

Jakub Vavřík, 5. Ročník

27.11.2009

World Digital Library

(www.wdl.org)

Esej v rámci předmětu PV070 - Digitální knihovny

O projektu

Projekt World Digital Library (v překladu Světová digitální knihovna), je projektem United States Library Of Congress (dále jako LOC), za podpory organizace UNESCO. Projekt si klade za cíle umožnit uživatelům zdarma přistupovat online (www.wdl.org), k významným materiálům kultur celého světa, jako jsou manuskripty, mapy, vzácné knihy, hudební díla, nahrávky, filmy, tiskoviny, fotografie, architektonické návrhy a mnoho dalších ve vícejazyčných formátech.

Hlavní milníky ve vývoji WDL

- **Červen 2005:** knihovník LOC James H. navrhuje organizaci UNESCO vybudování World Digital Library.
- **Prosinec 2006:** UNESCO a LOC se spolupodílí na setkání expertů s klíčovými partnery ze všech regionů světa. Setkání expertů vede k rozhodnutí vybudovat pracovní skupinu, která by navrhla standardy a postupy pro výběr obsahu.
- **Říjen 2007:** LOC a pět partnerských institucí prezentuje prototyp budoucího WDL na zasedání UNESCO
- **Březen 2009:** WDL je zpřístupněno mezinárodní veřejnosti s obsahem o každém členském státě UNESCO.

Cíle projektu

Hlavními cíly projektu jsou:

- Podporovat mezinárodní a mezikulturní porozumění;
- Rozšiřovat objem a variace kulturního obsahu na internetu;
- Poskytnout zdroje učitelům, školákům a široké veřejnosti;
- Budovat partnerské instituce za účelem odstranění digitální propasti mezi zeměmi a uvnitř jich samotných

Popis projektu a jeho výsledků

V květnu roku 2005 navrhl ve svém projevu k Národní komisi Spojených států pro UNESCO knihovník LOC James H. Billington ideu vybudovat World Digital Library. Základní myšlenkou bylo vytvořit webovou, snadno dostupnou sbírku světového kulturního bohatství, která by dokumentovala a zviditelnila úspěchy všech zemí a kultur, a podpořila tím mezikulturní povědomí a porozumění.

UNESCO přijala tuto myšlenku jako přínos k plnění hlavních strategických plánů této organizace. Organizace vyhradila své Oddělení pro komunikace a informace, aby spolupracovalo s LOC na rozvoji projektu. V prosinci 2006 uspořádalo UNESCO a LOC setkání expertů, aby celý projekt projednali. Probrali současné problémy, kdy jen malá část kulturního obsahu jednotlivých zemí je digitalizována. Část, která digitalizována je, postrádala požadovanou kvalitu, přístupnost nebo podporu vícejazyčnosti. Většina webů, zabývajících se touto tematikou, byla těžko použitelná, a zřídka se jí podařilo oslovit uživatele, zejména ty mladé. Prioritou tedy bylo vyvinout aplikaci obsahově zajímavou, snadno ovladatelnou a uživatelsky přívětivou jak po stránce vizuální přitažlivosti, tak po stránce navigace, přístupnou širokým masám díky vícejazyčnosti. S takto vytyčenými cíli pustili se do rozvoje projektu. Za spolupráce s partnerskými knihovnami byl vyvinut prototyp WDL a tento ukázán na konferenci UNESCO 2007. Úspěšná prezentace prototypu, podpořena mnoha knihovnami, vedla k vývoji veřejně volně přístupné verzi WDL, kterou UNESCO spustilo v březnu 2009. Více než dva tucty institucí se podílely na dodání obsahu.

Projekt WDL se snaží pojmout problematiku digitálních knihoven tak trochu jinak. Je zcela nesporné, že většina projektů se soustředí na kvantitu a kvalitu obsahu, byť u mnohých projektů se na úkor kvantity občas ustoupí od kvality, a nehlídá příliš na estetickou stránku formy prezentace obsahu. Výsledkem je tak sice velmi účelná a obsáhlá knihovna, která ale zejména mladou část čtenářské obce nemá příliš šanci zaujmout. Expertní skupina, stojící za celým projektem si tohoto byla od počátku vědoma, a tak navrhla celý portál velmi přístupný a vzhledný.

Vývoj projektu zaštilili vývojáři z LOC za asistence Bibliotheca Alexandrina (Egypt). Web samotný je naimplementován v jazyce Python, s použitím frameworku Django. Protože se vývojáři snažili o co největší přístupnost, je zde mnoho funkcí implementováno s využitím technologie ajax v kombinaci s hodně populární knihovnou jQuery pro okamžité zobrazování požadovaných změn (zejména pro okamžité filtrování na úvodní mapě). Pro přístup k datům využívá web kombinace technologií JSON, XML a mySql. Fulltextové vyhledávání v knihovně je pak implementováno s použitím knihovny Lucene.

Na úvodní stránce WDL nalezneme mapu světa se zastoupením všech kontinentů. Pod mapou je velmi pěkně udělaná pohyblivá časová osa, jež slouží pro filtrování obsahu podle data vzniku. Samozřejmostí je i políčko pro vyhledávání v knihovně, a možnost volného procházení dle místa, času, tématu, typu díla či instituce, v níž se původní dílo nachází. Pokud již zvolíme nějaké dílo, zobrazí se nám pro většinu děl podobná obrazovka prezentující metadata. Ta jsou velmi pěkně a přehledně zpracována a prezentována a působí pěkným dojmem. Celkově lze říci, že knihovna je, co do prezentační stránky, velmi líbivá. Do projektu je zapojeno mnoho světoznámých knihoven (mezi

nejznámějšími Alexandrijská knihovna, ale i mnohé další), a tak je obsah skutečně velmi bohatý. Sbíрка v současné době obsahuje 158 knih, 54 tiskovin, 129 manuskriptů, 305 map, 11 filmových záběrů, 553 plakátů a fotografií a 5 zvukových nahrávek z nejrůznějších koutů světa. Nejvíce položek pochází z evropských sbírek a ze sbírek latinské Ameriky a Karibiku. Zemí s nejvyšším počtem exponátů je Ruská federace a největší část sbírky se datuje do 19. a 20. století. Z České republiky obsahuje 5 dokumentů, mezi nimi například Kodex Gigas známý spíše jako ďáblová bible. Bible je velmi kvalitně nasnímána a lze jí prolistovat stránku po stránce.

Díla zahrnutá příspěvateli do knihovny musejí splňovat určitá kritéria. Příspěvatelé by měli vybírat díla, jež nejlépe reprezentují jejich národní kulturu. Díla či kolekce, které se vztahují k historii a kultuře jiných zemí. Komise pro výběr děl může vytyčit některá širší mezinárodní témata jako např. „historie písemnictví“ a požadovat o příspěvní partnerské instituce. Díla a kolekce, která jsou zahrnuta v Memory of the World Registru. Pro finální výběr je po příspěvatelích požadováno, aby dodali informace popisující kolekci či dílo, kterým mají v úmyslu přispět, spolu s informací o důležitosti, rozsahu, intelektuálním přínosu.

Knihovna WDL konzistentně popisuje obsažená díla stejnou množinou metadat, jež jsou na webu prezentovány. Tato množina metadat dovoluje propojovat jednotlivá díla do souvislejších celků, byť pocházejí z rozličných koutů světa a kultur. Velmi zdařilé je také zpracování popisů jednotlivých děl, jež se nesnaží pouze popsat dílo samotné, ale také fakt, proč je toto dílo považováno za tak význačné. Metadata jsou prozatím standardizována a každý digitální objekt v knihovně umístěný musí být těmito daty popsán. Do budoucna se počítá s upřesněním standardu a s vyvinutím katalogizačního nástroje, pro ulehčení tvorby metadat příspěvatelům.

Momentální standard metadat vypadá následovně:

Title – název zdroje. Povinný element.

Description – shrnutí obsahu zdroje, popis toho proč je zdroj tak význačný, že si zasloužil místo ve WDL, význam zdroje z historického hlediska. Povinný element.

Creator – entita zodpovědná za vytvoření zdroje. Nepovinný element.

Publisher – entita zodpovědná za zveřejnění zdroje. Nepovinný element.

Place of publication – místo publikace zdroje. Nepovinný element

Date Created – čas nebo období spojené se vznikem zdroje. Může být uvedeno jako přesné datum nebo rozsah let. Povinný element.

Language – jazyk původního zdroje. Povinný element.

Place – geografický původ zdroje. Může být uvedena země, stát, provincie, město. Povinný element.

Type of Place – typ místa, pokud je relevantní informací vzhledem k danému zdroji (např. řeka). Nepovinný element.

Time – časový obsah díla. Může být rok nebo rozsah let. Povinný element.

Topic - číslo kategorizace zdroje dle Deweyovy knihovnické soustavy. Povinný element.

Type of Item – typ zdroje – kniha, manuskript, mapa, tiskovina, fotka, hudební nahrávka, filmová nahrávka, atd. Povinný element.

Additional Subjects (Controlled and Uncontrolled) – termíny z revidovaných slovníků (např. Library of Congress Subject Headings) a nerevidovaných slovníků, popisující zdroj. Nepovinný element.

Notes – další informace o zdroji. Nepovinný element.

Physical Description – fyzický popis zdroje (počet stran, velikost, apod.). Nepovinný element.

Collection Title – patří-li zdroj do nějaké kolekce, pak je uveden její název. Nepovinný element.

Collection Description – popis kolekce. Nepovinný element.

Institution – instituce, jež zdrojem přispěla. Povinný element.

Related Web Site – je-li zdroj dostupný online, pak zde bude uveden odkaz na něj. Nepovinný element.

WDL má také další standard pro formáty vkládaných digitálních obrázků původních kusů děl do knihovny přidaných. Digitalizované objekty musejí být v nekomprimovaném tiff 6 formátu.

V závislosti na typu zahrnutého díla jsou kladeny různé nároky na barevnou paletu obrázku a jeho oříznutí. U všech obrázků je vyžadováno 300 DPI. Je zvažován přechod na 400 či 600 DPI, pro usnadnění rozpoznávání znaků pomocí OCR software, a také pro zvýraznění obrazových nedostatků původního díla. Zkrátka kvalita digitalizovaného díla by měla odpovídat maximální technologicky možné míře daného digitalizačního centra. Knihovně poskytnuté obrázky by neměly být výrazně upravovány (zostřovány apod.). Neměli by se ořezávat okraje – pro badatele je důležité, aby si mohli být jisti, že mají např. celou stránku s informacemi a ne jen její část. Při skenování z knih je požadováno co nejmenší možné zešikmení způsobené vazbou knihy. Tiff soubory by měly mít vyplněny informace v hlavičce a spolu s obrázkem by měl být dodán XML soubor, popisující zařízení použité k digitalizaci a též jeho konfigurace. Pokud není digitalizace technicky proveditelná, je možné požádat o asistenci LOC.

WDL podporuje všech sedm oficiálních jazyků organizace UNESCO – Arabštinu, Čínštinu, Angličtinu, Francouzštinu, Portugalštinu, Ruštinu a Španělštinu. Překlad do těchto jazyků nemalou měrou

prodloužil dobu vývoje WDL, nicméně nyní lze WDL považovat za skutečně celosvětovou digitální knihovnu. První verze WDL byla překládána poloautomaticky, tedy strojem s lidskou asistencí. Nyní je v rámci projektu založena Language and Translation working group starající se o budoucí vývoj překladu a standardizaci postupů pro překlad.

Vlastní zhodnocení projektu

Na tomto projektu se mi velmi líbí uživatelské rozhraní, které se snaží zpřístupnit co nejširší laické veřejnosti. Ovládání je intuitivní a jedinou výtkou, kterou bych vůči němu měl je ne zcela viditelně umístěný odkaz na samotné dílo v digitální formě. Metadata jsou prezentována pěkným způsobem, často doplněna o poutavý popis. Na úvodní stránce lze snadno vyhledat díla z příslušného kontinentu a doby. Za velmi pěkný prvek považuji pohyblivou časovou osu pod hlavní mapou na stránce, která okamžitě filtruje díla na mapě světa. Celkově je projekt velmi povedený a jedinou větší výtkou je pouze podpora oficiálních jazyků UNESCO, což je na jednu stranu pochopitelné, na druhou stranu je ochuzena spousta lidí o spoustu zajímavých informací.

Seznam zdrojů

<http://project.wdl.org/project/english/>

<http://www.wdl.org/en>

http://css.escwa.org.lb/ictd/850/AAbid_UNESCO_World_Digital_Library.pdf.

Metadata v DC

```
<link rel="schema.DC" href="http://purl.org/dc/elements/1.1/" />
<meta name="DC.Title" content="World Digital Library" />
<meta name="DC.Creator" content="Jakub Vavřík" />
<meta name="DC.Creator.address" content="jakub.vavrik@gmail.com" />
<meta name="DC.Subject" content="WDL" />
<meta name="DC.Subject" content="Digitální knihovny" />
<meta name="DC.Subject" content="wdl.org" />
<meta name="DC.Date.created" scheme="W3C-DTF" content="2009-11-28" />
<meta name="DC.Type" scheme="DCMIType" content="Text" />
<meta name="DC.Format" scheme="IMT" content="application/pdf" />
<meta name="DC.Format.medium" content="computerFile" />
<meta name="DC.Source" scheme="URL" content="http://project.wdl.org/project/english/" />
<meta name="DC.Source" scheme="URL" content="http://www.wdl.org/en" />
<meta name="DC.Language" scheme="RFC3066" content="cz" />
```