International Journal of Computer Applications (0975 – 8887)
Volume 180 – No.5, December 2017

Evaluation of Structural Ontology Alignment Techniques

Nouman Khan
Lecturer
University of Malakand
Chakdara, Pakistan

Sadaqat Jan
Associate Professor
University of Engineering & Technology
Mardan, Pakistan

M. Sohail Nisar
Lecturer
INU Peshawar
Pakistan

ABSTRACT
Specialists in the field of software engineering and particularly Ontology Engineers have proposed and executed different matching strategies to connect the semantic crevice between heterogeneous Ontologies. In any case, more change and examination is required in these strategies with a specific end goal to accomplish better recall & precision for the alignment sets. This thesis presents alignment of structural Ontologies which utilizes matcher to guide entities of two unique Ontologies. Falcon AO, RiMOM and few other matchers are utilized for comparing structural similarity among entities in mapping procedure. The results obtained are assessed for recall & precision and also compared with the given reference Ontologies. The result shows that RiMOM and LILY lead the list in alignment precision, but if Recall is considered then ASMOV is leading the list of matchers considered.

Keywords
Evaluation; Alignment; Ontology; Structural Alignment

1. INTRODUCTION
Aligning different Ontologies is a tedious & hectic job which still needs an investigation for getting better accuracy & precision. Alignment can be done either manually or automatically. In manual alignment it's the user who provides input but due to more time consumption & non-applicability to applications of automated frameworks, this isn't a feasible method. An automatic alignment is only conceivable approach to connect the semantic web. This strategy requires the members of experiments for gathering information, and uses distinctive Ontologies of a particular domain for alignment. Numerous researches with respect to structural alignment have been carried out. Each of it has its own advantages and disadvantages but no result that is semantically equivalent, yet exists for structural matching. Ontology alignments can be used for different tasks like answering queries, browsing semantic web and ontology merging etc.

2. RELATED WORK
The current web has a derivative called the semantic web where the data & information has been given a meaning that is very well defined. Semantic Web will enable machines and people to work in collaboration [1]. Getting knowledge from various Ontologies in different domains has also become easier with the arrival of semantic web. Ontology identifies or describes a concept [2]. Ontology can also be thought of as a representation of entities & relationship among them, in a certain domain [3]. It is obvious to have several Ontologies for the same domain over the internet hence there are possibilities of having the same entities with different representation in various Ontologies. Alignment therefore is a technique used for resolving this issue and to find entities that are semantically correct. Two Ontologies each consisting of classes & properties are given as in input to the alignment process & relationships between entities in the form of correspondences are shown as a result [4]. This result is called ontology alignment. Therefore, a set of correspondences among entities of a particular domain actually represents ontology alignment [5].

The matching techniques are basically classified into two categories that are Element-Level and Structure-Level. Many different techniques are there, which can be used for comparing entities from various Ontologies. For this purpose, different matchers are used for finding out similarities and correspondences between those entities. This similarity can be measured on the basis of string, linguistic and Structural correspondences between entities. The string based technique searches for sequences of letters or characters. For finding our similarities between entities from various Ontologies using homonyms & synonyms, technique of matching through linguistics is used [6] [7] [8]. Structural alignment techniques are used for measuring similarities among classes like super-type, sub-type relationships & properties of classes, further split into two types i-e internal (considers attributes and their types which is internal structure of entities) and relational (considers relation between the entities).

Structural alignment techniques are different from other linguistic and string based methods by not only considering a single concept at once, but by using the structural information of Ontologies for finding the correspondences. As Ontologies can be treated as graphs so graph matching techniques can be used for comparison of sub-graphs which belongs to different concepts. For instance, two or more concepts having similar children (or leaf nodes) sets must be matched, while the confidence could be the ratio of equal children (or leaves). For class hierarchy, Taxonomy structure can also be considered, for instance, the ratio of mutual super-classes. Similarity flooding [9] is another method utilizing an idea that nodes which are similar would have similar neighbours too, iteratively indicates similarity along the graphical structure. Ontological structure is used by a number of alignment systems, such as *Falcon AO, Cupid, COMA, Lily, ASMOV, Anchor-Prompt, RiMOM, OLA2*and many others.

Falcon AO is a well-known system for matching Ontologies &idealising interoperability in-between applications of semantic web, using the Ontologies which are although different but related to each other. Falcon-AO is a framework for Semantic Web applications, which goes for giving key advancements to finding, adjusting, realizing Ontologies, and at last, to capture information from the Web through a methodology which is ontology driven. Falcon AO, which is indeed a prominent tool of Falcon, is a programmed ontology coordinating framework that is used for finding interoperability among applications of semantic web that uses diverse yet related Ontologies. Recently, it has turned into an exceptionally useful and prominent decision for matching Ontologies that are presented by RDF(S) and OWL.

Falcon AO which is executed in Java, and nowadays, it is an open source venture under the license Apache 2.0.For structural alignment it uses an approach called GMO (Graph Matching for Ontologies). It also uses V-Doc which is a linguistic matcher and PBM which is partition-based Matcher and a few more.

3. STRUCTURAL MATCHERS

In order to produce ontology alignments effectively, ontology matchers play a vital role. Different kinds of matchers are available which can match Ontologies on the basis of their structures, language or strings etc. Multiple techniques are combined to carry out the matching process, and main techniques are usually structural and lexical (string based).We'll study and look into details of the structural matchers like Falcon AO [17], ASMOV [14], OLA2 [12], QOM [16], RiMOM [10], LILY [11] and few more.

3.1 RiMOM

RiMOM [10] produces alignment of Ontologies by utilizing a combination of different methodologies which outperform a single methodology used in many cases, but in some cases it underperforms i-e if we have two exactly same Ontologies in terms of taxonomical structure, one defined in English and other in Urdu; in this case a structure based method would be pretty much effective but using a language based method would be totally useless. There's no tuning of weights that up to what extent should be a single strategy used in combination. In earlier version of RiMOM there is static selection of single and multiple strategies which have been overcome by dynamic multi-strategy framework of RiMOM, which automatically determines the methods to be used in alignment & also tunes the thresholds of each method.

3.2 Falcon AO

Falcon AO is a well-known system for matching Ontologies & idealising interoperability in-between applications of semantic web, using the Ontologies which are although different but related to each other. Falcon-AO is a framework for Semantic Web applications, which goes for giving key advancements to adjusting, finding and realizing Ontologies, and at last, to capture information from the Web through a methodology which is ontology driven.

Falcon AO, which is indeed a prominent tool of Falcon, is a programmed ontology coordinating framework that is used for finding interoperability among applications of semantic web that uses diverse yet related Ontologies. Recently, it has turned into an exceptionally useful and prominent decision for matching Ontologies that are presented by RDF(S) and OWL.

Falcon AO which is executed in Java, and nowadays, it is an open source venture under the license Apache 2.0. It has four main components of matchers; for structural alignment it uses an approach called GMO (Graph Matching for Ontologies). It also uses V-Doc which is a linguistic matcher and PBM for making fragments of Ontologies whereas I-Sub for string matching.

3.3 LILY

LILY [11], an ontology alignment framework works on sub-graph matching technique. The procedure for alignment creation in LILY comprises of 3 phases.

In the initial phase, entities of Ontologies are converted into semantic sub-graphs. Sub-graph is basically the entity's context in its belonging ontology. In this tool, the entity's meaning is itself captured by the semantic sub-graph which simply means the meaning of each and every entity is defined by LILY before similarity computation.

Over the sub-graphs extracted, alignment similarity gets calculated. Descriptions of 2 kinds are used for interpretation of properties and concepts. Basic description is the first one i-e the document comprising of the identifier, name and comments. Description of semantics is the second one which contains the data about hierarchies of class, properties related and instances. String-based methodologies are used for the first document and structural similarity techniques for the second document. Similarities of corresponding parts are calculated for descriptions of various entities, which are then combined using experimental weights. But when the descriptions are not clear then a few alignments are returned. Third phase is of similarity propagation which is usually used in cases when only few mappings are returned by similarity computation in second phase.

3.4 OLA 2

OLA2 [12] is a framework for matching of Ontologies, that follows a similarity paradigm. Being the successor of OLA framework it uses the same features and principles as portrayed in [13].

OLA (DL Dialect) is basically addressed by OLA2. A same integrated approach is used by OLA2 for computing similarity for all type of entities. For matching purpose, match graphs are created from the given Ontologies, which represent the entities and their relationships among each other. OLA2 characterizes similarity as a straight blend of encompassing entities of Ontologies. As a result an equation system is produced and it gets solved by using algorithm named as approximate fix-point.

3.5 ASMOV

ASMOV [14][15] which is basically meant for automated mapping of Ontologies with validation of semantics. In order to facilitate the process of integrating heterogeneous systems, utilizing Ontologies of their data source, ASMOV ("Automated Systems Mapping of Ontologies with Validation") has been designed.

ASMOV matching process is performed in three phases which are preprocessing, running algorithm for computing iterative similarity and then validation of mappings using some rules-set. In the alignment validation phase, integrity of properties and classes is checked and as well as information from mappings & various Ontologies is combined. In case of violation or infringement detected, the process is restarted and prohibits the formation of violating alignment.

4. EVALUATION OF MATCHERS

Evaluation of the matchers basically aims to find out the strengths & weaknesses of various systems, which in long term can be used to make improvements to the systems. The evaluation process comprises of running matching over test suite and then compares the outcomes achieved with the fancied ones or we can say the reference alignments. These test suites consists of benchmark tests, made specifically for evaluation of the matchers.

For utilization in ontology alignment evaluation, information retrieval strategies have been adopted, especially precision & recall. Precision speaks about the correctness of discovered mappings i-e the ratio of correct alignments to total discovered alignments; whereas Recall measures the extent up to which the alignment is complete i-e the ratio of correct alignments to the total alignments. The exact definition of the terminologies can be as

Precision: $P(A,R) = \frac{|R \cap A|}{|A|}$ … equ no. (1)

Recall: $P(A,R) = \frac{|R \cap A|}{|R|}$ … equ no. (2)

Where A in *(A, R) in equations (1) and (2)* stands for the accurate alignments obtained from the matcher and R for total produced alignments. And there is inverse relation between precision and recall i-e for high precision recall will be less and vice versa.

5. RESULTS

For calculating precision and recall, a set of benchmark Ontologies is available online at the website of OAEI (Ontology Alignment Evaluation Initiative) http://oaei.ontologymatching.org, available in 3 groups i-e 1xx, 2xx & 3xx. 1xx group contains four simple Ontologies, 2xx group contains about 66 Ontologies from different domains where some information has been altered in them and 3xx group contains 4 Ontologies which are merely bibliographies.

5.1 RiMOM

The RiMOM matcher is applied over three groups of benchmark Ontologies i-e 1xx, 2xx and 3xx and the values of precision and recall have been calculated. The outcome is presented in Figure 1. Both precision and recall values for 1xx group of benchmark Ontologies is 1.00. Both the values are really good because in 1xx group reference Ontologies are compared to themselves. The precision and recall values for 2xx group of benchmark Ontologies are a bit less because some information such as entity names, classes and properties etc. in 2xx benchmark Ontologies is discarded.0.97 is the value of precision and 0.86 is the recall value in this group. As 3xx group of benchmark consists of real life bibliographic reference Ontologies so the precision and recall reduces a bit more here, & respectively precision& recall get the values 0.69 and 0.80.

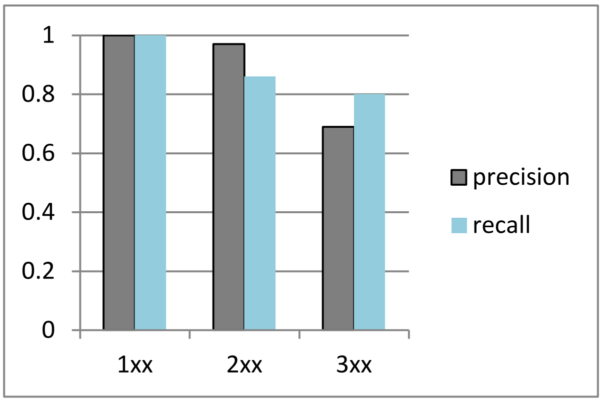

Figure 1 RiMOM Precision and Recall

5.2 Falcon AO

Falcon-AO is run for the three groups of benchmark Ontologies i-e 1xx, 2xx and 3xx and the values of precision and recall have been calculated. The outcome is presented in Figure 2. Both precision and recall values for 1xx group of benchmark Ontologies is 1.00. Both the values are really good because in 1xx group reference Ontologies are compared to themselves. The precision and recall values for 2xx group of benchmark Ontologies are a bit less because some information such as entity names, classes and properties etc. In 2xx benchmark Ontologies is discarded. 0.92 is the value of precision and 0.85 is the recall value in this group. As 3xx group of benchmark consists of real life bibliographic reference Ontologies so the precision and recall reduces a bit more here, & respectively precision & recall get the values 0.89 and 0.79.

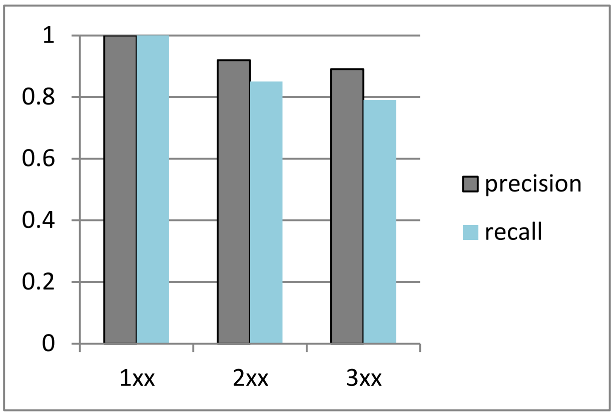

Figure 1 Falcon AO Precision & Recall

5.3 LILY

LILY in the same way is also run for the three groups of benchmark Ontologies i-e 1xx, 2xx and 3xx and the values of precision and recall have been calculated. The outcome is presented in Figure 3. Both precision and recall values for 1xx group of benchmark Ontologies is 1.00. Both the values are really good because in 1xx group reference Ontologies are compared to themselves. The precision and recall values for 2xx group of benchmark Ontologies are a bit less because some information such as entity names, classes and properties etc. in 2xx benchmark Ontologies is discarded. 0.97 is the value of precision and 0.89 is the recall value in this group. As 3xx group of benchmark consists of real life bibliographic reference Ontologies so the precision and recall reduces a bit more here, & respectively precision & recall get the values 0.81 and 0.80.

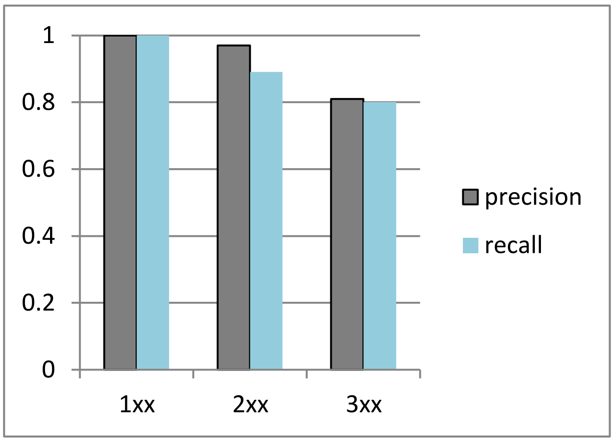

Figure 2 LILY Precision & Recall

5.4 ASMOV

ASMOV is short of "Automated Semantic Matching of

Ontologies with Verification”. Similarly ASMOV is applied for the three groups of benchmark Ontologies i-e 1xx, 2xx and 3xx and the values of precision and recall have been calculated. The outcome is presented in Figure 4. Both precision and recall values for 1xx group of benchmark Ontologies is 1.00. Both the values are really good because in 1xx group reference Ontologies are compared to themselves. The precision and recall values for 2xx group of benchmark Ontologies are a bit less because some information such as entity names, classes and properties etc. in 2xx benchmark Ontologies is discarded. 0.95 is the value of precision and 0.90 is the recall value in this group. As 3xx group of benchmark consists of real life bibliographic reference Ontologies so the precision and recall reduces a bit more here, & respectively precision & recall get the values 0.85 and 0.82.

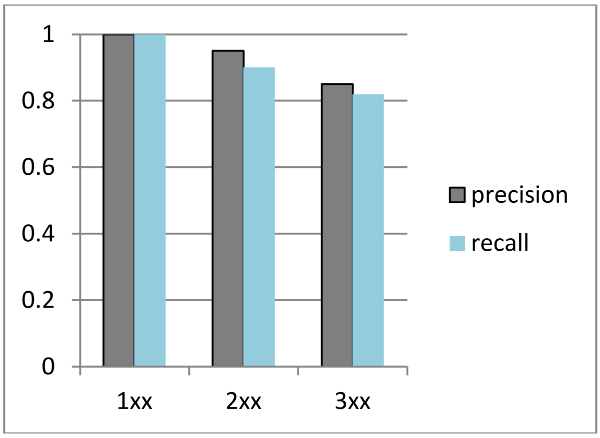

Figure 3 ASMOV Precision & Recall

5.5 OLA 2

OLA2 is applied for the three groups of benchmark Ontologies i-e 1xx, 2xx and 3xx and the values of precision and recall have been calculated. The outcome is presented in Figure 5. Both precision and recall values for 1xx group of benchmark Ontologies is 1.00. Both the values are really good because in 1xx group reference Ontologies are compared to themselves. The precision and recall values for 2xx group of benchmark Ontologies are a bit less because some information such as entity names, classes and properties etc. in 2xx benchmark Ontologies is discarded. 0.91 is the value of precision and 0.86 is the recall value in this group. As 3xx group of benchmark consists of real life bibliographic reference Ontologies so the precision and recall reduces a bit more here, & respectively precision & recall get the values 0.63 and 0.76.

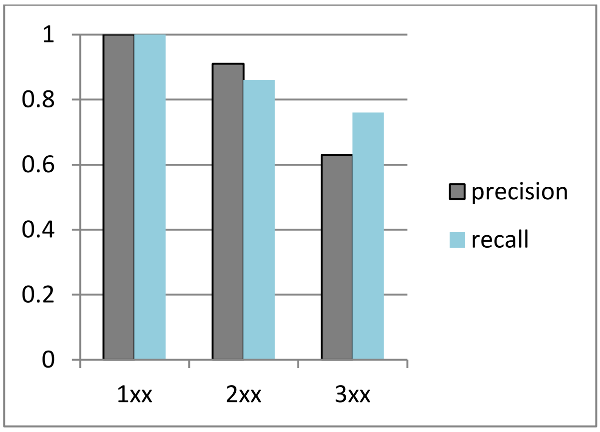

Figure 4 OLA 2 Precision & Recall

6. COMPARISON OF RESULTS OBTAINED

Various Structural matchers as discussed above are used for testing and evaluating over the available benchmark Ontologies at OAEI website. Values of precision & recall are calculated for each group respectively and the comparison of their outcomes is shown in Figures 6 and 7 below for benchmark group 2xx and 3xx respectively.

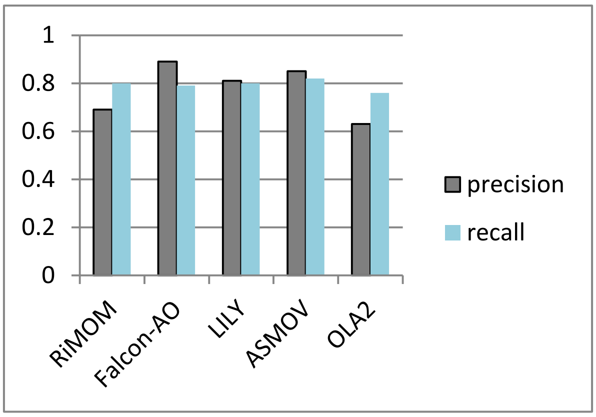

Figure 6 Precision & Recall for 2xx group of benchmark Ontologies

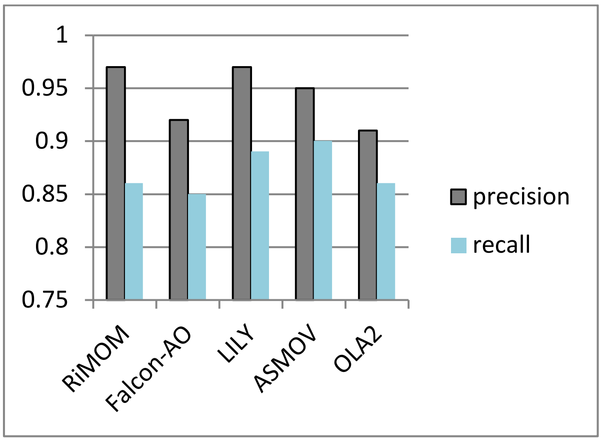

Figure 7 Precision & Recall for 3xx group of benchmark Ontologies

Graph in Figure 6 shows comparison of results of the matchers for 2xx group of Ontologies where there can be seen no particular order as some performed well in terms of precision and some performed well in terms of recall. If we look at precision then RiMOM and LILY are almost at the same level and ASMOV a little behind them. Falcon-AO and OLA2 lag behind marginally. But if values of recall are considered then ASMOV is leading the list, followed by LILY and then RiMOM. Falcon and OLA2 also have better rate in this race. The Recall overall in this group is relatively less because of the fact that some of the information in Ontologies is discarded in this group.

Figure 7 which is the result of matchers for 3xx group which contains real life bibliographic reference Ontologies, due to the reason values of precision and recall overall get reduces here. In terms of precision Falcon-AO has performed exceptionally here, but if precision and recall are considered collectively then surely ASMOV is the leading matcher. Recall is almost at same level for all the matchers but in terms of precision Falcon-AO leads and OLA2 comes the last.

7. CONCLUSION

Ontology Engineers have proposed and executed different matching strategies to connect the semantic gap between heterogeneous Ontologies. In any case, more change and examination is required in these strategies with a specific end goal to accomplish better recall & precision for the alignment sets. This document presents alignment of Structural Ontologies which utilizes matcher to guide entities of two unique Ontologies. The result shows that RiMOM and LILY lead the list in alignment precision, but if Recall is considered then ASMOV is leading the list of matchers considered. In Future more such structural matchers can be used for evaluation purposes and be compared in terms of precision, recall & F-measure, with the existing ones.

8. ACKNOWLEDGMENTS

Deepest appreciation and thanks to **Dr. Sadaqat Jan,** who has the attitude and substance of a genius. He has been instrumental in every stage, supervising this work.

9. REFERENCES

[1] Berners-Lee, Tim, James Hendler, and Ora Lassila. "The semantic web." *Scientific American* 284.5 (2001): 28-37.

[2] AH Doan, J Madhavan, P Domingos, A Halevy, "Learning to map between Ontologies on the semantic web.," in *In Proceedings of the 11th international conference on World Wide Web,pp. 662-673. ACM*, Chicago, 2002..

[3] Holi, Markus, and Eeru Hyvönen. "A method for modeling uncertainty in semantic web taxonomies." *Proceedings of the 13th international World Wide Web conference on Alternate track papers & posters*. ACM, 2004

[4] Decker, Stefan, Frank Van Harmelen, and Jeen Broekstra. "The semantic Web-on the respective Role of XML and RDF."

[5] McGuinness, Deborah L., and Frank Van Harmelen. "OWL web ontology language overview." *W3C recommendation* 10.10 (2004): 2004.

[6] Shvaiko, Pavel, and Jérôme Euzenat. "Ontology matching: state of the art and future challenges." *IEEE Transactions on knowledge and data engineering* 25.1 (2013): 158-176.

[7] Xu, Tianren, and Tianwei Xu. "An approach for ontology mapping based on semantic and structural information in Ontologies." *Information Management, Innovation Management and Industrial Engineering (ICIII), 2010 International Conference on*. Vol. 3. IEEE, 2010..

[8] Leung, Nelson KY, et al. "Ontology matching techniques: a 3-tier classification framework." (2009)..

[9] Del Vescovo, Chiara, et al. "The modular structure of an ontology: Atomic decomposition." *IJCAI Proceedings-International Joint Conference on Artificial Intelligence*. Vol. 22. No. 3. 2011

[10] Li, Juanzi, et al. "RiMOM: A dynamic multistrategy ontology alignment framework." *IEEE Transactions on Knowledge and Data Engineering* 21.8 (2009): 1218-1232.

[11] Wang, Peng, and Baowen Xu. "LILY: the results for the ontology alignment contest OAEI 2007." *Proceedings of the 2nd International Conference on Ontology Matching-Volume 304*. CEUR-WS. org, 2007.

[12] Kengue, Jean François Djoufak, Jérôme Euzenat, and Petko Valtchev. "OLA in the OAEI 2007 evaluation contest." *Proceedings of the 2nd International Conference on Ontology Matching-Volume 304*. CEUR-WS. org, 2007.

[13] J. Euzenat and P.Valtchev. "Similarity-based ontology alignment in OWL-lite"*In Proceedings of the European Conference on Artificial Intelligence (ECAI)*, 333–337, 2004.

[14] Jean-Mary, Yves R., and Mansur R. Kabuka. "ASMOV results for OAEI 2007." *Proceedings of the 2nd International Conference on Ontology Matching-Volume 304*. CEUR-WS. org, 2007.

[15] Jean-Mary, Yves, and Mansur Kabuka. "Asmov: Ontology alignment with semantic validation." *Joint SWDB-ODBIS Workshop on Semantics, Ontologies, Databases, Vienna, Austria*, 15–20, 2007.

[16] Ehrig, Marc, and Steffen Staab. "QOM–quick ontology mapping." *International Semantic Web Conference*. Springer Berlin Heidelberg, 2004.

[17] Jian, Ningsheng, et al. "Falcon-ao: Aligning Ontologies with falcon." *Workshop on Integrating Ontologies (K-CAP 2005)*. 2005.