

KNIHOVNA  
AKADEMIE VĚD ČR  
NÁRODNÍ 3, 115 22 PRAHA 1

# INFORMACE

1/2009



Jarmila Burgetová

## ÚVODEM

První letošní číslo INFORMACÍ je monotematické. Rozhodli jsme se, že jím budeme podrobněji informovat o výsledcích a plánovaném rozvoji Digitalizačního centra Knihovny Akademie věd (DC). Knihovna je buduje od roku 2003, kdy byl, po hrozné povodni v srpnu předchozího roku, převod dokumentů do digitální formy zvolen jako vhodný způsob pro jejich ochranu a zpřístupnění. Pro odstranění následků povodně bylo možné využít kromě mimořádných vlastních (v tomto případě akademických) prostředků také příspěvek Evropské komise. O vzniku, současném vybavení DC, zpracování a následném zpřístupnění digitalizovaných dokumentů se dozvíte z rozsáhlého příspěvku vedoucí DC Mgr. Anny Ryšánkové a jejího kolegy Michala Pšeničky. Na škodu jistě také nebudou kontaktní informace o DC, o jeho adrese, telefonech, e-mailovém spojení.

Ředitel knihovny ing. Martin Lhoták přispěl do čísla statí o rozvoji digitalizace KNAV v nejbližších letech, kdy bude dále budována *Digitální knihovna Akademie věd ČR*, a o participaci knihovny na řadě projektů, jejichž řešení má v řadě případů nejen tuzemský, ale i mezinárodní dosah.

Při převodu dokumentů do digitální formy, jejich popisu a zpřístupnění KNAV úzce spolupracuje s Národní knihovnou ČR, a to především v programu *Kramerius*, věnovaném digitálnímu zpřístupnění dokumentů ohrožených degradací kyselého papíru, ale také v *Národním programu digitálního zpřístupnění vzácných dokumentů* – „*Memoriae Mundi Series Bohemica*“.

Číslo jsme doplnili výběrem z novinek ze světa v oblasti digitalizace a dvěma přílohami. První obsahuje ceníky prací, které zajišťuje a nabízí naše DC, v druhé otiskujeme soupis titulů časopisů, které jsou již dostupné v digitalizované formě.



Anna Ryšánková  
Michal Pšenička  
**DIGITALIZAČNÍ  
CENTRUM  
KNIHOVNY AV ČR**

Digitalizační centrum Knihovny AV ČR (DC KNAV) je v současné době nejlépe vybaveným a nejproduktivnějším digitalizačním pracovištěm v rámci českých knihoven. Pracoviště, které se nachází v areálu hlavního depozitáře KNAV v Jenštejně, funguje pátým rokem a za tu dobu zde bylo digitalizováno již více než 2 000 000 stran. Unikátní v rámci celé České republiky je možnost skenování dokumentů až do formátu A0 při rozlišení 600 DPI v barevné hloubce 24 bitů. Na pracovišti je možné zajistit široké spektrum služeb: digitalizaci, grafické úpravy, konverze do různých formátů, popis, OCR, uložení na různé typy médií a archivaci na datových páskách. Typy materiálů se mohou velmi lišit. Je možné digitalizovat jednotlivé listy na rychlém automatickém skeneru s velmi nízkými náklady, vázané dokumenty, včetně vzácných historických exemplářů i velkoformátové předlohy – mapy, obrazy apod. To vše s možností výběru rozlišení i barevné hloubky.

### **Vznik**

Významným impulsem pro vznik digitalizačního pracoviště v KNAV byly povodně v roce 2002. Pro záchranu obsahu poškozených knih a časopisů byla digitalizace jedním z možných řešení. Při tak velkém objemu zničených publikací nebylo možné všechny restaurovat. V té době přispěla i Evropská unie z finančních prostředků určených na pomoc při odstraňování škod po povodních.

Zkušenosti a inspirace pro vybavení a celkové řešení digitalizačního pracoviště byly čerpány v evropských knihovnách se zavedenými digitalizačními pracovišti. Nejprínosnější byly informace získané v zavedeném Digitalizačním centru Univerzitní knihovny v německém Göttingenu, které je zaměřené na zpracování obdobných typů dokumentů – především vědeckých periodik a monografií. V českém prostředí byla navázána úzká spolupráce s Národní knihovnou ČR, která měla v době vzniku DC KNAV dvě digitalizační pracoviště, jedno pro rukopisy a staré tisky a druhé pro záchranu tisků (zejména periodik) z 18.–19. století tištěných na kyselém papíře. Bylo rozhodnuto od počátku přijmout stejné standardy, jaké jsou doporučovány Národní knihovnou i obdobné workflow. Vznikla rámcová představa, jak by mělo digitalizační pracoviště KNAV vypadat po technické, personální a organizační stránce.

V průběhu roku 2003 probíhal výběr hardware a software, k pořízení došlo na konci tohoto roku. Na počátku roku 2004 byli přijati 4 pracovníci, kteří měli za úkol podílet

se na vývoji postupů digitalizace a zajistit kvalitní a efektivní produkci. První polovina tohoto roku bylo proto obdobím testování. Byly vyzkoušeny různé metody digitalizace a následného zpracování, postupně se vyprofilovaly nejvhodnější způsoby. Vzhledem k vývoji nových technologií, softwaru a získávání nových poznatků z prostředí digitalizace se průběžně zlepšují postupy směrem k vyšší kvalitě a především k větší efektivitě práce. V současnosti pracuje v digitalizačním centru 8 zaměstnanců. S pracovištěm také úzce spolupracuje Oddělení informačních technologií KNAV.

## **Vybavení HW, SW**

### **Hardware**

Kvalitní technické zabezpečení digitalizačního centra je důležitým faktorem pro efektivní fungování celého pracoviště. Na podzim roku 2003 byly osloveny firmy s poptávkou na dodávku skenerů a softwaru pro následné zpracování. Výběrové řízení na základě ustanovení § 49a zákona č. 199/1994 Sb., o zadávání veřejných zakázek vypsal tehdejší vedení KNAV. Předmětem výběru bylo zakoupení dvou černobílých skenerů na formáty do rozměru A2 a jednoho barevného skeneru na dokumenty do formátu A1.

Výběrová komise doporučila barevné zařízení DigiBook a černobílé skenery Zeutschel. Skener Digibook RGB 10000 od francouzské firmy IzS dodala firma AIP Safe s.r.o. (<http://www.aipsafe.cz>), jeho servis zajišťuje firma XEROX Czech republic s.r.o. (<http://www.xerox.cz/cs>). Černobílé skenery od německé firmy Zeutschel GmbH (<http://www.zeutschel.de>) dodala firma Ampaco ČR s.r.o. (<http://www.ampacocr.cz>), která zajišťuje i servis.

Barevný knižní skener DigiBook 10 000 RGB je skenerem pro formáty do A1 včetně při rozlišení 400 dpi. Maximální rozlišení skeneru je 800 dpi při 1×A4. Ke skeneru bylo dodáno speciální úchytné zařízení s přítlačným sklem pro knihy, které nelze otevřít více než na 90° nebo 120°. Dalším příslušenstvím byla přídatná plochá deska na skenování velkých map a plochých neknižních předloh. Později bylo pořízeno ještě přítlačné sklo na klasickou vyrovnávací kolébku pro formáty do A3 včetně.

Dva černobílé skenery Zeutschel oS 7 000 jsou určeny pro formáty předloh do A2 při rozlišení 400 dpi a 800 dpi při formátu 1×A4. Součástí skenerů bylo již od samého počátku přítlačné sklo pro celou plochu skeneru.

Čtvrtým skenovacím strojem je Panasonic KV – S1025C High Speed Scanner – pořízený v roce 2007. Jedná se o velkokapacitní skener pro předlohy do A4 včetně. Skenuje v barvě (24bit), šedé škále nebo černobíle. Je určen pro skenování jednotlivých listů, nejedná se tedy o knižní skener.

Pátým skenerem, pořízeným v roce 2008, je ScannTech 600i od firmy ProServ GmbH. Toto zařízení je unikátní svým rozlišením 600 dpi při skenovaném formátu A0 a přítlačným sklem po celé ploše A0. Samozřejmě součástí skeneru je jak pohyblivá

vyrovnávací kolébka pro knižní předlohy, tak i deska pro plošné předlohy a speciální úchytné zařízení pro knihy, které nelze otevřít více než na 90° nebo 120°.

### **Software**

Nejpodstatnější grafické úpravy obrázků probíhají ve specializovaném softwaru Book Restorer™, která je produktem francouzské firmy IzS. Book Restorer™ je programem pro úpravy, opravy, změny a zpracování digitalizovaných knih. Poskytuje výkonné nástroje pro dosažení vysoce kvalitních výsledků bez ohledu na to, zda s ním pracuje začátečník nebo expert. Významnou vlastností tohoto softwaru je fakt, že dokáže pracovat s velkým množstvím obrázků najednou. Nejvyužívanější funkcí je zpravidla geometrická korekce (rovnání křivých řádků), ořezy stránek a různé kombinace filtrů pro zvýraznění písma a odstranění nečistot.

Pro tvorbu metadat a pro OCR je v Digitalizačním centru využíván software Sirius, který vyvinula firma Elsys Engineering ve Vyškově (<http://www.ee.cz>). Je určen pro grafické zpracování, tvorbu metadat a archivaci digitalizovaných dokumentů. Může být také samotným softwarem pro komunikaci skener – počítač (pro samotné skenování). Systém pracuje na principu klient – server, proto může být instalován jak v malých kancelářích, tak i ve velkých provozech o několika desítkách počítačů. Systém využívá databázi Microsoft SQL. V databázi jsou uloženy také veškeré indexy, které uživatel během zpracování dokumentům přiřadí a systémové informace o dokumentech. Kromě databáze je zde ještě samotné úložiště dokumentů, které jsou během zpracování ukládány.

### **Zpracování**

Typy obrázku resp. materiálu + návaznost na dotační programy.

Zpracování materiálu v Digitalizačním centru záleží na jeho povaze. Typy předloh lze rozdělit zhruba do těchto skupin:

- Periodika a monografie
- Staré tisky a rukopisy
- Mapy a ostatní velkoplošné předlohy

Periodika a monografie jsou digitalizovány jak z fondu Knihovny AVČR, tak i pro zájemce z jiných ústavů, knihoven nebo soukromých subjektů. Digitalizační centrum KNAV zpracovává také mnoho dokumentů v rámci národního programu VISK 7 určeného pro mikrofilmování a digitální zpřístupnění dokumentů ohrožených degradací kyselého papíru.

Standardní skenování těchto předloh probíhá v kvalitě 300 dpi, 4 bitové barevné hloubce a do grafického formátu tiff. Na vyžádání objednatele lze skenovat i v jiném rozlišení až do výše 800 dpi při formátu předlohy A4, v barevné hloubce 1 nebo 24 bitů. Po nasnímání předloh následuje grafické zpracování v softwaru Book Restorer™. Na všechny obrázky je použita funkce geometrická korekce, která srovná křivé řádky

a ořez, kdy je obrázek oříznut stejnoměrně kolem textu. Dále jsou pak používány různé grafické funkce pro vylepšení písma a pozadí stránek. Když jsou předlohy graficky upraveny, vytvoří se k tiskům metadata v programu Sirius. Propojí se zde jednotlivé obrázky s čísly stránek, vytvoří se popisy k výtiskům a ročníkům a pomocí OCR se nechá vyčíst text z jednotlivých stránek. Výstupem ze Siria je pak adresář s jednotlivými obrázky a textovými dokumenty (jeden textový dokument ke každému obrázku, vyčtený text slouží k pozdějšímu fulltextovému vyhledávání v systému Kramerius) a soubor XML s popisnými metadaty. Poté jsou obrázky převedeny do formátu JPG pomocí speciálně pro tento účel vytvořeným skriptem. Po této úpravě již nic nebrání importu do systému Kramerius, který publikace zpřístupní veřejnosti.

Rukopisům a starým tiskům je věnována při digitalizaci zvláštní péče. Knihy, které nelze bez poškození otevřít na 180 stupňů, jsou skenovány na speciálních držácích, které jsou součástí příslušenství ke skenovacím zařízením. Knihy lze tak polohovat až do otevření pouze na 90 stupňů. Digitalizace je tak šetrnější. Obvykle jsou knihy skenovány ve kvalitě 300 dpi a barevné hloubce 24 bitů. Vzácnější kusy skenujeme ve vyšší kvalitě 600 dpi. Po dohodě je možná i kvalita vyšší. Po skenování jsou vzácné tisky a rukopisy vždy uloženy do trezoru, který je přímo na pracovišti. Ke každé knize lze podle požadavků objednatele vytvořit ICC profil spolu s nasnímanou kalibrační tabulkou pro pozdější kontrolu kvality a možnost věrné reprodukce v tištěné i elektronické formě. Metadata k tomuto druhu dokumentů jsou vytvářena zejména v programu MTool. Tento software je doporučen provozovatelem Manuscriptoria, které zpřístupňuje rukopisy a staré tisky digitalizované v rámci národního programu VISK 6 pro digitální zpřístupnění vzácných dokumentů Memoriae Mundi Series Bohemica (dokumenty dochované na teritoriu České republiky).

Mapy a další velkoformátové předlohy jsou skenovány standardně na 300 dpi nebo 600 dpi a v barevné hloubce 24 bitů. Nová akvizice z roku 2008 – skener ScannTech 600i – umožňuje digitalizaci v rozlišení 600 dpi při maximální velikosti předlohy A0. Větší předlohy až do 4xA0 je možné skenovat po částech a následně je spojit pomocí speciálního softwaru.

## **Zpřístupnění**

### **Časopisy a monografie**

Zpřístupnění časopisů z produkce AV ČR bylo od počátku prioritou. Jako nejvhodnější řešení byla zvolena úzká spolupráce s Národní knihovnou ČR.

Ta se v roce 2003 zasloužila o realizaci systému Kramerius, který vznikl jako open source software a je k dispozici pod licencí GNU GPL. Jednalo se v rámci ČR o první krok k vytvoření kvalitní digitální knihovny, kterou mohou využít ke zpřístupnění svých dat i další instituce bez nutnosti vysokých výdajů za systém a jeho vývoj. Knihovna Akademie věd začala v roce 2003 spolupracovat s Národní knihovnou při definici vlastností systému Kramerius. Od roku 2004 se KNAV začala díky svým výzkumným projektům podílet i finančně na jeho vývoji. Systém, který je určen pro

zpřístupnění periodik a monografií je v současné době provozován ve většině velkých českých knihoven, pronikl ovšem i do menších institucí a dokonce i na Slovensko. Řádově je díky tomuto systému zpřístupněno již několik milionů stran. KNAV je připravena případně pomoci s implementací i dalším zájemcům.

Vývojový tým, který kromě KNAV tvoří Národní knihovna ČR, Moravská zemská knihovna v Brně a firma Qbizm technologies, a.s., začal intenzivně pracovat na vývoji nové verze systému založené na open source digitálním repozitáři Fedora, vyvíjeném v USA. Jedná se o propracovaný stroj umožňující správcům a programátorům využívat různé služby a volání. Repozitář může obsahovat v podstatě jakýkoli typ digitálních objektů. Tento přístup poskytující velkou flexibilitu však neumožňuje okamžité nasazení systému.

Prvním krokem realizovaným jako „Proof of Concept“ bylo vytvoření konverzního nástroje, který umožňuje převedení stávající datové struktury do nové, založené na FOXML, kterou používá digitální repozitář Fedora. Dále byl také spuštěn testovací server se systémem Fedora a bylo vytvořeno jednoduché webové rozhraní, aby existovala možnost kontroly importovaných dat.

### **Rukopisy a staré tisky**

Pro zpřístupnění těchto vzácných dokumentů je především pro výstupy spolufinancované z programu VISK 6 určeno Manuscriptorium provozované komerční firmou. Pracuje se i na zpřístupnění těchto dokumentů v systému Kramerius. V digitalizačním centru je možné opatřit metadata vhodná pro oba tyto systémy.

### **Mapy a velkoformátové dokumenty**

Ke zpřístupnění map a případně i jiných větších dokumentů je možné využít například software Zoomify, který umožňuje pohodlné prohlížení map v různých přiblíženích pomocí předem vygenerovaných dlaždic. Toto řešení je využíváno (u map digitalizovaných v Jenštejně) ve Vědecké knihovně v Olomouci a v Jihočeské vědecké knihovně v Českých Budějovicích.

### **Zálohování a dlouhodobá archivace**

Již krátce po vzniku digitalizačního centra jsme se začali potýkat s nedostatkem místa na diskových polích pro ukládání naskenovaných a zpracovaných materiálů. Bylo jasné, že s narůstajícím objemem dat se bude tento problém prohlubovat. Začali jsme tedy prověřovat možnosti dlouhodobé archivace dat. Řešení jsme našli v páskové mechanice HP StorageWork Ultrium 460 a datových páskách Ultrium II. Jedná se o pásky s kapacitou 200/400 GB s archivační životností až 30 let. Na tyto pásky jsou zálohována data v různé fázi zpracování. Prvním druhem dat jsou naskenované materiály ve formátu tiff, na nichž je z důvodu úspory místa provedena LZW komprese. Jedná se o bezeztrátovou kompresi dat. Druhým typem zálohovaných dat jsou materiály zpracované, na kterých již byly provedeny grafické úpravy a byla k nim v programu Sirius pořízena metadata. Každá páska je pořízena ve dvou identických kopiích. Jedna kopie se uchovává v Digitalizačním centru Jenštejn a je používána



v běžném provozu. Druhá je uložena v Knihovně Akademie věd na Národní třídě v Praze a slouží jako archivační kopie pro případ poškození pásky v Digitalizačním centru.

Hlavním problémem je velká těžkopádnost stávajícího řešení a to také vzhledem k očekávanému velkému nárůstu objemu dat. Svou roli zde bude hrát plánované nasazení dalších skenerů, zejména barevné skenování a využití robotického skeneru. Nyní existuje narůstající sada pásek, které však nejsou organizovány do sofistikovaného systému, jenž by zabezpečoval nějakou konsistentní formu správy metadat, která popisují co je kde zazálohováno. Záleží na svědomitosti obsluhy, jak bude udržovat popis a zařazení jednotlivých pásek, aby bylo možné v případě potřeby data obnovit. S tím také souvisí poměrně velká časová náročnost s prováděním záloh, která je kladena na obsluhu.

Posláním KNAV je mimo jiné také garance dlouhodobého uchování spravovaných dokumentů, včetně těch v elektronické formě. Případnou ztrátou dat by mohlo dojít k velkým finančním škodám, vzhledem k tomu, že náklady na digitalizaci a další zpracování dat jsou značné. Velmi důležitým aspektem je záruka zachování kulturního dědictví. Vzhledem k tomu, že řada dokumentů je, a čím dál více bude, dostupná pouze v elektronické formě, řeší v současné době knihovny po celém světě problematiku důvěryhodných uložišť pro dlouhodobé uchování dat.

KNAV chce řešit otázku ukládání digitalizovaných dat, vzhledem k jejich stále se zvětšujícímu objemu, komplexním systémem. Ten jednak postihne samotný problém s uložením dat na diskových polích a dále jejich zálohováním a archivací a v neposlední řadě i s jejich automatickou replikací do geograficky vzdálené destinace. Bude zajištěna existence 4 elektronických kopií na 3 různých geografických místech.

Vybrané řešení sestává jednak z diskových polí a dále z páskových knihoven, což jsou roboti, kteří dokáží automaticky manipulovat se sadou zálohovacích pásek bez zásahu lidské obsluhy. Tyto komponenty jsou organizovány do tzv. virtuálního úložného prostoru prostřednictvím další mezivrstvy, která již není hardwarové povahy, nýbrž je to programové vybavení. Výhodou je možnost průběžného zálohování dat z diskových polí na páskové knihovny a dále tzv. migrace a klonování dat. Systém umožňuje ekonomické využití dostupného místa tím, že data, ke kterým je častěji přistupováno jsou na rychlých (tedy dražších discích) a méně používaná data se postupně přesouvají na pomalejší disky, případně na pásky.

Migrace dat v úložišti probíhá transparentně na pozadí. Na páskových knihovnách lze vytvářet archivní zálohy, je možné odvézt např. do jiné geografické destinace a uložit na bezpečném místě. Systém umožňuje i replikaci dat po síti na jiné vzdálené úložiště.

Navržené datové úložiště by mělo řešit v následujících minimálně pěti letech problém s ukládáním digitalizovaného i born-digital obsahu v AV ČR, s možností dalšího postupného rozšiřování i v dalších letech.

Tak jak je celkově úložiště koncipováno, předpokládá existenci dvou samostatných identických jednotek umístěných fyzicky v různých lokalitách. Primární bude umístěna přímo v Jenštejně, kde vzniká největší objem dat, sekundární – záložní – bude umístěna v Praze na Národní 3. Mezi těmito jednotkami bude probíhat transparentně asynchronní replikace, takže obraz stavu primární jednotky bude uložen na sekundární jednotce.

Celková kapacita zvoleného řešení primárního datového úložiště představuje cca 40TB úložného prostoru na diskových polích a ve výchozí konfiguraci 40TB úložného prostoru v páskových knihovnách, které budou sloužit pro zálohování a migraci dat z diskových polí.

V souvislosti s datovým úložištěm, které bude uvedeno do provozu zřejmě ve 3. čtvrtletí roku 2009, byla v Jenštejně vybudována z části depozitáře serverovna s novými síťovými rozvody a klimatizací. Propojení nové serverovny s původní, kam jsou svedena strukturovanou kabeláží připojení počítačových stanic pracovníků digitalizačního centra, bylo realizováno optickým kabelem. Dále byly také zajištěny nové rozvody strukturované kabeláže v kancelářích digitalizačního centra a posíleno mikrovlnné spojení zajišťující připojení k Internetu přes CESNET.

### **Zakázky a reference**

Za dobu fungování Digitalizačního centra jsme zpracovali řadu velmi zajímavých zakázek, ať už od jednotlivců nebo institucí. Mezi největší zakázky lze zařadit například digitalizaci cca 140 kusů rukopisů a starých tisků pro Moravskou zemskou knihovnu v Brně za účelem vložení digitálních obrazů knih do digitální knihovny Manuscriptorium. Dalším vzácným materiálem byl například Portolánový atlas a staré tisky digitalizované pro Vědeckou knihovnu v Olomouci nebo několik stovek historických map skenovaných pro Jihočeskou vědeckou knihovnu v Českých Budějovicích. Neméně zajímavou předlohou pro digitalizaci byl zajisté také pětidílný Jungmannův slovník z Ústavu pro jazyk český AV ČR, v. v. i., Říšské zákoníky pro Právnickou fakultu UK nebo několik stovek kusů kvalifikačních prací pro Fakultu sociálních věd UK.

Již několik let spolupracujeme například s Institutem Bohuslava Martinů, pro který digitalizujeme rukopisné hudebniny skladatele Bohuslava Martinů. Divadelní ústav také již třetím rokem využívá našich služeb k převedení do digitální podoby divadelních scénářů, které jsou obzvláště poškozeným a náročným materiálem pro zpracování. V budoucnu se rýsují další významné spolupráce, a to především s Moravskou zemskou knihovnou v Brně a Vědeckou knihovnou v Olomouci. Doufáme, že si nás najdou i jiné instituce nebo jednotlivci, se kterými budeme rádi spolupracovat na digitalizaci jejich významných nebo vzácných materiálů.

Kromě dalšího vývoje open source systému Kramerius pro zpřístupnění digitalizovaných dokumentů je cílem vývoj systému pro evidenci digitalizace a digitalizačního workflow v rámci celé České republiky. Do budoucna má tento

důležitý projekt, jehož cílem bude mimo jiné zabraňovat duplicitám při digitalizaci, slibný potenciál v rámci mezinárodní spolupráce. Výstup nazvaný „Registr digitalizace CZ“ bude provozovat Národní knihovna ČR, která je partnerem projektu. Pracoviště, která digitalizaci provádějí, budou informace o svých digitalizovaných materiálech evidovat na centrálním serveru.

Mezi projekty spíše produkčního charakteru lze považovat jednoleté projekty financované z programu VISK MK ČR. Konkrétně se jedná o podprogramy VISK 6 a VISK 7.

Program VISK 6 – „Národní program digitálního zpřístupnění vzácných dokumentů Memoriae Mundi Series Bohemica“ je zaměřen na elektronické zpřístupnění a záchranu dokumentů dochovaných a uložených na území České republiky. Cílem programu je pomocí digitalizace uchovat a zpřístupnit vzácné dokumenty knihoven a sbírek, které jsou důležitou součástí našeho kulturního dědictví. Tento program navazuje na program UNESCO „Paměť světa“. Žadatelé o dotace mohou být jakékoliv tuzemské instituce vlastníci vzácné fondy rukopisů, starých tisků, map apod. Dotace jsou udělovány především s ohledem na míru ohrožení a možnosti co nejširšího využití. Podrobnější podmínky pro udělování dotací v rámci toho projektu lze nalézt na adrese <http://visk.nkp.cz/VISK6.htm>.

Cílem programu VISK 7 – „Národní program mikrofilmování a digitálního zpřístupňování dokumentů ohrožených degradací kyselého papíru – Kramerius“ je převedení nejvíce ohrožených dokumentů na mikrofilm, který garantuje jejich trvalé uchování. V druhé fázi je to vytvoření digitální kopie pro elektronické zpřístupnění dokumentů ať už na internetu, lokálních sítí či pomocí médií. Výstupem je tedy mikrofilm pro archivování a digitální kopie dokumentu pro zpřístupnění. O dotace v rámci tohoto programu mohou žádat knihovny, které mají vlastní mikrografické a digitalizační pracoviště za účelem jeho provozu a rozvoje, dále pak instituce vlastníci pracoviště reformátování (dotace jsou udělovány pouze na činnosti související např. s rozmnožováním mikrofilmů apod.) a instituce spravující knižní sbírky obsahující dokumenty ohrožené degradací papíru (dotace jsou poskytovány především na záchranu buď bohemikálních periodik nebo monografií nebo vzácných nebohemikálních periodik a monografií). Z dotace je možno hradit až 70% nákladů. Podrobnější informace o poskytování dotací v rámci tohoto projektu lze nalézt na adrese <http://visk.nkp.cz/VISK7.htm>.

Na pracovišti KNAV vznikají dokumenty v rámci obou těchto dotačních programů již několik let, a to jak při zpracování vlastních materiálů, tak i formou služeb pro ostatní knihovny a paměťové instituce.

# KONTAKTNÍ INFORMACE

Digitalizační centrum Knihovny AV ČR sídlí v Jenštejně u Prahy.

Adresa:

Jenštejn 26  
250 73 Praha–Východ

Tel.: 286 851 955, e-mail: [digitalizace@knav.cz](mailto:digitalizace@knav.cz)

Vedoucí DC: Mgr. Anna Ryšánková  
Tel.: 286 851 955, e-mail: [rysankova@knav.cz](mailto:rysankova@knav.cz)



Martin Lhoták

## **PŘEDPOKLÁDANÝ ROZVOJ DIGITALIZAČNÍHO CENTRA**

V souladu s činností v předchozích letech bude i nadále cílem zajistit špičkové zpracování kompatibilní s požadovanými standardy. Za mandatorní považujeme kompatibilitu s Národní knihovnou ČR. Na mezinárodní úrovni se chceme přizpůsobit standardům v EU projektech Europeana, DRIVER a Enrich.

Nadále budeme nabízet naše kapacity dalším zájemcům – zejména ústavům AV ČR, českým knihovnám a paměťovým institucím. Tato spolupráce nám, díky zkušenostem s různými typy materiálů, pomáhá zlepšovat pracovní postupy a finanční přínos umožňuje technický rozvoj pracoviště.

Po pořízení unikátního barevného skeneru ProServ pro formáty (až do) A0, bychom chtěli realizovat ještě v roce 2009 další významnou akvizici – robotický skener Treventus. Ten umožňuje velmi citlivou metodou barevné skenování vázaných předloh reálnou rychlostí kolem 1000 stran za hodinu. To povede k několikanásobně vyšší produktivitě a tím i výraznému snížení nákladů na jednu stranu.

Budeme se snažit zahájit intenzivnější spolupráci s dalšími ústavu AV ČR, které mají ve svých fondech množství hodnotných materiálů, ať už jde o vzácné tisky, historické mapy, archiválie, ale i novodobější publikace, které si zaslouží zpřístupnění i dlouhodobou archivaci a ochranu v digitální formě. Předpokládáme i digitalizaci části historické produkce nakladatelství Academia.

Součástí Digitalizačního centra se v letošním roce stane i centrální datové úložiště, které by mělo být zárukou pro bezpečné uchování elektronických materiálů. Automaticky je již v současné době archivována veškerá digitalizovaná produkce AV ČR a jejích předchůdkyň, která přejde na zmíněné úložiště.

Nejvýznamnějším výstupem Digitalizačního centra bude nadále Digitální knihovna AV ČR, která obsahuje archiv časopisů publikovaných v AV ČR v minulých více než 100 letech. V současné době je tímto způsobem on-line zpřístupněn více než 1 milion stran (<http://kramerius.lib.cas.cz/>). Celkově bude počet stran ještě cca o polovinu vyšší. Připravuje se změna workflow, která by měla zajistit členění časopisů na jednotlivé články. Nový postup bude vyžadovat značnou časovou náročnost. Z hlediska přístupu k vědeckým informacím je však členění na články nutné a přinese další žádoucí efekty při diseminaci vědeckých výstupů AV ČR.

Martin Lhoták  
**PROJEKTY**

Digitalizační centrum KNAV se podílí na řešení čtyř výzkumných projektů:

**Digitální knihovna Akademie věd ČR**

projekt: 1ET200830526 programu AV ČR „Informační společnost“ (2005–2009)

Účelem projektu je zkoumat, vyvinout a aplikovat metody, postupy a nástroje pro vytvoření infrastruktury a realizaci Digitální knihovny AV ČR. Cílem je retrospektivní zpřístupnění produkce vědeckých prací vzniklých v působnosti Akademie věd ČR a jejich předchůdkyň a další relevantní odbornou literaturu v plných textech a s fulltextovým prohledáváním. Podstatným výstupem projektu je systém pro zpřístupnění Kramerius, ve kterém je v současné době zpřístupněn 1 milion stran časopisů vydaných v Akademii věd.

<http://kramerius.lib.cas.cz/>

**Česká digitální matematická knihovna**

projekt: 1ET200190513 programu AV ČR „Informační společnost“ (2005–2009)

Cílem projektu je zkoumat, vyvinout a aplikovat postupy, metody a nástroje, které umožní vytvořit infrastrukturu a podmínky pro realizaci České matematické digitální knihovny zahrnující relevantní část odborné matematické literatury vydané v českých zemích, a pro její začlenění do světové matematické digitální knihovny (WDML). Výstupem projektu je digitální knihovna volně dostupných matematických článků a monografií DML-CZ: Czech Digital Mathematics Library na adrese <http://www.dml.cz/>.

**Optimalizace nástrojů pro digitalizaci tištěných dokumentů ohrožených degradací kyselého papíru**

projekt DCo6Po2OUK002 programu MK ČR „Zpřístupnění a ochrana kulturních, uměleckých a vědeckých zdrojů“ (2006–2010)

Cíle projektu:

1. Vybudování znalostní báze založené na slovnících a frazeologických bázích pro časové jazykové vrstvy z období let 1800 až 1989. Využití znalostní báze pro optimalizaci OCR u programového nástroje pro tvorbu dokumentů Sírius, používaného pro digitalizaci dokumentů v rámci Národního programu Kramerius.
2. Zvýšení produktivity systému Sírius a rozšíření popisu digitálních objektů zejména v oblasti jejich vnitřních částí (článků v časopisech a novinách, kapitol v monografiích)

se zaměřením na automatizaci pořizování těchto metadat. Vytvoření nástroje pro převod interního informačního aparátu dokumentů (obsahy, rejstříky, poznámky apod.) do textové podoby.

3. Vytvoření editačního a konverzního nástroje na tvorbu metadat pro digitální objekty rozšířením funkcionality externího modulu systému Kramerius.

### **Evidence digitalizovaných dokumentů, sledování procesu zpracování a vývoj systému pro zpřístupnění**

projekt DCo8Po2OUKoo8 programu MK ČR „Zpřístupnění a ochrana kulturních, uměleckých a vědeckých zdrojů“ (2008–2011).

Kromě dalšího vývoje open source systému Kramerius pro zpřístupnění digitalizovaných dokumentů je cílem vývoj systému pro evidenci digitalizace a digitalizačního workflow v rámci celé České republiky. Do budoucna má tento důležitý projekt, který má mimo jiné za cíl zabránit duplicitám při digitalizaci, slibný potenciál v rámci mezinárodní spolupráce. Výstup nazvaný „Registr digitalizace CZ“ bude provozovat Národní knihovna ČR, která je partnerem projektu. Pracoviště, která digitalizaci provádějí, budou informace o svých digitalizovaných materiálech evidovat na centrálním serveru.

Mezi projekty spíše *produkčního* charakteru lze považovat jednoleté projekty financované z programu VISK MK ČR. Konkrétně se jedná o podprogramy VISK 6 a VISK 7.

Program VISK 6

#### **Národní program digitálního zpřístupnění vzácných dokumentů Memoriae Mundi Series Bohemica**

je zaměřen na elektronické zpřístupnění a záchranu dokumentů dochovaných a uložených na území České republiky. Cílem programu je pomocí digitalizace uchovat a zpřístupnit vzácné dokumenty knihoven a sbírek, které jsou důležitou součástí našeho kulturního dědictví. Tento program navazuje na program UNESCO „Paměť světa“. Žadateli o dotace mohou být jakékoliv tuzemské instituce vlastníci vzácné fondy rukopisů, starých tisků, map apod. Dotace jsou udělovány především s ohledem na míru ohrožení a možnosti co nejširšího využití. Podrobnější podmínky pro udělování dotací v rámci toho projektu lze nalézt na adrese <http://visk.nkp.cz/VISK6.htm>.

Cílem programu VISK 7

#### **Národní program mikrofilmování a digitálního zpřístupňování dokumentů ohrožených degradací kyselého papíru – Kramerius**

je převedení nejvíce ohrožených dokumentů na mikrofilm, který garantuje jejich trvalé uchování. V druhé fázi je to vytvoření digitální kopie pro elektronické zpřístupnění dokumentů ať už na internetu, lokálních sítí či pomocí médií. Výstupem je tedy mikrofilm pro archivování a digitální kopie dokumentu pro zpřístupnění. O dotace

v rámci tohoto programu mohou žádat knihovny, které mají vlastní mikrografické a digitalizační pracoviště za účelem jeho provozu a rozvoje, dále pak instituce vlastní pracoviště reformátování (dotace jsou udělovány pouze na činnosti související např. s rozmnožováním mikrofilmů apod.) a instituce spravující knižní sbírky obsahující dokumenty ohrožené degradací papíru (dotace jsou poskytovány především na záchranu buď bohemikálních periodik nebo monografií nebo vzácných nebohemikálních periodik a monografií. Z dotace je možno hradit až 70 % nákladů. Podrobnější informace o poskytování dotací v rámci tohoto projektu lze nalézt na adrese <http://visk.nkp.cz/VISK7.htm>.

Na pracovišti KNAV vznikají dokumenty v rámci obou těchto dotačních programů již několik let, a to jak při zpracování vlastních materiálů, tak i formou služeb pro ostatní knihovny a paměťové instituce.



## NOVINKY Z DOMOVA I ZE SVĚTA

**Mezinárodní konference** Digitální zpřístupnění kulturního dědictví proběhla v úterý dne 26. května 2009 na Vysoké škole ekonomické v Praze. Zajistila ji společnost Albertina icome Praha s.r.o. pro Ministerstvo kultury ČR v rámci českého předsednictví EU. V programu, který zahájili ministr kultury ČR pan **Václav Riedlbauch**, rektor Vysoké školy ekonomické v Praze prof. **Richard Hindls** a ředitel Direktorátu E Digitální obsah a kognitivní systémy Evropské komise pan **Horst Forster**, zaznělo celkem 9 odborných příspěvků. Program konference sledovalo cca 250 účastníků z řady evropských zemí. Prezentace přednášejících jsou již dostupné na adrese [www.diginforum.cz/prezentace/](http://www.diginforum.cz/prezentace/).

**Digitalizační strategii Bavorské státní knihovny v Mnichově** se věnuje obsáhlá stať Klause Ceynowy, náměstka generálního ředitele této knihovny „ Mass Digitization for Research and Study: the digitization strategy of the Bavarian State Library, otištěná v časopise IFLA Journal.<sup>1</sup> Kromě tištěné formy je článek dostupný rovněž na internetu.<sup>2</sup> Bavorská státní knihovna patří k největším evropským vědeckým knihovnám se vzácnými knižními fondy. Jejím prioritním strategickým cílem je toto bohatství co nejdříve digitalizovat a představit veřejnosti. Stať popisuje projekty i konkrétní postupy, které umožní zpřístupnit v krátké době online více než 1,2 mil. knih.

Mezinárodní časopis **World Digital Libraries** hodlá vydávat indický Ústav pro energii a zdroje v New Delhi. Tento recenzovaný časopis, který bude vycházet 2x ročně, si klade prestižní cíle. V Ediční radě jsou zastoupeni kromě Indie odborníci z celého světa (Portugalsko, Austrálie, UK, USA, Singapur, Rakousko, Belgie, Německo, Norsko, Japonsko, Finsko, Nový Zéland). První číslo vychází v červnu, druhé vyjde v prosinci 2009. Obsah prvního čísla lze najít na webu [http://bookstore.teriin.org/journal\\_inside.php?material-id=477&qty=1](http://bookstore.teriin.org/journal_inside.php?material-id=477&qty=1).

**Mezinárodní konference „Towards a Digital Mathematic Library“ DML 2009** se bude konat ve dnech 8. a 9. července 2009 v kanadském Ontariu. Spolupořadatelem konference je Fakulta informatiky Masarykovy univerzity v Brně; na jejím webu najdete veškeré informace.

Viz: <http://www.fi.muni.cz/~sojka/dml-2009.html>

Po úspěchu DLM 2008 budou ve workshopech formulovány strategie a hlavní cíle globální matematické digitální knihovny, shrnuty dosavadní úspěchy a posouzeny používané technologie a související projekty.

---

<sup>1</sup> IFLA Journal, Vol. 35, No. 1, 17–24 (2009)

<sup>2</sup> <http://ifl.sagepub.com>

## PŘÍLOHA Č. 1

# CENÍK DIGITALIZAČNÍHO CENTRA

Po plánované akvizici robotického skeneru v 2. polovině roku 2009 lze očekávat velmi výrazné snížení ceny u formátů do A4.

## CENÍK PRO PRACOVIŠTĚ AKADEMIE VĚD ČR (bez DPH)

Ceny již od 1,80 Kč

### Ceník pro svázané předlohy

300 dpi	barva	256 šed'	16 šed'	BW
A1	15,00 Kč	12,50 Kč	—	11,00 Kč
A2	10,00 Kč	8,00 Kč	7,00 Kč	7,00 Kč
A3	7,00 Kč	4,80 Kč	4,00 Kč	4,00 Kč
A4–A5	4,00 Kč	2,20 Kč	1,80 Kč	1,80 Kč

600 dpi	barva	256 šed'	16 šed'	BW
A3	12,00 Kč	8,00 Kč	7,00 Kč	7,00 Kč
A4–A5	8,80 Kč	4,70 Kč	3,80 Kč	3,80 Kč

### Ceník pro nsvázané předlohy

300 dpi	barva	256 šed'	16 šed'	BW
A0	38,00 Kč	34,00 Kč	—	30,00 Kč
A1	18,00 Kč	15,00 Kč	—	14,00 Kč
A2	11,00 Kč	9,70 Kč	8,00 Kč	8,00 Kč
A3	7,50 Kč	6,00 Kč	4,70 Kč	4,70 Kč
A4–A5	4,00 Kč	2,80 Kč	2,50 Kč	2,50 Kč

600 dpi	barva	256 šed'	16 šed'	BW
A0	75,00 Kč	65,00 Kč	—	50,00 Kč
A1	38,00 Kč	30,00 Kč	—	28,00 Kč
A2	22,00 Kč	18,00 Kč	—	17,00 Kč
A3	13,50 Kč	11,00 Kč	8,50 Kč	8,50 Kč
A4–A5	10,30 Kč	6,00 Kč	4,60 Kč	4,60 Kč

## CENÍK PRO OSTATNÍ INSTITUCE A JEDNOTLIVCE (bez DPH)

Sleva až 40 % (viz tabulka na konci ceníku)

### Ceník pro svázané předlohy

Minimální cena zakázky je 1 000 Kč.

#### 300 dpi

Ceny již od 1,90 Kč

	barva	256 šed'	16 šed'	BW
A1	70,00 Kč	45,00 Kč	—	30,00 Kč
A2	45,00 Kč	30,00 Kč	25,00 Kč	25,00 Kč
A3	20,00 Kč	15,00 Kč	12,00 Kč	12,00 Kč
A4–A5	8,00 Kč	4,20 Kč	3,20 Kč	3,20 Kč

#### 600 dpi

Ceny již od 4,20 Kč

	barva	256 šed'	16 šed'	BW
A1	150,00 Kč	110,00 Kč	—	60,00 Kč
A2	80,00 Kč	65,00 Kč	—	40,00 Kč
A3	45,00 Kč	35,00 Kč	25,00 Kč	18,00 Kč
A4–A5	20,00 Kč	15,00 Kč	10,00 Kč	7,00 Kč

### Ceník pro nsvázané předlohy

#### 300 dpi

Ceny již od 1,60 Kč

	barva	256 šed'	16 šed'	BW
A0	130,00 Kč	100,00 Kč	—	85,00 Kč
A1	80,00 Kč	50,00 Kč	—	40,00 Kč
A2	55,00 Kč	40,00 Kč	30,00 Kč	30,00 Kč
A3	25,00 Kč	18,00 Kč	14,00 Kč	14,00 Kč
A4–A5	10,00 Kč	5,20 Kč	4,30 Kč	4,30 Kč

#### 600 dpi

Ceny již od 7,20 Kč

	barva	256 šed'	16 šed'	BW
A0	270,00 Kč	220,00 Kč	—	120,00 Kč
A1	170,00 Kč	140,00 Kč	—	78,00 Kč
A2	95,00 Kč	80,00 Kč	—	42,00 Kč
A3	55,00 Kč	42,00 Kč	30,00 Kč	28,00 Kč
A4–A5	25,00 Kč	20,00 Kč	15,00 Kč	13,00 Kč

## CENÍK PRO NESVÁZANÉ PŘEDLOHY – RYCHLOSKENER

Minimální cena zakázky je 1 000 Kč.

Ceny již od 0,42 Kč.

### 300 dpi

	Barva	256 šed'	BW
A4–A5	1,20 Kč	0,90 Kč	0,70 Kč

### 600 dpi

	Barva	256 šed'	BW
A4–A5	2,00 Kč	1,50 Kč	1,00 Kč

V případě špatné fyzické kvality dokumentu může být cena navýšena.

## SLEVY

	A1	A2	A3	A4–A5	Sleva z celkové částky zakázky
množství	> 50	> 80	> 120	> 300	10%
	> 150	> 180	> 220	> 600	20%
	> 250	> 280	> 320	> 1 000	30%
				> 5 000	40%

V případě špatné fyzické kvality dokumentu může být cena navýšena.

## PŘÍLOHA Č. 2

# SOUPIS DIGITALIZOVANÝCH ČASOPISŮ

### Hotové

Acta Comeniana  
Acta Entomologica Bohemoslovaca  
Acta Mathematica Et Informatica Universitatis Ostraviensis  
Acta Montana  
Acta Montana AB  
Acta Montana Serie A  
Acta Montana Serie B  
Acta Onomastica  
Acta Technika CSAV  
Aplikace Matematiky  
Archeologické Rozhledy  
Archiv J. A. Komenského  
Archiv J. A. Komenského  
Archiv Orientální  
Biologia Plantarum  
Biologické Listy  
Czechoslovak Journal of Physics  
Czechoslovak Mathematical Journal  
Časopis pro moderní filologii  
Časopis Pro Pěstování Matematiky a Fysiky  
Časopis Pro Pěstování Matematiky  
Československá Biologie  
Časopis pro pěstování matematiky  
Česká literatura  
Československá biologie  
Československá psychologie  
Československá rusistika  
Československý časopis historický  
Český časopis historický  
Český lid  
Entomologické Listy  
Estetika  
Filologica Pragensia  
Filozofický časopis  
Folia Historica Bohemica  
Folia Zoologica  
Historica

Historica Series Nova  
Historická demografie  
Historická geografie  
Hudební věda  
Journal of Mass Spectrometry  
Linguistica Pragensia  
Listy filologické  
Listy filologické a paedagogické  
Litteraria Pragensia  
Mathematica Bohemica  
Mathematica\_Bohemica  
Moderní dějiny  
Naše řeč  
Nový Orient  
Onomastický Zpravodaj CSAV  
Photosynthetica  
Právněhistorické studie  
Právník  
Physiologia Bohemoslovaca  
Physiologia Bohemoslovenica  
Rozpravy Čes. Akademie věd – řada společenských věd  
Rozpravy Čes. Akademie věd – řada technických věd  
Sborník historický  
Sborník Prací Přírodovědecké Fakulty University Palackého v Olomouci obory  
Matematika, Fyzika, Chemie  
Sborník Přírodovědecké Fakulty University Palackého v Olomouci  
Slavia  
Slavia  
Slezský sborník  
Slovanský Přehled  
Slovo a slovesnost  
Sociologický časopis  
Soudobé dějiny  
Studia Geophysica Et Geodaetica  
Umění  
Věstník matice opatské  
Zoologické Listy  
Zpravodaj Místopisné Komise CSAV

### **Rozpracované**

Byzantinoslavica  
Collection Czechoslovak Chemical Communication  
Československý časopis pro fyziku

Československá Fysiologie  
Československá Morfologie  
Folia Biologia  
Kybernetika  
Politická Ekonomie  
Právník

## **V plánu**

Bulletin ČSAV PsÚ Praha  
Collection Symposium Series  
Czech Sociological Review  
Czechoslovak Economic Papers  
Časopis České Společnosti Entomologické  
Časopis Pro Mineralogii a Geologii  
Eirene  
European Journal Of Entomology  
Folia Morphologica  
Funcional Developmental Morphology  
Kybernetika  
Physiologia Bohemoslovenica  
Psychologické Zprávy  
Slovanské Historické Studie  
Výzkumné Zprávy  
Zprávy Psychologické Laboratoře ČSAV  
Zprávy Psychologické Laboratoře v Brně ČSAV  
Zprávy Psychologický ústav AV ČR  
Monografie vydané ČSAV, její předchůdkyněmi a následovnicemi

# OBSAH

<b>Úvodem</b> (J. Burgetová) .....	3
<b>Digitalizační centrum Knihovny AV ČR</b> (A. Ryšánková – Michal Pšenička) .....	5
<b>Kontaktní informace</b> .....	13
<b>Předpokládaný rozvoj Digitalizačního centra</b> (M. Lhoták) .....	14
<b>Projekty</b> (M. Lhoták) .....	15
<b>Novinky z domova i ze světa</b> .....	18
<b>Přílohy</b>	
1. <b>Ceníky</b> .....	19
2. <b>Soupis digitalizovaných časopisů</b> .....	22







PRO PRACOVNÍKY STŘEDISEK VĚDECKÝCH INFORMACÍ  
A KNIHOVEN AV ČR VYDÁVÁ KNIHOVNA AV ČR, v.v.i.

REDIGUJE PhDr. JARMILA BURGETOVÁ

DUBEN 2009

ISSN 12 10-8502

