

iz naših knjižnica

Uređuje: Danko Škare

Zaštita i dugoročni pristup elektroničkom gradivu – projekt *Planets*

L. Škorić*

Središnja medicinska knjižnica
Medicinski fakultet Sveučilišta u Zagrebu
Šalata 3, 10 000 Zagreb

U svakodnevnom životu i radu svjedoci smo kako elektronički oblik dokumenata sve više istiskuje tiskane formate. Istodobno, neprekidno raste broj dostupnih informacija i mijenja se tehnološka osnovica za njihovo objavljivanje i pohranu. Jednako kao što se tiskane zbirke štite od štetnih utjecaja svjetla, vlage, temperature i različitih nepogoda, i elektroničke sadržaje potrebno je pravovremeno, prikladno i aktivno štiti. Pri tome, zaštita elektroničkog sadržaja treba biti trojaka – potrebno je očuvati fizičku razinu zapisa, odnosno razinu zapisa na neki medij, logičku razinu, to jest razinu zapisa objekta kao jedinice koju prepoznaje softver, ali i konceptualnu ili intelektualnu razinu – entitet kojega prepoznamo kao smislenu jedinicu informacije.¹ Ukratko, cilj zaštite je da korisnik u budućnosti razumije i koristi se sadržajem na jednako kvalitetan način kao i u vrijeme stvaranja sadržaja.

Hardver se razvija, softveri zastarijevaju, a nositelji podataka podložni su propadanju. Zbog toga je mogućnost gubitka velike količine digitalnog naslijeđa zaista realna, te smo danas suočeni s pravim izazovom – potrebno je sačuvati sve više informacija koje su sve sklonije propadanju, a sve na održiv način za budućnost.²

Projekt *Planets*

Europske nacionalne knjižnice i arhivi koji imaju zakonsku odgovornost čuvanja elektroničke građe i osiguranja dugotrajnog pristupa elektroničkom kulturnom i znanstvenom naslijeđu uočili su navedene probleme i potrebe te odlučili zajednički raditi na njihovu rješavanju. *Planets – Preservation and Long-term Access through Networked Services* projekt je koji se bavio, kako i sam naziv kaže, zaštitom i osiguranjem dugoročnog pristupa informacijama putem umreženih usluga.³ Projekt sufinanciran sredstvima Europske unije započeo je 1. lipnja 2006. godine, a završio 31. svibnja 2010. godine. Sudionici projekta okupljeni su u svojevrsan konzorcij sastavljen od evropskih nacionalnih knjižnica i arhiva, važnih znanstvenih sveučilišta i tehnoloških tvrtki, a koordinator projekta bila je knjižnica *British Library*. Glavni cilj projekta *Planets* bila je izgradnja alata i usluga kojima bi se omogućio dugoročni pristup elektroničkim kulturnim i znanstvenim sadržajima. Taj se cilj postiže unaprjeđenjem procesa donošenja odluka, kontrolom troškova jačanjem automatizacije i podesive infrastrukture, osiguranjem širokog prihvaćanja u korisničkoj zajednici te stvaranjem tržišta za servise i alate za zaštitu.⁴

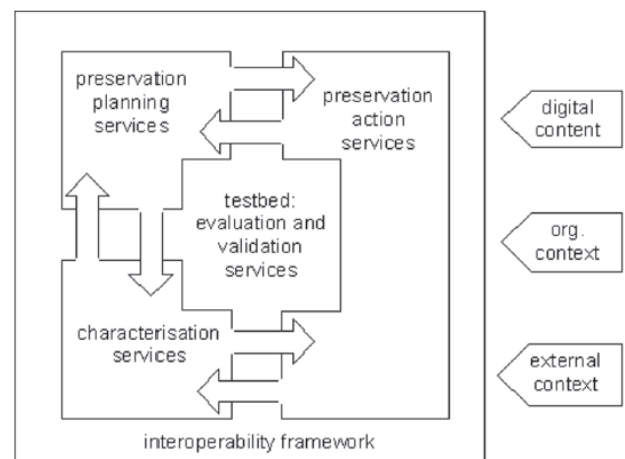
Uzimajući u obzir da je u Europi i svijetu posljednjih godina ostvaren velik napredak u istraživanju očuvanja digitalne građe te da već postoji određen broj rješenja i alata za očuvanje digitalnog gradiva, sudionici projekta nastojali su integrirati postojeća znanja,

rješenja i alate s novim spoznajama kako bi stvorili usluge i alate za primjenu u stvarnom životu i radu. Posebna pažnja posvećena je kompatibilnosti s referentnim modelom OAI. Osnovna je namjera projekta bila ponuditi integralna i sveobuhvatna rješenja uz pomoć kojih će biti moguće očuvati velike i kompleksne digitalne zbirke. Također, nastojalo se i kreirati ozračje koje će poticati stalni razvoj alata i usluga te okvir kojeg komercijalni proizvođači alata i usluga mogu primijeniti za kreiranje novih proizvoda.

Krajnji cilj projekta bila je izrada spektra alata i usluga u obliku softverskih paketa koje korisnici jednostavno mogu preuzeti i pokrenuti, a koji će omogućiti planiranje, administraciju, konfiguraciju, i sam postupak zaštite. Tako koncipirani alati i usluge trebali bi zadovoljiti zahtjeve velikih organizacija za sigurnom i mjerljivom implementacijom, potrebe manjih organizacija za jednostavnom primjenom, kao i zahtjeve prodavatelja softvera te ostalih neovisnih pružatelja usluga.⁵

Kako bi ispunila sve potencijalne zahtjeve korisnika, arhitektura sustava fleksibilno je ustrojena. Ključne sastavnice sustava, prikazane na slici 1, usluge su planiranja zaštite (*Preservation Planning Services*), usluge karakterizacije (*Characterisation Services*), usluge zaštite (*Preservation Action Services*), evaluacija i validacija usluga (*Testbed*) te interoperabilni okvir rada (*Interoperability Framework*).

Planets nije zamišljen kao repozitorij. Ustanova svoju zbirku treba čuvati u vlastitom arhivu ili repozitoriju. Zbog toga su softveri koje projekt razvija osmišljeni kako bi funkcionirali u različitim okruženjima, odnosno bili kompatibilni s različitim sustavima. Najčešće

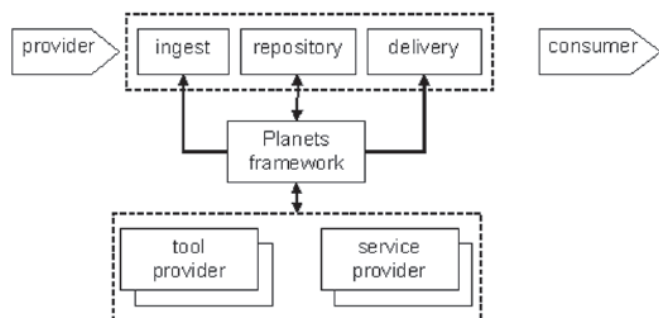


Slika 1 – Shematski prikaz arhitekture sustava, međuodnosa komponenti i vanjskih utjecaja⁶

* Lea Škorić, e-pošta: lea.skoric@mef.hr, tel.: 01-4566-937; faks: 01-4811-244

se to postiže primjenom takozvanih adaptora. Za svaki repozitorij ili arhiv definira se poseban adaptor, kako bi odgovarajući softver mogao ekstrahirati pohranjeni sadržaj i metapodatke, manipulirati njima te po potrebi pružiti novi sadržaj i nove metapodatke.⁶

Nakon što utvrde nemogućnost pristupa, institucije primjenjujući alate sustava *Planets* svojim korisnicima omogućuju pristup traženom sadržaju. To se najčešće postiže migracijom sadržaja, primjenom programskih dodataka ili emulacijom. Softver radi s arhivskim službama unosa, čuvanja i isporuke dokumenata, kao i s neovisnim pružateljima usluga i alata, tako da najčešće ne komunicira izravno niti sa stvarateljima niti s korisnicima sadržaja, kao što je prikazano na slici 2.



Slika 2 – Međusobni odnosi i djelovanja u sustavu *Planets*⁶

Projekt *Planets* nastoji zaštititi elektroničkog gradiva pristupiti sveobuhvatno, pa je rad na pojedinim aspektima organiziran kroz potprojeke. To su planiranje zaštite, karakterizacija, postupci zaštite, testiranje i evaluacija, interoperabilni okvir rada, širenje informacija i promocija rješenja.

Planiranje zaštite

Najvažniji proizvod potprojekta planiranja zaštite je *Plato* – *Planets preservation planning tool*, alat za planiranje zaštite.⁷ Za osiguranje dugoročnog pristupa i autentičnosti digitalnih sadržaja, pri planiranju zaštite potrebno je voditi računa o brojnim čimbenicima, kao što su politike zaštite, zakonske obveze, zahtjevi korisnika, karakteristike samog sadržaja, organizacijske i tehničke mogućnosti, troškovi i slično.

Plato vodi korisnike kroz složeni proces kako bi odredili koje potrebe i zahtjeve za zaštitom svaka organizacija ima, pomaže pri definiranju njihovih prioriteta tako da omogućuje usporedbu dostupnih strategija i mogućih postupaka, pomaže u izradi i procjeni plana zaštite te definiranju postupaka zaštite, automatski sastavlja plan i tijekom procesa zaštite, utvrđuje rezultate postupaka zaštite i uspoređuje svojstva obrađenih objekata s originalom te procjenjuje uspješnost postupaka zaštite i moguće kvarenje sadržaja.⁸

Od lipnja 2010. godine *Plato* je dostupan s mrežnih stranica *SourceForge*.⁹

Karakterizacija

Jedan od ciljeva potprojekta karakterizacije je razvoj sustava za kontrolu uspješnosti migracije. Ovaj postupak mora biti automatiziran, budući da korisnici nemaju vremena za ručnu kontrolu tisuća migriranih dokumenata. Pritom je presudno da sva, ili barem većina obilježja dokumenta ostanu nepromijenjena. Na primjer, ako se radi o slici, osim broja piksela prije i nakon konverzije, važno je i da podatak sadržan u pikselu bude uspješno konvertiran. Specifikacija svakog formata koristi se drugačijim rječnikom za opis osobina dokumenta, te pohranjuje taj opis u vlastitoj strukturi. Stoga je ovaj potprojekt razvio alate za opis formata na način koji omogućuje usporedbu informacija unutar dokumenata u

različitim formatima. Za ekstrakciju obilježja razvijen je softver *Extractor*, koji iz dokumenta ekstrahira bitna obilježja, a tako dobivenim podacima koristi se i *Comparator*, alat za usporedbu obilježja elektroničkih dokumenata prije i nakon provođenja postupaka zaštite.

Za potrebe rada *Extractora* i sličnih softvera razvijena su i dva nova formalna jezika bazirana na tehnologiji XML-a, nazvana *Extensible Characterisation Definition Language* (XCDL) i *Extensible Characterisation Extraction Language* (XCEL), koja opisuju formate i u njima sadržane informacije.¹⁰

Postupci zaštite

Cilj ovog potprojekta unaprjeđenje je mogućnosti zaštite digitalnog gradiva primjenom odgovarajućih učinkovitih i financijski isplativih alata. To se najčešće postiže migracijom i emulacijom. Primjenom postojećih tehnologija konverzije i migracije, kao što su *Dvips*, *PS2PDF*, *PDF2PDFa*, *BullZip*, *Ghostsript*, *GIMP*, *JMagik* i drugi, te razvojem novih rješenja, u projektu *Planets* u proces zaštite nastojali su uključiti što veći broj formata. Migracija sadržaja vrši se tako da softver razvijen kroz projekt *Planets* ekstrahira sadržaj iz repozitorija, karakterizira ga, bira odgovarajuću migracijsku uslugu, koristi se odabranim alatom za pripremu novog sadržaja, provjerava i ocjenjuje novonastali sadržaj te ga zatim, zajedno s pratećom informacijom o izvršenom procesu, unosi nazad u repozitorij. Institucijska služba korisniku pruža pristup ili isporučuje na taj način pripremljeni traženi sadržaj.

Emulacijom originalnog hardvera stvara se okruženje koje omogućava primjenu aplikacija i softvera stvorenih još osamdesetih godina prošloga stoljeća. U ovom potprojektu razvijena su četiri emulatora – *GRATE* – *Global Remote Access to Emulation Services*, *Dioscouri*, *QEMU* i *UVC* – *Universal Virtual Computer*.¹¹

Planets Testbed

Ovim potprojektom nastojalo se korisnicima i istraživačima omogućiti izvođenje pokusa u kontroliranom okruženju te prikupljanje podataka o zaštiti digitalnih sadržaja putem pravog znanstvenog pokusa. Ova mrežna aplikacija kombinira podatke o hardveru, softveru i ogleđne informacije, te pokazuje koji je pristup odnosno alat najbolji izbor za zaštitu različitih vrsta digitalnih sadržaja. Također, ova aplikacija pomaže u usporedbi različitih mogućnosti zaštite nudeći uvid u iskustva drugih korisnika i primjere najbolje prakse.¹² *Planets Testbed* osigurava metodologiju za vršenje sustavnih znanstvenih pokusa s mogućnošću reprodukcije pokusa i uključivanja više sudionika, dugoročno dostupnu strukturiranu dokumentaciju pokusa, automatiziranu evaluaciju rezultata te zajednički pristup pokusima uz robusnu i podesivu platformu. Radom u *Testbedu* korisnici na osnovi empirijskih dokaza sustavno analiziraju moguće postupke digitalne zaštite, u kontroliranom okruženju eksperimentiraju primjenjujući različite alate i usluge, mjere i uspoređuju rezultate s predefiniranim zahtjevima. Svojim radom ujedno doprinose i sustavu – vlastitim dokazima potvrđuju uspjeh primjene pojedinog alata za određene tipove objekata, čime potpomažu buduće planiranje zaštite i daju povratnu informaciju cjelokupnoj zajednici koja se bavi zaštitom.²

U ožujku 2010. godine beta verzija *Testbeda* postala je dostupna javnosti.¹³

Interoperabilni okvir

Takozvani interoperabilni okvir (*Planets Interoperability Framework*) čini softversku infrastrukturu cijelog projekta *Planets*. Poput ljepljive povezuje i drži na okupu sve aplikacije, alate i usluge. Unutar njegove infrastrukture inkorporirano je mnoštvo međunarodno priznatih standarda i slobodno dostupnih alata. Instalacijski paket interoperabilnog okvira sadrži prekonfiguriranu serversku aplikaciju koja omogućuje usluge poput *single-sign-on*, upravljanje ko-

risnicima, autorizaciju i autentifikaciju. To je okvir za mrežne aplikacije poput alata za planiranje, *Plato* i *Planets Testbed*. Alati za zaštitu pokreću se kao mrežne usluge smještene unutar distribuirane mreže.¹⁴

Zahvaljujući ovakvoj strukturi, korisnici se mogu koristiti cjelinom ili dijelovima ove distribuirane mreže usluga. Pretraživanjem registra usluga mogu birati željene usluge ali i razmjenjivati informacije o njima. Kroz obrasce radnih procesa jednostavne usluge, poput karakterizacije i migracije, mogu se objediniti u složene procese koji uključuju identifikaciju, validaciju, karakterizaciju i migraciju, pri čemu se mnogi složeni procesi odvijaju automatski. Također, interoperabilni okvir moguće je proširiti. Korisnici i stvaratelji usluga i alata svoje proizvode mogu testirati i putem standardiziranog sučelja integrirati u sustav *Planets*.

Interoperabilni okvir kompatibilan je s većinom postojećih sustava za pohranu dokumenata. Integracija alata i servisa projekta *Planets* u uobičajene arhivske i knjižnične sustave omogućena je putem posebnog adaptora. Adaptor se inkorporira u postojeći softver te korisnici preko njega iz vlastitog sučelja "pozivaju" željene usluge iz interoperabilnog okvira.

Širenje i prihvaćanje

Posljednji, ali svakako ne i najmanje važan potprojekt projekta *Planets* bavi se diseminacijom stečenih iskustava, znanja i rješenja te osiguravanjem njihovog prihvaćanja i primjene u praksi. To se ostvaruje u suradnji s ostalim stručnjacima zainteresiranim za digitalnu zaštitu i tvrtkama koje se bave razvojem aplikacija za zaštitu te organizacijom edukacije za rad s alatima razvijenim u projektu *Planets*.

Umjesto zaključka – pogled u budućnost

Projekt *Planets* završio je 2010. godine, no rezultati projekta i dalje se održavaju i unaprjeđuju radom organizacije nazvane *Open Planets Foundation*. To je neprofitna tvrtka registrirana u Velikoj Britaniji s ciljem razvoja praktičnih rješenja za zaštitu elektroničke građe. Smatraju kako je za ostvarivanje zaštite digitalnih sadržaja ključna suradnja i aktivno dijeljenje stečenih znanja, pa promiču izradu alata za zaštitu s licencijom za otvoreni pristup.¹⁵

Također, 18. ožujka 2010. godine, takozvana vremenska kapsula projekta *Planets* pohranjena je u švicarskom centru Swiss Fort Knox. U vremenskoj kapsuli pohranjeni su sadržaji za koje se procjenjuje da će s vremenom postati ugroženi – fotografija u formatu JPEG, poruka u Java-kodu, kratki film u formatu MOV, mrežna stranica u HTML-u i brošura u PDF-u. Svaki je objekt pohranjen u originalnom formatu i u novom za zaštitu pogodnijem formatu. Primijenjeni su i različiti mediji pohrane – papir, mikrofilm, disketa, CD, DVD, *flash*-disk, i HDD. Pohranjeni su i čitači navedenih medija, primijenjeni alati za konverziju, softveri potrebni za korištenje sadržaja, operacijski sustavi, opisi formata i načina kodiranja te opisi svih objekata i njihovih međusobnih odnosa te odnosa prema tehnologiji i standardima.¹⁶ Budući da je vremenska kapsula dostupna javnosti, znanstvenici će kroz deset, dvadeset ili više godina na najbolji način moći testirati autentičnost i upotrebljivost pohranjenog digitalnog sadržaja, a time i steći uvid u kvalitetu osmišljenih rješenja.

Treba se nadati da će zajednički trud međunarodne zajednice stručnjaka u očuvanju rastućeg korpusa elektroničkih informaci-

jskih objekata i primjena naprednih tehnologija pohrane i zaštite uroditi plodom, te će današnja digitalna baština zbog ovog i sličnih projekata ipak ostati dostupna budućim generacijama.

Literatura

1. K. Thibodeau, Overview of technological approaches to digital preservation and challenges in coming years, u: The state of digital preservation: an international perspective. Conference proceedings. (2005.) URL: http://chnm.gmu.edu/digitalhistory/links/pdf/preserving8_37e.pdf (25. 9. 2011.).
2. S. Bonin, Preservation and Long-term Access via NETWORKED Services: Keeping digital information alive for the future. Planets Consortium, 2009. URL: http://www.planets-project.eu/docs/comms/PLANETS_BROCHURE.pdf (11. 8. 2011.).
3. Planets – Preservation and Long-term Access through NETWORKED Services. URL: <http://www.planets-project.eu> (10. 8. 2011.).
4. C. Becker, Preservation and Long-term Access through NETWORKED Services. 2008. URL: http://www.planets-project.eu/docs/presentations/CeBIT_2008.pdf (11. 8. 2011.).
5. A. Farquhar, H. Hockx-Yu, Planets: Integrated Services for digital Preservation, The International Journal of Digital Curation. 2 (2) (2007) 88–99. URL: <http://www.ijdc.net/index.php/ijdc/article/view/45/31> (11. 8. 2011.).
6. A. Farquhar, H. Hockx-Yu. Planets: Integrated Services for digital Preservation. // Serials. 21 (2) (2008) 140–144. URL: <http://uksg.metapress.com/content/n8h8772626232h04/fulltext.pdf> (11. 8. 2011.).
7. Plato: The Preservation Planning Tool. URL: <http://www.ifs.tuwien.ac.at/dp/plato> (16. 8. 2011.).
8. C. Becker et al., Planets Planning Tool (Version 3). 2010. URL: <http://www.ifs.tuwien.ac.at/dp/plato/docs/PP4-D5.pdf> (25. 9. 2011.).
9. SourceForge.net Repository – [planets-suite] Index of /trunk/webapps/plato. URL: <http://planets-suite.svn.sourceforge.net/viewvc/planets-suite/trunk/webapps/plato> (15. 10. 2011.).
10. About XCL – eXtensible Characterization Language. URL: http://planetarium.hki.uni-koeln.de/planets_cms/about-xclxcdl (18. 10. 2011.).
11. Planets: Tools and services for digital preservation. URL: http://www.planets-project.eu/docs/comms/PLANETS_PRODUCT_SPECIFICATION.pdf (18. 10. 2011.).
12. B. Aitken, A. Lindley, The Planets Testbed: A Collaborative Research Environment for Digital Preservation. // ERCIM News. 80 (2010) 33–35. URL: <http://ercim-news.ercim.eu/images/stories/EN80/EN80-web.pdf> (20. 8. 2011.).
13. The Testbed v1.1 (Public Beta). URL: <http://testbed.planets-project.eu/testbed> (20. 8. 2011.).
14. R. King, The Planets Interoperability Framework. // ERCIM News. 80 (2010) 14–15. URL: <http://ercim-news.ercim.eu/images/stories/EN80/EN80-web.pdf> (13. 8. 2011.).
15. Open Planets Foundation. URL: www.openplanetsfoundation.org (10. 8. 2011.).
16. Time Capsule – A Showcase for Digital Preservation. URL: <http://www.ifs.tuwien.ac.at/dp/timecapsule/timecapsule.html> (28. 10. 2011.).