

Analiza odabranih sustava za digitalno arhiviranje

Kozarić, Korina

Undergraduate thesis / Završni rad

2020

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, University of Zagreb, Faculty of Humanities and Social Sciences / Sveučilište u Zagrebu, Filozofski fakultet**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:131:385589>

Rights / Prava: [In copyright](#)/[Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2024-07-28**



Sveučilište u Zagrebu
Filozofski fakultet
University of Zagreb
Faculty of Humanities
and Social Sciences

Repository / Repozitorij:

[ODRAZ - open repository of the University of Zagreb
Faculty of Humanities and Social Sciences](#)



SVEUČILIŠTE U ZAGREBU
FILOZOFSKI FAKULTET
ODSJEK ZA INFORMACIJSKE I KOMUNIKACIJSKE ZNANOSTI
Ak. god. 2019./2020.

Korina Kozarić

Analiza odabranih sustava za digitalno arhiviranje

Završni rad

Mentor: prof. dr. sc. Hrvoje Stančić

Zagreb, lipanj 2020.

Izjava o akademskoj čestitosti

Izjavljujem i svojim potpisom potvrđujem da je ovaj rad rezultat mog vlastitog rada koji se temelji na istraživanjima te objavljenoj i citiranoj literaturi. Izjavljujem da nijedan dio rada nije napisan na nedozvoljen način, odnosno da je prepisan iz necitiranog rada, te da nijedan dio rada ne krši bilo čija autorska prava. Također izjavljujem da nijedan dio rada nije korišten za bilo koji drugi rad u bilo kojoj drugoj visokoškolskoj, znanstvenoj ili obrazovnoj ustanovi.

Sadržaj

1. Uvod.....	2
2. Digitalno arhiviranje i očuvanje.....	3
2.1. Prijetnje digitalnim materijalima	4
2.2. Problem organizacije	5
2.3. Skladištenje.....	7
2.4. Alati za digitalno očuvanje	8
3. „Archivematica“	10
3.1. Tehničke specifikacije	10
3.2. Mikroservisi.....	11
3.3. Sučelje	12
3.4. Očuvanje.....	13
3.5. Klijenti	14
4. „Preservica“	16
4.1. Aplikacijska programska sučelja „Preservice“	17
4.2. Pristup sadržaju.....	18
5. Usporedba sustava „Archivematica“ i „Preservica“	20
6. Zaključak.....	24
7. Literatura.....	25
Popis slika	27
Popis tablica	27
Sažetak	28
Summary	28

1. Uvod

S pojavom pisanja javilo se i arhiviranje. Time se omogućilo očuvanje raznih ljudskih iskustava, otkrića i izuma. Bez arhiviranja i očuvanja ljudskih postignuća kroz povijest ne bi bio moguć napredak filozofije, umjetnosti, znanosti i religije jer se njihov napredak temelji na dijeljenju stečenog znanja sačuvanog na glinenim pločicama, pergamentima i knjigama. Kroz povijest se akumulirao sve veći broj dokumenata, informacija i znanja te izumom tiskarskog stroja glavni problem više nije bilo njihovo dokumentiranje. Novi izazov je kako njima upravljati, organizirati i danas kako ih digitalizirati. Kako se pojavljuje sve veći broj podataka nastalih u digitalnom okruženju pojavljuju se i brojni alati i sustavi za njihovo očuvanje i upravljanje što svojim konstantnim mijenjanjem i unaprjeđivanjem može predstavljati nove izazove u očuvanju digitalnih materijala. Treba se obratiti pažnja i na krhkost digitalnih materijala koji se lako mogu slučajno ili namjerno izbrisati ili može doći do fizičkog oštećenja medija na kojem su pohranjeni.

Nad digitalnim materijalima se provode određeni postupci poput identificiranja formata datoteka, provjere virusa, generiranje metapodataka itd. Svi ti postupci i promjene nad digitalnim materijalima su bitni te se moraju dokumentirati. Kako bi očuvanje bilo kvalitetno u njega se mora uračunati i planiranje kako bi se uvidjeli svi mogući rizici kao što je konstantni ubrzani napredak tehnologije, koje potrebe ima neka organizacija ili institucija te koje metode i tehnologije koristiti.

2. Digitalno arhiviranje i očuvanje

Digitalno arhiviranje je dugoročno očuvanje digitalnih sadržaja koji uključuju one nastale u digitalnom obliku te također i one nastale procesom digitalizacije. Konstantno mijenjajuća sveprisutna suvremena digitalna tehnologija omogućava sve veću suradnju arhivista s informatičarima kako bi se omogućilo bolje očuvanje i dostupnost elektroničkih zapisa. Arhivisti prikupljaju zapise ljudi i organizacija kako bi sačuvali izvore koji svjedoče najvećim ljudskim dostignućima i koji kasnije mogu poslužiti za istraživanje. Poduzeća također sadržavaju arhive u kojima čuvaju dokumente kao što su ugovori, računi, narudžbe, izvješća, zahtjevi, rješenja, izvodi, naredbe i slično.¹ Ti materijali se javljaju u svim vrstama formata te zbog toga digitalno očuvanje može obuhvatiti tekstove i slike, baze podataka i proračunske tablice, programe i aplikacije, e-poštu i društvene medije, igre, filmove, audio zapise, čitave web domene i pojedinačne objave. Digitalni materijali mogu dolaziti s laptopa, računala ili pametnih telefona i tableta te kao takvi mogu biti pohranjeni u repozitorijima, podatkovnim centrima, lokalnim računalima ili prijenosnim medijima. To znači da ne postoji digitalni objekt ili sustav koji nije povezan s pojmom digitalnog očuvanja. Digitalni materijali služe kao osnova za industriju, trgovinu, rad tijela državne i javne uprave, kulturnu djelatnost, a temelj su za istraživanja.

Glavna karakteristika digitalnih materijala je njihova ovisnost o informacijsko-komunikacijskoj tehnologiji (IKT). Informacijama se može pristupiti samo putem nekog uređaja koji koristi IKT. Kako tehnologija sve više napreduje ta ovisnost postaje sve veća i stvara sve složeniji lanac ovisnosti koje je teško pratiti i održavati. Sve dok IT sektor stvara nove alate i tehnologije, stručnjaci za digitalno očuvanje moraju smišljati nove učinkovite strategije za osiguranje trajnosti i upotrebljivosti digitalnih materijala te će tako digitalno očuvanje uvijek biti izazovno.

Digitalni materijali izvorno nastali u digitalnom obliku se često u takvom obliku jedino i mogu arhivirati jer ne postoji ne-digitalni ekvivalent poput papira koji bi zadržao sve informacije i funkcionalnost koju pruža digitalni oblik. Prečesto se javlja problem ispisivanja takvog gradiva kako bi se arhiviralo te kasnije ponovne digitalizacije ispisane kopije zbog nedostatka prostora u spremištu.

¹ Pećarina, G., (2014). Digitalno arhiviranje. InfoTrend. Dostupno na: <http://www.infotrend.hr/clanak/2014/10/digitalno-arhiviranje.81.1094.html> (22. lipnja 2020.)

Iako se može činiti da je danas moguće očuvati sav sadržaj to ne znači da bi se sve trebalo sačuvati. Glavno pitanje ne bi trebalo biti što sačuvati već što se ne smije izgubiti. Odabir, procjena, izlučivanje i arhiviranje su glavne aktivnosti vezane uz digitalno očuvanje.

2.1. Prijetnje digitalnim materijalima

Postoji niz prijetnji koje mogu utjecati na očuvanje digitalnih materijala. Mediji za pohranu s vremenom mogu propasti što dovodi do oštećenja datoteka ili pak mogu postati zastarjeli i ne podržavani od strane modernih računala i softvera. Uvijek se može dogoditi da se digitalni materijali slučajno izbrišu ili čak zlonamjerno unište. Prijenosni mediji na kojima su materijali pohranjeni mogu se ukloniti ili ostaviti na polici i zaboraviti, datoteke pohranjene na zajedničkim mrežnim uređajima mogu ostati zaboravljene, a neovisni proizvođač koji nudi uslugu pohranjivanja u računalnom oblaku (engl. *cloud*) može prestati poslovati.

Održavati sistematski proces očuvanja digitalnog gradiva je najvažnija potreba kojoj se treba udovoljiti kako bi se osigurala dugoročna digitalna pohrana što znači da se mediji za skladištenje moraju redovito nadgledavati i obnavljati. Potrebno je replicirati ili stvarati sigurnosne kopije datoteka kako bi se izbjegle katastrofalne posljedice uslijed nepogoda koje mogu zadesiti lokaciju na kojoj su datoteke čuvane. Mjesta gdje su digitalni materijali pohranjeni trebaju biti evidentirana te se treba dodijeliti odgovornost za njihovo čuvanje.

S vremenom se formati datoteka mogu promijeniti te softver koji se upotrebljava za njihovo korištenje može postati zastario. Iako je to dosta neuobičajeno za više korištene formate datoteka, oni manje korišteni mogu s vremenom zastarjeti jer softver koji ih stvara više nije podržan. Zato je potrebno shvaćanje tehnologije o kojoj ovise digitalni materijali što onda omogućava poduzimanje odgovarajućih mjera za njihovo očuvanje. Proces planiranja očuvanja često podrazumijeva konverziju digitalnih datoteka iz formata u format, emulacijom zastarjelog softvera ili korištenjem alternativnih softverskih aplikacija za prikazivanje podataka.

Digitalni materijali mogu biti lako promjenjivi što znači da mogu biti uređivani s lakoćom, oštećeni zbog kvara medija za pohranu ili pogrešno dešifrirani od strane nekog softvera. Kako bi krajnji korisnik imao pouzdanja u rezultate digitalnog očuvanja, potrebno je pažljivo razmotriti cjelokupni životni ciklus digitalnih materijala te tko je bio u interakciji s njima. Sustavi upravljanja informacijama moraju biti povezivi s kontekstom informacija koje čuvaju. Autentičnost i integritet digitalnih izvora mogu biti podjednako važni u ostalim sektorima. Na primjer, znanstvenici moraju biti sigurni da će reference koje citiraju s vremenom ostati iste.

Autentičnost digitalnih izvora ovisi i o pouzdanosti organizacije koja se bavi očuvanjem digitalnih materijala. Od presudne je važnosti održavanje procesa očuvanja na visokoj razini koja mora biti utemeljena na najnovijim vještinama.

Značenje digitalnih materijala može ovisiti o dodatnim informacijama koje su se možda podrazumijevale u kontekstu u kojem su prvobitno stvorene ili korištene, ali manje jasne kad su kasnije iznova korištene. Identifikacija, razumijevanje i bilježenje relevantnih kontekstualnih informacija može biti od presudne važnosti za uspješno očuvanje. Razumijevanje podataka, način na koji će se koristiti, njegove potrebe i kontekst omogućit će im njihovo očuvanje na odgovarajući način i razumijevanje njihovog sadržaja u budućnosti.

Davanjem prednosti aktivnostima vezanim uz digitalno očuvanje i njihova pravovremena primjena mogu biti presudni u izbjegavanju gubitka digitalnih materijala. Izbor formata datoteke, prikupljanje kritične dokumentacije ili opis ključnih odnosa u metapodacima mogu zahtijevati malo ulaganja unaprijed, ali i ostvariti znatnu uštedu kasnije. Rana intervencija radi sprječavanja tehnološke zastarjelosti može pružiti veću sigurnost u dugoročnu održivost, ali uz rizik da intervencija možda u konačnici i nije potrebna, a resursi su već potrošeni. Pravovremeno djelovanje može umanjiti nepotrebne aktivnosti te smanjiti napor potreban za istraživanje zastarjele tehnologije koje u određenim slučajevima mogu zahtijevati specijalističko znanje koje više nije aktualno.

Jednostavni procesi očuvanja koji učinkovito funkcioniraju na jednoj razini neće nužno raditi s vrlo velikim količinama podataka ili možda vrlo velikim pojedinačnim datotekama. Neki se repozitoriji još uvijek suočavaju s velikim izazovima u razvoju i održavanju postupaka za obradu sve veće količine podataka. Važno je zapamtiti da su odabir, procjena i izlučivanje značajne komponente u bilo kojoj aktivnosti vezanoj uz upravljanje digitalnim arhivskim gradivom.

2.2. Problem organizacije

Postoje zajednički izazovi za digitalno očuvanje s kojima se suočavaju sve organizacije, no i oni koji su drugačiji i ovisni o vrsti organizacije. Od vitalnog je značaja utvrditi pokretače organizacije i prilagoditi praktična rješenja za ispunjavanje tih potreba. Ne postoji jedan način digitalne pohrane koji bi odgovarao svima.

Odluka o tome hoće li digitalno očuvanje provoditi putem vanjskog pružatelja usluge ili će se ono odvijati unutar organizacije ili će to biti kombinacija toga dvoje je često složena. Digitalno očuvanje unutar organizacije se može organizirati ako postoji dovoljno osoblja i infrastrukture, no korištenje vanjskih usluga za neke aktivnosti i podršku može biti isplativo te može pozitivno utjecati na unutarnje mogućnosti i kapacitete. Posao obavljen od strane vanjskih pružatelja usluge mora biti provjeren, a to je omogućeno dizajniranjem specifikacija sa zahtjevima što je potrebno napraviti i naknadnim davanjem izvještaja o tome što je učinjeno.

Suvremeni digitalni svijet mjesto je brzih tehnoloških i organizacijskih promjena. Organizacije se sve više interno reorganiziraju, spajaju ili prestaju raditi. Digitalno očuvanje je dugoročna aktivnost i vjerojatnost da će promjene organizacije na nju utjecati s vremenom se povećava. To može utjecati na digitalno očuvanje ne samo promjenama u matičnoj organizaciji, već i promjenama glavnih ulagača i korisnika, dobavljača ili suradnika. Organizacijske promjene su stoga glavni rizik o kojem se treba brinuti.

Ogromna količina podataka koja se proizvodi digitalno, njihova upitna kvaliteta i ograničena sredstva onih koji imaju odgovornost očuvanja, selektivnost se čini neizbježnom ako je cilj očuvanje stalnog pristupa. U digitalnom okruženju ignoriranje važnosti očuvanja može gotovo sigurno značiti gubitak gradiva. U slučajevima kada može biti više verzija nekog gradiva, moraju se donijeti odluke o odabiru verzije koja je najbolja i najisplativija za očuvanje ili odabire li se više od jedne. Mora bit jasno tko će preuzeti odgovornost nad gradivom i za koje vremensko razdoblje. U obzir se mora uzeti da čak i ako je pohranjeno nekoliko kopija u raznim repozitorijima, svi ti repozitoriji mogu iz različitih razloga prestati održavati digitalni objekt u nekom trenutku.²

Mnoge vrste digitalnog materijala odabrane za dugoročno očuvanje mogu sadržavati povjerljive i osjetljive informacije koje moraju biti zaštićene kako bi se osiguralo da njima ne pristupaju neautorizirani korisnici. Također postoje i zakonske ili regulatorne obveze kojeg se repozitorij mora pridržavati koje utječu na pristup gradivu.³

² Digital Preservation Coalition, (2015). Digital Preservation Handbook: Preservation issues. Dostupno na: <https://www.dpconline.org/handbook/digital-preservation/preservation-issues> (31. kolovoza 2020.)

³ Ibid.

2.3. Skladištenje

„Digital Preservation Handbook“ navodi da učinkovito digitalno očuvanje zahtijeva neko osnovno postrojenje ili infrastrukturu u kojima se može vršiti obrada digitalnog materijala. Kada organizacija poduzme prve korake prema digitalnom očuvanju zahtijevat će se sve veća ulaganja u infrastrukture potrebne za podršku.

Osim kopiranja sačuvanih sadržaja kako bi se izbjegao gubitak, hardverski uređaji za pohranu ostaje među najvažnijim digitalnim uređajima za očuvanje. Tehnologija za skladištenje naglo se promijenila tijekom posljednjih desetljeća. Arhivi su se često koristili medijima kao što su CD-i ili DVD-i za dugoročno pohranjivanje, ali sve veći razvoj magnetskih medija donio je pouzdaniju pohranu. Iako imaju ograničeni vijek trajanja, obično oko 4-8 godina, lako ih je nadzirati i zamjenjivati kada im se bliži istek roka trajanja. Organizacije također mogu razmotriti cloud usluge za iznajmljivanje mjesta za pohranu. Fleksibilnost clouda omogućava relativno brzo i jeftino korištenje i navigiranje. Usluge u cloudu mogu pružiti lako, automatizirano kopiranje na više lokacija i pristup profesionalno upravljanoj digitalnoj pohrani i provjeri integriteta.⁴

Osim cloud servisa očuvanje digitalnih materijala se postiže i pomoću digitalnih sustava za očuvanje (kao što su i u radu kasnije predstavljeni sustavi „Archivemata“ i „Preservica“ koji također pruža cloud servise) ili pouzdanim digitalnim repozitorijima. Aplikacija za očuvanje će identificirati svaki digitalni objekt smješten u njoj te se pod time smatra upravljanje pohranom tog objekta, identifikacija njegovih karakteristika i pomaganje upravitelju skladišta da planira njegovo očuvanje. To će također olakšati pristup objektu.⁵

Povećavanje količine podataka zahtijeva ne samo više prostora za pohranu, već i jaču snagu računala. Procjena tehničkih karakteristika podataka, indeksiranje podataka radi pretraživanja i pristupa, provjera integriteta i niz drugih zadataka zahtijevaju bolje performanse računala.⁶

Nekoliko većih organizacija razvilo je laboratorijsko okruženje u kojem se može primijeniti niz starih i novih tehnologija za stabilizaciju ili uklanjanje podataka sa zastarjelih medija. Specijalni driveri za čitanje magnetskih medija, roboti za obradu velikog broja optičkih diskova i blokatori

⁴ Ibid.

⁵ Ibid.

⁶ Ibid.

pisanja koji omogućuju prikupljanje podataka bez njihovog mijenjanja, samo su neki od dijelova koji koriste u takvim laboratorijskim okruženjima.⁷

2.4. Alati za digitalno očuvanje

Korisnost alata za digitalno očuvanje ovisi o kontekstu njihovog korištenja. Ako se ne uskladi s potrebnom funkcijom ili vrstom rada kojim se organizacija bavi, tada postoji značajna šansa da alat neće uspjeti kvalitetno izvršiti zadane zadatke. Stoga je prije odabira digitalnih alata za očuvanje važno pažljivo razmotriti tijek rada i čime se institucija bavi.⁸

Neke organizacije, često one koje pripadaju visokim obrazovnim ustanovama i posebno repozitoriji ustanova, se više koriste softverima otvorenog koda posebno ako imaju internu grupu programera. Softver s otvorenim kodom omogućava slobodniju dopunu ili unaprjeđenje koda. Ostale organizacije koje nemaju pristup programerima često preferiraju komercijalna rješenja. Neki digitalni alati za očuvanje napravljeni su tako da omoguće proces od početka do kraja koji omogućava isporučivanje svih (ili većine) funkcija za digitalno očuvanje cijeloj organizaciji. Rješenja na razini poduzeća najčešće su izrađena spajanjem pojedinih alata integriranih u jedno sučelje.⁹

Ključno razmatranje prilikom odabira alata je kako se uklapaju u cjelokupnom tijeku rada, pa prije odabira alata dobro je isplanirati cijeli tijek rada. Biti jasan oko toga što se podrazumijeva pod tijek rada pomaže i u prepoznavanju suvišnih procesa. Jedan od čestih izazova je da alati rješavaju problem u dijelu tijeka rada, samo da bi stvorili problem u drugom. Opisivanje radnog tijeka stoga pruža osnovu za predviđanje budućih teškoća. Kako bi se procijenila korisnost i vrijednost mnogih dostupnih alata, puno pomaže unaprijed utvrditi izričite zahtjeve. Traženi zahtjevi se mogu promijeniti tijekom vremena te se zato digitalno očuvanje konstantno suočava s izazovima povezanim s tehnološkim promjenama. Dakle, potrebno je pratiti alate kako bi se osiguralo da ostanu prikladni za njihovu svrhu.¹⁰

Kako se zajednica za digitalno očuvanje povećava time se sve više javlja raznih alata i softvera koji poboljšavaju i omogućuju bolje digitalno očuvanje. Također sve veći odabir i raznolikost takvih alata može djelovati zbunjujuće novim korisnicima dok je put do tržišta sve teži za

⁷ Ibid.

⁸ Digital Preservation Coalition, (2015). Digital Preservation Handbook: Tools. Dostupno na: <https://www.dpconline.org/handbook/technical-solutions-and-tools/tools> (31. kolovoza 2020.)

⁹ Ibid.

¹⁰ Ibid.

programere. Sve češće se javljaju registri s alatima kao način da pomognu korisnicima da pronađu potrebne i korisne alate. Oni mogu dati detaljne opise, osvrte ili komentare na alate od ostatka zajednice. Osim što su korisni za korisnike, omogućavaju i stručnjacima pregled alata, procjenu njihove učinkovitosti te dobivanje povratne informacije o tome kako poboljšati svoje alate. Registri su uobičajeni način zajednice digitalnog očuvanja da razmjenjuje informacije.¹¹ Primjer jednog takvog registra je „COPTR (Community Owned digital Preservation Tool Registry) registry“. COPTR prikuplja alate korisne za dugoročno digitalno čuvanje. On prvenstveno djeluje kao alat za pronalaženje i procjenu kako bi pomogao stručnjacima otkriti alate potrebne za obavljanje određenih zadataka očuvanja. Prikuplja znanje zajednice digitalnog očuvanja o alatima na jednom mjestu. U COPTR-u je opisano 498 različitih alata.¹²

¹¹ Digital Preservation Coalition, (2015). Digital Preservation Handbook. Dostupno na: <https://www.dpconline.org/handbook> (24. lipnja 2020.)

¹² COPTR contributors. COPTR. Dostupno na: https://coptr.digipres.org/Main_Page (24. lipnja 2020.)

3. „Archivemática“

„Archivemática“ je besplatan digitalni sustav za očuvanje na bazi otvorenog koda (engl. *open-source*) koji je osmišljen kako bi se omogućio dugoročni pristup digitalnim zbirkama. On omogućuje obradu digitalnih zbirki u skladu s ISO-OAIS („Open Archival Information System“) modelom. Taj model je definiran kao arhiv koji se sastoji od organizacije ljudi i sustava koji su prihvatili odgovornost čuvanja podataka te ih dali na raspolaganje određenoj zajednici kojoj ti podaci nešto znače i koriste.¹³ „Archivemática“ je sustav za digitalno očuvanje koji automatizira proces pripreme digitalnih objekata za pohranu u repozitorij i pristupni sustav i pruža pristup arhiviranom materijalu. Može se pohraniti bilo kakav sadržaj uključujući slike, tekstualne datoteke, „office“ dokumente, publikacije, audiovizualne datoteke, baze podataka i skupovi podataka o istraživanjima. „Archivemática“ vrši radnje kao što su dodjeljivanje identifikatora, stvaranja kontrolne sume, skeniranja virusa, prepoznavanja formata i ekstrakcija metapodataka na svim pohranjenim objektima. Za određene vrste datoteka koriste se specijalizirani alati (poput „ImageMagick“ za rasterske slike, „ffprobe“ i „ffmpeg“ za AV datoteke, „Ghostscript“ za PDF datoteke, „Inkscape“ za vektorske slike, „bulk_extractor“ i „sleuthkit“ za slike forenzičkog diska) za ekstrakciju metapodataka, normalizaciju formata i druge zadatke potrebne za očuvanje. Novi alati, pravila i naredbe za očuvanje mogu se dodati koristeći Registar pravila formata (engl. *The Format Policy Registry ili FPR*) putem kartice „Planiranje očuvanja“ (engl. *Preservation Planning*) u web sučelju.

3.1. Tehničke specifikacije

„Archivemática“ koristi METS, PREMIS, Dublin Core, Baglt Kongresne knjižnice i ostale priznate standarde kako bi stvorila arhivski informacijski paket (engl. *Archival Information Package ili AIP*). Takav jedan AIP sadrži informacije o sadržaju (engl. *Content Information ili CI*), informacije o očuvanju (engl. *Preservation Description Information ili PDI*), informacije o pakiranju (engl. *Packaging Information ili PI*) i opisne informacije (engl. *Descriptive Information ili DI*)¹⁴. „Archivemática“ implementira mikroservisni pristup digitalnom očuvanju. Ti mikroservisi su detaljni sistemski zadaci koji obrađuju OAIS informacijske pakete, a to su: podneseni informacijski paket (engl. *Submission Information Package ili SIP*),

¹³ CCSDS Secretariat. (2012). Reference Model For An Open Archival Information System (OAIS). Dostupno na: <https://public.ccsds.org/Pubs/650x0m2.pdf> (22. lipnja 2020.)

¹⁴ McMaster University. Definitions AIP, DIP, SIP. Dostupno na: <https://digitalarchive.mcmaster.ca/node/56> (22. lipnja 2020.)

već prije spomenut arhivski informacijski paket (AIP) i informacijski paket za diseminaciju (engl. *Dissemination Information Package ili DIP*). Fizička struktura informacijskog paketa uključuje datoteke, kontrolne sume, zapise, dokumentaciju za podnošenje, XML metapodatke itd.

3.2. Mikroservisi

Kao što je već prije spomenuto, informacijski paketi su obrađeni pomoću raznih mikroservisa. Mikroservisi su stvoreni pomoću kombinacija kodova napisanih u programskom jeziku „Python“ ili od vanjskih alata (tablica 1) koji su napravljeni na temelju otvorenog koda te su ujedno ukomponirani u sustav „Archivematica“. Svaki izvršeni mikroservis može rezultirati uspjehom ili neuspjehom te će na temelju toga biti obrađen informacijski paket.

Tablica 1. Mikroservisi¹⁵

Alati	Opis alata
BagIt	Služi za standardno pakiranje digitalnih objekata i metapodataka za arhivsku pohranu.
bulk_extractor	Alat za analizu slike diska i sadržaja datoteka.
ClamAV	Antivirusni alat za UNIX.
Django	Django okvir (engl. <i>Framework</i>) potiče brzi razvoj.
ElasticSearch	Alat koji služi za pretragu.
ExifTool	Koristi se za čitanje, pisanje i uređivanje metapodataka u velikom broju datoteka.
FFmpeg	Alat za snimanje, pretvaranje i puštanje audio i video zapisa.
Fido (Format Identification for Digital Objects)	Služi za prepoznavanje formata datoteka digitalnih objekata.
File Information Tool Set (FITS)	Također se koristi za prepoznavanje formata datoteka, ali i kao alat za integraciju softvera za provjeru valjanosti.
Gearman	Gearman je alat koji prenosi zadatke drugim strojevima ili procesima koji su bolje namijenjeni za obavljanje tog zadatka.
Ghostscript	Tumači „PostScript“ jezik i PDF.

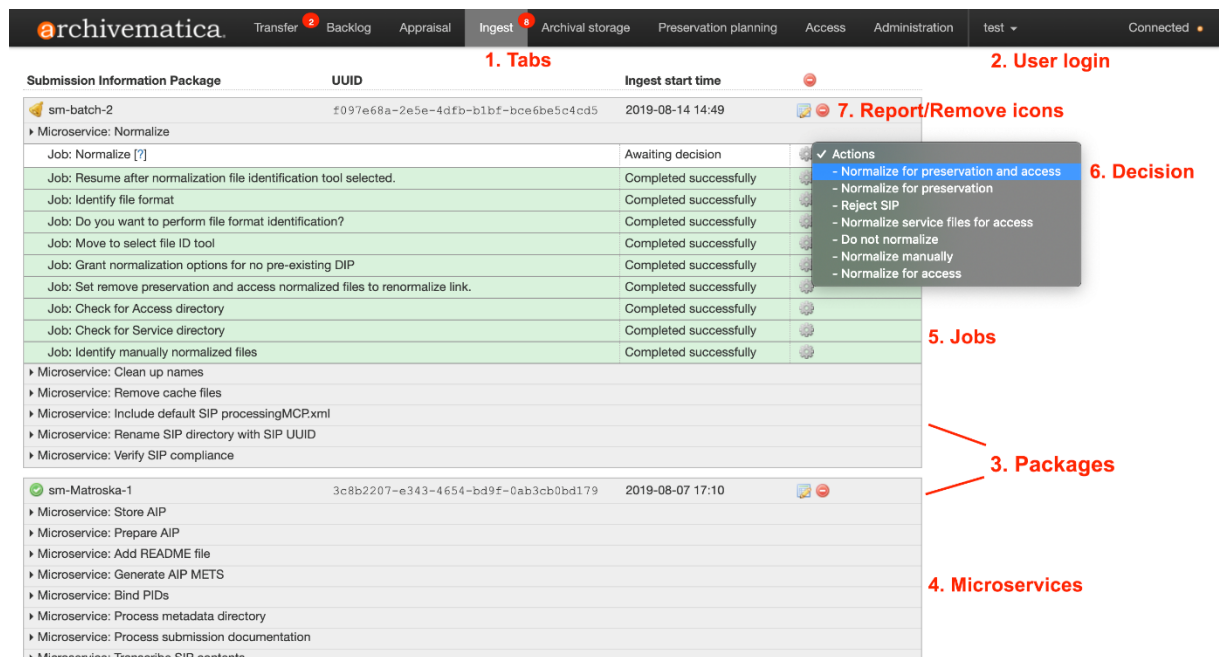
¹⁵ Archivematica: open-source digital preservation system. External tools. Dostupno na: <https://www.archivematica.org/en/docs/archivematica-1.11/getting-started/overview/external-tools/#external-tools> (29. kolovoza 2020.)

Imagemagick	Alat za prikazivanje, pretvaranje i uređivanje slika.
Inkscape	Pretvara vektorske slike u „Scalable Vector Graphic“ (SVG) format.
JHOVE	Alat za provjeru formata datoteka.
MediaConch	Provjerava implementaciju, izvjestitelja i popravljача audiovizualnih datoteka.
MediaInfo	Prikaz najrelevantnijih tehničkih podataka i podataka o oznakama za video i audio datoteke.
Nailgun	Klijent, protokol i poslužitelj za pokretanje Java programa iz naredbenog retka.
NFS-common (Network File System Access)	Omogućava pristup datotekama na mrežnim uređajima za pohranu podataka.
p7zip	Alat za arhiviranje datoteka s velikim stupnjem kompresije.
Python-lxml	Spajanje Pythona za libxml2 i libxslt.
rsync	Alat za kopiranje datoteka.
Siegfried	Još jedan alat za prepoznavanje formata datoteka.
Sleuthkit	Alat za upravljanje i izvlačenje slike diska.
Tesseract	Alat za optičko prepoznavanje znakova (OCR), čita slikovne datoteke te ih pretvara u tekst.
UUID	Naredbeno sučelje (CLI) za stvaranje DCE 1.1, ISO/IEC 11578:1996 i IETF RFC-4122 u skladu s univerzalnim jedinstvenim identifikatorom (UUID).
unar	Alat za raspakiranje datoteka.

3.3. Sučelje

Sve funkcije „Archivematica-e“ odvijaju se unutar sučelja kojem se može pristupiti prijavom putem web-preglednika. Sučelje predstavlja prostor za obradu digitalnog arhiva te omogućuje arhivistima lako premještanje paketa digitalnog materijala prije njihovog slanja na pohranu ili diseminaciju. Njen izgled se može vidjeti na slici 1. Pod brojem jedan na slici su prikazane kartice koje dijele sučelje. Broj dva označava gdje se vrši prijava i odjava korisnika. Broj tri prikazuje pakete koji se mogu nalaziti u „Prijenos“ (engl. *Transfer*) i „Unos“ (engl. *Ingest*) karticama gdje se prikazuje prijenos materijala ili podneseni informacijski paketi. Pod brojem četiri se nalaze mikroservisi i zadaci. Mikroservisi se mogu razgranati na zadatke unutar kojih se nalazi ikona zupčanika te pritiskom na nju otvara se nova kartica preglednika koja

prikazuje što se trebalo izvršiti za taj zadatak. Broj šest predstavlja ikone za prijavu i uklanjanje. Ikona za uklanjanje samo briše paket sa sučelja te se on može ukloniti bez obzira na to je li obrada završena ili nije. Preporuča se uklanjanje obrađenih paketa.¹⁶



Slika 1. Sučelje "Archivematica"

3.4. Očuvanje

„Archivematica“ održava izvorni format svih unesenih datoteka. Glavni zadatak je normalizirati datoteke u formate koji omogućuju očuvanje i pristup. „Archivematica“ grupira formate datoteka u tzv. politiku formata (engl. *Format policies*) (npr. tekst, audio zapis, video zapis, rasterska slika, vektorska slika itd.). Postoji registar pravila formata (engl. *The Format Policy Registry ili FPR*) te je on baza podataka koja omogućava korisnicima „Archivematica“ da definiraju koje bi radnje „Archivematica“ trebale poduzeti za određeni format datoteke, na primjer kako bi „Archivematica“ trebala normalizirati JPEG datoteku za dugoročno očuvanje. To je također primjer politike formata koja definira alate, pravila i naredbe koje će „Archivematica“ koristiti za izvršavanje željene akcije. Postupci očuvanja kao što su normalizacija, raspakiravanje paketa, karakterizacija, provjera valjanosti, identifikacija,

¹⁶ Archivematica: open-source digital preservation system. Web-based dashboard. Dostupno na: <https://www.archivematica.org/en/docs/archivematica-1.11/getting-started/overview/dashboard/#web-dashboard> (31. kolovoza 2020.)

provjera i transkripcija upravljaju se na kartici Planiranje očuvanja (engl. *Preservation Planning*) „Archivematica“.¹⁷

3.5. Klijenti

„Archivematicu“ je razvio „Artefactual Systems“ osnovan u siječnju 2000. godine kao konzultantska tvrtka za digitalno očuvanje koja se kasnije proširila i na razvoj softvera otvorenog koda, kao što je „Archivematica“, početkom 2006. godine. Razvijanje „Archivematice“ započeto je u suradnji s UNESCO-ovom „Memory of the World's Subcommittee on Technology“ i „City of Vancouver Archives“ te daljnji razvoj nastavlja i s partnerima „University of British Columbia Library“, „Rockefeller Archive Center“, „Simon Fraser University Archives and Records Management“, „Bentley Historical Library“ te brojnim drugim suradnicima.¹⁸

Neki od njihovih značajnijih klijenata je „University of British Columbia“ koji, osim što je jedan od ranijih financijera projekta „Archivematica“, ujedno i koristi „Archivematicu“ i „AtoM“ (ili „Access to Memory“ koji je internetska aplikacija otvorenog koda za stvaranje arhivskih opisa, također razvijena od strane „Artefactual Systems“¹⁹). „University of British Columbia“ koristi automatizirani alat za unos sadržaja razvijen za „Archivematicu“ za njihov repozitorij „cIRcle“ koji čuva niz digitalnih sadržaja kao što su: članci, radovi s konferencija i radionica, disertacije, završni i diplomski radovi, knjige, audiovizualni materijali, materijali za učenje, itd.²⁰

„Archivematicom“ se služi i „Provincial Archives of Saskatchewan“ kojemu „Artefactual“ pomaže prilikom migracije podataka u „AtoM“ (korišten za mrežni pristup sadržajima) radi uspostave programa digitalnog očuvanja koristeći se „Archivematicom“ (korištena za očuvanje sadržaja).²¹

¹⁷ Archivematica: open-source digital preservation system. Dostupno na: <https://www.archivematica.org/en/> (13. lipnja 2020.)

¹⁸ COPTR contributors. Archivematica. COPTR. Dostupno na: <https://coptr.digipres.org/index.php?title=Archivematica&oldid=3721> (30. kolovoza 2020.)

¹⁹ AtoM. Dostupno na: <https://www.accesstomemory.org/en/> (30. kolovoza 2020.)

²⁰ Artefactual. UBC Library running automated ingest from DSpace to Archivematica! Dostupno na: <https://www.artefactual.com/ubc-library-running-automated-ingest-from-dspace-to-archivematica/> (30. kolovoza 2020.)

²¹ Artefactual. Provincial Archives of Saskatchewan to adopt Archivematica and AtoM. Dostupno na: <https://www.artefactual.com/provincial-archives-of-saskatchewan-to-adopt-archivematica-and-atom/> (30. kolovoza 2020.)

Jedan od značajnijih partnera je i MoMA („Museum of Modern Art“) koji je zajedno s „Artefactualom“ razvio „Binder“, alat za upravljanje zbirkama u digitalnom repozitoriju.²²

²² Artefactual. MoMA and Artefactual launch new digital preservation management system! Dostupno na: <https://www.artefactual.com/moma-and-artefactual-launch-new-digital-preservation-management-system/> (30. kolovoza 2020.)

4. „Preservica“

„Preservica“ nudi mogućnost digitalnog očuvanja koje omogućava pohranjivanje i održavanje datoteka koje želimo sačuvati. Koristi pohranu temeljenu na cloudu i omogućuje korisničko sučelje putem web preglednika. Softver se temelji na arhivskim i digitalnim standardima očuvanja OAIS ISO 14721. „Preservica“ nudi pet različitih izdanja koja se mogu birati na temelju potreba organizacije: „Cloud Edition Essentials“ namijenjen manjim organizacijama, „Cloud Edition Professional“ za manje od srednje velike organizacije, „Enterprise Private Cloud“ za srednje do velike organizacije čiji je prioritet privatnost, „Enterprise Private Cloud Perform“ za veće organizacije sa zahtjevnom potražnjom i integracijskim potrebama i „Enterprise on Premise“ za srednje od velike organizacije s intrenom IT podrškom. „Preservica“ nije bazirana na tipu otvorenog koda te se ne može vidjeti kako funkcionira niti se može izmjenjivati iako su neke pojedinačne komponente razvijene i objavljene od strane „Preservica“ tima.²³

Glavne značajke koje sadrži su:

- Pohrana: lokalni alat za stvaranje SIP-ova i dodavanje metapodataka i prijenos na poslužitelj; tijekovi rada koji služe za prijenos datoteka: ZIP datoteke, DSpace, ContentDM ili SharePoint, datoteke e-pošte Outlook, Gmail ili Lotus Notes,
- Upravljanje podacima: uređivanje metapodataka koristeći bilo koju opisnu shemu, npr. Dublin Core, METS, EAD, itd.; sinkronizacija s katalogima, npr. CALM ili putem OAI-PMH; brisanje sadržaja,
- Administracija: stvaranje korisnika, postavljanje uloga, postavljanje pravila pristupa sadržajima; održavanje sporazuma o prijenosu; sinkronizacija s ActiveDirectory,
- Pristup: pregledavanje, pretraživanje i preuzimanje datoteka i zapisa pomoću sučelja i zasebno javno pristupno sučelje za sadržaje s odgovarajućim dozvolama,
- Pohrana: Veliki raspon adaptera za pohranu, uključujući NAS / File Server, HSM, Centera, Atmos, S3, Glacier, SFTP; može upravljati s više mjesta za pohranu i kopirati na udaljene lokacije neistovremeno; održava metapodatke i podatke o indeksima; može postaviti interval za provjeru ispravnosti,
- API (Aplikacijsko programsko sučelje ili engl. *application programming interface*) pristup: Sve verzije koriste OAI-PMH (engl. *Open Archives Initiative Protocol for Metadata Harvesting*) protokol, koji je razvijen kako bi prikupljao opise metapodataka

²³ Preservica: Active digital preservation. Dostupno na: <https://preservica.com/> (13. lipnja 2020.)

od zapisa iz arhiva kako bi se usluge mogle zasnivati i stvarati na osnovu prikupljenih metapodataka iz mnogih arhiva, i API za pristup metapodacima i sadržaju. Enterprise verzije imaju REST API za čitanje / pisanje i API za upravljanje procesima.²⁴

4.1. Aplikacijska programska sučelja „Preservice“

U tablici 2 je prikazan je niz aplikacijskih programskih sučelja (API) koje koriste vanjski sustavi za automatizaciju i korištenje sustava „Preservica“.

Tablica 2. API-ji „Preservice“²⁵

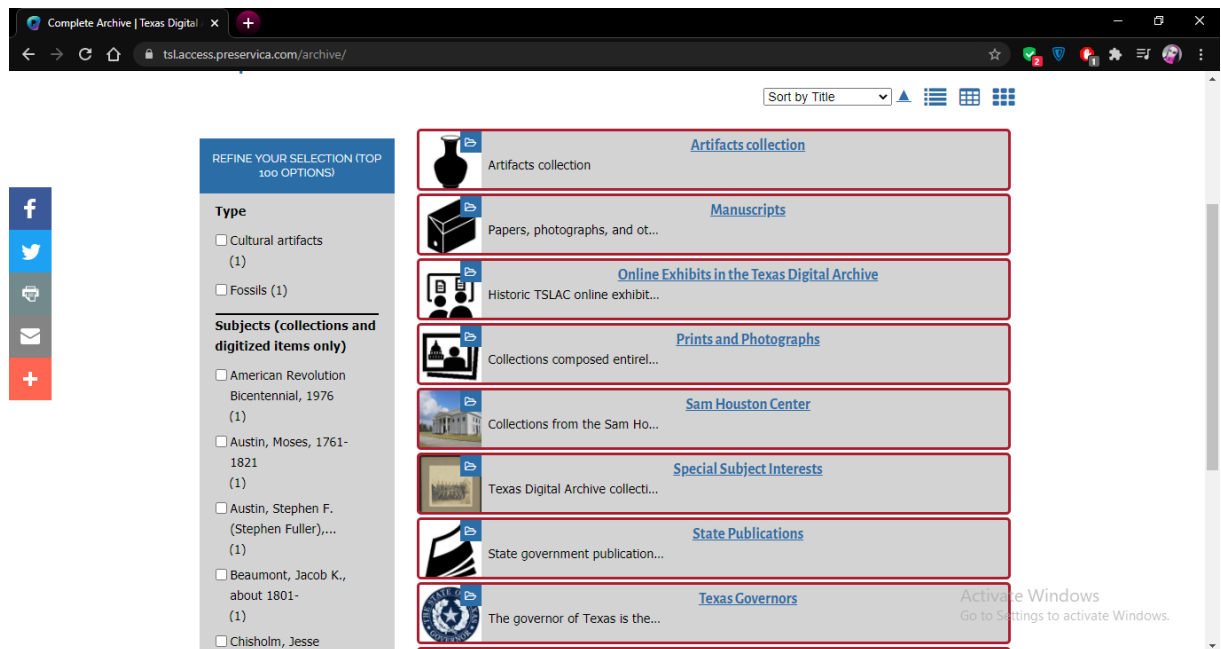
API	Opis
Preservica Entity REST API	Omogućuje čitanje „Preservicinog“ modela podataka i mogućnost stvaranja, izmjene i prijenos paketa.
Upload Content (S3 Compatible)	Implementacija S3 protokola za prijenos paketa.
Progress Token API	Nadzire dugotrajne zadatke kao što je prijenos datoteka putem Entity API-ja.
Process Monitoring API	API za dohvaćanje i ažuriranje nadzornih informacija o procesima.
Access Token API	API za provjeru autentičnosti.
Preservica CMIS Interface	Standardni protokol za interakciju sa sustavima za upravljanje sadržajem.
Content API	API s mogućnošću pretraživanja za pristup metapodacima i sadržajima.
SIP Creator Command Line	Program za kreiranje „Preservicinih“ paketa za predaju.
OAI-PMH Data Provider	Standardni protokol za skupljanje metapodataka.
Linked Data Registry REST Services	Pruža stvaranje, čitanje, ažuriranje i brisanje objekata u registru „Preservice“.

²⁴ COPTR contributors. Preservica. COPTR. Dostupno na: <https://coptr.digipres.org/index.php?title=Preservica&oldid=3350> (13. lipnja 2020.)

²⁵ Preservica: Active digital preservation. Preservica APIs - reference information. Dostupno na: <https://developers.preservica.com/api-reference> (31. kolovoza 2020.)

4.2. Pristup sadržaju

„Preservica“ pruža jedinstveni pristup (engl. *Universal Access*, UA) koji omogućuje pristup pohranjenim slikama, dokumentima, audiovizualnim datotekama, e-mailovima, itd. UA daje mogućnost odabira koji će se materijal odobriti za javni pristup dok se pristup povjerljivijem materijalu može ograničiti implementacijom lozinki. Audiovizualnim sadržajima se može pristupiti odmah bez dodatnog skidanja datoteke ili nekog web dodatka za gledanje takvih zapisa. Pohranjene slike se može pregledavati u punom zaslonu, približavati ili smanjivati te su uz njih priloženi i metapodaci i opis sadržaja. Sučelje, osim što dostupno na stolnim i prijenosnim računalima, optimizirano je i za pametne telefone i tablete. Izgled sučelja se može temeljiti na već gotovim predlošcima ili se za drugačije dizajne može koristiti CSS.²⁶



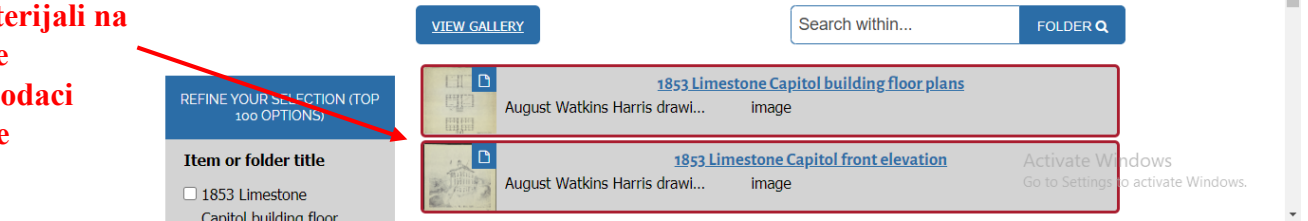
Slika 2. Sučelje za pristup materijalima "Texas Digital Archives"

²⁶ Preservica: Active digital preservation. Easy to customize, out-of-the box access and discovery. Dostupno na: <https://preservica.com/digital-archive-software-1/document-access-1> (1. rujna 2020.)

1. Metapodaci



2. Materijali na koje se metapodaci odnose



Slika 3. Sučelje za pristup materijalima "Texas Digital Archives"

1. Metapodaci



3. Slika na koju se odnose metapodaci

Slika 4. Sučelje za pristup materijalima "Texas Digital Archives"

5. Usporedba sustava „Archivematica“ i „Preservica“

I „Archivematica“ i „Preservica“ koriste isti pristup očuvanja, pa čak i iste alate u mnogim slučajevima. Oba sustava se temelje na OAIS referentnom modelu što uključuje: pakiranje bilo kakvih datoteka koje se žele sačuvati s nekim opisnim informacijama, premještanje ovog paketa u digitalni sustav za očuvanje i zadnje mogućnost sustava da izvodi niz manje ili više prilagodljivih radnji kako bi učinio zadatke poput identificiranja i provjere datoteka, izvlačenje ili generiranje metapodataka, izrade kopija u formatima za koje su održiviji, provjera virusa i drugih zadataka.

„Archivematica“ je na bazi otvorenog koda te ima dobro uspostavljenu rastuću zajednicu korisnika koji se uglavnom sastoje od visokoškolskih ustanova. Otvoreni kod omogućava da vidimo kako on točno funkcionira te ga se tako može razvijati prema vlastitim potrebama. Međutim, mnogi korisnici nemaju resurse ili vještine da to izravno iskoriste, pa se umjesto toga oslanjaju na vodećeg programera „Archivematice“, „Artefactual Systems“, za razvoj novih ažuriranja, pružanje obuke, podrške i dokumentacije. Otvoreni kod pomaže u ublažavanju rizika da se postane ovisan o jednoj organizaciji kako bi sustav funkcionirao. Kod sustava „Preservica“ kod nije objavljen i ne može se vidjeti kako funkcionira, iako je neke pojedinačne komponente poput DROID-a (engl. *Digital Record Object Identification*) besplatnog softverskog alata koji pomaže automatski prepoznati formate datoteka²⁷, razvio tim „Preservica“ i pustio u javnost. Također ima međunarodnu zajednicu korisnika, većinom u obrazovnim ustanovama, ali također i u korporativnim i poslovnim, regionalnim i nacionalnim arhivima. Također novi zahtjevi mogu se predložiti korisničkoj grupi „Preservica“ što onda „Preservica“ može, a i ne mora primijeniti što znači da se neke funkcije mogu razviti po mjeri. Kako je veliki broj korisnika, utjecaj jedne pojedinačne organizacije ili ustanove nije velik stoga ako se želi ostvariti neka nova funkcija mora se stvoriti neka vrsta udruženja s drugim institucijama kako bi se zahtjev za novom funkcijom izvršio. Rizik da se postane ovisan o jednoj organizaciji se ublažava uporabom otvorenih API-ja za izvoz podataka i metapodataka, mada općenito oslanjanje na samo jednu organizaciju i dalje predstavlja rizik kod „Preservica“.

Fokus „Archivematice“ je na pohrani te je pretpostavka da se koriste drugi alati za upravljanje i pružanje pristupa materijalom nakon pohrane dok se „Preservica“ više može smatrati cjelovitim sustavom za upravljanje i očuvanje, koji je glavni u zadacima vezanim uz očuvanje.

²⁷ The National Archives. File profiling tool (DROID). Dostupno na: <https://www.nationalarchives.gov.uk/information-management/manage-information/policy-process/digital-continuity/file-profiling-tool-droid/> (25. lipnja 2020.)

Za organizacije i institucije koje tek započinju sa svojim radom bi „Preservica“ bila bolji izbor zbog njegovog jednostavnijeg sučelja za korištenje dok je „Archivematica“ bolja za već iskusnije organizacije i institucije koje također cijene fleksibilnost te mogu iskoristiti otvoren kod koji nudi „Archivematica“.

Oba sustava mogu primiti razne korisnike s raznim dozvolama pristupa no za korištenje sustava „Archivematica“ preporuča se da ga koristi obučeno osoblje. S druge strane „Preservica“ je jednostavnija za pregledavanje čak i ako netko nije stručnjak te ima izgled „Windows Explorera“ što ga čini lakšim za korištenjem.

Jedna od glavnih razlika između „Archivematice“ i „Preservice“ je to što „Archivematica“ dijeli rad mikroservisa u dvije različite faze – faza prije pohrane i faza same pohrane dok „Preservica“ njihov rad odjednom za vrijeme pohrane. Takva podjela rada daje strukturiranu informaciju o tome što će se točno očuvati prije pohrane u sustav za očuvanje.

Identifikacija formata je jedna od najvažnijih značajki sustava za digitalno očuvanje te se u slučaju oba sustava javljao problem pri prepoznavanju formata, najčešće prilikom rada s podacima istraživanja. Oba sustava koriste isti registar za određivanje formata datoteke, „PRONOM“, koji je vrlo koristan, ali ima ograničenu podršku. Registar većinom pokriva formate datoteka koji se koriste u kontekstu vlade, ali lošije pokriva one nastale u visokoškolskom krugu. Uz dodatna sredstva bi se moglo pokrenuti više istraživanja o datotekama čiji se format ne može prepoznati te kontaktirati stvaratelje i dobiti informacije o softveru koji se koristio za izradu datoteka. Tada bi se format datoteke mogao službeno identificirati i dodati PRONOM-u. Oba sustava nude mogućnost ponovnog pokretanja alata za prepoznavanje formata datoteka u nekom trenutku nakon pohrane kako bi se iskoristile prednosti ažuriranja PRONOM-a ili drugih registara datoteka. Ovaj postupak je malo lakši u „Preservici“ jer omogućuje pokretanje tog postupak preko više zbirki odjednom, dok „Archivematica“ zahtijeva ponovnu pohranu pojedinih zbirki.

AIP je paket podataka i metapodataka koji se stvara i pohranjuje. Jedan je od ključnih proizvoda digitalnog sustava za očuvanja. Iako postoje standardi koji se koriste u kreiranju AIP-ova (poput METS i PREMIS), ne postoji standardni AIP. AIP-ovi stvoreni u „Archivematici“ imaju više složenih metapodataka i složeniju strukturu od onih koji su stvoreni u „Preservici“ koji su jednostavniji i razumljiviji za ljude no vidljivi samo u sustavu.

„Archivematica“ nije zamišljen kao pristupni sustav nego potiče stvaranje DIP-ova za pristup putem odvojenog sustava. „Atom“ bi bio najočitiji odabir za to jer dijele prodavača. Za pristup

„Preservica“ ima ugrađeni „Universal Access Viewer“ koji omogućuje pretraživanje i gledanje digitalnog sadržaja. Također može proizvesti DIP-ove za druge pristupne sustave.

I „Archivematica“ i „Preservica“ se mogu konfigurirati za upotrebu lokalne pohrane ili clouda ili kombinacije ovog dvoje.²⁸

Tablica 3. Usporedba sustava „Archivematica“ (A) i „Preservica“ (P)²⁹

		Opis	A	P
Prihvat	Kopiranje (copy)	Automatski stvara glavnu i kopiju za pristup	X	X
	Provjera ispravnosti (fixity check)	Provjerava da datoteka nije promijenjena koristeći digitalni potpis ili „checksum“	X	X
	Skeniranje virusa (virus scan)	Provjerava ima li zlonamjernih programa	X	X
	Dedupliciranje datoteka (file dedupe)	Provjerava postojanje dupliciranih datoteka		X
	Automatski jedinstveni ID (auto unique ID)	Automatski generira jedinstveni ID za neki objekt	X	X
Procesiranje	Automatsko stvaranje metapodataka (auto metadata creation)	Automatski stvara informacije o objektu na temelju podataka koji su ukomponirani u njemu	X	X
	Automatsko prikupljanje metapodataka (auto metadata harvest)	Automatski prikuplja informacije o objektu koje se nalaze u nekom vanjskom repozitoriju, katalogu itd.	X	X
	Ručni unos metapodataka (manual metadata)	Omogućuje ručno dodavanje informacija	X	X
	Upravljanje pravima (rights management)	Omogućuje uspostavljanje prava na pristup, korištenje i izmjenu objekta	X	X
	Metapodaci paketa (package metadata)	Metapodaci dodijeljeni na razini paketa	X	X
	Automatsko stvaranje SIP-a (auto SIP creation)	Automatski generira SIP usklađen s OAIS-om	X	X

²⁸ Ingram, C., (2019). Piloting institutional digital preservation. Jisc Research infrastructure and data. Dostupno na: <https://researchdata.jiscinvolve.org/wp/2019/11/07/piloting-institutional-digital-preservation/> (25. lipnja 2020.)

²⁹ Digital Powrr. Digital Preservation 101: Tool Grid. Dostupno na: <https://digitalpowrr.niu.edu/digital-preservation-101/tool-grid/#Preservica> (2. rujna 2020.)

Pristup	Sučelje za javni pristup (public interface)	Pružajući pristup objektima krajnjim korisnicima		X
	Automatsko stvaranje DIP-a (auto DIP creation)	Automatski generira DIP usklađen s OAIS-om	X	X
Pohrana	Automatsko stvaranje AIP-a (auto AIP creation)	Automatski generira AIP usklađen s OAIS-om	X	X
	Pouzdana, dugoročno očuvanje bitova (reliable, long-term bit preservation)	Očuvanje objekta na razini bita u razdoblju dužem od sedam godina		X
	Redundancija (redundancy)	Čuva više kopija objekta		X
	Geodisperzirani model pohrane podataka (geographically dispersed data storage model)	Čuva više kopija objekta u više geografskih regija		X
	Izlazna strategija (exit strategy)	Ima plan za masovno premještanje ili uklanjanje svih pohranjenih objekata		X
Održavanje	Migracija (migration)	Konverzija objekta iz jednog formata u drugi format za koji se smatra da ima veću stabilnost	X	X
	Praćenje (monitoring)	Automatski provjerava ima li oštećenja na objektu		X
	Automatski oporavak (auto recovery)	Automatski zamjenjuje oštećeni objekt neoštećenom kopijom		X
Drugo	Otvoreni kod (open source)	Postojanje otvorenog koda	X	
	Jasnoća dokumentacije (clear documentation)	Alat je razumljivo objašnjen na netehničkom jeziku	X	X
	Trošak (cost)	Cijena za jednu godinu ili jedan TB	Be spl atn o	Var ira

6. Zaključak

S vremenom se povećava broj podataka nastalih u digitalnom obliku, bilo da je stvoreno kao rezultat digitalizacije ne-digitalnih kolekcija ili kao dio svakodnevnog poslovanja organizacije ili pojedinca, sve više i više informacija stvara se digitalno. Ova se pojava odvija u okruženju u kojem raste svijest o značajnim izazovima povezanim s osiguravanjem kontinuiranog pristupa tim materijalima, čak i u kratkom roku. S jedne strane, velike su mogućnosti koje digitalna tehnologija nudi za brz i učinkovit pristup informacijama i njihovo dugoročno očuvanje. S druge strane, postoji rizik da će digitalni materijali biti stvoreni na takav način da im se ne može osigurati ni kratkoročna održivost, još manja vjerojatnost da će i buduće generacije imati pristup njima. Digitalne tehnologije i formati se mijenjaju s vremenom stoga se organizacije i institucije moraju pobrinuti da se sadržaj digitalnih formata očuva kako bi im se moglo pristupiti i kako bi se mogli koristiti u budućnosti čak i u slučaju zastarijevanja originalne tehnologije. Prilikom odabira sustava za digitalno očuvanje mora se obratiti pažnja na to može li prihvatiti novi materijal te kasnije njime upravljati u različitim verzijama te koje tehnologije i metode koristiti za njihovo očuvanje.

Jedni od takvih sustava su „Archivematica“ i „Preservica“ svaki sa svojim prednostima, ali nedostacima. Zato je na organizacijama ili institucijama, koje žele arhivirati i očuvati svoje materijale, da procijene koji sustav ima alate koji će im najviše koristiti i pomoći pri očuvanju digitalnih materijala.

7. Literatura

Archivemata: open-source digital preservation system. Dostupno na:

<https://www.archivemata.org/en/> (13. lipnja 2020.)

Archivemata: open-source digital preservation system. External tools. Dostupno na:

<https://www.archivemata.org/en/docs/archivemata-1.11/getting-started/overview/external-tools/#external-tools> (29. kolovoza 2020.)

Archivemata: open-source digital preservation system. Web-based dashboard. Dostupno na:

<https://www.archivemata.org/en/docs/archivemata-1.11/getting-started/overview/dashboard/#web-dashboard> (31. kolovoza 2020.)

Artefactual. MoMA and Artefactual launch new digital preservation management system!

Dostupno na: <https://www.artefactual.com/moma-and-artefactual-launch-new-digital-preservation-management-system/> (30. kolovoza 2020.)

Artefactual. Provincial Archives of Saskatchewan to adopt Archivemata and AtoM.

Dostupno na: <https://www.artefactual.com/provincial-archives-of-saskatchewan-to-adopt-archivemata-and-atom/> (30. kolovoza 2020.)

Artefactual. UBC Library running automated ingest from DSpace to Archivemata!

Dostupno na: <https://www.artefactual.com/ubc-library-running-automated-ingest-from-dspace-to-archivemata/> (30. kolovoza 2020.)

AtoM. Dostupno na: <https://www.accesstomemory.org/en/> (30. kolovoza 2020.)

CCSDS Secretariat. (2012). Reference Model For An Open Archival Information System

(OAIS). Dostupno na: <https://public.ccsds.org/Pubs/650x0m2.pdf> (22. lipnja 2020.)

COPTR contributors. COPTR. Dostupno na: https://coptr.digipres.org/Main_Page (24. lipnja 2020.)

COPTR contributors. Preservica. COPTR. Dostupno

na: <https://coptr.digipres.org/index.php?title=Preservica&oldid=3350> (13. lipnja 2020.)

Digital Powrr. Digital Preservation 101: Tool Grid. Dostupno na:

<https://digitalpowrr.niu.edu/digital-preservation-101/tool-grid/#Preservica> (2. rujna 2020.)

Digital Preservation Coalition, (2015). Digital Preservation Handbook. Dostupno na:

<https://www.dpconline.org/handbook> (24. lipnja 2020.)

Digital Preservation Coalition, (2015). Digital Preservation Handbook: Preservation issues. Dostupno na: <https://www.dpconline.org/handbook/digital-preservation/preservation-issues> (31. kolovoza 2020.)

Digital Preservation Coalition, (2015). Digital Preservation Handbook: Tools. Dostupno na: <https://www.dpconline.org/handbook/technical-solutions-and-tools/tools> (31. kolovoza 2020.)

Ingram, C., (2019). Piloting institutional digital preservation. Jisc Research infrastructure and data. Dostupno na: <https://researchdata.jiscinvolve.org/wp/2019/11/07/piloting-institutional-digital-preservation/> (25. lipnja 2020.)

McMaster University. Definitions AIP, DIP, SIP. Dostupno na: <https://digitalarchive.mcmaster.ca/node/56> (22. lipnja 2020.)

Pećarina, G., (2014). Digitalno arhiviranje. InfoTrend. Dostupno na: <http://www.infotrend.hr/clanak/2014/10/digitalno-arhiviranje,81,1094.html> (22. lipnja 2020.)

Preservica: Active digital preservation. Dostupno na: <https://preservica.com/> (13. lipnja 2020.)

Preservica: Active digital preservation. Easy to customize, out-of-the box access and discovery. Dostupno na: <https://preservica.com/digital-archive-software-1/document-access-1> (1. rujna 2020.)

Preservica: Active digital preservation. Preservica APIs - reference information. Dostupno na: <https://developers.preservica.com/api-reference> (31. kolovoza 2020.)

The National Archives. File profiling tool (DROID). Dostupno na: <https://www.nationalarchives.gov.uk/information-management/manage-information/policy-process/digital-continuity/file-profiling-tool-droid/> (25. lipnja 2020.)

Popis slika

Slika 1. Sučelje "Archivematice"	13
Slika 2. Sučelje za pristup materijalima "Texas Digital Archives"	18
Slika 3. Sučelje za pristup materijalima "Texas Digital Archives"	19
Slika 4. Sučelje za pristup materijalima "Texas Digital Archives"	19

Popis tablica

Tablica 1. Mikroservisi	11
Tablica 2. API-ji „Preservice“	17
Tablica 3. Usporedba sustava „Archivematica“ (A) i „Preservice“ (P)	22

Analiza odabranih sustava za digitalno arhiviranje

Sažetak

Digitalno arhiviranje je očuvanje digitalnih materijala koje neka tvrtka ili osoba treba sačuvati na duže vremensko razdoblje. Njime se omogućuje lakše sustavno pretraživanje arhiviranih podataka. Sustavi namijenjeni digitalnom arhiviranju pohranjuju zbirke digitalnih podataka poput dokumenta, zapisa, ali i video zapisa, slika, itd. u digitalnom formatu s namjerom dugoročnog pristupa. Digitalni arhiv može biti složena zbirka s višeslojnim sustavom za pohranu (npr. hijerarhijski sustav) ili se može nalaziti na sustavu s direktnim pristupom. U ovom radu će se pobliže pojasniti neka osnovna pitanja vezana uz digitalno arhiviranje i dugoročno očuvanje, analizirati sustave „Archivematica“ i „Preservica“ te usporediti njihove bitne karakteristike i svojstva kako bi se pobliže moglo pojasniti i predstaviti rad takvih sustava.

Ključne riječi: digitalno arhiviranje, digitalno očuvanje, sustav, Archivematica, Preservica

Analysis of selected digital archiving systems

Summary

Digital archiving is the preservation of digital materials that a company or person needs to preserve for a longer period of time. It makes it easier to systematically search archived data. Systems intended for digital archiving store collections of digital data such as documents, records, but also videos, images, etc. in digital format with the intention of long-term access. A digital archive can be a complex collection with a multi-layered storage system (e.g. a hierarchical system) or it can be located on a direct access system. This thesis discusses some basic issues related to digital archiving and long-term preservation, analyzes the systems "Archivematica" and "Preservica" and compares their important characteristics and properties in order to clarify and present the work of such systems.

Keywords: digital archiving, digital preservation, system, Archivematica, Preservica