

Digitalna transformacija ustanova u kulturi na primjeru nacionalnog projekta eKultura

Dančuo, Lena

Master's thesis / Diplomski rad

2023

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

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:131:850171>

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

Download date / Datum preuzimanja: **2024-05-17**



Repository / Repozitorij:

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



SVEUČILIŠTE U ZAGREBU
FILOZOFSKI FAKULTET
ODSJEK ZA INFORMACIJSKE I KOMUNIKACIJSKE ZNANOSTI
SMJER ARHIVISTIKA
Ak. god. 2022./2023.

Lena Dančuo

**Digitalna transformacija ustanova u kulturi na primjeru
nacionalnog projekta eKultura**

Diplomski rad

Mentor: prof. dr. sc. Hrvoje Stančić
Neposredni voditelj: dr. sc. Jozo Ivanović

Zagreb, lipanj 2023.

Izjava o akademskoj čestitosti

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

Sadržaj

1.	Uvod.....	1
2.	Digitalizacija kulturne baštine	2
2.1.	Digitalizacija, kulturna baština, digitalizacija kulturne baštine	2
2.2.	Specifičnosti digitalizacije knjižnične građe i arhivskog gradiva.....	2
2.3.	Koraci i postupci u procesu digitalizacije	4
2.3.1.	Odabir gradiva za digitalizaciju	4
2.3.2.	Izrada metapodataka	5
2.3.3.	Edukacija zaposlenika.....	6
2.3.4.	Priprema gradiva za digitalizaciju	6
2.3.5.	Snimanje – pretvorba gradiva iz analognog u digitalni oblik	7
2.3.6.	Kontrola kvalitete.....	11
2.3.7.	Vođenje dokumentacije	12
2.3.8.	Arhiviranje	12
2.3.9.	Objavljivanje.....	13
2.3.10.	Briga o trajnom očuvanju i dugoročno upravljanje	14
3.	Projekt eKultura	16
3.1.	Ciljevi.....	16
3.2.	Infrastruktura i alati.....	17
3.3.	Odabir gradiva za digitalizaciju	18
3.4.	Izrada metapodataka	21
3.5.	Edukacija zaposlenika.....	24
3.6.	Priprema gradiva za digitalizaciju	24
3.6.1.	Dnevnik pripreme gradiva za digitalizaciju	25
3.6.2.	Konkretni postupci u pripremi gradiva	27
3.7.	Snimanje	29

3.7.1.	Arhivska i prezentacijska datoteka	30
3.7.2.	Imenovanje datoteka	31
3.7.3.	Tehničke specifikacije, postupci i pohrana	31
3.8.	Kontrola kvalitete.....	32
3.8.1.	Stavke zastupljene u kontroli kvalitete	33
3.8.2.	Analiza učestalosti grešaka prilikom skeniranja na protočnom skeneru	34
3.9.	Arhiviranje	37
3.9.1.	Kriptografski sažetak i tablica za metapodatke	37
3.9.2.	Goobi – softverska aplikacija otvorenog koda za projekte digitalizacije	38
3.10.	Objavljivanje.....	39
3.11.	Briga o trajnom očuvanju i dugoročno upravljanje	40
4.	Kvalitativna i kvantitativna usporedna analiza koračnog i protočnog skenera na arhivskom gradivu	42
5.	Zaključak.....	47
6.	Literatura.....	48
7.	Popis ilustracija.....	51
7.1.	Popis slika	51
7.2.	Popis grafikona	51
	Sažetak	52
	Summary	53

1. Uvod

Digitalizacija kao proces pretvorbe fizičkih objekata u digitalni oblik aktualnija je nego ikad, stoga je za očekivati da se provodi i u hrvatskim baštinskim institucijama. U svjetskim se baštinskim institucijama ona sustavno provodi već dugi niz godina, a projektom eKultura zamišljeno je da se i u Hrvatskoj na taj način krene prakticirati.

U ovom se radu istražuju izazovi u digitalizaciji kulturne baštine u Republici Hrvatskoj u novije vrijeme, a glavni fokus rada jest integracija svih radnji unutar procesa digitalizacije kao najzahtjevnijeg mu aspekta. Taj je proces ukratko opisan, no veći dio rada obuhvaća analizu praktičnih iskustava u okviru sudjelovanja u projektu digitalizacije kulturne baštine eKultura, odnosno realizaciju teorije u praksi na konkretnom primjeru. Literatura korištena za potrebe pisanja ovog rada jest projektna dokumentacija, smjernice za digitalizaciju i norme, a tema je odabrana iz razloga što je istovremeno aktualna, ali i relevantna za informacijske i komunikacijske znanosti, točnije za arhivistiku i u Hrvatskoj i u svijetu. Dio rada u kojem se konkretno opisuje digitalizacija kulturne baštine u Hrvatskom državnom arhivu napisan je iz perspektive studentice arhivistike koja je sudjelovala na projektu eKultura. To je zapravo i svojevrsno istraživanje čiji je cilj opsežno iz prve ruke opisati sve korake digitalizacije te iznijeti vlastita nova saznanja s projekta i digitalizacije arhivskog gradiva, kao nekoga tko je s digitalizacijom bio upoznat isključivo na teorijskoj bazi. Usto je izvršena kvalitativna i kvantitativna usporedna analiza rada na protočnom i koračnom skeneru. Detaljno opisan proces digitalizacije kulturne baštine u arhivima od pripreme gradiva za digitalizaciju pa sve do savjesnog upravljanja i održavanja digitaliziranog gradiva, konkretni postupci, problemi, upravljanje projektom i praktično iskustvo rada na projektu u glavnoj arhivskoj instituciji u Hrvatskoj moći će poslužiti kao korisna referentna točka za sve buduće slične projekte.

2. Digitalizacija kulturne baštine

2.1. Digitalizacija, kulturna baština, digitalizacija kulturne baštine

Za početak valja pojasniti pojam digitalizacije. Hrvatska enciklopedija digitalizaciju u širem smislu riječi digitalizaciju definira kao „prevođenje analognoga signala u digitalni oblik“, a u užem smislu riječi podrazumijeva pretvorbu „teksta, slike, zvuka, pokretnih slika (filmova i videa) ili trodimenzijskog oblika nekog objekta u digitalni oblik, u pravilu binaran kôd zapisan kao računalna datoteka sa sažimanjem podataka ili bez sažimanja podataka, koji se može obrađivati, pohranjivati ili prenositi računalima i računalnim sustavima“. Osim toga, digitalizacija danas podrazumijeva i digitalizaciju modela poslovanja i poslovnih procesa.

Fokus ovoga rada jest digitalizacija kulturne baštine. Kako stoji na mrežnim stranicama Ministarstva kulture i medija Republike Hrvatske, ona je kolektivno bogatstvo čovječanstva upravo zbog svoje raznolikosti i posebnosti, a njezina zaštita je bitna da bi kulturni identitet bio potvrđen i prepoznat. Dakle, nedvojbeno je da je kulturna baština od neprocjenjive vrijednosti i samim tim je nužno da ostane očuvana i bude dostupna za javnost. Upravo to su dvije glavne stavke svrhe digitalizacije – očuvanje i dostupnost.

Kada govorimo o digitalizaciji kulturne baštine, iako postoji i 3D skeniranje, većina digitalizacije odnosi se na papirnatu građu kao što su knjige, karte, fotografije, razne tiskovine itd. Radi se o sveobuhvatnom trendu jer digitalizacija baštine ne samo da je štiti, nego je i popularizira među građanima kojima postaje dostupna u novom, možda i privlačnijem, obliku te jednostavnije nego ikad. Dakako, digitalizaciju treba i standardizirati te osigurati kvalitetu pa iz tog razloga postoje brojne smjernice, pravila, norme itd. koje taj proces olakšavaju. One pružaju i propisuju točan plan i slijed koraka, no isti proces mogu i otežati u slučaju nedostatka kompetentnog kadra i adekvatne opreme.

2.2. Specifičnosti digitalizacije knjižnične građe i arhivskog gradiva

Postupci u digitalizaciji pojedinih vrsta gradiva na papiru uglavnom su podudarni, jednako kao i osnovni tehnički zahtjevi i norme kojima se oni definiraju. Ipak, ima i određenih razlika.

S jedne strane je digitalizacija knjižnične građe koja se sve više i više provodi te se na taj način ona štiti – pogotovo kada se radi o vrlo staroj građi koja inače ne bi niti mogla biti dostupna javnosti – dok je njezina digitalizirana verzija dostupna svima i to čak u isto vrijeme. Među izazovima koji su naglašeni kod digitalizacije knjižne građe možemo spomenuti pitanje

autorskog i srodnih prava, dostupnost novije građe na mreži i uređivanje odnosa između javnih knjižnica i izdavača.

S druge su strane arhivi i digitalizacija arhivskog gradiva koja je kompleksnija iz tog razloga što uz brojne procese koji su prisutni u digitalizaciji knjižnične građe, za arhivsko je gradivo nužno osigurati njegovu „autentičnost, integritet, pouzdanost i iskoristivost“, kako stoji u Zakonu o arhivskom gradivu i arhivima NN 61/18, 98/19, 114/22, trećem članku, odjeljku n). Nužno je provoditi i neke dodatne kontrole kao što je to primjerice kontrola metapodataka koji su relevantni za očuvanje integriteta i autentičnosti dokumenata u obliku digitalnih kopija. Autentičnost podrazumijeva da je neki dokument upravo ono za što se tvrdi da on jest, dok integritet uvjetuje da je spis svojim sadržajem potpun te neizmijenjen, tako piše Ivanović (2010). Isti autor još napominje da dodatni zahtjevi za očuvanje autentičnosti i integriteta kod arhivskog gradiva proizlaze iz toga što su arhivski dokumenti često neposredno povezani sa stvaranjem, zaštitom i dokumentiranjem prava i obveza građana i pravnih osoba u pravnom sustavu te se ne može isključiti da će u pravnim ili upravnim postupcima biti potrebno dokazivati ili pobijati njihovu vjerodostojnost.

Integritet digitaliziranog gradiva može se utvrditi i pratiti na vrlo učinkovit način pomoću *hash* vrijednosti, odnosno kriptografskog sažetka, no o tome nešto kasnije. Spis se ISO 15489-1 normom karakterizira pouzdanim ako je njegov sadržaj potpun i točan prikaz aktivnosti, činjenice i transakcije koje on kao takav potvrđuje i na koje se može referirati prilikom svake sljedeće aktivnosti. Kada je riječ o iskoristivosti, tada spomenuta norma nalaže da se dotičan spis treba moći naći, „dohvatiti, predočiti i interpretirati“, a mora ga se moći i „identificirati u kontekstu širih poslovnih aktivnosti i funkcija“. Navedeni se kriteriji pri digitalizaciji arhivskog gradiva provjeravaju i dokumentiraju koristeći Obrazac za provjeru sukladnosti čiji je isječak prikazan na slici 1.

HRVATSKI DRŽAVNI ARHIV		Obrazac za provjeru sukladnosti pravila		
		Verzija: 1.0	Datum: 17.12.2020.	
Osnova	Zahtjev	Opis/Dokazi/Dokumenti	Nalaz	Ocjena
2.5. Osiguranje integriteta i validacija				
P25/1	M2.24 U sustavu moraju biti definirani informacijski objekti (paketi) koji se validiraju i ugrađeni mehanizam validacije. Validacija se može provoditi na različitim razinama: od pojedinačnog dokumenta do čitavog projekta. Validacijom se može smatrati svaki čin kojim se formalno ili neizravno potvrđuje da je određena cjelina, koju čine datoteke i pripadajući metapodaci, pregledana i da je u redu. Neizravna potvrda moguća je, primjerice, ako se dokumenti snimaju po zaprimanju i dalje proslijeđuju u radni proces.			
P25/2 i P25/3	M2.25 Za osiguranje integriteta gradiva koriste se pouzdane kriptografske metode. Obveza korištenja kriptografskih metoda (na primjer, vremenski žig, elektronički potpis) vezana je uz procjenu rizika od gubitka integriteta ili vjerodostojnosti podrijetla digitalnih kopija gradiva. Umjesto pouzdanih kriptografskih metoda mogu se koristiti i druga sredstva, no u tom slučaju treba pokazati da se sredstva koja se koriste mogu zadovoljavajućima.			
P25/2 i P25/3	M2.26 Za gradivo čiji se integritet štiti kvalificiranim elektroničkim potpisom predviđen je i koristi se mehanizam za očuvanje integriteta i po isteku valjanosti potpisa. Mogu se koristiti različite tehnike za očuvanje integriteta gradiva po isteku valjanosti kvalificiranog elektroničkog potpisa: naknadno potpisivanje prije isteka valjanosti potpisa, pohrana u sigurno i vjerodostojno okruženje prije isteka valjanosti potpisa.			
P25/2 i P25/3	M2.27 Za osiguranje integriteta gradiva koristi se kvalificirani elektronički potpis ili kvalificirani vremenski žig. Obveza korištenja kvalificiranog elektroničkog potpisa ili kvalificiranog vremenskog žiga vezana je uz procjenu rizika od gubitka integriteta ili vjerodostojnosti podrijetla digitalnih kopija gradiva. Umjesto ovih metoda za očuvanje integriteta gradiva mogu se koristiti i druga sredstva, no u tom slučaju treba pokazati da se sredstva koja se koriste mogu smatrati zadovoljavajućima.			
P25/1	M2.28 Sustav mora generirati podatke koji dokumentiraju validaciju informacijskog paketa. Ovi podaci mogu biti ugrađeni u informacijski paket, dnevnik aktivnosti u sustavu ili zapisani na drugi način.			
P25/3	M2.29 Provjera cjelovitosti i kvalitete pretvorbe gradiva obavlja se najmanje provjerom uzorka koji je utvrđen sukladno procjeni rizika. U sustavu mora biti određeno kako se utvrđuje uzorak za provjeru cjelovitosti i kvalitete pretvorbe gradiva, taj uzorak mora odgovarati utvrđenom riziku i ishodu prethodnih provjera, a ishodi provjera moraju biti dokumentirani.			
P25/2	M2.30 Ako je utvrđen visok rizik u pogledu očuvanja integriteta, podaci o validaciji moraju omogućiti provjeru integriteta informacijskog objekta. Podaci o validaciji mogu se generirati i pohraniti tako da omogućuju ne samo uvid u to tko je i kada validirao informacijski paket, nego i da se provjeri je li izvorni sadržaj informacijskog paketa ostao nepromijenjen od trenutka validacije.			
P25/2 i P25/3	M2.31 Ako je utvrđen vrlo visok rizik u pogledu očuvanja integriteta, provjera cjelovitosti i ispravnosti digitalizacije provodi se metodom dvostruke kontrole. Cjelovitost i ispravnost digitalizacije provjeravaju se u dvije različite točke u procesu i od strane dvije različite osobe, od kojih jedna nije sudjelovala u snimanju gradiva niti u obradi snimaka.			
P18/3	M2.32 Postupci u digitalizaciji moraju biti dokumentirani, tako da je moguće provjeravati tko je i kada digitalizirao gradivo, tko je obavio pojedine radnje i u kojem tehničkom i organizacijskom okruženju. Postupci u digitalizaciji mogu se dokumentirati pomoću metapodataka o snimkama, dnevnika aktivnosti ili na drugi način.			

Slika 1. Isječak iz Obrasca za provjeru sukladnosti.

Ustupio Hrvatski državni arhiv.

Iz navedenog je jasno da je digitalizacija arhivskog gradiva u dijelu koji se odnosi na očuvanje autentičnosti i integriteta dokumenata i spisa u obliku digitalnih kopija daleko kompleksnija od digitalizacije knjižnične građe upravo zbog obaveznog poštivanja četiriju spomenutih načela.

2.3. Koraci i postupci u procesu digitalizacije

2.3.1. Odabir gradiva za digitalizaciju

U skladu s ciljem i svrhom digitalizacije, odabire se gradivo za digitalizaciju što je ujedno i prvi korak procesa. Neki od kriterija u odabiru gradiva su: digitalizacija radi zaštite izvornika, radi poboljšanja dostupnosti gradiva, radi stvaranja novih usluga korisnicima, radi kompletiranja fonda ili zbirke, skeniranje na zahtjev. U obzir se trebaju uzeti djelatnost i zadaće institucije u kojoj se digitalizacija provodi te potrebe i želje korisnika njezinog gradiva. U

Smjernicama za digitalizaciju (2020) navode se još neki postupci koji bi se trebali uzeti u obzir pri odabiru gradiva za digitalizaciju: određivanje sadržaja i strukture digitalne zbirke koja se može sastojati od jednog ili pak više medija ili izgradnja zbirke koje sadrže digitalizirano i izvorno digitalno gradivo kao jedinstvenu cjelinu. Time se ponekad mogu smanjiti troškovi upravljanja gradivom i povećati vidljivost i iskoristivost izvorno digitalnog gradiva. Dalje je napomenuto da je još jedna vrlo bitna stavka zaštita autorskog i srodnih prava jer se takvo gradivo smije javno objaviti samo ako to njihov nositelj dopusti te se u tom slučaju utvrđuju prava korištenja i način pristupa djelu. Uglavnom se zbog jednostavnosti i manjeg troška digitalizira ono gradivo koje ne podliježe spomenutim pravima. U obzir zatim treba uzeti i opseg projekta, točnije predviđeno trajanje, raspoloživa financijska sredstva i okvirnu količinu. Primjerice, nije moguće uzeti preveliki fond ili prevelike fondove ako se to ne bi stiglo digitalizirati unutar vremenskog roka trajanja projekta. Pri odabiru je važna i vrsta gradiva, njegova veličina i količina te stanje (oštećenost i/ili osjetljivost). Sukladno tome se za digitalizaciju, na primjer, neće odabrati kartografska zbirka ako institucija u sklopu projekta nema skener za tako velike formate. Također se u tom slučaju neće skenirati osjetljivi i oštećeni papiri u protočnom skeneru koji ih uvlači.

2.3.2. Izrada metapodataka

Za odabrano gradivo se izrađuju metapodaci. Tri su glavne vrste metapodataka kako to opisuje University of Texas Libraries (2022), a to su opisni, administrativni i strukturalni. Opisni omogućuju pronalazak i identifikaciju te podrazumijevaju elemente kao što su naslov, autor i sadržaj. Administrativni služe za upravljanje zapisima i uključuju tehničke podatke, sva pridržana prava i uvjete korištenja, dok strukturalni opisuju veze između jedinica unutar veće strukture.

Tehnički metapodaci se koriste za detaljan opis gradiva te njegov lakši pronalazak, brzu pretraživost i za potvrdu da je ostalo neizmijenjeno i potpuno. Neke od stvari koje ti metapodaci uključuju su: identifikator tj. signatura, razina (npr. fond, zbirka, serija, podserija, predmet, komad), vrsta zapisa, sadržaj, napomene, jezik, pismo, nosač zapisa, količina, mjesto, autor, dostupnost i sl.

Neki od tehničkih metapodataka automatski nastaju skeniranjem te su ugrađeni u samu datoteku, a obuhvaćaju primjerice vrstu datoteke, rezoluciju, dimenzije, vrijeme kreiranja itd. no oni ne pripadaju u ovaj korak, nego u jedan od sljedećih.

2.3.3. Edukacija zaposlenika

Proces digitalizacije je slojevit, dugotrajan, zahtjevan i kompliciran s obzirom na to da postoje mnogi koraci, kriteriji, norme i pravila koja se moraju ispoštovati, što iziskuje i to da na digitalizaciji rade za to educirani stručnjaci. Sudeći po analizi tvrtke Ernst & Young Savjetovanje d.o.o., u njihovom Izvješću o analizi trenutnog stanja (2018) stoji da je „čak 50% institucija koje samostalno digitaliziraju građu kulturne baštine navelo kako njihovi zaposlenici angažirani na poslovima digitalizacije nisu pohađali nikakvu edukaciju vezanu uz digitalizaciju. To značajno ukazuje na potrebu za sistematičnim pristupom jačanja postojećih kapaciteta za provođenje aktivnosti digitalizacije.“ Treba još jednom napomenuti da se radi o analizi iz 2018. godine pa je za očekivati i ponadati se da su se stvari u međuvremenu promijenile na bolje.

Djelatnike treba uputiti u rad s opremom, prvenstveno sa skenerima, organizirati teorijske pa onda i praktične radionice na kojima će se simulirati skeniranje i obrada snimki. Također ih treba educirati o korištenju svih softvera koje će koristiti u digitalizaciji i učitavanju na druge sustave. Edukacije mogu voditi stručnjaci za digitalizaciju unutar ustanove ako ih ustanova ima ili pak vanjski stručnjaci koje će dotična ustanova unajmiti.

2.3.4. Priprema gradiva za digitalizaciju

Odabrano je gradivo prije skeniranja potrebno pripremiti na adekvatan način. The National Archives u svojim Smjernicama i uputama za digitalizaciju (2013) ističe da će u idealnom slučaju sljedeću pripremu obaviti konzervator: procijeniti stanje spisa kako bi se uvjerio da nisu preosjetljivi za skeniranje, pregledat će ima li na gradivu plijesni, stranica zalijepljenih jedna na drugu ili kakvih umetnutih objekata koji nisu arhivsko gradivo te se iz tog razloga ne skeniraju, zatim će pregledati je li narušena čitljivost teksta te ukloniti spajalice. Osim spajalica treba ukloniti i sve ostalo što se (greškom) nalazi u gradivu što bi prilikom skeniranja moglo oštetiti skener ili gradivo, naročito ako se radi o protočnom skeneru. Ako djelatnici koji će provoditi digitalizaciju nisu arhivisti koji već uspješno i stručno rukuju arhivskim gradivom, u smjernicama i uputama (2013) se spominje kako bi oni trebali proći obuku za rukovanje spisima koju bi provodili konzervatori. Prije skeniranja potrebno je odrediti na kakvom će se skeneru gradivo skenirati. To ovisi o starosti i općem stanju gradiva, veličini, odnosno formatu,

zatim o vrsti medija izvornika, radi li se o uvezanom gradivu te se naravno uzima u obzir kakva je oprema uopće na raspolaganju instituciji u kojoj se digitalizira.

K tome je potrebno osigurati ispravan redoslijed spisa, ispravno ih okrenuti i uvjeriti se da su predmeti potpuni. U slučaju da neki spisi nisu predani arhivu, na tim se mjestima skenira službena zabilješka koja potvrđuje da pojedini spis nedostaje. Gradivo je potrebno dopremiti iz spremišta u prostor gdje će se skenirati, treba ustanoviti redoslijed snimanja te dogovoriti način imenovanja datoteka.

2.3.5. Snimanje – pretvorba gradiva iz analognog u digitalni oblik

Za ovaj se korak ne može reći da je najvažniji jer su svi koraci podjednako važni, no može se reći da je skeniranje, odnosno pretvorba iz analognog u digitalni oblik, središnji korak u spomenutom procesu. Digitalizacija sa sobom donosi brojne smjernice i preporuke koje se donose u pojedinim zemljama ili institucijama zasebno, no one se često i preklapaju, naročito stavke koje su neophodne za trajno očuvanje gradiva. Postoje dva načina digitalizacije u arhivima, jedna je kada se nakon završene digitalizacije fizički izvornici unište, a digitalizirana verzija postane jedini vjerodostojan primjerak jedinice gradiva s pravnim učincima izvornika. Druga je kada se fizički izvornici ne unište, već se nastave čuvati, a digitalizirana verzija se daje na korištenje korisnicima.

2.3.5.1. Cilj i kriteriji

Cilj procesa digitalizacije jest, kako je to primjerice navedeno u Pravilima za digitalizaciju Finskog državnog arhiva (2018), proizvesti autentične i vjerodostojne digitalne manifestacije jedinica gradiva. Gradivo treba biti digitalizirano u cijelosti, dakle u svom punom sadržaju. To u praksi znači da se sve zbirke/serije/predmeti/komadi moraju digitalizirati kako nijedna informacija ne bi ostala nedigitalizirana zbog tehničke ili funkcionalne pogreške (naročito ako se nakon digitalizacije fizičko gradivo uništava). Dalje se u istim Pravilima (2018) propisuje da svaka pojedinačna slikovna datoteka mora sadržavati iste informacije kao i njezina analogna manifestacija te mora biti vizualno dostupna. Digitalizirani primjerak ne smije sadržavati elemente koji nisu uključeni u analogni zapis, a sve prazne stražnje stranice dokumenata treba ukloniti tijekom skeniranja. Prazna stražnja stranica je ona stranica dokumenta koja ne sadrži nikakve informacije. Stranice koje sadrže bilo kakve informacije ne smiju se uklanjati.

Slikovne datoteke nastale tijekom procesa digitalizacije moraju biti okrenute na pravu stranu, a smiju se eventualno nakon skeniranja okrenuti za devedeset stupnjeva.

2.3.5.2. Formati

Kada govorimo o formatima, bitno je odabrati adekvatan format za svaku vrstu gradiva. Smjernice za digitalizaciju koje se najčešće uvažavaju i koje se spominju u ovom radu su FADGI: Technical Guidelines for Digitizing Cultural Heritage Materials: Third Edition, Metamorfoze Preservation Imaging Guidelines te ISO/TR 19263-1 Photography – Archiving systems – Part 1: Best practices for digital image capture of cultural heritage material. One za svaku vrstu gradiva savjetuju optimalne i primjerene formate datoteka, a naravno svaka ustanova ili svaki projekt može zasebno izraditi svoje smjernice za digitalizaciju ili pak prihvatiti već postojeće.

Za pomisliti je možda da bi sve trebalo snimiti u najvišoj mogućoj kvaliteti, no to sa sobom povlači velike troškove, iznimno puno potrebnog prostora za pohranu, sporije skeniranje i dugotrajnije učitavanje datoteka u sustav te u konačnici vjerojatno i lošije rezultate optičkog prepoznavanja znakova (OCR), kako je to utvrdio Trbušić (2022).

Najveću količinu arhivskog gradiva obuhvaćaju tekstualni dokumenti i knjige te se za njih u spomenutim normama i smjernicama preporučaju formati JPEG2000, TIFF i PDF/A, no prihvatljivi su i JPEG, PNG i PDF. Što se tiče rezolucije, predlaže se ona od 300 dpi osim u slučaju iznimno malog teksta, a boja od 24 bita. Rezolucija od 300 dpi se pokazala optimalnom za tekstualne datoteke koja nudi idealan omjer između kvalitete i veličine datoteke. U slučaju da se namjerava na snimkama napraviti i proces optičkog prepoznavanja znakova (OCR), pokazalo se da su rezultati najbolji s TIFF datotekama. Ovdje treba napraviti distinkciju i pojasniti razliku između arhivske i prezentacijske datoteke. Prvenstveno se ove norme odnose na arhivsku datoteku (engl. *archival master file*). Ona prema FADGI-evom glosaru podrazumijeva datoteku koja predstavlja najbolju kopiju proizašlu iz digitalizacije koja se želi trajno očuvati bez gubitka ikakvih esencijalnih značajki. Datoteka smanjene kvalitete koja se izvodi iz nje zove se prezentacijska datoteka (engl. *derivative/access file*) te se prilagođava ovisno o namjeni i uglavnom se ne smatra trajnim sastavnim dijelom arhivske zbirke. Jedan od razloga za to jest zato jer njihovi formati mogu vrlo brzo zastarjeti, piše u spomenutom glosaru. Ukratko, arhivska datoteka verzija je u najboljoj kvaliteti koja sadrži sve informacije i trajno

se čuva, a prezentacijska se koristi primjerice na mrežnim stranicama gdje korisnici samostalno mogu pregledavati gradivo. Pritom je bitno da su datoteke čitljive te da se mogu brzo učitati.

Za fotografije *FADGI*, *ISO/TR* i *Metamorfoze* predlažu TIFF i JPEG2000, no prihvatljivi su i JPEG i PNG, boja u 24 bita te rezolucija od 400 do 600 dpi s obzirom na to da se na njima nalazi više detalja nego u tekstualnim datotekama. Za mikrofilm se predlažu isti formati kao i za sve ostalo spomenuto, a rezolucija od 4000 dpi jer se radi o iznimno malenom mediju te boja u samo 8 bita jer ni u izvorniku nije više. Za fotografske negative se također predlažu isti formati što nas dovodi do zaključka da TIFF i JPEG2000 omogućuju najvišu kvalitetu (nisu komprimirani) i najpogodniji su za dugotrajno čuvanje. Iz njih se mogu izvoditi drugi formati kao što su popularni JPEG, PNG, GIF i PDF koji su pogodni za korištenje i pregledavanje.

2.3.5.3. Tehnički kriteriji digitaliziranog materijala

Kvaliteta digitaliziranog materijala uvelike ovisi o analizi njihovih tehničkih kriterija. Budući da oni međusobno utječu jedni na druge, njihov je redoslijed ključan u procesu digitalizacije i kasnije kontrole kvalitete. Slijedi njihov redoslijed naveden u smjernicama *Metamorfoze Preservation Imaging Guidelines* (2012):

1. Prostor i dubina boje. Raspon boja se mora odrediti prije skeniranja. 8-bitne vrijednosti piksela za izražavanje stupnja ekspozicije i kontrasta mogu se razlikovati za svaki prostor boje. Datoteke bez prostora boja ne mogu se tehnički provjeriti.
2. Balans bijele boje. Za ispravnu procjenu ekspozicije i točnosti boje potrebne su neutralne postavke za cijeli spektar sive boje.
3. Ekspozicija. Ekspozicija se prvenstveno vrednuje pri visokom svjetlu. Ona se kod snimanja kamerom postavlja hardverski s odgovarajućim postavkama otvora blende i svjetla.
4. Zrnatost se može odrediti ako je ekspozicija u skladu s vrijednostima koje propisuju ove smjernice.
5. Ekspozicija i zrnatost. Nakon određivanja ekspozicije i zrnatosti na svjetlu, može se odrediti i ostatak spektra sive boje. Osvjetljenje se također može odrediti tek nakon određivanja vrijednosti zrnatosti slike.

Nakon tonske procjene slike opisane iznad, prema smjernicama (2012) se određuju sljedeći tehnički kriteriji:

6. Standardno odstupanje, šum
7. Osvjetljenje
8. Odstupanja u boji
9. Preciznost i točnost boje
10. Brzina uzorkovanja. Postavljanje ispravne i željene brzine uzorkovanja je prvi korak u digitalizaciji. Stvarna brzina uzorkovanja i ona navedena u metapodacima se potvrđuje pri tehničkoj provjeri nakon procjene točnosti boje.
11. Funkcija prijenosa modulacije (MTF10), učinkovitost uzorkovanja, MTF50, maksimalna modulacija, pogrešna interpretacija boje
12. Distorzija slike
13. Anomalije na snimkama, tzv. artefakti

2.3.5.4. Imenovanje datoteka

U Smjernicama za digitalizaciju kulturne baštine Australije i Novog Zelanda (2012), jedna od spomenutih stavki je imenovanje skeniranih datoteka te stoji da prije snimanja treba uspostaviti shemu imenovanja datoteke. Pritom treba uzeti u obzir zahtijeva li identifikator strojno ili ručno indeksiranje (ili oboje, u kojem slučaju slika može imati i više identifikatora). Imena datoteka mogu biti opisna (npr. usvajanje postojeće identifikacijske sheme koja povezuje digitalnu datoteku s izvornikom) ili neopisna (npr. numerički niz). Opisni nazivi datoteka već u imenu sadrže neke metapodatke, dok su neopisni nazivi datoteka povezani s metapodacima pohranjenima drugdje te im oni služe za identifikaciju datoteke. U istim se Smjernicama (2012) zaključuje da se manji projekti mogu služiti opisnim imenima datoteka koja olakšavaju pregledavanje i dohvaćanje; a velikim je projektima preporučljivo koristiti strojno generirana imena za datoteke koje se pronalaze naprednim pretraživanjem u bazi podataka zahvaljujući odgovarajućim metapodacima. Konkretni kriteriji koji se spominju su da naziv datoteka treba biti jedinstven i dosljedan, da u obzir treba uzeti okvirni broj datoteka te prema tome odrediti broj znamenki, korištenje vodeće nule kako bi se datoteke ispravno poredale i izbjegavanje razmaka u imenu te umjesto toga stavljati donju crtu. Osim toga, spominje se i izbjegavanje suviše dugih i kompleksnih naziva koji su podložni ljudskim pogreškama prilikom unošenja, izbjegavanje razmaka i simbola koji bi na raznim sustavima mogli stvarati probleme. Metapodatke u imenu datoteke (primjerice datum skeniranja, broj stranice i sl.) treba evidentirati još negdje jer je moguće da će se imena datoteka u budućnosti po potrebi mijenjati.

2.3.5.5. Metapodaci

Skeneri i fotoaparati prilikom snimanja generiraju razne tehničke metapodatke o slici pa su kod čestih formata kao što su JPEG, JPEG2000 i TIFF ti metapodaci ugrađeni u zaglavlje datoteke. Tehnički metapodaci, prema ISO/TR 19263-1:2017(E), omogućavaju softverima ispravnu obradu i renderiranje snimki te su korisni pri kontroli kvalitete i upravljanju snimkama. Što se tiče kontrole kvalitete snimke, ista norma preporučuje spremanje sljedećih tehničkih metapodataka zajedno s rezultatima analize kvalitete: datum i vrijeme snimanja slike, ime osobe koja ih je snimala (njezino ime ili svojevrsni ID), ime uređaja (proizvođač i model), ime i verzija softvera u kojem se skenira, postavke kamere: otvor blende, brzina zatvarača, ISO; podaci o slici: visina i širina, rezolucija, dubina bita, prostor boje i profil boje.

2.3.6. Kontrola kvalitete

U tehničkim smjernicama za digitalizaciju kulturne baštine koju je sastavio FADGI (2023) stoji da ovaj proces može uključivati automatizirani pregled usklađenosti s tehničkim specifikacijama. Nakon toga slijedi fizički pregled pri kojem se pregledava reprezentativni uzorak snimki kako bi se provjerilo jesu li u skladu sa specifikacijama projekta. Veličina uzorka određuje se u fazi planiranja projekta. Softveri kao što su JHOVE ili Jpylyzer mogu pomoći u tome.

Da bi se izbjeglo što više greški pri skeniranju (npr. stranice koje nedostaju, slaba čitljivost, nepotpunost predmeta, komprimirane datoteke), mora se izvršiti odgovarajuća kontrola kvalitete i po potrebi ponovno evidentirati i podesiti postavke skeniranja, piše u Tehničkim smjernicama 03138 (2017) njemačkog Saveznog ureda za informacijsku sigurnost (BSI). Detaljan slijed koraka kontrole kvalitete treba odrediti sukladno opsegu skeniranja te važnosti pojedinog gradiva. Dalje u istim Smjernicama BSI-ja (2017) piše da se „prilikom obrade velike količine spisa od „normalne“ važnosti, vizualni pregled može svesti na ravnomjerne nasumične uzorke s ciljem detektiranja eventualnih sustavnih pogrešaka (primjerice neodgovarajuće postavke skeniranja, greške u radu skenera).“ Frekventnost tih uzorka treba odrediti ovisno o stupnju važnosti zaštite skeniranih dokumenata i pouzdanosti skenera te je navesti u popratnoj dokumentaciji o kontroli kvalitete. Osim toga, mogu se koristiti softveri za automatsku kontrolu kvalitete s funkcionalnostima kao što su automatsko prepoznavanje praznih stranica, detektiranje nedovoljne kvalitete slike ili usporedba navedenih brojeva stranica s brojem snimaka (npr. prema navedenim metapodacima). Međutim, budući da su automatske provjere

kvalitete same po sebi također sklone pogreškama (npr. softver stranice s rukom pisanim komentarima slabo razlikuje od praznih stranica), svejedno treba izvršiti ručnu provjeru tih automatski identificiranih nepravilnosti. Ako se planira uništiti fizičke primjerke nakon skeniranja, to se nikako ne smije učiniti prije kontrole kvalitete.

BSI (2017) podsjeća da se naknadna obrada snimki (npr. promjena kontrasta/svjetline, promjena zasićenosti boje, obrezivanje, smanjenje šuma) ne smije provoditi osim s ciljem poboljšanja čitljivosti. Ako se to ipak radi, mora se adekvatno evidentirati i provesti s oprezom i pažnjom kako se potencijalno važne informacije ne bi uništile ili nehotice narušila autentičnost. U uputstvima za digitalizaciju treba navesti kakva je naknadna obrada snimki dopuštena i u kojim slučajevima.

2.3.7. Vođenje dokumentacije

Za pouzdano upravljanje procesom digitalizacije i za njegovu transparentnost treba postojati odgovarajuće regulirano dokumentiranje svih popratnih postupaka i procesa, objašnjava BSI (2017), a posebno obuhvaća sljedeće:

- informacije o promjeni kritičnih konfiguracijskih parametara te o autentifikaciji i autorizaciji,
- podatke o tome tko je, kada i na koji način provodio skeniranje,
- informacije o tome je li izvršena naknadna obrada snimki,
- izvještaj o neuspjeloj provjeri autentičnosti i drugim problemima i greškama koje su se dogodile.

Osim navedenih stavki treba voditi evidenciju o tome tko je, kada i koje gradivo dopremio u prostor za skeniranje, tko je i kada koje snimke pregledao te točno koje je promjene izvršio. Za digitalizirano gradivo trebaju postojati tablice s podacima o gradivu kao što su primjerice signatura, naziv, format, opis i ostali metapodaci.

2.3.8. Arhiviranje

Nakon što je završeno skeniranje i kontrola kvalitete, snimke je potrebno arhivirati. U ovoj se fazi još dodatno može provjeriti integritet snimki ako se zaključi da je potrebno. Nakon što se snimke skeniraju i provjeri zadovoljavaju li sve kriterije, za svaku se generira jedinstvena *hash* vrijednost, odnosno kriptografski sažetak. U slučaju ikakvih izmjena na datoteci, kriptografski

sažetak se mijenja te se na taj način može lako saznati je li autentičnost ugrožena. Snimke se pohranjuju na interni sustav za pohranu institucije te se obavezno radi i sigurnosna kopija. Sigurno se pohranjuju arhivske master datoteke koje će se trajno čuvati, ili pak do isteka roka čuvanja, a za davanje korisnicima gradiva na uvid i objavljivanje na mrežne sustave, koriste se prezentacijske datoteke izvedene iz master datoteka. Prednost arhivskih datoteka je, između ostalog, da se iz njih mogu izvesti brojni formati i to neograničeno mnogo puta pri čemu kvaliteta arhivske datoteke ostaje nepromijenjena. Uz gradivo koje zadovoljava načela integriteta, autentičnosti, pouzdanosti i iskoristivosti potrebno je čuvati apsolutno svu popratnu dokumentaciju nastalu u procesu digitalizacije jer bez toga gradivo ne bi ni ispunjavalo spomenuta načela.

2.3.9. Objavljivanje

FADGI (2023) u svojim tehničkim smjernicama napominje da arhivske datoteke nisu prikladne za prikaz, osim u iznimnim situacijama. Za objavljivanje ili prikazivanje, iz tih se datoteka izvode nove, tzv. prezentacijske datoteke u jednom od odgovarajućih formata, a moguće ih je izvesti u automatiziranom strukturiranom postupku u obradi nakon skeniranja ili u naknadno prema potrebi. Pritom se mora paziti da boja ostane vjerna onoj u arhivskoj datoteci.

Kada se datoteke učitavaju u softver koji se kasnije povezuje sa nekim sustavom gdje će digitalizirana građa biti dostupna za pregledavanje, izvođenje prezentacijskih datoteka iz arhivskih se vrši automatski prema točno određenim postavkama glede formata. U slučaju da se digitalizirano gradivo ne objavljuje, pojedino gradivo se izvodi u obliku prezentacijskih datoteka nakon što korisnik ispunji formular za zahtjev o uvidu u to gradivo te mu se te datoteke isporučuju u korisnički pretinac.

Osim samih snimki gradiva, u mrežne sustave za pregledavanje gradiva unose se i metapodaci o gradivu zbog lakšeg pronalaska i dodatnih informacija. Za to se najčešće koristi isti softver pomoću kojeg su učitane datoteke. Ti metapodaci, kao što su naziv jedinice gradiva, autor, razdoblje, pridržana prava i sl., upisuju se u točno strukturirane datoteke (npr. tablice u MS Excelu) kako bi se pri učitavanju u sustav one automatski povukle onamo kamo trebaju. Na taj se način ujedno izbjegava ručni unos za svaku kategoriju metapodataka koji bi za veći projekt digitalizacije bio skoro nemoguć, a čak i za manji bi bio iznimno dugotrajan.

2.3.10. Briga o trajnom očuvanju i dugoročno upravljanje

Smjernice za digitalizaciju kulturne baštine Australije i Novog Zelanda (2012), odnosno AS/NZS ISO 13028:2012 pojašnjavaju da digitaliziranim zapisima treba upravljati na način koji omogućuje njihovo kontinuirano postojanje onoliko dugo koliko ih je potrebno očuvati. Iznimno su osjetljivi na zastarjeli hardver, softver i medij te trebaju biti uključeni u postupke za upravljanje spisima koje je usvojila pojedina ustanova da bi uistinu ostali dostupni i upotrebljivi dok ih se god mora čuvati. Neki od tih postupaka su migracija s medija na medij i pretvorba iz jednog formata u drugi.

U Strategiji za očuvanje digitalnog gradiva (2022) Nacionalnog arhiva Sjedinjenih Američkih država (NARA) za razdoblje od 2022. do 2026. aspekti očuvanja digitalnog gradiva raščlanjeni su na četiri ključne sastavnice. One uključuju infrastrukturu, integritet, trajnost formata i medija te sigurnost informacija. Neki od infrastrukturnih preduvjeta za očuvanje digitalnog gradiva (pritom se navedeno odnosi i na digitalne izvornike i na digitalizirano gradivo) su dovoljno mjesta za pohranu, stabilna mreža, sustavi i alati za unos, obradu, aktivno upravljanje datotekama, očuvanje i izvoz između sustava digitalnih izvornika i digitaliziranog gradiva. Zatim bi trebala postojati jedna sigurnosna kopija na udaljenoj geografskoj lokaciji za dodatnu sigurnost (pohrana u oblaku), kao i alati za konverziju formata s obzirom na to da gradivo podliježe velikom riziku s vremenom kako formati zastarijevaju. Pritom se misli na konverziju svih datoteka određenog formata ili verzije formata u odabrani format. Kada je riječ o integritetu, NARA (2022) navodi da bi trebalo imati popis svih datoteka u digitalnom repozitoriju, zabilježiti podatke o svakom novom unosu i po mogućnosti sve kasnije intervencije kao što su konverzije formata, premještanje datoteka i revizija podataka. Unos datoteka trebao bi uključivati provjeru sadrže li datoteke kakav zlonamjerni softver i odgovara li kriptografski sažetak onom početnom kako bi se moglo potvrditi da ništa nije mijenjano. Zapise koji se trajno čuvaju trebalo bi kopirati na provjereni novi elektronički medij unutar deset godina od inicijalne pohrane. U vezi trajnosti formata i medija, poželjno je izraditi akcijski plan očuvanja digitalnog gradiva koji u obzir uzima formate u kojima je postojeće gradivo. Sukladno s tim on određuje postupanje u slučaju ako ti formati više neće biti podržani, primjerice ako im se više neće moći pristupiti putem postojećeg softvera. Ako do toga dođe, ugrožene datoteke potrebno je konvertirati sukladno akcijskom planu. Između ostalog bi trebalo kontinuirano analizirati vrste formata i medija sa svrhom prepoznavanja moguće zastarjelosti i pravovremeno migrirati gradivo s medija na medij. Sigurnost informacija

podrazumijeva primjerice da se zna tko ima pristup fizičkom gradivu i tko ima pristup sustavima za unos i obradu pohranjenih podataka.

Uz to valja napomenuti da bi bilo korisno da sustav može pratiti popis gradiva s rokovima čuvanja te automatski izlučiti gradivo nakon isteka roka te na taj način osigurati vrijedan prostor za pohranu novozaprimljenog gradiva. Automatsko izlučivanje se dakako ne odnosi na arhivsko gradivo od velikog značaja koje se trajno čuva.

3. Projekt eKultura

Važnost digitalizacije kao sredstvo u ostvarivanju kulturne politike, kao i razvoja kulture u Hrvatskoj već je prepoznalo Ministarstvo kulture i medija Republike Hrvatske iz kojeg je razloga 2005. pokrenut Nacionalni program digitalizacije arhivske, knjižnične i muzejske građe. Projekt eKultura – digitalizacija kulturne baštine (dalje u tekstu: eKultura ili Projekt) zamišljen je kao nastavak i nadovezivanje na prethodni. U Smjernicama za digitalizaciju (2020) istaknuto je da se ovim projektom nastoji „povećati pristup kulturnoj baštini u digitalnom obliku te je zaštititi sustavnom, standardiziranom i trajnom pohranom u digitalnom obliku na jednom mjestu“, a građani, znanstvenici, studenti i pripadnici AKM zajednice imat će jedinstven pristup i slobodno korištenje gradiva, što će se intenzivno ohrabrivati, naročito u obrazovanju, znanosti i turizmu. U Smjernicama stoji da je provoditelj projekta Ministarstvo kulture i medija, a kao partneri su navedeni Hrvatski državni arhiv (HDA), Hrvatska radiotelevizija, Muzej za umjetnost i obrt te Nacionalna i sveučilišna knjižnica.

3.1. Ciljevi

Unutar projekta definirani su ciljevi, što je jedan od elemenata koje svaki projekt mora imati kako bi se tijekom i uspješnost projekta mogli pratiti i evaluirati. Da bi oni bili smisleni, konkretni i da bi ih se moglo pratiti, trebali bi zadovoljavati takozvane SMART kriterije. Oni podrazumijevaju da ciljevi projekta budu specifični na način da se točno zna što se želi postići; mjerljivi tako da se jasno mogu kvantificirati; dostižni (engl. *achievable*) u smislu da su realno postavljani u skladu s mogućnostima; relevantni u pogledu neke stvarne koristi od njih te da imaju jasno određeni vremenski rok (engl. *time-bound*) unutar kojeg rezultati i zadani ciljevi trebaju biti isporučeni.

Kako stoji u Planu digitalizacije kulturne baštine 2020.-2025. sastavljenom povodom pripreme projekta eKultura, dva glavna cilja su unaprjeđenje mrežne dostupnosti digitalne kulturne baštine i uspostavljanje održivog umreženog središnjeg sustava za digitalnu kulturnu baštinu. Mrežnu dostupnost digitalne kulturne baštine zamišljeno je provesti uspostavljanjem umreženog sustava za digitalnu kulturnu baštinu i poticanjem korištenja te baštine. Održivost središnjeg sustava planira se provesti kontinuiranim praćenjem, vrednovanjem i izvještavanjem o rezultatima te razvojem novih usluga i proizvoda. To bi npr. bilo povezivanje s Europeanom, najznačajnijim projektom glede digitalizacije kulturne baštine u Europi. Jedna od aktivnosti koja se spominje u Planu jest edukacija djelatnika zaduženih za digitalizaciju i korisnika

sadržaja jer je to nužno za stjecanje više digitalnih stručnjaka, ali i za digitalno, odnosno informacijsko opismenjavanje građana. Zamišljena je na način da se predlože teme za spomenutu obuku djelatnika koji će biti zaduženi za digitalizaciju, a korisnicima sustava će se dati mogućnost da i sami sudjeluju u osiguravanju bolje kvalitete sadržaja kroz tagiranje, komentiranje, osvrt na metapodatke i sl. U suradnji s obrazovnim ustanovama razvit će se programi cjeloživotnog učenja. Ishod će se moći pratiti pomoću količine održanih edukacija, broja polaznika i programa cjeloživotnog učenja realiziranih do 2025. godine.

3.2. Infrastruktura i alati

Na projektu eKultura u Hrvatskom državnom arhivu za korištenje je predviđeno šest skenera i to dva koračna skenera za knjige modela Bookeye 5 V2 Automatic kojima se predlošci snimaju odozgo (engl. *overhead scanner*), dva ista takva koračna skenera za knjige modela Bookeye 5 V3 Automatic, jedan WideTEK 48 skener za velike formate te jedan protočni skener Canon imageFormula DR G2090. Na Canonovom skeneru instaliran je Canonov softver CaptureOnTouch Pro, a na ostalima Batch Scan Wizard (ImageAccess). Za upravljanje procesom digitalizacije koristi se prilagođena inačica aplikacije Goobi. Uspostavljen je Središnji sustav koji služi za upravljanje snimkama te će se ujedno na njemu i pristupati snimkama po završetku digitalizacije. Osim toga, korištene su tablice za metapodatke u programu MS Excel te razni obrasci za dokumentaciju procesa od kojih će neki biti spomenuti kasnije u radu.

Radi vizualne predodžbe priložena je fotografija (slika 2) koja redom prikazuje spomenute modele skenera kako slijedi: u gornjem redu su dva skenera za snimanje odozgo, a u donjem skener za velike formate (lijevo) te protočni skener (desno). Sa slike je vidljivo da prva tri imaju vlastite ekrane na kojima se reguliraju postavke, pokreće sam proces skeniranja te prikazuju snimke. Protočnim se skenerom upravlja putem računala na kojem je instaliran njegov softver te se snimke pregledavaju na računalnom monitoru.



Slika 2. Prikaz vrsta i modela skenera korištenih na projektu eKultura.

Preuzeto s <https://www.imageaccess.de/?page=ScannersBE5-V2A&lang=en>,
<https://www.imageaccess.de/?page=ScannersBE5-V3A&lang=en>,
<https://www.rpg.com/product/widetek-48-scanner-bundle/>, <https://pcchip.hr/printeri/canon-predstavlja-novu-generaciju-produkcijskih-skenera-imageformula-dr-g2090-dr-g2110-i-dr-g2140/> (21.4.2023.).

3.3. Odabir gradiva za digitalizaciju

Pri odabiru gradiva za digitalizaciju vođeno je računa o dvije glavne stavke, a to su vrijednost gradiva i iskoristivost. Izabrano je gradivo koje se naročito želi sačuvati i očuvati jer s vremenom čak i u adekvatnim uvjetima fizičko gradivo u nekoj mjeri propada. Druga glavna

stavka pri odabiru gradiva za digitalizaciju jest koliko je pojedino gradivo traženo među korisnicima. Logično je da bi bilo poželjno i praktično digitalizirati ono što korisnici često traže, a objavljivanjem tog gradiva na mrežne stranice i omogućavanjem korisnicima da samostalno, bilo kada i bilo gdje pristupe tom gradivu, zasigurno će se smanjiti opterećenje na djelatnike arhiva koji korisnicima isporučuju traženo gradivo. Osim toga, u obzir je uzet i opseg i trajanje projekta kako bi se adekvatno moglo odrediti koji fondovi/zbirke će se snimati i koliko njih jer je projekt ograničen i financijama i vremenskim rokom.

Na slici 3 prikazan je konkretan prijedlog za digitalizaciju arhivskoga gradiva u sklopu projekta eKultura. Na početku su navedeni opći podaci o predlagatelju, a zatim o gradivu. U tom se kontekstu spominju signatura, naslov, razdoblje nastanka, količina izvornika izražena u dužnim metrima te u čemu su pohranjeni, očekivani broj snimaka te napomene. Naposljetku se prijedlog obrazlaže na način da se odabranom gradivu dodjeljuju (pr)ocjene od jedan do deset za sljedeće kriterije:

1. fizičko-tehnička ugroza izvornika: u kojoj je mjeri gradivo staro, oštećeno i već teško čitljivo. Ako je to slučaj, naročito u kombinaciji s velikom traženosti među korisnicima, vrijedilo bi ga digitalizirati jer bi bilo spašeno od propadanja i mogao bi ga koristiti velik broj korisnika bez negativnih posljedica za fizički izvorNIK.
2. otežano i zahtjevno fizičko rukovanje izvornim gradivom prilikom korištenja: radi li se o velikim formatima, činjenici da je gradivo pohranjeno na udaljenoj lokaciji, oštećuje li se prilikom česte dopreme i dr.
3. interes korisnika
4. stvaranje dodatne vrijednosti za ustanovu kroz: kulturno-prosvjetne aktivnosti, objavu na mrežne portale te projektne suradnje
5. sređenost i cjelovitost. U obzir za digitalizaciju uzet će se gradivo koje je prilično cjelovito i sređeno jer je postupak sređivanja vrlo složen i dugotrajan pa je za očekivati da je bitno gradivo već sređeno i spremno za digitalizaciju (treba napomenuti da sređivanje gradiva i priprema gradiva za digitalizaciju nisu ista stvar).
6. dostupnost: u kojoj mjeri gradivo sadrži osobne podatke, oznake tajnosti i autorska prava jer ono ne bi moglo biti objavljeno na mrežnim portalima za javnost bez ograničenja
7. raspoloživost financija, kadra i opreme potrebnih za realizaciju. Primjerice neće se predložiti snimanje kartografske zbirke ako ustanova ne posjeduje skener pogodan za velike formate.

Prijedlog za digitalizaciju arhivskoga gradiva

PRIJEDLOG ZA DIGITALIZACIJU ARHIVSKOGA GRADIVA										
OPĆI PODATCI O PREDLAGATELJU										
Odjel :	Odjel za zaštitu i obradu arhivskog gradiva									
Odsjek:	Odsjek za arhivsko gradivo 1945.-1990.									
Osoba za kontakt:	[REDACTED]									
Kontakt podatci:	tel./mob.: [REDACTED]									
	e-mail: [REDACTED]@arhiv.hr									
PODATCI O GRADIVU										
Signatura (oznaka):	HR-HDA-306.									
Naziv (naslov):	Zemaljska komisija za utvrđivanje zločina okupatora i njihovih pomagača (ZKRZ) 2.2.1. Kartoteka i popisi žrtava 2.2.2. Kartoteka zločinaca									
Razdoblje nastanka (od – do):	1944-1947. (1950.)									
Količina izvornika:	12 d/m: 2.2.1. Kartoteka i popisi žrtava, 6 d/m (5,6 d/m kartica i 4 arhivske kutije) 2.2.2. Kartoteka zločinaca, 6 d/m (27.616 kartica)									
Očekivani broj snimaka:	78.000 snimaka									
Napomene:	Cijeli fond je snimljen na mikrofilmove (c/b) prema poretku u kutijama. Dio mikrofilmova je digitaliziran i još ih treba ih obraditi za korištenje sukladno strukturi fonda. Dio mikrofilmova još nije digitaliziran uključujući i mikrofilmove na kojima je u nekom ranijem razdoblju snimljena i kartoteka. Kartoteka i popisi žrtava je pomagalo izrađeno na temelju izvornih popisa žrtava po kotarevima i okruzima. Kartoteka zločinaca je pomagalo koje se prvenstveno koristi kako bi se moglo pristupiti gradivu fonda i zbog toga se predlaže digitalno snimanje direktno s izvornih kartica.									
OBRAZLOŽENJE PRIJEDLOGA										
Kriterij:	Procjena:									
1. Fizičko-tehnička ugroza izvornika: fizička oštećenost, loša kvaliteta predloška, blijeđenje zapisa, starost i sl.	1	2	3	4	5	6	7	8	9	10
2. Otežano i zahtjevno fizičko rukovanje izvornim gradivom prilikom korištenja: veliki formati, pohrana na izdvojenoj lokaciji, uništavanje izvornika i zaštitne ambalaže uslijed učestale opreme-dopreme i sl.	1	2	3	4	5	6	7	8	9	10
3. Interes korisnika – učestalo korištenje (vidi kriterije 1. i 2.)	1	2	3	4	5	6	7	8	9	10
4. Stvaranje dodane vrijednosti za HDA: korištenje u kulturno-prosvjetnim aktivnostima, objava na mrežnim portalima, uključivanje u projekte suradnje koji mogu ostvariti dodanu vrijednost za HDA i sl.	1	2	3	4	5	6	7	8	9	10
5. Stupanj središtenosti i cjelovitosti	1	2	3	4	5	6	7	8	9	10
6. Stupanj dostupnosti gradiva u odnosu na možebitna ograničenja: osobni podatci, oznake tajnosti, autorska prava	1	2	3	4	5	6	7	8	9	10
7. Raspoloživi financijski, kadrovski i tehnički resursi za realizaciju projekta digitalizacije	1	2	3	4	5	6	7	8	9	10
Dodatna obrazloženja / Napomene:										

Slika 3. Prijedlog za digitalizaciju arhivskoga gradiva.

Ustupio Hrvatski državni arhiv – Odsjek za arhivsko gradivo 1945.-1990.

Prijedlog za digitalizaciju se zatim pregledava, razmatra te se donosi odluka o tome hoće li se on prihvatiti. Ako je prihvaćen, kreće se u izradu metapodataka o gradivu.

3.4. Izrada metapodataka

Za odabrano i odobreno gradivo se zatim prikupljaju metapodaci. Oni se objedinjuju u tablice izrađene u programu Microsoft Excel jer je on nudi jednostavan, pregledan i sistematski prikaz podataka, a kasnije će poslužiti za automatsko ispunjavanje traženih podataka u sustavu u kojem će se na kraju objaviti. Podaci se većinom uzimaju i prepisuju iz prethodno postojećih obavijesnih pomagala te se nadopunjuju s dodatnim novim podacima. Excel tablice za metapodatke su standardizirane i sastoje se od tri lista u Excelu, a to su podaci, datoteke i list 2. Slike 4a i 4b ilustriraju konkretnu tablicu s metapodacima (razdijeljena je u dvije slike zbog bolje preglednosti), odnosno njezin prvi list. List dva popunjava se tek nakon završenog skeniranja i kontrole kvalitete.

	A	B	C	D	E	F	G	H
	Identifikator	Naslov	Razina	VisaID	Redoslijed	VrijemeOD	VrijemeDo	Sadržaj
1	HR-HDA-280	Izvršno vijeće Sabora Socijalističke Republike Hrvatske	Fond			1953.	1990.	Fond sadržava zapisnike sjednica Izvršnog vijeća Sabora, godišnje izvještaje o radu, uredbe, odluke, rješenja, upute i druge akte iz njegove nadležnosti, dokumentaciju u vezi s izradom prijedloga saveznih i republičkih propisa, izvještaje, elaborate i druge materijale koje su mu dostavljala savezna i republička tijela o poslovima iz svojega djelokruga, spise kabineta predsjednika Izvršnog vijeća Sabora, njegovih administrativnih i tehničkih stručnih službi, zapisnike i materijale sjednica te druge spise radnih tijela (odbori, komisije, radne grupe) i savjetodavnih tijela za pojedina poslovna područja (savjeti, komiteti, sekretarijati).
3	HR-HDA-280-1	Sjednice Izvršnog vijeća Sabora Socijalističke Republike Hrvatske	Serijska	HR-HDA-280	1	1953.	1990.	Serijska sadrži zapisnike sjednica i sastanaka IVS-a, te pojedine radne materijale.
4	HR-HDA-280-1-1	Zapisnici sjednica Izvršnog vijeća Sabora Socijalističke Republike Hrvatske	Podserijska	HR-HDA-280-1	1	1953.	1990.	Podserijska sadrži skraćene zapisnike sjednica IVS-a i pojedinih informativnih sastanaka
5	HR-HDA-280-1-1-1	Zapisnici sjednica, 1953.	Podserijska	HR-HDA-280-1-1	1	1953.	1953.	Sadrži zapisnike sjednica održanih u 1953. godini. Prva dva su zapisnici sjednica Vlade Narodne Republike Hrvatske (21.01. i 26.01.), od 06.02. do 16.12.1953. su zapisnici sjednica Izvršnog vijeća Narodne Republike Hrvatske.
6	HR-HDA-280-1-1-1-1	Zapisnik sjednice održane 21.01.1953.	Komad/Jedinica	HR-HDA-280-1-1-1	1	21.01.1953.	21.01.1953.	
7	HR-HDA-280-1-1-1-2	Zapisnik sjednice 26.01.1953.	Komad/Jedinica	HR-HDA-280-1-1-1	2	26.01.1953.	26.01.1953.	
8	HR-HDA-280-1-1-1-3	Zapisnik sjednice održane 06.02.1953.	Komad/Jedinica	HR-HDA-280-1-1-1	3	06.02.1953.	06.02.1953.	
9	HR-HDA-280-1-1-1-4	Zapisnik sjednice održane 19.02.1953.	Komad/Jedinica	HR-HDA-280-1-1-1	4	19.02.1953.	19.02.1953.	
10	HR-HDA-280-1-1-1-5	Zapisnik 3. sjednice održane 02.03.1953.	Komad/Jedinica	HR-HDA-280-1-1-1	5	02.03.1953.	02.03.1953.	
11	HR-HDA-280-1-1-1-6	Zapisnik 4. sjednice održane 09.03.1953.	Komad/Jedinica	HR-HDA-280-1-1-1	6	09.03.1953.	09.03.1953.	

Slika 4a. Tablica s metapodacima – prvi dio.

Ustupio Hrvatski državni arhiv – Odsjek za arhivsko gradivo 1945.-1990.

I	J	K	L	M	N	O	P	Q	R	S
Interni napomena	Količina	Tehnička jedinica	Izvorna oznaka/broj	VrstaZapisa	VrstaSadržaja	Institucija	Zbirka	StatusZapisa	ZaObjavu	Uvjeti korištenja
1. Gradivo je oblikovano u privremene cjeline prema vrstama dokumentacije i oznakama tajnosti (npr. zapisnici sjednica, spisi s oznakom „državna tajna“, strogo povjerljivi spisi, povjerljivi spisi, „trokut spisi“, opći spisi) te prema pojedinim stručnim, radnim i drugim tijelima.				TEXT	Tekstualni dokumenti	HR-HDA	hda	NEW	DatasetRecord	http://creativecommons.org/publicdomain/mark/1.0/
2.										
3.										
4.	118 knjiga, 28 kutija			TEXT	Tekstualni dokumenti	HR-HDA		NEW	Objavljen	http://creativecommons.org/publicdomain/mark/1.0/
5.	118 knjiga			TEXT	Tekstualni dokumenti	HR-HDA		NEW	Objavljen	http://creativecommons.org/publicdomain/mark/1.0/
6.	1 knjiga			TEXT	Tekstualni dokumenti	HR-HDA		NEW	Objavljen	http://creativecommons.org/publicdomain/mark/1.0/
7.	3 stranice	knjiga 1		TEXT	Tekstualni dokumenti	HR-HDA		NEW	Objavljen	http://creativecommons.org/publicdomain/mark/1.0/
8.	4 stranice			TEXT	Tekstualni dokumenti	HR-HDA		NEW	Objavljen	http://creativecommons.org/publicdomain/mark/1.0/
9.	2 stranice			TEXT	Tekstualni dokumenti	HR-HDA		NEW	Objavljen	http://creativecommons.org/publicdomain/mark/1.0/
10.	6 stranica			TEXT	Tekstualni dokumenti	HR-HDA		NEW	Objavljen	http://creativecommons.org/publicdomain/mark/1.0/
11.	5 stranica			TEXT	Tekstualni dokumenti	HR-HDA		NEW	Objavljen	http://creativecommons.org/publicdomain/mark/1.0/

Slika 4b. Tablica s metapodacima – drugi dio.

Ustupio Hrvatski državni arhiv – Odsjek za arhivsko gradivo 1945.-1990.

Prije skeniranja ispunjava se samo prvi list – podaci. Ovdje se unose strukturni podaci te deskriptivni metapodaci. U tom listu ima više vrsta rubrika, neke su obavezne, neke se ne moraju popuniti (te stoga nisu vidljive na slikama 4a i 4b), u nekima se samostalno pišu informacije, dok se u drugima bira između ponuđenih opcija u padajućem izborniku. Rubrike su:

- identifikator, tj. signatura po kojoj se gradivo evidentira. Radi se o standardiziranom načinu pisanja signature. Prva dva slova označavaju zemlju (u ovom slučaju HR), zatim se koristi kratica institucije (u ovom slučaju HDA), onda se koristi broj fonda pri čemu on ne prelazi četiri znamenke. Nakon njega slijede brojevi koji stoje za seriju, podseriju itd. i može ih biti onoliko koliko je potrebno. Sve vrijednosti međusobno su odvojene povlakom. Primjer jedne signature bio bi HR-HDA-1616-3-1-2.
- naslov
- razina. Ovdje je moguće birati između kategorija fond, zbirka, serija, podserija, predmet, komad/jedinica, podfond, klasa, grupa, podgrupa, sastavnica, drugo.
- visaID. Ovo je identifikator neposredno više arhivske jedinice i obuhvaća veću razinu unutar koje se nalazi jedinica koja se opisuje.
- redoslijed. Označava redni broj jedinice koja se opisuje unutar neposredno više arhivske jedinice.

- vrijeme od, vrijeme do
- sadržaj. U ovom se polju navodi kratki sadržaj koji se uglavnom piše samo na razini fonda i ponekih serija, a rijetko i predmeta. Nije obavezan za sve razine, no poželjan je.
- napomene. Riječ je pretežito o internim napomenama koje arhivistima služe za lakše snalaženje u gradivu i u pravilu se ne objavljuju ako ih se ne smatra relevantnima za korisnike.
- izvorna oznaka. Ona podrazumijeva oznaku gradiva koju je ono originalno imalo prije zaprimanja u arhiv i uvođenja novih standardiziranih signatura. Može pomoći u lakšem pronalasku gradiva.
- broj knjige/kutije
- fizička svojstva
- jezik
- pismo
- nosač zapisa
- količina nosača zapisa
- tehnika izrade
- mjesto
- autor
- vrsta gradiva
- dostupnost
- vrsta zapisa (tekstualni dokumenti, knjige/tisak, fotografije, crteži, grafike, novine, časopisi, rukopisi, karte i planovi, tehnički nacrti, zvučni zapisi, notni zapisi i dr.)
- vrsta sadržaja
- institucija
- zbirka
- status zapisa – novo (engl. *new*), ažurirano (engl. *update*)
- za objavu. Postoje dvije kategorije – objavljen i neobjavljen. Ova druga se upotrebljava za gradivo koje je u sklopu projekta digitalizirano, no zbog osobnih podataka, autorskog prava i sl. ne može biti bezuvjetno javno dostupno online.
- uvjeti korištenja. Obuhvaća sljedeće kategorije: zaštićeno autorskim pravom-EU djelo siroče, dopuštena upotreba u obrazovne svrhe, dopuštena nekomercijalna upotreba, nositelji prava se ne mogu identificirati ili locirati; nije zaštićeno autorskim pravom-ugovorna ograničenja, samo nekomercijalna upotreba, ostala poznata zakonska

ograničenja, Sjedinjene Američke Države; autorsko pravo nije istraženo, nije utvrđeno; nema poznatog autorskog prava; imenovanje, dijeli pod istim uvjetima, bez prerada, nekomercijalno: dijeli pod istim uvjetima, bez prerada; javno dobro. Ovdje valja napomenuti da se najčešće koristi oznaka javno dobro.

Tablica je iznimno korisna operateru na skeneru jer u svakom trenutku može provjeriti količinu gradiva, koje gradivo pripada kojoj jedinici i druge potencijalno korisne informacije.

3.5. Edukacija zaposlenika

Odabrani zaposlenici pohađali su edukaciju koju je provodio Ericsson Nikola Tesla d.d., tvrtka koja je izrađivala sustav eKultura, u okviru koje su upoznati s procesom digitalizacije, ponajprije s radom u sustavima Goobi i Središnjem sustavu. Goobi se koristi za povezivanje metapodataka sa skeniranim gradivom, sve se pakira u informacijske pakete te se oni potom prenose u Središnji sustav preko kojeg će se sa završetkom projekta digitaliziranom gradivu moći javno pristupiti.

Svi ostali zaposlenici zaduženi za rad na digitalizaciji prošli su obuku unutar ustanove. U okviru obuke upoznati su sa spomenutim dvama informacijskim sustavima. Najvažniji dio obuke bio je upoznavanje sa samim postupkom skeniranja, s načinom na koji se gradivo priprema za snimanje, s postupkom kontrole kvalitete snimaka te s načinom popunjavanja tablica s metapodacima. Obuku su provodili djelatnici Hrvatskog državnog arhiva odgovorni za projekt od kojih su neki prošli obuku u tvrtki Ericsson Nikola Tesla.

3.6. Priprema gradiva za digitalizaciju

Za studente je priprema gradiva za digitalizaciju bila prvo što su radili jer su prethodna tri koraka bila nadležnost djelatnika arhiva i nekih vanjskih osoba. Dio gradiva bio je pripremljen prije dolaska studenata te se s njim moglo odmah po dolasku studenata krenuti s digitalizacijom, dok su drugi dio gradiva pripremali studenti prema uputama voditelja. Studenti su bili raspodijeljeni na način da su neki skenirali, neki pripremali gradivo, neki provodili kontrolu kvalitete po završetku snimanja, a neki dopunjavali tablice s metapodacima. U pravilu je većina njih naizmjenice radila gotovo sve te poslove. Naravno, i taj je proces dokumentiran.

3.6.1. Dnevnik pripreme gradiva za digitalizaciju

Slike 5a i 5b prikazuju jedan takav konkretan dnevnik pripreme gradiva za digitalizaciju.

Dnevnik pripreme gradiva za digitalizaciju

DNEVNIK PRIPREME GRADIVA ZA DIGITALIZACIJU					
OPĆI PODATCI					
Odjel :	Odjel za zaštitu i obradu arhivskog gradiva				
Odsjek:	Odsjek za arhivsko gradivo 1945.-1990.				
Pripremio za snimanje:	Ime i prezime:	Kontakt podaci:			
	Studenti: Lena Dančuo, [redacted]	tel./mob.: [redacted]			
	[redacted]	e-mail: [redacted]@arhiv.hr			
Kontrolirao pripremu:	Ime i prezime:	Kontakt podaci:			
	[redacted]	tel./mob.: [redacted]			
	[redacted]	e-mail: [redacted]@arhiv.hr			
Trajanje pripreme:	Započeto: 17.5.2021.; 1.7.2022.		Završeno: 15.7.2021.; 22.7.2022.		
Trajanje kontrole pripreme:	Započeto: 5.7.2021., 25.7.2022.		Završeno: 21.7.2021., 29.7.2022.		
OPĆI PODATCI O GRADIVU					
Signatura (oznaka):	HR-HDA-306.				
Naziv (naslov):	Zemaljska komisija za utvrđivanje zločina okupatora i njihovih pomagača (ZKRZ) 2.2.1. Kartoteka i popisi žrtava 2.2.2. Kartoteka zločinaca				
Razdoblje nastanka (od – do):	1944-1947. (1950.)				
Količina izvornika:	12 d/m:				
	2.2.1. Kartoteka i popisi žrtava, 6 d/m (5,6 d/m kartica i 4 arhivske kutije)				
	2.2.2. Kartoteka zločinaca, 6 d/m (27.616 kartica)				
	Tehničke jedinice				
	Vrsta:	Količina:			
	<input type="checkbox"/> arhivske kutije	4 arhivske kutije			
	<input type="checkbox"/> knjige				
	<input type="checkbox"/> arhivske mape				
<input type="checkbox"/> fotografski albumi					
<input type="checkbox"/> kartice	11,6 d/m				
<input type="checkbox"/> _____					
PODATCI O PRIPREMI GRADIVA ZA SNIMANJE					
Broj pripremljenih arhivskih jedinica:	4 arhivske kutije i 11,6 d/m kartica				
Specifikacija pripremljenih predložaka					
Format	Broj predložaka prema vrsti				Ukupno predložaka
	Neuvezani listovi	Knjige / uvezane jed.	Fotografije	<input type="checkbox"/> _____	
<input type="checkbox"/> < A4	6.200				6.200
<input type="checkbox"/> A4 – A3					
<input type="checkbox"/> A3 – A2					

Slika 5a. Dnevnik pripreme gradiva za digitalizaciju – prvi dio.

Ustupio Hrvatski državni arhiv – Odsjek za arhivsko gradivo 1945.-1990.

<input type="checkbox"/> > A2					
<input type="checkbox"/> kartica žrtve (papir 18,4x10 cm)	44.200				44.200
<input type="checkbox"/> kartica zločinci (karton 20x14,5 cm)	27.600				27.600
Ukupno predložaka (očekivani broj snimaka):	78.500				78.500
DNEVNIK PRIPREME					
Opis postupka	Obavljena priprema	Obavljena kontrola	Napomena o pripremi (opisati specifičnosti, izuzetke od pravila, navesti informacije koje mogu biti korisne operateru prilikom snimanja gradiva)		
Provjera stupnja središtenosti i strukture arhivskih jedinica	DA	DA	U potpunosti sređen fond		
Formiranje tehničkih jedinica	DA	DA	Tehničke jedinice su formirane		
Uređivanje tehničkih jedinica (omota)	DA	DA	Sve tehničke jedinice su u omotima na kojima je ispisana signatura Bijeli omot na kojem je ispisana arhivska signatura HR-HDA-306		
Označavanje/izdvajanje multiplikata dokumenata	DA	DA	Multiplikati su odvojeni		
Označavanje tehničkih jedinica	<input type="checkbox"/>	<input type="checkbox"/>			
Poredak za snimanje (provjera, uspostava)	DA	DA	Kartice i popis su poredani redom za snimanje.		
Numeriranje stranica	NE	DA	Kartice nisu numerirane, poredane su abecednim redom i tim se poretком mogu snimiti		
Izrada popisa u Excel obrascu – Arhivske jedinice (list 1)	DA	DA			
Dnevnik sastavio (ime i prezime, potpis): XXXXXXXXXX					

Slika 5b. Dnevnik pripreme gradiva za digitalizaciju – drugi dio.

Ustupio Hrvatski državni arhiv – Odsjek za arhivsko gradivo 1945.-1990.

Osim općih podataka o odsjeku i o gradivu (signatura, naslov, razdoblje nastanka i količina izvornika), dnevnik pripreme uključuje podatke o pripremi gradiva kao što su broj pripremljenih arhivskih jedinica izražen u količini kutija i dužnim metrima, a nakon toga se specificira format i točan broj predložaka. Zatim se opisuju provedeni postupci kao što su formiranje tehničkih jedinica, uklanjanje multiplikata, uspostavljanje točnog poretka zapisa, izrada popisa u Excel tablici i sl. Navodi se je li za te stavke obavljena priprema i kontrola te se po potrebi dodaju i napomene u kojima se opisuju specifičnosti, izuzetci od pravila, informacije od koristi operateru prilikom snimanja i dr.

3.6.2. Konkretni postupci u pripremi gradiva

U nastavku je dodatno detaljnije opisano na koji se način pripremalo gradivo za snimanje. Prije skeniranja bilo je potrebno obaviti niz radnji.

Najprije se provjeravaju tablice s metapodacima u MS Excelu koje sadrže identifikator, opis, sadržaj (odgovara li broj stranica, opis sadržaja predmeta) i ostale podatke o jedinicama gradiva. Potom se provjerava na omotu svakog predmeta je li naveden ispravan identifikator – ako to nije slučaj, treba ga naknadno istaknuti jer se prema njemu određuje računalna mapa u koju će se pojedine snimke spremati i kasnije povezati s metapodacima o arhivskoj jedinici kojoj datoteke pripadaju. Imajući na umu format, starost i opće stanje gradiva, potrebno je u skladu s tim odrediti vrstu skenera na kojem će se snimati pojedino gradivo. Na primjer, HDA u sklopu ovog projekta raspolaže protočnim skenerom čiji je najveći podržani format A3. Za snimanje gradiva većeg formata koristi se široki skener za velike formate koji se u ovom slučaju većinom koristio za karte i nacрте. Na mjesto gdje je po redoslijedu bilo gradivo većeg formata, koje se nije moglo snimiti na određenom skeneru, skeniran je prazan list u boji. Njegova se snimka naknadno mogla vrlo lako uočiti te se po završetku skeniranja ta datoteka samo zamijenila datotekom tog gradiva snimljenog na drugom skeneru. Razlog ovog postupka je taj što datoteke dobivaju automatski generirana imena u numeričkom nizu te bi ih bilo nadasve komplicirano naknadno mijenjati. To bi ujedno i otvaralo mjesto za potencijalne pogreške. K tome nužno je pregledati gradivo „list po list“ i skinuti spajalice jer one mogu oštetiti skener i poderati papir. Obavezno je ukloniti i bilo što drugo što bi moglo oštetiti ili skener ili gradivo. Također treba ukloniti duplikate, a u slučaju da greškom nisu uklonjeni, naknadno ih ukloniti iz gradiva te izbrisati njihove snimke. Sve navedene postupke neophodno je evidentirati u popratnu dokumentaciju postupka kontrole kvalitete.

Ako je papir oštećen, pogotovo na rubovima gdje ga valjci uvlače u protočni skener, potrebno ga je minimalno obrezati (ako je moguće) da bi se spriječilo dodatno oštećenje. Druga je mogućnost te listove pažljivo zasebno umetati u skener. Također je moguće skener podesiti na opciju namijenjenu za vrlo tanke papire (engl. *extra thin paper*) koji ih tada malo sporije uvlači i ujedno umanjuje rizik od daljnjeg oštećenja. Ako je gradivo oštećeno u tolikoj mjeri da postoji veliki rizik da će ga skener poderati prilikom uvlačenja, preporuča se kopirati ga na fotokopirnom uređaju. Nakon toga je potrebno isprintati kopiju i potom nju skenirati. Tako se izvorNIK štiti od daljnjeg oštećenja. Kopija se skenira i prilaže u kutiju uz izvorNIK kao dokaz da odražava izvorNIK. Ako je pak gradivo pljesnivo ili znatno oštećeno na način da je ozbiljno narušena čitljivost i sam oblik lista, proslijeđuje se konzervatorima na restauraciju.

Važno je ustvrditi prevladavaju li papiri sa sadržajem otisnutim jednostrano ili obostrano. Ako je veći dio gradiva otisnut jednostrano, prilikom pripreme treba naznačiti pojedine obostrane stranice jer je u tom slučaju na skeneru potrebno promijeniti postavke skeniranja s jednostranog na obostrano. Ako je veći dio gradiva otisnut obostrano, ne treba gubiti vrijeme s označavanjem jednostranih stranica jer će se tada koristiti funkcija obostranog skeniranja te naknadnom provjerom izbrisati prazne stranice. Postavka skeniranja za uklanjanje praznih stranica (engl. *remove blank pages*) nije uvijek pouzdana jer su kod tanjih papira zbog njihove djelomične prozirnosti slova s jedne strane vidljiva i s druge te skener to automatski prepoznaje kao obostrano. Ta postavka nije korisna ni kod papira koji na poleđini imaju razne mrlje ili pak kod pečata koji su često vidljivi i s druge strane papira. Kada govorimo o papiru u boji, njih skener nikada ne prepoznaje kao prazne listove. Upravo zato se pokazalo da navedenu opciju nije preporučljivo stalno koristiti i pritom prividno uštedjeti vrijeme ukidanjem pregledavanja radi li se o papirima sa sadržajem otisnutim s jedne ili s obje strane. Znatno bi više vremena nakon snimanja bilo potrebno za ručno uklanjanje snimki praznih listova. Osim toga, bilo bi vrlo vjerojatno i nezanemarivo mnogo grešaka u njihovoj detekciji s obzirom na to da je u smanjenom prikazu naprosto nemoguće uočiti sve prazne listove iz ranije navedenih razloga, a pregledavanjem snimaka „list po list“ taj bi se proces uvelike usporio. Na taj bi se način također stavljal preveliko opterećenje na osobu koja skenira, kojoj bi produktivnost glede snimanja u velikoj mjeri opala. Razlog tome bio bi taj što bi ona trošila previše vremena na naknadnu kontrolu snimki prije njihovog spremanja umjesto da to radi osoba koja je ionako zadužena za pripremu gradiva.

Bilo bi poželjno označiti i gradivo većeg/manjeg formata unutar predmeta u kojem prevladava standardni A4 format zbog dodatnog opreza prilikom umetanja u skener i njegovih valjka koji ga uvlače. Jedan od načina na koje je označavano gradivo bilo je upotrebom samoljepivih *post-it* papirića pogodnih za arhivsko gradivo jer njihovo ljepilo na gradivu ne ostavlja nikakve tragove nakon odljepljivanja.

Poželjno je provjeriti redoslijed listova i njihovu orijentaciju. To je bitno zbog automatski generiranog naziva svake pojedine slikovne datoteke jer ga je naknadno komplicirano mijenjati, naročito kod velike količine snimaka.

Savijene rubove i zgužvane listove bi bilo poželjno ispraviti, pogotovo na dijelu gdje ih valjci uvlače u skener. Tako se postupa da bi se smanjio rizik od oštećenja te da bi snimka odražavala original i izgledala atraktivno.

U gradivu se ponekad nalaze interne bilješke i slični materijali koji nisu arhivsko gradivo pa se samim tim ne skeniraju. Njih je potrebno ukloniti.

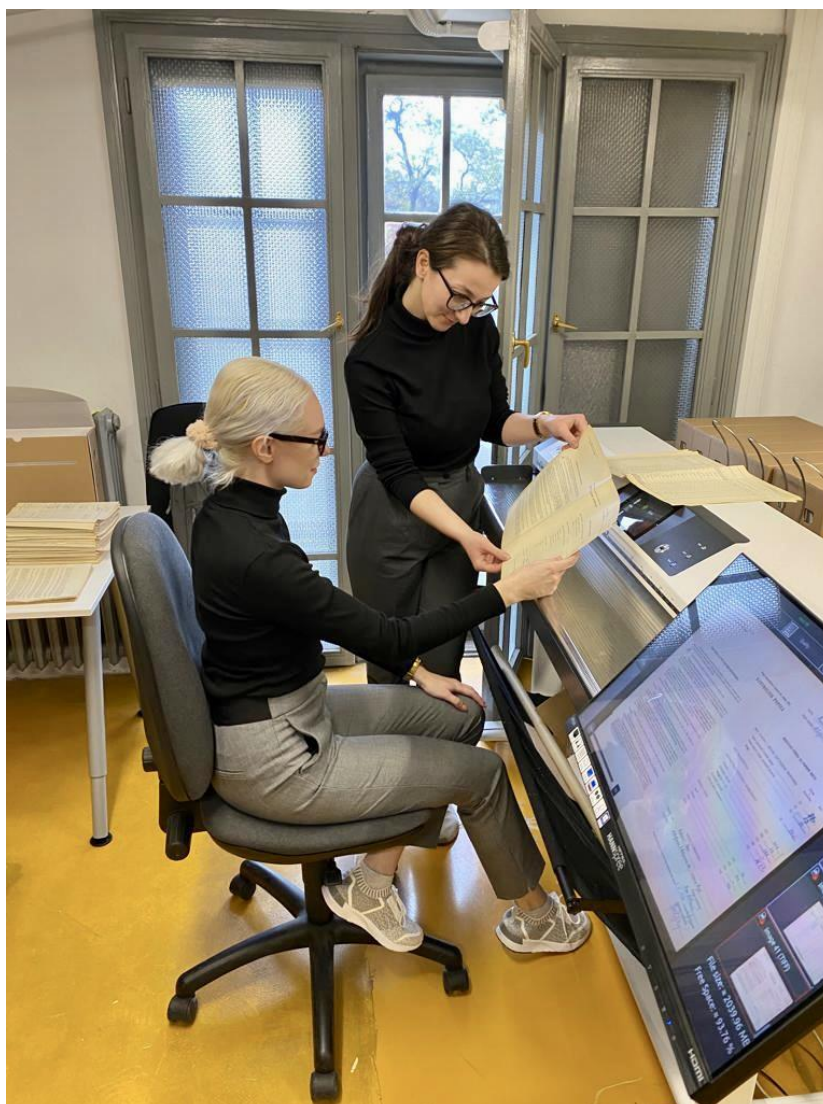
Nadalje, treba izdvojiti gradivo nestandardnog oblika koje se ne može skenirati te naći alternativan način digitalizacije tog gradiva. Odvojiti papire ako su se „zalijepili“ jedan za drugi. U protivnom će ih skener zajedno uvući i javiti grešku.

U slučaju da su na list papira namjerno zalijepljeni dodatni listovi (zbog uštede papira se često prije znalo naprosto prelijepiti dio teksta komadom papira na kojem je bio natipkan novi tekst), skener ih neće uvući jer automatski detektira uvlačenje dva ili više listova u isto vrijeme. Moguće je isključiti tu opciju, no obavezno je treba vratiti odmah nakon skeniranja spornog lista jer se ovdje radi o vrlo važnoj postavki.

Zaključno u ovom poglavlju valja istaknuti da je priprema gradiva za snimanje korak od iznimne važnosti u procesu digitalizacije jer kvalitetna i temeljita priprema rezultira lakšim i „bezbolnijim“ snimanjem. Ona također na minimum svodi potrebu za ponovnim snimanjem već snimljenog gradiva uslijed grešaka, oštećenje gradiva pri prolasku kroz skener te potencijalno oštećenje opreme za skeniranje zbog stranih objekata nehotice ubačenih u skener. Upravo zbog naznačenih razloga, ovaj se korak nipošto ne bi smjelo preskočiti niti bi se na njemu trebalo „štedjeti“ vrijeme. U praksi se pokazalo da površna priprema ili, još gore, njezin izostanak naposljetku kao posljedicu ima znatno veću količinu naknadno utrošenog vremena te, potencijalno, čak i financijskih sredstava zbog mogućih defekata na opremi te spram paradigme „vrijeme = novac“.

3.7. Snimanje

Pripremljeno gradivo napokon je spremno za snimanje. Snimanje su većinom odrađivali studenti prema uputama njihovih voditelja (slika 6) i stručnjaka iz središnjeg fotolaboratorija. Kada je riječ o formatima snimki, u Smjernicama za digitalizaciju kulturne baštine (2020) sastavljenim za potrebe projekta eKultura, zapravo su preuzete preporuke iz već citiranih dokumenata – *FADGI*, *ISO/TR* i *Metamorfoze*. Iz tog razloga neće biti ovdje opširno razrađene.



Slika 6. Studentica za skenerom uz podršku koordinatorice.
Preuzeto iz vlastite galerije fotografija.

3.7.1. Arhivska i prezentacijska datoteka

Ukratko, postoje dvije vrste snimki, arhivska datoteka i prezentacijska datoteka. Formati za svaku su korišteni u skladu sa smjericama, a ovisili su o vrsti i veličini gradiva. Najviše snimki snimljeno je na protočnom skeneru, što je i bilo za očekivati. Na njemu se snimalo novije gradivo koje je dobro očuvano i u kojem prevladava standardni A4 format. Na dugačkom skeneru namijenjenom za velike formate (slika 6) snimale su se karte, nacrti i gradivo veće od A3 formata. Sve ostalo snimalo se odozgo na koračnim skenerima, a to su knjige i stariji zapisi koji su preosjetljivi za protočni skener.

3.7.2. Imenovanje datoteka

Prije skeniranja dogovoren je točan način imenovanja datoteka koji je morao odgovarati sustavima u koje se gradivo kasnije učitavalo, što znači da je ime datoteka uključivalo isključivo brojeve i simbole koji se toleriraju u nazivu datoteke, ali nikako slova. Prve četiri znamenke stajale su za broj fonda, s tim da su sve znamenke morale biti iskorištene, a to se postizalo vodećim nulama. Primjerice fond 25 bilježio bi se kao 0025. Nakon broja fonda se uzeo u obzir okvirni broj budućih snimki u fondu te se na temelju njega određivao broj znamenki, s tim da je uvijek na početku trebala biti barem jedna vodeća nula kako bi se snimke automatski uspješno poredale. Ako fond sadrži količinu između 100.000 i 999.999 snimki, potreban broj znamenki je najmanje sedam tako da bi prvi broj snimke u tom fondu bio zapisan kao 0000001. Broj fonda i broj snimke odijeljeni su donjom crtom. Puni naziv neke snimke bio bi: 1616_0003684.

Bilo je slučajeva kada su se između dva naziva u nizu, npr. snimke „1616_0003684“ i „1616_0003685“ trebale dodati jedna ili više snimki jer je greškom bilo propušteno snimiti obostrano otisnut list, jer se naknadno pronašlo gradivo za umetanje i sl. Taj se problem rješavao dodavanjem tzv. podbroja koji je dolazio na sam kraj imena datoteke, također po principu da mora biti onoliko znamenki da ispred bude barem jedna vodeća nula. Na primjer, ako se između ove dvije snimke trebalo umetnuti pet novih datoteka, na sam kraj prve datoteke dodaje se donja crta i podbroj od (barem) dvije znamenke pa bi imena tih novih datoteka glasila: 1616_0003684_01, ..._02, ..._03, ..._04 i ..._05. Na taj je način između dvije datoteke moguće umetnuti naprosto beskonačno mnogo snimki ako bi to bilo potrebno.

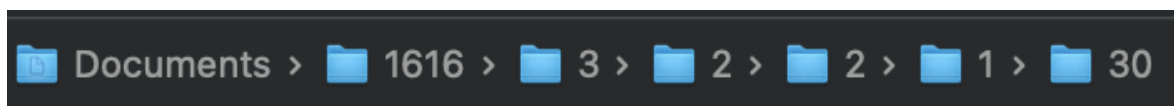
3.7.3. Tehničke specifikacije, postupci i pohrana

Gradivo je skenirano uvažavajući Smjernice za digitalizaciju kulturne baštine (2020) pri čemu je većina snimljena u TIFF formatu koji preporučuju sve već spomenute relevantne međunarodne smjernice i norme. Rezolucija nije smjela biti manja od 300 točaka po inču (engl. *dots per inch*; dpi), a dubina boje je morala biti 24-bitna. U softveru Goobi u koji se kasnije učitavaju slike, podešeno je da ne prihvaća snimke rezolucije manje od 300 dpi, one koje nisu propisanog formata i sl. Na taj je način osigurano da se snimke koje ne zadovoljavaju standarde uopće ne mogu učitati u sustav.

Na koračnim skenerima koji snimaju odozgo i protočnom skeneru za velike formate, snimke su snimljene s crnim obrubom debljine cca. 1 cm koji na taj način garantira da je zapis skeniran

u cijelosti i da su mu potpuno vidljivi svi rubovi. Novije gradivo formata A4 snimljeno na protočnom skeneru u većini slučajeva nema margina. Također, s obzirom na to da je najveća količina snimljenog gradiva snimljena na protočnom skeneru koji je najbrži, procijenjeno je da je za pretežito crno-bijele zapise u A4 formatu s tekstom natipkanim pisaćom mašinom sasvim dovoljan JPEG format u rezoluciji 300 dpi. On kreira datoteke veličine i do 50 puta manje u odnosu na one u TIFF formatu što znatno manje opterećuje kapacitete za pohranu Hrvatskog državnog arhiva.

Datoteke se pohranjuju u interni sustav ustanove spojen na sva računala koji omogućuje brz i jednostavan pristup datotekama i praktično upravljanje. Spremale su se na način da putanja direktorija odražava hijerarhiju arhivskih jedinica izvornog gradiva (vidi sliku 7). Npr. prvo je izrađen direktorij s nazivom fonda, onda su u njemu izrađeni direktoriji ovisno o tome koliko serija ima u fondu pa onda u svakom tom direktoriju budu i podserije itd. – ovisno o tome do koje jedinice je izvorno gradivo razrađeno.



Slika 7. Primjer ilustracije putanje direktorija.

Taj način pohrane gradiva pruža sistematičnu i vizualno jasnu pohranu digitaliziranog gradiva te omogućava njegovo jednostavno pronalaženje u kasnijoj fazi korištenja.

3.8. Kontrola kvalitete

Nakon što je gradivo skenirano i spremljeno u odgovarajuće foldere sukladno njegovoj signaturi, kreće postupak kontrole kvalitete. On je neophodan prije samog arhiviranja datoteka i objavljivanja za javnost jer ne možemo isključiti mogućnost nekog propusta osobe koja je skenirala ili greške u radu same opreme, u ovom slučaju skenera. Kontrolu su radili i djelatnici/arhivisti i studenti s tim da onaj tko je skenirao pojedino gradivo nije ga mogao ujedno i kontrolirati, već je to radila neka druga osoba. Korištene su evidencijske liste namijenjene kontroli kvalitete koje nisu bile strogo definirane, no sadržale su otprilike sljedeće podatke: ime fonda, signatura, naziv datoteke (tu je mogla biti samo jedna, dvije ili pak raspon datoteka), opis problema, ime i prezime osobe koja je pregledala pojedine snimke i zamijetila

nepravilnosti, datum pregledavanja, datum te ime i prezime osobe koja je ispravila naglašene nedostatke, kao i opis postupka na koji je način problem riješen ili pak zašto nije.

3.8.1. Stavke zastupljene u kontroli kvalitete

Snimke su uglavnom kontrolirane jedna po jedna, te su u pravilu u evidencijsku listu bile unošene samo one datoteke u kojima je uočena neka nepravilnost. Stavke koje su se pregledavale u ovom procesu bile su:

- ako je gradivo numerirano (naveden broj stranice pri dnu lista), ide li numeracija po redu i nedostaje li negdje možda neka stranica. Ako je ustanovljeno da je tek nekoliko stranica snimljeno obrnutim redoslijedom, problem se mogao riješiti na vrlo jednostavan način, odnosno preimenovanjem tih datoteka kako bi naposljetku sjele na svoje mjesto. Iako ih već nije bilo potrebno ponovno skenirati, bilo ih je potrebno i fizički staviti na pravo mjesto kako bi digitalizirano gradivo u potpunosti odražavalo fizičko gradivo. Ako je ustanovljeno da su stranice jako ispremiježane, ručno ih se ispravno posložilo u kutiji te skeniralo ponovno jer bi u tom slučaju predugo trajalo sve ih ručno preimenovati.

Bilo je slučajeva kada je uočena rupa u numeraciji, a imala je tri moguća scenarija: papir je bio otisnut obostrano, a skeniranje je bilo podešeno na jednostrano pa je naknadno skenirana greškom izostavljena stranica i spremljena pod podbrojem ili su sve druge datoteke pomaknute za jedan broj više; greška u izvornoj numeraciji koja čak i nije tako rijedak slučaj; te naprosto slučaj kada je neki list ili izgubljen ili uopće nije predan arhivu.

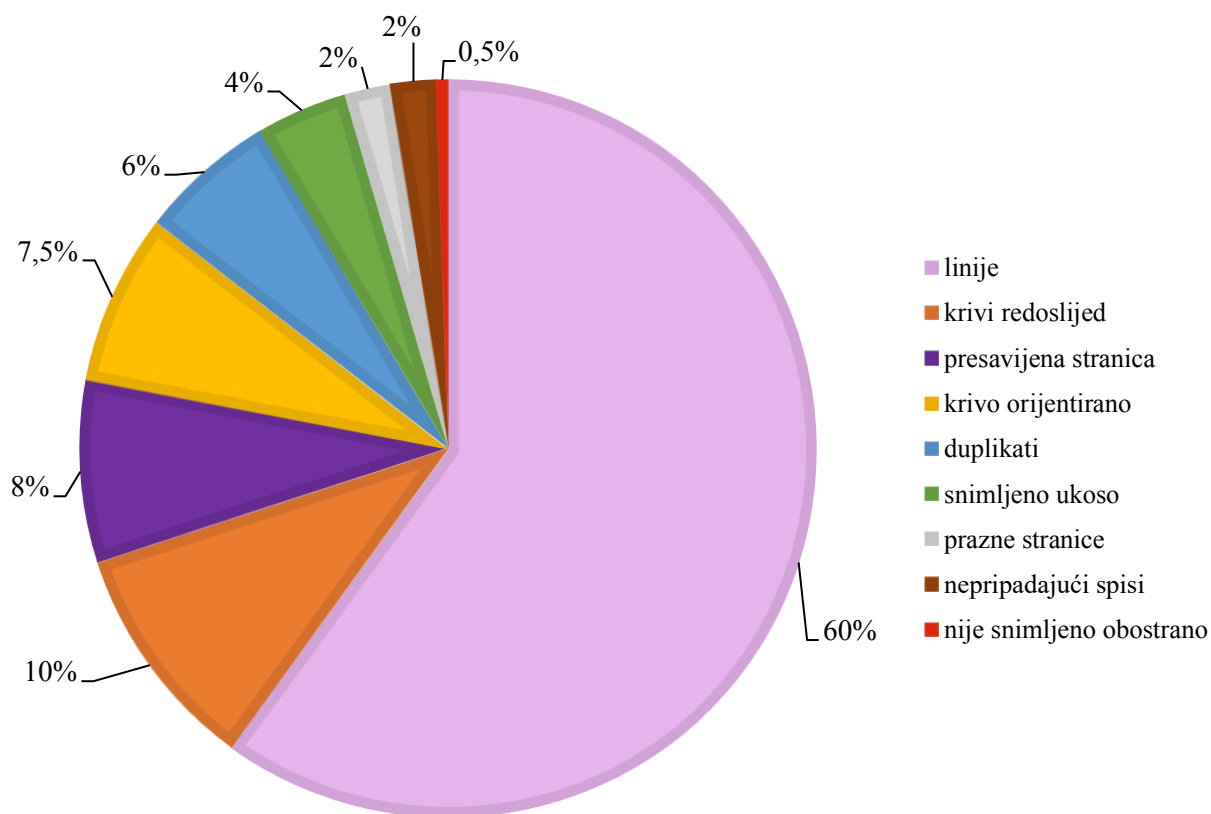
- provjeriti jesu li datoteke ispravno orijentirane. Ovo se također moglo lako ispraviti bez ponovnog skeniranja zahvaljujući toj funkciji u softveru, no svejedno je evidentirano radi transparentnosti jer bilo kakva izmjena na slici mijenja datum zadnje izmjene koji pripada metapodacima o gradivu.
- jesu li stranice presavijene na rubovima ili zgužvane na način da negativno utječu na čitljivost teksta. U ovom se slučaju gradivo prvo ispravlja te se onda ponovno skenira.
- nalazi li se unutar fonda/zbirke/podfonda/serije/podserije gradivo koje tamo ne pripada. Ono se briše iz direktorija, fizički uklanja iz kutija te smješta na adekvatno mjesto i ponovno se skenira ako pripada fondu koji je obuhvaćen planom za digitalizaciju u sklopu projekta.

- uočiti nečitljiv tekst. Tinta na pojedinom gradivu može jako izbljedjeti uslijed starosti, nekvalitetne tinte ili prevelike izloženosti svjetlosti te je u tom slučaju dopušteno, i čak preporučljivo, digitalno pojačati kontrast i time znatno poboljšati čitljivost. Ovo je primjer dopuštenog postupka koji tolerira odstupanje od izvornika u određenoj mjeri iz razloga što ovdje prioritet ima čitljivost i iskoristivost gradiva, no vrlo je bitno taj proces dokumentirati.
- detektirati osobne i/ili povjerljive podatke. Takvi se podaci ne smiju objaviti pa se oni zacrnjuju i time više nisu čitljivi (prezentacijska datoteka) što zapravo podrazumijeva postupak anonimizacije. Arhiv za sebe čuva i neanonimiziranu datoteku (arhivska master datoteka).
- izdvojiti snimke snimljene više od jednog stupnja nakoso. Iako se ovo većinom može digitalno ispraviti, preporučljivo je ipak sporne listove snimiti ponovno jer se prilikom softverskog rotiranja gubi ipak gubi (ne)značajan dio lista.
- ustvrditi ako na snimci nije vidljiv čitav tekst. Treba pregledati izvornik i vidjeti je li takav i izvornik (tada se upravo tako navodi u evidencijskom listu pod opisom postupka kako problem (ni)je riješen). Ako se greškom dogodilo da je pri softverskoj obradi izrezan dio teksta, to se gradivo snima ponovno.
- prepoznati strane objekte na snimkama koji se na fizičkom gradivu ne nalaze. Najčešće se ovdje radi o linijama koje ostavlja (pretežno protočni) skener na čije se staklo lijepe sitne čestice prašine kojim gradivo nažalost obiluje ili pak sitni komadići papira. Gradivo koje na digitalnim snimkama sadrži elemente koje fizičko gradivo ne sadrži obavezno se mora snimati ponovno dok se ne eliminiraju svi ti elementi.
- izbrisati snimljene prazne stranice i/ili duplikate. U slučaju da se nalaze i u fizičkom gradivu, maknuti i njih. Postoje slučajevi kada se može procijeniti da je duplikate ipak potrebno sačuvati u fizičkom obliku.

3.8.2. Analiza učestalosti grešaka prilikom skeniranja na protočnom skeneru

Analizirane su tablice s opaskama prilikom provođenja kontrole kvalitete i niže slijedi grafikon s njihovim prikazom. Grafikon 1 prezentira učestalost vrsta grešaka prilikom skeniranja. Iskazani postoci su okvirni te se odnose prvenstveno na analizu grešaka koje su se pojavljivale prilikom uporabe protočnog skenera. Dalje u radu bit će iznesena i usporedba učinkovitosti protočnog i koračnog skenera.

UČESTALOST GREŠAKA PRILIKOM SKENIRANJA



Grafikon 1. Analiza učestalosti vrsta grešaka prilikom skeniranja.

Odmah je na prvi pogled jasno vidljivo da greška koja se uvjerljivo najviše javlja i to sa čak 60% su linije, odnosno bilo kakve anomalije (engl. *artifacts*) na slici nastale prilikom prolaska papira kroz skener ili naknadnom digitalnom obradom u softveru za skeniranje. Linije (vidi sliku 10) su ujedno i najveća mana protočnog skenera, posebice za arhivsko gradivo koje uvijek sadrži daleko više čestica prašine na svojoj površini nego što to sadrži npr. uredski papir. Također, taj se skener ne može koristiti na osjetljivom gradivu jer postoji velika šansa da će ga oštetiti, možda čak i poderati, prilikom uvlačenja i prolaska kroz skener. Ove dvije činjenice razlog su zašto FADGI smjernice (2023) u većini slučajeva ne preporučuju upotrebu protočnih skenera za skeniranje arhivskog gradiva. Ova greška skenera ujedno je i među gorima jer ne postoji druga mogućnost rješavanja problema osim ponovnog skeniranja, ponekad čak i višestrukog.

Prije nastavka interpretacije grafa valja pojasniti da se postoci učestalosti ovih problema ne odnose na sveukupne snimke (linije se ne pojavljuju na 60% skeniranog gradiva), već o

postocima zastupljenosti pojedinih vrsta grešaka ako uopće do grešaka dođe (ako ipak dođe do greške, u 60% slučajeva je izgledno da će se raditi o linijama).

Čak ostalih osam vrsti grešaka raspoređeno je unutar 40% grafa, a najučestalija među njima su krivo poredane stranice sa 10%. Ta je greška najlakše uočljiva kod numeriranih listova, a kod nenumeriranih malo teže. S 8% su presavijene stranice, najčešće neznatno na samim rubovima što bi se čak i moglo tolerirati, no ako prekriva dio teksta, mora se ispraviti i snimiti ponovno. Zatim slijede krivo orijentirane stranice (7,5%) i snimljeni duplikati (6%). Krivo orijentirane stranice mogu biti greška skenera ili operatera za skenerom, ovisno o tome je li uključena opcija za automatsko prepoznavanje ispravne orijentacije slike pa ih sustav nije točno detektirao ili ih je operater zaboravio/propustio ispravno okrenuti prije spremanja. Duplikate je često teško uočiti prilikom pripreme gradiva za skeniranje, a lakše tijekom kontrole kvalitete snimaka. Tada se naknadno brišu snimke i miču fizičke kopije iz gradiva. Ove dvije greške nisu toliko ozbiljne jer se mogu lako riješiti bez ponovnog skeniranja. U 4% slučajeva kada je došlo do neke greške, skener blago ukrivo uvuče papir te snimka bude ukoso što nije vizualno atraktivno. Duplo rjeđe (2%) dogodi se da se skeniraju prazne stranice i da budu pronađeni spisi koji ne pripadaju tamo gdje su pohranjeni. Opcija na skeneru za preskakanje praznih stranica ne funkcionira uvijek najbolje te se onda događa da i prazne stranice budu snimljene. Zbog toga se prije spremanja snimki uvijek pregleda ima li snimljenih praznih stranica, ali rijetko i ovdje bude propusta koji se u naknadnoj kontroli isprave. Najrjeđa greška sa 0,5% jest kada se ne snimi otisnuta poledina lista, odnosno kada je skener greškom podešen na jednostrano skeniranje. To se najjednostavnije može uočiti kod numeriranih stranica u kojima stranice evidentno nedostaju. Kod onih koje nisu numerirane se uglavnom može prepoznati tako da sadržaj snimke završava usred rečenice, a sljedećoj nije njegov nastavak. Izostavljena stranica se skenira i pohranjuje na odgovarajuće mjesto, većinom pod podbrojem (postupak opisan u cjelini 3.7.2).

Iako je kontrola kvalitete dugotrajna i ponekad zamorna, ne može se preskočiti jer će se uvijek naći neke pogreške. Vrlo ih je bitno ispraviti ili barem svesti na minimum. U svezi s tim, također se može činiti dugotrajno i zamorno dokumentirati sve procese tijekom kontrole kvalitete što usporava sam proces kontrole snimki i ispravljanja nedostataka, no on je nužan zbog transparentnosti jer se ovdje ipak radi o arhivskom gradivu.

3.9. Arhiviranje

3.9.1. Kriptografski sažetak i tablica za metapodatke

Nakon što je potvrđeno da su snimke gotove i konačne, odnosno spremne za arhiviranje, generira se njihov jedinstveni kriptografski sažetak, za što se trenutno koristi algoritam SHA 256 (engl. *Secure Hash Algorithm - SHA*). Thakkar (2022) na jednostavan način opisuje kakav je to algoritam. Ovaj algoritam koristi 256-bitni ključ koji iz podataka datoteke stvara novi, neprepoznatljivi niz podataka fiksne duljine. Ovaj niz nasumičnih znakova i brojeva, koji se naziva *hash* vrijednost ili kriptografski sažetak, također je veličine 256 bita.

Kriptografski sažetak koristi se za provjeru integriteta datoteka: mijenja se u potpunosti čak i ako je datoteka minimalno izmijenjena pa se na taj način može garantirati da snimke nisu dirane nakon arhiviranja (ako je kriptografski sažetak odgovarajući). U tablicu za metapodatke se zatim ispunjava tzv. list 2 (vidi sliku 8), odnosno list s nazivom „Datoteke“, s dodatnim metapodacima o gradivu poznatim tek nakon snimanja. On će kasnije između ostalog služiti za svrstavanje datoteka u izvorne logičke strukture, odnosno odgovarajuću hijerarhiju.

	A	B	C	D	E	F
1	Identifikator	Slike	Redoslijed SHA-256		Velicina	Putanja
2						
3	HR-HDA-280-1-1-1-1	0280_00000001.jpg	1	33d2149c1a85664258870f63effd325180e02397cb6a3d66577905a357fe676f	166672	\\RS2421\etaza4\Digitalni repozitorij\280\1\1\1\0280_00000001.jpg
4	HR-HDA-280-1-1-1-1	0280_00000002.jpg	2	15c260d4efd78e572baec481e312e660d4461360ee48177eb4faa28478f0670	766570	\\RS2421\etaza4\Digitalni repozitorij\280\1\1\1\0280_00000002.jpg
5	HR-HDA-280-1-1-1-1	0280_00000003.jpg	3	6a5e5f008cbb214a35b99f6e317854cbe741ffcb66e4b86a6ec23a2bbaa1465	619921	\\RS2421\etaza4\Digitalni repozitorij\280\1\1\1\0280_00000003.jpg
6	HR-HDA-280-1-1-1-2	0280_00000004.jpg	1	b53ab810a50e561c8fd7e469c738a639318dd2ed3a86b354ee7d897f4cc16af	153253	\\RS2421\etaza4\Digitalni repozitorij\280\1\1\1\0280_00000004.jpg
7	HR-HDA-280-1-1-1-2	0280_00000005.jpg	2	a9a0e77dbd6e5f5e812045c491ef83a341010572d83611a448179514ee63d67b	818472	\\RS2421\etaza4\Digitalni repozitorij\280\1\1\1\0280_00000005.jpg
8	HR-HDA-280-1-1-1-2	0280_00000006.jpg	3	65c580c67e5b1b210af2ed8678c6373aa266ebfd72b53916bb217f5a0d1f69	582862	\\RS2421\etaza4\Digitalni repozitorij\280\1\1\1\0280_00000006.jpg
9	HR-HDA-280-1-1-1-3	0280_00000007.jpg	1	02d0b2cb938ec491126ae8c6dcda85dabdf5754ae1ba2555d2dc45e1a0fc969	187263	\\RS2421\etaza4\Digitalni repozitorij\280\1\1\1\0280_00000007.jpg
10	HR-HDA-280-1-1-1-3	0280_00000008.jpg	2	154fd8a73786c2b8aa8f89ac520122760ea4e3bce1808cd2eb35509424d473b	745358	\\RS2421\etaza4\Digitalni repozitorij\280\1\1\1\0280_00000008.jpg
11	HR-HDA-280-1-1-1-4	0280_00000009.jpg	1	e08a40e9dd3070709a9eba3d298fa0518b53dc3255af5c565c910c22aee7ad7f	180534	\\RS2421\etaza4\Digitalni repozitorij\280\1\1\1\0280_00000009.jpg
12	HR-HDA-280-1-1-1-4	0280_00000010.jpg	2	14ea3701047f614949144512c4275bf79632aaf66f6ed9cc5dbae7135975771	413834	\\RS2421\etaza4\Digitalni repozitorij\280\1\1\1\0280_00000010.jpg
13	HR-HDA-280-1-1-1-4	0280_00000011.jpg	3	2abbc6c8d25ef00860889730d7327052496ec537bf1cb1f98447305c563682	340363	\\RS2421\etaza4\Digitalni repozitorij\280\1\1\1\0280_00000011.jpg
14	HR-HDA-280-1-1-1-4	0280_00000012.jpg	4	07af1063a9499cd00b2e094564134ad27167abc5e64035c4bf9557f6796e071	579953	\\RS2421\etaza4\Digitalni repozitorij\280\1\1\1\0280_00000012.jpg
15	HR-HDA-280-1-1-1-4	0280_00000013.jpg	5	0a60ea18e5a8815533b938115dc446c415c0449976c8d8a98126ac3da34a9ef5	619581	\\RS2421\etaza4\Digitalni repozitorij\280\1\1\1\0280_00000013.jpg
16	HR-HDA-280-1-1-1-4	0280_00000014.jpg	6	487969474d5877da204b8235bf1b1977cc3446159e5b1dd10bfcc2e2db2645f8	353529	\\RS2421\etaza4\Digitalni repozitorij\280\1\1\1\0280_00000014.jpg
17	HR-HDA-280-1-1-1-5	0280_00000015.jpg	1	1b99722af94cc422fd201cae9988c929c1863a50471506992183ae051344dc3	179336	\\RS2421\etaza4\Digitalni repozitorij\280\1\1\1\0280_00000015.jpg
18	HR-HDA-280-1-1-1-5	0280_00000016.jpg	2	bb67fc149686ee540df3f4ac7652b2b5faebf77aa626639296de9bf721a39d3	868054	\\RS2421\etaza4\Digitalni repozitorij\280\1\1\1\0280_00000016.jpg
19	HR-HDA-280-1-1-1-5	0280_00000017.jpg	3	0a58db8e7e4d98cc0ce0ead50501a84d1c29b947bb5dd1b467d88e26d5315d	881830	\\RS2421\etaza4\Digitalni repozitorij\280\1\1\1\0280_00000017.jpg
20	HR-HDA-280-1-1-1-5	0280_00000018.jpg	4	42cf775377423e81ec3f1e738ba5a156d5348c08c639bd8e83965302921422c	940843	\\RS2421\etaza4\Digitalni repozitorij\280\1\1\1\0280_00000018.jpg
21	HR-HDA-280-1-1-1-5	0280_00000019.jpg	5	6063ab2d70818f2c314a02de917405e6ad91d7bbaf12af9430cf5084ca40dce	938556	\\RS2421\etaza4\Digitalni repozitorij\280\1\1\1\0280_00000019.jpg
22	HR-HDA-280-1-1-1-6	0280_00000020.jpg	1	5e20643b9340c5e668dbaecdc54f8bca21644cc83b668f5b1f6c5dbd190e5de	162702	\\RS2421\etaza4\Digitalni repozitorij\280\1\1\1\0280_00000020.jpg
23	HR-HDA-280-1-1-1-6	0280_00000021.jpg	2	7414811e87af11e44516baef485f5bd1ba44cc38b668f5b1f6c5dbd190e5de	810152	\\RS2421\etaza4\Digitalni repozitorij\280\1\1\1\0280_00000021.jpg
24	HR-HDA-280-1-1-1-6	0280_00000022.jpg	3	881f642cf9523b9f773c6828d63b61ed25a6728f593906c5b4eb6f4ccdd2d2	801518	\\RS2421\etaza4\Digitalni repozitorij\280\1\1\1\0280_00000022.jpg
25	HR-HDA-280-1-1-1-6	0280_00000023.jpg	4	131c4e416139c3a148d22931af4f4ebda5760a13ee3516432e4a24f097b6101	956701	\\RS2421\etaza4\Digitalni repozitorij\280\1\1\1\0280_00000023.jpg
26	HR-HDA-280-1-1-1-6	0280_00000024.jpg	5	ee6d51cf9a41edd49dd7dc3f52c5e9b480625f34f2a62126908d5a1f41c5ca377	813697	\\RS2421\etaza4\Digitalni repozitorij\280\1\1\1\0280_00000024.jpg
27	HR-HDA-280-1-1-1-6	0280_00000025.jpg	6	8ea1d9377195480105007245f3dd4401cbb9c416e9cac62ca74c2f97d93719	848095	\\RS2421\etaza4\Digitalni repozitorij\280\1\1\1\0280_00000025.jpg
28	HR-HDA-280-1-1-1-7	0280_00000026.jpg	1	23a0e3058fe773cb413392399a9349cd190e01f8be33294c2b1d2b47af87959	197315	\\RS2421\etaza4\Digitalni repozitorij\280\1\1\1\0280_00000026.jpg
29	HR-HDA-280-1-1-1-7	0280_00000027.jpg	2	5e1135378449892ee066266cd5b16e6154f2c5e9b3b40439a799dabf2180942	792363	\\RS2421\etaza4\Digitalni repozitorij\280\1\1\1\0280_00000027.jpg
30	HR-HDA-280-1-1-1-7	0280_00000028.jpg	3	ee33ce675a1e90d1735e7783d97e382a62378411bd93f31e84b92487ec94fca	860735	\\RS2421\etaza4\Digitalni repozitorij\280\1\1\1\0280_00000028.jpg
31	HR-HDA-280-1-1-1-7	0280_00000029.jpg	4	6d337a36ba62f6b9aaff5bac20256367a99e7dcd835f50509af9c548ab15dc5	869303	\\RS2421\etaza4\Digitalni repozitorij\280\1\1\1\0280_00000029.jpg
32	HR-HDA-280-1-1-1-8	0280_00000030.jpg	1	dfd641188e73718ad5efc93d56c1794221bfaa520c514e2df1752361c74a371	182127	\\RS2421\etaza4\Digitalni repozitorij\280\1\1\1\0280_00000030.jpg
33	HR-HDA-280-1-1-1-8	0280_00000031.jpg	2	d0f6e68747519a75c34d2e8f824d019a3f2e0c2e6bf6a6757e8d92c2d4fa	789621	\\RS2421\etaza4\Digitalni repozitorij\280\1\1\1\0280_00000031.jpg

Slika 8. Tablica s metapodacima – list 2.

Ustupio Hrvatski državni arhiv – Odsjek za arhivsko gradivo 1945.-1990.

Vidljivo je da se u list 2 upisuje svaka datoteka zasebno. Za svaku se datoteku navodi oznaka arhivske jedinice kojoj datoteka pripada, ime datoteke, redoslijed datoteke unutar arhivske

jedinice kojoj pripada, njezin jedinstveni kriptografski sažetak, veličina i putanja. Tablica može prihvatiti i druge metapodatke datoteke i podatke o procesu digitalizacije, no za sada se ti podaci ne upisuju (podaci o postupku digitalizacije bilježe se i čuvaju u aplikaciji Goobi). Nakon što je tablica ispunjena, arhivske datoteke su spremne za trajnu pohranu, a proces može krenuti dalje s objavljivanjem.

3.9.2. Goobi – softverska aplikacija otvorenog koda za projekte digitalizacije

Upute za korisnike – Sustav za podršku tijekom digitalizacije (2020) pojašnjavaju sljedeći korak: „Servis za podršku tijekom digitalizacije unutar projekta omogućava pojedinim djelatnostima zajednice (arhiv, knjižnice, muzeji, ministarstvo...) digitalizaciju kulturne baštine putem Web aplikacije na način da se pojedini kulturni sadržaj digitalizira u središnji informacijski paket (SIP; engl. *Submission Information Package*) koji sadrži metapodatke i multimedijalne komponente i koji se tada šalje u centralni sustav na daljnju obradu. Taj servis temeljen je na Goobi Web aplikaciji koja je prilagođena potrebama pojedinog segmenta zajednice. Svaki proces sastoji se od nekoliko koraka (zadataka) koji se odvijaju slijedno. Pojedini korisnik odrađuje zadatak u određenoj ulozi, ali korisniku također može biti dodijeljeno više korisničkih uloga. Krajnji rezultat svakog pojedinog procesa digitalizacije je kreiran arhivski informacijski paket (AIP; engl. *Archival Information Package*) koji se šalje u središnji sustav na daljnju obradu.“ Za rad na Goobiju nisu bili zaduženi studenti, već tek nekolicina djelatnika Hrvatskog državnog arhiva. U ovim uputama dalje stoji da automatski procesi obuhvaćaju „provjeru kvalitete snimke, automatski izvoz paketa snimanja, registraciju i validaciju datoteka prema uređaju snimanja, učitavanje tehničkih metapodataka datoteka, automatsko spajanje datoteka snimanja s logičkim strukturama, pripremu i automatsko slanje datoteka na OCR servis, kreiranje prezentacijskih datoteka te automatsko postavljanje parametara okoline“.

Kao što je već napomenuto, Goobi može [automatski] izvesti pakete snimki te ih oblikovati kao SIP-ove poštujući prethodno dogovorene standarde za dostavu podataka. Tehničke metapodatke sam učitava iz datoteka, a na temelju tablica s metapodacima, točnije s listom 2 – stranicom „Datoteke“ u Excel tablici, i imenom datoteka spaja datoteke s logičkim strukturama što je iznimno bitno za kasniju objavu na mrežnom portalu.

3.10. Objavljivanje

Nakon što su u Goobiju završeni svi prethodni procesi, posljednji je izvoz podataka. Informacijski paket (IP) može se preuzeti i spremi lokalno i poslati u Središnji sustav eKultura koji je implementiran na infrastrukturi Centra dijeljenih usluga. U Uputama za korisnike – Manualna izrada informacijskih paketa (2020) navedeno je da se jednom kada je IP pohranjen u Središnji sustav, moguće mu je radi ispravaka ili dopune postojećih datoteka ili podataka ažurirati metapodatke, dodati novi javni oblik s datotekama ili nove datoteke u postojeći javni oblik te zamijeniti postojeće datoteke s novima.

Informacijski paket koji je pohranjen u Središnjem sustavu može se objaviti na portalu eKultura i na Europeani, pod uvjetom da je u opisu jedinica gradiva u paketu navedeno da se mogu objaviti. Prvi korak u objavljivanju je tzv. agregacija: postupak pretvaranja podataka o gradivu u oblik koji je prihvatljiv portalu na kojem se podaci žele objaviti. Agregatoru trenutno mogu pristupiti ovlašteni djelatnici arhiva putem poveznice <https://ekultura-test.cdu.gov.hr/aggregator/>. Na slici 9 prikazana je naslovna stranica sustava na kojoj su vidljive trenutne i buduće mogućnosti. Središnja se administracija sastoji od središnjeg sustava u okviru kojeg se može pratiti trajna pohrana digitaliziranog sadržaja, pregledavati katalozi, upravljati IP-ima i sl., zatim od agregatora koji uključuje rad s korisnicima i institucijama te agregaciju sadržaja, od administracije javnog portala koji podrazumijeva uređivanje sadržaja javnog portala, virtualnih izložbi, novosti, e-usluga te naposljetku od samog javnog portala putem kojeg će se moći pregledavati sadržaj.



Slika 9. Naslovna stranica središnje administracije.

Preuzeto iz Uputa za korisnike – Središnja administracija.

U trenutku pisanja rada projekt još traje te cjelokupno gradivo nije još niti digitalizirano niti objavljeno. Međutim, jedan dio je već objavljen te mu se može pristupiti na <https://ekultura.hr>. Iako nije još sav sadržaj objavljen, već je sada vidljiv izgled portala.

3.11. Briga o trajnom očuvanju i dugoročno upravljanje

„Središnji sustav trajno čuva zapis o izvedenim aktivnostima nad informacijskim paketima kojima se pristupa kroz podizbornik:

- Procesi trajnog očuvanja – sadrži zapis o izvedenim procesima trajnog očuvanja kao što je konverzija formata datoteka.
- Interne radnje – interne radnje su akcije (procesi) izvedene od strane lokalnih administratora kao što su premještanje, brisanje, promjena podataka itd.

- Dnevnik aktivnosti – detaljan pregled svih aktivnosti izvedenih u sustavu kao što su prijave u sustav, pregled informacijskih paketa itd. (Upute za globalne administratore – Središnji sustav, 2020)“.

Dalje se u istim uputama navodi da administratori sustava u svrhu uspješnog trajnog očuvanja mogu uređivati dnevnik aktivnosti, voditi evidenciju rizičnih slučajeva, provjeravati kriptografski sažetak, izvoziti AIP-e, konvertirati slike, videozapise i zvuk, obnavljati indekse, provjeravati integritet datoteka te pristupiti tehničkim metapodacima. Osim toga, sustav pruža mogućnost naknadne implementacije novih softverskih funkcionalnosti.

Jedan od segmenata koje valja opširnije spomenuti jest kriptografski sažetak. U Središnjem sustavu moguće je iznova izračunati tu vrijednost bilo da se primjenjuje novi algoritam ili pak ako je narušen integritet dotičnog IP-a. Vrlo bitna funkcionalnost sustava jest i mogućnost konverzije formata jer oni s vremenom zastarijevaju, a konvertirati se mogu slika, zvuk i video. Postoji i mogućnost direktne provjere integriteta koja ispituje cjelovitost IP-a glede potpunosti i nepromijenjenosti, a korisno je to što se može podesiti njezino automatsko periodičko pokretanje. Osim toga, zbog transparentnosti se bilježe detalji o pokrenutim procesima, primjerice brisanje, premještanje, izmjena podataka i dr.

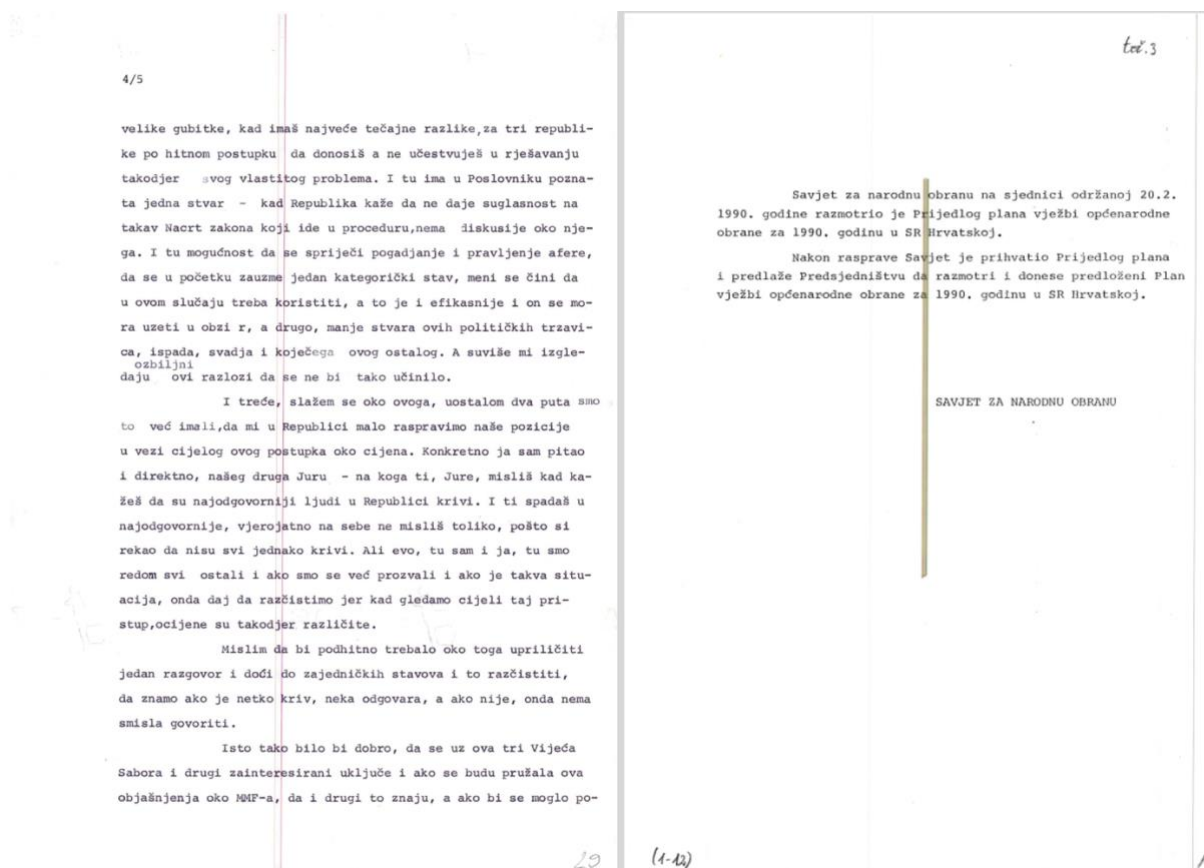
Iz svega prethodno navedenog evidentno je da proces digitalizacije ne staje nakon snimanja gradiva. Proces mora uključiti i digitalizaciju predložaka i digitalizaciju metapodataka, povezati metapodatke i digitalne kopije na način koji omogućuje učinkovito pretraživanje, korištenje i upravljanje digitalnim pojavnim oblicima jedinica gradiva te osiguranje kontinuirane dostupnosti gradiva putem kataloga ili javnog portala. Pored toga, dostupnost gradiva u digitalnom obliku i njegovo dugoročno očuvanje međusobno su povezani: ako gradivo u digitalnom obliku nije dostupno (osobama koje imaju pravo uvida), to se smatra ozbiljnim rizikom za dugoročno očuvanje gradiva, s druge strane, ako se ne provode primjerene aktivnosti dugoročnog očuvanja, to može utjecati na dostupnost gradiva.

4. Kvalitativna i kvantitativna usporedna analiza koračnog i protočnog skenera na arhivskom gradivu

Na gradivu Odsjeka za arhivsko gradivo 1945.-1990. izvršena je usporedna analiza dvaju vrsta skenera – koračnog i protočnog. Naime, usprkos činjenici da FADGI smjernice (2023) protočni skener preporučuju isključivo za novije tekstualne zapise otisnute na modernom uredskom papiru, on je u sklopu projekta korišten za skeniranje dijela gradiva Odsjeka za arhivsko gradivo 1945.-1990. Za takvo gradivo te smjernice preporučuju planetarne, koračne skenere koji snimaju odozgo i plošne skenere. Cilj ovog istraživanja jest utvrditi može li se ipak novije arhivsko gradivo uspješno snimiti na protočnom skeneru te ustanoviti pozitivne i negativne strane obaju skenera. I na jednom i na drugom skeneru snimljen je jednak broj snimaka te su pritom prikupljeni podaci o potrebnom vremenu za te radnje pri čemu je u obzir uzeto i vrijeme za eventualne ispravke i kontrolu kvalitete snimaka.

Na protočnom se skeneru u jednom radnom danu (8 sati; s pauzom od 30 minuta i još nekoliko kraćih pauzi) u prosjeku snimilo 6.000 snimki, dok se na koračnom skeneru snimilo približno 1.500 snimki dnevno što znači da je za isti broj snimaka na koračnom bilo potrebno četiri radna dana, odnosno četiri puta više vremena.

Nakon kontrole kvalitete, ustanovljeno je da je cca. 100-200 snimki snimljenih na protočnom skeneru bilo potrebno ponoviti zbog njihove nezadovoljavajuće kvalitete koja je u većini slučajeva podrazumijevala artefakte, koji su se uglavnom manifestirali kao jedna ili više linija uzduž snimki (vidi sliku 10).



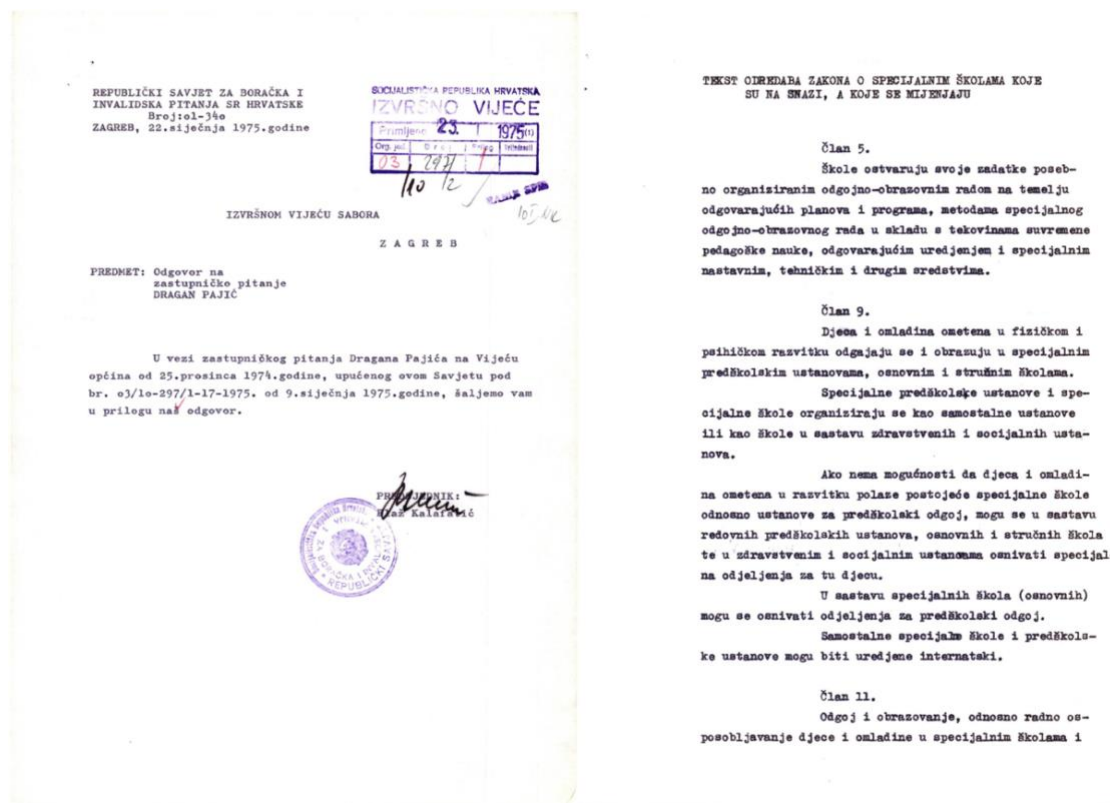
Slika 10. Prikaz artefakata nastalih na protočnom skeneru koji ne zadovoljavaju kriterije kvalitete.

Ustupio Hrvatski državni arhiv – Odsjek za arhivsko gradivo 1945.-1990.

To je ujedno i glavni razlog zašto FADGI (2023) ne preporučuje protočni skener za arhivsko gradivo. Za ponovno snimanje tih spornih snimki utrošeno je od četvrtine do polovine radnog dana. Uz navedeno treba naglasiti da se vrijeme potrebno za ispravke ne može točno odrediti jer ono uvelike ovisi o potrebnom broju ponavljanja skeniranja jednog lista da bi se dobila snimka zadovoljavajuće kvalitete. U to vrijeme ulazi i ponekad vrlo često čišćenje skenera – njegovih kotačića i valjaka od prljavih papira i tinte koja se prenosi na valjke te njegovog stakla koje je iznimno osjetljivo na čestice prašine i komadiće papira zbog kojih naposljetku i nastaju spomenuti artefakti, odnosno anomalije na snimkama. Pri prolasku kroz skener arhivsko se gradivo u nekoj mjeri i trusi te se zato često ti veći ili manji komadići zadržavaju na staklu ili se nakupljaju negdje unutar skenera. To je jedan od razloga zašto arhivsko gradivo nije idealno za protočni skener, već isključivo spisi na modernom uredskom papiru koji se ne trusi te nema oštećenja nastala uslijed starosti. Upravo zbog tih oštećenja, pri brzom prolasku papira kroz skener na otprilike dva papira na 6.000 snimki došlo je do daljnjeg oštećenja koje može biti

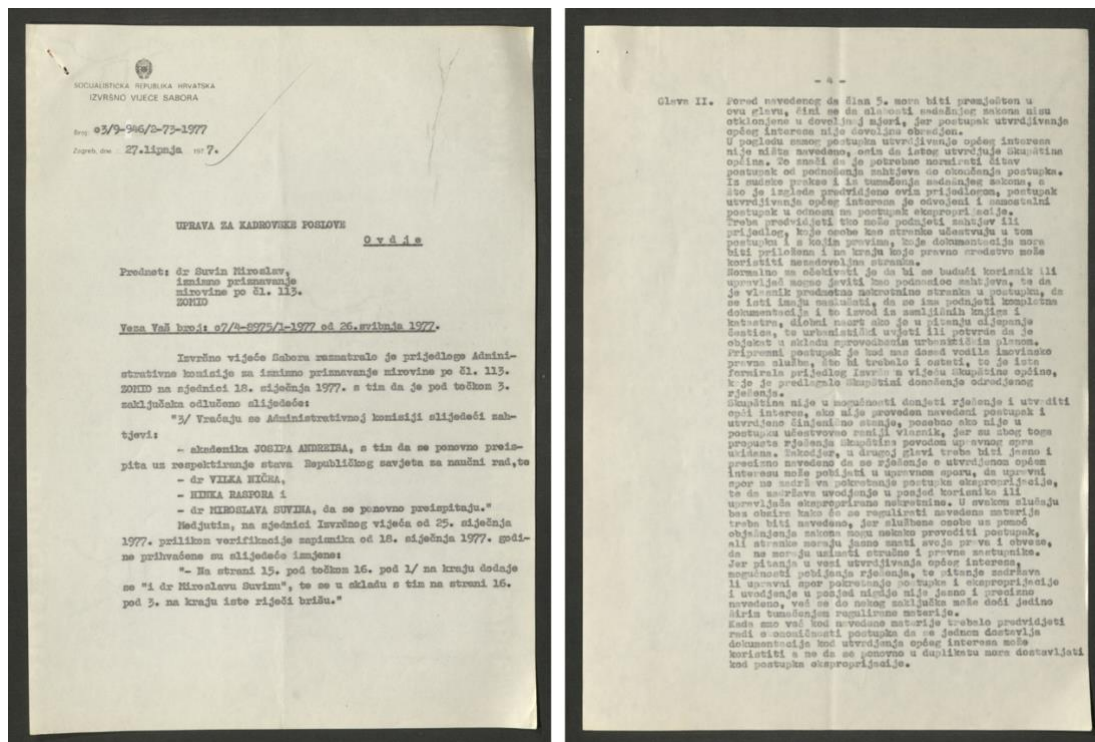
gužvanje, trganje ili pak kompletno kidanje lista te bi ga se zatim trebalo poslati na restauraciju. Ovakva situacija je, naravno, u potpunosti neprihvatljiva kad je riječ o vrijednom starom gradivu koje je jedinstveno, no ono se ionako ne bi niti skeniralo na protočnom skeneru.

No osim mana, protočni skener odlikuju i neke prednosti pri čemu je glavna prednost njegova brzina skeniranja. Ako u obzir uzmemo da je na protočnom za 6.000 snimki zajedno s ispravicima potrebno u najsporijem slučaju jedan i pol radni dan, dok je za tu istu količinu na koračnom potrebno četiri radna dana, dolazimo do zaključka da je protočni brži čak 3,75 puta od koračnog što je za duži projekt znatni broj snimaka više, odnosno manje. Ispravci snimaka snimljenih na koračnom skeneru nisu uzeti u obzir u ovu računicu s obzirom na to da njegove glavne greške – propust snimanja stražnje strane lista, krivi redoslijed, naopako snimljeno, snimljen neki objekt koji nije dio izvornog gradiva npr. ruka, i dr. – nisu greška skenera, već operatera za skenerom. Osim toga, protočni skener bez greške poreda listove upravo onako kako su i bili poredani te nema prostora za nepravilnosti očitovane u tome da gradivo bude naopako ili pogrešno poredano. To je iznimno korisno, naročito za veliku količinu gradiva, a također listovi budu s velikom preciznošću složeni na savršen kup papira iz kojeg ne vire papiri ukoso. Usto, to podrazumijeva da se gradivo može uredno pospremiti natrag u kutije i da se s vremenom neće dodatno oštetiti, odnosno savinuti. Također valja napomenuti da se blago savinuti i neravni listovi automatski uz pomoć valjaka pri prolasku kroz skener ravnaју te su snimke besprijekorne i nisu prisutne nikakve sjene (vidi sliku 11), dok se kod koračnog skenera svaka neravnina mora ručno ispraviti prije skeniranja te usprkos tome snimka nije potpuno bez ikakvih nabora i sjena (vidi sliku 12).



Slika 11. Arhivsko gradivo snimljeno na protočnom skeneru.

Ustupio Hrvatski državni arhiv – Odsjek za arhivsko gradivo 1945.-1990.



Slika 12. Arhivsko gradivo snimljeno na koračnom skeneru.

Ustupio Hrvatski državni arhiv – Odsjek za arhivsko gradivo 1945.-1990.

S jedne strane koračni skener ne uništava gradivo jer kod njega listovi ne prolaze kroz skener nego se samo polažu na površinu za skeniranje. On usto ima i mogućnost snimanja uvezene građe, kao što su knjige, i bilo kakvih dvolisnica u svom izvornom obliku, pri čemu softver razdvaja pojedine stranice u dvije zasebne snimke. S druge pak strane koračni skener ostavlja više prostora za operaterove greške koji u nepažnji može pogrešno složiti listove natrag u kutiju, nehotice ne skenirati neke listove zbog manjka koncentracije ili skenirati objekte koji nisu dio izvornika. Između ostalog, na koračnom se skeneru vide apsolutno sve neravnine na papiru (slika 12) pa su snimke vizualno i estetski manje privlačne u odnosu na protočni skener (slika 11), premda je dopušteno da se ovdje radi o pukom subjektivnom doživljaju.

Iz svega navedenog nameće se zaključak da nema jasnog „pobjednika“ između ova dva skenera jer svaki ima svoje prednosti i mane, kao i svoju namjenu. Na pojedinim je institucijama, kao i projektima da sami odluče je li im prioritet kvaliteta ili kvantiteta, pri čemu je važno istaknuti da korištenje protočnog skenera za arhivsko gradivo dolazi u obzir isključivo za ono novije (otprilike za ono na kojem je sadržaj ispisan pisaćom mašinom ili laserskim printerom). Za starije i daleko osjetljivije gradivo se uvijek koristi koračni skener ili fotoaparatus. Protočni skener pruža nenadmašnu brzinu skeniranja i zadovoljavajuću kvalitetu ako se redovito čisti, održava i kontrolira, no zahtijeva velik napor u nadziranju procesa koji se odvija vrlo brzo i općenito znatan trud da bi snimke ispunjavale standarde kvalitete. Isto tako, relativno često producira snimke s artefaktima koji nisu dopustivi te je ponekad potrebno isti list skenirati više puta da bi se dobila snimka koja zadovoljava smjernice za digitalizaciju. Zbog toga će neki procijeniti da im je bolja opcija skenirati „list po list“ iako će ona u svakom slučaju biti sporija čak i kada se uzme u obzir čišćenje skenera i ponovno snimanje spornih spisa na protočnom skeneru. No gradivo predodređeno za snimanje na protočnom skeneru sa sobom uvijek nosi dulju i temeljitiju pripremu jer je iz gradiva potrebno ukloniti sve što se neće snimati zbog toga što se u skener umeće velika količina gradiva odjednom te je uvezanu građu potrebno razvezati. Na koračnom skeneru nema umjetno dodanih linija uzduž spisa jer gradivo ne prolazi kroz skener te nema rizika od oštećenja gradiva što je za mnoge nepobitni argument kada je riječ o očuvanju arhivskog gradiva.

5. Zaključak

Iz svega navedenog vidljivo je koliko je proces digitalizacije kulturne baštine u arhivima slojevit, dugotrajan, zahtjevan i kompliciran kada se uzmu u obzir svi koraci, kriteriji, norme i pravila koja se moraju ispoštovati, no oni su nužni radi maksimalne iskoristivosti gradiva, kao i njegovog dugotrajnog očuvanja. Trenutno još ima podosta prepreka kao što su neiskustvo, manjak sredstava, manjak stručnog kadra i sl., ali što se prije krene u sustavnu digitalizaciju, to će biti više stručnog kadra koji će je provoditi na već uhodan način te će biti više adekvatne opreme, a time ujedno i veća učinkovitost i bolja kvaliteta odrađenog posla. S vješto sastavljenom i potpunom dokumentacijom za prijavu na projekte, moguće je dobiti i značajna sredstva iz Europske unije kao što je to bio slučaj i u Projektu eKultura – digitalizacija kulturne baštine.

Kao najzahtjevniji i najizazovniji aspekt procesa digitalizacije pokazala se njegoa integracija i dokumentiranje svih aktivnosti. Dokumentiranje svih aktivnosti ključno je zbog transparentnosti, odnosno kao garancija zadržavanja autentičnosti i integriteta gradiva. Također valja napomenuti da samim time što je gradivo digitalizirano, zadovoljilo kontrolu kvalitete, opisano metapodacima i učitano na neku vrstu digitalnog repozitorija, to nipošto ne znači da je time posao završio. Taj se sustav mora održavati, digitalizirani spisi se moraju provjeravati jesu li još uvijek čitljivi te se mora provoditi eventualna migracija s medija na medij u trenutku kada dođe do zastarijevanja postojećeg medija na kojima je sadržaj pohranjen i prijeti mu gubitak podataka ili konverzija u suvremeniji format zapisa.

Za kraj još valja reći da je ovim projektom tek počela ozbiljnija i sveobuhvatnija digitalizacija kulturne baštine ustanova na području kulture, što znači da se nakon njega digitalizacija svakako mora nastaviti, i to u što većem opsegu, jer je to budućnost svake ustanove, a naročito one čiji je zadatak očuvanje kulturne baštine i osiguranje njezine dostupnosti za korištenje. Upute, postupci, iskustva, analize i zaključci iz ovog rada mogu poslužiti kao orijentir za buduće slične projekte.

6. Literatura

AS/NZS ISO 13028 (2012). Australian/New Zealand Standard: Information and documentation – Implementation guidelines for digitization of records.

Bundesamt für Sicherheit in der Informationstechnik Bundesamt für Sicherheit in der Informationstechnik (2017). BSI Technische Richtlinie 03138 Ersetzendes Scannen.

digitalizacija (2021). Hrvatska enciklopedija, mrežno izdanje. Leksikografski zavod Miroslav Krleža <http://www.enciklopedija.hr/Natuknica.aspx?ID=68025> 11.1.2023.

FADGI (2023). Technical Guidelines for Digitizing Cultural Heritage Materials: Third Edition
https://www.digitizationguidelines.gov/guidelines/FADGI%20Technical%20Guidelines%20for%20Digitizing%20Cultural%20Heritage%20Materials_3rd%20Edition_05092023.pdf 17.6.2023.

FADGI (2023). Glossary: archival master file
<https://www.digitizationguidelines.gov/term.php?term=archivalmasterfile> 22.3.2023.

FADGI (2023). Glossary: derivative file
<https://www.digitizationguidelines.gov/term.php?term=derivativefile> 22.3.2023.

ISO 15489-1; Informacije i dokumentacija: upravljanje spisima (1. dio)
<http://www.das.hr/wp-content/uploads/2015/09/ISO-15489-1.pdf> 10.2.2023.

ISO/TR 19263-1:2017(E) (2017). Part 1: Photography – Archiving systems – Best practices for digital image capture of cultural heritage material.

Ivanović, Jozo (2010). Priručnik iz arhivistike 1. dio. Zagreb: Hrvatski državni arhiv.

Ministarstvo kulture (2018). Ugovor o pružanju usluga pripreme natječajne dokumentacije za projekt "Digitalizacija kulturne baštine": Izvješće o analizi trenutnog stanja
https://min-kulture.gov.hr/UserDocsImages/dokumenti/kulturna%20baština/Izvješće%20o%20analizi%20trenutnog%20stanja_Digitalizacija%20kulturne%20baštine%20ožujak2018..pdf 20.2.2023.

Ministarstvo kulture i medija. Kulturna baština [https://min-kulture.gov.hr/?id=349&pregled=1&datum=Wed%20Jan%2023%202019%2017:02:19%20GMT+0100%20\(srednjoeuropsko%20standardno%20vrijeme\)](https://min-kulture.gov.hr/?id=349&pregled=1&datum=Wed%20Jan%2023%202019%2017:02:19%20GMT+0100%20(srednjoeuropsko%20standardno%20vrijeme)) 11.1.2023.

NARA (2022). Digital preservation strategy 2022-2026
<https://www.archives.gov/files/preservation/electronic-records/digital-preservation-strategy-2022-2026.pdf> 25.3.2023.

Plan digitalizacije kulturne baštine 2020.-2025.

Smjernice za digitalizaciju kulturne baštine (2020).

Thakkar, Megha (2022). SHA 256 Algorithm Explained by a Cyber Security Consultant
<https://sectigostore.com/blog/sha-256-algorithm-explained-by-a-cyber-security-consultant/> 12.4.2023.

The National Archives (2013). Digitisation at The National Archives.

Trbušić, Željko (2022). Metode analize i optimizacije procesa optičkog prepoznavanja znakova u arhivskim informacijskim sustavima (Disertacija). Zagreb: Sveučilište u Zagrebu, Filozofski fakultet <https://urn.nsk.hr/urn:nbn:hr:131:696966> 17.6.2023.

University of Texas Libraries (2022). Metadata types <https://guides.lib.utexas.edu/metadata-basics/key-concepts> 20.2.2023.

Upute za globalne administratore – Središnji sustav (2020). Ugovor o javnoj nabavi Usluge razvoja i održavanja IKT sustava za projekt e-Kultura – „Digitalizacija kulturne baštine“ Ugovor br. 11/2020.

Upute za korisnike – Manualna izrada informacijskih paketa (2020). Ugovor o javnoj nabavi Usluge razvoja i održavanja IKT sustava za projekt e-Kultura – „Digitalizacija kulturne baštine“ Ugovor br. 11/2020.

Upute za korisnike – Sustav za podršku tijekom digitalizacije (2020). Ugovor o javnoj nabavi Usluge razvoja i održavanja IKT sustava za projekt e-Kultura – „Digitalizacija kulturne baštine“ Ugovor br. 11/2020.

Van Dormolen, Hans (2012). Metamorfoze Preservation Imaging Guidelines. National programme for the preservation of the paper heritage.

Zakon o arhivskom gradivu i arhivima NN 61/18, 98/19, 114/22

<https://www.zakon.hr/z/373/Zakon-o-arhivskom-gradivu-i-arhivima> 12.1.2023.

7. Popis ilustracija

7.1. Popis slika

Slika 1. Isječak iz Obrasca za provjeru sukladnosti.....	4
Slika 2. Prikaz vrsta i modela skenera korištenih na projektu eKultura.	18
Slika 3. Prijedlog za digitalizaciju arhivskoga gradiva.....	20
Slika 4a. Tablica s metapodacima – prvi dio.	21
Slika 4b. Tablica s metapodacima – drugi dio.	22
Slika 5a. Dnevnik pripreme gradiva za digitalizaciju – prvi dio.	25
Slika 5b. Dnevnik pripreme gradiva za digitalizaciju – drugi dio.	26
Slika 6. Studentica za skenerom uz podršku koordinatorice.	30
Slika 7. Primjer ilustracije putanje direktorija.	32
Slika 8. Tablica s metapodacima – list 2.....	37
Slika 9. Naslovna stranica središnje administracije.	40
Slika 10. Prikaz artefakata nastalih na protočnom skeneru koji ne zadovoljavaju kriterije kvalitete.....	43
Slika 11. Arhivsko gradivo snimljeno na protočnom skeneru.	45
Slika 12. Arhivsko gradivo snimljeno na koračnom skeneru.	45

7.2. Popis grafikona

Grafikon 1. Analiza učestalosti vrsta grešaka prilikom skeniranja.....	35
--	----

Digitalna transformacija ustanova u kulturi na primjeru nacionalnog projekta eKultura

Sažetak

Rad se bavi pojmovnim određenjem digitalizacije, odnosno digitalizacije kulturne baštine, kao i raščlambom i razradom tog procesa od početka pa do kraja. U radu su navedeni i svi izazovi i problemi na koje se nailazi u tom procesu. Koraci u tom postupku popraćeni su i konkretnim praktičnim iskustvima stečenim kroz sudjelovanje na nacionalnom projektu eKultura koji podrazumijeva transformaciju fizičkog arhivskog gradiva u digitalni oblik, s posebnim naglaskom na arhive. Cilj navedenog projekta jest pohrana kulturnog sadržaja u digitalnom obliku koji će korisnicima na jednom mjestu biti dostupan putem odgovarajućeg informacijskog sustava. Na taj se način ispunjavaju dva najvažnija zadatka procesa digitalizacije, a to su zaštita gradiva i njegova dostupnost. Osim toga, u radu je predstavljena statistička analiza učestalosti pojedinih grešaka koje se javljaju u radu na protočnom skeneru te kvalitativna i kvantitativna usporedna analiza rada na koračnom i protočnom skeneru.

Ključne riječi: digitalizacija, metapodaci, norme i smjernice, kulturna baština, digitalno očuvanje

Digital transformation of cultural institutions – example of the “eKultura” national project

Summary

The thesis deals with the conceptual definition of digitization, that is, the digitization of cultural heritage, as well as with the analysis and elaboration of this process from its start to finish. The thesis lists all the challenges and problems that have been encountered within this process. Also, accompanying the steps of this process, the concrete practical experiences gathered from the participation in the “eKultura” national project are given. They are dealing with the transformation of physical archival materials into digital form. The aim of the aforementioned project is to store cultural heritage content in digital form and make it available to users in one place via a corresponding information system. Thus, the two most important goals of the digitization process are fulfilled, namely preservation of archival materials and their accessibility. Furthermore, the thesis presents a statistical analysis of the frequency of individual errors that occur when working on a sheetfed scanner, as well as a qualitative and quantitative comparative analysis of working on a step scanner and a sheetfed scanner.

Key words: digitization, metadata, standards and norms, cultural heritage, digital preservation