

- <sup>9</sup> Boss R.W., Espo H. Standards, database design, & retrospective conversion // *Library journal*. - Vol. 112 (Oct. 1, 1987), p. 54.
- <sup>10</sup> Crismond L.F. Quality issues in retrospective conversion projects // *Library resources & technical services*. - Vol. 25 (Jan./Mar. 1981), p. 48-55.
- <sup>11</sup> Retrokonversion : Konversion von Zettelkatalogen in deutschen Hochschulbibliotheken. - Berlin : Dt. Bibliotheksinst., 1993. - 411 p.
- <sup>12</sup> Lux C. Vergleich der getesteten Konversionsmethoden // *Retrokonversion : Konversion von Zettelkatalogen in deutschen Hochschulbibliotheken*. - Berlin : Dt. Bibliotheksinst., 1993. - P. 233-284.
- <sup>13</sup> Vita S.H. PREMARC - a history and status report // *Cataloging & classification quarterly*. - Vol. 14, no. 3/4 (1992), p. 5-22.
- <sup>14</sup> McMillan G. Retrospective conversion of serials : implementing the MARC holdings format // *Technical services quarterly*. - Vol. 9, no. 3 (1992), p. 41-54.

- <sup>15</sup> USMARC format for holdings and locations. - Washington DC : Network Development and MARC Standards Office, Library of Congress, 1984.
- <sup>16</sup> USMARC format for holdings and locations. - Washington DC : Network Development and MARC Standards Office, Library of Congress, 1989.
- <sup>17</sup> ANSI : Z 39.44. Serial Holdings Statements. - New York : American National Standards Institute, 1986.
- <sup>18</sup> Zwink AE. Das DFG - Project "Maschinenlesbare Erfassung von Altebeständen 1501-1805". - Württembergische Landesbibliothek Stuttgart. - 12 p.
- <sup>19</sup> Willensen A.W. Retrospective national bibliography : Operational Projects and Possible Future Developments in the Netherland // *IFLA journal*. - Vol. 16, no. 1 (1990), p. 97-104.
- <sup>20</sup> Lietuvos integralios bibliotekų informacijos sistema : techninė užduotis kompiuterizuotai informacijos sistemai sukurti / parengė V. Kučiukas, G. Duobinienė, N. Kolesinskienė, R. Varnienė, T. Gustienė, A. Štreimikis. - Kaunas, 1996. - 180 p.

## Summary

### Retroconversion of Catalogues and National Bibliography

Regina VARNIENĖ

Start up of application of electronic cataloguing created perspectives for existing libraries to supply users with information on the old library holdings through the electronic catalogue.

Now, the question arises regarding the choice of adequate methods and approaches for installing into automated catalogues information on the existing old and valuable records, which have been created before the outset of electronic cataloguing.

Author of this article, who had an opportunity to get acquainted with the achievements of other countries in the field shares her opinion based on the experiences of USA, Germany, Holland, Check Republic.

Article concentrates on the analysis of the requirements for the structure of record conversion, as there has appeared in numerous literature on this topic some rather contradictory comments.

Based on the expertise and knowledge gained in German libraries and in reference to the publication "Retroconversion" prepared by the Deutsche Bibliotheksinstitut, different methods of recon are introduced in the article. These include its organizational matters, cost-effective-

ness, etc. Also, statistical data indicating to the material described is offered.

Two essential catalogue conversion projects of the scientific libraries of the United States are characterized, i.e., PREMARC/REMARC of the Library of Congress and VTLS - for serials.

Project titled "Retroconversion of Files of Old Materials" (1501-1850) of German Scientific Society is introduced, in which several scientific libraries are participating. Its implementation lasted for 6 years and was almost accomplished in 1994.

Also, a program of recon PICA is presented, which set forth the fundamentals for the Society's Union Catalogue and National Bibliography.

Subsystem of recon of the Lithuanian Integrated Library Information System (LIBIS) is surveyed specifying its main goal - conversion of enormous traditional library information products, i.e., printed bibliographic materials, catalogues, cards into the machine-readable media and creating possibilities for utilization of these data employing an on-line method.

UDK 022.9(437.1)+025.3/6:681.3(437.1)

## Retrospektyvioji konversija Čekijos bibliotekose

Bohdana STOKLASOVÁ

Čekijos nacionalinė biblioteka, Klementinum 190, 110 01 Prague 1, Czech Republic,  
e-mail: bohdana.stoklasova@nkp.cz

Miroslav BAREŠ

Comdat, Žirovnická 2389, 106 00 Praha 10, Czech Republic

### Įžanga

Stengdamiesi realizuoti vakarietiškas komunikacijos spereitkelių, bibliotekų bei virtualių bibliotekų idėjas, mes, buvę sovietiniame bloke, gerokai greičiau turėjome žengti neišvengiamą žingsnį - atverti bibliotekas skaitytojams. Viena iš pagrindinių atviros bibliotekos pralaidų - padaryti jos katalogą, sukauptos kolekcijos raktą, plačiai ir laisvai prieinamą. O šios pasaulio dalies bibliotekos dar nėra prie to įpratusios.

Šiais laikais naujai įgytų spaudinių automatizuotas apdorėjimas ir kompiuterinis katalogas yra veiksminga priemonė, leidžianti bibliotekos fondus padaryti prieinamus visam pasauliui. Išsivysčiusiose šalyse, turinčiose tvirtą ekonominį pagrindą, sukurti dideli, kartais net valstybės remiami, retrokonversijos projektai. Kelios Vidurio ir Rytų Europos pokomunistinių šalių bei buvusios Sovietų Sąjungos bibliotekos pradeda naudoti integruotas bibliotekines sistemas, o kai kurios net turi kompiuterinius katalogus, dirbančius dialogo režimu. Deja, šie katalogai atskleidžia tik menką dalelę kultūrinio palikimo, kuris lig šiol lieka paslėptas už privačių bibliotekų sienų, nežinomas ir neprieinamas pasauliui. Be to, pokomunistinių šalių dabartinė nacionalinė knygų leidyba tarptautinėse duomenų bazėse atskleidžiama labai menkai, tuo dar labiau ribojamas priėjimas prie kolekcijų. Dėl šių priežasčių didžioji šių rinkinių dalis iki šiol lieka nežinoma pasauliui. Apie 90% jų turi būti retrospektyviai konvertuota, o pokomunistinėms šalims šiuo metu tai atrodo beveik neišsprendžiama problema.

Katalogų yra labai daug, ir jie kelių šimtų metų senumo, todėl jų konversija finansiškai neįgyvendinama, nes be visa ko trūksta personalo ir laiko. Šių šalių ekonomika nepajėgi ir artimiausiu metu vyriausybė negalės efektyviai paremti savo bibliotekų. Skirtingai negu Vakaruose ir JAV, čia nėra ir pakankamai vietinių rėmėjų. Dėl šių priežasčių retrokonversijos perspektyva pokomunistinėse šalyse atrodo beviltiška, o bibliotekų atvėrimas pasauliui artimiausioje ateityje - nerealus.

Yra keletas pagrindinių veiksnių, trukdančių atlikti šių šalių bibliotekų katalogų retrokonversiją. Iš jų svarbiausi yra šie:

- \* lėšų stoka dabar ir artimiausioje ateityje,
- \* techninės įrangos stoka,
- \* žinių apie dabartines technologijas trūkumas,
- \* žinių apie tarptautinius standartus stoka dėl to, kad dešimtmečiais gyventa už geležinės uždangos,
- \* didelė katalogų apimtis ir jų turinys.

Paskutinis veiksnys yra lemiamas: dauguma bibliotekų turi labai senų spaudinių ir jų katalogų kortelės yra gana senos. Todėl tikimybė rasti atitinkamus įrašus egzistuojančiose duomenų bazėse labai menka. Katalogų kortelės nelotyniškais rašmenimis - dar viena papildoma problema, kadangi nėra vieno tarptautiniu mastu pripažinto transliteravimo standarto.

Dėl išvardytų priežasčių retrokonversija OCLC duomenų bazės pagrindu yra gana problematiška. Atrodytų, kad SAZTEC duomenų apdorėjimo būdas būtų tinkamesnis, tačiau Vidurio ir Rytų Europos bei buvusios Sovietų Sąjungos bibliotekoms tai labai brangi alternatyva. Kitas nepigus, bet gana produktyvus, visai neseniai pasiūlytas ir išbandytas metodas yra OCR/ICR naudojimas automatiškai katalogo kortelių konversijai. Tačiau šį metodą galima efektyviai naudoti tik kalboms, turinčioms gerai nusistovėjusį žodyną, arba jei kortelėse informacija pateikiama nuoseklia tvarka, t.y. įrašuose naudojama ISBD nurodyta skyryba. Tada galima lengvai pasinaudoti prieš tai sudarytu algoritmu transliteruojant į OCR/ICR. Dėl šių priežasčių dauguma bibliotekų linkusios pačios transkribuoti katalogų korteles. Tačiau šiuo atveju įstaiga turi nuveikti didžiulį darbą, o turint galvoje menką standartizacijos lygį, tokie įrašai dažnai bus tinkami naudoti tik toje bibliotekoje, kuri juos konvertavo, geriausiu atveju tik tam tikrame regione.

Esant ribotoms finansinėms galimybėms, nepakankamai technikos ir žmonių, ar galima rasti retrokonversijos metodą, kuriuo galima sukurti aukštos kokybės įrašus ir kuris tikrai padėtų atverti šios pasaulio dalies bibliotekų

fondus? Šiuo metu tokių metodų jau yra. Tam reikia naudoti pažangiausią technologiją, nereikalaujančią brangiai apmokamo profesionalų darbo, ir bibliotekoms bendradarbiaujant.

Pateiksime išvadas, prie kurių priėjome ir kurios galėtų būti naudingos kitoms šalims.

## Istorija

Čekijos nacionalinėje bibliotekoje 1994 m. retrokonversija vis dar buvo beviltiškoje padėtyje, nors per keletą pastarųjų metų buvome išnagrinę įvairiausių pasaulyje naudojamus retrokonversijos metodus, pvz., OCLC ir SAZTEC. OCLC siūlomi TAPECON ir RETROCON metodai atrodė tinkami keliems katalogo skyriams, bet ne visam masyvui, nes įrašų, atitinkančių viso bibliotekos fondo bendrojo katalogo korteles, procentas buvo labai nedidelis. SAZTEC metodas dėl technologijos pasirodė esąs tinkamesnis, tačiau dėl kainos (2,40 JAV dolerių už įrašą) jo teko atsisakyti, kadangi bendrame bibliotekos kataloge yra 3 mln. kortelių, o biblioteka nebuvo pajėgi surinkti ir dalelės reikiamų šiam darbui 7 mln. JAV dolerių.

Nors lėšų ir nebuvo, vis dėlto 1992 m. pradėjome retrokonversiją. Naudojamiesi informacijos apdorojimo kompanijos "Vodni stavby Tabor" paslaugomis, perrašėme katalogo korteles į kompiuterinį formatą. Kadangi "Vodni stavby Tabor" neturėjo jokios bibliotekinės veiklos patirties, mūsų darbuotojai turėjo mokytis jų personalą, analizuoti ir parengti kiekvieną kortelę prieš išsiųsdami jiems, patikrinti ir pataisyti beveik kiekvieną sugrįžusį įrašą. Tad projektas judėjo lėtai: kortelės sugrįždavo atgal negreit, o ir įrašų kokybė negerėjo.

Mūsų sutartis su "Vodni stavby Tabor" baigėsi 1995 m. birželio mėnesį.

Visi išvardyti retrokonversijos būdai turi tą patį trūkumą: reikia labai daug laiko, norint visą katalogą pertvarkyti į kompiuterinį formatą, be to, už baigtą darbą reikia sumokėti. Turint galvoje dabartinį biudžetą, negalime tokios didelės apimties darbo finansuoti greičiau nei per 10 metų.

Kitas šių metodų trūkumas, kad 3 mln. katalogo kortelių konversija reikalauja labai intensyvaus darbo. Be to, iškyla problemų dėl kainos. Mūsų vadovybė, Kultūros ministerija negalėjo skirti daugiau nei kelių šimtų tūkstančių kronų per metus, taigi, norint sukaupti 7 mln. JAV dolerių, tektų labai ilgai laukti. Laimė, gavome A.W. Mellono fondo dotaciją, bet tik technologijai įsigyti, o ne atlyginimams mokėti.

Kadangi netrukus bus baigta statyti nauja knygų saugykla Hostivaže (Prahos priemiestyje), neišvengiamai turėsime ten perkelti bibliotekos bendrąjį katalogą ir knygas. Tačiau negalime kelti katalogo dėžučių tol, kol neturėsime tinkamų naudojimui (ne tik archyvui) kito formato kopijų Klementinume, kur visuomenė yra įpratusi naudotis katalogu ir fondais. Buvome manę sukurti katalogo

mikrofišinių variantą, tačiau labai mažai bibliotekų panoro jį įsigyti ir neatsipirko išlaidos. Jeigu būtų buvę daug prenumeratorių, OLMS Verlag leidėjai ne tik padengtų išlaidas, bet dar ir gautume nemokamą kopiją. Turėjome rasti metodą, padėsiantį taupyti žmonių darbo sąnaudas ir naudoti pigesnę kompiuterinį apdorojimą, kur tik įmanoma, taip pat rasti būdų nustumti vis dar lemiamą "kamštį" - profesionalų darbą - į proceso pabaigą.

Mūsų idėja retrokonversijai panaudoti kitiems tikslams skirtą technologiją ir originali kompiuterinė programa, atitinkanti šį konkretų tikslą, gimė ieškant išieities iš susidariusios padėties ir atsižvelgiant į apribojimus.

Technologija yra modulinė ir universali, taikoma kiekvienoje bibliotekoje. Ji atsirado:

a) ČNB darbuotojams tikrinant ir nagrinėjant įvairius projektus;

b) iš patirties, kurią įgijo Čekijos mokslų akademijos biblioteka, skenuodama savo kortelių katalogą;

c) dėl A.W. Mellono fondo finansinės paramos;

d) iš COMDAT kompanijos informacijos apdorojimo patirties, naudojant tiek teksto, tiek vaizdo duomenis, iš jų didžiulio susidomėjimo šiuo projektu ir pagaliau jų kilniaširdiško noro patiems tobulinti reikalingą RETROCON technologiją be jokios finansinės paramos.

COMDAT prirėkė beveik metų naujai technologijai sukurti, patobulinti ir patikrinti, ir jie iki šiol turi nuostolių, kurių artimiausiu metu nesitiki padengti. Nacionalinė biblioteka irgi abejojo, ar gali pasirinkti net šį, pigesnę variantą. Žinoma, visos kitos Čekijos bibliotekos turi dėti visas pastangas, kad galėtų suteikti savo vartotojams bent būtiniausias paslaugas, turėdamos tokius ribotus finansinius išteklius. Ar galės jos skirti šiek tiek lėšų iš savo menkų išteklių retrokonversijai?

Nors tai brangus ir sudėtingas būdas, tačiau retrokonversija yra būtina, norint padaryti kolekcijas atviras ir prieinamas. Tai padėtų bibliotekoms efektyviai bendradarbiauti dabar ir ateityje bei susisiekti su tarptautiniais kompiuteriniais informacijos tinklais.

## Technologija

### I. Skenavimas

Pirmuoju etapu katalogo kortelės (arba spausdintos Čekijos nacionalinės bibliografijos puslapiai) skenuojami moderniu šiuolaikiniu įrenginiu, kuris garantuoja aukštos kokybės vaizdą, būtiną tolesniam apdorojimui. Specialiai sukurtos programinės įrangos sudaro sąlygas bibliotekos darbuotojams, lankytojams vartoti šį pagalbinių katalogų, vos tik nuskenavus, kad ir su tam tikrais apribojimais. Sistema leidžia pridėti pradmenį ir/ar kitą bibliografinę atžymą, jei to reikia.

### II. Transkripcija

Antruoju etapu skenuoti TIFF formato atvaizdai paverčiami ASC II ženklais. Jei vaizdas kokybiškas, procesas vyksta automatiškai, naudojant OCR, o migloti ar kiti atžvilgiai neaiškūs vaizdai konvertuojami rankiniu būdu.

### III. Kodavimas ir žymių rašymas

Trečiuoju etapu formatuoti įrašai verčiami UNIMARC (gali būti naudojami ir kiti MARC formatai) įrašais. Priklausomai nuo to, ar įrašas švarus, jo žymes užrašyti galima automatiškai, jei tik įmanoma taikyti reikiamą algoritmą, arba rankiniu būdu. Speciali programinė įranga padeda sukurti greitai ir aukštos kokybės įrašus, bibliotekos personalas apkraunamas labai nedaug, o naudotis labai nesudėtinga. Patirtis parodė, kad jei 80% (ar daugiau) skenuoto įrašo negalima užrašyti automatiškai, tai daug efektyviau jį konvertuoti rankiniu būdu.

## Prioritetai

Bendrojo katalogo kortelės yra labai nevienodos kokybės ir daugelį jų bus labai sunku konvertuoti automatiškai: žymių transkribavimas/užrašymas. Didelė kortelių dalis atitinka Čekijos knygų produkciją, o tiksliau - knygų leidybą politinių organizacijų, kurios daugelį šimtmečių buvo įsikūrusios toje vietoje, kur šiuo metu yra Čekijos Respublika. Kadangi esame įsitikinę, jog mūsų indėlis į tarptautinę duomenų bazę pirmiausia turi būti įrašai, atspindintys šalies leidybą, tai pagrindinis tikslas - pirmiausia juos konvertuoti. Kita priežastis ta, jog dauguma Čekijos bibliotekų turi daugiausia čekiškų knygų ir jei jos ketina atlikti retrokonversiją pagal pradėtą duomenų bazę, turime pajėgti aprūpinti jas daugybe tinkamų naudoti įrašų. Kai Čekijos knygų produkcija jau bus konvertuota, imsimės konvertuoti užsieninius leidinius, kurių dalį gali turėti ir kitos bibliotekos. Turime stengtis išvengti darbų dubliavimo ir nereikalingo lėšų švaistymo, kurios gali būti panaudotos kitiems tikslams ir uždaviniams.

Čekijos knygų leidyba iki 1922 m. atspindi spausdintoje Čekijos nacionalinėje bibliografijoje "Bibliografický katalog". Nuo 1983 m. ši informacija jau yra prieinama kompiuterijoje. Tačiau turime galvoje, kad reikia konvertuoti dar maždaug 400 000 kortelių, apimančių 1922-1982 m. Jei pridėsime ir šio šimtmečio pradžios publikacijas, kurias galima rasti tik bibliotekos bendrajame kataloge ir spausdintose bibliografijose, tektų konvertuoti apie pusę milijono įrašų.

ČNB bendrajame kataloge yra beveik 3 mln. katalogo kortelių, iš kurių, deja, tik nedidelę dalį bus galima konvertuoti per OCR ir automatiškai užrašyti į atitinkamus laukus. Pagrindinė kliūtis rašyti žymes automatiškai yra tai, kad kortelės buvo parengtos naudojant skirtingas katalogavimo taisykles. Taip aprašytas korteleles nelengva atpažinti net pačiai išradingiausiai programinei įrangai.

## Finansiniai poreikiai, terminai

Bendrasis katalogas bus prieinamas skaitytojams, kai baigsime pirmąjį etapą, t.y. skenavimą. Tačiau manome, kad prireiks dar 10 metų, kol visi įrašai bus paversti UNIMARC įrašais. Žinoma, jei turėtume daugiau lėšų, įrangos bei kvalifikuoto personalo, ši bauginanti pranašystė nepasitvirtintų.

Pirmajame etape (1995 m. rusėjis - 1996 m. birželis) gavome A.W. Mellono fondo stipendiją technologinei įrangai pirkti, o Kultūros ministerija finansavo skenavimo procesą. Kitas užduotis, t.y. konversiją ir žymių rašymą, pradėjome 1995 m. Kitiems dviems etapams įgyvendinti vėl reikės lėšų, ypač atlyginimams, ir tam, kad rezultatai būtų prieinami plačiajai visuomenei (1996 m.). Norime sutvarkyti visą XX a. Čekijos knygų produkciją iki 2000-ųjų metų.

Pirmajame etape (1995-1996 m.) išnaudojome A.W. Mellono fondo stipendiją technologijai. Vėliau skenuosime, konvertuosime į OCR ir rašysime žymes, nors visus 1995-uosius metus toliau tikrinome ir tobulinome technologiją. Ateityje tikriausiai prireiks lėšų skenavimui, konvertavimui ir žymių rašymui.

Mūsų paskaičiavimais, kiekvieno įrašo konversija kainuos maždaug 1,6 JAV dolerio:

- \* skenavimas ir OCR - 0,4 JAV dolerio
- \* žymių rašymas - 0,8 JAV dolerio
- \* redagavimas - 0,4 JAV dolerio.

Kaina beveik trečdaliu mažesnė negu SAZTEC nustatyta, o galutinis rezultatas geresnis. SAZTEC transkribuoja kortelę tokią, kokia yra, o mūsų projektas suteikia galimybę konvertuoti kiekvieną kortelę į ISBD įrašą, be to, mes atliksime autorizuotą patikrą. Taigi dirbame patys, todėl viskas kainuoja pigiau, galime greičiau sutvarkyti katalogą ir leisti jį naudoti.

## Bibliotekų bendradarbiavimas ir personalas

Bendrojo bibliotekos katalogo retrokonversija pirmiausia yra kiekvienos atskiros institucijos pareiga. Taip ir ČNB bendrojo katalogo retrokonversiją atliks pati biblioteka. Žinoma, kitoms bibliotekoms pasiūlysiame pasinaudoti mūsų rezultatais, nors gal kai kurioms bibliotekoms atrodys, kad iš to bus nedaug naudos. Kadangi dauguma Čekijos bibliotekų galės pasinaudoti konvertuotais Čekijoje leistų knygų įrašais, Nacionalinė biblioteka kviečia kitas bibliotekas bendradarbiauti.

Jei norime, kad "Bibliografický katalog" būtų konvertuotas kokybiškai, turime pasitelkti profesionalius darbuotojus, gerai išmanančius tarptautinius standartus. Kadangi tikimės kokybiško rezultato, iš bendradarbiaujančių bibliotekų reikalaujame:

- \* išmanymo apie UNIMARC, ISBD, AACR2 (visa tai rengiamės patikrinti),
- \* sugebėjimo identifikuoti atskirus kortelės laukus ir koduoti juos dialogo režimu,
- \* patirties dirbant su pele valdoma MS Windows sistema,

\* kruopštumo ir pastabumo smulkmenoms. Šios sąlygos, net jei atrodo labai griežtos, turėtų kiekvienai bibliotekai suteikti tokių privalumų:

- \* parengimą, konsultacijas, pagalbines literatūras,
- \* 0,8 JAV dolerio kreditą už kiekvieną įrašą,
- \* laisvą priėjimą prie šios duomenų bazės kiekvienai įstaigai, kuri sutvarkys 30 000 įrašų per trejus metus,

\* nemokamą kompiuterinę darbo vietą kiekvienai įstaiagai (jei ji finansiškai priklauso Kultūros ministerijai), kai ji sutvarkys 30 000 įrašų per trejus metus. (Pastaba: kitos įstaigos, t.y. nefinansuojamos Kultūros ministerijos, taip pat gali įgyti tokią teisę, jei sutinka su visomis retrokonversijos projekto sąlygomis),

\* į bibliotekas, kurios prisidės prie įrašų tvarkymo, atsižvelgs Kultūros ministerija, ateityje skirdama lėšas konvertavimui.

Žinoma, bibliotekos įgis neįkainojamos patirties ir susipažins su tarptautiniais standartais, tai padės lengviau patekti į tarptautines duomenų bazes.

Kelios bibliotekos jau tuo susidomėjo ir nori bendradarbiauti. Dažniausiai jos supranta, kad bendradarbiavimas padės įgyvendinti didesnį projektą, bet pirmiausia joms rūpi pertvarkyti savo katalogus. Visi turime apsvarstyti galimybes, kaip padaryti katalogą prieinamą vartotojams. Stengtis išvengti dvigubo darbo, atsirandančio, kuriant to paties leidinio įrašą, kai dirbdama keliose įstaigose. Norime maksimaliai išnaudoti turimus finansinius išteklius, personalą ir naują technologiją. Turime taikytis prie vieno "bazinio" įrašo, prie kurio kitos įstaigos galėtų pridėti savo atžymas.

Jei bibliotekos pirmiausia nori parengti savo rinkinių įrašus, į tai atsižvelgiame, bet griežtai reikalaujame kurti tuos įrašus pagal tarptautinius standartus, kad jais kiekvienas galėtų lengvai naudotis (anksčiau kai kurios bibliotekos ne visados naudodavo tas pačias katalogavimo taisykles). Jei bibliotekos nori ir savo katalogus padaryti prieinamus skaitytojams, minėtas skenavimo būdas gali joms padėti tai padaryti. Jos net gali parengti pirminę transkripciją, perduoti šiuos įrašus nacionalinei duomenų bazei redaguoti ir, baigusios kompiuterizuoti savo katalogą, ištraukti į bendradarbiavimo projektą.

## Išvada

Dabar bibliotekos gali pasinaudoti geriausio lygio technologija savo katalogams konvertuoti. Ši galinga technologija iš tikrųjų yra impulsas siekiant bendradarbiauti nacionaliniu ir tarptautiniu mastu. COMDAT savo technologijai gali pritaikyti ne tik UNIMARC, kaip kad mūsų projekte, bet ir USMARC ar bet kurią kitą MARC versiją. Atsižvelgdami į tai, kad svarbiausias tikslas yra tarptautinė standartizacija, nerekomenduojame naudoti ne MARC tipo versijos, nors tai ir įmanoma.

Didelis MS Windows programos pranašumas, kad ji gali pritaikyti ir ne lotyniškus rašmenis, tai ypač svarbu slavškiems rašmenims, nes jiems nėra pripažinta nė vieno tarptautinio transkribavimo standarto.

Svarbiausia turėti bendrą autorizotą failą, ir jei vartotojas naudodamasis atitinkamomis nuorodomis randa reikiamą pradžią, likusią įrašo dalį jis gali matyti užrašytą originalo raštu - slavišku, arabų ar paprasčiausiu lotynišku šriftu. Bendradarbiavimas visais etapais ir tarptautinių standartų pripažinimas yra vienintelis būdas sustabdyti įpras-

tinį menkų išteklių švaistymą, būdingą Vidurio ir Rytų Europos bei buvusios Sovietų Sąjungos šalims. Reikia liautis skūstis ir imtis efektyviai tvarkyti savo katalogus. Tai vienintelis būdas atverti buvusio sovietinio bloko bibliotekas. Jei kiekviena šalis, kaip Čekijos Respublika, skirs lėšų savai knygų produkcijai konvertuoti, tai bus reikšmingiausias indėlis į įvairias IFLA programas, pvz., UAP (Visuotinio spaudinių prieinamumo), UBC (Visuotinės bibliografinės apskaitos) ir kt. Kadangi turėsime efektyvų planą ir apčiuopiamus rezultatus, mūsų galimybės gauti tarptautinį finansavimą taip pat bus gerokai didesnės. Manome, jog mūsų žmonių skenavimo-užrašymo programa yra žingsnis link minėtų tikslų įgyvendinimo. Mes dar nesame šiuolaikiniame informacijos greiktelyje, bet sparčiai link jo veržiamės. Nenorėtume likti nuošalyje vien todėl, kad nebuvome pratę atverti savo išteklių. Jei mūsų katalogai bus prieinami, pasaulis supras, kad bibliotekose glūdi lobis, laukiantis tyrinėtojų bendram labui.

## COMDAT retrokonversijos technologija

COMDAT kompanija padeda bibliotekoms konvertuoti katalogus kompiuterine programine įranga. Programinę įrangą parengė COMDAT kartu su Čekijos nacionaline biblioteka.

Konversijos struktūra sudaryta iš atskirų dalių, taigi fazės vykdomos viena po kitos, todėl galima sukurti tinkamą naudoti kompiuterinį katalogą, pritaikius tik pirmąją konversijos proceso fazę. Tai padeda šiek tiek sumažinti pradines išlaidas, leidžia projektą finansuoti palaipsniui, per priimtina laiką tarpą. Išlaidos beveik 50% mažesnės nei naudojant tradicinius konvertavimo būdus. Visas darbo procesas planuojamas taip, kad galima būtų kuo efektyviau išnaudoti parengtą bibliotekos personalą ir jų darbo laiką. Procesas pradėdamas nuo katalogo kortelių, kaip jos beatrodytų, o baigiamas MARC įrašais, pvz., UNIMARC. Visas procesas susideda iš trijų dalių:

- I. Skenavimas ir archyvavimas.
- II. Skaitmeninio vaizdo perrašymas į tekstinį formatą.
- III. Tekstinio formato žymėjimas MARC turinio žymekliais.

### I. Skenavimas ir archyvavimas

Pagrindiniai pirmojo etapo uždaviniai:

- \* kuo mažiau trukdyti naudotis katalogu,
- \* katalogo kortelę paversti skaitmeniniu atvaizdu,
- \* skaitmeninį atvaizdą padaryti tokį pat aiškų kaip originalas,
- \* skaitmeninį atvaizdą perkelti į optinį diską,
- \* nuspręsti, kurį būdą pasirinkti atvaizdui perrašyti tekstinio formatu,
- \* naudoti archyvuotus įrašus kaip kompiuterinį katalogą, jei neketiname atlikti konvertavimo iki galo,
- \* sukurti bendrą įrašų leidiniui.

Dabar dokumento skenavimas ir gaunamas atvaizdas yra daug kokybiškesnis nei buvo mikrofišuojant. Kiekvieno įrašo archyvavimas skenuojant ir įrašant į magnetinį-optinį diską yra labai modernus duomenų apdorojimo ir įrašų išsaugojimo būdas.

Kadangi katalogo kortelės atrodo labai skirtingai, nes užrašytos įvairiaisiais būdais, skeneris turi būti labai kokybiškas, kad galėtų automatiškai prisitaikyti prie kortelių parametrų skenuojant.

Perkėlimas į magnetinius-optinius milžiniškos talpos diskus atliekamas taikant vadinamąją JUKE-BOX ("muzikinio automato") programą. Atvaizdas išsaugomas ne tik tam, kad jį galima būtų perdirbti kitam etapui. Tuo pat metu taikant WORM programą (vieną kartą užrašant į skaitomą atmintį), jis pasidaro visiškai apsaugotas, t.y. atvaizdo negalima sugadinti, bet galima padaryti tiek kopijų, kiek tik reikia už minimalią kainą. Galutinis archyvavimas retrokonversijai nėra toks reikšmingas, bet kai kurioms bibliotekoms jis gali pasirodyti naudingas, ypač jei jų kortelių katalogas turi istorinę vertę. Dauguma šiuo metu gaminamų magnetinių-optinių diskų turi garantiją 50-čiai metų, nors apskritai reikia manyti, kad tas laikas dvigubai ilgesnis. Nors ir negalime numatyti, kas pasikeis per 20 metų, manome, kad kiekvieną mūsų sukurtą įrašą bus galima perskaityti ir ateities priemonėmis.

Skenuojant ir archyvuojant daugiau nei 3 mln. įrašų, kiekvienam milijonui kortelių reikės vienos papildomos skenavimo įrangos. Kita vertus, norint skenuoti 500 000 ar mažiau kortelių turintį katalogą, nebūtina įrengti viso kompiuterinio tinklo. Tokiose vidutinio dydžio bibliotekose visai pakaks įrengti vieną COMDAT kompiuterinę darbo vietą su skeneriu ir išoriniu optiniu disku. Manome, kad Čekijos nacionalinės bibliotekos bendrajame kataloge esantiems 3 mln. kortelių skenuoti ir archyvuoti prireiks maždaug šešių mėnesių. Galima būtų perpus sutrumpinti šį laiką, jei skeneriu galėtų dirbti daugiau darbuotojų ir keliomis pamainomis. Kortelei skenuoti ir peržiūrėti reikia maždaug 5 sekundžių, t.y.: a) paimti kortelę iš katalogo, b) ją nuskenuoti, c) nustatyti iš atvaizdo, ar ją galima transkribuoti automatiškai, ar rankiniu būdu, d) grąžinti kortelę į katalogą. Toks trumpas laikas reikalauja, kad skeneris būtų prie pat katalogo arba bent tame pačiame pastate, kad kortelė nebūtų paimta iš katalogo ilgiau nei dešimčiai minučių. Baigus šį etapą, dirbama tik su skaitmeniniais kortelių atvaizdais, o pats katalogas nereikalingas. Taigi šitai dirbant ne tik katalogu beveik visą laiką galima naudotis, bet ir kortelės atvaizdas tuoj pat pasidaro prieinamas skaitytojui. Tai padaryti mums padėjo ARTIF ir KATALOG programos.

## COMDAT, ARTIF ir KATALOG programos

ARTIF, teksto formatavimo programa, yra pagrindinė viso skenavimo proceso dalis. Programa sudaro sąlygas dirbančiam su skeneriu pamatyti skenuojamą kortelę, nustatyti jos atvaizdo kokybę ir ją archyvuoti. Baigus šį procesą, bet kuriuo metu galima išsikviesti atvaizdus su

KATALOG programa ir operatorius gali dirbti vien tik su katalogo kortelių atvaizdais. Vienas iš svarbiausių pirmojo etapo uždavinių - nuspręsti, kuriuo būdu transkribuoti atvaizdą. Katalogo kortelei patekus į skenerį, aparatas trumpam sustoja ir operatorius, peržiūrėjęs ekrane skenuotą atvaizdą, turi nuspręsti, ar galima skenuoti automatiškai, OCR programine įranga, ar rankiniu būdu, jei atvaizdas miglotas ar neįskaitomas. Norint gerai atlikti šį darbą, reikia šiek tiek pasirengti arba turėti darbo su OCR programine įranga patirties.

Kai atvaizdai patenka į archyvą, KATALOG programa leidžia bibliotekos darbuotojams ir lankytojams juos pamatyti per viešuosius terminalus. Ši programa patraukli todėl, kad kiekvienos kortelės tuoj pat galima ieškoti, panašiai kaip ieškoma tradiciniame kortelių kataloge rankiniu būdu. Be to, galime ieškoti kortelės pagal pradžią ar pagal kokį kitą parametą, kurį galima įrašyti į optinį atvaizdą rankiniu būdu, priklausomai nuo įstaigos poreikių ir finansinių išteklių.

## II. Skaitmeninio vaizdo perrašymas į tekstinį formatą

Antrojo etapo uždavinys:

- \* transkribuoti skaitmeninį vaizdą į tekstinį formatą.
- Šios transkripcijos tikslas yra pervesti vaizdą į tekstinį formatą kuo greičiau ir kuo tiksliau. Jis dar nesiima įrašo struktūrizuoti, t. y. užrašyti žymių.

*Transkribavimas rankiniu būdu*

Norint kokybiškai transkribuoti, reikia turėti kokybišką pirminę medžiagą. Todėl darbuotojas ekrane peržiūri ne tik "šaltinį", t.y. katalogo kortelės atvaizdą, bet ir transkripcijos bandymą, atliktą pagal TEXTIF programą, kuri naudojama rankiniu būdu transkribuojant atvaizdus, kurių negalima transkribuoti įprastiniu OCR būdu.

*Transkribavimas naudojantis OCR*

Lengvai įskaitomą tekstą galima konvertuoti naudojant gerą OCR kompiuterinę programą. Patikrinę dabartinius įrašus, galime pasakyti, kad OCR konversija trunka vos kelias sekundes. Tačiau pastebėjome, kad atsižvelgiant į Čekijos nacionalinės bibliotekos bendrojo bibliotekos katalogo kortelių įvairovę ir kokybę, OCR programine įranga bus galima apdoroti tik apie 40% visų skenuotų kortelių. Dėl įvairiausių priežasčių likusias korteles teks perdirbti rankiniu būdu. Pavyzdžiui, ranka rašytas katalogo kortelės ar neryškias, nors ir perrašytas mašinėle, arba užpildytas gotišku šriftu, būtinai reikės transkribuoti rankiniu būdu. Reikia atkreipti dėmesį, kad šiame etape bibliotekos personalas nieko nenagrinėja ir netikrina.

## III. Tekstinio formato žymėjimas MARC turinio žymekliais

Trečiojo etapo uždaviniai:

- \* tekstinį formatą perdirbti į formatą su MARC žymėmis,
- \* kataloguoti ir patikrinti stilistiškai,
- \* galutinai peržiūrėti kiekvieną įrašą.

Bibliotekų specialistai pagal COMDAT Pro Tag programą parengia galutinį įrašo formatą su MARC žymėmis.

### Pro Tag programa

Programa buvo sukurta, norint padėti duomenis apdorojančiam personalui lengviau įveikti šį retrokonversijos etapą. Ji leidžia visus ekrane matomus duomenis, tiek sukurtus rankiniu būdu, tiek automatiškai, suskirstyti kategorijomis ir sugrupavus, paversti žymėtu MARC įrašu. Viso proceso metu viršutinėje ekrano dalyje matomas teksto formato įrašas. Operatorius kiekvieną segmentą, t.y. lauką, perkelia į "edit" (redaktooriaus) eilutę ir peržiūrėjęs bei aprobavęs specialiu klavišu, perkelia jį į atitinkamai pažymėtą lauką - "kirpk ir klijuok" principu. Šis metodas todėl patrauklus, kad operatoriui nereikia perrašinėti laukų, jis tiesiog kilnoja juos iš vienos vietos į kitą. Automatiškai konvertuoti tuo lengviau, kuo tinkamesnis atitinkamas OCR metodu apdoroto įrašo algoritmas. Todėl konvertuoti ISBD įrašą arba bet kurią kitą iš anksto nustatytą duomenų seką yra labai paprasta.

### Išlaidos

Įranga kainuoja apie 40 000 JAV dolerių. Maždaug tiek gali kainuoti bibliotekai, kurios kataloge reikia apdoroti apie 500 000 kortelių. Taigi vienos tipinės katalogo kortelės perdėrimas kainuoja apie dolerį, o sudėtingesnės kortelės kaina gali pakilti iki 1,6 JAV dolerio, bet ir tokia

kaina yra daug mažesnė negu įprastinė rinkos kaina (į kainą įtrauktos kompiuterinės technikos ir programinės įrangos kainos). Tačiau didesnėms įstaigoms, tokioms kaip Čekijos nacionalinė biblioteka, vien įrengimai kainuos gerokai daugiau: apie 150 000 - 190 000 JAV dolerių (taigi 3 mln. kortelių konversija turėtų kainuoti apie 3 mln. JAV dolerių). Kainos kilimą diktuoja ne tik didesnio skenerių kiekio poreikis, kad galima būtų kuo greičiau baigti pirmąjį etapą, bet visų pirma galingesnio "JUKE-BOX", pajėgiančio sukaupti ir apdoroti žymiai daugiau įrašų, kaina.

Mažesnėms bibliotekoms, kurjoms ir 40 000 JAV dolerių gali būti per brangu, geriau būtų sudaryti skenavimo ir OCR konversijos sutartį su COMDAT, kurio vykdytojai rūpintųsi konversija. Ateityje du pirmuosius etapus galėtų atlikti nebūtinai COMDAT kompanija. Tai galėtų būti bet kuri firma, geriau COMDAT parengta, kuri galėtų pagal sutartį dirbti atskiroms bibliotekoms. Jei būtų sudaryta kompanija, kurios darbuotojai būtų pasirengę bibliotekiniam darbui, galima būtų sudaryti sutartį visam procesui, nors veikiausiai kompanijos būtų prašoma atlikti tik pirmojo etapo darbus (skenuoti katalogą).

Iš viso to galima spręsti, kad tik paskutiniame etape prireiks kvalifikuotų bibliotekos darbuotojų, o pirmuosiuose dviejuose etapuose, nors ir jiems reikia bibliotekos paramos ir pasirengimo, didžioji dalis katalogo darbuotojų visai nedalyvauja. Darbas pagal sutartis turi dar ir tokį privalumą, kad visam projektui vadovauja, jį rengia ir prižiūri samdytas personalas, o bibliotekos darbuotojai gali dirbti savo įprastus darbus.

Iš anglų k. vertė N. Bulotaitė

UDK 022.9+025.3/6:681.3

## Nacionalinės ir universiteto bibliotekos retųjų knygų kolekcijos retrospektyvusis katalogavimas

Tinka KATIĆ

Nacionalinė ir universiteto biblioteka, Zagrebas, Kroatija

### 1. Istorija

Retųjų spaudinių kolekcijos istorija - tai Nacionalinės ir universiteto bibliotekos istorija. Biblioteka buvo įkurta 1606 m. kaip Collegium Societatis Iesu biblioteka. Ilgainiui, joje buvo sukaupta daug vertingų knygų ir rankraščių priklausiusių įžymiesiems kroatams ir įvairioms įstaigoms. XIX a. pabaigoje buvo įsigytos ir papildė kolekciją dvi labai reikšmingos privačios Ljudevito Gajaus ir Nikolo Zrinskio bibliotekos. Šis pirkinys paskatino tuometinį bibliotekos direktorių Ivaną Kostrenčičą paprašyti Vyriausybės leidimo atidaryti skaityklą, kurioje būtų kaupiamos retosios knygos ir rankraščiai. Kostrenčičo prašymas buvo patenkintas 1894 m. birželio 24 d., ir nuo tada kolekcija tapo atskiru padaliniu.

Kolekcijoje sukauptas didžiausias senųjų kroatų knygų nuo XV a. iki 1835 m. ir kitų tautų bei kultūrų senų retųjų knygų rinkinys. Kolekciją sudaro 9236 knygos, suskirstytos į 11 grupių. Zrinskių ir Kuševičių šeimų bibliotekos traktuojamos kaip atskiros kolekcijos.

### 2. Automatizacija

Kolekcijos knygos sukataloguotos ir vartotojai gali jas susirasti įvairiuose korteliniuose kataloguose: kartografiniame, abėcėliniame ir sisteminiame. Kataloguose sukaupta detali informacija ir jie atlieka savo pagrindines funkcijas.

Pastaraisiais metais padidėjęs susidomėjimas nacionaline kultūra ir palikimu atvėrė Retųjų knygų kolekciją naujos rūšies vartotojams. Šių vartotojų poreikiams tenkinti reikalinga informacija turi būti tiesiogiai ir lengvai prieinama.

Detaliau panagrinėsime bibliografinio įrašo funkcijas. Šiuo tikslu apibrėžėme bibliografinį įrašą, kuris turi atitikti pagrindines abėcėlinio katalogo funkcijas ir specifines atitinkamos medžiagos rūšies funkcijas. Tos specifinės funkcijos sujungia į vieną vietą bibliografinius įrašus, kurie yra:

- susiję su konkrečiu asmeniu,
- išleisti/išspausdinti tam tikroje vietoje arba tais pačiais metais ar tuo pačiu laikotarpiu,

- to paties fizinio ar intelektualinio žanro,
- panašūs fiziniiais požymiais,
- atskleidžiantys tą patį šaltinį.

1990 m. biblioteka idiegė naują informacinę sistemą CROLIST. Ši integrali sistema parengta Nacionalinės ir universiteto bibliotekos bei Kroatijos informacinio tinklo vartotojams. Ji pagrįsta UNIX operacine sistema ir ORACLE duomenų bazių valdymo sistema. CROLIST sudaro sąlygas naudotis autorizuotais įrašais ir kurti naujus UNIMARC laukus antikvarinei medžiagai<sup>1</sup>.

Duomenų paieškai ir atrankai bei perdavimui CROLIST sistemos retųjų knygų kompiuteriniais įrašais gali naudotis bibliotekos skaitytojai, Kroatijos bibliotekų ir informacinio tinklo vartotojai, užsienio vartotojai per Internet.

Adresas:

|                         |                      |
|-------------------------|----------------------|
| nsb2.nsb.hr 161.53.80.4 | X.25                 |
| login: opac1            | arba 621914110000303 |
| password: <return>      | login: opac1         |
|                         | password: <return>   |

### 3. Retrospektyvusis katalogavimas

Pirmasis retrospektyviojo katalogavimo tikslas - suvestinis Kroatijos bibliotekose esančių kroatų knygų ir senų knygų katalogas. Retųjų knygų kolekcija kataloguojama pagal chronologiją. Pagaliau šis katalogas taps pagrindine įrašų baze kitoms bibliotekoms, norinčioms kataloguoti savo fondus.

Antrasis tikslas - rankinio spausdinimo epochos nacionalinė bibliografija. Suvestinio katalogo įrašų forma yra nacionalinės bibliografijos branduolys, kurį būtina papildyti kitų šaltinių duomenimis.

Trečiasis tikslas - suvestinis Kroatijos bibliotekose esančių užsienio retųjų ir senų knygų katalogas. Jis bus formuojamas tokiais pat principais, kaip ir suvestinis Kroatijos bibliotekose esančių kroatų knygų ir senų knygų katalogas.

Ketvirtasis tikslas - keitimasis įrašais ir bendradarbiavimas įvairiuose tarptautiniuose projektuose:

- Europos mokslinių bibliotekų konsorciumo duomenų bazė - Europos retųjų knygų duomenų bazė<sup>2</sup>, ir