

# Digitalizace v Archivu MU – první krůčky

Miroslav Bartošek, Michal Hrabí,  
ÚVT MU

Archiv MU spravuje rozsáhlé sbírky dokumentů vztahujících se k historii i současnosti Masarykovy univerzity. Jde nejen o tradiční tištěné dokumenty, ale též o dokumenty obrazové, zvukové či videonahrávky, často cenné jak z historického tak i informačního hlediska. Obecné povědomí o existenci těchto dokumentů a přístup k nim nejsou však zatím příliš rozšířené. Je to dáno i tím, že soupisy archiválií, natož pak jejich obsah, nejsou k dispozici v digitální podobě. Situace se však začíná pomalu měnit. Ústav výpočetní techniky MU zahájil před časem s Archivem MU neformální spolupráci s cílem začít s pilotní digitalizací některých vybraných sbírek, a to v oblasti fotografií a videonahrávek. I když jde zatím spíše o izolované „ad-hoc“ projekty, přispívají tyto aktivity k širšímu využívání ICT technologií v univerzitním archivu a postupnému otevírání jeho sbírek.

## 1 Digitalizace fotografií

Již v roce 2001 začala neoficiální spolupráce ÚVT MU s Archivem MU při digitalizaci fotografií. Jejím cílem bylo ověřit postupy a nastavení (digitalizační parametry), odzkoušet dostupné technologie (skenery a programy) a vyvinout potřebné nástroje (aplikace pro popis, uchovávání a zpřístupnění digitalizovaných fotografií). Pro vlastní digitalizaci se používají běžně dostupné skenery střední třídy (dříve např. HP ScanJet 7400, nyní Microtek ScanMake i900). Digitalizační rozlišení se volí podle velikosti předlohy; u běžných rozměrů fotografií se používá rozlišení 600 dpi, u negativů malých rozměrů i vyšší. Pro správu a zpřístupnění digitálních fotografií je využíván vlastní systém ÚVT MU.

### 1.1 Digitální knihovna fotografií

Pro ukládání digitalizovaných fotografií, jejich popis, třídění, vyhledávání a zpřístupnění uživatelům je využívána webová aplikace *Digitální knihovna fotografií* (DKF-MU) vyvinutá na ÚVT

MU – viz <http://dkf.ics.muni.cz>. Tato aplikace slouží jak kurátorům fotosbírek tak koncovým uživatelům. Umožňuje jim pracovat pomocí standardního webového prohlížeče s digitálními objekty tří typů:

- *kolekce* (zastřešuje fotografie spadající do určité sbírky – například Lidé na MU);
- *složka* (sdružuje všechny fotografie vztahující se k jednomu objektu příslušné sbírky – například složka Karel Engliš v kolekci Lidé na MU);
- *fotografie* (digitální verze konkrétní fotografie – například Karel Engliš při inauguraci rektorem MU).

Pro každý z výše uvedených tří typů objektů je možné vytvořit v DKF-MU metadatový popis podle předdefinovaného schématu. Ten slouží pro popis příslušného objektu a současně i pro potřeby vyhledávání. Pro každý digitální objekt lze nastavit míru jeho dostupnosti koncovým uživatelům (volně dostupný na internetu, dostupný pouze v síti muni.cz, dostupný pouze ze zadaných stanic či pouze stanoveným autentizovaným uživatelům).

Po vložení fotografie do DKF-MU (povoleno pouze administrátorům a editorům příslušné kolekce) vygeneruje systém čtyři její digitální deriváty:

- *Archivní verze*: digital master, digitalizovaná fotografie v původní velikosti a originálním formátu – např. png nebo jpeg. Je možné uložit ji ve dvou instancích: jedna slouží pro uložení fotografie v původní formě tak, jak byla naskenována či digitálně nafotografována – bez jakýchkoliv úprav, zatímco druhá instance archivní verze se obvykle užívá pro uložení digitálně upravené (vylepšené) verze fotografie. Archivní verze je dostupná pouze kurátorům kolekce, není dostupná koncovým uživatelům.
- *Náhled*: thumbnail, zmenšená verze fotografie ve formátu jpeg generovaná tak, aby delší strana náhledu měla velikost 180 pixelů; používá se při přehledovém zobrazení obsahu složek a výsledků vyhledávání.
- *Menší přístupová kopie*: zmenšená verze fotografie ve formátu jpeg, s delší stranou o velikosti 640 pixelů; slouží pro zobrazení fotogra-

fie na monitorech s menší velikostí či nižším rozlišením.

- *Větší přístupová kopie*: zmenšená verze fotografie ve formátu jpeg, s delší stranou o velikosti 1024 pixelů; slouží pro zobrazení fotografie na monitorech s větším rozlišením.

## 1.2 Dostupné kolekce fotografií

V současnosti obsahuje DKF-MU na 3000 digitalizovaných fotografií z Archivu MU, dostupných v počítačové síti MU. Především jde o fotografie v kolekcích Lidé na MU (fotografie celkem 812 různých osobností a pracovníků univerzity), Čestné doktoráty (fotografie z udělování 95 čestných doktorátů) a Medaile MU (zatím jen fotografie z udělování 62 Velkých zlatých medailí a zlatých medailí MU).

Kromě Archivu MU využívají systém DKF-MU také další pracoviště univerzity. Především Ústav výpočetní techniky MU (fotodokumentace výpočetní techniky, osob a událostí), Fakulta informatiky MU (fotodokumentace z konferencí) a Rektoriát MU (Univerzita třetího věku, foto-galerie uměleckých děl). DKF-MU slouží i jako podpůrný nástroj pro jiné systémy zpřístupňující fotografie. Například nová webová prezentace MU přebírá automatizovaně z DKF-MU fotografie a metadata osobností a čestných doktorů MU, zpřístupňuje a prezentuje je ale již ve vlastní režii a vlastním grafickém designu.

## 1.3 Některé zkušenosti

Na příkladu některých kolekcí z univerzitního archivu se názorně ukázalo, že technické řešení digitalizace a zpřístupnění sbírek je často spíše tou jednodušší částí celého projektu. Mnohem obtížnější a déle trvající bývá někdy řešení doprovodných otázek, souvisejících spíše s lidskou a společenskou dimenzí problému. Například v souvislosti s digitalizací fotosbírkou Lidé na MU, a zejména pak s jejím zpřístupněním na webu, se vynořily problémy, které u původní sbírky (fotografií uložených v krabicích na regálech Archivu MU) nikoho zvláště netrápily: Kdo by vlastně měl být ve vystavené kolekci zachycen a kdo nikoliv? Měly by to být jen „osobnosti“? Kdo to ale je *osobnost MU*? Jak se vypořádat s historickými

hledisky a „vrstvami“ (v sedmdesátých či osmdesátých letech minulého století byly za osobnosti považovány, a tudíž v archivu dokumentárně zachyceny, zcela jiné typy osobností než dnes)? Kdo je povolán k tomu, aby o takovýchto otázkách vůbec rozhodoval? Chceme-li mít jako součást metadatového záznamu i charakteristiku osoby, její medailonek – kdo ho vytvoří a kdo posoudí míru objektivnosti takové charakteristiky? Ne vždy bývá snadné na takovéto otázky odpovědět.

## 2 Digitalizace a archivace videí

Zatímco technická stránka digitalizace fotografií je dnes již rutinní záležitostí, video poskytuje mnohem výživnější půdu pro zkoumání, ověřování technologií a hledání optimálních digitalizačních postupů. Archiv MU spravuje stovky videonahrávek o akcích na MU i několik komponovaných pořadů o univerzitě samotné, pořízených během dlouhé řady let – nejstarší z nich pochází z padesátých let minulého století. Stále akutnějším problémem se stává jednak technologické zastarávání nosičů a formátů dat (VHS, S-VHS, U-matic aj.) a jejich zpřístupňování uživatelům, jednak fyzické stárnutí vlastních médií s videonahrávkami. Trvanlivost videí na klasických analogových páskách je značně omezená. Časem dochází k samovolnému poškození záznamu různými vlivy, jako je magnetické pole, vlhkost, stárnutí pásku atd. Mnoho pořadů nahraných pro MU (či ještě UJEP) je archivováno pouze na jediném nosiči (obvykle typu VHS), a každým rokem tak dochází k nevratným ztrátám větších a větších částí videonahrávek.

S digitalizací videonahrávek Archivu MU začal Ústav výpočetní techniky v roce 2006, v podstatě jako s vedlejší aktivitou při přípravě informačního obsahu nové webové prezentace MU. Původní cíl – připravit několik zajímavých videoukázek pro veřejný web univerzity – přerostl postupně v něco mnohem zásadnějšího: ve snahu zachránit nejcennější záznamy obsažené na páskách v Archivu MU; digitalizovat je a zálohovat do takových formátů a na takové nosiče, které lépe zajistí jejich existence i do budoucna, a současně také usnadní přístup veřejnosti k videonahrávkám.

## 2.1 Digitalizační postup

Digitalizace pásek z Archivu MU probíhá v několika krocích; dále je stručně popíšeme. Při realizaci většiny kroků se využívá technického a programového vybavení Laboratoře pokročilých síťových technologií – společného pracoviště Fakulty informatiky MU, Ústavu výpočetní techniky MU a sdružení CESNET, z. s. p. o.

**Grabování.** Nejprve je video transformováno z analogového formátu (nejčastěji VHS a S-VHS) do digitálního (DV AVI) za pomoci AD/DA převodníku Canopus ADVC-100 a programu VirtualDub či Sony Vegas Studio. Tomuto procesu se říká grabování, či capturing videa. Video je nahráno dle normy PAL (720 × 576 pixelů, 25 snímků za sekundu, prokládaně). Soubory s videem ve formátu DV AVI jsou velké (1 minuta záznamu zabírá přes 210 MB prostoru), z čehož vyplývá i potřeba rozsáhlých diskových polí. Video je však v plné kvalitě a snadno zpracovatelné pomocí dalších aplikací.

**Úpravy videa.** Nad digitálním formátem je následně prováděna filtrace a maskování neduhů, jako je dynamický šum a drop out (drobné výpadky částí obrazu). Ve speciálních případech se mění i hlasitost zvukové stopy či opravuje jas a kontrast obrazu. Video je rovněž nutné oříznout o přebývajících minutách materiálu, která byla během grabování nahrána do počítače. Všechny tyto úpravy jsou prováděny v programu VirtualDub, stejně jako následné přidání titulků s jednoduchými stříhovými efekty (prolínačkami).

**Výroba archivačních a přístupových kopií.** Výsledkem předchozích kroků jsou DV AVI soubory určené k archivaci souběžně na diskových polích a na digitálních páskách. Hodinový pořad představuje DV AVI soubor o velikosti přibližně 13 GB dat. Jako sekvenční médium pro dlouhodobou archivaci videa se využívají pásky SONY PHDV-276DM (digital master), které mají kapacitu až 184 minut záznamu ve formátu DV. Odhadovaná životnost pásek se pohybuje kolem 30 let. Vzhledem k tomu, že průměrná délka digitalizovaného záznamu nepřevyšuje jednu hodinu, je na každé pásece archivováno několik pořadů. Veškeré operace až do tohoto okamžiku se provádějí s využitím bezztrátové komprese. Kromě archivačních médií (DVCAM pásek) jsou nyní vytvářena ještě

média určená pro běžnou práci s videonahrávkami, s možností jejich případného půjčování veřejnosti. Pro tento účel byl zvolen formát DVD. O převod z DV AVI na DVD-Video se stará program Canopus ProCoder. Následuje vypálení vzniklých souborů jako DVD-Video disk. Jako médium bylo pro svou deklarovanou dlouhou životnost vybráno DVD+R Emtec Gold. Každý nosič obsahuje pouze jeden záznam (komprimovaný již ztrátovou kompresí).

**Video-ukázky.** Poslední operací s videem před jeho katalogizací je vytvoření krátké ukázky (30 až 90 sekund), která by měla sloužit případným zájemcům o vypůjčení DVD z Archivu MU jako náhled. Toto video vzniká jako krátký sestřih celého záznamu a je kódováno do formátu Windows Media Video (WMV), který by měla být schopna přehrát většina uživatelů osobních počítačů. Videoukázka je zpřístupněna uživateli on-line, prostřednictvím univerzitní instalace volně dostupného systému DSpace<sup>1</sup> – viz další krok.

**Katalogizace a zpřístupnění metadat.** Posledním krokem ve zpracování videí je „katalogizace“ digitalizovaného záznamu videonahrávky. Z údajů poskytnutých Archivem je vytvořen metadatový záznam (jednoznačná identifikace, název videa, autor, datum vytvoření apod.), který je spolu s video-ukázkou uložen do kolekce „Archiv MU – Digitalizovaná videa“ v digitálním depozitáři DSpace, a jeho prostřednictvím zpřístupněn veřejnosti. Mimo jiné je záznam v DSpace využit i pro automatizované generování potisku obalu příslušných DVD médií.

## 2.2 Několik tipů na zajímavé záznamy

Digitalizace videí z Archivu MU stále pokračuje, a tak v DSpace postupně přibývají nové položky. Větší část tvoří záznamy udílení čestných doktorátů či jiných ocenění Masarykovy univerzity, dále byla digitalizována některá videa s historickými pořady o univerzitě. V archivu je tedy možné nalézt udělení čestného doktorátu Václavu Havlovi nebo udělení Velké zlaté medaile MU bývalé ministryni zahraničí USA Madeleine Albrightové. Zájemci o dokumenty pak ocení pořad Československé televize Brno z roku 1988

<sup>1</sup><http://dspace.muni.cz>

*Vědění je moc* nebo další videa, která vznikla u příležitosti kulatých výročí univerzity.

### **2.3 Co do budoucna**

V současné době obsahuje deponitář DSpace pouze ukázky jednotlivých záznamů. Do budoucna se však uvažuje o umístění vybraných pořadů na DSpace v nezkrácené délce. Zájemci tedy budou mít možnost zhlédnout některé zajímavější záznamy bez nutnosti fyzického vypůjčení video-nosiče v Archivu MU. Dalším krokem ve zpracování záznamů z Archivu MU bude digitalizace videí starších formátů, jako je např. U-matic. □