

**Univerzita Karlova v Praze**  
**Filozofická fakulta**  
**Ústav informačních studií a knihovnictví**

Studijní program: informační studia a knihovnictví  
Studijní obor: informační studia a knihovnictví

**Bc. Lucie Krouzová**

**Vývoj a současný stav digitalizace knihovních fondů v České republice a její klíčové otázky**

Diplomová práce

Praha 2008

Vedoucí diplomové práce: PhDr. Eva Bratková

Oponent diplomové práce:

Datum obhajoby:

Hodnocení:

**Hodnocení:**

**Prohlášení:**

Prohlašuji, že jsem diplomovou práci zpracovala samostatně a že jsem uvedla všechny použité informační zdroje.

V Praze, 12. prosince 2008

.....

podpis diplomanta

**Identifikační záznam:**

KROUZOVÁ, Lucie. *Vývoj a současný stav digitalizace knihovních fondů v České republice a její klíčové otázky [The Development and Current State of the Library Stock Digitalization in the Czech Republic and Its Key Issues]*. Praha, 2008. 101 s. Diplomová práce. Univerzita Karlova v Praze, Filozofická fakulta, Ústav informačních studií a knihovnictví 2008. Vedoucí diplomové práce PhDr. Eva Bratková.

**Abstrakt:**

Tématem této diplomové práce je problematika digitalizace knihovních fondů v České Republice a související klíčové otázky. V první části se práce zabývá procesem a nástroji digitalizace. Seznamuje s obrazem a jeho zpracováním, obrazovými formáty, metadaty, popisuje proces mikrofilmování dokumentů a zabývá se způsoby ochrany digitálních objektů. Další část práce stručně popisuje možnosti dotací pro projekty digitalizace v ČR a seznamuje se systémy Manuscriptorium a Kramerius, nejvýznamnějšími českými programy digitalizace. Následuje zásadní část práce, ve které jsou podrobně rozebrány vybrané projekty digitalizace. Další kapitola je věnována problematice duplicit digitalizovaných dokumentů a především nástrojům registrace digitalizovaných dokumentů, které umožňují předcházení vzniku těchto duplicit. Následuje popis projektu Evropské digitální knihovny a kroků, které jejímu vzniku předcházely. Závěr práce tvoří analýza a celkové zhodnocení popisovaných projektů digitalizace knihovních fondů.

**Klíčová slova:**

digitalizace, obrazové informace, ochranné reformátování, knihovní fondy, kulturní dědictví, rukopisy, staré tisky, novodobé tisky

<b>Obsah</b>	
<b>Předmluva.....</b>	<b>10</b>
<b>1.Úvod.....</b>	<b>12</b>
<b>2.Metodika sběru a zpracovávání informací pro diplomovou práci.....</b>	<b>14</b>
<b>3.Digitalizační proces.....</b>	<b>19</b>
<b>3.1.Obraz a jeho zpracování.....</b>	<b>19</b>
3.1.1.Barva a barevná hloubka.....	19
3.1.2.Druhy obrazu.....	19
3.1.3.Komprese obrazu.....	20
<b>3.2.Obrazové formáty.....</b>	<b>20</b>
3.2.1.JPEG.....	21
3.2.1.GIF.....	21
3.2.2.PNG.....	21
3.2.3.TIFF.....	22
3.2.4.Formáty typu wavelet.....	22
3.2.5.Formát typu smíšeného obrazu.....	23
<b>3.3.Metadata.....</b>	<b>23</b>
3.3.1.SGML.....	24
3.3.1.XML.....	25
3.3.2.Formát MASTER pro popis rukopisů.....	27
3.3.3.Formát METS.....	27
3.3.4.MODS Schema.....	30
<b>3.4.Ochranné mikrofilmování dokumentů.....</b>	<b>30</b>
3.4.1.Proces mikrofilmování.....	32
<b>3.5.Ochrana vytvořených dat .....</b>	<b>33</b>
<b>4.Programy dotací pro projekty digitalizace v České republice.....</b>	<b>36</b>
<b>4.1.VISK 6 Národní program digitálního zpřístupnění vzácných dokumentů Memoriae Mundi Series Bohemica.....</b>	<b>36</b>
<b>4.2.VISK 7 Národní program mikrofilmování a digitálního zpřístupňování dokumentů ohrožených degradací kyselého papíru – Kramerius.....</b>	<b>37</b>
<b>4.3.Finanční mechanismy Evropského hospodářského prostoru a Norský finanční mechanismus .....</b>	<b>38</b>
<b>5.Systémy Manuscriptorium a Kramerius.....</b>	<b>41</b>
<b>5.1.Manuscriptorium.....</b>	<b>41</b>
5.1.1. Standardy Manuscriptoria.....	42
<b>5.2.Kramerius.....</b>	<b>44</b>
5.2.1.Systém Kramerius.....	45
5.2.1.Standardy Krameria.....	46
<b>6. Programy a projekty digitalizace v ČR.....</b>	<b>48</b>

<b>6.1.Reformátování bohemikálních časopisů z historického fondu STK.....</b>	<b>48</b>
6.1.1.Průběh projektu Reformátování bohemikálních časopisů z historického fondu STK v jednotlivých letech.....	49
<b>6.2.České učební texty z fondu STK.....</b>	<b>50</b>
<b>6.3.Česká digitální matematická knihovna (DML-CZ).....</b>	<b>51</b>
6.3.1.Systém digitální knihovny DML-CZ.....	54
<b>6.4.Digitální knihovna Akademie věd ČR.....</b>	<b>56</b>
<b>6.5.Digitalizovaný archiv časopisů.....</b>	<b>57</b>
<b>6.6.Digitalizace historických dokumentů knihovny Vojenského historického ústavu.....</b>	<b>59</b>
<b>6.7.Digitalizace a zpřístupnění církevních listin v rámci evropského projektu Monasterium.....</b>	<b>62</b>
6.6.1.Popis webové stránky MOM Virtuální archiv listin střední Evropy.....	64
<b>6.8.HISPRA - Záchrana historických pragensijních a dalších vzácných dokumentů z fondu Městské knihovny .....</b>	<b>66</b>
<b>6.9. Společná česko-slovenská digitální parlamentní knihovna .....</b>	<b>68</b>
<b>6.10.České sněmy.....</b>	<b>71</b>
<b>6.11.Digitální knihovna historických fondů Vědecké knihovny v Olomouci. .</b>	<b>73</b>
<b>6.12.Digitální knihovna map Vědecké knihovny v Olomouci.....</b>	<b>75</b>
<b>6.13.Projekty digitalizace Moravské zemské knihovny v Brně.....</b>	<b>76</b>
6.13.1.Historické fondy Moravské zemské knihovny v Brně.....	76
<b>7.Problematika registrace digitalizovaných jednotek.....</b>	<b>79</b>
7.1.1.CZROMM Český registr reformátovaných dokumentů.....	79
7.1.2.Databázová aplikace RD.CZ.....	81
<b>8.Možnosti dalšího rozvoje a spolupráce.....</b>	<b>89</b>
<b>8.1.ENRICH Evropská digitální knihovna rukopisů.....</b>	<b>89</b>
<b>8.2.Evropská digitální knihovna a projekt TEL-ME-MOR.....</b>	<b>90</b>
8.2.1.Evropská knihovna.....	90
8.2.2.Projekt TEL-ME-MOR.....	91
8.2.3.Evropská digitální knihovna.....	92
<b>9.Závěr.....</b>	<b>93</b>
<b>9.1.Zhodnocení analyzovaných programů a projektů digitalizace.....</b>	<b>93</b>
.....	97
<b>Seznam použitých informačních zdrojů.....</b>	<b>98</b>

## Seznam obrázků

Obr. 1. Digitální knihovna DML-CZ se zobrazeným záznamem článku..	56
Obr. 2. Digitalizovaný archiv časopisů se zobrazeným seznamem vyhledaných článků obsahujících klíčové slovo „ kniha“ v názvu.....	59
Obr. 3. Úvodní stránka archivu MOM.....	64
Obr. 4. Zobrazení dokumentu v archivu MOM.....	65
Obr. 5. Titulní obrazovka Společné česko-slovenské digitální parlamentní knihovny.....	70
Obr. 6. Úvodní obrazovka digitální knihovny České sněmy.....	73
Obr. 7. Digitální knihovna historických fondů Vědecké knihovny v Olomouci se zobrazeným rukopisem „O svatém Jeronýmovi knihy troje“.....	74
Obr. 8. Abecední seznam map řazený dle formátu v Digitální knihovně map Vědecké knihovny v Olomouci.....	75
Obr. 9. Český registr reformátovaných dokumentů CZROMM, ukázka seznamu záznamů názvů periodik, začínajících písmenem L.....	81
Obr. 10.Obrazovka rozhraní pro jednoduché vyhledávání v databázi Registru Digitalizace RD.CZ s částí obrazovky helpu rešeršního systému Convera RetrievalWare, ze dne 4.12.2008.....	84
Obr. 11.Obrazovka rozhraní pro pokročilé vyhledávání v databázi RD.CZ s ukázkou možností hledání, ze dne 4. 12. 2008.....	85
Obr. 12.Obrazovka rozhraní v databázi RD.CZ, ukázka seznamu dokumentů obsahujících termín „Jungmann“, ze dne 4. 12. 2008.....	86
Obr. 13.Obrazovka rozhraní v databázi RD.CZ, ukázka náhledu vybraného dokumentu č. 1, ze dne 4. 12. 2008.....	87
Obr. 14.Obrazovka rozhraní v databázi RD.CZ, ukázka náhledu vybraného dokumentu č. 2, ze dne 4. 12. 2008.....	88



## **Předmluva**

Diplomová práce nese název *Vývoj a současný stav digitalizace knihovních fondů v České republice a její klíčové otázky* a zabývá se problematikou digitalizace knihovních fondů v knihovnách, archivech, muzeích a dalších paměťových institucích České Republiky.

Téma práce bylo zvoleno s ohledem na aktuálnost problematiky digitalizace vzácných či ohrožených dokumentů a na poměrně značné množství institucí v rámci České Republiky, které se na digitalizaci podílejí. K volbě tématu mne vedl rovněž můj osobní zájem o historii knihy a knižní kulturu, která byla také náplní mého bakalářského studijního programu *Knižní kultura na Ústavu umění a designu Západočeské Univerzity v Plzni*.

Technologie digitalizace je účinným nástrojem, jak dochované knižní poklady uchovat pro budoucí generace a zároveň umožnit jejich využití široké veřejnosti. Abych se blíže seznámila s problematikou digitalizace v knihovnách, absolvovala jsem v rámci příprav k diplomové práci kurz *Digitalizace v knihovnách aneb Co musí vědět knihovník o digitalizaci*. Tento kurz pořádala Národní knihovna ČR v říjnu až listopadu roku 2007 a lektorem byl Adolf Knoll.

Cílem práce je zmapovat, popsat a vyhodnotit projekty a programy digitalizace rukopisů a starých a novodobých tisků v paměťových institucích. Zároveň se práce zaměřuje na klíčové otázky spolupráce mezi různými subjekty, možnosti předcházení duplicit digitalizovaných dokumentů a registrace těchto dokumentů.

V rámci přípravy byla nejprve provedena rozsáhlá rešerše zaměřená na vyhledání, vytěžení a vyhodnocení odborné literatury k tématu. Dalším krokem byla rešerše zacílená na zjišťování konkrétních programů a projektů digitalizace knihovních fondů v rámci České Republiky. Jako hlavní zdroj pro získání přehledu projektů digitalizace v ČR posloužily seznamy výsledků dotačních řízení národních programů VISK 6 a VISK 7 za roky 2007 a 2008, které jsou volně dostupné na webu (<http://visk.nkp.cz/>) a databázová aplikace Informační systém výzkumu a vývoje IS VaV, která je volně online dostupná (<http://aplikace.isvav.cvut.cz/>). Z takto vzniklého přehledu byly vybírány projekty pro popis v této diplomové práci, a to s ohledem na typ digitalizovaných dokumentů, velikost či významnost projektu a v neposlední řadě na míru a dostupnost zveřejňovaných informací o projektu.

Poslední jmenované kritérium se ukázalo být značným problémem. Ačkoli jsou projekty digitalizace hrazeny v nemalé míře z veřejných finančních prostředků, povinnost informovat veřejnost o průběhu projektu mnozí příliš, či dokonce vůbec nesplňují. Mnohé instituce na svých webových sídlech nemají o probíhajících projektech žádnou zmínku, jiné podávají pouze velmi omezené informace. Výjimkou jsou instituce, které zveřejňují o projektech výroční zprávy, nebo informace o nich zahrnují do svých ročních zpráv. Z tohoto důvodu bylo sbírání zásadních informací o popisovaných projektech značně obtížné. Pokud to bylo možné, byl nedostatek zveřejněných informací kompenzován osobní konzultací přímo s řešiteli projektu. V případech, kdy nebylo možné uskutečnit konzultaci osobní, kontaktovala jsem řešitele projektu prostřednictvím elektronické pošty.

Práce je tematicky členěna následujícím způsobem: V první části se práce zabývá procesem digitalizace a vším co tento proces obnáší. Seznamuje s obrazem a jeho zpracováním, obrazovými formáty, metadaty, popisuje proces mikrofilmování dokumentů a zabývá se způsoby ochrany digitálních objektů. Další část práce stručně popisuje možnosti dotací pro projekty digitalizace v ČR a seznamuje se systémy Manuscriptorium a Kramerius, nejvýznamnějšími českými programy digitalizace. Následuje zásadní část práce, ve které jsou podrobně rozebrány vybrané projekty digitalizace. Další kapitola je věnována problematice duplicit digitalizovaných dokumentů a především nástrojům registrace digitalizovaných dokumentů, které umožňují předcházení vzniku těchto duplicit. Následuje popis ambiciózního projektu Evropské digitální knihovny a kroků, které jejímu vzniku předcházely. V závěru práce je provedena analýza a celkové zhodnocení popisovaných projektů digitalizace knihovních fondů a je nastíněna perspektiva dalšího rozvoje na poli digitalizace.

Na závěr jsou uvedeny použité informační zdroje, které byly citovány podle norem ISO 690 a ISO 690-2. Pro citace v textu práce byl zvolen Harvardský systém citování.

#### Poděkování

Na tomto místě bych chtěla poděkovat všem konzultantům mé diplomové práce, PhDr. Jiřímu Polišenskému, a zvláště PhDr. Evě Bratkové za vedení diplomové práce a poskytnutí mnoha cenných rad a připomínek.

## 1. Úvod

Historické dokumenty, které jsou kulturním dědictvím naší společnosti, jsou v knihovnách uloženy z důvodu jejich ochrany ve speciálních depozitářích s velmi omezenou, nebo vůbec žádnou možností přístupu uživatelů. Ukazuje se, že jedinou možností, jak tyto dokumenty zachovat budoucím generacím a zároveň umožnit uživatelům kvalitní přístup k nim, je vytvořit jejich digitální kopie a ty zpřístupnit široké vrstvě uživatelů v síti Internet.

Tato možnost se v důsledku vývoje informačních a komunikačních technologií objevila v 90. letech minulého století. V České Republice se jako první instituce začala digitalizací zabývat Národní knihovna ČR ve spolupráci s firmou AiP Beroun. Tyto aktivity začaly v rámci iniciativy UNESCO *Paměť světa (Memory of the World)* v roce 1992. V roce 1993 byl Národní knihovnou ČR vydán první CD-ROM tohoto programu s digitalizovanými vzácnými dokumenty. Digitalizační aktivity pokračovaly v roce 2000 vytyčením úkolů navazujícího českého národního programu *Memoriae Mundi Series Bohemica* řešením souvisejících aktuálních požadavků a praktickým provozem. Během let 2003-2004 byla zprovozněna digitální knihovna Manuscriptorium. V roce 2005 získala Národní knihovna ČR cenu [UNESCO Jikji](#), zejména pro aktivitu v rámci digitálního zpřístupnění vzácných sbírek [Manuscriptorium, 2006].

Příkladem Národní knihovny ČR následovala Vědecká knihovna v Olomouci, která začala s digitalizací historických dokumentů v roce 1996 převodem vzácného *Evangeliaře zábrdovického* do digitální podoby pod záštitou Ústavu pro klasická studia AV ČR v rámci projektu *Clavis monumentorum litterarum*. Digitalizační práce pak pokračovaly až od roku 1999 v rámci grantových programů Ministerstva kultury ČR *RISK* a *VISK*. Digitalizace probíhá i v současnosti a provádí ji firma AiP Beroun (<http://www.aipberoun.cz/>) v rámci programu *VISK 6 Memoriae Mundi Series Bohemica* [VOJNAR, 2002].

V roce 1997 se zrodil při Národní knihovně ČR program Kramerius v rámci pilotního projektu výzkumu a vývoje, kdy byla převedena na mikrofilm část závažně poškozených, významných bohemikálních periodik. Kramerius je označení pro Národní program ochranného mikrofilmování a digitálního zpřístupňování dokumentů ohrožených degradací kyselého papíru a v roce 2000 byl Kramerius

zařazen jako podprogram VISK 7 do programu Veřejné informační služby knihoven (<http://visk.nkp.cz>).

Po roce 2000 pak začínalo s digitalizací svých knihovních fondů stále více paměťových institucí. Smutným podnětem pro digitalizaci ohrožených dokumentů byly také záplavy v roce 2002, které v Praze zatopily depozitáře vzácných dokumentů hned několika knihoven.

V souvislosti se stále mohutnějším nástupem digitalizace dokumentů vývoj v této oblasti neustále pokračuje a přináší další zlepšování technologií digitalizace i zpřístupňování digitálních objektů.

## **2. Metodika sběru a zpracování informací pro diplomovou práci**

V rámci příprav pro vypracování této diplomové práce byla provedena rozsáhlá rešerše, která byla zaměřena jednak na zjišťování konkrétních programů a projektů digitalizace knihovních fondů v rámci České Republiky, jednak na zjišťování literatury pojednávající o různých aspektech tvorby a provozu těchto systémů a jejich projektech z teoretického i praktického hlediska. Podklady, informace a fakta nezbytná pro sepsání této diplomové práce byla sbírána několika různými způsoby. Některé údaje byly získány z publikovaných článků pojednávajících o dané problematice, některé informace byly vytěženy z webových sídel daných institucí, jiné z konzultací s odborníky, ať již formou osobního rozhovoru, nebo korespondencí.

Nejobtížnější částí práce bylo získat přehled projektů a programů digitalizace knihovních fondů, které v současnosti probíhají, jsou plánovány nebo byly v nedávné době dokončeny. Jako prvotní vodítko posloužily seznamy výsledků dotačních řízení národních programů VISK 6 a VISK 7, které jsou volně dostupné na webu (<http://visk.nkp.cz/>), kam lze vstoupit z webových sídel Národní knihovny ČR a Ministerstva kultury ČR. Ze seznamů za roky 2007 a 2008 jsem vybrala projekty, které splňují kritéria zadání této diplomové práce. Pro vyhledání některých projektů, které nejsou financovány z národních programů VISK, posloužila stejným způsobem databázová aplikace Informační systém výzkumu a vývoje IS VaV, která je volně online dostupná (<http://aplikace.isvav.cvut.cz/>). Prvotní část informací k takto vybraným projektům jsem vytěžila z údajů, článků a výročních zpráv, které mají instituce vystavené na svých webových sídlech. To se však nepodařilo zdaleka u všech institucí, neboť jen menšina z nich uvádí na svých webových sídlech alespoň základní informace o právě probíhajících projektech.

Níže je ke každému programu nebo projektu digitalizace, kterými se zabývá tato diplomová práce, uvedena metodika, jakým způsobem byly vytěženy klíčové informace.

### **System Manuscriptorium**

Informace o digitální knihovně Manuscriptorium a jejích cílech byly čerpány z dokumentů zveřejněných na webovém sídle Manuscriptoria (<http://www.manuscriptorium.com/Site/CZE/dokumentace.asp>). Zásadní informace

o standardech programu Manuscriptorium byly vytěženy z dokumentu *Manuscriptorium v. 1.0 : Výběr a popis dokumentů* [UHLÍŘ, 2006]. Dále bylo čerpáno z dalších dokumentů, ve kterých je tento projekt popisován [PSOHLAVEC, 2004, s.56], [KUČERA, 2006, s.4] .

### **System Kramerius**

Národní program Kramerius a jeho požadavky jsou podrobně popsány na webovém sídle programu Veřejné informační služby knihoven - VISK (<http://visk.nkp.cz/VISK7.htm>). Zásadní informace o standardech programu Kramerius a stejnojmenném open source systému byly vytěženy z článku, který byl zveřejněn na webovém sídle Národní knihovny České Republiky ([http://www.nkp.cz/pages/page.php3?page=weba\\_digitalizace.htm](http://www.nkp.cz/pages/page.php3?page=weba_digitalizace.htm)).

### **Společná česko-slovenská digitální parlamentní knihovna**

Informace o této digitální knihovně jsou volně dostupné online na webovém sídle Parlamentní knihovny (<http://www.psp.cz/eknih/info.htm>), odkud byly vytěženy pro potřeby této práce.

### **České sněmy**

Projekt České sněmy je také projektem Parlamentní knihovny, tudíž jsou podrobné informace o projektu rovněž volně dostupné na jejím webovém sídle (<http://www.psp.cz/eknih/snemy/projekt.htm>).

### **Digitální knihovna historických fondů Vědecké knihovny v Olomouci**

Digitální knihovna historických fondů Vědecké knihovny v Olomouci je volně dostupná online na webovém sídle knihovny (<http://dig.vkol.cz/>). Na téže stránce jsou uvedeny základní informace o tomto projektu.

### **Digitální knihovna map Vědecké knihovny v Olomouci**

Digitální knihovna map Vědecké knihovny v Olomouci je volně přístupná online (<http://mapy.vkol.cz/>). V úvodu stránky jsou uvedeny základní informace k tomuto projektu.

### **Historické fondy Moravské zemské knihovny v Brně**

Moravská zemská knihovna v Brně na svém webovém sídle zveřejňuje podrobné zprávy o svých výzkumných záměrech a projektech. Také výzkumný záměr

Historické fondy je popsán v průběžných zprávách za léta 2004 - 2007, které jsou volně online dostupné (<http://www.mzk.cz/projekty/histfondy/>). Na tyto zprávy se lze dostat z hlavní webové strany Moravské zemské knihovny. Obecné informace o projektu byly získány z databázové aplikace Informační systém výzkumu a vývoje IS VaV (<http://aplikace.isvav.cvut.cz/>).

### **Česká matematická digitální knihovna**

Základní informace o tomto projektu a o již plně funkční digitální knihovně byly získány u Masarykovy univerzity (<http://dml.muni.cz/documents.html>). V současné době je však tento link nefunkční. Na tuto webovou adresu bylo možné dostat se proklikem z webové adresy Digitalizačního centra Knihovny Akademie věd ČR (<http://digit.lib.cas.cz/index.php?cat=projekty>). Informace o aktuálním stavu projektu byly vytěženy z odborných příspěvků na konferencích. Obecná fakta byla získána z databázové aplikace Informační systém výzkumu a vývoje IS VaV (<http://aplikace.isvav.cvut.cz/>). Česká matematická digitální knihovna DML-CZ se nachází na nové webové adrese (<http://www.dml.cz>). Tato stránka je k dispozici pouze v anglickém jazyce.

### **Digitální knihovna Akademie věd ČR**

Informace o Digitální knihovně Akademie věd ČR jsou volně přístupné online na webové adrese Digitalizačního centra Knihovny AV ČR (<http://digit.lib.cas.cz/>), která je přístupná z domovské stránky Knihovny Akademie věd ČR. Proklikem na jednotlivé nabídky z menu po levé straně je možné získat přístup k informacím o organizaci a vybavení centra, projektech, zpracovaných titulech, spolupráci aj. Další informace byly získány z databázové aplikace Informační systém výzkumu a vývoje IS VaV (<http://aplikace.isvav.cvut.cz/>).

### **Digitalizovaný archiv časopisů**

Tento archiv je již plně funkční digitální knihovnou, volně dostupnou online (<http://archiv.ucl.cas.cz/>), kam lze vstoupit proklikem z hlavní webové strany zřizovatele archivu, Ústavu pro českou literaturu AV ČR. Odtud byly vytěženy také primární informace o archivu. Další potřebné údaje byly zjištěny elektronickou korespondencí s PhDr. Věrou Buriánkovou.

## **Digitalizace a zpřístupnění církevních listin v rámci evropského projektu Monasterium**

Projekt Monasterium, v České Republice prezentován Národním archivem v Praze, který je do projektu zapojen, prezentuje množství informací a seznam již digitalizovaných církevních dokumentů české provenience online (<http://www.nacr.cz/zpravy/monasterium.aspx>), přístupný z hlavní strany webového sídla Národního archivu ČR.

Dále je možné se proklikem dostat přímo na webové sídlo projektu Monasterium (<http://www.monasterium.net>), kde je archiv volně přístupný a navíc jsou zde k dispozici další vyčerpávající informace k projektu.

## **HISPRA Projekt záchrany vzácných historických tisků ve fondu Městské knihovny v Praze (MKP) prostřednictvím digitalizace**

O projektu HISPRA podrobně informuje Městská knihovna Praha na svém webovém sídle, kde jsou již k dispozici seznamy digitalizovaných dokumentů a odkazy na již publikované články s tematikou projektu HISPRA (<http://www.mlp.cz/hispra.htm>). Další podrobné údaje o stavu projektu byly zjištěny osobní konzultací s Mgr. Evou Měřínskou.

## **Digitalizace historických dokumentů knihovny Vojenského historického ústavu**

Tímto tématem se již zabývala diplomová práce M. Hokynka: Digitalizace starých tisků a rukopisů z fondu Knihovny Vojenského historického ústavu Praha [HOKYNEK, 2007]. Z této práce byly čerpány základní údaje projektu. Aktuální stav projektu byl zjištěn osobní konzultací s Bc. Zdeňkem Munzarem.

## **Reformátování bohemikálních časopisů z historického fondu STK**

Informace k programu digitalizace ve Státní technické knihovně v Praze byly vytěženy z Výročních zpráv a Ročních zpráv STK, které jsou volně dostupné na webovém sídle STK ([http://old.stk.cz/dalsi\\_informace.html](http://old.stk.cz/dalsi_informace.html)). Další podrobnosti byly získány osobní konzultací s PhDr. Petrou Večeřovou, řešitelkou tohoto projektu digitalizace v STK.



## **Digitalizace českých učebních textů z fondu STK**

Tento další projekt digitalizace ve Státní technické knihovně v Praze byl v době vzniku této práce v prvotním stadiu řešení, proto informace o něm mohly být získány pouze osobní konzultací s řešitelkou projektu PhDr. Petrou Pejšovou.

### **3. Digitalizační proces**

Procesem digitalizace rozumíme převod originálních dokumentů na papírovém médiu do bitmapové grafiky. V této grafice je celý obrázek popsán pomocí jednotlivých barevných bodů (pixelů), které jsou uspořádány do mřížky. Každý bod má určenu svou přesnou polohu a barvu. Poměr počtu pixelů, které jsou v mřížce uloženy vodorovně, a počtu pixelů uložených svisle určuje rozlišení obrazu (uváděno v počtu dpi). Základními technickými parametry digitálního obrazu je jeho rozlišení, barva a barevná hloubka, komprese a obrazové formáty.

#### **3.1. Obraz a jeho zpracování**

##### **3.1.1. Barva a barevná hloubka**

Barevná hloubka určuje množství barev, které slouží ke specifikaci hodnoty pixelů reprezentujících digitalizovanou předlohu. Tento údaj znamená počet možných barev v obrazu. Udává se v počtu barev nebo v bitech. Existují různé úrovně barevné hloubky: 1-bitový obraz, dvouvrstvý obraz (bilevel image), nebo černobílý (black-and-white) obraz. Paleta barev má pro každý pixel osm bitů a celkový počet barev je 256. Úroveň „High Color“ využívá 16 bitů a „True color“ 24 bitů, což představuje celkové množství 16 777 216 barev. Pro zobrazení na obrazovce je využit modelu RGB (Red, Green, Blue), kdy jsou všechny barvy smíchány z různého poměru tří základních barev: červené (red), zelené (green) a modré (blue). Při hloubce 24 bitů lze každou barevnou složku vyjádřit škálou 256 barev (8 bitů) - od nejtmaší po nejsvětější, což je celkem 16 777 216 možných barev [KNOLL, 2007a].

##### **3.1.2. Druhy obrazu**

Vektorový obraz obsahuje informace, které jsou uloženy ve formě vektorů. Tyto vektory se přepočítávají do bitmapy. Proto lze vektorový obraz snadno zvětšovat, aniž by utrpěla jeho kvalita. Vektorový obraz se užívá pro jednoduchou grafiku, např. nákresy a plány. Typickými formáty vektorového obrazu je EPS, AI, CDR, WMF, DXF, SVG [KNOLL, 2007a].

Rastrový obraz je bitmapovou grafikou. Struktura obrazu je složena z pixelů, což je znát při jeho zvětšování. Užívá se pro fotorealistické obrazy, při digitalizaci je tedy vždy produkován obraz rastrový. Typickými formáty rastrového obrazu jsou JPEG, TIFF, GIF, PNG [KNOLL, 2007a].

### **3.1.3. Kompresce obrazu**

Při digitalizaci je vytvořen primární obraz v nejvyšší kvalitě, který je uložen na lokálním úložišti. Pro zpřístupnění obrazu uživatelům a jeho přenos přes síť Internet (rychlé odeslání a stažení) je potřeba obraz komprimovat. Kompresce obrazu může být ztrátová nebo bezztrátová.

Při ztrátové kompresi je postupováno tak, že nepotřebná informace v obrazu je odstraněna a pouze relevantní část obrazu je zachována. Lidské smysly dokáží rekonstruovat originál díky své nedokonalosti – neschopnosti vnímat ztráty a naopak schopnosti dobudovat chybějící částičky v kontextu a analogii s obvyklými a očekávanými objekty. Ztrátová komprese bývá aplikována především v oblasti informačně bohatých barevných obrazů, zatímco u černobílého obrazu pouze v novějších přístupech. Nejrozšířenější dva barevné kompresní algoritmy jsou DCT (Discrete Cosine Transform) nazývaná též JPEG komprese, a Wavelet, používaná v novějších přístupech [KNOLL, 2007a].

Kompresce formátu JPEG je založená na metodě odstraňování pixelů o vysoké frekvenci (tj. pixely s výrazným jasnem - např. odstíny žluté, oranžové a červené), na které je lidské oko méně citlivé. Principem je rozložení obrazu do bloků osm krát osm pixelů, které se komprimují jako samostatné jednotky na základě tzv. kvantizačního koeficientu. Jeho hodnota je určována s pomocí tzv. kompresního poměru, který je hlavním parametrem pro redukci velikosti obrazu ve formátu JPEG. Při kompresi je vždy třeba nastavit kvalitu obrazu individuálně, neboť jakékoli numerické nebo procentuální vyjádření kompresních poměrů platí pouze v konkrétním prostředí daného obrazového editoru a hodnoty kompresních poměrů závisejí na vnímání objektů v obraze [KNOLL, 2007a].

Kompresce bezztrátová spočívá v dekomprimaci obrazu při otevření dokumentu, kdy se každý pixel znovu objeví na svém původním místě jako před kompresí. Původní informace tak zůstává zachována [KNOLL, 2007a].

## **3.2. Obrazové formáty**

Digitální obrazy jsou uloženy v počítačových formátech. Těchto formátů existuje značné množství, proto se zde zaměřím jen na několik nejvíce používaných. Jak již bylo popsáno výše, formáty obsahují informace o parametrech obrazu: rozlišení, barevné hloubce, kompresi a řadě dalších vlastností obrazu. Některé formáty (např.

TIFF) mohou podporovat různé hodnoty těchto parametrů, jiné zase pouze určité konkrétní hodnoty určitého parametru. Výběr konkrétních hodnot ovlivňuje kvalitu a charakter obrazu.

### **3.2.1. JPEG**

Formát JPEG vznikl v roce 1991 a byl představen pod názvem JPEG File Interchange Format (JFIF). Jedná se o formát webový užívaný pro fotorealistické obrazy. JPEG je velice rozšířený, známý standard pro předlohy se složitou barevnou strukturou, podporuje 16 milionů barev nebo škálu 256 odstínů šedi. Výhodou tohoto formátu je jeho široká podpora komerčními grafickými rastrovými editory (např. Adobe Photoshop, Corel Photo-Paint) a také webovými prohlížeči jako univerzálními nástroji pro příjem informací v prostředí WWW [VOJTÁŠEK, 2000a].

Formát JPEG se také vyznačuje příznivým poměrem mezi velikostí a kvalitou zobrazení. I při vysokém stupni komprese obraz zůstává vizuálně stále velmi kvalitní. Používá se pro ztrátovou kompresi, přičemž při každém dalším uložení obrázku se znovu komprimuje. Vyznačuje se také schopností progresivního kódování (tj. při otevírání se okamžitě načte celý, ale rozmazaný obraz, který se poté prokresluje), které je užitečné při větších velikostech obrazu a tím způsobeným pomalejším načítáním [VOJTÁŠEK, 2000a.].

### **3.2.1. GIF**

Formát GIF (Graphics Interchange Format) je nejstarší z formátů zde popisovaných. Podporuje pouze 256 barev, proto není vhodný pro fotorealistické obrazy. Používá se na webu pro jednoduchou grafiku, nákresy (perokresby) a jednoduchou sekvenční animaci. Používá bezeztrátovou kompresi LZW84 [KNOLL, 2007a].

### **3.2.2. PNG**

Formát PNG (Portable Network Graphics) byl vyvinut jako náhrada a vylepšení formátu GIF. Tento standard podporuje 16 milionů barev a navíc obsahuje osmibitovou průhlednost (tzv. alfa kanál), to znamená, že obrázek může být v různých částech různě průhledný. Ve formátu PNG je zabudována bezeztrátová komprese a některé programy umožňují nastavení kompresního poměru (např.

IrfanView). Lze konstatovat, že formát poskytuje nejúčinnější bezztrátovou kompresi mezi tradičními formáty [KNOLL, 2007a].

### **3.2.3. TIFF**

Formát TIFF podporuje jakékoli rozlišení a jakékoli množství barev a navíc nabízí řadu kompresních možností: bez komprese, LZW komprese (fotorealistické obrazy), CCITT Fax Group 3 a 4 komprese (černobílé obrazy) a teoreticky též kompresi vlastní JPG a další. Jeho využití je všude tam, kde předpokládáme další zpracování zejména barevného obrazu (v tomto případě se ukládá většinou nekomprimovaně). Tento standard poskytuje nejlepší klasické řešení pro černobílé obrazy při nastavení na kompresi CCITT Fax Gr. 4. Také nabízí podporu více obrazů uložených v jednom souboru [KNOLL, 2007a].

### **3.2.4. Formáty typu wavelet**

Mezi nejužívanější formáty typu wavelet patří JPEG2000 a MrSID. Tyto standardy nabízejí nejúčinnější bezztrátovou, ale také ztrátovou kompresi, avšak nejsou doporučeny pro web. Používají se ve speciálních aplikacích, kde lze garantovat budoucí komfort uživatele [KNOLL, 2007a].

JPEG2000 oproti formátu JPEG nabízí množství vylepšení. Vyšší odolnost proti chybám datového toku, možnost zpracovávat obrázky větší než 64000 krát 64000 pixelů, strukturu umožňující s obrázkem ukládat nejrůznější metadata, např. informace pro práci s nastaveními barev a ochranu autorských práv, možnost definice „zájmových oblastí“, které jsou od ostatních částí obrázku odlišeny prioritním umístěním v datovém toku a případně i vyšším rozlišením, možnost využití různých barevných módů (JPEG podporuje pouze RGB), speciální zpracování složených dokumentů, tedy především takových, které obsahují text a grafiku (JPEG byl k uvedenému účelu prakticky nepoužitelný) apod. Část JPEG2000 byla publikována jako standard ISO/IEC 15444-1:2000 [Krejčí, 2001].

MrSID (Multiresolution Seamless Image Database) je produktem firmy LizardTech. Využívá se pro digitalizaci předloh s extrémním rozlišením (kartografické dokumenty či družicové snímky). Byl navržen pro profesionály z oblasti geografických informačních systémů, ale velice se uplatnil také při digitalizaci a síťovém zpřístupnění historických map. Hlavní výhoda spočívá v tom, že uživatel

si nemusí ze serveru stahovat celý soubor o velikém objemu, nýbrž pouze vybraný detail v požadovaném zvětšení s pomocí plug-in nebo speciálního prohlížeče MrSID Viewer [HROTKOVÁ, 2006. s. 28].

### **3.2.5. Formát typu smíšeného obrazu**

Od roku 1998 výzkumné laboratoře americké telekomunikační firmy [AT&T](#) vyvíjely formát DjVu (v roce 2000 odkoupila technologii DjVu firma LizardTech), za účelem nahrazení formátu JPEG pro přenos obrazových dat po síti Internet. DjVu se řadí mezi formáty smíšeného rastrového obsahu. Jedná se o techniku komprese, obrazový formát a platformu pro doručování vícebarevných dokumentů. Užívá se pro obraz, který je rozdělen na popředí (komprimované jednobitovými kompresními algoritmy) a pozadí (komprimováno technologií wavelet). Typickým příkladem užití formátu DjVu jsou naskenované časopisy a noviny. DjVu se dekomprimuje podobně jako formát JPEG, ne naráz, ale přírůstkově. Standard DjVu je nepřekonatelný v účinnosti ztrátové komprese, je např. cca 10x účinnější než JPEG [VOJTÁŠEK, 2000a].

### **3.3. Metadata**

Digitální dokumenty jsou doprovázeny sekundárními informacemi, které jsou v souvislosti s elektronickými zdroji označovány jako metadata. Lze je definovat jako data o primárních datech ve formě strukturovaných textových informací. Jinými slovy lze říci, že smyslem metadat je poskytovat přidanou informační hodnotu k datům primárním.

Funkce metadat je popisná, selekční, archivační, administrativní, technická, právní, strukturální. V souvislosti s těmito funkcemi se rozlišují metadata pro účely popisu a identifikace, správy, právních nároků, technické funkčnosti, užití a archivace [CELBOVÁ, 2006].

Administrativní metadata jsou užívána pro řízení a správu informačních zdrojů, např. informace o umístění, údaje o době vzniku a poslední modifikaci, elektronický podpis aj. Archivační metadata se vztahují k procesům souvisejícím s dlouhodobou archivací digitálních informačních zdrojů a mají zajistit trvalou integritu a kontext dokumentu jako podmínku pro jeho zpřístupnění v budoucnosti. Metadata pro právní nároky dokumentují tyto nároky vztahující se k popisovanému zdroji nebo odkazující

na jiný zdroj uvádějící právní nároky spojené s přístupem nebo využíváním informačních zdrojů, např. autorská práva a oprávnění kopírování nebo reprodukce. Metadata pro užití charakterizující specifické podmínky pro užívání. Popisná metadata charakterizují původní objekt tak, aby byl uživatel schopen porozumět jeho obsahu. Podporují vyhledávání daného digitálního objektu. Strukturální metadata definují vnitřní organizaci digitálního objektu. Jsou nezbytná pro zobrazení a navigaci tohoto objektu. Technická metadata jsou vytvořena pro počítačový systém nebo vyrobená počítačovým systémem, která uvádějí, jak se systém nebo jeho obsah chová nebo co požaduje, aby mohl být provozován (protokol HTTP, parametry HW) [CELBOVÁ, 2006].

Oproti výše popsané definici metadat se poněkud liší jedna z nejnovějších definicí, kterou uvádí Katherine M. Wisser v glosáři webového sídla ECHO Project, North Carolina: “Metadata jsou strukturované informace, které popisují, objasňují, lokalizují a jinak usnadňují opětovné získání a užití informačního pramene [WISSER, 2007].”

Metadata vyjadřují identifikaci s originální předlohou, vyjadřují její strukturu fyzickou (celek, ročník, strana) i logickou (články, fotografie, ilustrace), umožňují zápis technických údajů o datových (obrazových, zvukových) souborech, které vznikají digitalizací, poskytují informace pro případnou komunikaci mezi systémy a další [KNOLL, 2006]. Údaje se obvykle vkládají přímo do zdroje (s výjimkou obrazových dokumentů), umísťují se např. v záhlaví dokumentu HTML [CELBOVÁ, 2006].

Použití vhodných metadat je velmi důležité pro vyhledávání a získávání dokumentů, zejména při jejich sdílení digitálními knihovnami. V současné době existuje mnoho metadatových formátů a každý projekt má možnost zvolit si formát nejlépe odpovídající jeho cílům. Praxe však ukázala, že je nezbytné používat formáty metadat s veřejně deklarovanou strukturou, uznávané standardy, a že je třeba dbát na důslednou nezávislost na jakýchkoliv konkrétních firemních nástrojích [KNOLL, 2006].

### **3.3.1. SGML**

SGML standard (Standard Generalized Markup Language) byl vytvořen pro potřeby sdílení dokumentů mezi různými platformami. Je normou ISO 8879-1986 [ISO

8879:1986, 1986]. Nenabízí přesně určenou sadu tagů, ale prostředky na vytváření tagů vlastních. Hlavním rysem SGML je velká flexibilita, kterou je však možno považovat zároveň za zápor, neboť nabízí tolik možností, že nástroje, které je dokáží všechny využívat, jsou příliš komplikované.

SGML je metajazyk, který umožňuje definování dalších značkovacích jazyků. Je základem pro další odvozené přístupy, které je můžeme nazývat značkovacími jazyky 2. generace: HTML (preskriptivní), XML (deskriptivní). Hlavní konstrukční jednotkou značkování na základě SGML je element, a každý element musí být definován externím pravidlem popisu obsahu. Pro toto definování jsou potřeba formální pravidla, která vyjádří standardy, popisující obsah. V SGML prostředí jsou stanoveny v DTD/Definice Typu Dokumentu (Document Type Definition) [KNOLL, 2007b].

### **3.3.1. XML**

XML (Extensible Markup Language) je podmnožinou SGML vytvořenou za účelem ukládání a výměny strukturovaných dokumentů v prostředí Internetu. Je to otevřený standard, který není vlastnictvím žádné společnosti. Byl přijat konsorciem W3C v roce 1998, kdy byla představena verze 1.0. XML je definován pomocí vlastních DTD. DTD pro XML je psáno pomocí syntaxe SGML, proto bylo zavedeno W3C Schema, aby jej nahradilo, takže XML dokument může být řízen DTD nebo Schématem [KNOLL, 2007b].

Formát XML je od počátku uzpůsoben pro potřeby všech jazyků, ne jen angličtiny. Jako znaková sada se používá ISO 10646 [ISO/IEC 10646:2003, 2003], což je 32bitová znaková sada, která dokáže pojmout všechny dnes používané znaky ze všech jazyků. Narozdíl od SGML je mnoho parametrů XML předem určeno a není možné je měnit (např. délka názvů značek, použité oddělovače a speciální znaky atd). Syntaxe zápisu dokumentů v XML je proti SGML přísná, což umožní mnohem snazší a levnější vývoj aplikací, které umožňují s XML pracovat [KOSEK, 2000].

XML jazyk umožňuje definovat a kontrolovat jakýkoli typ popisů, může je spojit s vnějšími daty, určuje strukturu digitalizovaných dokumentů a umožňuje jejich dlouhodobou čitelnost, zajišťuje validitu výstupních dokumentů [KNOLL, 2007b]. Značky umožňují v dokumentu zachytit důležité informace o struktuře a významu,



proto lze do XML dokumentu uložit nejenom text,ale např.obsah tabulky z databáze [KOSEK, 2000].

### **3.3.2. Formát MASTER pro popis rukopisů**

Formát MASTER (Manuscript Access through Standards for Electronic Records) vznikl na základě platformy TEI ([Text Encoding Initiative](#)) jako nástroj pro popis historických dokumentů (pimárně rukopisů, ale i starých tisků aj.), spočívající v aplikaci SGML. Z důvodu své obecnosti byl v projektu *Memoriae Mundi Series Bohemica* formát MASTER uznán jako metadatový standard pro popis historických dokumentů [UHLÍŘ, 2006].

Původní formát MASTER pro popis rukopisů byl zapracován do metadatového formátu TEI. Na základě tohoto formátu bylo definováno a v červnu 2001 předloženo DTD MASTER pro rukopisy [UHLÍŘ, 2006].

Standard MASTER a z něho i následně vycházející popis ve formátu TEI umožňuje popis strukturovat, je tedy možné v rámci popisu jednoho rukopisu popsat různé části popisovaného rukopisu a naznačit mezi nimi hierarchické vztahy. Základem bibliografického popisu je hlavička (TEI Header). MASTER je přizpůsobivější z hlediska variability popisovaného materiálu, orientace při zobrazení a postupů při vyhledávání. Standard je založen na hluboké strukturaci obsahových elementů a na relativně volném využívání funkčních elementů. Pevná jsou pouze pravidla syntaxe. Ve standardu MASTER lze tudíž pořizovat jak jednoduché, minimálně informačně bohaté záznamy, tak záznamy jdoucí do hloubky popisovaného originálního dokumentu. To znamená, že jeho praktické využití je velmi široké a flexibilní, je adaptovatelné pro různé účely i různou míru znalostí o materiálu, aniž je to na překážku využití v informačním systému [UHLÍŘ, 2006, s.5].

### **3.3.3. Formát METS**

Zařazování digitálních kopií dokumentů, jejich plných textů a dalších informací souvisejících s fyzickým dokumentem vedlo k formulování komplexního digitálního dokumentu. S tím souvisí také přechod z původně používaného schématu MASTER+ pro popis dokumentů na standardizovaný a více univerzální kontejnerový formát METS (Metadata Encoding and Transmission Standard), který se stal standardem pro rozsáhlé digitální archivy. Vznikl v roce 2001 a o údržbu a rozvoj standardu se stará Kongresová knihovna USA.

Standard METS je kontejnerový formát založený na schématu XML a je určen pro výměnu komplexních digitálních objektů mezi jednotlivými systémy. Skutečnost, že se jedná se o kontejnerový formát, znamená, že v něm mohou být obsažena data v různých metadatových formátech (např. metadata ve formátech MARC21, MARC XML, Dublin Core, MODS aj.) a formát tak slouží k propojení všech metadatových záznamů [VOJNAR, 2005].

Formát METS se vyznačuje náročnou datovou strukturou. METS obsahuje následující sekce metadat [VOJNAR, 2005]:

- o hlavičku, která obsahuje administrativní informace o METS dokumentu jako takovém (kdo a kdy ho vytvořil, kdo a kdy ho upravil apod.)
- o sekce popisných metadat, která mohou být vložena přímo v METS dokumentu anebo uvedena odkazem na externí zdroj
- o sekce administrativních metadat, která poskytují informace o jednotlivých objektech, o právech pro jejich zpřístupnění a šíření, o původním objektu, který byl vzorem pro digitalizaci apod.
- o sekce souborů, která obsahuje seznam všech fyzických souborů, z kterých se skládá popisovaný objekt nebo objekty, a jejich umístění
- o sekce strukturální mapy, která je klíčovou a povinnou částí každého METS dokumentu, zachycuje hierarchickou strukturu a vazbu mezi soubory, objekty a metadaty
- o sekce strukturálních odkazů, které umožňují odkazovat mezi jednotlivými uzly strukturální mapy
- o sekce pravidel chování, která lze využít pro definici akcí nebo událostí, jež mají nastat při manipulaci s částmi METS dokumentu.

V rámci standardu METS existuje také možnost definovat vlastní rozšíření, vzhledem k potřebě zápisu administrativních metadat, která nejsou standardizovaná. Proto byla na začátku existence formátu METS definována navíc některá rozšiřující schémata.

Tento původně americký formát se již celosvětově prosadil. Formát METS byl také implementován do systému Kramerius, což umožňuje exportovat digitální

dokumenty, a zejména jejich metadata, v tomto formátu [POLIŠENSKÝ, 2007].

### **3.3.4. MODS Schema**

Metadata Object Description Schema je metadatové schema, které obsahuje soubor bibliografických prvků navržených speciálně pro potřeby knihoven, lze je však použít i v jiných případech. Tímto systémem lze popsat prakticky jakékoliv médium (např. monografie, CD-ROM, mapy, webová sídla aj.). Standard MODS je založen na jazyce XML a díky tomu může převádět data z MARC či Dublin Core a opačným směrem, tedy z MODS do MARC či Dublin Core [Library of Congress, 2008].

MODS zahrnuje možnosti bohatého popisu jakéhokoliv dokumentu v jazyce XML, spravování složitých digitálních objektů a sjednocení metadat z knihovných databází založených na MARC s těmi, které využívají jiná schémata. Tento standard byl původně zamýšlen jako podmnožina MARC s jinými názvy prvků, nakonec vznikl jako samostatné schéma, které sice nezahrnuje celou sadu prvků MARC, ale obsahuje některé nové [Library of Congress, 2008].

První verze schematu MODS (verze 1.2) byla uvedena v červnu roku 2002. Postupně byly přidávány nové atributy a prvky a zdokonalována schopnost převoditelnosti dat z a do jiných formátů, aby se snížila ztrátovost a převedená data se co nejvíce blížila originálu. Systém byl celkově zdokonalován až k současné verzi MODS 3.3 [Library of Congress, 2008].

MODS ve verzi 3.3 obsahuje 20 základních prvků, kterými lze dokument popsat. Ty se dále dělí na podprvky a k jednotlivým prvkům i podprvkům lze přiřadit atributy. Všechny prvky i jejich atributy jsou nepovinné, prvky a podprvky jsou opakovatelné. Standard udržuje a dále rozvíjí MARC Standard Office a Kongresová knihovna také s pomocí uživatelů [Library of Congress, 2008].

### **3.4. Ochranné mikrofilmování dokumentů**

Kromě vytváření digitálních kopií dokumentů za účelem jejich uchování a zpřístupnění se dnes stále používá i technologie mikrofilmování dokumentů. Tato metoda je užívána u periodik, která jsou ohrožena degradací papíru a u kterých je nutné předpokládat, že se do budoucna nedochovají. Tyto dokumenty se reformátují na mikrofilm, který je z hlediska archivace jistou zárukou, neboť životnost mikrofilmu je cca 500 let, za předpokladu dodržení technických standardů.

Následující kapitola zásadně čerpá z dokumentu J.Hutaře: Ochrana novodobých knihovních fondů 19. a 20.století [HUTAŘ, 2007, s.141-145].

Technologie mikrofilmování se užívá již od 30. let 20. století, proto je dnes opravdu dobře propracovaná a standardizovaná. Za více než 70 let užívání metody mikrofilmování vzniklo množství standardů a mezinárodních norem, které určují pracovní postupy procesu mikrofilmování, vyvolávání a vlastnosti používaných filmů, definují požadavky na technické přístroje i na podmínky pro ukládání archivních mikrofilmů. Tvorba, kopírování a distribuce mikrofilmů a mikrofišů je ekonomicky výhodnější než digitalizace. Navíc je výhodou také skutečnost, že oproti digitálním médiím, pro jejichž přečtení je nezbytné využití techniky, mikrofilmy a mikrofiše lze přečíst pouhým okem nebo zvětšovací sklem. Vytvořené mikrofilmy a mikrofiše je také následně možné digitalizovat, což je výhodnější než vytvářet digitální kopii původní papírové formy dokumentu.

Využívání mikrofilmů a mikrofišů uživateli má i jisté nevýhody, například nutnost manipulace s filmem, který je třeba vložit do čtecího zařízení, a také nemožnost vyhledávání informací jiným způsobem než prostým čtením dokumentu.

Problém ochrany, uchování a zároveň i zpřístupnění dokumentů uživatelům je v dnešní době řešen užitím tzv. hybridní metody, která je kombinací digitalizace a mikrofilmování. Umožňuje archivování dokumentů na vhodném médiu (mikrofilm nebo mikrofiš) a zároveň jejich zpřístupnění v podobě digitálních objektů (databáze digitálních objektů na pevném disku, CD, online na síti apod.).

Digitalizace zlepšuje přístupnost, přenos a distribuci dokumentů, navíc umožňuje elektronicky retušovat nedokonalosti nasnímaného obrazu. Mikrofilm je v současnosti jediné médium skutečně vhodné pro archivní uložení velkého množství málo využívaných informací.

V současnosti se využívají dva odlišné postupy hybridní metody [HUTAŘ, 2007, s.148-150]:

1. **Metoda Film-first.** Podstata postupu Film-first spočívá v tom, že nejprve je dokument mikrofilmován a následně je mikrofilm skenován do digitální podoby. Výstupem jsou obrazové soubory v jednobitovém zobrazení nebo 16 bitové šedé škále, ve formátech TIFF, JPEG a GIF. Obrazové soubory ve formátu JPEG s minimálním rozlišením 300 dpi vyžadují následné úpravy a pro

zpřístupnění na síti konverzi do formátu DjVu. Tento obrazový formát je považován za základní, protože zachovává původní grafickou podobu dokumentu a zabezpečuje vyšší úroveň autenticity. Obrazové soubory jsou dále zpracovávány technologií OCR (optické rozpoznávání znaků), která konvertuje naskenovaný obraz (textovou stranu) v textový soubor, ve kterém se dá vyhledávat pomocí přirozeného jazyka. Textový formát neslouží pro zpřístupnění uživatelům, ale pouze pro fulltextové vyhledávání. Obrazové soubory jsou poté spojeny s metadaty, která umožňují orientaci i v rozsáhlém titulu.

2. **Metoda Scan-first.** Postup Scan-first je opačný než u metody Film-first, dokument je nejprve naskenován a výsledný digitální obraz je poté potom převeden na mikroformu (mikrofilm, mikrofiš). U tohoto postupu je nutná kvalita skenování, které je prvním krokem. Digitální obraz, naskenovaný v šedé škále, se elektronicky vylepší (retuší, zvyšováním kontrastu, úpravou barevnosti) a použije na vytvoření kvalitního filmu za použití elektronového paprsku nebo digitálním fotoaparátem s mikrofilmovým výstupem. Zároveň je obraz s co nejkvalitnějším rozlišením úrovní šedi převeden do vysoce kvalitního binárního obrazu s redukováním rozlišením, který je považován za nejvhodnější pro zpřístupnění uživatelům.

#### **3.4.1. Proces mikrofilmování**

Celý proces mikrofilmování dokumentů lze popsat jako postup o následujících bodech [HUTAŘ, 2007, s.146-147]:

- o **Výběr dokumentů.** Každá knihovna má vlastní koncepci výběru dokumentů pro mikrofilmování a také vlastní pravidla. Výběr konkrétních titulů provádí pracoviště zajišťující výpůjční služby, katalogizační a bibliografická pracoviště, oddělení speciálních sbírek. Pro mikrofilmování se vybírají dokumenty nejvíce ohrožené chemickou degradací papíru a nadměrným využíváním, především noviny a časopisy, rukopisy a další dokumenty.
- o **Příprava a kompletace předloh.** Dokument musí být zkontrolován z hlediska úplnosti (pokud např. stránka chybí, dá se doplnit z jiného exempláře i z jiné knihovny), z hlediska optické kvality a z hlediska způsobilosti pro mikrografické zpracování. Během přípravy se vytvoří

podrobná bibliografická předloha, která se spolu s identifikační a obsahovou předlohou snímkuje na počátek filmu. Předloha musí být čitelná i pouhým okem. Chyby ve stránkování, vazbě a další nepravidelnosti se značí dvojjazyčným návěštím, které se snímkuje v místě výskytu daného problému.

- o **Snímání předloh** je nejdůležitější částí celého procesu mikrofilmování, na které závisí kvalita i trvanlivost záznamů. Snímání se provádí snímacími kamerami na mikrografický film 35 mm. Snímáním se vytvoří tři generace mikrofilmů s odlišnými funkcemi: archivní negativ (slouží pro trvalé dochování obsahu dokumentu, zhotovuje se snímkováním předlohy na mikrografické kameře), matriční negativ (slouží pro zhotovování všech uživatelských kopií, vytváří se kopírováním archivního negativu) a uživatelská kopie v pozitivním tónovém zobrazení (nahrazuje originální dokument, zhotovuje se kopírováním matričního negativu).
- o **Kontrola kvality a archivního mikrofilmu.** Archivní mikrofilm se kontroluje z hlediska optické kvality a obsahové úplnosti. Kvalita mikrofilmu se vyhodnocuje podle zkušebních obrazců Test chart No. 1 a Test chart No. 2 dle normy ISO 4087 [ISO 4087. 2005], snímkaných na začátku mikrofilmu [Česko, příl. č. III, 2008].
- o **Uchovávání mikrofilmů.** Mikrofilmy jsou upevněny na cívkách z nekorodujících a neutrálních materiálů a vloženy do krabiček z nekyselých lepenky. Pro ukládání slouží většinou kovové skříně s plochými nebo vysokými zásuvkami, se speciálními parametry mikroklimatu (předepsaná vlhkost 40 % a teplota pod 19°C).
- o **Zabezpečení mikrofilmovaných dokumentů.** Mikrofilmované dokumenty jsou vyřazeny z běžného výpůjčního procesu. Z katalogizačního záznamu se uživatel dozví pouze o dostupnosti mikrofilmu. Mikrofilmované dokumenty jsou v depozitářích viditelně označeny, aby nedošlo omylem k zapůjčení originálního dokumentu namísto mikrofilmu.

### 3.5. Ochrana vytvořených dat

Digitální objekty, které vzniknou digitalizováním vzácným a ohrožených dokumentů, je potřeba dlouhodobě uložit a ochránit. Digitální dokumenty lze



poměrně snadno, rychle a flexibilně podle potřeby zpřístupnit a modifikovat a digitální záznam vydrží potenciálně neomezeně dlouho. Ochrana a archivace digitálního záznamu se definuje jako soubor vzájemně provázaných opatření a metod technické a organizační povahy týkajících se uložení, administrace a zpřístupnění digitálních objektů, přičemž smyslem je zabezpečit, aby bylo možné jejich přečtení v dlouhodobé perspektivě (po dobu, která není předem ohraničena) s vědomím, že vlastnosti technických prostředků, které budou k tomuto účelu aplikovány v budoucnosti, nelze v současnosti dostatečně popsat [VOJTÁŠEK, 2000b, kap.2].

Zprvu se předpokládalo, že nejkritičtější faktor pro uložení, ochranu a zpřístupnění digitálních dokumentů představuje právě trvanlivost médií pro uložení digitálního záznamu. Dřívější udávaná životnost CD-ROM byla až dvě stě let, což se ukázalo jako nereálné. Skutečná životnost se pohybuje v rozmezí 5 až 20 let podle frekvence využívání a způsobu uložení optických disků. Zkušenosti z realizace projektů digitalizace v konfrontaci s poznatky získanými z dosavadního vývoje softwaru a hardwaru prokázaly, že čitelnost digitálního záznamu významněji ovlivňují přístupový software, svázanost s původní softwarovou a hardwarovou platformou a datové a metadatové formáty. Přesto je samozřejmě životnost nosičů potřeba nadále brát v úvahu [VOJTÁŠEK, 2000b, kap.2].

Ohrožení pro zachování funkčnosti digitálních dokumentů tedy představuje především technika, která je zapotřebí k jejich dekodování, a která velmi rychle zastarává. Proto je nejdůležitější zachovat informace dostupné a čitelné v jejich formátu. Instituce, která se zabývá digitalizací dokumentů a tím produkuje vlastní digitální dokumenty, proto musí využít takové řešení vyhovující konkrétním podmínkám, které umožní dlouhodobé využívání těchto dokumentů.

V posledních letech se ustanovil konsensus o souboru základních principů ochrany, které jsou v digitálním světě stejné jako ve světě analogovém. V podstatě je určují priority pro rozšiřování užitečnosti a životnosti informačních zdrojů. Těmito základními principy jsou [HUTAŘ, 2007, s.156-157]:

- o Životnost. Životnost digitálních objektů do značné míry závisí na životnosti zařízení, která k nim zajišťují přístup. Knihovny musí být připraveny přenášet a přizpůsobovat data, aby byla čitelná i pro další generace systémů.

Knihovník může ovlivnit životnost např. kompaktního disku, ale vývoj technologií neovlivní.

- o Výběr. Zahrnuje určení hodnoty, rozpoznání a rozhodnutí. Výběr v digitálním světě není volbou jednou pro vždy, ale je to souvislý proces spojený s aktivním využíváním digitálních souborů. Vždy se tak musí dít v návaznosti na celý systém a instituci, její cíle.
- o Kvalita. V digitálním světě je podmíněna omezením kvality zachycení obrazu originálního dokumentu a poté i zobrazení digitálního dokumentu i. Cílem je zachytit intelektuální a vizuální obsah originálu v maximální možné kvalitě, tak jak dovoluje technika a stejným způsobem ho prezentovat i uživatelům.
- o Integrita. Má dvě roviny, fyzickou (stav) a intelektuální (pravost a hodnověrnost). Udržení fyzické integrity není podmíněno médii, na kterém jsou data uložena, ale hrozí ztráta informací během procesu vytváření objektu, jeho komprese či zasílání po síti. V oblasti intelektuální integrity se jedná také o propojení objektů s indexy, metadaty. Tyto popisné údaje musí být uloženy a ochraňovány společně se samotnými objekty.
- o Přístup. V digitálním prostředí je přístup vedlejším produktem ochrany dokumentů.

#### **4. Programy dotací pro projekty digitalizace v České republice**

Knihovny, které se rozhodly pro digitalizaci dokumentů se svých fondů, mají na výběr z několika možností získání dotace pro svůj digitalizační program.

Program VISK *Veřejné informační služby knihoven* byl schválen usnesením Vlády ČR ze dne 10. 4. 2000 č. 351 o Koncepci státní informační politiky ve vzdělávání [Česko, 2000]. Cílem programu se stala inovace veřejných informačních služeb knihoven na bázi informačních technologií (ICT). Program VISK připravuje Ministerstvo kultury ČR ve spolupráci s Národní knihovnou ČR a Svazem knihovníků a informačních pracovníků ČR. Program VISK je členěn na devět vzájemně provázaných podprogramů. Financování programu v jednotlivých letech se odvíjí od vývoje státního rozpočtu a rozpočtu Ministerstva kultury. Pro podporu digitalizování knihovnických fondů lze využít programy VISK 6 *Národní program digitálního zpřístupnění vzácných dokumentů Memoriae Mundi Series Bohemica* a VISK 7 *Národní program mikrofilmování a digitálního zpřístupňování dokumentů ohrožených degradací kyselého papíru – Kramerius* [VISK, 2005-01-11].

Další možností pro financování projektů digitalizace dokumentů je využití dotace prostřednictvím Finančních mechanismů Evropského hospodářského prostoru a také Norského finančního mechanismu.

##### **4.1. VISK 6 Národní program digitálního zpřístupnění vzácných dokumentů Memoriae Mundi Series Bohemica**

*Memoriae Mundi Series Bohemica* je program digitálního zpřístupnění a ochrany kulturního dědictví, které je obsaženo v dokumentech dochovaných na území České republiky. Od svého založení Národní knihovnou ČR ve spolupráci s firmou Albertina icome Praha s.r.o. roku 1995 patří do programu UNESCO *Paměť světa*, pro který bylo zpracováno několik pilotních projektů, například první CD-ROM celého programu a také první publikovaný digitalizovaný rukopis [KNOLL, 1997].

Následující kapitola zásadně čerpá z informací uvedených na webovém sídle programu VISK [VISK, 20058-06-11].

Program si již získal jméno po celém světě. Koordinátor programu je členem Subkomise pro technologii Paměti světa UNESCO. Zástupci programu se účastí různých mezinárodních projektů, jejichž cílem je vytvoření virtuálního badatelského prostředí. Program a jeho financování jsou součástí kulturní politiky vlády ČR,

v současné době jde zejména o využití technologií a know-how, kterými disponuje Národní knihovna ČR v důsledku úspěšného řešení projektů výzkumu a vývoje.

Základním cílem programu je zajistit metodou digitalizace ochranu a zpřístupnění vzácných dokumentů knihoven a dalších sbírek, které tvoří důležitou součást národního kulturního dědictví. Program zabezpečuje provoz technologií pro digitalizování vzácných dokumentů a vytváří podmínky pro co nejširší účast knihoven a dalších institucí, v jejichž vlastnictví se nacházejí vzácné fondy v souladu se zásadami programu UNESCO *Paměť světa*.

Převodem vzácných dokumentů do digitální podoby se zkvalitní jejich ochrana a zároveň budou k dispozici badatelům na celém světě. Realizací programu bude umožněno zajistit ochranu nenahraditelného národního a světového kulturního bohatství. Vzácné dokumenty dochované na území České republiky již nebude potřeba vystavit riziku poškození při kontaktu s uživateli. Některé z nich jsou v důsledku dlouholetého používání ve velmi špatném stavu, proto je jejich digitalizace často i jediným možným způsobem zpřístupnění. Tím, že se zabezpečí digitální přístup k dokumentům, je možné stav originálních dokumentů zakonzervovat a fyzicky je stabilizovat. Možností digitálního přístupu k dokumentům se zároveň se velmi zvyšuje okruh jejich možných uživatelů na celém světě.

#### **4.2. VISK 7 Národní program mikrofilmování a digitálního zpřístupňování dokumentů ohrožených degradací kyselého papíru – Kramerius**

Národní program Kramerius si klade za cíl záchranu a zpřístupnění bohemikálních dokumentů tištěných na kyselém papíru, který se postupem času rozpadá a tím se existence těchto dokumentů stává ohroženou. Tento špatný fyzický stav starších novin a časopisů výrazně omezuje využívání této součásti národního kulturního bohatství, u převážné části titulů znemožňuje meziknihovní výpůjční službu, znesnadňuje zhotovování kopií pro uživatele, zkracuje jejich životnost a zhoršuje možnosti uchování takto poškozených dokumentů pro budoucnost [VISK, 2005-01-13].

Program Kramerius umožňuje postupné převedení nejvíce ohrožených dokumentů na mikrofilm, který zajišťuje jejich trvalé uchování a následnou digitalizaci mikrofilmu, čímž se vytváří kopie pro zpřístupňování např. pomocí Internetu. Vznikají tak

archivní mikrofilmy zajišťující dlouhodobé dochování dokumentu, matriční negativy sloužící k výrobě uživatelských kopií mikrofilmů a uživatelské pozitivní mikrofilmy nahrazující originální dokument. Pokud se v rámci projektu mikrofilmy digitalizují, jsou zhotovovány první dvě generace mikrofilmu (archivní a matriční) a digitální kopie, která slouží jako náhrada ohroženého dokumentu.

Knihovny, které požádají o podporu v programu VISK 7, jsou povinny dodržovat formáty dat definované Národní knihovnou. Jako výrobní nástroj pro vytváření metadat je užíván systém Sirius. V současné době je využíván Národní knihovnou ČR, Knihovnou AV ČR a firmou Elsys Engineering, která tento systém vyvíjí. Vytvořená data jsou zpřístupněna v systému Kramerius. Vývoj tohoto systému byl iniciován Národní knihovnou ČR a je vyvíjen jako open source systém, který je kompatibilní s formátem dat pro VISK 7 a je k dispozici zdarma.

Ochranné reformátování umožňuje knihovnám zachránit publikace ohrožené degradací papíru a zachovat jejich obsah uživatelům i v budoucnosti. Kromě toho program zajišťuje koordinaci výběru titulů a tím předchází neekonomickým duplicitním činnostem v případě opakovaného reformátování stejného titulu na více místech. Koordinace se zajišťuje nejen v rámci České republiky, ale díky členství Národní knihovny ČR v evropském konsorciu EROMM (European Register of Microform Masters) nejen na evropské, ale také na mezinárodní úrovni.

#### **4.3. Finanční mechanismy Evropského hospodářského prostoru a Norský finanční mechanismus**

V květnu 2004 založily tři nečlenské státy EU Island, Knížectví Lichtenštejnsko a Norské království *Finanční mechanismus Evropského hospodářského prostoru (EHP)*. Norské království založilo i vlastní, *Norský finanční mechanismus*. Prostřednictvím finančního mechanismu státy EHP v letech 2004-2009 přispívají zemím přistupujícím do EHP a dále Španělsku, Portugalsku a Řecku na projekty v rozšířeném vnitřním trhu. Hlavním cílem finančních mechanismů je snížení sociálních a ekonomických nerovností v Evropském hospodářském prostoru v rámci vymezených [prioritních oblastí](#) (viz níže) [Česko, 2005a].

Dohodou z října 2003 (a následném podepsání memorand o porozumění v říjnu a prosinci 2004) byla Česká republika začleněna do Evropského hospodářského prostoru (EHP). Částka určená pro ČR je 110,91 milionů eur pro pětileté období

2004-2009, z toho 48,54 milionů z Finančního mechanismu EHP a 62,37 milionů z Norského finančního mechanismu.

Finanční mechanismus EHP má určeno šest prioritních oblastí:

1. Uchovávání evropského kulturního dědictví
2. Ochrana životního prostředí
3. Rozvoj lidských zdrojů
4. Zdravotnictví a péče o dítě
5. Podpora udržitelného rozvoje
6. Vědecký výzkum a vývoj.

Každá z těchto priorit má ještě určeno několik konkrétních oblastí zaměření priority. Oblasti kultury je přímo věnována priorita 1. *Uchovávání evropského kulturního dědictví*, jejíž oblast zaměření, prioritu 1.1 *Ochrana a obnova nemovitého kulturního dědictví* lze přímo využít pro finanční podporu projektů digitalizace dokumentů.

Uvedených šest priorit je identických i pro Finanční mechanismus Norska, který ale disponuje ještě dalšími dvěma prioritami 7 a 8 [Česko, 2005a]:

7. Implementace Schengenského práva, posilování justice
8. Technická pomoc.

Finanční mechanismus EHP řídí *Výbor pro řízení finančního mechanismu EHP (EEA Financial Mechanism Committee)*, který přijímá rozhodnutí o poskytování finanční pomoci. Stejnou roli v Norském finančním mechanismu zastává *Ministerstvo zahraničí Norského království*. Těmto nejvyšším orgánům s každodenním řízením pomáhá *Kancelář pro finanční mechanismus EHP/Norsko (Financial Mechanism Office)*, přidružená k sekretariátu *Evropského hospodářského prostoru* v Bruselu, která slouží zároveň jako kontaktní místo. Kancelář finančního mechanismu v Bruselu přijímá žádosti prostřednictvím *Národních kontaktních míst* ve státech příjemců. Tato kontaktní místa jednají jménem státu příjemce a nesou odpovědnost za řízení aktivit finančního mechanismu EHP/Norska, včetně finanční kontroly a auditu [JEDLIČKA, 2005].

Národním kontaktním místem v České republice je ustanoveno *Ministerstvo financí ČR*, kde jeho administrativní činnost zabezpečuje Centrum pro zahraniční pomoc (oddělení koordinace zahraniční pomoci) [JEDLIČKA, 2005].

Příspěvek z Finančního mechanismu EHP/Norska nesmí přesáhnout 60 % celkových nákladů na projekt, výjimkou jsou projekty spolufinancované ze státního, regionálního nebo místního rozpočtu, u nichž tento příspěvek nesmí přesáhnout 85 % celkových nákladů. Zároveň také platí, že kombinovaná podpora z finančního mechanismu EHP, Norského finančního mechanismu a grantů Evropského společenství nesmí překročit 90 % celkových výdajů [JEDLIČKA, 2005].

## 5. Systémy Manuscriptorium a Kramerius

Manuscriptorium a Kramerius jsou Národní programy ČR pro digitalizaci VISK 6 a VISK 7 a zároveň funkčními digitálními knihovnami. Jedná se o dva hlavní programy digitalizace v České Republice, na kterých se podílí množství knihoven, archivů, muzeí a dalších institucí se svými projekty digitalizace knihovních fondů.

Programy Manuscriptorium a Kramerius mají určeny své standardy, které jsou zúčastněné instituce povinny dodržovat. Tím je docíleno určité uniformity a přehlednosti digitálních záznamů a také usnadnění jejich implementace do digitálních knihoven Manuscriptorium a Kramerius.

### 5.1. Manuscriptorium

Systém digitální knihovny Manuscriptorium (<http://www.manuscriptorium.com>) shromažďuje a zpřístupňuje na síti Internet informace o historických knižních fondech, a je provázán s digitální knihovnou digitalizovaných dokumentů. Provoz Manuscriptoria je financován Národní knihovnou ČR a provozovatelem je firma AiP Beroun s.r.o. [Manuscriptorium, 2006, s. 3].

Cílem Manuscriptoria je výsledky programu *Memoriae Mundi Series Bohemica* v Národním programu VISK 6 (<http://visk.nkp.cz>) zpřístupnit uživatelům, shromáždit výsledky souvisejících aktivit týkajících se digitalizace a zpřístupnit historické dokumentů v integrovaném zdroji. Tím se vytvoří virtuální badatelské prostředí pro práci s historickými fondy. Základem Manuscriptoria je národní program *Memoriae Mundi Series Bohemica*, avšak postupně integruje data také dalších projektů a institucí nejenom v českém, ale i středoevropském měřítku [PSOHLAVEC, 2004, s.56].

Na projektu Manuscriptorium spolupracuje s Národní knihovnou ČR množství partnerů, ke dni 22. 10. 2008 bylo uvedeno na webovém sídle Manuscriptoria (<http://www.manuscriptorium.com/Site/CZE/reference.asp>) třicet tři partnerských institucí z České Republiky a jedenáct partnerských institucí zahraničních.

Těžištěm systému Manuscriptorium je otevřený katalog historických fondů (OKHF) a k němu připojených vlastních digitálních dokumentů, které jsou uloženy v datovém úložišti. OKHF soustřeďuje metadata o fyzických dokumentech ve formě evidenčních záznamů ve formátu XML, které jsou volně zpřístupněné široké



veřejnosti. Datové úložiště pak obsahuje digitální kopie z těchto katalogizovaných dokumentů ve formě komplexních digitálních dokumentů. Systém Manuscriptoria zajišťuje přístup k digitalizovaným dokumentům na datových úložištích provozovatele i na vzdálených úložištích partnerů. Manuscriptorium zpřístupňuje volně veškerá data (obrazová, zvuková, plné texty aj.) svým [partnerům](#), ostatním uživatelům pak prostřednictvím [licencí](#) [Manuscriptorium, 2006, s. 4] .

Databáze je nejen vstupem do digitální knihovny, ale také komplexním zdrojem umožňujícím různé způsoby práce s historickými dokumenty a směřuje k tomu, aby se pro tyto dokumenty a jejich uživatele stala mezinárodním badatelským prostředím. Manuscriptorium je dnes největší digitální knihovnou historických dokumentů (tj. do 18. století) na světě, obsahuje přes 2 000 plně digitalizovaných dokumentů, což činí přes 1 milion obrázků v nejvyšší kvalitě.

Manuscriptorium je výsledkem aktivit, které začaly v rámci iniciativy UNESCO „*Memory of the World*“ v roce 1992, kdy byly v Národní knihovně České Republiky zahájeny aktivity na poli digitalizace ve spolupráci s firmou AiP Beroun. Aktivity pokračovaly v roce 2000 vytyčením úkolů navazujícího českého národního programu *Memoriae Mundi Series Bohemica* řešením souvisejících aktuálních požadavků a praktickým provozem. Během let 2003-2004 byla zprovozněna digitální knihovna Manuscriptorium. V roce 2005 získala Národní knihovna ČR cenu [UNESCO Jikji](#), zejména pro aktivitu v rámci digitálního zpřístupnění vzácných sbírek [Manuscriptorium, 2006, s. 3].

### **5.1.1. Standardy Manuscriptoria**

Základní formou pro uchování informací v Manuscriptoriu jsou XML soubory se strukturou odpovídající standardu METS. Pro uchování popisných metadat v rámci METS je použit standard MASTER, speciálně vyvinutý mezinárodní pracovní skupinou ve spolupráci s Počítačovou službou Oxfordské univerzity (Oxford University Computing Services) pro oblast popisu rukopisů, starých tisků a dalších historických dokumentů.

Za přímo kompatibilní se považují metadata vytvářená libovolnými nástroji, jejichž výstupem jsou záznamy v XML, využívající DTD MASTER. Data pro evidenční záznamy musí být dodávána ve formátu XML a v kódování UNICODE UTF-8. Pro strukturu XML souboru je předepsán standard MASTER. Pro generování digitálních

dokumentů umožňujících přístup k distribuovaným datům na úložištích příspěvatelů bude použit formát XML MASTER+. Obrazová data musí být uložena ve formátech, které jsou přímo podporované internetovými prohlížeči. Jsou to formáty JPEG a JPEG 2000, GIF, TIFF a PNG [KUČERA, 2006, s. 4].

Požadavky na formát obrazů jsou rozlišovány podle toho, zda v něm jsou uloženy primární obrazy nebo obrazy odvozené. U primárních obrazů je doporučeno použití takových formátů, pro které platí: že definice formátu jsou standardizované a veřejně dostupné, že nástroje pro tvorbu a reprodukci obrazů v těchto formátech jsou součástí základního programového vybavení počítačů a že obrazy v těchto formátech jsou bez omezení konvertovatelné do jiných formátů .

V databázi Manuscriptorium se využívá formát TIFF pro snímání a zpracování obrazů, formát JPG s mírnou kompresí pro uložení primárních dat. Pro primární obrazy se nepoužívá formát PDF a formáty s odděleným zpracováním a oddělenou komprimací obrazu v předpokládaných informačních vrstvách (písmo a pozadí).

Popis dokumentů vždy musí předcházet vlastní digitalizaci. V počátcích digitalizace byl zaveden formát DOBM (Digitization of Old Books, Manuscripts, and Other Documents) využívající SGML, který byl postaven na principu pevně strukturovaného popisu. Jeho výhoda spočívala v jednoduchosti a tedy následnému usnadnění práce. Nevýhodou však bylo, že využívá pouze tvrdě strukturovaných dat na jedné nebo vůbec nestrukturovaných dat na druhé straně. Z toho důvodu byl princip pevně strukturovaného popisu v dalším vývoji opuštěn a byl nahrazen standardem MASTER (Manuscript Access through STandard for Electronic Records) vytvořeným v rámci evropského projektu, jehož řádným partnerem byla i Národní knihovna ČR.

Standard MASTER (nejprve na bázi SGML, posléze XML) umožňuje vytváření a využívání zejména semistrukturovaných dat, tzn. je přizpůsobivější jak variabilitě popisovaného materiálu, tak orientaci při zobrazení a postupům při vyhledávání. Je založen na strukturaci obsahových elementů do hloubky i na relativně volném využívání funkčních elementů, které se mohou vztahovat k různým horizontálním i vertikálním místům ve struktuře celého popisu. Pevná jsou pouze pravidla syntaxe. Ve standardu MASTER lze tudíž pořizovat jak zcela jednoduché, minimálně informačně nasycené záznamy, tak záznamy jdoucí do hloubky popisovaného

originálního dokumentu. To znamená, že jeho praktické využití je velice široké a flexibilní, adaptovatelné pro různé účely i různou míru znalostí o materiálu, aniž je to na překážku využití v informačním systému [UHLÍŘ, 2006, s. 4-5].

Pro využívání dat Manuscriptoria jinými systémy se využívají nástroje pro zpřístupnění a sdílení metadat ve standardních formátech (UNIMARC, MARC21, MASTER, Dublin Core ) pomocí standardních komunikačních rozhraní, jako je protokol Z39.50 a OAI-PMH [Manuscriptorium, 2006, s. 7].

## 5.2. Kramerius

Kramerius je označení pro Národní program ochranného mikrofilmování a digitálního zpřístupňování dokumentů ohrožených degradací kyselého papíru. Program se zrodil v roce 1997 v rámci pilotního projektu výzkumu a vývoje, kdy byla převedena na mikrofilm část závažně poškozených, významných bohemikálních periodik. V roce 2000 byl Kramerius zařazen jako podprogram VISK 7 do programu Veřejné informační služby knihoven (<http://visk.nkp.cz>).

Koncepce programu při Národní knihovně ČR předpokládá využití alternativních technologií a/nebo jejich kombinaci: digitální skenování, mikrofilmování, různé metody konzervace, program ochranných obalů apod.

Žadatelé o dotaci v rámci programu se zavazují ke splnění těchto koncepčních podmínek: respektování normy a doporučení, koordinace vlastní práce s činností ostatních pracovišť, poskytování bibliografických záznamů do společné centrální databáze záznamů reformátovaných dokumentů (CZROMM), zpřístupňování uživatelské kopie reformátovaných dokumentů a zajišťování služeb, napomáhání při ochranném reformátování ohrožených titulů dalším institucím (např. zápůjčkami dokumentů nebo poskytnutím bibliografických informací), a dlouhodobé uchovávání archivní kopie reformátovaných dokumentů [VISK, 2005-01-13].

Program VISK zajišťuje kromě vlastního reformátování také archivaci digitálních dokumentů v rámci podprogramu VISK 4 *Digitální knihovna* a archiv pro informační služby knihoven (<http://visk.nkp.cz/VISK4.htm>). Centrální archivace zdrojových dat je ekonomicky výhodná, neboť není nutné pořizovat drahá zařízení v ostatních institucích účastnících se programu VISK 7. Archivaci mikrofilmů zajišťují knihovny samy, pokud k tomu mají vhodné podmínky, tedy speciální sklad s předepsanými parametry mikroklimatu. Národní knihovna ČR zajišťuje také

centrální archivaci mikrofilmů pro instituce, které tyto podmínky k archivaci nejsou schopny samy splnit [VISK, 2005-01-13].

### 5.2.1. *Systém Kramerius*

Kramerius je ale také označení systému pro zpřístupnění digitalizovaných periodik, která vznikla v rámci Národního programu VISK 7 Kramerius. Vývoj systému Kramerius byl iniciován Národní knihovnou ČR a je vyvíjen jako open source systém, který metadata archivuje a je kompatibilní s formátem dat pro VISK 7. Systém Kramerius je k dispozici zdarma (<http://kramerius.nkp.cz>).

Jako výrobní nástroj pro vytváření metadat digitálních dokumentů se používá systém Sirius. V současné době je využíván Národní knihovnou ČR, Knihovnou AV ČR a firmou Elsys Engineering, která jej vyvíjí. Jedná se o komerční produkt, který si knihovny musí zakoupit. V systému Sirius je zabudován protokol Z3950 a k jeho instalaci je potřebný operační systém Linux [VISK, 2005-01-13].

[Systém Kramerius](#) představuje specifický typ systému CMS (Content Management System), jehož účelem je zpřístupnění archivních a cenných dokumentů široké veřejnosti tak, aby nedošlo k porušení autorského zákona. Metadata jsou proto přístupná všem, stejně jako obrazová data dokumentů, které nejsou chráněny autorským zákonem. Obrazová data chráněná autorským zákonem jsou přístupná uživatelům pouze z počítačů umístěných v prostorách Národní knihovny ČR [Národní knihovna, 2006-05-11].

Systém Kramerius je součástí širšího systému sestávajícího z více provázaných aplikací, které tvoří tzv. [výrobní linku](#). V jejím rámci se provádí digitalizace dokumentů, vytváří se primární metadata pomocí technologie OCR, následuje úprava metadat a obrazových podkladů a konečně zpřístupnění dokumentů pomocí WWW prohlížeče na síti Internet. Kramerius poskytuje intuitivní uživatelské rozhraní pro vyhledávání informací o archivních dokumentech i prohlížení jejich obsahu. Aplikace pro zpřístupnění digitalizovaných dokumentů může pracovat i ve více instancích v rámci různých institucí a poskytuje prostředky pro replikaci obsahové báze zpřístupňovaných titulů mezi těmito institucemi. Přes administrativní rozhraní aplikace lze řídit způsob replikace dat, spojovat a řadit jednotlivé objekty dokumentů (ročníky, stránky atd.) do větších celků, importovat/exportovat metadata a obrazová data, definovat přístupová práva k dokumentům pro aplikování autorského zákona

a tak postupně budovat a spravovat kompletní obsahovou bázi zpřístupňovaných dokumentů [Národní knihovna, 2006-05-11].

Systém Kramerius byl vytvořen v souladu s nutností standardizace metadat pro potřeby knihoven, archivů a dalších podobných institucí. Aplikace i externí moduly pro tvorbu metadat i pro řízenou konverzi byly vytvořeny na základě licence GNU GPL a využívá volně dostupné komponenty. Pro užívání modulu pro tvorbu metadat si uživatelé musí zakoupit licenci pro Corel XMetaL Author. Na realizaci projektu se podíleli partneři a subdodavatelé řešení, jimiž byli IBM, s.r.o., Elsys Engineering a INCAD, s.r.o. a také tým odborníků společnosti Qbizm technologies [Národní knihovna, 2006-05-11].

### **5.2.1. Standardy Krameria**

Knihovny, které požádají o podporu v programu VISK 7, jsou povinny dodržovat formáty dat a metadatové standardy definované Národní knihovnou ČR.

Dokumenty jsou digitalizovány pomocí skenerů mikromédií (tj. mikrofilmů a mikrofiší) ve spolupráci s aplikací Sirius, která zajišťuje potřebné úpravy obrazových dat a následné automatizované generování XML metadat pomocí dávkových operací typu zónového rozpoznávání, zónového čtení, OCR. Tento systém zefektivňuje proces digitalizace a snižuje náklady na jeho realizaci [Národní knihovna, 2004-09-23].

Výstupem jsou obrazové soubory v 16 bitové šedé škále, ve formátech TIFF, JPEG a GIF. Pro zpřístupňování dokumentů se používají v systému Kramerius kopie všech obrazových souborů konvertované do formátu DjVu, který má výhodnější kompresní parametry, je optimalizován pro archivaci a jeho plug-in je podporovým operačním systémem Windows. Metadata ve struktuře založené na XML. Pomocí systému lze provádět jednoduché operace s dokumenty, např. jejich spojování, nahrazování, mazání a replikace. Součástí systému jsou i externí moduly pro tvorbu metadat v XML a pro řízenou konverzi obrazových souborů i metadat. Pro tvorbu a úpravu XML metadat slouží XML editor Corel XMetaL Author. Pro potřeby Národní knihovny ČR bylo provedeno přizpůsobení XML editoru pro DTD Monografie a Periodika [Národní knihovna, 2004-09-23].

Pro realizaci systému Kramerius bylo navrženo použití těchto otevřených standardů [Národní knihovna, 2006-05-11]:

- o J2SE (Java 2 Standard Edition) poskytuje základní prostředí pro běh aplikací založených na programovacím jazyce Java
- o J2EE (Java 2 Enterprise Edition) poskytuje služby pro provoz robustních a škálovatelných aplikací založených na jazyce Java
- o Databáze kompatibilní se standardem SQL
- o Operační systém na bázi UNIXu.

V systému Kramerius byly použity následující technologie [Národní knihovna, 2006-05-11]:

- o server IBM xSeries 235 s diskovou kapacitou 3TB
- o Linux Debian
- o WWW server Apache
- o aplikační server Apache Tomcat
- o databázový server PostgreSQL
- o XML editor XMetaL Author.

## 6. Programy a projekty digitalizace v ČR

V této kapitole jsou podrobně rozebrány jednotlivé programy a projekty digitalizace knihovních fondů v České Republice, které byly v rámci provedené rešerše vybrány za účelem provedení jejich rozboru v této diplomové práci.

U každého projektu byly zjišťovány především následující zásadní skutečnosti:

- o Název projektu digitalizace/digitalizovaného fondu
- o Důvody pro digitalizaci fondu
- o Kdo je řešitelem projektu
- o Časové rozpětí projektu digitalizace
- o Instituce, o jejíž sbírky se jedná. Partneři, kteří na projektu spolupracují.
- o Financování (dotace, grant)
- o Kdo digitalizuje (vlastní digitalizační pracoviště nebo specializovaná firma)
- o Obsah fondu: typy dokumentů, celkové množství dokumentů, případně množství již zpracovaných dokumentů
- o Časový záběr sbírek
- o Formou zpřístupnění výsledků digitalizace a jejich lokalizace
- o Technické parametry: software, formát digitálního objektu, formát metadatového popisu digitálního objektu.

Ne vždy a ne u všech projektů se však podařilo získat informace ke všem výše uvedeným bodům. Důvodem bylo u některých projektů nedostatečné prezentování informací pro veřejnost, u jiných projektů nebyly některé skutečnosti ještě stanoveny.

### 6.1. Reformátování bohemikálních časopisů z historického fondu STK

Projekt *Reformátování bohemikálních časopisů z historického fondu Státní technické knihovny ČR* si klade za cíl digitalizaci dokumentů, které jsou ohroženy degradací kyselého papíru a tak zajišťuje jejich uchování pro budoucnost. Financování projektu se uskutečňuje z národního programu VISK 7 - Kramerius, který dotuje 70 % potřebné částky, přičemž 30 % částky knihovna financuje z vlastního rozpočtu.

Projekt digitalizace probíhá nepřetržitě od roku 2003 a plánován je i na roky příští. V současné době (rok 2008) je řešitelkou projektu Petra Večeřová, stejně jako v předchozích letech 2006-2007.

Digitalizovanými periodiky jsou dokumenty z fondu Státní technické knihovny v Praze, noviny a odborné časopisy. Zatím se jedná o kompletní ročníky osmi titulů časopisů, které budou blíže popsány v následující kapitole.

Digitalizaci provádí specializované firmy. Částečně se na digitalizaci podílí i reprografické středisko STK, kde se skenují některé obsahově útlé dokumenty, avšak pouze v menším množství, neboť středisko slouží primárně především uživatelům knihovny.

Výsledky digitalizace jsou zpřístupňovány v systému Kramerius, a to zároveň na webu STK i Národní knihovny. Projekt Kramerius má vytvořen vlastní metadatový systém, který jsou zúčastněné instituce povinny dodržovat. Jako obrazový formát digitalizovaných dokumentů jsou užívány formáty DjVu, PDF a JPEG.

#### **6.1.1. Průběh projektu Reformátování bohemikálních časopisů z historického fondu STK v jednotlivých letech**

##### **I. etapa 2003**

V rámci smlouvy o spolupráci mezi Státní technickou knihovnou a Národní knihovnou ČR a na základě sdělení MK ČR č. 3464 ze dne 11. 2. 2003 byly reformátovány a uloženy do ochranných obalů dva tituly časopisů: *Technický obzor* a *Hornické a hutnické listy*. Do navazujícího programu VISK – 7 Kramerius byl připraven a přihlášen projekt „Reformátování bohemikálních časopisů z historického fondu STK, II. etapa 2004“ [Státní technická, 2004, s.46].

##### **II. etapa 2004**

V roce 2004 pokračovala účast STK na řešení projektu, jehož řešitelkou byla Anna Patočková. V tomto roce byly kompletně reformátovány tituly *Technické listy* a *Průmyslník* a dále sedm svazků titulu *Chemické listy pro vědu a průmysl*. Reformátování provedly firmy Microna Kuřim a Elsys Engineering Vyškov. Originály reformátovaných dokumentů byly uloženy do ochranných obalů



zhotovovaných NK ČR. Archivaci mikrofilmů a digitálních kopií dokumentů zajišťuje centrálně NK ČR [Státní technická, 2005, s.42].

### **III. etapa 2005**

V tomto roce byla řešitelkou projektu do 31.3. Anna Patočková, po které vedení projektu převzala Štěpánka Žižková. Státní technická knihovna připravila a předala k reformátování dalších patnáct svazku titulu *Chemické listy pro vědu a průmysl* (sv. 8 - 22). Od roku 2003 byly reformátovány již čtyři tituly časopisu plně a pátý částečně, což představuje 72 svazků, tedy 22 459 vytvořených obrázků. Originály dokumentů byly v Centrálním depozitáři NK ČR v Hostivaři opatřeny ochrannými obaly z nekyselé archivní lepenky. Nosiče s reformátovanými dokumenty jsou uchovávány v referátu historického fondu STK [Státní technická, 2006, s. 44].

### **IV. etapa 2006**

Od roku 2006 se stala řešitelkou projektu Petra Večeřová, která jej povede i v letech následujících. Státní technická knihovna připravila a předala k reformátování sedmáct svazků titulu *Chemické listy pro vědu a průmysl* (1929-1945) a dvacet dva svazků titulu *Český mechanik* (1886-1907). Od roku 2003 bylo reformátováno již šest titulů časopisu, což představuje 111 svazků, tj. 33 504 obrázků. Originály dokumentů byly v Centrálním depozitáři NK ČR v Hostivaři opatřeny ochrannými obaly z nekyselé archivní lepenky. Nosiče s reformátovanými dokumenty jsou uchovávány v referátu historického fondu STK [Státní technická, 2007, s. 46].

### **V. etapa 2007**

Na reformátování titulu *Technische Blätter* (1869-1919) byla přidělena dotace ve výši 108 000 Kč neinvestičních prostředků . V průběhu řešení došlo ke změně oproti projektu a navýšení původně plánových nákladů, kdy byla navíc podle požadavků Národní knihovny ČR vytvářena administrativní a technická metadata, která byla hrazena ze spoluúčasti STK. Původně plánovaný rozpočet se také zvýšil oproti předpokladu vzhledem k nestandardnosti předlohy, kdy bylo nutno naskenovat vyšší počet mikrofilmů [Státní technická, 2008, s. 14].

## **6.2. České učební texty z fondu STK**

Ve Státní technické knihovně ČR se v současné době začíná realizovat i další projekt, a to digitalizování českých učebních textů<sup>1</sup> z knihovního fondu STK. Tento projekt je zatím v počátečním stadiu realizace a jeho řešitelem je Petra Pejšová.

V prosinci roku 2007 byl sepsán projektový záměr, který zahrnuje digitalizaci učebních textů z fondu STK z let 1978-2008, což nyní zahrnuje 13 500 titulů, přičemž tituly, které jsou již dostupné zároveň v elektronické podobě, skenovány nebudou. Pro projekt se nyní hledá vhodný dotační program.

Prozatím byla provedena digitalizace deseti titulů vybraných dokumentů. Digitalizaci provedlo Digitalizační centrum knihovny Akademie věd v Jenštejně. Dokumenty v digitální formě budou uživatelům dostupné v systému Kramerius, kam byl již importován vzorek 10 naskenovaných titulů. Jako obrazový formát pro prohlížení dokumentů byl zvolen formát PDF, pro plnotextové vyhledávání byl vybrán formát TXT.

### 6.3. Česká digitální matematická knihovna (DML-CZ)

Projekt *České digitální matematické knihovny* (Czech Digital Mathematics Library, ve zkratce DML-CZ) (<http://www.dml.cz/> nebo [www.dml.cz/](http://www.dml.cz/)) si klade za cíl digitálně uložit a volně zpřístupnit matematickou literaturu publikovanou v průběhu historie v Českých zemích, a to včetně bibliografie a metadat. Ve volně dostupné databázové aplikaci Informačního systému výzkumu a vývoje (IS VaV) (<http://aplikace.isvav.cvut.cz/projectDetail.do?rowId=1ET200190513>) jsou cíle řešení projektu stanoveny takto: „Projekt zahrnuje návrh a realizaci řešení vzájemně provázaných problémových okruhů v pěti oblastech: akvizice materiálů pro digitalizaci a ošetření vlastnických a autorských práv, digitalizace a pořízení metadat, tvorba strukturovaných digitálních dokumentů, vytvoření digitální knihovny a zajištění dlouhodobé archivace, začlenění digitální knihovny do WDML“

Projekt má identifikační číslo 1ET200190513 a je finančně podporován grantovou agenturou Akademie věd ČR v rámci národního výzkumného programu 1E-*Informační společnost*. Projekt DML-CZ byl zahájen 1.1. 2005 a ukončen bude k 1.12. 2009, v současné době tedy jeho řešení stále probíhá. Celkové náklady ze státního rozpočtu po celou dobu trvání projektu jsou stanoveny na 8 105 000 Kč.

Na projektu České digitální matematické knihovny se spolupodílí pět institucí:

---

<sup>1</sup> Dle elektronické korespondence s Petrou Pejšovou ze dne 21.4. 2008

- o Matematický ústav AV ČR Praha (<http://www.math.cas.cz/>) je koordinátorem projektu; řešitelem je Jiří Rákosník.
- o Knihovna AV ČR Praha (<http://www.lib.cas.cz/cs>) provádí vlastní digitalizaci dokumentů v Digitalizačním centru AV ČR v Jenštejně; řešitelem je Martin Lhoták.
- o Ústav výpočetní techniky Masarykovy univerzity v Brně (<http://www.ics.muni.cz/>) zajišťuje integraci digitálních dokumentů a implementaci digitální knihovny; řešitelem je Miroslav Bartošek.
- o Fakulta informatiky Masarykovy univerzity v Brně (<http://www.fi.muni.cz/>) zajišťuje vytváření OCR a začleňování aktuálních digitálních čísel časopisů; řešitelem je Petr Sojka.
- o Matematicko-fyzikální fakulta Univerzity Karlovy v Praze (<http://www.karlin.mff.cuni.cz/>) zajišťuje vytváření metadat; řešitelem je Oldřich Ulrych.

Dokumenty zařazované do DML-CZ jsou rozlišovány na tři různé typy ze tří různých časových období. Následující popis zásadně čerpá z článku M.Bartoška: Česká digitální matematická knihovna [BARTOŠEK, 2008].

1. Tištěné dokumenty: jde o časopisy, monografie a sborníky vydané před rokem 1990, které existují pouze v tištěné podobě. Objem dokumentů tohoto typu je v současnosti největší. Tyto dokumenty jsou skenovány a obrazy stránek jsou dále zpracovávány: aplikuje se OCR rozpoznávání textu, následuje vytvoření úvodního seznamu článků v příslušném vydání časopisu. Poté jsou pro články vytvářena metadata: nejprve se provádí autodetekce deskriptivních metadat z externích databází anebo z OCR, následuje manuální revize článků a jejich obsahů pomocí speciálně vyvinutého Metadatového Editoru. Ten umožňuje: 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, vytvářet nebo opravovat metadatové záznamy, 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 [BARTOŠEK, 2007].

2. „Retro-born-digital“ dokumenty: jsou materiály časového rozpětí od počátku devadesátých let do současnosti. Tyto dokumenty již existují v nějaké digitální formě, není tedy potřeba je skenovat. Tato digitální forma dokumentů však není jednotná a často se v průběhu doby několikrát měnila i v rámci jednoho dokumentu (časopisu), je tudíž nutné konvertovat ji do požadovaného jednotného tvaru. Metadata lze obvykle extrahovat přímo z digitálních dokumentů, pokud jsou přitom zohledňována specifika té které digitální formy.
3. Digitální dokumenty přebírané online: jedná se o přebírání aktuálně vydávaných článků z čísel časopisů. Cílem je vytvořit pro každý časopis mechanismus, aby byla z nově publikovaného čísla časopisu automaticky vygenerována i verze pro digitální knihovnu DML-CZ. Začleňování nově vydávaných časopiseckých čísel do DML-CZ pak bude probíhat automatizovaně, bez nutnosti dodatečné ruční práce.

V současnosti je do DML-CZ zařazeno 10 titulů časopisů české provenience, např. *Czechoslovak Mathematical Journal* (ISSN 0011-4642) a *Applications of Mathematics* (ISSN 0862-7940, vydává Matematický ústav AV ČR), *Archivum Mathematicum* (ISSN 0044-8753, tištěné periodikum, ISSN 1212-5059, online periodikum, Masarykova univerzita), *Commentationes Mathematicae Universitatis Carolinae* (ISSN 0010-2628, tištěné periodikum, ISSN 1213-7243, online periodikum vydává Matematicko fyzikální fakulta Univerzity Karlovy v Praze), sborníky z matematických konferencí (např. konference *International Conference on Differential Equations and their Applications*), monografie (např. *Bolzanova kolekce*), VŠ kvalifikační práce a samozřejmě také články. V červnu 2008 knihovna zahrnovala již kolem 16 tisíc matematických článků (tj. cca 170 000 stran textu) a v tomto rozsahu byla zpřístupněna veřejnosti. Předpokládaný rozsah digitální knihovny v roce 2009, kdy bude projekt dokončen, je minimálně 200 000 stran.

Postup při vytváření DML-CZ zahrnuje několik zásadních kroků prováděných v pořadí, v jakém jsou zde dále popisovány.

Nejprve se vybírají materiály pro digitalizaci, připravují se a je řešena otázka autorsko-právních aspektů. Následuje vlastní proces digitalizace, který zahrnuje skenování dokumentů, následné úpravy digitálních obrazů a pořízení základních stránkových a strukturálních metadat. U naskenovaných dokumentů se provádí dvoukrokové rozpoznávání OCR – nejprve běžné rozpoznání textu (FineReader), na které navazuje specializované rozpoznávání matematiky (InftyReader). Dalším krokem je pořízení tzv. referenčních metadat postupem získání a předzpracování základních článkových metadat z matematických referenčních databází *Mathematical Reviews* a *Zentralblatt MATH*. Proces integrace zahrnuje vytvoření strukturovaných digitálních dokumentů, kontrolu nebo vytvoření a doplnění veškerých popisných metadat a vygenerování článků ve formátu PDF. Posledním krokem je import digitálních objektů (dokumentů a jejich metadat) do digitální knihovny a jejich zpřístupnění uživatelům [BARTOŠEK, 2007].

### **6.3.1. Systém digitální knihovny DML-CZ**

Digitální knihovna DML-CZ je zcela integrována do mezinárodního matematického prostředí. Články v digitální knihovně a odkazy v seznamech referencí jsou provázány na záznamy a recenze v matematických referenčních databázích *MathSciNet (Mathematical Reviews)*, *Zentralblatt-MATH* a *Jahrbuch über die Fortschritte der Mathematik*. Knihovna DML-CZ nabízí anglické uživatelské rozhraní a pro neanglicky psané články jsou poskytována základní metadata v angličtině. Dodržují se mezinárodní standardy pro zápis matematických výrazů, stejně jako technické standardy pro interoperabilitu v rámci připravované světové matematické digitální knihovny WDML (<http://www.wdml.org/>), do které je již knihovna DML-CZ začleněna [BARTOŠEK, 2008].

Digitální knihovna DML-CZ je koncovým uživatelům zpřístupněna v open-source systému DSpace s nadstavbou Manakin. DSpace poskytuje veškeré základní funkce

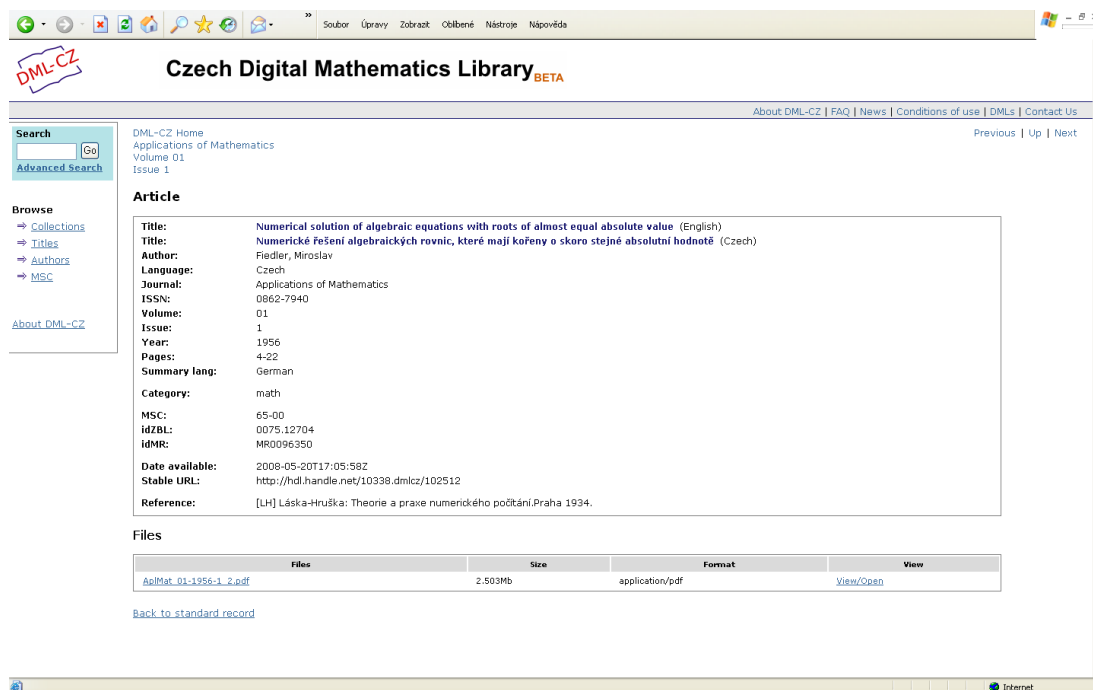
digitální knihovny (prezentace digitálního obsahu, procházení sbírek a hierarchických digitálních objektů, vyhledávání v metadatech i plných textech, trvalé identifikátory, metadata v Dublin Core, podpora protokolu OAI-PMH pro sklizení metadat, statistiky, podpora dlouhodobé archivace digitálního obsahu, aj.). Šablonovací systém Manakin umožňuje přizpůsobit vzhled, strukturu a obsah stránek systému DSpace potřebám knihovny DML-CZ [BARTOŠEK, 2008].

Digitální knihovna DML-CZ nabízí uživatelům jednoduché i pokročilé vyhledávání (v metadatech i v plných textech článků) a procházení jednotlivých kolekcí (časopisů, sborníků, monografií) a rejstříků (názvů, autorů a kódů MSC - Mathematics Subject Classification). Nalezené články jsou zobrazovány ve formátu PDF.

Zobrazení vyhledaného záznamu dokumentu v digitální knihovně je rozčleněno na pět úrovní. Poté, co si uživatel zvolí kliknutím periodikum pro vyhledávání, se zobrazí záznam časopisu ve vlastním citačním stylu digitální knihovny DML-CZ. Záznam obsahuje náhled titulní strany periodika a seznam všech jeho ročníků (volume) a čísel (issue). Po vybrání a kliknutí na konkrétní ročník periodika se zobrazí záznam druhé úrovně, který obsahuje seznam čísel daného ročníku. Na této obrazovce se již nabízí možnost vyhledávání v plném textu.

Poté, kdy si uživatel zvolí konkrétní číslo časopisu, je zobrazena třetí úroveň záznamu s nabídkou konkrétních článků daného časopisu. Po kliknutí na vybraný článek se zobrazí čtvrtá úroveň se záznamem článku v citačním stylu digitální knihovny, opatřený číslem MSC, kódem MR (MathSciNet) a kódem Zbl (Zentralblatt MATH), které jsou propojeny na záznamy v daných databázích. Záznam je také opatřen náhledem na obrázek první strany článku, nechybí článková bibliografie a především přístup k plnému textu článku (full entry).

Po zvolení funkce „full entry“ se zobrazí strukturovaný záznam článku v řádkovém formátu s anglickými návěštími (viz Obr. 1). Článek je opatřen trvalým identifikátorem HANDLE a URL. Na této úrovni, stejně jako již na předchozí, lze zvolit zobrazení plného textu článku ve formátu PDF. Každý článek je opatřen standardizovanou titulní stranou s ikonou systému a upozorněním na elektronické razítko.



Obr. 1. Digitální knihovna DML-CZ se zobrazeným záznamem článku

#### 6.4. Digitální knihovna Akademie věd ČR

Digitální knihovna Akademie věd ČR je projektem Knihovny AV ČR (<http://www.lib.cas.cz/cs>), která spolupracuje s Národní knihovnou ČR (<http://www.nkp.cz/>, je hlavním partnerem při definici standardů a při vývoji systémů pro zpracování a zpřístupnění digitálních dokumentů). Řešitelem projektu je Martin Lhoták. Následující informace byly vytěženy z volně dostupné databázové aplikace Informačního systému výzkumu a vývoje (IS VaV, <http://aplikace.isvav.cvut.cz/projectDetail.do?rowId=1ET200830526>). Projekt má identifikační číslo 1ET200830526 a je finančně podporován grantovou agenturou Akademie věd ČR v rámci národního výzkumného programu 1E-*Informační společnost*. Projekt byl zahájen 1.1. 2005 a je plánován do 1.12. 2009. Celkové finanční náklady ze státního rozpočtu na celou dobu trvání projektu jsou stanoveny na 5 500 000 Kč.

Cílem projektu Digitální knihovny AV ČR je zpřístupnit vědecké publikace, které vznikly na půdě Akademie věd ČR a jejich předchůdkyň a odrážejí tak výsledky jejich výzkumu, co nejširšímu okruhu vědeckých a odborných pracovníků.

Na webovém sídle Digitalizačního centra Knihovny Akademie věd ČR jsou cíle projektu definovány takto: „Cílem je zkoumat, vyvíjet a aplikovat metody, postupy a nástroje pro vytvoření infrastruktury a realizaci Digitální knihovny AV ČR.

Účelem digitální knihovny je zpřístupnit vědecké publikace vydávané Akademií věd ČR a další relevantní odbornou literaturu. Během projektu budou sledovány, zkoumány a aplikovány metody, standardy a nástroje používané v národním i mezinárodním měřítku, které jsou důležité pro zajištění kompatibility a sdílení dat s jinými systémy. Výsledky projektu budou využitelné při realizaci digitálních knihoven a archivů nejen v AV ČR, ale i v dalších institucích [Akademie Věd, 2005].“ Zmíněnou institucí je Národní knihovna ČR, která digitální knihovnu buduje v Centrálním depozitáři v Hostivaři.

Digitalizační práce probíhají v Digitalizačním centru Knihovny AV ČR (<http://www.lib.cas.cz/cs/sluzby/digitalizacni-centrum-knihovny-av-cr>), které zahájilo svoji činnost 1. 1. 2004. Předmětem digitalizace jsou monografie a časopisy z produkce AV ČR, jejich předchůdkyň a následovnic, obsahující výsledky výzkumu AV ČR.

Pro digitální obraz jsou používány obrazové formáty TIFF, JPG, DJVU. Při zpracování metadat, které probíhá v systému Sirius, je dodržována metadatová struktura definovaná Národní knihovnou ČR. Systém pro zpřístupnění digitálních materiálů a pro práci s nimi, který vzniká v rámci projektu, je vyvíjen pod licencí GNU General Public License. Bude tedy s minimálními náklady využitelný v kterémkoli z ústavů AV ČR i jiných institucích, což bude u některých publikací nutné z hlediska dodržování vydavatelských a autorských práv ke zpřístupňovaným materiálům [Akademie Věd, 2005].

Zdigitalizované dokumenty jsou již zpřístupněny v Digitální knihovně Akademie věd ČR, která je již volně dostupná online v systému Kramerius (<http://kramerius.lib.cas.cz/kramerius/Welcome.do>). Ke dni 3.11. 2008 bylo v Digitální knihovně zpřístupněno již 848 461 stran dokumentů, což představuje 58 titulů.

### **6.5. Digitalizovaný archiv časopisů**

*Digitalizovaný archiv časopisů* je projekt Ústavu pro českou literaturu AV ČR (<http://www.ucl.cas.cz/>), vzniká ve spolupráci s Divadelním ústavem (<http://www.divadelni-ustav.cz/>) a Masarykovým ústavem a Archivem AV ČR (<http://www.mua.cas.cz/>).



Tento digitální archiv časopisů je výjimkou v projektech a programech digitalizace, popisovaných v této práci. Nejedná se o projekt v pravém smyslu slova<sup>2</sup>. Tvorba digitálního archivu není financována z grantů ani dotačních programů, nemá určeno časové vymezení a nebylo pro ni vytvořeno zvláštní pracoviště.

Digitální archiv je vytvářen a zpracováván od roku 2001 a práce pokračují kontinuálně stále. Vytváření digitálního archivu je jednou složkou úkolů Střediska literárněvědných informací. Jediné finanční navýšení poskytuje Ústav pro českou literaturu ze svého rozpočtu na externí firmu, která se spolupodílí na fotografování některých velkých formátů periodik. Jinak se periodika skenují v rámci Střediska literárněvědných informací. V době vzniku myšlenky Ústavu pro českou literaturu AV ČR vytvořit digitální archiv časopisů ještě v České Republice nebyl vytvořen Národní program VISK 7 Kramerius. V pozdějších letech, kdy se Ústav pro českou literaturu AV ČR zajímal o možnosti dotace z tohoto programu, znamenala by účast v programu pro ústav neúnosnou finanční zátěž.

Tvorbu digitálního archivu částečně spolufinancoval Divadelní ústav, který financoval digitalizaci divadelních titulů. V současné době Masarykův ústav a Archiv AV ČR spolufinancuje digitalizaci periodika Čas. Pro digitalizaci jsou vybrána periodika 19. a 20. století z fondu Ústavu pro českou literaturu AV ČR. Důvodem k postupnému převodu dokumentů do elektronické podoby byla rychle postupující devastace časopiseckých svazků a zvyšující se požadavky na jejich využívání. Kritériem pro převedení jednotlivých titulů do digitální podoby je jejich fyzický stav a v souvislosti s ním frekvence jejich výpůjček.

---

<sup>2</sup> Dle elektronické korespondence s Věrou Buriánkovou ze dne 20. 11. 2008

Autor	Název článku	Časopis	Rubrika
	Kniha na okně	Literární noviny, ročník 13/1940, číslo 7-8, strana 169	(Verše)
	Večery s knihami	Literární noviny, ročník 13/1940, číslo 7-8, strana 172	(Verše)
	Stará kniha	Literární noviny, ročník 13/1940, číslo 7-8, strana 154	(Verše)
	Kniha o Praze (J.Kolářik)	Literární noviny, ročník 9/1940, číslo 6, strana 4	(Kniha (Literatura b/ o nových knihách))
	Rakos, Korvina čili Kniha o havranech	Literární noviny, ročník 4/1993, číslo 51-52, strana 27	(Kontroverze)(Recenzované knihy - výběr z noviny a č.)
(SL)	Osmašedesátý po třiceti letech (kniha 1968 - Magnus ve světě)	Literární noviny, ročník 9/1998, číslo 14, strana 2	(Fotografie)
-zj-	Na počátku všeho byla kniha	Literární noviny, ročník 9/1998, číslo 22, strana 2	(Publicistika)
Augustová Zuzana	Kniha o Evaldu Schormovi (kniha Radka Denemarková: Evald Schorm. Sám sobě nepřítel)	Literární noviny, ročník 10/1999, číslo 8, strana 13	(Film)
Bechyňová, Venceslava	Kniha svědectví a její přijetí	Literární noviny, ročník 16/1967, číslo 9, strana 4	(Literatura)
Begley Louis	Má "kniha století" (pretištěno)	Literární noviny, ročník 10/1999, číslo 16, strana 11	(Ankety)
Bendová Helena	Soukromé dějiny grotesky (kniha Petr Král: Groteska čili Morálka slehačkového dorth)	Literární noviny, ročník 10/1999, číslo 15, strana 13	(Film)
Blažek, Vladimír	Kniha stížnosti	Literární noviny, ročník 14/1965, číslo 38, strana 12	(Fejetony)
Boxbergerová, Jana	Česká kniha ve francii	Literární noviny, ročník 4/1993, číslo 22, strana 13	(Články o literatuře a umění)
Brdečka, Jiří	Kniha o AH	Literární noviny, ročník 16/1967, číslo 15, strana 4	(Výtvarnictví)
Brezovský, Boluslav	Malá kniha o Velké Číně (Adolf Hofmeister)	Literární noviny, ročník 4/1955, číslo 5, strana 6	(Studie a články, recenze)

Obr. 2. Digitalizovaný archiv časopisů se zobrazeným seznamem vyhledaných článků obsahujících klíčové slovo „ kniha“ v názvu

Archiv časopisů je volně online dostupný (viz Obr. 2, <http://archiv.ucl.cas.cz/>). Zpřístupněno uživatelům je čtyřicet šest titulů časopisů. V systému archivu lze vyhledávat prohlížením, [podle jména autora nebo názvu článku](#), [podle obvyklých bibliografických údajů](#) a [podle data](#) vydání.

## 6.6. Digitalizace historických dokumentů knihovny Vojenského historického ústavu

Knihovna Vojenského historického ústavu (<http://www.vhu.cz/cs/stranka/knihovna/>) v současnosti realizuje tři projekty digitalizace dokumentů<sup>3</sup> ze svého fondu, a to:

- o Historie a vojenství v dobovém tisku
- o Historie a vojenství v dobových písemnostech a
- o Digitalizace mapového fondu

Digitalizaci dokumentů pro prvé dva zmíněné projekty provádí specializované soukromé firmy, které zvítězily ve výběrovém řízení.

<sup>3</sup> Dle osobní konzultace s Zdeňkem Munzarem ze dne 24.9. 2008

## Historie a vojenství v dobovém tisku

Tento projekt je uskutečňován v rámci programu *Kramerius*, dotace tedy plyne z *Národního programu VISK 7*. Projekt digitalizace Historie a vojenství v dobovém tisku probíhá od roku 2006 a je plánován i na roky příští. Řešitelem projektu je Zdeněk Munzar. Důvodem pro digitalizaci je ohrožení dokumentů degradací kyselého papíru a jejich zachování po budoucnost. Digitalizovanými dokumenty jsou periodika z období 18. až počátku 20. století. Periodika k digitalizaci jsou vybírána podle svého stáří, rarity a možnosti využití.

V letech 2006-2008 bylo zdigitalizováno cca 34 tisíc stran textu, přičemž v roce 2006, kdy byl projekt zahájen, bylo vytvořeno 6 400 stran, o rok později 9 600 stran a za rok 2008 je sečteno 18 000 stran digitálního textu. Jelikož je zřejmé, že strany přibývají geometrickou řadou, předpokládané množství zdigitalizovaných stran pro rok 2009 je dvojnásobek množství z roku 2008.

V letech 2006-2008 byly vytvořeny digitální kopie těchto periodik: *Družina československých legionářů* - kompletní periodikum vydávané v letech 1920-1933. Časopis *Čechoslovák*, celé periodikum vydávané v letech 1939-1945. kompletní týdeník *Žižka*, vydávaný v letech 1870-71, který byl poté přejmenován na *Vlast*, vydávanou jako týdeník v letech 1871-1873. *Vojenské listy*, ročník 1868. *Österreichische militärische Zeitschrift*, ročníky 1808, 1811-1813 a 1818-1820. Celé periodikum *Zpravodaj československé jednotky na středním východě*, vycházelo v letech 1940-1943. *Československý voj*, ročníky 1939-1940. Časopis *Táborita*, ročník 1917. *Český kapesní vojenský kalendář*, ročníky 1899-1900. *Naše noviny*, ročníky 1940-1945. Časopis *Bratrství*, ročníky 1919-1927. Periodikum *Československá samostatnost*, ročníky 1906-1919 a časopis *Hlídka zajatců*, ročník 1917.

Digitalizaci těchto periodik provedla firma Ampaco ČR, která sídlí v Dobré Vodě u Českých Budějovic. Dodržovány jsou metadatové standardy a obrazové formáty určené projektem *Kramerius*. Výsledné dokumenty v digitální formě prozatím nejsou zpřístupněny uživatelům. Knihovna VHÚ plánuje do konce roku 2008 zprovoznit v rámci svého webového sídla systém *Kramerius*, ve kterém budou dokumenty uživatelům k dispozici.

## **Historie a vojenství v dobových písemnostech**

Tento projekt se účastní národního programu *Memoriae Mundi Series Bohemica*, pobírá tedy dotaci z *Národního programu VISK 6*. Ke zdigitalizování byl vybrán fond vzácných a ohrožených rukopisů starých tisků a monografií, s datací vzniku od počátku 17. století do poloviny 18. století.

Projekt Historie a vojenství v dobových písemnostech probíhá od roku 2007 a je plánován i na roky příští. Řešitelem projektu byl do srpna 2008 Michal Hokynek, nyní je řešitelem Zdeněk Munzar. V letech 2007-2008 bylo zdigitalizováno šestnáct dokumentů a vytvořeno cca 4 000 stran digitálního textu. Dokumenty k digitalizaci jsou vybírány podle svého stáří a rarity.

V roce 2007 provedla digitalizaci firma Ampaco ČR, v roce 2008 firma Albertina icome Praha. Dodržovány jsou metadatové standardy a obrazové formáty určené projektem *Manuscriptorium*. Výsledné digitální dokumenty jsou zařazeny do digitální knihovny *Manuscriptorium*, kde jsou zpřístupněny uživatelům.

## **Digitalizace mapového fondu**

Knihovna VHÚ má ve svém fondu také početnou sbírku kartografických materiálů pocházejících z rozmezí 17. až 19. století. Zastoupeny jsou rukopisné i tištěné mapy a plány, sbírka obsahuje také atlasy a grafiky.

Digitalizace tohoto fondu probíhá na digitalizačním pracovišti Vojenského ústředního archivu, které je vybaveno nejmodernější technikou, včetně kolíbkových knižních scannerů. Řešitelem projektu je Zdeněk Munzar. V současné době je projekt Digitalizace mapového fondu ve fázi snímání obrazů, výběr metadatového formátu a následné ukládání digitálních dokumentů není ještě zcela dořešeno. V současnosti je zdigitalizováno cca 300 kartografických dokumentů [HOKYNEK, 2007] .

Výsledné digitální objekty budou zpřístupněny v plánované digitální knihovně v systému *Kramerius* v rámci webového sídla Knihovny VHÚ. Knihovna zatím zvažuje, zda bude digitální knihovna *Kramerius* přístupná pouze z čtenářských stanic internetové studovny, nebo taktéž volně v síti Internet.

## 6.7. Digitalizace a zpřístupnění církevních listin v rámci evropského projektu Monasterium

Národní archiv ČR v Praze (<http://www.nacr.cz/>) se v roce 2005 zapojil do evropského projektu *Monasterium* [Národní archiv, 2008], který koordinuje Institut pro výzkum a zpřístupnění církevních pramenů (Institut zur Erforschung und Erschließung kirchlicher Quellen <http://www.dsp.at/dasp/>) se sídlem v St. Pölten v Rakousku. Národní archiv ČR se současně stal koordinačním členem konsorcia *Monasteria*, což je volné pracovní společenství archivů, knihoven, vědeckých a dalších podobných institucí spravujících kulturní dědictví.

Hlavním cílem projektu *Monasterium* je volně online zpřístupnit (<http://www.monasterium.net>) nejširší veřejnosti co nejvíce listin i jiných typů dokumentů (např. rukopisů a aktového materiálu) především ze středoevropských klášterů a biskupství. Na stránkách webového sídla projektu [Monasterium, 2002] jsou jeho cíle definovány takto:

- o „Vybudování virtuálního archivu listin jakékoliv provenience ze středověku a raného novověku.“
- o „Využití digitálního archivu pro bádání, výuku a širokou veřejnost zajímavící se o tuto problematiku.“
- o „Digitální propojení různých druhů historických pramenů.“

Během trvání projektu bylo zrušeno původní vymezení na listiny pouze církevní provenience a projekt byl rozšířen na listiny všech proveniencí. Projekt *Monasterium* je financován rakouským spolkovým ministerstvem pro vzdělání, vědu a kulturu, Evropskou unií (program *Culture 2000*), dále různými kláštery a dodatečně získávanými finančními prostředky ze státních i soukromých zdrojů zúčastněných zemí. Projektu *Monasterium* se účastní kromě České republiky a Rakouska také Bavorsko, Slovinsko, Maďarsko, Chorvatsko, Slovensko, Itálie a Švýcarsko.

První fáze digitalizace listin v I. oddělení Národního archivu proběhla v období od 4. září do 15. listopadu 2006. Digitalizovány byly listiny datované do roku 1526, které jsou součástí mnoha důležitých archivních fondů církevního původu. Do projektu bylo zahrnuto také několik fondů necírkevních, jejichž sbírky se ale církevní problematiky z větší části týkají. Listiny byly digitalizovány v rozlišení 400 dpi.

V tomto rozlišení zůstávají digitalizované dokumenty uloženy v diecézním archivu v St. Pölten v Rakousku, k prezentaci na webové stránky se vkládají obrázky již zmenšené.

Celkem bylo zdigitalizováno 7 118 listin z těchto 13 církevních a 5 necírkevních archivních fondů:

- o Archiv pražského arcibiskupství
- o Archiv kolegiální kapituly vyšehradské, Praha
- o Archiv kolegiální kapituly Všech svatých na Pražském hradě
- o Archivy českých klášterů zrušených za Josefa II.
- o Augustiniáni obutí – provincialát a konvent, Praha
- o Benediktini - klášter Břevnov, Praha
- o Benediktini - klášter Emauzy, Praha
- o Dominikáni – provincialát a konvent, Praha
- o Františkáni – provincialát a konvent, Praha
- o Kapucíni – provincialát a konvent, Praha
- o Křižovníci s červenou hvězdou – generalát a konvent, Praha
- o Maltézští rytíři – české velkopřevorství, Praha
- o Premonstráti – klášter Strahov, Praha
- o Česká finanční prokuratura Praha – listiny
- o České gubernium – guberniální listiny, Praha
- o Úřad desk zemských, Praha
- o Česká dvorská kancelář, Praha, Vídeň
- o České oddělení dvorské komory, Vídeň

Do roku 2011 je plánováno rozšíření projektu i na další archivy České republiky. Vybrány jsou listinné fondy státních oblastních archivů: Moravského zemského archivu v Brně, Zemského archivu v Opavě, Státních oblastních archivů v Třeboni, Plzni, Litoměřicích, Praze a Zámrsku. Osloven byl i Archiv Pražského hradu

a Archiv Národního muzea. V Moravském zemském archivu se digitalizovalo již v létě roku 2007, kdy bylo naskenováno 2 260 listin z 29 církevních fondů.

V období od 13. září do 11. října 2007 proběhla druhá fáze digitalizace listin v I. oddělení Národního archivu. Zdigitalizováno bylo 2 525 listin Archivu České koruny. V dalších letech Národní archiv plánuje dokončení digitalizace listinných částí fondů I. oddělení, ze kterých jsou již naskenovány listiny datované do roku 1526, a počítá i s listinami z fondů II. a VI. oddělení Národního archivu.

### 6.6.1. Popis webové stránky MOM Virtuální archiv listin střední Evropy

Webová stránka (<http://www.monasterium.net>) umožňuje uživatelům neomezený přístup ke všem listinným historickým pramenům z období od raného středověku až po novověk. Vytvoření webové stránky je založeno na spolupráci mnoha státních, zemských, diecézních a klášterních archivů. Archiv Monasterium (v titulu stránky je název archivu uveden jako: MOM Virtuální archiv listin střední Evropy) je proto ve své formě celosvětově jedinečný.

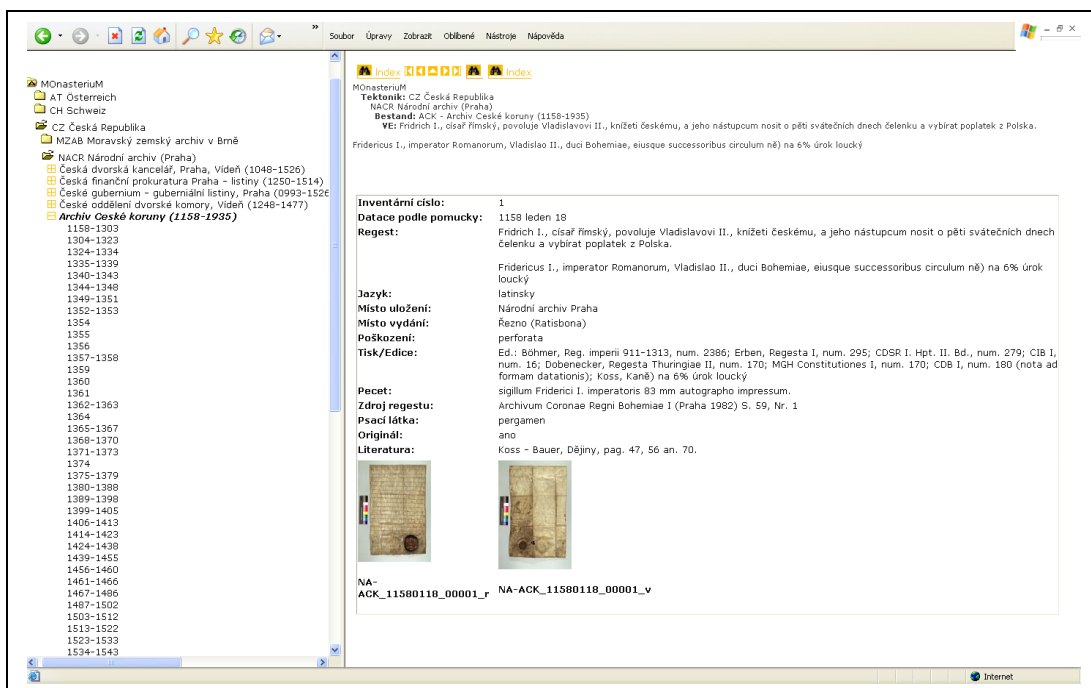


Obr. 3. Úvodní stránka archivu MOM

V úvodu stránky (viz Obr. 3) je zobrazen aktuální počet zpřístupněných listin, který čítal ke dni 1.10. 2008 103 456 titulů. Na webovém sídle archivu MOM je možné informovat se o projektu a jeho cílech, aktualitách, členech konsorcia, financování aj.

Hlavní funkcí webu je však zpřístupnění listin v digitální knihovně veřejnosti. V pravém menu je možné vybrat si z osmi nabízených jazykových mutací stránky - jedná se o jazyky, kterými se hovoří v zúčastněných zemích, samozřejmě je tedy i čeština.

Po kliknutí na odkaz Rešerše se otevře v novém okně úvodní stránka systému digitálního archivu. V menu po levé straně je možné vybrat dokument ze seznamu zemí podle jeho provenience. Po kliknutí na např. Českou Republiku se dále rozvine nabídka jednotlivých institucí a pod nimi seznam dokumentů z jejich provenience. Po zvolení konkrétního dokumentu se na pravé straně obrazovky zobrazí tzv. Předslav (anglicky Preface), který obnáší stručnější vědecký popis a komentář. Pod tímto popisem jsou vypsaná inventární čísla dokumentů, ze kterých je možno opět vybrat kliknutím.



Obr. 4. Zobrazení dokumentu v archivu MOM

Po tomto výběru se zobrazí Popis záznamové jednotky, který obsahuje stručný výtah z obsahu dokumentu společně s datací. Po kliknutí na zvolený Popis záznamové jednotky se zobrazí strukturovaný záznam té konkrétní listiny/dokumentu (viz Obr. 4), který uživateli přináší následující informace: inventární číslo, dataci dokumentu, regist a zdroj registu, jazyk dokumentu, místo jeho uložení, místo vydání dokumentu, jeho rozměry, psací materiál dokumentu, zda se jedná o originál a zda je dokument označen notářským znamením. Pod tímto strukturovaným



záznamem jsou náhledy obrázků dokumentu (podle počtu jeho stran), které se po kliknutí zobrazí v novém okně, kde je lze detailně prohlížet.

Systém prohlížení v digitální knihovně má tedy čtyři úrovně a je velmi intuitivní.

#### **6.8. HISPRA - Záchrana historických pragensijních a dalších vzácných dokumentů z fondu Městské knihovny**

*HISPRA* je projektem Městské knihovny v Praze (dále jen MKP) a jeho hlavním cílem je záchrana a alternativní zpřístupnění vzácných a historických dokumentů z fondu Městské knihovny v Praze. Na webovém sídle MKP a projektu *HISPRA* (<http://www.mlp.cz/hispra.htm>) jsou definovány cíle celkem čtyři [MĚŘÍNSKÁ, 2008]:

- o vybudovat vlastní digitalizační pracoviště
- o pořídit digitální kopie vytipovaných dokumentů
- o vytvořit online digitální knihovnu
- o rozšířit povědomí o těchto dokumentech a jejich využití.

Projekt *HISPRA* je financován prostřednictvím Finančních mechanismů Evropského hospodářského prostoru & Norských finančních mechanismů, které projekt podpořily 403 552 eury. Tato částka činí 80 % z celkových nákladů 507 740 euro na celou dobu trvání projektu. 20 % z této celkové částky hradí MKP sama. Projekt má číslo CZ0028, schválen byl 6. listopadu 2006 a je rozvržen na období duben 2007 až červen 2010 [Česko. 2005b]. V roce 2007 přitom probíhal pouze zkušební provoz. Vedoucí projektu a také vedoucí Oddělení digitalizace je Eva Měřínská. Projekt převzala po Zuzaně Kopencové, která vypracovala projekt ke schválení a přidělení grantu<sup>4</sup>.

K digitalizaci byl vybrán specializovaný fond pragensjí a dále fond divadelní, hudební, uměnovědné a další literatury se vztahem k Praze, se zaměřením na období konce 19. a počátku 20 století. Tyto monografie jsou důsledkem stáří, častého používání a degradace kyselého papíru, jenž je charakteristický pro dokumenty 19. století, ve špatném fyzickém stavu. Během projektu má být zdigitalizováno 350 000 stran těchto dokumentů z celkového počtu 20 000 svazků, které jsou svým obsahem unikátní. K digitalizaci v rámci projektu jsou navíc vybírány ty dokumenty,

---

<sup>4</sup> Dle osobní konzultace s Evou Měřínskou ze dne 12.11. 2008

kteří již nejsou zatíženy autorským právem [MĚŘÍNSKÁ, 2008]. Konkrétní hlediska pro výběr každého dokumentu ke zdigitalizování jsou fyzický stav dokumentu, zda je ohrožen degradací kyselého papíru, jeho čtenářská žádanost, jedinečnost dokumentu v rámci MKP a České republiky.

Městská knihovna v Praze již v minulých letech uskutečnila digitalizaci několika monografií a periodik ze svého fondu (cca 2 tituly za rok) v rámci Národních programů VISK 6 a VISK 7. V těchto případech byla digitalizace provedena externě formou služby od komerčních firem. Unikátní fond pragensií a dalších historických dokumentů MKP však nespĺňuje požadavky pro udělení dotace v rámci VISK 6 ani VISK 7. Proto se MKP rozhodla v roce 2005 zpracovat projekt *HISPR* do programu Finančních mechanismů Evropského hospodářského prostoru a Norska a vybudovat vlastní digitalizační pracoviště.

Tak vzniklo Oddělení digitalizace MKP, které je organizačně začleněno pod Odbor knihovních fondů. Toto oddělení bylo otevřeno 31. ledna 2008 a má sedm pracovních úvazků pro jedenáct zaměstnanců. Na podzim 2007 bylo pracoviště vybaveno serverem o velikosti 10 TB, a dvěma knižními skenery. Data jsou uchovávána ještě na dalších dvou odlišně lokalizovaných serverech (10 TB) se systémem zálohování, přičemž jeden slouží výhradně jako datové úložiště. Každý měsíc jsou také vytvářeny z těchto serverů kopie na externí pevné disky [MĚŘÍNSKÁ, 2008].

Zpracování neskenovaných obrázků probíhá v programu Sirius od firmy Elsys Engineering, vytvořená data jsou zpřístupněna přes open source systém Kramerius. Při digitalizaci jsou tedy dodržovány standardy Krameria. Dokumenty jsou skenovány v rozlišení 600 dpi, obvykle ve škálách šedi. Obrázky jsou uloženy pouze ve formátu JPG, který je používán jako archivní a rovněž jako výstupní. Následuje úprava (detekování okrajů a ořez obrázků, lineární vyrovnávání textu a rozdělení na jednotlivé strany) a další zpracování (tvorba metadat) v programu Sirius. Bibliografický popis je importován z katalogu Koniáš, jednotlivé strany jsou indexovány ručně. Nakonec se provádí OCR rozpoznání textu a vygenerování XML souboru, opatřeného administrativními metadaty [MĚŘÍNSKÁ, 2008].

Digitální dokumenty budou zpřístupněny v únoru 2009 v digitální knihovně na webovém sídle MKP. Volně dostupné přes síť internet budou dokumenty, které již

nejsou pod ochrannou Autorského zákona, ostatní dokumenty budou uživatelům přístupné pouze v síti MKP<sup>5</sup>.

Ke dni 13.11. 2008 bylo zdigitalizováno již 883 dokumentů, převážně monografií. Jejich úplný seznam je prozatím volně online dostupný na webovém sídle MKP (<http://www.mlp.cz/digifond.htm>). V tomto seznamu je již možné vyhledávat a také zobrazit tzv. katalogový lístek, kde je uveden katalogizační záznam dokumentu, anotace a klíčová slova.

Po ukončení projektu *HISPR*A bude na digitalizačním pracovišti nadále probíhat digitalizace vzácných a historických dokumentů z fondu Městské knihovny v Praze, neboť cílem je zdigitalizovat a uživatelům zpřístupnit všech 20 000 svazků z ohroženého fondu. V budoucnu bude pracoviště využíváno také pro digitalizaci komerční v rámci služeb nabízených jiným institucím. Během trvání projektu to není povoleno.

Oddělení digitalizace MKP také spolupracuje s Národní knihovnou ČR a Knihovnou AV ČR, a to především na bázi koordinace (předcházení vytváření duplicit digitálních dokumentů) a školení. MKP v současné době digitalizuje z 99 % pouze monografie, proto registr CZROMM (Český registr reformátovaných dokumentů) pro kontrolu duplicit využívá zřídka. Při výběru dokumentů k digitalizaci se MKP řídí seznamem dokumentů digitalizovaných Národní knihovnou ČR v rámci Norských fondů a ústní dohodou v rámci koordinační skupiny při NK ČR ohledně kontroly duplicit dokumentů, ve které bylo stanoveno, že MKP se při digitalizaci zaměří na Pragensia a NK ČR na další tituly.

## **6.9. Společná česko-slovenská digitální parlamentní knihovna**

*Společná česko-slovenská digitální parlamentní knihovna* je společný projekt českého a slovenského parlamentu. Projekt spoluvytváří: [Parlamentní knihovna](http://www.psp.cz/kps/knih/) (<http://www.psp.cz/kps/knih/>), [odbor informatiky](http://www.psp.cz/kps/inf/) (<http://www.psp.cz/kps/inf/>) [Kanceláře Poslanecké sněmovny](http://www.senat.cz/) a [odbor informačních technologií Kanceláře Senátu](http://www.senat.cz/) (<http://www.senat.cz/>) za stranu českou a [Parlamentní knihovna, parlamentní archiv a odbor informatiky Kanceláře Národní rady Slovenské republiky](http://www.nrsr.sk/default.aspx?sid=nrsr/kancelaria/orgstruc) (<http://www.nrsr.sk/default.aspx?sid=nrsr/kancelaria/orgstruc>) za stranu slovenskou. Tento projekt vznikl již v roce 1996 pod názvem Digitální knihovna Český

---

<sup>5</sup> Dle osobní konzultace s Mgr. Evou Měřínskou ze dne 12.11. 2008

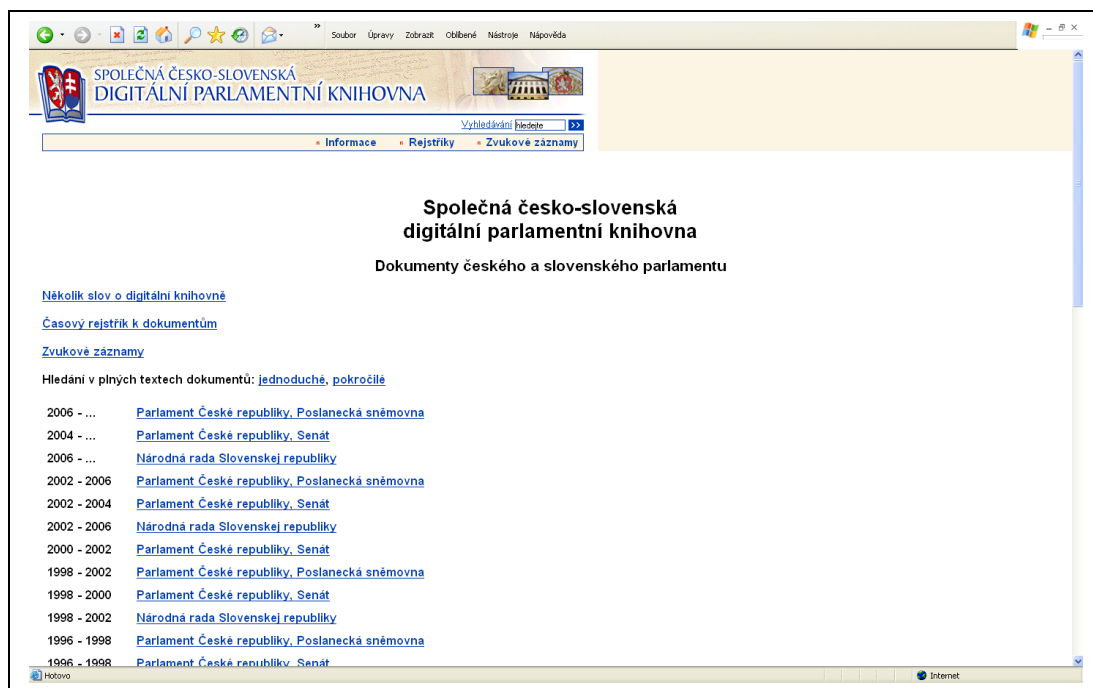
parlament, financován byl z prostředků [Kanceláře Poslanecké sněmovny](#). V roce 2002 se proměnil ve *Společnou česko-slovenskou digitální parlamentní knihovnu*, neboť se při digitalizaci starších parlamentárií ukázalo, že dokumenty připomínající společné dějiny obou národů je potřebné zkoumat jako jeden celek [TICHÝ, 2006]. V projektu se pokračuje i nadále, jeho řešitelem je ředitel Parlamentní knihovny Karel Sosna.

Projekt si klade za svůj cíl postupnou digitalizaci všech těsnopiseckých zpráv a parlamentních tisků od roku 1918 do současnosti, a výběrově i starších, a jejich zpřístupnění veřejnosti prostřednictvím Internetu.

Cíle projektu jsou následující [Parlamentní knihovna, 2003]:

- o Vytvořit informační systém o činnosti: zákonodárných orgánů České republiky a Slovenské republiky od jejich vzniku v roce 1993, společných zákonodárných orgánů Československé republiky v letech 1918 až 1992, a českých a slovenských poslanců v různých zákonodárných orgánech (Slovenská národní rada, Říšský sněm a Sněm království Českého) od roku 1848 do roku 1914.
- o Zpřístupnit tyto informace nejširší veřejnosti v objektivní podobě.
- o Umožnit rychlejší, jednodušší a efektivnější přístup k informacím z parlamentu.

Společná česko-slovenská digitální parlamentní knihovna (viz Obr. 5) je volně online přístupná (<http://www.psp.cz/eknih/>) a lze do ní vstoupit přes webové sídlo Parlamentní knihovny. Digitální knihovna obsahuje úplné texty parlamentních tisků (dokumenty, které jsou projednávány na schůzích Parlamentu: návrhy zákonů včetně důvodových zpráv a usnesení, mezinárodní smlouvy, státní rozpočty, interpelace), těsnopiseckých zpráv (stenoprotokoly-doslovné záznamy ze všech jednání a hlasování) a dalších parlamentárií (pozvánky na schůze: pozvánky na schůze Parlamentu s návrhy pořadu schůze), usnesení Parlamentu, usnesení Organizačního výboru Poslanecké sněmovny, usnesení Předsednictva České národní rady, dokumenty výborů: pozvánky na schůze výborů, zápisy z jednání výborů, usnesení výborů, dokumenty podvýborů a komisí [Parlamentní knihovna, 2003].



Obr. 5. Titulní obrazovka Společné česko-slovenské digitální parlamentní knihovny

Podle jednotlivých volebních období digitální knihovna obsahuje parlamentní texty těchto proveniencí [Parlamentní knihovna, 2003]:

- o Rakouský ústavodárný říšský sněm 1848 - 1849 (Vídeň, Kroměříž)
- o Sněm království Českého 1861 - 1913
- o Národní shromáždění ČSR a ČSSR (Poslanecká sněmovna a Senát) 1918 - 1968
- o Sněm Slovenské republiky 1939 - 1945
- o Slovenská národní rada 1944 - 1960
- o Česká národní rada 1969 - 1992
- o Usnesení předsednictva Slovenské národní rady 1970 - 1987
- o Federální shromáždění ČSSR a ČSFR (Sněmovna lidu a Sněmovna národů) 1969 - 1992
- o Parlament České republiky (Poslanecká sněmovna) 1993 - 2003
- o Národní rada Slovenské republiky 1993 – 2003.

Při výběru dokumentů k digitalizaci se postupuje tzv. „hnízdovým způsobem“. To znamená, že jsou vybrány klíčové momenty v českých dějinách a digitalizovány

materiály, které se jich týkají. Důvodem je předpokládaný větší zájem odborné i širší veřejnosti o tyto dokumenty. Vzniklé mezery mezi těmito historickými milníky jsou postupně podle finančních možností zaplňovány [TICHÝ, 2006].

Digitalizaci dokumentů provádí externě soukromá firma (dříve firmy dvě), která vyhrála výběrové řízení. Jako formát pro ukládání i zpřístupňování digitalizovaných textů se používá HTML. U tisků pocházejících z novějších volebních období se užívají formáty DOC, PDF a XLS. Pro stenoprotokoly, které nejsou starší než z ledna roku 2006, se používá zvukový formát MP3 [HOUDEK, 2007].

V digitální knihovně je možné vyhledávat několika způsoby. Jednoduchou možností je procházení rejstříků, časového, jmenného a věcného. Dále je možné využít hledání v plných textech dokumentů, jednoduché nebo podrobné, kdy však jediným kritériem pro omezení vyhledávání je časové období a zadání hledaného výrazu. Možností je samozřejmě i hledání procházením úplných textů jednotlivých období.

#### 6.10. České sněmy

Projekt *České sněmy* spoluvytváří [Parlamentní knihovna](http://www.psp.cz/kps/knih/) (<http://www.psp.cz/kps/knih/>), [odbor informatiky](http://www.psp.cz/kps/inf/) (<http://www.psp.cz/kps/inf/>) [Kanceláře Poslanecké sněmovny](http://www.senat.cz/) a [odbor informačních technologií Kanceláře Senátu](http://www.senat.cz/) (<http://www.senat.cz/>). *České sněmy* jsou druhým projektem digitalizace Parlamentní knihovny a navazují chronologicky sestupně na projekt *Společná česko-slovenská digitální parlamentní knihovna*. Bezprostředním impulsem pro vznik projektu bylo vydání publikace Zdeňka Veselého, CSc. *Dějiny českého státu v dokumentech* [VESELÝ, 2003], ve které autor navrhl základní pojetí projektu a historickou periodizaci. Projekt *České sněmy* byl pak zahájen v roce 2005 a jeho řešitelem je stejně jako u projektu *Společné česko-slovenské digitální parlamentní knihovny* ředitel Parlamentní knihovny Karel Sosna [Parlamentní knihovna].

Cílem projektu je zpřístupnit na internetu všechny dostupné pramenné materiály vztahující se k dějinám českého sněmovnictví. Nezbytná bude spolupráce s dalšími knihovnami a archivy v České Republice, které mají ve svých fondech uloženy tyto prameny. Dokumenty budou, obdobně jako v projektu *Společné česko-slovenské digitální parlamentní knihovny*, zpracovávány a zpřístupňovány tzv. "hnízdovým způsobem", a to podle své dostupnosti, stupně zpracovatelnosti a možnosti jejich využití pro digitalizaci současnými technologickými prostředky.

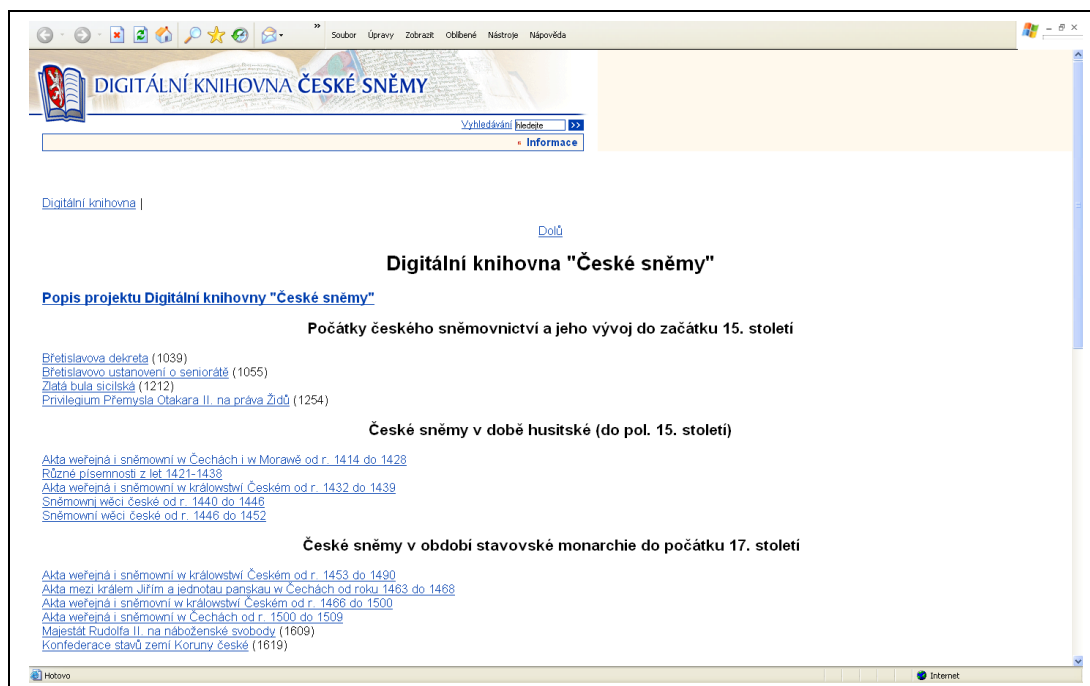
Předmětem digitalizace jsou základní dokumenty sněmovní, ústavní a státoprávní povahy z těchto historických období [Parlamentní knihovna, 2003]:

- o počátky českého sněmovnictví a jeho vývoj do začátku 15. století,
- o české sněmy v době husitské do poloviny 15. století,
- o české sněmy v období stavovské monarchie do počátku 17. století,
- o české sněmy v období habsburského absolutismu do roku 1848.

Digitální knihovna *České Sněmy* (viz Obr. 6.) je volně online dostupná (<http://www.psp.cz/eknih/snemy>) a lze do ní vstoupit přes webové sídlo Parlamentní knihovny.

Typologie dokumentů v digitální knihovně je velmi rozmanitá. Ve starších obdobích se zapisovaly jen výsledky jednání, usnesení, artikule, které se však všechny nedochovaly. V raném období existence českého státu rozhodnutí sněmů nahrazovala vyhlášená rozhodnutí knížat na shromážděních šlechty (např. *Břetislavova dekreta* z roku 1039). I v pozdějším období byla vydávána panovnická rozhodnutí s povahou zákonů, které podle dnešních zvyklostí přijímá parlament, ale sněmy je tehdy neschvalovaly, protože byly ve výlučné pravomoci panovníka (např. *Dekret Kutnohorský* z roku 1409). Některé dokumenty také původně navrhl sněm, ale ve známost vstoupily až výnosem od panovníka (např. *Inaugurační diplomy Jana Lucemburského* z roku 1310 a 1311) [Parlamentní knihovna, 2003].

Texty jsou v digitální knihovně zveřejňovány v češtině, němčině či latině. Ukládány a zpřístupněny jsou dokumenty v HTML nebo XML formátu a také jako digitální obrázky. Parlamentní knihovna také vstoupila do programu Manuscriptorium, jehož prostřednictvím zpřístupňuje uživatelům vzácné historické dokumenty v jejich původní podobě [TICHÝ, 2006].



Obr. 6. Úvodní obrazovka digitální knihovny České sněmy

V digitální knihovně České sněmy, narozdíl od Společné česko-slovenské digitální parlamentní knihovny, nejsou nabízeny rejstříky, které je možné využít pro vyhledávání. Vyhledávat je tedy možné prohlížením jednotlivých časových období, anebo lze využít, stejně jako u Digitální parlamentní knihovny, vyhledávání v plných textech, jednoduché nebo podrobné.

### 6.11. Digitální knihovna historických fondů Vědecké knihovny v Olomouci

Projekt *Digitální knihovny historických fondů* realizuje Vědecká knihovna v Olomouci (<http://www.vkol.cz/cs/>). Financování projektu je uskutečňováno prostřednictvím *Národních programů VISK a RISK* Ministerstva kultury ČR. Dále byl projekt spolufinancován Ústavem pro klasická studia AV ČR (<http://www.clavmon.cz/>) v rámci projektu *Clavis monumentorum litterarum* a dvěma sponzorskými firmami [Vědecká knihovna, 2005a]. Řešitelem projektu je Rostislav Krušínský.

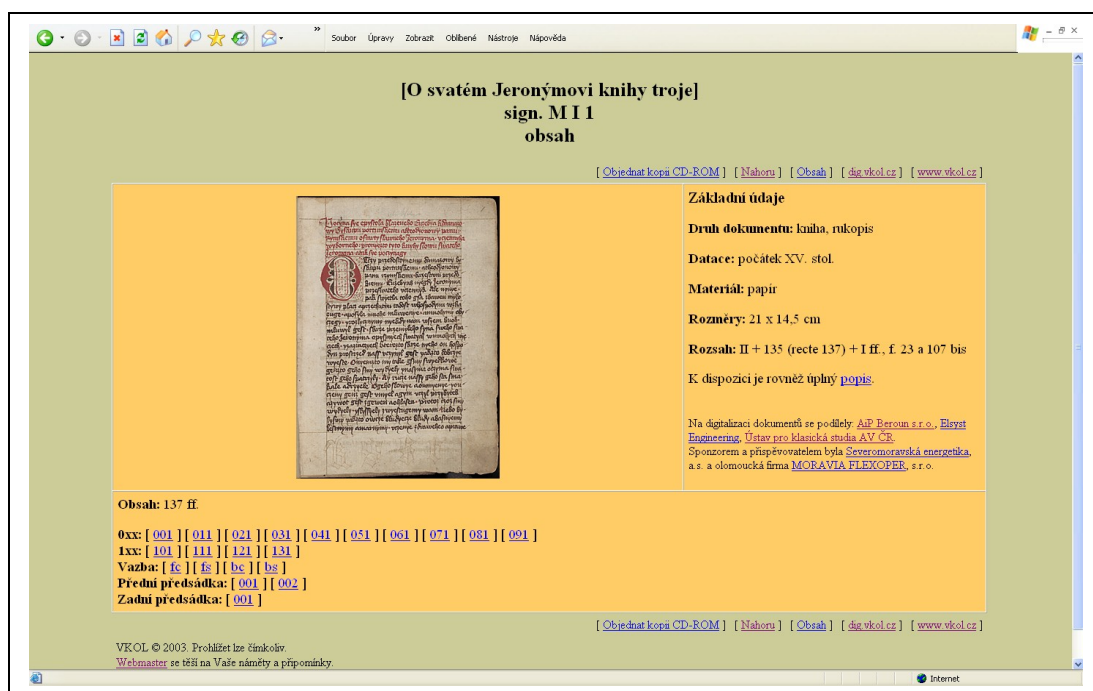
Předmětem digitalizace jsou vzácné rukopisy (cca 1 450), prvotisky (cca 1 900) a staré tisky (80 000) uložené ve fondu Vědecké knihovny v Olomouci [VOJNAR, 2002], které pocházejí z časového rozpětí 11. až 17. století. Důvodem pro digitalizaci fondu je potřeba ochrany historických fondů, vzácných nebo zajímavých map



a dalších ucelených sbírek a jejich zpřístupnění uživatelům [Vědecká knihovna, 2005a].

Digitalizace historických dokumentů začala v roce 1996 převodem vzácného *Evangeliaře zábrdovického* sign. M II 74, do digitální podoby pod záštitou Ústavu pro klasická studia AV ČR v rámci projektu *Clavis monumentorum litterarum*. Digitalizační práce pak pokračovaly až od roku 1999 v rámci grantových programů Ministerstva kultury ČR *RISK* a *VISK*. Digitalizace probíhá průběžně i v současnosti a je plánována i v letech příštích. Provádí ji firma AiP Beroun (<http://www.aipberoun.cz/>) v rámci programu *VISK 6 Memoriae Mundi Series Bohemica* [VOJNAR, 2002].

Jako formát digitálního obrazu se používá TIFF, JPEG a DjVu. Formátem metadatového popisu digitálního objektu je MARC 21 pro dokumenty typu staré tisky, kresby a mapy (s možností výstupu MARCXML)<sup>6</sup>.



Obr. 7. Digitální knihovna historických fondů Vědecké knihovny v Olomouci se zobrazeným rukopisem „O svatém Jeronýmovi knihy troje“

Zdigitalizované dokumenty jsou uživatelům zpřístupněny v digitální knihovně, která je volně online přístupná (<http://dig.vkol.cz/>), lze do ní vstoupit proklikem z webového sídla Vědecké knihovny v Olomouci. Ke dni 18. 11. 2008 bylo

<sup>6</sup> Dle elektronické korespondence s Jiřím Spáčillem ze dne 27. 11. 2008

v digitální knihovně vystaveno 83 rukopisů a prvotisků (viz Obr. 7.), 7 starých tisků a 3 mapy.

## 6.12. Digitální knihovna map Vědecké knihovny v Olomouci

Vědecká knihovna v Olomouci (VK) je také řešitelem projektu Digitální knihovny map. Tento projekt byl realizován v rámci programu Ministerstva kultury ČR [VISK 5 - Národní program retrospektivní konverze katalogů knihoven v ČR – RETROKON](#) [VKOL, 2005b]. Digitalizace začala roku 1999 a probíhá průběžně i v současnosti.

Digitalizovanými dokumenty jsou výhradně mapy uchovávané ve fondu Vědecké knihovny v Olomouci, které pocházejí z časového rozpětí 17. - 19. století [Vědecká knihovna, 2005b]. Formátem digitálního obrazu je, stejně jako u předchozího projektu VK v Olomouci TIFF, JPEG a DjVu. Rovněž formát metadatového popisu digitálního objektu je shodný s předchozím projektem (MARC 21 s možností výstupu MARCXML).

Signatura	Název	Vydáno	Obrazy
55.742	Mapy staré Prahy k letům 1200, 1345 a 1419 -- 1 mapa v 9 listech / stožil W. Wladislaw Tomek	Praha - Nakladem české akademie císaře Františka Josefa pro vědy, slovesnost a umění, 1892	
III 86.014	Visecí mapa Markrabství moravského a Vévodství slezského -- od B. Kozema ; nové přepracované vydání vzdelané od C. F. Baura a C. K. Prof. J. Havelky	Ve Vídni a Olomouci : Skladem Eduarda Hölzela, [1880-1883?]	
V 12.179	Ethnographische Karte der oesterreichischen Monarchie / -- entworfen von Karl von Czernig. Herausgegeben von der k. k. Direction der Administrativen Statistik	Wien : Ans der kaiserlich-koeniglichen Hof- und Staatsdruckerei, 1855	
V 13.314	Bemerkungen über den innern Kreislauf der Handlung in den oesterreichischen Erbstaaten : zur nöthigen Erläuterung der hydrographischen General- und Partikularkarten von diesen Ländern oder Hauptentwurf der zu eröffnenden schiffbaren Wasserstraßen von allen Meeren Europas bis nach Wien -- Von J. F. Maire hydraulisch- und hydrographischen Ingenieur	Strassburg und Leipzig : [s.n.], 1786	
V 13.315	Hydrographische Karte der oesterreichischen Erbstaaten desselbs des Rheins = Carte Hydrographique des Etats le Mison d' Autriche en deca du Rhin -- von F. J. Maire hydr. und Geogr. Ingenieur ; Grave par Mansfeld	A Vienne : aux frais de Mrs. Chr. et Fréd. Bargum et Comp. Banquiers, 1786	

Obr. 8. Abecední seznam map řazený dle formátu v Digitální knihovně map Vědecké knihovny v Olomouci

Digitální knihovna map Vědecké knihovny v Olomouci je volně dostupná online (<http://mapy.vkol.cz/>). Ke dni 18. 11. 2008 bylo v digitální knihovně vystaveno 556 mapových listů (viz Obr. 8). Digitalizaci všech mapových listů provedla firma [Elsyst](#)

[Engineering](#). Prohlížení mapových listů v digitální knihovně je zprostředkováno softwarem [Zoomify](#).

### **6.13. Projekty digitalizace Moravské zemské knihovny v Brně**

Moravská zemská knihovna v Brně (<http://www.mzk.cz/>) je na poli digitalizace dokumentů značně aktivní. V současné době je knihovna řešitelem jednoho výzkumného záměru a několika projektů, které nějakým způsobem souvisejí s digitalizací. Tato diplomová práce se podrobněji zabývá výzkumným záměrem *Historické fondy Moravské zemské knihovny v Brně*, který splňuje kritéria zadání této práce.

#### **6.13.1. Historické fondy Moravské zemské knihovny v Brně**

Historické fondy Moravské zemské knihovny v Brně jsou výzkumným záměrem číslo MK00009494301, jehož trvání je rozvrženo na období od 1. 3. 2004 do 31. 12. 2010. Hlavním řešitelem záměru je Petr Žabička. Výzkumný záměr je finančně podporován Ministerstvem kultury ČR, celkové finanční náklady ze celou dobu řešení projektu jsou vyčísleny na 40 579 000 Kč, přičemž celkové náklady ze státního rozpočtu činí 10 992 000 Kč. Cíle výzkumného záměru jsou v databázové aplikaci Informační systém výzkumu a vývoje IS VaV (<http://aplikace.isvav.cvut.cz/organizationDetail.do?rowId=ico%3A00094943>) definovány takto: „Hlavním cílem výzkumného záměru je získat a formou bibliografické databáze zveřejnit přesné a úplné informace o všech historických sbírkách MZK. Dokumenty z historických fondů MZK budou zpracovány po stránce bibliografické a jejich nejvýznamnější části (titulní listy, frontispisy, incipity apod.) digitalizovány. Digitalizace významných částí zpracovávaných dokumentů si vyžádá nasazení vhodného systému digitální knihovny. Jeho nalezení a implementace je jedním z cílů záměru. V oblasti digitalizace bude cílem výzkumného záměru nalezení takových technik, přístrojů a pracovních postupů, které by umožňovaly levnou a rychlou digitalizaci speciálních druhů dokumentů.“

Z uvedených cílů je zřejmé, že samotná digitalizace a následné zpřístupnění digitálních objektů v digitální knihovně je pouze součástí výzkumného záměru, a to částí podstatnou pro tuto diplomovou práci. Proto se tato práce bude dále zabývat pouze problematikou digitalizace.

Historické fondy MZK, které jsou předmětem výzkumného záměru, jsou značně obsáhlé. Čítají na 800 rukopisů, z nichž k nejvzácnějším patří převážně středověké bohemikální rukopisy mikulovské dietrichsteinské knihovny, 384 svazků prvotisků, 37 000 svazků starých tisků z kmenového fondu MZK a dvou menších samostatně vyčleněných celků, zednářské literatury a děl Jana Amose Komenského. K vzácným fondům se řadí i 25 000 svazků z historické sbírky (zámecké, klášterní i světské knihovny, které MZK získala v druhé polovině 20. století). Rozsáhlý je také fond starých map a grafík, zejména unikátní mapová sbírka B. P. Molla, kterou knihovna získala počátkem 19. století. Významný je též speciální fond hudebních rukopisů a starých tisků, čítající 6 000 svazků, který je obsahově zaměřen na skladby českých autorů 18. století [ŽABIČKA, 2004b, s. 78].“

Moravská zemská knihovna (dále jen MZK) zřídila vlastní digitalizační pracoviště, vybavené planetárním systémem Atlas a digitálním fotoaparátem Canon EOS20D od firmy Elsynt Engineering. Vzniklo tak kvalitní pracoviště pro operativní a šetrnou digitalizaci historických dokumentů, které je využíváno pro rutinní digitalizaci starých tisků, nikoli však vzácných či velkoformátových historických dokumentů. Digitalizace probíhá rychlostí 50-200 stran za hodinu v závislosti na kvalitě předlohy a požadavcích na následné úpravy [ŽABIČKA, 2005, s. 7]. Dále je digitalizační pracoviště vybaveno deskovými scannery HP 4670, Mustek ScanExpress A3 a Plustek OpticBook 3600 a digitálním kopírovacím a tiskovým strojem Minolta Bizhub C350 [ŽABIČKA, 2007, s. 3].

V oblasti digitalizace a digitálních knihoven byl výzkum zaměřen především na oblast digitalizace a zpřístupnění velkých grafických dokumentů s využitím technologií geografických informačních systémů. V rámci něj byl zprovozněn server [timemap.mzk.cz](http://timemap.mzk.cz). [ŽABIČKA, 2005, s. 11] a později server [www.staremapy.cz](http://www.staremapy.cz). Protože kvalitní zpřístupnění starých map vyžaduje mnoho specifických nástrojů, jejichž vývoj by přesahoval rámec výše zmíněného výzkumného záměru, podala MZK návrh projektu výzkumu a vývoje (VaV, Staré mapy online), který by řešil specifika kartografických dokumentů. Tento projekt se však nezabývá vlastním procesem digitalizace, ale vyvíjením nástrojů pro zpřístupnění digitalizované mapové sbírky uživatelům, proto jeho popis v této diplomové práci nebude věnován prostor.

MZK v rámci výzkumného projektu Historické fondy navázala spolupráci s několika institucemi. S Památníkem písemnictví na Moravě spolupracuje formou konzultací

z oblasti digitálních knihoven a informačních technologií, přes poskytnutí vybavení pro digitalizaci historického rukopisného lístkového katalogu knihovny a jeho umožnění zpřístupnit na Internetu. Historickému ústavu AV ČR byla poskytnuta konzultace týkající se digitalizace a zpřístupnění sbírky starých map z fondu ústavu. MZK také spolupráce s Národní knihovnou ČR v rámci Manuscriptoria, kam importuje záznamy o historických fondech MZK ve formátu MASTER. Tak byla v Manuscriptoriu zpřístupněna Mollova mapová sbírka, zdigitalizovaná již roku 2003, čítající 676 digitálních obrázků.

## 7. Problematika registrace digitalizovaných jednotek

Jedním z největších problémů v digitalizaci dokumentů jak v národním, tak i mezinárodním měřítku je vytváření duplicit digitalizovaných dokumentů. Duplicity vznikají kvůli absenci jednotné databáze, nebo rejstříku všech dokumentů již zdigitalizovaných, právě digitalizovaných a připravovaných k digitalizaci. V důsledku této situace knihovny, které digitalizují, musí dnes samy sledovat v nejrůznějších zdrojích, co již bylo zdigitalizováno, což je úkol velmi obtížný.

### 7.1.1. CZROMM Český registr reformátovaných dokumentů

Za účelem vyloučení duplicit je v Národní knihovně České Republiky, v rámci Odboru správy a ochrany fondů vytvářen a v současnosti již zveřejněn rejstřík CZROMM, který je volně dostupný online v rámci webu Národní knihovny ČR ([http://www.nkp.cz/pages/page.php3?page=weba\\_seznam.htm](http://www.nkp.cz/pages/page.php3?page=weba_seznam.htm)). Pod názvem CZROMM se skrývá *Český registr reformátovaných dokumentů*, tedy prozatím seznam novin a časopisů, které byly reformátovány v rámci *Národního programu VISK 7 Kramerius*.

V současnosti se jedná o pasivní rejstřík, který je možné pouze prohlížet. Prohlížení lze provádět podle abecedního seznamu názvů dokumentů, který je zobrazen na úvodní stránce rejstříku CZROMM. Vyhledávání jiným způsobem není doposud možné. Po kliknutí na vybrané písmeno abecedy se zobrazí seznam všech zaznamenaných dokumentů, jejichž název začíná vybraným písmenem. Tento seznam byl vytvořen v excelovém formátu, zpřístupněn je ve formátu PDF a ve formě tabulky (viz Obr. 9).

V této tabulce jsou zaznamenávány následující informace<sup>7</sup>:

- o název časopisu/novin: u některých titulů je v této kolonce uveden také podnázev a místo vydání
- o rok: jedná se o roky vydání periodik, která jsou již zdigitalizovaná
- o vlastník práv k reformátované kopii: každá instituce, která nechá reformátovat své fondy v Národní knihovně v rámci programu Kramerius, dává práva k reformátované kopii i NK ČR

---

<sup>7</sup> Vysvětlivky k některým položkám byly vytvořeny na základě konzultací s Janem Hutařem ze dnů 15. 9. a 27. 11. 2008

- o archivní negativ - počet mikrofilmovaných. polí: archivní negativ, nebo-li archivní mikrofilm, je negativ, který vznikl mikrofilmováním. Tento negativ slouží výhradně pro vytváření matriční kopie negativu, přičemž je povoleno jej kopírovat maximálně čtyřikrát. Z matričního negativu jsou pak vytvářeny uživatelské kopie. V tabulce je uvedeno, zda archivní negativ existuje (uvádí se „ANO“), či neexistuje (kolonka zůstává nevyplněna). Může zde být také uvedeno „ohlášeno“, což znamená, že je výroba archivního negativu plánována. Počet mikrofilmovaných polí u většiny dokumentů uveden není, v budoucnu by však měl být doplněn.
- o matriční negativ: z matričního negativu jsou vytvářeny uživatelské kopie. V tabulce je uvedeno, zda archivní negativ existuje (uvádí se „ANO“), či neexistuje (v tom případě kolonka zůstává nevyplněna).
- o mikrofiše (počet listů): tato položka uživatelům sděluje, zda je daný dokument reformátován také technologií mikrofišování, a pokud ano, kolik listů mikrofiše bylo vytvořeno. V ČR tato technologie není příliš užívaná. Pokud mikrofiš daného dokumentu existuje, je v kolonce číslem uveden počet jeho listů.
- o pozitivní kopie pro čtenáře: v této kolonce je uvedeno, zda byl vytvořen pozitiv pro čtenáře (uvádí se „ANO“), či vytvořen nebyl (kolonka zůstává nevyplněna).
- o digitální kopie: tato položka uvádí, zda je k dispozici také digitální kopie daného dokumentu. Pokud ano, je v kolonce uvedeno „ANO“, pokud ne, položka zůstává nevyplněna. Může zde být uveden také letopočet nebo časové rozmezí, které sděluje, které ročníky daného periodika jsou také v digitálním formátu. Pokud je tvorba digitálního obrazu plánována, nebo již vytvářena, je v kolonce uvedeno „ohlášeno“.

*Český registr reformátovaných dokumentů CZROMM* je prvním počinem v České Republice na poli registrace digitalizovaných dokumentů. Knihovnám, které se účastní Národního programu Kramerius, poslouží jistě jako vodítko k vyvarování se vytváření duplicit digitalizovaných dokumentů. Bohužel registr *CZROMM* má také ještě nedostatky.

První a hlavní nevýhodou je nemožnost interaktivního vyhledávání v rejstříku podle názvu periodika, roku vydání, vydavatele, vlastníka práv k reformátované kopii apod. Dalším nedostatkem je jistá nejasnost termínů uváděných v rejstříku a jejich význam. Odborné termíny bude informovaný uživatel obeznámený s problematikou reformátování a digitalizace pravděpodobně znát, laický uživatel bude kvůli absenci jakýchkoliv vysvětlivek nucen obrátit se na odborníka.

Z dosavadního vývoje však vyplývá, že Český registr reformátovaných dokumentů CZROMM je registr prozatímní a zmíněné nedostatky budou vyřešeny vytvořením a zprovozněním databázové aplikace RD.CZ, která je popsána v následující kapitole.

Název časopisu, novin	Rok	Vlastník práv k reformát. kopii	Archivní negativ-počet mř. polí	Matriční negativ	Mikrofiše (počet listů)	Pozitivní kopie pro čtenáře	Digitál. kopie
Levá fronta. Praha	1930 - 1932	NK ČR	ano				
Libussa. Eine vaterländische Vierteljahrschrift. Prag	1802, 1804	NK ČR	ano	ano			ano
Lidová demokracie. Orgán Československé strany lidové. Praha	V./1945 - VII./1994	NK ČR	ano				
Lidové noviny. Nezávislý deník založený 1893. Brno	1893 - 1952	MZK	ano	ano		ano	1893 - 1901, 1907 - 1914 ohlášeno 1915 - 1952
Lidové noviny. [Vydání Čechy]. Založeny 1893 - obnoveny 1988. Praha	1990 -	NK ČR	ano	ano			
Lipa slovanská. Časopis politicko-literární a obchodnický pro Slovaný na demokratickém základě. Praha	1848-1849	NK ČR MZK	ano	ano			ano
Listy. Časopis československé socialistické opozice. Řím	1989	NK ČR	ano				
Listy českého spolku pro ochranu ptactva v Praze. Praha	1881 - 1882	NK ČR			8		
Listy průmyslové. Praha	1896	NK ČR			4		
Listy chemické	1877 - 1906	STK	ohlášeno				
Listy. Týdeník Svazu českých spisovatelů. Praha	1968, 1969	NK ČR	ano				
Listy zahradnické. Troja	1877 - 1998	NK ČR			96		
Literární listy. Týdeník Svazu českých spisovatelů. Praha	1968	NK ČR	ano				ohlášeno
Listy veřejné správy. Jaroměř	1886 - 1893	NK ČR			441		

Obr. 9. Český registr reformátovaných dokumentů CZROMM, ukázka seznamu záznamů názvů periodik, začínajících písmenem L

### 7.1.2. Databázová aplikace RD.CZ

V Národní knihovně ČR ve spolupráci s Knihovnou AV ČR a firmou INCAD je současně vyvíjena databázová aplikace<sup>8</sup> s „pracovním“ názvem RD.CZ (Registr Digitalizace, pilotní databáze je lokalizována na serveru firmy INCAD [https://sluzby.incad.cz/projekty/registr\\_digitalizace\\_cz/doku.php](https://sluzby.incad.cz/projekty/registr_digitalizace_cz/doku.php)), koncipovaná jako

<sup>8</sup> Dle osobní konzultace s Jiřím Polišenským ze dne 6. 5. 2008



open source systém pro činnosti související s digitalizací, která bude obsahovat registr digitalizovaných dokumentů a ještě mnohem více.

Tato aplikace umožní [KOCOUREK, 2007]:

- o evidovat proces zpracování a provázat dílčí digitalizační práce
- o poskytovat průběžný přehled o zpracování a jeho průběhu
- o poskytovat celkový přehled o stavu a realizaci digitalizačních projektů

Evidovány budou tyto údaje:

- o digitalizované periodikum a digitalizovaná monografie (název, autor, rok vydání, počet stran, počet polí u mikrofilmů a mikrofišů, výška v cm, způsob digitalizace, financování)
- o zakázky (číslo zakázky, rozsah zakázky, URL)
- o zpracovatelé zakázky
- o přidělované úkoly
- o XML soubory
- o mikrofilmy

Aplikace se má stát nástrojem pro sledování digitalizace a činností s ní souvisejících, poskytne celkový přehled o stavu a realizaci digitalizačních projektů [KOCOUREK, 2007].

Aplikace *RD.CZ* je vyvíjena na aplikační platformě Relief III, která na základě XML definice datové struktury a formulářů poskytuje základní datové a obslužné operace, generuje uživatelské rozhraní a poskytuje základní tiskové výstupy. Data jsou do aplikace *RD.CZ* přebírána z několika zdrojů. Jako zdroje dat tzv. historických, tedy dat o digitalizaci již uskutečněné, jsou využívány: systém Kramerius, knihovní katalog Aleph NKC (později také knihovní katalogy každé z digitalizujících knihoven) a soupisy digitalizovaných dokumentů ve formátu XLS, které jsou vytvářeny v jednotlivých odděleních knihoven [PAVLOVÁ, 2008]. Takovýmto zdrojem je také registr *CZROMM*. Systém Relief dokáže přejímat základní bibliografický popis dokumentu z knihovního katalogu Aleph a zároveň přebírá popis dokumentu přímo ze systému Kramerius přes protokol OAI-PMH. Tato data

pak umístí do katalogu, čímž je provázán bibliografický záznam s digitální knihovnou<sup>9</sup>. Systém bude také propojen s digitální knihovnou, ve které bude daný digitální objekt uložen a zpřístupněn.

Přístupová práva k systému budou mít všechny knihovny, které digitalizují dokumenty. Uživatelé budou moci do databáze vkládat data o svých projektech digitalizace a sdílet je s ostatními knihovnami. Pro instituce, které digitalizaci financují z programů VISK bude zadávání dat o digitalizovaných dokumentech do systému povinné. Pro prohlížení bude databáze přístupná také nejširší veřejnosti. Systém nabízí několik uživatelských rolí, přičemž každý uživatel bude moci role kombinovat, aby pokryly všechny činnosti, které uživatel vykonává: corrector (zkušený uživatel, který může zpětně opravovat chyby v datech, role se nedá kombinovat s jinými), zadavatel, příprava mikrofilmů, fotograf, zpracovatel. Roli systemadmin bude zastávat pouze administrátor systému z firmy INCAD a roli appadmin bude zastávat pouze administrátor z Národní knihovny ČR [PAVLOVÁ, 2008].

Databázová aplikace *RD.CZ* by měla být zpřístupněna na webovém sídle Národní knihovny ČR na jaře roku 2009. V současné době již probíhá ověřovací provoz systému v interní síti Národní knihovny.

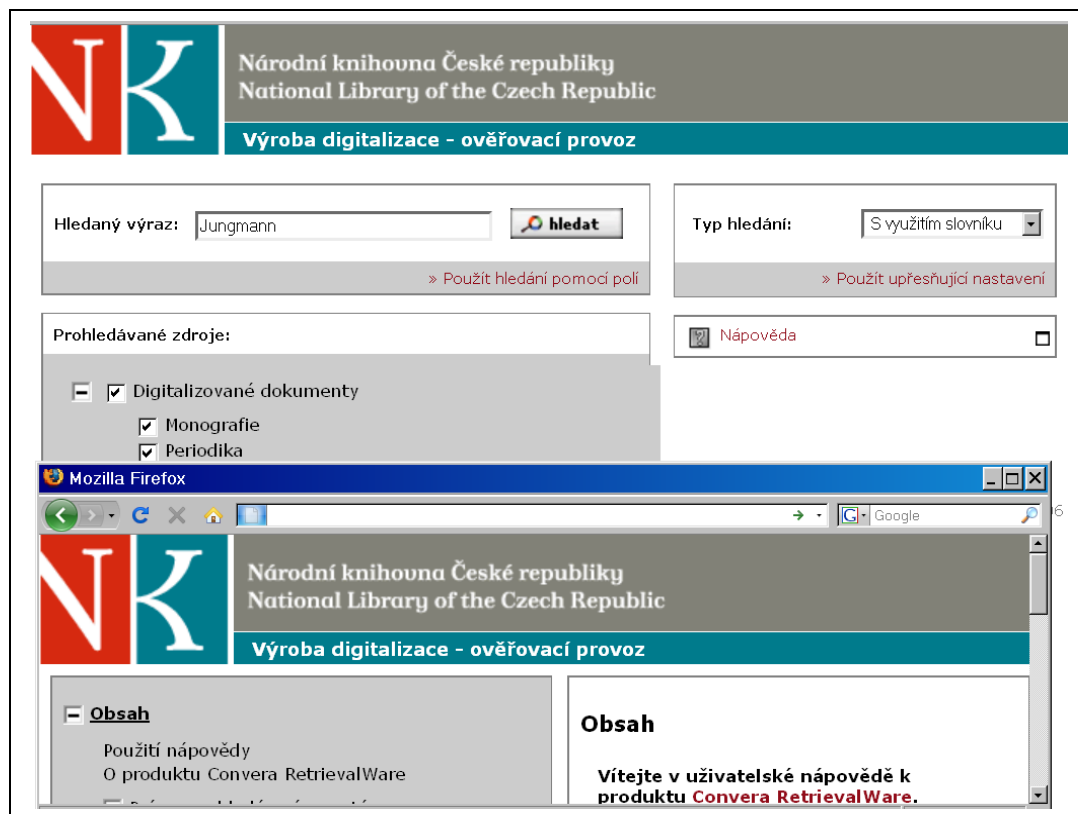
### **Popis vyhledávacího systému Registru Digitalizace RD.CZ**

Vyhledávacím systémem *Registru Digitalizace* je vytvořen na platformě znalostního systému *Convera RetrievalWare*. Výhody tohoto systému jsou tyto: vysoká úroveň relevance a rychlosti vyhledávání dat i z archivů o objemu terabytů, schopnost účinně prohledávat různé zdroje dat (např. elektronické dokumenty, skenované dokumenty), jednoduché rozhraní webovského prohlížeče, vysoká spolehlivost díky technologii „tolerance chyb“ *APRP*. Technologie *Adaptive Pattern Recognition Processing (APRP)* umožňuje efektivně prohledávat i data, která obsahují chyby a překlepy, jako například skenované dokumenty s provedeným optickým rozpoznáváním textu či dokumenty s nízkou kvalitou textu. Systém *Convera RetrievalWare* je dostupný na operačních systémech UNIX, LINUX a na platformě Windows NT .

---

<sup>9</sup> Dle osobní konzultace s Janem Hutařem ze dne 27. 11. 2008

V databázi RD.CZ je možné volit mezi dvěma stupni plnotextového vyhledávání, jednoduchým a pokročilým. Při vyhledávání jednoduchém (viz Obr. 10) je nabízen pouze příkazový řádek pro hledaný výraz a možnost zaškrtnutí zdrojů pro hledání (monografie a/nebo periodika).



Obr. 10. Obrazovka rozhraní pro jednoduché vyhledávání v databázi Registru Digitalizace RD.CZ s částí obrazovky helpu rešeršního systému Convera RetrievalWare, ze dne 4.12.2008

Z jednoduchého vyhledávání se stane pokročilé, pokud si uživatel v rolovacím menu zvolí typ hledání (viz Obr. 11), kde lze vybrat ze způsobů hledání s využitím slovníku, podobných slov nebo booleovské vyhledávání. Vyhledávání se slovníkem umožňuje formulovat dotaz přirozeným jazykem, který bude automaticky rozšířen o příbuzné a související pojmy. Při hledání s pomocí podobných slov budou vyhledány i dokumenty obsahující slova s podobnou stavbou nebo pravopisnými chybami. Počet podobných slov, o která se rozšíří každé slovo z dotazu lze nastavit v roletce "Podobná slova" (viz Obr. 11). Výhodou tohoto způsobu hledání je, že uživatel nemusí znát přesné znění hledaných termínů. Při zvolení booleovského vyhledávání lze využít následující booleovské a proximitní operátory zapsané pomocí znaků jim přidělených: **not ^, and & but, or |, within, between, adj, ( )**.

Dále je možné použít obvyklé zástupné znaky: \* zástupný znak pro cokoli, ? zastoupí 1 alfanumerický znak, [] zastoupí 1 či více znaků, ^ zakáže následující znak, \_ žádný nebo 1 znak, @ 1 alfabetský znak, # 1 numerický znak, ` předcházející znak je volitelný, \ následující znak je doslovný. Při využití tohoto typu vyhledávání je nutné, aby uživatel přesně věděl, jaká slova a v jaké kombinaci se v dokumentu vyskytují.

Další možností při vyhledávání je zvolit hledání pomocí polí. Systém nabídne sedm příkazových polí (viz Obr. 11): název, autor, rok vydání, číslo zakázky, čárový kód, signatura, číslo archivního negativu. Při tomto typu vyhledávání jsou, na rozdíl od plnotextového hledání, slova z dotazu vyhledávána pouze v takto zadané konkrétní oblasti dokumentu, což umožňuje snadno oddělit dokumenty nerelevantní. Při zadávání slov do polí je možné také využívat [booleovské a proximitní operátory](#).

Obr. 11. Obrazovka rozhraní pro pokročilé vyhledávání v databázi RD.CZ s ukázkou možností hledání, ze dne 4. 12. 2008

Po zadání vyhledávacího dotazu systém zobrazí seznam vyhledaných dokumentů (viz Obr. 12). Obrazovka s výsledky hledání obsahuje několik horizontálně zarovnaných bloků s ovládacími prvky, v následujícím pořadí od shora dolů: blok s navigací na úvodní obrazovku, roletkou s volbou počtu dokumentů v jednom listu

a roletkou s volbou způsobu řazení dokumentů, blok s navigací mezi listy s dokumenty, blok obsahující řádky s nalezenými dokumenty, blok pro hromadné označování vyhledaných dokumentů a ovládání [exportní funkce](#). Dále je v levé části každého řádku s nalezeným dokumentem pole, ve kterém je zobrazeno číslo dokumentu v seznamu, políčko pro označení záznamu, ukazatel relevance k dotazu a odpovídající procento relevance. Pokud byl dokument již otevřen, objeví se zde odpovídající ikona, stejně jako v případě, že byl dokument již exportován.



Obr. 12. Obrazovka rozhraní v databázi RD.CZ, ukázka seznamu dokumentů obsahujících termín „Jungmann“, ze dne 4. 12. 2008

Kliknutím na název dokumentu lze zobrazit v novém okně náhled dokumentu (viz Obr. 13). V závislosti na typu dokumentu by měl být otevřen originál dokumentu, pokud je k dispozici, nebo filtrovaný či generovaný dokument. Situaci, kdy by se otevřel originál dokumentu, jsem při analýze systému nezaznamenala, což je zřejmě způsobeno skutečností, že se jedná zatím o prototyp systému a jeho ověřovací provoz.

Obrazovka s náhledem dokumentu je rozdělena do tří horizontálně uspořádaných zón: funkční lišta obsahující ovládací prvky [aktivovaných funkcí](#) nebo [interaktivní tvorby dotazu](#), navigační lišta s číselným seznam odkazů na vyhledané dokumenty

a především blok s náhledem dokumentu (viz Obr. 13 a 14). Tento blok je tvořen tabulkami s bibliografickými údaji o dokumentu s barevně vyznačenými výskyty hledaných slov, informace o způsobu digitalizace a způsobu financování, a podrobnými údaji k digitalizaci (tabulka „Zakázka“).

The screenshot shows a web application interface in Mozilla Firefox. The address bar contains a URL from sluzby.incad.cz. The main content area displays a document record with the following details:

- Název:** **Dějepis** pro školy obecné a měšťanské. Díl třetí, Obrazy a příběhy z **dějín** středního a nového věku se zvláštním zřetelem k **dějínám** rakouským pro osmou třídu rakouských škol obecných a měšťanských / sepsal Antonín Gindely ; přeložil M. Kovář
- Autor(ři):** Gindely, Antonín, 1829-1892
- Rok vydání:** 1880
- Počet stran:** 0
- Počet polí:**
- Výška v cm:** 0
- Způsob digitalizace:** Přímá
- Financováno:** Norské fondy

Below the metadata, there is a section for 'Exempláře' (Copies) with a table:

Signatura:	Čarový kód:	Přívazek:
54 F 001368/D.3.	1001135666	

The 'Zakázka' (Order) section contains a table with the following data:

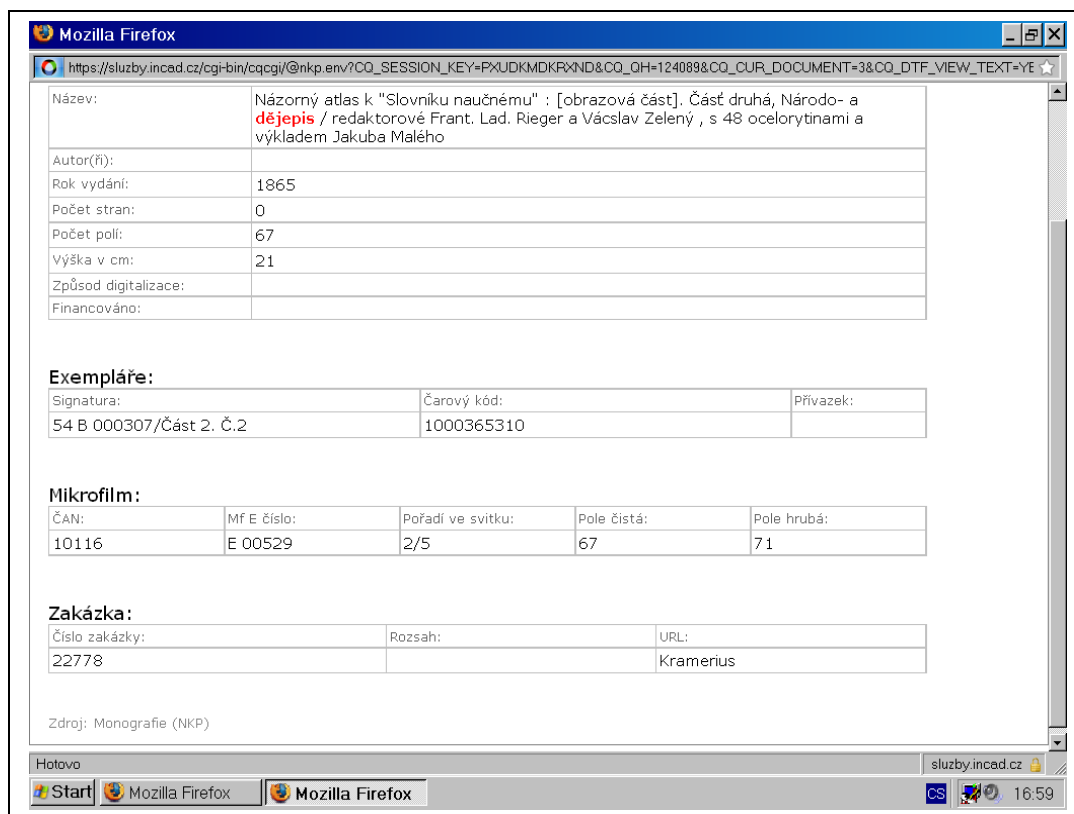
Číslo zakázky:	Rozsah:	URL:
41394		
Příprava dokumentů	Microna	Aktivní
Archivní negativ	Microna	Aktivní
Přímé skenování	Microna	Aktivní
Úprava dat	Elsyst	Aktivní
Konverze obrazových souborů	Elsyst	Aktivní
OCR	Elsyst	Aktivní
Tvorba metadat	Elsyst	Aktivní
Převzetí dat	NK CD Hostivař	Aktivní
Kontrola dat	NK CD Hostivař	Aktivní
Zveřejnění dokumentu	NK CD Hostivař	Aktivní
Archivace	NK CD Hostivař	Aktivní

The browser's taskbar shows the Start button and three instances of Mozilla Firefox. The system tray in the bottom right corner displays the time as 16:22.

Obr. 13. Obrazovka rozhraní v databázi RD.CZ, ukázka náhledu vybraného dokumentu č. 1, ze dne 4. 12. 2008

Při práci se systémem je k dispozici velmi podrobná nápověda, která umožní efektivní práci i méně zkušeným uživatelům. Z této nápovědy bylo čerpáno pro potřeby výše uvedené analýzy systému.

Databáze RD.CZ po svém zveřejnění zřejmě zcela nahradí Český registr reformátovaných dokumentů CZROMM. V budoucnu je navíc plánováno propojení databáze RD.CZ s Evropským registrem reformátovaných dokumentů EROMM (European Register of Microform Masters).



Obr. 14. Obrazovka rozhraní v databázi RD.CZ, ukázka náhledu vybraného dokumentu č. 2, ze dne 4. 12. 2008

## **8. Možnosti dalšího rozvoje a spolupráce**

### **8.1. ENRICH Evropská digitální knihovna rukopisů**

Mezinárodní projekt ENRICH (<http://enrich.manuscriptorium.com/>) byl založen jako součást programu eContentPlus Evropské Unie a jeho trvání je stanoveno na 24 měsíců, od 1. 12. 2007 do 30. 11. 2009. Cílem a smyslem projektu je vytvořit bezproblémový přístup k distribuovaným informacím o rukopisech a vzácných starých tiscích v Evropě na základě digitální knihovny Manuscriptorium (<http://www.manuscriptorium.eu>), spojit dílčí digitální knihovny, umožnit jejich provozovatelům ukládat svá data v Manuscriptoriu ve vlastním jazyce a ve společném rozhraní optimalizovaném dle potřeb uživatelů. Manuscriptorium je nyní se svým obsahem 2 800 zdigitalizovaných rukopisů a 1 milionem stran největší digitální knihovnou rukopisů v Evropě a spolupracuje s ním 50 partnerů z ČR i zahraničí.

Partneři projektu ENRICH mají ve svých sbírkách téměř 85% ze všech existujících digitalizovaných rukopisů evropských národních knihoven. Tyto sbírky budou doplněny o významné množství dat z univerzitních knihoven a dalších institucí. Konsorcium v průběhu projektu zpřístupní více než pět milionů stránek digitalizovaných dokumentů. ENRICH je historicky prvním projektem programu eContentPlus s českým koordinátorem, kterým je Národní knihovna České republiky [ENRICH, 2007].

Důvodů pro spolupráci knihoven na projektu ENRICH je více. Partnerská instituce získá přístup k celému dostupnému fondu digitálních dokumentů v Manuscriptoriu, sdílí prezentační prostředí (odpadá nutnost vytvářet prostředí vlastní), má možnost nahlížet dokumenty v nových souvislostech a možnost přístupu k dokumentům se otevírá více uživatelům.

Po spolupracujících knihovnách je vyžadováno pouze poskytnutí metadat k importaci do Manuscriptoria, kde se transformují do standardu MASTER. Metadata pro centrální databázi budou shromažďována pomocí protokolu OAI (Open Archives Initiative). Obrazové soubory (data) zůstávají ve správě spolupracující instituce, z Manuscriptoria jsou dosažitelná pomocí URL a výsledný digitální dokument je tak složen až v internetovém prohlížeči uživatele. Systém Manuscriptorium akceptuje



většinu metadatových formátů: MARC, UNIMARC a jejich verze, XML i proprietární formáty (přístup MS), a obrazové formáty JPG, PNG, GIF.

Výhodou spolupráce na projektu ENRICH je efektní cenové řešení ve srovnání s centralizací v evropském měřítku a vyšší systémová stabilita a bezpečnost, protože partnerská instituce nijak nepůsobí ve zbylé části systému [PSOHLAVEC, 2008].

Projekt ENRICH řídí konsorcium, které se skládá z 18 partnerů, projekt je však podporován i řadou dalších institucí včetně rozsáhlých digitálních archivů. Konsorcium projektu ENRICH bude úzce spolupracovat s Evropskou knihovnou (The European Library) a stane se součástí Evropské digitální knihovny (EDL) [ENRICH, 2007].

## **8.2. Evropská digitální knihovna a projekt TEL-ME-MOR**

Dne 30. 9. 2005 bylo uveřejněno sdělení Komise Evropskému parlamentu, Radě, Evropskému hospodářskému a sociálnímu výboru a výboru regionů iniciativy nazvané *i2010: Digitální knihovny*. V dokumentu jsou, mimo jiné, stanoveny tři základní pilíře, které mají vést k plné realizaci Evropské digitální knihovny: digitalizace analogových sbírek, archivace a ochrana digitálních zdrojů a jejich zpřístupnění veřejnosti online [i2010: Digitální knihovny, 2005].

### **8.2.1. Evropská knihovna**

Evropská knihovna (The European Library, [www.theeuropeanlibrary.org](http://www.theeuropeanlibrary.org)) byla vytvořena ve spolupráci devíti národních knihoven a Konference ředitelů evropských národních knihoven (CENL – The Conference of European National Librarians) v rámci projektu *TEL (Evropská knihovna: Brána ke znalostem Evropy)*. Projekt TEL započal v roce 2001, byl ukončen 31. 1. 2004 a byl částečně financován v rámci 5. rámcového programu *Evropské komise* [Evropská knihovna, 2008].

Portál Evropské knihovny je v současnosti hybridní digitální knihovnou, která uživatelům nabízí jednotné rozhraní pro prohlížení fondů evropských knihoven. Obsahem Evropské knihovny jsou zdroje digitální, zdroje digitalizované a zároveň i bibliografické záznamy. Portál Evropská knihovna poskytuje online přístup ke zdrojům dokumentů 47 evropských národních knihoven a to ve dvaceti jazycích. Všechny země, účastníci se projektu, jsou členy nevládní organizace CENL. Knihovny podílející se na vytváření portálu mohou mít dvojí statut: plné členství, či

základní členství. Rozdíl je v současné chvíli v přístupu ke zdrojům jednotlivých knihoven. Fondy knihoven s plným členstvím lze prohledávat v rámci Evropské knihovny podle různých hledisek, zatímco digitální sbírky ostatních členských knihoven jsou zatím přístupné pouze proklikem přes odkaz na webová sídla jednotlivých knihoven [Národní knihovna, 2008]. Evropskou knihovnu spravuje Nadace evropské digitální knihovny ([European Library Office team](#)), která sídlí v Královské knihovně Nizozemska (Koninklijke Bibliotheek) [Evropská unie, 2008].

### **8.2.2. Projekt TEL-ME-MOR**

Národní knihovny deseti členských států, připojených k Evropské unii v roce 2004, se staly členy Evropské knihovny v rámci projektu *TEL-ME-MOR Evropská knihovna - modulární rozšíření pro zprostředkovávání on-line zdrojů* (*The European Library - Modular Extensions for Mediating Online Resources*, <http://www.telmemor.net>). Projekt byl zahájen 1. února 2005, byl dvouletý a financován v rámci 6. Rámcového programu (FP6) EU technologie informační společnosti (IST). Cíle programu byly následující: rozvoj vědeckovýzkumného programu znalostní společnosti v deseti nových členských zemích, rozšiřování informací o příležitostech k účasti v IST aktivitách, podpora účasti institucí v budoucích IST výzvách z oblastí kulturního dědictví a vzdělávání a také rozšíření přístupu pro badatele, vědce, studenty a učitele v Evropě k výzkumným informačním zdrojům a materiálům (převážně v digitálním formátu), které jsou ve vlastnictví evropských národních knihoven. K docílení tohoto plánu projekt podporoval deset nových členských zemí k získání plného členství v Evropské knihovně, které mělo umožnit jejím uživatelům přístup k informacím a zdrojům z partnerských knihoven bez omezení místem a časem [MARVANOVÁ, 2008].

Mezi významné aktivity projektu TEL-ME-MOR patřilo zjišťování a analýza výzkumných potřeb a priorit v oblastech kultury a vzdělání v nových členských státech, zabezpečení jejich účasti ve formulaci budoucích výzkumných programů EU, podniknout kroky k realizaci integrace sbírek národních knihoven těchto členských států do Evropské knihovny a rozšiřování informovanosti o evropských aktivitách a příležitostech. Hlavní poslání projektu TEL-ME-MOR tedy bylo zajistit vzestup účasti těchto nových členských zemí v budoucích projektech z oblasti kulturního dědictví a vzdělání a rovněž na dalších výzkumných projektech financovaných Evropskou unií [MARVANOVÁ, 2008].

### **8.2.3. Evropská digitální knihovna**

Evropská knihovna byla rozšířena o projekt Evropské digitální knihovny (EDL), který odstartoval v září 2006. Tento projekt je financován Evropskou komisí v rámci programu eContentPlus a jeho trvání je stanoveno na osmnáct měsíců. Cílem EDL není výlučné šíření obsahu, ale koordinace přístupu k digitalizovaným dílům. Jednou z jejích částí bude Evropská digitální knihovna pro rukopisy a staré tisky (EDL-M). Základnou EDL-M se má stát systém Manuscriptorium jako místo pro společnou prezentaci rukopisů a starých tisků ze sbírek evropských národních knihoven [JEŽEK, 2005].

Evropská digitální knihovna byla spuštěna prozatím v beta verzi dne 20. 11. 2008 pod názvem Europeana ([www.europeana.eu](http://www.europeana.eu)). Tento multijazyčný portál poskytuje uživatelům přístup k téměř dvěma milionům digitalizovaných objektů typu knihy, fotografií, map, zvukových nahrávek, obrazů, filmů a archivních záznamů z více než tisícovky knihoven, archivů, muzeí a audiovizuálních sbírek dvaceti sedmi zemí Evropy. Do roku 2010 by počet zpřístupněných digitalizovaných objektů měl dosáhnout deseti milionů. Provoz Europeany bude financován částkou cca 2 000 000 eur ročně z rozpočtu EU. V letech 2009 - 2010 bude pro výzkum digitálních knihoven k dispozici 69 000 000 eur prostřednictvím výzkumných programů Evropské unie. Ve stejném období bude přiděleno z části týkající se informační společnosti programu Konkurenceschopnost a inovace cca 50 000 000 eur na zlepšení přístupu k evropskému kulturnímu a vědeckému dědictví [Evropská unie, 2008].

Každá z institucí, které se podílí na projektu Evropské digitální knihovny, bude mít přesto povinnost platit ročně členský poplatek, který je podle velikosti rozpočtu dané instituce stanoven na 10 000 až 50 000 eur.

## 9. Závěr

Tato diplomová práce se zabývala projekty a programy digitalizace knihovních fondů v České Republice, které v současné době probíhají, nebo byly nedávno ukončené. Zřetel byl brán především na možnosti financování, použité technologie a postupy v digitalizaci, možnosti spolupráce s dalšími institucemi, zpřístupnění digitalizovaných dokumentů v digitálních knihovnách. Prostor byl věnován také popisu digitalizačního procesu, obrazu a obrazových formátů, metadat, procesu ochranného reformátování a také problematice ochrany digitálních objektů. Popsány byly dva nejvýznamnější programy v oblasti digitalizace v naší republice, *Manuscriptorium* a *Kramerius*. Protože v rámci těchto programů spolupracuje množství digitalizujících institucí, byla pozornost věnována především vytvořeným standardům, které jsou instituce povinny v rámci interoperability dodržovat.

Jádrem práce byl popis jednotlivých projektů digitalizace, které jsem v rámci provedené rešerše za tímto účelem vybrala. Při výběru projektů byl zohledňován především předmět digitalizace, jímž podle zadání mohou být jen rukopisy a staré a novodobé tisky, dále dostupnost informací o projektech a také časové rozpětí projektu (kritériem bylo, aby projekt probíhal, byl zahájen nebo ukončen v současnosti).

Tato práce nepřináší nové objevy a skutečnosti, nýbrž na základě podrobného sběru podkladů poskytuje přehled o vybraných projektech digitalizace a popisuje současný stav digitalizace a související problematiku v rámci České republiky s navazujícími liniemi na dění v této oblasti v Evropské unii.

### 9.1. Zhodnocení analyzovaných programů a projektů digitalizace

Jednotlivé programy a projekty digitalizace dokumentů je třeba zhodnotit z několika zásadních hledisek, jimiž jsou:

- o způsob financování
- o spolupráce
- o digitalizační pracoviště (vlastní nebo externí firma)
- o zpřístupnění digitálních dokumentů.

## Způsob financování

Lze konstatovat, že až na výjimky instituce pro své projekty využívají některou z možností dotací státních, nebo Evropské unie.

Dotaci z Národního programu VISK 6 využívají tyto projekty:

- *Historie a vojenství v dobových písemnostech*, Vojenský historický ústav
- *Digitální knihovna historických fondů Vědecké knihovny v Olomouci*

Dotaci z Národního programu VISK 7 čerpají tyto projekty:

- *Reformátování bohemikálních časopisů z historického fondu STK*, Státní technická knihovna
- *Historie a vojenství v dobovém tisku*, Vojenský historický ústav

Dotaci z Národního programu [VISK 5](#) - *Národní program retrospektivní konverze katalogů knihoven v ČR – RETROKON* čerpá *Digitální knihovna map Vědecké knihovny v Olomouci*.

Z některého z dotačních programů Evropské unie čerpají tyto projekty:

- *Digitalizace a zpřístupnění církevních listin v rámci evropského projektu Monasterium*, Národní Archiv ČR, je financován z evropského programu *Culture 2000*
- *HISPRA - Záchrana historických pragensijních a dalších vzácných dokumentů z fondu Městské knihovny*, je financován prostřednictvím *Finančních mechanismů Evropského hospodářského prostoru & Norských finančních mechanismů*

Projekty finančně podporované grantovou agenturou Akademie věd ČR v rámci národního výzkumného programu 1E - *Informační společnost*:

- *Digitální knihovna Akademie Věd ČR*
- *Česká digitální matematická knihovna (DML-CZ)*

Projekty, které financují jejich zřizovatelé:

- *Společná česko-slovenská digitální parlamentní knihovna*, Parlamentní knihovna, financován je z prostředků [Kanceláře Poslanecké sněmovny](#)

- *České sněmy*, Parlamentní knihovna, financován je z prostředků [Kanceláře Poslanecké sněmovny](#)
- *Digitalizovaný archiv časopisů*, je vytvářen, provozován a financován Ústavem pro českou literaturu AV ČR

Na závěr je určitou výjimkou mezi popisovanými projekty výzkumný záměr *Historické fondy Moravské zemské knihovny v Brně*, který je finančně podporován Ministerstvem kultury ČR.

### **Spolupráce**

V souladu se zadáním této diplomové práce byla zkoumána míra spolupráce v oblasti digitalizace mezi jednotlivými institucemi. Z analýzy vyplynulo, že naprostá většina digitalizujících institucí mezi sebou nějakým způsobem spolupracuje.

Určitým druhem spolupráce je podílení se na programu VISK 6 *Manuscriptorium*, v jehož systému zpřístupňují své digitalizované historické dokumenty Moravská zemská knihovna v Brně a Vědecká knihovna v Olomouci, které se tím zároveň zavazují k dodržování určitých standardů. Obdobná je situace u programu VISK 7 *Kramerius*, kdy instituce, které získaly dotaci z tohoto národního programu jsou rovněž povinny dodržovat určené standardy. Jedná se o Vojenský Historický ústav a Státní technickou knihovnu.

Častá je rovněž spolupráce s Národní knihovnou ČR, která pořádá různá školení o digitalizaci, kterých se zúčastňuje Moravská zemská knihovna v Brně a Městská knihovna v Praze. S Národní knihovnou instituce rovněž spolupracují v rámci předcházení vytváření duplicit. Registr CZROMM využívají knihovny, které získaly dotaci na reformátování ohrožených periodik v rámci Národního programu VISK 7 *Kramerius*.

Některé projekty vznikají přímo ve spolupráci několika institucí:

- *Česká digitální matematická knihovna (DML-CZ)*, vzniká ve spolupráci Matematického ústavu AV ČR Praha, Knihovny AV ČR Praha, Ústavu výpočetní techniky Masarykovy univerzity v Brně, Fakulty informatiky Masarykovy univerzity v Brně a Matematicko-fyzikální fakulty Univerzity Karlovy v Praze

- *Společná česko-slovenská digitální parlamentní knihovna*, na její tvorbě se podílí [Parlamentní knihovna](#), [odbor informatiky](#), [Kancelář Poslanecké sněmovny](#) a [odbor informačních technologií Kanceláře Senátu](#) za stranu českou a [Parlamentní knihovna](#), [parlamentní archiv](#) a [odbor informatiky Kanceláře Národní rady Slovenské republiky](#) za stranu slovenskou
- *České sněmy* spoluvytváří [Parlamentní knihovna](#), [odbor informatiky Kanceláře Poslanecké sněmovny](#) a [odbor informačních technologií Kanceláře Senátu](#)
- *Digitalizace a zpřístupnění církevních listin v rámci evropského projektu Monasterium* koordinuje Institut pro výzkum a zpřístupnění církevních pramenů (Institut zur Erforschung und Erschließung kirchlicher Quellen) se sídlem v St. Pölten v Rakousku. Na projektu se podílí instituce z těchto zemí: Bavorsko, Slovinsko, Maďarsko, Chorvatsko, Slovensko, Itálie, Švýcarsko a Národní Archiv České Republiky

### **Digitalizační pracoviště**

Ze zkoumání projektů digitalizace vyplynulo, že řada institucí má již zřízeno vlastní digitalizační pracoviště, naproti tomu mnohé dávají přednost využívání služeb externích pracovišť jiných subjektů, nebo přímo komerčních firem.

Vlastní digitalizační pracoviště zřídily tyto instituce:

- Knihovna AV ČR Praha, provozuje vlastní digitalizační pracoviště v Jenštejně, nabízí službu digitalizace za úplaty i jiným institucím
- Městská knihovna v Praze, vytvořila Oddělení digitalizace
- Moravská zemská knihovna v Brně
- Národní archiv ČR

Externí digitalizační pracoviště nebo komerční firmy využívají tyto instituce:

- Parlamentní knihovna ČR
- Vojenský historický ústav
- Státní technická knihovna
- Vědecká knihovna v Olomouci

## Zpřístupnění digitálních dokumentů

Zpřístupnění výsledků digitalizace uživatelům je vyústěním celého projektu digitalizace. Se zpřístupňováním digitálních objektů v digitálních knihovnách jsou jednotlivé instituce různě daleko, v souvislosti s průběhem projektu.

Pro zpřístupnění digitalizovaných periodik se v českých zemích jednotně užívá open source systém Kramerius, přes nějž lze ovšem zpřístupnit i monografie. Systém Kramerius již používá jako digitální knihovnu Státní technická knihovna a Knihovna AV ČR. Tento systém plánují pro zpřístupnění dokumentů využít také Městská knihovna v Praze a Vojenský historický ústav.

Digitální knihovny v jiných systémech mají již zprovozněny Parlamentní knihovna ČR, Vědecká knihovna v Olomouci, Moravská zemská knihovna v Brně a Ústav pro českou literaturu AV ČR. Jako digitální knihovna rovněž funguje výstup evropského projektu *Monasterium* Archiv MOM a Česká matematická digitální knihovna.

Významným krokem k vyřešení problému duplicit digitálních dokumentů se stalo vytvoření databázové aplikace *RD.CZ* (Registr digitalizace). Tento open source systém vyvíjí Národní knihovna ČR ve spolupráci s Knihovnou AV ČR a firmou INCAD a jeho nasazení do provozu je plánováno na jaro roku 2009, v současnosti však již probíhá ověřovací provoz na interní síti NK ČR. Registr *RD.CZ* bude sloužit k evidenci digitalizačních procesů, k poskytování průběžného přehledu o průběhu těchto procesů a poskytne celkový přehled o stavu a realizaci digitalizačních projektů v Česku.

Z výše uvedené analýzy vyplývá současná situace okolo digitalizace dokumentů v České Republice, prezentována na vybraném vzorku třinácti digitalizačních projektů. Lze říci, že situace není nijak nepříznivá. Digitalizace dokumentů, a nikoliv jen vzácných či ohrožených, je v popředí zájmu kulturních institucí, které jsou podporovány v rámci státu i Evropské unie. Spolupráce, koordinace, technologie i standardy jsou neustále rozvíjeny a vyvíjeny za jediným účelem - zachování evropského kulturního dědictví.




## Seznam použitých informačních zdrojů

- [Akademie věd České republiky. Knihovna. Digitalizační centrum](#). 2005. *Digitální knihovna Akademie věd ČR* [online]. Praha : Akademie věd České republiky, Knihovna, 2005 [cit 2008-10-29]. Dostupný z WWW: <[http://digit.lib.cas.cz/index.php?cat=projekt\\_001](http://digit.lib.cas.cz/index.php?cat=projekt_001)>.
- BARTOŠEK, Miroslav. 2004. Digitální knihovny - teorie a praxe. *Národní knihovna* [online]. 2004, roč. 15, č. 4 [cit. 2008-05-29], s. 233-254. Dostupný také z WWW: <<http://knihovna.nkp.cz/NKKR0404/0404233.html>>. ISSN 1214-0678.
- BARTOŠEK, Miroslav. 2008. [Česká digitální matematická knihovna](#). In *INFORUM 2008 : 14. konference o profesionálních informačních zdrojích, Vysoká škola ekonomická v Praze, 28.-30.5. 2008* [online]. Praha : Albertina icome Praha, 2008 [cit. 2008-09-18]. Dostupný z WWW: <<http://www.inforum.cz/pdf/2008/bartosek-miroslav-cze.pdf>>.
- BARTOŠEK, Miroslav; KREJČÍŘ Vlastimil. 2007. Jak se dělá digitální matematická knihovna. In *Automatizace knihovnických procesů 2007 (AKP 2007), 11. ročník semináře, Liberec, 16.-17. 5. 2007* [online]. [cit. 2008-09-22]. Dostupný z WWW: <<http://www.akvs.cz/akp-2007/08-bartosek-krejcir.pdf>>. ISBN 978-80-01-03691-4.
- CELBOVÁ, Ludmila. 2006. Metadata. In *KTD : Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online databáze]. Praha : Národní knihovna České republiky, 2006 [cit. 2008-11-10]. Systém. č.: 000000543. Dostupný z WWW: <<http://sigma.nkp.cz/cze/ktid>>.
- Česko. Vláda. 2000. *Usnesení vlády České Republiky ze dne 10. dubna 2000 č. 351 o Koncepti státní informační politiky ve vzdělávání*. Praha 2000 [cit 2008-12-10]. Dostupný také z WWW: <[http://racek.vlada.cz/usneseni/usneseni\\_webtest.nsf/0/98B8F697B6F4630BC12571B6006CA75D](http://racek.vlada.cz/usneseni/usneseni_webtest.nsf/0/98B8F697B6F4630BC12571B6006CA75D)>.
- Česko. Ministerstvo financí. 2005a. *Ministerstvo financí České republiky : EU a zahraničí*. [online]. Praha 2005 [cit 2008-10-29]. Dostupný z WWW: <[http://www.mfcr.cz/cps/rde/xchg/mfcr/xsl/fm\\_norska\\_vo.html](http://www.mfcr.cz/cps/rde/xchg/mfcr/xsl/fm_norska_vo.html)>.
- Česko. Ministerstvo financí. 2005b. *Ministerstvo financí České republiky : EU a zahraničí : Schválené projekty* [online]. Praha 2005 [cit 2008-11-13]. Dostupný z WWW: <[http://www.mfcr.cz/cps/rde/xchg/mfcr/xsl/fm\\_norska\\_schvprojekty.html?year=2007](http://www.mfcr.cz/cps/rde/xchg/mfcr/xsl/fm_norska_schvprojekty.html?year=2007)>.
- Česko. Ministerstvo kultury. 2008. *Výběrové dotační řízení Veřejné informační služby knihoven : podprogram č. 7 Národní program mikrofilmování a digitálního zpřístupnění dokumentů ohrožených degradací kyselého papíru Kramerius* [online]. Praha 2008 [cit 2008-12-1]. 9 s., 3 příl. Dostupný z WWW:

<<http://www.mkcr.cz/assets/literatura-a-knihovny/granty-a-programy/visk7-podm2009.doc>>.

- ENRICH Project. 2007. *ENRICH – European Networking Resources and Information concerning Cultural Heritage* [online]. 2007 [cit 2008-11-23]. Dostupný z WWW: <<http://enrich.manuscriptorium.com/index.php?q=node/1>>.
- Evropská unie. 2008. *Europa* [online]. Brusel : Evropská Unie, [cit 2008-11-24]. Press releases RAPID. Evropská digitální knihovna „Europeana“ je nyní přístupná on-line. 20-11-2008. Dostupný z WWW: <<http://europa.eu/rapid/pressReleasesAction.do?reference=IP/08/1747&format=HTML&aged=0&language=CS&guiLanguage=en>>.
- Evropská knihovna. 2008. In *Wikipedie : otevřená encyklopedie* [online]. Posl.edit. 16:09, 17. 11. 2008. [cit 2008-11-24]. Dostupný z WWW: <[http://cs.wikipedia.org/wiki/Evropsk%C3%A1\\_knihovna](http://cs.wikipedia.org/wiki/Evropsk%C3%A1_knihovna)>.
- *Guidelines for Digitization* [online]. 2007. Edited by Katherine M. Wisser. 2007 revised edition. Updated May 16, 2007 [cit. 2008-07-03]. North Carolina ECHO, Exploring Cultural Heritage Online. Glossary. Dostupný z WWW: <<http://www.ncecho.org/dig/digguidelines.shtml>>.
- HOUDEK, Petr; PUKLOVÁ, Petra. 2007. Digitální parlamentní knihovna. In *Archivy, knihovny, muzea v digitálním světě 2007*. Praha : Národní technické muzeum, 2007 [cit 2008-11-19]. Dostupný z WWW: <<http://skip.nkp.cz/KeStazeni/Archivy07/Houdek.ppt>>.
- HOKYNEK, Michal. 2007. *Digitalizace starých tisků a rukopisů z fondů Knihovny Vojenského historického ústavu*. Praha, 2007. 81 s. Diplomová práce. Univerzita Karlova v Praze, Filozofická fakulta, Ústav informačních studií a knihovnictví.
- HROTKOVÁ, Zuzana. 2006. *Digitální knihovny historických dokumentů v prostředí WWW*. Praha, 2006. 108 s. Diplomová práce. Univerzita Karlova v Praze, Filozofická fakulta, Ústav informačních studií a knihovnictví.
- HUTAŘ, Jan. 2007. *Ochrana novodobých knihovných fondů 19. a 20. století* [online]. Verze 1.0. Praha : Univerzita Karlova v Praze, Filozofická fakulta, Ústav informačních studií a knihovnictví, 2007 [cit 2008-10-09], 180 s. Dostupný z WWW: <[http://texty.jinonice.cuni.cz/studijni-texty/hutar-jan/hutar\\_01.pdf](http://texty.jinonice.cuni.cz/studijni-texty/hutar-jan/hutar_01.pdf)>.
- ISO 8879:1986. *Information processing -- Text and office systems -- Standard Generalized Markup Language (SGML)*. 1 ed. 1986-10-23. 155 s. Dostupný také z WWW: <<http://www.iso.org/iso/search.htm?qt=ISO+8879%3A1986&sort=rel&type=simple&published=true>>.
- ISO/IEC 10646:2003. *Information technology-- Universal Multiple-Octet Coded Character Set (UCS)*. 1 ed. 2003-12-15. 1418 s. Dostupný také z WWW:


<[http://www.iso.org/iso/iso\\_catalogue/catalogue\\_tc/catalogue\\_detail.htm?csnumber=39921](http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=39921)>.

- ISO 4087:2005. *Micrographics -- Microfilming of newspapers for archival purposes on 35 mm microfilm*. 3 ed. 2005-04-13. 14 s. Dostupný také z WWW: <[http://www.iso.org/iso/iso\\_catalogue/catalogue\\_tc/catalogue\\_detail.htm?csnumber=26165](http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=26165)>.
- i2010: Digitální knihovny. *Knihovna plus* [online]. 2005, č. 2  [cit. 2008-11-24]. Dostupný z WWW: <[http://knihovna.nkp.cz/knihovnaplus51/digitalni\\_knihovnyEU.htm](http://knihovna.nkp.cz/knihovnaplus51/digitalni_knihovnyEU.htm)>. ISSN 1801-5948.
- JEDLIČKA, Jan. 2005. *Finanční mechanismus Evropského hospodářského prostoru a Norska* [online]. Praha : Česká spořitelna, EU Office České spořitelny, 2005 [cit. 2008-10-29]. 12 s. Dostupný z WWW: <[http://www.csas.cz/banka/content/inet/internet/cs/FM\\_EHPaNorska.pdf](http://www.csas.cz/banka/content/inet/internet/cs/FM_EHPaNorska.pdf)>.
- JEŽEK, Vlastimil ; KNOLL, Adolf. 2005. *Národní knihovna ČR chystá projekt Evropské digitální knihovny : tisková zpráva, 18. 10. 2005* [online]. Praha : Národní knihovna ČR, 2005 [cit. 2008-11-23]. Dostupný z WWW: <[http://www.nkp.cz/files/tz\\_cenl.doc](http://www.nkp.cz/files/tz_cenl.doc)>.
- KNOLL, Adolf. 2006. Digitální zpřístupnění knihovních sbírek. *Ikaros* [online]. 2006, roč. 10, č. 4 [cit. 2008-05-12]. Dostupný z WWW: <<http://www.ikaros.cz/node/3290>>.
- KNOLL, Adolf. 2007a. *Digitální obraz* [online prezentace]. Praha : Národní knihovna České republiky, 2007a [cit. 2008-05-10]. [49 snímků]. Školení Národní knihovny České republiky, část 2. Dostupný z WWW: <[http://www.digit.nkp.cz/Trainings/UNESCOTraining/Unit2\\_DigitalImage\\_cz.ppt](http://www.digit.nkp.cz/Trainings/UNESCOTraining/Unit2_DigitalImage_cz.ppt)>.
- KNOLL, Adolf. 2007b. *Mark-up* [online prezentace]. Praha : Národní knihovna České republiky, 2007b [cit. 2008-05-10]. [27 snímků]. Školení Národní knihovny České republiky, část 4. Dostupný z WWW: <[http://www.digit.nkp.cz/Trainings/UNESCOTraining/Unit4\\_Metadata\\_cz.ppt](http://www.digit.nkp.cz/Trainings/UNESCOTraining/Unit4_Metadata_cz.ppt)>.
- KNOLL, Adolf; PSOHLAVEC, Stanislav. 1997. *Memoriae Mundi Series Bohemica : digitální zpřístupnění vzácných dokumentů*. 1. vyd. Praha : Národní knihovna ČR, 1997. 27 s.
- KOCOUREK, P.; POLIŠENSKÝ, J. 2007. Evidence digitalizovaných dokumentů a sledování procesu zpracování v Národní knihovně ČR. In *INFORUM 2007: 13. konference o profesionálních informačních zdrojích, Vysoká škola ekonomická v Praze, 22.-24. 5. 2007* [online]. Praha : Albertina icome Praha, 2007 [cit. 2008-05-30]. [15 snímků]. Dostupný z WWW: <<http://www.inforum.cz/pdf/2007/kocourek-pavel.pdf>>.

- KREJČÍ, Richard. 2001. JPEG2000: revoluční formát pro kompresi obrázků? *Grafika on-line* [online]. 2001 [cit. 2008-11-07]. Dostupný na WWW: <<http://www.grafika.cz/art/polygrafie/jpeg2000.html>>.
- KUČERA, Karel; ŠIBRAVA, František; MAJER, Martin. 2006. *Manuscriptorium v. 1.0 : Technická kompatibilita* [online]. Verze 1.2, 2006-02-23. Praha : AiP Beroun s.r.o. [cit 2008-10-23]. Dostupný z WWW: <<http://www.manuscriptorium.com/Site/CZE/dokumentace.asp>>.
- LHOTÁK, Martin. 2004. Možnosti spolupráce při vytváření knihovních sbírek. In *Archivy, knihovny, muzea v digitálním světě 2004*. Praha : Národní technické muzeum, 2005, s. 95-96. ISBN 80-7037-138-2.
- [LHOTÁK, Martin](#). 2005. Digitalizace v českých zemích. *Ikaros* [online]. 2005, roč. 9, č. 12 [cit. 2008-05-12]. ISSN 1212-5075. Dostupný z WWW: <<http://www.ikaros.cz/node/2053>>.
- Library of Congress. 2008. *Library of Congress* [online]. Washington (DC) : The Library of Congress, [cit. 2008-12-12]. Standards at the Library of Congress. MODS Metadata Object Description Schema. Dostupný z WWW: <<http://www.loc.gov/standards/mods/>>.
- *Manuscriptorium-základy a kompatibilita* [online]. Verze 1.2, 2006-02-23 Praha, 2006 [cit 2008-10-23]. Dostupný z WWW: <<http://www.manuscriptorium.com/Site/CZE/dokumentace.asp>>.
- MARVANOVÁ, Eva. 2006. Evropská knihovna a projekt TEL-ME-MOR. In *Archivy, knihovny, muzea v digitálním světě 2005*. Praha : Národní technické muzeum, 2006, s. 9-12. ISBN 80-7037-149-8.
- MĚŘÍNSKÁ, Eva. 2008. Pragensijní dokumenty v digitální podobě. *Bulletin SKIP : Svaz knihovníků a informačních pracovníků ČR* [online]. 2008, roč. 7, č. 1 [cit. 2008-11-13]. Dostupný z WWW: <[http://skip.nkp.cz/Bulletin/Bull08\\_107.htm#ti](http://skip.nkp.cz/Bulletin/Bull08_107.htm#ti)>. ISSN 1213-5828.
- Monasterium-konsorciium. 2002. *Monasterium.net-Home* [online]. St. Pölten, 2002 [cit. 2008-10-17]. Dostupný z WWW: <<http://www.monasterium.net/>>.
- Národní archiv. 2008. *Digitalizace a zpřístupnění církevních listin v rámci evropského projektu Monasterium* [online]. Praha, [cit. 2008-10-01]. Dostupný z WWW: <<http://www.nacr.cz/zpravy/monasterium.aspx>>.
- Národní knihovna České republiky. 2004. *Digitalizace a digitální zpřístupnění dokumentů : Systém Kramerius* [online]. 2004-09-23. Praha : Národní knihovna České republiky [cit. 2008-11-03]. Dostupný z WWW: <[http://www.nkp.cz/pages/page.php3?page=weba\\_digitalizace\\_kramerius.htm](http://www.nkp.cz/pages/page.php3?page=weba_digitalizace_kramerius.htm)>.
- Národní knihovna České republiky. 2006. *Digitalizace a digitální zpřístupnění dokumentů : Národní program ochranného reformátování Kramerius* [online]. 2006-05-11. Praha : Národní knihovna České republiky [cit. 2008-11-03].

Dostupný z WWW: <[http://www.nkp.cz/pages/page.php3?page=weba\\_digitalizace.htm](http://www.nkp.cz/pages/page.php3?page=weba_digitalizace.htm)>.

- Národní knihovna České republiky. 2008. *Národní digitální knihovna* [online]. Praha : Národní knihovna České republiky, 2008 [cit. 2008-11-23]. Zahraniční DK a repozitáře, Portál „Evropská knihovna“ (The European Library). Autor Marek Melichar. Dostupný z WWW: <<http://www.ndk.cz/zahranicni-dk-repozitare/portal-201eevropska-knihovna201c-the-european-library/>>.
- Parlamentní knihovna (Praha, Česko). 2003. *Společná česko-slovenská digitální parlamentní knihovna: Informace* [online]. 2003-08-25. Praha: Parlamentní knihovna [cit. 2008-11-19]. Dostupný z WWW: <<http://www.psp.cz/eknih/info.htm>>.
- Parlamentní knihovna (Praha, Česko). 2008. *Digitální knihovna České Sněmy: Informace* [online]. Praha: Parlamentní knihovna [cit. 2008-11-19]. Dostupný z WWW: <<http://www.psp.cz/eknih/snemy/projekt.htm>>.
- PAVLOVÁ, Lucie. 2008. RD.CZ evidence digitalizovaných dokumentů a sledování procesu zpracování. In *Knihovny současnosti 2008, 16. ročník konference, 16.-18. 9. 2008, Seči u Chrudimi*. Brno 2008. 373 s. Dostupný také z WWW: <<http://www.sdruk.cz/sec/2008/b1/>>. ISBN 978-80-86249-49-0.
- POLIŠENSKÝ, J. 2007. Implementace formátu METS v Systému Kramerius. In *Automatizace knihovnických procesů 2007 (AKP 2007), 11. ročník semináře, Liberec, 16.-17. 5. 2007* [online]. [cit. 2008-09-22] Dostupný z WWW: <<http://www.akvs.cz/akp-2007/13-polisensky.pdf>>. ISBN 978-80-01-03691-4.
- PSOHLAVEC, Stanislav. 1999. Memoriae Mundi series Bohemica : digitalizace rukopisů a starých tisků : výsledky spolupráce NKČR s firmou Albertina icome Praha s.r.o. In *Středověké rukopisy a jejich prezentace : 7. odborná konference Problematika historických a vzácných knižních fondů Čech, Moravy a Slezska* [online]. Olomouc : Vědecká knihovna v Olomouci, 1999 [cit. 2008-04-04]. Dostupný z WWW: <<http://www.vkol.cz/konf/98hrad13.htm>>.
- [PSOHLAVEC, Stanislav](#). 2004. Projekt Memoria-zkušenosti z jednoročního zpřístupnění plných obrazů digitalizovaných rukopisů a starých tisků na internetu. In *Archivy, knihovny, muzea v digitálním světě 2004*. Praha : Národní technické muzeum, 2005, s. 56-63. ISBN 80-7037-138-2.
- [PSOHLAVEC, Stanislav](#). 2005. Přednosti a rizika digitálních dokumentů. *Ikaros* [online]. 2005, roč. 9, č. 12 [cit. 2008-05-12]. Dostupný z WWW: <<http://www.ikaros.cz/node/2064>>. ISSN 1212-5075.
- PSOHLAVEC, Stanislav. 2006. *Manuscriptorium Quality : Kvalita obrazových dat, definice a výklad* [online]. Verze 1.2, 2006-02-23. Praha : AiP Beroun [cit. 2008-09-18]. Dostupný z WWW: <<http://www.manuscriptorium.com/Site/CZE/dokumentace.asp>>.

- PSOHLAVEC, Stanislav. 2008. ENRICH - Evropská digitální knihovna rukopisů. In *INFORUM 2008: 14. konference o profesionálních informačních zdrojích, Vysoká škola ekonomická v Praze, 28.-30.5. 2008* [online]. Praha : Albertina icome Praha, 2008 [cit 2008-09-18]. [28 snímků]. Dostupný z WWW: <<http://www.inforum.cz/archiv/inforum2008/sbornik/80/>>. ISSN 1801-2213.
- Státní technická knihovna (Praha, Česko). 2004. *Výroční zpráva Státní technické knihovny 2003*. Odpovědný redaktor Miloslava Rupešová. Praha : Státní technická knihovna, 2004. 76 s. Dostupný také z WWW: <<http://www.stk.cz/cs/o-nas/pro-media/vyrocní-zpravy/>>. ISBN 80-86504-13-1. ISSN 1212-6268.
- Státní technická knihovna (Praha, Česko). 2005. *Výroční zpráva Státní technické knihovny 2004*. Odpovědný redaktor Miloslava Rupešová. Praha : Státní technická knihovna, 2005. 70 s. Dostupný také z WWW: <<http://www.stk.cz/cs/o-nas/pro-media/vyrocní-zpravy/>>. ISBN 80-86504-14-X. ISSN 1212-6268.
- Státní technická knihovna (Praha, Česko). 2006. *Výroční zpráva Státní technické knihovny 2005*. Odpovědný redaktor Miloslava Rupešová. Praha : Státní technická knihovna, 2006. 74 s. Dostupný také z WWW: <<http://www.stk.cz/cs/o-nas/pro-media/vyrocní-zpravy/>>. ISSN 1212-6268.
- Státní technická knihovna (Praha, Česko). 2007. *Výroční zpráva Státní technické knihovny 2006*. Praha : Státní technická knihovna, 2007. 78 s. Dostupný také z WWW: <<http://www.stk.cz/cs/o-nas/pro-media/vyrocní-zpravy/>>. ISSN 1212-6268.
- Státní technická knihovna (Praha, Česko). 2008. *Roční zpráva za rok 2007*. Zpracoval: Jan Bayer, Marie Vadlejchová. Praha : Státní technická knihovna, 2008. 25 s., 12 tab., 1 příl. Dostupný také z WWW <<http://www.stk.cz/cs/o-nas/rocní-zpravy/>>. ISBN 80-86504-16-6. ISSN 1212-6268.
- SYNKOVÁ, Veronika. 2007. *Aktuální problémy identifikace a popisu webovských informačních zdrojů*. Praha, 2007. vi, 161 s., 26 s. příloh. Rigorózní práce. Univerzita Karlova v Praze, Filozofická fakulta, Ústav informačních studií a knihovnictví.
- TICHÝ, Zdeněk A. 2006. Parlamentní knihovna míří na Letnou. *Knihovna* [online]. 2006, roč. 17, č. 1, s. 53-58  [cit. 2008-11-19]. Dostupný z WWW: <<http://knihovna.nkp.cz/knihovna61/tichy5.htm>>. ISSN 1801-5948.
- UHLÍŘ, Zdeněk. 2002. Standard MASTER : katalogizace rukopisů v XML. *Národní knihovna*. 2002, roč. 13, č. 2, s. 84-101. ISSN 0862-7487.
- UHLÍŘ, Zdeněk. 2006. *Manuscriptorium v. 1.0 : Výběr a popis dokumentů* [online]. Verze 1.2, 2006-02-23. Praha : Národní knihovna České republiky [cit 2008-10-23]. Dostupný z WWW: <<http://www.manuscriptorium.com/Site/CZE/dokumentace.asp>>.

- VESELÝ, Zdeněk. 2003. *Dějiny českého státu v dokumentech*. Praha, Epoque 2003. 495 s. ISBN 80-86328-26-0.
- Vědecká knihovna (Olomouc, Česko). 2005a. *Digitální knihovna historických fondů Vědecké knihovny v Olomouci* [online]. Posled. edit. 2008-02-25. Olomouc, [cit 2008-11-27]. Dostupný z WWW: <<http://dig.vkol.cz/>>.
- Vědecká knihovna (Olomouc, Česko). 2005b. *Digitální knihovna map Vědecké knihovny v Olomouci* [online]. Posled. edit. 2006-02-6. Olomouc, [cit 2008-11-27]. Dostupný z WWW: <<http://mapy.vkol.cz/>>.
- VISK veřejné informační služby knihoven : program ministerstva kultury [online]. 2005-01-11. Praha, [cit. 2008-05-10]. Cíle programu. Dostupný také z WWW: <<http://visk.nkp.cz/VISKcile.htm>>.
- VISK veřejné informační služby knihoven : program ministerstva kultury [online]. 2005-01-13. Praha, [cit. 2008-05-10]. VISK 7 Národní program mikrofilmování a digitálního zpřístupňování dokumentů ohrožených degradací kyselého papíru – Kramerius. Dostupný také z WWW: <<http://visk.nkp.cz/VISK7.htm>>.
- VISK veřejné informační služby knihoven : program ministerstva kultury [online]. 2008-06-11. Praha, [cit. 2008-05-10]. VISK 6 Národní program digitálního zpřístupnění vzácných dokumentů Memoriae Mundi Series Bohemica. Dostupný z WWW: <<http://visk.nkp.cz/VISK6.htm>>.
- VOJNAR, Martin. 2002. Digitální knihovna a automatizované zpracování dotazů aneb Týden knihoven ve Vědecké knihovně v Olomouci. *Ikaros* [online]. 2002, roč. 6, č. 10 [cit. 2008-11-18]. Dostupný z WWW: <<http://www.ikaros.cz/node/1192>>. ISSN 1212-5075.
- VOJNAR, Martin. 2005. Nové standardy digitálních knihoven pro dlouhodobou ochranu. *Knihovna* [online]. 2005, roč. 16, č. 2, s. 45-58 [cit. 2008-11-21]. Dostupný z WWW: <<http://knihovna.nkp.cz/knihovna52/vojnar.htm>>. ISSN 1801-3252.
- VOJTÁŠEK, Filip. 2000a. Využití grafických formátů JPEG a DjVu v digitalizaci. In *INFORUM 2000 : 6. ročník konference o profesionálních informačních zdrojích, Vysoká škola ekonomická v Praze, 23.-25. května 2000* [online]. Praha : Albertina icome Praha, 2000 [cit. 2008-10-02]. Dostupný z WWW: <<http://www.inforum.cz/inforum2000/prednasky/vyuzitigrafick.htm>>.
- VOJTÁŠEK, Filip. 2000b. Dlouhodobá archivace digitálních dokumentů. *Ikaros* [online]. 2000, roč. 4, č. 10 [cit. 2008-10-09]. Dostupný z WWW: <<http://www.ikaros.cz/node/675>>. ISSN 1212-5075.
- ŽABIČKA, Petr. 2004a. *Historické fondy Moravské zemské knihovny v Brně : Zpráva o řešení výzkumného záměru MK00009494301 řešeného Moravskou zemskou knihovnou v Brně : Průběžná zpráva za rok 2004*. Brno: Moravská

zemská knihovna v Brně, 2004 [cit. 2008-11-20]. Dostupný také z WWW:  
<<http://www.mzk.cz/projekty/histfondy/>>.

- ŽABIČKA, Petr; MACHOVÁ, Jitka. 2004b. Historické fondy a nové možnosti zpřístupnění informací. In *Archivy, knihovny, muzea v digitálním světě 2004*. Praha : Národní technické muzeum, 2005, s. 78 - 85. ISBN 80-7037-138-2.
- ŽABIČKA, Petr. 2005. *Historické fondy Moravské zemské knihovny v Brně : Zpráva o řešení výzkumného záměru MK00009494301 řešeného Moravskou zemskou knihovnou v Brně : Průběžná zpráva za rok 2005*. Brno: Moravská zemská knihovna v Brně, 2005 [cit. 2008-11-20]. Dostupný také z WWW:  
<<http://www.mzk.cz/projekty/histfondy/>>.
- ŽABIČKA, Petr a kol. 2006. *Historické fondy Moravské zemské knihovny v Brně : Zpráva o řešení výzkumného záměru MK00009494301 řešeného Moravskou zemskou knihovnou v Brně : Průběžná zpráva za rok 2006*. Brno: Moravská zemská knihovna v Brně, 2004 [cit. 2008-11-20]. Dostupný také z WWW:  
<<http://www.mzk.cz/projekty/histfondy/>>.
- ŽABIČKA, Petr a kol. 2007. *Historické fondy Moravské zemské knihovny v Brně : Zpráva o řešení výzkumného záměru MK00009494301 řešeného Moravskou zemskou knihovnou v Brně. Průběžná zpráva za rok 2007*. Brno: Moravská zemská knihovna v Brně, 2004 [cit. 2008-11-20]. Dostupný také z WWW: <<http://www.mzk.cz/projekty/histfondy/>>.



