *online* SPARQL-upitima zadovoljavajući odgovor i, time, povezivanje podataka.

Literatura

Bio: A vocabulary of biographical information [citirano: 2014-05-23]. Dostupno na: <http://vocab.org/bio/0.1/.html>

British Library Data Model – Book [citirano: 2014-05-23]. Dostupno na: <http://www.bl.uk/bibliographic/pdfs/bldatamodelbook.pdf>; British Library Metadata Services, Free Data Services [citirano: 2014-05-23]. Dostupno na: <http://www.bl.uk/bibliographic/datafree.html>

CERL Thesaurus [citirano: 2014-05-23]. Dostupno na: http://www.cerl.org/resources/cerl_thesaurus/main

DCMI International Conference on Dublin Core and Metadata Applications, Lisbon, September, 2013 [citirano: 2014-05-23]. Dostupno na: <http://dcevents.dublincore.org/IntConf/index/pages/view/rich-poor>

DCMI Metadata Terms [citirano: 2014-05-23]. Dostupno na: <http://dublincore.org/documents/dcmi-terms/>

Dunsire, Gordon; Mirna Willer; Predrag Perožić. Representation of the UNIMARC bibliographic data format in Resource Description Framework. // DC-2013, Lisbon, Portugal, [International Conference on Dublin Core and Metadata Applications]:

25 ISO 27729:2012 Information and documentation -- International standard name identifier (ISNI). Vidjeti i: [isni: International Standard Name Identifier \(ISO 27729\) \[Home\]](http://www.isni.org/) [citirano: 2014-05-23]. Dostupno na: <http://www.isni.org/>; MacEwan, Andrew; Anila Angjeli; Janifer Gatenby. The International Standard Name Identifier (ISNI): the evolving future of name authority control. // *Cataloging & Classification Quarterly* 51, 1-3 (2013), str. 55-71.

Linking to the Future, Lisbon, 2 - 6 September, 2013 [citirano: 2014-05-23]. Dostupno na: <http://dcevents.dublincore.org/IntConf/dc-2013/paper/view/173/91>

Dunsire, Gordon; Willer, Mirna. UNIMARC and Linked Data. // IFLA Journal 37, 4 (December 2011), 314-326 [citirano: 2014-05-23]. Dostupno na: http://www.ifla.org/files/hq/publications/ifla-journal/ifla-journal-37-4_2011.pdf

Europeana Data Model (EDM) documentation [citirano: 2014-05-23]. Dostupno na: <http://pro.europeana.eu/edm-documentation>

FOAF: Friend of a Friend [citirano: 2014-05-23]. Dostupno na: <http://www.foaf-project.org/>

Geo Names Ontology [citirano: 2014-05-23]. Dostupno na: <http://www.geonames.org/ontology/documentation.html>

GND Ontology [citirano: 2014-05-23]. Dostupno na: <http://d-nb.info/standards/elementset/gnd>

Hamilton, Gill. Nacionalne knjižnice i otvoreni i povezani podaci: nove usluge pronalaznja i dostupnosti. // 16. seminar Arhivi, knjižnice, muzeji: mogućnosti suradnje u okruženju globalne informacijske infrastrukture: zbornik radova / uredile Nives Tomašević i Ivona Despot. Zagreb: Hrvatsko knjižničarsko društvo, 2013. Str. 57-65.

Hillmann, Dianne; Dunsire, Gordon; Phipps, Jon. Maps and gaps: strategies for vocabulary design and development. // Proc. Int'l Conf. on Dublin Core and Metadata Applications 2013, 82-89 [citirano: 2014-05-23]. Dostupno na: <http://dcevents.dublincore.org/IntConf/dc-2013/paper/view/185/80>

IFLA UNIMARC Strategic Programme [citirano: 2014-05-23]. Dostupno na: <http://www.ifla.org/unimarc>

ISO 27729:2012 Information and documentation -- International standard name identifier (ISNI).

isni: International Standard Name Identifier (ISO 27729) [Home] [citirano: 2014-05-23]. Dostupno na: <http://www.isni.org/>

Jelenković, Leonardo; Willer, Mirna. Kako do povezanih podataka iz UNIMARC zapisa? // 16. seminar Arhivi, knjižnice, muzeji: mogućnosti suradnje u okruženju globalne informacijske infrastrukture: zbornik radova / uredile Nives Tomašević i Ivona Despot. Zagreb: Hrvatsko knjižničarsko društvo, 2013. Str. 30-56.

MacEwan, Andrew; Angjeli, Anila; Gatenby, Janifer. The International Standard Name Identifier (ISNI): the evolving future of name authority control. // *Cataloging & Classification Quarterly* 51, 1-3 (2013), str. 55-71.

MADS/RDF Primjer [citirano: 2014-05-23]. Dostupno na: <http://www.loc.gov/standards/mads/rdf/#t11>

An organization ontology [citirano: 2014-05-23]. Dostupno na: <http://www.epimorphics.com/public/vocabulary/org.html>

Preliminary version of RDA [citirano: 2014-05-23]. Dostupno na: <http://metadata-registry.org/schema/show/id/15.html>

VIAF: Virtual International Authority File [citirano: 2014-05-23]. Dostupno na: <http://viaf.org/>

Willer, Mirna; Dunsire, Gordon; Perožić, Predrag. The UNIMARC in RDF project: namespaces and linked data. // IFLA World Library and Information Congress, 79th IFLA General Conference and Assembly, 17-23 August 2013, Singapore. Session 222 — Expanding MARC metadata services with linked open data — UNIMARC Core Activity (UCA) [citirano: 2014-05-23]. Dostupno na: <http://library.ifla.org/156/1/222-willer-en.pdf>

W3C Semantic Web Interest Group. Basic Geo (WGS84 lat/lon) Vocabulary [citirano: 2014-05-23]. Dostupno na: <http://www.w3.org/2003/01/geo/>

W3C Semantic Web. SKOS: Simple Knowledge Organization System [citirano: 2014-05-23]. Dostupno na: <http://www.w3.org/2004/02/skos/>

w3schools.com. Introduction to RDF [citirano: 2014-05-23]. Dostupno na: http://www.w3schools.com/webservices/ws_rdf_intro.asp

Prilog 1. Usporedna analiza rječnika autoriziranih podataka:

UNIMARC/A		BIO ¹	DNB GND Ontology ²
Oznaka zapisa		Klasa	Klasa
Mjesto znaka 9 Vrsta entiteta	a = osobna odrednica	Class: foaf:Person	NameOfThePerson
	b = korporativna odrednica		CorporateBody ConferenceOrEvent
	c = političko-teritorijalni ili zemljopisni naziv		TerritorialCorporate Body
	e = obiteljsko ime		Family
	f = jedinstveni stvarni naslov		Work
	j = opći pojam		SubjectHeading SubjectHeadingSensoS
	k = pristupnica prema mjestu izdavanja		PlaceOrGeographicName
	l = oblik, žanr ili materijalna obilježja		
0-- Blok za identifikaciju			
001			gndIdentifier
1-- Blok kodiranih podataka			
2-- Blok odrednica			
2--			preferredName
200	\$a-\$g		preferredNameEntity ForThePerson
200 Ind. #0	\$a		
200 Ind. #1	\$a		surname
	\$b		forename
	\$c		
	\$d		
	\$f	bio:date	dateOfBirthAndDeath
	\$g		
	\$4		

UNIMARC/A		BIO ¹	DNB GND Ontology ²
	\$j-\$z; \$7, \$8		
210			
210 IND 0_			preferredNameForThe CorporateBody
210 IND 1_			preferredNameForThe ConferenceEvent
215			? Class: Territorial CorporateBody?
220			preferredNameForThe Family
230			preferredNameForThe Work
250			preferredNameForThe SubjectHeading
260			preferredNameFor PlaceOrGeographicName
280			
3-- Blok napomena			
300			?biographicalOrHistorical Information
340			biographicalOrHistorical Information
4-- Blok općih uputnica			
4--			variantName
400	\$a-\$g		variantNameEntityFor ThePerson
400 Ind. #0	\$a		
400 Ind. #1	\$a		surname
	\$b		forename
	\$c		
	\$d		
	\$f	bio:date	dateOfBirthAndDeath
	\$g		
	\$4		
	\$j-\$z; \$7, \$8		
410			
410 IND 0_			variantNameForThe CorporateBody

FOAF ³	DCT ⁴	RDA Namespace ⁵	SKOS/OWL/rdfs
foaf:name			
foaf:name			
			skos:altLabel
foaf:name			
foaf:givenName			
foaf:familyName			
foaf:givenName			

UNIMARC/A		BIO ¹	DNB GND Ontology ²
410 IND 1_			variantNameForThe ConferenceEvent
415			? Class: Territorial CorporateBody?
420			variantNameForThe Family
430			variantNameForThe Work
450			variantdNameForThe SubjectHeading
460			variantNameForPlace OrGeographicName
480			
5-- Blok unakrsnih uputnica			
5--			variantName
6-- Blok klasifikacijskih oznaka			
676			[relatedDdcwithDegree OfDereminacy...]
7-- Blok za povezivanje odrednica			
Ne koristi se			
8-- Blok izvora podataka			
Povezivanje s drugim izvorima			

Prilog 2. Potpuni RDF-zapis za osobno ime korištenjem imenskih prostora za format UNIMARC za autorizirane podatke (unima) i usporedo rječnika gnd, foaf i dc

```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:unimb="http://iflastandars.info/ns/unimarc/unimarcb/elements/"
xmlns:unima="http://iflastandars.info/ns/unimarc/unimarca/elements/"
xmlns:rdaRelGr2="http://metadataregistry.org/uri/schema/
RDARelationshipsGR2" xmlns:rdaGr2="http://rdvocab.info/ElementsGr2/"
xmlns:rdaGr3="http://rdvocab.info/ElementsGr3/" xmlns:edm="http://europe-
ana.eu/terms/" xmlns:gn="http://www.geonames.org/ontology#"
xmlns:gnd="http://d-nb.info/gnd/" xmlns:bio="http://vocab.org/bio/0.1/"
xmlns:foaf="http://xmlns.com/foaf/0.1/" xmlns:rel="http://purl.org/vocab/frbr/
core#" xmlns:bibo="http://purl.org/ontology/bibo/" xmlns:dc="http://purl.org/
dc/terms/" xmlns:owl="http://www.w3.org/2002/07/owl#" xmlns:skos="http://
www.w3.org/2004/02/skos/core#" xmlns:rda="http://RDVocab.info/Elements"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"><rdf:Description
rdf:about="http://opak.crolib.hr/auth/910306005182"><rdfs:isDefinedBy
rdf:resource="http://opak.crolib.hr/auth/910306005"/><foaf:page
rdf:resource="http://opak.crolib.hr/paga/910306005182"/><rdfs:label
xml:lang="en">Description of Authority record: 910306005</
rdfs:label><unima:P000>00000nx a2200000n 45 </unima:P000><rdf:type
rdf:resource="http://purl.org/dc/terms/Agent"/><rdf:type rdf:resource="http://
xmlns.com/foaf/0.1/Person"/><rdf:type rdf:resource="http://d-nb.info/gnd/Nam
eOfThePerson"/><unima:P001>910306005</unima:P001><dc:identifi
er>910306005</dc:identifier><gnd:gndidentifier>910306005</gnd:gndidentifier>
<unima:P100bba>19910306ahrvb0191 ba</unima:P100bba><dc:created
rdf:datatype="http://www.w3.org/2001/XMLSchema#date">1991-03-06</
dc:created><dc:language rdf:resource="http://id.loc.gov/vocabulary/iso639-2/
hrv"/><unima:P152bba>PPIAK</unima:P152bba><unima:P200b1a>Brlić-
Mažuranić</unima:P200b1a><foaf:familyName>Brlić-Mažuranić</foaf:familyNa
me><gnd:surname>Brlić-Mažuranić</gnd:surname><unima:P200b1b>Ivana</uni
ma:P200b1b><foaf:givenName>Ivana</foaf:givenName><gnd:forename>Ivana</
gnd:forename><skos:prefLabel>Brlić-Mažuranić Ivana</skos:prefLabel><gnd:pr
eferredName>Brlić-Mažuranić Ivana</gnd:preferredName><owl:sameAs
rdf:resource="http://dbpedia.org/resource/Ivana_Brlić-Mažuranić"/><unima:P
```

400b1a>Mažuranić</unima:P400b1a><foaf:familyName>Mažuranić</foaf:familyName><gnd:surname>Mažuranić</gnd:surname><unima:P400b1b>Ivana Brlić</unima:P400b1b><foaf:givenName>Ivana Brlić</foaf:givenName><gnd:forename>Ivana Brlić</gnd:forename><skos:altLabel>Mažuranić Ivana Brlić</skos:altLabel><gnd:variantName>Mažuranić Ivana Brlić</gnd:variantName><gnd:variantName rdf:parseType="Resource"><gnd:surname>Mažuranić</gnd:surname><gnd:forename>Ivana Brlić</gnd:forename></gnd:variantName><unima:P400b1a>Brlić Mažuranić</unima:P400b1a><foaf:familyName>Brlić Mažuranić</foaf:familyName><gnd:surname>Brlić Mažuranić</gnd:surname><unima:P400b1b>Ivana</unima:P400b1b><foaf:givenName>Ivana</foaf:givenName><gnd:forename>Ivana</gnd:forename><skos:altLabel>Brlić Mažuranić Ivana</skos:altLabel><gnd:variantName>Brlić Mažuranić Ivana</gnd:variantName><gnd:variantName rdf:parseType="Resource"><gnd:surname>Brlić Mažuranić</gnd:surname><gnd:forename>Ivana</gnd:forename></gnd:variantName><unima:P400b1a>Mažuranić</unima:P400b1a><foaf:familyName>Mažuranić</foaf:familyName><gnd:surname>Mažuranić</gnd:surname><unima:P400b1b>Ivana Brlić</unima:P400b1b><foaf:givenName>Ivana Brlić</foaf:givenName><gnd:forename>Ivana Brlić</gnd:forename><skos:altLabel>Mažuranić Ivana Brlić</skos:altLabel><gnd:variantName>Mažuranić Ivana Brlić</gnd:variantName><gnd:variantName rdf:parseType="Resource"><gnd:surname>Mažuranić</gnd:surname><gnd:forename>Ivana Brlić</gnd:forename></gnd:variantName><unima:P400b1a>Brlićova-Mažuranićova</unima:P400b1a><foaf:familyName>Brlićova-Mažuranićova</foaf:familyName><gnd:surname>Brlićova-Mažuranićova</gnd:surname><unima:P400b1b>Ivana</unima:P400b1b><foaf:givenName>Ivana</foaf:givenName><gnd:forename>Ivana</gnd:forename><skos:altLabel>Brlićova-Mažuranićova Ivana</skos:altLabel><gnd:variantName>Brlićova-Mažuranićova Ivana</gnd:variantName><gnd:variantName rdf:parseType="Resource"><gnd:surname>Brlićova-Mažuranićova</gnd:surname><gnd:forename>Ivana</gnd:forename></gnd:variantName><unima:P400b1a>Mažuranićova</unima:P400b1a><foaf:familyName>Mažuranićova</foaf:familyName><gnd:surname>Mažuranićova</gnd:surname><unima:P400b1b>Ivana Brlićova</unima:P400b1b><foaf:givenName>Ivana Brlićova</foaf:givenName><gnd:forename>Ivana Brlićova</gnd:forename><skos:altLabel>Mažuranićova Ivana Brlićova</skos:altLabel><gnd:variantName>Mažuranićova Ivana Brlićova</gnd:variantName><gnd:variantName rdf:parseType="Resource"><gnd:surname>Mažuranićova</gnd:surname><gnd:forename>Ivana Brlićova</gnd:forename></gnd:variantName>

e><unima:P400b1a>Berliq-Mazhuraniq</unima:P400b1a><foaf:familyName>Berliq-Mazhuraniq</foaf:familyName><gnd:surname>Berliq-Mazhuraniq</gnd:surname><unima:P400b1b>Ivana</unima:P400b1b><foaf:givenName>Ivana</foaf:givenName><gnd:forename>Ivana</gnd:forename><skos:altLabel>Berliq-Mazhuraniq Ivana</skos:altLabel><gnd:variantName>Berliq-Mazhuraniq Ivana</gnd:variantName><gnd:variantName rdf:parseType="Resource"><gnd:surname>Berliq-Mazhuraniq</gnd:surname><gnd:forename>Ivana</gnd:forename></gnd:variantName><unima:P400b1a>Mazhuraniq</unima:P400b1a><foaf:familyName>Mazhuraniq</foaf:familyName><gnd:surname>Mazhuraniq</gnd:surname><unima:P400b1b>Ivana Berliq</unima:P400b1b><foaf:givenName>Ivana Berliq</foaf:givenName><gnd:forename>Ivana Berliq</gnd:forename><skos:altLabel>Mazhuraniq Ivana Berliq</skos:altLabel><gnd:variantName>Mazhuraniq Ivana Berliq</gnd:variantName><gnd:variantName rdf:parseType="Resource"><gnd:surname>Mazhuraniq</gnd:surname><gnd:forename>Ivana Berliq</gnd:forename></gnd:variantName><unima:P675bba>886.2</unima:P675bba><dc:subject>UDC 886.2</dc:subject><unima:P675bbc>hrvatska književnost</unima:P675bbc><unima:P801b0a>HR</unima:P801b0a><unima:P801b0b>NSK</unima:P801b0b><dc:rights>HR, NSK</dc:rights><unima:P810bba>Zgode i nezgode šegrta Hlapića / Ivana Brlić-Mažuranić. Zagreb, 1995.</unima:P810bba><skos:editorialNote xml:lang="hr">Zgode i nezgode šegrta Hlapića / Ivana Brlić-Mažuranić. Zagreb, 1995.</skos:editorialNote><unima:P810bba>Podivuhodne pribehy ševcovskeho učne / Ivana Brlićova-Mažuranićova. V Praze, 1930.</unima:P810bba><skos:editorialNote xml:lang="hr">Podivuhodne pribehy ševcovskeho učne / Ivana Brlićova-Mažuranićova. V Praze, 1930.</skos:editorialNote><unima:P810bba>Peripetite e guditshme te shegert / Ivana Berliq-Mazhuraniq. Prishtine, 1973. RK/i</unima:P810bba><skos:editorialNote xml:lang="hr">Peripetite e guditshme te shegert / Ivana Berliq-Mazhuraniq. Prishtine, 1973. RK/i</skos:editorialNote><unima:P830bba>Hrvatska književnica, 1874.-1938.</unima:P830bba><skos:editorialNote xml:lang="hr">Hrvatska književnica, 1874.-1938.</skos:editorialNote><unima:P998bbp>0101TV</unima:P998bbp><unima:P998bbp>0311AR</unima:P998bbp><unima:P998bbp>0603ARi</unima:P998bbp></rdf:Description></rdf:RDF>