

METAPODATKI ZA TRAJNO OHRANJANJE ELEKTRONSKIH VIROV

Alenka Kavčič-Čolić

Pridobljeno: 11.05.2004 – Sprejetno: 28.09.2004

Izvirni znanstveni članek

UDK 025.31:004.738.5

Izvleček

Metapodatki nam posredujejo pomembne informacije o elektronskih virih in nam pomagajo pri njihovi lokalizaciji. Zaradi naraščanja števila digitalnih arhivov se v zadnjih letih vse večjo pozornost posveča metapodatkom za trajno ohranjanje, ki naj bi omogočili dostop do elektronskih virov tudi v prihodnosti. Namen tega prispevka je predstaviti pomen in funkcijo metapodatkov za trajno ohranjanje v okviru elektronskega arhiva, ki sledi modelu OAIS. Avtorica prispevka predstavlja nekatere pomembnejše mednarodne rešitve na tem področju in predlaga minimalni nabor metapodatkov za trajno ohranjanje elektronskih virov.

Ključne besede: metapodatki, elektronski viri, trajno ohranjanje, model OAIS, mednarodne primerjave

Original scientific article

UDC UDK 025.31:004.738.5

Abstract

Metadata provide information about electronic resources and are indispensable for their localization. In the last few years due to the proliferation of digital archives increased attention has been given to metadata for long-term preservation. These metadata provide information important for accessing electronic resources in the future. The paper presents the role and function of metadata for long-term preservation in electronic archives according to the OAIS model. The author gives an overview of important international metadata frameworks and proposes a minimal metadata set for the long-term preservation of electronic resources.

Key words: metadata, electronic resources, long-term preservation, OAIS model, international comparative studies,

KAVČIČ-ČOLIĆ, Alenka: Metadata for long-term preservation of electronic resources. Knjižnica, Ljubljana, 48(2004)3, str. 97-119

1 Uvod

Razvoj informacijske tehnologije in veliko navdušenje nad virtualnim dostopom do informacij sta v veliki meri vplivala na ustvarjanje številnih digitalnih in digitaliziranih zbirk. Sedaj, ko je omogočen dostop na daljavo do neštetih elektronskih informacij, se sprašujemo, ali bodo in kaj lahko storimo, da res bodo te dosegljive tudi v prihodnosti. Pri tem ne mislimo le na elektronske zbirke, ki jih ustvarjajo knjižnice, ampak tudi na tiste zbirke, ki so ustvarjene v muzejih, arhivih, javnih zavodih, kot je RTV, elektronskih založbah ter drugih organizacijah v javnem in privatnem sektorju.

Danes je težko odgovoriti na vprašanje, kam nas bo čez petdeset ali več let pripeljala informacijska tehnologija in ali smo v sedanjosti res sposobni predvideti način razmišljanja in delovanja prihodnjih generacij. Vendar večina ekspertov s področja elektronskega arhiviranja¹ trdi, da našim zanamcem lahko precej olajšamo delo, če poleg oblikovanja standardnega arhivskega sistema, elektronske vire opremimo z ustreznimi metapodatki, kar je tudi predmet tega prispevka.

Kaj pravzaprav so metapodatki? Znano je, da termin "meta-podatek" dobesedno pomeni "podatek o podatku". Anne J. Guillard-Swetland (2000) definira metapodatke kot "zbir vsega, kar lahko povemo o kateremkoli informacijskem objektu² na kateremkoli nivoju agregacije". Uveljavili so se z razvojem elektronskih virov in nam pomagajo pri njihovi identifikaciji, opisu in lokalizaciji. Poleg tega nam omogočajo kontrolo nad njimi tako, da nam posredujejo informacije o njihovi strukturi, delovanju, o okoljih in pogojih za dostop do le-teh. Obenem nam sporočajo, kdo in kako lahko dostopa do elektronskega vira.

V literaturi se omenjajo različne delitve metapodatkov, vendar se te ne izključujejo in različni metapodatki se medsebojno tudi prekrivajo. Pogosta delitev po funkciji uvršča metapodatke v deskriptivne (COMARC, *Dublin Core*, EAD...), strukturne (METS, MODS...), tehnične in administrativne. Vendar omenjeni metapodatki ne vsebujejo vse informacije, ki so pomembne za trajno ohranjanje elektronskih virov. Z razvojem večjih nacionalnih arhivskih sistemov ali repozitorijev konec devetdesetih let se vse večjo pozornost posveča metapodatkom, ki naj bi nas skušali rešiti od neberljivosti elektronskih virov v prihodnosti. To so tako ime-

1 *Projekti, ki so imeli za cilj rešiti problem trajnega ohranjanja elektronskih virov (CEDARS, NE-DLIB in drugi) so veliko pozornost posvetili metapodatkom za trajno ohranjanje. Njihov pomen je tudi potrjen na mednarodnih posvetovanjih, ki obravnavajo problematiko gradnje digitalnih knjižnic (pr. ECDL - European Conference on Research and Advanced Technology for Digital Libraries).*

2 *Informacijski objekt je samostojna enota (elektronski vir), ki je sestavljena iz enega ali več elementov (datotek) in katero lahko upravljamo s pomočjo človeka ali sistema. Vsi informacijski objekti imajo tri značilnosti: vsebino, kontekst in strukturo.*

novani metapodatki za trajno ohranjanje (*long-term preservation metadata*³) in so neločljivi od vsakega pomembnejšega elektronskega arhiva, ki ima namen preživeti vsaj 50 ali več let. Ti metapodatki vsebujejo informacije o formatu oz. o strukturi bitnega toka (*bit stream*) digitalnega objekta, značilnosti aplikacije, s katero lahko odpremo dokument, ali informacije o lastnosti in funkcionalnosti vira ipd.

Dobra osnova za razumevanje pomembnosti in ozadje, v katerem delujejo metapodatki za trajno ohranjanje, je referenčni model OAIS, ki obenem prikazuje v kakšni povezavi so ti z drugimi tipi metapodatkov.

Zaradi lažjega razumevanja definicije metapodatkov za trajno ohranjanje bo v tem prispevku na kratko predstavljen referenčni model OAIS. V nadaljevanju bo podan povzetek pomembnejših pristopov k definiciji metapodatkov za trajno ohranjanje elektronskih virov, ki so pogojeni s specifično implementacijo modela OAIS. Na koncu bomo predstavili skupino metapodatkov za trajno ohranjanje, ki smo jo pripravili na podlagi tujih izkušenj.

V besedilu uporabljamo termin "arhiviranje" za označevanje procesa bibliografske obdelave elektronskega gradiva in njegovo skladiščenje v primernem arhivskem sistemu. Posledično uporabljamo termin depozitni arhiv, ki predstavlja sistem za skladiščenje in ohranjanje elektronskih virov nacionalnega pomena.

2 Referenčni model OAIS

Na pobudo Mednarodne organizacije za standardizacijo je mednarodno združenje vesoljskih agencij (CCSDS - *Consultative Committee for Space Data Systems*) pripravilo predlog referenčnega modela elektronskega arhiva z namenom standardizacije gradnje tovrstnih arhivov. Tako je nastal referenčni model OAIS (*Open Archival Information System*) (Reference Model..., 2002), ki je leta 2003 bil registriran kot standard ISO 14721 za izdelavo arhivskih sistemov elektronskih virov. Večina nacionalnih depozitnih elektronskih arhivov temelji na osnovnih principih tega modela: elektronski viri se vnesejo v sistem v obliki informacijskega paketa in s pomočjo metapodatkov spreminjajo svojo obliko in funkcijo. Če si zamislimo elektronski vir kot katerokoli datoteko ali skup različnih datotek, ki so pakirane z določenimi podatki v informacijskem paketu, je ta pri vnosu v sistem še bibliografsko neobdelan oz. zgoščen z le nekaterimi podatki, ki identificirajo njegovo vsebino. To je vnosni informacijski paket (*Submission informa-*

3 Dobesedni prevod izraza "long-term preservation" je "dolgodobno oz. dolgoročno ohranjanje". Kljub temu smo se odločili za uporabo izraza "trajno ohranjanje", ker izhajamo iz dejstva, da želimo elektronske vire trajno ohraniti, čeprav je pojem trajnosti zelo relativen.

tion package - SIP). Bibliotekar ga prevzame, dodatno opremi z bibliografskimi in drugimi metapodatki in ga pošlje v sistem. Ta novi informacijski paket je arhivski informacijski paket (AIP), ki se ohranja ločeno od bibliografskega opisa oz. podatkov, ki so pomembni za njegovo lokalizacijo. V arhivu počiva, dokler ga uporabnik ne prikličje ali ga administrator arhiva, zaradi tehnoloških sprememb, ne pretvori v novejšo različico ali v drugo strukturo. Ko ga uporabnik prikličje, ker je način prikazovanja in dostopa do elektronskega vira prilagojen potrebam uporabnikov in obstoječe tehnologije, se ta ponovno spremeni v dostopni informacijski paket (*Dissemination information package - DIP*).

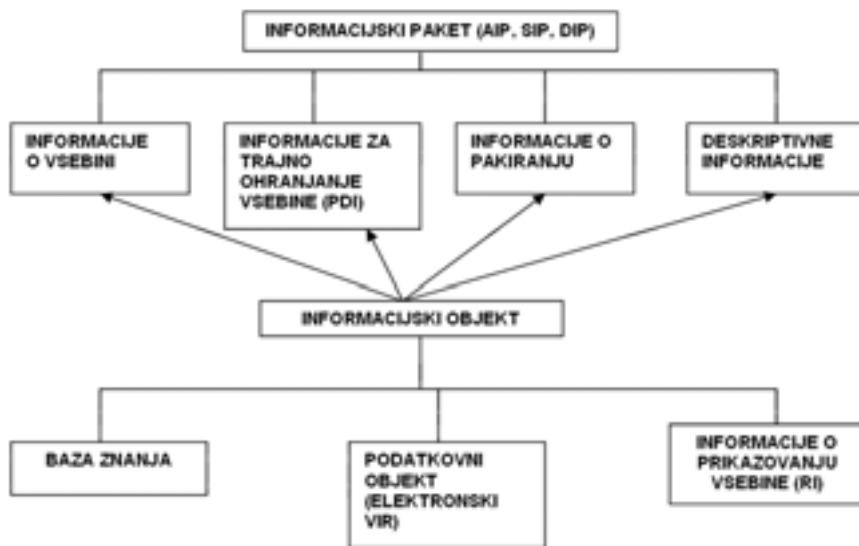
Informacijski model OAIS v arhivskem informacijskem paketu loči štiri zvrsti informacij oz. informacijskih objektov, med katerimi sta najpomembnejša:

1. **informacije, pomembne za trajno ohranjanje vsebine** (*Preservation description information - PDI*), ki so:
 - a. **referenčne informacije** - omogoča nedvomno identifikacijo vsebine elektronskega vira;
 - b. **informacije o kontekstu** - povedo zakaj je določen elektronski vir bil ustvarjen in kako je povezan z drugimi vsebinami;
 - c. **informacije o provenienci** - dokumentira zgodovino vsebine elektronskega vira npr. vir, prenos odgovornosti, dejanja pri ohranjanju in posledice;
 - d. **informacije o pristnosti** - dokumentira mehanizme za dokazovanje pristnosti informacij o vsebini in njene nespremenljivosti (*checksum*, digitalni podpis...).
2. **informacije o vsebini**, ki vsebujejo:
 - a. **informacije o informacijskem objektu** (elektronskem viru);
 - b. **informacije o načinu prikazovanja vsebine** (*Representation information - RI*) (struktura in semantika).

Poleg omenjenih sta še dva informacijska objekta:

3. **informacije o pakiranju**, ki povedo, kako je spravljen elektronski vir v arhivu in
4. **deskriptivne informacije**, ki nam omogočajo lokalizacijo vira v arhivu.

Vsak informacijski objekt je nastal v okviru določene **baze znanja**. Na primer, program v Javi je nastal s pomočjo poznavanja programskega jezika Java. Ker je to znanje zelo splošno, nam ne more natančno povedati, kaj je ohranjeno v določenem informacijskem objektu. Zato so potrebne dodatne informacije, ki nam pomagajo, da razumemo podatke, ki so shranjeni v digitalni obliki. (npr. da je številka 3548-4995 ISSN številka). Zato so informacijski objekti sestavljeni iz **podatkovnega objekta** in **informacije o načinu prikazovanja vsebine** (RI) (Slika 1).



Slika 1: Informacijski model OAIS (Vir: OCLC/RLG..., 2002, Slika 1, str. 6)

Informacijski model zelo na splošno določa tipe informacij ali metapodatkov, ki so potrebni za trajno zagotavljanje dostopa do skladiščenih informacij v elektronskem arhivu. Model se ne spušča v specifikacijo tehnologije, praktičnih standardov in metapodatkov. To je prepuščeno posamičnemu arhivu, glede na vsebino gradiva in funkcijo, ki jo ta ima v svojem okolju.

3 Pregled tujih rešitev

Model OAIS je dal pomembno osnovo pri izgradnji in razvoju elektronskih arhivov. V nekaterih nacionalnih knjižnicah so pripravili različne metapodatkovne okvire, ki so bili pogojeni z njihovimi specifičnimi pristopi k zbiranju elektronskih virov. Med temi so najbolj znane rešitve projekta NEDLIB, delovne skupine RLG (*Research Libraries Group*) in OCLC, projekta CEDARS ter izkušnje nacionalnih knjižnic Avstralije in Nove Zelandije. V nadaljevanju so povzete omenjene izkušnje.

3.1 Projekt NEDLIB (1998 - 2000)

Namen projekta NEDLIB je bil izdelati funkcionalni model depozitnega sistema elektronskih publikacij (*Deposit System for Electronic Publications - DSEP*) (Steenbakkens, 2000). Ta sistem je vzel za izhodišče model OAIS in ga dopolnil z dodatnimi procesi, ki se odvijajo v depozitnih knjižnicah, kot so nabava, katalogizacija, poizvedovanje itd.

DSEP operira z istimi identitetami, kot referenčni model OAIS, celo uporablja enako terminologijo. Ena od nalog na projektu je bila določiti minimalno število metapodatkov, ki so potrebni za trajno ohranjanje elektronskih virov. Izhodišče za izbiro metapodatkov so informacije o načinu prikazovanja vsebine (*RI*) in informacije, pomembne za trajno ohranjanje (*PDI*).

Ker trajno ohranjanje pomeni omogočanje trajnega dostopa do binarnih informacij, shranjenih na nekem nosilcu, so metapodatki o načinu prikazovanja vsebine (*RI*) tesno povezani z informacijami o tehničnih zahtevah za prikazovanje vsebine elektronskih publikacij in omogočanje njihovega upravljanja. Komponente, ki se morajo upoštevati v informacijah o načinu prikazovanja vsebine (*RI*) so razvidne iz informacijskega modela, ki je sestavljen iz petih plasti ali nivojev (Lupovici in Masanès, 2000). Vsak nivo ustreza določenemu vidiku elektronskega vira, in sicer:

fizični nivo - predstavlja standardizirani fizični nosilec zapisa, do katerega lahko dostopamo s pomočjo informacijske tehnologije (hardvera). Metapodatki, ki se nanašajo na ta del naj bi se hranili v okviru informacij o provenienci.

Metapodatki o hardveru so smiselni le v primeru računalniških programov, ki potrebujejo dodatna navodila ali v primeru multimedijskih aplikacij, ki uporabljajo dodatne naprave. V svojem poročilu Lupovici in Masanès (2000, str. 11) navajata, da bodo informacije o operacijskih sistemih čez 20-50 let neuporabne in da je že v okviru katalogizacijskih pravil nesmiselno navajanje minimalnih zahtev za delovanje določene aplikacije, ker vemo, da se tehnologija nenehno razvija;

binarni nivo je neodvisen od nosilca in se nanaša na logični binarni zapis in določeno organizacijo tega zapisa v blokih ali datotekah. V tem nivoju delujejo operacijski sistemi, ki točno "poznajo" datotečni sistem. Metapodatki vsebujejo informacije o tipu operacijskega sistema;

nivo strukture - biti so zbrani v primitivne strukture podatkov, katere upravljajo programi. Prevajalniki (kompajlerji) in predvajalniki (*interpreters*) programskih jezikov omogočajo funkcionalnost na tem nivoju. Metapodatki o teh so potrebni, če ohranjamo programske aplikacije, ki so se prenehale razvijati;

nivo objekta - podatki so strukturirani v objekte, ki imajo pomen za aplikacijo in tudi za uporabnika. Metapodatki nam dajejo informacijo o formatu tega objekta;

nivo aplikacije - programske aplikacije (softver) upravljajo z objekti iz prejšnjega nivoja in jih predstavljajo uporabnikom. Metapodatki nam dajejo informacijo o programski aplikaciji in njeni različici za dostop do določenega formata objekta.

Na projektu NEDLIB so ugotovili, da obstajata dva tipa elektronskih publikacij in dokumentov, za katere potrebujemo različne tipe metapodatkov, in sicer:

- dokumenti, ki delujejo na aplikacijah, ki so odvisne od specifičnih sistemov
 - pri tem so obvezni metapodatki o aplikacijskem in binarnem nivoju ter
- dokumenti, ki so kreirani v formatih, neodvisnih od specifičnih sistemov (XML, JPEG..) - v tem primeru so metapodatki o nivoju objekta nujno potrebni.

Poleg informacij o načinu prikazovanja vsebine (RI) upoštevajo tudi informacije, pomembne za trajno ohranjanje (PDI). Ker pa so se v projektu NEDLIB predvsem osredotočili na tehnološki vidik trajnega ohranjanja, te informacije delijo na referenčne informacije, informacije o pristnosti (*fixity information*) in informacije o zgodovini sprememb, ki združujejo tako informacije o provenienci kot informacije o kontekstu.

Na podlagi strukture metapodatkov so na projektu NEDLIB opredelili minimalni nabor, ki je potreben za trajno ohranjanje elektronskih virov (Lupovici in Masanès, 2000):

1. METAPODATKI O NAČINU PRIKAZOVANJA (RI):

1.1. SPECIFIČNE ZAHTEVE HARDVERA (HW)- opis nestandardnih konfiguracij platform ali zahtev HW:

- 1.1.1. specifične zahteve mikroprocesorja;
- 1.1.2. specifične multimedijske zahteve;
- 1.1.3. specifične periferne zahteve.

1.2. OPERACIJSKI SISTEM:

- 1.2.1. ime;
- 1.2.2. različica.

1.3. PREDVAJALNIK (*interpreter*) ALI PREVAJALNIK (kompajler):

- 1.3.1. ime;
- 1.3.2. različica;
- 1.3.3. navodila.

1.4. FORMAT OBJEKTA:

- 1.4.1. ime;
- 1.4.2. različica.

1.5. APLIKACIJA:

- 1.5.1. ime;
- 1.5.2. različica.

2. METAPODATKI, POMEMBNI ZA TRAJNO OHRANJANJE VSEBINE (PDI):

2.1. REFERENČNE INFORMACIJE:

- 2.1.1. kreator;
- 2.1.2. naslov;
- 2.1.3. datum kreiranja;
- 2.1.4. založnik;
- 2.1.5. identifikator:
 - 2.1.5.1. vrednost;
 - 2.1.5.2. način oblikovanja;
 - 2.1.5.3. odgovorna agencija.
- 2.1.6. URL:
 - 2.1.6.1. vrednost;
 - 2.1.6.2. datum vpisa.

2.2. INFORMACIJA O PRISTNOSTI:

- 2.2.1. checksum:
 - 2.2.1.1. vrednost;
 - 2.2.1.2. algoritem.
- 2.2.2. digitalni podpis.

2.3. ZGODOVINA SPREMEMB:

- 2.3.1. pomembnejše spremembe izvedene v digitalnem objektu:
 - 2.3.1.1. datum;
 - 2.3.1.2. stara vrednost;
 - 2.3.1.3. nova vrednost;
 - 2.3.1.4. orodje, s katerim so bile izvedene spremembe:
 - 2.3.1.4.1. ime;
 - 2.3.1.4.2. različica.
 - 2.3.1.5. opis postopka za povrnitev stare vrednosti.
- 2.3.2. metapodatki o drugih spremembah:
 - 2.3.2.1. stara vrednost;
 - 2.3.2.2. nova vrednost.

3.2 Avstralska nacionalna knjižnica (1999-2002)

Pri Avstralcih je pristop k arhiviranju elektronskih virov zelo selektiven. To je razumljivo, če upoštevamo, da je v tako širokem jezikovnem področju skorajda nemogoče uporabiti kriterije izčrpnosti, ki veljajo za manjše dežele, kot je denimo Slovenija. V ZDA pa že obstajajo podjetja (*Internet Archive*), ki neselektivno zbirajo vse, kar je dosegljivo na svetovnem spletu v angleščini.

V avstralski nacionalki se o vsakem viru odloča posebej. Določeno število strokovnjakov temeljito pretehta potrebo po ohranjanju ali ne. Zato je zelo pomembno, kdo je odgovoren za določeno odločitev in zakaj se je za njo opredelil. Zani-

mivo je, da Avstralci ločijo koncept dela (*work*), ki je stalnica in se nanaša na originalne izvode elektronskih virov, od njegove manifestacije, ki je odvisna od strategije ohranjanja oz. načinov dostopa in prikazovanja, glede na razpoložljivo tehnologijo. Pri tem se odločajo ali ohraniti le originalno delo ali tudi njegovo(e) manifestacijo(e).

Pri elektronskih virih ločijo tri nivoje granularnosti objekta: nivo datoteke, nivo logičnega objekta (zbir datotek, ki skupaj prikazuje eno celoto) in nivo zbirke (zbir samostojnih in nepovezanih datotek). Za posamezne zvrsti datotek, ki imajo skupne lastnosti, so predvideni dodatni metapodatki. To so slike, zvočni posnetki, premične slike, besedila, baze in sistemske ali programske datoteke. Zgodovina sprememb se lahko beleži v enem zapisu ali pa se za vsako spremembo odpre nov zapis.

Avstralci priporočajo 24 zvrsti metapodatkov za trajno ohranjanje elektronskih virov. Ti se lahko nanašajo na vse ali le nekatere nivoje objekta (Preservation metadata..., 1999):

- 1. STALNI IDENTIFIKATORJI;**
- 2. DATUM KREIRANJA;**
- 3. TIP STRUKTURE** (tip datoteke ali objekta, ki je lahko slika, zvočni posnetek, besedilo, multimedijaska datoteka...);
- 4. TEHNIČNA INFRASTRUKTURA KOMPLEKSNEGA OBJEKTA** - vsebuje strukturo oz. formate datotek v njegovi sestavi;
- 5. OPIS DATOTEKE** - za vsak tip datoteke (slike, zvočni posnetki, premične slike, besedila, baze in sistemske ali programske datoteke) so predvideni določeni atributi, ki jih opisujemo s pomočjo metapodatkov;
- 6. ZNANE SISTEMSKÉ ZAHTEVE** - operacijski sistem oz. okolje v katerem lahko dostopamo do vira oz. v katerem ta deluje;
- 7. ZAHTEVE ZA INSTALACIJO** - katerikoli specializiran postopek, ki je potreben za instalacijo;
- 8. INFORMACIJA O SKLADIŠČENJU** - velikost datoteke ali vira in lokacija v skladišču;
- 9. OMEJEVALCI DOSTOPA** - vse, kar lahko omejuje dostop (enkriptika ali digitalni podpis) in ki lahko vpliva na strategije za trajno ohranjanje;
- 10. ORODJA ZA ISKANJE IN POIZVEDOVANJE TER OLAJŠEVALCI DOSTOPA** (gesla) - s časom je potrebno tudi ta orodja vzdrževati, da bi se ohranil dostop do vira;
- 11. DOVOLJENJE ZA DEJANJA, KI SO POTREBNA ZA TRAJNO OHRANJANJE** - dovoljenje imetnikov avtorskih pravic za reprodukcijo, migracijo ali konverzijo vira z namenom trajnega ohranjanja;
- 12. PREVERJANJE PRISTNOSTI ALI VELJAVNOSTI VIRA** (*Validation*) - sistem,

- ki nam omogoča preverjanje, ali je originalni izvod bil spremenjen ali ne. Na nivoju datoteke se zelo uveljavljajo internetni *checksum*;
13. **POVEZAVE** - kažejo na to, ali je določena manifestacija vira v povezavi z drugimi oz. ali je določen vir del zbirke ali ne;
 14. **IZGUBE IN NAPAKE** - vsaka izguba v funkcionalnosti ali izgledu;
 15. **ODLOČITVE O ARHIVIRANJU DELA** - odločanje o tem, ali naj se arhivira, datum odločitve in datum predvidene revizije odločitve;
 16. **RAZLOGI ZA ODLOČITEV O ARHIVIRANJU DELA;**
 17. **INSTITUCIJA, KI JE ODGOVORNA ZA ODLOČANJE O ARHIVIRANJU DELA;**
 18. **ODLOČITEV O ARHIVIRANJU MANIFESTACIJE** - odločanje o tem, ali naj se arhivira vir oz. do kdaj naj stoji v zalogi ter datum odločitve in njene revizije;
 19. **RAZLOGI ZA ODLOČITEV O ARHIVIRANJU MANIFESTACIJE;**
 20. **INSTITUCIJA, KI JE ODGOVORNA ZA ODLOČANJE O ARHIVIRANJU MANIFESTACIJE;**
 21. **PREDVIDENA UPORABA DOLOČENE MANIFESTACIJE;**
 22. **INSTITUCIJA, ODGOVORNA ZA TRAJNO OHRANJANJE;**
 23. **PROCES** - vsi relevantni podatki o kakršnemkoli procesu, ki je bil izveden nad digitalnim objektom ali datoteko;
 24. **KREATOR ZAPISA;**
 25. **DRUGO** - katerakoli druga relevantna informacija o elektronskem viru.

Kot je razvidno iz tega seznama, se Avstralci osredotočajo na informacije izven sistema, ki jih potrebujemo za upravljanje s trajnim ohranjanjem. Metapodatki so enako uporabni za elektronske vire, ki so nastali v elektronski obliki, kakor za digitalne kopije fizičnih originalov. Ne ukvarjajo se s procesi, ki so potrebni za trajno ohranjanje (migracija ali emulacija). Seznam metapodatkov je neodvisen od tehnologije. Njihov sistem za arhiviranje ne izhaja povsem iz OAIS.

3.3 Projekt CEDARS (*CURL (The Consortium of University Research Libraries) Exemplars in Digital ARchiveS*) (april 1998 - marec 2002)

CEDARS je med prvimi projekti, ki je raziskoval strategije za trajno ohranjanje elektronskih virov. Glavni cilji projekta so bili (*Cedars Guide to preservation metadata, 2002*):

- opozoriti na pomen trajnega ohranjanja elektronskih virov;
- identificirati, dokumentirati in razširjati strateški okvir za razvoj upravljal-

- skih politik, ki imajo namen trajno ohranjanje elektronskih virov;
- raziskati, dokumentirati in promovirati metode za trajno ohranjanje različnih zvrsti elektronskih virov.

Za razliko od Avstralcev, nivo agregacije oz. granularnosti ni pomemben za določanje metapodatkov za trajno ohranjanje. To pomeni, da je vseeno ali je elektronski vir zbirka, objekt ali datoteka.

Trdijo, da odločitve o ohranjanju ne smejo biti vezane na razpoložljivost metapodatkov. So dokumenti, ki so zelo dragoceni, kljub temu, da nimamo informacije o avtorju ali naslovu.

Projekt CEDARS temelji na referenčnem modelu OAIS. Uporablja enako terminologijo pri določanju informacijskih objektov. Tako kot v OAIS so informacije o vsebini pakirane skupaj z referenčnimi informacijami.

Metapodatkovni elementi so določeni na nivoju informacijskega paketa (The Cedars Project Team & UKOLN, 2000):

1. INFORMACIJSKI PAKET

1.1. INFORMACIJE, POMEMBNE ZA TRAJNO OHRANJANJE (PDI):

1.1.1. REFERENČNE INFORMACIJE:

1.1.1.1. OPIS VIRA (katalogizacija v obstoječem sistemu)

1.1.1.2. OBSTOJEČI METAPODATKI (ki so že bili ustvarjeni v katalogu ali kjerkoli):

1.1.1.2.1. OBSTOJEČI ZAPISI O VIRU.

1.1.2. INFORMACIJE O KONTEKSTU:

1.1.2.1. POVEZAVE Z DRUGIMI INFORMACIJSKIMI OBJEKTI.

1.1.3. INFORMACIJE O PROVENIENCI:

1.1.3.1. ZGODOVINA NASTANKA:

1.1.3.1.1. RAZLOG ZA NASTANEK;

1.1.3.1.2. ZGODOVINA SKLADIŠČENJA;

1.1.3.1.3. ZGODOVINA SPREMEMB PRED ARHIVIRANJEM;

1.1.3.1.4. ORIGINALNO TEHNIČNO OKOLJE:

1.1.3.1.4.1. POGOJI ZA DELOVANJE (hardver/software);

1.1.3.1.4.2. POSTOPKI (instalacija opreme ali drugo);

1.1.3.1.4.3. DOKUMENTACIJA (priročniki...).

1.1.3.1.5. RAZLOGI ZA OHRANJANJE.

1.1.3.2. ZGODOVINA UPRAVLJANJA IN ROKOVANJA:

1.1.3.2.1. ZGODOVINA PROCESA VNOSA V SISTEM;

1.1.3.2.2. ZGODOVINA ADMINISTRIRANJA:

1.1.3.2.2.1. ZGODOVINA DEJANJ;

1.1.3.2.2.2. ZGODOVINA POLITIKE, KI VPLIVA NA DOLOČENE ODLOČITVE.

1.1.3.3. UPRAVLJANJE Z AVTORSKIMI PRAVICAMI:

1.1.3.3.1. ZGODOVINA POGAJANJ;

1.1.3.3.2. INFORMACIJE O PRAVICAH.

1.1.3.3.2.1. OPREDELITEV AVTORSKIH PRAVIC :

1.1.3.3.2.1.1. IME ZALOŽNIKA;

1.1.3.3.2.1.2. DATUM OBJAVE;

1.1.3.3.2.1.3. MESTO OBJAVE;

1.1.3.3.2.1.4. OPOZORILO NA OMEJITEV
DOSTOPA;

1.1.3.3.2.1.5. KONTAKTNE OSEBE ALI
IMETNIKI AVTORSKIH PRAVIC.

1.1.3.3.2.2. UPORABNIKI (kdo lahko dostopa do
elektronskega vira);

1.1.3.3.2.3. DEJANJA, KI SO DOVOLJENA:

1.1.3.3.2.3.1. DOVOLJENA DEJANJA NA
PODLAGI OBSTOJEČEGA
AVTORSKEGA PRAVA:

1.1.3.3.2.3.1.1. POVEZAVA NA
ZAKONSKO BESEDILO;

1.1.3.3.2.3.2. DOVOLJENA DEJANJA NA
PODLAGI LICENČNIH
POGODB:

1.1.3.3.2.3.2.1. POVEZAVA NA
BESEDILO LICENCE
ALI DRUGE POGODBE.

1.1.4. INFORMACIJA O PRISTNOSTI (avtentičnosti ali nespremenljivosti):

1.1.4.1. INDIKATOR PRISTNOSTI.

1.2. INFORMACIJA O VSEBINI:

1.2.1. INFORMACIJE O NAČINU PRIKAZOVANJA VSEBINE (RI):

1.2.1.1. INFORMACIJA O STRUKTURI:

1.2.1.1.1. KRATEK OPIS OBLIKE ;

1.2.1.1.2. OBJEKTI, KI OMOGOČAJO BRANJE
DOKUMENTA (*transformer objects*):

1.2.1.1.2.1. PLATFORMA;

1.2.1.1.2.2. PARAMETRI;

1.2.1.1.2.3. STROJI ZA PRIKAZOVANJE/ANALIZO
(softver);

1.2.1.1.2.4. IZHODNI FORMAT;

1.2.1.1.2.5. VNOSNI FORMAT.

1.2.1.1.3. OBJEKTI ZA PRIKAZOVANJE/ANALIZO/
KONVERZIZIJO.

1.2.1.2. SEMANTIČNE INFORMACIJE:

1.2.1.2.1. OBJEKTI ZA PRIKAZOVANJE/ANALIZO.

1.2.2. PODATKOVNI OBJEKT.

3.4 Delovna skupina OCLC (Online Computer Library Center, Inc.)/RLG (Research Library Group) (marec 2000-junij 2002)

Namen delovne skupine OCLC in RLG je bil določiti en univerzalni nabor metapodatkov za trajno ohranjanje. Za osnovo so vzeli dotedanje definicije metapodatkov za trajno ohranjanje, ki so nastale v okviru ali pod vplivom referenčnega modela OAIS, in oblikovali sintezo na podlagi štirih rešitev, in sicer: CEDARS, Nacionalne knjižnice Avstralije, NEDLIB in OCLC. Sintezo so dodatno dogradili in dodelali tako, da so uporabili entitete iz informacijskega modela OAIS, jih natančneje specifikirali in v drevesni hierarhiji opredelili različne tipe informacij, na podlagi katerih se ustvarjajo metapodatkovni elementi. Njihov predlog strukture metapodatkov je kar obsežen (OCLC/RLG Working Group on Preservation Metadata, 2002):

1. INFORMACIJE O VSEBINI:

1.1. VSEBINSKI PODATKOVNI OBJEKT.

1.2. INFORMACIJA O NAČINU PRIKAZOVANJA (RI):

1.2.1. OPIS VSEBINSKEGA PODATKOVNEGA OBJEKTA:

- o KRATEK OPIS VSEBINE (npr. datotečni sistem, povezave in kako jih lahko odpremo);
- o TIP STRUKTURE (slika, besedilo, podatkovna baza, zvočni zapis...);
- o TEHNIČNA INFRASTRUKTURA KOMPLEKSNIH OBJEKTOV (naštevane različnih formatov datotek, ki sestavljajo vir);
- o OPIS DATOTEK (tehnične specifikacije vsake datoteke);
- o ZAHTEVE IN POSTOPKI ZA INSTALACIJO;
- o VELIKOST OBJEKTA (v bitih);
- o ZAVIRALCI DOSTOPA (enkriptika, gesla in dr.);
- o OLAJŠEVALCI DOSTOPA (sistemi ali metode, ki nam omogočajo dostop do podatkovnega objekta, ki ga je treba s časom vzdrževati ali ne);
- o POMEMBNE LASTNOSTI PRI PRIKAZOVANJU PODATKOV (barve, oblika, zvok, struktura...);
- o FUNKCIONALNOST (način delovanja ali prikazovanja);
- o OPIS PRIKAZANE VSEBINE (kako se mora vsebina pregledati in kako jo morajo razumeti uporabniki - vsebuje tudi razlago nejasnih podatkov, definicije in opis podatkovnih struktur);
- o POMANJKLJIVOSTI IN NAPAKE - izgube v funkcionalnosti ali sprememba v prikazovanju;
- o DOKUMENTACIJA - vsa dokumentacija, potrebna za prikazovanje in razumevanje vsebine (priročniki, geslovniki...).

1.2.2. OPIS OKOLJA DELOVANJA:

o OKOLJE SOFTVERA:

- APLIKACIJE ZA PRIKAZOVANJE:

- PROCES TRANSFORMACIJE - softver, potreben za transformacijo bitnega toka v razumljivo vsebino:

- ORODJA ZA TRANSFORMACIJO:

- PARAMETRI;

- FORMAT VNOSA;

- IZHODNI FORMAT;

- LOKACIJA ORODJA ZA TRANSFORMACIJO V SISTEMU;

- DOKUMENTACIJA.

- APLIKACIJA ZA PRIKAZOVANJE/DOSTOP:

- VNOSNI FORMAT;

- IZHODNI FORMAT;

- LOKACIJA APLIKACIJE V SISTEMU;

- DOKUMENTACIJA.

- OPERACIJSKI SISTEM:

- IME;

- RAZLIČICA;

- LOKACIJA V SISTEMU;

- DOKUMENTACIJA.

o OKOLJE HARDVERA:

- LOKACIJA.

o RAČUNALNIŠKI VIRI:

- ZAHTEVE MIKROPROCESORAJA;

- ZAHTEVE POMNILNIKA;

- DOKUMENTACIJA.

o SKLADIŠČENJE:

- INFORMACIJA O SKLADIŠČENJU;

- DOKUMENTACIJA.

o PERIFERNE ENOTE:

- ZAHTEVE PERIFERNIH ENOT;

- DOKUMENTACIJA.

2. INFORMACIJE POMEMBNE ZA TRAJNO OHRANJANJE (PDI):

2.1. REFERENČNE INFORMACIJE:

2.1.1. IDENTIFIKACIJA ARHIVSKEGA SISTEMA

(LOKALNI IDENTIFIKATOR):

- VREDNOST;

- METODA USTVARJANJA;

- ODGOVORNA INSTITUCIJA.

2.1.2. SPLOŠNA IDENTIFIKACIJA (bibliografski identifikatorji, ki jih dodelijo uradne agencije):

- VREDNOST;
- METODA USTVARJANJA;
- ODGOVORNE AGENCIJE.

2.1.3. OPIS ELEKTRONSKEGA VIRA:

- OBSTOJEČI METAPODATKI;
- PRIMER ZAPISA.

2.2. INFORMACIJE O KONTEKSTU:

2.2.1. RAZLOGI ZA USTVARJANJA ELEKTRONSKEGA VIRA.

2.2.2. POVEZAVE:

- MANIFESTACIJA:
 - TIP POVEZAVE;
 - IDENTIFIKACIJA.
- INTELEKTUALNA VSEBINA:
 - TIP POVEZAVE;
 - IDENTIFIKACIJA.

2.3. INFORMACIJE O PROVENIENCI:

2.3.1. KAKO JE VIR BIL USTVARJEN (skeniran ali elektronsko ustvarjen ali drugo);

2.3.2. ZGODOVINA VIRA OD NASTANKA DO VNOSA V ARHIV;

2.3.3. OPIS PROCESA VNOSA VIRA V ARHIV;

2.3.4. OPIS PROCESOV PO VNOSU VIRA V ARHIV (vzdrževanje, spremembe v vsebini in druge posledice trajnega ohranjanja);

2.3.5. AVTORSKE PRAVICE;

2.3.6. DEJANJA, KI SO VPLIVALA NA VIR:

- IME DEJANJA;
- OPIS POSTOPKA;
- DATUM;
- ODGOVORNA AGENCIJA;
- REZULTATI POSTOPKA;
- OPOMBE.

2.4. INFORMACIJE O PRISTNOSTI:

2.4.1. DOKAZ PRISTNOSTI:

- METODA ALI SISTEM;
- POSTOPEK;
- DATUM;
- REZULTAT.

3.5 Novozelandska nacionalna knjižnica

Njihova strategija trajnega ohranjanja temelji na študiji delovne skupine OCLC/RLG iz leta 2001, vendar je njihov pristop veliko bolj pragmatičen in je neposredno povezan s prakso.

Za Novozelandce so metapodatki za trajno ohranjanje elektronskih virov zelo pomembni (Metadata standards framework, 2002). V njihovih smernicah in standardnih okvirjih, za razliko od drugih nacionalk, zelo jasno ločijo med originalnim izvodom elektronskega vira (*preservation masters*) in drugimi različicami, ki so prirejene za dostop oz. za širjenje ali upravljanje. Predmet njihovega ohranjanja je originalni izvod, ki se s časom lahko tudi spreminja ter prvotni original (pr. originalna izdaja na CD-ROM-u), ki je lahko tudi neberljiv zaradi razvoja informacijske tehnologije. Poleg tega imajo pri odločanju o ohranjanju prednost viri, ki so nastali v elektronski obliki, pred drugimi, ki predstavljajo konverzijo iz analognih oz. fizičnih oblik v digitalno (npr. skenirane slike ali besedila).

Po zgledu na avstralsko nacionalno knjižnico tudi Novozelandci upoštevajo stopnjo granularnosti digitalnih objektov in jih delijo na **enostavne**, ki so sestavljeni le iz ene datoteke; **kompleksne**, sestavljeni iz skupine nesamostojnih datotek, ki skupaj predstavljajo en logični objekt (npr. spletna stran ali podatkovna baza); in **skupine objektov**, sestavljeni iz skupine samostojnih datotek, ki niso odvisne ena od druge (npr. disketa z dvestotimi pismi znanega pisatelja v WordPerfectu in Wordu).

Njihov model zajema štiri entitete, ki so medsebojno povezane:

- **Objekt** - se nanaša le na originalni izvod logičnega objekta, ki je lahko enostavni, kompleksni objekt ali skupina objektov. Objekt definira 18 zvrsti metapodatkov, ki so pomembni za trajno ohranjanje. Med njimi so deskriptivni metapodatki, metapodatki, ki opisujejo tehnične zahteve za dostop do vira oz. za njegovo delovanje, informacijsko-tehnološko okolje, funkcionalnost vira in način dostopa.
- **Proces** - predstavlja katerokoli dejanje izvedeno nad originalnim izvodom elektronskega vira in se evidentira s pomočjo 13 elementov ali metapodatkov. Med njimi je tudi razlog, zakaj in kdaj je bil izveden proces, kdo ga je odobril, kako je bil izveden, opis opreme, ki je pri tem bila uporabljena in nazadnje še rezultati teh sprememb.
- **Datoteka** - vsebuje tehnične informacije o najnižjih nivojih digitalnih objektov. Devet metapodatkovnih elementov (npr. ime datoteke, ekstenzija, datum kreiranja, format...) je skupnih vsem tipom datotek. Specifične lastnosti se nanašajo na slike, zvočne posnetke, premične slike, in besedila.
- **Spremembe metapodatkov** - evidentira vse spremembe, ki so bile izvedene na zapisu metapodatkov s pomočjo petih elementov.

Metapodatke za trajno ohranjanje lahko prevzemamo tudi avtomatsko ali polavtomatsko, kajti mnogi podatki se že avtomatsko generirajo v trenutku skladiščenja v arhivski sistem.

Iz navedenih primerov je videti, da so Avstralci in Novozelandci veliko bolj pragmatični pri oblikovanju metapodatkov za trajno ohranjanje. Oboji dajejo večji pomen različnim nivojem granularnosti virov. Projekt NEDLIB, CEDARS in rezultati delovne skupine RLG/OCLC so pokazali, da je možna povsem različna interpretacija entitet referenčnega modela OAIS, vendar so elementi, ki jih nikakor ne smemo prezrti.

4 Rezultati raziskave v okviru projekta ASEP

Konec leta 2002 je Narodna in univerzitetna knjižnica v sodelovanju z Institutom "Jožef Stefan" začela z izvajanjem projekta "Metodologija zbiranja in arhiviranja slovenskih elektronskih publikacij na medmrežju" (ASEP). Projekt sofinancirajo Ministrstvo za šolstvo, znanost in šport, Ministrstvo za kulturo in Ministrstvo za informacijsko družbo v okviru Ciljnega raziskovalnega programa "Konkurenčnost Slovenije 2001-2006", njegov zaključek je predviden konec leta 2004. Namen projekta je izdelati izčrpno metodologijo arhiviranja elektronskih publikacij na medmrežju. Ena od pomembnih nalog je bila tudi izdelava osnovnega nabora metapodatkov za opis elektronskega gradiva, ki naj bi se ohranjalo za prihodnost. Pri tem smo skušali odgovoriti na naslednja vprašanja:

- kateri od navedenih metapodatkov so bistvenega pomena za trajno ohranjanje elektronskih virov in morajo biti vključeni v arhivski sistem;
- kateri od teh metapodatkov se lahko vnašajo v okviru bibliografske obdelave;
- kateri od navedenih metapodatkov se lahko generirajo sistemsko;
- katere metapodatke dodatno potrebujemo in niso nikjer zajeti;
- ali naj upoštevamo nivo granularnosti elektronskih virov, na katerem naj bi se ustvarjali metapodatki (datoteka, logični objekt ali zbirka)
- ali naj se metapodatki skladiščijo ločeno od elektronskega vira, se "spakirajo" v arhivski informacijski paket.

Pri raziskavi smo izhajali iz zgoraj opisanih rešitev in obenem upoštevali obstoječo prakso na področju tradicionalne bibliografske obdelave gradiva ter zmožnosti systemskega generiranja metapodatkov.

Pri vprašanju ali naj se metapodatki za trajno ohranjanje integrirajo oz. "spakirajo" v arhivski informacijski paket je bilo ugotovljeno, da je veliko bolje, če jih

ohranimo ločeno od informacijskega objekta (elektronskega vira). Na ta način bomo poenostavili in pospešili poizvedovanje po metapodatkih, preko katerih lahko nato dostopamo do samega vira. Vendar ni izključena možnost integriranja teh informacij v AIP, kot priporoča standard OAIS.

Pri bibliografski obdelavi elektronskih virov je razširjeni format za bibliografski opis COMARC. Pričakujemo, da se bo kmalu dopolnil s spremembami iz 3. in 4. izdaje UNIMARC-a. Kljub temu, da ni namen teh formatov opis gradiva za trajno ohranjanje, da ne bi podvajali vnosa podatkov je smiselno uporabiti že obstoječa polja (predvsem polja 135 in 856) in tem dodati nova.

Pri klasifikaciji metapodatkov smo se zgledovali po rešitvah v okviru projekta NEDLIB, ki enostavno in na zelo sistematičen način pokriva vse potrebe po informacijah za trajno ohranjanje elektronskega vira v arhivskem sistemu. Delitev je narejena na podlagi funkcije metapodatkov v referenčnem modelu OAIS, in sicer: metapodatki, ki vsebujejo informacije o informacijskem objektu in načinu njegovega prikazovanja (RI) ter metapodatki, ki vsebujejo informacije, pomembne za trajno ohranjanje vsebine (PDI). Standard omenja (Reference Model..., 2002, str. 2-6), da preden začnemo z definicijo informacij, ki so pomembne za trajno ohranjanje (PDI), moramo jasno definirati informacije o vsebini. Informacijski model OAIS (Reference Model..., 2002, priloga E) loči več ravni digitalnega objekta (elektronskega vira), ki ustrezajo različnim vidikom informacijske tehnologije (hardver, binarna raven, raven strukture, raven objekta in raven aplikacije), kar smo upoštevali pri definiciji tehničnih lastnosti elektronskega vira. Pri metapodatkih za trajno ohranjanje pa upoštevamo delitev na referenčne informacije, informacije o kontekstu, informacije o provenienci in informacije o pristnosti.

V nadaljevanju navajamo predlog metapodatkov za trajno ohranjanje elektronskih virov:

1. METAPODATKI, KI VSEBUJEJO INFORMACIJE O NAČINU PRIKAZOVANJA VSEBINE (RI)

1.1. OPIS NAČINA PRIKAZOVANJA ELEKTRONSKEGA VIRA:

1.1.1. KRATEK OPIS VSEBINE, STRUKTURE IN KAKO LAHKO DOSTOPAMO DO ELEKTRONSKEGA VIRA;

1.1.2. POMEMBNE LASTNOSTI PRI PRIKAZOVANJU PODATKOV (pr. delovanje specifičnih aplikacij, kot je JavaScript, ki so vgrajeni v samem viru, barva slik, izgled tabel, formati besedil...);

1.1.3. OPIS FUNKCIONALNOSTI ELEKTRONSKEGA VIRA

(pr. delovanje zvočnih zapisov v multimedijem viru, delovanje iskalnih funkcij v podatkovni bazi.);

1.1.4. OPIS PRIKAZANE VSEBINE (kako se mora vsebina brati in kako jo morajo razumeti uporabniki vključno z razlago nejasnih pojmov);

1.1.5. SPREMEMBE V PRIKAZOVANJU IN FUNKCIONALNOSTI SKOZI ČAS.

1.2. OPIS TEHNIČNIH ZNAČILNOSTI IN OKOLJA DELOVANJA ELEKTRONSKEGA VIRA:

1.2.1. APLIKACIJSKA RAVEN:

- Ime aplikacije:

- različica;
- lokacija;
- dokumentacija (navodila za instalacijo).

- Druge aplikacije za delovanje/dostop do elektronskega vira:

- lokacija;
- dokumentacija (navodila za instalacijo).

1.2.2. RAVEN OBJEKTA:

- format elektronskega vira (besedilo, grafični prikaz, zvočni posnetek...);

- format kompleksnih objektov (število in tipi datotek - glasbeni zapisi, besedilo, slike...);

- Opis podrejenih tipov datotek;

1.2.3. RAVEN STRUKTURE (Informacija o prevajalnikih (kompajlerjih) in predvajalnikih (interpreterjih) v primeru, da softver ni bil kompiliran):

- ime prevajalnik/predvajalnika;

- različica;
- lokacija v sistemu;
- dokumentacija.

1.2.4. BINARNA RAVEN:

- operacijski sistem:

- ime;
- različica;
- lokacija v arhivskem sistemu;
- dokumentacija.

1.2.5. HARDVER:

- lokacija v sistemu.

- računalniški viri:

- zahteve mikroprocesorja;
- zahteve pomnilnika;
- dokumentacija.

- skladiščenje:

- informacija o skladiščenju;
- dokumentacija.

- periferne enote:

- zahteve perifernih enot;
- lokacija;
- dokumentacija.

1.2.6. INFORMACIJA O SKLADIŠČENJU:

- skupna velikost elektronskega vira (vseh datotek);
- lokacija v elektronskem arhivu;
- druge lokacije;
- informacija o datotekah:
 - velikost datoteke;
 - tip kompresije;
 - navodila za sistem za shranjevanje;
 - struktura datotek in lokacija v elektronskem arhivu;
 - specifične lastnosti datotek (barva, rezolucija, zvok...).

2. METAPODATKI, KI VSEBUJEJO INFORMACIJE, POMEMBNE ZA TRAJNO OHRANJANJE (PDI):

2.1. REFERENČNE INFORMACIJE:

2.1.1. OPIS VIRA (deskriptivni metapodatki iz bibliografskega opisa);

2.1.2. IDENTIFIKATORJI:

- identifikator arhivskega sistema (lokalni identifikator);
- bibliografski identifikator (ISSN, ISBN);
- drugi (DOI, URN,...).

2.2. INFORMACIJE O KONTEKSTU:

- razlog za ustvarjanje elektronskega vira;
- razlog za ohranjanje elektronskega vira;
- datum ustvarjanja;
- povezave:
 - z drugimi različicami istega elektronskega vira;
 - z drugimi enotami zbirke;
 - z nadrejenimi enotami zbirke;
 - s podrejenimi enotami zbirke.

2.3. INFORMACIJE O PROVENIENCI:

2.3.1. ZGODOVINA NASTANKA IN ARHIVIRANJA ELEKTRONSKEGA VIRA:

- kako je elektronski vir bil ustvarjen (je nastal v elektronski obliki ali je reprodukcija drugega vira);
- prvotno tehnično okolje (pogoji in način delovanja, dokumentacija o delovanju);
- opis postopka prenosa vira v sistem;
- zgodovina od prenosa v sistem dalje:
 - imena bibliografov in vzdrževalcev elektronskega vira;
 - opis posegov v elektronski vir (konverzije, migracije...);
 - rezultati posegov;
 - avtor posegov;
 - datum posegov;
 - opombe.

2.4. INFORMACIJE O PRISTNOSTI:

- 2.4.1. IMETNIK(I) AVTORSKIH PRAVIC;**
- 2.4.2. ČAS ZAŠČITE;**
- 2.4.3. IMENA OSEB/ORGANIZACIJ, KI SO ODGOVORNE ZA TRAJNO OHRANJANJE ELEKTRONSKEGA VIRA;**
- 2.4.4. DOVOLJEN NAČIN DOSTOPA** (npr. dostop na lokaciji, dostop na daljavo, dostop določenim osebam...);
- 2.4.5. OMEJITVE** (gesla, enkriptika...);
- 2.4.6. LICENCE;**
- 2.4.7. OSEBE/ORGANIZACIJE, KATERIM JE DOSTOP DOVOLJEN;**
- 2.4.8. DOVOLJENI POSEGI V ELEKTRONSKI VIR** (migracija, konverzija, reprodukcija v drugo varno lokacijo...);
- 2.4.9. IMENA OSEB/ORGANIZACIJ, KI JIM JE DOVOLJENO POSEGATI V ELEKTRONSKI VIR ZARADI TRAJNEGA OHRANJANJA;**
- 2.4.10. REFERENČNI DOKUMENTI** (dokumenti, ki potrjujejo pridobljene pravice za arhiviranje, dostop in posege v elektronski vir, kot so pogodbe o arhiviranju, licenčne pogodbe, veljavna zakonodaja in dr.);
- 2.4.11. DRUGI POMEMBNI PODATKI**
- 2.4.12. INDIKATOR PRISTNOSTI** (potrjuje originalnost ali pristnost elektronskega vira, oz. da se ta ni spremenil od vnosa v depozitni arhiv):
 - (*Checksum* (vrednost ali algoritem));
 - Digitalni podpis;
 - *Watermark*....
- 2.4.13. DRUGI INDIKATORJI:**
 - Metoda;
 - Postopek;
 - Datum;
 - Rezultat.

Možna je tudi rešitev, da se za natančen opis elektronskih virov z vsemi naštetimi metapodatki dogradi obstoječi format v lokalni bazi (npr. dodatna polja v bloku IX v COMARC-u), v okviru katere bi se omogočila spletna povezava do vira v elektronskem arhivu. Pri večjih sistemih pa je verjetno bolje, če se ti podatki vodijo ločeno v posebni bazi v okviru elektronskega depozitnega sistema, za katero bi skrbeli določeni kadri, da ne bi prišlo do nalaganja obsežnega števila informacij, ki so pomembne le iz vidika trajnega ohranjanja. Kompatibilnost z drugimi sistemi je priporočljiva zaradi izmenjave podatkov.

5 Zaključek

Metapodatki za trajno ohranjanje so zelo pomembni za ohranjanje dostopa do elektronskih virov v prihodnosti. Mnoge tuje organizacije so vložile precejšnja sredstva, da bi proučile optimalne nabore teh metapodatkov, ki naj bi se upoštevali pri obdelavi. Vsekakor je to eno od področij, kateremu bi morali v Sloveniji posvetiti večjo pozornost. Zavedamo se, da je raziskava na projektu ASEP le začetek in da bo čas natančneje povedal, ali smo bili na pravi poti. Skrb za prihodnost elektronske dediščine zahteva veliko več dela s strani bibliografov. Verjetno je tudi napačno pričakovati, da nam z obdelavo vse večjega števila elektronskih virov ne bo treba spreminjati dosedanje prakse. Verjetno se bomo morali vse bolj posvečati podatkom, ki z bibliografskega vidika do sedaj niso bili toliko pomembni. Za razumevanje le-teh bomo potrebovali bolj poglobljena znanja s področja informacijske tehnologije in računalništva, kar za slovensko bibliotekarsko stroko predstavlja velik izziv. Po drugi strani metapodatki za trajno ohranjanje ne smejo biti le skrb bibliografov, temveč tudi vseh organizacij in posameznikov, ki ustvarjajo in ohranjajo elektronske vsebine. Brez njihovega sodelovanja ne bo mogoče ohraniti slovenske elektronske kulturne dediščine za prihodnost.

6 Citirani viri

1. *Cedars Guide to preservation metadata* (2002). Pridobljeno 5. 5. 2004 s spletne strani: <http://www.leeds.ac.uk/cedars/guideto/metadata>
2. The Cedars Project Team and UKOLN (2000). *Metadata for digital preservation: the CEDARS project outline specification. Draft for public consultation*. Pridobljeno 5. 5. 2004 s spletne strani: <http://www.leeds.ac.uk/cedars/MD-STR~5.pdf>
3. Guilliland-Swetland, A. J. (2000). *Introduction to Metadata: Setting the Stage*. 12 str. (Pathways to Digital Information). Pridobljeno 5. 5. 2004 s spletne strani: <http://www.getty.edu/research/institute/standards/intrometadata>
4. Lupovici C. & Masanès J. (2000). *Metadata for Long Term Preservation*. Den Haag: Koninklijke Bibliotheek, 2000. (NEDLIB Report series ; 2) Pridobljeno 5. 5. 2004 s spletne strani: <http://www.kb.nl/coop/nedlib/results/NEDLIBmetadata.pdf>
5. *Metadata standards framework - preservation metadata* (2002). Wellington: National Library of New Zealand Pridobljeno 5. 5. 2004 s spletne strani: http://www.natlib.govt.nz/files/4initiatives_metaschema.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. 5. 2004 s spletne strani: http://www.oclc.org/research/projects/pmwg/pm_framework.pdf
7. *Preservation metadata for digital collections : exposure draft (1999)*. Canberra: National Library of Australia Pridobljeno 5. 5. 2004 s spletne strani: <http://www.nla.gov.au/preserve/pmeta.html>
 8. *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
 9. Steenbakkens, J. (2000). *The Nedlib Guidelines : Setting up a Deposit System for Electronic Publications*. Den Haag : Koninklijke Bibliotheek. (NEDLIB Report series ; 5). Pridobljeno 5. 5. 2004 s spletne strani: <http://www.kb.nl/coop/nedlib/results/NEDLIBguidelines.pdf>

Uporabljene kratice

AIP - Archival Information Package

ASEP - Projekt NUK/IJS: "Metodologija arhiviranja slovenskih elektronskih publikacij na medmrežju"

CEDARS - CURL Exemplars in Digital Archives

DIP - Dissemination information package

DSEP - Deposit System for Electronic Publications

NEDLIB - Networked European Deposit Library

OAIS - Open Archival Information System

OCLC - Online Computer Library Center

PDI - Preservation Description Information

RI - Representation Information

RLG - Research Libraries Group

SIP - Submission Information Package

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