

KAKOVOST PODATKOV V KNJIŽNIČNIH KATALOGIH IN ARHIVSKIH PODATKOVNIH ZBIRKAH

Alenka Šauperl

*Oddelek za bibliotekarstvo,
informatijsko znanost in knjigarstvo,
Filozofska fakulteta, Univerza v
Ljubljani, Ljubljana, Slovenija*

Zdenka Semlič Rajh

*Pokrajinski arhiv Maribor, Maribor,
Slovenija*

KLJUČNE BESEDE:

*knjižnični katalog, arhivska
podatkovna zbirka, kakovost
zapisov, napake*

IZVLEČEK

Namen: Večini ukvarjanje s kakovostjo podatkov v knjižničnih katalogih in arhivskih podatkovnih zbirkah ni prijetno. Vendar smo po prehod iz listkovnih knjižničnih katalogov na računalniške in iz ročnih arhivskih evidenc za podatkovne zbirke soočeni z dejstvom, da le ustrezna kakovost zagotavlja tudi ustrezen priklic gradiva in informacij.

Metodologija/pristop: Pregled literature je strukturiran po sedmih vidikih kakovosti metapodatkov: popolnost, točnost, vir, skladnost s pričakovanji, doslednost, pravočasnost in zaostanek ter dostopnost.

Rezultati: Meniva, da je nujno treba do tega problema zavzeti aktivnejši odnos. Upava, da bova s svojim pregledom bralce in bralke prepričali, da bi mogli vsakdo prispevati svoj delež h kakovosti podatkov v zbirkah tako, kot prispevamo vsak svoj delež k recikliranju odpadkov.

Zaključki: Bibliografski vir ali arhivsko gradivo je kljub napaki v kataložnem zapisu ali arhivski podatkovni zbirki mogoče najti, kadar slučajno ustreza poizvedbi. Malo verjetno pa je, da takrat ustreza uporabničini oziroma uporabnikovi potrebi. Visoka kakovost podatkov v knjižničnih katalogih in arhivskih podatkovnih zbirkah seveda prinaša točen opis bibliografskih virov in arhivskega gradiva. Dosega pa tudi drugi cilj, ki je za obstoj knjižnic in arhivov morda celo pomembnejši, namreč zaupanje javnosti.

Uvod

Kakovost podatkov v knjižničnih katalogih in arhivskih podatkovnih zbirkah večinoma jemljemo kot samo po sebi umevno. Večinoma smo katalogizatorji, katalogizatorke, arhivisti in arhivistke same pri sebi prepričane, da smo nezmotljive in se napake zgodijo le drugim. To (zgrešeno) »teorijo« nam potrdijo napake, ki jih dnevno najdemo v zapisih drugih. Človeška hiba je pač, da svojih napak navadno ne opazimo. Seveda je delo, povezano z iskanjem in popravljanjem napak neprijetno, težavno in zamudno. Zahteva veliko znanja, raziskovanja in predvsem dolgotrajno in popolno pozornost. To so značilnosti, ki v sodobni družbi izgubljajo vrednost in veljavo. Razen tega, da za to prepotrebno delo ne moremo pričakovati pohvale ali nagrade, v tem času naše tekoče delo stoji in se nam zaradi tega celo obeta zaostanek. Zaskrbljujoče malo razlogov je, da bi se ukvarjali s popravljanjem napak, čeprav vodijo v dezinformacijo.

Poleg tega (ne nujno tudi dejansko izkušene) osebnega doživljanja je glede priljubljenosti te teme mogoče najti tudi nekaj dejstev. ProQuestova zbirka *Library & Information Science Collection*, ki je dostopna za Univerzo v Ljubljani, nam z iskalnima izrazoma »data quality« in »library catalogs« v izvlečkih za čas od 2008 do 2018 ponudi 15 člankov, zbirka *Library, Information Science and Technology Abstracts* 16 člankov in Emeraldova podzbirka *Library studies* 116 člankov. Branka Badovinac (2017) je v svojem izredno temeljitem pregledu literature uporabila 86 virov (ne čisto vseh na temo kakovosti podatkov). V podrobno vsebinsko analizo z ozirom na raziskovalne usmeritve na področju kakovosti (meta)podatkov v katalogizaciji, digitalnih knjižnicah in repozitorijih pa je vključila 27 virov, objavljenih v času 2003-2016.

Vendar neprijetnost ne pomeni tudi nepotrebnosti. Dogaja se, da je prav tisto, kar je za skupnost najvažnejše, na osebni ravni najmanj zaželeno. Ravnanje z odpadki odseva zelo podoben problem. Načeloma se vsi zavzemamo za ločevanje odpadkov, zmanjševanje količine embalaže in varčevanje z vodo. Ali vsi in vedno to tudi dosledno počnemo? Če za odpadke velja, da so lahko dobra surovina, pri podatkih in računalniških sistemih ni tako. Kot študentki so naju učili, da pri računalniških sistemih velja geslo: »garbage in – garbage out«. To pomeni, da slabi podatki ne morejo obroditi dobrih rezultatov.

V knjižnicah tema o kakovosti podatkov ni tuja, saj se s to problematiko ukvarjamo že od začetka delovanja sistema COBISS. Prvo poročilo je verjetno zapisala Zlata Dimec (1994). Izvedli smo tudi nekaj raziskav (npr. Likar (2003), Likar in Žumer (2004), Seljak (2006), Krstulović (2006), Pesjak in Petek (2010)). Za arhi-

ve pa je to problematika, ki se je pravzaprav do sedaj še nismo posebej lotevali. Zaradi vse večjega obsega arhivskih podatkovnih zbirk, ki so večinoma dostopne neposredno preko spleta, pa je zelo pereča. Eno pravih obširnejših tovrstnih raziskav sva izvedla Novak in Semlič Rajh (2013), deloma pa sva se tovrstne problematike v eni izmed svojih raziskav lotili tudi Semlič Rajh in Šaupperl (2013).

Namenoma se bova izognili navajanju rezultatov raziskav, ki pričajo o napakah v knjižničnih in arhivskih katalogih. Čeprav so številne in nam niso v čast, ne želiva opozarjati nanje. Želiva le postaviti okvir, v katerem bi imel vsakdo priložnost premisliti, kaj bi mogel sam ali sama prispevati k boljši kakovosti zapisov v knjižničnem katalogu in arhivski podatkovni zbirki. Kakovost kataloga in podatkovne zbirke ni odvisna le od vsakega posameznika in posameznice, vendar je osebni strokovni prispevek nujni del skupnega napredka. Nekateri prispevajo z razvojem sistema, drugi s požrtvovalno uporabo sistema, tretji morda celo z razlogi, zaradi katerih se sistemu izogibajo. V nadaljevanju bova najprej predstavili teoretično izhodišče sedmih kriterijev za oceno kakovosti. Nato bova kot vzorec ponudili svoj premislek v tem okviru.

Teoretično izhodišče za kakovost metapodatkov

Kakovosti ne pomeni le tega, da v metapodatkih ni napak. Bruce in Hillmann (2004) za okolje knjižnic in digitalnih zbirk predlagata sedem kriterijev. Predstavljava jih z vidika knjižničnih katalogov:

- Popolnost¹ (*completeness*): Metapodatki naj bi bili popolni v dveh ozirih. Prvič, podatki bi morali popolnoma opisati bibliografski vir. Vedno si lahko predstavljamo, da bi bilo mogoče nekaj podrobneje opisati. Toda to ni vedno mogoče ali ekonomično. Drugič, nabor podatkov, ki je predviden za opis bibliografskega vira, bi bilo treba v celoti izkoristiti. Če je določen podatek v zbirki redko prisoten, ni samo nesmiseln, temveč tudi zavajajoč. Uporabnice in uporabniki bi utegnili zaman pričakovati, da ga je mogoče uporabiti.
- Točnost (*accuracy*): Metapodatki bi morali odražati dejanske značilnosti bibliografskega vira. Tudi tu imamo po Burce in Hillmann (2004) dva vidika. Prvič, podatki bi morali pravilno opisovati značilnosti. Drugič, gre za kako-

1 Prevod izrazov le delno povzemava po Branki Badovinac (2017), ki je sicer uporabila kategorizacijo po Batini in Scannapiecco (2016). V pričujočem prispevku uporabljava nekoliko drugačno kategorizacijo Bruce in Hillmann (2004).

vostno lektoriranje. Slovnice in tipkarske napake se ne bi smele pojavljati. Imena oseb, družin, korporacij, krajev itd. bi morala biti zapisana pravilno. Kratice bi morale biti oblikovane po standardnem postopku.

- Vir (*provenance*): Včasih je mogoče kakovost metapodatkov oceniti na podlagi njihovega ustvarjalca. Izkušeni in razgledani strokovnjaki in strokovnjakinje imajo v tem smislu navadno večji ugled od novincev in novink v stroki. Upamo tudi, da imamo strokovnjaki in strokovnjakinje pri tem tudi večji ugled od laikov.
- Skladnost s pričakovanji (*conformance to expectations*): V knjižničnem katalogu pričakujemo, da bodo podatki skladni z mednarodnim standardom za bibliografski opis, katalogizacijskim pravilnikom in formatom COMARC oz. UNIMARC. Nasplošno gre po eni strani za skladnost z metapodatkovno shemo, po drugi pa za skladnost te sheme s pričakovanji in zahtevami uporabnic in uporabnikov. Metapodatkovna shema mora zajemati primerne podatke, take, ki so z vidika uporabe potrebni in so smiselni za dejavnosti, katerim je informacijski sistem namenjen. Bruce in Hillmann (2004) opozarjata na podobnost med skladnostjo s pričakovanji in popolnostjo. Če uporabnice in uporabniki menijo, da je nek metapodatek nepotreben, gre za problem skladnosti s pričakovanji, ne pa za problem popolnosti.
- Doslednost (*logical consistency and coherence*): Bruce in Hillmann (2004) menita, da je za dobro kakovost podatkov nujna dosledna uporaba enakih metapodatkov skozi čas. To ne služi le temu, da bodo vsi bibliografski viri v naši zbirki opisani enotno, z enakimi elementi, temveč tudi gladkemu spajanju oz. združevanju različnih podatkovnih zbirk. Pozorni pa moramo biti tudi na vrednosti, ki jih vključujemo v določen metapodatkovni element. Za iskanje in priklic bo velika ovira, če bomo nekaj časa, ustvarjalca oziroma ustvarjalko vpisovali v določeno polje v zaporedju priimek-ime, kasneje pa obratno, ime-priimek.
- Pravočasnost (*timeliness*): Po Bruce in Hillmann (2004) se je treba v zvezi s časom zavedati dveh vidikov: Eden je vidik sodobnosti. Ta pomeni, da je nek podatek vsak trenutek točen in veljaven. Kot primer navajata URL. Če se ta spremeni, zapis v podatkovni zbirki pa ne sledi spremembi, gre za neposodobljen podatek. Drugi je vidik zaostanka (*lag*), ki odraža čas, ki je npr. potreben od spremembe URL-ja do popravka v knjižničnem katalogu.
- Dostopnost (*accessibility*): Dostopnost ima številne vidike. Zgodi se lahko, da podatka ne moremo prebrati. Berljivost je lahko odvisna od pisave besedila (npr. če ne znamo brati ruske cirilice ali japonske pisave). Zgodi pa se tudi,

da format datoteke starejšega datuma ne omogoča odpiranja in berljivosti v sedanjem času. Ni nujno, da so digitalni podatki vsem dostopni, saj morda ponekod nimajo ustreznih naprav in omrežij. Podatki so lahko dostopni brezplačno ali le proti plačilu.

Ni mogoče vedno in v vseh primerih vnaprej določiti, kateri od opisanih kriterijev je pomembnejši. Bruce in Hillmann (2004) menita, da je pomembnost odvisna od mnogih dejavnikov, npr. od narave informacijskih objektov, podatkovne zbirke, uporabnic in uporabnikov, namena podatkov in zbirke itd. Najin prispevek ne namerava kateremu od kriterijev dati prednosti. Našteti so v vrstnem redu, kakršnega sta uporabila Bruce in Hillmann (2004). Tudi v nadaljevanju, kjer o kriterijih premišljujeta na podlagi osebnih izkušenj z iskanjem v slovenskem vzajemnem bibliografskem sistemu COBISS, bova obdržali enak vrstni red. Najin premislek je namenjen seznanjanju in ozaveščanju strokovne javnosti.

Pogled v slovenski vzajemni katalog s kategorizacijo napak po Bruce in Hillman (2004)

Popolnost

Neobstoječi podatki v elementih mednarodnega standardnega bibliografskega opisa in kataložnega zapisa v formatu COMARC

Tako ISBD(M) kot format COMARC/B² omogoča ustvarjanje različno popolnih kataložnih zapisov (*polje 001a—Status zapisa*). To je tudi nujno, saj nimamo vedno na razpolago potrebnih podatkov (kar pa sodi v poglavje 3.1.2). Manj popolni so zapisi za publikacije, ki so še v fazi nabave ali zapisi CIP (*cataloging in publication*). Najpopolnejši so popravljeni zapisi, v katerih so popravljene tudi tipkarske napake (COMARC/B, sep. 2018). Sistem, ki podpira celotno knjižnično poslovanje mora podpirati in omogočati nepopolne zapise za bibliografske vire, npr. za obdobje nabave. Čeprav je predvideno, da bo zapis le krajši čas v okrnjeni obliki, se včasih zgodi, da tak ostane za vedno. Manj kot ima zapis podatkov, manj je možnosti za njegov priklic. Če zapis ni priklican, je tudi malo možnosti, da bo bibliografski vir uporabljen.

2 Format COMARC/B je ustvarjen na podlagi formata UNIMARC.

Popoln kataložni zapis

Slovenske knjižnice do nadaljnjega upoštevajo »stare« standarde za bibliografski opis,³ npr. ISBD(M). Ta standard predvideva obvezne in neobvezne elemente bibliografskega opisa. Obvezna je npr. prva navedba odgovornosti, ne pa tudi naslednja. Prav tako je obvezna prva navedba založnika, ne pa tudi naslednjih. Ni obvezna navedba podatkov o tiskarju, kraju in času tiskanja ter navedba sprejete gradiva. Tudi v formatu COMARC/B niso vsi podatki obvezni, ker ni nujno, da vedno obstojijo (kar sodi v poglavje 3.1.1). Pričakovati pa je mogoče, da nekateri kataložni zapisi ne bodo vsebovali neobveznih podatkov, četudi bodo obstajali. Seveda to pomeni, da bo iskanje in najdenje mogoče samo po prvih oz. osnovnih navedbah.

Dve desetletji je vmesnik za iskanje COBISS/OPAC sicer omogočal izbiro kodiranih podatkov za vrsto vsebine (bibliografije, slovarji, učbeniki itd.) ter literarno vrsto (poezija, dramatika, romani, ljudske pesmi itd.), vendar je bila uporaba mogoča le v izbirnem iskanju, ki ga je uporabljal le zanemarljiv delež ljudi (Merica, 2011; Razpotnik, 2007). Ker podatek ni bil obvezen, se marsikomu ni zdelo vredno zapravljati časa za vpis. Veliko zapisov je ostalo brez tega podatka.⁴ Potem pa se je nenadoma zgodil preobrat. Uporabniški vmesnik Cobiss+, ki je začel delovati leta 2017, v fasetni navigaciji omogoča izbiro literarnega ali neliiterarnega žanra. Podatkov v zapisih pa ni bilo! Nimava podatka o tem, koliko uporabnic in uporabnikov to možnost v fasetni navigaciji uporablja. Vendar to niti ni pomembno. V vsakem primeru je namreč podatek zavajajoč, če ne odraža dejanskega stanja.

Te manjkajoče podatke bodo morda sčasoma dopolnili. Razprava o strošku sprotnega vnosa na videz nepotrebnega podatka in o dopolnjevanju starih zapisov z nekoč izpuščenim podatkom, bi bila sicer zanimiva. Ker pa nimava niti izkušenj niti podatkov, nimava osnove za izračun. Predvidevava, da bi bil sprotni vnos cenejši. Četudi ne gre za ročni vnos, temveč za nek algoritem, ki zmore tak podatek dopolniti, mora tak algoritem nekdo ustvariti.

Primer nas nauči dvoje: Prvič, vpisovati je treba vse podatke, ki jih predvideva standard in format za računalniški vnos. To se tudi ujema z mnenjem v Bruce in Hillmann (2004). Drugič, obstoj podatkov, razvoj podatkovne sheme, in razvoj uporabniškega vmesnika so tesno povezani in odvisni drug od drugega.

3 Glej pojasnilo Komisije za katalogizacijo pri NUK v poglavju Opisna katalogizacija. Dostopno na spletni strani Komisije: <https://www.nuk.uni-lj.si/nuk/komisija-za-katalogizacijo/opisna-katalogizacija>.

4 Dokaz o tem je mogoče najti v razlagi iskanja vzorca za raziskavo vsebinskega opisa leposlovja (Šauperl, 2016).

Willer, Šauperl, Petek in Tomić (2011) smo ugotovile, da v hrvaškem skupnem katalogu CROLIST in slovenskem COBISS velikokrat manjka podatke o izvirnem naslovu. Povezava izvirnega naslova in naslova prevoda je potrebna za združevanje kataložnih zapisov. Če bi ta podatek imeli, bi bilo v fasetni navigaciji v COBISS+ mogoče ponuditi izbiro dela glede na izdajo. Namesto dolgega seznama zadetkov različnih izdaj istega dela, bi se nam npr. lahko prikazal en sam zadek in se nam ponudila izbira različnih izdaj. V duhu funkcionalnih zahtev bi rekli, da bi nam vmesnik ob iskanju nekega dela zadetke razvrstil po izraznih in pojavnih oblikah.

Odsotnost nekaterih bibliografskih podatkov je z vidika priklica nedvomno škodljiva. Meniva pa, da utegne biti odsotnost vsebinskega opisa v kataložnem zapisu celo usodna. V eni od raziskav vsebinskega opisa književnih del v slovenskem vzajemnem katalogu je bilo ugotovljeno, da je bil edini element vsebinskega opisa vrstilec UDK pri kar dveh tretjinah romanov.⁵ Le ena tretjina jih je imela tudi predmetne oznake (Šauperl, 2016). Vemo, da uporabnice in uporabniki pri iskanju ne uporabljajo vrstilcev UDK (Merica, 2011; Razpotnik, 2007) in da imajo ne le javnost, temveč tudi strokovno knjižnično osebje težave z njim (Žaucer, 2007). V takem primeru je jasno, da uporabnega vsebinskega opisa pri dveh tretjinah romanov ni in iskanje po vsebini ni mogoče. Če v zapisu obstojijo vsaj kode za literarno vrsto, je ta vidik vsaj delno nadomeščen. Ker je ta koda uporabljena v fasetni navigaciji, imamo možnost uporabiti vsaj oznako oblike oz. žanra (*isness*), če že ne vsebine (*aboutness*⁶). Formata UNIMARC in COMARC imata tudi kodo za vrsto vsebine, ki prav tako opisuje neke vrste žanr. Z njo izrazimo, da je delo enciklopedija, slovar, učbenik itd. Čeprav je ta oznaka zelo informativna, ni dovolj, saj sama ne pove, s katerega področja je enciklopedija, priročnik ali slovar. Prav tako pa koda za literarno vrsto ne pove, o kom govori biografski roman ali o katerem dogodku piše zgodovinski roman.

Slovenski arhivi pri popisovanju arhivskega gradiva in njegovem opisu upoštevajo mednarodne arhivske standarde, npr. ISAD(G) (2000), ISAAR (CPF) (2004) in druge. Ti standardi, tako kot bibliotekarski, predvidevajo obvezne in neobvezne elemente opisa. Tako npr. standard ISAD(G), ki obsega 26 popisnih elementov, kot obvezne opredeljuje zgolj 6 osnovnih elementov.⁷ Vsi ostali so prepuščeni arhivistki oziroma arhivistu. Pri tem je potrebno omeniti, da je cela vrsta elementov v popisnem obrazcu sistema za popisovanje oblikovana v obliki

5 Vzorec je tvorilo 179 izvirnih slovenskih romanov iz leta 2014 leta 2015.

6 Za pojasnilo o pojmih *isness* in *aboutness* glej Hjørland (2016).

7 Obvezni elementi popisa so signatura popisne enote, naslov, čas nastanka, nivo popisa, količina in ustvarjalec (prim. ISAD(G), 2000).

spustnih seznamov, ker pa podatki niso obvezni, prevladuje prepričanje, da gre v primeru vnašanja teh podatkov za zapravljanje časa.

Novak in Semlič Rajh (2013) sva ugotovila, da imajo v skupni podatkovni zbirki vsi zapisi opredeljen nivo popisa, 99,84% zapisov ima opredeljen naslov, 99,96% zapisov signaturo ter čas nastanka popisne enote 99,45% zapisov. Količine popisne enote pa vsebuje komaj 60% zapisov glede na vse zapise v zbirki. Ker gre za obvezne elemente po standardu ISAD(G) bi teoretično morali znašati vnosi omenjenih elementov 100%. A teorija je eno, praksa pa velikokrat nekaj drugega. Razlike se lahko pojavljajo zaradi narave popisovanja gradiva, velikokrat pa je razlog predvsem v nedoslednosti popisovalcev in popisovalk ter odsotnosti končne redakcije popisov. Zaskrbljujoče pa je dejstvo, da se v zadnjih šestih letih ni veliko spremenilo, saj nam pogled v Statistiko popisnih enot v sistemu SIRAnet z dne 25. 3 2019 pokaže, da popisovalci in popisovalke ne upoštevajo niti pravila vnosa obveznih elementov, saj je še vedno edini element dosledno vpisani podatek element nivo popisa, ki pa ga sistem generira avtomatsko in ni odvisen od popisovalca oziroma popisovalke.

Na veliko težav glede popolnosti zapisov pa v arhivih naletimo v primerih, ko prihaja do prenosa že obstoječih popisov v nov informacijskih sistem. Težava je v tem, da ti popisi niso strukturirani na osnovi sedaj veljavnih standardov za popisovanje, kar seveda povzroča veliko težav pri prenosu, predvsem v smislu pomanjkanja določenih informacij, na osnovi katerih bi bilo možno izdelati popolnejši opis, kot npr. nivo popisa. Seveda takšno pomanjkljivo opisovanje arhivskega gradiva in nepravilno oblikovanje naslova popisne enote ter vsebine, predvsem pa tudi pomanjkljivo oblikovanje normativnih zapisov, na osnovi katerih bi bilo mogoče uspešno poizvedovati po podatkovni zbirki, pomeni, da bo iskanje in priklic relevantnega gradiva za uporabnico oziroma uporabnika, ki iščejo preko spleta, izjemno težavno. Nedvomno je za uporabnice in uporabnika to lahko usodno.

Točnost

Pravilnost opisa značilnosti

Tu bi lahko govorili o dveh vidikih:

- Če posamezni bibliografski podatek ni vpisan v pravi element bibliografskega opisa oz. pravo polje formata COMARC, potem ne opisuje realnega bibliografskega vira. Že zamenjava vloge med avtorjem in ilustratorjem

lahko uporabnice in uporabnike zmede in strokovnemu osebju knjižnice vzbudi vprašanje, ali gre za različni deli.

- Nepravilen opis bibliografskega vira je enostavno zagrešiti, če tvorimo nov kataložni zapis na osnovi obstoječega (kopiranje). V naglici nekaterih podatkov ne popravimo in zapis je napačen. Tu gre za pravo napako, npr. če bi za dvaindvajseti seminar *Arhivi, knjižnice, muzeji*, ki je bil v Poreču, zapisali, da je bil v Rovinju.⁸

Take napake lahko nastanejo tudi pri optičnem prepoznavanju besedil in pri prenosu podatkov iz enega v drug sistem. Pri optičnem prepoznavanju besedil izziv predstavljajo nekakovosten odtis, oblika črk in morebitne lise v papirju. Take napake je v nekaterih primerih lažje, drugih težje opaziti in popraviti.

Če gre za podatke v napačnih elementih bibliografskega opisa ali poljih formata COMARC, jih je včasih še mogoče opaziti v samem kataložnem zapisu, četudi ne poznamo bibliografskega vira. Za ugotovitev napačnih podatkov v pravih elementih in poljih pa je treba vir poznati. Take napake največkrat najdemo le po naključju.

Pravopis in slovnica

Tipkarska napaka ne odraža resničnosti. Priimek ene od avtoric je »Šauperk« ne »Šaupler«. Ne gre pa vedno za preproste tipkarske napake. Nekateri tuji viri ta priimek dosledno pišejo »Sauperl«, ker upoštevajo le črke angleške abecede. Katalogizacijski pravilnik in ISBD(M) od nas zahtevata prepis podatkov z vira, četudi imajo tipkarske napake. To moramo tudi posebej označiti.

Tudi v arhivih je situacija podobna v teoriji, praksa pa se razlikuje. Čeprav standard ISAD(G) (2000) od arhivistk in arhivistov zahteva, da za naslov popisne enote uporabijo naslov, ki se pojavlja v gradivu samem, vključno z vsemi tipkarskimi napakami, pa se v praksi to določilo večinoma ne uporablja. Naslov se oblikuje brez napak, izvirnik z napakami pa se vpiše kot drugi naslov. Pri tem pa je potrebno poudariti, da velika večina arhivskega gradiva nima izvirnega naslova, zato ga arhivisti večinoma oblikujejo po svoji presoji. (O tem je pisal Boštjan Zajšek (2012) in predlagal zanimive smernice za poenotenje.) Vse tovrstne napake pa imajo seveda velike posledice pri poizvedovanju po arhivskem gradivu, ki je prosto dostopno preko spleta. Normativne datoteke po-

8 Dostopno na spletni strani <http://akm.hkdrustvo.hr/> (29-03-2019).

magajo pri razreševanju in pojasnjevanju napak, kakršna je napačno zapisano osebno, geografsko ali drugačno ime. Nedvomno tudi možnost samodokončanja (*autocomplete*) v modulih za katalogizacijo pomaga pri pravilnem pisanju. Če gre za nenavadno zapisano besedo, nas utegne ta možnost tudi ovirati. S tako izkušnjo se včasih srečujemo tudi v zasebnem življenju na osebnih računalnikih, tablicah in telefonih.

Vir

Ker so vzajemni katalogi namenjeni izmenjavi kataložnih zapisov, ni zanemarljivo, kdo zapise ustvarja. Med slovenskimi katalogizatorji in katalogizatorkami imajo velik ugled zapisi, ki jih kreirajo v Narodni in univerzitetni knjižnici v Ljubljani. Ti so navadno tudi prvi, ker pri tej knjižnici deluje center za ISBN. Ob dodelitvi te številke izdelajo tudi CIP. Včasih druge knjižnice knjigo obdelajo prej in zapis dopolnijo, drugič s katalogizacijo čakajo, da jim zapisa ne bi bilo treba dopolnjevati.

V slovenskem vzajemnem katalogu se zapisi za tuje knjige prevzemajo iz OCLC. Pogosto taki zapisi vsebujejo tudi predmetne oznake v tujem jeziku. V nekaterih knjižnicah predmetne oznake v tujem jeziku brišejo in jih na novo oblikujejo v slovenščini. Druge knjižnice predmetnim oznakam v tujem jeziku dodajo oznake v slovenskem jeziku. Tretje pustijo le predmetne oznake v tujem jeziku. Čeprav je vir podatkov vreden vsega zaupanja, je vprašanje, če so take predmetne oznake uporabne. Prvič, je take bibliografske vire mogoče najti le s poizvedbami v različnih jezikih. Drugič, ne vemo, če se se javnost zaveda, da je treba iskati v več jezikih. Tretjič, je vprašanje, kako naj ljudje vedo, v katerih jezikih je treba iskati. Glede na običajno neučakanost pri iskanju, je malo verjetno, da bi isto poizvedbo ponavljali v različnih jezikih in da bi uporabnicam in uporabnikom to sploh prišlo na misel.

Skladnost s pričakovanji

V knjižničnih katalogih in arhivskih podatkovnih zbirkah pričakujemo, da bodo zapisi skladni s pravilniki in priročniki, ki veljajo v naših sistemih. Za slovenske katalogizatorje in katalogizatorke je bilo ugotovljeno, da redkeje uporabljajo pravilnik in priročnike ter da manj sledijo katalogizacijskim standardom in pravilnikom kot priročniku za format COMARC/B (Seljak, 2006 in 2000; Likar, 2003; Likar in Žumer, 2004; Krstulović, 2006; Pesjak in Petek (2010); Crnčić,

2010; Enci, 2011). Ne moremo pričakovati, da bo novi standard ali pravilnik sam po sebi rešil težave pri katalogizaciji in težavne katalogizacijske primere. Enako bi veljalo za strokovno osebje arhivov. Naše strokovno delo namreč vključujejo uporabo pravilnikov in standardov. Od taksista verjetno pričakujemo, da nas bo vozil varno in se, kot zahtevajo predpisi, ustavil pred rdečim semaforjem. Tako tudi od nas, katalogizatorjev in katalogizatorjk ter arhivistk in arhivistov, uporabnice in uporabniki pričakujejo, da bomo strokovno opravili svoje delo. Tako pričakovanje je zapisano tudi v etičnih kodeksih tako knjižničarjev⁹ kot arhivistov.¹⁰ V sodobni družbi je opaziti veliko prezira do predpisov in pravil. V takih razmerah je študentom in študentkam težko dopovedati, da imamo dostop do bibliografskih virov s skoraj vsega sveta (npr. Worldcat¹¹) prav zato, ker knjižnice že štirideset let sledijo skupnemu, mednarodnemu standardu.

Doslednost

Pomen sledenja enakim načelom in pravilom za opis gradiva skozi čas pona-zarjajo prav knjižnični katalogi. Ko je v začetku uvedbe računalniške obdelave knjižničnega gradiva prišlo do retrospektivne konverzije knjižničnih katalogov (med drugimi o tem piše tudi Žumer, 1998), se je po pričakovanju pokazalo, da obseg podatkov iz časa pred uvedbo ISBD ne zadostuje za sodoben računalniški katalog. Zato nekatere knjižnice svojega najstarejšega gradiva niso sistematično vključile v lokalne kataloge v sistemu COBISS. Npr. Narodna in univerzitetna knjižnica ima gradivo za obdobje 1774-1947 v posebni zbirki¹² in le nekateri bibliografski viri iz te zbirke so bili naknadno popisani tudi v sodobnem vzajemnem katalogu COBISS.

O doslednosti vsebinskega opisa je bilo v drugi polovici 20. stoletja veliko napisanega. Prva sta to temo načela Zunde in Dexter (1969). Mnogi so njuna spoznanja na različne načine potrjevali (npr. Sievert in Andrews (1991); Tonta, 1991; Svenonius in McGarry, 1993). Med zadnjimi so o tem pisali Olson in Wolfram (2006) ter Mai (2001 in 2005). Ugotovljeno je bilo, da isto delo različni katalogizatorji redko označujejo z enakimi vsebinskimi oznakami. Bolj se ujemajo v splošnih temah kot v podrobnostih. Prav tako bi posameznik, ki pripravlja vse-

9 Etični kodeks Hrvatskog knjižničarskog društva je dostopen na spletni strani https://www.hkdrustvo.hr/hr/etic-ki_kodeks/, Zveze bibliotekarskih društev Slovenije pa na spletni strani <http://www.zbds-zveza.si/?q=node3/20> (29-03-2019).

10 Kodeks etike, <http://www.arhiv.gov.si/fileadmin/arhiv.gov.si/pageuploads/zakonodaja/kodeks.pdf> (29-03-2019).

11 Worldcat, <https://www.worldcat.org/> (29-03-2019).

12 Katalog NUK 1774-1947, <http://old.nuk.uni-lj.si/moka/> (29-03-2019).

binski opis danes verjetno izbral nekoliko drugačne oznake kot čez čas. Popolno soglasje v tem smislu ni mogoče, kar je Mai (2000) utemeljil tudi v svoji disertaciji. Vendar položaj ni brezupen. Sezname vrednosti za vpis v posamezno polje (šifranti) in druge oblike normativne kontrole pripomorejo k večji doslednosti podatkov. Katalogizatorjem in katalogizatorkam so, kot kaže, v pomoč kataložni zapisi za podobno gradivo, ki je že v knjižnici. Nekateri opazovani katalogizatorji in katalogizatorke (Šauperl, 2002), so se namreč vedle, kakor da bi gradivo postavljale na isto polico (*collocation*).

Pomen sledenja enakim načelom in pravilom za popisovanje gradiva je tudi v arhivski stroki postal prezenten v času prehoda na popisovanje gradiva znotraj arhivske podatkovne zbirke. Gre za področje, ki je trenutno v slovenskem prostoru še dokaj neurejeno, kar je mogoče pripisati dejstvu, da slovenska arhivska stroka, v tem trenutku še nima izdelanih svojih lastnih pravil popisovanja arhivskega gradiva, kar pomeni, da strokovno osebje pravila ustvarja sprost in nekontrolirano. To seveda vodi v nekonsistentnosti arhivske podatkovne zbirke. Enako velja za oblikovanje normativnih zapisov korporativnih teles, fizičnih oseb in družin ter zemljepisnih imen in stvarnih gesel, kljub dejstvu, da je Semlič Rajh (2016) izdelala podrobne smernice za gesljenje vsebin arhivske vrednosti.

Pravočasnost in zaostanek

Pravočasnost ali sodobnost

Veliki angleško-slovenski slovar izraz »timeliness« prevaja v pravočasnost. Na pravočasnost morda najprej pomislimo v zvezi s tem, da uporabnica oziroma uporabnik dobi informacijo pravi čas. Npr. da študent ali študentka v knjižnico prideta pravočasno, da imata čas za iskanje in študij virov, ne pa da jih začneta iskati eno uro preden je treba nalogo oddati. Vendar Bruce in Hillmann (2004) mislita na sodobnost, saj jo ponazarjata z zastarelim URL-jem v zapisu, ki ga nismo pravočasno posodobili. V to kategorijo bi sodilo tudi pravočasno dopolnjevanje geslovníkov, klasifikacijskih sistemov, šifrantov in normativnih datotek, to je njihovo posodabljanje. Nekateri sistemi se posodablajo enkrat letno, npr. Univerzalna decimalna klasifikacija. Ko se spremeni ime države (npr. Severna Makedonija) ali korporativnega telesa (npr. ministrstva po volitvah) pričakujemo, da bo ta podatek takoj na voljo in uporaben. Kar najhitreje ga je treba dodati v normativno datoteko imen in predmetnih oznak.

Čeprav bo prejšnje ime ministrstva zastarelo, v smislu, da ne bo odražalo sodobnega imena korporativnega telesa, bo še vedno veljavno. Določeno število bibliografskih virov bo namreč za vedno označeno s tem imenom, in to pravilno. Ni pa dvoma, da je vrstilec za hrvaško književnost 886.2 po letu 1995 zastarel, veljaven pa je 821.163.4. Zastareli vrstilec je še vedno mogoče najti v slovenskem vzajemnem katalogu, čeprav bi ga verjetno bilo mogoče posodobiti z avtomatskim postopkom. Pri starih virih je izziv tudi starinski jezik, izrazi, ki jih danes ne uporabljamo več. To pride do izraza zlasti takrat, ko postanejo celotna besedila dostopna v digitalnih zbirkah.

Dostopnost

Dostopnost je mogoče razvrstiti v dve večji skupini (Bruce in Hillmann, 2004). Mehanska dostopnost oz. nedostopnost izvira iz strojne neberljivosti, ki je lahko rezultat nepravilne pretvorbe črk, nepravilnega prenosa polj, nekompatibilne računalniške in programske opreme in podobno. Omenili smo že (poglavje 3.5 o doslednosti), da pri retrospektivnih konverzijah zaradi prepogostih napak ni bilo mogoče izvesti prenosa vseh zapisov s kataložnih listkov v računalniške sisteme. Medtem ko človek med pomešanimi podatki razmeroma enostavno prepozna avtorja in naslov, za računalniške sisteme to ni vedno enostavno.

Intelektualna nedostopnost pa izvira iz našega nepoznavanja pisav in jezikov, tematike v bibliografskih virih, nerazumevanja klasifikacijskih in predmetnih oznak in podobno. Podatki, ki jih ne znamo prebrati ali dojeti, nam ostanejo nedostopni. Računalniški sistemi nam tu ne morejo čisto nič pomagati. Samokritičnost in učenje sta edino zdravilo proti tej vrsti napak.

Vse omenjeno lahko rečemo, da velja tudi za arhive, vendar je tukaj potrebno omeniti še nekaj, česar se uporabnice in uporabniki velikokrat ne zavedajo. Neprisotnost določenih podatkov v arhivskih podatkovnih zbirkah ne pomeni nujno, da arhiv s temi podatki ne razpolaga. Arhivi so namreč pri svojem gradivu vezani na varovanje osebnih in občutljivih podatkov, ki jih predvidevajo *Zakon o varstvu dokumentarnega in arhivskega gradiva ter arhivih* (2006, 2014), *Uredba o varstvu dokumentarnega in arhivskega gradiva* (2017) ter *Splošne uredbe Evropske unije o varstvu osebnih podatkov* (2017).

Katalog ni na pol poln napak, temveč na pol prazen

Bibliografski vir ali arhivsko gradivo je kljub napaki v kataložnem zapisu ali arhivski podatkovni zbirki mogoče najti, kadar slučajno ustreza poizvedbi. Malo verjetno pa je, da takrat ustreza uporabničini oziroma uporabnikovi potrebi. Visoka kakovost podatkov v knjižničnih katalogih in arhivskih podatkovnih zbirkah seveda prinaša točen opis bibliografskih virov in arhivskega gradiva. Dosega pa tudi drugi cilj, ki je za obstoj knjižnic in arhivov morda celo pomembnejši, namreč zaupanje javnosti (Bruce in Hillmann, 2004).

David Bade (2002 in 2006-7) je velik del svojega raziskovanja posvetil prav napakam v knjižničnih katalogih. V eseju iz leta 2002 (Bade, 2002, str. 1) je duhovito imenoval dve novi »parazitski vrsti«, in sicer »*Incompetentus linguisticus*« in »*Subjectus incorrectus*«. V eseju veliko prostora nameni razlagi njune narave in delovanja. Za prvo je značilno domišljanje, da dovolj poznamo jezik bibliografskega vira, da zmoremo pravilno prepisati podatke v primerne elemente kataložnega zapisa. Potem pa se ne izkaže samo to, da besedila nismo pravilno transliterirali, temveč tudi to, da smo imeli podatek o založbi za naslov, ali da smo celo nepravilno izbrali primarni vir podatkov. Za drugo vrsto, »*Subjectus incorrectus*«, pa je značilen nepravilen vsebinski opis, ki izhaja iz katalogizatorjeve oziroma katalogizatorjine površnosti ali neznanja. Kolegom katalogizatorjem in kolegicam kazalogizatorkam Bade (2002) razloži, da sta edina preventiva proti okužbi s tema dvema tegobama nenehna samokritičnost in stalno učenje. Takrat se Bade še ni ukvarjal z lažnimi informacijami, s katerimi smo oblegani danes. Samokritičnost in stalno učenje pa sta edino zdravilo tudi proti novemu »parazitu«, lažni informaciji (npr. *Falsus factus*).

Ariely (2010), eden od pionirjev vedenjske ekonomije, piše, da smo ljudje po naravi nagnjeni k delovanju z najmanjšim naporom in bomo za tak cilj pripravljene goljufati tudi če smo iskreni in pošteni ljudje. Napake v knjižničnih katalogih zelo verjetno izhajajo iz tega našega nagnjenja. Če je tako, potem je ena od možnih rešitev tudi ta, ki jo omenja Ariely (2010): Kakovost in poštenje je treba kot mantro, molitev ali etično prisego redno ponavljati. Raziskave, ki jih Ariely (2010) predstavi, so namreč pokazale, da npr. ponavljanje desetih zapovedi ugodno učinkuje na zmanjševanje obsega (namernih ali nenamernih) goljufij.

Sestavek torej obzirno ponujava v branje z željo, da bi ozavestil bralce in bralke ter pripomogel k dvigu kakovosti rezultatov njihovega dela. Le s kakovostnimi podatki bo namreč v dobro uporabnic in uporabnikov uspešno izkoriščen tudi neizmerni trud ljudi, ki sodelujejo, so sodelovali v preteklosti, ali še le bodo

sodelovali pri pripravi in vzdrževanju katalogizacijskih pravilnikov in sistemov za knjižnične kataloge ter pravilnikov za popisovanje arhivskega gradiva in arhivskih podatkovnih zbirk. Pri tem je posebnega občudovanja vreden projekt skupnega hrvaškega katalogizacijskega pravilnika.

NAVEDENI VIRI

- ARIELY, D. (2010). *Predvidljivo nerazumni*. Ljubljana: Mladinska knjiga.
- BADE, D. (2002). *The creation and persistence of misinformation in shared library catalogs: Language and subject knowledge in a technological era*. Urbana-Champaign, IL: Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign, <https://www.ideals.illinois.edu/bitstream/handle/2142/3874/gslisoccasionalpv00000i00211.pdf> (29-03-2019).
- BADE, D. (2006-7). *Responsible librarianship: Library policies for unreliable systems*. Duluth, Mn.: Library Juice Press, <https://books.google.si/books?id=rb84qjyEt3YC&printsec=frontcover&dq=responsible+librarianship&hl=sl&sa=X&ved=0ahUKewiOrtapx53gAhUDM-wKHYayCOcQ6AEIKDAA#v=onepage&q=responsible%20librarianship&f=false> (29-03-2019).
- BADOVINAC, B. (2017). Izhodišča za proučevanje kakovosti podatkov v bibliografskih in normativnih zapisih: kakovost podatkov v kontekstu in raziskovalne usmeritve v katalogizaciji. *Knjižnica*, 61 (1-2), 119-154, <https://knjiznica.zbds-zveza.si/knjiznica/article/view/6165> (29-03-2019).
- BRUCE, T.R. in HILLMAN, D.I. (2004). The continuum of metadata quality: defining, expressing, exploiting. V D.I. Hillmann in E.L. Westbrook (ur.) *Metadata in practice* (str. 238-256). Chicago: American Library Association.
- COMARC/B format za bibliografske podatke: priročnik za uporabnike. (1991-2018). Maribor: Inštitut informacijskih znanosti, https://izobrazevanje.izum.si/EntryFormDesktopDefault.aspx?tabid=38&type=manual&manual=1_Comarc_B_svn (29-03-2019).
- CRNČIĆ, M. (2010). *Kakovost bibliografskih zapisov v vzajemnem katalogi COBIB: diplomsko delo*. Ljubljana: Filozofska fakulteta.
- DIMEC, Z. (1994). Strokovna usposobljenost za katalogizacijo v COBISSu: analiza problematike in možni ukrepi za izboljšanje. *Knjižnica*, 38 (3/4), 75-92.
- ENCI, T. (2011). *Kakovost bibliografskih zapisov za serijske publikacije v vzajemnem katalogu COBIB.SI: diplomsko delo*. Ljubljana: Filozofska fakulteta.

- HJØRLAND, B. (2016). Subject (of documents). V B. Hjørland in C. Gnoli (ur.) *ISKO Encyclopedia of information science*, <http://www.isko.org/cyclo/subject> (29-03-2019).
- ISAAR(CPF): *International Standard Archival Authority Record for Corporate Bodies, Persons and Families*, <http://www.ica.org/sites/default/files/ISAAR2EN.pdf> (29-03-2019).
- ISAD (G): *General International Standard Archival Description*, http://www.ica.org/sites/default/files/isad_g_2e.pdf (29-03-2019).
- ISBD(M): *Mednarodni standardni bibliografski opis monografskih publikacij* (Pre-delana izd.). (1997). Ljubljana: Narodna in univerzitetna knjižnica.
- Kodeks etike (1997). *Arhivi XX* (str. 4-16). Ljubljana: Arhivsko društvo Slovenije, <http://www.arhiv.gov.si/fileadmin/arhiv.gov.si/pageuploads/zakonodaja/kodeks.pdf> (29-03-2019).
- KRSTULOVIĆ, Z. (2006). Katalogizacijska pravila in kakovost bibliografskih podatkov. *Organizacija znanja*, 11 (4), 215-218.
- LIKAR, T. (2003). Enotna obdelava knjižničnega gradiva. *Knjižnica*, 47 (1-2), 7-34.
- LIKAR, T. in ŽUMER, M. (2004). Mnenja katalogizatorjev o modulu za katalogizacijo v sistemu COBISS. *Knjižnica*, 48 (1-2), 83-122.
- MAI, J.-E. (2000). *The Subject Indexing Process: An Investigation of Problems in Knowledge Representation: Doctoral dissertation*. University of Texas at Austin.
- MAI, J.-E. (2001). Semiotics and indexing: an analysis of the subject indexing process. *Journal of Documentation*, 57 (5), 591-622.
- MAI, J.-E. (2005). Analysis in indexing: document and domain centered approaches. *Information Processing & Management*, 41 (3), 599-611
- MERICA, J. (2011). *Iskanje po vsebini v katalogu Mestne knjižnice Ljubljana: diplomsko delo*. Ljubljana: Filozofska fakulteta.
- NOVAK, M. in SEMLIČ RAJH, Z. (2013). Merjenje rezultatov uspešnosti izgradnje arhivskih podatkovnih zbirk s kvantitativno-kvalitativno metodo na primeru podatkovne zbirke SIRAnet. V J. Volčjak (ur.) *Standardizacija (p)opisov arhivskega gradiva in uskladitev strokovnih praks v slovenskih javnih in cerkvenih arhivih: zbornik referatov* (str. 19-39). Ljubljana: Arhivsko društvo Slovenije.
- OLSON, H.A. in WOLFRAM, D. (2006). Indexing consistency and its implications for information architecture: a pilot study. V *7th Information Architecture Summit*. Silver Spring, MA: American Society for Information Science and Technology.
- PESJAK, D. in PETEK, M. (2010). Kakovost bibliografskih zapisov v COBIB in

- uporaba katalogizacijskih priročnikov. *Knjižnica*, 54 (3), 15-33.
- Podatkovna zbirka SIRAnet. (2019). *Statistika popisnih enot*.
- RAZPOTNIK, Š. (2007). *Zajemanje uporabniškega besedišča kot osnova za razvoj geslovnikov: magistrsko delo*. Ljubljana: Filozofska fakulteta.
- SELJAK, M. (2000). Izvajanje določil pravilnika o izdaji dovoljenja za vzajemno katalogizacijo v sistemu COBISS.SI. *Organizacija znanja*, 11 (4), 208-214.
- SELJAK, M. (2006). Poti do konsistentnih katalogizacijskih pravil. *Organizacija znanja*, 5 (4), 29-49.
- SEMLIČ RAJH, Z. (2016). *Smernice za gesljenje, klasificiranje in izdelavo tezavrov pri vsebinah arhivske vrednosti: doktorska disertacija*. Ljubljana: Filozofska fakulteta.
- SEMLIČ RAJH, Z. in ŠAUPERL, A. (2013). Analiza oblikovanja vsebine zajetih podatkov v podatkovni bazi SIRAnet. V I. Fras in N. Gostenčnik (ur.). *Tehnični in vsebinski problemi klasičnega in elektronskega arhiviranja* (str. 145-157). Maribor: Pokrajinski arhiv, http://www.pokarh-mb.si/uploaded/datoteke/Radenci/Radenci2013/12_Semlic_Sauperl_2013.pdf (29-03-2019).
- SIEVERT, M.C. in ANDREWS, M.J. (1991). Indexing consistency in Information Science Abstracts. *Journal of the American Society for Information Science*, 42 (1), 1-6.
- SVENONIUS, E. in MCGARRY, D. (1993). Objectivity in evaluating subject heading assignment. *Cataloging & Classification Quarterly*, 16 (2), 5-40.
- ŠAUPERL, A. (2002). *Subject determination during the cataloging process*. Lanham (Ma.): Scarecrow.
- ŠAUPERL, A. (2016). Kratek opis leposlovja. *Knjižnica*, 60 (2-3), 101-126, <https://www.dlib.si/stream/URN:NBN:SI:DOC-84S7GN7W/4a2f35ec-cd5f-4182-b-298-ce5e446a2537/PDF> (29-03-2019).
- TONTA, Y. (1991). A study of indexing consistency between Library of Congress and British Library catalogers. *Library Resources and Technical Services*, 35 (2), 177-185.
- Uredba (EU) 2016/679 Evropskega parlamenta in Sveta z dne 27. aprila 2016 o varstvu posameznikov pri obdelavi osebnih podatkov in o prostem pretoku takih podatkov ter o razveljavitvi Direktive 95/46/ES. *Uradni list Evropske unije*, L 199, zvezek 59 z dne 4. maj 2016.
- Uredba o varstvu dokumentarnega in arhivskega gradiva. (2017). *Uradni list Republike Slovenije*, 42/17.
- WILLER, M., ŠAUPERL, A., PETEK, M. in TOMIĆ, M. (2011). Jedinostveni stvarni naslov: zašto nam je potreban više nego ikad? *Vjesnik bibliotekara Hr-*

- vatske, 54 (1/2), 93-119, [http://www.hkdrustvo.hr/datoteke/1148/vbh/God.54\(2011\),br.1/2](http://www.hkdrustvo.hr/datoteke/1148/vbh/God.54(2011),br.1/2) (29-03-2019).
- ZAJŠEK, B. (2012). Oblikovanje naslovov popisnih enot glede na mednarodne arhivske standarde, nove informacijske sisteme slovenskih javnih arhivov ter njihove uporabnike. V N. Gostenčnik (ur.) *Tehnični in vsebinski problemi klasičnega in elektronskega arhiviranja* (str. 581-604). Maribor: Pokrajinski arhiv, http://www.pokarh-mb.si/uploaded/datoteke/Radenci/Radenci2012/61_Zaj__ek_2012.pdf (29-03-2019).
- Zakon o varstvu dokumentarnega in arhivskega gradiva ter arhivih. (2006 in 2014). *Uradni list Republike Slovenije*, 30/2006 in 51/2014.
- ŽAUCER, M. (2007). Nujnost racionalizacije postopkov in dela v knjižnicah. *Knjižnica*, 51 (3-4), 195-210, <https://knjiznica.zbds-zveza.si/knjiznica/article/view/5513> (29-03-2019).
- ŽUMER, M. (1998). Retrospektivna konverzija starih katalogov Narodne in univerzitetne knjižnice. V J. Urbanija (ur.) *Zbornik razprav: 10 let Oddelka za bibliotekarstvo* (str. 215-229). Ljubljana: Filozofska fakulteta.

DATA QUALITY IN LIBRARY CATALOGUES AND ARCHIVAL DATA BASES

KEYWORDS:

library catalogues, archival databases, data quality, errors

ABSTRACT

Purpose: *Dealing with metadata quality in library catalogues and archival databases is a task we tend to avoid. However we have all found that data quality is not what we expected when we turned from card catalogues to computer databases. We realized that only good quality of data provides appropriate recall of resources and information.*

Methodology: *Our literature review is structured in seven aspects of metadata quality: completeness, accuracy, provenance, conformance to expectations, logical consistency and coherence, timeliness, and accessibility.*

Results: *We believe that it is urgent to take a more active approach to the problem of data quality. We hope that this review might persuade readers to make their contribution to data quality just as we all contribute in our small way to waste management in our cities.*

Conclusions: *Bibliographic and archival resources may be accessible even if their metadata contains some errors. Unfortunately, it is unlikely that it would suit the user's needs in such cases. High quality of data means accurate description in library and archival databases. In addition, it brings us closer to yet another very important goal, which is public trust.*