Promoting Government Controlled Vocabularies for the Semantic Web: the EUROVOC Thesaurus and the CPV Product Classification System

Luis Polo Paredes, Jose María Álvarez Rodríguez, and Emilio Rubiera Azcona

Fundación CTIC, Gijón, Asturias, Spain,
{luis.polo,josem.alvarez,emilio.rubiera}@fundacionctic.org,
WWW home page: http://www.fundacionctic.org

Abstract. The aim of the paper is to promote government controlled vocabularies for the Semantic Web in the context of the European Union. We will propose SKOS as the RDF/OWL common data model and an enclosed conversion method for the knowledge organization systems available in RAMON, the Eurostat Metadata Server. The study cases of this paper will be the Eurovoc Thesaurus and the Common Procurement Vocabulary, a product classification system.

1 Introduction

Knowledge Organization Systems (*KOS*), such as thesauri, taxonomies or classification systems, are developed by specific communities and institutions in order to organize huge collections of information objects: documents, texts, webpages, and multimedia resources as well. These vocabularies allow users to annotate the objects and easily retrieve them, promoting lightweight reasoning in the Semantic Web. Topic or subject indexing is an easy way to introduce machine-readable metadata for a resource's content description.

In the european eGovernment context, there are several conceptual/terminological maps of particular domains available in RAMON[1], the Eurostat's metadata server: in the Health field, the *European Schedule of Occupational Diseases* or the *International Classification of Diseases*; in the Education field, thesauri as *European Education Thesaurus* or the *European Glossary on Education*; in the Employment field, the *International Standard Classification of Occupations* among others. The structure and features of these systems are very heterogeneous, although some common aspects can be found in all of them: 1. Hierarchical relationships between terms or concepts. 2. Multilingual character of the information.

In this paper, we propose a common data model for RDF/OWL encodings of governmental controlled vocabularies and an enclosed generic method to allow straight-forward conversions. The SKOS vocabulary has been selected as

[1] http://ec.europa.eu/eurostat/ramon

the target common data model, thus we avoid to develop an ontology from scratch in order to metamodel thesauri and classification systems. Adopting a common "semantic" format will facilitate semantic interoperability and common understanding for information interchange between European digital libraries, governmental agencies and private third-parties. We will analyze and convert to SKOS two existing EU knowledge organization systems.

1. **The Eurovoc thesaurus**[2] is a multilingual, polythematic thesaurus focusing on the law and legislation of the European Union (EU). It is available in 21 official languages of the EU. Within the EU, the Eurovoc thesaurus is used in the Library of the European Parliament, the Publication Office as well as other information institutions of the EU. Moreover, the Eurovoc thesaurus is used in the libraries and documentation centers of national parliaments (e.g. Spanish Senate) as well as other governmental and private organizations of member (and non-member) countries of the EU.
2. **The Common Procurement Vocabulary**[3] (CPV) is a single product classification for describing the subject matter of public contracts, allowing companies to easily find public procurement notices[4] and increasing competitivity and business opportunities within European market. Its main goal is to standardize the codes used by contracting authorities. The use of the CPV is mandatory in the European Union from February 1, 2006 and it is regulated by Commission Regulation adopted on November 28, 2007 amending Regulation (EC) N^{o} 2195/2002 of the European Parliament.

This paper is structured as follows: in Section 2, we check different approaches for thesauri and product classification schemes conversions to RDF/OWL format. In Section 3, we propose the minimun set of common requirements for *KOS* sytems conversion, and we select a generic conversion method. In Section 4, we apply the method to the EUROVOC thesaurus and the CPV. Finally, we evaluate the results of our conversions and we present some conclusions.

2 Existing approaches for converting controlled vocabularies to RDF/OWL

This section discusses existing methods to convert *KOS* systems. We distinguish between RDF/OWL conversions methods for thesauri and product classification systems.

2.1 Thesauri Conversion Methods

A thesaurus is a controlled vocabulary, with equivalent terms explicitly identified and with ambiguous words or phrases (e.g. homographs) made unique. This set

[2] http://europa.eu/eurovoc/
[3] http://europa.eu/scadplus/leg/en/lvb/l22008.htm
[4] Published in *Tender Electronical Daily*.

of terms also may include broader-narrower or other relationships. Usually they are considered to be the most complex of controlled vocabularies.

Thesauri as a *KOS* system can be converted to RDF/OWL by means of different procedures. On one hand, there are methods, as the Soergel et al. one in [14] or Van Assem et al. in [15], that propose specific techniques for thesauri conversions into an ontology. However their method does not target a specific output format and it considers the hierarchical structure of thesauri as logical *is-a* relationships. On the other hand, there are some generic methods for thesauri conversions, as the step-wise method defined by Miles et al. in [10]. This method selects a common output data model, the SKOS vocabulary, and is comprised by the following steps: a) generation of the RDF encoding, b) error checking and validation and c) publishing the RDF triples on the Web. In addition, this method has been refined in [16], adding three new substeps for the generation of RDF encoding: 1. analyzing the vocabulary, 2. mapping the vocabulay to SKOS properties and classes and 3. building a conversion program .

2.2 Product Classification Systems Conversion Methods

Product Classification Systems (also known as PCSs) have been developed to organize the marketplace in several vertical sectors that reflect the activity (or some activities) of economy and commerce. They have been built to solve specific problems of interoperability and communication in e-commerce [9] providing a structural organization of different kind of products tied together by some economical criteria. The aim of a PCS is to be used as a *de facto* standard by different agents for information interchange in marketplaces [13,3].

Many approaches for product classification systems adaptation to the Semantic Web, like [2,7,8], present methods with the goal to convert them to domain-ontologies. The tree-based structure between product terms is interpreted then as a logical *is-a* hierarchy. From our point of view and following the discussion about [5,6], hierarchical links between the elements of each economic sector do not have the semantics of subsumption relationships. The next example taken directly from CPV ("term" and its code) shows how the relationship between the element "Parts and accessories for bicycles" (34442000-7) and its direct antecessor,"Bicycles" (34440000-3), does not seem as an *is-a* relation. In this case, an ontological property for object composition like *hasPart* would be much better. Moreover, there are further remarks against the idea of using the PCSs as domain-ontologies. It is difficult to assert that the CPV element, "Tin bars, rods, profiles and wire" (27623100), represents any specific product. Rather it should be regarded as a collection of products. To convert correctly this element into a domain ontology, it should be considered as equivalent to the union of several concepts (e.g. $TinBar \sqcup TinRod \sqcup TinProfiles \sqcup TinWire$).

Our approach instead do not consider PCSs as domain ontologies, but as a specific kind of *KOS* systems. Any PCS, as well as other classification systems (i.e. product classification systems, economic activities classification systems, occupation classification systems, etc.), are interpreted as a conceptual scheme comprised of conceptual resources. From this point of view, hierarchical relationships are not considered to be any more logical *is-a* relations, but broader/narrower ones.

3 "Greatest common divisor" of controlled vocabularies

As we have just introduced in the previous section, Knowledge Organization Systems are used for organizing large collections of information objects and efficient retrievery. Existing controlled vocabularies are currently available in several formats: XML files, spreadsheets or text. However promoting them to the the Semantic Web is not a mere process of RDF/OWL conversions of data. Conversions need to fullfil some requirements. Firstly, a common RDF/OWL representation is needed to ensure a) semantic compatibility between different vocabularies, b) processing vocabularies in a standard way and c) sharing vocabularies for third-parties adoption. SKOS, presented in the W3C SKOS Reference Working Draft [11], has been selected for these purposes. Secondly, although controlled vocabularies do not share some features, in practice a distinction between them is very hard to draw. We have identified a minimun set of common features for them. Therefore the data model should be expressive enough to preserve as much as possible the original semantics of primary sources for these common features. Thirdly, a generic method is needed to ensure the quality of data conversions to correct SKOS instances.

We have carried out a refinement of the methods [10,16] for thesauri conversions, by extending it to the PCSs field and taking into account their special features commented in section 2.2. These are the common features of *KOS* systems that have to be covered by the conversion method:

URI generation. Controlled structured vocabularies and conceptual resources are interpreted in SKOS as RDF resources: in particular, instances of `skos:ConceptScheme` and `skos:Concept`. Thus they are referenced by means of Uniform Resource Identifiers (URIs). Although namespaces are out of the scope of our analysis, one of the substeps of the method is the generation of the `rdf:IDs` of `skos:Concept` and `skos:ConceptScheme` from the original datasource. Controlled vocabularies usually provide unique identifiers for their terms or concepts. The options are the following:

1. Generating new identifiers for the elements of the vocabulary. This option introduces additional management. A mapping between elements of the original source and identifiers should be maintained for updating purposes.
2. Using the string of the preferred term. We would like to highlight here that multilingual sources introduce a factor of complexity that it is not present in monoligual systems. In european multilingual sources, this solution implies selecting a preferred term in a given natural language, thus promoting one language over the others with a possible non-desired political impact. In addition, a control procedure has to be established to ensure URI updating if the source term changes.

3. Using the identifier code of an element, if any. This solution avoids the problem of selecting one preferred language to encode the concept URIs. Moreover, codes are usually strings composed by a serial number (legal URI characters) and it preserves the original semantics of a multilingual vocabulary, where these codes identify unique terms or concepts and establish mappings between different languagues. This last option has been chosen for our method.

Hierarchy formalization. From our point of view, one of the common aspects shared by *KOS* is a hierarchy-based structure, at least by thesauri, taxonomies and by most of classification schemes [1]. Hierarchical relations establish links between conceptual resources, showing that the semantics of a resource is in some way more general ("broader") than other ("narrower"). In SKOS, the properties `skos:broader` and `skos:narrower` are only used to assert hierarchical statements between two conceptual resources. By the way, these properties are not currently defined ([11]) as transitive properties (as they were in [12]). Nevertheless, third-parties, if they consider valuable, can use an OWL reasoner to infer the transitive closure of the hierarchy by means of the transitive superproperties of `skos:broader` and `skos:narrower`: `skos:broaderTransitive` and `skos:narrowerTransitive` properties.

Multilingual and lexical features. Regarding European controlled vocabularies, multilinguism is a critical issue. Both CPV Vocabulary and Eurovoc Thesaurus are available in 21 official languages of the European Union (Bulgarian, Spanish, Czech, Danish, German, Estonian, Greek, English, French, Italian, Latvian, Lithuanian, Hungarian, Dutch, Polish, Portuguese, Romanian, Slovak, Slovene, Finnish and Swedish) and one more (Croatian). In SKOS conceptual resources are labelled with any number of lexical strings, in any given natural language identified by the `xml:lang` attribute, following normative RDF/XML syntax. One of these labels is selected as the `skos:prefLabel` for any given language, and the others as values of `skos:altLabel`.

In thesauri like Eurovoc, the `USE` and `UF` relations between descriptors and non-descriptors are language specific. Both, `USE` and `UF` relations, express term-equivalence relationships (synonymy, antonymy, holonymy, etc.). The W3C Semantic Web Deployment Working Group[5] is currently working on an extension to SKOS, SKOS-XL[6], for describing in detail these term relationships and for modeling lexical entities. A special class of lexical resources, called `xl:Label`, is defined. The use of this class would be much more accurate, as Eurovoc is a term-based thesaurus. However, there are mainly two problems 1) this information is not explicitely represented in the original datasources (thus the type of equivalence relation has to be detected by a human expert) and 2) the approach of the SWD Working Group is still in a preliminar stage. Moreover, notice that there is no additional benefits for the treatment of multilingual features: the language identification is still encoded using the `xml:lang` attribute.

[5] http://www.w3.org/2006/07/SWD/
[6] http://www.w3.org/2006/07/SWD/SKOS/xl/20080414

4 Case Studies

In this section, we apply our method to the *Eurovoc Theusarus* and the *CPV vocabulary*. Firstly, we generate the RDF/OWL encoding:

4.1 Adaptation of the EUROVOC Thesaurus to SKOS

Step 1: analyze controlled vocabulary. We used the XML version in our analysis. Eurovoc/XML structure is specified using a DTD associated with language specific files, providing a multilingual scheme.

The thematic scope of Eurovoc scope is defined by 21 thematic fields (identified by two-digit numbers and titles in words): e.g. *10 EUROPEAN COMMUNITIES*, divided into 127 microtheasuri (identified by four-digit numbers): *1011 COMMUNITY LAW*. The latest version contains 6,645 descriptors, 6,669 hierarchical and 3.636 associative relationships.

Step 2: map data vocabulary to SKOS. (See Table 1) Eurovoc has been developed using the standards ISO 2788 (for monolingual thesauri) and ISO 5964 (for multilingual theasuri). SKOS is compatible with the ISO 2788. In Eurovoc, descriptors are equivalent across the diversity of the 21 natural languages in which the thesaurus is encoded. The "descripteur_id" of descriptors is used to generate the `rdf:ID` of the instances of `skos:Concept`.

Hierarchical relationships between descriptors are expressed in Eurovoc using `BT` and `NT` relations. They are mapped to their equivalent elements in the SKOS data model: `skos:broader` and `skos:narrower` respectively. They are not defined as transitive properties, so there is no need of a transitive closure of the hierarchical relations in our conversion. In ISO 2788, polyhierarchies are not allowed, however certain descriptors in fields *72 GEOGRAPHY* and *76 INTERNATIONAL ORGANIZATIONS* have more than one broader term at the next higher level. SKOS data model allows this structure thus the correspondence will still be complete producing a correct SKOS. The `skos:related` property is used to map associative links between descriptors of hierarchical-trees (`RT` relationships). Both relationships (`skos:related` and `RT`) are symmetrical, not transitive and incompatible with the hierarchical relationship: if two conceptual resources are linked by a hierarchical relationship then there cannot be an associative relationship between them.

As we described above, Eurovoc descriptors are organized in two hierarchical levels: thematic fields and microthesauri. There is no direct translation into SKOS of these non-standard features of the thesaurus. Microthesauri and fields are also semantically related: each descriptor is assigned to a single microthesaurus, which belongs to a single thematic field. However, SKOS does not provide any property to create semantic links between concept schemes. On the one hand, we want to express that the descriptor *3062*-"executive body" is a top term of the microthesaurus *0436*-"executive power and public service" and the microthesaurus belongs to the thematic field *04*-"POLITICS". On the other hand, we want to represent complex internal structures of concept schemes. There are mainly two options here:

1. Interpreting microthesauri as `skos:ConceptScheme` instances. While fields are considered `skos:ConceptScheme` subclasses. Each Eurovoc microthesaurus is an instance of a single field class. Basically, there are two problems: i) it is not possible to link microthesauri and fields to the Eurovoc thesaurus RDF resource in a standard way and in addition ii) there is also no convincing way to relate top-hierarchy terms and fields. Triples of `skos:hasTopConcept` asserting direct links using is not DL compatible (fields are logical classes and descriptors instances) and formulas of the style $\texttt{Field} \equiv \forall \texttt{skos:hasTopConcept}.\ \{\texttt{Descriptor}_1\texttt{,..,}\texttt{Descriptor}_n\}$ introduce reasoning with nominals without evident benefits. This option has been discarded.
2. Interpreting microthesauri and fields as `skos:ConceptScheme` instances. This approach captures more accurately the semantics of microthesauri and fields resources. The `skos:hasTopConcept` and `skos:inScheme` properties can also be correctly used to assert local associations of descriptors to both microthesauri and fields without contradicting SKOS data model. However a new OWL object-property, `hasScheme`, has to be added (extending SKOS) to be able to assert hierarchical links between instances of `skos:ConceptScheme`. This property allows us to express *KOS* systems internal composition. The property has been defined as follows: its domain and range are `skos:ConceptScheme` and it is transitive from the whole to its parts: the Eurovoc "hasScheme" some fields, and every field "hasScheme" some microthesauri. We have chosen this option.

Step 3: convert the data. We have chosen XSL technology to convert the Eurovoc XML source into RDF/SKOS. In the first one, we built the basic skeleton of the document including all of microthesauri (concept schemes) and terms (concepts). Secondly, we decorated the definitions of concepts adding iteratively.

4.2 Adaptation of the CPV Vocabulary to SKOS

The main vocabulary of the CPV contains around 8,200 numerical codes, each one describing a single product term.

Step 1: analyze controlled vocabulary. The CPV consists of a main vocabulary, and a supplementary vocabulary that can be used for adding further qualitative information to the description of the subject of a contract. Only the main vocabulary will be considered in this analysis. This main vocabulary is composed of product terms identified by an alphanumeric code (an 8 digit code plus a check digit), see Table 2. The alphanumeric codes are shared between different versions of the CPV in each country, thus defining linguistic equivalence across languages of the European Union. The description "lemons" in the English version and the description "limones" in the Spanish one are both considered to be equivalent because both terms are identified with the same code: 01131210-9.

Step 2: map data vocabulary to SKOS. (See Table 3) Product terms have been considered `skos:Concept` and the code has been used to generate their `rdf:ID`. Specific-language literal description of product terms are mapped

Data Item	Feature	SKOS Element
$< Descripteur_i d$=X	Concept	skos:Concept with `rdf:ID`=X
$< Descriptor >$=Y in language=L	Preferred Term	`skos:prefLabel`= Y@xml:lang='L'
$< thesaurus_i d >$=X	Concept Scheme	`skos:ConceptScheme` with `rdf:ID`=X
Microthesaurus=Y in language=L	Concept Scheme Label	`skos:prefLabel`= Y@xml:lang='L'
$< domain_i d >$=X	Concept Scheme	`skos:ConceptScheme` with `rdf:ID`=X
Field=Y in language=L	Concept Scheme label	`skos:prefLabel`= Y@xml:lang='L'
Non-descriptor=Y in language=L, `UF` and `USE`	Equivalence term relation	`skos:altLabel`= Y@xml:lang='L'
CPV Non-descriptor=Y in language=L, `PERM`	Equivalence term relation	`skos:hiddenLabel`= Y@xml:lang='L'
`BT` Term	Broader Term	`skos:broader`
`NT` Term	Narower Term	`skos:narrower`
`RT` Term	Related Term	`skos:related`
`SN` Note in language=L	Scope Note	`skos:scopeNote`= Y@xml:lang='L'

Table 1. Mapping of Eurovoc Data Items to concept-based controlled vocabularies features and SKOS/OWL Classes and Properties.

to the SKOS property `skos:prefLabel` and the `xml:lang` attribute is used to identify the language.

Every product term is assigned to exactly one category. As we have considered product terms as `skos:Concept` instances, a decission had to be made about CPV categories formalization. They can not be straightforward mapped to any SKOS feature, thus we have introduced four new classes and we have declared them as `skos:Concept` subclasses using the RDF Schema property, `rdfs:subClassOf`. Product terms are also declared instances of their corresponding product category level. These statements can be realized parsing patterns in the product terms codes, see Table 2. There are 61 members of the top-level category, each one have been considered the top-term of a single vertical product sector comprised in a tree-structure.

The conversion of each subtree (product sector) has been made using a bottom-top transformation and the SKOS semantic property `skos:broader`. Parsing the alphanumeric code of each product term is sufficient to generate its direct parent in the hierarchy. E.g. Code "01112000-5" of the term "potatoes and dried vegetables", its broader term code is the string "01110000", which identifies the element "cereals and other crops" (the control digit can be easily generated then). However, if we try to generate the tree in the other direction (top-bottom), the transformation is computationally more complex. Given the previous code

Product Category	Identifier Code	Example
Division	XX000000-y	01000000-7
Group	XXX00000-y	01100000-8
Class	XXXX0000-y	01110000-1
Cat(L0)	XXXXX000-y	01112000-5
Cat(L1)	XXXXXX00-y	01112200-7
Cat(L2)	XXXXXXX0-y	01112210-0
Cat(L3)	XXXXXXXX-y	01112211-7

Table 2. Common Procurement Vocabulary Structure.

Data Item	Feature	SKOS Elements
Code=X	Concept	skos:Concept with `rdf:ID`=X
Product Term=Y in language=L	Preferred Term	`skos:prefLabel`= Y@xml:lang='L'
Product Category=C	Product Category	`rdf:ID`=C (subclass of `skos:Concept`)
Code=X for categorization	Product Category	X instance of Class C
Code=X for tree structure	Broader Term	`skos:broader`

Table 3. Mapping of CPV Data Items to concept-based controlled vocabularies features and SKOS/OWL Classes and Properties.

"01112000-5", 10 new alphanumeric codes (01112[0-9]00) will be generated as its possible descendants (`skos:narrover` in this case). Moreover another operation is necessary to check the existence of every generated code as the code generation process does not guarantee that the new codes are present in the CPV Vocabulary (e.g. the code "01112400-5" will be generated from "01112000-5", but it does not identify any product term in the CPV Vocabulary).

Therefore, a bottom-top algorithm has been choosen. The `skos:narrower` relations can be inferred using an OWL reasoner as `skos:narrower` is the inverse property of `skos:broader`.

Step 3: convert the data. In this case, we have used the RDF123 [4] tool and XSL to build the CPV in SKOS. Firstly, we created a new spreadsheet with the original MSExcel version of the CPV and all values for the mappings. Secondly, we loaded the new spreadsheet into RDF123 and created the mappings between cols and RDF nodes. Finally, as for Eurovoc, we applied several identity transformations with XSL to include all languages labels in the generated document.

After the conversions of Eurovoc thesaurus and CPV vocabulary, the execution of the complete method finished with the validation of both transformations with the *W3C SKOS Validator*[7]. Finally all the triples have been stored using the RDF repository, Sesame. The SKOS versions of the controlled vocabularies

[7] http://esw.w3.org/txamplesopic/SkosValidator

are publicly available on: Eurovoc (550, 707 triples)- http://idi.fundacionctic-.org/sparql_client/eurovoc and CPV (191, 472 triples)-http://idi.fundacionctic-.org/sparql_client/cpv. They can be queried using SPARQL.

5 Conclusions

We have defined the minimal set of features that a conversion method to RDF/OWL must cover to be applied for *KOS* systems in the EU context. Also, we have selected and tested SKOS as the common data model for the representation of these controlled vocabularies and we have carried out an existing method successfully.

The evaluation of the conversions must be checked from two different levels:

Correctness of SKOS conversions. Our approach has demonstrated to produce correct SKOS for the "greatest common divisor" of controlled vocabularies conversions: URIs generation, hierarchical relations and multilingual features.

Completeness of SKOS conversions. Our approach has demonstrated that just using SKOS does not produce complete conversions of the selected primary sources. Some specific properties and classes have to be added to preserve completely the original semantics. In the case of the Eurovoc, fields and microthesauri organize the set of descriptors. In this case, a new property, `hasScheme`, has to be added to be able to express the internal structure of the thesaurus. On the other hand, in the case of the CPV Vocabulary, every product term is associated with a product category. The output data model was extended with four new subclasses of `skos:Concept` to preserve the product categories hierarchy in the RDF/OWL conversion. Every product term has been interpreted as an instance of `skos:Concept` and as instance of its correspondent product category class.

Finally, we plan to apply the conversion method to other EU controlled vocabularies available in RAMON, the EUROSTAT metadata server. Our aim is to develop a common EU framework of *KOS* systems based on the SKOS vocabulary. A framework based on the same data model will promote semantic interoperability between European digital libraries, governmental agencies and other third-parties. The case studies of this paper, the Eurovoc Thesaurus and the CPV Vocabulary, show how the method can be put in practice.

In addition further research will allow us to convert to RDF/OWL the mappings between different classification systems. This is the case, for example, of product terms from the CPV Vocabularies and other PCSs. There exist some tables showing the correspondence between the CPV and the *Statistical Classification of Products by Activity* (CPA), the *General Industrial Classification of Economic Activities* within the European Communities (NACE Rev. 1) and the *Combined Nomenclature* (CN). The SKOS language provides a set of specific properties to represent semantic links between concepts belonging to different concept schemes. We will explore how this set of SKOS properties can indicate that a product term from the CPV is sufficiently similar with another product term from the NACE Classification to use them interchangeably in an information retrieval application.

We would like to report the results to the *Office for Official Publications of the European Communities* and the Linking Open Data[8] initiative, specially to the Riese Project[9] (*RDFizing and Interlinking the EuroStat Data Set Effort*).

Acknowledgements The work could not have been developed without the collaboration of the *Office for Official Publications of the European Communities* that granted us the license of the Eurovoc source, ref. 2008-COP-002, for academic purposes.

References

1. Richard Benjamins, Dieter Fensel, and Asunción Gómez Perez. Knowledge Management through Ontologies, 1998.
2. Oscar Corcho and Asunción Gómez Pérez. Integrating E-Commerce Standards and Initiatives in a Multi-Layered Ontology.
3. Dieter Fensel, Ying Ding, and Borys Omelayenko. Product Data Integration in B2B E-commerce. *IEEE-Intelligent E-Business*, pages 54–59, 2001.
4. Lushan Han, Tim Finin, Cynthia Parr, Joel Sachs, and Anupam Joshi. RDF123: A Mechanism to Transform Spreadsheets to RDF. Technical report, University of Maryland, Baltimore County, 1 2007.
5. Martin Hepp. A Methodology for Deriving OWL Ontologies from Products and Services Categorization Standards. (1):72–99, 2006.
6. Martin Hepp. The True Complexity of Product Representacion in the Semantic Web. In *Proceedings of the 14th European Conference on Information System (ECIS 006)*, Gothenburg, Sweden, 2006.
7. Martin Hepp. Possible Ontologies. *IEEE-Internet Computing*, (1):90–96, 2007.
8. Jörg Leukel, Volker Schmitz, and Frank-Dieter Dorloff. A Modeling Approach for Product Classification Systems. In *DEXA '02: Proceedings of the 13th International Workshop on Database and Expert Systems Applications*, pages 868–874, Washington, DC, USA, 2002. IEEE Computer Society.
9. Jörg Leukel, Volker Schmitz, and Frank-Dieter Dorloff. Exchange of Catalog Data in B2B Relationships - Analysis and Improvement. In *ICWI*, pages 403–410, 2002.
10. Brian Matthews, Alistair Miles, and Michael Wilson. Modelling Thesauri for the Semantic Web. In *Paper submission in to 3rd Workshop on Semantic Web and databases*, Humboldt-Universität Berlin, Germany, 2003.
11. Alistair Miles and Sean Bechhofer. SKOS Simple Knowledge Organisation Systems. W3C working draft, W3C, 2008. http://www.w3.org/TR/skos-primer/.
12. Alistair Miles and Dan Brickley. SKOS Simple Knowledge Organisation Systems. W3C working draft, W3C, 2005. http://www.w3.org/TR/2005/WD-swbp-skos-core-guide-20051102/.
13. Borys Omelayenko and Dieter Fensel. An Analysis of B2B Catalogue Integration Problems. In *Proceedings of the International Conference on Enterprise Information Systems (ICEIS-2001)*, Setúbal,Portugal, 2001.

[8] http://linkeddata.org/
[9] http://riese.joanneum.at/

14. D. Soergel, B. Lauser, A. Liang, F. Fisseha, J. Keizer, and S. Katz. Reengineering Thesauri for New Applications: the AGROVOC example. *Journal of Digital Information*, 2004.
15. M. van Assem, M. Menken, G. Schreiber, J. Wielemaker, and B. Wielinga. A Method for Converting Thesauri to RDF/OWL, 2004.
16. Mark van Assem, Véronique Malaisé, Alistair Miles, and Guus Schreiber. A Method to Convert Thesauri to SKOS. pages 95–109. 2006.