

Digitální archiv české archeologie

Specializované archivy mají ve struktuře archeologie jako oboru nezastupitelné místo. Při terénním výzkumu totiž původní archeologický kontext zpravidla zaniká a terénní dokumentace uložená v archivech pak pro další archeologický výzkum hraje roli primárního pramene.

■ Martin KUNA

Archeologický ústav AV ČR,
Praha, v. v. i.

Archeologické archivy a digitalizace

Na výše uvedeném tvrzení nemění samozřejmě nic skutečnost, že vlastně nejde o „archivy“ ve smyslu archivního zákona, nýbrž pouze o specializované sbírky terénní dokumentace archeologických výzkumů. V zásadě nevadí ani fakt, že pojem „archeologický archiv“ zahrnuje u nás pouze dokumentaci terénního výzkumu, zatímco v mnoha jiných zemích se do tohoto pojmu zahrnují i nálezy samy (Brown 2007). Možná je druhý přístup poněkud výhodnější, protože zdůrazňuje, že nálezy bez průvodní dokumentace ztrácejí většinu své odborné hodnoty a je tudíž nezbytné, aby byly uchovávány v pevném spojení s popisem nálezových situací.



■ Obr. 1. Archiv ARÚP, úsek nálezových zpráv. V popředí osm banánových krabic obsahujících dokumentaci z jednoho archeologického projektu – toto množství ilustruje stále rostoucí nároky na dokumentaci terénních situací. Foto M. Kuna.

Digitálním archeologickým archivem rozumíme v tomto příspěvku sbírku terénní dokumentace v podobě digitálních dat. V současné době mají digitální prvky terénní dokumentace archeologických výzkumů různou podobu. V zásadě jde buď o digitální ekvivalenty existujících analogových dokumentů (např. skenované výkresy, tabulky, texty apod.), nebo o digitální data ve formátech, které umožňují ve srovnání s analogovým dokumentem kvalitativně vyšší způsoby zpracování (např. textové soubory, databáze, vektorové kresby apod.). Některé části terénní dokumentace jsou dnes pořizovány primárně v digitální podobě a jejich všeobecný převod do analogového tvaru je již neúčelný až nemožný (srov. např. množství digitálních fotografií z jednoho terénního výzkumu, složitost výstupů geografických informačních systémů apod.).

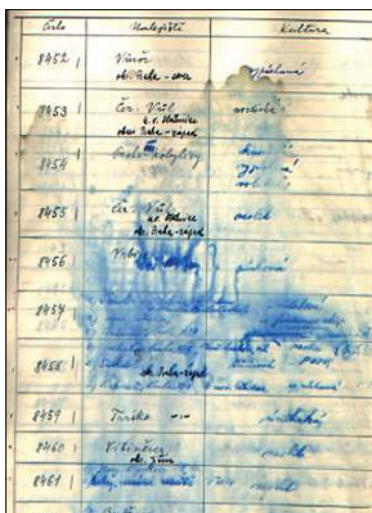
Podíl primárně digitálních („digital-born“) dat na terénním výzkumu stále stoupá (srov. Bibby 2010) a zřejmě není daleko doba, kdy nejen fotografie, ale kresby plánů a řezů, seznamy nálezů atd. budou přímo v terénu pořizovány elektronickými prostředky. Z hlediska archivace těchto dat stojí v současné době archeologie na rozcestí a je třeba si přiznat, že rozhodnutí o přechodu k plně a pouze digitálním archivům je pro většinu dnešních archeologů nesnadné. S převážně digitálními archivy (bez analogových ekvivalentů) je stále spojeno příliš mnoho otazníků – zdá se ovšem, že rychlé pronikání digitalizace do všech oblastí archeologické práce dané dilema nejspíše v krátké době rozhodne za nás. Jde tedy především o to, aby archeologie byla na takovou změnu připravena, to znamená, aby existovaly kvalifikované, z hlediska budoucnosti perspektivní a archeologickou obcí všeobecně sdílené zásady pořizování digitálních dat a následné péče o ně. Na základě dosavadních zkušeností lze tvrdit, že archivace digitálních dat je úkol nesrovnatelně složitější a nákladnější než archivace dat analogových, jakkoli by se na první pohled mohlo zdát, že tomu je naopak. Kvalitativně nové možnosti šíření a analýzy

dat v digitální podobě jsou však natolik žádoucí a z hlediska budoucnosti perspektivní, že jiné řešení, než věnovat nemalé prostředky na digitalizaci dat a průběžnou péči o ně, zřejmě neexistuje.

Archiv Archeologického ústavu v Praze

Archeologický ústav AV ČR v Praze (ARÚP) vytvořil za 90 let své existence rozsáhlý soubor terénní dokumentace (obr. 1). Jeho fondy obsahují dokumentaci především vlastních terénních výzkumů, kterých od 20. let minulého století ústav provedl stovky, v případě menších záchranných akcí tisíce. Kromě ní ale archiv obsahuje i řadu dalších dokumentů, např. hlášení o náhodných nálezech, rozpisů muzejních sbírek a soupisy nálezů podle druhů a období. Jednou z podstatných složek archivu ARÚP jsou také hlášení a kopie nálezových zpráv o výzkumech ostatních organizací na území Čech. Tyto dokumenty ARÚP shromažďuje na základě koordinační a garanční role v české archeologii, kterou mu předepisuje Památkový zákon. Výsledkem dlouholeté a dnes stále intenzivnější činnosti je soubor 3,5 tis. rozsáhlých nálezových zpráv, 65 tis. kratších zpráv, hlášení a expertních posudků, 220 tis. negativů a diapositivů (včetně snímků z leteckého průzkumu) a 10 tis. map a plánů. I když tento korpus zdaleka není ve vztahu ke známému archeologickému dědictví Čech úplný, představuje nejrozsáhlejší dostupný soubor dat o pravěkých až raně novověkých památkách našeho území a jako takový je základním zdrojem informací pro odborníky, studenty, projektanty i zájemce o kulturní památky. V širším slova smyslu lze tvrdit, že tento soubor patří k nejcennějším součástem našeho archeologického dědictví.

Dramatickou, ale významnou změnu v činnosti archivu ARÚP přinesla povodeň Prahy v r. 2002. Archiv sice nebyl postižen do té míry jako knihovna ARÚP; zejména jeho nejcenější část, textové dokumenty, byla a je uložena v patře, mimo dosah tehdejší povodňové vlny. Škody



■ Obr. 2. Poškozené inventární knihy fotoarchivu ARÚP po povodni v r. 2002. Foto M. Kuna.

způsobila povodeň především ve sbírce negativů a plánů, které byly zničeny nebo poškozeny zhruba ze třetiny celkového množství (obr. 2). Daná situace ovšem znamenala nejen ztrátu, ale i výstrahu, výzvu a příležitost k zásadní modernizaci archivu. V důsledku tohoto impulsu došlo během následujících několika let k rekonstrukci budovy, novému uložení archivních fondů (obr. 3) a realizaci zásadního kroku v ochraně a zpřístupnění dat, totiž kompletní digitalizaci archivu.

Program digitalizace archivu ARÚP

Digitalizace archivu ARÚP – zahájena ihned po povodni r. 2002 – sledovala tři hlavní cíle: (a) zachránit obsah dokumentů, které byly povodní poškozeny; (b) preventivně chránit obsah archivu před následky další podobné události, která by v budoucnu mohla nastat; (c) usnadnit výměnu informací v české archeologii. Digitalizace byla provedena skenováním dokumentů do obrazových formátů (*.tif, *.pdf) a v letošním roce byla víceméně dokončena. Zároveň s digitalizací analogových dat byly zpracovávány základní principy pro archivaci primárně digitálních dat, která v současné době začínají v mnoha oblastech dokumentace převládat (fotografie, databáze, vektorová grafika aj.) – na tomto poli však do budoucna zbývá ještě mnoho práce.

Digitální ekvivalenty analogových dokumentů a nová, primárně

digitální data v archivu ARÚP tvoří úhrnem Digitální archiv ARÚP. Počet digitálních souborů v archivu dosahuje cca 300 tis., počet jednotlivých stránek může být odhadnut až na milion, stávající objem dat činí přes 1 TB. K uložení dat je využíván jukebox Plasmon M104–910–X s kapacitou 104 magnetooptických disků (obr. 4), dále disková pole v počítačové síti ARÚP, přenosné pevné disky aj. Pro správu datového souboru a zejména pro vyhledávání dokumentů byla vytvořena počítačová aplikace Digitální archiv ARÚP, která umožňuje přístup k dokumentům prostřednictvím internetu. Aplikace byla zhotovena firmou NESS Czech s.r.o. podle návrhu ARÚP v r. 2006, a to na bázi softwaru *Documentum* americké firmy EMC, spadajícího do okruhu tzv. „systémů pro správu obsahu“ (*content management system*).

Digitální archiv ARÚP

Softwarová aplikace Digitální archiv umožňuje uchovávat, vyhledávat, prohlížet, třídit a editovat digitální dokumenty archivu ARÚP, exportovat popisné údaje o nich a získávat plné kopie digitálních dat (obr. 5), a to i z počítače vzdáleného uživatele. Do aplikace lze vstoupit z webových stránek ARÚP (http://www.arup.cas.cz/cz/archiv/archiv_digitalni.html); uživatelská příručka je k dispozici tamtéž, ale existuje i v tištěné podobě (Kuna et al. 2007; obr. 6). Touto cestou se uživatel dostane k popisným údajům



■ Obr. 3. Nová místnost archivu negativů, leteckých snímků a plánů v budově Archeologického ústavu v Praze. Foto M. Kuna.



■ Obr. 4. Jukebox Plasmon M104–910–X na 104 magnetooptických disků, jedno ze zálohovacích zařízení pro Digitální archiv ARÚP.

o dokumentech v Digitálním archivu (může zjistit, zda dokumenty v archivu existují a jejich základní vlastnosti), pro přístup k vlastnímu obsahu dokumentů je třeba ještě požádat administrátora systému v ARÚP o přidělení přístupového hesla. Tyto služby i obsah dokumentů jsou spolupracujícím organizacím poskytovány zdarma.

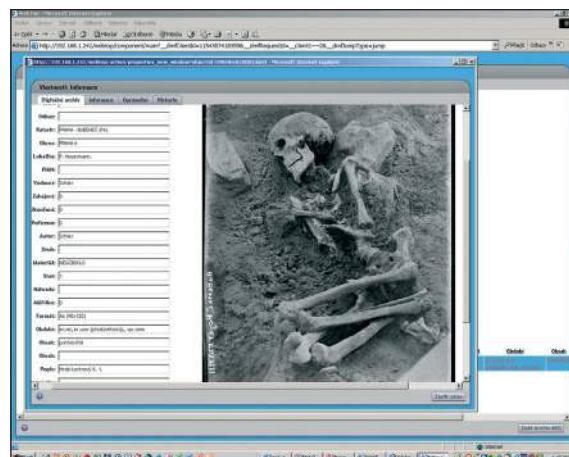
Digitální archiv svou datovou strukturou vychází z historicky vzniklého členění nálezové dokumentace. Skládá se z 12 skupin dokumentů, které mají vlastní číselné řady (obr. 7). Jde o negativy s terénními snímky (řada FT), snímky předmětů (FP), diapozitivy (FD), letecké negativy (LN), letecké diapozitivy (LD), primární digitální snímky terénu (DT) a předmětů (DP), digitální letecké fotografie (DL), jiné digitální soubory (databáze, vektorové kresby; DJ), plány (PY), nálezové zprávy (TX) a expertní posudky (TP). Každému dokumentu odpovídá záznam v databázi, shrnující údaje o dokumentu ve více než 20 polích. Popis zahrnuje informaci o druhu dokumentu, datu pořízení, autorovi, lokalitě, druhu archeologického nálezu apod.; této struktuře odpovídá i struktura vyhledávacího menu v internetové aplikaci. Vytváření popisné databáze k dokumentům je ovšem mnohem náročnější než samo skenování, a z tohoto důvodu není velká část dokumentů ještě popsána úplnými údaji. Nové dokumenty (digitální i analogové) jsou do archivu ARÚP přijímány pouze s kompletně vyplněnými údaji

databáze; příslušná databázová aplikace (Databáze dokumentů; **obr. 8**) je k dispozici na webových stránkách ARÚP a představuje komplexní nástroj ke správě souborů dokumentů i mimo Digitální archiv (*Křivánková – Kuna 2009*). Revize a doplnění údajů o starších dokumentech bude nutně trvat ještě několik let. Z tohoto důvodu v Digitálním archivu zatím nejsou uloženy všechny soubory, i když už byly v úplnosti digitalizovány (srov. **obr. 7**).

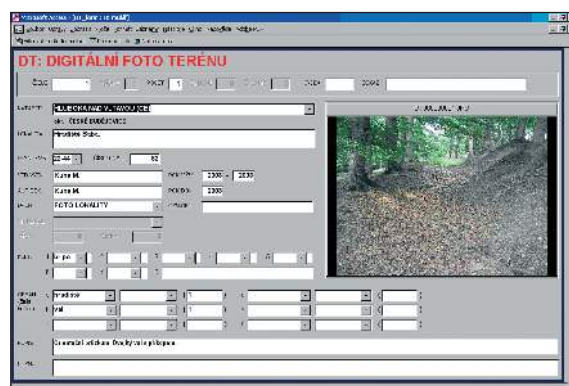
Nové nálezové zprávy se v Digitálním archivu uchovávají ve formátu *.pdf. Do tohoto formátu je převádějí sami jejich autoři (pokud není se správci archivu dojednáno jinak), a to spojením textových, obrazových a tabulkových dat do jednoho souboru. Starší nálezové zprávy byly digitalizovány do formátu *. (multi)tif, v budoucnu se počítá s jejich převodem také do *.pdf. U digitálních obrazových (rastrových) dokumentů jsou povoleny běžné formáty *.jpg a *.tif. Digitální fotografie jsou předávány v kvalitě, v jaké byly pořízeny, starší negativy, pozitivy či diapozitivy byly skenovány do formátu *.tif a velikosti 210×150 mm při rozlišení 300 dpi. Plány se skenují ve formátu 1:1, a to jako *.tif s 300 dpi (originál A0+ až A2), resp. 600 dpi (A3 a menší); LZV komprese je povolena. S archivací databázových souborů a vektorové grafiky dosud v ARÚP nejsou zkušenosti; podle současných znalostí však lze doporučit formát *.dgn pro vektorovou grafiku a samostatnou archivaci jednotlivých databázových tabulek ve formátech *.txt nebo *.csv. U uvedených formátů lze podle našeho názoru očekávat největší „trvanlivost“ a nejsnazší migraci do vyšších verzí softwaru v budoucnosti. V této souvislosti lze připomenout fakt,

že Digitální archiv má sloužit především k dlouhodobému uchování primárních dat, nikoliv jako „příruční“ sklad dat, která jsou průběžně používána a obměňována.

Přístup k dokumentům v archivu ARÚP i Digitálním archivu je diferencován, přičemž Digitální archiv v zásadě respektuje dlouhodobě zavedenou praxi v archivu ARÚP, autorský zákon i nutnost ochrany některých informací. Přístup k dokumentům v analogové i digitální podobě se řídí příslušnou směrnicí ředitele ARÚP (viz webové stránky ARÚP). Většina dokumentů je přístupná bez omezení (hlášení a krátké nálezové zprávy, náhledy fotografií a plánů, letecké snímky památek apod.), menší část jen se souhlasem ředitele ARÚP, příp. vedoucího výzkumu. Toto omezení se týká zejména nálezových zpráv z nových výzkumů velkého rozsahu; u obzvláště citlivých dat (např. letecké snímky zachycující polohu dosud neznámých lokalit) nejsou zpřístupněny ani náhledy. Smyslem tohoto opatření je omezit přístup k informacím těm, kteří je hodlají zneužít, zejména rabováním archeologických lokalit pomocí detektorů kovů. Celkově se ovšem utajování informací týká jen malé části datového souboru.



■ **Obr. 5.** Aplikace Digitální archiv ARÚP – stránka s náhledem dokumentu a popisnými údaji.



■ **Obr. 8.** Databáze dokumentů pro Digitální archiv – stránka z aplikace sloužící k popisu dokumentů.



■ **Obr. 6.** Obálky vydaných uživatelských příruček k počítačovým aplikacím souvisejícím s archivem ARÚP. Foto M. Kuna.

Skupina		Řada		Počet dokumentů		
				celkem (ks)	digitalizo- váno (%)	v Digitálním archivu (%)
F	fotoarchiv	FT	terénní snímky (negativy)	110000	100	90
		FP	fotografie nálezů (negativy)	80000	100	0
		FD	diapozitivy (terén i předměty)	15000	100	0
L	úsek letecké archeologie	LN	letecké snímky – černobílé negativy	10000	0	0
		LD	letecké diapozitivy (barevné)	5200	100	100
P	úsek map a plánů	PY	plány, mapy, měřičská dokumentace	10000	100	0
T	úsek náleзовých zpráv	TX	hlášení a krátké zprávy	58000	100	100
			nálezové zprávy (rozsáhlé)	3500	100	0
		TP	konzervační zprávy	2000	100	0
			expertní posudky	1500	100	0
D	úsek primárně digitálních dat	DT	digitální terénní fotografie	probíhá sběr	-	0
		DP	digitální fotografie předmětů	probíhá sběr	-	0
		DL	digitální letecké snímky	probíhá sběr	-	0
		DJ	databáze, tabulky, vektorové kresby a mapy	probíhá sběr	-	0
Celkem (bez skupiny D)				295200	96,6	54,9

■ **Obr. 7.** Základní skupiny dokumentů archivu ARÚP, jejich celkové počty, stav digitalizace a vložení do aplikace Digitální archiv (ke konci r. 2010). Sestavil M. Kuna a D. Křivánková.

Digitální archiv v kontextu informačního systému české archeologie

Digitální archiv ARÚP je pouze jedním z pilířů archeologického informačního systému, o jehož vybudování ARÚP dlouhodobě usiluje. Podle koncepce ARÚP by takový systém měl mít tři základní součásti: (1) databázi archeologických akcí – jednotek terénní činnosti, (2) databázi prostorových referenčních jednotek (PIAN – prostorová identifikace archeologických nálezů); (3) databázi a soubor (digitálních) jednotek terénní dokumentace, tj. jednotlivých dokumentů v archivu.

Do r. 2002 byla pozornost pracovníků ARÚP věnována především prvním dvěma tématům, spojeným v tzv. Archeologické databázi Čech (ADC). ADC je v ARÚP vedena již od první poloviny 90. let a tvoří ji především revidované Zprávy o archeologické akci, které jsou ústavu poskytovány ostatními oprávněnými organizacemi, ale i záznamy excerpané z jiných zdrojů (Kuna a kol. 1995; Kuna – Křivánková 2006; Herichová a kol. 2009). V loňském roce byl informační systém v ohledu terénních akcí zásadně inovován, a to zavedením Internetové databáze terénních archeologických výzkumů, k níž v současné době přistupují další organizace (Kuna a kol. 2009). Cílem inovace je zajistit včasnou (pokud možno předběžnou) informovanost o přípravě a/nebo zahájení terénního výzkumu – do budoucna se počítá s tím, že na záznam v této databázi bude přímo navazovat ADC. Prostorová

identifikace (tzv. PIAN) je v ARÚP od 80. let zpracovávána s velkou pečlivostí, a to prozatím na tištěných mapách ZM 1:10 000; již několik let také na intranetu ARÚP běží geografický informační systém, kde jsou data PIAN i s informacemi ADC zpřístupněna.

Hlavním problémem komplexního informačního systému, a to jak směrem do budoucna, tak směrem ke zpracování starších dat, je identifikace vzájemných vazeb mezi jednotkami tří základních datových souborů. Mezi objekty jednotlivých databází totiž vždy existují vztahy typu „n–n“. Je to důsledkem toho, že např. jedna terénní akce často proběhla ve více prostorových jednotkách, ale zároveň se v jedné a téže prostorové jednotce mohlo postupně odehrát několik různých terénních akcí. Totéž platí i ve vztahu k dokumentům, které se mohou týkat několika archeologických nálezů (to platí zejména u dokumentů starších, u nových dokumentů by se tomu autoři měli vyhýbat), přičemž každá akce vytváří samozřejmě sérii jednotlivých dokumentů. Celková revize informačního systému a popsání těchto vazeb mezi desítkami tisíc záznamů v různých databázích je úkolem, který bude pochopitelně vyžadovat velké nasazení a zvláště finanční prostředky – v návrzích ARÚP směrem k různým grantovým agenturám se s tímto úkolem počítá a je pro něj už připravena patřičná metodika.

Závěr

Digitální archiv ARÚP má za sebou první etapu své existence, která spočívala v digitalizaci starých fondů a vytvoření prostředí pro jejich elektronické zpřístupnění. „Popovodňový“ program digitalizace v loňském roce skončil, do budoucna však pro archiv ARÚP zůstává řada dalších úkolů. Na jedné straně jde o personální a technické zajištění dalšího provozu Digitálního archivu, na straně druhé o koncepční otázky spojené se sběrem a správou digitálních dat. Je třeba zajistit digitalizaci dalších dokumentů či jejich částí, které do archivu průběžně přicházejí v analogové podobě, řešit otázku vhodných formátů a jejich budoucí migrace do nových verzí softwarových produktů, dokončit indexaci starších

dokumentů a propojit Digitální archiv s dalšími informačními zdroji v ARÚP.

Literatura

- Bibby, D. 2010: Mit GPS, Laserstrahl und Mausclick: zur Digitalisierung moderner Ausgrabungstechnik. In: Bofinger, J. – Merkl, M. (eds.), Mit Hightech auf den Spuren der Kelten. Begleitheft zur gleichnamigen Sonderausstellung, Esslingen: Landesamt für Denkmalpflege Baden-Württemberg, 90–98.
- Brown, D. H. 2007: Archaeological archives. A guide to best practice in creation, compilation, transfer and curation. London: Archaeological Archives Forum.
- Herichová, I. a kol. 2009: Herichová, I. – Hrdlička, L. – Křivánková, D. – Kuna, M. 2009: Databáze archeologických dokumentačních bodů. Rozšiřující modul Archeologické databáze Čech v. 3.1. pro území Prahy. Uživatelská příručka. Praha (ARÚP).
- Křivánková, D. – Kuna, M. 2009: Databáze dokumentů pro Digitální archiv Archeologického ústavu v Praze. Uživatelská příručka, Praha (ARÚP).
- Kuna, M. a kol. 1995: Kuna, M. – Křivánková, D. – Krušinová, L. 1995: ARCHIV 2.0. Systém Archeologické databáze Čech. Uživatelská příručka, Praha (ARÚP – SÚPP).
- Kuna, M. a kol. 2007: Kuna, M. – Bžatek, M. – Křivánková, D. – Kundera, M. – Mazancová, M. 2007: Digitální archiv Archeologického ústavu v Praze. Uživatelská příručka. Praha (ARÚP).
- Kuna, M. a kol. 2009: Kuna, M. – Jindáček, M. – Křivánková, D. – Tomášek, T. – Vlašný, M. 2009: Internetová databáze terénních archeologických výzkumů, v. 1.0. Uživatelská příručka. Praha (ARÚP).
- Kuna, M. – Křivánková, D. 2006: ARCHIV 3.0. Systém Archeologické databáze Čech. Uživatelská příručka, Praha (ARÚP).

Summary

A digital archive of Czech archaeology

Specialized archives hold a highly valuable position in the discipline of archaeology. As the original archaeological context is typically lost during terrain excavations, terrain documentation stored in archives is a primary source of information for further archaeological research. During the 90 years of its existence, the Institute of Archaeology of the Academy of Sciences of the Czech Republic has systematically built an archive of terrain documentation containing data from the institute's own terrain excavations as well as documentation from the excavations conducted by the other organisations working in Bohemia. After part of the archive was damaged in the 2002 floods, the decision was made to convert all documentation into digital form. The actual digitalization of the archive (ca. 70,000 find reports, 220,000 negatives and slides, 10,000 maps) was essentially completed in 2010, and this data is gradually being transferred to the Digital Archive web application (see <http://digitalarchive.arup.cas.cz/webtop/>).

Článek vychází ze stejnojmenného grantového projektu podporovaného v rámci programu Informační společnost GA AVČR (č. 1ET200020405; 2004–2008).



■ Obr. 9. Kromě odborných dat uchovávají často archeologické archivy i kouzlo života dávno minulého... Výzkum v Zalužanech, okr. Příbram, v r. 1959. Foto A. Rybová, archiv ARÚP FT–25974.