

Projekt Elektronického archivu

► Petr Zeman

Na tiskové konferenci 10. července 2009 byla představena pilotní fáze projektu Elektronického archivu.¹ Je společným projektem Archivu bezpečnostních složek (dále Archiv) a Ústavu pro studium totalitních režimů (dále Ústav) vycházející ze zákonem dané povinnosti převodu dokumentů uložených v Archivu do elektronické podoby.² Vytváření digitálních kopií uložených archiválií má za cíl nejen jejich uchování a snadnější manipulaci s nimi v rámci vnitroarchivních agend, ale rovněž jejich zjednodušené předkládání badatelům. Elektronické kopie archiválií se pořizují na digitalizačním pracovišti Ústavu v otevřeném formátu JPEG v černobílé (pokud jsou předlohou mikrofiše) či barevné škále (ostatní dokumenty, negativy), v minimální tiskové kvalitě 300 dpi.

Značná technická náročnost celého projektu vedla k rozhodnutí zavádět celý systém postupně v několika krocích. První z nich spočíval ve zkušebním zprovoznění Elektronického archivu v badatelnách archivu (Praha – ul. Siwiewcova a Na Struze, Brno-Kanice). V této fázi bylo do Elektronického archivu vloženo 335 jednotek archivních materiálů. Během následujícího zkušebního provozu, který doposud probíhá, je testována schopnost systému snášet zátěž, postupně je také vyhodnocována funkčnost ovládání a funkčnost nově zaváděných prvků. Následovat bude plně zprovoznění systému Elektronického archivu v badatelnách Archivu, a následně umožnění dálkového přístupu, díky němuž bude možno prohlížet vybrané archiválie na internetu.

Řešení obtíží spjatých s technickým vývojem a administrací systému leží na bedrech Ústavu, nemalé výzvy přineslo ovšem i Archivu. Již při přípravě bylo nutno vyřešit několik základních problémů:

- 1) přípravu archiválií na digitalizaci,
- 2) vytvoření nové struktury archivních fondů, která by korespondovala se současným stavem zpracování, ale umožnila zároveň i případné změny v budoucnu,
- 3) uvedení digitalizovaných archiválií do takového stavu, aby byly naplněny zákonné normy, které upravují předkládání archiválií badatelům, tedy především zákon o archivnictví a spisové službě,
- 4) přístup badatelů do elektronického archivu, který by byl v souladu se zákonem danými povinnostmi.

Při přípravě archiválií na digitalizaci bylo nezbytné stanovit jasný postup práce s archiváliemi, kterým se budou zaměstnanci Archivu řídit a stanovit formu tzv. digitalizačního protokolu, do něhož se zaznamenají údaje nezbytné pro další práci jak s archiváliemi samotnými, tak s jejich elektronickou kopií.

1 Viz: <http://www.abscr.cz/cs/elektronicky-archiv>

2 § 4 zákona č. 181/2007 Sb.

V souvislosti se zpracováváním struktury archivních fondů, podfondů a dalších skupin archiválií, bylo nutné určit základní druhy metadat, která by umožnila elektronickou kopii dané archiválie jednoznačně zařadit a identifikovat. Metadata při re-spektování archivních zásad musí zároveň umožnit co nejjednodušší vyhledávání archiválií badatelskou veřejností. Základem pro metadataci se stal tzv. identifikátor – tj. jednoznačný kód, umožňující danou archiválii nalézt. U inventarizovaných či částečně inventarizovaných fondů se jako tohoto identifikátoru používá nejčastěji kombinace značky fondu a čísla inventární jednotky, u fondů alespoň částečně uspořádaných pak (dle druhu tohoto uspořádání) jeden z několika možných identifikátorů (dočasná signatura, archivní nebo registrační číslo, kombinace jména a data narození apod.). Doplnková metadata následně tvoří především jméno registrované osoby, krycí jméno svazku, obsah inventární jednotky apod. Tato metadata výhledově usnadní fulltextové vyhledávání v Elektronickém archivu.

Velkou překážkou při zpracování metadat se ukázala být digitalizace neuspořádaných archivních fondů. Archiv při svém vzniku delimitoval kilometry neuspořádaných archiválií, k nimž je (v lepším případě) vytvořen alespoň jednoduchý „přírůstkový soupis“; mnohdy ovšem chybí i toto základní vodítko. Elektronické kopie tak není možné zařadit do struktury archivních fondů a nezbývá, než s jejich zpřístupněním veřejnosti elektronickou formou ještě posečkat.

Veškeré činnosti související s vývojem systému jsou navíc vzájemně provázané, takže již při tvorbě norem, podle nichž se archiválie připravují k digitalizaci, se současně ukázala nezbytnost vytvoření takové formy předávacího protokolu, která by umožňovala jak zapsání metadat do paralelně vytvářené struktury fondů, tak rozlišení jednotlivých archivních jednotek dle způsobu anonymizace.

Tvorba Elektronického archivu a jeho naplňování elektronickými kopiemi je poměrně složitým vícefázovým procesem. Vznik elektronické kopie začíná přípravou konkrétní archivní jednotky na příslušném oddělení. Zaměstnanci Archivu ošetří archiválii předepsaným způsobem a připojí k ní digitalizační protokol, v němž jsou dány závazné pokyny jak pro oddělení digitalizace Ústavu, tak pro skupinu správy digitálního archivu. Protokol je prozatím vyplňován pouze na papír, časem však bude umožněno jeho vyplnění v elektronické podobě, což výrazně zrychlí metadataci elektronických kopií archiválií. Následně je archiválie převezena na digitalizační pracoviště, předepsaným způsobem digitalizována a navrácena zpět do depozitáře. Vytvoření elektronické kopie archiválie je však pouhým začátkem v procesu přípravy, na jehož konci může být archiválie v elektronické podobě předložena badateli.

Poté, co je archiválie digitalizována a nahrána do elektronického úložiště, nastává fáze kontroly a případné anonymizace údajů. Vzhledem k tomu, že Archiv může dle zákona o archivnictví a spisové službě předkládat badatelům pouze archiválie vzniklé před 1. lednem 1990, musí být údaje po tomto datu anonymizovány. Stejně tak podléhají anonymizaci údaje v archiváliích, na něž se vztahuje § 15 zákona o Ústavu

pro studium totalitních režimů a Archivu bezpečnostních složek.³ Zcela specifickou množinou jsou pak archiválie, které již byly dříve zpřístupněny badatelům na bývalém pracovišti Archivu v Pardubicích, dle zákona o zpřístupnění svazků vzniklých činností bývalé StB, pro něž platí zvláštní režim anonymizace. Vzhledem k tomu, že se objem anonymizovaných údajů (dle třicetileté lhůty dané archivním zákonem) v tomto typu materiálů každým rokem snižuje, a navíc k nutnosti rozlišovat, komu je dotyčná archiválie předkládána, bude zřejmě nezbytné vytvořit jakési „dočasné kopie“, které se budou každým rokem nahrazovat aktuální verzí. Ukáže-li se ovšem tato cesta z technických a personálních důvodů příliš náročnou, nezbude, než předkládat elektronické kopie – zpřístupňované dle tohoto zákona – ad hoc při konkrétním požadavku badatele. Samotná anonymizace elektronických kopií je řešena jednoduše. U předmětné elektronické archiválie je vytvořena kopie (jedna, popř. více), v níž se provedou prostým začerněním potřebné úpravy.

Po těchto vykonání výše popsaných úkonů, je archiválie ve stavu, který umožňuje zpřístupnění a může být tedy předložena badateli.

Všechny tři badatelé Archivu disponují minimálně jedním počítačem umožňujícím přímý přístup do datového úložiště elektronických kopií, přičemž přístup se děje formou uživatelského jména a hesla. Badatel, který má zájem o studium archiválií v Elektronickém archivu, požádá obsluhu badatelný o vydání formuláře s uvedenými údaji. Následně se může přihlásit do systému a volně prohlížet archiválie. Vyhledávání je možné dvěma způsoby – buď dle struktury archivních fondů, která je do značné míry totožná se strukturou již nyní zveřejněnou na webových stránkách Archivu (v levé části obrazovky); nebo druhým způsobem, což je vyhledávání v pravém horním rohu. Systém vyhledávání se stále rozvíjí a v budoucnu se možnosti vyhledávání podle různých kritérií ještě rozšíří. Badatel si může prohlížené stránky dle potřeby tisknout, později bude zprovozněno i přímé vypalování na CD/DVD.

Při současném způsobu přístupu do Elektronického archivu je největším problémem sloučení systému s informacemi o elektronických kopiích materiálů, které se v něm nacházejí. Na internetových stránkách Archivu je sice zveřejněna struktura archivních fondů i s příslušnými inventáři (pokud jsou k dispozici), ale z bezpečnostních důvodů prozatím nelze zavést internetové připojení přímo k počítačům v badatelárnách. Badatelé si tak musí potřebné informace zjistit ze stránek ABS doma, a potom s vypsanými signaturami navštívit badatelnu. Přestože je to výrazně rychlejší cesta, nežli tradiční zadání dotazu a následné čekání na odpověď a přípravu materiálu, jde z uživatelského hlediska o jisté nepohodlí, které se Archiv ve spolupráci s Ústavem snaží řešit.

Po skončení zkušební fáze provozu, která bude pravděpodobně běžet do počátku roku 2010, lze předpokládat, že Elektronický archiv bude při stávajících personálních podmínkách doplňován tempem 600–800 archivních jednotek měsíčně.

3 Podle ustanovení tohoto paragrafu zákona č. 191/2007 Sb. může Archiv odepřít nahlížení do archiválií a poskytování kopií, opisů a výpisů z nich, pokud obsahují informace i nadále důležité pro ochranu ústavního zřízení, významných ekonomických zájmů, bezpečnost a obranu České republiky. O odvolání proti rozhodnutí o odepření rozhoduje ředitel Archivu.

Další rozvoj Elektronického archivu bude záležet vcelku pochopitelně na financích, které se na tento projekt podaří v budoucnu uvolnit. Důležitý faktor sehráje i rychlost zpracování neuspořádaných fondů, což opět závisí především na personálních (a opět finančních) možnostech Archivu. Vzhledem k stávající finanční situaci není možné se vyhnout pečlivému výběru archivních jednotek, které budou zařazeny do Elektronického archivu, přičemž hlavními kritérii pro výběr jsou stupeň poškození archiválií a jejich celková žádanost badateli. Do budoucna lze naštěstí předpokládat, že po překročení určitého „kritického“ množství archiválií existujících v Elektronickém archivu, se objem práce sníží, protože bude možno zaměstnance zabývajících se v současné době správou, vyhledáváním a předkládáním originálů, uvolnit k přípravě archiválií na digitalizaci a k jiným činnostem.

Bohužel, však nelze dopředu odhadnout, kde se ono „kritické“ množství archiválií nachází... Doufejme, že se během roku 2010 objeví náznaky, že se k němu alespoň blížíme.



Tisková beseda k otevření Elektronického archivu

Zdroj: ABS.