Crowd sourcing through social gaming for community driven ontology engineering, results and observations

Alloy Martin Chua, Roland Christian Chua, Arthur Vincent Dychiching, Tinmon Ang, Jose Lloyd Espiritu, Nathalie Rose Lim, Danny C. Cheng

College of Computer Studies, De La Salle University Manila Philippines
{alchua01, arvolution7000, avd_arthur, angtinmon}@yahoo.com, espiritu.lloyd@gmail.com, nats.lim@delasalle.ph, danny.cheng@dlsu.edu.ph

Abstract. In developing ontology, expert driven approaches lack the scalability to accommodate the vast amount of data on the web. As such, the community is being tapped to build ontologies to cope with highly dynamic data sources. Common problems (like difficulty of the task, quality of output, and incentives needed to motivate the community), as discussed by other authors, are considered. In this paper, we discuss observations on our approach to improve the quality and sustain community ontology refinement though the use of social gaming and interaction. Current observations show that profile and knowledge of the concept in question, understanding and expressivity of the relationships play a key role in the quality of the result.

Keywords: Matcher selection, self-organization, community driven ontology engineering, social gaming, incentives.

1 Introduction

Given the disadvantages of expert driven ontology engineering [1], communities of stakeholders would have to be involved in the engineering process to allow the capture of emergent data and concepts and keep pace with ontology evolution. Social Networks are suitable for this process as the members all share common background knowledge, goals, and interests**.** Researches by [2][3][4][6] have explored community-driven approaches or games-with-a-purpose to develop the ontology. However, these works do not consider the familiarity and perception of the user and their ability to provide quality feedback (with regards to the concept in question). Also, the motivation or incentive for the user to continuously provide input to the system and its sustainability in terms of application propagation and social influence is not fully considered. Our methodology is to present the engineering task as a Facebook game to verify a lightweight ontology extracted from delicious and consider common background knowledge or familiarity to concept. Different aspects of social influence (both direct and indirect) [5] are used to allow for sustainability and scalability of the system. In terms of sustainability, we refer to direct influence such as friend requests to participate, while for indirect influence, we look at general awareness of peer activities via public postings. For scalability, the community is

allowed to engineer the ontology and version the ontology on a community basis (small world graphs) and it uses direct and indirect peer influence to allow self monitoring and propagation of the application. For both cases, activities that promote social influence are tightly integrated to incentive schemes to motivate the community to perform those actions.

2 Results and Observations

During our two-month testing with 110 users subdivided into 4 community groupings, we encountered several issues and discovered some interesting results. It was seen that groups who are familiar with the area (topic) were able to participate more. These were also the users who, generally, return more than once; as opposed to the community that is from a different background. This result serves as a promising response to the assertion that selection of participants is important and knowledge of the tag in question by the community is necessary for them to provide feedback. Also, users are more capable of identifying erroneous relationships as opposed to validating them. The results show that the majority of the players "disagreed" with the initial discovered ontology. Perception and interpretation affect how concepts will be organized. Limiting our current discussion to simply identifying hierarchical relationships such as "kind-of", it can be seen that connections validated and invalidated differ between groups. For the non-computer science community, some users perceived video as a "kind-of" learning mechanism (when they meant to express "videos can be used for learning"). This was due to the lack of expressivity of the relationships in the game. It should be noted that no "kind-of" relationship between "video" and "learning" was produced from inputs in the computer science user community. It was observed perspectives and interpretation play key roles and user profile and social affinity can serve as some of the bases to limit and select users for participation. Finally, the ontology is able to stabilize and converge; however, it is still not known whether it is the stability of the domain that allowed the ontology to converge or the process itself.

References

1. Dieter Fensel. Ontology Management Semantic Web, Semantic Web Services, and Business Applications. [ed.] Pieter De Leenheer, Aldo de Moor, York Sure Martin Hepp. New York : Springer Science+Business Media, LLC, 233 Spring Street, New York, NY, 2008.
2. Siorpaes, Katharina, Prantner, Kathrin and Bachlechner, Daniel. Class Hierarchy for the e-tourism Ontology Version 8. e-Tourism. [Online] November 2004. [Cited: January 6, 2009.] http://e-tourism.deri.at/ont/.
3. Cardoso, Jorge and Sheth, Amit P. Semantic Web Services, Processes and Applications. United States of America : Springer Science+Business Media, LLC, 2006.
4. Steve LEUNG, Fuhua LIN, Dunwei WEN. Towards a Data-driven Ontology Engineering. Taipei : The 16th International Conference on Computers in Education, 2008.
5. Eytan Bakshy, Brian Karrer,Lada A. Adamic. Social Influence and the Diffusion of User-Created Content. California : ACM, 2009. pp. 325-334. ISBN:978-1-60558-458-4.
6. Hepp, Katharina Siorpaes and Martin. OntoGame: Weaving the Semantic Web by Online Gaming. Tenerife : Springer, 2008.