

# Infomedia 98

Osnova přednášky dne 21.května, v 11:15 hodin. Astronomický sál.

## **Nová generace služeb elektronického vydavatelství**

ing Stanislav Psohlavec, Albertina icome Praha s.r.o.

### ***Česká národní bibliografie***

- Obsah a zdroje České národní bibliografie.
- Roční předplatné za 2999 Kč od července 1998.
- 4x čtvrtletní vydání na CD-ROM.
- Měsíční aktualizace na Internetu.

### ***Mezinárodní desetinné třídění na CD-ROM***

- konečně v anglo-české verzi

### ***Memoria mundi series Bohemica***

#### ***digitalizace rukopisů a starých tisků.***

- Digitalizace historických dokumentů - zaběhnutá rutina.
- CD-ROM disk představující stav projektu (rok 1997) je k dispozici.
- Využití pracovišť NKČR i výsledků dosavadní práce a vývoje mimo NKČR.
- Digitalizace vybraných dokumentů mimo NKČR zdarma (na náklady MKČR).

### ***CD-ROM v digitálním archivu***

- Všechny předpoklady pro dlouhou morální životnost formátu media.
- Kopie je lepší než originál.
- Měření rezervy v čitelnosti dat - jednoduchá cesta k zajištění 100% spolehlivosti.

Přímý kontakt na vydavatelství :

Albertina icome Praha,s.r.o.  
U stadionu 157  
266 01 Beroun  
Czech Republic

tel: 00 420 - 311 - 611 237  
fax: 00 420 - 311 - 611 238  
e-mail: Jana.Brezinova@aip.cz

## **Nová generace služeb elektronického vydavatelství**

ing Stanislav Psohlavec, Albertina icome Praha s.r.o.

### **Novinky AiP**

V příspěvku chci představit novinky v našich dlouhodobých aktivitách, které by pro Vás mohly být zajímavé. Zmíním se o Česká národní bibliografii, digitalizaci rukopisů a krátce o novém vydání mezinárodního desetinného třídění a o nově vzniklém digitálním archivu.

#### ***Česká národní bibliografie***

ČNB na CD-ROM obsahuje v současné době více než 200.000 kvalitních bibliografických záznamů knih, více než 400.000 záznamů o článcích v českých novinách a časopisech, relativně úplnou databázi novin a časopisů, bázi disertací. Letos, je připojena další databáze zahraničních bohemik. Databáze českých knih je každoročně obohacena o více než 100.000 záznamů pocházejících částečně z nové produkce a částečně z retrospektivní konverze.

V roce 1998 bude retrospektiva ČNB sahát již do roku 1965 a do roku 2000 bude zachycena produkce celého dvacátého století. CD-ROM obsahuje statisíce kvalitních záznamů, které doposud využívalo pouze cca 20 veřejných knihoven, jeho využití by však mohlo být širší

Byl vytvořen systém, který veškeré požadované výstupní formáty generuje z jediného standardního formátu, kterým byl původně 'Výměnný formát' a který je postupně nahrazován formátem UNIMARC.

Koncepce beroucí za základ jediný standardizovaný formát a potřeba poskytnout zákazníkům data v různých standardizovaných tvarech přinesla složitější zpracování a nutnost vytvoření zcela individuálních programových celků v rešeršní části programu.

Protože pro starší data z retrokonverze neexistovala jednoznačná pravidla, přešla AiP i na aktivní spoluúčast na vytváření konverzních pravidel. To přineslo vytvoření systému pro definování formátových konverzí a vytvoření podpůrných prostředků pro jejich testování na reálných datech v NKČR.

Obdobné je možno říci i o ostatních databázích ČNB.

Spolupráce při vydávání ČNB i rostoucí požadavky NKČR dávno překročily rámec prostého vydání bibliografických dat. Dlouhodobě vysoké náklady na tento projekt a trvalý malý počet zákazníků se obrazily ve vysoké ceně produktu, která byla doposud cca 30.000

Kč za roční předplatné včetně aktualizací. To je pro knihovny velmi mnoho a zvláště pro SVK, které přispívají svými daty do databáze ČNB, je tato situace i morálně neúnosná. Nicméně náklady na údržbu systému jsou trvale vysoké.

Systém byl původně budován jako DOS aplikace, pak byl postupně převeden pod Windows a nyní je nutno jej převést na 32-bit řešení. Kromě toho je nutno reagovat na rostoucí požadavky zákazníků.

Dosavadní koncepce financování projektu, vycházející pouze z prodeje finálního produktu uživatelům t.j. knihovnám, je nesprávná. Přenáší totiž břímě financování celého projektu na knihovny, přičemž je takto financováno zdaleka ne pouze samotné vydávání dat.

Neudržitelnost této situace nás vedla k žádosti o podání grantu, který by dovolil poskytnout knihovnám tento produkt zdarma. Grant jsme získali, avšak v poněkud jiné formě, nutno však říci, že ve formě účelnější. Poskytnuté prostředky byly sice nižší oproti našemu návrhu, avšak zároveň bylo stanoveno, že produkt se nebude poskytovat zadarmo, ale že základní cena ČNB na českém trhu bude do 3000 Kč včetně DPH a to pro veškeré zájemce. To dovolí nabízet produkt výrazně širšímu okruhu možných uživatelů. Protože tato data jsou opravdu tím nejlepším zdrojem informací o české literatuře, je to nepochybně významný kulturní přínos.

Mohu tedy říci, že počínaje červencovým vydáním je základní cena ročního předplatného ČNB, t.j. čtyř CD-ROM disků jednotně 2999 Kč pro veškeré zájemce o tuto databázi bez rozdílu.

Měsíční aktualizace připojitelné k tomuto CD budou zveřejňovány zdarma na internetu.

### ***Mezinárodní desetinné třídění na CD-ROM***

Podarilo se nám koupit licenci pro anglická data, takže nyní vychází již druhé vydání, dovolující současnou práci s anglickým originálem a českým překladem.

Software je doplněn o nové funkce, které usnadňují pohyb ve stromové struktuře třídění.

### ***Memoria mundi series Bohemica - digitalizace rukopisů a starých tisků.***

Rutinní digitalizace rukopisů probíhá v Národní knihovně intenzivně již třetí rok.

Digitalizace historických dokumentů je časově náročná a drahá záležitost, proto je velmi závažný požadavek na standardizaci a zajištění dlouhodobé použitelnosti vznikajících dat. Od samého počátku se spolu s NKČR snažíme přispět k překonání značné roztržitosti různých iniciativ v tomto oboru a nabídnout obecně akceptovatelnou metodiku.

Základní myšlenkou je využití SGML, HTML, Internetu a Internetovských prostředků k interpretaci a zpřístupnění dat.

Projekt Memoriae Mundi series Bohemica zahájený NKČR si vzal za cíl zpřístupnit nejvýznamnější historické dokumenty uložené na území české republiky. Je třeba, aby se tento cíl netýkal pouze Národní knihovny, ale i ostatních subjektů.

Proto jsme požádali o poskytnutí dotace z rozpočtu odboru umění a knižní kultury MKČR, v němž navrhujeme poskytnout vybraným subjektům pečujícím o významné dokumenty možnost **zdarma** zdigitalizovat a zpřístupnit je tak veřejnosti pro nekomerční účely. To přinese zpřístupnění těchto dokumentů badatelské veřejnosti i ochranu těchto dokumentů.

Inicializuje se tím též rozšíření standardizace digitalizace historických dokumentů.

Tuto dotaci jsme dostali, byla již stanovena komise, komise stanovila kritéria výběru a vybrala instituce, které již byly osloveny.

Podrobněji k dosavadním výsledkům:

V rámci projektu výzkumu a vývoje *'Zpřístupnění a archivace vzácných dokumentů s využitím digitální technologie'* dotovaného MKČR byl vytvořen systém přípravy dat a kompletace digitálních dokumentů, dovolující odborným pracovníkům bez větší průpravy a znalostí počítačů připravit data popisující a identifikující historický dokument.

Programy z připravených dat generují zjednodušeně řečeno WWW stránku zpřístupňující tento dokument. Tato data jsou pak spojována s digitalizovanými obrazy stránek a zapsána na CD-ROM disk.

Data z tohoto disku jsou interpretovatelná libovolným WWW prohlížečem. Tatáž data jsou přímo zveřejnitelná ve zvolené kvalitě na Internetu.

Jsou však také interpretovatelná programem ManuFret, který elektronická data převede do formy elektronické knihy s komfortními možnostmi vyhledávání a práce s obrazem.

Konvence, které v rámci výše zmíněného projektu prosazujeme, dovolují trvalou nezávislost vznikajících dat na dalším vývoji SW a HW, přičemž dovolují tohoto vývoje trvale využívat. Je to dáno tím, že data jsou založena tak, aby využívala výhradně základních, jasně standardizovaných, mezinárodně všeobecně uznávaných a využívaných pravidel tvorby HTML dokumentů.

Myšlenka postavit celý tento systém na HTML a tedy na úzké vazbě na Internet, který jej již celosvětovým a velmi stabilním standardem, nikoli na produktech konkrétních firem, vznikla sice před více než dvěma lety, ale vývoj potvrzuje její správnost.

Podrobné informace i ukázky programových prostředků (freeware) pro tuto práci jsou shrnuty na CD-ROM, který vážným zájemcům poskytneme zdarma na požádání.

Disk byl zaslán do všech významných knihoven světa, první reakce nyní přicházejí.

Současný stav považujeme na tolik hotový a životaschopný, že je vhodné začít oslovovat další knihovny a poskytnout jim možnost využít dosažených výsledků.

Samotným cílem projektu není vytvoření jediného dobrého pracoviště a digitalizace jediného fondu historických dokumentů, ale vytvoření obecně přijatelné a prakticky využívané koncepce zaručující sdílení vytvořených elektronických dokumentů.

Existence Internetu a jeho obecné přijetí dovoluje sdílení elektronických dat bez vnučování dalších firemně a komerčně závislých produktů.

Obracejí se na nás majitelé historických dokumentů s žádostmi o pomoc při digitalizaci a naopak pracovníci NKČR ví o významných dokumentech, které jsou významnou součástí národního kulturního bohatství, které jsou však v podstatě nedostupné.

Existující kapacita digitalizačního pracoviště je plně využívána v NKČR. Rozbíhá se však práce nadruhé směně. V letošním roce vznikne další digitalizační pracoviště, které bude koncipováno jako mobilní. Tím vším dojde ke zvýšení kapacity pracoviště digitalizace historických dokumentů.

Věřím, že se stane pomocníkem především pro další knihovny.

Dále je k dispozici pracoviště pro velmi šetrné fotografování velkoplošných dokumentů ( do rozměru 2x3 m) s unikátním fixovacím zařízením, na kterém byla fotografována část sbírky plakátů Národního filmového archivu.

Důležité je, že MKČR tím, že poskytlo prostředky pro financování digitalizace mimo Národní knihovnu, umožnilo mnohonásobně využít již vytvořené hodnoty.

### ***CD-ROM v digitálním archivu***

Velké množství digitálních dat, které je produkováno při digitalizaci nevyhnutelně spěje k nezbytnosti založit digitální archiv. To jsme v NKČR, byť ve skromném rozsahu, již učinili.

Funkce musí být stejné jako u klasického archivu. Musí být zajištěn přístup k dokumentům a potřebná životnost dokumentů. U digitálního archivu, jakožto u zcela nového řešení, je velmi závažná otázka vhodného archivačního media.

Při digitální archivaci v NKČR jsme se rozhodli z mnoha důvodů pro CD-ROM. Protože máme zajištěnu možnost měřit spolehlivost a rezervu uchování digitální informace na CD-ROM disku, můžeme zajistit skutečně 100% spolehlivost trvalého uchování digitálních dat.

Doposud byly provedeny testy týkající se vyhledání optimální kombinace media, záznamové jednotky a dalších parametrů záznamu. Pro ověření nutných dlouhodobých funkcí tohoto digitálního archivu připravujeme ve spolupráci se státním archivem zrychlené klimatické testy medií, Tím chceme dosavadní znalosti doplnit o informace popisující dlouhodobé chování tohoto media.

Protože jsem se při diskusích okolo CD-ROM setkal s mnoha 'moderními' pověrami, uspořádáme ještě v červnu v Národní knihovně a ještě v létě samostatný seminář na toto téma, kde bude vhodná příležitost k představení našich názorů i k diskusi na toto téma.

Děkuji za pozornost.