Multilingual Ontology Mapping in Practice: A Support System for Domain Experts

Mauro Dragoni(✉)

FBK–IRST, Trento, Italy
dragoni@fbk.eu

Abstract. Ontology mapping is a task aiming to align semantic resources in order to foster the re-use of information and to enable the knowledge and data expressed in the matched ontologies to interoperate. When ontology mapping algorithms are applied in practice, a manual refinement is furtherly necessary for validating the correctness of the resulted resource, especially, as it happens in real-world cases, when a gold standard cannot be exploited for assessing the generated mappings. In this paper, we present a suggestion-based mapping system, integrated as a component of a knowledge management platform, implementing an information retrieval-based (IR-based) approach for generating and validating, by experts, mappings between ontologies. The proposed platform has been evaluated quantitatively (i.e. effectiveness of suggestions, reduction of the user effort, etc.) and qualitatively (i.e. usability) on two use cases: Organic.Lingua and PRESTO, respectively an EU-funded and regional-funded projects. The results demonstrated the effectiveness and usability of the proposed platform in a real-world environment.

1 Introduction

Ontology mapping is a task aiming to enable interoperability between semantic resources by defining relationships that can be used for various tasks, such as ontology merging, data translation, query answering, or navigation on the web of data. This problem has been widely explored in the last years [1] from the research point of view, and several approaches in the literature have been proposed. However, issues that have to be addressed from the research point of view are amplified if ontologies have to be mapped manually. Indeed, the definition of such mappings is not a trivial task in real-world due to the different structures, contents, and sizes of the ontologies that have to be mapped. Knowledge experts have to deeply analyze the ontologies before creating new mappings, as well as, automatic systems need to take into account a lot of aspects in order to suggest suitable mappings between the concepts of the ontologies. This makes the scenario more complicated by considering, also, that experts knowledge is generally focused on particular domain/s and their level of expertise in knowledge management is, sometimes, inadequate for providing and validating mappings between ontology manually.

M. Arenas et al. (Eds.): ISWC 2015, Part II, LNCS 9367, pp. 169–185, 2015.
DOI: 10.1007/978-3-319-25010-6_10

For these reasons, it is necessary to provide an instrument that is able to support experts in the mapping activity by providing facilities avoiding them to spend a lot of time for exploring the structure of ontologies, but, at the same time, allow to easily define and validate mappings in a reasonable time.

In this paper, we present a suggestion-based mapping system aiming to improve the mapping activity experience in real-world scenarios. Our system is split in two components: (i) a back-end module implementing an IR-based techniques built with the scope of suggesting sets of candidate mappings, and (ii) a set of user interface facilities that have been integrated in a knowledge management tool for supporting experts in the mapping activity. A further peculiarity of the proposed approach is that it works in a multilingual environment either to support experts in a multi-language collaborative working environment, as well as, to exploit the multilingual information contained in the ontologies for the computation of suggestions. The effectiveness and usability of the proposed platform has been evaluated on two funded projects, described in Section 2, from a quantitative and qualitative points of view.

The paper is structure as follows. Section 2 describes the two use cases where the proposed platform has been adopted and validated. In Section 3, we present the back-end approach used for computing the mapping suggestions; while, in Section 4, we describe the facilities for supporting experts in the mapping task that have been implemented in the used knowledge management tool. Section 5 discusses the evaluation procedure, shows the obtained results, and presents general insights we inferred from this experience. In Section 6, we present a general overview of the literature about ontology mapping and knowledge management tools. Finally, Section 7 concludes the paper.

2 Use Cases

Below, we present the two projects used as test benches for the implemented platform and, for each of them, we introduce the ontologies used in the evaluation.

2.1 PRESTO Project

The objective of the PRESTO (Plausible Representation of Emergency Scenarios for Training Operations) research project is the creation of a system for the customization of serious games scenarios based on virtual reality. The advantage of this system, compared to the state of the art, resides in the richness and the ease of defining the behavior of artificial characters in simulated scenarios, and on the execution engines able to manage cognitive behaviors, actions, and perceptions within a virtual reality environment.

Within the project, ontologies have been used for modeling virtual reality items, actions, behaviors, etc. and the mapping task was needed for defining mappings between the core virtual reality ontology (namely "PRESTO Ontology") containing general 3d objects and classification schemata describing the set of items contained in 3d-specific libraries. In this use case, we focused the evaluation of the system on the mapping between the "PRESTO Ontology" and the XVR-library[1] classification schema representing the ontological view of the 3d items modeled in the XVR library. Concepts defined in both ontologies are modeled in English and Italian.

2.2 Organic.Lingua Project

Organic.Lingua (http://www.organic-lingua.eu) is an EU-funded project that aims at providing automated multilingual services and tools facilitating the discovery, retrieval, exploitation, and extension of digital educational content related to Organic Agriculture and AgroEcology. More in concrete, the project aims at providing, on top of a web portal, cross-lingual facility services enabling users to (i) find resources in languages different from the ones in which the query has been formulated and/or the resource described (e.g., providing services for the cross-lingual retrieval); (ii) manage meta-data information for resources in different languages (e.g., offering automated meta-data translation services); and (iii) contribute to evolve the content (e.g., providing services supporting the users in the content generation).

The accomplishment of these objectives is reached in the Organic.Lingua project by means of two components: on the one hand, a web portal offering software components and linguistic resources able to provide multilingual services and, on the other hand, a conceptual model (formalized in the "Organic.Lingua ontology") used for managing information associated with the resources provided to the final users and shared with other components deployed on the Organic.Lingua platform. In a nutshell, the usage of the Organic.Lingua ontology is twofold:

- Resource annotation: each time a content provider inserts a resource in the repository, the resource is annotated with one or more concepts extracted from the ontology.
- Resource retrieval: when web users perform queries on the system, the ontology is used, by the back-end information retrieval system, to perform advanced searches based on semantic techniques.

Concerning the specific ontology mapping task, one of the expected activity was the definition of the mappings between the Organic.Lingua ontology (described by using sixteen languages) and other ontologies related to the agricultural domain, in particular Agrovoc (28 languages) and Eurovoc (24 languages) with the aim of improving either the annotation and retrieval capabilities of the entire platform.

3 The Back-end IR-Based Approach For Multilingual Ontology Mapping

In this Section, we present the first component of the platform, i.e., the back-end system used for storing the structured representation of ontology concepts and for

[1] http://futureshield.com/xvr-esemble.shtml

generating and manipulating such representations for suggesting candidate mappings.

Before we present how information are structured, we want to introduce a formalization of what a *mapping* (or "match" or "alignment") is. A popular definition is to consider a "mapping" as a set of *correspondences* between entities asserting that a certain relation holds between these two entities. Formally, given two ontologies O_1 and O_2, a match M between O_1 and O_2 is a 5-tuple: $\langle id, e_1, e_2, R, c \rangle$ such that id is a unique identifier of the match, e_1 and e_2 are entities of O_1 and O_2 respectively, R is the matching relation between e_1 and e_2 (for example, equivalence ($\equiv$), more general ($\supseteq$), disjointness($\perp$)), and c is a confidence measure, typically in the $[0, 1]$ range.

While, the general definition of "mapping" includes different kind of relationships, in this work we focused on realizing a system aiming to suggest only equivalence relationships between concepts.

Below, we described which ontological information have been exploited, how the indexes used for storing ontological information are constructed, and how the messages for requesting and sending the suggestions about candidate mappings are composed.

Exploited Information. The representation of each concept in the system is based on the exploitation of textual information (i.e. the set of "labels") associated with each concept described in an ontology and its "context".

For explaining what we mean as "context", let's consider the following example. Figures 1 and 2 show excerpts of two ontologies about business processes representing two concepts that might be considered good candidates for defining a new mapping: "Activity" and "Executable Activity". As "context" of a concept C, we mean the set of concepts connected with C, where C is parent or child of another concept, or where C occurs in the domain or in the range of an object property. In particular, in this work we considered only the first degree of relationships of each concept C.

While, with the term "label", we mean a string identifying the concept associated with its language tag (i.e. "concept_label@lang_code") [2].

The usage of multilinguality is one of the key-aspects of the concept representation. When a label (independently by the language) is chosen by experts during the creation of an ontology, they implicitly inject in their choice the knowledge about the equivalence of meanings between different translations of each label.

Even if the definition of mapping between ontologies is typically performed on compatible domains ([2]), the use of approaches exploiting label-based techniques, leads to problems with effectiveness. First of all, is that different concepts, especially in case of ontologies representing different domains, could have similar labels without being similar in their meaning. For instance, given two ontologies O_1 and O_2, where O_1 describes the fruit domain O_2 the fishery one; if we consider the Italian concept "pesca", we can notice that such a label is polysemic because it can denote

[2] The format used for the structured representation of thesauri within the system follows the SKOS model http://www.w3.org/TR/skos-reference/

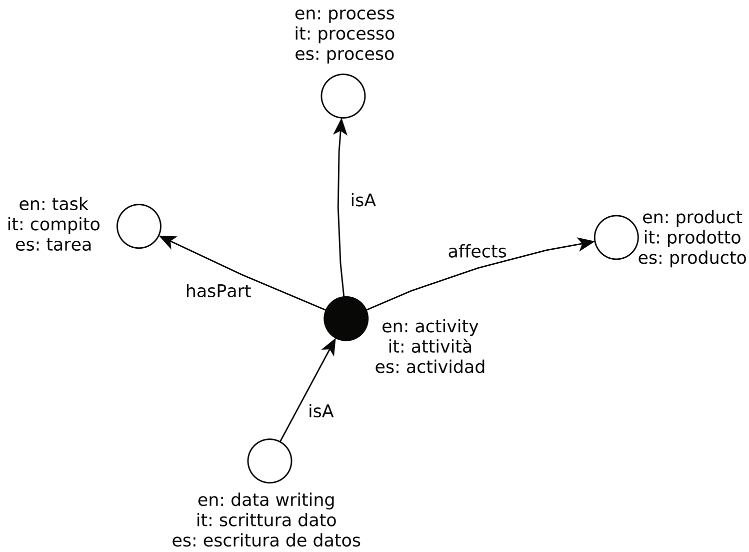

Fig. 1. Example of piece of Ontology 1 showing the context of the concept "Activity".

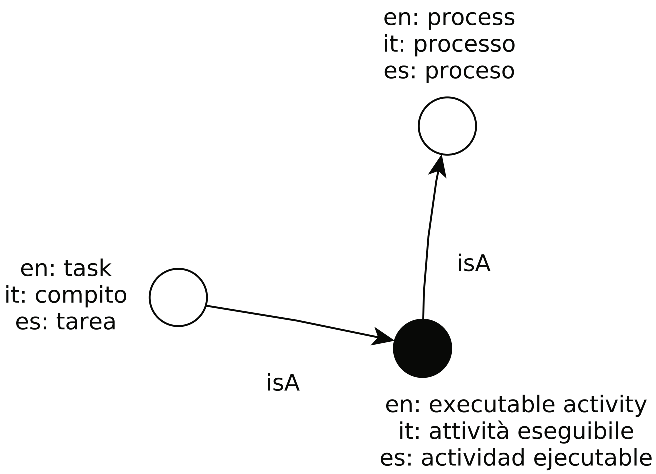

Fig. 2. Example of piece of Ontology 2 showing the context of the concept "Executable Activity".

a fruit (the peach) in the ontology O_1 and the activity of fishing in the ontology O_2. If O_1 and O_2 are multilingual, the difference between the two labels can be caught by considering their definition expressed in different languages. Particularly, for the example cited, the English labels "peach" and "fishing". This aspect helps to solve ambiguities during the computation of new suggestions.

Index Construction. By starting from the information described above, for each concept defined in the ontology, we created a set of triples of the kind "relation-label-language", and, in case of synonymy, we may have more triples for the same language. Such labels are tokenized, stemmed [3], and normalized by using natural language process libraries [4].

By taking into account the example shown in Figure 1, information related to the context of the concept "Activity" are the following[5]:

```
[self-prefLabel] "activity"@en
[self-prefLabel] "attività"@it
[self-prefLabel] "actividad"@es
[parent-prefLabel] "process"@en
[parent-prefLabel] "processo"@it
[parent-prefLabel] "proceso"@es
[child-prefLabel] "data writing"@en
[child-prefLabel] "scrittura dato"@it
[child-prefLabel] "escritura de datos"@es
[hasPart-prefLabel] "task"@en
[hasPart-prefLabel] "compito"@it
[hasPart-prefLabel] "tarea"@es
[affects-prefLabel] "product"@en
[affects-prefLabel] "prodotto"@it
[affects-prefLabel] "producto"@es
```

where with the relationship identifier "self", we indicate labels associated to the current concept description. Subsequently, such information are converted in the structured representation "field_name : value" as shown as follows:

```
self-label-en : activity
self-label-it : attività
self-label-es : actividad
parent-label-en : process
parent-label-it : processo
parent-label-es : proceso
child-label-en : data writing
child-label-it : scrittura dato
child-label-es : escritura de datos
hasPart-label-en : task
hasPart-label-it : compito
hasPart-label-es : tarea
affects-label-en : product
affects-label-it : prodotto
affects-label-es : producto
context-label-en : activity process data writing task product
context-label-it : attività processo scrittura dato compito prodotto
context-label-es : actividad proceso escritura de datos tarea producto
```

The fields "relationship-label-XX" are automatically generated during the creation of the record and they contain all label of the concept context. Such fields are useful when, as it will be shown in the explanation of how mappings are suggested, two concepts within the ontologies are linked with different relations.

---

[3] The list of languages supported by the used stemming algorithm are Italian, French, German, English, Greek, Spanish, Portuguese, Polish, Dutch, Norwegian, Hungarian, Swedish, Latvian, Turkish, Czech, Russian, Romanian, Bulgarian, Arabic, Hindi, Chinese, Japanese, Danish, Finnish, Armenian, Indonesian, Thai.

[4] The text processors included in the Lucene (http://lucene.apache.org) library have been used. In case of unavailability of libraries for a particular language, the original label is indexed as it is without being processed.

[5] SKOS notation is used.

After the construction of the structured representation for all concepts, they are indexed as documents by using the inverted index algorithm [3]. This operation is performed for each ontology stored in the platform with the result of having, for each ontology, a dedicated index. Indexes are stored in an architecture based on Apache Solr[6].

Matches Definition. Once the indexes are created, the suggestion of candidate mappings is done by performing queries using information extracted from the concept used as starting point for the mapping operation. Such a query is created by building a structured representation compliant with the one described above by using information defined in the concept used as starting point.

Therefore, similarly to the creation of the indexed records, by taking into account the example shown in Figure 2, the query for the concept "Executable Activity" is built as follows, firstly we extract all information related to the concept:

```
[self-prefLabel] "executable activity"@en
[self-prefLabel] "attività eseguibile"@it
[self-prefLabel] "actividad ejecutable"@es
[child-prefLabel] "task"@en
[child-prefLabel] "compito"@it
[child-prefLabel] "tarea"@es
[parent-prefLabel] "process"@en
[parent-prefLabel] "processo"@it
[parent-prefLabel] "proceso"@es
```

and, then, we create query:

```
proc(self-label-en:"executable activity" OR self-label-it:"attività eseguibile" OR
   self-label-es:"actividad ejecutable" OR child-label-en:task OR
   child-label-it:compito OR child-label-es:tarea OR
   parent-label-en:process OR parent-label-it:processo OR
   parent-label-es:proceso OR
   context-label-en:"executable activity task process" OR
   context-label-it:"attività eseguibile compito processo" OR
   context-label-es:"actividad ejecutable tarea proceso")
```

where proc() is the function representing the set of textual preprocessing activities performed on the terms contained in the query.

Here, we may see that in this ontology, the concept "Task" is represented as child of "Executable Activity"; while in the ontology used as example for the index construction, the concept "Task" is represented as part of the concept "Activity". In scenarios like these, the use of "relationship-label-XX" fields is important for avoiding loss of information during the query phase.

When the back-end component receives the request for suggesting new mappings, it performs a search operation in the index. As a result, a rank ordered by confidence score is produced and returned by the system. Such a score is computed by applying the scoring function shown in the Equation 1.

$$\mathrm{score}(R_{c_1}, R_{c_2}) = \mathrm{coord}(R_{c_1}, R_{c_2}) \cdot \sum_{x \in R_{c_1}} (\mathrm{tf}(x \in R_{c_2}) \cdot \mathrm{idf}(x)^2 \qquad (1)$$

[6] http://lucene.apache.org/solr/

where R_{c_1} and R_{c_2} are respectively the representations of concepts defined in the source and in the target ontologies, tf() and idf() are the standard "term frequency" and "inverse document frequency" functions used in the IR field [3], and coord(R_{c_1}, R_{c_2}) represents the number of label defined in the representation of c_1 occurring in the representation of c_2. The implemented version of the system returns the five candidate mappings with the highest score.

4 User Facilities of the Knowledge Management Tool

The back-end component described in the previous Section can be accessed through the user facilities that have been integrated as extensions of the MoKi [4] tool.

Figure 3 shows the process about how the entire suggestion-base mapping service works.

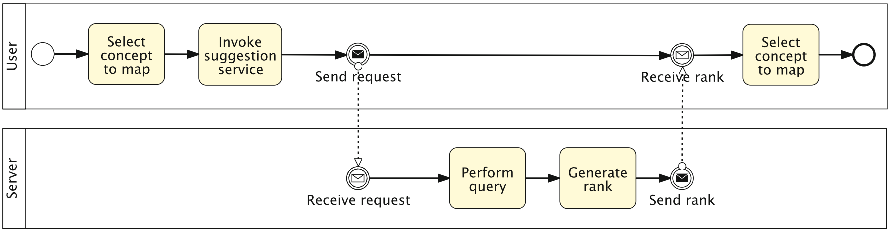

Fig. 3. Process describing the message exchange between the client and server sides

Once the expert selects the concept to map, a request message is sent to the server. Such a message contains the representation of the current concept and its context, structured as described in Section 3, and codified by using the JSON format. The server consumes the request and it generates a rank containing the five suggestions ordered by their confidence score. The rank is then encoded in JSON and it is sent back to the user interface that will show it to the user. Screens concerning the user facilities involved in this process are shown below.

For managing the mappings, a dedicated section in the concept modeling page as been integrated as shown in Figure 4. Here, the expert is able to see which are the concepts that have been already mapped with the current one and to decide if to maintain such mappings or to remove them. For creating a new mapping, the expert has to choose which ontology to use for requesting mapping suggestions, and then to click on the "Add New Mapping" button for invoking the suggestion service.

When the request is sent, on the background the structured representation of the current concept is converted into a query (as described in Section 3) which is performed on the index containing the concepts of the ontology specified by the expert. When the rank of the suggestions is composed, it is proposed to the expert as shown in Figure 5. For each suggestion, ordered by confidence score, the expert

Fig. 4. User facility for invoking the suggestion service.eps

is able to open the concept description page (if available) by clicking on the concept URI, and to eventually define a new mapping by clicking on the "Create Mapping" button.

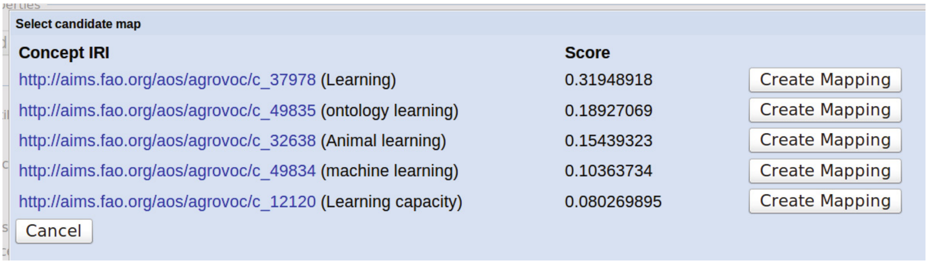

Fig. 5. Example of rank produced by the IR system containing the five suggestions for mapping the concept "Learning" defined in the Organic.Lingua ontology with a concept coming from the Agrovoc one.eps

Besides this, in a separated module, experts are able to upload new ontologies and creating the related indexes. Figure 6 shows the interface used by the experts for uploading a new ontology in the repository. This facility allows to convert the uploaded ontology in the structured representation described in Section 3 and to store it in a dedicated index. From this interface, experts are able to manage the ontologies already stored in the repository by viewing some basic information about them and, eventually, to delete one or more ontologies.

For adding a new ontology to the repository, experts have to select the file containing the ontology, write a description, decide an acronym for referring the ontology in the other sections of the tool, and, finally, press the "Save" button.

5 Evaluation

The presented platform has been evaluated in order to understand if the proposed suggestion-based mapping service provides an effective support to the mapping activity and if the facilities designed for supporting such an activity have been judged usable by the experts.

In detail, we are interested in answering two main research questions:

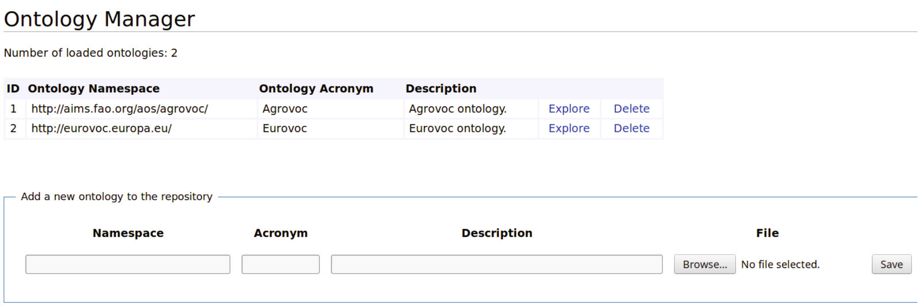

Fig. 6. Interface used for loading a new ontology within the system.

RQ1 Does the proposed approach provide an *effective* support, in terms of concept suggestions, to the mapping of multilingual ontologies?

RQ2 Are the provided interfaces *usable* for supporting the experts in the mapping activity between ontologies?

In order to answer these questions, we performed two types of analysis:

1. Quantitative: we collected objective measures concerning the effectiveness of the mappings suggested by the implemented approach and by providing a comparison between the time needed to map the resources manually, with respect to the time needed to define the mappings with our platform.
2. Qualitative: we collected subjective judgments from the experts involved in the use cases that have been asked to evaluate the general usability of the components and to provide feedbacks for future actions.

Concerning the qualitative evaluation of the suggestion-based mapping service, experts were asked to fill a questionnaire aiming at investigating their perception about the mapping process through the usage of the MoKi tool. Questions were organized in three parts: (i) one collecting information on the experts' background; (ii) one on the subjects' evaluation about MoKi and the usability of its different functionalities for accomplishing the mapping task; and (iii) a last one for retrieving impressions, and questions related to the work performed for the definition of new mappings between ontologies. Some of the questions were provided in the form of open questions, while most of them were closed questions. The latter type mainly concern the experts' evaluation of the tool usefulness on a scale from 1 to 5, varying according to the target of the evaluation (e.g., 1 = *extremely ease/useful/effective*, ... , 5 = *extremely useless/difficult/ineffective*).

It is important to highlight, that all experts were already familiar with the main functionalities of the MoKi tool, therefore, opinions about such functionalities were not collected through the questionnaire.

In Sections 5.1 and 5.2, we provide the description about how the evaluation on the two use cases has been organized; Section 5.3 presented the evaluation results, while in Section 5.4, we wrap up the consideration about what we experienced during the evaluation procedure.

5.1 Use Case 1: The Organic.Lingua Project

In the Organic.Lingua use case, six experts in the agricultural domain have been involved for defining the mappings between the Organic.Lingua ontology and the Agrovoc and Eurovoc ones. The Organic.Lingua ontology contains 291 concepts and it is focused on organic agriculture and sustainable agricultural processes. All concepts have been translated in 16 languages. The Agrovoc ontology is the wider available resource concerning food, nutrition, agriculture, fisheries, forestry, and environment. It is composed by around 32,000 concepts translated in 28 languages. While, the Eurovoc ontology is focused on modeling the general terminology used by the EU government including concepts related to agricultural policies that have to be adopted in the EU territory. It contains around 7,000 concepts translated in 24 languages.

Experts have been divided in two groups: three experts defined the mappings manually by using a well-known ontology editor tool (i.e. *Protégé* [5]); while, the other three, used the MoKi tool interface. The mapping operation has been performed in two steps: firstly, each expert mapped the Organic.Lingua ontology with the Agrovoc end Eurovoc ones independently by using the assigned tool; secondly, for cases where a full agreement were not found, experts discussed about the right mapping to provide. After the conclusion of the mapping activity, also the first group of experts performed the mapping between the ontologies by using the MoKi tool in order to allow them to compile the questionnaire about the tool.

First two rows of Table 1 contain information about the number of mapped concepts between the Organic.Lingua ontology and the Agrovoc and Eurovoc ones; while, rows 4 and 5 contain the measure of the inter-annotator agreement computed for each ontology. The averaged response times of the back-end component to requests performed by experts are shown in rows 7 and 8, respectively, for the Agrovoc and Eurovoc ontologies.

Finally, rows 1 and 2 of Table 2 show the results concerning the effectiveness of the ranks produced by the suggestion service, and the comparison of the time effort needed by the experts for completing the mapping operation.

5.2 Use Case 2: The PRESTO Project

In the PRESTO use case, three experts (that for this use case have an ontology engineering profile) have been involved for mapping the PRESTO core ontology with the classification schema extracted from the XVR 3d-library. The core PRESTO ontology contains 311 concepts described in 2 languages and it is focused on describing general items that can be used in virtual reality scenarios. The XVR classification schema contains around 1200 descriptions of 3d elements available in the environment defined in the XVR framework. Such elements provide descriptions in Italian and English languages.

For this use case, one expert has been assigned to the manual mapping task; while the other two performed the mapping activity by using the MoKi tool. Also in this case, all experts performed the mapping operation independently and, in case of disagreement, they discussed about the right mapping to write in the final

Table 1. Information about the number of mapped concepts, the measured inter-annotator agreement, and the average response time of the back-end component to mapping suggestion requests.

#	**Indicator**	**Value**
1	Number of concepts mapped with Agrovoc:	161
2	Number of concepts mapped with Eurovoc:	94
3	Number of concepts mapped with the XVR classification schema:	285
4	Inter-annotator agreement on Agrovoc mappings	94.41%
5	Inter-annotator agreement on Eurovoc mappings	97.87%
6	Inter-annotator agreement on the XVR classification schema	97.54%
7	Average response time for querying the Agrovoc repository (seconds)	1.27
8	Average response time for querying the Eurovoc repository (seconds)	0.94
9	Average response time for querying the XVR repository (seconds)	0.91

ontology. Moreover, the expert assigned to the manual mapping, performed the mapping activity also by using the MoKi tool in order to allow her to compile the evaluation questionnaire.

The number of mapped concepts of the PRESTO ontology is shown in the third row of Table 1, with the related inter-annotator agreement (row 6) and the average response time of the back-end component to the expert requests (row 9). While, the ranks effectiveness and the comparison of the time effort are shown in the third row of Table 2.

5.3 Quantitative and Qualitative Results

The registered average response time satisfies the experts need of being able to work quickly concerning the definition of new mappings. Indeed, by having an average response time around one second avoid downtime for users during the usage of the platform, aspect that has to be taken into account when web-based systems are used.

Concerning the analysis of the effectiveness results, in Table 2 we evaluate the precision of the ranks provided to experts by adopting standard information retrieval measures. In particular, here we evaluated if the correct mapping was placed respectively at the top of the rank, in the first three positions, or if it was at least provided in the rank (i.e. $Prec@5$, that in this case coincides with the Recall values). Results demonstrated the effectiveness of the back-end approach used for providing the suggestions to the experts. Indeed, almost in all cases the correct suggestions was presented to the users; moreover, encouraging results were obtained by observing the precision related to the top suggestion. This aspect allows to state that, in general, an approach based on IR techniques is a good direction for addressing the ontology mapping problem.

By considering the reduction of the time effort, we may see how the usage of the platform allows to strongly reduce the time necessary for completing the mapping between the ontologies, with a peak reduction of almost 75% of time. This result is

Table 2. Precision and recall values concerning the effectiveness of the suggestion retrieval system and the comparison (tool supported vs. manual) of the time effort needed by the experts for completing the mapping operations.

Mapped Ontology	**Suggested Ranks**				**Time Effort Comparison (minutes)**		
	Prec@1	Prec@3	Prec@5	Recall	Avg. Manual	Avg. With MoKi	Difference (%)
Agrovoc	0.81	0.91	0.97	0.97	193	49	-74.61%
Eurovoc	0.90	0.94	0.98	0.98	174	45	-74.14%
XVR	0.85	0.97	1.0	1.0	197	67	-65.99%

Table 3. MoKi functionality effectiveness in supporting the mapping activity

Ontology Loading	**Mapping Management**	**Mapping Browsing**
Effective	*Absolutely effective*	*Neither effective nor ineffective*

important for demonstrating the viability of the proposed platform for implementing it in real-world environments.

Concerning the qualitative evaluation, from the results collected through the questionnaire fulfilled by experts, in order to evaluate the statistical significance of the positivity/negativity of the collected results we applied the (one-tailed) Mann-Whitney test [6] verifying the hypothesis that $\tilde{F} \leq 3$, where $\tilde{F}$ represents the median of the evaluations for the factor F and 3 is the intermediate value in the 1 to 5 Likert scale. All the analyses are performed with a level of confidence of 95% (p-value < 0.05), i.e., there is only 5% of probability that the results are obtained by chance.

To better understand the relationship between the role of the tool in supporting the experts during the mapping activity, we asked them to express their evaluation about the effectiveness of the support provided by each functionality. Table 3 reports the corresponding evaluations.

The results show, in general, a good perception of the implemented functionalities, in particular concerning the procedure of defining a new mapping in the ontology. However, the browsing facility (i.e. the possibility of opening the description page of suggested concepts) did not convince the experts that asked for a graphical support able to quickly show the context of the suggested concepts. Indeed, they object about the fact that sometimes external pages describing suggested concepts might not be available and that the consult of the concept description might be time-consuming if context information are not clearly explained in the suggested concept pages.

5.4 Findings and Lesson Learned

The quantitative and qualitative results demonstrated the viability of the proposed platform in real-world scenarios and, in particular, its effectiveness in the proposed use cases. Therefore, we can positively answer to both research questions, **RQ1**: the back-end component provides effective suggestions for performing the mapping activity, and **RQ2**: the provided interfaces are usable and useful for supporting the experts in the mapping activity.

Besides these, there were other insights, either positive and negative, emerged during the subjective evaluation that we conducted.

The main positive aspect highlighted by the experts was related to the easy and quick way of defining a new mapping with respect to other available knowledge management tools (see details in Section 6) due to the missing, in them, of specific support for the mapping activity. The suggestion-based mapping service allowed to strongly reduce the amount of the most time-consuming activities, i.e., the navigation through the ontologies for analyzing candidate concepts for mappings. Indeed, while for relatively small (or less-deeper) ontologies, it is quite easy to detect which is the branch of the ontology containing candidate mappings. However, the same is not true for big (or high-deeper) ontologies where there is a significant time-overhead just for reaching the potential mapping candidates. Moreover, the ontology browsing aspect increases when the description of concept becomes more complex, i.e., when many relationships are modeled. In these cases, the visualization features implemented in knowledge management tools should be able to allow to show the context of each concept quickly. By adopting the proposed system, this problem can be avoided because the score computed for each suggestion already takes into account all relationships between the suggested concept and its directly connected ones.

However, even if from one side these suggestions effectively help the work of experts, they have been seen by them as a black box and, sometimes, when more than one suggestions are good candidates for defining a new mapping, experts requested to have a more immediate view of the context of the suggested concepts. As seen in Figure 5, the interface provides the possibility of opening the actual page containing details about the suggested concepts. This facility has been considered improvable by the experts for two reasons: (i) it may happen that the target page is not available, by blocking the work of the experts; and (ii) from the content of the target page may be not immediate to understand which concepts are in relationship with the suggested one and which kind of relationships they have. Experts proposal was the implementation of a graphical support showing the context (or a portion of the entire ontology branch) of each suggestion in order to a more clear picture of the "area" where the candidate mapping is placed. This feature has been judged valuable for improving the general overview of each suggested concept.

Finally, the only clearly negative aspect raised by the experts was the difficult to define a new mapping when the list of the suggested concept does not contain correct suggestions. Indeed, in this case, experts are not able to navigate through other suggestions unless they decide to open the ontology with the ontology management tool and start to navigate through it. This issue will be addressed by implementing the possibility of navigating through further set of suggestions that is a hypothesis that will be discarded in the beginning just for trying to find a good compromise between the number of suggestions and their effectiveness in order to avoid the consultation of a lot of suggestions by the experts.

6 Related Work

In this work either ontology mapping approaches, in a multilingual fashion, and knowledge management tools have been mentioned.

Literature about ontology mapping is very large and many systems and algorithms have been proposed. Surveys offering different perspectives can be found in [1] [7]. Concerning the use of multilinguality, research started to take it into account in the last fifteen years. First efforts on building and mapping multilingual ontologies have been conducted by using WordNet [8] [9] which kicked off several projects focused on the creation of different language versions of WordNet.

The two most significant ones are EuroWordNet and MultiWordNet. These projects adopted two different models of multilinguality: the one adopted in the EuroWordNet project [10] (EWN) consists in building language-specific wordnets independently from each other, and trying in a second phase to find correspondences between them. While the one adopted in the MultiWordNet project [11] (MWN), consists in building language-specific wordnets while keeping them aligned as much as possible to the synsets and semantic relations available in the Princeton WordNet (PWN).

After these, multilingual ontology mapping has been applied on several problems, as described in [12], where the authors address the problem of building conceptual resources for general multilingual applications. Examples of application fields in which multilingual ontology mapping has been applied are cross-language information retrieval [13], folksonomies [14], and specific domains-based applications [15] [16] [17] [18].

Concerning knowledge management tool, a lot of software born in the last decade supporting in different ways the modeling and knowledge sharing activities.

Knoodl [7] facilitates community-oriented development of OWL based ontologies and RDF knowledge bases. It also serves as a semantic technology platform, offering a Java service-based interface or a SPARQL-based interface so that communities can build their own semantic applications using their ontologies and knowledge bases.

Protégé [5] is an open source visual ontology editor and knowledge-base framework. Recently, Collaborative *Protégé* has been released as an extension of the existing *Protégé* system. It supports collaborative ontology editing as well as annotation of both ontology components and ontology changes. In addition to the common ontology editing operations, it enables annotation of both ontology components and ontology changes. It supports the searching and filtering of user annotations, also known as notes, based on different criteria.

NeOn [19]. It is a state-of-the-art, open source multi-platform ontology engineering environment, which provides comprehensive support for the ontology engineering life-cycle. The last version of the toolkit is based on the Eclipse platform and provides an extensive set of plug-ins covering a variety of ontology engineering activities.

[7] http://www.knoodl.com

While all these tools effectively support the ontology modeling activity, none of them provide specific facilities for supporting the mapping task in an easy and quick way, besides the classic axiomatization of the mapping relations, a discussion about the tool support for ontology mapping is presented in [20]. Indeed, as discussed in the previous section, the definition of a new mapping requires time-consuming activities by the experts that could be avoided by using facilities like the ones implemented in the discussed version of the MoKi tool.

7 Conclusions

In this paper, we presented a suggestion-based mapping service for supporting the ontology mapping problem in real-world scenarios with an emphasis on the importance of exploiting multilingual information provided by ontological models. Our platform, composed by an IR-based approach integrated in a knowledge management tool, namely MoKi, has been presented and its effectiveness and usability concerning the mapping process have been discussed. The platform has been quantitatively and qualitatively evaluated on two use cases, the Organic.Lingua and PRESTO projects, by demonstrating the usability and the effectiveness of the proposed suggestion-based mapping service. Therefore, by starting from the obtained results, the proposed system will be implemented in further uses cases in order to extend its overall evaluation. Finally, future work on the platform will be driven by the inferred lesson learned with the aim of improving either the quality and the usability of the entire system.

References

1. Euzenat, J., Shvaiko, P.: Ontology matching, 2nd edn. Springer-Verlag, Heidelberg (DE) (2013)
2. Choi, N., Song, I.Y., Han, H.: A survey on ontology mapping. SIGMOD Rec. **35**(3), 34–41 (2006)
3. van Rijsbergen, C.J.: Information Retrieval, Butterworth (1979)
4. Ghidini, C., Rospocher, M., Serafini, L.: Modeling in a wiki with moki: Reference architecture, implementation, and usages. International Journal On Advances in Life Sciences **4**(3&4), 111–124 (2012)
5. Gennari, J., Musen, M., Fergerson, R., Grosso, W., Crubézy, M., Eriksson, H., Noy, N., Tu, S.: The evolution of protégé: an environment for knowledge-based systems development. Int. J. Hum.-Comput. Stud. **58**(1), 89–123 (2003)
6. Wohlin, C., Runeson, P., Höst, M., Ohlsson, M.C., Regnell, B., Wesslén, A.: Experimentation in software engineering: an introduction. Kluwer Academic Publishers (2000)
7. Bellahsene, Z., Bonifati, A., Rahm, E. (eds.): Schema Matching and Mapping. Springer (2011)
8. Fellbaum, C.: WordNet: An Electronic Lexical Database. Bradford Books (1998)
9. Daudé, J., Padró, L., Rigau, G.: Mapping multilingual hierarchies using relaxation labeling. CoRR cs.CL/9906025 (1999)
10. Vossen, P.: Introduction to eurowordnet. Computers and the Humanities **32**(2–3), 73–89 (1998)

11. Magnini, B., Strapparava, C., Ciravegna, F., Pianta, E.: Multilingual lexical knowledge bases: applied wordnet prospects. In: Proceedings of the International Workshop on The Future of the Dictionary (1994)
12. Dorr, B.J., Levow, G.-A., Lin, D.: Building a chinese-english mapping between verb concepts for multilingual applications. In: White, J.S. (ed.) AMTA 2000. LNCS (LNAI), vol. 1934, pp. 1–12. Springer, Heidelberg (2000)
13. Zhang, L., Wu, G., Xu, Y., Li, W., Zhong, Y.: Multilingual collection retrieving via ontology alignment. In: Chen, Z., Chen, H., Miao, Q., Fu, Y., Fox, E., Lim, E. (eds.) ICADL 2004. LNCS, vol. 3334, pp. 510–514. Springer, Heidelberg (2004)
14. Jung, J.J.: Matching multilingual tags based on community of lingual practice from multiple folksonomy: a preliminary result. In: García-Pedrajas, N., Herrera, F., Fyfe, C., Benítez, J.M., Ali, M. (eds.) IEA/AIE 2010, Part II. LNCS, vol. 6097, pp. 39–46. Springer, Heidelberg (2010)
15. Galatescu, A.: Multilingual semantics in the request-service matchmaking. In: ECOWS, pp. 275–286. IEEE Computer Society (2006)
16. Stock, K., Cialone, C.: An approach to the management of multiple aligned multilingual ontologies for a geospatial earth observation system. In: Claramunt, C., Levashkin, S., Bertolotto, M. (eds.) GeoS 2011. LNCS, vol. 6631, pp. 52–69. Springer, Heidelberg (2011)
17. Jiun Chen, S., Hua Chen, H.: Mapping multilingual lexical semantics for knowledge organization systems. The Electronic Library **30**(2), 278–294 (2012)
18. Jung, J.J.: Exploiting multi-agent platform for indirect alignment between multilingual ontologies: A case study on tourism business. Expert Syst. Appl. **38**(5), 5774–5780 (2011)
19. Espinoza, M., Gómez-Pérez, A., Mena, E.: Enriching an ontology with multilingual information. In: Bechhofer, S., Hauswirth, M., Hoffmann, J., Koubarakis, M. (eds.) ESWC 2008. LNCS, vol. 5021, pp. 333–347. Springer, Heidelberg (2008)
20. Falconer, S.M., Noy, N.F.: Interactive techniques to support ontology matching. In: Bellahsene, Z., Bonifati, A., Rahm, E. (eds.) Schema Matching and Mapping. Data-Centric Systems and Applications, pp. 29–51. Springer (2011)