Products and Services Ontologies: A Methodology for Deriving OWL Ontologies from Industrial Categorization Standards

Martin Hepp, Digital Enterprise Research Institute (DERI), University of Innsbruck, Austria & Florida Gulf Coast University, Fort Myers, USA

ABSTRACT

Using Semantic Web technologies for e-business tasks, like product search or content integration, requires ontologies for products and services. Their manual creation is problematic due to (1) the high specificity, resulting in a large number of concepts, and (2) the need for timely ontology maintenance due to product innovation; and due to cost, since building such ontologies from scratch requires significant resources. At the same time, industrial categorization standards, like UNSPSC, eCl@ss, eOTD, or the RosettaNet Technical Dictionary, reflect some degree of consensus and contain a wealth of concept definitions plus a hierarchy. They can thus be valuable input for creating domain ontologies. However, the transformation of existing standards, originally developed for some purpose other than ontology engineering, into useful ontologies is not as straightforward as it appears. In this paper, (1) we argue that deriving products and services ontologies from industrial taxonomies is more feasible than manual ontology engineering; (2) show that the representation of the original semantics of the input standard, especially the taxonomic relationship, is an important modeling decision that determines the usefulness of the resulting ontology; (3) illustrate the problem by analyzing existing ontologies derived from UNSPSC and eCl@ss; (4) present a methodology for creating ontologies in OWL based on the reuse of existing standards; and (5) demonstrate this approach by transforming eCl@ss 5.1 into a practically useful products and services ontology.

Keywords: business ontologies; e-business; e-commerce; e-procurement; eCl@ss; ontology engineering; reuse; taxonomies; thesauri; UNSPSC

INTRODUCTION

Products and services categorization standards (PSCS), like the UNSPSC, eCl@ss, eOTD, or the RosettaNet Technical Dictionary (RNTD) form a valuable set of concepts from the products and services domain and reflect some degree of consensus. They are thus a promising foundation for the creation of product ontologies for the use in the Semantic Web, but
are not by themselves fully usable ontologies.

There already exist examples of products and services ontologies derived from PSCS, e.g., the RDF-S versions of UNSPSC by Klein (2002) and the DAML+OIL versions of UNSPSC by McGuinness (2001) and by Klein (2002). However, those ontologies do not properly reflect the specific semantics of the underlying standards. There also exists an early prototype of an RDF representation of eCl@ss 4.1 created by Bizer and Wolk (2003), which is based on an OWL ontology for representing taxonomies that itself is OWL Full, because it applies object properties to classes. All three ontologies are one-time captures of the reused standards, while the standards themselves have undergone significant change in the meantime (Hepp, Leukel, & Schmitz, 2005a, 2005b). These shortcomings in combination limit the usefulness of available ontologies.

Ontology engineering has been the subject of research for a long time, and multiple methodologies for the creation of ontologies have been proposed. However, the script-based (i.e., automated) reuse of concepts from hierarchically ordered standards, which were not created with the rigor of knowledge representation in mind, creates different requirements and demands novel approaches, mainly because we cannot partition the concept space into concepts at our own will. In other words, we must automatically capture as much semantics as possible out of the existing standards and have only limited control over the actual definition of the concepts. In general, we have to take the concepts as they are, because the size of those standards (between 20,000 and almost 60,000 classes and between 3,000 and 20,000 properties) renders manual steps in the process of importing the concepts unfeasible. On the other hand, we have to carefully analyze the constraints and dependencies resulting from the implicit assumptions of the standards creators.

Additionally, the intension of the reused concepts will often be influenced by our interpretation of the original standard. When transforming an informal categorization scheme into a formal ontology using a standard ontology language, we often narrow down the semantics, which includes some degree of freedom and thus several modelling decisions. Sometimes we might even want to have controlled changes in meaning between the original standard and the resulting ontology in order to create more useful ontologies. This, however, must be done with great care and we then must make respective decisions transparent to the ontology user.

Related Work

Related work to this paper can be classified into the following main groups:

Ontology engineering methodologies, implicitly or explicitly focusing on the manual creation of new ontologies based on knowledge engineering principles. A comprehensive discussion of all approaches in this field is beyond the scope of this paper; for an overview see, e.g., Fernández-López and Gómez-Pérez (2002) and de Bruijn (2003).

Analysis of the meaning of taxonomic relationships, especially the fundamental work of Brachman (1983). This yielded the insight that there are multiple types of taxonomic relationships, which should be represented distinctively.

Contributions regarding electronic commerce ontologies mainly from the
Related Content

A Broader View on Context Models towards Supporting Business Process Agility
www.igi-global.com/chapter/broader-view-context-models-towards/60069?camid=4v1a

Data Linking for the Semantic Web
www.igi-global.com/chapter/data-linking-semantic-web/76176?camid=4v1a

Interaction Models and Relevance Feedback in Image Retrieval
www.igi-global.com/chapter/interaction-models-relevance-feedback-image/28926?camid=4v1a

Adaptive Hybrid Semantic Selection of SAWSDL Services with SAWSDL-MX2
www.igi-global.com/articleadaptive-hybrid-semantic-selection-sawsdl/50497?camid=4v1a