Named Entities Workshop 2010

(NEWS 2010)

Held at ACL 2010

Uppsala, Sweden
16 July 2010

ISBN: 978-1-61738-818-7

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© (2010) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2010)

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

Report of NEWS 2010 Transliteration Generation Shared Task
Haizhou Li, A Kumaran, Min Zhang and Vladimir Pervouchine 1

Whitepaper of NEWS 2010 Shared Task on Transliteration Generation
Haizhou Li, A Kumaran, Min Zhang and Vladimir Pervouchine 12

Report of NEWS 2010 Transliteration Mining Shared Task
A Kumaran, Mitesh M. Khapra and Haizhou Li 21

Whitepaper of NEWS 2010 Shared Task on Transliteration Mining
A Kumaran, Mitesh M. Khapra and Haizhou Li 29

Transliteration Generation and Mining with Limited Training Resources
Sittichai Jiampojamarn, Kenneth Dwyer, Shane Bergsma, Aditya Bhargava, Qing Dou, Mi-Young Kim and Grzegorz Kondrak 39

Transliteration Using a Phrase-Based Statistical Machine Translation System to Re-Score the Output of a Joint Multigram Model
Andrew Finch and Eiichiro Sumita 48

Transliteration Mining with Phonetic Conflation and Iterative Training
Kareem Darwish 53

Language Independent Transliteration Mining System Using Finite State Automata Framework
Sara Noeman and Amgad Madkour 57

Reranking with Multiple Features for Better Transliteration
Yan Song, Chunyu Kit and Hai Zhao 62

Syllable-Based Thai-English Machine Transliteration
Chai Wutiwiwatchai and Ausdang Thangthai 66

English to Indian Languages Machine Transliteration System at NEWS 2010
Amitava Das, Tanik Saikh, Tapabrata Mondal, Asif Ekbal and Sivaji Bandyopadhyay 71

Mining Transliterations from Wikipedia Using Pair HMMs
Peter Nabende 76

Phrase-Based Transliteration with Simple Heuristics
Avinesh PVS and Ankur Parikh 81

Classifying Wikipedia Articles into NE's Using SVM's with Threshold Adjustment
Iman Saleh, Kareem Darwish and Aly Fahmy 85

Assessing the Challenge of Fine-Grained Named Entity Recognition and Classification
Asif Ekbal, Eva Sourjikova, Anette Frank and Simone Paolo Ponzetto 93

Using Deep Belief Nets for Chinese Named Entity Categorization
Yu Chen, You Ouyang, Wenjie Li, Dequan Zheng and Tiejun Zhao 102

Simplified Feature Set for Arabic Named Entity Recognition
Ahmed Abdul Hamid and Kareem Darwish 110

Think Globally, Apply Locally: Using Distributional Characteristics for Hindi Named Entity Identification
Shalini Gupta and Pushpak Bhattacharyya 116

Rule-Based Named Entity Recognition in Urdu
Kashif Riaz 126

CONE: Metrics for Automatic Evaluation of Named Entity Co-Reference Resolution
Bo Lin, Rushin Shah, Robert Frederking and Anatole Gershman 136