

Pracoviště elektronické archivace v Národním archivu a předpoklady jeho činnosti

PhDr. Tomáš Kalina - Ing. Miroslav Kunt

Otištěno: ISSS 2004 - Internet ve státní správě a samosprávě Hradec Králové. Sborník příspěvků. Triada, spol. s.r.o., Praha, 2004, s. 143-144.

Tímto příspěvkem navazujeme na naši loňskou stat' pro konferenci ISSS "Elektronická archivace – výzva pro odborníky více oborů", kde jsme se pokusili definovat z teoretického i praktického hlediska základy dlouhodobého ukládání elektronických dokumentů.

Národní archiv (do 31.12.2004 Státní ústřední archiv) má podle nového zákona č. 499/2004 Sb., o archivnictví a spisové službě mimo jiné plnit úkoly ústředního vědeckovýzkumného pracoviště a odborného metodického a školicího centra v oblasti dlouhodobého uchovávání elektronických dokumentů. Zmíněný zákon tak doplňuje dřívější usnesení vlády č. 11 ze 7. ledna 2004 k dlouhodobému uchovávání a zpřístupňování dokumentů v digitální podobě, které předpokládá, že se archiv stane trvalým úložištěm elektronických archiválií s celostátní působností.

Předpokladem pro řešení naznačené problematiky je kromě přípravy "podkladů pro projekt celostátního pracoviště" (z cit. vládního usnesení) především každodenní praxe, která spočívá v péči o tyto specifické archiválie - od evidence u původců, přes výběr, ukládání až po jejich zpřístupňování. Platná legislativa však nedává téměř žádné možnosti účinně ovlivňovat tvorbu elektronických dokumentů - především z hlediska definování softwarově nezávislých formátů vhodných pro dlouhodobé uchovávání a příslušných procesů, což je pro elektronickou archivaci věcí nezbytnou. Tady je namísto zmínit se o standardizaci v oblasti EDMS systémů, která v současné době hýbe IT Evropou ve veřejné správě, především o standardu MoReq (Model Requirements for the Management of Electronic Records), zpracovaném v rámci Evropské komise.

Nemáme tedy žádné možnosti něco "nařizovat", ale jak jsme si ověřili již v minulosti, je nejučinnější průběžná a vstřícná komunikace s příslušnými původci (většina ministerstev, ústředních úřadů a jimi zřizovaných organizací). Ti mají povinnost vyřazovat veškeré své dokumenty (tedy i elektronické) po schválení Národním archivem. Bez tohoto schválení není tedy možné nejen zlikvidovat "starý papír", ale ani např. vymazat databázi či internetové stránky. V tomto ohledu je nový archivní zákon daleko důslednější než předchozí, obsahuje výslovně instituty státní kontroly, správního a přestupkového řízení včetně pokut.

V oblasti elektronických dokumentů jsme se většinou setkali s pochopením a zájmem o spolupráci (Úřad vlády, Český telekomunikační úřad, Ministerstvo spravedlnosti, Výzkumný ústav vodohospodářský T.G.M. aj.). V první fázi se vedle průzkumů zaměřujeme na „osvětovou“ činnost: vysvětlujeme rozdíly mezi zálohováním a archivací, nutnost nezávislých formátů, problematiku migrace apod., získáváme první elektronické dokumenty. Povědomí o archivaci se v posledních letech zejména u IT pracovníků zvýšilo.

Z provedených průzkumů zjišťujeme, že veřejná správa skutečně disponuje gigantickým množstvím mimořádně cenných dat, z nichž část je ohrožena - aktualizacemi, přechodem na jiné systémy, zrušením či předáním agendy. Problematika přitom nesouvisí s informačními systémy veřejné správy jak je definuje zákon 365/2000 Sb. Naopak, i řadu dat tzv. provozních systémů lze kvalifikovat jako cenné archiválie. Z našeho pohledu je také zcela irelevantní, zda jsou systémy certifikovány - k jedinému pro nás důležitému problému, otázce archivace žádný národní standard neexistuje. Snahy o jeho definování, které jsme spíše jako neoficiální iniciativu spolu s odborníky z MV a firmy Relsie, spol. s r.o. v letech 2002-2004 vyvíjely, se neseťkaly ze strany Ministerstva informatiky s žádnou odezvou. Musíme si tedy poradit sami.

V současné době kromě konstituování nového pracoviště připravujeme vznik pracovní skupiny, která by standard a doporučení řešila. Standard by pak měl být využíván v duchu kompetence Národního archivu, jak bylo naznačeno na začátku.

K praktickým otázkám elektronické archivace dle nové legislativy.

Ustanovení §8 odst. 8 vyhlášky MV č. 646/2004 Sb., o podrobnostech výkonu spisové služby, která se jediná otázky v české legislativě dotýká, je poněkud mlhavé: *"Dokument v digitální podobě se skartačními znaky "A" a "V" vzniklý z činnosti určeného původce musí být zapsán ve formátu, který zaručí jeho neměnnost a umožní jeho následné čtení. Pokud tuto podmínku nemůže určený původce zabezpečit, převede takové dokumenty do analogové formy odpovídající době jejich vyřízení a opatří je náležitostmi originálu, a to nejpozději před jejich zařazením do skartačního řízení; obdobně postupuje při vyřizování dokumentů přijatých určeným původcem v digitální podobě včetně potvrzení o platnosti elektronického podpisu v době přijetí dokumentu."*

Za "umožnění následného čtení" považujeme dokument zásadně pouze v softwarově nezávislém formátu, který bude přejímajícím archivem definován (např. XML, TXT s oddělovači), jak je ostatně obvyklé i ve světě (např. směrnice Spolkového archivu SRN). Pokud tedy nebudou chtít původci dokumentů všechno tisknout (což samozřejmě nechceme ani my), musí zajistit, aby jejich informační systémy disponovaly příslušným rozhraním.

Na rozdíl od autorů vyhlášky nesouhlasíme s tím, že pouze to, co má "náležitosti originálu" je archiválií. Zaručený elektronický podpis (ani po doplnění časovými razítky) nelze považovat za dostatečný a vhodný k dlouhodobému uchovávání (na desítky a stovky let). Jako řešení navrhujeme využití hashovacích funkcí (známé a popsané algoritmy SHA-1 či MD5 používané i v rámci elektronického podpisu). Výsledný otisk dat (produkt hashovací funkce) by byl potom zaznamenán na protokolu, podepisovaném oprávněnou osobou zodpovědnou za autenticitu dat. Originál takového protokolu by pochopitelně existoval pouze na papíře. Tento postup je samozřejmě nutné uplatnit u dokumentů, které jsou dle vyhlášky označeny ve skartačních plánech "A" (archiválie) nebo "V" (teprve před skartačním řízením se rozdělí na dokumenty "A" a na ty určené ke zničení). Ze zákona je archiválií např. podací protokol jakožto základní evidenční pomůcka spisové služby.

U ostatních dokumentů, které se neevidují v rámci spisové služby – např. internetové stránky, databáze apod. postačí, když budou ve skartačním řádu uvedeny ve zvláštní příloze a dohodne se způsob předávání dat, opět s důrazem na nezávislý formát. To se týká i dokumentů, které jsou v "papírové" podobě, ale bylo by vhodné zachovat paralelně i podobu elektronickou (porady ministra, materiály vlády). Je to vhodnější také proto, že u elektronických dat není potřeba a často ani vhodné používat skartačních lhůt, které mohou přivodit zneprístupnění dokumentu (nepoužívaný formát na zastaralém médiu, bez dokumentace a osob "pamatujících" provoz).

V případě osamoceně vznikajících elektronických dokumentů (produktů kancelářských programů apod.) uložených v datových úložištích elektronických podatelen nebo nemajících vazbu na systémy spisové služby (DMS) považujeme za nejlepší řešení jejich tisk. Další práce s nimi je pak zabezpečena v papírové podobě (v rámci spisu). Totéž platí v případě, že část dokumentů evidovaných v systému spisové služby je elektronických a část analogových (papírových).

Pokud je nám známo, žádný ze současných systémů spisové služby/EDMS nemá rozhraní k migraci dokumentů na nezávislý datový formát. Inspirativním a perspektivním řešením je systém DOMEA realizovaný německou spolkovou vládou směřující k plně elektronickému úřadování.

Část veřejné správy neměla dosud vytvořeny předpoklady, aby mohla elektronicky úřadovat: chyběly především spisové plány nutné pro zatřídění úřední produkce, použitelné spisové řády a v neposlední řadě kvalifikované řízení spisové služby jako takové. Ostře to kontrastuje s prostředky, které se vydávají na budování příslušných, mnohdy značně diskutabilních, informačních systémů (z našeho pohledu např. systémy spisové služby neřešící ukládání evidovaných dokumentů, spisovnu). Dílčí řešení vyjádřené např. vyhláškou o elektronických podatelných problematiku archivace zcela opomíjejí.

Je tedy namístě znovu a důrazně konstatovat: bez e-archivace žádný e-governement nebude.

Příloha:

Představa činnosti pracoviště péče o elektronické dokumenty (ED) Národního archivu (NA)

Průzkumy ED u původců:

- úzce souvisí s klasickou předarchivní péčí, avšak vyžadují vedle znalostí archivnictví základní znalosti i v oboru informatiky
- metodika průzkumů je již delší dobu vypracována a ověřena, je třeba ji jen důsledně realizovat
- vedle kontaktů na pracovníky spisové služby je nutné udržovat vazby na odborníky - specialisty jednotlivých původců v oblasti informatiky.

Normalizační činnost:

- zde by mělo spočívat hlavní těžiště práce s původci
- Za účelem dosažení co nejširšího konsensu vytvořit koordinační skupinu ze zástupců archivu, Ministerstva vnitra, původců, vysokých škol a soukromých firem, která by koncipovala doporučení pro péči o ED v celém jejich životním cyklu při zohlednění evropských a světových řešení (MoReq, ISO 15489, doporučení Mezinárodní archivní rady atd.)

Rešeršní činnost:

- systematicky vyhledávat a vyhodnocovat dokumenty a odbornou literaturu týkající se ED zejména z oblasti Evropské unie; hlavním zdrojem je přitom Internet

Přejímání ED:

- přejímky nevázat na skartační lhůty běžné u papírových dokumentů a výrazně využívat výběru archiválií mimo skartační řízení
- zaměřit se prioritně na ED s vysokou informační hodnotou (databáze, GIS).

Migrace formátů:

- v praxi realizovat dosud u nás jen teoreticky posuzované metody dlouhodobého uchování dokumentů
- na pracovišti ED navázat na dosavadní testy prováděné ve Státním ústředním archivu a na zkušenosti zejména střeoevropské (podobná tradice spisové služby) směřující k využívání značkovacích jazyků, zvláště formátu XML a rastrové grafiky, stejně jako dalších softwarově nezávislých formátů.

Migrace médií:

- s ohledem na investiční náročnost sledování parametrů uložených médií v počátečním stádiu uvažovat s redundantním uložením dat na různých typech médií, aby byla eliminována pravděpodobnost poruchy
- sledovat možnosti měření médií, zejména v případě CD-R (spolupráce s vysokými školami a Národní knihovnou).

Autentizace a autorizace:

- s ohledem na mezinárodní doporučení zvážit, jaký stupeň autorizace je u různých druhů archiválií nutný a účelný s ohledem na to, že řešení prostřednictvím elektronického podpisu a časových razítek nelze považovat za trvalé; nabízí se využití hashovacích funkcí v kombinaci s administrativními opatřeními - jde o celostátně důležité téma pro výzkum, některé otázky řešit ve spolupráci s vysokými školami)
- za prioritní však považovat migraci na nezávislý formát při zachování nezměněného obsahu

Stahování internetových stránek:

- podle možnosti dohodnout s původci předávání jejich www stránek off-line k uložení (u ústředních úřadů např. před koncem volebního období vlády apod.)
- spolupracovat s Národní knihovnou na projektu WebArchiv (i když je možné provádět stahování vybraných stránek i vlastními silami).

Metadata:

- s ohledem na nemožnost ovlivňovat vznik ED, je tento problém z velké části teoretický; východiskem by se měl stát Dublin Core nebo řešení vycházející z MoReq
- v rámci migrace formátů zkoumat možnost automatického získávání a zatřídění metadat popisujících datové soubory a procesy s nimi

Ostatní

- důsledně shromažďovat a archivovat dokumentaci k jednotlivým druhům ED, elektronickým systémům původců a o migraci dat
- spolupráce se zahraničními pracovišti, vysokými školami, Národní knihovnou a Ministerstvem informatiky jako důležitými zdroji znalostí a informací; základní výzkum by měl být vyhrazen vysokým školám, jeho provádění v NA není efektivní.