

Digitálne dokumenty – otvorenie Pandorinej skrinky?

Fojtu, Andrea

Univerzita Karlova v Praze, Ústav výpočetní techniky

Dokumenty je možné definovať ako zakódovanú informáciu zapísanú na fyzický nosič. V prípade digitálnych dokumentov je binárne zakódovaná informácia oddeliteľná od fyzického nosiča; je teda možné hovoriť o kódovaní nezávislom od konkrétneho nosiča (Knoll, 1999). Digitálne dokumenty predstavujú veľmi komplexnú a často diskutovanú problematiku, ktorú možno lepšie pochopiť pomocou modelu životného cyklu dokumentov.

Životný cyklus digitálnych dokumentov

Jednotlivé fázy životného cyklu digitálnych dokumentov (objektov) definujeme podľa kľúčových oblastí problematiky. Prvý krok predstavuje tvorbu digitálnych dokumentov (primárny – pôvodný digitálny dokument a sekundárny – digitalizovaný dokument). V súčasnosti sa posúvame od problémov súvisiacich s technologickým riešením (retrospektívnej) digitalizácie ku kvalitatívnym problémom digitalizácie (súčasťou ktorých je výber materiálov pre transformáciu do digitálnej podoby (kvalita, reprezentácia, náhrada absencie)) a následne ku ďalším fázam životného cyklu digitálnych dokumentov - spracovaniu (výber najvhodnejšieho metadátového popisu, evidencie a katalogizácie), sprístupňovaniu (dostupnosť a prístupnosť), šíreniu a dlhodobej ochrany (long-term preservation) digitálnych dokumentov.

V tomto príspevku sa však chcem zamerať najmä na problematiku dlhodobej ochrany digitálnych dokumentov, ktorej sa ešte stále nedostáva zaslúženej a potrebnej pozornosti.

Dlhodobá ochrana (PREMIS)

Pod ochranou digitálnych dokumentov rozumieme súbor činností a zásahov, ktoré sú nevyhnutné pre zabezpečenie trvalého a spoľahlivého prístupu k digitálnym dokumentom bez ohľadu na čas (Pennock, 2006).

S narastajúcim počtom knižníc, archívov, vládnych agentúr, akademických inštitúcií a iných inštitúcií uchovávajúcich kultúrne dedičstvo, a ktoré investujú do budovania zbierok digitálnych dokumentov, sa zvyšuje i potreba ich dlhohobej ochrany. Nemalo by však len ísť o samoučelnú archiváciu, ale o skutočnú ochranu, sprístupnenie a šírenie.

Projekt PREMIS (The Preservation Metadata Implementation Strategies) vznikol ako odpoveď na potrebu širšieho rámca dlhodobej ochrany digitálnych dokumentov. Jeho výsledkom je veľmi dobre

štruktúrovaná sada metadátových prvkov a dátového slovníka. V neposlednom rade priniesol projekt definície rozličných možných stratégií a doporučení pre konkrétne spôsoby implementácie metadátového formátu.

Nespornou výhodou pri implementácii PREMIS je jednoduchosť jeho dátového modelu, ktorý pozostáva z piatich ťažiskových typov entít – Objekty (Objects), Udalosti (Events), Agenti (Agents), Práva (Rights) a zo samotnej Intelektuálnej entity (Intellectual Entity). Správcovia digitálnych zbierok a systémov digitálnych knižníc postupne obracajú pozornosť na skutočnosť, že dlhodobá ochrana nepozostáva len zo zachovania fyzického dátového toku, ale aj schopnosti reprezentovať obsah na všetkých úrovniach tak, ako ich definuje dátový model PREMIS. Avšak v jadre modelu nenájdete práva pre sprístupňovanie a šírenie; podrobné technické metadáta a dokumentáciu hardvérového riešenia je možné použiť podľa potreby a špecifikácie (Vojnar, 2006).

Vďaka DigiTool, systému na správu digitálnych zbierok, môžeme na Univerzite Karlovej v Prahe bezproblémovo implementovať niektoré najdôležitejšie prvky metadátového formátu PREMIS.

Digitálny univerzitný repozitár na Univerzite Karlovej v Prahe

Budovanie digitálneho univerzitného repozitára začalo aj vďaka rozvojovému projektu Ministerstva školství, mládeže a tělovýchovy „Archivace a zpřístupnění digitálních verzí kvalifikačních prací na UK (program 4a na rozvoj moderních technologií)“, ktorého cieľom je sprístupňovanie elektronických vysokoškolských záverečných prác. V rámci tohoto projektu boli definované nasledujúce rámcové ciele: stanovenie postupu pri zbere prác a metadát na jednotlivých fakultách, prepojenie digitálnej knižnice s informačným systémom univerzity, vytvorenie webového rozhrania pre vkladanie metainformácií a elektronických verzií prác s posudkami, definovanie zoznamu metadátových prvkov, návrh formátu dokumentov pre uchovávanie v digitálnej knižnici a vytvorenie všeobecných šablón pre záverečné práce (Pabianová, 2006). Momentálne sú záverečné práce vkladané prostredníctvom klientskej aplikácie (DigiTool – Meditor); každá práca je popisovaná pomocou šablóny (Dublin Core), má generované „history“, technické a PREMIS metadáta. V priebehu tohoto roku (druhá polovica 2007) prenecháme vkladanie prác študentom, ktorí tak budú činiť cez webové rozhranie. Svoje miesto si v repozitári už nachádzajú aj niektoré naše historické fondy (predovšetkým mapové zbierky).

Digitálny univerzitný repozitár na báze systému DigiTool bude predvedený v rámci vystúpenia na konferencii.

Čo sa týka dlhodobějších plánov (2007-2008), pripravujeme synchronizáciu dát s knižničným systémom Aleph (ExLibris) pre účely Jednotnej informačnej brány (JIB). Zároveň chceme prepojením systému DigiTool a Moodle zvýšiť úroveň eLearningu na univerzite. SFX nám posluží na generovanie

citácií. Vo výsledku by digitálny univerzitný repozitár mal obsahovať všetky digitálne a digitalizované dokumenty, ktoré vznikli na pôde univerzity alebo sa nejakým spôsobom dotýkajú inštitúcie (monografie (skriptá), iné študijné materiály, sylaby prednášok, videozáznamy (napríklad v prípade 3. Lekárskej fakulty videozáznamy operačných zákrokov)).

Bibliografické odkazy

1. HOROVÁ, I., CHVÁLA, R. 2003. *Digitalizace a zpřístupnění kvalifikačních prací* [online]. 2003. [cit. 2007-03-06]. Dostupné na internete: <http://knihovny.cvut.cz/akp2003/sbornik/09_horova.pdf>.
2. KNOLL, A. 1999. Digitalizace vzácných dokumentů. In *Ikaros* [online]. 1999, roč. 3, č. 9 [cit. 2007-03-06]. Dostupné na internete: <<http://www.ikaros.cz/node/2064>>. URN-NBN:cz-ik2064. ISSN 1212-5075.
3. KNOLL, A. 1999. Přednosti a rizika digitálních dokumentů. In *Ikaros* [online]. 1999, roč. 9, č. 12 [cit. 2007-03-06]. Dostupné na internete: <<http://www.ikaros.cz/node/405>>. URN-NBN:cz-ik405. ISSN 1212-5075.
4. LAVOIE, B.F. 2004. Implementing Metadata in Digital Preservation Systems: The Premis Activity. In *D-Lib Magazine* [online]. 2004, vol. 10, no. 4 [cit. 2007-03-06]. Dostupné na internete: <<http://www.dlib.org/dlib/april04/lavoie/04lavoie.html>>. ISSN 1082-9873.
5. PABIANOVÁ, I. Zveřejňování elektronických vysokoškolských kvalifikačních prací na Univerzitě Karlově. In *Digitální knihovny: provoz a budování: Sborník konference EUNIS-CZ* [online]. Špindlerův mlýn, 28. - 30. května 2006 [cit. 2007-03-06]. Dostupné na: <http://www.eunis.cz/Krkonose2006/sbornik/dokumenty/DIGITALNI_KNIHOVNY.pdf>. ISBN 80-7041-200-3.
6. PENNOCK, M. 2006. Digital preservation: continued access to authentic digital assets. In *JISC* [online]. November, 2006 [cit. 2007-03-06]. Dostupné na internete: <http://www.jisc.ac.uk/media/documents/publications/digital-pres-bp-v1-04-ab_web.pdf>.
7. VOJNAR, M. 2006. Nové standardy digitálních knihoven pro dlouhodobou ochranu. In *Knihovna plus* [online]. 2006, č. 1 [cit. 2007-02-05]. Dostupné na internete: <<http://www.knihovna.nkp.cz/knihovna61/vojnar.htm>>.