



DELIVERABLE

Project Acronym: thinkMOTION
Grant Agreement number: 250485
Project Title: Digital Mechanism and Gear Library goes Europeana

D2.4 - Intermediate report on DMG-Lib thesaurus via SKOS-format

Revision: 1.0

Authors:

Ulf Döring (Ilmenau University of Technology)
Sascha Falke (Ilmenau University of Technology)

Project co-funded by the European Commission within the ICT Policy Support Programme		
Dissemination Level		
P	Public	x
C	Confidential, only for members of the consortium and the Commission Services	

Revision History

Revision	Date	Author	Organization	Description
0.1	25.04.2012	U. Döring	IUT	Draft
0.2	25.01.2013	S. Falke	IUT	Revised Table of Content, SKOS introduction
0.3	11.02.2013	U. Döring	IUT	SKOSification
1.0	14.02.2013	S. Falke	IUT	Review and Final Version

Statement of originality:

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

Table of Contents

1	Introduction	4
2	Workflow and results.....	4
2.1	The DMG-Lib thesaurus	4
2.2	The Simple Knowledge Organization System (SKOS)	6
2.3	SKOSification of the DMG-Lib thesaurus.....	7
2.4	Conclusion.....	10

1 Introduction

In June 2010 partners from six European universities started the project thinkMOTION with the main objective of providing content from the field of motion systems via the Europeana online portal. Therefore, the currently available and all the DMG-Lib content collected in the proposed project must become accessible to Europeana - that means to Europeana users as well as to Europeana tools (e.g. the harvester).

To support a more intelligent search based on semantic information the DMG-Lib thesaurus will be made available via SKOS. That allows Europeana to transform search terms concerning language, synonyms, subterms etc. or to support interactive browsing in Europeana's semantic network. Making the DMG-Lib thesaurus accessible via SKOS-format is part of task 2.4 and described in this document.

Task 2.4 is part of WP2 – Adaptation of interfaces to Europeana – included in the “Description of Work” for thinkMOTION project.

2 Workflow and results

One of Europeana's objectives is to become “a network of inter-operating object surrogates enabling semantics based object discovery and use.”¹ To allow semantic interoperability with Europeana it is necessary to “SKOSify” terminology resources such as the DMG-Lib thesaurus. This deliverable describes the DMG-Lib thesaurus as semantic information resource and its “SKOSification” to support interactive browsing in Europeana's semantic network.

2.1 The DMG-Lib thesaurus

A thesaurus is a controlled vocabulary. It consists of terms which are connected with each other by relations.² These relationships can be

- hierarchical to indicate terms which are narrower and broader in scope,
- equivalent to connect synonyms and near-synonyms or
- associative to express relationships which are neither hierarchical nor equivalent.

Figure 1 shows an exemplarily thesaurus.

```
Animals  
Cats  
    UF (used for) domestic cats  
    RT (related term) wildcats  
    BT (broader term) animals  
    SN (scope note) used only for domestic cats  
Domestic cats  
    USE cats  
Wildcats
```

Figure 1. Thesaurus example³

¹ Dekkers, M.; Gradmann, S.; Meghini, C.. 25 May 2008. *Europeana Outline Functional Specification*. EDLnet D2.5 Version 1.0. Retrieved January 25, 2013, from <http://abm.ylm.se/europeanalocal/pdf/EuropeanaOutline08.pdf>

² *Thesaurus*. (n.d.). In: Wikipedia. Retrieved January 25, 2013, from <http://de.wikipedia.org/wiki/Thesaurus>

³ Isaac, A.. 21 September 2011. *SKOS (Simple Knowledge Organization System)*. Presentation held at Dublin Core 2011 conference. Retrieved January 25, 2013, from http://dublincore.org/resources/training/dc-2011/Tutorial_Isaac.pdf

According to the definition above, the DMG-Lib thesaurus is a controlled multilingual vocabulary which is used to improve access to information from the field of mechanism and gear science. The DMG-Lib thesaurus gathers knowledge in a logical, structured and organized way where relationships between single terms and items become clear. For optimal integration of the thesaurus in our production database ProDB, a new item type “Concept” was defined. The concept item provides a preferred name, synonyms, related concepts and definitions in up to seven languages each, as well as examples from real life to support user’s cross-language information retrieval and to provide a better understanding for technical terms in mechanism and gear science. Table 1 describes the structure of a concept item in detail.

Property	Description	Example
<i>ID</i>	Each concept item is identified by a unique numeric permanent ID	concept_15056_coupler
<i>Name</i>	Preferred name of a concept, one per language	coupler @en; Koppel @de; barre de couplage @fr; плавающее звено @ru
<i>Alternative name</i>	Alternative names may include synonyms, spelling variants, scientific and common forms or various forms of speech of a name	floating link @en; bielle @fr; member flottant @fr; шатун @ru
<i>Hidden name</i>	Hidden names may include misspellings or deprecated spellings of names	biéle @fr
<i>Definition</i>	Definition of a concept which describes the meaning of the concept in the context of mechanism and gear science	@en: Link that is not connected directly to the frame. @de: Glied, das nicht unmittelbar mit dem Gestell gelenkig verbunden ist. @fr: Membre qui n’est pas directement relié au bâti. @ru: Звено, которое не соединяется непосредственно со стойкой.
<i>Resource</i>	Original source of the concept, e.g. ID of a term within the IFToMM dictionary	IFToMM Dictionary, Chapter 1, 1.1.9
<i>Occurrence</i>	Examples from the DMG-Lib where the concept is described or its names are used. Such examples may be images, videos or animations illustrating a concept, related persons (e.g. inventor), as well as definitions or descriptions written in text documents	Image: http://www.dmg-lib.org/dmglib/handler?image=31547023 Document: http://www.dmg-lib.org/dmglib/handler?docum=19236009&style=pixo&phyPageNo=40
<i>Broader</i>	Hierarchical link to a parent concept with a more general meaning	concept_7056_link
<i>Narrower</i>	Hierarchical link to a child concept with a more specific meaning	concept_22056_connecting_rod
<i>Related</i>	Associative link to other concepts excluding equivalents and hierarchical relationships	concept_10056_frame; concept_114056_spatial_mechanism
<i>Data provider</i>	Creator and/or editor of the concept item within the ProDB	IFToMM, thinkMOTION
<i>Comment</i>	Comments within the change logs by the editor of the concept, including timestamps	Torsten Brix, 2013-01-10 at 14.30: Have to look for a better image

Table 1. Properties of a “Concept” item within the DMG-Lib thesaurus

The DMG-Lib thesaurus grows continuously. The International Federation for the Promotion of Mechanism and Machine Science (IFTToMM) gave permission to integrate the existing IFTToMM dictionary⁴ into the DMG-Lib thesaurus. The IFTToMM dictionary expresses itself as “a standard terminology (with definitions) for mechanism and machine science, including terms pertaining to the realization of motion in machines and its control, associated problems in dynamics and kinematics, and in machine components”⁵ and contains more than 1500 concepts available in English, German, French and Russian. Another resource is a database from the “Politehnica” University of Timisoara, our Romanian thinkMOTION partner, with about 400 concepts with specific terms from mechanisms science, available in Romanian, English, German and French as a complementary resource to the IFTToMM dictionary.

Currently, the DMG-Lib thesaurus is available via ProDB only. For end users, it will be available via the DMG-Lib portal too at the end of the thinkMOTION project.

2.2 The Simple Knowledge Organization System (SKOS)

Representing and exploiting a vocabulary such as the DMG-Lib thesaurus on the Web requires a format standard. In case of Europeana, they decided to format all the harvested metadata in SKOS. This allows Europeana for a homogeneous and effective exploitation of the resources, of the data and their related descriptions.

The Simple Knowledge Organization System (SKOS) is “a common data model for sharing and linking knowledge organization systems via the Web.”⁶ Knowledge organization systems, such as thesauri, taxonomies, classification schemes and subject heading systems, often share a similar structure. SKOS enables capturing and expressing of this similarity as machine-readable data, and exchanging it between diverse computer applications. SKOS is based on the RDF specification and enables migration towards OWL ontologies. Figure 2 shows the RDF XML serialization of the exemplarily thesaurus (see Figure 1) using SKOS.

```
<rdf:RDF>
<skos:Concept rdf:about="http://example.org/animals">
  <skos:prefLabel xml:lang="en">Animals</skos:prefLabel>
</skos:Concept>
<skos:Concept rdf:about="http://example.org/cats">
  <skos:prefLabel xml:lang="en">Cats</skos:prefLabel>
  <skos:altLabel xml:lang="en">Domestic cats</skos:altLabel>
  <skos:scopeNote>Used only for domestic cats</skos:scopeNote>
  <skos:broader rdf:resource="http://example.org/animals"/>
  <skos:related rdf:resource="http://example.org/wildcats"/>
</skos:Concept>
<skos:Concept rdf:about="http://example.org/wildcats">
  <skos:prefLabel xml:lang="en">Wildcats</skos:prefLabel>
</skos:Concept>
</rdf:RDF>
```

Figure 2. RDF XML serialization of a thesaurus example³ (see Figure 1)

⁴ IFTToMM. May 2012. *IFTToMM dictionaries online*. Retrieved January 25, 2013, from <http://www.iftomm.3me.tudelft.nl/>

⁵ IFTToMM. (n.d.). *IFTToMM PC on Standardization of Terminology*. Retrieved January 25, 2013, from http://www.iftomm.org/index.php?option=com_content&view=article&id=102

⁶ W3C. 18 August 2009. *SKOS Simple Knowledge Organization System Reference*. Retrieved January 25, 2013, from <http://www.w3.org/TR/2009/REC-skos-reference-20090818/>

In the following, the main features of the SKOS model are described in accordance to the SKOS reference⁶.

The SKOS data model is concept-oriented. A SKOS concept can be seen as an idea, a notion or a unit of thought. It is introduced as a class *skos:Concept* and can be used to describe the conceptual or intellectual structure of a knowledge organization system. SKOS concepts are identified by unique URI's to allow referencing and to avoid any ambiguity.

SKOS concepts can be labeled with any number of lexical Unicode strings in any given natural language. One of these labels in any given language can be indicated as the preferred label for that language, introduced in the SKOS data model as the *skos:prefLabel* property. Other labels, e.g. for synonyms, can be indicated as alternative labels, introduced as *skos:altLabel* property. Labels may also be hidden, introduced by the *skos:hiddenLabel* property, e.g. for mentioning misspellings or deprecated spellings of labels. Users will be able to find the relevant concept, but cannot see the hidden label to avoid further mistakes.

SKOS concepts can be assigned one or more notations, introduced as *skos:notation* property. A notation is a string of characters, mainly not recognizable or understandable in any natural language such as classification codes used in library catalogs, used to uniquely identify a concept within the scope of a given concept scheme.

Furthermore, SKOS concepts can be documented with notes of various types to give the most accurate additional information. The different types of notes are

- Note (*skos:note*)
- Change note (*skos:changeNote*)
- Definition (*skos:definition*)
- Editorial note (*skos:editorialNote*)
- Example (*skos:example*)
- History note (*skos:historyNote*)
- Scope note (*skos:scopeNote*)

The strength of the SKOS data model is the possibility of linking between concepts via semantic relation properties. SKOS supports hierarchical and associative links between concepts. Hierarchical links are introduced as *skos:broader* property to assert that a concept has more general meaning or as *skos:narrower* property to assert that it has a more specific meaning. Associative links are introduced as *skos:related* to link two concepts which are neither equivalent nor broader/narrower concepts.

Also, SKOS concepts can be mapped to other SKOS concepts in different concept schemes. Four basic types of mapping links are supported – hierarchical, associative, close equivalent and exact equivalent – which are described in detail in the SKOS reference⁶.

2.3 SKOSification of the DMG-Lib thesaurus

The aim of the Europeana Data Model (EDM) is an integration medium for collecting, connecting and enriching the descriptions provided by Europeana content providers. EDM uses a well-identified set of elements in order to include any element found in a content provider's description. Therefore, EDM re-uses existing namespaces, such as SKOS for all

entities from knowledge organization systems like thesauri.⁷ Table 2 shows the properties of the *skos:Concept* class, which are supported by EDM.

Property	Data type	Integrity Conditions
<i>skos:prefLabel</i>	literal	min 0, max 1 per lang tag
<i>skos:altLabel</i> , <i>skos:hiddenLabel</i>	literal	min 0, max unbounded
<i>skos:broader</i> , <i>skos:narrower</i> , <i>skos:related</i>	reference (to a Concept)	min 0, max unbounded
<i>skos:broadMatch</i> , <i>skos:narrowMatch</i> , <i>skos:relatedMatch</i>	reference (to a Concept)	min 0, max unbounded
<i>skos:exactMatch</i> , <i>skos:closeMatch</i>	reference (to a Concept)	min 0, max unbounded
<i>skos:note</i>	literal	min 0, max unbounded
<i>skos:notation</i>	string (+ rdf:datatype attribute)	min 0, max unbounded
<i>skos:inScheme</i> (URI should resolve to something meaningful)	reference (to a ConceptScheme)	min 0, max unbounded

Table 2. Properties of the *skos:Concept* class supported by EDM⁸

SKOSification means to transform the file entities of the various Europeana content providers into semantic web resources in terms of the SKOS standard with unique URIs to support Europeana in providing a semantic data layer to its content.⁹

The DMG-Lib thesaurus could be mapped relatively straightforward to adequate EDM properties via SKOS format. Table 3 shows the mapping of the DMG-Lib thesaurus to the SKOS standard and further to the Europeana Data Model.

Only two properties required adaptations. The *skos:definition* and *skos:example* properties are not supported by EDM. But, these more specific types of information are covered by the *skos:note* property in EDM for general documentation purposes.

⁷ Europeana. 24 February 2012. *Definition of the Europeana Data Model elements*. Version 5.2.3. Retrieved January 25, 2013, from <http://pro.europeana.eu/documents/900548/bb6b51df-ad11-4a78-8d8a-44cc41810f22>

⁸ Europeana Labs Wiki. (n.d.). *Classes for contextual resources*. Retrieved January 25, 2013, from <http://europeanalabs.eu/wiki/EDMObjectTemplatesProviders#skos:Concept>

⁹ Gradmann, S.. April 2010. *Europeana White Paper 1: Knowledge = Information in Context*. Retrieved January 25, 2013, from http://pro.europeana.eu/c/document_library/get_file?uuid=cb417911-1ee0-473b-8840-bd7c6e9c93ae&groupId=10602

DMG-Lib thesaurus property	SKOS property	Europeana EDM property
<i>ID</i>	skos:Concept	skos:Concept
<i>Name</i>	skos:prefLabel	skos:prefLabel
<i>Alternative name</i>	skos:altLabel	skos:altLabel
<i>Hidden name</i>	skos:hiddenLabel	skos:hiddenLabel
<i>Definition</i>	skos:definition	skos:note
<i>Resource</i>	skos:notation	skos:notation
<i>Occurrence</i>	skos:example	skos:note
<i>Broader</i>	skos:broader	skos:broader
<i>Narrower</i>	skos:narrower	skos:narrower
<i>Related</i>	skos:related	skos:related
<i>Data provider</i>	skos:note	-
<i>Comment</i>	skos:editorialNote	-

Table 3. DMG-Lib thesaurus mapping to SKOS and EDM

Furthermore, a verification interface was implemented within our production database ProDB, where editors immediately can check the accuracy of the SKOSification of a concept item (Figure 3). A final check against consistency conditions defined in the SKOS specification by an online SKOS validator¹⁰ revealed no inconsistencies.

The screenshot shows the ProDB interface for a concept named 'coupler'. At the top, there are navigation buttons: 'Test', 'Send', 'from DB', and 'End'. Below these, there is a small diagram of a coupler. The main content area displays the following information:

- Concept name: **coupler** (coupler | floating link | Koppel | barre de couplage | membre flottant | bielle | плавающее звено | шатун)
- Term IDs: **concept_15056_coupler | 1.1.9**
- Current state: hidden

Below this, there is a sidebar with tabs: Main, Names, Defin., Occu., Data prov., SKOS, raw data, Export, Import, and Logs. The main content area shows the following information:

- Main: @prefix skos: <http://www.w3.org/2004/02/skos/core#> .
- Names: @prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
- Defin.: @prefix dmg: <http://www.dmg-lib.org/dmglib/thesaurus/> .
- Occu.: @prefix dmgskos: <http://www.dmg-lib.org/dmglib/thesaurus/skos#> .
- Data prov.: dmg:concept_15056_coupler rdf:type skos:Concept ;
- SKOS: skos:prefLabel "coupler"@en ; skos:altLabel "floating link"@en ;
- raw data: skos:prefLabel "Koppel"@de ;
- Export: skos:prefLabel "barre de couplage"@fr ;
- Import: skos:altLabel "membre flottant"@fr ;
- Logs: skos:altLabel "bielle"@fr ; skos:prefLabel "плавающее звено"@ru ; skos:altLabel "шатун"@ru ; skos:hiddenLabel "bi le"@fr ;
- skos:definition "Link that is not connected directly to the frame."@en ;
- skos:definition "Glied, das nicht unmittelbar mit dem Gestell gelenkig verbunden ist."@de ;
- skos:definition "Membre qui n'est pas directement reli  au b ti."@fr ;
- skos:definition "Звено, которое не соединяется непосредственно со стойкой."@ru ;
- skos:example <http://www.dmg-lib.org/dmglib/handler?docum=19236009&style=pixo&phyPageNo=40> ;
- skos:example <http://www.dmg-lib.org/dmglib/handler?image=31547023> ;
- skos:narrower dmg:concept_22056_connecting_rod ;
- skos:related dmg:concept_114056_spatial_mechanism ;
- skos:broader dmg:concept_7056_link ;
- skos:related dmg:concept_10056_frame ;
- skos:notation "1.1.9"^^dmgskos:iftomm_dict_code .

Figure 3. ProDB interface for verification of SKOSified concepts

¹⁰ PoolParty. *Consistency Checks for SKOS Thesauri*. Retrieved February 14, 2013, from <http://demo.semantic-web.at:8080/SkosServices/check>

2.4 Conclusion

The semantic interoperability with Europeana was achieved by the SKOSification of the DMG-Lib thesaurus. The thesaurus can be used for metadata enrichment and for high quality disclosure of mechanism-related content within Europeana. Moreover, it will be made publicly available within the DMG-Lib portal, ready to be used by all interested.