

Hrvoje Stančić

digitalizacija

Hrvoje Stančić DIGITALIZACIJA



Zavod za informacijske studije



Hrvoje Stančić

Digitalizacija

Zavod za informacijske studije
Zagreb, veljača 2009.

Pregled sadržaja

Sadržaj	3
Predgovor	7
Uvod	9
1. Odabir gradiva za digitalizaciju	15
2. Digitalizacija gradiva	33
3. Obrada i kontrola kvalitete	71
4. Zaštita gradiva u elektroničkoj okolini	95
5. Pohrana i prijenos digitalnoga gradiva	113
6. Pregled i korištenje digitalnoga gradiva	139
7. Održavanje digitalnoga gradiva	141
Zaključak	159
Popis priloga	161
Literatura	165

Sadržaj

Predgovor	7
Uvod	9
1. ODABIR GRADIVA ZA DIGITALIZACIJU	15
1.1. Predlaganje gradiva.....	19
1.2. Procenjivanje gradiva	20
1.3. Određivanje prioriteta.....	20
1.3.1. Vrijednost.....	21
1.3.2. Rizik.....	21
1.3.3. Predviđeno korištenje	22
1.4. Digitalizacija unutar ili izvan institucije	25
1.4.1. Kako odabrati vanjskog pružatelja usluga?	26
1.4.2. Razvoj projekta i njegovih ciljeva	27
1.4.3. Zahtjev za informacijama	27
1.4.3. Zahtjev za ponudom	28
1.4.4. Kontrola kvalitete	32
2. DIGITALIZACIJA GRADIVA.....	33
2.1. Uvod	33
2.2. Uređaji za digitalizaciju.....	33
2.2.1. Skeneri.....	33
2.2.1.1. Koračni skeneri	34
2.2.1.1.1. Ručni skeneri.....	34
2.2.1.1.2. Plošni skeneri	34
2.2.1.1.3. Skeneri za mikrooblike	35
2.2.1.1.4. Rotacioni skeneri	37
2.2.1.1.5. Reprografski skeneri.....	39
2.2.1.1.6. 3D skeneri.....	39

2.2.1.1.7. Skeneri za knjige	40
2.2.1.2. Protočni skeneri	41
2.2.1.3. Karakteristike skenera	43
2.2.2. Digitalna kamera / fotoaparati	46
2.2.2.1. Vrste fotoaparata	49
2.2.2.1.1. Amaterski fotoaparati	49
2.2.2.1.2. Profesionalni fotoaparati	50
2.2.2.1.3. Studijske kamere	51
2.2.2.1.4. Profesionalni modularni sustav digitalnih kamera	51
2.2.3. Uređaji za digitalizaciju zvuka	51
2.2.4. Uređaji za digitalizaciju filma i videa	53
2.3. Postupci digitalizacije	55
2.3.1. Tekstualno gradivo	55
2.3.1.1. Prepisivanje gradiva	55
2.3.1.2. Skeniranje gradiva	55
2.3.1.3. Slikanje digitalnim fotografskim aparatom	57
2.3.2. Slikovno gradivo	58
2.3.2.1. Rezolucija	60
2.3.2.2. Bitna dubina točke	60
2.3.2.3. Boja	62
2.3.3. Zvučno gradivo	65
2.3.3.1. Uzorkovanje	66
2.3.3.2. Kvantizacija	68
2.3.4. Filmsko i video gradivo	69
2.3.5. Trodimenzionalno gradivo	70
3. OBRADA I KONTROLA KVALITETE	71
3.1. Tekstualno gradivo	71
3.1.1. Notni zapisi	73
3.2. Slikovno gradivo	75
3.2.1. Matematički modeli kodiranja	75
3.2.1.1. Kodiranje entropije	76
3.2.1.2. Huffmanovo kodiranje	76
3.2.1.3. Lemper-Ziv kodiranje	76
3.2.1.4. Kodiranje po principu dugih nizova	77
3.2.1.5. Pretprocesiranje filterom predviđanja	77
3.2.2. JPEG standard	77
3.2.3. Komprimiranje bez gubitaka	78

3.2.4. Komprimiranje s gubicima	78
3.3. Zvučno gradivo	82
3.3.1. MPEG standard	83
3.3.2. Komprimiranje bez gubitaka	83
3.3.3. Komprimiranje s gubicima	83
3.3.3.1. Komprimiranje nepromjenjivim brojem bitova	84
3.3.3.2. Komprimiranje promjenjivim brojem bitova	85
3.4. Filmsko i video gradivo	90
3.4.1. Prijenos zapisa s nepromjenjivim brojem bitova	92
3.4.2. Prijenos zapisa s promjenjivim brojem bitova	93
3.5. Trodimenzionalno gradivo	94
4. ZAŠTITA GRADIVA U ELEKTRONIČKOJ OKOLINI	95
4.1. Mehanizmi zaštite sustava	95
4.2. Šifriranje	97
4.2.1. Šifriranje simetričnim ključem	98
4.2.2. Šifriranje javnim ključem	99
4.3. Digitalni potpisi	101
4.4. Digitalni certifikati	104
4.5. Digitalni vodeni žigovi	106
4.5.1. Steganografija	109
5. POHRANA I PRIJENOS DIGITALNOGA GRADIVA	113
5.1. Izravni sustavi za pohranu i prijenos podataka	116
5.2. Poluizravni sustavi za pohranu i prijenos podataka	117
5.3. Hijerarhijski sustav za pohranu i prijenos podataka	118
5.3.1. Usporedba izravnih, poluizravnih i hijerarhijskih sustava za pohranu	120
5.4. Neizravni sustavi za pohranu i prijenos podataka	126
5.5. Mrežna pohrana	126
5.6. Mreža za pohranu	127
5.7. Sigurnosne kopije	129
5.7.1. Odabir modela skladištenja podataka	130
5.7.1.1. Nestrukturirana sigurnosna kopija	130
5.7.1.2. Potpuna sigurnosna kopija	130
5.7.1.3. Inkrementalna sigurnosna kopija	131
5.7.1.4. Diferencijalna sigurnosna kopija	131
5.7.1.5. Kontinuirana zaštita podataka	132

5.7.2. Model rotacije medija	132
5.7.2.1. Inkrementalni model.....	133
5.7.2.2. Model djed-otac-sin.....	134
5.7.2.3. Model hanojskih tornjeva	135
5.7.3. Odabir medija za pohranu	135
5.7.4. Obrada podataka sigurnosne kopije	136
5.7.4.1. Kompresija.....	137
5.7.4.2. Deduplikacija.....	137
5.7.4.3. Duplikacija.....	137
5.7.4.4. Šifriranje.....	138
5.7.5. Pohrana.....	138
6. PREGLED I KORIŠTENJE DIGITALNOGA GRADIVA.....	139
7. ODRŽAVANJE DIGITALNOGA GRADIVA	141
7.1. Definicija zapisa	141
7.2. Definicija problema	143
7.2.1. Problem pregleda gradiva.....	144
7.2.2. Problem kodiranja gradiva.....	144
7.2.3. Problem međusobne povezanosti sadržaja	145
7.2.3.1. Arhiviranje elektroničkih poruka	146
7.2.3.2. Arhiviranje internetskih stranica	148
7.2.4. Problem ovlasti za arhiviranje	149
7.2.5. Problem konverzije zapisa.....	151
7.3. Prijedlog rješenja.....	152
7.3.1. Osvježavanje medija.....	153
7.3.2. Migracija zapisa	155
7.3.3. Emulacija aplikacijske okoline.....	156
Zaključak.....	159
Popis priloga	161
Literatura	165

Predgovor

Ova se knjiga temelji na magistarskoj radnji autora, obranjenoj na Filozofskom fakultetu Sveučilišta u Zagrebu 2001. godine, i tekstu koji je autor napisao kao član radne skupine za digitalizaciju u okviru Nacionalnoga programa digitalizacije arhivske, knjižnične i muzejske građe, koji je 2006. godine pokrenulo Ministarstvo kulture Republike Hrvatske, tj. nacionalnoga projekta "Hrvatska kulturna bašina". Tekst magistarske radnje je osvježen i proširen, a primjeri su usklađeni s napretkom informacijsko-komunikacijske tehnologije i globalne informacijske infrastrukture. Svi mrežni izvori koji se spominju u knjizi bili su aktivni u trenutku predaje rukopisa u tisak.

Uvod

S izgradnjom globalne informacijske infrastrukture došlo je do globalizacije informacijskih sustava. Komunikacija putem računala postala je svakodnevicom, a količina dostupnih informacija sve je veća. Globalni razvoj interaktivnih sadržaja primjerenih novom mediju doveo je do natjecateljske atmosfere u kojoj se institucije moraju konstantno prilagođavati uvjetima poslovanja i, općenito, djelovanja. Neprestano se mijenjaju računalna i programska okolina, što uzrokuje nova konceptijska rješenja pa ponekad nije dovoljno ili jednostavno nije moguće prijeći na novu verziju bez znatnijih promjena u oblikovanju cijeloga informacijskoga sustava.

Arhivi, knjižnice i muzeji, kao ustanove od društvenog značaja, također su se našli u značajnim promjenama jer su i oni morali oblikovati bar neki dio svojeg gradiva i usluga primjereno globalnoj interaktivnoj komunikaciji te se djelomično i tržišno orijentirati. Mnoge takve ustanove, naime, nude pristup gradivu putem Interneta, ali to je gradivo često dostupno samo u najnižoj kvaliteti zapisa, dok pribavljanje kvalitetnije verzije nije moguće ili se dodatno naplaćuje, čime se barem djelomično nadoknađuju visoki troškovi digitalizacije.

Globalna komunikacija je u nekim granama dovela i do promjene nekih temeljnih načela djelovanja. Tako, na primjer, u postupanju s dokumentima, njihovo se očuvanje do pojave elektroničkog oblika zapisa provodilo tako da se sačuvao medij na kojem je bila zapisana informacija. Pojavom elektroničkih medija i usavršavanjem procesa digitalizacije "pojam očuvanja počinje se dijeliti na dva dijela: očuvanje informacijskog sadržaja, tj. informacije koju određeni dokument nosi, te očuvanje fizičkog objekta kao nositelja informacije. Informacijski sadržaj se, dakle, digitalizira i sprema odvojeno od objekta nositelja."¹ Stoga je bilo potrebno razviti sustave i metode za očuvanje gradiva u elektroničkom

¹ Stančić, Hrvoje, *Digitalizacija građe*, u: Willer, Mirna i Katić, Tinka (ur.), *2. i 3. seminar Arhivi, knjižnice, muzeji, Mogućnosti suradnje u okruženju globalne informacijske infrastrukture*, Zbornik radova, Hrvatsko muzejsko društvo, Zagreb, 2000., str. 64.

obliku, kako bi se u uvjetima neprestanih promjena sačuvao njegov informacijski sadržaj.

Općenito gledajući, digitalizacija arhivskoga, knjižničnoga i muzejskoga gradiva, ali i gradiva drugih tipova koji se mogu naći u bilo kojoj komercijalnoj instituciji, "provodi se radi zaštite izvornika, povećanja dostupnosti i mogućnosti korištenja građe, radi stvaranja nove ponude, odnosno usluga korisnicima ili pak radi upotpunjavanja postojećega fonda. Svaki od navedenih ciljeva digitalizacije postavlja određene zahtjeve koje treba imati u vidu pri planiranju i izvođenju projekata digitalizacije. Vrlo je važno da projekti digitalizacije uoče te zahtjeve, procijene njihovu razmjernu težinu za pojedini projekt i jasno definiraju čime, kako i u kojoj mjeri će njihov krajnji proizvod odgovoriti na pojedini zahtjev."² Noviji trend predstavlja i postupak digitalizacije na zahtjev.³

Digitalizacija radi zaštite izvornika provodi se prije svega zbog dva osnovna aspekta elektroničkih verzija gradiva. Prvi aspekt je taj da se korisnicima tada može ponuditi elektronička verzija umjesto klasične, čime se izvornik može bolje očuvati, jer je tada manje u upotrebi. Drugi aspekt je prezervacijski, tj. elektronička verzija može poslužiti kao sigurnosna kopija u slučaju oštećenja ili uništenja originala.

Digitalizacija radi povećanja dostupnosti provodi se iz razloga koji je jasan sam po sebi. Dok u slučaju posjedovanja jednog originala on u jednom trenutku može biti dostupan samo jednom korisniku, njegova elektronička inačica može biti istovremeno dostupna velikom broju korisnika. Zbog toga institucije koje nude elektroničko gradivo korisnicima mogu očekivati povećanje broja korisnika, a time i povećanje svoje vidljivosti i utjecaja u društvu. Postupcima digitalizacije ključnoga gradiva važnog za kulturnu baštinu neke zemlje može se postići i promocija države u cjelini na globalnoj razini.

Digitalizacija radi stvaranja nove ponude i usluga korisnicima jedan je od najvažnijih razloga za digitalizaciju gradiva. Naime, prebacivanjem u digitalni oblik otvara se čitav niz novih mogućnosti – od razmjene metapodataka između institucija, čime se znatno ubrzava proces obrade digitaliziranoga gradiva, preko pretraživanja punog teksta i raznih analiza svih vrsta gradiva, pa sve do virtual-

nog spajanja sadržaja raznih, fizički možda i vrlo udaljenih, izvora i stvaranja virtualnih zbirki ili izložbi.

Digitalizacija gradiva koje nije u vlasništvu neke institucije ili nabavka njegove digitalizirane verzije može poslužiti radi *upotpunjavanja fonda*. Razlozi za takvo djelovanje mogu biti različiti, od nepotpunog ili uništenog/stradalog dijela fonda, pri čemu je elektronički oblik gradiva jedini oblik u vlasništvu institucije, pa sve do izrade zbirki za koje institucija procijeni da su joj potrebne.

Digitalizacija na zahtjev ne bi smjela biti jedini i/ili najvažniji pristup digitalizaciji gradiva u nekoj instituciji jer ne osigurava osmišljenu izgradnju cjelovitih zbirki u elektroničkom obliku. Ona se po karakteru djelomično preklapa s postupkom digitalizacije radi stvaranja nove ponude i usluga, no najbolje dolazi do izražaja kao dodatna usluga institucije. Također, digitalizacija na zahtjev može se odlično iskoristiti kod postupaka određivanja prioriteta za digitalizaciju kad je gradivo već odabrano za digitalizaciju. Tada se prednost može dati onom gradivu koje korisnici najprije zatraže.

No, bez obzira na razloge digitalizacije, mora se konstatirati da razvoj informacijskih i komunikacijskih tehnologija i njihov sveobuhvatni utjecaj na upravljanje i posredovanje zabilježenog znanja otvara i pitanje kako odrediti zajedničke putove izgradnje informacijskog društva na temelju jedinstvene globalne informacijske infrastrukture. Stoga se ova knjiga bavi procesom digitalizacije i održavanja digitaliziranoga gradiva, ali i onog izvorno nastalog u elektroničkom obliku (engl. *born digitally*), kao početnog koraka u tom procesu. Ispravno postavljen i proveden proces digitalizacije postaviti će dobre temelje za kasniju nadgradnju u obliku osiguranja jednostavne i brze dostupnosti informacija i znanja na globalnoj razini te bolje organizacije njegova posredovanja i upravljanja njime.

Knjiga je koncipirana u sedam poglavlja, u kojima se opisuje cjelokupni proces digitalizacije. On se sastoji od sljedećih koraka: 1. odabir gradiva, 2. digitalizacija gradiva, 3. obrada i kontrola kvalitete, 4. zaštita, 5. pohrana i prijenos, 6. pregled i korištenje⁴ te 7. održavanje digitalnoga gradiva. To je vrlo složen postupak ako se želi sustavno provesti, i to prije svega zato što postoji više različitih vrsta gradiva, pa tako ovdje nije riječ samo o tekstualnom gradivu, već se također opisuju postupci obrade slikovnog, zvučnog, filmskog i video gradiva te trodimenzionalnih objekata kulturne baštine. Naravno, proces digitalizacije obuhvaća i njegovu naknadnu obradu, zaštitu i pohranu uz poštivanje specifičnosti

² *Prijedlog Nacionalnog programa digitalizacije arhivske, knjižnične i muzejske građe* (v. 2.0, konačni tekst), Ministarstvo kulture Republike Hrvatske, 20. listopada 2006., str. 10., <<http://daz.hr/bastina/NacionalniProgramDigitalizacije.pdf>>, 21. kolovoza 2007.

³ *Smjernice za odabir građe za digitalizaciju (radna verzija)*, Nacionalni projekt "Hrvatska kulturna baština", Ministarstvo kulture Republike Hrvatske, studeni 2007., <http://www.kultura.hr/hr/content/download/596/7925/file/smjernice_odabir.pdf>, 27. kolovoza 2008.

⁴ Stančić, *Digitalizacija građe*, n. dj., str. 65.

vezanih uz pojedinu vrstu gradiva, ali i dugoročno očuvanje usprkos neprestanim promjenama hardvera, softvera i normi. Konceptijski problemi i odabir prave tehnologije za digitalizaciju vrlo su kompleksni. Svaki stupanj u procesu digitalizacije zahtijeva zasebna konceptijska i hardversko-softverska rješenja koja izravno ovise o gradivu koje se obrađuje.

Prvo se poglavlje, tako, bavi problematikom procesa odabira gradiva koje će se digitalizirati i određivanja prioriteta za digitalizaciju. U njemu se, također, razmatra i dvojba je li bolje digitalizirati unutar ili izvan institucije.

Drugo poglavlje započinje klasifikacijom i detaljnim opisom uređaja za digitalizaciju svih vrsta gradiva. Za sve uređaje navedene su i njihove osnovne karakteristike koje su važne za proces digitalizacije. Nakon toga su objašnjeni postupci digitalizacije za svih pet vrsta gradiva (tekstualno, slikovno, zvučno, filmsko i video te trodimenzionalno).

U trećem poglavlju razlaže se problematika obrade digitaliziranoga gradiva. U njemu se specifičnosti obrade pojedinih vrsta gradiva razrađuju u potpoglavljima organiziranim, također, prema vrstama gradiva. Osim obrade nakon digitalizacije, vezane prije svega uz poboljšanje kvalitete digitaliziranoga gradiva, ovdje se obrađuju i postupci komprimiranja, tj. sažimanja gradiva u elektroničkom obliku koji su specifični za pojedine vrste gradiva.

Četvrto poglavlje razmatra postupke zaštite gradiva koje se nalazi u elektroničkom obliku. Od ovog koraka u procesu digitalizacije više nije važno je li ono digitalizirano ili je izvorno nastalo u elektroničkom obliku, jer se ovaj korak, kao i daljnji koraci podjednako odnose na obje vrste gradiva. Poglavlje razmatra mogućnosti zaštite sustava od neovlaštenog pristupa, ali i mogućnosti zaštite šifriranjem, digitalnim potpisima, certifikatima, vodenim žigovima itd.

Peto poglavlje obrađuje pitanja pohrane velike količine digitaliziranoga gradiva i njegova prijenosa do korisnika. Stoga je u tom poglavlju riječ o organizaciji složenih sustava za pohranu, njihovim karakteristikama te pozitivnim i negativnim stranama. Poglavlje obuhvaća sustave kao što su izravni, poluizravni, hijerarhijski i neizravni sustavi te sustavi mrežne pohrane i mreže za pohranu. Također se obrađuje i cjelokupna problematika izrade sigurnosnih (engl. *back-up*) kopija.

Šesto poglavlje, koje se odnosi na pregled i korištenje elektroničkoga gradiva, opsegom je najmanje, ali ono daje važne napomene koje treba uzeti u obzir kako bi krajnjim korisnicima uvijek moglo biti isporučeno ono gradivo koje oni traže, i to u onom obliku i formatu zapisa koji trenutačno čine standard.

Posljednje, sedmo poglavlje razmatra problematiku održavanja elektroničkoga gradiva. Elektroničko gradivo ima neke specifične probleme, počevši od očuvanja informacijskog sadržaja odvojenog od medija na kojem je ono zapisano, pa sve do nekih problema koji su toliko kompleksni da im se ne nazire neko konkretno i elegantno rješenje. Ti su problemi pobliže objašnjeni kako bi se upozorilo na njihovu stvarnu dubinu i složenost. Na kraju poglavlja predlažu se i neka od mogućih rješenja za dugoročno očuvanje elektroničkoga gradiva u uvjetima neprestanog napretka računalno-programске okoline, a to znači sustava u kojima se gradivo nalazi, medija na kojima je pohranjeno te formata u kojima je zapisano.

Na kraju ovog, uvodnog dijela važno je napomenuti da pokretanje projekata digitalizacije ima smisla samo ako se u te projekte uđe predano i s jasnom vizijom što se njima želi postići. Dodatno, potrebno je predvidjeti potrebna financijska sredstva ne samo za provedbu procesa digitalizacije, već i za kasnije održavanje toga gradiva, proširenje, tj. poboljšanje sustava i edukaciju djelatnika. Nitko ne bi trebao krenuti s projektom digitalizacije, a da nije siguran hoće li moći očuvati digitalizirano gradivo na dulji vremenski rok ili, barem, u čijoj će ovlasti biti preuzimanje toga gradiva u slučaju nedostatka potrebnih financijskih sredstava. Očuvanje u ovom kontekstu znači osiguranje pristupa sadržaju elektroničkoga gradiva, a postupak njegova očuvanja započinje stvaranjem gradiva, tj. digitalizacijom. Upravo zbog toga, ispravno planiran i dosljedno provođen proces digitalizacije od velikog je utjecaja na kasnije mogućnosti dugoročnog očuvanja digitaliziranoga gradiva. Ove je činjenice važno neprestano imati na umu tijekom analize svih koraka procesa digitalizacije.

Zbog svega navedenoga ova je knjiga namijenjena kako početnicima tako i stručnjacima u području digitalizacije upravo zbog zaokruženosti pristupa tematici. Njome će se moći koristiti svi djelatnici u arhivima, knjižnicama, muzejima, drugim informacijsko-dokumentacijskim institucijama i tvrtkama kako bi jednostavnije osmislili digitalizacijski projekt, bez obzira bio on manjega ili većega opsega. Ona će, također, poslužiti i studentima kao osnovna literatura u području digitalizacijske problematike.

1. ODABIR GRADIVA ZA DIGITALIZACIJU

Na početku procesa digitalizacije gradiva u nekoj ustanovi potrebno je odlučiti koje će se gradivo digitalizirati. Kada kažem *koje gradivo*, mislim *koje gradivo najprije* digitalizirati, jer se danas sve može digitalizirati uz više ili manje napore i veće ili manje troškove. Odabir, dakako, ovisi prije svega o vrsti ustanove, njezinu programu i ciljevima koji se žele postići digitalizacijom. Već je u uvodu rečeno i pomnije objašnjeno da se digitalizacija može provoditi radi zaštite izvornika, povećanja dostupnosti i mogućnosti korištenja građe, radi stvaranja nove ponude, tj. usluga korisnicima, radi upotpunjavanja postojećega fonda ili, pak, ona može biti uspostavljena kao usluga digitalizacije na zahtjev. U tom se kontekstu treba promatrati i proces odabira gradiva.

Cijeli proces odabira gradiva za digitalizaciju može se prikazati i matricom za donošenje odluka. Ona prikazuje redoslijed odlučivanja ostvaren u obliku stabla odlučivanja, tj. dijagrama toka, i dobar je pregled koraka/odluka koje je potrebno učiniti/donijeti prije nego što se krene u stvarni odabir gradiva, tj. još za vrijeme definiranja digitalizacijskog projekta. Poželjno je promisliti o detaljima koji se spominju u matrici te pokušati odgovoriti na tamo postavljena pitanja jer će tako projekt već u početku biti dobro pripremljen. Matrica za donošenje odluka djelomično zadire u koncepciju cijelog projekta digitalizacije, a djelomično se odnosi i na njegov prvi korak – odabir gradiva (vidi matricu 1).

Proces odabira gradiva je postupak kojim se, na temelju prethodne temeljite analize gradiva i primjene kriterija za odabir, određuje koje će gradivo biti digitalizirano te kojim redoslijedom. Kvalitetni kriteriji odabira vodit će se pravnim propisima te ići prema odabiru onoga gradiva za koje se smatra da će biti najčešće korišteno, a njegovo digitaliziranje najjeftinije. Dakle, najbolje je odabrati veće količine materijala koji je jednostavniji za digitalizaciju i za koji se očekuje manje problema prilikom obrade. Tako je, na primjer, bolje odabrati knjige, nego stare karte koje sadrže sliku, tekst i možda po koju rukom zabilježenu bilješku na margini već požutjelog papira. Digitalizirati najprije najpopularnija djela dobar je način za stjecanje šire popularnosti i potpore. U američkoj Kon-

```

graph TD
    Start([Kraj!]) -- ne --> Q1{Je li gradivo toliko vrijedno da njegova digitalna verzija može osigurati zanimanje korisnika?}
    Q1 -- da --> Q2{Hoće li se digitalizacijom znatno povećati pristup i korištenje?}
    Q2 -- ne --> End1([Kraj!])
    Q2 -- da --> Q3{Mogu li se predvidjeti i definirati korisnici?}
    Q3 -- ne --> End1
    Q3 -- da --> Q4{Koji se ciljevi žele ostvariti digitalizacijom?}
    Q4 --> Q5{Zaštita izrada vjernih preslika koje će zamijeniti gradivo u lošem stanju ili smanjiti uporabu krhkih i/ili vrijednoga gradiva}
    Q5 --> Q6{Bola intelektualna kontrola (npr. izrada elektroničkog pomagala za pronalaženje povezanog s digitalnim slikama; izrada digitalnih stranica sa sadržajem povezanih s metapodacima)}
    Q6 --> Q7{Bola funkcionalnost (mogućnost široke distribucije, pretraživanja i obrade teksta, proučavanja različitoga gradiva u istom kontekstu)}
    Q7 --> Q8{Ušteda (npr. izrada virtualnih zbirki različitih ustanova koje dijele troškove i odgovornost)}
    Q8 --> Q9{Može li se ustanova podnijeti troškove skeniranja i postupaka nakon skeniranja?}
    Q9 -- ne --> End1
    Q9 -- da --> Q10{Ima li ustanova potrebna sredstva i stručnjake za izradu i provedbu projekta?}
    Q10 --> Q11{Je li trenutna organizacijska i tehnička infrastruktura dovoljna za izradu, upravljanje i dostavu digitalnih produkata?}
    Q11 --> End1
    Q11 --> Q12{Postoji li već gradivo koje zadovoljava identificirane potrebe?}
    Q12 -- ne --> End1
    Q12 -- da --> Q13{Jesu li osigurana prava za izradu i raspačavanje?}
    Q13 -- ne --> End1
    Q13 -- da --> Q14{Omogućuje li trenutna raspoloživa tehnologija potrebnu kvalitetu slike da bi se ostvarili utvrđeni ciljevi (točan OCR, izradu visoko kvalitetnih primjeraka)?}
    Q14 -- ne --> Q15{Dopušta li tehnologija prikladan prijenos u digitalni oblik s posredničkog medija (fotografije)?}
    Q15 -- ne --> End1
    Q15 -- da --> Q16{Može li se ustanova podnijeti troškove skeniranja i potrebe infrastrukture?}
    Q16 -- ne --> End1
    Q16 -- da --> End2([Kraj!])
    End2 --> End3([Nastaviti planirati, primjenjivati i vrednovati projekt!])
    
```

gresnoj knjižnici je, na primjer, roman *Prohujalo s vihorom* 1996. godine zatražen u digitalnoj verziji mnogo više puta od bilo kojeg svečanog govora iz vremena I. svjetskog rata.⁵

- stručnjaci iz disciplina koje su tematski vezane uz gradivo koje se želi digitalizirati,
- edukacijski stručnjaci na razini ovisnoj o publici na koju projekt cilja,
- stručnjaci za digitalizaciju,
- knjižničari, arhivisti i muzeolozi,
- istraživači s iskustvom u radu s mrežnim izvorima,
- stručnjaci za konzervaciju i očuvanje te
- pravnici.⁶

*Priručnik za digitalne projekte*⁷, koji je razvio Sjeveroistočni centar za konverziju dokumenata iz Andovera u Massachusettsu, SAD, sugerira da je proces

⁶ Vogt-O'Connor, Diane, Selecting of Materials for Scanning, u: *Handbook for Digital Projects: A Management Tool for Preservation and Access*, ur. Sitts, Maxine K., Northeast Document Conversion Center, Andover, Massachusetts, 2000., str. 49, <<http://www.nedcc.org/oldnedccsite/digital/dighome.htm>>, 18. studenog 2008.

⁷ Vogt-O'Connor, *Selecting of Materials*, n. dj., str. 46.

odabira dokumenata za digitalizaciju u osnovi vrlo sličan principima odabira dokumenata u druge svrhe. Prema *Priručniku* odabir se odvija kroz tri faze – predlaganje gradiva, procjenjivanje gradiva i određivanje prioriteta.

Ono što *Priručnik* izričito ne spominje, a važno je naglasiti, jest korak koji treba prethoditi navedenim trima fazama, a to je uspostava kriterija za odabir. Kad su kriteriji jednom doneseni, tek tada se može pristupiti provedbi triju spomenutih faza odabira jer će se članovi povjerenstva za odabir tako moći voditi jednakim osnovnim mjerilima. Pritom treba imati na umu da će se različiti kriteriji primjenjivati na različito gradivo – arhivsko, knjižnično i muzejsko. *Smjernice za odabir građe za digitalizaciju* nastale u okviru Nacionalnog projekta “Hrvatska kulturna baština” donose preporuke za svako navedeno gradivo. Za arhivsko gradivo posebno ističu preporuke da “pri vrednovanju građe svako treba obratiti pozornost na njezinu informacijsku i evidencijsku vrijednost. Informacijska vrijednost odnosi se na obuhvat i cjelovitost informacija koju određena građa prenosi, njezin odnos prema informacijskoj vrijednosti drugih izvora u svezi s istom temom i na procjenu interesa određene korisničke skupine za tu temu. Evidencijska vrijednost odnosi se na sposobnost određene građe da pruži uvid i dokaze o nekoj aktivnosti ili događaju. Izvorni dokument imat će veću evidencijsku vrijednost od prepričanog i komentiranog sadržaja, koji pak može imati veću informacijsku vrijednost jer donosi pojašnjenja i druge informacije. Pri vrednovanju treba uzeti u obzir da digitalizacijom građa zadobiva dodatnu vrijednost jer takva građa stvara neke nove i bolje mogućnosti korištenja i upravljanja zbirkom ili fondom⁸.” Za knjižno gradivo “u postupku odabira luči se: a) knjižnična građa/knjižnične zbirke koje su već upisane u *Registar kulturnih dobara RH* (Lista preventivno zaštićenih dobara, Lista zaštićenih kulturnih dobara i Lista kulturnih dobara od nacionalnog značenja) i b) knjižnična građa/knjižnične zbirke za koje se predmnijeva da imaju svojstvo kulturnog dobra, ali još iz različitih razloga nisu upisane na neku od lista u *Registru kulturnih dobara RH*⁹.” Za muzejsko gradivo *Smjernice* navode da “postojeću legislativu u muzejskoj struci ne prate i standardi vrijednosnog sustava za kategorizaciju muzejskog predmeta odnosno kriteriji za vrednovanje muzejske građe ili muzejske zbirke/ustanove. Za razliku od arhivske i knjižnične građe, muzejska je građa kompleksnija i u načelu se sastoji od muzejskog predmeta koji je većim dijelom trodimenzionalan i često unikat te popraćen raznorodnom dokumentacijom.”¹⁰

⁸ Smjernice za odabir građe za digitalizaciju, n. dj. str. 9.

⁹ Smjernice za odabir građe za digitalizaciju, n. dj. str. 10.

¹⁰ Smjernice za odabir građe za digitalizaciju, n. dj. str. 10.

U nastavku je objašnjen postupak odabira gradiva prema spomenutom *Priručniku za digitalne projekte*.

1.1. PREDLAGANJE GRADIVA

Predlaganje gradiva je proces u kojem članovi povjerenstva za odabir daju mišljenje o određenim segmentima gradiva, tj. treba li ih uzeti u razmatranje za digitalizaciju ili ne te zašto. *Priručnik*, nadalje, navodi niz pitanja kojima se trebaju voditi članovi povjerenstva:

- Koliko je gradiva dobro i ažurno dokumentirano na pojedinačnoj razini u pouzdanim i potpunim indeksima i sustavima za pretraživanje te gdje se to gradivo nalazi?
- Koliko je gradiva u stabilnom ili dobrom stanju?
- Koliki dio papirnog gradiva je standardne i konzistentne veličine, normalnog kontrasta, koliko ih je u crno-bijeloj tehnici i/ili koliko ih je u tiskanom obliku te kojem tematskom području pripadaju? Uputno je izbje-gavati velike, neobične ili nejednake formate, kao i gradivo u boji ili rukopise, bar kad je riječ o početnim projektima.
- Koji su materijali lako dostupni istraživačima zbog svoje veličine i formata ili zbog tehničkih uvjeta pristupa (npr. dokumenti pisani pisačim strojem ili fotografije su dostupnije za pregled od 8 mm filma ili mikro-filma)?
- Koliki je postotak gradiva za koje institucija koja planira digitalizaciju ima autorska prava ili licence te koliko je gradiva u javnom vlasništvu?
- Koliki je postotak gradiva koje nema nikakvih ograničenja za korištenje niti je zbog sadržaja osjetljivog karaktera (primjerice dokumenti koji mogu ugroziti nečiju privatnost, ugroziti ili uzrokovati neželjeni publicitet, oklevetati ili osramotiti, kao i dokumenti koji imaju ograničenja donatora)?
- Koje gradivo ima najveću novčanu vrijednost te koje je najbolje osigurano?
- Za koje se gradivo procjenjuje da postoji najveći rizik od propadanja ili uništenja? Koje je od tog gradiva još uvijek dovoljno dobro očuvano da bi moglo biti podvrgnuto digitalizaciji bez opasnosti od oštećenja te koje je gradivo već više puta fotografirano?
- Koje je gradivo najčešće korišteno te kako ga se koristi?
- Koje se gradivo nalazi samo u toj instituciji, te je po tome jedinstveno?

Važno je da članovi povjerenstva za odabir gradiva imaju ova pitanja na umu prilikom predlaganja gradiva. Također treba napomenuti da povjerenstvo u sljedećem koraku procjenjuje sve prijedloge, kako one koji predlažu da se dotično gradivo digitalizira, tako i one koji predlažu da se to ne učini.

1.2. PROCJENJIVANJE GRADIVA

Nakon predlaganja gradiva povjerenstvo za odabir treba procijeniti koje bi gradivo trebalo digitalizirati. Stoga ono pregledava i uspoređuje listu gradiva koje je predloženo za digitalizaciju i listu gradiva koje je predloženo da se ne digitalizira. Kako svi članovi povjerenstva samostalno pregledavaju gradivo i sačinjavaju takve liste nužno je utvrditi koje je to gradivo koje su svi članovi komisije predložili za digitaliziranje, te postoji li gradivo koje se nalazi na obje liste. Ako postoji, onda o tom gradivu povjerenstvo treba diskutirati i donijeti zajedničku odluku. Tako, primjerice neki materijali mogu biti kvalificirani kao oni koje bi trebalo digitalizirati, ali za njih još nije dobiveno autorsko pravo. Ako se smatra da bi takva dozvola trebala uskoro biti dobivena, onda takve materijale treba ostaviti u pričuvi te ih digitalizirati kad se dozvola pribavi. Za takvo gradivo potrebno je sastaviti posebnu listu kako bi ga povjerenstvo moglo još jednom uzeti u razmatranje kad svi uvjeti budu zadovoljeni.

Nakon završetka ove faze gradivo koje je predloženo za digitalizaciju je procijenjeno te je odabrano ono koje treba digitalizirati.

1.3. ODREĐIVANJE PRIORITETA

Proces digitalizacije je dugotrajan te stoga treba odrediti kojim redoslijedom će se gradivo digitalizirati. Određivanje prioriteta treba sagledavati kroz prizmu vrijednosti, rizika i predviđenog korištenja gradiva.¹¹

¹¹ Vogt-O'Connor, Diane, Reformatting for Preservation and Access: Prioritizing Materials for Duplication, *Conserve O Gram*, National Park Service, Department of the Interior, Washington DC, SAD, srpanj 1995., br. 19/10, <<http://www.nps.gov/history/museum/publications/conserveogram/19-10.pdf>>, 29. siječnja 2001.

1.3.1. Vrijednost

Gradivo koje je predviđeno za digitalizaciju treba imati jednu od sljedećih vrijednosti, uz napomenu da svaka od njih ima tri stupnja vrijednosti – visoku, srednju i nisku:

- *Informacijska vrijednost* odnosi se na gradivo čiji je sadržaj izravno vezan uz cilj koji institucija želi postići cijelim projektom digitalizacije.
- *Administrativna vrijednost* je vrijednost koju ima gradivo koje se koristi za redovito poslovanje institucije.
- *Artefaktna ili stvarna vrijednost* je ona koju gradivo ima samo po sebi, zbog materijala od kojeg je napravljeno, veličine, unikatnosti ili čega drugoga.
- *Pridružena vrijednost* se odnosi na gradivo koje se povezuje s istaknutim osobama, mjestima slavnih događanja, poznatim grupama i udruženjima itd.
- *Dokazna vrijednost* je vrijednost onog gradiva koje može poslužiti kao pravni ili povijesni dokaz nekog događaja.
- *Novčana vrijednost* je ona vrijednost koju bi gradivo postiglo na tržištu. Naravno, ova se vrijednost mijenja ovisno o stanju na tržištu.

1.3.2. Rizik

Rizik se može promatrati kroz pravni, sociološki ili arhivistički aspekt. S obzirom na to da je povjerenstvo za odabir u prethodnim koracima razmotrilo pravni i sociološki aspekt gradiva, rizik se ovdje promatra s arhivističkog stajališta.

Pod materijale *visokog rizika* spadaju materijali koji su fizički i/ili kemijski nestabilni. To mogu biti materijali koji su u procesu raspadanja, koji prilikom raspadanja djeluju na gradivo smješteno u njihovoj neposrednoj blizini ili, pak, na ljude koji njime barataju, npr. celuloidni film na nitratnoj podlozi.

Oznaku *srednjeg rizika* nosi gradivo koje propada samim korištenjem jer trpi mehanička ili fizička oštećenja, npr. tiskane knjige, CD-ROM diskovi, indigo kopije pisama i slično.

Gradivo *niskog rizika* je ono koje je ispravno pohranjeno i ne prijete mu brzo propadanje ili je još k tome načinjeno od dugovječnog materijala, npr. vizualna građa izrađena pastelom, drvenim ugljenom, grafitom itd.

Gradivo visokog rizika uglavnom ima i visoku vrijednost, a na instituciji je da odluči hoće li digitalizirati takvo gradivo, tj. da procijeni kako će sam postupak digitalizacije utjecati na takvo gradivo, tj. može li se potencijalni rizik od oštećenja svesti na minimum ili čak posve ukloniti.

1.3.3. Predviđeno korištenje

Korištenje se predviđa u odnosu na ciljanu publiku. Materijali koji su najčešće traženi u analognom ne moraju biti najtraženiji i u digitalnom obliku. Ipak, postoji gradivo za koje se zna da je često u uporabi, bilo zbog istraživanja, citiranja ili nečeg drugog. U načelu, često korišteno gradivo ima i visoku vrijednost. Događa se, također, da neko gradivo iznenada postane vrlo zanimljivo iz prethodno nepredvidljivih razloga.

Postupak određivanja prioriteta trebao bi rezultirati tablicom koja usporedno prikazuje razine vrijednosti, rizika i predviđenog korištenja na nekoj, proizvoljno dodijeljenoj, skali. Skala bi svakako trebala imati dovoljno vrijednosti da se rezultati mogu kvalitetno rangirati (npr. 1-10), ali ipak ne previše kako bi se mogla uočiti razlika između susjednih vrijednosti. Teško bi bilo, na primjer, procijeniti razliku između vrijednosti 87 i 88 na skali 1-100. Tablica 1 prikazuje primjer takvih usporednih vrijednosti na jednostavnoj skali 1-3 (1=nisko, 2=srednje, 3=visoko) za svaku kategoriju, pri čemu se ukupne vrijednosti kreću od minimalnih 3 do maksimalnih 9. Ono što se na njoj može odmah uočiti jest da zamišljena zbirka 2 i zbirka 3 imaju jednaku ukupnu ocjenu (6) iako im se dodijeljene vrijednosti za pojedine kategorije razlikuju. Postavlja se pitanje kako u tom slučaju odrediti prioritet, tj. kako egzaktno odlučiti koju od te dvije zbirke digitalizirati najprije, a da pritom kriterij za donošenje takve odluke bude jednoznačan, dokumentiran i ponovljiv u drugim sličnim situacijama. Tom se problemu može prići tako da se odrede prioriteti kategorija: rizik kao najvažniji element, zatim vrijednost te na kraju korištenje. Tako bi prema tom kriteriju ona zbirka koja ima višu dodijeljenu vrijednost u kategoriji rizika, a to je zbirka 2, imala prednost. Njoj se čak može u tom slučaju dodijeliti za jedan stupanj viša vrijednost, iako bi to u ovom slučaju značilo izlazak izvan skale (dodavanje vrijednosti 4 na skali 1-3), ali bi numerička vrijednost u kategoriji ukupno tada bila 7 umjesto prethodnih 6. No, tada bi zbirka 2 zapravo izbila na prvo mjesto, jer bi postojale dvije zbirke s ukupnom vrijednošću 7, a zbirka 2 bi imala višu vrijed-

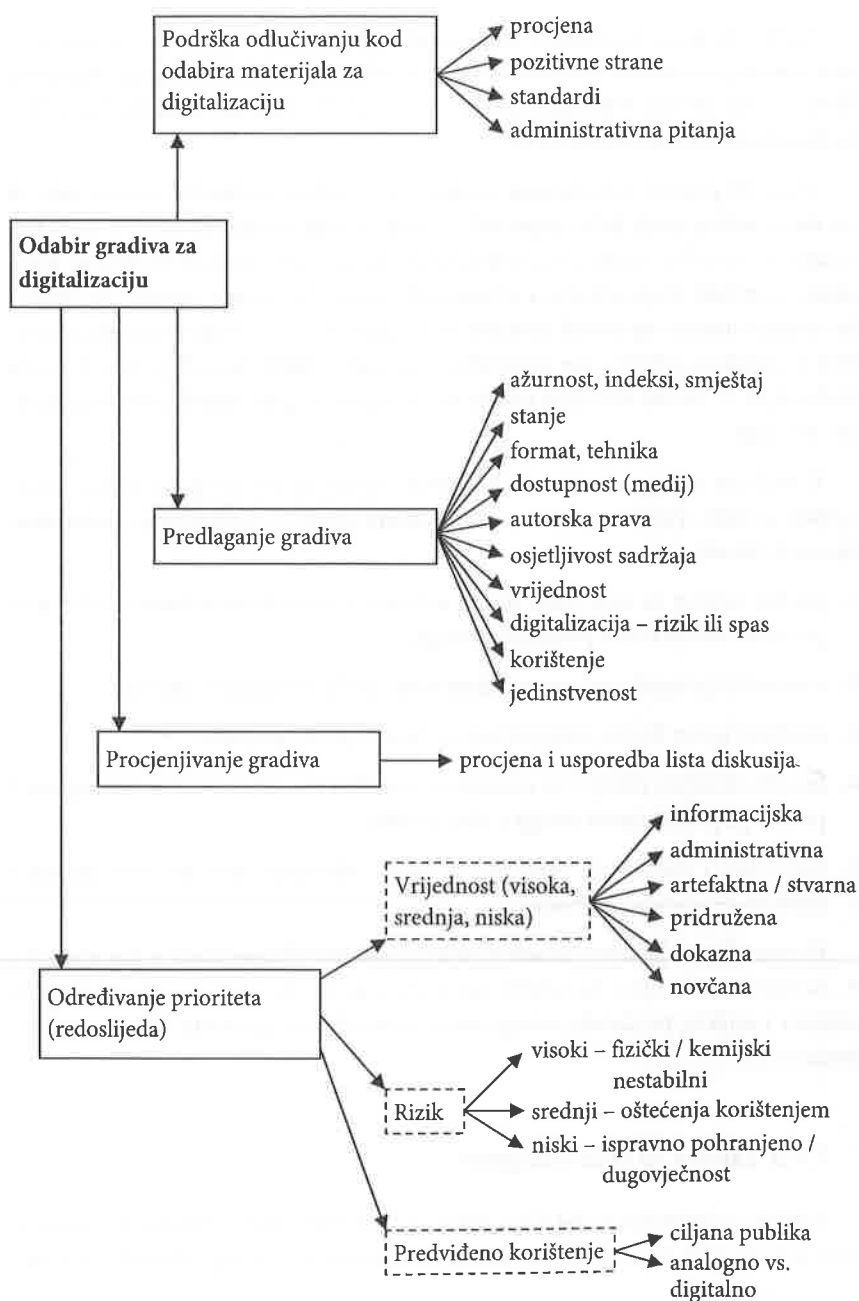
nost u kategoriji rizika. Stoga je dobro unaprijed predvidjeti ovakve situacije i odrediti procedure određivanja prioriteta u mogućim graničnim slučajevima.

Tablica 1: Tablica određivanja prioriteta za digitalizaciju

	vrijednost	rizik	korištenje	ukupno
zbirka 1	visoka (3)	srednji (2)	srednje (2)	7
zbirka 2	srednja (2)	visoki (3)	nisko (1)	6
zbirka 3	niska (1)	srednji (2)	visoko (3)	6
zbirka 4	srednja (2)	nizak (1)	srednje (2)	5
zbirka 5	niska (1)	nizak (1)	srednje (2)	4

Da zaključim, institucija bi trebala kombinirati procijenjene razine vrijednosti, rizika i predviđenog korištenja te odrediti redoslijed digitalizacije gradiva. Generalno gledano, gradivo koje svakako treba digitalizirati je ono za koje se može pretpostaviti da će biti često korišteno, dakle koje ima visoku vrijednost te čija se razina rizika može smanjiti na najnižu moguću zadovoljavajuću razinu. Ovime je završen proces odabira gradiva koji će se digitalizirati. Pregled cjelokupnog procesa odabira gradiva prikazan je na grafikonu 1.

Grafikon 1: Pregled procesa odabira gradiva za digitalizaciju



1.4. DIGITALIZACIJA UNUTAR ILI IZVAN INSTITUCIJE

Prije nego što se krene u samu digitalizaciju potrebno je odlučiti hoće li se ona obavljati unutar institucije ili će se povjeriti vanjskim davateljima tih usluga, tj. posebnim studijima koji se bave digitalizacijom. Bez obzira na to gdje će institucija obavljati digitalizaciju gradiva, ona treba prethodno točno odrediti koji je cilj digitalizacije i koji format digitalnoga gradiva zadovoljava taj cilj. Zbog toga svakako treba unutar institucije napraviti pilot projekt kojim će se eksperimentirati na malom uzorku gradiva bez obzira na to gdje će se poslije digitalizacija obavljati. Takav pilot projekt ne mora nužno dovesti do rezultata na temelju kojih se mogu donijeti preporuke za cijeli projekt. Naime, uvijek će se bar neki dio digitalizacije gradiva obavljati unutar institucije pa je zbog toga dobro na vrijeme steći dragocjeno iskustvo. Za vrijeme same digitalizacije treba pomno dokumentirati sve što je napravljeno kako bi se poslije olakšalo održavanje digitalnoga gradiva.

U načelu postoji dvojba o tom je li bolje digitalizirati unutar institucije ili izvan nje. Svaki od tih dvaju pristupa ima svoje pozitivne i negativne strane, a kako će u konačnici institucija organizirati postupak digitalizacije ovisi o nizu specifičnosti vezanih uz svaku pojedinu instituciju.

Digitalizacija unutar institucije ima nekih prednosti, kao što su viši stupanj izravne kontrole nad digitalizacijom, različitost poslova koji se mogu obavljati, učinkovitost i ekonomičnost. Ona je najuspješnija kad je riječ o relativno malom projektu koji je jednostavno provesti unutar predviđenog vremenskog roka ili, pak, koji je moguće podijeliti u manje cjeline, kad institucija ima stručnjake ili osoblje koje je zainteresirano za projekt i voljno učiti te ima dovoljno sredstava za njihovo obrazovanje i, na kraju, kad institucija već posjeduje opremu za digitalizaciju ili ima dobre izvore financiranja pa je može nabaviti, naravno, imajući na umu da hardver i softver vrlo brzo zastarijevaju.¹²

Digitalizacija izvan institucije prije svega ima prednosti s financijske i tehničke strane. Tako institucija, ako se odluči za ovu vrstu digitalizacije, neće morati osigurati prostor za skeniranje, neće se morati neprestano brinuti oko nabavke nove opreme, neće se morati baviti upošljavanjem ili školovanjem stručnjaka, studio za digitalizaciju brinut će se oko eventualnog pada sustava, kašnjenja u planu i ispravaka pogrešaka, institucija će profitirati produktivnošću

¹² Gertz, Janet, Vendor Relations, u: *Handbook for Digital Projects*, n. dj., str. 152.

unajmljenog digitalizacijskog studija, a cijena će biti poznata na početku cijelog procesa.¹³

1.4.1. Kako odabrati vanjskog pružatelja usluga?

Prije svega treba konstatirati da se vanjski pružatelji usluga razlikuju čak i kad se koriste istom opremom i metodama rada. Dobri pružatelji usluga trebali bi biti iskreno zainteresirani za to što je instituciji važno i kako zadovoljiti njezine potrebe, dok drugi to smatraju samo jednim od poslova u nizu, a treći pokušavaju prodati sustave koje su sami izradili umjesto da pružaju usluge digitalizacije.

Odabrati kvalitetnog vanjskog pružatelja usluga digitalizacije nije brz niti jednostavan proces, pogotovo kad je riječ o velikom ili kompleksnom projektu. Institucije ne moraju proći kroz sve dolje navedene korake, ali će ipak morati proći neke osnovne. Odabir bi trebao slijediti ove korake (neki će od njih poslije biti pobliže objašnjeni):

1. razviti početni koncept projekta i njegovih ciljeva;
2. pronaći potencijalne vanjske pružatelje usluga;
3. poslati zahtjev za osnovnim informacijama kako bi se odredili oni pružatelji usluga koji su zainteresirani i znaju ili imaju ideju kako odraditi zamišljeni projekt; u tom bi zahtjevu trebali biti jasno objašnjeni ciljevi projekta;
4. odrediti metodologiju projekta i razinu potrebne kvalitete;
5. izraditi uži popis potencijalnih vanjskih pružatelja usluga;
6. izraditi zahtjev za ponudom i poslati ga pružateljima usluga s užeg popisa. Uz zahtjev, ako je to dopušteno, priložiti uzorke koji će se skenirati;
7. komunicirati s pružateljima usluga za vrijeme njihove izrade ponuda u što mogu biti uključeni sastanci i vaš posjet mjestu gdje će se digitalizacija fizički provoditi;
8. usporediti i vrednovati primljene ponude te odabrati najbolju;
9. izraditi i potpisati ugovor;
10. surađivati s vanjskim pružateljem usluga za vrijeme odvijanja projekta.

¹³ Gertz, Vendor Relations, n. dj., str 152.

1.4.2. Razvoj projekta i njegovih ciljeva

Institucija mora odrediti što želi postići i kako. Na primjer, je li cilj digitalizacije očuvanje izvornika, izrada vizualnog indeksa slikovnih materijala, pretraživa elektronička verzija dokumenata koji se čuvaju ili, pak, neka kombinacija poboljšanja dostupnosti i očuvanja.

U načelu postoje dvije krajnje situacije: prva je kad institucija zna što želi, ali ne mora nužno znati kako to postići, a druga je kad institucija zna što želi i poznaje sve tehničke detalje do u najmanji detalj. U prvom slučaju institucija treba detaljno opisati svoje zahtjeve i od vanjskih pružatelja usluga zatražiti da predlože rješenje unutar općenitih zahtjeva same institucije. U drugom slučaju, potrebno je detaljno odrediti sve parametre i pronaći pružatelja usluge koji ih može zadovoljiti. U većini slučajeva institucije se nalaze negdje između ove dvije krajnje situacije.

U svakom slučaju, institucije bi trebale vanjskim pružateljima usluga jasno opisati što žele. Tada se opisani proces odabira svodi na modifikaciju prethodno opisanih koraka:

1. poslati zahtjev za osnovnim informacijama s općenito objašnjenim ciljevima projekta većem broju pružatelja usluga;
2. razmotriti prispjele informacije koje nude različite pristupe rješenju;
3. odabrati jedan ili dva pristupa koji se čine najprikladnijima;
4. izraditi detaljan zahtjev za ponudom s izraženim prihvatljivim pristupima i poslati ga pružateljima usluga s užeg popisa;
5. usporediti i vrednovati primljene ponude s detaljno opisanim postupcima i cijenom te odabrati najbolju.

U nastavku se detaljno razrađuje struktura zahtjeva navedenih u koracima 1 i 4. Zahtjevi su detaljno razrađeni, no svaka institucija koja se nađe u situaciji odabira vanjskog pružatelja usluga može odabrati one elemente koje smatra relevantnima.

1.4.3. Zahtjev za informacijama

Zahtjev za informacijama služi stjecanju osnovne ideje o mogućim pristupima i određivanju potencijalnih vanjskih pružatelja usluga digitalizacije. Ako in-

stitucija zna koji je cilj projekta digitalizacije, ali nije sigurna koja je najbolja metodologija provođenja tog postupka, ili koju je najbolju računalno programsku opremu koristiti, ili, pak, koje metapodatke pridružiti i prema kojem standardu, tad ona može iskoristiti zahtjev za informacijama kako bi se informirala o mogućnostima koje joj stoje na raspolaganju. Ovaj se pristup pokazao kao vrlo dobar u slučaju da se kontaktiraju oni vanjski pružatelji usluga s velikim iskustvom u provođenju takvih projekata. Ovakav bi se zahtjev trebao sastojati od:

- kratkog opisa predloženog projekta uključujući informaciju o (potencijalnom) budžetu, vremenu provođenja i željenim rezultatima,
- opisa bilo koje metodologije provođenja projekta o kojoj institucija razmišlja,
- zahtjeva vanjskom pružatelju usluga da komentira metodologiju,
- poziva vanjskom pružatelju usluga da predloži alternativna rješenja postizanja istih rezultata te
- zahtjeva vanjskom pružatelju usluga da se izjasni želi li kasnije predati ponudu.

Što se tiče zahtjeva za informacijama, sugerira se kontaktirati što više potencijalnih vanjskih pružatelja usluga za koje se smatra da bi mogli biti zainteresirani za zamišljeni projekt. Njihovi odgovori omogućavaju usporedbu različitih pristupa, procjenu kvalitete skeniranja u različitim uvjetima, ponovnu procjenu izvornih zamisli o rezultatima projekta u svjetlu novih saznanja te mogućnost preoblikovanja pojedinih segmenata projekta.

Postoje mnogi načini određivanja potencijalnih vanjskih pružatelja usluga. Jedan od najboljih načina jest zatražiti preporuke od institucija koje su se već odlučile za digitalizaciju izvan institucije na sličnim projektima. Rezultati projekata koji se provode ili koji su već završeni mogu pružiti mnogo važnih informacija. Prisustvovanje konferencijama ili radionicama slične tematike, diskusijske liste, web stranice institucija koje provode digitalizaciju ili su orijentirane prema očuvanju tim postupkom mogu biti značajan izvor informacija.

1.4.3. Zahtjev za ponudom

Za razliku od zahtjeva za informacijama, zahtjev za ponudom trebao bi biti osmišljen tako da potencijalnim vanjskim pružateljima usluga detaljno objasni zahtjeve i specifikacije projekta, kriterije prema kojima će se procjenjivati prijedloge ponude i specifikacije kako bi njihove ponude trebale izgledati. Izrada

ovakvoga zahtjeva može biti kompleksna i dugotrajna. Primjeri zahtjeva za ponudom od dvjestotinjak stranica mogu se naći na web stranicama američke Kongresne knjižnice i Research Libraries Group (RLG). Oni mogu poslužiti za izdavanje onih segmenata koje institucija smatra korisnim i primjerenim njihovom osmišljenom projektu. Dobra strana osmišljavanja detaljnog zahtjeva za ponudom jest ta da se on poslije može relativno jednostavno pretvoriti u ugovor s odabranim vanjskim pružateljem usluga.

Cilj dobro osmišljenog zahtjeva za ponudom trebao bi biti dobivanje što preciznijih i po zajedničkim traženim kategorijama međusobno što više usporedivih ponuda kako bi odabir najboljeg, ali ne nužno i najpovoljnijeg, vanjskog pružatelja usluga bio što objektivniji. Upravo zbog toga precizno izrađen zahtjev za ponudom može poslužiti institucijama da lakše izbjegniju odabir najpovoljnijeg ponuđača ako on ne može zadovoljiti tražene specifikacije.

Zahtjev za ponudom treba biti jasan i eksplicitan, s tehničkim opisom izgleda ponude i informacije o tome kako će se ona vrednovati. Zahtjev bi trebao biti dovoljno široko postavljen tako da ostavlja vanjskim ponuđačima usluga mogućnost prijedloga drugačijih rješenja od onih koja su zatražena, ali istovremeno dovoljno usko postavljen kako bi se osiguralo razumijevanje standarda i zahtjeva kojih se mora pridržavati. Zahtjev za ponudom treba razdijeliti u cjeline koje se tiču tehničkih zahtjeva, zahtjeva vezanih uz upravljanje projektom, cijene i slično. Gdje je to moguće, od ponuđača treba zatražiti ponudu u standardnom formatu kako bi se omogućila njihova lakša i jednostavnija usporedba.

Sadržaj zahtjeva za ponudu se može sastojati od sljedećih elementa:

- opis projekta i njegovih glavnih ciljeva,
- što detaljniji opis gradiva (objekata) koji se namjeravaju digitalizirati kako bi se potencijalnim vanjskim pružateljima usluga pomoglo u izradi ponude,
- količina, fizičko stanje i dimenzije gradiva,
- opis pojedinih grupacija gradiva: jesu li svi njezini dijelovi istog fizičkog stanja, istih dimenzija; ako je potrebno, detaljno opišite svaku pojedinu grupaciju,
- opis proporcija: s kojim je gradivom jednostavno/teško manipulirati,
- važnost sadržaja gradiva: postoji li redoslijed kojim se gradivo mora digitalizirati? Može li se gradivo grupirati prema veličini ili vrsti, ili se, pak, gradivo različitih veličina i vrsta mora digitalizirati određenim redoslijedom?

- poznavanje jezika: koje jezike vanjski pružatelj usluga mora poznavati kako bi mogao uspješno digitalizirati gradivo (npr. tekst ili brojeve stranica),
- detaljne upute o preferiranim rezultatima – rezolucija, tonalitet, bitna dubina, formati zapisa, kompresija, operativni sustav, sustavi ili mediji za pohranu,
- upute o stvaranju digitalne master-verzije i njezinih derivata u nižim rezolucijama,
- određenje zahtijevane razine točnosti i kako će je institucija procjenjivati,
- upute o imenovanju datoteka i metapodacima,
- upute o tome treba li u imena datoteka ručno upisivati neke identifikacijske podatke ili neku drugu vrstu informacija,
- upute o bilježenju tehničkih (procesnih) podataka – podatci o korištenoj opremi (npr. koji je skener korišten, njegove postavke, korišteni profili boje, datum skeniranja, vrsta filma u slučaju mikrofilmiranja i naknadne digitalizacije s mikrofilma itd.),
- upute o stvaranju baze podataka digitaliziranoga gradiva i pripadajućih metapodataka ili kodiranju digitaliziranoga gradiva i pripadajućih metapodataka na način da budu prihvatljivi za ulaz u digitalni arhiv (oblikovanje dostavljenog informacijskog paketa (engl. *Submission Information Package – SIP*) prema OAIS standardu – ISO 14721),
- upute o isporuci digitaliziranoga gradiva – tjedno, mjesečno, određivanje rokova pojedinih faza projekta,
- upute o rukovanju i transportu – sigurnost, osiguranje, specifični zahtjevi vezani uz prijevoz originala (npr. maksimalno dopušteno osvjetljenje i sl.)
- ime kontaktne osobe u instituciji koju vanjski pružatelj usluge može kontaktirati ako bude imao dodatnih pitanja, tj. kojoj detaljna ponuda mora biti poslana i u koliko primjeraka.

Dodatno se od vanjskog pružatelja usluge digitaliziranja može zatražiti:

- navođenje računalno-programске opreme koju namjerava koristiti,
- specificiranje procedure kontrole kvalitete koju primjenjuje,
- opis produkcijskog kapaciteta i dokumentiranje je li u mogućnosti izvršiti preuzeti posao u definiranoj kvaliteti unutar definiranih vremenskih rokova,

- opis mogućnosti prijevoza gradiva za digitalizaciju i zatim dostave digitaliziranoga gradiva (mediji za pohranu, FTP prijenos na server i sl.),
- opis kontrole okolinskih uvjeta ako to zahtijeva gradivo koje se namjerava digitalizirati,
- ime i kvalifikacije osobe koja će u njihovoj instituciji biti zadužena za vođenje projekta,
- reference – prijašnje uspješno završene poslove odrađene za druge arhive, knjižnice, muzeje ili slične institucije,
- skeniranje reprezentativnog uzorka svih vrsta gradiva koje se namjerava digitalizirati, uključujući primjerke gradiva koje je jednostavno i složeno digitalizirati (ako su originali vrlo vrijedni ili vrlo osjetljivi, uzorak bi se trebao sastojati od razumno sličnih primjeraka koji su manje vrijednosti ili potrošni),
- odgovor s konkretnim prijedlogom cijene,
- navođenje cijene u točno određenim jedinicama mjere, npr. po stranici, po slici ili što god je primjereno,
- navođenje cijene medija za pohranu, prijevoz, osiguranje i bilo koje druge dodatne troškove,
- izjavu o tome vrijede li cijene u ponudi za cijeli period odvijanja projekta,
- prijedlog alternativnih metoda kojima se projekt može izvršiti uz istu razinu kvalitete.

Komunikacija s potencijalnim pružateljima usluga svakako će zahtijevati komunikaciju s njima za vrijeme izrade ponude. Konkretna i oštroumna pitanja mogu pomoći u boljoj definiciji projekta digitalizacije.

Ovisno o uvjetima, a pogotovo ako pružatelji usluga budu digitalizirali vrijedne originale, vanjski pružatelji usluga mogu se pozvati na sastanak u instituciju kako bi vidjeli originale i postavili određena pitanja prije nego što pošalju svoju ponudu.

Ako je moguće, tražitelj usluge može posjetiti vanjske pružatelje usluga kako bi se uvjerio u postojanje potrebnog kapaciteta i dovoljnog broja ljudi za izvršenje posla te jesu li prostori u kojima će se digitalizacija odvijati primjereni i dobro održavani.

1.4.4. Kontrola kvalitete

Digitaliziranom gradivu potrebno je kontrolirati kvalitetu bez obzira na to je li ono digitalizirano unutar ili izvan institucije. Pregledavanje svakog slikovnog zapisa, preslušavanje svakog zvučnog zapisa i slično zahtijevalo bi prevelike troškove ljudske radne snage. Stoga je preporučljivo, ako je digitalizacija obavljena izvan institucije, pregledati sve datoteke samo u početku njihova pristizanja u instituciju kako bi se otklonile eventualne početne pogreške u postavkama, kvaliteti isporučenih materijala itd. Nakon inicijalne provjere, kvaliteta se može kontrolirati na uzorku od 10-15% ukupno digitaliziranih materijala postupkom slučajnog odabira. Ako se zamijeti ponavljanje greške određenog tipa na slučajnom uzorku, tada se može provesti temeljna analiza samo onog dijela digitaliziranoga gradiva na kojem se detektirana greška češće pojavljivala kako bi se utvrdio opseg gradiva za koji će eventualno biti potrebno ponoviti postupak digitalizacije.

2. DIGITALIZACIJA GRADIVA

2.1. Uvod

Gradivo koje je određeno za digitalizaciju, generalno gledano, može biti tekstualno, slikovno, zvučno, video ili trodimenzionalno (3D). Vrsta gradiva, njegove fizičke dimenzije i osjetljivost uvjetuju odabir opreme i postupke njegove digitalizacije. Stoga će u nastavku najprije biti obrađeni uređaji za digitalizaciju spomenutih vrsta gradiva. Potom, s obzirom na to da se postupci digitaliziranja prilagođavaju gradivu, analizirat će se postojeći postupci, tim više što ih postoji nekoliko za istu vrstu gradiva. Svaki od njih ima svoje prednosti i nedostatke ovisno o tome kako je postavljen cilj projekta digitalizacije te koliko je novčanih sredstava i vremena na raspolaganju. Postupci digitalizacije bit će objašnjeni za svaku vrstu gradiva posebno.

2.2. UREĐAJI ZA DIGITALIZACIJU

Tekstualno i slikovno gradivo digitalizira se skenerima i digitalnim fotoaparatima, dok se za zvučno i video gradivo koriste hardverski dodaci računalima ili, pak, zasebni uređaji specijalne namjene. Trodimenzionalni objekti mogu, pak, biti digitalizirani skenerima i digitalnim fotoaparatima, ako se kao rezultat želi dobiti njihova slika, ili 3D skenerima, ako se želi dobiti trodimenzionalni računalni model.

2.2.1. Skeneri

Skeneri se mogu podijeliti u dvije općenite skupine: koračni i protočni. *Koračni skeneri* su oni skeneri koji bez ljudske intervencije mogu skenirati samo ono pojedinačno gradivo koje je postavljeno na površinu za skeniranje. Nasuprot njima, *protočni skeneri* sami izmjenjuju gradivo za skeniranje (npr. raznim obli-

cima uvlakača papira ili tehničkim rješenjima za samostalno okretanje stranica) čime se proces digitalizacije automatizira.

Koračni skeneri se, nadalje, dijele na ručne, plošne, skenere za mikrooblike, rotacione, reprografske te 3D skenere. Treba, ipak, napomenuti da se neki koračni skeneri mogu nadograditi uvlakačima papira ili mikrooblika te tako postati protočni.

Prije početka postupka digitalizacije potrebno je zatražiti savjet konzervatora ili osobe zadužene za zaštitu građe kako bi se izbjegla moguća oštećenja za vrijeme digitalizacije – bilo zbog nestručnog rukovanja bilo zbog prevelikog izlaganja osvjetljenju.

2.2.1.1. Koračni skeneri

2.2.1.1.1. Ručni skeneri

Ručni (engl. *handheld*) skeneri, mogu biti različitih oblika, a sliku stvaraju prelaženjem preko izvorne građe. Za rukovanje njima potrebno je konstantno i jednolično pomicanje skenera, pri čemu je mirna ruka od presudnog značenja. Ovi skeneri daju relativno loše rezultate kada je riječ o boji. Iako su najjeftiniji, njih sve više zamjenjuju jeftini plošni skeneri te su danas rjeđe dostupni na tržištu.

2.2.1.1.2. Plošni skeneri

Plošni skeneri su najčešći uređaji na tržištu. Većina korisnika ovu vrstu skenera identificira s pojmom skener. Ponekad se naziva stolnim ili refleksnim skenerom.

Najčešći, A4 i A3 modeli relativno su jeftini i ne zahtijevaju veliko znanje za rukovanje njima. Nakon uspostavljanja i uhodavanja radnog procesa, s tim se skenerima može brzo raditi. Veći modeli (do A0 formata) vrlo su skupi te su isplativi jedino u dugoročnim projektima koji imaju velik opseg rada i digitaliziraju velike formate izvornika.

Plošni skeneri uspješno skeniraju sve vrste dvodimenzionalnih plosnatih predmeta: dokumente, fotografije, crteže i grafike, knjige, umjetnička djela te

blago trodimenzionalne predmete poput novčića (kovanica), nakita, botaničkih uzoraka i dr.

Plošni skeneri mogu se nadograditi dodatnom opremom za protočno skeniranje (uvlačać papira) ili skeniranje prozirnog gradiva/mikrooblika (prosvjetlivač). Dodavanjem adaptera za prozirno gradivo (engl. *Transparent Media Adapter – TMA*), koji se obično montira na plošni skener umjesto standardnog poklopca, ovim se skenerom može skenirati prozirno gradivo. Dijapozitivi ili negativi u tom se slučaju nalaze na staklenoj ploči, često u odgovarajućem okviru. Zbog male veličine izvornika, rezolucija koju postiže plošni skener obično je niža nego kod skenera za filmove, što daje i nižu kvalitetu rezultirajuće slike.



Slika 1: Plošni skener



Slika 2: Skener s prosvjetlivačem

2.2.1.1.3. Skeneri za mikrooblike

Gradivo na filmu dostupno je u različitim formatima, a najčešće se nalazi na 35 mm negativima i dijapozitivima te na mikrofilmu ili mikrofiševima. Stoga se pod pojmom skener za mikrooblike najčešće podrazumijeva uređaj namijenjen skeniranju gradiva na filmu, kao što su 35 mm dijapozitivi ili negativi. Neki skeneri mogu primiti filmove ili dijapozitive srednjeg (4" x 5") ili velikog formata.

Iako su skupi u odnosu na plošne skenere, skeneri za filmove i dijapozitive daju visokokvalitetnu digitalnu sliku. Zbog svojih malih dimenzija, 35 mm gradivo mora se skenirati u vrlo visokoj rezoluciji, što podrazumijeva skupe skenere koji nisu univerzalno primjenjivi. Alternativa skeneru za filmove je plošni skener s adapterom za prozirne medije – prosvjetlivačem (engl. *Transparent Media Adapter – TMA*). Njime se neće dobiti ista kvaliteta digitalne slike kao korište-

njem skenera za film, ali to može biti sasvim dovoljno za gradivo koje se namjerava učiniti dostupnim putem Interneta.

Nekim skenerima mogu se priključiti dodaci kao što su držač za neizrezane filmove i uvlakač (engl. *feeder*) za masovni prihvat dijapozitiva. Iako te naprave neće ubrzati proces skeniranja pojedinog snimka, ubrzat će proces ispravnog namještanja gradiva.

Kao jedna od tehnika digitalizacije dijapozitiva nameće se i snimanje digitalnim fotoaparatom spojenim izravno na leću dijaprojektora. Riječ je o uređaju koji automatizira cijeli proces – automatski izmjenjuje dijapozitive i za vrijeme prikaza automatski okida snimku na digitalnom fotoaparatu. Proces se može odvijati i ručno. Ovaj je pristup zanimljivo rješenje ako institucija već posjeduje kvalitetni digitalni fotoaparat. Sustav za digitalizaciju ovim pristupom prikazan je na sljedećoj slici.

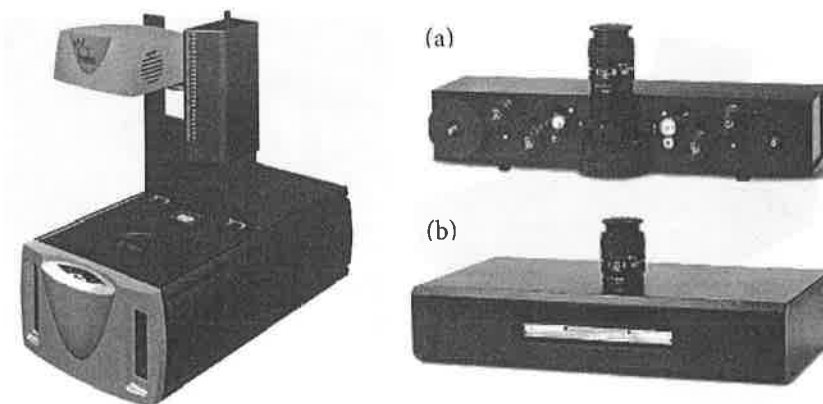


Slika 3: Digitalizacija dijapozitiva digitalnim fotoaparatom

Postoje i posebni skeneri za digitalizaciju mikrofilma ili mikrofiša s dodacima na koje se, na primjer, postavlja rola mikrofilma i koji automatiziranim prematanjem skeniraju cijelu rolu mikrofilma u jednom prolazu. Ovakvi uređaji usko specijalizirane namjene vrlo su skupi.

Općenito, skeneri za mikrooblike isplativi su ako postoji velika količina građe koju treba skenirati. Iako su skuplji od plošnih, skeneri za filmove dat će

veću kvalitetu slike jer je proces skeniranja prilagođen upravo mediju za koji su namijenjeni.



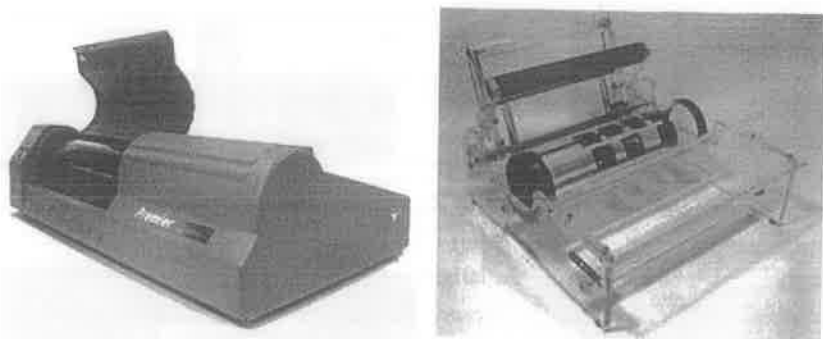
Slika 4: Skener za mikrooblike s dodacima za mikrofilm u roli (a) i mikrofiše (b)

2.2.1.1.4. Rotacioni skeneri

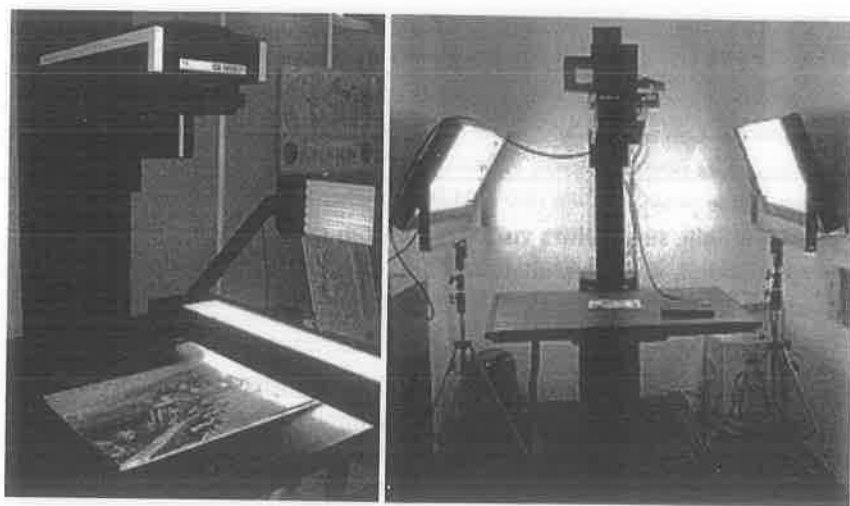
Rotacione skeneri (engl. *drum scanner*) većinom rabe profesionalni studiji za digitalizaciju. Oni daju slike visoke kvalitete iz kojih se mogu napraviti velika povećanja slika (ispis na velike formate). Izvorna se slika postavlja na cilindar ili bubanj koji se potom okreće oko središnjeg mehanizma. Svakim prolazom sken postaje sve finiji, što rezultira visokokvalitetnom slikom. Skeneri u obliku bubnja vrlo su skupi, ali daju rezultate koje niti približno ne nude stolni skeneri, kao što su izravna konverzija u CMYK, autoizoštravanje, veći dinamički raspon i veću površinu slike.

Ovom vrstom skenera moguće je skeniranje isključivo onog gradiva koje se nalazi u zasebnim, savitljivim listovima ili na filmu (dijapozitivi, negativi). Razlog tomu leži u činjenici da se gradivo prije skeniranja fiksira na bubanj, pri čemu ono mora čvrsto prilijegati uz njegovu zakrivljenu površinu zbog velike brzine vrtnje bubnja prilikom skeniranja. Stoga ono gradivo koje je uvezano, presavinuto ili se, pak, nalazi na tvrdim i debljim podlogama nije prikladno za ovakvu vrstu skeniranja.

Prije skeniranja ovakvim skenerima svakako je potrebno zatražiti stručno mišljenje konzervatora ili osobe zadužene za zaštitu gradiva.



Slika 5: Rotacioni skener i bubanj na postolju za postavljanje gradiva



Slika 6: Reprografski skeneri za digitalizaciju gradiva velikog formata

2.2.1.1.5. Reprografski skeneri

Reprografski skeneri namijenjeni su digitalizaciji gradiva velikog formata. Oni se sastoje od prostrane podloge, na koju se smješta gradivo prilikom digitalizacije, dvaju bočnih elemenata za osvjetljavanje gradiva, kako bi se uklonile moguće sjene i izbjegli odbljesci od reflektirajućih dijelova gradiva, te digitalne glave za snimanje (ponekad je na to mjesto moguće postaviti digitalni fotoaparat), koji se nalazi visoko iznad podloge i usmjeren je okomito na nju.

Ovi su skeneri vrlo praktični za digitalizaciju gradiva velikog formata ili, pak, gradiva koje je iz praktičnih razloga nemoguće skenirati drugačije, primjerice uokvirene umjetničke slike i dr.

2.2.1.1.6. 3D skeneri

3D skeneri su uređaji koji digitaliziraju predmete ili prostor analizirajući i prikupljajući podatke o njihovom obliku i, po mogućnosti, izgledu (npr. boji). Prikupljeni podatci služe za konstruiranje digitalnih 3D modela koji se mogu koristiti u različite svrhe: u industrijskom dizajnu, izradi prototipova, dokumentaciji kulturnih artefakata. Tehnologija 3D skeniranja je u pravilu vrlo skupa.

U praksi se jednim skeniranjem ne može izraditi potpuni model predmeta. Bit će potrebna brojna skeniranja iz mnogih različitih kutova da bi se dobili potpuni podatci o svim plohama predmeta.

Postoje dvije vrste 3D skenera: kontakti i nekontakti. *Kontakti 3D skeneri* sondiraju predmet fizičkim kontaktom. CMM (engl. *Coordinate Measuring Machine*) može biti vrlo precizan, no njegova je mana upravo doticaj s predmetom, što može izmijeniti ili oštetiti predmet. To je vrlo važno znati ako je riječ o skeniranju predmeta kulturne i spomeničke baštine. On je i relativno spor u usporedbi s drugim metodama.

Nekontakti 3D skeneri mogu se podijeliti u dvije glavne kategorije: aktivne i pasivne skenere. Svakoj pripada određena tehnologija. *Aktivni* skeneri emitiraju neku vrstu zračenja ili svjetla i detektiraju njegov odraz tako da sondiraju predmet ili okoliš. Mogu emitirati svjetlost (npr. laser ili bijela svjetlost s projiciranim uzorkom crnih pruga, pri čemu se predmeti snimaju stereofotogrametrijskom metodom), ultrazvuk ili rendgenske zrake. Lasersko skeniranje je metoda kojom se površina uzorkuje ili skenira korištenjem laserske tehnologije. Za digitalizaciju se koriste laseri male snage koji ne utječu na površinu predmeta.

Stoga je prilikom laserskog skeniranja potrebno paziti kolika je snaga lasera da ne dođe do oštećenja ili uništenja predmeta. *Pasivni* nekontaktni 3D skeneri mogu, na primjer, biti obični digitalni fotoaparati montirani tako da izrađuju niz uzastopnih fotografija predmeta koji se ispred njih nalazi i okreće. Nakon toga se te fotografije analiziraju u računalu i dobiva se 3D model stvarnog objekta. Nekonтактne je skenere bolje koristiti kod predložaka koji su vrlo vrijedni i/ili bi mogli biti oštećeni kontaktnim načinom skeniranja.

Nadalje, 3D skeneri mogu biti stacionarni ili prijenosni. *Stacionarni* uglavnom služe za skeniranje većih volumena koji se smještaju u njihovo radno područje, dok *prijenosni* skeneri mogu poslužiti ili za skeniranje manjih predmeta ili, pak, za digitalizaciju na terenu. Tako je, primjerice, 3D skeniranjem moguće izraditi računalne modele pokretnih i nepokretnih spomenika kulture ili, pak, arheoloških lokaliteta.

Neke od primjena:

- skeniranje arhitekture i spomenika,
- dokumentacija povijesnih, arheoloških i drugih lokaliteta,
- modeliranje i ocrtavanje lokaliteta,
- stvaranje GIS (engl. *Geographic Information System*) karti s različitim slojevima,
- utvrđivanje referentnih točaka za određivanje strukturalnih ili konstruktivnih promjena zbog izloženosti iznimnim situacijama poput potresa, požara, vibracija i sl.



Slika 7: Prijenosni 3D skener

2.2.1.1.7. Skeneri za knjige

Ovi uređaji mogu pripadati kako kategoriji koračnih, tako i kategoriji protočnih skenera. Koriste se za digitalizaciju uvezanih dokumenata koji su preveliki ili preosjetljivi da bi se mogli položiti na, primjerice, plošni skener. Kod njih se obično koriste posebne košare za polaganje knjiga, čime se izbjegava oštećenje hrpta. Neki skeneri koriste softver za uklanjanje distorzija slike koja nastaje zbog zakrivljenosti stranice.

Kod većine uređaja potrebna je ljudska intervencija za okretanje stranica. No, u novije vrijeme pojavile su se automatizirane verzije koje korištenjem vakuuma okreću stranice čak nježnije nego što to čovjek može učiniti. Pritom, korištenjem dviju digitalnih glava s dva digitalna fotoaparata visoke razlučivosti (21,1 megapiksela), postavljena okomito na predložak, baš poput onih kod reprografskih skenera, istovremeno digitaliziraju obje otvorene stranice knjige. Brzina im se kreće do 2.400 stranica po satu kontinuiranog rada, tj. 40 stranica u minuti. Riječ je o vrlo specijaliziranim uređajima te im je cijena, u skladu s tim, visoka.



Slika 8: Sustav za robotiziranu digitalizaciju uvezanog gradiva¹⁴

2.2.1.2. Protočni skeneri

Protočni skeneri nalikuju faks uređajima ili fotokopirnim aparatima budući da imaju automatski uvlakač stranica (engl. *Auto Document Feeder – ADF*). Obično se koriste za skeniranje dokumenata u kombinaciji s OCR (engl. *Optical Character Recognition*) softverom za automatsko pretvaranje slike teksta u obradivi tekst.

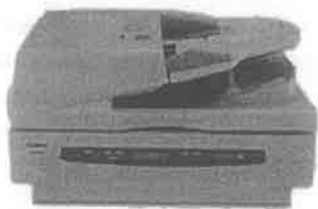
Jedna od osnovnih karakteristika ovih skenera jest njihova brzina, tj. broj stranica koje mogu obraditi u minuti. Budući da većina skenera koji se upotrebljavaju za poslovne svrhe skeniranja većih količina dokumenata ima automatski uvlakač dokumenata, brzine skeniranja kreću se od nekoliko desetaka dokumenata u minuti (takozvani *workgroup* skeneri) pa do nekoliko stotina u minuti (takozvani *production* skeneri).

Protočni skeneri imaju dvije glavne opcije – jednostrano (engl. *simplex*) ili obostrano (engl. *duplex*) skeniranje. Kod obostranog skeniranja, prilikom jednog prolaska skeniraju se obje strane papira, čime se znatno ubrzava proces skeniranja dvostranih dokumenata. Brzina skeniranja u karakteristikama skenera

¹⁴ Robotizirani sustav tvrtke Kirtas Technologies Inc., <<http://www.kirtastech.com>>, 30. travnja 2007.

navodi se u broju stranica u minuti (engl. *pages per minute – ppm*) ili broju slika u minuti (engl. *images per minute – ipm*). Kod skenera s mogućnošću jednostranog skeniranja ta su dva podatka identična, dok kod skenera s mogućnošću obostranog skeniranja broj slika u minuti može biti do dvostruko veći od izraženog broja stranica u minuti (podrazumijevaju se listovi).

Suvremeni skeneri s automatskim uvlakačem stranica imaju mogućnost detekcije uvlačenja više dokumenata odjednom, što bi inače dovelo do zastoja u skeniranju. Stariji mehanički sustavi nisu bili prikladni u situacijama u kojima se skeniraju dokumenti različitih dimenzija i debljina papira. Suvremeni sustavi temeljeni na ultrazvučnoj detekciji uspješno detektiraju uvlačenje više dokumenata odjednom, čak i kod dokumenata različitih fizičkih karakteristika. Nekorištenje ovih sustava ili njihova loša podešenost rezultirat će duljim vremenom potrebnim za skeniranje, budući da se ručno mora provjeravati broj dokumenata ubačenih u skener i konačan broj skenova te nalaženje i ponovno skeniranje dokumenata koji su eventualno preskočeni ili nisu bili pravilno skenirani. Suvremeni skeneri dodatno mogu biti opremljeni sustavima za detekciju spajalica koje nisu bile uklonjene prije ubacivanja u skener, a koje bi u slučaju dolaska do mehanizma za skeniranje mogle prouzročiti oštećenje skenera i dokumenata. Komponente sustava za uvlačenje dokumenata podložne su trošenju te ih je nakon određenog broja skeniranih dokumenata potrebno čistiti ili mijenjati. Učestalost održavanja tih komponenti, osim o količini skenova, ovisi i o kvaliteti papira koji se skenira. Predviđeni dnevni ili mjesečni opseg skenera jest veličina koja može unaprijed dati do znanja je li riječ o običnom protočnom skeneru ili sofisticiranim skeneru za masovno skeniranje.



Slika 9: Skener s uvlakačem

Ovdje treba napomenuti da se u grupaciju protočnih skenera, također, mogu ubrojiti i skeneri za knjige koji imaju mogućnost automatiziranog okretanja stranica, te skeneri za velike formate gradiva, primjerice prostornih planova, koji mogu imati slična rješenja za okretanje listova.

2.2.1.3. Karakteristike skenera

Brzina. Brzina varira ovisno o skeneru. Ako se mora skenirati veća količina materijala, brzina skeniranja bit će važan faktor koji valja uzeti u obzir. Podatak o brzini skeniranja odnosi se samo na određene uvjete pod kojima se ona može postići i zapravo daje podatak o brzini rada mehaničkih elemenata skenera. Na brzinu cjelokupnog procesa skeniranja dokumenata utječu i mnogi drugi faktori, kao što su priprema dokumenata, učestalost zaglavljivanja papira kod protočnih skenera, postojanje potrebe za promjenom orijentacije prilikom skeniranja, rezolucija, korisničko sučelje itd. Također, pod vremenom potrebnim za skeniranje podrazumijeva se ukupno vrijeme potrebno za prethodni pregled (engl. *preview*) i za potpuni proces skeniranja. Važan faktor je i brzina prijenosa podataka između uređaja za skeniranje i računala.

Razlučivost (rezolucija). Razlučivost je prostorna frekvencija uzimanja uzoraka iz okoline. Ona se odnosi na broj piksela koje skener može očitati u području skeniranja. Što je razlučivost viša, to je bolja kvaliteta skenirane slike. Razlučivost se obično mjeri u točkama po inču (engl. *dots per inch – dpi*). Odabir rezolucije ovisi o potrebi za kvalitetom slike, o kapacitetu medija za pohranu te o karakteristikama softvera za optičko prepoznavanje znakova (OCR). Osim toga, viša rezolucija rezultira i sporijim radom skenera. Neki modeli skenera omogućuju dvostruki izlaz slika: jedne u crno-bijelom formatu ili skali sivih boja (engl. *grayscale*) za potrebe daljnjeg procesiranja teksta, a druge u boji za potrebe arhiviranja.

Kod skenera je važno razumjeti razliku između optičke i interpolirane razlučivosti. Ta ista razlika pojavljuje se i kod digitalnih fotoaparata i digitalnih video kamera u kontekstu optičkog tj. digitalnog zuma. Naime, *optička razlučivost* je maksimalna razlučivost koju se može postići korištenjem optike ugrađene u sam uređaj. U trenutku kada se želi postići viša razlučivost, mora se interpolirati informacija o dodatnim pikselima, koji nisu fizički detektirani, već su izračunati na temelju svojstava susjednih piksela, pa tako dolazimo do termina *interpolirana razlučivost*. Kao rezultat dobiva se slika više razlučivosti, zbog većeg broja piksela od kojih se sastoji, no ona je u pravilu niže kvalitete zbog izračunatih, a ne stvarno detektiranih piksela. Postupak korištenja interpolirane razlučivosti bi, stoga, trebalo izbjegavati kad god je to moguće, jer on predstavlja svjesno unošenje šuma u elektroničku informaciju.

Dinamički raspon. Izražava se brojčano. U slikama koje se dobiju skeniranjem, dinamički se raspon odnosi na vjernost boje i kontrast slike. Što je dinamički raspon veći, to su veći kontrast i dubina boje. Dinamički raspon pokazuje sposobnost uređaja da bilježi razlike između tamnih i svijetlih tonova, tj. sposobnost *izvlačenja* detalja iz područja sjene u slikama. Uobičajeni plošni skeneri imaju dinamički raspon oko 2,4, što odgovara nižoj kvaliteti skeniranja. Kvalitetniji skeneri imaju dinamički raspon 3,0. Skeneri za film i mikrooblike obično imaju veći dinamički raspon, a rotacioni skeneri dostižu vrijednosti do 3,8.

Polje skeniranja. Odnosi se na najveći predmet koji skener može skenirati odjednom. Većina plošnih skenera ograničena je na građu veličine 18 x 24 cm. Neki softveri mogu spajati skenirane segmente predmeta u cjelinu, čime je omogućeno skeniranje predmeta većeg formata na manjem skeneru. Takav postupak spajanja zna biti neprecizan.

Vezni uređaji. Skenerski hardverski i softverski vezni uređaji omogućuju povezivanje skenera s računalom i kontrolu rada i rezultata. Hardverski vezni uređaj je fizička veza ili kabel koji skener povezuje s računalom. Prije odluke o kupnji važno je provjeriti kompatibilnost računala i skenera. Suvremeni skeneri spajaju se na računalu uglavnom putem USB, SCSI ili FireWire veze. Osim njih, može se koristiti paralelni ili serijski priključak. *USB* (Universal Serial Bus) postao je standardnim veznim uređajem zbog jednostavnosti instaliranja, niske cijene i zadovoljavajuće brzine veze. Koriste ga i PC i Macintosh. *SCSI* (Small Computer Serial Interface) priključak izabire se uglavnom zbog veće brzine prijenosa podataka na računalu. Obično je standardan za računala Macintosh, dok PC trebati nadograditi odgovarajućom adapterskom karticom. *FireWire* (IEEE 1394) tehnologija prihvaćena je kao standard za prenošenje velikih količina podataka između računala i perifernih jedinica (uglavnom skenera i digitalnih video kamera). *Paralelni priključak* uobičajen je vezni priključak na starijim modelima ili jeftinijim skenerima. Praktički svako računalo ima paralelni priključak, iako je on za današnje potrebe uglavnom spor. *Serijski priključak* mogao se pronaći na skenerima niže kvalitete, primjerice ručnim skenerima, ali oni su uglavnom prespori i ne preporučuju se.

Softver za skeniranje. Programi za skeniranje obično sadrže opcije za kontrolu ravnoteže boje (engl. *colour balance*), krivulje tonova (engl. *tonal curves*), svjetlinu i sjene (engl. *highlights and shadows*) te izoštravanje i čišćenje slike. Kvalitetniji programi za skeniranje imaju i neke napredne opcije, kao što su automatsko prepoznavanje i preskakanje prazne stranice kod skeniranja protočnim skenerima, zatim automatsko uklanjanje dviju crnih točaka koje se pojav-

ljuju kod skeniranja dokumenata koji su bili probušeni i umetnuti u registrator, prepoznavanja orijentacije sadržaja i njegovo zakretanje, korekcija ukošenosti, automatska detekcija i uklanjanje boje pozadine i drugo. Takav softver uvelike ubrzava proces skeniranja, budući da znatno smanjuje potrebu za ručnom intervencijom za vrijeme skeniranja. Ovisno o tipu dokumenata, on skraćuje proces njihove pripreme prije skeniranja, poboljšava kvalitetu naknadnog OCR-a te smanjuje konačnu veličinu datoteke. Ovakvi programi su, pod različitim tržišnim nazivima, često integrirani u pogonski program samih skenera, ali mogu biti i u obliku zasebnih programskih rješenja.

Opseg skeniranja. Osim brzine i razlučivosti, podjednako bitan podatak kod skenera jest njegov prosječni predviđeni dnevni opseg skeniranja (ova veličina se ponekad izražava i kao mjesečni opseg skeniranja). On se u karakteristikama skenera navodi u broju stranica koje skener može skenirati na dan. O opsegu skeniranja, predviđenim planom projekta digitalizacije, ovisiti će odabir skenera. Snižavanjem potrebne razine razlučivosti, broja korištenih nijansi boje ili, pak, skeniranjem u nijansama sive skale ili u crno-bijeloj varijanti može se ubrzati broj skeniranih dokumenata u jedinici vremena. Ipak, dnevni opseg skeniranja jest veličina koja više govori o kvaliteti skenera i njegovoj izdržljivosti u kontinuiranom radu, a manje o njegovoj brzini.

Praktične napomene:

- Na plošnom skeneru skenirajte samo ono gradivo koje se neće oštetiti prilikom pritiska poklopca na tvrdu površinu ili pritiska uvezanog gradiva o staklenu površinu za skeniranje. Ako niste sigurni u to, posavjetujte se sa relevantnim stručnjacima.
- Staklena površina za skeniranje uvijek mora biti posve čista. To osigurava bolju kvalitetu slike i štiti izvornik od onečišćenja.
- Ako je to moguće, skenirajte samo one predmete koji bez problema čitavi stanu na plošni skener ili skener opremljen košarom za knjigu.
- Ako je nužno skenirati predmet u više dijelova, činite to tako da na svakom snimku ima dovoljno površine koja se preklapa s prethodnim i sljedećim snimkom jer će to omogućiti bolje spajanje slike kod naknadne obrade (korištenjem posebnog softvera).
- Testirajte skener i njegove rezultate na gradivu koje nije osjetljivo prije nego što započnete skenirati izvornike. Prilikom obučavanja osoba koje će raditi sa skenerom, rabite također gradivo koje nije osjetljivo.

- Ustanovite konvenciju za imenovanje datoteka nastalih skeniranjem, na primjer rabeći postojeći inventarizacijski sustav ili dajući im nazive koji nešto znače. Naziv datoteke trebao bi omogućiti povezivanje datoteke i izvornika.
- Prije nego što ustanovite postupak rada ili prije nego što grupirate poslove, provedite nekoliko postupaka skeniranja i obrade slike od početka do kraja kako biste se uvjerali da će krajnji rezultat radnog procesa biti onakav kakav ste zamislili.
- Skenirajte na najvišoj rezoluciji i dubini boje koje su moguće imajući u vidu cilj projekta, ograničenja skenera, uvjete pohrane podataka, značajke izvornog gradiva, ali i predviđeno korištenje gradiva (istražite rezultate skeniranja na zaslonu, na papiru i u svakom drugom formatu za kojeg očekujete da će biti korišten, npr. na mobilnom telefonu).
- Vodite računa o dnevnoj ili tjednoj izradi pričuvnih kopija te njihovu eventualnom dislociranju zbog osiguranja u slučajevima požara ili elementarnih nepogoda.
- Kalibrirajte sve međusobno povezane uređaje kako biste osigurali maksimalnu predviđenu kvalitetu. Pritom vodite računa i o snazi i boji okolin-ske rasvjete.
- Master slike moraju biti napravljene s vidljivim kontrolnim uzorcima: standardiziranim kontrolnim uzorkom s nijansama sive, s uzorcima u boji i skalom s oznakama veličine u centimetrima i inčima.

2.2.2. Digitalna kamera / fotoaparati

Početna napomena: termin *digitalna kamera* u širem smislu obuhvaća digitalne fotoaparate raznih klasa te digitalne video kamere za snimanje pokretnih slika (engl. *camcorder*). Ovi se termini često miješaju pa je neophodna terminološka preciznost.

Uporaba digitalnih fotoaparata i digitalnih video kamera sve je češća u projektima digitalizacije. Razlog tome je njihova fleksibilnost, tj. činjenica da se njima mogu snimati dvodimenzionalni i trodimenzionalni predmeti.

Prednost digitalnih fotoaparata i video kamera je neposredni rezultat. Slika ili video zapis mogu se vidjeti odmah nakon snimanja te se odmah mogu napraviti neke izmjene i korekcije ili se može ponoviti snimanje.

Za rad s digitalnim fotoaparatom ili video kamerama u studiju, potrebni su jaki izvori svjetlosti, što može dovesti do pitanja vezanih uz zaštitu nekih vrijednih predmeta kulturne baštine. Stoga je prije početka rada potrebno savjetovati se s konzervatorom ili osobom zaduženom za zaštitu građe.

Kako rade?

Umjesto na fotoosjetljivi film, digitalni fotoaparat sliku bilježi koristeći CCD (engl. *Charged Coupled Device*) čip, koji je sastavljen od velikog broja fotoosjetljivih elemenata, koji detektirane fotone pretvaraju u elektrone, te tako stvara digitalni zapis – skup podataka o položaju, nijansi boje i stupnju osvijetljenosti osnovnih (gradivnih) elemenata slike (tzv. piksela). Kao medij za pohranu, digitalni fotoaparati rabe memorijske kartice raznih standarda, dok digitalne video kamere rabe magnetske trake, mini DVD (disk promjera 80 mm, za razliku od klasičnog DVD diska čiji je promjer 120 mm) ili čvrsti disk.

Značajke digitalnih fotoaparata i video kamera (u usporedbi s analognima):

- ne troše film,
- aparati koji rabe standardne baterije mogu rabiti i punjive baterije, dok su specijalne baterije uvijek punjive,
- fotografska ili video snimka odmah se vidi (nije potrebno razvijanje) i po potrebi se može ponoviti,
- nema opasnosti da se fotografska snimka ogrebe ili prekrije točkicama prašine,
- na memorijsku karticu mogu se snimati stotine snimaka,
- memorijska se kartica nakon pohrane snimaka na računalo ili optički disk briše i ponovno rabi,
- cijena digitalne fotografske snimke je zanemariva pa se obično snimaju brojne snimke i izabiru najuspjelije,
- snimke se mogu pregledavati u velikom formatu (projekcije, monitor ili TV-ekran) bez usluga fotografa,
- snimke se mogu retuširati, modificirati, montirati i dr. na računalu neusporedivo učinkovitije nego klasičnom tehnologijom,
- fotografije se mogu otisnuti,
- digitalni fotoaparati i kamere omogućuju brojne opcije koje klasični nemaju, primjerice serijsko snimanje po nekoliko (1-3, pa i više) snimaka u

sekundi sve do popunjenja memorijske kartice, snimanje kraćih video sekvenci fotoaparatom, različite vrste prijelaza (pretapanja) iz kadra u kadar kod snimanja digitalnom video kamerom, mogućnost snimanja u potpunom mraku bez vidljive rasvjete itd,

- u pravilu i jeftini digitalni fotoaparati imaju promjenjivu žarišnu duljinu ("zum"),
- digitalni fotoaparati automatski registriraju i spremaju podatke o vremenu i postavkama snimanja,
- digitalne snimke i video zapisi mogu se neograničeno puta kopirati bez gubitka kvalitete.

Specifičnosti i nedostaci digitalnih fotoaparata i video kamera:

- digitalni fotoaparati i kamere trebaju izvor energije – punjive baterije (tzv. "akubaterije") velikog kapaciteta često nisu standardne izvedbe te su stoga znatno skuplje (posebice baterije namijenjene digitalnim video kamerama);
- punjive baterije imaju svoj vijek trajanja;
- za automatsko izoštravanje i postavu elemenata snimanja potrebno je neko (iako ne dugo) vrijeme (okidač "kasni", tj. digitalni fotoaparati su pomalo "tromi" pri okidanju);
- za pohranu fotografija na memorijsku karticu u komprimiranom obliku potreban je djelić sekunde, dok kod primjene RAW (nekomprimiranog) formata, a posebno kod spremanja video sekvenci, proces spremanja može onesposobiti aparat za snimanje novog snimka i kroz nekoliko desetaka sekundi; u nekim situacijama i malo kašnjenje može uzrokovati propuštanje prilike za snimak (snimci pokreta); bolje aparate odlikuje manje kašnjenje okidača i kraće spremanje snimka, omogućeno *burst* tehnologijom;
- digitalne fotografije kod velikih povećanja pokazuju rubne nepravilnosti na ostrim granicama i pojavu neželjenih ornamenata na inače mirnim, jednobožnim površinama (tzv. "šum" slike), osobito kod komprimiranog oblika zapisa (npr. jpg format);
- snimke u visokoj rezoluciji (koja omogućuje bolje iskorištenje oštine objektiva), osobito u TIFF ili RAW formatu, memorijski su vrlo zahtjevne; veliko zauzeće memorije nije toliko problem zbog zauzeća kartice (koriste se kartice većeg kapaciteta, koje se po potrebi mogu mijenjati "na tere-

nu"), koliko zbog dužih vremena spremanja snimka i rapidnog rasta zauzeća potrebnog arhivskog prostora;

- video sekvence snimljene digitalnim fotoaparatom često su izrazito loše kvalitete; neki aparati, čak i srednje klase, imaju ograničeno vrijeme snimke neovisno o raspoloživom kapacitetu na memorijskoj kartici;
- jeftiniji digitalni aparati, pa čak i neki iz srednje klase, pokazuju izrazitu zrnatost snimaka snimljenih pri lošijim svjetlosnim uvjetima, ponekad uz pomak boja prema plavom dijelu spektra.

2.2.2.1. Vrste fotoaparata

Osnovnom podjelom mogla bi se smatrati podjela na amatersku i profesionalnu opremu, a prijelaznu kategoriju čine poluprofesionalni ili napredni amaterski aparati, pri čemu je teško odrediti granice pojedinih kategorija. Dodatnu kategoriju čini visokoprofesionalna, studijska oprema.

2.2.2.1.1. Amaterski fotoaparati

Obično su potpuno automatizirani, bez previše mogućnosti podešavanja. Stoga su jednostavni za rukovanje, koje se često svodi na pritisak na okidač. Budući da automatika manje-više uspješno optimizira parametre pri uobičajenim uvjetima, i amaterska klasa u načelu daje korektne, pa čak i vrlo dobre snimke pri pogodnim svjetlosnim uvjetima. Pri lošim svjetlosnim uvjetima slabija kvaliteta objektiva rezultirat će mutnijim snimkama. U takvim uvjetima često se koristi imitacija visoke osjetljivosti filma, što rezultira zrnatošću slike. Često se skromna svjetlosna jakost objektiva kompenzira i dužim ekspozicijama, što uzrokuje dodatan gubitak oštine zbog nehotičnog pomicanja aparata (ako se ne koristi stativ). Nije rijetkost naginjanje boja prema plavim tonovima. Ograničene su mogućnosti korekcija u uvjetima nedovoljnog osvjetljenja, nasuprotnog osvjetljenja, jakih sjena i sl.

Ovi fotoaparati obično imaju ugrađenu bljeskalicu i na njih nije moguće priključiti dodatnu ili vanjsku bljeskalicu.

U ovisnosti o namjeravanom korištenju fotografija i amaterska se oprema može koristiti u procesu digitalizacije, npr. za prikaz slika na webu.

2.2.2.1.2. Profesionalni fotoaparati

Profesionalna oprema uz potpunu automatiku omogućuje i ručne postavke ili zadavanje pojedinih elemenata (veličine relativnog otvora objektiva, imitirane osjetljivosti filma), nudeći bolje performanse i velik broj posebnih opcija, primjerice vrhunske sustave mjerenja svjetline slike u više točaka, veću snagu i posebne režime bljeskalice, mjerni predbljesak u cilju izračunavanja ispravne ekspozicije, makro snimanje sitnih objekata sa svega nekoliko centimetara udaljenosti, veće raspone zumiranja, veći izbor mogućih kvaliteta i rezolucija zapisa uključujući snimanje u TIFF i RAW formatima, snimanje panoramskih snimaka, automatsko uzastopno snimanje više slika u sekundi, snimanje istog motiva s tri različite ekspozicije od kojih se bira najbolja i dr. Profesionalna oprema se u pravilu izvodi u SLR (engl. *Single Lens Reflex*) izvedbi, s izmjenjivim objektivom, čime je uvjetovana i refleksna izvedba tražila.

Prednost SLR-a je da je unaprijed moguće precizno provjeriti (veličinom otvora blende i kontrolirati) dubinsku oštrinu budućeg snimka. Kako svjetlosni snop mora ulaziti kroz objektiv i pri kadriranju, izmjenjivi objektivni nemaju ugrađeni zatvarač, nego se on izvodi kao "zavjesni" neposredno ispred senzora slike (CCD ili CMOS čipa). Zbog sustava leća kojima se snop svjetla upućuje ili u tražilo ili na čip, prijašnji digitalni SLR fotoaparati nisu imali mogućnost kadriranja putem LCD tražila na pozadini fotoaparata. Noviji digitalni SLR fotoaparati zaobilaze i taj problem ugrađivanjem dodatnog čipa tik uz tražilo, zbog čega je moguće kadriranje i putem pozadinskog LCD tražila (engl. *Live View*).

Važna je činjenica da digitalni SLR fotoaparati mogu rabiti i izmjenjive objekte s klasičnih aparata. S obzirom da su dimenzije senzorskog čipa daleko manje nego korisni prostor 35 mm filma, slika zabilježena na čipu sadrži samo sredinu one slike koju bi zabilježio film s istim objektivom. Drugim riječima, normalan objektiv za 35 milimetarski film postaje teleobjektiv ako se priključi na digitalnu kameru. Zbog usporedbe žarišne dužine objektiva na digitalnom aparatu s onom kod analognog, 35 mm fotoaparata, potrebno ju je množiti s faktorom koji se kreće između 1,4 i 1,8, a poželjno je da bude što manji. Koji se množitelj koristi, ovisi o konkretnoj kombinaciji digitalnog fotoaparata i objektiva s klasičnog fotoaparata.

2.2.2.1.3. Studijske kamere

Pripadaju rangu visokokvalitetnih i skupih kamera. Digitalne studijske kamere imaju posebna leđa koja mogu zamijeniti leđa s filmom unutar normalnog profesionalnog sustava studijskih kamera.

Studijski rad s digitalnom kamerom vrlo je sličan onome s analognom. Takve kamere daju sliku visoke kvalitete, velike rezolucije s kompletnim podatcima o boji. Zbog veličine rezultirajućih digitalnih snimaka, studijske kamere su često priključene na računalo. Čim se slika snimi, odmah se putem veznog sklopa (SCSI, FireWire ili USB) prenosi na računalo.

2.2.2.1.4. Profesionalni modularni sustav digitalnih kamera

Riječ je o visokokvalitetnoj profesionalnoj opremi koja se može sastaviti od modularnih komponenti kako bi služila specifičnim potrebama. Obično je riječ o opremi koju proizvode Hasselblad i Mamiya. Sustav omogućuje bilježenje mnoštva detalja i velika povećanja slike. Zbog glomaznosti, obično se koristi u studiju.

2.2.3. Uređaji za digitalizaciju zvuka

Zvučni zapisi većinom se nalaze pohranjeni na magnetskim trakama ili vinilskim pločama. Magnetske trake mogu biti u obliku audio kasete, studijskih magnetskih traka i sl. Vinilske ili gramofonske ploče postoje u pet različitih oblika. To su:

- S (engl. *single*) – promjera 17 cm, 45 o/min, dužina snimke oko 3 minuta;
- MS (engl. *maxi-single*) – promjera 30 cm, 45 o/min, dužina snimke kao i kod single ploče;
- EP (engl. *extended play*) – to su zapravo single gramofonske ploče, ali s po dvije snimljene skladbe na svakoj strani ploče;
- MC (engl. *micro-compact*) – promjera 17 cm, 33 o/min, koristile su se uglavnom za govorne snimke;
- LP (engl. *long play*) – promjera 30 cm, 33 o/min, dužina snimke oko 45 minuta.

Za digitalizaciju gradiva s navedenih nosača zvučnih zapisa potrebno je na računalo imati ili dodatno spojiti potrebnu opremu za digitalizaciju. Prije svega

to su uređaji za reprodukciju izvornog zvučnoga gradiva, primjerice gramofona, kasetofona ili magnetofona. Zatim, s računalne strane to su, u većini slučajeva postojeće, interne zvučne kartice. No, one nisu pogodne za kvalitetnu digitalizaciju jer ponekad uzrokuju šum na digitaliziranom signalu pa rezultirajući digitalizirani zvučni zapis nije dovoljno dobre kvalitete za dugoročnu pohranu u obliku arhivskog zapisa. Zbog toga je, za kvalitetnu digitalizaciju zvučnih zapisa, potrebno na računalo priključiti opcionalno najprije miksetu, a zatim vanjski pretvarač visoke kvalitete. To su obično pretpojačala na koja se spajaju uređaji za reprodukciju izvornog zvučnoga gradiva, a ona se dalje spajaju ili na pojačalo ili izravno na računalo putem USB sučelja. Sljedeća slika prikazuje prednju i stražnju stranu pretpojačala.



Slika 10: Pretpojačalo zvučnog signala (prednja i stražnja strana)
(Izvor: Harmony Central, <http://news.harmony-central.com/>)

Za digitalizaciju vinilskih ploča vrlo je važno da zvučna kartica, vanjsko pojačalo ili pretpojačalo ima mogućnost RIAA¹⁵ ujednačavanja zvuka (engl. *RIAA equalization*). Riječ je o standardu kojim se omogućava poboljšanje kvalitete i povećanje duljine trajanja zvučnih zapisa na vinilskim pločama. Bez mogućnosti

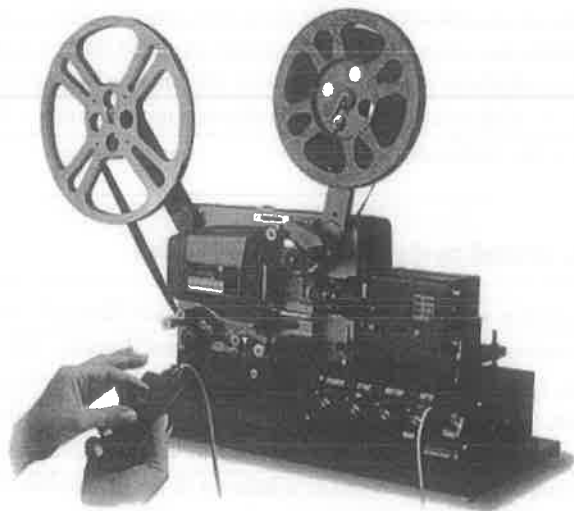
RIAA ujednačavanja zvuka, zvuk s gramofona priključenog izravno na zvučnu karticu bio bi digitaliziran uz vrlo slabe niže tonove i izrazito jake visoke tonove te bi se činilo kao da je prilikom digitalizacije došlo do nekih smetnji. Ipak, riječ je o načinu bilježenja zvuka na vinilskim pločama pa bez hardverskog ili softverskog RIAA ujednačavanja zvuka kvaliteta digitaliziranih snimaka s vinilskih ploča neće biti zadovoljavajuća.

2.2.4. Uređaji za digitalizaciju filma i videa

Na početku je važno znati da su se od pojave filmskih materijala do danas koristili najrazličitiji standardi celuloidnih filmskih traka. Tako su postojale one od 3 mm, koje je koristila NASA, pa sve do 70 mm i više. Stoga je realno očekivati da se, prilikom odabira uređaja za digitalizaciju ovakvih materijala, može naići na mnoge probleme vezane upravo uz format filmskog zapisa, iako će se većina gradiva, u pravilu, nalaziti u jednom ili tek u nekoliko formata. Važno je, na početku, testirati ispravnost reprodukcijских uređaja, tj. filmskih projektorā, prije nego što se krene u digitalizaciju. Naime, nemogućnost reproduciranja filmskog materijala vrlo će brzo pokazati da je jedina mogućnost digitalizacija izvan institucije kod nekog pružatelja takvih usluga. No, ako reprodukcijских uređaji ispravno funkcioniraju, tada će se morati odabrati tehnika i uređaj za digitalizaciju.

Postoji mnogo različitih tehnika za digitalizaciju filmskih materijala. Najkvalitetnija je, svakako, korištenje uređaja za digitalizaciju koji digitalizira gradivo spajanjem na leću projektorā koji se potom spaja na računalo. Ovisno o vrsti digitalizacijskog uređaja, on može biti i relativno skupo rješenje. Na računalo je potrebno imati grafičku karticu koja ima mogućnost prihvata i digitalizacije signala iz uređaja za digitalizaciju, kao i instaliran program za prihvāt signala. Ako je riječ o zvučnom filmskom materijalu, tada je potrebno na računalo imati i kvalitetnu zvučnu karticu za prihvāt tog dijela zapisa. Uređaj za digitalizaciju filmskog gradiva prikazan je na sljedećoj slici. U nedostatku ovog uređaja, postoje i druge tehnike digitalizacije, kao što su projiciranje snimke na bijelu pozadinu, ali nikako na projekcijsko platno, jer ono raspršuje svjetlost, te snimanje digitalnom video kamerom ili pak digitalizacija svake sličice filma zasebno uz naknadno spajanje. Posljednja spomenuta tehnika zahtijeva opsežnu naknadnu obradu (filmski zapis se sastoji od 24 sličice po jednoj sekundi materijala).

¹⁵ RIAA – Recording Industry Association of America.



Slika 11: Uređaj za digitalizaciju filmskih materijala

Digitalizacija video materijala, pak, zahtijeva drugačiji pristup. U ovom slučaju, naime, nisu potrebni projektori i relativno skupi uređaji za digitalizaciju spajanjem na leću projektor, već je dovoljno imati video uređaj za reprodukciju snimljenih materijala i grafičku karticu na računalu koja ima mogućnost prihvata i digitalizacije video signala. Osim toga, potrebno je imati i kvalitetnu zvučnu karticu za prihvata i digitalizaciju zvučnog dijela zapisa ili pak zaseban uređaj za digitalizaciju zvučnog zapisa (npr. pretpojačalo) koji se potom spaja na računal. Potrebno je, naravno, imati na računalu i program za prihvata digitaliziranog signala. Kvaliteta snimke na video kaseti određivat će kvalitetu njezine digitalizirane verzije.

2.3. POSTUPCI DIGITALIZACIJE

2.3.1. Tekstualno gradivo

Tekstualno se gradivo može unositi u računalno prepisivanjem i skeniranjem ili slikanjem digitalnim fotografskim aparatom, uz kasniju uporabu programa za optičko prepoznavanje slova (engl. OCR – Optical Character Recognition).

2.3.1.1. Prepisivanje gradiva

Prepisivanje je najjednostavniji oblik digitalizacije tekstualnog gradiva. To je dugotrajan, iscrpljujući i vrlo skup način prijenosa teksta u digitalni oblik. Ovaj oblik digitalizacije preporučljiv je ako su apsolutna vjernost digitalne verzije s originalom (bez tipkarskih pogrešaka) i pretraživanje teksta postavljeni kao uvjet ispunjenja cilja projekta. Prepisivanje gradiva može biti i najisplativiji oblik digitalizacije ako se radi o rukopisima, starim, požutjelim, nedovoljno kontrastnim stranicama, tekstovima s rukom nadodanim bilješkama na marginama ili, pak, označenim (engl. *mark-up*) tekstom. U većini tih slučajeva niti jedna automatizirana tehnika prepoznavanja teksta nije uopće upotrebljiva ili nije dovoljno učinkovita te zahtijeva dodatne korekcije. Takav postupak digitalizacije bio bi tada skuplji i vremenski zahtjevniji od digitalizacije postupkom prepisivanja gradiva.

2.3.1.2. Skeniranje gradiva

Tekstualno gradivo može se skenirati s originala ili već postojećeg mikrofilma. Skeniranje originala koji se nalazi na zasebnim listovima papira standardne veličine može se automatizirati uporabom stolnih skenera i uvlačka papira. Ostalo gradivo skenira se plošnim skenerom.

Za uspješno provođenje digitalizacije mikroformata potrebno je koristiti specijalne skenere. Neki će uređaji skenirati obje vrste mikroformata, mikrofilmove i mikrofiševe, ali obično je potrebna posebna vrsta skenera za svaki od tih dvaju medija. Zbog skupoće opreme, za takvu je vrstu posla bolje angažirati vanjskog pružatelja usluga skeniranja takvih oblika.

Za skeniranje mikrofiševa može se koristiti i plošni skener s prosvjetljivačem, ali on neće moći zabilježiti sve detalje izvornika.

Važno je konstatirati da se skeniranjem tekstualnog gradiva kao rezultat dobiva digitalna slika teksta, a ne tekst koji se može obrađivati. Stoga je kasnije potrebna dodatna obrada programom za optičko prepoznavanje slova koji sliku teksta pretvara u obradivi tekst.

Tekstualno gradivo skenira se na dva različita načina, ovisno o vrsti naknadne obrade digitalne slike. Ako se dokument skenira s namjerom da se tekst poslije provuče kroz program za optičko prepoznavanje slova, te tako dobije obradivi i pretraživi tekst, onda se on ne smije skenirati u boji, već kao crno-bijeli dokument zbog postizanja optimalnog kontrasta između, u većini slučajeva, svijetle podloge i tamnih slova, te rezolucijom od 300 dpi (400 dpi za mikroformate) ili većom. Upravo je dobar kontrast, uz kvalitetan izvornik koji se skenira, jedan od ključnih faktora za postizanje kvalitete i dobre učinkovitosti prilikom naknadne obrade. Iako se OCR programi stalno unapređuju, najučinkovitiji su kada rade s kontrastnim tekstom napisanim jednostavnim fontom crnom bojom na bijeloj podlozi. Digitalizirani tekst može se ostaviti u slikovnom formatu, može se provući kroz OCR program te spremi u nekom tekstualnom formatu zapisa, pri čemu je nemoguće izbjeći pojavljivanje određenog broja pogrešno prepoznatih znakova ili se, pak, može urediti i ispraviti tako da se dobije posve ispravna verzija teksta. Posljednja navedena opcija čini vremenski najzahtjevniji postupak jer se tekst mora ručno ispravljati. Neki ponudaci usluga digitalizacije u svoju uslugu skeniranja imaju uključen i OCR. Slike u sklopu dokumenta mogu se izrezati uporabom nekog programa za obradu slika. Tekstualno gradivo za koje je ovakav postupak preporučen uglavnom je tiskano gradivo, mikroformati i dokumenti uredno napisani pisačim strojem. Postupci obrade korištenjem OCR programa temeljitije su objašnjeni u trećem poglavlju koje govori o obradi i kontroli kvalitete.

Ako se, pak, dokument skenira s namjerom da stranice ostaju pohranjene kao slike, tj. digitalne fotokopije, tada se on može skenirati kao crno-bijeli dokument u skali sivih nijansi, ili kao dokument u boji. U tom slučaju tekst nije pretraživ. Iako je ovaj način skeniranja naizgled najbrži postupak pretvaranja gradiva u digitalni oblik, ipak se i takve digitalne slike teksta moraju poslije dodatno obraditi, tj. moraju im se ugraditi metapodatci kako bi postale pretražive. Pretraživanje arhiva digitaliziranih slika, teksta ili mikroformata moguće je, dakle, pomoću pretraživanja indeksnih metapodataka koji prate digitalnu verziju slike. To znači da će biti potrebno određeno vrijeme i trud da se digitalnoj verziji pridodaju prateći metapodatci. Metapodatci se sastoje od tehničkih metapodataka koje su sami uređaji za digitalizaciju spremili uz pojedinu datoteku (datum digitalizacije, oprema i njezine tehničke karakteristike i postavke korištene

prilikom digitalizacije i sl.), metapodataka koji se odnose na sadržajnu komponentu digitalnog zapisa te metapodataka pridodanih u svrhu jednoznačne identifikacije pojedinog zapisa u digitalnom arhivu. Zbog toga ovaj postupak nije znatno brži od prethodnog. Bilo koji tip gradiva može se digitalizirati na ovaj način, a preporučljiv je za sve tipove tekstualnog gradiva, pogotovo ako je cilj projekta digitalizacije prikazati izvorni izgled dokumenata, te za dokumente koji nisu obradivi programom za optičko prepoznavanje slova, a iz bilo kojeg razloga ne želi ih se prepisivati.

Praktične napomene:

- prostorna rezolucija trebala bi biti minimalno 300-400 dpi-a,
- poželjna je 8-bitna siva skala, TIFF 6.0 nekomprimiran,
- svaka datoteka treba sadržavati jednu sliku, a kad su snimljene dvije stranice odjednom, valja ih razdvojiti,
- ispraviti odstupanje zakrenutosti veće od 3 stupnja.

2.3.1.3. Slikanje digitalnim fotografskim aparatom

Digitalni fotografski aparati danas su već dostigli kvalitetu profesionalnih klasičnih fotoaparata, a cijenom ih nadmašuju. No, oni skraćuju proces digitalizacije gradiva koje se ne može obraditi klasičnim skeniranjem. Tako se, umjesto fotografiranja klasičnim fotoaparatom, izradom mikrofilmova, te njihova skeniranja, gradivo slika digitalnim fotoaparatom koji sliku odmah zapisuje u digitalnom obliku. Ova se tehnika koristi kod gradiva koje treba digitalizirati odozgo kako ne bi došlo do oštećenja, npr. uvezani svesci. Postoje i posebno dizajnirani skeneri vrhunske kvalitete koji robotizirano okreću stranice knjige. Robotska ruka proračunatom snagom usisa, koja točno odgovara vrsti gradiva kojim barata, prihvaća i okreće stranice.

Kao i kod skenera, digitalizacija uporabom digitalnog fotoaparata rezultira slikom izvornika koja se poslije mora dodatno obraditi. Sve opcije vezane uz način digitalizacije izvornika koje uvjetuje način buduće obrade, a koje su spomenute u prethodnom odlomku vezanom uz skeniranje gradiva, vrijede i prilikom slikanja digitalnim fotoaparatom.

2.3.2. Slikovno gradivo

Slikovno gradivo digitalizira se uporabom skenera s visokom razlučivošću ili digitalnih fotoaparata. Ako je gradivo prikladne vrste i veličine, mogu se koristiti i rotacioni skeneri. Njima se postiže znatno viša rezultirajuća kvaliteta digitalne slike nego uporabom klasičnih, plošnih skenera. Rotacioni skeneri pogodni su samo za digitalizaciju transparentnog ili netransparentnog gradiva koje nije uvezano, zahtijeva skeniranje u visokoj razlučivosti te smije biti presavijeno. Oni, naime, koriste bubanj oko kojeg se gradivo obavlja za vrijeme skeniranja. Ovakvi se uređaji uglavnom ne koriste za digitalizaciju tekstualnog gradiva, jer su dosta skupi, a dovoljno visoka kvaliteta može se postići i klasičnim skenerima.

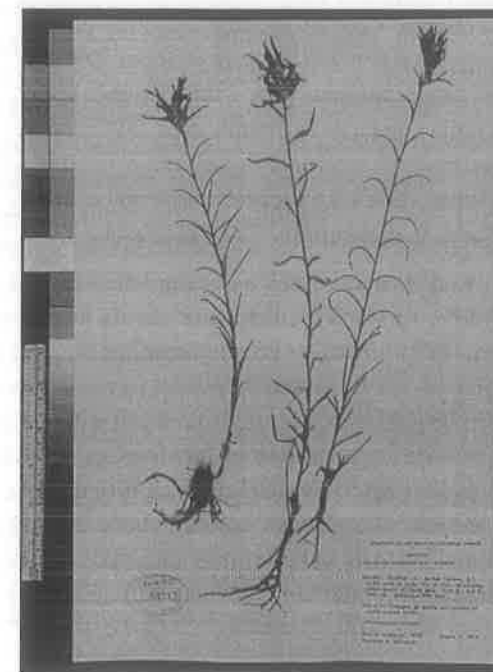
Za građu većih formata mora se koristiti visokokvalitetni digitalni fotoaparat kojim se predlošci digitaliziraju okrenuti licem prema gore. Rasvjeta prilikom digitalizacije fotoaparatom vrlo je bitna, pogotovo kad se digitaliziraju predlošci pri čijem se osvjetljenju javljaju zrcalni odbljesci. Tako je, na primjer, kod različitih srebrnih površina potrebno istodobno osvjetljenje s više strana.

Prilikom digitalizacije slikovne građe treba odrediti maksimalnu dopuštenu izloženost građe svjetlu te paziti prilikom odabira skenera i osvjetljenja za kameru da njihovi svijetleći elementi ne proizvode ultraljubičastu komponentu svjetlosti, jer je ona najzaslužnija za većinu blijeđenja i degradacije umjetničkih slika. Budući da se digitalizacijom želi sačuvati informacijski sadržaj predloška, digitalna slika mora biti kvalitetna. To znači da ona mora biti skenirana u visokoj razlučivosti (najmanje 600 dpi (engl. *dpi* – *dots per inch*, tj. broj točaka po kvadratnom inču) u 24-bitnoj boji za manje formate te 300 dpi za veće formate. Što je slika više razlučivosti, to zauzima više mjesta na disku i manje je prikladna za prijenos mrežom. Primjera radi, slika veličine A4 formata papira, skenirana u razlučivosti od 600 dpi koristeći 8 bitova (256 nijansi) za svaku od tri osnovne boje – crvene, zelene i plave, tj. 24-bitnu boju – u nekomprimiranom formatu zauzima diskovni prostor veličine 96 Mb.

Preporuča se izraditi tri verzije svake digitalizirane slike – jednu u boji, visoke razlučivosti (master), koja mora biti što vjernija originalu i neće biti komprimirana, drugu s manjim brojem nijansi osnovnih boja ili u skali sivih tonova, niže razlučivosti ili, pak, komprimirana, koja će biti pogodnija za prijenos mrežom, te sitnu, identifikacijsku sličicu (engl. *thumbnail*), koja će služiti kao vizualna referenca ili veza na bilo koju od prethodne dvije slike. Sve se slike poslije

dodatno obrađuju.¹⁶ Hoće li se izvornik skenirati tri puta kako bi se dobile sve tri preporučene verzije slike ili će se skenirati samo jednom, u najvećoj razlučivosti, pa naknadnom obradom komprimirati, smanjivati original i smanjivati broj boja, ostaje na instituciji da odluči. Ipak, praksa pokazuje da je najbolje jednom skenirati gradivo i naknadno ga obraditi zbog toga jer je time izvornik mnogo kraće vrijeme izložen osvjetljenju, a također se produžuje vijek trajanja opreme.

U svrhu smanjenja potrebne količine vremena za naknadnu obradu digitaliziranih slikovnih materijala i njezinu što konzistentniju provedbu, poželjno je prilikom digitalizacije zahvatiti i kontrolne uzorke koji se smještaju izvan sadržajnog područja predloška. "Uključivanjem kontrolnih uzoraka boja (npr. Q-13 kontrolni uzorak s nijansama sive i uzorcima u boji¹⁷ te ravnalom u centimetrima i inčima), koje se snima ili digitalizira zajedno s građom, moguće je unaprijediti naknadnu mogućnost korekcije boja."¹⁸



Slika 12: Digitalizirani botanički uzorak iz herbarija s pridodanim kontrolnim uzorkom s lijeve strane¹⁹

¹⁶ Stančić, Digitalizacija građe, n. dj. str. 65-66.

¹⁷ Engl. Color Separation Guide and Gray Scale. Najpoznatiji je onaj proizvođača Kodak.

¹⁸ Format datoteka za pohranu i korištenje (radna verzija), Nacionalni projekt "Hrvatska kulturna baština", Ministarstvo kulture Republike Hrvatske, studeni 2007., <http://www.kultura.hr/hr/content/download/597/7937/file/smjernice_formati.pdf>, 27. kolovoza 2008.

¹⁹ Izvor: Gilroy, Peter, *Herbarium Specimen Imaging Protocol*, Museum of Southwestern Biology, University of New Mexico, <<http://www.msb.unm.edu/herbarium/imaging.pdf>>, 5. siječnja 2009.

Kako bi se postigla što bolja kvaliteta digitaliziranoga gradiva, treba objasniti što utječe na kvalitetu digitalne slike. Postoje tri bitne odrednice svake digitalne slike – rezolucija, bitna dubina točke (engl. *Pixel Bit Depth*) i boja.

2.3.2.1. Rezolucija

Rezolucija ili prostorna frekvencija je broj uzorkovanja predloška tijekom skeniranja. Ona se iskazuje kao broj plošne ili linijske gustoće točaka – PPI (engl. *pixel per inch* – piksel po inču), DPI (engl. *dots per inch* – točke po inču) i LPI (engl. *lines per inch* – linije po inču). Veličina PPI upotrebljava se za označavanje rezolucije digitalnih slika, DPI za označavanje rezolucije pisača, a LPI za označavanje rezolucije nijansiranja (engl. *half-toning*) u tiskarstvu. Tako, na primjer, slika 4 x 5 inča rezolucije 600 dpi čini polje sastavljeno od 2.400 x 3.000 točaka, tj. sveukupno 7.200.000 točaka.

Kod skenera se obično ističu dvije vrste rezolucije – optička i interpolirana. *Optička rezolucija* je fizička rezolucija koju skener može postići koristeći CCD senzore²⁰ i povećanje kroz ugrađeni optički sustav. *Interpolirana rezolucija* je rezolucija izračunata matematičkom metodom interpolacije. Ona se koristi kako bi se matematičkim putem povećala rezolucija skenirane slike.²¹ Interpolirana rezolucija koristi se kao nadgradnja optičke rezolucije, ali zbog toga što se tom metodom fizički ne detektira veći broj točaka po inču, već se one izračunavaju matematičkim putem i umeću između onih fizički detektiranih točaka, rezultirajuća kvaliteta slike je niža, ponekad vidljivo zamućenija, od slike dobivene skeniranjem optičkom rezolucijom.

2.3.2.2. Bitna dubina točke

Kao što je već rečeno, svaka se digitalna slika sastoji od polja točaka. Svakoj točki u tom rasteru pridružen je binarni niz. Broj znamenka tog niza naziva se bitna dubina. Termin koji se također koristi je prostorna rezolucija (engl. *spatial resolution*). Točka s dubinom 1 bita je točka kojoj je pridružen niz dužine jedne znamenke, tj. vrijednost 0 ili 1. Točka s dubinom 2 bita je točka kojoj je pridru-

žen niz koji ima dvije znamenke. Vrijednosti u tom slučaju mogu biti 00, 01, 10, ili 11. Dakle, dubina bita određuje broj mogućih kombinacija jedinica i nula. Svaka od mogućih kombinacija čini određenu nijansu boje. O kojim se točno bojama radi ovisi o sustavu boja koji se koristi. Postupak dijeljenja analogne slike, kao cjeline sastavljene od beskonačno mnogo djelića, u mrežu s konačnim brojem točaka nazivamo prostornom kvantizacijom (engl. *spatial quantization*). Pretpostavimo da se radi o crno-bijeloj fotografiji. Tada govorimo o broju nijansi sive boje. Koliki je broj mogućih nijansi (n) X -bitne dubine može se izračunati formulom:

$$2^x = n$$

što se može prikazati i sljedećom tablicom (Tablica 2).²²

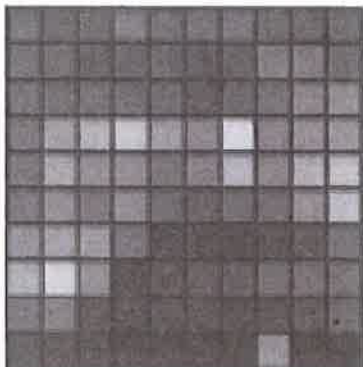
Tablica 2: Odnos dubine bita i broja nijansi

Dubina	Broj nijansi	Objašnjenje
1 bit	2	1 binarna znamenka od 2 moguće kombinacije, 1 ili 0, tj. bijelo ili crno
2 bita	4	2 binarne znamenke od 4 moguće kombinacije – crno, tamno sivo, svjetlo sivo, bijelo
3 bita	8	3 binarne znamenke od 8 mogućih kombinacija
4 bita	16	4 binarne znamenke od 16 mogućih kombinacija
5 bita	32	5 binarnih znamenki od 32 moguće kombinacije
6 bita	64	6 binarnih znamenki od 64 moguće kombinacije
7 bita	128	7 binarnih znamenki od 128 mogućih kombinacija
8 bita	256	8 binarnih znamenki od 256 mogućih kombinacija
10 bita	1.024	10 binarnih znamenki od 1.024 moguće kombinacije
12 bita	4.096	12 binarnih znamenki od 4.096 mogućih kombinacija
14 bita	16.384	14 binarnih znamenki od 16.384 moguće kombinacije
16 bita	65.536	16 binarnih znamenki od 65.536 mogućih kombinacija

²⁰ CCD senzori (engl. *Charged Coupled Device*) su nabojski vezani sklopovi osjetljivi na reflektiranu svjetlost koji se, uz dodatne elektroničke sklopove, koriste u skenerima i digitalnim fotoaparatima za konverziju analognog električnog signala koji oni proizvode u digitalni signal.

²¹ Puglia, Steven, Technical Primer, u: *Handbook for Digital Projects*, n. dj., str. 97.

²² Puglia, Technical Primer, n. dj., str. 99.



Slika 13: Mreža točaka 10 x 10 kod crno bijele fotografije

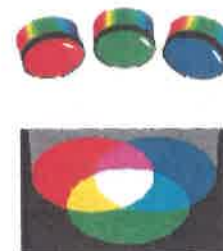
Bitna dubina točke uvelike utječe na kvalitetu digitalne slike. Naime, u prirodi postoji beskonačno mnogo nijansi, a vrijednosti koje opisuju svaku točku moraju biti konačne. Stoga se stvarne vrijednosti moraju zaokružiti na srednju vrijednost, što unosi određenu količinu greške u sliku (nereda u sustav). Dakle, što je dubina bita veća, to su pogreške sitnije, a slika kvalitetnija. Navedeni primjer govorio je o broju nijansi sive boje, dakle o crno-bijelim slikama. Kako se izračunava broj mogućih nijansi kod slika u boji? Postoji nekoliko sustava prikaza boje te je računanje dubine bita drugačije za svaki od tih sustava. U sljedećem odlom-

ku bit će više riječi o tim sustavima, ali ono što im je svima zajedničko je to da se boja prikazuje miješanjem nekoliko kanala zasebnih osnovnih boja. Koliko ima osnovnih boja i koje su to boje, ovisi upravo o upotrijebljenom sustavu. Broj nijansi računa se za svaku od osnovnih boja, pa tako 24-bitna boja po RGB sustavu (engl. *RGB* – *Red, Green, Blue*), koji se sastoji od crvene, zelene i plave osnovne boje, ima blizu 17 milijuna (točnije 16.777.216) mogućih kombinacija, jer se sastoji od triju 8-bitnih kanala osnovnih boja.

2.3.2.3. Boja

Prilikom digitalizacije slikovnog gradiva od velike je važnosti vjerna reprodukcija boje. Način izračunavanja boje ovisi o primijenjenom sustavu prikaza. Najčešće upotrebljavani sustavi su RGB, CMYK i CIELAB.

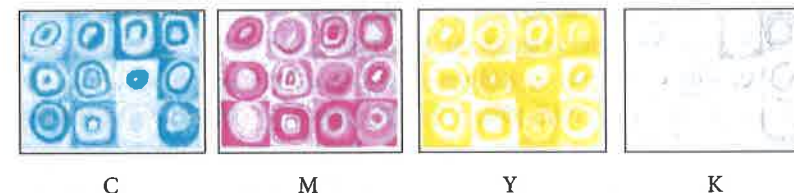
RGB sustav stvara sliku kombinirajući tri osnovne boje: crvenu, zelenu i plavu. On funkcionira po principu dodavanja boje crnoj pozadini. Sliku po RGB sustavu možemo zamisliti sačinjenu od triju zasebnih slojeva različitih koncentracija osnovnih boja preklapljivih jedan preko drugog. Preklapanje, tj. izračun finalne boje, izvodi se matematički za svaku pojedinu točku rastera od kojeg se slika sastoji. Najčešće se primjenjuje 24-bitna RGB slika koja se dobiva kombinacijom triju 8-bitnih kanala. Ovaj se sustav primarno koristi za prikaz boje na zaslonima računala.



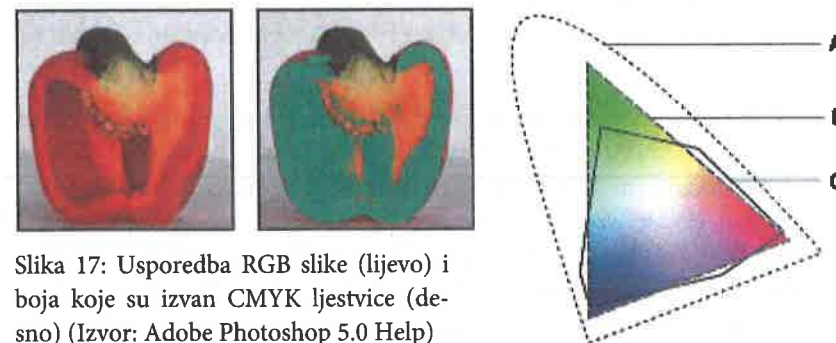
Slika 14: RGB sustav boja



Slika 15: CMYK sustav boja



Slika 16: Slika u CMYK sustavu i separacije četiriju osnovnih boja (C, M, Y, K)



Slika 17: Usporedba RGB slike (lijevo) i boja koje su izvan CMYK ljestvice (desno) (Izvor: Adobe Photoshop 5.0 Help)

Slika 18: Ljestvica boja kod (A) CIELAB, (B) RGB i (C) CMYK sustava (Izvor: Adobe Photoshop 5.0 Help)

CMYK (engl. *CMYK* – *Cyan, Magenta, Yellow, black*) sustav stvara sliku kombinirajući četiri osnovne boje: blijedo plavu, grimizno ljubičastu, žutu i crnu. Za razliku od RGB sustava, CMYK funkcionira po principu oduzimanja boje bijeloj pozadini. Slika se po ovom sustavu izračunava preklapanjem četiriju zasebnih slojeva osnovnih boja. Ako se svaki od tih četiriju kanala sastoji od 8-bitne boje, nastaje 32-bitna CMYK slika. Ovaj sustav primarno koriste pisači u boji.

Razlike između ovih dvaju sustava prikaza boja postaju značajne kada se iz bilo kojeg razloga pokuša na pisaču otisnuti skenirano slikovno gradivo. Naime, kako računalni zaslone i pisači koriste različite sustave prikaza boje, događa se da ono što dobro izgleda na zaslonu, ne izgleda isto tako dobro otisnuto na papiru, pogotovo ako se želi napraviti kvalitetan otisak u tiskari, za što je prethodno potrebno napraviti separaciju boja (vidi sliku 16). Separacija boja je postupak odjeljivanja svake od osnovnih boja. Nakon odjeljivanja zasebni se slojevi osvijetle na film od kojeg se izrađuju tiskarske ploče. Papir u tiskarskom stroju prolazi ispod svake ploče, a ona na njega nanosi boju. Slika je tiskana nakon nanošenja svih osnovnih boja jednih preko drugih. Upravo kod separacije boja dolazi do najvećih razilaženja, jer se područja boja koja pokrivaju ovi sustavi ne preklapaju u potpunosti, pa tako u svakom sustavu postoji određeni broj boja koje onaj drugi ne prepoznaje (vidi sliku 17). Zbog toga je potrebno izvršiti konverziju slikovnoga gradiva koje želimo otisnuti, a koje je zapisano po RGB sustavu radi izvornijeg prikaza na zaslonu računala, u CMYK sustav prije separacije boja za tisak. Naravno, uvijek moramo imati na umu da je takva konverzija matematička aproksimacija barem što se tiče onog dijela palete koje se nalazi izvan presjeka tih dvaju sustava. Kao rješenje ovog problema može poslužiti CIELAB sustav prikaza boja.

CIELAB sustav prikazuje boje koristeći tri vrijednosti kojima opisuje precizni smještaj boje unutar vidljivog prostora boja. Skraćenica CIE dolazi od *Commission Internationale de l'Eclairages*, (Međunarodna komisija istraživača boje), čiji standardi pridonose mogućnosti ispravne komunikacije informacijama o boji. Nadalje, L opisuje relativnu svjetlost (engl. *lightness*), A predstavlja relativan odnos između crvene i zelene, a B između žute i plave. Ovaj sustav prikaza boja ima mnogo širu paletu boja te obuhvaća sve boje s kojima barataju RGB i CMYK (vidi sliku 18). Zbog toga se slika zapisana po CIELAB sustavu može bez gubitaka konvertirati bilo u RGB ili CMYK sustav. Iako fizički zauzima nešto više diskovnog prostora, upravo bi zbog te mogućnosti ovaj sustav zapisa boja mogao vrlo lako prerasti u standard. To je prije svega stoga što je implementacija CIELAB sustava u PostScript razine 2, jezik koji koriste RIP procesori (engl.

Raster Image Processor – procesori rastera slike) koji se koriste za separaciju boja prije slanja na osvjetljivač, pridonijela tome da se slika zapisana u CIELAB sustavu može bez gubitaka najprije konvertirati u RGB za pregled, analizu i korekcije na zaslonu računala, zatim ponovo bez gubitaka konvertirati u CMYK za potrebe probnih ispisa na printerima u boji, te na kraju procesirati RIP procesorima prilikom separiranja boje.²³

Iz prethodnih odjeljaka vidi se da je digitalizacija slikovnoga gradiva vrlo kompleksan postupak. Kad se kreće u projekt digitalizacije ključno je na početku točno odrediti kako će se koristiti slikovno gradivo: hoće li se gradivo samo arhivirati, u vidu mastera, i biti namijenjeno isključivo pregledu na računalnim zaslonima, u umanjenoj verziji, bilo u instituciji ili putem Interneta, ili će institucija također nuditi mogućnost ispisa na zahtjev korisnika, ili možda predvidjeti izradu fizičke kopije originala. Cilj projekta će, dakle, uvjetovati odabir rezolucije, bitne dubine točke i sustava zapisa boja koji će se koristiti kao standard prilikom digitalizacije slikovnoga gradiva. Ponekad je vrlo teško predvidjeti na koje bi se sve načine moglo koristiti digitalno gradivo, pa je stoga najbolje odabrati ona rješenja koja nisu ograničavajuća.

Preporuke o tome koja svojstva trebaju imati datoteke arhivske kvalitete za trajnu pohranu u kontekstu preporučenog formata zapisa, veličine slike u pikselima, te potrebne rezolucije i dubine boje u odnosu na vrstu izvornika (tekstualni dokumenti, grafičke ilustracije i umjetničke grafike, zemljopisne karte, tlocrti, nacrti, fotografije, dijapozitivi i negativi) i njegova svojstva (raspon formata predloška, površina predloška) mogu se pronaći u prilogu A smjernica *Formati datoteka za pohranu i korištenje* nastalima u okviru Nacionalnog projekta "Hrvatska kulturna baština".

2.3.3. Zvučno gradivo

Zvučno se gradivo digitalizira tako da se zvučni izlaz uređaja za reprodukciju određenog medija, npr. gramofonske ploče, audio kasete, studijske magnetske trake itd., poveže s računalom. U računalu mora biti ugrađena kartica koja može prihvatiti zvučni ulaz, te ono mora biti opremljeno programom za prihvati i obradu zvučnog signala. Na tržištu postoji mnogo verzija zvučnih kartica i programa za obradu zvuka, od najjednostavnijih do profesionalnih. Prije nego što

²³ CIELAB Color Conversion, <<http://www.aols.com/colorite/cielabcolorconv1.html>>, 31. siječnja 2001.

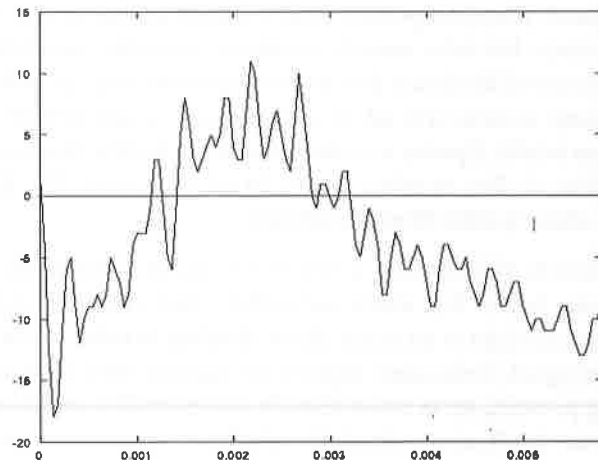
se krene u stvarnu digitalizaciju preporučljivo je testirati performanse hardvera i softvera u zajedničkom radu kako poslije ne bi došlo do neželjenih situacija.

Na početku je potrebno objasniti teoretske principe digitalizacije zvuka, te navesti praktična rješenja nekih problema koji se pritom javljaju i prihvaćene standarde.

Digitalizacija zvuka odvija se u dva koraka: 1. uzorkovanje i 2. kvantizacija.

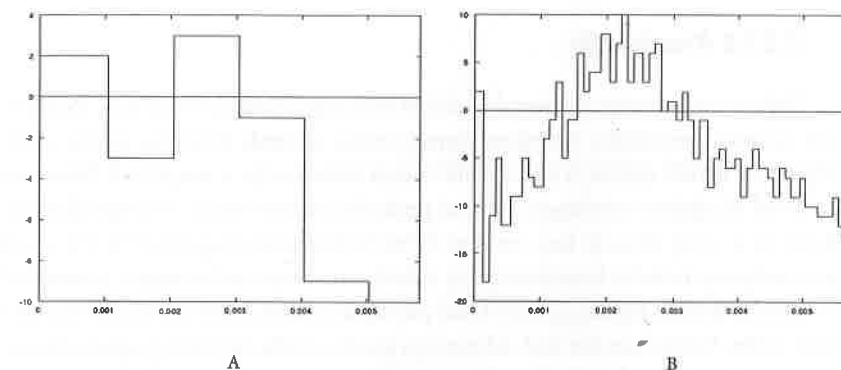
2.3.3.1. Uzorkovanje

Zvučni je signal kontinuiran kroz vrijeme. Kada se zvuk konvertira u električni napon, napon se mijenja u svakoj jedinici vremena (slika 19). S obzirom na to da postoji beskonačan broj vremenskih trenutaka na koje se može razdijeliti zvučni signal određenog (konačnog) trajanja, beskonačna je i količina informacija koju taj signal prenosi. Stoga je, u prvom koraku, potrebno odlučiti koliko često zapisivati informacije, a da to bude dovoljno za vjernu reprodukciju.



Slika 19: Uzorkovani valni oblik govora
(Izvor: Cyganski et. al., *Information Technology...*)

Prije svega je potrebno znati koji raspon frekvencije ljudsko uho može razlikovati. Iako rezultati variraju od osobe do osobe, smatra se da čovjek može razpoznati zvuk od 20 Hz, pa sve do 20.000 Hz (20 kHz), tj. ljudsko uho razaznaje frekvenciju od 20 do 20.000 podražaja u sekundi (za razliku od oka koje ne razaznaje više od 40 do 50 promjena u sekundi).²⁴ Ova činjenica nije strana proizvođačima zvučnih uređaja, pa stoga oprema standardno reproducira zvuk unutar navedenih frekvencija. Frekvencija uzorkovanja prilikom digitalizacije izravno utječe na kvalitetu zapisa, tj. na vjernost reprodukcije digitalnog signala analognom originalu. Slika 15 prikazuje odnos vjernosti dobivenog digitalnog signala uzorkovanog frekvencijama 1 kHz (1/1.000 s) i 10 kHz (1/10.000 s) na primjeru valnog oblika (engl. *waveform*) govora prikazanog na slici 20.



Slika 20: Uzorkovanje frekvencijama 1 kHz (A) i 10 kHz (B)
(Izvor: Cyganski et. al., *Information Technology...*)

Uzorkovanje frekvencijom od 1 kHz rezultira signalom koji ima vrlo malo sličnosti s originalom. Pitanje je bi li se uopće mogao prepoznati izvorni signal. Nedovoljno brzo uzorkovanje (engl. *undersampling*) rezultira stepeničastim signalom (engl. *aliasing*) te je ireverzibilan proces zbog prevelikog gubitka informacija. Uzorkovanje frekvencijom od 10 kHz, pak, daje bolje rezultate. U ovom slučaju signal je prepoznatljiv i nalik originalu. Dakle, što je frekvencija uzorkovanja viša, to je dobiveni signal kvalitetniji. Takav signal, naravno, zauzima više diskovnog prostora. Ipak, kao najniža frekvencija za dobivanje kvalitetnog sig-

²⁴ Cyganski, David, John A. Orr i Richard F. Vaz, *Information Technology: Inside and Outside*, Prentice-Hall, Inc., New Jersey, SAD, 2001.

nala uzima se 40 kHz, tj. dvostruka maksimalna frekvencija koju ljudsko uho razaznaje. U praksi se zvučni signal digitalizira frekvencijom koja je viša od dvostruke maksimalne, tj. frekvencijom od 44,1 kHz. To je iznos 10% veći od minimalno potrebnog. Iako viša frekvencija znači veći zapis, ona također znači kvalitetniji signal, tj. signal koji više nalikuje originalu, a za reprodukciju takvog signala, s obzirom na to da postoji dovoljan broj informacija za kvalitetnu reprodukciju, potreban je ne toliko složen algoritam, pa je dovoljna i manja procesorska snaga. Niža frekvencija uzorkovanja, pak, uvjetuje složeniji algoritam koji će pokušati reproducirati maksimalno kvalitetan signal iz uglavnom nedovoljne količine zabilježenih informacija. Stoga je frekvencija od 44,1 kHz idealan omjer cijene i kvalitete te se nametnula kao standard.

2.3.3.2. Kvantizacija

Nakon uzorkovanja vremenskog dijela zvučnog signala, potrebno je kvantizirati njegovu amplitudu. Uzorkovanjem je svaka sekunda zvučnog signala podijeljena na 44.100 djelića, a svaki djelić sadrži informaciju o amplitudi. Naravno, svaki od tih djelića amplitude može se podijeliti na beskonačno mnogo dijelova. Kako bi se u toj situaciji kao rezultat dobio beskonačno dugi binarni niz, mora se primijeniti tehnika kvantizacije, tj. tehnika svodenja na konačnu, prethodno određenu, dužinu binarnog niza. Ovaj postupak jednak je određivanju bitne dubine točke. Dakle, kao što kod slikovnoga gradiva treba odrediti moguću raznolikost svake točke, tako i kod zvučnoga gradiva treba odrediti broj bitova kojima je moguće opisati amplitudu u svakom uzorkovanom djeliću. Kako se svaki djelić amplitude može opisati s X bitova, tako i u ovom slučaju vrijedi formula $2^x = n$, gdje je n ukupan broj mogućih kombinacija. O broju bitova izravno ovisi kvaliteta, jer što je veći broj mogućih kombinacija, to je manja mogućnost pogreške. Za razliku od uzorkovanja, kvantizacijom se neminovno uvodi određena količina šuma, jer dolazi do određenog, svjesnog gubitka informacija. Uzrok toga gubitka nalazi se u tome što "navedena vrijednost statistički ne odgovara uvijek stvarnoj vrijednosti; pri dekodiranju izgubljeni se dio ne može ponovo rekonstruirati, pa nastaje pogreška; ta će pogreška biti to manja što je veći broj diskretnih stupnjeva kvantiziranja".²⁵ Naravno, ti gubici moraju biti tako mali da ljudskom uhu budu nezamjetljivi. Zvučno gradivo digitalizirano u skladu sa standardom koji se primjenjuje u CD uređajima za svaki kanal (stereo signal

ima razdvojen lijevi i desni kanal) koristi 16-bitni prikaz svakog uzorkovanog djelića zvučnog signala, tj. vrijednost amplitude prikazuje kao jednu, kvantiziranu vrijednost od 65.536 mogućih vrijednosti (2^{16}) za svaki od 44.100 djelića signala u sekundi.

Institucije bi prilikom izbora standarda za digitalizaciju zvučnoga gradiva trebale ovaj standard prihvatiti kao minimum.

2.3.4. Filmsko i video gradivo

Film i video trake, kao mediji za prijenos, prikaz i pohranu video gradiva u analognom obliku, uz pomoć uređaja za reprodukciju prikazuju niz statičnih slika na način koji se ljudskom oku doima kao neprekinuto gibanje. Pod terminom video gradiva podrazumijeva se svako gradivo koje svoj sadržaj prezentira pomičnom slikom i, ali ne nužno, zvukom.

Zbog tromosti ljudsko oko ne razaznaje promjene u okolini koje se događaju brzinom većom od 50 milisekundi. To znači da se pri brzini većoj od 20 promjena u sekundi promjene počinju stapati. Ako te promjene nisu nagle, stječe se dojam kretanja. Isto tako, uzme li se primjer bljeskalice, postavlja se pitanje do koje će učestalosti bljeskanja ljudsko oko registrirati treptaje. Nakon što učestalost bljeskanja postane brža od 80 u sekundi oko će interpretirati bljeskove kao neprekidno svjetlo. Upravo se ovim fenomenom koriste film i televizija. Danas je uobičajeno da film za prikaz koristi 24 sličice, a video 25 sličica u sekundi. Televizija obnavlja sliku 30 puta u sekundi.

Digitalizacija video gradiva u osnovi se svodi na digitalizaciju slike i digitalizaciju zvuka. Faktori koji utječu na kvalitetu slike i zvučnog signala objašnjeni su prije i primjenjuju se na svaku sličicu filmskog ili video zapisa te na pripadajući zvučni signal. Treba samo naglasiti da digitalni video zapis zauzima mnogo diskovnog prostora, jer se svaka sekunda takvog gradiva sastoji od najmanje 24 sličice. Ipak, korištenjem određenih tehnika komprimiranja mogu se postići mnogo manji zapisi, ponekad nauštrb kvalitete, ali o tome će više riječi biti u odjeljku koji se bavi obradom digitalnih zapisa.

Kao i kod svakog drugog gradiva potrebno je odrediti s kojim se ciljem gradivo digitalizira. Ako je razlog očuvanja video materijala u digitalnom obliku propadanje originalnog medija na kojem se gradivo nalazi, onda treba težiti postizanju maksimalne vjernosti originalu uz razumne troškove pohrane. Ako je, pak, digitalizirani materijal namijenjen prikazu s multimedijskog CD-ROM-a ili

²⁵ Kiš, Miroslav, *Englesko-hrvatski i hrvatsko-engleski informatički rječnik*, Naklada Ljevak, Zagreb, 2000., str. 765, s.v. quantization error.

distribuciji putem mreže, onda treba postići povoljan odnos veličine i kvalitete zapisa. Pojavom DVD medija, čiji je tipični kapacitet do 133 minute video zapisa, ta dvojba se smanjila, a dolaskom Blu-Ray medija, tipičnog kapaciteta do 13 sati video zapisa, ona se dodatno smanjuje. U budućnosti, s pojavom medija sve većeg kapaciteta, ovaj će se trend nastaviti.

2.3.5. Trodimenzionalno gradivo

Trodimenzionalno se gradivo u pravilu skenira posebnim skenerima namijenjenima skeniranju volumena. No, za digitalizaciju jednostavnijih trodimenzionalnih objekata mogu se koristiti i plošni skeneri ili digitalni fotoaparati. Naime, tim se postupcima može dobiti vjerna slika digitaliziranog predmeta. Tom se tehnikom mogu digitalizirati, primjerice, novčići, nakit, neke manje figurice, pa čak i botanički uzorci (npr. presjek gljive) itd. Ono na što je potrebno posebno obratiti pažnju jest činjenica da će prilikom korištenja plošnih skenera doći do gubitka plastičnosti uzoraka, jer će oni biti osvijetljeni pod pravim kutom u odnosu na njihovu površinu koja prijanja na skener. Dodatno, tim postupkom može doći do neželjenih odbljesaka kad je riječ o predmetima s refleksnim površinama. Tako, na primjer, reljefna slova na novčićima mogu biti vrlo slabo čitljiva, a motivi na njima, zbog odbljesaka, neprepoznatljivi. Ako se, pak, za taj postupak koristi digitalni fotoaparat ili reprografski umjesto plošnog skenera, tada se ovaj problem može zaobići korištenjem niskog bočnog osvjetljenja kako bi se stvaranjem sjena pridonijelo osjećaju plastičnosti tj. trodimenzionalnosti digitaliziranog objekta i izbjegli mogući odbljesci.

S druge strane, korištenjem specijalnih 3D skenera, nakon skeniranja, u računalu nastaju trodimenzionalni modeli skeniranih objekata. Ti skeneri detektiraju veliki broj točaka na predmetu, ali i njihovu udaljenost od skenera i referentnih točaka. Time se dobiva takozvani "oblak točaka", nakon čega se te točke spajaju i tako nastaju plohe. Nakon skeniranja 3D modela, svakako će biti potrebna daljnja obrada dobivenih 3D računalnih modela.

3. OBRADA I KONTROLA KVALITETE

Nakon uspješne digitalizacije digitalno gradivo treba dodatno obraditi i provjeriti zadovoljava li njegova kvaliteta postavljene standarde i nakon obrade. Svaka vrsta gradiva ima neke specifične načine dodatne obrade, dok su neki principi zajednički za više vrsta.

3.1. TEKSTUALNO GRADIVO

Kod tekstova koji su se ručno prepisivali potrebno je provesti korekturu kako bi se pogreške smanjile na najmanju moguću mjeru.

Tekst koji je digitaliziran skeniranjem ili slikanjem digitalnim fotografskim aparatom rezultira slikom svake stranice te ga nije moguće obrađivati, pretraživati ni indeksirati. Potrebno je, dakle, sliku pretvoriti u tekst. To se postiže uporabom OCR programa, tj. programa za optičko prepoznavanje slova (engl. *OCR – Optical Character Recognition*). Tada se takva slika otvara OCR programom koji na temelju razlike u kontrastu između otisnutog teksta i pozadine prepoznaje pojedine znakove i zapisuje ih u obliku obradivog, pretraživog i indeksibilnog teksta. Prepoznavanje teksta može se obavljati prilikom samog skeniranja, ali i naknadno, ovisno o tome kako je organiziran proces digitalizacije. OCR programi razlikuju bijele od crnih površina teksta te stvaraju tekstualnu datoteku. Problemi nastaju kod nedovoljno kontrastnih predložaka, tekstova sa čestim tipografskim promjenama (na primjer u rječnicima i enciklopedijama), tekstova koji sadrže znakove svojstvene drugim jezicima te tekstova sa zastarjelim ili stromodnim oblicima pisma. Tako je, na primjer, povijesne dokumente i stare novine, a da se ne govori o rukom pisanim dokumentima, gotovo nemoguće brzo, kvalitetno i ekonomski isplativo skenirati. Želi li ih se digitalizirati s naknadnom mogućnošću obrade, pretraživanja, indeksiranja, analize ili, pak, računalnog prevođenja teksta, takve je dokumente isplativije prepisati.

Mnogi projekti digitalizacije bave se digitalizacijom tiskanih dokumenata kao što su knjige i novine. To se najčešće (iako ne i isključivo) provodi skeniranjem. Uporaba OCR softvera popularan je način da se iz tako skenirane građe dobiju informacije koje se mogu uređivati i pretraživati, indeksirati, konvertirati u neki drugi format te se nad njima mogu obavljati razni drugi oblici procesiranja podataka.

Početna odluka prilikom svakog skeniranja odnosi se na vrstu slike koju se želi stvoriti (u sivoj skali ili u boji) i rezoluciju. Zapravo, mora se odrediti koliko će točaka po inču (engl. *dots-per-inch* – *dpi*) skener zabilježiti i koliko će informacija sadržavati svaka točka. Što je slika više rezolucije i ima više informacija u svakoj točki, to će rezultirajuća datoteka biti veća. Na te će odluke utjecati predviđeno korištenje digitalnog zapisa zbog kojeg se gradivo i digitalizira.

Današnji OCR programi pretvaraju kvalitetne predloške iz slike u tekst sa prosječnom točnošću od 99,95%, što je ujedno i donja granica isplativosti korištenja ovih programa. To još uvijek znači da ručno treba ispraviti jednu do dvije pogreške po stranici. Ako dolazi do više od 4 do 5 pogrešaka na 1.000 znakova, onda je isplativije gradivo ručno prepisati.²⁶ Treba napomenuti da se, ako se točnost želi podići na 99,99%, troškovi udvostručuju²⁷ zbog potrebe za povećanjem rezolucije kod skeniranja, čime se, pak, produljuje vrijeme postupka skeniranja.

OCR dobro funkcionira u većini latiničnih oblika pisama 19. i 20. stoljeća. S prije tiskanom građom ili loše otisnutim oblicima pisama bilo koje vrste, softver nailazi na prepreke koje mu oduzimaju mnogo vremena. Razbijena cjelovitost oblika slova, ligature, dvostruka slova, nejednolično crnilo tinte i starinski oblik slova softver može ne prepoznati ili krivo prepoznati, a svaki takav znak iziskuje vrijeme i trud za korekciju i verifikaciju sadržaja, što produljuje vrijeme potrebno za provedbu projekta.

Prije nego što se započne digitalizacija veće količine tekstualne građe, dobro je napraviti pokusno skeniranje. Rezultati bi trebali biti dobri kod većine modernih fontova, ali čak i s čistim tekstom i pristojnom veličinom slova, pojavit će se poneka pogreška. Njihov će broj biti to veći što je tekst gušći i nejasniji. Pode-

²⁶ *Digitization as a Means of Preservation?*, European Commission on Preservation and Access, Amsterdam, Nizozemska, listopad 1997., <<http://www.clir.org/cpa/reports/digpres/digpres.html>>, 10. lipnja 1999.

²⁷ Jann, Lynn-George, *Digitization: A Literature Review and Summary of Technical Processes, Applications and Issues*, 10. svibnja 1996., <http://web.archive.org/web/20020819040732/http://www.library.ualberta.ca/library_html/libraries/law/digit1.html>, 19. kolovoza 2002.

šavanje sjajnosti (engl. *brightness*) i rezolucije može popraviti rezultate, ali malo se može učiniti s bliedom fotokopijom ili tipografijom iz 17. i 18. stoljeća. Prepoznavanje rukopisa još uvijek je problem. Danas se za prepoznavanje rukopisa i manuskripata često primjenjuje ICR – Intelligent Character Recognition.

Nakon uspješnog optičkog prepoznavanja slova dobiva se tekst koji tek treba urediti tako da se u svakom trenutku može znati na kojoj se stranici u tiskanom izdanju nalazi onaj dio koji se u danom trenutku čita. Mnoge knjige imaju ilustracije, slike, grafikone, tablice, podnožne bilješke itd., koje također treba digitalizirati i staviti ih na njihovo mjesto u digitalnoj verziji knjige. Neki sustavi za optičko prepoznavanje znakova čak mogu približno reproducirati formatiranje izvorne stranice, uključujući slike, stupce ili druge netekstualne komponente.

Nadalje, ako je tekst namijenjen daljnjoj distribuciji putem mreže, on mora biti oblikovan kao hipertekstualni dokument. To znači da mu se moraju ugraditi poveznice (engl. *link*). One moraju biti organizirane tako da omogućuju što jednostavnije, nelinearno kretanje kroz dokument koje je ujedno i glavno obilježje hipermedija. Stoga, ako se neki tekst prebacuje u elektronički oblik, potrebno ga je i oblikovati u skladu s novim medijem u kojem se nalazi. Tako, na primjer, digitalizirati knjigu ne bi trebalo značiti samo provesti digitalizaciju i omogućiti dostupnost teksta u elektroničkom obliku. Naprotiv, u elektroničkoj verziji knjige trebalo bi hipervezama povezati tekst sadržaja, indeks, popise tablica, slika, grafikona, prikaza i sl. s pripadajućim pozicijama u knjizi te na taj način iskoristiti nelinearne mogućnosti elektroničkog medija. Upravo je takvo dodatno opremanje digitalizirane verzije nekog gradiva vremenski vrlo zahtjevno i dodatno poskupljuje proces digitalizacije.

Među najpoznatijim OCR programima su *Abby Finereader*, *Recognita* i drugi.

3.1.1. Notni zapisi

Za optičko prepoznavanje notnih zapisa koriste se OMR programi (engl. *Optical Music Recognition*). U ovom trenutku OMR softver ne prepoznaje rukom pisane note. Poseban problem kod notnog zapisa, u odnosu na tekstovni, je dvodimenzionalnost zapisa: notni tekst nema samo horizontalni/vremenski tijek (npr. jednoglasna melodija), već i vertikalna/istovremena, dakle superponirana događanja (npr. akord), što otežava razdvajanje jednog objekta od drugoga. Osim toga, u vokalno-instrumentalnoj glazbi notnom je zapisu priključen i tekst vokalne dionice tako da se mora koristiti i OCR program kao podzadatak OMR-a.

Među najpoznatijim OMR programima su *Optical Music Easy Reader* (OmeR), *Musiktek*, *VivaldiScan*, *Gamera*, *Cantor* i drugi.

Poseban je problem identifikacija glazbenog djela pomoću notnog teksta te mogućnost pretraživanja po istom principu. Budući da je notno pismo neprikladno za računalno procesuiranje, potrebno ga je nadomjestiti nekom drugom vrstom "koda" – baziranom na slovčanom i brojevnom sustavu znakova. Sustav nazvan *Plain and Easy Code*, koji se upotrebljava za bilježenje i pretraživanje notnih incipita, danas je prihvaćen kao standardni sustav za kodiranje nota i izradu notacijskih indeksa. Ovakav sustav vrlo je jednostavan i logičan, ali zahtjeva glazbenu pismenost.

Praktične napomene:

- Istražiti OCR/OMR softverske proizvode na tržištu prije odabira određenog proizvoda. Dok je OCR softver često uključen u cijenu skenera, složeniji softver obično se prodaje zasebno.
- Važan element svakog OCR projekta je identifikacija i ručno ispravljanje pogrešaka, nejasnoća i mjesta na kojima tekst nije mogao biti automatski procesiran. Stoga je važno da OCR paket ima sučelje koje je dobro prilagođeno korisniku, jer to može uštedjeti dosta vremena i truda.
- OCR najbolje radi s dokumentima koji su u dobrom stanju – presavijeni i zgužvani dijelovi, kao i gubitak boje izvorne građe povećat će broj pogrešaka. Gdje je to moguće, prije skeniranja bi trebalo izvornu građu obraditi tako da se isprave takvi dijelovi (restauriranjem ili nekim drugim postupkom).
- Prije nego što se pristupi uporabi OCR programa, kod građe koja nije u najboljem stanju, trebalo bi razmotriti prethodnu uporabu softvera za obradu slika radi uklanjanja mjesta koja su ostala bez boje i poboljšanja kontrasta.
- Prije nabave OCR softvera treba provjeriti postoje li u njegovu paketu rječnici za prepoznavanje onog jezika na kojemu je izvorna građa.

3.2. SLIKOVNO GRADIVO

Skeniranjem slikovnog gradiva se nakon digitalizacije mora kontrolirati kvaliteta zbog toga što uređaji za digitalizaciju, skeneri i digitalni fotografski aparati, ne moraju uvijek vjerno prenijeti boju. Također se događa da se prilikom digitalizacije zahvati veća površina gradiva od one koja je potrebna ili, pak, da je predložak skeniran ukoso. Zbog takvih je slučajeva slikovne zapise potrebno dodatno obraditi nekim od programa za obradu slika te ih učiniti što vjernijima originalu.

3.2.1. Matematički modeli kodiranja

Daljnja obrada digitaliziranog slikovnog gradiva odnosi se na njegovu pohranu. Slikovni se zapisi, zbog smanjenja memorijskih zahtjeva prilikom pohrane ili zbog smanjenja potrebnog prijenosnog kapaciteta tijekom distribucije putem mreže, moraju komprimirati. U takvom obliku postaju pogodniji za pohranu i prijenos. Komprimiranje je postupak pronalaženja zalihosti u binarnom zapisu. Zalihost se uklanja kodiranjem zapisa uz pomoć različitih matematičkih modela. Modeli se pri radu koriste predviđanjima koja se izvode iz temeljnih karakteristika strukture i svojstava slikovnog gradiva, čime se nadograđuju i proširuju osnovne tehnike komprimiranja binarnog zapisa koje ne uzimaju u obzir vrstu gradiva. Tako se, na primjer, dvije dodirujuće linije točaka na slici najvjerojatnije neznatno razlikuju, dok se to za dvije linije teksta ne može reći.

Dva su osnovna pristupa komprimiranju – komprimiranje bez gubitaka i komprimiranje s gubicima. Prvi oblik u potpunosti čuva informaciju, dok drugi složenim algoritmima proračunava koje dijelove može isključiti, a da pritom dođe do kontroliranog gubitka kvalitete. Zbog toga će odabir višeg stupnja kompresije značiti i nižu kvalitetu. Rezultati detaljnije analize ovih dvaju pristupa komprimiranju objašnjeni su nešto poslije.

Dodatno, komprimiranje se može promatrati i po tome na što se primjenjuje – na razinu binarnog zapisa, pri čemu se ne provodi analiza sadržaja gradiva, te na sam sadržaj, pri čemu se upravo na temelju analize sličnosti ili razlika u susjednim ili bliskim točkama sadržaja komprimira gradivo. Na primjer, algoritmi koji djeluju na razini binarnog zapisa uočavaju pravilnosti u nizovima jedinica i nula u binarnom zapisu te ih zapisuju skraćeno, dakle komprimirano. Drugi, na primjer, uočavaju koliko se puta uzastopce pojavljuje točka određene nijanse boje na slici te sliku komprimiraju tako da umjesto da zapišu nijanse svih tih to-

čaka, ponavljajući svaki put istu informaciju, zapisuju nijansu samo jednom uz zapisivanje broja koliko se puta ona ponavlja. Naravno, navedeni primjeri vrlo su jednostavni zbog boljeg načelnog razumijevanja samog postupka komprimiranja, dok su stvarni algoritmi znatno složeniji. U nastavku su, uz kratak opis, navedeni samo neki od glavnijih modela komprimiranja.

3.2.1.1. Kodiranje entropije

Ovaj model razvio je Claude E. Shannon, začetnik informacijske teorije. On je prvi shvatio temeljnu razliku između podatka, poruke i informacijskog sadržaja te ih definirao apsolutnim terminima koji postavljaju krajnje granice minimuma prosječne količine podataka koja je potrebna da se prenese poruka uz prethodno poznatu vjerojatnost pojavljivanja. Ovaj se model koristi činjenicom da je sadržaj poruke izravno povezan s vjerojatnošću njezina pojavljivanja. Što je vjerojatnost da se neki slučaj dogodi ili ne dogodi veća, to je komprimiranje bolje, jer se bilježi samo onaj podatak koji je manje vjerojatan, a takvih je malo, dok se za ostale podrazumijeva da su onakvi kakvim ih se predviđelo. Za upotrebljavanje entropije pri komprimiranju potrebno je prethodno poznavati statistiku promatranih događaja, iz čega se izračunava vjerojatnost njihova pojavljivanja.

3.2.1.2. Huffmanovo kodiranje

Huffmanovo kodiranje pronalazi najmanju moguću entropiju varijabilne duljine, povezanu s danim skupom događaja i njihovom vjerojatnošću pojavljivanja, koja se može kodirati. Zasnovano je na grupiranju događaja, određivanju njihove zajedničke vjerojatnosti te izradi binarnog stabla.

3.2.1.3. Lemper-Ziv kodiranje

Ovaj model kodiranja pripada grupi modela koji se koriste tehnikom *univerzalnog kodiranja*. To je kodiranje koje ne zahtijeva prethodno poznavanje statistike događaja, a zasniva se na pretpostavci da se bilo koji niz podataka s nekom mjerom zalihosti sastoji od ponavljanja sljedova podataka tipičnih za taj niz. Ono postiže najbolje rezultate kod velikih nizova, dok se kod malih nizova događa da konačni zapis bude veći od originala.

Lemper-Ziv kodiranje radi na principu kodiranja nizova 0 i 1 kao nekog prethodnog, prefiksnog niza uz dodatak jednog novog bita. Tako dobiveni niz postaje potencijalni prefiksni niz sljedećim nizovima. Što je početni niz duži, to je kompresija bolja, jer se dugački prefiksi tada mogu prikazati kao mali brožani indeksi.

3.2.1.4. Kodiranje po principu dugih nizova

Kodiranje po principu dugih nizova (engl. *run length encoding*) komprimira niz tako da niz točaka na slici koji ima istu vrijednost boje zapiše kao bajt oznaku (engl. *flag byte*) koja sadrži broj tih uzastopnih točaka, koju slijedi bajt sa zapisanom vrijednošću boje. Ako se dvije točke razlikuju, jednostavno ih se zapisuje kao zasebne točke. Ovaj model komprimiranja koristi PCX format zapisa slika.

3.2.1.5. Pretprocesiranje filterom predviđanja

Pretprocesiranjem filterom predviđanja (engl. *predictive filter preprocessing*) uz naknadnu upotrebu modela zasnovanih na kodiranju entropije postiže se ista razina kompresije kao i bez pretprocesiranja, ali uz velike uštede u vremenu kodiranja i potrebnoj procesorskoj snazi. Ovaj model preoblikuje niz podataka tako da se novi niz sastoji od zapisanih vrijednosti boje samo onih točaka slike koje nisu u skladu s predviđanjem da je svaka sljedeća točka jednaka prethodnoj. Ostalim se točkama dodjeljuje oznaka da su u skladu s predviđanjima. Time se dobiva niz u kojem ima mnogo istih oznaka. Takav je niz pogodan za komprimiranje, na primjer, Lemper-Ziv modelom kodiranja.²⁸

3.2.2. JPEG standard

JPEG standard stvorilo je Udruženje fotografskih stručnjaka (engl. *Joint Photographic Experts Group*) koje je oformljeno kao zajednička grupa Europske organizacije za telekomunikacijske standarde (CCITT – European Telecommunications Standards Organization) i Međunarodne organizacije za standardizaciju (ISO – International Standards Organization). JPEG standard ima 29 ugrađenih

²⁸ Opis matematičkih modela za komprimiranje prema: Cyganski et.al., Information Technology, n. dj.

različitih sustava kodiranja koji služe za komprimiranje slika. Oni se razlikuju po brzini kodiranja, postignutom stupnju kompresije te postignutom stupnju kvalitete, tj. vjernosti komprimirane slike originalu. Također je ugrađeno i 8 modela za predviđanje, koji se mogu svrstati u tri kategorije ovisno o načinu predviđanja vrijednosti boje točke. Ti modeli predviđaju da točka u liniji ima:

- istu vrijednost kao i njoj prethodna točka,
- istu vrijednost kao i točka iznad,
- vrijednost koja je kombinacija vrijednosti prethodne točke, točke iznad i točke prethodne točki iznad.²⁹

Kao što je već prije spomenuto, postoje dva načina komprimiranja: komprimiranje bez gubitaka i komprimiranje s gubicima. Rezultati analize obaju načina komprimiranja u nastavku su pobliže objašnjeni na primjeru JPEG standarda.

3.2.3. Komprimiranje bez gubitaka

Komprimiranje bez gubitaka čuva potpunu slikovnu informaciju. JPEG standard nudi samo dvije razine ovakvog komprimiranja. One se razlikuju jedino po modelu kodiranja entropije. U prosjeku je odnos komprimiranog i nekomprimiranog zapisa 2:1, što ipak čini znatno smanjenje.

3.2.4. Komprimiranje s gubicima

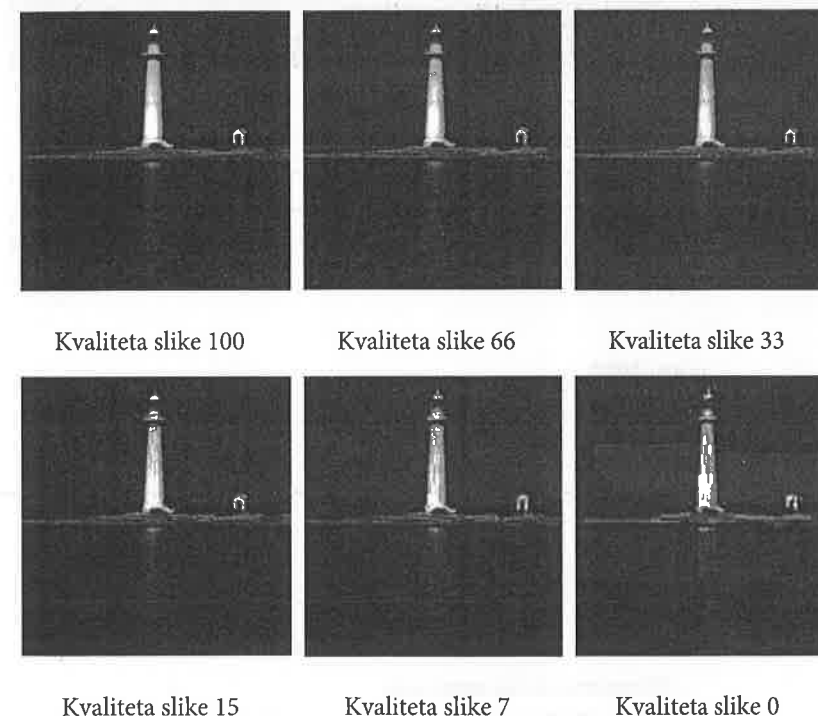
Kod kompresije s gubicima prema JPEG standardu može se birati željeni stupanj kompresije. Što je stupanj kompresije viši, to je rezultirajuća datoteka manja, ali je manja i kvaliteta slike, i obrnuto. Kompresija se provodi tako da se najprije slika pojednostavi uklanjajući kompleksnost na račun kvalitete, zatim se njezin zapis preoblikuje pretprocesiranjem filterom predviđanja, a potom komprimira upotrebom jednog od modela kodiranja entropije.

Za potrebe arhiviranja slikovnog gradiva odnos komprimiranog i nekomprimiranog zapisa ne bi smio biti veći od 20 ili 25:1 ako se želi zadržati vjerodostojnost originalu koristeći komprimiranje s gubicima. Ako je, pak, gradivo namijenjeno korištenju na multimedijском CD ili DVD-ROM-u ili prijenosu putem mreže, tj. stavljanju na Internet, mogu se postići i odnosi komprimiranja

čak do 200 ili 230:1. Naravno, pritom treba odrediti optimalni odnos kompresija – veličina – kvaliteta.

Za bolju ilustraciju svega što je do sada navedeno vezano uz komprimiranje slikovnog gradiva, u nastavku će biti prikazani rezultati provedenog testiranja. Skenirana fotografija bila je veličine 21,0 cm x 21,2 cm (širina je identična širini papira veličine A4). Skenirana je na plošnom skeneru UMAX Astra 1220S pri optičkoj rezoluciji od 800 dpi u 24-bitnoj boji po RGB sustavu. Za potrebe ovog testa korišten je softver ACDSee. Dobivena slika nije komprimirana, već je spremljena u TIFF formatu, a njezina je veličina iznosila 62.280 kb.

Fotografija u digitalnom obliku zatim je konvertirana u JPEG format, pri čemu je komprimirana. Testirana je kvaliteta komprimirane slike konvertirajući isti digitalni original svaki put s drugačijim omjerom kvalitete i kompresije, i to najprije bez optimizacije, a onda s Huffmanovom optimizacijom.



Slika 21: Odnos kvalitete i kompresije
(najbolja kvaliteta = 100, najbolja kompresija = 0)

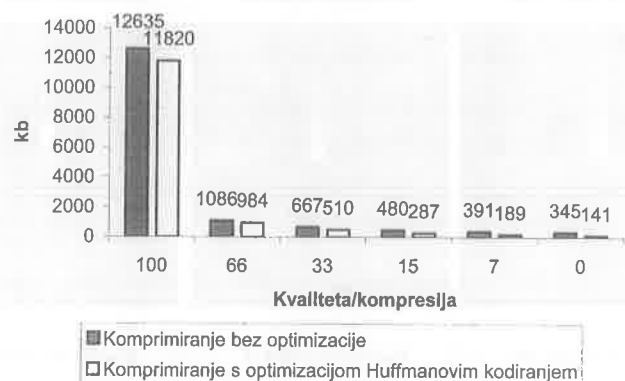
²⁹ Cyganski et.al., Information Technology, n. dj.

Na slici 21 može se vidjeti pad kvalitete uzrokovan povećanjem kompresije. Treba napomenuti da se komprimirane slike iste kvalitete komprimirane bez optimizacije i s optimizacijom Huffmanovim kodiranjem, prema vizualnoj procjeni, ne razlikuju. Razlike postoje u drugim analiziranim segmentima. Određena kvaliteta može zadovoljiti zahtjeve arhiviranja, a određena ne može. Isto tako, prilikom digitaliziranja gradiva s namjerom da se ono distribuira mrežom određena veličina datoteke je prevelika, a određena nije. Stoga je potrebno odrediti minimalnu kvalitetu koja je potrebna i maksimalnu dopuštenu veličinu zapisa, naravno uzimajući u obzir njegovu namjenu.

Tablica 3 prikazuje odnos veličina komprimiranih datoteka različitih kvaliteta, koristeći dvije od nekoliko ponuđenih opcija komprimiranja.

Tablica 3: Odnos veličina komprimiranih datoteka

Kvaliteta	Veličina komprimirane datoteke (kb)	
	Bez optimizacije	Optimiziranje Huffmanovim kodiranjem
100	12.635	11.820
66	1.086	984
33	667	510
15	480	287
7	391	189
0	345	141



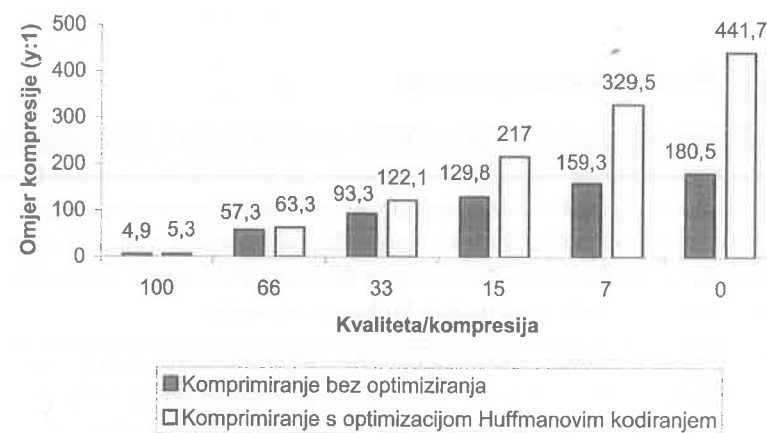
Grafikon 2: Odnos veličina komprimiranih datoteka

Iz tablice 3 može se vidjeti da komprimiranje slikovnog zapisa optimiziranjem Huffmanovim kodiranjem rezultira manjim komprimiranim zapisom. Pri višoj zadanoj kvaliteti razlika nije toliko velika koliko je kod niže zadane kvalitete, a što se može vidjeti i na grafikonu 2.

Tablica 4 prikazuje omjer postignute kompresije jednom od dviju korištenih opcija i izvornog, nekomprimiranog zapisa.

Tablica 4: Omjer kompresije

Kvaliteta	Omjer kompresije	
	Bez optimizacije	Optimiziranje Huffmanovim kodiranjem
100	4,9:1	5,3:1
66	57,3:1	63,3:1
33	93,3:1	122,1:1
15	129,8:1	217,0:1
7	159,3:1	329,5:1
0	180,5:1	441,7:1



Grafikon 3: Odnos omjera kompresije

Iz tablice 4 može se vidjeti pri kojoj se kvaliteti komprimirane slike postiže koji omjer kompresije u odnosu na nekomprimirani zapis. Da podsjetim, preporuka je da odnos kompresije slikovnoga gradiva koje se arhivira u svrhu očuvanja građe u digitalnom obliku nikako ne bi smio biti veći od 20 ili 25:1. Ako se još jednom pogleda slika 20, vidjet će se da su zahtjevi za vjernošću visoko postavljeni, a iz prethodnih tablica i grafikona se tada može zaključiti da se mogu očekivati datoteke koje zauzimaju dosta diskovnog prostora, premda znatno manje nego u nekomprimiranom obliku, te da se u predviđeno vrijeme potrebno da gradivo prođe kroz sve faze procesa digitalizacije, osim vremena potrebnog za digitalizaciju slikovnoga gradiva, mora uračunati i vrijeme potrebno za komprimiranje dobivenih zapisa. Postignuti omjeri kompresije mogu se također vidjeti na grafikonu 3.

Iz prethodnih je primjera vidljivo koje sve segmente prilikom obrade slikovnoga gradiva komprimiranjem s gubicima treba uzimati u obzir. U cijelom procesu digitalizacije komprimiranje zapisa bitan je segment. Ono izravno utječe na kvalitetu arhiviranog digitalnoga gradiva, a time i na veličinu koju digitalni zapis zauzima, što se neposredno odražava i na financijsku stranu cijelog projekta. Prethodno testiranje dalo je referentne vrijednosti koje mogu poslužiti za izračun potrebnih smještajnih kapaciteta za digitalizirane zbirke slikovnih materijala u ovisnosti o broju slika koje će biti digitalizirane, njihovoj veličini te omjeru kompresije koji se na njih namjerava primijeniti.

3.3. ZVUČNO GRADIVO

Digitaliziranom zvučnom gradivu također je potrebno kontrolirati kvalitetu te ga dodatno obraditi. S obzirom na način digitaliziranja i pritom korištene uređaje, digitalizacija zvučnoga gradiva ne pruža tolike mogućnosti za nastanak grešaka. Kod skeniranja slikovnoga gradiva, na primjer, može doći do greške prilikom skeniranja zbog optičkih svojstava samog skenera, zatim zaslon monitora može prikazivati za nijansu drukčiju boju itd. Nasuprot tome, ako se za digitalizaciju zvučnoga gradiva koristi kvalitetan uređaj za reprodukciju, kvalitetni kablovi za spoj sa računalom te kvalitetan uređaj (kartica) za digitalizaciju zvuka, malo je mjesta gdje može doći do zamjetnih grešaka. Kvalitetu zvučnog zapisa svakako treba kontrolirati, ali kad bismo usporedili ovu kontrolu s kontrolom kvalitete digitaliziranog slikovnoga gradiva, onda bismo vidjeli da kontrola slikovnog materijala gotovo uvijek rezultira potrebom za barem neznatnim dodatnim korekcijama, dok se kod zvučnog materijala to u načelu rjeđe događa. Do-

datna se obrada kod zvučnoga gradiva uglavnom odnosi na uklanjanje eventualne tišine na početku ili na kraju zapisa, do koje dolazi kad se s prihvatanjem zvučnog signala krene prerano ili završi prekasno, te na ujednačavanje glasnoće različitih snimaka u istoj zbirci.

3.3.1. MPEG standard

Zvučno gradivo, također, sadrži određenu količinu zalihosti te ga je moguće komprimirati bez gubitaka i s gubicima. Pritom se koriste matematički modeli koji rade na istoj teoretskoj osnovi kao i modeli koji se upotrebljavaju za kompresiju slikovnoga gradiva. Najpoznatiji je MPEG skup standarda koji je razvila Grupa stručnjaka za pokretne slike (engl. *Moving Picture Experts Group*), razvijen uz potporu Međunarodne organizacije za standardizaciju (ISO – International Standards Organization), a koji se koristi za kompresiju digitalnih zvučnih i video zapisa. Tim se standardom koriste svi programi koji mogu zapisivati zvučne zapise u popularnom MP3 formatu. MP3 je zapravo MPEG-1 ili MPEG-2 zvučni zapis sloja 3. MPEG skup standarda može se koristiti za komprimiranje bez gubitaka i s gubicima.

3.3.2. Komprimiranje bez gubitaka

Komprimiranje bez gubitaka čuva potpunu zvučnu informaciju. Testiranje programa za komprimiranje zvučnih zapisa, pri čemu su testirani zapisi bili frekvencije 44.1 kHz, te 16 bitnog stereo signala, pokazali su da se očekivani rezultati kompresije kreću od "60 do 70% nekomprimiranog zapisa za pop, rock, techno i drugu glasnu i bučnu muziku, a 35 do 60% za tiša pjevana i orkestralna djela".³⁰

3.3.3. Komprimiranje s gubicima

Zvučno gradivo može se komprimirati s gubicima u kvaliteti uz mogućnost odabira kvalitete, tj. stupnja kompresije. Kao i kod slikovnoga gradiva, što je stupanj kompresije viši, to je rezultirajući zapis manji, ali je niža i kvaliteta zvuka, i

³⁰ Whittle, Robin, *First Principles – Lossless Compression of Audio*, Centre for Signal Processing, Nanyang Technological University, Singapore, 17. siječnja 2001., <<http://www.firstpr.com.au/audiocomp/lossless/>>, 12. veljače 2001.

obrnuto. Koriste se različiti modeli komprimiranja. Tako se, na primjer, kompresija ostvaruje korištenjem jednog od dvaju kanala kod stereo zapisa za zapisivanje informacija koje su jednake kod lijevog i desnog kanala, a drugog kanala za zapisivanje samo onih informacija po kojima se razlikuju. Nadalje, kompresija se može postići kodiranjem samo onih bitova zvučnog zapisa koji se smatraju važnima za postizanje stereo zvuka. Također se koristi dinamičko kodiranje uzorkovanih djelića zvučnog signala, pri čemu se za svaki djelić odabire najbolji model kodiranja. Ovim načinom povećava se odnos komprimiranja, što rezultira manjim zapisom s višom kvalitetom digitaliziranoga gradiva.

Kod MP3 zapisa odnos kvalitete i veličine rezultirajućeg komprimiranog zapisa u načelu ovisi o broju bitova koji se prenose po jedinici vremena (engl. *bitrate*) – što je veći broj bitova u sekundi, to je stupanj kvalitete viši, a stupanj kompresije niži. U skladu s tim postoji komprimiranje nepromjenjivim brojem bitova (engl. *CBR – Constant Bitrate Encoding*) i promjenjivim brojem bitova (engl. *VBR – Variable Bitrate Encoding*).

3.3.3.1. Komprimiranje nepromjenjivim brojem bitova

Metoda komprimiranja nepromjenjivim brojem bitova rezultira konzistentnim brojem bitova po sekundi u cijelom zvučnom zapisu na račun kvalitete zvuka i/ili vremena potrebnog za komprimiranje. Prilikom kodiranja jednak se broj bitova dodjeljuje zahtjevnim, kao i jednostavnim dionicama zapisa, tj. jednak onima koje sadrže mnogo informacija kao i onima koje, poput tišine, sadrže malo informacija. Stoga je komprimirani zapis na zahtjevnijim dionicama nešto lošije kvalitete, dok na jednostavnim sadrži neke neiskorištene bitove, pri čemu je njegova zalihost nepotrebno velika. Ako postoji potreba za ograničavanjem i predviđanjem veličine rezultirajućeg zapisa, onda je ova metoda dobar izbor jer se njome može predvidjeti veličina zapisa tako da se pomnoži broj bitova koji se prenosi i dužina trajanja zvučnog zapisa.³¹

Tablica 5 prikazuje odnos broja bitova, kod korištenja metode komprimiranja nepromjenjivim brojem bitova, i kvalitete rezultirajućeg zvučnog zapisa.

Tablica 5: Komprimiranje nepromjenjivim brojem bitova

kbit/s	Algoritam	Opis kvalitete
16	MPEG-2	Telefonska kvaliteta zvuka – najmanja veličina zapisa
24		Telefonska kvaliteta zvuka – manja veličina zapisa
32		Kvaliteta AM radio signala – najmanja veličina zapisa
48		Kvaliteta AM radio signala – manja veličina zapisa
56		Kvaliteta FM radio signala – najmanja veličina zapisa
64		Kvaliteta FM radio signala – manja veličina zapisa
80		
96	MPEG-1	Kvaliteta približna CD kvaliteti – dobar izbor za reprodukciju na prijenosnim MP3 uređajima
112		CD kvaliteta – najbolji izbor za reprodukciju na prijenosnim MP3 uređajima
128		CD kvaliteta – najbolji izbor za većinu korisnika
160		Arhivirajuća kvaliteta – za reprodukciju na HI-FI uređajima visoke kvalitete
192		
224		Arhivirajuća kvaliteta – za reprodukciju na HI-FI uređajima najviše kvalitete
256		
320		

3.3.3.2. Komprimiranje promjenjivim brojem bitova

Metoda komprimiranja promjenjivim brojem bitova rezultira konzistentnom kvalitetom zvuka u cijelom zvučnom zapisu. Kvaliteta se postiže detekcijom broja potrebnih bitova za svaki djelić zapisa, ovisno o prethodno određenoj maksimalnoj kvaliteti. Kada program za vrijeme komprimiranja naiđe na zahtjevniju dionicu, njoj odredi veći broj bitova kako bi se zadržala kvaliteta, a jednostavnijoj dionici odredi manji broj bitova kako ne bi došlo do pojavljivanja praznih bitova, tj. bilježenja nepotrebne količine zalihosti. Stoga se rezultirajući komprimirani zapis sastoji od promjenjivog broja bitova po jedinici vremena. Ova metoda komprimiranja koristi se kada konstantna kvaliteta čini imperativ, a predvidljivost veličine zapisa i vremena potrebnog za komprimiranje nisu toliko važni.³²

Tablica 6 prikazuje odnos broja bitova, kod korištenja metode komprimiranja promjenjivim brojem bitova, i kvalitete rezultirajućeg zvučnog zapisa.

³¹ Prema: *Xing Audio Catalyst 2.10 Help*, s.v. Understanding Constant Bitrate Encoding.

³² Prema: *Xing Audio Catalyst 2.10 Help*, s.v. Understanding Variable Bitrate Encoding.

Tablica 6: Komprimiranje promjenjivim brojem bitova

broj kbit/s	Algoritam	Opis kvalitete
nizak	MPEG-1	Kvaliteta približna CD kvaliteti – dobar izbor za reprodukciju na prijenosnim MP3 uređajima
normalan-nizak		CD kvaliteta – najbolji izbor za reprodukciju na prijenosnim MP3 uređajima
normalan		CD kvaliteta – najbolji izbor za većinu korisnika
normalan-visok		Arhivirajuća kvaliteta – za reprodukciju na HI-FI uređajima visoke kvalitete
visok		Arhivirajuća kvaliteta – za reprodukciju na HI-FI uređajima najviše kvalitete

Za bolju ilustraciju u nastavku će biti navedeni rezultati testiranja provedenog na zvučnom zapisu Beethovenove 5. simfonije u C-molu, op. 67, 1. stavak Allegro con brio, u trajanju od 8 minuta i 9 sekundi, koji su izveli Bečki simfoničari, a koji se nalazio na CD mediju, dakle u digitalnom obliku. Za potrebe ovog testa korišten je softver Xing Audio Catalyst. Nekomprimirani je zapis bio zapisan u WAV formatu, a njegova je veličina iznosila 84.294 kb.

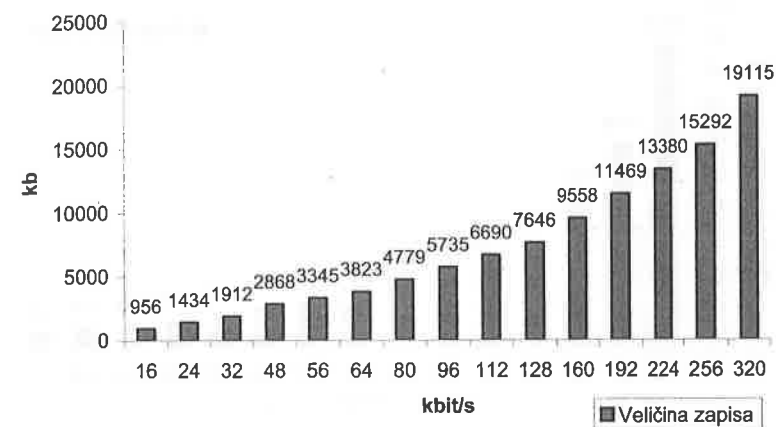
Tablica 7: Veličina komprimiranog zapisa kod CBR komprimiranja

CBR (kbit/s)	Veličina (kb)
16	956
24	1.434
32	1.912
48	2.868
56	3.345
64	3.823
80	4.779
96	5.735
112	6.690
128	7.646
160	9.558
192	11.469
224	13.380
256	15.292
320	19.115

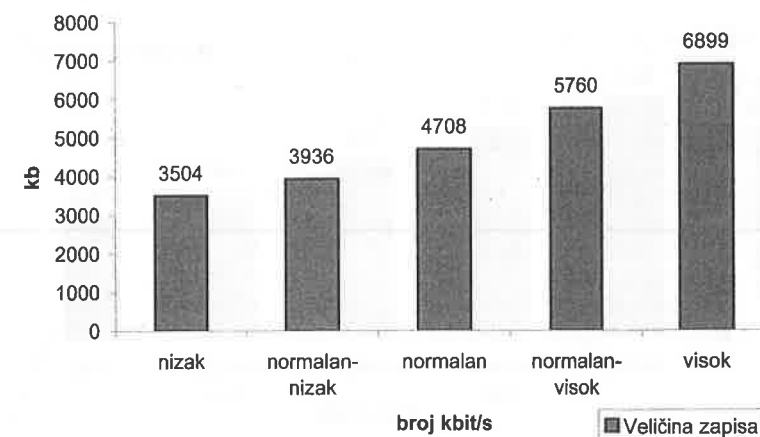
Tablica 8: Veličina komprimiranog zapisa kod VBR komprimiranja

VBR (kbit/s)	Veličina (kb)
nizak	3.504
normalan-nizak	3.936
normalan	4.708
normalan-visok	5.760
visok	6.899

Zvučni zapis u digitalnom obliku konvertiran je u MP3 format, pri čemu je komprimiran. Testirana je kvaliteta zvučnog zapisa konvertirajući isti digitalni original svaki put s drugačijim omjerom kvalitete i kompresije, i to najprije komprimiranjem nepromjenjivim brojem bitova (CBR), a onda promjenjivim brojem bitova (VBR). Tablice 7 i 8, te grafikoni 4 do 7 prikazuju rezultate testiranja.

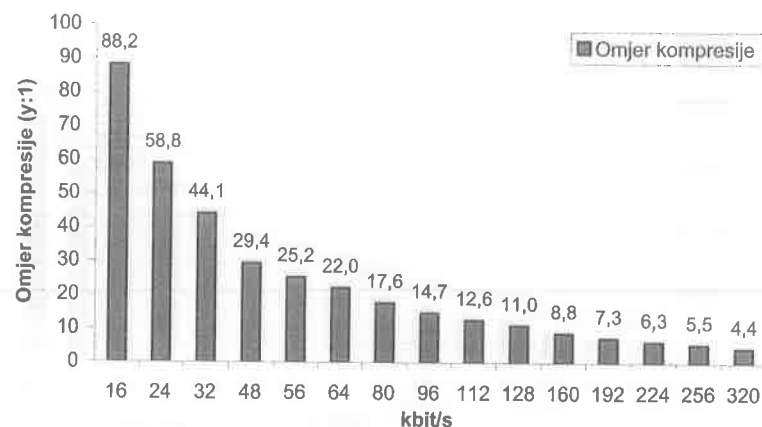


Grafikon 4: Veličina komprimiranog zapisa kod CBR komprimiranja

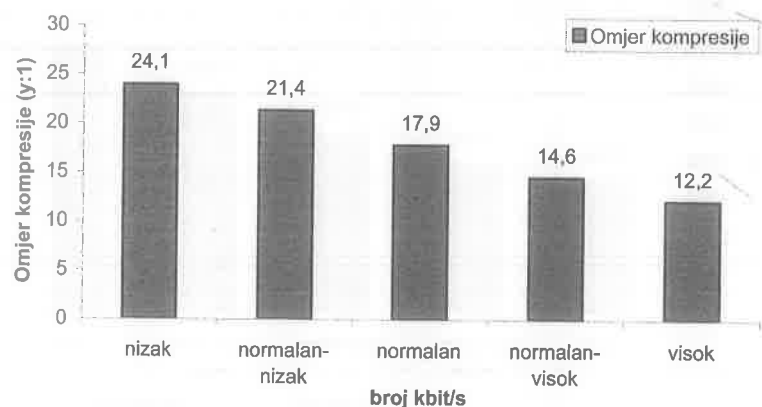


Grafikon 5: Veličina komprimiranog zapisa kod VBR komprimiranja

Iz grafikona 4 i 5 može se vidjeti da se komprimiranjem promjenjivim brojem bitova postižu manje veličine digitalnih zapisa. Rezultirajuća veličina zapisa najviše, arhivirajuće kvalitete tek je neznatno veća od one koja se postiže komprimiranjem pri 112 kbit/s s nepromjenjivim brojem bitova, što odgovara CD kvaliteti.



Grafikon 6: Omjer postignute kompresije kod CBR kompresije



Grafikon 7: Omjer postignute kompresije kod VBR kompresije

Grafikoni 6 i 7 prikazuju omjere postignute kompresije. Usporedbom ovih dvaju grafikona vidi se da je dobiveni omjer postignute kompresije arhivirajuće kvalitete kod zapisa s promjenjivim brojem bitova približno jednak omjeru kompresije najniže FM kvalitete koristeći metodu s nepromjenjivim brojem bitova.

Iz prethodnih tablica i grafikona vidljivo je što sve treba prilikom obrade zvučnoga gradiva komprimiranjem s gubicima uzimati u obzir. U procesu digitalizacije zvučnoga gradiva, baš kao i slikovnog, komprimiranje zapisa bitan je segment kako zbog kvalitete rezultirajućih zapisa, tako i zbog financijskog aspekta njihove pohrane. Iz tablica i grafikona također se mogu iščitati vrijednosti koje su preporučene za pojedine namjene digitaliziranoga gradiva. Tako za arhiviranje gradiva treba koristiti one opcije koje daju najbolju kvalitetu uz najbolji odnos kvaliteta – veličina zapisa, dok se za zvučno gradivo koje je namijenjeno prijenosu mrežom može koristiti one opcije koje daju nešto lošiju kvalitetu, ali znatno manji zapis. Na instituciji koja provodi digitalizaciju gradiva je da odluči kolika joj je kvaliteta gradiva potrebna, ovisno o njegovoj namjeni, te ako su joj sredstva ograničena, hoće li digitalizirati više gradiva nižom ili manje gradiva višom kvalitetom. Prethodno testiranje dalo je referentne vrijednosti koje mogu poslužiti za izračun potrebnih smještajnih kapaciteta za digitalizirane zbirke zvučnih materijala u ovisnosti o njihovu broju, njihovoj duljini te omjeru kompresije koji se na njih namjerava primijeniti. Tablica 9 prikazuje referentne vrijednosti koje proizlaze iz međusobne kombinacije frekvencije uzorkovanja, dubine bita i broja korištenih kanala.

Tablica 9: Referentne vrijednosti digitalizacije jednog sata zvučnog zapisa³³

Frekvencija uzorkovanja	Dubina bita	Broj kanala	Veličina rezultirajuće datoteke
44,1 kHz	16	2 (stereo)	591 MB (0.59 GB)
44,1 kHz	16	1 (mono)	296 MB (0.30 GB)
44,1 kHz	24	2 (stereo)	887 MB (0.87 GB)
44,1 kHz	24	1 (mono)	444 MB (0.44 GB)
96 kHz	24	2 (stereo)	1.931 MB (1.93 GB)
96 kHz	24	1 (mono)	966 MB (0.97 GB)

Napomena: U tablici nisu prikazane vrijednosti kod višekanalnih (engl. *surround*) zvučnih zapisa, no te se vrijednosti mogu relativno točno izračunati množenjem vrijednosti za osnovni, mono kanal s brojem kanala prisutnih u višekanalnoj snimci.

³³ Preuzeto i prilagođeno iz: *Formati datoteka za pohranu i korištenje*, n. dj., str. 11.

Za instituciju koja provodi digitalizaciju zvučnih zapisa javlja se još jedna važna dvojba. Riječ je, naime, o pitanju treba li zadržati izvorno digitalizirani zvučni signal sa svim šumovima koji su na njemu prisutni, kao što je, primjerice, "krckanje" prisutno kod reprodukcije melodija s vinilskih ploča, ili pročitati taj signal, pri čemu se može diskutirati o tome gubi li se time njegova autentičnost ili ne. Ova dvojba i ne mora postojati, jer se institucija, na primjer, može odlučiti arhivirati izvorno digitalizirani, nepročišćeni oblik i njegovu pročišćenu verziju. Ako se odluči za čišćenje zvučnog signala, to bi svakako trebala provoditi na nekomprimiranom obliku zapisa, a tek ga potom eventualno komprimirati. Programi koji se u tu svrhu mogu koristiti nazivaju se "de-klikeri", a oni rade na principu detekcije i uklanjanja raznih vrsta nepravilnosti u zvučnom signalu. Kad program pronade određenu nepravilnost, primjerice mjesto gdje se čuje "krckanje", on je briše i zamjenjuje zvukom iste duljine trajanja koji stvara na temelju zajedničkih frekvencijskih karakteristika zvuka koji je toj nepravilnosti prethodio i onoga koji je slijedio iza nje.³⁴

"Za trajnu pohranu zvučnog zapisa preporuča se Broadcast Wave Format (BWF),³⁵ format identičan Microsoftovom WAVE formatu, uz dodatak metapodataka. Format je odredila Europska unija radiotelevizija (engl. *European Broadcasting Union*) (...) Prihvatljivi formati za korisničke kopije su MPEG 1/2 Layer 3 (.mp3) i MPEG 4/AAC."³⁶ Kod trajne pohrane potrebno je voditi računa o pohrani svih prisutnih kanala – jedan kod mono zapisa, dva kod stereo zapisa i više kod višekanalnih (engl. *surround*) zapisa.

3.4. FILMSKO I VIDEO GRADIVO

Obrada i kontrola kvalitete filmskog i video gradiva u načelu bi trebala biti vrlo slična postupcima obrade i kontrole kvalitete kod slikovnog i zvučnoga gradiva, što je i logično s obzirom na to da se ono sastoji od niza slika i zvučnog signala. Ipak, komprimiranje tog gradiva znatno je zamršeniji postupak. Komprimiranje je s korisničke strane olakšano upotrebom intuitivnog korisničkog su-

³⁴ Princip na kojem radi ova metoda uklanjanja šumova vrlo je sličan digitalizaciji korištenjem interpolirane rezolucije, koja umjetno povećava rezoluciju umećući piksele čija je vrijednost izračunata na temelju prosječne vrijednosti susjednih piksela. Razlika je, naravno, u tome što je u jednom slučaju riječ o obradi digitaliziranog zvučnog signala, a u drugom o digitalizaciji slikovnih zapisa.

³⁵ Specifikacija BWF standarda javno je dostupna na: <http://tech.ebu.ch/docs/tech/tech3285s1.pdf>.

³⁶ *Formati datoteka za pohranu i korištenje*, n. dj., str. 13.

čelja, ali je matematička strana složena, pa stoga na ovom mjestu nije moguće dublje zalaziti u opis tih metoda.

Filmsko i video gradivo gotovo se uvijek komprimira zbog toga što bi nekomprimirano zauzimalo previše diskovnog prostora, a reprodukcija ne bi bila glatka. Najpoznatiji standard za komprimiranje video gradiva i pripadajućeg zvučnog signala je MPEG standard. O samom standardu i segmentu koji se odnosi na zvučni signal već je bilo riječi u odjeljku vezanom uz obradu digitaliziranog zvučnoga gradiva. No, ovdje valja napomenuti da MPEG standard u sebi sadrži algoritme za kodiranje i dekodiranje komprimiranih zapisa (CODEC) te da sintaksa MPEG video standarda omogućava učinkovit način prikaza niza slika u kompaktnijem, kodiranom obliku. Tehnike komprimiranja filmskog i video gradiva koje koriste ovaj standard analiziraju karakteristike koje su zajedničke nizovima slika, koriste ih kako bi detektirale prostornu zalihost, vremensku zalihost, jednolikost gibanja itd., te koristeći različite matematičke metode, jednokratnim zapisivanjem ponavljajućih segmenata komprimiraju zapis.

MPEG standard u sebi objedinjuje tri standarda:

MPEG-1 – standard koji propisuje kvalitetu prijenosa video signala preko cijelog zaslona, reproduciranog sa CD-ROM-a dvostruke brzine uz VHS kvalitetu;

MPEG-2 – standard koji propisuje kvalitetu prijenosa video signala putem satelita, kabela ili zemaljskih uređaja te DVD-a uz SVHS kvalitetu;

MPEG-4 – standard koji propisuje kvalitetu prijenosa video signala putem niskopropusnih kanala, primjerice Interneta, mobilnih telefona itd.³⁷, te kodiranje prema HDTV standardu³⁸.

Rad na razvoju MPEG-3 standarda je prekinut jer je zaključeno da je područje koje MPEG-2 standard opisuje tim standardom dovoljno dobro propisano.

Dva su osnovna načina komprimiranja filmskog i video gradiva s gubicima: komprimiranje nepromjenjivim brojem bitova (engl. *CBR – Constant Bitrate Encoding*) i promjenjivim brojem bitova (engl. *VBR – Variable Bitrate Encoding*), a uz svaki od ovih načina komprimiranja gradiva koristi se više složenih matematičkih metoda. S obzirom na to da se radi o istom standardu, principi koji vrijede pri kodiranju video gradiva identični su principima koji su opisani kod komprimiranja zvučnoga gradiva. No, ipak postoje neke posebnosti video

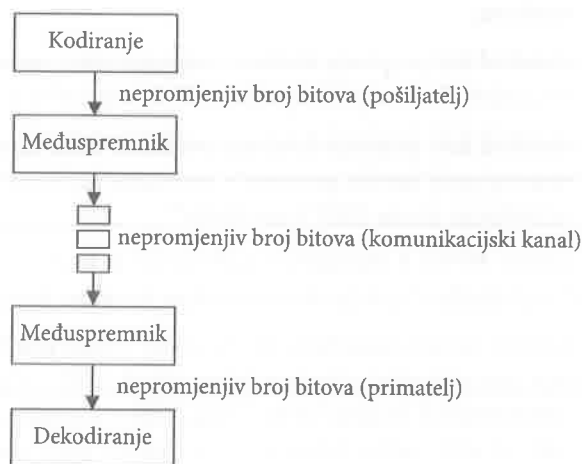
³⁷ *Moving Picture Experts Group (MPEG)*, <<http://www.chiariglione.org/mpeg/>>, 5. veljače 2001.

³⁸ Standard za prikaz televizijskog signala visoke razlučivosti.

gradiva, prije svega vezane uz veličinu zapisa. Naime, kod reprodukcije video gradiva postoje minimalni zahtjevi za kvalitetom slikovnog i zvučnog signala. Stoga se komunikacijskim kanalom, lokalno ili na daljinu, treba prenijeti veća količina podataka u jedinici vremena kako bi se održala stalna kvaliteta pri kojoj ne dolazi do istrzanosti slike ili neusklađenosti slike i zvuka.

3.4.1. Prijenos zapisa s nepromjenjivim brojem bitova

Kod zapisa s nepromjenjivim brojem bitova mora se osigurati neprekidan prijenos kodiranih podataka putem komunikacijskog kanala s unaprijed određenom nepromjenjivom propusnošću, ali bez uzrokovanja nedovoljne ili, pak, prevelike količine podataka u međuspremniku (engl. *buffer*) na prihvatnoj strani kanala (dijagram 1). Neprekidnost prijenosa komunikacijskim kanalom lako se postiže ako je riječ o prijenosu podataka sa, recimo, DVD medija na zaslon računala, ali ju je dosta teško postići kada je riječ o prijenosu putem mreže gdje je opterećenost i propusnost varijabilna. Ovaj način prijenosa video gradiva koristi se prilikom reprodukcije s DVD medija, digitalne video trake, čvrstog diska, ili prijenosa mrežom ako je osiguran zasebni kanal.³⁹



Dijagram 1: Prijenos s nepromjenjivim brojem bitova

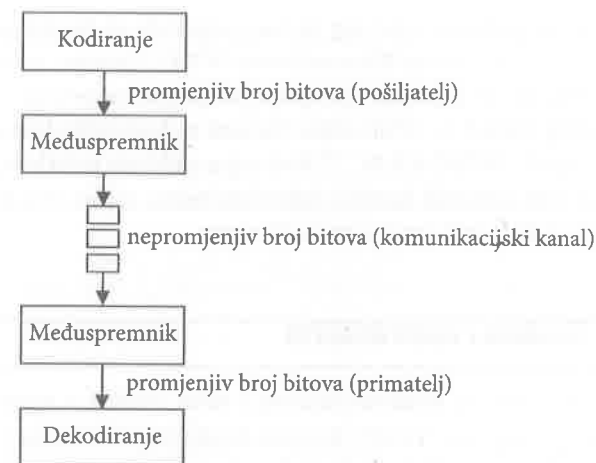
³⁹ MPEG-2 Frequently Asked Questions, Berkeley Multimedia Research Center, <<http://bmrc.berkeley.edu/frame/research/mpeg/mpeg2faq.html>>, 18. veljače 2001., s.v. What are constant and variable bitrate streams?

3.4.2. Prijenos zapisa s promjenjivim brojem bitova

Kod zapisa s promjenjivim brojem bitova postoji jednostavan i složen prijenos podataka. Oba načina prijenosa s lakoćom osiguravaju stalnu kvalitetu slike i zvuka, ali se ne pokoravaju pravilima koja postavljaju međuspremnici.

Jednostavan prijenos podataka s promjenjivim brojem bitova koristi se kada medij na kojem je pohranjeno digitalizirano filmsko ili video gradivo djeluje i kao međuspremnik.

Složen prijenos podataka s promjenjivim brojem bitova (dijagram 2) koristi se kod diskovnih uređaja kojima se može upravljati na način da upućuju na prijenos promjenjivu količinu podataka, zatim kod istovremenog prenošenja podataka putem više kanala ili kod mrežnih video sustava kod kojih je prosječan prijenos mrežom stalan, ali je različit broj korisnika koji zahtijevaju određeno filmsko ili video gradivo.⁴⁰



Dijagram 2: Prijenos s promjenjivim brojem bitova

Kao i kod svake do sada opisane vrste gradiva, tako je i kod obrade filmskog i video gradiva potrebno prethodno odrediti cilj digitalizacije i ciljanu skupinu korisnika te predvidjeti opremu kojom se oni služe, jer o tome ovisi kojom će kvalitetom i kojom vrstom komprimiranja gradivo biti obrađeno, što će opet bitno utjecati i na financijsku stranu cijelog projekta.

⁴⁰ MPEG-2 FAQ, n. dj.

S obzirom na to da se za svaku vrstu gradiva preporuča čuvanje arhivske kopije u nekomprimiranom obliku, tako je i za filmsko i video gradivo. No, ovo gradivo u elektroničkom nekomprimiranom obliku zahtijeva veliku količinu smještajnog kapaciteta pa se stoga preporuča čuvati ga u izvornom obliku (preporučljivo DV) na izvornom mediju u idealnim okolinskim uvjetima. Za izradu korisničkih kopija preporuča se korištenje formata MPEG-1 za zapise VHS kvalitete, MPEG-2 za zapise DVD kvalitete te MPEG-4 za zapise web-kvalitete i HDTV kvalitete.

3.5. TRODIMENZIONALNO GRADIVO

Nakon skeniranja 3D skenerima, kao što je to već i spomenuto, nastaju 3D modeli skeniranih predmeta ili prostora. Riječ je o takozvanim "žičanim" (engl. *wireframe*) modelima. Dodatna obrada nakon skeniranja podrazumijeva izradu 3D modela što sličnijeg originalu – po virtualnoj veličini (ispravnom uvećanom ili umanjenom mjerilu) i po izgledu. Apliciranjem određenih tekstura na osnovni žičani model tako se mogu dobiti vrlo vjerni modeli koji mogu vjerodostojno prikazivati skenirane izvorne objekte. Takvi, gotovi modeli mogu tada poslužiti u razne svrhe, od prezentacije kulturno-spomeničke baštine u elektroničkoj okolini putem mreže pa sve do osnove za izradu vjernih kopija originala. U potonjem slučaju koriste se 3D printeri, tj. uređaji koji iz raznih materijala mogu, uklanjanjem viška materijala, izraditi kopiju 3D objekta. Riječ je uglavnom o CNC glodalicama. Ipak, prije nego što se objekti puste u proces printanja, moraju se uvući u CAM (engl. *Computer Assisted Manufacturing*) program koji izračunava putanje kretanja alata za uklanjanje viška materijala. Ovim postupkom se, također, vrlo lako mogu izraditi negativni digitaliziranog objekta i zatim isprintati kalupi za lijevanje bronce. Upravo u tom slučaju digitalizacija i naknadna obrada 3D modela na računalu pokazuje svoju pravu snagu i vrijednost. Naime, poznato je da se bronca prilikom lijevanja i hlađenja skuplja pa je zbog toga klasičnim metodama izrade kalupa nemoguće dobiti, primjerice, bistu koja je po veličini identična izvorniku. Ona je u tom slučaju, zbog skupljanja materijala, uvijek nešto manja. No, korištenjem računala, računalni 3D model moguće je povećati upravo za omjer stiskanja bronce i tako dobiti odljev koji je iste veličine kao i izvornik. Iz svega navedenoga proizlazi da je, baš kao i za prethodne vrste gradiva, naknadna obrada digitaliziranoga gradiva vrlo važan korak u cijelom procesu digitalizacije.

4. ZAŠTITA GRADIVA U ELEKTRONIČKOJ OKOLINI

Zaštita digitaliziranoga gradiva ima dvojaku ulogu: zaštitu od neovlaštenog pristupa, kopiranja i daljnjeg distribuiranja te dokazivanje autentičnosti gradiva. Mehanizmi zaštite mogu se podijeliti u nekoliko skupina:

- mehanizmi koji se odnose na zaštitu i osiguranje identiteta računalnih operativnih sustava, kao što su dodjeljivanje prava pristupa određenim datotekama koja se dodjeljuju na razini sustava,
- mehanizmi vezani uz prava i obveze prema vlasnicima i distributerima koji na razini sustava određuju smiju li korisnici pristupiti određenim sadržajima bez povrede tih prava; ovi mehanizmi čine proširenje prethodnih mehanizama,
- mehanizmi šifriranja (zakrivanja) čine digitalno gradivo čitljivim samo onim korisnicima koji su legalno nabavili ključ za dešifriranje (raskrivanje),
- mehanizmi postojanog šifriranja (engl. *persistent encryption*) dopuštaju korisnicima upotrebu gradiva, dok sustav dešifrira samo one dijelove koji su trenutačno potrebni, a ostale drži šifriranim,
- mehanizmi digitalnih potpisa i digitalnih vodenih žigova ugrađuju informaciju o vlasniku ili vlasništvu u digitalno gradivo.⁴¹

4.1. MEHANIZMI ZAŠTITE SUSTAVA

Postoje mnogi mehanizmi zaštite sustava. Jedan od najosnovnijih je kvalitetno upravljanje razinama pristupa. Naime, određivanjem dopuštene razine pristupa svakom se korisniku može točno odrediti što smije, a što ne, tj. koji su mu podatci, servisi i usluge na mreži dostupni, a koji nisu. U tom kontekstu, uvijek

⁴¹ *The Digital Dilemma. Intellectual Property in the Information Age*, National Academy of Sciences, SAD, National Academy Press, 2000., str. 154 digitalne verzije, <http://books.nap.edu/html/digital_dilemma/>, 15. srpnja 2000.

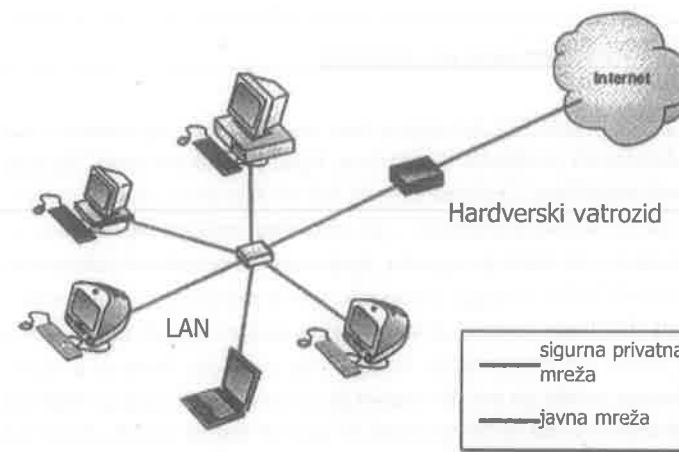
je dobro neprestano isticati koliko je važno (o)čuvati pristupnu lozinku tajnom, ne oblikovati je tako da ona može biti lako otkrivena (npr. datum rođenja korisnika ili neke njoj/njemu bliske osobe ili kombinacije tih datuma) te ne zapisivati je na lako dostupno mjesto (npr. na ceduljicu zalijepljenu na monitor ili na vidljivom mjestu u prvoj ladici). Treba uvijek imati na umu i da svatko tko neovlašteno koristi tuđu pristupnu šifru u virtualnome svijetu čini isti prekršaj kao da se u stvarnome svijetu predstavlja tuđim identitetom ili koristi identifikacijske dokumente s tuđim imenom.

Sljedeća razina zaštite je postavljanje antivirusne i slične zaštite protiv malicioznog djelovanja. Takvu zaštitu potrebno je redovito osvježavati, primjerice novim antivirusnim definicijama. Taj se postupak, također, odnosi i na redovito instaliranje ispravaka softvera koji se koriste na računalnim sustavima i softverske nadogradnje hardvera (engl. *fixes, patches, updates, service packs*).

Nikako ne treba zaboraviti i postavljanje vatrozida (engl. *firewall*). On može biti softverskog ili hardverskog tipa. Za zaštitu pojedinih računala uglavnom će se koristiti oni softverskog tipa, dok će se za zaštitu cjelokupne mreže neke institucije koristiti oni hardverskog tipa. U tom slučaju vatrozid je posebno "namjensko računalo koje sadrži posebne sigurnosne programe i posebno odabrane ulazno-izlazne priključke, a nadzire i upravlja signalima koji izlaze iz područne mreže i osobito onima koji trebaju ući⁴²." On se tada uglavnom postavlja između servera i vanjskog dijela mreže te štiti od raznih nepoželjnih postupaka, primjerice neovlaštenog udaljenog pristupa nekom računalu (engl. *remote login*), orkestriranih istovremenih zahtjeva za informacijama s mnogih, neovlašteno preuzetih, računala u svrhu namjernog zagušenja servera (engl. *(Distributed) Denial of Service – (D)DOS*), neželjene elektroničke pošte (engl. *spam*) i mnogih drugih malicioznih postupaka. Vatrozid funkcionira tako da filtrira ulaz informacija, ali i prema vanjskom, neštićenom, dijelu mreže ne propušta internetske (IP) adrese pojedinih računala u štijećenoj zoni. Upravo zbog toga nije moguće adresirati pojedino računalo te na nj ili s njega provoditi daljnje neovlaštene aktivnosti. Naravno, korištenje vatrozida pruža još mnoge dodatne sigurnosne mogućnosti koje prelaze okvire ovog orijentacijskog pregleda.

Na kraju, ipak, treba napomenuti da ne postoji savršeni sustav zaštite. Jedini način zaštite od napada izvan sustava jest potpuno ograničenje pristupa – kako onog putem mreže, dakle fizičkog nepovezivanja s mrežom, tako i onog fizičkog. No, takve mjere zaštite primjerene su nekim segmentima vojnih, obavje-

štajnih ili sličnih struktura, dok je, u kontekstu ove knjige, ipak cilj digitalizirati gradivo kako bi ono postalo dostupno ili dostupnije krajnjim korisnicima. Ispravno postavljanje višerazinske zaštite svakako će pridonijeti većoj sigurnosti sustava. U nastavku će biti objašnjene neke mjere dodatne zaštite (sadržaja) gradiva u elektroničkom obliku koje se mogu implementirati i kontinuirano provoditi neovisno o tome na koji je način zaštićen sustav u kojem je to gradivo pohranjeno.



Slika 22: Organizacija zaštite lokalne mreže (LAN) korištenjem vatrozida

4.2. ŠIFRIRANJE

Tehnike šifriranja čine vrlo važan segment upravljanja digitalnim gradivom i intelektualnim vlasništvom. Cilj šifriranja je zaštititi digitalno gradivo od neovlaštenog korištenja čineći ga nečitljivim sve dok se ne dešifrira. Mehanizmi šifriranja mogu se koristiti za različite primjene osiguranja, uključujući:

- osiguranje privatnosti i tajnosti,
- osiguranje integriteta podataka,
- mogućnost utvrđivanja autentičnosti ili identifikacije – utvrđivanje identiteta osobe, računalnog terminala, kreditne kartice itd.,
- mogućnost utvrđivanja autentičnosti poruke – utvrđivanje vjerodostojnosti izvora informacija,

⁴² Kiš, Informatički rječnik, n. dj., str. 390, s.v. firewall machine.

- mogućnost ugradnje digitalnog potpisa u poruku,
- mogućnost autorizacije – mogućnost prijenosa ovlasti na drugu fizičku ili pravnu osobu,
- mogućnost izdavanja digitalnih ovjera – potvrda da informacija dolazi iz provjerenog izvora,
- mogućnost svjedočenja – potvrda stvaranja ili postojanja određene informacije,
- mogućnost izdavanja računa – potvrda o primanju informacije,
- mogućnost potvrđivanja – potvrda o pružanju određene usluge,
- mogućnost dodjele vlasničkih prava – dodjela prava nekoj fizičkoj ili pravnoj osobi za korištenje i/ili daljnju prodaju gradiva,
- osiguranje anonimnosti,
- osiguranje nemogućnosti odbijanja neke prethodno dogovorene obveze,
- osiguranje mogućnosti opoziva autorizacije i ovjere.⁴³

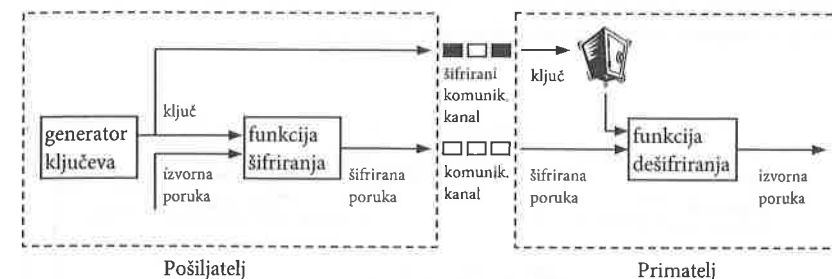
Kod šifriranja digitalnoga gradiva nije bitno o kojoj je vrsti gradiva riječ. Važno je da se radi o digitalnom zapisu. Princip šifriranja sastoji se u tome da se digitalni zapis uz pomoć ključa šifriranja preoblikuje tako da bude neprepoznatljiv te da ga se bez poznavanja ključa ne može vratiti u izvorni oblik. Dva su glavna načina šifriranja digitalnoga gradiva: šifriranje simetričnim ključem (engl. *symmetric-key encryption*) i šifriranje javnim ključem (engl. *public-key encryption*).

4.2.1. Šifriranje simetričnim ključem

Kod šifriranja simetričnim ključem isti se ključ koristi za šifriranje i dešifriranje poruke, tj. nekog digitalnoga gradiva koje se želi poslati primatelju putem komunikacijskog kanala. Takav sustav sastoji se od tri dijela: generatora ključeva, funkcije šifriranja i funkcije dešifriranja. Proces se odvija na sljedeći način. Najprije pošiljatelj pokreće generator ključeva – program koji mu dodjeljuje jedinstveni ključ za šifriranje poruke. Svaki se ključ koristi samo za jedno šifriranje. Zatim pokreće funkciju šifriranja koja kao ulazne vrijednosti ima izvornu poruku i ključ šifriranja. Ta funkcija preoblikuje poruku u skladu s ključem, a kao rezultat se dobiva šifrirana poruka. Ona se tada odašilje komunikacijskim kanalom do primatelja. Primatelj, koji mora poznavati isti onaj ključ kojim je poruka bila prvobitno šifrirana, tada pokreće funkciju dešifriranja, koja kao ulaz

⁴³ The Digital Dilemma, n. dj., str. 271 digitalne verzije.

ima šifriranu poruku i ključ šifriranja. Nakon unatražnog preoblikovanja funkcija rezultira izvornom porukom, koja tada primatelju postaje čitljiva (dijagram 3).



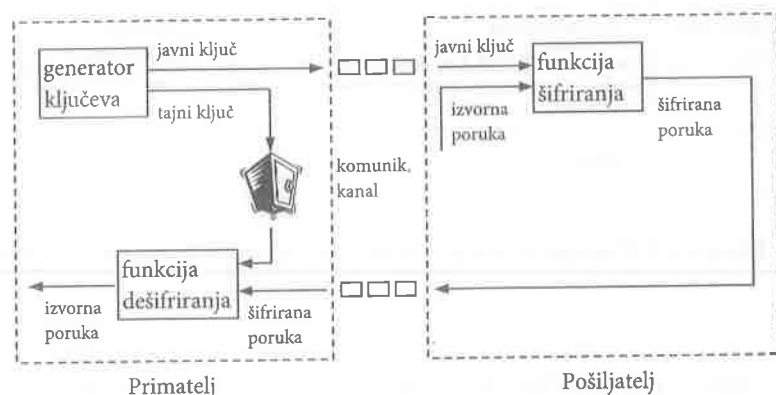
Dijagram 3: Postupak šifriranja i dešifriranja upotrebom simetričnog ključa

Tajnost ključa vrlo je bitna jer svatko tko ga posjeduje može dešifrirati poruku. Stoga je glavni problem kako dostaviti ključ određenom korisniku bez opasnosti da ga još netko prilikom prijenosa sazna. Jedan od načina dostave ključa može biti putem zaštićenog komunikacijskog kanala. Naravno, odmah se može postaviti pitanje zašto se onda cijela poruka ne prenese takvim kanalom. Odgovor se nalazi u brzini prijenosa. Zaštićeni komunikacijski kanal je obično kanal manje propusnosti. Stoga je njime lako poslati ključ šifriranja, jer je on vrlo malen. Cijela se poruka tada može nesmetano poslati javnim, nezaštićenim kanalom.

4.2.2. Šifriranje javnim ključem

Ova tehnika šifriranja koristi dvije vrste ključeva – javni ključ i privatni ključ. Ta dva ključa imaju jedinstveno svojstvo: poruka šifrirana javnim ključem može se dešifrirati jedino odgovarajućim privatnim ključem. Ovaj se sustav također sastoji od tri dijela: generatora ključeva, funkcije šifriranja i funkcije dešifriranja. Proces šifriranja i dešifriranja odvija se na sljedeći način. Najprije primatelj pokreće generator ključeva – program koji mu dodjeljuje jedinstveni par javnog i privatnog ključa. Tada primatelj javno objavi (na Internetu, u novinama i sl.) ili dostavi pošiljatelju svoj javni ključ, a privatni ključ strogo čuva. Pošiljatelj pokrene funkciju šifriranja koja kao ulazne vrijednosti ima izvornu poruku i javni

ključ primatelja. Šifrirana se poruka putem komunikacijskog kanala šalje primatelju. Tako šifriranu poruku može dešifrirati jedino vlasnik privatnog ključa koji odgovara javnom ključu kojim je poruka šifrirana. Primatelj pokreće funkciju dešifriranja koja kao ulazne vrijednosti ima šifriranu poruku i privatni ključ te kao rezultat dobiva izvornu poruku (dijagram 4).



Dijagram 4: Postupak šifriranja i dešifriranja upotrebom javnog ključa

Kad se usporede metoda simetričnog i metoda javnog ključa, može se zaključiti da je metoda javnog ključa mnogo sigurnija jer se ključ za dešifriranje nikamo ne prenosi pa je šansa za njegovo otkrivanje neznatna. Usto, ona omogućuje da, s obzirom na to da je javni ključ dostupan svima, bilo tko može poslati šifriranu poruku primatelju a da mu pritom ne treba slati i ključ za dešifriranje (možda mu i nije dostupan sigurnosni, šifrirani komunikacijski kanal potreban za dostavljanje simetričnog ključa). Logično je da je tajni ključ na neki način određen javnim ključem, ali otkrivanje tajnog samo na osnovi poznavanja javnog ključa upotrebom grube sile (engl. *brute force*), tj. načinom koji isprobava sve moguće kombinacije dok ne pogodi pravu, uz današnju bi veličinu ključeva (npr. 160-bitni ključ) i stupanj razvoja računala trajalo cijeli jedan ljudski vijek. Stoga se ove dvije tehnike mogu kombinirati. Tako se, na primjer, može upotrebljavati šifriranje simetričnim ključem, a ključ koji se dostavlja primatelju šifrira se tehnikom javnog ključa. S obzirom na to da je tehnika šifriranja javnim ključem procesorski zahtjevnija od šifriranja simetričnim ključem, ovakvom se kombinacijom dobiva na brzini uz istodobno povećanje sigurnosti.

Jedino pitanje koje ostaje neodgovoreno u kontekstu šifriranja simetričnim i javnim ključem jest pitanje utvrđivanja identiteta pošiljatelja šifrirane poruke. Naime, može li primatelj vjerovati da osoba koja tvrdi da je pošiljatelj šifrirane poruke doista to i jest. Sistem provjere i dokazivanja identiteta pošiljatelja poruke je u tom slučaju dodatni zahtjev koji mora biti posebno implementiran, tj. za to mora postojati dodatna infrastruktura. O rješavanju ovog problema i postupcima koji pritom moraju biti implementirani u sustav govori se u nekoliko narednih odjeljaka.

4.3. DIGITALNI POTPISI

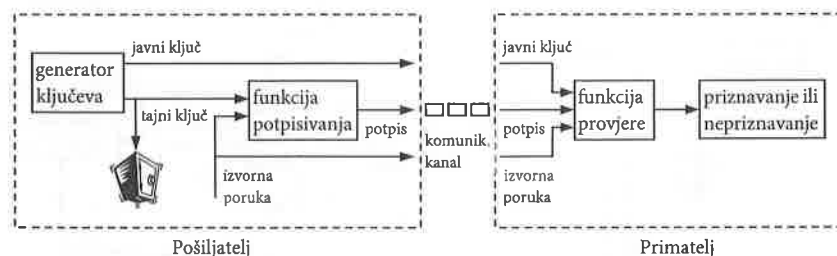
Digitalni potpis uspostavlja identitet učesnika u elektroničkoj razmjeni podataka i osigurava integritet podataka.⁴⁴ Sustav digitalnih potpisa zasniva se na tehnologiji šifriranja javnim ključem i na funkciji raspršenja (engl. *hash function*). Digitalni potpis je binarni niz koji se dodaje dokumentu kako bi se potvrdila njegova točnost i ispravnost. Binarni je niz izveden iz tajnog ključa potpisnika dokumenta.

Digitalni potpisi funkcioniraju na isti način kao i klasični potpisi na papirnatom dokumentu. Tako, u papirnatom obliku, određena osoba svojim potpisom jamči točnost i ispravnost nekog dokumenta. Potpis je jedinstveno obilježje svakog čovjeka i ovisan je o osobi koja potpisuje. Nasuprot tome, digitalne potpise možemo promatrati kao funkciju osobe koja potpisuje i potpisanog dokumenta. Razlika je u tome da kad jedna osoba potpisuje više digitalnih dokumenata svi se ti potpisi razlikuju, što nije slučaj kod potpisivanja papirnatih dokumenata. To mora biti tako zato što se digitalni potpis, kao binarni niz, šalje uz poruku, pa kad bi se isti potpis koristio za više dokumenata, svatko tko je dobio dokument s pridodanim binarnim nizom mogao bi taj niz nadodati nekom drugom dokumentu, tj. potpisati nekog drugog, i takav dokument poslati dalje.

Sustav digitalnog potpisivanja sastoji se od tri dijela: generatora ključeva, funkcije potpisivanja i funkcije provjere. Proces se odvija na sljedeći način. Osoba koja želi potpisati neki digitalni dokument najprije pokreće generator ključeva, čime dobiva jedinstveni set javnog i privatnog ključa. Zatim pokreće funkciju potpisivanja, koja kao ulazne vrijednosti ima digitalni dokument i tajni

⁴⁴ Prema: Strategija razvitka Republike Hrvatske "Hrvatska u 21. stoljeću", Informacijska i komunikacijska tehnologija, Vlada Republike Hrvatske, 16. svibnja 2002. <http://narodne-novine.nn.hr/clanci/sluzbeni/2002_09_109_1753.html>, 8. veljače 2003.

ključ, što rezultira digitalnim potpisom. Tako potpisani dokument, uz dodatak javnog ključa, pošiljatelj putem komunikacijskog kanala dostavi primatelju ili ih javno objavi. Primatelj dokumenta koji želi provjeriti autentičnost dokumenta mora pokrenuti funkciju provjere koja kao ulazne vrijednosti ima dokument, digitalni potpis i javni ključ. Funkcija provjere rezultira priznavanjem ili nepriznavanjem izvornosti digitalnog potpisa (dijagram 5).



Dijagram 5: Postupak digitalnog potpisivanja

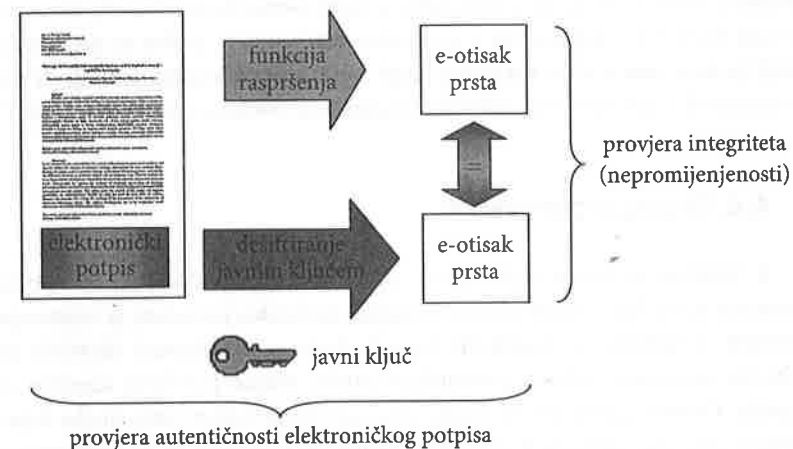
Drugi, znatno učinkovitiji oblik dodavanja i provjere digitalnog potpisa zasniva se na funkciji raspršenja (engl. *hash function*). Funkcijom raspršenja stvara se jedinstveni niz znakova za svaki dokument. Taj niz znakova znatno je manji od dokumenta na temelju kojeg je stvoren te se naziva elektroničkim otiskom prsta. Tada se e-otisk prsta šifrira upotrebom kombinacije javnog i tajnog ključa te se u obliku digitalnog potpisa pridružuje dokumentu. Princip izrade i šifriranja digitalnog potpisa korištenjem funkcije raspršenja prikazan je na sljedećoj slici.



Slika 23: Izrada digitalnog potpisa funkcijom raspršenja

Funkcija raspršenja vrlo je pogodna za korištenje jer se njome može provjeravati ne samo autentičnost poruke, već i njezin integritet, tj. nepromijenjenost. Naime, već je prije spomenuto da funkcijom raspršenja nastaje jedinstveni niz znakova za svaki dokument – e-otisk prsta. Upravo ta karakteristika, uz činjenicu da svaka, pa i najmanja promjena u dokumentu, uzrokuje stvaranje novog jedinstvenog niza, daje mogućnost provjere integriteta svakog dokumenta. Općenito gledajući, dvije su osobine funkcije raspršenja:

1. za funkciju raspršenja (H) vrijedi da je "računalno nemoguće (primjenjujući istu funkciju) pronaći bilo koje dvije poruke x i y takve da vrijedi $H(x)=H(y)$ "⁴⁵
2. iz $H(x)$ nije moguće dobiti x , tj. "računalno je nemoguće iz elektroničkog otiska prsta dobiti izvorni niz (objekt)".⁴⁶



Slika 24: Provjera autentičnosti i integriteta dokumenta korištenjem funkcije raspršenja

⁴⁵ Wikipedia, s.v. hash function, 10. rujna 2005. <http://en.wikipedia.org/wiki/Hash_function>, 13. rujna 2005.

⁴⁶ Jantz, Ronald i Giarlo, Michael J., Digital Preservation. Architecture and Technology for Trusted Digital Repositories, *D-Lib Magazine*, lipanj 2005., vol. 11, br. 6, <<http://www.dlib.org/dlib/june05/jantz/06jantz.html>>, 18. srpnja 2005.

Postupak provjere vrlo je jednostavan. Dešifriranjem e-otiska prsta provjerava se identitet potpisnika, tj. autentičnost poruke. Ponovnim generiranjem e-otiska prsta na temelju pristiglog dokumenta te njegovom usporedbom s dešifriranim e-otiskom prsta koji je pridodan dostavljenom dokumentu provjerava se integritet dokumenta. Ako su oba e-otiska prsta identična, tada je dokument nepromijenjen, a ako se razlikuju, to znači da je u međuvremenu došlo do neke promjene dokumenta. Slika 24 prikazuje postupke provjere autentičnosti i integriteta dokumenta korištenjem funkcije raspršenja.

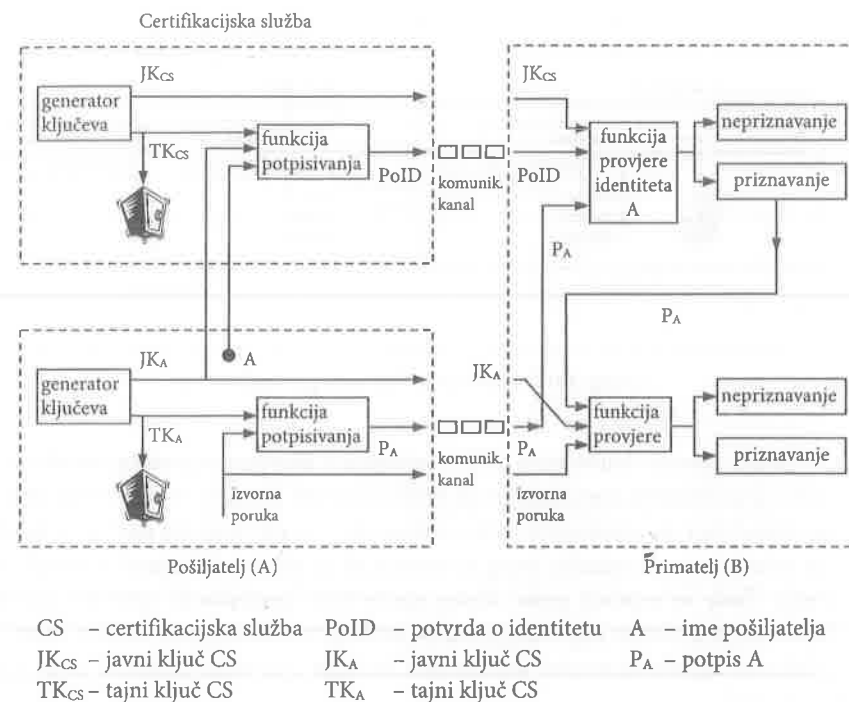
Kako se konkretno odvijaju postupci korištenja digitalnih potpisa unutar elektroničkog arhiva/repositorija Sveučilišta Rutgers u SAD-u, najbolje opisuju Jantz i Giarlo: "izračunati elektronički potpis za elektronički master (koristeći SHA1 [engl. *Secure Hash Algorithm*] koji primjenjuje 160-bitno šifriranje) i spremi ga u tehničke metapodatke objekta. Zatim izračunati potpis za cjelokupni objekt koji se sprema izvan repositorija. U okviru sustava za provjeru valjanosti, periodično se, kao pozadinski proces, ponovno izračunava elektronički otisak prsta [...] i uspoređuje s originalno izračunatim. Svaka se promjena događuje te se tada koriste neizravni (engl. *off-line*) sustavi za spremanje ili zrcalni repositoriji (engl. *mirrored repository*) kako bi se obnovio integritet objekta."⁴⁷

4.4. DIGITALNI CERTIFIKATI

S obzirom na to da, kako je već napomenuto, digitalni potpis uspostavlja identitet učesnika u elektroničkoj razmjeni podataka, potrebno je izdavanje digitalnih certifikata, tj. digitalnih potvrda kojima se dokazuje identitet (engl. *identity certificate*) kako bi primatelj podataka mogao provjeriti identitet pošilatelja. Certifikacijska služba (engl. *CA – certifying authorities*), služba koja dodjeljuje digitalne potvrde o identitetu, mora biti ovlaštena za tu djelatnost kako bi digitalne potvrde imale potreban kredibilitet.

Problem digitalnih potvrda o identitetu svodi se zapravo na razvoj infrastrukture za upravljanje javnim ključevima. Certifikacijska služba izdaje potvrdu o identitetu te je potpisuje svojim ključem za potpisivanje. Potvrda o identitetu neke osobe je, dakle, digitalno potpisan binarni zapis koji sadržava javni ključ i ime vlasnika, a može sadržavati i još neke podatke, kao npr. "rok upotrebe", tj. informaciju o tome u kojem je vremenskom periodu javni ključ pravovaljan. Stoga, kada primatelj posjeduje javni ključ certifikacijske službe (JK_{CS}), on tada

može na temelju povjerenja u certifikacijsku službu vjerovati u ispravnost potvrde o identitetu koju je ona izdala, a poznavajući javni ključ osobe koja se navodi u toj potvrdi, vjerovati u činjenicu da mu je upravo ta osoba poslala podatke koje je primio.⁴⁸



Dijagram 6: Postupak uporabe digitalnih certifikata

Na primjer, primatelj B prima dokument koji je digitalno potpisala osoba A zajedno s potvrdom o identitetu u kojoj se navodi ime A i pripadajući javni ključ (JK_A). Primatelj se tada služi javnim ključem certifikacijske službe kako bi provjerio ispravnost njezinog digitalnog potpisa na potvrdi o identitetu. Ako je ispravnost potvrđena onda primatelj može s povjerenjem iskoristiti javni ključ osobe A, kako bi utvrdio je li osoba A potpisala primljeni dokument. Primatelje-

⁴⁷ Jantz i Giarlo, Digital Preservation, n. dj.

⁴⁸ The Digital Dilemma, n. dj., str. 278 digitalne verzije.

vo povjerenje u certifikacijsku službu zapravo znači njegovo povjerenje u ispravnost njezina javnog ključa te povjerenje da je certifikacijska služba zaista osobi A pridijelila ispravan javni ključ.⁴⁹ Dijagram 6 prikazuje tijek dokumenta od pošiljatelja do primatelja uz korištenje digitalnih certifikata.

Segment verificiranog upravljanja javnim ključevima, tj. izdavanjem potvrda da iza određenog javnog ključa i imena stvarno stoji osoba koja tvrdi da stoji, vrlo je bitan kod primateljeve provjere je li mu baš navedena institucija dostavila određeni dokument (provjera putem JK_{CS}) te može li vjerovati da je dokument stvarno identičan originalu, tj. da ga netko neovlašten nije promijenio (provjera putem JK_A). Ovaj je segment također bitan kod potpisivanja digitalnih ugovora kad se dvije ugovorne strane mogu nalaziti u dva različita grada ili, pak, na dvjema različitim stranama svijeta, a moraju imati povjerenja da ona druga potpisna strana to stvarno i jest.

4.5. DIGITALNI VODENI ŽIGOVI

Digitalni vodeni žig (ili digitalna vodena oznaka – engl. *watermark*) je signal koji je dodan elektroničkom gradivu s namjerom da prenese određenu, malu količinu informacija. Detekcijom prisutnosti ili neprisutnosti digitalnog vodenog žiga može se dokazati autentičnost ili neautentičnost digitalnoga gradiva. Oni se uglavnom koriste za obilježavanje slikovnog, zvučnog ili video gradiva. Svako od tih gradiva može se smatrati dokumentom pa će se i u daljnjim opisima koristiti taj termin. Postoje različite primjene digitalnih vodenih žigova, među kojima su najčešće sljedeće:

- dokazivanje vlasništva nad nekim sadržajem,
- umetanje podataka o primatelju (engl. *fingerprinting*) kako bi se moglo ustanoviti odakle je potekla eventualna nelegalna kopija,
- provjera autentičnosti i integriteta,
- opisivanje sadržaja (engl. *content labeling*),
- kontrola korištenja te
- zaštita sadržaja.⁵⁰

⁴⁹ The Digital Dilemma, n. dj., str. 278-279 digitalne verzije.

⁵⁰ The Digital Dilemma, n. dj., str. 280-281 digitalne verzije.

Digitalni vodeni žigovi mogu korisnicima biti vidljivi ili nevidljivi, a po oblikovanju mogu biti krhki (engl. *fragile*) ili robusni (engl. *robust*). Vidljivi digitalni vodeni žigovi pojavljuju se u obliku logotipa ili neke poruke na vidljivom ili u čujnom području digitalnoga gradiva, a korisnicima služe kao informacija o vlasništvu ili dopuštenom korištenju. Neke institucije koriste ovu vrstu žigova za slobodnu distribuciju materijala nedovoljne kvalitete za profesionalnu upotrebu, a naplaćuju gradivo visoke kvalitete. Nevidljivi žigovi se, pak, mogu koristiti kao dokaz nelegalnog korištenja digitalnih dokumenata. Krhki digitalni vodeni žigovi nisu postojani pri obradi digitalnih dokumenata pa se koriste kako bi se detektirale eventualne izmjene dokumenata. Tako prilikom preuzimanja dijela originalnog dokumenta ili njegove obrade više nije moguće dokazati njegovo podrijetlo. Nasuprot tome, robusni se žigovi provlače kroz cijeli digitalni zapis te se njegovi dijelovi mogu detektirati i nakon što je izvorni dokument obrađen, smanjen ili povećan, izrezan i sl. te uključen u neki novi dokument.

Sustav digitalnih vodenih žigova sastoji se od dva dijela: funkcije umetanja žiga i funkcije detekcije žiga. Proces se odvija na sljedeći način. Pošiljatelj digitalnog dokumenta pokreće funkciju umetanja žiga koja kao ulazne vrijednosti ima izvorni digitalni dokument, digitalni vodeni žig i korisnički ključ, a kao rezultat digitalni dokument s umetnutim digitalnim vodenim žigom. Nakon toga pošiljatelj, zajedno s ključem, prosljeđuje takav dokument primatelju koji, ako to želi, tj. ako žig nije u



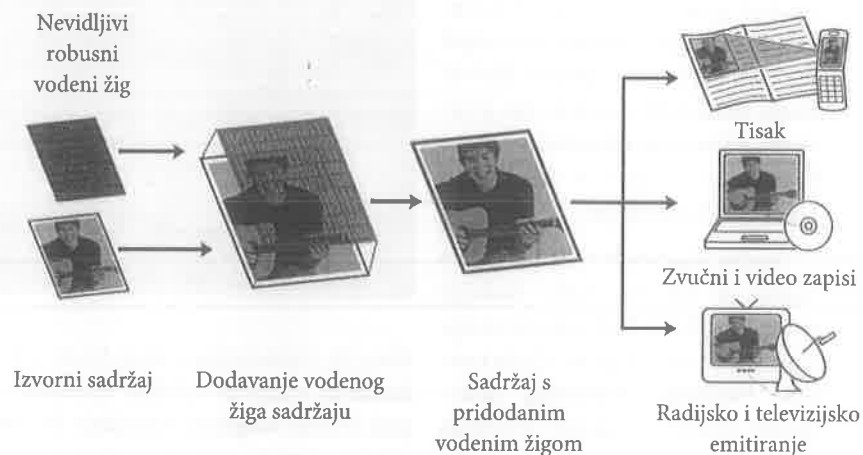
Slika 25: Digitalizirani detalj freske s pridodanim vidljivim krhkim vodenim žigom (Beato Angelico, Klanjanje Sv. Dominika raspelu (tal. *San Domenico in Adorazione del Crocifisso*), 15. st., samostan sv. Marka u Firenci, Italija)⁵¹

⁵¹ Izvor: http://www.memorynet.org/larger/?image=fm013.jpg&collection=florence&code=en_fm013

vidljivom ili čujnom dijelu dokumenta, može detektirati žig. Da bi ga detektirao, primatelj mora pokrenuti funkciju detekcije žiga. Ona kao ulazne vrijednosti ima korisnički ključ i primljeni dokument, a kao rezultat daje digitalni vodeni žig, čime potvrđuje ili opovrgava njegovu autentičnost (dijagram 7).

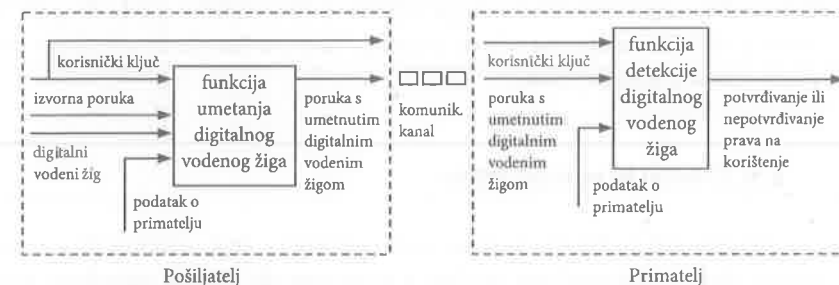


Dijagram 7: Sustav digitalnih vodenih žigova

Slika 26: Postupak dodavanja vodenog žiga⁵²

⁵² Izvor: *Digital Watermarking for the Digital Age: Empowering Consumers and Protecting Copyrights*, DIGIMARC, 2005., <https://www.digimarc.com/resources/docs/dmrc_content_id.pdf>, 5. siječnja 2009.

Kod procesa umetanja podataka o primatelju u digitalni vodeni žig postoji dodatna ulazna vrijednost, tj. šifra korisnika, kojom se na jedinstven način određuje primatelj, tj. korisnik kojem je dopuštena upotreba određenog digitalnog dokumenta. U tom se slučaju kao izlazna vrijednost funkcije umetanja žiga dobiva digitalni dokument s umetnutim digitalnim vodenim žigom i podatkom o primatelju. Funkcija detekcije žiga tada kao dodatnu vrijednost ima šifru korisnika, a rezultira potvrđivanjem ili opovrgavanjem prava na korištenje određenom primatelju (dijagram 8). Proces umetanja podataka o primatelju svojstven je za robusne digitalne vodene žigove.



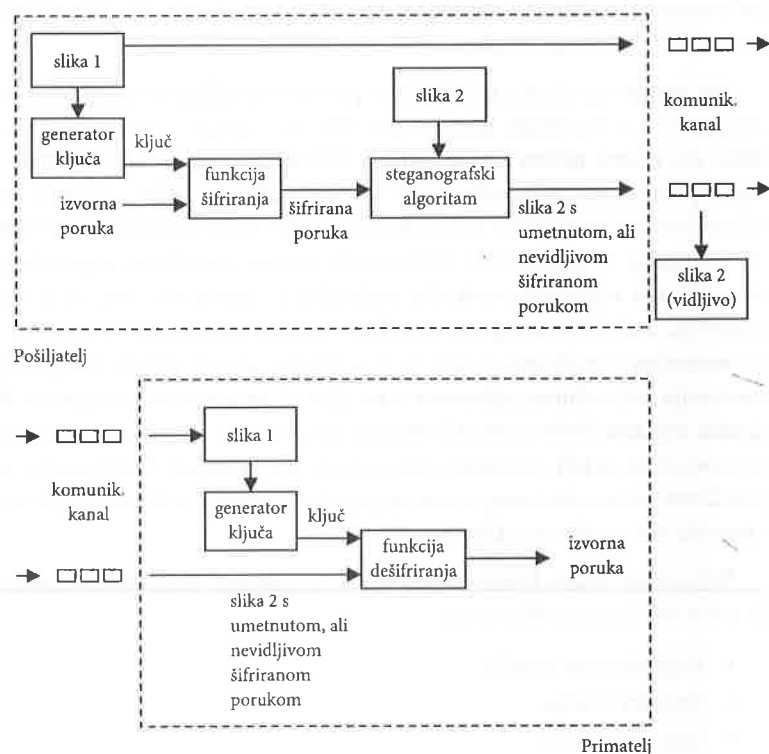
Dijagram 8: Sustav digitalnih vodenih žigova s umetanjem podataka o primatelju

Ako se u nekom budućem trenutku detektira postojanje elektroničkoga gradiva kod osoba kojima ono nije isporučeno, onda se ovim postupkom može utvrditi koji je korisnik nelegalno prosljedio to elektroničko gradivo trećim osobama na korištenje i tako prekršio vlasnička prava institucije u čijem se vlasništvu ono nalazi. Pritom korisnik može, ali i ne mora biti svjestan da je u gradivo umetnut podatak o primatelju.

4.5.1. Steganografija

Termin steganografija dolazi od grčkog *steganos*, što znači "prikriven", i *graphia* što znači "pisanje". Za razliku od šifriranja, gdje poruka nije čitljiva bez poznavanja ključa za dešifriranje, kod steganografije je riječ o skrivanju poruke na način da ona ne bude primijećena. U elektroničkom obliku, upotrebom po-

sebnih programa, neku tekstualnu, slikovnu, zvučnu, video poruku ili neku šifriranu poruku može se skriti u neku drugu poruku iste ili različite vrste. Za razliku od šifriranja, kod primjene steganografskih metoda, u slučaju presretanja poruke ona nije odmah vidljiva. Naime, u slučaju presretanja šifrirane poruke odmah je jasno da je riječ o poruci i da je ona šifrirana. Nasuprot tome, u slučaju presretanja poruke zaštićene steganografskim metodama, presretač uopće ne mora biti svjestan činjenice da je presreo poruku, na primjer neku sliku koja u sebi nosi skrivenu poruku. Cilj steganografije nije u tome da se poruka učini nečitljivom, kao što je to slučaj sa šifriranjem, već da se učini neprimjetnom. Metode šifriranja mogu se kombinirati sa steganografskim metodama na način da se šifrirana poruka steganografskim metodama umetne u neku drugu poruku.



Dijagram 9: Kombinacija metode šifriranja simetričnim ključem sa steganografskom metodom

Steganografske metode kombinirane s metodama šifriranja mogu postizati još bolje rezultate ako se, na primjer, koristi šifriranje simetričnim ključem, pri čemu bilo koja slika može poslužiti kao ključ (engl. *Image Based Encryption System*). Prednost ovih sustava je u tome da se ključ (u ovom slučaju neka slika) sastoji od velikog broja piksela (na primjer slika veličine 800 x 600 piksela ukupno ima 480.000 piksela), a svaki je piksel izražen u točno određenoj kombinaciji nijansi crvene, zelene i plave boje (RGB, 3 x 8 bita za svaku boju). Rezultirajući je ključ, dakle, $800 \times 600 \times 3 \times 8 = 11.520.000$ bitni (11,5 Mbitni) ključ.⁵³ Dijagram 9 prikazuje korake prenošenja poruke prilikom kombiniranja metode šifriranja simetričnim ključem sa steganografskom metodom.

Sasvim konkretna primjena steganografske metode nalazi se u postupcima nevidljivih vodenih žigova pri čemu pridodane poruke mogu nositi informacije o vlasničkim pravima, podatke o primatelju i slično.

4.6. Šifrirane omotnice

Šifrirane omotnice (engl. *cryptographic envelopes, cryptolopes*) nastale su kao sustav pojačane zaštite prijenosa i korištenja digitalnoga gradiva. Ovakvi sustavi koriste šifrirani digitalni spremnik koji sadrži izvornu poruku, izjavu o pravima pristupa i korištenja poruke, a također može sadržavati digitalni vodeni žig i digitalni vodeni žig s umetnutim podacima o primatelju. Primateljevo korištenje takve poruke mora odobriti softver koji provjerava pravo korištenja. Takav softver, spajanjem na poslužitelj za obračun korištenja, može istovremeno naplatiti korištenje gradiva. On bi također trebao provoditi dešifriranje samo onog dijela poruke koji se trenutno pregledava, čime se znatno povećava zaštita. Sustavom šifriranih omotnica omogućava se laka dostupnost gradiva bez ugrožavanja financijskih interesa institucije koja je vlasnik gradiva zbog mogućih izrada i distribucije nelegalnih kopija.

Danas su sustavi zaštite digitalnoga gradiva dovoljno uznapredovali da se mogu koristiti i za potrebe elektroničkog poslovanja, tj. za siguran prijenos do korisnika koji je spreman platiti pregledavanje ili korištenje gradiva. Naravno, i kod već razvijenih tehnika uvijek postoje mjesta gdje su potrebna daljnja istraživanja i razvoj, kao što je to, na primjer, kod šifriranih omotnica, gdje se diskutira o spajanju softverskih i hardverskih tehnika zaštite.

⁵³ Više informacija o ovoj metodi može se pronaći na: <http://www.eleversolutions.com/more.aspx>

5. POHRANA I PRIJENOS DIGITALNOGA GRADIVA

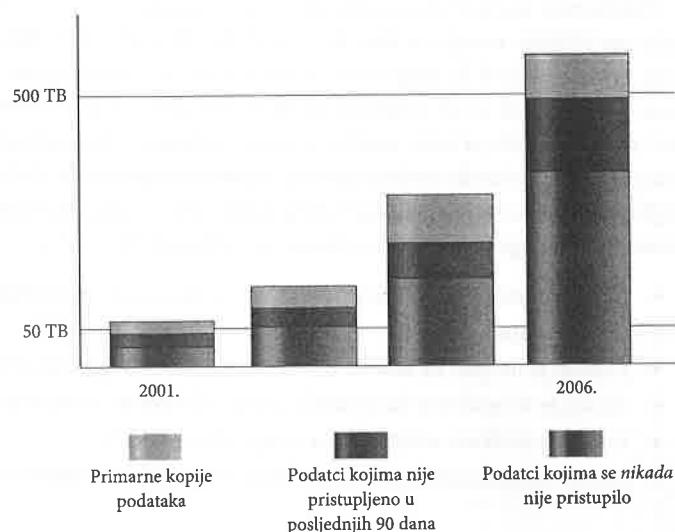
Problematicu pohrane i prijenosa digitalnoga gradiva treba razmatrati zajedno, jer bi cilj digitalizacije gradiva u institucijama kao što su arhivi, knjižnice i muzeji ili druge komercijalne institucije trebao biti postavljen tako da digitalizirano gradivo bude dostupno. Stoga je pitanje pohrane ujedno i pitanje prijenosa pohranjenog gradiva do korisnika jer sustavi za pohranu, između ostalog, uvjetuju način i brzinu pristupa gradivu.

No, problem pohrane ne bi trebalo promatrati isključivo u kontekstu digitalizacije, već i u kontekstu gradiva koje izvorno nastaje u elektroničkom obliku. Tako su, prema nekim istraživanjima IDC-a, kompanije koje su 2001. godine imale pohranjeno nešto više od 50 TB podataka, već 2006. godine i više nego udesetorostručile količinu tih podataka te tako imale pohranjeno približno 540 TB podataka. Ti podatci već sami za sebe govore o kolikom je povećanju riječ, ali mnogo je značajnija struktura podataka, tj. odnos onih koji su primarni ili najnoviji, onih kojima nije pristupljeno unutar posljednjih 90 dana u vrijeme provedenoga istraživanja i onih koji se čuvaju, ali im nikada nije pristupljeno. Ilustracija ovih odnosa prikazana je na slici 27. Iz nje se vidi da gotovo 80% podataka nije zatraženo unazad 3 mjeseca od trenutka njihove pohrane, dok 60% podataka koji su bili pohranjeni nisu nikada bili zatraženi. Dakle, važno je dobro procijeniti odnos uloženog novca za pojedino rješenje, dobivene funkcionalnosti i potrebe baš za takvim rješenjem.

Kriteriji za odabir kvalitetnog sustava za pohranu digitalnoga gradiva na duži rok u bilo kojoj instituciji su:

- dugovječnost medija,
- trajnost medija,
- visoki kapacitet,
- mala cijena,
- široka prihvaćenost te

- sustav mora biti izravan (engl. *on-line*) ili poluizravan (engl. *near-line*).⁵⁴



Slika 27: Odnos strukture podataka u kompanijama prema istraživanju IDC-a

Razlike između pojedinih institucija odnose se jedino na tempo digitalizacije, količinu digitalnoga gradiva kojom se barata te financijske mogućnosti.

Mediji na koje se pohranjuje digitalno gradivo mogu se podijeliti u nekoliko skupina prema načinu zapisa i vrsti medija:

- izmjenjivi magnetski diskovi (danas uglavnom već zastarjeli) – disketa, Iomega Zip i Jazz diskovi, SuperDisk i drugi,
- izmjenjivi optički diskovi – CD-ROM, DVD, BD – Blu-Ray, UDO⁵⁵ diskovi, HD-DVD⁵⁶ i razni drugi,

⁵⁴ Bell, Robert i Andrew Waugh, Digital Storage Media for VERS, u: *Standard for the Management of Electronic Records in the Victorian Government*, Specification 3 VERS Standard Electronic Record Format, prilog 4, Public Record Office Victoria, State of Victoria, Australija, travanj 2000., <<http://www.prov.vic.gov.au/vers/standard/version2.htm>>, 28. listopada 2008.

⁵⁵ Engl. Ultra Density Optical.

⁵⁶ HD-DVD medij se natjecao s Blu-Ray diskom za standard nove generacije. Potpisivanje ugovora za distribuciju filmskih materijala velikih filmskih studija s proizvođačima okupljenima oko Blu-Ray standarda u veljači 2008. godine naveo je grupaciju proizvođača okupljenih oko HD-

- magneto-optički diskovi – 130mm (5,25”), 90mm (1,44”)
- poluvodički mediji – razne USM prijenosne memorije, SmartMedia, CompactFlash, Memory Stick memorijske kartice i drugi,
- magnetske trake – DLT-S4, LTO-4, SAIT-2, T10000,
- optičke trake – TRAAMS, LOTS te
- urezivanje ionskom zrakom – HD-Rosetta.

Prema količini digitaliziranoga gradiva institucije se mogu podijeliti na:

- velike – stvaraju više od 1,5 Tb (terabajta) na godinu uz dvije dodatne sigurnosne kopije,
- srednje – stvaraju između 10 Gb i 1,5 Tb na godinu uz dvije dodatne sigurnosne kopije te
- male – stvaraju do 10 Gb na godinu⁵⁷ uz dvije dodatne sigurnosne kopije.

Ova podjela služi tek kao orijentacijska, jer napretkom tehnologije, povećanjem njezine brzine i propusnosti komunikacijske infrastrukture te padom cijene i istovremenim porastom kapaciteta medija za pohranu institucije mogu vrlo lako stvarati velike količine digitaliziranoga gradiva na godišnjoj razini. Ipak, u ovoj bi podjeli trebalo o digitaliziranom gradivu razmišljati kao o digitaliziranom i finalno obrađenom gradivu, a takvog će gradiva zasigurno biti manje.

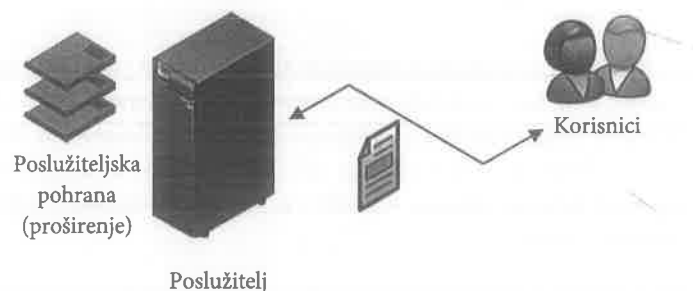
Pohrana i osiguranje dostupnosti digitalne građe putem mreže može se promatrati kroz šest osnovnih vrsta sustava. To su 1. izravni (engl. *on-line*), 2. poluizravni (engl. *near-line*), 3. hijerarhijski (engl. *Hierarchical Storage Management* – HSM), 4. neizravni (engl. *off-line*), 5. sustavi za mrežnu pohranu (engl. *Network Attached Storage* – NAS) i 6. mreža za pohranu (engl. *Storage Area Network* – SAN).

DVD standarda da odustanu od daljnje utrke. No, već i prije toga, baš kao što se to već dogodilo u vrijeme sličnog natjecanja između DVD+ i DVD- standarda, na tržištu su se pojavili uređaji koji su mogli čitati oba standarda (npr. LG Super Multi Blue Drive).

⁵⁷ 10 Gb na godinu možda se u današnjoj perspektivi čini malo, no ako neka institucija, primjerice, digitalizira crno-bijele dokumente A4 veličine, pri čemu je svaka slika takve stranice velika između 50 i 75 Kb, onda iz toga proizlazi da u 10 Gb stane između 150 i 200.000 dokumenata, što svakako nije zanemariva količina dokumenata (uz 250 radnih dana godišnje, to bi značilo produkciju od 600 do 800 skeniranih dokumenata dnevno).

5.1. IZRAVNI SUSTAVI ZA POHRANU I PRIJENOS PODATAKA

Izravni sustavi, kao što im i samo ime govori, omogućuju izravan pristup podacima (engl. *Direct Access Storage – DAS*). Za ovakve se sustave koriste čvrsti diskovi na kojima je gradivo pohranjeno. Oni funkcioniraju kao proširenje postojećem poslužitelju i, za razliku od sustava za mrežnu pohranu (engl. *Network Attached Storage – NAS*), nisu umreženi kao samostalna mrežna jedinica. Obično nije riječ o jednom disku, jer on kapacitetom ne bi mogao zadovoljiti, već o polju diskova (engl. *Redundant Array of Inexpensive Drives – RAID*) koji čine jednu logičku cjelinu i s korisničke se strane doimaju kao jedan disk. Performanse polja diskova bolje su od performansi svakog diska zasebno. Sustav, naime, digitalni zapis dijeli u blokove i svaki blok zapisuje na drugi disk, čime smanjuje vrijeme koje je potrebno za zapisivanje. Prilikom zahtjeva za čitanjem nekog zapisa, više diskova istovremeno čita manju količinu podataka, čime se povećava i ukupna brzina čitanja. Preporuča se korištenje onog sustava polja diskova koje radi sigurnosnu kopiju zapisa, tako da u slučaju kvara na nekom od diskova podatci ne budu izgubljeni. Iako su cijene diskova svakim danom sve niže, a njihov kapacitet sve veći, izravni sustavi za pohranu i prijenos podataka dosta su skupi te se koriste kada je potrebno osigurati brzi pristup gradivu. Sljedeća slika prikazuje organizaciju izravnih sustava.

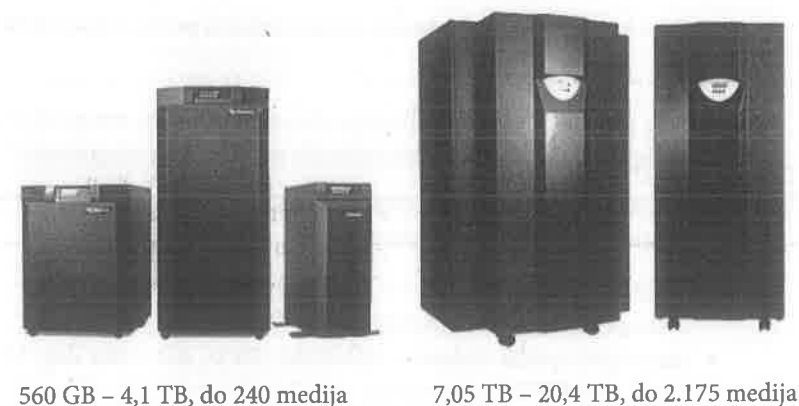


Slika 28: Izravni sustavi za pohranu i prijenos podataka

5.2. POLUIZRAVNI SUSTAVI ZA POHRANU I PRIJENOS PODATAKA

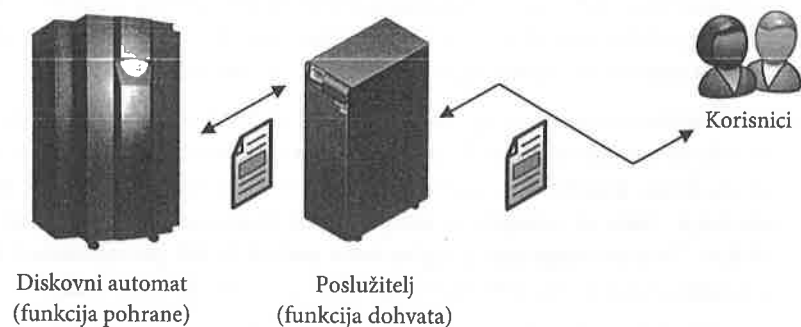
Poluizravni sustavi imaju odličan odnos kapaciteta i cijene. Oni čuvaju podatke na jeftinim medijima kao što su CD-ROM, DVD, BD (Blu-Ray) diskovi, drugi optički diskovi ili magnetske trake, nauštrb brzine pristupa. Svaki poluizravni sustav sastoji se od smještajnog dijela, nekoliko čitača i robotske ruke koja služi za automatski prihvat medija i njegov prijenos od smještajnog mjesta do čitača i natrag. Robotski sustavi manjeg kapaciteta nazivaju se diskovni automati (engl. *jukebox*), a oni većeg smještajnog kapaciteta – silosi. Korištenje robotskih sustava ima mnoge prednosti u odnosu na ručni prihvat medija:

- manji su troškovi korištenja jer se ne mora plaćati poslužiteljsko osoblje,
- brži je pristup gradivu,
- pristup je moguć 24 sata na dan, 7 dana u tjednu (pristup 24-7),
- manja je mogućnost da se medij zagubi ili vrati na krivo mjesto,
- manji su troškovi osvježavanja i migracije zapisa te
- postoji mogućnost rada u hijerarhijskom sustavu za pohranu.⁵⁸



Slika 29: Diskovni automati

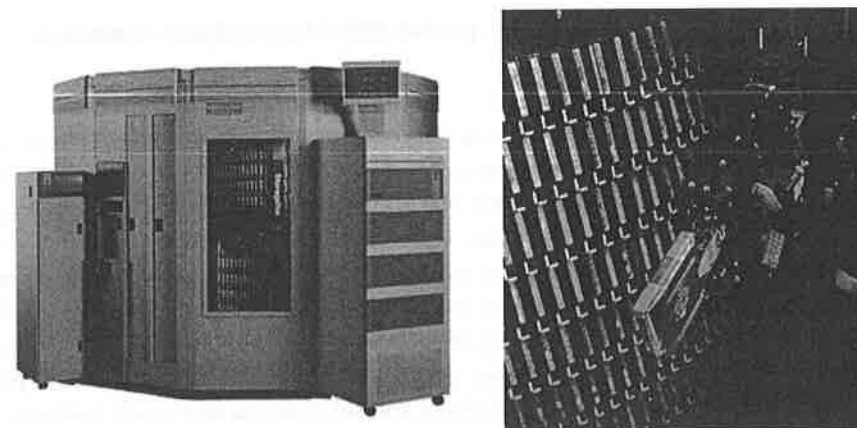
⁵⁸ Bell i Waugh, Digital Storage Media, n. dj.



Slika 30: Poluizravni sustavi za pohranu i prijenos podataka

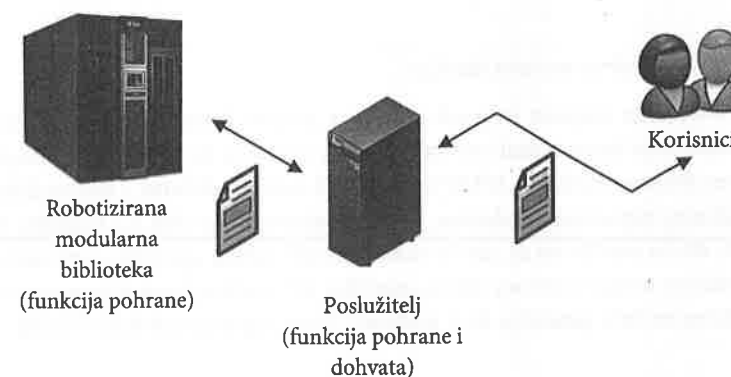
5.3. HIJERARHIJSKI SUSTAV ZA POHRANU I PRIJENOS PODATAKA

Ovaj je sustav kombinacija izravnog i poluizravnog sustava. Hijerarhijski sustav sastoji se od čvrstog diska ili polja diskova, koji su skuplji medij, ali omogućavaju veliku brzinu pristupa, i magnetskih traka kao jeftinijeg medija, ali sporijeg pristupa. Tehnika hijerarhijskog sustava zasniva se na statistici korištenja digitalnih zapisa. Novi se zapisi uglavnom najčešće traže pa se stoga oni najprije smještaju na čvrsti disk. Nakon određenog vremena, kada se disk skoro popuni, oni zapisi koji su se najrjeđe koristili sele se na magnetske trake, magnetno-optičke ili neke druge izmjenjive medije. Kada netko zatraži ono gradivo koje se nalazi na izmjenjivom mediju, sustav ga tada s njega iščita. S korisnikove strane ovaj je postupak potpuno transparentan te on ne zna koje se gradivo gdje nalazi. Cijela digitalna kolekcija doima se kao jedna cjelina te se ovakvi sustavi često opisuju kao virtualni diskovi. Jedina razlika vezana uz fizički smještaj gradiva koju će korisnik osjetiti jest ta da će nešto dulje morati pričekati kad postavi zahtjev za gradivom koje se nalazi na izmjenjivome mediju.



Do 144.000 traka, svaka traka do 200 GB,
ukupni kapacitet 28.800 PB (petabajta).

Slika 31: Robotizirana modularna biblioteka magnetskih traka kao dio hijerarhijskog sustava za pohranu



Slika 32: Hijerarhijski sustav za pohranu i prijenos podataka

5.3.1. Usporedba izravnih, poluizravnih i hijerarhijskih sustava za pohranu

Ovo poglavlje donosi usporedbu karakteristika triju "konkurentnih" rješenja za pohranu velikih količina podataka – izravni sustav s RAID poljem diskova, poluizravni sustav s magnetskim trakama i hijerarhijski sustav s UDO (engl. *Ultra Density Optical*) diskovima. Sustav s UDO diskovima izabran je za ovu analizu upravo zbog procijenjene trajnosti medija od 50 godina,⁵⁹ iako bi se umjesto tih diskova u hijerarhijskim sustavima mogli koristiti i CD, DVD ili BD diskovi. Na početku ipak treba napomenuti kako i jest i nije riječ o konkurentnim rješenjima, jer će odabir izravnog, poluizravnog ili hijerarhijskog sustava za pohranu prije svega ovisiti o predviđenoj količini podataka koji se žele pohraniti i cijeni koju je neka institucija spremna platiti, ali i mnogim drugim faktorima. Stoga ova usporedba ima za cilj analizirati karakteristike ovih rješenja i pokazati u kojim situacijama bi koje rješenje moglo biti najpogodnije.

Hijerarhijski sustavi, po svojoj definiciji, čine kombinaciju izravnih i poluizravnih sustava, no ovdje će se analizirati samo onaj njihov element pohrane na izmjenjivim medijima, jer su karakteristike elementa zaduženog za izravnu pohranu identične onima navedenima za izravne sustave. Navedena tri rješenja usporedit će se prema osam kriterija. To su: 1. predviđena trajnost medija, 2. kapacitet sustava, 3. osiguranje autentičnosti, 4. mogućnost uništenja podataka, 5. brzina pristupa, 6. izrada sigurnosnih kopija, 7. cijena i 8. potrošnja energije.

1. Predviđena trajnost medija

Prosječna trajnost izravnih sustava s poljem čvrstih diskova izražava se u prosječnom broju radnih sati između dva kvara, tj. do prvog kvara (engl. *Mean Time Between Failure* – *MTBF*). Kod nekih novijih diskova, s perpendikularnim načinom zapisivanja podataka, ta se vrijednost kreće oko 1,2 milijuna radnih⁶⁰ sati, što bi značilo da se, ako je sustav u 24/7 načinu upotrebe, prvi potencijalni problemi mogu očekivati nakon otprilike 137 godina (unatoč tome proizvođači uglavnom daju garanciju do 5 godina)! Treba napomenuti kako bi sustavi s po-

⁵⁹ Pritom se mora uzeti u obzir da iako zapisi teoretski mogu biti fizički čitljivi tijekom 50 godina i nakon toga, to ne garantira da će ih se moći pročitati zbog promjena koje u međuvremenu nastaju kako u području razvoja hardvera za podršku takvom sustavu, tako i, možda značajnije, softvera kojim bi se spremljeni podaci trebali pročitati (zastarijevanje formata zapisa).

⁶⁰ Na primjer, Seagate Barracuda ES.2 SATA, 750 GB, 3 GB/s.

ljem diskova trebali imati redundanciju tako da kvar jednog od diskova ne dovede do gubitka podataka. Usto, kod izravnih sustava trebala bi biti omogućena izmjena diskova za vrijeme rada sustava (engl. *hot swap*).

Poluizravni sustavi s magnetskim trakama su takozvani kontakti sustavi jer za vrijeme čitanja dolazi do fizičkog kontakta medija i čitača. Takav način čitanja podataka dugotrajnom upotrebom dovodi do trošenja materijala i njegova oštećenja. Trake su osjetljive na temperaturne promjene i promjene vlažnosti u okolini. Trajnost magnetskog zapisa treba periodički biti provjeravana i po potrebi zapisi kopirani na novu traku (tzv. postupak osvježavanja medija).

UDO diskovi, kao dio hijerarhijskog sustava, baš kao i čvrsti diskovi, mediji su s beskontaktnim načinom čitanja podataka što sasvim sigurno produljuje njihovo procijenjeno vrijeme korištenja koje iznosi 50 godina. Proizvođači garantiraju toliku dugotrajnost čitljivosti spremljenih podataka na temelju vrlo visoke kvalitete materijala iskorištenih za izradu medija, ali i načina zapisivanja – zapisivanje faznom izmjenom (engl. *Phase Change Recording*). Ovakav je sustav neosjetljiv na prosječne promjene temperature i vlage u okolini.

2. Kapacitet sustava

Kad je u pitanju kapacitet nekog sustava, onda tu nema neke dvojbe. Sustavi koji koriste magnetske trake za pohranu podataka uvjerljivo su sustavi s najvećim mogućim kapacitetom. Izravni sustavi koriste se većim brojem brzih, čvrstih diskova i dodavanjem novih diskova ili novih elemenata sustava s prostorom za nove diskove može im se proširivati kapacitet. Povećanje kapaciteta sustava, zbog definicije izravnog sustava, ne može se postići dodavanjem medija koji nisu aktivni (engl. *off-line*). Snaga poluizravnih i hijerarhijskih sustava leži upravo u proširivosti njihova kapaciteta dodatnim medijima (trakama ili diskovima) te u korištenju robotiziranog sustava za dohvrat tih medija, njihovo prenošenje do čitača i vraćanje na mjesto.

3. Osiguranje autentičnosti

Osiguranje autentičnosti u ovoj će se analizi ograničiti na (ne)mogućnost promjene zapisa koji je jednom spremljen na medij za pohranu, ne ulazeći u pitanje stvarne autentičnosti sadržaja tih zapisa. Kod ovog kriterija po karakteristikama prednjače UDO diskovi (a mogli bi i bilo koji CD ili DVD diskovi koji bi se u hijerarhijskim sustavima također mogli koristiti) u odnosu na čvrsti disk

i magnetsku traku. Naime, jedino oni imaju mogućnost WORM (engl. *Write Once Read Many*) načina zapisivanja. Dakle, podatci se zapišu samo jednom i nakon toga ih se samo može čitati, ali ne i mijenjati, tj. brisati, što je najbolja dokazna vrijednost nepromijenjenosti izvorno pohranjenih zapisa. Čvrsti disk je po svojim osnovnim karakteristikama magnetski medij i samim time izbrisiv i podložan utjecaju snažnih vanjskih magnetskih polja baš kao i magnetska traka. Kod izravnih sustava, stoga, postoji softverska emulacija WORM načina zapisivanja, dok kod poluizravnih sustava s magnetskim trakama ona može biti implementirana u uređaj za čitanje i zapisivanje. Tada je riječ o hardverskoj emulaciji WORM načina zapisivanja. Važno je napomenuti da je u oba slučaja riječ o emulaciji zaštitnih mehanizama, a ne o njihovoj fizičkoj implementaciji. Softverskim uklanjanjem takve zaštite moguće je mijenjati ili brisati zapisane podatke. Treba napomenuti da u slučaju korištenja izbrisivih UDO medija (engl. *Re-writable* – RW) spomenuta funkcionalnost nije uspostavljena.

4. Mogućnost uništenja podataka

Mogućnost uništenja podataka važna je u trenutku kad prođe obvezni rok čuvanja dokumentacije. Rok čuvanja ovisi prije svega o vrsti dokumenata i može biti od nekoliko godina do trajno. Stoga je važno da kad jednom dođe trenutak za izlučivanje, dokumentacija može biti uništena tako da joj se više ne može pristupiti. To je važno i s aspekta sigurnosti. Kod izravnih sustava, tj. čvrstih diskova, brisanje uz naknadnu nemogućnost povratka izbrisanih podataka provodi se višestrukim brisanjem fizičke pozicije na disku ili njezino višestruko prepisivanje novim podacima (bilo stvarnima ili onima koji služe samo za postupak brisanja). Kod poluizravnih sustava s magnetskim trakama, zbog slijednog načina zapisivanja, nije moguće brisanje pojedinih podataka, već samo cijele trake ili svih podataka od nekog mjesta do kraja trake. Hijerarhijski sustavi s UDO diskovima, ili nekim drugim optičkim medijima, omogućuju potpuno uništenje potrebnih podataka bilo uništenjem samog medija (u slučaju WORM načina zapisivanja) ili brisanjem podataka (u slučaju korištenja izbrisivih (RW) medija). Pritom je, naravno, važno uzeti u obzir preporuku o prethodnoj organizaciji sadržaja na način da se na jednom fizičkom disku nalazi dokumentacija s istim rokom čuvanja kako ne bi došlo do situacije da se na istom mediju nađu podatci koje je potrebno izlučiti i oni koje je potrebno još čuvati.

5. Brzina pristupa

Kao što su magnetske trake neprikosnovene kad je u pitanju kapacitet, tako su i čvrsti diskovi rješenje za najbrži pristup pohranjenim podacima⁶¹. Stoga su i izravni sustavi koji se na njima temelje rješenja za najbržu pohranu podataka. Nasuprot njima, magnetske su trake, zbog vremena potrebnog za premotavanje trake do željene pozicije i slijednog načina čitanja podataka, najsporije rješenje kod kojeg se mogu dogoditi i duža čekanja do trenutka dobivanja traženih podataka zbog nagomilavanja prethodnih upita za podacima. Zbog toga su poluizravni sustavi najsporije rješenje. Hijerarhijski sustavi su u sredini po pitanju brzine, jer je pristup podacima na UDO diskovima nešto sporiji od brzine pristupa podacima na čvrstim diskovima, no ipak daleko brži od magnetskih traka – tek nekoliko sekundi je potrebno za pristup bilo kojem zapisu na mediju. S obzirom na to da je riječ o hijerarhijskom sustavu koji prati statistiku korištenja podataka i one traženije drži u segmentu s izravnim pristupom, ovi sustavi mogu postići jednaku brzinu kao i izravni sustavi, ako je riječ o traženijim podacima. Ipak, zbog činjenice da se većina podataka nalazi na, u ovom primjeru, UDO diskovima, ovi su sustavi ipak nešto sporiji od onih u potpunosti utemeljenih na principu izravnog pristupa.

6. Izrada sigurnosnih kopija

Kad je riječ o sigurnosnim kopijama, najprije valja napomenuti kako bi ih trebalo redovito izrađivati. Potrebne su najmanje dvije, od kojih bi jedna trebala biti pohranjena lokalno, ali ne u istoj prostoriji sa sustavom za pohranu, a druga bi trebala biti dislocirana. Usto bi te dvije kopije trebale biti na različitim medijima za pohranu. Više o načinima izrade sigurnosnih kopija može se pronaći u poglavlju 5.6. Izravni sustavi se zbog izrade sigurnosnih kopija u skladu s navedenim uvjetima moraju oslanjati ili na vrlo skupu on-line dislociranu pohranu ili na izmjenjive medije, kao što su magnetske trake ili optički mediji. Stoga, u ovom kontekstu, izravni sustavi za pohranu ne nude kvalitetno samostalno rješenje, već se uglavnom moraju oslanjati na druge vrste medija. S druge strane, poluizravni sustavi s magnetskim trakama i hijerarhijski sustavi, u ovom primjeru, s UDO diskovima podjednako mogu zadovoljiti uvjete izrade sigurnosnih kopija koristeći se istim medijima koji se koriste i za glavno skladištenje podata-

⁶¹ Danas im jedino po pitanju brzine mogu konkurirati diskovi bez pokretnih dijelova koji za pohranu koriste memorijske čipove (engl. *Solid State Disk* – SSD), no oni još uvijek nisu dovoljno velikog kapaciteta i dovoljno niske cijene kako bi bili ozbiljna konkurencija.

ka. Jedina razlika može eventualno postojati u brzini povratka podataka u slučaju nenadanog kvara. Ona je na strani UDO diskova, no u takvim situacijama je ipak važnije vratiti sustav u neko, posljednje zabilježeno stanje, nego inzistirati na brzini tog povratka – iako i to u nekim slučajevima može biti također važno.

7. Cijena

U kontekstu usporedbe triju sustava za pohranu treba analizirati i njihovu ukupnu cijenu rada (engl. *Total Cost of Ownership* – TCO). Ta se cijena odnosi na nabavku rješenja, njegovo održavanje, cijenu medija i dodatnih medija i uređaja potrebnih za nadogradnju i proširenje sustava, potrošnju električne energije u radu itd. Izravni sustavi u ovoj kategoriji prednjače kako cijenom samog sustava tako i cijenom održavanja i proširenja kapaciteta, te su najskuplje rješenje za pohranu. Poluizravni sustavi imaju nižu nabavnu cijenu, a u troškove održavanja treba ubrojiti i nabavku uređaja i novih medija zbog proširenja, ali i zamjenske medije onim koji se istroše uporabom. Hijerarhijski sustavi s UDO diskovima također su jeftinije rješenje, jer praktički nema potrebe za zamjenu medija zbog njihove istrošenosti, već je samo potrebno uračunati uređaje i medije potrebne za proširenje kapaciteta sustava.

8. Potrošnja energije

Potrošnja energije potrebne za rad nekog sustava postaje sve važnija stavka ne samo zbog njezine sve više cijene, već ponekad i zbog imidža same tvrtke. Sustavi s izravnim pristupom najveći su potrošači energije, jer se diskovi moraju neprestano vrtjeti. Doduše, sustav može biti podešen i tako da se nakon nekog kraćeg vremena kada nema zahtjeva za podacima diskovi prestanu vrtjeti, no to će, zbog ponovnog postizanja potrebne brzine vrtnje diskova, usporiti vrijeme odziva sustava prilikom prvog zahtjeva za podacima nakon što je sustav prešao u stanje mirovanja. Ovi sustavi imaju i visoke zahtjeve za hlađenjem diskova. Poluizravni sustavi s magnetskim trakama, pak, troše manje energije jer su svi mediji u mirovanju do postavljanja zahtjeva za nekim podatkom. Ipak, traku je potrebno premotati do pozicije na kojoj su spremjeni traženi podatci i nakon njihova čitanja natrag na početak, a tu je još i robotski sustav koji je zadužen za prihvat i vraćanje trake nazad na njezino mjesto. Onaj dio hijerarhijskih sustava na kojem se nalaze podatci u izravnom pristupu troši više energije od ostatka sustava koji troši znatno manje i nema tolike potrebe za hlađenjem. Dio sustava

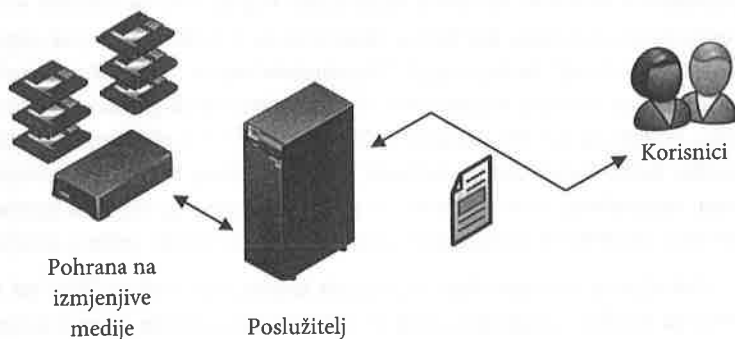
s UDO diskovima ima minimalne energetske zahtjeve i u stanju je mirovanja dok se s njega ne zatraži neki podatak. A i tada se pokrene čitanje samo s onog diska na kojem se traženi podatci nalaze.

Tablica 10: Usporedba izravnih, poluizravnih i hijerarhijskih sustava za pohranu

Karakteristike	SUSTAVI		
	Izravni (RAID sustav)	Poluizravni (magnetske trake)	Hijerarhijski (UDO diskovi)
trajnost medija	vrlo dugotrajni, beskontaktni, neosjetljivi na okolinske uvjete	trajni, kontaktni, osjetljivi na okolinske uvjete	dugotrajni, beskontaktni, neosjetljivi na okolinske uvjete
kapacitet	visok, skupo proširenje	vrlo visok, jeftino proširenje	relativno visok, jeftino proširenje
osiguranje autentičnosti	softverskom emulacijom WORM načina zapisivanja	hardverskom emulacijom WORM načina zapisivanja	WORM način zapisivanja – najpogodnije rješenje
uništenje podataka	višestrukim brisanjem fizičke pozicije na disku	nije moguće brisanje pojedinih podataka	uništenje medija (WORM) ili brisanje (RW)
brzina pristupa	vrlo visoka	relativno sporo rješenje	visoka
sigurnosne kopije	on-line dislocirana pohrana ili korištenje izmjenjivih medija	isti mediji kao i za skladištenje podataka, pogodno rješenje	isti mediji kao i za skladištenje podataka, pogodno rješenje
cijena	najskuplje rješenje	jeftinije proširenje, potreba za zamjenskim medijima	jeftinije proširenje
potrošnja energije	vrlo visoka	srednja	niža

5.4. NEIZRAVNI SUSTAVI ZA POHRANU I PRIENOS PODATAKA

Kod neizravnih sustava digitalnom se gradivu ne može izravno pristupiti, iako su metapodatci o tom gradivu dostupni putem mreže. Vrlo je važno da baza metapodataka (metabaza) bude usklađena s podacima spremljenim u digitalnom arhivu. Ovakvi sustavi ubrzavaju pretraživanje i pronalaženje digitalnoga gradiva, iako mu ograničavaju pristup. Neizravni se sustavi vrlo često koriste kao sustavi za izradu i održavanje sigurnosnih kopija (engl. *backup*).⁶² Glavna karakteristika neizravnih sustava jest činjenica da je potrebna ljudska intervencija za pristup pohranjenim podacima.



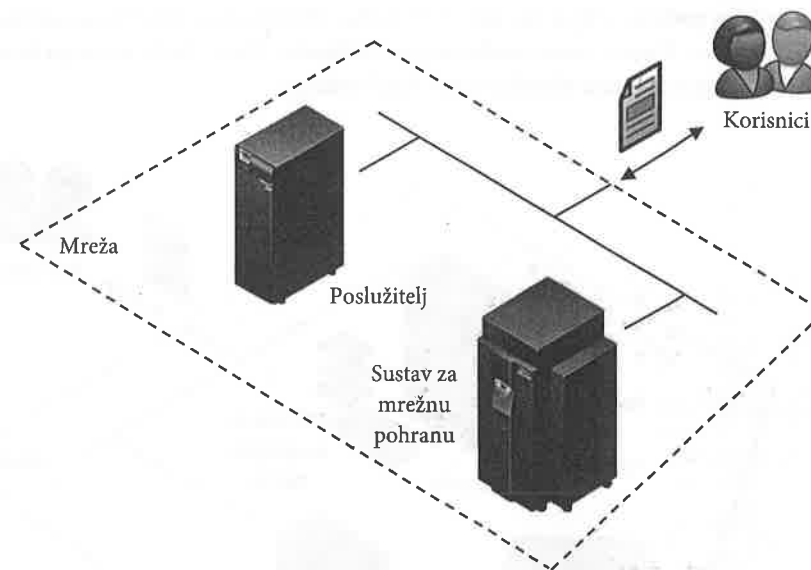
Slika 33: Neizravni sustavi za pohranu i prijenos podataka

5.5. MREŽNA POHRANA

Mrežna pohrana (engl. *Network Attached Storage – NAS*) je koncept pohrane velike količine podataka vrlo sličan izravnim sustavima za pohranu, iako se od njih ipak razlikuje. Razlika je prije svega u tome što se sustavi za mrežnu pohranu priključuju na mrežu kao samostalne i neovisne jedinice, dok se kod izravnih sustava slični uređaji spajaju na poslužitelj i funkcioniraju kao proširenje nje-

⁶² *Archiving and Managing Digital Information*, Commission of the European Communities, 4. studenog 1997., <http://www.kfs.oeaw.ac.at/publications/1999_deutsch_remote_access_and_transfer_of_audio_recordings/d3x3a1.htm>, 29. rujna 2008.

gova smještajnog kapaciteta. Ovi su sustavi, dakle, spojeni izravno na mrežu, uglavnom nemaju priključen monitor niti tastaturu te im se za potrebe konfiguracije pristupa mrežnim preglednikom spajajući se izravno na njihovu mrežnu adresu. Vrlo često je riječ o RAID sustavu diskova za pohranu podataka.⁶³



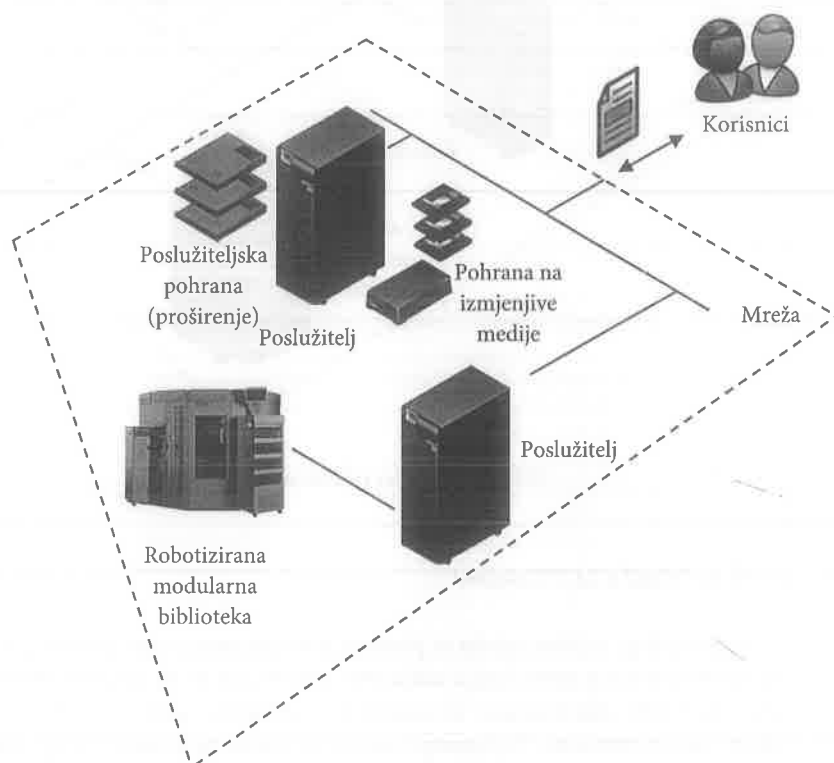
Slika 34: Mrežna pohrana podataka

5.6. MREŽA ZA POHRANU

U posljednje vrijeme razvio se poseban koncept namijenjen pohrani podataka na institucijskoj razini koji je usklađen s činjenicom da su današnje institucije sve više fizički distribuirane, ali međusobno povezane globalnom informacijskom infrastrukturom. Taj koncept naziva se *mreža za pohranu* (engl. *Storage Area Network – SAN*). Mreža za pohranu povezana je s računalnom mrežom institucije, ali zapravo čini podmrežu koja povezuje sve uređaje za trajnu pohranu i izradu sigurnosnih kopija. Sve do nedavno je pohrana bila organizirana tako da su na više mjesta postojala poslužiteljska računala s implementiranim nekim

⁶³ *Wikipedia*, s.v. Network-attached storage, 30. srpnja 2008., <http://en.wikipedia.org/wiki/Network-attached_storage>, 4. kolovoza 2008.

od prije navedenih rješenja za pohranu. Mreža za pohranu ujedinjuje sve te sustave te se njima može upravljati s jednog mjesta. Ovakvo rješenje ne samo da je jednostavnije za administriranje, već se njime optimalnije iskorištava prostor za pohranu.⁶⁴ Vrlo često se ovi sustavi uspoređuju sa sustavima za mrežnu pohranu (NAS), no razlika je, između ostaloga, u tome što SAN prilikom pohrane organizira podatke u blokove, dok NAS koristi datoteke kao osnovnu organizacijsku jedinicu. Unatoč ovim međusobnim razlikama, NAS i SAN se mogu kombinirati, čime se stvara hibridno SAN-NAS rješenje.



Slika 35: Mreža za pohranu

⁶⁴ Wright, Marilyn, Trendy (and Tried) Ways to Secure Your Knowledge, *Knowledge Management World*, Vol. 7, br. 12, 1998., <[http://www.kmworld.com/Articles/Editorial/Feature/Trendy-\(and-tried\)-ways-to-secure-your-knowledge-9045.aspx](http://www.kmworld.com/Articles/Editorial/Feature/Trendy-(and-tried)-ways-to-secure-your-knowledge-9045.aspx)>, 19. travnja 2001.

Pitanje odabira optimalnog sustava za pohranu i prijenos podataka prije svega ovisi o odnosu zahtijevane brzine, potrebnoj kapacitetu, količini financijskih sredstava predviđenih za ovu stavku projekta digitalizacije te tempu digitalizacije. Institucije koje digitaliziraju mnogo dokumenata, bilo da se radi o arhivima, knjižnicama i muzejima, ili pak nekim drugim institucijama koje stvaraju vlastite digitalne fondove, odabrat će različiti sustav za skladištenje digitalnoga gradiva od onih institucija koje ne digitaliziraju tako brzim tempom.

5.7. SIGURNOSNE KOPIJE

Nevezano uz ukupnu godišnju količinu proizvedenog digitalnoga gradiva, potrebno je redovito izrađivati sigurnosne kopije. One se također nazivaju rezervnim ili pričuvnim kopijama, sigurnosnim preslikama ili, pak, sigurnosnim duplikatima (engl. *backup copy*). "Sigurnosna kopija ... jest kopija podataka koja se izrađuje u svrhu osiguranja u slučaju oštećenja ili gubljenja izvornih podataka. Podatci mogu biti datoteke ili/i programi."⁶⁵ Do oštećenja ili gubljenja podataka može doći iz raznih razloga, primjerice zbog hardverskog kvara na uređaju, računalnog virusa, iznenadne promjene strujnog napona, nestručnog rukovanja, otuđenja ili elementarne nepogode poput poplave, požara, zemljotresa i sl.

Uobičajena je izrada dviju odvojenih kopija, i to preporučljivo na različitim vrstama medija za pohranu, čime se vjerojatnost gubitaka znatno smanjuje. Dodatno, barem jedna sigurnosna kopija bi trebala biti dislocirana. Osim ovih, relativno jednostavnih preporuka postoje i znatno složeniji modeli izrade i rotacije sigurnosnih kopija. Upravo će o njima biti riječi u nastavku ovog poglavlja. Gledajući s pozicije organizacije procesa izrade sigurnosnih kopija, svaki takav proces trebao bi se sagledati i definirati kroz pet osnovnih koraka:

1. odabir modela skladištenja podataka,
2. model rotacije medija,
3. odabir medija za pohranu,
4. obrada podataka sigurnosne kopije,
5. pohrana.

Neki elementi navedenih osnovnih koraka već su prije objašnjeni, no ovdje će biti ponovo spomenuti kako bi se mogao bolje shvatiti kontekst u kojem se oni pojavljuju.

⁶⁵ Kiš, Informatički rječnik, n. dj., str. 92, s.v. backup.

5.7.1. Odabir modela skladištenja podataka

Odabir modela skladištenja podataka prvi je korak u organizaciji procesa izrade sigurnosnih kopija. U načelu postoji pet različitih modela skladištenja podataka: 1. nestrukturirana sigurnosna kopija, 2. potpuna sigurnosna kopija, 3. inkrementalna sigurnosna kopija, 4. diferencijalna sigurnosna kopija i 5. kontinuirana zaštita podataka. Svaka institucija treba odabrati onaj model koji njoj najviše odgovara ili koji je najviše u skladu s njezinim načinom poslovanja i implementiranim poslovnim procesima. Odabir modela skladištenja podataka često je povezan s odabirom modela rotacije medija pa se u praksi ova dva koraka vrlo često rješavaju paralelno.

5.7.1.1. Nestrukturirana sigurnosna kopija

Izrada nestrukturirane sigurnosne kopije najjednostavniji je model izrade sigurnosnih kopija. Ovaj model predviđa kopiranje podataka na rezervne medije uz dodavanje informacije što se na kojem mediju nalazi i kada je tamo bilo pohranjeno. Već je na prvi pogled jasno da je ovaj model vrlo jednostavno implementirati u praksi, no, također, da rezultira vrlo niskom razinom sigurnosti u slučaju potrebe za povratkom podataka. On je, stoga, pogodan za kućnu i manje zahtjevnu uredsku upotrebu, potencijalno samo za razinu pojedinih računala u slučaju kad institucija na nekom drugom, centralnom mjestu provodi neki kompleksniji model izrade sigurnosnih kopija.

5.7.1.2. Potpuna sigurnosna kopija

Ovaj model predviđa izradu potpune kopije nekog sustava – svih podataka koji se na njemu nalaze, uključujući i radnu okolinu, tj. instalirani operativni sustav i sve na njemu instalirane programe. Također, moguće je napraviti i sigurnosnu kopiju samo određenog dijela podataka na računalu (npr. samo podataka, a ne i operativnog sustava s programima). Naravno da ovakav pristup izradi sigurnosnih kopija s jedne strane zahtijeva mnogo prostora na medijima na kojima se kopija sprema, dok s druge strane pruža mogućnost vrlo brzog povratka cjelokupnog sustava dokle god se taj sustav vraća na isto računalo nepromijenjene konfiguracije. Model izrade potpune sigurnosne kopije praktičan je kod manjih sustava i onih koji se ne koriste 24 sata dnevno. Kad je riječ o nekim većim i složenijim sustavima, ovaj se model može kombinirati s nekim drugim

modelima izrade sigurnosnih kopija, na primjer inkrementalnim ili diferencijalnim modelom, a također će biti vidljivo i kroz model rotacije medija.

5.7.1.3. Inkrementalna sigurnosna kopija

Model izrade i korištenja inkrementalne sigurnosne kopije jest model za koji je potreban veći broj medija. Naime, on se vrlo često implementira s postupkom izrade potpune sigurnosne kopije, ali čini nadgradnju toga modela. Najprije je potrebno izraditi potpunu sigurnosnu kopiju, a u nekom sljedećem trenutku, kad je potrebno ponovno napraviti sigurnosnu kopiju, sprema se samo razlika u odnosu na prethodno zabilježeno stanje. Dakle, druga sigurnosna kopija inkrementalnog tipa zahtijevat će znatno manje prostora, jer su u njoj zabilježene samo nove datoteke ili one koje su se, u odnosu na prethodno stanje, promijenile. Ipak, za povratak cjelokupnog sustava bit će potrebno iskoristiti sve pohranjene inkrementalne kopije, počevši od prve prema zadnjoj, kako bi se u potpunosti vratio cjelokupni sustav. Treba napomenuti da se ovaj model može koristiti i bez početne izrade potpune sigurnosne kopije. U tom se slučaju odredi početni set podataka koji se želi osigurati izradom sigurnosne kopije te se on zabilježi, a daljnji su koraci jednaki kao i kad je kao početak uzeta potpuna sigurnosna kopija.

5.7.1.4. Diferencijalna sigurnosna kopija

Model izrade diferencijalne sigurnosne kopije također se vrlo često koristi u kombinaciji s modelom izrade potpune, ali i inkrementalne sigurnosne kopije. Naime, model je vrlo sličan inkrementalnom modelu jer također bilježi samo one datoteke koje su nove ili su se promijenile u odnosu na prethodnu izradu sigurnosne kopije. Prednost ovog modela jest u tome da je za potpuni oporavak sustava u slučaju havarije potrebno imati na raspolaganju samo prijašnju potpunu i posljednju diferencijalnu sigurnosnu kopiju, dok se međukopije u tom postupku neće koristiti. One mogu poslužiti, ako je to potrebno, za povratak u neko od zabilježenih međustanja. Ovaj je model izrade sigurnosnih kopija, u odnosu na prethodno objašnjene modele, najbrže i najelegantnije rješenje.

5.7.1.5. Kontinuirana zaštita podataka

Model kontinuirane zaštite podataka najnapredniji je model izrade sigurnosnih kopija. Za razliku od prethodno objašnjenih modela koji su pratili promjene na razini pojedinih datoteka, ovaj model prati promjene na znatno nižoj razini – razini pojedinih bajtova, tj. memorijskih blokova. I kod ovog je modela potrebno imati početni set podataka, bilo da je riječ o potpunoj sigurnosnoj kopiji ili nekom određenom setu podataka. Kod ovog modela sustav neprekidno prati i bilježi samo promjene do kojih je došlo na navedenoj, nižoj razini te ih sprema na pričuvenu lokaciju ili medije. Takva kopija zauzima znatno manje kapaciteta na medijima za pohranu sigurnosnih kopija. U slučaju potrebe, moguće je vrlo brzo oporaviti cijeli sustav ili sigurnosnom kopijom obuhvaćene podatke, pri čemu su potrebni izvorni set podataka i posljednji zabilježeni set promjena. S obzirom na to da sustav bilježi sva međustanja, moguće je cjelokupan sustav ili tako osigurani set podataka vratiti u bilo koje zabilježeno stanje između početnog i posljednjeg zabilježenog.

5.7.2. Model rotacije medija

U praksi se vrlo često za izradu sigurnosnih kopija koristi više pojedinačnih medija. Tako se, primjerice, jedan može koristiti za spremanje svih promjena učinjenih u ponedjeljak, drugi za one učinjene u utorak itd. Zbog toga su i osmišljeni razni modeli rotacije medija koji predviđaju korištenje određenog broja medija i određuju način njihovoga korištenja u svrhu osiguranja sigurnosti podataka, kao i mogućnosti povratka u neko prijašnje stanje. Modeli se međusobno razlikuju prema broju medija koje predviđaju, složenosti rotacije medija i vremenu čuvanja podataka na njima prije nego što budu prebrisani i zamijenjeni novima. Svakako treba napomenuti da se kod svih spomenutih modela, na mjestu gdje se spominje korištenje jednog medija, u tu svrhu može koristiti i više adekvatno obilježenih medija ako sadržaj koji se sprema u sigurnosnu kopiju prelazi kapacitet odabrane vrste medija za pohranu. Tri najčešća modela rotacije medija za izradu sigurnosnih kopija su: 1. inkrementalni model, 2. model djed-otac-sin i 3. model hanojskih tornjeva.

5.7.2.1. Inkrementalni model

Inkrementalni model rotacije medija najjednostavniji je model izrade sigurnosnih kopija. Uglavnom ga koriste nezahtjevni korisnici. Uzmimo za primjer da neki sustav radi 7 dana u tjednu i da se na kraju svakog dana izrađuje sigurnosna kopija. Neka se kao mediji na koje se spremaju sigurnosne kopije koriste DVD mediji. Ovaj model predviđa korištenje 7 DVD medija, pri čemu se svaki koristi za jedan dan u tjednu. Sljedećeg tjedna ciklus kreće iz početka. Svaka napravljena sigurnosna kopija ostaje dostupna 7 dana od trenutka izrade. Sljedeća tablica prikazuje ovaj pristup.

Tablica 11: Osnovni inkrementalni model izrade sigurnosnih kopija

	Pon	Uto	Sri	Čet	Pet	Sub	Ned
1. tjedan	DVD 1	DVD 2	DVD 3	DVD 4	DVD 5	DVD 6	DVD 7
2. tjedan	DVD 1	DVD 2	DVD 3	DVD 4	DVD 5	DVD 6	DVD 7
3. tjedan	DVD 1	DVD 2	DVD 3	DVD 4	DVD 5	DVD 6	DVD 7

Tablica 12: Složeni inkrementalni model izrade sigurnosnih kopija

	Pon	Uto	Sri	Čet	Pet	Sub	Ned
1. tjedan	DVD 1	DVD 2	DVD 3	DVD 4	DVD 5	DVD 6	DVD 7
2. tjedan	DVD 2	DVD 3	DVD 4	DVD 5	DVD 6	DVD 7	DVD 8
3. tjedan	DVD 3	DVD 4	DVD 5	DVD 6	DVD 7	DVD 8	DVD 9
4. tjedan	DVD 4	DVD 5	DVD 6	DVD 7	DVD 8	DVD 9	DVD 10
5. tjedan	DVD 5	DVD 6	DVD 7	DVD 8	DVD 9	DVD 10	DVD 11
6. tjedan	DVD 6	DVD 7	DVD 8	DVD 9	DVD 10	DVD 11	DVD 1

Postoji, također, i modifikacija inkrementalnog modela. Ona je nešto složenija i za sedmodnevnu izradu sigurnosnih kopija kroz pet tjedana predviđa 11 medija. Prvi je tjedan identičan osnovnom inkrementalnom pristupu, no već se drugog tjedna ne započinje ciklus s prvim medijem, već s drugim, a završava s novim – osmim. Sljedećeg tjedna proces se ponavlja i tek se nakon 5 tjednih ciklusa, tj. u 6. tjednu, ponovo koristi prvi medij, ali kao posljednji. Dakle, nakon što se medij jednom iskoristi za spremanje podataka u ponedjeljak, oni na njemu ostaju pohranjeni 40 dana. Prednost ovakvog, složenog inkrementalnog modela jest u tome da se po potrebi mogu dodavati, ili oduzimati, mediji koji se

koriste u ciklusima pohrane i tako produžavati vrijeme u kojem se čuva početna sigurnosna kopija zabilježena u ponedjeljak. Tablica 12 prikazuje ovaj pristup.

5.7.2.2. Model djed-otac-sin

Model djed-otac-sin (engl. *Grandfather-Father-Son*) rotacije medija vrlo je popularni model. Model će ovdje biti objašnjen na primjeru petodnevno radnog tjedna unutar mjesec dana. Cijeli proces započinje izradom potpune pričuvene kopije cijelog sustava i to u petak. Zatim se svakog dana od ponedjeljka do četvrtka izrađuju inkrementalne kopije, a u petak se ponovo izrađuje potpuna sigurnosna kopija, ali na zasebnom mediju u odnosu na prethodni petak. Svaka potpuna sigurnosna kopija izrađena petkom može se čuvati lokalno dok se njezina kopija može slati na udaljenu lokaciju zbog povećanja sigurnosti. Isti se postupak ponavlja sljedećeg tjedna s razlikom da se za dnevne inkrementalne kopije koriste mediji od prethodnog tjedna, dok se za petak ponovo koristi novi medij. Posljednjeg petka u mjesecu koji, primjerice, ima 4 tjedna također se izrađuje takozvana mjesečna potpuna sigurnosna kopija koja se čuva godinu dana. Dakle, za ovaj su model potrebna 4 medija za inkrementalne kopije od ponedjeljka do četvrtka, zatim, u mjesecima s četiri tjedna, 3 medija za svaki petak i 1 medij za mjesečnu potpunu sigurnosnu kopiju, tj. ukupno 12 mjesečnih medija godišnje. Drugim riječima, za provođenje ovog modela treba najmanje 19 medija. S obzirom na to da neki mjeseci imaju i po 5 tjedana, broj potrebnih medija se time povećava. Sljedeća tablica prikazuje ovaj pristup.

Tablica 13: Model djed-otac-sin (prva dva mjeseca godišnjeg ciklusa)

		Inkrementalne kopije				Potpune kopije	
		Pon	Uto	Sri	Čet	Pet	
Mjesec 1	1. tjeđan					PET 1	
	2. tjeđan	PON 1	UTO 1	SRI 1	ČET 1	PET 2	
	3. tjeđan	PON 1	UTO 1	SRI 1	ČET 1	PET 3	
	4. tjeđan	PON 1	UTO 1	SRI 1	ČET 1		MJESEC 1
Mjesec 2	1. tjeđan	PON 1	UTO 1	SRI 1	ČET 1	PET 1	
	2. tjeđan	PON 1	UTO 1	SRI 1	ČET 1	PET 2	
	3. tjeđan	PON 1	UTO 1	SRI 1	ČET 1	PET 3	
	4. tjeđan	PON 1	UTO 1	SRI 1	ČET 1		MJESEC 2

5.7.2.3. Model hanojskih tornjeva

Model hanojskih tornjeva (engl. *Tower of Hanoi*) je, u odnosu na prethodno objašnjene, najsloženiji model izrade sigurnosnih kopija. On se temelji na rekurzivnoj matematičkoj funkciji, pri čemu se prvi medij za pohranu koristi svaki drugi dan, drugi svaki četvrti dan, treći svaki osmi dan itd. Iz toga proizlazi da se formulom 2^{n-1} , gdje n predstavlja broj korištenih medija, može izračunati koliko se dugi ciklusi izrade sigurnosnih kopija mogu ostvariti. Na primjer, ako se koriste 4 medija, s njima se može ostvariti osmodnevni ciklus. Isto tako, u slučaju korištenja 8 medija tada, nakon početnog ciklusa, uvijek postoji 128 dana stara sigurnosna kopija, što znači vrlo dobru sigurnost u slučaju pada sustava, zaraze neke važne datoteke virusom, slučajnog brisanja itd. Sljedeća tablica prikazuje ovaj pristup.

Tablica 14: Izrada sigurnosnih kopija prema modelu hanojskih tornjeva⁶⁶

		Dani															
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Mediji	A	A		A		A		A		A		A		A		A	
	B		B				B				B				B		
	C				C								C				
	D							D									
	E																E

5.7.3. Odabir medija za pohranu

Kao mediji za pohranu mogu se koristiti bilo koji mediji koji su primjereni zbog svojeg kapaciteta, brzine, trajnosti, pouzdanosti, jednostavnosti rukovanja itd. Za pohranu su, dakle, iskoristivi svi mediji nabrojani na početku ovog poglavlja (v. podjelu prema načinu zapisa i vrsti medija). No, oni će ovdje ipak biti spomenuti zbog toga jer će uz njih biti spomenute i neke njihove karakteristike.

⁶⁶ Prema: *Backup Basics: Reliability, Redundancy and Recovery*, White Paper, Quantum Corporation, 2006., <http://www.datasecurityprovider.com.br/anexos/backup_basics.pdf>, 4. kolovoza 2008.

Za izradu sigurnosnih kopija iskoristivi su:

- izmjenjivi diskovi
 - CD, DVD, BD – od 800 MB do 50 GB, relativno jeftini, rasprostranjeni, brzi pristup podacima, jednostavni za rukovanje, relativno kratkog vijeka trajanja, relativno osjetljivi na vanjska oštećenja
 - UDO – do 120 GB, skuplji, nisu široko rasprostranjeni, brzi pristup podacima, jednostavni za rukovanje, dugotrajni, pouzdani, neosjetljivi na vanjska oštećenja
- trake – magnetske (do 800 GB nekomprimirano) i optičke (znatno višeg kapaciteta), rasprostranjene, spor pristup podacima, jednostavne za rukovanje, manje osjetljive na fizičke vanjske utjecaje, ali zato osjetljive na okolinske uvjete
- tvrdi diskovi
 - magnetski – danas do nekoliko terabajta (s tendencijom povećanja kapaciteta), vrlo brzi pristup, osjetljivi zbog korištenja vrlo preciznih mehaničkih dijelova, mogu se koristiti i kao izmjenjivi mediji upotrebom ladica za jednostavno umetanje i vađenje medija
 - memorijski (SSD) – danas manje od 100 GB (s tendencijom brzog povećanja kapaciteta), najbrži pristup, još uvijek skupo rješenje, prenosivi, neosjetljivi na vanjske utjecaje, spajanje putem USB sučelja
- udaljeni servisi za izradu sigurnosnih kopija – mrežna izrada pričuvnih kopija, SAN – pristupačni zbog postojanja mrežne veze visoke propusnosti, netko drugi brine o sigurnosti sustava, potrebno logički organizirati spremanje kopija prema nekom od modela za izradu sigurnosnih kopija, jer se ne koriste izmjenjivi mediji, osim u slučaju da udaljeni servis koristi robotizirani sustav s trakama ili optičkim diskovima, ali i tada udaljeni sustav sam određuje kako će ih koristiti

5.7.4. Obrada podataka sigurnosne kopije

Sljedeći korak u pokretanju postupka izrade sigurnosnih kopija jest odabir jednog od načina obrade podataka sigurnosne kopije. Odluka da će se podatci spremljeni u sigurnosne kopije ili, pak, cijela sigurnosna kopija dodatno obrađivati može povoljno utjecati na optimiziranje cijelog procesa, ali i donijeti koju poteškoću. Ovdje treba napomenuti da prilikom spomenute obrade ne dolazi do bilo kakve promjene podataka koji su pohranjeni te njihov integritet i autenti-

čnost nisu dovedeni u pitanje. Postoji nekoliko načina obrade podataka prilikom izrade sigurnosnih kopija, a najvažniji među njima su: 1. kompresija, 2. deduplikacija, 3. duplikacija i 4. šifriranje.

5.7.4.1. Kompresija

Kompresija podataka, kako joj i samo ime govori, služi za komprimiranje podataka spremljenih na neki medij. Prilikom korištenja magnetskih traka, to je opcija koja se može uključiti već i prilikom izrade same sigurnosne kopije. Uređaj za pohranu samostalno provodi kompresiju. Naravno da se tim postupkom smješta više podataka na manji prostor, tj. kapacitet medija se skoro udvostručuje, ali u slučaju gubitka podataka zbog fizičkog oštećenja gubi se duplo više podataka. Spremanje podataka na traku ponekad je sporije. Komprimiranje, također, može biti postavljeno kao automatsko u slučaju korištenja čvrstih diskova ili polja čvrstih diskova. Prilikom korištenja optičkih medija ne postoji mogućnost automatskog komprimiranja podataka, osim ako tu opciju nema program s kojim se provodi izrada pričuvnih kopija. No, s obzirom na današnju procesorsku i memorijsku snagu računala, kompresija podataka preporučljiva je zbog uštede prostora i količine medija koji su potrebni za izradu pričuvnih kopija.

5.7.4.2. Deduplikacija

Do duplikacije datoteka koje se spremaju u sigurnosne kopije može doći u slučaju kada se podatci s više računala, radnih stanica ili poslužitelja spremaju u obliku sigurnosne kopije na jedno mjesto. Tada se može provesti postupak deduplikacije kojim se, eventualno postojeće, identične datoteke brišu na način da ostanu zabilježena samo jedna njihova kopija. Naravno, umjesto ostalih kopija ostaju zabilježene pozicije na koje se one, u slučaju potrebe, moraju vratiti. Na taj se način, napose prilikom izrade potpunih sigurnosnih kopija, može uštedjeti na prostoru, ali i vremenu potrebnom za izradu pričuvnih kopija.

5.7.4.3. Duplikacija

Postupak duplikacije se, u ovom kontekstu, odnosi na umnažanje već izrađenih sigurnosnih kopija. Razlozi za to mogu biti različiti, ali oni najčešći su zbog sigurnosnih razloga. Naime, takva se kopija tada može dislocirati, tj. poslati na

udaljenju lokaciju, kako to i nalažu sigurnosne i arhivske smjernice. Drugi razlog umnažanja sigurnosne kopije može biti ubrzanje oporavka sustava u slučaju havarije. Tada je vrlo vjerojatno, pogotovo kod havarije većih sustava, da će takvim postupkom doći do ubrzavanja oporavka sustava, jer se neki postupci mogu odvijati paralelno.

5.7.4.4. Šifriranje

Šifriranje (kriptiranje, enkripcija) je postupak osiguranja sigurnosne kopije od neovlaštenog pristupa u njoj zabilježenim podacima. Iako pruža dodatnu zaštitu, važno je napomenuti da sam postupak šifriranja ipak usporava izradu sigurnosne kopije. Osim toga, u slučaju šifriranja podataka i njihovog čuvanja tako šifriranih na dulji rok, treba imati na umu mogućnost zaboravljanja šifre, mogućnost prelaska na korištenje drugog računalnog rješenja, operativnog sustava ili programskog rješenja koje ne mora biti kompatibilno s ranije šifriranim sigurnosnim kopijama. Ako se ova situacija ne uzme u obzir, sačuvane sigurnosne kopije mogu u jednom trenutku postati nečitljive. Stoga ovakve situacije treba spriječiti dobrim planiranjem postupaka obrade sigurnosnih kopija.

5.7.5. Pohrana

Pohrana sigurnosnih kopija posljednji je korak u organizaciji procesa izrade sigurnosnih kopija. Već je prije bilo riječi o medijima koji se u tu svrhu mogu koristiti baš kao i o sustavima – izravnim, poluizravnim, hijerarhijskim, neizravnim, mrežom za pohranu. No, na ovome mjestu je potrebno ipak dodati i još jednu mogućnost, a to je pohrana sigurnosnih kopija na udaljenom mjestu, ali u trezoru (engl. *Off-Site Vault*). Ova je metoda najsigurniji oblik zaštite sigurnosnih kopija, jer se one nalaze u fizički zaštićenoj okolini. Takve kopije mogu biti i dodatno zaštićene nekim od postupaka šifriranja. Ipak, smještanje sigurnosnih kopija u trezor na udaljenom mjestu je skup i zahtjevan postupak koji će se primjenjivati u rjeđim, ali zasigurno opravdanim situacijama kada su u pitanju vrlo povjerljivi podatci.

6. PREGLED I KORIŠTENJE DIGITALNOGA GRADIVA

Prilikom razrade projekta digitalizacije mora se razraditi i način pregledavanja i korištenja, jer je to jednako važan korak kao i ostali. Naime, pitanja pregleda i korištenja digitaliziranoga gradiva ovise o tome hoće li se gradivo samo pregledavati na zaslonu ili će se također i ispisivati na pisačima, te hoće li korisnik pretraživati i koristiti građu lokalno ili putem Interneta ili će mu, pak, biti omogućeno samo pretraživanje metapodataka i korištenje tekstualnog gradiva putem Interneta, a korištenje ostale vrste gradiva isključivo lokalno putem radnih stanica dostupnih javnosti u samoj instituciji i sl.

Ako se omogućiti ispis gradiva na zahtjev, tada se moraju predvidjeti i uređaji namijenjeni različitoj vrsti gradiva. Ako je riječ o tekstualnom gradivu, većina današnjih pisača ima kvalitetan crno-bijeli ispis. No, ako se radi o slikovnom gradivu, tada posebnu pažnju treba obratiti na odabir pisača, jer on mora davati zadovoljavajuće rezultate. Ako se radi o ispisu slika u nijansama sive boje, onda razlučivost ispisa mora biti jednaka ili veća od razlučivosti slikovnog zapisa, premda ispisi pri razlučivosti od 600 dpi daju tek kvalitetu usporedivu sa novinskim slikama. Za ispis u boji kvalitetne rezultate daju pisači koji rade na principu sublimacije,⁶⁷ laserski pisači u boji te foto ink-jet pisači. Također se mora predvidjeti i mogućnost ispisa većih formata, te odabrati oprema za ispis takvog gradiva. Specifičan je problem 3D ispis trodimenzionalnih elektroničkih objekata nastalih 3D skeniranjem objekata kulturne baštine. Uređaji za 3D ispis vrlo su specifična i skupa rješenja pa se ne može očekivati njihova masovna nabava. Stoga će usluga 3D ispisa u većini slučajeva biti obavljana izvan institucije u kojoj se nalazi izvorno i/ili digitalizirano 3D gradivo.

Dakle, potrebno je razmišljati i iz perspektive krajnjeg korisnika i načina njegova pristupa gradivu i mogućnosti pružanja dodatnih (komercijalnih) usluga,

⁶⁷ Sublimacija je proces prelaženja čvrste tvari zagrijavanjem izravno u plinovitu i ponovno kondenziranje para u čvrstu tvar.

primjerice ispisa na zahtjev, koje cijelom procesu digitalizacije daju “završni glanc”.

Uz ovaj aspekt treba svakako razmišljati i o tome da nisu korisnici digitalnih i digitaliziranih zbirki isključivo ljudi, već su to, u komunikacijskom procesu, i programi i programski agenti. Važno je, u tom kontekstu, razmišljati i kojim se operativnim sustavima, programima, preglednicima mrežnih sadržaja i protokolima koriste ciljani korisnici. U suprotnom, postoji mogućnost da im ponuđeni sadržaj neće biti dostupan ili barem ne u željenom obliku. Provjeru i preispitivanje ovog tipa potrebno je dodatno provoditi s vremena na vrijeme kako bi ponuđeni sadržaj ostao svima dostupan. Upravo sljedeće poglavlje iscrpnije razlaže ovu problematiku.

7. ODRŽAVANJE DIGITALNOGA GRADIVA

Iako je na početku ovog rada održavanje digitalnoga gradiva navedeno kao posljednji korak procesa digitalizacije, on to i jest i nije. Naime, on nije izravno povezan s postupcima digitalizacije, ali cijeli proces digitalizacije bez njega ne bi imao smisla, jer bi gradivo vrlo skoro postalo zastarjelo i više ne bi bilo dostupno korisnicima. Upravo je zbog tog dvojnog svojstva ovaj korak izdvojen kao zasebni problem. Prije svega on se sadržajno može sagledati kroz dva, vrlo slična, problema: problem održavanja digitaliziranoga gradiva i problem održavanja gradiva izvorno nastalog u elektroničkom obliku. Razlike u održavanju, tj. dugoročnom očuvanju obiju vrsta gradiva gotovo da i nema. Jedini aspekt u kojem se, u ovom koraku, ove dvije “vrste” gradiva razlikuju jest činjenica da jedna ima, a druga nema izvorni analogni oblik, pa je u slučaju gubitka elektroničke verzije jedne moguće ponovo digitalizirati, a druge, pak, nije. Ipak, takve postupke ponavljanja već obavljenoga posla potrebno je izbjeći kad god je to moguće. U narednom će dijelu, dakle, biti objašnjeni problemi do kojih dolazi prilikom zastarijevanja sustava, medija i zapisa.⁶⁸

7.1. DEFINICIJA ZAPISA

Zbog boljeg razumijevanja ovih problema najprije je potrebno definirati osnovne pojmove. S tradicionalnog stajališta, arhiviranje dokumenata zasnivalo se na očuvanju fizičkog objekta, dakle medija, kao nositelja informacije. Tako definicija *zapisa* iz 1980-ih glasi:

“Zapis je dokument (uključujući bilo koji rukom pisani ili tiskani materijal) ili objekt (uključujući zvučni zapis, magnetsku traku ili disk, mikrofilm, fotografiju, film, mapu, plan, model, sliku ili neko drugo slikarsko ili

⁶⁸ Opisi prema: Stančić, Hrvoje, Arhiviranje digitalnih dokumenata, u: Willer, Mirna i Katić, Tinka (ur.), 4. seminar Arhivi, knjižnice i muzeji. Mogućnosti suradnje u okruženju globalne informacijske infrastrukture, Hrvatsko knjižničarsko društvo, Zagreb, 2001., str. 209-214.

grafičko djelo) koji jest ili je bio namjerno sačuvan zbog bilo koje informacije ili sadržaja koje posjeduje, ili značenja koje može biti iz njega izvedeno, ili zbog njegove veze s nekim događajem, osobom, okolnošću ili stvari.”⁶⁹

Danas, kada arhiviranje dokumenata sve više podrazumijeva i arhiviranje elektroničkih dokumenata, bilo da su oni nastali u digitalnom obliku ili da čine elektroničku verziju klasičnih dokumenata, ovakva je definicija manjkava iz više razloga. Naime, u elektroničkoj verziji, zapis nije niti samo dokument niti samo objekt, nezavisan je od medija na kojemu se nalazi te nije fizički zapisan u računalnom sustavu u logičkom slijedu, već je fragmentiran i može se nalaziti čak i na fizički različitim jedinicama (na primjer, kod RAID sustava). Dakle, postoji razlikovanje fizičke i logičke strukture. Stoga je bilo potrebno osuvremeniti definiciju zapisa te ugraditi generički set karakteristika i naglasiti neovisnost o formatu. Tako, prema novoj definiciji:

- Zapis je ono što je stvoreno i sačuvano kao dokaz funkcija, aktivnosti i transakcija neke tvrtke ili pojedinca.
- Da bi se smatrao dokazom, zapis mora imati sadržaj, strukturu i kontekst te mora biti dio sustava za arhiviranje.⁷⁰

Ova definicija dozvoljava postojanje zapisa nezavisnog od medija. Definiranjem zapisa preko njegovih temeljnih karakteristika omogućeno je njegovo postojanje neovisno o trenutačnoj tehnologiji, čime zapisi postaju virtualni.

Imajući na umu primarno aspekt informacijske tehnologije (IT), zapis se definira kao:

- skup objektnih podataka, po mogućnosti s različitim atributima, koji imaju pridijeljene identifikatore, ili
- skup podataka koji se smatra jedinicom ili skup koji se sastoji od jedne ili više jedinica podataka grupiranih radi obrade.⁷¹

⁶⁹ *The Archives Act*, Australia, 1983., <http://www.austlii.edu.au/au/legis/cth/consol_act/aa198398/>, 10. veljače 1999.

⁷⁰ *Digital Recordkeeping. Guidelines for Creating, Managing and Preserving Digital Records*, National Archives of Australia, svibanj 2004., <http://www.naa.gov.au/Images/Digital-recordkeeping-guidelines_tcm2-920.pdf>, 6. svibnja 2000.

⁷¹ *IBM Dictionary of Computing*, McGraw Hill, New York, 1994., str. 561.

7.2. DEFINICIJA PROBLEMA

Prethodne su definicije bitne kako bi se bolje shvatili problemi koji se javljaju prilikom održavanja elektroničkoga gradiva. Naime, izvor problema nalazi se u tome što će gradivo u klasičnom obliku biti onoliko dugovječno koliko je to medij od kojeg je gradivo načinjeno, i ako se ne poduzme nikakav poseban postupak arhiviranja, gradivo će i nakon 100 godina biti čitljivo, naravno pod uvjetom da se nije nalazilo u uvjetima koji bi na njega djelovali deteriorirajuće. S druge strane, ako se u uvjetima neprestanog razvoja informacijsko-komunikacijske tehnologije svjesno ne poduzmu koraci za očuvanje, čitljivost elektroničkoga gradiva postat će upitna već i nakon 10-ak godina. Upravo zbog ove razlike dolazi do promjene paradigme očuvanja i održavanja gradiva. Stoga pristup očuvanju i održavanju elektroničkoga gradiva mora postati proaktivan kako bi se osigurala izvorna čitljivost kroz dulji vremenski period. Za ilustraciju problema do kojih dolazi prilikom primjene tradicionalnih metoda arhiviranja na elektroničko gradivo poslužit će sljedeći primjer.

Pretpostavimo da želimo pročitati neki dokument stvoren u elektroničkom obliku prije dvadeset godina. Ako je on zapisan na magnetskoj traci, prije svega moramo imati uređaj koji čita tu vrstu magnetskih traka. Pretpostavimo da posjedujemo takav čitač. Njega moramo spojiti na računalo koje danas koristimo. No, hoće li današnji operativni sustav automatski prilikom priključivanja prepoznati tako stari čitač? Vjerojatno ne. Dakle, trebamo nabaviti pokretački program (engl. *driver*). Na Internetu ćemo ga možda pronaći ako proizvođač tog hardvera još posluje. Ako ga je preuzela druga tvrtka, što nije rijedak slučaj, pitanje je hoće li se ta tvrtka brinuti za razvoj pokretačkih programa za stari hardver preuzete tvrtke namijenjenih novom softveru. Pretpostavimo da smo i to riješili te da je naš stari čitač proradio na suvremenom sustavu. Sljedeći je korak pokretanje programa za čitanje zapisa sa magnetskih traka. Takav program najvjerojatnije ne postoji u Windows verziji, već radi samo pod DOS-om. S tim programom možemo, također, imati problema, jer pitanje je hoće li on raditi u Windows prozoru ili će zahtijevati samostalan pristup magnetskoj traci i čvrstom disku. Morat ćemo ponovo pokrenuti računalo, i to u DOS modu (ako to instalirana verzija još uopće podržava) te ponovno pokrenuti program za čitanje. Sada na red dolazi pitanje kvalitete magnetske trake. Poznato je da magnetski zapis slabi s vremenom pa tako ni zapis na ispravno uskladištenim magnetskim trakama neće trajati niti približno toliko dugo kao zapis na papiru. Na kraju, nakon eventualno uspješnog prijenosa zapisa s trake na čvrsti disk, preostaje još samo jedan korak, a to je pokretanje programa kojim je zapis stvoren

kako bi mogli doći do sadržaja. Ako ne posjedujemo takav program, izvorni zapis ćemo morati konvertirati u neki format zapisa koji je čitljiv nekim programom koji imamo na računalu, prilikom čega postoji mogućnost da izvorno oblikovanje zapisa ne bude sačuvano, a da o autentičnosti i vjerodostojnosti i ne govorimo. Dakle, za čitanje starog zapisa potreban je stari hardver i stari softver koji treba raditi na novom hardveru i na novom operativnom sustavu, što je često vrlo teško spojivo.

Općenito, problemi koji se javljaju prilikom održavanja digitalnoga gradiva mogu se grupirati u sljedeće skupine:

- problem pregleda gradiva,
- problem kodiranja gradiva,
- problem međusobne povezanosti sadržaja,
- problem ovlasti za arhiviranje te
- problem konverzije zapisa.⁷²

7.2.1. Problem pregleda gradiva

Gradivu u klasičnom obliku ipak se može prepoznati barem struktura, bez obzira na to koliko ono bilo staro, te ako je riječ o tekstualnom gradivu, eventualno prepoznati jezik kojim je ono napisano. S druge strane, prilikom pregleda starijeg elektroničkoga gradiva može doći do problema zbog toga što program novije verzije ne može ispravno protumačiti stariji zapis, jer ne prepoznaje njegov format. Stoga će program ili javiti grešku prilikom otvaranja zapisa ili će se na zaslону pojaviti gradivo u neprepoznatljivom obliku. Problemi se također javljaju ako se gradivo nalazi na medijima starijeg formata ili ako je stvoreno na nekom starijem ili, pak, različitom operativnom sustavu.

7.2.2. Problem kodiranja gradiva

Problem kodiranja gradiva odnosi se prije svega na probleme dugovječnosti vezane uz komprimiranje. Uzmimo za primjer slikovno gradivo. Kao što je već rečeno, komprimiranje može biti bez gubitaka, pri čemu je komprimirana slika identična originalnoj, i s gubicima, pri čemu se, među ostalim, uklanja onaj dio nijansi boje koje ljudsko oko ne razaznaje. Gledajući s današnjeg stajališta to je

korisno jer se znatno smanjuje veličina zapisa, što omogućava znatno brži prijenos mrežom. Ipak, danas još nije poznato hoće li možda baš taj eliminirani dio slikovne informacije biti potreban nekim budućim aplikacijama za analizu i usporedbu slikovnog materijala.

Razvijene su i tehnologije kao što je Epsonova "Print Image Matching" tehnologija, "koja pamti podatke izgubljene pri komprimiranju u JPEG zapis te omogućuje kompatibilnim kolor pisačima da ih ispišu".⁷³ Riječ je zapravo o tehnologiji koja se ugrađuje u digitalne fotoaparate i printere. Postavke u fotoaparatu, poput onih za portret, pejzaž, makro snimanje, sport, panoramu i sl., omogućavaju da se poslije na printeru ispišu znatno bolje fotografije jer postavke boje, kontrasta i oštine bivaju zadržane i ispravno interpretirane. Iako nije riječ o vraćanju izgubljene informacije prilikom komprimiranja, ovaj postupak ipak poboljšava kvalitetu ispisanih fotografskih materijala, ali je također oblik dodatnog kodiranja gradiva.

No, treba naglasiti i da komprimiranje dodaje još jedan stupanj kompleksnosti, što dodatno utječe na problem pregleda gradiva. Primjer, nadalje, postaje drastičniji ako se primijeni neki od oblika zaštite gradiva, primjerice šifriranje ili digitalne omotnice. Stoga bi, kod primjene nekog oblika zaštite gradiva, trebalo osigurati dostupnost javnih ključeva kroz dulji vremenski period, što je opet zaseban problem.

7.2.3. Problem međusobne povezanosti sadržaja

Svaki digitalni dokument može biti zasebna cjelina, ali je čest slučaj da se ono što se korisniku doima kao jedinstvena cjelina zapravo sastoji od nekoliko dijelova. Tako, na primjer, ako korisnik pregledava slikovno gradivo u arhivu digitalnoga gradiva, slika se može nalaziti u jednom zapisu, a naslov i ostali metapodatci u drugom. Stoga takvu organizaciju gradiva treba vjerno preslikati na novi sustav kako bi se međusobne veze zadržale.

Ovaj je problem naročito izražen prilikom arhiviranja dokumenata koji su nastali u elektroničkom obliku. Naime, većina dokumenata koji se danas stvaraju u procesu modernog poslovanja nastaju u elektroničkom obliku. Gotovo da nema novog dokumenta koji je pisan rukom ili pisačim strojem. Ti se dokumenti uglavnom otisnu na papir, ali uvijek negdje postoji njihova elektronička

⁷² Besser, Howard, Digital Longevity, u: *Handbook for Digital Projects*, n. dj., str. 165.

⁷³ Računala & tehnologija i telekomunikacije, prilog u: *Večernji list*, 11. ožujka 2001., str. 32, članak "Za ljepše boje".

verzija. Elektroničkim dokumentima se, s obzirom na to da se uglavnom radi o dokumentima koji nisu međusobno povezani, relativno jednostavno upravlja koristeći danas već dobro razvijene sustave za upravljanje elektroničkim dokumentima (engl. *EDMS – Electronic Document Management System*). Veći su problem elektroničke poruke, koje često u privitku sadrže sliku, zvuk ili video zapis, i Web stranice koje su međusobno povezane, te je vrlo teško osigurati izvornost gradiva s obzirom na to da se stranice na koje se neka stranica poziva mogu ugasiti. Smatram da su pitanja koja se postavljaju prilikom arhiviranja elektroničkih poruka i Web stranica ključna za razvoj moderne arhivistike pa stoga zahtijevaju pomniju razradu.

7.2.3.1. Arhiviranje elektroničkih poruka

Na početku je potrebno naglasiti da se *elektronička poruka* definira kao bilo koja komunikacija putem računala koja može biti u formi elektroničke pošte, elektroničke izmjene dokumenata (e-faks), elektroničke izmjene podataka i multimedijske komunikacije, primjerice glasovne pošte, video pošte te telekonferencija i videokonferencija.⁷⁴

Elektroničke poruke su način poslovne komunikacije baš kao što su to telefon, pošta ili telefaks. Stoga ih je potrebno spremati u sustav za arhiviranje digitalnih dokumenata. Takvo bi arhiviranje trebalo biti definirano samim računalnim sustavom tvrtke koji bi trebao taj postupak učiniti korisniku transparentnim. Dakle, s obzirom na to da su elektroničke poruke dokaz poslovne aktivnosti, funkcija arhiviranja mora biti ugrađena kao automatizam u računalni sustav ili elektroničke poruke moraju, zajedno s kontekstom, biti otisnute na papiru i arhivirane na klasičan način. Posljednja mogućnost u praksi baš i nije potpuno izvediva jer se problem javlja s multimedijским porukama ili razmjenom hipermedijских podataka. Prema tumačenju Nacionalnih arhiva Australije, poruke spremljene u računalu, dakle u okolini koja nije dio sustava za arhiviranje digitalnih dokumenata, također su zapis ili dio zapisa, mogu biti predmet policijske istrage i subpoene te trebaju dozvolu da bi bile obrisane.

⁷⁴ *Managing Electronic Messages as Records*, National Archives of Australia, <http://pandora.nla.gov.au/pan/22371/20011105-0000/www.naa.gov.au/recordkeeping/er/elec_messages/contents.html>, 6. svibnja 2006.

Elektroničke se poruke stoga mogu podijeliti u tri grupe:

- poslovne – kao što su odobrenje za neki poslovni potez, poslovna komunikacija između tvrtki,
- informativne – imaju poslovni kontekst, ali nisu dio poslovanja, na primjer obavijest o sastanku, te
- privatne – nemaju poslovni kontekst.

Poslovne poruke moraju biti arhivirane, dok informativne i privatne poruke uglavnom ne moraju.

Postoji tendencija u svijetu da se elektroničke poruke počinju prihvaćati na sudu kao dokaz. Ipak, sudovi rado prihvaćaju zapise nastale klasičnim metodama, dok elektroničke zapise još uvijek gledaju s nepovjerenjem jer uglavnom ne postoji standardizirana upotreba razrađenih normi i definiranih postupaka arhiviranja koji bi zapise činili vjerodostojnima na sudu. Stoga bi tvrtke i vladine institucije trebale pri uvođenju sustava za arhiviranje elektroničkih dokumenata napose paziti kako osigurati da elektronički zapisi imaju arhiviran sadržaj, strukturu i kontekst, što bi ih činilo pravovaljanim dokazom na sudu. Takav sustav arhiviranja bi zapravo trebao biti integriran u sustave za upravljanje elektroničkim dokumentima (engl. *Electronic Document Management System – EDMS*), sustave za upravljanje elektroničkim zapisima (spisovodstveni sustavi ili engl. *Electronic Records Management System – ERMS*) ili sustave za upravljanje poslovanjem na razini cijele institucije (engl. *Enterprise Management System*). Danas su takvi sustavi često u upotrebi u većim institucijama, a poznatiji su (abecednim redom) Documentum, Meridio, SAP, SharePoint i mnogi drugi.

Kako bi se olakšao postupak takvog oblikovanja i klasificiranja elektroničkih poruka, sustav može biti oblikovan tako:

- da korisnik definira odrednice elektroničke oмотnice prema kojoj se poruka može identificirati bez da se otvara,
- da korisnik popunjava određene podatke o poruci bez kojih poruka ne može biti odaslana,
- da su korisnicima ponuđeni određeni predlošci koje mogu koristiti za pisanje elektroničkih poruka. Tako im, na primjer, mogu biti ponuđeni predlošci za poslovnu poruku, zakazivanje sastanka i sl. prema kojima se onda poruke klasificiraju u sustavu za arhiviranje.⁷⁵

⁷⁵ *Managing Electronic Messages as Records*, n. dj.

Uporabom sustava oblikovanog prema jednom od navedenih primjera zapisi elektroničkih poruka se standardiziraju, što rezultira znatnim povećanjem mogućnosti njihova prihvaćanja kao dokaznog materijala na sudu.

7.2.3.2. Arhiviranje internetskih stranica

Mnoge organizacije i vladine institucije održavaju svoje javne Web stranice. Koriste ih za pružanje informacija, komunikaciju i međusobnu suradnju. Isto tako koriste i druga rješenja za mrežnu komunikaciju, kao što su intranet, ektranet, virtualne privatne mreže (VPN) te osobne Web stranice. Na takvim se Web stranicama može naći sve više službenih dokumenata koji nemaju svoje papirnato izdanje, a činjenica je da se s takvim zapisima još uvijek ne postupa kao s tradicionalnim dokumentima u smislu arhiviranja. Oni su, naime, konceptualni problem. Pred vladine se institucije, jer one bi trebale predvoditi u definiranju rješenja, postavljaju ključna pitanja, kao što su:

- Jesu li Web stranice (engl. *Web page*) kao cjelina, unutar jednog Web mjesta (engl. *Web site*), publikacija ili zapis?
- Ako je Web mjesto zapis, kako ga se može unijeti u sustav za arhiviranje elektroničkih dokumenata, opisati i sačuvati te koliko ga je dugo potrebno držati pohranjenog?
- Treba li pratiti promjene na Web stranicama koje su određene za arhiviranje, koliko često promjene treba arhivirati te treba li sačuvati i prethodne verzije kako bi se poslije mogao analizirati razvoj pojedinih Web mjesta?
- Postoje li tehnička rješenja za upravljanje arhiviranjem Web mjesta i ažuriranjem njihovih promjena u sustavu za arhiviranje?
- Kako upravljati zapisima ako postoje jedino na Web stranicama?⁷⁶

Očigledno je da se sva ova pitanja ne mogu objediniti jednim pravilom!⁷⁷

⁷⁶ *Archiving Web Resources: A Policy for Keeping Records of Web-based Activity in the Commonwealth Government*, National Archives of Australia, siječanj 2001., <http://www.naa.gov.au/Images/archweb_policy_tcm2-902.pdf>, 6. svibnja 2007.

⁷⁷ U Hrvatskoj se dvije institucije istaknutije bave arhiviranjem sadržaja mrežnih stranica. To su Nacionalna i sveučilišna knjižnica u okviru projekta DAMP (Digitalni arhiv mrežnih publikacija) koji je dostupan na <http://www.nsk.hr/digarhiv> i <http://damp.nsk.hr/> te Hrvatska informacijsko dokumentacijska referalna agencija (HIDRA) koja prikuplja gradivo "čiji su autori i/ili nakladnici tijela javne vlasti (...) i trajno ga pohranjuje u DAMIR (Digitalni arhiv mrežnih izvora Republike Hrvatske)", a koja je dostupna na <http://www.hidra.hr/>.

7.2.4. Problem ovlasti za arhiviranje

Kad je riječ o gradivu u klasičnom obliku, uglavnom su određene ovlasti za arhiviranje pojedinih vrsta gradiva. Problem se javlja kod gradiva koje je nastalo u digitalnom obliku gdje ovlasti još uvijek nisu definirane. Na primjeru Australije i preporuka Nacionalnih arhiva Australije – NAA (engl. *NAA – National Archives of Australia*) i Nacionalne knjižnice Australije – NKA (engl. *NLA – National Library of Australia*) za rješavanje pitanja ovlasti ukazat će na kompleksnost cijelog problema.

Kao što je to i s drugim elektroničkim zapisima, očuvanje zapisa Web stranica zajednička je odgovornost obiju ustanova. Nacionalni arhivi Australije i Nacionalna knjižnica Australije zajedno su odgovorni za očuvanje Web sadržaja od nacionalne arhivske vrijednosti.

Nacionalni arhivi Australije osiguravaju ne samo načela oblikovanja, već i arhiviranje (Web) zapisa Australije koji zadovoljavaju određene poslovne potrebe, zahtjeve za odgovornošću i javna očekivanja te su ih dužni čuvati sve dok ti zapisi imaju vrijednost.

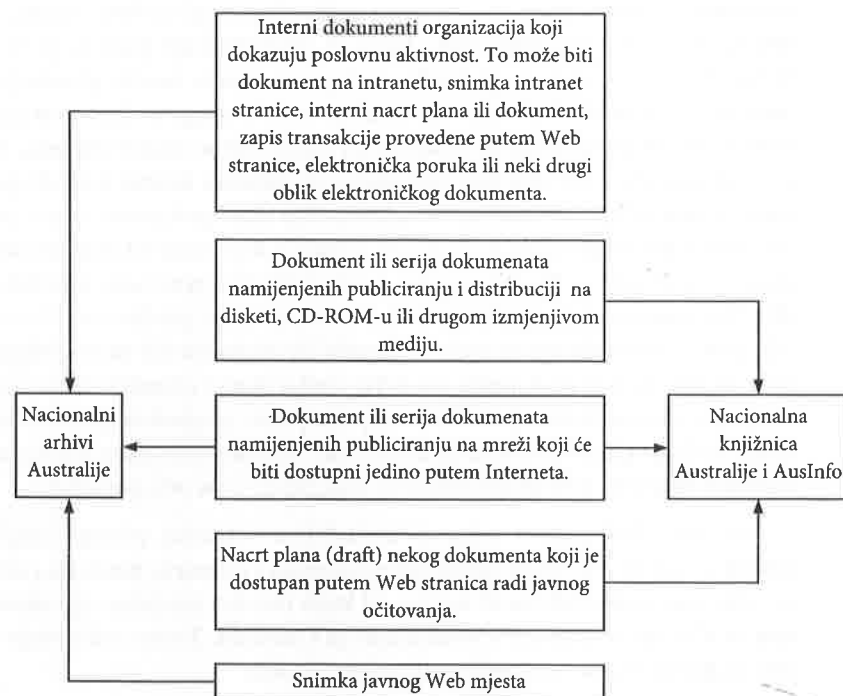
Nacionalna knjižnica Australije je 1996. godine pokrenula projekt PANDORA (engl. *Preserving and Accessing Networked Documentary Resources of Australia*), kojem je zadatak očuvati i omogućiti pristup umreženim izvorima dokumenata u Australiji. Danas se on odvija u suorganizaciji s devet drugih australskih knjižnica i institucija koje čuvaju kulturnu baštinu.

Za neke su sadržaje ovlašteni NAA, za neke NKA, a postoje sadržaji za koje se ovlasti preklapaju. Dijagram 10 prikazuje ovlasti NAA i NKA za arhiviranje elektroničkih sadržaja.

Nacionalni arhivi Australije ovlašteni su, među ostalim, za određivanje načina arhiviranja digitalnih dokumenata vladinih institucija. Te institucije održavaju svoje Web stranice koje imaju važnu ulogu u posredovanju između Vlade Australije i australskih građana. One su izravno odgovorne za sadržaj objavljen na njihovim Web stranicama te su dužne sačuvati njihov sadržaj, strukturu i kontekst u sustavu za arhiviranje digitalnih dokumenata kako bi ih na zahtjev mogle podastrijeti kao dokaz na sudu.

S obzirom na javni karakter Web mjesta vladinih institucija Nacionalni arhivi Australije smatraju ih publikacijama. Prema tome, kao što su papirnate publikacije zapisi koje je potrebno arhivirati, tako su i Web mjesta također zapisi. Stoga ih je potrebno u cijelosti sačuvati kako bi se u bilo kojem trenutku mogao

rekonstruirati sadržaj, struktura i kontekst svakog pojedinog Web mjesta u bilo kojem trenutku njegova postojanja.



Dijagram 10: Ovlasti Nacionalnih arhiva Australije i Nacionalne knjižnice Australije za arhiviranje digitalnih sadržaja

Iznimka iz ovog pravila su Web portali. Oni su sačinjeni od mnoštva veza na druge Web stranice, prije svega one koje nisu fizički smještene pod istom domenom, te uglavnom nemaju ili imaju vrlo malo dodatnog sadržaja. Kao takvi, Web portali nemaju veliku arhivističku vrijednost, ali ih je ipak potrebno pohraniti kao elektroničke zapise sve dok aktivno djeluju i dok postoje aktivne veze prema njima.

Preporuke Nacionalnih arhiva Australije i Nacionalne knjižnice Australije ne definiraju vremenski rok koji treba proći između dva arhiviranja istog Web

mjesta, već ostavljaju institucijama da to same procijene ovisno o učestalosti i opsegu promjena.

S obzirom na karakter elektroničkih dokumenata i potrebu za očuvanjem njihove pravne vjerodostojnosti, njihovo arhiviranje mnogo je složenije i tehnički zahtjevnije nego kod klasične građe. Neprestane tehnološke inovacije uzrokuju česte promjene u tehnikama upravljanja arhiviranim digitalnim zapisima. Novi oblici mrežne komunikacije (napose elektroničke poruke i Web mjesta) zahtijevaju konceptualizaciju sustava za njihovo arhiviranje. Općenito, potreban je sustav arhivskih institucija koje bi se proaktivno brinule za ujednačavanje tehnike arhiviranja i njezina razvoja.

7.2.5. Problem konverzije zapisa

Prilikom konverzije digitalnog zapisa iz jednog u drugi format dolazi do problema na razini interpretacije, dok fizička razina zapisa (binarni zapis) ostaje identična. Konverzija zapisa provodi se onda kada se ne može osigurati postojanost aplikacijske okoline u kojoj je zapis stvoren, što se događa zbog neprestanog razvoja novih verzija programa koji, pak, moraju raditi na novim verzijama operativnih sustava. Tako, na primjer, kada se konvertira tekstualni dokument stvoren u starijem programu za obradu teksta WordStar u MS Word 2000, vjerojatno će se tekst dokumenta prenijeti, ali je pitanje hoće li se ispravno prenijeti i sve ostalo oblikovanje dokumenta, kao što su oznake za oblik slova (masna, kurziv, podcrtana), podnožne bilješke itd. Kasnije verzije MS Word programa nisu niti imale mogućnost konvertiranja tekstova iz tako starog (dvadesetak godina) programa. Treba naglasiti da je pritom riječ o jednom od najrasprostranjenijih programa za obradu teksta današnjice. No, pitanje je kako stvari stoje u slučajevima elektroničkoga gradiva za čije su pregledavanje i korištenje potrebni neki manje poznati tipovi programa. Upravo je zato vrlo bitno da se gradivo koje nastaje u elektroničkom obliku odmah oblikuje tako da se uz njega zapišu i neki dijelovi aplikacijske okoline, kao što su metapodatci i podatci o strukturi, kako bi se naknadno, po potrebi, zapis mogao lakše rekonstruirati.

7.3. PRIJEDLOG RJEŠENJA

Problemi koji su objašnjeni u prethodnom odjeljku rješavaju se postupkom migracije.⁷⁸ Zapisi moraju biti migrirani nakon svakog značajnijeg unapređenja hardvera ili softvera na kojem su stvoreni, i to tako da zadrže svoju funkcionalnost i integritet. Naime, bez obzira na procijenjenu trajnost medija i zapisa na njemu, medij ili softver za čitanje zapisa vrlo brzo zastarijevaju zbog, takoreći, svakodnevnog razvoja novih medija sve većeg kapaciteta ili već spomenutog razvoja novih verzija programa i operativnih sustava. U takvoj okolini migracija služi kako bi se starim zapisima moglo i dalje pristupiti te kako bi oni ostali *vjerodostojni, potpuni, autentični* i imali sačuvano dovoljno *konteksta*.

Vjerodostojan je onaj zapis koji dolazi iz pouzdanog izvora. Vjerodostojnost je zapravo karakteristika izvedena iz činjenice da je određeni zapis stvorila institucija ili osoba od autoriteta, te ne proizlazi iz individualnih karakteristika samog zapisa.

Potpun je onaj zapis koji ima tako oblikovane karakteristike da im je pridruženo vrijeme i mjesto nastajanja zapisa, detalji o očekivanom korisniku, naslov, predmet te, naravno, sadržaj.

Autentičan je onaj zapis koji ima očuvanu povijest nastanka, prijenosa, korištenja i očuvanja kroz vrijeme. Pitanje autentičnosti u elektroničkoj okolini u kojoj se bilo koji zapis može nebrojeno puta kopirati iz prethodne kopije, a svaka nova kopija je identična po sadržaju i kvaliteti originalu, mora se promatrati na drugačiji način od autentičnosti na kakvu smo navikli kod papirnatih zapisa. Naime, u elektroničkoj okolini *originalom* se smatra zapis čiji sadržaj, struktura i kontekst odgovaraju izvornom zapisu. Naravno, prilikom migracije nedvojbeno će doći do određenih gubitaka, ali je bitno da se sačuvaju one karakteristike zapisa koje ga čine dokazom transakcija koje opisuje.

Kontekst se odnosi na međusobne veze pojedinih zapisa te okolinu u kojoj je zapis stvoren. Bez konteksta autentičnost bi bila ugrožena.⁷⁹

⁷⁸ Dodatni postupci koji se također mogu primjenjivati su vrlo složeni i u ovoj se raspravi neće zalaziti u njihovu analizu. Više informacija o njima može se pronaći u Thibodeau, Kenneth, Overview of Technological Approaches to Digital Preservation and Challenges in Coming Years, u: *The State of Digital Preservation: An International Perspective*, Council on Library and Information Resources (CLIR), Washington, D.C., SAD, srpanj 2002., str. 4-31.

⁷⁹ *Managing Electronic Records*, The Australian Archives, 1997., <http://pandora.nla.gov.au/pan/52389/20050913-0000/www.naa.gov.au/recordkeeping/er/manage_er/contents.html>, 6. svibnja 2006.

Jedino pitanje koje ostaje neodgovoreno jest što točno znači “nakon svakog značajnijeg unapređenja hardvera i softvera”. Svaka institucija će ipak morati samostalno o tome donijeti odluku i procijeniti kada je došao taj trenutak, jer su tehničko-vremenski uvjeti za migraciju u svakoj instituciji različiti. Naime, neka institucija može posjedovati veliku količinu elektroničkoga gradiva, pa će njoj trebati i više vremena da ga fizički prebaci na medije novije generacije, na primjer sa CD na DVD medije. Istovremeno će neka druga institucija sa znatno manje elektroničkoga gradiva trebati za isti posao znatno manje vremena. Stoga prva spomenuta institucija mora započeti s postupkom znatno prije druge institucije, dok druga, teoretski, može čekati zadnji tren i procijeniti treba li prelaziti, u navedenom primjeru, s CD na DVD medije ili će moći još malo pričekati i onda sve prebaciti na Blu-Ray medije, preskačući jednu generaciju optičkih medija. Naravno, ona si to može dopustiti, jer nema mnogo gradiva na CD medijima. Ipak, u ovom slučaju ne treba zaboraviti da migracija nije samo prebacivanje s medija na neki novi medij, već treba uzeti u obzir i pitanje razvoja programa i operativnih sustava potrebnih za pristupanje, pregledavanje i korištenje elektroničkoga gradiva, čime je cijela procjena trenutka kada treba krenuti s postupkom migracije ipak znatno složenija nego što se to na prvi pogled čini.

Migracija, dakle, mora osigurati korisnicima nesmetan pristup, pregled i korištenje zapisa u uvjetima neprestanog napretka i promjene hardvera i softvera. Migracija se provodi na tri načina, od kojih načelno niti jedan nije dovoljan sam za sebe, već se preporuča kombinirati ga s ostalima. To su osvježavanje medija, migracija zapisa i emulacija aplikacijske okoline.

7.3.1. Osvježavanje medija

S obzirom na to da su mediji koji se koriste za pohranu digitalnoga gradiva podložni fizičkom propadanju prilikom dugotrajnog korištenja, tj. kemijskom propadanju, u slučaju optičkih diskova uslijed kemijske nestabilnosti sloja diska u koji se laserom upisuju podatci, zapise je potrebno periodički osvježiti (engl. *refresh*) na način da ih se:

- kopira na novi medij iste tehnologije (na primjer, s CD-a na CD) zbog istrošenosti medija ili
- kopira na novi medij naprednije tehnologije (na primjer, s CD-a na DVD) zbog uspostave novog standarda.

Ručno osvježavanje zapisa je isplativo ako se radi o malim količinama medija. No, ako je riječ o velikom sustavu, tada je vjerojatno da je u njega implementiran neki oblik robotskog sustava za prihvata medija kao poluizravni sustav za pohranu i prijenos podataka. Robotski sustav maksimalno pojednostavljuje proces osvježavanja medija. Takav sustav može obavljati automatsko osvježavanje po zadovoljenju uvjeta aktiviranja poput premašenog određenog broja čitanja medija nakon čega se, prema statistici, može pretpostaviti da može doći do fizičkog oštećenja medija, premašenog pretpostavljenog roka trajanja medija ili, pak, premašivanja određenog broja grešaka prilikom čitanja zapisa s medija. Sustav čuva podatke koji su bitni za zadovoljavanje uvjeta aktiviranja. S obzirom na to da je sustav potpuno automatiziran i radi bez prekida, može ga se programirati tako da obavlja osvježavanje u vrijeme kada nema mnogo zahtjeva za čitanjem digitalne građe. Trošak osvježavanja kopiranjem na novi medij iste tehnologije robotom je zapravo trošak zamjenskog medija. Potrebna količina zamjenskih medija može se izraziti sljedećom formulom:

$$\text{broj medija koje treba godišnje zamijeniti} = \frac{\text{ukupni broj korištenih medija}}{\text{prosječni vijek trajanja medija}}$$

Kada je riječ o kopiranju na novi medij naprednije tehnologije, ono je također isplativije ako već postoji implementiran robotski sustav na koji se može spojiti modul koji je opremljen čitačem medija novije generacije. Kopiranje zapisa na naprednije medije znači smanjenje broja medija, jer su napredniji mediji u pravilu većeg kapaciteta. Prelazak na naprednije medije stoga znači znatno proširenje kapaciteta. Trošak takvog prijelaza primarno čini nabavka novog modula za čitanje. Iako postoji i trošak nabavke novih medija, on je manji od troška nabavke zamjenskih medija iste tehnološke razine ako se računa odnos cijene i kapaciteta.⁸⁰

U slučaju da cijeli sustav nije automatiziran, troškovima je potrebno dodati i troškove ljudskog rada.

7.3.2. Migracija zapisa

Migracija se može definirati kao dokumentirana konverzija kojom se mijenja fizički zapis uz očuvanje logičke strukture i sadržaja dokumenta.⁸¹ Migracija zapisa je, dakle, pristup koji podrazumijeva prijelaz iz jednog formata zapisa u drugi, noviji format zapisa. Primjer migracije tekstualnih zapisa bio bi migriranje teksta iz formata zapisa već spomenutoga programa Word Star u Word for Windows 2.0, pa zatim redom u Word for Windows 6.0, Word 95 (verzija 7.0), Word 97, Word 2000, Word 2002, Word 2003 te naposljetku u Word 2007 – najnoviju inačicu namijenjenu za rad na operativnom sustavu MS Vista. Prije početka migracije potrebno je:

- odlučiti koje zapise institucija treba sačuvati, a koji joj više nisu potrebni. Ovo napominjem zbog toga što pod digitalnim arhivom ne smatram samo arhive u posjedu institucija, kao što su arhivi, knjižnice i muzeji, gdje se gradivo digitalizira s namjerom, već općenitije, arhiv digitalnoga gradiva bilo koje institucije ili tvrtke gdje se arhiviraju i elementi elektoničkog poslovanja. Naravno da bi i neke od takvih dokumenata trebao preuzimati središnji arhiv, ali to nažalost u većini zemalja uglavnom još nije slučaj;
- pribaviti dopuštenje za brisanje ili uništenje nepotrebnih zapisa;
- odabrati softver za migraciju, tj. softver koji će funkcionirati kao novi standard unutar institucije;
- testirati migraciju zapisa prije nego što se krene u migraciju svih zapisa.

Nakon migracije zapisa potrebno je:

- provjeriti uspješnost postupka migracije prije nego što se krene s brisanjem starih zapisa ili uništenjem starih medija,
- provjeriti jesu li novim zapisima pridruženi svi potrebni metapodatci te
- provjeriti jesu li potpuni i ispravno zapisani podatci o cijelom procesu migracije zapisa.⁸²

Kao što je već i spomenuto, migraciju je potrebno provesti nakon svakog značajnijeg unapređenja hardvera ili softvera. Na svakoj instituciji je da sama odluči što za nju znači "značajnije unapređenje", ali bi migraciju svakako trebalo provoditi sa svakom promjenom generacije tehnologije kao standarda unutar

⁸⁰ Bell i Waugh, Digital Storage Media, n. dj.

⁸¹ Rubčić, Darko, Josip Šaban i Jozo Ivanović, *Vodič za arhiviranje dokumentacije u trgovačkim društvima i ustanovama*, Informator, Zagreb, 1999., str. 61-62.

⁸² *Managing Electronic Records*, Appendix 3 – Preserving Electronic Records through Migration, n. dj.

institucije, jer bi se preskakanjem generacije zbog, na primjer, uštede moglo dogoditi da ne bude moguće zadržati vjerodostojnost, potpunost, autentičnost i dovoljno konteksta migriranih zapisa.

7.3.3. Emulacija aplikacijske okoline

Emulatori – programi koji emuliraju neku aplikacijsku okolinu – ponekad su jedino rješenje za čitanje starih zapisa koji nisu na vrijeme migrirani. Ako postoje takvi zapisi, potrebno je izraditi ili, ako već postoji, koristiti emulator – program koji će imitirati ili operativni sustav na kojem je zapis stvoren ili aplikaciju koja je korištena za izradu tog zapisa. Dakle, ako aplikacija s kojom je izvorno stvoren zastarjeli zapis ne radi na novom operativnom sustavu, tada je potrebno koristiti emulator starijeg operativnog sustava i uz pomoć njega pokrenuti aplikaciju. Ako se, pak, aplikacija neispravno izvršava na novom hardveru, tada je potrebno koristiti emulator starijeg hardvera.

Kako bi se omogućilo da dokumenti koji se danas arhiviraju budu čitljivi na budućim operativnim sustavima koji će raditi na zasad nepoznatom budućem hardveru, potrebno ih je prirediti na takav način da se budućim korisnicima što je više moguće olakša buduće čitanje i rukovanje takvim dokumentima te izrada potrebnih emulatora. Zapis se tada sastoji od tri vrste podataka:

- izvornog dokumenta te podataka o aplikacijskoj okolini i operativnom sustavu,
- specifikacije emulatora hardvera kako bi se opisao izvorni prikaz dokumenta kad se aplikacija kojom je stvoren pokrene pod emulatorom te
- metapodataka o izvornom dokumentu, dokumentacije o korištenom hardveru i softveru te anotacijskih oznaka. Ovi bi podatci trebali biti zapisani u najjednostavnijem obliku, kao što je to najobičniji ASCII tekstualni dokument, kako bi ih budući korisnici mogli pročitati bez potrebe za starijim hardverom ili softverom. Ako dođe do znatnijih promjena u anotacijskom sustavu, tada je prilikom osvježavanja medija potrebno anotacijske oznake transliterirati u novi sustav oznaka.

Svi se ovi podatci zajedno zatvaraju u zamišljenu kapsulu (engl. *encapsulation*) te spremaju. Ovakav pristup ima nekoliko prednosti. Metapodatci koji su spremljeni zajedno s dokumentom omogućuju jednostavno upravljanje dokumentom (klasificiranje, kopiranje i distribuiranje) čak i kad je platforma na kojoj je on izrađen već zastarjela. U tom je slučaju emulacija potrebna jedino kad

se dokument želi pročitati. Nadalje, specifikacije za izradu emulatora zastarjelog hardvera potrebno je zapisati samo jednom, jer su zajedničke za sve dokumente čija je aplikacija koristila taj hardver. Naposljetku, emulator za pojedini zastarjeli hardver ili operativni sustav je potrebno izraditi samo jednom za svaku napredniju hardversko-operativno-softversku platformu. Kad je jednom izrađen, emulator se može koristiti na svim računalima iste, nove generacije. Stoga se proces očuvanja digitalnog dokumenta u zamišljenoj kapsuli može podijeliti u četiri koraka:

1. pridružiti dokumentu anotacijske oznake,
2. zatvoriti dokument zajedno sa svim potrebnim podacima u zamišljenu kapsulu,
3. po potrebi transliterirati anotacijske oznake prilikom osvježavanja medija te
4. u budućnosti, otvoriti zamišljenu kapsulu, izraditi potreban emulator i pokrenuti ga na budućem računalu.

Ovaj proces omogućava čitanje izvornog dokumenta koristeći izvornu aplikaciju kojom je on stvoren na izvornom računalu, a sve to na računalu neke novije generacije.⁸³

Da zaključim, održavanje digitalnoga gradiva vrlo je bitan segment procesa digitalizacije iako ga se može promatrati i kao zaseban problem. Propisno održavanje osigurava dugovječnost zapisâ čime se ostvaruje njihova vjerodostojnost, potpunost, autentičnost i kontekst u kojem su stvoreni. S obzirom na neprestani razvoj, o očuvanju digitaliziranoga gradiva treba razmišljati u relativnim, a ne u apsolutnim okvirima, tj. kao o ciklusima pohrane podataka, jer vrijeme trajanja jednog medija jest zapravo vrijeme do novog ciklusa migracije podataka.

⁸³ Opis postupka emulacije prema: Rothenberg, Jeff, *Avoiding Technological Quicksand: Finding a Viable Technical Foundation for Digital Preservation*, Council on Library and Information Resources, Washington, SAD, 1998., <<http://www.clir.org/pubs/reports/rothenberg/contents.html>>, 28. listopada 2000.

Zaključak

Institucije se sve više oslanjaju na komunikaciju i djelovanje putem globalne informacijske infrastrukture. Sve više informacija biva pretraženo, zatraženo, preneseno i isporučeno u elektroničkom obliku. Postalo je svakodnevicom očekivati dostupnost informacija putem bilo kojeg računala spojenog na Internet. Upravo u tom kontekstu digitalizacija gradiva doista postaje područje koje ima vrlo velike perspektive za razvoj. Postupci digitalizacije mogu se primjenjivati u svim područjima društvenih djelatnosti gdje postoji veća količina dokumenata u digitalnom obliku, počevši od državne uprave, javnih institucija kao što su arhivi, knjižnice i muzeji, pa sve do institucija čije se poslovanje zasniva na donošenju kritičnih odluka zbog potrebe za stalnim prilagođavanjem brzim promjenama na tržištu.

Na postupke digitalizacije ne treba isključivo gledati kao na postupke koji će osigurati prisutnost nekog gradiva i u elektroničkoj okolini i njegovu bržu pretraživost. Mnogo je važnija činjenica, pogotovo kad je riječ o gradivu koje ima status nacionalne kulturne baštine, da je to gradivo dostupno i u elektroničkom obliku, te adekvatno prezentirano. Digitalizacija zapravo svima pruža mogućnost jednake prisutnosti na Internetu. Tako se digitalizacijom može djelovati i u svrhu očuvanja nacionalnog identiteta i dostojnog predstavljanja kulturne baštine neke nacije. Naravno, dobro osmišljenim projektom digitalizacije ujedno će nastati i zaštitne, arhivske kopije izvornog gradiva u elektroničkom obliku pa će se vrlo često smanjiti upotreba osjetljivih i vrijednih originala koji će moći biti nesmetano pohranjeni bez pretjeranog korištenja. Treba imati na umu da se danas bilo koje gradivo može digitalizirati – tekstualno, slikovno, zvučno, filmsko i video gradivo, ali isto tako i trodimenzionalni objekti. Ipak, nemoguće je osmisliti kvalitetan digitalizacijski projekt bez podjele njegovih aktivnosti u nekoliko koraka. Zbog toga je u ovoj knjizi proces digitalizacije podijeljen u sedam koraka: (1.) odabir gradiva kao pretpoces; (2.) digitalizacija gradiva (objašnjeni su svi aspekti problematike vezani uz tekstualno, slikovno, zvučno, video i 3D gradivo); (3.) obrada i kontrola kvalitete svakog tipa gradiva (objašnjene su i metode njihovoga komprimiranja); (4.) zaštita kao važan element za sprječava-

nje neovlaštenog pristupa, kopiranja i distribuiranja te dokazivanja autentičnosti gradiva; (5.) pohrana i prijenos (objašnjeno nekoliko glavnih sustava); (6.) pregled i korištenje (dan je pogled na rezultat procesa digitalizacije iz perspektive korisnika) te (7.) održavanje digitalnoga gradiva. Premda je održavanje digitalnoga gradiva navedeno kao posljednji korak procesa digitalizacije on to i jest i nije zbog činjenice da se proces digitalizacije doista može završiti bez tog koraka. Održavanje digitalnoga gradiva nije korak koji je potrebno učiniti kako bi gradivo bilo u potpunosti digitalizirano i pripremljeno za korištenje, no bez njega čitav posao oko digitalizacije ne bi imao smisla. Moderni pristup očuvanju elektroničkoga gradiva govori o očuvanju pristupa elektroničkom gradivu i očuvanju njegove funkcionalnosti. To je zbog toga jer se problem očuvanja gradiva sagledava sa stajališta očuvanja informacijskoga sadržaja, a za to je potrebno poznavati te adekvatno i proaktivno pristupiti problemima vezanima uz trajnost medija, prikladnost hardverskih i softverskih rješenja te migraciju zapisa.

Zbog svega navedenoga, potrebno je prilikom promišljanja o projektu digitalizacije uzeti u obzir sve navedene objašnjene aspekte digitalizacijskog postupka. Njih bi svatko, pojedinac ili institucija, trebao poznavati, ili ih barem biti svjestan, prije nego što krene u sam proces digitalizacije. Posebno treba naglasiti važnost osiguranja neprestanog dotoka financijskih sredstava ne samo za provedbu procesa digitalizacije, već i za buduće održavanje toga gradiva, proširenje, tj. poboljšanje sustava te edukaciju djelatnika. Najvažnije od svega je, ipak, u digitalizacijski projekt ući predano i s jasnom vizijom što se njime želi postići. Ova knjiga bi, stoga, trebala pomoći svima koji razmišljaju o digitalizaciji, nalaze se pred konkretnijim projektom digitalizacije ili su već u nekoj njegovoj (podmakloj) fazi, kako bi mogli što bolje osmisliti njegovo provođenje ili poboljšati njegovo odvijanje.

Popis priloga

Slike

Slika 1: Plošni skener.....	35
Slika 2: Skener s prosvjetlivačem.....	35
Slika 3: Digitalizacija dijapozitiva digitalnim fotoaparatom.....	36
Slika 4: Skener za mikrooblike s dodacima za mikrofilm u roli i mikrofiše.....	37
Slika 5: Rotacioni skener i bubanj na postolju za postavljanje gradiva.....	38
Slika 6: Reprografski skeneri za digitalizaciju gradiva velikog formata.....	38
Slika 7: Prijenosni 3D skener.....	40
Slika 8: Sustav za robotiziranu digitalizaciju uvezanog gradiva.....	41
Slika 9: Skener s uvlakačem.....	42
Slika 10: Pretpojačalo zvučnog signala (prednja i stražnja strana).....	52
Slika 11: Uređaj za digitalizaciju filmskih materijala.....	54
Slika 12: Digitalizirani botanički uzorak iz herbarija s pridodanim kontrolnim uzorkom s lijeve strane.....	59
Slika 36: Mreža točaka 10 x 10 kod crno bijele fotografije.....	62
Slika 14: RGB sustav boja.....	63
Slika 15: CMYK sustav boja.....	63
Slika 16: Slika u CMYK sustavu i separacije četiriju osnovnih boja.....	63
Slika 17: Usporedba RGB slike (lijevo) i boja koje su izvan CMYK ljestvice.....	63
Slika 18: Ljestvica boja kod CIELAB, RGB i CMYK sustava.....	63
Slika 19: Uzorkovani valni oblik govora.....	66
Slika 20: Uzorkovanje frekvencijama 1 kHz i 10 kHz (B).....	67
Slika 21: Odnos kvalitete i kompresije.....	79
Slika 22: Organizacija zaštite lokalne mreže (LAN) korištenjem vatrozida.....	97
Slika 23: Izrada digitalnog potpisa funkcijom raspršenja.....	102
Slika 24: Provjera autentičnosti i integriteta dokumenta korištenjem funkcije raspršenja.....	103
Slika 25: Digitalizirani detalj freske s pridodanim vidljivim krhkim vodenim žigom.....	107

Slika 26: Postupak dodavanja vodenog žiga.....	108
Slika 27: Odnos strukture podataka u kompanijama prema istraživanju IDC-a.....	114
Slika 28: Izravni sustavi za pohranu i prijenos podataka	116
Slika 29: Diskovni automati	117
Slika 30: Poluizravni sustavi za pohranu i prijenos podataka.....	118
Slika 31: Robotizirana modularna biblioteka magnetskih traka kao dio hijerarhijskog sustava za pohranu	119
Slika 32: Hijerarhijski sustav za pohranu i prijenos podataka.....	119
Slika 33: Neizravni sustavi za pohranu i prijenos podataka.....	126
Slika 34: Mrežna pohrana podataka	127
Slika 35: Mreža za pohranu	128

Tablice

Tablica 1: Tablica određivanja prioriteta za digitalizaciju	23
Tablica 2: Odnos dubine bita i broja nijansi	61
Tablica 3: Odnos veličina komprimiranih datoteka	80
Tablica 4: Omjer kompresije.....	81
Tablica 5: Komprimiranje nepromjenjivim brojem bitova	85
Tablica 6: Komprimiranje promjenjivim brojem bitova.....	86
Tablica 7: Veličina komprimiranog zapisa kod CBR komprimiranja	86
Tablica 8: Veličina komprimiranog zapisa kod VBR komprimiranja.....	86
Tablica 9: Referentne vrijednosti digitalizacije jednog sata zvučnog zapisa.....	89
Tablica 10: Usporedba izravnih, poluizravnih i hijerarhijskih sustava za pohranu.....	125
Tablica 11: Osnovni inkrementalni model izrade sigurnosnih kopija	133
Tablica 12: Složeni inkrementalni model izrade sigurnosnih kopija.....	133
Tablica 13: Model djed-otac-sin (prva dva mjeseca godišnjeg ciklusa)	134
Tablica 14: Izrada sigurnosnih kopija prema modelu hanojskih tornjeva.....	135

Grafikoni

Grafikon 1: Pregled procesa odabira gradiva za digitalizaciju.....	24
Grafikon 2: Odnos veličina komprimiranih datoteka	80
Grafikon 3: Odnos omjera kompresije	81
Grafikon 4: Veličina komprimiranog zapisa kod CBR komprimiranja.....	87
Grafikon 5: Veličina komprimiranog zapisa kod VBR komprimiranja.....	87
Grafikon 6: Omjer postignute kompresije kod CBR kompresije.....	88
Grafikon 7: Omjer postignute kompresije kod VBR kompresije.....	88

Dijagrami

Dijagram 1: Prijenos s nepromjenjivim brojem bitova	92
Dijagram 2: Prijenos s promjenjivim brojem bitova	93
Dijagram 3: Postupak šifriranja i dešifriranja upotrebom simetričnog ključa.....	99
Dijagram 4: Postupak šifriranja i dešifriranja upotrebom javnog ključa	100
Dijagram 5: Postupak digitalnog potpisivanja.....	102
Dijagram 6: Postupak uporabe digitalnih certifikata.....	105
Dijagram 7: Sustav digitalnih vodenih žigova	108
Dijagram 8: Sustav digitalnih vodenih žigova s umetanjem podataka o primatelju.....	109
Dijagram 9: Kombinacija metode šifriranja simetričnim ključem sa steganografskom metodom	110
Dijagram 10: Ovlasti Nacionalnih arhiva Australije i Nacionalne knjižnice Australije za arhiviranje digitalnih sadržaja.....	150

Matrice

Matrica 1: Redoslijed donošenja odluka o digitalizacijskom projektu prije postupka odabira gradiva	16
---	----

Literatura

Adobe Photoshop 5.0 Help

Archives Act, The, Australia, 1983., <http://www.austlii.edu.au/au/legis/cth/consol_act/aa198398/>, 10. veljače 1999.

Archiving and Managing Digital Information, Commission of the European Communities, 4. studenog 1997., <http://www.kfs.oeaw.ac.at/publications/1999_deutsch_remote_access_and_transfer_of_audio_recordings/d3x3a1.htm>, 29. rujna 2008.

Archiving Web Resources: A Policy for Keeping Records of Web-based Activity in the Commonwealth Government, National Archives of Australia, siječanj 2001., <http://www.naa.gov.au/Images/archweb_policy_tcm2-902.pdf>, 6. svibnja 2000.

Arms, William Y., Roger Adkins, Cassy Ammen, Allene Hayes, Collecting and Preserving the Web: The Minerva Prototype, *RLG DigiNews*, 15. travnja 2001., vol. 5, br. 2, <<http://worldcat.org/arcviewer/1/OCC/2007/08/08/0000070519/viewer/file1302.html>>, 24. travnja 2007.

Authenticity in a Digital Environment, Council on Library and Information Resources, Washington D.C., svibanj 2000., <<http://www.clir.org/pubs/reports/pub92/pub92.pdf>>, 21. listopada 2000.

Backup Basics: Reliability, Redundancy and Recovery, White Paper, Quantum Corporation, 2006., <http://www.datasecurityprovider.com.br/anexos/backup_basics.pdf>, 4. kolovoza 2008.

Batnožić, Slaven, Digitalizacija starih izdanja, *Vijenac*, godište VIII, br. 172, 21. rujna 2000., str. 47

Being Fluent with Information Technology, National Academy of Sciences, National Academy Press, Washington, D.C., SAD, 1999., <http://www.nap.edu/openbook.php?record_id=6482>, 15. srpnja 2000.

Beagrie, Neil, Daniel Greenstein, *A Strategic Policy Framework for Creating and Preserving Digital Collections*, Arts and Humanities Data Service, King's College, London, 14. srpnja 1998., <<http://ahds.ac.uk/strategic.pdf>>, 5. srpnja 2008.

Bell, Robert i Andrew Waugh, Digital Storage Media for VERS, u: *Standard for the Management of Electronic Records in the Victorian Government*, Specification 3 VERS Standard Electronic Record Format, prilog 4, Public Record Office Victoria, State of Victoria, Australija, travanj 2000., <<http://www.prov.vic.gov.au/vers/standard/version2.htm>>, 28. listopada 2008.

- Besser, Howard, *Digital Longevity*, u: *Handbook for Digital Projects: A Management Tool for Preservation and Access*, ur. Sitts, Maxine K., Northeast Document Conversion Center, Andover, Massachusetts, 2000., str. 164-176, <<http://www.nedcc.org/oldnedccsite/digital/dighome.htm>>, 18. studenog 2008.
- CIELAB Color Conversion, <<http://www.aols.com/colorite/cielabcolorconv1.html>>, 31. siječnja 2001.
- Clavel-Merrin, Genevieve, *NEDLIB List of Terms*, NEDLIB (Networked European Deposit Libraries) Consortium, Den Haag, prosinac 2000., <<http://nedlib.kb.nl/results/NEDLIBterms.pdf>>, 26. ožujka 2001.
- Cleveland, Gary, *Digital Libraries: Definitions, Issues and Challenges*, ožujak 1998., <<http://www.ifla.org/VI/5/op/udtop8/udtop8.htm>>, 29. rujna 1998.
- Colorado Digitization Project: General Guidelines for Scanning, Metadata Working Group, SAD, proljeće 1999., <<http://chnm.gmu.edu/digitalhistory/links/cached/chapter3/link3.45.CDPscanningguidelines.htm>>, 5. srpnja 2000.
- Converting RGB to CMYK, <<http://www.aols.com/colorite/convertingrgbcmyk1.html>>, 31. siječnja 2001.
- Cyganski, David, John A. Orr i Richard F. Vaz, *Information Technology: Inside and Outside*, Prentice-Hall, Inc., New Jersey, SAD, 2001.
- D'Amato, Donald P., *Requirements and Options for the Digitization of the Illustration Collections of the National Museum of Natural History*, ožujak 1996., <<http://web.archive.org/web/20071116070317/http://www.nmnh.si.edu/cris/techrpts/imagopts/contents.html>>, 25. siječnja 2009.
- Digital Dilemma, The Intellectual Property in the Information Age*, National Academy of Sciences, National Academy Press, Washington, D.C., SAD, 2000., <http://books.nap.edu/html/digital_dilemma/>, 15. srpnja 2000.
- Digital Recordkeeping. Guidelines for Creating, Managing and Preserving Digital Records*, National Archives of Australia, svibanj 2004., <http://www.naa.gov.au/Images/Digital-recordkeeping-guidelines_tcm2-920.pdf>, 6. svibnja 2000.
- Digital Watermarking for the Digital Age: Empowering Consumers and Protecting Copyrights*, DIGIMARC, 2005., <https://www.digimarc.com/resources/docs/dmrc_content_id.pdf>, 5. siječnja 2009.
- Digitization as a Means of Preservation?*, European Commission on Preservation and Access, Amsterdam, Nizozemska, listopad 1997., <<http://www.clir.org/pubs/reports/digpres/digpres.html>>, 10. lipnja 1999.
- Documentum*, <<http://www.emc.com/products/family/documentum-family.htm>>, 26. siječnja 2009.
- Dollar, Charles M., *Arhivistika i informacijske tehnologije. Utjecaj informacijske tehnologije na arhivsku teoriju i praksu*, Hrvatski državni arhiv, Zagreb, 1999.
- Feenstra, Bendert, *Standards for the Implementation of a Deposit System for Electronic Publications (DSEP)*, NEDLIB (Networked European Deposit Libraries) Consortium, Den Haag, 2000., <<http://nedlib.kb.nl/results/NEDLIBstandards.pdf>>, 26. ožujka 2001.
- Formati datoteka za pohranu i korištenje (radna verzija)*, Nacionalni projekt "Hrvatska kulturna baština", Ministarstvo kulture Republike Hrvatske, studeni 2007., <http://www.kultura.hr/hr/content/download/597/7937/file/smjernice_formati.pdf>, 27. kolovoza 2008.
- Funding Information and Communication Technology (ICT) in the Heritage Sector: Policy Recommendations to the Heritage Lottery Fund*, Humanities Advanced Technology and Information Institute, University of Glasgow, 2. veljače 2000. <<http://www.hatii.arts.gla.ac.uk/research/HLFICT/>>, 25. ožujka 2000.
- Gertz, Janet, *Vendor Relations*, u: *Handbook for Digital Projects: A Management Tool for Preservation and Access*, ur. Sitts, Maxine K., Northeast Document Conversion Center, Andover, Massachusetts, 2000., str. 150-163, <<http://www.nedcc.org/oldnedccsite/digital/dighome.htm>>, 18. studenog 2008.
- Gilheany, Steve, *Preserving Information Forever and a Call for Emulators*, izlaganje na: Digital Libraries Conference and Exhibition *The Digital Era: Implications, Challenges and Issues*, Singapore, 18.-20. ožujka 1998., <<http://www.archivebuilders.com/whitepapers/22010p.pdf>>, 19. lipnja 1999.
- Gilheany, Steve, *The Promise of DVDs for Digital Libraries*, *Research Libraries Group DigiNews*, 1998. Vol. 2, br. 3., <<http://worldcat.org/arcviewer/1/OCC/2007/08/08/0000070511/viewer/file585.html>>, <<http://www.archivebuilders.com/whitepapers/22012p.pdf>>, 19. lipnja 2007.
- Gilroy, Peter, *Herbarium Specimen Imaging Protocol*, Museum of Southwestern Biology, University of New Mexico, <<http://www.msb.unm.edu/herbarium/imaging.pdf>>, 5. siječnja 2009.
- Graham, Peter S., *Intellectual Preservation: Electronic Preservation of the Third Kind*, Commission on Preservation and Access, 27. travnja 1998., <<http://www.clir.org/pubs/reports/graham/intpres.html>>, 18. studenog 2000.
- Green, Ann, JoAnn Dionne, Martin Dennis, *Preserving the Whole*, The Digital Library Federation, Council on Library and Information Resources, Washington D.C., lipanj 1999., <<http://www.clir.org/pubs/reports/pub83/pub83.pdf>>, 27. siječnja 2001.
- Hakala, Juha, *Collecting and Preserving the Web: Developing and Testing the NEDLIB Harvester*, *RLG DigiNews*, 15. travnja 2001., vol. 5, br. 2, <<http://worldcat.org/arcviewer/1/OCC/2007/08/08/0000070519/viewer/file1302.html>>, 24. travnja 2007.
- Handbook for Digital Projects: A Management Tool for Preservation and Access*, ur. Sitts, Maxine K., Northeast Document Conversion Center, Andover, Massachusetts, 2000., <<http://www.nedcc.org/oldnedccsite/digital/dighome.htm>>, 18. studenog 2008.
- Hazen, Dan, Jeffrey Horrell, Jan Merrill-Oldham, *Selecting Research Collections for Digitization*, Council on Library and Information Resources, kolovoz 1998., <<http://www.clir.org/pubs/reports/hazen/pub74.html>>, 27. siječnja 2001.
- HD Rosetta Archival Preservation Technologies and Services*, Norsam Technologies, 1997., <<http://www.norsam.com/rosetta.html>>, 20. studenog 1998.

- Hedstrom, Margaret, *Digital preservation: a time bomb for Digital Libraries*, School of Information and Library Studies, University of Michigan, SAD, <<http://www.uky.edu/~kiernan/DL/hedstrom.html>>, 20. siječnja 2001.
- IBM Dictionary of Computing, McGraw Hill, New York, 1994.
- In *Desktop Publishing, what is the difference between RGB and CMYK?*, Indiana University Knowledge Base, 30. lipnja 2008., <<http://kb.indiana.edu/data/aeon.html>>, 31. listopada 2008.
- Issues in Digitization, A Report Prepared for the Washington State Library Council, Washington State Library, SAD, 5. siječnja 1999., <<http://web.archive.org/web/200101217220600/http://www.statelib.wa.gov/projects/Digitize/Digitization10.html>>, 20. listopada 2002.
- Jann, Lynn-George, *Digitization: A Literature Review and Summary of Technical Processes, Applications and Issues*, 10. svibnja 1996., <http://web.archive.org/web/20020819040732/http://www.library.ualberta.ca/library_html/libraries/law/digit1.html>, 19. kolovoza 2002.
- Jantz, Ronald i Giarlo, Michael J., Digital Preservation. Architecture and Technology for Trusted Digital Repositories, *D-Lib Magazine*, lipanj 2005., vol. 11, br. 6, <<http://www.dlib.org/dlib/june05/jantz/06jantz.html>>, 18. srpnja 2005.
- Kenney, Anne R., *Digital to Microfilm Conversion: A Demonstration Project 1994-1996*, Department of Preservation and Conservation, Cornell University, Ithaca, SAD, <<http://www.library.cornell.edu/preservation/com/comfin.html>>, 8. listopada 1998.
- Kirtas Technologies Inc., <<http://www.kirtastech.com/>>, 30. travnja 2007.
- Kiš, Miroslav, *Englesko-hrvatski i hrvatsko-engleski informatički rječnik*, Naklada Ljevak, Zagreb, 2000.
- Kuny, Terry, An Introduction to Digitization Technologies and Issues, *Network Notes*, br 14, National Library of Canada, 1. listopada 1995., <<http://epe.lac-bac.gc.ca/100/202/301/netnotes/netnotes-h/notes14.htm>>, 26. siječnja 2009.
- Lesk, Michael, Going Digital, *Scientific American*, 1997., br. 3, <http://www.sciamdigital.com/index.cfm?fa=Products.ViewIssue&ISSUEID_CHAR=00B8E369-1805-4A27-A331-9D727FEAC21>, uz pretplatu, 27. rujna 1998.
- Lynn, Stuart, *Preservation and Access Technology. The Relationship Between Digital and Other Media Conversion Processes: A Structured Glossary of Technical Terms*, Commission on Preservation and Access, Washington D. C., SAD, 1990., <<http://www.clir.org/pubs/reports/lynn/index.html>>, 8. listopada 1998.
- Managing Electronic Messages as Records, National Archives of Australia, <http://pandora.nla.gov.au/pan/22371/20011105-0000/www.naa.gov.au/recordkeeping/er/elec_messages/contents.html>, 6. svibnja 2006.
- Managing Electronic Records, The Australian Archives, 1997., <http://pandora.nla.gov.au/pan/52389/20050913-0000/www.naa.gov.au/recordkeeping/er/manage_er/contents.html>, 6. svibnja 2006.
- Meridio, <<http://www.meridio.com/>>, 26. siječnja 2009.

- Microsoft Servers – SharePoint Portal Server, <<http://www.microsoft.com/Sharepoint/default.aspx>>, 26. siječnja 2009.
- Model Requirements for the Management of Electronic Records – MoReq Specification, CECA-CEE-CEEA, Bruxelles- Luxembourg, 2001., <<http://www.cornwell.co.uk/moreqdocs/MoReq%20Specification%20v5-2.4.doc>>, 14. travnja 2007.
- Model Requirements for the Management of Electronic Records – MoReq2 Specification. Update and Extension, European Commission, Luxembourg, 2008., <http://www.cornwell.co.uk/moreq2/MoReq2_typeset_version.pdf>, 26. studenog 2008.
- Moving Picture Experts Group (MPEG), <<http://www.chiariglione.org/mpeg/>>, 5. veljače 2001.
- MPEG-2 Frequently Asked Questions, Berkeley Multimedia Research Center, <<http://bmrc.berkeley.edu/frame/research/mpeg/mpeg2faq.html>>, 18. veljače 2001.
- Network-attached storage, Wikipedia, 30. srpnja 2008., <http://en.wikipedia.org/wiki/Network-attached_storage>, 4. kolovoza 2008.
- Noerr, Peter, *The Digital Library Tool Kit*, Sun Microsystems, California, SAD, 1998., <http://typhon.perseus.tufts.edu/typhon/Flashy/Documents/Documents.2000/ACM%25_2FDL%202000/sundlib.pdf>, 13. veljače 2008.
- Position Paper on Copyright in the Electronic Environment, IFLA, kolovoz 1996., <<http://www.ifla.org/V/press/pr961002.htm>>, 29. rujna 1998.
- Pravilnik o korištenju arhivskoga gradiva, NN 67/99, <<http://zagreb.arhiv.hr/hr/standardi/fs-ovi/provedbeni-propisi.htm>>, 10. ožujka 2001.
- Preserving Digital Information, Report of the Task Force on Archiving of Digital Information, The Commission on Preservation and Access and The Research Libraries Group, Inc., 1. svibnja 1996., <<http://www.ifla.org/documents/libraries/net/tfadi-fr.pdf>>, 5. ožujka 2001.
- Prijedlog Nacionalnog programa digitalizacije arhivske, knjižnične i muzejske građe (v. 2.0, konačni tekst), Ministarstvo kulture Republike Hrvatske, 20. listopada 2006., <<http://daz.hr/bastina/NacionalniProgramDigitalizacije.pdf>>, 21. kolovoza 2007.
- Puglia, Steven, Technical Primer, u: *Handbook for Digital Projects: A Management Tool for Preservation and Access*, ur. Sitts, Maxine K., Northeast Document Conversion Center, Andover, Massachusetts, 2000., str. 93-112, <<http://www.nedcc.org/oldnedccsite/digital/dighome.htm>>, 18. studenog 2008.
- Računala & tehnologija i telekomunikacije, prilog u: *Večernji list*, 11. ožujka 2001., str. 32, članak "Za ljepše boje"
- Ross, Seamus, Maria Economou, Information and Communications Technology in the Cultural Sector. The Need for National Strategies, *D-Lib Magazine*, lipanj 1998., <<http://www.dlib.org/dlib/june98/06ross.html>>, 4. srpnja 2000.
- Rothenberg, Jeff, *Avoiding Technological Quicksand: Finding a Viable Technical Foundation for Digital Preservation*, Council on Library and Information Resources, Washington, SAD, 1998., <<http://www.clir.org/pubs/reports/rothenberg/contents.html>>, 28. listopada 2000.

- Rubčić, Darko, Josip Šaban i Jozo Ivanović, *Vodič za arhiviranje dokumentacije u trgovačkim društvima i ustanovama*, Informator, Zagreb, 1999.
- SAP, <<http://www.sap.com/croatia/solutions/index.epx>>, 26. siječnja 2009.
- Sikora, Thomas, *MPEG-1 and MPEG-2 Digital Video Coding Standards*, Image Processing Department, Heinrich-Hertz-Intitut, Berlin, <<http://citeseerx.ist.psu.edu/icons/pdf.gif?sessionid=D3F9CD92254BB840653093C84B5B8552>>, 4. travnja 2003.
- Smjernice za odabir građe za digitalizaciju (radna verzija)*, Nacionalni projekt "Hrvatska kulturna baština", Ministarstvo kulture Republike Hrvatske, studeni 2007., <http://www.kultura.hr/hr/content/download/596/7925/file/smjernice_odabir.pdf>, 27. kolovoza 2008.
- Specification of the Broadcast Wave Format. A format for audio data files in broadcasting*, Supplement 1 – MPEG audio, European Broadcasting Union, Ženeva, Švicarska, srpanj 1997., <<http://tech.ebu.ch/docs/tech/tech3285s1.pdf>>, 8. siječnja 2009.
- Stančić, Hrvoje, Arhiviranje digitalnih dokumenata, u: Willer, Mirna i Katić, Tinka (ur.), 4. seminar Arhivi, knjižnice i muzeji. *Mogućnosti suradnje u okruženju globalne informacijske infrastrukture*, Hrvatsko knjižničarsko društvo, Zagreb, 2001, str. 209-214.
- Stančić, Hrvoje, Digitalizacija građe, u: Willer, Mirna i Katić, Tinka (ur.), 2. i 3. seminar Arhivi, knjižnice, muzeji, *Mogućnosti suradnje u okruženju globalne informacijske infrastrukture*, Zbornik radova, Hrvatsko muzejsko društvo, Zagreb, 2000., str. 64-70.
- Standard for the Management of Electronic Records in the Victorian Government*, Public Record Office Victoria, State of Victoria, Australija, travanj 2000., <<http://www.prov.vic.gov.au/vers/standard/version2.htm>>, 17. rujna 2008.
- Strategija razvitka Republike Hrvatske "Hrvatska u 21. stoljeću"*, Informacijska i komunikacijska tehnologija, Vlada Republike Hrvatske, 16. svibnja 2002. <http://narodne-novine.nn.hr/clanci/sluzbeni/2002_09_109_1753.html>, 8. veljače 2003.
- Süsstrunk, Sabine, Robert Buckley, Steve Swen, *Standard RGB Color Spaces*, Proceedings of IS&T/SID's 8th Color Imaging Conference, 2000., <<http://infoscience.epfl.ch/record/34089/files/SusstrunkBS99.pdf>>, 31. siječnja 2006.
- Tangley, Laura, Whoops, There Goes Another CD-ROM, *U.S. News Online*, 16. veljače 1998., <http://www.usnews.com/usnews/culture/articles/980216/archive_003232.htm>, 25. studenog 2000.
- Vodič za upravljanje elektroničkim gradivom s arhivističkog stajališta*, Hrvatski državni arhiv, Zagreb, 1999. (prijevod izdanja *The Guide for Managing Electronic Records from an Archival Perspective*, International Council on Archives, Paris 1997.)
- Vogt-O'Connor, Diane, Care Of Archival Digital And Magnetic Media, *Conserve O Gram*, National Park Service, Department of the Interior, Washington D.C., USA, br. 19/20, rujan 1996., <<http://www.cr.nps.gov/museum/publications/consveogram/19-20.pdf>>, 29. siječnja 2001.
- Vogt-O'Connor, Diane, Reformating for Preservation and Access: Prioritizing Materials for Duplication, *Conserve O Gram*, National Park Service, Department of the Interior, Washington D.C., SAD, br. 19/10, srpanj 1995., <<http://www.nps.gov/history/museum/publications/consveogram/19-10.pdf>>, 29. siječnja 2001.
- Vogt-O'Connor, Diane, Selecting of Materials for Scanning, u: *Handbook for Digital Projects: A Management Tool for Preservation and Access*, ur. Sitts, Maxine K., Northeast Document Conversion Center, Andover, Massachusetts, 2000., str. 45-73, <<http://www.nedcc.org/oldnedccsite/digital/dighome.htm>>, 18. studenog 2008.
- Waters, Donald J., *Electronic Technologies and Preservation*, Commission on Preservation and Access, Washington D. C., SAD, <<http://www.clir.org/pubs/reports/waters/waters2.html>>, 8. listopada 1998.
- Whittle, Robin, *First Principles – Lossless Compression of Audio*, Centre for Signal Processing, Nanyang Technological University, Singapore, 17. siječnja 2001., <<http://www.firstpr.com.au/audiocomp/lossless/>>, 12. veljače 2001.
- Wikipedia, s.v. hash function, 10. rujna 2005. <http://en.wikipedia.org/wiki/Hash_function>, 13. rujna 2005.
- Wikipedia, s.v. Network-attached storage, 30. srpnja 2008., <http://en.wikipedia.org/wiki/Network-attached_storage>, 4. kolovoza 2008.
- Woodyard, Deborah, *Digital Preservation: The Australian Experience*, 3rd Digital Library Conference: Positioning the Fountain of Knowledge, State Library of Sarawak, Malaysia, 2.-4. listopada 2000., <<http://www.nla.gov.au/nla/staffpaper/dw001004.html>>, 20. siječnja 2001.
- Woodyard, Deborah, *Farewell my Floppy: A Strategy for Migration of Digital Information*, National Library of Australia, 20. siječnja 2001., <<http://www.nla.gov.au/nla/staffpaper/valadw.html>>, 20. siječnja 2001.
- Wright, Marilyn, Trendy (and Tried) Ways to Secure Your Knowledge, *Knowledge Management World*, Vol. 7, br. 12, 1998., <[http://www.kmworld.com/Articles/Editorial/Feature/Trendy-\(and-tried\)-ways-to-secure-your-knowledge-9045.aspx](http://www.kmworld.com/Articles/Editorial/Feature/Trendy-(and-tried)-ways-to-secure-your-knowledge-9045.aspx)>, 19. travnja 2001.
- Xing Audio Catalyst 2.10 Help
- Zagami, Bob, Legacy Data in the Engineering Environment: What It Is, and What You Can Do With It, *Document Management Magazine*, rujan/listopad 1998., <<http://www.docmanage.com/magazine/backissues/8-5/Zagami.htm>>, 11. siječnja 1999.
- Zakon o arhivskom gradivu i arhivima*, 29. rujna 1997., <<http://www.nn.hr/clanci/sluzbeno/1997/1617.htm>>, 10. ožujka 2001.

Radovi Zavoda za informacijske studije

Knjiga 1.

Informacijske znanosti i znanje

Knjiga 2.

M. Tudman

Obavijest i znanje

Knjiga 3.

A. Stipčević

Cenzura u knjižnicama

Knjiga 4.

S. Jelušić

Struktura i organizacija
knjižničnih sustava

Knjiga 5.

Obrada jezika i prikaz znanja

Knjiga 6.

I. Maroević

Uvod u muzeologiju

Knjiga 7.

T. Aparac-Gazivoda

Teorijske osnove knjižnične znanosti

Knjiga 8.

J. Lasić-Lazić

Znanje o znanju

Knjiga 9.

B. Tepeš

Računarska lingvistika

Knjiga 10.

Z. Dovedan

Formalni jezici: sintaksna analiza

Knjiga 11.

Zbornik radova "Težakovi dani"

Knjiga 12.

Modeli znanja i obrada
prirodnog jezika

Knjiga 13.

D. Kovačević, J. Lasić-Lazić,

J. Lovrinčević

Školska knjižnica – korak dalje

Knjiga 14.

Odabrana poglavlja
iz organizacije znanja

Knjiga 15.

Informacijske znanosti
u procesu promjena

Knjiga 16.

J. Lovrinčević, D. Kovačević,

J. Lasić-Lazić, M. Banek Zorica

Znanjem do znanja

Knjiga 17.

J. Lasić-Lazić, M. László, D. Boras

Informacijsko čitanje

Knjiga 18.

S. Špiranec, M. Banek Zorica

Informacijska pismenost:

Teorijski okvir i polazišta

Knjiga 19.

H. Stančić

Digitalizacija