

Zpřístupňování digitálních dokumentů na UK v Praze

Lucia Tesařová, Pavel Krbec

Univerzita Karlova v Praze

Ústav výpočetní techniky

lucia.tesarova@ruk.cuni.cz

krbec@cuni.cz

INFORUM 2006: 12. konference o profesionálních informačních zdrojích
Praha, 23. - 25.5. 2006

Abstrakt. V roce 2006 zahájila Univerzita Karlova v Praze pilotní projekt budování centrálního digitálního repozitáře, který bude zastřešovat jednotlivé podprojekty a aktivity v oblasti digitalizace a zpřístupňování digitálních dokumentů na UK.

Centrální digitální repozitář UK je chápán jako soubor služeb, které univerzita členům své komunity nabízí, a to za účelem správy, zpřístupňování a rozšiřování digitálních (born digital) a digitalizovaných dokumentů, které Univerzita Karlova a členové její komunity (tj. studenti, fakulty, vědecko-pedagogičtí pracovníci, externí vědečtí pracovníci, zaměstnanci apod.) vytvořili a vytvoří.

Významným přínosem digitálního repozitáře je jednak usnadnění komunikace v oblasti vědecko-pedagogické činnosti a jednak v oblasti vědeckého publikování. Digitální repozitář zároveň umožňuje zvýšit prestiž univerzity zviditelněním výsledků práce členů její komunity.

Digitální repozitář na UK v Praze

Současný stav problematiky zpřístupňování digitálních dokumentů na UK neodpovídá aktuálním potřebám studentů a vědecko-pedagogických pracovníků a technologickým možnostem. Nejen jednotlivé fakulty, ale i různá pracoviště na fakultách přistupují k možnostem elektronické komunikace vědeckých poznatků individuálně, bez jasné koncepce sdílení a dalšího využití vytvářených digitálních dokumentů. Je zřejmé, že tyto digitální dokumenty dnes představují obrovské intelektuální bohatství univerzity a jeho racionální využití je silně žádoucí. Vyžaduje si to však koordinované úsilí celé řady pracovníků z jednotlivých pracovišť univerzity pro vybudování společného akademického komunikačního prostoru, sloužícího k efektivnímu sdílení a archivaci digitálních dokumentů a jejich zpřístupnění v rámci centrálního repozitáře při současném zabezpečení právní ochrany duševního vlastnictví.

V rámci Transformačních a rozvojových projektů pro rok 2006, podporovaných MŠMT ČR, zahájila Univerzita Karlova v Praze pilotní projekt budování centrálního digitálního repozitáře, který bude zastřešovat jednotlivé podprojekty a aktivity v oblasti digitalizace a zpřístupňování digitálních dokumentů na UK. Řešitelem projektu je Ústav výpočetní techniky UK.

Cíle projektu

- postupné budování centrálního digitálního repozitáře UK
- podpora tvorby a sběru elektronických vysokoškolských prací na UK
- zlepšení dostupnosti digitálních dokumentů vytvořených na UK
- zkvalitnění elektronické komunikace vědeckých poznatků
- zlepšení a obohacení možností vzdělávání na UK

- zpřístupňování digitálních dokumentů prostřednictvím jednotného rozhraní
- archivace a dlouhodobé uchovávání digitálních dokumentů
- zabezpečení interoperability a otevřeného přístupu
- zvýšení prestiže univerzity zviditelněním výsledků výzkumu a vědy

Postup řešení

Pro budování, správu a zpřístupňování digitálního repozitáře UK byl vybrán programový produkt izraelské společnosti Ex Libris (<http://www.exlibrisgroup.com>) – **DigiTool verze 3.0** (Digital Assets Management System), kterým Univerzita Karlova doplnila svou řadu produktů této společnosti. DigiTool je komplexní, modulární a flexibilní systém, který obsahuje všechny nástroje potřebné pro budování digitálního repozitáře. Architektura systému je založena na moderních technologiích a standardech, např. webové služby (SOAP), XML, Unicode, JPEG 2000, MARC21, kvalifikovaný Dublin Core, METS, OAI-PMH, Z39.50 a další.

Centrální digitální repozitář bude budován postupně a na jeho rozvoji a obsahovém zaměření se budou podílet nejen pracovníci z jednotlivých součástí univerzity, ale také sami studenti. Do repozitáře budou vkládány nejrůznější typy digitálních dokumentů, jako např. studijní materiály; zvukové nebo audio-vizuální nahrávky přednášek, seminářů; příspěvky z konferencí; výroční zprávy univerzity; soubory různých dat, fotografií, uměleckých prací; historicky vzácné prameny (rukopisy, listiny, matriky, mapy apod.); odborné vědecké články; vysokoškolské kvalifikační práce atd. V jistém smyslu tedy vše, co bude či již je v digitální formě na půdě univerzity vyprodukováno (jak vědecko-pedagogickými pracovníky, tak i studenty). Digitální dokumenty budou do repozitáře přidávány průběžně, a to buď hromadným importem nebo individuálně, podle doporučených pracovních postupů.

Vzhledem na různorodost digitálních dokumentů produkovaných na UK budou postupně stanoveny doporučené formáty pro ukládání digitálního obsahu a pravidla pro jejich zpřístupňování a sdílení, a manipulaci s nimi.

Pro popis digitálních dokumentů jsme zvolili formát Dublin Core. Ten je pro svou jednoduchou a přehlednou strukturu vhodný jak pro zkušené informační pracovníky a knihovníky, tak pro ostatní členy akademické obce (vědecko-pedagogické pracovníky, studenty, administrativní pracovníky apod.)

Na každé ze spolupracujících součástí UK bude určen koordinátor, který bude koordinovat a zodpovídat za pracovní postupy na příslušném pracovišti při vkládání digitálních dokumentů a metadat do repozitáře. Za celkovou koordinaci, metodiku a technickou podporu zodpovídá ÚVT UK.

Digitální sbírky

V první fázi projektu, ve spolupráci s **Ústavem dějin a Archivem Univerzity Karlovy (ÚDAUK)**, bude do repozitáře vložena **digitální sbírka nejcennějších historických pramenů** uložených v Archivu UK, tj. listiny, matriky, protokoly ze 14.-19. století. Digitalizace pramenů v AUK byla zahájena již v roce 2004 a je prováděna externí

specializovanou firmou. V současné době je zdigitalizovaných celkem 50 listin+pečetě a 4 matriky. Digitalizace dalších historických pramenů a jejich vkládání do repozitáře bude pokračovat průběžně. Odborný metadatový popis ve formátu MASTER vytváří externí pracovník Archivu UK v průběhu procesu digitalizace. Vzhledem k tomu, že pro popis digitálních dokumentů v repozitáři jsme zvolili formát Dublin Core, je potřebné metadatové záznamy ve formátu MASTER převést do formátu Dublin Core. Původní záznam (MASTER) zůstane v repozitáři zachován a bude připojen k příslušnému zkonvertovanému záznamu. Ve spolupráci s PhDr. Zdeňkem Uhlířem z Národní knihovny ČR byla vytvořena základní konverzní tabulka. Pracovníci ÚVT spolu s Archivem UK připravili konverzní pravidla, na základě kterých byl pak vyvinut konverzní program. V současné době probíhá testování a ladění konverze.

Digitální sbírka historických pramenů bude zpočátku podle požadavků ÚDAUK zpřístupňována badatelům, odborné vědecké veřejnosti, studentům a pracovníkům Archivu UK pouze v prostorách ÚDAUK. Časem bude vybraná část sbírky i veřejně přístupná.

Další součástí digitálního repozitáře bude **digitální sbírka elektronických vysokoškolských kvalifikačních prací na UK**, která je řešena jako samostatný projekt Ústavu výpočetní techniky a knihoven 11 fakult UK. Ze zástupců participujících knihoven a ÚVT byla vytvořena pracovní skupina pro řešení celé problematiky tvorby, sběru a zpřístupňování elektronických vysokoškolských prací na UK, tj. definování pracovních postupů, tvorba šablony pro psaní kvalifikačních prací, návrh legislativní úpravy vnitřního předpisu UK pro zpřístupňování kvalifikačních prací, návrh licenční smlouvy s autorem, tvorbu metadat, definování pravidel pro zpřístupňování a skupin uživatelů atd. Vzhledem k tomu, že prvotní záznam o kvalifikační práci vzniká již v aplikaci Informačního systému UK, data z IS UK budou vyexportována a vložena do digitálního repozitáře. Metadatový záznam bude v souladu s českým národním standardem, který definuje 22 metadatových prvků pro popis elektronických kvalifikačních prací.

Digitální sbírka studijních materiálů (např. přednášek, video záznamů přednášek, manuskript, záznamů experimentů, vědeckých článků a publikací atd.) bude začleněna do repozitáře v spolupráci s Matematicko-fyzikální fakultou UK a jejími pracovišti.

Předpokládáme, že v další fázi projektu bude repozitář obohacen o další zajímavé sbírky a dokumenty.

Digitální sbírky centrálního repozitáře UK budou přístupné na internetu prostřednictvím www rozhraní. S ohledem na charakter digitálního dokumentu a ochranu duševního vlastnictví bude přístup k dokumentům kontrolován podle stanovených pravidel a požadavků vlastníků dokumentů. Přístup tedy bude možné omezit např. jen pro vyjmenované osoby; osoby příslušející fakultě; osoby příslušející UK; bez omezení.

Závěrem

Budování digitálního repozitáře na UK je proces dlouhodobý, při kterém narazíme i na řadu problémů, které bude potřebné řešit. Předpokladem úspěchu projektu je dobrá spolupráce a koordinované úsilí pracovníků z jednotlivých pracovišť univerzity. Doufáme, že se nám podaří naplnit cíl, který jsme si stanovili, tj. vybudování společného akademického komunikačního prostoru, sloužícího k efektivnímu sdílení a archivaci digitálních dokumentů a jejich zpřístupnění v rámci centrálního digitálního repozitáře.