

MULTIMODAL COMMUNICATION IN THE 21ST CENTURY: PROFESSIONAL AND ACADEMIC CHALLENGES. 33rd Conference of the Spanish Association of Applied Linguistics (AESLA), XXXIII AESLA CONFERENCE, 16-18 April 2015, Madrid, Spain

Lexicalizing Ontologies: The Issues Behind The Labels

Guadalupe Aguado-de-Cea*, Elena Montiel-Ponsoda, María Poveda-Villalón, Olga Ximena Giraldo-Pasmin

Ontology-Engineering Group, Universidad Politécnica de Madrid, Campus de Montegancedo, Boadilla del Monte 28660, Madrid, España

Abstract

In information science, ontologies are used to capture knowledge about some domain of interest, by formally naming and defining the types, properties and interrelationships of the concepts that describe that domain. They are the building blocks of the Linked Data initiative in which datasets of related domains are linked to each other, and also to more general datasets, resulting in a huge space of interconnected data. It is precisely in this linking step that the linguistic descriptions used to name or label ontology entities (i.e., concepts, properties, attributes) become undeniably significant. This has also an impact in the subsequent process of ontology localization, thus turning it into a critical process. Based on our experience in the lexicalization and localization of several well adopted ontologies (FOAF, GoodRelations, the Organization ontology, among others) from English into Spanish, we propose a preliminary set of guidelines with a twofold goal. First, to guide users in the process of assigning labels and descriptions to ontology entities; second, to help terminologists and translators in the translation of these specific resources by providing them with coherent, user-friendly examples of how to apply the above mentioned guidelines

© 2015 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Peer-review under responsibility of the Scientific Committee of the XXXIII AESLA CONFERENCE

Keywords: ontology lexicalization; ontology localization; ontology entity labels

* Corresponding author. Tel.: +34913367415; fax: +34913524819.
E-mail address: lupe@fi.upm.es

1. Introduction

In information science, ontologies are used to capture knowledge about some domain of interest, by formally naming and defining the types, properties and attributes of the concepts that describe that domain. Ontology entities are generally represented using machine-readable languages such as RDFs - Resource Data Framework Schema (Brickley & Guha, 2014), SKOS - Simple Knowledge Organization System (Miles & Bechhofer, 2008), or OWL - Web Ontology Language (McGuinness & van Harmelen, 2004). These languages provide mechanisms to associate natural language descriptions to ontology entities, a process termed “ontology lexicalization” (Cimiano et al., 2014). For instance, in Fig. 1 we can see an excerpt from the GoodRelations ontology, an ontology about e-commerce, in which the concept `gr:BusinessEntity` is labelled “Business entity” by means of the `rdfs:label` annotation, and is also provided with a definition by means of the `rdfs:comment` annotation, both in the English language, as specified by the XML annotation, `xml:lang="en"`. (A table with all the prefixes related to the ontologies mentioned in the paper is included in the Appendix.)

```
<owl:Class
  rdf:about="http://purl.org/goodrelations/v1#BusinessEntityType">
  <rdfs:label xml:lang="en">Business entity</rdfs:label>
  <rdfs:comment xml:lang="en">An instance of this class represents the
  legal agent making (or seeking) a particular offering. This can be a
  legal body or a person. A business entity has at least a primary
  mailing address and contact details. (...)
  </rdfs:comment>
</owl:Class>
```

Fig. 1. Excerpt from the GoodRelations ontology

In the same way, other ontology languages or vocabularies have specific annotations for these natural language descriptions (e.g., `skos:prefLabel` in SKOS or `lemon:LexicalEntry`, in the *lemon* -Lexical Model for Ontologies- vocabulary, accessible at <http://lemon-model.net/>). Such descriptions not only facilitate the understanding of the knowledge represented by the ontology, but they also contribute to a quick adoption of the ontology by application developers.

Recently, ontologies have also become the building blocks of the Linked Data initiative (Bizer et al., 2009). This initiative is based on the publication, sharing and interlinking of data, in the same way as we publish and link documents on the Web. The vocabularies used to publish data sets are represented as ontologies, and they allow data sets of related domains to be linked to each other, and also to more general data sets, resulting in a huge space of interconnected data.

It is precisely in the linking step that the natural language descriptions used to term or label ontology entities become undeniably significant. Moreover, those data sets may not be in the same language, nor the ontologies used to structure and represent them. Consequently, some vocabularies have to be translated or localized to ensure a smooth linking process. The process of adapting an ontology to a particular linguistic and cultural community is known as “ontology localization” (Espinoza et al. 2012, Gracia et al. 2012).

Therefore, be it for the purposes of contributing to the linking process in the context of Linked Data or for the adoption by users from different linguistic and cultural communities, ontology lexicalization and ontology localization have become a priority in the current ontology engineering research. In this sense, there is a lack of guidelines or recommendations for the performance of both tasks from a linguistic perspective that needs to be addressed to meet the challenges mentioned.

Based on our experience in the lexicalization of ontologies (AEMET, <http://aemet.linkeddata.es/ontology/>, El Viajero, <http://webenemasuno.linkeddata.es/ontology/OPMO>) and the localization of several well adopted ontologies (FOAF, <http://xmlns.com/foaf/spec/>, GoodRelations, <http://www.heppnetz.de/projects/goodrelations/>, and the ORG Organization ontology, <http://www.w3.org/TR/vocab-org/>, amongst others) from English into Spanish, as well as on previous literature in the field (Noy & McGuinness, 2001; Schober et al., 2009; Montiel-Ponsoda et al., 2011), we propose a preliminary set of principled reflections with a twofold goal. Firstly, to guide users in the label assignment task within the ontology lexicalization process. Secondly, to help general users, terminologists and translators in the translation/localization of these specific resources by providing them with coherent, user-friendly

Download English Version:

<https://daneshyari.com/en/article/1109625>

Download Persian Version:

<https://daneshyari.com/article/1109625>

[Daneshyari.com](https://daneshyari.com)