Named Entities Workshop 2011

(NEWS 2011)

Held at IJCNLP 2011

Chiang Mai, Thailand
12 November 2011

ISBN: 978-1-62276-453-2

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© (2011) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2012)

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 2011 Machine Transliteration Shared Task
Min Zhang, Haizhou Li, A Kumaran and Ming Liu 1

Whitepaper of NEWS 2011 Shared Task on Machine Transliteration
Min Zhang, A Kumaran and Haizhou Li 14

Integrating Models Derived from non-Parametric Bayesian Co-segmentation into a Statistical Machine Transliteration System
Andrew Finch, Paul Dixon and Eiichiro Sumita 23

Simple Discriminative Training for Machine Transliteration
Canasai Kruengkrai, Thatsanee Charoenporn and Virach Sornlertlamvanich 28

English-Korean Named Entity Transliteration Using Statistical Substring-based and Rule-based Approaches
Yu-Chun Wang and Richard Tzong-Han Tsai 32

Leveraging Transliterations from Multiple Languages
Aditya Bhargava, Bradley Hauer and Grzegorz Kondrak 36

Comparative Evaluation of Spanish Segmentation Strategies for Spanish-Chinese Transliteration
Rafael E. Banchs 41

Using Features from a Bilingual Alignment Model in Transliteration Mining
Takaaki Fukunishi, Andrew Finch, Seiichi Yamamoto and Eiichiro Sumita 49

Product Name Identification and Classification in Thai Economic News
Nattadaporn Lertcheva and Wirote Aroonmanakun 58

Mining Multi-word Named Entity Equivalents from Comparable Corpora
Abhijit Bhole, Goutham Tholpadi and Raghavendra Udupa 65

An Unsupervised Alignment Model for Sequence Labeling: Application to Name Transliteration
Najmeh Mousavi Nejad and Shahram Khadivi 73

Forward-backward Machine Transliteration between English and Chinese Based on Combined CRFs
Ying Qin and GuoHua Chen 82

English-to-Chinese Machine Transliteration using Accessor Variety Features of Source Graphemes
Mike Tian-Jian Jiang, Chan-Hung Kuo and Wen-Lian Hsu 86

The Amirkabir Machine Transliteration System for NEWS 2011: Farsi-to-English Task
Najmeh Mousavi Nejad, Shahram Khadivi and Kaveh Taghipour 91

English-Chinese Personal Name Transliteration by Syllable-Based Maximum Matching
Oi Yee Kwong 96

Statistical Machine Transliteration with Multi-to-Multi Joint Source Channel Model
Yu Chen, Rui Wang and Yi Zhang 101

Named Entity Transliteration Generation Leveraging Statistical Machine Translation Technology
Pradeep Dasigi and Mona Diab 106