I-Match and OntoIdea results for OAEI 2017

Abderrahmane Khiat and Maximilian Mackeprang

Human-Centered Computing Lab, Freie Universität Berlin, Germany
abderrahmane.khiat@fu-berlin.de, maximilian.mackeprang@fu-berlin.de

Abstract. Presenting a set of similar or diverse ideas during the idea generation process leads ideators to come-up with more creative and diverse ideas. However, to better assess the similarity between the ideas, we designed two matching systems, namely I-Match and OntoIdea. In the context of the idea generation process, each idea is represented by a set of instances from DBpedia describing the main concepts of the idea. Then, the developed matching systems are applied to compute the similarity between a set of instances that represent the ideas. The purpose of our participation at OAEI is to evaluate our designed instance matching algorithm in order to apply it to assess the similarity between ideas. The results obtained for the first participation of I-Match and OntoIdea systems at OAEI 2017, on different instance matching tracks are so far quite promising.

Keywords: Collaborative Ideation, Semantic Annotation, Ontology, Instance Matching, OAEI.

1 Introduction

The idea generation process is the key part of innovation. This process aims to generate ideas to solve problems and challenges. A promising approach for supporting such process is the ”brainstorming method” [3]. This method seeks to increase the number of ideas based on ideas of collaborating individuals while restricting criticism.

In addition to leveraging the crowd [10], prior work has shown that generating ideas that are both creative and diverse can be greatly enhanced through presenting inspirational examples [6]. However, a major issue is ”how to find inspiring ideas from hundreds” [9]. To overcome this challenge, research has shown three ways of selecting a set of inspiring examples systematically [4, 5]: (1) presenting diverse ideas, (2) presenting similar ideas and (3) visualizing all ideas.

Our work is in line with approaches that assess the diversity (i.e. low similarity rating) of inspiring examples automatically [8]. However, assessing similarity between ideas is challenging due to the form of the ideas, i.e. the ideas are described in a short unstructured text.

To solve this problem, we propose another strategy from our prior work proposed in [2]. This strategy consists of two main parts: (1) concepts annotation and (2) an instance matching mechanism. Firstly, we annotated the main concepts of an idea with instances from DBpedia, a validation through user-based selection of images are carried out in order to obtain the right meaning of the identified concepts. Secondly, these annotated concepts with a set of instances are used as a support to calculate the similarity between ideas using an instance matching system. Using our approach, we can

assess the similarity of two ideas, which can then be used further to select (1) a set of diverse ideas (low similarity rating), (2) a set of similar ideas (high similarity rating) that inspire the user to generate more creative ideas. Furthermore, we use the similarity ratings obtained to provide a visualisation of the solution space to give ideators an overview of the collaborative effort.

In this paper, we focus on the matching part of the proposed solution by describing our two instance matching systems I-Match and OntoIdea. The designed systems implement an enhancing algorithm that we proposed in our previous work [1]. The proposed algorithm extracts first all information about the two instances to be matched and normalizes them using NLP. Then, it applies edit distance as a matcher to calculate the similarities between the normalized information. Finally, the approach selects the equivalent instances based on the maximum of shared information between the two instances.

2 Instance Matching Algorithm

We summarize the algorithm of our developed systems to provide a general idea of the proposed solution. It consists of the following successive phases:

2.1 Extraction and Normalization

The system extracts from each individual Ii P_1 m_1; P_2 m_2,... a set of information m_1, m_2, ... using different properties P_1, P_2, ... Then, NLP techniques are applied to normalize these information. In particular, three pre-processing steps are performed: (1) case conversion (conversion of all words in same upper or lower case) (2) lemmatization stemming and (3) stop word elimination. Since String based algorithm is used to calculate the similarities between information, these steps are necessary.

2.2 Similarity Calculation

In this step, the system calculates the similarities between the normalized informations using edit distance as string matcher. Our system selects the maximum similarity values calculated between different informations by edit distance. If two informations are the same (based on maximum similarity values) the counter is incremented to 1, etc.

2.3 Identification

Finally, we apply a filter on maximum counter values in order to select the correspondences which mean that the selected correspondences (equivalent individuals) are those who share maximum informations.

3 Experimentation

The I-Match and Ontoidea systems participated only for instance matching tracks of OAEI 2017 evaluation campaign. For the results Please refer to the following website:
`http://oaei.ontologymatching.org/2017/results/index.html`.

4 Conclusion

In this paper, we have introduced I-Match and OntoIdea, two systems specially designed to compute similarity between instances. The proposed algorithm is useful, especially when the instances contain terminological information. The developed systems provide a quite promising results, thus, we will be applied in the context of the idea generation process to asses similarity between ideas.

As future perspective, we attempt to apply enhance our instance matching algorithm especially for DORUMUS track.

Acknowledgements The authors acknowledge the financial support of the Federal Ministry of Education and Research of Germany in the framework of Ideas to Market (project number 03IO1617).

References

1. A. Khiat, M. Benaissa and M. A. Belfdhal, "STRIM results for OAEI 2015 instance matching evaluation". In Proceedings of the 10th International Workshop on Ontology Matching co-located with the 14th International Semantic Web Conference (ISWC 2015), USA, 2015.
2. A. Khiat, M. Mackeprang, C. Mller-Birn, "Semantic Annotation for Enhancing Collaborative Ideation", Semantics 2017, Netherlands.
3. Alex F. Osborn. 1963. Applied imagination; principles and procedures of creative problem-solving. Scribner, NY, 1963.
4. Pao Siangliulue, Kenneth Arnold, Krzysztof Gajos and Steven Dow. 2015. Toward collaborative ideation at scale-leveraging ideas from others to generate more creative and diverse ideas". In Proceedings of CSCW, 2015.
5. Pao Siangliulue, Joel Chan, Steven Dow and Krzysztof Gajos. 2016. IdeaHound: improving large-scale collaborative ideation with crowd-powered real-time semantic modeling. In Proceedings of UIST.
6. Richard Marsh, Joshua Landau and Jason Hicks. 1996. How examples may (and may not) constrain creativity. Memory and Cognition, 24 (5), pp. 669–680.
7. Nicholas Kohn and Steven Smith, 2011. Collaborative fixation: Effects of others' ideas on brainstorming. Applied Cognitive Psychology 25 (3), pp. 359–371.
8. Joel Chan, Pao Siangliulue, Denisa McDonald, Ruixue Liu, Reza Moradinezhad, Safa Aman, Erin Solovey, Krzysztof Gajos and Steven Dow. 2017. Semantically Far Inspirations Considered Harmful? Accounting for Cognitive States in Collaborative Ideation. In Proceedings of C& C'17.
9. Elahe Javadi, Judith Gebauer and Joseph Mahoney. 2013. The impact of user interface design on idea integration in electronic brainstorming: an attention based view. Journal of AIS 14, pp. 1–21.
10. Victor Girotto, Erin Walker and Winslow Burleson. 2017. The effect of peripheral micro-tasks on crowd ideation. In Proceedings of CHI '17, pp. 1843–1854.
11. Osvald Bjelland and Robert Wood. 2008. An inside view of IBM's' innovation jam'". MIT Sloan Management Review 50 (1), pp. 32–40.
12. Joel Chan, Steven Dang and Steven Dow. 2016. IdeaGens: enabling expert facilitation of crowd brainstorming. In Proceedings of CSCW, pp. 13–16.