### Lecture Notes in Computer Science

8797

Commenced Publication in 1973
Founding and Former Series Editors:
Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

### **Editorial Board**

David Hutchison

Lancaster University, UK

Takeo Kanade

Carnegie Mellon University, Pittsburgh, PA, USA

Josef Kittler

University of Surrey, Guildford, UK

Jon M. Kleinberg

Cornell University, Ithaca, NY, USA

Alfred Kobsa

University of California, Irvine, CA, USA

Friedemann Mattern

ETH Zurich. Switzerland

John C. Mitchell

Stanford University, CA, USA

Moni Naor

Weizmann Institute of Science, Rehovot, Israel

Oscar Nierstrasz

University of Bern, Switzerland

C. Pandu Rangan

Indian Institute of Technology, Madras, India

Bernhard Steffen

TU Dortmund University, Germany

Demetri Terzopoulos

University of California, Los Angeles, CA, USA

Doug Tygar

University of California, Berkeley, CA, USA

Gerhard Weikum

Max Planck Institute for Informatics, Saarbruecken, Germany

Peter Mika Tania Tudorache Abraham Bernstein Chris Welty Craig Knoblock Denny Vrandečić Paul Groth Natasha Noy Krzysztof Janowicz Carole Goble (Eds.)

# The Semantic Web – ISWC 2014

13th International Semantic Web Conference Riva del Garda, Italy, October 19-23, 2014 Proceedings, Part II



### Volume Editors

Peter Mika, Yahoo Labs, Barcelona, Spain

E-mail: pmika@yahoo-inc.com

Tania Tudorache, Stanford University, CA, USA

E-mail: tudorache@stanford.edu

Abraham Bernstein, University of Zurich, Switzerland

E-mail: bernstein@ifi.uzh.ch

Chris Welty, IBM Research, Yorktown Heights, NY, USA

E-mail: welty@us.ibm.com

Craig Knoblock, University of Southern California, Los Angeles, CA, USA

E-mail: knoblock@isi.edu

Denny Vrandečić, Natasha Noy, Google, USA

E-mail: vrandecic@google.com; natashafn@acm.org

Paul Groth, VU University Amsterdam, The Netherlands

E-mail: p.t.groth@vu.nl

Krzysztof Janowicz, University of California, Santa Barbara, CA, USA

E-mail: jano@geog.ucsb.edu

Carole Goble, The University of Manchester, UK

E-mail: carole.goble@manchester.ac.uk

ISSN 0302-9743 e-ISSN 1611-3349

ISBN 978-3-319-11914-4 e-ISBN 978-3-319-11915-1

DOI 10.1007/978-3-319-11915-1

Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014949616

LNCS Sublibrary: SL 3 – Information Systems and Application, incl. Internet/Web and HCI

© Springer International Publishing Switzerland 2014

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. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in ist current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

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.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

### **Preface**

The Semantic Web is now a maturing field with a significant and growing adoption of semantic technologies in a variety of commercial, public sector, and scientific fields. Linked Data is pervasive: from enabling government transparency, to helping integrate data in life sciences and enterprises, to publishing data about museums, and integrating bibliographic data. Significantly, major companies, such as Google, Yahoo, Microsoft, and Facebook, have created their own "knowledge graphs" that power semantic searches and enable smarter processing and delivery of data: The use of these knowledge graphs is now the norm rather than the exception. The schema.org effort led by the major search companies illustrates the industry interest and support of the Semantic Web. Commercial players such as IBM, Siemens, BestBuy, and Walmart are seeing the value of semantic technologies and are regular presenters at Semantic Web conferences. The papers and the research topics covered in these proceedings follow directly from the requirements of this large adoption, and contribute greatly to the continuing success of the field.

The International Semantic Web Conference is the premier forum for Semantic Web research, where cutting-edge scientific results and technological innovations are presented, where problems and solutions are discussed, and where the future of this vision is being developed. It brings together specialists in fields such as artificial intelligence, databases, social networks, distributed computing, Web engineering, information systems, human–computer interaction, natural language processing, and the social sciences for tutorials, workshops, presentations, keynotes, and ample time for detailed discussions.

This volume contains the main proceedings of the 13th International Semantic Web Conference (ISWC 2014), which was held in Riva del Garda, Trentino, Italy, in October 2014. We received tremendous response to our calls for papers from a truly international community of researchers and practitioners. Indeed, several tracks of the conference received a record number of submissions this year. The careful nature of the review process, and the breadth and scope of the papers finally selected for inclusion in this volume, speak to the quality of the conference and to the contributions made by researchers whose work is presented in these proceedings. As such, we were all honored and proud that we were invited to serve the community in the stewardship of this edition of ISWC.

The proceedings include papers from four different tracks: the Research Track, the Semantic Web In-Use Track, the newly added Replication, Benchmark, Data and Software (RBDS) Track, and a selection of Doctoral Consortium papers. For the first time since we started publishing the LNCS proceedings, the papers are organized by their topic rather than by their track and correspond closely to the sessions in the conference schedule. The topics of the accepted papers reflect

the broad coverage of the Semantic Web research and application: Linked Data, its quality, link discovery, and application in the life sciences; data integration, search and query answering, SPARQL, ontology-based data access and query rewriting and reasoning; natural language processing and information extraction; user interaction and personalization, and social media; ontology alignment and modularization; and sensors and streams.

Creating the program for ISWC 2014 would not have been possible without the tireless and fantastic work of the Senior Program Committees (SPC), the Program Committees (PC), as well as of the many sub-reviewers in the different tracks, several of whom volunteered to provide high-quality emergency reviews. To acknowledge this work, the Research Track and the Semantic Web In-Use Track each offered a best reviewer award. The decision on the awards was taken with the input of the SPC members, of the fellow reviewers from the PC, of the authors, and also using objective measures about the reviews provided by EasyChair, the conference management system.

The Research Track of the conference attracted 180 submissions, 38 of which were accepted, resulting in a 21% acceptance rate. Each paper received at least three, and sometimes as many as five, reviews from members of the PC. After the first round of reviews, authors had the opportunity to submit a rebuttal, leading to further discussions among the reviewers, a metareview and a recommendation from a member of the SPC. The SPC held a 10-hour virtual meeting in order to select the final set of accepted papers, paying special attention to papers that were borderline or had at least one recommendation for acceptance. In many cases, additional last-minute reviews were sought out to better inform the SPC's decision.

The best paper nominations for the Research Track reflect the broad range of topics that were submitted to this track:

Best paper nominations:

- AGDISTIS "Graph-Based Disambiguation of Named Entities Using Linked Data" by Ricardo Usbeck, Axel-Cyrille Ngonga Ngomo, Michael Röder, Daniel Gerber, Sandro Athaide Coelho, Sören Auer and Andreas Both
- "Expressive and Scalable Query-based Faceted Search over SPARQL Endpoints" by Sébastien Ferré
- "Explass: Exploring Associations Between Entities via Top-K Ontological Patterns and Facets" by Gong Cheng, Yanan Zhang and Yuzhong Qu
- "Querying Factorized Probabilistic Triple Databases" by Denis Krompaß, Maximilian Nickel and Volker Tresp

### Best student paper nominations:

- "OBDA: Query Rewriting or Materialization? In Practice, Both!" by Juan
   F. Sequeda, Marcelo Arenas and Daniel P. Miranker
- "SYRql: A Dataflow Language for Large Scale Processing of RDF Data" by Fadi Maali, Padmashree Ravindra, Kemafor Anyanwu and Stefan Decker
- "Pushing the Boundaries of Tractable Ontology Reasoning" by David Carral,
   Cristina Feier, Bernardo Cuenca Grau, Pascal Hitzler and Ian Horrocks

The Semantic Web In-Use Track received 46 submissions. Fifteen papers were accepted – a 33% acceptance rate. The papers demonstrated how semantic technologies are applied in a variety of domains, including: biomedicine and drug discovery, smart cities, sensor streams, multimedia, visualization, link generation, and ontology development. The application papers demonstrated how semantic technologies are applied in diverse ways, starting from using linked data in mobile environments to employing fully fledged artificial intelligence methods in real-time use cases. At least three members of the In-Use PC provided reviews for each paper. After the first round of reviews, authors had the opportunity to submit a rebuttal, leading to further discussions among the reviewers, a metareview and a recommendation from a member of the SPC.

The best paper nominations for the Semantic Web In-Use Track are:

- "Web Browser Personalization with the Client Side Triplestore" by Hitoshi Uchida, Ralph Swick and Andrei Sambra
- "Semantic Traffic Diagnosis with STAR-CITY: Architecture and Lessons Learned from Deployment in Dublin, Bologna, Miami and Rio", by Freddy Lecue, Robert Tucker, Simone Tallevi-Diotallevi, Rahul Nair, Yiannis Gkoufas, Giuseppe Liguori, Mauro Borioni, Alexandre Rademaker and Luciano Barbosa
- "Adapting Semantic Sensor Networks for Smart Building Analytics" by Joern Ploennigs, Anika Schumann and Freddy Lecue

This year we introduced the Replication, Benchmark, Data and Software (RBDS) track that provides an outlet for papers of these four categories. It extended and transformed last year's evaluations and experiments track to incorporate new categories of contributions. The four types of papers had very clearly specified scope and reviewing criteria that were described in the Call for Papers: (1) Replication papers focus on replicating a previously published approach in order to shed light on some important, possibly overlooked aspect; (2) benchmark papers make available to the community a new class of resources, metrics or software that can be used to measure the performance of systems in some dimension; (3) data papers introduce an important data set to the community; and (4) software framework papers advance science by sharing with the community software that can easily be extended or adapted to support scientific study and experimentation. The RBDS track received 39 submissions (18 benchmark studies, eight data papers, eight software framework papers, and four replication studies), and accepted 16 papers (five benchmark studies, five data papers, four software framework papers, and two replication studies), corresponding to an acceptance rate of 41%. Each paper was reviewed by at least three members of the PC and discussed thoroughly. The papers address a range of areas, such as linked stream data, federated query processing, tag recommendation, entity summarization, and mobile semantic web.

The Doctoral Consortium is a key event at the ISWC conference. PhD students in the Semantic Web field get an opportunity to present their thesis proposals and to interact with leading academic and industrial scientists in the field, who act as their mentors. The Doctoral Consortium received 41 submissions, a record number compared to previous years. Each paper received two reviews, one from an SPC member, and one from a co-chair. Out of 41 submissions, six were selected to be both included in these proceedings and for presentation at the Doctoral Consortium, while an additional 11 were selected for presentation. The Doctoral Consortium day is organized as a highly interactive event, in which students present their proposals and receive extensive feedback and comments from mentors as well as from their peers.

A unique aspect of the ISWC conference is the Semantic Web Challenge, now in its 12th year, with the goal of demonstrating practical progress toward achieving the vision of the Semantic Web. The overall objective of the challenge is to apply Semantic Web techniques in building online end-user applications that integrate, combine, and deduce information needed to assist users in performing tasks. Organized this year by Andreas Harth and Sean Bechhofer, the competition enables practitioners and scientists to showcase leading-edge real-world applications of Semantic Web technology. The Semantic Web Challenge is advised by a board of experts working at universities and in industry. The advisory board also acts as a jury and awards the best applications at the conference.

The keynote talks given by leading scientists or practitioners in their field further enriched the ISWC program. Prabhakar Raghavan, Vice-President of Engineering at Google, discussed "Web Search – From the Noun to the Verb." Paolo Traverso, Director of the Center for Information Technology at Fondazione Bruno Kessler, talked about "To Be or to Do?: The Semantics for Smart Cities and Communities." Yolanda Gil, Associate Director of the Intelligent Systems Division at ISI University of South California, discussed the "Semantic Challenges in Getting Work Done" addressing the application of semantics to scientific tasks. The industry track featured a plenary keynote on "The Semantic Web in an Age of Open Data" by Sir Nigel Shadbolt, Chairman and Co-Founder of the UK's Open Data Institute and Professor of Artificial Intelligence at the University of Southampton.

As in previous ISWC editions, the conference included an extensive tutorial and workshop program. Johanna Völker and Lora Aroyo, the chairs of this track, created a stellar and diverse collection of eight tutorials and 23 workshops, where the only problem that the participants faced was which of the many exciting workshops and tutorials to attend. This year, we hosted for the first time the Developers' Workshop, a dedicated event for software developers discussing implementations, methods, techniques, and solutions to practical problems of Semantic Web and Linked Data. The main topic of the Developers' Workshop was "Semantic Web in a Browser."

We would like to thank Matthew Horridge, Marco Rospocher, and Jacco van Ossenbruggen for organizing a lively poster and demo session. This year, the track got a record 156 submissions, a 50% increase compared with previous years. Moreover, 71 posters and 50 demos were introduced in a "minute madness session," where each presenter got 45 seconds to provide a teaser for their poster or demo. Axel Polleres, Alexander Castro, and Richard Benjamins coordinated an exciting Industry Track with presentations both from younger companies focusing on semantic technologies and from large enterprises, such as British Telecom, IBM, Oracle, and Siemens, just to name a few. With a record number of 39 submissions (seven of which were selected for full presentations and 23 for short lightning talks) in the industry track this year, the mix of presentations demonstrated the success and maturity of semantic technologies in a variety of industry- and business-relevant domains. The extended abstracts for posters, demos, and industry talks will be published in separate companion volumes in the CEUR workshop proceedings series.

We are indebted to Krzysztof Janowicz, our proceedings chair, who provided invaluable support in compiling the volume that you now hold in your hands (or see on your screen) and who put in many hours of additional work to create a completely new structure for these proceedings based on the topic rather than the tracks, as in previous years. Many thanks to Oscar Corcho and Miriam Fernandez, the student coordinators, for securing and managing the distribution of student travel grants and thus helping students who might not have otherwise attended the conference to come to Riva. Roberta Cuel, Jens Lehmann, and Vincenzo Maltese were tireless in their work as sponsorship chairs, knocking on every conceivable virtual "door" and ensuring an unprecedented level of sponsorship this year. We are especially grateful to all the sponsors for their generosity.

As has been the case in the past, ISWC 2014 also contributed to the Linked Data cloud by providing semantically annotated data about many aspects of the conference. This contribution would not have been possible without the efforts of Li Ding and Jie Bao, our metadata chairs.

Mauro Dragoni, our publicity chair, tirelessly tweeted, sent old-fashioned announcements on the mailing lists, and updated the website, creating more lively "buzz" than ISWC has had before.

Our very special thanks go to the local organization team, led by Luciano Serafini and Chiara Ghidini. They did a fantastic job of handling local arrangements, thinking of every potential complication way before it arose, often doing things when members of the Organizing Committee were only beginning to think about asking for them. Many thanks to the Rivatour Agency for providing great service for local arrangements.

### X Preface

Finally, we would like to thank all members of the ISWC Organizing Committee not only for handling their tracks superbly, but also for their wider contribution to the collaborative decision-making process in organizing the conference.

October 2014

Peter Mika
Tania Tudorache
Abraham Bernstein
Chris Welty
Craig Knoblock
Denny Vrandečić
Paul Groth
Natasha Noy
Krzysztof Janowicz
Carole Goble

### Conference Organization

### Organizing Committee

General Chair

Carole Goble University of Manchester, UK

**Local Chairs** 

Chiara Ghidini Fondazione Bruno Kessler, Italy Luciano Serafini Fondazione Bruno Kessler, Italy

**Program Committee Chairs** 

Peter Mika Yahoo Labs, Spain

Tania Tudorache Stanford University, USA

Abraham Bernstein University of Zurich, Switzerland

Chris Welty IBM Research, USA

Craig Knoblock University of Southern California, USA

Denny Vrandečić Google, USA

Research Track Chairs

Peter Mika Yahoo Labs, Spain

Tania Tudorache Stanford University, USA

In-Use Track Chairs

Craig Knoblock University of Southern California, USA

Denny Vrandečić Google, USA

**Industry Track Chairs** 

Richard Benjamins Telefonica, Spain

Alexander G. Castro LinkingData I/O LLC, USA

Axel Polleres Vienna University of Economics and

Business, Austria

Replication, Benchmark, Data and Software Track Chairs

Abraham Bernstein University of Zurich, Switzerland

Chris Welty IBM Research, USA

### Workshops and Tutorials Chairs

Lora Aroyo VU University Amsterdam,

The Netherlands

Johanna Völker University of Mannheim, Germany

### Posters and Demos Chairs

Matthew Horridge Stanford University, USA

Marco Rospocher Fondazione Bruno Kessler, Italy

Jacco van Ossenbruggen Centrum voor Wiskunde en Informatica,

The Netherlands

### **Doctoral Consortium Chairs**

Paul Groth VU University Amsterdam,

The Netherlands

Natasha Noy Google, USA

### **Proceedings Chair**

Krzysztof Janowicz University of California, Santa Barbara,

USA

### Semantic Web Challenge Chairs

Sean Bechhofer The University of Manchester, UK Andreas Harth Karlsruhe Institute of Technology.

Germany

### Sponsorship Chairs

Roberta Cuel University of Trento, Italy
Jens Lehmann University of Leipzig, Germany
Vincenzo Maltese DISI - University of Trento, Italy

### **Publicity Chairs**

Mauro Dragoni Fondazione Bruno Kessler, Italy

### **Metadata Chairs**

Jie Bao Memect, USA Li Ding Memect, USA

### **Student Coordinators**

Oscar Corcho Universidad Politécnica de Madrid.

Spain

Miriam Fernandez KMI, Open University, UK

### Senior Program Committee – Research

Harith Alani The Open University, UK

Lora Aroyo VU University Amsterdam, The Netherlands

Sören Auer Bonn Universität, Germany Philipp Cimiano Bielefeld University, Germany

Oscar Corcho Universidad Politécnica de Madrid, Spain

Philippe Cudré-Mauroux University of Fribourg, Switzerland

Claudio Gutierrez Chile University, Chile
Jeff Heflin Lehigh University, USA
Ian Horrocks University of Oxford, UK

Lalana Kagal MIT, USA David Karger MIT, USA

Spyros Kotoulas IBM Research, Ireland Diana Maynard University of Sheffield, UK

Natasha Noy Google, USA

Jeff Pan University of Aberdeen, UK Terry Payne University of Liverpool, UK

Marta Sabou MODUL University Vienna, Austria
Uli Sattler The University of Manchester, UK
Steffen Staab University of Koblenz-Landau, Germany
Hideaki Takeda National Institute of Informatics, Japan

### Program Committee - Research

Karl Aberer Paolo Bouquet Sudhir Agarwal Loris Bozzato Faisal Alkhateeb John Breslin

Pramod Anantharam Christopher Brewster

Sofia Angeletou Paul Buitelaar Kemafor Anyanwu Gregoire Burel Marcelo Arenas Andrea Calì Manuel Atencia Diego Calvanese Medha Atre Amparo E. Cano Isabelle Augenstein Iván Cantador Nathalie Aussenac-Gilles Soumen Chakrabarti Jie Bao Pierre-Antoine Champin

Payam Barnaghi Gong Cheng Sean Bechhofer Key-Sun Choi

Klaus Berberich Smitashree Choudhury
Christian Bizer Michael Compton
Roi Blanco Isabel Cruz

Eva Blomqvist Bernardo Cuenca Grau Kalina Bontcheva Claudia D'Amato

### XIV Conference Organization

Mathieu D'Aquin Danica Damljanovic

Stefan Decker Stefan Dietze John Domingue Jérôme Euzenat Lujun Fang Nicola Fanizzi Anna Fensel

Miriam Fernandez Lorenz Fischer Achille Fokoue Enrico Franconi

Fabien Gandon Aldo Gangemi Raúl García-Castro Anna Lisa Gentile

Yolanda Gil Birte Glimm

Jose Manuel Gomez-Perez

Olaf Görlitz Alasdair Gray Tudor Groza Michael Gruninger Christophe Guéret Giancarlo Guizzardi

Peter Haase Armin Haller Harry Halpin

Siegfried Handschuh

Lynda Hardman Manfred Hauswirth

Sandro Hawke Tom Heath Cory Henson Martin Hepp Pascal Hitzler Rinke Hoekstra Aidan Hogan Matthew Horridge Andreas Hotho Geert-Jan Houben

Wei Hu Bo Hu

Eero Hyvönen

Krzysztof Janowicz

Mustafa Jarrar Hanmin Jung Jaap Kamps

Yevgeny Kazakov Maria Keet Hong-Gee Kim Matthias Klusch Jacek Kopecky

Manolis Koubarakis Markus Krötzsch Werner Kuhn

Agnieszka Lawrynowicz

Freddy Lecue

Lei Li
Yuefeng Li
Juanzi Li
Chengkai Li
Vanessa Lopez
Pasquale Lops
David Martin
Robert Meusel
Alessandra Mileo

Iris Miliaraki Riichiro Mizoguchi Dunja Mladenic Luc Moreau Boris Motik Enrico Motta

Mark Musen

Ekawit Nantajeewarawat Axel-Cyrille Ngonga Ngomo

Nadeschda Nikitina Andriy Nikolov Lyndon Nixon Kieron O'Hara Beng Chin Ooi Massimo Paolucci Bijan Parsia Carlos Pedrinaci Silvio Peroni

H. Sofia Pinto Dimitris Plexousakis Valentina Presutti

Guilin Qi

Yves Raimond Ganesh Ramakrishnan Maya Ramanath

Chantal Reynaud Riccardo Rosati

Marco Rospocher Matthew Rowe Sebastian Rudolph Harald Sack Satya Sahoo Hassan Saif

Manuel Salvadores François Scharffe

Ansgar Scherp Stefan Schlobach Daniel Schwabe Juan F. Sequeda Amit Sheth Michael Sintek Milan Stankovic Markus Strohmaier

Rudi Studer Gerd Stumme Jing Sun Vojtěch Svátek Valentina Tamma Kerry Taylor

Krishnaprasad Thirunarayan

Ioan Toma Nicolas Torzec Thanh Tran Raphaël Troncy Giovanni Tummarello Anni-Yasmin Turhan Jacopo Urbani

Victoria Uren

Jacco van Ossenbruggen Maria-Esther Vidal Johanna Völker Haofen Wang Kewen Wang Zhichun Wang Yong Yu Fouad Zablith Qingpeng Zhang Antoine Zimmermann

### Additional Reviewers – Research

Nitish Aggarwal Muhammad Intizar Ali Carlos Buil Aranda Mihael Arcan

Estefanía Serral Asensio

Andrew Bate Wouter Beek Konstantina Bereta

David Berry Nicola Bertolin Leopoldo Bertossi Dimitris Bilidas Stefano Bortoli Adrian Brasoveanu

Volha Bryl

Jean-Paul Calbimonte Diego Calvanese Delroy Cameron Michele Catasta Sam Coppens Julien Corman Jérôme David Jana Eggink

Jae-Hong Eom

Ju Fan

Jean-Philippe Fauconnier Daniel Fleischhacker Andre Freitas Riccardo Frosini

Irini Fundulaki

Stella Giannakopoulou

Jangwon Gim Kalpa Gunaratna Andreas Harth

### XVI Conference Organization

Julia Hoxha Gao Huan Armen Inants

Ernesto Jimenez-Ruiz

Zoi Kaoudi

Patrick Kapahnke Mario Karlovcec Robin Keskisärkkä Taehong Kim

Magnus Knuth

Haridimos Kondylakis

Aljaz Kosmerlj Egor V. Kostylev Adila A. Krisnadhi Sarasi Lalithsena Christoph Lange Sungin Lee Zhixing Li

Juanzi Li Nuno Lopes Yongtao Ma

Andre de Oliveira Melo Albert Meroño-Peñuela

Ankush Meshram Hyun Namgoong Radim Nedbal Hai Nguyen

Quoc Viet Hung Nguyen

Vinh Nguyen Inna Novalija

Andrea Giovanni Nuzzolese

Francesco Osborne Matteo Palmonari Bianca Pereira Jorge Pérez Catia Pesquita Robert Piro Camille Pradel Freddy Priyatna Jav Pujara

Jean-Eudes Ranvier David Ratcliffe Martin Rezk Laurens Rietveld Petar Ristoski Giuseppe Rizzo

Víctor Rodríguez-Doncel

Yannis Roussakis Tong Ruan Brigitte Safar Andreas Schwarte Kunal Sengupta Chao Shao Philipp Singer

Panayiotis Smeros Kostas Stefanidis Andreas Steigmiller Nenad Stojanovic Aynaz Taheri

Vahid Taslimitehrani Veronika Thost Alberto Tonon Trung-Kien Tran Tran Quoc Trung

Jung-Ho Um
Joerg Waitelonis
Simon Walk
Zhigang Wang
Zhe Wang
Guohui Xiao
Nasser Zalmout
Eva Zangerle

Dmitriy Zheleznyakov Zhiqiang Zhuang

### Senior Program Committee – Semantic Web In-Use

Yolanda Gil University of South California, USA

Paul Groth VU University Amsterdam, The Netherlands

Peter Haase fluid Operations, Germany

Siegfried Handschuh Digital Enterprise Research Institute (DERI),

Ireland

Andreas Harth Karlsruhe Institute of Technology, Germany Krzysztof Janowicz University of California, Santa Barbara, USA

Natasha Noy Google, USA

Matthew Rowe Lancaster University, UK
Uli Sattler University of Manchester, UK

### Program Committee - Semantic Web In-Use

Dean Allemang Pavel Klinov
Phil Archer Matthias Klusch
Isabelle Augenstein Christoph Lange
Sean Bechhofer Yuan-Fang Li
Christian Bizer Thorsten Liebig
Sebastian Blohm Antonis Loizou

Dan Brickley Markus Luczak-Roesch

Iván Cantador Pablo Mendes Vinav Chaudhri Hannes Mühleisen Michelle Cheatham Lvndon Nixon Paolo Ciccarese Daniel Oberle Michael Compton Massimo Paolucci Carlos Pedrinaci Oscar Corcho Mathieu D'Aguin Héctor Pérez-Urbina Brian Davis Edoardo Pignotti Mike Dean Yves Raimond Leigh Dodds Marco Rospocher Manuel Salvadores Basil Ell Fabian Floeck Miel Vander Sande

Alasdair Grav Marc Schaaf Tudor Groza Michael Schmidt Giancarlo Guizzardi Oshani Seneviratne Armin Haller Juan F. Sequeda Sebastian Hellmann Evren Sirin Martin Hepp William Smith Milan Stankovic Matthew Horridge Wei Hu Thomas Steiner Prateek Jain Nenad Stojanovic

Anja Jentzsch Pedro Szekely
Benedikt Kämpgen Ramine Tinati

### XVIII Conference Organization

Nicolas Torzec Zhe Wu
Giovanni Tummarello Fouad Zablith
Jacco van Ossenbruggen Amrapali Zaveri
Ruben Verborgh Qingpeng Zhang
Holger Wache Amal Zouaq
Jesse Jiaxin Wang

### Additional Reviewers - Semantic Web In-Use

Ana Sasa Bastinos Gwendolin Wilke Anila Sahar Butt Zhe Wang

Charles Vardeman II

Kewen Wang

### Program Committee – Replication, Benchmark, Data and Software

Sören Auer Siegfried Handschuh
Jie Bao Matthew Horridge
Cosmin Basca Jens Lehmann
Chris Biemann Alejandro Llaves
Philipp Cimiano Axel Polleres
Oscar Corcho Marta Sabou

Philippe Cudré-Mauroux Thomas Scharrenbach Richard Cyganiak Heiner Stuckenschmidt Claudia D'Amato Lamie Taylor

Claudia D'Amato Jamie Taylor
Victor de Boer Jürgen Umbrich
Gianluca Demartini Gerhard Weikum
Jérôme Euzenat Stuart Wrigley

Javier D. Fernández Josiane Xavier Parreira

Alfio Gliozzo Martin Grund

### Additional Reviewers – Replication, Benchmark, Data and Software

Judie AttardMichael LuggenMihaela BorneaChristian MeilickeYu ChenNiels OckeloenSouripriya DasMatthew PerrySarthak DashRoman ProkofyevMichael GlassStephan Radeck-Arneth

Tatiana Lesnikova Eugen Ruppert

Joerg Schoenfisch Claus Stadler Wei Tai Alberto Tonon Joerg Unbehauen

Han Wang Marcin Wylot Seid Muhie Yimam

### Program Committee – Doctoral Consortium

Harith Alani Lalana Kagal
Lora Aroyo Diana Maynard
Abraham Bernstein Enrico Motta
Oscar Corcho Natasha Noy
Philippe Cudré-Mauroux Terry Payne
Fabien Gandon Guus Schreiber
Paul Groth Elena Simperl
Pascal Hitzler

### **Sponsors**

### Student Travel Award Sponsor

Semantic Web Science Association (SWSA) National Science Foundation (NSF)

### **Invited Speakers Sponsor**

Artificial Intelligence Journal

### Semantic Web Challenge Sponsor

Elsevier

#### Platinum

Yahoo! Labs Elsevier

#### Gold

Telefonica Parlance project Google IBM Research Microsoft Research LinkedUp project The Open PHACTS Foundation Systap Semantic Valley

#### Silver

OpenLink software Poolparty fluid Operations News Reader project Synerscope Okkam SICRaS Bpeng Centro Ricerche GPI Expert System Informatica Trentina s.p.a. Institute of Cognitive Sciences and Technologies - CNR

### **Doctoral Consortium Sponsor** iMinds

### Video Sponsor

Videolectures

### Organizing Institution/Foundation

Fondazione Bruno Kessler Semantic Web Science Association



Microsoft Research

































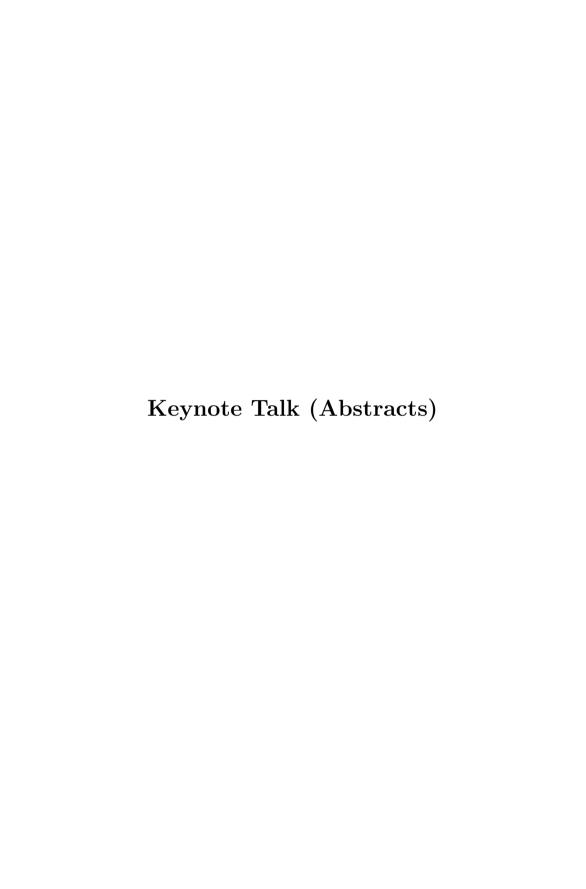












# Web Search - From the Noun to the Verb (Keynote Talk)

Prabhakar Raghavan

Vice President Engineering Google, USA

**Abstract.** This talk examines the evolution of web search experiences over 20 years, and their impact on the underlying architecture. Early web search represented the adaptation of methods from classic Information Retrieval to the Web. Around the turn of this century, the focus shifted to triaging the need behind a query - whether it was Navigational, Informational or Transactional; engines began to customize their experiences depending on the need. The next change arose from the recognition that most queries embodied noun phrases, leading to the construction of knowledge representations from which queries could extract and deliver information regarding the noun in the query. Most recently, three trends represent the next step beyond these "noun engines": (1) "Queryless engines" have begun surfacing information meeting a user's need based on the user's context, without explicit querying; (2) Search engines have actively begun assisting the user's task at hand - the verb underlying the noun query; (3) increasing use of speech recognition is changing the distribution of queries.

### "To Be or to DO?": The Semantics for Smart Cities and Communities (Keynote Talk)

Paolo Traverso

Director Center for Information Technology Fondazione Bruno Kessler, Italy

**Abstract.** The major challenge for so-called smart cities and communities is to provide people with value added services that improve their quality of life. Massive individual and territorial data sets – (open) public and private data, as well as their semantics which allows us to transform data into knowledge about the city and the community, are key enablers to the development of such solutions. Something more however is needed. A "smart" community needs "to do things" in a city, and the people need to act within their own community. For instance, not only do we need to know where we can find a parking spot, which cultural event is happening tonight, or when the next bus will arrive, but we also need to actually pay for parking our car, buy a bus ticket, or reserve a seat in the theater. All these activities (paying, booking, buying, etc.) need semantics in the same way as data does, and such a semantics should describe all the steps needed to perform such activities. Moreover, such a semantics should allow us to define and deploy solutions that are general and abstract enough to be "portable" across the details of the different ways in which activities can be implemented, e.g., by different providers, or for different customers, or for different cities. At the same time, in order to actually "do things", we need a semantics that links general and abstract activities to the possibly different and specific ICT systems that implement them. In my talk, I will present some of the main problems for realizing the concept of smart city and community, and the need for semantics for both understanding data and "doing things". I will discuss some alternative approaches, some lessons learned from applications we have been working with in this field, and the still many related open research challenges.

# Semantic Challenges in Getting Work Done (Keynote Talk)

#### Yolanda Gil

Associate Director Information Sciences Institute and Department of Computer Science University of Southern California, USA

**Abstract.** In the new millennium, work involves an increasing amount of tasks that are knowledge-rich and collaborative. We are investigating how semantics can help on both fronts. Our focus is scientific work, in particular data analysis, where tremendous potential resides in combining the knowledge and resources of a highly fragmented science community. We capture task knowledge in semantic workflows, and use skeletal plan refinement algorithms to assist users when they specify high-level tasks. But the formulation of workflows is in itself a collaborative activity, a kind of meta-workflow composed of tasks such as finding the data needed or designing a new algorithm to handle the data available. We are investigating "organic data science", a new approach to collaboration that allows scientists to formulate and resolve scientific tasks through an open framework that facilitates ad-hoc participation. With a design based on social computing principles, our approach makes scientific processes transparent and incorporates semantic representations of tasks and their properties. The semantic challenges involved in this work are numerous and have great potential to transform the Web to help us do work in more productive and unanticipated ways.

# The Semantic Web in an Age of Open Data (Keynote Talk)

Nigel Shadbolt

Professor of Artificial Intelligence
The University of Southampton
and
Chairman of the Open Data Institute

Abstract. The last five years have seen increasing amounts of open data being published on the Web. In particular, governments have made data available across a wide range of sectors: spending, crime and justice, education, health, transport, geospatial, environmental and much more. The data has been published in a variety of formats and has been reused with varying degrees of success. Commercial organisations have begun to exploit this resource and in some cases elected to release their own open data. Only a relatively small amount of the data published has been linked data. However, the methods and techniques of the semantic web could significantly enhance the value and utility of open data. What are the obstacles and challenges that prevent the routine publication of these resources as semantically enriched open data? What can be done to improve the situation? Where are the examples of the successful publication and exploitation of semantically enriched content? What lessons should we draw for the future?

### Table of Contents - Part II

Ontology Alignment and Modularization	
Detecting and Correcting Conservativity Principle Violations in Ontology-to-Ontology Mappings	1
Towards Annotating Potential Incoherences in BioPortal Mappings  Daniel Faria, Ernesto Jiménez-Ruiz, Catia Pesquita,  Emanuel Santos, and Francisco M. Couto	17
Conference v2.0: An Uncertain Version of the OAEI Conference Benchmark	33
Fast Modularisation and Atomic Decomposition of Ontologies Using Axiom Dependency Hypergraphs	49
A Study on the Atomic Decomposition of Ontologies	65
Querying	
Structural Properties as Proxy for Semantic Relevance in RDF Graph Sampling	81
Holistic and Compact Selectivity Estimation for Hybrid Queries over RDF Graphs	97
Querying Factorized Probabilistic Triple Databases	114
Ontology Search: An Empirical Evaluation	130

### Reasoning

Pushing the Boundaries of Tractable Ontology Reasoning	148
Effective Computation of Maximal Sound Approximations of Description Logic Ontologies	164
Abstraction Refinement for Ontology Materialization	180
Goal-Directed Tracing of Inferences in EL Ontologies Yevgeny Kazakov and Pavel Klinov	196
Semantic Infrastructures and Streams	
Semantic Web Application Development with LITEQ  Martin Leinberger, Stefan Scheglmann, Ralf Lämmel, Steffen Staab, Matthias Thimm, and Evelyne Viegas	212
Semantic-Based Process Analysis	228
Efficient RDF Interchange (ERI) Format for RDF Data Streams Javier D. Fernández, Alejandro Llaves, and Oscar Corcho	244
Sensors	
Knowledge-Driven Activity Recognition and Segmentation Using Context Connections	260
A Use Case in Semantic Modelling and Ranking for the Sensor Web  Liliana Cabral, Michael Compton, and Heiko Müller	276
Semantic Traffic Diagnosis with STAR-CITY: Architecture and Lessons Learned from Deployment in Dublin, Bologna, Miami and Rio  Freddy Lécué, Robert Tucker, Simone Tallevi-Diotallevi, Rahul Nair, Yiannis Gkoufas, Giuseppe Liguori, Mauro Borioni, Alexandre Rademaker, and Luciano Barbosa	292

Adapting Semantic Sensor Networks for Smart Building Diagnosis Joern Ploennigs, Anika Schumann, and Freddy Lécué	308
Social Media	
Semantic Patterns for Sentiment Analysis of Twitter	324
Stretching the Life of Twitter Classifiers with Time-Stamped Semantic Graphs	341
Linked Open Data Driven Game Generation	358
SPARQL Extensions	
On the Semantics of SPARQL Queries with Optional Matching under Entailment Regimes	374
Strategies for Executing Federated Queries in SPARQL1.1	390
Toward the Web of Functions: Interoperable Higher-Order Functions in SPARQL	406
User Interaction and Personalization	
Explass: Exploring Associations between Entities via Top- $K$ Ontological Patterns and Facets	422
Expressive and Scalable Query-Based Faceted Search over SPARQL Endpoints	438
Querying Heterogeneous Personal Information on the Go	454
The Web Browser Personalization with the Client Side Triplestore	470

### XXX Table of Contents – Part II

CrowdTruth: Machine-Human Computation Framework for Harnessing Disagreement in Gathering Annotated Data	486
Doctoral Consortium	
Joint Information Extraction from the Web Using Linked Data	505
Entity Linking with Multiple Knowledge Bases: An Ontology Modularization Approach	513
Populating Entity Name Systems for Big Data Integration	521
Semantic Complex Event Processing for Decision Support	529
Enriching Ontologies with Encyclopedic Background Knowledge for Document Indexing	537
Detecting and Correcting Conservativity Principle Violations in Ontology Mappings	545
Author Index	553

### Table of Contents – Part I

Data Integration and Link Discovery	
CAMO: Integration of Linked Open Data for Multimedia Metadata Enrichment	1
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	17
SAKey: Scalable Almost Key Discovery in RDF Data  Danai Symeonidou, Vincent Armant, Nathalie Pernelle, and Fatiha Saïs	33
Introducing Wikidata to the Linked Data Web	50
Web-Scale Extension of RDF Knowledge Bases from Templated Websites	66
Data Integration and Link Discovery in Life Sciences	
EPCIS Event-Based Traceability in Pharmaceutical Supply Chains via Automated Generation of Linked Pedigrees	82
Scientific Lenses to Support Multiple Views over Linked Chemistry  Data	98

Linked Biomedical Dataspace: Lessons Learned Integrating Data for Drug Discovery	114
Ali Hasnain, Maulik R. Kamdar, Panagiotis Hasapis, Dimitris Zeginis, Claude N. Warren Jr., Helena F. Deus, Dimitrios Ntalaperas, Konstantinos Tarabanis, Muntazir Mehdi, and Stefan Decker	
Drug-Target Interaction Prediction Using Semantic Similarity and Edge Partitioning	131
Large-Scale RDF Processing and Dataset Availability	
SYRql: A Dataflow Language for Large Scale Processing of RDF	
Data	147
Sempala: Interactive SPARQL Query Processing on Hadoop	164
Querying Datasets on the Web with High Availability	180
Diversified Stress Testing of RDF Data Management Systems	197
Linked Data	
LOD Laundromat: A Uniform Way of Publishing Other People's Dirty Data	213
Dutch Ships and Sailors Linked Data	229
Adoption of the Linked Data Best Practices in Different Topical Domains	245
Analyzing Schema.org	261

The WebDataCommons Microdata, RDFa and Microformat Dataset	
Series	277
On Publishing Chinese Linked Open Schema	293
Linked Data and Data Quality	
Discovery and Visual Analysis of Linked Data for Humans	309
Col-Graph: Towards Writable and Scalable Linked Open Data Luis-Daniel Ibáñez, Hala Skaf-Molli, Pascal Molli, and Olivier Corby	325
Transferring Semantic Categories with Vertex Kernels:  Recommendations with SemanticSVD++	341
Detecting Errors in Numerical Linked Data Using Cross-Checked Outlier Detection	357
Noisy Type Assertion Detection in Semantic Datasets	373
Mobile Reasoning and SPARQL Updates	
A Cross-Platform Benchmark Framework for Mobile Semantic Web Reasoning Engines	389
A Power Consumption Benchmark for Reasoners on Mobile Devices Evan W. Patton and Deborah L. McGuinness	409
Dynamic Provenance for SPARQL Updates	425
Updating RDFS ABoxes and TBoxes in SPARQL	441

## Natural Language Processing and Information Extraction

AGDISTIS - Graph-Based Disambiguation of Named Entities Using Linked Data	457
Ricardo Usbeck, Axel-Cyrille Ngonga Ngomo, Michael Röder, Daniel Gerber, Sandro Athaide Coelho, Sören Auer, and Andreas Both	
M-ATOLL: A Framework for the Lexicalization of Ontologies in Multiple Languages	472
Towards Efficient and Effective Semantic Table Interpretation Ziqi Zhang	487
Semano: Semantic Annotation Framework for Natural Language	
Resources	503
Ensemble Learning for Named Entity Recognition	519
OBDA and Query Rewriting	
OBDA: Query Rewriting or Materialization? In Practice, Both! Juan F. Sequeda, Marcelo Arenas, and Daniel P. Miranker	535
Answering SPARQL Queries over Databases under OWL 2 QL	
Entailment Regime	552
kyrie2: Query Rewriting under Extensional Constraints in $\mathcal{ELHIO}$ Jose Mora, Riccardo Rosati, and Oscar Corcho	568
Schema-Agnostic Query Rewriting in SPARQL 1.1	584
How Semantic Technologies Can Enhance Data Access at Siemens	
Energy	601
Author Index	621