WoLE2012
ISWC2012 Workshop

Web of Linked Entities

Workshop in conjunction with the
11th International
Semantic Web Conference
Boston USA, November 11th 2012
http://wole2012.eurecom.fr/

Edited by:
Giuseppe Rizzo
Pablo N. Mendes
Eric Charton
Sebastian Hellmann
Aditya Kalyanpur

Introduction

The WoLE 2012 workshop envisions the Semantic Web as a Web of Linked Entities (WoLE), which transparently connects the World Wide Web (WWW) and the Giant Global Graph (GGG) using methods from Information Retrieval (IR) and Natural Language Processing (NLP).

Topics of interest to WoLE 2012 include:
1. Improvements upon the state of the art in NLP using information in the Linked Open Data (LOD) space;
2. Knowledge extraction from text and HTML documents (or other structured and semi-structured documents) on the Web, with a special focus on scalability, evaluation of precision & recall and/or real-time systems;
3. Representation of NLP tool output and NLP resources as RDF/OWL and especially connections to linked data;
4. Novel applications to search and browse the WWW with the help of extracted knowledge and the Web of Data.

The focus of this workshop is to reconcile the communities of Information Retrieval, Semantic Web and NLP. The primary goal is to strengthen research techniques that provide access to textual information published on the Web to further improve the adoption of Semantic Web technology.

Motivation

Most of the knowledge available on the Web is present as natural language text enclosed in Web documents aimed at human consumption. A promising approach to have programmatic access to such knowledge uses information extraction techniques in order to reduce texts written in natural languages to machine readable structures, from which it is possible to retrieve entities and relations. The Natural Language Processing (NLP) community has been approaching this crucial task for the past few decades, with two major

guidelines: establishing standards for various tasks, and metrics to evaluate the performance of algorithms. Scientific evaluation campaigns, starting in 2003 with CoNLL, ACE (2005, 2007), TAC (2009, 2010, 2011, 2012), and ETAPE in 2012 were proposed to involve and compare the performance of various systems in a rigorous and reproducible manner. Various techniques have been proposed along this period to recognize entities mentioned in text and to classify them according to a small set of entity types.

Recently, an increasing number of researchers have investigated information extraction techniques in the context of Semantic Web research. Working in the intersection with the NLP community, researchers have used fine grained ontologies to classify entities and proposed disambiguation techniques to map these pieces of information to real world entities. Therefore, the Web represents a vital lookup space where entities extracted from textual documents can be disambiguated. Moreover, it offers a broad range of relationships that already exist among entities. The landscape of available techniques vary on their approaches and performance, leading to new evaluation campaigns being proposed -- focusing on the extraction of richer information as compared to previous work. The final results of such information extraction tasks may potentially be consumed in the LOD cloud, as exemplified by efforts of the NLP2RDF/NIF community.

The focus of this workshop is to strengthen the connection between the communities of Information Retrieval, Semantic Web and NLP -- reconciling research and techniques that provide access to textual information published on the Web to further improve the adoption of Semantic Web technology.

Programme Committee

The following colleagues kindly served in the workshop's program committee. Their joint expertise covers all of the questions addressed in the workshop, and they reflect the range of relevant scientific communities.

- Caroline Barriere, Centre de Recherche Informatique de Montréal, DETI, Canada

- Frédéric Béchet, Université d'Aix-Marseille, LIF, France
- Andreas Blumauer, Semantic Web Company
- Paul Buitelaar, DERI/National University of Ireland, Galway, Ireland
- Philipp Cimiano, CITEC, University of Bielefeld, Germany
- Eric de la Clergerie, INRIA, France
- Christian Dirschl, Wolters Kluwer, Germany
- Benoit Favre, Université d'Aix-Marseille, LIF, France
- Michel Gagnon, École Polytechnique de Montréal, Canada
- Daniel Gerber, AKSW, Universität Leipzig, Germany
- Claudio Giuliano, Fondazione Bruno Kessler, Italy
- Jiafeng Guo, Institute of Computing Technology, China
- Daniel Hladky, Ontos AG
- Guy Lapalme, Université de Montréal, RALI-DIRO, Canada
- Paul McNamee, Johns Hopkins University, USA
- Marie-Jean Meurs, Semantic Software Lab, CSFG, Concordia University, Canada
- Meenakshi Nagarajan, IBM Research, USA
- Axel-C Ngonga Ngomo, AKSW, Universität Leipzig, Germany
- Cartic Ramakrishnan, ISI, University of Southern California, USA
- Ganesh Ramakrishnan, IIT Bombay, India
- Harald Sack, HPI, University of Potsdam, Germany
- Benoit Sagot, INRIA, France
- Felix Sasaki, DFKI-LT, German Research Center for Artificial Intelligence (DFKI), Germany
- Tomas Steiner, Universitat de Catalunya, Spain
- Fabian Suchanek, Max-Planck Institute for Informatics, Saarbrücken, Germany
- Vojtech Svatek, University of Economics, Prague
- Krishnaprasad Thirunarayan , Wright State University, USA
- Andraz Tori, Zemanta
- Ruben Verborgh, IBBT, Ghent University, Belgium
- Mateja Verlic, Zemanta
- Wouter Weerkamp, University of Amsterdamm, Netherlands
- Gerhard Weikum, Max-Planck Institute for Informatics, Saarbrücken, Germany
- René Witte, University of Montréal, Canada

- Feiyu Xu, DFKI-LT, German Research Center for Artificial Intelligence (DFKI), Germany

Additional reviewers:

- Magnus Knuth, University of Potsdam, Germany
- James Mayfield, Johns Hopkins University, USA
- Nadine Steinmetz, University of Potsdam, Germany

Organizing Committee

The WoLE 2012 workshop was organised by:
- Giuseppe Rizzo, EURECOM, France
- Pablo N. Mendes, Freie Universität Berlin, Germany,
- Eric Charton, Centre de Recherche Informatique de Montréal, Canada
- Sebastian Hellmann, Universität Leipzig, Germany
- Aditya Kalyanpur, IBM, USA