

# Using Linguistic Resource for Cross-Lingual Ontology Alignment

<http://dx.doi.org/10.3991/ijes.v1i1.2956>

Marouen Kachroudi<sup>1</sup>, Sami Zghal<sup>1,2</sup>, and Sadok Ben Yahia<sup>1</sup>

<sup>1</sup> Université de Tunis El Manar, Tunis, Tunisie;

<sup>2</sup> Université de Jendouba, Jendouba, Tunisia

**Abstract**—In the Semantic Web register, ontology alignment process can be seen as a cornerstone solution for the data heterogeneousness by allowing their interoperability. However, the most of the existing alignment methods assume that all ontologies to be aligned are described by identical languages. Indeed, very few approaches paid attention to the thriving challenge of multilingual ontology alignment. This paper introduces a new alignment method for multilingual ontologies. The proposed method implements a strategy of a direct alignment based on an external resource. Results obtained after extensive carried experiments are very encouraging and highlight many useful insights about the new proposed method.

**Index Terms**—Semantic Web, Interoperability, Ontology Alignment, Multilingualism, External Resource.

## I. INTRODUCTION

Multilingualism has become an issue of major interest for the Semantic Web community. This process has been accelerated due to a few initiatives which encourage all the active participants to make their data available to the public. These actors often publish their data sources in their own respective languages, in order to make this information interoperable and accessible to members of other linguistic communities [1].

As a solution, the ontology alignment process aims to provide semantic interoperable bridges between heterogeneous and distributed information systems. Indeed, the informative volume reachable via the Semantic Web stresses needs of techniques guaranteeing the share, reuse and interaction of all resources [2]. The explicitation of the associated concepts related to a particular domain of interest resorts to ontologies, considered as the kernel of the Semantic Web [3]. According to Gruber [4], an ontology can be defined in the context of computer and information sciences, as *a set of representational primitives with which to model a domain of knowledge or discourse*. On the other hand, the open and dynamic resources of the future Web endow it with a heterogeneous aspect, which reflects at once the formats or the languages variety of its description. This characteristic implies that the ontologies used for the description and the structuralization of the resources will be expressed in diverse formats, in particular in different natural languages.

Indeed, multilingual knowledge representation, access and translation are an impending need. For that purpose, the task of ontology alignment becomes particularly important by authorizing the reconciliation of resources described by different or multilingual ontologies. Multilingualism is identified as one of the six challenges of the Semantic Web. Consequently, some solutions were proposed at the ontology level, annotation level and the interface level [5]. At the ontology level, the support should be conceived by the ontology designers to create knowledge representations in diverse natural languages. At the annotation level, tools should be developed to help users to annotate the ontologies independently of the natural languages adopted in their descriptions. At the interface level, users should be able to have access to the information in natural languages of their own choice, without any linguistic constraint.

The absence of the multilingual aspect coverage can be a real handicap during the information exchange in between various services offered by the Semantic Web. So, application fields are more and more numerous and they put in front very specific difficulties. Moreover, the multilinguality coverage allows the reasoning on the context intersections of various ontological representations. In this context, the task of reasoning about overlapping context domains led to support multilingual information retrieval and digital content management. Multilingual ontologies alignment is still a little investigated domain in spite of the multiplicity of the alignment methods which remain restricted to monolingual ontologies, such as: OMAP [6], H-MATCH [7], OLA2 [8], RiMOM [9], to cite but a few.

This paper meets challenges strictly bound at the annotation level. Indeed, it proposes a new idea for multilingual ontology alignment called DCLOA (Direct Cross Lingual Ontology Alignment). The main idea is to align ontologies expressed in the OWL-DL language and written in different natural languages which already exist within such a semantic environment. Therefore, both ontologies  $O_s$  and  $O_t$  treated in the alignment process contain entities described in two different natural languages. The DCLOA method presents an originality on dealing with the multilingual aspect in ontology alignment. Furthermore, the new proposed approach appeals an external resource, *i.e.*, to assure and establish

equivalence between the ontological descriptors expressed in two different languages.

The outline of this paper is as follows. Section 2 reviews the existing methods in the field of multilingual ontology alignment and defines some terminologies and notations for the rest of this paper. Section 3 supplies a detailed description of the DCLOA method, its foundation and its various steps as the main contribution of this work. Section 4 presents the experimental results obtained with the considered test base as well as a comparative study with the pioneering methods of the literature. Finally, Section 5 draws the conclusion and the future issues of this paper.

## II. SCRUTINY OF RELATED WORK

Ontology alignment is considered as an evaluation of the degrees of resemblance or the differences detected on them [10]. Besides, the process of alignment can be defined as follows: being given two ontologies  $O_s$  and  $O_t$ , an alignment between  $O_s$  and  $O_t$  is a set of correspondences, (*i.e.*, a quadruplet):  $\langle e_s, e_t, r, Conf_n \rangle$ , with  $e_s$  in  $O_s$  and  $e_t$  in  $O_t$ ,  $r$  is a relation between two given entities  $e_s$  and  $e_t$ , while  $Conf_n$  represents the confidence level in this relation [10].

In the literature, a few methods were interested in multilingual ontologies alignment. A theoretical idea was also presented for building indirect alignments between multilingual ontologies [12]. The basic principle of this method is the reuse of already existing and stored alignment files. An intermediary alignment should be done between source and target ontology to compose a new alignment using such objects. Beforehand, equivalence between multilingual entities belonging in the two distinct ontologies should be discovered and established by a human expert. Then, a process of alignment composition is applied using alignment algebra [11]. Indeed, each composed relation is obtained thanks to two correspondence relations, *e.g.*, equivalence and inclusion. An API for multi-lingual ontology alignment was developed, which specifies a minimal interface based on few strategies [13]. Following a direct translation-based strategy, one source ontology is translated into a new one. TranslateOnto module reads the source ontology, translates it, and writes the resulting ontology. The translation step relies on an URI translation labels strategy. The TranslateStrategy module implements the OWLEntityURICConverterStrategy method of the OWL-API<sup>1</sup>. Such conversion is carried out using an external resource, *i.e.*, the Google-Translator-API<sup>2</sup> to provide the translations. In the stage of reading the source ontology and rendering the translated one, some tools are provided by the Alignment API<sup>3</sup> [14] and OWL-API are

respectively used. Consequently, the translated and targeted ontologies, can be matched with regard to a chosen matcher.

The DAMO method [12] implements a direct alignment strategy based on two phases and the use of two external resources. The alignment process contains two complementary components. The first one is the string based similarity module which allows to obtain a compound terminological similarity between the descriptors of the ontological entities to be aligned. The second one, is the structure based similarity module.

The recent evaluation campaign<sup>4</sup> for alignment systems was marked by the release of a new multilingual test base for the community working on ontology alignment. This has led to the emergence of new methods dedicated to the multilingual aspect. The AUTOMSV2 [16] alignment algorithm is composed of four complementary modules. The first module, synthesizes the alignments of two string-based similarity distance methods distributed with the Alignment API<sup>5</sup>. The second one, treats the alignments of two WordNet-based string-based similarity distance methods of the Alignment API. The third task is a single method that is implemented based on a string similarity distance approach. Finally, two more methods, a structure-based and an instance-based method, are integrated, based on the general principle of neighborhood similarity. AUTOMSV2 is using a free Java API called WebTranslator<sup>6</sup> to solve the multilingual problem. AUTOMSV2 translation method is converting the labels of classes and properties that are found to be in a non-English language and creates a copy of an English-labeled ontology file for each non-English ontology. The core idea of the WESEE approach [12] is to use a web search engine for retrieving web documents that are relevant for concepts in the ontologies to match. Concepts, labels, comments, and URI fragments are used as search terms. The search results of all concepts are then compared to each other. The more similar the search results are, the higher the concepts' similarity score. First, stop words like 'and', 'or', and 'so on' are inherently filtered, because they occur in the majority of documents. Second, terms that are common in the domain and thus have little value for disambiguating mappings are also weighted lower. For multi-lingual ontologies, we first translate the fragments, labels, and comments to English as a pivot language, using the Bing Search API's translation capabilities. The translated concepts are then processed as described above. In YAM++ approach [18], multiple working strategies have been implemented in order to deal with both terminological and conceptual heterogeneity of ontologies. Input ontologies are loaded and parsed. Then, information of entities in ontologies are indexed by the annotation indexing and the structure indexing components. The terminological matcher

<sup>1</sup> <http://owlapi.sourceforge.net/>

<sup>2</sup> <http://code.google.com/p/google-api-translate-java>

<sup>3</sup> <http://alignapi.gforge.inria.fr/>

<sup>4</sup> <http://oaei.ontologymatching.org/2012/>

<sup>5</sup> <http://alignapi.gforge.inria.fr/>

<sup>6</sup> <http://webtranslator.sourceforge.net/>

component produces a set of mappings by comparing the entities annotations. The instance-based matcher component supplements new mappings through shared instances between ontologies. In YAM++, matching results of the terminological matche and the instance-based matcher are aggregated into an element level matching result. Finally, the semantic verification component refines those mappings in order to eliminate the inconsistent ones. In the case where input ontologies use different languages to describe the annotations of entities, Bing <sup>7</sup> multilingual translator is used to translate those annotations to English. GOMMA [18] includes a generic component to semantically align ontologies.

GOMMA's matching component allows for direct and indirect ontology matching. Direct match strategies involve internal ontology knowledge like concept-associated or structural information. By contrast, the indirect matching is based on the composition of existing mappings to intermediate ontologies. The postprocessing stage deal with the combination or aggregation of the directly and indirectly obtained mappings and to select the most likely correspondences from the combined mapping. This approach returns for each concept only correspondences with the maximal similarity value or those within a small delta distance to the maximal value, *i.e.*, only the best correspondences for each source and target concept are kept. In case of multilingual ontologies a free translation API <sup>8</sup> is used to automatically translate non-English terms and produce a temporary list of equivalent terms. In the context of multilingual ontologies alignment, the alignment process should establish a semantic link between a source ontology  $O_s$  and a target ontology  $O_t$ . A source ontology  $O_s$ , expressed in a language  $L_s$ , will be aligned to a target ontology  $O_t$ , expressed in another language  $L_t$ . Indeed, both entities descriptors of the ontologies  $O_s$  and  $O_t$  presented by their corresponding graph will be translated from  $L_s$  language to the  $L_t$  language.

### III. THE DCLOA METHOD

The DCLOA method aligns a source ontology  $O_s$  to a targeted ontology  $O_t$ . The DCLOA alignment method operates on OWL-DL ontologies. In a pretreatment stage, both considered ontologies in entry are transformed into a graph structure. Consequently, each ontology is loaded in memory only once and transformed into a graph structure. This stage of parsing, is a technique which allows analyzing a given characters flow supplied in entry. This means segmenting the supplied flow following a model. Consequently, *parsing* is very useful to extract a targeted information in some data. For the DCLOA method, the parsing is realized through the OWL API <sup>9</sup>. Indeed, all the informative wealth of every ontology is described by a corresponding graph, *i.e.*, classes, relations and instances. Nodes of each graph are

classes and instances, whereas arcs represent links between the ontological entities. Each entity of an ontology is expressed with the RDF formalism :  $\langle \text{subject}, \text{predicate}, \text{Object} \rangle$  [20] and described thanks to OWL-DL constructors. The subject corresponds to a class or a relation. The predicates are the OWL-DL primitives or the RDF properties. Indeed, each property used in a triplet, enriches the knowledge of a described entity. The arrangement of all these knowledge constitutes the definition of an entity. The representation of an OWL-DL ontology under the shape of a graph allows to store it in central memory only once, so reducing disk accesses to the OWL-DL file. Also, we note that the alignment process is restricted only to entities names.

#### A. Cross-Lingual Similarity

From the corresponding graphs of both ontologies  $O_s$  and  $O_t$ , we get back entity names. As a first treatment, every target entity name will be considered as *release mechanism*. Then, from the external linguistic resource, *i.e.*, Bing <sup>10</sup>, we obtain the list of  $n$  equivalent words to the source word as depicted by figure 1. Assume that we want to compute the cross-lingual similarity (CLS) between two given strings  $s$  and  $t$ . These strings can be a single word or a text containing several statements.

At first, we tokenize each set of strings using delimiters and then convert it to a bag of words. Any non-alphabetical character in the given strings will be deleted. For example, if a given string  $s$  contains two non-alphabetical characters then we consider these two as delimiters and remove them from the string  $s$ . As a result,  $s$  will be tokenized into three words.

After converting each of the strings  $s$  and  $t$  to a bag of words, every word that is common to the two bags will increase the similarity score. In addition, from each set we remove stopwords, and this is achieved based on a list that contains stopwords for each supported language during the alignment process. For each extracted entity name  $N_t$  in  $O_t$ , we run a search test over the external resource Bing to create a list of the correlated wors, *i.e.*, a list in the same language of  $s$ .

If a given word  $d$  appears in this list, we conclude that both entities belonging respectively to  $O_s$  and  $O_t$  are similar, otherwise they aren't. We consider that  $\alpha$  is the number of words in  $s$ ,  $\beta$  the number of words in  $t$  and  $\gamma$  the number of common words between the two considered strings  $s$  and  $t$ . In this case, the cross-lingual based similarity using external linguistic resource is computed as follows:

$$\text{Sim}_{\text{CLS}}(s,t) = \gamma / \max(|\alpha|, |\beta|)$$

<sup>7</sup> <http://www.microsoft.com/enus/translator/>

<sup>8</sup> <http://mymemory.translated.net/>

<sup>9</sup> <http://owlapi.sourceforge.net/>

<sup>10</sup> <http://www.microsoft.com/en-us/translator/>

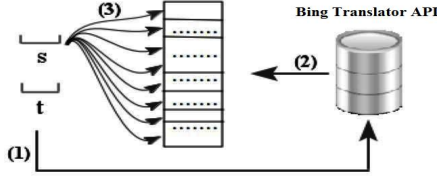


Figure 1. The translation stage in the DCLOA method

### B. Context Based Similarity

The idea of the context-based similarity is based on the assumption that, when two entities are similar, there is a big chance that the concepts that surround it are also similar. Here, by surrounding concepts (which define the semantic context) we mean super-concepts, sub-concepts and siblings concepts. Therefore, in the context based similarity, the description of a concept is based on its context. The value of this similarity is equivalent to computing the ratio of the number of nodes that are linguistically similar, to the total number of nodes belonging to the neighborhood. If we denote  $N_{CLS}$  the number of nodes that are linguistically similar and  $N_{NEIGHBORS}$  the total number of nodes belonging to the neighborhood, then the context based similarity formula is expressed by :

$$Sim_{CBS}(N_{Es}, N_{Et}) = |N_{CLS}| / |N_{NEIGHBORS}|$$

#### Example1:

If a given source node  $n_s$ , as depicted by Figure 2, is aligned to a targeted node  $n_t$ , then if we admit that :  $n_{s1}$ ,  $n_{t1}$ ,  $n_{s2}$ ,  $n_{t4}$  and  $n_{s3}$ ,  $n_{t2}$  are pairwise linguistically similar, then the context based similarity value is computed as follows :

$$Sim_{CBS}(N_{Es}, N_{Et}) = |N_{CLS}| / |N_{NEIGHBORS}| = 6/8 = 0.75$$

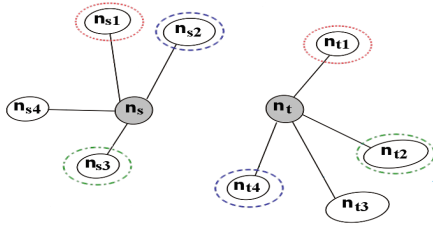


Figure 2 . Illustrating the three predefined neighborhood sets of two given nodes.

### C. Global Similarity Value

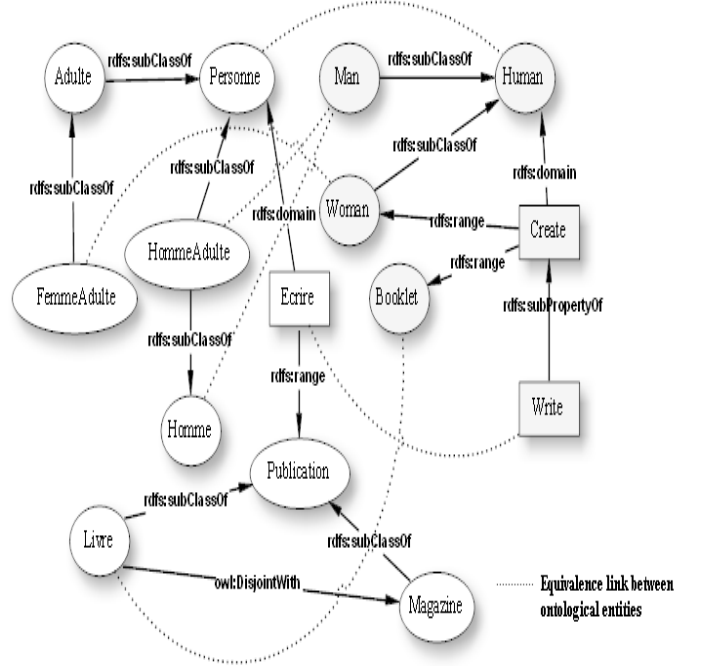
The alignment process ends by aggregating the various stemming values of the two modules: the linguistic and the context based one. The aggregation is realized through a fair weighty combination. With the sum of various weights is equal to 1, *i.e.* ( $\pi_1 + \pi_2 = 1$ ), the aggregative correspondence value,  $V_{Agg.Corr}$ , is computed as follows:

$$V_{Agg.Corr} = \pi_1 * Sim_{CLS}(n_s, n_t) + \pi_2 * Sim_{CBS}(n_s, n_t).$$

Such a setting can sometimes not be adequate, since it does not take into account the intrinsic nature of ontologies to align. This nature is described by its hierarchy and its information content. The values of  $\pi_1$  and  $\pi_2$  cannot be universally optimal. In other words, a particular parameters configuration can produce an excellent alignment result for a pair of ontologies and may provide poor results in other cases. In context we used a novel approach for the automatic adaptation of an ontology alignment method parameters. This approach is based on the use of the Choquet integral as an advanced aggregation operator [21]. Indeed, the aggregation problem is returned to the resolution of a linear system, in a manner to maximize the sum of the similarity values. This was done using the Kappalab R package <sup>11</sup>.

### IV. ILLUSTRATIVE EXAMPLE

The first stage of the DCLOA method consists in the opening of the OWL-DL files representing the ontologies to be aligned. Each ontology is loaded with in memory only once and transformed into a graph structure. The alignment process can operate using the fixed weights assigned to every entity name as well as for the various stages of the DCLOA method. The example uses two mockups of two ontologies represented by the Figure 3 and whose source code is represented in what follows.



The various stages of the alignment process are based on the equivalence links which can exist between the entities of both considered ontologies. This task supplies two vectors corresponding to the similarity computation in the DCLOA method.

Table 1 recapitulates the similarity values during each treatment. The weights assigned to the various values

<sup>11</sup> <http://cran.r-project.org>



stemming from each component treated by the Choquet integral. Doing so, the global correspondence value  $V_{\text{Agg.Corr}}$  is computed according to the pre-introduced expression.

TABLE I.  
 SIMILARITY VALUES FOR ENTITIES HANDLED IN THE ILLUSTRATIVE  
 EXAMPLE

$N_{E_s}$	$N_{E_t}$	$\text{Sim}_{\text{CLS}}$	$\text{Sim}_{\text{CBS}}$	$V_{\text{Agg.Corr}}$
Human	Personne	0.57	1.00	0.83
Write	Ecrire	1.00	0.50	0.79
Booklet	Livre	0.43	0.33	0.40
Man	Homme	1.00	0.75	0.85
Man	HommeAdulte	0.50	1.00	0.78
Woman	FemmeAdulte	0.50	0.75	0.68

Code of the first ontology :

```

<!DOCTYPE owl [
<!ENTITY onto1 "http://wwwsop.
inria.fr/acacia/ontologies/onto1.owl">
<!ENTITY
owl"http://www.w3.org/2002/07/owl">]
<rdf:RDF xmlns:owl
xmlns:rdf
="http://www.w3.org/1999/02/22-
rdf-syntax-ns"
xmlns:rdfs="http://www.w3.org/
2000/01/rdfschema"xml:base ="onto1;" >
<owl:Class rdf:ID="Personne" />
<owl:Class rdf:ID="Adulte">
<rdfs:subClassOf
rdf:resource="Personne"/>
</owl:Class>
<owl:Class rdf:ID="Femme Adulte">
<rdfs:label>Femme Adulte</rdfs:label>
<rdfs:label>Femme</rdfs:label>
<rdfs:subClassOf
rdf:resource="Adult"/>
</owl:Class> <owl:Class rdf:ID="Homme"
/>
<owl:Class rdf:ID="Homme Adulte">
<rdfs:subClassOf
rdf:resource="Personne"/>
</owl:Class> <owl:Class
rdf:ID="Livre">
<rdfs:subClassOf
rdf:resource="Publication"/>
<owl:disjointWith
rdf:resource="Magazine"/>
</owl:Class>
<owl:Class rdf:ID="Magazine">
<rdfs:subClassOf
rdf:resource="Publication"/>
</owl:Class>
<owl:Class rdf:ID="Publication" />
<owl:ObjectProperty rdf:ID="Ecrire">
<rdfs:domain rdf:resource="Personne"/>
<rdfs:range
rdf:resource="Publication"/>
</owl:ObjectProperty>
</rdf:RDF>
    
```

Code of the second ontology :

```

<!DOCTYPE owl [
<!ENTITY onto2 "http://wwwsop.
inria.fr/acacia/ontologies/onto2.owl">
<!ENTITY
owl"http://www.w3.org/2002/07/owl">
]
<rdf:RDF xmlns:owl
xmlns:rdf
="http://www.w3.org/1999/02/22-
rdf-syntax-ns"
xmlns:rdfs="http://www.w3.org/
2000/01/rdfschema"xml:base ="onto2;" >
<owl:Class rdf:ID="Human" />
<owl:Class rdf:ID="Woman">
<rdfs:label>Woman</rdfs:label>
<rdfs:label>Adult Female</rdfs:label>
<rdfs:subClassOf
rdf:resource="Human"/>
</owl:Class>
<owl:Class rdf:ID="Man">
<rdfs:subClassOf
rdf:resource="Human"/>
</owl:Class>
<owl:Class rdf:ID="Booklet" />
<owl:ObjectProperty rdf:ID="write">
<rdfs:subPropertyOf
rdf:resource="create"/>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="create">
<rdfs:domain rdf:resource="Human"/>
<rdfs:range rdf:resource="Booklet"/>
<rdfs:domain rdf:resource="Human"/>
<rdfs:range rdf:resource="Booklet"/>
</owl:ObjectProperty>
</rdf:RDF>
    
```

Commonly, ontologies developers adopt syntactical rules for some entity names, such as the hungarian notation. This notation is the practice of writing compound words or phrases in which the elements are joined without spaces, with each element's initial letter capitalized within the compound and the first letter either upper or lower case.

As flagged by the illustrative example, there is two compound entities names written using the hungarian notation, namely, FemmeAdulte et HommeAdulte. Worthy of mention, in the DCLOA method, such entities names are transformed to be treated separately as a set of words.

As depicted by Figure 2, HommeAdulte entity is aligned to Man entity. Doing so, the task means determining the statistical correspondence which can exist between the following both sets : {Man} and {Homme, Adulte} using the linguistic external resource. For example, Table 1 show that the two entities

Man and Homme have the highest similarity value, with regard to the obtained similarity values.

## V. EXPERIMENTAL STUDY

In what follows we will present the experimental study, based on the metrics of Precision and Recall. Subsequently, a comparative study is conducted with the pioneering methods.

### A. Test Cases

The carried out experimental evaluation uses the battery of test files provided by the OAEI (Ontology Alignment Evaluation Initiative Campaign) <sup>12</sup> <sup>13</sup>. This dataset is composed of a subset of the Conference track <sup>14</sup>, translated in eight different languages (*i.e.*, Chinese (cn), Czech (cz), Dutch (nl), French (fr), German (de), Portuguese (pt), Russian (ru), and Spanish (es)). With a special focus on multilingualism, it is possible to evaluate and compare the performance of alignment approaches through these test cases.

### B. Results and Discussion

The table I summarizes the values of precision and recall for the pioneering methods and the DCLOA method.

TABLE II.  
 PRECISION AND RECALL VALUES FOR DCLOA METHOD VS THE OTHER METHODS

TESTS	AUTOMSV2		GOMMA		WESEE		YAM++		DCLOA	
	Pre.	Rec.	Pre.	Rec.	Pre.	Rec.	Pre.	Rec.	Pre.	Rec.
en-fr	0.56	0.22	0.33	0.41	0.69	0.37	0.56	0.51	<b>0.71</b>	<b>0.55</b>
en-es	0.60	0.33	0.44	0.37	0.65	0.44	0.58	0.15	<b>0.70</b>	<b>0.52</b>
en-pt	0.62	0.26	0.53	0.39	0.63	0.43	0.58	0.54	<b>0.69</b>	<b>0.50</b>
fr-de	0.91	0.20	0.15	0.37	0.81	0.27	0.50	0.43	<b>0.85</b>	<b>0.17</b>
fr-pt	0.62	0.24	0.26	0.43	0.65	0.41	0.53	0.54	<b>0.73</b>	<b>0.62</b>

These metrics are presented in a condensed formula <sup>15</sup>, *i.e.*, each pair of languages reflects the mapping of eight ontologies in the source language ontologies and eight ones in the target language. Indeed, the DCLOA method supports five language pairs where it clearly exceeds other methods, except the couple (fr-de), DCLOA is exceeded by AUTOMSV2. The accuracy of the DCLOA method is highlighted by the recall metric which reflects the number of entity pairs correctly aligned. This is explained by two aspects: the first is that the DCLOA method adopts a dynamic and flexible aggregation operator which allows it to react to similarity values changes; the second is that the context-based similarity eliminates ambiguities at the level of aligned entity pairs.

<sup>12</sup> <http://oaei.ontologymatching.org/>  
<sup>13</sup> <http://oaei.ontologymatching.org/2012/multifarm/>  
<sup>14</sup> <http://oaei.ontologymatching.org/2012/conference/>  
<sup>15</sup> <http://oaei.ontologymatching.org/2012/multifarm/result/>

Therefore, enhancing the recall values by reducing the number of false positive entities. Unlike other methods mentioned above, the DCLOA method does not use string-based similarity measures that are not totally reliable, but at the DCLOA linguistic similarity is based on a semantic and boolean test. During the translation stage, the DCLOA method does not create a new translated ontology, which is an expensive process, but the translation starts and ends on a temporary lists.

## VI. CONCLUSIONS AND OUTLOOKS

Multilingual ontology alignment is an important task in the ontology engineering field. In this paper we introduced the DCLOA method for OWL-DL multilingual ontologies alignment. In addition, the results obtained by DCLOA method are satisfactory. In this frame, it is important to highlight the external resource contribution.

The proposed method showed a good performance compared to other methods, but still requires some improvements. In the near future, we intend also to enhance the performance of the DCLOA so that it can handle a wider range of natural languages. Also, the integration of new external resources can provide a wider choice of translation which is beneficial to the task of alignment. Besides, a graphical user interface (GUI) is needed to assist ordinary users. In addition, we work on the automatic detection of hierarchical trends paths in the considered ontologies even in the multilingual context, *e.g.*, logical links are usually independent of linguistic details. In addition, we aim to integrate the DCLOA method in a complete ontology localization system in interactive environments in the Semantic Web domains.

## REFERENCES

- [1] Berners-Lee, T.: Designing the web for an open society. In: Proceedings of the 20<sup>th</sup> International Conference on World Wide Web (WWW2011), Hyderabad, India (2011) 3-4. <http://dx.doi.org/10.1145/1963405.1963408>
- [2] Suchanek, F.M., Varde, A.S., Nayak, R., Senellart, P.: The hidden web, xml and semantic web: A scientific data management perspective. Computing Research Repository (2011) 534\_537.
- [3] Berners-Lee, T., Hendler, J., Lassila, O.: The semantic web. Scientific American (2001) <http://dx.doi.org/10.1038/scientificamerican0501-34>
- [4] Gruber, T.: Ontology. In: Encyclopedia of Database Systems. Springer-Verlag (2009) 1963-1965
- [5] Benjamins, V., Contreras, J., Corcho, O., Gómez-Pérez, A.: Six challenges for the semantic web. In: Special Interest Group on Semantic Web and Information Systems (SIGSEMIS Buelletin). (2004)
- [6] Straccia, U., Troncy, R.: oMPA: Combining classi\_ers for aligning automatically OWL ontologies. In: Proceedings of the 6th International Conference on Web Information System (WISE), Budva, ME (2005) 133-147
- [7] Castano, S., Ferrara, A., Montanelli, S.: Matching ontologies in open networked systems: techniques and applications. Journal on Data Semantics V (2006) 25-63
- [8] Kengue, J., J. Euzenat, P.V.: OLA in the OAEI 2007 evaluation contest. In: Proceedings of the 2<sup>nd</sup> International Workshop on Ontology Matching (OM-2007), Busan, Korea (2007) 188-195
- [9] Zhang, X., Zhong, Q., F. Shi, Li, J., Tang, J.: RiMOM results for OAEI 2009. In: Proceedings of the 4th International Workshop on

PAPER  
USING LINGUISTIC RESOURCE FOR CROSS-LINGUAL ONTOLOGY ALIGNMENT

- Ontology Matching (OM-2009), Washington, USA (2009) 208-215
- [10] Ehrig, M.: Ontology alignment: bridging the semantic gap. Springer-Verlag, New-York (2007)
- [11] Euzenat, J., Shvaiko, P.: Ontology Matching. Springer-Verlag, Heidelberg (DE) (2007)
- [12] Jung, J.J., Håkansson, A., Hartung, R.: Indirect alignment between multilingual ontologies: A case study of korean and swedish ontologies. In: Agent and Multi-Agent Systems: Technologies and Applications. Volume 5559 of LNCS. (2009) 233-241
- [13] Dos Santos, C.T., Quaresma, P., Vieira, R.: An api for multilingual ontology matching. In: Proceedings of the International Conference on Language Resources and Evaluation. (2010) 3830-3835
- [14] Euzenat, J.: An api for ontology alignment. In: Proceedings of the 3<sup>rd</sup> International Semantic Web Conference. (2010) 698-712
- [15] Kachroudi, M., Ben Yahia, S., Zghal, S.: Damo - direct alignment for multilingual ontologies. In: Proceedings of the 3<sup>rd</sup> International Conference on Knowledge Engineering and Ontology Development (KEOD), 26-29 October, Paris, France (2011) 110-117
- [16] Kotis, K., Katasonov, A., Leino, J.: Automsv2 results for oaei 2012. In: Proceedings of the 9th International Workshop on Ontology Matching (OM-2012) Colocated with the 11th International Semantic Web Conference (ISWC-2012). Volume 946 of CEUR-WS., Boston, USA (2012) 124-132
- [17] Paulheim, H.: Wesee-match results for oaei 2012. In: Proceedings of the 9th International Workshop on Ontology Matching (OM-2012) Colocated with the 11th International Semantic Web Conference (ISWC-2012). Volume 946 of CEUR-WS., Boston, USA (2012) 213-219
- [18] Ngo, D., Bellahsene, Z.: Yam++ results for oaei 2012. In: Proceedings of the 9th International Workshop on Ontology Matching (OM-2012) Colocated with the 11th International Semantic Web Conference (ISWC-2012). Volume 946 of CEUR-WS., Boston, USA (2012) 226-233
- [19] Groß, A., Hartung, M., Kirsten, T., Rahm, E.: Gomma results for oaei 2012. In: Proceedings of the 9th International Workshop on Ontology Matching (OM-2012) Colocated with the 11th International Semantic Web Conference (ISWC-2012). Volume 946 of CEUR-WS., Boston, USA (2012) 133-140
- [20] Klyne, G., Carroll, J.J.: Resource Description Framework (RDF): Concepts and Abstract Syntax. Technical report, W3C: World Wide Web Consortium, <http://www.w3.org/TR/2004/REC-rdf-concepts-20040210/> (05/24/2010) (2004)
- [21] Kachroudi, M., Zghal, S., Ben Yahia, S.: Paramétrage intelligent de l'alignement d'ontologies par l'intégrale de choquet. In: Actes des 13<sup>eme</sup> Journées Francophones Extraction et Gestion des Connaissances (EGC 2013), Toulouse, France (2013) 377-382

AUTHORS

**Marouen Kachroudi** and **Sadok Ben Yahia** are with Université de Tunis El Manar, Faculté des Sciences de Tunis, LIPAH, Programmation Algorithmique et Heuristique, 2092, Tunis, Tunisie, {marouen.kachroudi,sadok.benyahia}@fst.rnu.tn

**Sami Zghal** is with Université de Jendouba, Faculté des Sciences Juridiques, Economiques et de Gestion de Jendouba, 8189 Jendouba, Tunisia, sami.zghal@planet.tn

Submitted 26 June 2013. Published as re-submitted by the authors 23 July 2013.