

## СЕМАНТИЧКИ ВЕБ И LINKED (OPEN) DATA МОГУЋНОСТИ И ПЕРСПЕКТИВЕ ЗА БИБЛИОТЕКЕ

**Станислава Гардашевић**, stanislava.gardasevic@nb.rs, Народна библиотека Србије,  
Одељење за развој дигиталне библиотеке и микрографију, Београд, Србија

### Апстракт

Једна од новијих тема у библиотекарству јесте примена технологија семантичког веба и Linked (Open) Data у циљу управљања библиотечким метаподацима. У првом делу овог рада дат је преглед основних принципа и модела семантичког веба (XML, RDF, URI, онтологије итд.), као и примена неких од тих принципа у библиотекама и другим културним институцијама, са нарочитим освртом на примену у Еуропеани, тзв. Европској дигиталној библиотеци, архиву и музеју. Други део рада осврће се на Linked (Open) Data, као саставни и неизоставни део семантичког веба. Осим основних принципа, наведене су и кључне предности објављивања библиотечких метаподатака као Linked (Open) Data, како за организације тако и за кориснике. Коначно, побројане су неке од иницијатива које су већ кренуле овим путем и дата је препорука за преиспитивање досадашње библиотечке праксе, као и њено унапређивање тиме што ће се пратити ови, актуелни трендови, у циљу побољшања услуга и боље видљивости библиотечких података.

### Кључне речи

семантички веб, Linked (Open) Data, метаподаци, онтологија, URI

### 1. Увод

Чак и пре настанка интернета, библиотекари су се трудили да корисницима олакшају претрагу библиотечких каталога. Изумом машински читљивих формата за каталогизацију (MARC) крајем шездесетих година и развојем Међународног стандарда за библиографски опис (ISBD) почетком

седамдесетих година XX века, почиње прелазак са лисног на аутоматизовани каталог. (McCallum 2002) Библиотеке су унапредовале толико да је данас могуће претражити велики део библиотечких фондова у свету преко сервиса за обједињену претрагу каталога попут WorldCat или COBISS.NET.

Како би омогућили услуге обједињене

претраге, библиотекарима су морали да постигну компатибилност међу стандардима за размену података и комуникацију. Ови стандарди (тзв. шеме за енкодирање метаподатака) служе успостављању метода којима дескриптивни подаци могу да се деле међу институцијама и на тај начин омогуће обједињену претрагу. (Fox 2001) У библиотечком свету UNIMARC, MARC и Даблинско језгро су примери оваквих стандарда. Убрзани пораст броја ресурса на интернету и дигиталних колекција пропраћен је и развојем великог броја шема метаподатака, од којих је свака базирана на захтевима различитих корисничких профила, заједница, врста материјала, предметних оквира, потреба пројеката итд. Ипак, још увек није постигнута потпуна компатибилност међу шемама за енкодирање метаподатака. Проблем при грађењу великих дигиталних библиотека или репозиторијума са метаподацима настаје углавном због нестандардизованости описа, нарочито у мањим институцијама. (Chan and Zeng 2006)

Без обзира на овај проблем, који представља први корак у постизању компатибилности у раду са библиотечким метаподацима, стручна и научна јавност већ извесно време ради на постизању семантичке повезаности не само библиотечких ресурса, већ и оних који долазе из различитих културних, научних и сродних институција. Поента овог рада је да се омогући што лакше истраживања одређене теме. У овом случају, корисник не мора да зна тачну врсту грађе на којој се налази жељена информација, у којој институцији се она чува, адресу и различите процедуре коришћења и проналажења тих докумената. Крајњи циљ је обједињена претрага којом се групишу све сродне јединице, које се притом повезују и са релевантним изворима са веба. Овај сценарио могуће је остварити путем сервиса које нам пружа интернет, а помоћу технологија на којима почивају Семантички веб и његова

апликација Linked (Open) Data.

Овај рад покушава описати на који начин горе поменуте технологије функционишу и у којој мери би њихова имплементација допринела библиотечкој струци.

## 2. Семантичка повезаност

Питањем семантичке повезаности бави се дисциплина под називом представљање знања (Knowledge Representation). Ова дисциплина бави се решавањем неких од кључних проблема информационог друштва: како конструисати и чувати информације и како их пронаћи и преузети на најпрецизнији и најефикаснији начин. (Weller 2010) Поред додељивања тзв. формалних или библиографских метаподатака, у библиотечком свету постоји добро утврђена пракса предметизације и класификације, којом се одређена јединица описује предметним одредницама, представљеним садржајно-описним кључним речима или се смешта у неку од класификационих шема. (Cleveland and Cleveland 2001) За ове потребе развијени су системи за управљање знањем (Knowledge Organisation Systems-KOS), који се користе у сврху организације докумената, са циљем њиховог лакшег проналажења и организације колекција.

У зависности од садржаја, структуре и начина организације ови системи као термилошки ресурси могу бити листе термина (нормативне датотеке, речници итд), класификације и категорије (предметне одреднице, шеме за класификацију и категоризацију и таксономије), релационе листе (тезауруси, семантичке мреже и онтологије). Ови системи могу имати улогу контролираних речника. Они могу да дефинишу и односе у којима одређени концепти могу да се нађу (углавном у односу хијерархије или другим семантичким везама). Што се више користе семантичке

везе у систему организације знања, то је он комплекснији по својој структури. (Hoge 2000)

Хармонизацијом система за управљањем знањем које одржавају различите институције, односно компатибилношћу формата семантичком компатибилношћу класификационих шема, могуће је успоставити семантичку повезаност метаподатака који описују њихову грађу. Као једно од потенцијалних решења за постизање семантичке компатибилности класификационих шема користе се принципи на којима се темељи Семантички веб (Semantic Web или Семантичка мрежа).

### 3. Семантички веб

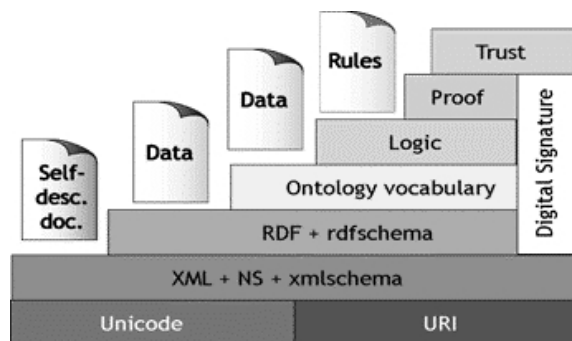
Визију семантичког веба изнео је творац World Wide Web-а (WWW) Тим Бернерс Ли, још давне 1994. године на првој WWW конференцији. (Shadbolt et al. 2006) Бернерс Ли је данашњи веб у ствари замислио као семантички тј. као „мрежу података“, како га он још дефинише, а не као „мрежу докумената“, какав је данас. Бернерс Ли је и даље један од највећих заговорника ове визије, само што је сада подржан од стране великог броја научника и професионалаца из најразличитијих сфера.

„У овом сценарију визија интероперабилности података заснована је на изазову и идеји интеграције информација: све информације које кориснику требају, можда нису доступне на једном месту (нпр. на истом веб сајту) већ могу бити сакупљене са неколико страница, а затим презентоване на интегрисани начин.“ (Shadbolt et al. 2006) Семантички веб нуди могућност издизања механизма за проналажење информације на “семантички” ниво, на којем је могућ њихов софистициран опис, али и успостављање вишег нивоа разумевања између корисника и сервиса, путем достављања информација кроз

размену онтологија, које пружају неопходни вокабулар за дијалог. (Fensel et al. 2005) Често називан веб 3.0, Семантички веб би требало да омогући машинама да уместо сирових података обрађују знање које ти подаци носе, у процесима попут људског дедуктивног резонувања и закључивања. Услов за остваривање тога је огромна количина јавно доступних података у стандардизованом формату. (Republički zavod za statistiku)

#### 3.1 Саставни делови семантичког веба

Најпознатија илустрација семантичког веба (Semantic Web Layer Cake, слика 1) приказује делове-слојеве Семантичког веба, од којих ће неки бити укратко представљени.



Слика 1: Слојеви на којима почива технологија Семантичког веба, тзв. Semantic Web Layer Cake (Miller 2001)

XML представља синтаксички слој и то је метајезик који дозвољава корисницима да дефинишу ознаке за своја документа, уз помоћ етикета. NS (namespaces) су простори имена, који означавају из ког простора имена потиче одређена XML етикета, нпр. етикета dc:title је ознака за наслов из Даблинског језгра (Dublin Core), стандарда за метаподатке који се састоји од сета елемената за опис широког спектра извора на мрежи. Ипак, XML није средство за изражавање семантике (т.ј. значења) података.

За ове сврхе користи се Оквир за описивање ресурса (Resource Description Framework, RDF). Помоћу RDF-а извори могу бити описани семантички значајним међусобним везама, путем представљања структуриране информације о било ком ресурсу у облику једноставног исказа у форми уређене тројке: субјекат, предикат, објекат. (RDF primer)

Пример RDF тројке може бити:

<пр:беснило>, <пр:има\_аутора>, <пр:борислав\_пекић>, која описује књигу „Беснило“.

Поред исказа, остала два фундаментална концепта RDF-а су извори (resources) и својства (properties). Извори су “ствари” о којима желимо да говоримо, које у изјави могу бити и субјекат и предикат и објекат.

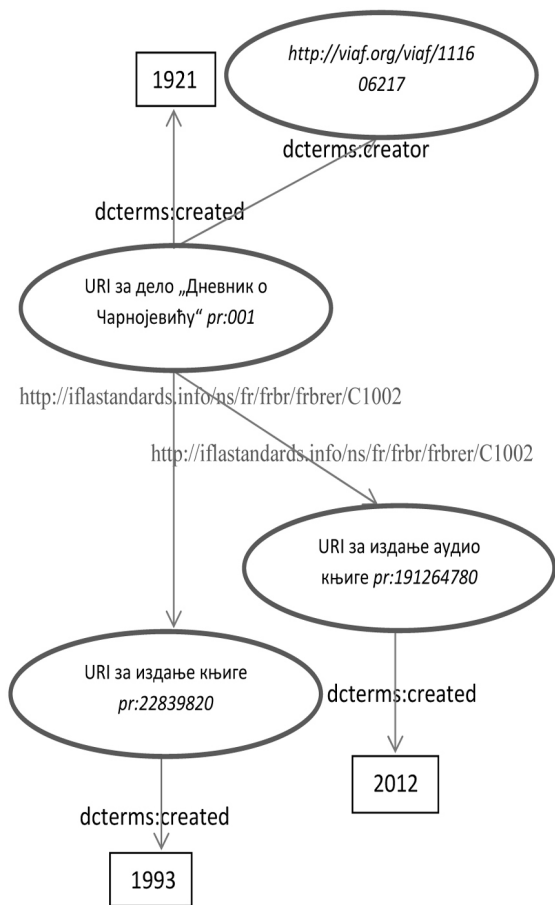
За идентификацију извора користе се Јединствени идентификатори извора (Uniform Resource Identifiers, URI). Главна одлика URI-ја јесте да је сваки глобално јединствен и да различите особе или установе могу да их дефинишу и користе за идентификацију ресурса. (Doeff at al. 2010)

Својства су посебна врста ресурса који описују односе међу њима, могу бити различитог типа, нпр. “има аутора”, “има издавача”, “има наслов” итд. и увек су у функцији предиката у исказу. Пошто објекат једне изјаве може бити субјекат неке друге, RDF омогућава да се једноставне изјаве представе путем графа у којем субјекти и предикати представљају чворове (nodes), који су повезани луковима (својствима). Док својства увек морају бити URI-ји, чворови у RDF графу могу имати ниску за вредност, могу бити празни (у случајевима када немају форму и идентификацију независно од локалног графа, што није препоручљиво) или могу бити URI. (RDF primer)

Слика 2 представља пример RDF графа, који описује дело „Дневник о Чарнојевићу“ (pr:001). Уз помоћ својства из dcterms<sup>2</sup>

простора имена исказани су аутор дела (URI из Virtual International Authority Files, VIAF базе<sup>3</sup>) и када је оно настало (ниска 1921). Такође, исказује две манифестације овог дела (по Функционалним захтевима за библиографске записе-FRBR вокабулару<sup>3</sup>), од којих је једна књига издата 1993. године, а друга аудио књига из 2012. Путем једноставних исказа, овај граф би теоретски могао бесконачно да се шири, док се не забележе све релевантне информације.

Даље, уз помоћ језика као што је RDF шема (RDF Schema, RDFS, на слици 1 трећи ред одоздо) могуће је дефинисати вокабуларе који ће се користити у изјавама. (RDF primer) Сви ресурси који се описују на овај начин могу бити типизирани, а за субјекте и објекте то се ради тако што се декларишу као инстанце неке класе, којом се дефинише тај њихов тип (нпр. pr:22809820 и pr:19126780 могу бити декларисани као инстанце класе Библиографски ресурс). Шта више, за класе из одређене области знања, задају се њихови могући међусобни односи (својства) путем декларације ограничења и правила. Ти односи међу класама се дефинишу путем онтологија.



Слика 2: Пример описа дела и његова два издања

### 3.2 Онтологије

Најкоришћенија дефиниција онтологије (у рачунарском, не у оригиналном филозофском смислу) јесте Груберова, који је назива “спецификацијом концептуализације.” (Gruber 1993) Тачније, онтологија је формална, експлицитна спецификација неке заједничке концептуализације. Концептуализација представља апстрактни модел неке појаве, који идентификује релевантне концепте те појаве. Експлицитност у овој дефиницији подразумева експлицитно дефинисање врста коришћених концепата и ограничења њиховог коришћења. Коначно, појам формалности

у овој дефиницији односи се на чињеницу да би онтологија требало бити у машински читљивом формату. (Studer et all. 1998)

Онтологије се углавном дефинишу помоћу већ поменутих RDF шеме (RDFS), и Web Ontology Language (OWL) стандарда. За потребе постављања упита над RDF подацима развијени су различити језици, а најпознатији је SPARQL (Simple Protocol and RDF Query Language) који дозвољава комплексније упите него његова претеча SQL. Даље, аутоматски механизми расуђивања (reasoners) анализирају структуру, исказе дескриптивне логике онтологије, и изводе информацију која није унесена директно. Дакле, на основу експлицитно кодираног знања, изводи се ново, имплицитно знање. (Gomez-Perez et al. 2003)

Из овако структурираних података, било би могуће добити информацију нпр. о учесталости издавања дела Милоша Црњанског, док је он био у емиграцији, уколико нам је потребна за анализу. Пример упита који би могао да се постави је: врати информацију где и на колико је година у просеку је објављено издање дела pr:001, од првог до 1941. године (када је имигрирао), потом од 1941. до 1965. године (током имиграције) и коначно од 1965. до данас (након имиграције).

### 4. Примена у библиотекама и другим културним институцијама

Онтологије су препознате као један од најсвеобухватнијих система за опис грађе која представља културно, историјско и научно наслеђе, управо зато што дозвољавају тзв. концептуални приступ. Постоји неколико вокабулара-онтологија које могу да се користе за потребе библиографског описа (Даблинско језгро, BIBO, ISBD, RDA, FRBR итд.) и најчешће се употребљавају за потребе дигиталних библиотека.

Пример успешног коришћења онтологије као шеме маетаподатака јесте Europeana Data Model (EDM), развијен за потребе Европеане<sup>4</sup>, европске дигиталне библиотеке, музеја и архива. Мотив иза стварања оваквог модела био је да се омогући компатибилност у раду са описима јединица које потичу из различитих културно-историјских институција, а да се притом задржи пуноћа основног формата у којима се ти описи бележе. Аутори EDM модела га описују као онтологију вишег нивоа, која је по структури веома општа, али са којом могу да се повежу специфичне онтологије које покривају одређени домен (архиве, библиотеке, музеје аудио-визуалне збирке итд). EDM у себи садржи и већ постојеће вокабуларе-онтологије (Даблинско језгро, SKOS итд.). Осим техничке компатибилности у раду, визија иза имплементације новог модела у Европеани је стварање могућности богате семантичке концептуализације за објекте чији се описи налазе у њеној бази (за сада 25 милиона), како би могле да се изводе комплексније операције над њима.

Једна од познатијих онтологија је и Simple Knowledge Classification System (SKOS), веб стандард за представљање тезауруса у RDF-у. (Doerr et al. 2010) Пошто је пракса коришћења контролисаних вокабулара у библиотекама веома заступљена, овај стандард се користи, између осталог, за спајање семантички еквивалентних концепата из различитих контролисаних вокабулара помоћу skos:exactMatch својства. Европеана користи SKOS како би изградила семантички слој података, којим би се повезала сродна грађа из различитих институција. Ипак, чињеница је да многе институције које учествују у Европеани немају своје контролисане вокабуларе и самим тим ће овај процес бити знатно отежан или барем ограничен. (Olensky 2010)

Ипак, ове технологије су нашле више

примене у музејима. Оне дозвољавају опис који се односи не само на објекат, већ и на догађај који је на неки начин везан за њега, а за опис музејске грађе су битне информације везане за догађаје проналаска артефакта или када је коришћен. Најпознатија онтологија створена за потребе описа музејске грађе је CIDOC-Conceptual Reference Model (CIDOC-CRM).

Тимови из Европеане и сродних пројеката, који су везани за истраживања на тему олакшаног приступа информацијама, сматрају да је могуће да се ресурси који долазе из различитих институција културе међусобно повежу уједначавањем контролисаних вокабулара који се користе у овим институцијама, као и систематским повезивањем са информацијама из Linked (Open) Data, о чијим квалитетима ће бити речи у наредном тексту. Европеана са EDM моделом настоји да “обогати” прикупљене метаподатке њиховим повезивањем са LOD-ом. То се изводи тако што се уместо ниске, која је обично вредност елемента-поља описа, користите линкови објављени као LOD. На пример, имена аутора се замењују са URI-има из VIAF базе, која представља унификацију ауторских нормативних датотека из 16 библиотека, а географска имена са онима из GeoNames<sup>5</sup> базе за географске локације, који у себи садрже много више података него што је само име те географске локације

## 5. Linked (Open) Data

Као што је већ поменуто, Семантички веб би требало да представља мрежу података. Ипак, да би ова мрежа постала стварност, неопходно је да се велики број података на веб објави у стандардном формату, тако да може да се обрађује помоћу технологија на којима почива Семантички веб. Шта више, да би се ова мрежа података успоставила, неопходно је и то да подаци буду и у неком

међусобном односу (за разлику од обичне колекције скупова података). Ова база међусобно повезаних скупова података на вебу се назива Linked (Open) Data или повезани подаци. (Linked Data) Linked Data или, како га још називају поборници покрета слободног приступа информацијама, Linked Open Data (LOD) је скуп технологија и стандарда који омогућују Семантички веб. Ова иницијатива проистекла је из World Wide Web (W3) конзорцијума<sup>6</sup> са фокусом на повезивање података на вебу и циљем да се да значења подацима на вебу. (Јурић 2011)

Сам појам Linked Data односи се на скуп примера добре праксе или начела за објављивање и међусобно повезивање структурисаних података на вебу. Ова начела је представио Тим Бернерс Ли у раду “Linked Data”, а они су:

1. Користи URI-је као имена за појмове (при томе се не мисли само на појмове који означавају људе и места, већ и оне апстрактније попут веза међу њима).
2. Користи HTTP URI-је, тако да људи могу претраживати та имена.
3. При претрази одређеног URI-ја, омогућити корисне информације везане за њега, користећи већ дефинисане стандарде (RDF, SPARQL).
4. Укључи линкове на друге URI-је, тако да људи могу открити више. (Berners-Lee 2006)

У Linked (Open) Data, URI-ји се користе аналогно идентификаторима за нормативну контролу у традиционалном библиотекарству. URI може бити универзални локатор ресурса (Uniform Resource Locator, URL), интернационализовани ресурсни идентификатор (International Resource Identificaton, IRI), или било која друга врста јединственог идентификатора. (Antoniou and Van Harmelen 2008)

Уз појам Linked (Open) Data се често додаје именица облак (тзв. LOD cloud<sup>7</sup>) и графички је представљен као дијаграм међусобно повезаних скупова података. Број скупова података се повећава до те мере да врло вероватно ускоро неће ни моћи бити приказан у једном графу, док је DBpedia графички приказана као највећи део облака. DBpedia<sup>8</sup> је јавна иницијатива при којој се извлаче структурисане информације из Википедије, како би се омогућило постављање комплекснијих упита. DBpedia<sup>8</sup> је тренутно у верзији 3.7 и описује 3.64 милиона “ствари”, од чега је 1.83 милиона класификовано конзистентном онтологијом.

Битно је нагласити да се путем Linked (Open) Data повезују структурисани подаци, а управо оно чиме библиотеке располажу јесу велике количине добро структурисаних и богатих података који су везани за одређени ресурсе и изворе људског знања. Из тог разлога, библиотеке представљају једно од најбољих поља за имплементацију Linked (Open) Data технологије. Та чињеница препозната је и од стране W3 конзорцијума, у оквиру којег је маја 2010. године отворена група под именом Library Linked Data Incubator Group. Неколико престижних међународних институција основале су ову групу која окупља заједницу заинтересовану за библиотечке метаподатке и LOD са циљем да се направи извештај који ће помоћи побољшање “глобалне интероперабилности библиотечких података на вебу.” (Baker at al. 2010) Извештај је објављен у октобру 2011. године, а у наредном тексту је преведен и представљен први део овог документа. У њему се наводе предности објављивања метаподатака као LOD, како би се илустровао значај предузимања даљих корака од стране библиотекара.

## 5.1 Предности Linked Data приступа

Овај приступ нуди значајне предности над тренутним праксама за стварање библиотечких метаподатака, а са друге стране представља природни наставак за моделе који се традиционално користе у библиотекама за њихово међусобно дељење. Linked Data (нарочито Linked Open Data -open се у овом контексту посматра са правног аспекта) омогућава дељење, проширивање и поновно коришћење података. Подржава и функционалност мултилингвалности (нпр. једаном URI-ју се може доделити више имена на различитим језицима коришћењем `skos:prefLabel` и `skos:altLabel` својства). Ресурси могу бити описани у колаборацији са другим библиотекама и повезани са подацима које су допринеле друге заједнице, чак и појединци. Слично повезивању које се одвија на вебу међу документима, Linked Data дозвољава било коме да допринесе јединственом експертизом у форми у којој она може бити поново искоришћена и комбинована са експертизом других. Коришћење идентификатора дозвољава различите описе исте ствари. Путем повезивања са комплементарним подацима из веродостојних извора, библиотеке могу побољшати вредност сопствених података. Коришћењем URI-ја за означавање дела, места, људи, догађаја, предмета и других објеката или концепта од интереса, библиотеке ће дозволити да њихови извори буду цитирани од стране великог броја извора података и тиме учинити њихове метаподатке доступнијим и видљивијим. Систем именовања интернет домена (Internet Domain Name System) осигурава стабилност и надзор, тиме што се идентификатори постављају у регулисан и разумљив контекст својине и одржавања. Ово је у складу са дугорочним циљем библиотека, да обезбеде поуздане метаподатке за важне изворе на вебу.

Још једна предност изнова употребљених

јединствених идентификатора јесте могућност каталогизатора да допринесу делове каталогског записа као исказе. У тренутном систему базираном на документима, подаци се размењују у форми целих записа који се преузимају као комплетни описи. Супротно томе, у систему базираном на графу, организација може да допринесе индивидуалном изјавом о извору, а све остале изјаве достављене о том јединствено идентификованом извору могу бити интегрисане у глобални граф. На пример, једна библиотека доприноси предметизацијом, док друга доприноси преводом наслова. У Linked Data окружењу ни један допринос није сувише мали, а један атрибут чини могућим да значајне везе “испливају” из претходно непознатих извора. Библиотечке нормативне датотеке за имена и предмете ће помоћи да се смањи редувантност библиографских описа на вебу, тиме што ће се јасно идентификовати ентитети који су подељени путем Linked Data. Овим ће се потпомоћи и смањење поновљених описа који представљају библиотечки фонд.

## 5.2 Предности за истраживаче, студенте и кориснике

Коришћење Linked Data технологија можда није очигледно за наше кориснике, јер се одиграва у позадини. Ипак корисници могу приметити побољшане могућности за проналажење и коришћење података услед промене саме структуре података, који сада бивају повезани и са другим изворима. Навигација међу библиотечким и неблиотечким информационим изворима постаће софистициранија. Обједињене претраге биће побољшане коришћењем линкова којима се шири индекс, а корисници ће имати веће могућности за прегледање садржаја.

Вредност Linked Data за кориснике библиотека изводи се из основних принципа



навигације, где ће свако моћи да прати успостављене линкове зависно од сопствених интереса. Линкови међу библиотечким и неблиотечким сервисима као што су Википедија, GeoNames, MusicBrainz, BBC, New York Times, повезиваће локалне колекције са ширим универзумом информација на вебу.

Linked Data није стварање друкчијег веба, већ његово унапређење кроз додавање структурираних података. Овим се утиче на функције прикупљања (crawling) и алгоритама релевантности које обављају машине за претрагу и социјалне мреже, уједно омогућавајући побољшану видљивост библиотека кроз њихову оптимизацију. Структурирани подаци уграђени у HTML стране ће олакшати поновно коришћење библиотечких података у сервисима за претрагу информација: управљање референцама може бити једноставно колико и “copy-paste” URI-ја. Аутоматизација преузимања референци из Linked Data или стварање линкова из веб извора ка библиотечким изворима подразумева да ће библиотечки подаци бити у потпуности интегрисани у истраживачке радове и библиографије. Linked Data ће побољшати интердисциплинарно истраживање путем обогаћивања знања кроз линковање међу базама које припадају различитим дисциплинама.

### **5.3 Предности за организације**

Промовисањем овог приступа објављивању података (тзв. bottom-up приступ), Linked Data ствара прилику да се побољша опис извора у библиотекама. Традиционални приступ додељивања библиотечких метаподатака, тј. производња каталожских записа као засебних описа за библиотечки материјал (тзв. top-down приступ), углавном је наметнут буџетским ограничењима; библиотеке немају средства за производњу информација на вишем нивоу грануларности. Са Linked Data, различите

врсте података о истој јединици могу се произвести на децентрализован начин, затим припојити у исти граф. Linked Data технологија може да потпомогне побољшање унутрашњег одржавања података и одржавање линкова међу нпр. дигитализованих објеката и њихових описа. Може да побољша процес објављивања података у организацијама у којима подаци нису у слободном приступу. (W3C Incubator Group 2011)

Linked Data може бити први корак ка приступу заснованом на облаку (cloud-based) за управљање културним информацијама, који би могао бити финансијски повољнији од засебних система који се одржавају у институцијама. Мање библиотеке или засебни пројекти би овим приступом могли да се учине видљивијим и боље повезаним, истовремено смањујући инфраструктурне трошкове.

Коришћењем ових технологија библиотеке повећавају своје присуство на вебу, где се налазе бројни потенцијални корисници информација. Фокус на идентификаторе омогућа састављање описа за специфичне заједнице као што су музеји, архиве, галерије и аудио-видео архиве. Отвореност података представља прилику више него претњу.

Разјашњењем услова лиценцирања дескриптивних метаподатака олакшава се њихово коришћење и побољшава се институционална видљивост. Тако изложени подаци биће искоришћени на неочекиване начине. Као што једна изрека каже: “The coolest thing to do to your data will be thought of by someone else.”

### **5.4 Предности за библиотекарe, архивисте и кустосе**

Коришћење веба и идентификатора базираних на вебу омогућиће ажурне описе јединица доступне за директне референце од стране каталогизатора. Коришћење дељених

идентификатора омогућиће обједињавање описа извора из различитих домена, из свих појединачних база институција које чувају културно наслеђе, чак и са великом већином садржаја на вебу. Каталогизатори ће моћи да се концентришу на домен њихове локалне експертизе, уместо на репродукцију постојећих описа који су већ разрађени од стране других.

Историја нам је показала да су све технологије пролазне, а историја информационе технологије сугерише и то да су појединачни формати метаподатака нарочито кратког века. Linked Data описује значење података (семантику) одвојено од појединих структура података (синтаксе или формата), са резултатом да се то значење задржава и поред промена формата. У том смислу Linked Data је дуготрајнији и робуснији од формата метаподатака, који су зависни од појединих структура података. (W3C Incubator Group 2011)

## 6. Закључак

Иако све ово можда звучи као далека будућност и научна фантастика, заправо многе библиотеке су заиста препознале потенцијал и већ имплементирале ову технологију на својим подацима. Број тих пројеката експоненцијално расте. Дакле, будућност је стигла!

Еуропеана је објавила око 3,5 милиона записа као Linked (Open) Data, међу којима су и неки из doiSerbia сервиса, који пружа Народна библиотека Србије. Подаци из Шведског узајамног каталога LIBRIS, који садржи око 6 милиона библиографских записа из 175 библиотека, објављени су помоћу вокобулара FOAF (Friend of a Friend), SKOS, BIBI и Даблинско језгро. У овом каталогу URI је додељен свакој од јединица, а описи су даље повезани са LOD изворима DBPедје и Предметних

одредница Конгресне библиотеке. Сличне иницијативе дешавају се у већини великих библиотека, нпр. у Конгресној библиотеци, Британској, Француској, Мађарској и многим другим националним библиотекама. Поред горепоменутог VIAF пројекта, неки од вокабулара који су доступни као LOD су: Дјуијева децимална класификација, Предметне одреднице Конгресне библиотеке, Предметне одреднице Француске националне библиотеке итд.

“Студенти, чак и наставно особље радије ће почети своју претрагу за информацијама на Гуглу него у библиотеци” каже Том Бејкер, председник Иницијативе за метаподатке Даблинског језгра (Dublin Core Metadata Initiative, DCMI). Уколико библиотеке желе да задрже своју улогу кустоса друштвене интелектуалне продукције, оно чиме оне располажу мора бити део веба и иницијалне претраге наших корисника. (Kelley 2011)

Иако са овом темом можда није упозната шира библиотечка јавност, изван број професионалаца сматра да је за библиотечку будућност кључно мигрирати податке ван институционалних база у којима су складиштени (популарна метафора за такве базе јесу “издвојени силоси”) у отворену, глобалну средину у којој се подаци деле. Повезивање библиотечких метаподатака са неблиотечким кроз LOD механизам види се као обећавајући пут ка постизању постојаности, доступности и могућности поновног коришћења у веб-базираном, интегрисаном, информационом универзуму. Такође, се сматра да би форматирање и повезивање библиотечких метаподатака RDF-ом учинило да наш садржај чешће избија на површину, међу резултате веб претраге и да би их тиме учинили доступним путем веб протокола као што је HTTP, уместо само оних створених строго за библиотечке потребе, као што је Z39.50. (Kelley 2011)

Из иницијатива које су велике библиотеке широм света већ преузеле у овом смеру, можемо закључити да је овај процес већ отпочео и да би било пожељно да се руководимо тим примерима уколико желимо да будемо у току са захтевима друштва који свакодневно расту. Ипак, да би пружили крајње жељене услуге нашим корисницима, осим иницијативе библиотекара за успостављање оваквог система, неопходна је и ангажованост програмера и инжењера информационог система, који ће на првом месту разумети начине описа јединица у библиотекама и другим културним институцијама, а потом развити апликације за најоптималније коришћење ових важних података.

### Захвалност

Рад је проистекао из стручног рада под насловом „Предности примене технологије „повезивања података“ (Linked Data) на библиотечке метаподатке“, приложеног као део Стручних испита у библиотечно-информационој делатности, јуна 2011. године. Овом приликом се захваљујем Народној библиотеци Србије, која је финансирала моје похађање ових Стручних испита и мр. Биљани Косановић, супервизору оригиналне верзије рада.

### Белешке:

1. <http://dublincore.org/documents/dcmi-terms/>
2. <http://viaf.org/>
3. <http://www.ifla.org/publications/functional-requirements-for-bibliographic-records>
4. <http://europeana.eu/>
5. <http://www.geonames.org/>
6. <http://www.w3.org/>
7. <http://lod-cloud.net/>
8. <http://dbpedia.org/About>

Текст од одељка 5.1-5.4 представља

модификовани превод пасуса 2-2.4 документа: W3C Incubator Group. 2011. Library Linked Data Incubator Group Final Report. <http://www.w3.org/2005/Incubator/lld/XGR-lld-20111025/> (приступљено 24.VII 2013)

### Литература:

Antoniou, Grigoris and Frank Van Harmelen. 2008. *A Semantic Web Primer*. USA: Massachusetts Institute of Technology.

Baker, Tom, Emmanuelle Bermes and Antoine Isaac. 2010. Library Linked Data Incubator Group Charter. W3C incubator activity. <http://www.w3.org/2005/Incubator/lld/charter> (приступљено 20.VII 2013)

Berners-Lee, Tim. 2006. Linked Data. World Wide Web Consortium. <http://www.w3.org/DesignIssues/LinkedData.html> (приступљено 20.VII 2013)

Chan, Lois Mai and Marcia Lei Zeng. 2006. Metadata Interoperability and Standardization – A Study of Methodology, Part I. *D-Lib Magazine*, 12(6): стр. 4. <http://www.dlib.org/dlib/june06/chan/06chan.html#CC-DA> (приступљено 20.VII 2013)

Cleveland, Donald B. and Ana D. Cleveland. 2001. *Introduction to Indexing and Abstracting*. Englewood, Colorado: Greenwood Press

Doerr, Martin, Stefan Gradmann, Steffen Henniecke, Antoine Isaac, Carlo Meghini and Herbert van de Sompel. 2010. The Europeana Data Model (EDM). Рад представљен на конференцији IFLA 149, Information Technology, Cataloguing, Classification and Indexing with Knowledge Management. <http://www.ifla.org/files/hq/papers/ifla76/149-doerr-en.pdf> (приступљено 20.VII 2013)

Fox, Michael J. 2001. Stargazing: Locating EAD in the Descriptive Firmament. *Journal of Internet Cataloging* 4(3): 61-74.

Gomez-Perez, Asuncion, Mariano. Fernandez-Lopez and Oscar Corco. 2003. *Ontological Engineering*. Springer-Verlag

Gruber, Thomas Robert. 1993. Toward principles for

the design of ontologies used for knowledge sharing. *International Journal of Human-Computer Studies* 43 (5-6): 907-928.

Hoge, Gail. 2000. *Systems of Knowledge Organization for Digital Libraries: Beyond Traditional Authority Files*. Washington, DC: the Council on Library and Information Resources <http://www.sims.monash.edu/subjects/ims2603/resources/Assignment2Papers/SKOforDigLib.pdf>.

Jurić, Andrija. 2011. *Linked Data u ekonomskim aplikacijama*, diplomski rad. Ekonomski fakultet u Osijeku.

Kelley, Michael. 2011. How the W3C Has Come to Love Library Linked Data. *Library Journal*. [http://www.libraryjournal.com/lj/home/891826-264/how\\_the\\_w3c\\_has\\_come.html.csp](http://www.libraryjournal.com/lj/home/891826-264/how_the_w3c_has_come.html.csp) (приступљено 15.7.2012)

McCallum, Sally H. 2002. MARC: Keystone for Library Automation. *Annals of the History of Computing, IEEE* 24 (2): 34-49.

Miller, Eric. 2001. Layer Cake. World Wide Web Consortium. <http://www.w3.org/2001/09/06-ecdl/slide17-0.html> (приступљено 24.VII 2013)

Neches, Robert, Richard E. Fikes, Tim Finin, Thomas Gruber, Ramesh Patil, Ted Senator and William Swartout 1991. Enabling technology for knowledge sharing. *AI Magazine* 12(3): 16-36.

Olensky, Marlies. 2010. Semantic interoperability in Europeana: An examination of CIDOC in digital cultural heritage documentation. *Bulletin of IEEE Technical Committee on Digital Libraries* 6(2). <http://www.ieee-tcdl.org/Bulletin/v6n2/Olensky/olensky.html> (приступљено 24.VII 2013)

RDF Primer. 2004. World Wide Web Consortium. <http://www.w3.org/TR/rdf-primer/> (приступљено 24.VII 2013)

Republički zavod za statistiku. Semantički veb. [http://www.google.rs/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CCYQFjAA&url=http%3A%2F%2Fwebrzs.stat.gov.rs%2FWebSite%2FuserFiles%2Ffile%2FAktuelnosti%](http://www.google.rs/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CCYQFjAA&url=http%3A%2F%2Fwebrzs.stat.gov.rs%2FWebSite%2FuserFiles%2Ffile%2FAktuelnosti%2F)

[2FSemanticki%2520veb\\_cyr.doc&ei=5SK-T-3VM5DE4gTTmYw\\_&usg=AFQjCNGj-wZ\\_Q42Q2\\_bfCFSTkbTtmK541A&sig2=jblacBN1GrZ1EclKo\\_pgJw](http://www.google.rs/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CCYQFjAA&url=http%3A%2F%2Fwebrzs.stat.gov.rs%2FWebSite%2FuserFiles%2Ffile%2FAktuelnosti%2F) (приступљено 24.VII 2013)

Shadbolt, Nigel, Tim Berners-Lee and Wendy Hall. 2006. The Semantic Web Revisited. *IEEE Intelligent Systems*. 21(3): 96-101.

Studer, Rudi, Benjamins, Richards V. and Fensel, Dieter. 1998. Knowledge Engineering: Principles and methods. *Data & Knowledge Engineering* 25 (1-2)

Fensel, Dieter, James Hendler, Henry Lieberman and Wolfgang Wahlster. 2005. *Spinning the Semantic Web: bringing the World Wide Web to its full potential*. Cambridge: MIT Press

Weller, Katrin. 2010. *Knowledge representation in the social semantic Web*. New York; Berlin: De Gruyter Saur. стр. 53

W3C Incubator Group. 2011. Library Linked Data Incubator Group Final Report. World Wide Web Consortium. <http://www.w3.org/2005/Incubator/lld/XGR-lld-20111025/> (приступљено 24.VII 2013)

Примљено: 13. III 2013.

Прихваћено: 25. VI 2013.