

**7th Workshop on Language
Technology for Cultural Heritage,
Social Sciences, and Humanities**

(LaTeCH 2013)

**Sofia, Bulgaria
8 August 2013**

ISBN: 978-1-62748-998-0

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© (2013) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2013)

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: 845-758-2634
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Contents

<i>Generating Paths through Cultural Heritage Collections</i>	
Samuel Fernando, Paula Goodale, Paul Clough, Mark Stevenson, Mark Hall and Eneko Agirre ..	1
<i>Using character overlap to improve language transformation</i>	
Sander Wubben, Emiel Kraemer and Antal van den Bosch	11
<i>Comparison between historical population archives and decentralized databases</i>	
Marijn Schraagen and Dionysius Huijsmans	20
<i>Semi-automatic Construction of Cross-period Thesaurus</i>	
Chaya Liebeskind, Ido Dagan and Jonathan Schler	29
<i>Language Technology for Agile Social Media Science</i>	
Simon Wibberley, David Weir and Jeremy Reffin	36
<i>Morphological annotation of Old and Middle Hungarian corpora</i>	
Attila Novak, György Orosz and Nóra Wenszky	43
<i>Argument extraction for supporting public policy formulation</i>	
Eirini Florou, Stasinou Konstantopoulos, Antonis Koukourikos and Pythagoras Karampiperis ..	49
<i>Towards a Tool for Interactive Concept Building for Large Scale Analysis in the Humanities</i>	
Andre Blessing, Jonathan Sonntag, Fritz Kliche, Ulrich Heid, Jonas Kuhn and Manfred Stede ..	55
<i>Learning to Extract Folktale Keywords</i>	
Dolf Trieschnigg, Dong Nguyen and Mariët Theune	65
<i>Towards Creating Precision Grammars from Interlinear Glossed Text: Inferring Large-Scale Typological Properties</i>	
Emily M. Bender, Michael Wayne Goodman, Joshua Crowgey and Fei Xia	74
<i>Using Comparable Collections of Historical Texts for Building a Diachronic Dictionary for Spelling Normalization</i>	
Marilisa Amoia and José Manuel Martínez	84
<i>Integration of the Thesaurus for the Social Sciences (TheSoz) in an Information Extraction System</i>	
Thierry Declerck	90
<i>The (Un)faithful Machine Translator</i>	
Ruth Jones and Ann Irvine	96
<i>Temporal classification for historical Romanian texts</i>	
Alina Maria Ciobanu, Anca Dinu, Liviu Dinu, Vlad Niculae and Octavia-Maria Şulea	102
<i>Multilingual access to cultural heritage content on the Semantic Web</i>	
Dana Dannells, Aarne Ranta, Ramona Enache, Mariana Damova and Maria Mateva	107