ACL 2014

52nd Annual Meeting of the Association for Computational Linguistics

Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations

June 23-24, 2014
Baltimore, Maryland, USA

©2014 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 978-1-941643-00-6

Preface

Welcome to the proceedings of the system demonstration session. This volume contains the papers of the system demonstrations presented at the 52nd Annual Meeting of the Association for Computational Linguistics, on June 23-24, 2014 in Baltimore, Maryland, USA.

The system demonstrations program offers the presentation of early research prototypes as well as interesting mature systems. The system demonstration chairs and the members of the program committee received 39 submissions, of which 21 were selected for inclusion in the program (acceptance rate of 53.8%) after review by three members of the program committee.

We would like to thank the members of the program committee for their timely help in reviewing the submissions.

Co-Chairs:

Kalina Bontcheva, University of Sheffield
Zhu Jingbo, Northeastern University

Program Committee:

Beatrice Alex, University of Edinburgh
Omar Alonso, Microsoft
Thierry Declerck, DFKI
Leon Derczynski, University of Sheffield
Mark Dras, Macquarie University
Josef van Genabith, Dublin City University
Dirk Hovy, University of Copenhagen
Degen Huang, Dalian Univeristy of Technology
Xuanjing Huang, Institute for Media Computing
Brigitte Kren, OFAI
Oi Yee Kwong, City University of Hong Kong
Maggie Li, The Hong Kong Polytechnic University
Mu Li, Microsoft Research Asia
Wenjie Li, The Hong Kong Polytechnic University
Maria Liakata, University of Warwick
Yuji Matsumoto, Nara Institute of Science and Technology
Nicolas Nicolov, Amazon
Patrick Paroubek, LIMSI
Stelios Piperidis, Athena RC/ILSP
Barbara Plank, University of Copenhagen
Kiril Simov, Bulgarian Academy of Science
Koenraad De Smedt, University of Bergen
Lucia Specia, University of Sheffield
John Tait, johntait.net Ltd.
Keh-Yih Su, BDC
Le Sun, ISCAS
Marc Vilain, MITRE
Leo Wanner, ICREA and Pompeu Fabra University
Derek F. Wong, University of Macau
Wei Xu, New York University
Liang-Chih Yu, Yuan Ze University
Jiajun Zhang, Chinese Academy of Sciences
Yue Zhang, Singapore University of Technology and Design
Tiejun Zhao, Harbin Institute of Technology
Guodong Zhou, Soochow University

Table of Contents

Cross-Lingual Information to the Rescue in Keyword Extraction
Chung-Chi Huang, Maxine Eskenazi, Jaime Carbonell, Lun-Wei Ku and Ping-Che Yang 1

Visualization, Search, and Error Analysis for Coreference Annotations
Markus Gärtner, Anders Björkelund, Gregor Thiele, Wolfgang Seeker and Jonas Kuhn 7

Open-Source Tools for Morphology, Lemmatization, POS Tagging and Named Entity Recognition
Jana Straková, Milan Straka and Jan Hajič 13

Community Evaluation and Exchange of Word Vectors at wordvectors.org
Manaal Faruqui and Chris Dyer 19

WINGS:Writing with Intelligent Guidance and Suggestions
Xianjun Dai, Yuanchao Liu, Xiaolong Wang and Bingquan Liu 25

DKPro Keyphrases: Flexible and Reusable Keyphrase Extraction Experiments
Nicolai Erbs, Pedro Bispo Santos, Iryna Gurevych and Torsten Zesch 31

Real-Time Detection, Tracking, and Monitoring of Automatically Discovered Events in Social Media
Miles Osborne, Sean Moran, Richard McCreadie, Alexander Von Lunen, Martin Sykora, Elizabeth Cano, Neil Ireson, Craig Macdonald, Iadh Ounis, Yulan He, Tom Jackson, Fabio Ciravegna and Ann O'Brien 37

The Excitement Open Platform for Textual Inferences
Bernardo Magnini, Roberto Zanoli, Ido Dagan, Kathrin Eichler, Guenter Neumann, Tae-Gil Noh, Sebastian Padó, Asher Stern and Omer Levy 43

WELT: Using Graphics Generation in Linguistic Fieldwork
Morgan Ulinski, Anusha Balakrishnan, Bob Coyne, Julia Hirschberg and Owen Rambow 49

The Stanford CoreNLP Natural Language Processing Toolkit
Christopher Manning, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven Bethard and David McClosky 55

DKPro TC: A Java-based Framework for Supervised Learning Experiments on Textual Data
Johannes Daxenberger, Oliver Ferschke, Iryna Gurevych and Torsten Zesch 61

WoSIT: A Word Sense Induction Toolkit for Search Result Clustering and Diversification
Daniele Vannella, Tiziano Flati and Roberto Navigli 67

A Rule-Augmented Statistical Phrase-based Translation System
Cong Duy Vu Hoang, AiTi Aw and Nhung T. H. Nguyen 73

KyotoEBMT: An Example-Based Dependency-to-Dependency Translation Framework
John Richardson, Fabien Cromières, Toshiaki Nakazawa and Sadao Kurohashi 79

kLogNLP: Graph Kernel–based Relational Learning of Natural Language
Mathias Verbeke, Paolo Frasconi, Kurt De Grave, Fabrizio Costa and Luc De Raedt 85

Automatic Annotation Suggestions and Custom Annotation Layers in WebAnno
Seid Muhie Yimam, Chris Biemann, Richard Eckart de Castilho and Iryna Gurevych 91

Web Information Mining and Decision Support Platform for the Modern Service Industry
Binyang Li, Lanjun Zhou, Zhongyu Wei, Kam-fai Wong, Ruifeng Xu and Yunqing Xia 97

FAdR: A System for Recognizing False Online Advertisements
Yi-jie Tang and Hsin-Hsi Chen 103

lex4all: A language-independent tool for building and evaluating pronunciation lexicons for small-vocabulary speech recognition
Anjana Vakil, Max Paulus, Alexis Palmer and Michaela Regneri 109

Enhanced Search with Wildcards and Morphological Inflections in the Google Books Ngram Viewer
Jason Mann, David Zhang, Lu Yang, Dipanjan Das and Slav Petrov 115

Simplified Dependency Annotations with GFL-Web
Michael T. Mordowanec, Nathan Schneider, Chris Dyer and Noah A. Smith 121

Conference Program

23 June 2014

18:50–21:30 Session A

Cross-Lingual Information to the Rescue in Keyword Extraction
Chung-Chi Huang, Maxine Eskenazi, Jaime Carbonell, Lun-Wei Ku and Ping-Che Yang

Visualization, Search, and Error Analysis for Coreference Annotations
Markus Gärtner, Anders Björkelund, Gregor Thiele, Wolfgang Seeker and Jonas Kuhn

Open-Source Tools for Morphology, Lemmatization, POS Tagging and Named Entity Recognition
Jana Straková, Milan Straka and Jan Hajič

Community Evaluation and Exchange of Word Vectors at wordvectors.org
Manaal Faruqui and Chris Dyer

WINGS:Writing with Intelligent Guidance and Suggestions
Xianjun Dai, Yuanchao Liu, Xiaolong Wang and Bingquan Liu

DKPro Keyphrases: Flexible and Reusable Keyphrase Extraction Experiments
Nicolai Erbs, Pedro Bispo Santos, Iryna Gurevych and Torsten Zesch

Real-Time Detection, Tracking, and Monitoring of Automatically Discovered Events in Social Media
Miles Osborne, Sean Moran, Richard McCreadie, Alexander Von Lunen, Martin Sykora, Elizabeth Cano, Neil Ireson, Craig Macdonald, Iadh Ounis, Yulan He, Tom Jackson, Fabio Ciravegna and Ann O'Brien

The Excitement Open Platform for Textual Inferences
Bernardo Magnini, Roberto Zanoli, Ido Dagan, Kathrin Eichler, Guenter Neumann, Tae-Gil Noh, Sebastian Padó, Asher Stern and Omer Levy

WELT: Using Graphics Generation in Linguistic Fieldwork
Morgan Ulinski, Anusha Balakrishnan, Bob Coyne, Julia Hirschberg and Owen Rambow

The Stanford CoreNLP Natural Language Processing Toolkit
Christopher Manning, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven Bethard and David McClosky

24 June 2014

16:50–19:20 Session B

DKPro TC: A Java-based Framework for Supervised Learning Experiments on Textual Data
Johannes Daxenberger, Oliver Ferschke, Iryna Gurevych and Torsten Zesch

WoSIT: A Word Sense Induction Toolkit for Search Result Clustering and Diversification
Daniele Vannella, Tiziano Flati and Roberto Navigli

A Rule-Augmented Statistical Phrase-based Translation System
Cong Duy Vu Hoang, AiTi Aw and Nhung T. H. Nguyen

KyotoEBMT: An Example-Based Dependency-to-Dependency Translation Framework
John Richardson, Fabien Cromières, Toshiaki Nakazawa and Sadao Kurohashi

kLogNLP: Graph Kernel–based Relational Learning of Natural Language
Mathias Verbeke, Paolo Frasconi, Kurt De Grave, Fabrizio Costa and Luc De Raedt

Automatic Annotation Suggestions and Custom Annotation Layers in WebAnno
Seid Muhie Yimam, Chris Biemann, Richard Eckart de Castilho and Iryna Gurevych

Web Information Mining and Decision Support Platform for the Modern Service Industry
Binyang Li, Lanjun Zhou, Zhongyu Wei, Kam-fai Wong, Ruifeng Xu and Yunqing Xia

FAdR: A System for Recognizing False Online Advertisements
Yi-jie Tang and Hsin-Hsi Chen

lex4all: A language-independent tool for building and evaluating pronunciation lexicons for small-vocabulary speech recognition
Anjana Vakil, Max Paulus, Alexis Palmer and Michaela Regneri

Enhanced Search with Wildcards and Morphological Inflections in the Google Books Ngram Viewer
Jason Mann, David Zhang, Lu Yang, Dipanjan Das and Slav Petrov

Simplified Dependency Annotations with GFL-Web
Michael T. Mordowanec, Nathan Schneider, Chris Dyer and Noah A. Smith