

# Open Library

<http://openlibrary.org>

Marek Tuhársky, študent Bc., 5. Semester

Spracované dňa 3. decembra 2011

## ***Stručná charakteristika projektu***

Open Library je projekt, ktorý má za cieľ vytvoriť internetovú stránku pre každú knihu, ktorá bola kedy napísaná. Toto je hlavná myšlienka projektu. Vo svojej podstate je to katalóg kníh, kde sú popísané a sú tu uvedené aj odkazy na stránky, kde sa dajú pozrieť, prípadne kúpiť ich elektronické verzie. V súčasnosti obsahuje viac ako 20 miliónov záznamov on-line, z toho 1,7 milióna naskenovaných verzií kníh. [1]

## ***História projektu***

Open Library je projekt iniciatívy Internet Archive [2], ktorý vytvoril v roku 2007 tím na čele s Aaronom Swartzom. Motivovalo ho k tomu to, že keď sa v tomto období niekto chcel odkázať na nejakú knihu on-line, hlavné miesto kam ľudia chodili bol Amazon. Podľa neho myšlienka, že hlavným zdrojom informácií o knihách je komerčná stránka, nebolo správne, preto sa rozhodol vytvoriť stránku, ktorá zozbiera informácie od vydavateľov, čitateľov, recenzentov a knižníc. „This site will become the place where you can find interesting books and information about them, whether they're in print, out of print, out of copyright or whatever“ [3]. Na začiatku teda začali kontaktovať knižnice, aby im poslali ich vlastné katalógy. Toto samotné predstavovalo technický problém, pretože každá knižnica mala na reprezentáciu dávať rôzne formáty, prípadne rôzne jazyky. Keďže im samotným by to trvalo niekoľko ľudských životov, rozhodli sa, že projekt bude ako analógia k Wikipedii s tým, že každý môže editovať autora, informácie o knihe, prípadne písať recenzie. V máji 2010 bol projekt prerobený pod vedením Georgie Oatesovej.

## ***Ciele projektu***

Hlavným cieľom projektu je vytvoriť katalóg pre všetky knihy, ktoré boli kedy vydané, čím sa líši napríklad od projektu Gutenberg [4] tým, že ten ponúka len stiahnutie ebooku ak je voľne dostupný. Takisto nie je cieľom knihy predávať, ale len poskytnúť odkaz, kde sa dá daná kniha kúpiť. Sekundárnym cieľom je dostať užívateľa čo najbližšie ku knihe, ktorú hľadá, či už je to naskenovaná verzia alebo odkaz, kde sa dá požičať, respektíve kúpiť. Po technickejšej stránke sa projekt snaží vytvoriť tzv. Uniform Resource Identifier [5].

## ***Pridanie záznamu do katalógu***

Pridať záznam o knihe do katalógu môže prakticky hocikto, dokonca ani nemusel danú knihu čítať, pretože minimum informácií, ktoré treba vyplniť sú: meno autora, názov diela, vydavateľstvo, rok vydania a prípadne ISBN danej knihy. Ďalšie informácie, ktoré sa dajú o knihe vyplniť sú: kde bola kniha

vydaná (mesto, štát), podtitul, obsah, mená spoluautorov, grafikov, editorov, v akom jazyku je kniha napísaná, odkazy kde sa dá kniha kúpiť resp. požičať, aký je to tip knihy (pevná väzba, paperback, ...), koľko má strán, aké má rozmery, koľko váži, a nakoniec pridať obálku knihy. Na druhú stranu odstrániť nejaký záznam nemôže nikto okrem moderátorov. Ak by sa napríklad vytvoril duplicitný záznam, na konci stránky sa nachádza tlačidlo problém, kde sa zadá email a napíše sa o aký problém sa jedná.

Pre pridanie autora treba najprv pridať aspoň jednu knihu ktorú napísal. Potom sa dá samotný autor vyhľadávať a dajú sa o ňom pridávať tieto informácie: meno, dátum narodenia (prípadne úmrtia), krátky životopis alebo externé odkazy na stránky o danom autorovi, prípadne jeho umelecké mená.

## **Požičanie knihy**

Existujú tri spôsoby ako sa dá požičať knihy z Open Library. Prvým z nich je cez službu OverDrive [6], ktorá ponúka požičanie ebookov ako jednu zo svojich služieb. Požiadavkou je, že záujemca má preukaz v danej knižnici, s ktorou OverDrive spolupracuje. Druhým spôsobom je požičanie fyzickej kópie knihy tak, že vďaka spolupráci s projektom WorldCat [7], ktorý zisťuje dostupnosť danej knihy v čo najbližšej knižnici v okolí ktoré užívateľ zadá. Posledným spôsobom je možnosť stiahnutia ebooku z Internet Archive alebo inej knižnice, ktoré digitalizovali vybrané knihy z ich kolekcí. Je tu však obmedzenie, že zadarmo ju môže mať len jeden človek v daný čas.

## **Použitá technológia**

Pri vytváraní Open Library bolo niekoľko hlavných problémov. Bola potrebná databáza, ktorá bude obsahovať milióny záznamov, umožní modifikovať obsah užívateľom a bude obsahovať históriu každej zmeny. Za týmto účelom bola vytvorená tzv. *Infobase* [8], ktorá má dostatočnú flexibilitu aby tieto tri požiadavky splnila. Funguje na princípe, že všetky objekty databázy, či sa jedná o knihu, autora, alebo žánr, sú evidované ako tzv. „things“. Každá „thing“ má povinné dvojice kľúč-hodnota ako svoje vlastnosti. Napríklad „thing“ kniha môže mať kľúč „nadpis“ a hodnotu „20 tisíc míľ pod morom“ a tiež kľúč „žánr“ a hodnotu „fantasy“. Každá dvojica kľúč-hodnota je potom uložená ako verzia spolu s časom, kedy bola uložená a menom kto ju uložil. Tým sa dajú uchovávať dáta v štruktúrovanej forme, a zároveň sa dajú prechádzať v časovej postupnosti. Ďalším problémom bolo, ako umožniť užívateľom editovať jednotlivé záznamy, pretože klasické wiki [9] prostredie umožňuje pridávať len bloky textu. Za týmto účelom bolo vytvorené tzv. *Infogami* [10]. Infogami, tak ako Infobase, umožňuje ukladať štruktúrované údaje ako aj využívať dotazy z Infobase na ich triedenie. Každá Infogami stránka (s konkrétnou URL) má priradený typ. Každý typ obsahuje schému na to, ktoré polia môžu byť editované a aký formát majú tieto polia. Toto je potom použité na vygenerovanie šablón pre editovanie a prezeranie, ktoré môžu byť ďalej prispôsobené potrebám konkrétnych typov. V Infogami existujú 2 typy: jednoduché a zložené. Jednoduché typy:

- Type/int
- Type/boolean

- Type/string –jeden riadok textu
- Type/text- viac riadkov textu

Zložené typy špecifikujú, aké vlastnosti musí obsahovať „thing“ daného typu. Vytvárajú sa pomocou wiki stránky typu type/type, ktorá reprezentuje všetky typy. Definuje tieto vlastnosti:

- description typu Type/text – jednoduchý popis daného typu
- is\_multiple typu Text/boolean- rozlišuje medzi primitívnymi a zloženými typmi
- properties typu Type/property- množina vlastností „thing“ daného typu. Táto vlastnosť môže obsahovať viaceré hodnoty. Typ Type/property sa skladá z mena, jednoduchého popisu a očakávaného typu vlastnosti.

### ***Dostupnosť kníh pre zrakovo postihnutých***

Niektoré knihy sú označené ako accessible. Tento pojem znamená, že boli špeciálne vytvorené vo formáte DAISY (Digital Accessible Information SYstem) [11] a slúži k vytvoreniu digitálne nahovorených kníh pre ľudí so zrakovým postihnutím, ktorí by si radi vypočuli nejakú knihu. V Open Library existujú dva typy DAISY kníh: otvorené a chránené. Otvorené DAISY sú takisto dostupné aj vo formátoch TXT,PDF a ePub [12]. Hoci kto má prístup k týmto knihám a pri vyhľadávaní sa dá zapnúť možnosť zobrazovať len výsledky s otvorenými typmi. Na prečítanie takejto knihy slúži napríklad add-on DDReader pre Firefox [13]. Naopak chránené DAISY sa dajú momentálne otvoriť len pomocou kľúča vydávaného v rámci programu NLS Kongresovej knižnice [14]. Samotná stránka Open Library bola vytvorená podľa štandardov W3C pre zrakovo postihnutých, aj keď obsah vytvorený niektorými užívateľmi nemusí byť úplne validný. Toto sa snažia autori opravovať a vďaka tomu je stránka vhodná aj pre zrakovo postihnutých.

### ***Vlastné zhodnotenie projektu***

Open Library je neziskový projekt zameraný hlavne na katalogizáciu kníh, čím sa podľa mňa líši od ostatných podobných projektov ako sú Projekt Gutenberg alebo Google Books [15], ktoré sa snažia hlavne o naskenovanie kníh. Takisto jeho hlavnou výhodou oproti ostatným je, že užívatelia môžu tvoriť obsah, a teda má potenciál rásť rýchlejšie ako ostatné. Samozrejme, že je to aj jeho hlavná nevýhoda, pretože ľudia môžu mať nepresné informácie, alebo zámerne poskytujú chybné informácie. Prínosom sú aj voľne dostupné zdrojové kódy, ktoré umožňujú komunite vytvárať užitočné rozšírenia. Tým, že sa snaží priblížiť používateľa čo najbližšie k danej knihe bez toho aby musel vyhľadávať rôzne informácie na internete, či už o autorovi, alebo podrobnosti o vzniku diela, sa z neho stáva užitočný zdroj informácií pri hľadaní konkrétnej informácie. Slúži tiež ako dobré miesto pre odkazovanie na literatúru. Hlavný prínos projektu je, že vďaka podporám rôznych knižníc má projekt dobrú katalogizáciu starších diel, čím ich chráni pred upadnutím do úplného zabudnutia a aj užívatelia, aj keď len náhodou, sa o nich môžu niečo dozvedieť, prípadne si dané diela prečítať.

## Použité zdroje

1. – Wikipedia *Open Library* [online]. 2011 [cit. 2011-12-03]. Dostupné z WWW: <[http://en.wikipedia.org/wiki/Open\\_Library](http://en.wikipedia.org/wiki/Open_Library)>.
2. Internet Archive [online] 2011 [cit. 2011-12-03]. Dostupné z WWW: <<http://www.archive.org/about/about.php>>
3. BBC *A library bigger than any building [online] 2007* [cit. 2011-12-03]. Dostupné z WWW: < [http://news.bbc.co.uk/2/hi/uk\\_news/magazine/6924022.stm](http://news.bbc.co.uk/2/hi/uk_news/magazine/6924022.stm)>.
4. Wikipedia *Project Gutenberg* [online]. 2007 [cit. 2011-12-03]. Dostupné z WWW: <[http://en.wikipedia.org/wiki/Project\\_Gutenberg](http://en.wikipedia.org/wiki/Project_Gutenberg)>.
5. Wikipedia. *Uniform Resource Identifier* [online]. 2011 [cit. 2011-12-03]. Dostupné z WWW: <[http://en.wikipedia.org/wiki/Uniform\\_Resource\\_Identifier](http://en.wikipedia.org/wiki/Uniform_Resource_Identifier)>.
6. OverDrive *About the OverDrive* [online]. 2011 [cit. 2011-12-03]. Dostupné z WWW: <<http://www.overdrive.com/About/>>.
7. WorldCat. *About the WorldCat* [online]. 2011 [cit. 2011-12-03]. Dostupné z WWW: <<http://www.worldcat.org/whatis/default.jsp>>.
8. Open Library. *About the technology* [online]. 2011 [cit. 2011-12-03]. Dostupné z WWW: <<http://openlibrary.org/about/tech>>.
9. Wikipedia. *Wiki* [online]. 2011 [cit. 2011-12-03]. Dostupné z WWW: <<http://en.wikipedia.org/wiki/Wiki>>.
10. Open Library. *About the Infogami* [online]. 2008 [cit. 2011-12-03]. Dostupné z WWW: <<http://openlibrary.org/dev/docs/infogami#anchor5.2>>.
11. DAISY. *About the DAISY* [online]. 2011 [cit. 2011-12-03]. Dostupné z WWW: <[http://www.daisy.org/about\\_us](http://www.daisy.org/about_us)>.
12. International Digital Publishing Forum. *ePub* [online]. 2011 [cit. 2011-12-03]. Dostupné z WWW: <<http://idpf.org/epub>>.
13. Mozilla Firefox. *DDReader* [online]. 2011 [cit. 2011-12-03]. Dostupné z WWW: <<https://addons.mozilla.org/en-US/firefox/addon/ddreader/>>.
14. Library of Congress. *NLS program* [online]. 2011 [cit. 2011-12-03]. Dostupné z WWW: <<http://www.loc.gov/nls/>>.
15. Wikipedia. *Google books* [online]. 2007 [cit. 2011-12-03]. Dostupné z WWW: <[http://en.wikipedia.org/wiki/Google\\_Books](http://en.wikipedia.org/wiki/Google_Books)>.

## Metadáta v DC

<link rel="schema.DC" href="http://purl.org/dc/elements/1.1/" />

<meta name="DC.Title" content="Open Library" />

<meta name="DC.Creator" content="Marek Tuhársky" />

<meta name="DC.Language" content="sk-SK" />

<meta name="DC.Date" content="2011-12-03" />

<meta name="DC.Description" content="Informácie o Open Library, bezplatnom on-line katalógu tvorenom komunitou užívateľov" />

<meta name="Type" content="Text" />

<meta name="DC.Format" content="application/pdf" />

<meta name="DC.Source" content="[http://en.wikipedia.org/wiki/Open\\_Library](http://en.wikipedia.org/wiki/Open_Library)" />

<meta name="DC.Source" content="<http://www.archive.org/about/about.php>" />

<meta name="DC.Source" content="[http://news.bbc.co.uk/2/hi/uk\\_news/magazine/6924022.stm](http://news.bbc.co.uk/2/hi/uk_news/magazine/6924022.stm)" />

<meta name="DC.Source" content="[http://en.wikipedia.org/wiki/Project\\_Gutenberg](http://en.wikipedia.org/wiki/Project_Gutenberg)" />

<meta name="DC.Source" content="[http://en.wikipedia.org/wiki/Uniform\\_Resource\\_Identifier](http://en.wikipedia.org/wiki/Uniform_Resource_Identifier)" />

<meta name="DC.Source" content="<http://www.overdrive.com/About/>" />

<meta name="DC.Source" content="<http://www.worldcat.org/whatis/default.jsp>" />

<meta name="DC.Source" content="<http://openlibrary.org/about/tech>" />

<meta name="DC.Source" content="<http://en.wikipedia.org/wiki/Wiki>" />

<meta name="DC.Source" content="<http://openlibrary.org/dev/docs/infogami#anchor5.2>" />

<meta name="DC.Source" content="[http://www.daisy.org/about\\_us](http://www.daisy.org/about_us)" />

<meta name="DC.Source" content="<http://idpf.org/epub>" />

<meta name="DC.Source" content="<https://addons.mozilla.org/en-US/firefox/addon/ddreader/>" />

<meta name="DC.Source" content="<http://www.loc.gov/nls/>" />

<meta name="DC.Source" content="[http://en.wikipedia.org/wiki/Google\\_Books](http://en.wikipedia.org/wiki/Google_Books)" />