

10th SIGHUM Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities (LaTeCH 2016)

Held at ACL 2016

Berlin, Germany
11 August 2016

ISBN: 978-1-5108-2773-8

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© (2016) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2016)

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

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2633
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Contents

<i>Brave New World: Uncovering Topical Dynamics in the ACL Anthology Reference Corpus Using Term Life Cycle Information</i> Anne-Kathrin Schumann	1
<i>Analysis of Policy Agendas: Lessons Learned from Automatic Topic Classification of Croatian Political Texts</i> Mladen Karan, Jan Šnajder, Daniela Sirinic and Goran Glavaš	12
<i>Searching Four-Millenia-Old Digitized Documents: A Text Retrieval System for Egyptologists</i> Estíbaliz Iglesias-Franjo and Jesús Vilares	22
<i>Old Swedish Part-of-Speech Tagging between Variation and External Knowledge</i> Yvonne Adesam and Gerlof Bouma	32
<i>Code-Switching Ubique Est - Language Identification and Part-of-Speech Tagging for Historical Mixed Text</i> Sarah Schulz and Mareike Keller	43
<i>Dealing with word-internal modification and spelling variation in data-driven lemmatization</i> Fabian Barteld, Ingrid Schröder and Heike Zinsmeister	52
<i>You Shall Know People by the Company They Keep: Person Name Disambiguation for Social Network Construction</i> Mariona Coll Ardanuy, Maarten van den Bos and Caroline Sporleder	63
<i>Deriving Players & Themes in the Regesta Imperii using SVMs and Neural Networks</i> Juri Opitz and Anette Frank	74
<i>Semi-automated annotation of page-based documents within the Genre and Multimodality framework</i> Tuomo Hiippala	84
<i>Nomen Omen. Enhancing the Latin Morphological Analyser Lemlat with an Onomasticon</i> Marco Budassi and Marco Passarotti	90
<i>How Do Cultural Differences Impact the Quality of Sarcasm Annotation?: A Case Study of Indian Annotators and American Text</i> Aditya Joshi, Pushpak Bhattacharyya, Mark Carman, Jaya Saraswati and Rajita Shukla	95
<i>Combining Phonology and Morphology for the Normalization of Historical Texts</i> Izaskun Etxeberria, Iñaki Alegria, Larraitz Uria and Mans Hulden	100
<i>Towards Building a Political Protest Database to Explain Changes in the Welfare State</i> Çağıl Sönmez, Arzucan Özgür and Erdem Yörük	106
<i>An Assessment of Experimental Protocols for Tracing Changes in Word Semantics Relative to Accuracy and Reliability</i> Johannes Hellrich and Udo Hahn	111
<i>Universal Morphology for Old Hungarian</i> Eszter Simon and Veronika Vincze	118

<i>Automatic Identification of Suicide Notes from Linguistic and Sentiment Features</i>	
Annika Marie Schoene and Nina Dethlefs	128
<i>Towards a text analysis system for political debates</i>	
Dieu-Thu Le, Ngoc Thang Vu and Andre Blessing	134
<i>Whodunit... and to Whom? Subjects, Objects, and Actions in Research Articles on American Labor Unions</i>	
Vilja Hulden	140
<i>An NLP Pipeline for Coptic</i>	
Amir Zeldes and Caroline T. Schroeder	146
<i>Automatic discovery of Latin syntactic changes</i>	
Micha Elsner and Emily Lane	156
<i>Information-based Modeling of Diachronic Linguistic Change: from Typicality to Productivity</i>	
Stefania Degaetano-Ortlieb and Elke Teich	165