

OBLIKOVANJE BIBLIOGRAFSKE PODATKOVNE ZBIRKE ZA GRADIVO IZ OBDOBJA ROČNEGA TISKA: PROJEKT CERL-a*

Zlata Dimec

Izvirni znanstveni članek

UDK 025.3:094.1(4)

Izvleček

Knjige in drugo tiskano gradivo iz obdobja ročnega tiska v Evropi (do ca. 1830) predstavljajo specializirano področje raziskovanja za različne znanstvene discipline. Konzorcij evropskih znanstvenih knjižnic (Consortium of European Research Libraries - CERL) je bil ustanovljen l. 1992 z namenom oblikovati skupno bibliografsko podatkovno zbirko za to gradivo. Cilja projekta sta predvsem dva: raziskovalcem ponuditi dostop do podatkov za to gradivo na enem mestu in s podatki o lokaciji posameznih izvodov (funkcija centralnega kataloga), knjižnicam pa omogočiti prevzemanje zapisov v svoje lokalne kataloge in s tem zmanjšati stroške obdelave. Za oblikovanje skupne bibliografske podatkovne zbirke z imenom Hand Press Book (HPB) so potrebni številni dogovori v zvezi z obdelavo, formatom zapisov, uporabo normativnih seznamov itd. Potrebno je najti način, ki na eni strani omogoči sožitje različnih katalogizacijskih pravil, obenem pa še vedno zagotavlja uporabnost bibliografskih zapisov. Zato so določeni minimalni kriteriji za obseg zapisov in evidentirane posebnosti posameznih prispevanih datotek. Uporabnik ima poleg zapisov samih na voljo tudi osnovne informacije o značilnostih gradiva iz zbirk posameznih knjižnic in njihovih zapisov ter dodatna navodila za učinkovitejše poizvedovanje. Poleg tega CERL sistematično zbira podatke o drugih projektih, ki obravnavajo tovrstno in drugo starejše gradivo (vključno z rokopisi), njegovo varovanje, pa tudi razvoj knjige in založništva oz. tiskarstva. Povezovanje in sodelovanje evropskih knjižnic je tudi ena od glavnih prednosti tega projekta, v katerem od samega začetka sodeluje tudi Narodna in univerzitetna knjižnica v Ljubljani.

Ključne besede: bibliografske podatkovne zbirke, stari tiski, Evropa, Consortium of European Research Libraries

* *Prispevek je bil objavljen v publikaciji »Tradicionalni mediji v sodobni informacijski družbi«. Zveza bibliotekarskih društev Slovenije, strokovno posvetovanje, Maribor, 11. do 13. oktober 2000 (str. 126-137). Ljubljana : Zveza bibliotekarskih društev Slovenije, 2000.*

DIMEC, Zlata: Establishing a bibliographic database for hand press books : CERL project. Knjižnica, Ljubljana, 46(2002)3, 105-117

Abstract

Books and other printed materials dating to the Hand Press period (ca. 1830) in Europe represent a special field of research. Consortium of European Research Libraries – CERL was founded in 1992 with an objective to establish a machine-readable database for this material. The goals of the project are twofold: to offer one site access to information about the Hand Press books, including holdings data (providing thus the function of a union catalogue), and enable downloading of records into local catalogues, reducing processing costs. To organize a union database named Hand Press Book (HPB), detailed arrangements are necessary to harmonize different cataloguing practices and machine-readable formats, establish use of thesaurus files etc., and make the records as useful as possible. Minimal level requirements have been defined and the characteristics of contributed files are described. The user is also provided with information on characteristics of the material from respective library collections, as well as with detailed instructions for successful searching. Besides, CERL systematically collects information on other related projects and institutions, covering the Hand Press period, manuscripts, together with the history of the book, printing and publishing. Cooperation and project coordination of European libraries is one of the main benefits of this project in which the Slovenian National and University Library participates from the very beginning.

Key words: bibliographic databases, hand press books, Europe, Consortium of European Research Libraries

1 Uvod

Knjige in drugo tiskano gradivo iz obdobja ročnega tiska v Evropi (do ca. 1830) predstavljajo specializirano področje raziskovanja za različne znanstvene discipline. Podatki o tem gradivu so seveda že vključeni v klasične ali računalniške kataloge posameznih knjižnic, pa tudi v večje nacionalne in mednarodne bibliografske podatkovne zbirke oz. vzajemne kataloge. Vendar so ti podatki razpršeni, shranjeni skupaj s podatki o drugem gradivu in za poizvedovanje različno dostopni, odvisno od določenega sistema. Določeno število tega gradiva pa v knjižnicah celo sploh še ni obdelano. Na primeru bibliografske podatkovne zbirke za tovrstno gradivo z imenom HPB (Hand Press Book), ki jo gradi Konzorcij evropskih znanstvenih knjižnic (Consortium of European Research Libraries - CERL), so v prispevku prikazani cilji tega projekta in načela, ki so pri tem upoštevana. Dodano je razmišljanje o prednostih in pomanjkljivostih izbranega načina za doseg tega cilja, posebej tudi s stališča stanja v Sloveniji.

2 Ozadje

Novembra 1990 sta Britanska knjižnica in Bavarska državna knjižnica v Münchnu organizirali mednarodno konferenco na temo retrospektivne katalogizacije oz. retrospektivne konverzije starejšega tiskanega gradiva (15.-19. stoletje) v evropskih knjižnicah. Kot omenja F. G. Kaltwasser v predgovoru h konferenčnemu zborniku (1, str. 9), je bil glavni povod za to konferenco dejstvo, da so v tem času v Evropi potekali ali bili načrtovani številni projekti retrospektivne konverzije ali katalogizacije, vendar med seboj niso bili usklajeni, zato knjižnice niso poznale izkušenj in rezultatov drugih in jih tudi niso mogle uporabiti. Podana je bila ocena, da je potrebno s stališča bodoče združene Evrope to stanje preseči in poiskati način, kako omogočiti čimbolj enoten dostop do te kulturne dediščine.

Na konferenci so bila predstavljena poročila o stanju in projektih v posameznih državah. Ob tem so se jasno pokazale razlike v obsegu oz. natančnosti zapisov, predvsem pa razlike v oblikah imen, ki zaradi različnih katalogizacijskih pravil otežujejo izmenjavo podatkov (2, str. 27-29). Zaključki konference, ki so vzpodbudili nadaljnje aktivnosti, so se glasili (1, str. 190):

- pripravi naj se projekt za sodelovanje večjih evropskih znanstvenih knjižnic, ki naj bi podatke o svojem gradivu ponudile v širšo uporabo prek medsebojno usklajene in s tem povečane dostopnosti svojih bibliografskih podatkovnih zbirk;
- nacionalne bibliografske ustanove nosijo odgovornost za normativne podatkovne zbirke imen s svojega območja;
- raziskati možnosti za povezovanje med različnimi sistemi in podatkovnimi zbirkami.

Ti zaključki so torej pripeljali do ustanovitve CERL-a leta 1992. Cilja vzpostavitve posebne bibliografske podatkovne zbirke za gradivo iz obdobja ročnega tiska (Hand Press Book – HPB) sta predvsem dva: raziskovalcem in znanstvenikom ponuditi kar najustreznejši dostop do podatkov za to gradivo na enem mestu in s podatki o lokaciji posameznih izvodov (funkcija centralnega kataloga), knjižnicam pa obenem omogočiti prevzemanje zapisov v svoje lokalne kataloge in s tem zmanjšati stroške obdelave.

3 Značilnosti gradiva

Gradivo iz tega obdobja ima nekaj posebnosti, po katerih se razlikuje od ostalega knjižničnega gradiva. Določeno raznolikost mu sicer daje že časovni razpon nastanka, ki obsega štiri stoletja: od prve tiskane knjige sredi 15. stoletja (vk-

ljuje torej prvotiske), do ca. 1830, ko se predvsem v Zahodni Evropi zaključijo obdobje ročnega tiska. Na ozemlju današnje Slovenije pa lahko obdobje zaključimo nekoliko kasneje, okrog leta 1850. To obdobje v primerjavi s sodobnimi publikacijami zaznamuje pomembnejša vloga tiskarjev, za knjige pa so se postopoma ustalili značilni elementi, ki nam služijo pri obdelavi. Tako je npr. predvsem za inkunabule značilno, da nimajo naslovne strani, kot jo poznamo danes. Ko pa se naslovna stran uveljavi, na njej najdemo v primerjavi z današnjimi izjemno dolge naslove, pri katerih je velikokrat težko določiti, kje se določen element bibliografskega opisa začne oz. konča.

Za bibliografski opis tega gradiva je bil ravno zaradi teh značilnosti pripravljen poseben standard: ISBD(A) (3). Nekatere posebnosti standarda v primerjavi z drugimi ISBD-ji so:

- doslednejše upoštevanje podatkov v obliki in zaporedju z naslovne strani;
- pogosta navedba tiskarja na mestu založnika;
- natančnejša navedba fizičnega obsega in navedba naravnega formata (tj. števila pregibov tiskarske pole);
- obvezna navedba bibliografskega citata pri inkunabulah;
- praviloma številnejše in natančnejše opombe, predvsem opombe k posameznemu izvodu (privezki, posebna vezava, provenienca, rokopisni dodatki itd.);
- tudi prstni odtis kot nadomestilo standardne številke je zaradi pomanjkljivih ali poškodovanih izvodov velikokrat vezan na posamezen izvod in ne na celotno izdajo.

Za ilustracijo primer, kako je na prvi pogled težko z običajnimi bibliografskimi elementi ISBD identificirati dve nesporno različni izdaji:

OSEM inu shestdeset sveteh pesm, katire so na proshnje, inu poshelenje vezh brumneh dush ... na svitlobu dane : Sraven dveh normalskeh per pejtj, inu per gmej Mashi navadneh Peisem. – Na novizh vun-dane. – V' Lublani : se najdejo per Marii Anni Raab, 1800. – 245, [11] str. ; 8š (16 cm)

Sestavljaec antologije Maksimiljan Redeski ugotovljen po: F. Simonič: Slovenska bibliografija, Ljubljana 1903-1905. – Boh. – 1. izd. 1775. – Rokopisni dodatek 13 svetih pesmi na 48 uvezanih str. zapisal Matija Naglič: NUK sign. 222182

OSEM, inu shestdeset sveteh pesm, katire so na proshnje, inu poshelenje vezh brumneh dusch ... na svitlobu dane. – Na novizh vun-dane. – V' Lublani : se najdejo per Marii Anni Raab, 1800. – 241, [7] str. ; 8š (17 cm)

Sestavljaec antologije Maksimiljan Redeski ugotovljen po: F. Simonič: Slovenska bibliografija, Ljubljana 1903-1905. – Boh. – 1. izd. 1775

Kot vidimo, je potreben zelo dosleden in podroben opis, da so razvidne razlike med obema izdajama, ki sta izšli v istem letu. Najočitnejša razlika je podnaslov pri prvem opisu, ki bi ga bil katalogizator zaradi že tako dolgega naslova lahko celo izpustil, pa različno napisani besedi v naslovu (dush : dusch), ki bi ju tudi

lahko pripisali pomoti pri prepisovanju; razlika pri fizičnem opisu bi lahko nastala zaradi dodatnih listov, uvezanih že po izidu. Če pa vidimo podatke z naslovnih strani, nam je na prvi pogled jasno, da gre za dve različni izdaji: uporabljene so drugačne črke, razlikujeta se razpored besed ter raba velikih in malih črk. Zato bi bila za identifikacijo zelo koristna povezava zapisov z naslovno stranjo (za uporabnika pa seveda tudi s celotnim besedilom):



Bibliografski opis tega gradiva je torej lahko precej dolg in zapleten, s številnimi opombami, ki zahtevajo dobro poznavanje gradiva in iskanje podatkov o njem. Ravno npr. podatki o posameznem izvodu so lahko izjemno zanimiv vir za znanstvenike, ki se ukvarjajo z zgodovino knjige in tiskarstva ali raziskujejo vpliv pisane besede v določenem zgodovinskem okolju. Vemo pa, da struktura formatov za bibliografske podatke pri opombah ni zelo natančna in opombe tudi niso vedno indeksirane, zato je iskanje teh podatkov oteženo.

Posebna bibliografska podatkovna zbirka za to gradivo naj bi zato uporabnikom zagotavljala kar najustreznejše možnosti za pozvedovanje, nudila naj bi pregled nad izvodi posamezne izdaje in omogočila natančno lociranje zaželenega izvoda (ime knjižnice in signatura). Knjižnicam pa bi prek prevzemanja zapisov omogočila prihranek pri zahtevni obdelavi tega gradiva.

4 CERL (Konzorcij evropskih znanstvenih knjižnic)

4.1 Organizacija CERL-a

Zagon v delovanju Konzorcija je predvsem od ustanovnih članic, pa tudi od kasneje pridruženih, zahteval velik prispevek, tako glede sodelovanja pri podrobnejši vsebinski zasnovi projekta kot v obliki članarine. Konzorcij ima zdaj ca. 60 članic, od tega 35 polnopravnih, ostale so pridružene ali imajo poseben status, ker ne gre vedno za knjižnice, ampak ustanove, ki so drugače povezane z osnovnim ciljem Konzorcija (npr. inštituti za zgodovino knjige). Članica je navadno posamezna knjižnica, možno pa je tudi skupinsko članstvo, za kar so se doslej odločile nekatere italijanske in francoske knjižnice. Le polnopravne članice lahko sodelujejo pri upravljanju Konzorcija in uporabljajo podatkovne zbirke (poleg HPB jim je na voljo še nekaj drugih).

Konzorcij ima sedež v Londonu in zato deluje po regulativi britanskega prava. Trenutno (julij 2000) je redno zaposlena le izvršna direktorica, ki se ji bo kmalu pridružil strokovnjak za področje informacijske tehnologije, za določene naloge pa se sklene pogodba s konzultantom. Oblikovana so bila tudi pravila za imenovanje predsednika, članov izvršnega odbora in ekspertne skupine, v katerih sodelujejo članice. Enkrat letno je organizirano srečanje vseh članic konzorcija, na katerem se sprejmejo smernice za nadaljnje delovanje. Ob tej priložnosti je organiziran tudi strokovni seminar, ki naj bi pritegnil zanimanje znanstvenikov za podatkovno zbirko HPB in vzpodbudil njeno uporabo.

Podatkovna zbirka HPB je dostopna na sistemu RLIN, ki ga vzdržuje ameriški RLG (Research Libraries Group). Odločitev, da je izrazito evropski projekt, torej podatkovna zbirka evropskih knjižnic, dostopna prek ameriškega ponudnika, ni naletela na splošno odobravanje vseh članic in je imela tudi nekaj dodatnih posledic. Vendar je prevladala pragmatična ugotovitev, da so pokazali vso pripravljenost za prilaganje zahtevam Konzorcija, in da imajo nekatere članice Konzorcija z njimi dolgoletne in pozitivne izkušnje pri sodelovanju. Konzorcij je v pogajanjih z njimi tudi dosegel, da članice poleg podatkovne zbirke HPB lahko v okviru sistema RLIN uporabljajo tudi njihovo osnovno bibliografsko zbirko.

4.2 Načela delovanja

Vzpostavitev evropske podatkovne zbirke za gradivo iz obdobja ročnega tiska je zahtevala nekaj načelnih odločitev in pripravljanih postopkov. Glede na evropsko orientiranost projekta in zahteve DG XIII Evropske komisije, ki je financira-

la preliminarno študijo (4), je bila logična odločitev za format UNIMARC. V tej študiji je bil analiziran vzorec zapisov, ki jih je prispevalo šest članic: nekateri zapisi so bili preneseni v format UNIMARC iz izvornih formatov, nekateri pa so bili prvotno izdelani v formatu UNIMARC. Zapisi so se razlikovali tudi po tem, da so nekateri nastali z retrospektivno konverzijo klasičnih katalogov, drugi pa s katalogizacijo de visu. Na podlagi analize tega dovolj raznolikega vzorca in osnovnih ciljev projekta so nastala osnovne smernice, ki so jih članice sprejele kot vodilo pri delovanju. Po določenih popravkih teh prvotnih usmeritev so bila za delovanje sprejeta naslednja načela:

1. Podatkovna zbirka naj bo kar najširše dostopen informacijski vir za znanstvenike, ki jih to gradivo zanima.
2. Dostop do podatkovne zbirke naj bo možen prek telekomunikacijskih ali drugih distribucijskih sistemov.
3. Podatkovna zbirka naj vsebuje podatke o lokaciji in deluje kot centralni katalog.
4. Možno bo prevzemanje zapisov v kataloge posameznih knjižnic.
5. Zapisi za zbirko naj bodo po možnosti posredovani v formatu UNIMARC, ne glede na izvorni format pa bodo v formatu UNIMARC dostopni za prevzemanje.
6. Zaželeni standard pri obliki zapisov je sicer ISBD(A), vendar je možno prispevati tudi po drugačnih pravilih izdelane zapise, ki so lahko rezultat retrospektivne konverzije ali pa katalogizacije de visu.
7. O postopkih za dopolnjevanje, popravljanje in brisanje zapisov odloča Konzorcij.
8. Zaželeno je, da dostopni elementi (tj. značnice) ustrezajo nacionalnim normativnim oblikam, vendar so sprejemljive tudi ostale.

Kot je videti iz zgornjih načel, je Konzorcij pripravljen sprejeti praktično kakršnekoli zapise. Prvotni proevropski entuziazem o obvezni uporabi formata UNIMARC je splahnel ob dejstvu, da so zaradi ameriškega gostitelja morali biti zapisi za podatkovno zbirko HPB preoblikovani v format USMARC, ki pa ga tudi nekatere članice CERL-a uporabljajo kot delovni format. Konverzija iz formata USMARC v UNIMARC zaradi prvotnega načela Konzorcija in nato spet v USMARC zaradi sistema RLIN pa seveda ni smiselna.

5 HPB (podatkovna zbirka za gradivo iz obdobja ročnega tiska)

5.1 Struktura podatkovne zbirke

Podatkovna zbirka HPB ni homogena, ker se zapisi po obsegu, nastanku in uporabljenih pravilih zelo razlikujejo in določeno enotnost kažejo le v okviru posameznih datotek, kot jih prispevajo članice. Zato zbirko sestavljajo te datoteke, trenutno jih je dvanajst (6). Vseh zapisov je ca. 865.000, posamezne datoteke pa se po obsegu zelo razlikujejo: od 3.500 zelo natančnih zapisov Nacionalne i sveučilišne knjižnice v Zagrebu, izdelanih de visu, pa do prek 500.000 večinoma z retrospektivno konverzijo izdelanih zapisov Bavarske državne knjižnice. Prispevek Narodne in univerzitetne knjižnice v Ljubljani sodi s ca. 19.000 zapisi med srednje velike in je nastal v okviru retrospektivne konverzije njenega najstarejšega kataloga.

5.2 Navodila za pripravo zapisov

Za vse članice je na voljo tudi osnovna dokumentacija z navodili, ki naj jih članice upoštevajo pri pripravi zapisov. Konzorcij sicer načelno sprejema vse zapise, ne glede na uporabljena katalogizacijska pravila, da le izpolnjujejo minimalne zahteve glede obsega podatkov v računalniške formatu. Gre torej za priporočila v zvezi s katalogizacijo, formatom UNIMARC in postopki pri pošiljanju zapisov (7):

- Na področju katalogizacijskih pravil je bil ISBD(A) izbran kot priporočeni in zaželeni standard za bibliografski opis tega gradiva. Navodila obsegajo predvsem podrobnejša določila za tiste primere, kjer standard sam dopušča alternativne rešitve.
- V okviru formata UNIMARC so bila poleg že obstoječih obveznih polj dodatno določena še:
 - nekatera priporočena polja za podatke, ki so za tovrstno gradivo pomembni (npr. kraj izida v normativni obliki v polju 620)
 - dodatna polja za lokalno rabo oz. za rabo v HPB (torej polja 9—, -9-, -9, npr. za alternativne oblike imen in za podatke o lokaciji – podobno kot v formatu COMARC).

Poleg tega se je istočasno začelo zelo plodno sodelovanje s Stalnim odborom za UNIMARC (Permanent UNIMARC Committee - PUC). Konzultant CERL-a je pripravil obsežen predlog za določitev novih polj, ki so za to gradivo ključnega pomena. PUC je pri nadaljnjem razvoju UNIMARC-a

večino predlogov upošteval in s tem dokazal, da je UNIMARC format, ki lahko polno zaživi upoštevajoč potrebe in predloge svojih uporabnikov.

- Določeni so bili postopki, ki jih morajo članice upoštevati pri pošiljanju zapisov (načini pošiljanja, spremna dokumentacija itd.).

5.3 Bibliografski podatki in računalniški format

Čeprav je ISBD(A) priporočen kot standard za bibliografski opis tega gradiva, pa to priporočilo lahko seveda velja le za zapise, izdelane na novo. Zato zbirka obsega paleto zapisov, izdelanih po najrazličnejših pravilih (npr. münchenska pravila, pruske inštrukcije, Panizzijeva pravila, RAK, AACR itd.). To dejstvo ni pomembno le za bibliografski opis, ampak ključno vpliva na izbor in obliko značnic, ki se zato za ista imena lahko zelo razlikujejo. Ekspertna skupina Konzorcija je tako pričela z načrtovanjem normativne kontrole, pri tem pa je upoštevala osnovno načelo, da Konzorcij članicam ne želi vsiljevati pravil za oblikovanje značnic v smislu edino veljavne enotne značnice. Klasična normativna datoteka, za katero je značilna struktura nadrejene enotne značnice, povezane s podrejenimi alternativnimi oblikami, ob upoštevanju takega načela ni primerna. Zato je bil izbran drugačen model, ki povezuje vse alternativne oblike določenega imena kot enakovredne dostopne elemente. To je obenem pristop, o katerem je zadnje čase veliko govora tudi v okviru Ifle in bo verjetno dobil veljavo mednarodnih priporočil. Zaenkrat je v pripravi tako urejen seznam imen za kraje izida (za katerega je v polju 620 formata UNIMARC predviden vnos v obliki značnice). Sledila bo normativna kontrola imen, med katerimi za to gradivo niso pomembni le avtorji, pač pa tudi tiskarji, bivši lastniki itd., zato je zaželeno, da zapisi vsebujejo tudi njihova imena v obliki dostopnih elementov, tj. značnic.

Omenili smo, da je kraj izida v obliki enotne značnice zaželen element v zapisih. S tem je povezana tudi uporaba šifre za državo izida v polju 102 formata UNIMARC. UNIMARC določa, da je potrebno uporabiti šifro za državo glede na sedanje stanje, kar pa je za tovrstno gradivo zaradi nastanka novih držav, sprememb obstoječih meja itd. seveda lahko zavajajoč podatek. Zato je bilo veliko diskusij na temo, ali je ta podatek sploh ustrezen ali pa bi bilo bolje uporabo polja opustiti. Na koncu je prevladalo mnenje, da je podatek z določenega stališča le informativen, da pa morajo biti uporabniki opozorjeni na njegov pomen.

Za publikacije v več zvezkih UNIMARC omogoča hierarhično povezovanje zapisov in kar nekaj knjižnic tak način tega povezovanja tudi uporablja. Tu je CERL naletel na problem prenosa teh podatkov v USMARC, ki takega povezovanja ne pozna. Narejena je bila zasilna rešitev, ki pa za iskanje ni prijazna.

5. 4 Uporaba podatkovne zbirke

Ker se zapisi med seboj razlikujejo po obsegu podatkov, jeziku katalogizacije, katalogizacijskih pravilih itd., je to za uporabnike lahko precej moteče. Zato je bil pripravljen priročnik za uporabo podatkovne zbirke (8), ki je uporabnikom na voljo vgrajen kot pomoč pri iskanju. Priročnik opisuje obseg in vsebino posamezne datoteke, podatke o uporabljenih katalogizacijskih pravilih, o (ne)uporabljenih poljih v računalniškem formatu in njihovem pomenu pri iskanju, o načinu opisovanja publikacij v več zvezkih, daje pa tudi nekaj uporabnih navodil za iskanje.

Ob tem je zanimivo omeniti, da v sistemu RLIN nasploh velja pravilo, da ne izločajo podvojenih zapisov, ampak s postopkom, ki ga imenujejo 'clustering', tvorijo skupine zapisov, ki jih družijo enake vrednosti v določenih poljih/podpoljih in zato predvidoma predstavljajo isto bibliografsko enoto. Tako knjižnice lahko za prevzemanje izberejo tisti zapis, ki najbolje ustreza njihovi katalogizacijski praksi. Algoritem tvorijo podatki iz impresuma, naslov (oz. najkrajša različica naslova) brez upoštevanja velikih začetnic in ločil, glavna paginacija in podatek o izdaji. G. Jonsson (9) je ta algoritem analizirala na primerih zapisov iz podatkovne zbirke HPB in ugotovila, da algoritem z manjšimi popravki ustreza, zaradi značilnosti gradiva in s tem povezanimi posebnostmi pri katalogizaciji pa zadnja dva elementa (paginacija in podatek o izdaji) ne prideta v poštev. Njeno mnenje je tudi, da morda le ni tako pomembno, da ta postopek omogoči natančno ločevanje med različnimi izdajami. Ne le da je to pri pri tem gradivu izjemno težko, kot je tudi razvidno iz zgoraj prikazanega primera, ampak uporabnik morda pri tem gradivu raje dobi seznam več kot pa manj možnih zapisov, ki ustrezajo njegovi poizvedbi. Ob tem opozarja na izraz 'bibliotype' ki ga je skoval švedski bibliograf Rolf Du Rietz za zapis, ki ne opisuje natančno določene izdaje nekega dela, ampak t.i. prototip tega dela. Glede na novo shemo, ki jo uvaja študije Ifle Funkcionalne zahteve za bibliografske zapise (10), je to torej zapis, ki opisuje izrazno obliko dela in ne njene pojavne oblike, torej publikacije same.

6 Razvojni načrti CERL-a

Vzpostavitev podatkovne zbirke je Konzorciju vzela bistveno več časa, kot je bilo spočetka predvideno, obenem je bilo treba veliko časa posvetiti organizaciji Konzorcija samega. Članice so v projekt vložile visoka finančna sredstva in delo svojih strokovnjakov, kot rezultat pa so šele leta 1997 dobile možnost za uporabo zbirke. Tiste, ki dotlej še niso bile članice sistema RLIN, so pridobile več, saj pogodba omogoča tudi uporabo njihove osrednje bibliografske zbirke in nekaj

specializiranih zbirk. Vendar pa Konzorcij čaka še nekaj nujnih razvojnih nalog, med katerimi so najpomembnejše:

1. vzpostavitev normativne kontrole
2. omogočanje prevzemanja zapisov
3. izboljšano obveščanje potencialnih uporabnikov

Prva naloga je, kot smo že omenili, v teku. Ravno kar potekajo tudi priprave za izvedbo druge naloge, torej prevzemanja zapisov. Morda še najbolj šepa pridobivanje uporabnikov: izkazalo se je, da članice same svoje uporabnike relativno slabo obveščajo o možnosti uporabe podatkovne zbirke. Konzorcij zato poskuša okrepiti neposreden stik z uporabniki: njegova domača stran (<http://www.cerl.org>) je zelo informativna in vsebuje številne povezave s sorodnimi projekti in ustanovami, od antikvarnih ponudnikov in projektov v zvezi z rokopisnim gradivom do specializiranih inštitutov za zgodovino knjige. Konzorcij je začel izdajati tudi tiskane novice, strokovni seminar ob letnem srečanju pa dodatno privablja interesente. Profesionalizacija kadrovske zasedbe je verjetno dobra naložba za uspeh celotnega projekta.

7 Zaključek

7.1 Prednosti in pomanjkljivosti specializirane podatkovne zbirke

Kot je bilo omenjeno na začetku, podatkovna zbirka HPB vsebuje zapise, ki so sicer že dostopni prek drugih knjižničnih sistemov, gre torej v določeni meri za podvajanje ponudbe. Kljub temu lahko naštejemo številne prednosti specializirane podatkovne zbirke:

- ponudba podatkov na enem mestu z upoštevanjem značilnosti gradiva
- ta podatkovna zbirka ima načrtan časovni okvir glede na leto izida, zato je možno celo optimistično predvideti, da bo nekoč popolna, da bodo torej v njej zbrani podatki o vsem gradivu iz tega obdobja
- projekt s skupnim ciljem povezuje evropske knjižnice in jih vzpodbuja h krepitvi sodelovanja in usklajevanja pri pridobivanju, obdelavi in ponudbi tega gradiva (CERL zanje uspešno organizira tudi seminarje o obdelavi gradiva)
- CERL je pomemben povezovalni element z drugimi projekti in ustanovami s tega področja

Seveda pa ne gre zamolčati pomanjkljivosti:

- visoki zagonski stroški
- vsaj trenutno relativno majhne koristi glede na vložek

S stališča Slovenije oz. Narodne in univerzitetne knjižnice kot članice pa še:

- število potencialnih uporabnikov te specializirane zbirke je majhno (čeprav je res, da tudi pri nas nismo dovolj pozornosti posvetili promoviranju zbirke)
- zapisi, ki jih je v zbirko prispeval NUK, so na voljo tudi v slovenskem vzajemnem katalogu in velja ocena, da pokrivajo večino potreb pri katalogizaciji tega gradiva
- možnost prevzemanja zapisov iz podatkovne zbirke HPB tudi za NUK ni aktualna, saj bo še letos s konverzijo glavnega kataloga pridobil tudi ostale še manjkajoče zapise za to gradivo.

7.2 Gradivo iz obdobja ročnega tiska v slovenskem vzajemnem katalogu

Za Narodno in univerzitetno knjižnico je vključitev v ta projekt v mednarodnem merilu pomenila predvsem promocijo slovenske države in slovenskega knjižničarstva. Obenem je prispevala te zapise v slovenski vzajemni katalog COBIB in so tako na voljo za prevzemanje. Izkušnje pri sodelovanju v CERL-u pa lahko izkoristimo pri načrtovanju ponudbe tega in tudi drugega specializiranega gradiva v okviru slovenskega vzajemnega kataloga.

Podatkovna zbirka COBIB namreč obsega zapise za zelo raznovrstno gradivo: monografske, serijske publikacije in članke na vseh medijih. Poleg razdelitve po bibliografskih nivojih in fizičnih nosilcih pa bi lahko pripravili in ponudili dodatne vsebinsko določene segmente iz celotne zbirke: razen gradiva iz obdobja ročnega tiska npr. tudi domoznansko gradivo, učbenike, leposlovje, mladinsko literaturo itd. Razen za domoznansko gradivo je segmentacija možna glede na podatke v šifriranih poljih, tako bi te podatke tudi bolje izkoristili. Možnosti za poizvedovanje po posameznem segmentu bi morale biti prilagojene značilnostim gradiva in vsaj za gradivo iz obdobja ročnega tiska nam zbirka HPB daje kar nekaj namigov.

Izkušnje pri sodelovanju v Konzorciju lahko strnemo v nekaj točk:

- pomembno je natančno določiti cilje takega projekta
- nujen je vnaprejšen dogovor o uporabljeni metodologiji in vrstnem redu postopkov
- določiti je treba medsebojne obveznosti med konzorcijem in drugimi sodelujočimi partnerji.

Citirani viri

1. *Retrospective cataloguing in Europe: 15th to 19th century printed materials: Proceedings of the international conference, Munich, 28th - 30 November 1990* (1992). München: K.G. Saur
2. Kaltwasser, F. G. (1992). Language barriers in data exchange. V *Retrospective cataloguing in Europe: 15th to 19th century printed materials: Proceedings of the international conference, Munich, 28th - 30 November 1990*. München: K.G. Saur
3. ISBD(A). (1998). Ljubljana: Narodna in univerzitetna knjižnica
4. *Application of UNIMARC to multinational databases*. (1999). München: K.G. Saur
5. Consortium of European Research Libraries. (19.4.2000). *Functional principles for gathering and inputting information to Hand Press Book database*. London: CERL. Pridobljeno 12.7.2000 s svetovnega spleta: <http://www.cerl.org/hpbfunc.htm>
6. Consortium of European Research Libraries. (8.6.2000). *The Hand Press Book (HPB) database*. London: CERL. Pridobljeno 12.7.2000 s svetovnega spleta: <http://www.cerl.org/hpb.htm>
7. Consortium of European Research Libraries. (24.5.2000). *CERL file procedures (FILPROC) and cataloguing rules (UNICAT)*. London: CERL. Pridobljeno 12.7.2000 s svetovnega spleta: <http://www.cerl.org/unadintr.htm>
8. Consortium of European Research Libraries. (2000). *Manual for the Hand Press Book Database via Eureka on the Web*. London: CERL
9. Jonsson, G. (1997). *The HPB file and clustering*. London: CERL. Pridobljeno 12.7.2000 s svetovnega spleta: <http://www.cerl.org/jons97.htm>
10. *Funkcionalne zahteve za bibliografske zapise* (2000). Ljubljana: Oddelek za bibliotekarstvo, Filozofska fakulteta

MERGING PAST AND FUTURE: TRANSFERRING OLD CATALOGUE RECORDS INTO AUTOMATED SYSTEMS (SLOVENIAN EXAMPLE)

Zlata Dimec

Maja Žumer

Izvirni znanstveni članek

UDC 027.54(497.4 Ljubljana):025.354:004.3

Abstract

Retrospective conversion as a conversion of data from an old to a newer format is usually realized as a transfer of card catalogue into a computer form in a fast and simple way. The National and University Library in Ljubljana, Slovenia, started its first retro-conversion project in 1997 with the oldest, partly handwritten catalogue, covering the period 1774-1947. The description of the catalogue is given, taking into account both cards' characteristics (size, handwritten/penwritten, scripts used etc.) and different cataloguing rules applied at that period (Austrian Library Instruction, Prussian Instruction etc.). The whole process was divided into separate phases according to technologies and human resources available, and specific procedures were designed (transcription, manual and automatic structuring, control etc.). The results are ca. 70,000 records in MARC format, which were included into existing library catalogue. Records for items published until 1830 were contributed to Consortium of European Research Libraries (CERL) Hand Press Books database hosted by Research Libraries Group (RLG). NUL is currently working on retroconversion of remaining catalogues (ca. 300,000 records) where different procedures have to be designed due to different characteristics. The experience gained during these projects may be found useful by other libraries that are either starting or considering retrospective conversion.

Key words: National and university library Ljubljana, card catalogues, retrospective conversion

* *Prispevek je bil objavljen v publikaciji »Libraries and associations in the transient world: New technologies and new forms of cooperation : Conference proceedings.« Sudak, Ukraine, June 5-13, 1999 (str. 156-162). Simferopol' : Tavrida, 1999.*

DIMEC, Zlata; MAJA ŽUMER: Merging past and future: transferring old catalogue records into automated systems (Slovenian example). Knjižnica, Ljubljana, 46(2002)3, 119-127

1. Introduction

Each library planning its retrospective conversion should consider carefully the reasons for and goals of such an expensive project, and then analyze the existing resources: printed bibliographies, inventory lists, and most often, the card catalogues. The National and University Library in Ljubljana, Slovenia, chose to start the project with its oldest card catalogue. The reasons were twofold: this catalogue itself represents a cultural heritage as the cards date in the early 19th century, and many of them are identified as a handwriting of distinguished scholars who ran the Library. The cards are deteriorating and it was necessary to remove the catalogue from even the restricted use, but make it available for historical research. The second reason was the extent of the catalogue, of which both the size and the concept were considered a closed entity.

2. Description of the catalogue

The old card catalogue covers the period 1774-1947, i.e. from the establishment of the Library until a couple of years after the end of the World War II, when new cataloguing rules were introduced. The catalogue holds ca. 95,000 cards for monographs and serials in different languages, from which Latin, German and Slovenian are most heavily represented, followed by other Slavic languages, Italian and English, not excluding Greek and Hebrew.

Ca. 60 % of these cards are handwritten, many of them in black letter. An example is in Fig. 1. The scripts are often very difficult to read. Most of these cards have a form of a template, divided into eight fields, each designed for a set of data: shelfmark, subject term, format, heading, bibliographic data, imprint, key to the subject catalogue, and the old shelfmark. Bibliographic data in the biggest field may consist of the title proper, statement of responsibility, edition statement, physical description (except format), series, and notes which may be very extensive. These templates however show us a well considered structure, easily comparable with the structure of modern bibliographic records. Newer cards, some still handwritten and later typed, lack the template form, the subject term, and have in general less elements of bibliographic description (Fig. 2).

For historical reasons, the language of cataloguing in older cards (templates) is German, and in others Slovenian. For the same reasons, the concept of the cata-

logue followed the Austrian Library Instruction¹. The cataloguing rules applied in this catalogue vary from the Schrettinger's rules² to the famous Prussian Instructions³, both of which were widely adapted to the needs of the Library and to the librarians' judgments – which, if not recorded, may prevent a successful and adequate retroconversion – a very topical lesson for modern librarians, being aware of future changes.

271.	Literat. clasp. Roman. Poetis	4 ^a
Horatius Q. Flavius.	Quinti Horatii Flacci Opera, interpre- tatione et notis illustravit Ludovicus Fodprea, jussu Christianissimi Regis, et usum Universitatis Delphensis, ac Universi- tatum Principum Burgundiae, Andaniae, Biturigenae.	Massani, 1777. Ad pressum Hactenus opac Romanorum
II 127	jussu Universitatis	I cont. 11
10714		I. H. C. C. L.

Fig. 1

- 1 Instruction vorgeschrieben für alle Universitäts- und Lycealbibliotheken mit Hofdecrete vom 30. April 1778, and the later edition entitled Instruction für die k.k. Universitäts- und Studienbibliotheken, provisorisch erlassen mit Stud. Hof.-Comm.-Decrete vom 23. Juli 1825
- 2 Schrettinger, M: Versuch eines vollständiges Lehrbuchs der Bibliothek-Wissenschaft. München, 1829
- 3 Instruktionen für die alphabetischen Kataloge der preussischen Bibliotheken vom 10. Mai 1899. 2. Ausg. 1909



Fig. 2

3. Retroconversion process

Funding for the project was received from the Ministry of Culture of Slovenia and Regional Library Programmes of Open Society Institute in equal parts and the work began in 1996.

As the first phase all catalogue cards were scanned. Colour scanning was chosen because we wanted to keep an exact digital copy of the catalog for preservation and archiving purposes. The process took place in the library to ensure both the security of the cards and minimal interference with the everyday use of the catalog. At the same time additional data were added to the scanned images: shelf-mark and an indication of the type of each card (pre-1900 handwritten form or typed).

The whole process for each type was different. The oldest cards are well structured and could be transcribed and structured at the same time even by non-librarians. A group of students with good knowledge of German (including black letter) and Latin transcribed the data from each scanned image using a special utility program. The interface is shown in Figure 3.



Fig. 3

The other half of the catalogue cards were first keyed-in, keeping both the layout and punctuation. The process was out-sourced. OCR had been considered first, but the poor quality of the paper and typewriters used resulted in too many manual corrections to justify that approach.

All these records were created following the same cataloguing rules, and after extensive sampling which showed that the records were both consistent and relatively error-free, we decided to use automatic structuring. An algorithm was designed which enabled a computer programme to identify elements of bibliographic records from their relative position and punctuation. The results were encouraging: 80% of the records were structured directly, the rest needed only minimal manual intervention to enable automatic structuring in the next step.

At that stage structured records from both parts were combined for the last phase: control. Indexes of all fields were systematically reviewed and errors and inconsistencies were identified and corrected using a utility programme. The records were then ready for the conversion into any MARC format.

4. Results of the retroconversion

As already mentioned, the electronic copy of the catalogue was one of the goals of this project. The electronic catalogue (Fig. 4-5) is a faithful copy as it keeps the original number of drawers, including the number and arrangement of cards in them. The cards can be rotated (to accommodate reading the notes written perpendicularly), and even their backs can be viewed if written. Additional benefit is the option of enlargement which makes reading easier. The catalogue is now available via the Library server, ready for scholars who wish to do any historical research.

After achieving the uniform structure of records, the conversion into UNIMARC format (Fig. 6) was prepared for records covering books printed in the hand-press period in Europe (until 1830). 17,000 records were contributed to the Hand Press Books database established by the Consortium of European Research Libraries (CERL) and available via RLIN system. All the records derived from the retroconversion (ca 65,000) were then converted into the Slovenian COMARC format (based on UNIMARC - Fig. 7) and included into the Library local database. After deduplication procedure, they were loaded into the Slovenian shared cataloguing system COBISS (COoperative Bibliographic Systems and Services), available to other Slovenian libraries for downloading and use.

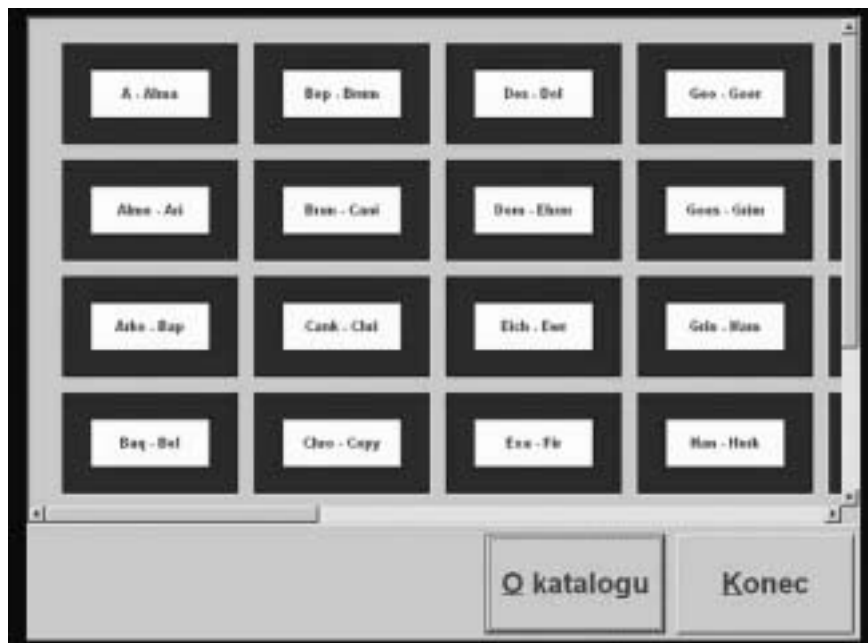


Fig. 4



Fig. 5

Label: 00753nam0#22002051n#450
 #001SI\NUK\RK\007287
 #00519980716143511.0
 100##\$a19971018d1777#### || || 0gery0103####ba
 101 | # \$alat
 2000#\$aQuinti Horatii Flacci Opera, interpretatione et notis illustravit
 Ludovicus Desprez. jussu Christianissimi Regis, in usum Serenissimi
 Delphini, ac Serenissimorum Principum Burgundiae, Andium, Biturigum
 210##\$aBassani\$cSed Prostant Venetiis Apud Remondini\$d1777
 215##\$d4°
 300##\$aSehr brauchbar
 6100#\$alateinische Literatur
 675##\$a821.124.09
 700#1\$aHoratius Flaccus\$bQuintus\$4070
 801#0\$aSI\$bNUK
 801#1\$aSI\$bNUK\$c19971018
 801#2\$aSI\$bNUK\$c19980716\$2unimarc
 899#\$aSI\NUK\$j971

Fig. 6

```

ID=92073216 LN=0000444238 M V1 10.10.1997 NUK::RETRORK
001 <a>n - novi zapis <b>a - tekstovno gradivo - tiskano <c>m - zaključ-
na
    publikacija <d>0 - ni hierarhičnega odnosa <g>1 - podnivo 1 (podatki,
    prevzeti s kataložnega listka) <h>n - zapis ni v ISBD obliki
100 <b>d - monografska publikacija, zaključena <c>1934 <g>0 - nemod-
ificiran
    stavek <h>slv - slovenski <l>ba - latinica
101 | <a>ita - italijanski
200 <a>Sulla tubercolosi miliare del polmone. Osservazioni Cliniche
210 <a>Foligno <c><s.n.> <d>1934
215 <d>4o
300 <a>Estr. da: Rinnovamento Medico, Il Giornale di Tisiologia, 1934.
    adligat k 66031
700 1<a>Leproni <b>Federico <4>070 - avtor
996 1<d>n66054

```

Fig. 7

5. Conclusions

The results of the first phase of the retroconversion are encouraging. The concept of old cataloguing rules has showed an admirable consistency. It's been proved that the rules in past 200 years have not changed dramatically, keeping to a certain extent effectiveness in the modern environment of MARC formats. This fact also indicates that the experiences gained in any library or environment may be valuable to others, taking into account that the methods of the whole process may differ, depending on the starting points, reasons and goals of the project.

6. Future plans

The retrospective conversion of the old catalogue was the first step. During its final stages the next part, the conversion of the catalogues for the period 1948-1988 was planned. There are several catalogues covering this period, and after a careful analysis the systematic (UDC) catalogue was chosen as the basis. This was the only way to ensure at least a broad subject access to the converted records. We are estimating that the final result will be about 350,000 bibliographic records.

All the catalogue cards have already been scanned. Shelfmarks, subject headings and UDC notations were entered at the same time. By the fall 1999, all cards will be transcribed using OCR as much as possible. At that time the records will be publicly available as a full text (unstructured) database enabling searching on shelfmark, subject heading, UDC, and keyword.

Funding for that part was also received from Ministry of Culture and Open Society Institute. Upon approval of the funding for the final stage, the full text records will be structured automatically, controlled and converted into MARC format.

Although this is not strictly a part of the retrospective conversion project, some future plans for further processing of the resulting records have to be mentioned as well. All the records which belong to the national bibliography will be identified and enhanced, most of the respective items will be recatalogued. National and University Library will thus fulfill its duty as the national bibliographic agency.

7. Literature

Kodrič-Dačić, Eva: Matija Čop in njegov prispevek k slovenskemu bibliotekarstvu. *Knjižnica* 41: 2/3 (1997) 17-31

Pirjevec, Avgust: *Knjižnice in knjižničarsko delo*. V Celju: Družba sv. Mohorja, 1940

Žumer, Maja: Retrospektivna konverzija starih katalogov Narodne in univerzitetne knjižnice. In: *Zbornik razprav: 10 let Oddelka za bibliotekarstvo (1987-1997)*. Ljubljana: FF, 1998. P. 215-229