

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

Závěrečná zpráva projektu (9/2009–4/2010)

Identifikační kód projektu

M	U	N	I	3	3	/	0	3	2	0	0	9
---	---	---	---	---	---	---	---	---	---	---	---	---

Základní údaje	
Garant projektu (jméno, příjmení, email)	doc. RNDr. Petr Sojka, Ph.D., sojka@fi.muni.cz
Vedoucí projektu (jméno, příjmení, email)	
Název projektu (česky)	Rozšíření redakčního systému odborného časopisu s podporou exportu do digitální knihovny
WWW stránka projektu	http://www.fi.muni.cz/lemma/math-journal/
Doba trvání projektu v měsících	8

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 byl další rozvoj na typografickém systému TeX založeného redakčního systému pro vydávání odborného časopisu s podporou exportu do digitální knihovny (v návaznosti na projekt úspěšně řešený v rámci prvního kola Programu děkana FI MU pro podporu studentských výzkumných a vývojových projektů).

Cíl se podařilo naplnit.

Úkolem redakčního systému je maximálně automatizovat přípravu vydávání jednotlivých čísel časopisů pro ulehčení práce redakce, a zároveň připravovat kvalitně označovaná metadata pro potřeby digitálních knihoven s využitím moderních technologií.

V první etapě projektu došlo k implementaci samostatného validačního modulu pro provoz na straně digitální knihovny, který umožňuje provádět automatizovanou validaci dat předávaných redakcemi časopisů k zařazení do repositáře knihovny. Validována je struktura předávaných dat, jejich úplnost a formální správnost.

Ze strany redakcí některých časopisů vyvstal požadavek na odlehčenou verzi redakčního systému, která by umožňovala připravovat výstupy pro digitální knihovnu s pouze minimálními zásahy do dosavadních pracovních postupů. Byla proto implementována odlehčená verze redakčního systému, která tyto požadavky splňuje. Součástí je také implementace systému pracujícího na platformě MS Windows.

Odlehčený redakční systém byl prezentován jako součást přednášky „TeXové technologie pro digitální matematickou knihovnu“ pořádané CSTUGem (<http://www.cstug.cz/>) 21. listopadu 2009 na FI MU, která byla řešitelem projektu proslovena společně s doc. Sojkou.

Dále byl zahájen průzkum možností zvýšení přístupnosti výstupů redakčních systémů zařazením rozšíření, které umožní z PDF verzí článků produkovaných systémem extrahovat původní TeXové

zdrojové texty matematických výrazů. Průběžně byla také prováděna údržba a vylepšování redakčního systému, který je úspěšně využíván redakcí časopisu Archivum Mathematicum [1] vydávaného Masarykovou univerzitou.

Ve druhé etapě projektu došlo k rozšíření funkcionality samostatného modulu pro provoz na straně digitální knihovny, který umožňuje provádět automatizovanou validaci dat předávaných redakcemi časopisů k zařazení do repozitáře knihovny. Byl rozšířen okruh kontrovaných parametrů XML metadat, dále byl validátor integrován s rozhraním knihovny pro příjem dat od redakcí.

Odlehčená verze redakčního systému byla upravena dle potřeb redakce dalšího časopisu a nasazena do zaváděcího provozu v unixovém prostředí. Redakční systémy a validátor byl rozšířen o podporu některých nových TeXových matematických prostředí využívaných některými redakcemi časopisů.

Pokračovala práce na integraci rozšíření redakčních systémů pro zvýšení přístupnosti jejich výstupů obohacením PDF dokumentů produkovaných systémem o možnost extrakce TeXového zdrojového textu matematických výrazů. Do odlehčeného redakčního systému byla experimentálně přidána podpora generování XML metadat se zápisem matematických výrazů v jazyce MathML. Průběžně byla také prováděna údržba a vylepšování redakčního systému, který je úspěšně využíván redakcí časopisu Archivum Mathematicum.

V závěrečné dvouměsíční etapě projektu došlo ke stabilizaci a dalšímu vylepšení produktů předchozích etap. Redakční systémy jsou nyní nasazeny nejen v redakci časopisu Archivum Mathematicum, ale také v redakcích časopisů Acta Universitatis Palackianae Olomucensis vydávaného na Univerzitě Palackého v Olomouci [2] a Acta Mathematica Universitatis Ostraviensis vydávaného Ostravskou univerzitou [3]. Validační modul je mimo uvedených redakcí využíván také redakcemi časopisů Kybernetika [4] (vydávaného Ústavem teorie informace a automatizace Akademie věd České republiky) a Commentationes Mathematicae Universitatis Carolinae [5] (který vydává Univerzita Karlova v Praze). Jeho využití nadále poroste s připojením dalších redakcí k born-digital publikování. Validační modul byl také využit pro zpětné prověření starších retro-born-digital a born-digital dat, na základě těchto výstupů byla data zpětně upravena.

Některé výsledky projektu jsou prezentovány ve dvou člancích, které jsou v současnosti v recenzním řízení pro přijetí na DML 2010 Workshop [6] (July 7th–8th, 2010, Paris, France), který je přidružen k mezinárodní konferenci Conferences on Intelligent Computer Mathematics 2010 [7] (July 5th–10th, 2010, Paris, France).

Redakční systém je založen na volně dostupných nástrojích, především pak na typografickém systému TeX s makronadstavbou LaTeX. Pro přípravu metadat je použit open-source nástroj Tralics [8]. Pro generování PDF výstupů se používá program pdfTeX – rozšířená implementace původního systému TeX. Zdrojové texty systému jsou dostupné na domovské stránce projektu. Všechny části systému jsou licencovány pod některou ze svobodných licencí (konkrétní typ licence se u jednotlivých použitých nástrojů liší). Zázemí projektu poskytly laboratoře NLP [9] a LEMMA [10].

Odkazy:

- [1] Elektronická edice časopisu Archivum Mathematicum. <http://www.emis.de/journals/AM/>
- [2] Domovská stránka Univerzity Palackého v Olomouci. <http://www.upol.cz/>
- [3] Domovská stránka Ostravské univerzity. <http://www.osu.cz/>
- [4] Domovská stránka časopisu Kybernetika. <http://www.kybernetika.cz/>
- [5] Domovská stránka časopisu Commentationes Mathematicae Universitatis Carolinae. <http://www.karlin.mff.cuni.cz/cmuc/cmucemis/cmucemis.html>
- [6] Domovská stránka DML 2010 workshop. <http://www.fi.muni.cz/~sojka/dml-2010.html>
- [7] Domovská stránka Conferences on Intelligent Computer Mathematics 2010. <http://cicm2010.cnam.fr/>
- [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	25,6	25,6
2	Cestovné	0	0
3	Literatura, materiál, drobný majetek	0	0
4	Celkem	25,6	25,6

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

Finanční prostředky byly čerpány dle plánu v měsíčních stipendiích ve výši 3 200 Kč po dobu osmi měsíců.

Vyjádření garanta (případně vedoucího) projektu

Se způsobem řešení, výsledky a závěrečnou zprávou souhlasím.

Datum:

podpis garanta