

Konceptuální rámec virtuálního badatelského prostředí pro práci s historickými fondy

Zdeněk Uhlíř

Národní knihovna ČR, Praha

Zdenek.Uhlir@nkp.cz

INFORUM 2004: 10. konference o profesionálních informačních zdrojích

Praha, 25. – 27.5. 2004

Digitalizace originálních dokumentů již zhruba deset patnáct let probíhá relativně rychlým tempem. Lze s uspokojením říci, že Česko tento vývoj úspěšně zachytilo a že se i v mezinárodním a nadnárodním rámci podílí na výzkumu a vývoji a dosahuje také v této oblasti zdařilých výsledků. Nyní se ukazuje, že byla už dávno překročena úroveň prosté ochrany originálů, ale že digitalizace začíná postupovat i za pouhé zpřístupnění prostřednictvím komunikace na dálku a elementárních nepřímých služeb vzdáleným uživatelům. Digitalizace kulturního dědictví či historických fondů již pokročila v komunikační rovině natolik, že se skrze ni začíná měnit sama povaha informace a její pojem: není už nějakou pevnou formou, ani dokumentem či artefaktem, ani rovná datům nebo obsahem či významem dat, ale stává se naopak symbolickou formou kontextuálního významu. Digitalizace a informačně komunikační technologie zkrátka způsobily, že informace se přesunula ze sféry objektivního do sféry virtuálního; proto můžeme mluvit o virtuálním informačním, komunikačním a znalostním prostředí. Již z elementární zkušenosti je zřejmé, že toto prostředí se podstatným způsobem odlišuje od tradičního prostředí tištěného, jež dominovalo v novověku. Vstupujeme do jiné doby, která je nesena informační vlnou a zanechává za sebou vlnu industriální s jejími sociálními modely, jimž se dnes říká evropské. Jelikož v této situaci nemůžeme vystačit s tím, co se tzv. osvědčilo, musíme se pokusit o vytvoření nového konceptuálního rámce, který by zohlednil nově vyrostlé skutečnosti. O to se – byť jenom dílčím způsobem – budu snažit v tomto příspěvku.

Při vytváření konceptuálního rámce je třeba mít poměrně přesnou a určitou představu, pro který obor skutečnosti je určen, tzn. pro který obor skutečnosti je použitelný. A tak tedy konceptuální rámec, který chci předestřít na tomto fóru, se týká virtuálního badatelského prostředí pro práci s historickými fondy. Tím jsou vyjádřeny tři podmínky, pro něž má platit. Za prvé pro prostředí virtuální. Je to prostředí tvořené informačně komunikačními technologiemi, sítí počítačových sítí, tedy to, co se běžně nazývá prostředím elektronickým. Mimo okruh svého zájmu tedy odsunují tradiční prostředí tištěné, poněvadž důvodně předpokládám, že již v blízké budoucnosti se stane marginálním vedle dominantního prostředí elektronického. Na tradiční prostředí za těchto okolností není třeba brát zřetel. Za druhé pro prostředí badatelské. Badatelské prostředí je pouhou částí či výsečí prostředí virtuálního, tzn. zatímco první podmínka je obecná, druhá je specifická. Badatelské prostředí je charakterizováno dvěma rysy, jež musí naplňovat: jednak dodatečnou informací, jednak přidanou hodnotou. Dodatečná informace vykazuje těsnou návaznost na informaci základní, které je sdělována, a to buď na základě jinde se nacházející zhmotnělé, akumulované informace, nebo na základě informace nové, pro příslušný účel přímo vytvořené. To ovšem znamená, že nelze prostě říci, jako by nebo že by informace (ať už základní nebo dodatečná) prostě někde jenom „byly“. Musí být se vzájemným ohledem na sebe interpretovány – a z toho plyne, že cílem badatelského prostředí je poskytovat informaci nějakým způsobem novou, tj. pragmatickou informaci, transinformaci. Přidaná hodnota znamená, že badatelské

prostředí zároveň musí poskytovat technické nástroje umožňující nejenom základní vyhledávání či třídění informace, ale i některé vybrané nadstavbové nástroje umožňující další zpracování jak základní informace, tak dodatečné informace, tj. metainformace za účelem získání transinformace. Virtuální badatelské prostředí není jen tradiční knihovnou podrobenou elektronicko-digitální konverzi, nýbrž sférou znalostní práce. Za třetí pro historické fondy. Historické fondy jsou dochovanou zhmotnělou kulturou, jak ji organizačně spravují, zpracovávají a zpřístupňují paměťové instituce, tj. knihovny, muzea a archivy, a to jak veřejné (státní či samosprávné), tak soukromé. Odhlédneme-li od právních a administrativních záležitostí, což je při tvorbě konceptuálního rámce docela dobře možné, na konkrétní povaze institucí nezáleží. Jistě může vyvstat problém, nakolik jsou historické fondy totožné s kulturním dědictvím, jaký je jejich specifický vztah k pozůstatkům minulosti a pramenům dějin, ale to už bychom se posunuli do jiného konceptuálního rámce.

Pokud jsme takto vymezili základní podmínky, jež je třeba naplnit, lze pokročit dále k formulaci základních rozdílů mezi tradičním prostředím tištěným a novým elektronickým. Rozdíly mezi oběma prostředími jsou takové a tak výrazné, že je nutno chápat je jako změnu kvantity v kvalitu, jak se dříve říkalo, takže nakonec přerůstají v rozdíl paradigmatický. Projevuje se to ve dvou základních dimenzích, totiž v komunikaci informace a znalosti na jedné a v povaze informace a znalosti na druhé straně. Rozdíly v případě komunikace informace a znalosti mezi prostředím tištěným a elektronickým se projevují ve třech základních sférách – týkají se času, prostoru i množství a odrážejí rozdíl v povaze mentálního vnímání závislého buď na konkrétních věcech, nebo na abstraktních objektech. Z hlediska vnímání a následného pojímání času je pro tištěné prostředí typická lineárnost a diachronnost (jasné odlišení intervalů, během nichž v následné sekvenci vstupují do vědomí jednotlivé nosiče informace, tj. intervaly jsou diskrétní), zatímco pro prostředí elektronické je typická simultánnost a synchronnost (intervaly, během nichž vstupují do vědomí jednotlivé reprezentace informace, jsou ve své neurčitosti je stěží rozlišitelné, tj. intervaly jsou kontinuální). Z hlediska vnímání a následného pojímání prostoru – což je ostatně jen důsledek odlišného vnímání a pojímání času – je pro tištěné prostředí typická lokálnost (aktuálně vstupuje skrze vnímání do vědomí jen ten nosič informace, který je reálně místně přítomen), zatímco pro prostředí elektronické je typická globálnost (aktuálně vstupují skrze vnímání do vědomí různé reprezentace informace, přičemž není nezbytné, aby byly reálně místně přítomny, tzn. mohou být místně rozptýleny). Z hlediska vnímání a následného pojímání množství – i to je nakonec jen důsledkem odlišného vnímání a pojímání času – je pro tištěné prostředí typická jednotlivostnost (postup od vědeckým pomyslem reprezentované přirozené jednotliviny – věci, procesu, činnosti – ke kumulaci takových jednotlivin), zatímco pro prostředí elektronické je typická hromadnost (postup od simultánně či synchronně do vědomí vstupujících jednotlivin k analytickému rozlišování a následnému kombinování jak přirozených jednotlivin, tak vědeckých pomyslů). Úhrnem vzato, komunikace tištěného prostředí je založena na striktním odlišení subjektu a objektu ve vztahu k informaci, neboť je v něm zřejmá sekvenční diskrétnost intervalů vnímání a pojímání informace, přičemž jasná odlučitelnost nosičů a reprezentací informace vede také k projekci objektu z vnitřní mentální sféry do vnějšího světa, na základě čehož je přirozená představa věci jako substance nahrazována vědeckým pomyslem jakožto faktem. Elektronická komunikace naproti tomu neumožňuje vzhledem k simultánnosti či synchronnosti, globálnosti a hromadnosti reprezentací informace jasně rozlišit mezi subjektem a objektem. Reprezentace informace není věc, nejsou věci, ale významy věci, významy věcí a informace sama je významem významu věci, významem významů věcí. Informace je jednak kontextuální interpretací význam významů, jednak je intersubjektivní a nikoli objektivní srozumitelností. Není zde jednoduchá projekce do vnějšího světa, jež by mohla být pochopena či vyložena jako objektivní fakt, ale symbolická analýzy informace jakožto vztahu její reprezentace,

přirozených představ věcí a vědeckých pomyslů a konečně pragmatické a finalitní zacílení této symbolické analýzy.

Tak se od odlišné komunikace informace a znalosti ve dvou různých prostředích – tištěném a elektronickém – odvíjí i jejich různá povaha. Zejména pět rozdílných rysů je tu nutno zdůraznit. Za prvé, nejzřetelnější je rozdíl v reprezentaci: zatímco v tištěném prostředí je typickou reprezentací informace dokument, v prostředí elektronickém je to celý zdroj dokumentů. To je analogií mezi jednotlivostností a hromadností z hlediska komunikačního, třebaže tyto dvě dimenze nelze ztotožňovat. Za druhé, rozdíl se projevuje v tom, že typická reprezentace tištěného prostředí, tj. dokument, je charakteristická svou uzavřeností, zatímco typické reprezentaci prostředí elektronického je vlastní otevřenost: k dokumentu nelze nic přidat, aniž přestane být tímž dokumentem, zatímco rozšiřování zdroje je dobře možné, a přece zůstane stále tímž zdrojem. Za třetí, podstatný rozdíl mezi oběma prostředími spočívá v metodologii tvorby a dalšího zpracování informace. Zatímco tištěné prostředí preferuje redukcionismus spočívající v jednoaspektovém pozitivním vidění informace, prostředí elektronické se soustřeďuje především na komplexitu zakládající se na prolínání různých aspektů informace. Čtvrtý zásadní rozdíl vyplývá z předchozího: v důsledku svého redukcionismu se tištěné prostředí pohybuje pouze v jedné dimenzi, takže v něm dochází k prosté kumulaci informace, třebaže kumulovaná informace může být následně různým způsobem strukturována; prostředí elektronické naproti tomu v důsledku svého aspektového pohledu na informaci se rozkládá do neurčitého počtu dimenzí, jež mohou být sledovány jak v rámci pouze jedné z nich, tak napříč, třeba i všemi. Za páté, úhrymem tedy tištěné prostředí tenduje k předmětnosti, zatímco elektronické prostředí je svou povahou zaměřeno spíše k nepředmětnosti, předmětnost v jeho rámci je teprve jakýmsi následným zurčněním. Základní povaha informace a znalosti v tištěném prostředí je tedy taková, že se dokáže pohybovat, pokud je využívána uvědoměle a kvalifikovaně, buď jen v rámci přirozených věcí-substancí, nebo naopak pouze v rámci vědeckých pomyslů (přičemž není vyloučeno využití neuvědomělé a nekvalifikované, v němž přirozené věci-substance a vědecké pomysly nekontrolovaně až nekontrolovatelně splývají), zatímco v prostředí elektronickém se vědomě podává nějaká mixáž přirozených věcí-substancí spolu s vědeckými pomysly. Povaha informace a znalosti v tištěném prostředí je založena na správnosti, která je více méně ztotožňována s pravdivostí, přičemž pravdě je přiřítán ontologický status, zatímco povaha informace a znalosti v prostředí elektronickém se zakládá na vhodnosti, které je právě tak více méně ztotožňována s pravdivostí, pravdě je ovšem při tom přikládán status axiologický. V elektronickém prostředí se pak rozlišuje virtualita pravá od nepravé-faktoidu, to však už je mimo tento konceptuální rámec.

A tak se dostáváme k tomu, jak by mělo být virtuální badatelské prostředí pro práci s historickými fondy prakticky organizováno. Tu je třeba rozlišit dvě sice různá, ale souvislá hlediska – organizaci informace na jedné a organizaci znalosti na druhé straně; informace v tomto smyslu vyjadřuje strukturování obsahu, zatímco znalost představuje práci s informačním obsahem. Organizace informace ve virtuálním prostředí musí být nejprve strukturována do postupných horizontálních dimenzí, protože jen tak může zůstat jednoduše přehledná a lehce srozumitelná. Zcela postačující jsou v tomto prvním plánu pouze dva horizonty představované centrem a periferií. Je tu ovšem jistý problém, protože koncepty centra a periferie, pokud mají být chápány vskutku virtuálně, jsou v obecném smyslu věčně zaměnitelné, tzn. jsou pouze relativní, nelze v absolutní smyslu říci, co je centrem a co periferií *an sich*. Jsou relativní vzhledem k tématu. To ovšem znamená, že nelze hovořit o totálním virtuálním prostředí, ale o množině dílčích virtuálních prostředí. Pluralita takových prostředí musí být dána nikoli pluralitou předmětů, nýbrž pluralitou témat. Volba toho, co je centrum a co periferie, je tudíž relativní vzhledem k tématu, což znamená, že příslušná k němu se vztahující informace nikoli „je“ taková či onaká, nýbrž že se „jeví“ právě tak a ne

onak, zatímco při jiném tématu se může „jevit“ jinak. Organizace znalosti ve virtuálním badatelském prostředí spočívá na základním rozlišení mezi statickou externí pamětí na jedné a základními i nadstavbovými nástroji pro práci s ní na druhé straně. Z hlediska znalosti jsou obsahem statické externí paměti pouhá data, ještě nikoli informace. Kdybych chtěl být striktní, měl bych říci *capta* jakožto zaznamenaná data, ale to by v těchto souvislostech znělo až příliš subtilně. Z dat statické externí paměti činí informaci (mateinformaci, transinformaci) až základní i nadstavbové nástroje informačně komunikačních technologií. Znalost je právě schopnost využívat těchto nástrojů s ohledem na dané téma, tj. nikoli s rigidní metodickou apriorností. Rozhodující tudíž je, že data a nástroje nesmějí splývat, protože v tom případě by došlo nikoli k organizaci, nýbrž k dezorganizaci.

Za centrum informace je nutno pokládat otevřený katalog historických fondů skládající se z jednotlivých evidenčních záznamů. Z technického hlediska je vhodné rozlišit řekněme elektronickou databázi evidenčních záznamů a prezentační databázi, přičemž prezentačních databází může být nakonec více, kdybychom jim dokázali přiložit různý smysl. Ale to v tuto chvíli není podstatné. Jelikož katalog je vždy katalogem něčeho určitého, je třeba definovat katalogizační jednotku, které je předmětem, resp. denotátem jednotlivého katalogového záznamu. Je to materiální originál, kterému je připsán status jednotky historického fondu, tedy jednotka pohybující se na škále mezi krajními póly pozůstatku minulosti a pramene dějin. Prakticky vzato je to prvek množiny takových jednotek, která je představována sbírkami participujících veřejných, případně i soukromých institucí. Periferie informace naproti tomu je představována souvisejícími digitálními dokumenty navázanými na ten který evidenční záznam. Tyto související digitální dokumenty mohou být rozličné povahy: digitální obrazové kopie originálních dokumentů, digitální zvukové kopie originálních dokumentů, digitální zvukové dokumenty reprezentující originální dokument jiné formy, elektronické plné texty originálních dokumentů, elektronické plné texty sekundárních dokumentů, elektronické terciární dokumenty, multimodální dokumenty atd. Vazba evidenčního záznamu a jednoho každého souvisejícího digitálního dokumentu musí být jednoznačná, tzn. musí být jasné, ke kterému evidenčnímu záznamu se váže ten který související digitální dokument. Zároveň však není jedno-jednoznačná, protože k jedinému evidenčnímu záznamu se může vázat více než jeden související digitální dokument. Evidenční záznam se souvisejícími digitálními dokumenty, jež jsou na něj navázány, pak vytváří komplexní virtuální dokument. A jestliže výsledkem základního vyhledávání v otevřeném katalogu historických fondů je více než jeden evidenční záznam a potažmo komplexní virtuální dokument, objevuje se přechodný sdružený dokument, který už je jasně virtuální povahy.

Jsou tu tedy dimenze, které se prolínají napříč: jednak centrum a periferie, jednak komplexní virtuální, resp. přechodné sdružené dokumenty. Ty tvoří jeden strukturovaný zdroj, jemuž ještě nelze v plné míře přiřknout virtuální charakter, třebaže některé rysy virtuality už nepochybně má. Je třeba připustit, že takový zdroj je pouhou výsečí celkového elektronického, resp. elektronicko-digitálního prostředí a že teprve toto celkové prostředí nabývá plně virtuálních rysů. Je ovšem ve své univerzálnosti v zásadě neurčitě a bez prostředků orientace a nástrojů navigace se jeví jako chaotické. Je tudíž nutno je nějakým způsobem strukturovat, aby mohla vystoupit na povrch jeho virtuální a nikoli jen chaotická povaha. A tak dospíváme k rozlišení řekněme vlastního zdroje od zdrojů v síti. Vzhledem k tomu, že vlastní zdroj se prakticky, aktuálně jeví jako komplexní virtuální dokument, potažmo přechodný sdružený dokument, lze potom za jistých podmínek mluvit o dimenzi komplexního virtuálního zdroje, a to tehdy, je-li komunikace vlastního zdroje a cizích zdrojů postavena na jasných principech. Lze pak říci, že vlastní zdroj představuje centrum vyšší úrovně a ostatní zdroje v síti jsou jeho periférií. Zdroje v síti lze rozlišit na tzv. viditelný web, tj. řekněme webové stránky a tzv. neviditelný web, tj. řekněme databáze. Komplexní virtuální

zdroj je tvořen zdrojem vlastním a těmi zdroji neviditelného webu, k nimž je nejenom volný či licencovaný přístup, ale dohodnutý způsob komunikace na základě příslušných komunikačních protokolů. Virtuální prostředí je však okruh ještě mnohem širší než komplexní virtuální zdroj, takže i na této úrovni je možno rozlišit centrum a periferii. Viditelný web je součástí právě této periferie a tato periferie zahrnuje i ty zdroje neviditelného webu, jež jsou sice z vlastního zdroje přístupné, ale nikoli na bázi zvláště dohodnutých komunikačních protokolů. Teprve v tomto rozsahu lze mluvit o virtuálním badatelském prostředí. Naopak ty zdroje neviditelného webu, jež nejsou z vlastního zdroje přístupné, do virtuálního badatelského prostředí zařadit nelze než jen potenciálně.

Záleží tedy podstatným způsobem na prostředcích orientace a nástrojích navigace, neboť jedině tak lze realizovat znalost jakožto schopnost využít informaci ve virtuálním prostředí. Znalosti jsou obvykle – stále ještě – ztotožňovány s pouhými vědomostmi, tj. s obsahem interní paměti. Jestliže však znalost pochopíme jako schopnost, takové vymezení evidentně přestává stačit. Jednak je třeba znalost rozšířit také o obsah externí paměti, jednak je nutno ji nějakým způsobem dynamizovat, tzn. aktualizovat ji jak mentálními nástroji a prostředky, tak vnějšími technickými nástroji. Důsledkem toho je, že znalost není pouhou reprodukcí obsahu, ale tu jeho aplikací, tu inovací. V tomto smyslu nemá znalost pouze faktickou, ale také symbolickou povahu: není manifestací akumulované informace, ale symbolickou analýzou virtuálního prostředí, tj. komplexních virtuálních dokumentů, komplexních virtuálních zdrojů, jakož i dalších komplementárních zdrojů a jednotlivých dokumentů. Pro práci s historickými fondy to znamená, že centrum vlastního zdroje nemůže být postaveno jako tzv. *catalogue raisonné*, tj. nemůže primárně poskytovat detailní sofistikovanou deskripci vnějších znaků originálních dokumentů, ale musí být pojat tak, aby umožňoval interpretaci vnitřních znaků textového, tj. symbolického kontinua, jehož je ten který originální dokument pouhým prvkem. To vyžaduje strukturální definici vnitřních znaků dokumentů, jež mají být zahrnuty do otevřeného katalogu historických fondů, a to definici relativní a orientovanou, tj. vztaženou k tématu či tematickému okruhu, s ohledem na nějž je vlastní zdroj budován, potažmo je předpokládáno využití dalších zdrojů v celém virtuálním prostředí. Katalogový záznam tvořící centrum vlastního zdroje tak musí být charakterizován jako analytický, nikoli blokový rozpis obsahu originálního dokumentu, k němuž se vztahuje, přičemž ideálem je, aby analýza pokročila až do nitra základních textových jednotek, tzn. k jejich kompozici a k manifestaci jejich konkrétních rysů. Ty pak lze považovat za traktory nebo bifurkační body významového pole jako potenciálního předmětu symbolické analýzy. Zajisté je nutno, aby strukturální definice originálních dokumentů popisovaných v evidenčních záznamech byla komplementární jak s obsahem souvisejících digitálních dokumentů, tak s jejich formální strukturou. To však je otázka nikoli obecného, nýbrž specifického řešení, tudíž je pouhým konceptuálním rámcem nepostižitelná a musí být zodpovězena až v konkrétním datovém modelu.

Nastíněný konceptuální rámec virtuálního badatelského prostředí pro práci s historickými fondy do sebe integruje jak informaci, tak znalost aniž je vzájemně zaměňuje nebo naopak vylučuje jednu z nich. Zdá se, že v tom spočívá jeho podstatný přínos. Integraci informace a znalosti pak lze pochopit jako kontinuitu idejí v diskontinuitě diskursů, což je podstatný rozměr historie: znemožňuje jak infinitezimální fakticistní kumulaci, tak permanentní paradigmatickou revoluci a tím umožňuje vnímání historie i sám dějinný pohyb, třebaže nikoli ve smyslu lineárního pokroku. Kontinuita idejí v diskontinuitě diskursů je mentálním protějškem kulturního dědictví, tj. originálních historických dokumentů ve smyslu materiálního substrátu historie. Virtuální badatelské prostředí pro práci s historickými fondy přirozeně není kulturním dědictvím samým, nýbrž jeho reprezentací, historie však už od svého počátku měla a má co do činění s pouhými reprezentacemi, takže už v tomto základu je pevná kontinuita. Diskontinuita je v diskursu virtuality oproti předchozímu diskursu

objektivitu typickému pro prostředí tištěné. Virtuální reprezentace kulturního dědictví je reprezentací synchronní, tzn. přítomnost je sice oddělena od minulosti, nikoli však od jejích pozůstatků a tím ani od dějin. Dějiny nejsou nic odlišného od přítomnosti a v tom smyslu z hlediska přítomnosti eventuálně zbytného, jsou naopak jejím rozšířením, potencialitou. Kulturní dědictví je tedy rozšířením aktuálního, pouze bezprostředního, výlučně faktického. Pokud předchozí pozitivistické odbornictví vytvořilo ve svém důsledku z kulturního dědictví a jeho chápání jakýsi nadstandard, který je mimo běžné lidské možnosti, protože jsou pouze ignorantské, informační vlna se svým virtuálním prostředím vymaňuje kulturní dědictví z takového omezení. Pokud pozitivistické odbornictví dovedlo samo sebe i ostatní k chápání dějin jako pouhého dění a tím zrušilo zájem o ně (protože co s takovým děním, které sice bylo, ale už není a které není univerzální, nýbrž pouze specifické), virtuální prostředí a jeho paradigma tím, že kulturní dědictví reprezentuje jako rozšířenou přítomnost, zakládá tak po neblahém 19. a zejména 20. století pochopení dějin jako jednání, vytváří tak možnost takového jednání. A to jistě není málo.