A Web Application Towards Semiotic-based Evaluation of Biomedical Ontologies

Muhamamd "Tuan" Amith, MS and Cui Tao, PhD

University of Texas Health Science Center,
School of Biomedical Informatics,
Houston, Texas, United States
{muhammad.f.amith,cui.tao}@uth.tmc.edu

Abstract. With the emerging importance of biomedical ontology research impacting Big Biomedical Data, there will be a need for knowledgebase evaluation that is both systematic and also engage a community of experts. This paper will introduce a prototype in production to evaluate the quality of formal ontologies through an online web tool, using the semiotic-influenced metrics to grade ontology quality. Here we introduce the Semiotic-based Evaluation Management System (SEMS), which is designed for (1) automatic generation of various quality scores of an uploaded ontology and recommendations for improvement for the ontology, and (2) a GUI for experts to conduct manual review and provide feedback. In this paper, we will discuss the current status of the tool as well as the course for its continued development.

Keywords: Ontology, Ontology Evaluation, Big Data, Semiotics, Knowledge Engineering, Knowledge Management

1 Introduction

An article from Scientific America [7] described Tim Berners Lee's vision of "Semantic Web" or linked meaningful data on the web. Most interesting, the use case that elaborated his vision was of health care scenario where patients can access health information through software agents. While the semantic web vision ("Web 3.0") may (or may not) be possible in the foreseeable future, the copious amount of health information on the World Wide Web is growing.

The massive growth of information have ushered a new discipline called Big Data. Big Data, according to the International Data Corporation, are "technologies describe a new generation of technologies and architectures, designed to economically extract *value* from very large *volumes* of a wide *variety* of data, by enabling the high-*velocity* capture, discovery, and/or analysis."[12] Healthcare has been dramatically affected by these new technologies, saving $300 billion dollars from analytics of Big Data [10], mitigating diseases, and affecting patient health behaviors [13]. Biomedical ontologies can and will play an important role in Big Data, specifically with consolidating *variety* in Big Data and introducing reasoning and analytical functions, from its success in "Small Data".

1.1 Ontology Evaluation

While this success is undisputed with the vast amount of biomedical literature highlighting biomedical ontologies for encoding knowledge and machine reasoning, the evaluation of ontologies is not settled [3][15]. Ontology evaluation "is the problem of assessing a given ontology from the point of view of a particular criterion of application, typically in order to determine which of several ontologies would best suit a particular purpose" [8]. For the last decade several ideas emerged addressing ontology evaluation [8], but none have appeared to be adopted universally by ontologists [15] [4]. Commonly, subject matter expert (SME) reviewers are sought to evaluate an ontology. However, this effort is a time and resource intensive approach, especially if the reviewers need to acclimate themselves on the topic of ontology and ontology-related tools, like Protégé [1]. A brief review of 200 randomly selected biomedical ontologies hosted on the National Center of Biomedical Ontologies' (NCBO) BioPortal reveal that only 17 out of 200 have a formal assessment described in a corresponding design paper, and the remaining do not have any explicit documented evaluation. With ontologies helping to further research in the biomedical domain, this highlights a strong need for evaluation for biomedical ontologies.

1.2 Semiotics in Ontology Evaluation

Ontologies are sometimes alluded as symbolic representations of a domain space where the terms signify the entities contained within the domain space. Likewise, semiotics is a study of meaning behind signs and symbols or representations, divided by three aspects - *pragmatic*, *syntactic*, and *semantic*. Burton-Jones, et al. introduced an ontology evaluation framework based on the theories of semiotics that utilized various metrics formulated within the three branches of semiotics, along with an additional branch called "*social*" [9]. Each evaluation criteria, based on the branches, asks if the ontology is "useful" (*pragmatic*), can it be "read" (*syntactic*), can it be "understood" (*semantic*), and can it be "trusted" (*social*). Each of these branches are decomposed to additional aspects that derive their values from data acquired from the ontology and external sources.

The authors of this paper introduce a Java web application, Semiotic Evaluation Management System (SEMS), to assist ontologists and reviewers to measure the qualities of their ontologies based on semiotic-inspired metrics. Previously, the authors have successfully utilized this framework in a previous study for patient-centered vaccine ontology [5], all while discovering ways to streamline the process in an all-in-one tool. The remaining sections introduce the implementation of the application and discuss further development for public release.

2 SEMS - Semiotic-based Evaluation Management System

SEMS is a web application tool designed to assist knowledge engineers to assess the strengths and weakness of their ontologies using the evaluation framework proposed by Burton-Jones, et al. The advantages of using their framework is that it is designed to accommodate various users, it is domain independent, and it is both uncomplicated and comprehensive. Their work included a C-based software, but the SEMS application would be the first public tool of its type that fuses semiotic-driven evaluation for ontologies, and an online platform to generate rapid evaluations of formal ontologies for the aim of promoting uniformed ontology evaluation. SEMS is an online software that is developed in Java with a modern HTML5 interface and hosted through an Apache Tomcat application server. It will permit ontologists to log in to their account and upload their encoded ontology file. SEMS will then calculate the various scores and allow for the user to invite SMEs to participate in a formal review process to verify the truthfulness from the ontology.

2.1 Prototype

Currently, the authors have implemented an operational prototype. Figure 1 captures the initial view when activating the application. First, an ontologist can specify the preferred evaluation criteria (i.e. *pragmatic*, *syntactic*, *semantic*) for assessing the ontology. Other options include identifying how the preprocessing should handle the annotations from the ontology file. These options include breaking camel cases; removing determiners, brackets, underscores, and dashes; and determining whether the ontology is using the labels annotation or unique identifiers.

Figure 2 shows a screen where the user uploads the ontology file to the server and the aftermath of the preprocessing functions. After the application alerts the user that the ontology has been uploaded, the user can click "Preprocess" and the server will output each of terms extracted, and its corresponding cleaned term based on the configuration. In addition, each term will have the number of word senses calculated from the WordNet component. For example, a term such as "Mild" would have three word senses, and a compounded term, like "Mild Fever" would have five word senses [1]. Concurrently, the server will also extract the statements evoked from the ontology in simple natural language statements for SMEs to evaluate the truthfulness.

After the ontology has been processed, the user can navigate using the left column to tab through the various scores available. Figure 3 shows one of the screens associated with the *syntactic* aspect with tabs to navigate to a sub-score. For each of the scores, the user can choose to include or exclude parts of the scoring depending on the purpose of the evaluation.

Also implemented, the software tool facilitates a knowlegebase review to determine the truthfulness of the ontology (Figure 4). A SME can scroll through all of the statements evoked from the ontology and denote whether the statement is true, false, or other. This is demonstrated by clicking the "Add Assessment" button beside the statement, which will display a pop-up to assess it (Figure 4).

[1] *Mild* has three word senses and *fever* has two. This adds to a total of five word senses for *mild fever*.

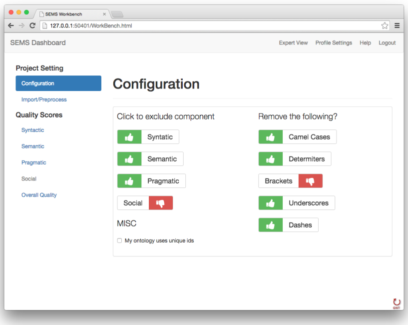

Fig. 1. Screenshot of processing configuration.

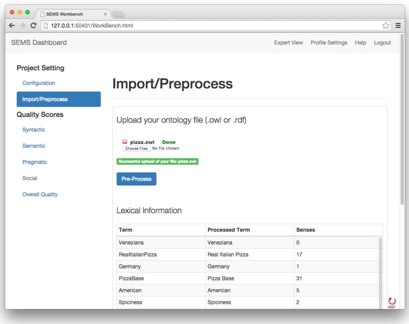

Fig. 2. Screenshot of import and process of ontology file.

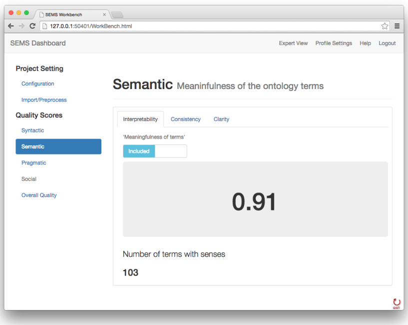

Fig. 3. One of the screens corresponding with an evaluation score.

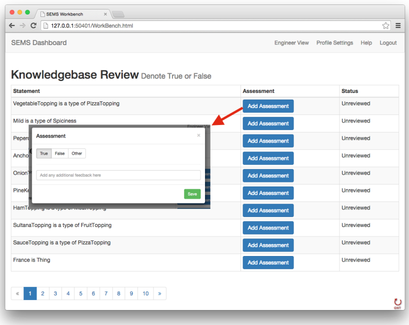

Fig. 4. Knowledgebase review screen for subject matter expert evaluation.

For a brief test demonstration, we utilized the Wine [2] and Pizza [11] ontology. For the configuration, the tool was set to remove camel cases, determiners, underscores, and dashes from the labels from both ontologies. We excluded the *social* criteria, not only because the component was under-development, but due to insufficient data to collect on the number of ontologies extending to it and the number of times it has been downloaded. We excluded *accuracy* aspect of *pragmatic* quality, due to lack of expertise to evaluate the specifics of pizzas and wines and the review interface is still under-development. The following figures (5 and 6) shows the final overall quality score calculated with the available data from the the *pragmatic*, *semantic*, and *syntactic* scores. While both of these ontologies are "toy" ontologies used by knowledge engineers, it would be trivial to examine specific scores, but potentially, the knowledge engineer can view the scores and determine areas of improvement. For example, while both ontologies exhibited nearly equal quality in both *syntactic* and *semantic* aspects, the *comprehensiveness* (under *pragmatic* quality) for the Pizza ontology was relatively lower than the Wine ontology - 0.23 and 0.50 respectively. Here, this would reveal that the Pizza ontology lacks enough classes based on a proportional ratio and may need to define additional, since *comprehensiveness* measures the size of the ontology, as indication of whether it covers the domain completely. Another example would be *richness*, which is the proportion of the number of ontology features utilized, where both ontologies account for using nearly half of the available ontology features - 0.44 and 0.56. This would reveal that the ontologist may need to consider incorporating more ontological features. Understandably, knowing the specifics of each measurement may be cumbersome for knowledge engineers and perhaps defeat the purpose of generating on-demand evaluation scores. One of the future possibilities is to include automated suggestions and description of the quality scores for the knowledge engineer to improve their ontology and to learn more about the specifics of the metric suite.

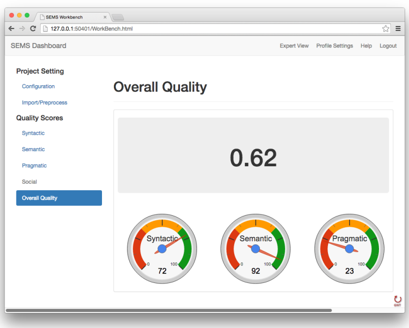

Fig. 5. Sample screen of overall quality score for the Pizza ontology.

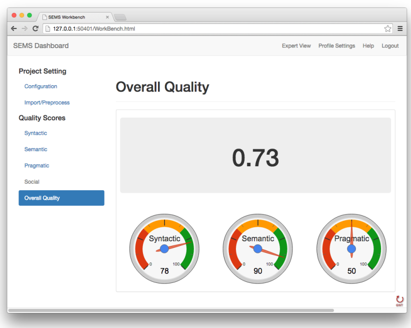

Fig. 6. Sample screen of overall quality score for the Wine ontology.

3 Upcoming Development and Conclusion

SEMS was developed to address the needs of ontology evaluation for authors and experts to encourage ontology usability, and help systematize and streamline the evaluation process. Currently, SEMS is in developmental status and further testing and enhancements are underway, which includes testing the tool with biomedical ontologies and employing the help of biomedical SMEs to determine *accuracy*. Some of the immediate areas under development were alluded to in this paper - user account management, expert interface for SMEs' evaluation of statements, flexibility to handle diverse labeling, suggestions for users to improve their ontology, etc. Also, possible future development may include integrating with NCBO BioPortal's REST service to directly access ontologies and community metrics, and we intend to investigate conformity with The Open Biological and Biomedical Ontologies Foundry's standardization requirements and addressing alignment with upper ontologies. While SEMS is a prototype experimenting with a particular theoretical ontology evaluation framework, there is much room to explore.

Since the semiotic evaluation suite is adaptable, certain aspects for some of the criteria were excluded (like the *relevance* aspect from the *pragmatic* criteria [2]), or they need to be overhauled for the present ontology community. The authors have also considered the possibility of further extending the metric suite to evaluate other features of the ontology that have been supported in ontology evaluation literature - structural assessment [14], ontology question and answering, and perhaps differential semantics to evaluate entities' parent and sibling relationships [6]. As ontologies play an important role in biomedical research for Big Data and analytics, there exist a need to validate and evaluate ontologies, a role that the SEMS application could accommodate.

Acknowledgments. Research was partially supported by the National Library Of Medicine of the National Institutes of Health under Award Number R01LM011829 and the Cancer Prevention Research Institute of Texas (CPRIT) Training Grant #RP140103

References

1. protégé, http://protege.stanford.edu/
2. Wine ontology, http://www.w3.org/TR/owl-guide/wine.rdf
3. Alani, H., Brewster, C.: Metrics for ranking ontologies (2006), http://eprints.ecs.soton.ac.uk/12603
4. Almeida, M.B.: A proposal to evaluate ontology content. Applied Ontology 4(3), 245–265 (2009), http://iospress.metapress.com/index/B2L2XT606156H141.pdf
5. Amith, M., Gong, Y., Cunningham, R., Boom, J., Tao, C.: Developing VISO: Vaccine information statement ontology for patient education. Journal of Biomedical Semantics 6(1), 23 (May 2015)
6. Bachimont, B., Isaac, A., Troncy, R.: Semantic commitment for designing ontologies: a proposal. In: Knowledge Engineering and Knowledge Management: Ontologies and the Semantic Web, pp. 114–121. Springer (2002), http://link.springer.com/chapter/10.1007/3-540-45810-7_14
7. Berners-Lee, T., Hendler, J., Lassila, O., others: The semantic web. Scientific american 284(5), 28–37 (2001)
8. Brank, J., Grobelnik, M., Mladenić, D.: A survey of ontology evaluation techniques (2005), http://eprints.pascal-network.org/archive/00001198/
9. Burton-Jones, A., Storey, V.C., Sugumaran, V., Ahluwalia, P.: A semiotic metrics suite for assessing the quality of ontologies. Data & Knowledge Engineering 55(1), 84–102 (Oct 2005)
10. Chen, M., Mao, S., Liu, Y.: Big data: A survey. Mobile Networks and Applications 19(2), 171–209 (Apr 2014)
11. Drummond, N., Horridge, M., Stevens, R., Wroe, C., Sampaio, S.: Pizza ontology. The University of Manchester (2007)

[2] See [9] for details.

12. Gantz, J., Reinsel, D.: Extracting value from chaos. IDC iview (1142), 9–10 (2011), https://www.emcgrandprix.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf
13. Hansen, M.M., Miron-Shatz, T., Lau, A.Y.S., Paton, C.: Big data in science and healthcare: A review of recent literature and perspectives: Contribution of the IMIA social media working group. IMIA Yearbook 9(1), 21–26 (2014)
14. Ning, H., Shihan, D.: Structure-based ontology evaluation. In: e-Business Engineering, 2006. ICEBE'06. IEEE International Conference on. pp. 132–137. IEEE (2006), http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4031643
15. Obrst, L., Ceusters, W., Mani, I., Ray, S., Smith, B.: The evaluation of ontologies. In: Semantic Web, pp. 139–158. Springer (2007), http://link.springer.com/chapter/10.1007/978-0-387-48438-9_8