

ZÁVĚREČNÁ ZPRÁVA
o řešení a výsledcích projektu
programu 'Informační společnost'

Rok
2009

Identifikační kód
1ET200190513

Řešitel projektu: RNDr. Jiří Rákosník, CSc.

Příjemce: Matematický ústav AV ČR, v. v. i.

Název projektu: DML-CZ: Česká digitální matematická knihovna

Doba řešení: 01/2005 - 12/2009

Charakteristika dosažených výsledků - česky

Výsledkem projektu je plnohodnotná dálkově volně přístupná digitální knihovna zahrnující 275 000 stran odborných matematických textů. Řada funkcionalit umožňuje mj. prohledávat a stahovat metadata a plné texty dokumentů a dále rozvíjet repozitář.

Charakteristika dosažených výsledků - anglicky

The project resulted in a fullfeatured digital library with a free remote access containing 275 000 pages of scholarly mathematical texts. Various functionalities enable search and download metadata and fulltexts and further develop the repository.

Statutární zástupce příjemce svým podpisem potvrzuje souhlas se zněním zprávy a údaji o financování projektu.

Datum: 13.1.2010

Podpis řešitele:



Podpis a razítko statutárního zástupce příjemce:

MATEMATICKÝ ÚSTAV AV ČR, v.v.i.
Žitná 25, 115 07 Praha 1
tel.: 222 000 711
©



Akademie v d R

ZÁVĚREČNÁ ZPRÁVA

Závěrečná zpráva je uvedena v příloze na následující straně.

1. Výsledky dosažené v roce 2009

Práce v posledním roce řešení projektu probíhala v souladu s návrhem projektu a postupu prací pro rok 2009. Pro řešitele projektu jsme uspořádali tři organizačně technické workshopy. Na prvním, který se uskutečnil v únoru v Brně, jsme stanovili podrobný seznam úkolů a postup práce v průběhu roku. Na červnovém workshopu v Praze jsme monitorovali stav a upřesnili harmonogram a zodpovědnost za jednotlivé úkoly. Na začátku prosince v Brně jsme provedli kontrolu stavu DML-CZ a určili nezbytné kroky pro úspěšné ukončení projektu a předání DML-CZ do rutinního provozu. V průběhu roku se podle potřeby uskutečňovala ad hoc vzájemná setkání a konzultace účastníků projektu. V prvních deseti měsících jsme se zaměřili na doplňování a aktualizaci obsahu digitální knihovny, dokončení vývoje funkcionalit a postupů a jejich ověřování v praxi, kontrolu a odstraňování nedostatků a přípravu přechodu na běžný provoz knihovny v dalším období. Všechny úkoly s malou výjimkou (viz bod 1.10 níže) se podařilo splnit.

1.1. Doplnění obsahu DML-CZ

V oblasti retrodigitalizace jsme dokončili úpravu dokumentů skenovaných v předchozím období. Digitalizovali jsme obsáhlou sborníkovou řadu *Winter School „Geometry and Physics“* a více než 50 ročníků časopisu *Pokroky matematiky, fyziky a astronomie* určeného širší veřejnosti včetně středoškolských učitelů. Zpracování časopisu a jeho prezentace v DML-CZ bude dokončena v rámci rutinního provozu v roce 2010. K zpřístupnění jsou připraveny první čtyři svazky ediční řady *Dějiny matematiky* vydávané Jednotou českých matematiků a fyziků. Pomocí postupů vyvinutých v roce 2008 jsme zpracovali velké množství článků z oblasti retro-born-digital (dokumenty, které jsou alespoň zčásti k dispozici v elektronické podobě, takže zpravidla odpadá skenování a zjednodušuje se pořizování metadat).

1.2. Import born-digital dokumentů

Ve spolupráci s technickými redakcemi časopisů *Archivum Mathematicum* a *Commentationes Mathematicae Universitatis Carolinae*, které pro sazbu používají dvě základní prostředí (LaTeX a AMSTeX) byly vytvořeny a odzkoušeny postupy umožňující v průběhu přípravy tiskových předloh automaticky vytvářet vstupy pro import do DML-CZ (viz [2]). Těmito postupy, přizpůsobenými technickým podmínkám a zvyklostem jednotlivých redakcí, bude v rutinním provozu průběžně aktualizován časopi-secký obsah knihovny.

1.3. Úprava prezentačního systému DSpace

Na základě zkušeností z provozu DML-CZ a prvních připomínek uživatelů byl upraven prezentační systém [7] tak, aby poskytoval informace přesnějším a přehlednějším způsobem. Zabudovaná funkce umožňuje uživatelům, aby snadno pomocí tlačítka „Feedback“ zaslali své návrhy na úpravy a zlepšení. Unikátní službu poskytuje funkcionalita nabízející výběr až 30 obsahově nejbližších článků (viz 1.6 níže).

1.4. Prezentace výsledků projektu

Výsledky projektu byly prezentovány na několika domácích i mezinárodních odborných akcích. J. Rákosník měl přednášku [18] v rámci přípravného workshopu pro Forward Look ESF v Santiagu de Compostela a DML-CZ prezentoval [8] na semináři, který pro vyučující na vysokých školách uspořádaly ČVUT a JČMF. P. Sojka byl hlavním organizátorem workshopu *DML 2009* [5] a editorem sborníku [11]. O poznat-

cích z tvorby DML-CZ přednášel na mezinárodních seminářích CASLIN 2009 [12] a RASLAN 2009 [14] a na konferenci ECDL 2009 prezentoval poster [10]. Na workshopu [16], na jehož přípravě se podílel, měl zvanou přednášku [19]. Technologii DML-CZ prezentoval na Univerzitě v Kyushu, kde japonští odborníci vyvíjejí specializovaný systém OCR pro rozeznávání matematických formulí (viz [13]). M. Šárfa měl na konferenci Knihovny současnosti 2009 přednášku [15]. O. Ulrych a J. Veselý publikovali článek [17] v časopise Pokroky matematiky, fyziky a astronomie.

Součástí přípravy na rutinní provoz byla podzimní systematická prezentace DML-CZ veřejnosti. V rámci listopadových Dnů vědy a techniky byla pro odbornou komunitu i širší veřejnost DML-CZ prezentována v Brně na Masarykově univerzitě (M. Bartošek) a v Praze na Matematicko-fyzikální fakultě a v Matematickém ústavu AV ČR (J. Rákosník). Další prezentace se uskutečnily na konci listopadu na Technické univerzitě v Liberci (J. Veselý), na MFF UK v Praze (J. Rákosník) a na Valné hromadě společnosti CSTUG v Brně (M. Růžička a P. Sojka). K rozšíření informace o DML-CZ byly a distribuovány propagační letáky.

1.5. Konsolidace báze autorit

Součástí Metadatového editoru je strukturovaná databáze autorit, ve které jsou systematicky ukládány informace o autorech všech dokumentů zařazených do DML-CZ. Databáze umožňuje vyhledávání autorů a jejich identifikaci podle nejrůznějších forem jejich jmen vyskytujících se v digitalizovaných dokumentech i v referenčních databázích Zentralblatt MATH a MathSciNet. Velká nejednotnost v prezentaci jmen autorů, neúplnost jejich forem v člancích, absence obecně přijaté jednoznačné transkripce kyrilice do latinky a pravidel pro uvádění složitých jmen autorů např. španělského nebo východoasijského původu představuje velmi obtížný problém a zároveň posiluje potřebu jeho řešení. Pro zvýšení komfortu uživatelů jsou jména autorů na předávkách dokumentů zveřejněných v DML-CZ uváděna v neúplnější a nejpřesnější známé formě.

1.6. Podobnost článků

Mezinárodní matematická komunita používá od sedmdesátých let minulého století vlastní tematické třídění odborné literatury. Systém *Mathematics Subject Classification* označuje jednotlivé tematické okruhy pětímístným kódem složeným z dvojčíferného označení hlavního oboru, jednopísmenného označení podoboru a dvojčíferného označení tématu. Zatřídění konkrétního článku v tomto systému vyžaduje znalost oboru a obsahu článku. V rámci projektu byla vyvinuta metoda analýzy textového obsahu článků, která na základě porovnání s dostatečně velkou databází spolehlivě zatříděných prací umožňuje navrhnout pravděpodobnou klasifikaci daného článku a nabídnout uživateli seznam obsahově nejpodobnějších prací v DML-CZ, popř. v dalších digitálních matematických archivech (např. ve francouzském archivu NUMDAM).

1.7. Rekompresse

Jedním z důležitých nástrojů pro tvorbu DML-CZ, jehož vývoj byl roce 2009 doveden do fáze využitelnosti v praxi, je technologie pro automatizaci ztrátové (re)kompresse skenovaných bitmap do standardu JBIG2 s adaptivní optimalizací míry ztrátovosti a komprese.

1.8. Mezinárodní aktivity

V souladu s charakterem matematiky a s původním plánem projektu jsme pokračovali v úsilí zařadit DML-CZ do širších mezinárodních struktur. Účastnili jsme se pří-

pravy tzv. Forward Look – nástroje, kterým European Science Foundation podporuje zpravidla roční přípravu podkladů, vytváření sítí vzájemných vazeb mezi potenciálními partnery a zajištění politické podpory pro financování vhodných aktivit na evropské úrovni. J. Rákosník spolu s T. Bouchem a U. Rehmannem za Komisi pro elektronické publikace Evropské matematické společnosti připravili materiál *Proposal for a Forward Look on European Virtual Library of Mathematics*. Mezitím hodnotícím sítím úspěšně prošel návrh mezinárodního projektu *The European Digital Mathematics Library (EuDML)* v Competitiveness and Innovation Framework Programme – ICT Policy Support Programme a ESF usoudila, že již není nutné realizovat příslušný Forward Look. Cílem projektu EuDML, jehož řešení v letech 2010–2012 se spolu s dalšími dvanácti zahraničními partnery účastní Masarykova univerzita a Matematický ústav AV ČR, je vytvořit multifunkční digitální knihovnu, která spojí digitální matematický obsah v jednotlivých distribuovaných repozitářích včetně DML-CZ.

1.9. Moving wall

U téměř všech časopisů se podařilo domluvit rozumný interval v rozmezí 0 až 5 let mezi publikováním časopiseckého čísla a zpřístupněním jeho obsahu v DML-CZ (tzv. moving wall). Výjimkou jsou časopisy *Czechoslovak Mathematical Journal* a *Applications of Mathematics*, kde vydavatel je vázán smlouvou o distribuci časopisů uzavřenou se Springerem, se kterým jednání o moving wall zatím nijak nepokročila.

1.10. Propojení DML-CZ s databázemi Zentralblatt MATH a MathSciNet

Tento cíl se nepodařilo naplnit v celém plánovaném rozsahu. Články a bibliografické reference v DML-CZ jsou propojeny na odpovídající položky v databázích Zentralblatt MATH a MathSciNet. To je důležité zejména v případě referencí, kde uživatel odskokem do referenční databáze může získat doplňující informace o citované práci. Navzdory oboustrannému zájmu se nepodařilo zajistit prolinkování položek v databázích Zentralblatt MATH a MathSciNet na plné texty obsažené v DML-CZ. V průběhu jednání s odpovědnými technickými pracovníky obou databází vyšla najevo dosavadní nejednotnost přístupů k zacházení s metadaty. Protože řešení tohoto problému na mezinárodní úrovni bylo zařazeno mezi cíle projektu EuDML, jehož řešení se účastní dva partneři projektu DML-CZ a do něhož bude DML-CZ integrována, rozhodli jsme se neplýtvat silami na prozatímní kroky a v rámci našeho projektu jsme otázku prolinkování Zentralblatt MATH a MathSciNet do DML-CZ neřešili.

1.11. Příprava na běžný provoz

Součástí přípravy na rutinní provoz DML-CZ od 1.1.2010 bylo dopracování technické dokumentace a uživatelských manuálů k hlavním komponentám systému DML-CZ. Po vyhodnocení různých možností bylo dohodnuto, že další provoz, udržování a rozvoj knihovny bude zajišťován Ústavem výpočetní techniky MU v Brně ve spolupráci s Fakultou informatiky MU. K tomu účelu byly připraveny odpovídající smlouvy mezi Matematickým ústavem AV ČR a těmito subjekty. Na konci listopadu byl po všech kontrolách a úpravách proveden celkový reimport dat znamenající přechod od pilotní beta-verze DML-CZ k standardní oficiální verzi, která byla k 1.1.2010 předána do správy Matematickému ústavu AV ČR a předložena veřejnosti k užívání na webové adrese <http://dml.cz>.

1.12. Použití finančních prostředků

Všichni účastníci použili v roce 2009 finanční prostředky v souladu se schváleným plánem projektu. Protože si náplň práce v posledním roce řešení (digitalizace, vytvá-

ření a úpravy metadat atd.) vyžadovala značné množství externí spolupráce studentů, významná část prostředků byla čerpána na ostatní osobní náklady. Využili jsme k tomu části prostředků nespotřebovaných na cestovné v Matematickém ústavu AV ČR. Část těchto ušetřených prostředků byla v rámci věcných výdajů použita k úhradě nákladů na digitalizaci starších ročníků časopisu *Pokroky matematiky, fyziky a astronomie*, které tak bude možné zpřístupnit v roce 2010, a na posílení technické infrastruktury knihovny. Ve fondech účelově určených prostředků žádného z příjemců nebyly uloženy prostředky, které by bylo nutno v roce 2009 dočerpat. Nebyly také provedeny žádné změny ve složení řešitelských týmů. H. Severová setrvala na mateřské dovolené, kterou nastoupila v průběhu druhého roku řešení projektu. Její úkoly v rámci týmu MU AV ČR i v roce 2009 plnil J. Rákosník.

1.13. Hodnocení průběhu a úrovně spolupráce

Řešení projektu a složení řešitelských týmů byly navrženy tak, aby role a úkoly byly důsledně distribuovány mezi jednotlivé účastníky. Podmínkou úspěchu byla jejich dobrá souhra a spolupráce, které na velmi dobré úrovni pokračovaly i v roce 2009.

1.14. Zapojení studentů

Tak jako v předchozích letech byli do řešení projektu zapojeni studenti Fakulty informatiky MU v Brně. Dílčí úkoly byly řešeny v rámci bakalářských prací [20], [21] a v rámci projektů [22], [23] a [24]. Pro odbornou přípravu studentů měla význam i jejich práce s metadaty, která jim umožnila hlouběji pochopit struktury odborné matematické literatury a získat představu o obsahu a rozsahu matematické literatury vydávané v českých zemích.

1.15. Výsledky, přednášky, prezentace

Výsledky pro uplatnění v RIV

- [1] *Česká digitální matematická knihovna DML-CZ*. Výsledek typu Z.
- [2] *DML-CZ Born-Digital Import*. Výsledek typu Z.
- [3] *DML-CZ OCR*. Výsledek typu Z.
- [4] *DML-CZ Similarity&Classification*. Výsledek typu Z.
- [5] *DML 2009*. 2. mezinárodní workshop. Ontario, 8.–9.7.2009. P. Sojka předsedal programovému výboru. <http://www.fi.muni.cz/~sojka/dml-2009.html>. Výsledek typu W.
- [6] *Metadatový editor*. Výsledek typu Z.
- [7] *Prezentační systém DML-CZ*. Výsledek typu Z.
- [8] J. Rákosník: *DML-CZ: Česká digitální matematická knihovna*. (L. Herrmann, ed.) 8. seminář Matematika na vysokých školách. Herbertov, 31.8.–2.9.2009, Jednota českých matematiků a fyziků a České vysoké učení technické, Praha, 2009, str. 140–154. ISBN 978-80-01-04375-2. Výsledek typu D.
- [9] *Rekomprese skenovaných bitmap v PDF*. Výsledek typu Z.
- [10] P. Sojka: *Workflow of a digital mathematics library*. European Conference on Digital Libraries ECDL 2009 v Corfu 27.9.–2.10. Poster. Výsledek typu O.
- [11] P. Sojka (editor): *DML 2009*. Sborník 2. mezinárodního workshopu. Ontario, 8.–9.7.2009. Masarykova univerzita, Brno, 2009. ISBN 978-80-210-4781-5. http://is.muni.cz/publikace/publikace_simple.pl?lang=en;id=839287 Výsledek typu O.

- [12] P. Sojka: *An experience with building digital open access repository*. In: Proceedings of 16th International Seminar CASLIN 2009, Institutional Online Repositories and Open Access. University of West Bohemia, Plzeň, 2009, str. 74–78. ISBN 978-80-7043-806-0. <http://www.knihovna.zcu.cz/Caslin/Caslin09.pdf>
Výsledek typu D.
- [13] P. Sojka: *Digitisation Workflow in the Czech Digital Mathematics Library*. Math-for-Industry. COE Lecture Note Series: Kyushu University (2009), str. 272–280. ISSN 1881-4042. Výsledek typu J.
- [14] P. Sojka: *Languages of Mathematics -- Random Walking in the Mathematics of Languages*. In: Proceedings of RASLAN 2009. Masaryk University, 2009, str. 127–133. ISBN 978-80-210-5048-8. Výsledek typu D.
- [15] M. Šárfa: *Metadatový editor pro digitální knihovny*. Knihovny současnosti 2009. 17. konference Sdružení knihoven ČR. Seč u Chrudimi, 23.–25. června 2009. str. 140–154. ISBN 978-80-86249-54-4. <http://www.sdruk.cz/sec/2009/sbornik>.
Výsledek typu D.
- [16] *UK Mathematics Content Workshop*. The Open University, Milton Keynes, UK, 9.7.2009. P. Sojka členem programového výboru.
<http://groups.google.com/group/uk-math-content-2009/web/home>
Výsledek typu W.
- [17] O. Ulrych, J. Veselý: *DML-CZ – současnost a budoucnost*. Pokroky matematicky, fyziky a astronomie, 54 (2009), no. 3, 224–231. Výsledek typu J.

Přednášky, prezentace

- [18] J. Rákosník: *DML-CZ: Asks and Bids*. Přednáška na přípravném workshopu European Science Foundation pro Forward Look European Virtual Library of Mathematics. Santiago de Compostella, 12.–15.3.2009.
- [19] P. Sojka: *Semantics, search and digital libraries*. Přednáška na mezinárodním workshopu UK Mathematics Content Workshop. The Open University, Milton Keynes, UK, 9.7.2009.
<http://www.fi.muni.cz/usr/sojka/presentations/sojka-uk-math-pres2009.pdf>

Další související výsledky

- [20] R. Hatlapatka: *JBIG2 komprese*. Bakalářská práce. Fakulta informatiky Masarykovy univerzity v Brně, 2009.
- [21] L. Lalinský: *Hledání citačních záznamů*. Bakalářská práce. Fakulta informatiky Masarykovy univerzity v Brně, 2009.
- [22] M. Růžička (řešitel): *Redakční systém odborného časopisu s podporou exportu do digitální knihovny*. Projekt MUNI/33/03/2009 v Programu děkana Fakulty informatiky MU pro podporu studentských výzkumných a vývojových projektů, 2008/2009. <http://www.fi.muni.cz/~xruzick7/journal-processing-system/>.
- [23] P. Sojka, M. Růžička: *Rozšíření redakčního systému odborného časopisu s podporou exportu do digitální knihovny*. Pokračování projektu [22], 2009–2010.
- [24] R. Řehůrek: *Podobnost matematických vzorců*. Projekt MUNI/E/0084/2009 v Programu rektora Masarykovy univerzity na podporu tvůrčí činnosti studentů – Kat. E – Podpora výzkumné činnosti studentů v oborech lékařství, zdravotnictví, přírodovědy a informatiky. 2009–2010.

2. Zhodnocení řešení projektu

Můžeme konstatovat, že hlavního cíle projektu – vytvoření České digitální matematické knihovny – bylo dosaženo, přičemž v některých ohledech výsledek významně překonává původní plány při zachování rozpočtu. Výsledkem společné pětileté činnosti pěti partnerů je plnohodnotná dálkově volně přístupná digitální knihovna <http://dml.cz> zahrnující 275 000 stran odborných textů (26 000 článků a kapitol) téměř 10 000 autorů, vybavený řadou funkcionalit a umožňující

- prohlížet a stahovat úplná metadata obsažených dokumentů,
- prohlížet a stahovat plné texty článků a knižních kapitol, které nejsou blokovány podmínkami vydavatelů (tzv. moving wall),
- vyhledávání v rejstřících autorů, názvů, MSC klasifikací,
- jednoduché a pokročilé vyhledávání v metadatech i v plných textech,
- prohlížet nabídku obsahově příbuzných článků sestavenou podle míry podobnosti vypočtené na základě textové analýzy.

Plné texty v DML-CZ jsou prezentovány ve formátu PDF s digitálním podpisem prokazujícím autenticitu dokumentu. Každý článek je opatřen předšádkou obsahující základní metadata a copyrightovou doložku označující vlastníka autorských práv a upozorňující uživatele, že dokument může využívat jen pro osobní potřebu.

Tabulka. Přehledná informace o obsahu DML-CZ k datu ukončení projektu

Zdroj	Od-do	čísel/ svazků	stran	článků / kapitol
<i>-- časopisy</i>				
Acta Mathematica Universitatis Ostraviensis	1993-2008	16	1 711	169
Acta Universitatis Palackianae Olomucensis. Mathematica	1960-2008	47	9 684	769
Applications of Mathematics	1956-2008	313	27 258	2 259
Archivum Mathematicum	1965-2009	175	13 645	1 334
Časopis pro pěstování matematiky a fyziky	1872-1950	353	35 511	4 060
Časopis pro pěstování matematiky	1951-1990	160	19 186	2 119
Commentationes Mathematicae Universitatis Carolinae	1960-2009	201	36 276	3 248
Czechoslovak Mathematical Journal	1951-2008	230	42 947	3 510
Kybernetika	1965-2008	293	30 038	2 413
Mathematica Bohemica	1991-2008	69	8 556	745
Mathematica Slovaca	1951-2007	238	24 922	2 423
Celkem časopisy		2 095	249 734	23 049
<i>-- sborníky</i>				
EQUADIFF	1962-2005	15	5 212	646
Nonlinear Analysis, Function Spaces and Applications	1978-2006	8	1 958	85
TopoSym	1961-1976	6	2 362	459
Winter School on Abstract Analysis	1976-2009	36	4 699	693
Winter School "Geometry and Physics"	1981-2009	29	4 245	396
DML - Digital Mathematics Library	2008-2009	2	307	29
Celkem sborníky		96	18 783	2 308
<i>-- monografie</i>				
Bolzano - kolekce knih	1804-1963	25	5 287	335
Dějiny matematiky - kolekce knih	1994-1996	4	786	55
ostatní knihy	1874-1979	3	630	37
Celkem monografie		32	6 703	427
Celkem DML-CZ		2 223	275 220	25 784

DML-CZ nyní obsahuje prakticky celou českou matematickou odbornou časopiseckou produkci (8 odborných časopisů včetně jejich případných předchůdců) a jeden slovenský časopis, sborníky pěti významných řad mezinárodních konferencí pořádaných českými a slovenskými institucemi a 28 monografií včetně kolekce 25 děl Bernarda Bolzana (nebo o něm). Dalších téměř 20 000 stran textů je rozpracováno a bude zpřístupněno v DML-CZ během příštího roku.

V DML-CZ se vyskytují texty v celkem 12 jazycích, přičemž nejsou výjimkou dokumenty zahrnující texty dvou až čtyř různých jazyků včetně ruštiny. Tato mnohojazyčnost představovala řadu zajímavých problémů, které bylo třeba vyřešit (automatické rozlišování jazyka, překlady názvů do angličtiny, transkripce jmen do latinky apod.).

Starší časopisy se zpravidla vyznačují složitou strukturou a zahrnují odborné nejen matematické články, ale také články s fyzikální, chemickou, astronomickou tematikou, recenze, zprávy a spolkové informace i články s historickým či politickým zaměřením.

Časopisecké články jsou opatřeny strukturovanými seznamy bibliografických referencí představujících celkem téměř 160 000 položek.

Za **nejdůležitější výsledky** považujeme sedm komplexních aplikací typu ověřených technologií, které společně tvoří funkční jádro České digitální matematické knihovny, umožňují její průběžnou aktualizaci a další rozvoj. Zároveň představují účinné nástroje, které lze využít při vytváření jiných digitálních repozitářů. Teoretické a praktické aspekty těchto nástrojů a další zkušenosti z řešení projektu byly prezentovány v řadě publikací, které vznikly v průběhu řešení projektu.

Česká digitální matematická knihovna DML-CZ [1] představuje komplexní řešení typu ověřené technologie, které samo zahrnuje následujících šest nezávislých prvků typu ověřené technologie.

Metadatový editor [6] je komplexní systém pro podporu všech činností při tvorbě článkově orientované digitální knihovny, umožňující distribuovaným strukturovaným dálkovým přístupem provádět integraci digitalizovaných stránek do hierarchické struktury, tvorbu, úpravu a automatizovanou verifikaci popisných metadat včetně strukturovaných a vzájemně prolinkovaných bibliografických referencí a generování pdf souborů pro prezentaci. Metadatový editor současně slouží pro strukturovanou archivaci veškerých digitálních dat.

Prezentační systém DML-CZ [7] je komplexní systém pro prezentaci digitální knihovny a její plné využívání prostřednictvím dálkového přístupu, vytvořený na základě open source systému DSpace.

DML-CZ OCR [3] představuje technologii vyvinutou pro optické rozpoznávání matematických textů založená na komerčním produktu FineReader 8 SDK a na systému InftyReader. DML-CZ OCR provádí rozpoznávání textů s optimalizací pro jazyky daného odstavce.

DML-CZ Born-Digital Import [2] je komplexní systém založený na volně dostupném software (TeX, Tralics-CEDRICS), umožňující technickým redakcím matematických časopisů sázených typografickým programem TeX automatizovanou tvorbu optimalizovaných a validovaných výstupů pro digitální knihovnu ve formě PDF, článkových metadat a bibliografických referencí.

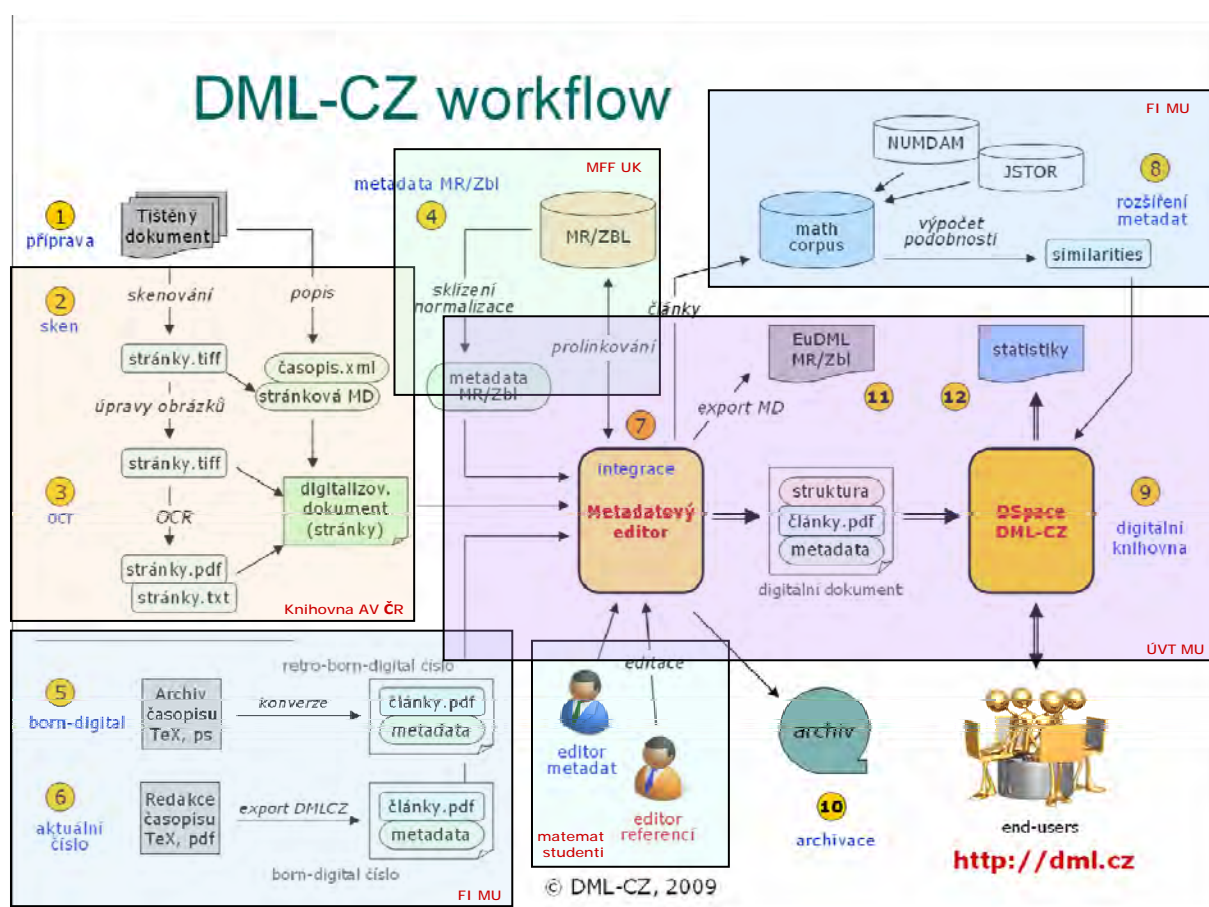
DML-CZ Similarity&Classification [4] je robustní technologie umožňující metodami strojového učení na základě plných textů článků naučit top-level klasifikátor třídit

texty podle Mathematics Subject Classification a počítat sémantickou podobnost matematických článků.

Rekomprese skenovaných bitmap v PDF [9] je technologie pro automatizaci ztrátové (re)komprese skenovaných bitmap do standardu JBIG2 s adaptivní optimalizací míry ztrátovosti a komprese.

Metadatový editor a Prezentační systém DML-CZ jsou dílem Ústavu výpočetní techniky MU, další čtyři technologie vznikly na Fakultě informatiky MU. Celá DML-CZ je pak společným dílem všech pěti spoluřešitelů projektu, z nichž každý měl při jeho tvorbě nezastupitelnou úlohu.

Následující schéma představuje přehledné znázornění činností, vazeb a rolí v rámci řešení projektu:



Systém DML-CZ včetně veškerých dat byl k 1.1.2010 předán Matematickému ústavu AV ČR, který bude mít v budoucnosti na starosti správu, udržování a rozvoj této digitální knihovny včetně zajištění potřebných finančních prostředků. V průběhu projektu rozvinutá výborná spolupráce všech partnerů však bude pokračovat i nadále. Ústav výpočetní techniky MU a Fakulta informatiky MU v Brně budou technicky zajišťovat správu, udržování a rozvoj DML-CZ na základě smluv uzavřených s Matematickým ústavem AV ČR. Skenování a grafické úpravy bude v případě potřeby zajišťovat Digitalizační středisko Knihovny AV ČR, po odborné matematické stránce bude s MÚ AV ČR spolupracovat tým Matematicko-fyzikální fakulty UK v Praze. Podle smluvního ujednání o poskytování obsahu uzavřených s vydavateli časopisů budou v roce 2010 vyhodnoceny finanční nároky na provoz DML-CZ a dohodnuta finanční

participace poskytovatelů obsahu. Odborným dohledem a propagací bude DML-CZ podporovat také Česká matematická společnost, sekce Jednoty českých matematiků a fyziků.

Systematická propagace od zveřejnění pilotní části projektu v r. 2008 spolu s rozšiřujícím se obsahem i nabídkou služeb vede k rostoucímu zájmu veřejnosti o DML-CZ. Přispívá k tomu i skutečnost, že DML-CZ indexuje Google, kde odkazy na položky v DML-CZ bývají na předních místech. V období 1.1.–27.12.2009 bylo v rámci statistik Google Analytics zaregistrováno přes 73 tisíc přístupů k DML-CZ ze 177 zemí celého světa (nejčastější přístupy jsou z ČR, USA, Německa, Indie, Slovenska, Číny, Francie, Spojeného království, Polska a Ruska). Denní návštěvnost má trvale vzestupný trend, v posledních týdnech roku 2009 se pohybovala kolem 500. Pozitivní je i zájem vydavatelů časopisů a sborníků konferencí. Na návrh Společnosti Otakara Borůvky je zvažováno zařazení sekce věnované odkazu předních osobností české matematiky.

V průběhu roku 2010 bude do DML-CZ zařazen časopis *Pokroky matematiky, fyziky a astronomie* vydávaný Jednotou českých matematiků a fyziků. V dalším období se počítá se zařazením odborných pedagogických časopisů *Matematika-Fyzika-Informatika* a *Učitel matematiky*, dalších řad sborníků konferencí, dalších svazků edice *Dějiny matematiky* a v závislosti na postoji vlastníků autorských práv také důležité monografie. V souladu s deklarovanou politikou EU bude připravováno zařazování doktorských disertací do DML-CZ. Již nyní je zřejmé, že DML-CZ je přínosem nejen pro odbornou matematickou komunitu, ale i pro pedagogy, kteří mohou snadného přístupu k literatuře využít k efektivnější práci se studenty, a v neposlední řadě i pro knihovníky.

Budoucí zapojení DML-CZ do Evropské digitální matematické knihovny v rámci projektu EuDML zmíněného v části 1.8 nepochybně ještě zvýší zájem o českou matematiku a odborné texty publikované institucemi v ČR.

Souhrnné informace, materiály a webové odkazy týkající se průběhu celého projektu, jsou uvedeny na webových stránkách projektu <http://projekt.dml.cz>.

VÝKAZ
o použití p id lených finan ních prost edk v roce 2009
na øešení projektu
programu 'Informaèní spoleènost'

Identifika ní kód projektu: 1ET200190513

P íjemce: Matematický ústav AV ÈR, v. v. i.

1. Investi ní náklady (IN)

specifikace položky	výkaz 2009 v tis. K
celkem investi ní náklady	0

2. Mzdové náklady (MZD)

a) mzdy a platy pro pracovníky

jméno a p íjmení pracovníka	výkaz 2009 v tis. K
celkem mzdy a platy	0

b) odm òny

jméno a p íjmení pracovníka	výkaz 2009 v tis. K
Jiøí Rákosník	30
Helena Severová	0
Technicko-administrativní personál	0
celkem odm òny	30

3. Ostatní osobní náklady (OON)

specifikace položky OON	výkaz 2009 v tis. K
vytváøení digitálních dat a metadat pro born-digital èasopisy	15
pøíprava digitálních podkladù èasopisu Pokroky MFA	2
tvorba, kontrola a úprava metadat èasopisù a sborníkù	41
celkem ostatní osobní náklady	58

4. V èné náklady (VN)

položky v èných náklad	výkaz 2009 v tis. K
Drobný majetek	3
Provoz a údržba	1
Další provozní náklady	0
Služby	24
Zveøejnìní výsledkù	9
Cestovné	9

Doplňkové náklady	36
Povinné zákonné odvody	11
celkem věcné náklady	93

Účelové finanční prostředky na řešení projektu (v tis. Kč)

rok	IN	MZD	OON	VN	celkem
přiděleno 2009	0	30	52	99	181
čerpání 2009	0	30	58	93	181

Celkové uznané náklady na řešení projektu ze všech zdrojů financování

Účelové prostředky od AV ČR (v tis. Kč)					
2005	2006	2007	2008	2009	celkem
437	181	381	181	181	1361
Ostatní veřejné zdroje (v tis. Kč)					
0	0	0	0	0	0
Neveřejné zdroje financování (v tis. Kč)					
0	0	0	0	0	0
Celkové uznané náklady na řešení projektu (v tis. Kč)					
437	181	381	181	181	1361

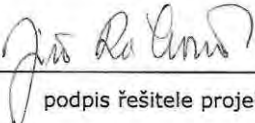
Povinnou přílohou výkazu je tištěná sestava jednotlivých účetních položek za rok 2009 potvrzená odpovědnou osobou!

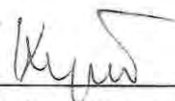
Datum: 11. 1. 2010

Zpracoval: 

Tel: 222 090 762

E-mail: RAKOSNIK@MATH.CAS.CZ


podpis řešitele projektu

MATEMATICKÝ ÚSTAV AV ČR, v.v.i.
Žitná 25, 115 67 Praha 1
tel.: 222 090 711

podpis a razítko statutárního zástupce příjemce dotace

VÝKAZ
o použití p id lených finan ních prost edk v roce 2009
na øešení projektu
programu 'Informaèní spoleènost'

Identifika ní kód projektu: 1ET200190513

P íjemce: Knihovna AV ÈR, v. v. i.

1. Investi ní náklady (IN)

specifikace položky	výkaz 2009 v tis. K
celkem investi ní náklady	0

2. Mzdové náklady (MZD)

a) mzdy a platy pro pracovníky

jméno a p íjmení pracovníka	výkaz 2009 v tis. K
celkem mzdy a platy	0

b) odmìny

jméno a p íjmení pracovníka	výkaz 2009 v tis. K
Martin Duda	25
Roman Chýla	0
Martin Lhoták	30
Zdeník Tichý	0
Technicko-administrativní personál	0
celkem odm ny	55

3. Ostatní osobní náklady (OON)

specifikace položky OON	výkaz 2009 v tis. K
digitalizace, grafické úpravy, vytváøení metadat	91
celkem ostatní osobní náklady	91

4. V cné náklady (VN)

položky v cných náklad	výkaz 2009 v tis. K
Drobný majetek	0
Provoz a údržba	0
Další provozní náklady	0
Služby	150
Zveøejnìní výsledkù	0
Cestovné	40

Doplňkové náklady	250
Povinné zákonné odvody	20
celkem věcné náklady	460

Účelové finanční prostředky na řešení projektu (v tis. Kč)

rok	IN	MZD	OON	VN	celkem
přiděleno 2009	0	55	91	460	606
čerpání 2009	0	55	91	460	606

Celkové uznané náklady na řešení projektu ze všech zdrojů financování

Účelové prostředky od AV ČR (v tis. Kč)					
2005	2006	2007	2008	2009	celkem
519	606	651	651	606	3033
Ostatní veřejné zdroje (v tis. Kč)					
0	0	0	0	0	0
Neveřejné zdroje financování (v tis. Kč)					
0	0	0	0	0	0
Celkové uznané náklady na řešení projektu (v tis. Kč)					
519	606	651	651	606	3033

Povinnou přílohou výkazu je tištěná sestava jednotlivých účetních položek za rok 2009 potvrzená odpovědnou osobou!

Datum: 13.1.2010

Zpracoval: Markus Lhoták

Tel: 221 403 260

E-mail: Lhotak@KNAV.CZ



podpis řešitele projektu



podpis a razítko statutárního zástupce příjemce dotace

KNIHOVNA AV ČR, v.v.í.
Národní 3, 115 22 Praha 1

VÝKAZ
o použití p id lených finan ních prost edk v roce 2009
na øešení projektu
programu 'Informaèní spoleènost'

Identifika ní kód projektu: 1ET200190513

P íjemce: Univerzita Karlova v Praze, Matematicko-fyzikální fakulta

1. Investi ní náklady (IN)

specifikace položky	výkaz 2009 v tis. K
celkem investi ní náklady	0

2. Mzdové náklady (MZD)

a) mzdy a platy pro pracovníky

jméno a p íjmení pracovníka	výkaz 2009 v tis. K
celkem mzdy a platy	0

b) odmìny

jméno a p íjmení pracovníka	výkaz 2009 v tis. K
Oldøich Ulrych	26
Jiøí Veselý	25
Technicko-administrativní personál	0
celkem odm ñny	51

3. Ostatní osobní náklady (OON)

specifikace položky OON	výkaz 2009 v tis. K
tvorba, kontrola a úprava metadat èasopisù a sborníkù	25
celkem ostatní osobní náklady	25

4. V ñné náklady (VN)

položky v ñných náklad	výkaz 2009 v tis. K
Drobný majetek	16
Provoz a údržba	0
Další provozní náklady	0
Služby	0
Zveøejnìní výsledkù	0
Cestovné	1
Doplòkové náklady	17
Povinné zákonné odvody	18

celkem věcné náklady	52
-----------------------------	-----------

Účelové finanční prostředky na řešení projektu (v tis. Kč)

rok	IN	MZD	OON	VN	celkem
přiděleno 2009	0	51	25	52	128
čerpání 2009	0	51	25	52	128

Celkové uznané náklady na řešení projektu ze všech zdrojů financování

Účelové prostředky od AV ČR (v tis. Kč)					
2005	2006	2007	2008	2009	celkem
128	128	128	128	128	640
Ostatní veřejné zdroje (v tis. Kč)					
0	0	0	0	0	0
Neveřejné zdroje financování (v tis. Kč)					
0	0	0	0	0	0
Celkové uznané náklady na řešení projektu (v tis. Kč)					
128	128	128	128	128	640

Povinnou přílohou výkazu je tištěná sestava jednotlivých účetních položek za rok 2009 potvrzená odpovědnou osobou!

Datum: 11.1.2010

Zpracoval: *Ulych*

Tel: 222191 3216

E-mail: ULRYCH@KARLIN.MFF.CUNI.CZ

Ulych
podpis řešitele projektu

[Signature]
podpis a razítko statutárního zástupce příjemce dotace



VÝKAZ
o použití p id lených finan ních prost edk v roce 2009
na øešení projektu
programu 'Informaèní spoleènost'

Identifika ní kód projektu: 1ET200190513

P íjemce: Masarykova univerzita, Fakulta informatiky

1. Investi ní náklady (IN)

specifikace položky	výkaz 2009 v tis. K
celkem investi ní náklady	0

2. Mzdové náklady (MZD)

a) mzdy a platy pro pracovníky

jméno a p íjmení pracovníka	výkaz 2009 v tis. K
Petr Sojka	71
celkem mzdy a platy	71

b) odm òny

jméno a p íjmení pracovníka	výkaz 2009 v tis. K
Petr Sojka	39
Technicko-administrativní personál	0
celkem odm òny	39

3. Ostatní osobní náklady (OON)

specifikace položky OON	výkaz 2009 v tis. K
Bc. Michal Rùžíèka: pøedzpracování dat DML-CZ	15
Mgr. Radim Øehùøek: Mìøení podobností èlánkù DML-CZ	15
Bc. Peter Boèák: Aktualizace aplikace pdfsign	1
Radim Hatlapatka: Aplikace pro rekompresi PDF v JBIG2	2
Miha Filej: Návrh regresních testù metadatového editoru	3
celkem ostatní osobní náklady	36

4. V òcné náklady (VN)

položky v òcných náklad	výkaz 2009 v tis. K
Drobný majetek	0
Provoz a údržba	0
Další provozní náklady	0
Služby	0

Zveřejnění výsledků	5
Cestovné	54
Doplňkové náklady	36
Povinné zákonné odvody	32
celkem věcné náklady	127

Účelové finanční prostředky na řešení projektu (v tis. Kč)

rok	IN	MZD	OON	VN	celkem
přiděleno 2009	0	110	36	127	273
čerpání 2009	0	110	36	127	273

Celkové uznané náklady na řešení projektu ze všech zdrojů financování

Účelové prostředky od AV ČR (v tis. Kč)					
2005	2006	2007	2008	2009	celkem
334	273	273	273	273	1426
Ostatní veřejné zdroje (v tis. Kč)					
0	0	0	0	0	0
Neveřejné zdroje financování (v tis. Kč)					
0	0	0	0	0	0
Celkové uznané náklady na řešení projektu (v tis. Kč)					
334	273	273	273	273	1426

Povinnou přílohou výkazu je tištěná sestava jednotlivých účetních položek za rok 2009 potvrzená odpovědnou osobou!

Datum: 19.1.2010

Zpracoval: PETR SOJKA


Tel: 549496966

E-mail: sojka@fh.muni.cz



podpis řešitele projektu

MASARYKOVA UNIVERZITA
Žerotínovo nám. 9
601 77 BRNO


podpis a razítko statutárního zástupce příjemce dotace

VÝKAZ
o použití p id lených finan ních prost edk v roce 2009
na øešení projektu
programu 'Informaèní spoleènost'

Identifika ní kód projektu: 1ET200190513

P íjemce: Masarykova univerzita, Ústav výpoèetní techniky

1. Investi ní náklady (IN)

specifikace položky	výkaz 2009 v tis. K
celkem investi ní náklady	0

2. Mzdové náklady (MZD)

a) mzdy a platy pro pracovníky

jméno a p íjmení pracovníka	výkaz 2009 v tis. K
Miroslav Bartošek	37
Petr Kováø	36
Vlastimil Krejèíø	22
Martin Šárfy	46
celkem mzdy a platy	141

b) odmìny

jméno a p íjmení pracovníka	výkaz 2009 v tis. K
Miroslav Bartošek	5
Petr Kováø	3
Vlastimil Krejèíø	5
Martin Šárfy	3
Technicko-administrativní personál	0
celkem odm òny	16

3. Ostatní osobní náklady (OON)

specifikace položky OON	výkaz 2009 v tis. K
studenti - poøizování metadat	25
celkem ostatní osobní náklady	25

4. V òcné náklady (VN)

položky v òcných náklad	výkaz 2009 v tis. K
Drobný majetek	45
Provoz a údržba	0

Daší provozní náklady	6
Služby	1
Zveřejnění výsledků	0
Cestovné	1
Doplňkové náklady	43
Povinné zákonné odvody	51
celkem věcné náklady	147

Účelové finanční prostředky na řešení projektu (v tis. Kč)

rok	IN	MZD	OON	VN	celkem
přiděleno 2009	0	157	25	147	329
čerpání 2009	0	157	25	147	329

Celkové uznané náklady na řešení projektu ze všech zdrojů financování

Účelové prostředky od AV ČR (v tis. Kč)					
2005	2006	2007	2008	2009	celkem
329	329	329	329	329	1645
Ostatní veřejné zdroje (v tis. Kč)					
0	0	0	0	0	0
Neveřejné zdroje financování (v tis. Kč)					
0	0	0	0	0	0
Celkové uznané náklady na řešení projektu (v tis. Kč)					
329	329	329	329	329	1645

Povinnou přílohou výkazu je tištěná sestava jednotlivých účetních položek za rok 2009 potvrzená odpovědnou osobou!

Datum: 11.1.2009

Zpracoval: M. BARTOŠEK

Tel: 549 49 3296

E-mail: bartosek@icc.muni.cz


podpis řešitele projektu


podpis a razítko statutárního zástupce příjemce dotace

Celkové uznané náklady na řešení projektu
ze všech zdroj financování za všechny příjemce

Identifikační kód projektu: 1ET200190513

Účelové finanční prostředky na řešení projektu za všechny příjemce (v tis. Kč)

rok	IN	MZD	OON	VN	celkem
přiděleno 2009	0	403	229	885	1517
čerpání 2009	0	403	235	879	1517

Celkové uznané náklady na řešení projektu ze všech zdroj financování za všechny
příjemce

Účelové prostředky od AV R (v tis. Kč)					
2005	2006	2007	2008	2009	celkem
1747	1517	1762	1562	1517	8105
Ostatní veřejné zdroje (v tis. Kč)					
0	0	0	0	0	0
Neveřejné zdroje financování (v tis. Kč)					
0	0	0	0	0	0
Celkové uznané náklady na řešení projektu (v tis. Kč)					
1747	1517	1762	1562	1517	8105

P EHLED UPLATN ÝCH VÝSLEDK ZA CELÝ PROJEKT

Publikace		Publikováno	P ijato k publikaci
J	ělánek v odborném periodiku	5	0
B	odborná kniha	0	0
C	kapitola v odborné knize	1	0
D	ělánek ve sborníku	16	0

Chrán ěné výsledky		Ud ěno	Podáno
P	patent nebo jiný výsledek podle zvláštních právních pøedpisù	0	0

Aplikované výstupy		Uplatn no	
Z	poloprovod, ovìøená technologie, odrùda, plemeno	7	
S	prototyp, uplatnìná metodika, funkèní vzorek, autorizovaný software, výsledek aplikovaného výzkumu promítнутý do právních pøedpisù a norem, užitný vzor, specializovaná mapa s odborným obsahem	0	

Ostatní výsledky		Uplatn no	
V	oponovaná výzkumná zpráva urèená pro státní zprávu	0	
A	audiovizuální tvorba, elektronický dokument	0	
M	uspoøádání konference	1	
W	uspoøádání workshopu	8	
E	uspoøádání výstavy	0	
O	ostatní výsledky	8	

CELKEM		46	
--------	--	----	--