Elektronický sborník info media PRAHA'95

Perspektivy budování knihovnických bází dat v ČR a jejich zpřístupňování

ing. Aleš Brožek,
Státní vědecká knihovna, Ústí nad Labem

Klesající ceny výpočetní techniky, uživatelsky přívětivější software i vyšší počítačová gramotnost pracovníků informačních středisek a knihoven patří mezi hlavní důvody, proč na těchto pracovištích vzniká řada výynamných bází dat. Není snadné odhadnout jejich počet, protože neexistuje žádný jejich vyčerpávající soupis (prakticky není možné jej ani sestavit, jelikož řada bází se užívá pouze interně a jejich provozovatelé nemají zájem -ke škodě možných uživatelů- je propagovat.). Publikace "Přehled bází dat využitelných v knihovnách" (1), vydaná na podzim loňského roku, jich uvádí 105. Zahrnuje však i báze vznikající mimo informační střediska a knihovny. Báze AVIS jich uvádí dokonce 428, zachycuje však i některé báze neknihovnické povahy. Nebudeme proto daleko od pravdy, budeme-li hovořit již o stovkách knihovnických bází dat v České republice. (Na tomto místě je vhodné připomenout, že převážná část knihovnických bází je stále u nás ještě vytvářena pomocí micro CDS/ISIS, tj. software vyvinutého UNESCO a distribuovaného v 80. letech do českých knihoven a informačních pracovišť Inorgou prakticky zdarma. Toto software se ukázalo jako účinný databázový nástroj a řada knihovníků se na něm mnohé naučila.)

Za knihovnické databáze můžeme označovat i ty, které sice nevznikají v knihovnách a informačních střediscích, ale jsou v nich často využívány. Je potěšitelné, že jejich producenti, většinou soukromé firmy, si uvědomují, že knihovny jsou pro ně potenciálním odběratelem. Počet takových bází roste a navíc některé firmy začínají knihovnám nabízet i software, které zlepšuje vyhledávání v bázích dat.

Chceme-li hovořit o trendech budování knihovnických bází dat, je nutné nejprve se pokusit tyto báze klasifikovat a analyzovat. Do určité míry k tomu poslouží rozbor bází uvedených v již zmiňovaném Přehledu.

Velkou skupinu knihovnických bází dat (téměř 30 %) tvoří báze adresářového typu. Možná že je to poněkud zavádějící název, protože se v nich najdou nejen adresy, ale v případě bází např. podnikatelských subjektů i údaje o počtech zaměstnanců, aktivitách aj. V této skupině nalezneme jak báze s 1 300 000 záznamy (Registr organizací), tak jen s 44 záznamy (Banky a jejich obchodní místa). Kvalitnější jsou většinou právě menší specializované databáze, kdy jejich producent je schopen snáz udržovat údaje v aktuálnějším a úplnějším stavu. Specializace přitom může být vedena např. z hlediska činnosti subjektu (Báze čistíren odpadních vod) nebo z hlediska regionálního (Katalog podnikatelů vsetínského okresu).

Do skupiny bází adresářového typu lze zařadit i podskupinu biografických bází, tj. bází s biografickými údaji. Díky REOSu, nadstavbě micro CDS/ISIS, je tato podskupina zastoupena bázemi, které vytvářejí téměř všechny Státní vědecké knihovny. Zachycují v nich údaje o osobnostech, které žily nebo žijí na území toho kterého regionu. I v tomto případě dochází k duplicitnímu až multiplicitnímu zpracování údajů o stejné osobě. Není to ale tak neekonomické jako např. v případě sestavování bází s adresami. Kromě "všeobjímajícího" Registru organizací lze nalézt např. adresy knihoven v 5(!) specializovaných bázích.

Druhá největší skupina knihovnických bází (27 %) zachycuje báze vzniklé excerpcí údajů z knih, periodického tisku včetně Sbírky zákonů. Příkladem může být báze Beletrie (která obsahuje i krátké úryvky z knih), báze ANAL Národní knihovny v Praze, do níž přispívají státní vědecké knihovny bibliografickými záznamy z regionální periodické produkce, nebo báze Vyhlášky.

V souvislosti s touto skupinou můžeme hovořit i o fulltextové skupině (9%), jako jsou báze ASPI, Jurix CD, CS Legsys-Text (všechny tři báze s plnými texty zákonů), báze Biblia Sacra, Encyklopedický slovník či jazykové slovníky. Bohužel dosud není nabízena fulltextová báze novin či časopisů.

Třetí nejpočetnější skupinou (23 %) jsou báze s údaji o fondech knihoven. Využitím bází Národní knihovny se uživatel dozví, co vychází v ČR v knižní formě (báze ČNB-Česká kniha však registruje pouze knižní publikace došlé do Národní knihovny, nevyužívá spolupráce s ostatními příjemci povinných výtisků a je tedy značně neúplná), co je vydáváno periodicky (díky spolupráci s regionálními knihovnami je báze PE relativně úplná), co v podobě norem, disertací a hudebnin.

Poměrně málo bází registruje diplomové práce. Báze s údaji o fondech menších knihoven se na první pohled mohou zdát zajímavé jen pro uživatele, kteří uvedené knihovny navštěvují. (Výjimku tvoří báze se specializovanými fondy, jako je např. báze Sudetika nebo Rakouská studovna, o které bude zájem celostátní.) Jakmile se však objeví jako součást báze označované jako souborný katalog (zatím např. báze KAKS), bude celá báze nepostradatelnou pomůckou při meziknihovní výpůjční službě. Z toho pohledu je významná již nyní báze Národní knihovny nazvaná ZP.

Z hlediska zpřístupňování knihovnických bází je nejběžnější (v 90 %) získání disket s daty a jejich překopírování na pevný disk počítače. Distribuce dat na CD-ROM (10 %) teprve postupně nabývá na významu a přístup ke knihovnickým bázím vystavovaným na hostitelském počítači je zatím dost ojedinělý. Jako příklad lze uvést vystavování bází ČTK po komutované telefonní lince v ČTK nebo jízdního řádu autobusů a letadel pomocí Videotexu.

Situace při budování knihovnických bází dat a jejich zpřístupňování se bude nadále výrazně měnit. Předpokládám, že se jednak podstatně změní poměr mezi jednotlivými skupinami bází dat, jednak ve způsobu zpřístupňování. Vzhledem k pracnosti bází adresářového typu očekávám jejich úbytek. Nebude vznikat tolik bází dat s duplicitními údaji, a pokud ano, budou to specializované báze a nikoliv "všezahrnující". U biografických bází naproti tomu počítám, že se knihovny dohodnou na spolupráci při vytváření společné báze. Očekávám, že se u těchto bází uplatní v budoucnu nové technologie a že kromě údajů vyhledaných hypertextem uvidí uživatel na obrazovce fotografii osobnosti, o kterou se zajímá, případně uslyší i její hlas.

Dále počítám i s úbytkem bází, které vznikají excerpcí periodického tisku. Bude to souviset s postupným vytvářením plnotextových bází nejčtenějších periodik, takže nebude nutné již excerpovat články z nich do speciálních bází.

Výrazný nárůst lze očekávat u bází, které obsahují údaje o fondech knihoven. Tato skutečnost je závislá na tom, jak se i menší knihovny budou vybavovat výpočetní technikou a budou převádět své katalogy do elektronické podoby. Patrně vzniknou i další souborné katalogy v elektronické podobě, a to s fondy veřejných knihoven, které využívají systém Lanius.

Vzhledem ke známé situaci s českými telekomunikacemi a s cenami za jejich využívání předpokládám, že alespoň v nejbližších několika letech bude značný počet bází distribuován na CD-ROM (případně na magnetooptických discích?). Tomuto nosiči dají zřejmě přednost před přístupem do databázových center menší veřejné knihovny. Snad se to stane významným signálem pro pronajímatele datových spojů, aby tyto služby zlevňovali a zkvalitňovali. Získávání údajů z bází dat v databázových centrech má však pro knihovny tu přednost, že tato data jsou vysoce aktualizovaná. A informace, které uživatelům v knihovnách poskytujeme, by měly být ty nejaktuálnější.

Poznámka:
(1) Přehled bází dat využitelných v knihovnách. Brno: Sdružení knihoven ČR, 1994. Tato publikace uvádí stručné charakteristiky bází dat, jež jsou v této přednášce často uváděny pouze akronymem.