

ACL 2012

**50th Annual Meeting of the
Association for Computational Linguistics**

**Proceedings of the
3rd Workshop on the People's Web Meets NLP:
Collaboratively Constructed Semantic Resources
and their Applications to NLP**

July 8 - 14, 2012
Jeju, Republic of Korea

©2012 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 978-1-937284-36-7

Introduction

Recent emergence of Collaboratively Constructed Semantic Resources has led to a set of new research topics that became visible at major computational linguistics conferences. Researchers have studied methods to utilize collaboratively constructed resources to substitute or supplement conventional lexical semantic and linguistically annotated resources. At the same time, Natural Language Processing techniques have been applied to enhance the collaboration process and the overall quality of the collaboratively constructed resources.

This volume contains papers accepted for presentation at the 3rd Workshop on Collaboratively Constructed Semantic Resources and their Applications to NLP, which took place on July 13, 2012, as part of the ACL 2012 conference in Jeju, Republic of Korea. We received submissions from a broad spectrum of research topics, going beyond the coverage of the previous workshops. After careful review by our program committee, three long papers and three short papers were finally accepted for presentation.

The preceding "People's Web meets NLP" workshops at ACL-IJCNLP 2009 and COLING 2010 have successfully gathered researchers from different areas and allowed an interdisciplinary exchange of research outcomes and ideas. The workshop has contributed to the creation of valuable semantic resources and tools based on Collaboratively Constructed Semantic Resources and the consolidation of this young research field.

As the third one in the series, this workshop aims to intensify research that demonstrates the effectiveness of the resources mined from Collaboratively Constructed Semantic Resources in diverse Natural Language Processing tasks. We issued a call for both long and short papers. We especially encouraged submissions that show the benefit of Collaboratively Constructed Semantic Resources in diverse Natural Language Processing tasks.

The aim of the People's Web Meets NLP workshop series is to bring together researchers with different backgrounds: from computational linguistics to various Natural Language Processing areas that benefit from collaboratively constructed semantic resources. We hope this is well reflected in the proceedings.

Jeju, July 2012

Iryna Gurevych, Nicoletta Calzolari Zamorani, and Jungi Kim

Organizers:

Iryna Gurevych, Ubiquitous Knowledge Processing Lab, Technische Universität Darmstadt
Nicoletta Calzolari Zamorani, Istituto di Linguistica Computazionale, CNR
Jungi Kim, Ubiquitous Knowledge Processing Lab, Technische Universität Darmstadt

Program Committee:

Andras Csomai, Google Inc.
Andreas Hotho, Julius-Maximilians-Universität Würzburg
Anette Frank, Heidelberg University
Benno Stein, Bauhaus University Weimar
Christian M. Meyer, Technische Universität Darmstadt
David Milne, University of Waikato
Delphine Bernhard, University of Strasbourg
Diana McCarthy, Lexical Computing Ltd, UK
Donald Metzler, Information Sciences Institute, University of Southern California
Emily Pitler, University of Pennsylvania
Ernesto William De Luca, Technische Universität Berlin
Florian Laws, University of Stuttgart
Gerard de Melo, UC Berkeley
German Rigau, University of the Basque Country
Graeme Hirst, University of Toronto
Günter Neumann, DFKI Saarbrücken
Ido Dagan, Bar Ilan University
John McCrae, University of Bielefeld
Jong-Hyeok Lee, Pohang University of Science and Technology
Judith Eckle-Kohler, Technische Universität Darmstadt
Magnus Sahlgren, Swedish Institute of Computer Science
Manfred Stede, Universität Potsdam
Massimo Poesio, University of Essex
Omar Alonso, Microsoft Bing
Paul Buitelaar, DERI, National University of Ireland, Galway
Rene Witte, Concordia University Montréal
Roxana Girju, University of Illinois at Urbana-Champaign
Saif Mohammad, National Research Council Canada
Shuming Shi, Microsoft Research
Sören Auer, Leipzig University
Tat-Seng Chua, National University of Singapore
Tonio Wandmacher, SYSTRAN, Paris, France
Zornitsa Kozareva, Information Sciences Institute, University of Southern California

Invited Speaker:

Massimo Poesio, University of Essex

Title: Phrase Detectives: the First Three Years

Abstract: Phrase Detectives, one of the first games-with-a-purpose for corpus annotation (www.phrasedetectives.org) went officially online on December 1st 2008, and one of its very first presentations in front of an NLP audience took place at the first edition of the “People’s Web Meets NLP” workshop in Singapore in 2009. The option of annotating Italian documents was added in 2010, and a Facebook version went live in January 2012. Although the project that funded its creation ended in September 2009, the game has stayed very much alive, in fact it is getting more active all the time - we recently passed the 11,000 players mark and are about to reach 200,000 words of fully annotated documents, with a goal of annotating at least 1 million. In the talk I will discuss recent developments and analyze the results so far in terms of quality and quantity of annotated data and annotation costs.

References: Massimo Poesio, Jon Chamberlain, Udo Kruschwitz, Livio Robaldo, and Luca Ducceschi, In Press - Phrase Detectives: Utilizing Collective Intelligence for Internet-Scale Language Resource Creation, ACM Transactions on Interactive Intelligent Systems.

Short bio: Massimo Poesio is a Professor of Computer Science at the University of Essex and the Director of the Language Interaction and Computation Lab at the Centre for Mind/Brain Sciences, University of Trento. He is best known for his work in anaphora resolution, corpus annotation, and the acquisition of common sense knowledge.

Table of Contents

<i>Sentiment Analysis Using a Novel Human Computation Game</i> Claudiu Cristian Musat, Alireza Ghasemi and Boi Faltings	1
<i>A Serious Game for Building a Portuguese Lexical-Semantic Network</i> Mathieu Mangeot and Carlos Ramisch	10
<i>Collaboratively Building Language Resources while Localising the Web</i> Asanka Wasala, Reinhard Schäler, Ruvan Weerasinghe and Chris Exton	15
<i>Resolving Task Specification and Path Inconsistency in Taxonomy Construction</i> Hui Yang	20
<i>EAGER: Extending Automatically Gazetteers for Entity Recognition</i> Omer Farukhan Gunes, Tim Furche, Christian Schallhart, Jens Lehmann and Axel-Cyrille Ngonga Ngomo	29
<i>Extracting Context-Rich Entailment Rules from Wikipedia Revision History</i> Elena Cabrio, Bernardo Magnini and Angelina Ivanova	34

Workshop Program

Friday July 13, 2012

- 09:15–09:30 Opening remarks
- 09:35–10:05 *Sentiment Analysis Using a Novel Human Computation Game*
Claudiu Cristian Musat, Alireza Ghasemi and Boi Faltings
- 10:10–10:30 *A Serious Game for Building a Portuguese Lexical-Semantic Network*
Mathieu Mangeot and Carlos Ramisch
- 10:30–11:00 Coffee break
- 11:00–11:20 *Collaboratively Building Language Resources while Localising the Web*
Asanka Wasala, Reinhard Schäler, Ruvan Weerasinghe and Chris Exton
- 11:25–12:30 Invited talk: Phrase Detectives: The First Three Years by Massimo Poesio
- 12:30–14:00 Lunch break
- 14:00–14:30 *Resolving Task Specification and Path Inconsistency in Taxonomy Construction*
Hui Yang
- 14:35–15:55 *EAGER: Extending Automatically Gazetteers for Entity Recognition*
Omer Farukhan Gunes, Tim Furche, Christian Schallhart, Jens Lehmann and Axel-Cyrille Ngonga Ngomo
- 15:00–15:30 *Extracting Context-Rich Entailment Rules from Wikipedia Revision History*
Elena Cabrio, Bernardo Magnini and Angelina Ivanova
- 15:30–16:00 Coffee break
- 16:00–17:00 Panel discussion: Collaboratively Looking Ahead: How to Make Sustainable Goods out of Collaboratively Constructed Semantic Resources?

