

TEORETIČNI MODEL DIGITALNIH ARHIVOV

Alenka Kavčič-Čolić

Pridobljeno: 07.04.2004 - Sprejeto: 25.10.2004

Pregledni znanstveni članek

UDK 004.9:930.251

Izvleček

Nacionalne organizacije, ki prevzemajo odgovornost za trajno ohranjanje elektronskih virov, morajo izpolnjevati določene pogoje. Med ostalim mora njihov skladiščni sistem temeljiti na referenčnem modelu odprtega arhivskega informacijskega sistema (OAIS). V prispevku je predstavljen model OAIS in nekatere pomembnejše mednarodne izkušnje pri njegovem uvajanju. Omenjen je tudi projekt Narodne in univerzitetne knjižnice (NUK) in Instituta "Jožef Stefan", ki ima za cilj zgraditi tovrstni digitalni arhiv v NUK-u, in ki temelji na omenjenem modelu.

Ključne besede: digitalni arhivi, informacijski modeli, OAIS, elektronski viri, trajno ohranjanje

Review article

UDC 004.9:930.251

Abstract

There are certain conditions that national organizations should fulfill to become trusted digital repositories. Among other things, their deposit system should be in compliance with the reference model for an Open Archival Information System (OAIS). In this paper the model OAIS is presented as well as some important international experiences regarding its implementation. The joint project of the National and University Library and The Jozef Stefan Institute, which is aimed at developing such a digital archive based on OAIS, is also mentioned.

Key words: digital repositories, information models, OAIS, electronic sources, long-term preservation

KAVČIČ-ČOLIĆ, Alenka: Theoretical model of digital archives. Knjižnica, Ljubljana, 48(2004)4, str. 63-75

1 Uvod

Vse večja produkcija elektronskega gradiva nalaga organizacijam in posameznikom potrebo po iskanju rešitve za njihovo skladiščenje in ohranjanje. Komerzialna softverska podjetja že vrsto let nudijo poslovnemu svetu sisteme za upravljanje z dokumenti (*document management systems*), ki uporabnikom omogočajo shranjevanje in lokalizacijo iskanega gradiva. Večina teh sistemov so navadne relacijske baze, ki ne omogočajo vseh potrebnih funkcij za trajno ohranjanje (*long-term preservation*¹) elektronskih virov, delujejo le na določenih platformah ali pri večjih kapacitetah in so zelo počasne.

Poleg tega mnogi menijo, da je ohranjanje elektronskega gradiva zelo enostavno in pri tem zadostuje le aplikacija za upravljanje z datotekami ter močnejši strežnik z večjo kapaciteto pomnilnika. Vendar izkušnje organizacij, ki so že začele sistematično zbirati nacionalno elektronsko produkcijo², kažejo na to, da trajno ohranjanje zahteva širši pristop v smislu informacijsko-tehnološkega, kadrovskega, prostorskega, varnostnega ter finančnega načrtovanja. Obenem je potrebno predvideti vse mogoče scenarije uporabe elektronskega gradiva v prihodnosti in razčistiti vse avtorskopravne dileme v zvezi z dostopom do le-teh. Vemo, da se informacijska tehnologija dokaj hitro razvija in bomo mnoge dokumente, do katerih danes lahko dostopamo z obstoječimi računalniškimi aplikacijami, jutri lahko prebirali le s pomočjo emulatorjev oz. jih bomo morali pretvarjati v druge računalniške formate. Za to bomo potrebovali dovoljenje avtorjev. Trajno ohranjanje elektronskega gradiva je kompleksna naloga, za katero mora organizacija izpolnjevati določene pogoje.

V svojem poročilu (*Trusted Digital Repositories*, 2002) ameriško združenje raziskovalnih knjižnic RLG (*The Research Library Group*) v sodelovanju z OCLC definira lastnosti organizacij, ki naj bi prevzele odgovornost za varovanje elektronskega gradiva (*trusted digital repositories*):

- njihov sistem ohranjanja mora biti usklajen z referenčnim modelom OAIS (*Open Archival Information System*) - o tem bo govora pozneje;
- biti morajo pripravljene za prevzemanje odgovornosti za upravljanje s tem sistemom v skladu z obstoječimi nacionalnimi oz. mednarodnimi standardi

1 Miroslav Novak (Novak, 2002) uporablja termin "dolgodobno varovanje", ki je zelo primeren prevod izraza "long-term preservation". V tem prispevku smo se odločili za uporabo termina "trajno ohranjanje", ker želimo poudariti skrb za trajni dostop do gradiva, ne glede na spremembe informacijske tehnologije, kar je značilno ne le za nacionalno knjižnico, ampak tudi za druge večje arhivske organizacije.

2 S to problematiko se že mnoga leta ukvarjajo nacionalne knjižnice. Prve med njimi, ki so trajno ohranjanje elektronskih virov implementirale v praksi, so skandinavske nacionalne knjižnice, Kongresna knjižnica v Washingtonu in avstralska nacionalna knjižnica. Danes vzdržujejo elektronske arhive mnoge druge depozitne institucije, kot so nacionalni arhivi, muzeji, RTV itd.

- pri tem so mišljeni vsi procesi v zvezi s prevzemanjem, opisom, vzdrževanjem, varovanjem in zaščito gradiva;
- biti morajo zanesljive – odločitev za ohranjanje ne sme biti pogojena s trenutnim komercialnim interesom, organizacije pa morajo prevzeti to odgovornost tudi za prihodnost;
- zagotavljati morajo finančno uravnovešeno – imeti morajo stalni vir financiranja in biti sposobne kriti vse stroške, ki jih trajno ohranjanje zahteva;
- imeti morajo ustrezne strokovne kadre in tehnologijo za ohranjanje tovrstnega gradiva;
- zagotavljati morajo varnost pred vdorom v sistem, zaščito avtorskih pravic, zaščito pred nesrečami in skrbeti za trajno ohranjanje.

Odgovornost za trajno ohranjanje elektronskega gradiva lahko prevzame več institucij glede na vrsto gradiva, vendar le če zadovoljujejo posebne pogoje za ohranjanje. V Sloveniji bi poleg nacionalne knjižnice in nacionalnega arhiva tovrstno odgovornost lahko prevzele tudi nekatere druge knjižnice oz. arhivi, muzejske, kulturne in javne državne institucije. Pomembno je, da te institucije razumejo zahteve svojih uporabnikov in skupnosti, da delujejo koordinirano, in da so pripravljene trajno vlagati v vzdrževanje informacijske tehnologije in digitalnega fonda.

V omenjenem poročilu skupine RLG in OCLC se priporoča usklajenost sistema za arhiviranje z referenčnim modelom odprtega arhivskega informacijskega sistema (*Open Archival Information System - OAIS*). Na njegov velik pomen kaže dejstvo, da je ta model bil sprejet kot standard za oblikovanje vseh elektronskih arhivov. Standard definira OAIS, kot "arhiv, ki je sestavljen iz organizacije ljudi in sistemov, in ki je sprejel odgovornost, da trajno ohranja informacije in omogoča določeni skupnosti dostop do le-teh" (Reference Model..., 2002, str. 1). Referenčni model OAIS je teoretični model, ki zajema vse procese v celovitem sistemu za trajno ohranjanje elektronskega gradiva.

Namen tega prispevka je prikazati vsebino omenjenega modela ter dosedanje pomembnejše mednarodne izkušnje pri njegovem uvajanju v tujini in pri nas, in sicer v Narodni in univerzitetni knjižnici (NUK).

2 Referenčni model OAIS

2.1 Zgodovina

Na prošnjo Mednarodne organizacije za standardizacijo (ISO) je CCSDS (*Consultative Committee for Space Data Systems*), mednarodno združenje vesoljskih agencij, ki je namenjeno razvoju standardov za upravljanje s podatki, ki so v podporo vesoljskim raziskavam, začel s poskusom usklajevanja razvoja arhivskih standardov za trajno arhiviranje gradiva v elektronski obliki. CCSDS je pripravil referenčni model, ki naj bi določil terminologijo in koncepte za opis in primerjavo podatkovnih modelov in arhivskih arhitektur, identificiral pomembne entitete in njihove medsebojne odnose v arhivskem okolju, pojasnil ključne funkcionalne in informacijske komponente arhivskega sistema in nazadnje ponudil okvir, v katerem naj bi standardizacija potekala.

Maja leta 1999 je CCSDS dokončal referenčni model OAIS (*Open Archival Information System*), ki opisuje okolje, v katerem deluje arhiv³, funkcionalne komponente arhiva in informacijsko infrastrukturo, ki podpira procese arhiva. Kot osnutek predloga ISO (Reference Model..., 2002) je bil referenčni model razširjen izven vesoljskih organizacij, da bi postal predmet širše strokovne obravnave. Osnutek je hitro pritegnil tudi knjižnice in druge institucije s kulturno vsebino, vladne agencije in privatni sektor (OCLC/RLG..., 2002, str. 5). Leta 2001 je izšla prenovljena izdaja referenčnega modela in leta 2003 je predlog registriran kot ISO standard št. 14721.

Celotni sistem OAIS temelji na ohranjanju informacij, skupaj s kontekstom oziroma bazo znanja, ki je potrebna za razumevanje vsebine teh informacij.

Referenčni model OAIS sestavljata informacijski model, ki definira vrste informacij ali informacijskih objektov v sistemu, in funkcionalni model, ki opisuje procese v sistemu.

2.2 Informacijski model

Informacijski model OAIS loči štiri vrste informacijskih objektov, ki se nanašajo na različne zvrsti informacij, ki so potrebne za trajno ohranjanje elektronskih virov:

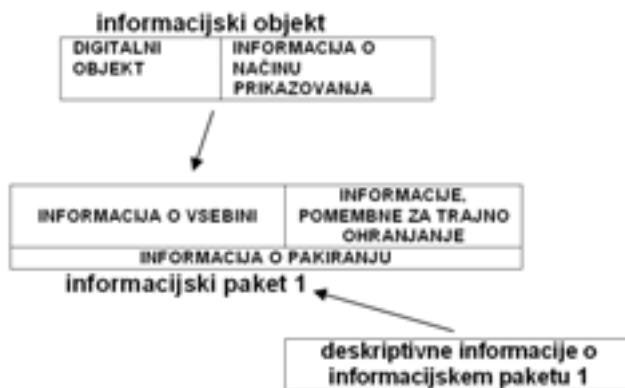
3 Termin "arhiv" v prispevku uporabljamo za označevanje sistema za skladiščenje oz. skladišča elektronskih virov.

- **informacija o vsebini** – je informacija, ki jo želimo ohraniti. To je pravzaprav elektronski vir ali, po terminologiji v standardu, digitalni objekt;
- **informacije, pomembne za trajno ohranjanje vsebine** (*Preservation description information* – PDI) – vse informacije, ki nam bodo omogočile, da bomo razumeli informacije o vsebini v prihodnosti;
- **informacije o pakiranju** – kako so shranjene vse komponente digitalne vsebine;
- **deskriptivne informacije** – vsebujejo bibliografski opis, ki nam pomaga pri lokalizaciji in dostopu do vira.

Vsi informacijski objekti so sestavljeni iz podatkovnega ali digitalnega objekta (seznam bitov v obliki niza števil) in informacij o načinu prikazovanja (*Representation information*), ki omogočajo interpretacijo in razumevanje vsebine digitalnega objekta (na primer ali je elektronski zapis slika, besedilo ali zvočni posnetek). Informacije o načinu prikazovanja so lahko strukturne ali semantične. Strukturne nam povedo o formatu podatkov in lahko tudi opisujejo strojno in programsko okolje, potrebno za dostop do podatkov (npr. besedilo je v ASCII). Semantične pa dodajo pomen podatkovnim strukturam, ki jih določajo strukturne informacije (primer, besedilo je v angleščini, je razlaga, kako moramo interpretirati informacije o pakiranju itd.).

Osnovni koncept modela OAIS je informacijski paket, ki naj bi vseboval vse štiri tipe informacijskega objekta. Sestavljata ga **informacija o vsebini** in **informacije, ki so pomembne za trajno ohranjanje digitalne vsebine** (PDI). Oba dela sta enotno povezana in »spakirana« v informacijski paket. Spremljajo ju **informacije o pakiranju**, ki navajajo način pakiranja (uporabljena aplikacija ali drugo). Informacija o vsebini je najpomembnejši element informacijskega paketa in je cilj trajnega ohranjanja. Vsebuje elektronski vir⁴ (digitalni objekt), ki ga želimo ohraniti, ter informacijo o načinu prikazovanja (Slika 1).

4 OAIS kot teoretični model obravnava elektronske vire, vendar ne izključuje tudi vire na fizičnih nosilcih, predvsem tiste, ki prinašajo dodatne informacije o digitalnih virih (pr. priručniki za uporabo računalniških aplikacij, tiskana slika vmesnikov, podatkovnih baz, umetniških del itd.).



Slika: Informacijski paket (Vir: Reference Model..., 2002, Slika št. 2-3, str. 2-5)

Informacije, pomembne za trajno ohranjanje (PDI) so:

- **informacije o provenienci**, ki opisujejo zgodovino vira od njegovega nastanka in lastništva oz. spremembe in posege na samem viru (npr. ali je dokument elektronska različica originala ali je uporabljen poseben postopek kompresije digitalne slike...);
- **informacije o kontekstu** oz. kako se vir povezuje z drugimi informacijami izven informacijskega paketa (npr. iz katerega originala je slika ali dokument digitaliziran ali je to slika iz elektronske monografije ali je članek del serijske publikacije);
- **referenčne informacije**, ki nedvoumno identificirajo vir (so lahko bibliografski identifikatorji ISBN, ISSN...)
- **informacije o pristnosti vira** (*fixity information*), ki nam zagotavljajo, da se vir ni spremenil od njegovega prevzema (to je lahko digitalni podpis ali checksum).

Deskriptivne informacije (metapodatki) nam pomagajo, da v arhivu najdemo določen informacijski paket.

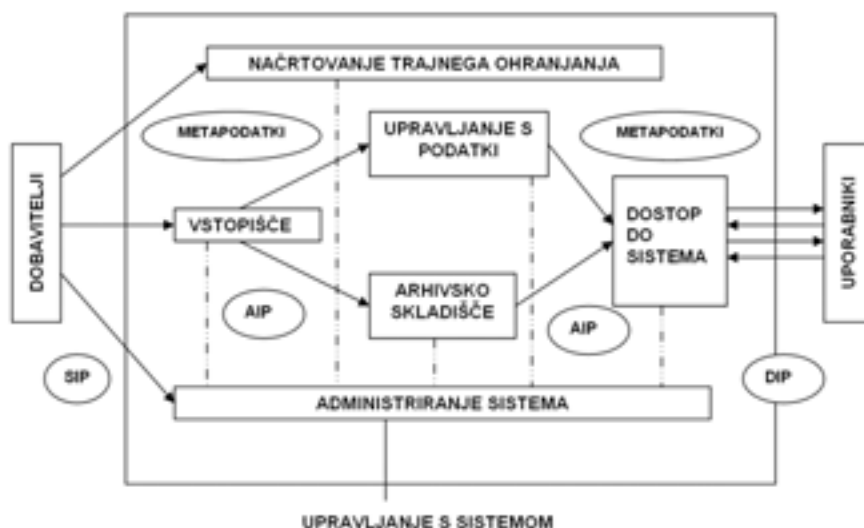
2.3 Funkcionalni model

Informacijski paket se od svojega vstopa v arhivski sistem do njegovega izhoda spreminja in lahko prevzema tri različne oblike, odvisno od arhivskega procesa, v katerem se nahaja. Ti procesi se odvijajo v šestih funkcionalnih entitetah, in sicer:

1. **vstopišče** – predstavlja vstopno entiteto elektronskih virov v sistem. Sistem prevzema elektronski vir od proizvajalcev, avtorjev, posrednikov ali na kak drug način. Spremljajo ga ali pa ne osnovne informacije o njem. To je takomenovani vstopni informacijski paket ali *Submission information package* (SIP). SIP lahko vsebuje vse štiri omenjene informacijske objekte ali le informacijo o vsebini. Po preverjanju, ali v sistemu že obstaja kopija ali starejša različica vira, se preveri njegova kvaliteta, celovitost in pristnost in se opremi z novimi metapodatki, ki so pomembni za arhivsko trajno ohranjanje. Iz SIP-a nastane eden ali, v primeru publikacij, ki so sestavljene iz več samostojnih delov, več arhivskih informacijskih paketov (AIP), ki nadaljujejo svojo pot do skladišča sistema. V tej enoti se poleg AIP pripravijo deskriptivni metapodatki (bibliografski opis), ki se ločeno hranijo v posebni podatkovni zbirki (npr. v katalogu);
2. **arhivsko skladišče** – prevzame AIP iz vstopišča in ga shrani v svoje skladišče na podlagi obstoječe hierarhije. Skladišče skrbi za osvežitev nosilcev, rutinsko preverjanje napak, omogočanje možnosti za povrnitev podatkov v primeru nesreč in omogočanje dostopa do AIP na zahtevo sistema;
3. **upravljanje s podatki** – ta entiteta omogoča dostop do deskriptivnih podatkov (do kataloga) fonda in do administrativnih podatkov, ki so pomembni za upravljanje z arhivom. Tu se izvajajo vse funkcije podatkovnih baz (indeksiranje, brskanje, poizvedovanje, prikazovanje rezultatov in poročil...) in njihovo vzdrževanje oziroma prenavljanje;
4. **administriranje** – v tej entiteti se izvajajo storitve in funkcije v zvezi s celovitim delovanjem arhivskega sistema. Tu so zahtevki in pogajanja v zvezi s prevzemom gradiva od avtorjev, posrednikov, založnikov, preverjanje kvalitete gradiva ter vzdrževanje delovanja oz. spreminjanje konfiguracije strojne in programske opreme sistema. V tem delu se izvaja kontrola nad delovanjem celotnega arhiva, od prevzemanja, skladiščenja do uporabniških storitev v zvezi z dostopom do gradiva. Obenem se odloča o prenavljanju oz. migraciji v nove različice ali uvajanje novih standardov. Skratka, v tej entiteti se definira politika upravljanja celotnega arhivskega sistema;
5. **načrtovanje trajnega ohranjanja** – namen te entitete je zagotoviti, da bodo informacije oz. digitalne vsebine, ki so skladiščene v arhivu, v prihodnosti trajno dostopne uporabnikom, ne glede na to ali bo prvotna informacijska tehnologija še uporabna ali dostopna. To pomeni, da se v tem delu spremlja digitalni fond, potrebna infrastruktura za dostop do njega, potrebe skupnosti in se določajo standardi ter strategije za trajno ohranjanje elektronskih virov. V primeru razvoja novih računalniških platform ali verzij aplikacij se odloča o emulacijah, migracijah ali konverzijah, ki morajo biti zelo natančno načrtovane. Informacije, ki nastanejo v tem delu, se uporabljajo v vstopišču zaradi spremljanja in prevzemanja novega fonda;
6. **dostop** – v tej entiteti se izvajajo storitve za uporabnike tako, da se jim omogoča iskanje po zbirki podatkov (katalogu), njihovo lokalizacijo in dostop do

fonda oz. se omejuje dostop do avtorsko zaščenega gradiva. Na podlagi bibliografskega opisa lahko uporabnik priključuje AIP iz arhivskega skladišča in, če ni omejitev dostopa, si ga lahko ogleda ali prebere v celoti. Pri tem se AIP odpackira in spremeni v obliko, ki je primerna za vpogled na zaslonu. To je informacijski paket za vpogled ali *Dissemination information package* (DIP).

Na Sliki 2 je prikazan celotni model.



Slika 2: Funkcionalne entitete OAIS (Vir: Reference Model..., 2002, Slika št. 4-1, str. 4-1)

3 Mednarodne implementacije OAIS

Preden je bil OAIS sprejet kot ISO standard, je bilo narejenih nekaj projektov, ki so potrdili njegovo ustreznost pri gradnji elektronskih arhivov. Med njimi so najbolj znani CEDARS (*CURL Exemplars in Digital Archives*) iz leta 1998-2002 (OCLC/RLG..., 2002), avstralski projekt PANDORA in evropski projekt NEDLIB (1998-2000). Sistem je univerzalen, ne pogojuje določene infrastrukture, temveč se osredotoča na model funkcij in procesov, ki so lahko sprejemljive za ohranjanje informacij in gradiva na vseh nosilcih. Vse več knjižnic, ki so zadolžene za ohranjanje pisne digitalne dediščine, gradi svoje arhive na osnovi referenčnega modela OAIS. Skupina ameriških raziskovalnih knjižnic (RLG) in OCLC ga priporočajo kot izhodišče za razvoj digitalnih arhivov (Trusted Digital Repositories, 2002).

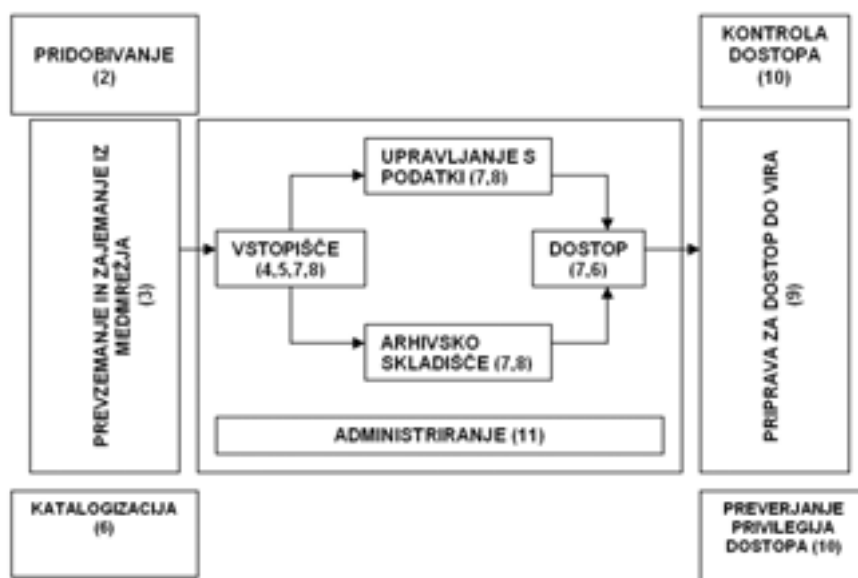
V projektu NEDLIB (*NEtworked Digital LIBrary*) so sodelujoči partnerji, večinoma nacionalne knjižnice, uporabili referenčni model OAIS kot temelj teoretičnega modela svojih arhivov, vendar so ga dodelali in prilagodili procesom knjižnice. Projekt je opisan v zelo obširnih poročilih, ki so zelo uporabna za razvoj podobnih arhivskih knjižničnih sistemov. Zato ga v nadaljevanju podrobneje predstavljamo.

Sodelujoče knjižnice projekta NEDLIB so entitetam iz OAIS dodale procese, ki so značilni za bibliografsko obdelavo oz. za uporabniški dostop do gradiva. Tako je nastal DSEP (*Deposit System for Electronic Publications*), depozitni sistem za elektronske publikacije, ki je vseboval enajst funkcionalnih entitet (Steenbakkens, 2000):

1. **izbiro gradiva** – proces, ki se opravlja ročno in na podlagi nacionalne politike odločanja o ohranjanju gradiva ter obstoječe zakonodaje oziroma sklenjenih pogodb z avtorji, založniki ali posredniki;
2. **pridobivanje** – zajema vsa administrativna dela, ki so potrebna za pridobivanje elektronskega gradiva, vključno s prejemanjem informacij o novih publikacijah, sklepanjem pogodb, računovodstvom itd.;
3. **prevzemanje, zajemanje iz medmrežja** – predvideva vsa dejanja, ki so potrebna, da bi se pridobila kopija elektronskega vira v arhiv. Pri tem lahko prevzemamo gradivo na fizičnem nosilcu (CD, disketa), preko medmrežja (FTP, e-pošta) ali ga zajemamo s pomočjo robota neposredno iz medmrežja selektivno ali na podlagi domen;
4. **registracijo** – elektronska publikacija se vnaša v depozitni sistem. Sledijo naloge registracije, odločanje in obveščanje dobaviteljev o sprejetju publikacije;
5. **preverjanje** – rutinska kontrola za preverjanje celovitosti in kvalitete pridobljenega elektronskega vira. V to entiteto sodijo procesi instalacije, ugotavljanje napak pri delovanju in obveščanje nabavne službe;
6. **opis vira** – elektronski vir se katalogizira, uporabijo se osnovne informacije, pridobljene od dobaviteljev, kar je lahko tudi avtomatsko ali polavtomatsko, vnos metapodatkov za trajno ohranjanje, dodajanje identifikatorjev (URN ali druge), vsebinsko obdelavo itd.;
7. **skladiščenje vira** – elektronski vir se prenaša v arhivski sistem. Pri tem procesu se lahko opravi migracija vira na drug nosilec ter njegova reprodukcija in prenos na drugo lokacijo zaradi zaščite pred nesrečami. Občasno se preverja kontrola kvalitete;
8. **trajno ohranjanje elektronskega vira** – vključuje vsa dejanja, ki so povezana s trajnim ohranjanjem, kot so: načrtovanje strategije trajnega ohranjanja, migracije, ugotavljanje potrebe po emulaciji aplikacij ali konverziji zastarelih formatov, kontrola kvalitete idr.;

9. **pripravo za dostop do vira** (*delivery*) – gre za entiteto, ki zajema proces prenosa kopije originalnega vira iz skladišča v vmesni delovni prostor na računalniku, kjer se pripravi tako, da bo uporabnik lahko dostopal do te. Če je na primer delo iz več poglavij, vsebuje dele, ki jih je treba ponovno formatirati oz. do tega lahko dostopamo le z emulacijo, potem se vse to opravi v okviru te entitete;
10. **dostop** – zajema vse funkcije uporabniškega vmesnika, s katerim končni uporabnik poišče, odda zahtevek za vpogled in dostopa do gradiva. Poseben del predstavlja kontrola dostopa in preverjanje privilegija dostopa, če je gradivo avtorsko zaščiteno;
11. **upravljanje s sistemom** – predvideva upravljanje, kontrolo in spremljanje vseh desetih procesov.

Na Sliki 3 je prikazan model DSEP v primerjavi z referenčnim modelom OAIS. Opažamo, da v središču ostaja struktura OAIS. Pred entiteto "vstopišče" in za entiteto "dostop" so dodane nove entitete. Za lažjo primerjavo smo vse entitete označili s številkami opisanih entitet iz DSEP.



Slika 3: Depozitni sistem za elektronske publikacije (Vir: Werf, 2000, Slika št. 6, str. 17)

Do sedaj je bila največja pozornost pri tem sistemu posvečena metapodatkom za trajno ohranjanje. Težko je predvideti spremembe informacijske tehnologije in kateri podatki nam bodo v prihodnosti pomagali razumeti informacije, ki jih

sedaj želimo ohraniti. Zanimivo je, da so v različnih nacionalnih knjižnicah pristopili k definiciji potrebnih metapodatkov za trajno ohranjanje zelo različni⁵.

4 Implementacija modela OAIS v nacionalnem elektronskem arhivu

Prednost Narodne in univerzitetne knjižnice (NUK) pri ustvarjanju nacionalnega elektronskega arhiva je, da lahko upošteva tuje izkušnje ter pozitivne strani in pomanjkljivosti različnih pristopov pri arhiviranju elektronskih virov. Leta 2002 je NUK v sodelovanju z Institutom "Jožef Stefan" (IJS) pridobil sredstva iz Ministrstva za kulturo, Ministrstva za informacijsko družbo ter Ministrstva za šolstvo, znanost in šport za izvajanje projekta iz Ciljnega raziskovalnega programa "Metodologija arhiviranja slovenskih elektronskih publikacij na medmrežju" (ASEP). V okviru tega projekta je bil NUK zadolžen za pripravo metodologije arhiviranja elektronskih publikacij na medmrežju, medtem ko je bil IJS zadolžen za izdelavo zmogljivega robota za zajemanje teh publikacij na medmrežju.

Teoretični model arhiviranja, ki ga razvijamo v NUK-u, temelji na rezultatih projekta NEDLIB in drugih mednarodnih izkušnjah. Zaradi načrtovanja implementacije različnih pristopov pri selekciji in zajemanju elektronskih virov, smo v okviru arhivskega skladišča predvideli spletni arhiv, ki bo deloval ločeno od arhiva selektivno zajetih dokumentov. Tako bodo spletne strani, ki so zajete na podlagi predefiniranih domen, dostopne preko ločenega brskalnika po zgledu *Internet Archive*⁶ (Kavčič-Čolić, 2001). Dokumenti, ki bodo pomembni iz nacionalnega, kulturnega, znanstvenega in zgodovinskega vidika, in ki sodijo v sloveniko, bodo bibliografsko obdelani po veljavnih strokovnih standardih in ohranjeni ločeno. Do le-teh bo omogočen dostop preko lokalnega oz. vzajemnega kataloga, odvisno od avtorskopравnih pogodb in dovoljenj avtorjev ali založnikov. Tudi mi smo posebno pozornost posvetili metapodatkom za trajno ohranjanje.

5 Medtem ko so se metapodatki za trajno ohranjanje elektronskih virov v okviru projekta NEDLIB določili na podlagi komponent sistema OAIS, se v avstralski nacionalni knjižnici določajo na podlagi informacij, pridobljenih izven sistema. Avstralci uporabljajo 24 vrst metapodatkov, ki se nanašajo na različne nivoje objekta (datoteko, elektronski vir ali zbirko). Tretji primer je model novozelandske nacionalne knjižnice, ki temelji na štirih medsebojno povezanih entitetah, ki opredeljujejo potrebne metapodatke za trajno ohranjanje: objekt, proces, datoteko in spremembe metapodatkov.

6 The Internet Archive je neprofitna organizacija v San Franciscu, Združene Države Amerike, ki hrani vse spletne strani na html od leta 1996. Gradivo jim s šestmesečnim zamikom podarja komercialno podjetje Alexa Internet, ki ga za svoje potrebe zbira mesečno.

V ozadju bo elektronski arhivski sistem, ki ga bomo razvili s pomočjo drugega aplikativnega projekta, ki ga financira Ministrstvo za šolstvo, znanost in šport in katerega nosilec je IJS: "Oblikovanje in analiza zbirke slovenskih digitaliziranih in elektronskih publikacij nacionalnega pomena". Sistem bo praktična implementacija teoretičnega modela iz projekta ASEP. Predvidevamo, da bomo prve rezultate dobili že konec leta 2004.

5 Zaključek

V prispevku smo predstavili referenčni model OAIS, ki je postal standard za gradnjo večjih sistemov za trajno ohranjanje elektronskih virov. Pomembno vlogo je treba pripisati evropskemu projektu NEDLIB, ki je omenjeni referenčni model dogradil in prilagodil potrebam nacionalnih depozitnih elektronskih arhivov. Objavljena poročila ob zaključku projekta so zelo uporabna kot izhodišča pri razvoju lastnih arhivov.

Model je zelo široko zastavljen in uporaben ne glede na že obstoječe sisteme v knjižnicah. Obenem opisuje procese, ki jih nikakor ne smemo opustiti pri trajnem ohranjanju elektronskih virov in ki so nevtralni glede na način njihovega izvajanja.

Naslednja naloga po vzpostavitvi nacionalnega elektronskega depozitnega arhiva v NUK-u bo koordinacija in standardizacija postopkov obdelave in trajnega ohranjanja med različnimi slovenskimi depozitnimi elektronskimi arhivi.

6 Citirani viri

1. Day, M. (1998). CEDARS: Digital preservation and metadata. *Sixth DELOS Workshop: Preservation of Digital Information, Tomar, Portugal, 17-19 June 1998*. Pridobljeno 5.4.2004 s spletne strani: <http://www.ukoln.ac.uk/metadata/presentations/delos6/cedars.html>
2. *INTERNET ARCHIVE*. Pridobljeno 5.4.2004 s spletne strani: <http://webdev.archive.org/>
3. Kavčič-Čolić, A. (2001). Elektronsko arhiviranje spletnih strani: nov izziv za nacionalne knjižnice. V Ambrožič, M. (Ur.). *Digitalna knjižnica : Zbornik posvetovanja ZBDS* (str. 211-224). Ljubljana: ZBDS, 2001.

4. NEDLIB (*Networked European Deposit Library*) [informacije o projektu]. Pridobljeno 5.4.2004 s spletne strani: <http://www.kb.nl/coop/nedlib/>
5. Novak, M. (2002). Arhivsko in dokumentarno gradivo ter sistemi zagotavljanja kakovosti. *Sodobni arhivi*, 24, 32-45. Pridobljeno 5.4.2004 s spletne strani: <http://www.cimrs.uni-mb.si/novak.pdf>
6. OCLC/RLG Working Group on Preservation Metadata (2002). *Preservation Metadata and the OAIS Information Model. A metadata framework to support the preservation of digital objects*. Pridobljeno 5.4.2004 s spletne strani: http://www.oclc.org/research/projects/pmwg/pm_framework.pdf
7. PANDORA [informacije o projektu]. Pridobljeno 15.4.2003 s spletne strani: <http://pandora.nla.gov.au/index.html>
8. PANDORA [dokumeti in priročniki]. Pridobljeno 15.4.2003 s spletne strani: <http://pandora.nla.gov.au/documents.html>
9. *Reference Model of an Open Archival Information System (OAIS) : Recommendation for space data system standards : CCSDS 650.0-B-1.Blue Book* (2002). Washington: CCSDS (Consultative Committee for Space Data Systems) Secretariat. Pridobljeno 5.4.2004 s spletne strani: http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html
10. Steenbakkens, J. (2000). *The Nedlib Guidelines : Setting up a Deposit System for Electronic Publications*. Den Haag: Koninklijke Bibliotheek. (NEDLIB Report series ; 5). Pridobljeno 5.4.2004 s spletne strani: <http://www.kb.nl/coop/nedlib/results/NEDLIBguidelines.pdf>
11. *Trusted Digital Repositories: Attributes and responsibilities : an RLG-OCLC Report* (2002). Mountain View, CA: RLG. Pridobljeno 5.4.2004 s spletne strani: <http://www.rlg.org/longterm/repositories.pdf>
12. Werf, T. van der (2000). *The Deposit System for Electronic Publications: A process Model*. Den Haag: Koninklijke Bibliotheek. (NEDLIB Report Series ; 6). Pridobljeno 5.4.2004 s spletne strani: <http://www.kb.nl/coop/nedlib/results/DSEProcessmodel.pdf>

Mag. Alenka Kavčič-Čolić je zaposlena v Narodni in univerzitetni knjižnici v Ljubljani kot vodja Bibliotekarskega raziskovalnega centra.

Naslov: Turjaška 1, 1000 Ljubljana

Naslov elektronske pošte: alenka.kavcic@nuk.uni-lj.si