

Joint Workshop on NLP&LOD and SWAIE 2013

**Proceedings of the
Joint Workshop on NLP&LOD and
SWAIE:
Semantic Web, Linked Open Data and
Information Extraction**

associated with

**The 9th International Conference on
Recent Advances in Natural Language Processing
(RANLP 2013)**

12 September, 2013
Hissar, Bulgaria

Joint Workshop on NLP&LOD and SWAIE:
Semantic Web, Linked Open Data and Information Extraction
associated with THE INTERNATIONAL CONFERENCE
RECENT ADVANCES IN
NATURAL LANGUAGE PROCESSING'2013

PROCEEDINGS

Hissar, Bulgaria
12 September 2013

ISBN 978-954-452-025-0

Designed and Printed by INCOMA Ltd.
Shoumen, BULGARIA

Introduction

The Linked Open Data (LOD), understood as published structured data, which is interlinked and which builds upon standard Web technologies, such as HTTP and URIs, as well as on RDF-presented world facts datasets in various domains, has become a necessary component within all modern NLP-related tasks and applications since it provides large quantities of useful knowledge about people, facts, organizations, events, etc.

At the same time, the Information Extraction community specialises in mining the nuggets of information from text: such techniques could, however, be enhanced by annotated data or domain-specific resources. The Semantic Web community has already taken great strides in making these resources available through the Linked Open Data cloud, which are now ready for uptake by the Information Extraction community.

The Joint NLP&LOD and SWAIE Workshop presents papers that discuss the above mentioned topics from various perspectives. Enjoy reading them!

The Organizers

Organizers:

NLP&LOD organizers:

Petya Osenova, Sofia University
Kiril Simov, Bulgarian Academy of Sciences
Georgi Georgiev, OntoText Lab
Preslav Nakov, Qatar Computing Research Institute, Qatar Foundation, Qatar

SWAIE organizers:

Diana Maynard, University of Sheffield
Marieke van Erp, VU University Amsterdam
Brian Davis, DERI Galway

Program Committee:

NLP&LOD PC:

Eneko Agirre (University of the Basque Country, Spain)
Isabelle Augenstein (Sheffield University, UK)
Guido Boella (Università di Torino, Italy)
Kalina Boncheva (Sheffield University, UK)
António Branco (University of Lisbon, Portugal)
Nicoletta Calzolari (Istituto di Linguistica Computazionale, Italy)
Thierry Declerck (DFKI, Germany)
Georgi Dimitroff (Germany)
Kuzman Ganchev (Google, the USA)
Valia Kordoni (Humboldt University in Berlin, Germany)
Jarred McGinnis (King's College London, UK)
Pavel Mihajlov (Ontotext AD, Bulgaria)
Maciej Piasecki (Wroclaw University of Technology, Poland)
Laura Tolosi (Ontotext AD, Bulgaria)
Gertjan van Noord (University of Groningen, the Netherlands)
Piek Vossen (Vrije Universiteit Amsterdam, the Netherlands)

SWAIE PC:

Eneko Agirre (University of the Basque Country, Spain)
Paul Buitelaar (DERI, Galway, Ireland)
Matje van de Camp (Tilburg University, The Netherlands)
Philipp Cimiano (CITEC University of Bielefeld, Germany)
Hamish Cunningham (University of Sheffield, UK)
Thierry DeClerck (DFKI, Germany)
Antske Fokkens (VU University, the Netherlands)
Dirk Hovy (ISI, USA)
Georgeta Bordea (DERI Galway, Ireland)
Phil Gooch (City University London, UK)

Siegfried Handschuh (DERI, Ireland)
Véronique Malaisé (Elsevier, The Netherlands)
Laurette Pretorius (University of South Africa, South Africa)
German Rigau (University of the Basque Country, Spain)
Marco Rospocher (Fondazione Bruno Kessler, Italy)
Sara Tonelli (Fondazione Bruno Kessler, Italy)
Piek Vossen (VU University, The Netherlands)
René Witte (Concordia University, Montreal, Canada)
Birgit Proell (Johannes Kepler University Linz)
Killian Levacher (KDEG, CNGL)
Chris Welty (IBM, IBM T.J. Watson Research Center in New York)
D.J McCloskey (IBM Watson, IBM Dublin)

Table of Contents

<i>Linguistic Linked Open Data (LLOD) – Building the cloud</i>	
Christian Chiarcos	1
<i>Evaluation of SPARQL query generation from natural language questions</i>	
K. Bretonnel Cohen and Jin-Dong Kim	3
<i>Mining translations from the web of open linked data</i>	
John Philip McCrae and Philipp Cimiano.....	8
<i>Porting Elements of the Austrian Baroque Corpus onto the Linguistic Linked Open Data Format</i>	
Ulrike Czeitschner, Thierry Declerck and Claudia Resch.....	12
<i>Towards a System for Dynamic Language Resources in LOD</i>	
Kiril Simov	16
<i>Applications of Semantic Publishing</i>	
Borislav Popov	23
<i>Literature-driven Curation for Taxonomic Name Databases</i>	
Hui Yang, Alistair Willis, David Morse and Anne De Roeck	25
<i>Exploring the inference role in automatic information extraction from texts</i>	
Denis de Araujo, Sandro Rigo, Carolina Muller and Rove Chishman.....	33
<i>Combined analysis of news and Twitter messages</i>	
Mian Du, Jussi Kangasharju, Ossi Karkulahti, Lidia Pivovarova and Roman Yangarber	41
<i>Linguistically analyzed labels of knowledge objects: How can they support OBIE? Lessons learned from the Monnet and TrendMiner projects</i>	
Thierry Declerck	49

Conference Program

12.09.2013

Session 1: NLP+LOD (Part 1)

(9:15-10:15 (Invited Talk))

Linguistic Linked Open Data (LLOD) – Building the cloud
Christian Chiarcos

(10:15-10:45)

Evaluation of SPARQL query generation from natural language questions
K. Bretonnel Cohen and Jin-Dong Kim

Session 2: NLP+LOD (Part 2)

(11:15-11:45)

Mining translations from the web of open linked data
John Philip McCrae and Philipp Cimiano

(11:45-12:15)

Porting Elements of the Austrian Baroque Corpus onto the Linguistic Linked Open Data Format
Ulrike Czeitschner, Thierry Declerck and Claudia Resch

(12:15-12:45)

Towards a System for Dynamic Language Resources in LOD
Kiril Simov

12.09.2013 (continued)

Session 3: SWAIE (Part 1)

(14:00-15:00 (Invited Talk))

Applications of Semantic Publishing

Borislav Popov

(15:00-15:30)

Literature-driven Curation for Taxonomic Name Databases

Hui Yang, Alistair Willis, David Morse and Anne De Roeck

(15:30-16:00)

Exploring the inference role in automatic information extraction from texts

Denis de Araujo, Sandro Rigo, Carolina Muller and Rove Chishman

Session 4: SWAIE (Part 2)

(16:30-17:00)

Combined analysis of news and Twitter messages

Mian Du, Jussi Kangasharju, Ossi Karkulahti, Lidia Pivovarova and Roman Yangarber

(17:00-18:00 (Invited Talk))

Linguistically analyzed labels of knowledge objects: How can they support OBIE? Lessons learned from the Monnet and TrendMiner projects

Thierry Declerck