

# Preface

This volume contains the papers presented at LD4IE2014: Linked Data for Information Extraction, held on October 20, 2014 in Riva Del Garda, in conjunction with the 13th International Semantic Web Conference (ISWC2014). Additional workshop details can be found at <http://oak.dcs.shef.ac.uk/ld4ie2014>

The program committee selected 5 full papers for the program of LD4IE2014. In addition, the workshop, at its second edition, features an Information Extraction (IE) challenge, described in the paper “IE challenge overview”. The winning IE challenge system is presented in the short paper “ $\mu$ Raptor: A DOM-based system with appetite for hCard elements”. The invited speaker for LD4IE2014 is Pablo N. Mendes, with his talk “Where is the Semantics and where do we want it to be? Let the flaming begin!”.

- 9:00-9:05 Opening. Anna Lisa Gentile
- 9:05-09:55 Invited talk. Pablo Mendes  
Where is the Semantics and where do we want it to be?  
Let the flaming begin!
- 9:55-10:05 IE challenge overview. Robert Meusel and Heiko Paulheim
- 10:05-10:10 IE challenge lightening talk.  
 $\mu$ Raptor: A DOM-based system with appetite for hCard elements.  
Emir Muñoz, Luca Costabello and Pierre-Yves Vandenbussche
- 10:10-10:30 Cross-Document Coreference Resolution using Latent Features.  
Axel-Cyrille Ngonga Ngomo, Michael Röder and Ricardo Usbeck
- 10:30-11:00 Coffee break
- 11:00-11:20 Learning Regular Expressions for the Extraction of Product  
Attributes from E-commerce Microdata.  
Petar Petrovski, Volha Bryl and Christian Bizer
- 11:20-11:40 Learning Content Patterns from Linked Data.  
Emir Muñoz
- 11:40-12:00 Inductive Entity Typing Alignment.  
Giuseppe Rizzo, Marieke Van Erp and Raphaël Troncy
- 12:00-12:20 Online Index Extraction from Linked Open Data Sources.  
Fabio Benedetti, Sonia Bergamaschi and Laura Po
- 12:20-12:30 Closing and awards

October 15, 2014  
Sheffield

Anna Lisa Gentile

## Table of Contents

Where is the Semantics and where do we want it to be? Let the flaming begin! .....	1
<i>Pablo Mendes</i>	
Linked Data for Information Extraction Challenge 2014: Tasks and Results	3
<i>Robert Meusel and Heiko Paulheim</i>	
Online Index Extraction from Linked Open Data Sources .....	9
<i>Fabio Benedetti, Sonia Bergamaschi and Laura Po</i>	
Learning Content Patterns from Linked Data .....	21
<i>Emir Muñoz</i>	
Cross-Document Coreference Resolution using Latent Features .....	33
<i>Axel-Cyrille Ngonga Ngomo, Michael Röder and Ricardo Usbeck</i>	
Learning Regular Expressions for the Extraction of Product Attributes from E-commerce Microdata .....	45
<i>Petar Petrovski, Volha Bryl and Christian Bizer</i>	
Inductive Entity Typing Alignment .....	55
<i>Giuseppe Rizzo, Marieke Van Erp and Raphaël Troncy</i>	
$\mu$ Raptor: A DOM-based system with appetite for hCard elements .....	67
<i>Emir Muñoz, Luca Costabello and Pierre-Yves Vandenbussche</i>	

## Program Committee

Rabeeh Ayaz Abbasi	Quaid-i-Azam University, Islamabad
Isabelle Augenstein	University of Sheffield, UK
Payam Barnaghi	University of Surrey, UK
Pierpaolo Basile	University of Bari, Italy
Christopher Brewster	Aston University, UK
Paul Buitelaar	INSIGHT@NUI Galway, Ireland
Amparo Elizabeth Cano Basave	The Open University, UK
Claudia D'Amato	University of Bari, Italy
Brian Davis	INSIGHT@NUI Galway, Ireland
Ernesto William De Luca	Fachhochschule Potsdam, Germany
Nico Fanizzi	University of Bari, Italy
Christina Feilmayr	University of Linz, Austria
Miriam Fernandez	Open University, UK
Anna Lisa Gentile	University of Sheffield, UK
Katja Hose	Aalborg University, Denmark
Craig Knoblock	University of Southern California, USA
Petr Knoth	Open University, UK
Vanessa Lopez	IBM Dublin, Ireland
Diana Maynard	University of Sheffield, UK
Pablo Mendes	IBM Research, USA
Andrea Moro	Sapienza University of Rome, Italy
Varish Mulwad	UMBC ebiquity, USA
Matthias Nickles	INSIGHT@NUI Galway, Ireland
Andrea Giovanni Nuzzolese	University of Bologna, Italy
Heiko Paulheim	University of Mannheim, Germany
Simone Paolo Ponzetto	University of Mannheim, Germany
Jay Pujara	University of Maryland, USA
Disheng Qiu	Università Roma Tre, Italy
Will Radford	Xerox XRCE Grenoble, France
Martin Riedl	TU Darmstadt, Germany
Thomas Roth-Berghofer	University of West London, UK
Volker Tresp	Siemens AG, Germany
Victoria Uren	Aston University, UK
Xingxing Zhang	University of Edinburgh, UK
Ziqi Zhang	University of Sheffield, UK