

MANUSCRIPTORIUM: EVROPSKÁ DIGITÁLNÍ KNIHOVNA, RUKOPISŮ S PLNÝMI TEXTY

Zdeněk Uhlíř

Hned jak v roce 2003 vzniklo Manuscriptorium jako otevřený katalog historických fondů a jak se v roce 2004 rozšířilo v digitální knihovnu obrazových kopií rukopisů a posléze i inkunábulí, starých tisků a dalších historických dokumentů, uvažovali jeho tvůrci o tom, že úroveň jediné instituce, ba ani úroveň národní nebude stačit k naplnění cílů, které si přitom kladli. A tak se od počátku více či méně úspěšně snažili kontaktovat i partnery zahraniční, aby se Manuscriptorium prosadilo i na úrovni mezinárodní a nadnárodní. Když se to začalo dařit v měřítku středoevropském a když se rozšířily kontakty i na Balkán a dalších zemí jihovýchodní Evropy, jakož i do zemí středoasijských, začalo být jasné, že Manuscriptorium pomalu, ale jistě stává globálním hráčem, a že se tudíž musí pokusit prosadit se jako dominantní evropská digitální knihovna rukopisů a starších historických knižních fondů vůbec.

To se začalo rýsovat jako reálná možnost, když Národní knihovna České republiky jakožto koordinátor získala grant na evropský projekt ENRICH, jehož cílem je integrovat výsledky digitalizace rukopisů v Evropě do jediného uživatelského rozhraní a jediného distribuovaného systému, jehož základem bude právě Manuscriptorium. Projekt bude zahájen 1. prosince 2007, potrvá dva roky a jeho řádnými partnery bude osmnáct institucí z dvanácti zemí (Česko, Polsko, Litva, Maďarsko, Island, Dánsko, Německo, Francie, Rakousko, Španělsko, Itálie, Velká Británie) a asociovanými partnery další instituce z devíti zemí (Švédsko, Srbsko, Rumunsko, Moldávie, Makedonie, Slovinsko, Slovensko, Kazachstán, Turecko).

Při bližším zkoumání je samozřejmě možné namítnout, že např. velké evropské země jako Velká Británie, Německo a Francie jsou zastoupeny pouze technickým a nikoli obsahovým partnerem, tzn. univerzitním pracovištěm oboru *computing in humanities* či *hi-tech company*, na druhou stranu však je třeba zdůraznit, že právě tyto tři země v digitalizaci rukopisů v žádném případě nevynikají, ba rozhodně patří k těm slabším. Naopak země jako Česko, Island či Srbsko patří v evropském měřítku k velmocím v digitalizaci rukopisů a z velkých starých evropských zemí s nimi dokáže jakž takž držet krok pouze Španělsko. Z toho plyne, že projekt ENRICH, pokud naplní své předpoklady, bude integrovat asi 85 % v Evropě dosud zdigitalizovaného rukopisného materiálu. Zajímavé je zjištění, že v digitalizaci rukopisů jednoznačně dominují ty národy, které byly významnější ve středověku

než v novověku (Čechové, Islandané, Srbové), zatímco ty národy, jež dosáhly hlavního významu v novověku (Britové, Němci, Francouzi), digitalizaci rukopisů vcelku opomíjejí. To je jistě zajímavý problém, ale na tomto místě se nemohu pokoušet jej řešit.

Zároveň s rozšiřováním geografického záběru se rozšiřuje i věcný a materiálový rozměr Manuscriptoria. Kromě katalogových záznamů (otevřený katalog historických fondů) a digitálních obrazů (digitální knihovna obrazových kopií) byly po testování v roce 2006 v letošním roce 2007 zpřístupněny první plné texty (elektronická knihovna plných textů), přičemž všechny tyto typy digitálních objektů jsou vzájemně koordinovány až na úroveň jednotlivých dílčích jednotek dokumentů-souborů, takže pro koncového uživatele komfortní a vřidný komplexní digitální dokument. A vzhledem k tomu, že komplexní digitální dokument umožňuje přiřadit k jedinému fyzickému objektu nejenom více typů digitálních objektů (katalogový záznam, sekvenci digitálních obrazů, plný text, audio, video, multimodální, resp. multimediální dokument), ale dovoluje i alternativní reprezentace v rámci jednotlivých typů digitálních objektů (více katalogových záznamů soustřeďujících se na deskripci originálního fyzického objektu z různých hledisek, více sekvencí digitálních obrazů, více plných textů atd. atp.), lze vytvářet také virtuální dokumenty, které jsou kombinacemi částí komplexních digitálních dokumentů korespondujících jednoznačně s konkrétními fyzickými objekty a vytváření takových virtuálních dokumentů lze poměrně snadno personalizovat, v důsledku čehož je vytvoření virtuálního badatelského prostředí pro práci s historickými fondy, jež je cílem týmu budujícího Manuscriptorium, opravdu na dosah ruky.

Zařazení plných textů do standardní nabídky Manuscriptoria je tedy důležitým krokem vpřed na cestě k tomu, čemu se někdy poněkud kontroverzně říká digitální historie, resp. digitální filologie, a co je adaptací někdy nemálo konzervativních humanitních disciplín na nepřilíš určitou představu informační a znalostní společnosti. Je proto důležité, aby nabídka plných textů v Manuscriptoriu byla nejenom standardní, ale také druhově variabilní tak, aby postihovala různé vrstvy zájmu o historické fondy a reagovala na reálnou práci s nimi, přičemž nebude od věci, aby se tím otvírala cesta novým a v tradičním tištěném prostředí neznámým, protože nerealizovatelným postupům a pravděpodobně také dosud nemyslitelným představám. Snadná kombinace různých reprezentací téhož typu dílčího digitálního dokumentu v rámci komplexního digitálního dokumentu to snadno dovoluje, a to jak po stránce věcně obsahové, tak formálně technické či metodické. Zejména v tomto ohledu je tudíž důležitá velká variabilita.

Poněvadž účelem digitálních knihoven dostupných po síti je usnadnit přístup k fondům a dokumentům, je tomu nejenom možno, ale přímo nutno rozumět

různým způsobem. Zajisté nejjednodušší je poskytnutí prosté digitální kopie – přirozeně s popisnými a technickými metadaty, protože digitalizace není jen obyčejné dělání obrázků, jak všichni dávno víme. Digitální obrazová kopie může být snadno přenášena po síti, neomezeně kopírována, technicky zpracovávána (komprimace, kontrastování, inverze, černo-bílé optimalizace apod.), což na jedné straně může přinášet a přímo vytvářet problémy, na druhé straně se na tom však jasně ukazuje, že digitalizace není jen prostředek ochrany, ale také – a pravděpodobně hlavně – snadná cesta zpřístupnění. Jestliže to platí pro digitální obrazy, tím spíše to platí pro elektronické texty. Zcela nasnadě tu připadá v úvahu vytvářené edice originálních historických dokumentů, a to v různé podobě, která je alternativní, nikoli hierarchicky uspořádaná, zejména jako prostých přepisů či transkripcí, pragmatických či kritických edic. Naopak jako pseudoprobém se v tomto kontextu ukazují diplomatické přepisy či transkripce, které dříve měly svou funkci a podle toho byly oceňovány. Bylo totiž obtížné a ve větší míře dokonce nemožné pořizovat tzv. zrcadlové edice zpřístupňující simultánně jak obrazovou kopii originálu, tak jeho přepis. V digitálním světě je však něco takového velice snadné, takže není nutno plýtvat silami na tak náročnou práci, která je v novém kontextu takřka úplně zbytečná, a je možno se systematicky věnovat jiným věcem.

Základem tudíž je zpřístupňování plných textů primárních dokumentů v originálním, prvotním, původním vyjádření. Aby nedošlo k omylu, tím nemíním ani archetyp, ani tzv. *Urtext*, ani tzv. text poslední ruky či kanonický text, nýbrž prostě konzistentní a intaktní, tj. dobré a úplné znění v původním jazyce opírající se o konkrétní zápis v příslušném rukopise, ovšem edičně upravený, tzn. s provedenými emendacemi (ale raději bez konjektur) a s určenými *testimoniis* (označujícími zjevné i skryté citace, jakož i aluze a konečně i prosté odkazy bez citací). Pro různocnění a jiný kritický aparát ve smyslu tradiční kritické edice tu tudíž nejspíše nebude místa. Přirozeně mám v této souvislosti na mysli nově vznikající edice, protože se nebude možné vyhnout – alespoň ze začátku – přebírání starších standardních kritických edic, jejichž nahrazení jinou formou by trvalo neúnosně dlouho.

Musíme zde totiž myslet na již zmíněné virtuální badatelské prostředí, jež sice ještě není ani v dílčím náznaku realizováno, ale které už takřkajíc klepe na dveře. Ideálem tu je představa tzv. verzování známá ze studia textů romantického období, tzn. ediční zpřístupnění všech dochovaných, resp. dostupných zápisů a jejich postupné zpracování za pomoci nějakého korpusového manažeru a případně dalších externích nástrojů počítačnické lingvistiky, a to na základě předchozího určení markantů nacházejících se v každém zápisu (zejména specifická syntagmata či slovní komplexy vůbec a motivy) a následného vyhodnocení příbuznosti textů podle míry shod markantů, tzn. podle shlukování. Vnitřní znaky (v tomto případě markanty)

Lze potom korelovat s vnějšími znaky určenými paleografickým a kodikologickým, ale i dalším pomocněvědným zkoumáním, a tak se pokusit zjišťovat reálnější vztahy mezi rukopisnými zápisy, než jak je představují lachmanniánská či bédierovská stemmata. Takové koordinované zpřístupnění digitální knihovny obrazových kopií a elektronické knihovny plných textů sice představuje oproti tradičním postupům paradigmatickou změnu (právě proto se mluví o digitální historii a digitální filologii), nicméně přece je relativně neutrální, protože představuje jednotný podklad pro tak rozdílné metodologie, jako jsou *New Philology* na jedné a tzv. fluidní text na druhé straně. Zdá se, že teprve rozvoj digitálních knihoven obsahujících opravdové komplexní digitální dokumenty a rozšíření různých reprezentací elektronických textů, jakož i různých postupů při jejich zpracování dovolí prakticky rozvinout tyto nové metodologie a neuvážovat o nich více méně jen teoreticky.

Cílem zpřístupnění na dálku skrze digitální knihovny je nepochybně také demokratizace přístupu k vzácnému kulturnímu dědictví a odmítnutím obav z tzv. zneužití neodborníky. Je sice takřka zcela jistě iluzí, že kulturní dědictví se takovým způsobem dostane až k tzv. člověku z ulice, nicméně k těm, které lze nazvat tzv. vzdělanými laiky, má nepochybně šanci se dostat. A právě zde se nejzřejměji ukazuje důležitost toho, že je možné zařazovat do komplexního digitálního dokumentu různé reprezentace téhož typu dílčího digitálního objektu, v tomto případě alternativní plné texty. Kromě edic různého typu tak je možno dodat i překlad a učinit tak kulturní dědictví přístupné širšímu publiku. A důležité je, že v tomto případě nejde o tzv. popularizaci, která leckdy nemusí mít daleko k indoktrinaci, ale o relativně neutrální prostředek k vymanění kulturního dědictví z odbornického (záměrně uvádím odbornického a nikoli odborného) monopolu. Emancipace od odbornictví (znovu podotýkám, že nehovořím o bezpříznakové odbornosti) může být přinejmenším dobrou zbraní proti ortegovské vzpouře davů, což je dnes stále ještě – nebo možná už zase znovu – aktuální.

Jako plné texty ovšem mohou být zpřístupňovány i texty sekundárních dokumentů, tzn. v historickém smyslu dokumentů pojednávajících o primárních dokumentech čili tzv. literatura na rozdíl od pramenů. Zde jsou zajisté možnosti nepřehledné, jak dokazuje bouřlivý rozvoj internetu v posledních deseti patnácti letech a agregovaných informačních zdrojů v jeho rámci. Jak dokazují odborné texty a vůbec věcná próza na webových stránkách i v databázích, nemusíme být daleko jak od změny dominantních žánrů, kterými jsou dodnes především články v časopise a příspěvek z konference, tak od změny způsobů publikační kontroly, kterými jsou recenzní posudky, a tudíž od přehodnocení celého procesu evaluace ve všech jeho stupních od autora přes distributora (záměrně neříkám nakladatele)

až ke koncovému uživateli. Toto je velká výzva a považuji za nesmírně důležité, že se Manuscriptorium jako celek pokusí na ni podat, byť jenom dílčí, odpověď.

Mnohosti funkcí, které naplňuje zařazování plných textů – a zvláště edic – do Manuscriptoria, odpovídá i mnohost forem a technických postupů při jejich utváření. Nepochybně je nesmírně důležité, aby přírůstek plných textů v Manuscriptoriu byl nejenom pravidelný, ale také co možno největší. Kombinace různých postupů tak rozhodně nabývá na důležitosti. A tak přichází v úvahu snaha o automatický postup. Ten však lze v nynější chvíli ekonomicky zvládnout pouze u časově nejmladší vrstvy starých tisků, třebaže technologie se rychle vyvíjejí a perspektivy automatického postupu i pro starší až nejstarší tištěný materiál nejsou pravděpodobně příliš vzdálené a jistě nebude trvat dlouho a dojde i na rukopisy. Rozvoj na tomto poli bude v průběhu jedné jediné generace určitě velice prudký, takže nelze předkládat žádné věrohodné odhady, každopádně však je nutné, abychom byli připraveni a nenechali si uniknout příležitost. Jak se ukazuje, postup lze vskutku snadno industrializovat od scannování přes OCR a kódování až k závěrečné korektuře. Podobně je možno postupovat i při přebírání starších edic, ale v tomto případě se někdy ukazuje, že je – vzhledem k množství a různorodosti poznámkového aparátu, který vzdoruje automatickému kódování – efektivnější manuální přepis.

Konečně je záhodno pořízovat i původní edice, a to jak v podobě edic pragmatických, o nichž již byla řeč, tak i prostých prepisů bez výkladového aparátu a tzv. *testimonií*. Vzhledem k velkému výběru možností kódování textu je v tomto případě nutno hledět na účel zpřístupnění a případně postupně diferencovat mezi různými formálními reprezentacemi téhož plného textu, tzn. mezi jeho různým vnitřním strukturováním a tomu odpovídajícím kódováním. V tomto případě jsou tvůrci plných textů pro Manuscriptorium teprve v počátcích, takže je předčasné mluvit o autoritativních metodikách a technikách, nicméně jedno podstatné pravidlo je jasné již dnes, totiž že je nutno udržovat vyváženost mezi odborností a profesionalitou na jedné a zdravým rozumem na druhé straně. Soudným lidem se to možná zdá být natolik samozřejmé, že je hanba se o tom zmiňovat, ale nelze si nepovšimnout, že každý člověk, který pronikl do hloubi nějaké problematiky, zdravý rozum snadno ztrácí.

Je zřejmé, že pracovníci paměťových institucí – mám zde nyní na mysli hlavně knihovny, ale v různé míře se to dotýká a archivů a muzeí – budou muset v důsledku budování digitálních knihoven více či méně změnit svůj profesionální profil. Prinejmenším bude nutné, aby informační gramotnost mnohem více než dosud spojovali s kompetencemi v oblasti *computing in humanities*. A pokud konkrétně jde o pracovníky obírající se rukopisy, budou ve světle zkušeností z budování Manuscriptoria potřebovat rozšířit tradicionalistický a poněkud omezený rejstřík

pomocných věd historických o další obory tak, aby mohli nejenom spolehlivě katalogizovat a odborně pojednávat o vnějších znacích rukopisů a o rukopisech jakožto nosičích, ale také zpřístupňovat plné texty v prostých přepisech i pragmatických edicích a odborně pojednávat o vnitřních znacích rukopisů a o rukopisech jakožto záznamech. Možná že to bude největší problém, který před námi a celým našim oborem stojí.