Ontology Localization

Mauricio Espinoza
Ontology Engineering Group,
Dpto. de Inteligencia Artificial,
Facultad de Informática, UPM
28660 Boadilla del Monte
Madrid, Spain
jespinoza@fi.upm.es

Elena Montiel-Ponsoda
Ontology Engineering Group,
Dpto. de Inteligencia Artificial,
Facultad de Informática, UPM
28660 Boadilla del Monte
Madrid, Spain
emontiel@fi.upm.es

Asunción Gómez-Pérez
Ontology Engineering Group,
Dpto. de Inteligencia Artificial,
Facultad de Informática, UPM
28660 Boadilla del Monte
Madrid, Spain
asun@fi.upm.es

ABSTRACT

International organizations (e.g., FAO[1], WHO[2], etc.) are increasingly expressing the need for multilingual ontologies for different purposes, e.g., ontology-based multilingual machine translation, multilingual information retrieval. However, most of the ontologies built so far have mainly English or another natural language as basis. Since multilingual ontology building is a very expensive and time-consuming undertaking, we propose methods for guiding users in the localization of ontologies, and provide tools for supporting the process. The main contributions of this paper are: i) the description of a generic Ontology Localization Activity and a methodology for guiding in the localization of ontologies; ii) the description of a tool built according to the guidelines proposed for an automatic localization of ontologies; and iii) a set of experiments used to evaluate the methodological and technological aspects of the Ontology Localization Activity.

Categories and Subject Descriptors

I.2.4 [**Knowledge Representation Formalisms and Methods**]: [semantic networks]

General Terms

Algorithms, Documentation

Keywords

Ontology Localization, Multilingual Ontologies

[1]www.fao.org
[2]www.who.int

K-CAP'09, September 1–4, 2009, Redondo Beach, California, USA.

1. INTRODUCTION

In the context of the Semantic Web, a great effort has been done in the construction of ontologies. However, although access to top-quality ontologies (e.g., Galen[3], CYC[4], or AKT[5]) is in some cases free and unlimited for users all around the world, most of these ontologies are available only in English. Due to this language barrier, non-native English users often encounter problems when trying to access and use ontological knowledge in other languages. Moreover, the development of some ontology-based systems built for multilingual applications such as information retrieval [8], question answering [14] or knowledge management [15] has increased the need for multilingual ontologies. Despite this situation, the number of multilingual ontologies available on the Web is still insignificant.

Conscious of this problem, in the NeOn project[6] we have identified Ontology Localization as one of the activities within the ontology network development process to support the construction of multilingual ontologies. In this context, Ontology Localization has been defined as "the adaptation of an ontology to a particular language and culture" [16]. As in Software Localization, in which localization "involves taking a product and making it linguistically and culturally appropriate to the target locale (country/region and language) where it will be used and sold" [7], in Ontology Engineering, this adaptation is also needed if we aim to express the ontology labels[7] represented in the ontology in a natural language different from the one in which the ontology has been conceived.

The rest of this paper is structured as follows. Section 2 describes the ontology localization problems. Section 3 presents different strategies for representing multilingual information in ontologies and for solving translation problems. Then, with the aim of guiding users in

[3]www.co-ode.org/galen/
[4]http://www.opencyc.org/downloads
[5]http://www.aktors.org/publications/ontology/
[6]http://www.neon-project.org/web-content/
[7]Name of an ontology term.

the development of multilingual ontologies, we propose some guidelines (section 4), and show how localization is performed automatically with LabelTranslator [6]. Section 5 presents a set of experiments used to evaluate the methodological and technological aspects of the Localization Activity. Finally, the main conclusions are discussed in Section 6.

2. CHARACTERIZATION OF THE ONTOLOGY LOCALIZATION PROBLEM

When dealing with the localization on an ontology, several dimensions have to be taken into account:

Translation Problems. Since cultures parcel the world in a different way, when translating ontologies we come across various situations:

i. Existence of an exact equivalent. This is specific of highly specialized technical and engineering fields such as mechanics, in which there is a direct/complete equivalence among the terms in different languages referring to a certain object or process. For example, "hotel" in English is translated as "hotel" in Spanish.

ii. Existence of several context-dependent equivalents. When one term in a language can be translated by several equivalents in another language, and the user has to choose the most suitable one depending on the context of the ontology, word connotations, the sociolinguistic register in which the ontology will be used, etc. For example, the English term "accommodation" can be translated into Spanish as "alojamiento" (Spain), or "hospedaje" (South America). Each translation represents certain nuances of the concept, so it is necessary to find the most approximate or suitable equivalent for the term "accommodation".

iii. Existence of a conceptualization mismatch. When a certain culture categorizes reality with a degree of granularity that does not correspond to the granularity degree of the other culture. This may result in a lexical gap or a lack of direct equivalence in the target language. For example, in French there is a difference between big rivers that flow into the sea, which are called "fleuves", and rivers that flow into other rivers, "rivières". In most topography ontologies in English this distinction is not made.

Management Problems. Whereas the translation process of ontology labels 'per se' implies certain difficulties, the maintenance and updating of translated ontology labels throughout the ontology life cycle also requires special attention. The main difficulty is to identify policies for managing changes in ontology terms and their translated labels. Up to now, none of the works on managing ontology changes [13, 17] deals with changes on ontology elements with multilingual information. Thus, several situations could happen:

i. An ontology term is added, then the ontology label should be translated to all supported languages.

ii. An ontology term disappears, then all its translations should be removed.

iii. An ontology term is renamed, then all multilingual labels should be reviewed.

In all cases this process can be performed using two operation modes: instant mode or batch mode. The former is executed when changes are applied, while the latter can be executed at the end of the user's session, for instance.

Multilinguality Representation Problems. As a result of the Localization Activity, the ontology obtained has labels represented in different natural languages. According to the state-of-the-art, we can identify three main models of representing multilinguality in an ontology:

i. Inclusion of multilingual information in the ontology by means of the rdfs:label *and* rdfs:comment *properties* (Model 1). This has been the approach most used by the Ontology Engineering Community until now since it allows multilingual labels to be associated to ontology terms.

ii. Creation of one conceptualization per culture and language involved, and establishment of mappings among the different conceptualizations (Model 2). Each conceptualization will reliably reflect the categorization of the reality that each culture makes. However, the effort required in the development of the various conceptualizations and the linkage among conceptualizations is by no means trivial. A representative example of this approach is the well-known EuroWordNet[8] lexicon.

iii. Association of external multilingual information to the ontology (Model 3). Different models have been proposed to associate linguistic data to ontologies: a) the Linguistic Information Repository (LIR) [11], specially designed to account for cultural and linguistic differences among languages; b) LingInfo [1] and, c) LexOnto [2], which is focused on the linguistic enrichment of ontologies from a morphosyntactic viewpoint. The main advantage of this third approach is that it does not require the effort of creating additional conceptualizations and can take advantage of the ontologies already available on the Web to create multilingual ontologies.

The choice among the three models will be mainly determined by two factors: 1) the type of domain of knowledge represented by the ontology, and 2) the amount of linguistic information required for the final application. Regarding the first factor, we mainly consider here two

[8]www.illc.uva.nl/EuroWordNet/

types of domains of knowledge: domains whose categorization usually finds consensus among different cultures, and culturally-dependant domains, i.e., domains whose categorization is normally influenced by a certain culture. The selection of the representation model and how translation problems are solved will be explained in more detail in the following section.

3. MULTILINGUAL ONTOLOGIES

We can agree that ontologies are consensual models that explicitly represent and organize knowledge, but experience shows that certain conceptualizations are prone to reflect cultural particularities. When dealing with the localization issue, the type of domain being categorized becomes crucial. If the conceptualization is shared among all the cultures implied in the Localization Activity (e.g., an ontology of the human genome), the localization will only affect the terminological layer, that is, the labels that name ontology terms. On the contrary, regarding culturally-dependant domains (e.g., the judicature) in which categorizations tend to reflect the particularities of a certain culture, the localization may affect the conceptual layer, i.e. the conceptualization. Optionally, we may be able to account for cultural differences at the terminological layer. According to the domain typology and quantity of linguistic information required, we propose the following **multilinguality representation models**:

i. If the conceptualization represents a consensual domain, there are two options: we can opt for the inclusion of multilingual information in the ontology (Model 1), or for the association of an external model with the ontology (Model 3). The decision between these two options will depend on the linguistic needs of the final application. If morphosyntactic data is needed for the purpose of Information Retrieval or Information Extraction, for example, the most suitable option will be the association of an external model such as LingInfo, which enriches the ontology with a great amount of morphosyntactic information.

ii. If the conceptualization represents a culturally-dependant domain, and conceptualization mismatches among different cultures are faced, we have again two options: the creation of one conceptualization per language and culture involved (Model 2), or the association of an external model that permits to account for those cultural divergences at the terminological layer (Model 3). In this sense, from the models presented so far for the association of multilingual information to ontologies, the LIR is the only one that permits explaining cultural divergences among languages at the terminological layer.

In the following, we show how we aim to solve the **Translation problems** introduced in section 2 by means of the LIR model.

Exact and Context-dependant Equivalents. In Figure 1(a), we represent a consensual conceptualization (upper side) between the English and the Spanish cultures. In this case, there exists an exact equivalent (hotel) and a context-dependent equivalent (accommodation) between English and Spanish. The country (e.g., Spain) in which the ontology is to be used determines which is the most suitable translation of "accommodation" into Spanish, which will be translated as "alojamiento", instead of "hospedaje", that would be rather used in South America. The linguistic information associated to the ontology (lower side) is stored in an external model (Model 3). In these examples, we have used a highly simplified example of the LIR model, in which two of its classes (LexicalEntry and Lexicalization) are linked to the ontology concepts. In this case, the LIR contains information in two natural languages: the original language of the conceptualization (left side) and the target language resulting from the Localization Activity (right side).Such a model allows enriching a conceptualization with as much linguistic information as we wish in one or several natural languages. If only multilingual labels were required, we could have opted for Model 1.

Conceptualization mismatches. Figures 1(b), 1(c), and 1(d) represent ontologies of culturally-dependant domains (upper side) whose categorizations do not always reflect the needs of all cultures. In Figure 1(b) we represent a conceptualization of Spanish rice dishes, from which a subclass would be *paella*. Assuming that this concept does not exist in the English culture, two localization strategies would be possible. One would consist in lending the word from the Spanish language and adding an intensional description of the concept in English. A different strategy would consist in going to the translation of the hypernym (rice dish) and reuse it. These two strategies pursue different objectives. The first one has the objective of "documenting" the target culture about a certain knowledge parcel in the original culture. The second has the objective of looking for the most approximate equivalent in the target culture (for more on this see [12] on *documental* vs. *instrumental* translation).

Figures 1(c) and 1(d) reflect conceptualizations accepted in the English culture that do not reflect usual categorizations in the French culture. In Figure 1(c), when trying to localize *ape* and *monkey* into French we discover that both are translated as *singe* because, for some reason, the usual categorization accepted in the French culture makes no difference at this level if the animal is small and has a tale (monkey) or big and has no tail (ape). On the contrary, when referring to rivers (see figure 1(d)), English generally talks about rivers without noticing if they end up flowing into the sea or into another river if there is no specific need for making this distinction explicit.

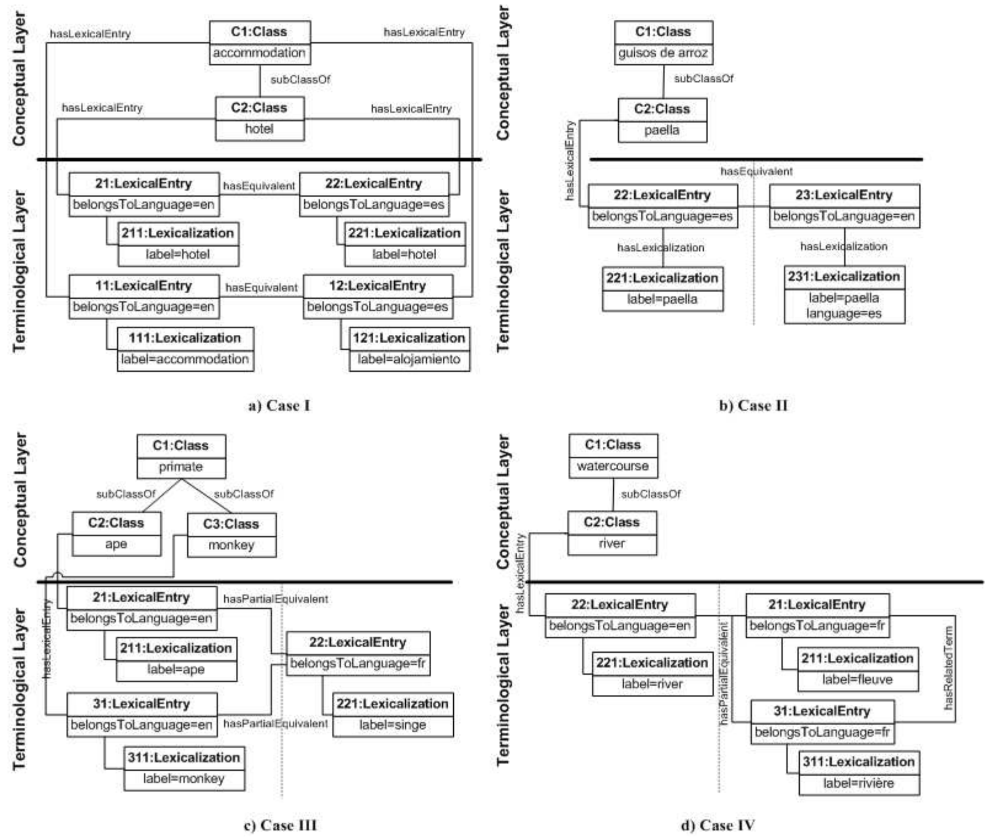


Figure 1: Modelling Multilinguality in Ontologies.

Once we have proposed different strategies for solving translations and multilinguality representation problems, we present preliminary guidelines for the localization of ontologies whenever a conceptualization is available.

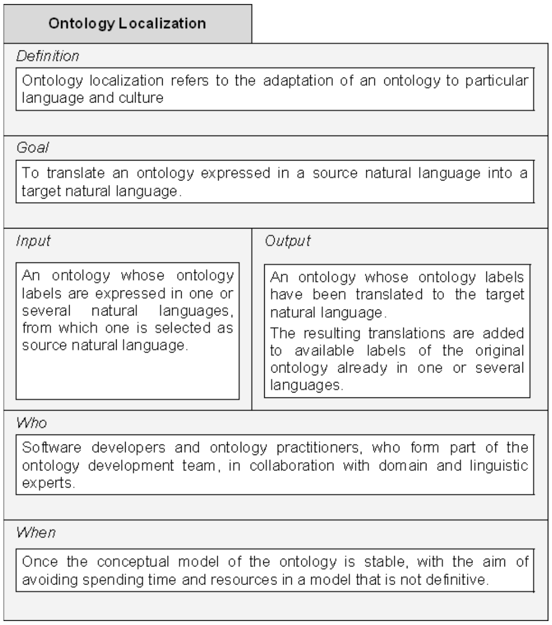
4. GUIDELINES PROPOSED FOR ONTOLOGY LOCALIZATION

In this section, we explain the guidelines we propose to help ontology practitioners in the Localization Activity, which are inspired in the human translation process. Taking into account the methodological work included in [3], we provide a filling card for the Ontology Localization Activity in Figure 2(a), in which the activity is explained. Then, we describe the tasks for performing Ontology Localization in a prescriptive manner. Figure 2(b) shows how the activity should be carried out, with inputs, outputs and actors involved.

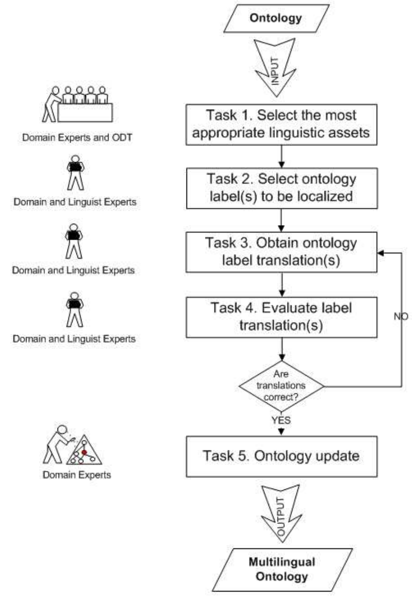
Task 1. Select the most appropriate linguistic assets. The goal of this activity is to select the most appropriate linguistic assets that help in the Localization Activity. User and domain experts carry out this activity taking as input the ontology to be localized. The activity output is a set of linguistic assets that can help to reduce the cost, improve the quality, and increase the consistency of the Localization Activity. The selection of a specific resource is performed manually and looking for linguistic assets with the following characteristics i) *Consensus.* Used resources should contain multilingual terminology already agreed on by the community, thus the effort and time spent in finding out the right translation labels for the ontology terms will decrease considerably; ii) *Broad coverage.* The resources should cover translation information from general to specific domain labels. Besides, since each resource supports different feature and language sets, the selected resources should cover all target languages for current and possible future ontology localization projects; and iii) *High precision.* The resources used for ontology localization should be able to identify the lexicographical differences existing between different natural languages.

Task 2. Select ontology label(s) to be localized. The goal of this task is to select the ontology label(s) to be localized. Domain experts and the ontology development team perform this task taking as input an ontology whose ontology labels, expressed in a source natural language, need to be localized to a target language. The task output is a set of ontology labels and its context[9]. The context describes the meaning of a spe-

[9]In NLP, context is the environment in which a word is used, and context, viz. word usage, provides the only information we have for figuring out the meaning of a new or a



(a) Ontology Localization Filling Card.



(b) Proposed tasks for the Ontology Localization.

Figure 2: Ontology Localization Activity

cific label in the ontology and consist of a small excerpt of ontology labels around the ontology label itself (e.g., direct hypernyms labels, hyponyms labels). For example, in the case of the ontology label "primate" shown in Figure 1(c), the context will be the labels "ape" and "monkey".

Since there are no methodological guidelines for guiding the selection of the ontology labels, we believe that the user is the one who has to choose the space of the candidates to be localized. At this stage, the user may choose to localize the complete ontology or only certain labels.

Task 3. Obtain ontology label translation(s). The goal of this task is to obtain the most appropriate translation in the target language for each ontology label. Domain and linguist experts carry out this task taking as input the ontology label(s) to be localized. Different techniques can be used to perform this task: i) *cross-language term extraction* to discover translation equivalents; ii) *word sense discovery* to discover the possible senses or definitions of the translations; and iii) *word sense disambiguation* to select the most appropriate translation of each ontology label(s).The task output is a ranked set of labels in the target language for each ontology label(s).

Task 4. Evaluate label translation(s). The goal of this task is to evaluate the label translations in the target language. Domain and linguist experts carry out this activity taking as input the labels in the target language. The output of this task is a set of labels with its corresponding evaluation. Different linguistic criteria can be used for the evaluation of the label translations. We propose two levels of evaluation criteria and for each level a set of tests, which can be automated as far as possible.

Semantic fidelity evaluation. The aim is to control that the label translation be conceptually equivalent to the ontology label in the source language. A way of evaluating the semantic fidelity is to perform a backward translation test, which provides a quality-control step demonstrating that the quality of the translation is such that the same meaning is derived when the translation is moved back into the source language.

Stylistic evaluation. The aim is to control the clarity and syntax of the target language, which depends on the style of the source language and on the peculiarities of the individual idiolect. Some tests that could be performed here are whether a translation is totally blank, or if there are repeated words in the translation.

Task 5. Ontology update. The goal of this task is to update the ontology with the label translations obtained for each localized label. Domain experts carry out this task taking as input the selected label translations. The activity output is an ontology enriched with labels in the target language associated to each localized term. Depending on the amount of linguistic information required to enrich the ontology we can use the models proposed in section 2 for storing the label

polysemous word.

translations and additional linguistic information associated to each ontology term, if required.

Ontology Localization with LabelTranslator

LabelTranslator [6] has been designed to support ontology localization by automating the main tasks described above and with the aim of reducing human intervention. The tool has been implemented in the ontology editor NeOn Toolkit as a plugin[10], and in its current version, it can localize ontologies in English, German and Spanish. In the following, we briefly describe how the tasks are performed by our system, and the techniques and tools used in each task.

Task 1. The linguistic assets that the current version of the LabelTranslator NeOn plugin accesses to perform ontology localization are multilingual linguistic resources (EuroWordNet, Wiktionary, or IATE) and translation web services (GoogleTranslate, BabelFish, etc.). The addition of further domain specific resources as AGROVOC[11], MESH[12], etc., is foreseen in the near future.

Task 2. Once an ontology has been created or imported in NeOn Toolkit, LabelTranslator allows users and domain experts to manually/automatically sort out the ontology labels that should undergo localization. For each ontology label, LabelTranslator retrieves its *local context* (e.g. direct hypernyms, hyponyms), which is used by the system to describe the meaning of the ontology label in the ontology. Consider, for example, the word 'plant', which depending on the context can be translated into Spanish as 'planta' in the sense of "living organism" or 'fábrica' in the sense of "industrial plant".

Task 3. To obtain the most appropriate translation for each ontology label, LabelTranslator uses the following techniques in the indicated order:

-Using *cross-language term extraction*, the system obtains equivalent translations for all selected labels by accessing the linguistic assets listed in Task 1.

-The system uses *sense discovery technique* to retrieve a list of semantic senses for each translated label, querying different third-party knowledge pools: Watson[13], which indexes many ontologies available on the Web, and remote lexical resources as EuroWordNet.

-The senses of each context label are also discovered by querying the resources identified in the previous step.

-The system uses *word sense disambiguation* to sort the label translations according to their context. LabelTranslator carries out this task with the senses of each translated label and the senses of the context labels. At this stage, domain and linguist experts may decide to choose the most appropriate translation from the ones in the ranking. In default of this, the system will consider the one in the highest position.

Task 4. In the current version of the tool, the evaluation of the translations of each ontology term is manually performed by the user. We are working to automatize this task.

Task 5. The ontology is updated with the resulting label translations, which are stored in the LIR model, an external model adopted by the LabelTranslator NeOn plugin for organizing and relating linguistic information within the same language and across languages (see section 1).

5. EXPERIMENTAL EVALUATION

In this section we describe a set of experiments that were conducted with the objective of evaluating the methodological and technological aspects of the localization activity. First, we present the usability evaluation of the methodological guidelines using a manual translation with independence of the utilized software (section 5.1). In section 5.2 we describe the experiments used to evaluate some aspects related to the translation ranking techniques, where the task is to select the most appropriate translation for ontology labels. Section 5.3 deals with the study used to assess the usability of the LabelTranslator system for carrying out the ontology localization activity.

5.1 Methodological Evaluation

This example refers to the manual localization of the Pest control ontology that uses the guidelines proposed in this paper. Basically, the ontology localization activity is carried out by the Information Management specialist of the FAO with the contribution of domain and linguist experts. The description of the sample localization is not intended to be exhaustive, it just summarizes the most important points. A detailed and complete ontology localization sample is described in [4].

The objective of this FAO example is to localize the Pest control ontology from English to French and Italian. The input ontology is a module of the AGROVOC Concept Server[14] containing English terms identifying one or more concepts. Next, we show the tasks that have been performed to localize the Pest control ontology to the selected languages.

Task 1. Select the most appropriate linguistic assets. In general, FAO experts make use of several

[10] http://www.neon-toolkit.org/wiki/index.php/LabelTranslator
[11] http://www.fao.org/aims/agintro.htm
[12] http://www.nlm.nih.gov/mesh/
[13] http://watson.kmi.open.ac.uk/WatsonWUI/
[14] http://naist.cpe.ku.ac.th/agrovoc/

computer aided translation tools to perform the localization of their ontologies. FAO experts used mainly FAOTERM, the institutional multilingual terminological system[15]. However, this glossary only covers the six official languages of the FAO: English, Spanish, French, Arabic, Chinese, and Russian. In addition to FAOTERM, another important asset used in the localization activity was the Google define functionality. Finally, for this use case, FAO experts used some cataloguing systems such as AGRIS[16] or FAODOC[17].

Task 2. Select ontology label(s) to be localized. From the Pest control ontology, they manually extracted the ontology labels to be localized. Due to space limitations, let us suppose that the selected ontology label is "pest control" and their related terms to be localized into French and Italian. The related terms are postharvest sparring, product protection, postharvest control, and postharvest treatment.

Task 3. Obtain ontology label translation(s). For each ontology label, experts used a manual process to discover translation equivalents, possible senses or definitions for each translation, and to disambiguate translation senses.

Cross-language term extraction. Translations in French were obtained using FAOTERM, but for Italian they resorted to specialized dictionaries, online or printed. Some candidate translations obtained for the label "pest control" were "lutte contre les ravageurs", "lutte phytosanitaire",...(French), and "nebulizzazione postraccolta", "difesa dei prodotti immagazzinati", ...(Italian).

Word Sense Discovery. In order to discover the definitions, domain experts used the Google define functionality. For example, the search [define:pest control] shows a list of definitions for "pest control" gathered from various online sources. Additionally, the experts checked the use of the term "pest control" and possibly translated documents that make use of its translations in the desired languages using AGRIS.

Word Sense Disambiguation. With the information obtained in the previous steps, they used a manual disambiguation process to rank the translations of each ontology term. For example, for the sample term "pest control", the most appropriate translations obtained for each target language were lutte contre les ravageurs (french), and nebulizzazione postraccolta (italian).

Task 4. Evaluate label translation(s). Based on the proposed guidelines, they identified the following situations:

[15]http://www.fao.org/faoterm/
[16]http://www.fao.org/agris/search/search.do
[17]http://www.fao.org/Documents/

Semantic fidelity evaluation. In order to evaluate the semantic fidelity of the translation, they implemented the "Backward Translation" criteria. In many cases, the translation did not match exactly the original meaning, but in a deeper analysis, taking in consideration the context and the topics (agriculture), they verified that the semantic fidelity was covered 100% while the syntactic fidelity was not ensured.

Stylistic evaluation. In this case, they checked elements such as acronyms, the use of multiple words, capitalizations, etc. For the use case mentioned, no particular problems arised but the use of the parenthesis, for example, the English term "Product protection (stored)" appear to be translated in Italian as "Difesa dei prodotti immagazzinati", and they proposed the label "Difesa dei prodotti (immagazzinati)". In other cases instead, the translations proposed were consistent.

Task 5. Ontology update. In this task domain experts stored the translated labels in a external module linked to the AGROVOC Concept Server.

5.2 Performance Evaluation

To evaluate the quality of translation of the LabelTranslator system we conducted in March 2008 a preliminary experiment [6] involving PhD students. The main goal of the experiment was to evaluate the translation ranking techniques used by the system to select the most appropriate translation for each ontology label. This was done by comparing the translations provided by an expert (gold standard) with the translations provided by the ranking algorithm used in LabelTranslator. The ontology corpus used for the evaluation was selected from the set of KnowledgeWeb[18] ontologies used to manage EU projects. The experimental results showed that our system suggested the correct translation 72% of the times. Also, the values of recall obtained suggested that a high percentage of correct translations were part of the final translations shown to the user. One of the main limitations was the low quality of the translations of compound labels (e.g., railroad_transportation). We implemented some improvements to the algorithm as i) a recursive function that attempts to match the bi/tri-tokens of a compound label with the lexical templates[19] stored in the database, or ii) a method that learns new lexical templates from the translations supplied by the user.

To test the improvements implemented, a new experiment was performed in the "Artificial Intelligence (AI)" Master course at the Facultad de Informática (Universidad Politécnica de Madrid) with 17 Master students. We decided to use a questionnaire that allowed collecting the assessments of the students about the capacity

[18]http://knowledgeweb.semanticweb.org/
[19]The notion of lexical template refers to text correlations found between a pair of languages.

of the translation algorithm in providing correct translations according to the context. As a general conclusion of the results obtained, we can mention that 33% of the students identified the level of correctness of the translations greater than 80%. The rest of students believed that the translations obtained had a level of correctness greater than 90%. Basically, the main strength provided by this experiment was the improvement in the quality of translation of compound labels. However, some weaknesses were detected: i) the misuse or omission of the definite article; ii) the incorrect translation of acronyms; and iii) the erroneous identification of the part of speech (POS) of single words or compound labels. The final aspect is very important, because we use POS tagging as a first mechanism of disambiguation to discard candidate translations. We are now working to solve this problems. More details about this experiment in [5].

5.3 Usability Evaluation

To asses the usability of the LabelTranslator system we conducted an experiment following the Software Usability Measurement Inventory (SUMI) method [10]. The SUMI questionnaire includes 50 items for which the user selects one of three responses ("agree", "don't know", "disagree"). The questionnaire is designed to measure the affect, efficiency, learnability, helpfulness, and control of a software product [9]. The experiment involved 10 participants, most of whom were PhD students with a good command on ontology engineering. The experimenters met with all participants for 10 minutes to explain the purpose of the evaluation session and present the methodology of SUMI evaluation. Then, the participants had 20 minutes to translate an ontology using the guidelines and LabelTranslator, and 10 minutes to fill in the SUMI questionnaire for user-interaction satisfaction. During these two phases of the experiment users were not allowed to ask questions to the evaluators.

The most important findings of the experiment are related to the high level of learnability shown by LabelTranslator, especially in the case of a novice user. There was only one evidence about the need of making minor modifications in the LabelTranslator user interface to improve affect and efficiency with better navigation and informative functions. Additional details about user perception with respect to the goals of each SUMI dimension can be found in [5].

6. CONCLUSION

The main objective of the research presented in this paper is the definition of the Ontology Localization Activity and the identification of its main challenges and open problems. To the best of our knowledge, the study presented here is the first attempt to offer guidelines for the localization of ontologies. In a first stage, we analyze the implications of localizing an ontology, and the different strategies that can be followed to solve translation problems and to store the multilingual information resulting from the Localization Activity. Then, we present the preliminary guidelines for the localization of ontologies whenever a conceptualization is available. The guidelines are used in the design of LabelTranslator, a tool for automatically localizing ontologies. Finally, we have described a set of experiments used to evaluate the methodological and technological aspects of the Ontology Localization Activity.

7. ACKNOWLEDGEMENTS

This work is supported by the European Commission's Sixth Framework Program under the project name: Lyfecycle support for networked ontologies (NeOn) (FP6-027595), and the National Project "GeoBuddies" (TSI-2007-65677C02). We would like to kindly thanks Rosario Plaza Arteche for her review of the document.

8. REFERENCES

[1] P. Buitelaar, M. Sintek, and M. Kiesel. A multi-lingual/multimedia lexicon model for ontologies. In *Proc. ESWC'06, Budva, Montenegro.*, 2007.

[2] P. Cimiano, P. Haase, M. Herold, M. Mantel, and P. Buitelar. Lexonto: A model for ontology lexicons for ontology-based nlp. In *OntoLex'07, South Corea*, 2007.

[3] Deliverable 5.4.1. (2008). In *http://www.neon-project.org/*.

[4] Deliverable 5.4.2. (2009). In *http://www.neon-project.org/*.

[5] Deliverable 5.6.2. (2009). In *http://www.neon-project.org/*.

[6] M. Espinoza, A. Gómez-Pérez, and E. Mena. Enriching an ontology with multilingual information. In *Proc. of 5th European Semantic Web Conference (ESWC'08), Tenerife, (Spain), June 2008.*

[7] B. Esselink. *A practical Guide to Localization*. John Benjamins, 2000.

[8] J. Guyot, S. Radhouani, and G. Falquet. Ontology-based multilingual information retrieval. In *CLEF*, 2005.

[9] J. Dumas and J. Redish. A Practical Guide to Usability Testing. Exeter, UK: Intellect, 1993.

[10] J. Kirakowski and M. Corbett. Sumi: The software usability measurement inventory. *British Journal of Educational Technology*, 1993.

[11] E. Montiel-Ponsoda, G. Aguado, A. Gómez-Pérez, and W. Peters. Modelling multilinguality in ontologies. In *Coling 2008: Companion volume - Posters and Demonstrations, Manchester, UK*, 2008.

[12] C. Nord. Translating as a purposeful activity. Functionalist approaches explained. UK: St. Jerome, 1997.

[13] R. Palma, P. Haase, O. Corcho, A. Gómez-Pérez, and Q. Ji. An editorial workflow approach for collaborative ontology development. In *3rd Asian Semantic Web Conference. ASWC 08. Bangkok, Thailand*, 2008.

[14] M. Pazienza, A. Stellato, F. Zanzotto, L. Henriksen, and P. Paggio. Ontology mapping to support ontology based question answering. In *Proceedings of the Meaning 05 Workshop*, 2005.

[15] A. Segev and A. Gal. Enhancing portability with multilingual ontology-based knowledge management. *Decis. Support Syst.*, 2008.

[16] M. C. Suárez-Figueroa and A. Gómez-Pérez. First attempt towards a standard glossary of ontology engineering terminology. In *Proc. of 8th International Conference on Terminology and Knowledge Engineering (TKE'08)*, 2008.

[17] T. Tudorache, N. Noy, S. Tu, and M. Musen. Supporting collaborative ontology development in protege. In *International Semantic Web Conference*, 2008.