

UNIVERZA V LJUBLJANI
EKONOMSKA FAKULTETA

DIPLOMSKO DELO

JAKOB KOTNIK

UNIVERZA V LJUBLJANI
EKONOMSKA FAKULTETA

DIPLOMSKO DELO

**PRIMERJALNA ANALIZA TEHNIK
ARHIVIRANJA DIGITALNIH DOKUMENTOV**

Ljubljana, junij 2003

JAKOB KOTNIK

IZJAVA

Študent **Jakob Kotnik** izjavljam, da sem avtor tega diplomskega dela, ki sem ga napisal pod mentorstvom **dr. Jurija Jakliča** in dovolim objavo diplomskega dela na fakultetnih spletnih straneh.

V Ljubljani, dne 13. 6. 2003

Podpis:
Jakob Kotnik

KAZALO

1. UVOD	1
2. DOKUMENTI V SODOBNEM POSLOVANJU	2
3. ARHIV	4
3.1 Opredelitev in pomen arhiviranja	4
3.2 Elektronski arhiv	5
3.2.1 Elektronski arhiv v praksi	6
4. ZAKONODAJA IN STANDARDI	9
4.1 Zakonodaja.....	9
4.2 Standardi	10
4.2.1 Uporaba standardov za dolgoročno hranjenje dokumentov.....	10
5. ODPRT ARHIVSKI INFORMACIJSKI SISTEM	14
5.2 Informacije, ki sodelujejo pri procesu arhiviranja	14
5.2.1 Informacijski paketi	15
5.2.2 Različice informacijskih paketov	17
5.3 Model OAIS.....	17
5.3.1 Podrobni opis funkcionalnih entitet.....	19
6. DOLGOROČNO HRANJENJE INFORMACIJ	25
6.1 Papirna kopija in mikrofilm	26
6.2 Tehnološki muzeji.....	26
6.3 Migracija	27
6.4 Emulacija	29
6.5 Enkapsulacija	31
7. ANALIZA TEHNIK ARHIVIRANJA DOKUMENTOV	32
7.1 Prednosti in slabosti tehnik digitalnega arhiviranja	34
7.2 Izbira najustreznejše tehnike.....	36
7.2.1 Opis kriterijev	37
7.2.2 Ocena tehnik	39
8. ZAKLJUČEK	43
9. LITERATURA	44
10. VIRI	46
11. PRILOGA	

1. UVOD

Že od nekdanj je človek čutil potrebo po zapisovanju dogodkov, zamisli in ugotovitev z namenom, da bi ohranil svoje znanje za kasnejše rodove. Tako so Mezopotamci že v tretjem tisočletju pr. n. št. pisali na glinaste ploščice, ki so še danes shranjene v kakšnih muzejih. S pojavom papirja so ljudje ugotovili, da je ta veliko bolj praktičen od predhodnih medijev, a tudi manj obstojen. Zaradi večje možnosti in praktičnosti uporabe se je število dokumentov izredno povečalo, prav tako pa se je povečala tudi skrb za njihovo ohranjanje. Tako so v 15. st. nastali prvi arhivi, z njimi pa tudi veda, ki se ukvarja z ohranjanjem in urejanjem dokumentov – arhivistika.

Nič drugače ni danes. V dobi elektronskih informacij papir vse bolj izgublja na veljavi. Elektronski dokumenti skupaj z moderno tehnologijo, v primerjavi s tradicionalnim papirjem, omogočajo nove možnosti in praktičnosti uporabe, zaradi česar pa se eksponentno hitro povečuje količina teh dokumentov. Problem je v tem, da danes ne moremo uporabiti nekaj let starega elektronskega dokumenta, če ta ni ustrezno ohranjen, po drugi strani pa še vedno lahko beremo zapise izpred nekaj tisočletij. Prav dejstvo hitre zastarelosti elektronskih dokumentov postavlja današnje elektronske arhive pred veliko nalogo. Tako kot Sumeri moramo tudi mi iznajti tehniko, ki bo omogočala hranjenje današnjih dokumentov tisočletja dolgo.

Vendar pa je implementacija arhiva, ki bi dolgoročno hranila elektronske dokumente, vse prej kot enostavna. Ta ni odvisna le od tehnoloških rešitev, pač pa tudi od zakonskih zahtev in standardov, ki tehnološke rešitve usmerjajo in so zato prav tako zelo pomembni za razvoj sodobnih arhivov. Danes sicer obstaja več strategij dolgoročnega arhiviranja, a žal zaenkrat še nobena ni popolna.

V tej diplomski nalogi želim skozi zgoraj omenjeni problem podrobno predstaviti sodobni arhiv, zakonodajo, ki ga omejuje, sodobne standarde, ki zagotavljajo splošno sprejetje teh arhivov, ter aktualne tehnike dolgoročnega arhiviranja, ki nam puščajo upanje, da bo naše znanje prav tako ohranjeno za naslednje rodove, kot so ga za nas ohranili naši predniki. Cilj naloge je, da s primerjalno analizo izberem tisto tehniko, ki napoveduje najbolj svetlo prihodnost stroki arhiviranja, kljub temu, da so si mnjenja strokovnjakov o tem, katera je najbolj primerna, danes še precej neenotna.

2. DOKUMENTI V SODOBNEM POSLOVANJU

Dokument je lahko sinonim za pismeno potrdilo ali dokazilo o nekem dejstvu, hkrati s tem pa tudi dokaz obstoja nekega upravnno-opravnega ali drugega dokumentiranega dejanja ali stanja (Novak, 2002, str. III-60). Dokumenti in informacije pridobivajo vse večjo vlogo tudi v poslovanju v podjetjih. Podjetja jih razumejo kot ključ do kakovosti in poslovnega uspeha ter posledično konkurenčnosti posameznega podjetja.

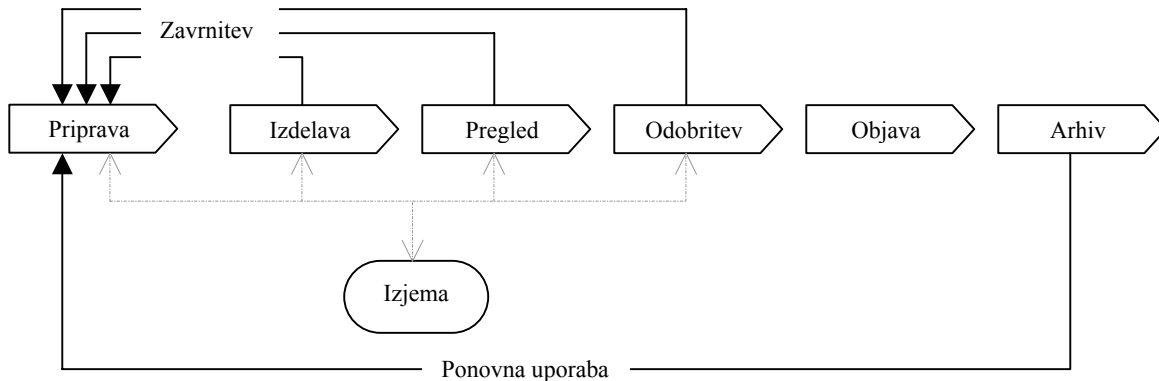
Upravljanje dokumentov je v slovenskem prostoru bolj prisotno zadnjih nekaj let. Pojavlja se na različnih aplikativnih področjih. Med drugimi mednje spadajo upravljanje dokumentacije sistema kakovosti, upravljanje poslovne dokumentacije ter kompleksnejši, kot so obvladovanje registracijskih dosjejev itd. Ta področja so podprta z bolj ali manj kompleksnimi sistemi za upravljanje dokumentov. Cilj upravljanja dokumentov je zagotovitev ustreznih informacij oz. dokumentov končnim uporabnikom s čim manjšimi stroški.

Vsak dokument ima svoj življenjski cikel. Pri realizaciji dokumenta morata sodelovati vsaj njegov tvorec in njegov naslovnik. Pri nastajanju tega pa nastopajo fizične ali pravne osebe v najmanj treh različnih vlogah. Prvo vlogo opravlja fizična ali pravna oseba, ki ima pooblastilo za izdelavo dokumenta v svojem imenu ali v imenu nekoga drugega. Drugo vlogo opravlja sam ustvarjalec dokumenta, ki ima odgovornost, da sestavi dokument v vsebinsko celoto. Tretjo vlogo pa opravlja fizična oseba, ki dokument sprejema kot rezultat ali sled zahteve po njegovem nastanku.

Zaradi narave poslovanja se zahteve po določenem dokumentu povečujejo, razširjajo in dopolnjujejo ter s tem postajajo vse bolj kompleksne. Z njimi pa postajajo vse bolj kompleksni tudi življenjski cikli dokumenta. Te so strokovnjaki skušali razdeliti v tri skupine. Prvo skupino sestavljajo aktivnosti upravljanja aktivnih ali tekočih zapisov (ustvarjanje, sprejemanje in razdelitev dokumentov), drugo skupino sestavljajo aktivnosti, vezane na upravljanje polaktivnih in poltekočih zapisov (urejanje dokumentov po vnaprej dogovorjenem sistemu), tretjo skupino pa tvorijo aktivnosti upravljanja neaktivnih ali nedejavnih oz. netekočih zapisov (hranjenje in izločanje dokumentov) (Novak, 2002, str. III-61).

V poslovnih sistemih dokumenti prehajajo iz ene faze v drugo fazo življenjskega cikla na podlagi sprožitvenih pravil, ki so vezana na izpolnitev določenih pogojev (npr. pogoj, da dokument napreduje iz faze »odobritev« v fazo »odobren«, je ta, da vsi odobritelji odobrijo dokument). Z življenjskimi cikli je definirana pot dokumentov od priprave do končne uporabe oz. arhiviranja preko različnih procesov kot so pregledi, odobritve, distribucija itd. Primer takega življenjskega cikla je prikazan na Sliki 1.

Slika 1: Primer preprostega življenjskega cikla dokumenta



Vir: Dular, Jakovljević, 2000, str. V-42.

Življenjski cikel dokumenta kot dokumentarnega gradiva se običajno zaključi z arhiviranjem v arhiv, kjer se dokumenti dolgodobno varujejo in ohranjajo kot arhivsko gradivo. Pomembno je, da je dolgoročno shranjevanje dokumentov zanesljivo in varno. V ta namen je treba zagotoviti: (Dobnikar, Žužek, 2002, str. 270)

- zanesljivost vsebine dokumenta,
- avtentičnost dokumenta in subjekta, ki je dokument ustvaril,
- celovitost oz. garancijo, da dokument ni bil spremenjen delno ali v celoti,
- uporabnost oz. berljivost dokumenta ter v primeru podpisanih dokumentov tudi verifikacijo podpisa v daljšem ali trajnem časovnem obdobju,
- pravno veljavnost arhiviranega dokumenta.

Za zagotavljanje teh pogojev pa tako v svetu kot pri nas še ni enostavnih rešitev oz. standardov, s tem pa tudi ne pravnih oz. zakonskih podlag, ki bi na splošno oz. globalno določale načine in pravno veljavo arhiviranja elektronskih dokumentov. Po drugi strani pa se že izvajajo različne iniciative in priporočila tako za tehnološki kot pravni vidik, ki bi reševala problem dolgoročnega arhiviranja elektronskih dokumentov.

V nadaljevanju bom ta problem predstavil bolj podrobno. Opozoril pa bom tudi na danes znane predloge in rešitve tega problema.

3. ARHIV

3.1 Opredelitev in pomen arhiviranja

Arhiviranje je postopek prevzemanja, hranjenja, vzdrževanja, strokovne obdelave in uporabe dokumentarnega¹ in arhivskega gradiva v zbirki dokumentarnega gradiva oziroma v arhivu ustanove ali posameznika. Arhivira se dokumentarno gradivo, ki je rešeno oz. zaključeno (*ad acta*) in ni več neposredno (stalno) potrebno za tekoče poslovanje. Dokumentarno gradivo arhiviramo zaradi različnih potreb in hranimo v arhivu ustanove, dokler ne potečejo roki hranjenja, ki jih narekujejo predpisi in potrebe poslovanja, ali, dokler dela dokumentarnega gradiva, ki ima značaj arhivskega gradiva (trajni pomen za zgodovino, znanost ali kulturo), ne odberemo ali izločimo pristojnemu javnemu ali zasebnemu arhivu (Žumer, 2001, str. 48). Arhiv pa lahko razumemo kot zbirko listin in dokumentov, ki imajo vrednost kot zgodovinsko gradivo, ali kot prostor oz. ustanovo, kjer so shranjene te listine in dokumenti (Reference Model for an Open Archival Information System, 2002).

Arhivirana pisna gradiva ali zapisi vsebujejo pomembna dejstva, podatke, informacije, risbe, slike, glasove, glasbo, umetniške stvaritve, v zadnjem času pa tudi virtualno resničnost. Te s splošnim izrazom imenujemo tudi dokumenti. Arhiv jih prejme v trajno last, s tem pa sprejme obvezo, da zagotavlja vsebino teh dokumentov tako, da je ta ustrezna in dostopna, razumljiva ter na voljo ciljnim (javnim in zasebnim) skupinam uporabnikov. Hkrati pa mora zagotavljati tudi dolgoročno nespremenljivost in verodostojnost hranjenih dokumentov.

V zgodovini so se taki zapisi pojavljali v formatih, kot so knjige, papir, zemljepisne karte, fotografije, filmi itd. Ti so imeli neposredno vrednost za ljudi, saj so jih lahko uporabljali s pomočjo raznih preprostih povečevalnih pripomočkov oz. celo brez teh.

Tradicionalni arhivi so se osredotočali predvsem na hranjenje informacij na medijih z dolgoročno stabilnostjo in kontroliranim dostopom. Z razvojem tehnike so nosilci zapisov postajali vse bolj raznovrstni: od najstarejših zapisov na kamnu in glinastih ploščicah prek papirusa, pergamenta, papirja, fotografskega in filmskega traku ter mikrofilma do najnovejših magnetnih in optičnih oziroma elektronskih nosilcev, ki so začeli nastajati z razvojem informacijske tehnologije. Na nosilce (medije) so zapisi vtisnjeni z najrazličnejšimi orodji oz. strojno ter programsko opremo - od peres in pisalnih strojev do računalniških programov - ter z najrazličnejšimi sredstvi - od črnila do magnetnega in laserskega načina zapisa. Glede na to, v kakšni obliki arhiv hrani informacije, lahko tako ločimo tradicionalne arhive ter novodobne - digitalne arhive.

¹ Dokumentarno gradivo obsega vse vrste dokumentov, ki nastajajo pri poslovanju in delu ustanov ter posameznikov, in vse prejete dokumente. Arhivsko gradivo je hranjeno v arhivih in je praviloma le del dokumentarnega gradiva. Je izvorno in reproducirano in služi kot zgodovinski vir za preučevanje preteklosti in sedanjosti.

S tem v zvezi lahko delimo tudi dokumente. Te delimo na tradicionalne, digitalne ter elektronske. Tradicionalni dokumenti so tisti, ki jih uporabnik uporablja v njihovi fizični obliki. Digitalni dokumenti vsebujejo zapis v digitalni obliki, a so lahko predstavljeni tudi v neelektronski obliki. Medtem pa so elektronski dokumenti ustvarjeni s pomočjo računalnika in jih lahko uporabljamo samo z njegovim posredovanjem.

Pri ugotavljanju odnosa med elektronskimi dokumenti in celoto arhivskega in dokumentarnega gradiva je treba opozoriti, da so elektronski dokumenti praviloma del celote arhivskega in dokumentarnega gradiva, ki nastaja pri ustvarjalcih v nekem času in prostoru. Praktično to pomeni, da elektronskih dokumentov s stališča vrednotenja arhivskega gradiva ni mogoče izvzeti iz celote nastalega gradiva ali ga obravnavati ločeno samo zaradi njegove drugačne oblike (Novak, 2002, str. III-3).

Tako bo v tej nalogi glavna beseda namenjena digitalnim oz. elektronskim dokumentom in hranjenju teh, ki jim po besedah Terryja Kunyja grozi »izumrtje«. Kuny pravi, da se danes nahajamo v mračnem srednjem veku digitalne dobe, ko bo večina znanja, ki ga hranimo v elektronski obliki, za vedno izgubljena. Zato je pred nami velika naloga, podobna nalogi nekdanjim samostanskih opatov, ohraniti čimveč elektronskih zapisov (Kuny, 1998).

3.2 Elektronski arhiv

Elektronski arhiv se od tradicionalnega razlikuje v tem, da ta hrani dokumente v elektronski obliki. Danes živimo v času, ko je uporaba elektronskih dokumentov postala naša vsakdanjost. Tako so se začele pojavljati zahteve po ohranjanju elektronskih zapisov za nedoločeno prihodnost. S tem pa se postavljajo vprašanja, kako brati te dokumente čez petdeset, sto in več let in v kakšni obliki jih hraniti, da bodo razumljivi v prihodnosti.

S pojavom elektronskih dokumentov in s tem elektronskih arhivov se je spremenila tudi logika arhiviranja. Ta ne temelji več izključno na vzdrževanju in ohranjanju nosilcev in na njih zapisanih sporočilih, ampak predvsem na vzpostavljanju, vzdrževanju in razumevanju kontekstov, v katerih so zapisi nastali ali bili uporabljeni (Novak, 2002, str. III-6). Brez teh digitalni dokumenti izgubijo svoj pomen in tako predstavljajo le še množico zaporedij bitov, ki sami po sebi ne povedo nič. Ta problem bom podrobneje izpostavil, ko bom opisoval zgradbo digitalnih objektov in njim pripadajočih opisnih informacij, problem razumevanja konteksta pa sem s primerom prikazal v prilogi C.

Kaj pa se sploh dogaja znotraj digitalnih arhivov? Digitalni dokumenti se tako kot tisti v drugih oblikah pomikajo skozi življenjski cikel. So ustvarjeni, urejeni, opisani in indeksirani, razposlani, pridobljeni, uporabljeni, podvrženi reviziji, ponovnemu ustvarjanju, spremenjeni in hranjeni za prihodnost ali uničeni s strani kompleksne združbe ustvarjalcev in drugih lastnikov, distributerjev ter institucionalnih in individualnih uporabnikov. Življenjski krog

informacij se vrta v začaranem krogu ustvarjenja, pridobitev, katalogizacije/identifikacije, shranitve, hranjenja in dostopa.

3.2.1 Elektronski arhiv v praksi

Digitalna tehnologija ima svoje prednosti in slabosti. Na eni strani odpira nove možnosti uporabe, hkrati pa predstavlja nevarnost nedelovanja. Njena funkcionalnost je povezana z njeno kompleksnostjo. Branje in razumevanje digitalnih informacij zahteva strojno in programsko opremo, ki se z neprestanim razvojem tehnologije nenehno spreminja (Garret, Waters, 1996).

Prednost dokumentov v digitalni obliki (Garrett, 1996) so:

- možnost ustvarjanja popolnih kopij digitalnih zapisov,
- možnost njihovega oglaševanja na številnih medijih,
- možnost distribucije in širjenja na omrežjih,
- možnost preoblikovanja v drugo obliko,
- možnost njihovega lociranja,
- možnost iskanja po vsebini,
- lahko jih povrnemo nazaj, če se izgubijo, ter
- lahko jih obdelamo z avtomatiziranimi (avtomatskimi) in polavtomatiziranimi orodji.

Pri hranjenju in zagotavljanju dostopa do arhiviranih informacij digitalna tehnologija pomaga reševati tradicionalni konflikt med hranjenjem in dostopom do gradiv. Glavni problem tradicionalnih arhivov je njihova okornost, saj je upravljanje s papirnimi ali drugimi analognimi dokumenti v primerjavi z digitalnimi precej težavno in nepregledno, še posebej, ko razpolagamo z ogromnimi količinami informacij, poleg tega pa so zanje potrebne tudi ogromne tehnične ter prostorske kapacitete, ki zahtevajo visoke stroške hranjenja.

S pojavom elektronskega arhiva so glavne pomanjkljivosti tradicionalnega odpravljene. Elektronski arhiv na tem mestu omogoča preprosto iskanje zelenih vsebin po raznovrstnih atributih dokumentov, prostorsko so ti neomejeni, dostopni časi do vsebin pa so v primerjavi s tradicionalnim arhivi zanemarljivi. Torej funkcionalnost govori v prid elektronskim arhivom. Poleg tega digitalni podatki omogočajo tudi integracijo informacij iz različnih virov.

Na drugi strani se elektronskem arhiviranju postavlja kar nekaj bistvenih vprašanj. Številni tehnični, socialni in pravni vidiki povzročajo precejšnje težave specialistom na področju arhiviranja že nekaj desetletij.

Bistvena vprašanja oz. problemi glede digitalnega arhiviranja so:

Porast števila digitalnih objektov

Število digitalnih dokumentov se zaradi prednosti, ki jih prinaša digitalna tehnologija, vztrajno povečuje. Nadzor nad temi nam vse bolj uhaja iz rok. Vprašanje je, ali ohraniti vse digitalne dokumente ali določiti kriterije, po katerih bi določili dokumente, ki jih je treba hraniti.

Kompleksnost digitalnih dokumentov

Danes se poleg tekstovnih dokumentov pojavljajo novi z zahtevnejšo strukturo. To so multimedijske datoteke (avdio, video), zahtevne programske kode, grafični prikazi itd. Veliko vprašanje je, kako jih hraniti dolgoročno - tudi potem, ko bo izvorno okolje, v katerem so dokumenti nastali, že zastarelo. Dolgoročna rešitev hranjenja dokumentov mora biti neodvisna od tehnologije.

Odgovornost nad dokumenti

Odgovornost za dokument si v življenjskih ciklih dokumenta deli več udeležencev (ustvarjalec, lastnik, uporabnik, ...), vendar ni jasno, kdo je za dokument odgovoren v celoti. Knjižnice bi to odgovornost sicer lahko prevzele, vendar so običajno samo posredniki dokumentov. Arhivi za kaj takega niso primerno organizirani. Posamezni avtorji dokumentov pa zaradi ostalih problemov česa takega ne bi zmogli. To nalogo lahko prevzamejo komercialna podjetja, ali pa se v ta namen organizira javna inštitucija.

Financiranje

Celo življenjsko obdobje dokumenta (vključno z arhiviranjem) zahteva precejšna finančna sredstva. Stroškov ne moremo globalno ločiti, temveč moramo predvideti, da bodo stroški arhiviranja povezani z vsemi fazami življenjskega cikla digitalnega dokumenta (Fenney, 1999). Denarni vidik je eden pomembnejših pri iskanju ustrezne rešitve za ohranitev dokumentov.

Različni standardi in različni formati

Problem hitrega tehnološkega razvoja je tudi v poplavi informacij, ki niso kontrolirane s strani knjižnic ali arhivov. Soočamo se s problemom, ki narašča hitreje, kot ga lahko rešujemo. Žal še ni veliko sprejetih standardnih formatov, ki bi upočasnili zastarevanje informacij. S sprejetjem takih standardov bi bil omogočen tudi lažji dostop do arhiviranih vsebin.

Problem avtentičnosti in integritete

Potrebe po dokazovanju avtentičnosti in integritete naraščajo zaradi velike odgovornosti do digitalnega objekta, ki ga zelo lahko kopiramo in spreminjamo. Pri procesu arhiviranja obstaja nevarnost sprememb in poškodb dokumenta. Graham (1998) predlaga izbiro tehnike, ki varuje dokument pred nepooblaščenim kreiranjem in dostopom do dokumenta. Ena od

rešitev je tako imenovana »*Digital Time-stamping*« (DTS), ki je podobna sistemu poštnega žiga.

Pomanjkanje razumevanja terminov v zvezi s hranjenjem

Novak (2002) pravi, da logika ohranjanja digitalnih dokumentov temelji na vzpostavljanju, vzdrževanju in razumevanju okolja (kontekstov), v katerem so zapisi nastali in bili uporabljeni. Prvi pogoj za dojetanje teh procesov in aktivnosti pa je razumevanje arhivske strokovne terminologije.

Pomanjkanje pozornosti za ohranitvene zahteve

Večjo pozornost zahtevam hranjenja dokumentov je treba nameniti že pri nakupu oz. kreaciji digitalnih objektov, saj je od tega precej odvisna samo hranjenje in dostopnost do sedanjih digitalnih objektov.

Vprašanje avtorskih pravic in intelektualne lastnine

Pomembno vprašanje je tudi, kdo ima avtorske pravice nad arhivirano vsebino ter čigava je ta intelektualna lastnina. Temu vprašanju je bilo do danes namenjeno premalo pozornosti, tako da to še ni rešeno v celoti. Rešitev problema pa bi lahko precej vplivala na možnosti prezervacije digitalnih objektov.

Tehnološka zastarelost

Tehnološka zastarelost je rezultat razvoja tehnologije: novejša tehnologija izrivajo uporabo obstoječe. Novejša verzija programov neprestano povzročajo zastarevanje starejših verzij, temu pa se mora prilagoditi tudi strojna oprema, ki je pogoj za delovanje novih programov. Posledično postanejo informacije, ki so temeljile na zastarelih tehnologijah, nedostopne.

Zaradi hitrega tehnološkega razvoja je življenjska doba programske in strojne opreme zelo kratka. Poleg tega pa so tudi mediji, na katerih so hranjeni podatki, zelo nestabilni in zato presenetljivo življenjsko omejeni. Glavno vprašanje je, kako v takem okolju ohraniti arhivirano vsebino tako, da bo na voljo v nekem trenutku v prihodnosti, ko bo na razpolago povsem druga tehnologija.

S tem namenom se razvija več dolgoročnih strategij hranjenja digitalnih dokumentov. Te so npr. migracija, emulacija, enkapsulacija, omeniti pa velja tudi tehnološke muzeje ter papirno tehniko. Te strategije bom bolj podrobno opisal v poglavju o strategijah dolgoročnega arhiviranja. Te bom med seboj tudi primerjal in poskušal ugotoviti, katera je za dolgoročno arhiviranje najbolj primerna.

Odgovore na mnoga postavljena vprašanja bi lahko v prihodnosti zagotovila tudi ustrezna zakonodaja in sprejetje ustreznih standardov. Sprejetje teh je zelo pomembno za nadaljnji

tehnološki razvoj in implementacijo ter širšo uporabo novih oz. do danes razvitih sistemov za arhiviranje.

4. ZAKONODAJA IN STANDARDI

Zaradi vseh potreb, ki jih ima sodobna družba glede dolgoročnega arhiviranja sodobnih dokumentov, so elektronski arhivi postali nujnost. Ker pa gre za precej novo področje, kjer ni enotnih standardov, zakonodaja pa je o problemu začela razmišljati šele pred kratkim, moramo najprej pogledati, kaj od tega je na voljo, da bi se elektronski arhivi sploh razvili in zaživel kot hranilec informacij.

4.1 Zakonodaja

Elektronski dokumenti dobijo dimenzijo kulturne in poslovne dediščine, ki jo je treba zaščititi na dolgi rok, šele v povezavi z zakonom o arhivih in arhivskemu gradivu – ZAGA (Ur. l. RS, št. 20/1997), ki je bil v Sloveniji sprejet leta 1997. Ta ureja odnose in pravilno poslovanje z arhivskim gradivom, zakonsko določa varstvo arhivskega gradiva in uporabo tega, podrobneje pa opisuje tudi pristojnosti in naloge posameznih arhivov.

V letu 2000 je bil sprejet Zakon o elektronskem poslovanju in elektronskem podpisu – ZEPEP (Ur. List, št. 57), ki je prinesel veliko novosti na področju arhiviranja na elektronskih medijih. Ta podpira uporabo elektronskih medijev in dokumentov ter jim daje večjo pravno-formalno enakopravnost v primerjavi s papirnim dokumentom, če le omogoča preverjanje zgodovine oz. nastanka dokumenta. Ta zakon je zelo pomemben predvsem za arhive zasebnih gospodarskih družb, podjetij, zavodov, društev itd. (Klasinc, 2001, str. I-68).

Skupne točke in rešitve v ZAGA in v ZEPEP lahko z vidika slovenske arhivistike pomenijo neke skupne rešitve. Na eni strani omogočata zavarovanje pomembnih dokumentov za njihovo vsestransko dolgodobno uporabo, na drugi strani pa določata pomembnim dokumentom tako vrednost, da jih lahko razglasimo za arhivsko gradivo.

Tak primer je npr. zakonsko določena oblika dokumenta, ki ga želimo hraniti. Oblike in kategorije arhivskega gradiva so v slovenski arhivski teoriji in praksi zelo dobro opredeljene. Zakon zahteva, da jih je v nadaljnjo hrambo v pristojni arhiv treba predati v izvorniku, s čimer sta zagotovljeni integriteta in avtentičnost dokumenta (Novak, 2001, str. I-72).

Na tem mestu se mi zdi prav, da na kratko omenim tudi sredstva za zagotavljanje avtentičnosti elektronskih dokumentov. Gre za elektronski podpis, ki predstavlja alternativo klasičnemu podpisu. Za arhiviste in podjetja je zelo zanimiv, saj zagotavlja edinstvene podatke, nedvoumno identiteto imetnika digitalnega potrdila, nezatajljivost lastništva

podpisanih e-podatkov ter celovitost (integriteto) e-podatkov, kar pomeni, da ni mogoče spremeniti ali drugače popraviti podatkov brez podpisnika (Dobnikar, Žužek, 2002, str. 172).

Pojavlja pa se vprašanje življenjske dobe digitalnih potrdil in šifrirnih ključev. Stopnja varnosti ključa se s starostjo zmanjšuje, zato imajo omejeno obdobje veljavnosti. To pa vpliva na varnost elektronskih dokumentov, ki jih je treba hraniti dlje, kot je obdobje veljavnosti digitalnih potrdil.

Pomembna je tudi revizijska sled, ki je še bolj kot vodenju elektronskih arhivov namenjena elektronskemu obvladovanju dokumentov. Je skupek vseh informacij, ki beležijo aktivnosti nad hranjenim dokumentom skozi ves življenjski krog dokumenta. Je kontrolno okolje, ki omejuje grožnje ponarejanja in dejanskega delovanja načrtovanega sistema. Pomembno pri arhiviranju je, da so vsi shranjeni dokumenti verodostojna kopija originala, kar zahteva tudi ZEPEP (Novak, 2001, str. I-72).

Sama zakonodaja pa ni dovolj za razvoj dolgoročnega sistema za arhiviranje. Ta postavlja le smernice, ki lahko usmerjajo nastajanje, pogosto pa so zakoni sprejeti šele po implementaciji prvih rešitev. Za zagotavljanje pogojev za trajno hranjenje elektronskih dokumentov tako v svetu kot pri nas še ni tehnološko dorečenih rešitev oz. standardov, posledično pa tudi ne pravnih oz. zakonskih podlag, ki bi na splošno določale načine in pravno veljavo arhiviranja elektronskih dokumentov (Novak, 2002, str. III-57). Razkorak med teorijo, to je sprejeto zakonodajo, in prakso pokaže čas. Sprejeta zakonodaja je pogosto rezultat priporočil in standardov (in obratno). Standardi tako predstavljajo široko sprejete temelje, ki jih je treba upoštevati pri razvoju novih proizvodov.

4.2 Standardi

Dolgoročno hranjenje digitalnih dokumentov je danes postala »vroča tema«. Kot odgovor na mnogo zahtev in potreb različnih organizacij, ki so spoznale potrebo po dolgoročnem hranjenju dokumentov, je predlaganih nekaj možnih rešitev. Zelo velik, če ne celo glavni problem, ki ga prinaša hiter razvoj informacijske tehnologije in informacijskih sistemov, je zastarevanje strojne in programske opreme ter formatov, ki jih obstoječa tehnologija uporablja. Standardizacija proces staranja upočasnjuje.

Slaba stran standardov je, da njihova implementacija in uporaba zahtevata veliko denarja. Poleg tega se pogosto dogaja, da standardi med seboj niso združljivi, ali pa so združljivi samo delno.

4.2.1 Uporaba standardov za dolgoročno hranjenje dokumentov

Za ohranjanje digitalnih dokumentov skozi čas so pristojne naslednje vrste standardov (Lorist, Van der Meer, 2001):

- standardi, ki služijo kot referenčni model; ti opisujejo funkcionalnost ter delovanje, postopke in koncepte digitalnih arhivov,
- standardi, ki omogočajo ohranjanje digitalnih dokumentov in njihovih predstavitev,
- metapodatkovni standardi, ki omogočajo dostop do hranjenih vsebin, tako da opisujejo kontekst, strukturo, izvor ter anotacijo hranjenih vsebin,
- standardi za usklajeno delovanje.

Standardi kot referenčni model

Standardi so nujno potrebni za nadaljnji razvoj arhivov. Ena glavnih organizacij na področju standardiziranja je zagotovo ISO, ki je do sedaj razvila že mnogo standardov na različnih področjih. Mnogokrat so standardi ISO osnova za razvoj novih standardov tudi drugih organizacij.

ISO 15489. V širokem digitalnem okolju upravljanje digitalnih zbirk ne more biti odgovornost le ene centralne organizacije, zato je pomembno, da se v takem okolju sprejme koncepte, definicije ter postopke, ki veljajo na splošno. Ta standard omogoča organizacijam, da standardizirajo pojme in definicije, ki so povezane s področjem upravljanja z dokumenti, reguliranim okoljem, politiko in odgovornostmi, postopki in kontrolami na tem področju.

AS 4390. Standard, razvit v Avstraliji, je primer standarda ISO 15489, ki je sprejet v Severni Ameriki in Evropi. Ustanovljen je bil leta 1996 in kot ISO standard predstavlja strategije in delovne napotke najboljše prakse arhiviranja. Nanaša se na vse vrste zapisov (Lorist, Van der Meer, 2001, str. 3).

DoD 5015.2-STD. Podoben standard je razvit s strani ameriškega obrambnega ministrstva. Predlaga smernice implementacij in postopkov, namenjenih sistemom za upravljanje z dokumenti. V svojo zakonodajo ga vključuje vse več držav, ki razširjajo svoje zakone na tem področju.

OAIS. V zadnjem času je velika pozornost usmerjena v predlog za ISO standard, ki ga ponuja NASA. Gre za odprt arhivski informacijski sistem. Opisuje informacijske tokove in zahteve, ki naj bi jih izpolnjeval sodoben arhivski sistem. Z implementacijo tega se ukvarja mnogo pomembnih organizacij (NEDLIB, PANDORA, UK University Library) (Kavčič-Čolič, 2001, str. 216).

Zadnje čase OAIS zbuja veliko pozornosti med strokovnjaki, zato ga bom v nadaljevanju tudi podrobneje predstavil.

Standardi za ohranjanje vsebin

Standardi naj bi upočasnili proces zastarevanja dokumentov tako v semantičnem, kot tudi fizičnem smislu. Standard pa lahko tako dolgoročnost zagotovi le, če se ne spreminja oz. če omogoča kompatibilnost s prejšnjimi verzijami. Danes najbolj primerna standarda za dolgoročno ohranjanje vsebine kot oblika dokumentov v digitalnem arhivu sta *Portable Document Format* (PDF) ter *eXtensible Markup Language* (XML).

PDF. Je produkt podjetja Adobe, postal pa je »de facto« dokumentni standard. Uporablja image model Postscript jezika, ki preslika tekst in slike v natančni kopiji originala. Žal uporabniki nimajo formalnega vpliva na izboljšave novih verzij, zato ta ne zagotavlja kompatibilnosti s prejšnjimi verzijami. Iskalne možnosti zaenkrat še ne izpolnjujejo zahtev digitalnih arhivov.

XML. Kljub razvoju in veliki popularnosti jezika XML ta ni idealna izbira za hranjenje zapisov. Vzratna kompatibilnost je - podobno kot pri PDF formatu - omejena: njegov predhodnik (SGML) pogosto le s težavo prebira XML datoteke. Poleg tega XML še ne more zagotoviti ohranjanja originalnega izgleda, kar pa je nujno za zagotavljanje dokazne vrednosti digitalnih dokumentov.

Sicer pa tako PDF kot tudi XML nista bila razvita za namene zagotavljanja digitalnega dolgoročnega obstoja. Zato hranjenje avtentičnih zapisov še naprej ostaja problem.

Nosilci podatkov. Danes imamo velik seznam standardov za nosilce podatkov. Standardi, ki omogočajo podaljševati fizično hrambo bitnih podatkov, so npr. CD-ROM, magnetni trak, optično/magnetni nosilci, DVD standardi itd. Velika izbira nosilcev je posledica tržnih sil, ki vse prevečkrat uvajajo nove standarde zaradi velikih dobičkov, kar pa ima na kupce tako pozitivne kot negativne učinke.

Standardi ohranjanja dostopa do vsebin

Hramba digitalnih dokumentov v eni obliki bitov ni dovolj za hranjenje teh za neomejeno časovno obdobje. Potrebujemo tudi opise digitalnih objektov, o čemer bom več pisal, ko bom opisoval referenčni model OAIS. Gre za metapodatke, ki morajo vključevati informacije o kontekstu, ki je pomemben za nadzor, obnavljanje in interpretacijo digitalne informacije skozi čas. Hranjenje metapodatkov je bistveno. Danes obstaja nekaj standardiziranih metapodatkovnih shem.

Dublin core. Razvit je bil z namenom raziskovanja mrežnih virov. Uporablja 15 skrbno izbranih elementov, ki tehnično opisujejo vsebino dokumentov. Ti elementi so bili izbrani z namenom lažjega iskanja želene vsebine. Po mnenju nekaterih Dublin Core ne omogoča

zadovoljivega dostopa do zbirk podatkov, kljub temu da je široko uveljavljen. Dobra stran tega pa je, da je lahko dostopen in poceni (Kavčič-Čolić, 2001, str. 218).

MARC. Je drugi podoben standard. MARC zagotavlja predstavitev in komunikacijo knjižničnih in s tem povezanih informacij v strojno-berljivi obliki. Prednost uporabe metapodatkov MARC je v tem, da ni treba razviti specifičnih metod organiziranja knjižničnih informacij, kar prihrani ogromno dela, omogoča pa sodelovanje in izmenjavo teh podatkov z drugimi knjižnicami. Je industrijsko razširjen standard, ki zagotavlja trajen dostop do zapisov skozi čas.

ISAD(G). Tretji standard vsebuje splošna navodila za arhivsko popisovanje (Kernel-Umek, 2000, str. 26), ki zagotavljajo dosledne opise teh zapisov ter kooperacijo, skupno rabo in integracijo opisov v enoten informacijski sistem. Ta standard temelji na dolgoletnem proučevanju teorije arhiviranja. Uporablja 26 elementov, ki se delijo na več nivojev. Prednost tega standarda je dejstvo, da ga arhivisti razumejo in da predstavlja strukturo arhiva v povezavi s kontekstom opisa.

Z39.50. Je predhodnik ISO 23950, ki omogoča iskanje v heterogenih informacijskih sistemih na različnih platformah v različnih okoljih. Arhitektura Client/Server omogoča komunikacijo med dvema različnima računalnikoma. Poizvedbe lahko vrne v več sintaksah, kot so SUTRS², XML ali MARC. Kljub uspehu, ki ga je dosegal, pa se pri uporabi občasno pojavljajo napake, ko različni informacijski sistemi ne »govorijo« istega jezika.

Standardi za usklajenost delovanja

Pomemben je tudi standard, ki povezuje sisteme za hranjenje zapisov, čeprav se ne nanaša neposredno na lastnosti digitalnih dokumentov.

ODMA. Omogoča integracijo nekaterih namiznih aplikacij (npr. MS Word) v ODMA sistem za upravljanje z dokumenti.

DMA. Ta standard skuša reševati problem otokov informacij, ki nastajajo znotraj organizacij kot posledica uporabe različnih uporabniških aplikacij za upravljanje z dokumenti.

WebDAV. Od drugih se razlikuje. Kot brskalnik, ki uporablja HTTP protokol, lahko bere mrežne razširjene dokumente, hkrati pa avtorjem omogoča, da zapisujejo oz. popravljajo (*upload*) elemente teh dokumentov.

² Simple Unstructured Text Record Syntax, ki temelji na ASCII

Vsi omenjeni standardi so razviti z namenom implementacije splošno sprejetega sistema za dolgoročno arhiviranje, ki bi zagotovil varno in trajno hranjenje digitalnih informacij. S tem perečim problemom se ukvarjajo številne institucije. Te se pri implementaciji svoje rešitve naslanjajo na OAIS (*Open Archival Information System*), ki danes predstavlja enega največjih potencialov na tem področju. Ta teoretični model odprtega informacijskega sistema za arhiviranje je razvil *Consulative Committee for Space Data System* (CCSDS) organizacije NASA kot odgovor na standard ISO TC20/SC 13 z namenom, da bi ga v prihodnosti obravnavali kot osnovo za nadaljnjo standardizacijo na področju dolgoročnega arhiviranja.

5. ODPRT ARHIVSKI INFORMACIJSKI SISTEM

Vse hitrejši računalniki, vse večji prenosi podatkov po omrežjih in vse boljše povezanost povzročajo, da vse več organizacij uporablja informacije v digitalnih oblikah, ki so začele nadomeščati bolj tradicionalne medije kot npr. papir. Eksplozivna rast informacij v digitalni obliki pa je povzročila velik izziv ne samo za tradicionalne arhive, pač pa tudi za javni, profitni in neprofitni sektor. Nekatere organizacije že ugotavljajo, druge pa še bodo, da bodo morale tudi v prihodnosti ohraniti tradicionalni način arhiviranja, saj ugotavljajo, da se digitalne informacije z lahkoto izgubijo ali pokvarijo in jih je zato veliko težje ohraniti v primerjavi s tradicionalnimi, kot so papir in filmi. Neverjetna hitrost tehnološkega razvoja povzroča zastarelost strojnih in programskih sistemov v nekaj letih, kar pa postavlja velika vprašanja o tem, ali bodo sedanje oblike zapisov še naprej lahko učinkovito zagotavljale zelene informacije. Zaradi tega, in še posebej zaradi dejstva, da je informacije najlažje oz. edino možno shraniti in ohraniti le takrat, ko so te v svoji originalni obliki, je nujno, da se organizacije aktivno vključijo v izdelavo arhivov, ki bi omogočili ohranjanje informacij tudi na dolgi rok, neodvisno od strojne in programske infrastrukture.

Odprt arhivski informacijski sistem (angl. *Open Archival Information System* – v nadaljevanju OAIS) predstavlja ogrodje funkcionalnega in informacijskega koncepta, ki je primeren za razvoj vsakega arhiva. Še posebej je primeren za organizacije, katerih potreba je zagotoviti dolgoročno hranjenje digitalnih informacij. Pri tem dolgoročno pomeni obdobje hranjenja, ki ni odvisno od sprememb v tehnologiji, zaradi česar današnji elektronski arhivi postanejo zastareli ali celo neuporabni. Sama arhitektura rešitve temelji na XML (Novak, 2001, str. I-6).

5.2 Informacije, ki sodelujejo pri procesu arhiviranja

Za lažjo predstavo o delovanju tako OAIS kot tudi veliko preprostejšega arhiva moramo najprej spoznati nekaj temeljnih pojmov s področja arhiviranja. OAIS za hranjenje informacij zahteva njihovo točno definicijo, saj le tako lahko shrani prave informacije. Razumevanje teh je bistvenega pomena za razumevanje arhiva!

V nadaljevanju bom s tem namenom predstavil naslednje vrste podatkov: vsebinske informacije, ohranitvene opisne informacije, informacije o paketiranju ter opisne informacije (Reference Model for an Open Archival Information System, 2002, str. 4-30).

5.2.1 Informacijski paketi

Za delovanje arhiva sta bistveni dve vrsti prenosa podatkov. V prvem primeru gre za prenos podatkov od ustvarjalca informacij (avtorja) v arhiv, druga vrsta prenosa pa poteka iz arhiva do uporabnika arhiviranih informacij. V prvem primeru govorimo o sprejemu podatkov, v drugem pa o posredovanju informacij. Prenesene podatke oz. informacije, ki sodelujejo pri prenosu, in tiste, ki so shranjeni v arhivu, bomo poimenovali *informacijski paketi* (angl. *Information package*).

Informacijski paket konceptualno vsebuje dva tipa informacij – *vsebinsko informacijo* in *ohranitveno opisno informacijo*, slednja pa vsebinsko informacijo loči od drugih sorodnih informacij. Te pa dodatno opisujejo še *informacije o paketiranju* ter *opisne informacije*.

Vsebinska informacija (angl. *Content information*) predstavlja niz informacij, ki so cilj arhiviranja v OAIS. Ta predstavlja vsebino, ki jo hranimo. Hranjena informacija vsebuje digitalne in tudi fizične objekte. Digitalni objekt (angl. *Digital object*) je izvorno zaporedje bitov, ki predstavlja hranjeni dokument v binarnem zapisu. Zaporedje bitov se šele s pomočjo dodatnih Predstavitev informacij (angl. *Representation information*), ki služijo kot slovar, poveže v splošno prepoznani tip podatkov, kot je npr. številka, črka in v skupino teh, ki da celotnemu izrazu pomen na višji ravni.

Opisna informacija o ohranjanju (angl. *Preservation Description Information* – v nadaljevanju PDI) vsebuje informacije, ki zagotavljajo razumevanje vsebinskih informacij. Ta informacija je značilna za vse vrste arhivov (tudi tradicionalnih), ki omogočajo kontekstualna razmerja posameznih dokumentov. Te skrbijo za to, da se vsebina skozi čas ne izgubi med množico vseh informacij, da je jasno definirana in da razumemo okolje, v katerem je shranjena.

Ohranitvene opisne informacije vsebujejo štiri tipe informacij, ki omogočajo ohranjanje vsebine:

- *Izvor* (angl. *Provenance*), ki opisuje izvor ali vir vsebine (ta jo je imel v lasti ob njenem nastanku) ter njeno zgodovino in morebitne spremembe od njenega nastanka.
- *Kontekst* (angl. *Context*) povezuje vsebino z okoljem, v katerem se nahaja. Pove, kako je vsebina povezana z drugimi informacijami zunaj informacijskega paketa. Na primer - pojasni, zakaj je nastala in kako se povezuje s sorodnimi vsebinami, ki so na voljo.

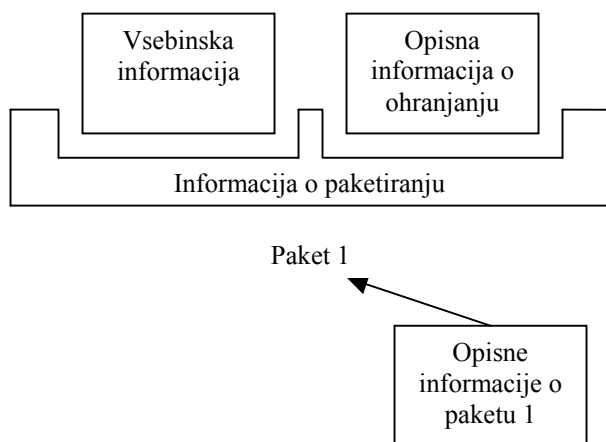
- *Vir* (angl. *Reference*) zagotavlja enega ali sistem identifikatorjev, s katerimi je vsebina enoznačno določena. Praktični primer je lahko ISBN številka knjige ali serija števil, ki loči eno vsebino od druge.
- *Stabilnost* (angl. *Fixity*), ki varuje vsebino pred nedovoljeno spremembo. Vsebuje preverjanje podatkovne integritete ali verifikacijske ključe, ki zagotavljajo, da vsebina ni bila spremenjena na nedovoljen način.

Informacija o paketiranju (angl. *Packaging information*) dejansko ali logično povezuje in identificira komponente informacijskega paketa v celoto na posameznem mediju. Na primer - če so vsebina in njene ohranitvene informacije hranjene v obliki določenih datotek na CD-ROM-u, potem informacija o paketiranju teh dveh vključuje ISO 9660 strukturo na CD-ROM-u, kot tudi imena datotek in informacij direktorija, v katerem se te nahajajo. Informacije, shranjene v imenih datotek ali direktorijski strukturi, so z zastarelostjo informacij o paketiranju lahko izgubljene.

Opisne informacije (angl. *Description information*) so v bistvu metapodatki. Vsebujejo informacije o informacijskih paketih, ki hranijo želeno vsebino. Glede na nastavitve se lahko te pojavljajo samo kot opisni naslov informacijskega paketa, lahko pa zajema vrsto atributov, ki služijo kot pomoč potencialnim uporabnikom za iskanje vsebine po raznih kriterijih v iskalnikih in katalogih.

Razmerja vseh teh informacij lahko prikažemo tudi s Sliko 2. Bolj podrobne povezave med zgoraj opisanimi informacijami znotraj informacijskih paketov pa so prikazane v prilogi B.

Slika 2: Povezava vseh tipov informacij: vsebinska informacija, ohranjanje opisnih informacij, paketne informacije in opisne informacije



Vir: Reference Model for an Open Archival Information System, 2002, str. 2-5.

5.2.2 Različice informacijskih paketov

Nujno je razločevati med informacijskimi paketi, ki so shranjeni znotraj arhiva, in med informacijskimi paketi, ki so sprejeti in oddani s strani arhiva. Tako govorimo o sprejetih oz. predloženih informacijskih paketih (angl. *Submission Information Package* – v nadaljevanju SIP), arhiviranih informacijskih paketih (angl. *Archival Information Package* – v nadaljevanju AIP) ter diseminacijskih oz. razposlanih informacijskih paketih (angl. *Dissemination Information Package* – v nadaljevanju DIP).

Predložen informacijski paket (SIP) preide v last arhiva od ustvarjalca oz. avtorja dokumentarnega gradiva. Forma in podrobna vsebina SIP-a je ponavadi določena med ustvarjalcem podatkov ter OAIS. SIP je v fazi zajema pretvorjen v AIP, ki se shrani v arhiv. Za izgradnjo AIP ne zadostuje nujno le en SIP, temveč jih je lahko potrebnih več. Prav tako pa lahko en SIP vsebuje informacijo, ki bo vključena v več AIP-jev.

Arhivski informacijski paket (AIP) je torej transformacija enega ali večih SIP-ov znotraj OAIS. Ustvarjen je z namenom, da poleg vsebine v hrambo vključi vse informacije, ki so potrebne za dolgoročno arhiviranje informacijskih objektov. AIP vsebuje celoten set potrebnih ohranitvenih opisnih informacij (PDI) povezanih z vsebino. AIP se trajno shrani v arhiv. Vsak AIP je povezan z opisno informacijo o paketu (angl. *Package description*), ki omogoča naročniku, da poišče informacijo, ki ga zanima, jo analizira in naroči.

Kot odgovor na zahtevo po arhivirani vsebini pa OAIS pretvori AIP v obliko diseminacijskega informacijskega paketa (DIP), ki je razumljiv naročniku. V kakšni obliki bo ta dostavljen, je odvisno od potreb uporabnika in medija, na katerem bo DIP dostavljen. Ponavadi so skupaj z DIP dostavljene tudi opisne informacije, ki naročniku omogočijo razlikovanje med večimi podobnimi DIP.

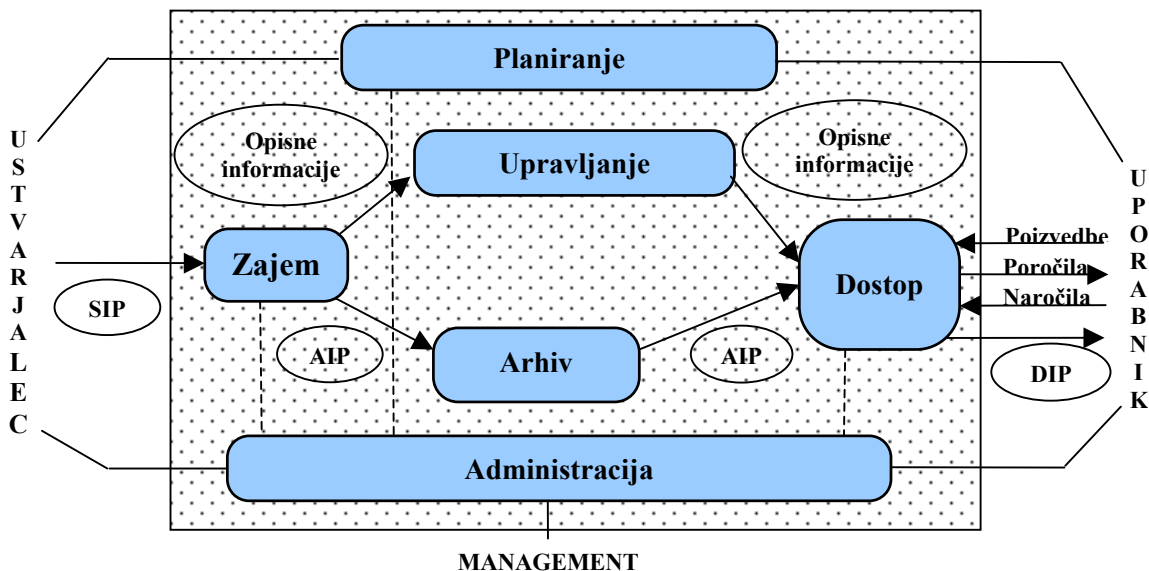
5.3 Model OAIS

S Sliko 3 želim prikazati glavne podatkovne tokove, ki so vključeni v OAIS model, katerega sestavlja šest funkcionalnih entitet, in s tem omogočiti širši pogled na arhiv³. S črtami in puščicami so prikazane komunikacijske poti. Smer puščice kaže smer podatkov oz. informacij, črta pa pomeni, da potekajo informacije v obe smeri. Črte, ki povezujejo entitete z administracijo in planiranjem, so tukaj zaradi boljše preglednosti izpuščene. V nadaljevanju bom vsako od notranjih entitet tudi opisal.

³ Na arhiv, ki ga prikazuje OAIS, lahko gledamo tudi bolj splošno. Ostali arhivi se v svojem bistvu ne razlikujejo mnogo od OAIS arhiva, saj za njimi stoji enaka logika arhiviranja. OAIS se od drugih arhivov razlikuje predvsem s svojo dolgoročno usmerjenostjo.

Slika prikazuje tudi transformacije informacijskih paketov in z njimi povezanih objektov znotraj OAIS s tem, ko sledijo življenjskim ciklom od ustvarjalca do OAIS in od OAIS do uporabnika.

Slika 3: Funkcionalne entitete OAIS modela



Vir: Reference Model for an Open Archival Information System, 2002, str. 4-1.

Zajem

Ta entiteta sprejema SIP od ustvarjalca (ali od funkcij znotraj administracije) ter ga pripravi in preoblikuje v ustrezno obliko (AIP), da bodo ti lahko shranjeni in obravnavani znotraj arhiva. Hkrati proizvede tudi opisne informacije o arhivirani vsebini, ki se shranijo v relacijsko bazo podatkov (entiteta upravljanje podatkov).

Arhiv

Glavna naloga entitete Arhiv je, da shrani, vzdržuje in priskrbi arhivirane informacijske pakete (AIP-je). Ta prejme AIP-je pripravljene v fazi Zajema in jih doda v trajno hrambo znotraj arhiva. Na podlagi zahtev s strani kupcev oz. uporabnikov jih pošlje entiteti Dostop.

Upravljanje podatkov

Entiteta Upravljanje podatkov shranjuje, vzdržuje in omogoča dostop do opisnih podatkov, ki identificirajo arhivirano vsebino, hrani pa tudi nekatere administrativne podatke, ki so pomembni za upravljanje arhiva. Na zahtevo zunanjih in notranjih uporabnikov dela poizvedbe na relacijski bazi, generira rezultat in ga v obliki poročila dostavi naročniku.

Administracija

Administracija skrbi za pravilno delovanje celotnega OAIS sistema. Zagotavlja, da celotni proces arhiviranja zadovoljuje standarde, ki so pomembni na področju dolgoročnega arhiviranja. Neprestano spremlja delovanje arhiva in izboljšuje operacije arhiviranja. Odgovorna je za uveljavljanje novih standardov in politik. Omogočati pa mora tudi podporo kupcem oz. uporabnikom.

Planiranje

Funkcija planiranja je, da spremlja okolje, v katerem se OAIS nahaja, in zagotavlja, da so informacije na voljo dolgoročno, četudi je prvotno računalniško okolje že dolgo zastarelo. Predlaga možne posodobitve arhiva, migracijo arhiviranih podatkov, nastanek novih standardov in politik, spremljati pa mora tudi spremembe v tehnologiji in zahtevah uporabnikov.

Dostop

Entiteta Dostop omogoča podporo uporabnikom v iskanju informacij o obstoju dokumentov, shranjenih v OAIS glede na opis, lokacijo, dostopnost itd. Navadno komunikacija OAIS-a in uporabnika poteka preko uporabniških vmesnikov. Želen zapis se iz arhiva (AIP) prenese v Dostop, kjer se transformira v Diseminacijski informacijski paket (DIP) glede na uporabnikove zahteve.

Poleg teh entitet pa so znotraj OAIS prisotne še nekatere druge funkcije. Za te bi lahko rekli, da tvorijo novo entiteto v tem modelu, poimenoval pa jih bom Skupne funkcije.

5.3.1 Podrobni opis funkcionalnih entitet

V nadaljevanju bom predstavil specifične tokove informacij znotraj vsake od entitet. Opisal bom tudi skupne funkcije, ki so značilne tudi za druge informacijske sisteme.

Skupne funkcije

Danes je na voljo ogromno podpornih funkcij, ki so že vgrajene v sodobne aplikacije in zagotavljajo varnost, varnostno shranjevanje, začasno shranjevanje, medprocesno komuniciranje, upravljanje z direktoriji, ... Delimo jih lahko na sistemske, mrežne ter varnostne funkcije.

Sistemske funkcije zagotavljajo osnovne funkcije, ki so pomembne za delovanje in administracijo platforme, na kateri temeljijo aplikacije, poleg tega pa ponujajo tudi vmesnik med programsko opremo in to platformo. Sem spadajo kernel operacije, ki upravljajo procese in izvršujejo programe, razni komandni ukazi in orodja, sistemske operacije in varovanje sistema, ki zagotavlja dostop do sistemskih podatkov, funkcij, strojnih in programskih virov le pooblaščenim uporabnikom.

Mrežne funkcije omogočajo mehanizme za podporo skupnim aplikacijam razmeščenim na večih mestih, ki potrebujejo dostop do skupnih podatkov in usklajeno delovanje v heterogenih mrežnih okoljih. Sem lahko prištevamo podatkovne komunikacije, ki temeljijo na raznih protokolih, transparenten dostop do ponujenih datotek znotraj omrežja, mrežne varnostne funkcije, ki vključujejo avtentikacijo uporabnikov, integriteto, zaupnost, upravljanje komunikacije med pošiljateljem in prejemnikom informacij znotraj omrežja itd.

Varnostne funkcije omogočajo mehanizme za varovanje občutljivih informacij znotraj informacijskega sistema. Primerna stopnja varnosti je določena glede na vrednost informacije za končnega uporabnika aplikacije. Med te lahko štejemo identifikacijo/avtentikacijo uporabnikov, ki želijo upravljati z informacijami; dostopne kontrole, ki onemogočajo nepooblaščen uporabo informacij; integriteto podatkov, ki skrbi, da podatki ne zastarijo oz. niso uničeni po nepooblaščen poti (velja predvsem za trajno hranjene podatke ter podatke v komunikacijskih sporočilih) in ki zagotavlja zaupnost podatkov ter omogoča dokazno funkcijo, ki dokazuje, da je bila informacija poslana in sprejeta.

Zajem

V Zajemu se pripravi primeren prostor za sprejem SIP od ustvarjalca dokumentarnega gradiva (ali administracije). Z raznih medijev (npr. disk, CD-ROM), ali preko elektronskega transfera (npr. FTP), se na ta prostor prenesejo digitalni paketi. Nedigitalni paketi so lahko dostavljeni kot pošiljka preko vsakdanjih poti. Funkcija sprejemanja podatkov lahko zahteva posebne dostopne kontrole do vsebine paketov.

Preveri se uspešnost prenosa paketov na začasno področje. Prenos podatkov lahko spremljajo mehanizmi, kot so CRC (*Cyclic Redundancy Checks*)⁴ in *checksums*⁵, lahko pa uporabimo tudi log datoteke⁶, ki identificirajo in zabeležijo tiste prenose, pri katerih je prišlo do napake bodisi pri branju ali zapisovanju digitalnih informacij.

Generira se AIP. Gre za pretvorbo enega ali več SIP-ov v enega ali več AIP-ov. Pretvorba mora potekati v skladu s standardi o formatiranju (oblikovanju) in dokumentiranju podatkov. Entiteta dobi poročilo o tem, s kakšnimi opisnimi podatki mora opremiti nastajajoč AIP pri shranjevanju. Poleg tega pa lahko pošlje SIP ali AIP v pregled v Administracijo, ki preveri, če ti ustrezajo standardom.

⁴ CRC je metoda preverjanja morebitnih napak med komunikacijskim prenosom podatkov. S prenesenimi podatki se prenese tudi koda (*cyclic redundancy code*). Če je ta enaka kot na začetku, je prenos potekal brez napak.

⁵ *Checksum* je podobna metodi CRC, le da ta preverja, ali ima prenesen paket na koncu prenosa enako število bitov, kot jih je bilo poslanih na začetku. Je manj zanesljiva od CRC metode.

⁶ Log datoteka hrani zapise o procesih, ki so se vršili na računalniku, in opozarja na morebitne napake.

V skladu s poročilom o potrebnih opisnih informacijah entiteta iz AIP in drugih virov zbere opisne informacije, ki vsebujejo informacije o arhivirani vsebini (AIP). Gre za metapodatke AIP-jev, ki omogočajo iskanje in pridobitev želenih AIP-jev.

AIP se shrani v arhiv in opisne informacije v bazo podatkov znotraj entitete Upravljanja podatkov. Prenos AIP-ja v arhiv sproži proces, ki vrne identifikacijsko informacijo o shrambi. To je treba vključiti v opisno informacijo pripadajočega AIP, bazo v katero se prenovljena opisna informacija shrani, pa posodobiti.

Arhiv

Entiteta Arhiv predstavlja arhiv v ožjem smislu in je bistvena v procesu arhiviranja. Sem se shranijo podatki, ki jih želimo arhivirati (AIP). Arhiv od entitete Zajem prejme AIP in zahtevo po njegovi shranitvi na trajen medij. Od pričakovane frekvence uporabe podatkov je odvisno, kakšen medij bo izbran za arhiviranje. V okviru te funkcije se pripravi medij in izvede fizični zapis. Po uspešno opavljenem prenosu se entiteti Zajem pošlje sporočilo o zapisu skupaj z identifikacijo zapisa AIP.

Elementi AIP-jev se preko ukazov razporedijo na primerne medije. Zagotovljena je določena raven varnosti (določeno je maksimalno dovoljeno število napak zapisov, izvajajo se backup procedure, ...), s spremljanjem napak (*error logs*) pa se zagotavlja, da med prenosom ni prišlo do okvar AIP-jev.

Zelo pomembna funkcija za OAIS je zamenjava medija znotraj Arhiva, ki omogoča dolgoročno hranjenje AIP-jev (migracija). To pomeni, da se na dolgi rok vsebina in ohranitvene opisne informacije skozi čas ne spreminjajo. Strategija migracije izbere nov medij, na katerega bodo prenešeni podatki, pri tem pa mora razmisliti o možnih napakah, ki se pojavljajo pri različnih izbirah tipov medijev, o zmogljivostih novih medijev, stroškov nakupa itd.

Za arhiv je pomembno, da zagotavlja obnovitev stanja pred uničenjem (angl. *Disaster recovery*). Gre za varno shranjevanje oz. podvajanje že shranjenih digitalnih podatkov, ki so kot kopija originala shranjeni na fizično ločen medij, ki je običajno prenosljiv (npr. digitalni linearni trak ali zgoščenka). Podrobnosti te politike določa administracija.

Glede na zahtevo uporabnikov po določenem dokumentu entiteta Arhiv posreduje kopijo AIP entiteti Dostop, kjer se transformira v DIP.

Upravljanje podatkov

Entiteta Upravljanje podatkov je relacijska baza podatkov, ki vsebuje opisne in sistemske podatke. Opisni podatki identificirajo in opisujejo vsebino arhiva, sistemske informacije pa so pomembne za upravljanje in delovanje arhiva.

Entiteta ima nalogo definiranja tabel in shem, ki so potrebne za podporo entitete Upravljanja podatkov, omogoča izdelovanje, uporabo in dostopanje do prikrojjenih uporabniških pogledov (angl. *User views*) nad vsebino arhiva ter zagotavlja referenčno integriteto podatkov.

Glede na zahteve uporabnikov entiteta opravi poizvedbo po vnaprej določenih kriterijih. Izvede jo na bazi, v kateri so shranjeni opisni podatki, rezultat pa vrne v Dostop. Rezultat so generirana poročila, ki so poslana v entitete Zajem, Dostop in Administracija. Običajna poročila vključujejo povzetke o arhivirani vsebini po kategorijah, ali pa statistiko dostopanja in izkoriščanja arhiviranih vsebin.

Podatki v bazi niso stalni, zato je pomembno, da entiteta Upravljanja podatkov z dodajanjem, spreminjanjem in brisanjem informacij iz baze zagotavlja posodabljanje baze. Glavni vir sprememb sta entiteta Zajem, ki izdeluje Opisne informacije za nove, v arhivu shranjene AIP-je, in entiteta Administracija, ki skrbi za posodobitev sistema. Slednja vključuje vse informacije, povezane z delovanjem sistema (statistika delovanja, informacije o uporabnikih, status zahtev). Pregled nad posodobitvami se opravlja s periodičnim pregledovanjem in posodabljanjem informacij, kot so npr. kontaktna imena, naslovi uporabnikov itd.

Administracija

Za celotni arhiv so pomembna pogajanja med OAIS-om ter ustvarjalcem podatkov, ki jih želi OAIS prejeti v trajno last. Gre za pogajanja o vročitveni pogodbi (angl. *Submission agreement*). Dogovorita se o razporedu vročitev, ko bo prenešen eden ali več SIP-ov v OAIS.

Glede na pogodbo mora entiteta preveriti, ali SIP in AIP izpolnjujeta zahteve, določene v vročitveni pogodbi. Pregled mora preveriti, če podatki ustrezajo nivoju kvalitete, kot jo zahtevata arhiv in komisija, ki jo sestavljajo strokovnjaki iz več področij. V primeru, da nekateri SIP-i ne ustrezajo dogovorjenim zahtevam, morajo biti ponovno vročeni ali pa izločeni. Potem, ko je proces zaključen, se izdela končno poročilo o zajemu, ki ga prejmeta tako OAIS kot ustvarjalec.

Administracija neprestano spremlja funkcionalnost celotnega arhivskega sistema. To vključuje spremljanje operacij sistema, njegove učinkovitosti ter obremenjenosti. Na podlagi informacij o sistemu, ki jih Administracija dobi v obliki poročil iz entitete Upravljanja podatkov ter Arhiva, redno pripravlja informacije o delovanju OAIS ter poročila o

skladiščenju vsebine znotraj arhiva. Glede na informacije o delovanju se znotraj Administracije uvedejo novi standardi in politike.

Na podlagi informacij o proračunu, politike poslovanja, obsega poslovanja, politike zaračunavanja storitev itn., ki jih določa *Management*, ter na podlagi predlogov in pobud za nove standarde s strani Planiranja Administracija osnuje nove standarde in politike⁷, ki veljajo za arhiv. V skladu z varnostno politiko sistema mora biti zagotovljen mehanizem omejitve fizičnega dostopa do arhiva.

V preteklosti so bili lahko nekateri zahtevani podatki nedostavljeni uporabniku, ker ti v trenutku naročila še niso bili na voljo. S pojavom teh entiteta generira zahtevo po diseminaciji, ki jo pošlje entiteti Dostop. Generira lahko tudi periodična naročila, katerih časovni razmak med naročili je določen s strani stranke, ali pa temelji na nekem dogodku (npr. posodobitev baze podatkov).

Administracija skrbi tudi za zaračunavanje storitev strankam. Ta prejme informacije o opravljenih storitvah, na podlagi katerih je kupcu izstavljen račun. Spremlja pa se tudi zadovoljstvo strank s ponujenimi storitvami.

Planiranje dolgoročnega hranjenja

Zelo pomembno za OAIS je, da se pravočasno odziva na zunanje spremembe. Za to skrbi entiteta Planiranje dolgoročnega hranjenja. Ta spremlja zahteve svojih deležnikov, ki so v večini ustvarjalci podatkov ter uporabniki. OAIS mora spremljati spremembe njihovih potreb ter tehnologijo, ki je na voljo. Pojavljajo se potrebe po novih formatih podatkov, novejši izbiri medijev, programski opremi ali novih platformah, novih mehanizmih za komuniciranje z arhivom. Izdelajo se poročila ter opozorila novih zahtevah in standardih.

Spremljajo se vzhajajoče nove tehnologije, informacijski standardi in računalniške platforme (hardver in programje), ki lahko povzročijo zastarelost obstoječih tehnologij vključenih v okolje arhiva, kar bi posledično lahko onemogočilo dostop do nekaterih vsebin arhiva. Za ocenjevanje pojavljajočih se novih tehnologij lahko uporablja prototipe.

Ta entiteta je z razvijanjem strategij in standardov odgovorna za razvoj in predloge novih strategij in standardov, ki bi arhivu omogočili lažje napovedovanje sprememb v prihodnosti. Gre za spremembe v zahtevah okolja in tehnoloških trendih, ki bi potrebovali migracijo nekaterih shranjenih vsebin ali pa zahtevali nove podatke. Ti predlogi so poslani v odobritev Administraciji.

⁷ Gre za razne politike delovanja sistema kot je npr. politika upravljanja s shranjeno vsebino, politika migracije (ta skrbi, da sistem ne zastari), politika upravljanja z bazo opisnih podatkov, razne varnostne politike itd.

V primeru odobritve novih standardov in migracijskih ciljev so razviti novi informacijski paketi, detajlni migracijski plani ter prototipi. Sprejeti standardi vključujejo standardizirane formate, metapodatkovne standarde ter dokumentacijske standarde. Na podlagi teh so izdelane nove šablone (angl. *Templates*) za AIP in SIP. Sprejeti migracijski cilji pomenijo transformacije AIP-jev vključno z njihovo vsebino, da bi se izognili nedostopnosti teh zaradi tehnološkega zaostanka. Tako ti vključujejo razvoj novih AIP oblik, prototipnih programov, testiranj, plan implementacij. V proces je ponavadi vključeno tudi ekspertno znanje drugih področij.

Potem, ko je migracijski načrt zaključen in so oblikovani novi AIP-ji, entiteta pošlje migracijske pakete v Administracijo, ki določi in izvede migracijo.

Dostop

Entiteta Dostop ima v sistemu arhiva zelo pomembno funkcijo, saj predstavlja uporabniški vmesnik, preko katerega uporabniki iščejo in naročajo zeleno gradivo. Do uporabniškega vmesnika se običajno dostopa preko računalniške mreže ali *dial-up* povezave (*on-line* storitve), naročilo pa je možno izvesti tudi preko katalogov, faksa, telefona ipd. Ločimo tri vrste zahtev: poizvedbe, ki so izvedene preko entitete upravljanja podatkov, te pa nemudoma vrnejo rezultat; poročila, ki lahko vključujejo več poizvedb, rezultat pa je oblikovano poročilo; naročila, ki z dostopom do baze podatkov in arhiva vrnejo stranki diseminacijski informacijski paket (DIP). Preveri se, če so zahtevani podatki na voljo, če je uporabnik avtoriziran za dostop in prejem, in obvesti stranko, ali je bilo naročilo odobreno ali zavrnjeno.

Entiteta Dostop prejme zahtevo po diseminaciji, zbere AIP iz arhiva in jo pošlje v začasni prostor za nadaljnje procesiranje. Iz entitete Upravljanje podatkov dobi opisne informacije, ki pripadajo zbranemu AIP in ki so potrebne za generiranje DIP.

Možna je tako *on-line* kot *off-line* dostava vseh zahtev (DIP, rezultati poizvedb, poročila in podpora) uporabniku. Informacija o finančnih terjatvah do kupca je poslana Administraciji.

OAIS se pri problemu dolgoročnega hranjenja informacij ne osredotoča le na metodo migracije, ki je v povezavi z njim največkrat omenjena. Odprt je tudi za druge strategije dolgoročnega arhiviranja, čeprav jih v svojem osnutku predloga ne omenja. V nadaljevanju se zato ne bom omejil le na omenjeno strategijo, pač pa bom opisal tudi druge, ki so danes v uporabi.

6. DOLGOROČNO HRANJENJE INFORMACIJ

Ohranjanje informacij je ena najbolj pomembnih tem v zgodovini človeka, kulturi in ekonomiji, poleg tega pa ima tudi velik vpliv na razvoj naše civilizacije. Prvotni zapisi so bili hranjeni na kamnih, keramiki, lesu. Razvoj civilizacije je iznašel nove medije za hranjenje informacij ter nove tehnike pisanja nanje. Pojavilo se je pisanje na svilo in papir. Do revolucionarnih sprememb na področju hranjenja informacij pa je ponovno prišlo z iznajdbo elektronskih medijev za shranjevanje digitalnih informacij.

Kakšne prednosti prinašajo digitalne informacije, smo že spoznali in zato se zdi, da je uvajanje teh neizogibno. Prav zato se danes v procesu digitalizacije nahaja mnogo podjetij ter drugih institucij kot npr. knjižnice ali razni arhivi idr. Ostaja pa dejstvo, da nekatera ključna vprašanja, povezana z dolgoročnim hranjenem digitalnih vsebin, še zdaleč niso rešena.

V ospredju sta predvsem dva: neobstojnost digitalnih medijev, na katerih hranimo digitalne podatke (v primerjavi s tradicionalnim papirjem) ter hitrost tehnološkega razvoja, ki povzroči zastarelost strojne in programske opreme in s tem dokumentov. V nasprotju z zapisi na tradicionalnih medijih (papir), za razumevanje katerih ni potrebna posebna programska oprema, pač pa le uporaba naših čutov, zapisi na novejših medijih ne morejo ostati nedotaknjeni desetletja ali stoletja in biti še naprej berljivi. Potrebne so prilagoditve opreme novim zahtevam okolja.

Poleg tehničnih vprašanj je treba opozoriti tudi na organizacijske in pravne probleme hranjenja informacij. Predvsem gre tu za problem ohranjanja avtentičnosti (verodostojnosti) in zagotavljanja integritete (neokrnjenosti) informacij. Potrebe po dokazovanju avtentičnosti in integritete naraščajo zaradi velike odgovornosti do digitalnega objekta, saj dokument zelo lahko kopiramo ali spreminjamo in tako ne predstavlja več originala.

V zvezi z dolgoročnim hranjenjem digitalnih dokumentov je danes aktualnih pet strategij oz. tehnik, ki ne smejo biti obravnavane izključno kot vzajemno izključljive, pač pa lahko nastopajo tudi v povezavi druga z drugo. To so (Lee et al., 2002, str. 93-106):

- papirna kopija in mikrofilm,
- tehnološki muzeji,
- migracija,
- emulacija in
- enkapsulacija.

Metodi papirne kopije ter tehnoloških muzejev nista najboljše metodi za dolgoročno arhiviranje digitalnih dokumentov, zato jih bom predstavil bolj površno. Bolj podrobno pa

bom predstavil tri različne pristope. To so migracija, emulacija in enkapsulacija. Vsaka od treh tehnik dolgoročnega arhiviranja ima veliko možnost, da bo izbrana kot najbolj primerna.

Pomembno je, da razumemo, da teh tehnik ne moremo neposredno primerjati med seboj, saj jih različni uporabniki v različnih primerih vrednotijo različno. Vse je odvisno od tega, kakšne zahteve imajo glede arhiviranja. Arhivi, ki imajo nalogo kratko- oz. srednjeročno hraniti preprostejše dokumente (npr. preprosta tekstovna oblika), bodo izbrali metodo migracije. Arhivi, katerih cilj je zagotoviti popolnoma identično kopijo originalu, ki je še vedno v preprosti obliki, bodo izbrali strategijo enkapsulacije, kompleksni dokumenti pa bodo najverjetneje arhivirani z metodo emulacije (YEA - The Yale Electronic Archive, 2002, str. 23).

6.1 Papirna kopija in mikrofilm

Strategija papirne kopije zagovarja nizkotehnične rešitve (*»low tech«*) hranjenja dokumentov na tradicionalnih produktih, kot sta papir ali mikrofilm. Ti zagotavljajo relativno dolgo življenjsko dobo arhiviranih zapisov, saj še vedno predstavljajo najobstojnejši medij ob določenih pogojih, čeprav tudi ta še zdaleč ni trajen. Strategija ima dve različici: prva je kopija dokumenta na papir ali mikrofilm, druga pa kopija oz. tiskanje bitnega niza v obliki (0, 1). Kopija na papir poleg ohranjanja vsebine in oblike omogoča tudi hranjenje organizacije dokumenta (YEA - The Yale Electronic Archive, 2002, str. 22).

Po drugi strani pa hranjenje informacij na papirju vodi k veliki izgubi funkcionalnosti v primerjavi z digitalnimi oblikami dokumentov, saj ne omogoča tiskanja dinamičnih oz. multimedijskih vsebin, kot so zvok, film, dinamične spletne strani itd. Ti izgubijo tudi vse interaktivne lastnosti v smislu hipertekstnih povezav in izgubijo lastnosti, ki izhajajo iz računalniške čitljivosti (npr. indeksacija). Poleg tega je poznejša restavracija v digitalno obliko praktično nemogoča, saj danes še ni tehnologije, ki bi preslikala analogni dokument v digitalno obliko s 100 % natančnostjo. Takšna strategija ima predvsem kratkoročni pomen, kar pomeni, da je na ta način primerno reševati tiste digitalne dokumente, ki so v neposredni nevarnosti, da bodo za ciljnega uporabnika postali nečitljivi in nedostopni v zelo kratkem času.

6.2 Tehnološki muzeji

Ideja pristopa tehnoloških muzejev je ustanovitev in vzdrževanje centrov, kjer bi bila za vsako elektronsko publikacijo nameščena izvirna strojna in programska oprema. Prednost tega pristopa je v tem, da lahko edino z originalno programsko opremo hranjeni dokumenti ohranijo vso funkcionalnost.

Ta metoda je - tako kot papirna kopija - nepraktična, saj je vzdrževanje delujočega tehnološkega sistema dolgoročno neizvedljivo. Ta metoda zahteva vzdrževanje računalniških

sistemov, ki z leti zastari. Vzdrževanje je otežkočeno, potrebovali bi stalno zalogo rezervnih delov in arhivskih medijev, ki niso več v proizvodnji. Četudi bi te imeli, bi se postavljal vprašanje primerne diagnostične in testne opreme, specifikacij komponent itd. Tako bi lahko zastarele medije, ki jim je potekla življenjska doba, zamenjevali le z novejšimi mediji, za katere pa star računalnik nima ustreznih vmesnikov ter programske opreme za pravilno delovanje. Za vsako generacijo programske opreme bi potrebovali kompletno opremo in strategijo vzdrževanja (YEA - The Yale Electronic Archive, 2002, str. 22).

Poleg teh problemov pa bi bilo takih centrov s tehnološkimi muzeji le nekaj v svetu, kar nasprotuje zahtevi arhiviranih elektronskih dokumentov, da naj bi ti zagotavljali stalno dostopnost. V takem muzejskem okolju bi lahko resnično dvomili o dolgoročnem obstoju arhiviranih vsebin in dostopa do njih. Vsekakor bi lažje govorili o muzeju starih računalnikov kot pa o arhivu digitalnih gradiv.

6.3 Migracija

Migracija je periodičen prenos digitalnih gradiv iz enega strojnega okolja na drugo oz. iz ene generacije računalniške tehnologije na kasnejšo (Garret, Waters, 1996, str. 17). Namen migracije je ohranjanje celovitosti digitalnih objektov in hkrati omogočanje, da so ti na voljo uporabnikom kljub neprestanemu spreminjanju tehnologije. Današnji mediji, ki hranijo digitalne podatke, lahko delujejo največ nekaj let, preden verjetnost nepopravljivih izgub podatkov postane prevelika, da je ne bi upoštevali (Reference Model for an Open Archival Information System (OAIS), 2002, str. 5-3).

Poznamo več strategij migracij za prenos digitalnih dokumentov iz zastarelih sistemov na novejša. Nobena strategija ne ustreza vsem formatom digitalnih informacij in nobena od obstoječih metod hranjenja ni popolnoma zadovoljiva. Migracijske strategije in z njimi povezani stroški variirajo glede na različna aplikacijska okolja, različne formate digitalnih dokumentov in različno stopnjo razvitosti uporabljene tehnologije. Bolj so uveljavljene metode migriranja digitalnih informacij v relativno preprostih datotekah, medtem ko bo treba migracijam kompleksnih digitalnih objektov nameniti še veliko dodatnih raziskav, ki bodo testirale tehnične zmožnosti posameznih pristopov, določile s tem povezane stroške in uveljavile najboljšo prakso.

Ena od tehnik migracije je tehnika osveževanja (angl. *Refreshment*), ki enostavno na nov medij kopira obstoječe podatke v enako obliko. Ta strategija lahko premesti težave nestabilnosti medijev, vendar ponavadi ni dovolj, da bi odpravila problem tehnološke zastarelosti. Migracija na nova tehnološka okolja pogosto pomeni, da kopija ne bo popolnoma identična njeni originalni informaciji. Uspešna migracija bo vsebovala informacije (metapodatke) o spremembah digitalnih objektov.

Kompleksnost migracijskega procesa je odvisna od narave digitalnih virov, ki variira od najpreprostejšega teksta do interaktivnih multimedijskih vsebin.

Najpreprostejša je migracija kot kopiranje digitalne informacije na stabilnejši nedigitalni medij kot npr. papir ali mikrofilm. V tej obliki informacije sicer zagotavljajo boljšo dolgoročno zanesljivost hranjenja, vendar po drugi strani ne predstavijo informacije v pravi obliki in tako izgubijo funkcionalnost ter izgled izvornega digitalnega objekta. Prenos na stabilnejši digitalni medij (npr. CD-R) pomeni kratko- do srednjeročno rešitev hranjenja in omogočanja dostopa do hranjenih informacij, saj še vedno zahteva migracijo, ko pride do tehnoloških sprememb (Migration. PADI – Preserving Acces to Digital Information. <http://www.ula.gov.au/padi/topics/topic21.html>, 13.2.2003).

Druga rešitev - programska oprema - ki je vzratno kompatibilna, lahko poenostavi migracijo. Zadnje verzije najpopularnejših urejevalnikov besedil na primer omogočajo dekodiranje starih formatov v nove (npr. dokument iz aplikacije Word 6 v Word XP), vendar ta strategija ni primerna za dolgoročno hranjenje ali za bolj kompleksne digitalne vire, ki niso bili ustvarjeni na eni od široko sprejetih aplikacij.

Naslednji pristop hranjenja in omogočanja dostopa do digitalnih informacij vključuje migracijo številnih formatov v nove, standardizirane. Ti so bolj stabilni kot nestandardne oblike podatkov, še vedno pa se imajo možnost preoblikovati v izvorni dokument. Odločitev, v kakšen format naj se dokument transformira, je odvisna od strukture digitalnega vira, ciljev naročnika ter potreb uporabnikov. Na primer, ena od pomembnih odločitev je, ali dati večji pomen ohranjanju zmožnosti procesa transformacije ali ohranjanju oblike ter vizualne prezentacije digitalne informacije. Pri tem pristopu je treba upoštevati, da se tehnološki standardi hitro razvijajo, zato hranjenje digitalnih vsebin ne more temeljiti le na tej strategiji. Pred kratkim je bil na področju digitalnega arhiviranja in elektronskega poslovanja široko sprejet nov univerzalni standardni format XML. Pričakuje se, da bo bistveno pripomogel k dolgoročnemu hranjenju digitalnih dokumentov tudi v procesu migracije.

OAIS pozna naslednje štiri vrste strategij migracij: (Reference Model for an Open Archival Information System (OAIS), 2002, str 5-4)

- **Osveževanje (angl. *Refreshment*):** Je digitalna migracija, ko obstoječi medij zaradi nevarnosti izgub podatkov zamenjamo z medijem istega tipa. Vsebina (eden ali več AIP-jev ali deli teh) je v nespremenjeni obliki prekopirana na nov medij.
- **Podvajanje (angl. *Replication*):** Je digitalna migracija, kjer pri prenosu podatkov na nov medij ne prihaja do sprememb v Vsebinskih informacijah, Informacijah o paketiranju ter PDI. Prenešeni biti, ki opisujejo informacijski objekt, se v nespremenjeni obliki prenesejo na enak ali nov tip medija. Osveževanje je tudi lahko

Podvajanje, vendar lahko slednje zahteva spremembe v Infrastrukturi označevanja skladišča⁸.

- **Prepakotiranje (angl. *Repackaging*):** Migracija vsebuje prepakotiranje, ko med prenosom Vsebine in PDI na nov medij prihaja do sprememb v Informacijah o paketiranju. Ta ima glavno vlogo v razmejevanju in povezovanju Vsebine in PDI.
- **Transformacija (angl. *Transformation*):** S splošno uporabo pojma migracija je pogosto enačena tehnika transformacije, čeprav je ta le ena od tehnik migracije. Transformacija je digitalna migracija, pri kateri prihaja do sprememb v nizu bitov, ki opisujejo Vsebino ali PDI, v procesu prenosa teh na nov medij oz. v novo tehnološko okolje. S tem se spremeni tudi Opisna informacija. V vseh primerih migracij je trud namenjen maksimalnemu ohranjanju informacije. Novi AIP je v večini primerov tako popolna zamenjava za transformirani AIP. O novem AIP-ju lahko govorimo tudi kot o novi verziji prejšnjega AIP. Prvotna verzija je obravnavana kot original in je pogosto obdržana kot verifikacija procesa ohranjanja informacij.

Od naštetih podtehniki migracije najmanjše tveganje izgube informacij predstavlja proces osveževanja ter imitacije, saj se niz bitov, ki določajo AIP, ne spremeni. Pri prepakotiranju bo sicer prišlo do sprememb v zaporedju bitov, vendar samo v Informacijah o paketiranju, ki pa ne vplivajo na spremembe hranjenih digitalnih virov (Vsebina ter PDI). Transformacija predstavlja največje tveganje, kajti spremenjena sta tako Vsebina kot PDI.

Po Rothenbergu (1998) je proces migracije zelo tvegan v smislu izgube podatkov. Migracija je lahko v kompleksnih primerih (kompleksni multimedijски dokumenti) zelo zamuden in drag proces. Poleg tega ni očitno, kako oz. ali sploh lahko migriramo vse vrste podatkov (npr. računalniški programi). Tukaj pridejo na vrsto zagovorniki emulacije, po mnenju katerih naj bi v določenih situacijah ta zagotavljala boljšo rešitev dolgoročnega hranjenja podatkov.

6.4 Emulacija

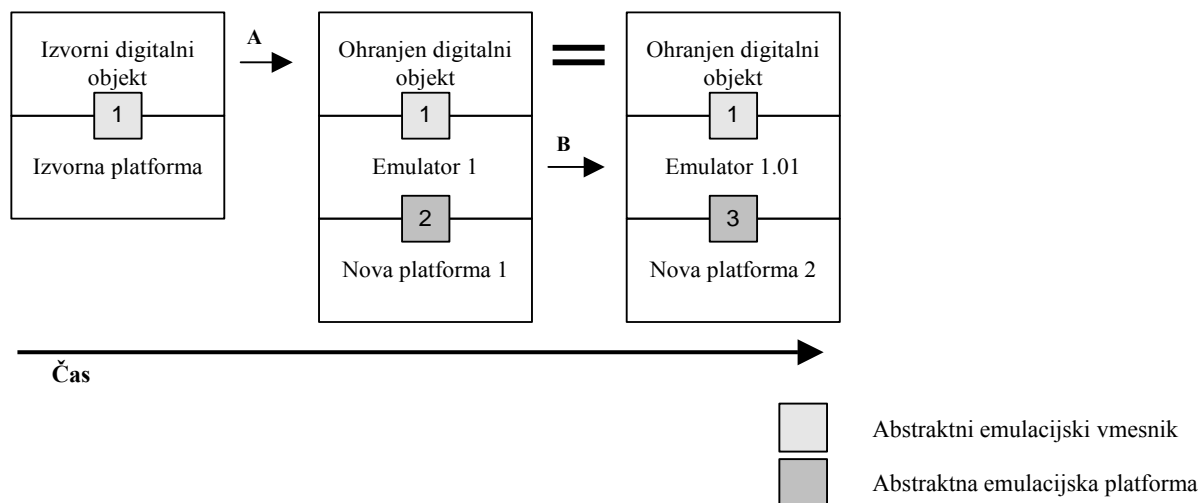
Emulacija, katere velik zagovornik je Rothenberg (2000), je dobra alternativa tehniki tehnoloških muzejev, saj je za organizacije skoraj nemogoče ohraniti delujoče primere računalnikov ter pripadajoče programske opreme, da bi zagotovili dostop do informacij, saj tega ne dovojujejo visoki stroški. Poleg tega pa tako kot tehnološki muzeji, tudi emulacija ohranja izvirno programsko opremo za dostopanje do hranjenih dokumentov. Poenostavljeno bi lahko idejo emulacije prikazali s Sliko 4.

Emulacija s pomočjo emulatorjev⁹ poustvarja staro tehnično okolje na trenutni ali prihodnji tehnologiji oz. platformi, ki omogoča uporabo oz. prikaz digitalnih dokumentov iz preteklih

⁸ Infrastruktura načrtovanja arhivskega skladišča (angl. *Archival Storage mapping Infrastructure*) označuje položaj AIP-ja, shranjenega na mediju, glede na njegovo identifikacijsko oznako (AIP ID), ter omogoča dostop do tega.

časovnih obdobj. Z drugimi besedami, gre za poustvarjanje delovanja zastarelega računalnika in operacijskega sistema na današnjem ali bodočem operacijskem sistemu, ki omogoča zagon zastarele originalne opreme in s tem prikazovanje starih elektronskih dokumentov, ki jih sicer v novo okolje ne bi bilo možno priklicati. Gre za kreiranje nove programske opreme, ki posnema starejše programske ali strojne specifikacije.

Slika 4: Pretvorba originalnega digitalnega objekta v novi objekt na novi platformi



Vir: Holdsworth, Wheatley, 2003, str. 3.

Zgornji diagram prikazuje transformacijo (puščica A) izvornega digitalnega objekta v ohranjen digitalni objekt, ki nam v emuliranem okolju omogoči prikaz lastnosti, kot jih je imel original. Ta korak omogoča neoviran dostop do zapisov, ne glede na to, da je prvotna platforma, na kateri je izvirni dokument nastal, že zastarela. Sčasoma bo zastarela tudi sedanja platforma (platforma 1). Zato bo potrebna transformacija, ki jo na diagramu prikazuje puščica B. Zastarelost platforme 1 bo sprožila potrebo po posodobitvi emulatorja tako, da bo lahko deloval na novi platformi (platformi 2). V tem primeru ostane niz bitov digitalnega objekta nespremenjen, kar je na sliki prikazano kot enačaja.

V zadnjem času se pojavljajo tudi alternativne rešitve emulaciji. Primer sta npr. razvoj »emulacijskega navideznega stroja« (angl. *Emulation Virtual Machine*), katerega avtor je Jeff Rothenberg, Raymond Lorie je razvil univerzalno navidezni stroj (angl. *Universal Virtual Machine - UVM*), ki omogoča implementacijo emulacije idr. (Holdsworth, Wheatley, 2003, str. 4).

⁹ Emulator je opis lastnosti izvorne strojne opreme, na kateri je tekel operacijski sistem - dovolj natančen, da lahko novi operacijski sistem »poustvari« delovanje starega računalnika. Emulator ne more biti program, ker ne moremo vnaprej poznati lastnosti bodočih operacijskih sistemov, na katerih bi tekel (Dimec, 2000).

Ena od prednosti emulacije pred migracijo je, da izvirne informacije ne zastarijo. Namesto teh se spreminja emulacija tehnološkega okolja. Zaradi tega ta strategija omogoča ohranitev funkcionalnosti ter izvirnega izgleda (*look and feel*) digitalnih informacij, ki bi se drugače izgubila znotraj tehnološke zastarelosti. V primerjavi z drugimi prezervacijskimi tehnikami je emulacija primerna za dolgoročneje hranjenje bolj kompleksnih digitalnih dokumentov. Naslednja prednost implementacije emulacije pa je učinkovitost, saj ko so podatki enkrat shranjeni skupaj z ustreznimi metapodatki in programsko opremo, ni potrebna nobena druga akcija, razen preprostega kopiranja na nove medije, dokler ni podana zahteva po hranjenih objektih. En emulator lahko uporabimo kot rešitev za nekaj podatkovnih objektov, ki za prikaz potrebujejo enako operativno okolje.

Kljub svojim prednostim emulacija ne more in ne sme biti prva in zadnja izbira procesa dolgoročnega hranjenja digitalnih dokumentov. Nanjo lahko gledamo kot na del orožja, potrebnega za zaščito digitalne zapuščine pred uničevalnim časom - v svetu, kjer so inovacije in zato spremembe visoko cenjene. Omeniti pa velja, da obstaja veliko standardnih dokumentov, ki bodo s standardno programsko in strojno opremo ter s pomočjo drugih prezervacijskih tehnik uporabni še dolgo v prihodnosti.

6.5 Enkapsulacija

To metodo je v večjem delu razvila organizacija CEDARS (Lee et al., 2002, str. 99) in temelji na OAIS metodologiji. V CEDARS predvidevajo, da bo proces migracije datotek pomenil izgubo originalnih lastnosti digitalnega objekta. Prepričani so, da je varneje hraniti preprost objekt kot niz bitov neodvisno od medija, na katerem je shranjen. Zato izvirno programsko opremo lahko zamenja nova, ki bo prav tako lahko prebrala ta niz bitov.

Enkapsulacija je v kontekstu ohranjanja digitalnih vsebin tehnika, ki združuje digitalni objekt in vse potrebne informacije, ki so potrebne za dostop do tega objekta. S pomočjo vseh teh zajetih podatkov je metoda sposobna ustvariti originalno aplikacijo, preko katere je uporabnik prvotno dostopal do hranjene informacije, na neki novi platformi. Enkapsulacija je lahko dosežena z uporabo fizičnih ali logičnih struktur, ki jih poimenujemo »kontejnerji«, ki zagotovijo povezavo med vsemi informacijami, kot so digitalni objekt in podperne informacije vključno z metapodatki.

Referenčni model OAIS opisuje tipe podpornih informacij, ki naj bi bile vključene v enkapsulacijo. To so opisne informacije (*Representation information*), katerih naloga je, da pravilno interpretirajo bite ter informacije, ki so vključene v PDI (izvor, kontekst, vir, stabilnost).

OAIS razlaga metodo enkapsulacije s konceptom informacijskih paketov (IP so v našem primeru kontejnerji), ki zajemajo Vsebino (*Content information*) ter PDI (*Preservation*

description information), ki jih povezuje Informacija o paketiranju (*Packaging information*). Tipi informacij so prikazani na Sliki 2.

Ta metoda je znana tudi kot »migracija na zahtevo« (angl. *Migration on Request*). Ko uporabnik zahteva digitalni objekt iz arhiva, se avtomatično izvrši migracija niza bitov. V bistvu se ločita vsebina in format. Prav zato pa lahko objekt zagotavlja avtentičnost, saj je vedno shranjen v obliki izvornega niza bitov.

Do neke mere lahko ta proces obravnavamo kot kombinacijo emulacije z migracijo, kajti treba je hraniti tudi orodje, ki pretvori bite v dokument. Orodje, ki je opisano v metapodatkih, je lahko programska specifikacija ali emulator, ki oponaša strojno ali programsko okolje. Figurativno bi proces lahko opisali kot »počitek« digitalnega objekta v obliki bitov do trenutka, ko ga na uporabnikovo zahtevo preoblikuje programska oprema.

Primeren jezik za shranjevanje metapodatkov in navodil za prezerviranje digitalnih objektov je jezik XML. Enkapsulacija ga uporablja kot ogrodje, na katerega »obesi« dokumente ali dele dokumentov. V tem primeru XML tvori hrbtenico digitalnega arhiviranega zapisa (XML and Digital Preservation, 2002, str. 27).

Velik zagovornik enkapsulacije (in tudi emulacije) je Rothenberg. Na drugi strani Bearman oporeka njegovi teoriji z razlago, da ni povsem jasno, kako bi strategijo metapodatkovne enkapsulacije lahko izvedli tudi v praksi (Rothenberg, 1999).

Ena od metod, ki je bila razvita za namen dolgoročnega ohranjanja informacij in ki temelji na teoriji enkapsulacije, je metoda *Universal Preservation Format* (UPF) (Lee et al., 2002, str. 98). Je tehnologija hranjenja, ki deluje na samopopisnem principu. Ta vključuje kontejner, ki hrani digitalni objekt ter metapodatke skupaj, da jih obvaruje pred tehnološko zastarelostjo. Druga, »*The Digital Rosetta Stone*«, je podobna metoda, ki shranjuje Opisne informacije, da interpretirajo digitalno vsebino objekta zunaj enkapsulacije, da ne pride do podvajanj podatkov in s tem neučinkovite zasedenosti hrambenega prostora.

7. ANALIZA TEHNIK ARHIVIRANJA DOKUMENTOV

Do sedaj sem v nalogi predstavil pomen dokumentov, arhivov, omenil slovensko zakonodajo v zvezi z arhiviranjem ter standarde, ki se pojavljajo na tem področju. Predstavil sem model odprtega informacijskega sistema za arhiviranje (OAIS) kot enega potencialnih standardov, ki bi lahko omogočil nadaljnji razvoj arhivov za hranjenje digitalnih informacij na dolgi rok. Opisal sem tudi problem dolgoročnega arhiviranja ter pet aktualnih tehnik, ki delno rešujejo ta problem. V nadaljevanju bom naredil analizo teh tehnik in s pomočjo ocen posameznih

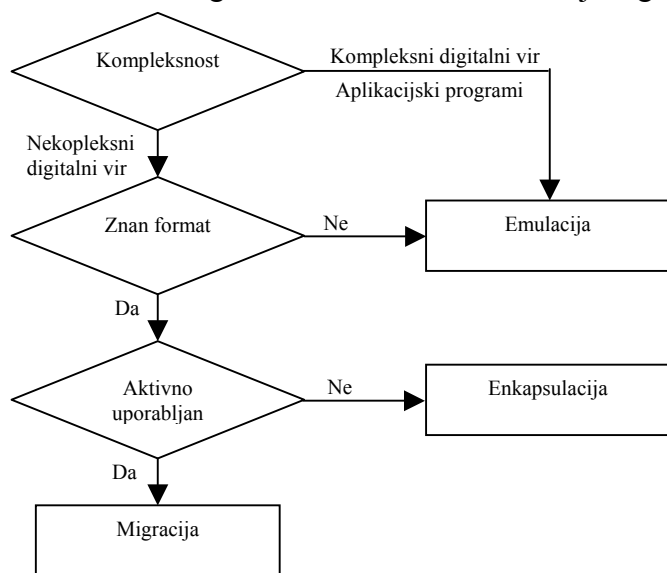
kriterijev poskusil ugotoviti, kateri arhiv je najbolj uporaben za dolgoročno hranjenje digitalnih vsebin.

Mnenja strokovnjakov o tem, katera tehnika je bolj primerna za dolgoročno arhiviranje, se precej razlikujejo. Mnogokrat se zdi, da naklonjenost eni ali drugi tehniki temelji na načelu »vsak hvali svoje«. Pomembno je, da razumemo, da teh metod ne moremo popolnoma primerjati med seboj na podlagi dobrih in slabih strani, saj je treba upoštevati, da so nekatere bolj primerne za hranjenje ene vrste digitalnih dokumentov, medtem ko so nekatere boljše za druge. Lahko bi rekli, da nimajo skupnega imenovalca. Na podlagi njihovih predstavitev v prejšnjem poglavju lahko sedaj opredelimo, v kakšnih pogojih je posamezna metoda najbolj uporabna.

V primeru hranjenja kompleksnih in aplikacijskih programov, kot so npr. igre ali izvršilne datoteke, je najbolj primerna metoda emulacije. Poleg tega je ta metoda primerna tudi za arhiviranje digitalnih formatov, ki niso pogosto na voljo in o katerih nimamo veliko znanja (nestandardni). Ustrezna je tudi v primerih, ko je pomembno, da ohranimo izvorni izgled digitalnih informacij. Migracija in enkapsulacija sta bolj primerni za arhiviranje digitalnih podatkov, katerih format poznamo dovolj dobro. Bolj primerni sta za shranjevanje preprostejših digitalnih virov. Migracija je bolj primerna za vire, do katerih dostopamo pogosto in so pogosto v uporabi. V nasprotju s tem pa je enkapsulacija bolj primerna za tiste vire, ki niso pogosto v uporabi.

Način izbire primerne tehnike shranjevanja digitalnih informacij glede na tip in kompleksnost digitalne informacije, dostopnost do formata ter njihove uporabe želim prikazati s Sliko 5.

Slika 5: Shematični diagram izbire tehnike arhiviranja digitalnih informacij



Vir: Lee et al., 2002, str. 103.

7.1 Prednosti in slabosti tehnik digitalnega arhiviranja

V nadaljevanju bom v tabelah od 1 do 5 predstavil prednosti in slabosti strategij dolgoročnega hranjenja digitalnih dokumentov. Ta primerjava bo kasneje služila za izbiro najbolj perspektivne tehnike arhiviranja, ki jo bom s pomočjo ocene ustreznosti glede na postavljene kriterije poskusil izbrati ob določenih predpostavkah.

Tabela 1: Prednosti in slabosti papirne tehnike ter mikrofilma

Prednosti	Slabosti
<ul style="list-style-type: none"> Ni več tako zelo občutljiva na tehnološko zastarelost, če predvidevamo, da shranjujemo na kvaliteten mikrofilm ali trajni papir Le enkratni strošek za konverzijo Zagotavlja uporabnost informacij stoletja dolgo, pod predpostavko, da je konvertiran v arhivsko standardno obliko Je dobra (pragmatična) začasna rešitev, dokler ni razvita bolj primerna strategija digitalnega arhiviranja 	<ul style="list-style-type: none"> Izgubi funkcionalnost izvirnega digitalnega vira Je razumna opcija le za digitalne vire, ki ne potrebujejo ohranitve popolne funkcionalnosti digitalne tehnologije Povzroča težave, četudi je uporabljena le preprosta elektronska pošta Ni primerna za kompleksnejše digitalne vire, kjer izguba funkcionalnosti zmanjša, če ne celo uniči, uporabnost ter celovitost virov Izguba prednosti digitalne tehnologije kot npr. udobnost uporabe, učinkovita uporaba prostora, ... Vse večji stroški konverzije v standardiziran format in vse večji stroški zasedenega prostora (kumulativni stroški bodo čez čas zelo pomembni)

Vir: Lee et al., 2002, str. 93-106.

Tabela 2: Prednosti in slabosti tehnoloških muzejev

Prednosti	Slabosti
<ul style="list-style-type: none"> Skladiščenje ohrani funkcionalnost ter izgled originala Skladiščenje odloži čas, potreben za uvedbo novih strategij arhiviranja Skladiščenje je lahko najbolj praktična vmesna strategija za kompleksne digitalne vire 	<ul style="list-style-type: none"> Uporabljamo jo lahko le kot kratko- do srednjeročno strategijo Tehnična podpora bo neizogibno ukinjena v relativno kratkem času Dostop bo s časom postajal vse bolj problematičen

Vir: Lee et al., 2002, str. 93-106.

Tabela 3: Prednosti in slabosti migracije

Prednosti	Slabosti
<ul style="list-style-type: none"> • Dobro uveljavljeni postopki za preprosto migracijo • Trenutno najbolj uporabljana strategija v večini digitalnih arhivov • Z razvojem tehnologije in novimi platformami se bo postopek migracije poenostavil • Ni potrebno ohranjanje izvirnih aplikacij • Podpora aktivnemu dostopu in upravljanju arhiviranih informacij • Prenos digitalnih informacij iz manj stabilnega na bolj stabilen medij 	<ul style="list-style-type: none"> • Pri zahtevnih migracijah so potrebni posebni programi, zaradi česar lahko stroški postanejo enormni • Migracija je lahko zelo zamudna in zelo kompleksna • Verjetna delna izguba funkcionalnosti ter izgleda (<i>look and feel</i>) originala • Visoki stroški pri dolgoročnem shranjevanju • Kompromis s celovitostjo originalov, razen če so zahtevani strogi kontrolni postopki • Manj primerna za dolgoročno hranjenje • Bolj kompleksne digitalne vire lahko migriramo z veliko izgubo na funkcionalnosti • Migracija mora potekati v rednih časovnih razmikih skozi življenjski cikel digitalnega vira • Pomanjkanje metapodatkov o arhiviranju

Vir: Lee et al., 2002, str. 93-106.

Tabela 4: Prednosti in slabosti emulacije

Prednosti	Slabosti
<ul style="list-style-type: none"> • Ponovno pridobljena funkcionalnost ter izgled izvirnega dokumenta • Izogibanje ponavljajočim se stroškom, ki nastajajo v primeru migracije • Zelo primerna za bolj kompleksne digitalne vire • Omogoča ohranjanje velike količine informacij 	<ul style="list-style-type: none"> • V fazi razvoja, zato je potrebnih še mnogo testiranj • Včasih lahko le delno emulira funkcionalnost in izgled originala • Zelo draga tehnika, razen če pride do ekonomij obsega. Potrebna je izdelava novih emulatorjev za nova okolja. Ti stroški lahko presežejo ponavljajoče se migracijske stroške • Stroge zahteve pri izdelovanju dokumentacije za staro strojno in programsko opremo (do take mere so bile le redko izdelane v preteklosti) • Precej tvegana strategija za dolgoročno hranjenje informacij (ni garancije za podroben opis, ki bi omogočil izdelavo emulatorja v prihodnosti)

Vir: Lee et al., 2002, str. 93-106.

Tabela 5: Prednosti in slabosti enkapsulacije

Prednosti	Slabosti
<ul style="list-style-type: none"> • Vse potrebne informacije za dostop se hranijo znotraj ene entitete • Lahko prekosi nekatere glavne pomanjkljivosti alternativnih strategij • Ohrani arhivirane informacije • Osredotočena na elemente, ki so potrebni za dostop 	<ul style="list-style-type: none"> • S podvajanjem lahko ustvari ogromne datoteke • Programska oprema, namenjena enkapsulaciji, je še naprej odprta za hitro tehnološko zastarelost • Znanje o formatu mora biti ohranjeno • Potrebni sistemi, ki zajamejo potrebne digitalne podatke • Ni še popolnoma jasno, kako praktično implementirati to strategijo

Vir: Lee et al., 2002, str. 93-106.

7.2 Izbira najustreznejše tehnike

Zanimiva bi bila ugotovitev, katera tehnika dolgoročnega arhiviranja ima trenutno največji potencial za uveljavitev kot standard. Izvedbo take ocene sem si zamislil tako, da najprej določim kriterije, po katerih bom ocenjeval vsako od v nalogi opisanih tehnik.

Podobno metodo izbire najbolj optimalne strategije opisuje tudi prof. Pučko v knjigi *Strateško upravljanje* (1996, str. 245). Omenja metodo ocenjevanja ustreznosti ideje za nov proizvod, ki se vsebinsko precej pokriva z mojim problemom.

Vsakemu kriteriju bom določil ponderje oz. uteži, ki opredeljujejo pomen posameznega kriterija v sklopu vseh kriterijev. Kriterije bom ocenil z oceno od 1 do 5, ki je relativna glede na druge omenjene tehnike arhiviranja. Ocena 5 pomeni, da tehnika najbolj izmed vseh izpolnjuje kriterij, medtem ko ocena ena pomeni, da tehnika najslabše izpolnjuje ta kriterij oz. ga sploh ne. Poleg teh pa so možne tudi vmesne ocene med 1 in 5. Vsak kriterij bom podprl z opombo, če bo ta potrebna. Končna ocena vsake od tehnik bo tako sestavljena iz vsote ocen vseh kriterijev, vsaka ocena bo pomnožena z ustrežno utežjo. Končna ocena bo variirala od 1 do 5. Višja ocena bo pomenila, da ima tehnika arhiviranja v povprečju boljšo uporabnost v primerjavi z drugimi tehnikami.

Vendar pa je na tem mestu nujno opozoriti, da končna ocena ne pomeni, da je ena tehnika na splošno bolj primerna od druge. Prav tukaj je problem ocenjevanja teh tehnik. Vsaka od tehnik ima svoje prednosti in slabosti, o katerih sem že govoril. Nekatere so bolj primerne v enih situacijah, druge pač v drugih (npr. za arhiviranje dokumentov v preprostem formatu ne bomo uporabili tehnike emulacije, čeprav bi ta lahko imela višjo končno oceno od papirne tehnike). Prav zato se v realnosti pogosto uporablja kombinacija več tehnik.

Kot odgovor na ta problem pa lahko iz vsake spodaj navedene tabele razberemo, kateri kriterij bolje zadovoljuje posamezna tehnika. To lahko služi tudi kot pomoč pri izbiri posamezne tehnike, če je za nas pomemben prav določen kriterij (npr. želimo izbrati tehniko, ki zahteva najnižje stroške).

7.2.1 Opis kriterijev

Kriteriji so izbrani na podlagi prebrane literature, vendar so kljub temu določeni po subjektivni presoji. Z njimi poskušam zajeti najpomembnejše lastnosti arhivov.

Integriteta in avtentičnost

Ne glede na metodo mora biti glavni cilj arhiviranja ohranitev integritete, ki definira in ohranja tiste funkcije digitalnega dokumenta, ki ga loči od drugih. V digitalnem okolju funkcijo integritete določajo kontekst, izvor, vir in stabilnost (glej OAIS). Z ohranjanjem teh kontekstualnih razmerij je zagotovljena tudi verodostojnost dokumenta. Mnenja o tem, ali mora hranjeni dokument sploh zagotavljati verodostojnost, se med strokovnjaki razlikujejo (Južnič, 2000, str. 80). V poslovnem svetu je to vsekakor nujnost, saj le tako lahko arhivi služijo svojemu namenu in omogočajo revizijsko varnost. Avtentičnost pa je pomembna tudi za zagotavljanje zgodovinske vrednosti arhivov in njihovih vsebin.

Tveganje izgub informacij

Vsaka pretvorba podatkov ima posledično svoje tveganje, da kopija ne bo več enaka originalu, ali da se zaradi kakršnekoli nepričakovane napake podatki izgubijo. Večje tveganje je pri nestandardiziranih, v praksi manj preizkušenih tehnikah.

Funkcionalnost

Funkcionalnost dokumenta je ohranjena, če dokument zagotavlja enake funkcije, kot jih je imel pred arhiviranjem. Za primer bi lahko dali spletno stran s svojimi povezavami. Funkcionalnost teh bo ohranjena, če bodo tudi v arhiviranem stanju ohranile zmožnost povezav.

Izvirni izgled (*look and feel*)

S funkcionalnostjo se dopolnjuje tudi izgled. Pomembna lastnost arhiva je, da omogoči enak izgled arhiviranega dokumenta, kot ga ima njegov izvirnik. Splošno uveljavljena standarda sta sicer TIFF in PDF format, ki sta namenjena statičnim dokumentom (Hodge, 2000).

Stroški (kratko-, srednje-, dolgoročni)

Stroški predstavljajo prav posebno poglavje znotraj opisanih kriterijev. Ti so glede na izbrano tehniko v času različni. K temu precej pripomorejo tudi ekonomije obsega. Iz tega razloga sem ločil tri vrste stroškov: kratkoročni, srednjeročni ter dolgoročni. Pri ocenjevanju tehnik

pa sem izračunal povprečne stroške vseh treh obdobj s preprosto aritmetično sredino¹⁰. Stroške je vnaprej zelo težko predvideti, saj ne moremo natanko vedeti, kdaj bo potreben prenos ali nova pretvorba podatkov.

Obdobje hranjenja

Pomembno za dolgodobne arhive je, da niso odvisni od uporabljene tehnologije, saj le v takem primeru lahko omogočamo hranjenje dokumentov na dolgi rok. Obdobje hranjenja pove, za koliko časa je posamezna tehnika sposobna ohranjati dokument, preden je potreben nov prenos ali pretvorba dokumentov¹¹.

Hranjenje kompleksnih formatov

Hranjenje kompleksnih formatov je sposobnost tehnike arhiviranja, da poleg preprostih dokumentov, kot so npr. dokumenti v tekstovni obliki, shranjuje tudi dokumente s kompleksnim formatom, kot so npr. multimedijски dokumenti ali programske izvorne kode itd.

Obvladovanje količine informacij

Vsak arhiv zaradi fizičnih ali kakšnih drugih ovir ni sposoben obvladovati velikih količin dokumentov. Preobilje informacij je eden bistvenih problemov sodobnega arhiva, zato ta kriterij ni nepomemben. Digitalni arhivi imajo pred tradicionalnimi v tem smislu veliko prednost.

Oskrbovanje arhiva

Oskrbovanje arhiva je kriterij, ki pove, kolikšna sredstva (delo, oprema, ...) so potrebna za ohranjanje delovanja arhiva ter kakšni stroški so s tem povezani. Zagotovo je treba tradicionalnim arhivom v primerjavi z modernim (digitalnim) nameniti veliko več časa in sredstev.

Prostor

Ta kriterij pove, koliko prostora je treba nameniti arhivom. Tukaj sicer ne moremo neposredno ločiti med tradicionalni arhivi, ki potrebujejo ogromne površine za oskrbo arhiviranih vsebin, ter digitalnimi arhivi, ki zahtevajo veliko prostora na digitalnem mediju. Posredno bi ta dva lahko primerjali s stroški.

¹⁰ Povprečni stroški so aritmetična sredina vseh treh vrst stroškov deljena s 3. Kratkoročno omenjene tehnike hranjenja ne predstavljajo velikih stroškov, saj še ni potrebe po hranjenju teh. Z daljšim obdobjem pa se ti spreminjajo različno glede na izbrano tehniko.

¹¹ Kratkoročno obdobje je obdobje od implementacije arhiva do pojavitve nove generacije tehnologije. Srednjeročno obdobje je obdobje od implementacije arhiva do zastarelosti obstoječe tehnologije. Dolgoročno obdobje je obdobje od zastarelosti obstoječe tehnologije dalje.

Preprostost izvedbe

Za mnoga podjetja ali kakšno drugo institucijo lahko ta kriterij predstavlja utež, ki lahko tehtnico odločitve o najbolj uporabni tehniki arhiviranja prevesi na eno ali drugo stran. Tukaj ocenjujem tako preprostost implementacije sistema kot tudi preprostost kasnejše uporabe.

7.2.2 Ocena tehnik

Z oceno tehnik torej želim izbrati tehniko arhiviranja, ki ima na splošno največjo uporabno vrednost. Končna ocena je vsota ponderiranih ocen, pri čemer je treba opozoriti, da so tabele narejene na podlagi osebnega razumevanja problema arhiviranja in so tako uteži kot ocene določene po lastni presoji.

Tabela 6: Ocena splošne uporabnosti papirne kopije ter mikrofilma kot tehnike digitalnega arhiviranja glede na kriterije

Kriterij	Ocena	Ponder	Ponderirana ocena	Opomba
Integriteta, avtentičnost	2	0,120	$2 \times 0,120 = 0,240$	Integriteta in avtentičnost niso vedno zagotovljene in se pogosto izgubita pri kompleksnejših dokumentih
Tveganje izgub informacij	4	0,110	0,440	Ob upoštevanju pravil in politike arhiviranja je tveganje majhno
Funkcionalnost	1	0,100	0,100	Izgubi funkcionalnost izvirnega digitalnega vira
Izvirni izgled (<i>look and feel</i>)	3	0,100	0,300	Mogoče ga je ohraniti le za statične dokumente
Stroški (kratko, srednje, dolgo)	(4,5,4) 4,3	0,090	0,390	Stroški se pojavljajo pri kopiranju dokumentov na »trajni« medij ter pri hranjenju medijev
Obdobje hranjenja	5	0,090	0,450	Papir ali mikrofilm zagotavljata dolgo življenjsko dobo arhiviranih zapisov
Hranjenje kompleksnih formatov	1	0,085	0,085	Hrani lahko le preproste oblike dokumentov
Obvladovanje količine informacij	2	0,085	0,170	Obvladujemo lahko relativno malo informacij zaradi problema prostora
Oskrbovanje	3	0,075	0,225	Pomembna pravilna oskrba arhivskega gradiva ter arhivskega prostora.
Prostor	1	0,075	0,075	Arhivi lahko zasedejo ogromne površine
Preprostost izvedbe	5	0,070	0,350	Najpreprostejši arhiv
SKUPAJ			2,825	

Tabela 7: Ocena splošne uporabnosti tehnoloških muzejev kot tehnike dolgoročnega arhiviranja glede na kriterije

Kriterij	Ocena	Ponder	Ponderirana ocena	Opomba
Integriteta, avtentičnost	5	0,120	0,600	Popolno ohranjanje kontekstualnih razmerij
Tveganje izgub informacij	2	0,110	0,220	Tehnološko zastarevanje opreme
Funkcionalnost	5	0,100	0,500	Popolno ohranjanje funkcionalnosti
Izvirni izgled (<i>look and feel</i>)	5	0,100	0,500	Popolno ohranjanje izvirnega izgleda
Stroški (kratko, srednje, dolgo)	(4,3,1) 2,7	0,090	0,240	Stroški z zastarelostjo opreme strmo naraščajo
Obdobje hranjenja	2	0,090	0,180	Kratko- do srednjeročna tehnika
Hranjenje kompleksnih formatov	5	0,085	0,425	Popolno ohranjanje kompleksnih formatov
Obvladovanje količine informacij	2	0,085	0,170	Obvladuje lahko majhno količino informacij zaradi neprestane zastarelosti opreme
Oskrbovanje	1	0,075	0,075	Nesmotno zaradi zastarevanja opreme
Prostor	2	0,075	0,150	Arhivi lahko zasedejo velike površine
Preprostost izvedbe	4	0,070	0,280	Sama izvedba ni zahtevna, če izvzamemo vzdrževanje
SKUPAJ		3,34		

Tabela 8: Ocena splošne uporabnosti migracije kot tehnike dolgoročnega arhiviranja glede na kriterije

Kriterij	Ocena	Ponder	Ponderirana ocena	Opomba
Integriteta, avtentičnost	5	0,120	0,600	Na zahtevo je možna ohranitev integritete in avtentičnosti
Tveganje izgub informacij	4	0,110	0,440	Tveganje izgub je majhno, sploh, če gre za standardizirane dokumente
Funkcionalnost	4	0,100	0,400	Funkcionalnost ni nujno 100 % ohranjena
Izvirni izgled (<i>look and feel</i>)	4	0,100	0,400	V večini primerov tega ohrani
Stroški (kratko, srednje, dolgo)	(4,3,2) 3	0,090	0,270	Pojavljajo se periodično, odvisni od kompleksnosti formata
Obdobje hranjenja	3	0,090	0,270	kratko- do srednjeročna tehnika, redko dogoročna
Hranjenje kompleksnih formatov	3	0,085	0,255	Ni najbolj primerna tehnika za hranjenje kompleksnih formatov
Obvladovanje količine informacij	4	0,085	0,340	Obvladuje lahko veliko količino podatkov, vendar ne toliko kot emulacija ali enkapsulacija
Oskrbovanje	4	0,075	0,300	Urejanje baze, oskrbovanje strojne opreme zanemarljivo
Prostor	4	0,075	0,300	Učinkovita uporaba prostora
Preprostost izvedbe	3	0,070	0,210	Zahtevnejše od prejšnjih dveh in manj zahtevno od preostalih dveh tehnik
SKUPAJ		3,785		

Tabela 9: Ocena splošne uporabnosti emulacije kot tehnike dolgoročnega arhiviranja glede na kriterije

Kriterij	Ocena	Ponder	Ponderirana ocena	Opomba
Integriteta, avtentičnost	5	0,120	0,600	Popolno ohranjanje kontekstualnih razmerij
Tveganje izgub informacij	2	0,110	0,220	zahtevna izvedba, nepoznavanje prihodnje tehnologije, nepreizkušeno
Funkcionalnost	5	0,100	0,500	Popolno ohranjanje funkcionalnosti
Izvirni izgled (<i>look and feel</i>)	5	0,100	0,500	Popolno ohranjanje izvirnega izgleda
Stroški (kratko, srednje, dolgo)	(4,2,1) 2,3	0,090	0,180	Zelo zahtevna implementacija predstavlja glavni strošek
Obdobje hranjenja	5	0,090	0,450	Dolgoročna tehnika
Hranjenje kompleksnih formatov	5	0,085	0,425	Namenjena predvsem kompleksnim nestandardnim podatkom
Obvladovanje količine informacij	5	0,085	0,425	Podatki ohranjajo izvirno obliko, zato lahko obvladuje veliko količino teh
Oskrbovanje	4	0,075	0,300	Oskrbovanje strojne opreme zanemarljivo
Prostor	4	0,075	0,300	Učinkovita uporaba prostora
Preprostost izvedbe	1	0,070	0,070	Zelo zahtevna implementacija
SKUPAJ		4		

Tabela 10: Ocena splošne uporabnosti enkapsulacije kot tehnike dolgoročnega arhiviranja glede na kriterije

Kriterij	Ocena	Ponder	Ponderirana ocena	Opomba
Integriteta, avtentičnost	5	0,120	$4 \times 0,120 = 0,480$	Popolno ohranjanje kontekstualnih razmerij
Tveganje izgub informacij	3	0,110	0,440	Zahtevna izvedba in nepreizkušeno
Funkcionalnost	5	0,100	0,100	Popolno ohranjanje funkcionalnosti
Izvirni izgled (<i>look and feel</i>)	5	0,100	0,300	Popolno ohranjanje izvirnega izgleda
Stroški (kratko, srednje, dolgo)	(4,2,1) 2,3	0,090	0,390	Implementacija in vmesna migracija
Obdobje hranjenja	5	0,090	0,450	Dolgoročna tehnika
Hranjenje kompleksnih formatov	4	0,085	0,085	Na dokument gleda kot na zaporedje bitov, zato tudi kompleksnejše
Obvladovanje količine informacij	4	0,085	0,170	Veliko količino informacij, a zaradi migracije ne toliko kot emulacija
Oskrbovanje	4	0,075	0,225	Oskrbovanje strojne opreme zanemarljivo
Prostor	4	0,075	0,075	Učinkovita uporaba prostora
Preprostost izvedbe	1	0,070	0,070	Zahtevna implementacija, ni vedno potrebna emulacija
SKUPAJ		3,94		

Tabela 11: Primerjava ocen tehnik dolgoročnega arhiviranja

	papirna tehnika	tehnološki muzej	migracija	Emulacija	enkapsulacija
Končna ocena	2,83	3,34	3,79	4	3,94

Ugotovimo lahko, da so za dolgoročno arhiviranje res bolj primerne tehnike enkapsulacije, emulacije (ki sta po rezultatu najbolj ocenjeni) ter migracije, medtem ko sta papirna tehnika ter strategija tehnoloških muzejev nekoliko manj uporabni. Papirna tehnika predvsem zaradi izgube integritete ter funkcionalnosti, tehnološki arhivi pa zaradi hitre tehnološke zastarelosti in zahtevnega vzdrževanja.

Glavni cilj vsakega arhiviranja mora biti ohranitev integritete, ki jo enkapsulacija, emulacija ter migracija tudi zagotavljajo. Še enkrat je treba omeniti, da je izbira tehnike odvisna tudi od kompleksnosti in standardiziranosti formatov, obdobja hranjenja, stroškov ter tveganja izgube informacij, ki so zaradi zahtevnejše implementacije in nepreizkušenosti tehnik večji pri emulaciji ter enkapsulaciji. Tako moramo v Tabeli 11 predstavljene končne ocene razumeti kot subjektivne ocene, odvisne od različnih potreb v povezavi z arhiviranjem.

Moja analiza temelji predvsem na teoretičnih člankih mnogih strokovnjakov s področja arhiviranja, ki pa se pogosto precej razlikujejo, saj so mnenja o učinkovitosti in uporabnosti posameznih strategij med strokovnjaki precej neenotna in pogosto temeljijo na premalo praktičnih izkušnjah. Za bolj točno oceno bi bilo treba izvesti praktične kvantitativne teste učinkovitosti tehnik ali orodij, ki so za nekatere tehnike že razvite. Taki rezultati bi lahko dali dobre smernice za izbiro in razvoj najbolj primernega orodja.

Glede mojih rezultatov sem z veseljem ugotovil, da so skoraj identični mojim začetnim pričakovanjem. Moje mnenje je, da imata predvsem metodi emulacije in enkapsulacije v bližnji prihodnosti veliko možnost prevzeti glavno vlogo kot strategijo dolgoročnega arhiviranja, če le ne bo prevelikih težav z implementacijo. Pri tem je seveda treba upoštevati, da bodo zaradi hitrega tehnološkega razvoja neprestano predstavljane nove, boljše, hitrejše, zanesljivejše strategije. Zato se mogoče poraja vprašanje, ali je v tem trenutku sploh smiselno reševati problem dolgoročnega arhiviranja, če pa bodo jutri na voljo nove in boljše tehnike arhiviranja. Ali ne bi v tem primeru informacije hranili le začasno, jutri pa prevzeli nove metode. Ali ne bi začasno hranili informacij v nedogled.

8. ZAKLJUČEK

Arhiviranje predstavlja za sodobno družbo izjemno pomembno področje, ki se precej povezuje tudi z mnogimi drugimi področji, ki za razumevanje arhiva in njegovega okolja niso nepomembna. Zelo pomembno je za podjetja tako v zasebnem kot javnem sektorju, ki hranijo dokumente v arhivu, zadnje stanje v življenjskem ciklu dokumenta. Nič manj ni pomembno za mnoge institucije, ki hranijo dokumente kot kulturno dediščino, itd. Rad bi posebej poudaril, da je področje arhiviranja zelo zanimivo za proučevanje, vendar žal okviri te naloge ne predvidevajo bolj poglobljenega znanja o le-teh.

Dolgoročno arhiviranje digitalnih vsebin je razmeroma novo področje, na katerem bo potrebnih še veliko raziskav, preden bo uveljavljena dokončna rešitev. Glavni problem arhiviranja digitalnih informacij je hiter tehnološki razvoj, ki povzroča zastarelost strojne in programske opreme, na katerih so bili prvotno ustvarjeni dokumenti. Razvoj povzroči, da dokumenti v novem tehnološkem okolju pogosto niso več dostopni za uporabnike hranjenih informacij. Z drugo besedo, informacije postanejo neuporabne.

Možnih je nekaj strategij dolgoročnega hranjenja dokumentov, ki imajo svoje prednosti in slabosti. Pogosto je najboljša praksa kar kombinacija nekaterih od teh. Danes je predstavljenih tudi nekaj standardov, ki te strategije vključujejo v svoje delovanje. Eden teh je Odprt Arhivski Informacijski Sistem (OAIS), ki je primeren predvsem za strategijo migracije ter enkapsulacije. Poleg emulacije pa sta ti dve strategiji trenutno najboljši potencial za rešitev vprašanja dolgoročnega arhiviranja, kar sem tudi uspel prikazati z analizo teh strategij.

Mnenja glede najboljših trenutnih rešitev dolgoročnega arhiviranja se med strokovnjaki na tem področju precej razlikujejo (čeprav jih ne smemo obravnavati absolutno, saj so nekatere tehnike bolj primerne za arhiviranje ene vrste arhivskega gradiva, druge pač za drugo). V prihodnosti bo potrebnih še veliko testiranj, ki bodo tudi kvantitativno lahko podala oceno učinkovitost orodij za dolgoročno hranjenje digitalnih vsebin. S temi rezultati bo lažje postaviti smernice za izbiro prave strategije in s tem pravega orodja za dolgoročno prezervacijo različnih formatov digitalnih virov in posledično tudi za razvoj novih standardov. Seveda pa je treba upoštevati tudi razvoj tehnologije, ki bo neprestano prinašal nove ideje. Vprašanje je le, ali bomo v tej neprestani tekmi s časom imeli še toliko časa, da se ustavimo pri eni in zagotovimo njeno uresničitev.

9. LITERATURA

1. Dobnikar Aleš, Žužek Alenka: Arhiviranje digitalno podpisanih dokumentov. Tehnični in vsebinski problemi klasičnega in elektronskega arhiviranja. Maribor: Pokrajinski arhiv Maribor, 2002, str. 269-277.
2. Dular Dušan, Jakovljević Čedo: Upravljanje dokumentov in postopkov – Splošna rešitev za farmacevtsko industrijo. DOK_SIS 2001 – Portorož, 24. do 26. maj 2000: Sistemi za upravljanje z dokumenti. Ljubljana: Media.doc, 2000, V-39 – V-42.
3. Fenney Mary: Towards a National Strategy for Archiving Digital Materials. Alexandria, 11 (1999) 2, str. 107-121.
4. Garrett John, Waters Donald: Preserving Digital Information: Report of the Task Force on Archiving of Digital Information. [URL: <http://www.rlg.org/ArchTF/tfadi.index.htm>], 8. 2. 2003
5. Graham Peter S.: Long-Term Intellectual Preservation. [URL: <http://www.ifla.org/documents/libraries/net/dps.htm>], 14. 2. 2003
6. Hodge Gail M.: Best Practices for Digital Archiving - An Information Life Cycle Approach. [URL: <http://www.dlib.org/dlib/january00/01hodge.html>], 22. 2. 2003
7. Južnič Primož: Arhivi in specialne knjižnice, k večji povezanosti? Sodobni arhivi 2000: XXII. posvetovanje o strokovnih in tehničnih vprašanjih v arhivih. Maribor: Pokrajinski arhiv Maribor, 2000, str. 77-88.
8. Kavčič-Čolič Alenka: Elektronsko arhiviranje spletnih strani: nov izziv za nacionalne knjižnice. Digitalna knjižnica: Strokovno posvetovanje – Radenci, 10. - 12. oktober 2001, Hotel Radin. Ljubljana: Zveza bibliotekarskih društev Slovenije, 2001, str. 211-224
9. Klasinc Peter Pavel: Slovenska arhivistika in zakon o elektronskem poslovanju in elektronskem podpisu (ZEPEP). DOK_SIS 2001 – Portorož, 23. do 25. maj 2001: Sistemi za upravljanje z dokumenti. Ljubljana: Media.doc, 2001, I-67 – I-74.
10. Krnel-Umek Duša: Standardi za popisovanje arhivskega gradiva na primeru urejanja sodobnih fondov. Sodobni arhivi 2000: XXII. posvetovanje o strokovnih in tehničnih vprašanjih v arhivih. Maribor: Pokrajinski arhiv Maribor, 2000, str. 26-42
11. Kuny Terry: "The Digital Dark Ages? Challenges in the Preservation of Electronic Information.". [URL: <http://www.ifla.org/VI/4/news/17-98.htm#2>], 3. 3. 2003
12. Lee Kyong-Ho et al.: The State of the Art and Practice in Digital Preservation. Journal of Research of the National Institute of Standards and Technology, Vol. 107, No.1. Gaithersburg ND: National Institute of Standards and Technology, 2002, str 93-106.

13. Lorist H. H. Jeroen, Meer, Kees van der.: Standards for digital libraries and archives: digital longevity. Delft: Delft University of Technology, 2001. 9 str.
14. Novak Miroslav: Arhivska stroka in e-dokumenti. DOK_SIS 2002 – Portorož, 22. do 24. maj 2002: Sistemi za upravljanje z dokumenti. Ljubljana: Media.doc, 2002, III-1 – III-10.
15. Novak Miroslav: Dokument kot izziv sodobnega poslovanja. DOK_SIS 2002 – Portorož, 22. do 24. maj 2002: Sistemi za upravljanje z dokumenti. Ljubljana: Media.doc, 2002, III-56 – III-66.
16. Novak Miroslav: Sodobni trendi sistemov za upravljanje z dokumenti v arhivih. DOK_SIS 2001 – Portorož, 23. do 25. maj 2001: Sistemi za upravljanje z dokumenti. Ljubljana: Media.doc, 2001, I-1 – I-9.
17. Pučko Danijel: Strateško upravljanje. Ljubljana: Ekonomska fakulteta, 1996. 233 str.
18. Reference Model for an Open Archival Information System (OAIS). CCSDS 650.0-R-2, Consultative Committee for Space Data Systems, 2002. 111 str., 44 pril.
[URL <http://www.ccsds.org/documents/pdf/CCSDS-650.0-R-2.pdf>], julij 2002.
19. Rothenberg Jeff: An Experiment in Using Emulation to Preserve Digital Publications. Hague: Koninklijke Bibliotheek, 2000. 85 str.
20. Rothenberg Jeff: Avoiding technological quicksand: Finding a Viable Technical Foundation for Digital Preservation, Council on Library and Information Resources. Washington, DC., 1999.
21. Selan Milan: Problemi berljivosti e-zapisov v stalni zbirki orgnov državne uprave. DOK_SIS 2001 – Portorož, 23. do 25. maj 2001: Sistemi za upravljanje z dokumenti. Ljubljana: Media.doc, 2001, II-19 – II-25.
22. Vodopivec Jedert: Standardi in predpisi na področju zaščite in hrambe arhivskega in knjižničnega gradiva. Digitalna knjižnica: Strokovno posvetovanje – Radenci, 10.- 12. oktober 2001, Hotel Radin. Ljubljana: Zveza bibliotekarskih društev Slovenije, 2001, str. 225- 232.
23. YEA - The Yale Electronic Archive: Report on the Digital Preservation Planning Project. New Haven (CT): Yale University Library and Elsevier Science, 2002. 136 str.
24. Žumer Vladimir: Arhiviranje zapisov: priročnik za ravnanje z dokumentarnim in arhivskim gradivom. Ljubljana: GV Založba, 2001. 479 str.

10. VIRI

1. Dimec Jure: Zapiski predavanj pri predmetu Podatkovne zbirke za šolsko leto 1999/2000. [URL: http://www.mf.uni-lj.si/~jure/pred_bib/prosojnice/pz2/d-arhiviranje/D-ARHIVIRANJE.PPT], 23. 2. 2003
2. Emulation. PADI – Preserving Acces to Digital Information. National Library of Australia. [URL: <http://www.ula.gov.au/padi/topics/topic19.html>], 8. 2. 2003.
3. Encapsulation. PADI – Preserving Acces to Digital Information. National Library of Australia. [URL: <http://www.ula.gov.au/padi/topics/topic20.html>], 8. 2. 2003.
4. Holdsworth David, Wheatley Paul: Emulation, Preservation and Abstraction. [URL: <http://www.rlg.org/preserv/diginews/diginews5-4.html#feature2>], 12. 2. 2003.
5. Metadata. PADI – Preserving Acces to Digital Information. National Library of Australia. [URL: <http://www.ula.gov.au/padi/topics/topic30.html>], 8. 2. 2003.
6. Migration. PADI – Preserving Acces to Digital Information. National Library of Australia. [URL: <http://www.ula.gov.au/padi/topics/topic21.html>], 8. 2. 2003.
7. Preserving Digital Information: Report of the Task Force on Archiving of Digital Information. [URL: <ftp://ftp.rlg.org/pub/archtf/final-report.pdf>], 25. 2. 2003.
8. Standards. PADI – Preserving Acces to Digital Information. National Library of Australia. [URL: <http://www.ula.gov.au/padi/topics/topic43.html>], pregledano 8. 2. 2003.
9. Technological obsolescence. PADI – Preserving Acces to Digital Information. National Library of Australia. [URL: <http://www.ula.gov.au/padi/topics/topic13.html>], 8. 2. 2003.
10. XML and Digital Preservation. Digital Preservation Testbed Whitepaper. [URL: http://www.digitaleduurzaamheid.nl/bibliotheek/docs/white-paper_xml-en.pdf], 17. 2. 2003
11. Zakon o arhivskem gradivu in arhivih (Uradni list RS, št. 20/1997).
12. Zakon o elektronskem poslovanju in elektronskem podpisu (Uradni list RS, št. 57/2000).

11. PRILOGA

Priloga A

Okrajšave

AIP	Archival Information Package
ASCII	American Standard Code for Information Interchange
CCSDS	Consultative Committee for Space Data Systems
CDO	Content Data Object
CD-ROM	Compact Disk - Read Only Memory
CRC	Cyclical Redundancy Check
DBMS	Data Base Management System
DMA	Document Management API
DTS	Digital time-stamping
DIP	Dissemination Information Package
DVD	Digital Video Disk
FTP	File Transfer Protocol
IEEE	Institute of Electrical and Electronic Engineers
ISAD	Information System for Assembly and Disassembly
ISBN	International Standard Book Number
ISO	International Organization for Standardization
MARC	Machine-Readable Cataloging
NARA	National Archives and Records Administration
NASA	National Aeronautics and Space Administration
NEDLIB	Networked European Deposit Library
OAIS	Open Archival Information System
ODMA	Open Document Management API
PANDORA	Preserving and Accessing Networked Documentary Resources of Australia
PDF	Portable Document Format
PDI	Preservation Description Information
SIP	Submission Information Package
SUTRS	Simple Unstructured Text Record Syntax
XML	eXtensible Markup Language
UML	Unified Modeling Language
UNICODE	Universal Code
WebDAV	Web-based Distributed Authoring and Versioning
WWW	World-Wide Web

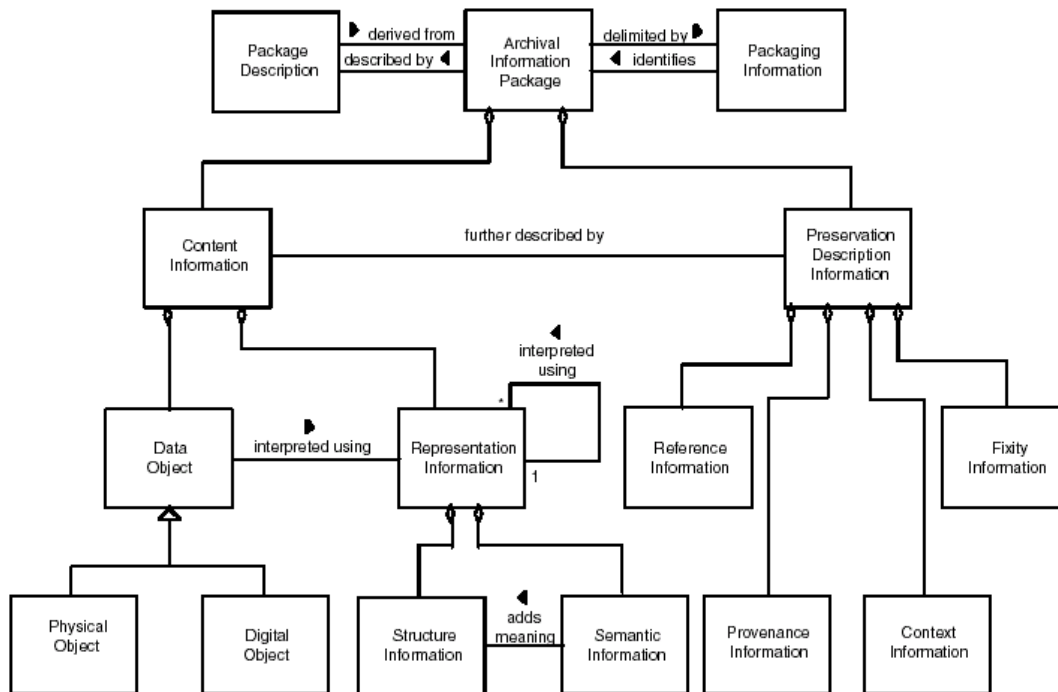
Prevodi tujih besed

Archive	Arhiv, arhivsko gradivo
Archival Storage mapping Infrastructure	Infrastruktura načrtovanja arhivskega skladišča
Archival information package	Arhivski informacijski paket
Content information	Vsebinska informacija
Context	Kontekst
Description information	Opisna informacija
Digital information	Digitalna informacija
Digital object	Digitalni objekt
Digital time-stamping	Digitalni časovni žig
Disaster recovery	Obnovitev stanja pred uničenjem/izbrisom
Dissemination	Oddaja, razpošiljanje, diseminacija
Dissemination information package	Oddan (diseminacijski) informacijski paket
eXtensible Markup Language (XML)	Razširljiv označevalni jezik
Fixity	Stabilnost
Information package	Informacijski paket
Open Archival Information System	Odpri Arhivski Informacijski Sistem
Package description	Opis paketa
Packaging information	Informacija o paketiranju
Preservation description information	Opisna informacija o ohranjanju
Provenance	Izvor
Refreshment	Osveževanje podatkov
Replication	Podvajanje podatkov
Repackaging	Prepaketiranje podatkov
Submission	Sprejetje, predložitev, potrditev
Submission information package	Sprejet (predložen) informacijski paket
Reference	Vir, referenca
Submission information package	Prejet informacijski paket
Submission Agreement	Pogajanja o vročitveni pogodbi
Transformation	Pretvorba podatkov, preoblikovanje, transformacija

Priloga B

Strukturna povezava informacij

Slika 6: Podrobnejši pogled na členitev AIP



Vir: Reference Model for an Open Archival Information System, 2002, str. 4-35.

Komentar:

Arhivski informacijski paket (AIP) je sestavljen iz Vsebinskih informacij ter Ohranitvenih opisnih informacij (PDI). Informacije o paketu hranijo metapodatki znotraj informacije Opis paketa. AIP identificira Informacija o paketiranju, ki vsebuje podatke o razporeditvi vsebine znotraj AIP.

Vsebinsko informacijo sestavljajo Podatkovni objekti, ki so lahko v digitalni (niz bitov) ali fizični (papir) obliki, pomen pa jim dajejo Predstavitvene informacije. Predstavitvene informacije vključujejo semantične informacije, ki dajejo pomen strukturnim informacijam. Skupaj pa predstavijo objekt v uporabnikom razumljivi obliki.

Pomen Vsebinskim informacijam dodatno dajejo Ohranitvene opisne informacije. Te vključujejo naslednje vrste informacij: Izvor, Kontekst, Vir, Stabilnost.

Priloga C

Pomen konteksta pri arhiviranju

Pri pretvorbi digitalnih dokumentov je zelo pomembno, da ti ohranijo kontekst, ki dokumentu še naprej zagotavlja pomen, kot ga je ta imel pred pretvorbo v novo obliko. V primeru, ko v novo obliko pretvorjen dokument izgubi kontekst, postane ta za uporabnika neuporaben.

Prvotna oblika

arhiviranje
digitalno je izrednega *na kasnejše*
Dolgoročno pomena. prenesli rodove?
Le kako drugače
bomo znanje

Pretvorjena oblika z izgubo konteksta

arhiviranje
digitalno je izrednega na kasnejše
Dolgoročno pomena. prenesli rodove?
Le kako drugače
bomo znanje