

KVALITNÍ KATALOG – NEZBYTNÝ ZDROJ PRO POSKYTOVÁNÍ SLUŽEB KNIHOVNY

Radka, Římanová, Státní technická knihovna v Praze

Katalog knihovny je jedním ze základních zdrojů pro poskytování knihovnicko-informačních služeb. Knihovny, které proklamují kvalitu své katalogizace, obvykle argumentují (jistě správně) dodržováním standardů. Razítko úrovně katalogizace lze získat v České republice při přispívání do souborného katalogu CASLIN, který „známkuje“ zasílané záznamy přidělením váhy, jež určuje právo přepisu záznamu v katalogu při duplicitním odběru ve více knihovnách. Do zcela jiných dimenzí se hodnocení kvality katalogizace dostane tehdy, pokud katalog hodnotí koncový uživatel, ať již z řad profesionálů nekatalogizátorů nebo laik. A ještě dramatictější situace nastává při vzájemném hodnocení kvality katalogizace různými katalogizátory například v rámci kooperativních projektů nebo při kopírování záznamů z různých knihoven.

Co je však opravdu důležitým měřítkem kvality? Významným prvkem procesu hodnocení kvality katalogizace je porovnávání záznamů a jejich částí katalogu se studií Functional Request for Bibliographic Format (FRBR). Pravdou je, že poměrně vysoké hodnocení kvality záznamů vyplývající z dostupných studií (např. dále zmiňovaný test knihovníků British Musea [MCEWAN, 2004]) jako by katalogizátory uklidňovalo, že je stávající praxe naprosto v pořádku a v souladu s trendy a uživatelskými potřebami [DROBÍKOVÁ, 2003]. Z testů vyplývá závažnost typografické chyby [BALLARD, 2005] a odchylky od standardů. Ze své praxe však vím, že právě standardy nás někdy nutí ne zcela logicky (nahlíženo pohledem uživatele vyhledávajícího Google) popsat dokument. Ve standardech vidím především důležitý prvek unifikace a udržení kontinuity databází v prostředí vnitřním i vnějším. Víme však, že na revizi anglo-amerických pravidel se pracuje, AACR3 bude reflektovat změny vycházející z možností elektronických databází oproti listkovým katalogům. Vedle knihovnického přístupu ke katalogu, který reprezentuje studie FRBR, začínají stále výrazněji vystupovat zastánci přizpůsobování knihovních katalogů internetovým vyhledávacím typu „Google“, jako uživatelsky nepřijatelnější podobě katalogu.

Vytvoření a uvádění kvalitativních ukazatelů katalogizace by se mělo co nejdříve stát dalším ukazatelem atributu konkrétního katalogu, vedle dosud uváděných statistických kvantitativních položek.

Proč nás zajímá kvalita katalogizace:

1. zkvalitnění služeb – uživatelské hledisko;
2. měřítko kvality vlastního personálu;
3. měřítko vynaložených nákladů;
4. měřítko kvality výuky;
5. ověření kvality a smysluplnosti uplatnění technologických postupů;
6. ověření kvality softwaru;
7. ověření zdroje pro sdílenou katalogizaci (copy cataloging);
8. ověření kvality pracoviště (při žádosti o grant, žádosti o zaměstnání, praxi);
9. měřítko úspěšnosti projektů (retrospektivní katalogizace, souborné katalogy).

Uživatel katalogu

Jeden z testů sledujících kvalitu katalogizace a popisovaných ve studii autorů A. McEwana a T. Younga je byl nazván „User story“ a vypráví příběh nesmírně trpělivého čtenáře a nesmírně chybného záznamu. Uživatel chce najít knihu v katalogu. Popisována jsou všechna úskalí, která se před uživatelem objeví, když hledá záznam s chybným zápisem ISBN a dalšími chybnými nebo nepřesnými údaji. Studie zjišťuje, jaké výsledky získává uživatel při postupném zadávání selekčních termínů z konkrétního záznamu. Selekční termíny jsou vybírány v rámci doporučení FRBR. V tomto příběhu mne velmi zaujalo, že mimo jiné zmiňuje důležitost uvedení kontaktu na vydavatele nebo distributora přímo do katalogizačního záznamu, jako formu cesty pro získání publikace (uživatel si knihu koupí, ne vypůjčí). Dosud jsem za požadavkem „získat“ chápala pouze výpůjčku nebo MVS, nikoli nákup. Údaje o akvizici se obvykle v záznamech českých knihoven tají jako „top secret“ information!

Při vyhodnocování kvality katalogu z hlediska spokojenosti uživatelů je problematické získat zpětnou vazbu od uživatelů. Přitom právě uživatelské hledisko je doporučováno sledovat při opravách katalogu, vyhledávání chyb [MACEWAN, 2004]. Jednou z možností jsou rozhovory s uživateli, uživatelské ankety a dotazníky.

Druhým zdrojem informací o uživateli je knihovnicko-informační systém knihovny, včetně elektronického katalogu. Tam se dozvíme, co si čtenáři půjčují, jak často, co hledají a které termíny pro vyhledávání používají.

Ve Státní technické knihovně jsme zprovoznili statistické sledování používání vyhledávacích polí. Výsledky k 20. 7. 2005 dokumentuje následující tabulka

Statistika vyhledávání v bázi za období 21. 4. 2005 – 21. 7. 2005:

Index	WWW	Z39
Prohlížení abecedních rejstříků		
Název	6204	0
Autor	3310	0
Předmětové heslo	2536	0
Název produktu	302	0
Heslo PSH	188	0
MDT	75	1
Vydavatel	204	0
Edice	77	0
ISBN	430	1212
ISSN	79	56
Signatura	3350	0
Vyhledávání v plných textech položek		
Všechny položky	46939	0
Autor	8377	3369
Název	7461	2716
Předmětová hesla	2884	107
Vydavatel	730	0

Ve sloupci WWW jsou zachyceny přímé přístupy do OPAC (z knihovny i mimo knihovnu), ve sloupci Z39 jsou přístupy přes protokol Z39.50 (převážně Jednotná informační brána). Výsledky statistiky jsou velmi inspirativní pro další rozvoj interface OPAC, názvy vstupních prvků vyhledávacích obrazovek (např. „Heslo PSH“ – pro uživatele zavádějící termín) atd. Velmi nás překvapil nízký počet vyhledávání v rejstříku „Vydavatel“, domnívali jsme se, že se jedná o důležitější údaj, právě z hlediska vyhledávání.

Knihovník ve službách, pomocník při vyhledávání v katalogu

Profesionální knihovník je prvním pomocníkem uživatele v případě, že selže samoobslužný způsob vyhledávání v katalogu. Zatím jsem následující tvrzení nijak experimentálně neověřovala, ale domnívám se, že poměr mezi

kliknutím na „help“ elektronického katalogu a dotazem na pracovníka služeb při potřebě pomoci s hledáním v katalogu je minimálně 1:1. Pouze kvalifikovaný pracovník se znalostí vnitřního prostředí katalogu může spolehlivě splnit požadavek najít, určit, vybrat a získat informace o hledané entitě v katalogu, což je jedním ze základních požadavků na funkční katalog [Funkční požadavky na bibliografický záznam, 2001]. Takový pracovník služeb také dokáže plnit nezbytný úkol zpětné vazby pro katalogizátora při zjištění, že katalog knihovny tyto základní požadavky neplní. Tuto zpětnou vazbu velmi zřídka dostaneme přímo od laického uživatele katalogu, ten spíše rezignuje na služby katalogu i knihovny.

Samozejmě v případě malých a středních knihoven, kdy knihovník je univerzálním profesionálem, lze získat většinu znalostí pro profesionální orientaci v katalogizaci při jeho budování.

Inspirována americkými osnovami pro vzdělávání katalogizátorů pro práci s metadaty [HSIEH-YEE, 2003] jsem vytvořila malý přehled toho, co by tedy měl vědět každý knihovník o katalogizaci (vypustila jsem oblasti týkající se průniku katalogizace a metadat, o těch až příště):

1. prvky a struktura jmenného a věcného popisu, zákonitosti výběru selekčních termínů, soubory autorit a jejich význam pro verifikaci termínů pro vyhledávání,
2. interpretace katalogizačních záznamů – formáty zobrazení, importů, exportů,
3. sémantická a syntaktická pravidla katalogizace,
4. standardy v oblasti katalogizace,
5. znalost různých systémů katalogů (souborné katalogy, oborové katalogy, národní bibliografie...),
6. volby a možnosti rešeršní strategie v knihovních katalozích.

Pracovník služeb by měl být také obeznámen s historií budování a koncepcí katalogu vlastní knihovny.

Rozhraní OPAC a kvalita katalogizace

Kvalitu práce s katalogem ovlivňuje mimo úroveň zpracování dat i kvalita přístupového rozhraní elektronického katalogu.

Dochází k situaci, že data jsou z hlediska požadavků na kvalitu v pořádku, ale navigace uživatele pro vyhledávání a zobrazování nikoli.

Poté co byla zpracována funkční analýza mapující vztah mezi studií FRBR a formátem MARC21, navázaly na tuto práci další projekty, které jsou zaměřeny na zjišťování problémů dnešního pojetí katalogizace a navrhuje změny standardů (např. revize AACR), ale také hledají nové možnosti práce s již existujícími daty katalogů v souladu s požadavky FRBR. Jak již bylo řečeno, proměny koncepce zpracování významně ovlivňují konzistenci

databáze, která je kvalitativním ukazatelem katalogu. Výrazné změny standardů katalogizace bez navázání na postupy používané při budování stávajících knihovních katalogů, bez existence převodních můstků, mapování a zpětné konfrontace již uložených dat by znamenaly podstatné znehodnocení katalogu knihovny. V podstatě se při všech změnách musí vycházet ze základního předpokladu, že stávající záznamy nelze přepracovávat jednotlivě ve smyslu nové koncepce [2005, RAJAPATIRANA]. Naopak chyby je třeba opravovat i zpětně.

Jeden z experimentů inspirovaný FRBR byl proveden v databázi OCLC a zkoumal možnosti identifikace všech vyjádření a provedení díla. Z této studie (HICKEY, 2002) vyplývá důležitost užívání principu „authority control“ vůči selekčním prvkům záznamu (unifikovaný název, jméno autora, věcný popis...) při katalogizaci i při verifikaci dotazu při vyhledávání. Tento problém nastává u děl, která patří mezi často komentovaná a doplňovaná a na jejichž základě vznikají další samostatná díla (např. ilustrace Adolfa Kašpara knihy B. Němcové Babička mohou být vydána samostatně nebo být vystavena), revidované práce, sebrané a vybrané spisy, posvátná díla, díla často prezentovaná (např. Jih proti severu), díla často překládaná a upravovaná [2002, HICKEY].

Tuto problematiku řeší vyvážení autoritních záznamů typu Jméno/Název. Příklad záznamu takové autority je uveden zde.

Janáček, Leos, 1854-1928. Moravské lidové písně

Heading

- Janáček, Leos, 1854-1928. Moravské lidové písně

Used for:

- Janáček, Leos, 1854-1928. Patnáct moravských lidových písní
- Janáček, Leos, 1854-1928. Patnáctkrát moravských národních písní
- Janáček, Leos, 1854-1928. Moravské národní písně
- Janáček, Leos, 1854-1928. Fifteen Moravian folk songs
- Janáček, Leos, 1854-1928. Moravian folk songs
- Janáček, Leos, 1854-1928. 15 moravských lidových písní
- Janáček, Leos, 1854-1928. 15 moravských národních písní
- Janáček, Leos, 1854-1928. 15 Moravian folk songs
- Janáček, Leos, 1854-1928. Mährische Volkslieder
- Janáček, Leos, 1854-1928. Chansz traditionnels de Moravie

Source data found

- His Patnáct moravských lidových písní, 1978. 1 p. (Patnáct moravských národních písní = Patnáctkrát moravských národních písní = Fifteen Moravian folk songs)
- New Grove dict. (Moravské lidové písně)



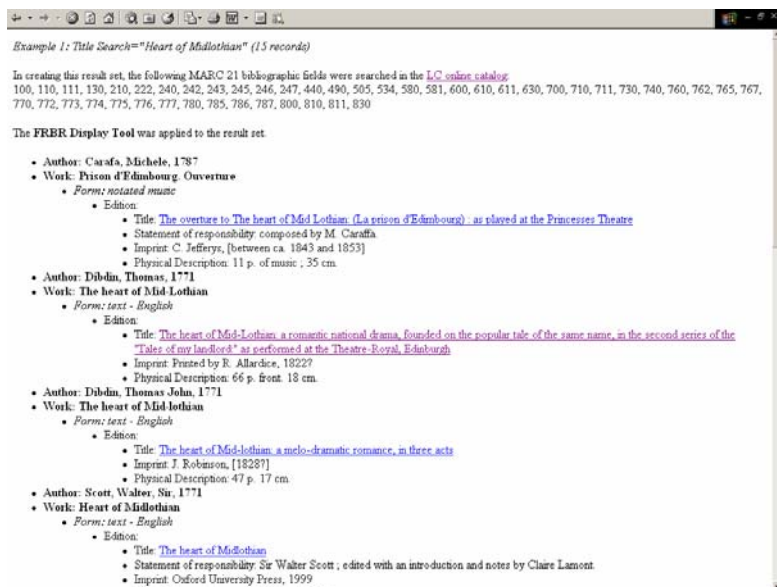


Please send comments, suggestions, or bug reports to webmaster@vww.dra.com

Pro analýzy použitelnosti zobrazení dat v intencích FRBR je na stránkách Library of Congress vystaven „FRBR Display Tool“.

Tento volně dostupný program umožňuje knihovně otestovat kvalitu své katalogizace vůči FRBR a vyzkoušet možnosti zobrazování svých katalogizačních dat v souvislostech vzájemných vztahů [2004, FRBR Display Tool] a nezávisle na komerčních softwarech OPACů. FRBR Display Tool pracuje nad bibliografickou (nikoli přírůstkovou/holdingovou) částí záznamu. Software předpokládá data ve formátu MARC21, před zpracováním je nutné převést data do formátu XML (MARCXML). Veškeré technické podrobnosti k tomuto softwaru lze získat na stránce „FRBR Display Tool“.

Jak takové zobrazení může vypadat, ukazuje příklad získaný v manuálu webové prezentace „FRBR Display Tool“.



Tento software byl použit australskými knihovníky k testování záznamů Australian National Bibliographic Database (ANBD), která čítá celkem 37 500 000 záznamů a je souborným katalogem, obsahujícím záznamy tradičních i elektronických i on-line dostupných dokumentů [RAJA-PATIRANA, 2005]. Projekt měl za cíl analyzovat záznamy ANBD vzhledem k vytváření skupin ve smyslu vztahů mezi entitami FRBR, podrobně zanalyzovat možnosti zobrazení populárního australského díla (Waltzing Matylda) a zjistit další možnosti rozvoje ANBD v intencích

FRBR. V závěru práce je uvedeno, že získané zkušenosti budou použity pro úpravu interface OPAC.

Z komerčních knihovnických softwarů nabízí možnost podobného zobrazování podle požadavků FRBR VTLS - Delivering Visionary Technology in Library Solutions.

Měření hodnoty záznamů podle FRBR

Další využití FRBR proběhlo v testu provedeném odborníky British Museum [MACEWAN, 2004].

V této studii je kvalita záznamu určována:

1. stupněm úrovně zpracování (verifikace v souborech autorit, atd.);
2. přesností zpracování (dodržování standardů, typografické chyby).

Oba tyto aspekty směřují k třetímu a zároveň hlavnímu měřítku:

3. naplnění cíle vytvoření záznamu (záznam slouží uživateli – najít, určit, vybrat, získat entitu).

Měření kvality záznamu je teď především měření stupně dosažení cíle, který má záznam v katalogu splnit. Měříme tedy kvalitu záznamu porovnáváním údajů v záznamu, které jsou pro uživatele důležité.

Hodnoceny byly záznamy vytvořené pro Britskou národní bibliografii a v rámci CIP (cataloging in publication).

Při testování byla vytvořena tabulka mapující důležitost jednotlivých částí bibliografického záznamu vůči cílům FRBR. Jednotlivé elementy, které se vztahují k vyjádření a provedení díla, byly dále klasifikovány ve stupni 1-3.

Příklad některých polí (MARC21):

Vstupní vlastnost/vztah	Najít	Určit	Vybrat	Získat
Datum 1 (008, 07-10)	1	3	3	3
Datum 2 (008, 11-14)	1	3	3	3
Místo vydání (008, 15-17)	0	1	1	3
Ilustrace (008, 18-21)	2	2	3	0

Při poměřování hodnoty záznamu se opět postupovalo měřením chyb nebo nesprávností nebo absencí údajů v testovaných polích (negativní přístup – forma „ztráty květinčky“).

Samostatně se hodnotila část popisná i předmětová. Sledován byl počet chyb vyskytujících se v polích důležitých pro konkrétní uživatelský požadavek (najít, určit, vybrat, získat). Tento počet se procentuálně vyjadřoval jako část z cílového ideálního a bezchybného stavu. Nakonec byly hodnoty získané testováním v části jmenného a věcného popisu sečteny a znovu procentuálně vyjádřen počet vyhovujících a nevyhovujících záznamů. Vý-

sledné hodnoty nejsou nijak průkazné vůči kvalitě zdrojových databází (BNB 92 % celkové úspěšnosti, CIP 72 % celkové úspěšnosti), ale z pohledu testujících byla ověřena schopnost FRBR statisticky vyhodnotit kvalitu katalogu. Další částí testu je již popsaná „User story“. Studie Quality versus Quantity končí doporučením, že knihovna má mít jako zdroj katalogizačních záznamů ten, který prošel testem kvality, nebo si jej před zahájením stahování otestovat.

Katalogizátor a jeho vliv na katalog

Nelze zavírat oči před faktem, že katalogizace jako samostatná, vážená a velmi rozšířená specializace knihovníka je v současné době ohrožena. Trendem je získávat katalogizační záznamy již hotové, pouze kopírováním do lokální báze, nebo sdílet virtuální katalogy s přidáváním údajů o vlastnictví exempláře. Zkušenosti knihoven, které s touto praxí pracují, však upozorňují na dva okruhy problémů:

1. kvalita získávaných záznamů, ve smyslu topografických a věcných chyb, včetně chyb vůči standardům;
2. konzistence vlastního katalogu, dodržování vlastních „zvyklostí“ úpravy záznamů.

Situace je zhruba taková, že až 55 % stahovaných záznamů je katalogizátory po stahování upravováno [HIDER, 2005]. Tyto úpravy jsou vedeny jednak snahou odstranit chyby v záznamu, ale i přizpůsobit záznam koncepci vlastní databáze. Tyto koncepce by měly odrážet historický vývoj katalogu (nejkřiklavějším příkladem jsou změny v koncepci věcného popisu) a zároveň sledovat uživatelské potřeby profesionálů i laiků. Počtem úprav roste však i čas na zpracování jednoho katalogizačního záznamu, a tím i jeho cena. Smyslem testování kvality tedy je i obhájení nákladů na tvorbu a redakci katalogů vůči managementu knihovny.

Knihovny v současné době se stávají účastníky projektů, i když v ČR méně, než bychom si přáli, které rozšiřují popis dokumentu za hranice tradičního popisu tištěného textového dokumentu směrem k popisu digitálních informačních zdrojů v různých formátech metadat. Metadata nejsou „výnosem“ katalogizátorů. Ti však mohou do dnes široké škály různorodých struktur vnést pořádek ve smyslu maximální unifikace a standardizace, včetně vzájemné propojitelnosti [HSIEH-YEE, 2003]. Český katalogizátor, který prošel v posledních letech několika změnami formátů dat (výměnný formát, UNIMARC, MARC21, vlastní formáty pro objekty zdánlivě nepopsatelné ve standardech...) má s převody bohaté zkušenosti. Jakékoli pořádání informací vycházející z analýzy obsahu a formy dokumentu je základní dovedností, kterou musí každý katalogizátor mít.

O tuto práci se však musí katalogizátoři přihlásit sami, ať již přímým iniciováním projektů nebo propagováním a nabídkou svých vlastních schopností mezi tvůrci digitálních knihoven a podobných projektů.

Při vyhledávání podkladů jsem našla článek „10 ways to improve data quality“ [BEALL, 2005]. Autor uvádí desatero doporučení rad pro knihovníky, katalogizátory, tvůrce knihovnických softwarů i management knihovny, jak postupovat při redakcích katalogů, s cílem vytvořit z nich opravdu účinný prostředek orientace ve fondu. Seznam je však přesným výčtem všeho, co jsem našla v dalších zde citovaných článcích.

Texty jednotlivých bodů jsou v zájmu radikalizace proklamací zkráceny, pořadí, které považuji také za zajímavé, ponecháno:

1. každé záhlaví bibliografického záznamu verifikuj a propoj se soubory autorit;
2. zásadně studuj a dodržuj národní katalogizační standardy pro popis všech druhů dokumentů;
3. překlep – nepřítel katalogu;
4. pro každou životopisnou knihu utvoř předmětové personální záhlaví;
5. věcný popis = aktuální terminologie s odkazy na starší verze;
6. cíleně vyhledávej chyby, náhodně nalezené pomohou najít další podobné;
7. nenechávej systém vylučovat mluvnické členy automaticky – i „DER“ může něco znamenat;
8. hlídej údaje o exemplářích a jednotkách – nebudeš vzbuzovat falešné naděje ve svých čtenářích;
9. revizi polic zautomatizuj, načítej čárové kódy, porovnávej skutečnost s počítačem – pobavíš se a budeš překvapen;
10. hledání chyb je společné – hledají všichni knihovníci a jejich rodní příslušníci, kdo nalezne dříve než uživatel, má první cenu.

Děkuji svým kolegům ze Státní technické knihovny, Mgr. Štěpánce Žižkové a Petru Novákovi za cenné rady, které mi poskytli při práci na tomto článku.

Použité prameny:

1. BALÍKOVÁ, M., KUBALOVÁ, H., SVOBODOVÁ, J. *Katalogizace ve formátu MARC21 : stručná instrukce a příklady pro knihy a některé typy pokračujících zdrojů*. Praha : Národní knihovna České republiky. 2004. 156 s.

2. BALLARD, T. Typographical Errors in Library Databases [online]. 2005 [cit. 2005-07-20]. Dostupné na www: <http://faculty.quinnipiac.edu/libraries/tballard/typoscomplete.html>
3. BEALL, J. 10 ways to improve data quality. *American libraries*. 2005, vol. 36, no. 3, s. 36-37.
4. DROBÍKOVÁ, B. Nový model bibliografického univerza podle FRBR a česká katalogizační praxe [online]. In: *Automatizace knihovnických procesů : sborník příspěvků*. Praha : ČVUT. 2003. [cit. 2005-07-21]. Dostupné na www: http://platan.vc.cvut.cz/akp2003/sbornik/12_drobikova.pdf
5. HICKEY, T. B., O'NEILL, E. T., TOVES, J. Experiments with the IFLA Functional Requirements for Bibliographic Records (FRBR) [online]. *D-Lib magazine*. 2002, vol. 8, no. 9, 15 s. [cit.2005-07-20]. Dostupné na www: www.dlib.org/dlib/september02/hickey/09hickey.html
6. HIDER, P. , MOHAMMED, R. Ensuring cataloging students meet the Mark : testing the validity of a cataloging worksheet marking scheme. *Journal of education for library and information science*. Winter 2005, vol. 46, no. 1, s. 3-20. [HIDER, 2005, s. 3-20]
7. HSIEH-YEE, I. *Cataloging and metadata education : a proposal for preparing cataloging professionals of the 21st century*. 51 s. Dostupné na www: <http://www.loc.gov/catdir/bibcontrol/CatalogingandMetadataEducation.pdf>
8. Library of Congress. Network Development and MARC Standards Office. FRBR Display Tool : version 2.0 [online]. Washington : Library of Congress. 2004. Dostupné na www: <http://www.loc.gov/marc/marc-functional-analysis/tool.html>
9. MACEWAN, A., YOUNG, T. Quality vs. Quantity : developing a systematic approach to a perennial problem. *Catalogue and index*. Summer 2004, no. 152, s. 1-7.
10. RAJAPATIRANA, B., MISSINGHAM, R. The Australian National Bibliographic Database and the Functional Requirements for the Bibliographic Database (FRBR). *The Australian library journal*. February 2005, s. 31-42.