

INTERNATIONAL WORKSHOP

**ADAPTATION OF LANGUAGE RESOURCES
AND TECHNOLOGY TO NEW DOMAINS
(AdaptLRTtoND)**

*held in conjunction with the International Conference
RANLP - 2009, 14-16 September 2009, Borovets, Bulgaria*

PROCEEDINGS

Edited by
Núria Bel, Erhard Hinrichs, Kiril Simov and Petya Osenova

The Workshop is endorsed by FLaReNet Project



<http://www.flarenet.eu/>

The Workshop is related to CLARIN Project

<http://www.clarin.eu/>

Borovets, Bulgaria
17 September 2009

International Workshop
ADAPTATION OF LANGUAGE RESOURCES
AND TECHNOLOGY TO NEW DOMAINS

PROCEEDINGS

Borovets, Bulgaria
17 September 2009

ISBN 978-954-452-009-0

Designed and Printed by INCOMA Ltd.
Shoumen, Bulgaria

Workshop Organisers

Núria Bel

Pompeu Fabra University, Spain

Erhard Hinrichs

Tuebingen University, Germany (co-chair)

Petya Osenova

Bulgarian Academy of Sciences and Sofia University, Bulgaria

Kiril Simov

Bulgarian Academy of Sciences, Bulgaria (co-chair)

Workshop Programme Committee

Núria Bel, Pompeu Fabra University

Gosse Bouma, Groningen University

António Branco, Lisbon University

Walter Daelemans, Antwerp University

Markus Dickinson, Indiana University

Erhard Hinrichs, Tübingen University

Josef van Genabith, Dublin City University

Iryna Gurevych, Technische Universität Darmstadt - UKP Lab

Atanas Kiryakov, Ontotext AD

Vladislav Kubon, Charles University

Sandra Kübler, Indiana University

Lothar Lemnitzer, DWDS, Berlin-Brandenburgische Akademie der Wissenschaften

Bernardo Magnini, FBK

Detmar Meurers, Tübingen University

Paola Monachesi, Utrecht University

Preslav Nakov, National University of Singapore

John Nerbonne, Groningen University

Petya Osenova, Bulgarian Academy of Sciences and Sofia University

Gabor Proszeky, MophoLogic

Adam Przepiorkowski, Polish Academy of Sciences

Marta Sabou, Open University - UK

Kiril Simov, Bulgarian Academy of Sciences

Cristina Vertan, Hamburg University

Table of Contents

<i>Exploiting the Russian National Corpus in the Development of a Russian Resource Grammar</i> Tania Avgustinova and Yi Zhang	1
<i>Maximal Phrases Based Analysis for Prototyping Online Discussion Forums Postings</i> Gaston Burek and Dale Gerdemann	12
<i>LEXIE – an Experiment in Lexical Information Extraction</i> John J. Camilleri and Michael Rosner	19
<i>Adapting NLP and Corpus Analysis Techniques to Structured Imagery Analysis in Classical Chinese Poetry</i> Alex Chengyu Fang, Fengju Lo and Cheuk Kit Chinn	27
<i>Cross-lingual Adaptation as a Baseline: Adapting Maximum Entropy Models to Bulgarian</i> Georgi Georgiev and Preslav Nakov	35
<i>Mix and Match Replacement Rules</i> Dale Gerdemann	39
<i>Enabling Adaptation of Lexicalised Grammars to New Domains</i> Valia Kordoni and Yi Zhang	48
<i>QALL-ME needs AIR: a portability study</i> Constantin Orăsan, Iustin Dornescu and Natalia Ponomareva	50
<i>Personal Health Information Leak Prevention in Heterogeneous Texts</i> Marina Sokolova, Khaled El Emam, Sean Rose, Sadrul Chowdhury, Emilio Neri, Elizabeth Jonker and Liam Peyton	58

Conference Program

Session 1

- 09:20-09:30 Opening
- 09:30-10:30 TBA
Atanas Kiryakov (Invited Speaker), OntoText AD, Sofia, Bulgaria
- 10:30-11:00 *Enabling Adaptation of Lexicalised Grammars to New Domains*
Valia Kordoni and Yi Zhang

Session 2

- 11:30-12:00 *Personal Health Information Leak Prevention in Heterogeneous Texts*
Marina Sokolova, Khaled El Emam, Sean Rose, Sadrul Chowdhury, Emilio Neri, Elizabeth Jonker and Liam Peyton
- 12:00-12:30 *Mix and Match Replacement Rules*
Dale Gerdemann

Session 3

- 14:00-14:30 *Cross-lingual Adaptation as a Baseline: Adapting Maximum Entropy Models to Bulgarian*
Georgi Georgiev and Preslav Nakov
- 14:30-15:00 *Exploiting the Russian National Corpus in the Development of a Russian Resource Grammar*
Tania Avgustinova and Yi Zhang
- 15:00-15:30 *LEXIE - an Experiment in Lexical Information Extraction*
John J. Camilleri and Michael Rosner
- 16:00-16:30 *Maximal Phrases Based Analysis for Prototyping Online Discussion Forums Postings*
Gaston Burek and Dale Gerdemann
- 16:30-17:00 *QALL-ME needs AIR: a portability study*
Constantin Orăsan, Iustin Dornescu and Natalia Ponomareva
- 17:00-17:10 Closing Remarks

