

Preface

Supporting new forms of scholarly data publication and analysis is a crucial task for researchers, publishers, and companies working in innovative solutions for scholarly communication. Currently, most research papers are published as portable document format (PDF) and/or poorly annotated, only with simple metadata provided as a set of keywords or topic categories, if at all, which makes it hard to extract information from the full text. In addition, not always are other research outcomes, such as research data or software, made available nor do they include rich metadata. This hinders the discoverability, reuse, and reproducibility of research data and findings. A more structured and semantically rich representation of the research outcomes could bring significant advantages to various areas: linking more effectively research and industrial efforts, supporting researchers' work, fostering cross-pollination of ideas and methods across different areas, driving research policies, and acting as a source of information for a variety of applications. The program of the Semantics, Analytics, Visualization: Enhancing Scholarly Data (SAVE-SD) 2016 Workshop highlighted topics in these areas.

The first edition of the SAVE-SD workshop took place on May 19, 2015, and was co-located with the 24th International World Wide Web Conference in Florence, Italy. After the success of the first edition, the second edition, presented in this volume, took place on April 11, 2016, in Montreal, Canada, co-located with the 25th International World Wide Web Conference.

SAVE-SD 2016 opened with the keynote speech by Alex Wade, Director of Scholarly Communications at Microsoft Research, whose work focuses on Microsoft Academic and involves aspects of knowledge acquisition, knowledge representation, intentionality, dialog systems, semantic search, and intelligent agents. His talk was entitled "The Microsoft Academic Graph: New Applications and Research Opportunities" and described the new entity graph of research publications, authors, venues, organizations, and topics that is developed by Microsoft Research and drives new features in Bing, Cortana, and Microsoft Academic. We are grateful to Alex Wade for his inspiring talk.

The scientific program of SAVE-SD 2016 comprised 11 papers: five full papers, selected out of 11 submissions, which corresponds to an acceptance rate of 45%; two position papers and six poster or demo papers, selected out of five submissions, plus three of the full papers that instead were accepted as one position, one demo, and one poster paper. The workshop received a total of 16 submissions from authors of 15 countries in three continents (Europe, Asia, Americas) and was attended by about 50 people.

The topics in this edition demonstrate current research on semantic publishing and cover the extraction of semantic information from research papers or pre-existing datasets and the use of semantic techniques for characterizing citations and analyzing research topics and trends. SAVE-SD provided a forum for researchers, publishers, and companies interested in enhancing scholarly data to come together and discuss challenges and innovative solutions.

SAVE-SD accepts submissions not just in the traditional PDF format, but also in HTML, encouraging authors to provide semantically rich papers themselves. In order to support authors willing to submit in HTML, SAVE-SD encouraged the use of the RASH format (<https://github.com/essepuntato/rash>) for submissions. RASH stands for Research Articles in Simplified HTML (RASH) and can be produced from Open Office documents, Microsoft Word documents, and other formats.

SAVE-SD 2016 awarded two prizes: one for best paper and another for best RASH paper. The latter was sponsored by Springer Nature.

The criteria for selecting the best paper award considered the reviewers' scores and selected the paper with the best score. The best paper award was given to: "*Detection of Embryonic Research Topics by Analysing Semantic Topic Networks*" by Angelo Salatino and Enrico Motta.

The best RASH paper award is given to the paper that makes best use of the RASH format. This is chosen by an automatic score system that rates all the RASH submissions considering:

1. The quality of the markup (i.e., considering the number of errors in the document compared with the RASH grammar)
2. The quality of HTML (i.e., how many errors the document has compared with HTML5)
3. The number of Resource Description Framework (RDF) statements defined
4. The number of RDF links to Linked Open Data (LOD) datasets

The best RASH paper of SAVE-SD 2016 was awarded to: "*Citation Functions for Knowledge Export—A Question of Relevance, or, Can CiTO Do the Trick?*" by Joakim Philipson.

As SAVE-SD aims to address the gap between the theoretical and practical aspects of scholarly data, the review process ought to consider both perspectives. Thus, SAVE-SD has three different Program Committees (PCs):

- An Industrial PC, which mainly evaluates the submissions from an industrial perspective by assessing how much the theories or applications described in the papers (may) influence (positively or negatively) the publishing and technological domain and whether they could be concretely adopted by publishers and scholarly data providers
- An Academic PC, which evaluates the papers mainly from an academic perspective by assessing the quality of the research described in such papers
- A Senior PC, whose members act as meta-reviewers and have the crucial role of balancing the scores provided by the reviews from the other two PCs

We are very grateful to all the members of the three PCs, listed herein, for their high-quality reviews and constructive feedback, which improved significantly the quality of the papers contained in these proceedings.

Last but certainly not least, we want to thank our sponsors:

- Springer Nature (<http://www.springernature.com/>), who provided a 150-euro voucher to buy Springer Nature's products for the best RASH paper award

- Pensoft Publishers (<http://www.pensoft.net/>), who hosted a free-of-charge special collection for selected position/poster/demo papers in the *Research Idea and Outcomes* (RIO) journal
- *GigaScience* journal (<http://www.gigasciencejournal.com/>), which provided cool “Data is coming” t-shirts for the workshops attendees

March 2017

Alejandra González-Beltrán
Francesco Osborne
Silvio Peroni

Semantics, Analytics, Visualization. Enhancing Scholarly
Data

Second International Workshop, SAVE-SD 2016,
Montreal, QC, Canada, April 11, 2016, Revised Selected
Papers

Gonzalez-Beltran, A.; Osborne, F.; Peroni, S. (Eds.)

2016, XII, 153 p. 39 illus., Softcover

ISBN: 978-3-319-53636-1