



HAL
open science

Initiating Cross-Lingual Ontology Alignment with Information Retrieval Techniques

Marouen Kachroudi, Gayo Diallo, Sadok Ben Yahia

► **To cite this version:**

Marouen Kachroudi, Gayo Diallo, Sadok Ben Yahia. Initiating Cross-Lingual Ontology Alignment with Information Retrieval Techniques. 6èmes Journées Francophones sur les Ontologies, Oct 2016, Bordeaux, France. hal-01451016

HAL Id: hal-01451016

<https://hal.science/hal-01451016>

Submitted on 31 Jan 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Initiating Cross-Lingual Ontology Alignment with Information Retrieval Techniques

Marouen Kachroudi* — **Gayo Diallo**** — **Sadok Ben Yahia***

* *Université de Tunis El Manar, Faculté des Sciences de Tunis,
LIPAH-LR IES14, Informatique Programmation Algorithmique et Heuristique,
2092, Tunis, Tunisie ;*

{marouen.kachroudi,sadok.benyahia}@fst.rnu.tn

** *ERIAS, INSERM1219
Université de Bordeaux, France
gayo.diallo@isped.u-bordeaux2.fr*

RÉSUMÉ. L'alignement d'ontologies est une tâche qui a attiré une attention considérable ces dernières années. La majorité des travaux dans ce contexte ont porté sur le développement d'algorithmes d'appariement monolingues. Comme de plus en plus de ressources sont disponibles dans plus d'une langue naturelle, de nouveaux algorithmes sont nécessaires. Ceux-ci sont censés être capables d'aligner des ontologies qui partagent plus d'une langue. Pour relever ce défi, nous proposons dans ce papier une nouvelle méthode d'alignement inter-langues qui opère suivant une stratégie d'alignement basée sur l'utilisation des ressources externes et les techniques de recherche d'information. Les résultats obtenus montrent l'efficacité de la nouvelle méthode.

ABSTRACT. Ontology alignment is a task that has attracted considerable attention in recent years. The majority of works in this context focused on the development of monolingual matching algorithms. As more and more resources become available in more than one natural language, novel algorithms are required. The latter should be able to align ontologies which share more than one language. To meet this challenge, we propose in this paper a new cross-lingual alignment method that operates in a strategy based on the use of external resources and information retrieval techniques. The obtained results show the effectiveness of the new proposed method.

MOTS-CLÉS : Alignement d'ontologies, Inter-linguisme, Techniques de recherche d'information

KEYWORDS: Ontology Alignment, Cross-lingualism, Information Retrieval Techniques

1. Introduction

Ontologies, as specifications of conceptualisations (Gruber, 1995), are recognised as a "*basic component of the Semantic Web*" (Berners-Lee *et al.*, 2001) and have been widely used in knowledge management (Jurisica *et al.*, 2004). Indeed, actors of this new Web generation publish their data sources in their own respective languages, in order to make this information interoperable and accessible to members of other linguistic communities (Berners-Lee, 2011). 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 (Suchanek *et al.*, 2011). The explicitation of the associated concepts related to a particular domain of interest resorts to ontologies, considered as the kernel of the Semantic Web (Berners-Lee *et al.*, 2001). In addition, the open and dynamic resources of the future Web endow it with a heterogeneous aspect, which reflects at once the formats or the language varieties of its description. Consequently, ontologies that are labelled in diverse natural languages are increasingly evident. Given ontologies that are likely to be authored by different actors using different terminologies, structures and natural languages, ontology alignment has emerged as a way to achieve semantic interoperability. Recent researches in ontology alignment field has largely focused on dealing with ontologies labelled in the same natural language. Still a few researches have focused on alignment scenarios where the involved ontologies are labelled in different natural languages. However, current monolingual alignment techniques often rely on lexical comparisons made between resource labels, which limits their deployment to ontologies in the same natural language or at least in comparable natural languages. Given the existing monolingual alignment tools limitations, there is a pressing need for the development of alignment techniques that can deal with ontologies in different natural languages. One approach is direct cross-lingual ontology mapping. In this paper, direct cross-lingual ontology alignment refers to the process of establishing relationships among ontological resources from two or more independent ontologies where each ontology is labelled in a different natural language.

A well known approach (Dos Santos *et al.*, 2010) to achieve cross-lingual ontology alignment is to use translation techniques with the goal of converting a direct cross-lingual alignment problem into a monolingual alignment problem, which can then be solved by monolingual ontology alignment tools. Various techniques (Clark *et al.*, 2010) such as statistical machine translation and rule-based machine translation have been developed with the aim of improving the translations quality through word sense disambiguation (Navigli, 2009). Otherwise, machine translation tools are intended to assign an accurate meaning to a phrase in a specific natural language while limiting possible ambiguity, which is not necessarily a requirement in direct cross-lingual ontology alignment however. Therefore, to achieve direct cross-lingual ontology alignment, translations should lead to the correct alignments generation, but it is not of interest whether these translations are the most accurate localisations in the specific natural language. Consequently, translating the ontology labels in the context

of direct cross-lingual ontology alignment is not solely concerned with finding translated equivalents in the target natural language, but also is finding translations that can lead to correct mappings. There can be various ways to express the same concept in many natural languages. A simple example of this is : Ph.D. candidate and doctoral student both describe someone who is pursuing an academic degree of Doctor in any research axis. Envision this in the context of direct cross-lingual ontology alignment, assuming the target ontology is labelled in English and the source ontology is labelled in a natural language other than English. For a considered source ontology concept, its English translation can be Ph.D. candidate or doctoral student. Which one is more appropriate in the given alignment scenario ? To solve this problem, we would ideally like to know which candidate translation will lead to a correct mapping given that an equivalent concept is also presented in the target ontology. This translation selection process differs from traditional word sense disambiguation, as word sense disambiguation is the association of a given word in a text or discourse with a definition or meaning (sense) which is distinguishable from other meanings potentially attributable to that word (Ide *et al.*, 1998).

The remainder of this paper is organised as follows. A state of the art scrutinizes on direct cross-lingual ontology alignment approaches and related backgrounds are presented in section 2. To tackle this challenge, CLONA (Cross Lingual ONtology Alignment) is proposed in section 3. An overview of the evaluation is presented in section 4. Finally, section 5 concludes this paper with some outlooks for future work.

2. Context and literature review

This section presents the related knowledge background and a review on current approaches proposed in cross-lingual ontology alignment. Ontologies have gained increasing interest for structured meaning modeling among the Semantic Web community (Sure *et al.*, 2004). However, in decentralised environments such as the Semantic Web, the heterogeneity issue occurs when ontologies are created by different authors. To address the heterogeneity issue arising from ontologies on the Semantic Web, ontology alignment has become an important research field (Euzenat *et al.*, 2007).

Ontology alignment is considered as an evaluation of the resemblance degrees or the differences detected on them (Ehrig, 2007). Besides, the alignment process can be defined as follows : being given two ontologies \mathcal{O}_1 and \mathcal{O}_2 , an alignment between \mathcal{O}_1 and \mathcal{O}_2 is a set of correspondences, (*i.e.*, a quadruplet) : $\langle e_1, e_2, r, Conf_n \rangle$, with $e_1 \in \mathcal{O}_1$ and $e_2 \in \mathcal{O}_2$, r is a relation between two given entities e_1 and e_2 , while $Conf_n$ represents the confidence level in this relation (Euzenat *et al.*, 2013).

Ontologies are widely used in knowledge-based systems and the ontological applications crosses several disciplines, discussed next. In agriculture, the Food and Agriculture Organization (FAO)¹ provides reference standards for defining and structuring agricultural terminologies. Since all FAO official documents must be made available

1. <http://www.fao.org/home/fr/>

in five official languages including Arabic, Chinese, English, French and Spanish. For this purpose, a large amount of research has been carried out on large multilingual agricultural thesauri translations, their alignment methodologies and a requirements definition to improve the interoperability between these multilingual resources (Caracciolo *et al.*, 2007). In higher education domain, the Bologna declaration² has introduced an ontology-based framework for qualification recognition (Kismihók *et al.*, 2012) across the European Union (EU). In an effort to better reconcile labour markets with employment opportunities, an ontology is used to support the degrees and qualifications recognition within the EU. In e-learning, educational ontologies are used to enhance learning experiences and to empower system platforms with high adaptivity. In finance, ontologies are used to model knowledge in the stock market domain and portfolio management. In medicine, ontologies are used to improve knowledge sharing and reuse, such as work presented by Fang *et al.* (Fang *et al.*, 2006) that focuses on a traditional Chinese medicine ontology creation. A key observation from ontology-based applications such as those mentioned above is that the ontology development is closely associated with natural languages. Given the diversity of natural languages and the different conceptual models of ontology authors, the heterogeneity issue is ineluctable with ontologies that are built on different conceptualisation models and natural languages. Ontologies vast existence in several natural languages provides an impetus to discover solutions to support semantic interoperability.

The recent evaluation campaign³ for alignment systems was marked by the release of a new multilingual alignment systems which we are going to review the most relevant. AgreementMakerLight (AML) (Faria *et al.*, 2015) is presented as an automated ontology matching system that heavily relies on element-level matching and on the use of some external resources as background knowledge. The AML workflow comprises nine steps, *i.e.*, *ontology loading and profiling*, *translation*, *baseline matching*, *background knowledge matching*, *word and string matching*, *structural matching*, *property matching*, *selection*, and finally a *repair module*. AML employs the OWL API to read the input ontologies and extract the necessary information to populate its own data structures. The translation is achieved by querying Microsoft Translator for the full name in order to help provide the described context. AML employs an efficient, and generally precise weighted *string-equivalence* algorithm through the *Lexical Matcher*, to obtain a baseline class alignment between the input ontologies. AML has available four background knowledge sources used as mediators between the input ontologies. AML also employs the *Multi-Word Matcher*, which matches closely related multi-word names that have matching words and/or words with common WordNet synonyms or close hypernyms. For small and medium-sized ontologies, AML also employs a structural matching algorithm, called *Neighbor Similarity Matcher*. This algorithm computes similarity between two classes by propagating the similarity of their matched ancestors and descendants. AML employs a greedy selection algorithm, *i.e.*, the *Ranked Selector*, to reduce the alignment cardinality. Depending on the size of the input ontologies, one of three selection strategies is used : strict, permissive, or

2. <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:c11088>

3. <http://oaei.ontologymatching.org/2015/>

hybrid. AML employs a heuristic repair algorithm to ensure that the final alignment is coherent.

XMAP (Djeddi *et al.*, 2015) is a highly scalable ontology alignment system, which is able to deal with voluminous ontologies in efficient computation time. It is a fast and effective highly precised system able to perform large ontologies alignment. A semantic similarity measure is defined by XMAP using UMLS and WordNet to provide a synonymy degree between two entities from different ontologies, by exploring both their lexical and structural contexts. XMAP relies on the Microsoft Translate API to translate ontologies into many languages. The terminological layer is in charge of carrying out the similarity computing process between the entity names, combining linguistic similarity with the semantic elements according to the entities context. The structural layer performs two key tasks related to the ontologies structure. The first is the similarity computation between the concepts taking into account the taxonomic hierarchy, the second is the the similarity computation using the internal concepts structure information, *i.e.*, their properties, types and cardinality restrictions. The alignment layer is the final layer and its aim is to provide the final similarity matrix between the concepts taking into account the properties number influence and the similarity value that properties supply to the final similarity score.

LYAM++ alignment system (Tigrine *et al.*, 2015) take as an input a source ontology, given in a natural language and a second target ontology. The overall processes consists of three main components : a terminological multilingual matcher, a mapping selection module and, finally, a structural matcher. One of the original contributions of LYAM++ is the choice orchestrating of the pre-described components. The multilingual terminological matching module, acts on the one hand as a preprocessing component and, on the other hand as a light-weight terminological matcher between cross-lingual labels. The Mapping selection is a module that transforms the initial many mappings to a 1 : 1 alignment based on the principle of iteratively retaining the concept pairs with maximal similarity value. Finally, the structural matcher component filters the trustworthy pairs of the aligned concepts by examining the similarity values produced for their *parents* and their *descendants* in the ontology hierarchies.

LOGMAP (Jiménez-Ruiz *et al.*, 2015) is also a highly scalable ontology matching system that implements the consistency and locality principles. LOGMAP also supports real-time user interaction during the matching process, which is essential for use cases requiring very accurate mappings. LOGMAP is one of the few ontology matching system that can efficiently match semantically rich ontologies containing tens or even hundreds of thousands of classes. LOGMAP incorporates sophisticated reasoning and repair techniques to minimise the logical inconsistencies number, and provides support for user intervention during the matching process. LOGMAP extended a multilingual module to use both Google translate and Microsoft translator. Additionally, in order to split Chinese words, LOGMAP relies on the ICTCLAS⁴ library developed by the Institute of Computing Technology of the Chinese Academy of Sciences.

4. <https://github.com/xylvxy/node-ictclas>

To sum up, a key observation from the literature review is that, using machine as a means to bridge the gap between natural languages is a feasible approach to achieve cross-lingual ontology alignment as shown in the literature. In addition, the alignment process can be enhanced by the integration of powerful treatments inherited from other related specialties, such as Information Retrieval (IR) (Diallo, 2014). This paper presents the CLONA system and demonstrates its ability to align ontologies expressed in different natural languages using an external linguistic resource. CLONA as a few methods (Groß *et al.*, 2012, Kachroudi *et al.*, 2011) meets challenges strictly bound at the linguistic level in the context of multilingual ontology alignment. The driven idea of our new method is to cross the natural language barrier. CLONA presents a novel view to improve the alignment accuracy that draws on the information retrieval techniques.

3. CLONA method

The CLONA workflow comprises six different complementary steps, as flagged by Figure 1 : (i) Parsing and Pretreatment, (ii) Translation, (iii) Indexation, (iv) Candidate Mappings Identification and (vi) Alignment Generation. CLONA is an alignment system which aims through specific techniques to identify the correspondences between two ontologies defined in two different natural languages. Indeed, it starts with a pretreatment stage to model the input ontologies in a well defined format for the rest of the process. The second phase is that of translation into a chosen pivot language achieved by the Microsoft Bing⁵ translator. Thereafter, our method continues with an indexing phase over the considered ontologies. Then these indexes are queried to supply the candidate mappings list as a primary result. Before generating the alignment file, CLONA uses a filtering module for recovery and repair.

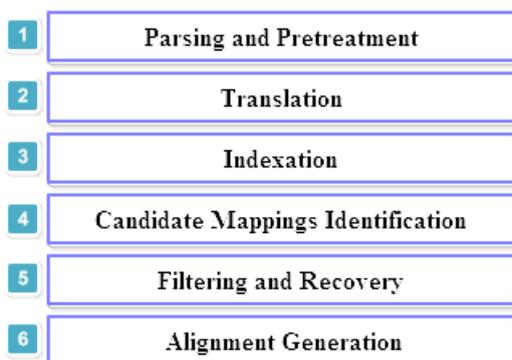


Figure 1. CLONA workflow.

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

Parsing and Pretreatment : This phase is crucial for ontologies pretreatment. It is performed using the OWL API. Indeed, it transforms the considered ontologies represented initially as two OWL files in an adequate format for the rest of the treatments. In our case, the goal is to substitute all the existing information in both OWL files so that each entity is represented by all its properties. Indeed, the parsing module begins by loading two ontologies to align described in OWL. This module allows to extract the ontological entities initially represented by a set of primitives. In other words, at the parsing stage, we seek primarily to transform an OWL ontology in a well defined structure that preserves and highlight all the information contained in this ontology. Furthermore, the resulting informative format, has a considerable impact on the similarity computation results thereafter. Thus, we get couples formed by the entity name and its associated label. In the next step we add an element to such couples to process these entities regardless of their native language.

Translation : The main goal of our approach is to solve the heterogeneity problem mainly due to multilingualism. This challenge brings us to choose between two alternatives, either we consider the translation path to one of the languages according to the two input ontologies, or we consider the translation path to a chosen pivot language. At this stage, we must have a foreseeable vision for the rest of our approach. Specifically, at the semantic alignment stage we use an external resource such as WordNet⁶. The latter is a lexical database for the English language. Therefore, the choice is governed by the use of WordNet, and we will prepare a translation of the two ontologies to the pivot language, which is English. To perform the translation phase we chose Bing Microsoft⁷ tool.

Indexation : Whether on the Internet, with many search engine or local access, we need to find documents or simply sites. Such research is valuable to browse each file and the analysis thereafter. However, the full itinerary of all documents with the terms of a given query is expensive since there are too many documents and prohibitive response times. To enable faster searching, the idea is to execute the analysis in advance and store it in an optimized format for the search. Indexing is one of the novelties of our approach. It consists in reducing the search space through the use of effective search strategy on the built indexes. In fact, we no longer need the sequential scan because with the index structure, we can directly know what document contains a particular word. To ensure this indexing phase we use the Lucene⁸ tool. Lucene is a Java API that allows developers to customize and deploy their own indexing and search engine. Lucene uses a suitable technology for all applications that require text search. Indeed, at the end of the indexing process, we get four different indexes to everyone of the two input ontologies depending on the type of the detected entities (*i.e.*, concepts, data types, relationships, and instances). More precisely, we get a first index file \mathcal{I}_1

6. <https://wordnet.princeton.edu/>

7. <https://www.bing.com/translator>

8. <https://lucene.apache.org/>

that corresponds to concepts index, a second one \mathcal{I}_2 dedicated to relationships, a third one \mathcal{I}_3 where we find a *datatypes* file index, and the last is \mathcal{I}_4 a file indexing instances. The documents at the indexes represent semantic information about any ontological entity. These semantic information is enriched by means of the external resource (*i.e.*, WordNet). Indeed, for each entity, CLONA keeps the entity name, the label, the translated label in English and its synonyms in English. So with Lucene, we created a set of indexes for the two ontologies, a search query is set up to return all the mapping candidates.

Candidate Mappings Identification : With Lucene, `TermQuery` is the most basic query type to search through an index. It can be built using one term. In our case, `TermQuery`'s role is to find the entities in common between the indexes. Indeed, once the indexes are set up, the querying step of the latter is activated. Thus, the query implementation satisfies the terminology search and semantic aspects at once as we are querying documents that contain a given ontological entity and its synonyms obtained via WordNet. It is worthy to mention that indexes querying is done in both senses. In other words, if we have two indexes \mathcal{I}_1 and \mathcal{I}'_2 respectively belonging to \mathcal{O}_1 and \mathcal{O}_2 , querying is outlined as follows : $(\mathcal{I}_1 \rightarrow Query \rightarrow \mathcal{I}'_2)$ and $(\mathcal{I}'_2 \rightarrow Query \rightarrow \mathcal{I}_1)$. Indeed, this junction greatly increases the CLONA performance. The result of this process is a set of documents sorted by relevance according to the Lucene score assigned to each returned document. Thus, for each query, CLONA keep the first five documents returned and considers them as candidate mappings for the next phase.

Filtering and Recovery : The filtering module consists of two complementary sub-modules, each one is responsible of a specific task in order to refine the set of primarily aligned candidates. Indeed, once the list of candidates is ready, CLONA uses the first filter. Indeed, we should note that indexes querying may includes a set of redundant mappings. Doing so, this filter eliminates the redundancy. Indeed, it goes through the list of candidates and for each candidate, it checks if there are duplicates. If this is the case, it removes the redundant element(s). At the end of filtering phase, we have a candidates list without redundancy, however, there is always the concern of *false positives*, indeed, there was the need to establish a second filter. Once the redundant candidates are deleted, CLONA uses the second filter that eliminates *false positives*. This filter is applied to what we call *partially* redundant entities. An entity is considered as *partially* redundant if it belongs to two different mappings (*i.e.*, being given three ontological entities e_1 , e_2 and e_3 . If on the one hand, e_1 is aligned to e_2 , and secondly, e_1 is aligned to e_3 , this last alignment is qualified as doubtful. We note that CLONA generates (1 : 1) alignments. To overcome this challenge, CLONA compares the topology of the two suspicious entities (e_3 neighbors with e_1 neighbors, e_2 neighbors with e_1 neighbors) with respect to the redundant entity e_1 , and retains the couple having the highest topological proximity value. All candidates are subject of this filter, and as output we have the final alignment file.

Alignment Generation : The result of the alignment process provides a set of mappings, which are serialized in the RDF format.

4. Experimental Study

CLONA system has been developed with a unique focus on multilingual ontologies. The carried out experimental evaluation uses the battery of test files provided by the OAEI (Ontology Alignment Evaluation Initiative Campaign)⁹. The Multifarm dataset is composed of a subset of the Conference track, translated in nine different languages (*i.e.*, Chinese (cn), Czech (cz), Dutch (nl), French (fr), German (de), Portuguese (pt), Russian (ru), Spanish (es) and Arabic (ar)). The considered ontologies are relatively small. With a special focus on multilingualism, it is possible to evaluate and compare the performance of alignment approaches through these test cases.

Since the CLONA method participated in the OAEI'2015 campaign, the evaluation process was conducted following a well established procedure. Therefore, CLONA was performed to determine $45 * 25$ alignments. In this context, we distinguish two sub-cases, namely, "*Same Ontologies*" and "*Different Ontologies*". The first sub-case seeks alignment between same ontologies but expressed in two different natural languages. The second sub-case concerns alignment between all possible combinations of different ontologies pairs in different languages. By comparing the pioneering methods participating in the OAEI'2015 campaign and which developed specific techniques in multilingual processing, CLONA ranks second. It is anticipated by AML and succeeded by LogMap, LYAM ++ and XMAP. Table 1 summarizes the Precision (P), F-Measure (F-M) and Recall (R) metric values for all methods.

| | Different Ontologies | | | Same Ontologies | | |
|--------|----------------------|------|------|-----------------|------|------|
| | P | F-M | R | P | F-M | R |
| AML | 0.53 | 0.51 | 0.50 | 0.93 | 0.64 | 0.50 |
| CLONA | 0.46 | 0.39 | 0.35 | 0.91 | 0.58 | 0.42 |
| LOGMAP | 0.75 | 0.41 | 0.29 | 0.95 | 0.45 | 0.30 |
| LYAM++ | 0.14 | 0.14 | 0.14 | 0.26 | 0.19 | 0.15 |
| XMAP | 0.22 | 0.24 | 0.27 | 0.66 | 0.37 | 0.27 |

Table 1. Evaluation metrics average values for CLONA and other methods

These values are an average of 1125 made alignments. The obtained values in the sub-case "*Same Ontologies*" are clearly better than those of "*Different Ontologies*" sub-cases. This is explained by the fact that ontologies have the same structure in the first scenario (*i.e.*, "*Same Ontologies*"), which improves the process of the filtering and

9. <http://oaei.ontologymatching.org/>

recovery module through enhancing similarity between the neighbor entities. Compared to the other participating methods, CLONA provides the second better values, especially in terms of Recall, *i.e.*, 35% and 42% respectively for the first and second sub-cases. The F-Measure values flagged by Table 2, confirm the CLONA performance which ranks second in the average metrics for all mixed sub-cases.

| Method | F-Measure |
|--------|-----------|
| AML | 0.56 |
| CLONA | 0.43 |
| LOGMAP | 0.40 |
| XMAP | 0.38 |
| LYAM++ | 0.15 |

Table 2. *F-Measure average values for CLONA and other methods*

In summary, and as shown in Figure 2, the CLONA method, and compared to other competing methods of the OAEI'2015 campaign, dominates 28.88% of tests occupying the first place. In addition, CLONA ranks second on 46.66% of the tests, third on 22.22% and fourth in 2.22% of tests. It should be emphasized also that in the sub-case "Same Ontologies", and over 45 treated language pairs, CLONA ranked first out of 15 couples.

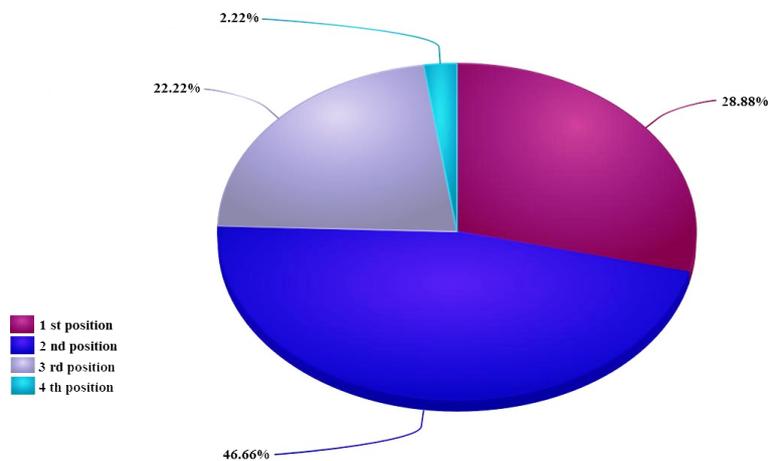


Figure 2. *Percentages of occupied positions by CLONA over all tests*

The performance of the CLONA method is detailed on the link below as well as the alignment resulting files ¹⁰. CLONA benefited from the power of the Lucene search and

10. <http://oaei.ontologymatching.org/2015/results/multifarm/index.html>

query engine, for detecting candidate alignments. This has been an asset for CLONA compared to other methods. Especially, the fact of bringing all ontological entities labels in a pivot language (*i.e.*, English), endowed CLONA with a sharpness to cross the natural language barrier. In this context, it should be noted that the enrichment of the label entities translated lists with synonyms from WORDNET armed CLONA with a semantic layer to increase its ability to detect similar entities.

5. Conclusions

Multilingual ontology alignment is an important task in the ontology engineering field. In this paper, we introduced the CLONA method for cross-lingual ontology alignment. In addition, the results obtained by CLONA method are satisfactory. In this frame, it is important to highlight the external resource contribution. CLONA has implemented a technique for determining alignment candidates across the power of Lucene search engine. In addition, during the translation phase, we have set up a local translator that is built during the alignment process. This treatment reduces the translation time cost and access to the external resource. CLONA participation in OAEI'2015 campaign was encouraging, as it supplies good F-measure values in the two considered scenarios. Results reflects some strengths and some positive aspects. In addition, we aim to exploit more than one external resource during the alignment process and deploy CLONA in a semantic Web service.

6. Bibliographie

- Berners-Lee T., « Designing the web for an open society », *Proceedings of the 20th International Conference on World Wide Web (WWW2011)*, Hyderabad, India, p. 3-4, 2011.
- Berners-Lee T., Hendler J., Lassila O., « The semantic Web », *Scientific American*, 2001.
- Caracciolo C., Sini M., Keizer J., « Requirements for the Treatment of Multilinguality in Ontologies within FAO », *Proceedings of the OWLED 2007 Workshop on OWL : Experiences and Directions, Innsbruck, Austria, June 6-7, 2007*, vol. 258 of *CEUR Workshop Proceedings*, 2007.
- Clark A., Eyraud R., Habrard A., « Using Contextual Representations to Efficiently Learn Context-Free Languages », *Journal of Machine Learning Research*, vol. 11, p. 2707-2744, 2010.
- Diallo G., « An effective method of large scale ontology matching », *Journal of Biomedical Semantics*, vol. 5, p. 44 (Electronic Edition), 2014.
- Djeddi W. E., Khadir M. T., Yahia S. B., « XMap : results for OAEI 2015 », *Proceedings of the 12th International Workshop on Ontology Matching (OM-2015) Colocated with the 14th International Semantic Web Conference (ISWC-2015)*, vol. 1545 of *CEUR-WS*, Bethlehem (PA US), p. 216-221, 2015.
- Dos Santos C. T., Quaresma P., Vieira R., « An API for Multi-lingual Ontology Matching », *Proceedings of the International Conference on Language Resources and Evaluation*, p. 3830-3835, 2010.

- Ehrig M., *Ontology alignment : bridging the semantic gap*, Springer-Verlag, New-York, 2007.
- Euzenat J., Shvaiko P., *Ontology Matching*, Springer-Verlag, Heidelberg (DE), 2007.
- Euzenat J., Shvaiko P., *Ontology Matching (Second Edition)*, Springer-Verlag, Heidelberg (DE), 2013.
- Fang K., Chang C., Chi Y., « Leveraging Ontology-Based Traditional Chinese Medicine Knowledge System : Using Formal Concept Analysis », *Proceedings of the 2006 Joint Conference on Information Sciences, JCIS 2006, ROC, October 8-11, 2006*.
- Faria D., Martins C., Nanavaty A., Oliveira D., Sowkarthiga B., Taheri A., Pesquita C., Couto F., Cruz I., « AML results for OAEI 2015 », *Proceedings of the 12th International Workshop on Ontology Matching (OM-2015) Colocated with the 14th International Semantic Web Conference (ISWC-2015)*, vol. 1545 of *CEUR-WS*, Bethlehem (PA US), p. 116-123, 2015.
- Groß A., Hartung M., Kirsten T., Rahm E., « GOMMA Results for OAEI 2012 », *Proceedings of the 9th International Workshop on Ontology Matching (OM-2012) Colocated with the 11th International Semantic Web Conference (ISWC-2012)*, vol. 946 of *CEUR-WS*, Boston, USA, p. 133-140, 2012.
- Gruber T. R., « Towards Principles for the Design of Ontologies used Knowledge Sharing », *International Journal of Human-Computer Studies*, vol. 43, p. 907-928, 1995.
- Ide N., Véronis J., « Introduction to the Special Issue on Word Sense Disambiguation : The State of the Art », *Computational Linguistics*, vol. 24, n° 1, p. 1-40, 1998.
- Jiménez-Ruiz E., Grau B. C., Solimando A., Cross V., « LogMap family results for OAEI 2015 », *Proceedings of the 12th International Workshop on Ontology Matching (OM-2015) Colocated with the 14th International Semantic Web Conference (ISWC-2015)*, vol. 1545 of *CEUR-WS*, Bethlehem (PA US), p. 171-175, 2015.
- Jurisica I., Mylopoulos J., Yu E. S. K., « Ontologies for Knowledge Management : An Information Systems Perspective », *Knowledge and Information Systems*, vol. 6, n° 4, p. 380-401, 2004.
- Kachroudi M., Ben Yahia S., Zghal S., « DAMO - Direct Alignment for Multilingual Ontologies », *Proceedings of the 3rd International Conference on Knowledge Engineering and Ontology Development (KEOD), 26-29 October, Paris, France*, p. 110-117, 2011.
- Kismihók G., Szabó I., Vas R., « Six Scenarios of Exploiting an Ontology Based, Mobilized Learning Environment », *International Journal of Mobile and Blended Learning*, vol. 4, n° 1, p. 45-60, 2012.
- Navigli R., « Word sense disambiguation : A survey », *ACM Computing Surveys*, 2009.
- 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*, 534-537, 2011.
- Sure Y., Staab S., Studer R., *Handbook on Ontologies*, Springer-Verlag, chapter On-To-Knowledge Methodology (OTKM), p. 117-132, 2004.
- Tigrine A. N., Bellahsene Z., Todorov K., « LYAM++ results for OAEI 2015 », *Proceedings of the 12th International Workshop on Ontology Matching (OM-2015) Colocated with the 14th International Semantic Web Conference (ISWC-2015)*, vol. 1545 of *CEUR-WS*, Bethlehem (PA US), p. 176-180, 2015.