

# **REFORMÁTOVÁNÍ, DIGITÁLNÍ KNIHOVNA, KRAMERIUS – PRO VÁS A TROCHU JINAK**

*Františka Vrbenská*

Kramerius

Úložiště

Problematika autorského práva (google)

Norské fondy

## **1 Reformátování novodobých dokumentů**

### *1.1 Začátky záchrany*

Před dvanácti lety jsem měla čest a příležitost přednést na této konferenci příspěvek o ochranném reformátování; tehdy ještě o mikrofilmování knihovnických dokumentů, ať už se jednalo o tituly historicky cenné, nebo svazky tzv. „novodobé“, ale kulturně významné a ohrožené fyzickou destrukcí.

V té době již probíhal projekt „Memoria mundi“, zaměřený na přímou digitalizaci nejvzácnějších dokumentů světového kulturního dědictví. V porovnání s „Paměti světa“ působily naše sny o digitalizaci kriticky ohrožených sbírek starých novin a časopisů – navíc s výhledem na možnost fulltextového vyhledávání – jako fantazie. Obecné mínění odborníků znělo: „Periodika? Příliš náročné, rozsáhlé, technicky komplikované, příliš nákladné.“ Ale skutečnost, že Francouzská národní knihovna v roce 1994 přikročila ke skenování části významných monografií z devatenáctého století, např. sebraných spisů Julese Vernea, byla povzbuzující.

Dokumenty, které byly publikovány po roce 1800, nepostrádají historickou, kulturní a informační hodnotu. V mnoha případech se jedná o tituly pozoruhodné, vyhledávané a přitom soliterně dochované. Ačkoliv mají své nezastupitelné místo mezi literárními památkami a slouží stále velkému počtu uživatelů, jejich degradace způsobená vnějšími vlivy i vnitřními příčinami je nevyhnutelná. Jejich přenesení na trvanlivější archivní médium, mikrofilm, však neposkytovalo optimální nástroj ke zpřístupnění. Řešení nabízela hybridní metoda reformátování – digitalizace mikromédií.

## 1.2 Od mikrofilmování k digitalizaci

Od r. 1997 bylo započato v rámci národního programu ochrany knihovních sbírek s mikrofilmováním významných bohemikálních periodik 19. století. Projekt byl nazván na počest slavného českého buditele Václava Matěje Krameria jeho jménem. Bezprostředně navázalo reformátování dalších periodických titulů, důležitých z hlediska celospolečenského, odborného nebo regionálního. Většina periodik nasnímaných na mikrofilm byla následně digitalizována. Současně se však otevíraly další odborné a technické otázky.

Byly stanoveny zásady pro výběr dokumentů k reformátování a pro volbu vhodné technologie, která odvisí od stavu a míry využití konkrétního dokumentu. Řešili jsme problematiku evidencí, katalogizace, přípravu dokumentů k reformátování, začleňování reformátovaných kopií do fondu a do automatizovaného knihovního systému. Neopomněli jsme ani zajistit bezpečné uložení již reformátovaných svazků s využitím ochranných obalů.

Byla vybudována infrastruktura, zahrnující i centrální archivaci mikrofilmů s pravidelnými kontrolami chemické stability. Archivaci zdrojových digitálních dat včetně migrací a konverzí prospělo úspěšné vyřešení standardizace metadat ve formátu SGML, nověji v XML.

Proces mikrofilmování a digitalizace mikrofilmu bylo – a stále je – nutné dolaďovat a optimalizovat. Hybridní metoda a důraz na zálohování kopií reformátovaných dokumentů jako ideová východiska zůstávají. K technologii digitalizace mikrofilmů připojily výstupy z hybridní kamery, která současně mikrofilmuje a skenuje. Digitální repozitáře začínají získávat na bezpečnosti a spolehlivosti, což v brzké budoucnosti ovlivní požadavky na mikrofilm jako výhradní archivní médium.

V souvislosti s vývojem knihovnických systémů a softwarových i hardwarových nástrojů prochází program „Kramerius – reformátování ohrožených a vzácných dokumentů“ neustálým rozvojem a zdokonalováním. Těto výzvě a potřebě, která by měla prospět knihovním sbírkám i čtenářům, se snažíme dostát.

## 2 Digitální knihovny

### 2.1 Zpřístupnění a uchování elektronických dokumentů

S rozmachem informačních technologií se zrodily systémy pro správu a dlouhodobé využívání digitálních dokumentů – digitální knihovny. V první fázi před-

stavovaly souhrn dokumentů či souborů informací, které již byly generovány přímo pro virtuální prostředí; pak se jejich prostřednictvím šířily rovněž elektronické verze tištěných materiálů. Poměrně brzy však přibýly digitalizované tradiční dokumenty ze sbírek knihoven, muzeí, archivů.

Digitální knihovny poskytly nejen badatelům či případně studujícím, ale celé široké čtenářské obci přístup k rukopisům, unikátním periodikům, k archivním zvukovým materiálům. V roce 1999 byly uspořádány jedny z prvních mezinárodních konferencí o digitálních knihovnách – dnes už hovoříme o digitálním věku. Ve světě funguje více než 70 projektů digitálních knihoven, nejméně čtyřicet pracovišť a výzkumných programů aktivních v dané oblasti, jíž se věnuje na půl tuctu specializovaných periodik.

V rozmachu programů, které si vzaly za cíl zpřístupnění obsahů paměťových institucí, hrozilo Evropě, že se ocitne pozadu oproti jiným teritoriím, dokonce i některým zemím třetího světa (Brazílie, Čína). Ani Evropská unie nehodlá zanedbat virtuální zpřístupnění svého velkolepého kulturního dědictví, v rámci digitalizace však věnuje pozornost také šíření vědeckých informací. Program „i2010: Digitální knihovny“ staví na dostupnosti on-line, na digitalizaci analogových sbírek, které by tak mohly být plně využívány, a klade důraz na jejich uchování a ochranu nejen v materiální, nýbrž také elektronické podobě. Primární digitalizace by měla být financována z národních rozpočtů.

## *2.2 Národní digitální knihovna*

Uchování a ochrana digitálních dokumentů je ve světě obvykle pokládána za úkol národních paměťových institucí: národní knihovny, národního muzea, národního archivu.

Nestačí mít k dispozici nástroje moderních informačních technologií a ambiciózní projekty.

V České republice digitalizační aktivity podepřely strategické materiály: „Státní informační a komunikační politika“, „Koncepce rozvoje knihoven v České republice na léta 2004 až 2010“ a „Koncepce trvalého uchování knihovních sbírek tradičních a elektronických dokumentů v knihovnách ČR do roku 2010“, kterou zpracovala Národní knihovna ČR ve spolupráci s Ministerstvem kultury ČR. Jinou záležitostí jsou však dostatečné finanční zdroje, které přes dobře míněné záměry pokulhávají před hrozivým tempem stárnutí a degradace jedinečných dokumentů.

Přesto se podařilo v širším kontextu České digitální knihovny, složené zejména z množství spíše užitných digitálních dokumentů oborového, regionálního, insti-

tucionálního i jiného charakteru, vybudovat Národní digitální knihovnu, jejímž iniciátorem a hlavním realizátorem byla Národní knihovna ČR.

Národní digitální knihovna postupně a průběžně shromažďuje základ národního kulturního dědictví, který tvoří digitální i digitalizované dokumenty; uchovává je a zpřístupňuje uživatelům. Tyto elektronické materiály pocházejí ze tří národních programů, realizovaných Národní knihovnou ČR.

### 2.2.1 *WebArchiv*

Relativním, ale velmi agilním nováčkem je program WebArchiv, zaměřený na shromažďování, uchovávání a zpřístupňování elektronických dokumentů publikovaných prostřednictvím internetu v rámci národní domény cz.

### 2.2.2 *Manuscriptorium*

Nejstarší a nejuctyhodnější národní program představuje Memoria – Manuscriptorium <http://www.manuscriptorium.com>, jež staví na digitalizaci vzácných rukopisů a starých tisků. Je otevřeným katalogem historických fondů, které u vybraných rukopisů a starých tisků doprovázejí obrazové kopie. V Manuscriptoriu je rovněž budován souborný katalog rukopisů. Díky využití speciálních nástrojů a elektronické edice dovoluje Manuscriptorium zařazování prezentovaných informací do kontextu. Provoz je financován Národní knihovnou ČR, provozovatelem této služby je firma AiP Beroun, s. r. o.

S Manuscriptoriem spolupracuje a přispívá do něj svými digitalizovanými vzácnými dokumenty řada institucí, včetně významných kulturních subjektů, jimiž jsou např. Knihovna Národního muzea v Praze, Vědecká knihovna v Olomouci, Muzeum východních Čech v Hradci Králové, Moravská zemská knihovna v Brně; od počátku roku 2007 přibyl zahraniční účastník – Národní knihovna Turecka v Ankaře. V současné době je do digitální podoby převedeno 1700 rukopisů, což představuje 800 000 obrazových souborů.

## 3 **Kramerius**

Národní program Kramerius přispívá do Národní digitální knihovny digitalizovanými novodobými dokumenty (od r. 1801 až po sedmdesátá léta dvacátého století), ohroženými degradací kyselého papíru. Uchovávaná periodika v elektronické formě nepocházejí pouze z Národní knihovny, jsou výsledkem úsilí několika desítek českých paměťových institucí, které se snažily důležité tituly zachránit prostřednictvím reformátování.

Od počátku digitalizačního procesu přes budování Digitální knihovny se jako zásadní mezníky dalšího vývoje ukazovaly standardizace dokumentů (zejména metadat), začlenění reformátovaných dokumentů do knihovního systému, tvorba fulltextu; především pak klíčový problém digitalizace: archivace dat, vázaná na dostatečnou interoperabilitu a prostředky pro archivaci dokumentů.

Archivace digitálních dokumentů byla řešena systémem, který se skládal z robotické magnetopáskovou knihovny ADIC Scalar 1000, diskového subsystému a souborového systému (HSM) SAM FS. Obrazové soubory jsou uchovávány ve file systému na třech magnetických páskách (technologie AIT-3), z nichž knihovna obsahuje dvě identické on-line kopie. Třetí páska je uložena mimo pracoviště v trezoru. Metadata jsou uložena na diskovém poli. Systém zabezpečuje centrální ošetření dat, např. konverze do nových formátů, migrace na nové prostředky apod. V současné době prochází archivační systém upgradem, neboť již kapacitně ani technicky nevyhovuje.

### *3.1 Změny a projekty*

Digitální knihovna NK ČR – později Národní digitální knihovna – je předmětem podprogramu VISK 4, v jehož rámci je realizována řada programů výzkumu a vývoje, zaměřená na získávání, strukturaci, archivaci a zpřístupnění digitálních dokumentů. K těmto výchozím funkcím se přidružuje výzkum v oblasti stability médií.

Reformátovaná periodika uložená v Digitální knihovně byla nyní propojena se souborným katalogem a s Jednotnou informační bránou (JIB). Napomohlo tomu přidělení ISBN, ačkoliv se leckdy jedná o tituly velmi staré a z bibliografického hlediska spletité. Díky pochopení kolegů z národní agentury pro ISBN při Státní technické knihovně se daří mnohdy velmi komplikované případy řešit.

Co se týče fulltextového vyhledávání, dokumenty procházejí konverzí do OCR, a to i zpětně (nebylo zpočátku digitalizace možné aplikovat). V současné době je realizován projekt na vybudování znalostní báze založené na slovních a frazeologických bázích pro časové jazykové vrstvy z období let 1800 až 1989. Vznikne rovněž editační a konverzní nástroj na tvorbu metadat pro digitální objekty, takže bude možné přejímat a využívat data vytvořená v jiných systémech. Jiným cílem je implementace kontejnerového formátu METS pro zveřejňování metadat uložených v systému Kramerius.

Jediné, co se v dohledné době nepodaří vyřešit, je omezení daná respektováním autorského zákona. Ačkoliv se v mnoha případech jedná o dokumenty publiko-

vané před více než sto lety či naprosté unikáty, mohou být zpřístupněny pouze na intranetu NK ČR. Možná však, že na tuto nesnáz bude lékem konsorcium paměťových institucí a dohoda s autorskými svazy...

### 3.2 Stěžejní projekty reformátování v Národní knihovně ČR

Kramerius I. – projekt ochranného mikrofilmování. 1997–1998

Digitalizace mikromédií. 1997–1999

Kramerius – Digitalization of Microforms. 1998

Digitální knihovna: produkce, ochrana a zpřístupnění digitálních dokumentů. 1999–2003

Optimalizace hybridní technologie ochranného reformátování. 2000–2001

Automatizace tvorby metadat při digitalizaci periodických dokumentů. 2002

Vytvoření virtuálního badatelského prostředí pro zpřístupnění a ochranu digitálních dokumentů. 2004–2010

Archivace a zpřístupnění dokumentů Národní digitální knihovny prostřednictvím centrálního datového úložiště. I. fáze. 2006

Záchrana neperiodických bohemikálních dokumentů 19. století ohrožených degradací papíru – 2007–2010

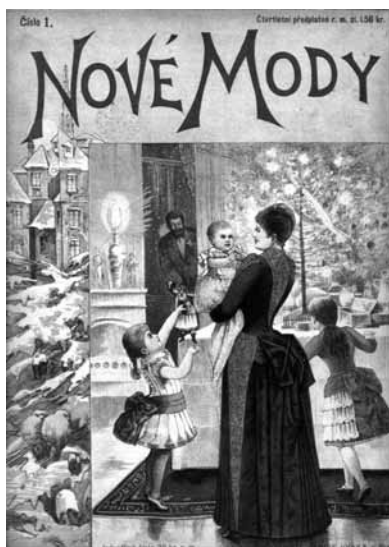
Projekty v rámci podprogramů VISK 4 – Digitální knihovna, a VISK 7 – Kramerius. 2001–2007

***Od r. 1992 bylo mikrofilmováno více než 600 bohemikálních novin a časopisů a 300 raritních periodik ze sbírek Slovanské knihovny; k nim přibude do r. 2010 přes 6 tisíc monografií a nejméně padesátka dalších periodik. Na 400 titulů – téměř 2 mil. obrazových souborů – včetně průběžně doplňovaných Národních listů je nyní přístupno v Digitální knihovně NK ČR prostřednictvím Systému Kramerius: <http://kramerius.nkp.cz/kramerius/Welcome.do>.***

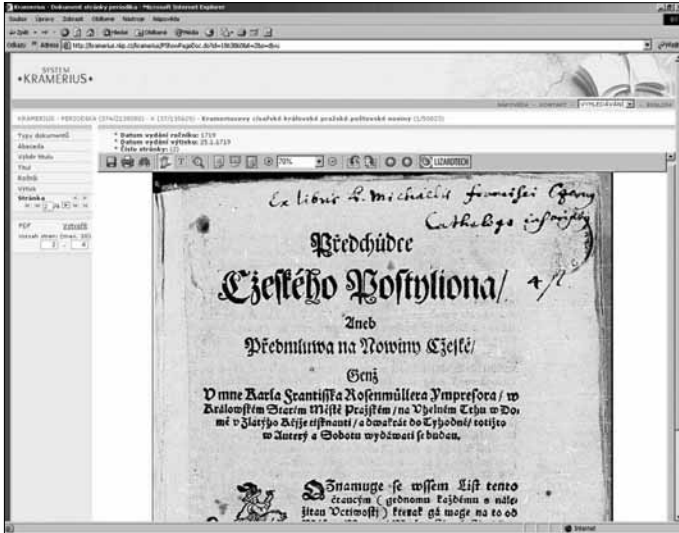
## Obrazová příloha



Obr. 1



Obr. 2



Obr. 3



Obr. 4