

Razvoj normativne datoteke tematskih predmetnih odrednica u elektronskom katalogu rukopisa i arhivalija AtoM Narodne biblioteke Srbije

Ivan Obradović, Ana Stevanović, Nebojša Kovačević

Elektronski katalog rukopisa

Tijekom 2015. godine stručnjaci Narodne biblioteke Srbije intenzivirali su rad na pitanjima elektronske obrade Fonda mlađih književnih rukopisa i arhivalija Odjela posebnih fondova, kroz sveobuhvatnu analizu postojećih obavijesnih pomagala za arhivsku građu i rukopise koji se čuvaju u Odjelu, s osnovnim ciljem uvođenja opisnih normi Međunarodnoga arhivskoga vijeća (ICA) i EAD standarda.

Stari katalog rukopisa i arhivske građe na kataložnim listićima, formiran krajem četrdesetih godina 20. stoljeća na temelju bibliografskog opisa, s vremenom se pokazao kao nekonzistentan i nedovoljno informativan. Iso tako, kroz rad na projektu „Europeana zbirke 1914-1918“ na pripremi digitalne biblioteke „Veliki rat“ i izrade shema metapodataka za digitalne objekte navedene biblioteke, stručnjaci Narodne biblioteke Srbije suočili su se s posebnim zahtjevima i problemima obrade ove vrste građe. Bilo je neophodno prilagoditi se novom online okruženju i novim navikama, potrebama i zahtjevima korisnika.

Pridržavanje međunarodnih standarda, prvenstveno ISAD(G)-a, i izbor odgovarajuće aplikacije koja ga podržava nametnulo se kao osnovni preduvjet budućega rada. Također, nakon detaljne analize postojećih dostupnih sustava za arhivsku (i bibliografsku) obradu građe, fokus je usmjeren na aplikacije u otvorenom pristupu (*open source*). Ovakav pristup ostavio je slobodnu perspektivu za dalji razvoj sustava na razini nacionalne biblioteke, ali i na nacionalnom nivou. Također, omogućavao je različite konverzije u pogledu izvoza i uvoza podataka iz najrazličitijih aplikacija i baza podataka, ukoliko su usklađene sa

zajedničkim standardima za arhivski i/ili bibliografski opis. Odmah su se morala rješavati i pitanja lokalizacije, pri čemu je, u skladu s jezikom, bila neophodna i podrška za ćirilično pismo.

Najpogodnijim rješenjem pokazala se aplikacija AtoM kanadske tvrtke *Artefactual Systems*, koja je nastala u okviru aktivnosti ICA-e za izradu aplikacije koja bi omogućila dostupnost arhivskih fondova u online okruženju.



Početni ekran Elektroniskog kataloga

Može se reći da je već samo ime aplikacije na neki način odredilo i put kojim treba krenuti. U pogledu sadržajne obrade, to je bio put prilagođavanja i korištenja bibliografskih instrumenata kao što su normativne pristupnice: tematske, geografske, vremenske ili one osobnih imena. Sam naziv programa, Pristup sjećanju (*Access to Memory*), kao da je svjedočio o nužnosti kombinacije arhivskih i bibliografskih instrumenata, jer je cilj bio ono što daje smisao čitavom poduhvatu. Na ovaj ili onaj način, bavili smo se očuvanjem (ili oživljavanjem) prošlosti, kroz teme, lica, događaje, intimne i manje intimne zapise, povijesne događaje i sudbinu pojedinaca u njima. Kao da je čitav poduhvat tako bio svedočanstvo istinitosti iskaza velikog bibliotekara Nikolaja Fjodorova iz biblioteke Rumjanceva, koji u djelu *Filozofija zajedničkog djela* (Философия общего дела) govori o zajedničkom djelu očeva i sinova, predaka i potomaka,

koji nije ništa drugo do „oživljavanje“, jer: „Muzej (arhiv/biblioteka – prim.) nije skup stvari, već sabor ličnosti ... vraćanje života svemu što je odživjelo svoj vijek“. Tako je ovaj „pristup sjećanju“ zaista svjedočio o bliskosti i istoj usmjerenosti posla i zadatka arhivista i bibliotekara, u ovom slučaju klasifikatora.

Kao rezultat svih gore navedenih aktivnosti Narodna biblioteka Srbije ustanovila je Elektronski katalog rukopisa i arhivalija AtoM NBS (EKRAAN) čime je riješen dugododišnji raskorak između dostupnosti različitih vrsta građe koji se čuvaju u Odjelu posebnih fondova. EKRAAN se još uvek nalazi u testnoj fazi i njegov konačni završetak se očekuje tijekom 2018. godine, a od rujna 2017. je dostupan korisnicima online.

Normativna datoteka predmetnih odrednica

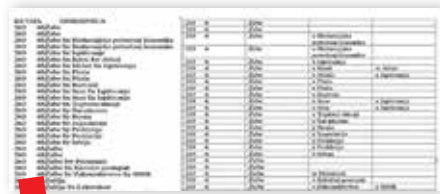
Nakon rešavanja inicijalnih problema opisa građe na razini inventara i osnovnog arhivskog opisa, na red je došlo pitanje sadržajne obrade građe. U pogledu normatizacije osobnih imena i dalje je otvorena mogućnost povezivanja s normativnom bazom osobnih imena CONOR.RS koja se formira za potrebe bibliografske baze COBIB. U testnoj fazi je prema načelima za normativne pristupnice formirana posebna baza osobnih imena za EKRAAN, što je svakako teži i neekonomičniji put. Pitanje geografskih imena riješeno je povezivanjem elektronskog kataloga s bazom geografskih naziva *GeoNames*.

Sadržajno najzahtjevniji zadatak je bio oblikovanje nekog oblika normativne, kontrolirane liste općih tematskih predmetnih odrednica za bazu EKRAAN. Razlozi su bili vrlo

jednostavni: potrebna je što šira i sveobuhvatnija lista tematskih predmetnih odrednica, jer se rukopisna i arhivska građa Fonda mlađih književnih rukopisa i arhivalija odnosi na najrazličitije povijesne, osobne, umjetničke i znanstvene teme. Kako Narodna biblioteka Srbije još uvijek nema normativnu listu tematskih predmetnih odrednica, bilo je neophodno formirati takvu listu, pa je izbor pao na već pomalo zaboravljenu listu predmetnih odrednica koja je nastala kao rezultat retrospektivne konverzije tiskanog *Kataloga knjiga na jezicima jugoslovenskih naroda 1868-1972*, dio s općim tematskim odrednicama iz starog kataloga na listićima.

Lista je nastala kroz veliki posao konverzije predmetnog kataloga na listićima (od 1949. do 1989. godine) u bazu podataka u SQL programu. Napravljena je redakcija ukupno 900 ladica predmetnog kataloga i dobivena je baza koja sadrži 605.869 jedinica – tematskih odrednica (uključujući duplikate), a rezultati konverzije su bili dostupni u jednostavnim, neformatiranim tekstualnim dokumentima.

Kako bi se pristupnica mogla koristiti u okviru aplikacije AtoM bilo je neophodno uvesti uređenu listu predmetnih odrednica u SKOS (*Simple Knowledge Organization System*), u .xml formatu. Iz tog su razloga postojeće liste prvo konvertirane u upotrebljivi .txt oblik, pri čemu je kao najpogodnija usvojena konverzija u tablicu. Uz pomoć niza operacija (pretraživanja i algoritma) tematske odrednice su razdvojene od različitih vrsta pododrednica uz koje se javljaju, a potom su iz dobijene liste uklonjeni duplikati, tako da je konačno iz taksonomije tematskih predmetnih odrednica kataloga na listićima dobijeno oko 5.500 kontroliranih tematskih predmetnih odrednica.



Izgled liste prije i posije konverzije u tablicu

Za slova DŽ i LJ, koja su iz nepoznatih razloga preskočena u retrospektivnoj konverziji, proveden je sličan postupak, ali iz javno dostupne baze COBISS.

NBS (za razdoblje od 1989. do 2016. godine). Nakon niza simultanih pretraživanja dobivena je tablica usvojenih tematskih predmetnih odrednica, iz koje su kasnije uklonjene pododrednice, a potom i duplikati. Na temelju dobivenih rezultata zaključeno je da bi bilo potrebno obaviti identičan posao izlučivanja usvojenih tematskih odrednica iz baze NBS i za sva ostala slova.


Potom, po završetku posla, rezultati su ukršteni i umjesto postojećih 5.500 dobijeno je 21.801 jedinstvenih tematskih odrednica (tematske predmetne odrednice iz retrospektivne konverzije kataloga na listićima i javno dostupnog elektronskog kataloga COBIB.RS. U tijeku je postupak provjere i redakcije u okviru kojeg je procijenjeno da će rezultat biti jedinstvena lista s 16.000 - 18.000 tematskih predmetnih odrednica. Nakon osnovne redakcije i prebacivanja tematskih predmetnih odrednica u ćirilčno pismo, izvršene su usporedbe i testiranja, a dobivena .txt datoteka s odrednicama se lako konvertira u .xml dokument po SKOS standardu. U idućim bi fazama rada ova normativna lista/taksonomija u SKOS formatu mogla komunicirati s drugim formatima i bazama te se povezivati s bazama na engleskom i drugim jezicima i s web aplikacijama. Već u testnoj fazi dijelovi normativnih lista odrednica uspješno su uvezeni u AtoM i dostupni su katalogizatoru preko „padajućih“ menija i unosa početnog slova.



Izgled .txt datoteke konvertirane u SKOS

Prevedene normativne liste mogu se uspješno povezati s normativnim bazama na Internetu (npr. za opće pojmove *Getty*, za osobe *DBpedia*, za vremenska razdoblja *Semium Time*, za geografske pristupne tačke *Geonames*), koje su dio *Linked Open Data* i imaju jedinstvene identifikatore - URI. U testnom razdoblju su u AtoM, probe radi, uspješno uvezene neke od postojećih baza odrednica za društvene

znanosti kao što je npr. UNESCO tezaurus. Tezaurus je u .rdf formatu i obuhvaća sedam osnovnih područja i niz mikrotezaurusa: Obrazovanje; Kultura; Prirodne znanosti; Društvene i humanističke znanosti; Informacije i komunikacija; Politika, pravo i ekonomija i Države i grupe država. Kroz rad na redakciji uočen je visok stupanj podudaranja UNESCO tezaurusa s dobivenom kontroliranom listom odrednica, koja je višestruko obimnija (4.500 vs. 21.801 termina).

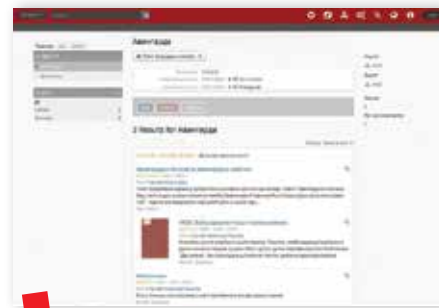


Višejezični UNESCO tezaurus na ruskom i engleskom jeziku i EKRAAN lista odrednica u sredini

U testnim uvjetima uspješno je izvršena proba konverzije i uvoza podataka iz UNESCO tezaurusa u EKRAAN, što je potvrdilo komunikaciju aplikacije AtoM sa odgovarajućim bazama podataka. Čitav ovaj postupak je testiran i na postojećoj elektronskoj bazi rukopisa i arhivalija Narodne biblioteke Srbije, za sada na ograničenom broju dokumenata, jer je baza u razvojnoj fazi.



Podaci iz UNESCO tezaurusa uvezeni u EKRAAN



Predmetna odrednica s povezanim arhivskim opisima

Na primjeru odrednice za Prvi svetski rat prikazano je moguće povezivanje osobe i odgovarajućeg povijesnog događaja, kao jedan segment mogućeg semantičkog bogatstva koje se otvara kroz primjenu bogate i normativno uređene liste tematskih predmetnih odrednica. Naravno, ova iskustva otvorila su i nova pitanja i nove mogućnosti, pa se arhivistima kroz bibliografski i klasifikatorski pristup opisu građe nude nove mogućnosti, preko razvoja drugih kontroliranih taksonomija, koje je takođe moguće izlučiti iz bibliografskih baza.

Elektronski katalog rukopisa i arhivalija ujedno je poslužio i kao pogodan alat za šira i dalekosežnija saznanja vezana uz opis heterogene i specifične građe. Prije svega, kombinacijom već postojećih sadržajnih instrumenata poput tematskih odrednica, moguće je opis višestruko semantički obogatiti i kroz online predstavljanje učiniti dostupnim svim vrstama korisnika. Isto tako, ovaj pristup omogućavaju ne samo komunikaciju između različitih sustava za opis građe, nego i nadilazi ograničenja vrsta građe.



Predmetne pristupne točke u arhivskom opisu

Primjer baze EKRAAN pokazuje da je moguće dobiti listu kontroliranih predmetnih odrednica svih vrsta (tematske, obiteljske korporativne, vremenske, stvarnog naslova...) i učiniti ju interoperabilnom u različitim sustavima i na webu. Sličnim postupkom moguće je dobiti i kontrolirane liste različitih

tematskih pododrednica, uz svaku odrednicu. Na taj bi se način dobio složeni semantički instrument, koji svojom informativnošću, u zavisnosti od broja normiranih pristupnica, može zadovoljiti i najspecifičnije potrebe korisnika i stručnjaka iz različitih područja.

Sve ovo je isprobano i funkcionalno kroz uvoz normativnih datoteka u formatima .rdf (UNESCO tezaurus) i .csv (lista tematskih odrednica). Lista je uspješno konvertirana u upotrebljiv SKOS/RDF format. Osim tehničkih problema s jezikom i pismom, koji su bili minimalni i koji su riješeni u novijim verzijama programa, sve funkcionira kroz "padajuće menije" i tako je dostupan klasifikatorima u EKRAAN-u.

Tijekom rada na testiranju elektronske baze arhivalija i dokumenata postojala je svijest o tome da se radi o ograničenom broju zapisa, ali i nastojanje da se u pogledu tematskih odrednica uvoz u aplikaciju AtoM obavlja za što veći, reprezentativan broj normativnih jedinica (odrednica). Baza je izdržala ove testove i kada se radilo o više stotina ili više tisuća jedinica koje su automatski „importirane“ u bazu i to preko različitih, široko prihvaćenih formata.

Konačno, kroz formiranje semantičkih instrumenata za elektronski katalog EKRAAN Narodne biblioteke Srbije suočili smo se sa nekom vrstom hermeneutičkog bibliografsko-arhivističkog kruga. I pored ogromnog napretka automatizacije sustava za organizaciju znanja, u određenom trenutku je nužno osloniti se na rad i nasljeđe generacija stručnjaka-klasifikatora, naših prethodnika, što je prevladavanje onoga što Elaine Svenonius naziva „semantičkom preprekom“: „...automatizacija može napredovati samo dok se ne dođe do semantičke prepreke. Na toj prepreci pitanja značenja i značaja dolaze u prvi

plan, a to znači da rad na automatizaciji organiziranja informacija mora zastati i osloniti se na postojeće leksičke informacije, sakupljene i strukturirane intelektualnim radom“.

U slučaju naše baze, pomalo paradoksalno, ne samo zbog sugestivnog naziva aplikacije, *Access to Memory*, već zbog prirode samog procesa, među zaključcima iz iskustva rada na pokušaju normiranja sadržajnog opisa rukopisa i arhivalija ili normativnoj listi tematskih odrednica je i ovaj: najpouzdanije i najtrajnije je ono što je sačuvano u sjećanju ili kroz ponovno sjećanje (ili oživljavanje sećanja) na intelektualni rad generacija kolega na organiziranju jednog koherentnog sustava znanja, u ovom slučaju arhivističko-bibliografskog.

Time je praktični posao, nastao iz nasušne potrebe za opisom konkretne zbirke, onima koji su u njemu učestvovali poslao i jednu dalekosežniju poruku: budućnost opisa i prije svega pretraživanja (na Internetu ili u bazama podataka) je semantička, time u suštini više ljudska i pored privida automatike. U srži te semantike uvijek je neki intelektualni, ljudski rad. Ili, drukčije rečeno, neka želja ili sjećanje na nekoga ili nešto. Sjećanje nestalno i varljivo kao i sama Mreža, ali izgleda i trajno, poput one fine materije „od koje su sačinjeni snovi“.



■ Saznajte više:

Informacije o aplikaciji AtoM možete naći na: Artefactual Systems, "Providing access to memory since 2007": <https://www.accesstomemory.org/en/>.

EKRAAN je dostupan na: <http://arhivalije.nb.rs/>.

Više o bazi podataka *GeoNames* možete saznati na: GeoNames, „About GeoNames“, <http://www.geonames.org/about.html>