

Program děkana Fakulty informatiky MU pro podporu studentských výzkumných a vývojových projektů

Závěrečná zpráva projektu (7/2008–1/2009)

Identifikační kód projektu

M	U	N	I	3	3	/	0	3	2	0	0	8
---	---	---	---	---	---	---	---	---	---	---	---	---

Základní údaje	
Garant projektu (jméno, příjmení, email)	RNDr. Petr Sojka, Ph.D., sojka@fi.muni.cz
Vedoucí projektu (jméno, příjmení, email)	
Název projektu (česky)	Redakční systém odborného časopisu s podporou exportu do digitální knihovny
WWW stránka projektu	http://www.fi.muni.cz/~xruzick7/journal-processing-system/
Doba trvání projektu v měsících	7

Výstupy projektu (uved'te, zda a jak byly naplněny/překročeny cíle projektu uvedené v návrhu a jakým způsobem byly výsledky projektu zveřejněny)

Hlavním cílem projektu bylo připravit komplexní na typografickém systému TeX založený redakční systém pro vydávání odborného časopisu. Ukořelem systému je maximálně automatizovat přípravu vydávání jednotlivých čísel pro ulehčení práce redakce, a zároveň souběžně generovat podklady pro vkládání nových čísel časopisu do digitální knihovny. Neméně důležitou součástí projektu bylo také zpracování archivních elektronicky publikovaných článků do podoby vhodné pro digitální knihovnu. Výše uvedené cíle se podařilo naplnit.

V první tříměsíční etapě projektu došlo v rámci pilotního testovacího nasazení nového systému k převedení archivních elektronicky publikovaných čísel časopisu Archivum Mathematicum do podoby vhodné pro potřeby projektu DML-CZ a vyladění a rozšíření nového redakčního systému při přípravě dalších čísel časopisu. Vznikla také první verze dokumentace.

Archivní elektronicky publikovaná čísla časopisu Archivum Mathematicum (ročníky 28–43 z let 1992–2007) byla zpracována a pro potřeby projektu DML-CZ vznikly zrestaurované PDF verze archivních čísel. Kde to bylo možné, tam byla provedena náhrada bitmapových fontů fonty vektorovými, opraven byl formát souborů. Dále byla z dostupných zdrojových textů archivních čísel extrahována metadata a převedena do formátu vhodného pro projekt DML-CZ. Vzniklé podklady byly předány týmu projektu DML-CZ.

Na konci první etapy projektu byl automatizovaný pracovní postup pro zpracování archivních čísel časopisu Archivum Mathematicum nasazen na retro-born-digital datech časopisu Applications of Mathematics vydávaného Matematickým ústavem Akademie věd České republiky. Naplánováno bylo jeho přizpůsobení a použití i na další odborné časopisy.

Nový redakční systém pro vydávání odborného časopisu byl při pilotním provozu v redakci časopisu Archivum Mathematicum odladen a obohacen o některé nové funkce (např. o generování recenzního formuláře pro Zentralblatt) a k systému vznikla prvotní verze dokumentace.

O projektu byl prezentován článek [1] v rámci DML 2008 Workshop (Birmingham, UK, 27 July 2008) [2], který byl přidružen k mezinárodní konferenci Mathematical Knowledge Management 2008 (Birmingham, UK, 28–30 July 2008) [3].

Ve druhé tříměsíční etapě projektu byl rozšířen automatizovaný pracovní postup pro zpracování retro-born-digital čísel časopisů i na publikace dalších redakcí. Průběžně byla také prováděna údržba a vylepšování nového redakčního systému pro born-digital publikování, který je úspěšně využíván redakcí časopisu Archivum Mathematicum vydávaného Masarykovou univerzitou.

Automatizovaným pracovním postupem pro zpracování archivních elektronicky publikovaných čísel časopisů byly zpracovány ročníky 39–53 (z let 1994–2008) časopisu Applications of Mathematics vydávaného Matematickým ústavem Akademie věd České republiky. Vznikly zrestaurované PDF verze archivních čísel, z článků byla získána bibliografická metadata a převedena do formátu vhodného pro digitální knihovnu.

Dále byly obdobným způsobem zpracovány ročníky 42–47 (z let 2003–2008) časopisu Acta Universitatis Palackianae Olomucensis vydávaného na Univerzitě Palackého v Olomouci a ročníky 117–127 (z let 1992–2002) časopisu Mathematica Bohemica vydávaného Matematickým ústavem Akademie věd České republiky.

Ve třetí jednoměsíční etapě projektu byly automatizovaným systémem zpracovány retro-born-digital ročníky 42–52 (z let 1992–2002) časopisu Czechoslovak Mathematical Journal vydávaného Matematickým ústavem Akademie věd České republiky. Průběžně byla prováděna údržba a vylepšování nového redakčního systému pro born-digital publikování, který byl úspěšně použit redakcí časopisu Archivum Mathematicum k vydání všech pěti čísel ročníku 2008 jmenovaného časopisu. V rámci toho byla rozšířena dokumentace nového systému v podobě uživatelské příručky pro redakci uvedeného časopisu. Kopie příručky je přiložena jako příloha této zprávy.

Uvedená příručka je k dispozici také na [4]. Tamtéž je k dispozici také obecná verze automatizovaného redakčního systému odborného časopisu s podporou exportu do digitální knihovny a automatizovaný systém pro zpracování retro-born-digital TeXem sázených článků pro import do digitální knihovny. Zpracovaná data předávaná týmu projektu DML-CZ jsou načítána a dále zpracovávána v Metadatovém editoru projektu [5], po zpracování jsou výsledky k dispozici ve veřejném repozitáři projektu DML-CZ [6].

Systém je založen na volně dostupných nástrojích, především pak na typografickém systému TeX s makronadstavbou LaTeX a vychází z redakčního systému projektu CEDRAM [7]. S přípravou metadat pomáhá open-source nástroj Tralics [8]. Zázemí projektu poskytly laboratoře NLP [9] a LEMMA [10].

Odkazy:

- [1] Růžička, Michal. Automated Processing of TeX-Typeset Articles for a Digital Library. In *DML 2008: Towards Digital Mathematics Library*. First edition. Brno : Masaryk University, 2008. s. 167–176. ISBN 978-80-210-4658-0.
- [2] Domovská stránka DML 2008 workshop. <http://www.fi.muni.cz/~sojka/dml-2008.xhtml>
- [3] Domovská stránka Mathematical Knowledge Management 2008. <http://events.cs.bham.ac.uk/cicm08/mkm08/>
- [4] Domovská stránka projektu redakční systém odborného časopisu s podporou exportu do digitální knihovny. <http://www.fi.muni.cz/~xruzick7/journal-processing-system/>
- [5] Metadatovém editoru projektu DML-CZ. <http://editor.dml.cz/>
- [6] Domovská stránka projektu DML-CZ. <http://dml.cz/>
- [7] Domovská stránka projektu CEDRAM. <http://www.cedram.org/>
- [8] Domovská stránka programu Tralics. <http://www-sop.inria.fr/apics/tralics/>
- [9] Domovská stránka laboratoře NLP. <http://nlp.fi.muni.cz/>
- [10] Domovská stránka laboratoře LEMMA. <http://www.fi.muni.cz/lemma/>

Čerpání finančních prostředků (uvádějte v tisících Kč na jedno desetinné místo)

	položka	přiděleno	čerpání
1	Stipendia pro členy řešitelského týmu	22,4	22,4
2	Cestovné	0	0
3	Literatura, materiál, drobný majetek	0	0
4	Celkem	22,4	22,4

Komentář k položkám předchozí tabulky (vysvětlete zejména případné rozdíly v plánovaném čerpání)

Finanční prostředky byly čerpány dle plánu v měsíčních stipendiích ve výši 3 200 Kč.

Vyjádření garanta (případně vedoucího) projektu

Projekt byl realizován v plánovaných termínech a v rozšířeném rozsahu. Redakční systém byl ověřen na pěti číslech časopisu Archivum Mathematicum ke spokojenosti jeho redakce a bude patrně použit redakcemi dalších časopisů (např. Kybernetika vydávaného Ústav teorie informace a automatizace Akademie věd České republiky).

Projekt splněn v plném rozsahu.

Datum:

podpis garanta