

Bohdana Stoklasová - Jaroslava Jeřábková
Národní knihovna ČR

Úvod

Na konci 20. století vydává Národní knihovna České republiky CD-ROM zachycující ve formě bibliografických záznamů většinu knižní produkce vydané v Česku (a část knižní produkce vydané na Slovensku) v průběhu celého 20. století. Záznamy jsou přístupné také na internetu. Převod 400 000 záznamů z tištěné do digitální podoby (včetně jejich částečné úpravy podle mezinárodních standardů) je výstupem projektu známého pod názvem „Zpřístupnění záznamů české knižní produkce 20. století prostřednictvím Internetu a CD-ROM“ řešeného v letech 1997-2000 v rámci grantového programu pro výzkum a vývoj.

Obsah báze dat na CD-ROM

V průběhu řešení projektu byly převedeny do digitální podoby tištěné bibliografické soupisy za léta 1901-1982. Od roku 1983 již probíhalo zpracování České národní bibliografie automatizovaně a tato časová vrstva nebyla předmětem řešení projektu.

Báze dat na CD-ROM však pokrývá celé 20. století. Následující text je nezbytnou pomůckou pro orientaci v bázi dat vzniklé spojením několika velmi odlišných časových vrstev a zároveň ilustruje vývoj české národní bibliografie v průběhu 20. století.

Časové vrstvy - charakteristika: 1983-2000

Záznamy tištěných monografií od roku 1983 již vznikaly v automatizované podobě, převod z tištěné do digitální podoby (retrospektivní konverze) tedy nebyl nutný. Záznamy byly zpracovány na základě povinného výtisku, nebyla zahrnuta firemní literatura a drobné tisky. V důsledku neujasněného vztahu monografie-seriál obsahuje báze dat i menší počet záznamů seriálů.

Záznamy dokumentů byly v letech 1983-2000 vytvářeny v různých automatizovaných systémech, což se projevuje zřetelnými rozdíly ve výsledné podobě záznamů různých vrstev báze dat. Záznamy z let 1983-1992 byly zpracovány v systému ASN BK (Automatizovaný systém národní bibliografie - knihy), později byly převedeny do systémů CDS/ISIS a ALEPH, neboli do Výměnného formátu a následně do formátu UNIMARC; záznamy z let 1993-1995 (srpen) byly zpracovávány v systému CDS/ISIS ve Výměnném formátu a převedeny do formátu UNIMARC.

Převody mezi různými formáty mají vždy negativní vliv na výslednou kvalitu záznamů. Záznamy z období 1995 (září) – 2000 již byly vytvářeny přímo ve formátu UNIMARC v systému ALEPH. Vzhledem k požadavku na co nejúplnější pokrytí české knižní produkce obsahuje báze dat na konci devadesátých let i méně úplné záznamy vznik-

lé na základě programu Kooperativní katalogizace: české knihy, na kterém se podílí Moravská zemská knihovna a všechny státní vědecké knihovny.

Rozdíly v záznamech z těchto časových vrstev jsou ovšem také důsledkem přirozeného vývoje bibliografického popisu, změn pravidel používaných pro jmenný popis i vývoje používaného systematického a předmětového selekčního jazyka. Prvním mezinárodním normativem pro jmenné zpracování bylo od počátku devadesátých let mezinárodní doporučení ISBD (International Standard Bibliographic Description), od poloviny roku 1995 navíc AACR2 (Anglo-American Cataloging Rules).

Pro věcné zpracování dokumentů se do roku 1989 používaly úplné notace MDT (Mezinárodní desetinné třídění), které byly pro potřeby CD-ROM upraveny tak, aby bylo možné záznamy vyhledávat podle jednotlivých hlavních znaků. Později dochází k přechodu na hrubší vybrané znaky MDT. Od roku 1999 se používají notace MDT v rozsahu MRF (Master Reference File).

Záznamy jsou opatřeny předmětovými hesly, jejichž charakter se v průběhu let měnil v návaznosti na postupnou aplikaci mezinárodních standardů v této oblasti. V letech 1993-1995 se v záznamech vyskytovala volně tvořená klíčová slova, od roku 1996 byla jejich aplikace silně omezena. U předmětových hesel do roku 1996 zůstala zachována substantivní inverze, od roku 1996 se předmětová hesla zapisují v přirozeném pořádku slov. Na konci devadesátých let po aplikaci principů LCSH (Library of Congress Subject Headings) se charakter předmětových hesel podstatně změnil.

1945-1982

V bibliografických soupisech za roky 1945-1982 byla uvedena produkce, která byla do Národní knihovny České republiky (dále NK) dodána jako povinný výtisk (Národní knihovna ČR je současný název, název instituce se několikrát změnil v průběhu let, pro přehlednost je v textu používána zkratka současného názvu). V Bibliografickém katalogu (dále BK) se neuváděla produkce, která vyšla dříve než 2 roky před zpracováním daného ročníku. BK obsahuje českou produkci včetně jazykových mutací, v malé míře slovenika a zahraniční bohemika. Do retrospektivní konverze byly zařazeny monografie a nepravá periodika, vynechány byly normy, hudebniny, mapy a pravá periodika.

Vzhledem k tomu, že tyto dokumenty byly v BK popsány podle stejné metodiky jako knihy, nelze jejich záznamy bez primárního dokumentu v ruce převést do podoby vyhovující mezinárodním standardům (bude nutné doplnění údajů). Záznamy převzaté z těchto ročníků jsou velmi podrobné, opatřené anotacemi, předmětovými hesly, třídnicí MDT a signaturami národního konzervačního fondu NK.

1929-1944

Vrstva 1929-1944 obsahovala české (v některých ročnících i slovenské) spisy, německé, maďarské, slovanské, latinské, románské a další jinojazyčné spisy, tedy vlastně zahraniční bohemika v nejširším slova smyslu (u nás, o nás, od nás). Opět platí, že byly vypuštěny normy, hu-

debniny, mapy a pravá periodika. Záznamy jsou stručnější, bez anotací, předmětových hesel, třídniců MDT a signatur. Signatury národního konzervačního fondu NK ČR budou postupně doplněny.

1926-1928

Soupisy z let 1926-1928 obsahovaly převážně „spisy československé“, v malé míře spisy cizojazyčné a časopisy. Záznam o časopisu, almanachu, kalendáři byl zpracován jako monografický na základě 1. uvedeného čísla. Ze zpracování byly vypuštěny hudebniny. Záznamy jsou stručné, bez předmětových hesel, třídniců MDT a signatur.

1901-1925

Za základ pro zpracování produkce z počátku století byl zvolen ucelený Soupis československé literatury za léta 1901-1925 (K. Nosovský, V. Pražák). Soupis obsahuje české a slovenské spisy, v rámci Soupisu jsou zařazeny i almanachy, ročenky, kalendáře, v menší míře i mapy. Soupis byl převeden do digitální podoby jako celek. Protože jde v podstatě o nakladatelský soupis, záznamy jsou mnohdy velmi stručné, bez předmětových hesel, třídniců MDT a signatur. Konvertované záznamy jsou porovnávány na úplnost pokrytí s ročníky České bibliografie (1920, 1921 J. Gotthard; 1902-1911 Z. V. Tobolka).

Produkce, která byla získána v rámci retrospektivního doplňování z jiných zdrojů než povinným výtiskem apod., jinými slovy, celá produkce uložená v národním konzervačním fondu NK, bude do báze české knižní produkce 20. století postupně doplněna. Dále bude báze obohacena o českou knižní produkci z dalších fondů NK a z fondů ostatních českých knihoven. Intenzivní obohacování báze z těchto zdrojů bude probíhat v několika příštích letech. Velký objem dodatečné práce nás čeká také na sjednocení selekčních prvků (jmen autorů, korporací, názvů knih, názvů edic atd., případně i vybraných věcných selekčních prvků). Jde o tzv. problematiku autorit, která je poměrně mladá a její efektivní řešení umožňují až nejmodernější knihovnické systémy aplikované v posledních letech. Báze dat obsahuje záznamy celého dvacátého století, tedy převážně z doby, kdy nástroje pro kontrolu autorit byly velmi slabé.

Historie projektu

Projekt „Zpřístupnění záznamů české knižní produkce 20. století prostřednictvím Internetu a CD-ROM“ byl oficiálně zahájen v roce 1997, jeho historie však začíná již na počátku roku 1994, kdy NK na základě provedených analýz připravila funkční zadání pro komplexní technologii retrospektivní konverze katalogů a bibliografických soupisů a stanovila priority pro jejich postupný převod. Hlavní prioritou se stala bohemikální produkce.

V roce 1994 proběhla soutěž na realizaci technologie s mezinárodní účastí, ze které vyšla vítězně česká firma Comdat. V roce 1995 a 1996 pak nastala kuriózní situace: měli jsme k dispozici technologii špičkové světové úrovně, na kterou se jezdily dívat zahraniční návštěvy, dvakrát jsme byli dokonce požádáni o uspořádání mezinárodního semináře o problematice retrospektivní konverze, naše tech-

nologie se začala prakticky uplatňovat (ve více či méně modifikované podobě) v zahraničí, u nás však došlo k jejímu praktickému nasazení až za několik let. Důvod byl prostý: absence finančních zdrojů, která dlouho nedovolovala překročit linii testování. Proto je třeba vyzdvihnout jako klíčový moment podporu Mellonovy nadace v roce 1995.

V závěru roku 1994 navštívil Národní knihovnu prof. Andrew Lass, kterého velmi zaujal náš přístup k retrospektivní konverzi a předložil Mellonově nadaci projekt s názvem RETROCON. Projekt byl přijat a díky němu mohla Národní knihovna nakoupit finančně velmi náročné technické vybavení pro zahájení prací. Stále však chyběly finanční prostředky na zaplacení lidských zdrojů, bez nichž se retrospektivní konverze ani při maximálním využití špičkové techniky neobejde. NK se snažila projekt dotovat ze svého (v oblasti provozních prostředků stále se tenčícího) rozpočtu, Ministerstvo kultury se snažilo přispět účelovými dotacemi. Přes veškeré snahy tohoto druhu stále nebylo možné hovořit o dohledném časovém horizontu pro ukončení projektu. Dalším mezníkem bylo schválení projektu výzkumu a vývoje „Zpřístupnění záznamů české knižní produkce 20. století prostřednictvím Internetu a CD-ROM“, který zaručil potřebné finanční prostředky pro roky 1997-2000, umožnil zahájení spolupráce s velkými českými knihovnami, stanovení přesného harmonogramu a rozdělení prací.

Spolupráce deseti velkých českých knihoven

NK by nebyla schopna vlastními silami zpracovat během několika let 400 000 rozsáhlých bibliografických záznamů, proto se obrátila s žádostí o pomoc na Moravskou zemskou knihovnu (dále MZK) a všechny státní vědecké knihovny (dále SVK), tedy knihovny, v nichž je tradičně vysoká úroveň katalogizace a mají univerzální skladbu fondů, takže se dalo předpokládat, že výsledek projektu v budoucnu výrazně usnadní jejich vlastní retrospektivní konverzi. Spolupráce na společném projektu totiž znamenala dočasné zbrzdění jejich vlastní retrospektivní konverze nebo její omezení na tvorbu zkrácených záznamů, jinými slovy podřízení vlastních cílů cíli společnému s tím, že omezením multiplicitních převodů těchto záznamů v jednotlivých knihovnách se uspoří značné množství peněz, na druhé straně ovšem časový posun převodu vlastních katalogů poněkud zkomplikuje lokální situaci zejména v souvislosti s nasazením automatizovaného výpůjčního systému. Všechny oslovené knihovny bez výjimky se do spolupráce zapojily a mají tak významný podíl na konečném výsledku. V mezinárodním kontextu je kromě použité technologie velmi oceňována právě kooperační rovina projektu.

V poloprovozním ověřování technologie retrospektivní konverze zpracoval tým 7 pracovníků z NK ročníky 1980-1982. V roce 1997 byla zahájena spolupráce s MZK a všemi SVK, v konečné fázi na projektu pracoval tým 27 vyškolených pracovníků.

Na počátku projektu byla ověřena technologie RETROKON. Postupně vznikala velmi komplikovaná pravidla pro převod, která musela pokrýt převod záznamů vytvořených podle několika zcela odlišných pravidel a zaručit maxi-