Facilitating access to cultural heritage content in Czechia: National Authority Files and INTERMI project

Marie Balíková
National Library of the Czech Republic, Prague, Czechia
E-mail address: Marie.Balikova@nkp.cz

Miroslav Kunt
National Archives, Prague, Czechia
 E-mail address: Miroslav.Kunt@nacr.cz

Jana Šubová
Cosmotron Bohemia, s. r. o., Hodonín, Czechia
 E-mail address: subova@cosmotron.cz

Nadežda Andrejčíková
Cosmotron Bohemia, s. r. o., Hodonín, Czechia
 E-mail address: andrejcikova@cosmotron.cz

Abstract:

The paper informs about the project Interoperability in memory institutions (INTERMI) in Czechia realized by the National Library of the Czech Republic and National Archive within Applied research and development of national and cultural identity Programme"(NAKI) program funded by the Ministry of Culture of the Czech Republic. The purpose of the project is to create scientific and technological infrastructure to support the processing, sharing and use/reuse of the cultural heritage content in the form of metadata about information objects preserved in memory institutions, it means to create a conceptual model of knowledge which meets the criteria and user's needs of all memory institutions in Czechia. The project is based on a new paradigm, which is focused on processing of entities enlarged by necessary semantic information and on expression of complex relationships among them. The role of National Authority files in INTERMI knowledge base will be evaluated.

Keywords: memory institutions, national authorities, culture heritage, knowledge base, conceptual models, entities, event-centric approach

Introduction and context

Libraries, archives, galleries, and museums as memory institutions collect, organize and make available information objects/resources of national cultural, industrial, scientific, and natural heritage domain. Cultural heritage domain includes documentary and archival heritage as well.

Now, in the period of the mass digitisation, the main goal of memory institutions activities is to support effective selection, aggregation and easy use/reuse of the digital content.

In INTERMI project, this common objective depends on the specification of user community needs, on the quality of metadata related to the information objects (unambiguous identification of objects on the metadata level included), on application of object-oriented programming principles and event-centric approach, on application of semantic interoperability technologies, it means on conceptual model and data structure applied, on formal representation tools suitable for Semantic Web environment (e.g. XML, RDF, OWL), and on application of Linked Open Data principles.

Specification of user's community needs and their expectations from INTERMI project

Users of memory institutions in Czechia (general public, cataloguers, indexers, information specialists, curators, archivists, and professionals from all subject domains) support the creation and development of the infrastructure for cultural heritage database; they prefer to make cultural heritage content available on the Web in both human readable and machine understandable form, it means to share cultural heritage content according to Semantic Web and Linked Open Data principles. They are aware that it supposes development and application of standards, rules and metadata schemas of high quality and to be appropriate for the Semantic Web. They need more information about entities in current authority records and they agree that it is necessary to move from authorities to entities of value on the web, from terms to concepts.

Information objects and their identification on metadata level

In INTERMI project, the information objects represent the real-world entities, e.g. persons, institutions, three-dimensional objects (i.e. artistic and technical objects, objects of inanimate nature), activities, events and performances, artistic and other achievements, places, etc. which represent topics of their collection items and are subject to their activities. At the time being, the identification of the information objects on the metadata level is based on both name and subject authorities.

The role of National Authority files in making available digital content in Czechia

National Authority Files contain Personal name file (639 472 items), Corporate and Meeting name file (140 002), Uniform title (2 454), Topical term file (36 954), Geographical name file (27 170) and Genre/form name file (1 783).

National Authority files were originally developed with the aim to support the share cataloguing and the indexing of library materials. Name authorities were created according to AACR2[1] rules while subject authorities were originally created according to LCSH rules.

---

[1] Since May, 2015 RDA Rules have been introduced.

Later, with advent of WWW, some new principles were introduced: post-coordination, the authorisation of isolated lexical units (not entire subject headings strings), the limited application of subject heading strings in bibliographic records. The main goal of authority records has been to establish authorised forms of names/appellations, their variants and evidence of these forms; they have been encoded using MARC formats.

Strengths of traditional authority files

In traditional authority files, standardized access points are established, they contain well documented best practices and are shared among libraries and other institutions as well. Libraries in Czechia participate in "Cooperative creation and use of national authority files", while museums and galleries participate in "Museum Authorities" project and "Register of Fine Art collections" project. All these projects are based on national authorities; Museum authorities and Register of Fine Art collections are limited to the name (personal and corporate body) authorities.

Weaknesses of traditional authority files

Record based system
Traditional library authority files are record-based systems. They are created above all for humans, they are human understandable; but they are not available in machine understandable format, they are machine readable only.

Lack of data granularity
AACR2 and MARC format do not allow sufficient granularity of encoded information. MARC format doesn't offer enough granularity in the data elements that should be clearly defined for Semantic Web and LOD purposes. "The MARC record was not created as a set of data elements but as a format for the storage and display of the text of library catalogue records."[2]

Entities of the same type in different authority files
Another problem is that some entities are classified into various/different authority files in direct relation with MARC fields in which they are encoded. Some entities like events form part of two authority files: Meeting names file (conferences, seminars, workshops) and topical term file (wars, battles etc.). The named buildings form part of Corporate name file, the unnamed ones are classified in Topical file. These entities should be classified according to their attributes and ability to express the relationships to other entities rather than according to a MARC field.

Little information in authority records
Authority records contain only little information about the entity: they are concentrated on the preferred forms of entities name/appellation entered as main headings and their variants entered as see references. There was no possibility to enter all information needed for expressing properties/attributes and relationships of the entity in a structured way. On the contrary, it was possible to mention some information in note fields as a free text only.

---

[2] http://futurelib.pbworks.com/w/page/29114548/MARC%20elements

Term describing type of corporate body

For corporate body headings, it is not possible to indicate a term describing type of corporate body (e.g.: ltd, plc), however for archivists, type of corporate body is one of the most important attributes that distinguishes an individual entity from other similar ones.

National Authority files applied in the INTERMI project

The national authorities were analyzed in depth and it was decided using them according following principles:

- The standardized access points as preferred or variant forms of entities appellations shall be used.

- All information that provides context to entities in authority records shall be accepted in INTERMI system as well as entered in structured form.

- The tendency to make decisions based purely on the format (MARC format) should be avoided whenever possible.

- Proprietary conceptual model and data structure format have to be created to meet all specific needs of user´s communities.

- The INTERMI conceptual model shall be based on entities (not on authorities), it means there will be a strong tendency to move from the authorities to entities of value on the web.

INTERMI conceptual model

The INTERMI conceptual model is based on some principles of object-oriented paradigm, event-centric approach, and is designed as convertible to CIDOC CRM.

Principles of object-oriented paradigm
- data granularity – the data are divided into smaller meaningful elements which allow to define better type of data and to generalize them, so that the flexibility of data is supported

- data inheritance – the hierarchical structure of elements requires and supports inheritance of attributes or properties

- data reuse – the elements are designed to be reused in more than one location in the model and to enable further expansion of the model

- polymorphism – the elements can work as different elements depending on their use

Event-centric approach

Events represent specific actions that occur at a specific time and place and are encoded as information objects. We agree that "event modelling is so abstract that it can be used to describe cultural items and documentation of scientific observations"[3].

There are two types of events in INTERMI conceptual model: events as class of entities and events that are used to express complex attributes and relationships by describing other entities, so called "descriptive entities". Events (class of entities) include a named temporary event (long-term or short-term; one-time or repeated) and also named entities which are related to human activity such as culture, folk habits. "Descriptive entities" are connected to events in the life or history of entity (e.g. birth, establishment, creation, destruction).

Event-centric approach means that as many entity attributes as possible are described by event; we think that the event-centric approach reflects the concept of Linked Data.

Information technologies applied in INTERMI project

The main goal of INTERMI project is to create knowledge database and ontology for information objects that occur within cultural heritage domain. The need to cope with technological questions is self-evident, so we have to deal with two main problems:

- How to represent data covered by INTERMI conceptual model and to provide preservation of this data? This means to find internal data model for data storage that supports openness of INTERMI conceptual model, quick access to data and flexible presentation. We decided to use proprietary XML based structure that supports relations among entities. The base for INTERMI also contains powerful object database.

- How to represent data for external systems (library and management systems, data management systems used in archives and museums, and the Web as well)? Concepts of Semantic Web and Linked Data represent suitable space for INTERMI because they provide technologies for presentation of data in context not for presentation of content only.

Technologies of Semantic Web consist of:

- URI (uniform resource identifier) used for identification of objects and abstract concepts;

- RDF and RDF Schema for representation of information using abstract metadata model, RDF allows to registration of simple statements with triplets (subject – predicate – object);

- OWL – language for ontology creation;

- SPARQL – query language.

Various standards related to Linked Data concept and Semantic Web were considered to find one or more suitable standards for data presentation. As to INTERMI project, following standards should be mentioned:

- SKOS – for representation of thesauri – can be applied on entities from class "General concept",

---

[3] http://www.cidoc-crm.org/docs/fin-paper.pdf

- FOAF – ontology for describing persons, activities and relations – can be applied on entities from class "Person/creature".

For maintaining continuity, it is necessary to provide data from INTERMI knowledge database in current structure that is frequently used in libraries – it means MARC21/XML for authorities.

Presentation of INTERMI data in several specific structures requires a selection of data that can be used in specific structure (or standard). This also requires more data transformation from internal to external structure.

For external systems, access to INTERMI data will be provided by using web services (e.g. to log-in of user, to find record, to get record).

Figure 1. INTERMI conceptual model

Key elements of INTERMI conceptual model: entities and relationships

Entity is the most important element of INTERMI conceptual model.

Entities represent real-world identifiable objects. An entity is a concept (individual and general) that has an identifier, an entity type that describes the category of the entity and is defined by a set of characteristics which allow describing any entity used in the model. One of the most difficult tasks of the project was to define classes of entities and their characteristics. Based on an intensive interdisciplinary discussion it was decided that INTERMI conceptual model would be realized in 7 classes of entities: personal, family, corporate body, geographic object, work, event, and general work entities. Each class is divided to subclasses that are meaningful subgroupings of its entities and need to be represented explicitly because of their significance to the INTERMI knowledge base application.

The definition of a class is based on common characteristics of the entities. The main goal of definition of classes was to easily identify and describe characteristics of entities

belonging to specific class. However it is allowed to use characteristics from one class to describe entity primarily categorized in another class (examples are castles, dams etc.)

Class of entities	Number	Examples of subclasses of entities
Person/Creature	4	real persons, persons named in religious works, fictional and legendary persons, real or fictional non-human entities
Family	3	families, branches of the family , fictional families
Corporate body	12	territory with its own administration, political parties and movements, associations, firms …
Geographical object	9	states, countries, historical regions, geomorphological features
Events	3	organized actions and events (conferences, workshops), important days, wars, battles
Work	7	literary works, artworks, generally known documents (laws, constitutions), structures, named buildings (castles, churches…), programs, projects, grants
General concept	7	objects and their physical parts, categories and groups of no named persons and corporate bodies, abstract entities, materials, techniques

Table 1 Classes of entities and examples of subclasses of entities

Brief characteristics of class of entities used in INTERMI project

Class of entities: Person/Creature
> This class of entities includes real persons and persons named in religious works, fictional and legendary persons, and real or fictional non-human entities.

Person with more biographical and other identities
> There are two possibilities to describe a complex person with more identities in INTERMI project. The first one (and preferred by INTERMI rules) is to describe the person with more identities in one record with one preferred appellation (more used, more known etc.), while other appellations are recorded as variant appellations. The second one mentioned in INTERMI rules is to describe the person with real and fictional identity separately – in two records. It seems to be useful when different biographical data used by alternative identity (as fictional curriculum vitae) exist. It is necessary to create relationships between all records connected to one entity.

Collective pseudonym, anonymous and unknown creator
> A group using collective pseudonym is included in this class of entities if the name they use is similar to that of person. Specific entities as anonymous creator and unknown creator are included in the "person/creature" class of entities as well.

Class of entities: Family
> "Family" class of entities includes families and parts (branches) of families. In the INTERMI project, it was decided to create separate class of these entities to describe specific characteristics of families in their complexity. As to this class, it is to be mentioned that different approach is applied in archives that concede a concept of families as corporate bodies (as named and organized group of people).

Class of entities: Corporate body

"Corporate body" class of entities was one of the most discussed classes regarding to different methods of description applied in archives and libraries. INTERMI rules adopt principles from archives and include the type of corporate body as one of attributes which determine the corporate body entity. Yet, this principle does not comply with the RDA rules.

Class of entities: Geographic object

Geographic object as entity is identified with stable set of geographic coordinates. The stages of development of geographic object are usually described in one record with variant appellations specified by dates indicating the time when the appellation was used. If necessary, INTERMI allows the application of separated record principle for every significant stage of entity development (with corresponding relations among records).

Class of entities: Work

This class represents the most complex class of entities in INTERMI project. The concept "work" within memory institutions communities varies greatly. *Functional Requirements for Bibliographic Records* (hereinafter FRBR) defines "work" as „a distinct intellectual or artistic creation. A work is an abstract entity; there is no single material object one can point to as the work. We recognize the work through individual realizations or expressions of the work, but the work itself exists only in the commonality of content between and among the various expressions of the work".[4]

According to Cataloguing Cultural Objects "a work is a distinct intellectual or artistic creation limited primarily to objects and structures made by humans, including built works, visual art works, and cultural artifacts. …Works include architecture, landscape architecture, other built works, objects such as paintings, sculptures, murals, drawings, prints, photographs, furniture, ceramics, tools, costume, textiles, other decorative or utilitarian objects, or any other of thousands of types of artistic creations and other cultural remains. Performance art, installations, and site-specific works are included. *Excluded are literary works, music, performing arts, language arts, culinary arts, science, religion, philosophy, and other intangible culture*"[5].

As these are two very different approaches, it was very difficult to define this class of entities. In museums and galleries, the FRBR concept of "work" met with misunderstanding because of their different meaning of the concept (as "art work" – product of art creative power).

When identifying the class of entities „work" we define basic characteristics of the entity „work" as follows: "Work" is the result of intentional human activity, and it doesn´t represent a process. Using this high level abstraction of description we are allowed to identify the class of entities „work" acceptable for all memory institutions communities. The description of entity "work" is concentrated to its current state. One work in its complexity is described in one record. If necessary, it is possible to create separated records on significant historical stage or part of work (with relations among records). "Work" class of entities consists of entities that have some material representation, but also of entities which do not exist in representation any more. This class include art and literary works, legal documents (laws etc.), standard documents, products, trademarks, games, grants, projects, buildings etc.

---

[4] http://www.ifla.org/files/assets/cataloguing/frbr/frbr_2008.pdf
[5] http://cco.vrafoundation.org/downloads/PartOne_GeneralGuidelines.pdf

Class of entities: Event

"Event" class of entities includes a named temporary event (long-term or short-term; one-time or repeated) and also named entities which are related to human activity such as culture, folk habits.

According to INTERMI rules, every repetition or interpretation of event is described as separated entity. However, it is possible to create a collective (cover) record that consists of information related to complex history of events. This record is related to all repetitions/interpretations.

Class of entities: General concept

This class includes general concepts for specific entities such as categories of people, things, animals, plants, and also general concepts for abstract entities such as characteristics, scientific areas, art styles, etc.

Relationships between entities in INTERMI conceptual model

A relationship models an association between two or more entities to which all of the occurrences of those entities must conform. Frequently, a meaningful relationship exists between two different types of entity (e.g. creator – work, family – member of family, person - geographical object (place of birth). There are two types of relationships in INTERMI conceptual model: simple and complex.

To create a simple relationship in INTERMI knowledge model rules means to create a link to another entity (e.g. a person to person link) with possibility to specify role of linked entity (e.g. sibling, wife). It is possible to enter additional information about the relationship such as dates and note.

Some relationships are complex: expressing the complex relationships in its entirety requires to define and use one „descriptive" event in the life or history of entity (e.g. birth, establishment, creation, destruction) and a set of relationships that are connected to.

INTERMI entity record - structure of information about an entity

INTERMI entity is described/documented in record that contains a set of related data elements that are stored and processed together. Typical INTERMI entity record includes following information:

Identifier	**INTERPI ID** - unique, obligatory; more identifiers are allowed, e. g. ID of National Authorities, Museums authorities, VIAF ID
Class/subclass	obligatory
Language	language in which the notes and additional information are entered – Czech language
Rules	used in descriptions of attributes and information about the entity, e. g. RDA, AACR2, CCO, Basic Rules (Archives), INTERPI rules
Dates	simple in the form of the text and complex (full information) which enables both: display and search
Notes	public and nonpublic
Brief characteristics, description, history	brief characteristics: maximum two sentences; description, history: maximum 12 000 signs
Appellation	preferred and variant forms, each form is identified by rules; maximum one preferred form for one rule

Parts of appellation	main part, another part of the appellation, additional part of appellation (topical, geographical, chronological information)
Classification	characteristics of the entity expressed through a topical term, category
Relationships and events used in description of the entity	simple relationship is expressed trough link to other entity (with expression of the role of related entity, dates of relationship and note) complex relationship is expressed through events linked to more types of entities in different roles, dates and notes Three types of „descriptive" entities are identified: beginning, end of the existence of the entity, change of the entity
Visual representation of the entity	e.g. picture
Resource for information about the entity	source of information is given to the whole record or to a specific property; standardized description is applied

Table 2 INTERMI entity record

INTERMI knowledge model rules

INTERMI knowledge model rules provide instructions and guidelines on formulating data for memory institutions in Czechia. The aim was not to create proprietary INTERPI rule. At the same time, we wanted to keep the best practices and uniqueness of approach of memory institutions when it is required. We have carefully analyzed the rules used in the memory institutions (AACR2, ISAAR CPF, CCO, RDA) and we came to the conclusion that it was not possible to apply any existing rules in its entirety. We decided to apply RDA whenever possible. However, for specific areas of entities description, e.g. description of entity appellation, description of dates or definition of rules for mapping of controlled vocabularies used in memory institutions, specific instructions have been formulated. Specific instructions are complying with e.g. the Cataloguing Cultural Objects rules (e. g. description of dates) or Basic rules for Archives in Czechia (e.g. description of preferred and variant forms of corporate bodies).

INTERMI rules for entities appellation

Different memory community user groups in Czechia want to have different preferred forms for the appellations of the same entity (e.g. general object entity, corporate body entity); this idea is of key importance to the INTERMI rules: it allows to create more than one preferred appellation of entity in one entity record. However, it is necessary to assign each appellation to those rules according which it was created. This principle enables to provide multiple views on data of entity according to rules used in each memory institution and it is also helpful in thesaurus mapping process. We hope that the principle of multiple preferred appellation of an entity is main forthcoming solution, but in practical application the communities of memory institutions will be able to use one preferred appellation – because they will find it easy and reasonable.

INTERMI rules for dates in entity description

According to using dates in entity description we have decided to adopt principles of CCO. These principles are based on entering dates in two forms – for user access and for computer access. For user access, uncertain dates are described in text form (e.g.

ca/approximately 1894), for computer access in format of ISO 8601, where the first date delimits potential initial date, the second one potential closing date (e.g. 1890–1900).

INTERMI rules for mapping of controlled vocabularies, registries etc. used in memory institutions

The aim of these INTERMI rules is to summarize methods used for mapping various vocabularies to INTERMI general object entities. We suppose that INTERMI controlled vocabularies mapping rules would be important for those institutions creating their own specific controlled vocabulary in a way that fits to Semantic Web.

Mapping/harmonization process depends on:

- the choice of a controlled system of terms that may serve as a basis for comparing and evaluating terms from other terminological resources

- the choice of a tool for formal representation of tedrminological resources (SKOS),

- the description and characteristics (formal and semantic) of terms used in terminological resources,

- the choice of methods applied in process of mapping/harmonization, such as lexical-based, concept-based and instance based-mapping.

The purpose of lexical-based mapping is to ascertain the degree of similarity of text strings.

Pojem	Identifikátor		
	Národní autority	EuroVoc	AGROVOC
politika	PSH8309	131558	6062
informatika	PSH6548	100223	3864
přírodní vědy	PSH11969	141614	4318
trh práce	PSH1264	142776	28713

Figure 2 Example of identical terms of three resources obtained by lexical-based method

Concept-based mapping is applied when terminology of specialized thesauri and controlled vocabularies is compared. It is realized as common graphs method or vector similarity method.
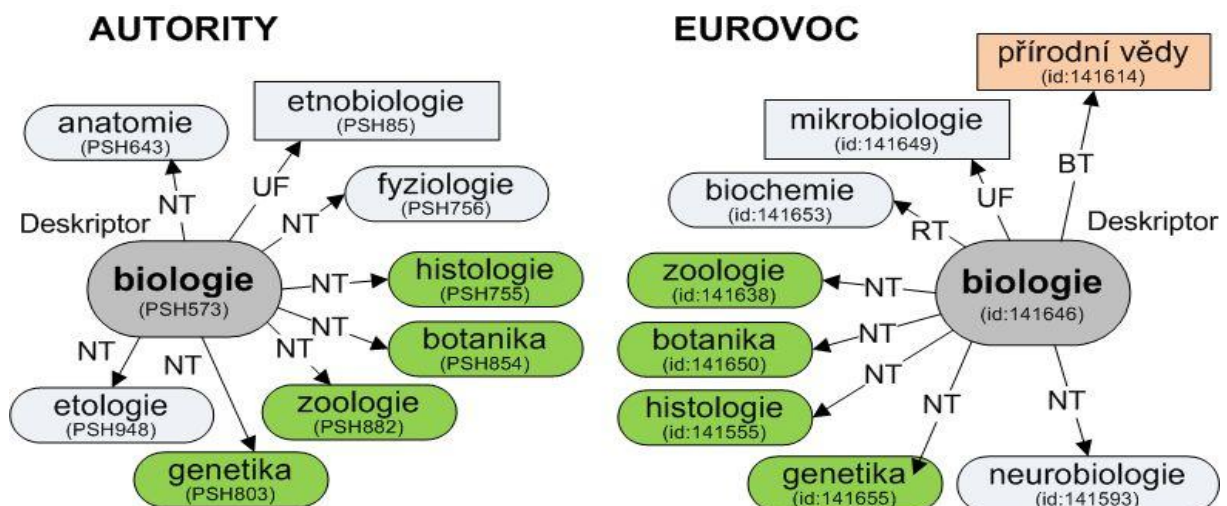
Figure 3 Example of Common graphs method - subgraphs of concept "biology" from terminological resources: National Authorities and EUROVOC
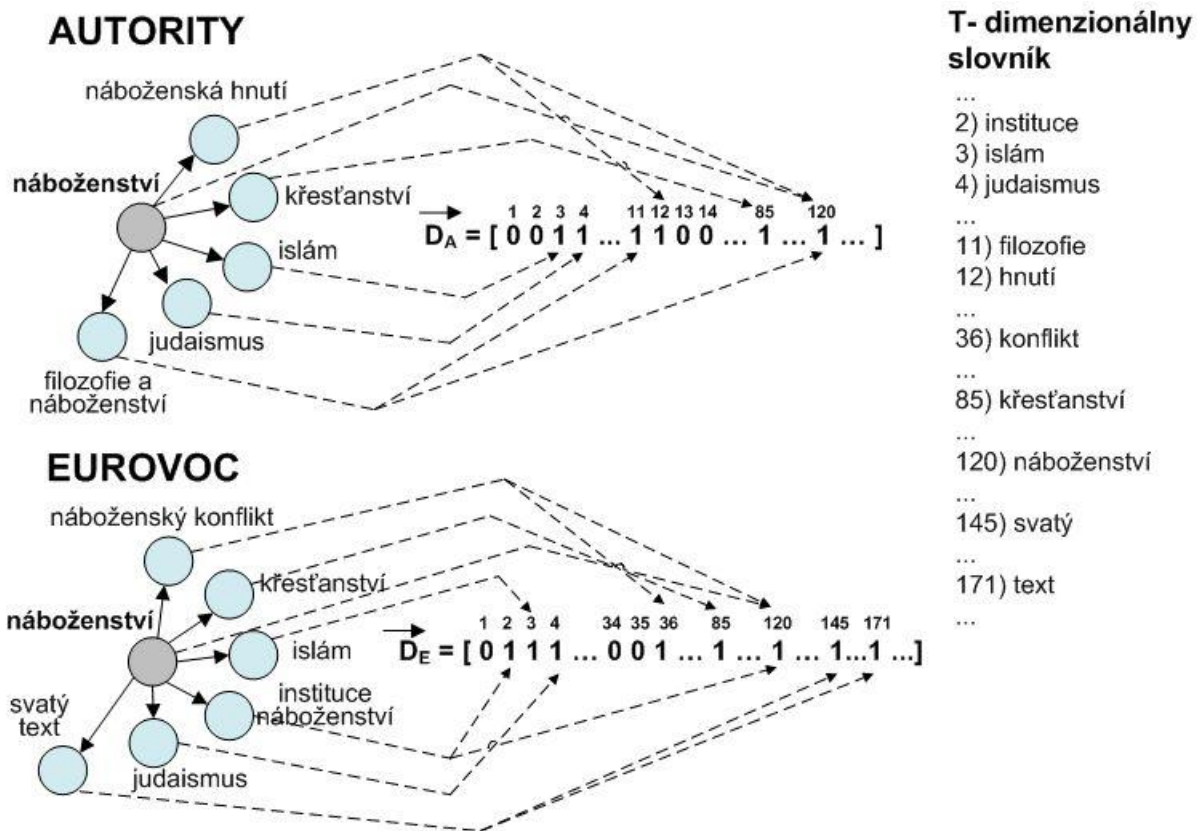

Figure 4 Example of Vector similarity method – T-dimensional vector of concept "religion" in terminological sources: National authorities and Eurovoc

When using the method of instance-based mapping, equivalent, broader, narrower, and related concepts based on the lexical and semantic similarities between source and target thesaurus are searched.
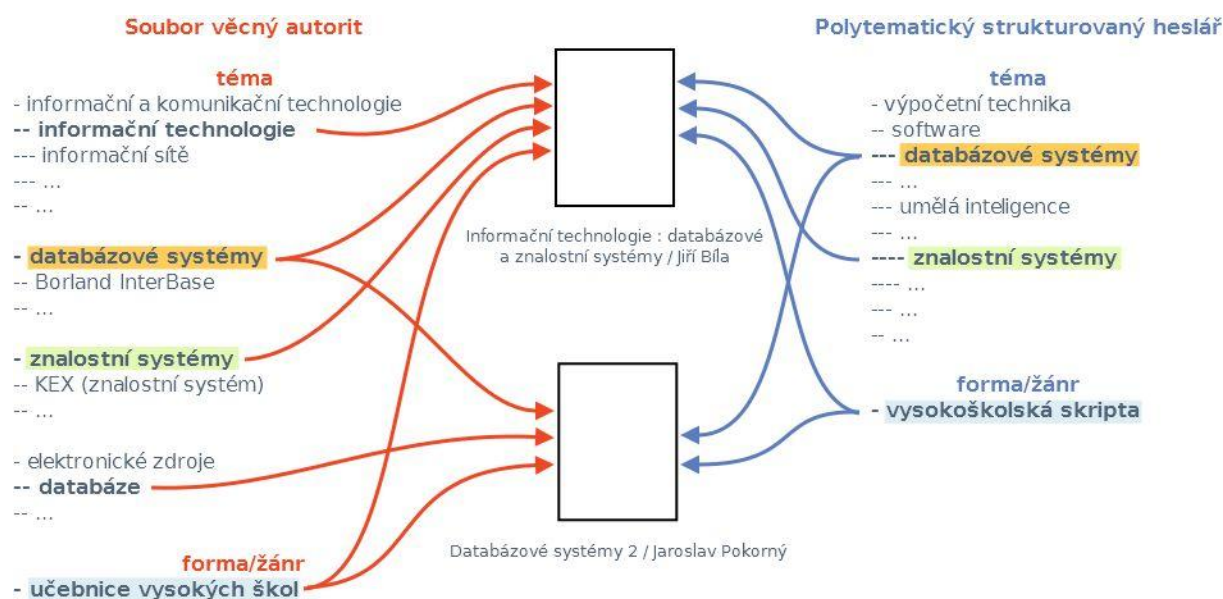
Figure 5 Example of instance-based mapping

The role of Topical authorities

Topical authorities file has been chosen as a basis for mapping/harmonization of terminological resources used in memory institutions. Topical authorities file represents a general set of terms with content covering all fields at general and specific level as well. Concerning Czech subject authority file, our intention is to create a tool enabling efficient mapping/harmonization, it menas to transform it into a „general" ontology. It supposes

- to define the meaning of topical terms exactly by adding qualifiers, UDC notations and scope notes, so that the topical terms are able to represent isolated well defined concepts,

- to express hierarchical structure exactly to be complete and consistent,

- to use permanent unique identifier which shall be languague independent to avoid the necessity to change ID in conjuction with the change of preferred form applied in description of the concept,

- to express topical terms file in machine-understandable way within the framework of the Semantic Web, when using RDF application SKOS; it enables data to be linked and merged with other RDF data by Semantic Web applications.

Conclusion

Importance of the INTERMI project

INTERMI project provides a space for creation and preservation of data about entities used in memory institutions in Czechia. It represents an user-friendly tool for the access and share of the cultural heritage content across the memory institutions ensuring semantic interoperability at the conceptual level. The aim of the project is to create conceptual model and ontology for making Czech national cultural heritage content available on the web. The project builds on earlier research goals and projects, e.g. National authority files and Museum authorities, but in addition it aims for a comprehensive construction of an infrastructure for building a

knowledge based model of the cultural heritage content and for the opening of its use in the form of working pilot operation. The project is complementary to the projects such as the National digital library, the National digital archive.

Impact of collaboration between memory institutions

It should be noted that the project has an impact on discussion among different communities from memory institutions. Experts have had opportunity to meet together and express their needs dealing with the identification and description of entities from their collections. At the end of INTERMI project we foresee an effective collaboration between professionals which would have a practical impact on their everyday practice when describing information resources using INTERMI entities. For realization of this goal, a set of INTERMI web services for application in local information system is provided.

Impact of user access to information

Using INTERMI entities in description of memory institutions collections will improve quality of user access to information on national heritage content published on World Wide Web; at the same time INTERMI project will demonstrate ways to present information about entities with applying technologies of semantic web.

References

Ansgar Scherp , Thomas Franz , Carsten Saathoff , Steffen Staab, A core ontology on events for representing occurrences in the real world, Multimedia Tools and Applications, v.58 n.2, p.293-331, May 2012

American Library Association et al. *Anglo-American Cataloguing Rules.* 2nd Edition. Edited by Michael Gorman and Paul W. Winkler. Chicago: American Library Association, 1978.

Bountouri, L., Papatheodorou, C., & Gergatsoulis, M. (2010, January). Modelling the public sector information through CIDOC conceptual reference model. In *On the Move to Meaningful Internet Systems: OTM 2010 Workshops* (pp. 404-413). Springer Berlin Heidelberg.

Constantia Kakali, Irene Lourdi, Thomais Stasinopoulou, Lina Bountouri, Christos Papatheodorou, Martin Doerr, Manolis Gergatsoulis (2007): Integrating Dublin Core Metadata for Cultural Heritage Collections Using Ontologies. In: Proceedings of the International Conference on Dublin Core and metadata Applications (DC-2007), pp. 128-139. 27-31 August, 2007. Singapore.

Coyle, Karen. "Metadata mix and match." *Information Standards Quarterly* 21, no. 1 (2009): 9-11.

Doerr, M., Gradmann, S., Hennicke, S., Isaac, A., Meghini, C., & van de Sompel, H. (2010, August). The europeana data model (EDM). In *World Library and Information Congress: 76th IFLA general conference and assembly* (pp. 10-15).

Gaitanou, P., & Gergatsoulis, M. (2011). A semantic mapping of VRA Core 4.0 to the CIDOC Conceptual Reference Model. In *Metadata and Semantic Research* (pp. 387-399). Springer Berlin Heidelberg.

Gergatsoulis, M., Bountouri, L., Gaitanou, P., & Papatheodorou, C. (2010). Mapping cultural metadata schemas to CIDOC conceptual reference model. In *Artificial Intelligence: Theories, Models and Applications* (pp. 321-326). Springer Berlin Heidelberg.

Harper, C. A.(2006). Encoding library of congress subject headings in SKOS: authoritym control for the semantic web, Proceedings of the 2006 international conference on Dublin

Core and Metadata Applications: metadata for knowledge and learning, 89-94. Manzanillo, Colima, Mexico: Dublin Core Metadata Initiative.

IFLA Study Group on the Functional Requirements for Bibliographic Records. *Functional Requirements for Bibliographic Records: Final Report.* München: K.G. Saur, 1998.

ISO 25964-1:2011 = International Standards Organization (ISO). Information and documentation -- Thesauri and interoperability with other vocabularies -- Part 1: Thesauri for information retrieval. 1st ed. Geneva: ISO, 2011 (ISO 25964-1- 2011-08-05)

Joint Steering Committee for Development of RDA. *RDA: Resource Description and Access.* Chicago: American Library Association, 2010.

Kramer, R., R. Nikolai, and C. Habeck (1997) "Thesaurus federations: loosely integrated thesauri for document retrieval in networks based on Internet technologies". In *International Journal on Digital Libraries* (1), 122-131

Landry, P. (2000) "The MACS Project: Multilingual Access to Subjects (LCSH, RAMEAU, SWD)". *Classification and Indexing Workshop, 66th IFLA Council and General Conference*, Meeting No. 181 http://www.ifla.org/IV/ifla66/papers/165-181e.pdf

Lourdi, Irene, Christos Papatheodorou, and Martin Doerr. "Semantic integration of collection description." *D-Lib Magazine* 15.7/8 (2009): 1082-9873.

MERIMEE, "THESAURUS ARCHITECTURE" for the indexing of complexes, buildings and built works described in the national database "Merimee" about the French Heritage http://www.culture.gouv.fr/documentation/thesarch/pres.htm

Mitchell, E. T. (2013). Chapter 1: Metadata Developments in Libraries and Other Cultural Heritage Institutions. *Library Technology Reports*, *49*(5), 5-10.

Soergel, D. (1995) *The Art and Architecture Thesaurus (AAT): A critical appraisal*. Visual Resources, X, pp. 369-400

Svensson, Lars G. "Are Current Bibliographic Models Suitable for Integration with the Web." *Information Standards Quarterly* 25 (2013): 7-12.
Thomas R. Gruber (1993): Toward Principles for the Design of Ontologies Used for Knowledge Sharing. In: Nicola Guarino and Roberto Poli (eds.): Formal Ontology in Conceptual Analysis and Knowledge Representation. Thomais Stasinopoulou, Lina Bountouri, Constantia Kakali, Irene Lourdi, Christos Papatheodorou, Martin Doerr, Manolis Gergatsoulis (2007): Ontology-based Metadata Integration in the Cultural Heritage Domain. In: Proceedings of the 10th International Conference on Asian Digital Libraries, pp.165-175. Hanoi, Vietnam, December 10-13, 2007.

Skinner, J. (2014). Metadata in Archival and Cultural Heritage Settings: A Review of the Literature. *Journal of Library Metadata*, *14*(1), 52-68.