

Editorial

Mining Social Semantics on the Social Web

Andreas Hotho^a, Robert Jäschke^b, Kristina Lerman^c

^a *University of Würzburg, Germany*

^b *University of Sheffield, United Kingdom*

^c *University of Southern California, United States*

In recent years the amount of data available on the social web has grown massively. Consequently, researchers have developed approaches that leverage this social web data to tackle interesting challenges of the semantic web. Among these are methods for learning ontologies from social media or crowdsourcing, extracting semantics from data collected by citizen science and participatory sensing initiatives, or for better understanding and describing user activities. The rich data provided by the social web can be used to build the semantic web. This task includes learning basic semantic relationships, e.g., between entities, or by employing more sophisticated methods to construct a complete knowledge graph or ontology. There are additional synergies between the social web and the semantic web. For example, content from the social web could be enriched and linked to the semantic web using named entity recognition and linking, as well as sentiment analysis. These topics were covered previously in the Special Issue on The Personal and Social Semantic Web.¹

This special issue attracted six submissions. Each submission was peer-reviewed by three reviewers. Three submissions were judged to be appropriate and, after revisions with subsequent reviews, were accepted for publication in this special issue. These three submissions share a common theme: the extraction of meaning from user-generated texts and the challenges associated therein. The submissions also show the importance of Twitter, the most popular microblogging service to date, since the methods described in the sub-

missions are particularly suitable for the short texts of tweets. Within that area, the works are diverse, covering sentiment detection, named entity recognition for Turkish, and an overview of named entity recognition and linking approaches.

- *Sentiment Lexicon Adaptation with Context and Semantics for the Social Web* by Hassan Saif, Miriam Fernandez, Leon Kastler, and Harith Alani: Sentiment analysis tools analyze the moods and feelings expressed in social media text. The tools rely on lexicons that encode the emotional meanings of popular words. While sentiment analysis has grown in popularity, the static nature of lexicons cannot account for the context-dependent variations in the sentiment of words. To address this challenge, authors propose a general, unsupervised method to adapt sentiment lexicons to any domain or context of social media posts. The method automatically enriches the lexicon with semantic concepts of words and relations between them. Authors show that this improves performance of sentiment analysis of social media text.
- *Extending a CRF-based Named Entity Recognition Model for Turkish Well Formed Text and User Generated Content* by Gökhan Şeker and Gülşen Eryiğit: The detection of named entities is still a major challenge and a highly active research area, as it is a basic method for many social web applications. While most of the research in this area is focused on less complex languages, the authors of this work present an approach concentrating on a set of well-defined features and feature templates used with a standard conditional

¹<http://www.semantic-web-journal.net/blog/special-issue-personal-and-social-semantic-web>

random field (CRF) for the morphologically rich Turkish language. To demonstrate the strength of the approach, existing well-formed datasets are re-annotated with additional entity types and new social web dataset are introduced and used in the evaluation. Both the data and the code of the system are made available to the research community. The results are very promising and can be the starting point for further research with comparably complex languages.

- *Lessons Learnt from the Named Entity rEcognition and Linking (NEEL) Challenge Series* by Giuseppe Rizzo, Bianca Pereira, Andrea Varga, Marieke van Erp, and Amparo Elizabeth Cano Basave: The identification, classification and linking of named entities in tweets is particularly challenging due to the brevity of tweets. This paper presents a comprehensive overview on the four editions of the NEEL challenge series from

2013 to 2015. The authors first explain the creation process of the annotated tweet corpora and then highlight their strengths and weaknesses in a detailed analysis. The paper also provides a detailed overview on the submitted approaches, including the systems, features, strategies and knowledge bases that were used by the different approaches. Finally, the evaluation setups and measures are presented and the results compared.

We would like to thank all authors for their submissions and the reviewers for their great work which contributed to the high quality of this special issue. With the selected three contributions the special issue addresses only a small fraction of the overall topic of Mining Social Semantics on the Social Web, but we hope that you will enjoy reading it.

Andreas, Kristina, and Robert