

# Problemi popisovanja digitalnega in digitaliziranega arhivskega gradiva<sup>1</sup>

ZDENKA SEMLIČ RAJH, PH.D., ASSIST. PROF.

Archival councilor at Regional Archives Maribor, Glavni trg 7, SI-2000 Maribor

e-mail: zdenka.semlic@pokarh-mb.si

## Problems with Describing E-born and Digitized Archival Records

### ABSTRACT

National Program for Culture defines that the digitization and preservation of digital content in the field of culture, which include archives and libraries are among the basic tasks of public institutions dealing with cultural heritage. However, it would be reasonable to think how the digitization of library and archival material and free access to the online content are reflected in the description of digital content. Libraries and archives implemented in the last ten years various digitization projects. However, they did not pay attention to the description of digitized material, which raises the question of the importance of context. This raises also the question of the evidential value of the digitized archives as well as authenticity, integrity and accountability of the digitized material in a broader context, which archives provide on the basis of the principles of provenance and original order. The digitization of archives is not specific only from the standpoint of the digitization process itself but also from the standpoint of the description of digitized archives. Adequate description enables namely the wide use of digitized archival material. A short presentation of literature and up to now accomplished research is followed by an analysis of online accessible descriptions of digitized archival material. The analysis was conducted in 48 foreign archival information systems and their databases. These are large systems, some of which operate as a common database of several smaller archives, but some systems are individual systems of the large national archives. The analyses was conducted in publicly accessible user interfaces, and was based on a direct applicability of the descriptions of digitized archival material within individual database. This is the query that can be performed by average user of archival material with access to the search engine knowing at least the basics of the web browser and of searching and sorting of the results.

**Key words:** Archives, digitalization, description, evidence, context

## Problematiche nella descrizione di documenti archivistici nativi digitali e digitalizzati

### SINTESI

Il programma nazionale per la cultura definisce che digitalizzazione e conservazione dei contenuti digitali nel campo della cultura, che includano i archivi e biblioteche, sono tra i compiti fondamentali delle istituzioni pubbliche che si occupano del patrimonio culturale. Tuttavia, sarebbe ragionevole pensare che la digitalizzazione della biblioteca e del materiale archivistico ed il libero accesso ai contenuti online si riflettano nella descrizione dei contenuti digitali. Biblioteche e archivi hanno portato avanti negli ultimi dieci anni diversi progetti di digitalizzazione. Tuttavia, non hanno prestato attenzione alla descrizione del materiale digitalizzato, che solleva la questione dell'importanza del contesto e pone anche la questione del valore probatorio degli archivi digitalizzati, nonché dell'autenticità, integrità e responsabilità del materiale digitalizzato in un contesto più ampio, che gli archivi forniscono sulla base dei principi della provenienza e dell'ordine originale. La digitalizzazione degli archivi è specifica non solo dal punto di vista del processo, ma anche dal punto di vista della descrizione degli archivi digitalizzati. Una descrizione adeguata consente di utilizzare in modo ampio il materiale archivistico digitalizzato. Una breve presentazione della letteratura e della ricerca finora effettuata è seguita da un'analisi delle descrizioni accessibili online di materiale archivistico digitalizzato. L'analisi è stata condotta in 48 sistemi informativi internazionali d'archivio e nelle loro banche dati. Questi sono grandi sistemi, alcuni dei quali funzionano come un database comune di diversi archivi più piccoli, ma alcuni sono sistemi individuali dei grandi Archivi nazionali. Le analisi sono state condotte nelle interfacce utente pubbliche ed accessibili, e sono state basate su un'applicazione diretta delle descrizioni di mate-

1. Članek je bil v razširjeni obliki pod naslovom »Izzivi popisovanja digitalnega in digitaliziranega arhivskega gradiva: Analiza izvedenih popisov izvirnega e-arhivskega gradiva in digitaliziranih v spletno dostopnih arhivskih podatkovnih zbirkah« prvič objavljen v zborniku Tehnični in vsebinski problemi klasičnega in elektronskega arhiviranja 2017, ter predstavljen na istoimenski konferenci v Radencih, aprila 2017.

riale archivistico digitalizzato all'interno di un singolo database. Questa interrogazione può essere eseguita dall'utente medio di materiale d'archivio con accesso al motore di ricerca, conoscendo almeno le basi del browser web e di ricerca e ordinamento dei risultati.

**Parole chiave:** archivi, digitalizzazione, descrizione, evidenza, contesto

## Problemi popisovanja digitalnega in digitaliziranega arhivskega gradiva

### IZVLEČEK

Nacionalni program za kulturo opredeljuje, da digitalizacija in hranjenje digitalnih vsebin s področja kulture sodita med osnovne naloge javnih ustanov, ki se ukvarjajo s kulturno dediščino, torej tudi arhive in knjižnice. Vsekakor pa velja malce razmisliti o tem, kako se digitalizacija knjižničnega in arhivskega gradiva in prost dostop do teh vsebin na spletu odražajo u popisovanju digitalnih vsebin. Tako knjižnice kot arhivi smo v zadnjih nekaj desetih letih izvajali različne projekte digitalizacije, brez da bi ob tem razmišljali tudi o popisovanju takšnega digitaliziranega gradiva, ki tako knjižnice kot arhive postavlja pred vprašanje pomena konteksta digitaliziranega gradiva. Pri tem se postavlja vprašanje pravne veljave digitaliziranega arhivskega gradiva oziroma zagotavljanje avtentičnosti, celovitosti in integritete gradiva v širšem kontekstu, ki jih arhivi zagotavljamo na osnovi načela provenience in prvotne ureditve. Digitalizacija arhivskega gradiva ni posebna samo s stališča samega postopka temveč tudi s stališča popisovanja digitaliziranega arhivskega gradiva. Predvsem ustrezen popis digitaliziranega arhivskega gradiva omogoča tudi njegovo široko uporabo. Po krajši predstavitvi literature sledi analiza popisovanja digitaliziranih gradiv v spletno dostopnih arhivskih podatkovnih zbirkah. Analiza in raziskava je bila izvedena na 48 tujih arhivskih informacijskih sistemih in njihovih podatkovnih zbirkah. Gre za velike sisteme, med katerimi nekateri delujejo kot skupna podatkovna zbirka več manjših arhivov, nekateri pa kot sistemi velikih nacionalnih arhivov. Raziskava je bila izvedena na javno dostopnih uporabniških vmesnikih, njeno osnovo pa predstavlja neposredna uporabnost popisov digitaliziranih gradiv v posameznih podatkovnih zbirkah. Pri tem gre za poizvedovanje, ki ga lahko izvede povprečen uporabnik arhivskega gradiva z dostopom do spletnega iskalnika in ki pozna vsaj osnovne zakonitosti delovanja spletnega brskalnika ter iskanja in razvrščanja rezultatov poizvedovanja. Pri tem mora poznati tudi vsaj osnove poizvedovanja s pomočjo krnjenja besed.

**Ključne besede:** arhivi, digitalizacija, popisovanje, javna vera, kontekst

## 1 Uvod

Med osnovne naloge javnih ustanov, ki se ukvarjajo s kulturno dediščino, torej tudi arhivov in knjižnic, sodi tudi digitalizacija in hranjenje digitalnih vsebin s področja kulture, kar opredeljuje tudi Nacionalni program za kulturo 2014-2017 (Zupan in Černigoj Blažko, 2014). Po podatkih, ki jih najdemo v omenjenem programu so omenjen institucije skupaj z drugimi institucijami, ki prav tako digitalizirajo in hranijo digitalizirano gradivo<sup>2</sup>, do leta 2013 digitalizirale prek 650.000 različnih vsebin (Zupan in Černigoj Blažko, 2014, str. 130). Ker pa je smisel digitaliziranega gradiva poleg hranjenja in varovanja, tudi njegova širša dostopnost, ga je mnogo dostopnega tudi preko svetovnega spleta (podrobneje glej Zupan in Černigoj Blažko, 2014, str. 130).

Kot navaja Nacionalni program za kulturo (Zupan in Černigoj Blažko, 2014) pa namen digitalizacije zbirk ni samo njihovo varno in trajno hranjenje, temveč tudi omogočanje čim večje dostopnosti digitalnih kulturnih vsebin preko spleta z namenom služenja izobraževalnim, ustvarjalnim, znanstvenim in tudi drugim nameram.

Tako knjižnice kot arhivi smo v zadnjih nekaj desetih letih izvajali različne projekte digitalizacije. Šlo je predvsem za digitalizacijo najbolj ogroženega in najbolj uporabljenega arhivskega gradiva.

Vendar se je stroka projektov digitalizacije lotila brez, da bi ob tem razmišljali tudi o popisovanju takšnega digitaliziranega gradiva, ki tako knjižnice kot arhive postavlja pred vprašanje pomena konteksta digitaliziranega gradiva. Pri tem se postavlja vprašanje pravne veljave digitaliziranega arhivskega gradiva oziroma zagotavljanje avtentičnosti, celovitosti in integritete gradiva v širšem kontekstu.

2. Digitalizirano gradivo je tisto gradivo v fizični obliki, ki je šlo skozi proces digitalizacije. Nasprotje tega je »ne digitalizirano gradivo«, ki še ni šlo skozi proces digitalizacije ter izvorno digitalno arhivsko gradivo. Produkt digitalizacije pa so »digitalizati«, »digitalne kopije« ali »digitalne vsebine«.

## 2 Pomembnejša literatura in dosedanje raziskave

Arhivska stroka, ne samo domača, ampak tudi tuja, se je z vprašanjem digitalizacije ker precej ukvarjala, vendar so se večinoma vse razprave ukvarjale predvsem s tehnološkimi vidiki digitalizacije, njenimi pomanjkljivostmi in prednostmi, etičnimi in zakonskimi vidiki, ter vprašanjem avtorskih pravic na digitaliziranem arhivskem gradivu (prim.: McKay, 2003; Hughes, 2003; Guidelines for Digitization..., 2003; Buetman et al. 2005; Hirtle et al., 2009; Owens, 2013; Miller, 2013; Moravec, 2014; McCracken, 2014 in drugi). Avtorjev in člankov, ki se nanašajo na tehnološke vidike digitalizacije v tem prispevku ne bomo posebej predstavljali, saj nas je pri naši raziskavi zanimalo predvsem popisovanje oziroma vprašanje in pomen konteksta pri popisovanju digitaliziranih gradiv.

Izjemno redki pa so članki, v katerih so se avtorji ukvarjali s popisovanjem digitaliziranega arhivskega gradiva, oziroma z vprašanjem in pomenom konteksta za digitalizirano arhivsko gradivo. Eno izmed del, ki jih velja posebej omeniti je delo Marka Vojcnerja z naslovom *The Importance of Context for Digitized Archival Collections* (2008). Vojcner (2008) navaja, da je že leta 1981 R. Krauss v svojem prispevku »The Originality of Avant-Garde« poudarila pomen javne vere arhivskega gradiva. Za arhive pa je »ohranjanje izvirnika« (certification of original) ostalo osnovna naloga v teoriji in praksi (Vojcner, 2008, str. 1). In prav zagotavljanje avtentičnosti digitaliziranega arhivskega gradiva predstavlja za arhive velik izziv.

Arhivsko gradivo je že od nekdaj imelo dokazno vrednost in je zagotavljalo pravice, kot navaja Vojcner (2008). Za doseg tega cilja, pa so arhivisti razvili dva osnovna principa, ki danes predstavljata temelj arhivske stroke. Prvi je princip provenience, drugi pa princip prvotne ureditve. Predvsem drugi prepoveduje obširnejše preurejanje gradiva znotraj fonda in rušenje prvotne ureditve. Oba principa namreč zagotavljata, da uporabnik gradivo vidi in uporablja v kontekstu, v katerem je nastalo, saj mu le ta zagotavlja možnost pravilne interpretacije gradiva, kot pravi Vojcner (2008, str. 2).

Danes je mnogo arhivskega gradiva dostopno preko spleta, izven arhivov, kjer hranijo izvirno arhivsko gradivo. Gre za celo vrsto različnih institucij, ki pa to gradivo objavljajo preko spleta izven njegovega zgodovinskega konteksta in izven konteksta prvotne ureditve. Največjo težavo pa predstavlja dejstvo, da gre pri digitalizaciji za selektiven proces, saj se digitalizirajo zgolj posamezni dokumenti, ne pa celote arhivskega gradiva nekega arhivskega fonda. Takšna selektivna digitalizacija arhivskega gradiva pa ima vsekakor velike posledice na arhivske koncepte, predvsem na vprašanje konteksta in zagotavljanje javne vere.

Že Cook (1997) je ugotavljal, da je prišel skrajni čas, da se arhivisti pričnejo ukvarjati z vprašanji, ki se nanašajo na arhivsko gradivo v digitalnem svetu. Do tedaj so se arhivisti namreč ukvarjali le z izvirniki na klasičnih nosilcih informacij in pojavljati se je pričelo vprašanje, kako se spopasti z porajajočim se elektronskim arhivskim gradivom, ki se je izkazalo za veliko občutljivejše kot papir. V nekem trenutku se je morala arhivska stroka spopasti s prevrednotenjem svojih temeljnih teorij in jih uporabiti na novih medijih (Cook, 1997 v Vojcner, 2008).

Pred tem, kako sodobne tehnologije grozijo »da bodo spremenile način razmišljanja« arhivov in tudi kako se navezujejo na arhivske koncepte resnice in javne vere, sviri tudi Koltun (1999, str. 115 v Vojcner, 2008). Vojcner (2008) pa poudarja, da arhivisti arhivsko gradivo pridobivajo, pri tem pa stremijo za tem, da zagotovijo njegovo celovitost in ga uredijo in popišejo tako, da ne ogrozijo njegove javne vere, nato pa gradivo pustijo, da govori samo zase. Pri tem pa tudi poudarja, da so bili tradicionalni uporabniki arhivskega gradiva predvsem raziskovalci in uporabniki, ki jim je kontekst predstavljal pomembno dejstvo. Že raziskave iz leta 1984 so namreč pokazale, da je tradicionalne uporabnike zamenjala druga skupina uporabnikov, ki pa nimajo pretiranega interesa za raziskovanje primarnih virov, kar pa je vodilo tudi v nepoznavanje in v zmanjšano zanimanje za kontekst (Freeman, 1984 v Vojcner 2008).

Zanimivo je bilo mnenje Hickersona (2001), ki je med izzive arhivske stroke uvrstil dostopnost, ki bi jo po njegovem bilo mogoče zagotoviti z digitalizacijo. Vendar pa je bil prepričan, da glede na to, da raziskovalcev kontekst ne zanima, temu arhivisti ne bi smeli posvečati prevelike pozornosti (Hickerson, 2001 v Vojcner 2008).

Na problematiko konteksta pri popisovanju arhivskega gradiva in objavljanje posameznih vrst gradiva izven arhivskih informacijskih sistemov pa je opozoril tudi Popovici v svojem prispevku *Medievalia: A Place Where Archival Science and History Meet and Then Separate Again* (2016).

Med slovenskimi avtorji sta se s problematiko digitalizacije arhivskega gradiva poglobljeno ukvarjala Novak in Horvat (2014), ki sta prva opozorila na dejstvo, da je vprašanje digitalizacije kompleksno vprašanje, povezano s strokovno obravnavo digitalizatorjev. Tako kot prej omenjeni tudi avtorji sta prepoznala razliko med digitalizati, nastalimi pri ustvarjalcih in tistimi, ki nastajajo v arhivih kot rezultat digitalizacije arhivskega gradiva, predvsem pa s tem povezano vprašanje avtentičnosti in javne vere (Novak in Horvat, 2014: 174).

### 3 Metoda

Da bi analizirali podatke in prišli do rezultatov, je bila za namen naše raziskave kot glavna uporabljena metoda analize vsebine. V slovenski arhivski teoriji in praksi je metoda analize vsebine dokaj nepoznana in se do nedavnega ni uporabljala, čeprav gre za raziskovalno metodo, ki se zelo pogosto uporablja tako v družboslovnih, kot tudi naravoslovnih raziskavah. Prvo pilotno raziskavo v slovenski arhivistiki sta izvedli Šauperl in Semlič Rajh (2012)<sup>3</sup>, nato pa leto kasneje še drugo<sup>4</sup> (*Semlič Rajh in Šauperl, 2013*).

Kot navaja Šauperl (2008, str. 155), je potrebno analizo vsebine (ang. *content analysis*) ločiti od vsebinske analize (ang. *subject analysis*). To vsekakor velja tudi v arhivistiki, saj arhivisti dejansko vsebinsko analizo opravljajo, ko pri popisovanju arhivskega gradiva, bodisi na nivoju fonda ali dokumenta, v skladu z ISAD(G) (2000) ugotavljajo njegovo vsebino in ko to vsebino izražajo tudi z določanjem ustreznih deskriptorjev.

Analiza vsebine je raziskovalna metoda, s katero iščemo in ugotavljamo vzorce vedenja ali dogodkov v zbranih podatkih (Šauperl, 2008; Peräkylä, 2005; Content Analysis, 2012; Stemler, 2001; Elo in Kyngäs, 2008; Roblek, 2009).

Izvedbo analize izvedenih popisov izvirnega e-arhivskega gradiva in digitalizatorjev v tujih in slovenskih arhivskih informacijskih sistemih bomo predstavili po korakih, ki jih predlagata White in Marsh (2006, str. 30) z dodatnimi vprašanji, ki si jih moramo zastaviti po Krippendorf (2004).

Raziskava je bila izvedena na javno dostopnih uporabniških vmesnikih, njeno osnovo pa predstavlja neposredna uporabnost popisa izvirnega e-arhivskega gradiva in digitalizatorjev v posameznih podatkovnih zbirkah. Poizvedovanje, ki smo ga izvedli lahko izvede povprečen uporabnik arhivskega gradiva z dostopom do spletnega iskalnika. Pri tem mora poznati vsaj osnovne zakonitosti delovanja spletnega brskalnika ter iskanja in razvrščanja rezultatov poizvedovanja.

Naše raziskovalno vprašanje je bilo: Ali je popis digitaliziranih vsebin arhivskega gradiva izvedeno tako, da omogoča uporabnikom samostojno poizvedovanje po arhivskih podatkovnih zbirkah in tudi hitrejši in lažji dostop do arhivskega gradiva, predvsem pa njegovo razumevanje v kontekstu in provenien-  
ci? Izvedena analiza je temeljila na ugotavljanju sledečih dejstev:

- ali je v sistemu implementirano iskanje po online dostopnem arhivskem gradivu;
- po katerih vrstah digitalizatorjev je možno iskanje v sistemu (spisovno gradivo, razglednice, fotografije, listine, izvorno e-arhivsko gradivo, načrti). To je pomemben podatek, ki kaže na raznolikost možnosti dostopanja do digitalizatorjev in poizvedovanja znotraj arhivske podatkovne zbirke;
- ali je iz dodanih metapodatkov razvidno, da gre za digitalizate;
- ali je v popisu navedena količina in obseg za fizično gradivo, ali za priložene oziroma objavljene digitalizate;

3. Šauperl in Semlič Rajh sta v raziskavi Značaj področja identifikacije popisnih enot i sadržaja prema standardu ISAD (G) u slovenskoj bazi podataka Sira\_net skušali ugotoviti, ali se področje identifikacije popisne enote in področje vsebine in ureditve iz ISAD(G)2 dosledno uporabljata. Področji sta pomembni, saj uporabniku omogočata ugotoviti značilnosti in vsebino popisne enote ter s tem presoditi o primernosti za konkretno informacijsko potrebo. Zanimalo nas je tudi, ali poleg teh arhivisti popisujejo tudi katere druge značilnosti popisnih enot. S postopkom analize vsebine smo pregledali zapise iz dveh slovenskih arhivov v javno dostopni slovenski podatkovni bazi Sira\_net.

4 V raziskavi Analiza oblikovanja vsebine zajetih podatkov v podatkovni bazi SIRAnet smo na podlagi naslova in vsebine popisnih enot, ki uporabniku omogočajo ugotoviti, ali je arhivsko gradivo ustrezno za njegovo uporabo, kako se v podatkovni bazi SIRAnet oblikujeta ta dva elementa, ki jih predpisuje ISAD(G)2. S postopkom analize vsebine smo pregledali 100 zapisov iz baze SIRAnet. Štiri različne fonde enega arhiva smo izbrali zato, ker v ostalih arhivih ni bilo ustreznih vzorcev s podrobneje opisano vsebino.

- ali so pravilno oblikovani naslovi popisnih enot glede na prezentirano gradivo;
- ali iskalnik omogoča iskanje po digitaliziranem arhivskem gradivu in ali je način iskanja uporaben za povprečnega uporabnika.

### 3.1 Podatki za analizo

Po Krippendorffu (2004) si je treba v postopku analize vsebine zastaviti še ta vprašanja:

- Kakšne podatke analiziramo?
- Kako so definirani?
- Kakšna je populacija, na katero se nanašajo?
- V kakšnem kontekstu so podatki, ki jih analiziramo?

Podatki, ki jih za našo analizo potrebujemo, so besedilni in slikovni. Gre za posamezne popisne enote in digitalizate v arhivskih podatkovnih zbirkah. Gre za posamezne besede ali besedne zveze, ki opredeljujejo popisna polja v skladu s standardom ISAD(G) 2 (2000) in jih arhivisti vpisujejo v ustrezna polja, namenjena popisovanju in dodajanju metapodatkov.

Za raziskavo smo v tistih izbranih podatkovnih zbirkah, ki uporabljajo različna programska orodja, opazovali posamezne zapise in metapodatke. Ker gre za raziskavo, ki je bila narejena na razmeroma velikih podatkovnih zbirkah, smo opazovali naključno izbrane zapise, in sicer prvih deset zapisov vsake posamezne zvrsti digitaliziranega arhivskega gradiva in pri tem oblikovali priložnostni vzorec. Ker za veliko večino analiziranih podatkovnih zbirk ne poznamo obsega zbirke, tudi nismo mogli oblikovati statističnega vzorca.

Kako smo za analizo izbrali arhive in njihove podatkovne zbirke? Za analizo smo si prizadevali dobiti dostop do čim večjega števila arhivov, ne samo v Sloveniji, temveč tudi v tujini. Dobili smo dostop do vseh podatkovnih zbirk v slovenskih javnih arhivih in v izbranih tujih arhivih. Raziskava je bila izvedena v 48 arhivskih informacijskih sistemih in njihovih podatkovnih zbirkah. 60% obravnavanih sistemov deluje na programskem orodju scopeArchiv in aplikaciji za spletno poizvedovanje scopeQuery, 40% obravnavanih arhivskih sistemov pa uporablja druge rešitve. Gre za velike sisteme, med katerimi nekateri delujejo kot skupna podatkovna zbirka več manjših arhivov (npr. škotska, hrvaška, romunska mreža arhivov), nekateri pa kot sistemi velikih nacionalnih arhivov (npr. Italija, ZDA, Anglija, Poljska, Nemčija itd.)<sup>5</sup>. Izbrali smo arhive, katerih arhivski informacijski sistemi so dostopni online in omogočajo poizvedovanje po arhivskih podatkovnih zbirkah. Gre za arhive, ki uporabljajo različne prakse popisovanja arhivskega gradiva. Prav tako gre za arhive iz različnih jezikovnih področij, tako anglo-ameriškega, german-

5. Za namen naše raziskave smo izbrali sledeče arhivske informacijske sisteme:

- arhivski informacijski sistemi, ki delujejo v okviru programskega orodja scopArchive in preko spletnega iskalnika scopeArchive Query: Avstrijski državni arhiv na Dunaju (Österreichisches Staatsarchiv), Arhiv Univerze na Dunaju (Archiv der Universität Wien), Nacionalni arhiv Luksemburga (Archives nationales de Luxembourg), Magyar Nemzeti Levéltár, Mestni arhiv Budimpešta (Budapest Főváros Levéltára), Mestni arhiv Saarbrücken (Stadtarchiv Saarbrücken), Podatkovna zbirka za regijo Bukarešta, Podatkovna zbirka za regijo Brasov, Podatkovna zbirka za regijo Temišvar, Podatkovna zbirka za regijo Cluj, Podatkovna zbirka za regijo Iasi, Državni arhiv kantona Jura (Archives cantonales jurassiennes), Državni arhiv kantona Vaud (Archives cantonales vaudoises), Državni arhiv Freiburg (Archives de l'Etat de Fribourg), Državni arhiv Valais (Archives de l'Etat du Valais), Knjižnica Bern (Burgerbibliothek Bern), Državni arhiv Aargau (Staatsarchiv Aargau), Helvetica Archives – podatkovna zbirka Nacionalne knjižnice (Helvetic Archives – the archive database of the Swiss National Library), Zgodovinski arhiv in knjižnica švicarskega PTT (Historisches Archiv und Bibliothek PTT), Podatkovna zbirka švicarskih železnic (SBB Historic), Švicarski zvezni arhiv (Schweizerisches Bundesarchiv), Državni arhiv kantona Basel – mesto (Staatsarchiv des Kantons Basel-Stadt), Državni arhiv kantona Bern (Staatsarchiv des Kantons Bern), Državni arhiv kantona Lucern (Staatsarchiv des Kantons Luzern), Državni arhiv kantona St. Gallen (Staatsarchiv des Kantons St. Gallen), Državni arhiv kantona Thurgau (Staatsarchiv des Kantons Thurgau), Državni arhiv kantona Zürich (Staatsarchiv des Kantons Zürich), Mestni arhiv Zürich (Stadtarchiv Zürich) in Arhiv samostana St. Gallen (Stiftsarchiv St. Gallen), Urad Združenih narodov v Ženevi (United Nations Office at Geneva UNOG), Svetovni svet cerkva v Ženevi (World Council of Churches WCC),
- Arhivski informacijski sistemi, ki delujejo na osnovi drugih aplikacij: Zvezni arhiv Koblenz (Bundesarchiv Koblenz), Štajerski deželni arhiv Gradec (Steiermärkisches Landesarchiv Graz), Koroški deželni arhiv Celovec (KLAIS), Češki državni arhiv Praga, Danski državni arhiv (DAISY), Hrvaški državni arhiv (ARHINET), Irski državni arhiv Dublin (Databases), Portal italijanskih državnih arhivov (Guida generale), Kanadska državna knjižnica in arhiv (Library and Archives Canada), podatkovne zbirke Poljskega nacionalnega arhiva (NAC), Nacionalni arhiv ZDA (podatkovne zbirke), Nacionalni arhiv Združenega kraljestva, Katalog škotske arhivske mreže (SCAN), Vzajemni katalog arhivov kolidžev in univerz (ArchivesHub), Evropski arhivski spletni portal (Archives Portal Europe), UNESCO, NATO.

skega in slovanskega jezikovnega področja. Izbrani arhivi pri popisovanju gradiva uporabljajo tudi različne standarde za popisovanje gradiva (mednarodne in domače standarde). Veliko število različnih arhivov z različno prakso popisovanja daje naši raziskavi posebno težo.

Kvalitativnih raziskav namreč ni mogoče izvajati avtomatsko. Ves postopek je izveden bodisi intelektualno ali ročno, kar pomeni, da so vzorci običajno majhni in homogeni.

Za razliko od teh je naš vzorec izredno velik in heterogen. V našo raziskavo smo zajeli vse slovenske javne arhive, se pravi celotno populacijo, ki arhivsko gradivo popisujejo na osnovi istega programskega orodja, kar nam je zagotovilo izjemno možnost primerjave. Izjemna prednost je bila tudi ta, da smo raziskavo lahko izvajali ne samo preko spletno dostopnega iskalnika, temveč znotraj same produkcijske podatkovne zbirke, v kateri se popisuje gradivo.

### 3.2 Določitev kodirne sheme, kodiranje podatkov in preverjanje zanesljivosti

Kodirno shemo smo določili vnaprej, saj smo vedeli, kaj iščemo. Pri raziskavi smo uporabili prosto dostopni program »Weft. QDA«, vendar le-ta ne omogoča razvoja kompleksnejše kodirne sheme, ki je potrebna pri bolj poglobljenem raziskovanju in se razvije v več krogih analize. Analiza namreč nujno poteka v več krogih, saj s časom opazimo nove značilnosti, ki zahtevajo nove kode. Tudi podpora postopkom analize so v omenjenem programu skromnejši od profesionalnih programov za kvalitativno analizo besedil (npr. Nudist ali Atlas).

Analiza besedila je potekala z natančnim opazovanjem zapisanega besedila. Pozorni smo bili na pojme ter jih označevali z izbranimi kodami. Samo kodirno shemo so delno narekovale že posamezne zvrsti arhivskega gradiva, ki smo jih opazovali v posameznih podatkovnih zbirkah. Prav zaradi tega se kodirna shema med delom ni spreminjala in dopolnjevala in je bila določena pravzaprav že na samem začetku raziskave.

Naslednji korak analize je bilo ugotavljanje pogostosti pojavljanja posameznih kod in njihovih kombinacij. To smo opravili z natančnim poznavanjem podatkov, nismo pa uporabili statistične metode, ki jo lahko uporabimo le, če imamo dovolj veliko količino podatkov. V našem primeru je zaradi nepoznavanja obsega več kot petdesetih analiziranih podatkovnih zbirk nismo mogli izvesti.

Zanesljivost kodiranja bi lahko zagotovili z neodvisnim kodiranjem dveh naključno izbranih oseb. Vendar prav zaradi izjemne velikosti vzorca, smatramo, da raziskava zagotavlja veliko mero zanesljivosti<sup>6</sup>.

Analizo rezultatov kodiranja predstavljamo v nadaljevanju. Za statistično testiranje se nismo odločili oziroma ga nismo mogli izvesti iz razloga, ker nismo imeli podatkov o velikosti velike večine analiziranih podatkovnih zbirk.

## 4 Analiza rezultatov in razprava

V analiziranih informacijskih sistemih obstaja cela paleta različnih izvedb iskalnikov po vsebinah arhivskih podatkovnih zbirk. Le ti omogočajo iskanja tako po polnem tekstu,<sup>7</sup> napredna iskanja po posameznih poljih<sup>8</sup>, po tektoniki arhiva<sup>9</sup> ter seveda po deskriptorjih. Za namen naše raziskave smo se poslužili različnih strategij iskanja. Predvsem smo morali biti inovativni v tistih podatkovnih zbirkah, ki delujejo na različnih iskalnih platformah, medtem ko smo se v tistih podatkovnih zbirkah, ki delujejo v programskem orodju scopeArchiv omejili na iskanje po posameznih poljih, kjer smo iskanje zožili na online dostopne slike in posamezne zvrsti arhivskega gradiva.

6. Gre za raziskavo, ki je bila izvedena na živih podatkovnih zbirkah, kar pomeni, da bodo rezultati ob morebitnem preverjanju zanesljivosti raziskave zagotovo drugačni, saj verjetno vse podatkovne zbirke beležijo prirast arhivskega gradiva.

7. Pri poizvedovanju po podatkovnih zbirkah govorimo o iskanju po polnem tekstu v primeru, ko iskalnik poizveduje v okviru celotne tekstovne podatkovne zbirke. Pri tem iskalnik primerja posamezne besede ali tekst, ki ga je določil uporabnik, znotraj vsakega posameznega zapisa ali dokumenta znotraj podatkovne zbirke.

8. Nekateri iskalniki po arhivskih podatkovnih zbirkah omejujejo iskanje po polnem tekstu na posamezna polja, znotraj katerih so shranjeni podatki, kot npr. naslov, signatura, zvrst gradiva, datum itd.

9. Gre za hierarhično strukturo v obliki klasifikacije ali pa strukturo, urejeno po nekem drugem logičnem sistemu, na osnovi katere so razvrščeni arhivski fondi in zbirke znotraj sistema.

Analiza vseh podatkovni zbirk je pokazala, da ima kar 71% obravnavanih arhivskih informacijskih sistemov v zbirko vključeno online dostopno arhivsko gradivo v obliki digitalizator ali v nekaterih primerih celo izvirnega e-arhivskega gradiva, vendar veliko večino le-tega predstavlja slikovno gradivo, medtem ko je izvirnega spisovnega e-gradiva izjemno malo.

Kar v 47% arhivov najdemo v podatkovno zbirko vključene digitalizate spisovnega gradiva, 17% arhivskih podatkovnih zbirk pa vključuje razglednice. Tukaj je potrebno pripomniti, da so razglednice velikokrat v nekaterih podatkovnih zbirkah označene kot fotografije. Le 11% arhivov v svoje podatkovne zbirke vključuje digitalno nastale fotografije novejšega datuma, medtem ko je procent arhivskih podatkovnih zbirk, v katerih najdemo objavljene digitalizate klasičnih fotografij kar 70%. Digitalizati listin se pojavljajo v 32% analiziranih podatkovnih zbirk, medtem ko je procent arhivov, ki omogočajo spletni dostop do izvirnega e-arhivskega gradiva le 3%. Digitalizate različnih načrtov pa najdemo kar v 56% analiziranih arhivskih podatkovnih zbirk.

Le 29% sistemov je takšnih, ki zaenkrat še ne vključujejo digitaliziranega in izvirnega digitalnega arhivskega gradiva.

Visok odstotek arhivskih podatkovnih zbirk, ki vključujejo online dostopno digitalizirano arhivsko gradivo pravzaprav ni presenetljiv in je zelo v skladu z novimi uporabniškimi trendi. Kot so namreč pokazale nekatere študije, uporabniki danes čedalje pogosteje posegajo po online dostopnem digitaliziranem arhivskem gradivu (prim.: Vilar et al, 2017).

Analiza je tudi pokazala, da velika večina digitaliziranega gradiva zajema slikovno gradivo, in sicer fotografije, razglednice in različne oblike načrtov. Pri tem gre seveda v veliki večini za klasično arhivsko gradivo, izvirne digitalne oblike fotografij zasledimo le v štirih arhivskih informacijskih sistemih.

Med digitalizirano arhivsko gradivo, ki se pogosto pojavlja v arhivskih podatkovnih zbirkah sodijo nedvomno listine, ki predstavljajo, po prepričanju mnogokaterega arhivista, kot tudi uporabnikov, tisto najdragocenejše gradivo, ki ga hranijo posamezne arhivske institucije. Razlog za pojavljanje digitalizator omenjene zvrsti arhivskega gradiva je verjetno iskati v dejstvu, da gre za izjemno občutljivo arhivsko gradivo, ki ga je prav zaradi možnosti digitalizacije na ta način mogoče zaščititi pred poškodbami originala, saj je uporabnikom za raziskovanje zato na voljo digitalizat.

Presenetljivo visok je odstotek arhivskih podatkovnih zbirk, v katerih najdemo online dostopno spisovno gradivo (47%). Predvsem gre za gradivo, ki ga v veliki meri uporabljajo rodoslovci, se pravi vse vrste matičnih knjig, ter predvsem starejše arhivsko gradivo, tisto, ki je na nek način najzanimivejše za uporabnike, kot npr. zemljiško knjižne zadeve, oporoke, gradivo prve in druge svetovne vojne itd.

Zanimivo je, da pa smo v dveh arhivskih podatkovnih zbirkah naleteli celo na izvirno e-arhivsko spisovno gradivo novejšega datuma, ki je v enem primeru dostopno na osnovi Zakona o prostem dostopu do informacij (Freedom of Information Act, 1967). Gre za gradivo, ki ga hrani NARA v Washingtonu (National Archives and Records Administration) in sicer za elektronsko pošto Bele Hiše oziroma predsedniško arhivsko gradivo, ter za gradivo Ameriškega kongresa.

V nobeni izmed analiziranih podatkovnih zbirk pa nismo naleteli na avdio-vizualno gradivo.

Namen analize izbranih arhivskih informacijskih sistemov je bil ugotoviti, kot smo povedali že v drugem poglavju, Ali je popis digitaliziranih vsebin arhivskega gradiva izvedeno tako, da omogoča uporabnikom samostojno poizvedovanje po arhivskih podatkovnih zbirkah in tudi hitrejši in lažji dostop do digitaliziranega in digitalnega arhivskega gradiva, predvsem pa njegovo razumevanje v kontekstu in provenienci?

Odgovor na to vprašanje je le deloma pritrديلen, saj le nekateri analizirani arhivski informacijski sistemi omogočajo iskanje po digitaliziranih vsebinah arhivskega gradiva ter omogočajo tudi iskanje po zvrsti arhivskega gradiva (kombinirano iskanje). Tam, kjer takšno iskanje ni omogočeno, pa seveda digitalizatorov ni bilo mogoče najti, razen v primerih, ko je bil posebej omogočen dostop do posameznih digitalnih zbirk arhivskega gradiva.

Pri tem smo naleteli tudi na posamezne primere, ko je bilo digitalizate arhivskega gradiva posameznih arhivov potrebno iskati v drugih bazah podatkov, ne v arhivskih. To velja predvsem za tisto arhivsko

gradivo, katerega digitalizacijo so opravili zunanji ponudniki in ne sami arhivi. Seveda pa tovrstne rešitve za uporabnike nikakor niso optimalne, saj je v veliki večini primerov tako digitalizirano arhivsko gradivo komercialni produkt, za uporabo katerega je potrebno plačati.

Kot že povedano, je največ digitaliziranega arhivskega gradiva, ki je dostopno online, slikovnega. Vzroke za to moramo seveda iskati v dejstvu, da gre za zvrsti arhivskega gradiva, pri katerih je le v izjemnih primerih potrebno upoštevati določila o nedostopnosti arhivskega gradiva, saj gre za starejše gradivo, kjer v skladu z zakonodajo ni potrebno varovati občutljivih osebnih podatkov. Istočasno pa je to tudi gradivo, ki je vizualno vsečno in po katerem uporabniki velikokrat zelo radi posegajo. Velja pa na tem mestu opozoriti še na eno težavo, na katero lahko naletimo pri objavi tovrstnega gradiva na spletu. Gre za vprašanje avtorskih pravic predvsem na fotografskem gradivu, razglednicah in plakatih novejšega datuma. Arhivi smo na tovrstna vprašanja vse premalo pozorni, in se bomo o tem verjetno pričeli bolj spraševati šele takrat, ko bomo naleteli na prvi večji tovrsten problem.

Največja težava, ki se kaže v analiziranih arhivskih informacijskih sistemih pa je popisovanje digitalizatorov, ki ni v skladu s provenienco, niti v skladu z veljavnim standardom za popisovanje arhivskega gradiva ISAD(G) 2 (2000).

Iz večine obravnavanih in objavljenih ter online dostopnih popisov iz naslova popisne enote ni razvidno, ali gre za popis digitalizata in kakšne vrste je ali za originalne dokumente. Popisani so kot fizični originali z elementi, ki so lastni popisu originalnega dokumenta, digitalizati pa so zgolj pripeti kot slikovno gradivo in so v funkciji metapodatkovnega elementa popisa. S stališča formalne obdelave tako naj opozorimo na prisotne elemente kot so obseg in količina gradiva, pri čemer se digitalizat še vedno opisuje kot 1 fotografija, papir, mere v centimetrih itd., namesto, da bi obseg in kvaliteto opredelili v pikslih, velikost v kilobajtih, pojavnost v formatih kot npr. jpg, ali tif., integriteto vsebine v obliki kontrolne vsote itd.

Zgolj v eni podatkovni zbirki smo naleteli na delno ustrezne opise omenjenih elementov, vendar pa je popis vseboval celo vrsto drugih pomanjkljivosti, predvsem tistih kontekstualnih.

Naslednja težava ki smo jo zasledili pri popisovanju digitalizatorov je njihovo izvzemanje iz konteksta samega gradiva, ki je popisano. Naleteli smo na dva zanimiva primera. V prvem gre za digitalizat rodbinskega grba, ki je sestavni del popisne enote na nivoju združenega dokumenta. Gre za spisovno gradivo o podelitvi plemstva. Uporabnik bi glede na naslov popisne enote in umestitev v kontekst fonda pričakoval digitalizat celotnega spisa, vendar je digitaliziran in objavljen samo grb. Vidimo lahko, da se sama popisna enota ne sklada z digitalizatom, ki je objavljen. Gre za izjemno pogost pojav objave tovrstnega gradiva, saj predstavljajo grbi potencialno izjemno zanimivo gradivo tako za rodoslovce, kot tudi za zgodovinarje. Nikjer v popisni enoti uporabniki niso opozorjeni, da gre za digitalizat, ki je sestavni del neke večje zaključene enote.

V drugem predstavljenem primeru je situacija podobna. Uporabnik tudi tukaj, glede na naslov popisne enote in njeno umestitev v strukturo določenega fonda, verjetno glede na naslov pričakuje poročilo oziroma spis o sami prometni nesreči, a namesto tega najde samo digitalizat fotografije prometne nesreče, ne pa tudi pripadajočega spisa. Za uporabnika zavajajoča je tudi zvrst arhivskega gradiva, saj uporabnika obvešča, da gre za spis.

Pri objavah digitalizatorov spisovnega gradiva, pa smo v mnogih primerih naleteli tudi na objave, ko je namesto celotnega združenega dokumenta, ki je sestavljal določeno popisno enoto, pripet digitalizat predstavljal samo naslovno stran spisa, vsebina pa ni bila digitalizirana. V nekaterih primerih pa je bilo v arhivski podatkovni zbirki navedeno, da zbirka vsebuje digitalne vsebine, vendar le-teh z obstoječim iskalnikom ni bilo mogoče najti.

Digitalizati so v vseh analiziranih podatkovnih zbirkah v tektoniki arhiva umeščeni k popisom fizičnega gradiva (v posamezen ustrezen fond, ki mu pripadajo), ne da bi se pri tem upoštevalo dejstvo, da gre zgolj za izbor digitalizatorov (izrazita implementacija pertinenčnega načela) in ne za digitalizacijo celotnega fonda oziroma zbirke (prikrita implementacija pertinenčnega načela). Digitalizate je namreč potrebno obravnavati enako, kot druge vrste kopij npr. fotokopije. Teh praviloma ne uvrščamo med gradivo že obstoječih fondov, temveč v posebne zbirke dopolnilnih fotokopij.

To je seveda lahko zelo zavajajoče za uporabnike, saj gradivo na ta način ni pravilno umeščeno v

kontekst, kar lahko onemogoča tudi pravilno interpretacijo tako uporabljenega arhivskega gradiva. Velika večina uporabnikov, pa tudi arhivistov, je prepričana, da to ne igra pomembne vloge. Če bi takšno razmišljanje lahko sprejeli s strani uporabnikov, ki jih po večini zanima zgolj vsebina dokumenta, pa bi to moralo predstavljati veliko težavo za arhiviste.

Velike težave pa so se pokazale predvsem v tistih zbirkah digitalizatorov arhivskega gradiva, ki se ne nahajajo v arhivskih informacijskih sistemih, temveč v sistemih ponudnikov ali posebej za to namenjenih podatkovnih zbirkah. Takšen primer so na primer Evropski arhivski portal, Monasterium.net, slikovni arhiv Unesca, digitalizirane podatkovne zbirke Madžarskega nacionalnega arhiva in zbirke, ki jih upravljajo različni ponudniki rodoslovnih in drugih podobnih institucij, kot na primer Ancestry.com in Family Search.org. Težava pri teh digitaliziranih zbirkah se kaže v tem, da so nastale po pertinenčnem principu, predvsem pa, da objavljeni digitalizati nimajo izdelane ustrezne povezave na institucijo izvora, metapodatki, s katerimi pa so opremljeni, uporabnikom ne zagotavljajo ustreznega konteksta, saj so iz njega iztrgani.

O tem smo se prepričali, ko smo v eni izmed velikih podatkovnih zbirk, namenjenih rodoslovnim raziskavam, pregledovali zapise, ki se nanašajo na kraje v Sloveniji in na arhivsko gradivo, ki ga v izvirniku hranijo slovenski arhivi. Ugotovili smo, da posamezne popisne enote nimajo izdelane ustrezne povezave na popis izvirnika, ki ga hrani slovenski arhiv.

Poleg tega popis digitalizata ne ustreza standardom za popisovanje arhivskega gradiva, saj ne vsebuje predpisanih popisnih elementov. Prav tako pa iz popisa ni razviden širši kontekst, v katerega sodi popisani digitalizat.

Namesto tega je ob naslovu zbirke izdelana povezava na posebno spletno stran, kjer najdemo opisne podatke o sami zbirki, kaj vsebuje, kako po njej poiščemo in še kaj. Ponovno pa ne najdemo povezave na izvirno arhivsko gradivo, niti tega, od kod je bilo digitalizirano gradivo pridobljeno.

Pregled omenjenih podatkovnih zbirk pa je pokazal še na eno veliko težavo in sicer se je le-ta pokazala pri zajemanju podatkov v podatkovno zbirko oziroma njihovem gesljenju. Ko smo preverili digitalizate, ki se nanašajo na slovensko arhivsko gradivo, ki je objavljeno v eni izmed prej omenjenih rodoslovnih institucij, smo naleteli na celo vrsto zapisov, ki so napačno gesljeni, oziroma pred tem celo napačno zajeti.

To je samo nekaj nepravilnosti, ki jih je v prvem koraku pokazala naša analiza in ki pod vprašaj postavljajo ustreznost popisovanja digitaliziranega arhivskega gradiva in njegovega umeščanja v kontekst.

## 5 Zaključek

Dejstvo je, da se velika večina arhivov digitalizacije ni lotila arhivsko strokovno celovito in sistematično ob upoštevanju temeljnih načel arhivske teorije in prakse, in da je vse bolj ali manj prepuščeno naključju ter podrejeno posameznim tipom uporabnikov, kot je ugotovi že Popovici (2016) ter finančnim sredstvom, ki jih imajo arhivi na voljo za digitalizacijo.

Avtentičnost, originalnost in izvor predstavljajo pomemben dejavnik v arhivski teoriji in praksi, predvsem v smislu zagotavljanja javne vere arhivskega gradiva. Zato je v procesu digitalizacije arhivskega gradiva kontekst izjemnega in bistvenega pomena za dolgoročno hrambo in zagotavljanje dostopnosti. Lahkota, s katero se digitalizirano arhivsko gradivo lahko spreminja, objavlja ali kako drugače obdeluje kaže na izjemno pomembnost ohranjanja in obvladovanja originalov in avtentičnosti vsebin in kontekstov.

Arhivi bodo morali v prihodnje prevzeti bolj aktivno vlogo pri procesih digitalizacije predvsem v smislu uveljavljanja temeljnih arhivskih principov, saj digitalno arhivsko gradivo arhive postavlja pred velike izzive. Samo na ta način bo možno zagotoviti javno vero digitaliziranega arhivskega gradiva v času in prostoru.

Kot je pokazala izvedena analiza, arhivisti vse bolj pogosto pozabljamo na naše temeljne principe, ki pa bistveno vplivajo na popisovanje digitaliziranega arhivskega gradiva in na zagotavljanje javne vere nam zaupanega arhivskega gradiva.

## Viri in literatura

- Berelson, B. (1952). *Content Analysis in Communication Research*. Glencoe, Ill: Free Press.
- Bultmann, Barbara, Rachel Hardy, Adrienne Muir & Clara Wictor (2005). Digitised Content in the UK Research Library and Archives Sector. A report to the Consortium of Research Libraries and the Joint Information Systems Committee. Dostopno na [http://www.rluk.ac.uk/wp-content/uploads/2014/02/Loughborough\\_Digitization\\_Report.pdf](http://www.rluk.ac.uk/wp-content/uploads/2014/02/Loughborough_Digitization_Report.pdf) (28.12.2016).
- Content analysis (2012). V: Wikipedia: The free encyclopedia. Sneto 28.01.2017 s spletne strani [http://en.wikipedia.org/wiki/Content\\_analysis](http://en.wikipedia.org/wiki/Content_analysis).
- Cook, Terry (1997). What is Past is Prologue: A History of Archival Ideas since 1898, and the Future Paradigm Shift. *Archivaria* 43 (Spring), str. 16-63.
- Digitising accumulated physical records..... Dostopno na <http://www.naa.gov.au/information-management/managing-information-and-records/capturing/digitising-accumulated-records/index.aspx>* (28. 12. 2016).
- Elo, S., Kyngäs H. (2008). The Qualitative Content Analysis Process. V *Journal of Advanced Nursing* 62(1), str. 107-115. Blackwell Publishing Ltd.
- Freeman, Elsie (1984). In the Eye of the Beholder: Archives Administration from the User's Point of View. *American Archivist* 47(2), str. 111-123.
- GAO - U.S. General Accounting Office (1996). *Content Analysis. A Methodology for Structuring and Analysing Written Material*. Washington: Programme Evaluation and Methodology Division, United States General Accounting Office.
- Guidelines for Digitization Projects for collections and holdings in the public domain. Draft (2003). Dostopno na <http://www.ifla.org/files/assets/preservation-and-conservation/publications/digitization-projects-guidelines.pdf> (28.12.2016).
- Hickerson, Thomas H. (2001). Ten Challenges for the Archival Profession. *American Archivist* 64 (Spring/Summer), str. 11.
- Hirtle, Peter B., Hudson, Emily, Kenyon, Andrew T. (2009). Copyright and Cultural Institutions. Guidelines for Digitization for U.S. Libraries, Archives and Museums. Ithaca, new York: Cornell University Library. Dostopno na [https://ecommons.cornell.edu/bitstream/handle/1813/14142/Hirtle-Copyright\\_final\\_RGB\\_lowres-cover1.pdf;jsessionid=C44B9A9624BC68AC3CB0B959C3247FBB?sequence=2](https://ecommons.cornell.edu/bitstream/handle/1813/14142/Hirtle-Copyright_final_RGB_lowres-cover1.pdf;jsessionid=C44B9A9624BC68AC3CB0B959C3247FBB?sequence=2)* (28.12.2016).
- Holsti, O. R. (1969). *Content Analysis for the Social Sciences and Humanities*. Reading, MA: Addison-Wesley.
- Hughes, Lorna M. (2003). Digitizing Collections. Strategic issues for the information manager. Facet Publishing. Dostopno na [http://www.facetpublishing.co.uk/downloads/file/sample\\_chapters/digitizing%20collections%20chapter%201.pdf](http://www.facetpublishing.co.uk/downloads/file/sample_chapters/digitizing%20collections%20chapter%201.pdf) (28.12.2016).
- ISAD (G) – General International Standard Archival Description. Pridobljeno 21. 08. 2009 s spletne strani [http://www.ica.org/sites/default/files/isad\\_g\\_2e.pdf](http://www.ica.org/sites/default/files/isad_g_2e.pdf).
- Koltun, Lilly (1999). The Promise and Threat of Digital Options in an Archival Age. *Archivaria* 47 (Spring), str. 114–135.
- Krippendorff, K. (2004). *Content analysis: an introduction to its methodology*. 2nd ed. Thousand Oaks, CA: Sage.
- Mark Vajcner (2008). The Importance of Context for Digitized Archival Collections. *Journal of the Association for History and Computing*, vol. 11, no. 1, April 2008. Dostopno na <http://hdl.handle.net/2027/spo.3310410.0011.102> (28.12.2016).
- McCraken, Krista (2014). Archival Digitization and the Struggle to Create Useful Digital Reproductions. Dostopno na <http://activehistory.ca/2014/01/archival-digitization-and-the-struggle/> (28.12.2016).
- McKay, Sally (2003). Digitization in an Archival Environment. *Electronic Journal of Academic and Special Librarianship*, v. 4, no. 1 (Winter 2003). Dostopno na [http://southernlibrarianship.icaap.org/content/v04n01/Mckay\\_s01.htm](http://southernlibrarianship.icaap.org/content/v04n01/Mckay_s01.htm) (28.12.2016).
- Miller, Larisa (2013). All Text Considered: A Perspective on Mass Digitizing and Archival Processing. *The American Archivist: Fall/Winter, Vol. 76, No. 2*, pp. 521-541.
- Moravec, Michelle (2016). How digitized changed historical research. Dostopno na <https://medium.com/on-archivy/how-digitized-changed-historical-research-d77c78540878#.tiwln09mj> (28.11.2016)
- Novak, M. in Horvat, M. (2014). Posebnosti digitalizacije arhivskega gradiva in strokovna obravnava tovrstnih digitalizatorov. *Knjižnica*, 58 (2014) 1-2, str. 151–177.

## Zdenka SEMLIČ RAJH: Problemi popisovanja digitalnega in digitaliziranega arhivskega gradiva, 193-204

Owens, Trevor (2013). Mass Digitization, Archives, and a Multiplicity of Orders & Arrangements. Dostopno na <http://www.trevorowens.org/2013/12/mass-digitization-archives-and-a-multiplicity-of-orders-arrangements/> (28.11.2016).

Peräkylä, A. (2005). Analyzing talk and text. V *Denzin, N. K., Lincoln, Y. S. (ur.), The Sage handbook of qualitative research*, str. 869–886. Thousand Oaks, Ca.: Sage.

Roblek, V. (2009). Primer izpeljave analize besedila v kvalitativni raziskavi. V *Management*, letnik 4, št. 1, str. 53–69). Sneto 5. 1. 2013 s spletne strani <http://www.dlib.si/?URN=URN:NBN:SI:doc-34HZGHX>.

Semlič Rajh, Z., Šabotič, I., Z., Šauperl, A. (2013). Znanstveno raziskovalno delo v arhivistiki: značilnosti uporabe dveh raziskovalnih metod. V Gostenčnik, N. (ur.), *Tehnični in vsebinski problemi klasičnega in elektronskega arhiviranja*, str. 125–144. Maribor: Pokrajinski arhiv Maribor.

Semlič Rajh, Z., Šauperl, A. (2013). Analiza oblikovanja vsebine zajetih podatkov v podtkovni bazi SIRAnet. V Gostenčnik, N. (ur.), *Tehnični in vsebinski problemi klasičnega in elektronskega arhiviranja*, str. 145–157. Maribor: Pokrajinski arhiv Maribor.

Stemler, S. (2001). An Overview of Content Analysis. V *Practical Assessment, Research & Evaluation*, 7(17). Sneto 5.1.2013 s spletne strani <http://pareonline.net/getvn.asp?v=7&n=17>.

Šauperl, A. (2008). Kvalitativne raziskovalne metode. V Šauperl, A. (ur.), *Raziskovalne metode v bibliotekarstvu, informacijski znanosti in knjigarstvu*, str. 149–162). Ljubljana: Oddelek za bibliotekarstvo, informacijsko znanost in knjigarstvo, Filozofska fakulteta, Univerza v Ljubljani.

Šauperl, A., Semlič Rajh, Z. (2012). Značaj področja identifikacije popisnih jedinic i sadržaja prema standardi ISAD(g) u slovenskoj bazi podataka SIRAnet. V *Arhivska praksa*, št. 15(2012), str. 323–334. Tuzla: Arhiv Tuzlanskog kantona. Sneto 5.1.2013 s spletne strani <http://www.arhivtk.com.ba/ARHIVSKA%20PRAKSA%2015.pdf>.

Viler, P., Šauperl, A., Semlič Rajh, Z., Robinson, L., Bowden, D. (v tisku). Information competencies of historians as archive users: A Slovenian/UK comparison. V Serap et al (ur.), *Information Literacy: Key to an Inclusive Society*. Springer.

Weber, R. P. (1990). *Basic Content Analysis*, 2nd ed. Newbury Parc, CA: Sage.

White, Marilyn Domas in Marsh, Emily (2006). Content Analysis: A Flexible Methodology. V Baker, Linda M. (ur.), *Library Trends*, vol. 55, No.1, Summer 2006, str. 22–45.

Zupan, M., Černigoj Blažko, L. (2014). Nacionalni program za kulturo 2014–2017: pot do novega modela kulturne politike. Ljubljana: Ministrstvo za kulturo.

## SUMMARY

The digitization of archival records is not specific only from the standpoint of the digitization process itself, but also from the standpoint of the description of digitized archival records. Adequate description namely enables the wide use of digitized archival records. The author first presents the literature and accomplished research in the field of digitization and description of digitized archival records. It is necessary to emphasize that the archival profession, not only domestic, but also foreign, was very much engaged with the issue of digitization, however, discussions mostly dealt with technological aspects of digitization, its weaknesses and strengths, ethical and legal aspects, as well as issues of copyright of digitized archival records. Authors and articles relating to technological aspects of digitization are not specifically mentioned in this presentation, as, in our study, we were interested in the description and the importance of context in describing digitized archival records. However, articles in which authors deal with the description of digitized archival records or with questions related to the meaning and context of digitized archival records are extremely rare. As the main method for the purposes of our research, we have used Content Analysis. It is a relatively unknown method in the Slovenian archival theory and practice, which, until recently, was not applied, even though it is a research method widely used both in social as well as natural science research. The analysis and the results of the performed research of description of digitized archival records in online assessable archival information systems are presented according to steps proposed by White and Marsh (2006, p. 30) with additional questions to be asked according to Krippendorff (2004). The analysis was conducted in 48 archival information systems and their databases. These are large systems, some of which operate as a common database of several smaller archives, but some systems are individual systems of large national archives. The analysis was conducted in publicly accessible user interfaces, and was based on a direct applicability of descriptions of digitized archival records within an individual database. Such query can be performed by an average user of archival records with access to the search engine knowing at least the basic functions of the web browser and searching and sorting of results. With our research, we tried to answer the following questions: Is the description of digitized contents of archival records carried out in the way that enables users to search within the database independently? Does the

description enable easier access to digitized archival records and in particular, the understanding of the context and provenance? The conducted analysis was based on the determination of the following facts:

- Does the system enable the search of online accessible archival records?
- According to which genres is it possible to search in the system (files, postcards, photographs, charters, born digital records, plans etc.)? This is important information that shows a variety of options for accessing digitized archival records and searching within the archival database;
  - Does added metadata indicate that we describe digitized archival records?
  - Are the described content and medium of the unit of description given for physical archives or attached digital images?
  - Are the titles of the units of description properly designed and enable proper identification of the described records?
  - Does a search engine enable the search of digitized archival records and can the search engine be used by an average user?

For the analysis, we chose the archives, where archival information systems are available online and allow retrieval of archival databases. Analysed Archives use different description practices, as they belonging to different linguistic areas, both Anglo-American, Germanic and Slavic language areas. Selected Archives also use different description standards.

Typology: 1.02 Review Article

Submitting date: 20.05.2017

Acceptance date: 31.05.2017