

# **25th International Conference on Computational Linguistics**

## **(COLING 2014)**

**System Demonstrations**

**Dublin, Ireland**  
**23 - 29 August 2014**

**Editors:**

**Lamia Tounsi**

**Rafal Rak**

**ISBN: 978-1-63439-196-2**

**Printed from e-media with permission by:**

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571



**Some format issues inherent in the e-media version may also appear in this print version.**

Copyright© (2014) by the Association for Computational Linguistics  
All rights reserved.

Printed by Curran Associates, Inc. (2014)

For permission requests, please contact the Association for Computational Linguistics  
at the address below.

Association for Computational Linguistics  
209 N. Eighth Street  
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006  
Fax: 1-570-476-0860

[acl@aclweb.org](mailto:acl@aclweb.org)

**Additional copies of this publication are available from:**

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571 USA  
Phone: 845-758-0400  
Fax: 845-758-2634  
Email: [curran@proceedings.com](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

## Table of Contents

<i>An Error Analysis Tool for Natural Language Processing and Applied Machine Learning</i> Apoorv Agarwal, Ankit Agarwal and Deepak Mittal . . . . .	1
<i>Claims on demand – an initial demonstration of a system for automatic detection and polarity identification of context dependent claims in massive corpora</i> Noam Slonim, Ehud Aharoni, Carlos Alzate, Roy Bar-Haim, Yonatan Bilu, Lena Dankin, Iris Eiron, Daniel Hershcovich, Shay Hummel, Mitesh Khapra, Tamar Lavee, Ran Levy, Paul Matchen, Anatoly Polnarov, Vikas Raykar, Ruty Rinott, Amrita Saha, Naama Zwerdling, David Konopnicki and Dan Gutfreund . . . . .	6
<i>Copa 2014 FrameNet Brasil: a frame-based trilingual electronic dictionary for the Football World Cup</i> Tiago Torrent, Maria Margarida Salomão, Fernanda Campos, Regina Braga, Ely Matos, Maucha Gamonal, Julia Gonçalves, Bruno Souza, Daniela Gomes and Simone Peron . . . . .	10
<i>Creating Custom Taggers by Integrating Web Page Annotation and Machine Learning</i> Srikrishna Raamadhurai, Oskar Kohonen and Teemu Ruokolainen . . . . .	15
<i>How to deal with students’ writing problems? Process-oriented writing support with the digital Writing Aid Dutch</i> Lieve De Wachter, Serge Verlinde, Margot D’Hertefelt and Geert Peeters . . . . .	20
<i>Processing Discourse in Dislog on the TextCoop Platform</i> patrick saint-dizier . . . . .	25
<i>UIMA Ruta Workbench: Rule-based Text Annotation</i> Peter Kluegl, Martin Toepfer, Philip-Daniel Beck, Georg Fette and Frank Puppe . . . . .	29
<i>Discourse Relations in the Prague Dependency Treebank 3.0</i> Jiří Mírovský, Pavlína Jínová and Lucie Poláková . . . . .	34
<i>Lightweight Client-Side Chinese/Japanese Morphological Analyzer Based on Online Learning</i> Masato Hagiwara and Satoshi Sekine . . . . .	39
<i>MultiDPS – A multilingual Discourse Processing System</i> Daniel Anechitei . . . . .	44
<i>Sanskrit Linguistics Web Services</i> Gérard Huet and Amba Kulkarni . . . . .	48
<i>TICCLops: Text-Induced Corpus Clean-up as online processing system</i> Martin Reynaert . . . . .	52
<i>Trameur: A Framework for Annotated Text Corpora Exploration</i> Serge Fleury and Maria Zimina . . . . .	57
<i>TweetGenie: Development, Evaluation, and Lessons Learned</i> Dong Nguyen, Dolf Trieschnigg and Theo Meder . . . . .	62
<i>A Sentence Judgment System for Grammatical Error Detection</i> Lung-Hao Lee, Liang-Chih Yu, Kuei-Ching Lee, Yuen-Hsien Tseng, Li-Ping Chang and Hsin-Hsi Chen . . . . .	67

<i>CLAM: Quickly deploy NLP command-line tools on the web</i>	
Maarten van Gompel and Martin Reynaert .....	71
<i>CRAB 2.0: A text mining tool for supporting literature review in chemical cancer risk assessment</i>	
Yufan Guo, Diarmuid Ó Séaghdha, Iona Silins, Lin Sun, Johan Högberg, Ulla Stenius and Anna Korhonen .....	76
<i>Nerdle: Topic-Specific Question Answering Using Wikia Seeds</i>	
Umar Maqsd, Sebastian Arnold, Michael Hülfenhaus and Alan Akbik .....	81
<i>NTU-MC Toolkit: Annotating a Linguistically Diverse Corpus</i>	
Liling Tan and Francis Bond .....	86
<i>RDF Triple Stores and a Custom SPARQL Front-End for Indexing and Searching (Very) Large Semantic Networks</i>	
Milen Kouylekov and Stephan Oepen .....	90
<i>What or Who is Multilingual Watson?</i>	
Keith Cortis, Urvesh Bhowan, Ronan Mac an tSaoir, D.J. McCloskey, Mikhail Sogrin and Ross Cadogan .....	95
<i>A Marketplace for Web Scale Analytics and Text Annotation Services</i>	
Johannes Kirschnick, Torsten Kilius, Holmer Hemsén, Alexander Löser, Peter Adolphs, Heiko Ehrig and Holger Düwiger .....	100
<i>DKPro Agreement: An Open-Source Java Library for Measuring Inter-Rater Agreement</i>	
Christian M. Meyer, Margot Mieskes, Christian Stab and Iryna Gurevych .....	105
<i>Distributional Semantics in R with the wordspace Package</i>	
Stefan Evert .....	110
<i>Method51 for Mining Insight from Social Media Datasets</i>	
Simon Wibberley, David Weir and Jeremy Reffin .....	115
<i>MT-EQuAl: a Toolkit for Human Assessment of Machine Translation Output</i>	
Christian Girardi, Luisa Bentivogli, Mohammad Amin Farajian and Marcello Federico .....	120
<i>OpenSoNaR: user-driven development of the SoNaR corpus interfaces</i>	
Martin Reynaert, Matje van de Camp and Menno van Zaanen .....	124
<b>THE MATECAT TOOL</b>	
Marcello Federico, Nicola Bertoldi, Mauro Cettolo, Matteo Negri, Marco Turchi, Marco Trombetti, Alessandro Cattelan, Antonio Farina, Domenico Lupinetti, Andrea Martines, Alberto Massida, Holger Schwenk, Loïc Barrault, Frederic Blain, Philipp Koehn, Christian Buck and Ulrich Germann .....	129