

Communications in Computer and Information Science

641

Commenced Publication in 2007

Founding and Former Series Editors:

Alfredo Cuzzocrea, Dominik Ślęzak, and Xiaokang Yang

Editorial Board

Simone Diniz Junqueira Barbosa

*Pontifical Catholic University of Rio de Janeiro (PUC-Rio),
Rio de Janeiro, Brazil*

Phoebe Chen

La Trobe University, Melbourne, Australia

Xiaoyong Du

Renmin University of China, Beijing, China

Joaquim Filipe

Polytechnic Institute of Setúbal, Setúbal, Portugal

Orhun Kara

TÜBİTAK BİLGEM and Middle East Technical University, Ankara, Turkey

Igor Kotenko

*St. Petersburg Institute for Informatics and Automation of the Russian
Academy of Sciences, St. Petersburg, Russia*

Ting Liu

Harbin Institute of Technology (HIT), Harbin, China

Krishna M. Sivalingam

Indian Institute of Technology Madras, Chennai, India

Takashi Washio

Osaka University, Osaka, Japan

More information about this series at <http://www.springer.com/series/7899>

Harald Sack · Stefan Dietze
Anna Tordai · Christoph Lange (Eds.)

Semantic Web Challenges

Third SemWebEval Challenge at ESWC 2016
Heraklion, Crete, Greece, May 29 – June 2, 2016
Revised Selected Papers

Editors

Harald Sack
IT Systems Engineering
Hasso-Plattner Institute
Potsdam
Germany

Stefan Dietze
Leibniz Universität Hannover
Hannover
Germany

Anna Tordai
Elsevier B.V.
Amsterdam
The Netherlands

Christoph Lange
Universität Bonn
Bonn
Germany

ISSN 1865-0929

ISSN 1865-0937 (electronic)

Communications in Computer and Information Science

ISBN 978-3-319-46564-7

ISBN 978-3-319-46565-4 (eBook)

DOI 10.1007/978-3-319-46565-4

Library of Congress Control Number: 2016951976

© Springer International Publishing Switzerland 2016

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

This Springer imprint is published by Springer Nature

The registered company is Springer International Publishing AG Switzerland

Preface

Common benchmarks, established evaluation procedures, comparable tasks, and public datasets are vital to ensure reproducible, evaluable, and comparable scientific results. To assess the current state of the art and foster the systematic comparison of contributions to the Semantic Web community, open challenges are now a key scientific element of the Semantic Web conferences. Following the success of the previous years, reflected by the high number of high-quality submissions, we organized the third edition of the “Semantic Web Challenges” as an official track of the ESWC 2016 conference (held in Anissaras, Crete, Greece, from May 30 to June 2, 2016), one of the most important international scientific events for the Semantic Web research community. The purpose of challenges is to validate the maturity of the state of the art in tasks common to the Semantic Web community and adjacent academic communities in a controlled setting of rigorous evaluation, thereby providing sound benchmarks, datasets, and evaluation approaches, which contribute to the advancement of the state of the art. This third edition included five challenges: Open Knowledge Extraction (OKE 2016), Semantic Publishing (SemPub 2016), Semantic Sentiment Analysis (SSA 2016), Question Answering over Linked Data (QALD 6) and Top-K Shortest Paths in RDF Datasets. A total of 31 teams competed in the different challenges. The event attracted attendees from across the conference, with a high attendance for all challenge-related activities during ESWC 2016. This included the dedicated conference track and participation of challenge candidates during the ESWC poster and demo session. The very positive feedback and resonance suggests that the ESWC challenges provided a central contribution to the ESWC 2016 program.

This book includes the descriptions of all methods and tools that competed at the “Semantic Web Challenges 2016,” together with a detailed description of the tasks, evaluation procedures, and datasets, offering to the community a snapshot of the advancement in those areas at that moment in time and material for replications of the results. The editors have divided the book content into five chapters, each dedicated to one area, or challenge. Each chapter includes an introductory section by the challenge chairs providing a detailed description of the challenge tasks, the evaluation procedure, and associated datasets and peer-reviewed descriptions of the participants’ methods, tools, and results.

We would like to thank all challenge chairs, whose hard work during the organization of the 2016 edition of the Semantic Web Challenges helped in making it a fruitful event. Thanks to their work, we experienced a successful and inspiring scientific event, and we are now able to deliver this book to the community.

June 2016

Harald Sack
Stefan Dietze
Anna Tordai

Organization

Organizing Committee (ESWC)

General Chair

Harald Sack Hasso Plattner Institute (HPI), Germany

Program Chairs

Mathieu d'Aquin Knowledge Media Institute KMI, UK
Eva Blomqvist Linköping University, Sweden

Workshops Chairs

Dunja Mladenic Jožef Stefan Institute, Slovenia
Sören Auer University of Bonn and Fraunhofer IAIS, Germany

Poster and Demo Chairs

Giuseppe Rizzo Istituto Superiore Mario Boella, Italy
Nadine Steinmetz Technische Universität Ilmenau, Germany

Tutorials Chairs

Tommaso di Noia Politecnico di Bari, Italy
H. Sofia Pinto INESC-ID, Instituto Superior Técnico, Universidade de Lisboa, Portugal

PhD Symposium Chairs

Simone Paolo Ponzetto Universität Mannheim, Germany
Chiara Ghidini Fondazione Bruno Kessler, Italy

Challenge Chairs

Stefan Dietze L3S, Germany
Anna Tordai Elsevier, The Netherlands

Semantic Technologies Coordinators

Andrea Giovanni Nuzzolese University of Bologna/STLab ISTC-CNR, Italy
Anna Lisa Gentile University of Mannheim, Germany

EU Project Networking Session Chairs

Erik Mannens Data Science Lab – iMinds – Ghent University, Belgium

Mauro Dragoni Fondazione Bruno Kessler, Italy
Lyndon Nixon Modul Universität Vienna, Austria
Oscar Corcho Universidad Politécnica de Madrid, Spain

Publicity Chair

Heiko Paulheim University of Mannheim, Germany

Sponsor Chairs

Steffen Lohmann Fraunhofer IAIS, Germany
Freddie Lecue IBM, Ireland

Web Presence

Venislav Georgiev STI International, Austria

Proceedings Chair

Christoph Lange University of Bonn and Fraunhofer IAIS, Germany

Treasurer

Ioan Toma STI International, Austria

Local Organization and Conference Administration

Katharina Haas YouVivo GmbH, Germany

Challenges Organization

Challenge Chairs

Stefan Dietze L3S, Germany
Anna Tordai Elsevier, The Netherlands

Open Knowledge Extraction Challenge

Andrea Giovanni Nuzzolese STLab-CNR, Italy
Anna Lisa Gentile University of Mannheim, Germany
Valentina Presutti STLab-CNR, Italy
Robert Meusel University of Mannheim, Germany
Heiko Paulheim University of Mannheim, Germany
Aldo Gangemi Université Paris 13, France

ESWC-16 Challenge on Semantic Sentiment Analysis

Mauro Dragoni Fondazione Bruno Kessler, Italy
Diego Reforgiato Recupero University of Cagliari, Italy

6th Open Challenge on Question Answering over Linked Data

Christina Unger	Universität Bielefeld, Germany
Axel-Cyrille Ngonga Ngomo	University of Leipzig, Germany
Elena Cabrio	University of Nice Sophia Antipolis, France

Top-K Shortest Path in Large Typed RDF Graphs Challenge

Ioannis Papadakis	Ionian University, Greece
Michalis Stefanidakis	Ionian University, Greece
Phivos Mylonas	Ionian University, Greece
Brigitte Endres-Niggemeyer	Hannover University for Applied Sciences, Germany
Spyridon Kazanas	Ionian University, Greece

Semantic Publishing Challenge 2016

Angelo Di Iorio	University of Bologna, Italy
Anastasia Dimou	Ghent University, Belgium
Christoph Lange	University of Bonn/Fraunhofer IAIS, Germany
Sahar Vahdati	University of Bonn, Germany

Sponsoring Institutions**Gold Sponsors**

Silver Sponsors



<http://byte-project.eu/>



<http://entropy-project.eu/>

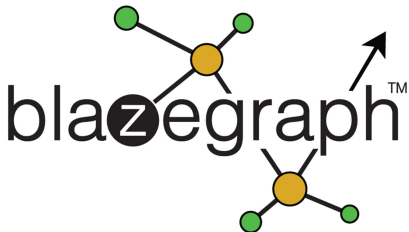


<http://project-hobbit.eu/>



<http://www.springer.com/lncs>

Sponsors of Specific Challenges



<http://www.blazegraph.com/>

Contents

Open Knowledge Extraction Challenge

The Second Open Knowledge Extraction Challenge	3
<i>Andrea Giovanni Nuzzolese, Anna Lisa Gentile, Valentina Presutti, Aldo Gangemi, Robert Meusel, and Heiko Paulheim</i>	
Enhancing Entity Linking by Combining NER Models	17
<i>Julien Plu, Giuseppe Rizzo, and Raphaël Troncy</i>	
Collective Disambiguation and Semantic Annotation for Entity Linking and Typing.	33
<i>Mohamed Chabchoub, Michel Gagnon, and Amal Zouaq</i>	
DWS at the 2016 Open Knowledge Extraction Challenge: A Hearst-Like Pattern-Based Approach to Hypernym Extraction and Class Induction	48
<i>Stefano Faralli and Simone Paolo Ponzetto</i>	
Entity Typing and Linking Using SPARQL Patterns and DBpedia	61
<i>Lara Haidar-Ahmad, Ludovic Font, Amal Zouaq, and Michel Gagnon</i>	

Challenge on Semantic Sentiment Analysis

Challenge on Fine-Grained Sentiment Analysis Within ESWC2016.	79
<i>Mauro Dragoni and Diego Reforgiato Recupero</i>	
App2Check Extension for Sentiment Analysis of Amazon Products Reviews	95
<i>Emanuele Di Rosa and Alberto Durante</i>	
Sentiment Polarity Detection from Amazon Reviews: An Experimental Study	108
<i>Efstratios Sygkounas, Giuseppe Rizzo, and Raphaël Troncy</i>	
Exploiting Propositions for Opinion Mining	121
<i>Andi Rexha, Mark Kröll, Mauro Dragoni, and Roman Kern</i>	
The IRMUDOSA System at ESWC-2016 Challenge on Semantic Sentiment Analysis	126
<i>Giulio Petrucci and Mauro Dragoni</i>	
A Knowledge-Based Approach for Aspect-Based Opinion Mining.	141
<i>Marco Federici and Mauro Dragoni</i>	

Aspect-Based Sentiment Analysis Using a Two-Step Neural Network Architecture. 153
Soufian Jebbara and Philipp Cimiano

Challenge on Question Answering over Linked Data

6th Open Challenge on Question Answering over Linked Data (QALD-6) . . . 171
Christina Unger, Axel-Cyrille Ngonga Ngomo, and Elena Cabrio

SPARKLIS on QALD-6 Statistical Questions 178
Sébastien Ferré

Top-k Shortest Paths in Large Typed RDF Datasets Challenge

Top-K Shortest Paths in Large Typed RDF Datasets Challenge. 191
Ioannis Papadakis, Michalis Stefanidakis, Phivos Mylonas, Brigitte Endres Niggemeyer, and Spyridon Kazanas

Top-k Shortest Paths in Directed Labeled Multigraphs. 200
Sven Hertling, Markus Schröder, Christian Jilek, and Andreas Dengel

Modified MinG Algorithm to Find Top-K Shortest Paths from large RDF Graphs. 213
Zohaib Hassan, Mohammad Abdul Qadir, Muhammad Arshad Islam, Umer Shahzad, and Nadeem Akhter

Using Triple Pattern Fragments to Enable Streaming of Top-k Shortest Paths via the Web. 228
Laurens De Vocht, Ruben Verborgh, and Erik Mannens

Semantic Publishing Challenge

Semantic Publishing Challenge – Assessing the Quality of Scientific Output in Its Ecosystem 243
Anastasia Dimou, Angelo Di Iorio, Christoph Lange, and Sahar Vahdati

Reconstructing the Logical Structure of a Scientific Publication Using Machine Learning 255
Stefan Klampfl and Roman Kern

Automatically Identify and Label Sections in Scientific Journals Using Conditional Random Fields. 269
Sree Harsha Ramesh, Arnab Dhar, Raveena R. Kumar, Anjaly V., Sarath K.S., Jason Pearce, and Krishna R. Sundaresan

ACM: Article Content Miner for Assessing the Quality of Scientific Output . . . 281
*Andrea Giovanni Nuzzolese, Silvio Peroni,
and Diego Reforgiato Recupero*

Information Extraction from PDF Sources Based on Rule-Based System
Using Integrated Formats 293
Riaz Ahmad, Muhammad Tanvir Afzal, and Muhammad Abdul Qadir

An Automatic Workflow for the Formalization of Scholarly Articles’
Structural and Semantic Elements 309
Bahar Sateli and René Witte

Author Index 321