



PAM Pokrajinski
arhiv
Maribor

Moderna
arhivistika

Časopis arhivske teorije in prakse
Journal of Archival Theory and Practice

Letnik 4 (2021), št. 2 / Year 4 (2021), No. 2

Maribor, 2021

Pokrajinski arhiv Maribor

Moderna arhivistika

Časopis arhivske teorije in prakse
Journal of Archival Theory and Practice

Letnik 4 (2021), št. 2 / Year 4 (2021), No. 2

Maribor, 2021

VSEBINA

- Tanja MARTELANC** 114
Pokrajinski arhiv Nova Gorica /Regional Archives Nova Gorica, Slovenia
Obdelava podatkov v arhivskih podatkovnih zbirkah z uporabo nekaterih metod analize vsebine
Data Processing in Archival Databases Using Certain Methods of Content Analysis
- Dr. Žiga KONCILIJA, dr. Gregor JENUŠ, dr. Tatjana HAJTNIK** 129
Arhiv Republike Slovenije, Slovenija / Archives of the Republic of Slovenia, Slovenia
Virtualna arhivska čitalnica (VAČ) in izzivi digitalizacije arhivskih čitalniških postopkov
Virtual Archival Reading Room and Challenges of Digitalization of Reading Room Services
- Dr. Gregor JENUŠ, dr. Žiga KONCILIJA, dr. Tatjana HAJTNIK** 149
Arhiv Republike Slovenije, Slovenija / Archives of the Republic of Slovenia, Slovenia
Avtomatizirano prekrivanje z arhivskim zakonom varovanih osebni podatkov - anonimizacija
Automated Processing of Personal Data Protected by Archival Law - Anonymisation
- Mag. Tatjana STIBILJ, Primož TANKO** 169
Arhiv Republike Slovenije, Slovenija / Archives of the Republic of Slovenia, Slovenia
Digitalni filmski arhiv - sistem za upravljanje in dostopnost do filmskih in avdiovizualnih vsebin e-arhivskega gradiva
Digital Film Archives – a System for the Management and Access to Film and Audiovisual Content of E-Archival Records
- Mojca KOSI, Antun SMERDEL, Mateja CIGLAR** 179
Arhiv Republike Slovenije, Slovenija / Archives of the Republic of Slovenia, Slovenia
Rešitve e-ARH.si – prijazne in uporabne tudi ranljivim skupinam
e-ARH.si Solutions: Friendly and Useful Even for Handicapped

- Jože GLAVIČ, Vesna GOTOVINA, Klavdija KRIVEC, dr. Žiga KONCILIJA, dr. Tatjana HAJTNIK** 192
Arhiv Republike Slovenije, Slovenija / Archives of the Republic of Slovenia, Slovenia
- Postopek prevzema in problematika oblikovanja SIP paketov na primeru zvočnih zapisov Državnega zbora Republike Slovenija**
Ingest Procedure and the Challenges of Creating Submission Information Packages (SIP) on the Case of Audio Records of the National Assembly of the Republic of Slovenia
- Mag. Boštjan ZAJŠEK, dr. Miroslav NOVAK** 208
Pokrajinski arhiv Maribor / Regional Archives Maribor, Slovenia
- Arhivski strokovni izzivi dolgoročne hrambe elektronskih sporočil**
Archival Professional Challenges of Long-Term Storage of Electronic Messages
- Dr. Jože ŠKOFLJANEC, mag. Boris DOMAJNKO** 224
Arhiv Republike Slovenije, Slovenija / Archives of the Republic of Slovenia, Slovenia
- Izročitev gradiva evidenc Inženirske zbornice Slovenije**
Acquisition of Registers of the Ingeneering Chamber of Slovenia
- Nataša MAJERIČ KEKEC** 241
Zgodovinski arhiv na Ptuju / Historical Archives in Ptuj, Slovenia
- Pilotski prevzem video posnetkov in digitalnih fotografij**
Pilot Ingest of Video Clips and Digital Photographs I

Prejeto / Received: 19. 08. 2021

1.04 Strokovni članek

1.04 Professional Article

PILOTSKI PREVZEM VIDEO POSNETKOV IN DIGITALNIH FOTOGRAFIJ

Nataša MAJERIČ KEKEC

Zgodovinski arhiv na Ptuj, Slovenija

natasa.majeric@arhiv-ptuj.si

Izvleček:

V prispevku je predstavljen potek pilotskega prevzema video posnetkov in digitalnih fotografij gledaliških predstav Mestnega gledališča Ptuj in sicer vrednotenje in priprava popisa gradiva ter priprava dodatnega strokovne tehničnega navodila v izvedbi Zgodovinskega arhiva na Ptuj.

Ključne besede:

video posnetki, digitalna fotografija, Mestno gledališče Ptuj

Abstract:

Pilot Ingest of Video Clips and Digital Photographs

The article presents the pilot ingest of video clips and digital photographs created at the Ptuj City Theatre. It outlines the appraisal process, the preparation of the inventory and the creation of additional professional technical instructions by the Historical Archives in Ptuj.

Key words:

video clips, digital photographs, Ptuj City Theatre

1. Priprave na pilotski prevzem gradiva Mestnega gledališča Ptuj

Mestno gledališče Ptuj je bilo izbrano kot kandidat za izvedbo pilotskega prevzema video posnetkov in digitalnih fotografij iz dveh razlogov in sicer zaradi primerne količine omenjenega gradiva in zaradi strokovno podkovanega skrbnika dokumentarnega in arhivskega gradiva. Brez primerne osebe bi pilotski prevzem težko izvedli oz. se bi moral arhiv še bolj angažirati.

Priprave na pilotski prevzem so se pričele pri ustvarjalcu gradiva, kjer sta arhivistka in skrbnik digitalnega gradiva pregledala stanje urejenosti in ohranjenosti arhivskega gradiva, določenega za pilotski prevzem (video posnetki in digitalne fotografije gledaliških predstav). Ob pregledu se je izkazalo, da je gradivo pomanjkljivo urejeno in ni opremljeno z meta podatki. Vsaka mapa je bila sicer poimenovana z naslovom posamezne gledališke predstave, vendar se je izkazalo, da so bile nekatere mape prazne, v nekaterih so bile tudi fotografije drugih predstav, pri nekaterih predstavah je bilo fotografij tudi po 100 in več, posamezne fotografije pa so bile označene le z zaporedno številko. Najprej je bilo potrebno določiti vzorec oziroma izbor predstav (v začetku smo izbrali naključno deset predstav, na koncu jih je ostalo le šest) v nekem daljšem (poljubnem) časovnem obdobju. Izbrane so bile predstave, ki so imele tudi video posnetke.

Ustvarjalcu so bila že pred pripravami na pilotski prevzem izdana pisna strokovna navodila s strani pristojnega arhiva. Pred pričetkom pilotskega prevzema je arhiv pripravil Načrt pilotskega prevzema kot neke vrste kontrolni seznam dejavnosti, ki jih bo potrebno izvesti, da bi se lahko pilotski prevzem izvedel (v njem so osnovni podatki o prevzemu, kot so npr. predmet prevzema, odgovorne osebe ter načrt izvedbe prevzema s časovnico). Prav tako je arhiv s pomočjo ustvarjalca izpolnil vprašalnik kot pripomoček za predajo gradiva. Z njegovo pomočjo se je definirala natančna vsebina pilotskega prevzema, predmet prevzema, status gradiva ob predaji, oblike zapisa (tu so mišljeni predvsem različni formati video posnetkov, fotografije so bile v glavnem v jpg obliki), podatki o pretvorbi podatkov, podatki o informacijskem okolju, poimenovanju datotek, ...).

Na podlagi vprašalnika se je ugotovilo, da ima ustvarjalec avdio video gradivo od leta 1998 dalje. Najstarejše predstave so bile posnete na VHS kasete, kasneje so jih snemali na DVD zgoščenke, oboje pa so pretvorili zaradi lažje uporabe v digitalno obliko in shranili na trdem ali zunanjem disku v formatu VOB. Predstave so se snemale zaradi promocijskih potreb (za izdelavo kratkih predstavitev klipov za youtube kanal ali facebook za potencialne kupce predstav, včasih tudi za obnovitvene vaje). Ob pripravah na pilotski prevzem se je v okviru kompetenčnega centra 1 sprožila debata ali vzeti tudi izvorno obliko in mnenje je bilo da ja (vendar ne za potrebe pilotskega prevzema). Pretvorbe formatov (npr. iz VHS v VOB) ustvarjalec ni dokumentiral tako, da dokumentacija o pretvorbah ne obstaja. Ohranjeni pa so izvorniki avdio video posnetkov. Uporaba digitalnega gradiva ni bila posebej definirana, saj je običajno bila le ena oseba zadolžena za njegovo uporabo. Ustvarjalec ima eno varnostno kopijo celotnega digitalnega gradiva na zunanjem disku. Popisa vsega digitalnega gradiva (video posnetkov in fotografij posameznih predstav) ustvarjalec ob pripravah na pilotski prevzem ni imel. Ustvarjalec do pilotskega prevzema ni sprejel notranjih pravil. Arhiv ni pripravil dodatno strokovno tehničnega pisnega navodila do zaključka pilotskega prevzema.

2. Odbiranje in priprava gradiva

Ustvarjalec je pričel odbirati gradivo izbranih gledaliških predstav na podlagi pisnih navodil arhiva. Najprej se je na prenosnem računalniku ustvarila mapa, znotraj mape so bile ustvarjene podmape z imenom posamezne predstave, v to mapo pa so bile odložene izbrane fotografije (ca. deset za posamezno predstavo), ki so bile poimenovane z dogovorjenimi meta podatki (naslov predstave, producent, sezona, imena igralcev po vrstnem redu). Fotografije so bile v jpg obliki zato jih ni bilo potrebno spreminjati, video posnetke pa je bilo potrebno spremeniti iz oblike VOB v MP4 (običajno je bil video posnetek predstave v več delih zato jih je bilo potrebno s posebnim programom zlepliti). Lepljenje video posnetkov je za ustvarjalca precejšen časovni (za eno predstavo je bilo potrebno vsaj tri ure) in finančni vložek (za nakup licenc za lepljenje in spreminjanje formatov, čeprav so na spletu prosto dostopne tudi poskusne licence, vendar so časovno omejene na 30 dni). Ker program za kreiranje SIP-a še ni bil dokončno pripravljen za uporabo (bil je še v tesni fazi) se je arhiv odločil, da se ustvarjalcu pripravi Excelova tabela v katero bi nato ustvarjalec vpisoval meta podatke o gradivu, ki so potrebni za (avtomatiziran) prenos v program SIP. Tabela je imela kar precej elementov popisa in je bila za ustvarjalca preveč okorna in preobsežna. Tako ob prevzemu gradiva tabela ni bila izpolnjena v celoti.

3. Izzivi glede odbiranja in priprave gradiva

Pri ustvarjalcu se je pojavil prvi izziv s količino digitalnih fotografij. Za nekatere novejšje predstave je fotografij tudi več kot sto in precej časa vzame sama valorizacija fotografij, ki bi bile primerne za predajo arhivu. Izbor bo moral biti prepuščen ustvarjalcu saj arhivisti tega dela ne bodo utegnili opravljati. Za pilotski prevzem smo se odločili, da se odbere le nekaj fotografij, pojavlja pa se vprašanje ali se bo ob resničnem prevzemu kdo ukvarjal z izborom fotografij ali bodo prevzete kar vse? Vse (ali le izbrane fotografije) je smiselno ob shranjevanju na računalnikih ali zunanjih diskih razporediti v mape z naslovom, ki naj vsebuje vsaj podatka o dogodku ter čas nastanka fotografije. Dodatni podatki, ki so lahko dobrodošli, pa so še lokacija dogodka in osebe na fotografiji. Z kako obsežnimi podatki opremiti posamezne fotografije? Tu bo verjetno prišlo do razhajanj v mnenjih oz. željah posameznega arhivista, na našo odločitev pa bo vplival program za kreiranje SIP-a, ki ima pri naslovu popisne enote omejitve znakov (255).

Drugi izziv za ustvarjalca so predstavljali video posnetki, ki jih je bilo potrebno najprej zlepiti, če so bili v več delih in nato spremeniti v formate za dolgoročno hrambo (Seznam oblik zapisa za dolgoročno hrambo, ARS, 2021). Pri tej vrsti dela se pokaže, da bo pri ustvarjalcu potreben strokovno podkovan delavec, ki bo znal izvesti vsa omenjena opravila.

4. Potek prevzema digitalnega gradiva

Osnova za pripravo SIP paketa s strani ustvarjalca je bil strukturiran repozitorij, kjer se zbira bodoče arhivsko gradivo. Ta mora biti pravilno zgrajen in čim bolj enoten. Pristop k pripravi SIP paketa se je pričel z inštalacijo programskega orodja na prenosni računalnik ustvarjalca. Podatki, torej repozitorij ustvarjalca, pa se je nahajal na zunanjem prenosnem disku.

Inštalacija programskega orodja je bila uspešna, nato je sledilo pregledovanje samega repozitorija, postavitve strukture in formatov datotek, da bi ugotovili, kakšen bo optimalen pristop za pripravo sheme profila. Ker je ustvarjalec imel poimenovane direktorije in imena datotek opisno, se je izvedel prenos strukture repozitorija kot model za pripravo sheme. Določiti je bilo potrebno veljavne formate in obrazce za popisovanje SIP-a razširiti z dodatnimi elementi popisa.

Sledilo je prvo presenečenje in to je, da sistem ni dovolil kreiranja SIP-a na zunanjem disku, prav tako ni bilo možno izpeljati pretoka podatkov iz zunanjega diska v SIP urejevalnik. Kaj bi lahko bil vzrok te težave ni bilo mogoče ugotoviti. Nato je sledilo pretakanje podatkov na prenosni računalnik ter izdelava SIP paketa, ki je znašal 31 GB. Struktura SIP paketa je bila ista kot je bila struktura ustvarjalčevih direktorijev.

Vmes se je pojavila nevšečnost – namreč sistem je pri izdelavi sheme zaznal v enem poddirektoriju format zapisa, ki se je pri avtomatskem prenosu strukture spregledal, kar je pomenilo delo od začetka, saj to ni bila veljavna oblika formata zapisa in je bil tak SIP neveljaven. Ob tem bi bilo vredno opozoriti, da je pred pričetkom kreiranja SIP urejevalnika, pravzaprav že pri pripravi sheme, potrebno z 100% natančnostjo vedeti, ali so vsi formati datotek v pravilni obliki zapisa.

Glede na velikost SIP- a (31 GB) se je pri ponovni izgradnji SIP-a izbralo le nekaj predstav, ki so se ponovno uvozile z izdelano shemo profila, v SIP urejevalniku so se dodali vsebinski metapodatki (žal jih ustvarjalec ni imel popisanih). Sledila je validacija SIP paketa in kreiranje paketa, ki je v končni obliki bil hranjen na prenosnem računalniku. Velikost SIP paketa je bila 13,36 GB, sama izdelava je trajala cca. 30 minut, nato je sledilo kopiranje na zunanji disk.

Naslednji korak je predstavljal prenos na omrežje arhiva za pripravo Ingesta. Informacijsko varnostna politika zahteva, da je potrebno vsak zunanji disk preveriti z antivirusno programsko opremo, preden se naredi vpogled vanj. Pregled zunanjega diska je trajal skoraj 3 ure, saj je imel ustvarjalec na njem ogromno raznovrstnih podatkov, programske inštalacijske datoteke...

Na koncu so zaradi vseh omenjenih težav v Arhivu Republike Slovenije terminalsko uvozili SIP paket v Ingest (v njihovo notranje omrežje). Zajem paketa je bil uspešno izveden (sam prenos celote 13,36 GB je trajal 17 minut), ni pa bil speljan celotni proces zajema do konca. Ustavilo se je pri kreiranju AIP paketov, saj bi nadaljnja operacija pomenila nepotrebno obremenjevati Fedore z neuporabnimi podatki. Kljub pomislekom vendar v želji, da se pilotski prevzem izvede do konca, se je pričelo shranjevanje AIP paketa in tukaj se je pojavilo novo presenečenje. Pri hranjenju na Fedoro je prišlo do kritične napake 7ZIP, zato se hramba v Fedori in prenos popisa strukture AIP paketa v Scope kot novonastale PE ni izvedla. Poizkusno so zadevo poskušali izpeljati tudi v ARS-u vendar neuspešno.

Predlogi glede na izveden projekt za naslednje podobne prevzeme:

1. Ustvarjalec mora k zadevi pristopiti z določenim predznanjem.
2. Ustvarjalec mora imeti zbirne podatke že strukturirane v svojem arhivskem direktoriju (ni smiselno, da bi arhivi delali shemo profila, predpisali formate zapisa, nato pa ustvarjalcu dali prosto pot). Pristop mora biti sledeč – ustvarjalec in arhivist valorizirata dokumentarno digitalno gradivo, arhivske kopije se skopirajo (prenesejo) v arhivski repozitorij, tam se vse te datoteke pretvorijo v format za dolgoročno hrambo in šele nato se dela struktura s strani ustvarjalca, kar pa je osnova za izdelavo dSTN in sheme profila. Vsak drug pristop pri nestrukturiranih datotekah vodi v kaos, repetitije popisovanja SIP-a....
3. Arhivist mora do najmanjše datoteke poznati strukturo in formate zapisov, saj lahko šele takrat pripravi zahtevo preko sheme profila in definira same formate zapisa.
4. Ustvarjalec mora pripraviti »prostor« na osebнем računalniku, kjer se bo kreiral SIP paket (ne priporoča se izdelava SIP paketa direktno na zunanjem disku).
5. Računalnik, ki ga bo uporabljal ustvarjalec, naj ne bo star računalnik ali osebni prenosni računalnik, ker je tak pristop neprofesionalen in ne pritiče ustanovi, povzroča zamude in težave.
6. Opozoriti je potrebno, da se naj v dodatkih strokovno tehničnih navodilih zahteva disk le s SIP-om in ničemer drugim (torej čisti disk).
7. Največja težava pri izdelavi SIP paketov leži na arhivistih. Sama dSTN ne pripravijo ustvarjalca do tega, da bi lahko začel logično razmišljati o tem, kaj sploh počne in kakšno znanje in razumevanje potrebuje, preden se loti same priprave SIP paketa.
8. Definitivno ustvarjalec brez pomoči informatika in arhivista ne bo zmožel/znal izdelati SIP paketa, saj ne razumejo kaj je to shema profila, za kaj to potrebuje, kako prenašati. Res je, da obstajajo navodila, a ustvarja se dvom, da se bo ustvarjalec poglobil v študiranje programske opreme.
9. Optimizacija komunikacijskih kanalov (poiskati nekoga, ki bo znal poiskati ozka grla in jih znotraj ARS, HKOM-a in Ingest strežnika odpraviti), ker je trenutno stanje sicer pogojno delujoče in pogojno uporabno.
10. Ugotoviti, kaj povzroča oziroma, kje je časovna omejitev 25-ih minut (možnih več scenarijev – Scope, Ingest strežnik, ruterji, swichi, HKOM, client-server) in to omejitev odpraviti.
11. Ugotoviti dejansko kaj povzroča napako na Fedori.

5. Sklepni zaključki pilotskega prevzema

V tej fazi razvoja pilotski projekt Mestnega gledališča ni bil uspešno zaključen, saj ni bilo možno preveriti celotnega workflowa oz. inštalacije do izdelave DIP-a, je pa bila uspešno izvedena izdelava SIP paketa, kar pa je bil namen projekta. Programsko orodje za izdelavo SIP-a nima resnih pomanjkljivosti z izjemo prijavljene napake opisovanja podmap uvoza podatkov preko Excelove tabele.

Šibke točke so na poznavanju in predpripravi gradiva s strani ustvarjalca (tukaj bodo potrebne delavnice o pripravah SIP paketov, o opisovanju, o pretvarjanju v obliko za dolgoročno hrambo, komuniciranje med ustvarjalcem – arhivistom – informatikom), pa tudi poznavanje SIP programskega orodja in koncept delovanja SIP urejevalnika s strani arhivistov. Gre za nov način dela in razumevanja koncepta elektronskega arhiviranja, ki se bo sčasoma spremenil na boljše odvisno od vloženega truda na strani arhivov.

Največ težav je bilo ob zaključku pilotskega prevzema na prezahtevnih povezavah, ki nam lahko povzročajo nadaljnje težave vse dokler se ne odpravijo zaznane komunikacijske Inges, Fedora, client-server težave.

6. Viri

Seznam oblik zapisa, primernih za dolgoročno hrambo gradiva v digitalni obliki. Pridobljeno 18. 09. 2021 s spletne strani <https://www.gov.si/assets/organi-v-sestavi/Arhiv-RS/Zakonodaja-2019/Pravilnik-o-enotnih-tehnoloskih-zahtevah-za-zajem-in-hrambo-gradiva-v-digitalni-obliki/Seznam-oblik-zapisa-za-dolgorocno-hrambo-ver-1.0.pdf>.

SUMMARY

PILOT INGEST OF VIDEO CLIPS AND DIGITAL PHOTOGRAPHS

Nataša MAJERIČ KEKEC

Historical Archives Ptuj, Slovenia

natasa.majeric@arhiv-ptuj.si

The Ptuj City Theatre was chosen as a candidate for the pilot ingest of videos and digital photographs for two reasons: it created the appropriate amount of records for ingest and employs a professionally educated administrator of records. Preparations began with the review of archival records designated for the pilot ingest. For each theatre performance there was a folder with the title of the performance, which contained videos and photographs. Firstly, it was necessary to select sample performances for a longer period. The selection of photographs and videos was done based on written professional instructions, issued by the competent archives.

The initial step was the creation of a folder on a portable computer, with subfolders equipped with a name of a performance. These folders were filled with selected photographs (approximately ten for each performance). File names were created with defined metadata (performance title, producer, season, names of performers). The photographs were in .jpg format, videos had to be converted from .vob to .mp4 format (usually the videos were in several parts and had to be merged together). This activity required a significant amount of time (at least three hours for one performance) and a financial investment for the appropriate software.

In this phase of development, the pilot project was not successfully completed because the entire workflow could not be checked. But the production of the SIP package was successfully finished, which was the purpose of the project. The software tool for creating SIPs performed successfully and did not show any serious shortcomings.