

SVEUČILIŠTE U ZAGREBU
FILOZOFSKI FAKULTET
ODSJEK ZA INFORMACIJSKE I KOMUNIKACIJSKE
ZNANOSTI

Dina Mužina

**Modeli i sheme metapodataka za dugoročno očuvanje u
institucijskim repozitorijima: Implementacija PREMIS-a u
Islandora digitalnom repozitoriju**

Diplomski rad

Mentor: dr. sc. Goran Zlodi, doc.

Zagreb, 2017.

Sadržaj

1. Sažetak.....	4
Summary	5
2. Uvod	6
3. Dugoročno očuvanje	7
3.1. Informacijski i elektronički informacijski objekt.....	7
3.1.1. Fizička razina.....	7
3.1.2. Logička razina	8
3.1.3. Konceptualna razina	8
3.2. Obilježja ustanova koje rade na dugoročnom očuvanju	8
3.3. Prijetnje dugoročnom očuvanju	9
3.3.1. Tehnološke prijetnje.....	9
3.3.2. Organizacijske prijetnje	10
3.3.3. Financijske prijetnje.....	10
3.3.4. Osoblje.....	11
4. O metapodacima	11
4.1. Što su metapodaci?	11
4.2. Razvoj metapodataka	12
4.3. Vrste metapodataka	14
4.3.1. Administrativni metapodaci	14
4.3.2. Deskriptivni metapodaci.....	14
4.3.3. Metapodaci za zaštitu.....	15
4.3.4. Tehnički metapodaci	15
4.3.5. Metapodaci vezani uz korištenje.....	15
4.4. Važnost i upotreba metapodataka	15
5. O repozitorijima.....	19
5.1. Zašto postoje repozitoriji; obilježja i funkcije.....	19
5.2. Vrste repozitorija	20
5.2.1. Tematski repozitoriji.....	20
5.2.2. Istraživački repozitoriji.....	21
5.2.3. Nacionalni repozitoriji	21
5.2.4. Institucijski repozitoriji	21
6. O PREMIS-u.....	22
6.1. Nastanak i uspostava PREMIS-a	23
6.1.1. OAIS referentni model.....	25

6.1.2.	Aktivnosti vezane uz očuvanje	27
6.2.	PREMIS podatkovni rječnik.....	28
6.2.1.	Veze	30
6.2.2.	Semantičke jedinice.....	30
6.2.3.	Kontejneri, prošireni kontejneri	30
6.3.	PREMIS podatkovni model	32
6.3.1.	Intelektualni entiteti.....	33
6.3.2.	Objekti	33
6.3.3.	Događaji.....	35
6.3.4.	Agenti	36
6.3.5.	Prava.....	37
6.3.6.	Veze među objektima.....	38
6.4.	PREMIS verzija 3.0	39
6.5.	PREMIS u XML	40
7.	Islandora.....	41
7.1.	Drupal	42
7.2.	Fedora.....	43
7.3.	Apache Solr.....	43
7.4.	Primjena Islandore	44
8.	Zaključak.....	47
9.	Literatura.....	48
	Prilozi	51
	Prilog 1: Popis slika	51

1. Sažetak

Devedesetih godina dvadesetog stoljeća pojavljuje se pojam dugoročnog očuvanja. Taj pojam obuhvaća cijeli niz procesa od same organizacije očuvanja, odabira (digitalne ili digitalizirane) građe, financija, uspostave sustava, odabira i edukacije osoblja pa sve do odabira tehnologije. Zbog zahtjevnosti cijelog procesa i mnogih potencijalnih prijetnji, arhivska i bibliotekarska struka su te koje se najčešće odlučuju za tu aktivnost. Kad se govori o dugoročnom očuvanju, govori se o određenoj građi ili objektima koje želimo sačuvati. Obavljanje aktivnosti nad određenom građom ili objektom olakšavaju metapodaci jer oni daju niz strukturiranih informacija o njima. Međunarodna radna skupina nazvana PREMIS je 2005. godine objavila standard za metapodatke koji pomaže u očuvanju digitalnih objekata. Svojom širokom primjenom u zajednici dugoročnog očuvanja kao i fleksibilnošću ovaj se standard našao u raznim svjetskim projektima dugoročnog digitalnog očuvanja. Mjesta gdje se pohranjuju digitalni objekti nazivaju se repozitoriji. Iako se pohranjivati može bilo što, vrsta građe koja se najčešće odabire za pohranu i očuvanje su znanstveni i istraživački radovi, obrazovni materijali ili radovi vezani uz određenu instituciju. U radu se spominje i Islandora kao primjer programskog okvira koji pomaže institucijama, organizacijama i njihovim korisnicima u procesima pohrane i upravljanja digitalnim objektima. Cilj ovog rada je definirati ključne pojmove vezane uz dugoročno očuvanje te objasniti kako dolazi do procesa dugoročnog očuvanja s najvećim naglaskom na PREMIS standardu.

Ključne riječi: dugoročno očuvanje, metapodaci, repozitorij, PREMIS, Islandora

Summary

At the end of the twentieth century the concept of long-term preservation appeared. This concept includes the whole process from preservation management, selection of (born digital or digitized) material, financing, system establishment, choosing and educating the staff, to selecting the technology. This process is very challenging and has a lot of potential threats and because of this very few other than archivists and librarians decide to undertake this challenge. Mostly specific materials or objects are included in long-term preservation. Carrying out activities on a specific material or object is easier when there are metadata. Metadata are giving a lot of structured information about a material or object they are describing. An international working group called PREMIS published in 2005 a metadata standard which helps to preserve digital objects. It can be widely used for the long-term preservation purposes. It is very flexible so it can be found in various international long-term digital preservation projects. Places where digital objects are stored are called repositories. Although everything can be stored, scientific and research work, education materials and work linked to a specific institution are mostly chosen. Islandora is a software framework which helps institutions, organizations and their users during the storage process and management of digital objects. The aim of this thesis is to define key terms which are used with long-term preservation and to explain the whole long-term preservation process while placing emphasis on PREMIS standard.

Key words: long-term preservation, metadata, repository, PREMIS, Islandora

2. Uvod

Dvadeseto je stoljeće period razvitka i otkrića mnogih lijekova, izuma, strojeva, teorija i dr. pa i samog računala. S obzirom na sve veću količinu informacija koje su nastale zbog navedenih otkrića, zbog povećanja načina i mogućnosti komunikacije, često je teško odrediti relevantnost informacija. Ipak nakon određenja vjerodostojnosti i autentičnosti informacija, bilo bi poželjno one važne sačuvati. Iako su se informacije u početku bilježile na papiru (knjige, novine, časopisi itd.), danas postoji sve više informacija u digitalnom obliku. Zbog brzog napretka tehnologije, svjesni smo da računala koja trenutno koristimo, kao ni drugi mediji poput diskete, CD-a, DVD-a, optičkog diska na kojima su pohranjene informacije, neće biti upotrebljivi zauvijek. Tu dolazi do pojave raznih procesa (dugoročnog) očuvanja pomoću kojih će se informacije moći sačuvati. Još jedan važan pojam je i pojam repozitorija koji predstavlja mjesto gdje se nalazi sve ono što netko želi očuvati. U središtu ovog rada nalazi se PREMIS Rječnik podataka za metapodatke za zaštitu kao standard koji pomaže u očuvanju digitalnih objekata.

Iako je ovaj rad pisan samostalno, potrebno je spomenuti rad kolegice Marine Krstanović „*Sustav digitalnih akademskih arhiva i repozitorija DABAR: platforma Islandora i sheme deskriptivnih metapodataka*“ zbog bliskosti teme i međusobnog nadopunjavanja i nadovezivanja s ovim radom.

Najveće zahvale za podršku i usmjeravanje ovog rada želim uputiti mentoru dr. sc. Goranu Zlodiju, ali i Sveučilišnom računskom centru (SRCE), točnije Dabar timu (koji se bavi Digitalnim akademskim arhivima i repozitorijima).

3. Dugoročno očuvanje

Govoreći o dostupnosti i očuvanju informacija, pojmovi koji se uz njih najčešće spominju su pouzdanost i dugoročnost. Dostupnost označava mogućnost pristupa informacijama u bilo koje vrijeme te mogućnost njihove upotrebe. Pouzdanost znači da su informacije sigurno pohranjene te da su autentične, a dugoročnost znači osiguranje dostupnosti i očuvanosti informacije bez određenog vremenskog roka.¹ Iz navedenog proizlazi definicija dugoročnog očuvanja kao planiranog postupka pohrane informacija na neodređeni vremenski rok ili do tog trenutka kada informacije više neće biti potrebne. Pritom je bitno da su informacije pouzdane, da je zadržan njihov integritet (kompletnost i nepromjenjivost) te da su upotrebljive i autentične.

3.1. Informacijski i elektronički informacijski objekt

Dugoročno očuvanje može se odnositi na fizički i na digitalni oblik neke građe ili objekta, a u kontekstu ovog rada govori se o digitalnom, tj. elektroničkom obliku. Ovdje je bitno još spomenuti i pojmove informacijski objekt i elektronički informacijski objekt. „Informacijski objekt je bilo koje gradivo koje pruža informaciju bez obzira nalazi li se u analognom ili digitalnom (elektroničkom obliku).“ Elektronički informacijski objekt ili elektroničko gradivo je bilo koje gradivo nastalo upotrebom elektroničke i informacijske tehnologije neovisno o tome je li to njegov izvorni oblik ili je u tom obliku zbog postupka digitalizacije. Za svaki informacijski objekt bitne su i njegove tri razine – fizička, logička i konceptualna od kojih svaka ima svoja svojstva, ali se one i međusobno upotpunjuju.²

3.1.1. Fizička razina

„Fizička razina jest razina zapisa elektroničkog informacijskog objekta na neki medij.“ Odnosi se na binarni sustav znakova za računalni zapis i obradu. Zbog različitih fizičkih svojstava medija, svaki bilježi zapis na svoj način. „Na fizičkoj se razini ne odvija nikakva

¹ Hasenay, D., & Krtalić, M. (2009). *Dostupnost i očuvanje informacija – što znači pouzdano i dugoročno?*. 12. seminar Arhivi, knjižnice, muzeji: mogućnost suradnje u okruženju globalne informacijske infrastrukture. uredila Sanjica Faletar Tanacković. Zagreb: Hrvatsko knjižničarsko društvo. str. 2

² Stančić, H. (2006). *Arhivsko gradivo u elektroničkom obliku: mogućnost zaštite i očuvanja na dulji vremenski rok*. Arhivski vjesnik, Vol. 49 No. 1. Zagreb: Hrvatski državni arhiv. str. 2

značenjska interpretacija; fizička razina ne sadrži morfološke, sintaktičke ni semantičke informacije o zapisanom objektu.“³

3.1.2. Logička razina

„Logička razina određuje način na koji će sadržaj biti fizički organiziran i zapisan, a zanemaruje vrstu medija i način zapisa na njega, ali mora biti prisutna i na fizičkoj razini, tj. mora biti fizički zabilježena. Na logičkoj razini objekti mogu biti jednostavni i složeni.“ Jedna e-knjiga može biti jedan PDF dokument (jednostavan objekt), ali jedna e-knjiga može biti pohranjena u više PDF dokumenata ako je svaki PDF dokument npr. jedno poglavlje (složeni objekt). U tom se slučaju na fizičkoj razini nalazi informacija o njihovom povezivanju i redoslijedu te o softveru koji ih ispravno procesira.⁴

3.1.3. Konceptualna razina

„Konceptualna razina predstavlja onu razinu na kojoj se objekt prepoznaje kao smisljena cjelina, odnosno kao informacijska jedinica (npr. dokument, knjiga, slika, melodija itd.). Ova se razina još može nazivati i intelektualnom razinom, a njezin sadržaj i struktura moraju biti sadržani u logičkoj razini.“⁵

3.2. Obilježja ustanova koje rade na dugoročnom očuvanju

Institucije ili ustanove koje rade na dugoročnom očuvanju imaju određene karakteristike i moraju ispunjavati određene uvjete. Prema modelu Pouzdanih digitalnih repozitorija (engl. *Trusted Digital Repositories*, TDR) kojeg su razvili *Research Libraries Group* (RLG) i *Online Computer Library Center* (OCLC) 2002. godine, ustanova koja radi na dugoročnom očuvanju elektroničkog gradiva mora:⁶

- „prihvatiti odgovornost za dugoročno očuvanje,

³ Stančić, *Arhivsko gradivo u elektroničkom obliku: mogućnost zaštite i očuvanja na dulji vremenski rok*, n. dj., str. 3

⁴ Isto.

⁵ Stančić, *Arhivsko gradivo u elektroničkom obliku: mogućnost zaštite i očuvanja na dulji vremenski rok*, n. dj., str. 4

⁶ Stančić, H. (2009). *Značaj planiranja postupaka dugoročnog očuvanja elektroničkoga gradiva: značajne norme i metodologije*. 12. seminar Arhivi, knjižnice, muzeji: mogućnost suradnje u okruženju globalne informacijske infrastrukture. uredila Sanjica Faletar Tanacković. Zagreb: Hrvatsko knjižničarsko društvo. str. 14-15

- imati prikladan informacijski sustav (repozitorij),
- pokazati financijsku odgovornost i održivost,
- oblikovati sustav prema važećim normama,
- uspostaviti metodologiju za procjenu sustava,
- biti odgovorna pouzdana ustanova,
- imati izrađene politike i propisane prakse te ih provoditi mjerljivo i provjerljivo te
- zadovoljavati propisane odgovornosti.“⁷

Iz toga se izvode značajke takvih ustanova koje su:

1. „administrativna odgovornost,
2. organizacijska održivost,
3. financijska održivost,
4. prikladnost tehnologije i postupaka,
5. sigurnost sustava,
6. proceduralna odgovornost i
7. udovoljavanje zahtjevima OAIS referentnog modela.“⁸

3.3. Prijetnje dugoročnom očuvanju

Očuvanje digitalnih objekata može se odnositi na bilo koji tip objekta bio to tekst, slika, baza podataka, program, aplikacija, film, glazba ili nešto drugo. Iako postoji mogućnost da digitalni objekt „živi vječno“, nepoduzimanjem određenih mjera i postupaka očuvanja on je svakodnevno u opasnosti od trajnog nestanka. Postoji nekoliko područja na kojima se mogu pojaviti prijetnje i opasnosti, ali postoje i rješenja, tj. radnje koje možemo poduzeti kako ne bi došlo do gubitka ili propadanja građe.⁹

3.3.1. Tehnološke prijetnje

Neke od tehnoloških prijetnji za digitalni objekt su: raspadanje medija na kojem se nalaze podaci, zastarijevanje medija i/ili nepodržavanje medija na novijim računalima ili novijim

⁷ Isto.

⁸ Isto.

⁹ Digital Preservation Coalition. (2016). *Why digital preservation matters*. [citirano: 2016-12-13] Dostupno na: <http://www.dpconline.org/digital-preservation/why-digital-preservation-matters>

softverima i dr. Sve većim razvojem i napretkom tehnologije institucije se odlučuju za pohranu u oblaku, ali i tu postoji opasnost od propasti pružatelja usluga u oblaku.¹⁰

Postoje mnoge aktivnosti koje možemo poduzeti kako bismo spriječili propadanje podataka. Medij na kojem su pohranjeni podaci treba provjeravati i ažurirati, a podatke je potrebno migrirati (kad uočimo fizičko propadanje medija ili kad dođe do znatnijeg napretka u tehnologiji). Osim toga, potrebno je imati rezervne kopije podataka, tzv. *back up* te je preporučljivo provjeravati podatke kako ne bi došlo do gubitka u zapisu ili gubitka samog podatka. U slučaju pohrane materijala na više lokacija, lokacije zapisa bi trebale biti zabilježene itd.¹¹

3.3.2. Organizacijske prijetnje

Organizacija uključuje odabir građe, način organiziranja i obavljanja procesa očuvanja te proces promjene tih radnji kroz vrijeme i životni ciklus digitalnih materijala. Postoji nekoliko kriterija odabira građe (kada govorimo o postupku digitalizacije), a prioriteti najčešće ovise o nadležnoj instituciji. Prema tome, prednost može imati građa koja je najčešće korištena, građa koja je fizički u najboljem stanju (ako se radi o građi koja nije digitalna ili digitalizirana), građa koja sadrži povjerljive podatke, građa (vezana uz) određenog autora i dr. Iako postoje sličnosti u procesu digitalnog očuvanja između organizacija i ustanova, ovisno o kontekstu, svaka dolazi do individualnih prepreka zbog čega jedno rješenje ne odgovara svima. Jedna od važnijih odluka jest odrediti tko će obavljati proces digitalnog očuvanja. Postoji li dovoljno osoblja s potrebnim znanjem i infrastruktura unutar institucije, institucija može sama obavljati taj proces. Neke će se institucije ipak odlučiti za *outsourcing* jer im je to isplativije, ali tad institucija mora obavezno provjeriti obavljene radnje.^{12, 13}

3.3.3. Financijske prijetnje

S obzirom na to da je proces dugoročnog očuvanja (kao što i sam naziva govori) dugoročan, takvi su i troškovi. Novac nije potreban samo u početku, već za cijeli period očuvanja, o čemu treba voditi računa. Novac je potreban za medije na kojima su pohranjeni materijali, za

¹⁰ Digital Preservation Coalition. (2016). *Preservation issues*. [citirano: 2016-12-14] Dostupno na: <http://www.dpconline.org/digital-preservation/preservation-issues>

¹¹ Isto.

¹² Hasenay & Krtalić, *Dostupnost i očuvanje informacija – što znači pouzdano i dugoročno?*, n. dj., str. 3-6

¹³ Digital Preservation Coalition, *Preservation issues*, n. dj.

softver (ovisno o tome koji se upotrebljava), za osoblje (i s vremenom za njihovo usavršavanje), za ispunjavanje budućih ciljeva (npr. povećanje zbiraka, tj. repozitorija) itd.¹⁴

3.3.4. Osoblje

Za obavljanje procesa digitalnog i dugoročnog očuvanja potrebne su raznovrsne vještine i organizacijske sposobnosti. Ako se radi o većim institucijama (i repozitorijima), najčešće se radi i o većem broju osoblja. U manjim institucijama često dolazi do *outsourcinga* barem nekog dijela jer je teško pronaći osobe koje će imati sva potrebna znanja i vještine. Kategorija osoblja može se svrstati u financijske i organizacijske prijetnje jer se na njih troši dio financija, a sudjeluju i u organizaciji procesa očuvanja.¹⁵

4. O metapodacima

U današnje doba okruženi smo mnoštvom informacija i podataka sa svih strana i medija. Zbog velike količine podataka, pojavila se potreba za dodatnim opisom podataka radi lakšeg pronalaska, strukturiranja i dr. Ni jedan sustav danas ne bi funkcionirao da ne postoje podaci koji opisuju druge podatke, tj. metapodaci. U nastavku rada bit će detaljnije opisani i objašnjeni metapodaci te njihov razvoj i uloga u raznim okruženjima.

4.1. Što su metapodaci?

Iako se metapodaci najčešće opisuju kao podaci o podacima, ta definicija je vrlo općenita i zahtijeva proširenje. Prema autorima knjige „*Metadata for Digital Resources*“, određena definicija metapodataka jest ta da su metapodaci strukturirane informacije o informacijskom izvoru na bilo kojem mediju ili u bilo kojem formatu, koje služe korisniku (čovjeku ili računalu) u interakciji s objektima, pri čemu je naglasak na riječi strukturirano.¹⁶ Uz to je i važno „da u određenom kontekstu metapodaci moraju biti na višoj razini apstrakcije nego što su podaci koje oni opisuju.“ U današnje se vrijeme pojavljuju složenije situacije u kojima

¹⁴ Isto.

¹⁵ Isto.

¹⁶ Foulonneau, M., & Riley, J. (2008). *Metadata for Digital Resources*. Oxford: Chandos Publishing. str. 3

moramo upravljati zapisima pa u određenim okolnostima na metapodatke gledamo kao na podatke.¹⁷

Metapodaci mogu biti ugrađeni u digitalnom objektu što je čest slučaj u zaglavljima slikovnih datoteka ili u dokumentima hipertekstualnog jezika za označavanje (engl. *HyperText Markup Language*, HTML) kojim se određuje struktura, sadržaj i funkcija nekog HTML dokumenta (web stranice). Pohranjujući metapodatke s objektom koji opisuju, sprječavamo gubitak metapodataka, izbjegavamo probleme povezanosti između podataka i metapodataka te osiguravamo da će metapodaci i objekt biti zajedno ažurirani. Dakako, postoje objekti (kao što su artefakti) kod kojih nije moguće ugraditi metapodatke pa su u tom slučaju metapodaci pohranjeni zasebno. Pohrana metapodataka s objektom može olakšati upravljanje samim metapodacima te njihovu pretragu i dohvaćanje. Zbog toga se metapodaci često pohranjuju u bazama podataka, ali su i povezani s objektom koji opisuju. Uz sve to, oni i tumače, lociraju i olakšavaju dohvaćanje, korištenje i upravljanje informacijskim izvorima.¹⁸

4.2. Razvoj metapodataka

Upotreba metapodataka zbog digitalizacije postaje sve šira ne samo na području arhivistike, već i na području bibliotekarstva, muzeologije i dr., a metapodaci imaju ključnu ulogu u svakodnevnom radu tih institucija. Iako metapodaci postoje već jako dugo, nisu oduvijek poznati pod tim nazivom.

Razvoj metapodataka kreće od pojave katalogizacije, tj. od knjižničnih kataloga. S obzirom na to da katalozi sadrže podatke o knjigama i drugim izvorima, možemo na njih gledati kao na metapodatke. Ideja o katalogizaciji informacija pojavljuje se već u antičkom Egiptu, točnije u Aleksandrijskoj knjižnici gdje je sastavljena prva kompilacija metapodataka.¹⁹

¹⁷ Zlodi, G. (2007). *Mogućnosti uspostavljanja interoperabilnosti među shemama metapodataka u muzejskom okruženju*. Doktorska disertacija. Zagreb: Filozofski fakultet u Zagrebu. Odsjek za informacijske i komunikacijske znanosti. str. 97

¹⁸ National Information Standards Organization. (2004). *Understanding Metadata*. Bethesda: NISO Press. str. 1 [citirano: 2016-08-31] Dostupno na: <http://www.niso.org/publications/press/UnderstandingMetadata.pdf>

¹⁹ Haynes, D. (2004). *Metadata for Information Management and Retrieval*. London: Facet Publishing. str. 2 [citirano: 2016-08-15] Dostupno na: http://www.facetpublishing.co.uk/downloads/file/sample_chapters/HaynesMetadataach1.pdf

U zapadnoj Europi, katalogizacija se razvila kasnije počevši u crkvenim, a zatim i u akademskim knjižnicama. Sredinom 1960-ih godina računala se počinj u koristiti u svrhu katalogizacije, prilikom čega je stvoren novi standard strojno čitljiva katalogizacija (engl. *Machine-Readable Cataloging*, MARC) za formate podataka u kataložnom zapisu.²⁰ MARC definira formate podataka koji su se pojavili s inicijativom 1960. godine koju je vodila Kongresna knjižnica. On pruža mehanizam prema kojem računala razmjenjuju, koriste i interpretiraju bibliografske informacije, a njegovi elementi podataka čine temelj većine knjižničnih kataloga koji se danas koriste. MARC je 1980-ih postao USMARC, a u kasnim 1990-im MARC 21.²¹

Iako su se metapodaci u početku koristili u knjižničnim katalozima, danas se osim u već navedenim područjima koriste i u upravljanju zapisima, izdavaštvu, glazbenoj industriji, vladi, statistici i sve više u IT aplikacijama i još mnogim drugim područjima. Metapodaci se često koriste jer pružaju alate za opis elektroničkih informacijskih izvora, omogućuju konzistentnije dohvaćanje, bolje upravljanje izvora podataka i razmjenu podatkovnih zapisa između aplikacija i organizacija.²²

Sredinom 1990-ih godina iznijeta je ideja o jezgrovitom setu semantike za web izvore kako bi se kategorizirao web i kako bi se povećalo dohvaćanje. To je postalo poznato kao Dublin Core metapodatkovna inicijativa (engl. *Dublin Core Metadata Initiative*, DCMI) koja je stvorila jednostavan i sažet standard nazvan *Dublin Core* za opisivanje web sadržaja neovisno o disciplini ili jeziku. DCMI definira set podatkovnih elemenata koji se mogu koristiti kao spremnici metapodataka. Metapodaci su ugrađeni u izvore tako da im mogu pristupiti korisnici i sustavi koji koriste web izvore. Izvorni cilj Dublin Corea bio je definirati skup elemenata koje će koristiti autori kako bi opisali vlastite web izvore, a taj cilj se kasnije prenamijenio u proces definiranja nekoliko elemenata i nekoliko jednostavnih pravila koje će moći primijeniti oni koji nisu upoznati s katalogiziranjem. Originalnih 13 elemenata jezgre kasnije se povećalo na 15: naslov, tema (sadržaj), opis, izvor, jezik, odnos, područje/razdoblje, stvaratelj, nakladnik, suradnik, vlasnička prava, datum, tip, format i

²⁰ Haynes, *Metadata for Information Management and Retrieval*, n. dj., str. 3

²¹ The Library of Congress. (2006). *MARC Standards Frequently Asked Questions*. [citirano: 2016-10-26]

Dostupno na: <https://www.loc.gov/marc/faq.html>

²² Haynes, *Metadata for Information Management and Retrieval*, n. dj., str. 4

identifikator. DCMI je inicijativa koja je u tijeku i nastavlja sa razvojem i razradom metapodatkovnih standarda i za web aplikacije koje zahtijevaju kompleksnost.^{23, 24}

4.3. Vrste metapodataka

U knjizi „*Introduction to Metadata*“ navedeno je i definirano pet vrsta metapodataka. Ti metapodaci su: administrativni, deskriptivni, tehnički, metapodaci za zaštitu te metapodaci vezani uz korištenje o kojima će više riječ biti u nastavku.

4.3.1. Administrativni metapodaci

Administrativni metapodaci su oni koji se koriste u upravljanju, administraciji, primjeni i pronalasku zbiraka i informacijskih izvora. To mogu biti informacije o nabavi, praćenje prava i reprodukcije (kao i prava intelektualnog vlasništva – autorska prava, licence), dokumentacija o legalnim zahtjevima pristupa, informacije o lokaciji te selekcijskim kriterijima digitalizacije, informacije o vremenu i načinu stvaranja izvora (uključujući i izmjene kao što su migracije i sl.).^{25, 26} Neki autori kao podvrste administrativnih metapodataka navode metapodatke za zaštitu te korištenje, ali u ovom će radu oni biti opisani zasebno.

4.3.2. Deskriptivni metapodaci

Deskriptivni metapodaci se koriste kako bi identificirali i opisali zbirke i srodne informacijske izvore npr. kataložne zapise, obavijesna pomagala, diferencijaciju među verzijama, specijalizirane indekse, kustoske informacije, poveznice među izvorima te anotacije stvaratelja i korisnika. Elementi koji mogu biti uključeni u deskriptivnim metapodacima su naslov, sažetak, autor, ključne riječi itd.^{27, 28}

²³ Haynes, *Metadata for Information Management and Retrieval*, n. dj., str. 5

²⁴ National Information Standards Organization, *Understanding Metadata*, n. dj., str. 3

²⁵ Gilliland, A. J. (2008) *Setting the Stage*. Introduction to Metadata. uredila Murtha Baca. Los Angeles: The Getty Research Institute. str. 9

²⁶ National Information Standards Organization, *Understanding Metadata*, n. dj., str. 1

²⁷ Gilliland, *Setting the Stage*, n. dj., str. 9

²⁸ National Information Standards Organization, *Understanding Metadata*, n. dj., str. 1

4.3.3. Metapodaci za zaštitu

Metapodaci za zaštitu odnose se na upravljanje očuvanjem zbiraka i informacijskih izvora. Neki od primjera su dokumentacija o fizičkom stanju izvora i dokumentacija o radnjama vezanima uz očuvanje fizičkih i digitalnih verzija izvora. Te se radnje odnose na migraciju te dokumentaciju o promjenama koje su se pojavile tijekom digitalizacije ili tijekom procesa očuvanja, što je bitno za podrijetlo i vjerodostojnost digitalnog objekta.²⁹

4.3.4. Tehnički metapodaci

Tehnički metapodaci su oni koji se odnose na način funkcioniranja sustava i na ponašanje metapodataka, npr. dokumentacija o hardveru i softveru, tehničke informacije o digitalizaciji kao što su format, stupanj sažimanja i način skaliranja, vremensko praćenje odgovora sustava, podaci o korištenju te podaci o autentičnosti i sigurnosti kao što su ključevi za enkripciju i lozinke.³⁰ Oni „se mogu smatrati i podskupom metapodataka za zaštitu jer su podaci o izradi digitalne građe ključni dio skupa podataka potrebnih za trajnu pohranu koji omogućuju očuvanje i „rekonstrukciju“ građe čak i kad nestanu pomagala za pregled datoteke.“³¹

4.3.5. Metapodaci vezani uz korištenje

Metapodaci vezani uz korištenje ukazuju nam na to kako su objekti sastavljeni, a odnose se i na razinu i način korištenja zbirke ili informacijskog izvora. Primjer toga su: kruženje zapisa, fizički i digitalni zapisi o izložbama, korištenje i praćenje korisnika, ponovno korištenje sadržaja, zapisi o pretraživanjima i metapodatkovna prava.³²

4.4. Važnost i upotreba metapodataka

Osim što opisuju sadržaj, metapodaci opisuju i format i attribute podatkovnog zapisa ili informacijskog izvora. Koriste se za opis vrlo strukturiranih izvora, ali i za nestrukturirane informacije poput tekstualnih dokumenata. Važni su i zbog povećanja dohvata relevantnih

²⁹ Gilliland, *Setting the Stage*, n. dj., str. 9

³⁰ Isto.

³¹ Klarin, S. (2008.) *Tehnički metapodaci i metapodaci za zaštitu u normama ANSI/NISO Z39.87 i PREMIS*. 11. seminar Arhivi, knjižnice, muzeji: mogućnost suradnje u okruženju globalne informacijske infrastrukture. uredila Mirna Willer. Zagreb: Hrvatsko knjižničarsko društvo. str. 151

³² Gilliland, *Setting the Stage*, n. dj., str. 9

informacija te zbog mogućnosti povećanja dohvaćanja stvaranjem konteksta za individualne deskriptore. Metapodaci se mogu primijeniti i na opise elektroničkih izvora, digitalnih podataka (uključujući i digitalne slike), ali i na fizičke zapise kao što su knjige, novine i izvješća. Mogu biti ugrađeni u informacijskom izvoru (što je čest slučaj kod web izvora) ili mogu biti odvojeni od izvora u bazi podataka.³³

Prikladne metapodatkovne oznake vezane uz različite elemente podataka dozvoljavaju pretraživačima pretragu informacija na ograničen način. Metapodaci omogućavaju i upravljanje elektroničkim digitalnim objektima. Sustav za upravljanje sadržajem (engl. *Content Management System*, CMS) „je računalni program koji se koristi za stvaranje, izmjenu, upravljanje i objavljivanje sadržaja na dosljedno organiziran način“³⁴ koji, na primjer, koristi metapodatke kako bi pratio kada je digitalni objekt posljednji puta ažuriran ili verificiran, tko je odgovoran za pojedine aktivnosti na objektu te postoje li posebni uvjeti pristupa. Metapodaci koji su povezani s digitalnim objektom pružaju zajednički format za upravljanje i manipulaciju resursima. Oni mogu pomoći pri utvrđivanju autentičnosti podataka – elektroničkog dokumenta ili slike, što je vrlo važno organizacijama koje ovise o elektroničkim zapisima više nego o fizičkim (na papiru). Postaje neophodno dokazati da je elektronički dokument bio držan na sigurnom, da je zapis potpun i da nije bio mijenjan. Iz toga proizlazi da metapodaci daju dokaz za integritet elektroničkog dokumenta. Oni su ključ za interoperabilnost, koja ovisi o razmjeni metapodataka među sustavima za uspostavu tipa podataka koji su premješteni. Pristup metapodacima pomaže uspostaviti protokole za razmjenu podataka i utvrditi na koji način se to može iskoristiti. Metapodaci su budućnost jer sve veći broj softvera i dobavljača sustava rade na metapodatkovnim standardima ili izrađuju vlastite odgovarajuće metapodatkovne standarde. To omogućuje bolju suradnju i razmjenu podataka među industrijama.³⁵

Michael Day u svojem modelu iz 2001. godine izlaže sedam najvažnijih uloga metapodataka. Te uloge su: opis izvora, pronalazak izvora, administracija i upravljanje resursima, zapis o intelektualnim vlasničkim pravima, dokumentacija softverskog i hardverskog okruženja, očuvanje digitalnih izvora te pružanje informacija o kontekstu i autentičnosti.³⁶

³³ Haynes, *Metadata for Information Management and Retrieval*, n. dj., str. 8

³⁴ CARNET (2008). *CMS sustavi i sigurnost*. Zagreb: Hrvatska akademska i istraživačka mreža. str. 6
[citirano: 2016-10-26] Dostupno na:
<http://www.cis.hr/www.edicija/LinkedDocuments/CCERT-PUBDOC-2008-12-249.pdf>

³⁵ Haynes, *Metadata for Information Management and Retrieval*, n. dj., str. 11

³⁶ Haynes, *Metadata for Information Management and Retrieval*, n. dj., str. 13

Nakon proučavanja prvog izdanja knjige „*Introduction to Metadata*“ iz 1998. godine te nakon proučavanja modela i uloga metapodataka koje je definirao Michael Day, David Haynes osmišljava vlastiti suvremeniji model metapodataka koji u 5 točaka objašnjava svrhu metapodataka.³⁷

Prva od tih točaka je opis izvora. To je osobito važno u organizacijama koje trebaju opisati informacije o vlastitim aktivama. Registri s informacijama o aktivama, koje sastavljaju javne vlasti i sve više korporativni sektor, također zahtijevaju opise informacijskih repozitorija i izvora.³⁸

Sljedeća točka je dohvaćanje informacija. U akademskom sektoru sve se više truda ulaže u pronalazak izvora na internetu. Neke institucije i agencije osmislile su tematske portale koji katalogiziraju relevantne web izvore prema određenim temama što usmjerava korisnike na put do mjerodavnih izvora informacija. Podaci vezani uz katalogizaciju obično uključuju opis izvora, kontrolirane indeksirane pojmove te klasifikaciju naslova. To su izvori metapodataka koji mogu izvesti metapodatke direktno s ciljnih web stranica ili elektroničkih izvora.³⁹

Treća točka je upravljanje informacijskim izvorima. Porast sustava za upravljanje elektroničkim dokumentima i zapisima (engl. *Electronic Document and Records Management System*, EDRMS) dogodio se zbog nastalih zahtjeva za većim organizacijama koje će učinkovito upravljati dokumentacijom u papirnatom i elektroničkom obliku. Ti sustavi trebaju pristup informacijama o katalogizaciji individualnih dokumenata kako bi mogli upravljati životnim ciklusom zapisa. Primjeri uključuju autorstvo, vlasništvo, podrijetlo dokumenta (u pravne svrhe) te datume stvaranja i datume promjene. Navedeni podaci, uz druge elemente pružaju bazu za ekonomično i konzistentno upravljanje dokumentacijom. CMS se također koristi za upravljanje izvorima podataka, a to uključuje i materijale koji su objavljeni na intranetima i web stranicama.⁴⁰

³⁷ Haynes, *Metadata for Information Management and Retrieval*, n. dj., str. 15-16

³⁸ Isto.

³⁹ Isto.

⁴⁰ Isto.

Sljedeća točka je dokumentacija vlasništva i autentičnosti digitalnih izvora. Metapodaci omogućavaju deklariranje vlasništva intelektualnog sadržaja i izgled dokumenta. Pružaju i zapis o autentičnosti dokumenta omogućavajući pregled tragova. Jedan od preduvjeta za široko prihvaćanje elektroničkih dokumenata kao originalnog dokaza je taj da elektronički sustavi postaju preferencijalni mediji za dugoročno očuvanje dokumenata.⁴¹

Posljednja točka je interoperabilnost. „Interoperabilnost je mogućnost da više sustava, koji imaju različite hardverske i softverske platforme, strukture podataka i sučelja, mogu razmjenjivati podatke s minimalnim gubitkom sadržaja i funkcionalnosti.“⁴² Metapodaci omogućavaju prijenos informacija i podataka među sustavima i kao takvi su ključna komponenta interoperabilnosti. Da bi samostalno dizajniranim aplikacijama dozvolili propusnost podataka, potreban nam je zajednički okvir za opis prenijetih podataka kako bi svatko znao kako na najprikladniji način njima rukovati. To može biti na različitim razinama među različitim jezicima ili u obliku razumijevanja različitih formata podataka. Interoperabilnost omogućuje e-trgovinu. Kada dio podatka prolazi s jednog sustava na drugi popratni ili ugrađeni, metapodaci omogućuju novoj aplikaciji da daje smisao podatku i da ga koristi na odgovarajući način.⁴³

Ovaj model od pet točaka razdvaja opis od dohvaćanja. Neka područja ujedinjuje, kao što su upravljanje izvorima i upravljanje očuvanjem (koje je predstavljeno kao podvrsta upravljanja) te upravljanje pravima koje je vezano uz podrijetlo i autentičnost. Ovaj model čini razliku između svrhe i svojstava metapodatkovnih elemenata. Čineći to, postaje jasno da se svaki element podatka može koristiti na više načina i da ima više od jedne namjene.^{44, 45}

Vidljivi primjeri metapodataka kreću od metaoznaka HTML-a na web stranicama do MARC zapisa koji se koriste za razmjenu katalogiziranih podataka između sustava za upravljanje u knjižnicama. Metapodaci mogu biti izraženi u obliku strukturiranog jezika kao što je proširiv jezik za označavanje (engl. *Extensible Markup Language*, XML) i mogu slijediti upute ili

⁴¹ Isto.

⁴² National Information Standards Organization, *Understanding Metadata*, n. dj., str. 2

⁴³ Haynes, *Metadata for Information Management and Retrieval*, n. dj., str. 16-17

⁴⁴ Isto.

⁴⁵ National Information Standards Organization, *Understanding Metadata*, n. dj., str. 1

sheme za određene aktivnosti domene.⁴⁶ XML je osmišljen za strukturiranje, pohranjivanje i razmjenu informacija te je kreiran s namjerom da bude malen, brz, jeftin, jednostavan za učenje i prilagođen korištenju na internetu. Njegov je cilj da bude čitljiv strojevima i ljudima.

5. O repozitorijima

Svakodnevnim povećanjem broja računala i upotrebe interneta počinje se stvarati sve veća količina informacija. Shodno tome povećava se i broj istraživanja, završnih i diplomskih radova, doktorskih disertacija i dr., pa se zbog toga 1990-ih godina počinju razvijati repozitoriji koji omogućuju trajnu i strukturiranu pohranu akademskih i znanstvenih radova. Govoreći o repozitorijima, Lasić-Lazić ih definira kao „zbirka digitalnog obrazovnog materijala koja omogućuje sustavno upravljanje procesima objavljivanja, pristupa i pohrane nastavnog/obrazovnog sadržaja.“⁴⁷ S druge strane, Vrana navodi širu definiciju Johnsona koji opisuje digitalni repozitorij ustanove kao „digitalni arhiv koji je rezultat intelektualnog rada koje su stvorili nastavnici, istraživačko osoblje i studenti određene ustanove i koji je dostupan krajnjim korisnicima unutar i izvan ustanove s manjim ili nikakvim preprekama u pristupu.“⁴⁸

5.1. Zašto postoje repozitoriji; obilježja i funkcije

Radovi koji su otvorenog pristupa su oni koji su digitalni, nalaze se na mreži, besplatni su te nisu licencama ograničeni. Svrha postojanja otvorenog pristupa radovima i postojanja repozitorija ima mnogo, a neke od njih su: povećanje vidljivosti i čitljivosti autora, besplatan i slobodan pristup strukturiranim znanstvenim radovima, izbjegavanje plaćanja znanstvenih

⁴⁶ Foulonneau & Riley, *Metadata for Digital Resources*, n. dj., str. 50

⁴⁷ J. Lasić-Lazić. (2005). *Repozitoriji digitalnog obrazovnog materijala kao sastavnica kvalitete suvremenih koncepta obrazovanja*. [citirano: 2016-11-25] Dostupno na: <http://edupoint.carnet.hr/casopis/33/clanci/1.html>

⁴⁸ R. Vrana. (2010). *Digitalni repozitoriji ustanova i njihova uloga u znanstvenom komuniciranju Istraživanje o digitalnim repozitorijima u visokoškolskim knjižnicama Sveučilišta u Zagrebu*. 13. seminar Arhivi, knjižnice, muzeji: mogućnost suradnje u okruženju globalne informacijske infrastrukture. uredile Mirna Willer i Sanjica Faletar Tanacković. Zagreb: Hrvatsko knjižničarsko društvo. str. 281

časopisa, ukoliko sveučilište ili fakultet imaju svoj repozitorij – povećanje njihove vidljivosti, dugoročna pohrana i upravljanje radovima i mnoge druge.⁴⁹

Bosančić u svom članku navodi osnovna obilježja i funkcije repozitorija ustanova koje je preuzeo od Gibbonsa (a primjenjuju se i na druge vrste repozitorija). On tvrdi da se „preko obilježja najbolje može uočiti svrha repozitorija, a ta obilježja su: pohrana digitalnog sadržaja (tekst, slika, audio, video itd.), namjena posebnoj zajednici korisnika, podrška ustanove u kojoj se nalaze, osiguranje trajne pohrane digitalnih sadržaja (kao i trajnosti hiperveza preko kojih se može pristupiti digitalnom sadržaju unutar repozitorija) te osiguranje pristupa sadržaju članovima i nečlanovima zajednice korisnika. S druge strane, tu su i funkcije koje bi svaki repozitorij trebao podržavati: unos digitalnog sadržaja u repozitorij, opis digitalnog sadržaja jednom od shema metapodataka, raspačavanje (diseminacija) digitalnih sadržaja iz repozitorija na korisnički zahtjev, određivanje prava unosa, uređivanje i pristup digitalnim sadržajima u repozitoriju, mogućnost pretrage i pregleda digitalnih sadržaja u repozitoriju na razini metapodataka i potpunog teksta te zaštita digitalnih sadržaja na dulji vremenski rok.“⁵⁰

5.2. Vrste repozitorija

Kad se govori o repozitorijima, spominje se više vrsta repozitorija koji se najčešće dijele na: tematske repozitorije, istraživačke repozitorije, nacionalne repozitorije i institucijske repozitorije.

5.2.1. Tematski repozitoriji

Tematski repozitoriji (engl. *subject-based repositories*) najčešće su prilagođeni široj zajednici, a uspostavljaju ih članovi zajednice. Ovdje se pojavljuje i termin samoarhiviranja (postupak samostalne pohrane vlastitog rada u repozitorij) jer takva vrsta repozitorija ima veliku vrijednost za znanstvenike. Autori radova mogu objaviti svoje radove i početna istraživanja u obliku radnih verzija ili preprintova (nerecenziranih radova) koji mogu rezultirati korisnošću u obliku testiranja vrijednosti ideje ili rezultata, poboljšanja rada prije objave, dobivanjem priznanja, dostizanjem međunarodne pažnje i dr. Teme tematskih repozitorija dobro su

⁴⁹ P. Suber. (2015). *Open Access Overview: Focusing on open access to peer-reviewed research articles and their preprints*. [citirano: 2016-11-25] Dostupno na: <http://legacy.earlham.edu/~peters/fos/overview.htm>

⁵⁰ B. Bosančić. (2009). *Repozitoriji ustanova u teoriji i praksi*. 12. seminar Arhivi, knjižnice, muzeji: mogućnost suradnje u okruženju globalne informacijske infrastrukture. uredila Sanjica Faletar Tanacković. Zagreb: Hrvatsko knjižničarsko društvo. str. 74

definirane, a statistika upotrebe je značajna za korisnike zajednice.⁵¹ Najpoznatiji primjer takvog repozitorija je arXiv (<https://arxiv.org/>) koji sadrži radove iz područja fizike, matematike, informatike, kvantitativne biologije, kvantitativnih financija i statistike.⁵²

5.2.2. Istraživački repozitoriji

Istraživački repozitoriji (engl. *research repositories*) su oni koje najčešće sponzorira neka zaklada ili organizacija kod koje postoji interes za istraživanja. Rezultati su često značajni i od velikog interesa, a objavljuju se čak i u obliku knjiga. Zbog vrlo važnih znanstvenih zapisa koji se nalaze u radovima, postoje i strogi standardi vezani uz pohranu i očuvanje radova. U većini su slučajeva radovi u tim repozitorijima visoke kvalitete zbog višestrukih pregleda i recenzija te zbog količine uloženog novca.⁵³

5.2.3. Nacionalni repozitoriji

Nacionalni repozitoriji (engl. *national repositories*) prikupljaju radove poput završnih i diplomskih, doktorske disertacije, ali mogu sadržavati i obrazovne materijale, tj. materijale korištene u nastavi. Radovi su najčešće objavljeni na matičnom jeziku repozitorija. Dostupnost tih radova povećava broj obrazovnih izvora za viši stupanj obrazovanja te javni pristup znanju.⁵⁴ Primjer takvog repozitorija na našem području je Nacionalni repozitorij završnih i diplomskih radova – ZIR Nacionalne i sveučilišne knjižnice u Zagrebu (<https://zir.nsk.hr/>).

5.2.4. Institucijski repozitoriji

Institucijski repozitoriji (engl. *institutional repositories*) „digitalne su zbirke raznovrsnih dokumenata koji su nastali kao rezultat znanstveno-istraživačkog rada neke institucije. Osim prikupljanja, oni arhiviraju i distribuiraju radove nastale kao produkt intelektualnog rada

⁵¹ Armbruster, C., & Romary, L. (2010). *Comparing Repository Types – Challenges and barriers for subject-based repositories, research repositories, national repository systems and institutional repositories in serving scholarly communication*. International Journal of Digital Library Systems. IGI Global. str. 62 [citirano: 2016-11-26] Dostupno na: <https://hal.inria.fr/inria-00465197v2>

⁵² arXiv.org. (1991). *General Information About arXiv*. [citirano: 2016-11-29] Dostupno na: <https://arxiv.org/help/general>

⁵³ Armbruster & Romary, *Comparing Repository Types – Challenges and barriers for subject-based repositories, research repositories, national repository systems and institutional repositories in serving scholarly communication*, n. dj., str. 62

⁵⁴ Armbruster & Romary, *Comparing Repository Types – Challenges and barriers for subject-based repositories, research repositories, national repository systems and institutional repositories in serving scholarly communication*, n. dj., str. 63

institucije. Najčešće zastupljene vrste dokumenata jesu: ocjenski radovi (doktorati i magisteriji te diplomski, odnosno završni radovi), knjige, poglavlja u knjigama, članci u časopisima, članci u novinama, radovi u zbornicima skupova, nastavni materijali.⁵⁵ U većini slučajeva se institucijski repozitoriji vežu uz neku akademsku instituciju poput fakulteta, sveučilišta ili instituta. Osim što postojanje institucijskog repozitorija pridonosi razvoju same institucije, pridonosi i regionalnom razvoju područja u kojem se nalazi dajući tvrtkama i tijelima javne vlasti uvid u područja kojima se bave ljudi iz zajednice.⁵⁶ Primjer takvog repozitorija je Repozitorij Filozofskog fakulteta u Zagrebu (<http://darhiv.ffzg.unizg.hr/>).

6. O PREMIS-u

Strategije primjene metapodataka za zaštitu (engl. *PREservation Metadata: Implementation Strategies*, PREMIS) je naziv međunarodne radne skupine koju su osnovali i sponzorirali OCLC i RLG 2003.-2005. godine.⁵⁷ Ta radna skupina predstavila je PREMIS Rječnik podataka za metapodatke za zaštitu (engl. *PREMIS Data Dictionary for Preservation Metadata*) - međunarodni standard za metapodatke, koji pomaže u očuvanju digitalnih objekata, čime omogućava njihovu dugoročnu upotrebu. PREMIS je implementiran u brojne projekte digitalnog očuvanja diljem svijeta, a podrška za nj inkorporirana je u mnoge digitalne komercijalne alate i sustave, ali i u one koji su otvorenog pristupa. PREMIS-ov urednički odbor koordinira revizije i implementacije standarda koji se sastoji od podatkovnog rječnika, XML shema i popratne dokumentacije.⁵⁸

Zadatak PREMIS-ove radne skupine bio je:

- definirati implementacijski jezgrovit skup elemenata metapodataka koji su široko primjenjivi u zajednici digitalnog očuvanja;
- izraditi podatkovni rječnik koji će podržavati jezgrovit skup elemenata metapodataka za očuvanje;

⁵⁵ Krajina, T. (2008). *Slobodan pristup informacijama: institucijski repozitoriji*. Polimeri: časopis za plastiku i gumu, Vol. 28 No. 3. str. 199

⁵⁶ Armbruster & Romary, *Comparing Repository Types – Challenges and barriers for subject-based repositories, research repositories, national repository systems and institutional repositories in serving scholarly communication*, n. dj., str. 63

⁵⁷ Caplan, P. (2009). *Understanding PREMIS*. Washington, D. C.: The Library of Congress. str. 4 [citirano: 2016-09-20] Dostupno na: <https://www.loc.gov/standards/premis/understanding-premis.pdf>

⁵⁸ The Library of Congress. (2016). *PREMIS: Preservation Metadata Maintenance Activity*. [citirano: 2016-09-20] Dostupno na: <http://www.loc.gov/standards/premis/>

- ispitati i evaluirati alternativne strategije za kodiranje, pohranu i upravljanje metapodacima za očuvanje unutar digitalnog sustava za očuvanje, kao i za razmjenu metapodataka za očuvanje među sustavima;
- provesti pilot programe za testiranje preporuka radne skupine u raznim postavkama sustava;
- istražiti mogućnosti za kooperativno stvaranje i razmjenu metapodataka za očuvanje.⁵⁹

Skica tog izvješća i podatkovni rječnik završeni su u veljači 2005. godine. Rječnik su vrednovali PREMIS-ov savjetodavni odbor i malen broj pozvanih ocjenjivača. Radna skupina dobila je mnogo vrijednih povratnih informacija u inicijalnom periodu ocjenjivanja te je utrošila mnogo vremena razmatrajući svaki komentar i radeći revizije iz čega je proizašla velika korist za oba dokumenta.⁶⁰

Nakon objave PREMIS podatkovnog rječnika verzije 1.0, sljedeći je korak bio usredotočiti se na implementaciju i interoperabilnost. Operativni repozitoriji mogu koristiti PREMIS kao popis s kojim mogu usporediti vlastite specifikacije za očuvanje metapodataka. Repozitoriji u razvoju mogu služiti kao testne platforme za implementaciju PREMIS-ove semantike te koristiti svoja iskustva za buduće revizije podatkovnog rječnika. XML veze za podatkovni rječnik su razvijene kako bi predstavljale PREMIS-ove metapodatke za razmjenu s arhivskim informacijskim paketima među repozitorijima za očuvanje.⁶¹

Radna skupina želi naglasiti da namjera podatkovnog rječnika nije da bude nepromjenjiv, već da omogući polaznu točku za napredak i poboljšanja koja će se temeljiti na iskustvu i povratnim informacijama zajednice.⁶²

6.1. Nastanak i uspostava PREMIS-a

PREMIS je uspostavljen kako bi se nadogradio na raniji rad druge inicijative OCLC-a i RGL-a, točnije na Okvir radne skupine za primjenu metapodataka za zaštitu (engl. *Preservation Metadata Framework working group*). Od 2001. do 2002. godine, ta je radna skupina

⁵⁹ The PREMIS Working Group. (2005). *Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group, Version 1.0*. str. vii [citirano: 2016-10-07] Dostupno na: <http://www.loc.gov/standards/premis/v1/index.html>

⁶⁰ Isto.

⁶¹ Isto.

⁶² Isto.

istaknula tipove informacija koje bi trebale biti povezane s arhiviranim digitalnim objektom. Njihov izvještaj *Metapodatkovni okvir za podršku pri očuvanju digitalnih objekata* (engl. *A Metadata Framework to Support the Preservation of Digital Objects*) ponudio je listu prototipa elemenata metapodataka iako je trebalo uložiti još rada kako bi taj prototip bilo moguće implementirati. PREMIS-ova radna skupina je odlučila ići korak dalje – razviti podatkovni rječnik s jezgrovitim elementima metapodataka koji će se moći primijeniti na arhivirane objekte, dati upute za implementaciju skupa elemenata metapodataka u sustave za očuvanje te predložiti najbolju praksu za popunjavanje tih elemenata.⁶³

S obzirom na to da je fokus PREMIS-a bio više praktični nego teorijski, članovi radne skupine su odabrani iz institucija za koje je bilo poznato da razvijaju repozitorijske sustave za očuvanje unutar kulturnog nasljeđa te informacijske industrije. Radna skupina sastojala se od predstavnika akademskih i nacionalnih knjižnica, muzeja, arhiva, vlade te komercijalnih poduzeća iz šest različitih zemalja. PREMIS je pozvao i međunarodni savjetodavni odbor stručnjaka koji će promatrati napredak.⁶⁴

Kako bi postigli što je više moguće u realnom vremenskom okviru, radna skupina se podijelila u dvije podskupine. Podskupina za implementacijske strategije (engl. *The Implementation Strategies Subgroup*) ispitivala je različite strategije za kodiranje, pohranu i upravljanje metapodacima za očuvanje, unutar digitalnih sustava za očuvanje dok je Podskupina za temeljne elemente (engl. *The Core Elements Subgroup*) bila odgovorna za odabir temeljnih elemenata metapodataka za očuvanje te za izradu podatkovnog rječnika.⁶⁵

Kako bi saznali na koji način repozitoriji za očuvanje implementiraju metapodatke za očuvanje, podskupina za implementacijske strategije ispitala je sedamdesetak organizacija koje su aktivne ili koje pokazuju interes za digitalnim očuvanjem. Istraživanje je osiguralo mogućnost istraživanja stanja općenito digitalnog očuvanja, a pitanja su bila sastavljena tako da se izvuku informacije o politici, upravljanju i financiranju, arhitekturi sustava, strategijama očuvanja i o metapodatkovnoj praksi. U prosincu 2004. godine PREMIS-ova je radna skupina objavila izvještaj koji se temeljio na istraživanjima digitalnih repozitorija pod nazivom „Implementacija repozitorija za očuvanje digitalnih materijala: Trenutna praksa i trendovi u nastajanju u zajednici kulturne baštine“ (engl. *Implementing Preservation Repositories for*

⁶³ The PREMIS Working Group, *Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group, Version 1.0.*, n. dj., str. viii

⁶⁴ Isto.

⁶⁵ Isto.

Digital Materials: Current Practice and Emerging Trends in the Cultural Heritage Community).⁶⁶

PREMIS definira metapodatke za očuvanje kao informaciju koju repozitorij koristi kako bi podržao proces digitalnog očuvanja. Radna skupina je tražila metapodatke koji podržavaju funkcije rada, renderiranja, razumijevanja, autentifikacije i identifikacije u kontekstu očuvanja. Tako metapodaci za očuvanje obuhvaćaju više kategorija metapodataka: administrativne (uključujući prava i dozvole), tehničke te strukturalne metapodatke. Dio pažnje je posvećen i dokumentaciji o digitalnom podrijetlu kao i dokumentaciji koja govori o vezama (osobito među različitim objektima unutar repozitorija za očuvanje).⁶⁷

Vodeće načelo skupine pri sastavljanju podatkovnog rječnika bilo je da implementirajuća shema metapodatka što strože definira svaku semantičku jedinicu i poveže ju s tipom entiteta koji opisuje.⁶⁸

6.1.1. OAIS referentni model

Okvir radne skupine za očuvanje metapodataka i PREMIS podatkovni rječnik izgrađeni su na Referentnom modelu za otvoreni arhivski informacijski sustav (engl. *Open Archival Information System reference model*, OAIS (ISO14721)). Referentni model otvorenog arhivskog informacijskog sustava razvio je NASA-in Savjetodavni odbor za svemirske podatkovne sustave (engl. *Consultative Committee for Space Data Systems*, CCSDS) 1999. godine. To je bio začetak generiranja formalnih standarada za dugoročno arhiviranje svemirskih znanstvenih podataka. Referentni model je konceptualni okvir arhivskog sustava koji se bavi dugoročnim očuvanjem i održavanjem pristupa digitalnim informacijama. Svrha referentnog modela jest povećati svijest i razumijevanje koncepata koji su važni za arhiviranje digitalnih objekata, objasniti terminologiju i koncepte za opis i usporedbu podatkovnih modela i arhivske arhitekture, povećati konsenzus elemenata i endemskih procesa u očuvanju digitalnih informacija i njihovom pristupu te stvoriti okvir koji će voditi identifikaciju i razvoj standarada. Arhiv organiziran prema OAIS modelu ima šest (minimalnih) zadataka:⁶⁹

- pregovarati i prihvaćati odgovarajuće podatke za stvaratelje podataka

⁶⁶ Isto.

⁶⁷ The PREMIS Working Group, *Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group, Version 1.0.*, n. dj., str. ix

⁶⁸ The PREMIS Working Group, *Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group, Version 1.0.*, n. dj., str. x

⁶⁹ Lavoie, B. (2000). *Meeting the challenges of digital preservation: The OAIS reference model*. str. 3 [citirano: 2016-11-15] Dostupno na: <http://www.oclc.org/research/publications/library/2000/lavoie-oais.html>

- održati dovoljnu kontrolu nad podacima kako bi se osiguralo dugoročno očuvanje
- odrediti ciljne korisničke skupine
- osigurati da podaci budu jasni ciljnim korisničkim skupinama
- pratiti dokumentirane politike i procedure te time osigurati očuvanje podataka od slučajnosti i omogućiti diseminaciju podataka kao autenticiranih kopija originala ili onih kojima se može provjeriti original
- omogućiti ciljnim korisničkim skupinama dostupnost podataka⁷⁰

Okolina OAIS referentnog modela izvedena je iz interakcije četiri entiteta: proizvođači, potrošači, menadžment i sam arhiv. Proizvođači stvaraju podatke koje arhiv treba očuvati, a potrošači koriste očuvane podatke. Potrošači su ciljna korisnička skupina od kojih se očekuje da će razumjeti arhivirane podatke, dok je entitet menadžment odgovoran za određivanje jasnih ciljeva politike arhiva.⁷¹

Struktura referentnog modela je takva da sastoji se od informacijskog modela (koji je osnovni model unutar OAIS-a), modela transformacija informacijskih paketa te od funkcionalnog modela. Informacijski model se sastoji od logičkog modela arhivskih informacija i logičkog modela informacija u OAIS-u. Logičkim modelom arhivskih informacija se definira struktura informacijskog objekta, a on unutar sebe sadrži informacijski objekt (engl. *Information Object*) koji je najmanja informacijska jedinica. Informacijski objekt sastoji se od podatkovnog objekta i informacija za prikaz. Logički model informacija u OAIS-u definira pojam, strukturu i vrste informacijskih paketa. On se sastoji od informacijskog paketa (engl. *Information Package*) koji logički objedinjuje četiri osnovne vrste informacijskih objekata. Te vrste informacijskih objekata su: informacije o sadržaju, informacije o opisu zaštite, informacije o pakiranju te opisne informacije. Informacijski paketi kasnije se dijele u tri vrste – dostavljeni informacijski paket (engl. *Submission Information Package*, SIP), arhivski informacijski paket (engl. *Archival Information Package*, AIP) i diseminacijski informacijski paket (engl. *Dissemination Information Package*, DIP). Kod modela transformacija informacijskih paketa dolazi do točnog određenja mogućih i dozvoljenih promjena informacijskih paketa koje se odvijaju po koracima – prihvata, arhivska pohrana i upravljanje podacima te pristup. Funkcionalni model sadrži šest funkcionalnih entiteta i to su: prihvata (engl. *Ingest*), arhivsku pohranu (engl. *Archival Storage*), upravljanje podacima (engl. *Data Management*), administraciju (engl. *Administration*), planiranje procesa očuvanja (engl.

⁷⁰ Isto.

⁷¹ Lavoie, *Meeting the challenges of digital preservation: The OAIS reference model*, n. dj., str. 5

Preservation Planning) i pristup (engl. *Access*). Svaki od navedenih entiteta funkcionalnog modela ima određenu funkciju, a svi imaju zajednički cilj očuvanja elektroničkih informacijskih objekata na dulji vremenski rok.⁷²

OAIS model pruža konceptualnu osnovu osiguravajući taksonomiju informacijskih objekata i paketa za arhivirane objekte te strukturu njihovih povezanih metapodataka. Na okvir se može gledati kao na razradu OAIS modela objašnjenog kroz mapiranje metapodataka za očuvanje za tu konceptualnu strukturu. Na PREMIS se može gledati kao na prijevod okvira u skup implementacijskih semantičkih jedinica u podatkovnom rječniku. PREMIS i OAIS ponekad upotrebljavaju različitu terminologiju jer PREMIS-ove semantičke jedinice zahtijevaju više specifičnosti nego OAIS definicije što je i za očekivati kada se prelazi s konceptualnog okvira u implementaciju.⁷³

6.1.2. Aktivnosti vezane uz očuvanje

PREMIS podatkovni rječnik definira metapodatke za očuvanje kao informacije koje koristi repozitorij u svrhu podupiranja procesa digitalnog očuvanja. Neki od primjera aktivnosti koji su vezani uz očuvanje njihove metapodatkovne podrške su sljedeći:

- Izvor mora biti sigurno pohranjen kako ga nitko ne bi mogao nehotice ili zlonamjerno izmijeniti. Postoje informacije o provjerama, koje su pohranjene u obliku metapodataka, a one sadrže podatke o mogućim promjenama u određenom periodu na pohranjenim podacima.
- Dokumenti moraju biti pohranjeni na mediju koji mogu čitati aktualna računala. Ako dođe do oštećenja ili zastarijevanja medija (npr. 8-inčna disketa iz 1970-ih godina) proces oporavka podataka postaje težak ili nemoguć. Metapodaci mogu pomoći u upravljanju medijima bilježenjem podataka o vrsti i godini medija na kojem su podaci pohranjeni te bilježenjem datuma o posljednjem osvježavanju podataka.
- Čak i najpopularniji podatkovni formati mogu kroz dulji period zastarjeti što dovodi do toga da ih ni jedna postojeća aplikacija ne može čitati. Kako bi se to spriječilo, najčešće se rade procesi migracije (pretvorba starih formata u novije

⁷² Stančić, H. (2005). *Teorijski model postojanog očuvanja autentičnosti elektroničkih informacijskih objekata*. Doktorska disertacija. Zagreb: Filozofski fakultet u Zagrebu. Odsjek za informacijske i komunikacijske znanost. str. 25-38

⁷³ The PREMIS Working Group, *Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group, Version 1.0.*, n. dj., str. ix

odgovarajuće) i emulacije (oponašanje stare renderirane okoline na novom hardveru i softveru). Oba procesa zahtijevaju metapodatke o originalnom formatu podataka kao i metapodatke o hardveru i softveru koji su podržavali taj format.

- Postupci očuvanja mogu zahtijevati promjenu originalnih izvora ili mogu mijenjati način njihovog vraćanja što dovodi do pitanja autentičnosti izvora. Metapodaci mogu potvrditi autentičnost dokumentiranjem digitalnog podrijetla izvora i poviješću ovlaštenih promjena.⁷⁴

6.2. PREMIS podatkovni rječnik

PREMIS podatkovni rječnik i njegova popratna dokumentacija opsežni su i praktični izvori za implementaciju očuvanja metapodataka u digitalnim sustavima arhiviranja. Podatkovni je rječnik izgrađen na modelu podataka koji definira pet entiteta: intelektualne entitete, objekte, događaje, prava i agente. Svaka semantička jedinica definirana u podatkovnom rječniku svojstvo je jednog od entiteta u podatkovnom modelu.⁷⁵

PREMIS podatkovni rječnik definira jezgroviti skup semantičkih jedinica koje bi repozitoriji trebali znati kako bi mogli izvršiti ulogu očuvanja. Te uloge mogu varirati od repozitorija do repozitorija, ali generalno uključuju postupke koji osiguravaju vidljivost i čitljivost digitalnih objekata kao i sigurnost od nehotične izmjene digitalnih objekata unutar repozitorija te dokumentaciju legitimnih izmjena objekata.⁷⁶

Namjena podatkovnog rječnika nije definirati sve postojeće elemente za očuvanje metapodataka, već samo one koje će repozitoriji najviše koristiti. Nekoliko kategorija metapodataka je isključeno uključujući: metapodatke koji se odnose samo na jedan format ili skupinu formata kao što su audio, video ili vektorske grafike; metapodatke koji opisuju politiku i praksu individualnog repozitorija; deskriptivne metapodatke (iako je opis izvora relevantan za očuvanje, u tu svrhu se koriste samostalni standardi poput MARC21, MODS-a i Dublinske jezgre); detaljne informacije o mediju odnosno hardveru (iako su ti podaci relevantni za očuvanje, ostavljeni su drugim zajednicama da ih definiraju); informacije o

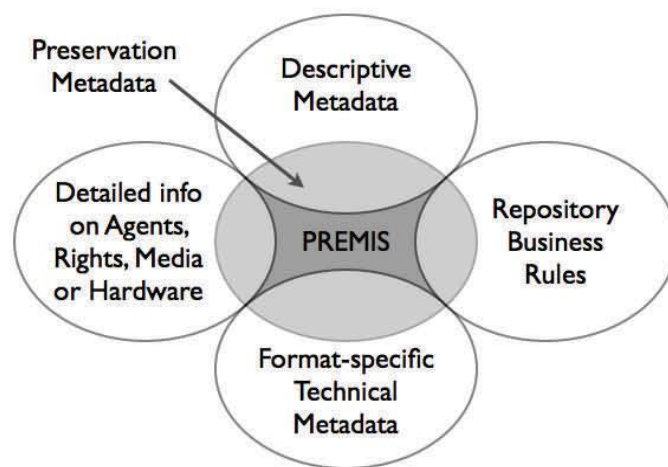
⁷⁴ Caplan, *Understanding PREMIS*, n. dj., str. 3

⁷⁵ The Library of Congress, *PREMIS: Preservation Metadata Maintenance Activity*, n. dj.

⁷⁶ Caplan, *Understanding PREMIS*, n. dj., str. 4-5

agentima (osim onih minimalnih); informacije o pravima i dozvolama (osim onih koje se odnose direktno na očuvanje).⁷⁷

Na PREMIS možemo gledati kao na podskup svih metapodataka koji se koriste za očuvanje. PREMIS ne brine o pronalasku i pristupu, a nije mu ni namjera definirati detaljne formate specifičnih metapodataka. Definiira samo one metapodatke koji su najpotrebniji za očuvanje svih vrsta materijala.⁷⁸



Slika 1: PREMIS kao jezgra metapodataka za očuvanje.

Izvor: <https://www.loc.gov/standards/premis/understanding-premis.pdf>

Slika prikazuje sve metapodatke koji su važni za očuvanje i PREMIS kao jezgru. Ovdje su uključeni neki deskriptivni metapodaci, neka pravila poslovanja, neki detaljni tehnički metapodaci te neke detaljne informacije o agentima, pravima, medijima i hardveru. PREMIS je manja jezgra u srcu očuvanja metapodataka koja isključuje sve ostale tipove.⁷⁹

PREMIS podatkovni rječnik definiira što bi repozitorij za očuvanje trebao raditi. Bitno je imati na umu da je u središtu sustav repozitorija i njegovo upravljanje, a ne autori digitalnih sadržaja ili osoblje koje evaluira i odobrava komercijalne elektroničke izvore. Primarna

⁷⁷ Isto.

⁷⁸ Isto.

⁷⁹ Isto.

upotreba PREMIS-a je za dizajn i evaluaciju repozitorija te za razmjenu arhiviranih informacijskih paketa među repozitorijima za očuvanje.⁸⁰

6.2.1. Veze

Veza (engl. *relationship*) je povezanost među instancama entiteta i to između dva ili više objekta ili između različitih tipova entiteta kao što su objekt i agent.⁸¹

6.2.2. Semantičke jedinice

PREMIS podatkovni rječnik definira semantičke jedinice, a ne elemente metapodataka. Razlika je malena, ali značajna. Semantička jedinica je dio informacije ili znanja te je svojstvo entiteta i ima svoju vrijednost. Najčešće je određena semantička jedinica svojstvo jednog tipa entiteta, ali postoje slučajevi kad je semantička jedinica svojstvo dvaju ili više tipova entiteta. U tom se slučaju veže samo s jednim tipom entiteta podatkovnog rječnika. Model se oslanja na povezanost među različitim entitetima kako bi veze bile jasne. Element metapodatka definiran je način predstavljanja informacije u metapodatkovnom zapisu, shemi ili bazi podataka. PREMIS ne određuje kako bi metapodaci trebali biti predstavljeni u pojedinom sustavu, već definira što sustav treba znati i što bi trebao moći izvesti u druge sustave. PREMIS-ove semantičke jedinice imaju direktno mapiranje elemenata metapodataka koji su definirani u PREMIS XML shemi, dok to nije uvijek tako kada govorimo o mapiranju metapodataka drugih shema. Nazivi PREMIS-ovih semantičkih jedinica nizovi su spojenih riječi (bez bjelina) prilikom čega je svako početno slovo riječi pisano velikim slovom (npr. `objectIdentifier`, `relatedEventIdentification`).^{82, 83}

6.2.3. Kontejneri, prošireni kontejneri

Neke semantičke jedinice su definirane kao kontejneri što znači da samostalno nemaju vrijednost, već kao grupa povezanih semantičkih jedinica.⁸⁴

Prošireni kontejner posebna je vrsta kontejnera koji nema podjedinice. Dizajniran je kako bi metapodaci koji nisu iz PREMIS-a na tom mjestu mogli biti zabilježeni. Na taj način PREMIS

⁸⁰ Isto.

⁸¹ The PREMIS Working Group, *Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group, Version 1.0.*, n. dj., str. 2

⁸² Isto.

⁸³ Caplan, *Understanding PREMIS*, n. dj., str. 7

⁸⁴ Isto.

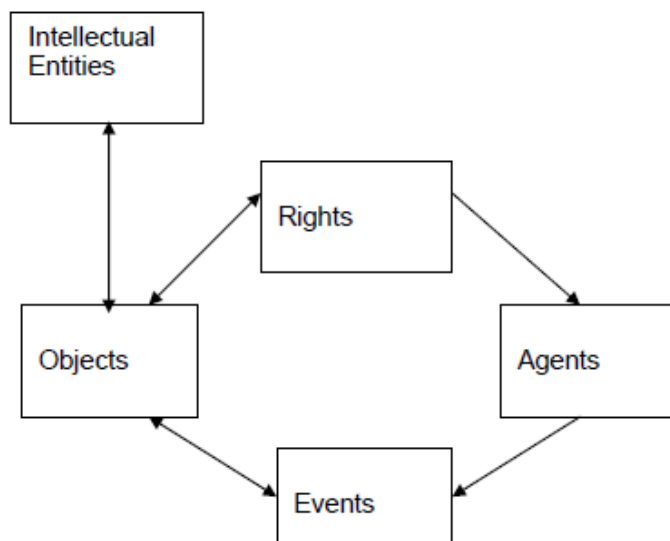
može biti proširen i može uključiti metapodatke koji su izvan spektra ili nisu uključeni u podatkovni rječnik. Npr. tehnički metapodaci o specifičnim formatima nisu uključeni u PREMIS, ali čine važne informacije za digitalno očuvanje. Prošireni kontejner *objectCharacteristicsExtension* daje mjesta za zapis tehničkim metapodacima koji su definirani prema drugim podatkovnim rječnicima kao što je Z39.87 standard za opis bitmapiranih slika.⁸⁵

PREMIS podatkovni rječnik dizajniran je da bude kompatibilan s XML-om te njegove semantičke jedinice mogu biti implementirane kao XML elementi. Kontejneri su elementi koji uzimaju samo druge elemente kao sadržaj, a prošireni kontejneri su kontejneri za elemente koji su definirani prema vanjskoj shemi.⁸⁶

⁸⁵ Caplan, *Understanding PREMIS*, n. dj., str. 8

⁸⁶ Isto.

6.3. PREMIS podatkovni model



Slika 2: PREMIS podatkovni model.

Izvor: <http://www.loc.gov/standards/premis/v1/index.html>

Kao što je već spomenuto, PREMIS podatkovni model (engl. *PREMIS Data Model*) definira pet tzv. entiteta: intelektualne entitete, objekte, agente, događaje i prava (slikovni prikaz). U dijagramu podatkovnog modela entiteti su smješteni u kutije dok su veze među njima prikazane u obliku strelica. Smjer u kojem strelice pokazuju pokazuje smjer veze definirane u podatkovnom rječniku. Na primjer strelica koja kreće iz kutije prava (engl. *rights*) prema agentima (engl. *agents*) znači da metapodaci koji su definirani za prava uključuju semantičke jedinice za identifikaciju povezanih agenata. Strelice koje pokazuju u oba smjera označavaju postojanje definirane veze za oba smjera.^{87, 88} „Cilj i svrha primjene takvog modela jest pridružiti metapodatke entitetima i digitalnim objektima u repozitoriju, uspostaviti odnose među digitalnim objektima, pružiti digitalnim objektima podatke o događajima (poduzetim postupcima zaštite) te pridružiti digitalnim objektima metapodatke o pravima i ograničenjima prava vezanim uz zaštitu.“⁸⁹

⁸⁷ Isto.

⁸⁸ The PREMIS Working Group, *Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group, Version 1.0.*, n. dj., str. 1

⁸⁹ Klarin, S. (2008.) *Tehnički metapodaci i metapodaci za zaštitu u normama ANSI/NISO Z39.87 i PREMIS*. 11. seminar Arhivi, knjižnice, muzeji: mogućnost suradnje u okruženju globalne informacijske infrastrukture. uredila Mirna Willer. Zagreb: Hrvatsko knjižničarsko društvo. str. 155-156

6.3.1. Intelektualni entiteti

PREMIS definira intelektualni entitet kao skup sadržaja koji se smatra zasebnom intelektualnom jedinicom koja služi za upravljanje i opis, npr. knjiga, karta, fotografija ili baza podataka. Intelektualni entiteti mogu obuhvaćati druge intelektualne entitete; na primjer web stranica može obuhvaćati i fotografije prema čemu se zaključuje da intelektualni entitet može imati jedan ili više digitalnih prikaza. Intelektualni entiteti su konceptualni te se mogu nazivati i „bibliografski entiteti“. PREMIS ne definira metapodatke koji se odnose na intelektualne entitete jer za to postoje mnogi deskriptivni metapodatkovni standardi. PREMIS tvrdi da bi objekt u sustavu očuvanja trebao biti povezan s intelektualnim entitetom koji ga predstavlja tako da postoji identifikator intelektualnog entiteta u metapodacima objekta.^{90, 91}

6.3.2. Objekti

Objekti su oni entiteti koji su pohranjeni i kojima se upravlja u repozitoriju očuvanja. To su jedinice informacija u digitalnom obliku. Većina PREMIS-a je posvećena opisu digitalnih objekata. Informacije koje mogu biti zabilježene uključuju:⁹²

- jedinstveni identifikator objekta (tip i vrijednost),
- veličinu objekta,
- format objekta koji može biti direktno specificiran ili povezan s registrom formata,
- originalni naziv objekta,
- informacije o stvaranju objekta,
- informacije o inhibitornima,
- informacije o značajnim svojstvima,
- informacije o okruženju objekta,
- gdje i na kojem mediju je objekt pohranjen,
- informacije o digitalnom potpisu,
- povezanost s dugim objektima i drugim vrstama entiteta.⁹³

⁹⁰ Caplan, *Understanding PREMIS*, n. dj., str. 9

⁹¹ The PREMIS Working Group, *Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group, Version 1.0.*, n. dj., str. 1

⁹² Caplan, *Understanding PREMIS*, n. dj., str. 9

⁹³ Isto.

Neke semantičke jedinice su definirane kako bi zabilježile okolinu objekta, točnije koji hardver i softver je potreban za renderiranje te kako one ovise o drugim objektima. Npr. PDF dokument može biti prikazan pomoću nekoliko verzija Adobe Acrobat i Adobe Readera kao i pomoću drugih komercijalnih programa otvorenog pristupa. Svakog od njih podržavaju različiti operacijski sustavi te zahtijevaju određene minimalne specifikacije hardvera (brzina procesora, memorija diska). S obzirom na to da Adobe Reader nije samostalna aplikacija, već se veže uz preglednik, ovisi o određenim verzijama određenih preglednika; npr. Reader 9 za Mac operacijski sustav zahtijeva preglednik Safari od verzije 2.0.4 nadalje.⁹⁴

Kada se detaljne informacije primjenjuju na klasama objekata više nego na lokalnim instancama objekata, kao što je slučaj s informacijama o okolini, informacije mogu biti agregirane u registrima te ih mogu svi repozitoriji dijeliti. PREMIS dopušta da se repozitoriji povezuju s informacijama koje su pohranjene u vanjskim registrima (kad je to bolje rješenje od lokalne pohrane). PREMIS definira tri različite podvrste objekata te zahtijeva da ih se razlikuje. Te podvrste objekata su: datoteka (engl. *file*), reprezentacija (engl. *representation*) i niz bitova (engl. *bitstream*).⁹⁵

6.3.2.1. Datoteka

Pod pojmom datoteka misli se na uobičajenu računalnu datoteku u npr. PDF ili JPEG formatu. Datoteka se može čitati, pisati i kopirati te ima svoje ime i format.^{96,97}

6.3.2.2. Reprezentacija

Reprezentacija je skup svih datoteka objekta koje su potrebne za ponovno oblikovanje (renderiranje) intelektualnog entiteta. Želimo li npr. očuvati određenu web stranicu, vrlo je vjerojatno da se početna stranica, koju vidimo pomoću preglednika, sastoji od mnogo različitih datoteka. To mogu biti jedna ili više HTML datoteke, mnogo GIF ili JPEG slika, Flash animacije itd. Pregledniku su potrebne sve navedene datoteke kako bi oblikovao početnu stranicu. Želimo li očuvati i pohraniti stranicu koju ćemo moći ponovno oblikovati u repozitorij, moramo očuvati i pohraniti sve navedene datoteke te ih moramo znati ponovno

⁹⁴ Isto.

⁹⁵ Isto.

⁹⁶ Caplan, *Understanding PREMIS*, n. dj., str. 10

⁹⁷ The PREMIS Working Group, *Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group, Version 1.0.*, n. dj., str. 2

sastaviti. Reprerentacija omogućava repozitoriju da identificira skup povezanih datoteka, ali i da opiše karakteristike web stranice kao cjeline. U repozitoriju za očuvanje može se dogoditi da postoji više od jedne reprezentacije za jedan intelektualni entitet. Npr. postoji slika u jednom formatu, a kasnije se iz toga napravi drugi derivirani format. U tom se slučaju zadržavaju oba formata, a svaki od njih predstavlja reprezentaciju te slike.^{98,99}

6.3.2.3. Nizovi bitova

Nizovi bitova su podskupovi datoteka. Nizovi bitova su definirani kao podaci (bitovi) unutar datoteke koji imaju zajednička svojstva vezana uz očuvanje te ne mogu biti samostalni ako nemaju dodano zaglavlje datoteke ili neku drugu strukturu. Primjer toga je datoteka u AVI formatu kod koje možemo (ako želimo) razdvojiti audio niz bitova od video niza bitova te ih opisati kao odvojene nizove bitova.¹⁰⁰

Neki se nizovi bitova mogu transformirati u samostalne datoteke bez dodavanja dodatnih informacija, ali postoji mogućnost da će se prije toga trebati obaviti proces dekompresije ili dekriptiranja prilikom ekstrakcije. U PREMIS podatkovnom modelu ti se nizovi bitova nazivaju nizovi datoteka (engl. *filestreams*). Za razliku od nizova bitova, nizovi datoteka imaju karakteristike datoteka.¹⁰¹

Neke se semantičke jedinice definirane u PREMIS podatkovnom rječniku mogu primijeniti na sve tri podvrste objekta dok se neke mogu primijeniti na jednu ili dvije podvrste. Postojanje različitih podvrsta objekata navodi nas na razmišljanje o tome što opisujemo te nas potiče na maksimalnu preciznost što je važno za strojnu obradu.¹⁰²

6.3.3. Događaji

Entitet događaji agregiraju informacije o radnjama koje utječu na barem jedan objekt ili na jednog agenta u repozitoriju za očuvanje. Točnost i pouzdanost zapisa su važni kako bi se dokazala autentičnost objekta što je ponekad teško održati.¹⁰³

Informacije o događajima koje mogu biti zabilježene uključuju:

⁹⁸ Caplan, *Understanding PREMIS*, n. dj., str. 10

⁹⁹ The PREMIS Working Group, *Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group, Version 1.0.*, n. dj., str. 2

¹⁰⁰ Caplan, *Understanding PREMIS*, n. dj., str. 10

¹⁰¹ The PREMIS Working Group, *Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group, Version 1.0.*, n. dj., str. 3

¹⁰² Caplan, *Understanding PREMIS*, n. dj., str. 10

¹⁰³ Isto.

- jedinstveni identifikator događaja (tip i vrijednost),
- vrstu događaja (stvaranje, migracija, itd.),
- datum i vrijeme odvijanja događaja,
- detaljan opis događaja,
- kodirani rezultat događaja,
- detaljniji opis rezultata,
- agente uključene u događaj i njihove uloge,
- objekte uključene u događaj i njihove uloge.¹⁰⁴

Svaki sustav repozitorija mora donijeti vlastite odluke o tome koje će događaje zabilježiti kao trajne dijelove povijesti objekta. PREMIS preporučuje bilježenje svih radnji koje mijenjaju objekt. Unos u polje tip događaja (engl. *eventType*) podatkovnog rječnika daje popis važnih tipova događaja koji potiču repozitorije na dosljedno bilježenje događaja. Uz to postoji i mogućnost bilježenja događaja koji su se dogodili prije pohrane objekta u repozitorij očuvanja kao npr. derivacija iz prijašnjeg objekta.^{105, 106}

6.3.4. Agenti

Agenti mogu biti ljudi, organizacije ili softverske aplikacije koji su povezane s očuvanjem objekta, tj. imaju ulogu u događajima i u pravima. PREMIS definira minimalan broj semantičkih jedinica koje su potrebne za identifikaciju agenata zbog postojanja nekoliko vanjskih standarda koji mogu bilježiti više informacija. Repozitorij može koristiti zaseban standard za bilježenje dodatnih informacija o agentima ili može koristiti identifikator agenta koji može nekoga usmjeriti na informacije zabilježene izvana.¹⁰⁷

Podatkovni rječnik uključuje:

- jedinstveni identifikator agenta (tip i vrijednost),
- ime agenta,
- oznaku tipa agenta (osoba, organizacija, softver).¹⁰⁸

Kada je agent povezan s događajem ili pravima, trebala bi se zabilježiti i uloga agenta. Svaki agent može imati više uloga. Netko može, na primjer, biti autor i nositelj prava jednog djela,

¹⁰⁴ Isto.

¹⁰⁵ Isto.

¹⁰⁶ The PREMIS Working Group, *Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group, Version 1.0.*, n. dj., str. 5-6

¹⁰⁷ Caplan, *Understanding PREMIS*, n. dj., str. 11

¹⁰⁸ Isto.

autor (ali ne i nositelj prava) drugog djela te ulagač (pokrovitelj, sponzor) trećeg djela. U PREMIS-ovom modelu, repozitorij dodjeljuje jedinstveni identifikator toj osobi te bilježi taj identifikator u svakom zapisu događaja ili pravima u kojem je ta osoba agent zajedno s ulogom te osobe u tom određenom kontekstu. Uloga agenta vezana uz događaj ili prava smatra se vlasništvom događaja ili prava, a ne samog agenta.¹⁰⁹

6.3.5. Prava

Većina strategija očuvanja uključuje stvaranje identičnih kopija i izvedenih verzija digitalnih objekata - radnje koje su ograničene zakonom o autorskim pravima za nositelje prava. Entitet prava agregira informacije o pravima i dozvolama koje su izravno vezane uz očuvanje objekata u repozitorijima.¹¹⁰

Informacije koje mogu biti zabilježene unutar prava uključuju:

- jedinstveni identifikator prava (tip i vrijednost),
- ukoliko su osnova za potraživanje prava autorska prava, licenca ili statut, potrebno je više informacija o statusu autorskih prava, odredbama licence ili statuta,
- radnje koje prava dozvoljavaju,
- ograničenja radnji,
- termine odobravanja ili vremenski period primjene prava,
- objekte na koje se primjenjuju prava,
- agente koji su uključeni u pravima te njihove uloge.¹¹¹

Većina informacija zabilježena je u kontroliranoj formi tako da mogu raditi u skladu s računalnim programom. PREMIS-ova prava nisu zabilježene informacije prema kojima se prava određuju, već su to navodi prava. Kod PREMIS-a je specifično i to da ne definira detaljne informacije o autorima, datumima i mjestima izdavanja te o autorskim pravima koja su definirana u specifikacijama Kalifornijske digitalne knjižnice (engl. *California Digital Library's copyright specification*). Svrha tih specifikacija je da pomaže ljudima prilikom donošenja odluka o pravima na trajnoj osnovi, a svrha PREMIS-ovih entiteta prava je

¹⁰⁹ Isto.

¹¹⁰ Isto.

¹¹¹ Isto.

osigurati informacije koje će biti djelotvorne za očuvanje sustava repozitorija očuvanja ostavljajući prava vezana uz pristup i distribuciju po strani.¹¹²

6.3.6. Veze među objektima

Strukturalne veze ukazuju na veze među dijelovima objekata, dok strukturalne veze među datotekama, koje sadrže reprezentaciju intelektualnog entiteta, imaju ulogu važnih metapodataka za očuvanje. Ako repozitorij za očuvanje ne može ponovno sastaviti dijelove digitalnog objekta, to znači da objekt nije očuvan.¹¹³

Derivirane su veze rezultat kopija ili transformacija objekata. U tom je slučaju intelektualni sadržaj dobivenog objekta isti, ali je instanca objekta, kao i njegov format, drugačija.¹¹⁴

Strukturalne veze među objektima mogu biti uspostavljene postupkom deriviranja prije ulaska objekata u repozitorij.¹¹⁵

Ne postoji način modeliranja svih postojećih strukturalnih i deriviranih informacija. Umjesto specificiranja određenog pristupa, skupina je odredila nužne informacije koje moraju biti zabilježene. Strukturalne i derivirane veze povezuju objekte koji moraju biti identificirani. Vrsta se veze na neki način mora identificirati, a ona može biti povezana s događajem koji je stvorio tu vezu. Implementatori mogu odabrati pristup koji najviše odgovara sadržaju koji treba biti očuvan upotrebom npr. METS-ove strukturalne mape.¹¹⁶

Zavisna veza postoji kada je jednom objektu potreban drugi kako bi podržao svoju funkciju, dostavu ili koherentnost sadržaja. Podatkovni rječnik odnosi se prema zavisnim vezama kao prema dijelovima informacija o okolini. Na taj su način zahtjevi za hardver i softver spojeni s zahtjevima zavisnih datoteka te tvore potpunu sliku informacija ili aktiva koje su potrebne za renderiranje i razumijevanje objekta.¹¹⁷

¹¹² Isto.

¹¹³ The PREMIS Working Group, *Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group, Version 1.0.*, n. dj., str. 8

¹¹⁴ Isto.

¹¹⁵ Isto.

¹¹⁶ The PREMIS Working Group, *Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group, Version 1.0.*, n. dj., str. 9

¹¹⁷ Isto.

6.4. PREMIS verzija 3.0

Prva verzija PREMIS podatkovnog rječnika 1.0 je nastala u razdoblju od 2003. do 2005. godine. Nakon toga je slijedilo još nekoliko verzija PREMIS podatkovnog rječnika i to verzije 2.0 2008. godine, 2.1 2011. godine, 2.2 2012. godine i 2.3 2014. godine u kojoj je nastala samo promjena sheme, ali ne i podatkovnog rječnika. Najnovija verzija PREMIS podatkovnog rječnika je 3.0 iz 2015. godine.

U najnovijoj verziji PREMIS podatkovnog modela došlo je do nekih promjena. Intelektualni entitet postaje druga kategorija objekta prilikom čega ima jednake semantičke jedinice kao reprezentacije. U verzijama 1 i 2 intelektualni su entiteti bili izvan dosega, a semantičke jedinice koje su ih opisivale nisu bile uključene u PREMIS podatkovni model (osim za identifikatora kako bi se omogućilo povezivanje s opisom). PREMIS-ov urednički odbor dobio je nekoliko zahtjeva za proširenje podatkovnog rječnika kako bi se mogli intelektualni entiteti opisati i kako bi se omogućilo njihovo povezivanje s drugim PREMIS-ovim entitetima. Kao podrška za to razvili su se slučajevi upotrebe.¹¹⁸

Jedan slučaj upotrebe je kada repozitorij želi prikazati skupove kao što su zbirke, rad funkcionalnih zahtjeva za bibliografske zapise, fondove ili serije. Namjera repozitorija je dohvaćanje deskriptivnih metapodataka, zadovoljavanje zahtjeva djelatnosti (poput značajnih karakteristika, definicija rizika, smjernica za očuvanje, itd.), podržavanje strukturalne i derivirane veze te omogućavanje veze s pravima i događajima. Drugi slučaj upotrebe je kada repozitorij želi dohvatiti različite verzije informacija i metapodatke najnovijih događaja na razini intelektualnog entiteta za intelektualne entitete kao što su članci ili izdanja.¹¹⁹

¹¹⁸ The Library of Congress. (2015). *PREMIS Data Dictionary for Preservation Metadata: Approved Changes for version 3.0 (from previous version 2.2)*. [citirano: 2016-10-27]

Dostupno na: <http://www.loc.gov/standards/premis/changes-3-0.html>

¹¹⁹ Isto.

6.5. PREMIS u XML

Najčešće se očekuje da se PREMIS prilikom razmjene prikaže u XML formatu. PREMIS aktivnost održavanja (engl. *PREMIS Maintenance Activity*), koju održava Kongresna knjižnica, osigurava XML shemu koja odgovara podatkovnom rječniku te omogućava direktan opis objekata, događaja, agenata i prava.¹²⁰

U praksi većina sustava za očuvanje koriste XML format za unošenje i izvođenje podataka. Mnogo ih koristi Metapodatkovni standard za kodiranje i prijenos (engl. *The Metadata Encoding and Transmission Standard, METS*). To je standard kojeg održava Kongresna knjižnica kao XML kontejner za skupljanje različitih tipova metapodataka. PREMIS se može koristiti unutar METS-a, ali ne u potpunosti. METS dijeli informacije na različite sekcije prema tehničkim metapodacima, metapodacima o pravima i metapodacima o podrijetlu. PREMIS-ova shema (prema podatkovnom rječniku) ima sekcije za objekte, prava, događaje i agente. Između tih dviju struktura postoji korespondencija, ali ona nije savršena (naročito za agente). PREMIS i METS se ponekad preklapaju pa je potrebno odlučiti (u slučaju upotrebe i PREMIS-a i METS-a zajedno), hoće li se bilježiti ti preklapajući elementi u sekciji PREMIS-a, sekciji METS-a ili oboje.¹²¹

Kad bi svaki repozitorij koji se bavi očuvanjem donosio svoje odluke, došlo bi do razlika u prikazu podataka što bi otežalo interoperabilnost. Zbog toga su u tijeku neki procesi koji će pomoći u definiranju najboljeg rješenja za upotrebu PREMIS-a i METS-a zajedno.¹²²

```
-->  
<xs:element name="premis" type="premisComplexType"/>  
<xs:element name="object" type="objectComplexType"/>  
<xs:element name="event" type="eventComplexType"/>  
<xs:element name="agent" type="agentComplexType"/>  
<xs:element name="rights" type="rightsComplexType"/>  
- <!--
```

Slika 3: Deklaracija elemenata u PREMIS-u.

Izvor: <http://www.loc.gov/standards/premis/premis.xsd>

¹²⁰ Caplan, *Understanding PREMIS*, n. dj., str. 14-15

¹²¹ Isto.

¹²² Isto.

7. Islandora

Islandora je programski okvir za softver otvorenog koda izrađen s ciljem da pomaže institucijama, organizacijama i njihovim korisnicima u zajedničkom upravljanju i otkrivanju digitalne imovine. Robertson knjižnica Sveučilišta Prince Edward Island (engl. *University of Prince Edward Island's Robertson Library*) je uspostavila Islandoru, ali trenutnom razvoju sve više doprinosi međunarodna zajednica. Zajednica radi na primjeni i razvoju što je moguće transparentnije s ciljem da dokumentacija bude što bolja kao i kod otvorenog pristupa. Islandora se temelji na Drupalu, Fedori i Apache Solru, a rješenja koja izdaje potiču korisnike na rad s raznim tipovima podataka poput slika, videa i PDF-ova te s raznim domenama znanja. Paketna rješenja često nude integraciju s dodatnim preglednicima, uređivačima te aplikacijama za obradu podataka.^{123, 124}

Dozvole (engl. *permissions*) se u Islandori određuju prema ulogama (engl. *role*), a svaku ulogu čini određena skupina korisnika. Kad se pojavljuje novi korisnik, njemu se dodjeljuje uloga, a time i određene dozvole. Dozvole se manifestiraju kroz Drupal, a dozvole određenih korisnika se mogu vidjeti prijavom osobe u Islandoru, odabirom kartice *People* zatim *Permissions* gdje su vidljive različite uloge te njihove dozvole (tko može što raditi).¹²⁵

Struktura (engl. *structure*) u Islandori jest takva da se sve naziva objektom (npr. zbirka, PDF, slika, video itd.). Objekti se sastoje od podatkovnih sljedova (engl. *datastream*), a primjer toga su deskriptivni metapodaci. Nakon odabira neke slike iz kolekcije, odabire se kartica *Manage*, a zatim kartica *Datastream* nakon čega je vidljiva lista komponenata od kojih se određeni objekt sastoji. Neki od podatkovnih sljedova su *OBJ* što označava glavni objekt za očuvanje, *MODS* podatkovni slijed deskriptivnih metapodataka koji se automatski uparuje s *DC-om*, derivati slika koji se najčešće automatski izvode (*JPG*, *TN*) i drugi.¹²⁶

U Islandori postoji nekoliko metoda umetanja (engl. *ingest methods*). Jedna od metoda je metoda individualnog umetanja putem weba (Drupal sučelja). Druga je serijsko umetanje putem weba (Drupal sučelja) koja se dijeli na: a) umetanje ZIP datoteke (koja se sastoji od više JPG-ova, TIFF-ova ili MP4 i oni moraju biti odvojeni s obzirom na tip datoteka, te

¹²³ Islandora. *About*. [citirano: 2016-12-01] Dostupno na: <http://islandora.ca/about>

¹²⁴ Discoverygarden Inc. (2016, May 30). *Islandora Webinar: A „How To“ Guide for Islandora*. [citirano: 2016-12-07] Dostupno na: <https://www.youtube.com/watch?v=I06UOfHevtQ>

¹²⁵ Isto.

¹²⁶ Isto.

opcionally XML datoteka koja opisuje te datoteke pri čemu je bitno da su nazivi npr. JPG i XML datoteke isti), b) serijsko umetanje *Four Batch Citation Uploaders* koje se odnosi na znanstvene publikacije s jedinstvenim identifikatorom. Treća metoda umetanja odvija se direktno u Fedori ili putem skriptiranih naredbi, a koristi se najčešće za velike serije ili datoteke veće od 2 GB.¹²⁷

7.1. Drupal

Drupal je softver otvorenog pristupa za upravljanje sadržajem, a koristi se za izradu mnogih web stranica i aplikacija današnjice. U Islandori on predstavlja prezentacijski sloj. Neke od karakteristika Drupala su jednostavno stvaranje sadržaja, pouzdan rad te visoka razina sigurnosti, a karakteristike po kojima se ističe od drugih su fleksibilnost i modularnost. Drupal sadrži i mnoge alate koji pomažu u oblikovanju fleksibilnog i strukturalnog sadržaja koji čine dinamičan web. Drupal je dobar odabir pri stvaranju integriranih digitalnih okvira koji se mogu proširiti raznim dodacima. Modulima se može proširiti funkcionalnost Drupala dok se temama prilagođava prezentacija sadržaja. Drupal se može i integrirati u vanjske usluge i druge aplikacije postojeće infrastrukture. Drupal zajednica je jedna od najvećih zajednica otvorenog pristupa koja ima veliku ulogu u razvoju softvera, a čine ju programeri, dizajneri, stratezi, koordinatori, urednici te sponzori koji rade zajedno. Oni izgrađuju Drupal, nude podršku, stvaraju dokumentaciju i dr. Drupal 8 je posljednja aktualna verzija na kojoj su se desila najveća ažuriranja. Olakšano je stvaranje sadržaja, svaka ugrađena tema je responzivnog dizajna, dostupan je na više od sto jezika, sadrži mnogo integriranih alata, a rad više od 4500 ljudi pridonio je njegovom razvoju što je sve dovelo do više od dvjesto novih karakteristika Drupala.^{128, 129}

¹²⁷ Isto.

¹²⁸ Drupal. *About*. [citirano: 2016-12-04] Dostupno na: <https://www.drupal.org/about>

¹²⁹ Discoverygarden Inc., *Islandora Webinar: A „How To“ Guide for Islandora*, n. dj.

7.2. Fedora

Arhitektura fleksibilnog proširivog repozitorija digitalnih objekata (engl. *Flexible Extensible Digital Object Repository Architecture*, Fedora) nastala je 1997. godine na Sveučilištu Cornell, a od 1998. godine je dostupna javnosti. Fedora je robustan i modularan sustav repozitorija koji je otvorenog pristupa, a služi za upravljanje, diseminaciju, pohranu, dugoročno očuvanje, migraciju i pristup digitalnom sadržaju. Posebno je prilagođen za digitalne knjižnice i arhive kojima je svrha pristup i očuvanje sadržaja. Koristi se i za omogućavanje specijaliziranog pristupa velikim i kompleksnim digitalnim zbirkama povijesnih, kulturnih i znanstvenih podataka. Fedora se koristi diljem svijeta i to u akademskim i kulturnim zajednicama, sveučilištima, istraživačkim institucijama, sveučilišnim i nacionalnim knjižnicama te vladinim agencijama. Projekt Fedora vodi *Fedora Leadership Group* uz pomoć neprofitne organizacije *DuraSpace* koja nudi vodstvo i inovacije za tehnološke projekte otvorenog pristupa, financijsku podršku te rješenja kojima je cilj ostvariti trajni i stalni pristup digitalnim podacima. Posljednja verzija Fedore je 4.0.0 koja je 2014. godine puštena u produkciju.^{130, 131}

7.3. Apache Solr

Apache Solr je platforma za pretraživanje cjelovitog teksta koja je otvorenog pristupa. Neke od karakteristika Solra su fleksibilnost, pouzdanost, skalabilnost, tolerancija na pogreške, indeksiranje u gotovo stvarnom vremenu, izrada kopija, filtriranje, automatizirano preuzimanje sustava kad se kod glavnog sustava dogodi neka greška i dr. Solr ima karakteristične mehanizme za brzo pretraživanje pomoću kojih pretražuje ono što se nalazi u Fedori.^{132, 133}

¹³⁰ Fedora. *About Fedora*. [citirano: 2016-12-05] Dostupno na: <http://www.fedora-commons.org/about>

¹³¹ Discoverygarden Inc., *Islandora Webinar: A „How To“ Guide for Islandora*, n. dj.

¹³² Solr. *Learn more about Solr*. [citirano: 2016-12-07] Dostupno na: <http://lucene.apache.org/solr/>

¹³³ Discoverygarden Inc., *Islandora Webinar: A „How To“ Guide for Islandora*, n. dj.

7.4. Primjena Islandore

U Kolekciju PDF-ova dodan je PDF dokument. Odabirom kartice Pregled (engl. *View*) vidljiv je dodani dokument te je vidljivo u kojoj se kolekciji nalazi. Proširenjem mogućnosti Detalji (engl. *Details*) vidljive su neke informacije o dokumentu te se ovdje (ovisno o ulozi koju pojedini član ima) dokument može i preuzeti.

Unutar kartice Upravljanje (engl. *Manage*) nalaze se druge kartice s raznim mogućnostima. Kartica Pregled (engl. *Overview*) nudi mogućnost migracije odabranog objekta u drugu kolekciju i mogućnost dijeljenja objekta s drugom kolekcijom (u tom se slučaju objekt nalazi u obje kolekcije). Osim toga kartica Pregled prikazuje koja je roditeljska (engl. *parent*) kolekcija objekta, odnosno u kojoj se kolekciji objekt nalazi.

Odabirom kartice Podatkovni sljedovi (engl. *Datastream*) vidljivi su svi zapisi koji se ovdje nalaze (Slika 4) i neke informacije o njima. Datoteka RELS-EXT sadrži sve veze među objektima u Fedori.

Arhivsko gradivo u elektroničkom obliku: mogućnost zaštite i očuvanja na dulji vremenski rok - kmffzg:62

ID	Label	Type	Mime type	Size	Versions	Operations				
RELS-EXT	Fedora Object to Object Relationship Metadata.	Inline XML	application/rdf+xml	914 B	1		download			
MODS	MODS Record	Managed	application/xml	718 B	1	replace	download	edit		
DC	DC Record	Inline XML	text/xml	528 B	1	replace	download			
OBJ	Arhivsko gradivo.pdf	Managed	application/pdf	134.56 KiB	1	replace	download			
POLICY	XACML Policy Stream	Managed	application/xml	3.54 KiB	1	replace	download			
TN	TN	Managed	image/jpeg	2.49 KiB	1	replace	download			regenerate
PREVIEW	PREVIEW	Managed	image/jpeg	31.39 KiB	1	replace	download			regenerate
FULL_TEXT	FULL_TEXT	Managed	text/plain	39.45 KiB	1	replace	download			regenerate

Slika 4: Popratni dokumenti objekta pohranjenog u Islandori.

Izvor: <https://kolegij-metapodaci.ffzg.unizg.hr/islandora/object/kmffzg%3A62/manage/datastreams>

Datoteka MODS sadrži MODS zapis objekta na temelju danih informacija prilikom unosa samog objekta i određenih informacija o njemu. MODS zapis može se uređivati pa je nakon

odabira mogućnosti Uređivanja (engl. *Edit*) MODS zapisa te učinjenih izmjena i pohrane tih izmjena, vidljivo i u prikazu da je došlo do promjena (sada postoje dvije verzije MODS i DC zapisa) (Slika 5). Klikom na broj verzija, vidljive su sve verzije i moguće je odabrati i vratiti zapis na neku od prijašnjih verzija. Datoteka DC sadrži DC zapis objekta (gotovo isto kao i MODS zapis), ali se DC zapis ne može uređivati. Ipak, prilikom uređivanja MODS zapisa dolazi do pojave i druge verzije DC zapisa jer su oni međusobno povezani.

Arhivsko gradivo u elektroničkom obliku: mogućnost zaštite i očuvanja na dulji vremenski rok - kmffzg:62

ID	Label	Type	Mime type	Size	Versions	Operations				
RELS-EXT	Fedora Object to Object Relationship Metadata.	Inline XML	application/rdf+xml	914 B	1		download			
MODS	MODS Record	Managed	application/xml	806 B	2	replace	download	edit		
DC	DC Record	Inline XML	text/xml	567 B	2	replace	download			
OBJ	Arhivsko gradivo.pdf	Managed	application/pdf	134.56 KiB	1	replace	download			
POLICY	XACML Policy Stream	Managed	application/xml	3.54 KiB	1	replace	download			
TN	TN	Managed	image/jpeg	2.49 KiB	1	replace	download			regenerate
PREVIEW	PREVIEW	Managed	image/jpeg	31.39 KiB	1	replace	download			regenerate
FULL_TEXT	FULL_TEXT	Managed	text/plain	39.45 KiB	1	replace	download			regenerate

Slika 5: Promjene nakon uređivanja MODS zapisa.

Izvor: <https://kolegij-metapodaci.ffzg.unizg.hr/islandora/object/kmffzg%3A62/manage/datastreams>

OBJ označava sam objekt (u ovom slučaju PDF dokument), a klikom na nj vidljiv je cijeli objekt, tj. dokument. Datoteka „politika“ (engl. *Policy*) je u XACML obliku, a daje informacije o tome tko može upravljati objektom, kome je objekt vidljiv te tko ima pristup određenim podatkovnim sljedovima. TN označava Thumbnail sliku, tj. smanjenu verziju prikaza objekta, dok PREVIEW prikazuje objekt u veličini koja nije smanjena. Odabirom FULL_TEXT prikazuje se potpun (u ovom slučaju) tekst iz PDF dokumenta, ali ne u PDF obliku, već u pregledniku u nesređenom obliku. Povrh svega navedenog, u ovoj se kartici mogu dodati i željeni podatkovni sljedovi.

U kartici Svojstva (engl. *Properties*) može se promijeniti naziv objekta koji mu je dan u Islandori, tzv. vlasnik objekta te status objekta (aktivan, neaktivan ili obrisan).

U kartici Spajanje (engl. *Compound*) objektu se može dodati roditeljski objekt, tj. odrediti kojem objektu pripada.

U kartici Politika objekta (engl. *Object Policy*) mogu se uključiti zabrane za upravljanje i pregled objekta te zabrane koje kontroliraju pristup korisnika individualnim podatkovnim sljedovima određenog objekta ili kolekcije. Ipak, prilikom uključivanja zabrana može se odrediti koji korisnici ili koje uloge imaju dozvolu, tj. na koga se zabrane ne odnose.

Kartica PREMIS nudi mogućnost preuzimanja PREMIS dokumenta u XML obliku, a taj dokument predstavlja PREMIS metapodatke za objekte u Islandora repozitoriju. Samim otvaranjem kartice vidljiva je podjela na polja vezana uz PREMIS objekt, PREMIS događaje, PREMIS agente te PREMIS prava. Svako polje od navedenih sadrži unutar sebe tablicu s Poljima (engl. *Field*) te pripadajućim odgovarajućim vrijednostima (engl. *Value*). PREMIS datoteka nije pohranjena kao podatkovni slijed u Fedori i ne može biti ručno uređivana.

8. Zaključak

Iako je proces digitalnog dugoročnog očuvanja prilično nov i tek kreće uzlaznom putanjom, pojam očuvanja (u kontekstu svakodnevice) je svima poznat. Ljudi spremaju svoje (ili tuđe) stvari ili dokumente kako bi ih oni sami ili neke nadolazeće generacije mogle ponovno upotrijebiti. Svakom je ponaosob neki predmet ili dokument važan, a važnost akademskih, znanstvenih i istraživačkih radova za širu populaciju još je veća. To dovodi do važnosti i samog procesa kvalitetnog dugoročnog očuvanja, postojanja repozitorija kao mjesta pohrane željenih dokumenata itd. Iako u ovom procesu postoji još mjesta za razvoj i poboljšanje, PREMIS kao međunarodni standard za metapodatke, koji pomaže u očuvanju digitalnih objekata i time omogućava njihovu dugoročnu upotrebu, te Islandora kao programski okvir za softver otvorenog koda izrađen s ciljem da pomaže institucijama, organizacijama i njihovim korisnicima u zajedničkom upravljanju i otkrivanju digitalne imovine su dobar odabir za početak procesa dugoročnog očuvanja.

9. Literatura

1. Armbruster, C., & Romary, L. (2010). *Comparing Repository Types – Challenges and barriers for subject-based repositories, research repositories, national repository systems and institutional repositories in serving scholarly communication*. International Journal of Digital Library Systems. IGI Global. Dostupno na: <https://hal.inria.fr/inria-00465197v2>
2. arXiv.org. (1991). *General Information About arXiv*. Dostupno na: <https://arxiv.org/help/general>
3. Bosančić, B. (2009). *Repozitoriji ustanova u teoriji i praksi*. 12. seminar Arhivi, knjižnice, muzeji: mogućnost suradnje u okruženju globalne informacijske infrastrukture. uredila Sanjica Faletar Tanacković. Zagreb: Hrvatsko knjižničarsko društvo.
4. Caplan, P. (2009). *Understanding PREMIS*. Washington, D. C.: The Library of Congress. Dostupno na: <https://www.loc.gov/standards/premis/understanding-premis.pdf>
5. CARNet (2008). *CMS sustavi i sigurnost*. Zagreb: Hrvatska akademska i istraživačka mreža. Dostupno na: <http://www.cis.hr/www.edicija/LinkedDocuments/CCERT-PUBDOC-2008-12-249.pdf>
6. Digital Preservation Coalition. (2016). *Preservation issues*. Dostupno na: <http://www.dpconline.org/digital-preservation/preservation-issues>
7. Digital Preservation Coalition. (2016). *Why digital preservation matters*. Dostupno na: <http://www.dpconline.org/digital-preservation/why-digital-preservation-matters>
8. Discoverygarden Inc. (2016, May 30). *Islandora Webinar: A „How To“ Guide for Islandora*. Dostupno na: <https://www.youtube.com/watch?v=I06UOfHevtQ>
9. Drupal. *About*. Dostupno na: <https://www.drupal.org/about>
10. Fedora. *About Fedora*. Dostupno na: <http://www.fedora-commons.org/about>
11. Foulonneau, M., & Riley, J. (2008). *Metadata for Digital Resources*. Oxford: Chandos Publishing.
12. Gilliland, A. J. (2008). *Setting the Stage*. Introduction to Metadata. uredila Murtha Baca. Los Angeles: The Getty Research Institute.
13. Hasenay, D., & Krtalić, M. (2009). *Dostupnost i očuvanje informacija – što znači pouzdano i dugoročno?*. 12. seminar Arhivi, knjižnice, muzeji: mogućnost suradnje u

- okruženju globalne informacijske infrastrukture. uredila Sanjica Faletar Tanacković. Zagreb: Hrvatsko knjižničarsko društvo.
14. Haynes, D. (2004). *Metadata for Information Management and Retrieval*. London: Facet Publishing. Dostupno na: http://www.facetpublishing.co.uk/downloads/file/sample_chapters/HaynesMetadatach1.pdf
 15. Islandora. *About*. Dostupno na: <http://islandora.ca/about>
 16. Klarin, S. (2008.) *Tehnički metapodaci i metapodaci za zaštitu u normama ANSI/NISO Z39.87 i PREMIS*. 11. seminar Arhivi, knjižnice, muzeji: mogućnost suradnje u okruženju globalne informacijske infrastrukture. uredila Mirna Willer. Zagreb: Hrvatsko knjižničarsko društvo.
 17. Krajina, T. (2008). *Slobodan pristup informacijama: institucijski repozitoriji*. Polimeri: časopis za plastiku i gumu, Vol. 28 No. 3.
 18. Lasić-Lazić, J. (2005). *Repozitoriji digitalnog obrazovnog materijala kao sastavnica kvalitete suvremenih koncepta obrazovanja*. Dostupno na: <http://edupoint.carnet.hr/casopis/33/clanci/1.html>
 19. Lavoie, B. (2000). *Meeting the challenges of digital preservation: The OAIS reference model*. Dostupno na: <http://www.oclc.org/research/publications/library/2000/lavoie-oais.html>
 20. National Information Standards Organization. (2004). *Understanding Metadata*. Bethesda: NISO Press. Dostupno na: <http://www.niso.org/publications/press/UnderstandingMetadata.pdf>
 21. Solr. *Learn more about Solr*. Dostupno na: <http://lucene.apache.org/solr/>
 22. Stančić, H. (2006). *Arhivsko gradivo u elektroničkom obliku: mogućnost zaštite i očuvanja na dulji vremenski rok*. Arhivski vjesnik, Vol. 49 No. 1. Zagreb: Hrvatski državni arhiv.
 23. Stančić, H. (2005). *Teorijski model postojanog očuvanja autentičnosti elektroničkih informacijskih objekata*. Doktorska disertacija. Zagreb: Filozofski fakultet u Zagrebu. Odsjek za informacijske i komunikacijske znanost.
 24. Stančić, H. (2009). *Značaj planiranja postupaka dugoročnog očuvanja elektroničkoga gradiva: značajne norme i metodologije*. 12. seminar Arhivi, knjižnice, muzeji: mogućnost suradnje u okruženju globalne informacijske infrastrukture. uredila Sanjica Faletar Tanacković. Zagreb: Hrvatsko knjižničarsko društvo.

25. Suber, P. (2015). *Open Access Overview: Focusing on open access to peer-reviewed research articles and their preprints*. Dostupno na: <http://legacy.earlham.edu/~peters/fos/overview.htm>
26. The Library of Congress. (2006). *MARC Standards Frequently Asked Questions*. Dostupno na: <https://www.loc.gov/marc/faq.html>
27. The Library of Congress. (2016). *PREMIS: Preservation Metadata Maintenance Activity*. Dostupno na: <http://www.loc.gov/standards/premis/>
28. The Library of Congress. (2016). *PREMIS Preservation Metadata XML Schema Version 3.0*. Dostupno na: <http://www.loc.gov/standards/premis/premis.xsd>
29. The PREMIS Working Group. (2005). *Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group, Version 1.0*. Dostupno na: <http://www.loc.gov/standards/premis/v1/index.html>
30. Vrana, R. (2010). *Digitalni repozitoriji ustanova i njihova uloga u znanstvenom komuniciranju Istraživanje o digitalnim repozitorijima u visokoškolskim knjižnicama Sveučilišta u Zagrebu*. 13. seminar Arhivi, knjižnice, muzeji: mogućnost suradnje u okruženju globalne informacijske infrastrukture. uredile Mirna Willer i Sanjica Faletar Tanacković. Zagreb: Hrvatsko knjižničarsko društvo.
31. Zlodi, G. (2007). *Mogućnosti uspostavljanja interoperabilnosti među shemama metapodataka u muzejskom okruženju*. Doktorska disertacija. Zagreb: Filozofski fakultet u Zagrebu. Odsjek za informacijske i komunikacijske znanosti.

Prilozi

Prilog 1: Popis slika

<i>Slika 1: PREMIS kao jezgra metapodataka za očuvanje.</i>	<i>29</i>
<i>Slika 2: PREMIS podatkovni model.</i>	<i>32</i>
<i>Slika 3: Deklaracija elemenata u PREMIS-u.</i>	<i>40</i>
<i>Slika 4: Popratni dokumenti objekta pohranjenog u Islandori.</i>	<i>44</i>
<i>Slika 5: Promjene nakon uređivanja MODS zapisa.</i>	<i>45</i>