

INTERNETOVÁ DATABÁZE KRAMÁŘSKÝCH TISKŮ

Iva Bydžovská

Když jsem před dvěma lety ve svém příspěvku mluvila o své představě o zpracování fondu kramářských tisků v Knihovně Národního muzea,¹ netušila jsem, že nebude trvat dlouho a budu moci představit probíhající projekt.

Grantový projekt „Vytvoření badatelského prostředí pro kramářské tisky“² byl zahájen na začátku roku 2008 pod záštitou Ministerstva kultury České republiky a v roce 2011 by jeho výstupem měla být databáze kramářských tisků pocházejících převážně z knihovny Národního muzea. Digitalizované tisky budou opatřeny metadaty, která umožní vyhledávání podle množství kritérií a poslouží zájemcům o tento druh literatury – badatelům, knihovníkům i veřejnosti – k orientaci ve zpřístupněném materiálu.

Jak je patrné, název projektu je trochu zavádějící, protože „badatelské prostředí“ v zadání grantu má sloužit především svým uživatelům, tedy odborníkům i laikům, kteří mají zájem o studium kramářských písní, modliteb a dalších tisků vzniklých v rozmezí 17. až 20. století.

To, že sbírka bude digitalizována, ale zároveň poslouží i tiskům samotným, které jsou často velmi poničeny následkem špatného skladování, použitím nekvalitního papíru či tisku a u tisků po polovině 19. století kyselostí dřevitého papíru, tak jak je tomu i u dalších tiskovin z tohoto období. Digitální kopie dokumentů budou k dispozici všem zájemcům bez toho, aby byly originály dále poškozovány při nezbytné manipulaci.

Hlavní cíle projektu jsou tedy jasné: Jedním je ochrana sbírkových předmětů, aby byly zachovány ve své fyzické podobě pro další generace. Přínos kvalitní digitalizace celých tisků v tomto ohledu není myslím třeba příliš rozvádět. Tisky budou v plánované databázi zpřístupněny v takové kvalitě, aby je bylo možno pohodlně číst a také jednotlivé tisky srovnávat, nebude tedy až na výjimečné případy třeba studovat originál ve fyzické podobě.

¹ STRAŠILOVÁ, Iva. Kramářské písně v Knihovně Národního muzea: Cíle a prostředky zpracování. In *Problematika historických fondů Čech, Moravy a Slezska 2006. Sborník z 15. odborné konference. Olomouc, 22.-23. listopadu 2006*. Ed. Rostislav Krušínský. Olomouc : Sdružení knihoven ČR Brno, Vědecká knihovna v Olomouci, 2007. s. 191–199.

² Grantový projekt: Vytvoření otevřeného badatelského prostředí pro kramářské tisky (Věda a výzkum, MK ČR, kód DC08P02OUK002); trvání projektu: březen 2008–prosinec 2011; řešitelé: Iva Bydžovská, Michal Klacek.

Druhým a mnohem diferencovanějším cílem je zpřístupnění vybraného souboru kramářských tisků (v rámci samotného grantu to bude cca 4000 exemplářů) v samostatné internetové databázi, která umožní jejich využití v širší míře než jejich zveřejnění společně s jinými druhy tiskovin, například knihami či novinami.

Před schválením grantového projektu jsme spolupracovali s firmou AiP Beroun a Národní knihovnou na případném zpřístupňování kramářských tisků v rámci projektu Manuscriptoria, naráželi jsme ale na několik problémů. Prvním je časové období, které je u kramářských tisků třeba pokrýt (převážně 17.–19. století). Manuscriptorium je určeno spíše rukopisům a starým tiskům, tedy období do přelomu 18. a 19. století.

Druhým problémem, který jsme si uvědomovali, je specifická povaha kramářských tisků. Naše zkušenosti z práce s fondem ukazují, že pro efektivní využití zvláště rozsáhlejších souborů těchto tisků je třeba vytvářet poměrně podrobné bibliografické záznamy se zvláštními údaji, jakými jsou strukturované věcné rejstříky, které umožní vyhledávat kramářské tisky podle údajů, které obsahuje text tisků. Jen málo badatelů přichází s konkrétním názvem tisků, které ostatně byly dost stereotypní (jak se lze přesvědčit na příkladě *Knihopisu*,³ jehož celý svazek je věnován názvu „Píseň“), nebo chtějí dohledat tisky podle jejich prvního verše (začátku textu).⁴ Autor je uveden pouze výjimečně, stejně jako nakladatel či knihkupec, a ani tiskař či místo a rok tisku nejsou vždy uvedeny. Přesto jsou kramářské tisky vzhledem k frekvenci, v jaké byly vydávány, zajímavým pramenem pro zjišťování působení jednotlivých tiskařů. Hlavní zájem laické i badatelské veřejnosti je ale nalézt tisky s určitým obsahem. Takové, které se vztahují k určité události, k regionu, období nebo k určitému společenskému či přírodnímu jevu.

Vzhledem k těmto požadavkům na vyhledávání se samostatná databáze „ušitá“ přímo na míru žánru kramářských tisků (a především kramářských písní) jeví jako nejlepší a nejpřehlednější řešení.

Výběr tisků k digitalizaci

Nelehkým úkolem k řešení je způsob výběru tisků do databáze, která při cca čtyřech tisících tisků, které budou v rámci projektu digitalizovány, nemůže zdaleka obsáhnout sbírku Knihovny Národního muzea, natož zahrnout známou produkci kramářských tisků.

³ *Knihopis českých a slovenských tisků od doby nejstarší až do konce XVIII. století: Díl 2, Tisky z let 1501–1800: Část 6, Písmo P – píseň*. Praha : ČSAV, 1956. Čís. 6688–12820.

⁴ První verš (incipit) je hlavním třídícím znakem stávajícího katalogu kramářských tisků. Volba incipitu jako hlavního třídícího znaku umožnila větší diferenciaci než třídění použité v *Knihopise českých a slovenských tisků*, ale pro využití sbírky z badatelského hlediska jsou velmi důležité hlavně rejstříky odkazující na obsah tisků – vlastní témata, lokalizaci, postavy, dataci.

V první variantě projektu jsme uvažovali o zpracování části sbírky kramářských písní (ty budou tvořit většinu zpřístupněných textů) podle fondu, ale později zvítězila varianta skutečně výběrové databáze, která by obsáhla hlavní typy a témata kramářských tisků v celém časovém průřezu, tedy vytvořit jakýsi jejich reprezentativní vzorek. Podle dosavadních zkušeností jsme se rozhodli, že budeme postupovat převážně podle kritérií obsahových a vytvoříme databázi na základě tematických celků, v jejichž rámci se pokusíme ukázat místní a časový rozptyl výskytu jednotlivých tisků.

Hlavní tradiční rozdělení na tisky se světským a náboženským obsahem pro takový výběr není dostatečně konkrétní ani přesné (protože se de facto jedná o rozdělení mechanické, nikoli organické – světské a náboženské prvky se u mnohých tisků prolínají, stejně jako tomu bylo v životě recipientů tohoto druhu literatury).

Dosud jsme při přípravě použili v případě tisků s převážně náboženskou (katolickou) tematikou okruhy tisků vztahující se k jednotlivým světcům – takové soubory mohou být poměrně obsáhlé, jako v případě oblíbeného svatého Jana Nepomuckého (výběr čítá něco kolem 350 tisků a znamená poměrně mravenčí a občas téměř detektivní práci se sbírkou).

V případě mariánských písní, které jsou ještě početnější, je při výběru třeba jít ještě do větších podrobností – zde se naskytá možnost zmapovat tisky s vazbou na v Čechách a na Moravě oblíbená poutní místa. U tisků, které jsou věnovány úctě k Bohu či Ježíši Kristu, je pak třeba zvolit opět jiná kritéria výběru, protože při omezeném počtu tisků, které se objeví v databázi, nelze podat jejich úplný přehled, ani v rámci jediné sbírky.

V případě světských písní jsme se rozhodli pro tematické celky. Příkladem mohou být moritáty – tisky o vraždách, sebevraždách; dále tisky o katastrofách, válkách, povodních (ty jsou, stejně jako dnes, předmětem zvláštního zájmu), o společenských jevech (manželství, móda, slavnosti, alkoholismus, vztahy rodičů s dětmi atd.).

Na jednu stranu jsou kramářské tisky poměrně stereotypní, zvláště ty, jejichž hlavním námětem je náboženská úcta – texty se často opakují v identické, nebo téměř identické podobě, ačkoli byly vydávány často a obvykle byly přebírány více tiskárny. Zdálo by se tedy, že stačí postupovat dle textového kritéria. Na druhé straně stojí ovšem fakt, že i místní a časové rozšíření textů, a tedy i stereotyp sám, je též zajímavou výpovědí o společnosti lidí, kteří si kramářské tisky kupovali a kramářské písně zpívali. I k tomu se snažíme při výběru textů přihlížet.

Dalším cílem, který jsme měli při tvorbě obsahu databáze do budoucna na mysli, je možnost srovnávat různá vydání stejného textu – tato možnost je důležitá hlavně pro pracovníky různých institucí, kteří kramářské tisky zpracovávají.

Jednotlivá vydání se často liší pouze málo a podle samotného katalogizačního záznamu není možno ověřit, zda se jedná o různá vydání, či nikoli.

To je ovšem funkce, kterou výběrová databáze nemůže zcela splnit – věříme ovšem, že databáze neustrne v podobě, kterou získá ve své počáteční fázi stávajícího grantu, ale bude se nadále rozšiřovat nejen z našich, ale i z dalších zdrojů a začne postupně plnit i tuto funkci – v podstatě funkci bibliografie kramářských tisků vydaných v našich zemích.

Dalším úkolem do budoucna, jehož splnění se zatím nedočkáme, je uveřejnění plných textů digitalizovaných dokumentů. To by umožnilo efektivnější plnotextové vyhledávání a znamenalo lepší přístupnost textů, jak pro vědecké účely (např. pro jazykový a literární rozbor textů), tak pro laickou veřejnost, a využitelnost písní například jako doplňkového materiálu při výuce ve školách – většina textů hluboko do devatenáctého století je tištěna švabachem, a tedy pro neškoleného člověka obtížněji čitelná.

Totéž je bohužel zatím pravda pro možnost využití převodu textů pomocí OCR programů, ruční přepis textů je příliš časově náročný. Je to ovšem cesta, se kterou se do budoucna musí počítat. Je to ovšem cesta, se kterou se do budoucna musí počítat a na niž musí být databáze připravena.

Tvorba metadat

K tomu, aby databáze mohla fungovat a sloužit a v budoucnu se případně rozšířit o další možnosti, je třeba zvolit vhodný způsob zpracování dat – tedy vhodný způsob zápisu metadat, dostatečně diferencovaný, aby umožnil po vytvoření databáze možnost efektivního vyhledávání ze strany uživatelů databáze a zároveň podal jasnou informaci o digitalizovaném tisku.

Zvolený zápis tedy musí respektovat zvláštnosti tohoto druhu tisků a umožňovat dostatečnou variabilitu záznamu.

Dalším požadavkem, jak již bylo zmíněno výše, je možnost rozšíření záznamu o další položky.

Zároveň ovšem musí vytvořená metadata korespondovat s pravidly bibliografického záznamu a v zájmu širšího využití institucemi také umožňovat převod do různých knihovnických a případně i muzejních systémů.

Zpracování metodiky popisu je součástí grantového projektu.

Z důvodu větší volnosti při výběru formy záznamu jsme zvolili zápis ve značkovacím jazyce XML dle formátu MASTER, který využívá badatelské prostředí Manuscriptorium. Jde o formát pro tvorbu záznamů o rukopisech a starých tiscích,

založený na mezinárodním metadatovém formátu TEI (Text Encoding Initiative),⁵ který definuje XML elementy, sloužící k popisu textu (strukturování plných textů, zápisu metadat, připojení obrazových dat). Formát MASTER, který v současné době používáme, je založen na verzi TEI 4, v budoucnu ale budeme uvažovat o přechodu k aktuální verzi TEI 5.

XML a na něm založené formáty zápisu metadat umožňují poměrnou volnost při strukturování zápisu, to ale s sebou nese jen klady. Z hlediska zpracování databáze je naopak výhodnější pevná struktura dat a důsledné používání stejných základních prvků popisu (elementů) a jejich atributů.

Další nezanedbatelnou překážkou je při zápisu přímo v XML (což je pomocí jednotlivých prvků strukturovaný text) nepřehlednost a náročnost tvorby záznamů pro katalogizátory. Je tedy třeba vyvážit ohled k samotnému materiálu, k využití dat – tedy databázovému zpracování – a praktická hlediska katalogizace.

Řešením do budoucna je vytvoření uživatelsky příjemnějšího rozhraní na základě vytvořené šablony popisu s možností přepnutí do textového způsobu zápisu, aby byl umožněn pohodlnější zápis katalogizátorům a zároveň byla zachována volnost, kterou umožňuje textový zápis přímo do struktury XML.

Tvorba takového nástroje je součástí projektu a jeho výsledná verze bude bezplatně přístupná případným dalším zájemcům.

Tvorbu databáze v rámci projektu „Vytvoření badatelského prostředí pro kramářské tisky“ zajišťuje firma AiP Beroun a bude fungovat na podobném základu jako Manuscriptorium. Kladem tvorby metadat ve formátu XML je kromě širokého pole použití a možnosti přihlídnutí k povaze materiálu také to, že nejsou závislá na jediném programu, který data zpracovává, ale lze použít či vytvořit databázi vlastní, vytvořenou přímo „na míru“, a pro zobrazení a případný tisk lze použít nástroje pro práci s HTML nebo XML, jako jsou kaskádové styly či XSLT.

⁵ Viz <http://www.tei-c.org/>.





