

Iva Melinščak Zlodi  
Knjižnica Instituta "Ruđer Bošković", Zagreb  
imelinsc@irb.hr

Martina Melinščak  
Hrvatski povijesni muzej, Zagreb  
martina.melinscak@zg.htnet.hr

## OAI@AKM

### *Sažetak*

Kad se prije pet godina govorilo o interoperabilnosti digitalnih zbirki, svi su neizostavno spominjali protokol Z39.50. Ako pak danas razgovaramo o interoperabilnosti, obvezno ćemo spomenuti *Protokol za pobiranje metapodataka Inicijative za otvorene arhive* (Open Archives Initiative Protocol for Metadata Harvesting - OAI-PMH). To ipak ne znači da će OAI-PMH "istisnuti" Z39.50, jer svaki od njih ima svoja polja primjene koja se samo djelomično preklapaju.

Z39.50 je vrlo sofisticiran (i zahtjevan) standard koji omogućuje distribuirano pretraživanje baza podataka. To znači da se informacijski upit s jednog mjesta istodobno šalje u različite baze podataka, a dobiveni odgovori/rezultati pronalaženja (metapodaci) obrađuju i prikazuju na mjestu upita. S druge strane, koncept interoperabilnosti kakvog omogućuje OAI-PMH-a počiva na mehanizmu pobiranja metapodataka i izgradnje baze podataka u kojoj su pohranjeni pobrani metapodaci koji se potom pretražuju. OAI-PMH dijeli svijet na davatelje podataka (repozitorije) i pružatelje usluga. Davatelji podataka izlažu metapodatke o građi koju sadrže, a pružatelji usluga ih pobiru, organiziraju i grade na temelju njih nove usluge. Shema metapodataka koju OAI-PMH nužno zahtijeva jest DC (Dublin Core Metadata Element Set) iskazan u XML-u, ali je uporaba ostalih shema metapodataka također moguća (i poželjna). Sam je protokol lako primjenljiv i nezahtjevan. Izvorno, protokol je nastao u kontekstu reformiranja procesa znanstvene komunikacije. Trebalo je omogućiti interoperabilnost unutar mreže elektroničkih arhiva znanstvenih radova. Radilo se o podacima koji nisu pretražljivi i dohvatljivi pomoću postojećih mrežnih pretraživača, već se nalaze skriveni u tzv. "nevidljivom webu".

Vrlo se brzo pokazalo da je protokol primjenljiv ne samo u kontekstu znanstvene komunikacije, već i u brojnim i raznolikim projektima vezanim za knjižnice, muzeje i arhive. Među onima koji primjenjuju protokol nalazimo danas, osim tematskih i institucijskih arhiva

znanstvenih radova, i predmetne pristupnike, knjižnične WebPAC-e, zbirke elektroničkih magistarskih i doktorskih radova, izdavače, muzejske i arhivske zbirke. Zastupljeni sadržaji mogu biti tekstualni dokumenti (potpuni ili sažeci), slike, video, statistički podaci, softver i dr. No, iako je popis davatelja podataka koji primjenjuju protokol svakim danom sve duža, pravi test primjenljivosti i svrhovitosti protokola predstavlja razvoj što većeg broja pružatelja usluga u svim gore spomenutim područjima.

*Ključne riječi:* interoperabilnost, OAI-PMH, Protokol za pobiranje metapodataka Inicijative za otvorene arhive, pobiranje metapodataka

### *Summary*

Five years ago, every discussion on interoperability issues in digital collections was at the same time the discussion on Z39.50. While if we talk about interoperability today, we will certainly mention *Open Archives Initiative Protocol for Metadata Harvesting* (OAI-PMH). Still, it does not mean that OAI-PMH could “replace“ Z39.50, because each of those has its own area of application which only partially overlap.

Z39.50 is very sophisticated (and demanding) standard which enables distributed or “broadcasting“ searching. It means that information query is simultaneously sent from one place to multiple databases. On the other hand, interoperability concept which is enabled by OAI-PMH is based on metadata harvesting mechanism. OAI-PMH divides the world on “data providers“ and “service providers “. Data providers expose metadata about materials they contain, while service providers harvest those metadata, organize them and build new services on top of them. Metadata schema that OAI-PMH requires as obligatory is Dublin Core Metadata Element Set (coded in XML), but other metadata schemas employment is also possible (and suggested). The protocol itself is easily implemented and “low-barrier “. Originally, it is born in the context of reforming scholarly communication. Interoperability in the network of eprint archives needed to be established. The data that those archives contained were irretrievable via existing search engines; they were hidden inside “the deep web“.

Very soon it became clear that the protocol is useful not only in the context of scholarly communication, but also in numerous and various projects in libraries, museums and archives. Among implementers of OAI-PMH we can find, besides subject based and institutional eprint archives, subject gateways, library WebPACs, electronic theses and dissertations collections, publishers, museum and archives collections. Materials contained can be text (abstracts or full texts), images, video, raw data, software etc. But, although the list of data providers that have

implemented the protocol is longer every day, the real test of applicability and usefulness of the protocol is development of new innovative service providers, in all the above mentioned areas.

*Keywords:* interoperability, OAI-PMH, Open Archives Initiative Protocol for Metadata Harvesting, metadata harvesting

## 1 Uvod

O *Protokolu za pobiranje metapodataka Inicijative za otvorene arhive* (Open Archives Initiative Protocol for Metadata Harvesting – OAI-PMH) bilo je riječi i na šestom AKM-seminaru, ali spominjao se isključivo u kontekstu novih stremljenja u znanstvenoj zajednici, u svojstvu pomagala koje će pomoći bržem i efikasnijem protoku slobodno dostupnih digitalnih znanstvenih informacija. Iako je to zaista bila prvotna namjena protokola, njegova je svrhovitost uskoro prepoznata i u drugim kontekstima, te u drugim zajednicama koje se također brinu o digitalnoj građi. Kao što kaže jedan od tvoraca protokola, Herbert Van de Sompel, “u početku je namjera *Inicijative za otvorene arhive* bila promijeniti komuniciranje u znanosti. Osnovna je zamisao bila povećati utjecaj komunikacije putem *preprinta* na taj način da se njihovi repozitoriji učine interoperabilnima. Kasnije je naš rad postao općenitiji. Repozitoriji *preprinta* nisu bili jedini otoci na mreži. To su bile i knjižnične digitalne zbirke. Postojala je jasna potreba za interoperabilnošću na razini pronalaženja izvora za sve vrste građe.”<sup>1</sup>

Namjera je ovog rada opisati razloge i okolnosti nastanka OAI-PMH-a, način njegovog funkcioniranja te moguće primjene, posebno u okviru muzeja, knjižnica i arhiva.

## 2 Povijest

*Inicijativa za otvorene arhive* (Open Archives Initiative – OAI) osnovana je na sastanku u Santa Feu u listopadu 1999. godine. U to se vrijeme smatralo da će globalna mrežna povezanost omogućiti mnogo efikasniju i bržu diseminaciju znanstvenih rezultata te slobodnu dostupnost znanstvene literature. Kao glavni i najetabliraniji primjer nove paradigme objavljivanja znanstvenih rezultata prepoznat je *Los Alamos Physics Eprint Archive* koji je još 1991. pokrenuo Paul Ginsparg u Nacionalnom laboratoriju u Los Alamosu. Slični arhivi pojavili su se i u drugim

---

<sup>1</sup> Van de Sompel, Herbert. Developing new protocols to support and connect digital libraries. 2003. <http://www.oclc.org/news/e-newsletter/n261/interview.htm> (2003-09-05)

znanstvenim područjima, npr., *CogPrints*, *NCSTRL*, *RePEC* što je potaklo raspravu o načinu njihove interoperabilnosti. Na spomenutom sastanku radne grupe u Santa Feu, koja se u to vrijeme još nazivala Sveopća *preprint* usluga (Universal Preprint Service – UPS), smatralo se da će interoperabilnost elektroničkih arhiva osigurati njihov povećani utjecaj time što će potaknuti razvoj inovativnih usluga unutar arhivske mreže, poput povezivanja citata, stvaranja jedinstvenih sučelja za pretraživanje, sinkronizacije arhiva, usluge tekućih upozorenja i tsl. Međutim, ostvarivanje interoperabilnosti zahtijevalo je rješavanje određenih tehničkih i organizacijskih problema, kao što su jedinstveno nazivlje, formati metapodataka, modeli dokumenata i pristupni protokoli. Sudionici sastanka znali su da je ključno pronaći jednostavno rješenje s malo prepreka, pa su za pristupne protokole odabrali pobiranje metapodataka, a ne distribuirano pretraživanje protokola Z39.50,<sup>2</sup> odnosno ISO 23950.<sup>3</sup> (O razlikama u funkcioniranju ta dva protokola i razlozima odabira modela pobiranja metapodataka bit će riječi kasnije u radu.) Na taj bi način elektronički arhivi mogli na jednostavan način, putem otvorenog sučelja, izložiti svoje metapodatke alatima za pobiranje, a pobrani bi metapodaci potom postali osnova za izgradnju i razvoj usluga s dodanom vrijednošću. Rezultat sastanka bio je niz tehničkih i organizacijskih dogovora ujedinenih u dokumentu nazvanom *Santa Fe Convention*. Uskoro nakon njegovog objavljivanja u veljači 2000., dokument je izazvao velik interes i izvan *eprint*-zajednice, naročito među knjižničarima, muzealcima i izdavačima. U skladu s tako široko izazvanim zanimanjem, iznova su razmotrene odluke iz 1999. godine, a poslanje *Inicijative* izmijenjeno: “*Inicijativa za otvorene arhive* razvija i promiče standarde interoperabilnosti kojima je cilj omogućiti efikasnu diseminaciju sadržaja. *Inicijativa za otvorene arhive* ima svoje korijene u naporu da se unaprijedi pristup elektroničkim arhivima kao sredstvima povećanja dostupnosti znanstvene komunikacije. [...] Temeljni tehnološki okvir i standardi koji su razvijeni u tu svrhu, ipak, ne ovise ni o vrsti sadržaja ni o ekonomskim mehanizmima u kojima nalazimo taj sadržaj, i obećavaju da će imati daleko veću relevantnost u otvaranju pristupa čitavom nizu digitalnih sadržaja.”<sup>4</sup>

Godine 2000. objavljena je alfa-verzija protokola pod nazivom *Open Archives Initiative Protocol for Metadata Harvesting* – OAI-PMH, a u njenom testiranju sudjelovali su predstavnici

---

<sup>2</sup> ANSI/NISO Z39.50-2003 Information retrieval (Z39.50) : Application service definition and protocol specification. <http://lcweb.loc.gov/z3950/agency/Z39-50-2003.pdf>

<sup>3</sup> ISO 23950-1998 Information documentation -- Information retrieval (Z3950) -- Application service definition and protocol specification. Genève : International Organization for Standardization, 1998.

<sup>4</sup> Lagoze, Carl; Herbert Van de Sompel. The Open Archive Initiative : building a low-barrier interoperability framework. 2001. <http://www.openarchives.org/documents/oai.pdf> (2002-04-08)

eprint-zajednice, muzeja, znanstvenih knjižnica, izdavača i sabirača metapodataka.<sup>5</sup> Verzija protokola 1.0 (beta) objavljena je 2001., a verzija 2.0 (za sada posljednja) 2002. godine.

### 3 Kako funkcionira OAI-PMH?

Sudionike u komunikaciji pomoću OAI-PMH-a dijelimo na **davatelje podataka** (*data providers*) i **pružatelje usluga** (*service providers*). Davatelji podataka izlažu metapodatke o građi koju sadrže, a pružatelji usluga pobiru metapodatke koje potom koriste kao osnovu za izgradnju usluga s dodanom vrijednošću.

Davatelje podataka možemo nazvati i “otvorenim arhivima”, ali treba napomenuti da se ovdje izrazi “otvoren” i “arhiv” ne koriste u svojim konvencionalnim značenjima. Otvorenost je u ovom kontekstu ponajprije svojstvo arhitekture sustava i znači da davatelj podataka izlaže metapodatke putem strojnog sučelja čiji je protokol otvoreno dokumentiran. Otvorenost arhiva ne pretpostavlja nikakva pravila u vezi s cijenom podataka ili eventualnim ograničenjima pristupa podacima (“otvoren” ne smijemo izjednačiti sa “slobodan”, “besplatan” ili “neograničen”).<sup>6</sup> Izraz “arhiv” ovdje je također ogoljen od svih implikacija koje ima u arhivskoj zajednici, pa tako OAI-PMH ne implicira dugotrajnu zaštitu i ostale značajke institucije arhiva već isključivo ima značenje spremišta podataka. Osim strojnog sučelja otvorenog za pobiranje metapodataka, davatelj podataka često nudi i korisničko mrežno sučelje. Najčešće je to usluga koju pružaju elektronički arhivi.

Za davatelje usluga češće upotrebljavani sinonim jest “repozitorij,” pa će se taj izraz koristiti u radu usporedo s izrazom davatelj usluga. Ključni su izrazi vezani uz repozitorij:

- izvor (*resource*) - predmet na koji se odnose metapodaci; narav predmeta nije definirana OAI-PMH-om;
- jedinica (*item*) - komponenta repozitorija iz koje metapodaci o izvoru mogu biti diseminirani; ima jedinstveni identifikator;

---

<sup>5</sup> U testiranju tehničkih specifikacija alfa-verzije protokola sudjelovale su sljedeće ustanove: CIMI, Ex Libris, Los Alamos National Laboratory, NASA, OCLC, UKOLN Resource Discovery Network, Virginia Tech, te sedam američkih i britanskih sveučilišta (Cornell, Old Dominion, Illinois Urbana Champaign, North Carolina, Pennsylvania, Southampton i Tennessee).

<sup>6</sup> U svojim začetcima *Inicijativa za otvorene arhive* imala je gotovo identične ciljeve kao i brojne inicijative koje danas prepoznajemo pod zajedničkim nazivom *Pokret za otvoreni pristup* ili *Inicijativa za otvoreni pristup* (Open Access Movement ili Open Access Initiative, odatle i jednaki akronimi). Iako se često radi o srodnim idejama, čiji su zagovornici često iste osobe, a OAI-PMH je važan preduvjet za ostvarenje ciljeva *Inicijative za otvoreni pristup*, važno je razlikovati ih.

- zapis (*record*) - metapodaci o izvoru u jednom od formata metapodataka; uz jednu jedinicu može biti vezano više zapisa (više formata metapodataka);
- identifikator (*identifier*) - jedinstvena oznaka jedinice u repozitoriju;
- skup (*set*) - neobvezna tvorevina koja služi grupiranju jedinica u repozitoriju (i omogućuje selektivno pobiranje metapodataka).<sup>7</sup>

Zapis koji se prenosi OAI-PMH-om sastoji se od tri dijela, a to su zaglavlje (čiji su obvezni elementi jedinstveni identifikator i oznaka datuma posljednje promjene metapodataka - *datestamp*), metapodaci i dio "o" (*about*, podaci o metapodacima, npr., podaci o uvjetima korištenja, autorstvu i sl.).

Metapodaci moraju biti iskazani u XML-u i nekvalificiranom DC-u kao obaveznom zajedničkom formatu. OAI-PMH-, međutim, omogućuje i potiče uporabu drugih shema metapodataka specifičnih za pojedine zajednice (MARC, EAD, MODS itd.) pod uvjetom da se mogu iskazati pomoću XML-a.

Definiranje skupova unutar repozitorija nije obvezno, ali je od velike koristi pružateljima usluga pri pobiranju metapodataka, naročito kad se pobiru metapodaci vrlo velikih repozitorija, jer se pomoću skupova omogućuje selektivno pobiranje. Skupovi se mogu formirati ovisno o vrsti izvora (članak, magisterij, doktorat, poglavlje u knjizi itd.), vrsti dokumenta (tekst, slika, zvučni zapis itd.), pripadnosti ustanovi ili njejoj organizacijskoj jedinici ili, najčešće, pripadnosti nekom znanstvenom području ili predmetu unutar korištene klasifikacije. Pripadnost neke jedinice jednom skupu ne isključuje mogućnost njena pripadanja i drugim skupovima.

Komunikacija između repozitorija i pružatelja usluga odvija se putem HTTP-a. OAI-PMH definira šest različitih vrsta zahtjeva koji se pri pobiranju upućuje repozitoriju, a to su:

- Identify
- ListMetadataFormats
- ListSets
- ListIdentifiers
- ListRecords
- GetRecord

Pri pobiranju nije obvezno upotrijebiti svih šest zahtjeva, ali repozitorij mora moći odgovoriti na svih šest. Odgovori repozitorija kodirani su u XML-u. Primjer jednoga takvog

---

<sup>7</sup> Müller, Uwe. OAI-PMH implementation : tutorial. 2004.  
<http://eprints.rclis.org/archive/00000984/01/tutorial3muller.pdf> (2004-04-03)

odgovora na zahtjev "Identify" možemo vidjeti u prilogu. Zahtjev je postavljen repozitoriju E-LIS, a sam zahtjev izgleda ovako: <http://eprints.rclis.org/perl/oai2?verb=Identify>.

Selektivno pobiranje podataka moguće je na temelju nekoliko kriterija:

- prema skupovima: pobiranje metapodataka koji pripadaju nekom skupu problematično je zbog toga što se unutar pojedinih repozitorija skupovi neujednačeno oblikuju jer nema općeg pravila prema kojem se to čini; uočavaju se nastojanja razvijanja ujednačenih postupaka zajedničkih nekim zajednicama repozitorija odnosno oblikovanja skupova u suradnji i dogovoru određenog repozitorija i sabirača;
- prema datumu posljednje promjene: pobiru se samo oni zapisi koji su dodani ili izmijenjeni nakon određenog datuma pa sabirač ne mora svaki put iznova pobirati sve postojeće jedinice, nego samo ažurira svoju bazu promjenama koje su se dogodile nakon njegova posljednjeg posjeta i
- prema formatu metapodataka.

Često se naglašava da je OAI-PMH mehanizam s vrlo malo prepreka. Preciznije je, međutim, ustanoviti da se ta tvrdnja odnosi samo na jednog člana para "davatelja podataka-pružatelja usluga". Vrlo je jednostavno postati repozitorij OAI-a i ponuditi svoj digitalni sadržaj. Upravo je to i bio cilj *Inicijative za otvorene archive*: omogućiti svakom vlasniku digitalnog sadržaja da, bez obzira na njegovu spretnost u rukovanju tom građom ili na veličinu digitalnih zbirki, uz što manje truda izloži svoj sadržaj i omogući njegovo korištenje od strane pružatelja usluga, a onda posredno i od strane krajnjih korisnika.<sup>8</sup> Međutim, sam protokol ne utječe ni najmanje na način na koji će pružatelj usluga jednom pobrane metapodatke iskoristiti, koliku će im vrijednost dodati, koliko će biti inventivan pri osmišljavanju nove usluge i, na kraju, na koji će način uslugu pružiti krajnjem korisniku.

Usporedimo sada pristupne protokole pobiranja metapodataka OAI-PMH-a i distribuirano pretraživanje protokola Z39.50. U distribuiranom modelu Z39.50 korisnik (izvor) šalje upite na udaljene lokacije gdje lokalni Z39.50 poslužitelji (ciljevi) pretražuju baze podataka i vraćaju korisniku (izvoru) odgovore (strukturirane (meta)podatke) na postavljeni upit.<sup>9</sup> Pri upotrebi OAI-PMH-a sabirač pružatelja usluga pobire (meta)podatke s udaljene lokacije davatelja podataka i

---

<sup>8</sup> Velik je i broj otvoreno dokumentiranih softverskih rješenja koja omogućuju relativno jednostavnu implementaciju protokola odnosno izgradnju repozitorija koji podržavaju protokol. Najpoznatiji i najčešće korišteni primjeri su *Eprints*, nastao na Sveučilištu u Southamptonu i *Dspace*, nastao u suradnji Tehnološkog instituta u Massachusettsu i tvrtke Hewlett Packard.

<sup>9</sup> Vidjeti npr.: Miller, Paul. Z39.50 za sve. // *Vjesnik bibliotekara Hrvatske* 43, 1-2(2000), 69-83.

pohranjuje ih u svoju bazu podataka koja se potom pretražuje. Distribuirano pretraživanje postavlja veće zahtjeve pred udaljene poslužitelje, ali je jednostavnije za sustav koji šalje upit (korisnik/izvor), dok je kod pobiranja metapodataka obratno.

Preglednu usporedbu sličnosti i razlika između Z39.50 i OAI-PMH-a daje T. Cole:<sup>10</sup>

	<b>Z39.50</b>	<b>OAI</b>
sadržaj (predmeti)	distribuirani	<b>distribuirani</b>
pogled na svijet	bibliografski	<b>bibliografski</b>
predstavljanje predmeta	davatelj podataka	<b>davatelj podataka</b>
pretraživanje je	distribuirano	<b>centralizirano</b>
pretraživanje obavlja	davatelj podataka	<b>pružatelj usluga</b>
pretraženi su metapodaci	ažurni	<b>Zastarjeli</b>
semantičko mapiranje	kod pretraživanja	<b>kod dostave metapodataka</b>

Razlozi zbog kojih su sudionici inicijalnog sastanka u Santa Feu dali prednost pobiranju metapodataka višestruki su.<sup>11</sup> Iskustva postojećih digitalnih knjižnica pokazala su da je distribuirano pretraživanje unutar mreža s velikim brojem udaljenih poslužitelja nepouzdana, jer je uspješnost čitavog sustava ograničena uspješnošću najsporijeg poslužitelja. Prisutan je i jezični problem pri postavljanju upita: sintaksa varira između različitih poslužitelja. Postupak spajanja rezultata pristiglih od raznih poslužitelja i njihova rangiranja također se u praksi pokazao problematičnim, a isto je tako teško i na temelju rezultata distribuiranog pretraživanja graditi sučelja za pregledavanje (*browse interface*).

OAI-PMH daleko je jednostavnije implementirati na razini davatelja podataka nego što je to slučaj s protokolom Z39.50. Ipak, pri usporedbi ova dva modela interoperabilnosti treba imati na umu da oni nisu u potpunosti konkurentni i da svaki ima svoje područje primjene. Na primjer, Z39.50 nezamjenljiv je pri pretraživanju knjižničnih kataloga gdje je ažurnost i preciznost rezultata od ključne važnosti. U budućnosti ćemo vjerojatno imati priliku vidjeti sve veći broj usluga koje kombiniraju oba pristupa.

<sup>10</sup> Cole, Timothy W. Using OAI-PMH to aggregate metadata describing cultural heritage resources. [http://dli.grainger.uiuc.edu/Publications/TWCCole/ALA2003OAI/ALA2003\\_OAI.ppt](http://dli.grainger.uiuc.edu/Publications/TWCCole/ALA2003OAI/ALA2003_OAI.ppt) (2003-10-20)

<sup>11</sup> Müller, Uwe. OAI-PMH implementation : tutorial. 2004.

<http://eprints.rclis.org/archive/00000984/01/tutorial3muller.pdf> (2004-04-03)



#### 4 Primjena OAI-PMH-a

Među korisnicima OAI-PMH-a, kako među davateljima podataka tako i među pružateljima usluga, ponajčešće nalazimo pripadnike *eprint*-zajednice i baštinske ustanove (arhive, knjižnice i muzeje), ali i izdavače i mrežne usluge poput predmetnih pristupnika. Iako bi bilo vrlo zanimljivo vidjeti zastupljenost protokola u pojedinim gore navedenim domenama, takvi nam pokazatelji nisu poznati na globalnoj razini. Međutim, zahvaljujući projektu *Open Archives Forum* poznati su nam približni pokazatelji o uporabi OAI-PMH-a i njegovoj zastupljenosti u pojedinim domenama u Europi.

##### 4.1 Europski projekt *Open Archives Forum*

Dvogodišnji projekt *Open Archives Forum* pokrenut je 2002. s ciljem prikupljanja informacija o korištenju OAI-PMH-a u Europi, razmjene iskustava korisnika, te osiguravanja podrške postojećim davateljima podataka i pružateljima usluga, kao i onima u izradi. U okviru projekta provedeno je istraživanje o korištenom softveru pri primjeni OAI-PMH-a, troškovima njegove implementacije i sadržajima repozitorija te dobivanja uvida u iskustva i očekivanja unutar različitih zajednica i različitih zemalja.. Preliminarni rezultati projekta objavljeni su početkom 2003.<sup>12</sup>

Zanimljivo je pogledati podjelu sudionika ispitivanja u odnosu na pripadnost pojedinim domenama. Većina korisnika protokola pripada knjižničnoj zajednici (29%), zatim arhivskoj i znanstvenoj *eprint*-zajednici (oboje po 12%). Muzejskoj zajednici pripada 5%, a ostalim zajednicama 10%. Čak 32% nije specificiralo svoju pripadnost zajednici.

Istraživanje je, uz već iskusne korisnike protokola, obuhvatilo i one davatelje podataka koji razmatraju njegovo implementiranje (od 33 sudionika na temelju čijih odgovora su izvedeni preliminarni rezultati, njih 11 još nisu implementirali protokol). Davatelji podataka razvrstani su prema sljedećoj tipologiji: predmetni pristupnici (*subject gateways*) u koje su uključeni i tematski elektronički arhivi, što možda nije najsretnije rješenje, institucionalni repozitoriji, knjižnični

---

<sup>12</sup> Dobratz, Susanne; Birgit Matthaei. Open archives activities and experiences in Europe : an overview by the Open Archives Forum. // D-lib magazine 9, 1(2003).  
<http://www.dlib.org/dlib/january03/dobratz/01dobratz.html> (2003-03-02)

WebPAC-i, komercijalni izdavači i posrednici, arhivi različitih medija i muzeji. Pružatelji usluga analizirani su s aspekta pružanja usluga pretraživanja i proširenih usluga.

Rezultati ispitivanja potvrđuju ranije izrečene tvrdnje o jednostavnosti implementacije protokola: većina implementacija obavljena je unutar tri mjeseca i uglavnom je bio dovoljan angažman samo jednog programera. Repozitoriji najčešće sadrže samo metapodatke ili potpune tekstove, a rjeđe slikovnu i video građu, statističke podatke i softver. Sadržajno prevladavaju različiti oblici znanstvenih publikacija (disertacije, *preprinti*, članci iz časopisa, izlaganja na kongresima itd.), dok ostale sadržaje poput knjižničnih kataloga nalazimo rijetko. Najzastupljeniji je format metapodataka DC, a zatim MARC.

## 4.2 Primjeri primjene OAI-PMH-a

Sljedeći niz primjena OAI-PMH-a nije iscrpan popis projekata koji su implementirali protokol, pa čak niti izbor onih najrelevantnijih. Svrha ovakva odabira i opisa projekata ponajprije je pokušaj slikovita prikazivanja raznolikosti namjena kojima OAI-PMH može služiti.<sup>13</sup>

### 4.2.1 Davatelji podataka

#### 4.2.1.1 *Eprints in Library and Information Science - E-LIS*

E-LIS<sup>14</sup> je repozitorij slobodno dostupnih tekstova iz područja knjižničarstva i informacijskih znanosti, a namjera mu je okupiti različite vrste objavljenih i neobjavljenih dokumenata iz tog područja, iz svih zemalja i na svim jezicima. Repozitorij je pokrenut 2002. godine, a u studenom 2003. sadržavao je oko 300 radova, uglavnom iz talijanskog i španjolskog govornog područja. U pozadini je E-LIS-a softver *Eprints*, otvoreno dokumentirani softver nastao na Sveučilištu u Southamptonu, a koji trenutno koristi najveći broj elektroničkih repozitorija u svijetu. Osim što omogućuje laganu izgradnju strojnog sučelja repozitorija primjenom OAI-PMH-a, *Eprints* nudi i korisničko mrežno sučelje, preko kojeg krajnji korisnici mogu pristupiti metapodacima i građi koju repozitorij sadrži. Tako svaki korisnik E-LIS-a može

---

<sup>13</sup> Među odabranim primjerima zanemareni su davatelji podataka i pružatelji usluga iz arhivske zajednice. Detaljan pregled moguće primjene protokola i opis dosad razvijanih primjena u arhivima vidjeti u: Prom, Christopher J. Reengineering archival access through the OAI protocols. // *Library hi tech* 21, 2(2003), 199-209 i MacKenzie, George; Göran Kristiansson. How real archivists can learn to love OAI : a review of the potential for using the Open Archive Initiative Protocol for Metadata Harvesting in conventional archives. A report for the Open Archives Forum. 2003.

[http://www.oaforum.org/otherfiles/oaf\\_d44\\_cser2\\_kenzie\\_krist.pdf](http://www.oaforum.org/otherfiles/oaf_d44_cser2_kenzie_krist.pdf) (2003-09-03)

<sup>14</sup> <http://eprints.rclis.org>

pretraživati i pregledavati repozitorij, prema godini i prema područjima iz JITA klasifikacije<sup>15</sup>, a registrirani korisnik može i pohranjivati vlastite dokumente ili se pretplatiti za službu tekućih upozorenja.

#### 4.2.1.2 Kongresna knjižnica

Čini se da su knjižnice prepoznale svrhovitost OAI-PMH-a ponajprije u izlaganju metapodataka o građi u specijalnim zbirka, i to u pravilu u zbirka digitalne/digitalizirane građe (npr., digitalnih slika, notnih zapisa, elektroničkih ocjenskih radova). Tek u rijetkim slučajevima knjižnice putem OAI-PMH-a osiguravaju pristup katalogima cjelovitog fonda.

Među najranijim korisnicima protokola nalazimo Kongresnu knjižnicu<sup>16</sup> koja je pomoću OAI-PMH-a omogućila integriranje svoje digitalizirane povijesne zbirke u druge usluge. Otkad je zbirka *American Memory* 1994. godine stavljena na mrežu, izazvala je velik interes pa su uskoro počeli stizati zahtjevi različitih organizacija koje su željele uključiti čitavu zbirku ili njene veće dijelove u vlastite online izvore. Takvi su zahtjevi predstavljali opterećenje za Kongresnu knjižnicu. Osim toga, za većinu ne-knjižničnih korisnika, MARC zapisi nudili su previše informacija i, zapravo, predstavljali smetnju. Stvaranjem repozitorija OAI-a za svoje povijesne zbirke Kongresna je knjižnica svakome omogućila preuzimanje zapisa bez njezina posredovanja. Izradom zapisa za građu u DC-u Knjižnica je ujedno olakšala novim pružateljima uslugu korištenje prikupljenih metapodataka. Danas Kongresna knjižnica svojim repozitorijem OAI-a nudi preko 120.000 zapisa (karata, filmova, nota, plesnih priručnika, fotografija) iz 21 zbirke.<sup>17</sup>

#### 4.2.1.3 Internet Archive

*Internet Archive*<sup>18</sup> implementirao je OAI-PMH radi sudjelovanja u razmjeni informacija s ostalim digitalnim knjižnicama. U repozitoriju su raspoloživi metapodaci za više od 1.700 filmova, 8.500 tekstova (prikupljenih iz različitih projekata, uključujući projekte *Gutenberg*, *Million Books* i *Open Source Texts*), 3.700 koncerata i 8.400 naslova softvera.

---

<sup>15</sup> Naziv klasifikacije je akronim osobnih imena autora: *Jose Manuel Barrueco Cruz*, *Imma Subirats Coll*, *Thomas Krichel* i *Antonella De Robbio*

<sup>16</sup> <http://lcweb2.loc.gov/ammem/ammemhome.html>

<sup>17</sup> Arms, Caroline R. Available and useful : OAI at the Library of Congress. // *Library hi tech* 21, 2(2003), 129-139.

<sup>18</sup> <http://www.archive.org>

#### 4.2.2 Pružatelji usluga

##### 4.2.2.1 Networked Digital Library of Theses and Dissertations - NDLTD

NDLTD<sup>19</sup> je međunarodni skupni katalog doktorskih, magistarskih i diplomskih radova u elektroničkom obliku, u okviru kojeg nalazimo i *Electronic Thesis/Dissertation OAI Union Catalog*. Radi se o usluzi izgrađenoj na temelju metapodataka pobranih putem OAI-PMH-a iz otvorenih arhiva koji prikupljaju ocjenske radove u elektroničkom obliku. Pobrani metapodaci pohranjeni su u skupnom katalogu koji je opremljen korisničkim sučeljem za pretraživanje i pregledavanje, ali i strojnim sučeljem koje jednom pobrane metapodatke izlaže putem OAI-PMH-a drugim sabiračima. OCLC održava skupni katalog, a Virginia Tech korisničko sučelje.

##### 4.2.2.2 Resource Discovery Network - RDN

U okviru RDN-a,<sup>20</sup> danas najveće i najetabliranije mreže predmetnih pristupnika za pojedina područja znanosti, započela je primjena OAI-PMH-a kao mehanizma za razmjenu metapodataka između pojedinih pristupnika, a u svrhu izgradnje usluga za pronalaženje izvora iz različitih područja. Svaki pristupnik unutar mreže RDN mora podržavati OAI-PMH, a obavezan je zajednički format zapisa nekvalificirani DC, unatoč tome što je u njega nemoguće potpuno prevesti postojeće bogate zapise RDN-a. Mreža RDN osigurava strojno sučelje za pristup metapodacima, međutim, metapodatkovni zapisi "otvoreni" su ovdje samo unutar RDN mreže.<sup>21</sup>

##### 4.2.2.3 Australian Museums and Galleries On Line - AMOL Open collections

AMOL<sup>22</sup> je mrežno mjesto izgrađeno s namjerom da pomogne australskim baštinskim ustanovama pri pružanju informacija o svojim zbirkama svjetskoj publici. Putem baze podataka *Open Collections* osiguran je pristup zapisima o predmetima iz čitave Australije. Baza podataka koja trenutno sadrži 500.000 zapisa iz 92 različite zbirke služi za pronalaženje izvora i osnovnih informacija o muzejskim predmetima. Baza podataka nastaje pobiranjem metapodataka uložениh

---

<sup>19</sup> <http://www.ndltd.org>

<sup>20</sup> <http://www.rdn.ac.uk>

<sup>21</sup> Powell, Andy. An OAI approach to sharing subject gateway content. (s.a.). <http://www.rdn.ac.uk/publications/www10/oaiposter.pdf> (2003-10-16)

<sup>22</sup> <http://amol.org.au>

u HTML-dokumente, ali u tijeku je prijelaz na OAI-PMH i XML, zbog jednostavnijeg dijeljenja i veće mogućnosti korištenja metapodataka u nove svrhe.<sup>23</sup>

#### 4.2.2.4 *The University of Illinois Open Archives Initiative Metadata Harvesting Project*

Dvogodišnjem projektu Sveučilišta u Illinoisu ciljevi su bili razviti alate za pobiranje metapodataka i izgradnja usluge za pretraživanje prikupljenih metapodataka iz područja kulturne baštine. Rezultat je projekta pristupnik *UIUC Digital Gateway to Cultural Heritage Materials*<sup>24</sup> kojim se može pretraživati više od dva milijuna metapodataka pobrana s 39 davatelja podataka. Među davateljima podataka nalazimo muzeje, akademske knjižnice i različite digitalne zbirke.

#### 4.3.6 *OAIster*

*OAIster*<sup>25</sup> je projekt Sveučilišta u Michiganu kojem je cilj izgraditi jednostavno pretražljivu zbirku slobodno dostupnih akademskih digitalnih izvora. U studenom 2003. godine u *OAIsteru* se moglo pretraživati oko dva milijuna zapisa pobranih s 239 davatelja podataka. Radi se o izvorima koji su prethodno bili dio “nevidljivog weba” i teško dostupni korisniku. Načelo koje ističu u *OAIsteru* nepostojanje je slijepih ulica (*no dead ends*), jer korisnik kao rezultat pretraživanja neće dobiti samo informaciju o izvoru (metapodatkovni zapis), nego će ga pretraživanje uvijek dovesti do čitavog dokumenta.

#### 4.3.7 *Citebase*

*Citebase*<sup>26</sup> omogućuje pretraživanje slobodno dostupnih potpunih tekstova znanstvenih radova, pri čemu je moguće rezultate rangirati prema nekoliko kriterija: datumu stvaranja, datumu posljednje promjene, citiranosti rada, citiranosti rada u nekoj godini, citiranosti autora, broju pristupanja nekom radu ili broju pristupanja nekom autoru. Također je omogućena citatna analiza te navigacija kroz literaturu putem citatnih veza.

---

<sup>23</sup> Dewhurst, Basil. Enabling interoperability : Australian Museums Online (AMOL) & the Open Archives Initiative Protocol for Metadata Harvesting. [http://www.cimi.org/wg/oai/MW\\_2002\\_OAI\\_AMOL.pdf](http://www.cimi.org/wg/oai/MW_2002_OAI_AMOL.pdf) (2003-10-15)

<sup>24</sup> <http://oai.grainger.uiuc.edu>

<sup>25</sup> <http://oaister.umld.umich.edu>

<sup>26</sup> <http://citebase.eprints.org>

*Citebase* je razvijen u okviru projekta *Open Citation* na Sveučilištu u Southamptonu. Za sada je tek u eksperimentalnoj fazi i obuhvaća pobiranje radova samo iz repozitorija *arXiv.org*.

## 5 Zaključak

Osnovni je cilj *Inicijative za otvorene arhive* rješavanje pitanja interoperabilnosti digitalnih zbirki, a OAI-PMH predstavlja samo dio rješenja. Naime, treba naglasiti da je postizanje interoperabilnosti informacijskih sustava složeno pitanje koje, prema C. R. Arms,<sup>27</sup> treba rješavati na tri razine: tehničkoj, sadržajnoj i organizacijskoj. Na tehničkoj razini problem je uvelike riješen objavljivanjem OAI-PMH-a i razvojem velikog broja softverskih rješenja koja pomažu pri njegovoj implementaciji. Ipak, i na ovoj razini za očekivati je pojavljivanje sve većeg broja novih specifičnih alata. Primjeri za to su ZMARCO i DP9. ZMARCO je davatelj podataka koji omogućuje da se MARC zapisi dostupni putem Z39.50 poslužitelja relativno jednostavno učine dostupnim i putem OAI-PMH-a. DP9 posrednički je otvoreno dokumentirani softver koji tražilicama poput Googlea omogućuje indeksiranje repozitorija OAI-a, a da same ne podržavaju OAI-PMH. Na taj je način omogućeno indeksiranje “nevidljivog weba” sadržanog unutar repozitorija OAI-a.

Daleko su složeniji zahtjevi koji se u cilju postizanja interoperabilnosti postavljaju na sadržajnom i organizacijskom planu. Davateljima podataka još uvijek predstoje veliki naponi oko usvajanja novih shema metapodataka koje će na prikladniji i bogatiji način opisivati posebne vrste građe, kao i ujednačivanja i standardiziranja vrijednosti metapodatkovnih elemenata korištenjem klasifikacijskih i predmetnih sustava te kontroliranog nazivlja. No, s obzirom na raznorodnost repozitorija i njihovu geografsku rasprostranjenost, nerealno je očekivati da će takva ujednačivanja uskoro biti ostvarena na globalnoj razini. Vjerojatnije je očekivati pojavu stvaranja interesnih ili regionalnih konzorcija, unutar kojih će se u interakciji repozitorija i pružatelja usluga rješavati i sadržajna pitanja. Stvaranjem i razvojem svojevrsnog tržišta metapodataka, pitanja vlasništva i autorskih prava nad metapodacima (koja su odnedavno u žarištu zanimanja *Inicijative za otvorene arhive*), ali i pitanja jamstva kvalitete metapodataka dobivat će sve veću važnost.

No, pravi test upotrebljivosti i svrhovitosti OAI-PMH-a predstavlja tek razvoj većeg broja što inovativnijih pružatelja usluga.

---

<sup>27</sup> Arms, Caroline R. Nav. dj.

## LITERATURA

Arms, Caroline R. Available and useful : OAI at the Library of Congress. // Library hi tech 21, 2(2003), 129-139.

Cole, Timothy W. Using OAI-PMH to aggregate metadata describing cultural heritage resources. [http://dli.grainger.uiuc.edu/Publications/TWCOLE/ALA2003OAI/ALA2003\\_OAI.ppt](http://dli.grainger.uiuc.edu/Publications/TWCOLE/ALA2003OAI/ALA2003_OAI.ppt) (2003-10-20)

Dewhurst, Basil. Enabling interoperability : Australian Museums Online (AMOL) & the Open Archives Initiative Protocol for Metadata Harvesting. [http://www.cimi.org/wg/oai/MW\\_2002\\_OAI\\_AMOL.pdf](http://www.cimi.org/wg/oai/MW_2002_OAI_AMOL.pdf) (2003-10-15)

Dobratz, Susanne; Birgit Matthaei. Open archives activities and experiences in Europe : an overview by the Open Archives Forum. // D-lib magazine 9, 1(2003). <http://www.dlib.org/dlib/january03/dobratz/01dobratz.html> (2003-03-02)

Lagoze, Carl; Herbert Van de Sompel. The Open Archive Initiative : building a low-barrier interoperability framework. 2001. <http://www.openarchives.org/documents/oai.pdf> (2002-04-08)

MacKenzie, George; Göran Kristiansson. How real archivists can learn to love OAI : a review of the potential for using the Open Archive Initiative Protocol for Metadata Harvesting in conventional archives. A report for the Open Archives Forum. 2003. [http://www.oaforum.org/otherfiles/oaf\\_d44\\_cser2\\_kenzie\\_krist.pdf](http://www.oaforum.org/otherfiles/oaf_d44_cser2_kenzie_krist.pdf) (2003-09-03)

Miller, Paul. Z39.50 za sve. // Vjesnik bibliotekara Hrvatske 43, 1-2(2000), 69-83.

Müller, Uwe. OAI-PMH implementation : tutorial. 2004. <http://eprints.rclis.org/archive/00000984/01/tutorial3muller.pdf> (2004-04-03)

Powell, Andy. An OAI approach to sharing subject gateway content. (s.a.) <http://www.rdn.ac.uk/publications/www10/oaiposter.pdf> (2003-10-16)

Prom, Christopher J. Reengineering archival access through the OAI protocols. // Library hi tech 21, 2(2003), 199-209.

Van de Sompel, Herbert. Developing new protocols to support and connect digital libraries. 2003. <http://www.oclc.org/news/e-newsletter/n261/interview.htm> (2003-09-05)

## **Prilog**