

Preface of the Workshop Organizers

The First International Workshop on “Challenges and Experiences from Data Integration to Knowledge Graphs” (DI2KG 2019) hosted cross-disciplinary contributions on data integration and knowledge graph construction. Such complex processes have been studied by different communities, including data management, machine learning, statistics, data science, natural language processing and information retrieval, typically in isolation. As holistic solutions are emerging, we claim the need for a more cross-disciplinary community that pushes research toward the creation of the next generation of methods. In order to stimulate advances in this direction we have also hosted the DI2KG challenge, a set of fundamental integration tasks leading to the construction of a knowledge graph of product specifications extracted from the Web with their own manually-checked ground truth. The workshop encouraged multi-disciplinary submissions on a broad spectrum of topics: source selection and discovery, data and information extraction, data cleaning and fusion, schema extraction and alignment, algorithmic and statistical techniques for entity resolution, machine learning methods for data integration, benchmarking and performance measurement, knowledge graph augmentation and embedding. The submitted papers underwent a zero-blind review process and were evaluated on the basis of originality, relevance, quality, and technical contribution. Authors of accepted papers took into account the feedback received at the workshop for preparing their camera ready. The workshop attendance included more than 60 registered delegates, both from industry and academia, with geographical, gender and seniority diversity. The workshop program featured 2 invited talks (AnHai Doan and Andrew McCallum), a panel with 4 speakers (AnHai Doan, Renée Miller, Luna Dong, and Andrew McCallum) moderated by Paolo Merialdo and presentation of 7 papers, including the benchmark challenge winner. The detailed workshop program is available on-line at <http://di2kg.dia.uniroma3.it>. We believe that future editions of DI2KG will foster new research on holistic data integration methods leading to the construction of knowledge graphs, while building a community that can contribute to the evolution of the benchmark for the DI2KG challenge.

We are indebted to the members of the program committee for their meticulous work in the selection process: Denilson Barbosa, Valter Crescenzi, Xin Luna Dong, Laura Haas, Colin Lockard, Paolo Merialdo, Renée Miller, Mourad Ouzzani, and Themis Palpanas. Finally, we would like to thank the authors who submitted their work to this workshop and all the challenge and workshop participants who contributed to the success of the event.

November 2019

Donatella Firmani
Valter Crescenzi
Andrea De Angelis
Xin Luna Dong
Maurizio Mazzei
Paolo Merialdo
Divesh Srivastava