Ontology Matching: Current trends among practitioners

Lorena Otero-Cerdeira, Francisco J. Rodríguez-Martínez and Alma
Gómez-Rodríguez

Departamento de Informática,
Ed. Politécnico, Campus As Lagoas,
Universidad de Vigo
Ourense, Spain,
{locerdeira, franjrm, alma}@uvigo.es

Abstract Ontology matching is an important and prolific field of research, reflecting the growing interest of the community. This work makes a study over the different approaches and lines of research within the field of Ontology Matching, in order to help and guide new practitioners to approach the field and to determine possible lines of research. The study is based on a survey which has being conducted among the practitioners and which addresses different questions including the background, the future goals and the main achievements of the researcher. The purpose of such survey is to gain knowledge about the current state of art within the ontology matching field and on how such techniques are applied to real-life environments. From the results, the main conclusion achieved is that most researchers share the same concerns about the necessity of practical application of ontology matching techniques, due to the existence of too many theoretical solutions but few applied ones.

1 Introduction

Ontology matching is a complex process that helps in reducing the semantic gap between different overlapping representations of the same domain [3]. The existence of such different representations obeys to the natural human instinct to have different perspectives and hence to model problems differently. When these domains are represented using ontologies, the solution typically involves the use of ontology matching techniques. Ontologies and ontology matching techniques are an increasing trend as ontologies provide probably the most interesting opportunity to encode meaning of information and integrate systems sharing different semantics or knowledge.

Last decade constitutes a period of extensive research in this field. Nowadays, far from dying down, the activity seems to be increasing and new publications in the field are continuously being released. This reflects the general interest in ontology matching and constitutes a challenge for new researchers, because of the huge amount of information available.

The general approach for dealing with such overwhelming volume of information is the use of papers that successfully presented the state-of-the-art in the field, such as, [5], [4], [8], [1], [9], [6]. However, these approaches are usually influenced by their authors, and therefore, present a particular and personal, and, in general, very valuable, view of the field. In addition, these works try to provide a general view of the field, including fields of applications, tools, algorithms, evaluation techniques, etc, but, in general, they do no address specifically what are the main concerns of the researchers in the field.

This paper tries to approach in a different way the current state of art in the field. That is, by asking the researches on what are their expectations and challenges in Ontology Matching field. The answers of the research community have been stored and analyze to describe the general concerns and challenges shared by the people who answer the questionnaire back. We consider this approach valuable, because it make a photograph, not influenced by the authors, of the current trends in the research field.

The remainder of the paper is organized as follows. In section 2 we describe the basic characteristics of the survey, that is, how, to whom and when it has been made, and, of course, what has been asked. Next, in section 3, the results of the study are presented and detailed. Section 4 introduced the limitations and analyzes the conclusions that may be achieved from the results. Finally, in section 5 we present the concluding remarks and directions for future work.

2 Survey Description

As stated before this work is based on a survey made in the past months which try to clarify general questions related to ontology matching field of research. Such concerns were mainly related to the current state of the research on ontology matching and its application in real-life projects.

2.1 Participants

The participants in the survey were selected among those taking part at the Ontology Alignment Evaluation Initiative (OAEI) contests [7]. OAEI is a co-ordinated international initiative to forge consensus for evaluation of ontology matching methods. They propose a benchmark to evaluate the methods and organize a yearly evaluation event [2].

The researchers were individually contacted by email and asked to answer the questionnaire shown in Table 1. The e-mails have been sent in a two month period, from December 2013 to February 2014. Even though the participants were directly contacted, their identities and responses were strictly confidential and only available to the team conducting the survey. Out of the 288 experts contacted, we received 46 replies.

2.2 Survey Design

The survey was composed of 8 short open-ended questions. We consider better a this approach than introducing multiple-choice questions, because we did not want to influence at any degree the answers provided by the participants.

The questions can be classified into three groups.

- *Background questions.* They are questions from 1 to 3. These questions are included in the questionnaire to assess the suitability of the participants and contextualize the answers they may provide.
- *Research field questions.* The questions 4, 5 and 7 are of this kind. They are designed to gain knowledge about the current fields and topics that have become more attractive to the research community,
- *Future challenges questions*, which include questions 6 and 8 and identify, according to practitioners' point of view, the main challenges to be addressed and the potential expansion fields of ontology matching.

Number	Question	Type
1.	How long have you been researching in Ontology Matching?	Background
2.	What are your main purposes to do it?	Background
3.	How many research papers have you written on topics related to Ontology Matching?	Background
4.	Within the Ontology Matching field, in which particular topic are you currently working on?	Research field
5.	From your point of view, which are the main fields where the research on Ontology Matching is currently being applied?	Research field
6.	According to your expertise, which are the main challenges that are still to be addressed?	Future challenges
7.	Will you continue to research in Ontology Matching?Why?	Research field
8.	In which fields do you believe that Ontology Matching could also be used?	Future challenges

Table 1. Questionnaire used for the survey

3 Survey Results

Out of the 46 answers that we received, only 5 declined the participation in the survey, one answered it partially and 13 have not been researching on ontology matching for a while. Some of the researchers that have stopped working in the field recently, answered the questionnaire anyhow, as their contribution was still relevant. Others suggested a more appropriate contact within their groups to redirect the requests, half of which were answered back. In the end, the final amount of replies was cut down to 33 actual answers with profitable information.

3.1 Background questions

The background of the participants in the research is quite broad and varied, and it includes different types of researchers in the field. From a temporal point of view, there are those who have started in the field more than a decade ago but also those who have recently started to work in this field. Nevertheless, the main part (79%) of researches who answer this question have been researching in the field for five years or more. A more detailed study of this value is introduced in Figure 1, where the number of years doing research in the field is shown in relation with the number of researchers. The average number of years working in the field is slightly superior than 7.
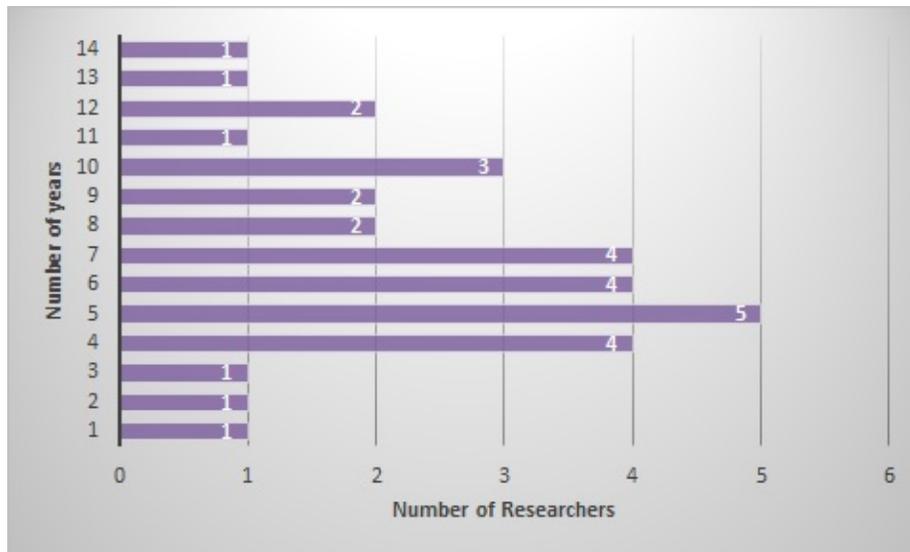

Figure 1. Number of researchers in relation with the number of years working in the field

The way of starting in the research in the field has a great variety of answers, some researchers have directly tackled the ontology matching problem by focusing on very specific topics such as *assessing the impact of using different similarity metrics in different ontology matching tasks*, *aligning large ontologies* or *improving ontology matching by using reasoning*, while others arrived to ontology matching as a support tool for matters such as *data integration*, *semantic interoperability* or *telecommunications systems interoperability*.

Anyhow the suitability of the participants is more than satisfactory as they account on their own for above 460 publications of different types linked to this subject.

As stated before, the main objective of this kind of questions is to have a general view of the suitability of the population that answer the questionnaire.

In general terms, we consider that the participants are qualified, because of the important number of papers they have produced and have enough expertise, as they refer, in majority, more than five years of expertise in the field.

3.2 Research field questions

The research field questions included in the questionnaire have as main aim to obtain a quick view of the more relevant working fields, as well as, other fields where the experts think that ontology matching could be applied. Out of the answers sent by the participants, we could determine that a high percentage of them (63%) is working in any of these four topics: *instance matching*, *user involvement in the matching process*, *data interlinking* and *discovery of different types of correspondences, not 1:1 equivalence relations*. The rest of the respondents mentioned other more specific topics such as *parallel ontology matching*, *large-scale ontology matching*, *ontology matching negotiation* and *mapping reuse*. However, besides these main research topic, most of the answers include also other topics such as *similarity metrics* and *combinations of methods to improve the coverage of ontology matching* as a way to enhance or support their basic research interests.

Regarding the question about the fields where ontology matching is being applied, the answers achieve an important consensus among researchers. Attending the practical application of ontology matching, the answers point to *the medical* and *life science domain*. Other more general or theoretical answers mention *data integration* and *interoperability* as the fields where ontology matching is being applied.

It is remarkable that several (roughly speaking a 20%) researchers pointed out that nowadays the use of ontology matching techniques is reserved to spot cases and that the research at this time seems merely foundational. However, they also agree that ontology matching can be applied in any field where there are two parties that need to communicate and that employ potentially different protocols, being this way the list of use cases potentially long. From this answers, we think that it can be concluded that, researches still think that the discipline is quite immature and has a big possibility of projection in future years.

Finally, when questioned about whether they would continue to research on this field, the majority, 63.64%, confirmed they would follow with present or related research lines. The reasons for continuing in this line of research include answers such as: *there are still plenty of challenges to address* and *the development of new domains will sparkle new matching problems*. On the contrary, 30.30% of the respondents stated that they would, if not yet, change subjects. Among the reasons to do so, some mentioned they have moved to other related fields such as *linking open information systems* or *knowledge transformation*, while others definitely quit the field claiming the *lack of usefulness for real applications* or *the little incentive coming from the application side*. A small percentage, 6.06% were still considering whether to change subject or not. These results may be interpreted as a confirmation of the idea previously stated, that is, that in the

opinion of researchers there is still a broad spectrum of works to be done in the field.

3.3 Future challenges questions

These group of questions were included in the questionnaire to gain knowledge about how practitioners see the future evolution of the field and the main challenges still to be addressed. We consider, the answers provided as really useful, because they identify a variety of challenges to address and quote several new fields where matching techniques could also be used. In consequence, these ideas could be used to guide the research lines adopted by different research groups.

Regarding the main challenges still to be addressed in the ontology matching field, most respondents agree on the need to *automatically discover complex relations, instead of 1:1*, to *correctly align large ontologies* and to *focus on applying automatically created mappings to practical applications*.

Other topics that arose in the responses were not supported by so many respondents but they point out anyway challenges that need to be addressed, such as:

– Automated acquisition of reference alignment for evaluating large scale matching systems
– Creating large datasets to asses matching algorithms
– Construction of good tools that are easy to use for non-experts
– Develop high quality and fast intelligent combinations of string-based and new semantic-similarity measures
– Holistic ontology matching
– How to effectively complement automatic computation with human validation
– How to minimize involvement of users when turning matches into mapping
– Human readable explanations for matches
– Improving the mapping process through semi-automatic machine learning
– Integration of domain knowledge into alignment techniques
– Learning what metrics to choose in which scenario
– Precision and Recall of automatic methods
– Scalability and parallelization of the matching
– Semantic mapping

Some practitioners, used the answer to this question to mention certain current situations that they consider a mistake. For instance, some answers point out that the fact that most approaches to ontology matching focus only on lexicographic and structural information, while they consider that language is more complex than that. And, as a consequence, achieving a perfect precision and recall is impossible in real life applications. Other aspect they complained about was benchmarks habitually used to test the matching systems, that is OAEI benchmarks. They claim that even if they are really useful, their main drawback is that there are yet too many artificial datasets and tasks in it. Other practitioners, use the answers as a general complain about current situation science

culture, pointing out that tasks such as creating and maintaining one tool are not rewarded at all. This produces the effect that in most cases, a prototype is created focused in a particular solution or paper and after it is abandoned.

Finally, regarding the fields where ontology matching could also be applied we have obtain two types of answers. Some practitioners gave a fuzzy answer mentioning that matching techniques could be used *practically anywhere where is no standard for information exchange and where the domains are open to adopt ontology approaches* or in broad sense in *any information related field*.

On the contrary, others actually mentioned specific fields to apply these techniques, such as: *bioinformatics, information systems, e-commerce, web services, intrusion detection systems, cultural heritage, library science, government, education, banking, personal and social data management, law, etc.* Most of them agree in the fact that if in these fields ontologies and ontology matching techniques are not already in use is due to a lack of information on the potential benefits.

4   Discussion

4.1   Limitations of the Survey

Although we consider the results obtained by this study very valuable, there are, of course, limitations. The foremost is the sampling size and the population. Although we feel that our 33 final responses offered a wide variety of useful remarks and points of view, it is true that the sample is still small and hence our analysis may be biased.

Besides, in an effort to prevent the questions from influencing the answers and to obtain as much information as possible, we have defined the questionnaire with 8 open-ended questions. This fact, possibly together with the way some questions were posed, led us to obtaining answers that, however interesting, did not exactly match what we expected. Besides, these open questions may make that the researchers perceive the questionnaire as a time consuming task, so that, they decide not to fill it, and the number of answers has result not to be very high.

Also, the participants targeted were obtained from the participants at the OAEI contests, most of which are academically oriented, therefore our survey may be biased towards academical researchers rather than a balance between academical and industrial researchers.

4.2   Results Discussion

Despite the limitations previously exposed, we consider the results of this survey valuable and we think that it provides a first high level abstraction view of the field of research. The answers provided by researchers who fill up the questionnaire were accurate and introduce a high level of detail, suggesting a detailed

work of the respondents. Moreover, the reliability of data is supported by researchers experience and by their knowledge of the field, as can be proved, for instance, by the huge amount of papers produced.

In general terms, we consider that the results show that researchers think that ontology matching is still a field in its first stages of development and there is a lot of work to do. This idea is supported by the big number of researchers who have the intention of going on investigating in the field, and the number of different fields of challenges in research introduced. Moreover, the researchers have the impression that there is still many challenges to face in the field, as it is reflected in their answers. Attending the people who do not go on making research in the field, in many cases they have change to a new field as an evolution of that investigation, and, in general, they do not justify the change in the lack of interests or promising results in the field.

The results introduce also a wide range of different and specific fields the researchers are approaching within matching ontologies. The broad spectrum of such research, with important practical applications, guarantees the future of the field, in our opinion.

4.3   Correlation between responses

As a final point of the study we have revisited these answers looking for deeper connections between the answers. We try to obtain relationships between responses, to address questions such as: *Is there any relation between how long a researcher has been working in a subject and the amount of publications?*, *Which type of researchers are more prone to quitting?*, etc.

Regarding first question, although a statistical analysis must be done, it seems to be a direct relation between the number of papers produced by a researcher and the number of years in the field, which is not surprising. Nevertheless, there are some short-term researchers who has an important production. This may be due to the previous expertise of the research group they have joined.

Attending people who has leaved ontology matching as field of research, there seems to be a direct relationship between the answer to the challenges to be addressed in the field and their leaving; many of those researchers have less confidence on the future of the field.

Despite this first qualitative approach in the relation among responses, we plan to make a deeper analysis, including quantitative and statistical values.

5   Conclusions and Future Work

By means of this work we have provided a general overview of the ontology matching field in the present and what are the expected future trends for practitioners.

The approach chosen in this paper to clarify the ideas related to matching ontology field is to ask openly to the research community, by means of a practitioner-oriented survey. The purpose of such survey is to gain knowledge

about the current state of the ontology matching field and the application of such techniques to real-life environments. We have noticed that most researchers share the same concerns about the practical application of the ontology matching techniques, and the problem of having too many theoretical solutions but few applied ones.

The results of the survey can be used as a starting point for new practitioners to get a general idea of the research field but also to help in deciding on research lines, hopefully by tackling some of the challenges highlighted in the survey.

Besides, as stated before, we consider that using a multiple choice questionnaire will obtain a higher number of responses. In the future, we plan to complete this study in this way. The main results obtained from this survey can be used as a guide to design the new study, using the topics posed by the researchers as basis for making multiple choice questions. Maybe, using this kind of questions, which can be perceived by researchers as easier and quicker to answer will laid to a bigger number of participants.

Due to the nature of the survey, with open-ended questions, a deeper analysis of the information obtained can still be done, addressing, for instance, the interrelations between different answers. Although a first qualitative attempt has been made in this work, we plan to do it in a more formal and quantitative way, using the mathematical tools to prove our conclusions.

References

1. Jérôme Euzenat. State of the art on Ontology Alignment. *Knowledge Web*, 2, 2004.
2. Jérôme Euzenat, Christian Meilicke, Heiner Stuckenschmidt, Pavel Shvaiko, and Cássia Trojahn. Ontology alignment evaluation initiative: Six years of experience. In Stefano Spaccapietra, editor, *Journal on Data Semantics XV*, volume 6720 of *Lecture Notes in Computer Science*, pages 158–192. Springer Berlin Heidelberg, 2011.
3. Jérôme Euzenat and Pavel Shvaiko. *Ontology matching*. Springer-Verlag, Heidelberg (DE), 2nd edition, 2013.
4. Fabien Gandon. Ontology engineering: A survey and a return on experience. Technical report, Institut de Recherche en Informatique et Automatique, 2002.
5. Asunción Gómez-Pérez. Ontological engineering: State of the art. Technical report, Ontology Engineering Group, 1999.
6. Y. Kalfoglou and M. Schorlemmer. Ontology mapping: The state of the art. *Science*, 2:3, 2003.
7. *Ontology Alignment Evaluation Initiative*, 2014 (accessed July 17, 2014).
8. M. Shamsfard and A. Barforoush. The State of the Art in Ontology Learning: A Framework for Comparison. *The Knowledge Engineering Review*, 18, 2003.
9. P. Shvaiko and J. Euzenat. Ontology Matching: State of the Art and Future Challenges. *Knowledge and Data Engineering, IEEE Transactions on*, 25(1):158–176, 2013.