

# Primena OAIS referentnog modela u digitalnim arhivama

## Application of OAIS Reference Model in Digital Archives

Aleksandra Bradić-Martinović, Aleksandar Zdravković, Institut ekonomskih nauka, Beograd

**Apstrakt** – U većini naučnih oblasti kvalitet empirijskog istraživanja značajno zavisi od kvaliteta raspoloživih podataka i mogućnostima istraživača da pristupe postojećim kolekcijama podataka. Skladištenje podataka u digitalnim arhivama, uz mogućnost kontrolisanog pristupa, pokazalo se kao najsigurniji metod za njihovo trajno očuvanje. Na taj način istraživači podatke, dokumentaciju o svojim istraživanjima i rezultate stavljaju na raspolaganje široj naučnoj zajednici. Koristi od digitalnih arhiva su brojne, a među najvažnije spadaju dostupnost podacima u čije su prikupljanje već uložena sredstva, zatim samo-korigujući faktor prilikom objavljivanja naučnih radova, kao i podsticaj za raznolikost ideja i mišljenja. U ovom radu izložen je referentni model OAIS (*Open Archive Information System*), koji predstavlja apstraktni okvir za organizaciju i funkcionisanje informacionog sistema digitalne arhive. Predstavljen je proces nastanka modela i osnovni pojmovi (informacija i okruženje). Detaljno su opisana dva osnovna mehanizma OAIS referentnog modela (funkcionalni i informacioni), uz tok podataka, kao i njihove implikacije na održivost koncepta.

**Ključne reči** – digitalne arhive, OAIS, informacioni paketi, SIP, DIP, AIP.

**Abstract** - In most scientific field quality of empirical research is largely dependent on the quality of available data and the possibility for researchers to access existing data sets. Data storage in a digital archive, with the possibility of controlled access, has proven to be the safest method for their permanent preservation. As an outcome of the effort the data, documentation and results generated in scientific research can be available to the wider scientific community. The benefits of digital archives have been numerous, and the most important are the availability of data in which collection the funds had already been invested, then self-correcting factor when publishing scientific papers, and incentive for diversity of ideas and opinions. In this paper we presented the OAIS reference model (*Open Archive Information System*), which is an abstract framework for the organization and operation of digital archives information system. We introduced the brief history of reference model and basic concepts (information and environment). We also described the two mechanisms of OAIS reference model (functional and informational), the data flow, as well as their implications for the sustainability of the concept.

**Index terms** - digital archives, OAIS, information package, SIP, DIP, AIP.

---

dr Aleksandra Bradić-Martinović – Institut ekonomskih nauka, Zmaj Jovina 12, 11000 Beograd, Srbija (e-mail: abmartinovic@ien.bg.ac.rs).

ma Aleksandar Zdravković – Institut ekonomskih nauka, Zmaj Jovina 12, 11000 Beograd, Srbija (e-mail: aleksandar.zdravkovic@ien.bg.ac.rs).

## 1. UVOD

Eksplozivni rast informacija u digitalnom obliku predstavlja ozbiljan izazov za sve organizacije koje čuvaju podatke. Informacije i podaci u digitalnoj formi se primarno vezuju za istraživanja u sektoru nauke i obrazovanja, ali su vremenom postali i analitička osnova u procesu donošenja odluka vladinih tela i komercijalnog i neprofitnog sektora. Ovom procesu doprinosi i jednostavan način konvertovanja u digitalni oblik različitih dokumenata, podataka, zapisa, informacionih i kulturnih artefakata, u slučaju da nisu originalno generisani u digitalnoj formi. Rezultat ovog procesa je pojava velikog broja organizacija koje su u stanju da pružaju usluge arhiva ili repozitorijuma sa otvorenim pristupom.

Zbog svega navedenog fenomen digitalnih arhiva i repozitorijuma dobio vrlo široku primenu, ali se u savremenom društvu i mnogo veći broj različitih organizacija sreće sa problemima vezanim za dugoročno očuvanje (*Long-term preservation*) podataka. Loše upravljanje podacima, zavisnost od tehnologije, degradacija medija i promena formata u kojima se čuvaju podaci predstavljaju najveće pretnje u procesu čuvanja podataka. Imajući u vidu da je proces prikupljanja i diseminacije podataka u digitalnoj formi počeo u tehnološki razvijenim zemljama već pedesetih godina dvadesetog veka, problemi očuvanja podataka imaju relativno dugu istoriju. Jedno od rešenja problema predstavljalo je uspostavljanje odgovarajućih standarda u toj oblasti. Inicijalno proces standardizacije u oblasti čuvanja digitalnih podataka započele su dve vrste organizacija. Pre svega to su bili muzeji i razni tipovi arhiva, ali su se ubrzo u proces uključili tzv. centri podataka (*data centers*) koji su ubrzano nastajali kako u nauči, tako i u korporativnom sektoru. Napor da se razvije odgovarajući standard, odnosno u ovom slučaju referentni model, kao apstraktni okvir za uspostavljanje komunikacije u jednom sistemu, došao je u vreme kada su već razvijene odvojene aktivnosti i tokovi podataka. To je podrazumevalo da se aktivnosti povežu i da se obezbedi neometani tok podataka.

Da bi digitalni arhiv ili repozitorijum ostvario svoju svrhu on mora da osigura poverenje i provajdera i korisnika, time što će zadovoljiti sledeće zahteve. a) autentičnost – obezbeđenje pouzdanost da je digitalni sadržaj stvorila ona osoba koja zaista na njega polaže autorsko pravo, a da bi se obezbedila autentičnost koriste se digitalni potpisi i digitalni vodeni pečati; b) integritet – sposobnost da se očuva tačnost i celovitost informacija onemogućavanjem nenamernih i zlonamernih izmena; c) pouzdanost i dostupnost – pouzdanost se odnosi na mogućnost da hardver i softver ne obavljaju operacije prema svojim specifikacijama bez problema i grešaka, a dostupnost predstavlja vreme u kome sistem, program ili komponenta sistema pravilno funkcionišu u odnosu na ukupno vreme predviđeno za njihovo funkcionisanje i d) mogućnost ponovne upotrebe – koja se odnosi na mogućnost prikladnog pretraživanja koje je obezbeđeno u dužem vremenskom periodu, uz upotrebu postojanih identifikatora i održavanja medija i formata aktuelnim. Osim navedenog, čuvanje digitalnog sadržaja mora uključivati i druge elemente infrastrukture, meta podatke, identifikatore i bazu podataka zaduženu za usklađivanje adresa, odnosno likova prema određenom setu podataka.

Predstavljanje koncepta OAIS motivisano je naporima autora da u okviru FP-7 finansiranog projekta SERSCIDA kreiraju arhiv, odnosno servis podataka u društvenim naukama u Srbiji, koji bi u potpunosti bio zasnovan na implementaciji OAIS referentnog modela. Prihvatanje OAIS referentnog sistema i njihovih principa omogućuje zadovoljenje navedenih zahteva, kao i izgradnju funkcionalne digitalne arhive koja obezbeđuje uzajamno poverenje i osiguranje dugoročnog očuvanja digitalnih objekata.

U ovom radu predstavljen je kompletan koncept OAIS. Dat je prikaz okvira referentnog modela i kratak istorijat njegovog nastanka. Definisani su osnovni pojmovi i okruženje arhivskog informacionog sistema, pri čemu je izdvojen eksterni tok podataka. Najveći deo rada posvećen je objašnjenju strukture OAIS i ulozi dva centralna modela: funkcionalnog i informacionog, koji se sastoji od dostavljenog (SIP), arhivskog (AIP) i distribucionog (DIP) informacionog paketa. Nakon detaljne analize upravljačke strukture OAIS, predstavljen je i interni tok podataka. U zaključku je istaknut značaj ovog referentnog modela.

## 2. OKVIR OAIS-a U PROCESU ČUVANJA PODATAKA

*Open Archival Information System*<sup>1</sup> referentni model (OAIS) je osnova za razumevanje celine i kompleksnosti očuvanja elektronskih objekata na duži vremenski rok, uz istovremeno očuvanje njihove autentičnosti. Praktično, u pitanju je apstraktni model usvojen ISO 14721 standardom, koji je razvio *Consultative Committee for Space Data Systems (CCSDS)*, pri američkoj agenciji NASA 1999. godine. Uprkos tome što je apstraktan ovaj model detaljno određuje sve korake u postupku očuvanja informacija, delove sistema elektronskog arhiva, uz razradu njihovih međusobnih veza. Prema tome, on se može koristiti kao primer informacione, odnosno logičke, i funkcionalne povezanosti delova modela prema kome se gradi stvarni sistem. [3] Osnovna prednost OAIS modela nije u pokušajima rešavanja problema, nego u njegovom prepoznavanju. Tu se, pre svega misli na kretanje trendova u informacionim sistemima koji neprekidno menjaju elektronske medije i komunikacione kanale.

OAIS je informacioni sistem za arhive koga čine „organizacija, stručnjaci (HR) i sistemi koji su prihvatili odgovornost da očuvaju informacije i da ih učine dostupnim zajednici korisnika, odnosno ciljanoj zajednici korisnika (*Designated Community*)<sup>2</sup>“. [5] OAIS je odgovoran za digitalne informacije dugoročno, što je dovoljno dugo da se prevladaju problemi vezani za uticaj promene tehnologije, uključujući i podršku za nove medije i formate, ali i promenu korisničkog okruženja. Navedena definicija naglašava i dve primarne funkcije arhiva i repozitorijuma: prva je da čuva informacije, uz uslov obezbeđenja dugoročne postojanosti, a druga je da obezbedi pristup do arhiviranih podataka na način koji je u skladu sa potrebama primarnih korisnika sistema, zajednice korisnika. Prema tome, OAIS mora da: [1]

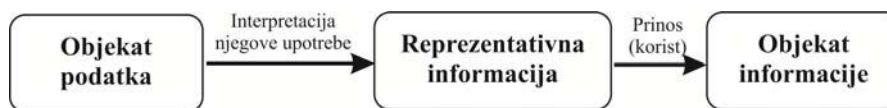
- pregovara u vezi informacija koje će prihvatiti od strane provajdera,
- obezbedi dovoljnu kontrolu nad informacijama u cilju zadovoljenja svih ciljeva dugoročnog očuvanja,
- odredi okvir korisnika (zajednice korisnika),
- proveri da li je sačuvana informacija razumljiva zajednici korisnika, odnosno da li je zajednica korisnika u stanju da koristi informaciju samostalno, bez pomoći onoga ko je obezbedio tu informaciju (provajdera),
- prati odgovarajuće politike i procedure kako bi omogućio da informacije budu postajane u svim predvidivim okolnostima, kao i da se omogući diseminacija autorizovanih kopija sačuvanih informacija u svom originalnom formatu, ili u formatu koji se može pratiti do originala i
- učini da sačuvane i postojeće informacije budu raspoložive zajednici korisnika.

### 2.1. OAIS informacija

Da bi se shvatio koncept funkcionisanja OAIS veoma je važno definisati pojam informacije. Informacija predstavlja bilo koju vrstu znanja koja može biti razmenjena, pri čemu je ona uvek predstavljena kao određeni tip podatka u razmeni. Na primer, informacija koja se nalazi u fajlu izraženom bitovima, koji kada se kombinuju sa reprezentativnom informacijom za te bitove daju razumljivu informaciju onome ko poseduje bazu znanja. Konkretno, podaci o kretanju vrednosti indikatora u tehničkoj analizi predviđanja cena akcija, sačuvani u fajlu dobijaju svoju korist (prinos) samo u slučaju da njima raspolaže osoba koja poznaje metode tehničke analize, odnosno osoba koja ima odgovarajuću bazu znanja.

<sup>1</sup> Odrednica „otvoreni“ (*open*) odnosi se na činjenicu da je ovaj referentni model razvijen učešćem javnog foruma.

<sup>2</sup> *Designated Community* je prihvaćen izraz koji podrazumeva informacionu i arhivsku zajednicu, odnosno grupu potencijalnih korisnika, koji su u stanju da razumeju određeni set podataka. U daljem tekstu ovaj termin će biti označen kao „zajednica korisnika“.



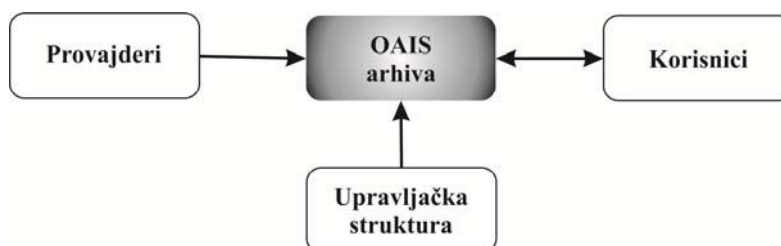
Grafik 1: Generisanje informacije iz podatka  
 Izvor: [5]

OAIS referentni model naglašava značaj očuvanja sadržaja informacije. Kao posledica procesa razvoja digitalne i multimedijalne tehnologije, kao i složene zavisnosti između podataka i tehnologije njihovog predstavljanja mnoge organizacije će se u budućnosti suočiti sa izazovom očuvanja originalnog izgleda sačuvane informacije.

## 2.2. Okruženje OAIS

OAIS ne predstavlja izolovani sistem, već se nalazi u okruženju sa kojim je u interakciji. Najšire posmatrano to okruženje čine stvaraoci digitalnog sadržaja (provajderi), korisnici tog sadržaja i upravljačka struktura samog arhiva. Okruženje OAIS možemo detaljnije posmatrati kao skup sledećih elemenata: [4]

- *Provajder* je jedna ili više osoba ili klijentski sistem koji obezbeđuje informacije koje će biti sačuvane u sistemu, u unapred dogovorenom formatu;
- *Upravljačka struktura* je grupa koja uspostavlja politike, procedure i standarde OAIS, ali ne utiče i ne obavlja njegove redovne funkcije i
- *Korisnici* (zajednica korisnika) su one osobe ili klijentski sistemi koji po osnovu direktne interakcije sa OAIS servisima pronalaze i preuzimaju sačuvane informacije koje su im od značaja, pri čemu uvažavaju ograničenja pristupa pojedinim vrstama informacija.



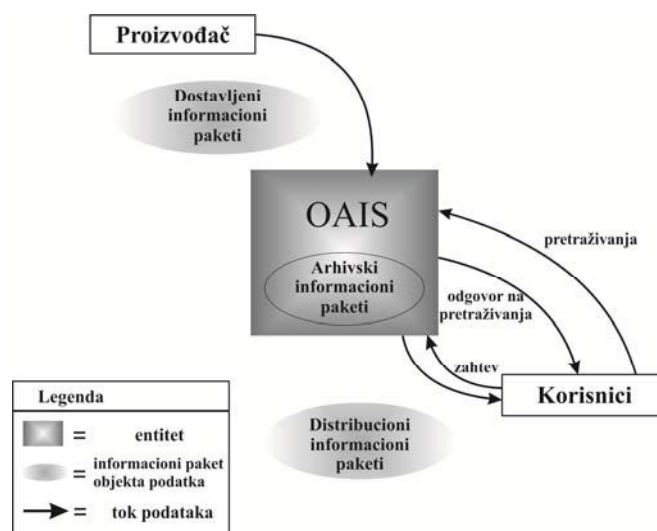
Grafik 2: Okruženje OAIS [5]

Odnosi predstavljeni grafikom 1 mogu biti modifikovani u zavisnosti od konkretne situacije. U pojedinim slučajevima OAIS arhiva može biti i provajder informacija koje se čuvaju u okviru njene infrastrukture ili provajder drugim OAIS arhivama. Takođe, OAIS arhiva može biti i korisnik podataka koji se čuvaju u drugim OAIS arhivama.

Paket informacije konceptualno sadrži dva tipa informacija koje se nazivaju: sadržaj informacije (*Content Information*) i opis za očuvanje informacije (*Preservation Description Information – PDI*). Sadržaj informacije i opis za očuvanje informacije posmatra se kao zatvorena celina, a identifikuje se pomoću paketa informacije. Dobijeni paket je vidljiv pomoću opisne informacije.

Postoje tri varijante informacionog paketa:

- Dostavljeni informacioni paket (*Submission Information Package – SIP*) koje OAIS dobija od provajdera;
- Arhivski informacioni paket (*Archival Information Package – AIP*) koji nastaje iz SIP i
- Distribucionni informacioni paket (*Dissemination Information Package – DIP*) je odgovor OAIS na upućeni zahtev korisnika.



Grafik 3: Eksterni podaci OAIS [5]

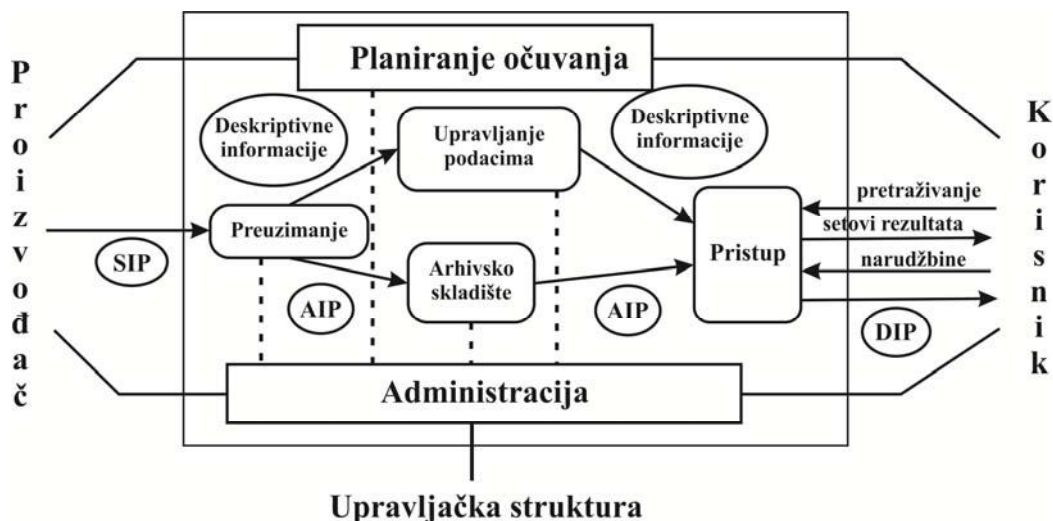
Na grafiku 3. predstavljen je uopštena šema veza koje nastaju između entiteta identifikovanih u okruženju OAIS. Grafik ujedno predstavlja i dijagram toka podataka u razmeni sa okruženjem.

### 3. SADRŽAJ I KONCEPT OAIS

Osnovni mehanizmi OAIS referentnog modela su: funkcionalni model (*Functional Model*) i informacioni model (*Informational Model*). Uopšteno rečeno prvi pokazuje šta OAIS mora da uradi, a drugi šta mora da ima u svojim kolekcijama.

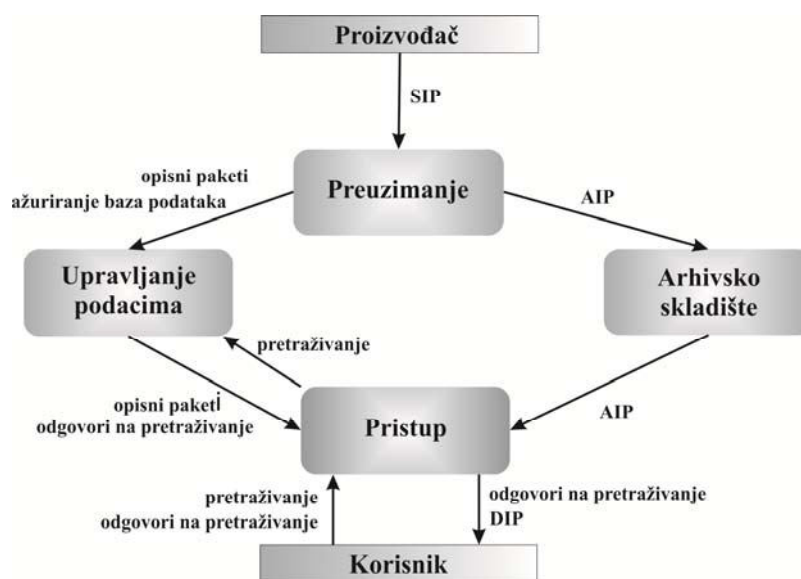
Funkcionalni model, prikazan na grafiku 4, sastavljen je od sedam glavnih funkcionalnih entiteta i veza koje se uspostavljaju između njih: pristup, administracija, arhivsko skladište, opšti servisi, upravljanje podacima i planiranje preuzimanja i očuvanja podataka (*Ingest and preservation planning*).

Mnogi aspekti ovog funkcionalnog modela počivaju na bazičnom razlikovanju SIP, AIP i DIP informacionih paketa. Preuzimanje je entitet koji prima SIP, obezbeđuje kontrolu kvaliteta, generiše AIP i koordinira ažuriranje prema Arhivskom skladištu i Upravljanju podacima. Arhivsko skladište odgovorno je za prijem AIP od Preuzimanja, kao i prosleđivanje istog u stalno (permanentno) skladište, uz upravljanje hijerarhijom skladišta, osvežavanje medija na kojima se arhiva sačuvana, sprovođenje rutinskih procedura i specijalnih kontrola da bi se izbegle greške, obezbeđenje oporavka u slučaju katastrofe i dr. Upravljanje podacima podržava popunjavanje, održavanje i pristup kako Deskriptivnim informacijama, tako i administrativnim podacima koji su sastavni deo OAIS, što uključuje administraciju baza podataka, sprovođenje upita nad podacima i proizvodnju izveštaja koji proizilaze iz upita. Administracija je odgovorna za opšte operacije arhivskog sistema, što uključuje obezbeđenje sadržaja i pregovaranje o uslovima preuzimanja, reviziju upravljanja konfiguracijom, sistemski inženjering, arhiviranje sačuvanih zahteva i uspostavljanje i održavanje standarda i politika. Planiranje očuvanja prati okruženje sa ciljem da blagovremeno uoči promene u tehnologiji ili potrebama zajednice korisnika, procenjuje implikacije tih promena na arhivu, dizajnira šablone za informacione pakete, razvija migracione planove, prototipove softvera i planove za testiranje. Pristup obezbeđuje i adekvatno ograničava sposobnost korisnika da otkrije, zahteva i primi informacije od arhiva. [2] Ograničenja se uvode u cilju zaštite određenih kategorija podataka.



Grafik 4: OAIS funkcionalni model [2]

Na grafiku 4. Opšti servisi nisu direktno prikazani, a oni predstavljaju usluge koje podržavaju rad celog sistema, što podrazumeva ispravne i funkcionalne računarske sisteme povezani stabilnim vezama, mrežu, bekap, zaštitu sistem i sl. Uprkos tome što su opšti servisi neophodni u strukturi OAIS, oni nisu u centralnom fokusu Referentnog Modela, već se podrazumevaju. Osim toga, OAIS model predstavljan grafikom 4. identifikuje Administraciju i Upravljačku strukturu, ali ne elaborira njihove funkcije.



Grafik 5: OAIS tok podataka [5]

Informacioni model definiše i opisuje tipove informacija koje se razmenjuju i kojima se upravlja u okviru OAIS. Bazira se na prepoznavanju činjenice da dugoročno očuvanje digitalnih informacija zahteva arhiv koji će biti u stanju da ih sačuva sa svim sadržajima neophodnim za njeno održanje u vremenu i mogućnost ponovne upotrebe. Referentni model koristi termin „informacioni paket“ da bi opisao logičku jedinicu koja obuhvata i sam digitalni objekat i druge vrste informacija koje treba da budu povezane sa digitalnim objektom sa ciljem njegovog očuvanja i obezbeđenja pristupa u dužem vremenskom periodu. Logički model za arhivske informacije podrazumeva informacioni objekat koji je sastavljen od podatka i reprezentativne informacije. Objekat podatka može biti fizički objekat (npr. uzorak stene), ali i digitalno objekat (sekvencu bitova). Reprezentativna informacija služi da bi digitalnom objektu obezbedila dodatno značenje. Na primeru JMBG, reprezentativna informacija bi bila značenje rasporeda brojeva (datum, mesec, poslednje tri cifre godine rođenja itd.).

Centralna tema OAIS je transformacija informacionih paketa, koji omogućuju tokove od proizvođača do korisnika informacija i čine da čitav sistem bude celishodan. Tok podataka predstavljen je na grafiku 5.

Transformacija podataka vrši se svim segmentima OAIS. Inicijalna se dešava u entitetu proizvođača, pri čemu proizvođač pregovara sa arhivom o tome šta će informacija sadržati, u kom će formatu biti i u koje vreme će SIP biti isporučen. Kada je SIP postao deo OAIS njegova forma i sadržina mogu se menjati. To se najčešće odnosi na dodeljivanje jedinstvenog identifikatora (*Unique identifier*). Nakon toga Upravljanje podacima preuzima opis paketa, dok se istovremeno novonastali AIP dodaje sadržaju Arhivskog skladišta. Kada korisnik poželi da preuzme sadržaj sačuvan u OAIS on u prvom koraku upotrebljava pretraživač (*Finding Aid*) da bi locirao informacije koja ga zanimaju. Nakon što korisnik pronade željeni sadržaj, on ga naručuje pomoću opcije narudžbine (*OAIS-supplier Ordering Aid*). Narudžbina može sadržati i uputstvo za bilo koju transformaciju koju korisnik želi da primeni na AIP u procesu kreiranja DIP. Funkcija pristupa tada beleži narudžbinu u funkcionalnu oblast upravljanja podacima. Nakon što su zadovoljeni svi zahtevi narudžbine funkcija pristupa daje odgovor korisniku. Pristup tada kontaktira upravljanje podacima i zahteva odgovarajući AIP, kako bi ga u obliku DIP isporučila korisniku. Konačno, arhivsko skladište i funkcionalna oblast upravljanja podacima kreiraju duplikat traženog objekta u privremeno spremište. Time je završen tok podataka u okviru OAIS.

### 3. ZAKLJUČAK

Primena OAIS koncepta ima potencijalno veliki značaj za razvoj društvenih nauka u Srbiji, s obzirom da podaci prikupljeni u istraživačkim projektima u oblasti društvenih nauka nisu nikada bili predmet procesa sistematskog arhiviranja u skladu sa standardima koji obezbeđuju njihovu sekundarnu upotrebu i međunarodnu uporedivost. Predstavljanje OAIS koncepta u ovom radu prevashodno je usmereno ka podizanju svesti naučne zajednice o značaju sistematskog arhiviranja istraživačkih podataka i osnivanja budućeg nacionalnog centra za podatke u društvenim naukama Republike Srbije, imajući u vidu rudimentarni karakter postojećih znanja i nedostatak relevantne literature u ovoj oblasti.

### LITERATURA

- [1.] B.F. Lavoie, The Open Archival Information System Reference Model: Introductory Guide, Digital Preservation Coalition - DPC Technology Watch Series Report 04-01, 2004.
- [2.] C.A. Lee, Open Archival Information System (AOIS) Reference Model, (2010), *Encyclopedia of Library and Information Sciences, Third Edition*, Taylor & Francis, str. 4020-4030.
- [3.] H. Stančić, Teorijski model postojanog očuvanja autentičnosti elektroničkih informacijskih objekata, doktorska disertacija, Filozofski fakultet, Sveučilište u Zagrebu, 2005.
- [4.] Model View of an OAIS's Environment, (pristup: 10.2.2014.), dostupno na:  
[<http://nost.gsfc.nasa.gov/isoas/dads/presentations/OAISRMTutorial/sld014.htm>].
- [5.] Reference Model for an Open Archival Information System (OAIS), Recommendation for Space Data System Practices, The Consultative Committee for Space Data Systems (CCSDS), 2012.