

NÁRODNÍ PROGRAM RETROSPEKTIVNÍ KONVERZE ČR

(Historie a současný stav, priority, finanční a časové nároky, spolupráce)

Bohdana Stoklasová

Úvod

V případě retrospektivní konverze katalogů českých knihoven můžeme bez nadsázky hovořit o národním programu. Díky iniciativě Národní knihovny ČR, podpoře Ministerstva kultury ČR, existenci originální špičkové technologie vyvinuté českou firmou Comdat a pozitivnímu postoji českých knihoven ke spolupráci přestává být retrospektivní konverze katalogů záležitostí jednotlivých knihoven a stává se programem společným, národním. Jedině tento přístup dává českým knihovnám naději na zvládnutí důležitého, ale finančně, časově i personálně nesmírně náročného úkolu v prostředí, kde všechny zmíněné zdroje jsou silně limitovány.

Příspěvky zástupců Ministerstva kultury ČR na této konferenci vytvářejí obecný rámec, z něhož vychází tento příspěvek. Technologie vyvinutá pro národní program retrospektivní konverze katalogů českých knihoven je zde popsána jen velmi obecně, detailní popis i potřebné technické vybavení jsou obsahem příspěvků zástupců firmy COMDAT.

Historie a současný stav

NK ČR byla ještě v minulém roce v oblasti retrospektivní konverze svých katalogů v těžké situaci. V průběhu posledních let zde byly analyzovány ve světě nejrozšířenější metody retrospektivní konverze, kterými jsou metoda **OCLC** a **SAZTEC**. Procento záznamů obsažených v generálním katalogu univerzálního knihovního fondu (dále jen GK UKF), které byly nalezeny v bázi OCLC, bylo nízké. Metoda bude vhodná pro určité části katalogu, nikoli pro katalog jako celek. Metoda SAZTEC se jevila jako vhodnější. Při částce 60 Kč za záznam by však retrospektivní konverze GK UKF v rozsahu 3.000.000 lístků stála 180.000.000 Kč. NK ČR neměla k dispozici ani zlomek této částky a neexistovala reálná naděje na její získání. Retrospektivní konverze katalogu byla zahájena v roce 1992 metodou pro NK ČR finančně přijatelnou – lístky byly připravovány v NK ČR, přepisovány ve **Vodních**

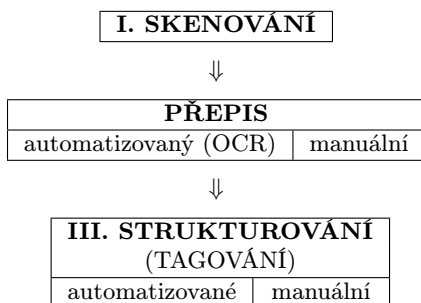
stavbách Tábor a kontrolovány opět v NK ČR. Na NK ČR zůstávaly stále dvě třetiny práce, poměr se ani po letech (navzdory očekávání) nezlepšil. Práce postupovala pomalu, části katalogu byly po dlouhou dobu vyraženy z provozu. Spolupráce s Vodními stavbami Tábor byla ukončena v červnu 1995. **Všem jmenovaným metodám je společné jedno: katalog lze ve strojem čitelné podobě zpřístupnit až po dokončení celé akce, což zároveň znamená vyplacení celé částky.** S ohledem na silně limitované finanční zdroje nelze počítat se získáním tak vysoké částky dříve než v horizontu zhruba deseti let. „Úzké hrdlo“ představuje u všech zmíněných technologií obrovský objem lidské práce, kterou je třeba v souvislosti s retrospektivní konverzí 3.000.000 lístků vykonat a pochopitelně zaplatit. Nebyla naděje na získání finančních prostředků na práci najednou, zřizovatel přislíbil příspěvek po menších částech (řádově několik miliónů – v žádném případě desítek či stovek miliónů/rok). Existoval přislíb grantu od Mellonovy nadace – ovšem nikoli na práci, ale na zakoupení technického vybavení. S blížícím se dokončením výstavby depozitáře NK ČR v Hostivaři nabyla na aktuálnosti nutnost přestěhovat GK UKF z Klementina do Hostivaře. Podmínkou pro jeho odstěhování byla existence provozně využitelné (nejen čistě archivní) kopie na náhradním nosiči. Přitom naděje na pořízení mikrofilmové kopie zdarma se rozplynuly v důsledku nedostatku zájemců o její zakoupení. Bylo nezbytně nutné odstěhovat katalog a zpřístupnit ho na náhradním médiu, které zabere minimum prostoru. Jedinou nadějí na získání finančních prostředků byl přislíb Mellonovy nadace – grant omezený na nákup základní techniky. **Bylo nutné najít metodu, která umožní ušetřit lidskou práci a nahradit ji levnější prací strojovou a zároveň přesunout úzké hrdlo, které představuje kvalifikovaná lidská práce, jíž se nelze vyhnout, na konec, za fázi zpřístupnění katalogu ve strojem čitelné podobě.**

Z této složité situace a zmíněných limitů vzešla **originální technologie pro retrospektivní konverzi katalogů. Technologie je stavebnicová a univerzální, využitelná v kterékoli české nebo zahraniční knihovně jako celek nebo po jednotlivých modulech.** Je výsledkem analyticko-projekčních zkušeností pracovníků NK ČR (funkční zadání, testování, připomínky), zkušeností se skenováním katalogu v knihovně AV ČR, finanční podpory Mellonovy nadace a pomoci prof. Andrew Lasse při jejím získání a především zkušeností firmy COMDAT se zpracováním a využitím textových i obrazových dat a ochoty vyvinout na vlastní náklady komplexní technologii pro retrospektivní konverzi katalogů. Téměř roční vývoj technologie představoval pro firmu značné riziko ve smyslu návratnosti značného objemu vložených prostředků a energie. Technologie je natolik revoluční, že bylo těžké odhadnout, zda ji knihovny přijmou. V případě NK ČR není stále

jasné, zda si vzhledem ke svému omezenému rozpočtu bude moci plnou realizaci projektu dovolit. S finančními zdroji jsou na tom ostatně podobně i další české knihovny. Většina z nich se svým přiškrceným rozpočtem těžko zvládá nejnужnější provoz. Budou mít vůbec prostředky a síly vrátet se do minulosti? **Retrospektivní konverze ale není jen náročnou a obtížnou cestou do minulosti. Je nezbytnou podmínkou zpřístupnění fondů a efektivní spolupráce knihoven v přítomnosti a budoucnosti i jejich plnohodnotnému připojení ke stále častěji skloňovaným informačním superdálnicím.**

Technologie

Technologie bude detailně popsána v příspěvku firmy COMDAT. Pro účely tohoto textu postačí velmi hrubé schéma a charakteristika jednotlivých kroků:



I. SKENOVÁNÍ

V první fázi jsou lístky (nebo tištěné bibliografie) naskenovány špičkovými skenery, které zaručují kvalitní výsledek nezbytný pro další zpracování. Díky speciálním programům je možné katalogy využívat již po naskenování. K záznamům je možné doplnit do pomocné báze dat záhlaví, popřípadě další soupisné a popisné údaje (rozsah pomocné báze závisí na velikosti katalogu, personálně-finančních možnostech knihovny a prioritách – časových horizontech).

II. PŘEPIS

Ve druhé fázi jsou lístky převedeny z obrazové podoby (soubory TIFF) do nestrukturované textové (ASCII) podoby. Přepis probíhá u kvalitních záznamů automaticky (OCR), u méně kvalitních manuálně.

III. STRUKTUROVÁNÍ (TAGOVÁNÍ)

Ve třetí (závěrečné) fázi jsou nestrukturované textové soubory převáděny do

podoby strukturované (UNIMARC) pomocí speciálního programu. V závislosti na kvalitě zdrojového materiálu probíhá strukturování buď automatizovaně (je-li možné převod spolehlivě algoritmizovat) nebo manuálně. Speciální programové vybavení poskytuje maximální komfort, šetří kvalifikovanou lidskou práci a umožňuje dosahovat vysoké rychlosti i kvality.

Priority

GK UKF obsahuje lístky různé kvality, jejichž retrospektivní konverze bude různě náročná. Podstatnou část katalogu tvoří **záznamy české knižní produkce**. Retrospektivní konverzi těchto záznamů je třeba provést v první fázi. Pro zapojení do mezinárodní spolupráce bychom měli mít v pořádku vlastní produkci. Dalším důvodem je skutečnost, že české knihy tvoří největší část fondů většiny českých knihoven, které rovněž plánují (nebo již zahájily) retrospektivní konverzi svých katalogů. Hrozí zde zbytečná multiplicitní práce a plýtvání personálními i finančními zdroji, které by měly být využity racionálněji. Po české produkci bude následovat produkce zahraniční, postup bude od jednodušších záznamů ke složitějším.

Česká knižní produkce obsažená v konzervačním fondu NK ČR je od roku 1922 zachycena v Bibliografickém katalogu. Od roku 1983 existuje ve strojem čitelné podobě. Jedná se tedy o převod zhruba 400 tisíc záznamů z let 1922-1982. Přidáme-li počátek století, celkový objem bude zhruba půl miliónu záznamů. Je samozřejmé, že se nejedná o úplné zachycení české knižní produkce 20. století, ale o snadno dostupný základ, na němž je možné dále (v úzké kooperaci s dalšími českými knihovnami) stavět.

Při převodu se bude pracovat téměř výhradně se soubory vytvořenými automatickým přepisem pomocí OCR, možnost algoritmicke pro strukturování se v jednotlivých časových vrstvách výrazně liší.

GK UKF obsahuje 3 milióny lístků, z nichž pouze část bude možné přepsat pomocí OCR, možnost algoritmicke závisí, podobně jako u Bibliografického katalogu, na použitých pravidlech popisu. V porovnání s Bibliografickým katalogem, kde díky chronologii vydávání určitá časová vrstva odpovídá určitým pravidlům, je u GK UKF možnost algoritmicke podstatně snížena prolínáním záznamů z různých časových vrstev vytvořených na základě různých pravidel v rámci jednotlivých písmen abecedy.

Finanční a časové nároky

GK UKF bude zpřístupněn po naskenování (zhruba za 1 rok) v jednoduché podobě, v úplné (UNIMARCové) podobě zhruba do 10 let. Odhad vychází z objemu finančních prostředků (a počtu pracovníků, s nimiž lze pro akci v dnešní době reálně počítat). Bude-li k dispozici více finančních prostředků (a pracovníků ochotných se na akci podílet), celková doba se zkrátí.

Česká knižní produkce 20. století by měla být zpřístupněna v úplné podobě do roku 2.000. Pro odhad finančních a personálních možností platí totéž, co pro GK UKF. Zvýšení objemu dostupných finančních prostředků a počtu pracovníků bude znamenat i zde urychlení akce.

Spolupráce

Především záznamy české knižní produkce 20. století budou v českých knihovnách široce využitelné. Proto se na ně NK ČR obrací se žádostí o spolupráci.

Nároky jsou (s ohledem na potřebnou vysokou kvalitu výsledného produktu) značné:

- znalost standardů (UNIMARC, ISBD, AACR2) – bude prověřeno testem
- schopnost strukturovat záznam (rozeznat v záznamu jeho jednotlivé části a přidělit jim ekvivalentní pole/podpole v UNIMARCu, popřípadě je modifikovat podle doporučení ISBD/AACR2)
- schopnost práce se speciálním programem (pod MS Windows, nutnost práce s myší)
- pečlivost, přesnost

Zajímavá nabídka Národní knihovny ČR a Ministerstva kultury ČR odpovídá vysokým nárokům:

- zaškolení, příručky, konzultace
- 15-20 Kč/záznam (s ohledem na nutnost odečtení daně, sociálního a zdravotního pojištění z celkového příjmu se nejedná o čistý výdělek!)
- pro knihovny, které otagují minimálně 30 tisíc záznamů do tří let celý výsledek retrospektivní konverze české knižní produkce 20. století zdarma
- pro knihovny, které otagují minimálně 30 tisíc záznamů do tří let a jejichž zřizovatelem je Ministerstvo kultury ČR, kompletně vybavená pracovní stanice, kterou si budou moci po ukončení akce ponechat. Rozšíření této podmínky i na knihovny spadající pod jiné zřizovatele závisí na přijetí/nepřijetí vládního programu.
- knihovny, které se zúčastní kooperace, budou mít prioritu při přidělování finančních prostředků na další aktivity související s retrospektivní konverzí (skenování vlastních katalogů atp.). Toto je prozatím strategie Ministerstva kultury – věříme, že osloví i zřizovatele ostatních knihoven.

Podmínky kooperace byly vyhlášeny na semináři o retrospektivních konverzích katalogů v Hradci Králové v červnu 1995. Většina knihoven v resortu Ministerstva kultury ČR se již ke spolupráci přihlásila. Z jejich dopisů vyplývá, že chápou výhodnost i nutnost spolupráce při retrospektivní konverzi katalogů českých knihoven a šanci využít optimálním způsobem omezené finanční i personální zdroje, které jsou u nás pro retrospektivní konverzi k dispozici.

Většina knihoven chce však co nejdříve dát čtenářům k dispozici vlastní katalog, čekání na výsledky spolupráce jim připadá zdlouhavé. **Je nutné zvážit, v jaké podobě čtenářům katalog zpřístupnit, je třeba usilovat o omezené multiplicitní práci, které zbytečně odčerpávají již tak omezené finanční i personální zdroje. Nová technologie díky své modulárnosti může poskytnout řešení, pokud ji dokážeme plně využít. K největšímu plýtvání penězi i energií dochází tam, kde jsou multiplicitně vytvářeny záznamy na tentýž dokument. Čím rozsáhlejší záznamy vytvářené bez ohledu na jednotná pravidla a možnost jejich využití v ostatních knihovnách, tím větší plýtvání v nedostatku. Optimálním modelem je spolupráce, společné vytvoření „jádra“ (podle jednotných pravidel), jeho obohacení o vlastní lokační údaje a záznamy, které „jádro“ přesahují.**

Existuje celá řada variant spolupráce při využití popsané technologie. Volba varianty záleží na tom, zda určitá knihovna s retrospektivní konverzí svého katalogu již začala, či zda se na ni teprve chystá. Naskenování katalogu a jeho doplnění základními údaji umožní čtenářům katalog využívat. Až se naše úzké a klikaté informační pěšinky změní v informační superdálnice, bude katalog v této podobě přístupný i mimo knihovnu. Uložení základních údajů není časově, finančně ani personálně příliš náročné, propojení těchto základních údajů s naskenovaným katalogizačním lístkem přináší čtenářům dostatek informací. Postupně lze základní údaje nahradit úplným katalogizačním záznamem vytvářeným pouze jednou v rámci kooperace knihoven (nejen českých).

Závěr

Nedostatek financí, času a lidských zdrojů, technické zkušenosti, které získali čeští knihovníci v začátcích automatizace, kdy si knihovny nemohly dovolit nákup specializovaných knihovnických systémů a knihovníci museli univerzální systémy dostupné zdarma nebo za minimální cenu sami upravit pro využití v knihovnách, i nápady a znalosti nejnovějších technologií pracovníků firmy COMDAT, byly základními stavebními kameny moderní, špičkové technologie, která nám získala mezinárodní respekt. Její modulární charakter a jednoduché využití jsou dobrým základem pro národní program retrospektivní konverze katalogů českých knihoven. Ochota knihoven ke spo-

lupráci, koordinační a metodické snahy Národní knihovny ČR a v neposlední řadě podpora Ministerstva kultury ČR jsou příslibem do budoucna. Přijme-li vláda program, o němž informovali zástupci Ministerstva kultury ČR, budou k dispozici i finanční prostředky nezbytné na uvedení národního programu retrospektivní konverze katalogů českých knihoven v život (a jeho udržení při životě v horizontu několika následujících let). Úvodní fáze národního programu retrospektivní konverze katalogů českých knihoven je hodnocena (nejen doma, ale i v zahraničí) jako modelová ukázka spolupráce zřizovatele, národní knihovny a ostatních knihoven, fundraisingu a vytvoření i využití špičkové domácí technologie, o kterou jeví zájem i knihovny zahraniční. Doufáme, že schválení vládního programu a zapojení knihoven mimo resort Ministerstva kultury ČR přispějí k tomu, že nejen technologie, ale i způsob provedení celé akce budou sloužit jako model pro ostatní země.