Digitalizace matrik v archivech České republiky s přihlédnutím k situaci ve Státním oblastním archivu v Plzni

Mgr. Ludmila NOVOTNÁ, Mgr. Jana KOPEČKOVÁ

Státní oblastní archiv v Plzni

lnovotna@soaplzen.cz

kopeckova@soaplzen.cz

INFORUM 2014: 20. ročník konference o profesionálních informačních zdrojích

Praha, 27. – 28.5.2014

Abstrakt

Matriční knihy jsou jedním z nejdůležitějších a nejvyužívanějších pramenů uchovávaných v archivech. Záznamy v nich obsažené si vždy zasluhovaly zvláštní ochranu. Digitalizace matrik je moderní způsob ochrany těchto pramenů a možnost, jak je zpřístupnit co nejširšímu okruhu zájemců. Všechny archivy v České republice, které matriky ve svých depozitářích mají, přistoupily k jejich digitalizaci. Příspěvek se týká historie matrik jako úředních knih, počátků jejich ochrany a tvorby zabezpečovacích kopií, počátků digitalizace a srovnání způsobu zpřístupnění digitalizovaných matrik v konkrétních archivech v České republice.

Státní oblastní archiv v Plzni začal s digitalizací matrik v roce 2005. Postupně bylo digitalizováno všech 11 500 svazků matričních knih. Nejprve byly zveřejňovány spolu s matrikami dalších archivů na portále www.actapublica. eu., od roku 2013 na vlastním portále www.portafontium.eu . Kromě matrik Státní oblastní archiv v Plzni, podobně jako ostatní archivy v ČR, digitalizuje i jiné typy archiválií: listiny, kroniky, fotografie. Webový portál „Porta fontium" je výsledkem projektu „Bavorsko-česká síť digitálních historických pramenů", financovaného z prostředků programu přeshraniční spolupráce Cíl 3 Česká republika –Svobodný stát Bavorsko 2007 – 2013 a České republiky. Zpřístupňuje badatelům v digitální podobě archiválie ze západních Čech i Bavorska.

Matriční knihy jsou jedněmi z nejdůležitějších archiválií uložených v archivech České republiky. Obsahují záznamy o narození, sňatcích a úmrtí osob od 2. poloviny 16. století.

Vedení matrik bylo dlouhá staletí záležitostí církve. Tridentský koncil v roce 1547 nařídil evidenci udělených svátostí – nejprve křtu jako svátosti nejdůležitější, vzápětí se připojila i evidence sňatků a úmrtí. V českých zemích se matriční knihy více rozšířily až ve 30. letech 17. stol. Nejstarší dochovanou matrikou u nás je matrika oddaných pro město Jáchymov z r. 1531, shodou okolností uložená ve Státním oblastním archivu v Plzni. Obecně nejstarší matriky u nás pocházejí ze severozápadního pohraničí, kam se rozšířily vlivem luterství a hospodářských a kulturních styků se Saskem. Zpočátku jednotlivé církve vedly matriky pro

své členy, stát začal zasahovat až za josefínských reforem: 1787 byly katolické matriky uznány jako veřejné knihy (nekatolické až 1829, židovské 1868). Roku1870 byly zavedeny knihy pro osoby bez vyznání u tehdejších okresních úřadů. Zákonem ze 7. 12. 1949 byly zavedeny jednotné státní matriky, dosavadní církevní byly prohlášeny za majetek socialistického státu. Část tzv. živých matrik byla předána příslušným orgánům samosprávy, matriky starší 100 let (v případě narození) a 75 let (matriky oddaných a zemřelých) byly předány státním archivům.

Matriky jsou jedněmi z nejpečlivěji chráněných úředních knih, potažmo archiválií, které vždy zasluhovaly zvláštní pozornost. Již v 18. století lidé chápali, že, řečeno dnešní terminologií, takto významná data musejí mít záložní kopie, takže v minulosti vedly farní úřady druhopisy, někdy i triplikáty matričních záznamů. V našem archivu máme dochovány nejstarší druhopisy z konce 18. stol. Na matričních úřadech nejnižší instance jsou i v současnosti vedeny matriční knihy v prvopisech a druhopisech. Druhopisy jsou předávány obvykle centrálně do úschovy pověřeným samosprávným úřadům vyšší instance, kde jsou uloženy i starší druhopisy, kterým dosud neuplynula skartační lhůta 100 respektive 75 let pro předání do příslušných státních oblastních a zemských archivů. Prvopisy se předávají archivům po uplynutí skartační lhůty, některé archivy je vyžadují např. v pětiletých cyklech, u nás přejímáme každoročně – cca 60-80 svazků za celý bývalý Západočeský kraj, nyní kraj Karlovarský a Plzeňský. Druhopisy musejí být z bezpečnostních důvodů i v archivech uloženy v oddělených depozitářích – v našem případě v Klášteře u Nepomuka. Prvopisy matrik jsou uloženy v Plzni.

Po převzetí matrik státními archivy v roce 1949 byla samozřejmě prioritou ochrana jejich fyzického stavu. Staly se archiváliemi, které si prioritně zasluhují ochranu konzervátorů. Ale písemnostem nesvědčí ani časté používání – předkládání badatelské veřejnosti. Jako první technologické řešení se objevilo snímkování na mikrofilmy. Pak byl zájemci o studium předložen pouze mikrofilm na čtecím přístroji a originál zůstal ochráněn.

Moderní technologie 21. století poskytly do té doby netušené možnosti – digitalizaci matričních knih ve všech archivech v České republice. Digitalizace matrik začala v ČR cca v letech 2005-2007. Nejprve byla digitalizace považována pouze za formu ochranné kopie podobně jako mikrofilm. Reálné výsledky přinesla během roku 2008 - jako první předložil veřejnosti digitální prohlížeč matrik SOA v Třeboni. Nyní již jsou zveřejněny digitální kopie naprosté většiny matrik narozených, oddaných a zemřelých všech církví i matriky civilní od 16. do 20. století. Toto dílo čítá 70 tisíc svazků s více než 6 miliony listů. Pro zájemce o vyhledávání vlastních předků je to úžasný zdroj informací, ale především je to situace dobrá pro samotné matriční knihy, které zůstávají uloženy v optimálních podmínkách archivních depozitářů. Množství digitalizovaných matrik není konečné – každý rok uplyne skartační lhůta pro další knihy – je to 100 let pro knihy narozených a 75 pro knihy oddaných a zemřelých, takže letos přešly do péče archivů knihy z let 1913-14 a 1938-1939. Prohlížeče matrik tak musejí být živými databázemi, které musí být schopny průběžné aktualizace.

Dlouhodobé projekty digitalizace matrik většinou iniciovala americká genealogická společnost (Family Search, což je organizace zřízená církví The Church of Jesus Christ of Latter-day Saints, známé také jako Mormoni), která nabídla archivům techniku, pracovníky i

zpracování dat. Navázané projekty zpřístupnění matrik byly prováděny samostatně po oblastních archivech a v různém rozsahu podle možností a záměru jednotlivých archivů. Nepodařilo se, tak jako třeba na Slovensku v roce 1991, dosáhnout generální dohody mezi státem a mormonskou církví o systematickém kopírování archiválií. Část českých a moravských archivů projekty digitalizace řídí zcela ve vlastní režii bez spolupráce s mormonskou církví.

Archivů v ČR, které disponují matrikami, je 10: 5 státních oblastních archivů v Čechách, 2 zemské archivy na Moravě a Archiv hlavního města Prahy. Židovské matriky jsou soustředěny v Národním archivu a vojenské matriky, které byly vedeny jednotlivými pluky pro vojáky a jejich rodinné příslušníky, jsou jednak v oblastních a zemských archivech spolu s církevními matrikami, jednak ve Vojenském historickém archivu a Národním archivu. Digitální snímky matrik nejsou v České republice zpřístupněny jednotným systémem.

Jednu z databází, ve které jsou matriky zpřístupněny veřejnosti, vyvinula firma Bach systems s.r.o. Archivy, které ji používají, si ji nechaly částečně modifikovat podle svých aktuálních potřeb – jedná se o SOA v Litoměřicích, Zemský archiv v Opavě a Archiv hlavního města Prahy.

Přístup k matrikám těchto archivů je zajištěn přes jejich webové stránky:

Litoměřice: http://www.soalitomerice.cz/content/matriky,

Opava:http://vademecum.archives.cz/vademecum/searchlink?fcAr=217000010-0&fcAr=217000010-1&fcDb=10041

Praha: http://www.ahmp.cz/index.html?wstyle=2&catalogue=1

Východočeský oblastní archiv v Zámrsku také spolupracuje s mormonskou církví, Sbírku matrik východočeského kraje je možné studovat buď na webu Family Search International (https://familysearch.org), nebo si snímky stáhnout do svého počítače ve formě souboru ZIP, v němž jsou zabaleny reprodukce všech popsaných listů příslušné matriky ve formátu JPG – pro prohlížení pak doporučují freeware XnView.

Zámrsk: http://vychodoceskearchivy.cz/zamrsk/matriky/

Židovské matriky z Národního archivu v Praze jsou přístupné na adrese http://www.badatelna.eu/fond/1073 . Portál Badatelna je produktem firmy LemonBone s.r.o a je společný pro Národní archiv, Památník národního písemnictví a Knihovnu Národního muzea.

Státní oblastní archiv v Třeboni je známý svým vlastním prohlížečem: https://digi.ceskearchivy.cz/ . Byl první v české republice a stále se vylepšuje.

Pro společný prohlížeč digitalizovaných matrik se rozhodly Státní oblastní archiv v Praze a Moravský zemský archiv v Brně. Je přístupný na adrese http://actapublica.eu/index.php. Tento projekt vznikl s podporou Evropské unie v rámci Projektu přeshraniční spolupráce"Cíl Evropská územní spolupráce" Rakousko - Česká republika 2007-2013. Zmíněné Rakousko je

důležité – na tomto prohlížeči lze studovat též matriky z Dolnorakouských archivů – konkrétně Bischöfliches Archiv St. Pölten.

Do prosince roku 2013 používal tento prohlížeč i náš SOA v Plzni. Nyní poskytujeme digitalizované archiválie včetně sbírky matrik na portálu http://www.portafontium.eu/

Ačkoli jsou prohlížeče ne zcela jednotné, zájemce o studium v nich nalezne všechny potřebné údaje – základní popis matriční knihy, její časový rozsah, seznam obcí, jichž se matrika týká a především samotné digitální snímky. Pro jejich prohlížení je samozřejmě vhodné rychlejší a stabilnější připojení k internetu, pro komfortnější vizuální dojem velký monitor.

Záložní kopie všech mormony nasbíraných dat putují do skalního trezoru v Salt Lake City, ale jsou v dalších třech exemplářích na dalších místech. Matriky jsou také paralelně uloženy na stránkách https://familysearch.org, kde může badatel nalézt i data z dalších zemí. Nevýhodou pro české badatele, zvyklé na systém členění matrik v rámci farností, je menší přehlednost americké databáze a těžší orientace při vyhledávání. Situace v ČR, co se týče zálohy dat, není tak dramatická, skalní trezor prozatím nemáme, ale samozřejmě se zálohování digitálních snímků věnuje v jednotlivých archivech zvýšená pozornost.

Digitalizační projekty ve Státním oblastním archivu v Plzni

S digitalizací archiválií započal Státní oblastní archiv v Plzni na přelomu roku 2004 a 2005. V té době došlo na pracovišti v Plzni k nákladné rekonstrukci a kompletnímu novému vybavení konzervátorských dílen. K digitalizaci a současně k restaurování byly vybrány nejstarší listiny z fondu Premonstráti Teplá. Mezi těmito listinami se nachází archivní kulturní památka, tzv. Hroznatova závěť z roku 1197, která je nejstarší šlechtickou listinou v České republice. Prvním výstupem digitalizačních prací bylo CD s těmito listinami a bohatým doprovodným materiálem včetně hudebního, které bylo vydáno při příležitosti krátkodobé výstavy originálů listin, která se konala v srpnu 2005 v klášteře v Teplé. Po tepelských listinách byly postupně digitalizovány další listiny z fondů velkostatků a měst. Ve spolupráci s rakouským partnerem z biskupského archivu v St. Pölten dochází k postupnému zveřejňování listin na portálu Monasterium. Digitalizace listin stále pokračuje.

Ještě v roce 2005 archiv započal s digitalizací matričních knih jako nejvyužívanějšího a tím i nejohroženějšího materiálu. Digitální výstupy byly zamýšleny pouze pro interní potřebu, pro předkládání v badatelnách archivu místo originálů matričních knih. Se zveřejněním na internetu se ještě nepočítalo. Digitalizace matrik trvala od roku 2005 do roku 2013, současně s digitalizací se matriky i podrobně popisovaly. V současné době jsou již digitalizovány všechny matriční knihy, probíhají opravy méně kvalitně nafocených knih z počátků projektu.

Od roku 2010 byla zahájena digitalizace dalšího velmi žádaného materiálu – kronik všeho druhu. I tento projekt je již dokončen, stejně jako projekt na digitalizaci fotografií. Pro badatele v oblasti lázeňství jsou velmi zajímavé tzv. Kurlisty, seznamy lázeňských hostů zejména z Karlových Varů, které se dochovaly od konce 17. století. Zejména genealogičtí badatelé čekají na zveřejnění sčítacích operátů z období před rokem 1900, policejních přihlášek k pobytu pro Plzeň po roce 1919 a urbářů a pozemkových knih velkostatků a měst.

Tento materiál je již rovněž digitalizován, ovšem vzhledem k jeho velkému množství budou projekty trvat ještě řadu let.

Záhy po roce 2005 se ukázalo, že udržet digitalizovaný materiál jen pro interní archivní použití není možné. Různé digitalizační projekty po celém světě i v Česku, zejména velmi úspěšný projekt Státního oblastního archivu v Třeboni, způsobil značný tlak na vedení našeho archivu, které nakonec rozhodlo o otevření databází široké veřejnosti. Již při přípravě CD k výstavě v Teplé archiv zahájil spolupráci se soukromou plzeňskou firmou, se kterou byl vypracován rozsáhlý plán projektu zveřejňování archiválií. Prvním krokem po stanovení strategie zveřejňování bylo vytvoření podrobné databáze obcí Plzeňského a Karlovarského kraje. Než však došlo k zpřístupnění prvních výsledků tohoto projektu, proběhla v archivu v Plzni v roce 2009 personální výměna na funkci ředitele a spolupráce s firmou byla ukončena. Ve stejné době zahájily další archivy v ČR své vlastní nezávislé digitalizační projekty. Moravský zemský archivem v Brně a SOA Praha spolupracovaly se stejnou firmou, měly tedy databáze původně ze stejného zdroje, a rovněž v přibližně stejné době spolupráci s firmou ukončily. V roce 2009 jsme uskutečnili několik cest po jednotlivých oblastních archivech, kde jsme hledali inspiraci a případně i možnou spolupráci. Ukázalo se, že nejschůdnějším a nejrychlejším řešením bude spolupráce s Moravským zemským archivem v Brně, který v rámci projektu Acta publica použil původní databázovou strukturu z období spolupráce s plzeňskou firmou. Na počátku roku 2010 jsme tedy na základě smlouvy přistoupili jako hosté do brněnské databáze, kde byla do konce roku 2010 zveřejněna asi polovina našich matrik. Se zapojením našeho archivu byly různé problémy, zejména s částečnou nekompatibilitou naší databáze obcí, která u nás vycházela ze starší verze z roku 2008, s databází Moravského zemského archivu v Brně. Tato nekompatibilita způsobovala například nespolehlivé vyhledávání menších lokalit, vyhledávat bylo nakonec možno pouze podle názvů původců. Přesto jsme do konce roku 2013 uveřejnili přes Acta Publica všechny naše matriky (9). Současně jsme však pokračovali ve vývoji vlastního prohlížeče, zejména proto, že jsme chtěli zveřejňovat i další již zdigitalizované archiválie, což Acta publica neumožňovala.

V roce 2010 byla zahájena mezinárodní spolupráce s bavorskými státními archivy na projektu Bavorsko-české síti digitálních historických dokumentů Porta fontium. Web byl slavnostně představen 8.4.2011 ve Waldsassenu a dnes je již bohatě naplněn archivními dokumenty, hojně navštěvován a snad i vděčně přijímán badatelskou veřejností. K poslednímu prosinci roku 2013 byla současně ukončena naše spolupráce s projektem Acta publica a matriční knihy se zcela přesunuly na web Porta fontium.

Značnými změnami prošly v našem archivu prostředky digitalizace. Na samém počátku jsme digitalizovali staré zabezpečovací mikrofilmy pomocí speciálního, poměrně drahého přístroje, který byl pro tuto práci přímo určen. Výstupy však byly bohužel, zejména kvůli nekvalitním originálním mikrofilmům, velmi nekvalitní a v některých případech skoro nepoužitelné. Výsledný snímek byl v každém případě černobílý, nebylo ho možno dostatečně zvětšit, byl málo ostrý a zrnitý. Paralelně jsme dosud nezmikrofilmované matriky fotografovali digitálním fotoaparátem. Kvůli omezení finančních prostředků nešlo o nijak špičkový fotoaparát, výsledky však byly poměrně dobré. Kvalita snímků závisela hodně na pečlivé práci obsluhy. Pro zveřejnění na webu Acta publica se však snímky zmenšovaly a "odbarvovaly". Zvolená

technologie zobrazování totiž i při potřebě pouhého náhledu na snímek nebo jeho část načítala do prohlížeče kompletně celý snímek, takže přes internetové připojení byla rychlost zobrazování velmi nízká. Druhý fotoaparát, který jsme zakoupili po poruše prvního, byl bohužel velmi málo kvalitní a výsledky byly podstatně horší, než s prvním fotoaparátem. Teprve v prosinci 2010 zakoupil archiv knižní skener formátu A2, jehož výstupy mají konečně velmi pěknou a stabilní kvalitu i při změnách operátorů a brigádní práci. Po nafocení všech matričních knih jsme přistoupili k opravám starších nekvalitních snímků, zejména černobílých, které postupně na Portě vyměňujeme.

Druhé velké digitalizační pracoviště v kompetenci Státního oblastního archivu v Plzni se dvěma knižními skenery je ve Státním okresním archivu v Chebu. Zde se skenují listiny z okresních archivů, kroniky, periodika z 19. století i ty matriky, které se nám pro své větší rozměry nevejdou do skeneru v Plzni. Chebské skenery mají totiž rozměry stolů A1 a A2. Některé okresní archivy skenují vlastními silami neknižní archiválie, např. sčítací operáty a fotografie.

Data jsou z primárních diskových polí zálohována na další disková pole a na pásky. Některá data jsou též zároveň umístěna v geograficky oddělených lokalitách – v Plzni a Chebu.

Systém zobrazování snímků je založen na rozšířených open-source technologiích, které jsou využívány např. v knihovnách. K zveřejňování tak používáme zdrojové snímky ve formátu JPEG 2000 v pyramidové struktuře, který umožňuje prostřednictvím image serveru vytvářet dlaždice ve formátu JPEG s různými velikostmi a přiblíženími snímku nebo jeho části. Do prohlížeče se tak nepřenáší celý snímek, ale pouze ta část, která odpovídá uživatelem právě zobrazované velikosti a přiblížení snímku. Tím je zajištěn dostatečný uživatelský komfort a rychlost prohlížení, a to i při nižších rychlostech internetového připojení.

Použitý obrazový server umožňuje komunikovat s klienty v různých formátech (např. Zoomify nebo Deepzoom). V současné době používáme na webu Porta fontium pro zobrazování v prohlížeči formát IIP.

Pro zobrazování a ukládání metadat používáme rozšířené open-source databáze a systémy pro správu obsahu, pro které vyvíjíme vlastní moduly.

Ohlasy ze strany badatelů nás přesvědčují, že digitalizace archiválií, a především matrik, byl krok správným směrem. Musíme však nadále sledovat vývoj technologií a naše prohlížeče digitalizovaných archiválií jim přizpůsobovat. Pro archiválie samotné představuje zveřejnění pomocí digitálních kopií velmi dobrý způsob ochrany, mohou být uloženy v optimálních podmínkách depozitářů a nejsou poškozovány častým předkládáním.