

ÚČAST NK ČR V PROJEKTU NORSKÉ FONDY – DIGITALIZACE BOHEMIKÁLNÍCH MONOGRAFIÍ OHROŽENÝCH DEGRADACÍ PAPÍRU

Jiří Polišenský, Národní knihovna ČR

Projekt „Záchrana neperiodických bohemikálních dokumentů 19. stol. ohrožených degradací papíru“, realizovaný v rámci podpory EHP a Norska, patří mezi největší digitalizační projekty v České republice. V rámci projektu bude digitalizováno a zpřístupněno prostřednictvím internetu cca 20 000 monografií 19. stol., což představuje 2 400 000 stran dokumentů. Takový objem již představuje poměrně slušnou speciální knihovnu. Realizace tak rozsáhlého projektu byla možná díky předcházejícím aktivitám NK ČR na poli digitalizace.

Národní knihovně ČR se podařilo poměrně záhy vyřešit standardizaci digitalizace, a to nejprve u vzácných rukopisů, po té i u novodobých dokumentů. Koncem 90. let se v NK ČR podařilo vybudovat pracoviště digitalizace založené na hybridní metodě využívající mikrofilm. Od r. 2000 pak bylo digitalizováno mnoho zejména novinových titulů, které byly nejvíce poškozeny dlouholetým využíváním. Z průzkumů fondů NK ČR, ale i z vlastních zkušeností jsme věděli, že degradace papíru postihuje i monografické publikace, z nichž mnohé již byly na pokraji úplného zničení, proto bylo rozhodnuto využít výzvy v 1. kole tohoto programu a zaměřit cíle připravovaného projektu právě na tento typ dokumentu.

V rámci přípravy projektu v r. 2005 byly odhadnuty objemy digitalizace, které zvládnou dodavatelské firmy danou technologií a byly odhadnuty objemy dat, které bude možné archivovat prostřednictvím předpokládané kapacity datového úložiště. Na základě toho byly kalkulovány objemy digitalizace a z toho vyplývající náklady ve výši 1,2 mil. EUR (1 mil. dotace, 0,2 mil. spoluúčast ČR). Celková doba realizace projektu byla stanovena na 48 měsíců, do konce r. 2009. Protože se ale schvalování projektů zdrželo byla doba realizace zkrácena na 28 měsíců.

Do schválení projektu byly vkládány velké naděje, protože dosud neexistoval program, v rámci kterého by bylo možné hradit rutinní digitalizaci většího objemu. Projekty VaV a výzkumné záměry byly zaměřeny na vývoj a optimalizaci technologií, evropské projekty, se kterými již měla NK ČR

bohaté zkušenosti, umožňovaly hradit spíš pořádání konferencí, příp. také financovat vývoj, nebo koordinační aktivity, ale nebylo možné využít finanční prostředky z těchto zdrojů na běžnou digitalizaci. Prostředky v rámci programu VISK 7 byly omezené.

V rámci programu bylo vyhlášeno celkem 8 prioritních oblastí, z nichž naší problematice odpovídala 1. PO „Uchování evropského kulturního bohatství“. Pro podání projektu bylo nutné zvolit typ projektu (individuální projekt, program, blokový grant) a vyplnit formulář s rozsáhlými přílohami, v rámci kterého bylo nutné uvést nejen specifikaci předmětu projektu a finanční rozpočet, ale také legislativu související s projektem, plán publicity projektu, řízení rizik a tzv. horizontální témata (udržitelný rozvoj, rovnost příležitostí, zásady dobré správy atd.).

Projekty se podávaly na kontaktní místo pro podávání projektů (KM), což v našem případě bylo Ministerstvo kultury ČR. Po zapracování připomínek byly vybrané projekty postoupeny Národnímu kontaktnímu místu (NKM), což bylo Centrum pro zahraniční pomoc Ministerstva financí. Po zapracování připomínek z této instance bylo nutné vypracovat anglickou verzi, která se postupovala na schválení Kanceláři finančního mechanismu Norska (KFM). Úspěšné schválení projektu koncem r. 2006 bylo dáno pravděpodobně tím, že byly jasně vymezené cíle a indikátory projektu, přehledně zpracované výdaje, dodány všechny přílohy a zapracovány všechny připomínky.

Projekt záchranu ohrožených monografií 19. stol. je postaven na využití hybridní technologie. K dispozici je technologie skenování mikrofilmů a hybridní snímkování na hybridní kameře, umožňující zhotovovat současně mikrofilm i digitální záznam. Druhým charakteristickým rysem projektu bylo dodavatelské zajištění všech prací. To znamenalo, že NK ČR nemusela řešit problémy s nákupem a instalací technického vybavení, zajištěním prostoru pro vybudování pracoviště, získáváním pracovníků příslušné specializace atd. Bylo nutné zorganizovat výběrová řízení a v rámci nich podrobně specifikovat požadavky na mikrofilmy a digitální dokumenty.

Ze čtyř výběrových řízení vzešly následující firmy jako dodavatelé:

1. AMPACO ČR pro zhotovování mikrofilmů a skenování mikrofilmů
2. MICRONA pro hybridní snímkování
3. Elsys Engineering pro zpracování dat

Všechny vybrané firmy musely přizpůsobit své kapacity pro objem činností vyplývajících z projektu, což znamenalo doplnění technického vybavení a dočasné rozšíření počtu pracovníků. Firma MICRONA musela před zahájením prací stihnout nakoupit a zprovoznit hybridní kameru. Výběrová řízení se podařilo realizovat v 1. pololetí 2007 a krátce po té i uzavřít smlouvy s dodavateli, vlastní činnost byla oficiálně zahájena v červnu 2007 a postupně byly do procesu realizace zapojovány všechny vybrané firmy.

Neoficiálně však již od konce r. 2005 probíhaly přípravné práce na realizaci projektu. Bylo nutno vytipovat velké množství monografií pro mikrofilmování a digitalizaci, ověřit jejich úplnost a zajistit katalogizaci a zpracování jednotek exempláře v případě, že příslušná kniha neprošla ještě procesem retrokonverze a neměla záznam v elektronickém katalogu NK ČR a nebyla vybavena čárovým kódem. Zpočátku bylo třeba část dokumentů mikrofilmovat a digitalizovat na vlastních pracovištích, aby bylo možné co nejdříve využít i firmu Elsynt Engineering, která navazovala na obě firmy zajišťující mikrofilmování a skenování, a byla tudíž na konci celého procesu.

Finanční prostředky na realizaci poskytuje Ministerstvo kultury ČR, a to jak na vlastní dotaci, tak i na spoluúčast, a to na základě uzavřené smlouvy. Realizace projektu je rozdělena na čtvrtletí, pro která jsou plánovány objemy produkce a čerpání finančních prostředků. Po skončení daného čtvrtletí je nutné zpracovat čtvrtletní zprávu (Project Implementation Report), která se dává schválit na NKM a po té posílá do Kanceláře finančního mechanismu. V této zprávě se vyčíslí veškeré náklady a doloží výpisem z banky jejich čerpání. Po schválení KFM jsou uhrazeny oprávněné náklady Ministerstvu kultury ČR, a to prostřednictvím Ministerstva financí ČR.

Firma AMPACO ČR zhotovuje z dodaných publikací archivní a matriční negativní mikrofilmy a provádí jejich skenování. Výsledkem jsou zdrojové obrazové soubory ve formátu jpg. Firma MICRONA zhotovuje archivní negativy a skeny ve stejném formátu. Firma Elsynt Engineering provádí ořez a narovnání, OCR, konverze do uživatelského formátu DjVu a vytváří popisná a administrativní metadata.

Požadavky na proces mikrofilmování a digitalizace jsou identické s těmi, které platí pro program VISK 7. V případě mikrofilmu to znamená respektování normy ISO 4087. V případě digitalizace jsou používané formáty pro archivaci JPEG, pro zpřístupnění DjVu. Metadata se vytváří podle národního standardu, který definuje popis pro jednotlivé typy dokumentů (DTD pro monografie)¹. V rámci zpracování dat se vytváří OCR soubory ve formátu TXT a administrativní metadata. Tento přístup umožňuje snadné další využití jako jsou replikace dat jiným institucím, harvesty pomocí OAI PMH, automatizované přenašení některých údajů mezi Systémem Kramerius, Registrem digitalizace cz a katalogy knihoven.

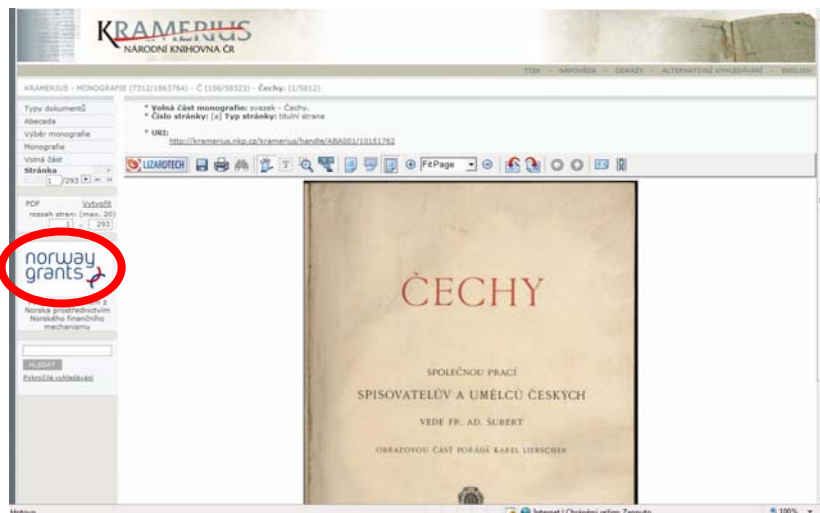
Pro úspěšné zvládnutí velkých objemů digitalizace bylo nutné změnit řadu zavedených postupů. Největší změnou bylo vytvoření speciální aplikace Registr digitalizace cz (RD.CZ), který slouží pro evidenci digitalizovaných dokumentů a současně se využívá i jako základní work flow pro proces digitalizace. Tento nástroj však bohužel nebyl k dispozici od začátku, takže bylo nutné práce zahájit s pomocí jednoduchých prostředků. Všechny

¹ <http://digit.nkp.cz/techstandards.html>

dokumenty vybrané pro reformátování v rámci tohoto projektu jsou evidovány pomocí výpůjčního modulu knihovního systému Aleph a současně jako výstup ze systému se vytváří excelový soubor, který se postupně kompletuje a zveřejňuje pod názvem „Norské fondy“ na webové stránce NK ČR pod záložkou Portály (je dostupný i prostřednictvím Infoportálu Systému Kramerius). Tento seznam slouží jako první informace o vybraných dokumentech pro ostatní digitalizační pracoviště. V této úvodní fázi se přenáší vybrané bibliografické údaje z knihovního systému do aplikace RD.CZ.

Vybrané dokumenty se na pracovišti digitalizace v CDH zkontrolují a roztrídí podle způsobu zpracování a vyhotoví se předávací seznamy pro jednotlivé dodavatele. Tyto činnosti již probíhají v prostředí aplikace RD.CZ. Po zhotovení mikrofilmu a skenů se obrazové soubory dále předávají ke zpracování firmě Elsystem Engineering. Zde se provádí ořez a narovnání, konverze do formátu DjVu, OCR a vytváří se popisná a administrativní metadata. V průběhu zpracování se pro jednotlivé mikrofilmy a obrazové soubory generují identifikátory, které umožňují částečně automatizovat uvedené procesy.

Po dodání hotových digitálních dokumentů se provádí jejich validace, archivace zdrojových dat na úložiště a import uživatelských digitálních kopií dokumentů do Systému Kramerius. Po dobu zpracování se uchovávají na centrálním datovém úložišti dodané obrazové soubory, které je možné odstranit až po provedení importů a validace zhotovených digitálních dokumentů. Dodavatelé jsou smluvně vázáni uchovávat zdrojová nebo zpracovaná data ještě po několik měsíců pro případ reklamace.



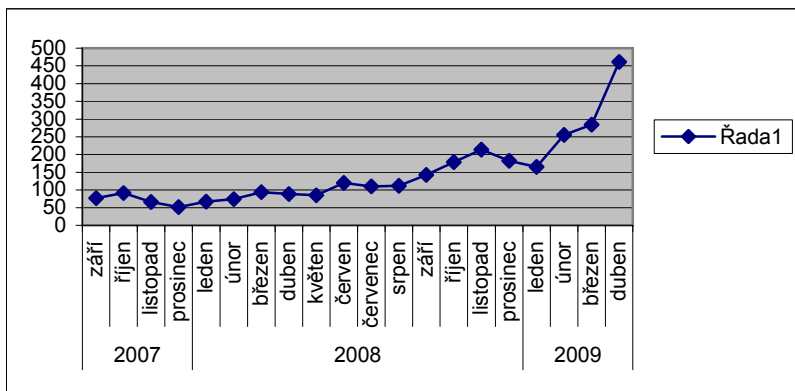
NK ČR se zavázala v rámci propagace projektu označovat všechny mikrofilmy i digitální dokumenty logem Norských fondů, které je umístěno na levé straně obrazovky. Logo doprovází dokument nejen na stránce, kde jsou zveřejněna metadata nebo titulní list, ale je umístěno u všech stran dokumentu.

Po dokončení importů jsou doplněny některé údaje ze Systému Kramerius do RD.CZ, což umožňuje přenést odkaz na digitalizovaný dokument do záznamu v bázi NKC. Uživatelé tak budou mít možnost se dostat z katalogu přímo na hlavní stránku digitalizovaného dokumentu.

Součástí projektu je zhotovování ochranných obalů pro nejvíce zničené dokumenty. Pro tyto knihy jsou vyráběny krabice z archivní lepenky na pracovišti péče o novodobé fondy. Plánujeme tímto způsobem ošetřit asi 4 000 svazků.

V současné době je digitalizováno nebo připraveno pro zpracování více než 16 000 svazků publikací 19. stol. Dokončení projektu předpokládáme v pol. r. 2010, do té doby budou probíhat úpravy dat a hlavně importy hotových digitálních dokumentů na úložiště a do Systému Kramerius.

Během realizace projektu bylo nutno řešit i řadu problémů. Některé příručky obsahující pravidla pro použití finančních prostředků byly vyhotoveny a distribuovány již během realizace, na začátku nebyla kapacita dodavatelských pracovišť dostatečná a musela být nahrazena produkcí oddělení digitalizace NK ČR, aby bylo možné dodat závazné objemy souborů ke zpracování firmě Elsys Engineering. Během celého řešení jsme čelili technickým problémům s provozem datového úložiště.



Postupně zveřejňování digitalizovaných dokumentů v Systému Kramerius se projevilo v nárůstu počtu návštěvníků. Zatímco v době zahájení projektu se průměrný počet návštěv pohyboval okolo sedmdesáti za den, v současné

době přesahuje 400. Vývoj průměrné denní návštěvnosti Systému Kramerius od září 2007 je znázorněn na předcházejícím grafu.

Na závěr je možné konstatovat, že realizace projektu obohatila již nyní prostředí internetu o velké množství dokumentů 19. stol., které slouží badatelům i široké veřejnosti. K oblibě těchto dokumentů, která se projevuje růstem návštěvnosti, přispívá i skutečnost, že se jedná o publikace, které až na výjimky již nepodléhají ochranné lhůtě autorského zákona a jsou volně k dispozici.