

ČTYŘI ROKY REÁLNÉHO FUNGOVÁNÍ CASLIN – SOUBORNÉHO KATALOGU ČR A JEHO BUDOUCNOST

Gabriela Krčmařová, Národní knihovna ČR

CASLIN – Souborný katalog ČR (dále jen souborný katalog) je největší reálný souborný katalog v České republice a stal se významným informačním zdrojem pro uživatele knihoven i pro knihovníky. Tomuto konstatování předcházely čtyři roky velmi náročné práce jak správce souborného katalogu, tak všech jeho účastníků.

1. Lokalizace dokumentů a sdílení záznamů dokumentů

Na počátku bylo nezbytné odpovědět si na otázku, komu má souborný katalog sloužit, zda knihovníkům nebo uživatelům v nejšířším slova smyslu. Souborný katalog je samozřejmě třeba budovat tak, aby si zachoval ve vyvážené poloze možnost sloužit oběma skupinám stejně. Není správná cesta sběru pouze bibliograficky dokonalých záznamů od úzkého okruhu ověřených knihoven, abychom tak dosáhli skvělého zdroje pro sdílenou katalogizaci. Na konci této cesty je výrazné okleštění informační funkce souborného katalogu. Správná cesta však není ani sběr co největšího počtu záznamů bez ohledu na jejich kvalitu. Tato cesta v podstatě znemožňuje knihovníkům využití sdílené katalogizace a ve svém důsledku velmi znesnadňuje vyhledávání všem běžným uživatelům, a tudíž oslabuje také informační funkci souborného katalogu.

Správce souborného katalogu zvolil onu příslovecnou zlatou střední cestu, která nebyla jednoduchá. Je založena na jedné straně **na otevřenosti všem knihovnám**, které mají zájem zasílat své záznamy, a na druhé straně **na pevně stanovených standardech**, které musí daná knihovna dodržet.

Pozn.: Jak obtížné bylo stanovit jednotný obsah a strukturu záznamu pro souborný katalog, vypovídá skutečnost, že návrh Záznamu pro souborný katalog vydala Národní knihovna ČR v červnu 1994 a definitivní instrukci teprve po dvou letech v květnu 1996.

K 30.6.1999 obsahuje CASLIN – Souborný katalog ČR téměř **270.000 záznamů** porovnaných na duplicitu, do souborného katalogu dodává záznamy 30 českých a moravských knihoven. Záznamy mají **shodnou strukturu** (formát UNIMARC), která je nezbytná z hlediska jejich vzájemné kompatibility. Po obsahové stránce se mohou lišit z hlediska úplnosti jednotlivých údajů, jejichž minimální rozsah však zaručuje jednoznačnou identifikaci dokumentu. **Kvalita jednotlivých záznamů z hlediska úplnosti je ohodnocena tzv. vahou**, čím je vyšší, tím je záznam kvalitnější.

Co to znamená v praxi?

Vstoupí-li uživatel (čtenář či pracovník MVS) do souborného katalogu jen proto, aby lokalizoval konkrétní dokument, nezajímá se o váhu nalezeného záznamu, i u záznamu s nejnižší vahou se zajímá pouze o sigly knihoven, které hledaný dokument vlastní. Naproti tomu knihovník – katalogizátor v souborném katalogu vyhledává záznamy k dokumentům, které má bibliograficky zpracovat, a chce si je stáhnout do svého lokálního katalogu (copy cataloging). Katalogizátor naopak velmi pozorně sleduje váhu záznamů a má možnost stahovat jen ty nejuplněnější. Má ovšem také možnost stahovat záznamy s nižší vahou a po doplnění je kopírovat do lokálního katalogu své knihovny.

Souborný katalog je přístupný na Internetu a v plném rozsahu poskytuje informačně – lokální služby. V současné době probíhá ve spolupráci s knihovnami ČR testování kopírování záznamů i vlastní sdílená katalogizace prostřednictvím vstupního formuláře a pro uživatele bude tato služba k dispozici do konce roku 1999.

2. Kvantitativní naplňování souborného katalogu

V červnu 1995 byl souborný katalog zpřístupněn na Internetu v ověřovacím provozu jako báze KZK – Katalog zahraničních knih.

Z hlediska kvality souborného katalogu je jistě velmi významná skutečnost, že **na samém počátku sběru dat byla stanovena jednotná instrukce**, která specifikovala strukturu a obsah záznamu pro souborný katalog. V září 1994 byl vydán návrh instrukce Záznam pro souborný katalog, na jejímž základě v lednu 1995 správce zahájil první kontakty s knihovnami, které projevíly zájem přispívat svými záznamy do souborného katalogu. V červnu 1995 byly do souborného katalogu importovány první záznamy zahraničních monografií a do konce roku 1995 bylo importováno celkem 25.000 záznamů osmi knihoven. **Ověřovací provoz souborného katalogu**, kdy byly přijímány záznamy pouze zahraničních monografií bez možnosti kontroly duplicit, probíhal **od 1.1.1996 do 31.1.1997**, na konci ověřovacího provozu souborný katalog obsahoval téměř 44.000 záznamů čtrnácti účastníků. V květnu 1996 vydala Národní knihovna ČR schválenou instrukci Záznam pro souborný katalog: UNIMARC. Tištěné monografie. V lednu 1997 bylo dokončeno testování programu na kontrolu duplicit pro souborný katalog a všechny záznamy do té doby dodané do souborného katalogu byly porovnány na duplicitu.

1.2.1997 zahájil souborný katalog příjem záznamů nejen zahraničních, ale také českých tištěných monografií a zahájil tak svůj **rutinní provoz**. Na konci roku 1997 souborný katalog obsahoval téměř 65.000 záznamů (z toho 1.780 duplicitních) devatenácti českých a moravských knihoven.

K 31.12.1998 souborný katalog obsahoval téměř **243.000** na duplicitu porovnaných záznamů dvaceti osmi účastníků.

Kvantitativní údaje naplňování souborného katalogu v uplynulých čtyřech letech ukazuje následující tabulka:

	1995	1996	1997	1998
Celkový počet dodaných záznamů	24.820	40.210	105.160	318.947
Záznamy dodané v příslušném roce	24.820	15.390	64.950	213.787
Duplicitní záznamy	–	–	1.780	76.132
Počet účastníků	8	14	19	28

3. Zpracování a import dat do souborného katalogu

Účastníci souborného katalogu zasílají data na ftp server nebo na disketě klasickou poštou. Zasilání dat prostřednictvím e-mailu se neosvědčilo, velmi často docházelo k narušení struktury i obsahu dat.

Proces zpracování záznamů se skládá z těchto fází:

- kvalitativní analýza dat
- konverze dat
- rozlišení záznamů českých a zahraničních dokumentů
- přidělení kvalitativní váhy
- řízený import dat do báze souborného katalogu

Kvalitativní analýza dat spočívá v ruční kontrole struktury a obsahu dat u každé dodané dávky. Pokud se jedná o účastníka, který již dodává data pravidelně, je prováděna pouze namátková kontrola u každé nové dávky dat. Avšak záznamy všech nových účastníků jsou prověřovány velmi pečlivě v několika kolech, kdy dochází k aktivní komunikaci písemně, telefonicky i osobní mezi správcem souborného katalogu a účastníkem, který má zájem o dodávání svých záznamů do souborného katalogu.

Souborný katalog je provozován v systému ALEPH, který však neumožňuje provádět **import záznamů s kontrolou na duplicity** dle potřeb souborného katalogu a dále nedovoluje zásahy do záznamů v bázi jinými než vlastními prostředky. Další modifikace pomocí programů je nutno provádět mimo bázi. Správce souborného katalogu používá externě vytvořený program pro řízený import s kontrolou duplicit. Toto řešení však vyžaduje aktivaci procedury kontroly duplicit mimo bázi v systému ALEPH, a tudíž je možné zpracovávat pouze data získaná off-line.

Řízený import dat s kontrolou duplicit do souborného katalogu trvá nejméně osm hodin, pokud se vyskytnou komplikace i déle. Přitom nezáleží na velikosti jedné dávky dat, protože program „čte“ celou bázi, ale záleží na velikosti báze, která pochopitelně stále roste. Zpracování dávky dat jedné knihovny (bez ohledu na to, zda obsahuje 100 či 10 000 záznamů) trvá dva pracovní dny a správce je nucen během tohoto procesu spustit osm pomocných programů:

- první den probíhá kvalitativní analýza dat, konverze dat, rozlišení záznamů českých a zahraničních dokumentů, přidělení kvalitativní váhy
- druhý den probíhá řízený import dat do báze souborného katalogu.

Od počátku roku 1998 je souborný katalog zdrojem pro **kooperativní zpracování České národní bibliografie**, proto správce musí pravidelně každý měsíc přednostně importovat záznamy MZK v Brně a ostatních státních vědeckých knihoven, které se na kooperativním zpracování ČNB podílejí.

Lze si snadno spočítat, jak málo pracovních dnů zbývá správci souborného katalogu na zpracování záznamů ostatních účastníků. S rostoucím počtem záznamů v souborném katalogu probíhá vlastní import jednotlivých dávek stále pomaleji. Správce řeší situaci tak, že pokud je to nezbytné, importuje záznamy ostatních účastníků po dvou až třech dávkách najednou. Každé řešení tohoto stavu je na úkor aktuálnosti souborného katalogu.

4. Vývoj nového softwaru pro souborný katalog

Při různých příležitostech bylo konstatováno, že zajištění sdílené katalogizace pro knihovny v ČR je hlavním důvodem změny systému pro souborný katalog. Mnohem méně se však hovořilo o tom, že vývoj vlastního softwaru pro provoz a správu souborného katalogu v systému ORACLE je pro správce nezbytný z důvodu zásadní racionalizace zpracování a celkové správy dat. K výše uvedenému je třeba zdůraznit, že zpracování a import dat bude v rámci nového systému probíhat stokrát rychleji s lepší kvalitou (celý proces zpracování a importu dat, který v současné době probíhá dva pracovní dny, bude trvat v ideálním případě deset minut).

Od března 1998 probíhá vývoj aplikací pro provoz souborného katalogu v systému ORACLE. Dojde ke značné automatizaci všech činností spojených se správou bázi souborného katalogu včetně analýzy vstupních dat, a to s maximálním využitím k tomu účelu dosud vytvořených softwarových prostředků. Správce měl na mysli **maximální racionalizaci pracovních postupů**, která spočívá v eliminaci lidského zásahu všude tam, kde je to možné a efektivní.

Zpracování záznamů probíhá zcela automaticky:

- příjem a identifikace souboru dat (včetně konverze)
- formálně logické kontroly dat
- import dat
- statistiky
- problémy k řešení pro správce

Účastník umístí svá data v přiděleném prostoru na **ftp serveru** (pokud je dodá na disketě, nahraje je na ftp server správce). Program bude v pravidelných intervalech kontrolovat, zda na ftp server nepřibyla nová data. Stáhne je a dle názvové konvence zjistí jejich vlastníka, v jakém jsou formátu a jaká byla použita znaková sada.

Všechny nové i editované záznamy budou před importem do souborného katalogu testovány. Součástí **automatické kontroly** je test na UNIMARC, přidělení kvalitativní váhy a test na duplicitu záznamů.

Po úspěšném dokončení formálně logických kontrol dojde automaticky k vlastnímu **importu záznamů** do báze souborného katalogu. Informace o nevyhovujících záznamech budou prostřednictvím e-mailu jako statistika s příslušným komentářem zaslány zpět příslušné knihovně k opravě.

Celý **proces zpracování záznamů** má správce možnost spustit **jako celek**, takže jednotlivé kroky proběhnou automaticky v návaznosti na sebe, **nebo po jednotlivých krocích** tak, že má možnost kontroly výstupů z kaž-

dého kroku. V obou variantách správce může nastavit datum a čas spuštění celého procesu nebo kroků.

Pro uživatele i pro správce se vytvářejí **statistiky**. Uživatel má možnost zobrazit si statistiku a zprávy systému vztahující se k jeho činnosti v bázi na konci své aktivity. Tyto statistiky jsou archivovány po dobu stanovenou správcem a přístupné uživateli.

Problémy k řešení pro správce souborného katalogu:

- stejně jako dodavatel dat obdrží správce přesnou statistiku chyb v záznamech
- pokud v průběhu automatického zpracování dat dojde k chybnému zakončení procesu, zabývá se správce hledáním příčiny, proč k tomu došlo a hledá nápravu
- pokud kvalitativní váha přidělená nové dávce dat neodpovídá dříve přidělené váze, systém na to správce upozorní a ten provede hlubší analýzu dat
- systém správce upozorní, že vlastníkem dodané dávky dat je knihovna, která ještě nemá přidělenou váhu, správce provede ruční kvalitativní analýzu dat.

Základní činností správce souborného katalogu již nebude zdlouhavé a náročné zpracování a import dat, uvolněné pracovní kapacity bude možné věnovat systematickému čištění dat v souborném katalogu.

5. Testování nového softwaru pro souborný katalog

Důležitou součástí vývoje každého softwaru je jeho testování externími uživateli. „*Nultou*“ verzi vyvinutého systému budou testovat členové Týmu technického rozvoje Souborného katalogu ČR, z toho vyplývá, že do nového systému souborného katalogu dostanou v první fázi přístup tyto instituce:

- ČVUT Praha
- Masarykova univerzita Brno
- Moravská zemská knihovna Brno
- Národní knihovna Praha
- Státní technická knihovna v Praze
- Univerzita Palackého Olomouc
- Vysoká škola ekonomická Praha

Beta verzi softwaru pro souborný katalog budou testovat členové Pracovní skupiny pro souborný katalog.

Připomínky budou správci zasílány prostřednictvím e-mailu.

Při vstupu do souborného katalogu je uživateli nabídnuto vyhledávání v bázi **SKK** (souborný katalog knih), v bázi **ADR** (adresář knihoven a informačních institucí) a propojení do báze **KZP** (katalog zahraničních periodik). V bázi SKK je uživateli nabídnuto:

- vyhledávání
- vytváření tiskových souborů
- stahování vyhledaných záznamů
- přepisování sigly k záznamům, které vlastní
- editace vlastních záznamů

- vytváření nových záznamů přímo v souborném katalogu.

Do konce roku 1999 bude souborný katalog v systému ORACLE přístupný pro veřejnost.

6. Další rozvoj souborného katalogu

Správce souborného katalogu v současné době připravuje technické podmínky pro realizaci 1. fáze **meziknihovní výpůjční služby** prostřednictvím souborného katalogu.

1. fáze:

Služba je poskytována pouze registrovaným uživatelům resp. účastníkům souborného katalogu (uživatel – jednotlivec je vždy zastoupen knihovnou). Po volbě „**Objednávka MVS**“ a zalogování obdrží uživatel **formulář objednávky**. K vyplnění jsou určena pouze silně orámovaná pole. Povinné pro vyplnění je pouze pole požádaná instituce a částka, kterou je uživatel ochoten zaplatit. Vyplnění ostatních polí závisí pouze na uživateli. Obsah pole „Uživatele zastupuje“ je přebrán z báze ADR. Obsah pole „Objednané dílo“ a „Požádaná instituce“ je přebrán z báze SKK.

Vyplněný formulář je **e-mailem zaslán** na adresu vybrané „požádané“ knihovny. E-mailová žádanka MVS může být odeslána pouze té knihovně, která s jejím přijetím bude souhlasit. Tím je služba MVS pro souborný katalog ukončena.

2. fáze:

Realizace 2.fáze MVS v rámci souborného katalogu předpokládá, že většina transakcí souvisejících s MVS probíhá na serveru souborného katalogu a je založena na existenci schráněk.

1. fázi MVS v rámci souborného katalogu bude možné realizovat v krátké době bez větších problémů na základě jednoduchých formálních dohod mezi knihovnami.

Realizace 2. fáze vyžaduje nejen detailní jednání mezi účastnickými knihovnami, jehož účelem je vytvoření jasné dohody po organizační i legislativní stránce, ale také vývoj dalších náročných aplikací v systému ORACLE umožňujících technickou realizaci MVS probíhající na serveru souborného katalogu.

V současné době souborný katalog přijímá pouze záznamy tištěných monografií. V rámci řešení úkolů projektu výzkumu a vývoje správce řeší vývoj aplikací v systému ORACLE pro seriály a speciální druhy dokumentů. Dalšími důležitými úkoly je vývoj protokolu Z39.50 a technické propojení záznamů souborného katalogu s autoritními záznamy souboru národních autorit. Projekt Rozvoje souborného katalogu byl nahlášen jako výzkumný záměr Radě vlády pro výzkum a vývoj do Centrální evidence výzkumných záměrů na léta 1999 – 2003, což dává jistý předpoklad určité nadstandardní finanční podpory budování a rozvoje CASLIN – Souborného katalogu ČR.

Závěr

Závěrem je třeba zdůraznit, že vývoj vlastního softwaru dává správci souborného katalogu i do budoucna možnost úprav stávajících aplikací a **vývoj nových aplikací v souladu s vývojem informačních technologií.**