

Ontolex 2010

**23rd International Conference on  
Computational Linguistics**

**Proceedings of the  
6th Workshop on Ontologies and Lexical  
Resources**

Alessandro Oltramari, Piek Vossen, and Qin Lu

22 August 2010

Beijing International Convention Center  
Beijing, China

Produced by  
*Chinese Information Processing Society of China*  
*All rights reserved for Coling 2010 CD production.*

To order the CD of Coling 2010 and its Workshop Proceedings, please contact:

Chinese Information Processing Society of China  
No.4, Southern Fourth Street  
Haidian District, Beijing, 100190  
China  
Tel: +86-010-62562916  
Fax: +86-010-62562916  
cips@iscas.ac.cn

## Introduction

Welcome to the Coling Workshop on *Ontologies and Lexical Resources (OntoLex 2010)*.

As human linguistic practice reveals, accessing to concepts through natural language is the implicit pathway for enabling mutual comprehension and effective meaning negotiation between agents in a community. But, in order to exchange knowledge, we need to share the *conceptual models underlying the lexicon*, namely ontologies. These remarks become even more crucial when focusing on human-computer interaction. In this context, computational ontologies and human-language technologies converge in the task of providing the semantic description of knowledge contents (e.g. multimedia, web resources, services, etc.): underlying intended models need to be made explicit in order to become accessible by artificial agents and sharable with humans. According to this picture, 1) computational lexicons, whose aim is to make lexical-content machine-understandable, constitute a fundamental component to foster the (mono- and multi-linguistic) access to any knowledge content; 2) computational ontologies, on the other side, are necessary to capture the logical structure of those knowledge contents: both contribute to dig out the basic elements of a given semantic space (domain-dependent or general), characterizing the different relations holding among them.

In this general framework, the contributions presented under the scope of OntoLex 2010 (Ontologies and Lexical Resources) show in fact a variety of approaches under many respects. Some of the papers are oriented to describe the different construction processes of semantic resources (e.g., Daoud et al. and Nagata deal with two approaches based on Wikipedia), other papers are especially concerned with specific tasks and applications. Regarding the latter aspect, some contributions present proposals to enhance interoperability within the various standardization formats for linguistic and terminological descriptions (Peters, Vossen et al.) as well as exploiting specific algorithms for ontology matching. Some papers also focus on formal ontology, both at the level of theoretical analysis and at the level of specific categories and relations (see for example the paper by Bogulaslavsky). The investigated domains span from bio-surveillance (Conway et al.) through medicine; sentiment/opinion mining confirms to be an emergent area of interest too (see Cadilhac et al.). Automatic techniques and algorithms to extract terms and taxonomies are also introduced (Van der Plas, Nagata et al., vor der Brueck).

Originating in 2000, OntoLex is recognized as a common *meeting place* by a constantly growing interdisciplinary community of lexicographers, ontologists and computational linguists. Traditionally represented by researchers and practitioners from a variety of backgrounds (acquisition of lexical knowledge, ontology-based approaches to information extraction, ontology learning, ontology matching, etc.), OntoLex 2010's contributions confirm this trend in the Sixth edition of the workshop too, hosted by COLING conference for the first time. We think that the comprehensive perspective emerging from the 10 articles collected in these proceedings can help in progress towards next-generation knowledge systems based on the integration between ontologies and lexical resources.

July, 2010

Alessandro Oltramari, Laboratory for Applied Ontology (ISTC-CNR) & Department of Management and Engineering (University of Padua) (Italy)

Piek Vossen, Faculty of Arts, VU University Amsterdam (The Netherlands)

Qin Lu, Department of Computing, The Hong Kong Polytechnic University (HKSAR)

**Workshop Organizers:**

Alessandro Oltramari, Laboratory for Applied Ontology (ISTC-CNR) & Department of Management and Engineering (University of Padua) (Italy)

Piek Vossen, Faculty of Arts, VU University Amsterdam (The Netherlands)

Qin Lu, Department of Computing, The Hong Kong Polytechnic University (HKSAR)

**Program Committee:**

Christiane Fellbaum, Princeton University(U.S.A.)

Paul Buitelaar, DERI, National University of Ireland(Ireland)

Philipp Cimiano, Delft University of Technology(The Netherlands)

Emanuele Pianta, Fondazione Bruno Kessler(Italy)

Massimo Poesio, University of Trento(Italy)

Ed Hovy, University of Southern California(U.S.A.)

Bolette Pedersen, University of Copenhagen(Denmark)

John Bateman, University of Bremen(Germany)

Armando Stellato, Universit di Roma "Tor Vergata"(Italy)

Chu-Ren Huang, The Hong Kong Polytechnic University(HKSAR)

Guido Vetere, Center for Advanced Studies, IBM(Italy)

Laurent Prevot, University of Provence(France)

Kiril Simov, Bulgarian Academy of Sciences(Bulgaria)

Alessandro Lenci, University of Pisa(Italy)

Ernesto De Luca, Technische Universität Berlin(Germany)

## Table of Contents

### *KYOTO: an open platform for mining facts*

Piek Vossen, German Rigau, Eneko Agirre, Aitor Soroa, Monica Monachini and Roberto Bartolini  
1

### *Using Goi-Taikai as an Upper Ontology to Build a Large-Scale Japanese Ontology from Wikipedia*

Masaaki Nagata, Yumi Shibaki and Kazuhide Yamamoto ..... 11

### *Multilingual Lexical Network from the Archives of the Digital Silk Road*

Hans-Mohammad Daoud, Kyo Kageura, Christian Boitet, Asanobu Kitamoto and Mathieu Mangeot..... 19

### *Finding Medical Term Variations using Parallel Corpora and Distributional Similarity*

Lonneke van der Plas and Jorg Tiedemann ..... 28

### *Learning Semantic Network Patterns for Hypernymy Extraction*

Tim vor der Bruck ..... 38

### *Intrinsic Property-based Taxonomic Relation Extraction from Category Structure*

DongHyun Choi, Eun-Kyung Kim, Sang-Ah Shim and Key-Sun Choi ..... 48

### *Developing a Biosurveillance Application Ontology for Influenza-Like-Illness*

Mike Conway, John Dowling and Wendy Chapman ..... 58

### *Interfacing the Lexicon and the Ontology in a Semantic Analyzer*

Igor Boguslavsky, Leonid Iomdin, Victor Sizov and Svetlana Timoshenko..... 67

### *Ontolexical resources for feature-based opinion mining: a case-study*

Anais Cadilhac, Farah Benamara and Nathalie Aussenac-Gilles ..... 77

## Conference Program

### Sunday, August 22, 2010

- 8:30–9:30 Workshop registration
- 9:30–10:30 Keynote by Prof. Huang Chu-Ren Huang(HK Polytechnic University)
- 10:30–10:50 Coffee break
- 10:50–11:15 *KYOTO: an open platform for mining facts*  
Piek Vossen, German Rigau, Eneko Agirre, Aitor Soroa, Monica Monachini and Roberto Bartolini
- 11:15–11:40 *Using Goi-Taikai as an Upper Ontology to Build a Large-Scale Japanese Ontology from Wikipedia*  
Masaaki Nagata, Yumi Shibaki and Kazuhide Yamamoto
- 11:40–12:05 *Multilingual Lexical Network from the Archives of the Digital Silk Road*  
Hans-Mohammad Daoud, Kyo Kageura, Christian Boitet, Asanobu Kitamoto and Mathieu Mangeot
- 12:05–12:30 *Finding Medical Term Variations using Parallel Corpora and Distributional Similarity*  
Lonneke van der Plas and Jorg Tiedemann
- 12:30–14:30 Lunch break
- 14:30–14:55 *Learning Semantic Network Patterns for Hypernymy Extraction*  
Tim vor der Bruck
- 14:55–15:20 *Intrinsic Property-based Taxonomic Relation Extraction from Category Structure*  
DongHyun Choi, Eun-Kyung Kim, Sang-Ah Shim and Key-Sun Choi
- 15:20–15:45 *Developing a Biosurveillance Application Ontology for Influenza-Like-Illness*  
Mike Conway, John Dowling and Wendy Chapman
- 15:45–16:00 Coffee break
- 16:00–16:25 *Interfacing the Lexicon and the Ontology in a Semantic Analyzer*  
Igor Boguslavsky, Leonid Iomdin, Victor Sizov and Svetlana Timoshenko

**Sunday, August 22, 2010 (continued)**

16:25–16:40 Conclusions

*Ontolexical resources for feature-based opinion mining: a case-study*

Anais Cadilhac, Farah Benamara and Nathalie Aussenac-Gilles