

VRLO KRATAK (I POPULARAN) UVOD U METAPODATKE¹

Boris Bosančić

*Odsjek za informacijske znanosti, Filozofski
fakultet, Sveučilište J. J. Strossmayera u
Osijeku, Osijek, Hrvatska*

KLJUČNE RIJEČI:

*metapodaci, sheme metapodataka,
informacijski objekti,
interoperabilnost, aplikacijski profili*

SAŽETAK

U radu je izložen ontološki pristup metapodacima, današnjem području koje se bavi opisom (pretežito) digitalnih izvora informacija, proizišlog iz knjižničnog područja bibliografske organizacije i kontrole. Rad nastoji odgovoriti na pitanja: „Što su (to) metapodaci?“, „Što njima opisujemo?“ i „Kakvu korist imamo od toga?“. Objašnjena su i dva ključna procesa u vezi sa samom srži koncepta – proces izrade metapodatkovnog opisa i metapodatkovnog zapisa. S jedne strane, knjižnični standardi za opis građe evoluiraju u sheme metapodataka za opis informacijskih objekata svih vrsta, koje se mogu mapirati s drugim shemama metapodataka te kombinirati u aplikacijskim profilima. S druge strane, knjižnični zapisi označeni putem strojno čitljivih formata pohrane, kao što je to RDF, evoluiraju u metapodatkovne zapise koji se, uz pretraživanje i pregledavanje, sada mogu i pobirati (putem OAI-PMH) te konvertirati iz jednog formata zapisa u drugi. Na koncu, dan je osvrt i na interoperabilnost metapodataka u mrežnom okruženju, ističući ulogu metapodataka u okviru semantičkog weba, na čijem je daljnjem razvoju položen i daljnji razvoj metapodataka u cjelini.

1 Ovaj vrlo kratak i popularan uvod u metapodatke proizišao je iz nastavnih materijala kolegija Metapodaci i identifikatori koji sam s prof. Mirnom Willer, kao njezin asistent, izvodio paralelno na Sveučilištu u Osijeku i Sveučilištu u Zadru između 2007. i 2012.

Uvodna razmatranja

Metapodaci su jedna vrsta podataka koja nešto kaže o drugim vrstama podataka. Zbog toga se u običnom, kolokvijalnom govoru za metapodatke kaže da su „podaci o podacima“. No, s filozofskog stajališta, za metapodatke se tvrdi da posjeduju mogućnost opisivanja stvarnosti (Floridi, 2010), a tu se prije svega misli na opisivanje *stvari* ili *entiteta* koji postoje u stvarnosti. U tom smislu, entitet je svaka *stvar* koja se jednoznačno (i postojano) razlikuje od neke druge *stvari*, a što je najvažnije obilježje jednog drugog pojma, usko povezanog s pojmom metapodatka – pojam identifikatora. Za identifikatore možemo reći da jednoznačno i postojano identificiraju entitete, odnosno stvari u svijetu, kako bi ih čovjek, u ovom slučaju – specijalist metapodataka (*metadata specialist*), osoba koja se bavi izradom i upravljanjem metapodacima – mogao razlikovati.

U vezi s metapodacima, je li doista riječ o *opisu podataka* koji posredno zrcale svojstva *stvari* ili je riječ o neposrednom opisu samih *stvari*, čini se da nikada nećemo biti sigurni. Prešutno ćemo prihvatiti da metapodacima opisujemo *stvari* koje su unaprijed, već po svojoj prirodi *informativne* i tu zaustaviti svaku daljnju filozofsku raspravu.

Danas metapodaci prožimaju kompletno mrežno okruženje. Od tražilica, preko društvenih mreža do Wikipedije, metapodaci su nezamjenjiv kotačić postupaka indeksiranja, pretraživanja, pregledavanja, identificiranja i pobiranja (*harvesting*). Od najobičnijih opisnih ili deskriptivnih metapodataka koji se rabe za opis videa na YouTubeu, knjižne i druge vrste građe u knjižničnim katalogima, preko strukturnih koje pogone Googleov *Knowledge Graph*, jedan od najvećih repozitorija metapodataka na mreži o ljudima, mjestima, stvarima i odnosima među njima, ili administrativnih, poput metapodataka za upravljanje pravima pristupa (*digital rights management*) te, na koncu, uporabnih, koje svakodnevno generiraju korisnici *Facebooka* i *Instagrama*, metapodaci su sveprisutni entiteti na internetu.

Dakako, prostor samog rada ne dopušta da svoju pozornost posvetimo širokoj lepezi uporabe metapodataka na mreži danas, nego se njegova svrha ogleđa u prikazu njihove ontološke dimenzije. U skladu s tim, pitanja: „Što su (to) metapodaci?“, „Što njima opisujemo?“ i, na koncu, „Kakvu korist imamo od toga?“ jesu ona na koja ćemo pokušati pronaći odgovor. Nakon kratke povijesti metapodataka uslijedit će opis dvaju ključnih procesa povezanih s njihovom najužom prirodom – proces izrade metapodatkovnog opisa i proces izrade metapodatkovnog zapisa. Na koncu, interoperabilnost metapodataka u mrežnom

okruženju posljednja je tema kojoj će biti posvećena pažnja u radu, a na čijem je daljnjem razvoju položen i daljnji razvoj područja u cjelini.

U funkcionalnom smislu, nema razlike između metapodatkovnog i bibliografskog odnosno kataložnog opisa. No, u praksi, metapodaci se uglavnom povezuju s metapodatkovnim zapisima različitih oblika digitalnih izvora informacija. Zbog velike količine takvih izvora informacija, metapodatke u velikom broju slučajeva kreiraju sami autori, a u novije vrijeme čak i strojevi.

Granica između metapodataka i podataka u novije vrijeme postaje nejasna: metapodatak za jedan informacijski objekt može biti podatak u drugom! Tako, vrijednosti navedene u zaglavlju tablica skupova podataka (*datasets*), iako čine metapodatke za podatke pohranjene u tablici, podaci su i za metapodatke koji opisuju promatrani skup podataka u cijelosti!

Kratka povijest metapodataka

U posljednjih stotinu pedeset godina metapodacima su se uglavnom bavili knjižničari u okviru vlastitih postupaka katalogizacije, klasifikacije i indeksiranja. Za MARC – prvi strojno čitljiv format zapisa metapodataka – obično se tvrdi da je prvi metapodatkovni standard uopće (Bosančić, 2018). No uporaba termina *metapodaci* ne pronalazi svoj korijen u knjižničarstvu. Prema P. Caplan (2003), termin *metapodaci* porijeklo vuče iz računalnih znanosti. Prvo spominjanje termina *metapodaci* u današnjem značenju javlja se u NASA-inu dokumentu *Directory Interchange Format Manual* objavljenom 1988.² Prije sredine 1990-ih godina, termin *metapodaci* rabi se za potrebe upravljanja i održavanja geoprostornih informacija odnosno skupova podataka istraživanja u priručniku *Federal Geographic Data Committee's Content Standard for Digital Geospatial Metadata*, objavljenom 1994. Sa stupanjem Dublin Core sheme metapodataka na scenu 1995. započinje prava povijest metapodataka.

Povod nastanku Dublin Core sheme metapodataka ogleda se u potrebi definiranja univerzalnog skupa elemenata metapodataka za opis mrežnih izvora čiji je broj u to vrijeme enormno rastao. Pored toga, Dublin Core shema metapodataka bila je namijenjena samim autorima. No, ubrzo se pokazalo da se, uz mrežne izvore informacija, DC shema metapodataka može rabiti i za opis i drugih vrsta

2 Zanimljivo, sâm termin „metapodaci“ (*metadata*), prema P. Caplan, skovao je Jack E. Myers u kasnim 1960-im godinama za potrebe naziva svoje tvrtke (*Metadata Company*) koju je registrirao 1986., ali koja nije imala nikakve poveznice s današnjim značenjem tog pojma (Caplan, 2003).

izvora informacija. U skladu s tim, Inicijativa Dublin Core metapodataka (*Dublin Core Metadata Initiative* – DCMI), koja stoji iza DC skupa elemenata metapodataka, propisala je DCMIType – kontrolirani rječnik termina koji sadrži popis svih vrsta izvora informacija koji se mogu opisati DC skupom elemenata.³ Sama inicijativa nastavila se baviti i razvojem drugih metapodatkovnih standarda (npr. DCMI Apstraktnog modela, Dublin Core aplikacijskog profila i sl.) te smjernica za njihovu uporabu, općenito osiguravajući podršku najboljim praksama korištenja metapodataka u mrežnom okruženju.

Valja naglasiti i da situacija oko knjižničnih standarda za opis elektroničke građe krajem stoljeća nije bila blistava. U mrežnom okruženju MARC više nije mogao ispuniti svoju funkciju. Ni MARC21, novija inačica MARC-a, namijenjena knjižničnoj zajednici i izvan SAD-a, a koja se pojavila na prijelazu stoljeća, nije to bila u stanju. Roy Tennant 2002. objavljuje članak indikativnog naziva „MARC mora umrijeti“ (*MARC must die*) (Tennant, 2002), a iste godine Kongresna knjižnica u Washingtonu, potaknuta Tennantovim člankom ili ne, razvila je dva nova metapodatkovna standarda koja će iscertati konture nove politike razvoja knjižničnih standarda u mrežnom okruženju. Riječ je bila o MARCXML-u i MODS-u (*Metadata Object Description Schema*).

W3 Konzorcij (*World Wide Web Consortium*) 2001. objavljuje preporuku o korištenju RDF-a (*Resource Description Framework*) strojno čitljivog formata za opis mrežnih izvora koji će u godinama koje slijede postati najvažniji format pohrane metapodataka u mrežnom okruženju.

Tijekom 2000-ih godina metapodaci svoju primjenu sve više nalaze u digitalnim repozitorijima, koji utjelovljuju nov način komunikacije znanstvenika te čine praktični ishod inicijative za otvoreni pristup znanstvenim informacijama, kao novoj mogućnosti dijeljenja znanja u mrežnom okruženju. U tom smislu, Mreža repozitorija otvorenog pristupa (*Open Access repositories* – *OpenAIRE*) započinje objavljivati razvojne smjernice za administratore institucijskih repozitorija u kojima metapodaci igraju veoma važnu ulogu.

Konačno, 10-ih godina razvojem otvorenih povezanih podataka (*Linked Open Data*), metapodaci pronalaze još jednu važnu primjenu – u okviru semantičkog *weba*. U tom smislu, može se govoriti i o otvorenim povezanim metapodacima (Bosančić, 2018).

3 DCMIType uključuje sljedeće termine: zbirku (*Collection*), podatke istraživanja (*Dataset*), događaj (*Event*), sliku (*Image*), interaktivni izvor (*Interactive Resource*), pokretnu sliku (*MovingImage*), fizički objekt (*PhysicalObject*), uslugu (*Service*), program (*Software*), zvuk (*Sound*), platna/cртеže/mape (*StillImage*) i tekst (*Text*).

Informacijski objekt

Sa stajališta metapodatkovnog stručnjaka ili specijalista metapodataka, *stvar* koja se opisuje metapodacima može poprimiti najrazličitije nazive:

- informacijski izvor (npr. na *webu*)
- digitalni objekt (npr. u digitalnoj knjižnici)
- resurs (npr. u općenitom smislu) itd.

Kombinacijom navedenog nazivlja mogu se dobiti varijacije naziva opisiva-nog entiteta: informacijski objekt, digitalni izvor, informacijski resurs, digitalni resurs itd.

Anne J. Gilliland u poglavlju knjige *Uvod u metapodatke*, koju je objavio Gettyjev istraživački institut (*The Getty Research Institute*) 2008., odabrala je naziv „informacijski objekt“ za *stvar* koja se opisuje metapodacima. Pa tako, informacijski objekt može biti bilo što što se može identificirati i njime upravljati, odnosno čime čovjek ili stroj mogu manipulirati kao diskretnim entitetom (Gilliland, 2008). Informacijski objekt može biti obična fotografija, članak, knjiga, zbirka knjiga ili čak kompletan sustav baza podataka!

Pored toga, svaki informacijski objekt može sadržavati više drugih informacijskih objekata. Primjerice, u kontekstu bibliografskog zapisa, nakladnička cjelina kao nadređeni informacijski objekt može sadržavati više knjiga kao zasebnih informacijskih objekata. Čak je i autor knjige zaseban informacijski objekt o kojem se izrađuje normativni zapis. U tom smislu, metapodacima je moguće opisati nadređeni informacijski objekt (nakladničku cjelinu) i sve njegove dijelove (knjige) kao zasebne informacijske objekte.

Svrhe metapodataka

Pored one „podaci o podacima“, u literaturi se navode i brojne druge definicije metapodataka. Udruženje američkih nakladnika (*Association of American Publishers*) metapodatke definira kao informacije koje opisuju sadržaj (Caplan, 2003). Za W3 Konzorcij metapodaci su strojno čitljive informacije za *web* (*Metadate and Resource Description*, 2019), a za P. Caplan (2003), to su strukturirane informacije o informacijskom izvoru na bilo kojem mediju i u bilo kojem formatu.

Dakako, nema općeprihvaćene definicije metapodataka. Definicija pojma umnogome ovisi o zajednici i kontekstu u kojem se metapodaci rabe. Jedna od najpoznatijih definicija metapodataka potječe iz NISO priručnika *Razumijevanje metapodataka* autorice Jenn Riley i glasi:

Metapodaci su strukturirane informacije koje opisuju, objašnjavaju, lociraju ili na neki drugi način olakšavaju pretraživanje, korištenje ili upravljanje informacijskim izvorom (Riley, 2017).

Zapažamo da se navedena definicija sastoji od dvaju dijelova. U prvom se govori o tome što metapodaci *čine*, dok se u drugom nabraja za koje se sve *svrhe* to radi. U tom smislu, moguće je razlikovati sljedeće svrhe metapodataka:

- indeksiranje metapodataka za potrebe postupka *pretraživanja* i lakšeg otkrivanja informacijskih objekata
- organiziranje metapodataka za potrebe postupka *pregledavanja* informacijskih objekata
- mapiranje metapodataka za potrebe postupka *pretraživanja* (Z39.50) i *po-biranja* (OAI-PMH) metapodatkovnih zapisa.

Dakako da se ovdje svrhe metapodataka ne iscrpljuju i da ih je moguće nastaviti nabrajati: arhiviranje i zaštita informacijskih objekata, njihovo korištenje, upravljanje, osiguranje prava pristupa itd.

U novije vrijeme, jedna od najvažnijih svrha metapodataka ogleda se u uključivanju metapodatkovnih zapisa u semantički *web* (*semantic web*) povezanih podataka (*linked data*). Povezani podaci odnose se na najbolju praksu dijeljenja i povezivanja podataka, informacija i znanja u okviru semantičkog *weba* pomoću URI/IRI-ja i RDF-a (*Resource Description Framework*). Prema riječima samog tvorca semantičkog *weba* i koncepta povezanih podataka Tima Bernersa-Leeja povezani su podaci – *semantic web done right* (Berners-Lee, 2008).

Da bi se ostvarile sve navedene svrhe, u širem kontekstu metapodataka, potrebno je razlikovati dva procesa:

- izradu metapodatkovnog opisa informacijskog objekta (katalogizacija, izrada bibliografskog zapisa) – najčešći je slučaj da autor isporuči osnovne metapodatke, a da postupak kreiranja metapodataka dovrši odgovarajući

stručnjak (knjižničar, informacijski specijalist ili specijalist metapodataka)

- izradu metapodatkovnog zapisa – odnosi se na označivanje/enkodiranje metapodatkovnog opisa u jedan ili više formata pohrane (XML, HTML, RDF/XML) kako bi on postao „razumljiv“ i stroju.

Metapodatkovni opis

Svaki metapodatkovni opis započinje tako što se za promatrani informacijski objekt ustvrdi da posjeduje neka svojstva. Za informacijski objekt, primjerice, knjigu *Majstor i Margarita* može se ustvrditi da je crna, napravljena od papira, objavljena 2012. itd. No u ovom, najjednostavnijem mogućem, opisu promatrane knjige nije navedeno upravo ono bitno s metapodatkovnog stajališta – značenja navedenih svojstava, a koja su sama srž metapodatkovnog opisa. Što znači da je knjiga „crna“? „Knjiga je crna“ znači da je knjiga crne boje. To ujedno znači da svako svojstvo knjige koje spominjemo u običnom govoru čini dihotomiju: nositelja značenja svojstva (*boja*) i vrijednosti tog svojstva (*crna*). Ovu terminološku začkoljicu, koju je lako previdjeti u običnom govoru, metapodatkovni stručnjaci odlučili su razriješiti na sebi svojstven način. Nositelje značenja svojstva nazvali su metapodatkovnim svojstvima ili samo svojstvima (*property*), dok su svojstva informacijskih objekata koja se spominju u običnom govoru odlučili nazvati *vrijednostima* metapodatkovnih svojstava ili samo *vrijednostima svojstava* (*value of property*).

Ogledni, ali još nestandardizirani, metapodatkovni opis knjige *Majstor i Margarita* tako bi glasio:

Informacijski objekt knjiga *Majstor i Margarita* crne je boje, napravljena je od (materijala) papira, te objavljena 2012. godine.

Pojednostavljeno rečeno, svaki metapodatkovni opis nekog informacijskog objekta sastavljen je od jednog ili više parova metapodatkovnih svojstava i njihovih vrijednosti koje, onda, zajedno, nazivamo metapodacima:

boja = crna,
materijal = papir,
godina = 2012.

U slučaju informacijskog objekta koji se javlja kao knjižna, mrežna ili neka druga vrsta građe, metapodaci mogu biti sljedeći parovi metapodatkovnih svojstava s njihovim vrijednostima:

naslov = *Majstor i Margarita*

autor = *Mihail Bulgakov*

godina izdavanja = 2012.

nakladnik = *Šareni dućan*

itd.

I kao što je iz primjera vidljivo, neke od vrijednosti metapodatkovnih svojstava mogu postati, ili to već jesu, zasebni informacijski objekti s vlastitim metapodatkovnim opisom (Mihail Bulgakov, Šareni dućan).⁴

Sheme metapodataka

Metapodatkovna svojstva koja se rabe za opisivanje određene vrste informacijskog objekta (poput knjiga, muzejskih eksponata, arhivskog gradiva, mrežne stranice i sl.) i/ili ih kontinuirano rabi pojedina zajednica istraživača mogu pripadati unaprijed definiranom skupu metapodatkovnih svojstava ili tzv. shemi metapodataka (*metadata schema*). Sheme metapodataka skupovi su metapodatkovnih svojstava (elemenata), ali i pravila za njihovu uporabu koji su definirani za posebnu svrhu (Caplan, 2003). Najpoznatiji je takav skup elemenata metapodataka Dublin Core shema metapodataka. Jednostavni (nekvalificirani) DC skup elemenata metapodataka ili jednostavna DC shema metapodataka sastoji se od petnaest izvornih elemenata.⁵

Od ostalih shema metapodataka vrijedi izdvojiti tzv. sheme metapodataka knjižnične zajednice – MODS (*Metadata Object Description Schema*), MADS (*Metadata Authority Description Standard*) i METS (*Metadata Encoding and Transmission Standard*) koje održava Kongresna knjižnica u Washingtonu; zatim, muzejske zajednice – CDWA (*Categories for the Descriptions of Works of Art*)

4 Ovdje izložena ontološka dimenzija metapodatkovnog opisa teorijski je zahvaćena Modelom izvora ili resursa Inicijative Dublin Core metapodataka u okviru Dublin Core apstraktnog modela (<http://dublincore.org/documents/abstract-model/>), čiji prikaz premašuje granice ovog rada.

5 Petnaest izvornih elemenata Dublin Core sheme metapodataka jesu: Naslov (*Title*), Tema (*Subject*), Opis (*Description*), Izvor (*Source*), Jezik (*Language*), Odnos (*Relation*), Obuhvat (*Coverage*), Stvaratelj (*Creator*), Nakladnik (*Publisher*), Suradnik (*Contributor*), Vlasnička prava (*Rights*), Datum (*Date*), Tip (*Type*), Format (*Format*), Identifikator (*Identifier*).

i VRA Core (*Visual Resources Association Core*); te, na koncu, arhivske zajednice – EAD (*Encoded Archival Description*). Tu su i druge sheme metapodataka: LOM (*Learning Object Metadata*) za opis objekata učenja, ONIX (*Online Information Exchange*) koju rabe nakladnici, PREMIS (*PREservation Metadata Implementation Strategies*) za izradu metapodataka za zaštitu, TEI Header za izradu metapodataka digitalne građe označene pomoću TEI standarda (*Text Encoding Initiative*) i dr.

Sheme metapodataka uobičajeno smjernicama/uputama preciziraju naziv metapodatkovnog svojstva (npr. naslov) te njegovu definiciju ili značenje. Pored smjernica/uputa, većina shema metapodataka nudi pristup i zasebnom XML Schema dokumentu u kojem su deklarirana sva metapodatkovna svojstva, odnosno elementi te atributi promatrane sheme metapodataka za strojnu validaciju metapodatkovnih zapisa. Opcionalno sheme metapodataka mogu specificirati:

- pravila za formiranje sadržaja (*content rules*) vrijednosti metapodatkovnih svojstava (npr. kako pisati datum, 2014-10-23 ili 23. 10. 2014.), odnosno uporabu sintaktičkih „enkodirajućih shema“ (*syntax encoding scheme*) poput W3CDTF-a, URI-ja i sl.
- unaprijed definirane vrijednosti metapodatkovnih svojstava, odnosno uporabu termina iz rječničkih „enkodirajućih shema“ (*vocabulary encoding scheme*), kao što su DCMIType, UDK, ili iz kontroliranog rječnika termina kao što je LCSH (*Library of Congress Subject Headings*) ili čak tezaurusa poput MeSH-a (*Medical Subject Headings*).

Kada se za opis određene vrste informacijskog objekta rabe metapodatkovna svojstva iz više shema metapodataka, onda je riječ o aplikacijskom profilu.

Metapodatkovni zapis

Kao što je već navedeno, kada se govori o metapodacima u općenitom smislu, misli se, zapravo, na parove metapodatkovnih svojstava i njihovih vrijednosti. U danom primjeru, metapodatkovno svojstvo ili metapodatkovni elementi jesu *naslov*, *autor*, *godina izdavanja* i *nakladnik*, a njihove vrijednosti „Majstor i Margarita“, „Mihail Bulgakov“, „2012.“ i „Šareni dućan“. Takav metapodatkovni opis izražen na bilo kojem mediju i u bilo kojem formatu naziva se metapodatkovnim zapisom.

Metapodatkovni zapis (*metadata record*), dakle, opis je informacijskog objekta nastao korištenjem elemenata metapodataka određene sheme metapodataka, ali koji je uvijek označen ili *enkodiran* u određenom formatu pohrane (XHTML, XML, RDF/XML).

Označivanje (enkodiranje) metapodataka

Tri su najčešća formata pohrane metapodatkovnih zapisa: XML, HTML (XHTML) i RDF/XML. Na mrežnoj stranici DCMI-ja mogu se pronaći detaljne upute/smjernice za provedbu ovih postupaka.⁶ No, ovdje je važno još naglasiti da je RDF (*Resource Description Framework*) ujedno i model podataka za opis mrežnih izvora. RDF je prepoznatljiv po tome što uobičajene metapodatkovne parove svojstvo-vrijednost svojstva zamjenjuje tripartitnom izjavom u kojoj se svaki put spominje informacijski objekt, svojstvo informacijskog objekta i vrijednost svojstva. Umjesto da se kaže: za zapis identificiran brojem tim i tim, metapodatkovni opis glasi:

naslov = Majstor i Margarita

autor = Mihail Bulgakov

u RDF se modelu podataka „kaže“:

Ovaj zapis (navodi se identifikator) – ima naslov – Majstor i Margarita

Ovaj zapis – ima autora – Mihaila Bulgakova.

Jedna od serijalizacija, odnosno način označivanja RDF modela podataka proveden je i u XML-u, pa je „jezik označivanja“ dobio naziv RDF/XML. RDF je možda najvažniji gradivni blok semantičkog *weba* i daleko najvažnija tehnologija koncepta povezanih podataka. Drugim riječima, RDF je gotovo idealan sintaktički metapodatkovni standard koji omogućuje strojno razumijevanje metapodatkovnih zapisa.

6 Poveznice na sintaktičke smjernice za „enkodiranje“ Dublin Core metapodataka – DC-Text, DC-HTML, DC-DS-XML i DC-RDF – nalaze se na adresi <http://dublincore.org/specifications/>.

Interoperabilnost metapodataka

Općenito o interoperabilnosti

Prema Rječniku Merriam Webster, biti *interoperabilan* (*interoperate*) znači – ‘koji može zajedno djelovati/funkcionirati’ (*to operate together*) (Interoperate, 2019). U načelu, interoperabilnost se odnosi na mogućnost sustava ili proizvoda da djeluje/funkcionira s drugim sustavima ili proizvodima bez pomoći čovjeka (Interoperability, 2019). M. Willer pritom naglašava važnost interoperabilnosti sadržaja metapodataka (Willer, 2001). No, iako se pojam interoperabilnosti uglavnom povezuje sa stupnjem učinkovitosti funkcioniranja različitih sustava i uređaja tehničke prirode, P. Miller pojam interoperabilnosti rabi i u kontekstu učinkovite razmjene i korištenja informacija među različitim organizacijama odnosno ustanovama. Za Millera interoperabilnost znači aktivno se baviti upravljačkim procesom koji – sustavima, procedurama i kulturom neke ustanove – upravlja na način koji dovodi do intenzivnije razmjene i ponovnog korištenja informacija, unutar te ustanove ili među takvim ustanovama (Miller, 2000).

U kontekstu interoperabilnosti metapodataka, L. Mai Chan i M. Lei Zeng (2006) razlikuju tri razine interoperabilnosti:

- razina shema metapodataka (*schema level*) – ostvaruje se postupkom mapiranja (*metadata mapping*) kojem je cilj usklađivanje semantike elemenata iz različitih shema metapodataka (npr. dc. title = 245 \$a MARC21)
- razina metapodatkovnog zapisa (*record level*) – ostvaruje se postupkom konverzije (*metadata conversion*) metapodatkovnih zapisa iz jednog formata pohrane u drugi (npr. iz XHTML-a u XML ili RDF/XML) i kojem je cilj usklađivanje sintakse metapodatkovnog zapisa
- razina repozitorija (*repository level*) – ostvaruje se usklađivanjem protokola za komunikaciju između pobirača ili harvestera i repozitorija (npr. OAI-PMH, SOAP) koji omogućuje postupak pobiranja (*harvesting*) zapisa iz repozitorija putem posebnog programa koji se naziva pobirač ili harvester (*harvester*).

Uz navedene tri razine, M. Willer (2001) navodi i četvrtu koja uključuju uporabu tzv. sintaktičkih i rječničkih „enkodirajućih shema“ u kreiranju vrijednosti metapodatkovnih svojstava.

Mapiranje shema metapodataka

U načelu, mapiranje metapodataka odnosi se na postupak pridruživanja elemenata jedne sheme metapodataka semantički ekvivalentnim elementima druge sheme metapodataka. Prema W. E. Moenu (2005), mapiranje je intelektualna aktivnost koja identificira semantički ekvivalentan element u drugoj shemi metapodataka. U tom smislu, engleski izraz *metadata crosswalk* odnosi se na dokumentaciju koja proizlazi iz postupka mapiranja. Drugim riječima, *metadata crosswalk* čini specifikaciju postupka mapiranja, a najčešće je riječ o usporednim tablicama s dvama ili više stupaca, ovisno o broju shema metapodataka koje se mapiraju.⁷

Konverzija metapodatkovnih zapisa

Kada, pak, metapodatkovne zapise konvertiramo iz jednog formata pohrane u drugi (npr. iz XHTML u XML), tada kažemo da je riječ o postupku konvertiranja ili konverzije metapodatkovnih zapisa. U literaturi se postupak konverzije ponekad naziva mapiranjem metapodataka na sintaktičkoj razini.

Pobiranje metapodataka

Pobiranje ili harvestiranje (*harvesting*) metapodataka temelji se na izravnom dohvatanju metapodataka zapisa pohranjenih u repozitorijima putem odgovarajućeg protokola. Najpoznatiji protokol koji omogućuje pobiranje jest OAI-PMH (*Open Archives Initiative Protocol for Metadata Harvesting*).⁸ OAI-PMH temelji se na HTTP-u, XML-u, XML Schemi i XML imenskim prostorima te omogućuje harvesteru ili pobiraču (posebnom programu) da zatraži i dobije neke ili sve zapise iz repozitorija koji podržava komunikaciju OAI-PMH-om.⁹ Postupak pobiranja razlikuje se od postupaka pretraživanja po tome što se za potrebe pretraživanja sadržaj prethodno indeksira, dok u slučaju pobiranja indeksiranje sadržaja repozitorija nije potrebno.

7 Jedno od najsveobuhvatnijih *mapiranja shema metapodataka* provedeno je na mrežnoj stranici *Metadata Standards Crosswalk* Gettyjeva istraživačkog instituta.

8 OAI-PMH kreirali su Herbert Van de Sompel i Carl Lagoze u okviru Inicijative za otvorene arhive (*Open Archives Initiative*) u kasnim 1990-im godinama.

9 OAI-PMH omogućuje pobiraču pobiranje zapisa iz repozitorija putem točno šest zahtjeva: *Identify* (zahtjev za identifikiranjem repozitorija), *ListMetadataFormats* (pobiranje svih podržanih shema metapodataka u repozitoriju), *ListSets* (pobiranje skupova odnosno kategorija sadržaja repozitorija) *ListIdentifiers* (izlistanje ili pobiranje identifikatora metapodatkovnih zapisa u repozitoriju), *ListRecords* (izlistanje ili pobiranje samih metapodatkovnih zapisa u repozitoriju) i *GetRecord* (izlistanje ili pobiranje pojedinog metapodatkovnog zapisa u repozitoriju).

Aplikacijski profili

I pored velikog broja shema metapodataka za opis različitih vrsta resursa odnosno informacijskih objekata, u praksi se rijetko događa da elementi jedne sheme metapodataka zadovolje sve potrebe opisa konkretnog resursa na lokalnoj razini. Zbog toga se javila potreba za kombiniranjem elemenata iz više shema metapodataka kako bi informacijski objekt na lokalnoj razini bio učinkovito opisan. Ako je ovaj postupak kombiniranja elemenata iz više shema metapodataka standardiziran, onda govorimo da je riječ o aplikacijskom profilu. U skladu s navedenim, aplikacijski profil jest shema metapodataka koja se sastoji od elemenata iz različitih imenskih prostora koji se kombiniraju i optimiziraju za potrebe lokalne aplikacije (Caplan, 2003).

Zašto ne rabimo samo jedan metapodatkovni standard? Možebitno što imamo različite zahtjeve u pogledu metapodatkovnog opisa informacijskih objekata, a ponekad – i različit pogled na svijet: *The days of 'one size fits all' standards are over!* (Rust, 2005). Upravo nemogućnost dogovora oko jednog univerzalnog standarda za opis informacijskih objekata možda je i najveći izazov interoperabilnosti metapodataka u budućnosti!

Svaki aplikacijski profil, u načelu, sastoji se od:

- ljudski čitljive dokumentacije koja se sastoji od opisa načina uporabe elemenata metapodataka iz različitih shema metapodataka u aplikacijskom profilu
- strojno čitljivog XML Schema dokumenta u kojem su deklarirani svi elementi i atributi aplikacijskog profila; pritom se deklaracije elemenata postojećih shema metapodataka uvoze iz njihovih XML Schema dokumenata.

Primjeri okvira/uputa za izradu aplikacijskih profila čine *Dublin Core Application Profile* i *CanCore Application Profile*, dok se kao primjeri konkretnih aplikacijskih profila mogu navesti *UKETD (UK ETheses OAI-PMH format)*, izdavački aplikacijski profil s Dublin Core elementima te *e-GMS (e-Government Metadata Standard)*, metapodatkovni standard vlade Velike Britanije za opis dokumenata.

Zaključak

Metapodaci su nov naziv za „staro“ područje bibliografske organizacije i kontrole, odnosno opisa informacijskih izvora ili objekata. Kataložni i bibliografski zapisi u mrežnom okruženju danas se sve više tretiraju kao metapodatkovni zapisi, jer se svrha samog postupka – opisa informacijskih izvora ili objekata – nije izmijenila, dočim su način izrade bibliografskog (kataložnog) zapisa za građu u digitalnom obliku te prateća terminologija ipak doživjeli bitne promjene. Knjižnični standardi za opis građe postali su sheme metapodataka za opis informacijskih objekata svih vrsta, koje se mogu mapirati s drugim shema metapodataka te kombinirati u aplikacijskim profilima. Knjižnični zapisi označeni putem strojno čitljivih formata pohrane kao što je to RDF postali su metapodatkovni zapisi koji se, uz pretraživanje i pregledavanje, mogu i pobirati putem OAI-PMH-a te konvertirati iz jednog formata zapisa u drugi.

Očekuje se da će svoj potencijal metapodatkovni standardi nastaviti razvijati i u okruženju semantičkog *weba*, u kojem se već danas isprepliću s razvojem računalnih ontologija, strojno čitljivih tezaurusa, pravila i drugih srodnih standarda na kojima se temelji razvoj ovog područja.

LITERATURA

- BERNERS-LEE, T. (2008). *The Linked Open Data Movement*, [\(https://www.w3.org/2008/Talks/0617-lod-tbl/#\(3\)\)](https://www.w3.org/2008/Talks/0617-lod-tbl/#(3)) (01-02-2019).
- BOSANČIĆ, B. (2018). Otvoreni povezani podaci i metapodaci. U I. Hebrang Grčić (ur.) *Otvorenost u znanosti i visokom obrazovanju* (str. 181–197). Zagreb: Školska knjiga.
- CAPLAN, P. (2003). *Metadata fundamentals for all librarians*. Chicago: American Library Association.
- CHAN, L. M. i ZENG, M. L. (2006). Metadata interoperability and standardization: a study of methodology part 1. *D-Lib Magazine*, 12 (6), <http://www.dlib.org/dlib/june06/chan/06chan.html> (01-02-2019).
- CanCore Application Profile*, <http://cancore.athabascau.ca/en/docs.html> (01-02-2019).
- Dublin Core Metadata Initiative – DCMI, <http://dublincore.org> (01-02-2019).
- FLORIDI, L. (2010). *Information: a very short introduction*. Oxford: Oxford Uni-

- versity Press.
- Getty. *Metadata Standards Crosswalk*, http://www.getty.edu/research/publications/electronic_publications/intrometadata/crosswalks.html (05-02-2019).
- GILLILAND, A. J. (2008). Setting the stage. U M. Baca (ur.) *Introduction to metadata*. Los Angeles: Getty Research Institute, <http://www.getty.edu/publications/intrometadata/> (01-02-2019).
- Guidelines for Dublin Core Application Profile*, <http://dublincore.org/documents/profile-guidelines/> (01-02-2019).
- Interoperability. (2019). *WhatIs.com*, <https://whatis.techtarget.com/search/query?q=interoperability> (01-02-2019).
- Interoperate. (2019). *Merriam-Webster's online dictionary*, <https://www.merriam-webster.com/dictionary/interoperate> (01-02-2019).
- MILLER, P. (2000). Interoperability. What is it and why should i want it? *Ariadne*, 24, <http://www.ariadne.ac.uk/issue24/interoperability/> (01-02-2019).
- MOEN, W. E. (2005). *Metadata Interaction, Integration, and Interoperability: MODS, MARC and Metadata Interoperability*. Chicago: American Library Association, <https://www.slideserve.com/geri/metadata-interaction-integration-and-interoperability> (01-02-2019).
- Metadata and Resource Description*. (2019), <https://www.w3.org/Metadata/> (01-02-2019).
- Open Archives Initiative, <http://www.openarchives.org/> (01-02-2019).
- Resource Description Framework (RDF)*, <https://www.w3.org/RDF/> (01-02-2019).
- RILEY, J. (2017). *Understanding metadata: what is metadata, and what is it for?: a primer*, http://groups.niso.org/apps/group_public/download.php/17446/Understanding%20Metadata.pdf (01-02-2019).
- RUST, G. (2005). *Thoughts from a different planet*, https://www.oclc.org/research/events/frbr-workshop/presentations/rust/050502_Godfrey_Rust_FRBR_presentation (01-02-2019).
- TENNANT, R. (2002). MARC Must Die. *Library Journal*, 127 (17), 26–27.
- WILLER, M. (2001). Interoperabilnost sadržaja metapodataka. U M. Willer i T. Katić (ur.) *4. seminar Arhivi, knjižnice, muzeji: mogućnosti suradnje u okruženju globalne informacijske infrastrukture: zbornik radova* (str. 57–72.). Zagreb: Hrvatsko knjižničarsko društvo.

VERY SHORT (AND POPULAR) INTRODUCTION TO METADATA

KEYWORDS:

metadata, metadata scheme, information objects, interoperability, application profiles

ABSTRACT

This paper presents an ontological approach to metadata, the area which currently deals with the description of (mainly) digital information sources, derived by the well-known library area of bibliographic organization and control. The paper tries to answer the questions “What are metadata?”, “What do we describe by them?”, and finally, “What are the benefits?”. Two key processes related to the very core of the metadata concept are explained: the process of creating a metadata description and process of creating a metadata record. On the one hand, library standards for describing sources evolve into metadata schemes for describing information objects of all kinds that can be mapped to other metadata schemes and combined in application profiles. On the other hand, library catalogue records encoded by the machine-readable formats such as RDF evolve into metadata records that, apart from enabling searching and browsing, can now also be harvested (via OAI-PMH) and converted from one format to another. Finally, there is also a review of the metadata interoperability in the network environment, emphasizing the role of metadata within the semantic web, whose development is also the basis for the further metadata development.