Exploring the Relationships between OWL and SKOS

Nor Azlinayati Abdul Manaf

Bio-Health Informatics Group (BHIG), School of Computer Science
The University of Manchester, United Kingdom
norazlinayati.abdulmanaf@postgrad.manchester.ac.uk

Abstract. Knowledge organization and representation schemes vary from simple structures such as classification schemes to more complex axiomatisations of a domain such as ontologies. These representations are conceptual and need to be specified through concrete representation using languages. Web Ontology Language (OWL) and the Simple Knowledge Organization System (SKOS) are amongst the languages used in the Semantic Web for knowledge representation. Each of these languages have different representation characteristics so that one language may have advantages over another in a particular situation. Thus an understanding of the patterns of use of these languages provides a framework that helps us explore the relationships between these languages, their characteristics and the problems they are used to solve. In this paper, we present a proposal for exploring the relationships between OWL and SKOS using this framework that aims to provide a transformation from one formalism to the other.

1 Background and Research Problem

Knowledge can be organized and represented in various forms ranging from simple conceptual structures with less semantics such as taxonomies to more complex conceptual structures and richer semantics like ontologies. Different knowledge organization and representation schemes have different characteristics. The organization and representation choice depends on the application problem requirements in hand. For example, some applications require the domain knowledge to be organized and represented in the form of ontologies, while other applications require the domain knowledge to be organized in the form of taxonomies or thesaurus-like representations. Sometimes, domain knowledge represented in a certain form for one application can be re-used by another application. It will often be the case, however, that one representational form of domain knowledge wil need to be transformed to another representational form, rather than being merely translated. Having some ways to transform from one representation scheme to another can help us to re-use existing knowledge resources.

There are a variety of languages for representation of conceptual models with different characteristics such as expressiveness, ease of use and computational complexity [1]. In this research, we study the characteristics of the Web

Ontology Language (OWL)[1] and the Simple Knowledge Organization System (SKOS)[2] that are among the languages used to support Semantic Web knowledge representation. Furthermore, we are also exploring the relationship between the two languages and patterns for their use together.

OWL and SKOS are intended for different purposes. OWL, for instance, is a Semantic Web knowledge representation language for expressing richly axiomatised logic based ontologies. In Computer Science, the term *ontology* has been adopted to refer to a set of precise descriptive statements about a domain that is described in terms of *individuals*, *classes* and *properties* [2]. OWL is intended to express complex conceptual structures that can be used to generate rich metadata and support inference tools. OWL has rich and strict semantics that enable automated reasoning and allows the explicit modeling and description of a domain.

SKOS, on the other hand, is a language designed for representation of thesauri, classification schemes, taxonomies, subject-headings, controlled structured vocabularies or any other knowledge organization systems. These are structures whose representation has weak semantics that are used for simple retrieval and navigation tasks. The basic element in SKOS is *concept* that refers to the unit of thought—ideas, meanings or objects—that exist in the mind as abstract entities and independent of the terms used for their labels. Each concept is given one or more *label* to refer to them in natural language, through *prefLabel* or *altLabel*. Besides, the terms are semantically linked to each other through hierarchical *broader (BT)/narrower(NT)* and associative *related (RT)* relations. The SKOS data model offers a standard, low-cost migration path for transferring existing knowledge organization systems to the Semantic Web technology context allowing better re-usibility, interoperability and sharing.

1.1 Research Problem

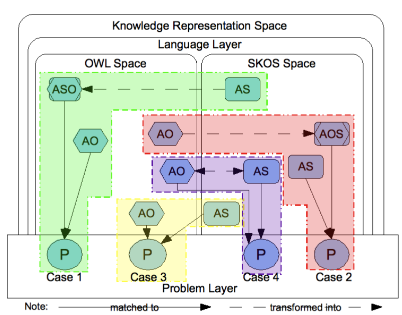

Fig. 1. Research problem illustration

[1] http://www.w3.org/2004/OWL/
[2] http://www.w3.org/2004/02/skos/

Figure 1 illustrates research area with which we are going to deal. The problem layer consists of sets of problems to be solved. For each problem, a set of *requirements* can be defined to determine the effectiveness of the solution to the problem. If all the problem requirements are satisfied, then the solution can be said to be effective even though it may not be an optimal solution. On top of the problem layer is the knowledge representation space which contains among others, a layer called 'Language Layer' that consists of various knowledge representation languages. The two languages (OWL and SKOS) are represented as two separate layers within the language layer, called 'OWL Space' and 'SKOS Space'.

Each representation language has some *characteristics*, which determines the usability in addressing an application problem *requirements* in a particular *domain of interest*, where the representation characteristics are used to *match* the problem requirements. Once we can determine this match, we can perform *mapping* of the problem into the representation space of choice. In general problems can be mapped into more than one knowledge representation space depending on whether the characteristics of the knowledge representation can match the problem requirements.

An *artefact* is a *solution* to the problem when the problem requirements are addressed by the knowledge representation characteristics. This is the result of addressing the problem using the knowledge representation language that in the end produces an artefact which is the solution to the problem represented within the selected knowledge representation space.

We have identified four cases of problems as illustrated in Figure 1. Case 1 represents a situation where for a problem, P, in the Problem Layer for a particular domain, the problem requirements can be matched by the OWL characteristics, resulting in the problem being mapped into OWL space that will produce an artefact in terms of an OWL ontology, AO. There, however, also exists an artefact in SKOS space in terms of a SKOS vocabulary, AS, that has been published to address the same domain problem. Therefore, the existing SKOS vocabulary, AS, can be re-used and thus needs to be transformed into the OWL space producing a new OWL ontology, ASO, to solve the same domain problem.

For example, a problem in the medical domain requires intentional representation in the form of an OWL ontology to represent the domain. The MeSH (Medical Subject Headings)[3] vocabulary, a controlled vocabulary-like representation, already exists and represents the same domain. Therefore, re-using this resource and transforming this representation into OWL an ontology can fulfill the problem requirement. In this case, our aim is that by using the existing SKOS vocabulary as a starting point and enriching the semantics between the concepts we can produce an OWL ontology that can fulfill the requirements of the application.

In Case 2, the situation is basically the converse of what we have in Case 1. A problem, P, in the Problem Layer for a particular domain can be mapped into SKOS space since the problem requirements are matched by SKOS char-

[3] http://www.nlm.nih.gov/mesh/

acteristics, producing a SKOS vocabulary, *AS*. There, however, also exists an artefact in OWL space in term of OWL ontology, *AO* representing the knowledge for the same domain. Therefore, the existing OWL ontology, *AO*, can be reused and needs to be transformed into the SKOS space producing a new SKOS vocabulary, *AOS*, to solve the same domain problem.

For example, a problem in the biomedical domain requires some representation for navigation and browsing purposes. This requirements are matched by SKOS characteristics and requires a SKOS vocabulary that represents the knowledge for this domain. The OBO (Open Biomedical Ontologies)[4] exist as representations of this domain. The OWL representation of these OBO artefacts do not have the appropriate structure for a simple navigation and browsing application. Therefore, transforming the definitions in the original ontology to produce a light-weight vocabulary is seen to be helpful as compared to the amount of effort needed to build the vocabulary from the beginning. In this case, we aim that by using the existing OWL ontology as a starting point and transforming this ontology will produce a SKOS vocabulary that can fulfill the requirements of the application.

Case 3 represents the situation where a problem, *P*, in the Problem Layer has different problem requirements that can be matched by both OWL and SKOS characteristics. Therefore, it is possible to map the problem into both language spaces. As for this case, further research and investigation is needed to determine if such a problem exists and which transformation procedures can help in reducing the amount of effort to fulfill the requirements of the system.

Deciding which representation space is more appropriate as a match for a problem can help in reducing the amount of effort to fulfill the requirements of the system. We predict that if the problem is mapped into one space, producing one artefact, then, we might transform that artefact to fulfill remaining problem requirements.

As for Case 4, a problem, *P*, in the Problem Layer for a particular problem domain requires both a representation as an OWL ontology and as a SKOS vocabulary for the domain due to a subset of problem requirements being matched by the characteristics of either languages. There exist solutions in both of the representation spaces for different aspects of the problem. In this case, no transformation is needed; instead, both artefacts may need to co-exist in an application with some alignment of the domain knowledge. Having some systematic procedures of using both representations at the same time in the same application is considered necessary.

Thus, the aim of this research is to explore the relationships between the two representations, OWL and SKOS, by looking at ways in which these two formalisms can be used to support each other, taking the advantages of each formalism, especially when we want to apply them to the above cases. In doing this, there are questions to be answered:

– *What are the principle ways to determine which knowledge representation space to choose for matching the problems from the Problem Layer?* – How

[4] http://www.obofoundry.org/

can we determine the characteristics of the knowledge representation, and decide that the characteristics of the knowledge representation match the requirements of a particular problem?

– *How can we enrich the SKOS vocabularies to match the OWL representation of the problem requirements in OWL space?*—What are the principle ways to transform SKOS vocabularies into the OWL space to match the problem requirements; how can we determine what sort of information that is sufficient and needed to be add to, retain or not to keep from the SKOS vocabulary?
– *How can we "reduce" the OWL ontologies to match the SKOS representation of the problem requirements in SKOS space?*—What are the principle ways to transform OWL ontologies into the SKOS space to match the problem requirements; how can we determine what sort of information that is sufficient and needed to be add to, retain or not to keep from the OWL ontology?

2 Related Work

Uschold and Grüninger [3], presented a comprehensive introduction to the field concerned with the design and use of ontologies. One of the aspects in which we are interested is their discussion on characterizing the ontologies outlining the nature of an ontology and what purposes they serve. Another work in characterizing the knowledge representation has been presented by Stevens et. al [1], which introduced the ontology-based representation within bioinformatics field. The authors focused on the type of knowledge held in ontology and provided several examples taken from bioinformatics and molecular biology to show the importance of using ontology in representing the knowledge of these fields.

In [4], a conceptual structure and transition procedure for transferring from an existing knowledge organization system into a semantically rich knowledge organization system is presented. The authors have chosen AGROVOC, an existing thesaurus by the Food and Agriculture Organization (FAO) of the United Nations, as a case study to explore the re-engineering of a traditional thesaurus into a well-defined ontology. However, the processes involved in the re-engineering process demand user intervention and intuition in deciding the appropriate and suitable relationships between the entities being described. Another work in deriving OWL ontologies from traditional knowledge organization systems has been presented in[5]. However, the authors focused only on three types of knowledge organization systems such as hierarchical classifications, thesauri and inconsistent taxonomies. A new methodology for automatically deriving RDF-S and OWL ontology from a traditional knowledge organization scheme has been presented with two e-business categorization standards Standardized Material and Service Classification (eCl@ss)[5] and United Nations Standard Products and Services Code (UNSPSC)[6] have been chosen to demonstrate the approach.

In [6], the authors explore how different knowledge representations may effect Web navigation with comparison between ontologically formal and semantically

[5] http://www.eclass-online.com/
[6] http://www.unspsc.org/

rich formalisms such as OWL ontology, and ontologically informal and semantically weak formalisms such as SKOS vocabulary. The authors also demonstrated how SKOS can be used to take advantage of the vast amount of existing ontological representation in Web navigation, especially in supporting the semantic linking of Web-based information and facilitating information retrieval.

3 Approach

In this research we study the characteristics of the two languages, OWL and SKOS, and explore the relationships between OWL and SKOS, with the aim of producing systematic methods and procedures of transforming from one formalism to the other. In order to achieve this aim, there are three main tasks that need to be carried out as discussed in the next three subsections. Each of these tasks is independent of each other and can be carried out in parallel or series.

3.1 Task 1: Problem requirements identification and language characterization

The initial task in this project is to identify the problem requirements and characterize the languages. This is done by analyzing the use cases for both OWL and SKOS languages [7–9] and also looking at their designs and implementations. Based on these use cases, we gathered and listed the requirements of each use case presented. These use cases also highlight the key features of the applications which can be used to characterize the languages. Furthermore, a few related works on characterizing the knowledge representation language as described in Section 2 are also referred to assist us in identifying the characteristics of the languages.
Evaluation: At the end of this task, we expect for each language, OWL and SKOS, to come out with a list of general problem requirements and set of characteristics which characterize the languages. A tool for assisting the choice of language or representation schemes between OWL and SKOS will be implemented. The gathered information will act as background knowledge which will be represented in OWL ontology. This tool will assist the user by asking a set of yes or no questions and based on the answers provided by the users, a query will be formulated upon the ontology and suitable representation scheme will be suggested.

3.2 Task 2: OWL to SKOS transformation procedures

For this particular task, we perform an initial work of studying the existing standards for thesaurus construction by referring to the works in [10–14]. The aim of this initial work is to fully understand the standard convention of thesaurus, since SKOS is developed based on these standard. This also will help with the process of transforming from OWL ontology to SKOS vocabulary. There are some aspects that we particularly concerned with. First, how does a term in thesaurus

relates to other terms. In this case, we performed an analysis on the use of semantic relations in thesaurus since they defined three types of semantic relations that can be used between concepts which represent equivalence relationships through the *USE* and *USED FOR (UF)* relations, hierarchical relationships through the *Broader Term (BT)* and *Narrower Term (NT)* relations, and associative relationships through the *Related Term (RT)* relation. There are general guidelines in determining which relationship to be used to relate between terms and we hypothesize this information can assist use in determining the transformation procedures to transform an OWL ontology into SKOS vocabulary. Furthermore, we are also looking at how the thesaurus deals with compound terms, terms that consist more that one terms connected by 'and' or 'or' conjunction. We hoped that by having some general guidelines on dealing with the compound terms may help in transforming the compound terms in OWL class descriptions into SKOS vocabulary, since SKOS language did not have standard way to handle the compound terms. Based on the gathered information, we try to define a standard procedures for transforming from OWL ontology to SKOS vocabulary the aim to produce a useful and faithful vocabulary that reflects the content of the original ontology.
Evaluation: The transformation procedures defined is then implemented to evaluate the performance. The implementation will be in term of tool for transforming from OWL ontology to SKOS vocabulary. This tool will be given to the domain experts to be tested where several OWL ontologies will be used to be transformed into SKOS vocabulary and a few criteria such as the faithfulness of the content and does the transformed vocabulary sustains the important information of the original ontology will be evaluated.

3.3 Task 3: SKOS to OWL transformation procedures

In this task, we are defining a standard principles to transform from SKOS vocabulary into OWL ontology with the aim to automate the process and have as minimum user intervention as possible. Two of the works defined in Section 2 have presented the transformation from the traditional knowledge organization systems into ontologies. However, one of them is only limited to a few types of knowledge organization systems only, while the other needs major users intervention and intuition. By combining, one or more of these works may result in a proper transformation procedures from SKOS vocabulary to OWL ontology.
Evaluation: The transformation procedures defined is then implemented to evaluate the performance. The implementation will be in term of tool for transforming from SKOS vocabulary to OWL ontology. This tool will be given to the domain experts to be tested where several SKOS vocabulary will be used to be transformed into OWL ontologies and a few criteria such as the faithfulness of the content and does the transformed ontology reflects the important information of the original vocabulary will be evaluated.

4 Current Status and Future Plan

The initial work for this PhD research started in October 2008 with identifying and defining the research problem. We are currently doing Task 1, the work on characterizing the languages, and Task 2 to define the procedure for transforming from OWL ontology to SKOS vocabulary, as described in the previous section. These two tasks are interleaved between one another. We are currently implementing the transformation procedures to evaluate the effectiveness and the correctness of the procedures[7]. Once this implementation is completed, the tool will be given to domain experts to be evaluated. Then, the feedback and comment from the experts based on the evaluation will be used to further refine the procedures with further iterations taking place. While performing this optimization of Task 2, we also will start the work on Task 3 followed by the implementation and evaluation of the implemented tool.

5 Contributions

1. Lists of problem requirements and language characteristics and tool for assisting in choosing the suitable knowledge representation language.
2. Systematic procedures and tool for transforming OWL ontologies into SKOS vocabularies.
3. Systematic procedures and tool for transforming SKOS vocabularies into OWL ontologies.

References

1. Stevens, R., Goble, C.A., Bechhofer, S.: Ontology-based knowledge representation for bioinformatics. Briefings in Bioinformatics **1**(4) (2000) 398–414
2. Horrocks, I.: Ontologies and the semantic web. Commun. ACM **51**(12) (2008) 58–67
3. Uschold, M., Grüninger, M.: Ontologies: principles, methods, and applications. Knowledge Engineering Review **11**(2) (1996) 93–155
4. Soergel, D., Lauser, B., Liang, A., Fisseha, F., Keizer, J., Katz, S.: Reengineering thesauri for new applications: The AGROVOC example. Journal of Digital Information **4**(4) (2004)
5. Hepp, M., Bruijn, J.D.: Gentax: A generic methodology for deriving owl and rdf-s ontologies from hierarchical classifications, thesauri, and inconsistent taxonomies. In: 4th European Semantic Web Conference (ESWC2007)
6. Jupp, S., Stevens, R., Bechhofer, S., Yesilada, Y., Kostkova, P.: Knowledge representation for web navigation. In: Semantic Web Applications and Tools for the Life Sciences (SWAT4LS 2008) Workshop. (2008)
7. Heflin, J.: Web ontology language (OWL) use cases and requirements. World Wide Web Consortium, Recommendation REC-webont-req-20040210 (February 2004)
8. Golbreich, C., Wallace, E.K.: OWL 2 web ontology language: New features and rationale. World Wide Web Consortium, Working Draft WD-owl2-new-features-20081202 (December 2008)

[7] Due to limited space, this discussion could not be presented here.

9. Isaac, A., Phipps, J., Rubin, D.: SKOS use cases and requirements. World Wide Web Consortium, Working Draft WD-skos-ucr-20070516 (May 2007)
10. BSI: BS 8723-2: Structured Vocabularies for Information Retreival – Guide: Thesauri. (2005)
11. ISO: ISO 2788: Documentation – Guidelines for the Establishment and Development of Monolingual Thesauri. (1986)
12. ANSI/NISO: ANSI/NISO Z39.19 - Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies. (2005)
13. Aitchison, J., Gilchrist, A., Bawden, D.: Thesaurus Construction and Use: A Practical Manual. Europa Publisher (2000)
14. Broughton, V.: Essential Thesaurus Construction. Facet Publishing (2006)