

ZPRÁVA O VÝSLEDKÁCH ŘEŠENÍ PROJEKTU V ROCE 2011

INTERPI – Interoperabilita v paměťových institucích

Program aplikovaného výzkumu a vývoje národní a kulturní identity (NAKI) (DF11P01OVV023)

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



INTERPI

13. 11. 2011 | verze 1.7

Obsah

A	Konstatační část	5
A.1	Rešerše	5
A.2	Současný stav řešení ve světě a v Česku	5
A.2.1	Paměťové instituce	5
A.2.2	Sémantická interoperabilita - současný stav	6
A.2.2.1	Konceptuální modely.....	8
A.3	Vstupní data a cíl	8
B	Analytická část	9
B.1	Analýza, ověřování, doplňování a úprava dat.....	9
B.1.1.1	Ověřování a doplňování autoritních termínů - úprava AUT báze.....	9
B.1.1.2	Identifikace v oblasti autorit	9
B.1.1.3	Aplikace nástrojů pro vyjádření vztahů, informace v souvislostech.....	9
B.1.1.4	Vztah předchůdce - následník	9
B.1.1.5	Hierarchická struktura	10
B.1.1.6	Poznámky v autoritních záznamech	11
B.2	Korporace	11
B.2.1.1	Postup prací	11
B.3	Geografické entity	13
B.3.1.1	Rozšíření stávajících geografických autorit o potřebné kódované informace, např. kód_cz.....	13
B.3.1.2	Prostorová identifikace aplikovaná v projektu INTERPI	13
B.3.1.3	Záznamy geografických entit typu „sídlo“	14
B.3.1.4	Pracovní postup	14
B.3.1.5	Analýza zdrojů záznamů geografických entit.....	15
B.3.1.6	Počet zpracovaných záznamů	18
B.4	Entity typu války, bitvy	18
B.5	Analýza potřeb paměťových institucí	18
B.5.1	Regionální muzeum v Litomyšli	19
B.5.1.1	Personální jména	19
B.5.1.2	Jména rodin/rodů	19
B.5.1.3	Jména korporací	20
B.5.1.4	Geografická jména	20
B.5.1.5	Tematické termíny	21
B.5.1.6	Obecná úvaha na závěr	21

B.5.2	Geografické názvy - lokality v Demusu	21
B.5.2.1	Geografické názvy v přírodních vědách	22
B.5.2.2	Typologie geografických názvů v přírodních vědách	22
B.5.2.3	Fytochoriony (botanika)	22
B.5.2.4	Katastry, Lokality (všechny moduly)	22
B.5.2.5	Jednotky (geologie)	23
B.6	Potenciální aplikace klasifikačního systému MDT jako nástroje propojení různých systémů	23
B.6.1	Paměťové instituce - rozdíly v systémech zpřístupnění	24
B.6.2	Klasifikační systémy a konceptuální model FRASAD	25
B.7	Terminologická analýza	25
B.7.1	Autorita a související pojmy	26
B.7.2	Typy autorit	27
B.7.3	Objekt a související pojmy	27
B.7.3.1	Objekt, vztahy	27
B.7.4	Třídy, hierarchie a dědičnost	28
B.7.4.1	Entita a související pojmy	28
B.7.5	Terminologie - závěry	29
B.8	Specifikace požadavků na datovou strukturu	29
B.8.1	Entity v projektu INTERPI	29
B.8.1.1	Oblast vlastní identifikace, typ entity, stav záznamu	29
B.8.1.2	Oblast jména	30
B.8.1.3	Určení profesního zařazení a jazykové oblasti, do které daná entita spadá, ..	30
B.8.1.4	Detailnější informace uvádějící entitu do potřebných souvislostí, vztahů a vazeb	31
B.8.1.5	Životopis/historie	32
B.8.1.6	Zaměření, činnost	32
B.8.1.7	Informační prameny / dokumenty	32
B.8.1.8	Poznámka (veřejná, neveřejná)	32
B.9	Analýza metodologických postupů	32
B.9.1	GARR	33
B.9.1.1	Karta standardu	33
B.9.1.2	Cíle standardu	33
B.9.2	FRBR	34
B.9.2.1	Karta standardu	34
B.9.2.2	Cíle standardu	34

B.9.3	FRAD	34
B.9.3.1	Karta standardu.....	34
B.9.3.2	Cíle a hlavní myšlenky standardu.....	35
B.9.4	FRSAD	36
B.9.4.1	Karta standardu.....	36
B.9.4.2	Cíle a hlavní myšlenky standardu.....	37
B.9.5	RDA.....	38
B.9.5.1	Karta standardu.....	38
B.9.5.2	Cíle standardu.....	39
B.9.6	CCO	39
B.9.6.1	Karta standardu.....	39
B.9.6.2	Cíle standardu.....	40
B.9.7	ISAAR (CPF)	40
B.9.7.1	Karta standardu.....	40
B.9.7.2	Cíle standardu.....	41
B.9.8	ISAD(G)	41
B.9.8.1	Karta standardu.....	41
B.9.8.2	Cíle standardu.....	42
B.9.9	CIDOC-CRM.....	42
B.9.9.1	Karta standardu.....	42
B.9.9.2	Cíle standardu.....	43
B.10	Návrh datové struktury	44
B.10.1	Prvky datové struktury	44
B.11	Informační stránka projektu	46
B.11.1	Design.....	46
B.12	Systémová analýza současných technologií a metodologických postupů.....	47
B.12.1	Technické posouzení platform vhodných datových úložišť	48
B.12.2	Teoretická analýza různých typů datových úložišť	48
B.12.3	Stanovení základních požadavků na vhodné datové úložiště	48
B.12.4	Výběr konkrétních databázových platform pro testování.....	49
B.12.5	Výběr reprezentativních modelů použitelných v konkrétní databázové platformě a jejich realizace	49
B.12.6	1. model.....	49
B.12.7	2. model.....	49
C	Návrhová část	56

C.1	Definování všech typů entit (tříd)	56
C.1.1	Typologie entit	56
C.1.2	Specifikace vlastností a vztahů entit	57
C.2	Analýza, ověřování a úprava dat	57
C.2.1	Jednoznačná identifikace	57
C.2.2	Geografické entity	57
C.2.3	Entity typu války, bitvy	57
C.2.4	Personální entity	57
C.2.5	Entity rodina/rod	57
C.2.6	Korporativní entity.....	57
C.3	Potenciální aplikace klasifikačního systému MDT.....	57
C.4	Analýza aktuálních potřeb paměťových institucí	58
C.5	Analýza metodologických postupů	58
C.6	Aktivity v oblasti informačních technologií	58
C.6.1	Výběr databázové platformy, technického a datového modelu	58
C.6.2	Dokončení návrhu datové struktury a její kritické hodnocení	58
C.6.3	Zpracování funkčního konceptu softwarového řešení pro základní skupinu entit ..	59
C.6.4	Technická podpora komunikace a prezentace výsledků projektu	59
	Seznam bibliografických odkazů	60
	Seznam zkratk	65
D	Použití účelové podpory v členění dle rozpočtu projektu	
E	Resumé	
F	Přílohy	
F.1	Použití účelové podpory v členění dle rozpočtu projektu - komentář	
F.2	Dodatky, rozhodnutí	
F.3	Cestovní zprávy	
F.4	Zpráva z analýzy metodologických postupů, technologií sémantického webu, datových úložišť a návrh datové struktury pro základní skupinu entit	

A Konstatační část

A.1 Rešerše

BALÍKOVÁ, Marie. Focusing on User Needs: New Ways of Subject Access in Czechia. In *Subject Access : Preparing for the Future*. Berlin ; Munich : De Gruyter Saur, 2011, s. 7-24.

BALÍKOVÁ, Marie. Věcné autority v Česku : význam, aplikace, budoucí vývoj ... In *Zborník z 36. medzinárodného informatického sympózia o nových výzvach 21. storočia pre pamäťové inštitúcie : Nové rozmery informačného univerza*, 9. - 12. máj 2011, Stará Lesná - Vysoké Tatry. Bratislava : Spolok slovenských knihovníkov, 2011. s. 25-36. Dostupné na [www: <http://vili.uniba.sk/AK/INFOS2011.pdf>](http://vili.uniba.sk/AK/INFOS2011.pdf).

BALÍKOVÁ, Marie. UDC MRF Translation Tool and Multilingual Translation Database : how does it work for Czech translation. In *UDC Editorial Workshop, The Hague, 21 September 2011*. Dostupné na [www: <http://authority.nkp.cz/vecne-autority/mdt/udc-editorial>](http://authority.nkp.cz/vecne-autority/mdt/udc-editorial).

BALÍKOVÁ, M. - KUNT, M. - ANDREJČÍKOVÁ, N. - ŠUBOVÁ, J. Interoperabilita v paměťových institucích : vize a realita. In *12. konference Archivy, knihovny, muzea v digitálním světě 2011, Praha, 30.11.-1.12.2011*. Dostupné na [www: <http://www.skipcr.cz/akce-a-projekty/akce-skip/12.-konference-archivy-knihovny-muzea-v-digitalnim-svete-2011>](http://www.skipcr.cz/akce-a-projekty/akce-skip/12.-konference-archivy-knihovny-muzea-v-digitalnim-svete-2011).

ŠUBOVÁ, Jana. Spracovanie trochu inak. In *INFOS 2011*. 10.05.2011. Prezentace na konferenci.

ŠUBOVÁ, Jana. Cultural heritage information system as ecosystem = Informačný systém kultúrneho dedičstva ako ekosystém. In *Information ecology and library*. 11.10.2011. Prezentace na konferenci.

ŠUBOVÁ, Jana. Tam, kde autority končia, interdisciplinárna spolupráca začína. In *12. Slovenská bibliografická konferencia*. 7.11.2011. Prezentace na konferenci.

ANDREJČÍKOVÁ, Nadežda - ŠUBOVÁ, Jana. Regionálna bibliografia z elektronických zdrojov alebo Keď technológie pracujú s nami. In *12. Slovenská bibliografická konferencia*. 9.11.2011. Prezentace na konferenci.

BARTL, Zdeněk. České národní autority jubilující : soubory národních jmenných autorit jako výsledek kooperace českých knihoven In *Zborník z 36. medzinárodného informatického sympózia o nových výzvach 21. storočia pre pamäťové inštitúcie : Nové rozmery informačného univerza*, 9. - 12. máj 2011, Stará Lesná - Vysoké Tatry. Bratislava : Spolok slovenských knihovníkov, 2011. s. 37-40. Dostupné na [www: <http://vili.uniba.sk/AK/INFOS2011.pdf>](http://vili.uniba.sk/AK/INFOS2011.pdf).

A.2 Současný stav řešení ve světě a v Česku

A.2.1 Paměťové instituce

Prostředí paměťových institucí je specifické svou různorodostí a současně stejným předmětem zájmu, kterým je kulturní dědictví. Paměťové instituce jsou definovány jako: „knihovny,

archivy, muzea, výzkumné ústavy, univerzity, jejichž cílem je ochrana a zpřístupňování dokumentů kulturního dědictví“¹.

Teoretický pohled na rozdíly je možné nabídnout pomocí porovnání předmětu zájmu, použitých metod a cílů, které paměťové a také jiné informační instituce chtějí dosáhnout (Tab. 1).

Tab. 1: Druhy institucí a jejich primární úkoly²

instituce	primární medium (skupina)	primární medium (druh, vlastnosti)	primární metoda	primární cíl
muzeum	přírodní výtvar a artefakt	předmět konkrétní, jedinečné	výstava	porozumění a zážitek
archív	mentefakt (informační pramen)	záznam konkrétní, jedinečné	zpráva	poznatek
knihovna	mentefakt (informační pramen)	záznam abstraktní, namnoženo	vypůjčení	vědomost
databanka	mentefakt (informační pramen)	záznam virtuální, namnožitelné	zpráva	informace
výzkumný ústav	mentefakt (informační pramen)	záznam abstraktní, namnožitelné	publikace	vědomost
škola	mentefakt (informační pramen)	vědomost abstraktní, namnožitelné	učení	vědomost

A.2.2 Sémantická interoperabilita - současný stav

Současná úroveň sémantické interoperability a standardizace se v jednotlivých typech paměťových institucí značně liší. Základními pilíři sémantické interoperability v knihovnách jsou soubory autorit včetně řízených slovníků, tezaurů a metadatová schémata, která plně pokrývají potřeby sémantické interoperability z hlediska uživatelů knihoven. Samotná sémantika je skryta v metadatových schématech - výměnných formátech MARCového typu, kterou však nelze využít ve strojovém zpracování dat. Navíc rozsah údajů obsažených v autoritních záznamech plně neodpovídá požadavkům nejen regionálních knihoven z hlediska plnění jejich regionálních funkcí, ale ani dalších paměťových institucí, kdy zejména v archivech s ohledem na klíčové hledisko provenience je i na mezinárodní úrovni vyžadováno podrobnější zpracování autoritních záznamů (popisný standard ISAAR(CPF)³ a technický standard EAC⁴). Nedostačující rozsah údajů byl zapříčiněn především tím, že struktura stávajících standardů pro autoritní záznamy v knihovnách neumožňovala zařazení požadovaných prvků do datové struktury. Sémantická interoperabilita v archivech, muzeích a galeriích je ve stadiu zkoumání, prvotního ověřování standardů a zvolených postupů.

¹ MARVANOVÁ, Eva. Paměťová instituce.

² WAIDACHER, F. Příručka všeobecné muzeologie, s. 108.

³ International council on archives. ISAAR (CPF) : International standard archival authority record for corporate bodies, persons and families [online].

⁴ EAC-CPF : Encoded archival context, corporate bodies, persons, and families [online].

Na základě mezinárodních projektů, např. DESIRE⁵, DELOS (projekt Semantic Interoperability in Digital Library Systems)⁶ a VIAF⁷, můžeme usoudit, že interoperabilita je řešena v zásadě dvěma způsoby:

1. Mapováním jednotlivých systémů organizace poznání - souborů autorit, tezaurů, řízených slovníků⁸, které se aplikuje:
 - 1.1. na jeden centrální systém organizace poznání za pomoci jednoho metadatového schématu. Příkladem tohoto postupu je projekt HILT⁹ MSAC¹⁰;
 - 1.2. vzájemným mapováním všech autoritních systémů, např. MACS¹¹.
2. Budováním společného znalostního modelu organizace poznání napříč pamětovými institucemi.

Proto byla v roce 2006 přijata mezinárodní norma ISO21127¹², která vytváří základní rámec sémantické interoperability. Na jejím základě je v celosvětovém měřítku realizováno množství výzkumných projektů např. BRICKS¹³, EPOCH¹⁴ (konkrétně nástroj AMA - Archive Mapper for Archeology), SCULPTEUR¹⁵, OCHRE¹⁶, CHIOS¹⁷, SeMuSe¹⁸), které poukazují na význam objektově orientovaného přístupu při zpracování a zpřístupnění kulturního dědictví. Problematika však nebyla doposud na mezinárodní úrovni řešena komplexně.

Aplikace konceptuálního rámce přináší nové paradigma zpracování, které se posouvá od lineárního k objektovému. Zaměřuje se především na obsah současných záznamů o informačních objektech - tedy na zpracování entit (tříd) a jejich vztahů a tím na zabezpečení interoperability fondů, sbírek a souborů archiválií zpřístupňovaných v pamětových institucích na konceptuální úrovni bez ztráty jakýchkoliv dat s cílem omezit neefektivní duplicitní tvorbu dat jejich sdílením. Rešerše v prvním roce projektu prokázaly, že tento přístup, tedy tvorba společného znalostního modelu organizace poznání napříč pamětovými institucemi je efektivní a vyhovuje všem pamětovým institucím participujícím na projektu INTERPI

⁵ DESIRE Project : development of a european service for information on research and education [online].

⁶ PATEL, Manjula et al. Semantic interoperability in Digital Library Systems. [online].

⁷ VIAF : virtual international authority file [online].

⁸ eGovernment Resource Centre. Taxonomies & Thesauri [online].

⁹ HILT : high-level thesaurus [online].

¹⁰ BALÍKOVÁ, Marie. Multilingual Subject Access to Catalogues of National Libraries (MSAC) : Czech Republic's collaboration with Slovakia, Slovenia, Croatia, Macedonia, Lithuania and Latvia.

¹¹ ISAAC, Antoine - CHAMBERS, Sally. Prototype integrating MACS initial data and new alignment into TEL framework [online].

¹² ICOM. The CIDOC Conceptual Reference Model [online].

¹³ The BRICKS Community : the BRICKS Project [online].

¹⁴ EPOCH : European Network of Excellence in Open Cultural Heritage [online].

¹⁵ SCULPTEUR : Semantic and content-based multimedia exploitation for European benefit [online].

¹⁶ OCHRE : Online Cultural Heritage Research Environment [online].

¹⁷ CHIOS : Cultural Heritage Interchange Ontology Standardisation.

¹⁸ NEUMANN, Marco. SeMuSe the future of semantic museum data.

A.2.2.1 Konceptuální modely

Existují dva typy konceptuálních modelů:

- entitně-relační: FRBR, FRAD, FRSAD,
- objektově orientované: CIDOC-CRM, FRBR₀₀.

Uvedené modely je možné hodnotit různě. Z jednoho úhlu pohledu jsou důležitým krokem ke změně přístupu k zpracování, z jiného úhlu zůstávají teoretickými modely, které je možné do plně uvést do praxe jenom experimentálně v určité oblasti a v omezeném rozsahu. Model FRBR je úzce specializován na zpracování v knihovnách a ostatním paměťovým institucím nevyhovuje. CIDOC/CRM má ambice splnit požadavky všech typů paměťových institucí, ale je natolik obecný, že interpretace a zavedení do praxe je velmi náročné.

A.3 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. V rámci základního výzkumu, který spolu s aktivitami v oblasti vývoje, je charakteristický pro první etapu projektu (2011-2012) jsme se v roce 2011 zaměřili na:

- ověřování a doplňování autoritních termínů v rámci stávající databáze národních autorit jako jednoho ze základních předpokladů tvorby znalostní báze paměťových institucí: stávající databáze autorit představuje základ identifikace jednotlivých entit, která je však srozumitelná jen pro lidského uživatele,
- geografické entity,
 - rozšíření stávajících geografických autorit o potřebné kódované informace, např. kód_cz,
 - prostorovou identifikaci aplikovanou v projektu INTERPI,
 - problematiku pořizování dat o vybraných geografických entitách, zvláště entitě typu sídlo,
- analýzu potřeb paměťových institucí,
- potenciální aplikaci klasifikačního systému MDT jako nástroje propojení,
- terminologickou analýzu termínů aplikovaných v projektu INTERPI,
- výběr, studium a analýzu standardů relevantních pro řešení projektu,
- specifikaci požadavků paměťových institucí v oblasti základních entit, tj. personálií, korporací a geografických entit,
- tvorbu návrhu datové struktury pro základní skupinu entit,
- testování databázových platforem pro cílové řešení.

B Analytická část

B.1 Analýza, ověřování, doplňování a úprava dat

B.1.1.1 Ověřování a doplňování autoritních termínů - úprava AUT báze

Ověřování a doplňování autoritních termínů je v této fázi projektu jedním ze základních předpokladů tvorby znalostní báze paměťových institucí INTERPI.

B.1.1.2 Identifikace v oblasti autorit

Stávající databáze autorit představuje základ identifikace jednotlivých entit, která je srozumitelná pro lidského uživatele. Nyní je nutné oblast jednoznačné identifikace rozšířit tak, aby byla srozumitelná, interpretovatelná a zpracovatelná i softwarovými agenty. Z tohoto důvodu bude ID autority připojeno k preferované formě i variantním, tj. nepreferovaným formám termínů v záznamech stávajících autorit a v provázaných bibliografických záznamech. Přípravné práce pro technickou realizaci probíhaly v průběhu tohoto roku - akce bude realizována v prosinci (důvod: akce vyžaduje uzavření BIB i AUT báze NKC pro veřejnost, proto budou k tomuto účelu využity prosincové sváteční dny).

B.1.1.3 Aplikace nástrojů pro vyjádření vztahů, informace v souvislostech

Ve stávající autoritní bázi je nově umožněno vyjádření vztahů předchůdce - následník, je explicitně vyjádřena hierarchická struktura a je možno připojit další potřebné informace formou vhodných poznámek.

B.1.1.4 Vztah předchůdce - následník

V rámci sjednocování postupů a prostředí v rámci databáze národních autorit je nyní umožněno i v případě věcných autorit, tj. geografických, tematických a formálních autoritních termínů v poli 55X (tedy 551, 550, 555) rozlišit, jde-li o dřívější či novější záhlaví. Současná úprava se týká převážně geografických termínů.

Jedná se o možnost připojení podpole "w" s hodnotou "a" pro dřívější záhlaví a s hodnotou "b" pro novější záhlaví v uvedených polích 55X (příklad na Obr. 1).

Obr. 1: Příklad zápisu vztahu „předchůdce - následník“

AUT - Úplné zobrazení záznamu

Zvolte formát: [Standardní](#) -- [MARC](#)

Záznam 1 z 1

Ident. číslo	ge128585
Záhlaví	● Sri Lanka
Kód geogr. oblasti	a-ce---
Odkaz. forma	Demokratická socialistická republika Sri Lanka
Předch. záhlaví	Cejlon
MDT	●(292.563) (fyziogeografické určení) ●(548.7) (místní určení)
Angl. ekvivalent	Sri Lanka
GPS	7°52'23"N, 80°46'18.5"E Zobrazit na mapě
Poznámka	Stát na stejnojmenném ostrově, název Sri Lanka používán od roku 1972.
Status	správní celek
Zdroj	LC (Names)
	www(Getty)
Další informace	Wikipedie (Sri Lanka)

AUT - Úplné zobrazení záznamu

Zvolte formát: [Standardní](#) -- [MARC](#)

Záznam 1 z 1

Ident. číslo	ge229696
Záhlaví	● Ceylon
Odkaz. forma	Ceylon
Násled. záhlaví	Sri Lanka
Poznámka	Používá se pro tvorbu korporativních záhlaví (do roku 1972). Při věcném zpracování se používá poslední varianta názvu Sri Lanka.
Systém. číslo	000229696

B.1.1.5 Hierarchická struktura

Bylo také připojeno speciální pole pro vyjádření hierarchické struktury u GEO názvů sídel.

Struktura pole 951:

- \$a - stát
- \$b - oblast
- \$c - kraj
- \$d - okres
- \$e - obec/město
- \$f - část obce/města
- \$g - lokalita v rámci obce/města

Obr. 2: Příklad zápisu obce Loučky s kompletní hierarchickou strukturou

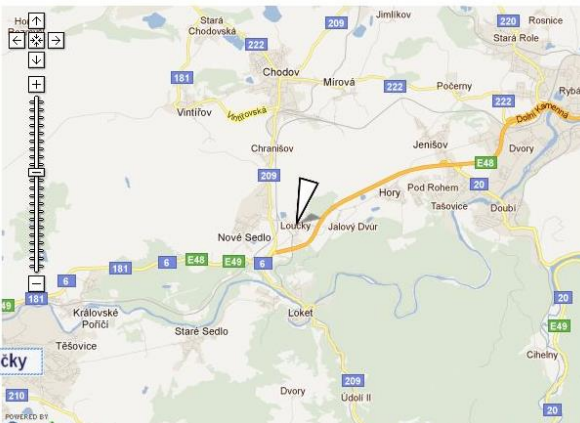
Zvolte formát: [Standardní](#) -- [MARC](#)

Záznam 1 z 1

Ident. číslo	ge369152
Záhlaví	● Loučky (Nové Sedlo, Sokolov, Česko)
Kód geogr. oblasti	e-xr--- e-xr-ka
Odkaz. forma	Grúnlas (Nové Sedlo, Sokolov, Česko)
Nadřazený termin	Nové Sedlo (Sokolov, Česko)
MDT	●(437.318)
Angl. ekvivalent	Loučky (Nové Sedlo, Sokolov, Czech Republic)
GPS	50°12'24.7"N, 12°45'9.8"E Zobrazit na mapě
Poznámka	Část obce Nové Sedlo v okrese Sokolov .
Status	správní celek
Zdroj	www(Adresy ČR-MVČR)
	Ortslexikon Sudetenland

951 |a Česko |c Karlovarský kraj |d Sokolov : okres |e Nové Sedlo |f Loučky

Systém. číslo 000369152

[http://aleph.nkp.cz/F/?func=direct&doc_number=000369152&local_base=AUT]

Akce proběhla v měsíci září, v této fázi bylo připojeno pole v záznamech 8645 názvů sídel. Záznamy jsou průběžně kontrolovány a informace v nich obsažené jsou upravovány tak, aby byly srozumitelné pro softwarové agenty¹⁹.

B.1.1.6 Poznámky v autoritních záznamech

Poznámky v autoritních záznamech obsahují definice a komentáře k rozsahu autoritního termínu uvedeného v záhlaví. Zapisují se v poli 680.

V současné době je možno rozlišit dva typy poznámek:

- definice, které se zapisují v poli 680 s hodnotou druhého indikátoru „0“,
- pokyny k použití, které se uvádějí v poli 680, v němž pozice druhého indikátoru není obsazena.

Obr. 3: Příklad zápisu poznámek v záznamu autority

FMT	GA
001	ge487180
003	CZ-PrNK
005	20101005154305.0
008	081120 n acznnaabn-----n-aja-----
040	a ABA001 b cze
151	a Rakousko (1526-1804)
451	a Habsburská monarchie (1526-1804) 0 o
451	a Rakouská monarchie (1526-1804) 0 o
551	a Rakousko
551	a Rakousko (1804-1867)
551	a Rakousko (1867-1918)
551	a Rakousko-Uhersko
670	a Evropa v proměnách staletí
6800	i Používá se pouze pro korporace zapsané pod správním celkem s působností pro území Rakouska v letech 1526-1804.
68000	i Označuje historický státní útvar, jemuž vládla rakouská větev habsburské dynastie a jejich následnická habsbursko-lotrinská dynastie v letech 1526-1804; známé také pod neoficiálními názvy Habsburská monarchie, popř. Rakouská monarchie.
75107	a Austrian monarchy (1526-1804) 2 eczenas
906	a wp20081030 b hedo
9501	a správní celek
SYS	000487180

B.2 Korporace

Jedním z hlavních úkolů projektu INTERPI je také cílené a komplexní pokrytí problematiky korporací a vazeb mezi nimi. Řešení tohoto problému je velmi důležité zvláště z pohledu archivů, které tyto entity nutně potřebují při popisu archiválií a zároveň disponují adekvátními odbornými předpoklady.

B.2.1.1 Postup prací

Pro analýzu korporací byl jako prvotní zdroj dat zvolen výstup (MS Excell) z celostátní evidence archivních souborů. Zároveň byl navržen následující postup práce:

¹⁹ V nejednoznačných/sporných případech je nutné posoudit a připojit údaje hierarchické struktury individuálně, případně podle jednotlivých typů.

- dle výstupu z PEvA zhodnotit u původců zápis jejich názvů (chybějící původní cizojazyčné názvy, transkripce, úpravy oproti původním názvům),
- vybrat typické a netypické představitele jednotlivých skupin původců, přičemž se zaměřit zejména na problematiku původců, které nejsou „právníckou osobou“, resp. jsou součástí jiného původce (= na ty nejvíce komplikované) a nejsou třeba ani uvedeni v PEvA,
- u vybraných původců zjistit, zda zápis původce v PEvA přibližně odpovídá skutečnosti a doplnit cizojazyčné názvy, shromáždit podklady o původci (úvod k inventáři - jeho naskenování, přepis apod.),
- sejít se nad výsledky a probrat problematiku, zjistit způsob zápisu původců dle AACR2R a českých specifik.

Protože je tato fáze projektu zaměřena především na základní výzkum a testování, a není zatím k dispozici konkrétní softwarové řešení (to je předmětem 2. fáze projektu), byly pro zpracování informací o korporacích využity jednoduché tabulky. Struktura dat o korporacích byla zároveň navržena striktně podle standardu ISAAR(CPF) a to z toho důvodu, že i při aplikaci jiných datových modelů a návrhu datové struktury je potřeba respektovat potřeby archivů formulované v ISAAR(CPF). Model CIDOC-CRM, který má být základem řešení datové struktury pro projekt, představuje pouze obecný rámec, ve kterém je nutné aplikovat požadavky všech zúčastněných paměťových institucí a odborných komunit, včetně archivářů.

Ukázka popisu jedné korporace, červeně jsou vyznačeny povinné údaje (čísla vlevo jsou kapitoly ISAAR(CPF)):

5.1.1	Korporace
5.1.2	NÁMI autorizovaná forma jména tj. oficiální resp. nejčastěji užívaný název v „originální“ jazykové mutaci: jména v jazyce, v kterém je korporace obecně známa; v případě pochybností (utrakvistické instituce v 19. století, mezinárodní organizace) má přednost čeština, pokud v ní název prokazatelně dobově existuje. Zpravidla se uvede jméno nejčastěji užívané v pramenech (nemusí nutně odpovídat oficiálnímu). Součástí názvu je případné vyjádření korporativnosti (n.p., s.p., a.s., s.r.o.)
5.1.3	Název v jiných jazykových mutacích, plný název např. s „k.k.“ / „k. und k.“
5.1.5	Uvádí se všechna další zjištěná nebo pravděpodobná jména, popř. též v invertované podobě, včetně zkratk, akronymů, jazykových a gramatických variant názvu a změn názvu, pokud se nejedná o jinou korporaci.
5.2.1	Faktické fungování instituce, přesnost na roky
5.2.2	Pokud je k dispozici elektronický text inventáře, tak překlomit z něj (viz níže ostatní náměty 4a+4b)
5.2.3.1	Geografické sídlo: Uvádí se geografické určení hlavního nebo nejčastěji užívaného sídla. Provádí se vazbou na příslušnou geografickou entitu.
5.2.3.2	Geografická působnost: Uvádí se geografické vymezení působnosti (jurisdikce) korporace.
5.2.4	Typ korporace
5.2.5	Stručná hlavní náplň činnosti
5.2.6.1	Normy, které vymezují působnost, kompetence etc.
5.2.6.2	Normy vzniklé z činnosti instituce v oboru její působnosti
5.2.7	Stručně
5.4.8	Jen skutečně použitá literatura v pořadí dle míry použití, citace přiměřeně dle příkladu z ISAAR(CPF) bez použití kapitálek v příjmení.
5.4.9	Vytvořil: XY, CZ-A100000010 (Národní archiv)

5.3.1	5.3.3	5.3.4
Název z 5.1.2, u předchůdců a nástupců, kteří mimo pracovní vymezeném období nutná konzultace s tím, který je má přidělené	Předchůdce Nástupce Nadřízený Podřízený Součást	Rok zlomu Rok zlomu Od-do Od-do Od-do

Spolu s tímto návodem bylo stanoveno, že budou sledovány zejména problémy a komplikace, nejednoznačnosti apod. Na základě těchto zjištění proběhne následná diskuse a upřesnění metod, tvorby názvů a datové struktury INTERPI.

B.3 Geografické entity

B.3.1.1 Rozšíření stávajících geografických autorit o potřebné kódované informace, např. kód_cz

V roce 2011 byla vypracována studie Význam Databáze sídelních lokalit Čech, Moravy a Slezska CZ_RETRO pro projekt INTERPI: aplikace kódu CZ v souboru Národních geografických autorit²⁰; která doporučuje využití této databáze v projektu INTERPI:

„Pro potřeby projektu INTERPI je více než vhodné využívat databázi CZ_RETRO nejen v její tabelární podobě, ale zejména v podobě geografického informačního systému, tedy nahlížením, vizualizací a vyhledáváním v mapových projektech a službách. Toto prostředí poskytuje jistotu správného ztotožnění a jednoduché vizuální kontroly hledané lokality. Navíc lze vizualizovat též atributy vedené v tabulkách propojitelných pomocí kódu CZ s body a polygony lokalit v mapě.

Pro projekt INTERPI, tedy pamětové instituce, je zásadně důležité prostorová identifikace všech typů entit specifických pro předmět jejich zájmu a to v jejich časovém, resp. historickém kontextu a vývoji. Přitom celá řada těchto entit (například zaniklé nebo administrativně zrušené vesnice, všechny drobné lokality) není ve veřejných databázích vůbec evidována. Nejlepší a v podstatě jediný základ pro prostorovou identifikaci prvků a atributů aktuální i historické sídelní struktury, sledovaných pamětovými institucemi, poskytuje v současné době právě CZ_RETRO, existující v rámci Integrovaného informačního systému památkové péče. Standardy pro prostorovou identifikaci CZ_RETRO je možno pro projekt INTERPI s výhodou využít a dále podle potřeby rozšířit a prohloubit“²¹.

B.3.1.2 Prostorová identifikace aplikovaná v projektu INTERPI

V tomto roce řešení projektu byla také vypracována studie o potenciální aplikaci prostorové identifikace²², která doporučuje:

²⁰ KUČA, Karel. Význam Databáze sídelních lokalit Čech, Moravy a Slezska CZ_RETRO pro projekt INTERPI : aplikace kódu CZ v souboru Národních geografických autorit.

²¹ Cit. 20.

²² SYROVÝ, Jiří. Prostorová identifikace v projektu INTERPI.

...“pro projekt INTERPI, tedy pamětové instituce, je zásadně důležitá prostorová identifikace všech typů entit specifických pro předmět jejich zájmu včetně časového, resp. historického kontextu a vývoje. Přitom celá řada těchto entit (například objekty kostelů, zřícenin hradů) není ve veřejných databázích vůbec evidována, resp. nemusí být evidována s potřebnými atributy (v katastru nemovitostí nebo v RSO nalezneme definiční bod budovy pouze s informací o jejím čísle popisném či orientačním, nikoli však o tom, že se jedná o zámek).

Solidní základ pro prostorovou identifikaci prvků a atributů sledovaných pamětovými institucemi poskytuje v současné době Integrovaný informační systém památkové péče. V něm definované standardy pro prostorovou identifikaci je možno pro projekt INTERPI s výhodou využít a dále podle potřeby rozšířit a prohloubit. IISPP je důležitý i existující technologickou základnou umožňující efektivní správu a publikování prostorových identifikátorů, jejichž součástí je i generování prostorových relací s potřebnými datovými zdroji státní správy a dalších subjektů, zajišťující vazby na aktuální údaje těchto zdrojů“²³.

B.3.1.3 Záznamy geografických entit typu „sídlo“

V roce 2011 byly vzhledem k legitimnímu požadavku konsorciálního partnera Národního archivu (dále NA ČR) preferovány geografické entity, a to konkrétně entity typu sídlo. Dohodli jsme se, že se v první fázi omezíme na sídla v Moravskoslezském kraji, v Česku, na Slovensku a na Zakarpatské Ukrajině. Výběr těchto oblastí nebyl zvolen náhodně, např. Moravskoslezský kraj byl zvolen z důvodů: „Mnohé názvy v Moravskoslezském kraji se v průběhu času měnily, lze se domnívat, že ovšem častěji než v mnohých jiných krajích republiky. Kraj jako příhraniční region měl pohnutou historii. Hranice (hlavně jeho slezské části) se posouvaly, na čas některé oblasti patřily jiným státním útvarům, což mělo zejména na místní jména jistě vliv. Dalším velmi významným faktorem je národnostní složení kraje, resp. jazyk, k němuž se obyvatelé hlásí. Jako všude v našem pohraničí jde pochopitelně až do r. 1945 především o němčinu, německé názvy. V tomto kraji ale má zřetelný vliv ještě polština, polské názvy jsou aktuální dodnes“²⁴.

B.3.1.4 Pracovní postup

Dále bylo rozhodnuto, že se bude vycházet ze stávajícího souboru národních geografických autorit, jehož záznamy autoritních termínů z uvedených oblastí budou ověřeny a rozšířeny o informace potřebné pro pamětové instituce. NK ČR připravila export údajů z uvedených oblastí formou tabulky MS Excell, tyto údaje jsou ověřovány v odpovídajících zdrojích (pro NA ČR byly stanoveny zdroje Profous a Hosák²⁵); záznamy, které nejsou obsaženy v souboru národních geografických autorit, jsou zaznamenávány ve speciálním souboru a jsou doplňovány o potřebné údaje včetně záhlaví podle knihovnických pravidel, tedy je v záhlaví uváděn i kvalifikátor a v záznamu geografický kód kraje (Tab. 3 a Tab. 4 na konci části B). V souborech takto

²³ Cit. 22.

²⁴ HRAZDIL, Aleš. Geografické termíny Moravskoslezského kraje - specifikace požadavků uživatelské komunity jednotlivých pamětových institucí daného regionu.

²⁵ Informaci o obou zdrojích viz níže v kapitole B.3.1.5.

ověřených geografických názvů jsou uváděny německé, variantní české názvy z 19. století a starší, v případě Moravskoslezského kraje i polské variantní názvy - max. 7 názvů, včetně jejich datace, která se uvádí k názvu za vyhrazený znak „%“. Pro potřeby archivů se uvádějí i varianty, které nejsou v 1. pádě, jsou označovány symbolem „***“; tyto variantní formy nebudou importovány do databáze Národních autorit, budou integrovány až do databáze INTERPI. U nepreferovaných názvů se nepoužívá geografický kvalifikátor vyššího územního celku, tj. označení státu, zejm. „Česko“. Případné autoritní záznamy vzniklé v mezidobí bude možné vyloučit, pokud jméno obce bude upraveno podle pravidel pro tvorbu autoritního záhlaví (v případě totožných jmen). Současný okres se uvádí v samostatném poli podle Statistického lexikonu z r. 2008.

Pro připojení takto vytvořených nových autoritních termínů do autoritní databáze vypracovala NK ČR v tomto roce řešení projektu speciální postup a program.

Tab. 2: Ukázka tabulky pro doplnění informací geografických autorit

ID	Záhlaví	Poznámka	Česko	kraj	okres 05	obec 05	část obce 05	Neprefer.1	Neprefer.2	zdroj
ge11 4756	Holice (Olomouc, Česko)	Část města Olomouc.	e-xr---	e-xr-o	Olomouc	Olomouc		Olomouc-Holice (Česko)	Holitz%1846	Hosák
ge13 0299	Olomouc (Česko)	Krajské a okresní město.	e-xr---	e-xr-o	Olomouc			Olmütz (Česko)	Olmütz%1872	Hosák
	Lidlovy Dvory (Kašperské Hory, Česko)	Osada 2 km jižně od Kašperských Hor.		Plz	Klatovy	Kašperské Hory		Liedlhöfen%1840	Dvory Lidlovy%1854	Profous

Pro potřeby INTERPI, zvláště pro potřeby archivů - po vypracování seznamu Národním archivem - budou v roce 2012 integrovány do souboru národních autorit i záznamy historických územních celků, např. Tábor (okres politický : 18XX-1949). Tábor (okres soudní : 18xx-1949).

B.3.1.5 Analýza zdrojů záznamů geografických entit

V roce 2011 byla provedena analýza jednotlivých zdrojů geografických názvů sídel.

Kol.: Historický lexikon obcí České republiky 1869-2005. Český statistický úřad, 2003.²⁶

Pramen zachycuje změnu názvu v letech 1869-2005, neuvádí důsledně německé názvy (spíše jen nahodile - např. není Wittingau). Zachycuje příslušnost k (politickým) okresům a obcím, vychází z Retrospektivního lexikonu a je uspořádán abecedně.

²⁶ konzultací s ředitelem odboru ČSÚ zjištěno, že je možné získat databázi

Kol.: Retrospektivní lexikon obcí Československé socialistické republiky 1850-1970. Federální statistický úřad, 1978.²⁷

Zdroj zachycuje příslušnost k (politickým) okresům a obcím, neuvádí důsledně jinojazyčné názvy a je uspořádán abecedně.

Sturm, Heribert [Hrsg.]: Ortslexikon der böhmischen Ländern 1910-1965. R. Oldenbourg Verlag München, 1995

V tomto zdroji jsou obce členěny podle období: a) před 1. válkou do 1910; b) první republika 1930; c) Protektorát/Sudety 1938, resp. 1939; d) do 1949; e) do 1955; po 1961; členění je uspořádáno podle politických okresů v r. 1910, v jejich záhlaví uvedeny změny příslušnosti obcí, je opatřen rejstříky obsahující německé názvy, české názvy, polské názvy.

Profous, Antonín: Místní jména v Čechách. Jejich vznik, původní význam a změny. Nakladatelství Československé akademie věd, 1954-1960, díl I-IV + dodatky (V).

Zdroj využívá historických výpisků z pramenů (Palacký, Sedláček), jména jsou uváděna včetně citací podle dobových pramenů s dobou užití, zaznamenává se i lidová výslovnost. Zdroj obsahuje odkazy viz, přídavné jméno nefunguje jako pořadatel místního jména (např. Lídlový Dvory uváděny pod Dvory, ale v přímém pořadí), je uváděno vysvětlení vzniku jména (filologicky), zdroj obsahuje přesné citace s výskytem jména a uvedením pramene (zkratkou - komplikované pro pochopení); v řadě případů se jedná o skloněnou variantu (Holonohe - „1143/8 Golonozech villa donatur monio Strahoviensi“).

Roubík, František: Soupis a mapa zaniklých osad v Čechách. Nakladatelství Československé akademie věd, 1959

Geografické názvy jsou členěny podle okresů a uváděny s přibližným určením místa a vročením.

Hosák, Ladislav - Šrámek, Rudolf: Místní jména na Moravě a ve Slezsku I, II. Academia Praha, 1970, 1980.

Ve zdroji je podrobně popsána struktura hesla: 1a) charakteristika sídliště; 1b) bližší lokalizace; 1c) územní příslušnost ke Slezsku; 2) historické doklady místního jména; 3a) příslušnost k panství 1848; 3b) nejdůležitější údaje místních dějin; 3c) národnost obyvatelstva 1930; 3d) odkud přišlo obyvatelstvo po 1945 (dosídlení pohraničí); 4a) lidová podoba jména; 4b) údaje o místním nářečí; 5) výklad místního jména; 6a) samoty v katastru obce; 6b) výběr historických názvů částí obcí uliv aj., zdroj obsahuje odkazy viz.

Turek, Adolf - kol.: Místopisný rejstřík obcí českého Slezska a severní Moravy. Zemský archiv v Opavě, 2004²⁸

Zdroj důsledně uvádí odkazy viz.

²⁷ 4 svazky, I/1 historický lexikon ČSR, II/1 SSR.

²⁸ upravené vydání starší publikace z roku 1974 (je k dispozici v TXT elektronicky)

Šulc, Ivo - kol.: Databáze prvních zmínek [o sídelních jednotkách]

Zdroj obsahuje strukturovaná data pro okres Chrudim, je zaměřen na všechny pomístní názvy, neobsahuje odkazy viz (nejedná se o publikaci).

Kol.: Vlastivedný slovník obcí na Slovensku I-III. VEDA, vydavateľstvo Slovenskej akadémie vied, Bratislava 1977-1978

Ve zdroji je uvedena přesná struktura záznamů:

1. Obecné údaje:
 - 1.1. Staré a jinojazyčné názvy obcí a obyvatelské jména;
 - 1.2. Administrativní začlenění v minulosti a současnosti;
 - 1.3. Části obcí podle Statistického lexikonu 1965;
 - 1.4. Místní části obcí;
 - 1.5. Počty obyvatelů;
 - 1.6. Výměra území patřící k obci;
2. Přírodní poměry:
 - 2.1. Nadmořská výška;
 - 2.2. Poloha;
 - 2.3. Povrch;
 - 2.4. Podnebí;
 - 2.5. Vodstvo;
 - 2.6. Půda;
 - 2.7. Flóra;
 - 2.8. Fauna;
 - 2.9. Rezervace;
3. Dějiny:
 - 3.1. Předhistorie;
 - 3.2. Feudalismus a kapitalismus;
 - 3.3. Dějiny od r. 1918;
4. Kultura:
 - 4.1. Školství;
 - 4.2. Divadlo;
 - 4.3. Hudební život;
 - 4.4. Literární život;
 - 4.5. Tiskárny;
 - 4.6. Časopisy;
 - 4.7. Výzkum a věda;
 - 4.8. Výtvarní život;
 - 4.9. Památky;
 - 4.10. Osvěta;
 - 4.11. Národopis;
 - 4.12. Osobnosti;
5. Zdravotnictví;
6. Sport a turistika;
7. Literatura.

Kol.: Istorija gorodov i sel Ukrainskoj SSR. Zakarpatskaja oblast'. Institut istorii Akademii nauk USSR, Kiev 1982

Jedná se o encyklopedii převážně historického vývoje obcí, neobsahuje vývoj jmen obcí, jména obcí jsou uvedena pouze v ruštině.

B.3.1.6 Počet zpracovaných záznamů

Celkem byl zkontrolován či zpracován následující počet záznamů. Údaj není přesný, v době psaní této zprávy práce stále probíhají.

Země	záznamů doplněných (přítomných v národních autoritách)	záznamů nových
Podkarpatská Rus ²⁹	-	1441
Slovensko	nepoužito	727
Čechy, Morava a Slezsko	822	463
Moravskoslezský kraj	505	
Polsko	701	

Celkem bylo tedy zpracováno 4659 záznamů různé složitosti.

B.4 Entity typu války, bitvy

V pojetí národních věcných autorit jde o „specifické entity“, které jsou „provizorně“ zařazeny do souboru tematických autorit, formáty rodiny MARC jiné řešení neumožňují. Původně měla být tato problematika řešena až v další etapě projektu, ale vzhledem k aktuálním potřebám specializované knihovny Vojenského historického ústavu jsme přikročili k jejímu řešení již v této fázi projektu. V současné době je posuzováno a bude do AUT báze do připojeno XYZ návrhů těchto entit; záznamy těchto entit budou opatřeny poznámkami, definicemi a budou v rámci stávajících omezených možností AUT báze vzájemně propojeny, plnohodnotně budou propojeny až v databázi INTERPI.

Ukázka tabulky s informacemi o entitách typu války, bitvy se nachází na konci části B (Tab. 5).

B.5 Analýza potřeb paměťových institucí

V prvním roce projektu byla také vypracována analýza potřeb paměťových institucí a definovány požadavky na identifikaci entit a množství potřebných informací jasně deklarujících jejich vlastnosti a vztahy/vazby, do kterých vstupují.

²⁹ Specifickým příkladem jsou sídla na bývalé Podkarpatské Rusi (Zakarpatská oblast Ukrajiny), kde bylo místo historické literatury zatím využito úředního statistického lexikonu z roku 1937 a doplnkově soupisu politické správy z roku 1922. Současně příslušná pracovnice zvládla i psaní názvů cyrilicí.

B.5.1 Regionální muzeum v Litomyšli

Internetová databáze sbírek Regionálního muzea v Litomyšli (Sbírky RML online, www.sbirky.rml.cz) obsahuje přes 30.000 sbírkových předmětů v 11 oborových modulech; každý z nich je opatřen jedním až pěti slovníky, tj. soubory selekčních prvků.

Analýza je rozdělena na pět částí: personální jména, jména rodin/rodů, jména korporací, geografická jména, tematické termíny.

B.5.1.1 Personální jména

Výsledek: současné národní autority jsou pro účely sbírek RML online **vyhovující**.

Komentář k výsledku:

- struktura záznamů personálních autorit je mnohem bohatší, uvádí se zde např. více (všechny) obory působení, odkazové formy atd.,
- propojení/nahrazení personálních jmen jejich záznamy z národních autorit by bylo pro sbírky RML online přínosné (zejména u jmen, u nichž sbírky online neuvádí kromě jména žádné další údaje).

Poznámka k obecnému užití v paměťových institucích:

- jako ryze selekční prvky jsou záznamy personálních jmen dostatečné,
- obecně vhodné by bylo konkrétnější (normované) odkazování na zdroje informací,
- paměťové instituce budou směřovat k obohacení strohých selekčních prvků o další údaje a jejich vzájemné propojení (znalostní databáze) - v tomto smyslu jsou tvořeny tzv. Muzejní autority,
- projekt INTERPI by měl zřejmě převzít/zakomponovat Muzejní autority do své struktury.

B.5.1.2 Jména rodin/rodů

Výsledek: současné NA jsou pro účely sbírek online **vyhovující**.

Komentář k výsledku:

- struktura záznamu jmen rodin/rodů je bohatší, velmi cenné jsou především odkazové formy,
- propojení/nahrazení jmen rodin/rodů jejich záznamy z národních autorit by bylo ve sbírkách RML online možné; přínos tohoto kroku by byl pouze v odkazových formách.

Poznámka k obecnému užití v paměťových institucích:

- jako ryze selekční prvky jsou národní jmenné autority dostatečné,
- obecně vhodné by bylo doplnění časového určení existence rodu do charakteristiky a konkrétnější (normované) odkazování na zdroje informací,
- zejména archiváři budou zřejmě usilovat o doplnění dalších informací - kromě časového určení existence rodu minimálně např. hierarchizace rozrodů jednotlivých rodů, přiřazení/propojení jmenných autorit členů rodu na jméno rodu, příp. obsáhlejší „životopis“ rodu.

B.5.1.3 Jména korporací

Výsledek: současné NA jsou pro účely sbírek online částečně vyhovující.

Komentář k výsledku:

- nedostatkem řady (většiny?) korporací v národních autoritách je chybějící časové určení, někdy (často?) chybí i geografické určení (mělo by být ve stručné charakteristice v poli Poznámka),
- není také zřejmé, jestli je v národních autoritách možné užití skupinové autority typu „vysočinské sklárny“,
- kladem národních autorit jsou odkazové formy a odkazy na předchůdce či následníka korporace,
- propojení/nahrazení jmen korporací jejich záznamy v národních autoritách - po doplnění časového a geografického určení - by bylo pro sbírky online přínosné.

Poznámka k obecnému užití v paměťových institucích:

- jako ryze selekční prvky jsou národní autority jen částečně dostatečné - zcela nedostatečné jsou v případech, kdy různé korporace, působící v různých obdobích, měly stejné pojmenování - typické jsou např. okresní úřady - problém je v chybějícím časovém určení (archiváři to budou požadovat zcela nepochybně),
- obecně vhodné by bylo doplnění konkrétnější (normované) odkazování na zdroje informací,
- možno je uvažovat o obsáhlejším „životopisu“ korporace.

B.5.1.4 Geografická jména

Výsledek: současné NA jsou pro účely sbírek online částečně vyhovující.

Komentář k výsledku:

- existující geografická jména v národních autoritách jsou vyhovující, řada druhů geografických jmen ovšem v národních autoritách není, např. sochy v exteriéru, pomníky, drobné církevní památky atd.; toto se týká i jednotlivých domů, resp. čísel popisných,
- kladem národních autorit jsou odkazy viz a viz též a rovněž geolink,
- propojení/nahrazení geografických jmen jejich záznamy v národních autoritách - po doplnění chybějících druhů geografických jmen - by bylo pro sbírky RML online přínosné.

Poznámka k obecnému užití v paměťových institucích:

- různé vědní obory mohou chápat různě geografická jména se stejným názvem - je proto nutné důsledně používat rozlišovací kvalifikátor (zároveň tento kvalifikátor zřejmě značně zpřesnit/rozšířit),
- historické názvy lokalit (a obecně všechny názvy) bude nutné pevně navázat na časové určení a na zdroj informací,
- řada druhů lokalit (např. hrady, kostely, nádraží, hřbitovy atd.) jsou dnes v tematických či jmenných autoritách, což je problematické - jejich záznamy nemají stejnou strukturu jako „pravé“ geografické autority (není geolink; problém bude markantnější, pokud se struktura geoautority bude dále specializovat, obohacovat o ryze „geografická“ pole),

- možno je uvažovat o obsáhlejší „životopisu“ geoautority.

B.5.1.5 Tematické termíny

Výsledek: současné národní tematické jsou pro účely sbírek RML online **nevyhovující**.

Komentář k výsledku:

- problém je v podřazených termínech (např. hodiny - hodiny sluneční, věžní; různé ikonografické typy Matky Boží atd.) - uživatel sbírek online (i jakékoli obdobné databáze) potřebuje vyhledat konkrétní typ hodin atd.,
- podřazené termíny (resp. některé současné) by musely být samostatnými autoritami (věžní hodiny jsou zcela něco jiného než sluneční) - podřazené by měly být pouze synonyma, lokální názvy téhož apod.

Poznámka k obecnému užití v paměťových institucích:

- tematické autority by měly mít vysvětlivku významu (charakteristiku) - v tomto případě to je zcela nezbytné,
- různé vědní obory mohou chápat různě tytéž termíny (názvy) - je proto nutné používat rozlišovací kvalifikátor,
- možno uvažovat o připojení/propojení na obrázky či další doplňující informace na webu.

B.5.1.6 Obecná úvaha na závěr

Každá „autorita“ (selekční prvek) by měla mít následující okruhy informací, tj. údajů, které tytéž termíny (slovo či slovní spojení) vzájemně jednoznačně odliší:

- časová informace (životopisná data osob, období historického názvu lokality atd.),
- geografická informace (místo/stát působení osoby, lokalizace geoautority atd.),
- oborové zařazení (obor působení osoby, rozlišovací kvalifikátor atd.),
- synonyma (odkazové formy jmen atd.),
- obecné vysvětlení (charakteristika).

Národní autority, které mají všechny uvedené oblasti informací, jsou vyhovující, dostatečné (pro sbírky RML online, pro jakékoli jiné obdobné databáze i pro obecné užití v paměťových institucích); jakékoli další oblasti informací jsou nadstavbové.

Národní autority bez jedné či více uvedených oblastí jsou nevyhovující³⁰.

B.5.2 Geografické názvy - lokality v Demusu

1. Administrativní územní jednotky: státy, katastry, okresy, obce, kraje

³⁰ KLIMEŠ, René. Analýza selekčních prvků Sbírek RML online - srovnání s Národními autoritami.

U všech schází časové vymezení platnosti názvu, vzájemné provázání různých názvů a různé podřízenosti v průběhu času.

2. Geografické oblasti vymezené vágně, bez topologie (mohou se všelijak překrývat).

Jen v entomologii existuje většinově uznávaný systém lokalit na území ČR, v ostatních oborech žádný společný systém není.

Pole název a volný slovní popis (schází zejména časové vymezení platnosti názvu, vzájemné provázání různých názvů téhož území).³¹

B.5.2.1 Geografické názvy v přírodních vědách

Motivací pro používání geografických názvů je zkratkovitá a zároveň dostatečně jednoznačná identifikace prostorových vazeb popisovaného objektu nebo fenoménu. Podle standardu CIDOC-CRM reprezentují geografické názvy instance třídy E44 *Označení místa*, pomocí kterých jsou identifikovány instance třídy E53 *Místo* - konkrétní oblasti ve vesmíru, zejména však na povrchu Země, čistě ve fyzikálním smyslu (nezávisle na dočasných jevech a hmotné podstatě).

B.5.2.2 Typologie geografických názvů v přírodních vědách

Dále jsou popsány dosud identifikované typologické skupiny, do kterých lze používané geografické názvy začlenit. Jistě lze navrhnout i další:

- podle vztahu k Zemi - pozemské, mimozemské,
- podle polohy vůči povrchu - podzemní, povrchové,
- podle temporální stability - trvalé, pomíjivé (efemérní),
- podle interpretační stability - pevné, kulturně podmíněné, vágně interpretované.

B.5.2.3 Fytochoriony (botanika)

Obvykle víceslovné názvy pro fytogeograficky ucelené jednotky, které představují paralelu např. ke geomorfologickému členění. V praxi se používá dvoustupňové dělení na okresy a podokresy.

B.5.2.4 Katastry, Lokality (všechny moduly)

Pro zápis zejména nálezových okolností se hojně využívá administrativního členění. I přes rozmach GPS navigace zůstává lokalizace pomocí blízké obce nejdůležitější a nejvyužívanějším postupem. Zejména v disciplínách, které hojně pracují s historickými údaji se lze zcela běžně setkat s dnes nepoužívanými místními jmény a jejich jazykovými variantami (hlavně v němčině a latině) - *Gottwaldov*, *Mährisch Weißkirchen*, *Prešpurk*, *Leningrad*, *Vindobona*.

³¹ CIGÁNEK, David. Geografické názvy v přírodních vědách : pracovní materiál pro řešitelskou skupinu projektu INTERPI. PODOLNÍKOVÁ, Jarmila - MAZAČOVÁ, Vladka. Lokality v systému Demus.

Názvy vyšších administrativních a státních celků bývají používány střídavě ve tvaru místního jazyka, střídavě ve tvaru bližšího jazykové příslušnosti badatele (Čechy = Böhmen, Kladsko = Klodsko), totéž platí i pro místní názvy, u kterých je navíc poměrně běžné komolení původního názvu při opakovaném převodu mezi jazyky a vytváření umělých názvů (Pulgarn = Pulgary = Bulhary, Schildberg = Šilperk = Šimperk = Štítý = Žalkov). Zvláště u často používaných německých názvů může být obtížné rozhodnout, zda se jedná o lokalizaci spjatou s alternativním názvem poměrně známého sídla nebo naopak jinde lokalizovanou, zcela obskurní polní trati téhož (podobného) jména.

U historických údajů je lokalizace nezřídka nejistá buďto v důsledku nejednoznačnosti jména (různé Lhoty a Újezdy !) nebo v důsledku použití jména, jehož geografická vazba později upadla v zapomnění (Cikánská kuchyně, Padělky) případně byla relativizována změnou ve vymezení administrativního útvaru (sloučené a odštěpené obce, změny státních hranic atd.). Zcela výjimečné je zneprášení lokalizace v důsledku změn přírodních tvarů (břehová čára vodních nádrží, průběh toků, hranice lesních porostů). Nezřídka se lze setkat s velmi povšechnými údaji, které by z dnešního pohledu zcela spolehlivě vedly k jejich vědecké nevyužitelnosti (křoví u Prahy, Morava, Alpy...). Pro eliminaci těchto problémů se v současné přírodovědné praxi používá buď doplňková lokalizace na úrovni vyšších správních jednotek (obec - okres) nebo připojení absolutních zeměpisných souřadnic (z GPS). Oba postupy však skrývají drobné záludnosti:

1. Lokalita identifikovaná pomocí nejbližší obce se nemusí nutně nacházet na jejím katastru. Zvláště u asymetrických katastrálních území s excentricky umístěnou zastavěnou částí může být neshoda výrazná;
2. Katastrální hranice se v minulosti měnily a stále mění. Totéž platí o okresech, které byly v průběhu administrativně-správních reforem mnohokrát reorganizovány. V případě krajů došly změny tak daleko, že doplňková identifikace lokality pomocí kraje je bez znalosti historického kontextu nepraktická až zavádějící;
3. U zeměpisných souřadnic je třeba uvádět i použitý souřadnicový systém. Vhodné je rovněž doplnění údaje o přesnosti měření, zvláště pokud není zřejmá metoda použitá při odečtu.

B.5.2.5 Jednotky (geologie)

Zařazení do stratigrafické jednotky vnáší do geografické lokalizace další dva vzájemně provázané rozměry. Prvním z nich je čas (geologické stáří), druhým hloubka pod povrchem jako složitá funkce stáří jednotky. Názvy bez přímé vazby na nadzemní geomorfologické členění se obvykle píšou s malým počátečním písmenem.

Jak bylo naznačeno dříve, lze diskutovat o tom, zda názvy těchto jednotek vůbec mají povahu geografických názvů. Lze je totiž nahlížet i jako atributy konkrétní lokality nebo vzdálenou paralelu techniky ve výtvarném umění.

B.6 Potenciální aplikace klasifikačního systému MDT jako nástroje propojení různých systémů

V tomto roce řešení projektu jsme se zabývali i teoretickým průzkumem a potřebnou úpravu klasifikačního systému MDT jako potenciálního propojovacího jazyka při řešení terminologické

různorodosti termínů aplikovaných v jednotlivých paměťových institucích a jako nástroje pro tvorbu ontologií.

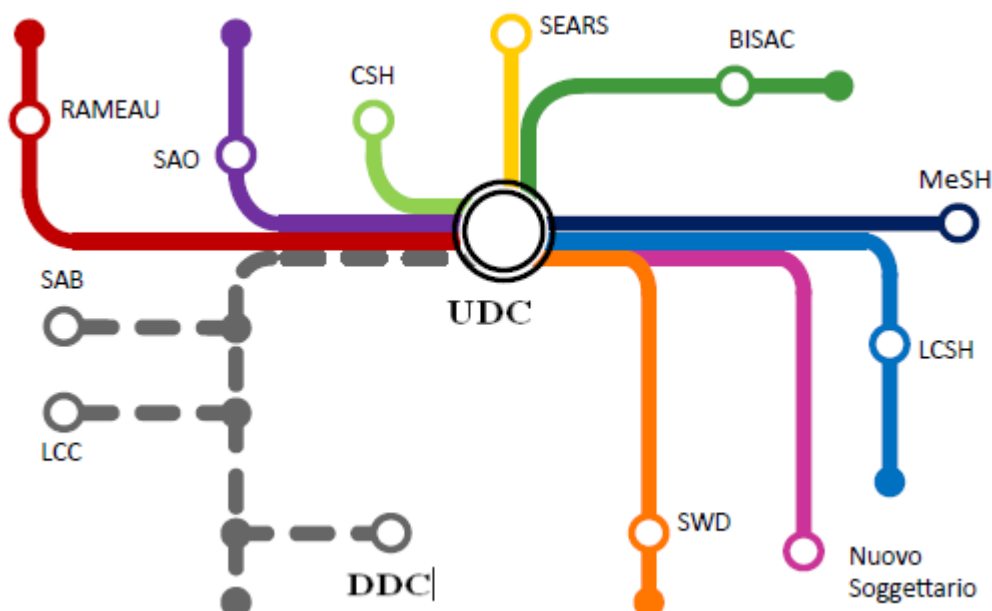
B.6.1 Paměťové instituce - rozdíly v systémech zpřístupnění

Je všeobecně známo, že notací klasifikačních systémů, tedy i MDT je možné využít jako symbolů propojovacího jazyka při tvorbě univerzálního a multilingválního systému organizace poznání. Jednotlivé systémy zpřístupnění aplikované v paměťových institucích se mohou lišit:

- ve struktuře, tj. odlišnosti ve vyjádření vztahů mezi jednotlivými koncepty,
- v oborech, které pokrývají: jednotlivé termíny v terminologických soustavách různých oborů jsou sémanticky závislé na hierarchických strukturách těchto oborů (akustika ve fyzice, hudbě a fonetice),
- v hloubce a rozsahu významu termínů použitých v jednotlivých systémech (živočich v zoologii a v psychologii),
- ve formální stránce použitých termínů - singulár, plurál, rozdílný pravopis,
- v syntaktické rovině - substantivní inverze či neinverze, uvádění předložek a jiných slovních propojovacích elementů,
- v principu prekoordinace či postkoordinace uplatněném v terminologické rovině, tedy představují-li tematické selekční prvky jednoduché termíny, případně fráze,
- v jazyce, systémy zpřístupnění jsou vázány na národní jazyky.

Odlišnosti jednotlivých systémů mohou být překonány pomocí notací MDT, které poslouží jako propojovací elementy.

Obr. 4: UDC/MDT jako nástroj mapování a propojení



B.6.2 Klasifikační systémy a konceptuální model FRSAD

Nyní stojíme před zásadní otázkou, je-li možné transformovat konceptuální model FRSAD - Funkční požadavky na předmětová autoritní data - na modelování klasifikačních systémů.

Základní schéma konceptuálního modelu FRSAD, ve kterém „nomen“ v případě klasifikačních systémů, např. MDT označuje notaci MDT, slovní vyjádření a URI.

Obr. 5: FRSAD - struktura prvků



Názory na transformaci modelu FRSAD na modelování klasifikačních systémů se liší, někteří odborníci považují oba systémy za natolik rozdílné, že aplikovat model FRSAD na klasifikační systémy není možné; jiní se domnívají, že oba systémy vykazují jisté podstatné společné rysy umožňující aplikaci konceptuálního modelu FRSAD i na klasifikační systémy. Řešení této otázky předpokládá další výzkum³².

B.7 Terminologická analýza³³

Předmětem terminologické analýzy je v první řadě pojem autorita, protože je v souvislosti s entitami, které jsou předmětem zájmu projektu, výchozí. Projekt se opírá o zkušenosti knihoven při vytváření záznamů autorit. Principy pro zpracování informací o autoritách jsou z historického hlediska prvními principy pro popis personálních, korporativních a geografických entit.

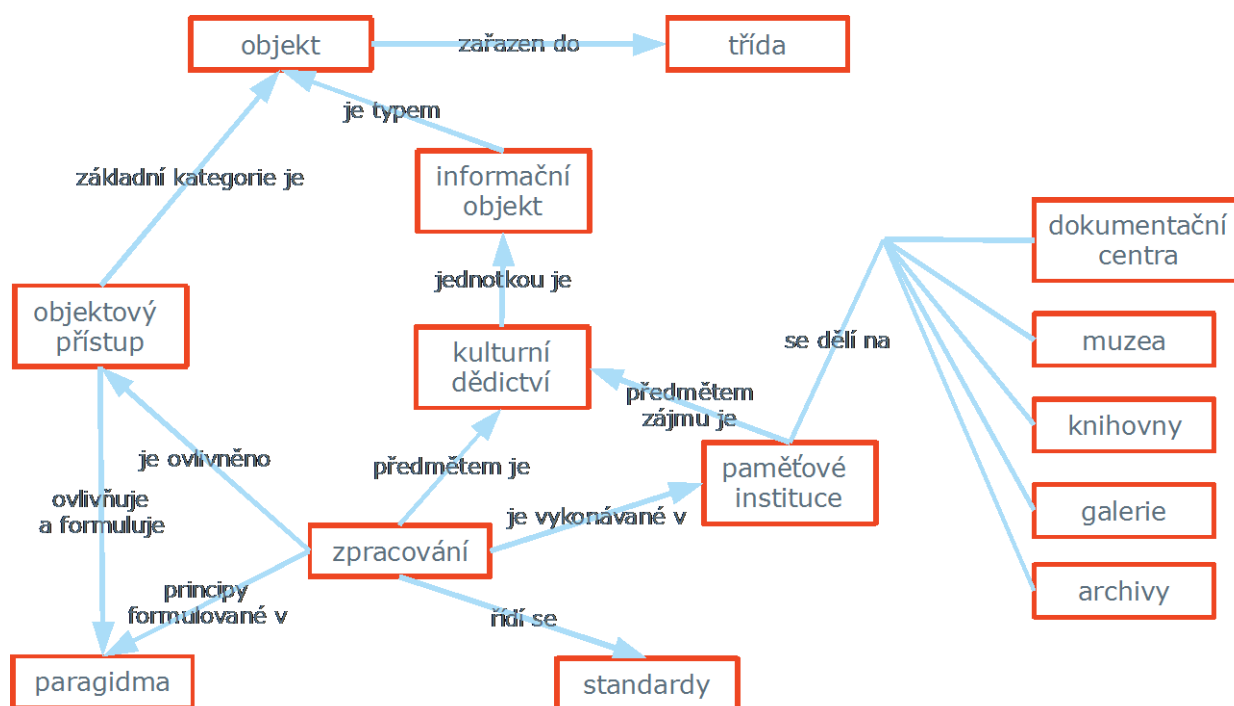
Dalšími pojmy, jejichž chápání je třeba sjednotit, jsou ty, které souvisejí s pojmem objekt. Důležitý je filozofický pohled i pohled informačních technologií na tuto oblast, protože se v projektu prolínají.

Vztahy mezi některými pojmy jsou značeny v pojmové mapě, kde je jako východisko pojem zpracování, který není součástí analýzy, ale je klíčovou odbornou činností v paměťových institucích (Obr. 6).

³² MITCHELL, Joan S. - ZENG, Marcia Lei - ŽUMER, Maja. Extending models for controlled vocabularies to classification systems : modelling DDC with FRSAD.

³³ Zpracováno podle Zprávy z analýzy metodologických postupů, technologií sémantického webu, datových úložišť a návrh datové struktury pro základní skupinu entit – viz: příloha F.4.

Obr. 6: Pojmová mapa zpracování a souvisejících pojmů



B.7.1 Autorita a související pojmy

Termín *autorita* souvisí s oblastí unifikace základních selekčních prvků vycházející z katalogizační praxe a ve své podstatě znamená unifikaci selekčních prvků, přístupových bodů. Jedna z forem záhlaví se označí za formu preferovanou, tj. *ustanovené*, *preferované*, *autorizované* nebo *autoritní* případně *autoritativní* záhlaví. Toto autorizované záhlaví tvoří nedílnou součást *záznamu autority*, který bývá definován: „záznam v souboru autorit, jehož určujícím prvkem je autorizované záhlaví pro entitu“³⁴. Skupina záznamů autorit tvoří *soubor autorit* tj. „soubor ověřených a unifikovaných jmenných a/nebo věcných selekčních údajů, určených pro zpracování a vyhledávání dokumentů s nezbytným odkazovým a poznámkovým aparátem“³⁵. Pro osobní jména, jména korporací, geografická jména, unifikované názvy, názvy edicí a obecné tematické termíny se většinou vytvářejí samostatné soubory autorit. Soubor autorit zahrnuje též odkazy z ustanoveného záhlaví.“³⁶

³⁴ IFLA. Guidelines for Authority Records and References [online], s. 2.

³⁵ VODIČKOVÁ, H. Soubor autorit.

³⁶ REITZ, J. M. ODLIS : Online dictionary for library and information science [online].

B.7.2 Typy autorit

Personální autorita (nebo také *osobní autorita*) je považována za základní typ. Jde o druh autority „obsahující jméno osoby jakožto autora nebo osoby, která je předmětem pojednání, nebo již se dokument týká.“³⁷ Základ záznamu personální autority je tvořen *personální záhlaví*.

S personálními autoritami úzce souvisí *autority rodiny/rodu*. Je to autorita obsahující jméno rodiny nebo rodu, přičemž rodina nebo rod představuje dvě nebo více osob, spojených narozením, svatbou, adopcí nebo stejným právním statusem. Základ záznamu je tvořen *záhlavím rodiny nebo rodu*.

Dalším nejrozšířenějším typem autorit je *autorita korporace* (též *korporativní autorita* nebo *autorita akce a korporace*). Je definovaná jako druh autority „obsahující jméno korporativního autora“³⁸. Základ záznamu je tvořen *korporativním záhlavím*.

Geografická autorita se používá pro informace o skutečných nebo fiktivních pozemských geografických nebo mimozemských lokalitách. Základ záznamu je tvořen *geografickým záhlavím*.

B.7.3 Objekt a související pojmy

Pojem *objekt* je širokospektrální a zasahuje do několika oblastí. V nejobecnější rovině se pod objektem chápe předmět nebo věc. Z hlediska informační vědy je pojem objekt používán v smyslu *informační objekt* a definován jako „jakýkoliv fragment skutečnosti, který se nachází ve vztahu se subjektem a je pro něho zdrojem informací“³⁹. Informační objekt je tedy ve vztahu k paměťovým institucím cokoli, co je předmětem jejich zájmu, odborných činností a má pro ně význam jako zdroj informací.

Jiný pohled na pojem objekt přináší informatika. Objekt je zde chápán více ve vztahu k informačním systémům.

B.7.3.1 Objekt, vztahy

V objektovém přístupu je objekt považován za entitu „ kterou je možné odlišit od jiných a která obsahuje data i možné operace s nimi“⁴⁰. Objekt je reálná entita, instance třídy⁴¹ a vztahuje se na něj abstrakce. O abstrakci se hovoří tehdy, pokud „klient modulu nepotřebuje

³⁷ VODIČKOVÁ, H. Personální autorita.

³⁸ VODIČKOVÁ, H. Korporativní autorita.

³⁹ ČABRUNOVÁ, A. Informačný proces, s. 144.

⁴⁰ KITJONGTHAWONKUL, Somkiat - KHOSLA, Rajiv. Modeling information systems using objects, agents, and task-based problemsolving adapters, s. 475.

⁴¹ MADAN, Alok. Object-oriented paradigm in programming for computer-aided analysis of structures, s. 228.

více, než je známo z rozhraní“⁴². Odlišení jednoho objektu od druhého se označuje jako identita⁴³. Mezi objekty vznikají vztahy - statické a dynamické. Statické vztahy znamenají dlouhodobé spojení objektů, dynamické vztahy představují komunikaci mezi objekty⁴⁴. Vztahy je též možné rozdělit na asociativní vztah, vztah agregace a vztah dědičnosti⁴⁵. Asociativní vztah znamená, že dva objekty spolu souvisejí - může jít o jednosměrný vztah nebo obousměrný (jeden objekt odkazuje na druhý a opačně). Pod vztahem agregace se rozumí, pokud jeden objekt pozůstává z více jiných objektů. Vztah dědičnosti zahrnuje vztah zobecnění a specifikace. Kromě vyjmenovaných vztahů mezi objekty je možné vymezit i vztahy mezi objektem a třídou.

B.7.4 Třídy, hierarchie a dědičnost

Objekty se začleňují do tříd - třídy určují typ objektů. Třída určuje vlastnosti objektů a činnosti, které je možné s objekty vykonávat. Objekty, které patří do konkrétní třídy, jsou označovány jako instance. Možné je rozlišovat abstraktní, konkrétní nebo smíšenou třídu⁴⁶. Abstraktní třída neobsahuje žádné objekty, může však být zapojena a použita pro dědičnost. Prvotním účelem konkrétní třídy je seskupení objektů, může též podporovat dědičnost.

Hierarchie jako další vlastnost objektového přístupu umožňuje nadřadit a podřadit jednu třídu jiné. Pokud má třída jako instance pouze jiné třídy označuje se jako „metatřída“ (z anglického „metaclass“). Metatřída je vždy abstraktní. S třídami se spojuje také dědičnost, která je též základní charakteristikou objektového přístupu. Dědičnost podporuje princip znovupoužití (reuse). Teoreticky je možné všechny vlastnosti dědit. Třída, která dědí vlastnosti a činnosti jiné třídy, se označuje jako podtřída, odvozená třída nebo specializovaná třída⁴⁷. Třída, ze které se vlastnosti dědí, je nadtřída, supertřída nebo základní, abstraktní nebo obecná třída⁴⁸. Dědičnost může být víceúrovňová. Slabým místem dědičnosti je fakt, že každá změna v nadtřídě může ovlivnit všechny podtřídy a současně při modelování široké oblasti je velmi náročné identifikovat vlastnosti podtřídy, pokud je dědí z více nadtříd.

B.7.4.1 Entita a související pojmy

Pojem *entita* je podobně jako pojem objekt zatížen množstvím různých významů od čistě filozofického přes matematiku po programování a informatiku. Nejširší je význam pojmu entita jako něčeho co existuje a souvisí s vědou o bytí. V informatice je často spojován s tvorbou informačních systému - konkrétně s entitně-relačním modelem.

⁴² STEVENS, P - POOLEY, R. Using UML : software engineering with objects and components, s. 10.

⁴³ PAETAU, Patrik. On the benefits and problems of the object-oriented paradigm including a finnish study, s. 27.

⁴⁴ Cit. 43, s. 27.

⁴⁵ Cit. 43, s. 37-38.

⁴⁶ Cit. 43, s. 33.

⁴⁷ Cit. 43, s. 39.

⁴⁸ Cit. 43, s. 39.

B.7.5 Terminologie - závěry

Za informační objekt se bude považovat obecně pro tuto fázi projektu osoba (ve smyslu fyzické osoby - reálná nebo fiktivní), korporace (ve smyslu instituce, dočasné korporace, sdružení apod.) a geografická entita. Obecně se všechny tyto kategorie označují jako entity. Výklad pojmu entita se neshoduje s pojmem autorita, tak jak tomu často bývá v knihovnické praxi. Výklad pojmu entita se pro projekt bude používat ve všeobecném filozofickém významu. Také není spojován s entitně-relačním modelem z informatiky. Pojem třída se bude používat pro označení skupiny entit se stejnými vlastnostmi - tj. pro osoby jako skupinu entit (jde o třídu osob), pro korporace (tj. jde o třídu korporací) a pro geografické entity (tj. jde o třídu geografických entit). Pojem třída se používá především tehdy, kdy je nutné zdůraznit rozdílnost jednotlivých entit.

Pojem autorita se kvůli jejímu specifickému významu pro oblast knihoven nebude v projektu využívat pro označení informačních objektů nebo souvisejících informací vůbec.

B.8 Specifikace požadavků na datovou strukturu⁴⁹

Specifikace požadavků se opírá o potřeby všech typů paměťových institucí, pro úspěšnou realizaci projektu je nutné se dohodnout na společných rysech, aplikaci odpovídajících standardů a hloubce a způsobu zpracování informací o jednotlivých typech entit.

Pro specifikaci požadavků na datovou strukturu byly ustanoveny některé základní principy:

- pro dataci je žadané upřednostňovat princip stanovený standardem CCO - tj. možnost zapsat dataci v textové podobě a také ve strojem čitelné a zpracovatelné formě,
- entity budou mít vlastní ID nezávislé na systému Národní autorit ČR, ale současně bude možné zapsat ID jiných systémů.

Požadavky jsou v každém směru otevřeny a předpokládá se jejich doplňování v rámci odborné diskuse.

B.8.1 Entity v projektu INTERPI

Požadavky u jednotlivých typů entit lze rozdělit do několika oblastí uvedených níže.

B.8.1.1 Oblast vlastní identifikace, typ entity, stav záznamu

1. ID /identifikační číslo entity
2. Typ entity
 - Typ entity INTERPI N- personální, korporativní, GEO
 - Typ personální entity INTERPI - N: osoba, rod/rodina

⁴⁹ Zpracováno podle Zprávy z analýzy metodologických postupů, technologií sémantického webu, datových úložišť a návrh datové struktury pro základní skupinu entit – více viz: příloha F.4.

- Typ autoritního záznamu NA
- 3. Stav záznamu INTERPI: rozpracovaný - definitivní

B.8.1.2 Oblast jména

1. Identifikace- oblast jména;
2. Určení jazykové oblasti, do které daná entita spadá, pohlaví, profesního zařazení;
3. Poskytnutí detailnějších informací týkajících se života, profesní kariéry, členství v různých organizacích apod., tedy informací nutných k získání globálního pohledu na danou personální entitu a uvádějí tuto entitu do potřebných souvislostí (vztahů a vazeb).

Identifikace entity - oblast jména

Preferovaná forma jména -O (podle různých pravidel)

- Preferovaná jména INTERPI - povinná

Rozsah/struktura údajů:

- Jméno - Forma jména
- Struktura zápisu
- Datace- datum od - datum do
- Jazyk jména (pro určení zda je pro popis použita forma jména v konkrétním jazyku - např. při jménech se slovním doplňkem - šlechtickými tituly, označeními svatých apod.)
- Zdroj údajů
- Pravidla
- Poznámka
- V případě potřeby uvádět také preferovanou formu jména podle jiných pravidel, např. AACR2/RDA

Variantní forma jména - O / Paralelní forma jména - O

Rozsah a struktura údajů jako u preferované formy jména.

Povinně uvádět všechny variantní formy. Podle potřeby uvádět paralelní formy jména.

K diskusi: rozlišování paralelních a variantních forem jména nebo jejich uvádění společně.

B.8.1.3 Určení profesního zařazení a jazykové oblasti, do které daná entita spadá,

- Jazykové určení (jazyk používaný entitou)
- Příslušnost k zemi původu, působnosti (národní komunitě) - formulace vychází z potřeby nahradit nevhodné označení „národnost“(u personálních entit)
- Pohlaví (u personálních entit)
- Zařazení
 - Klíčové slovo
 - Profese
 - Styl
 - Obor

B.8.1.4 Detailnější informace uvádějící entitu do potřebných souvislostí, vztahů a vazeb

Související informace je možné vyjádřit pomocí souvisejících entit propojených prostým vztahem asociace (viz též) nebo diferencovaně pomocí komplexních vazeb a vztahů, prostřednictvím jednotlivých událostí významných pro danou entitu, na nichž se podílejí všichni aktéři.

Za významné události se u entity považují: narození/vznik, úmrtí/zánik, studium, profesní kariéra/zaměstnání, členství v různých organizacích, významné aktivity (proslavení), rodinné a partnerské vztahy, status korporace, typ korporace, logo, používaná výrobní značka apod.

Obecná struktura události

- Datace
- Místo
- Osoba
- Instituce (korporace)
- Poznámka
- Zdroj

Doporučená struktura jednotlivých typů události (personálie):

- Narození - datace, místo, osoba (otec, matka), poznámka, zdroj
- Úmrtí - datace, místo, osoba (pachatel, apod.), poznámka, zdroj
- studium - datace, místo, korporace (instituce), osoba (učitel), poznámka, zdroj
- profesní kariéra/zaměstnání - datace, místo, korporace (instituce), osoba (nadřízený, podřízený), poznámka, zdroj
- členství v různých organizacích - datace, místo, korporace (instituce), osoba (člen, vedení organizace), poznámka, zdroj
- významné aktivity (proslavení) - datace, místo, osoba, dílo/výtvor/čin, poznámka, zdroj
- rodinné a partnerské vztahy (uzavření manželství, partnerství, ukončení manželství partnerství) - datace, místo, osoba (svědek), korporace, poznámka, zdroj

Doporučená struktura pro jednotlivé typy události (korporace):

- vznik: datace, místo, typ (založení, odštěpení, sloučení apod.), osoba (vlastník, obchodní partner apod.), korporace, poznámka, zdroj
- zánik: datace, místo, typ (prodej, úpadek, katastrofa, konkurs, sloučení, zrušení apod.), osoba (vlastník, obchodní partner, věřitel, akcionář, vlastník apod.), korporace, poznámka, zdroj
- přearžení: datace, místo, osoba, korporace, poznámka, zdroj
- změna vlastnictví: datace, místo, osoba, korporace, poznámka, zdroj
- změna právní formy, funkce, zaměření nebo sídla: datace, místo, osoba, korporace, poznámka, zdroj
- členství: datace, místo, osoba (člen, zástupce), korporace, poznámka, zdroj

Související entity - čerpat z databáze (odkaz na příslušné entity)

Kategorie vztahu (podle kategorie související entity níže). Pro zápis např. rodinných příslušníků (sourozenci, děti, vnuci apod.).

- Datace
- Poznámka
- Zdroj
- Entita

Typy souvisejících entit

- Personální entita (osoba/rod)
- Entita korporace
- Entita dočasná korporace (akce)
- Geografická entita
- Entita předmět (pojem/fyzický objekt)

B.8.1.5 Životopis/historie

Volný text

B.8.1.6 Zaměření, činnost

Volný text. Např. v současnosti součást biografické poznámky pro NA ČR.

B.8.1.7 Informační prameny / dokumenty

K diskusi: způsob uvádění informačních pramenů, dokumentů jako samostatných záznamů, nebo jako součást popisu entity a způsob použití v souvislosti s uváděním zdroje při konkrétních datech.

B.8.1.8 Poznámka (veřejná, neveřejná)

B.9 Analýza metodologických postupů

V této zprávě uvádíme pouze analýzu standardů, které jsou z hlediska řešení projektu klíčové, analýza všech standardů je obsažena ve Zprávě z analýzy⁵⁰

⁵⁰ Viz: příloha F.4 Zpráva z analýzy metodologických postupů, technologií sémantického webu, datových úložišť a návrh datové struktury pro základní skupinu entit

B.9.1 GARR

Informace o standardu jsou zpracovány v původním jazyce - tj. v angličtině protože není k dispozici oficiální překlad. Je možné do budoucna uvažovat o překladu důležitých částí standardu. Standard GARR je jedním z výchozích pro další standardy jako FRBR, FRAD, FRASD.

B.9.1.1 Karta standardu

Zkratka	GARR
Originální název standardu	Guidelines for authority records and references
Odpovědná instituce	IFLA
Verze	2nd edition Guidelines for authority records and references. – 2nd ed. / revised by the IFLA Working Group on GARE Revision. – München : K.G. Saur, 2001. – xii, 46 pages ; 25 cm. – (UBCIM publications ; new series, vol. 23). – ISBN 978-3-598-11504-2.
Datum	2001
www standardu	http://archive.ifla.org/VII/s13/garr/garr.pdf
Typ standardu	popisní
Překlad standardu	
Oficiální překlad názvu	--
Odpovědná organizace	--
Datum vydání	--
Verze	--
Překlad z verze	--
www překladu	--
Souvislosti	
Související standardy	GSARE - Guidelines for Subject Authority and Reference Entries (2nd ed. - 1998)
Jiné dokumenty	

B.9.1.2 Cíle standardu

GARR určuje požadavky pro zobrazení informací týkajících se autorizovaných záhlaví a odkazů v různých formách prezentace v bibliografiích, katalozích apod. Pravidla pokrývají autorizovaná záhlaví a odkazy pro osoby, korporace a díla. Cílem je též vytvořit obecné pravidla pro mezinárodní výměnu a využití informací o autorizovaných záhlavích a odkazech. Důležitou částí GARR jsou definice základních pojmů z oblasti autorit, které mohou být východiskem pro všechny ostatní standardy.

B.9.2 FRBR

Pozn.: Informace o standardu jsou převzaty z českého překladu standardu.

B.9.2.1 Karta standardu

Zkratka	FRBR
Originální název standardu	Functional Requirements for Bibliographic Records
Odpovědná instituce	IFLA
Verze	Z 2.2.2009 - s úpravami v paragrafe 3.2, pdf soubor 1,2MB 142stran
Datum	2.2.2009
www standardu	http://www.ifla.org/publications/functional-requirements-for-bibliographic-records
Typ standardu	Entitno-relačný model
Překlad standardu	
Oficiální překlad názvu	Funkční požadavky na bibliografické záznamy
Odpovědná organizace	NK ČR, překlad L. Celbová
Datum vydání	2001
Verze	
Překlad z verze	
www překladu	http://www.ifla.org/files/cataloguing/frbr/frbr-cs.pdf
Souvislosti	
Související standardy	FRAD, FRASAD, FRBRoo, CIDOC-CRM
Jiné dokumenty	

B.9.2.2 Cíle standardu

Studie má dva primární cíle. Prvním je poskytnout jasně definovaný, strukturovaný rámec, uvádějící do souvislosti údaje zaznamenané v bibliografických záznamech a vyhovující potřebám uživatelů těchto záznamů. Druhým cílem je doporučit základní úroveň funkčnosti pro záznamy vytvořené národními bibliografickými agenturami.

B.9.3 FRAD

B.9.3.1 Karta standardu

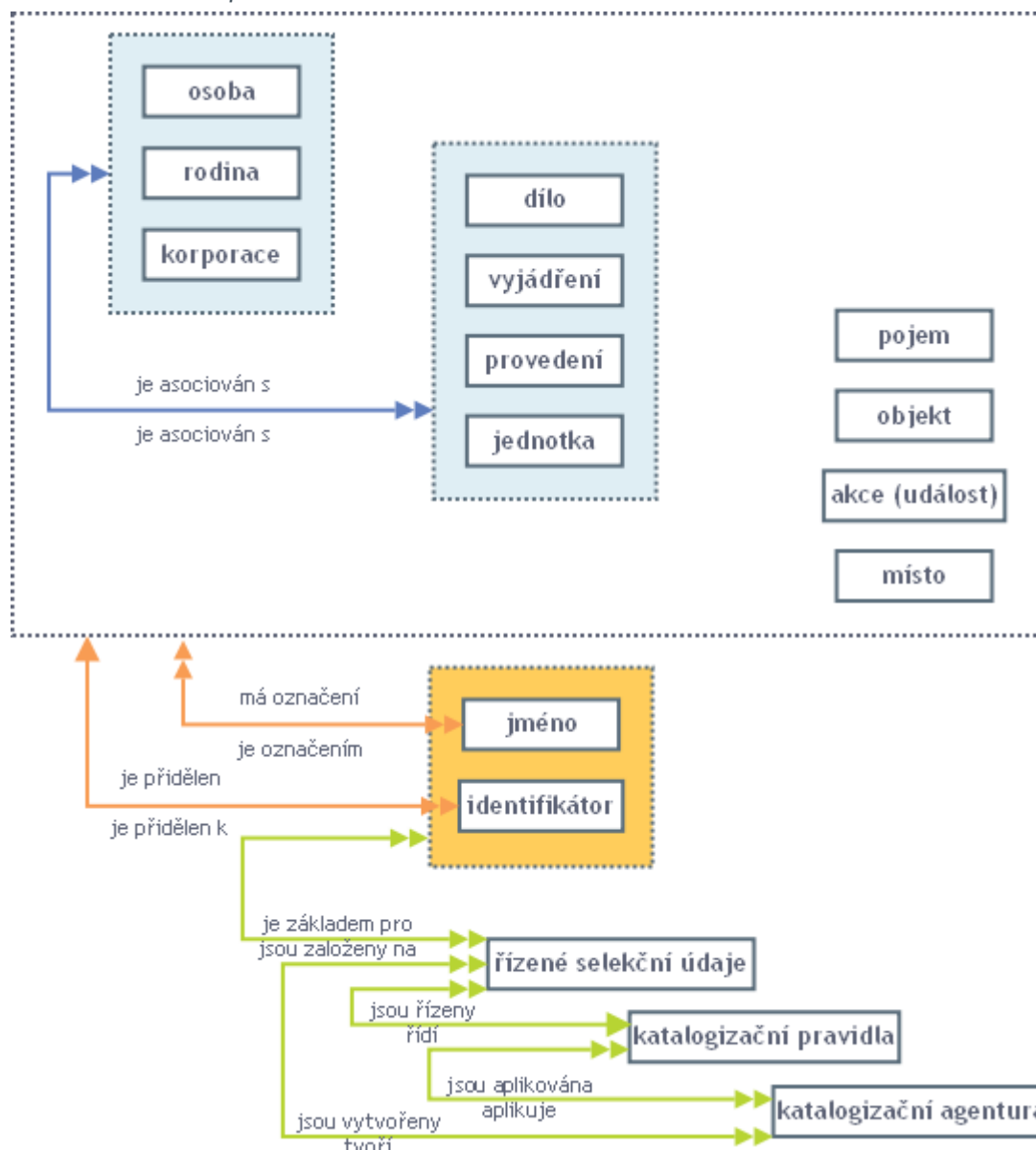
Zkratka	FRAD
Originální název standardu	Functional Requirements for Authority Data
Odpovědná instituce	IFLA Working Group on Functional Requirements and Numbering of

	Authority Records
Verze	Final Report
Datum	Prosinec 2008, vydáno 2009
www standardu	http://www.ifla.org/publications/functional-requirements-for-authority-data
Typ standardu	Konceptuální model
Překlad standardu	
Oficiální překlad názvu	bez oficiální názvu; podle překladu: Funkční požadavky na autoritní data
Odpovědná organizace	zatím nestanovena
Datum vydání	v českém jazyce nevydáno
Verze	--
Překlad z verze	--
www překladu	--
Souvislosti	
Související standardy	FRBR - Funkční požadavky na bibliografické záznamy FRSAD - Funkční požadavky na předmětová autoritní data
Jiné dokumenty	--

B.9.3.2 Cíle a hlavní myšlenky standardu

Konceptuální model FRAD - Funkční požadavky na autoritní data byl vytvořen v úzké souvislosti k modelu FRBR - Funkční požadavky na bibliografické záznamy. Podrobněji modeluje reálný svět fyzických osob a korporací, které byly definovány v modelu FRBR v tvůrčím nebo vlastnickém vztahu vůči entitám: *dílo*, *vyjádření*, *provedení* a *jednotka*. Vedle světa reálných fyzických a právnických osob se snaží model též reflektovat svět bibliografických dat. Patrné je to zejména v oblasti bibliografických identit. Model uvažuje možnost vztahu jedné reálné osoby a více různých jmen. Tvůrci modelu zdůrazňují, že model je obecný a lze jej aplikovat na širší doménu autoritních dat nejen na systémy autoritních záznamů budovaných v národních bibliografiích. Model navíc zobrazuje v základních entitách a vztazích katalogizační proces, čímž se liší jak od FRBR, tak i od návazného modelu FRSAD.

Obr. 7: FRAD - model prvků



B.9.4 FRSAD

B.9.4.1 Karta standardu

Zkratka	FRSAD
Originální název standardu	Functional Requirements for Subject Authority Data
Odpovědná instituce	IFLA Working Group on the Functional Requirements for Subject Authority Records (FRSAR)
Verze	Final Report
Datum	Červen 2010
www standardu	http://www.ifla.org/node/1297

Typ standardu	Konceptuální model
Překlad standardu	
Oficiální překlad názvu	bez oficiální názvu; podle překladu: Funkční požadavky na předmětová autoritní data
Odpovědná organizace	zatím nestanovena
Datum vydání	v českém jazyce nevydáno
Verze	--
Překlad z verze	--
www překladu	--
Souvislosti	
Související standardy	FRBR - Funkční požadavky na bibliografické záznamy FRAD - Funkční požadavky na autoritní data
Jiné dokumenty	--

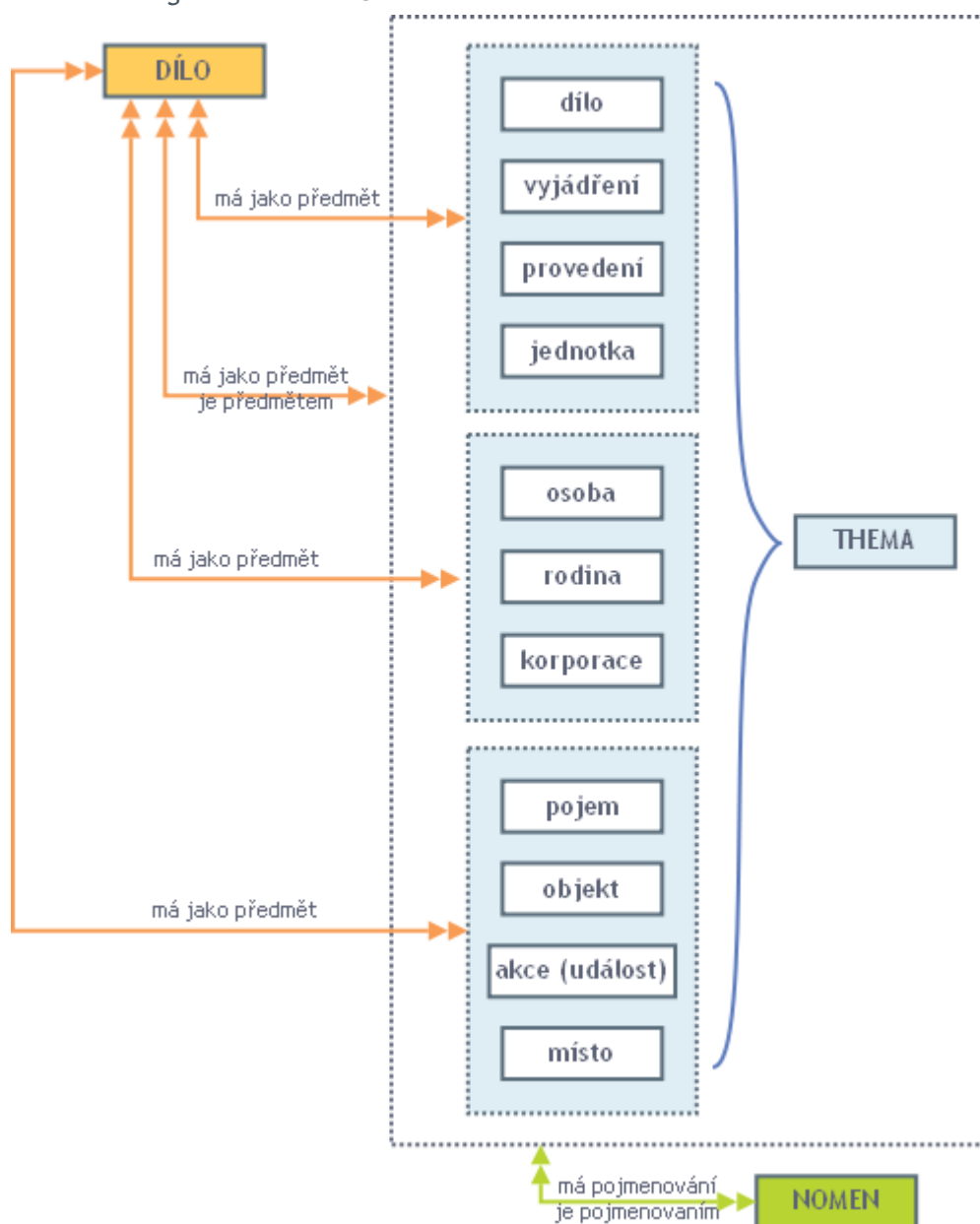
B.9.4.2 Cíle a hlavní myšlenky standardu

Konceptuální model FRSAD - Funkční požadavky na předmětová autoritní data je třetím doplňujícím modelem k modelům FRBR a FRAD. Reflektuje třetí skupinu entit FRBR: *pojem, objekt, místo a událost (akce)*. Cílem je vytvořit model klíčových entit předmětových autoritních databází tedy klíčových entit, které pokrývají předmětové (tematické) zaměření děl (definice podle FRBR). Autoři modelu FRSAD rozhodli, že entity definované reflektují pouze oblast předmětu díla (aboutness), tj. o čem dílo pojednává, nikoliv žánrové zařazení (isness), tj. že je dílo např. románem, sonatinou, životopisem nebo sochou či obrazem.

Obr. 8: FRSAD - struktura prvků



Obr. 9: Základní diagram modelu FRSAD



B.9.5 RDA

Informace o standardu jsou zpracovány v původním jazyce - tj. v angličtině protože není k dispozici oficiální překlad. Překlad standardu je klíčový.

B.9.5.1 Karta standardu

Zkratka	RDA
Originální název standardu	Resource Description and Access
Odpovědná instituce	American Library Association, Canadian Library Association,

	CILIP: Chartered Institute of Library and Information Professionals
Verze	
Datum	2010
www standardu	http://www.rdatoolkit.org/
Typ standardu	popisný
Překlad standardu	
Oficiální překlad názvu	--
Odpovědná organizace	--
Datum vydání	--
Verze	--
Překlad z verze	--
www překladu	--
Souvislosti	
Související standardy	AACR2; ISBD; MARC 21 Formát pro autority; MARC 21 Bibliografický formát, Dublin Core
Jiné dokumenty	Schválené české interpretace AACR2

B.9.5.2 Cíle standardu

RDA je nový standard pro popis zdrojů a vstupních prvků, vytvořený pro digitální prostředí. RDA přímo navazuje na katalogizační pravidla AACR2, kromě způsobu zápisu údajů se však více orientuje na to, které údaje kam patří podle Funkčních požadavků na tvorbu autoritních záznamů (FRAD).

Z pohledu národních autorit je podstatné, že RDA definuje rozsah údajů, které jsou důležité pro vymezení autority a nezaměřuje se přitom na autoritu pouze z hlediska selekčních prvků, ale také na to, aby záznam autority obsahoval dostatečnou informační hodnotu.

Z pohledu objektově orientovaného přístupu je podstatné, že RDA definují nejen rozsah údajů důležitých pro identifikaci a vymezení entity z hlediska jejich informační hodnoty, ale i oblast vzájemných vztahů, které jsou pro tvorbu znalostní báze podstatné.

B.9.6 CCO

B.9.6.1 Karta standardu

Zkratka	CCO
Originální název standardu	Cataloging cultural objects
Odpovědná instituce	Visual Resources Association Foundation, Cataloging Cultural Objects Committee

Verze	2006 (pdf verze dostupná online: http://cco.vrafoundation.org/index.php/toolkit/cco_pdf_version/)
Datum	2006
www standardu	http://cco.vrafoundation.org/
Typ standardu	popisní
Překlad standardu	
Oficiální překlad názvu	--
Odpovědná organizace	--
Datum vydání	--
Verze	--
Překlad z verze	--
www překladu	--
Souvislosti	
Související standardy	CDWA VRA Core AARC Describing Archives : A Content Standard (DACS)
Jiné dokumenty	Thesaurus of Graphic Materials AAT NISO. A Framework of Guidance for Building Good Digital Collections

B.9.6.2 Cíle standardu

Standard zaměřen na pravidla pro obsah dat v popisní katalogizaci - pravidla, které pomáhají vybrat termíny, a které vymezují pořad, syntax a formu ve které jsou tyto termíny, fráze, hodnoty a narativní popisy zaznamenávány. Primárním cílem standardu je popsat metadata a principy autorit - údaje, které jsou potřebné pro popis kulturních děl, data používány pro tvorbu katalogizačních záznamů pro tyto díla a jejich zobrazení (obrázky). Standard pokrývá mnoho typů kulturních děl, včetně architektury, malby, sochařství, tisku, rukopisů, fotografií a jiných vizuálních médií, živé umění (divadlo, tanec, hudba), archeologické naleziště a artefakty a rozličné předměty hmotné kultury.

B.9.7 ISAAR (CPF)

B.9.7.1 Karta standardu

Zkratka	ISAAR (CPF)
Originální název standardu	International Standard Archival Authority Record for Corporate Bodies, Persons, and Families
Odpovědná instituce	International Council on Archives / Mezinárodní archivní rada
Verze	2nd edition / 2ème édition
Datum	2004

www standardu	http://www.wien2004.ica.org/en/node/30230
Typ standardu	Popisný
Překlad standardu	
Oficiální překlad názvu	ISAR(CPF): Mezinárodní standard pro archivní autoritní záznamy korporací, osob a rodů
Odpovědná organizace	odbor archivní správy a spisové služby MV ČR, Milady Horákové 133, 166 21 Praha 6
Datum vydání	2009
Verze	2. vydání
Překlad z verze	2nd edition / 2ème édition
www překladu	http://www.wien2004.ica.org/en/node/39464
Souvislosti	
Související standardy	ISAD(G) - General International Standard Archival Description, Second edition; EAC - Encoded Archival Context; EAD - Encoded Archival Description
Jiné dokumenty	Možnosti a formy zpřístupnění archivních fondů nebo jejich součástí veřejnosti v elektronické podobě. Závěrečná zpráva projektu. Dostupné z WWW < http://www.nacr.cz/G-vyzk/moznosti.aspx >

B.9.7.2 Cíle standardu

ISAAR(CPF) definuje strukturu a obsah autoritních záznamů zejména v případě původců archiválií - fyzických osob, korporací a rodů. Nezabývá se tvorbou autoritního záhlaví. Uvádí také příklady použití a způsoby definice vazeb mezi autoritními záznamy vzájemně i mezi záznamy a archiváliemi. K použití uvádí: „Archivní autoritní záznamy mohou být použity: a) k popisu korporace, osoby nebo rodu jako jednotek v systému archivního popisu nebo b) k řízení vytváření a využívání přístupových bodů v archivním popisu, c) k dokumentování vztahů mezi různými původci dokumentů, vztahů mezi nimi a dokumenty či jinými zdroji jimi vytvořenými nebo vytvořenými o nich.“ [ISAAR 1.2, s. 62]. Překlady do národních jazyků reflektují příslušné národní prostředí.

B.9.8 ISAD(G)

B.9.8.1 Karta standardu

Zkratka	ISAD (G)
Originální název standardu	General International Standard Archival Description
Odpovědná instituce	International Council on Archives / Mezinárodní archivní rada
Verze	2nd edition / 2ème édition
Datum	2000
www standardu	http://www.wien2004.ica.org/en/node/30000

Typ standardu	Popisný
Překlad standardu	
Oficiální překlad názvu	ISAD(G) : Všeobecný mezinárodní standard pro archivní popis
Odpovědná organizace	odbor archivní správy a spisové služby MV ČR, Milady Horákové 133, 166 21 Praha 6
Datum vydání	2009
Verze	2. vydání
Překlad z verze	2nd edition / 2ème edition
www překladu	http://www.wien2004.ica.org/en/node/39463
Souvislosti	
Související standardy	ISAAR(CPF) - International Standard Archival Authority Record for Corporate Bodies, Persons, and Families; EAC - Encoded Archival Context; EAD - Encoded Archival Description
Jiné dokumenty	Možnosti a formy zpřístupnění archivních fondů nebo jejich součástí veřejnosti v elektronické podobě. Závěrečná zpráva projektu. Dostupné z WWW < http://www.nacr.cz/G-vyzk/moznosti.aspx >

B.9.8.2 Cíle standardu

ISAD(G) definuje víceúrovňový hierarchický popis archiválií (skupin archiválií - fondů, sérií apod.) a je zaměřen na tvorbu archivní pomůcky jako prostředku pro vyhledávání archiválií. Nestanovuje přesný popis jednotlivých druhů archiválií (proto všeobecný) ani na jaké úrovni se jednotlivé položky popisu používají. Zahrnuje také v minimální míře popis původců (osob, korporací, rodů). Výslovně se hovoří o tom, že „Sada všeobecných pravidel archivního popisu je součástí procesu, který (...) c. umožní sdílení databáze autorit“ [I.5, s. 10]. Překlady do národních jazyků reflektují příslušné národní prostředí.

B.9.9 CIDOC-CRM

B.9.9.1 Karta standardu

Oficiální český překlad standardu není k dispozici, použit (s laskavým svolením) nepublikovaný překlad modelu CIDOC CRM Z. Lenharta z r. 2006.

Zkratka	CIDOC-CRM
Originální název standardu	CIDOC-CRM (CIDOC Conceptual Reference Model), ISO 21127 (Information and documentation - A reference ontology for the interchange of cultural heritage information)
Odpovědná instituce	CIDOC Documentation Standards Working Group CIDOC CRM SIG
Verze	5.0.2
Datum	2010
www standardu	http://www.cidoc-crm.org/

Typ standardu	model
Překlad standardu	
Oficiální překlad názvu	
Odpovědná organizace	
Datum vydání	
Verze	
Překlad z verze	
www překladu	
Souvislosti	
Související standardy	FRBRer, FRBEoo, EAD
Jiné dokumenty	

B.9.9.2 Cíle standardu

CRM představuje pojmový referenční rámec určený k podpoře integrace, zveřejňování a výměny heterogenních informací o kulturním dědictví; je vyjádřen jako objektivně objektivě orientovaný sémantický model.

Podrobný výběr definicí pro CIDOC-CRM bude zpracován v další části této fáze projektu. Ukazuje se, že CIDOC-CRM je příliš obecný standard deklarující univerzálnost a obsahující velmi obecné principy, které je nutné podrobit detailní analýze a blíže definovat, mají-li být důsledně aplikovány na popis objektů měnících se v prostoru a čase (což je hlavní charakteristikou pro objekty, na které je projekt INTERPI zaměřen. Důkazem náročnosti a obtížnosti důsledné aplikace CIDOC-CRM na popis objektů měnících se v prostoru a čase je okolnost, že zatím se jen málo institucí a odborných skupin (i v zahraničí) věnovalo tomuto úkolu.

„Samotnou strukturou a formalismem je CIDOC-CRM rozšiřitelný a vybízí uživatele k vytváření nástaveb (extensions) pro potřeby specializovaných společenství a aplikací.

Ve snaze využít obsahové součinnosti, již CRM poskytuje, budou asi uživatelé chtít vytvářet části svých datových struktur v souladu s CRM. Takové části se musí týkat buď vztahů, podle nichž by uživatelé chtěli mít svá data přístupná v integrovaném prostředí, nebo obsahu, jež má být přenášén do jiných prostředí, aby význam zakódovaný strukturou byl zachován i v cílovém prostředí.

V tomto smyslu není cílem CRM navrhovat úplné přizpůsobení dokumentačních struktur uživatele podle CRM, ani tlačit uživatele, aby vždy zabudoval všechny pojmy a vazby z CRM. CRM je veden snahou ponechat prostor pro nejruznější doplňky k pokrytí košatosti kulturních informací, ale také pro zjednodušení z důvodů ekonomických“ Lenhart, 2006.

V této fázi projektu CIDOC-CRM zůstává klíčovým zdrojem inspirace, navržená datová struktura je plně kompatibilní s modelem CIDOC-CRM.

B.10 Návrh datové struktury

Návrh datové struktury představuje jedinečné řešení uspořádání dat pro zpracování personálních entit, korporací a geografických entit v komplexnosti a rozsahu, jaké doposud nebylo navrženo a publikováno. Ojedinělost návrhu spočívá v aplikaci všeobecných principů (viz níže) s ohledem na požadavky odsouhlasené všemi typy paměťových institucí.

Datová struktura současně respektuje mezinárodní standardy a je kompatibilní s modelem CIDOC-CRM.

Pro návrh datové struktury jsou formulovány následující základní principy, z kterých především znovupoužitelnost, dědičnost, polymorfismus a zapouzdření vycházejí z objektového principu nebo objektového paradigmatu využívaného v informatice:

- Atomizace - data jsou v dostatečné míře rozděleny na menší části, resp. části, které už není efektivní dál dělit. V jednotlivých menších částech je možné lépe rozpoznat základní datové typy, je možné je více zevšeobecňovat a v datové struktuře pak maximálně využít principy dědičnosti a znovupoužitelnosti;
- Znovupoužitelnost - jednotlivé prvky musí být navrženy tak, aby se efektivně daly využít na více místech, případně umožňovaly další rozvoj datové struktury v budoucnosti;
- Dědičnost - hierarchická struktura jednotlivých prvků předpokládá a podporuje dědění vlastností směrem k prvkům na nižší úrovni;
- Zapouzdření - definice prvků v sobě obsahuje jednak hodnotu (data) a operace, které je s nimi možno vykonávat - v případě datové struktury jsou to především vlastnosti, které jsou pro daný prvek definovány;
- Polymorfismus - jednotlivé prvky mohou v závislosti od způsobu použití působit navenek jako několik různých prvků - všeobecně polymorfismus znamená „jedno rozhraní, více funkcí“⁵¹.

B.10.1 Prvky datové struktury

Popis jednotlivých prvků datové struktury se skládá z následujících částí:

- název prvku - je vybrán s ohledem na celou datovou strukturu, měl by naznačovat obsah a účel prvku a zároveň se nesmí v rámci datové struktury vyskytovat více než jeden prvek s daným názvem;
- popis prvku - stručná charakteristika prvku s případným vysvětlením použití;
- typologie prvku - určuje, zda jde o metaprvek, základní prvek nebo hodnotu a také zda je prvek možné při použití opakovat nebo ne. Opakovatelné prvky je možné vždy pomocí příslušné vlastnosti, pro kterou tvoří obor hodnot, spojit s jiným prvkem, který je jejím definičním oborem, několikrát (resp. v tolika opakováních, kolik zpracování vyžaduje). Neopakovatelné prvky je naopak možné použít pouze jednou;

⁵¹ MADAN, Alok. Object-oriented paradigm in programming for computer-aided analysis of structures, s. 231.

- nadřazený prvek - obsahuje odkaz na nejbližší nadřazený prvek (nebo prvky) podle hierarchické struktury;
- podřazený prvek - obsahuje odkaz na nejbližší podřazený prvek (nebo prvky podle hierarchické struktury);
- vlastnosti a zděděné vlastnosti - obsahuje přehled vlastností, které jsou pro daný prvek definovány, nebo které prvek zdědil z nadřazeného prvku (nebo nadřazených prvků ve více úrovních). V definici prvků nejsou uváděny tzv. zpětné vlastnosti - tj. vlastnosti, které spojují prvek z oboru hodnot s prvkem definičního oboru vlastnosti, označení zpětných vlastností je uvedené pouze v definici vlastnosti;
- pravidla - určuje, jaká pravidla se využívají při výběru, formě a struktuře hodnot;
- hodnoty - určuje povolenou strukturu hodnot;
- případy použití - definuje, jaké hodnoty může prvek obsahovat v závislosti na použití prvku, nebo určuje jiné specifiky použití prvku pro jednotlivé třídy entit; pokud nejsou případy použití vymezeny, je možné prvek použít libovolným způsobem, ale v souladu s jinými vlastnostmi (např. pravidly pro zápis prvku).

Pro smysluplný popis je důležité dodržet minimální strukturu entity - proto jsou prvky, které se v záznamu entity musí vyskytnout označeny v názvu hvězdičkou (*).

Vlastnosti, vycházející z principu CIDOC-CRM, jsou navrženy jako nepovinné, protože je cílem akceptovat alternativní názory a neúplný popis, je však v zájmu srozumitelnosti popisu některé vlastnosti považovat jako za nevyhnutelné a doporučené - tyto jsou v definici prvku taktéž označeny hvězdičkou (*).

Jak již bylo zmíněno výše - definice prvků taktéž určuje opakovatelnost jednotlivých prvků, co není v souladu s CIDOC-CRM, ale tento princip nebyl zachován z důvodu předcházení nejednoznačnosti popisu a určení hranic pro možnou interpretaci datové struktury, které zabezpečí větší interoperabilitu.

V této fázi projektu není do datové struktury podrobněji zapracována problematika umělých výtvorů nebo výkonů související s proslavením a také symbolů, znaků a objektů nehmotného kulturního dědictví souvisejících s entitou. Tato specifická a metodologicky ojedinělá oblast si vyžaduje širší odbornou diskusi pro zachování všech specifik různých typů paměťových institucí.

Datová struktura taktéž prozatím neobsahuje prvky pro zápis služebních údajů - tyto budou doplněny v následující fázi.

Na počátku realizace projektu bylo řešitelským týmem vymezeno, že účelem projektu obecně nebude řešit popis informačních pramenů. Při tvorbě datové struktury se však ukázalo, že je naopak účelné navrhnout prvky alespoň v podobě minimálních požadavků na záznam dokumentu nebo jiného informačního pramenu. V této fázi jsou teda navrženy obecné prvky související s informačními prameny a v další se počítá s jejich dopracováním.

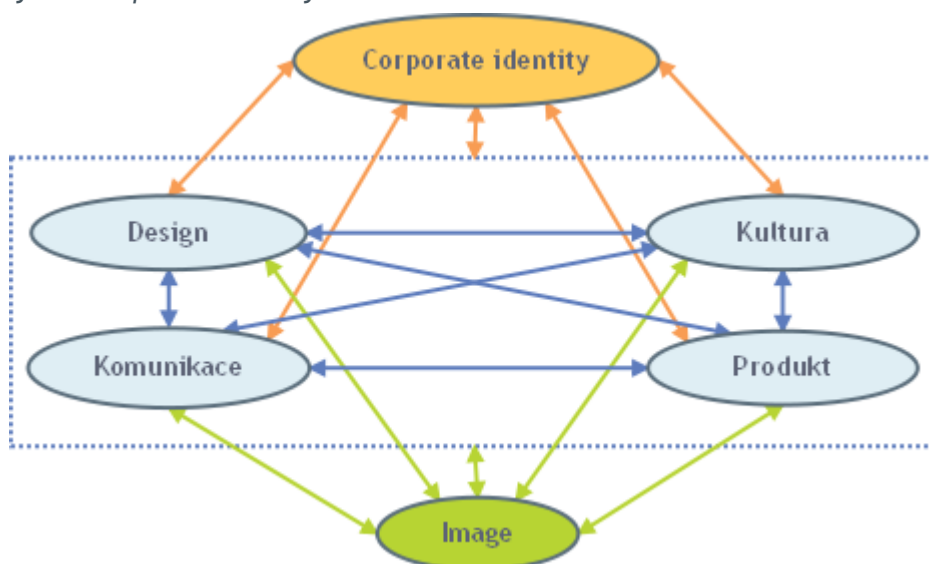
Grafická reprezentace datové struktury je na Obr. 12 na konci dokumentu. Pro více informací viz: příloha Zpráva z analýzy⁵².

B.11 Informační stránka projektu

Po dohodě s řešiteli projektu se v této fázi projektu zaměřila činnost společnosti Cosmotron na přípravu materiálů souvisejících s tzv. corporate identity projektu a současně konceptu technické podpory komunikace mezi řešiteli projektu.

Corporate identity, nebo v češtině užívaný termín firemní identita je spojován především s komerční sférou, ale stále častěji se jednotlivé složky corporate identity aplikují i v nekomerční sféře. Corporate identity není zaměřena pouze na vytvoření image firmy, ale představuje řadu nástrojů, které vytvářejí ucelený obraz o firmě. V přenesení do nekomerční sféry jde o ucelený obraz o společnosti, organizaci, instituci, nebo jako v našem případě ucelený obraz o projektu. Corporate identity zahrnuje historii, filozofii a vizi, lidi a hodnoty a tvoří součást komunikační politiky. Obecný systém corporate identity je znázorněn na Obr. 10.

Obr. 10: Systém corporate identity



V této fázi projektu se činnost společnosti Cosmotron zaměřila na dvě části: design a komunikaci.

B.11.1 Design

V první řadě byl řešen koncept loga projektu jako značky projektu a následně zabezpečeno zpracování tzv. design manuálu jako reprezentace jednotného vizuálního stylu. „Jde o soubor

⁵² příloha F.4 Zpráva z analýzy metodologických postupů, technologií sémantického webu, datových úložišť a návrh datové struktury pro základní skupinu entit.

vizuálních konstant, které jsou využívány v komunikaci uvnitř, ale především vně⁵³ prostředí projektu.

Součástí přípravy a realizace loga bylo řešení:

- koncepce loga, ochranné zóny a použití loga,
- rozměry loga, nejmenší povolený rozměr pro web a tisk,
- varianty loga,
- barevnost, černo-bíle, šedé a inverzní provedení.

Dalšími prvky design manuálu jsou:

- grafické předlohy pro textové dokumenty,
- hlavičkový papír,
- grafická předloha pro e-maily,
- grafická předloha pro prezentace,
- grafický návrh obálky,
- grafický návrh zakladače na dokumenty,
- grafický návrh pozvánky.

Kompletní design manuál je na přiloženém DVD.

B.12 Systémová analýza současných technologií a metodologických postupů

Činnost společnosti Cosmotron se v předmětném období soustředila na doplnění analýzy současných technologií a metodologických postupů o:

- teoretickou analýzu technologií sémantického webu,
- analýzu metodologických postupů s ohledem na navrženou datovou strukturu,
- testování datové struktury vzhledem k reálným datům,
- grafickou reprezentaci datové struktury.

Výsledky jsou v samostatném dokumentu v příloze⁵⁴.

Důležitou součástí technické podpory a analýzy technologií je také zvolení nástrojů pro grafickou reprezentaci datové struktury, příkladů a modelů. Cílem je nalézt a implementovat opensource nástroje potřebné k těmto účelům. V prvním roce řešení projektu jde především o výběr, instalaci a implementaci nástroje umožňujícího editovat a prezentovat navrženou ontologii pro popis jednotlivých entit (tj. datovou strukturu). Vzhledem k tomu, že návrh prvků pro popis jednotlivých typu entit musí být v souladu s ontologickým rámcem CIDOC-CRM a s

⁵³ VYSEKALOVÁ, Jitka - MIKEŠ, Jiří. Image a firemní identita. Praha : Grada Publishing, 2009. ISBN 978-80-247-2790-5, s. 40.

⁵⁴ Příloha F.4 Zpráva z analýzy metodologických postupů, technologií sémantického webu, datových úložišť a návrh datové struktury pro základní skupinu entit.

jeho rozšířením, které vzniklo na základě harmonizace tohoto modelu s konceptuálním modelem FRBR_{ER}, byl k tomuto účelu zvolen nástroj Protégé (<http://protege.stanford.edu>).

V současné době je realizována implementace, přenesení dat datové struktury a tvorba grafické reprezentace. Výsledky implementace budou prezentovány na připravovaném workshopu projektu.

B.12.1 Technické posouzení platform vhodných datových úložišť

Posouzení databázových platform pro cílové řešení se řídí následujícími metodologickými postupy:

1. teoretická analýza různých typů datových úložišť,
2. stanovení základních požadavků na vhodné datové úložiště,
3. výběr konkrétních databázových platform pro testování,
4. výběr reprezentativních modelů použitelných v konkrétní databázové platformě,
5. realizace vybraných modelů,
6. stanovení kritérií pro porovnání jednotlivých modelů,
7. stanovení hodnotící škály pro porovnání databázových platform.

B.12.2 Teoretická analýza různých typů datových úložišť

Základem testování databázových platform pro cílové řešení byla analýza různých typů datových úložišť, která je součástí Zprávy z analýzy⁵⁵.

B.12.3 Stanovení základních požadavků na vhodné datové úložiště

Jako základní požadavek na vhodné datové úložiště bylo stanoveno, že datové úložiště musí plně respektovat vlastnosti modelu CIDOC-CRM (vlastnosti jsou uvedeny ve Zprávě z analýzy).

Kromě základních požadavků, které do určité míry a určitým způsobem splňují téměř všechny typy datových úložišť, je pro CIDOC-CRM typický požadavek zabezpečující vícenásobnou dědičnost. A to nejenom z hlediska dědění vlastností jednotlivých entit, ale také z hlediska uložení instancí. Cílem je, aby instance podtřídy, která dědí vlastnosti z dvou nadtříd, byly vidět také u obou těchto nadtříd. Tuto vlastnost neposkytuje žádná ze standardních funkcí datových úložišť a proto je potřeba ji řešit na úrovni aplikace, stejně jako vyhledávání a jakékoliv další operace s takovými daty.

Mezi další požadavky pro vhodné datové úložiště byly zařazeny:

⁵⁵ Příloha F.4 Zpráva z analýzy metodologických postupů, technologií sémantického webu, datových úložišť a návrh datové struktury pro základní skupinu entit.

- podpora technologií sémantického webu,
- obecná podpora technologií pro práci s ontologiemi,
- podpora dotazovacích jazyků sémantického webu a přístupu k ontologiím.

Jelikož k základním technologickým nástrojům pro práci s ontologiemi patří především jazyk XML a RDF, rozhodli jsme se od jednotlivých modelů realizovatelných v prostředí konkrétních databázových požadovat využití právě těchto základních nástrojů.

B.12.4 Výběr konkrétních databázových platforem pro testování

V rámci testování datových úložišť pro práci se znalostními databázemi jsme v prvním kroku začali hledat řešení pro objektovou databázi Caché, která se v NK ČR dnes využívá pro projekt „Kooperativní tvorba národních autorit v heterogenním prostředí on-line“, také pro zpřístupnění a přebírání záznamu autorit z Library of Congress. Stejná databázová platforma je v současné době použita pro řešení muzejní vrstvy autorit v projektu „Národní autority v prostředí muzeí a galerií - interoperabilita s NK ČR“, kde jsou zpracovávány především personální autority. Z toho důvodu jsme na základě dohody s hlavním řešitelem projektu INTERPI začali hledat způsob uložení a možnosti práce s ontologií založené na CIDOC-CRM právě v této objektové databázi. Objektová databáze Caché je jedním z hlavních produktů společnosti Intersystems corporation.

B.12.5 Výběr reprezentativních modelů použitelných v konkrétní databázové platformě a jejich realizace

Na základě vzájemných konzultací s analytiky firmy Intersystems, jsme navrhli a začali realizovat dva na sobě nezávislé modely pro uložení a práci s ontologiemi v Caché.

B.12.6 1. model

První model je postaven na XML, kde jsou pro jednotlivé entity a vlastnosti ontologie navrženy samostatné třídy, a je potřeba definovat abstraktní třídy pro všechny vícenásobné dědičnosti, které daná ontologie definuje. Na přiloženém DVD jsou dosud realizované zdrojové kódy. Na tomto řešení se dále pracuje, aby bylo možné ho použít pro další testování. Cílem je porovnání s druhým datovým modelem a implementace dotazovacího jazyka SPARQL do Caché a to jak na úrovni syntaxe tak i na úrovni protokolu, ale také porovnání s jinými datovými úložišti.

B.12.7 2. model

Datový model založený na RDF je navržen více obecně, kde ontologie, se kterou bude datové úložiště pracovat, bude nadefinovaná v jedné základní tabulce, čímž by mělo být dané řešení vysoce flexibilní a využitelné pro použití nad libovolnou ontologií. Na druhé straně si však toto řešení vyžaduje kompletní implementaci dotazovacího jazyka. S ohledem na stručnou analýzu dotazovacích jazyků byl vybrán právě dotazovací jazyk SPARQL, jehož základní prvky se nám již podařilo implementovat také do navrženého řešení. Na přiloženém DVD jsou taktéž zdrojové kódy k tomuto datovému modelu, ve kterých se nachází také potřebné komentáře.

Obě uvedená řešení budou detailněji zdokumentována po jejich ukončení. V současné době probíhají práce na dokončení obou modelů (řešení) a vytváření podmínek pro objektivní testování, stanovení kritérií pro porovnání modelů a hodnocení databázových platforem.

**Standardizovaný archivní popis podle ISAD(G)
nebo kompatibilních národních pravidel**

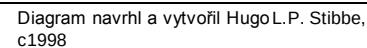
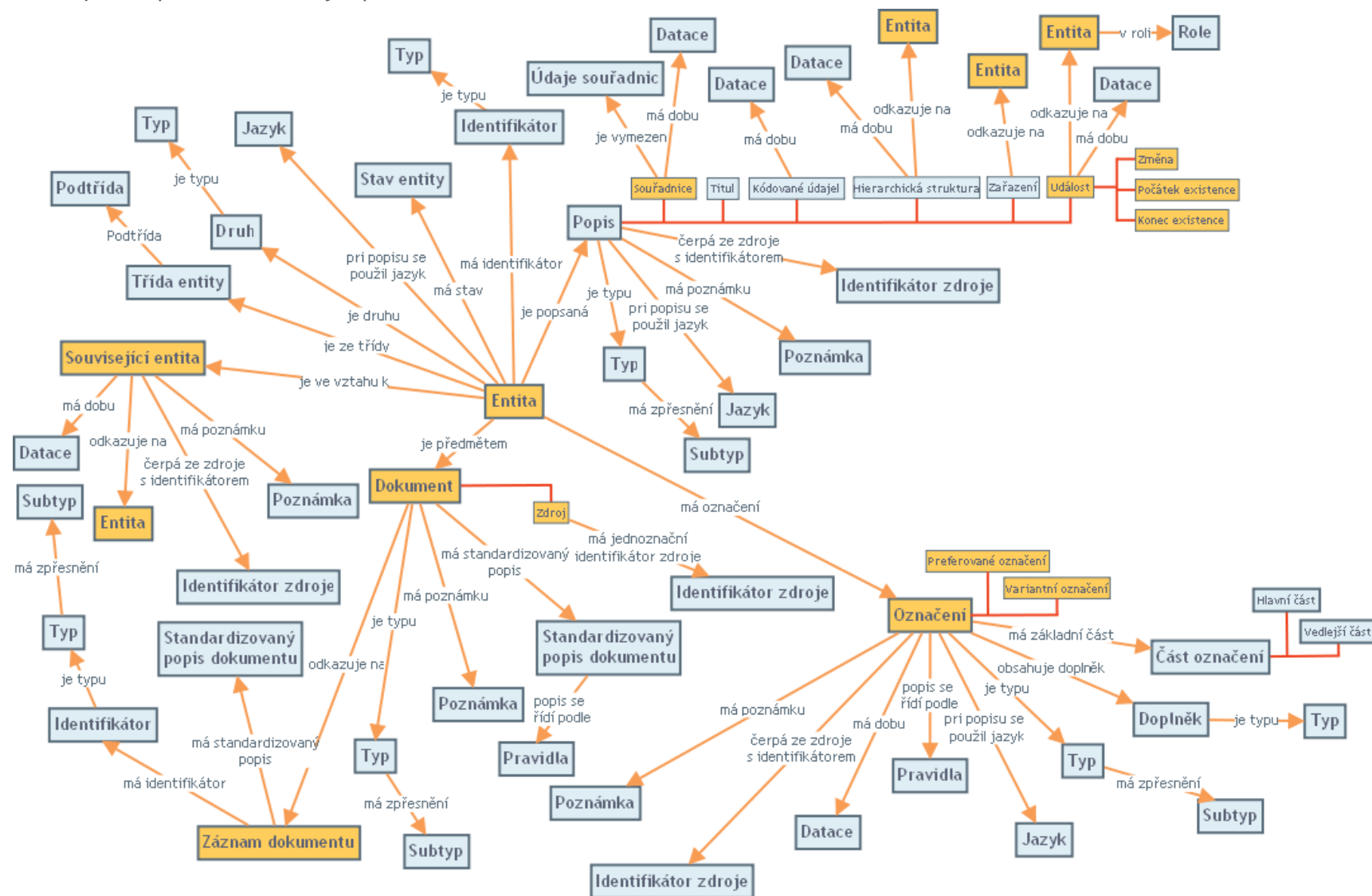


Diagram navrhl a vytvořil Hugo L.P. Stibbe,
c1998

Obr. 12: Grafická reprezentace datových prvků



Tab. 3: Příklad záznamu obce neuvedené v databázi národních autorit ČR - Čechy a Morava (první dva záznamy jsou vzor)

Region - MDT	Poznámka	Záhlaví	Okres 2008	obec 2008	část obce 2008	zdroj (Profous, Hosák)	viz 1	viz 2	viz 3	viz 4	viz 5	viz 6
(437.317)	Lídlovy Dvory, osada obce Kašperské Hory v okrese Klatovy, 2 km jižně od Kašperských Hor.	Lídlovy Dvory (Kašperské Hory, Česko)	Klatovy	Kašperské Hory		Profous	Liedlhöfen%1840	Dvory Lidlovy%1854	Ober-Liedlhöfen%1840	Unter-Liedlhöfen%1840	Lidlhoff%1654	Dwory Lidlowske%1584
(437.318)	Doubí, část obce Třebeň, okres Cheb.	Doubí (Třebeň, Česko)	Cheb	Třebeň		Profous	Doubí%1923	Aag%1923	Aag%1847	Trag%1847	Aag%1785	Drag%1785
(437.324)	Dolní Bobrová, část obce Bobrová, okres Žďár nad Sázavou	Dolní Bobrová (Bobrová, Česko)	Žďár nad Sázavou	Bobrová	Dolní Bobrová	Hosák	Unter Bobrau%1872	Dolní Bobrová%1872	Dolnj Bobrowa%1846	Unter Bobrowa%1751	Unter Obrowa%1718	**v Dolny Bobrowy%1528
(437.324)	Horní Bobrová, část obce Bobrová, okres Žďár nad Sázavou, 10 km jihovýchodně od Žďáru nad Sázavou.	Horní Bobrová	Žďár nad Sázavou	Bobrová	Horní Bobrová	Hosák	Ober Bobrau%1872	Horní Bobrová%1872	Hornj Bobrowa%1846	Ober Bobrau%1846	Ober Bobrowa%1751	Obrowa%1720

Tab. 4: Příklad záznamu obce - Slovensko

Záhlaví	část obce	obec	kód obce	Okres	Kraj	Maďarsky	Německy	viz 1	viz 2	viz 3	viz 4
Ábelová (Slovensko)		Ábelová	511226	Lučenec	Banskobystrický	Ábelfalva		Abelfeuld%1275	Abellehota%1460	Ábelová%1808	
Urpín (Banská Bystrica, Slovensko)	Urpín	Banská Bystrica	508438	Banská Bystrica	Banskobystrický			Urpín%1520			
Kráľová (Banská Bystrica, Slovensko)	Kráľová	Banská Bystrica	508438	Banská Bystrica	Banskobystrický	Garamkirályfalva		**Villa Reginalis Querali%1264	Khunigldorf%1402	Khunigsdorf%1503	Kyralowa%1522
Abovce (Slovensko)		Abovce	557757	Rimavská Sobota	Banskobystrický	Abafalu		Abafalwa%1339	alia Hanua%1382	Abovce%1920	
Jánošovce (Abovce, Slovensko)		Abovce	557757	Rimavská Sobota	Banskobystrický			Janosfalua%1427			
Abrahám (Slovensko)		Abrahám	503673	Galanta	Trnavský	Abrahám		Scent Abraam%1266	Sz. Abraham%1773	Abraham%17	
Abrahámovce (Slovensko)		Abrahámovce	519014	Bardejov	Prešovský	Abrahámfalva, Ábrahámfalva		Abranfalua%1427	Abran%1476	Abrahamovce%1786	Abranowece%1808
Bencova (Abrahámovce, Slovensko)	Bencova	Abrahámovce	519014	Bardejov	Prešovský						
Dolina (Abrahámovce, Slovensko)	Dolina	Abrahámovce	519014	Bardejov	Prešovský						

Tab. 5: Příklad tabulky s informacemi o entitách typu války, bitvy

Záhlaví	MDT	Odkaz	BT NT RT	Angl.	Pozn.
burgundské války, 1474-1477	355.483(434.5+494+444.1/.4)"1474/1477"		NT Neuss, obléhání, 1474 NT Héricourt, bitva, 1474 NT Planta, bitva, 1475 NT Grandson, bitva, 1476 NT Murten, bitva, 1476 NT Nancy, bitva, 1477	Burgundian Wars, 1474-1477	nová
Neuss, obléhání, 1474	355.483:355.44 (430.119)"1474"	obléhání města Neuss, 1474	BT burgundské války, 1474-1477	Neuss, siege of, 1474	nová
Héricourt, bitva, 1474	355.483(444.1/.4)"1474"	bitva u Héricourtu, 1474	BT burgundské války, 1474-1477	Héricourt, battle of, 1474	nová
Planta, bitva, 1475	355.483(494)"1475"	bitva u Planta, 1475	BT burgundské války, 1474-1477	Planta, battle on the, 1475	nová
Grandson, bitva, 1476	355.483(494)"1476"	bitva u Grandsonu, 1476	BT burgundské války, 1474-1477	Grandson, battle of, 1476	nová
Murten, bitva, 1476	355.483(494)"1476"	bitva u Murtenu, 1476 Morat, bitva, 1476	BT burgundské války, 1474-1477	Murten, battle of, 1476	nová
Nancy, bitva, 1477	355.483(434.5)" 1477"	bitva u Nancy, 1477	BT burgundské války, 1474-1477	Nancy, battle of, 1477	nové

C Návrhová část

Obsahuje cíle projektu INTERPI pro rok 2012; aktivity v tomto roce spadají do kategorie základního výzkumu a vývoje a jejich konkrétním cílem je:

- definování všech typů entit (tříd),
 - typologie entit,
 - specifikace vlastností a vztahů entit nejdůležitějších z hlediska sémantické interoperability v paměťových institucích (personálií, korporací a GEO),
 - specifikace požadavků na datovou strukturu,
- analýza, ověřování a úprava dat,
 - jednoznačná identifikace,
 - geografické entity,
 - entity typu války, bitvy,
 - personální entity,
 - entity rodina/rod,
 - korporativní entity,
- potenciální aplikace klasifikačního systému MDT jako nástroje propojení různých systémů,
- analýza aktuálních potřeb paměťových institucí,
- analýza metodologických postupů,
- aktivity v oblasti informačních technologií,
 - výběr databázové platformy, technického a datového modelu,
 - dokončení návrhu datové struktury a její kritické hodnocení,
 - zpracování funkčního konceptu softwarového řešení pro základní skupinu entit,
 - technická podpora komunikace a prezentace výsledků projektu.

Výsledkem bude množina společných entit základní skupiny včetně jejich specifikace a vzájemných vztahů, budou definovány požadavky na SW pro základní skupinu entit a základní metodika, která bude představena na semináři uspořádaném v závěru dílčí etapy.

C.1 Definování všech typů entit (tříd)

C.1.1 Typologie entit

Analytické studie zástupců paměťových institucí jednoznačně prokázaly, že koncepce jednotlivých typů entit je v jednotlivých paměťových institucích různá: muzea a oblastní galerie považují za geografickou entitu různé stavby či objekty a jejich části (hrady, zámky, 1.-nádvoří, dvůr pivovaru atd.), kdežto knihovnická obec považuje tyto „entity“ za komplexní informační objekty skládající se z jednotlivých entit, např. korporace a dílčího objektu, tedy kombinace korporativní a tematické entity.

Tuto otázku lze řešit v první fázi definováním obecného modelu entity a příslušných vazeb a vztahů, ve druhé fázi se pokusit definovat jednotlivé typy entit a jejich subkategorie, což bude cílem této fáze projektu INTERPI.

C.1.2 Specifikace vlastností a vztahů entit

V roce 2012 bude nutné ověřit navrženou strukturu a specifikaci základní skupiny entit na konkrétních souborech příslušných dat z databází jednotlivých paměťových institucí.

C.2 Analýza, ověřování a úprava dat

C.2.1 Jednoznačná identifikace

V této „přechodné,, fázi řešení projektu se zaměříme na tvorbu nástrojů pro postupné doplňování informací do stávajících AUT záznamů a propojování těchto informací způsobem adekvátním pro vyhledávání v prostředí sémantického webu. Zajistíme tím lepší dostupnost dat a jejich propojení (např. prostřednictvím nevýznamového identifikátoru - ID) a podpoříme tím přímou integraci stávajících údajů do znalostní databáze INTERPI. Naší snahou je jednoznačná identifikace všech relevantních údajů obsažených v AUT záznamech, která podpoří reprezentaci těchto údajů, jejich propojování a jejich přímou integraci v hybridním prostředí sémantického webu.

C.2.2 Geografické entity

Budou pokračovat práce na přípravných souborech dat, jejich ověřování

C.2.3 Entity typu války, bitvy

Bude nutné vyřešit zařazení těchto specifických entit v rámci typologie entit

C.2.4 Personální entity

V oblasti personálních entit bude nutné se zaměřit a vyřešit entitu „anonymní tvůrce/osoba“.

C.2.5 Entity rodina/rod

V této oblasti je potřebné vyřešit problematiku identifikace a potřebných informací týkajících se rodin/rodů, zvláště z hlediska archivů

C.2.6 Korporativní entity

Prvořadým úkolem pro tento rok řešení projektu je v oblasti korporativních entit sladit navrženou obecnou datovou strukturu se strukturou odpovídající standardu ISAAR(CPF)

C.3 Potenciální aplikace klasifikačního systému MDT

V roce 2012 je také naším cílem zabývat se teoretickým průzkumem a potřebnou úpravu klasifikačního systému MDT jako potenciálního propojovacího jazyka při řešení terminologické různorodosti termínů aplikovaných v databázích paměťových institucí.

C.4 Analýza aktuálních potřeb paměťových institucí

Nadále je potřeba sledovat, analyzovat konkrétní požadavky této komunity uživatelů, aby přijatá řešení v projektu INTERPI - pokud možno - v plné míře byla schopna uspokojit jejich reálné potřeby.

C.5 Analýza metodologických postupů

Je žádoucí v tomto roce základního výzkumu a vývoje dokončit komplexní metodologickou analýzu standardů a jejich transformaci do praktičtější (srozumitelnější) roviny, zvláště v případě standardů, např. CIDOC CRM, které jsou definovány příliš obecně. Dále je nutné pro potřeby budoucích školení připravit alespoň pracovní překlady vybraných pasáží těchto standardů.

C.6 Aktivity v oblasti informačních technologií

C.6.1 Výběr databázové platformy, technického a datového modelu

Ve druhém roce 1. etapy projektu bude probíhat další testování databázových platform pro výběr vhodného datového úložiště tak, aby bylo možné v závěru 1. etapy vybrat vyhovující řešení.

Z hlediska testování databázových platform je potřebné podle navrženého postupu především:

- stanovit kritéria pro porovnání jednotlivých modelů,
- stanovit hodnotící škálu pro porovnání databázových platform.

Probíhat bude testování jednak dvou realizovaných modelů na databázové platformě Caché, které jsou popsány výše a současně výběr jiných databázových platform, které splňují základní kritéria pro vhodnou databázovou platformu.

C.6.2 Dokončení návrhu datové struktury a její kritické hodnocení

Pro další postup v přípravě softwarového řešení, které je hlavním cílem navazující 2. etapy projektu je nevyhnutné dokončení návrhu datové struktury pro základní skupinu entit a jeho kritické hodnocení po věcné stránce, ale také po stránce technologické realizovatelnosti.

Pro dokončení návrhu datové struktury je potřebné:

1. doplnit prvky pro zápis informací souvisejících s umělými výtvary nebo výkony pro zaznamenání proslavení osob a korporací,
2. doplnit prvky pro zápis informací o symbolech, znacích a objektech nehmotného kulturního dědictví souvisejících s entitou,
3. doplnit prvky pro zápis služebních údajů,

4. rozšířit prvky pro popis informačních pramenů, kterých zpracování se ukázalo jako nevyhnutelné pro vyhovující komplexní řešení problematiky.

Kritické hodnocení je nevyhnutelné, protože datová struktura musí splňovat požadavky paměťových institucí pro základní skupinu entit. Současně musí být kompatibilní s CIDOC-CRM a musí respektovat v současnosti využívané standardy a také jejich očekávané změny.

C.6.3 Zpracování funkčního konceptu softwarového řešení pro základní skupinu entit

Na základě formulace požadavků odborných komunit a s ohledem na datovou strukturu pro základní skupinu entit je potřebné v následujícím období 1. etapy projektu zpracovat funkční koncept softwarového řešení pro základní skupinu entit. Pro dosažení tohoto cíle je potřeba:

1. formulovat obecné požadavky na softwarové řešení - tj. hlavní funkce, které bude poskytovat,
2. vymezit standardy, které softwarové řešení musí respektovat,
3. definovat požadavky na softwarové řešení vzhledem na vybranou databázovou platformu, datový a technický model.

Součástí je také ověřování komunikačních prostředků pro integraci s existujícími systémy, které v této fázi projektu představuje především analýzu technologických řešení tak, aby bylo v dalších etapách možné nabídnout existujícím systémům vhodné řešení pro spolupráci s projektem.

C.6.4 Technická podpora komunikace a prezentace výsledků projektu

V návaznosti na předchozí aktivity bude cílem další části 1. etapy realizace technického řešení pro podporu komunikace a prezentace výsledků projektu. Výsledkem bude softwarové řešení pro evidenci informací souvisejících s projektem (úkolů, kontaktních údajů řešitelů a spolupracovníků, informací o akcích, evidence dokumentů), které sice není stanoveno v cílech projektu, ale poskytuje přiměřené zázemí na jejich plnění.

V rámci technické podpory prezentace výsledků projektu je také cílem vytvoření vlastní informační stránky projektu.

Seznam bibliografických odkazů

ANDREJČÍKOVÁ, Naděžda - ŠUBOVÁ, Jana. Nevyhnutné podmienky pri tvorbe digitálnych knižníc. In *ITlib*. Roč. 14, 2010, č. 2.

ANDREJČÍKOVÁ, Nadežda. *Autority ako nástroj interoperability v pamäťových a fondových inštitúciách* [Dizertační práce]. Univerzita Komenského v Bratislave. Filozofická fakulta; Katedra knižničnej a informačnej vedy. Školiteľ: Doc. PhDr. Jaroslav Šušol, PhD. Bratislava : FFUK, 2010.

ANGLES, R. - GUTIERREZ, C. Survey of graph database models. In *ACM Computing Surveys* [online]. 2008 [cit. 2011-04-14]. Dostupné z www: <http://swp.dcc.uchile.cl/TR/2005/TR_DCC-2005-010.pdf>.

Archiv. In *Wikipedie* [online]. 2010-08-04 [cit. 2010-08-09]. Dostupné z www: <<http://cs.wikipedia.org/wiki/Archiv>>.

ATKINSON, M., et al. The object-oriented database system manifesto. In *Proceedings of the First International Conference on Deductive and Object-Oriented Databases*. 1989.

BALÍKOVÁ, Marie. Multilingual Subject Access to Catalogues of National Libraries (MSAC) : Czech Republic's collaboration with Slovakia, Slovenia, Croatia, Macedonia, Lithuania and Latvia. In *IFLA. World Library and Information Congress : 71th IFLA General Conference and Council* [online]. 2005 [cit. 2010-06-24]. Dostupné z www: <<http://archive.ifla.org/IV/ifla71/papers/044e-Balikova.pdf>>.

CIDOC. *International guidelines for museum object information : the CIDOC information categories* [online]. June 1995 [cit. 2011-02-01]. Dostupné z www: <<http://www.cidoc-crm.org/docs/guide.htm>>.

CIGÁNEK, David. *Geografické názvy v prírodných vedách : pracovný materiál pro řešitelskou skupinu projektu INTERPI*. 2011. Interní zpráva k projektu INTERPI. Dostupné z www: <<http://authority.nkp.cz/interpi>>.

ČABRUNOVÁ, Anna. Informačný proces. In KATUŠČÁK, Dušan - MATTHAEIDESOVÁ, Marta - NOVÁKOVÁ, Marta. *Informačná výchova : terminologický a výkladový slovník*. Bratislava : Slovenské pedagogické nakladateľstvo, 1998, s. 143 - 149.

DESIRE Project : development of a european service for information on research and education [online]. 2000-08-07 [cit. 2010-06-24]. Dostupné z www: <<http://www.desire.org/>>.

EAC-CPF : Encoded archival context, corporate bodies, persons, and families [online]. Berlin : Staatsbibliothek [cit. 2010-06-24]. Dostupné z www: <<http://eac.staatsbibliothek-berlin.de/>>.

eGovernemnt Resource Centre. Taxonomies & Thesauri [online]. 2009-11-24 [cit. 2010-06-24]. Dostupné z www: <<http://www.egov.vic.gov.au/website-practice/taxonomies-thesauri/thesauri-archive.html>>.

EPOCH : European Network of Excellence in Open Cultural Heritage [online]. [cit. 2010-06-24]. Dostupné z www: <<http://www.epoch-net.org/>>.

Galerie umění. In *Wikipedie* [online]. 2010-07-18 [cit. 2010-08-10]. Dostupné z www: <http://cs.wikipedia.org/wiki/Galerie_%28expozice%29>.

HASLHOFER, B., et al. Europeana RDF Store Report. In *Library* [online]. 2011 [cit. 2011-04-28]. Dostupné z www: <http://www.europeanaconnect.eu/documents/europeana_ts_report.pdf>.

HERTEL, A. - BROEKSTRA, J. RDF Storage and Retrieval Systems. In *Handbook on Ontologies* [online]. 2009 [cit. 2011-04-14]. Dostupné z www: <<http://ki.informatik.uni-mannheim.de/fileadmin/publication/Hertel08RDFStorage.pdf>>.

HILT : high-level thesaurus [online]. 2009-06-22 [cit. 2010-06-24]. Dostupné z www: <<http://hilt.cdlr.strath.ac.uk/>>.

HRAZDIL, Aleš. *Geografické termíny Moravskoslezského kraje - specifikace požadavků uživatelské komunity jednotlivých paměťových institucí daného regionu*. 2011. Interní zpráva k projektu INTERPI. Dostupné z www: <<http://authority.nkp.cz/interpi>>.

CHIOS : Cultural Heritage Interchange Ontology Standarisation. In *CORDIS : Information & Communication Technologies* [online]. [cit. 2010-06-24]. Dostupné z www: <http://cordis.europa.eu/fetch?CALLER = PROJ_ICT&ACTION = D&DOC = 2&CAT = PROJ&QUERY = 01286358ca51:f323:7897c63e&RCN = 58093>.

ICOM. *Profesní kodex muzeí* [online]. Aktualizováno v lednu 2003 [cit. 2010-08-01]. Dostupné z www: <<http://www.cz-icom.cz/doc0008.html>>.

ICOM. *The CIDOC Conceptual Reference Model* [online]. [cit. 2010-06-24]. Dostupné z www: <<http://cidoc.ics.forth.gr/>>.

IFLA. *Guidelines for Authority Records and References* [online]. 2nd ed. München : K. G. Saur, 2001. Dostupné z www: <<http://archive.ifla.org/VII/s13/garr/garr.pdf>>.

International council on archives. *ISAAR (CPF) : International standard archival authority record for corporate bodies, persons and families* [online]. 2nd ed. ICA, 2004 [cit. 2009-07-30]. Dostupné z www: <<http://www.ica.org/sites/default/files/ISAAR2EN.pdf>>. V češtině: International council on archives. *ISAAR (CPF) : Mezinárodní standard pro archivní autoritní záznamy* [online]. 2. vyd. Praha, 2009 [cit. 2009-07-30]. Dostupné z www: <<http://www.ica.org/en/node/39464>>.

ISAAC, Antoine - CHAMBERS, Sally. *Prototype integrating MACS initial data and new alignment into TEL framework* [online]. 4 January 2010 [cit. 2010-06-24]. Dostupné z www: <http://www.theeuropeanlibrary.org/portal/organisation/cooperation/telplus/documents/TEL_plus_D3.4_04012010.pdf>.

KITJONGTHAWONKUL, Somkiat - KHOSLA, Rajiv. Modeling information systems using objects, agents, and task-based problem solving adapters. In *Proc. 10th Australasian conference on information systems* [online]. 1999 [cit. 2011-02-02], s. 474-481. Dostupné z www:

<<http://ebookbrowse.com/ts-6-modeling-information-systems-using-objects-agents-task-based-problem-solving-adapters-pdf-d40286619>>.

KLIMEŠ, René. *Analýza selekčních prvků Sbírek RML online - srovnání s Národními autoritami*. 2011. Interní zpráva k projektu INTERPI. Dostupné z www: <<http://authority.nkp.cz/interpi>>.

KUČA, Karel. *Význam Databáze sídelních lokalit Čech, Moravy a Slezska CZ_RETRO pro projekt INTERPI : aplikace kódu CZ v souboru Národních geografických autorit*. 2011. Interní zpráva k projektu INTERPI. Dostupné z www: <<http://authority.nkp.cz/interpi>>.

LEAVITT, N. Will NoSQL Databases Live Up to Their Promise?. In *Computer* [online]. 2010, vol. 43 [cit. 2011-04-14], s. 12-14. Dostupné z www: <<http://www.leavcom.com/pdf/NoSQL.pdf>>.

LENHART, Zdeněk. *Národní autority v prostředí muzeí a galerií - interoperabilita s NK ČR* [online]. Prezentace na: Seminář o spolupráci knihoven a paměťových institucí JMK, 27.5.2008 [cit. 2010-0813]. Dostupné z www: <<http://www.mzk.cz/region/Autority.pdf>>.

MADAN, Alok. Object-oriented paradigm in programming for computer-aided analysis of structures. In *Journal of computing in civil engineering*. July 2004, vol. 18, issue 3, s. 226-236.

MARVANOVÁ, Eva. Paměťová instituce. In *KTD : Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha : Národní knihovna České republiky, 2003 [cit. 2010-08-01]. Dostupné z www: <<http://aleph.nkp.cz/cze/ktd>>.

MITCHELL, Joan S. - ZENG, Marcia Lei - ŽUMER, Maja. Extending models for controlled vocabularies to classification systems : modelling DDC with FRSAD. In *Classification and ontology: formal approaches and access to knowledge: the International UDC Seminar, 19-20 September 2011, The Hague, The Netherlands*. Dostupné z www: <http://www.udcs.com/seminar/2011/media/slides/UDCSeminar2011_Mitchell_Zeng_Zumer.pdf>.

NEUMANN, Marco. SeMuSe the future of semantic museum data. In *Semantic universe* [online]. 2009-03-02 [cit. 2010-06-25]. Dostupné z www: <<http://www.semanticuniverse.com/articles-semuse-future-semantic-museum-data.html>>.

NISO. *Metasearch Initiative* [online]. [cit. 2010-08-13]. Dostupné na www: <<http://www.niso.org/workrooms/mi>>.

OCHRE : *Online Cultural Heritage Research Environment* [online]. [cit. 2010-06-24]. Dostupné z www: <http://ochre.lib.uchicago.edu/index_files/Page845.htm>.

PAETAU, Patrik. *On the benefits and problems of the object-oriented paradigm including a finnish study* [online]. Helsingfors, 2005 [cit. 2011-02-02]. ISBN 951-555-894-8. 294 s. Dostupné z www: <<http://dhanken.shh.fi/dspace/bitstream/10227/123/2/151-951-555-894-8.pdf>>.

PATEL, Manjula et al. *Semantic interoperability in Digital Library Systems* [online]. DELOS, March 2005 [cit. 2010-06-24]. Dostupné z www: <<http://www.ukoln.ac.uk/ukoln/staff/t.koch/publ/SI-in-DLs.doc>>.

PODOLNÍKOVÁ, Jarmila - MAZAČOVÁ, Vlad'ka. *Lokality v systému Demus*. 2011. Interní zpráva k projektu INTERPI. Dostupné z www: <<http://authority.nkp.cz/interpi>>.

REITZ, Joan M. *ODLIS : Online dictionary for library and information science* [online]. Last updated 2007-11-19 [cit. 2009-07-29]. Dostupné z www: <<http://lu.com/odlis/index.cfm>>.

SCULPTEUR : Semantic and content-based multimedia exploitation for European benefit [online]. [cit. 2010-06-24]. Dostupné z www: <<http://www.sculpteurweb.org/>>.

SODOMKOVÁ, Jana. Knihovna. In *KTD : Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha : Národní knihovna České republiky, 2003 [cit. 2010-08-01]. Dostupné na www: <<http://aleph.nkp.cz/cze/ktd>>.

STEVENS, P - POOLEY, R. Using UML : software engineering with objects and components. Harlow : Addison-Wesley, 2000, s. 10. Zdroj: PAETAU, Patrik. *On the benefits and problems of the object-oriented paradigm including a finnish study* [online]. Helsingfors, 2005 [cit. 2011-02-02]. ISBN 951-555-894-8, s. 24. Dostupné z www: <<http://dhanken.shh.fi/dspace/bitstream/10227/123/2/151-951-555-894-8.pdf>>.

SVÁTEK, Vojtěch - LABSKÝ, Martin. Objektové modely a znalostní ontologie - podobnosti a rozdíly. In *Objekty 2003* [online]. 2003 [cit. 2011-10-20]. Dostupné z www: <<http://nb.vse.cz/~svatek/obj03fi.pdf>>.

SYROVÝ, Jiří. *Prostorová identifikace v projektu INTERPI*. 2011. Interní zpráva k projektu INTERPI. Dostupné z www: <<http://authority.nkp.cz/interpi>>.

The BRICKS Community : the BRICKS Project [online]. [cit. 2010-06-24]. Dostupné z www: <<http://www.brickscmmunity.org/prj/>>.

VIAF : virtual international authority file [online]. [cit. 2010-06-24]. Dostupné z www: <<http://www.viaf.org/>>.

VODIČKOVÁ, Hana. Autorita. In *KTD : Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha : Národní knihovna České republiky, 2003 [cit. 2009-07-30]. Dostupné z www: <<http://sigma.nkp.cz/cze/ktd>>.

VODIČKOVÁ, Hana. Bibliografický záznam. In *KTD : Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha : Národní knihovna České republiky, 2003 [cit. 2010-08-01]. Dostupné na www: <<http://aleph.nkp.cz/cze/ktd>>.

VODIČKOVÁ, Hana. Geografická autorita. In *KTD : Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha : Národní knihovna České republiky, 2003 [cit. 2009-09-25]. Dostupné z www: <<http://sigma.nkp.cz/cze/ktd>>.

VODIČKOVÁ, Hana. Korporativní autorita. In *KTD : Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha : Národní knihovna České republiky, 2003 [cit. 2009-09-25]. Dostupné z www: <<http://sigma.nkp.cz/cze/ktd>>.

VODIČKOVÁ, Hana. Personální záhlaví. In *KTD : Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha : Národní knihovna České republiky, 2003 [cit. 2009-10-29]. Dostupné z www: <<http://sigma.nkp.cz/cze/ktd>>.

VODIČKOVÁ, Hana. Soubor autorit. In *KTD : Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online]. Praha : Národní knihovna České republiky, 2003 [cit. 2009-07-30]. Dostupné z www: <<http://sigma.nkp.cz/cze/ktd>>.

Slovník zkratk

AACR2	Anglo-American Cataloguing Rules [Anglo-americká katalogizační pravidla] mezinárodně využívaná základní příručka pro bibliografický a katalogizační popis, druhé vydání
CCO	Cataloging Cultural Object [Katalogizace kulturních objektů] mezinárodní pravidla pro popis a katalogizaci kulturních objektů
CDWA	Categories for the Description of Works of Art [Kategorie pro popis uměleckých děl] rámec pro popis a zpřístupnění informací o uměleckých dílech, architektuře, materiální kultuře a sbírek děl a souvisejících obrazových dokumentech
CIDOC	International Committee for Documentation [Mezinárodní výbor pro dokumentaci] organizace při ICOM (Mezinárodní radě muzeí) zabývající se muzejní dokumentací
CIDOC-CRM	CIDOC Conceptual Reference Model [Konceptuální referenční model vytvořen CIDOC] poskytuje definice a formální strukturu pro popis pomocí pojmů a vztahů využívaná v dokumentaci kulturního dědictví
CITeM	Centrum pro informační technologie v muzejnictví založeno při Moravském zemském muzeu, metodické centrum pro informační technologie v muzejnictví
FRAD	Functional Requirements for Authority Data [Funkční požadavky na data autorit] dokument určující požadavky na data shromažďovány o autoritách
FRBR (FRBR_{ER})	Functional Requirements for Bibliographic Records [Funkční požadavky na bibliografické záznamy] dokument určující požadavky, které musí splňovat bibliografické záznamy
FRBR_{OO}	FRBR - object oriented [FRBR - objektově orientované] úprava FRBR a jejich přizpůsobení objektově orientovanému modelování
FRSAD	Functional Requirements for Subject Authority Data [Funkční požadavky na data předmětových autorit] dokument vycházející z FRAD, ale zaměřen na předmětové autority
GARR	Guidelines for authority records and references [Pravidla pro záznamy autorit a odkazy] dokument určující požadavky pro zobrazení informací o autorizovaných záhlavích a odkazech
ISO	International Organization for Standardization [Mezinárodní organizace pro normalizaci] mezinárodní síť spolupracujících národních normalizačních organizací, zabývá se tvorbou mezinárodních norem
MARC	MAchine Readable Cataloging [strojem čitelná katalogizace] typ katalogizačního formátu
MARC21	MAchine Readable Cataloging pro 21 století katalogizační formát podporovaný Library of Congress, který vznikl harmonizací formátů USMARC a CAN/MARC do jednoho dokumentu
NA ČR	Národní archiv ČR správní a ústřední archiv státu
NK ČR	Národní knihovna ČR centrum systému knihoven v ČR, vykonává koordinační, odborné, informační, vzdělávací, analytické, výzkumné, standardizační, metodické a poradenské činnosti
RDA	Resource Description and Access [Popis zdrojů a přístup k nim] aktualizace AACR2 - tedy revidovaná pravidla pro popis dokumentů
UNIMARC	Universal Machine-readable cataloging [univerzální stroj čitelná katalogizace] katalogizační formát vytvořený s cílem sjednocení různých formátů do jednoho
XML	eXtensible Markup Language [rozšiřovatelný značkovací jazyk] technologie, standard zápisu textu pro webové publikování

D Použití účelové podpory v členění dle rozpočtu projektu

E Resumé

První rok 1. etapy projektu je zaměřen na základní výzkum, proto se převážná část aktivit opírala o základní metody vědecké práce - tj. analýzu a syntézu. Analýza probíhala hned v několika oblastech. Jednak se analyzovali potřeby paměťových institucí v oblasti základní skupiny entit, také se prováděla terminologická analýza a analýza standardů a technologií relevantních pro projekt. Současně se ověřovali a doplňovali autoritní termíny v rámci stávající databáze národních autorit a pozornost byla věnována především geografickým autoritám. Potřebné bylo také analyzovat systém MDT jako potenciální nástroj propojení.

Za nejdůležitější výsledky prvního roku 1. etapy lze považovat specifikaci požadavků paměťových institucí v oblasti základních entit, tj. personálií, korporací a geografických entit a také návrh datové struktury pro základní skupinu entit. Oba výsledky jsou jedinečné v tom, že spojují hlediska a požadavky více typů paměťových institucí.

Po technické stránce začalo testování databázových platforem vhodných pro datové úložiště. V prvním roce první etapy byly vytvořeny dva modely v databázovém prostředí Caché.

Uplatněný výsledek - druh výsledku podle RIV: C (kapitola v odborné knize):

BALÍKOVÁ, Marie. Focusing on User Needs: New Ways of Subject Access in Czechia. In *Subject Access : Preparing for the Future*. Berlin ; Munich : De Gruyter Saur, 2011, s. 7-24.

F Přílohy

- F.1 Použití účelové podpory v členění dle rozpočtu projektu - komentář
- F.2 Dodatky, rozhodnutí
- F.3 Cestovní zprávy
- F.4 Zpráva z analýzy metodologických postupů, technologií sémantického webu, datových úložišť a návrh datové struktury pro základní skupinu entit

F.1 Použití účelové podpory v členění dle rozpočtu projektu - komentář

F.2 Dodatky, rozhodnutí

F.3 Cestovní zprávy

F.4 Zpráva z analýzy metodologických postupů, technologií sémantického webu, datových úložišť a návrh datové struktury pro základní skupinu entit