

# Joint Conference on Digital Libraries

## Ottawa, Kanada

13. – 17. června 2011

### Zpráva ze zahraniční cesty

*Ing. Lukáš Budínský, Knihovna Univerzity Tomáše Bati ve Zlíně*

V polovině června 2011 se v hlavním městě Kanady – Ottawě konala již jedenáctá výroční konference věnovaná problematice digitálních knihoven (The 11. Annual ACM/IEEE Joint Conference on Digital Libraries – JCDL2011). Určujícím tématem tohoto ročníku bylo „Bringing Together Scholars, Scholarship and Research Data“ což je aktuální problematika i ve většině českých univerzitních knihoven.

Hostitelem byla **University of Ottawa** a jednotlivé akce se konaly v prostorách univerzitního kampusu.



Program konference byl rozdělen do tří hlavních částí.

- Tutoriály a Doktorandské symposium (pondělí 13. června)
- Hlavní konference (úterý 14. června – čtvrtek 16. června)
- Workshopy (pátek 17. června)

Nedílnou součástí konference je prezentace posterů, které přehlednou formou představují zajímavá témata, která se bohužel nevešla do hlavního programu. Jejich informační hodnota je však vysoká, protože rozšiřují pohled na danou problematiku a přinášejí zajímavá řešení úzce související s tématy konference.

Zajímavou a ne příliš rozšířenou součástí byly tzv. „demonstrations“, kdy byly během přestávek živě prezentovány různé projekty v praxi.

## Konference

Program konference byl rozprostřen do tří dnů a pokrýval nejrůznější aspekty problematiky digitálních knihoven. Přednášky probíhaly ve dvou až třech paralelních sekcích. Mnohdy bylo složité vybrat z bohaté nabídky zajímavých témat jedno, kterého se zúčastnit.

### Session 1 – Automated methods to help our Understanding of texts

Příspěvky v této sekci měli společného jmenovatele ve snaze pochopit obsah dokumentu a textu na základě strojové analýzy. Tento postup ve výsledku umožňuje automatickou indexaci obsahu nejen na základě plného textu, ale například i přikládá váhy nadpisům nebo analyzuje strukturu dokumentu. V principu umožňuje také vytvořit tématické celky nebo popsat obsah dokumentu.

První příspěvek „*Measuring Historical Word Sense Variation*“ od Davida Bammana z Turfts University představil velmi úspěšný projekt analýzy latinských textů s využitím korpusu z téměř 400 milionů slov v textu. Při strojovém zpracování významu textu došli k pozoruhodné korelaci ve srovnání s testovací skupinou zpracovávanou manuálně.

Druhý příspěvek „*Structure Extraction from PDF-based Book Documents*“ byl zaměřen praktičtěji a věnoval se hloubkové analýze PDF dokumentů. Vytvořený program umožňuje získat z PDF dokumentů maximum informací, vytvořit metadatový záznam a obsah dokumentu. Při rostoucí oblibě PDF jako primárního formátu elektronických dokumentů jistě najde uplatnění.

### Session 5 – Finding web content again – is it possible?

Sekce věnovaná archivaci informací v Internetu. V době, kdy jsme odkázáni na dostupnost elektronických dokumentů dostupných v Internetu, je otázka persistence odkazů a zachování obsahu velmi důležitá.

Příspěvek „*How Much of the Web is Archived?*“ od Scotta Ainswortha z Old Dominion University představil projekt Memento, který uchovává a umožňuje uživatelům procházet historické kopie Internetových stránek.

Praktické možnosti využití analýzy webu v digitálních knihovnách ukazoval příspěvek „*On Identifying Academic Homepages for Digital Libraries*“ od Sujatha Dase z The Pennsylvania State University. Autoři provedli studii na základě tří otázek: (1) Kolik akademiků provozuje osobní stránky? (2) Lze jednoznačně oddělit stránky akademiků od ostatních stránek? (3) Kolik informací lze získat ze stránek vědeckých pracovníků? V projektu se autorům povedlo nejen rozpoznat s vysokou pravděpodobností domovské stránky akademiků, ale i získat strukturovaný obsah se zájmy a vědeckými tématy, kterými se daný výzkumník zabývá. Informace mohou následně sloužit k obohacení lokálních repositářů a digitálních knihoven.

### Session 9 – Show me a new way to view and discover

Aktuálním tématem v knihovnách je přizpůsobení prostředí knihovních systému potřebám uživatelů. Cílem je nabídnout kvalitní služby a dokumenty atraktivním způsobem. Klasické OPACy jsou pomalu nahrazovány moderními prostředími s mnoha nadstavbovými službami, které zajistí spokojenost uživatele a nabízí vlastnosti webu 2.0.

Zajímavým příkladem byla přednáška „*A Link-based Visual Search Engine for Wikipedia*“ od Davida Milneho z University of Waikato. Jednalo se sice o nadstavbu nad Wikipedií, nicméně systém využíval velice atraktivního prostředí, které uživatele přímo vybízelo k dalšímu objevování souvislostí a informací, které jinak zůstávají skryty.

### Session 11–Improving impact by understanding users information needs and strategie

Další sekce byla věnovaná uživatelským vzorcům chování a přizpůsobení knihoven pro naplnění jejich očekávání. Nejzajímavější byla jednoznačně přednáška George Buchanana z City University London: „*In the Bookshop: Examining Popular Search Strategies*“. Autor nebyl knihovník, ale psycholog a na své studii ukázal stále zakořeněné omyly v myšlení

knihovníků. Uživatelé nezajímá dokonalý MARC záznam ani Conspect. Jediné co je zajímavá je nalezení požadované dokumentu o kterém většinou vědí naprosté minimum informací.

## **Session 12 – Continuing the work on improving recommendation**

Tato sekce byla dalším důkazem, že knihovny se intenzivně snaží přiblížit svým uživatelům. Digitální knihovny a repositáře k tomu poskytují vhodné prostředí a zejména možnost budovat nadstavbové služby. Jednou z rychle se rozvíjejících služeb je doporučování dokumentů, které by mohly čtenáře také zajímat.

Zajímavou cestu nastínil příspěvek „*Serendipitous Recommendation for Scholarly Papers Considering Relations Among Researchers*“ od Kazunari Sugiyameho z National University of Singapore. Autoři analyzují vyhledávání napříč celou univerzitou a výsledky využívají pro doporučování vhodných dokumentů. Výzkumní pracovníci tak mohou rozšířit své obzory zejména při multioborových výzkumech a využít zkušenosti svých kolegů.

## **Sekce posterů**

Postery byly představeny v minutových příspěvcích a následně vystaveny během recepce v nádherné Tabaret Hall. Tradičně přinesly zajímavé pohledy na problematiku související s tématy konference. Celkem bylo přihlášeno a vystaveno 32 a mnohé stojí alespoň za krátkou zmínku.

Poster „*Visual Interfaces for Stimulating Exploratory Search*“ nabídl výsledky studie sledující chování uživatelů v různých prostředích. Snahou studie bylo maximálně podpořit „průzkum“ zajímavých dokumentů pomocí vizuálně laděných uživatelských rozhraní.

Zajímavý pohled na uživatelské chování při výběru dokumentů prezentoval poster „*How Children Find Books for Leisure Reading: Implication for the Digital Library*“. Jak název napovídá, autoři se inspirovali strategiemi dětí při výběru vhodných knih pro čtení a aplikovali je na prostředí digitální knihovny.

Problematikou podpory kolaborace vědců při tvorbě a sdílení dokumentů se zabýval poster „*Building a Research Social Network from a Individual Perspective*“. Je patrné, že využití sociálních sítí v akademickém prostředí je stále aktuálnější a vznikají nové a nové projekty na jejich podporu. Tento brazilský příspěvek rozšiřuje řady již existujících systémů jako je BibApp nebo MyExperiment.

## **Shrnutí**

Konference JCDL2011 byla jedinečnou příležitostí nahlédnout na nejaktuálnější dění v oboru Digitálních knihoven. Příspěvky vycházely zejména z výsledků vědeckého výzkumu a s mnohými se jistě setkáme v budoucnosti v praxi.

Zajímavé bylo sledovat široký záběr témat, která s problematikou digitálních knihoven souvisí. Velmi patrný byl trend opouštění již relativně probádané oblasti samotného ukládání elektronických dokumentů a naopak enormní snaha přiblížit digitální knihovny co nejvíce uživatelům. Velký důraz byl kladen na nová atraktivní rozhraní, efektivní možnosti vyhledávání a doporučování dokumentů pro zvýšení jejich využití.

Účast na konferenci byla přínosná zejména pro získání přehledu v projektech, které jsou řešené v zahraničí. Cenné jsou také kontakty navázané s kolegy ze zahraničních univerzit a univerzitních knihoven.

Ve Zlíně

Ing. Lukáš Budínský