

SDÍLENÁ KATALOGIZACE V PROSTŘEDÍ CASLIN - SOUBORNÉHO KATALOGU ČR

Zuzana Hájková, Státní vědecká knihovna České Budějovice

Eva Svobodová, Národní knihovna ČR

Od 1.1. 2000 je na Internetu zpřístupněn CASLIN - Souborný katalog ČR - knihy pod systémem ORACLE. Každému uživateli, který vstoupí na adresu www.caslin.cz je umožněno vyhledávat v bázi Souborného katalogu ČR - knihy (dále jen SK) a výsledky dotazů ukládat do schránky nebo vytisknout. Možné je též získat základní informace o vlastníkovi nalezeného dokumentu a propojit se na jeho www stránky, popřípadě odeslat zprávu na jeho e-mailovou adresu.

Pokud má uživatel zájem využívat další funkce souborného katalogu, ke kterým patří

- stahování (export) nalezených záznamů
- připisování lokační značky k nalezeným záznamům
- editace záznamů, ke kterým má uživatel vlastnické právo
- vytváření nových záznamů uvnitř souborného katalogu

je třeba, aby se stal registrovaným uživatelem, tj. aby mu bylo správcem SK přiděleno: **uživatelské jméno a heslo, kategorie**, která ho opravňuje k využívání konkrétního spektra služeb, a pokud se rozhodne dodávat do katalogu svoje data také **váha**, která vyjadřuje kvalitu dodávaných záznamů a rozhoduje v případě, že do katalogu přijdou současně dva stejné záznamy, který bude do katalogu přijat a ze kterého bude převzata pouze lokační značka (popř. signatura) vlastníka dokumentu. Uživatel též získává přístup na FTP server SK.

Váha se uživateli přiděluje na základě analýzy vzorku jeho dat. Původně se tato analýza prováděla pouze hlavoručně, v tuto chvíli je k ní možno již využít řadu automatizovaných formálně logických kontrol, které dopomohou k tomu, aby všechny nedostatky v datech byly odhaleny a popsány najednou a zmenšil se tak počet vzorků, které si nový dodavatel se správcem musí vyměnit, než bude knihovna přijata mezi stálé dodavatele.

Přechodem na ORACLE se rozšířily nejen možnosti pro uživatele, ale především se zjednodušila samotná správa SK a několikanásobně urychlily procesy související s přijímáním a importováním dat. Zpracování dávkově dodávaných záznamů probíhá zcela automaticky. Systém je schopen bez zásahu správce přijmout a identifikovat soubor dat, provést formálně logické kontroly, importovat vyhovující data a vytvořit statistiku jak pro správce tak pro dodavatele dat. Nová data může dodavatel předat do SK na disketě, umístit do příslušného pro-

storu na FTP server nebo předat pomocí volby Zaslání záznamů do SK (stažením souboru ze svého počítače).

Jedinou podmínkou pro to aby, naimportování dat mohlo proběhnout zcela bez zásahu správce, kterou je třeba ze strany dodavatel dodržet při předání dávky, je pojmenování zasláního souboru v souladu s názvovou konvencí. Délka názvu datového souboru je 8+3 tedy 11 znaků. Prvních 6 znaků názvu tvoří sigla instituce, znaky 7 a 8 popisují použitou znakovou sadu a znaky 9, 10 a 11 popisují formát dat (např. aba001kg.vfi). Pokud dodavatel zasílá současně více souborů může je odlišit číslem na konci názvu souboru (např. aba001kg.vfi10 nebo aba001kg.vfi20). Pokud tato podmínka dodržena není, systém vyrozumí správce, ten soubor přejmenuje a vpustí zpět do procesu zpracování. Správce zasahuje také v případě, že zaslání data neodpovídají váze, která byla knihovně přidělena - a to jak v případě, že jsou data méně nebo naopak více kvalitní. (Podrobnější informace o názvové konvenci, systému přidělování vah apod. lze nalézt v Zadáání pro vývoj aplikací v systému ORACLE na adrese <http://www.caslin.cz:7777/caslin/zadani2.html>).

K 26.6.2000 má Souborný katalog ČR - knihy celkem 605048 originálních záznamů ze 42 knihoven. V tabulce 1. je zpracován přehled kolik záznamů, která knihovna doposud dodala. Nízká čísla u některých knihoven jasně ukazují, které knihovny jsou teprve na počátku spolupráce s SK ČR. O spolupráci od začátku roku projevíly zájem další čtyři knihovny, které v tabulce zatím nejsou zahrnuty.

Tab. I - přehled počtu dodaných záznamů od konkrétních knihoven

ABA001	405 955	BOA003	26 815
ABA003	1 980	CBA001	7 252
ABA004	4 781	CBB001	1 298
ABA006	18 309	CBD001	724
ABA007	3 760	CBD005	74
ABA008	3 810	CHD001	5
ABA009	11 134	HKA001	4 808
ABA012	3 222	KLK001	2 666
ABA013	8 733	LIA001	8 565
ABA601	945	OLA001	14 340
ABB033	77	OPD001	1 435
ABB040	512	OSA001	9 689
ABB076	340	OSD002	45
ABB083	2 414	OSK017	105
ABC004	213	PAD001	177
ABC023	2 958	PNA001	1 733
ABC163	3 192	PND002	396
ABD027	640	PND003	389
ABE310	511	PND007	203
BOA001	40 550	ULG001	5 276
BOA002	4 469	ZLD002	546

V souborném katalogu může uživatel najít mezi vlastníky hledaných dokumentů i sigly řady jiných institucí než těch, které jste našli v Tabulce I. Jedná se celkem o 342 dalších institucí mezi nimiž je i 102 institucí ze Slovenska. Důvodem je skutečnost, že základem báze SK se stala báze konferencí, která se po řadu let před rozdělením Československa budovala tehdy ještě v odd. CEZL v systému ISIS. Představovala více než 7000 titulů materiálů ze zahraničních konferencí a více než 2000 z nich vlastnily slovenské instituce (celkem ve 2262 případech můžete nalézt jako vlastníka vyhledaného dokumentu slovenskou institucí). Sigly těchto institucí se bohužel v SK nedají dešifrovat - z katalogu však existuje odkaz na www adresu UK v Bratislavě, kde je možno v příslušném adresáři vyhledat informace o sigle.

Zbývajících 240 sigel českých knihoven se do báze dostalo, buď stejným způsobem jako sigly slovenských knihoven nebo v procesu testování jedné z nových funkcí souborného katalogu - připisování sigel. V rámci tohoto testu byly připsány sigly patřící 38 různým českým knihovnám k cca 360 záznamům již existujícím v SK. Funkce „připis sigly“ včetně převzetí záznamu po připsání sigly tak byla vyhodnocena jako zcela spolehlivá.

SK je katalogem heterogenním. Stejně jako dodávat tak i přebírat (exportovat) záznamy uživatel může v různých formátech a v různém kódování. Pro ty, kteří by měli zájem pouze o přebírání dat, je jistě důležitá informace v jakých formátech a kódováních je možné z SK data přebírat:

Formát souborů dat:

- Řádkový UNIMARC (manuál UNIMARC) používá pro označení stopslov v Latin 2 PC znaky l ě
- UNIMARC ISO 2709 (manuál UNIMARC, Záznam pro souborný katalog UNIMARC) používá pro označení stopslov v Latin 2 PC znaky l ě
- Výměnný formát ISO 2709 (manuál CDS/ISIS) používá pro označení stopslov znaky <>
- Exportní soubor z ALEPH používá pro označení stopslov znaky <<>>
- Výměnný formát, exportní soubor ze systému CDS/ISIS (je totožný s ISO 2709, ale rozsegmentovaný na řádky po 80 znacích). používá pro označení stopslov znaky <>

Znakový repertoár:

Abecedy na základě latinky, přípustné kódy jsou:

- ISO 646 nebo ISO 5426 (manuál UNIMARC ,Příloha J)
- veškerá diakritika kódována pomocí GIZMO notace (tabulka GIZMO notace instrukce Záznam pro souborný katalog, příloha H)
- kód PC Latin 2 (Microsoft Code Page 852) + GIZMO
- kód Kamenických + GIZMO
- UNICODE UTF 8
- ISO 8859-2 + GIZMO

Exportovaná data si může uživatel převzít e-mailem, nechat vystavit na příslušném místě FTP serveru SK (uvažuje se i o možnosti vystavení na FTP serveru konkrétní knihovny, ale tato možnost dosud nefunguje), popřípadě může soubor pro export stáhnout přímo do vybraného adresáře na svém počítači. Exportovat lze jak jednotlivé záznamy, tak prostřednictvím k tomu určené schránky i větší počet záznamů najednou. V praxi tuto možnost - exportovat záznamy - zatím běžně používá pouze NK ČR. Pro potřeby tvorby České národní bibliografie došlo od počátku roku 2000 k vyexportování 15291 záznamů. Ostatní knihovny si (podle statistiky pro správce) exportování záznamů pouze vyzkoušely. Větší počet záznamů zatím z SK převzal pouze Obor knihovnictví NK ČR - celkem 41 záznamů.

Poslední funkcí, která neodlučně patří ke sdílené katalogizaci je editace již dodaných záznamů popř. vytváření nových záznamů přímo v prostředí souborného katalogu. K tomu slouží speciální formuláře, vytvořené v JAVA. Pro editaci je k dispozici vstupní formulář, který zobrazí kompletní obsah editovaného záznamu a který zároveň umožní zpracovateli přidat do záznamu libovolné další povolené pole. Před uložením editovaného záznamu, má uživatel možnost zkontrolovat záznam na UNIMARC a následně opravit případné nedostatky. Také při vytváření nového záznamu nabízí SK uživateli speciální formulář, který obsahuje všechna pole minimálního záznamu. Uživatel ale tento formulář není povinen využívat a má možnost vytvořit si formulář vlastní - s takovým výběrem polí a podpolí, které bude chtít při katalogizaci svého fondu vyplňovat, tento formulář si libovolně pojmenovat a uložit pro pozdější využití. Možné je i to, že část údajů, které předpokládá, že se v jeho záznamech budou opakovat si může předvyplnit. Takto může vytvořit libovolný počet vlastních formulářů. Jedinou podmínkou pro uložení vlastního formuláře je dodržení pravidla, že ve formuláři jsou přítomna všechna pole minimálního záznamu pro SK.

Před uložením nově vytvořeného záznamu do SK prochází záznam testem na UNIMARC, formálně logickými kontrolami a testem na duplicitu, stejně jako záznam, který přichází dávkově. V případě, že je během ukládání záznam vyhodnocen jako duplicitní - jsou porovnány váhy přidělené jednotlivým knihovnám, které duplicitní záznamy zpracovaly a je přijat záznam s vyšší vahou. V případě, že záznamy mají váhu shodnou, zůstává v bázi záznam, který byl přijat jako první a z nového záznamu se přebírají pouze lokální údaje. K této situaci, ale bude zřejmě docházet jen ojediněle - neboť předpokládáme, že katalogizátor před vytvořením nového záznamu prověří dostupnost záznamu v SK a dá přednost připsání své sigly k již existujícímu záznamu a jeho převzetí než k jeho opětovné katalogizaci.

V tuto chvíli SK poskytuje služby všem uživatelům, kteří potřebují vyhledávat konkrétní monografické dokumenty, získat informaci o jejich vlastníkovi, zviditelnit své fondy v prostředí SK, zjednodušit práci svým katalogizátorům, či využít prostředky SK ke katalogizaci vlastních fondů. Další funkcí, kterou se do konce roku 2000 SK chystá nabídnout je MVS.