

ZPRÁVA O VÝSLEDKÁCH ŘEŠENÍ PROJEKTU V ROCE 2014

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

Mgr. Marie Balíková
odpovědná řešitelka - koordinátorka
Národní knihovna ČR

Ing. Tomáš Böhm
generální ředitel
Národní knihovna ČR



INTERPI

15. 11. 2014

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	Aplikace souboru národních autorit jako terminologického základu znalostní báze INTERPI	4
B.1.2	Aplikace souboru národních autorit v procesu mapování/harmonizace terminologických zdrojů	4
B.1.3	Úprava záznamů Národních autorit v souladu s budoucími požadavky znalostní báze INTERPI	5
B.1.4	Příprava záznamů Národních autorit, export pro transformaci dat	6
B.1.5	Úprava, doplňování a analýza entit pro potřeby archivů	6
B.1.6	Aktualizace klasifikačního systému MDT	8
B.2	Příprava metodických a metodologických postupů, jejich zavedení v archivech	8
B.3	Transformace dat stávajících systémů	8
B.4	Ověřování postupů mapování a harmonizace terminologických zdrojů	10
B.5	Formulace funkčních požadavků na softwarové řešení zpřístupnění entit a prototyp rozhraní	10
B.6	Propagace a informování o projektu	11
C	Návrhová část	12
C.1	Informace o uplatněných výsledcích projektu v roce 2014	12
C.1.1	Metodika tvorby znalostního modelu	12
C.1.2	Metodika mapování a harmonizace rejstříků, číselníků a řízených slovníků aplikovaných v paměťových institucích	12
C.2	Informace o neuplatněném výsledku projektu v roce 2014 - zdůvodnění	13
C.2.1	Rozhraní pro zpracování druhé (doplňující) skupiny entit	13
C.3	Cíle projektu pro rok 2015	14
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 Metodika tvorby znalostního modelu (včetně oponentských posudků)
- F.6 Metodika mapování a harmonizace rejstříků, číselníků a řízených slovníků aplikovaných v paměťových institucích (včetně oponentských posudků)
- F.7 Zpráva o činnosti Národní knihovny v projektu INTERPI za rok 2014
- F.8 Zpráva o činnosti Národního archivu v projektu INTERPI za rok 2014

A Konstatační část

A.1 Rešerše

BALÍKOVÁ, Marie - KUNT, Miroslav. Projekt INTERPI v roce 2013 : entity sem, entity tam, entity všude, kam se podívám. In *14. konference Archivy, knihovny, muzea v digitálním světě*. 2013. Příspěvek na konferenci. Dostupné na: <http://bulletin.skipcr.cz/prezentace/archivy-2013/Balikova.pdf>.

ŠUBOVÁ, Jana - ANDREJČÍKOVÁ, Nadežda. Projekt INTERPI v roce 2013 : ten zapisuje to, ta zas tohle a všichni dohromady uděláme moc. In *14. konference Archivy, knihovny, muzea v digitálním světě*. 2013. Příspěvek na konferenci. Dostupné na: <http://bulletin.skipcr.cz/prezentace/archivy-2013/Subova.pdf>.

BALÍKOVÁ, Marie - KUNT, Miroslav - ŠUBOVÁ, Jana - ANDREJČÍKOVÁ, Nadežda. Interoperability in Memory Institutions in Czechia: INTERMI Project. In *URLA 2014 : international congress on management of cultural heritage and cultural memory institutions*. 2014. Příspěvek na konferenci. Dostupné na: <http://bulletin.skipcr.cz/prezentace/archivy-2013/Balikova.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 2014 je v projektu vymezen jako třetí etapa. Navázali jsme na výsledky předešlých etap a s ohledem na cíle stanovené v projektovém záměru pro druhou etapu jsme se v rámci aplikovaného výzkumu a vývoje zaměřili na:

1. dokončení metodiky tvorby znalostního modelu v projektu INTERPI a předložení k certifikaci,
2. vytvoření metodiky mapování a harmonizace rejstříků, číselníků a řízených slovníků aplikovaných v paměťových institucích a předložení k certifikaci,
3. definování požadavků na software pro zpřístupnění entit,
4. vytvoření prototypu webového rozhraní pro zpřístupnění entit,
5. ověření prototypu software (webového rozhraní) pro zpracování druhé skupiny entit, testování webového rozhraní pro zpracování první skupiny entit a jeho doplnění podle metodiky tvorby znalostního modelu,
6. mapování stávajících zdrojů na základní skupinu entit,
7. transformace dat ze stávajících systémů, testování nástrojů harmonizace rejstříků, číselníků a řízených slovníků.

B Analytická část

Analytická část zprávy obsahuje informace o dílčích výsledcích dosažených v roce 2014 v oblasti aplikovaného výzkumu a vývoje v souladu s cíli projektu. I v této etapě jsme se zaměřili na analýzu, ověřování a úpravu dat v stávajících systémech - především pak v souboru národních autorit. Dále byly ověřována a doplňována data pro potřeby archivů.

B.1 Analýza, ověřování, doplňování a úprava dat

V roce 2014 NKČR pokračovala v pracích spojených s analýzou, exportem a úpravu autoritních údajů jako základního předpokladu pro úspěšný vývoj znalostní databáze INTERPI. Byla vypracována studie o možnosti aplikace souboru národních autorit jako terminologického základu znalostní báze INTERPI. Zvláštní pozornost byla věnována analýze a úpravě tematických autorit a autoritních termínů označujících formu/žánr.

B.1.1 Aplikace souboru národních autorit jako terminologického základu znalostní báze INTERPI

Knihovnické obci slouží soubor národních autorit jako prostředek k unifikaci selekčních prvků bibliografických záznamů a jako takový má svou vnitřní strukturu a určité množství konkrétních informací, odpovídající katalogizačním pravidlům AACR2 od roku 2015 katalogizačním pravidlům RDA.

V souvislosti s řešením Metodiky mapování a harmonizace rejstříků, číselníků a řízených slovníků aplikovaných v paměťových institucích jsme analyzovali možnosti využití souboru národních autorit (především dílčího souboru tematických autorit) jako terminologického základu znalostní báze INTERPI. Soustředili jsme se na strukturu autoritního záznamu a možnost využití údajů z autoritního záznamu pro vytvoření záznamu entity v báze INTERPI a také na pravidla, které se na výběr termínu aplikují.

Výsledky analýzy umožnily formulovat pravidla výběru preferovaných termínů ve zmiňované Metodice mapování a harmonizace rejstříků, číselníků a řízených slovníků a zároveň definovat strukturu údajů entity obecný pojem v Metodice tvorby znalostního modelu.

Více informací v příloze F.7.

B.1.2 Aplikace souboru národních autorit v procesu mapování/harmonizace terminologických zdrojů

Všechny složky souboru autorit, tedy jednotlivé dílčí soubory, se budou podílet na harmonizaci rejstříků, které budou obsahovat označení všech entit, tj. běžně používané typologicky nerozlišené indexy klíčových slov či jednotlivých segmentů řetězců předmětových hesel.

Filtrací těchto rejstříků obsahujících všechny termíny pomocí dílčích souborů NAČR získáme oddělené soubory potenciálních typologicky rozlišených deskriptorů: soubory personálních jmen, korporativních, soubor geografických termínů, obecných pojmů, atd.

Role Souboru věcných tematických autorit v procesu mapování/harmonizace terminologických zdrojů

1. Soubor věcných tematických autorit je univerzální - jsou v něm zahrnuty termíny, případně terminologické soustavy ze všech oborů lidského poznání. Jednotlivé termíny jsou provázány s ekvivalentní notací MDT, která umožní expanzi do jednotlivých sémantických domén a zároveň usnadní tvorbu mikrotezaurů v jednotlivých vědních oborech;
2. Soubor věcných tematických autorit obsahuje aktuální terminologii; vytváří se induktivní metodou (metodou zdola), jednotlivé termíny jsou do něj vřazovány podle potřeb zpracovatele. Na jedné straně to představuje úskalí, protože SVAT neobsahuje všechny termíny jednotlivých terminologických soustav všech oborů. Na druhé straně to přináší velké výhody. Do souboru tematických autorit jsou neustále vřazovány jako preferované či nepreferované formy termínů současně používané termíny. Můžeme tedy konstatovat, že terminologie používaná v rámci SVAT je aktuální;
3. Soubor věcných tematických autorit používá víceslovné deskriptory - v praxi to znamená, že jsou používány termíny, které jsou specifické a detailní do té míry, že reprezentují daný (nově se objevivší) pojem adekvátním způsobem a zasazují ho tak do sémantického rámce daného vědního oboru;
4. Autoritní termíny v souboru tematických autorit jsou propojeny s ekvivalentní notací MDT, což v praxi umožní zařazení i specifických termínů používaných jednotlivými paměťovými institucemi, které tak mohou být federativní metodou navázány na obecnější termín, který tvoří součást univerzálního souboru;
5. Autoritní termíny jsou opatřeny anglickými ekvivalenty, což v praxi podpoří aplikaci tohoto souboru na mezinárodní úrovni.

Více informací v příloze F.7.

B.1.3 Úprava záznamů Národních autorit v souladu s budoucími požadavky znalostní báze INTERPI

Úprava souboru tematických autorit

V roce 2014 byly provedeny následující úpravy v souboru tematických autorit: do autoritních záznamů byly připojeny invertované tvary víceslovných termínů. Byl také vytvořen soubor obsahující označení kategorie osob.

Úprava souboru geografických autorit

V souboru geografických autorit bylo v roce 2014 dořešeno propojení této databáze s Databází sídelních lokalit Čech, Moravy a Slezska cz_retro pomocí pole 052, kde se uvádí kód cz_retro. Byla vytvořena aplikace, která automaticky doplní kód cz_retro v poli 052 a propojí tak GEO termíny s bází cz-retro s pomocí hierarchické struktury uvedené v poli 951; takto lze postupovat v případech jednoznačných termínů. V záznamech geografických autorit, ve kterých nedošlo - z důvodů rozdílnosti údajů - k automatickému propojení na bázi údajů v poli 951, se postupně doplňuje pole 052 intelektuálně.

Více informací v příloze F.7.

B.1.4 Příprava záznamů Národních autorit, export pro transformaci dat

Podle kritérií definovaných v roce 2013, na základě kterých byly exportovány záznamy Národních autorit pro testování transformace dat, byl i v této etapě vytvořen export stávajících autorit. Bylo exportováno:

- 589764 personálních autorit označujících osoby,
- 14635 autorit označujících kombinace jména aurora a název jeho díla,
- 94782 korporativních záhlaví,
- 3489 kombinací korporativních záhlaví a názvů díla,
- 26250 geografických autorit,
- 37132 autorit akcí,
- 2339 autorit unifikovaných názvů.

B.1.5 Úprava, doplňování a analýza entit pro potřeby archivů

Tvorba geografických entit typu sídlo

V roce 2014 pokračovaly práce na pořizování záznamů o entitách typu sídlo, zejména jejich historických názvů.¹ Kompletně je zrevidován soupis obcí na Slovensku,² do konce roku 2014 bude připravena databáze sídel Moravy a Slezska (v současných státních hranicích).³ V případě Čech bude do konce roku 2014 dokončen soupis, revize proběhne v roce následujícím tak, aby všechna data mohla být importována do databáze INTERPI. Při prezentacích byly zaznamenány příznivé ohlasy na tvorbu zmiňovaných dat, několik pracovišť o ně projevilo zájem.

Více informací v příloze F.8.

Entita dílo/výtvor

I s ohledem na úkoly projektu v roce 2014 byla věnována zvýšená pozornost mimořádně komplikovanému popisu entit dílo/výtvor. Složitost spočívá především v široké paletě a různorodosti podtříd této entity. V souvislosti s tvorbou Metodiky tvorby znalostního modelu INTERPI byl na příkladech zkoušen popis různých entit: staveb, liniových staveb a tras, právních dokumentů. Na základě toho byla metodika následně doplňována a upravována.

Další zkoumanou skupinou entit dílo/výtvor byly právní normy. Byl zpracován komplexní rozbor „Právní dokument jako entita dílo/výtvor“. U většiny pramenů práva lze určit územní jurisdikci (působnost). Problematická je tam, kde se již nemusí jednat vždy o normativní pramen práva

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 Vlastivedný slovník obcí na Slovensku, Veda, vydavateľstvo Slovenskej akadémie vied, Bratislava : 1977-1978, díl I - III

³ Hosák, c.d.

(vnitrostátní smlouvy), kde jurisdikce může být vyjádřena jinak (korporací) nebo je násobná (mezinárodní smlouvy). Prakticky vždy je možné určit vydavatele, kterých však může být více. Naproti tomu zdroj zveřejnění, ač se zprvu zdál bezproblémovým, nemusí být u některých pramenů práva jasný nebo dohledatelný.

Druhým problémem je vlastní citace pramene práva. Citace používaná v NAČR se radikálně liší od řešení běžného v právní a správní praxi, ale odchyluje se i od běžného citačního úzu.⁴

Zároveň je potřeba vyrovnat se s novelizacemi. Např. není únosné citovat „Zákon č. 223/2013 Sb., kterým se mění zákon č. 258/2000 Sb., o ochraně veřejného zdraví a o změně některých souvisejících zákonů, ve znění pozdějších předpisů“, když se jedná pouze o (jednu) dílčí novelizaci, která navíc pro vyhledání aktuálního znění postrádá smysl.

Na základě výše uvedeného rozboru byla navržena pravidla INTERPI pro zápis unifikovaných názvů normativních pramenů práva, judikátů a interních úředních dokumentů.

Více informací v příloze F.8.

Entita korporace

Práce na shromáždění dat o korporacích, klíčovém příspěvku archivů do projektu, pokračovaly několika směry:

1. strukturovaný popis současných korporací ústřední státní správy
2. vytvoření databáze názvů mezinárodních organizací podle publikace Klepacki, Zbigniew M., ed. Slovník mezinárodních organizací. 1. vyd. Bratislava: Pravda, 1979. 756 s. (ukázalo se, že pro vytváření vztahů mezi korporacemi jsou mezinárodní organizace důležité)
3. Popis původců archivního fondu okresního archivu (typové korporace - obecní úřady a národní výbory, školy, okresní soudy, okresní úřady, farnosti atd.)
4. V oblasti popisu odvětví byly sledovány dvě linie (s ohledem na zapojené odborníky):
 - a) vývoj korporací z oblasti zestátněné kinematografie (navazuje na loňskou práci)
 - b) vývoj korporací z oblasti hospodářství, zejména dopravy

V případě bodu 4 se předpokládá dokončení výstupu a jeho publikování.

Více informací v příloze F.8.

Entita osoba

K publikování byly připraveny Slovník představitelů soudní správy v Čechách a pro databázi INTERPI připraveny údaje z rejstříku edice Vojáček, Milan, ed., Marie Červinková-Riegrová, Zápisky II (1885-1886), Praha, 2013. S ohledem na relativně jednoduchou a již propracovanou metodiku popisu (fyzických) osob nebyla tato činnost pro Národní archiv zásadní, úkolem řešitelů bylo však prověřit prameny pro zjišťování osobních údajů a možnosti jejich zpracování.

⁴ ČSN ISO 690:2011 Bibliografické citace

B.1.6 Aktualizace klasifikačního systému MDT

V tomto roce (2014) 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 2014 pokračovaly práce na vývoji vícejazyčné překladové databáze MDT - UDC online.

Více informací v příloze F.7.

B.2 Příprava metodických a metodologických postupů, jejich zavedení v archivech

Základní pravidla pro zpracování archiválií (dále též „Základní pravidla“) vydaná v roce 2014 Ministerstvem vnitra a schválená jako závazná pro použití v českých archivech⁵ byla prezentována na „roadshow“ jednotlivým veřejným archivům čtyřčlennou skupinou, z níž tři osoby jsou zapojeny do projektu INTERPI. Samostatná část prezentací byla věnována popisu původců archiválií a tvorbě přístupových bodů s ohledem na projekt INTERPI. Lze říci, že všichni archiváři v České republice jsou s projektem INTERPI seznamováni a nastavené metodické pokyny ho činí v této komunitě již nezbytným.

Národní digitální archiv

V souladu s celkovou koncepcí projektu INTERPI byly zpracovány podklady pro zadání archivního portálu s ohledem na popis původců archiválií a komunikaci software pro zpracování archiválií. Financování projektu ze evropských strukturálních fondů bylo však s ohledem na problémy v dalších veřejných soutěžích (rozhodování Úřadu pro ochranu hospodářské soutěže a nedostatek času) ukončeno. V projektu digitálního archivu Národního archivu je pokračováno vlastními silami tak, aby byly naplněny zákonné povinnosti (mj. i popis původců). Klíčovým prvkem je definice rozhraní mezi databázemi INTERPI a jinými systémy.

Více informací v příloze F.8.

B.3 Transformace dat stávajících systémů

V návaznosti na testování transformace dat stávajících systémů z předcházející etapy projektu a v souladu s dokončením metodiky tvorby znalostního modelu byla provedena transformace údajů ze souboru národních autorit.

Při transformaci bylo nutné pro každou skupinu dat:

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

- určit třídu a podtřídu entit,
- vyčlenit prvky, které jsou v datové struktuře zpracovány samostatně a v původních datech jsou součástí delšího řetězce - například sídlo korporace, datace korporace, doplňky různých typů při původních geografických entitách,
- odstranit interpunkci podle pravidel AACR2 vkládanou do záznamů v MARC21,
- odkazy mezi autoritami transformovat do příslušných vztahů a událostí (např. datum narození na událost narození, geografický doplněk korporace na sídlo korporace).

Transformovány byly záznamy patřící do tříd entit:

- osoba/bytost - z původních osobních autorit,
- dílo/výtvor - z původních autorit autor/název, korporace/název, korporací (staveb - hrady, zámky, kostely, kaple apod.), geografických autorit (mosty, tratě, naučné stezky, technické památky, vodní nádrže apod.) a unifikovaných názvů,
- událost - z původních autorit akcí,
- korporace.

Zadání pro transformaci dat z roku 2013 bylo nutné přepracovat, tak aby výsledek byl v souladu s metodikou tvorby znalostního modelu v její podobě předložené k certifikaci. Toto se především týkalo entit dílo/výtvor, kde byl v metodice vyřešen soubor typů částí označení a doplňků.

Entity dílo/výtvor tak představovaly z hlediska transformace dat nejnáročnější skupinou údajů. Kromě jiného jsme při transformaci údajů narazili na to, že v původních autoritách typu autor/název, korporace/název, nebo unifikovaný název se uvádí jinojazyčná (originální) forma názvu, ovšem bez uvedení jazyka, což snižuje uplatnění takové informace.

V třídě entit osoba/bytost je testována metoda pro převod informací z biografické poznámky na události narození a úmrtí. Na základě strukturalizace informací v biografické poznámce je možné vybrat informace o datu a místě narození nebo úmrtí. Výhodou je důsledná strukturalizace biografické poznámky, nevýhodou ale je nestandardní zápis názvů míst (upřesnění místa uvedením většího města po předložce „u“), skloňování názvů míst, zkracování názvů (místo „nad“ pouze „n.“) a používání výrazů „tamtéž“ v případě shodného místa narození a úmrtí. Z výsledků testování této metody vyplývá, že ve čtvrtině záznamů o entitách osoba/bytost, které vznikly transformováním souboru národních autorit, je možné v biografické poznámce identifikovat údaj o narození a v desetině záznamů údaj o úmrtí. Metodu pro převod informace z biografické poznámky na události narození a úmrtí je možné použít nezávisle na transformaci dat (tj. i dodatečně).

Transformované údaje budou importovány do testovací databáze, která je využívána v rozhraní pro zpracování základní skupiny entit, a budou zpřístupněny ve webovém rozhraní pro zpřístupnění entit.

B.4 Ověřování postupů mapování a harmonizace terminologických zdrojů

Pro potřeby zpracování Metodiky mapování a harmonizace rejstříků, číselníků a řízených slovníků aplikovaných v paměťových institucích byla vytvořena desktopová aplikace s využitím relačního databázového systému pro testování metod porovnávání termínů z terminologických zdrojů. Prvotní výsledky průzkumu pro ověření základních algoritmů podobnosti jsou zpracovány v studii, která je součástí přílohy F.7.

Ověření algoritmů podobnosti jsme provedli pomocí transformace získaných korpusů číselníků, slovníků a tezurů používaných v našem prostředí pro zpracování objektů kulturního dědictví. Následně jsme vytvořili prototypovou aplikaci pro porovnání těchto zdrojů se souborem věcných autorit. Samotné srovnání bylo z časového hlediska realizováno na lexikální úrovni. Výsledek tohoto porovnání, což je patrné i v příloze, potvrzuje skutečnosti, které jsme uvedli v metodice a které se týkaly předzpracování vstupních dat za účelem zvýšení automatizovaného způsobu harmonizace.

Pro účely porovnání dat jsme zvolili data z těchto systémů:

1. **Authority** - byly brány jako základ pro srovnávání a s těmito daty jsme srovnávali data ze všech níže uvedených zdrojů. Ze souboru Národních autorit jsme vybrali, jak jsme již uvedli, jenom tematické termíny. Data autorit sestávala ze samotného unifikovaného termínu a všech termínů, které k němu mají nějaký vztah. Brali jsme termíny, které jsou v záznamech autorit uvedeny v polích MARC21 1XX, 4XX, 5XX.
2. **DEMUS** - tento systém je poměrně značně zastoupen v oblasti muzeí. Jednalo se o 2 tezaury a 4 číselníky:
 - Demus_VUM_materiály,
 - Demus_VUM_specifikace_materiálu,
 - Demus_VUM_specifikace_techiky
 - Demus_VUM_techiky.Nadřazený termín obsahovalo 1790 termínů. Všechny termíny obsahovaly také anglický a německý překlad.
3. **LANius, Clavius** - systémy zastoupeny především ve veřejných knihovnách, kde používají jednoduchá klíčová slova bez jakýchkoliv hierarchických vazeb.
4. **PSH - Polytematický strukturovaný heslář**, který spravuje Národní technická knihovna a je používán i jinými organizacemi v různých systémech.

Více informací v příloze F.7.

B.5 Formulace funkčních požadavků na softwarové řešení zpřístupnění entit a prototyp rozhraní

Pro vytvoření softwarového řešení týkajícího se zpřístupnění entit byly formulovány následující základní funkční požadavky:

- dostupnost rozhraní bez potřeby přihlášení - webové rozhraní,

- vyhledávání pomocí jednoduchého zadání vyhledávacího termínu (jednoduché vyhledávání)
- vyhledávání pomocí kombinace vyhledávacích termínů (rozšířené vyhledávání),
- vyhledávání v jednotlivých třídách entit, ve vybraných třídách entit i ve všech třídách současně,
- zobrazení výsledků ve zkrácené podobě,
- zobrazení výsledků v detailu - rozdělení informací na logické části, rozbalování a skrývání informací,
- upřesnění vyhledávání pomocí fazet,
- odkazování mezi entitami,
- služby pro uživatele: odeslání informace o entitě e-mailem, vytištění / uložení informací o vybraných entitách.

Konkretizace těchto základních funkčních požadavků bude předmětem programu semináře a další odborné diskuse; pro tento účel byl vytvořen jednoduchý prototyp, který prezentuje základní funkčnost rozhraní a poskytuje konkrétnější představu o jeho možnostech.

B.6 Propagace a informování o projektu

Hlavními činnostmi v oblasti zpřístupňování výsledků projektu je v roce 2014:

- organizace semináře (uskuteční se 19. - 10. 11. 2014),
- prezentace projektu na konferenci Archivy, knihovny a muzea v digitálním světě,
- prezentace projektu na mezinárodní konferenci URLA v Istanbulu.

Seminář ke třetí etapě projektu bude zaměřen na představení Metodiky tvorby znalostního modelu a Metodiky mapování a harmonizace řízených slovníků v té podobě, v jaké byly předloženy k certifikaci. Dále na práci s webovým rozhraním pro zpracování entit (informace o formulářích pro doplňující skupinu entit) a diskuzi o funkčních požadavcích na rozhraní pro zpřístupnění entit a jeho prototypu.

Na webové sídlo projektu byly doplněny pracovní verze obou metodik. V následujícím období budou na webové sídlo soustředěny všechny materiály související s projektem (prezentace o projektu a pod.), tak aby sídlo sloužilo i po skončení projektu jako základní informační zdroj o projektu.

C Návrhová část

C.1 Informace o uplatněných výsledcích projektu v roce 2014

C.1.1 Metodika tvorby znalostního modelu

V tomto roce byla dokončena Metodika tvorby znalostního modelu a předložena k certifikaci.

Základem Metodiky tvorby znalostního modelu INTERPI je vymezení 7 tříd entit a jejich rozdělení do podtříd. Z požadavků odborných komunit na rozsah informací uchovávaných o entitách z jednotlivých tříd vyplynula potřeba zpracovat podrobněji doporučení pro výběr, zpracování a uchování těchto informací. Doporučení, která by neomezovala současnou praxi paměťových institucí, ale naopak rozšiřovala ji s ohledem na rozvoj informačních technologií a potřeby uživatelů.

Předkládaná Metodika obsahuje doporučení pro jednotlivé třídy a podtřídy entit. Při jejím zpracování jsme narazili na problematiku pomalého vývoje mezinárodních a v návaznosti i národních standardů a pravidel a jejich národních interpretací. S tímto problémem se setkávají všechny obdobné projekty. Protože je cílem projektu INTERPI vytvořit model budoucí kooperace různých paměťových institucí snažili jsme se při přípravě metodiky minimalizovat vytváření proprietárních pravidel, která by bylo nutno v budoucnu upravovat. Zvolili jsme cestu hledání kompromisů a vytvoření otevřeného systému, ve kterém lze uplatnit různé přístupy, pokud není možno nalézt shodu mezi odbornými komunitami.

Metodiku průběžně aktualizujeme v rámci webového sídla www.intepri.cz tak, aby poskytovala zpracovatelům informace potřebné při testování Rozhraní pro zpracování entit. Metodika bude až do její certifikace zpřístupněna v rámci webového sídla projektu pouze registrovaným uživatelům.

Metodika tvorby znalostního modelu včetně oponentských posudků tvoří přílohu F.6.

C.1.2 Metodika mapování a harmonizace rejstříků, číselníků a řízených slovníků aplikovaných v paměťových institucích

V tomto roce byla vytvořena Metodika mapování a harmonizace rejstříků, číselníků a řízených slovníků aplikovaných v paměťových institucích a předložena k certifikaci.

V předkládané Metodice jde především o pochopení a respektování rozmanitosti. Poskytuje přehled metod, které při mapování a harmonizaci terminologických zdrojů lze využít a rovněž doporučený postup pro realizaci konkrétních záměrů v této oblasti. V konečném důsledku je také doporučením jak postupovat při integraci specifických terminologických zdrojů do znalostní databáze INTERPI a to i v případě, že tyto terminologické zdroje doposud nebyly reprezentovány pomocí žádného technologického nástroje.

I tato metodika je zveřejněna na webovém sídle www.intepri.cz, aby zpracovatelé měli k dispozici informace potřebné při testování Rozhraní pro zpracování entit (především entit ze třídy obecný pojem). Metodika bude až do její certifikace zpřístupněna v rámci webového sídla projektu pouze registrovaným uživatelům.

Metodiky mapování a harmonizace rejstříků, číselníků a řízených slovníků aplikovaných v paměťových institucích tvorby znalostního modelu včetně oponentských posudků tvoří přílohu F.6.

C.2 Informace o neuplatněném výsledku projektu v roce 2014 - zdůvodnění

C.2.1 Rozhraní pro zpracování druhé (doplňující) skupiny entit

Podle projektové dokumentace měl být předložen jako uplatněný výsledek R- software pro druhou skupinu entit. Při předkládání projektu jsme - vzhledem k nedokončenosti, nehotovosti mezinárodních standardů a pravidel - předpokládali, že výsledná pravidla pro tvorbu této skupiny entit, zvláště pak pro tvorbu entity dílo/výtvor a entity obecný pojem, se budou výrazně odlišovat od standardů a pravidel pro první skupinu entit. V průběhu projednávání těchto pravidel, na kterém jsme se také podíleli, došlo k výraznému zjednodušení původních představ o komplexnosti zpracování tohoto typu entit. To se projevilo i ve výrazném zjednodušení rozhraní pro jejich zpracování: ukazuje se, že není nutné ani vhodné (vzhledem ke zpracovatelům), aby rozhraní pro zpracování druhé skupiny entit se odlišovalo od rozhraní pro zpracování základní skupiny entit.

Proto bylo rozhraní pro druhou skupinu entit realizováno v podobě samostatných formulářů pro tyto entity v již stávajícím webovém rozhraní.

Rozhraní pro zpracování entit druhé skupiny tak představují formuláře pro zpracování entity ze třídy:

- událost,
- dílo/výtvor - realizováno jako 3 formuláře - zvláštní formuláře pro autorská a umělecká díla, stavby a ostatní díla,
- obecný pojem.

Formuláře navržené v minulém roce v podobě prototypu byly testovány a v tomto roce doplněny podle funkčních požadavků a podle Metodiky tvorby znalostního modelu.

Dospěli jsme k závěru, že realizace rozhraní pro zpracování druhé skupiny entit, které spočívá v doplnění formulářů pro další skupinu entit, nespĺňuje požadavky pro uplatnění výsledku typu software, proto nebude v tomto roce tento výsledek uplatňován.

Uživatelská příručka původně zpracovaná pro rozhraní pro základní skupinu entit je aktualizovaná a rozšířená o informace potřebné pro zpracování druhé skupiny entit; její aktuální podoba je zveřejněna na webovém sídle projektu www.interpi.cz.

V průběhu prací na projektu, zvláště pak nyní v průběhu testování využitelnosti Metodiky pro tvorbu znalostního modelu v měsíci září a v následné diskusi v odborných komunitách, se ukázalo, že bude vhodné oddělit rozhraní pro zpřístupnění entit od rozhraní pro jejich zpracování. Toto oddělené rozhraní pro zpřístupnění entit bude uživatelsky vstřícnější nejenom svou funkcionalitou, ale především i tím, že bude volně dostupné pro širokou veřejnost bez povinné registrace uživatele. Bude také více odpovídat požadavkům obsaženým v konceptuálních modelech, např. FRBR. Proto bude rozhraní pro zpřístupnění entit realizováno samostatně a uplatněno v roce 2015 jako druhý výsledek typu „software“ místo původně plánovaného výsledku pro software určeného ke zpracování druhé vedlejší skupiny entit a uplatněného v roce 2014. V tomto roce je k dispozici jeho prototyp, který bude představen na semináři k projektu.

C.3 Cíle projektu pro rok 2015

Pro následující čtvrtou a závěrečnou etapu projektu - tj. rok 2015 - byly v projektové dokumentaci stanoveny tyto cíle:

1. kontrola a ověření správnosti zvolených postupů,
2. návrh a realizace případných úprav,
3. závěrečné ověření v režimu poloprovozu.

Vzhledem k činnostem, které probíhaly a které jsou vysvětleny v části B, je potřeba upravit také cíle pro rok 2015 tak, aby aktivity projektu na sebe vhodně navazovaly. Jde o doplnění těchto aktivit:

1. vytvoření rozhraní pro zpřístupnění entit pro širokou veřejnost s ohledem na strukturu dat vymezenou Metodikou tvorby znalostního modelu;
2. propagace projektu - v oblasti propagace projektu a informování spolupracovníků projektu o dosažených výsledcích bude v roce 2015 realizován seminář, projekt bude prezentován na mezinárodní konferenci; bude průběžně doplňována samostatná stránka projektu.

D Použití účelové podpory v členění dle rozpočtu projektu

E Resumé

Třetí etapa projektu byla zaměřena na aplikovaný výzkum a vývoj. Hlavním výsledkem jsou dvě certifikované metodiky, které poskytují obecný rámec pro tvorbu znalostního modelu INTERPI a řeší rozsáhlou problematiku mapování a harmonizaci terminologických zdrojů.

První z uvedených metodik - Metodika tvorby znalostního modelu (vytvořená v minulém roce v podobě návrhu) byla v tomto roce dopracovaná do podoby, která představuje kompromisní přístup k řešení sporných otázek a současně je otevřená pro zavedení předpokládaných budoucích požadavků jednotlivých odborných komunit.

Druhá metodika - Metodika mapování a harmonizace číselníků, rejstříků a řízených slovníků aplikovaných v paměťových institucích přináší analýzu typů terminologických systémů využívaných v jednotlivých odborných komunitách, přehled metod využívaných v procese mapování a harmonizace termínů a návrh postupu při začleňování terminologických systémů do znalostní báze INTERPI. Metodika obsahuje také doporučení pro výběr termínů, které lze využít i v případě, že odborná komunita s budováním terminologického systému začíná. Metody porovnávání termínů shrnuté v Metodice byly experimentálně ověřovány na výběru termínů z různých systémů.

Dalším důležitým výsledkem projektu v této etapě je doplnění softwaru pro zpracování entit o formuláře pro druhou skupinu entit - tj. o třídu entit událost, třídu entit obecný pojem a třídu entit dílo/výtvar. Webové rozhraní pro zpracování entit tak představuje jednotný nástroj, pomocí kterého lze zpracovat všechny typy entit v rozsahu údajů daném Metodikou tvorby znalostního modelu.

V tomto roce pokračovaly také práce na úpravě, doplňování a analýze dat v prostředí souboru národních autorit i v prostředí archivů. Součástí byla také další zkušební transformace dat v souladu se závěrečnou podobou Metodiky tvorby znalostního modelu.

V oblasti propagace projektu jsme i v tomto roce připravili seminář, který se bude konat 19. - 20. 11. 2014 a také jsme projekt prezentovali na mezinárodním kongresu URLA v Istanbulu a připravujeme již tradiční prezentaci projektu na konferenci Archivy, knihovny, muzea v digitálním světě.

Uplatněné výsledky - druh výsledku podle RIV:

W (uspořádání workshopu): INTERPI. *Seminář projektu INTERPI. 19. 11. 2014 - 20. 11. 2014. Velké Bílovice.* Pozn.: výsledek bude uplatněn po uskutečnění semináře.

N_{met} (certifikovaná metodika): BALÍKOVÁ, Marie - KUNT, Miroslav - ŠUBOVÁ, Jana - ANDREJČÍKOVÁ, Nadežda. *Metodika tvorby znalostního modelu. 2014.* Pozn.: výsledek bude uplatněn po zpracování připomínek oponentů.

N_{met} (certifikovaná metodika) : BALÍKOVÁ, Marie - KUNT, Miroslav - ŠUBOVÁ, Jana - ANDREJČÍKOVÁ, Nadežda. *Metodika mapování a harmonizace rejstříků, číselníků a řízených slovníků aplikovaných v paměťových institucích. 2014.* Pozn.: výsledek bude uplatněn po zpracování připomínek oponentů.

F Přílohy

- F.1 - Seznam fyzicky předložených výsledků projektu
- F.2 - Použití účelové podpory v členění dle rozpočtu projektu - komentář
- F.3 - Dodatky, rozhodnutí - v roce 2014 nejsou
- F.4 - Cestovní zprávy
- F.5 - Metodika tvorby znalostního modelu (včetně oponentských posudků)
- F.6 - Metodika mapování a harmonizace rejstříků, číselníků a řízených slovníků aplikovaných v paměťových institucích (včetně oponentských posudků)
- F.7 - Zpráva o činnosti Národní knihovny v projektu INTERPI za rok 2014
- F.8 - Zpráva o činnosti Národního archivu v projektu INTERPI za rok 2014

F.1 Seznam fyzicky předložených výsledků projektu

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

F.3 Dodatky, rozhodnutí

V roce 2014 nejsou.

F.4 Cestovní zprávy

F.5 Metodika tvorby znalostního modelu (včetně oponentských posudků)

F.6 Metodika mapování a harmonizace rejstříků, číselníků a řízených slovníků aplikovaných v paměťových institucích (včetně oponentských posudků)

F.7 Zpráva o činnosti Národní knihovny v projektu INTERPI za rok 2014

F.8 Zpráva o činnosti Národního archivu v projektu INTERPI za rok 2014