

Jak se dělá digitální matematická knihovna

Miroslav Bartošek*

bartosek@ics.muni.cz

Vlastimil Krejčíř*

krejcir@ics.muni.cz

Abstrakt: Článek přibližuje postupy a řešení projektu České digitální matematické knihovny (DML-CZ), v jehož rámci jsou zkoumány a vyvíjeny technologie pro digitalizaci matematických textů. Ty jsou pak následně aplikovány při vytváření digitální matematické knihovny. V úvodu je uveden rámcový postup při tvorbě DML-CZ a poté jsou podrobněji popsány jeho dvě finální fáze: (a) vytváření digitálních článků a jejich metadat s využitím Metadatového editoru vyvinutého na ÚVT MU, (b) implementace digitální knihovny DML-CZ s využitím volně dostupného repozitářového systému DSpace.

Klíčová slova: Digitalizace, metadata, matematická digitální knihovna, DSpace.

1 Úvod

Česká digitální matematická knihovna DML-CZ (<http://dml.muni.cz/>) je projekt¹ programu Informační společnost AV ČR. Jeho cílem je vytvořit digitální knihovnu obsahující matematickou literaturu mezinárodní úrovně vydanou v České republice a tuto následně začlenit do připravované světové digitální matematické knihovny WDML (World Digital Mathematics Library, <http://www.wdml.org>). Projekt má jak svou výzkumnou náplň (zkoumání a vývoj pokročilých technologií pro digitalizaci matematiky), tak i praktickou část (digitalizace a zpřístupnění jádra české matematické literatury v rozsahu alespoň 200 tisíc stran textu).

K řešení projektu se sdružilo pět partnerů pokrývajících svými znalostmi a zkušenostmi všechny potřebné odbornosti: matematickou část projektu a specifikaci uživatelských požadavků mají na starosti Matematický ústav AV ČR (hlavní řešitel J. Rákosník) a MFF UK (spoluřešitel J. Veselý); digitalizaci materiálů zajišťuje Knihovna AV ČR (spoluřešitel M. Lhoták); technologie pro OCR matematických textů a specializované vyhledávání zkoumá Fakulta informatiky MU (spoluřešitel P. Sojka); realizaci vlastní digitální knihovny řeší Ústav výpočetní techniky MU (spoluřešitel M. Bartošek). Projekt směřuje k vytvoření obecné matematické digitální knihovny, umožňující začlenit a zpřístupnit různé typy dokumentů – časopisy, sborníky, knihy, kvalifikační práce a samostatné články. V současné etapě řešení jsou dokončeny postupy a softwarová podpora pro zpracování časopiseckých dokumentů a je zdigitalizováno zhruba 70 tisíc stran časopisů.

Postup při praktické realizaci DML-CZ zahrnuje následující vzájemně navazující okruhy činností (fáze):

1. *Příprava:* výběr a příprava materiálů k digitalizaci, řešení autorsko-právních aspektů.
2. *Digitalizace:* vlastní skenování, úpravy digitálních obrazů, pořízení základních stránkových a strukturálních metadat.

* Masarykova univerzita, Ústav výpočetní techniky, Botanická 68a, 602 00 Brno

¹ DML-CZ: Česká digitální matematická knihovna, projekt IET200190513 programu Informační společnost AV ČR, řešený s grantovou podporou AV ČR v letech 2005-2009.

3. *OCR*: dvoukrokové rozpoznávání – nejprve běžné rozpoznání textu (FineReader), na ně navazuje specializované rozpoznávání matematiky (InftyReader).
4. *Referenční metadata*: získání a předzpracování základních článkových metadat z matematických referenčních databází Mathematical Reviews a Zentralblatt MATH.
5. *Integrace*: vytvoření strukturovaných digitálních dokumentů, kontrola/vytvoření a doplnění veškerých popisných metadat, vygenerování článků ve formátu PDF.
6. *Digitální knihovna*: import digitálních objektů (dokumentů a jejich metadat) do digitální knihovny a jejich zpřístupnění uživatelům.

Jádrem tohoto příspěvku je popis posledních dvou z výše uvedených fází. Jejich náplní je integrace výstupů z předchozích fází zpracování a vytvoření cílové digitální knihovny. Pro tyto účely vyvíjí ÚVT MU vlastní podpůrné nástroje (Metadatový editor) a adaptuje obecně dostupné technologie (digitální repozitář DSpace).

2 Integrace: od naskenovaných stránek ke článkům

Hlavními výstupy prvních čtyř fází tvorby DML-CZ jsou kostra strukturovaného dokumentu (v případě časopisu jde o strukturu titul-ročník-číslo), sady obrázků jednotlivých stránek zdigitalizovaného dokumentu, OCR texty stránek a množina metadat získaná z referenčních matematických databází. Z těchto podkladů je nyní třeba vytvořit strukturu jednotlivých časopiseckých čísel (seskupit stránky do článků), zkompletovat potřebná strukturální, popisná a administrativní metadata, a vygenerovat články.

Na podporu těchto aktivit vyvinul ÚVT MU speciální webovou aplikaci **Metadatový editor**. Pomocí Metadatového editoru lze provádět následující druhy činností:

- založit nový časopis a nastavit jeho strukturu (počty ročníků, počty čísel v ročnících);
- importovat obrázky naskenovaných stran časopisu a jejich základní metadata vytvořená během fáze digitalizace;
- importovat referenční metadata článků;
- vytvořit strukturu časopiseckého čísla; tj. vytvořit digitální obsah (Table of Contents) čísla a seskupit naskenované obrazy stran do příslušných článků;
- vytvářet/opravovat metadatové záznamy pro
 - časopisecký *titul*,
 - jednotlivé *ročníky* daného časopisu,
 - jednotlivá *čísla* daného ročníku časopisu,
 - jednotlivé *články* v daném časopiseckém čísle;
- vytvářet či opravovat *seznamy referencí* u jednotlivých článků;
- procházet strukturu časopisu (titul-ročník-číslo-článek) a zobrazovat jednotlivé články (jejich metadata a plné texty v podobě PDF);
- vyhledávat články v databázi Metadatového editoru podle různých kritérií, zobrazovat statistiky a přehledy stavu zpracování u ročníků, čísel a článků zvoleného časopisu.

Jak již bylo uvedeno, klíčovými činnostmi integrační fáze při tvorbě DML-CZ je editace struktury časopiseckého čísla, vytvoření článkových metadat a generování samotných článků. Metadatový editor se snaží všechny tyto činnosti v maximální možné míře automatizovat, aby se co nejvíce snížilo množství potřebné ruční práce.

2.1 Editace struktury čísla

Na základě referenčních metadat a autodetekce hranic článků provede metadatový editor nejprve automatické seskupení naskenovaných stran časopiseckého čísla do jednotlivých článků – vytvoří obsah čísla. Referenční metadata však často obsahují chyby v údajích o stránkování a identifikaci ročníků a čísel časopisu; stejně tak autodetekce hranic článků není stoprocentní. Proto je třeba, aby obsluha editoru strukturu čísla vygenerovanou automaticky vizuálně překontrolovala a případné chyby opravila. Metadatový editor k tomu nabízí velmi efektivní nástroj, který zobrazí náhledy všech stránek čísla uspořádaných jako karty na stole do skupin podle předpokládaných článků – viz následující obrázek.

The screenshot shows the 'DML-CZ: Metadata editor (serial)' interface in a Mozilla Firefox browser. The address bar shows the URL: <http://dmlcz.ieia.ics.muni.cz:9999/edit/issue/9/contents>. The page title is 'DML-CZ: Metadata editor (serial)' and the breadcrumb is 'DML-CZ / CZECHOSLOVAK MATHEMATICAL JOURNAL / Volume 04 / Issue 3 /'. Below the breadcrumb, there is a table of articles with columns for article title and page range:

(#1) Über zwei neue ebene Konfigurationen \$(12_4, 16_3)\$ (4-29)	193-218
(#2) The theory of characters of finite commutative semigroups (30-58)	219-247
(#3) System of congruence relations on lattices (59-93)	248-282
(#4) Sur les espaces à connexion affine partiellement projectifs (94-101)	283-(290)
(#5) Characters of commutative semigroups as class functions (102-103)	(291)-(292)

Below the table are buttons: 'Delete Articles', 'Change Ranges', and 'Save Contents'. A red box highlights a preview area with two thumbnails: '(193a) [2]' showing a journal cover and '(193b) [3]' showing a table of contents. Below these are 'edit ocr scan' links and 'Move Pages' and 'Create Article' buttons. Below the preview, the first article is expanded, showing its title '(#1) Über zwei neue ebene Konfigurationen \$(12_4, 16_3)\$' and page range '193-218'. Underneath is an 'Obsah' (Table of Contents) section with a green background, displaying thumbnails for pages 193 through 202, each with an 'edit ocr scan' link.

Jednotlivé náhledy stránek lze podle potřeby snadno zvětšovat, přesouvat, spojovat do nových skupin či naopak rozpojovat – a takto měnit obsahy článků, vytvářet nové články či naopak chybně identifikované články rušit. Náhledy zobrazené v červených polích obsahují stránky vyřazené z článkového zpracování (titulní a závěrečné strany čísla, vakáty, obsahové stránky), náhledy zobrazené v zelených polích reprezentují příslušné články.

2.2 Editace metadat

Dalším krokem po zafixování struktury časopiseckého čísla je editace popisných metadat článků. Ve většině případů je metadatový záznam článku již automaticky předvyplněn údaji z referenčních metadat, a stačí ho pouze zkontrolovat či doplnit. V případech, kdy referenční metadata článku nejsou k dispozici, je nutné vytvořit kompletní metadatový záznam v editačním formuláři:

The screenshot displays the 'DML-CZ: Metadata editor (serials)' interface. On the left, there is a metadata form with various input fields and dropdown menus. On the right, a preview of the article page is shown, featuring the journal title 'ЧЕХОСЛОВАЦКИЙ МАТЕМАТИЧЕСКИЙ ЖУРНАЛ' and the article title 'О НЕКОТОРЫХ УСЛОВИЯХ ЖЕСТКОСТИ ПОВЕРХНОСТЕЙ ПОЛОЖИТЕЛЬНОЙ КРИВИЗНЫ' by I. N. VEKVA.

Editační formulář se skládá ze dvou částí: v levé části okna je formulář pro vyplňování dat, v pravé části je zobrazen naskenovaný text první stránky článku. Mezi oběma částmi jsou odkazy na všechny stránky článku – kliknutím na zvolené číslo strany se v pravém okně zobrazí její naskenovaný text (uživatel tak může procházet jednotlivé stránky článku a ověřovat si potřebné údaje, aniž by musel opustit editační formulář). Namísto obrázku stránky si operátor může zobrazit její text vytvořený programem OCR, a použít jej například pro vyplňování metadatového záznamu metodou cut-and-paste.

Kromě popisných metadat článků umožňuje Metadatový editor vytvářet metadata i pro časopisecký titul, ročník a číslo.

Vedle editace popisných metadat umožňuje Metadatový editor opravovat i seznam referencí (seznam literatury) u jednotlivých článků. Editační formulář je rozdělen opět na dvě části, jak ukazuje další obrázek:

The screenshot shows the 'DML-CZ: Metadata editor (serials)' interface. The top navigation bar includes 'Users | Roles | Register' and 'mirek (logout)'. The main header displays 'DML-CZ / CZECHOSLOVAK MATHEMATICAL JOURNAL / Volume 07 / Issue 3 / Mathematische Theorie der Torsions- und Biegeschwingungen anisotroper Stäbe / PDF | OCR [3 / 9]'. The interface is split into two main sections:

- Left Section (Editor):** Contains a list of references with line numbers 374-412. The text is partially visible, showing a list of works by Collatz, Courant-Hilbert, Kanke, and others. There are 'Save' and 'Save and Next' buttons at the top and bottom of this section.
- Right Section (Preview):** Shows a preview of the article text in German, including the title 'Das am Anfang dieser Arbeit aufgestellte Problem ist damit bei weitem noch nicht vollständig gelöst...' and a list of references under the heading 'LITERATUR'. Below the references is a 'Резюме' (Summary) section in Russian, starting with 'МАТЕМАТИЧЕСКАЯ ТЕОРИЯ КОЛЕБАНИЙ ПРИ КРУЧЕНИИ И ИЗГИБЕ АНИЗОТРОПНЫХ СТЕРЖНЕЙ'.

V pravé části obrazovky je zobrazen obrázek stránky se seznamem referencí, v levé části obrazovky se pak nachází editační textový blok. V něm je zobrazen OCR-text seznamu referencí do něhož operátor vkládá dohodnuté značky a provádí příslušné opravy. Opravený seznam referencí se automaticky stává součástí metadat příslušného článku.

2.3 Vygenerování článků

Podle údajů ze strukturálních metadat spojí Metadatový editor pdf-soubory příslušných stránek (vytvořených ve fázi OCR) do jediného pdf-souboru článku. Tento soubor obsahuje dvě vrstvy: první vrstvou jsou naskenované obrázky stránek tvořících článek (tato vrstva je zobrazována koncovému uživateli), druhou vrstvou je text článku vytvořený pomocí OCR (nezobrazuje se, slouží pro vyhledávání v textu článku).

Vygenerováním článků je dokončena integrační fáze a veškerá data a metadata daného dokumentu (časopisu) jsou připravena pro import do digitální knihovny.

3 Digitální knihovna: DML-CZ v systému DSpace

V rámci projektu DML-CZ je paralelně zkoumáno a prakticky ověřováno několik přístupů pro implementaci digitální knihovny, jejímž prostřednictvím budou matematické dokumenty zpřístupněny uživatelům. Jedním z nich je využití systému DSpace.

3.1 Systém DSpace

Systém DSpace [1] je volně dostupný repozitářový systém. Hlavní těžiště vývoje nese v současnosti MIT (Massachusetts Institute of Technology) ve Spojených státech amerických – významnou měrou se však na tvorbě systému DSpace podílí široká komunita vývojářů z celého světa (přidáváním vlastních kusů kódu, které rozšiřují možnosti systému DSpace, lokalizacemi pro národní prostředí, opravami chyb, diskusí o budoucím směřování apod.).

Systém DSpace slouží k ukládání informací různého typu – textů, videa, zvukových nahrávek, obrázků aj. Je elegantní především pro archivaci, správu a prezentování samotných souborů s již zpracovanou informací (např. článků ve formátu PDF). Kromě samotného úložiště poskytuje i *kvalitní uživatelské rozhraní*, které přináší možnost *vyhledávání* v uložených objektech (i fulltextově včetně prohledávání souborů typu MS Doc a Adobe PDF), *procházení* repozitářem, možnost sledovat změny zasíláním e-mailů a další. Systém DSpace má také podporu pro *autentizaci a autorizaci uživatelů* – autentizovaným uživatelům na základě jejich práv dovoluje vkládat nové digitální objekty a upravovat jejich metadata. Pro vybrané uživatele (či skupiny uživatelů) je také možné definovat workflow proces schvalování vkládaných dokumentů (nový dokument vložený uživatelem se nejdříve předá vybrané skupině uživatelů, kteří si jej mohou prohlédnout, zkontrolovat/opravit vložená metadata a následně celý objekt zamítnout nebo přijmout do repozitáře).

Ze standardů a protokolů, které DSpace podporuje zmiňme ty nejvýznamnější. Popisná metadata jsou ukládána ve formátu Dublin Core (je však možné si nadefinovat a používat jiný standard). Perzistence identifikátorů je řešena pomocí systému CNRI Handles a interoperabilitu zajišťuje server OAI přes protokol OAI-PMH.

Digitální objekty v systému DSpace jsou uloženy ve stromové struktuře. Jejím nejvyšším prvkem je *komunita*. Každá komunita může obsahovat libovolný počet dalších komunit. Zároveň také může obsahovat i *kolekce*, ve kterých jsou uloženy samotné digitální objekty, kterým se říká *položky*. Každý digitální objekt se sestává z metadat a příložených souborů.

Systém DSpace postaven na technologii Java a je naprogramován v třívrstvé architektuře. Na každé vrstvě je pak definována množina metod, které může programátor při rozšiřování a vylepšování systému používat. Všechny tyto metody jsou zároveň velmi dobře popsány v dokumentaci. Uživatelské rozhraní je generováno v HTML + CSS – lze tak poměrně snadno měnit jeho vzhled. Rozsah změn je omezen možnostmi technologie CSS – můžeme tak měnit například barvy, fonty a celkový vzhled rozhraní, nikoli však strukturu informací na stránce. Strukturu informací uživatelského rozhraní lze měnit úpravou kódu stránek JSP (ze kterých jsou výsledné HTML stránky generovány) a samotného kódu DSpace. Jinou alternativou pro změnu a přepracování uživatelského rozhraní je nový šablonovací systém **Manakin** [2]. Ten využívá technologie XSLT ke generování výsledných stránek uživatelského rozhraní a lze pomocí něj elegantně dosáhnout totéž, co přímou úpravou stránek JSP.

Všechny výše uvedené vlastnosti dávají dobrý základ a prostředky k vybudování digitální matematické knihovny nad systémem DSpace.

Zájemce o podrobnější informace můžeme odkázat na [www stránky Czech DSpace](http://www.czechdspace.org) [3] v českém jazyce, které mohou posloužit jako výchozí bod pro seznámení se systémem DSpace.

3.2 Úprava systému DSpace pro DML-CZ

Jednotlivé články v projektu DML-CZ jsou řazeny do struktury časopis-ročník-číslo-článek. Pro reprezentaci této struktury v systému DSpace jsme využili toho, že komunita nemusí být jen „komunitou“ v obvyklém významu tohoto slova, ale může posloužit i jako prvek s odlišným smyslem. V systému DSpace, který je speciálně upraven pro DML-CZ, slouží kořenové komunity pro reprezentaci samotných časopiseckých titulů (například Czechoslovak Mathematical Journal). Každý časopisecký titul obsahuje komunity, které reprezentují jeho jednotlivé ročníky. Každý ročník obsahuje kolekce odpovídající číslům v daném ročníku. Konečně každá kolekce má v sobě digitální objekty, což jsou články, které v daném čísle časopisu vyšly. Tímto způsobem lze poměrně elegantně mapovat strukturu uložení článků v systému DSpace.

Základním cílem, kterého chceme dosáhnout, je vložení článků, které byly připraveny pomocí Metadatového editoru, do systému DSpace. Metadatový editor ukládá vytvořené články v PDF i příslušné soubory s metadaty a referencemi do vlastní adresářové struktury.

Systém DSpace má podporu hromadného importování pouze jednotlivých objektů do vybrané kolekce a nedokáže importovat objekty spolu s vytvářením nových komunit a kolekcí.

The screenshot shows the DML-CZ interface within DSpace. At the top left is the DML-CZ logo. The main header reads "Czech Digital Mathematics Library". Below this, the page is divided into several sections:

- Search DML-CZ:** Includes a search input field with a "Go" button, and links for "Advanced Search" and "Subject Search".
- Browse:** A list of navigation links: Home, Resources, Titles, Authors, Subjects, and By Date.
- Sign on to:** Links for "Receive email updates", "My DSpace authorized users", "Edit Profile", "Administrator", "Help", and "About DSpace".
- DML-CZ in DSpace >** A section with a blue background containing:
 - DML-CZ is live!** A message stating: "Collection of digitized mathematical journals, proceedings, books and other literature published by Czech and Slovak institutions."
 - Search:** A section with the text "Enter some text in the box below to search DML-CZ." and a search input field with a "Go" button.
 - Resources in DML-CZ:** A section with the text "Choose a journal/book/proceedings to browse." and a link for "Czechoslovak Mathematical Journal [2297]".
- News:** A section titled "DML-CZ Czech Digital Mathematics Library" with a sub-section "News" containing the text "Jan 2007 - CMJ imported".

At the bottom of the page, there is a footer: "DSpace Software Copyright © 2002-2006 MIT and Hewlett-Packard - Feedback".

Přirozeně má i vlastní formát uložení objektů pro import, který je odlišný od struktury, se kterou pracuje a kterou vytváří Metadatový editor. Protože je nativní importovací nástroj systému DSpace nevyhovující, vytvořili jsme nový importovací nástroj, který dokáže importovat časopisy přímo ze struktur vytvořených Metadatovým editorem. Velkou výhodou tohoto přístupu je možnost importovat data přesně dle našich požadavků – včetně výchozího

nastavení přístupových práv apod. Při tvorbě (jakéhokoli) skriptu lze samozřejmě s výhodou využít metod rozhraní, které systém DSpace poskytuje (např. metody pro vytvoření komunity, nastavení přístupových práv objektu apod.).

Bylo také nutné rozšířit základní množinu prvků Dublin Core o několik nových kvalifikátorů elementů, protože výchozí základní množina nedostačovala potřebám DML-CZ. Rozšíření se týká například identifikátorů článků v matematických referenčních bázích Zentralblattmath a MathSciNet, metadat pro popis rozsahu stran článku v daném čísle, pole pro kategorizaci dle MSC (Mathematics Subject Classification) apod.

Pro zajímavost lze uvést, že import časopisu Czech Mathematical Journal, tedy asi 2 290 článků ve 168 číslech rozložených do 42 ročníků, trvá přibližně 3 hodiny (na dvoujádrovém procesoru Intel Pentium D 3.4 GHz). Zde si vybírá svoji daň použití jazyka Java a zpracovávání velkého množství souborů XML (soubory s metadaty a referencemi). Zaindexování pro fulltextové vyhledávání po importu pak trvá přibližně další 3 hodiny (systém DSpace používá volně dostupný indexovací nástroj Apache Lucene).

Search DML-CZ

[Advanced Search](#)
[Subject Search](#)

⇒ [Home](#)

Browse

⇒ [Resources](#)
⇒ [Titles](#)
⇒ [Authors](#)
⇒ [Subjects](#)
⇒ [By Date](#)

Sign on to:

⇒ [Receive email updates](#)
⇒ [My DSpace](#)
authorized users
⇒ [Edit Profile](#)

⇒ [Administrator](#)

⇒ [Help](#)
⇒ [About DSpace](#)

[DML-CZ in DSpace >](#)
[Czechoslovak Mathematical Journal >](#)
[1964 \(Volume 14\) >](#)
[3 >](#)

Please use this identifier to cite or link to this item: http://dspace.muni.cz/handle/dml_cz/1116

Title: Über Verbandsgruppen mit zwei Erzeugenden

Authors: Jakubík, Ján

Language: ger
rus

Pages range: 444-454

MSC: [06.75](#)
[20.00](#)

MathSciNet id: [MR0170962](#)

Zentralblattmath id: [0135.06202](#)

Journal: Czechoslovak Mathematical Journal

Volume-Year-Issue: 14-1964-3

References: [1] **G. Birkhoff:** *Lattice Theory*. Rev. edition. New York 1948. [ZB: 0033.10103](#), [MR: 0029876](#)
[2] **P. Conrad:** *Some structure theorems for lattice ordered groups*. Trans. Amer. Math. Soc. 99, 1961, 212-240. [ZB: 0099.25401](#), [MR: 0121405](#)
[3] **Ян Якубик:** *О главных идеалах в структурно упорядоченных группах*. Чех. мат. ж. 9 (84), 1959, 528-543.

Files in This Item:

File	Description	Size	Format	
CMJ_14-1964-3_8.html	fulltext	31 kB	HTML	View/Open
CMJ_14-1964-3_8.pdf	fulltext and images	1383 kB	Adobe PDF	View/Open

[Recommend this item](#)
[Print version](#)

Po provedení importu je možné zpřístupnit systém DSpace uživatelům, kteří mohou s digitální knihovnou matematických časopisů okamžitě začít pracovat. Uživatelské rozhraní

však na specializovaná data není uzpůsobeno a např. zobrazení metadat u článku není zcela vyhovující. Proto bylo nutné upravit i uživatelské rozhraní.

Snadnější část se týkala výchozího vzhledu, který byl více přizpůsoben DML-CZ. Úvodní obrazovka tak již tolik nepřipomíná systém DSpace. Dále jsme změnili některé položky v bočním menu, popisky u komunit a kolekcí (tak aby odpovídaly tomu, co komunita/kolekce reprezentuje). Těchto jednoduchých změn lze dosáhnout pomocí úpravou stylů CSS, stránek JSP a editace souboru s texty, které se na generovaných www stránkách vyskytují.

Pořadí zobrazení metadatových záznamů u článků umožňuje systém DSpace měnit pomocí konfiguračního souboru. Neumožňuje však nastavit pro jednotlivé metadatové položky způsob jejich zobrazení. Například je žádoucí, aby identifikátory do matematických referenčních bází byly zároveň hypertextovými odkazy, které povedou na záznam o článku v dané bázi. Kód MSC pak může odkazovat na seznam článků, které jsou uloženy v repozitáři a mají daný MSC. Dosáhnout těchto vlastností lze pouze naprogramováním vlastního kódu, který zpracuje metadata a u příslušných polí a provede žádanou úpravu (dodá odkazy apod.). Totéž platí pro další provedené úpravy – vypsání seznamu názvů článků (které odkazují daný článek v repozitáři) na úrovni čísla časopisu, úpravy předmětových řazení aj.

Po provedení výše uvedených úprav dostáváme funkční verzi digitální knihovny DML-CZ. Lze v ní fulltextově vyhledávat, procházet ji po ročnících a číslech, prohlížet si články. Úsilí, které by bylo nutno vynaložit na vybudování podobné knihovny „na zelené louce“ by pravděpodobně několikanásobně překročilo množství práce potřebné pro úpravu systému DSpace. Navíc u systému DSpace dostáváme v podstatě téměř zadarmo funkční OAI-PMH server a v neposlední řadě i podporu rozsáhlé komunity uživatelů a vývojářů, se kterými lze problémy a zkušenosti s provozem i vývojem systému DSpace diskutovat.

4 Závěr

Dosavadní řešení projektu DML-CZ prokázalo, že vhodnou kombinací volně dostupných systémů a speciálně vyvinutých nástrojů lze výrazně zefektivnit proces zpracování digitalizovaných dokumentů a vytvoření digitální knihovny. Projekt bude pokračovat pracemi na zařazení zbývajících typů dokumentů, rozsáhlejší digitalizací vybraných matematických dokumentů v ČR a perspektivně i na Slovensku, a začleněním vznikající DML-CZ do světové digitální matematické knihovny.

WWW odkazy

1. DSpace, <http://www.dspace.org/>
2. DSpace Manakin, <http://di.tamu.edu/projects/xmlui/manakin>
3. Czech DSpace, <http://www.ics.muni.cz/dspacecz/>