

Александра Поповић
Универзитетска библиотека
„Светозар Марковић“
Милица Шевкушић
Институт техничких наука САНУ
Ђорђе Стакић
Математички факултет Универзитета у Београду

UDC: 02:004.738.5
DOI: 10.18485/dh.2015.1.ch12

БИБЛИОТЕКЕ И ВИКИПЕДИЈА ЗАЈЕДНО НА ВЕБУ: СЛОБОДНО ЗНАЊЕ ЗА СВЕ

Сажетак

Дигитализација фондова и постављање садржаја у отворени приступ данас је приоритет библиотека, поборника отворене науке и слободног приступа садржаја на интернету.

Универзитетска библиотека „Светозар Марковић“ заједно са Математичким факултетом Универзитета у Београду и Викимедијом Србије у току 2015. године реализује пројекат *Вики-библиотекар*. Идеја пројекта је да се постави што више квалитетних садржаја на Википедију, најважнији пројекат Викимедија фондације (Wikimedia). Викимедија Србије је настала још давне 2005. године као пети огранак Викимедије. То је невладино, непрофитно и нестраначко удружење са основном идејом да сви људи имају једнак приступ знању и образовању. На пројекту учествује и Институт техничких наука САНУ као значајан партнер у реализацији задатих циљева.

Библиотеке поседују непроцењива блага у својим фондовима која нису свима на располагању. За дела којима су истекла ауторска права или она чији аутори желе да их поставе у слободан приступ, постављање на Викимедијину оставу, Вики-изворник или Вики-књиге, у зависности од тога којој категорији дела припадају, представља прави начин промовисања. Библиотекари су као заједница добро повезани, на радним местима им је доступна референсна литература, вешти су у проналажењу валидних и проверљивих информација, познају богате фондове и збирке које се налазе у њиховим установама, едукују кориснике различитих нивоа узраста и образовања и прави су избор за квалитетно уређивање Википедије и других Вики-пројеката.

У раду ће бити речи о методама и механизмима уз помоћ којих се садржаји унутар Вики-пројеката, а нарочито дигитализовани садржаји на Викимедијиној остави, Вики-изворнику и Вики-књигама, могу учинити видљивим и лако налажљивим. Те методе, унутар Википедије, обухватају, пре свега: додељивање категорија, повезивање са садржајима на Википедији на другим језицима,

успостављање међусобних веза унутар пуног текста чланака, повезивање садржаја из Викимедијине оставе, Вики-изворника и Вики-књига са чланцима у Википедији, а за садржаје на Викимедијиној остави и коришћење детаљних и прецизних описа на више језика. Биће речи и о могућностима повезивања садржаја са Вики-пројеката и онлајн каталога, дигиталних библиотека и репозиторијума и сајтова библиотека. Индексирање садржаја на Википедији донекле се разликује од индексирања унутар конвенционалних библиотечких каталога, а логика којом се руководи сличнија је индексирању унутар дигиталних репозиторијума. Ослањајући се на своја стручна знања, библиотекари могу знатно да допринесу бољем структурисању података и квалитетнијем индексирању садржаја унутар Вики-пројеката.

Кључне речи: библиотеке, Википедија, Викимедија Србије, дигитализација, отворени приступ

Увод

Између библиотека и Википедије, односно пројеката Фондације Викимедије (Wikimedia), постоји природна веза, а њихова мисија је у суштини иста: теже структурисаном знању које је јавно доступно. Иако се индексирање садржаја на Википедији донекле разликује од индексирања унутар конвенционалних библиотечких каталога, а логика којом се руководи сличнија је индексирању унутар дигиталних репозиторијума, систем категорија на ком почива организација материјала у оквиру Вики-пројеката лако се може упоредити са системима класификације који се користе у библиотекарству (Voss). Библиотеке располажу провереним и структурираним информацијама, као и богатом грађом (књижном, фотографском, картографском, звучним записима итд.) која се, када су дела прешла у јавни домен или када носиоци ауторских права то желе, може ставити на увид јавности. У том контексту, Вики-пројекти представљају функционалну софтверску платформу која омогућава дисеминацију информација и грађе и обезбеђује њихову видљивост на глобалној мрежи, те тако помаже библиотекама у заступању јавног интереса и обављању релевантних друштвених функција.

Правци сарадње између библиотека и Вики-пројеката

Сарадња између библиотека, широм света, и Викимедија

фондације до сада се развијала у два главна правца. Најшира и најдинамичнија сарадња остварена је у оквиру тзв. ГЛАМ (GLAM) пројеката, односно пројеката сарадње са галеријама, библиотекама, архивима и музејима (“Wikipedia Loves Libraries: Outreach Wiki”). Два основна сегмента ових пројеката су: похрањивање дигитализованих садржаја из фондова библиотека на Викимедијину оставу и писање и уређивање текстова који су у вези са библиотекама у којима се пројекат реализује и њиховим збиркама (нпр. “Wikipedia”). Ипак, ГЛАМ пројекти нису нужно ограничени на такве акције, па њихов обим и садржај често зависе од интересовања и капацитета локалних носилаца. Тако су, на пример, у оквиру пројекта који се реализује у сарадњи са италијанским библиотекама успостављене везе између појмова у Новом предметном тезаурусу (*Nuovo soggettario*)¹ на сајту Централне народне библиотеке у Фиренци (Biblioteca Nazionale Centrale di Firenze) и Википедије (Lucarelli). Други важан правац сарадње представљају пројекти везани за јавне доступне податке (Linked Open Data), семантички веб, аутоматску анализу текста и употребу вештачке интелигенције. Они се базирају на екстракцији структурисаних података из Википедије и њихову интеграцију у библиотечке каталоге или репозиторијуме. То су, на пример, *DBpedia* (Voss), *Spotlight*, *YAGO*, *AIDA*, *WordNet* и *Cycontology* (Pohl; Le Vœuf). Значајну улогу у таквим пројектима играју Вики-подаци, односно централни репозиторијум структурисаних података из свих Вики-пројеката, који је заживео крајем 2012. и почетком 2013. године (Müller-Birn et al.). Википедија и други Вики-пројекти представљају погодан извор зато што омогућавају потпуно отворен приступ свим садржајима и структурисаним подацима. Поред тога, Вики-пројекти су изузетно богат извор података на више језика проверен од стране Вики-заједница, у којима се ауторитет успоставља одоздо нагоре.

Викимедија Србије и библиотеке

Сарадња између библиотека у Србији и Вики-пројеката

1 Доступан на следећој адреси: <http://thes.bncf.firenze.sbn.it/>

успостављена је и за сада се реализује у оквиру ГЛАМ иницијативе. Током 2015. године, Универзитетска библиотека „Светозар Марковић“, заједно са својим партнерима реализује пројекат *Вики-библиотекар*. У реализацији пројекта учествује и Институт техничких наука САНУ. Иако у основи садржи све главне карактеристике ГЛАМ пројеката (уређивање текстова и депоновање дигитализованих садржаја из библиотечких фондова), приликом дефинисања пројекта *Вики-библиотекар* водило се рачуна о локалном контексту, односно о обиму и карактеру Википедије на српском језику и карактеристикама локалне Вики-заједнице, па су у складу са тиме дефинисани циљеви и приоритети пројекта.

Википедија на српском језику

Википедија на српском језику основана је 2003. године: у фебруару је постављена само насловна страна, а у септембру су објављени и први чланци. Од тада, број чланака константно расте, а запажа се и да се сваки чланак временом развија и добија на квалитету (Табела 1).

Табела 1. Број чланака на Википедији Србије у различитим годинама (септембар 2015)

| Година | Број чланака (у хиљадама) |
|--------------|---------------------------|
| 2004. | 5 |
| 2009. | 100 |
| 2013. | 200 |
| 2015. | 328 |

Носилац Вики-пројеката у Србији је Викимедија Србије, која на годишњем нивоу расписује конкурс за доделу микрогрантова најбољим идејама за обогаћивање Википедије на српском.

Википедија на српском језику је у септембру 2015. имала нешто преко 167 хиљада регистрованих корисника. Од тога је тек око 670 корисника било активно, при чему се под активним корисницима подразумевају они који су извршили бар једну измену током последњих 30 дана. Свака измена се прати и о томе се брину патролери и администратори. У септембру 2015. године, Википедија на српском језику

имала је 53 патролера и 22 администратора (“Статистике”).

Табела 2. Преглед ранга Википедија на јужнословенским језицима – број чланака, администратора и активних уредника (септембар 2015)

| Ранг | Језик | Број чланака (у хиљадама) | Број администратора | Број активних уредника |
|------|-----------------|------------------------------|---------------------|------------------------|
| 19 | српско-хрватски | 427 | 10 | 260 |
| 28 | српски | 328 | 22 | 669 |
| 36 | бугарски | 207 | 27 | 772 |
| 41 | хрватски | 157 | 27 | 486 |
| 42 | словеначки | 148 | 25 | 324 |
| 58 | македонски | 83 | 15 | 171 |

Из Табеле 2 се види да је Вики-заједница која уређује Википедију на српском језику релативно мала, а нарочито је мали број патролера и администратора, који проверавају техничке стандарде и квалитет написаних чланака и депонованих материјала. Добар део уредника Википедије на српском језику чине средњошколци и студенти који учествују у пројектима Вики-гимназијалац (“Википедија:Вики гимназијалац”) и Семинарски радови (“Википедија:Семинарски радови”). Имајући у виду да се ради о младим људима који још увек немају искуства у коришћењу научног апарата и структурисању енциклопедијских јединица, често се дешава да у тексту није дата дефиниција појма који се обрађује, да нису наведени валидни извори података или референце, да је одредница настала неселективним и неовлашћеним преузимањем текста из неке штампане публикације или електронског извора. Резултат таквог приступа најчешће је текст дугачког и недовољно омеђеног наслова који више личи на есеј него на енциклопедијску јединицу. Такве текстове је понекад тешко повезати са другим чланцима, као и са садржајима Википедија на другим језицима и у оквиру других Вики-пројеката.

Поред тога, често се дешава да уредници забораве да напишу текст сврстају у одговарајуће категорије и повежу са чланцима на другим језицима, па је такве одреднице тешко пронаћи.

Један део чланака, а углавном су то одреднице о насељеним местима, унет је помоћу програмираних ботова, односно масов-

ним уносима. Таквим текстовима су неопходни додатна обрада и проширивање, а нарочито усклађивање са језичким стандардима. Исто важи и за текстове преведене са Википедија на другим језицима.

Пројекат Вики-библиотекар

Пројекат *Вики-библиотекар* фокусиран је на обогаћивање Википедије, као најважнијег пројекта Викимедија фондације, са што више квалитетних садржаја. Посебна пажња је посвећена Википедији на српском језику, а учесници пројекта нису се ограничили на теме које имају везе само са њиховим матичним кућама и њиховим фондовима, иако је и овим садржајима посвећена значајна пажња, нарочито у случају сарадника из библиотека у унутрашњости Србије. У овој фази, главни циљ је био помоћи библиотекарима да савладају технику рада у Вики-тексту, али и подстаћи их да се активно укључе у уређивање Википедије и да то знање пренесу и својим корисницима.

Током реализације пројекта написано је или знатно допуњено око три стотине текстова и депонован је велики број фотографија на Викимедијину оставу. Запажено је да је, како је реализација пројекта одмицала, квалитет текстова који су уређивани бивао све бољи, како у погледу садржаја, тако и у погледу техничких стандарда. Поред тога, дигитализоване су три публикације за Вики-изворник. Приликом обраде публикација у програмима за препознавање текста наишло се на много проблема. Изабране публикације су из 1865. (Аћим Медовић, *Судска медицина, део први или припремни*) из 1871. (Аћим Медовић, *Судска медицина, део други*) и из 1869. године (Аћим Медовић, *Мале хирургијске услуге*). Програми за препознавање текста имају потешкоће да правилно препознају ћириличне знакове нашег тадашњег писма, тако да препознати текст мора ручно да се поправља. То изискује додатно ангажовање запослених у Библиотеци или волонтера.

Интеграција садржаја унутар Википедије и Вики-пројеката

Методe интегрисања садржаја унутар Википедије подразумевају,

пре свега: додељивање категорија, повезивање са садржајима на Википедији на другим језицима, успостављање међусобних веза унутар пуног текста чланака, повезивање садржаја из Викимедијине оставе, Вики-изворника и Вики-књига са чланцима у Википедији, а за садржаје на Викимедијиној остави и коришћење детаљних и прецизних описа на више језика.

Систем категорија у Википедији ослања се на постојеће таксономије, али добрим делом почива на консензусу великог броја корисника. Да је то тако показује чињеница да се систем категорија на Википедијама на различитим језицима разликује, те да зависи и од покривености одређених области и тема. Тај проблем је посебно изражен у Википедијама на тзв. „малим“ језицима, па корисници који обрађују тему из слабије покривене области често немају на располагању одговарајуће предефинисане категорије које би приписали својој одредници. Мање искусни уредници, који не познају довољно технику рада у оквиру Вики-пројеката, најчешће ће одустати од покушаја да садржај који су креирали инкорпорирају у постојећи корпус. Управо на овом плану библиотекари – који свакодневно у свом раду користе различите индексе, таксономије, тезаурусе и сл. – могу да пруже значајну помоћ, пре свега кроз ревизију постојећих система категорија у оквиру Вики-пројеката и интервенције које имају за циљ њихово побољшавање (креирање нових категорија, преименовање, брисање, спајање, дефинисање хијерархијских односа). Искуство које на тај начин стичу може се применити и у ревизији и оптимизацији тзв. званичних термилошких ресурса који се користе у библиотекарству.

Примећено је и да релативно велики број текстова на Википедији на српском језику није повезан са текстовима на другим језицима.² Некада то није ни могуће урадити зато што не постоје одговарајуће одреднице на другим језицима. У великом броју случајева ради се, међутим, о томе да аутор одреднице на српском језику није знао назив те одреднице на другим језицима. Дешава се, такође, и да су одреднице на српском језику погрешно повезане са одредницама на другим језицима. Библиотекари се приликом каталогизације

² Дешава се, међутим, и да текстови на енглеском језику нису повезани са другим језицима, нарочито када еквивалентна одредница постоји само на неком „малом“ језику.

и претраживања каталога и база података често срећу са литературом на језицима које не познају. Такве ситуације разрешавају коришћењем референсне литературе, која им је доступна на радном месту. На исти начин они могу да идентификују еквивалентне одреднице на Википедијама на различитим језицима и да препознају и исправе грешке у повезивању. Имајући у виду да се управо ови подаци похрањују у централни репозиторијум Викимедијиних пројеката – Вики-податке, исправљањем грешака у овом домену директно се утиче на квалитет структурисаних података који се користе у напредним апликацијама које екстрахују податке из Вики-пројеката и користе их у контексту семантичког веба.

Библиотекари могу да обогате постојеће текстове новим референцама, као и да провере релевантост и тачност постојећих. Поред тога, они могу да побољшају и допуне описе материјала који већ постоје на Викимедијиној остави, а које користе у чланцима, а по потреби могу и да их сврстају у нове категорије како би их лакше повезали с тим текстовима.

Реализовани ГЛАМ пројекти показали су како се садржаји са Вики-пројеката могу успешно укључити у музејске изложбе посредством QR кодова (“Wikipedia: GLAM/Derby/QR Code Experiment”). На исти начин се могу обрадити и садржаји из библиотеке – значајне публикације на полицама доступним корисницима или материјали приказани на повременим, најчешће тематским изложбама. Припрема једне такве изложбе подразумевала би, поред уобичајених активности, и уређивање чланака који су повезани са темом изложбе и депоновање одговарајућих материјала на Викимедијину оставу или Вики-изворник. С друге стране, на тај начин би се обезбедио редован прилив квалитетних чланака, који би се, у зависности од профила очекиваних посетилаца изложбе, писали и на страним језицима.

Иако у Србији у овом тренутку не постоје услови за систематско повезивање библиотечких каталога и репозиторијума са подацима из Википедије, рад на том пољу могао би се започети додавањем у каталог линкова ка дигитализованим садржајима депонованим на Викимедијину оставу и у Вики-изворник, те повезивањем садржаја са сајта библиотеке са садржајима у оквиру Вики-пројеката.

Закључак

Можемо очекивати да сарадња између библиотеке и Викимедије Србије још неко време остане у оквирима ГЛАМ пројеката. Библиотеке у Србији још увек немају капацитете неопходне за креирање и вођење значајнијих пројеката сарадње са Викимедијом у сфери семантичког веба. С друге стране, ни обим и квалитет садржаја на Википедији на српском језику не могу се поредити са Википедијама на тзв „великим“ језицима.

Рад на пројекту *Вики-библиотекар* показао је да библиотекарима могу значајно да допринесу обогаћивању Википедије и других Вики-пројеката квалитетним садржајима. Поред тога, радом на интеграцији садржаја у оквиру Вики-пројеката библиотекарима утичу на квалитет структурисаних података који ће касније бити коришћени као повезани јавно доступни подаци (Linked Open Data), чиме се практично припрема терен за неке будуће пројекте у домену семантичког веба.

Литература

- “Википедија: Вики гимназијалац.” *Википедија, слободна енциклопедија*. 2015. Web. 13. дец. 2015. <https://sr.wikipedia.org/wiki/Vikipedija:Viki_gimnazijalac>.
- “Википедија:Семинарски радови.” *Википедија, слободна енциклопедија*. 2015. Web. 13. дец. 2015. <https://sr.wikipedia.org/wiki/Vikipedija:Seminarski_radovi>.
- Voss, Jakob. “Wikipedia as Knowledge Organization System.” 2009. Web. 22 Sept. 2015. <<http://www.slideshare.net/nichtich/wikipedia-as-knowledge-organization-system-2374373>>.
- Le Bœuf, Patrick. “Customized OPACs on the Semantic Web: The OpenCat Prototype.” 2013. Web. 1 Oct. 2015. <<http://files.dnb.de/svensson/UILLD2013/UILLD-submission-3-formatted-final.pdf>>.
- Lucarelli, Anna. “«Wikipedia Loves Libraries»: In Italia è un amore corrisposto...” *AIB studi* 54.2/3 (2014). Web. <<http://aibstudi.aib.it/article/view/10108>>.
- Müller-Birn, Claudia, Benjamin Karran, Janette Lehmann, and Markus Luczak-Rösch. “Peer-Production System or Collaborative Ontology Engineering Effort: What Is Wikidata?” *Proceedings of the 11th International Symposium on Open Collaboration*. New York, NY, USA: ACM, 2015. 20:1–20:10. *ACM Digital Library*. Web. 13 Dec. 2015. <<http://doi.acm.org/10.1145/2788993.2789836>>. OpenSym '15.

- Pohl, Aleksander. "Classifying the Wikipedia Articles into the OpenCyc Taxonomy." *Proceedings of the Web of Linked Entities Workshop in Conjunction with the 11th International Semantic Web Conference (ISWC 2012)*. Ed. Giuseppe Rizzo et al. Vol. 906. Boston, USA: CEUR, 2012. 5–16. Web. 5 Oct. 2015. <<http://ceur-ws.org/Vol-906/paper2.pdf>>. CEUR Workshop Proceedings.
- "Статистике." *Википедија, слободна енциклопедија*. 2015. Web. 13. дец. 2015. <<https://sr.wikipedia.org/wiki/Posebno:Statistike>>.
- "Wikipedia:GLAM/British Library." *Wikipedia, the free encyclopedia*. 2015. Web. 13 Dec. 2015. <https://en.wikipedia.org/w/index.php?title=Wikipedia:GLAM/British_Library&oldid=665179997>.
- "Wikipedia:GLAM/Derby/QR Code Experiment." *Wikipedia, the free encyclopedia*. 2011. Web. 13 Dec. 2015. <https://en.wikipedia.org/wiki/Wikipedia:GLAM/Derby/QR_code_experiment>.
- "Wikipedia Loves Libraries: Outreach Wiki." 2015. Web. 13 Dec. 2015. <https://outreach.wikimedia.org/wiki/Wikipedia_Loves_Libraries>.
- Libraries and Wikipedia together on the Web: Open Knowledge for Everyone

Aleksandra Popović

University Library "Svetozar Markovic", Belgrade

Milica Ševkušić

Institute of Technical Sciences of SASA

Đorđe Stakić

University of Belgrade

Faculty of Mathematics

**LIBRARIES AND WIKIPEDIA TOGETHER ON THE WEB:
OPEN KNOWLEDGE FOR EVERYONE****Summary**

The digitization of library holdings and enabling open access to various types of content are nowadays major priorities for libraries, as well as for the supporters of Open Science and Open Access.

In 2015, the University Library Svetozar Marković, the Faculty of Mathematics of the University of Belgrade and Wikimedia Serbia have been implementing the Wiki-Librarian project. The guiding idea of the project is to enrich Wikipedia, as the most im-

portant project of the Wikimedia Foundation, with as much quality content as possible. Wikimedia Serbia was established back in 2005 as Wikimedia's fifth national chapter. It is a non-governmental, non-profit and non-political organization supporting the idea that everyone should have equal access to knowledge and education. The Institute of Technical Sciences of SASA is also involved in the project as an important partner in achieving these objectives.

Libraries hold invaluable treasures that are not easily available to everyone. Uploading content to Wikimedia Commons, Wikisource or Wikibooks (depending on content type) is an ideal way of promoting works with expired copyright or those whose authors wish to make them freely available. Librarians – as a well-connected community whose members have access to reference literature at their workplace, are skilled in locating valid and verifiable information, familiar with their institutions' rich holdings and collections and experienced in training users of different ages and education levels – are a perfect choice as good Wikipedia editors and participants in other Wiki projects.

The paper discusses methods and mechanisms to make Wikipedia-related content (especially digitized materials on Wikimedia Commons, Wikisource or Wikibooks) visible and easily retrievable. Inside Wikipedia, these methods include: category assignment, linking content between Wikipedia's in different languages, interlinking within the full text of articles and linking content in Wikimedia Commons, Wikisource or Wikibooks with articles in Wikipedia, as well as detailed and precise descriptions in multiple languages for content uploaded to Wikimedia Commons. Attention will also be drawn to the possibilities for linking content from Wiki projects and online catalogues, digital libraries, repositories and library websites. Indexing procedures in Wikipedia are somewhat different from subject indexing in conventional library catalogues and they conform to logic more similar to indexing in digital repositories. Relying on their expertise, librarians can significantly contribute to better data structuring and more effective content indexing in Wikipedia projects.

Key words: libraries, Wikipedia, Wikimedia Serbia, digitization, Open Access