

ZPRÁVA O VÝSLEDKÁCH ŘEŠENÍ PROJEKTU V ROCE 2013

INTERPI – Interoperabilita v paměťových institucích

Program aplikovaného výzkumu a vývoje národní a kulturní identity (NAKI) (DF11P010VV023)

Předkládá:

Mgr. Marie Balíková, odpovědná řešitelka - koordinátorka
Národní knihovna České republiky
Klementinum 190
110 00 Praha 1

Spolupracovali:

Mgr. Jana Šubová, Ing. Naděžda Andrejčíková PhD.,
Ing. Miroslav Kunt

Ing. Tomáš Böhm
generální ředitel
Národní knihovna ČR



INTERPI

15. 11. 2013

Obsah

A	Konstatační část	3
A.1	Rešerše	3
A.2	Vstupní data a cíl	3
B	Analytická část	4
B.1	Analýza, ověřování, doplňování a úprava dat.....	4
B.1.1	Příprava záznamů Národních autorit pro testování transformace dat ze stávajících systémů	4
B.1.2	Úprava, doplňování a analýza entit pro potřeby archivů.....	5
B.1.3	Aktualizace klasifikačního systému MDT	6
B.2	Vymezení tříd a podtříd entit.....	6
B.2.1	Třída entit dílo/výtvor.....	6
B.2.2	Třída entit obecný pojem	7
B.2.3	Třída entit rod/rodina	7
B.3	Příprava metodických a metodologických postupů, jejich zavedení v archivech	7
B.4	Testování transformace dat stávajících systému.....	8
B.5	Formulace požadavků na softwarové řešení pro doplňující skupinu entit	8
B.6	Propagace a informování o projektu	9
C	Návrhová část	10
C.1	Informace o uplatněných výsledcích projektu v roce 2013.....	10
C.1.1	Rozhraní pro zpracování základní skupiny entit.....	10
C.1.2	Studie v recenzovaném periodiku	10
C.2	Informace o neuplatněných výsledcích projektu - zdůvodnění	11
C.2.1	Metodika tvorby znalostního modelu.....	11
C.3	Cíle projektu pro rok 2014	13
D	Použití účelové podpory v členění dle rozpočtu projektu.....	15
E	Resumé	16
F	Přílohy	
F.1	Seznam fyzicky předložených výsledků projektu	
F.2	Použití účelové podpory v členění dle rozpočtu projektu - komentář	
F.3	Dodatky, rozhodnutí	
F.4	Cestovní zprávy	
F.5	Rozhraní pro zpracování základní skupiny entit - uživatelská příručka	
F.6	Metodika tvorby znalostního modelu	

F.7 Zpráva o činnosti Národní knihovny v projektu INTERPI za rok 2013

F.8 Zpráva o činnosti Národního archivu v projektu INTERPI za rok 2013

A Konstatační část

A.1 Rešerše

BALÍKOVÁ, Marie - KUNT, Miroslav - ANDREJČÍKOVÁ, Nadežda - ŠUBOVÁ, Jana. INTERPI: nástroj pro zpřístupnění národního kulturního dědictví. In *Knihovna*. 2013, roč. 24, č. 1, s. 28-44. Dostupné též na: <http://knihovna.nkp.cz/knihovna131/13128.htm>.

BALÍKOVÁ, Marie - ANDREJČÍKOVÁ, Nadežda - KUNT, Miroslav - ŠUBOVÁ, Jana. Znalostní báze pro paměťové instituce v Česku: projekt INTERPI v roce 2013. In *INFOS 2013 : zborník príspevkov z 37. medzinárodného infromatického sympózia*. Bratislava : Spolok slovenských knihovníkov, 2013. ISBN 978-80-89586-07-3, s. 12-25.

KUNT, Miroslav. Mezinárodní archivní standardy a příprava nových základních pravidel pro zpracování archiválií v České republice. In *Slovenská archivistika*. 2012, roč. 47, č. 2, s. 64-79.

ANDREJČÍKOVÁ, Nadežda - BALÍKOVÁ, Marie - KUNT, Miroslav - ŠUBOVÁ, Jana. Znalostní báze pro paměťové instituce v Česku: projekt INTERPI v roce 2013. In *Duha [online]*. 2013, č. 2. Dostupné na: <http://duha.mzk.cz/clanky/znalostni-baze-pro-pametove-institute-v-cesku-projekt-interpi-v-roce-2013>.

BALÍKOVÁ, Marie - KUNT, Miroslav - ŠULC, Ivo. Projekt INTERPI: harmonizace pravidel, standardů a postupů aplikovaných v paměťových institucích. In *13. konference Archivy, knihovny, muzea v digitálním světě*. 2012. Příspěvek na konferenci. Dostupné na: <http://www.skipcr.cz/akce-a-projekty/dokumenty/akm-2012/balikova.pdf>.

ANDREJČÍKOVÁ, Nadežda - ŠUBOVÁ, Jana. Projekt INTERPI: prototyp nástrojů pro zpracování entit v paměťových institucích. In *13. konference Archivy, knihovny, muzea v digitálním světě*. 2012. Příspěvek na konferenci. Dostupné na: <http://www.skipcr.cz/akce-a-projekty/dokumenty/akm-2012/Andrejckikova.pdf>.

A.2 Vstupní data a cíl

Základním cílem projektu, který naplňuje tematickou prioritu TP 3.6 - Infrastruktura pro vědeckou a technologickou podporu zachování národní kulturní identity, je tvorba znalostní databáze paměťových institucí na bázi sémantické interoperability.

Rok 2013 je v projektu vymezen jako druhá etapa. Navázali jsme na výsledky první etapy a s ohledem na cíle stanovené v projektovém záměru pro druhou etapu jsme se v rámci aplikovaného výzkumu a vývoje zaměřili na:

1. definování požadavků na softwarové aplikace pro zpracování základní skupiny entit,
2. vytvoření softwarové aplikace pro zpracování základní skupiny entit v podobě webového rozhraní,
3. přípravu metodiky tvorby znalostního modelu v projektu INTERPI,
4. testování možností transformace stávajících dat v Národních autoritách do struktury entit definované v projektu INTERPI,
5. doplnění prototypu softwaru pro zpracování druhé skupiny entit.

B Analytická část

Analytická část zprávy obsahuje informace o dílčích výsledcích dosažených v roce 2013 v oblasti aplikovaného výzkumu a vývoje v souladu s cíli projektu. Podobně jako v první etapě jsme se zaměřili na analýzu, ověřování, doplňování a úpravu dat jako přípravu pro transformaci dat ze stávajících systémů. Testování transformace dat proběhlo na základě připravených dat exportovaných ze souborů Národních autorit.

V návaznosti na výsledky první etapy pokračovali práce vymezení tříd entit a také definice požadavků na software pro zpracování entit.

B.1 Analýza, ověřování, doplňování a úprava dat

Cílem činnosti v této oblasti v roce 2013 bylo ověřit a doplnit údaje, které jsou vyžadovány pro projekt na základě vymezení potřeb paměťových institucí a z pohledu specifikace vlastností jednotlivých typů entit.

B.1.1 Příprava záznamů Národních autorit pro testování transformace dat ze stávajících systémů

V roce 2013 se NKČR zaměřila na analýzu, export a úpravu autoritních údajů jako základního předpokladu pro úspěšný vývoj znalostní databáze INTERPI. Údaje autoritní databáze byly podrobeny důkladné analýze z hlediska jejich budoucí integrace do tříd a podtříd entit znalostní báze INTERPI.

Byla definována kritéria pro export jednotlivých typů autorit v souladu s jejich budoucím zařazením.

Záznamy autorit, u nichž došlo k zásadní změně v jejich zařazení, např. záznamy ozbrojených střetů, byly podrobeny důkladnější analýze z pohledu formy jejich zápisu. Na základě této analýzy bylo přistoupeno k jejich úpravě již v rámci stávajícího souboru autorit: forma jejich zápisu byla přizpůsobena formě zápisu obvyklému pro třídu entit, do které budou integrovány, byly opatřeny stručnou definující poznámkou a dalšími údaji tak, aby jejich zařazení do „nové“ třídy entit bylo pokud možno bezproblémové.

Takto bylo upraveno 10 500 autoritních záznamů různých typů.

Více informací v F.7.

Export autoritních údajů

Podle definovaných kritérií byly vytvořeny exporty stávajících autorit.

Bylo exportováno:

- 562016 personálních autorit označujících osoby,
- 13299 autorit označujících kombinace jména aurora a název jeho díla,

- 128103 korporativních záhlaví,
- 3296 kombinací korporativních záhlaví a názvů díla,
- 25210 geografických autorit.

Více informací v F.7.

B.1.2 Úprava, doplňování a analýza entit pro potřeby archivů

Tvorba geografických entit typu sídlo

Již třetím rokem pokračovala práce na pořizování geografických entit typu sídlo ze dvou základních zdrojů.¹ Ukázalo se, že náročnost úkolu byla v počátcích podceněna. Názvy sídel, kromě gramatických odchylek, nejsou velmi často v prvním pádě, ale mají podobu citace z pramene (mnohdy jediný výskyt názvu). Komplikace také působí výskyt entit, které nezahrnujeme mezi geografické, zejména hrady, zámky (entita typu stavba). V těchto případech bude nutné provést revizi, stejně jako zkontrolovat formulaci názvu dle AACR2R použitelného pro databázi NA ČR. Následovat musí doplnění identifikace databáze CZ RETRO, čímž dojde k napojení na další údaje entitu charakterizující (např. historická příslušnost k okresům apod.).

Více informací v F.8.

Entita stavba - přehled úseků a vzniku železničních tratí Česka

Byla dokončena příprava dat z rukopisu uloženého v tzv. Archivu Českých drah v Praze-Libni.² Jedná se o unikátní soubor dat o vzniku nejen jednotlivých železničních tratí, ale i stanic a zastávek, stejně jako historii jejich doplňování (obr. 1). Následně práce pokračovaly dohledáním změn názvů od poválečné doby do současnosti. Liniové stavby jsou v mnoha ohledech problematické a uvedený soubor dat demonstruje dobře vazby i na další entity (stavitel - korporace, osoba; událost).

Více informací v F.8.

Entita korporace

Pokračovaly práce na shromáždění dat o historických korporacích zejména z oblasti veřejné správy. Pořízena byla data o všech ústředních a ústředně řízených institucích a podnicích dle stavu v roce 1986 (obr. 2)³

¹ Profous, Antonín: Místní jména v Čechách. Jejich vznik, původní význam a změny. Nakladatelství Československé akademie věd, 1954-1960, díl I-IV + dodatky (V); Hosák, Ladislav - Šrámek, Rudolf: Místní jména na Moravě a ve Slezsku I, II. Academia Praha, 1970, 1980

² Josef Panáček: Historická příručka ČSD 1838 - 1938. České kraje.

³ Jednotný číselník organizací v ČSSR. 1 Organizace řízené ústředními orgány. Federální statistický úřad, Praha : 1986

Bylo pokračováno ve shromažďování dat o vývoji korporací veřejné správy. Jedním z výstupů výzkumů korporativních entit z období před rokem 1918 byl referát na semináři v Národním technickém muzeu.⁴

Více informací v F.8.

B.1.3 Aktualizace klasifikačního systému MDT

V tomto roce (2013) řešení projektu jsme pokračovali v teoretickém průzkumu o aplikaci klasifikačního systému MDT jako potenciálního propojovacího jazyka při řešení terminologické různorodosti termínů aplikovaných v jednotlivých paměťových institucích a jako nástroje pro tvorbu ontologií.

V roce 2013 pokračovaly práce na vývoji vícejazyčné překladové databáze MDT - UDC online.

Více informací v F.7.

B.2 Vymezení tříd a podtříd entit

V roce 2013 se pozornost soustředila především na vymezení tříd obecný pojem a dílo/výtvar. Tyto třídy jsou z hlediska rozdílnosti řešení v jednotlivých komunitách nejvíc komplikované a nejméně zmapované.

Proto jsme přistoupili k analýze navrhovaného vymezení třídy dílo/výtvar s ohledem na řešení problematiky anonymních děl a děl neznámých autorů a také s ohledem na překryv entit této třídy s entitami ostatních tříd - např. geografický objekt, korporace resp. událost.

Vymezení ostatních tříd bylo ověřováno na příkladech.

Více informací v F.7.

B.2.1 Třída entit dílo/výtvar

Třída entit dílo/výtvar byla vyčleněna na základě společných vlastností tohoto typu entit, především na základě nejdůležitější charakteristiky, a to že se jedná o umělý výtvar, intencionálně lidskou bytostí vytvořené dílo. Proto do této kategorie spadají i stavby jako intencionálně člověkem vytvořený artefakt.

Abychom zohlednili všechny charakteristiky všech lidských výtvarů zařazených do této kategorie, bylo nutné vybrat typické příklady. Ukázka těchto příkladů je uvedena v příloze F 7.

⁴ Kunt, Miroslav: Ústřední správa železnic Předlitavska ve vazbě na jeho (nejen) dopravní politiku. Seminář z dějin železniční dopravy „Železnice v nás“ 15. října 2013.

B.2.2 Třída entit obecný pojem

Třída obecný pojem představuje „tvrdý oříšek“ ve všech znalostních bázích. Bývá velmi obtížné stanovit přesné hranice mezi individuálním a obecným pojmem (např. formalizované pojmenované entity typu bitvy, války. Termíny označující pojmy jsou vázány na přirozený jazyk - problematika nejednoznačnosti termínů - homonymie, synonymie a především jsou vázány na národní jazyky, jejichž systémy se někdy výrazně odlišují.

Více informací v F.7.

B.2.3 Třída entit rod/rodina

V problematice této třídy bylo nutné sjednotit přístupy komunity archivů a komunity knihoven v uvádění preferované podoby jména větve rodu. Pro zápis preferované formy větve rodu bylo použito schéma pro zápis korporativního záhlaví s podřízenou složkou.

Více informací v F.7.

B.3 Příprava metodických a metodologických postupů, jejich zavedení v archivech

Důležitým krokem k dosažení cílů projektu je postupné zavádění připravovaných postupů v jednotlivých odborných komunitách. Předpokladem pro úspěšné zavedení výsledků projektu je dosažení určité úrovně standardizace zpracování a zpřístupnění sbírek a fondů. Nástrojem v oblasti archivů jsou Základní pravidla pro zpracování archiválií (dále též „Základní pravidla“), která byla připravená v minulém roce a v tomto roce byla vydána a Ministerstvem vnitra schválena jako závazná pro použití v českých archivech.⁵ Jejich geneze, vazba na mezinárodní standardy a projekt INTERPI byla popsána v minulé průběžné zprávě, stejně jako opatření, kterými se promítl projekt INTERPI do platné archivní legislativy (zákon č. 499/2004 Sb., o archivnictví a spisové službě, ve znění pozdějších předpisů a jeho prováděcí předpisy).

Národní digitální archiv

V rámci opakované veřejné soutěže na dodávku technologií projektu Národního digitálního archivu bylo uchazečům zodpovězeno několik dotazů týkajících se popisu původců a návaznosti na projekt INTERPI (obsah zadávací dokumentace viz minulá průběžná zpráva). Časové zpoždění projektu je však již tak velké, že lze předpokládat - pokud bude docházet k dalším odvoláním uchazečů - přijetí jiného řešení celé problematiky (a pravděpodobně přímé zapojení databáze INTERPI jako integrální součásti řešení).

⁵ Metodický návod č. 1/2013 odboru archivní správy a spisové služby Ministerstva vnitra, kterým se vydávají nová Základní pravidla pro zpracování archiválií, č. j. MV-46913/AS-1/AS-2013 ze dne 17. dubna 2013

Více informací v F.8.

B.4 Testování transformace dat stávajících systému

Na základě exportů připravených dat ze souborů národních autorit probíhalo testování transformace dat.

Při transformaci bylo potřebné pro každou skupinu dat:

- určit třídu a podtřídu entit,
- vyčlenit prvky, které jsou v datové struktuře zpracovány samostatně a v původních datech jsou součástí delšího řetězce,
- odstranit interpunkci podle pravidel AACR2 vkládanou do záznamů v MARC21,
- odkazy mezi autoritami transformovat do příslušných vztahů a událostí.

Transformovány byly záznamy patřící do tříd entit:

- osoba/bytost,
- dílo/výtvar - z původních autorit autor/název, korporací (staveb), geografických autorit (mosty, tratě...),
- událost,
- korporace.

Jako nejvíce problematická se jevila transformace záznamů typu autor/dílo na záznamy entit dílo/výtvar. Komplikace způsobuje také zápis doplňků ve formě souvislého řetězce, kdy je potřebné doplněk rozdělit do zvláštních prvků v záznamu entity.

Transformována data byla nahrána do testovací databáze, která je využívána v rozhraní pro zpracování základní skupiny entit. Současná podoba záznamů již neodpovídá úplně aktuální verzi Metodiky tvorby znalostního modelu a transformace bude provedena ještě jednou s novým zadáním. Testování transformace bude pokračovat v následující etapě projektu, tak aby bylo dosaženo co největší přesnosti a úplnosti při převodu existujících záznamů.

B.5 Formulace požadavků na softwarové řešení pro doplňující skupinu entit

V tomto roce řešení projektu byly zpracovány funkční požadavky na rozhraní pro doplňující skupinu entit - tj. třídy entit: událost, dílo/výtvar a obecný pojem. Funkční požadavky vycházely z požadavků na rozhraní pro základní skupinu entit a z návrhu Metodiky tvorby znalostního modelu.

V rozhraní pro doplňující skupinu entit je kladen důraz na respektování rozdílných přístupů ve tvorbě označení pro tyto entity, tj. rozhraní musí umožňovat zápis označení podle různých pravidel. Pro třídu entit obecný pojem musí být zachována možnost zapsat deskriptor (označení) existujícího tezauru. V této třídě je také důležité poskytnout nástroje podporující proces začleňování vysoce specializovaných odborných termínů do databáze INTERPI.

B.6 Propagace a informování o projektu

Hlavními činnostmi v oblasti zpřístupňování výsledků projektu je v roce 2013:

- organizace semináře (uskuteční se 18. - 19. 11. 2013),
- prezentace projektu na konferenci Archivy, knihovny a muzea v digitálním světě
- publikování recenzované studie o projektu v periodiku Knihovna.

Seminář ke druhému roku projektu bude zaměřen na představení základní metodiky a práci s webovým rozhraním.

Kromě těchto aktivit se pozornost věnovala také zpřístupnění informací o projektu v rámci samostatného webového sídla. Byla registrována doména www.interpi.cz a vytvořen základ webového sídla. V současnosti slouží především pro přístup k Rozhraní pro zpracování základní skupiny entit a pro zveřejnění metodiky a uživatelské příručky tak, aby měl uživatel Rozhraní k dispozici aktuální informace. Na stránky metodiky v rámci webového sídla projektu jsou také odkazovány jednotlivé části Rozhraní.

C Návrhová část

C.1 Informace o uplatněných výsledcích projektu v roce 2013

C.1.1 Rozhraní pro zpracování základní skupiny entit

Rozhraní pro zpracování základní skupiny entit je vytvořeno jako webová aplikace a je určeno zpracovatelům, kteří spolupracují (případně) budou spolupracovat na vytváření znalostního modelu kulturního dědictví paměťových institucí.

Pro práci s webovým rozhraním pro zpracování základní skupiny entit je nutné respektovat, že:

- přístup k aplikaci je přes zabezpečený protokol HTTPS, tedy přenášená data jsou šifrována pomocí SSL (platnost certifikátu zabezpečuje řešitelská organizace),
- do aplikace se přistupuje na základě uživatelského jména a hesla, které je vázáno na konkrétní osobu.

Uživatel, kterému bylo přidělené uživatelské jméno, je obeznámen s používáním rozhraní a také s Metodikou tvorby znalostního modelu.⁶

Rozhraní pro zpracování základní skupiny entit je vázáno na obsah Metodiky tvorby znalostního modelu a v případě změn v této Metodice bude aktualizováno i rozhraní.

Do spuštění poloprovozu, který je výsledkem projektu INTERPI v roce 2015 jsou všechna data zpracována v rozhraní považována za testovací a není tudíž zaručena jejich správnost s ohledem na aktuální Metodiku tvorby znalostního modelu ani jejich uchování.

Uživatelská příručka je v příloze F.5.

C.1.2 Studie v recenzovaném periodiku

Začátkem roku 2013 jsme připravili studii do recenzovaného periodika Knihovna s následující anotací:

Příspěvek informuje o projektu Interoperabilita v paměťových institucích (INTERPI) DF11P010VV023 Programu aplikovaného výzkumu a vývoje národní a kulturní identity (NAKI) financovaném Ministerstvem kultury ČR.

⁶ pro potřeby posouzení rozhraní pro jeho uplatnění jako výsledku projektu INTERPI v RIV v roce 2013 jsou vytvořeny následující přístupové údaje (platné do 31. 3. 2014) - pro zpracovatele: uživatelské jméno: rivZ / heslo: riv2013Z ; pro supervizora: uživatelské jméno rivS / heslo: riv2013S

Fondy a sbírky paměťových institucí tvoří jádro národního kulturního, industriálního a přírodního dědictví. Cílem projektu INTERPI je tvorba společné ontologie a znalostního modelu odpovídajícího potřebám všech paměťových institucí a obohaceného o potřebné sémantické informace umožňující i strojové zpracování dat. Přináší také nové paradigma zpracování dat na základě objektového přístupu, které se zaměřuje na zpracování entit (tříd) a komplexních vztahů mezi nimi, a jehož cílem je zabezpečení sémantické interoperability.

Studie přináší teoretická východiska projektu a také informace o výsledcích a prostředcích pro jejich dosažení.

Příspěvek byl vydán v čísle 1 (roč. 24, 2013) na stranách 28-44. Je dostupný také na: <http://knihovna.nkp.cz/knihovna131/13128.htm>.

C.2 Informace o neuplatněných výsledcích projektu - zdůvodnění

C.2.1 Metodika tvorby znalostního modelu

V projektovém záměru bylo pro 2. etapu projektu vymezeno jako cíl vytvoření a certifikace Metodiky tvorby znalostního modelu INTERPI (dále též Metodika).

Komplexní Metodika se skládá ze dvou částí:

- základní metodiky, která obsahuje popis struktury znalostní báze INTERPI,
- jednotlivých metodik pro zpracování jednotlivých tříd entit.

Základní metodika je dokončena: v rámci projektu INTERPI bylo dosaženo dohody a definováno 7 tříd entit, které se dále dělí na podtřídy. Více informací v F.6.

Při tvorbě jednotlivých metodik pro třídy a podtřídy entit jsme narazili na problematiku pomalého vývoje mezinárodních a v návaznosti i národních standardů a pravidel a jejich národních interpretací.

Znalostní báze INTERPI je budována na objektovém principu a na bázi konceptuálních modelů. Jedná se o poměrně nové postupy a metody, jejichž vývoj není ustálený a navíc probíhá (či donedávna probíhal) odděleně v rámci jednotlivých odborných komunit; náznaky společných postupů knihovníků, muzejníků a archivářů jsou patrné až v posledních letech a jsou předmětem dosud neukončených projektů, např. vývoj technického standardu EAC.

Typickým příkladem postupného vývoje jsou konceptuální modely tvořené zpočátku pro zpracování bibliografického univerza, určené tedy primárně pro knihovnickou obec. Konceptuální model FRBR (Functional Requirements for Bibliographic Records - Funkční požadavky na bibliografické záznamy) publikovaný v roce 1997 definuje 10 základních entit (...), důležitá entita rod/rodina zde však zastoupena není. K její integraci dochází až v rámci konceptuálního modelu FRAD (Functional Requirements for Authority Data - Funkční požadavky na autoritní údaje) publikovaného v roce 2009 (poslední revize v roce 2013), který se zabývá tvorbou autoritních forem personálních jmen, jmen rodů/rodin, korporativních názvů a definováním vazeb mezi těmito entitami. Neřeší však problematiku označení a vztahů

geografických entit, které jsou spolu s problematikou „pojmu“, „objektu“ a „události“ řešeny v konceptuálním modelu FRASAD (Functional Requirements for Subject Authority Data - Funkční požadavky na předmětové autoritní údaje) publikovaném v roce 2010. Všechny uvedené modely jsou považovány za entitně relační, které definují základní entity, jejich atributy a vztahy mezi těmito entitami. Zabývají se převážně tvorbou selekčních prvků/přístupových bodů a obsahují z pohledu muzejníků a archivářů málo kontextuálních informací. Jak již bylo uvedeno, jsou určeny primárně pro knihovnickou obec a nelze je bez úprav používat např. při popisu archiválií, muzejních exponátů a kolekcí.

Na tento neuspokojivý stav reaguje archivní a muzejnická obec publikováním „svých“ standardů: archiváři publikují v roce 2004 standard pro tvorbu autoritních záznamů korporativních a personálních záhlaví ISAAR CPF (geografické autority jsou opět opominuty), protože „Archivní autoritní záznamy ... musejí podporovat mnohem širší soubor požadavků, než jaký existuje u záznamů knihovnických autorit. Tyto dodatečné požadavky jsou spojeny s významem dokumentování informací o původcích a kontextu vytváření archiválií v systémech archivního popisu. Jako takové jdou záznamy archivních autorit mnohem dále a obvykle obsahují mnohem více informací než záznamy knihovnických autorit⁷.”

V roce 2006 publikuje muzejnická obec standard CIDOC CRM (CIDOC Conceptual Reference Model) schválen jako ISO 21127:2006, který je založen na objektovém principu - definuje třídy entit a jejich vlastnosti; v přímé návaznosti na tuto normu vzniká a je v letech 2006 a 2008 publikována objektově orientovaná verze modelu FRBR, tedy FRBRoo.

Podobně je tomu i při tvorbě pravidel týkající se zápisu formy označení/pojmenování jednotlivých typů entit. Pravidla AACR (Anglo-American Cataloguing Rules) vznikla v roce 1967 v přímé návaznosti na výměnný formát MARC. V roce 1978 byla publikována revidovaná verze AACR2R, která se prakticky používá (u nás) dodnes. O vývoji nových pravidel se jedná od roku 1997, v roce 2004 bylo rozhodnuto, že pravidla AACR2 nebudou aktualizována a budou se vyvíjet pravidla nová - RDA (Resource Description and Access), která budou odpovídat konceptuálním modelům a potřebám webu, v současné době sémantického webu. Bylo přislíbeno, že v roce 2009/2010 bude publikována 1. verze těchto pravidel. Pravidla byla publikována a dána k dispozici až v květnu 2013. Navíc tato verze stále neobsahuje řešení pro zápis entit „událost“, „pojem“ a „objekt“. Jde o složitou problematiku, jejíž řešení není jednoduché, protože se částečně odvíjí od existence jednotlivých národních jazyků. Tuto skutečnost jsme předvíдали a připravili pro projekt INTERPI dílčí řešení v intencích pravidel RDA, např. zápis označení událostí typu bitvy, války apod.

Pravidla pro tvorbu a zápis řízených selekčních prvků - AACR2R i nově implementovaná pravidla RDA jsou důležitá nejenom pro knihovnickou obec; představují důležitý standard i pro muzejnickou obec: pravidla CCO (Cataloguing of Cultural Objects) přímo vycházejí z těchto pravidel. Nepřímo na ně navazují i standardy určené pro archivní komunitu, i když tato se na ně neodvolává přímo. Ve standardu ISAAR CPF se uvádí „Tento standard byl vytvořen pro sdílení archivních autoritních záznamů na podporu přípravy konzistentních, vhodných a samo

⁷ http://asta.nacr.cz/data/fil_1831.pdf

vypovídajících popisů korporací, osob a rodů, které vytvářejí dokumenty. Byl vytvářen tak, aby mohl být použit jako doplněk existujících národních standardů nebo jako základ pro vývoj národních standardů⁸. K nejrozšířenějším národním standardům určených pro tvorbu popisných údajů patří právě pravidla AACR2R a nově implementovaná pravidla RDA. I nově vznikající národní pravidla, jako např. „Základní pravidla pro zpracování archiválií“⁹ v části „Tvorba přístupových rejstříků a přístupových bodů“ v podstatě vycházejí z principů pravidel AACR2 či RDA.

Nyní je potřeba připravit národní interpretace pravidel RDA. Byla založena pracovní skupina pro implementaci pravidel RDA v Česku, jejímž členem je i řešitel INTERPI, byl připraven harmonogram přechodu na aplikaci těchto pravidel: bylo rozhodnuto, že v Česku začneme používat pravidla RDA v květnu v roce 2015.

Podstatou metodiky jednotlivých tříd entit je definování preferované a variantních forem daných entit. Tyto formy jména však nelze stanovit bez rámcových pravidel, kterými jsou pravidla RDA.

Z těchto objektivních důvodů nemohla být dokončena tvorba dílčích metodik jednotlivých tříd entit. Tato část metodiky bude dokončena v roce 2014 a metodika bude v tomto roce předložena k certifikaci. Předpokládáme však, že vazby mezi paměťovými institucemi, které vznikly při tvorbě Metodiky, zůstanou aktivní i po jejím dokončení a v případě nových požadavků ze strany kterékoliv komunity, proběhne společná diskuse, kterou považujeme za základní podmínku interoperability. Pro publikaci Metodiky chceme využít progresivních nástrojů a publikovat ji v elektronické podobě. Zabezpečí se tím větší interaktivita s budoucími požadavky odborné komunity.

Metodiku již teď průběžně aktualizujeme v rámci webového sídla www.intepri.cz tak, aby se o ní mohli opírat zpracovatelé při testování Rozhraní pro zpracování základní skupiny entit.

Aktuální verze Metodiky tvorby znalostního modelu tvoří přílohu F.6.

C.3 Cíle projektu pro rok 2014

Pro následující třetí etapu projektu - tj. rok 2014 - byly v projektové dokumentaci stanoveny tyto cíle:

1. mapování stávajících zdrojů na základní skupinu entit,
2. transformace dat ze stávajících systémů,
3. testování software pro základní skupinu entit,
4. ověřování software pro druhou skupinu entit v průběhu jeho přípravy,
5. zpracování metodiky mapování řízených slovníků aplikovaných v paměťových institucích.

⁸ http://asta.nacr.cz/data/fil_1831.pdf

⁹ <http://www.mvcr.cz/clanek/metodiky.aspx>

Vzhledem k činnostem, které probíhaly a které jsou vysvětleny v části B, je potřeba upravit také cíle pro rok 2014 tak, aby aktivity projektu na sebe vhodně navazovaly:

1. dokončení Metodiky tvorby znalostního modelu INTERPI a její podání k certifikaci - metodika bude uplatnitelná jako výsledek do RIV pro rok 2014;
2. doplnění současného webového rozhraní o rozhraní pro zpřístupnění zpracovaných dat pro širokou veřejnost s ohledem na strukturu dat vymezenou Metodikou tvorby znalostního modelu, včetně definování a analýzy požadavků na rozhraní pro zpřístupnění dat ze strany paměťových institucí;
3. zpracování modelu kooperace paměťových institucí při vytváření znalostního modelu INTERPI;
4. analýza, ověřování, doplňování a úprava dat;
5. propagace projektu - v oblasti propagace projektu a informování spolupracovníků projektu o dosažených výsledcích bude v roce 2014 realizován seminář a naplněna samostatná stránka projektu.

D Použití účelové podpory v členění dle rozpočtu projektu

E Resumé

Druhá etapa projektu navazuje na výsledky základního výzkumu z roku 2012 a hlavní pozornost věnuje aplikovanému výzkumu a vývoji. Hlavním výsledkem je rozhraní pro zpracování základní skupiny entit, které navazuje na požadavky formulované při tvorbě a testování prototypu v roce 2012 a požadavky vzniklé při přípravě Metodiky tvorby znalostního modelu. Rozhraní je řešeno jako webová aplikace, což umožňuje zpracovávání entit odkudkoliv a kdykoliv podle možností zpracovatelů. Současně s webovým rozhraním bylo vytvořeno i webové sídlo projektu - www.interpi.cz, jehož cílem je integrovat všechny výsledky projektu. V současné podobě obsahuje průběžně aktualizovanou Metodiku tvorby znalostního modelu, manuál pro práci s rozhraním, základní informace o projektu a samozřejmě odkaz na Rozhraní pro zpracování základní skupiny entit.

Poznatky získané v základním výzkumu jsme v této etapě transformovali do návrhu Metodiky tvorby znalostního modelu, která bude s přihlédnutím na situaci ve vývoji klíčových mezinárodních standardů a přípravě jejich národních interpretací, dokončena v následujícím roce a předložena k certifikaci. Cílem metodiky není kopírovat aktuální pravidla a standardy využívané v paměťových institucích, cílem je různá pravidla sjednocovat a propojovat tak, aby ve znalostním modelu INTERPI mohly koexistovat především ke spokojenosti všech jeho potenciálních uživatelů. Za hlavní přínos metodiky (i když v této neukončené verzi) považujeme, že v mnohých sporných bodech bylo kompromisu dosaženo. Některé problémové okruhy (např. dílo/výtvar, rod/rodina, událost) jsou takto komplexně a s ohledem na dosavadní praxi všech odborných komunit řešeny poprvé.

Pokračovaly taktéž práce na kontrole a přípravě dat, přičemž některé již byly letos využity pro testování transformace dat ze současných systémů.

V oblasti propagace projektu se pozornost soustředila kromě přípravy zmíněného webového sídla na přípravu semináře, který se bude konat 18. - 19. 11. 2013. Projekt byl (a do konce roku ještě bude) prezentován na odborných akcích. O projektu byla také publikována studie v recenzovaném periodiku.

Uplatněné výsledky - druh výsledku podle RIV:

R (software): INTERPI. *Rozhraní pro zpracování základní skupiny entit*. Software.

W (uspořádání workshopu): INTERPI. *Seminář projektu INTERPI*. 18. 11. 2013 - 19. 11. 2013. Brno. Pozn.: výsledek bude uplatněn po uskutečnění semináře.

J_{rec} (recenzovaný odborný článek): BALÍKOVÁ, Marie - KUNT, Miroslav - ANDREJČÍKOVÁ, Nadežda - ŠUBOVÁ, Jana. INTERPI: nástroj pro zpřístupnění národního kulturního dědictví. In *Knihovna*. 2013, roč. 24, č. 1, s. 28-44. Dostupné též na: <http://knihovna.nkp.cz/knihovna131/13128.htm>.

D (článek ve sborníku): BALÍKOVÁ, Marie - ANDREJČÍKOVÁ, Nadežda - KUNT, Miroslav - ŠUBOVÁ, Jana. Znalostní báze pro paměťové instituce v Česku: projekt INTERPI v roce 2013. In *INFOS 2013 : zborník príspevkov z 37. medzinárodného informatického sympózia*. Bratislava : Spolok slovenských knihovníkov, 2013. ISBN 978-80-89586-07-3, s. 12-25.

F Přílohy

- F.1 - Seznam fyzicky předložených výsledků projektu
- F.2 - Použití účelové podpory v členění dle rozpočtu projektu - komentář
- F.3 - Dodatky, rozhodnutí
- F.4 - Cestovní zprávy
- F.5 - Rozhraní pro zpracování základní skupiny entit - uživatelská příručka
- F.6 - Metodika tvorby znalostního modelu
- F.7 - Zpráva o činnosti Národní knihovny v projektu INTERPI za rok 2013
- F.8 - Zpráva o činnosti Národního archivu v projektu INTERPI za rok 2013

F.1 Seznam fyzicky předložených výsledků projektu

F.2 Použití účelové podpory v členění dle rozpočtu projektu - komentář

F.3 Dodatky, rozhodnutí

F.4 Cestovní zprávy

F.5 Rozhraní pro zpracování základní skupiny entit - uživatelská příručka

F.6 Metodika tvorby znalostního modelu

F.7 Zpráva o činnosti Národní knihovny v projektu INTERPI za rok 2013

F.8 Zpráva o činnosti Národního archivu v projektu INTERPI za rok 2013