A Pattern-Based Method for Re-Engineering Non-Ontological Resources into Ontologies

Boris Villazón-Terrazas, Universidad Politécnica de Madrid, Spain
Mari Carmen Suárez-Figueroa, Universidad Politécnica de Madrid, Spain
Asunción Gómez-Pérez, Universidad Politécnica de Madrid, Spain

ABSTRACT
To speed up the ontology development process, ontology developers are reusing all available ontological and non-ontological resources, such as classification schemes, thesauri, lexicons, and so forth, that have already reached some consensus. Non-ontological resources are highly heterogeneous in their data model and storage system (or implementation). The reuse of these non-ontological resources involves their re-engineering into ontologies. This paper presents a method for re-engineering non-ontological resources into ontologies. The method is based on so-called re-engineering patterns, which define a procedure that transforms the non-ontological resource components into ontology representational primitives using WordNet for making explicit the relations among the non-ontological resource terms. The paper also provides the description of NOR2O, a software library that implements the transformations suggested by the patterns. Finally, it depicts an evaluation of the method, patterns, and software library proposed.

Keywords: Non-Ontological Resources, Ontologies, Ontology Development, Patterns, Re-Engineering

INTRODUCTION

Research on Ontology Engineering methodologies has provided methods and techniques for developing ontologies from scratch. Well-recognized methodological approaches such as METHONTOLOGY (Gómez-Pérez, Fernández-López & Corcho, 2003), On-To-Knowledge (Staab, Schnurr, Studer & Sure, 2001), and DILIGENT (Pinto, Tempich & Staab, 2004) give guidelines that help researchers to develop ontologies. However, researchers face an important limitation: no guidelines are provided for building ontologies by re-engineering existing knowledge resources widely used by a particular community.

During the last decade, specific methods, techniques and tools were proposed for building ontologies from existing knowledge resources. First, ontology learning methods and tools have been proposed to extract relevant concepts and relations from structured, semi-structured, and non-structured resources (Gómez-Pérez & Man-
zano-Macho, 2004; Maedche & Staab, 2001) in order to form a single ontology. One important constraint to these methods and tools is that they propose ad-hoc solutions to transforming such resources, mainly texts, into ontologies. Hepp et al. (Hepp, 2006; Hepp & Brujin, 2007; Hepp, 2007) stated that employing methods and techniques when ontologizing non-ontological resources to the level of ontologies is key for the success of semantic technology for two main reasons: (1) if the use of semantic technologies for real-world data integration challenges is required, it is possible to refer to the original conceptual elements, and (2) for many domains, the existing category systems, XML schemas, and normative entity identifiers are the most efficient resources for engineering ontologies.

The ontologization of non-ontological resources has led to the design of several specific methods, techniques and tools (Hepp & Brujin, 2007; Hyvönen, Viljanen, Tuominen & Seppälä, 2008; Gangemi, Guarino, Masolo & Oltramari, 2003; García & Celma, 2005). These are mainly specific to a particular resource type, or to a particular resource implementation. Thus, everytime ontology engineers face a new resource type or implementation, they develop ad-hoc solutions for transforming such resource into a single ontology.

In parallel, and within the context of the NeOn project\(^1\), a novel scenario-based methodology for building ontology networks\(^2\) has been proposed: the NeOn Methodology (Suárez-Figueroa, 2010; Gómez-Pérez & Suárez-Figueroa, 2009). One of these novel scenarios is Building Ontology Networks by Reusing and Re-engineering Non-Ontological Resources. As opposed to custom-building silos of single ontologies from scratch, this new scenario emphasizes the re-engineering of knowledge resources for building ontologies that are connected with other ontologies in the ontology network.

The motivation of this paper lies in this scenario of the NeOn Methodology and the use of re-engineering patterns to transform the non-ontological resources components into ontology representational primitives. Along this paper we will try to demonstrate that the use of re-engineering patterns for transforming non-ontological resources into ontologies has several advantages: (1) they embody expertise about how to guide a re-engineering process, (2) they improve the efficiency of the re-engineering process, and (3) they make the transformation process easier for ontology engineers.

In this paper we present first our proposed categorization of non-ontological resources. Then, we provide a framework for comparing methods for re-engineering non-ontological resources into ontologies, and the conclusions drawn on the comparative study. After that, we analyze the role of patterns in software engineering and ontology engineering with particular emphasis on re-engineering patterns. Then, we present our pattern-based method for re-engineering non-ontological resources into ontologies. Then, we explain NOR\(_2\)O, a software library that performs the transformation automatically. Next, we present an evaluation of the method, patterns and software library. Finally we draw the conclusions and provide future lines of work.

### NON-ONTOLOGICAL RESOURCES

**Non-Ontological Resources** (NORs) are knowledge resources whose semantics have not yet been formalized by an ontology (García-Silva, Gómez-Pérez, Suárez-Figueroa & Villazón-Terrazas, 2008). There is a great number of NORs that embody knowledge about some particular domains and that represent some degree of consensus. These resources are present in the form of textual corpora, classification scheme, thesauri, lexicons, etc. NORs have usually implicit semantics that allows interpreting the knowledge they contain. Regardless of whether the semantics is explicit or not, the main problem is that the semantics of NORs is not always formalized, and this lack of formalization prevents them from being used as ontologies. Using non-ontological resources that already have reached a consensus for build-
Enhancing Visibility in EPCIS Governing Agri-Food Supply Chains via Linked Pedigrees
[www.igi-global.com/article/enhancing-visibility-in-epcis-governing-agri-food-supply-chains-via-linked-pedigrees/124924?camid=4v1a](www.igi-global.com/article/enhancing-visibility-in-epcis-governing-agri-food-supply-chains-via-linked-pedigrees/124924?camid=4v1a)

A New Similarity Measure for Automatic Construction of the Unknown Word Lexical Dictionary
[www.igi-global.com/article/new-similarity-measure-automatic-construction/4110?camid=4v1a](www.igi-global.com/article/new-similarity-measure-automatic-construction/4110?camid=4v1a)