

Univerzita Karlova v Praze

Filozofická fakulta

Ústav informační studií a knihovnictví

Studijní program: Informační studia a knihovnictví:

Bc. Kateřina Vojřová

PROJEKT HISPRA ODDĚLENÍ DIGITALIZACE MĚSTSKÉ KNIHOVNY V PRAZE

THE HISPRA PROJECT OF THE MUNICIPAL LIBRARY OF PRAGUE'S
DIGITALISATION DEPARTMENT

Diplomová práce

ZADÁNÍ DIPLOMOVÉ PRÁCE

Poděkování

Na tomto místě bych ráda poděkovala vedoucí práce PhDr. Evě Bratkové, PhD. za podporu a odborné připomínky, které mi poskytla při realizaci této práce. Dále bych chtěla poděkovat vedoucí Odboru knihovních fondů PhDr. Zuzaně Kopencové, Mgr. Evě Měřínské, Mgr. Věře Benyovszké a dalším kolegyním a kolegům z Městské knihovny v Praze. V nespolední řadě děkuji své rodině za trpělivost a hektolitry čaje.

Prohlášení:

Prohlašuji, že jsem diplomovou práci vypracovala samostatně, že jsem řádně citovala všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

V Praze dne 1.1.2012

podpis

Anotace

Diplomová práce analyzuje projekt HISPRA Oddělení digitalizace Městské knihovny v Praze – vývoj projektu, harmonogram a financování. Další část práce představuje Oddělení digitalizace, předpoklady založení, vývoj, organizační a personální zajištění a technické a programové vybavení. Dále práce představuje předmět digitalizace, zde je jádrem prezentace podstatných rysů fondu historických pragensijních dokumentů. Poté je uvedena strategie výběru dokumentů. Práce prezentuje také jednotlivé procesy digitalizace včetně zpřístupnění v digitální knihovně Kramerius.

Klíčová slova

digitalizace, Městská knihovna v Praze, metadata, digitální knihovna, digitalizační centra

Abstract

The thesis analyses project HISPRA of The Municipal Library of Prague's Digitalisation Department – development of the project, schedule and funding. Other part of thesis presents Digitalisation Department, assumptions of establishment of the department, development, organization and staff, and hardware and software. Furthermore, there is a presentation of an object of digitization. The most important in this part of thesis is presentation of features of the historical documents about Prague. Furthermore, there is an introduction of document's selection strategy. Thesis also presents the various processes of digitization, including access to The Digital Library Kramerius.

Keywords

digitalisation, Municipal Library of Prague, metadata, digital library, digitalization centers

Obsah

SEZNAM POUŽITÝCH ZKRATEK	9
PŘEDMLUVA	11
1 SPECIFIKACE PROJEKTU HISPRA	13
1.1 VÝVOJ PROJEKTU A JEHO CÍLE	13
1.1.1 <i>Snahy MKP o digitalizaci před projektem HISPRA</i>	13
1.1.2 <i>Počátky projektu HISPRA</i>	13
1.1.3 <i>Cíle projektu</i>	14
1.2 HARMONOGRAM A FINANCOVÁNÍ PROJEKTU	15
1.2.1 <i>Harmonogram</i>	15
1.2.2 <i>Financování</i>	16
2 ODDĚLENÍ DIGITALIZACE.....	17
2.1 PŘEDPOKLADY JEHO ZALOŽENÍ	17
2.2 VÝVOJ	18
2.3 ORGANIZAČNÍ A PERSONÁLNÍ ZAJIŠTĚNÍ	20
2.4 ORGANIZACE PRÁCE	20
2.5 TECHNICKÉ A PROGRAMOVÉ VYBAVENÍ.....	21
2.5.1 <i>Pracovní stanice</i>	21
2.5.2 <i>OCR klíč</i>	22
2.5.3 <i>Skenery</i>	22
2.5.3.1 <i>Zeutschel</i>	22
2.5.3.2 <i>Minolta</i>	24
2.5.3.3 <i>Výběr zařízení jako náhrada za skener Minolta</i>	25
2.5.3.4 <i>Bookeye</i>	26
2.5.4 <i>Servery</i>	27
2.5.5 <i>Software</i>	28
2.6 SPOLUPRÁCE S JINÝMI SUBJEKTY	28
3 PŘEDMĚT DIGITALIZACE	29
3.1 FOND DIVADELNÍHO A FILMOVÉHO ÚSEKU ÚK MKP	29
3.1.1 <i>Divadelní a filmový úsek ÚK MKP</i>	30
3.1.2 <i>Dokumenty divadelního a filmového úseku určené k digitalizaci</i>	30
3.2 FOND HUDEBNÍHO ÚSEKU.....	31
3.2.1 <i>Hudební úsek MKP</i>	32
3.2.2 <i>Dokumenty z hudebního fondu určené k digitalizaci</i>	33
3.3 PRAGENSIJNÍ FOND	33

3.3.1	<i>Studovna pragensíí</i>	34
3.3.2	<i>Dokumenty z historického pragensijního fondu určené k digitalizaci</i>	35
4	SPECIFIKACE FONDU HISTORICKÝCH PRAGENSIJNÍCH DOKUMENTŮ	37
4.1	POJEM PRAGENSÍÍ A PRAGESNIJNÍ LITERATURY	38
4.2	VYMEZENÍ PROFILU PRAGENSIJNÍHO FONDU	40
4.3	ULOŽENÍ FONDU PRAGENSÍÍ	49
5	DIGITALIZACE V MKP	50
5.1	STRATEGIE VÝBĚRU DOKUMENTŮ	50
5.2	PROCES DIGITALIZACE V MKP	52
5.2.1	<i>Příprava fondu</i>	54
5.2.2	<i>Skenování</i>	56
5.2.2.1	Grafický formát.....	58
5.2.2.2	Rozlišení.....	60
5.2.2.3	Barevná hloubka.....	60
5.2.2.4	Postup při skenování	61
5.2.2.4.1	Zeutschel.....	62
5.2.2.4.2	Bookeye.....	65
5.2.3	<i>Tvorba metadat</i>	67
5.2.3.1	Popisná metadata.....	68
5.2.3.2	Strukturální metadata	73
5.2.3.3	Administrativní metadata.....	78
5.2.4	<i>Úprava naskenovaných obrázků</i>	80
5.2.4.1	Načtení obrazů do Siria.....	80
5.2.4.2	Ořez	82
5.2.4.2.1	Ruční kontrola okrajů	82
5.2.4.2.2	Manuální kontrola vyrovnání.....	84
5.2.4.3	Další nástroje pro úpravu skenů	85
5.2.5	<i>Optické rozpoznávání znaků</i>	86
5.2.6	<i>Kontrola dat</i>	87
5.2.7	<i>Export dat</i>	89
5.2.8	<i>Zveřejnění pro uživatele</i>	90
5.2.9	<i>Archivace dat</i>	93
5.2.10	<i>Kontrola fyzických svazků a určení následné péče</i>	94
6	PROPAGACE	96
6.1	ONLINE PREZENTACE	96

6.2	PLAKÁT A LETÁKY, DVD	97
6.3	PRAHA V KNIHOVNĚ	98
6.4	KONFERENCE DIGITALIZACE ANEB KONEC OSLÍCH UŠÍ	100
7	DIGITÁLNÍ KNIHOVNA KRAMERIUS	102
7.1	TECHNICKÁ SPECIFIKACE	102
7.2	UŽIVATELSKÉ ROZHRANÍ SYSTÉMU KRAMERIUS	103
7.2.1	<i>Vyhledávání</i>	104
7.2.2	<i>Vývoj systému Kramerius</i>	107
	ZÁVĚR	110
	SEZNAM POUŽITÉ LITERATURY	115
	SEZNAM VYOBRAZENÍ.....	123
	SEZNAM PŘÍLOH	124
	PŘÍLOHY.....	I

Seznam použitých zkratek

AACR2	Anglo-American Cataloguing Rules
AKS	Automatizovaný knihovní software
AV ČR	Akademie věd České republiky
CCD	Charge-Coupled Device
CD	Compact Disc
CMS	Content Management System
DILIA	Divadelní, literární, audiovizuální agentura
DPI	Dots per inch
DTD	Document Type Definition
DVD	Digital Video Disc
EHP	Evropský hospodářský prostor
EUR	Euro
Fedora	Flexible Extensible Digital Object Repository Architecture
FOXML	Fedora Object XML
GB	Gigabyte
GNU GPL	GNU General Public License
GPS	Global Positioning System
HDD	Hard Disc Drive
HISPRA	Historická pragensia
HW	Hardware
IR	Infrared
JPEG	Joint Photographic Experts Group
LCD	Liquid Crystal Display
LED	Light-Emitting Diode

METS	Metadata Encoding and Transmission Standard
MKP	Městská knihovna v Praze
MZK	Moravská zemská knihovna
ND	Národní divadlo
NK ČR	Národní knihovna České republiky
OAI-PMH	Open Archives Initiative - Protocol for Metadata Harvesting
OCR	Optical Character Recognition
OD	Oddělení digitalizace
OCH	Obsahová charakteristika
OIT	Oddělení informačních technologií
PNG	Portable Network Graphics
PREMIS	Preservation Metadata Maintenance Activity
RD.CZ	Registr digitalizace České republiky
SGML	Standard Generalized Markup Language
S-PVA	Super-Patterned Vertical Alignment
SQL	Structured Query Language
SW	Software
TB	Terrabyte
TIFF	Tagged Image File Format
TXT	Text file
ÚK	Ústřední knihovna
URI	Uniform Resource Identifier
UV	Ultraviolet
VISK	Veřejná informační služba knihoven
XML	Extensible Markup Language

Předmluva

Digitalizace a s ní spojené další termíny jsou skloňovány v posledních letech v prostředí knihoven velmi často. V této práci jsem se rozhodla popsat a zhodnotit projekt HISPRA Oddělení digitalizace Městské knihovny v Praze. Důvodem výběru právě tohoto tématu je jeho propojení s mou pracovní činností. Do Oddělení digitalizace jsem nastoupila v první vlně nabírání pracovních sil v době vzniku oddělení na konci roku 2007 jako řadový pracovník. A v Oddělení digitalizace pracuji s půlroční přestávkou v roce 2009 v době studia v rámci programu Erasmus ve Finsku dodnes. V roce 2010 jsem po odchodu bývalé vedoucí Oddělení digitalizace vyhrála výběrové řízení a stala vedoucí oddělení. Na základě osobních zkušeností jak dříve z pozice zaměstnance oddělení, tak nyní z pozice vedoucí Oddělení digitalizace, mohu zhodnotit projekt HISPRA z obou těchto hledisek. Především cenné se ukázaly být moje postřehy ze začátku existence Oddělení digitalizace – osobně jsem prošla všechny popisované digitalizační procesy a vycházím tedy z vlastních zkušeností. Protože jsem byla v Oddělení digitalizace od jeho počátku, mohla jsem se podílet na vzniku pracovních postupů a jejich uzpůsobení potřebám oddělení. Také jsem během svého působení sepsala pracovní manuál zaměstnance Oddělení digitalizace a své poznatky jsem předávala při zaškolování nově přijatých zaměstnanců. Po svém návratu do oddělení ze studijního pobytu jsem se kromě digitalizace historických dokumentů také ujala role vedení projektu digitalizace beletrie.

Cílem této práce je analyzovat a zhodnotit projekt digitalizace a zpřístupňování historických dokumentů z fondu Městské knihovny v Praze. Pro lepší orientaci je text rozdělen do kapitol na několika úrovních.

V první kapitole je charakterizován samotný projekt HISPRA – historie a cíle projektu a s nimi související harmonogram a plán projektu a nezbytnou část každého projektu tvořící plán financování. Další část práce představí nově vzniklé Oddělení digitalizace, od jeho programového a technického vybavení přes personální obsazení až k charakterizaci vývoje oddělení.

Další část práce tvoří kapitola představující jednotlivé úseky a služby, které provozují, a samotný předmět digitalizace, tedy popis vybraných částí fondu vyhrazených k digitalizaci. Klíčová část práce představuje specifikace historického pragensijního fondu a porovnání celkového fondu s částí fondu určenou k digitalizaci. Práce pokračuje prezentací strategie výběru dokumentů k digitalizaci, po níž následuje prezentace jednotlivých procesů

digitalizace. Předposlední kapitola popisuje propagaci projektu HISPRA a jeho výsledků. V poslední kapitole je vyhodnoceno zpřístupnění digitalizovaných dokumentů v digitální knihovně MKP v systému Kramerius.

V rámci této práce jsem se rozhodla pro lepší srozumitelnost a jednotnost textu popisovat pouze zpracování monografií. Periodické dokumenty byly během realizace projektu co do počtu dokumentů v menšině a byly také digitalizovány později. Zpracování periodik je komplikovanější a pravidla jejich popisu jsou natolik rozsáhlá, že by v rámci rozsahu této práce nebylo možné jim věnovat dostatečný prostor. Z důvodů srozumitelnosti textu byly také vynechány některé podrobné technické detaily související především s HW a SW vybavením digitalizačního pracoviště.

V rámci přípravy práce byla provedena rešerše na téma historických pragensií a procesu digitalizace, vybrané tituly jsou uvedeny v seznamu literatury. Jedná se spíše o teoreticky zaměřené dokumenty, informace pro větší část práce a dokumenty praktičtější povahy byly získávány z Oddělení digitalizace a dalších oddělení Městské knihovny v Praze, především ze Studovny pragensií a z Hudebního úseku a z Divadelní a filmového úseku Ústřední knihovny Městské knihovny v Praze. Tyto dokumenty mají povahu šedé literatury a z velké části se jedná o interní materiály, ať již zápisy z porad nebo podklady pro vypracování projektu. V práci vycházím z osobních pracovních zkušeností z praktické digitalizace, také z teoretických poznatků spojených s návštěvou odborných seminářů a v neposlední řadě i z prezentací zkušeností s digitalizací právě v rámci těchto odborných seminářů.

Diplomová práce byla vypracována podle Pravidel a metodických pokynů pro zpracování a odevzdání závěrečných prací Filozofické fakulty Univerzity Karlovy v Praze.

Pro citování je v textu využit tzv. Harvardský styl citování, jedná se o běžnou metodu citování pomocí prvního údaje záznamu a data vydání, v případě potřeby jsou údaje doplněny také o stránkování. Všechny tyto údaje jsou uvedeny v hranatých závorkách. Použitá literatura je citována podle normy ISO 690:2010. Seznam použité literatury je řazen podle prvního údaje záznamu, pokud se první údaje shodují, pak jsou k řazení použity další údaje.

1 Specifikace projektu HISPRA

1.1 Vývoj projektu a jeho cíle

1.1.1 Snahy MKP o digitalizaci před projektem HISPRA

Městská knihovna v Praze (dále také jako MKP) začala s ochranou svých fondů pomocí jejich digitalizace ještě před projektem HISPRA¹, od roku 2001 se ochraně svých fondů intenzivně věnovala. Hlavním přínosem bylo vypracování plánu digitalizace historických fondů v roce 2002.

Během let 2002–2005 bylo zdigitalizováno několik dokumentů díky využití grantových prostředků z Ministerstva kultury ČR v rámci programů Veřejné informační služby knihoven (dále VISK) č. 6 (*Národní program digitálního zpřístupnění vzácných dokumentů Memoriae Mundi Series Bohemica*) a VISK č. 7 (*Národní program mikrofilmování a digitálního zpřístupňování dokumentů ohrožených degradací kyselého papíru – Kramerius*).

Mezi tyto dokumenty patřila *Pražská bible* z roku 1488, *Orientační plán hlavního města Prahy* z roku 1938 a pražské adresáře. Zpracování těchto dokumentů bylo zadáno digitalizačnímu pracovišti Akademie věd ČR. Za období od roku 2002 do roku 2005 bylo zdigitalizováno pouhých pět dokumentů, zatím co sbírky ohrožených dokumentů byly rozsáhlé.

Městská knihovna proto začala hledat jiný způsob financování záchrany svých sbírek. Jako jedna z možností, jak fondy ohrožené především degradací kyselého papíru zachránit, se ukázaly být programy tzv. norských fondů.

1.1.2 Počátky projektu HISPRA

V roce 2005 MKP začala zpracovávat projekt do programu Finančních mechanismů Evropského hospodářského prostoru a Norska (v prioritní oblasti *Záchrana evropského kulturního dědictví*, 1. výzva 18. 5. 2005).

Práce na projektu trvala díky náročnosti na zpracování téměř dva roky a za její realizací stojí především vedoucí odboru knihovnických fondů Mgr. Zuzana Kopencová. Žádost byla

¹ Název projektu HISPRA je odvozen od hlavní části fondu určeného k digitalizaci, od fondu historických pragensií. Jedná se tedy o akronym.

podána 19. 1. 2006 a byla schválena k 31. 1. 2007. Projekt byl nazván *Záchrana historických pragensijních a dalších vzácných dokumentů z fondu Městské knihovny v Praze*.

1.1.3 Cíle projektu

Hlavní cíl projektu je deklarován v žádosti o grant jako „trvalé uchování a zpřístupnění historických pragensijních a dalších vzácných dokumentů“ [MĚSTSKÁ, 2005c].

Tyto pragensijní dokumenty, teatraia a hudebniny obsahují jedinečné informace a jsou v některých případech dokonce ojedinělými exempláři na území České republiky. Díky častému využívání čtenáři, vlivem degradace papíru a působením špatného životního prostředí jsou informace v těchto dokumentech ohroženy. Vzhledem k povaze dokumentů byla jako jedna z možných metod užívaných k záchraně dokumentů postižených degradací papíru zvolena metoda ochranného reformátování. Ke splnění toho hlavního cíle bylo třeba ustanovit menší podcíle, které na sebe navazují a společně vedou k jeho naplnění.

Prvním podcílem se zcela logicky stalo vybudování odpovídajícího digitalizačního pracoviště, bez jehož existence by bylo zcela nemožné digitalizovat v projektu vymezený počet dokumentů – tedy 350 000 stránek (asi 3 000 dokumentů). Tato etapa byla pro zjednodušení nazvána jako „*zázemí*“.

Druhý podcíl byl definován jako samotné pořízení digitálních kopií dokumentů, a to metodou přímé digitalizace (tedy bez použití mezičlánku nejčastěji ve formě mikrofilmu). Lze jej jednoslovně vyjádřit heslem „*záchrana*“.

Současně na podcíl záchrany navazuje třetí podcíl v pořadí definovaný jako zlepšení přístupu veřejnosti k pragensijním fondům vybudováním digitální knihovny. Jednoduše řečeno bylo třeba digitalizované dokumenty nabídnout uživatelům v takovém prostředí, které by umožnilo rychlý a jednoduchý přístup k těmto dokumentům. Jako zapamatovatelné heslo bylo vybráno slovo „*zpřístupnění*“.

A poslední, ale rozhodně ne zanedbatelný cíl zněl rozšířit povědomí o zdigitalizovaných dokumentech a jejich využití, a to různými způsoby jak v očích odborné tak neodborné veřejnosti. Mezi komunikační kanály byly vybrány především upozornění na webu MKP, letáky a brožury, výukové DVD a závěrečná konference a z ní vyplývající vícejazyčný sborník. Tento poslední podcíl byl vyjádřen slovem „*propagace*“.

1.2 Harmonogram a financování projektu

1.2.1 Harmonogram

Harmonogram realizace celého projektu byl rozdělen do čtyř hlavních etap, které bylo třeba realizovat v rámci tříletého trvání projektu.

První etapa nazvaná *Vybudování digitalizačního pracoviště* trvala od dubna 2007 do července 2007. Původně bylo zamýšleno vybudovat digitalizační pracoviště v prostorách Ústřední knihovny na Mariánském náměstí v Praze a pouze upravit tyto prostory tak, aby odpovídaly nárokům kladeným na digitalizační pracoviště. Vzhledem k následným změnám a celkové reorganizaci v Ústřední knihovně MKP bylo zjištěno, že prostory pro digitalizační pracoviště již nejsou adekvátní, a realizační tým začal hledat jiné možnosti. Po dlouhých jednáních bylo nakonec nalezeno odpovídající místo v Provaznické ulici, které patřilo Magistrátu hl. m. Prahy. MKP tak za výhodných podmínek získala v centru Prahy nedaleko od Ústřední knihovny prostor pro nově zbudované digitalizační pracoviště. Stavební úpravy zahrnovaly položení nové elektroinstalace, položení optického kabelu k datovému úložišti, nainstalování klimatizace atd.

Druhá etapa v období od srpna 2007 do července 2008 měla trvat 12 měsíců a její součástí bylo vybavení digitalizačního centra, zkušební provoz a následná digitalizace. V prvních třech měsících této etapy mělo být digitalizační pracoviště vybaveno mobiliářem a digitalizační technikou (dvěma planetárními skenery). Další tři měsíce mělo trvat vybavení hardwarem a softwarem – tedy stolními počítači a vhodným softwarem pro zpracování, indexování a ukládání digitalizovaných dat do archivu. Tento proces byl reálně zkrácen o 2 měsíce, takže vyškolení odborných pracovníků mohlo být provedeno dříve a zkušební provoz mohl být zahájen již v prosinci roku 2007 a již od dubna 2008 mohl probíhat rutinní provoz digitalizačního pracoviště.

Třetí etapa od srpna 2008 do července 2009 pokrývala oblasti digitalizace, zpřístupnění digitálních dokumentů a především medializaci projektu zahrnující tvorbu plakátů, publikace a DVD pro děti a mládež.

Čtvrtá etapa projektu od srpna 2009 do dubna 2010 zahrnovala pokračování digitalizace dokumentů a realizaci konference s mezinárodní účastí, která uzavírala celý projekt.

1.2.2 Financování

Náklady projektu byly vyčísleny na 507 740 eur, z toho 3 300 eur potřebných na pořádání konference nebylo možno zahrnout do oprávněných nákladů. Z finančních mechanismů EHP/Norska získala MKP 80 % nákladů, tedy 403 552 eur. Zbýlých 20 %, tedy 104 188 eur přislíbil spolufinancovat Magistrát hl. m. Prahy.

V době vypisování tohoto grantu byl měnový kurz české koruny k euru 30,186 Kč za 1 EUR. Během tří let došlo vlivem hospodářské situace k posilování české koruny a kurz klesl až na pouhých 24 Kč. Vzhledem k tomu, že finanční náklady byly vykazovány čtvrtletně, peníze z grantu tak ubývaly i bez toho, že by byly čerpány. Tento problém byl částečně vyřešen tím, že od napsání projektu po výslednou realizaci výběrových řízení na HW a SW klesla jejich cena natolik, že se podařilo ztráty poklesem kurzu vyrovnat.

Rozpočet byl rozvržen do jednotlivých let trvání projektu (viz Tabulka 1). V roce 2006 byly vypočtené náklady jen minimální, zatím se jednalo pouze o přípravu samotného projektu. V roce 2007 tvořilo největší položku vybavení digitalizačního pracoviště, především z hlediska HW a SW. V letech 2008 a 2009 se největší položkou staly platy zaměstnanců.

	Roky					celkem
	2006	2007	2008	2009	2010	
Příjem finančních prostředků						
Celkové finanční zdroje	29 600 €	274 100 €	114 130 €	89 910 €	0 €	507 740 €
Čisté příjmy						0 €
Celkový příjem finančních prostředků	29 600 €	274 100 €	114 130 €	89 910 €	0 €	507 740 €
Výdej finančních prostředků						
Pracovní síla	2 590 €	71 450 €	75 020 €	62 080 €		211 140 €
Služby	26 560 €		37 380 €	25 630 €		89 570 €
Vybavení		201 240 €				201 240 €
Administrativní náklady	220 €	370 €	690 €	1 590 €		2 870 €
Suroviny						0 €
Získání budov						0 €
Získání půdy						0 €
Energie (elektrina, vytápění, paliva)	160 €	640 €	640 €	480 €		1 920 €
Další						0 €

Finanční služby	70 €	400 €	400 €	130 €		1 000 €
Celkový výdej finančních prostředků	29 600 €	274 100 €	114 130 €	89 910 €	0 €	507 740 €
Čisté finanční prostředky	29 600 €	274 100 €	114 130 €	89 910 €	0 €	0 €
Výsledné cashflow	29 600 €	303 700 €	417 830 €	507 740 €	507 740 €	0 €

Tabulka 1: finanční náklady projektu HISPRA [MĚSTSKÁ, 2005c]

„Projekt byl plánován jako neziskový a v rámci jeho řešení nesměly být uživatelům výstupy tohoto projektu žádným způsobem speciálně zpoplatňovány“ [MĚSTSKÁ, 2005c].

2 Oddělení digitalizace

2.1 Předpoklady jeho založení

Městská knihovna v Praze vlastní unikátní sbírky vzácných a historických dokumentů, které jsou svou povahou jedinečné. Především se jedná o sbírku pragensií, dále o fond teatrální, hudební i uměnovědné literatury. Vzhledem k využívanosti těchto fondů je většina z nich ve špatném stavu, ať již stářím, pouhým opotřebením či degradací nekvalitního papíru a vlivy okolního prostředí. Nejvýhodnější metodou, jak tyto dokumenty zachovat a informace v nich poskytovat uživatelům, je pořízení digitální kopie těchto dokumentů.

Digitalizace se také stala jedním ze strategických směrů MKP a vyplývá z plnění regionálních funkcí krajské knihovny ze zákona č. 257/2001 Sb. a ze Strategického plánu hlavního města Prahy, kde je jedním z hlavních úkolů záchrana historického a kulturního dědictví“ [STRATEGICKÝ, 2000].

Vzhledem k malému počtu digitalizačních pracovišť na území České republiky (především Národní knihovna v Praze a Knihovna Akademie věd), které by pokryly digitalizaci alespoň nejohroženějších dokumentů², bylo třeba vymyslet jiné řešení. Tím se stalo vybudování vlastního digitalizačního pracoviště, které bude moci po skončení tohoto projektu pokračovat v záchraně dalších dokumentů³.

Výhodou vlastního digitalizačního pracoviště je také záruka šetrné manipulace nejen při samotné digitalizaci, ale také při návazných procesech (restaurování, zvláštní režim). Další výhodou vlastního pracoviště je přehlednost a možnost vlastního režimu digitalizačního

² Podle propočtů MKP by při využití kapacity digitalizačních pracovišť NK ČR a AV ČR, kterými disponovaly obě knihovny v roce 2006, trvala digitalizace tohoto fondu několik desetiletí.

³ Projekt HISPRA zahrnuje pouze dvacetinu vzácných a historických dokumentů ve fondu MKP.

procesu a usnadnění kontroly samotného procesu tvorby digitálního dokumentu, ale i návazných procesů.

Dalším předpokladem byla existence digitalizačního pracoviště přímo v Ústřední knihovně na Mariánském náměstí, která by zajišťovala velmi šetrnou manipulaci s dokumenty umístěnými převážně v budově Ústřední knihovny.

2.2 Vývoj

Jak již bylo uvedeno výše, digitalizační pracoviště mělo být umístěno v prostorách Ústřední knihovny MKP. V rámci reorganizace MKP došlo k přesunu několika oddělení a kvůli prostorovým nárokům na budoucí pracoviště digitalizace nebylo na Mariánském náměstí pro toto oddělení místo. Městská knihovna proto začala hledat náhradní prostory, kam by toto oddělení mohlo být umístěno. Jiné prostory MKP v okolí Mariánského náměstí nevyhovovaly a tržní ceny za pronájem v centru Prahy zase nebyly započítány do nákladů na projekt. Nakonec byly magistrátem hl. m. Prahy nabídnuty k pronájmu za velmi výhodných podmínek prostory bývalých kanceláří v Provaznické ulici, uvolněné po *Územním pracovišti Pražské správy sociálního zabezpečení* (viz Obrázek č. 1).



Obrázek 1: prostory v Provaznické ulici před rekonstrukcí

Poté již mohly začít nutné stavební úpravy – od elektroinstalace, pokládání podlah, výmalby až po zavedení vzduchotechniky. Vzhledem k tomu, že tento dům je památkově chráněný, nebylo možné nijak zasahovat do velikosti a počtu místností. Z tohoto důvodu také musela být nová elektroinstalace vedena v lištách při podlaze.

Nově vzniklé Oddělení digitalizace tak získalo dvě místnosti pro samotný proces digitalizace, místnost vedoucí oddělení, serverovnu, šatny a jiné sociální zázemí. Celá rekonstrukce a vybavení místností probíhalo od července do listopadu 2007, kdy začal zkušební provoz Oddělení digitalizace MKP. Na obrázku č. 2 je znázorněna podoba Oddělení digitalizace po rekonstrukci.



Obrázek 2: Vybavené Oddělení digitalizace po rekonstrukci (autor fotografie: Václav Jiroušek)

Třicátého prvního ledna 2008 bylo uspořádáno slavnostní otevření a oficiální zahájení provozu Oddělení digitalizace, kterého se zúčastnil také norský velvyslanec v ČR Peter Ræder. Na tomto setkání bylo poprvé představeno celé Oddělení digitalizace veřejnosti a vystoupil zde kromě zástupců Městské knihovny v Praze – ředitele RNDr. Tomáš Řeháka, vedoucí odboru knihovnických služeb PhDr. Zuzana Kopencové a vedoucí Oddělení digitalizace Mgr. Eva Měřínské, také Milan Richter, radní hl. m. Prahy pro oblast kultury, památkové péče a cestovního ruchu.

2.3 Organizační a personální zajištění

Nově vzniklé Oddělení digitalizace je v rámci organizační struktury MKP začleněno do *Odboru knihovních fondů* a spadá tedy přímo pod vedoucí odboru PhDr. Zuzanu Kopencovou, která je zodpovědná celkově za projekt, jeho implementaci a medializaci.

V samotném oddělení bylo vytvořeno 7 pracovních úvazků. „Vedoucí oddělení odpovídá za všechny činnosti spojené s digitalizací, indexováním, ukládáním v digitálním archivu a další činnosti spojené se zpřístupněním digitálních dokumentů“ [MĚSTSKÁ, 2005a].

Procesu digitalizace se věnují 2 zaměstnanci na plný úvazek a 8 zaměstnanců na zkrácený poloviční úvazek. Dalších 5 pracovníků působí v rámci projektu na dohodu o pracovní činnosti. Dvě pracovnice ze specializovaného úseku MKP ze Studovny pragensií zajišťují výběr a přípravu dokumentů k digitalizaci. V rámci ekonomického oddělení je v rámci jednoho úvazku zvlášť zpracováno ekonomické řízení projektu HISPRA a dva pracovníci z Oddělení informačních technologií se starají o údržbu technologií a celkovou technickou podporu Oddělení digitalizace⁴.

2.4 Organizace práce

Organizace práce se v Oddělení digitalizace naprosto liší od ostatních digitalizačních pracovišť v ČR. Zatímco například v *Digitalizačním centru AV ČR* se jednotliví zaměstnanci věnují pouze jedné části digitalizačního procesu – například pouze skenují či pouze zpracovávají naskenované stránky, v Oddělení digitalizace MKP byla zavedena jiná politika organizace práce: každý zaměstnanec si zpracovává svůj dokument od počátku až do konce.

Každý zaměstnanec je proto odpovědný za určitý dokument, čímž je zajištěna rychlá a efektivní kontrola. Pokud je například nalezena při kontrole práce chyba v počtu stran, je rychle objeven zaměstnanec, který dokument zpracovával a může jej přímo ve zdrojovém souboru opravit. Tímto způsobem je eliminována také možná jednotvárnost při vykonávání pouze určité části celého procesu digitalizace. Ten je velmi náročný na soustředění a prověřuje také trpělivost a pozornost každého zaměstnance. Proto je velmi důležité pomoci odbourávat možný rutinní způsob zpracování střídáním jednotlivých činností v průběhu pracovní doby.

⁴ Po ukončení projektu HISPRA jsou digitalizaci historických dokumentů věnovány v průměru 3 pracovní úvazky.

V oddělení je kladen velký důraz na kvalitu zpracování, která ovšem vyžaduje stoprocentní nasazení pracovníků a jejich kreativitu, která by mohla být při rutinní práci potlačena. Všichni zaměstnanci jsou proto vyškoleni na všechny typy činností vykonávaných při procesu digitalizace – přípravu dokumentů ke skenování, následné skenování, ořez dokumentů, zpracování skenů a indexování.

Pro efektivní výměnu informací byly pro kontrolování dokumentů zvoleny tzv. „kontrolorské dvojice“, kdy každý zaměstnanec má určeného kolegu, který kontroluje jím zdigitalizované dokumenty a stejný kolega zase kontroluje dokumenty vytvořené jím samotným. Kontrola je prováděna v SW pro zpracování dat Sirius ještě před importem digitalizovaných děl do digitální knihovny. Tak mohou být případné chyby odhaleny včas, ještě před zveřejněním dokumentu. Na počátku projektu byli vytipováni tři pečliví zaměstnanci, kteří se věnovali zpřístupňování dokumentů v digitální knihovně a následně kontrole těchto dokumentů. S narůstajícím počtem digitalizovaných dokumentů a se vzrůstající rychlostí zpracování (po zavedení rutinního provozu) bylo zjištěno, že počet tří pracovníků na kontroly není dostatečný, takže byl vytvořen systém kontrolorských dvojic.

Zpřístupňováním dokumentů v digitální knihovně Kramerius byl pověřen jediný zaměstnanec, který současně s importem digitální kopie řeší autorskoprávní otázky – tedy určuje, zda bude dílo v případě volných majetkových autorských práv volně k dispozici nebo bude v případě trvání majetkových autorských práv zpřístupněno pouze v síti MKP.

Navíc je třeba vykonávat další činnosti, které udržují v chodu celé oddělení – koordinaci a kontrolu toku knih z místa jejich původního uložení do OD MKP, administrativní agendu atd. Tyto činnosti jsou rozděleny mezi zaměstnance tak, aby každý vykonával kromě rutinní digitalizace také další aktivitu, čímž je podpořena diverzita pracovních činností jednotlivých členů oddělení.

2.5 Technické a programové vybavení

2.5.1 Pracovní stanice

Oddělení digitalizace je vybaveno osmi pracovními stanicemi, sedm počítačů je používáno jako běžná pracovní stanice, osmý z nich slouží jako server, na kterém probíhají dávkově procesy identifikace okrajů a následný ořez. Tyto počítače jsou vybaveny výkonnou grafikou, aby byly schopny následného zpracování datově objemných obrazových souborů

a doplněny o 21“ LCD monitor⁵ s technologií S-PVA⁶ zajišťující věrné podání barev. Osmý počítač je potom rozšířen o větší diskový prostor.

Během realizace projektu byla vyzkoušena také varianta pracovní stanice se dvěma monitory, která rozšiřovala možnosti zobrazení hlavně při manuálním kontrolování detekovaných okrajů stránek. Tato možnost ovšem nevyhovovala všem pracovníkům, a nakonec tak byly zachovány pouze dvě původně zkušební stanice s přídatným monitorem a k realizaci dalších se již nepokročilo.

2.5.2 OCR klíč

Pro potřeby optického rozeznání textu umožňující fulltextové vyhledávání slouží dva *OCR⁷ klíče*, jejichž možná kapacita je 50 000 vyčtených stránek textu za měsíc. Tyto klíče byly součástí nabídky firmy *Elsyst Engineering⁸* a vyčtené texty se v podobě textového souboru přímo importují do SW Sirius.

2.5.3 Skenery

Pro potřeby archivního uchovávání digitálních kopií dokumentů byly ve výběrovém řízení vybrány dva face-up knižní skenery, které jsou nejvhodnější pro knihy, které tvoří největší procento fondu vybraného k digitalizaci.

2.5.3.1 Zeutschel

Skener *Zeutschel OS5000 TT⁹* snímá předlohy pouze ve škálách šedi a to do velikosti formátu A2, maximální rozlišení¹⁰ závisí na velikosti skenované plochy, a to od 200 do 600

⁵ Displej z tekutých krystalů (angl. *Liquid Crystal Display*, zkratka LCD) je tenké a ploché zobrazovací zařízení skládající se z omezeného (velikostí monitoru) počtu barevných nebo monochromatických pixelů seřazených před zdrojem světla nebo reflektorem [Displej, 2011].

⁶ S-PVA (angl. *Super-Patterned Vertical Alignment*) je typ technologie využívány v LCD monitorech, LCD panely s technologií S-PVA mají lepší pozorovací úhly, kontrast a zobrazení černé barvy [KOVAČ, 2011].

⁷ OCR (angl. *Optical Character Recognition*) je metoda převodu tištěného nebo rukou či strojem psaného písma na strojem čitelný a tudíž zpracovatelný text [ELSYST, 2007b, s. F-2].

⁸ Elsys Engineering je firma se sídlem ve Vyškově, založená v roce 1991. Firma se zabývá především programováním grafických a databázových aplikací, například software Sirius určený pro práci s dokumenty a daty. Nabízí také digitalizaci na zakázku, služby v oblasti DTP a předtiskové přípravy a vývoj softwarových aplikací na zakázku [ELSYST, ©2003b; ELSYST ©2003d].

⁹ Skener *Zeutschel Omniscan OS5000TT* je produktem společnosti Zeutschel GmbH, která se věnuje speciálním digitálním a analogovým systémům pro ukládání dat pro dokumentaci a archivy [ZEUTSCHEL, [2007]].

¹⁰ Rozlišení (angl. *resolution*) je počet pixelů (nebo maximální rozlišení obrazu), které může být zobrazeno na zobrazovači, médiu. Často se udává jako počet sloupců (horizontálně, „X“), které se uvádí vždy jako první, a počet řádků (vertikálně, „Y“) [VYCHODIL, 2007, s. 95].

DPI¹¹. Tento přístroj je vybaven pohyblivou skenovací hlavou, kterou je nutno nastavit podle velikosti předlohy, dále je nutné nastavit clonu regulující světlost skenu a zaostření (viz Obrázek č. 3).



Obrázek 3: Skener Zeutschel OS5000 TT(autor fotografie: Eva Měřinská)

Velkou výhodou je velmi dobře řešená kolébka na knihy umožňující šetrné skenování knih s problémovou vazbou. Manipulaci s kolébkou zajišťují dva otočné mechanismy – jeden umožňuje nastavení jednotné polohy obou částí kolébky dohromady, druhý případný pohyb obou částí kolébky najednou. V případě objemných svazků je tak možné docílit toho, že kolébka je v určité vzdálenosti od přitlačného skla a současně je stále umožněna rozdílná vzájemná poloha obou částí kolébky, která se mění v závislosti na snímané stránce. Vzdálenost obou částí kolébky od sebe lze samozřejmě nastavit pouhým přitlačením, korekce vzdálenosti desek od sebe slouží k přizpůsobení tloušťce knižní vazby.

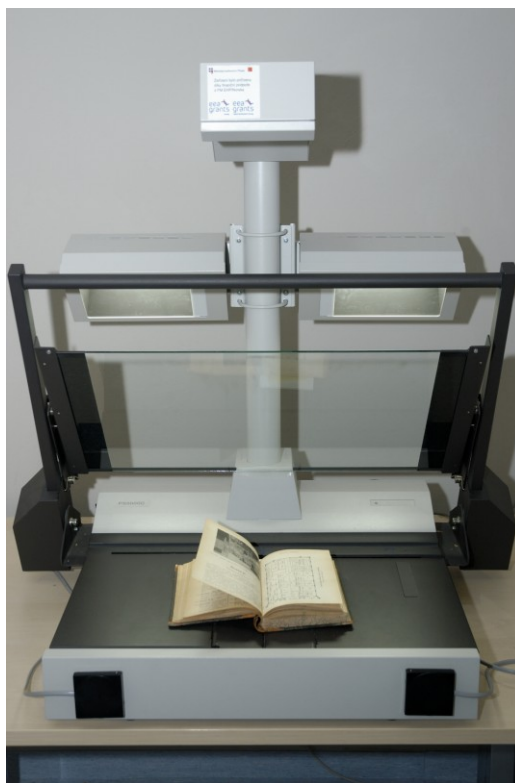
Skener včetně přitlačného skla je možné ovládat jak ručně, tak automaticky pomocí nožních pedálů. Skener se proto velmi dobře osvědčil při snímání velkoformátových knih a map.

¹¹ DPI (dots per inch = počet bodů na palec) je jednotka míry kvality zobrazení a tisku vyjádřená v počtu bodů, které je zařízení schopno zobrazit na šířku jednoho palce (2, 54 cm) [BALVÍNOVÁ, 2003-].

2.5.3.2 *Minolta*

Minolta *PS 5000 C*¹² je plnobarevný skener s automatickým ovládáním kamery a ručním nastavením jak kolébky, tak přítlačného skla. Skenovat lze formáty do velikosti A3 s maximálním rozlišením 600 dpi.

Oproti dobře řešené kolébce u skeneru Zeutschel je kolébka modelu Minolta tvořena jednou kolébkou, která se skládá z více navzájem pohyblivých částí, které jsou upevněny na pružinách. Bohužel při skenování dokumentů dochází kvůli předimenzování síly pružin k samovolnému posouvání kolébky – v případě objemnějších knihy do stran, naopak v případě tenčích svazků nahoru, a to samovolně i během procesu skenování. Přítlačné sklo je proto třeba preventivně přidržovat během skenovacího procesu a tím udržet kolébku a potažmo i předlohu ve stejné poloze (viz Obrázek č. 4).



Obrázek 4: Skener Minolta PS 5000 C (autor fotografie: Eva Měřínská)

Během práce s tímto skenerem ovšem byly zjištěny také problémy v kvalitě skenovaných obrazů, především s obrysovou ostrotí. Dalším problémem byla nestálost věrnosti barevného obrazu a občasné „rozsypání skenovaného obrazu“, kdy je chyba způsobena skládáním jednotlivých částí skenu během jeho ukládání. Jednou z příčin nestálosti věrnosti barevného

¹² Skener Minolta PS 5000 C je produktem společnosti Konica Minolta, která se věnuje oblasti tiskových a kopírovacích řešení, mimo jiné nabízí skenovací zařízení [KONICA, 2011].

obrazu je kombinace nerovnoměrně nasvícené předlohy způsobené statickou polohou světelných trubíc a nepříliš vhodných světelných podmínek vzhledem k celkové poloze pracoviště. Velká okna a přítomnost dalších zaměstnanců vykonávajících jiné činnosti vyžadující stálé světelné podmínky neumožňují regulovat světlo v místnosti tak, jak by bylo třeba, a přídatné lampy na tomto skeneru nestačily stálé podmínky udržet. Vzhledem k tomu bylo nutné velmi často kalibrovat barvy – především nastavit vyvážení bílé, přesto nebyly barvy věrně zachyceny. Věrnost zachycení barev byla velmi důležitá, vzhledem k tomu, že druhý skener Zeuschel umožňuje pořizovat obrazy dokumentu pouze ve stupních šedi.

Poté, co byly zjištěny tyto chyby, byla kontaktována další pracoviště s otázkou, zda se u nich tyto problémy také vyskytují. Bylo nutné zjistit, zda se především u obrysově ostrosti jedná o ojedinělou vadu pouze tohoto přístroje nebo zda se jedná o vlastnost celého modelu. Většina institucí, která tento přístroj vlastní, jej používá k jinému účelu, tudíž nemohly dát jasnou odpověď. Po testech provedených na ostatních pracovištích v Praze se ukázalo, že označené problémy se v různé míře objevovaly u všech skenerů. Oba tyto problémy byly výrobcem označeny za charakteristickou vlastnost skeneru a vedly nakonec k rozhodnutí vyměnit tento skener, který se ukázal jako nevyhovující z hlediska pořizování snímků archivní kvality. Po dlouhých jednáních byl tento přístroj nakonec vyjmut z finančního plánu a uvolněné finance mohly být použity na nákup novějšího a kvalitnějšího skeneru.

2.5.3.3 Výběr zařízení jako náhrada za skener Minolta

Za téměř rok od nákupu skeneru Minolta se český trh rozrostl o další modely a došlo k rapidnímu snížení pořizovacích cen – díky tomu si knihovna mohla dovolit vybírat z větší nabídky skenerů než na počátku projektu. Nejprve byly vybrány v rámci českého trhu modely, které se nejvíce hodí pro potřebu Oddělení digitalizace. Vzhledem k předchozím zkušenostem se skenerem Minolta bylo rozhodnuto, že tyto modely skenerů budou podrobeny vlastnímu testování, aby se zabránilo případným pozdějším komplikacím. Vzhledem k tomu, že v té době již probíhal rutinní provoz na pracovišti digitalizace, mohly požadavky na parametry skeneru vycházet také z praktických zkušeností s obsluhou skeneru. Pro účely skenování v OD bylo nutné pořídit plnobarevný skener, který by vhodně doplnil černobílý model.

Díky dobrým zkušenostem se skenerem Zeuschel, jehož hloubka ostrosti odpovídá požadavkům na archivaci, se do užšího výběru dostal plnobarevný model skeneru *Zeuschel*

*OS 12000 HQ*¹³ a přístroj *Bookeye 3 A2*¹⁴ – oba tyto skenery jejich dodavatelé umožnili vyzkoušet přímo na pracovišti digitalizace. Bylo tak možné posoudit vlastnosti přístroje v závislosti na prostředí a detailně prozkoumat různé vlastnosti obou skenerů a na místě je porovnat. Bylo vyhlášeno nové výběrové řízení a na základě již praktických zkušeností s různými typy skenerů byl vybrán model *Bookeye 3 A2*. Jedním z parametrů, které tento skener splnil, byly právě pohyblivé lampy, které opravdu velmi dobře osvětlí plochu snímání. Argumenty pro tento přístroj byly také podpořeny tím, že se jedná o jiného výrobce a na pracovišti se tak ocitnou různé typy přístrojů, jejichž vlastnosti se mohou doplňovat.

2.5.3.4 Bookeye

Skener *Bookeye 3 A2* je samozřejmě plnobarevný skener, který snímá dokumenty až do formátu A2, a to s maximálním rozlišením 600 dpi. Tento skener je vybaven kvalitními *CCD senzory*¹⁵ s vysokým rozlišením. Jeho velkou výhodou jsou dvě *osvitové LED lampy*, které jsou umístěny vhodně po stranách skeneru a které si velmi dobře poradí s rozdílnou intenzitou okolního osvětlení. Tím, že je použita technologie tzv. studeného LED osvětlení, která nevyzařuje UV a IR záření, lampy jsou aktivní pouze při snímání předlohy a osvětlují pouze vymezený (nastavený) prostor, nedochází při práci se zařízením k oslňování. „Zajímavostí je, že světlený paprsek se pohybuje po předloze spolu s procesem snímání“ [MĚŘÍNSKÁ, 2008].

¹³ Skener OS 12000 HQ je produktem společnosti Zeutschel GmbH, která se věnuje speciálním digitálním a analogovým systémům pro ukládání dat pro dokumentaci a archivy [ZEUTSCHEL, 2011].

¹⁴ Skener Bookeye 3 A2 je produktem společnosti ImageWare Components GmbH, která se zabývá komplexním řešením pro digitalizaci a ochranu kulturního dědictví [IMAGEWARE, 2011].

¹⁵ CCD (angl. Charge-Coupled Device) je zařízení s vázanými náboji, je elektronická součástka používaná pro snímání obrazové informace [VYCHODIL, 2007, s. 31].



Obrázek 5: skener Bookeye 3 A2 (autor fotografie: Václav Jiroušek)

Skener je vybaven motorizovanou kolébkou s ručně nastavitelnou šířkou hřbetu a ručně ovládaným přitlačným sklem (viz Obrázek č. 5). Výhodou je možnost manipulovat s kolébkou i po spuštění přitlačného skla, což usnadňuje proces skenování různých typů dokumentů – např. umožňuje šetrné rozevření dokumentu s tužší vazbou a kvalitnější nasnímání tisku hluboko do hřbetu. Celý proces skenování je usnadněn také automatickou detekcí vzdálenosti předlohy od snímače pomocí laseru. Skener je možné ovládat jak pomocí dodaného nožního pedálu, tak tlačítka umístěnými na přístroji nebo dodaným *softwarem* BSC-2¹⁶. Na rozdíl od skeneru Zeutschel nebo Minolta je přední deska přístroje vybavena několika okruhy ovládacích prvků, takže je možné rychle a pohodlně měnit parametry skenování – velikost snímané plochy, barevnou škálu použitou při skenování a jas.

2.5.4 Servery

Také na servery bylo vydáno výběrové řízení, jehož vítězem se stala společnost *Lamals*¹⁷, která dodala jak jednoprocesorový datový server s procesorem Pentium 4 a 30 GB diskem, stejně jako dvouprocesorový server na zpracování obrázků. Celkem má oddělení digitalizace k dispozici 3 servery rozmístěné v různých lokalitách s celkovou kapacitou 30 TB, jeden slouží jako datové úložiště a další dva jsou aplikačními servery.

¹⁶ Software BSC-2 určený k nastavení skenovacího procesu je produktem společnosti ImageWare Components GmbH.

¹⁷ Společnost Lamals se zabývá prodejem a servisem výpočetní techniky a příslušenství [LAMALS. 2006].

2.5.5 Software

Ke každému ze skenerů nabízel výrobce také speciální software umožňující jednoduše upravovat parametry skenování a případnou úpravu naskenovaných obrázků. Skener Zeuschel je ovládán programem *Omniscan 11*, ke skeneru Minolta byl dodán *Photoshop Adobe* a ke skeneru Bookeye je možné využít dodávaný SW *BCS-2*.

Pro zpracování a archivaci dat byl vybrán modulový program *Sirius* od fy Elsys Engineering, který v sobě jako jediný na trhu zahrnoval zpracování všech procesů digitalizace – lze v rámci něj skenovat, upravovat naskenované obrázky (včetně ořezu dokumentů), indexovat jednotlivé stránky, vkládat technická a administrativní metadata¹⁸ a archivovat vytvořené digitální kopie dokumentů. V rámci SW *Sirius* byly pořízeny dva *OCR klíče*, které rozšířily využitelnost *Siria* o možnost vyčtení textu.

Na úpravu metadat byl zvolen program *XMetal*, který umožňuje také validaci dat. Podrobnější informace o jednotlivých softwarech budou následovat v dalších kapitolách vždy u popisu jednotlivých procesů digitalizace.

2.6 Spolupráce s jinými subjekty

Ještě než vzniklo Oddělení digitalizace, byla navázána spolupráce s *Digitalizačním centrem Knihovny Akademie věd* (dále v textu také *KnAV*) – právě toto centrum posloužilo svým vybavením, personálním zajištěním a digitalizačními postupy jako inspirace pro Oddělení digitalizace. Dále OD velmi úzce spolupracuje s *Národní knihovnou České republiky* (dále v textu také *NK ČR*) a jejím digitalizačním pracovištěm, hlavním styčným bodem je používání stejného softwaru pro zpracování dat (*Sirius*), také toto centrum bylo především v počátcích při psaní projektu velkou inspirací. S *NK ČR*, *KnAV* a *Moravskou zemskou knihovnou* (dále v textu také *MZK*) pojí OD také skutečnost, že všichni využívají pro zpřístupňování svých digitálních kopií open source SW *Kramerius*, na jehož vývoji se tyto tři instituce spolu s firmou *INCAD* podílí.

¹⁸ Metadata jsou strukturovaná data, která nesou informace o primárních datech [CELBOVÁ, 2003d-].

3 Předmět digitalizace

Již od počátku projektu bylo jasné, že digitalizace se týká většího množství titulů, než bude možné do projektu zahrnout. Proto byly po mnoha konzultacích k záchraně pomocí procesu reformátování vybrány historické tituly uložené ve třech specializovaných úsecích MKP: ve Studovně pragensií, v Divadelním a filmovém úseku a nakonec v Hudebním úseku. Největší množství poškozených titulů, historicky nejstarší a co do objemu svazků nutných k digitalizaci největší fond je uložen v úseku pragensií. Tento fond se také zasloužil o název celého projektu. Bylo proto rozhodnuto, že historický pragensijní fond bude digitalizován přednostně a zahájí tak druhou a třetí etapu projektu. Z celkového počtu zdigitalizovaných titulů vyplynulo, že počet titulů v historickém pragensijním fondu se vyšplhal na necelé dva tisíce dokumentů, počet titulů z fondu hudebního a filmového úseku se ustálil na čísle tisíc dokumentů a fond hudebního úseku dosáhl pouze počtu 180 dokumentů.

První část kapitoly představuje všechny úseky, jejich historii, služby, které poskytují, a podává informace o celém fondu. V další části je v krátkosti představen historický fond těchto úseků určený k digitalizaci. U pragensijního úseku jsou uvedeny pouze informace o Studovně pragensií a je představen historický pragensijní fond určený k digitalizaci, podrobná specifikace celého historického pragensijního fondu je poté popsána v následující kapitole.

3.1 *Fond Divadelního a filmového úseku ÚK MKP*

Fond tohoto úseku byl vybrán k digitalizaci jako druhý fond po historickém pragensijním fondu.

Fond *Divadelního a filmového úseku* tvoří:

- divadelní a filmové encyklopedie, české i cizojazyčné dějiny divadla a filmu, publikace o teorii divadla a filmu, herectví, divadelní a filmové režii, kostýmnictví, scénografii a další specifická literatura s obory souvisejícími
- historicky cenná vydání divadelních her a publikací z dějin divadla a filmu
- úplná produkce divadelních her z vydavatelství DILIA od roku 1949
- knižní vydání divadelních, loutkových a rozhlasových her, filmových a televizních scénářů

- divadelní a filmové časopisy, divadelní ročenky a kalendária z 19. a 20. stol.
- cizojazyčná literatura (divadelní hry, teorii divadla a filmu, bohatá sbírka publikací o anglickém dramatu převzatá z fondu knihovny Britské rady) [MĚSTSKÁ, 2005a, s. 1].

Historický fond tohoto úseku tvoří 60 000 svazků z oborů divadla, filmu, rozhlasu, televize, kostýmnictví a scénografie. Vzácný fond pak tvoří několik částí, první část fondu tvořící velké procento signatury B je uložena v prvním patře na galerii, vzácné tisky od 17. století do 1. pol. 20. století jsou uloženy v Oddělení vzácných tisků na Brusnici [Skladba, 2011].

3.1.1 Divadelní a filmový úsek ÚK MKP

Divadelní a filmový úsek tvoří fond divadelní a filmové knihovny určené pro laickou i odbornou veřejnost. Byl založen 1. 11. 1942 díky péči *Jaroslava Tumlíře* a *Artura Salavy* a od roku 1982 funguje jako samostatný specializovaný úsek MKP. Úsek sídlí v přízemí Ústřední knihovny MKP a je ukázkou původního díla architekta Roitha [MĚSTSKÁ, 2005a, s. 1].

Uživatelskou základnu tvoří zájemci o divadlo, film, rozhlas, televizi, kostýmnictví a scénografii, dále dramaturgové, režiséři, divadelní a filmoví historici a publicisté a herci.

V rámci svých služeb poskytují pracovníci divadelního a filmového úseku informace z oborů divadla, filmu, rozhlasu, televize, kostýmnictví a scénografie (kartotéky článků z odborného tisku, kartotéky divadelních monologů pro muže a ženy), a to i pomocí lístkových katalogů (katalogu osobností, předmětového, jmenného, názvového, tematického, cizojazyčných her, loutkových her, her pro děti a mládež, rozhlasových her, filmových scénářů, divadelních a filmových časopisů), mezi další služby patří prezenční půjčování divadelních a filmových časopisů a zpřístupňování divadelních, filmových a televizních recenzí z výstřížkové služby [Tamtéž, s. 1].

3.1.2 Dokumenty divadelního a filmového úseku určené k digitalizaci

Při výběru monografií a periodik byl kladen důraz na to, aby výběr zohledňoval zejména díla týkající se dějin českého divadla a filmu, a vycházel z existence historického fondu MKP [MĚSTSKÁ, 2006a, s. 1].

K digitalizaci tak bylo vybráno několik vrstev dokumentů z celého historického fondu.

V rámci monografií byly vybrány tři následující vrstvy:

1. První vydání divadelních her klasiků českého dramatu od 18. do 20. století, odhadem asi 10 000 stran formátu A5 a 5 000 stran formátu A4.
2. Filmové scénáře klasických děl z 30. až 50. let 20. století, jejichž množství bylo odhadnuto na 5 000 stran.
3. Odborná literatura z oborů divadlo, film a scénografie od 18. do 20. století v množství asi 50 000 stran formátu A5 a 6 000 stran formátu A4.

Druhou rovinu titulů určených k digitalizaci tvoří periodika:

1. Padesát tři titulů divadelních časopisů s rozmezím let vydávání od roku 1866 do roku 1970¹⁹, například: *Almanach královského českého ND*, *České divadlo*, *Divadelní svět*, *Loutkář* atd.
2. Devět titulů filmových časopisů s rozmezím let vydávání od roku 1927 do roku 2002²⁰, například: *Filmová práce*, *Kino*, *Illustrierten Film.Kurier* atd. [Tamtéž, s. 1-2].

Na konci projektu HISPPRA bylo digitalizováno více než 1 000 svazků (divadelní hry, divadelní časopisy, filmové scénáře), z toho 188 svazků vydaných do roku 1900, které byly zařazeny do evropského divadelního projektu Europeana [Skladba 2011].

3.2 Fond hudebního úseku

Dokumenty z tohoto fondu byly ponechány jako poslední z titulů vybraných k digitalizaci.

Fond hudebního úseku tvoří:

- fond hudebnin (kolem 150 000 sv.)²¹ ve formě partitur, klavírních výtahů a jednotlivých sólových partů

¹⁹ V tomto případě bylo rozhodnuto nedodržet časovou linii historických dokumentů stanovenou pro účely této práce na rok 1950. Daný divadelní titul periodika má svou cenu jako ucelený soubor vydaných čísel, u kterého není možné od sebe odlišit a trvat na umělé hranici dané rokem 1950. Titulů, které překročily vydáním některého ze svých čísel hranici roku 1950, jsou tyto: *Acta scaenographica* (1961-1971), *Divadelní revue* (1989-2005), *Divadlo* (1903-1914, 1920-1947, 1949-1970), *Městská divadla pražská* (1940, 1951-1964), *Národní divadlo* (1930-1962), *Naše divadlo* (1945-1951), *Naše loutky* (1945-1951), *RDZN* (1958-1973), *Repertoár malé scény* (1963-1971), *Svět a divadlo* (1990-2005) a *Umělecký měsíčník* (1948-1951).

²⁰ I zde platí stejné pravidlo jako pro historické tituly divadelních časopisů.

- specializovaný fond naučné, populární, referenční i zábavné literatury o hudbě v češtině i v hlavních světových jazycích, kolem 30 000 knižních svazků, např. hudební encyklopedie, dějiny české a světové hudby, životopisy o skladatelích či interpretech atd.
- starší i současné tituly specializovaných hudebních časopisů: 33 českých titulů a 8 zahraničních titulů a více než 100 titulů (reprezentovaných cca 2 350 svazky) ucelených ročníků vázaných historických časopisů
- audio fond – CD, magnetofonové kazety, gramofonové desky (asi 20 000 nahrávek)
- audiovizuální fond – DVD, BluRay [MĚSTSKÁ, 2005b, s. 1].

3.2.1 Hudební úsek MKP

Hudební úsek Ústřední knihovny MKP je největší a nejstarší veřejná hudební knihovna, která se již více než sto let zaměřuje na získávání, shromažďování a půjčování hudebních materiálů. Hudební oddělení bylo zřízeno již dva roky po založení první Veřejné obecní knihovny královského hlavního města Prahy v roce 1983. V roce 1922 se na výzvu tehdejšího ředitele knihovny *PhDr. Jana Thona* ujal vedení Hudebního oddělení *JUC. Josef Plzeňský*, za jeho vedení bylo oddělení v souvislosti s blížícím se 100. výročím narození skladatele *Bedřicha Smetany* nazváno *Knihovnou Bedřicha Smetany*. V roce 1998 po otevření zrekonstruovaných prostor ÚK se úsek přestěhoval do nově řešeného interiéru, úsek nyní sídlí v přízemí a na galerii v prostorách vlevo od hal služeb v ÚK [ČERMÁKOVÁ, 2003; MĚSTSKÁ, 2005b, s. 1].

Tento specializovaný úsek nabízí půjčování hudebních materiálů (hudebniny, knihy, audio), poslech hudby různých žánrů a mluveného slova na CD, magnetofonových kazetách a gramofonových deskách a možnost přehrávání zapůjčených notových materiálů na digitálním pianě Yamaha [MĚSTSKÁ, 2005 b, s. 1].

²¹ V tomto fondu jsou zastoupeny: noty pro sólové nástroje, notové materiály pro tradiční i netradiční komorní ansámby, partitury komorních, orchestrálních i vokálně-orchestrálních děl, klavírní výtahy oper, operet, muzikálů a dalších vokálních a instrumentálních skladeb, vokální literatura pro všechny hlasové obory a pro sólový i sborový zpěv, zpěvníky lidových písní a písní pro děti, populárních písní českých i zahraničních interpretů a hudebních skupin.

3.2.2 Dokumenty z hudebního fondu určené k digitalizaci

K digitalizaci byly vybrány různé vrstvy fondu. První vrstvu tvoří jedinečné monografie v rámci „fondu uchovávaných v knihovnách v České republice a zároveň tituly akutně ohrožené kvůli špatné kvalitě tisku i papíru [MĚSTSKÁ, 2006b, s. 1]“. Tuto vrstvu tedy tvoří původní česká nebo do češtiny přeložená libreta oper a operet a hudebních scénických útvarů z konce 19. století a z první třetiny 20. století, v počtu přibližně 200 kusů, odhadem asi 2 500 – 3 000 stran formátu A5.

Další vrstvu tvoří periodika, především všechny význačné nebo v jiných knihovnách či kulturních institucích nedostupné tituly hudebních časopisů vycházejících od 60. let 19. století a tituly česky a německy tištěných časopisů z období první republiky a z doby 2. světové války do roku 1948, odhadem asi 60 000 stran formátu A4 či A5. Příkladem jsou tituly *Česká hudba*, *Dalibor*, *Der Auftakt*, *Hudební revue* atd.

Poslední vrstvu tvoří hudebniny, pro něž zatím ani v České republice, ani ve světě neexistuje standardizovaný formát pro přepis různých druhů tištěných hudebnin. Z důvodů ochrany a jejich dalšího budoucího zpřístupnění bylo pro začátek vybráno několik desítek titulů české hudební provenience z fondu vzácných hudebních tisků. Kromě vzácných tisků pak fond hudebního úseku obsahuje asi 50 000 titulů z let 1880 až 1930, řada z nich patří k vzácnějším bohemikálními hudebninám.

V průběhu projektu HISPRA bohužel došlo ke komplikacím ve výběru a zpracování titulů z hudebního fondu, celý proces digitalizace tak byl opožděn téměř o jeden rok, místo hudebního fondu tak byly digitalizovány další dokumenty z divadelního a filmového úseku. V době určené a vyhrazené pro realizaci projektu HISPRA tak byla zdigitalizována část libret a některé již zpracované tituly periodik. Digitalizace hudebního fondu tak pokračuje i po oficiálním skončení projektu HISPRA v létě roku 2010, byla digitalizována další libreta a zpracovány další tituly periodik. Digitalizace již ovšem postupuje pomaleji, Oddělení digitalizace se již plně nevěnuje pouze digitalizaci historických titulů a v hudebním úseku již nejsou vyčleněni zaměstnanci na zpracování titulů a zpracování titulů tak probíhá v rámci běžné pracovní náplně.

3.3 *Pragensijní fond*

Fond Studovny pragensíí tvoří přibližně 7 000 svazků titulů knižní povahy, 200 plánů a map, více než 1 000 svazků vázaných periodik (celopražský deník *Večerní(k) Praha*, tisk

obvodních, městských a místních úřadů a dalších regionálních vydavatelů, částečně odborné časopisy), 50 CD-ROM a cca 2 000 obálek s pragensijními výstřižky [MĚSTSKÁ, 20??, s. 2]. Podrobnější informace o pragensijním fondu jsou uvedeny v kapitole 4.2 Vymezení profilu pragensijního fondu.

3.3.1 Studovna pragensií

Pragensia byla v Městské knihovně v Praze součástí fondu, teprve později byla vyčleněna a již několik desítek let tvoří samostatný úsek. O založení úseku pragensií uvažoval již ve 30. letech 20. století tehdejší ředitel MKP *Jan Thon* v souvislosti s chystaným katalogem topografické literatury. Myšlenku vyčlenění fondu zrealizoval o 30 let později tehdejší ředitel *Dr. Rudolf Málek*. Studovna pragensií byla slavnostně otevřena v budově Ústřední knihovny na Mariánském náměstí v roce 1962, při té příležitosti byl uspořádán literární večer, na kterém účinkoval například *František Hrubín* či *Rudolf Deyl starší* [POMIKÁLKOVÁ, 2008, s. 30]. Studovna se během své existence několikrát stěhovala, v roce 1965 byla přemístěna do větších prostor v domě č.p. 21 v Mostecké ulici, v roce 1981 se stěhovala jen naproti přes ulici do domu č.p. 40. V roce 1998 se Studovna pragensií vrátila na Mariánské náměstí, kde sdílí prostory s Informačním střediskem.

Rozsah fondu je uzpůsoben potřebám specializovaného pracoviště typu studoven zřízeného ve veřejné knihovně a určeného pro odbornou i laickou veřejnost [MĚSTSKÁ, 2002b, s. 2]. Prezenční režim půjčování tak „plně odpovídá charakteru pracoviště v podstatě s konzervační funkcí [Tamtéž s. 3]“.

Velká část fondu je umístěna ve volném výběru, tudíž umožňuje lepší a samostatnější přístup uživatelů. Bohužel se zvýšilo riziko poškození fondu a krádeží. Z tohoto důvodu jsou vzácnější tituly umístěny v příručním skladu a jsou dostupné pouze na vyžádání. Přes tato opatření docházelo a dochází k poškozování fondu uživateli, tudíž je digitalizace historických pragensijních dokumentů jediným východiskem z této situace – dokumenty zůstanou ochráněny a jejich obsah může být zpřístupněn uživatelům [Tamtéž, s. 2-3].

Studovna pragensií poskytuje metodické a poradenské služby nejen uživatelům, ale také knihovnám v síti MKP a místním knihovnám v Praze. „Seznamuje pracovníky knihoven o fondu a službách studovny, poskytuje informace o věcném zpracování dokumentů v elektronickém katalogu a o vyhledávání pragensijní literatury v katalogu. Poskytuje možnost konzultace odborných pragensijních dotazů. Vypracovává bibliografie dokumentů pro místní

knihovny o jejich lokalitě a upozorňuje je na publikace týkající se jejich místa působení“ [POMIKÁLKOVÁ, 2008, s. 40].

3.3.2 Dokumenty z historického pragensijního fondu určené k digitalizaci

Dokumenty určené k digitalizaci tvoří i v tomto případě několik vrstev, podkladem pro tuto část textu je příloha k podávanému grantu na projekt HISPra [MĚSTSKÁ, 2006c, s. 1].

První vrstvu tvoří vybrané vzácné pragensijní tisky vydané do konce 19. století v češtině, jedná se o 56 svazků různých formátů, převážně A5. Celkem podle odhadu činí přibližně 9 000 stran, např. *Bartošova Kronika pražská od léta páně 1524*, *Pamětihodnj dnowé Prahy*, *Horopisné obrazy okolí pražského* atd.

Druhá vrstva je tvořena vzácnými tisky vydanými do r. 1860, jako hlavní jazyk dokumentů převažuje němčina, dále latina a čeština. Jedná se o 202 svazků různých formátů, převážně A5, podle odhadu je celkový počet stran roven přibližně 25 000.

Třetí vrstvu reprezentují pražské adresáře, ukazatele domů a orientační knihy z let 1859 až 1940. Jedná se o velmi specifický druh pragensijních dokumentů, hlavní charakteristikou je zvláštní struktura dokumentu, která je odlišuje od klasických monografických děl. Celkem bylo vybráno 47 svazků různých formátů, A4 a A5 i menší, celkový počet stran byl odhadnut na 18 000 stran, např. *Adressář královského hlavního města Prahy*, *Ukazovatel domů v městě Smíchově 1912*, *Chytilův adresář hl. města Prahy* atd.

Čtvrtá vrstva představuje pragensijní periodika z konce 19. století a první poloviny 20. století. Byly vybrány tituly, které se neobjevili v seznamech digitalizovaných děl v rámci projektu VISK č. 7 – Kramerius. Celkem se jedná o 30 svazků převážně formátu A4 i A3, dohromady bylo odhadnuto asi na 9 800 stran, např. *Naše Praha*, *Za starou Prahu* atd.

Pátou vrstvu tvoří periodické pragensijní publikace, které se svým obsahem vztahují k pražské městské správě a její činnosti. I zde byly pečlivě vybrány tituly, které již nebyly digitalizovány v rámci projektu VISK č. 7 – Kramerius. Jedná se o 81 svazků převážně formátu A4 a A5 s celkovým počtem stran asi 32 000, například *Administrativní zpráva obce královského hlavního města Prahy*, *Almanach hlavního města Prahy 1922*, *Statistická příruční knížka královského hlavního města Prahy* atd.

Šestá vrstva zastupuje staré pražské mapy vydané v 19. a počátkem 20. století. Celkem bylo vybráno 24 plánů formátu A3 a většího, např. *Mapa okresu smíchovského v Pražském*

*kraji, Polohopisný plán Malé Strany a Hradčan, Rozhled s Petřína v Praze s vrchole
Rozhledny atd.*

Předposlední vrstvu tvoří zajímavé a někdy unikátní pragensijní dokumenty 1. poloviny 20. století. Většina z nich je vtištěna na méně kvalitním papíře než výše jmenované historicky starší dokumenty. Odhadem by se mohlo jednat o 200 knih s přibližně 50 000 stranami, např. *Památník města Královských Vinohradů, Posvátná místa království Českého, Základy starého místopisu pražského atd.*

Poslední vrstva reprezentuje starší vybrané výstřižky z periodik, především z denního tisku, jejichž kvalita již není vhodná na běžné půjčování čtenářům. Počet stran byl odhadnut na 2 000 stran formátu A5 a A4.

V rámci projektu HISPRA byly digitalizovány nejstarší tituly, tedy vybrané vzácné pragensijní tisky vydané do konce 19. století v češtině a vzácné tisky vydané do r. 1860, dále pražské adresáře a vybrané periodické pragensijní publikace. Poslední vrstva, která byla digitalizována v rámci projektu HISPRA jsou pragensijní dokumenty 1. poloviny 20. století, jejichž stav byla často rozsahem poškození horší než některé starší tisky do roku 1860.

Na digitalizaci ještě čekají vybraná pragensijní periodika a staré pražské mapy a plány. V Oddělení digitalizace je možné převést do digitální podoby pouze některé z map a plánů, jejichž velikost nepřesahuje A2. Na větší plány není v OD dostupný vhodný skener, tyto plány a mapy tak bude nutné pravděpodobně naskenovat na zakázku v některém z jiných digitalizačních pracovišť. Poslední vrstva dokumentů, tedy výstřižky z periodik je zatím v plánu digitalizace přesunuta na pozdější dobu. I přes unikátnost této sbírky byl dohodnut odklad digitalizace, jejíž průběh bude v případě výstřižků z periodik komplikovaný.

4 Specifikace fondu historických pragensijních dokumentů

Fond historických pragensijních dokumentů byl vybrán jako hlavní ze všech fondů Městské knihovny pro digitalizaci. Tato kapitola by měla zodpovědět hlavní otázky, proč tomu tak bylo.

Hlavním důvodem je unikátnost této sbírky daná celistvostí a pečlivostí výběru a akviziční politiky. Pragensijní fond je svou podstatou fondem regionálním. Městská knihovna v Praze se snaží zachytit jako veřejná knihovna a současně krajská knihovna s regionální působností²² svou regionální literaturu. Definice pojmu *region* a *regionální literatura* jsou různorodé, liší se především úhlem pohledu na tyto pojmy. Podle L. Lederbuchové [2002, s. 275-276] je *regionální literatura* „literární tvorbou čerpající látku z určitého kraje nebo spjatá s ním autorovým původem či místem pobytu“.

Region Praha je v případě pragensijní literatury ovšem regionem specifickým. Významnost tohoto regionu tkví v tom, že Praha je hlavním a největším městem České republiky, tedy významným kulturním, politickým a správním centrem. Praha byla také svědkem důležitých historických událostí, které se výrazně dotýkaly českého národa. A velký význam má Praha také jako turistický cíl, metropole byla v roce 2007 dokonce na 7. místě nejnavštěvovanějších měst Evropy [Praha, 2001-]. Vzhledem k tomu, jak významný region se fond snaží obsáhnout, je jeho přesah výrazný.

V této práci je vzhledem k povaze titulů určených k digitalizaci probírána pouze část tohoto fondu, a to fond historický pragensijní. Pojetí historického fondu se také různí. Výklady a především časové rozmezí vymezující historický fond jsou velmi rozdílné. Pro účely knihovního zákona²³ byl historický fond vymezen takto: „jedná se o knihovní fond sestávající z knihovních dokumentů, které vznikly do roku 1860 nebo mají v daném oboru pro svou jedinečnost historickou hodnotu, popřípadě jiný knihovní fond mající zvláštní historickou a kulturní hodnotu“ [ČESKO, 2001, § 2, odst. e]. Pro účely této práce je významná především druhá část výkladu historického fondu, neomezená datací. Historický pragensijní fond totiž zahrnuje nejen dokumenty vzniklé před rokem 1860, ale také další dokumenty z 19. století, které se svou kulturní a historickou hodnotou mohou zařadit pod pojem historická pragensia. Pro účely této diplomové práce byla datace historického

²² Od 1. ledna 2002 se MKP stala na základě knihovního zákona č. 257/2001 Sb. krajskou knihovnou s regionálními funkcemi pro Hlavní město Prahu.

²³ zákon č. č. 257/2001 Sb.

fondu pragensií omezena rokem 1950. Toto datum bylo zvoleno záměrně, aby bylo možné vyčlenit a porovnat historický pragensijní fond s fondem určeným k digitalizaci. Historické pragensijní dokumenty určené k digitalizaci byly vybírány do projektu HISPRA především z důvodů špatné kvality papíru vlivem klimatických podmínek, zejména špatného prostředí, ve kterém byly uloženy. Tato skutečnost se masivně týká dokumentů datovaných do roku 1950, pozdější tituly již není nutné masově digitalizovat. Zde byly vybrány některé tituly ve špatném fyzickém stavu daném především častou manipulací s dokumentem a jeho opotřebením. Za historické pragensijní dokumenty je tedy možné považovat všechny dokumenty vydané před rokem 1950 včetně. Pragensijní fond má totiž velkou hodnotu právě díky historii budování tohoto fondu, který byl velmi pečlivě revidován, poslední velká revize proběhla před nastěhováním fondu do prostor Ústřední knihovny v roce 1990. Díky této revizi byly vyčleněny především multiplikáty a dokumenty nevyhovující z obsahového hlediska. Rok 1950 vymezující tento fond byl zvolen také z důvodů dostatečného nadhledu při revizi fondu, který umožňuje podle mého názoru pouze dostatečně dlouhé časové období. Náhled nutný při utváření celkové hodnoty fondu je podle mého názoru dostatečně zajištěn obdobím 40 let, takže může být dodržena celková koncepce budování fondu. U novějších titulů vydaných po roce 1950 již podle mého názoru není možné určit hodnotu, která je většinou prověřena až delším časovým úsekem. Díky revizím je udržována celistvost pragensijního fondu a odpovídající šíře pohledu, která činí tento fond natolik unikátním, že je možné jej díky zvláštní historické a kulturní hodnotě nazývat fondem historickým.

4.1 Pojem pragensií a pragensijní literatury

Aby mohlo dojít k představení historického pragensijního fondu, který je v současné době uložen jak ve Studovně pragensií, tak v Oddělení vzácných tisků, je třeba vysvětlit pojem *pragensií*.

Jaké dokumenty se tedy dají označit jako *pragensia*? V tomto případě je možné vycházet z dělení regionální literatury. Regionální literaturu můžeme dělit podle vztahu k regionu do tří hlavních skupin:

- z obsahového hlediska tvoří regionální dokumenty díla pojednávající o regionu celým svým rozsahem (monografie) nebo jen z části (studie ve sbornících, články v seriálech, hesla v encyklopediích apod.)

- z územního hlediska tvoří regionální dokumenty díla vydaná na území regionu bez ohledu na jejich obsah. „Dokumenty vydané na území regionu jsou cenným pramenem pro historický, politický, hospodářský a kulturní výzkum regionu. Podle tohoto formálního hlediska se považují za regionální dokumenty kompletní produkce nakladatelství a tiskáren působících na území regionu” [BÍNOVÁ, 1992, s. 4]
- z autorského hlediska tvoří regionální dokumenty díla těch autorů, kteří jsou s regionem spjati svojí činností nebo původem (buď se v regionu narodili, žili nebo v něm působili) [POMIKÁLKOVÁ, 2008, s. 9].

Analogicky můžeme *pragensia* neboli *pragensijní dokumenty* rozdělit do tří skupin:

- z obsahového hlediska tvoří *pragensia* dokumenty pojednávající o Praze celým svým rozsahem nebo jen z části
- z územního hlediska tvoří *pragensia* díla vydaná na území Prahy bez ohledu na jejich obsah
- z autorského hlediska tvoří *pragensia* díla těch autorů, kteří jsou s Prahou spjati svojí činností nebo původem.

V Městské knihovně v Praze jsou *pragensia* chápána jako „dokumenty nejrůznější povahy, které pojednávají o Praze nebo ji zobrazují“ [BENYOVSZKÁ, 2008, s. 34]. *Pragensia* jsou tedy chápána v užším pojetí pojmu, hlavní zřetel je dán na obsahové hledisko. Ve starších dokumentech koncepční povahy je termín *pragensií* vysvětlen takto: „*pragensia* neboli *literatura o Praze* je literatura s tematikou z minulosti, současnosti i budoucnosti Prahy ze všech vědních oborů, se vztahem k Praze, periodika, krásná literatura a jednotlivé tisky“ [MĚSTSKÁ, 1990, s. 1].

Pokud bychom ovšem nevztahovali *pragensia* pouze na dokumenty tištěné povahy, je možné použít tuto zajímavou definici pojmu *pragensií* z diplomové práce V. Novákové. *Pragensia* je možné definovat jako „souhrn psaných, obrazových a ostatních hmotných pramenů, které jsou výsledkem hospodářského, kulturního a politického života v procesu historického vývoje a které se svým obsahem a učením týkají Prahy, přičemž není vždy podmínkou, aby vznikaly na jejím území“ [NOVÁKOVÁ, 1978, s. 18].

4.2 Vymezení profilu pragensijního fondu

Vymezit profil pragensijního fondu je náročné, výchozím materiálem pro zpracování jsou údaje o pragensijních dokumentech z AKS Koniáš a informace získané z materiálu Kateřiny Richterové [MĚSTSKÁ, 2002a], Věry Benyovszké [MĚSTSKÁ, 2006c] a z práce Soni Pomikálkové [POMIKÁLKOVÁ, 2008]. V této části práce jsem se rozhodla pro zjednodušení a výstižnější srovnání uvést kromě představení profilu historického pragensijního fondu také porovnání celého historického pragensijního fondu s částí fondu vybranou k digitalizaci.

Pragensijní fond byl vytvořen vyčleněním publikací s pražskou tematikou z fondu MKP a dokoupením publikací v antikvariátech. Veškerý fond uložený ve Studovně pragensii byl již od jejího založení určen výhradně k prezenčnímu studiu [POMIKÁLKOVÁ, 2008, s. 31].

Fond Studovny pragensii tvoří cca 7000 svazků titulů knižní povahy, 200 plánů a map, více než 1000 svazků vázaných periodik (celopražský deník *Večerní(k) Praha*, tisk obvodních, městských a místních úřadů a dalších regionálních vydavatelů, částečně odborné časopisy), 50 CD-ROM a cca 2000 obálek s pragensijními výstřižky [MĚSTSKÁ, 20??, s. 2]. Historický pragensijní fond reprezentuje přes 1 900 svazků titulů knižní povahy, 2 plány a mapy a 2 tituly periodik²⁴. Z toho vzácné historické dokumenty představují 230 dokumentů vydaných do roku 1860 a dalších více než tisíc dokumentů do 20. let 20. století.

Dokumenty jsou ve Studovně pragensii obvykle k dispozici v jednom exempláři, v případě mimořádně žádaných dokumentů nebo v případě dokumentů, které jsou nezbytné pro odbornou práci knihovníka, jsou tituly ve dvou exemplářích. Další exempláře byly po stěhování fondu z Mostecké ulice na Mariánské náměstí umístěny do centrálního skladu Ústřední knihovny, případně na volný výběr úseku naučné literatury pro dospělé i pro děti a mládež v Ústřední knihovně [MĚSTSKÁ, 2002b, s. 2].

Typy dokumentů, které se vyskytují napříč celým historickým pragensijním fondem:

- knihy v pevné vazbě a brožury (monografie, sborníky)
- drobné tisky (regionální a příležitostné tisky)
- periodika (např. *Světovzor*, *Večerní Praha*)
- regionální periodika

²⁴ Tato statistika byla odvozena z dávek dokumentů určených k digitalizaci, odráží informace z pole *Typ dokumentu* v AKS Koniáš.

- kartografické dokumenty (plány a mapy)
- skripta a další studijní a metodické texty
- katalogy výstav
- adresáře od konce 19. stol. do poloviny 20. stol.
- články a výstřižky z tisku (archiv výstřižků) [POMIKÁLKOVÁ, 2008, s. 31].

V rámci profilace fondu byly vyčleněny typy dokumentů, které se nezačleňují do fondu a které tedy Studovna pragensii nesoustřeďuje. Patří mezi ně archiválie a listiny, které v podmínkách MKP není možné uchovávat a za jejichž uložení je zodpovědný Archiv hlavního města Prahy. Dalšími dokumenty jsou fotografie, diapozitivy, plakáty, kalendáře a pohlednice, které jsou většinou předmětem zájmu muzeí a které by z hlediska množství nebylo možné v prostorách MKP ukládat a opatrovat. Také komunální dokumenty a informační materiály s omezenou časovou platností nejsou shromažďovány, v rámci historických dokumentů byly vyčleněny pouze dokumenty s širším záběrem, které zapadají do celkové profilace fondu [MĚSTSKÁ, 2002b, s. 1].

Tematické vymezení fondu je rozděleno na dvě části: na *vlastní pragensia* a *pragensia v širším slova smyslu*. *Vlastní pragensia* jsou obsahově vymezená *pragensia* neboli „dokumenty, u nichž je pragensijní námět zvolen jako hlavní a ústřední téma, jež stojí v centru pozornosti autora“ [MĚSTSKÁ, 2002a, s. 5], patří mezi ně oborové monografie a další specializované pragensijní dokumenty. *Pragensia v širším slova smyslu* jsou dokumenty „širšího záběru přesahující do jiných oborů a témat, které ve své části významně pojednávají o Praze v souvislosti s její úlohou hlavního města státu a centra politického, kulturního i společenského života. Ve fondu jsou zastoupeny odborné, úzce specializované práce vědeckého charakteru i populárně naučná literatura určená spíše pro laickou veřejnost, součástí fondu je výběrově také krásná literatura s pragensijní tematikou“ [MĚSTSKÁ, 2002b, s. 1]. U podrobnějšího členění tematických okruhů lze vycházet z *obsahové charakteristiky dokumentů*²⁵. Přestože je obsahová charakteristika platná pro celý pragensijní fond, je možné použít vyčlenění také na část fondu, tedy *historická pragensia*. Dřívější

²⁵ *Obsahová charakteristika* (OCH) je vlastní klasifikace naučné literatury v MKP vytvořená ve 30. letech minulého století. S mírnými změnami je užívána dodnes. Každému oboru je přiřazen určitý znak, notaci tvoří písmeno a číslice, např. D5, kterým jsou označovány dokumenty pojednávající o Praze. Stojí na pomezí předmětové a systematické klasifikace. Třídění naučné literatury sloužilo pro budování systematického katalogu a nyní pouze při stavění knih na volném výběru [POMIKÁLKOVÁ, 2008, s. 27-28].

základní rozdělení fondu na *pražské dějiny, umění v Praze a průvodce a plány Prahy* nevystihovalo dostatečně tematickou různorodost, proto jsou nyní pragensia členěna na pět hlavních tematických okruhů:

- *pražské dějiny, politika a ekonomika*
- *pražská kultura, umění a společnost*
- *pražský místopis*
- *pražská příroda, věda a technika*
- *beletrie o Praze* [MĚSTSKÁ, 20??, s. 1]

Podrobnější členění udávající větší představu o hlavních pragensijních tématech bylo sice zpracováno dříve, ale lépe ilustruje obsah fondu. Pro účely této práce byly vyjmuty tematické části nekorrespondující s historickým zaměřením fondu. Co je tedy pragensijní tematika:

- dějiny Prahy (včetně pomocných věd jako archeologie, heraldika, epigrafika atd.)
- Praha z hlediska přírodních věd – geografie, geologie, klimatické podmínky, příroda a její ochrana
- místopis Prahy – měst, čtvrtí, obcí, ulic, domů, staveb atd.
- architektura a umění v Praze
- obecné i jednotlivé otázky vědy, techniky a průmyslu v Praze
- komunální záležitosti, doprava, obchod, služby v Praze
- kultura, sport, školství, zdravotnictví, sociální péče v Praze
- osobnosti
 - vážící se svou působností k Praze
 - slavní návštěvníci Prahy [MĚSTSKÁ, s. 1]

Pro účely této práce bylo vybráno 5 hlavních tematických okruhů, které jsou zastoupeny v rámci obsahové charakteristiky dokumentů *pěti hlavními OCHy*:

D5a *pražské dějiny, politika a ekonomika*

D5b *pražská kultura, umění a společnost*

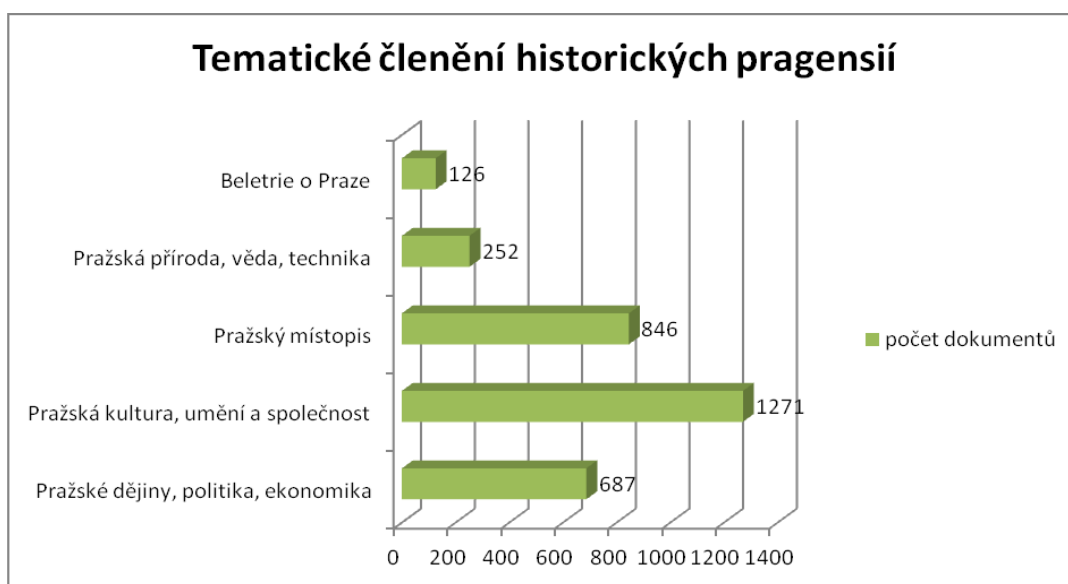
D5c pražský místopis

D5d pražská příroda, věda a technika

D5e beletrie o Praze

Hlavní obsahové charakteristiky jsou dále členěny na podrobnější témata²⁶. Seznam všech používaných pragensijních obsahových charakteristik lze nalézt v Příloze č. 1.

Jak je vidět na grafu č. 1, jsou v dokumentech historických pragensii zastoupeny nejvíce ty pojednávající o *Pražské kultuře, umění a společnosti, kulturních dějinách a společensko-kulturních událostech a slavnostech*. Tyto dokumenty svým počtem výrazně převyšují ostatní tematické celky. Další v pořadí vzhledem k počtu dokumentů jsou srovnatelně dvě obsahové charakteristiky: *Pražský místopis* a *Pražské dějiny, politika a ekonomika*.

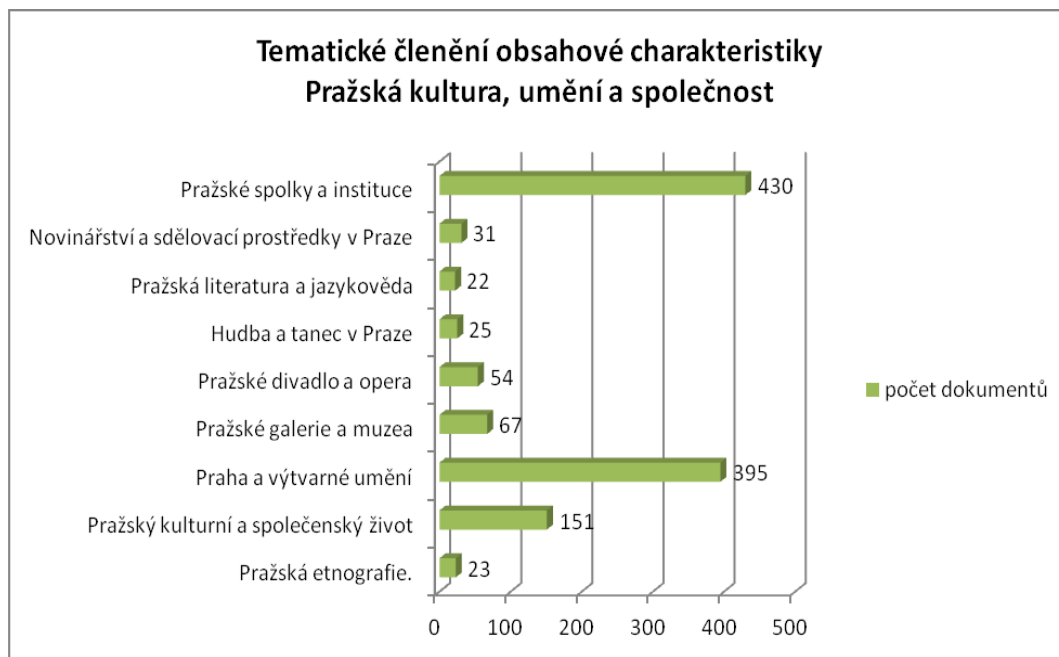


Graf 1: tematické členění fondu historických pragensii

Je zcela přirozené, že největší podíl ve fondu mají témata kultury a umění. Praha vždy byla kulturním centrem Čech v průběhu věků i různých státních celků a knihovna zcela logicky shromažďovala dokumenty mapující kulturní a společenský život. Pokud srovnáme

²⁶ Notace tohoto třídění má následující charakter: D5 je hlavní skupiny označující pragensia, po ní následuje písmeno odkazující na hlavní tematický celek (a-e). Další v pořadí je číslice (v notaci není použita číslice vyšší než 9), další podskupinu tvoří opět písmena abecedy (poslední využitě písmeno je j), poslední podskupinu tvoří znovu číslice. Například pokud rozebereme notaci D5a4a. D5 značí *pragensijní tematiku*, a náležitost k hlavní tematické skupině *Pražské dějiny, politika a ekonomika*. Číslice 4 poukazuje na podskupinu označenou celým názvem jako *Praha novověká, přehledy pražských dějin 16.-21. století, Praha za vlády Habsburků, české stavovské povstání v Praze, pražská defenestrace, bitva na Bílé hoře v Praze, počátky občanské společnosti a dělnického hnutí v Praze*. Poslední a pak odkazuje na další podskupinu, která je tematicky užší a odvozená od vyšší podskupiny, její název je *Praha 19. století, Praha a české národní hnutí, pražské revoluční události let 1848-1849, rozvoj občanské společnosti v Praze*.

podskupiny v rámci tohoto hlavního tématu, docházíme při pohledu na graf č. 2 k závěru, že nejpočetnějšími kategoriemi jsou *Pražské spolky a instituce* a *Praha a výtvarné umění*.

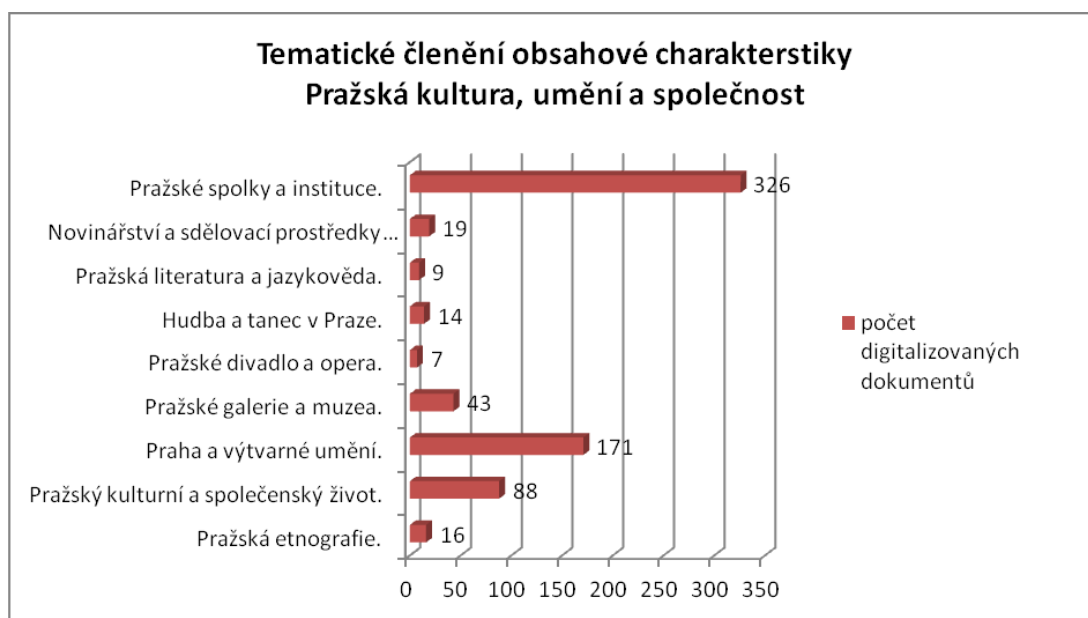


Graf 2: Tematické členění OCHu Pražská kultura, umění a společnost

Při pohledu blíže do obsahu podkategorie *Pražských spolků a institucí* zjistíme, že nejvíce dokumentů je označeno obsahovou charakteristikou D5b9a. Ta pojednává o pražských školách a pražském školství, dále obsahuje dějiny, almanachy a výroční zprávy pražských škol a akademií a samozřejmě zahrnuje dokumenty o Univerzitě Karlově a o životě pražských studentů. Další dvě témata, *Pražské knihovny a archivy* a *Církev a náboženství v Praze*, jsou zastoupena již v daleko menší míře.

Vzhledem k unikátnosti dochovaných almanachů a výročních zpráv pražských škol a akademií je zcela jasné, že právě tyto dokumenty tvoří z celé obsahové charakteristiky *Pražské kultury, umění a společnosti* největší podíl titulů vybraných k digitalizaci. Almanachy a výroční zprávy tvoří ucelenou sbírku a jsou tak cenným historickým materiálem, jehož důležitost je dána žádostmi o informace ze stran badatelů a studentů tvořících hlavní uživatelskou základnu historických pragensií. Jak je patrné na grafu č. 3, bylo z celkového množství 430 titulů digitalizováno 75 % z nich, tedy 326 titulů. Také druhá nejpočetnější podkategorie zůstala u digitalizovaných titulů stejná, *Pražské výtvarné umění*. V rámci historických dokumentů činila velikost podkategorie celkem 34 %, ve vybraných titulech k digitalizování klesla její četnost v celkovém počtu na 24 %. Ostatní podkategorie zůstaly v téměř totožném poměru jako u celého historického pragensijního fondu, výjimku tvoří

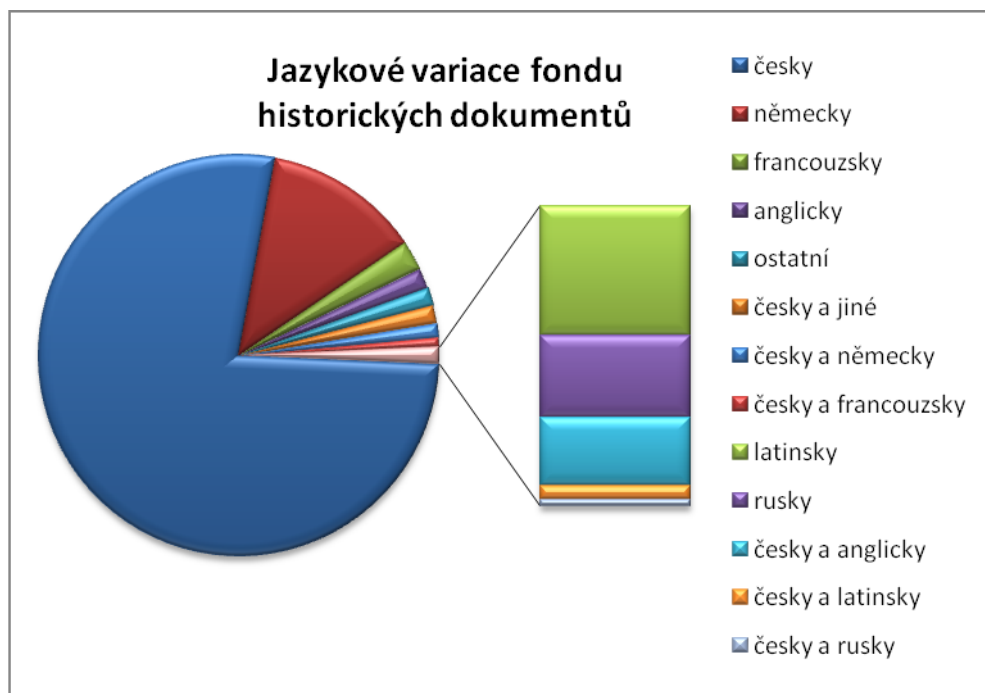
podkategorie *Pražské divadlo a opera*, kde bylo k digitalizaci vybráno pouze 7 titulů z celkového počtu 54 titulů. I tento významný pokles má své opodstatnění, vzhledem k tomu, že dalšími částmi fondu MKP určenými k digitalizaci byly fond divadelní a hudební, jejichž přesah a vymezení vůči fondu pragensijnímu je stále nedořešený, byl výběr titulů vhodných k digitalizaci z těchto oblastí přenechán divadelnímu a filmovému úseku a hudebnímu úseku.



Graf 3: Tematické členění OCHu Pražská kultura... (digitalizované dokumenty)

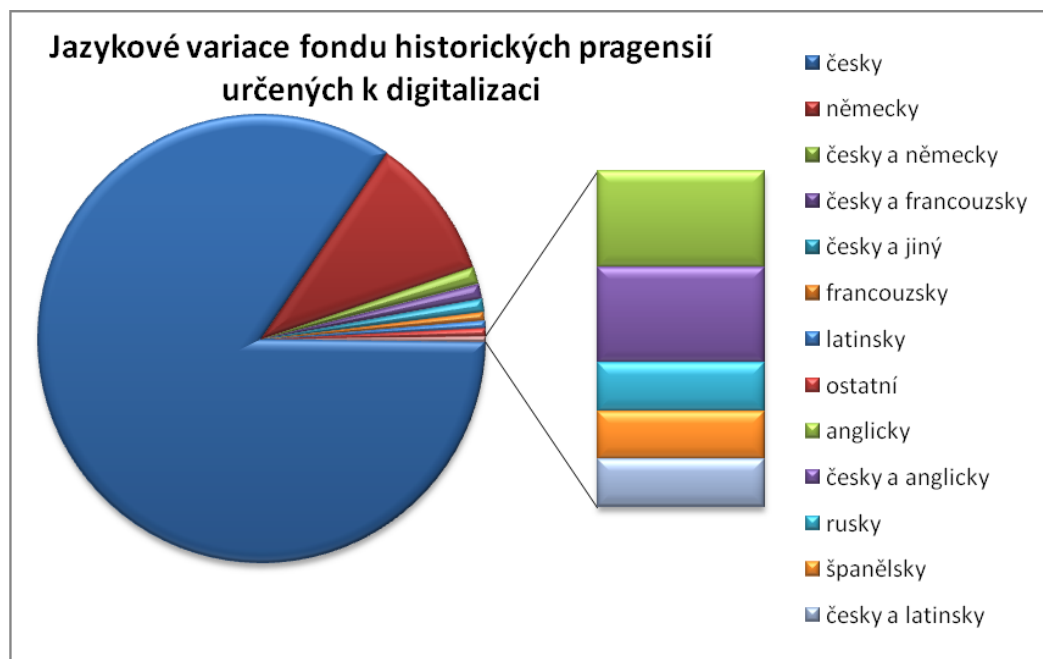
Druhou největší kategorií v rámci pražské kulturní scény je *Praha a výtvarné umění*, která zahrnuje pražskou architekturu a urbanismus, sochařství, památkovou péči, malířství, kresbu a grafiku, užité umění a umělecké řemeslo a také fotografii a film. *Pražská architektura a urbanismus*, *Pražské malířství*, *kresba a grafika* a *Pražská fotografie a film* jsou nejvíce zastoupená témata zahrnující přibližně stejný počet dokumentů.

Jazykové rozložení fondu historických pragensíí je znázorněno v Grafu č. 4. V historickém pragensijním fondu převládají dokumenty v českém jazyce. Po nich následují dokumenty psané německy. Třetím nejčastějším jazykem je francouzština, za ní následuje angličtina. Další v pořadí jsou v podobném množství dokumentů kombinace českého jazyka a dalších jazyků, v tomto pořadí podle četnosti: jiné, německy, francouzsky, anglicky a okrajově latinsky a rusky. V kategorii ostatní jsou v grafu uvedeny minoritní jazyky, počet dokumentů psaných těmito jazyky se pohybuje v řádu jednotek. Objevují se zde obvyklé jazyky jako španělština, italština či polština, také kombinace latiny a jiného jazyka. V historickém fondu jsou ovšem zastoupeny také méně obvyklé jazyky jako srbochorvatština, rumunština a umělé jazyky jako esperanto.



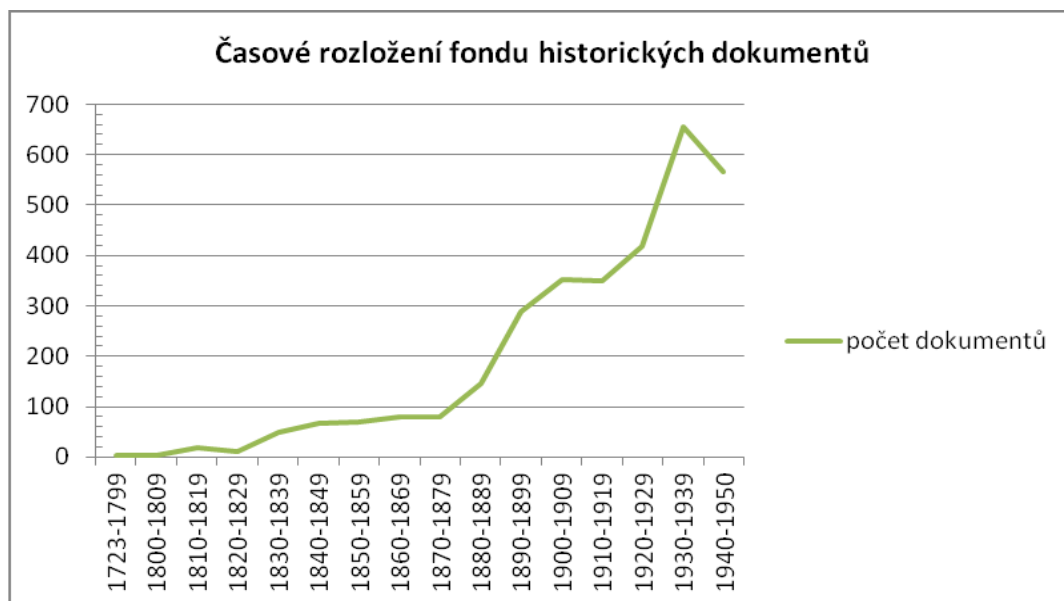
Graf 4: Jazykové variance fondu historických pragensijních dokumentů

Pokud srovnáme rozložení jazyků s digitalizovaným historickým pragensijním fondem, můžeme potvrdit podobnou tendenci rozložení jako je tomu u celého fondu historických pragensijních (viz Graf č. 4 a Graf č. 5). Jasně převládajícím jazykem je čeština tvořící 84 % z celkového počtu dokumentů, dalším jazykem je němčina tvořící 11 %. Zatímco v celém fondu historických pragensijních je procentuální poměr těchto jazyků rozdílný: česky 77 %, německy 13 %. Poměr mezi češtinou a němčinou je tedy v případě dokumentů určených k digitalizaci 8:1, zatímco v případě celého historického fondu pragensijních je poměr mezi češtinou a němčinou nižší, 6:1. Celkově se snížil výskyt jiných jazyků. Například počet anglicky psaných dokumentů klesl z celkové hodnoty 1,6 % na hodnotu 0,1 %, podobný pokles postihl také francouzštinu – z 2,4 % na 0,6 %. Ostatní jazyky jsou zastoupeny pouze slovenštinou, rumunštinou a esperantem.



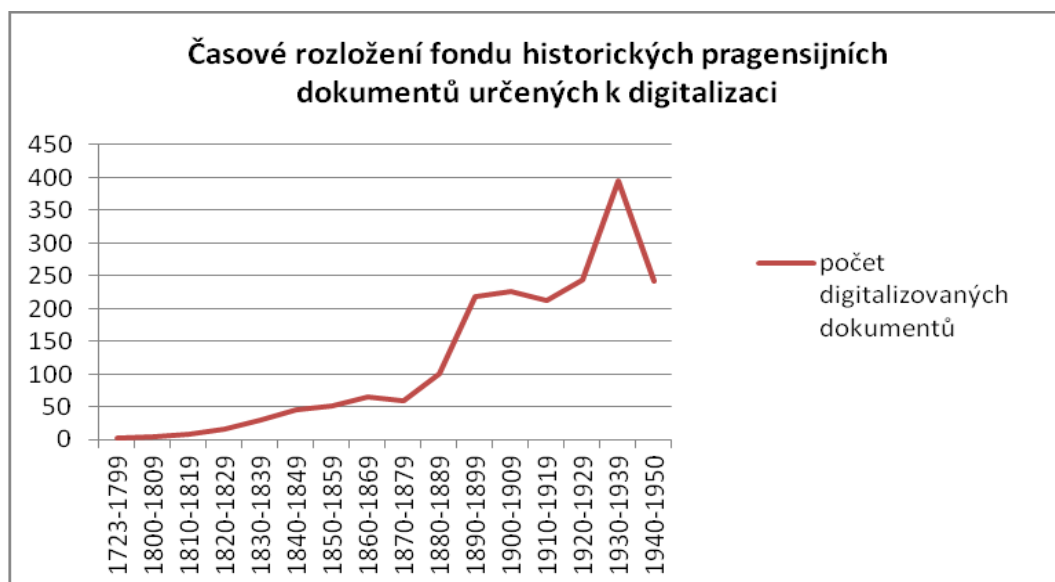
Graf 5: Jazykové variace fondu historických pragensií určených k digitalizaci

Časové rozložení fondu historických dokumentů vyobrazené pomocí časové osy je naznačeno v grafu č. 6. Po vyčlenění vzácných historických fondů a jejich přesunu do Oddělení vzácných tisků je „nejstarším dokumentem ve Studovně pragensií Prodomus Gloriam Pragensem (Posel slávy pražské) J. Hammerschmidta z roku 1723, který byl poškozen nevhodnou převazbou“ [POMIKÁLKOVÁ, 2008, s. 33]. Do roku 1829 je fond tvořen pouze desítkou titulů, od roku 1830 pak počet titulů v dalších dekádách stoupá. Do roku 1880 však počet dokumentů za celé desetiletí nepřesahuje číslo 100, mezi léty 1890–1899 je velký rozdíl v počtu dokumentů, přibylo téměř dvojnásobné množství titulů oproti 80. létům 19. století.



Graf 6: Časové rozložení fondu historických pragensijních dokumentů

V letech 1900–1919 byl růst počtu pragensijních dokumentů pozastaven, celkový počet titulů z těchto dvou dekád je téměř shodný, liší se pouze o tři tituly. Další strmý nárůst dokumentů je evidován ve 30. letech 20. století. Ve 40. letech 20. století nastává pokles počtu titulů. Oba poklesy – ve druhé dekádě 20. století i ve 40. letech 20. století je možné pokles počtu dokumentů vysvětlit konáním první a druhé světové války.



Graf 7: Časové rozložení fondu digitalizovaných historických pragensijí

Pokud srovnáme časový úsek celého fondu historických pragensijí s částí fondu vybranou k digitalizaci, můžeme z grafu č. 7 vidět, že nejstarší historické tituly byly většinou vybrány v podobném poměru k celkovému fondu, do 1830 byly téměř všechny tituly. Od roku 1830

do 1889 byly vybrány asi dvě třetiny titulů, Od konce 19. století začíná pokles vybraných titulů. V 90. letech 19. století byly zvoleny téměř čtyři pětiny fondu, od počátku 20. století do roku 1939 tři pětiny fondu a ve 40. letech 20. století pouze dvě pětiny fondu. Je tedy vidět jasný trend, čím novější byly dokumenty, tím pečlivější a procentuálně větší výběr byl realizován.

4.3 Uložení fondu pragensií

Fond pragensií je uložen nejen ve specializovaném úseku MKP ve Studovně pragensií, ale také v jiných pobočkách v síti MKP²⁷. Ve všech pobočkách v síti je možné najít v různém rozsahu fond regionální literatury, tedy pragensijní fond. Většinou je uložen ve volném výběru, kde je řazen podle obsahové charakteristiky. Velmi důležité je, zda je možné daný titul zařadit do skupiny *vlastních pragensií*, tedy dokumentů pojednávajících o Praze celým obsahem, nebo *pragensií v širším slova smyslu*. Vlastní pragensia mají primární obsahovou charakteristiku pragensijní a stojí tedy na regálu vedle sebe. Zatímco u druhé skupiny pragensií, kde je primární obsahová charakteristika určená podle hlavního oboru, kterému se dokument věnuje nejvíce, je dokument zařazen do určeného oboru na regále. Tyto dokumenty pak nejsou spolu a jsou roztroušeny ve volném výběru. Ve Studovně pragensií jsou ovšem začleněny podle daného pragensijního OCHu. V některých pobočkách je regionální literatura uložena ve skladech.

Studovna pragensií je umístěna spolu s *Informačním střediskem* v přízemí Ústřední knihovny, má vyhrazen dvoupodlažní prostor. Většina pragensijního fondu, asi dvě třetiny, jsou umístěny v tematicky řazeném volném výběru na galerii. V příručním skladu a na vázaném výběru na galerii jsou vyčleněny historické a vzácné tisky, mapy, výstřižky a atypické formáty. A nejstarší a historicky necennější část fondu označená signaturami H až HE je chráněna v uzamykatelných skřínkách v přízemí Informačního střediska [MĚSTSKÁ, 20??, s. 2]. „Z prostorových důvodů nebylo možné uložit veškerý fond vydaný před rokem 1900 a všechn vzácný fond (celkově cca 1/4 fondu). V roce 1998 byly nejstarší tituly převedeny do péče Oddělení vzácných tisků“ [POMIKÁLKOVÁ, 2008, s.33].

²⁷ Knihovní síť MKP je tvořena 43 pobočkami včetně Ústřední knihovny a 2 bibliobusy [MĚSTSKÁ, 2011].

5 Digitalizace v MKP

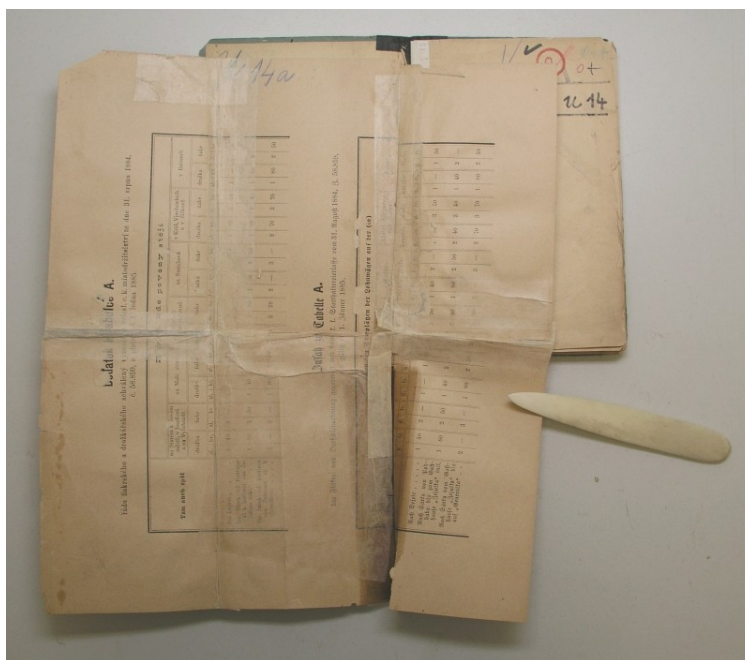
5.1 Strategie výběru dokumentů

Již během kompletace žádosti o poskytnutí grantu z „norských fondů“ byly vytipovány ohrožené dokumenty ze všech vybraných úseků – ze Studovny pragensí, z Divadelního a z Filmového úseku MKP. Seznam typů dokumentů s odhadovanými celkovými součty stran a s příklady konkrétních titulů již byl přiložen jako příloha k žádosti o grant.

Mezi hlavní kritéria výběru dokumentů patří:

- unikátnost titulu v rámci ČR
- neexistence digitální kopie tohoto titulu
- špatný fyzický stav dokumentů (vlivem stáří, častého používání, degradací nekvalitního papíru, špatného životního prostředí)
- odkaz na historický fond MKP
- stav knižní vazby
- atraktivita nebo také žádanost titulů uživateli.

Na Obrázku č. 6 je vyobrazen dokument splňující všechna výše uvedená kritéria.



Obrázek 6: poškozený svazek určený k digitalizaci (autor fotografie: Eva Měřinská)

Výběr dokumentů provádějí odborníci z řad knihovníků působících v jednotlivých úsecích. Právě oni mají nejlepší přehled o titulech ve svém fondu, jsou schopni posoudit jak fyzický stav dokumentu, tak také atraktivitu pro uživatele – jsou to oni, kdo jsou s uživateli neustále ve styku, znají jejich informační potřeby a poskytují uživatelům požadované informace.

V první fázi jsou v rámci procesu výběru dokumentů vhodných k digitalizaci také zhodnoceny *kvalitativní a kvantitativní znaky*, které mohou mít vliv na manipulaci s dokumenty. Základním znakem je *stupeň degradace papíru* vlivem jeho kyselosti. Digitalizace těchto dokumentů je, bohužel, vyvážena snahou o zachování předlohy. V rámci procesu přípravy je proto rozhodnuto, zda je vůbec možné tuto předlohu zpracovat – u některých dokumentů stupeň degradace postoupil natolik, že po pouhém doteku se list papíru úplně rozpadne. Pokud je to tedy prakticky možné, dokumenty i v takto pokročilém stadiu degradace papíru se digitalizují, proces digitalizace je většinou pro tyto tituly téměř likvidační – proto je při zpracování, především při skenování, nutné dodržet zásady šetrného zacházení s dokumentem a minimalizovat nutné manipulační zásahy. Tyto dokumenty jsou po velmi šetrném naskenování a zpracování přesunuty do klimatizovaného depozitáře v Brusnici (viz kapitola č. 5.2.10 Kontrola fyzických svazků a určení následné péče).

Dalším znakem, který je třeba při vybírání dokumentů posoudit, je *knižní vazba*. Původní knižní vazby většinou při opatrné manipulaci nečiní při skenování na skeneru potíže. Bohužel, Městská knihovna v Praze začala s kompletním převazováním dokumentů s uvolněnou vazbou. Kvůli této převazbě jsou v dokumentu většinou zúženy vnitřní okraje tisku a skenování je velmi obtížné. U dotčených titulů je nutné prakticky na skeneru vyzkoušet, zda je možné dokument naskenovat svázaný nebo zda je nutné dokument rozešít. Pokud to stav dokumentu dovolí, je upřednostňována varianta naskenování svázané předlohy. U některých dokumentů to není možné, proto musí být odeslány do knihvazačské a knihtiskařské dílny, kde jsou odborně rozešity. Pokud je to uskutečnitelné a kvalita degradovaného papíru to dovolí, pak jsou po ukončení procesu digitalizace znovu svázané, a vráceny zpět na regály jednotlivých úseků nebo jsou svazky opatřeny ochranným obalem z nekyselého papíru a ponechány v prostředí klimatizovaného depozitáře.

Protože podmínkou přidělení dotace bylo zamezení duplicitní digitalizace, bylo nutné ještě před samotným zahájením digitalizace zkontrolovat, zda vybrané tituly z jednotlivých úseků již nejsou k dispozici v digitální podobě, nebo se některá z institucí nechystá je

digitalizovat. Na počátku projektu, kdy ještě neexistoval *Registr digitalizace*²⁸ (dále v textu je užívána také zkratka RD.CZ), se kontrola možných duplicit prováděla velmi obtížně. Protože jedinou knihovnou, která digitalizovala podobný fond, byla Národní knihovna ČR. Po konzultaci s NK ČR byla zvolena metoda ruční kontroly pomocí seznamu digitalizovaných titulů. Ještě před počátkem digitalizace tak byly předány seznamy titulů ke kontrole. Po proběhnutí této kontroly mohly být dokumenty postoupeny dále k digitalizaci.

Po vzniku *Registru digitalizace* je kontrola možných duplicit ulehčena. Stačí pouze vyhledat dokument plánovaný v digitalizaci v RD.CZ podle identifikačních prvků, jakými jsou název dokumentu, autor atd. Pokud se dokument nepodaří najít, je možné jej digitalizovat. Městská knihovna v Praze pomocí protokolu OAI-PMH dodává záznamy do RD.CZ. Registr digitalizace umožňuje také knihovnám sdílet informaci, které dokumenty se chystají k digitalizaci, nebo dokonce umožňuje knihovně provádět kontrolu digitalizačního workflow. MKP zatím využívá pouze možnost přispívat do RD.CZ záznamy digitalizovaných titulů, výhledově je naplánováno zapojit se také do sdílení záznamů titulů určených k digitalizaci, tzv. *Plánované digitalizace*.

Výhodou využití RD.CZ pro kontrolu duplicit je možnost porovnání digitalizovaných titulů nebo ohlášených titulů se záznamy velkého množství institucí, které do RD.CZ přispívají. Je tak zvýšena pravděpodobnost, že se podaří případnou duplicitní digitalizaci odhalit včas, ještě před její realizací.

5.2 Proces digitalizace v MKP

Jednotlivé fáze digitalizačního workflow v Městské knihovně v Praze byly konzultovány především s *Oddělením digitalizace Národní knihovny ČR* a *Digitalizačním centrem knihovny Akademie věd*, které se více velikostí a zaměřením podobalo velikosti a typu vznikajícího OD MKP. Tato fáze přípravy a modelování byla velmi důležitá, na reálných příkladech fungování ostatních digitalizačních pracovišť se bylo možné poučit z případných chyb a minimalizovat rizika při volbě špatné strategie digitalizace. Text je dále rozdělen do podkapitol, které reflektují jednotlivé fáze procesu digitalizace.

²⁸ Projekt *Registr digitalizace* je financován v rámci *Výzkumu a Vývoje Ministerstva kultury*. Na realizaci se podílí NK ČR a KnAV ČR spolu se společností INCAD, která se stará o vývoj systému. Mezi další partnery projektu patří MZK, NTK a MKP. Cílem projektu je vytvořit národní registr digitalizovaných dokumentů [JÍLKOVÁ, 2010, s. 26; Registr]. Aplikace je dostupná přes webové rozhraní, je umožněno prohledávat databázi pomocí typů dokumentů, podle digitální knihovny či vlastníka dokumentu.

Celé digitalizační pracoviště bylo vybaveno a připraveno na ostrý proces dříve, než bylo původně plánováno, takže od listopadu 2007 do února 2008 mohl být zahájen zkušební provoz. Tato fáze harmonogramu se ukázala být velmi důležitou součástí celého projektu. Ještě před reálným ostrým provozem tak mohlo být odzkoušeno celé pracovní workflow digitalizace a mohly být vyladěny některé aspekty, které při teoretickém plánování nelze odhadnout. V této fázi došlo také k zaškolení zaměstnanců OD – jak při práci se skenery, tak pro úpravu skenů a jejich zpracování v příslušných programech. Aby si zaměstnanci mohli osvojit proces digitalizace a nabyli tak potřebné zkušenosti, vyzkoušeli si proces ochranného reformátování nejprve na vyřazených titulech z fondu MKP. Díky tomu se naučili celý proces bez strachu z možného poškození exemplářů určených pro *záchrannou digitalizaci*. V další fázi byly ke zpracování vybrány již tituly zahrnuté do projektu HISPPRA, tyto exempláře však musely splňovat další kritéria: nižší počet stránek a z toho vyplývající tenká vazba, jednoduchá struktura. Zpracování těchto dokumentů je jednodušší a přitom již mají specifické rysy historického fondu – např. odlišnou paginaci, různorodé typy písem užitých při tisku, atd. Po oficiálním otevření digitalizačního pracoviště byl zahájen již rutinní provoz a k digitalizaci byly dodávány dokumenty podle připraveného plánu.

Pro jakékoli zpracování je nezbytné dokumenty jednoznačně identifikovat. V tomto případě slouží jako vnitřní identifikátor signatura. Po mnoha uvažováních byla jako identifikátor zvolena právě signatura, jako jedinečný generovaný kód se specifickou syntaxí. Pro požadavky v rámci ICT dochází k mírné úpravě syntaxe signatury – mezery mezi jednotlivými částmi signatury, případně znak lomítka, jsou nahrazeny znakem podtržítka. Pokud se v signatuře vyskytuje pomlčka, je zachována. Výhoda užití signatury jako identifikátoru oproti novému generovanému kódu je jasná, není třeba vymýšlet nový složitý systém, pokud již určitá identifikace v rámci knihovny existuje. Signatura je také dobrým odlišovacím prvkem, je lépe zapamatovatelná a v rámci zpracování se s ní lépe pracuje než při použití delšího generovaného kódu, obzvláště kódu tvořeného číselnou řadou.

Bohužel se ukázalo, že ne všechny signatury splňují podmínku unikátnosti v celém fondu MKP. Bylo tedy nutné vytvořit seznam dokumentů, které jsou identifikovány shodnou signaturou, a tyto dokumenty v rámci celého fondu MKP znovu designovat. To bohužel znamenalo změnu signatury i v generálním katalogu. Naštěstí se změna signatury netýkala velkého množství dokumentů.

Signatura je tedy použita jako název složky obsahující skeny daného dokumentu. Skeny jsou označeny signaturou jako prefixem, sufix tvoří pořadové číslo skenu. V rámci Siria jsou signaturou nazvány také pořadače i zakázky, signatura tvoří prefix názvu výsledného XML souboru určeného k importu do digitální knihovny Kramerius.

Zpracování obrazových dokumentů je po nasnímání realizováno především v softwaru Sirius. „Software Sirius je modulární systém určený pro budování komplexních aplikací. Základem systému je dvouúrovňová architektura klient/server. Systém se sestává z jednotlivých modulů, které mají jednotný vzhled“ [ELSYST, c2003a]. Protože se nejedná o software vytvořený speciálně pro knihovny, mohou být některé použité termíny nezvyklé. Software Sirius je založen na stromové struktuře tzv. *Pořadačů*, které zastupují dokumenty jako takové. Ke každému Pořadači je přiřazena tzv. *Zakázka*, která umožňuje spravovat libovolné množství dávek, které mohou být rozpracovány současně.

5.2.1 Příprava fondu

Na přípravě fondu k digitalizaci se podílí několik oddělení včetně specializovaných úseků. Dokumenty je nutné nejprve vyjmout z regálu. Poté je zkontrolován fyzický stav dokumentu. Dochází k vyřazení duplicit a nestandardních dokumentů – tedy dokumentů fyzicky poškozených nebo neúplných. V případě multiplicity je možné vybrat z více svazků ten nejméně poškozený, vhodně doplnit poškozené stránky stránkami z jiného, zachovalejšího exempláře. V tomto případě se potvrdilo, jak velmi dobré je využít pro přípravu dokumentů pracovníky ze specializovaných úseků, kteří svou část fondu bezpečně znají a kteří jsou schopni určit, zda máme ve fondu jiný exemplář v lepším stavu. Dále bylo většinu dokumentů ze specializovaných úseků nutné rekatologizovat. Městská knihovna v Praze byla založena v roce 1891 a po celou dobu jejího fungování byly knihy katalogizovány vždy podle platných pravidel té doby. Během historie byla používána různá katalogizační pravidla, od roku 1998 se začalo katalogizovat podle pravidel AACR2. Bylo rozhodnuto, že zpětně se záznamy podle nových pravidel opravovat hromadně nemohou. Jedním z důvodů byl fakt, „že jakákoli změna uložená do systému znamená nutnost fyzické opravy – tedy všechny svazky daného titulu ze všech poboček shromáždit, přelepit a přeřadit“ [FERDINANDOVÁ, 2011]. Vzhledem k velikosti MKP a především k počtu jednotlivých exemplářů uložených v různých pobočkách rozmístěných po celém území hlavního města, komplikacím spojeným s nutností jejich shromážděním bylo rozhodnuto, že většina fondu se podle nových katalogizačních pravidel zpětně nezpracovává.

Ovšem bibliografické údaje ze záznamů dokumentů určených k digitalizaci, které většinou nebyly podle výše uvedeného pravidla rekatalogizovány, jsou importovány do softwaru Sirius a použity jako základ údajů popisných metadat. Proto bylo rozhodnuto o zpětné rekatalogizaci záznamů těchto dokumentů podle nových pravidel. Této práci se ujaly jak pracovníci z *Oddělení katalogizace* (v případě jmenného zpracování), tak především pracovníci specializovaných úseků – tedy pragensijního, divadelního a hudebního (v případě věcného zpracování). Teprve poté, co byl bibliografický popis dokumentů v pořádku, bylo možné dokumenty připravit pro transport do Oddělení digitalizace. Realizace transportu se ukázala být jedním z nedostatků plynoucích z jiné lokalizace OD. V původní představě byl transport velmi jednoduchý – pouze v rámci jedné budovy na Mariánském náměstí. S přesunem pracoviště do Provaznické ulice, která je, co se týče obslužnosti, ve velmi komplikované lokalitě – automobilem je možné se v ulici pohybovat pouze do 8 hodiny ráno, bylo nutné znovu zvážit otázku přesunu dokumentů. Nakonec bylo rozhodnuto, že vzhledem ke krátké vzdálenosti mezi Mariánským náměstím, kde byly dokumenty fyzicky umístěny, a Oddělením digitalizace, se transport bude realizovat jen s využitím lidských zdrojů, po menších dávkách než bylo původně předpokládáno a s vyšší frekvencí. Výhodou tohoto zvoleného způsobu přesunu je skutečnost, že doba, po kterou dokumenty opustily regály, a délka trvání procesu digitalizace byly zkráceny. Mezi zpracováním dokumentů nenastaly dlouhé prodlevy, které jsou zákonitým výsledkem transportu většího množství exemplářů, byla umožněna průběžná rekatalogizace a tím plynulá výměna dokumentů. Nevýhodou takto zvoleného systému přenosu představuje nutnost častého kontrolování množství dostupných dokumentů připravených v OD k digitalizaci. Za dostatečné množství dokumentů odpovídá jeden pracovník. Zpočátku bylo nutné zjistit, jaké množství dokumentů se průměrně zpracuje při plném nasazení za jeden den a podle toho včas informovat příslušné úseky, že je třeba připravit další dávku dokumentů.

Vzhledem k tomu, že dokumenty opouštěly budovu *Ústřední knihovny*, bylo nutné vést jejich evidenci a při transportu předkládat předávací protokoly o výměně exemplářů. A to jak při přijetí, tak při odevzdání jednotlivých titulů ve specializovaných úsecích. Pro lepší přehlednost byl zaveden systém dávek²⁹, přijímají a odevzdávají se pouze kompletní dávky. Důvody k rozhodnutí o použití kompletnosti dávek jsou jasné: je zachována kontinuita v transportu, každý exemplář je zaevidován a je zřejmé, kde se v určitou chvíli nachází.

²⁹ Jednu dávku označenou určitým klíčem tvoří vždy dokumenty z jednoho transportu.

Vzhledem k velkému množství exemplářů a k tomu, že v určitém období byly k digitalizaci posílány paralelně tituly z více úseků najednou, udržoval tento systém pořádek a včas informoval o uplynulé době, kterou dokument strávil v OD. Především šlo o preventivní opatření, aby nedocházelo ke zbytečnému prodlužování doby, po kterou jsou exempláře vyjmuty z regálu a připraveny na digitalizaci. Výjimku tvoří dávky, které obsahují dokumenty nutné k restaurování či dalšímu prozkoumání, v tomto případě je při odevzdání dávky zpět do jednotlivých úseků označeno, které tituly chybí a důvod jejich absence. Tyto tituly jsou poté většinou po určité době shromážděny a vráceny zpět ve zvláštní dávce.

Po přesunu do Oddělení digitalizace dochází k exportu seznamu přijatých dokumentů, který slouží pro místní evidenci. Z tohoto seznamu je možné rychle odečíst, kolik dokumentů z dané dávky je právě v procesu zpracování a kolik nezpracovaných titulů v dávce zbývá. Poté, co je v papírovém seznamu dokument označen za rozpracovaný, je nutné zadané údaje zaznamenat také do elektronické tabulky, ze které je poté možno vyčíst, kdy začalo zpracování daného dokumentu, kdo jej zpracoval, počet vyexportovaných stránek a datum dokončení digitalizačního procesu. Lze tak snadno ověřit, kdo zpracovával daný dokument a jak dlouho tento proces trval, případně, zda se vyskytly nějaké těžkosti – např. dokument byl rozdělen na velké množství kapitol, takže doba potřebná na jeho analytické zpracování byla úměrně delší než dokumentu, jež je tvořen spojitým textem.

5.2.2 Skenování

Skenování dokumentů probíhá na knižních face-up skenerech, které již byly představeny v kapitole 2.5.3 Skenery. Na počátku byli všichni pracovníci pro práci se skenery proškoleni odbornou firmou, která ukázala, jaké vlastnosti tyto skenery mají a jaké možnosti se při skenování jednotlivých typů dokumentů dají využít, samotný postup při skenování bylo možné vyzkoušet během pilotního provozu pracoviště.

Aby skenery mohly sloužit svému účelu, projdou čtyřikrát ročně servisními zásahy, které zahrnují technickou kontrolu, kalibraci³⁰ barev u modelu Bookeye, vyčištění a promazání pohyblivých součástí, aktualizace softwaru atd. V případě potřeby je v OD k dispozici kalibrační sada a je možné skener Bookeye zkalibrovat. Z dosavadních zkušeností vyplývá, že kalibrace v rámci pravidelného servisu je pro účely digitalizace v MKP dostačující.

³⁰ Kalibrace je soubor úkonů, kterými se stanoví přesná pozice snímací hlavy vůči chladné katodové lampě a skenované předloze [VYCHODIL, 2007, s. 110].

Při skenování se používají černé nebo bílé matné papíry, kterými jsou jednotlivé stránky dokumentu podkládány. Pokud je papír skenovaného dokumentu natolik tenký, že tisk prosvítá na druhou stranu, je dobré použít pro podložení bílé papíry, které minimalizují nežádoucí efekty plynoucí z prosvícení stránek. U skenování bez podložení bílým nebo černým papírem je patrná nejen druhá strana naskenovaného listu, ale mnohdy prosvítají také následující stránky knižního bloku. U takového dokumentu je vhodné nasnímat zkušební skeny a otestovat, zda je lepší použít na podložení černý nebo bílý papír.

U ostatních dokumentů, u nichž nehrozí prosvícení stránek, jsou vždy listy podkládány černým papírem, který usnadňuje proces automatické detekce okrajů stránky v SW Sirius. Při používání černých papírů je automatická detekce okrajů přesnější, než když se skenuje bez podkládání papíry. Při manuální kontrole následující po automatické detekci je díky tomu minimalizováno množství ručních oprav u programem navržených okrajů. Vždy platí, že čím lépe je dokument naskenovaný, tím menší množství práce je třeba vynaložit při následné úpravě v softwaru Sirius.

Bílý kancelářský papír má při skenování ještě další využití – slouží k vyrovnání složených map a plánů vevázaných do dokumentu. Naskenování rozloženého materiálu, který je součástí knižního bloku, je velmi obtížné. Pokud je mapa větších rozměrů, které několikanásobně přesahují velikost knižního bloku, bývá složena/přeložena několikrát. Tyto sklady jsou pak velmi výrazné a ani vhodně zvolená poloha knižní kolébky neumožní předlohu správně nasnímat. Výsledný sken pak odhaluje, že sklady papíru jsou často na nevhodných místech, deformují linii tisku a tím znemožňují orientaci v plánu či v mapě. Pokud není možné z jakéhokoli důvodu předlohu rozešít, bylo nutné hledat jiné způsoby, jak ji naskenovat³¹. Během pokusů o způsob nasnímání vevázaných map byla objevena metoda podkládání stohem papíru, který umožňuje vyrovnat sklady na papíře. Tuto skutečnost jsem objevila při snaze naskenovat mapu vevázanou do dokumentu. Při použití přitlačného skla při snímání předlohy se deformace minimalizují. Tento způsob lze samozřejmě využít jen u menších map a plánek, jejichž velikost nepřesahuje formát A2³².

³¹ Vzhledem k tomu, že vevázané mapy a plány se v dokumentech z 19. a 1. pol. 20. století vyskytují poměrně sporadicky, bylo rozhodnuto, že tyto předlohy se skenují vlastními silami. Jedině v případě, že se jedná o větší množství dokumentů nebo je nemožné naskenovat předlohu na dostupných skenerech, je nutné objednat skenování na zakázku.

³² Formát A2 je standardizovaný formát papíru, splňuje normu ČSN EN ISO 216, papír má velikosti stran v poměru $1:\sqrt{2}$ (přibližně 1: 1,4142) a jeho rozměry jsou 420 × 594 mm [TYPO.CZ, 2009].

Při manipulaci s některými tituly poškozenými degradací papíru je nutné být maximálně opatrní, snažit se s dokumentem zacházet co nejšetrněji, některé exempláře již byly v takovém stavu, že při skenování zůstávali zaměstnancům v ruce celé kusy papíru, například při obrácení stránek se papír tak drotil, že po obrácení zůstala v knize díra odpovídající velikostí a tvarem prstům člověka, který listy obracel.

U některých starších dokumentů byly nalezeny ještě nerozřezané listy. Když byl dokument nepoškozený a jednalo se jen o jeden či dva listy, pak byly listy rozřezány v Oddělení digitalizace. Pokud byl titul příliš poškozený anebo se jednalo o větší množství listů, bylo nutné ještě před skenováním dopravit dokumenty do knihvazačské dílny, aby listy odborně rozřezali.

5.2.2.1 Grafický formát

Velmi diskutovaným problémem se v rámci procesu digitalizace po celém světě stává výběr vhodného grafického formátu. Vzhledem k tomu, že neustálý vývoj na poli digitalizace s sebou přináší změny ve standardizaci, je rozhodování, jaký formát zvolit, velmi těžké. Také rozhodnutí digitalizovat přímou metodou bez mezičlánku v podobě mikrofilmu bylo v době plánování projektu nezvyklé a neslo s sebou jiný způsob uvažování nad použitým grafickým formátem.

Každý formát má své specifické vlastnosti, a proto je velmi obtížné a prakticky nemožné určit nejlepší formát. Vždy záleží na kritériích, která jsou pro danou instituci a její účely podstatná [Vychodil, 2007, s. 97]. V případě MKP byly hlavními kritérii univerzálnost, datová náročnost, přesnost, požadavky na dlouhodobou ochranu (standardizace, zastaralost formátu, jeho podpora, typ dokumentace).

Na základě zkušeností z ČR i zahraničí postoupily do užšího výběru dva formáty – *JPEG* a *TIFF*³³. V době výběru grafických formátů vhodných pro účely Oddělení digitalizace MKP ještě nebylo v odborných kruzích řešeno ve větším měřítku nasazení formátu *JPEG 2000*, jehož velkou výhodou je nabídka většího množství dekomprese, umožňuje několik možností zobrazení a je odolnější proti chybám. Tento formát má ovšem velkou nevýhodu, není podporován všemi grafickými programy. Pro prohlížení obrázku ve formátu *JPEG 2000* je většinou nutné nainstalovat do počítače rozšíření umožňující jeho zobrazení.

³³ Formát PNG se ukázal být pro OD nepoužitelným, není podporován softwarem Sirius.

Formát *JPEG* (*Joint Photographic Experts Group*) se používá především pro fotorealistické obrazy, podporuje jen 16 miliónů barev nebo 256 odstínů šedi. Využívá ztrátovou kompresi za použití rovnoměrného rozdělení bloků složených z 8x8 pixelů a výsledná kvalita se liší v závislosti na nastavení kompresního poměru. Rušící prvky způsobené ztrátovou kompresí mají podobu čtvercových plošek. Používán je především na internetu a při běžné práci s digitálními fotografiemi. Formát *TIFF* (*Tagged Image File Format*) je široce využívaný formát pro tisk grafiky, podporuje jakékoli rozlišení a jakékoli množství barev. Výhodou je možnost využití řady kompresních možností (i bez komprese) a podpora více obrazů uložených v jednom souboru [KNOLL, 2008].

Oba formáty splňují kritérium univerzálnosti formátu. Pokud srovnáme datovou náročnost obou formátů, je velmi důležité vzít v úvahu jejich kompresní poměry. Při nejvhodnějším poměru kvalita/velikost, který je dosahován při kvalitě 70–80 %, je datová náročnost na výpočet u formátu *JPEG* menší než při využití nekomprimovaného formátu *TIFF*, který je z hlediska dlouhodobé archivace doporučován.

Co se týče přesnosti, je samozřejmě lepší formát *TIFF*, který umožňuje různé nastavení rozlišení a jakékoli množství barev.

Oba formáty splňují kritéria pro dlouhodobou archivaci, která doporučují použít formát se standardizovanou specifikací³⁴ [KNOLL, 2008]. Dále je důležitým kritériem, jaké softwarové aplikace či nástroje s daným formátem pracují. V tomto případě je rozšířenější formát *JPEG*, je také masivně podporován internetovými prohlížeči. Nezbytný prvek při rozhodování tvoří také podpora vývoje formátu, i zde *JPEG* vychází ze srovnání lépe, vývoj formátu *TIFF* se zastavil již v roce 1992. Z hlediska dlouhodobé archivace je nejdůležitějším kritériem dostupnost otevřené dokumentace formátu, toto kritérium splňují oba formáty.

Po shrnutí výše uvedených kritérií a především ujasnění účelu digitalizace dokumentů v projektu HISPRA, jehož hlavním účelem je zachování obsahové informace, byl vybrán formát *JPEG* jak pro archivaci, tak pro zpřístupnění. Vzhledem k tomu, že většina dokumentů určených k digitalizaci nenesou významné obrazové informace, jsou pro digitalizaci dostačující vlastnosti formátu *JPEG*.

³⁴ V případě formátu *JPEG* se jedná o normu ISO/IEC IS 10918-1:2005, v případě formátu *TIFF* je platná norma ISO 12639:1998.

5.2.2.2 Rozlišení

Pro nasnímání titulů byly vybrány následující hodnoty rozlišení. Zpočátku byly všechny exempláře snímány na 600 dpi. Později se při zpracování ukázalo, že velikost skenu u předloh, jež přesahují velikost A3³⁵, natolik vzroste, že tyto soubory není software Sirius používáný pro zpracování skenů schopen zpracovat. Při zpracování se vyskytly dva druhy chyb – buď byl Sirius schopen zpracovat pouze část obrazů a zbytek nezpracovaných souborů ponechal v původním stavu, nebo při zpracování ohlásil chybové upozornění a celou dávku poškodil, takže již nešla otevřít. Tato skutečnost byla konzultována s firmou *Elsyst Engineering*, která software Sirius vyvíjí. Po prozkoumání zaslaných logů³⁶ bylo techniky z firmy *Elsyst Engineering* zjištěno, že velikost zpracovaných skenů byla při použití dané výpočetní techniky pro software Sirius příliš velkou zátěží. Doporučením bylo dosáhnout menší velikosti zpracovávaných obrazů.

Proto bylo rozhodnuto, že předlohy, jejichž velikost přesáhne formát A3, je nutné skenovat pouze v rozlišení 300 dpi. Pokud je tedy do OD doručen dokument, který svou velikostí přesahuje formát A3, je po naskenování několika stránek v rozlišení 600 dpi vyzkoušeno, zda proces zpracování v softwaru Sirius probíhá korektně. Pokud ano, je možné pokračovat ve snímání předlohy na 600 dpi. Při rozhodování o rozlišení je třeba brát v úvahu nejen velikost dokumentu, ale také typ obrazu. U velikosti dokumentu se posuzuje, o kolik je předloha větší než formát A3. Pokud je rozdíl zanedbatelný, je pravděpodobné, že i při nasnímání ve větším rozlišení, je možné skeny v softwaru Sirius zpracovat. U typu obsahu je důležité odlišit, zda v dokumentu převažuje text či obraz. Pokud je v dokumentu více obrazových částí, bude celková velikost souboru po naskenování větší. A čím větší je velikost skenu, tím pravděpodobnější je výskyt potíží při zpracování v softwaru Sirius.

5.2.2.3 Barevná hloubka

Vzhledem k povaze dokumentů určených k digitalizaci bylo rozhodnuto skenovat v režimu škály šedi, a to v barevné hloubce³⁷ 8 bitů využívajících až 256 stupňů šedi.

³⁵ Formát A3 je standardizovaný formát papíru, splňuje normu ČSN EN ISO 216, papír má velikosti stran v poměru $1:\sqrt{2}$ (přibližně 1: 1,4142) a jeho rozměry jsou 297×420 mm [Typo.cz, 2009].

³⁶ Log je název pro záznam nebo soubor záznamů, které si některé programy vytvářejí pro ukládání informací o své činnosti a běhu. Logy slouží při zpětné analýze k rozpoznání, zda došlo k nějaké chybě a pakliže ano, pak pomáhají určit, k jaké chybě došlo a proč [Log, 2011].

³⁷ Barevná hloubka je termín používaný v počítačové grafice, který popisuje počet bitů použitých k popisu určité barvy nebo pixelu v bitmapovém obrázku nebo rámečku videa. Toto pojetí je také známé jako počet bitů na pixel, zejména je-li uvedeno spolu s počtem použitých pixelů. Větší barevná hloubka zvětšuje škálu různých barev a přirozeně také paměťovou náročnost obrázku či videa [VYCHODIL, 2007, s. 86].

Převážná většina dokumentů určených k digitalizaci v rámci projektu HISPPRA pochází z 19. a 1. poloviny 20. století a obsahuje především texty. Pokud předlohy obsahují obrazové informace, jedná se hlavně o černobílé fotorealistické obrazy či v případě kartografie o jednoduché plány. Protože hlavní zásadou při digitalizaci je zachování obsahu dokumentu, pro většinu předloh je dostatečné nasnímání v šedivé škále. V případě, že se v dokumentu vyskytují barevné obrazové informace, je posuzováno, jak velkou část z celku představují. V případě digitalizovaných dokumentů z fondu MKP není třeba zachovávat barevnou věrnost digitální kopie, jako například u starých tisků, které jsou zpřístupňovány v systému Manuscriptorium. Pokud se jedná o jednu či několik izolovaných objektů, pak jsou stránky obsahující barevné objekty naskenovány barevně. Jestliže tvoří barevné obrazové informace většinu dokumentu, pak je celý dokument nasnímán barevně, ovšem v rozlišení 300 dpi. Důvodem je velikost souboru, která by mohla činit potíže při zpracování v softwaru Sirius stejně jako dokumenty přesahující svou velikostí formát A3.

Jas a kontrast jsou individuálně nastavovány během skenování.

5.2.2.4 Postup při skenování

Při skenování jsou platná obecná pravidla, která upravují zásady, jaké části dokumentu skenovat. Samozřejmě i zde je třeba individuálně posuzovat jednotlivé případy, vzhledem k variabilitě a různorodosti skenovaných dokumentů je doporučeno konzultovat s kolegy a případně s vedoucí oddělení, zda sporné strany skenovat či nikoli. Toto pravidlo je nutné dodržovat především u nově přijatých zaměstnanců, kteří nemají takové zkušenosti jako ti, co již naskenovali a popsali několik set dokumentů.

Skenuje se první strana obsahující text, tedy první strana, která není prázdná. Desky včetně hřbetu se neskenují, většina z nich prošla převazbou, takže desky nejsou originální, a novodobé desky, které nebyly součástí původního dokumentu, není třeba digitalizovat. Neskenuje se ani přední ani zadní předsádka, stejně tak není nutné snímat první prázdné listy, které jsou součástí knižního bloku. Často se na prvním prázdném listu před titulním listem nebo před pativalem či impresem objevuje razítko knihovny. Zde je na úvaze toho, kdo skenuje, jestli zahrnout stránku s razítkem do skenování – většinou se nezahrnuje, pouze pokud obsahuje i jiné informace – například poznámky knihovníka, zde se rozhoduje o skenování individuálně. Například pokud je podle rukopisu či podpisu identifikována známá osobnost (*J. Thon* či *A. Sova*), tak se tato strana nasnímá. Analogicky je stejné pravidlo aplikováno i na konec dokumentu – skenují se jen poslední stránky s textem, pokud jsou

za nimi vložené další vakáty, tak ty se stejně jako zadní desky neskenují. Pokud jsou v dokumentu mezi tištěnými stránkami prázdné stránky zahrnuté do číslování, pak se tyto strany skenují. Nejčastěji se vakáty³⁸ vykytují jako oddělovač jednotlivých kapitol. Pokud dokument obsahuje více po sobě jdoucích prázdných stran a jsou zahrnuty do číslování, pak se skenují. Pokud nejsou zahrnuty do číslování, pak je možné je neskenovat, ovšem musí být dodrženo pravidlo střídání levé a pravé stránky.

Když kniha obsahuje záložku, pak je při skenování záložka vyjmuta z knižního bloku a umístěna mimo knižní blok tak, aby při skenování nepřekážela a na výsledném obrazu nebyla patrná.

Pokud dokument tvoří volné listy, které informaci nesou pouze na jedné rubové straně listu a u kterých je líc obvykle tvořen prázdnou stránkou, pak je třeba zjistit, zda jsou prázdné stránky číslovány nebo zahrnuty do číslování. Pokud tuto podmínku nesplní, je možné prázdné stránky naskenovat. Tyto znaky se objevují hlavně u velkoformátových fotografií, grafík a jiných kreseb. Například z fondu pragensií je najdeme nejčastěji u panoramatických fotografií Prahy a okolí, dále u některých typů teatrálií – především u fotografií interiéru divadel a portrétů umělců.

V následující části kapitoly jsou prezentovány postupy při skenování dokumentů na přístrojích Zeutschel a Bookeye, skener Minolta jsem již nepopisovala, vzhledem k tomu, že byl z projektu vyřazen již v roce 2008.

5.2.2.4.1 *Zeutschel*

Skener *Zeutschel OS5000 TT* umožňuje, jak již bylo popsáno v kapitole 2.5.3.1 *Zeutschel*, snímat předlohy pouze ve škálách šedi nebo v černobílém režimu. Původně byl tento skener využíván více, po zjištění problémů se skenerem Minolta byl dokonce používán na skenování všech dokumentů. Pouze stránky, které bylo nutné nasnímat barevně, byly naskenovány na skeneru Minolta. Naopak po výměně skenerů byl pro jednodušší obsluhu oblíbenější skener Bookeye, který se nyní využívá více.

Před začátkem skenování je nutné skener *Zeutschel* připravit a nastavit. Nejprve je třeba skener zapnout. Spouštění skeneru trvá delší dobu, pokud není spuštěn včas, je zobrazeno upozornění, že skener nebyl nalezen a dotaz, zda by měl být zapnut skener virtuální.

³⁸ Vakát je prázdná stránka dokumentu.

V případě, že se tak stane, je nutné zavřít program a znovu jej spustit. Jinak může dojít také k zaseknutí programu, poté je třeba postup několikrát opakovat.

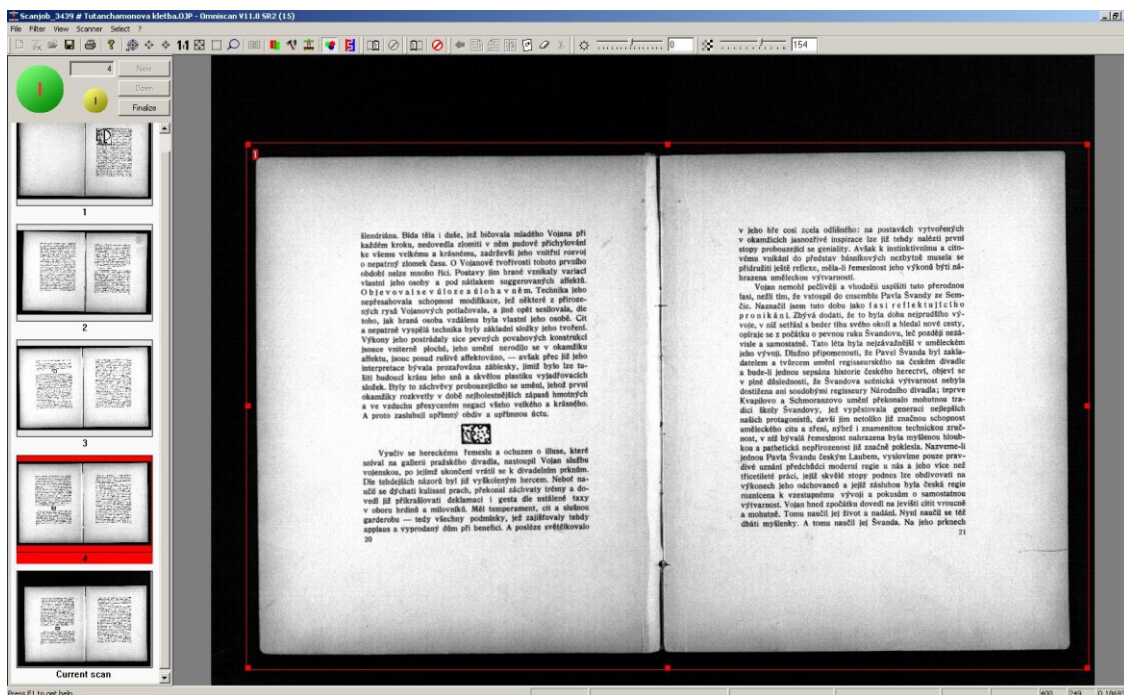
Po spuštění skeneru je důležité určit velikost snímané plochy podle velikosti předlohy. Dokument je položen na kolébku, kde je měřítkem označena velikost plochy v intervalu od 1 do 10. V tomto případě je důležité odečíst správnou velikost, pro jistotu je dobré zvolit číslo o jedno větší, aby zbylo dost prostoru pro manipulaci s dokumentem při skenování. Hodnota tohoto čísla se použije k nastavení vzdálenosti posuvné kamery od snímané plochy a také k nastavení ostřicího kroužku na kameře. Při nastavování ostřicího kroužku je zkontrolováno také nastavení hodnoty u clony. Ta je automaticky nastavena na hodnotu osm, která odpovídá většině dokumentů. Pokud je předloha světlejší, je hodnota clony upravena směrem dolů, pokud je tmavší, pak je třeba nastavit hodnotu vyšší než osm.

Po nastavení kamery následuje úprava kolébky, která se skládá ze dvou částí. Ty lze od sebe oddálit a tím uzpůsobit desky pro šířku hřbetu. K tomuto účelu slouží dvě aretační kolečka umístěná vpředu na skeneru. Menší kolečko slouží k nastavení celkové výšky kolébky a tím také vzdálenosti předlohy od kamery. Stále je ovšem možné upravovat výšku obou desek, která se při použití přítlačného skla přizpůsobí objemu rozevřené předlohy – výška desky se přizpůsobí měnící se výšce jednotlivé strany rozevřené předlohy podle toho, jak během skenování přibývají otáčené stránky na pravé desce kolébky a ubývají na desce levé. Pokud se při ručním ovládní skla tato výška nenastaví správně, pak je nutné vynaložit velkou sílu k posunutí kolébky na určenou vzdálenost. Při utážení většího kolečka zůstanou desky stabilně v jedné poloze.

Přítlačné sklo je možné přiklápět k deskám kolébky ručně nebo automaticky. Ruční přítlak má výhodu ve větší přesnosti, zatímco při automatickém přítlačení skla je vyvinut větší tlak a stránky dokumentu se více narovnají. Použití automatického nebo ručního přítlaku skla je individuální – záleží na jednotlivém zaměstnanci a na stavu předlohy. Pokud je například hřbet předlohy tuhý a text je vytištěn blízko hřbetu, je lepší využít automatický přítlak skla, který umožní větší rozevření hřbetu. Někdy je ovšem tlak příliš velký a je lepší pomalá ruční manipulace se sklem. Při zapnutí automatického přítlaku se sklo ovládá pomocí přidavných pedálů. Pedál se šipkou nahoru sklo zvedá a pedál s šipkou dolů sklo spouští. Při spouštění skla je důležité vyzkoušet reakci rychlosti spouštění na sešlápnutí pedálu. Sešlápnout pedál je nutné s citem, jinak je sklo spuštěno příliš rychle a ten, kdo skener ovládá,

nestačí vytáhnout ruce a sklem si přiklepe prsty. Třetí pedál s názvem *Aufnahme Exposure* slouží ke snímání.

Samotné nastavení snímání a jeho realizace je v prostředí programu *Omniscan* náročnější než v programu *BCS-2*, který umožňuje nastavení parametrů skenování i ukládání následováním menšího počtu kroků. Při skenování je možné využít přednastavené profily. Při běžném skenování je používán profil nastavený na snímání ve stupních šedi, dochází ke snímání celé plochy najednou (není rozlišována pravá a levá strana) a je použit jeden ořezový rámeček pro celou předlohu. Po výběru tohoto profilu je v rozšířeném nastavení vybráno číslo použité na ostřícím kroužku kamery. Poté je nasnímán zkušební sken a podle výsledku snímání jsou upraveny hodnoty jasu a kontrastu a je nastavena velikost ořezového rámečku³⁹ (viz Obrázek č. 7). Při zkoušení neoptimálnějšího nastavení je využívána možnost přeskenování již nasnímaného obrazu.



Obrázek 7: Skenování v SW Omniscan na přístroji Zeuschel [Omniscan]

Po ukončení skenování se nastavují parametry ukládání. Výsledné soubory jsou ukládány na lokální disk počítače, odkud jsou při dokončení procesu skenování kopírovány na určené

³⁹ Velikost ořezového rámečku je uzpůsobena stavu vazby a objemu předlohy. Pokud je předloha objemná, je nastaven ořezový rámeček větší, aby byl ponechán prostor pro manipulaci a při každém novém nasnímání nebylo nutné znovu upravovat velikost rámečku. Nevýhoda při úpravě velikosti rámečku nebo jeho posouvání je dána tím, že program po každé z těchto úprav ukládá sken znovu a tato operace trvá několik sekund. Při časté změně se výrazně prodlužuje doba snímání.

místo do adresáře *Skeny*⁴⁰. Zde se nastavuje cílová složka, skeny se přejmenovávají podle výše popsaných pravidel a jako koncový formát je zvolen *JPEG*.

5.2.2.4.2 *Bookeye*

Skener *Bookeye 3 A2* je v současnosti nejoblíbenějším skenerem v Oddělení digitalizace. Je také nejuniverzálnějším skenerem, na rozdíl od modelu Zeutschel umožňuje snímat i v barevném spektru. Velmi věrně také zachycuje fotografie a jiné grafické prvky v textu.

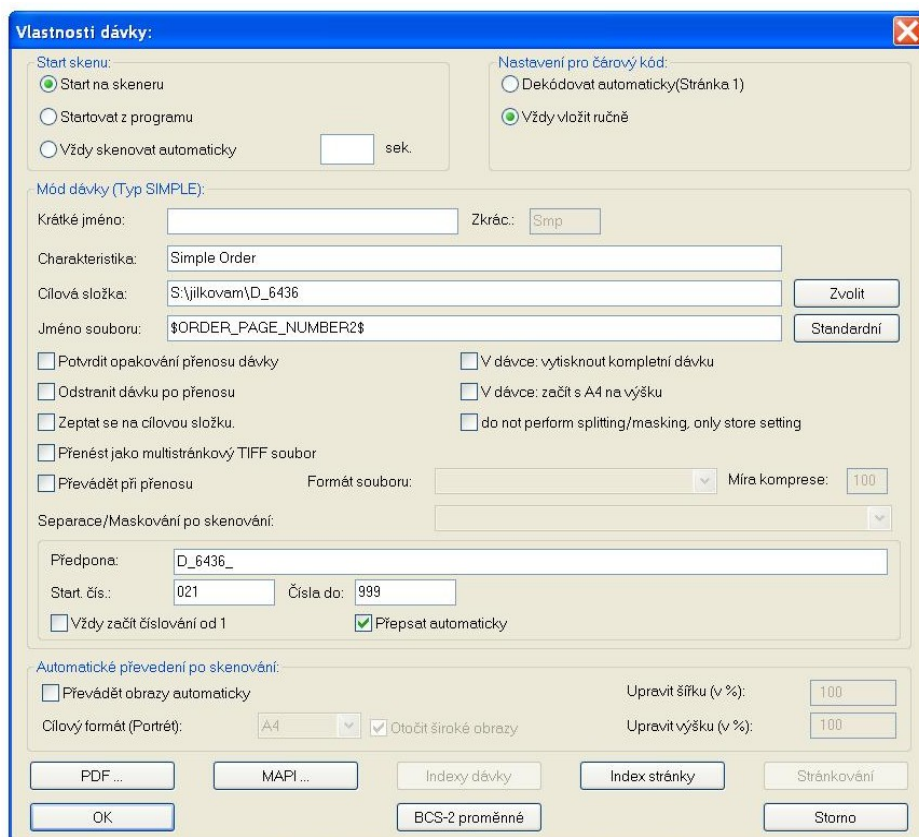
Před skenováním je třeba stejně jako u jiných skenerů nastavit jak skener, tak parametry snímání. Na rozdíl od skeneru Zeutschel, který je vybaven posuvnou kamerou, u modelu *Bookeye* je snímací zařízení umístěno nad snímanou plochou ve stabilní poloze. Vzdálenost od snímané plochy je detekována automaticky pomocí laseru – odpadá tedy nutnost ručního nastavování. I u tohoto přístroje trvá start od zapnutí do stavu připravenosti poměrně dlouhou dobu, velmi důležité je nemanipulovat při startu s tlačítky, může dojít k resetování síťového nastavení.

Po zapnutí je možné nastavit kolébku podle předlohy. Kolébka se skládá ze dvou desek, které jsou ovládány nikoli ručně, ale elektronicky pomocí sady tlačítek umístěných na přední straně skeneru. Uprostřed je čtyřsměrné tlačítko, kterým jsou regulovány obě desky najednou, tlačítka po stranách je ovládána každá deska zvlášť. Desky je možné posouvat vertikálně či horizontálně. Nelze ale měnit úhel sklonu desek, je umožněno pouze skenování rozevřené předlohy pod úhlem 180°. Při snímání předlohy pod nižším úhlem rozevření je nutné předlohu přidržovat v požadované poloze ručně. Výšku desek je vhodné přizpůsobit po sklopení přitlačného skla tak, aby se listy předlohy co nejvíce vyhladily a předloha byla co největší plochou přitlačena ke sklu. Vpředu na skeneru je možné nastavit některé z parametrů skenování: formát neboli velikost snímané plochy (*Format*), typ snímání plochy – vcelku nebo samostatné stránky (*Pages*) a volbu módu (*Mode*) – tedy barevně, stupně šedi, černobíle a mód pro fotografie. U všech alternativ je možné skenovat s přitlačným sklem či bez něj. A v neposlední řadě je možné na skeneru ovládat také jas.

Po spuštění programu *BCS-2* je nutné na jednom místě nastavit parametry pojmenování a ukládání skenu ještě před začátkem skenování. I zde jsou skeny ukládány primárně

⁴⁰ Právě nastavování parametrů kopírování je zdlouhavé, ovšem v několika případech byl postup výhodný. Například při zpracování skenů téhož dne, kdy byla předloha nasnímána, došlo omylem k jejich smazání, ovšem ještě nebyly vytvořeny zálohy. Předlohu nebylo nutné skenovat znovu, protože soubory byly uloženy na lokálním disku.

na lokální disk, odkud jsou po skončení skenování kopírovány do příslušné složky v adresáři *Skeny*. Ovšem u modelu Bookeye je jednodušší a rychlejší nastavení těchto parametrů.



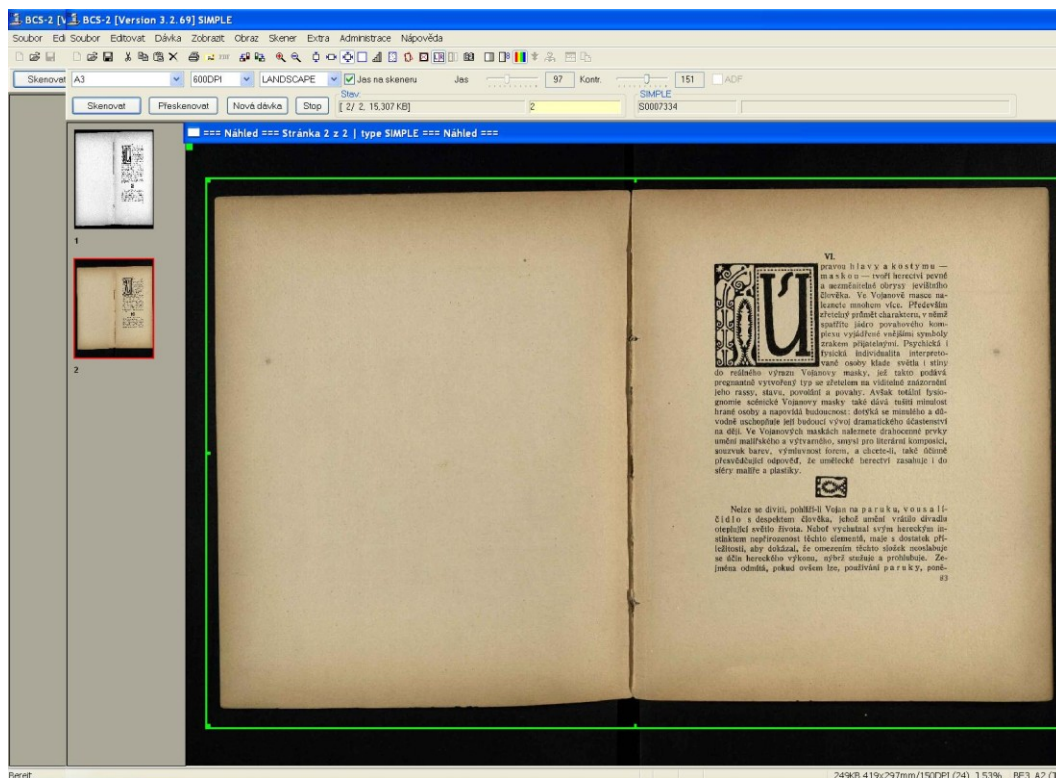
Obrázek 8: Okno s nastavením parametrů skenování v SW BCS-2 [BCS-2]

Každý pracovník si může vybrat, zda je mu příjemnější ovládat proces skenování tlačítky na skeneru nebo pouze tlačítky z programu (viz Obrázek č. 8). Pokud je zvolena možnost Startovat z programu, tlačítka vpředu na skeneru nereagují. Také je možné zvolit spuštění snímání automaticky po vybrané době v sekundách, ovšem tato možnost není příliš využívaná, doba manipulace a nastavování se u stejné předlohy liší stránku od stránky⁴¹.

Následuje nastavení parametrů skenování. Primárně je nastaven mód skenování na hodnotu *Simple*. Dále je určena velikost snímané plochy. Kromě klasických formátů je možné definovat vlastní uživatelský formát. Tato možnost je využívána u atypických formátů předloh, kdy je zbytečné, aby byla při každém skenování snímána větší plocha, než je nezbytně nutné. U klasických formátů je možné nastavit orientaci snímání – na výšku či na šířku. Dále je u každého skenování posuzováno nastavení rozlišení. Do velikosti A3 včetně se

⁴¹V nastavení je nutné věnovat pozornost číslování skenů, kde je možné zvolit variantu *Vždy číslovat od 1*. Při skenování stejné předlohy v několika dávkách je nutné zkontrolovat nastavení číslování, jinak jsou u druhé dávky všechny skeny číslovány znovu od čísla jedna a skeny z první dávky jsou novými skeny přepsány.

skenuje na 600 dpi, jinak na 300 dpi (viz kapitola 5.2.2.2 Rozlišení). V neposlední řadě jsou nastaveny hodnoty jasu a kontrastu.



Obrázek 9: skenování v SW BCS-2 na přístroji Bookeye [BCS-2]

Samotné skenování je velmi jednoduché – stačí provést snímání stisknutím tlačítka *Start* na skeneru nebo při ovládání skenování pomocí programu stačí stisknout příslušnou ikonu v programu BCS-2. Přeskenování je možné provést pouze pomocí ikony ve skenovacím programu, samotné přeskenování je realizováno pomocí tlačítka na skeneru. Poté, co je přeskenování provedeno a pracovník chce pokračovat ve skenování dál, je nutné znovu v programu BCS-2 označit ikonu Skenovat (viz Obrázek č. 9). Po vzniku prvního ostrého skenu je nastavena velikost ořezového rámečku a moment, kdy bude ořez aplikován – po skenování nebo před následujícím skenem. Parametry skenování – například barevnost skenu, je možné měnit i během skenování.

Po ukončení snímání stačí zvolit ze seznamu dávek poslední z nich, kterou je třeba přenést, a soubory jsou přeneseny na zvolené místo v adresáři *Skeny*.

5.2.3 Tvorba metadat

Tvorba metadata je nedílnou součástí procesu digitalizace. Metadata jako strukturované informace o jakémkoli informačním zdroji jakéhokoli typu či formátu [STATE, 2007] jsou

klíčovým konceptem pro jakoukoli práci s daty obecně [CUBR, 2009, s. 82]. Existují různé typologie metadat, jedna z nich rozděluje metadata na popisná, strukturální a administrativní [Gartner, 2008, s. 5]. Popisná nebo také deskriptivní metadata „charakterizují původní objekt tak, aby byl uživatel schopen porozumět jeho obsahu. Podporují vyhledávání daného digitálního objektu“ [CELBOVÁ, 2003c-]. Strukturální metadata zachycují vnitřní strukturu digitálního objektu a tím umožňují lepší orientaci uživatele, často jsou nezbytná pro digitální objekty složené z velkého množství souborů [Gartner, 2008, s. 6]. Administrativní metadata jsou „používána pro řízení a správu informačních zdrojů“ [CELBOVÁ, 2003a-]. Často jsou ještě rozdělena na metadata technická, právní a archivační.

5.2.3.1 Popisná metadata

Bibliografické údaje jsou importovány z automatizovaného knihovního systému (dále AKS) Koniáš do softwaru Sirius. Tento software umožňuje importovat metadata přímo z knihovního systému pomocí protokolu Z 39.50⁴². AKS Koniáš bohužel tento protokol nepodporuje, takže bylo nutné vymyslet jiný způsob, jak bibliografické údaje importovat. Bylo tedy nezbytné vytvořit konverzní tabulky⁴³, které převod umožní. Na vytvoření těchto tabulek se podíleli E. Měřinská, D. Krpata, Z. Kopencová a I. Ferdinandová. Cílem bylo určit, která pole v *AKS Koniáš* odpovídají polím v SW Sirius. Problémové bylo například převedení pole Autor, po dohodě bylo nastaveno, že deskriptory z AKS Koniáš se importují do SW Sirius jako klíčová slova, z importu jsou vynechány formální deskriptory⁴⁴.

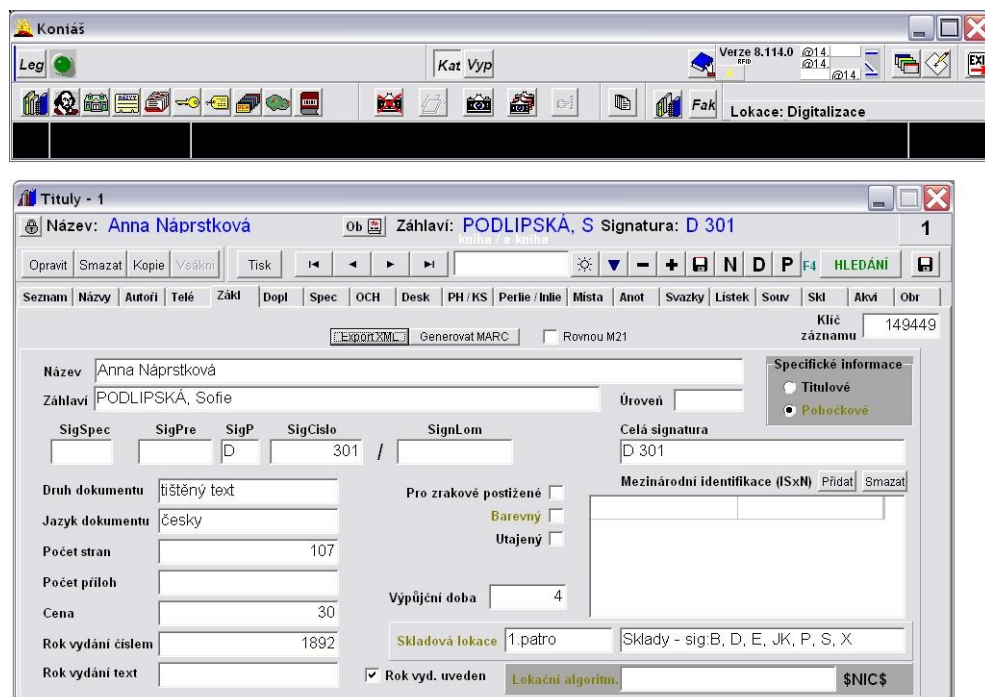
Konverzní tabulky byly začleněny do *AKS Koniáš*, což se projevilo přidáním ikony *Export XML* do záložky Základní (viz Obrázek č. 10). Po stisknutí této ikony je na pozadí provedena konverze, jejímž výsledkem je soubor *XML* obsahující bibliografické údaje ve formátu *DTD*⁴⁵ pro monografie (verze 2.10). Tento soubor je ukládán do paměti PC na zvolené místo, V programu Sirius je tento soubor předdefinován jako cílový soubor, odkud jsou příslušná metadata importována.

⁴² Protokol Z 39.50 je určen „pro vyhledávání informací v rámci knihovních aplikací, speciálně pro podporu vyhledávání z distribuovaných serverů. Umožňuje provádět z aplikace na jednom počítači dotazy do databáze na jiném počítači“ [CELBOVÁ, 2003e-].

⁴³ Konverzní tabulky slouží k přepisu dokumentu z jednoho kódování do druhého. Cílem převodu je tedy nahradit kód zdrojového kódování vyjadřující určitý znak kódem, který má stejný význam v cílovém kódování [Konverzní, 2010].

⁴⁴ Formální deskriptory vyjadřují formu dokumentu, například výklady, příručky, slovníky, romány, povídky atd.

⁴⁵ *DTD (Document Type Definition)* představuje „ve formátech SGML a XML formální popis složek (součástí) určitého dokumentu nebo třídy dokumentů. DTD pomocí formální syntaxe přesně popisuje, které elementy a entity se mohou v dokumentu vyskytovat a jaký je obsah a atributy těchto elementů“ [CELBOVÁ, 2003b-].



Obrázek 10: Bibliografický záznam dokumentu Anna Náprstková v AKS Koniáš [Koniáš]

Import je možné realizovat buď před začátkem skenování, nebo po naskenování dokumentu. Záleží na každém pracovníkovi, jakému způsobu dává přednost. Často je dobré importovat metadata ještě před skenováním, kdy je z bibliografického popisu možné zjistit podrobnější informace o knize. Například, že se jedná o jeden svazek z vícesvazkového díla. Většinou jsou do digitalizace zařazeny všechny části vícesvazkového díla, tedy pokud jsou ve fondu k dispozici. Zde je potom vhodné, aby všechny díly zpracovával jeden zaměstnanec tak, aby konzistence ve způsobu zpracování zůstala zachována.

Poté, co je dokument označen za rozpracovaný jak v papírovém seznamu, tak v elektronické tabulce, je podle signatury, čárového kódu nebo kombinací slov z názvu a autora vyhledán v knihovním systému Koniáš. Jsou procházeny všechny relevantní záložky, úplnost bibliografického popisu je zkontrolována zobrazením velkého katalogizačního lístku. Pokud nejsou shledány vážné nedostatky, jsou bibliografické údaje importovány z AKS Koniáš pomocí ikony *Export XML* v záložce *Základní*. Export probíhá na pozadí, je ovšem třeba zkontrolovat, zda export proběhl v pořádku. Pokud se při zpracování vyskytne nějaká chyba, je třeba znovu zkontrolovat záznam dané signatury v AKS Koniáš a zkusit export znovu. Pokud ani opakovaný export není úspěšný, je nezbytné kontaktovat Oddělení informačních technologií MKP (dále v textu OIT) a závadu nahlásit. Častou příčinou

nezpracovaných exportů je chyba v nastavení konverzních tabulek⁴⁶ nebo výpadek v činnosti serveru. V Příloze č. 2 je uveden bibliografický záznam ve formátu XML vyexportovaný z AKS Koniš.

Po provedení importu do databáze softwaru Sirius, je vytvořen nový pořadač s názvem signatury, syntax signatury je nutné upravit podle pravidel – mezera mezi jednotlivými částmi a lomítko se nahrazují znakem podtržítka, pomlčka zůstává zachována. Po úpravě signatury je třeba zkontrolovat přesnost a úplnost nahrání bibliografických údajů do softwaru Sirius. Zde je nutné především ověřit správnost importu údajů týkajících se fyzického popisu, zejména rozsah stran dokumentu. Pokud jsou v rámci rozsahu uvedeny kromě počtu stran také další fyzické údaje, rozměr nebo doprovodný materiál, stává se, že tyto informace nejsou importovány. Dalším problematickým údajem jsou chybné autorské role. Zde je rozdíl mezi pojmenováním vedlejších autorských rolí v programu Sirius a v AKS Koniáš. Oba programy využívají odlišnou bázi rolí. Například v AKS Koniáš nejsou autorské role tak odlišené jako v programu Sirius, který nabízí širokou škálu těchto rolí. Třetí nejčastější údaj určený k revizi je souběžný název dokumentu. Se souběžným názvem se nesetkáváme často, ovšem při převodu údajů se chybně importuje do pole podnázev, odkud je třeba jej v softwaru Sirius přesunout do pole určeného pro *Souběžný název*. V pilotní fázi projektu jsem byla zástupcem ostatních zaměstnanců a z pozice studentky oboru Informační studia a knihovnictví jsem řešila s vedoucí OD problémy při importu údajů a především správnost údajů v bibliografickém záznamu.

Další úroveň popisu jednotlivých stran přichází na řadu po úpravě obrazu, tedy ořezání skenů. V softwaru Sirius se tato fáze nazývá indexování. V této fázi se popisují jednotlivé strany předlohy a je zde řešena také struktura dokumentu (viz další podkapitola). Jednotlivé strany dokumentu jsou označeny číslem strany a typem strany. Čísla stránek je možné vložit jednotlivě nebo hromadně u označených stránek. Při ručním vpisování čísel stránek do označeného pole je nezbytné číslo strany potvrdit stisknutím klávesy Enter, jinak se číslo v daném poli nezobrazí a při kontrole je možné přehlédnout již vyplněné pole a číslo vyplnit znovu.

Pro lepší přehlednost uvádím v následující části této kapitoly obecná pravidla popisu dokumentů v Oddělení digitalizace. I já jsem se podílela na úpravě těchto pravidel.

⁴⁶ Například nová role tvůrce díla, která byla v původním nastavení opomínuta.

Obecná pravidla popisu vychází z Pravidel popisu pro tvorbu metadat periodik a monografií používaných v NK ČR, „která jsou určena pro velkoobjemový automatizovaný způsob zpracování dokumentů v systému Sirius“ [NÁRODNÍ, 2007, s. 4]. Tato pravidla byla pro účely digitalizace v MKP upravena tak, aby více vyhovovala vlastnostem a znakům, které reflektují odlišnosti fondu MKP určeného k digitalizaci.

Pravidla popisu jsou určena pro tvorbu metadat podle DTD verze 2.1, administrativní metadata jsou vytvářena podle verze 1.2.

Mezi hlavní zásady patří zákaz používání uvozovek – přestože jsou někdy uvozovky uvedeny v tištěné předloze, při popisu se nepoužívají. U všech textových informací je nutné dodržet kontinuitu názvů – ty jsou uváděny stejně jako v předloze, tedy češtinou užívanou v době tisku dokumentu. Zde je nutné dávat pozor při přebírání údajů z *AKS Koniáš*, kde je v některých případech uveden název podle pravidel současného jazyka, ten je třeba nahradit názvem podle předlohy. Názvy se neukončují žádným znakem, za poslední slovo z názvu tedy tečka nepatří. Pokud je název složitější, členěn na více částí, pak je možné mezi jednotlivými částmi uvést znak použitý v předloze, tedy například tečku. Zde je ale třeba dbát na dodržování pravidel při použití znaků (viz výše – nepoužívání uvozovek). Součástí názvu dokumentu může být v některých případech také jméno autora, pokud je tak jednoznačně možné určit dle předlohy, je v metadatech uveden v názvu také autor.

Jiný způsob popisu je užit u pole autor, kde je jméno autora uvedeno ve formátu podle současných pravidel, tedy v *kanonické* neboli *autoritní podobě*.

Při popisu monografií jsou využívány pro popis stránek tyto typy stran:

- a) *Titulní strana* obsahuje základní informace o monografii. Pokud tyto informace nejsou uvedeny jinde než na deskách dokumentu, jsou desky označeny jako titulní strana. Při výskytu titulních stran v různých jazycích je uvedena česká jazyková verze jako titulní strana, ostatní jazykové verze se označí jako normální strany. Grafický titulní list nebo patitul jsou uvedeny také jako normální strana.
- b) *Obsah*
- c) *Prázdná strana*
- d) *Inzerce* je užitá v případě, že strana obsahuje vloženou reklamu
- e) *Normální strana* je použita pro všechny ostatní strany, které nejsou popsány jiným typem stránky.

- f) *Rejstřík*
- g) *Seznam ilustrací, seznam map, seznam tabulek*
- h) *Tabulka* a další méně obvyklé typy stran.

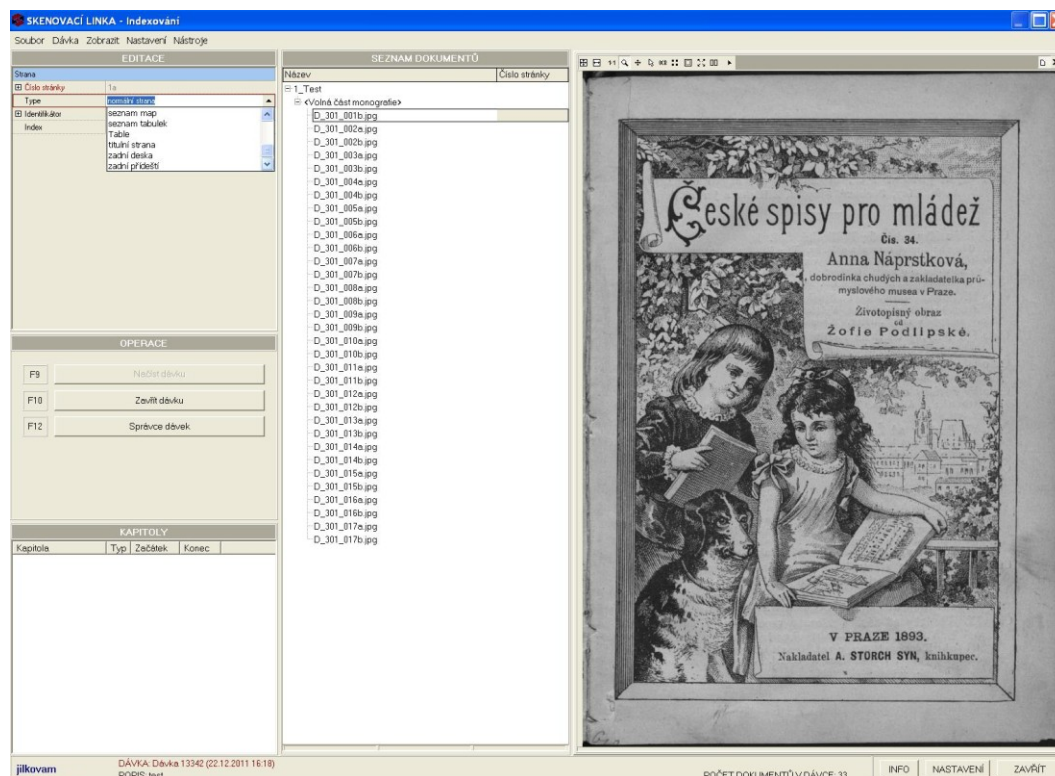
Číslování stránek, kapitol, dílů, obrázků či tabulek je striktně popisováno podle předlohy. Ne vždy je možné čerpat číslování pouze z předlohy, často se vyskytují v číslování variace či jiné vady, takže i pro vkládání čísel je možné sestavit všeobecná pravidla:

- všechna čísla stránek, která nejsou vytištěna v předloze a která je nutné doplnit, se uvádí v hranatých závorkách
- v dokumentu je nalistována první číslovaná strana, poté se zpětně čísluje do jedné, pokud se před stranou označenou jedna vyskytují ještě další strany, používá se označení kombinace čísla a písmena – začíná se od počátku číslování i abecedy, tedy 1a, 1b, 2b atd. (když dojde abeceda, pokračuje se přidáním druhého písmena – 1aa, 1ab, 1ac...1ba, 1bb, 1bc, ...) ⁴⁷ (viz Obrázek č. 11)
- při nesprávně uvedené paginaci se do hranatých závorek za číslo strany podle předlohy následované mezerou přidá dopočítaný údaj ⁴⁸
- číslo strany není natištěno a nechybí v číselné řadě, použije se kombinace čísla z předchozí nebo následující strany (podle obsahu popisované stránky, když je obsahově spjatá s následující stranou, je uplatněno číslo následující strany, v jiných případech se použije číslo předchozí strany) a písmene abecedy v hranatých závorkách ⁴⁹
- u obsahově neoddělitelných dvoustran jsou uvedena dvě čísla stránek, každé do vlastního řádku
- číslování pod obrázky, tabulkami a mapkami, obvykle uvedené malými římskými či arabskými čísly nenavazujícími na paginaci stránek, není bráno jako čísla stránek, se stranou se zachází jako s vakátem

⁴⁷ Například v praxi je první číslovaná strana 4, před ní je šest nečíslovaných stran, výsledné číslování je tedy: [1a], [1b], [1c], [1], [2], [3], 4.

⁴⁸ Například když je v předloze vytištěné číslo 24, ale podle návaznosti paginace je jasné, že správný údaj by měl být 26, tak je v poli číslo strany uvedeno 24 [26].

⁴⁹ Po stránce číslo 117 následují dvě neočíslované stránky a číslování dále pokračuje číslem 118. Tyto dvě neočíslované strany nejsou obsahově spjaty s předchozí ani následující stranou, proto bude uvedena paginace [117a] a [117b].



Obrázek 11: Vložení čísel stránek v SW Sirius [Sirius]

Samozřejmě i u číslování stránek platí pravidlo číslovat co nejlogičtěji, zachovat konzistenci číslování a co nejvíce ulehčit uživateli orientaci v dokumentu.

5.2.3.2 Strukturální metadata

Strukturální metadata zachycují strukturu daného digitálního dokumentu. Díky nim je možné se snáze orientovat v dokumentu. Jejich význam roste především u dokumentů zachycujících rozsáhlé dokumenty s bohatou vnitřní strukturou. Příkladem jsou sborníky s jednotlivými příspěvky. Strukturální metadata jsou vytvářena také v softwaru Sirius, který je využíván pro úpravu zpracování digitálních obrazů. Po importu do systému Kramerius je možné metadata zobrazit ve formátu *METS*⁵⁰.

Při popisu struktury dokumentu je nejdůležitější rozlišení dvou základních úrovní – *vnitřní části monografie a reference*. Rozdíl je patrný především v rovině zpřístupňování, kde je každý typ struktury zobrazen v digitální knihovně Kramerius odlišně. *Reference* se vkládají do dokumentu tehdy, pokud je dokument členěn na logické části, které nestojí samostatně (jednotlivé kapitoly, předmluva významné osobnosti atd.) Tyto části se podrobněji

⁵⁰ METS (angl. *Metadata Encoding & Transmission Standard*) je metadatový formát určený pro výměnu digitálních objektů mezi systémy jako kontejnerový formát umožňuje export části obsahu ze systému [Polišenský, 2007, s. 1; FOLTÝN, 2008-2011].

popíší a přidá se k nim rozsah stran. V digitální knihovně jsou zobrazeny pod výpisem čísel stránek a slouží k lepší orientaci v obsahu dokumentu. *Volné části monografie* se při popisu uplatňují tehdy, když je dokument členěn na díly nebo je dokument součástí většího celku, většinou vícesvazkového díla (popisovaný svazek je jedním z více svazků tvořících dohromady celek). V digitální knihovně je pod základními bibliografickými údaji popisujícími dílo jako celek uveden seznam volných částí monografie, ze kterých se tento celek skládá. Volné části jsou označeny jako svazek, jsou také popsány údaji identifikujícími tuto část a teprve za nimi následuje výčet stránek obsažených v této části. V počátcích pilotního provozu OD jsem objevila existenci volné části monografie a také jsem provedla šetření, jakým způsobem se volné části monografie v SW Sirius popisují a upravila pravidla popisu pro potřeby OD.

Podle identifikovaných úrovní rozlišujeme strukturu monografií na tři základní typy:

1. Monografie, která nevycházela v dílech, není rozdělena na kapitoly a má průběžnou paginaci (stránkování). U této monografie není nutné popisovat žádné části, popisuje se pouze paginace a typ stránky.

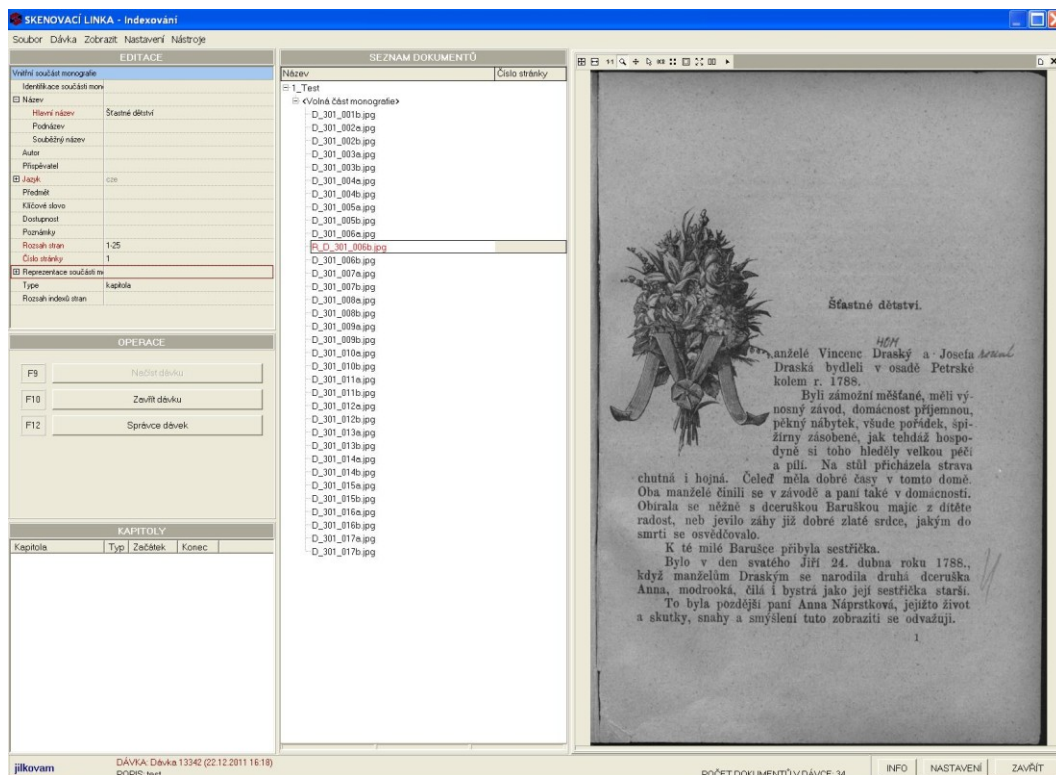
2. Monografie, která nevycházela v dílech a je členěna po kapitolách (bez ohledu jaká je paginace kapitol). U vícevrstvé struktury je na úvaze jednotlivého zaměstnance, jakou úroveň vybere pro popis základní části.⁵¹ Za kapitolu jsou považovány i jinak pojmenované struktury plnící stejnou funkci, mohou či nemusí být součástí obsahu. Je s nimi při popisu zacházeno jako s kapitolami. Každá kapitola neboli reference se popisuje zvlášť:

- vyplňuje se identifikace součásti – údaj identifikující kapitolu v rámci předlohy, odpovídá tomu, co je napsáno v předloze
- označení části – např. číselné, za číslem není udávána tečka

⁵¹ Zde je důležité, jakou strukturu a obsah má dokument. Ne vždy je při popisu zvolena ta nejvyšší úroveň. Pokud je struktura několika set stránkového dokumentu dvouvrstvá či třívrstvá, je nutné se zamyslet, jakou informaci chceme uživateli poskytnout. První vrstva je často tvořena kapitolami nazvanými obvykle římskými či arabskými číslicemi a množství kapitol rozčleněných v první vrstvě je malé – kolem 3 až 5 kapitol. Poté následuje druhá vrstva obvykle označená názvem upřesňujícím obsah dokumentu, je většinou členěna na více částí než první vrstva. Tuto druhou vrstvu obvykle označujeme referencemi. Je to nejvýhodnější řešení z pohledu uživatele, který má díky tomu představu, jaký je obsah jednotlivých kapitol, a snadněji se v textu orientuje. V případě, že je přítomna ještě třetí či čtvrtá vrstva, bere se v úvahu znovu hledisko uživatelské, ovšem většinou jsou již tyto vrstvy velmi obsáhlé a omezují přehlednost dokumentu, tudíž nejsou častými částmi, které se označují referencemi. Vnitřní součást monografie se tedy popisuje vždy, kdy to rozvrstvení textu dovolí. Opačná situace nastává u starších titulů z 2. poloviny 19. století, které jsou členěny na velké množství kapitol – jedna kapitola odpovídá jedné či dvěma stranám textu. Třicetistránkový dokument je tak tvořen dvaceti kapitolami, často uvedenými v obsahu. V tomto případě je vkládání referencí přímo na škodu, v dokumentu je obsah označen odpovídajícím typem strany a kapitoly nejsou označeny referencemi.

- hlavní název, případně podnázev kapitoly – zde je za hlavní považován název kapitoly tak, jak je uveden v obsahu, delší názvy či podnázvy je možné zkrátit a doplnit mezerou s třemi tečkami (nesmí být ovšem porušena srozumitelnost názvu)
- autor – je-li znám a je-li odlišný pro každou kapitolu
- třípísmenný kód jazyka – je-li odlišný pro každou kapitolu, jazyk je standardně nastaven na češtinu, existuje možnost vyplnit více jazyků
- rozsah stran kapitoly – od které do které strany je daná kapitola, pokud se číslování neshoduje, je dána přednost reálnému rozsahu než rozsahu uvedenému v obsahu (viz Obrázek č. 12)

Reference se vkládá vždy na stránku, na které daná kapitola nebo obdobná popisovaná část začíná. Když kapitola začíná na samostatné straně, vkládá se *reference typu A*. Někdy ovšem na jedné straně končí jedna kapitola a na té samé straně začíná kapitola jiná. Pro tyto účely je možné vložit *referenci typu B*. Reference je od běžné stránky odlišena červenou barvou písma. Základem pojmenování je název stránky, na kterou je reference vložena, doplněný o písmeno R na počátku. Syntaxe názvu je tedy *R_název_stránky.jpg*.



Obrázek 12: Okno modulu Indexování v SW Sirius [Sirius]

3. Monografie, která má samostatné svazky nebo díly spojené do jednoho dokumentu, kdy každý díl nebo svazek se popisuje jako volná součást monografie. Tento samostatný díl nebo svazek může být ještě členěn do kapitol. Při popisu dokumentu je třeba se soustředit na to, zda je svazek, který má zpracovatel v ruce, součástí většího celku či nikoli. Tato skutečnost se někdy určuje velmi obtížně. Je tedy velmi důležité si všimnout jakýchkoli indicií ukazujících na vícesvazkové dílo. Přestože jsou tituly kontrolovány, bohužel se stalo, že každý svazek byl poslán jinou dávkou a tím pádem nebyly svazky identifikovány jako volná část monografie. Tuto skutečnost lze většinou odečíst při důsledné kontrole správnosti údajů v Koniáši. Skutečnost, že některé tituly je nutné zpracovat jako volné části monografie, jsem při zpracování objevila já. Pro lepší podchycení této možnosti jsem také navrhla, aby vícesvazkový titul popisoval jediný pracovník tak, aby konzistence údajů byla zachována. Dalšími příklady využití *volné části monografie* jsou přitisky⁵² a přívazky⁵³, které se vyskytují především u staršího fondu. Toto pravidlo popisu bylo přidáno až během rutinního provozu pracoviště, při prvním výskytu přitisku v dokumentech určených k digitalizaci. Přítisk byl dokonce nalezen až při skenování, do té doby nebyl v dokumentu identifikován – jednalo se o pár stránek textu, na první pohled bylo možné jej zaměnit s kapitolou. Při bližším zkoumání v rámci popisu této *kapitoly*, bylo objeveno, že jde o přítisk. Obě varianty jsou v rámci popisu řešeny jako *volné části monografie* tak, aby byla zachována celistvost původního díla. Nevýhodou je, že přívazek nebo přítisk není možné vyhledat klasickou cestou prohlížení podle názvu monografie, oba tituly je možné najít pouze přes fulltextové vyhledávání.

V programu Sirius je automaticky u každé dávky načtených skenů vkládána volná část monografie, teprve do ní jsou vnořeny jednotlivé stránky neboli skeny. Při popisu je důležité pamatovat na odstranění volné části monografie tam, kde není potřeba. Smazat volnou část monografie je možné kdykoli během procesu zpracování před uvolněním dávek do Siria. Při mazání je velmi důležité přesunout všechny skeny neboli stránky nad *volnou část monografie* a teprve poté ji smazat. Při nedodržení tohoto postupu dojde s volnou částí monografie ke smazání také všech vložených stránek. Naštěstí po zavření dávky a nevyplnění povinných polí se zobrazí okno s upozorněním na absenci informací v těchto polích

⁵² Přítisk je obsahově uzavřené dílo, přitištěné v jedné tiskovině k jinému samostatnému dílu, ale nespojené s ním souborným názvem. Zvláštní název přitisků může být uveden na společné titulní stránce nebo na zvláštním titulním listu [UHLÍŘ, 2003-].

⁵³ Přívazek je samostatná bibliografická jednotka, rukopisná nebo tištěná, přivázaná k jinému, mnohdy obsahově odlišnému dílu, s nímž tvoří jeden svazek označovaný jako konvolut [RICHTEROVÁ, 2003-].

a software neumožní dávku uzavřít. Povinná pole jsou v softwaru Sirius označena červenou barvou písma.

Při popisu se postupuje podle toho, zda předloha má či nemá obsah:

- nemá obsah a není na první pohled vidět jednotlivé části – dokument se zpracovává bez toho, že by byla vkládána reference
- nemá obsah a je rozdělena na mnoho částí – např. na jedné stránce jsou 2 kapitoly – v tom případě se posoudí z uživatelského hlediska, zda má či nemá smysl, aby kapitoly byly označeny
- nemá obsah, ale je tam přiměřené množství kapitol – vkládá se *reference*, případně je rozdělena na díly – vkládá se *volná část monografie*
- má obsah, ale je rozdělen na velké množství kapitol – zpracovává se, jako by kniha neměla žádné kapitoly, při existenci více úrovní v obsahu je rozdělení řešeno podle uživatelského hlediska, je vybrána jedna úroveň (většinou střední) a na ni se vkládají *reference*
- má obsah – postupuje se podle obsahu – vkládají se *reference* na hlavní kapitoly, případně je použita *volné části monografie*.

Pokud jsou jakékoli pochybnosti ohledně interpretace obsahu, je dobré se poradit s ostatními zaměstnanci případně s vedoucí oddělení, nejdůležitějším pravidlem je uživatelské hledisko, tedy usnadnění orientace v dokumentu.

Po ukončení indexování a definitivním uzavření dávky do systému, které se nazývá také uvolnění dávky do systému, se vkládá index stránky. Slouží k identifikaci jednotlivých stránek. Číslo stránek k identifikaci nestačí, i v jednom dokumentu se mohou objevit stránky se stejnou paginací – například vlivem chyby v tisku. Index stránky je tvořen arabskou číslicí, číslování začíná od nuly a pokračuje podle potřeby, tedy objemnosti titulu. Vložení indexu je sice proces automatický, ovšem automatické vložení je možné pouze u jednotlivých stránek. Aby byl proces identifikace úplný, je nezbytné vyplnit také rozsah indexů stran u každé reference. U reference zjistíme reálný rozsah stránek v rámci paginace, tedy na jaké straně je začátek a konec reference. U těchto určitých stránek se odečte, jaký má každá z těchto stran index stránky, teprve tato čísla tvoří rozsah indexů

stran⁵⁴. Tímto způsobem se ručně musí vypsát rozsahy indexů stran u každé kapitoly. Korektně vyplněný index stran hraje významnou roli při vizualizaci v digitální knihovně. Index stran umožňuje zobrazení správného počtu stran v rámci reference – díky tomu každá kapitola začíná i končí na náležité stránce a rozsah stran v rámci kapitoly zobrazuje jen strany obsažené v kapitole.

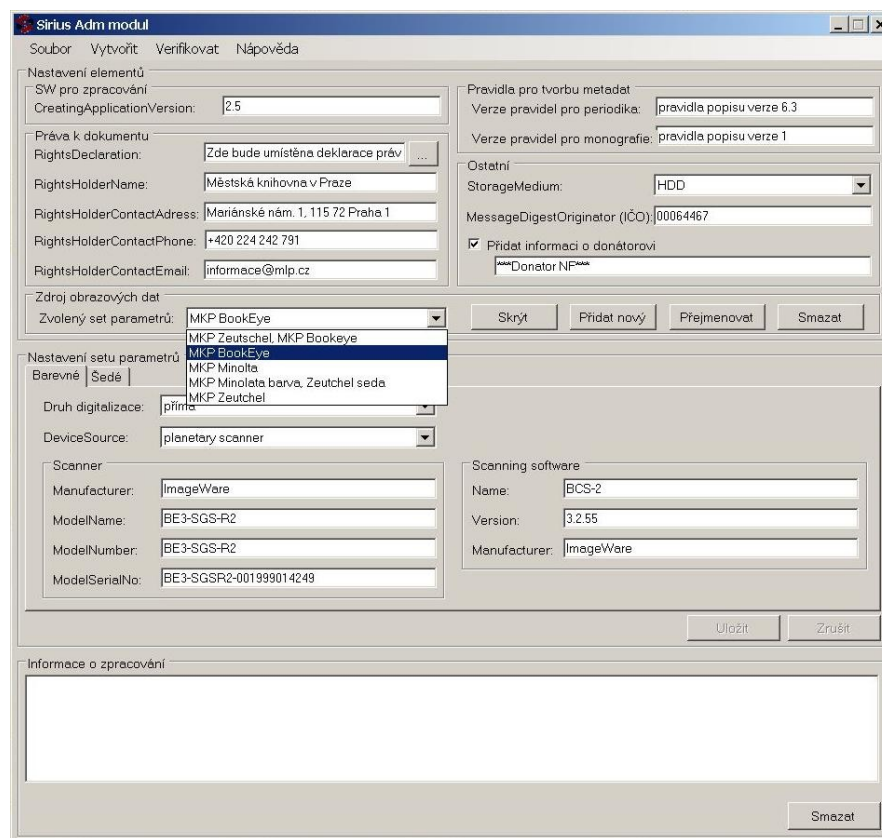
5.2.3.3 *Administrativní metadata*

Administrativní metadata jsou vytvářena v samostatném modulu softwaru Sirius nazvaném ADM a přidávají se do již vyexportovaného XML. Přidání administrativních metadat je tak poslední fází tvorby metadat a ukončení jedné z fází vlastního digitalizačního procesu. Podrobnější představu o parametrech, které je možné v programu ADM spravovat, lze nalézt na obrázku č. 13.

Právní metadata informují o právech k dokumentu. V programu jsou uvedena v následujícím pořadí: deklarace práv, informace o vlastníkově práv – *název instituce, kontaktní adresa, telefon a e-mailová adresa, IČO*.

Dále jsou do XML přidávány informace o verzích pravidel popisu a archivační metadata obsahující informaci, jak jsou data ukládána, v případě MKP se data nachází na HDD.

⁵⁴ V našem případě může kapitola začínat na straně 9 a končit na straně 13. U strany s číslem 9 je index stránky roven číslu 10 a straně s číslem 13 odpovídá index stránky 14. Výsledný rozsah indexů stran je tedy 10–14. Někdy se čísla stran a indexy stran velmi rozcházejí, obzvláště u titulů, které obsahují více vakátů nebo samostatně číslovaných obrazů – např. fotografií nebo výkresů. V těchto případech se index strany může od paginace odlišovat až o jeden či více číselných řádů.



Obrázek 13: náhled základního okna modulu Sirius ADM [Sirius]

Technická metadata popisující informace o skeneru a typu digitalizace jsou již implementována do programu Sirius ADM, takže stačí pouze vybrat, na kterém skeneru byly digitální kopie pořízeny. Nabídka výběru obsahuje i volbu kombinace skenerů – Zeutschel a Bookeye, kdy je možné skenovat část dokumentu na jednom a další část na jiném skeneru. U skenerů jsou uváděny tyto parametry: výrobce skeneru, název a číslo modelu skeneru, výrobní číslo skeneru. Parametry jsou doplněny informacemi o softwaru zahrnujícími jméno a verzi programu a výrobce. Dále je zvoleno, jak byly skeny pořízeny, zda ve stupních šedi či barevně. Automaticky je nastaven způsob digitalizace přímou metodou a zařízení označeno jako planetární skener.

Po výběru využitého skeneru a tím pádem i nastavení parametrů je třeba pouze zvolit umístění a název souboru XML, do kterého budou příslušná technická metadata přidána.

Modul Sirius ADM také nabízí další funkci, po doplnění technických metadat je možné spustit validační funkci, která kontroluje, zda jsou data v XML validní a tím pádem odpovídají nárokům a požadavkům na digitalizaci.

5.2.4 Úprava naskenovaných obrázků

Nejdůležitější částí procesu digitalizace je správně naskenovaná předloha. Pokud je totiž naskenovaná předloha dobře nasnímána, není třeba ji nijak výrazně upravovat. U dobře naskenované předlohy je třeba pouze oříznout okraje a případně vyrovnat sklon písma.

Vzhledem k tomu, že hlavní cíl digitalizace v MKP byl zachování obsahu dokumentu a primárně byly předlohy snímány ve stupních šedi, nebylo ve většině případů třeba obraz předlohy dodatečně upravovat v jiných grafických programech. K základní úpravě obrazu postačil software Sirius. Pokud bylo třeba výrazněji upravovat některé barevné obrazy, byl použit software *Adobe Photoshop* (verze 5.1), který byl dokoupen jako software pro skenování u skeneru Minolta. Po vyřazení skeneru Minolta byl tento přístroj umístěn po nějakou dobu v OD, takže software mohl být používán dále. V polovině roku 2010 byl skener Minolta přesunut do Ústřední knihovny, aby mohl být využíván pro účely skenování obálek a obsahů u velkoformátových dokumentů, místo softwaru *Photoshop* byl dočasně používán grafický software *Zoner Photo Studio* (verze 11). V roce 2011 byl znovu pořízen program *Photoshop* od Adobe, který je nyní využíván v případě potřeby úprav barevných skenů.

„Editor je opatřen vlastní oknem a poskytuje všechny základní funkce, jimiž jsou vybaveny grafické programy pro zpracování bitmapových obrázků. Jedná se o funkce, které umožňují změnit velikost a natočení obrazového dokumentu, ale především o nástroje pro úpravu vzhledu dokumentu, tzn. jas, kontrast, sytosti barev atd. K dispozici jsou také histogram a uživatelské filtry. Úpravy se provádí vždy pouze s jedním dokumentem, zvoleným v hlavním okně programu Sirius“ [ELSYST, 2003d, s. H-3].

5.2.4.1 Načtení obrázků do Siria

Poté, co je předloha nasnímána na některém ze skenerů a výsledné obrazy jsou uloženy v příslušné složce pojmenované podle signatury, dochází ke zpracování obrázků tak, aby mohli být uživatelům zveřejněny v digitální knihovně. Již při naskenování jsou obrazy vizuálně kontrolovány, takže dotyčný zaměstnanec již na počátku skenování tuší, zda je nutné skeny výrazně upravovat. Pokud není úprava nijak zásadní nebo se jedná o úpravu, kterou umožňuje během zpracování software Sirius, jsou obrazy načteny do Siria k příslušnému pořadači reprezentujícímu daný dokument a obsahujícímu již importované a zkontrolované bibliografické údaje. Software Sirius obsahuje Editor grafických dokumentů, který „slouží pro úpravu dokumentu, barevných i v odstínech šedé, uložených v některém z podporovaných

bitmapových formátů“ [Tamtéž, s. H-2]. Editor je tvořen vlastním oknem a umožňuje úpravy pouze jednoho skenu. Pro základní úpravy je tedy dostačující, ovšem zásadnější úpravy je pohodlnější vykonávat v jiném grafickém programu – tedy například v programu Photoshop nebo Zoner, který přece jen nabízí uživatelsky komfortnější prostředí a více nástrojů, díky kterým je samotná úprava snazší a rychlejší.

Načítání skenů je proces velmi jednoduchý a je uskutečněn v modulu nazvaném Skenovací linka. Program Sirius umožňuje načíst do libovolné dávky libovolné množství dokumentů. Zde záleží na každém pracovníkovi, jaký způsob práce mu vyhovuje. Množství souborů načtených do dávky je také ovlivněno celkovou velikostí souboru a tím pádem náročností na výkon. Pokud celková velikost dokumentu nepřesahuje formát A3 a dokument se sestává z menšího počtu stran, jsou do jedné dávky obvykle načteny všechny soubory najednou. Tento způsob usnadní práci se soubory, dokument je popisován najednou a lépe se označují jednotlivé reference neboli kapitoly. U dokumentů, jejichž velikost se blíží velikosti A3 nebo je větší, je vždy lepší do dávky načítat menší množství dokumentů. Zde potom závisí na odhadu a zkušenostech daného zaměstnance. Tyto skeny kladou při zpracování velký nárok na výkon. V těchto případech Sirius není příliš stabilní, často se při zpracování zasekne na úpravě jednoho obrazu, a to především při ořezu skenu. Pokud se tato situace vyskytne a například při zpracování skenů v dávce nastane nějaký problém, dochází často k poškození celé dávky, dávku nelze otevřít nebo se při jejím otevření celý software Sirius zasekne a je nutné vynutit jeho restartování. Na tento jev Sirius upozorní oznámením v novém okně. Mnohdy tento problém způsobují jeden nebo dva obrázky v dávce, kvůli kterým je pak celá dávka zničená a všechny dosud vykonané práce na této dávce byly zbytečné. Proto je lepší pracovat u velkoformátových dokumentů pouze s dávkami obsahujícími malé množství skenů.

Po nahrání dokumentů do dávky je nutné zkontrolovat nastavení. Především se nastavuje typ detekce okrajů, a to na hodnoty *Obě* či *Automaticky*. Pokud je nastavena hodnota *Obě*, Sirius detekuje obě stránky. Při nastavení hodnoty *Automaticky* se vybírá jedna určená strana z několika možností. Pokud je účelem, aby software Sirius detekoval u dvoustránkové předlohy obě strany, je vybrána možnost *Obě*. Pokud je například dokument tvořen volnými listy, vybere se možnost *Automaticky*. Volba automaticky je vhodná také u skenů, u kterých není žádoucí rozřezání stránek. Příkladem mohou být tabulky, fotografie nebo plány vyplňující celou dvoustranu. V tom případě je vhodné tyto skeny načíst do jedné samostatné dávky a nezapomenout zkontrolovat správnost nastavení typu detekce okrajů. Protože Sirius

si v paměti uchovává poslední nastavení programu, je třeba po načtení dávky tuto možnost zkontrolovat v nastavení. Pokud se tak nestane, je nutné po správném nastavení dávku načíst znovu. Poté již stačí dávku zavřít a tím spustit proces *Automatické detekce okrajů*.

Průběh procesu zpracování je možné sledovat v malém okně, během zpracování skenů není možné s programem nijak pracovat. Po skončení procesu se zobrazí informace o zpracování dávky – název dávky, datum, úspěch/neúspěch při ukládání dávky a doba zpracování. Poté je možné okno se *Skenovací linkou* uzavřít.

5.2.4.2 Ořez

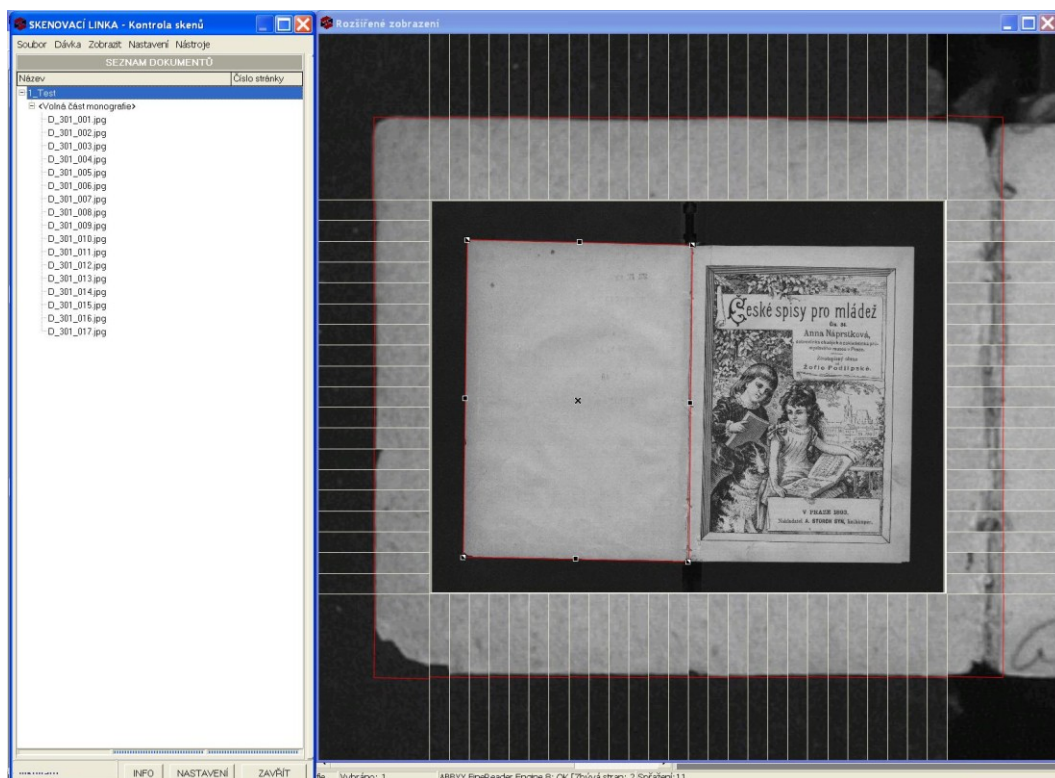
Po načtení skenů do dávky a po spuštění automatické detekce okrajů je třeba okraje zkontrolovat a manuálně vyrovnat nakřivo naskenované obrazy. K regulaci automatické detekce okrajů slouží modul nazvaný *Kontrola dokumentů*, který je zastoupen vlastním oknem a umožňuje upravit okraje jedné vybrané dávky.

5.2.4.2.1 Ruční kontrola okrajů

Pro usnadnění kontroly automaticky detekovaných okrajů je lepší využít doplňkové okno s Rozšířeným zobrazením. „Jedná se o takovou vizualizaci dokumentu, která využívá dostupnou plochu na monitoru s maximální efektivitou [ELSYST, 2007a, s. 1].“ Toto okno je nezávislé na okně *Kontroly dokumentů*, kde je obraz předlohy velikostí přizpůsoben oknu pro kontrolu. Lze s ním zacházet jako se samostatným oknem, tedy například zvětšovat či zmenšovat bez ohledu na velikost hlavního okna. Výhodu samostatného okna lze velmi dobře uplatnit při použití dvou monitorů – na jednom monitoru ponechat pro kontrolu seznamu skenů v dávce, zatímco na druhém monitoru pracovat s oknem s *Rozšířeným zobrazením* zvětšeným na celou plochu obrazovky.

Rozšířené zobrazení se skládá ze dvou částí – uprostřed je znázorněn celý sken spolu s ořezovým rámečkem červené barvy označujícím okraje nalezené automatickou detekcí, ve vnější části je přiblíženo bezprostřední okolí kolem tohoto rámečku (viz Obrázek č. 6). Při kontrole automaticky detekovaných okrajů je důležité mít na paměti zásadu, že z dokumentu musí být patrné, že se žádná jeho část při našem zpracování někde neztratila. Kontroluje se tedy, zda je okolo hranice konce dokumentu ponechán přibližně jeden či dva pixely černého pozadí od okrajů stránky a u svázaných titulů i to, zda je vidět alespoň kousek protější stránky u hřbetu. Velikost rámečku je možné upravit tažením jeho rohů nebo tažením hran. Tažením rohů vznikne libovolný útvar (lichoběžník), tažením hran jsou zachovány

pravé úhly rámečku. Posun celého rámečku je možný ve vnějším i vnitřním okně. Ve vnitřním okně slouží k posouvání černý křížek uprostřed skenu, ve vnější části stačí umístit kurzor myši dovnitř ořezového rámečku. Umístěním kurzoru myši vně rámečku umožní rotovat s celým rámečkem podle jeho středu. U starších titulů je předloha natolik nepravidelná, že software Sirius detekuje okraje velmi přibližně či dokonce špatně. Pokud tato situace nastane, je rychlejší vytvořit si vlastní ořezový rámeček než použít rámeček navržený softwarem Sirius.



Obrázek 14: Rozšířené zobrazení v SW Sirius [Sirius]

Rámeček lze vytvořit kliknutím na libovolnou část skenu a roztaháním na požadovanou velikost rámečku. V *Rozšířeném zobrazení* je možné zpracovávat pouze dávku o minimálním počtu dvou skenů. Při načtení skenů se totiž první obrázek nemusí načíst v *Rozšířeném zobrazení* a je nutné kliknout v seznamu skenů na další v pořadí a potom zpět. Teprve po tomto manévru se první sken načte a je možné pokračovat dál. Ovšem pokud by v dávce byl pouze jeden obrázek a při otevření *Rozšířeného zobrazení* by se neotevřel, nebylo by možné okraje zkontrolovat.

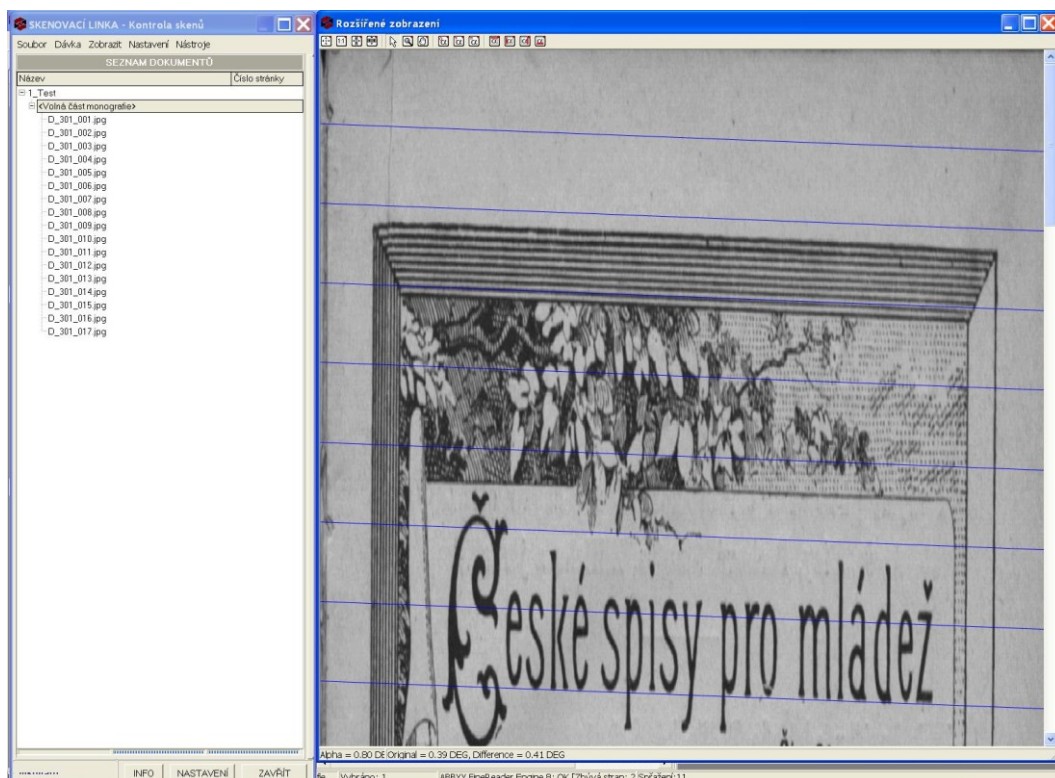
5.2.4.2.2 *Manuální kontrola vyrovnání*

V rámci kontroly okrajů po automatické analýze okrajů jsou u jednoho skenu ve dvou krocích revidovány dvě části procesu digitalizace: nejen ruční kontrola řezu ale i manuální kontrola vyrovnání řádků textu. Okno rozšířené zobrazení pro manuální vyrovnání umožňuje „optimální zobrazení celé šířky stránky při zachování zřetelných mezer mezi jednotlivými řádky spolu s grafickým znázorněním požadovaného úhlu natočení“ [ELSYST, 2007a, s. 5].

Ruční kontrola řádků může či nemusí být nastavena pro všechny skeny v dávce. Pokud je nastavena pro všechny skeny, po kontrole a opravě ořezového rámečku stačí stisknout mezerník a ve druhém kroku se upraví natočení skenu. Pokud není nastaven, kontrola natočení se vyvolá stisknutím klávesy *F5*. Pro manuální vyrovnání slouží vodící čáry zobrazující sklon textu dokumentu.

Pokud nejsou vodící čáry rovnoběžné se sklonem písma, je možné upravit pomocí myši sklon vodících čar (viz Obrázek č. 15). Klepnutím pravým tlačítkem myši a příslušným pohybem posouváme čáry nahoru a dolů, aby lícovaly s konkrétním místem v obraze. Nejlepší je vybrat si v obraze předlohy jeden řádek a podle něj určovat sklon textu. Někdy je sklon textu nepravidelný a řádky nejsou rovnoběžné, v tom případě je lepší vyzkoušet sklon na více místech a přizpůsobit vodící čáry, jak nejlépe to půjde. Sklon vodících čar se mění posunem kolečka myši nebo tažením dané přímky pomocí levého tlačítka myši. Použití levého tlačítka myši je ovšem rychlejší a méně přesné. Pokud je sklon výrazně jiný, je dobré použít pro přibližný posun vodící čáry levé tlačítko myši a na jemnější a přesnější manipulaci kolečko myši.

Funkce mezerníku pro posun mezi skeny funguje pouze při práci v rozšířeném zobrazení. Pokud se například mezitím přepneme do okna Kontroly dokumentů, není možné přepínat mezi skeny mezerníkem. Pro pohyb mezi jednotlivými skeny slouží kromě mezerníku také klávesy *F3* (o jeden vpřed) a *F4* (o jeden zpět).



Obrázek 15: Manuální kontrola vyrovnání v SW Sirius [Sirius]

Kontrolu okrajů i sklonu textu je možné kdykoli přerušit a zavřít dávku pomocí ikony Zavřít. Dosavadní kontrola je uložena, stačí pouze zaznamenat, kde kontrola skončila a po znovuotevření pokračovat v práci.

Poté, co jsou všechny skeny v dávce kontrolovány, je možné dávku uzavřít stiskem klávesy *F10*. Zobrazí se nové okno *Zpracovat dávku*, kde je kontrolováno nastavení procesu ořezu zahrnující možnost zpracování skenů a ořezání skenů s úpravou obrazu. Po potvrzení nabídky se zobrazí okno s průběhem procesu zpracování stejně jako při uzavření dávky v procesu analýzy okrajů.

Po ukončení procesu ořezu je třeba zkontrolovat, zda se skeny korektně ořezaly. U dvoustránkové předlohy jsou skeny po ořezání označeny příznakem *a* a *b*. Tedy například *D_9939_001a.jpg*, *D_9939_001b.jpg*.

5.2.4.3 Další nástroje pro úpravu skenů

K úpravě jednotlivých skenů i k hromadné úpravě je možné využít také nástroj pro manuální ořez i vyrovnání. Manuální vyrovnání je možné uplatnit pouze u černobílých dokumentů, nelze tedy použít pro barevné skeny. Nástroj manuální vyrovnání je možné spustit jak kdykoli z hlavního okna označením skenů v tabulce záznamů tak i během

indexování. Po označení jednoho či více dokumentů je po klepnutí pravým tlačítkem a výběrem položky *Manuální vyrovnání* otevřeno nové dialogové okno obsahující seznam dokumentů v levé části okna a náhled dokumentu v pravé části okna. Vyrovnání je realizováno úpravou v poli *Úhel*, kde se zadává hodnota, o kterou se má dokument otočit. Natočení dokumentu je možné kontrolovat v náhledu. Po dokončení úprav stačí klepnout na ikonu *Opravit*.

U skenů velkoformátových dokumentů se může stát, že se ořez ve skenovací lince neprovede a je nutné skeny oříznout ručně. Manuální ořez je stejně jako manuální vyrovnání možné spustit z tabulky záznamů v hlavním okně nebo během indexování. Po označení ikony *Výběr* je možné začít vytvářet ořezový rámeček. Umístěním kurzoru myši na vybraný roh, stisknutím levého tlačítka a tažením se označí vybraná oblast. Pokud je označeno více skenů, doplní se rámeček i na náhledy ostatních skenů. Pro úpravu velikosti a požadovaného tvaru rámečku slouží ikona *Upravit výběr*. Poté, co jsou skeny označeny, stačí klepnout na tlačítko *Zpracovat* a potvrzením volby je celá operace spuštěna.

5.2.5 Optické rozpoznávání znaků

Pro převod bitmapového obrazu písma do zpracovatelného textu pomocí metody OCR je používán modul programu Sirius nazvaný *Celostránkové OCR*. Ten využívá dva OCR klíče se softwarem ABBY FineReader s celkovou kapacitou 50 000 stránek za měsíc.

Výhoda modulu integrovaného OCR spočívá v jednoduchosti nastavení parametrů a ve společném exportu, který zaručuje automatické přiřazení vyčtených textů k odpovídajícím skenům.

OCR je možné spustit kdykoli po ukončení indexování, v praxi je vhodné provést OCR ihned po vlastní kontrole dokumentů. Všechny dokumenty určené na OCR se přesouvají pod příslušný pořadač označený podle jazyka textu. Pokud je psán dokument více jazyky, je lepší provést OCR ihned, pokud je to možné a před spuštěním procesu vybrat v nastavení příslušné jazyky. Optické rozpoznání znaků probíhá na určených počítačích, kde je zasunut OCR klíč, obvykle se spouštělo při odchodu posledního pracovníka a celý proces se nechával běžet přes noc. To bylo nutné především z kapacitních důvodů, při vyčítání textu není možné se softwarem nijak pracovat, takže pokud běží proces OCR, není možné program Sirius využít na nic jiného. Pokud byly všechny počítače přes den obsazeny, spouštělo se OCR přes noc.

Pokud je ovšem během dne počítač obsahující software Sirius volný, je možné na něm spustit OCR kdykoli.

OCR je možné spustit nad jakýmkoli pořadačem pomocí funkce *Rozpoznat všechny dokumenty v pořadači (OCR)*. Pokud je z nějakého důvodu nutné proces vyčítání přerušit, je možné jej znovu spustit funkcí *Rozpoznat všechny nerozpoznané dokumenty v pořadači (OCR)*, která rozpozná dosud nevyčtené skeny a proces vyčítání spustí pouze nad nimi.

Výsledkem procesu OCR je dokument ve formátu prostého textu, který slouží pro účely fulltextové vyhledávání v digitální knihovně. Každý sken má svůj vyčtený text v jednom souboru.

Metoda automatického rozpoznání textu funguje na principu porovnání znaků a složených slov s vlastním slovníkem. „Program analyzuje strukturu obrazu dokumentu a označí textové bloky, tabulky, obrázky a jiné části stránky. Po rozlišení jednotlivých znaků je software porovná se sadou vzorových obrazů a určí varianty, čím by tento znak mohl být. Na základě toho analyzuje různé varianty slov“ [ABBYY]. Čím poškozenější a nekvalitnější je předloha, tím menší je pravděpodobnost správného rozpoznání znaků. U dokumentů poškozených degradací papíru, kde se projevují znaky stárnutí papíru a skladování dokumentů v nevhodných podmínkách, jsou samozřejmě chyby při vyčtení textu pomocí metody OCR ve větším množství. Bohužel z kapacitních důvodů není možné kontrolovat každou stranu textu, proto jsou jednotlivé textové dokumenty ponechány v původním stavu a nedochází k jejich úpravám. V dalších úvahách o rozvoji služeb po skončení projektu HISPPRA je jedním z bodů také oprava vyčtených textů vybraných dokumentů.

5.2.6 Kontrola dat

Kontrola dat je nedílnou součástí procesu digitalizace a je velmi důležité nepodcenit tuto fázi a včas nastavit dostatečný mechanismus kontrol, který zamezí výskytu chyb ve zpracovaných digitálních dokumentech. Samotná kontrola probíhá v několika krocích.

První fáze kontroly dat je automatická, v rámci procesu indexování dat je softwarem hlídána úplnost vyplněných povinných polí. Pokud je některé povinné pole nevyplněno, nelze dávku vypustit do systému a dále s ní pracovat.

Druhá fáze je ruční a provádí ji pracovník, který dokument v softwaru Sirius zpracoval. Pro účely rychlé a efektivní kontroly se nejlépe hodí funkce Přeuspořádání dokumentů. V novém okně je vedle sebe zobrazen seznam zpracovaných skenů a jejich vyobrazení.

Výhoda spočívá v rychlejší načtení náhledu jednotlivých obrazů, v hlavním okně trvá po kliknutí na daný sken jeho načtení minimálně dvakrát déle. Také lze velmi jednoduše přesouvat více skenů najednou a hromadně upravovat pole indexů. v tomto režimu pracovník zkontroluje jím doplněné údaje, projde návaznost skenů a jejich číslování.

Třetí fáze kontroly je také ruční, provádí ji ovšem jiný pracovník než ten, který dokument zpracoval. Zpočátku realizace projektu HISPRÁ byla kontrola prováděna až po nahrání dokumentu do digitální knihovny Kramerius, ještě před zveřejněním dokumentu uživatelům. Tuto kontrolu prováděli tři vybraní pracovníci, ovšem časově ani objemově nebylo možné provádění kontroly tímto způsobem trvale udržet. Nebyl také takový výskyt chyb nebo nepřesností, protože se pracovníci při pilotním provozu pracoviště zaškolili, případné nejasnosti při zpracování byly odhaleny dříve. Ovšem s výměnou pracovníků bylo nutné reorganizovat tento zaběhnutý systém.

Nový systém kontrol je založen na kontrolorských dvojicích. Před exportem dat ze Siria jsou dané dokumenty přesunuty do zvláštního pořadače na kontrolu, který je označen jménem kontrolovaného, druhý z dvojice dokumenty zkontroluje, teprve po kontrole dané dokumenty exportuje a doplní technická metadata. Byla vybrána metoda kontroly správnosti doplněných údajů i celkového zpracování dokumentu ještě před exportem ze Siria. Důležité je, že dokumenty vidí někdo jiný, jinýma očima, má jinou perspektivu, takže vidí i chyby nebo může odhalit komplikovaná místa v dokumentu, jejichž zpracování je lepší prodiskutovat. Také se rychle odhalí případné neporozumění některým pravidlům zpracování. U nových pracovníků je při zaškolování a později určitou překlenovací dobu přidělen na kontrolu jeden prověřený pracovník, který perfektně ovládá pravidla zpracování. U nových pracovníků je velmi důležité včasné odhalení případné špatné interpretace pravidel. Jinak hrozí, že si tato nesprávně vyložená pravidla osvojí a jejich napravení je proces dlouhodobý.

V první části kontroly se pracovník zaměří na správnost zpracování skenů a jejich popis. Tato část je většinou realizována v režimu Přeuspořádání dokumentů. Při kontrole se prochází skeny, reviduje se paginace, oříznutí skenu a vyrovnání textu a vhodně vložené reference. Kontrola správnosti bibliografických údajů je prováděna v hlavním okně Siria, údaje jsou revidovány de visu i v AKS Koniáš.

Mezi nejčastější chyby patří nevyplněné indexy a rozsah indexů stran, uvozovky v názvu dokumentu, překlapy v názvových údajích, nedoplněné údaje o rozsahu stran a chybně uvedené autorské role po importu bibliografických údajů, tečky za názvy kapitol,

identifikační čísla monografií chybně vyplněná do názvu a ne do pole Identifikace součásti monografie, nedoplnění rozsahu stran u kapitol. Mezi zásadní chyby při ořezu dokumentu patří neotočení skenů. V digitální knihovně bohužel nelze skly otáčet, takže pokud jsou v dokumentu umístěny obrázky s popisky a pracovník je neotočí korektně, snižuje uživatelský komfort při prohlížení dokumentu. Další zásadní chybou je rozřezání stránek i v případech, kdy je to vyloženě nevhodné – zejména u tabulek či obrazových informací umístěných na dvoustraně.

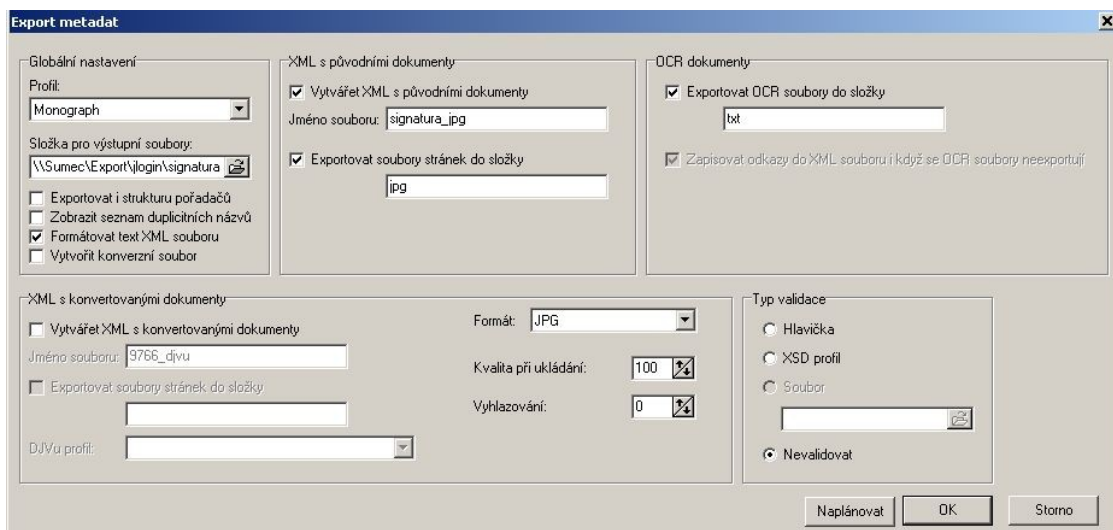
Třetí fáze kontroly probíhá při nahrání dokumentu do digitální knihovny, kdy je náhodně zkontrolována kvalita skenů i popisu.

5.2.7 Export dat

Export dat je prováděn po důsledné kontrole kvality skenů, korektnosti ořezu, správnosti popisu obrazů a referencí a importu bibliografických údajů z AKS Koniáš. Export do XML je možné provést, pokud je dokument indexovaný, má vyčtené OCR a naplněné pole index stránky a dokument je zařazen v pořadači *Na Export*. Z jiného pořadače není možné export do XML provést.

Export je prováděn automaticky, stačí pouze nastavit dané parametry (viz Obrázek č. 16). Je určen adresář pro ukládání výstupních souborů (každý pracovník má určený adresář nazvaný jeho jménem), dále se nastavuje jméno generovaného XML souboru, zda jsou exportovány skeny a název složky, kam jsou exportovány, a zda se exportují soubory s vyčteným textem a do jaké složky. Software Sirius umožňuje také export konvertovaných souborů, tato možnost ale není v Oddělení digitalizace MKP využita.

Výsledkem exportu dat je adresář s názvem signatury obsahující soubor XML pojmenovaný *signatura_jpg.xml*, složku se skeny nazvanou JPEG a v případě realizace procesu optického rozpoznání textu také složku s textovými soubory označenou TXT. Příklad vyexportovaného digitalizovaného dokumentu ve formátu XML je uveden v Příloze č. 3



Obrázek 16: Nastavení exportu dat v SW Sirius [Sirius]

Po úspěšném vyexportování dat se příslušný pořadač přesune do pořadače *Na archivaci*, kde jsou zpracované dokumenty po nějaký čas archivovány. Poté je pořadač se zpracovanými soubory vymazán. V případě nutnosti zásahů do zpracovaného dokumentu jej lze kdykoli znovu importovat pomocí původně exportovaných dat zpět do Siria.

Po doplnění technických metadat pomocí modulu *Sirius ADM* (viz kapitola 5.2.3.3 Administrativní metadata) je možné zveřejnit digitalizovaný dokument v digitální knihovně Kramerius.

5.2.8 Zveřejnění pro uživatele

Po provedení exportu dat z programu Sirius, následného doplnění administrativních metadat a validování dat v modulu *Sirius ADM* je možné digitální dokument nahrát do digitální knihovny Kramerius. K tomuto účelu slouží administrátorské rozhraní systému Kramerius (viz Obrázek č. 17).

Nahrání dat do systému Kramerius je možné pouze z určitého adresáře, v našem případě je to adresář s názvem *Dokrameria*. Tituly vybrané ke zveřejnění jsou nahrány do adresáře, přes noc proběhne spojení adresáře s databází systému Kramerius přes protokol OAI-PMH. Teprve poté je možné digitální dokumenty nahrát.

Je vybrán import dat a po dobu určenou k nahrávání tak, aby bylo možné dokumenty také zálohovat, v našem případě je to od 10 do 12 hodin, za tu dobu je proveden pouze určitý počet import, je proto důležité obsáhlejší monografie skládající se z několika volných částí importovat zvlášť.

The screenshot shows the administrative interface of the Kramerius system. At the top, it identifies the system as 'SYSTEM • KRAMERIUS •' and the user as 'Digitální knihovna Městské knihovny v Praze'. The main content area displays the metadata for a document titled 'Památník Divadelní Jednoty v Opavě'. The metadata includes:

- Internal identifier:** 4973
- Access rights:** neviditelný (invisible)
- Import date:** 20.12.2011 13:26:27
- File name:** D:5932_pg.xml
- Main title:** Památník Divadelní Jednoty v Opavě
- Document type:** monografie
- Publisher:** Divadelní jednota
- Publication year:** 1927
- Place of publication:** Opava
- Physical description:** 60 s.
- Language:** cze
- Keywords:** 1927, české divadlo (o něm), divadlo (o něm), forma: text, nosič: papír, ochotnické divadlo, ochotnické spolky, Opava (město), spolkové divadlo, úroveň zpracování: jednoduchá, způsob vydávání: monografie, divadelní almanachy, divadelní soupisy, památníky (knihy)
- Notes:** Památník Divadelní Jednoty v Opavě vydaný při příležitosti 40. výročí jeho založení a spolkové činnosti v roce 1927.
- Location:** Městská knihovna v Praze
- Signature:** D 5932

At the bottom, there is a section for 'Vnitřní součást monografie:' with a list of chapters and their page ranges, such as 'kapitola - Naše jubileum - s. 3 - 4; kapitola - Čtyřicet let Divadelní Jednoty v Opavě - s. 5 - [32b]; kapitola - Členové sdružení Divadelní Jednoty - s. 33 - 34; kapitola - Soupis her provedených Divadelní Jednotou v Opavě - s. 35 - 49; kapitola - Tabulární přehled divadelní činnosti a peněžního obrátu (vzdání) Divadelní Jednoty - s. [50] - [51]; kapitola - K odhalení Smetanova pomníku v městských sadech opavských - s. 52 - [52a]; kapitola - Vzpomínky členů Divadelní Jednoty - s. 53 - 60;'. A sidebar on the left contains administrative menu items like 'Typy dokumentů', 'Administrativní menu provoz', 'Nastavit přístupnost dokumentů', etc.

Obrázek 17: Administrátorské rozhraní systému Kramerius [Systém, c2003-2009]

Po importu jsou dokumenty označeny jako neviditelné, jsou tedy nahledatelné pouze z administrátorského rozhraní, v uživatelském rozhraní není možné je vyhledat a nejsou zobrazeny ani v seznamu dokumentů. Po jejich zviditelnění, tedy zviditelnění v rámci uživatelského rozhraní je nutné nastavit u digitálního dokumentu příslušnou platnost či neplatnost majetkových autorských práv podle zákona č. 121/2000 Sb. (autorský zákon). Existují dva zobrazovací režimy: titul může být veřejný nebo neveřejný. Veřejný titul je titul, k němuž již majetková autorská práva nemají platnost. U tohoto dokumentu je možné zobrazit jak bibliografický popis, tak jeho digitální obraz. V případě druhého režimu, který se nazývá neveřejný, stále trvají majetková autorská práva, případně právo nakladatele na daný titul. To znamená, že titul není možné plně zveřejnit mimo síť Městské knihovny v Praze. Na počítačích připojených do sítě MKP si tak uživatelé mohou prohlédnout plný digitální dokument, mimo síť MKP je dostupný pouze bibliografický popis. Posuzování existence majetkových autorských práv je individuální, provádí se u každého titulu zvlášť. To je velký rozdíl oproti jiným velkým knihovnám, které se zabývají digitalizací historického fondu. Například Národní knihovna ČR má určenou časovou hranici, která je platná plošně

pro všechny dokumenty, tuto hranici tvoří pro periodika rok 1890, u monografií je tato hranice tvořena rokem 1880.

V rámci právního posudku, který byl pro účely projektu HISPRA vypracován vyplývají pro zveřejňování historických děl následující pravidla.

U monografií a dokumentů monografické povahy jsou v kombinaci s knihou v ruce a údaji z AKS Koniáše zjištěny všechny osoby zastávající autorské role. V jednoznačných případech je možné určit datum zveřejnění dokumentu podle data úmrtí posledního z autorů, datum je zapsáno do určeného pole v AKS Koniáš. V případech, kdy je autor znám trvají majetková práva autora 70 let po úmrtí autora a dílo může být zveřejněno 1. ledna následujícího roku. Pokud je tedy dílo již volné, je možné nastavit v AKS Koniáš trvání autorských práv na hodnotu NE, po úspěšném importu je tato skutečnost promítnuta i do digitální knihovny Kramerius, kde může být daný titul zveřejněn. Pokud ještě neuplynula doba, kdy je možné titul zveřejnit, запиše se příslušné datum do daného pole v AKS Koniáš a současně je v systému Kramerius titul označen jako neveřejný.

Pokud autor není znám, trvají majetková práva 70 let od oprávněného zveřejnění díla, poté je dílo označeno za volné a je možné s ním nakládat stejně jako s jiným volným dílem, viz výše.

Periodika, sborníky, časopisy a jiné pokračující zdroje se považují za dílo souborné, autorem je uspořadatel. Toto právo není odvozeno od práv jednotlivých autorů ke svým dílům, a není jím dotčeno. Jednak tedy běží lhůta 70 let k jednotlivým článkům, jednak lhůta 70 let k dílu soubornému. U periodických děl většího rozsahu není možné zjišťovat všechna majetková autorská práva, proto bylo vybráno podobné řešení, jaké má Národní knihovna – v našem případě bylo jako mezní vybráno také datum 1890, v letošním roce došlo k jeho posunu na rok 1895.

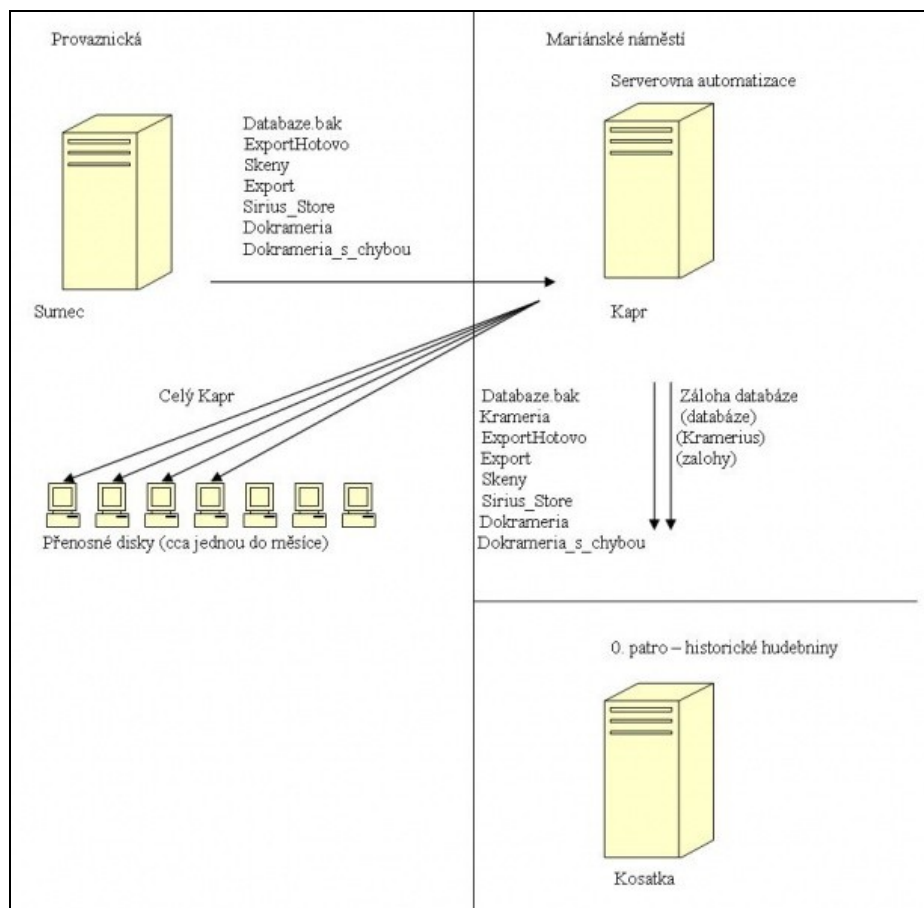
Po označení hodnoty v AKS Koniáš je přes noc k záznamům původních tištěných dokumentů doplněn příznak *d* v případě zveřejněných titulů a příznak *d*© v případě titulů neveřejných. Pomocí příznaku jsou tedy označeny digitalizované tituly. V online katalogu se potom tato skutečnost projeví odkazem *Prohlížejte online*.

5.2.9 Archivace dat

Archivace dat byla také jednou z podmínek realizace projektu HISPRA, proto bylo již od prvních úvah o projektu počítáno se zálohami na servery s dostatečnou kapacitou. V případě ztráty nebo poškození dat tak data mohou být obnovena.

Celý systém zálohování je sice promyšlený, ale nejedná se o standardní systém archivace dat pomocí repositáře. Proto existují jisté pochybnosti, zda je možné splnit podmínky dlouhodobé archivace dat. Již půl roku je tedy v MKP připravována revize archivace dat v souvislosti s nutností archivace výstupů z dalších činností Oddělení digitalizace po skončení projektu HISPRA. Tento nedostatek jsem si uvědomovala ihned po nástupu do funkce vedoucí Oddělení digitalizace, proto je přehodnocení archivace nebo spíše zálohování dat v Oddělení digitalizace jedním z dlouhodobých úkolů.

Data jsou uložena na třech serverech, s názvy *Sumec*, *Kapr* a *Kosatka*. Server *Sumec* je umístěn v Oddělení digitalizace v Provaznické ulici, na tomto serveru je spuštěný software Sirius a jsou na něj ukládána všechna data vyprodukovaná jakoukoli činností v Oddělení digitalizace (viz Obrázek č. 18). Přes noc jsou tato data zálohována na server *Kapr* umístěný v serverovně automatizace na Mariánském náměstí, odkud jsou data ještě zrcadlově zálohována na server *Kosatka* v přízemí Ústřední knihovny v depozitáři historických hudebnin. Do budoucna bychom ovšem chtěli tento třetí server přesunout na jinou lokalitu tak, aby všechny servery byly umístěny v jiných lokalitách a dělila je určitá vzdálenost. Tím bude zajištěna větší bezpečnost, aby v případě neočekávaných událostí, jakými byly například povodně v roce 2002, nedošlo ke ztrátě dat.



Obrázek 18: Schéma zálohování dat v Oddělení digitalizace MKP [MĚSTSKÁ, 2009b, s. 1]

Na obrázku č. 18 jsou znázorněny také adresáře vstupující do procesu archivace. Archivovány jsou původní skeny v adresáři *Skeny*, databáze softwaru Sirius obsahující upravené skeny i metadata (*Sirius_Store*), dále exporty z databáze softwaru Sirius (*Export*) a adresář *Dokrameria* a *Dokrameria_s_chybou*, které obsahují data nutná k importu digitálních dokumentů do systému Kramerius. Mezi další archivovaná data patří adresář *Všichni* zahrnující soubory obsahově významné pro všechny pracovníky Oddělení digitalizace (mimo jiné také seznamy digitalizovaných titulů a pravidla popisu). Dále je zálohován i osobní prostor na disku každého uživatele.

5.2.10 Kontrola fyzických svazků a určení následné péče

Již po povodních v roce 2002 a s nimi souvisejícím přesunu Oddělení vzácných tisků na Brusnici bylo vydáno rozhodnutí, že unikátní tituly ve specializovaných úsecích mají být postupně přesouvány do klimatizovaného depozitáře. Tyto tituly byly vybrány přednostně k digitalizaci, po skončení digitalizace nebyly vráceny zpět na regály ve specializovaných

úsecích, ale mohla jim být poskytnuta náležitá péče spočívající především v uložení ve vhodnějším prostředí klimatizovaného depozitáře.

Kontrola fyzických svazků probíhá již během příprav dokumentů k digitalizaci, kdy je po vyjmutí svazku z regálu zkontrolován jeho stav. Pokud je stav dokumentu špatný, je v seznamu titulů určených k digitalizaci dokument označen jako poškozený. Další kontrola fyzických svazků probíhá během procesu digitalizace, špatný fyzický stav je odhalen během procesu skenování – dokumenty se často nachází v tak špatném stavu, že jsou i po šetrné manipulaci při digitalizaci nenávratně poškozeny. Tato situace nastává především u dokumentů poškozených degradací kyselého papíru, u některých z nich se papír drotil při pouhém doteku. Naskenování takového dokumentu je možno označit přímo uměním.

V rámci zjednodušení postupu byly vyčleněny tři vrstvy míry poškození:

- dokumenty fyzicky poškozené, které by měly být v knihařské dílně opraveny, po opravě by bylo vhodné je ochránit obalem z lepenky archivní kvality, poté mohou být vráceny zpět na regály
- dokumenty poškozené degradací papíru, po stránce knihařské není možná jejich oprava, musí být opatřeny ochrannou krabicí, restaurátorky je nedoporučují dál půjčovat, ovšem konečný verdikt vysloví knihovnice z dotčeného úseku
- dokumenty, u kterých je vhodný restaurátorský zásah, zde záleží na posouzení daného úseku [MĚSTSKÁ 2008, s. 1].

Vybrané poškozené svazky jsou ponechány v OD, kde jsou posouzeny restaurátorkou a rozděleny do výše uvedených tří skupin. Svazky z prvních dvou skupin se přesouvají do knihařské dílny na Mariánském náměstí. Na dokumenty z druhé skupiny, které není možné opravit, jsou vyrobeny ochranné obaly. Poté jsou svazky vráceny zpět do úseků, kde je posouzeno, zda dokumenty z poslední dávky zůstanou na regálech či budou umístěny do Oddělení vzácných tisků do klimatizovaného depozitáře.

6 Propagace

Nedílnou součástí každého projektu je propagace výsledků. I v případě projektu HISPRA je tomu stejně. Především v případě grantů ze zahraničí je propagace výsledků dosažených s pomocí dotačních financí důležitou součástí. Šíření povědomí o zdigitalizovaných dokumentech a jejich využívání je dokonce jedním z cílů projektu. V následující kapitole je uvedeno, jakými prostředky bylo tohoto cíle dosahováno.

6.1 Online prezentace

Na webových stránkách knihovny se nachází informace o Oddělení digitalizace i projektu HISPRA. Na počátku roku 2011 byl oficiálně spuštěn nový web Městské knihovny v Praze, kde jsou informace o projektu HISPRA již minimalizovány, větší propagace byla na původních webových stránkách MKP. Nyní na stránkách Oddělení digitalizace uživatelé najdou základní informace o procesu digitalizace a odkazy na prezentace projektu v médiích. Na původních stránkách byly také tematické dávky usnadňující orientaci v množství digitalizovaných dokumentů. Tyto dávky se v jiné podobě postupně převádí na nový web. Odkaz na digitální knihovnu Kramerius i na alternativní webové stránky *Praha v knihovně* je na nových webových stránkách pod projekty *e-knihovny*, která sdružuje všechny dokumenty MKP dostupné v elektronické podobě. Během realizace projektu byla na webových stránkách také tematická upozornění.

Zpřístupnění dokumentů je realizováno nejen v digitální knihovně Kramerius a na webových stránkách www.prahavknihovne.cz, ale také prostřednictvím AKS Koniáš a online katalogu. V AKS Koniáš i v původním online katalogu byla možnost prohlížení digitálního obrazu dokumentu označena příznakem *d* v případě volných děl nebo *d©* v případě děl pod ochranou autorského zákona. Informace o digitální verzi je pouze připojena k původnímu bibliografickému záznamu. Tato možnost byla zvolena v počátcích projektu, dnes se jeví jako méně vhodná, není možné vyhledávat pouze ve zdigitalizovaných dokumentech nebo při vyhledávání vyfiltrovat digitální dokumenty. Výhoda tohoto systému byla patrná při vyhledávání na starém online katalogu, kde bylo možné z výpisu výsledků hledání okamžitě odlišit digitalizované dokumenty. Po kliknutí na příznak DIGI byl uživatel přesměrován na záznam dokumentu v digitální knihovně Kramerius.

Na novém webu byla dlouho zachována stejná funkcionalita, výjimkou byla absence odlišení digitalizovaných dokumentů přímo z výpisu výsledků hledání. Novým prvkem

v online katalogu je nyní možnost filtrovat digitalizované dokumenty přímo v rámci vyhledávání.

6.2 Plakát a letáky, DVD

V rámci propagace byl vytištěn letáček o projektu HISPRA, který zahrnoval základní informace o projektu: cíle a obsah projektu, hlavní aktivity projektu a informace o digitalizačním pracovišti. Tato propagační brožurka byla vytištěna ke slavnostnímu otevření Oddělení digitalizace. Další letáky malého formátu informovaly o webové stránce www.prahavknihovne.cz. K příležitosti uspořádání konference byly vytištěny také letáky a plakát, obojí v česko-anglické verzi, upozorňující na elektronické texty MKP.

Výukové CD nebo DVD cílené na mládež vyústilo z projektu *Pragensia pro 21. století*, na kterém se podílelo *Soukromé reálné gymnázium – Přírodní škola*. Hlavním cílem projektu bylo „podpořit inovativním a kreativním způsobem využívání digitální knihovny pragensií MKP pro potřeby dětí a mládeže“ [MĚŘÍNSKÁ, 2009]. Projekt byl realizován od konce roku 2008 do poloviny roku 2010, kdy byly jeho výsledky prezentovány na konferenci. Přírodní škola byla do projektu přizvána z několika důvodů. Studenti této školy jsou zvyklí na projektové vyučování, takže je tento typ práce nezaskočí. Jsou vedeni ke spolupráci, praktičnosti a samostatnosti, jsou také zvyklí formulovat vlastní názory. V rámci pilotní fáze byla vyhlášena témata, která měli studenti za úkol zpracovat⁵⁵. Při práci na tématech studenti absolvovali odbornou exkurzi ve *Studovně pragensií*, aby se seznámili s fondem pragensií, zkusili si vyhledat dokumenty v knihovním katalogu a z vyhledaných pragensií získat informace. Poté následovala vyučovací hodina s názvem *Digitalizace informací*, kde se studenti od svých pedagogů dozvěděli podrobnější informace o digitalizaci a porovnávali výhody a nevýhody digitalizace. S povědomím o digitalizaci pak studenti navštívili také Oddělení digitalizace, kde se seznámili v praxi s procesy digitalizace, především se snímáním obrazových informací na knižních skenerech. Součástí exkurze bylo také vyhledávání v digitální knihovně Kramerius. Poslední část pilotní fáze pak tvořila samostatná práce studentů, kteří využili poznatky z exkurzí a vypracovávali zadaná témata s pomocí informací získaných z digitální knihovny Kramerius. Studenti a učitelé po vypracování témat zhodnotili přínos projektu, důležitá byla zpětná vazba především

⁵⁵ V pilotní fázi byla vyhlášena tato témata: Srovnání obchodního využití domů na Smíchově v 30. letech 20. století a v současnosti, Sledování nemocnosti pražských dětí v minulosti a dnes, Srovnání povodní v letech 1890 a 2002, Pražská zoologická zahrada dříve a dnes, Pražská nádraží v proměnách času [MĚŘÍNSKÁ, 2009].

na vyhledávání v systému Kramerius. Na základě připomínek vznikla nová webová prezentace digitalizovaných dokumentů nazvaná Praha v knihovně (viz kapitola 6.3 Praha v knihovně). Součástí projektu byla také soutěž *Výlety pana Broučka po Praze*, která se skládala ze dvou částí. První část tvořila *Literární soutěž pana Broučka* určená pro žáky z věkové kategorie 11-18 (20) let, tématem práce byla *Praha za sto let*. Soutěžící byli vyhodnoceni ve dvou kategoriích, z každé kategorie byly vybrány tři nejlepší práce, které byly oceněny [MĚSTSKÁ, 2009a, s. 1]. Druhá část soutěže s názvem *Pan Brouček se ptá* byla realizována prostřednictvím komunitního webu K4U, cílem bylo odpovědět na otázky v kvizech tematicky zaměřených na Prahu [Tamtéž, s. 1]. Odpovědi na otázky se skrývaly v dokumentech dostupných na www.prahavknihovne.cz. Za správnou odpověď získávali soutěžící body, výhercem byl soutěžící s nejvyšším počtem bodů. Celkovým cílem celé soutěže bylo ukázat mladší věkové skupině uživatelů, co vše se dá najít na webových stránkách www.prahavknihovne.cz.

6.3 Praha v knihovně

Z realizace projektu *Pragensia pro 21. století* vyplynula jistá nespokojenost při vyhledávání v systému Kramerius, nejen co se týká uživatelského rozhraní systému, ale také například různého zobrazení obrázků ve formátu JPEG v různých prohlížečích. Mezi hlavní připomínky patřily tyto: vylepšení fulltextového vyhledávání a menu na vyhledávání, přizpůsobení na vyhledávání jen grafických informací [HAVLŮJOVÁ, 2010, s. 90].

Na základě požadavků nejen od učitelů a studentů zapojených do projektu *Pragensia pro 21. století*, ale také požadavků uživatelů získaných v MKP vznikla alternativní webová prezentace digitalizovaných dokumentů s názvem *Praha v knihovně*.

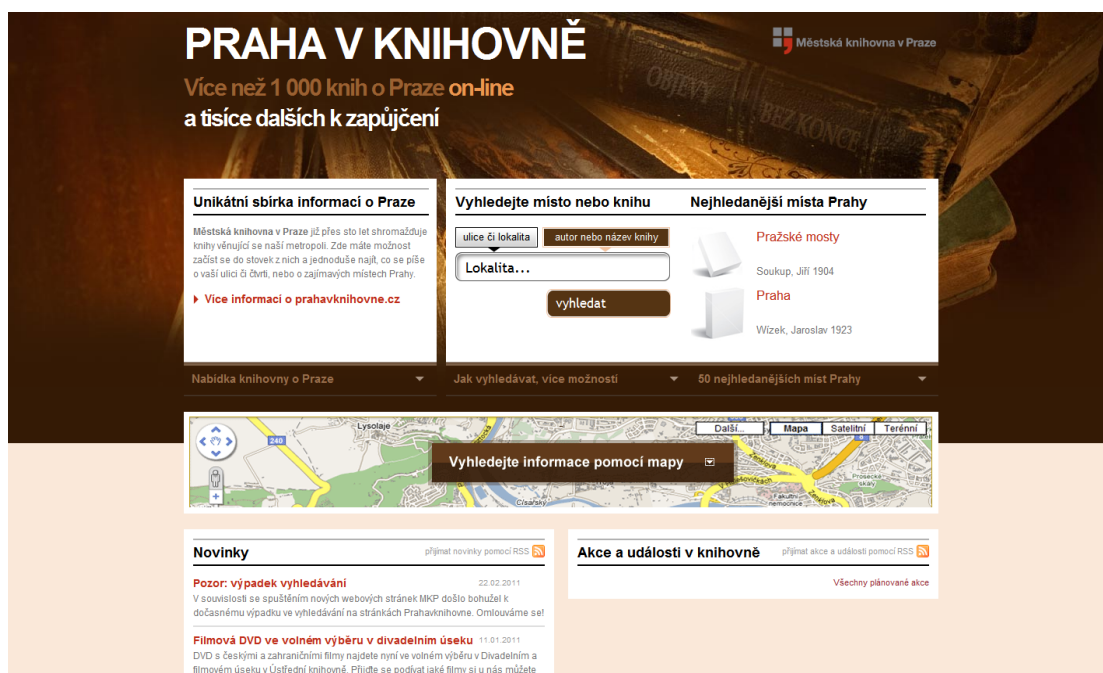
Webová stránka *Praha v knihovně* obsahuje informace o projektu *Praha v knihovně* i o digitalizovaných dokumentech. Výhodou je umístění tematických seznamů vybraných digitálních dokumentů.

Způsob hledání se přizpůsobil očekávání uživatelů, kteří jsou zvyklí na uživatelská rozhraní a prostředí zahrnující prvky webu 2.0. Dnešní mladší generace je zvyklá na vyhledávání typu *Google*, očekává tedy jedno textové pole, ve kterém formuluje své dotazy. Na základě podobných podnětů nabízí hlavní strana vyhledávání textové pole s možností výběru vyhledávání ulice či lokality nebo vyhledávání v názvu titulu či jménu

autora. Velkou výhodou je možnost vyhledávání v interaktivní mapě, a to pomocí místopisných deskriptorů, které byly obohaceny o GPS souřadnice. Z prvků webu 2.0 byl integrován například tag cloud 50 nejhledanějších míst Prahy. Vyhledávání je realizováno nejen v digitalizovaných dokumentech, ale pod výsledky vyhledávání jsou uživatelům nabídnuty i nedigitalizované tedy analogové dokumenty z fondu MKP.

Výsledky hledání jsou zobrazeny v seznamu obsahujícím *Název dokumentu*, *Autora* a *Rok vydání*. U dokumentů chráněných autorským zákonem, u kterých je možné zobrazit pouze bibliografický záznam je v seznamu výsledků hledání uvedena tato věta: „Z důvodu autorskoprávní ochrany může být elektronická verze knihy prohlížena pouze v pobočkách MKP přes elektronický katalog Koniáš“ [Praha, 2009].

Zobrazení konkrétního dokumentu je uspořádáno jinak než v systému Kramerius. Ústředním motivem je obrazová informace dostupná v zobrazovacím okně určené velikosti. U skenů je možné měnit velikost, k přiblížování a oddalování slouží posuvník nad oknem zobrazení. Vedle posuvníku se nachází možnost zobrazit sken ve v reálné velikosti, tedy 1:1. Po stranách jsou šipky umožňující listování v dokumentu. V nástrojové liště je také zobrazeno, na které straně v dokumentu se uživatel nachází, například na straně 5 z 500 stran. První údaj o místě, kde se uživatel nachází, je uveden v podobě zapisovatelného pole, které slouží pro skokový přesun ve stránkách dokumentu (viz Obrázek č. 19).



Obrázek 19: Uživatelské rozhraní webové stránky Praha v knihovně [Praha, 2009]

Po pravé straně jsou dostupné informace z bibliografického popisu, a to *Název dokumentu, Autor a Rok vydání*. Pod nimi se nachází popis knihy a odkaz do digitální knihovny Kramerius, kde jsou uvedeny podrobnější údaje z bibliografického popisu. Pod odkazem je seznam *Podobných knih*.

Webové rozhraní *Praha v knihovně* slouží tedy především pro prohlížení obrazových informací, popis dokumentu je minimalizován na anotaci, pro další informace o knize slouží odkaz na digitální dokument do systému Kramerius. Nevýhoda webové stránky spočívá v tom, že neumožňuje fulltextové vyhledávání.

Po ohlášení realizace systému Kramerius 4 a především po první reálné instalaci systému Kramerius 4 byl vývoj tohoto webového rozhraní pozastaven.

6.4 Konference Digitalizace aneb konec Oslich uší

Na závěr projektu HISPRA byla uspořádána konference s názvem *Digitalizace aneb konec Oslich uší*, která projekt symbolicky uzavřela. V rámci konference byl také vydán česko-anglický sborník s texty příspěvků.

Konference se zahraniční účastí se konala ve dnech 14. – 16. 6. 2010 v prostorách Ústřední knihovny Městské knihovny v Praze. Na pořádání konference se kromě MKP podílel také Archiv hlavního města Prahy a spolupráce byla navázána i se Sdružením knihovníků a informačních pracovníků.

Na programu konference byly kromě příspěvků také exkurze na digitalizační pracoviště MKP, Archivu hl. m. Prahy, NK ČR a KnAV ČR.

Konferenci zahájili ředitel MKP Tomáš Řehák, ředitel Archivu hl. města Prahy Václav Ledvinka, a velvyslanec Norského království Jens Eikaas. Na konferenci byly kromě projektu HISPRA prezentovány i výsledky dalších projektů podpořených z norských fondů, a to projekt *Ad fontes* (Tomáš Hanousek, Archiv hlavního města Prahy), projekt *Digimon – digitalizace a mikrofilmování moravských novin* (Lubomír Novotný a Radka Vítková, Vědecká knihovna Olomouc) a projekt *Záchrana neperiodických bohemikálních dokumentů 19. století ohrožených degradací papíru* (Jiří Polišenský a Ondřej Lehrl, Národní knihovna ČR). V další části konference zazněly příspěvky představující digitalizační projekty ze zahraničí: *KAMRA, portál místního kulturního dědictví* (Breda Karun, Slovinsko) a *Slovenská digitálna knižnice - detašované pracoviisko SNK, Vrútky* (Katarína Krištofová, Slovensko).

Z dalších příspěvků jmenuji například příspěvek J. Hubičky z Muzea Šechtl a Voseček v Táboře s názvem *Digitalizace historických fotografií* nebo příspěvek o *Langweilově modelu Prahy* od Kateřiny Bečkové z Muzea hlavního města Prahy [MĚSTSKÁ, [2010]].

Na závěr byla připravena panelová diskuze na téma *Kam směřuje digitalizace v knihovnách*, které se účastnili zástupci knihoven, soukromého sektoru i z řad uživatelů⁵⁶.

⁵⁶ Pro lepší přehlednost uvádím účastníky panelové diskuze: Luci Vývrová - PALMINKIHY; Patrick Zandl, LUPA; Tomáš Řehák, Městská knihovna v Praze; Petr Žabička, Moravská zemská knihovna; Pavla Švástová - Moravská zemská knihovna; Pavel Kocourek, INCAD; Matouš Bičák - student Přírodní školy [MĚSTSKÁ, 2010].

7 Digitální knihovna Kramerius

Open source software Kramerius byl vyvinut v rámci stejnojmenného programu VISK č. 7 (*Národní program mikrofilmování a digitálního zpřístupňování dokumentů ohrožených degradací kyselého papíru – Kramerius*), který probíhá od roku 2001.

Potřeba vzniku takové typu systému byla zesílena po rozsáhlých povodních v roce 2002, kdy bylo třeba zachránit a zpřístupnit poškozené tituly. Povodně tak znamenaly velký krok ve vývoji v oblasti digitalizace. Počátky vývoje systému jsou spojeny s Národní knihovnou ČR, především se jmény Adolf Knoll a Jiří Polišenský. Poté se přidaly také další knihovny – velmi brzy se do projektu zapojila Knihovna Akademie věd, později se přidala také Moravská zemská knihovna, Státní vědecká knihovna v Olomouci a další.

Do vývoje systému byla zapojena firma *Qbizm*, jakožto vítěz vyhlášeného výběrového řízení. „Jednou z podmínek výběrového řízení bylo, aby byl systém vyvíjen jako open source, což je počítačový software s otevřeným zdrojovým kódem“ [POLIŠENSKÝ, 2009]. Původně byl systém vyvíjen pouze pro periodické publikace, ovšem v roce 2007 přibyla možnost importovat do databáze systému také monografie.

Samotný systém Kramerius představuje specifický typ CMS⁵⁷ (*Content Management System*) jehož primárním účelem je zpřístupnění archivních a cenných dokumentů široké veřejnosti v souladu s autorským zákonem [Uživatelský, 2003].

„Celý systém sestává z více provázaných aplikací, které tvoří výrobní linku, v rámci které je zajištěna digitalizace dokumentů, vytváření primárních metadat pomocí OCR, následná úprava metadat a obrazových podkladů a finální zpřístupnění dokumentů pomocí WWW prohlížeče“ [QBIZM, c2005-20009]. Aplikace je zdarma dostupná pod licencí GNU GPL⁵⁸. V této části kapitole bude probírána pouze digitální knihovna Kramerius dostupná přes webové rozhraní prohlížeče, ostatní procesy již byly popsány v předchozích kapitolách.

7.1 Technická specifikace

Základní technologie, které aplikace Kramerius využívá, jsou: JAVA, OS Linux (běh systému), Apache (webový server), Tomcat (aplikační server), Postgres SQL (databázový

⁵⁷ CMS je systém pro správu obsahu (angl. Content Management System) je software zajišťující správu dokumentů [Systém, 2001-].

⁵⁸ GNU GPL (angl. General Public License) je licence pro svobodný software [GNU, 2001-].

server), Lucene (vyhledavač). Tedy technologie, které jsou rovněž zdarma a nepředstavují pro případného zájemce o instalaci žádné dodatečné náklady [LJUBKA, 2008, s. 92].

Do systému jsou digitální dokumenty implementovány pomocí formátů XML, v roce 2005 byl do systému začleněn také protokol OAI-PMH umožňující sklizení metadat a tím pádem i spolupráci s jinými projekty. K importu digitálních dokumentů slouží formát XML, TXT a obrazové soubory v různých formátech (podpora DjVu, JPG, PNG, PDF). K identifikaci dokumentu slouží identifikátor zdroje URI, který umožňuje propojení pomocí vazby na daný digitální dokument přímo z online katalogu knihovny.

V roce 2007 byl do systému integrován také formát METS (*Metadata Encoding & Transmission Standard*), který je určen pro výměnu digitálních objektů mezi systémy [Polišenský, 2007, s. 1]. Další formáty metadat jsou PREMIS a MIX. Bibliografická data jsou exportována ve formátu MARC XML a Dublin Core, administrativní metadata ve formátu PREMIS [FOLTÝN, 2008-2011].

7.2 Uživatelské rozhraní systému Kramerius

Zpřístupňování dokumentů je v softwaru Kramerius realizováno pomocí dvou rozhraní – administrátorského a uživatelského.

Administrátorské rozhraní nabízí neomezený přístup do systému Kramerius a slouží k nahrávání dokumentů, umožňuje správu práv, snadné úpravy atd. Tituly je možné exportovat, vymazat, zviditelnit a zveřejnit, případně realizovat replikaci dat. Administrátorské rozhraní bylo dříve popsáno v kapitole 5.2.8 Zveřejnění pro uživatele.

Uživatelské rozhraní a vyhodnocení zpřístupnění a vyhledávání v uživatelském rozhraní systému Kramerius je popsáno v následující části této práce.

Uživatelské rozhraní umožňuje vyhledávat digitální dokumenty, zobrazovat metadata a tisknout je. V případě autorsky volného díla nebo v případě díla chráněného autorským zákonem v prostorách sítě MKP je možné prohlížet také digitální obrazy dokumentu.

Digitální knihovna je pro uživatele dostupná buď z přímého odkazu <http://kramerius.mlp.cz/kramerius/> nebo prostřednictvím odkazu z webových stránek Městské knihovny v Praze, konkrétně je odkaz na digitalizované dokumenty umístěn na stránkách e-knihovny sdružujících všechny elektronické texty MKP.

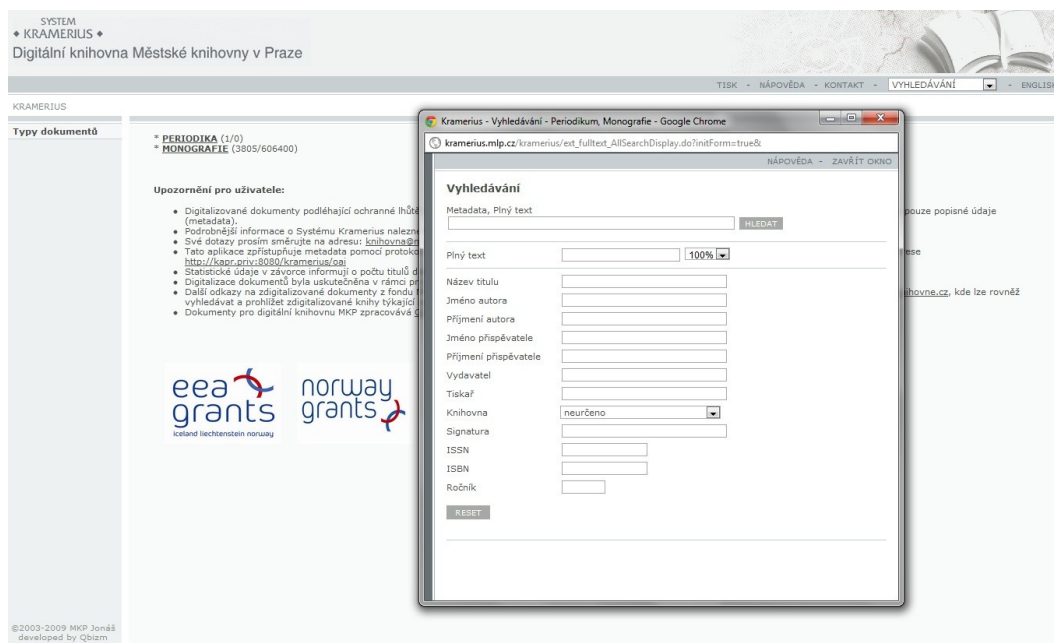
Úvodní stránka Digitální knihovny Městské knihovny v Praze obsahuje přehled dostupných digitalizovaných dokumentů: monografií a periodik a Upozornění pro uživatele. Upozornění obsahují základní informace o podmínkách zobrazení celého digitálního dokumentu, odkaz na nápovědu, emailovou adresu pro případné dotazy, informaci o zpřístupnění metadat pomocí protokolu OAI-PMH, informace o projektu HISPRA a odkazy na další elektronické dokumenty MKP včetně alternativního rozhraní pro uživatele z řad široké veřejnosti – www.prahavknihovne.cz.

V pravé podélné liště se nachází hlavní menu, jehož obsah se mění v závislosti na aktuálně zobrazené stránce. Hlavní lišta pak obsahuje odkaz na *Tisk*, *Nápovědu*, pole pro vyhledávání a volbu jazyka stránek. Pod hlavní lištou se nachází stavový řádek, kde je během procházení digitální knihovny uživatel informován, v jaké části digitální knihovny se nachází.

7.2.1 Vyhledávání

Uživatelé mohou nalézt digitální dokumenty prostřednictvím fulltextového vyhledávání, kde je integrován vyhledávač LUCENE, případně vyhledávání podle data. Obě možnosti jsou k dispozici v kontextu hlavní lišty.

Vyhledávání je v systému Kramerius realizováno nad popisnými metadaty a v případě, že je k dispozici také plný text ve formátu TXT, také nad plným textem dokumentu. Vyhledávat je možné dvěma různými způsoby: vyhledávání ve všech dokumentech a vyhledávání nad množinou typů dokumentů, tedy nad monografiemi či periodiky.



Obrázek 20: Fulltextové vyhledávání v systému Kramerius [Systém, c2003-2009]

Při výběru fulltextového vyhledávání ve všech dokumentech, které je možné spustit pouze z úvodní stránky digitální knihovny, se otevře vlastní dialogové okno s nabídkou polí, které lze prohledávat. První v pořadí je textové pole pro zadání vyhledaného textu. Pokud je do tohoto boxu zadán řetězec slov, jsou automaticky spojena booleovským operátorem OR, a jsou tedy vyhledány všechny dokumenty, které obsahují alespoň jedno z těchto slov. Vyhledávání v tomto boxu je realizováno nad plným textem a následujícími metadaty: *Název, Jméno a příjmení autora, Jméno a příjmení přispěvatele, Vydavatel, Tiskař a Signatura* (viz Obrázek č. 20).

Dále jsou v rámci vyhledávání ve všech dokumentech nabízena pole pro filtrování záznamů, při použití více parametrů pro filtrování se ve výsledcích vyhledávání zobrazí pouze dokumenty, které splňují všechny tyto parametry. Pro *Vyhledávání v plném textu* jsou zadána slova vyhledávána nad plným textem, více slov je automaticky spojeno operátorem OR. K vyhledávání v plném textu lze použít i jiné logické spojky AND a NOT. Existuje také možnost výběru míry shody hledaných slov, kterou je možné nastavit na hodnotu 10 % až 100 %. Čím více procent je označeno, tím větší musí být míra shody hledaných slov.

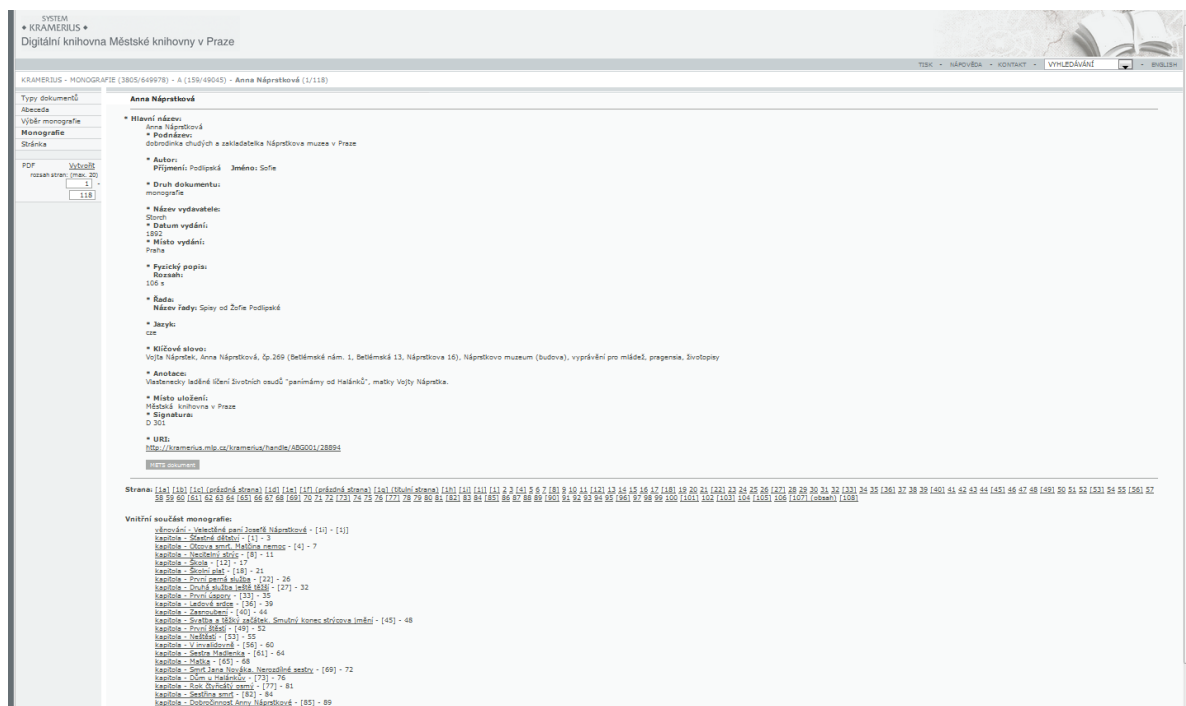
Další pole pro filtrování záznamů jsou již jednoduše vybrána z popisných metadat: *Název titulu, Jméno autora, Příjmení autora, Jméno přispěvatele, Příjmení přispěvatele, Vydavatel, Tiskař, Knihovna vlastníci dokument, Signatura, ISSN, Ročník*. Poslední dvě pole omezují výběr titulů pouze na režim *Periodika*.

Pro vyhledávání nad monografiemi a periodiky se nabídky prohledávaných polí zužuje pouze na pole používaná u daného typu dokumentů. Například při vyhledávání nad monografiemi nejsou k dispozici pole ISSN a ročník.

Po dokončení procesu vyhledávání se zobrazí seznam výsledků seřazený podle relevance, hledané výrazy jsou odlišeny pomocí žluté barvy. Pro možnost odečtení kontextu výrazu je kromě hledaného výrazu zobrazena také věta, ve které se vyskytuje. Jednotlivé záznamy seznamu výsledků hledání odkazují na stránku s informacemi o celém dokumentu. Velkou nevýhodu vidím v nemožnosti seskupování záznamů ze stejného dokumentu. Pokud se hledaný výraz vyskytuje v dokumentu víckrát, každý výskyt se odrazí v novém záznamu. Fulltextové vyhledávání v metadatech a v plném textu dokumentu bohužel nefunguje tak, jak by mělo. V některých případech se na dotaz vrací relevantní dokumenty, ovšem většinou nejsou do výsledků vyhledávání zahrnuty i dokumenty, o kterých je možné říci, že musí být zahrnuty.

Po výběru možnosti *Vyhledávání podle data* se nachází v rozbalovacím menu na hlavní liště a umožňuje vyhledat dokumenty podle roku, měsíce nebo dne vydání či tisku. Povinné je pole *Rok*, ostatní dvě pole *Měsíc* a *Den* mohou zůstat nevyplněna. Výčet let tedy neobsahuje celou datovou řadu, ale pouze roky vydání dokumentů dostupných v systému Kramerius.

Další možnosti, jak nalézt digitální dokument, je prostřednictvím volby typů dokumentů: monografií nebo periodik. Vyhledáváním nad periodiky se v tomto textu dále nebudu zabývat, stejně jako není výše v textu popsáno zpracování periodik. Po výběru monografií je k dispozici v dalším kroku abeceda názvů dokumentů a abeceda autorů. Po zvolení příslušného písmene z názvu je zobrazen seznam dokumentů vyhovujících tomuto výběru. „Vybrané písmeno je na všech stránkách nižších úrovní zobrazeno ve stavovém pruhu aplikace“ [Nápověda]. Po zvolení vhodného dokumentu je na další úrovni zobrazen bibliografický popis monografie, výpis všech dostupných stránek, vnitřních součástí nebo volných částí monografie. Například při hledání titulu s názvem Anna Náprstková stačí vybrat z abecedy názvu monografií písmeno A, v seznamu zobrazených dokumentů vyhledat ten s názvem Anna Náprstková. Po kliknutí na tento název se uživatelům zobrazí bibliografický popis díla (viz Obrázek č. 21).



Obrázek 21: Uživatelské rozhraní systému Kramerius [Systém, c2003-2009]

Výběrem počátečního písmene jména autora se zobrazí seznam autorových publikovaných děl dostupných v systému Kramerius. Čísla v závorce za písmeny udávají počet monografií a počet obrazových souborů s autory daného písmena. Po výběru zvoleného díla jsou uživateli k dispozici údaje bibliografického popisu a seznam stránek dokumentu.

Po zvolení periodik je zobrazena abeceda názvů periodik, po výběru příslušného názvu je zobrazen bibliografický popis celého titulu a seznam dostupných ročníků.

7.2.2 Vývoj systému Kramerius

Vzhledem k rostoucímu počtu knihoven používajících ke zpřístupňování svých digitalizovaných dokumentů systém Kramerius a vzhledem k rozvoji ICT bylo rozhodnuto o masivnější inovaci systému Kramerius. Dalším možným důvodem byla kritika uživatelského rozhraní, které není možné nazvat „user friendly“ tedy uživateli přívětivé. V roce 2009 tak bylo zadáno výběrové řízení na vývoj open source softwaru nové generace, nazvaném Kramerius 4. Výběrové řízení vyhrála firma INCAD, s.r.o., řešiteli projektu jsou Národní knihovna ČR, Knihovna Akademie věd ČR a Moravská zemská knihovna v Brně.

Hlavní rozdíl oproti starší verzi systému Kramerius je v repozitáři Fedora Commons, repozitář podporuje vlastní formát FOXML. V roce 2008 probíhal projekt s názvem Proof of Content, který měl za úkol prověřit možnost využití repozitáře Fedora pro další vývoj

systemu Kramerius, výsledkem projektu byl také konverzní nástroj pro převod stávajícího DTD pro monografie a periodika do formátu FOXML. Stejně jako předchozí verze je i systém Kramerius 4 k dispozici pod licencí GNU-GPL. Vyhledávání zajišťují nástroje Apache SOLR a LUCENE, nová verze je založena na technologii Java 1.6 a rozhraní je založeno na javascriptové knihovně JQuery. Metadata mohou být sklížena pomocí protokolu OAI-PMH, který nyní podporuje metadatový formát ESE využívaný digitální knihovnou Europeana. Systém byl představen v polovině roku 2010, od té doby stále probíhá jeho vývoj.

Systém Kramerius 4 je zatím reálně implementován a zveřejněn v Moravské zemské knihovně, v Knihovně akademie věd a v Národní technické knihovně je sice implementován, ovšem zatím je využit pouze pro malou skupinu dokumentů. V Městské knihovně v Praze se uvažuje o přechod na nový systém již od roku 2010. Implementace systému Kramerius 4 a s ním související konverze dokumentů ze starší verze je prioritním úkolem. Bohužel vzhledem k problémům při realizaci konverze starších digitalizovaných dokumentů ze systému Kramerius 3 do systému Kramerius 4 je tento úkol zatím posouván. Již bylo uvažováno o možnosti spuštění nové verze systému Kramerius, ve které by mohly být zpřístupněny nově vznikající digitální dokumenty a o paralelním běhu stávající verze systému Kramerius s původními dokumenty. Tato alternativa se mi ovšem z uživatelského hlediska jeví jako nešťastná, takže z pozice vedoucí digitalizačního pracoviště zatím trvám na spuštění nové verze i s původními dokumenty. Pokud se ovšem tato alternativa bude jevit jako nereálná i v druhé polovině roku 2012, budu pravděpodobně nucena k povolení dočasného chodu obou verzí systému Kramerius vedle sebe.

Hlavní výhoda nového systému tkví v možnosti přidání dalších typů dokumentů – kromě monografií a periodik například jednotlivé články nebo mapy. Další výhoda je v příjemnějším uživatelském rozhraní a v kvalitnější prezentaci výsledků. Systém Kramerius 4 podporuje dvouvrstvé PDF, takže je možné jej využít nejen pro digitalizované dokumenty, ale také pro *born digital dokumenty*, tedy dokumenty, které již vznikly v digitálním prostředí. Uživatelské rozhraní nabízí nové prvky – z obrazových souborů je možné generovat náhledy, zobrazit je pomocí filmového pásu a tím zvýšit komfort při prohlížení dokumentů (viz Obrázek č. 22).



Obrázek 22: Uživatelské rozhraní systému Kramerius 4 [SKOLKOVÁ, 2011a]

„Uživatelské rozhraní Kramerius 4 více orientováno na fulltextové vyhledávání dokumentů, doplněné o rychlý výběr skupin dokumentů pomocí faset a časové osy. Další změnou je způsob zobrazení vlastního obsahu vybraného dokumentu, kde je vždy vidět jeho celá struktura namísto „zanořování“ do nižších úrovní dokumentu – uživatel má vždy k dispozici přehlednou informaci o celkovém kontextu“ [LAHODA, 2011]. Užitečnou novinkou je také možnost třídit výsledky vyhledávání nejen podle typu dokumentu, ale také podle názvu či časového období. Velký dopad na uživatelskou základnu má také možnost vytřídit dokumenty autorsky volně a tím pádem dostupné pro všechny uživatele odkudkoli, a dokumenty chráněné autorským zákonem, jejichž obrazové soubory je možné zobrazit jen v síti dané knihovny.

Závěr

V knize Čtenáři a internauti Jiří Trávníček [2011] považuje digitalizaci za třetí informační revoluci v dějinách lidstva. Jiří Trávníček sice ve své knize nepoužívá pojem digitalizace pouze ve smyslu převodu od analogového k digitálnímu, ale poukazuje na schopnost přijímat informace prostřednictvím obrazovky, schopnost číst z monitoru. Nedává tak do souvislostí digitalizaci pouze s digitálními dokumenty. Přesto má v mnohém pravdu. Digitalizace může být v prostředí knihoven a potažmo i celé společnosti chápána jako hybatel další informační revoluce.

V éře digitalizace dominuje snaha o zachování kulturních hodnot informací obsažených v dokumentech často skladovaných v nevhodných podmínkách. Všechny tyto faktory vedou k potřebě uchránit nejen obsah, tedy informace uložené v těchto dokumentech, ale také formu, tedy vzhled dokumentu a u vzácných tisků také umění tiskaře a knihvazače.

Protože Městská knihovna v Praze má ve svém fondu také unikátní vzácné historické dokumenty, jejichž obsah stojí za to uchránit pro příští generace, bylo rozhodnuto informace v těchto dokumentech zachránit právě prostřednictvím digitalizace. Nejprve se MKP ucházela o granty z programu Veřejných informačních služeb knihoven, především o VISK č. 7. Ovšem rychlost s jakou byly dokumenty digitalizovány, nezbuzovala velkou naději – zatímco byl zachráněn jeden dokument, dalších deset jich podlelo větší zkáze. Degradace papíru a především nevhodné podmínky skladování v kombinaci s povodněmi si vybraly svou daň, bylo třeba hledat jiné cesty k záchraně dokumentů.

Proto byla v roce 2005 podána žádost o grant v rámci tzv. norských fondů. Přestože práce na vyhotovení grantu stála velké úsilí a mnoho hodin práce, vyplatila se a peníze z grantu byly získány. Mohlo se začít s přípravami. Velkou výhodou zisku finančních prostředků z norských fondů byla možnost vytvoření a vybavení vlastního digitalizačního pracoviště. Bylo třeba se poučit u jiné instituce, která se digitalizací zabývala. V té době nebylo příliš na výběr, hlavními vzory se stala pracoviště NK ČR a KnAV, obě v Praze a tedy v blízkosti MKP, s možností kdykoli konzultovat.

Nejdůležitější byl výběr správných skenerů. Od počátku bylo jasné, že oba skenery by bylo výhodné koupit od různých výrobců tak, aby splňovaly podmínky diverzity. Bohužel v době výběru skenerů se plnobarevné modely pohybovaly v nesrovnatelně vyšší cenové hladině. Proto byl vybrán model Minolta, na který byly získány dobré reference

z několika zdrojů. Od počátku se objevovaly potíže s udržením kvality barevného snímání, lampy nedokázaly přesvítit změny světelných podmínek v místnosti, kde byl skener umístěn. Později byly objeveny potíže s obrysovou ostroť. Zpětně bylo zjištěno, že instituce, které podaly reference na tento skener, jej využívali k jiným účelům a popsané vady jim nevadily. Po náročném vyřízení formalit byl pečlivě vybrán další skener tak, aby zůstala zachována různost značek. Nový model skeneru má název Bookeye 3 A2. Výběr skeneru je velmi důležitý, ovšem záleží také na kvalitě naskenování předlohy. Vždy je důležité mít na paměti, že čím lépe je předloha nasnímána, tím méně práce je s následnými úpravami.

Po překonání potíží s umístěním nově vznikajícího oddělení bylo pokročeno k úpravám vybraného prostoru. Ten není příliš vhodný pro snímání, jedna stěna je tvořena pouze okny. Muselo být tedy vyřešeno zatemnění místnosti speciálními neprůsvitnými žaluziemi.

Současně s rekonstrukcí a vybavováním prostoru byly upřesňovány a konzultovány jednotlivé složky procesu digitalizace.

Jako zásadní se ukázal být výběr vhodného formátu. Bylo rozhodnuto použít formát JPEG nejen pro prezentaci výsledků digitalizace, ale také pro archivaci. Rozhodnutí vyvolalo v prostředí českých knihoven vlnu připomínek. Po několika letech byl ovšem i formát JPEG označen jako vhodný formát pro archivaci, přestože samozřejmě formát TIFF s bezztrátovou kompresí, případně JPEG 2000 jsou v čele seznamu formátů vhodných pro archivaci. V souvislosti s plánováním Národní digitální knihovny byla v České republice zkoumána také možnost využití formátu JPEG 2000. Rozhodli jsme se ovšem zatím formát JPEG 2000 nenasadit, hlavním důvodem je nutnost časové i finanční investice na změnu vybavení podporujícího tento formát. Například verze softwaru Sirius užívaného k úpravě a popisu obrazových dokumentů dostupná v Oddělení digitalizace tento formát nepodporuje.

K úpravám a popisu obrazových dokumentů byl vybrán modulární software Sirius. Jeho neskonalá výhoda je v integraci všech procesů souvisejících se zpracováním naskenovaných dokumentů. K uživatelské přívětivosti přispívá také jednotné rozhraní všech modulů. Z mého pohledu má ovšem systém Sirius hned několik vážných nedostatků. Základní nedostatek spočívá v exkluzivním postavení firmy Elsys Engineering na českém trhu. S tím souvisí nepříliš pružné reakce na úpravy systému Sirius nejen pro potřeby Oddělení digitalizace MKP, ale také na obecné požadavky na úpravu softwaru. Nevýhodná je také poloha sídla firmy vzhledem k nutným servisním zásahům. Dnes bych po těchto zkušenostech raději zvolila jednotlivé programy na každý proces zpracování zvlášť. Pokud je

již software zastaralý či nevyhovuje stále se měnícím požadavkům na parametry zpracování a jiných změn, které jsou spojené s vývojem digitalizace, je daleko jednodušší vyměnit jeden z programů. A tato změna ve většině případů neovlivní programy ostatní.

Zatímco u modulárního systému jsou moduly propojeny a úzce spolu souvisí.

Výběr vhodného softwaru nebo kombinaci softwarů na zpracování obrazových dokumentů je tak dalším zásadním krokem k úspěšnému zvládnutí a hladkému průběhu zpracování výsledků snímání analogových předloh.

Podstatný prvek provázející celý proces digitalizace je kontrola – kontrola kvality obrazových souborů, provedených úprav, metadat atd. Dostatečně prováděná kontrola v průběhu celého procesu digitalizace je jeho nezbytnou součástí. Zpětné zásahy a opravy již zveřejněných digitálních dokumentů jsou časově, finančně i personálně náročné. V začátcích projektu HISPPRA byl podle mého názoru systém kontrolování lehce podceňen, především co se týče personálních kapacit. Tato situace měla za následek zpoždění kontrol a tím pozdní odhalení některých často se opakujících chyb. Po zavedení nového systému kontrol ubylo špatných návyků při zpracování dokumentů a možné chyby jsou odhaleny včas, než je dokument zpřístupněn v digitální knihovně.

Nezbytnou součástí procesu digitalizace je systém dlouhodobé archivace pořízených dat. Dlouhodobá ochrana dat je nesmírně důležitá, digitální dokumenty jsou totiž v podstatě ohroženy více než tradiční dokumenty sloužící jako jejich předloha. Pokud totiž není dostatečně podchycena problematika dlouhodobé ochrany, hrozí nenávratná ztráta informací v těchto dokumentech obsažených. Při nesprávně zvolené strategii paradoxně mohou informace z papírových nosičů přežít déle než informace v digitalizovaných dokumentech. V případě dat vzniklých z činností v rámci projektu HISPPRA se jedná spíše o promyšlený systém záloh než o dlouhodobou archivaci.

Zakončením celého procesu digitalizace je zpřístupnění digitalizovaného dokumentu v digitální knihovně Kramerius. Volba vhodného prostředí zpřístupnění dokumentů má zásadní vliv na četnost využívání digitálních dokumentů. Pro ztraktivnější poskytování informací obsažených v digitálních dokumentech uživatelům jsou digitální dokumenty nabízeny nejen v systému Kramerius, ale také v uživatelsky přívětivějším rozhraní. Zobrazení jednotlivých stránek v digitální knihovně Kramerius je možné až po minimálně čtyřech krocích ze strany uživatele. Pro uživatele zvyklé prvky webu 2.0 a na vyhledávání typu Google je vyhledávání v systému Kramerius verze 3.X návratem o několik let zpět.

Pro uživatele z řad široké veřejnosti, kteří tvoří uživatelskou základnu veřejných knihoven, byl proto spuštěn paralelní systém www.prahavknihovne.cz. Uživatelská přívětivost, ale také funkcionality systému Kramerius se výrazně zlepšily s nástupem nové verze nazvané Kramerius 4.

K digitalizaci byly vybrány dokumenty ze tří specializovaných úseků umístěných v Ústřední knihovně. První byly digitalizovány pragensia, tedy dokumenty pojednávající o Praze. Výběr dokumentů k digitalizaci byl možná trochu volnější, byly vybrány nejen tituly vydané do roku 1860, ale i novější dokumenty, jejichž historická a kulturní hodnota je prokazatelná. Mezi nejzajímavější digitalizované tituly jistě patří pražské adresáře a ukazatele domů, jejich využití je široké a široká je také uživatelská základna, kterou tyto typy dokumentů zajímají. Podle mého názoru mohly být digitalizovány přednostně v první fázi všechny nejvíce poškozené dokumenty ze všech oddělení namísto separace podle úseků. Na druhou stranu byla digitalizace po úsecích výhodná v tom, že při procházení fondu a výběru dokumentů podle digitalizačního plánu mohly být ještě nalezeny další dokumenty, které v první fázi nebyly vybrány, přesto si zasloužily převod do digitální podoby. Také evidence titulů nebyla tak obtížná. Po několika měsících digitalizace pouze pragensijních dokumentů začal výběr a proces digitalizace dokumentů z divadelního a filmového úseku. Zatímco ze Studovny pragensíí byly digitalizovány převážně dokumenty monografické povahy, většinu dokumentů z Divadelního a filmového úseku tvořily divadelní časopisy, tedy publikace periodické. Digitalizaci dokumentů z Hudebního úseku bylo nutné z personálních důvodů odložit, takže během trvání projektu HISPPRA bylo reálně digitalizováno jen malá část z původně vybraných dokumentů, většinu z nich tvořila původní česká libreta oper a operet.

Projekt HISPPRA přinesl Městské knihovně v Praze nejen plnohodnotně vybavené pracoviště Oddělení digitalizace a digitalizované tituly, ale také vysoké postavení v oblasti digitalizace v prostředí nejen veřejných ale i odborných knihoven. Digitalizace historických dokumentů byla počátkem elektronického publikování MKP. Již během projektu HISPPRA se začala rozrůstat oblast aktivit Oddělení digitalizace, k historickým dokumentům především odborného rázu přibývalo elektronické publikování beletrie. Po dokončení digitalizace dokumentů z vybraných úseků se nyní připravuje digitalizace dokumentů z Oddělení vzácných tisků na Brusnici, tedy i těch poškozených při povodních v roce 2002. Dále jsou k digitalizaci vybrány také dokumenty z depozitního fondu Městské knihovny v Praze.

Projekt HISPRA odstartoval proces vlastní digitalizační činnosti Městské knihovny v Praze, díky kterému byly zachráněny informace z více než 3 000 dokumentů, které jsou nyní pro uživatele k dispozici prostřednictvím digitální knihovny 24 hodin denně 7 dní v týdnu.

Seznam použité literatury

ABBYY (Kyjev, Ukrajina). *Co je OCR* [online]. Abby [cit. 2011-11-21]. Dostupné z: http://www.abbyy.cz/products/document_conversion/about_ocr/whatis_ocr/.

BALVÍNOVÁ, Alena. 2003-. Dpi. In: *KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha: Národní knihovna ČR [cit. 2011-10-23]. Dostupné z: http://aleph.nkp.cz/F/?func=direct&doc_number=000000167&local_base=KTD.

BALVÍNOVÁ, Alena. 2003-. Dpi. In: *KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha: Národní knihovna ČR [cit. 2011-11-28]. Dostupné z: http://aleph.nkp.cz/F/?func=direct&doc_number=000000167&local_base=KTD.

BCS-2 [software]. Verze 5.10. [Bonn]: ImageWare.

BENYOVSZKÁ, Věra. 2008. O Studovně pragensií. *Biblio*. Roč. 2, č. 5, s. 34. ISSN 1802-3320.

BÍNOVÁ, Jiřina. 1996. Regionální dokumenty ve veřejných knihovnách. Část 1. *Čtenář*, roč. 48, č. 1, s. 2-8. ISSN 0011-2321

BOOKEYE 3 : A2 Book Scanners. [2008]. ImageWare Components. Bonn. 6 s. Informační brožura. Dostupný také na WWW: http://www.incad.cz/public/91/30/7e/a8/1491_11673_BE3_R2_Web_E_1.pdf.

CELBOVÁ, Ludmila. 2003a-. Administrativní metadata. In: *KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha: Národní knihovna ČR [cit. 2011-10-27]. Dostupné z: http://aleph.nkp.cz/F/?func=direct&doc_number=000000512&local_base=KTD.

CELBOVÁ, Ludmila. 2003b-. DTD. In: *KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha: Národní knihovna ČR [cit. 2011-11-21]. Dostupné z: http://aleph.nkp.cz/F/?func=direct&doc_number=000000526&local_base=KTD.

CELBOVÁ, Ludmila. 2003c-. Popisná metadata. In: *KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha: Národní knihovna ČR [cit. 2011-10-27]. Dostupné z: http://aleph.nkp.cz/F/?func=direct&doc_number=000000549&local_base=KTD.

CELBOVÁ, Ludmila. 2003d-. Metadata. In: *KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha: Národní knihovna ČR [cit. 2011-10-28]. Dostupné z: http://aleph.nkp.cz/F/?func=direct&doc_number=000000543&local_base=KTD.

CELBOVÁ, Ludmila. 2003e-. Z39.50. In: *KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha: Národní knihovna ČR [cit. 2011-11-01]. Dostupné z: http://aleph.nkp.cz/F/?func=direct&doc_number=000000575&local_base=KTD.

CUBR, Ladislav. 2009. *Strategie ochrany digitálních dokumentů*. Praha. 141 s. Diplomová práce (Mgr.). Univerzita Karlova, Filozofická fakulta, Ústav informačních studií

a knihovnictví. Dostupná také z univerzitního digitálního archivu (URI):
http://digitool.is.cuni.cz/R/-?func=dbin-jump-full&object_id=89760&silos_library=GEN01.

ČERMÁKOVÁ, Zdeňka. 2003. *Městská knihovna v zrcadle století (1891 – 1991)*. Praha: Městská knihovna v Praze. 39 s.

ČESKO. 2000. Zastupitelstvo hl. m. Prahy. *Strategický plán hl. m. Prahy. Útvar rozvoje hl. města Prahy*. Odbor strategické koncepce. Praha. SWOT analýza, s. 12-16. Dostupný také na WWW: <http://www.monet.cz/strategplan/PDF/cesky/cz_3.pdf>.

ČESKO. 2001. Zákon č. 257 ze dne 29. Června 2001 o knihovnách a podmínkách provozování veřejných knihovnických a informačních služeb. In: *Sbírka zákonů České republiky*. Částka 98, s. 5683-5688. Dostupný také na:
<http://aplikace.mvcr.cz/archiv2008/sbirka/2001/sb098-01.pdf>. ISSN 1211-1244.

ČESKO. Zákon č. 216 ze dne 26. dubna 2006, kterým se mění zákon č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, a některé další zákony. In: *Sbírka zákonů České republiky*. 2006. Částka 72, s. 2707-2726. Dostupný také z WWW:
<http://aplikace.mvcr.cz/archiv2008/sbirka/2006/sb072-06.pdf>. ISSN 1211-1244.

Displej z tekutých krystalů. In: *Wikipedie: otevřená encyklopedie* [online]. San Francisco (Kalifornie): Wikimedia Foundation, 2001-, poslední editace 13. října 2011 [cit. 2011-11-21]. Dostupné z: http://cs.wikipedia.org/wiki/Displej_z_tekut%C3%BDch_krystal%C5%AF.

ELSYST ENGINEERING. ©2003a. *Architektura* [online]. Vyškov: Elsynt Engineering [cit. 2011-11-19]. Dostupné z: <http://www.ee.cz/sirius/architektura.htm>.

ELSYST ENGINEERING. ©2003b. *O firmě* [online]. Vyškov: Elsynt Engineering [cit. 2011-11-19]. Dostupné z: <http://www.ee.cz/ofirme.htm>.

ELSYST ENGINEERING. ©2003c. *Sirius: aplikace pro pořizování dat* [online]. Vyškov: Elsynt Engineering [cit. 2011-11-19]. Dostupné z: <http://www.ee.cz/sirius/index.htm>.

ELSYST ENGINEERING. [2003d]. *Sirius: Uživatelská příručka*. [Vyškov]: Elsynt Engineering. 184 s. Verze 2.0, dokument ve formátu PDF.

ELSYST ENGINEERING. ©2003e. *Služby* [online]. Vyškov: Elsynt Engineering [cit. 2011-11-19]. Dostupné z: <http://www.ee.cz/sluzby/index.htm>.

ELSYST ENGINEERING. [2007a]. *Rozšířené zobrazení*. [Vyškov]: Elsynt Engineering. 5 s. Dokument ve formátu DOC.

ELSYST ENGINEERING. [2007b]. *Sirius: Postup při zpracování dokumentů*. [Vyškov]: Elsynt Engineering. 22 s. Dokument ve formátu PDF.

FERDINANDOVÁ, Ivana. 2009. *Katalog : sporné otázky zpracování*. In: Intranet [online]. Městská knihovna v Praze: Praha [cit. 2010-12-11]. Dostupné z intranetu MKP.

FOLTÝN, Tomáš. 2008-2011. Systém Kramerius. In: *Kramerius System* [online]. Poslední změna 10. 03. 2009 [cit. 2011-06-03]. Dostupné z: <http://kramerius-info.nkp.cz/digitalizace-v-nk/system-kramerius/>.

GARTNER, Richard. 2008. *Metadata for digital libraries: state of the art and future directions*. [online] Version 1.0. Bristol: JISC [cit. 2011-10-09]. 19 s. Dostupné z: http://www.jisc.ac.uk/media/documents/techwatch/tsw_0801pdf.pdf.

GNU GPL. 2001-. In: *Wikipedie: otevřená encyklopedie* [online]. San Francisco (Kalifornie): Wikimedia Foundation, poslední editace 26. prosince 2011 [cit. 2011-12-27]. Dostupné z: http://cs.wikipedia.org/wiki/GNU_GPL.

HAVLŮJOVÁ, Hana. 2010. Zpráva o projektu Pragensia pro 21. století. In: *Digitalizace, aneb, Konec oslích uší: sborník vydáný u příležitosti stejnojmenné konference konané 14.-16. 6. 2010 v Městské knihovně v Praze*. Editor Eva Měřínská. Praha: Městská knihovna v Praze. s. 76-91. ISBN 978-80-85041-05-7.

HOKYNEK, Michal. 2007. *Digitalizace starých tisků a rukopisů z fondů Knihovny Vojenského historického ústavu*. Praha. 81 s. Diplomová práce. Univerzita Karlova v Praze, Filozofická fakulta, Ústav informačních studií a knihovnictví 2007. Vedoucí diplomové práce Jan Hutař.

IMAGEWARE. 2011. *About us* [online]. [Bonn]: ImageWare [cit. 2011-11-30]. Dostupné z: <http://www.imageware.de/en/unternehmen/details/>.

ISO 10918-1:1994. 2011. *Information technology - Digital compression and coding of continuous-tone still images: Requirements and guidelines*. 1st ed. Geneva: International Organization for Standardization, 2011-09-22.

ISO 12639:1998. 2007. *Graphic technology - Prepress digital data exchange - Tag image file format for image technology (TIFF/IT)*. 2nd ed. Geneva: International Organization for Standardization, 2007-09-07.

JANSOVÁ, Linda a Martin LHOTÁK. [2011a]. *Kramerius 4: nové řešení pro digitální knihovnu* [prezentace]. Praha: Knihovna Akademie věd ČR. Dostupné také z: <http://www.slideshare.net/lindask/kramerius-4-nov-eeen-pro-digitln-knihovnu-linda-jansov-martin-lhotk>.

JANSOVÁ Linda a Martin LHOTÁK. [2011b]. *Uživatelské rozhraní systému Kramerius 4. Čtenář* [online]. Roč. 63, č. 6. ISSN 0011-2321. Dostupný z: <http://ctenar.svkk1.cz/clanky/2011-roc-63/06-2011/uzivatelske-rozhrani-systemu-kramerius-4-85-933.htm>.

JÍLKOVÁ, Marta. 2010. *Historie a současnost projektu Kramerius*. Praha. 48 s, 6 s. příl. Bakalářská práce (Bc.). Univerzita Karlova v Praze, Filozofická fakulta, Ústav informačních studií a knihovnictví.

KNOLL, Adolf. 2008. *Digitální obraz* [prezentace]. Praha: Národní knihovna v Praze [cit. 2011-10-29]. 49 s. Část 2, Prezentace v rámci kurzu Digitalizace v knihovnách aneb Co musí vědět knihovník o digitalizaci. Dostupná také z: http://digit.nkp.cz/Training_en.html.

- KOCOUREK, Pavel. 2010. *Kramerius 4 : pohled pod pokličku* [online prezentace]. [cit. 2011-06-30]. Dostupné z WWW: < <http://www.inforum.cz/pdf/2010/kocourek-pavel.pdf>>
- Koniáš* [software]. Verze 8.114.0. Praha: Městská knihovna v Praze. Automatizovaný knihovní systém.
- KONICA MINOLTA EUROPE. 2011. *O společnosti* [online]. Konica Minolta Europe [cit. 2011-12-01]. Dostupné z: <http://www.konicaminolta.cz/business-solutions/company.html>.
- Konverzní tabulky. 2000. In: *Čeština.cz: Pověstná Velká stránka o kódování češtiny* [online]. 1. 3. 2000 [cit. 2011-11-02]. Dostupné z: <http://www.cestina.cz/tabulky/>.
- KOPENCOVÁ, Zuzana. 2008. *Digitální záchrana a zpřístupňování informací o Praze*. Biblio. Roč. 2, č. 5, s. 35-37. ISSN 1802-3320.
- KOVAČ, Pavel. 2011. Technologie LCD panelů. In: *Svět hardware: vše ze světa počítačů* [online]. oXy Online [cit. 2011-12-10]. Dostupný z: http://www.svethardware.cz/art_doc-7ADDD23432464B19C12571BD002A4AC4.html.
- LAHODA, Vladimír. 2011. UI: uživatelské rozhraní. In: *kramerius* [online]. Updated 7 November 2011 [cit. 2011-11-25]. Google Project Hosting. Dostupné z: <http://code.google.com/p/kramerius/wiki/UI>.
- LAMALS. 2006. *Lamals akciová společnost* [online]. [cit. 2011-10-12]. Dostupný z: <http://www.lamals.cz/index.html>.
- LEDERBUCHOVÁ, Ladislava. 2002. *Průvodce literárním dílem: výkladový slovník základních pojmů literární teorie*. 1. vyd. Jinočany: H & H. 355 s. ISBN 80-7319-020-6.
- LHOTÁK, Martin. 2010. Kramerius 4 - nová generace systému pro digitální knihovnu. In *11. konference Archivy, knihovny, muzea v digitálním světě 2010* [online]. Praha : Svaz knihovníků a informačních pracovníků ČR, 2010 [cit. 2011-11-25]. Dostupný z: <http://skip.nkp.cz/KeStazeni/Archivy10/den2/Lhotak.ppt>.
- LJUBKA, Ivan. Kramerius : vývoj aplikace pro zpřístupnění. In: *Knihovny současnosti 2008*. Brno : Sdružení knihoven ČR, 2008, s. 91-94. Dostupné také z: <http://www.svkos.cz/data/xinha/sdruk/2008-1-091.pdf>. ISBN 978-80-86249-49-0.
- Log. 2001-. In: *Wikipedie: otevřená encyklopedie* [online]. San Francisco (Kalifornie): Wikimedia Foundation, poslední editace 26. července 2011 [cit. 2011-09-21]. Dostupné z: <http://cs.wikipedia.org/wiki/Log>.
- MĚŘÍNSKÁ, Eva. 2008. Základem je kvalitní digitální kopie. *Ikaros* [online]. Roč. 12, č. 12 [cit. 2010-10-08]. Dostupné z: <http://www.ikaros.cz/node/5115>. URN-NBN:cz-ik5115. ISSN 1212-5075.
- MĚŘÍNSKÁ, Eva. 2009. *Pragensia v knihovně a ve škole* [online prezentace]. [cit. 2011-12-12]. Dostupná také z: <http://www.inforum.cz/pdf/2009/merinska-eval-cze.pdf>.
- MĚŘÍNSKÁ, Eva; VOJÍŘOVÁ, Kateřina. 2008. Projekt HISPRA - záchrana historických pragensií z fondu Městské knihovny v Praze (MKP). *Knihovnický zpravodaj Vysočina*

[online]. Roč. 10, č. 2 [cit. 2010-08-10]. Dostupné z:
<http://kzv.kkvysociny.cz/Default.aspx?id=1085>. ISSN 1213-8231.

MĚSTSKÁ KNIHOVNA (Praha, Česko). 1990. *Koncepce Studovny pragensii*. Alena Pavlíková. [Praha]: Městská knihovna v Praze. 8 s. Interní materiál.

MĚSTSKÁ KNIHOVNA (Praha, Česko). [20??]. *Profil fondu Studovny pragensii*. Věra Benyovszká [Praha]: Městská knihovna v Praze. 4 s. Interní materiál, dokument je ve formátu DOC.

MĚSTSKÁ KNIHOVNA (Praha, Česko). 2002a. *Budování fondů regionální literatury*. Kateřina Richterová. Praha: Městská knihovna v Praze. 8 s. Interní materiál.

MĚSTSKÁ KNIHOVNA (Praha, Česko). 2002b. *Současný stav budování fondů regionální literatury a jeho hodnocení*. Kateřina Richterová. Praha: Městská knihovna v Praze. 4 s. Interní materiál.

MĚSTSKÁ KNIHOVNA (Praha, Česko). 2004. *Obsahové charakteristiky fondu MKP*. [Praha]: Městská knihovna v Praze. 110 s. Interní materiál, dokument je ve formátu DOC.

MĚSTSKÁ KNIHOVNA (Praha, Česko). [2005a]. *Divadelní a filmový úsek ÚK MKP*. Praha: Městská knihovna v Praze. 1 s.

MĚSTSKÁ KNIHOVNA (Praha, Česko). [2005b]. *Hudební úsek ÚK MKP*. Praha: Městská knihovna v Praze, 1 s.

MĚSTSKÁ KNIHOVNA (Praha, Česko). [2005c]. *Projekt HISPRA : záchrana historických pragensijních a dalších vzácných dokumentů z fondu Městské knihovny v Praze*. [Praha]: Městská knihovna v Praze. [31] s., [33] s. příl. Žádost o grant z finančních mechanismů EHP/Norska.

MĚSTSKÁ KNIHOVNA (Praha, Česko). [2006a]. *Digitalizace tištěných dokumentů MKP: Divadelní a filmový úsek MKP*. Marie Valtrová. [Praha]: Městská knihovna v Praze. 2 s. Příloha č. 18 k žádosti o grant.

MĚSTSKÁ KNIHOVNA (Praha, Česko). [2006b]. *Digitalizace tištěných dokumentů MKP: Hudební úsek ÚK MKP*. Jana Navrátilová. [Praha]: Městská knihovna v Praze. 2 s. Příloha č. 17 k žádosti o grant.

MĚSTSKÁ KNIHOVNA (Praha, Česko). [2006c]. *Digitalizace tištěných dokumentů MKP: Studovna pragensii ÚK MKP*. Věra Benyovszká. [Praha]: Městská knihovna v Praze. 5 s. Příloha č. 16 k žádosti o grant.

MĚSTSKÁ KNIHOVNA (Praha, Česko). 2008. *Postup při ochraně pragensijních digitalizovaných dokumentů MKP*. Praha: Městská knihovna v Praze. 1 s. Interní materiál, dokument ve formátu DOC.

MĚSTSKÁ KNIHOVNA (Praha, Česko). 2009a. *Pravidla pro soutěž Výlety pana Broučka po Praze*. Praha: Městská knihovna v Praze. 2 s. Interní materiál, dokument ve formátu DOC.

MĚSTSKÁ KNIHOVNA (Praha, Česko). 2009b. *Zálohování dat*. Praha: Městská knihovna v Praze. 1 s. Interní materiál, dokument ve formátu DOC.

MĚSTSKÁ KNIHOVNA (Praha, Česko). [2010]. Program. In: *Digitalizace aneb Konec Oslích Uší* [online]. Praha: Městská knihovna v Praze [cit. 2011-11-09]. Dostupné z: <http://osliusi.mlp.cz/doku.php/program>.

MĚSTSKÁ KNIHOVNA (Praha, Česko). 2011. *Pobočky* [online]. Praha: Městská knihovna v Praze [cit. 2011-11-29]. Dostupné z: <http://www.mlp.cz/cz/kontakty/pobočky/>.

MĚSTSKÁ KNIHOVNA (Praha, Česko). *Studovna pragensií*. Praha: Městská knihovna v Praze. 3 s. Interní materiál.

Nápověda [online]. [cit. 2011-10-28]. Nápověda Digitální knihovny MKP Kramerius. Dostupný z: http://kramerius.mlp.cz/kramerius/ext_fulltext_help_cs.jsp#M1.

NÁRODNÍ KNIHOVNA ČESKÉ REPUBLIKY. 2007. *Pravidla popisu pro tvorbu metadat periodik a monografií*. Praha: NK ČR. Dostupné také z: http://digit.nkp.cz/projekty/VZ-2004_2010/2007/Prilohy/3.pdf/.

NEJEZCHLEBOVÁ, Jana, Kateřina POJEROVÁ, Linda SKOLKOVÁ a Lucie VAVŘÍKOVÁ. 2007. Zpráva z konference Knihovny současnosti 2007. *Ikaros* [online]. Roč. 11, č. 10 [cit. 2011-12-21]. Dostupný na World Wide Web: <<http://www.ikaros.cz/node/4349>>. URN-NBN:cz-ik4349. ISSN 1212-5075.

NOVÁKOVÁ, Vladimíra. 1978. *Sbírky pragensií v pražských institucích*. Praha, 70 s. Diplomová práce. Univerzita Karlova v Praze, Filozofická fakulta, Ústav informačních studií a knihovnictví.

Omniscan [software]. Verze 11. [Tübingen-Hirschau]: Zeutschel GmbH.

POLIŠENSKÝ, Jiří. 2009. Účast NK ČR v projektu Norské fondy : digitalizace bohemikálních monografií ohrožených degradací papíru. In *Knihovny současnosti 2009 : sborník ze 17. konference, konané ve dnech 23.-25. června 2009 v Seči u Chrudimi*. Brno : Sdružení knihoven ČR, s. 167-172. Dostupný z: <http://www.svkos.cz/data/xinha/sdruk/2009-6-167.pdf>. ISBN 978-80-86249-54-4.

POLIŠENSKÝ, Jiří. Implementace formátu METS v Systému Kramerius. In *Automatizace knihovnických procesů – 11 : sborník z 11. ročníku semináře pořádaného ve dnech 16.–17. května 2007 v Liberci*. Editor Daniela Tkačiková; Barbora Ramajzlová. Praha : ČVUT, 2007, s. 1-8. Dostupné z WWW: <<http://www.akvs.cz/akp-2007/13-polisensky.pdf>>. ISBN 978-80-01-03691-4.

POMIKÁLKOVÁ, Soňa. 2008. *Práce s regionálními dokumenty ve veřejných knihovnách se zaměřením na Městskou knihovnu v Praze*. Brno. 61 s. Masarykova univerzita, Filozofická fakulta, Ústav české literatury a knihovnictví, Kabinet knihovnictví. Dostupná také z: http://is.muni.cz/th/110909/ff_b/BC_prace_text.pdf.

Praha v knihovně [online]. 2009. Praha: Městská knihovna v Praze. [cit. 2011-10-29]. Dostupné z: <http://www.prahavknihovne.cz/>.

Praha. 2001-. In: *Wikipedie: otevřená encyklopedie* [online]. San Francisco (Kalifornie): Wikimedia Foundation, poslední editace 20. října 2011 [cit. 2011-11-21]. Dostupné z: <http://cs.wikipedia.org/wiki/Praha>.

PS 5000C. 2004. Konica Minolta. [Ramsey]. 2 s. Informační brožura. Dostupný také z: <http://kmbs.konicaminolta.us/main/content/products/attachments/productdocs/ps5000c.pdf>.

QBIZM TECHNOLOGIES. c2005-2009. *Systém Kramerius pro tvorbu a zpřístupnění digitalizovaných dokumentů* [online]. [cit. 2010-04-25]. Dostupný z: http://www.qbizm-technologies.cz/Reference/pripadove_studie/narodni_knihovna/index.html.

Registr digitalizace : evidence digitalizovaných dokumentů a sledování procesu zpracování [online databáze]. [cit. 2011-11-03]. Dostupné z: <http://www.registrdigitalizace.cz/rdcz/>.

RICHTEROVÁ, Alena. 2003-. Přívazek. In: *KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha: Národní knihovna ČR [cit. 2011-10-29]. Dostupné z: http://aleph.nkp.cz/F/?func=direct&doc_number=000002271&local_base=KTD.

RYŠÁNKOVÁ, Anna. 2007. *Řešení výstavby digitalizačního centra na příkladu Knihovny AV ČR, v.v.i.* Praha. 94 s. Diplomová práce. Univerzita Karlova v Praze, Filozofická fakulta, Ústav informačních studií a knihovnictví.

Sirius [software]. c2007-2009. Verze 2.5.0.3189. Vyškov: Elsys Engineering.

Skladba knihovního fondu specializovaných úseků ÚK. 2011. In: *Intranet* [online]. Městská knihovna v Praze: Praha [cit. 2011-11-10]. Dostupné z intranetu MKP.

STATE LIBRARY OF NORTH CAROLINA. [2007]. Digitization Guidelines: Chapter 5: Metadata. In: *North Carolina Echo: Exploring Cultural Heritage Online* [online]. Edited by Catherine M. Wisser. 2007 Revised Edition. Raleigh (Severní Karolína): State Library of North Carolina, Department of Cultural Resources [cit. 2011-11-02]. Dostupné z: http://www.ncecho.org/dig/guide_5metadata.shtml.

Strategický plán hl. m. Prahy. 2000. Útvar rozvoje hl. města Prahy. Odbor strategické koncepce. Praha. SWOT Analýza, s. 12-16. Dostupné také z: http://www.monet.cz/strategplan/PDF/cesky/cz_3.pdf.

Systém Kramerius: Digitální knihovna Městské knihovny v Praze [online]. c2003-2009 [cit. 2011-11-30]. Dostupné z: <http://kramerius.mlp.cz/kramerius/Welcome.do>.

Systém pro správu obsahu. 2001-. In: *Wikipedie: otevřená encyklopedie* [online]. San Francisco (Kalifornie): Wikimedia Foundation, poslední editace 20. října 2011 [cit. 2011-11-21]. Dostupné z: http://cs.wikipedia.org/wiki/Syst%C3%A9m_pro_spr%C3%A1vu_obsahu.

TRÁVNÍČEK, Jiří. 2011. *Čtenáři a internauti: obyvatelé České republiky a jejich vztah ke čtení (2010)*. Vyd. 1. Brno: Host ve spolupráci s Národní knihovnou České republiky. 191 s. ISBN 978-80-7294-515-3.

TYPO.CZ. 2009. Standardní velikosti papíru ISO. In: *Databáze: Pravidla a názvosloví* [online]. Aktualizováno 1. 10. 2009 [cit. 2011-11-02]. Dostupné z: <http://www.typo.cz/databaze/pravidla-a-nazvoslovi/standardni-velikosti-papiru-iso/>.

UHLÍŘ, Zdeněk. 2003-. Přítisk. In: *KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha: Národní knihovna ČR [cit. 2011-10-28]. Dostupné z: http://aleph.nkp.cz/F/?func=direct&doc_number=000002569&local_base=KTD.

Uživatelský portál systému Kramerius [online]. QBIZM, c2003-20010 [cit. 2010-04-25]. Dostupný z: <<http://kramerius.qbizm.cz/>>.

VOJTÁŠEK, Filip. 2000. *Role digitalizace v systému ochrany historických fondů se zvláštním zřetelem na projekt Národní knihovny ČR "Memoriae Mundi Series Bohemica" realizovaný v rámci programu UNESCO "Memory of the World"*. Praha. 154 s., 17 s. příl. Diplomová práce. Univerzita Karlova v Praze, Filozofická fakulta, Ústav informačních studií a knihovnictví.

VYCHODIL, Bedřich. 2007. *Obrazové dokumenty*. Praha. 131 s., V s. příl, 1 DVD. Diplomová práce. Univerzita Karlova v Praze, Filozofická fakulta, Ústav informačních studií a knihovnictví. Dostupná také z univerzitního digitálního archivu (URI): http://digitool.is.cuni.cz/R/-?func=dbin-jump-full&object_id=95071&silos_library=GEN01.

ZEUTSCHEL. [2011]. *Company* [online]. [Tübingen-Hirschau]: Zeutschel GmbH [cit. 2011-09-28]. Dostupné z: <http://www.zeutschel.com/company.html>.

Zeutschel OS 5000 TT. 2007. [Tübingen-Hirschau]: Zeutschel GmbH. 2 s. Informační brožura.

Seznam vyobrazení

Seznam tabulek

TABULKA 1: FINANČNÍ NÁKLADY PROJETKU HISPRA (PŘEVZATO Z [MĚSTSKÁ, 2007])	17
--	----

Seznam obrázků

OBRÁZEK 1: PROSTORY V PROVAZNICKÉ ULICI PŘED REKONSTRUKCÍ	18
OBRÁZEK 2: VYBAVENÉ ODDĚLENÍ DIGITALIZACE PO REKONSTRUKCI (AUTOR FOTOGRAFIE: VÁCLAV JIROUŠEK)	19
OBRÁZEK 3: SKENER ZEUSCHEL OS5000 TT(AUTOR FOTOGRAFIE: EVA MĚŘÍNSKÁ)	23
OBRÁZEK 4: SKENER MINOLTA PS 5000 C (AUTOR FOTOGRAFIE: EVA MĚŘÍNSKÁ).....	24
OBRÁZEK 5: SKENER BOOKEYE 3 A2 (AUTOR FOTOGRAFIE: VÁCLAV JIROUŠEK)	27
OBRÁZEK 6: POŠKOZENÝ SVAZEK URČENÝ K DIGITALIZACI (AUTOR FOTOGRAFIE: EVA MĚŘÍNSKÁ).....	50
OBRÁZEK 7: SKENOVÁNÍ V SW OMNISCAN NA PŘÍSTROJI ZEUSCHEL [OMNISCAN]	64
OBRÁZEK 8: OKNO S NASTAVENÍM PARAMETRŮ SKENOVÁNÍ V SW BCS-2 [BCS-2]	66
OBRÁZEK 9: SKENOVÁNÍ V SW BCS-2 NA PŘÍSTROJI BOOKEYE [BCS-2].....	67
OBRÁZEK 10: BIBLIOGRAFICKÝ ZÁZNAM DOKUMENTU ANNA NÁPRSTKOVÁ V AKS KONIÁŠ [KONIÁŠ].....	69
OBRÁZEK 11: VLOŽENÍ ČÍSEL STRÁNEK V SW SIRIUS [SIRIUS]	73
OBRÁZEK 12: OKNO MODULU INDEXOVÁNÍ V SW SIRIUS [SIRIUS]	75
OBRÁZEK 13: NÁHLED ZÁKLADNÍHO OKNA MODULU SIRIUS ADM [SIRIUS]	79
OBRÁZEK 14: ROZŠÍŘENÉ ZOBRAZENÍ V SW SIRIUS [SIRIUS]	83
OBRÁZEK 15: MANUÁLNÍ KONTROLA VYROVNÁNÍ V SW SIRIUS [SIRIUS].....	85
OBRÁZEK 16: NASTAVENÍ EXPORTU DAT V SW SIRIUS [SIRIUS]	90
OBRÁZEK 17: ADMINISTRÁTORSKÉ ROZHRAŇÍ SYSTÉMU KRAMERIUS [SYSTÉM, C2003-2009].....	91
OBRÁZEK 18: SCHÉMA ZÁLOHOVÁNÍ DAT V ODDĚLENÍ DIGITALIZACE MKP [MĚSTSKÁ, 2009B, S. 1]	94
OBRÁZEK 19: UŽIVATELSKÉ ROZHRAŇÍ WEBOVÉ STRÁNKY PRAHA V KNIHOVNĚ [PRAHA, 2009]	99
OBRÁZEK 20: FULLTEXTOVÉ VYHLEDÁVÁNÍ V SYSTÉMU KRAMERIUS [SYSTÉM, C2003-2009].....	105
OBRÁZEK 21: UŽIVATELSKÉ ROZHRAŇÍ SYSTÉMU KRAMERIUS [SYSTÉM, C2003-2009]	107
OBRÁZEK 22: UŽIVATELSKÉ ROZHRAŇÍ SYSTÉMU KRAMERIUS 4 [SKOLKOVÁ, 2011A].....	109

Seznam grafů

GRAF 1: TEMATICKÉ ČLENĚNÍ FONDU HISTORICKÝCH PRAGENSÍ	43
GRAF 2: TEMATICKÉ ČLENĚNÍ OCHU PRAŽSKÁ KULTURA, UMĚNÍ A SPOLEČNOST	44
GRAF 3: TEMATICKÉ ČLENĚNÍ OCHU PRAŽSKÁ KULTURA... (DIGITALIZOVANÉ DOKUMENTY)	45
GRAF 4: JAZYKOVÉ VARIANCE FONDU HISTORICKÝCH PRAGENSIJNÍCH DOKUMENTŮ.....	46
GRAF 5: JAZYKOVÉ VARIANCE FONDU HISTORICKÝCH PRAGENSÍ URČENÝCH K DIGITALIZACI	47
GRAF 6: ČASOVÉ ROZLOŽENÍ FONDU HISTORICKÝCH PRAGENSIJNÍCH DOKUMENTŮ	48
GRAF 7: ČASOVÉ ROZLOŽENÍ FONDU DIGITALIZOVANÝCH HISTORICKÝCH PRAGENSÍ.....	48

Seznam příloh

PŘÍLOHA 1: SEZNAM PRAGENSIJNÍCH OBSAHOVÝCH CHARAKTERISTIK [MĚSTSKÁ, 2004].....	I
PŘÍLOHA 2: PŘÍKLAD EXPORTU BIBLIOGRAFICKÉHO ZÁZNAMU VE FORMÁTU XML Z AKS KONIÁŠ [KONIÁŠ]...	VII
PŘÍLOHA 3: PŘÍKLAD EXPORTNÍHO XML ZE SW SIRIUS (PRO ZJEDNODUŠENÍ KRÁCENO) [SIRIUS]	VIII