

КАДА НАЦИОНАЛНЕ БИБЛИОТЕКЕ ОТВАРАЈУ ПОДАТКЕ...

Тамара Бушићан Вучај
Народна библиотека Србије,
Београд

УДК: 027.54:004(4-672EU)
004.738.5:02

Сажетак

Покрет отварања података актуелан је у Европи већ низ година, а у Србији од 2016. Библиотеке, као значајни актери у производњи и управљању подацима, заинтересоване су за отворене податке, а још више за повезане отворене податке. Да ли су националне библиотеке пронашле своје место у процесу отварања података? Који су бенефити за њих саме, а који за најшире корисничке групе? Које сетове података су националне библиотеке већ отвориле? Одговоре на ова питања потражићемо у доброј пракси националних библиотека земаља Европске уније, са тенденцијом да постојеће добре праксе применимо у Србији.

Кључне речи: отворени подаци, повезани отворени подаци, националне библиотеке, сетови података.

Производња и управљање подацима дубоко су утемељени у библиотекарству. Националне библиотеке, које би требало да воде све остале и планирају трасе пута за развој библиотекарства, имају за императив да прате и препознају нове тенденције које се тичу подручја података. Трѐвало би да су већ препознале *повезане погашке* (linked data) и *повезане отворене погашке* (linked open data). Библиографски подаци доступни путем онлајн каталога под неком од отворених лиценци, базични су пример отворених података, односно њима би Тим Бернерс Ли (Tim Berners Lee), човек коме се приписује проналазак веба, доделио једну звездицу¹ за отварање података. Пет звездица доделио би повезаним отвореним подацима. У овом раду ћемо инспирацију за отварање података потражити управо међу националним библиотекама са пет звездица.

У Србији је посебна пажња посвећена отвореним подацима 2016. године, када је Дирекција за електронску управу покренула национални портал отворених података – <https://data.gov.rs/sr/> и формирала Радну групу за отворене податке. Народна библиотека Србије део је националне Радне групе за отворене податке од 2017. године. Од друге половине 2017. године о порталу и радној групи брине владина Канцеларија за информационе технологије и електронску управу, <http://www.deu.gov.rs/>.

Покрет отварања података веома је активан у земљама Европе, преваходно у чланицама Европске уније (ЕУ), чије законодавство уређује ову област. Отворени подаци и њихова расположивост за поновну употребу носе потенцијал за бизнис и профит, али су и изузетан ресурс за грађане јер омогућавају увид у систематизоване податке из различитих области. Легислатива ЕУ прописује обавезност достављања података за поновну употребу, чиме су од 2013. године обухваћене и библиотеке, музеји и архиви. Многе националне библиотеке земаља чланица ЕУ отвориле су своје податке и притом показале висок степен креативности у одабиру и приказу података. Елизабет Нигеман (Elisabeth Niggemann), дугогодишња директорка Немачке националне библиотеке, у свом раду изложеном на Ифлиној конференцији у Финској, 2012. године, објашњава значај отварања података за националне библиотеке кроз два бенефита. Први је да објављивањем и повезивањем података на вебу, библиотеке граде семантички веб и на тај начин обогаћују своје податке и унапређују сервисе за своје кориснике. Други, можда и значајнији бенефит је у томе што библиотеке отварањем података досежу до много већег броја корисника путем социјалних мрежа, портала и претраживача, скрећу пажњу на себе и тиме доказују своју релевантност у универзуму информација. Жилда Илијен (Gildas Illien), директор за библиографске и дигиталне информације Националне библиотеке Француске,² у свом раду изложе-

¹ Портал: 5 звездица отворених података (5 star open data), <http://5stardata.info/en/>.

² Жилда Илијен је од 2016. године на челу библиотека Националног природњачког музеја у Паризу.

ном на истој Ифлиној конференцији, у новој ери веба, у вебу података, препознаје значајну улогу библиотекара и библиотека, као произвођача квалитетних структурираних података. Он види, пре свега, националне библиотеке као лидере у коришћењу RDF тројки³, примени FRBR-а⁴ и учешћу у семантичком вебу. Он наводи: „Укратко, због своје публице, ауторитета, одговорности и репутације, националне библиотеке више од других утичу на јавно мњење, јавну политику и најбоље праксе и позваније су да преузму улогу у заговарању за отварање података“. Парадигму отварања података у традицији склоним библиотекама, Жилда пореди са преласком са националне валуте на евро. Отварање података донело је Националној библиотеци Француске генерално већу видљивост и немало интересовање медија. Наједном, сетови метаподатака су били поређени и мерени са подацима из јавне администрације, јавног здравља или саобраћаја, што је, на неки начин, променило и слику библиотечких каталога, приказујући их као потенцијално корисне за грађане.

Кренућемо на пут кроз отворене библиотечке податке земаља ЕУ, полазећи од најближе Словеније, преко Француске, Велике Британије и других земаља, како бисмо потражили примере добре праксе за отварање података из библиотека у Републици Србији, земљи кандидованој за чланство у ЕУ.

Портал отворених података Републике Словеније, <https://podatki.gov.si/>, у области образовања, културе и спорта, има више десетина сектора података, од тога, највише из државне управе, тј. Министарства за образовање, науку и спорт и Министарства за културу. Национална и универзитетска библиотека Словеније отворила је податке са портала Словеначка библиографија, <http://sb.nuk.uni-lj.si/>. Портал окупља библиографске записе о књигама, часописима, чланцима у часописима, музикалијама, звучним записима, картографској и сликовној грађи, у неколико стандардних формата, CSV, XLS, XML⁵, као и слике у форматима PNG, JPG, GIF⁶. Такође, Дигитална библиотека Словеније, <https://dlib.si/>, отворила је своје метаподатке, који се могу преузети са странице Националног агрегатора е-садржаја, www.aggregator.si. Словенија иначе заузима високо 7. место

3 RDF (Resource Description Framework) – Формат за изражавање семантике података. Информације о било ком извору дају се у формату уређене тројке: субјекат–предикат–објекат.

4 FRBR (Functional Requirements for Bibliographic Records – Функционални захтеви за библиографске записе), концептуални модел који представља библиографске ентитете и везе између њих, настао у циљу задовољавања потреба корисника при коришћењу библиографских записа у односу на најразличитије носиоце информација.

5 CSV (Comma-separated values), XLS (eXcel Spreadsheet) – формати за табеларни приказ података; XML (Extended Markup Language) – формат података истовремено читљив за људе и машине.

6 PNG (Portable Network Graphics), JPG (Joint Photographic Expert Group), GIF (Graphics Interchange Format) – формати слика који се најчешће користе на вебу.

међу државама ЕУ у области отворених података из јавног сектора, према извештају Европске комисије из новембра 2011. године⁷.

Портал отворених података Републике Француске, <http://www.data.gouv.fr/fr/>, у области културе нуди читав низ веома различитих сетова отворених података, које доставља Министарство културе и комуникација, као и јавне установе на државном и локалном нивоу. Национална библиотека Француске има седам сетова отворених података на овом порталу, <http://www.data.gouv.fr/fr/organizations/bibliotheque-nationale-de-france/>. На порталу је доступна статистика о броју поновних употреба за сваки од сетова, период ажурирања података, као и подаци о корисничким групама заинтересованим за поједине сетове. Сетови отворених података су следећи:

1. Тематске веб-збирке

У оквиру своје мисије, Национална библиотека Француске води бригу о обавезном примерку интернета, тако да редовно прикупља узорак француског веба, како неселективни годишњи, тако и циљани узорак, који садржи две врсте збирки: пројектне (у вези са актуелним догађајима) и текуће (референтни веб-извори по дисциплинама). Свака URL⁸ адреса у оквиру овог сета, садржи описне и техничке метаподатке, као што су кључне речи, тема, дубина, фреквенција роботског харвестовања и историја адресе. Овај сет садржи 57 табела у CSV формату, које садрже 57.538 URL адреса⁹, прикупљаних од 2011. године.

2. Веб-збирке о изборима

Национална библиотека Француске прикупља узорак француског веба из домена политике. Од 2002. године, током сваког гласања, сакупљају се веб странице политичких странака и кандидата на изборима, блогови, прилози у онлајн медијима и активности на социјалним мрежама. Циљ ове збирке јесте да документује развој и диверзитет политичког живота, као и онлајн дебате грађана. Партнери у овом пројекту су регионалне библиотеке које су укључене у прикупљање обавезног примерка. И у овом кумулативном сету података могу се наћи URL адресе, заједно са описним и техничким метаподацима од 2002. до 2017. године.

7 Open Data Maturity Report 2017, <https://www.europeandataportal.eu/en/highlights/open-data-maturity-europe-2017>.

8 URL (Uniform Resource Locator) – Јединствени локатор извора.

9 Сви подаци са портала отворених података преузети су у фебруару 2018. године.

3. Евиденција о обавезном примерку

Законском обавезом прикупљања обавезног примерка, Национална библиотека Француске у прилици је да отвори скупове података о обавезном примерку за све врсте грађе, од веб-страница до карата и фотографија. Ови сетови података садрже статистичке податке из домена јавног издаваштва: о издавачима, штампаријама, медијима и дисциплинама. Сет садржи 36 докумената од 2011. до 2015. године.

4. Портал data.bnf.fr¹⁰

На посебним веб-страницама посвећеним повезаним отвореним подацима, Национална библиотека Француске пружа приступ систематизованим подацима о ауторима, делима и предметним одредницама. За једног аутора, приказују се подаци о свим делима, сва издања једног дела, као и онлајн приступ документу на Галици (*Gallica*)¹¹, уколико је већ дигитализован. Овим путем се промовишу класици француске књижевности, историје и права. У овом пројекту користе се технике семантичког веба, који омогућава лакше и интуитивније претраживање различитих врста извора. Коришћење формата RDF, структурираног за веб, олакшава индексирање, дељење и поновну употребу података.

5. Комплетни подаци из каталога

Од 2014. године, Национална библиотека Француске донела је одлуку да отвори библиографске податке из својих каталога под отвореном лиценцом, која омогућава поновну употребу свих података под условом да се наведе извор. У оквиру овог сета података доступни су следећи документи: библиографски записи у MARC формату и формату DC¹² из општег каталога, записи о дигитализованој грађи у формату DC из општег каталога и каталога архива и рукописа, као и подаци у RDF формату са портала data.bnf.fr и reliures.bnf.fr.

6. База дигитализованих повеза

Национална библиотека Француске поседује једну од најдрагоценијих збирки повеза на свету. На посебним веб-страницама, www.reliures.bnf.fr, налази се избор дигитализованих повеза, за-

¹⁰ Пројекат отворених података Националне библиотеке Француске, <http://data.bnf.fr/> заједно са пројектом дигиталне библиотеке <http://gallica.bnf.fr/>, награђен је 2013. године Наградом за иновације у истраживачким библиотекама, коју додељује Универзитет Стенфорд.

¹¹ Gallica, www.gallica.fr, Дигитална библиотека Француске.

¹² DC (Dublin Core) – Даблинско језгро, <http://dublincore.org/> – један од најчешће коришћених стандарда за метаподатке у библиотекама.

једно са детаљним описима који садрже податке о радионицама, власницима, као и библиографије извора. Подаци о око 200 повеза доступни су у RDF формату на самом порталу.

7. Библиографија париских издања из 16. века

На посебним страницама, <http://bp16.bnf.fr>, Национална библиотека Француске представља податке о издавачкој продукцији града Париза током 16. века. Само за прву половину 16. века, постоји 14.200 записа. На порталу су за преузимање доступни подаци у RDF формату.

Статистика о поновној употреби сетова података у први план ставља податке са портала <http://data.bnf.fr/>. Овде су забележена два случаја поновне употребе и то:

1. Апликација <http://www.human-music.eu/> коју је израдила Федерација специјализованих вокалних и инструменталних састава. Ова асоцијација за циљ има демократизацију и демистификацију света музике. На основу података о делима и композиторима са портала data.bnf.fr, креирана је база података звучних и видео записа, документарних филмова и снимака концерата чланова Федерације. Ова дигитална платформа за музику својеврсна је друштвена мрежа у којој учествују и бивају вредновани сами музичари.
2. Калкулатор јавног домена, креиран од стране Фондације отвореног знања Француске, користи податке са портала data.bnf.fr како би аутоматски одредио да ли је неко дело заштићено ауторским правима или је у јавном домену. Калкулатор се налази на адреси calculateurdomainepublic.fr, а више информација се може добити на видеу који потписује Министарство културе и информација <https://vimeo.com/78740619>.

Велика Британија, одлазећа чланица ЕУ, једна је од најнапреднијих земаља у области отварања података. На националном порталу <https://data.gov.uk/> долазимо до податка да је Британска библиотека заступљена са четири сета података, <https://data.gov.uk/data/search?publisher=the-british-library&q=british+library>. За разлику од Француске, Велика Британија показује много прагматичнији однос према отварању података, као и настојање британске владе да њен рад и финансије буду транспарентни. Тако Британска библиотека кроз три сета података пружа податке о утрошеним буџетским средствима, детаље уговора који прелазе одређену суму новца, даје организациону шему са приходима свог менаџмента и запослених. Четврти сет података је Британска национална библиографија са отвореним и повезаним подацима. Овај сет је део националне информационе инфраструктуре. Скуп од 3.9 милиона записа бележи издавач-

ку активност у Уједињеном Краљевству и Републици Ирској од 1950. године. Поред података о штампаним публикацијама однедавно су додати подаци о електронским публикацијама. Региструју се књиге, оне које су објављене и које ће се тек објавити, као и серијске публикације. Посебне веб-странице Британске библиотеке посвећене су овим подацима, <http://bnb.data.bl.uk/>. Такође, Британска библиотека има и странице посвећене отвореним подацима, <http://www.bl.uk/bibliographic/datafree.html>. На овим страницама постоји читав низ нових сетова података, а из уводног документа сазнајемо да се политика отварања података Британске библиотеке шири и ка небiblioteчким заједницама, и то: ка генералној публици, академским истраживањима, заједници архива, издавачкој индустрији Уједињеног Краљевства, заједници програмера и ка школама. За различите заједнице резервисани су различити стандарди метаподатака, методе за приступ и испоруку података, док је подразумевана лиценца СС0¹³. Књиге које долазе – најновији је сет података, а ради се о књигама које су прошле каталогизацију у публикацији, а сређене су тематски.¹⁴ Посебно је занимљив приступ Британске библиотеке отварању података за истраживаче. Наиме, постоји могућност да се креирају сетови података у CSV формату, како би истраживачи могли да их анализирају посебним алатима као што је Open Refine¹⁵. Сваки сет садржи пет докумената од којих сваки пружа другачији приступ подацима са фокусом на: библиографски запис, наслов, ауторство, предметну одредницу или Дјуијеву класификацију. Интересантно је да су ови документи креирани или на захтев истраживача или да би били објављивани уз изложбе Британске библиотеке. Илустративан је пример сета података о панку, <http://www.bl.uk/bibliographic/download.html#csvpunk>, који обухвата податке о делима објављеним од 1976. године, из националне библиографије и из збирки штампане и рукописне грађе. Овај сет је креиран уз изложбу Британске библиотеке „Punk 1976–78”, која је била отворена од 13. маја до 2. октобра 2016. године, <https://www.bl.uk/events/punk-1976-78>, у част четрдесетог рођендана овог јединственог музичког феномена. Иначе, сви подаци Британске библиотеке доступни су корисницима преко протокола Z39.50 у MARC21 формату, за коришћење у некомерцијалне сврхе и уз претходну регистрацију за овај сервис.

У наставку наводимо још неколико примера добре праксе националних библиотека у отварању података. На порталу отворених података

¹³ СС0 (Creative Commons Zero) – Кријетив Комонс нула, опредељено за јавни домен.

¹⁴ Народна библиотека Србије је отворила податке о публикацијама са СИП-а још 2006. године, објављиване по месецима, а сређене по УДК класификацији, https://www.nb.rs/info_publishers/cip.php (прим. аут.).

¹⁵ Open Refine је бесплатан, моћан алат за рад са неуређеним подацима, <http://openrefine.org/>.

Немачке, налазимо да је Немачка дигитална библиотека¹⁶, са седиштем у Националној библиотеци Немачке у Франкфурту, отворила своје метаподатке, <https://www.govdata.de/web/guest/suchen/-/details/deutsche-digitale-bibliothek>, а поред метаподатака, отворени су и подаци у вези са апликационим програмским интерфејсом¹⁷, користећи податке из дигиталне библиотеке. Такође, сама Немачка национална библиотека има свој портал повезаних отворених података, доступан на адреси: <http://www.dnb.de/DE/Service/DigitaleDienste/LinkedData/linkedata.html>.

На националном порталу отворених података Шпаније, Национална библиотека Шпаније има 12 сетова података, углавном се ради о библиографијама за поједине врсте грађе, као и за дигиталне збирке, http://datos.gob.es/en/catalogo?q=biblioteca+nacional+de+espana&theme_id=cultura-ocio&sort=metadata_modified+desc. Национална библиотека Шпаније има посебне странице посвећене отвореним библиографским подацима на адреси: <http://datos.bne.es/inicio.html>.

Библиотеке у Италији отвориле су 38 сетова података, према националном порталу отворених података Италије. Углавном се ради о регионалним библиотекама које пружају увид у своје библиографске записе или дигиталне збирке. Краљевска и уједно национална библиотека Холандије отворила је четири сета података, https://data.overheid.nl/data/dataset?q=Koninklijke+Bibliotheek&theme_facet=http%3A%2F%2Fstandaarden.overheid.nl%2Fowms%2Fterms%2FCultuur_en_recreatie&sort=score+desc%2C+modified+desc%2C+metadata_modified+desc, и то у вези са рано штампаним књигама, инкунабулама и рукописном грађом.

Европеана, европски портал за приступ дигитализованом културном наслеђу читаве Европе, својом политиком отварања података много је учинила и за националне библиотеке, своје партнере, да отварању података дају приоритет у својим агендама. Метаподаци за све објекте на Европеани су отворени, и за све важи лиценца за јавни домен (CC0), дакле сви подаци се могу преузети преко API-ја, могу се употребити и поново употребити у различите сврхе. На отварање метаподатака и додељивање лиценце јавног домена све партнере Европеане обавезује Уговор о размени података.¹⁸ Овај уговор је, 2015. године, потписала и Народна библиотека Србије, као партнер и агрегатор метаподатака за Европеану.

Отварање података са више од једне звездице, према класификацији Тима Бернерса Лија, домаћи је задатак за националне библиотеке. Такође и активније укључивање у семантички веб путем повезаних отворених података. Народној библиотеци Србије потребна је помоћ за ове задат-

¹⁶ Deutsche Digitale Bibliothek, <https://www.deutsche-digitale-bibliothek.de/>.

¹⁷ API (Application Programming Interface).

¹⁸ DEA (Data Exchange Agreement), <https://pro.europeana.eu/page/the-data-exchange-agreement>.

ке. Када освестимо проблем, на добром смо путу и да га решимо. Нисмо сами, на једној страни имамо породицу европских националних библиотека, која кроз своју организацију CENL19 има одличан канал за сарадњу свих европских националних библиотека. На другој страни имамо решеност државе Србије и велику подршку да се подаци отворе. Ако добро повежемо ове две стране, наши задаци биће написани.

19 CENL (Conference of European National Librarians) – Конференција директора националних библиотека Европе, <https://www.cenl.org/>.

ЛИТЕРАТУРА

1. Дакић, Наташа, Андоновски, Јелена. Linked Open Data и Bibframe: нови облици организовања метаподатака у библиотекарству, www.unilib.rs/wp-content/uploads/2016/03/Metapodaci-seminar_4.pptx (27. фебруар 2018).
2. Illien, Gildas. "Are you ready to dive in? A case for Open Data in national libraries". Рад представљен на Ифлиној конференцији у Хелсинкију, Финска, 15. августа 2012. <https://www.ifla.org/past-wlic/2012/181-illien-en.pdf> (27. фебруар 2018).
3. Niggemann, Elisabeth. "The importance of open data to national libraries". Рад представљен на Ифлиној конференцији у Хелсинкију, Финска, 15. августа 2012. <https://www.ifla.org/past-wlic/2012/181-niggemann-en.pdf> (27. фебруар 2018).

Tamara Butigan Vučaj
National library of Serbia,
Belgrade

WHEN NATIONAL LIBRARIES OPEN THEIR DATA...

Summary

The open data movement has been present in Europe for many years, while in Serbia it is the case since 2016. Libraries, as important factors in data production and administration, are interested in the concept of open data, especially when linked open data are in question. Have national libraries found their places in the process of opening data? What are the benefits for themselves, and what for the broadest user groups? Which sets of data have already been opened by national libraries? We will try to find the answers to these questions in a good practice of the national libraries from the European Union's countries, with the tendency to apply that kind of practice in Serbia.

Kew words: open data, linked open data, national libraries, datasets.