

# PROFILES '17

## 4th International Workshop on Dataset PROFiling & fEderated Search for Web Data

### Preface

This volume of proceedings gathers papers from the 4th International Workshop on Dataset PROFiling and fEderated Search for Web Data (PROFILES '17), co-located with the 16th International Semantic Web Conference (ISWC 2017) held on the 22nd of October, 2017 in Vienna, Austria.

The Web of Data, including Linked Data and knowledge graphs, has seen tremendous growth recently. In addition, new forms of structured data have emerged in the form of Web markup, such as schema.org, and entity-centric data in Web tables. Considering these rich, heterogeneous and evolving data sources which cover a wide variety of domains, exploitation of Web Data becomes increasingly important in the context of various applications, including federated search, entity linking, question answering and fact verification. These applications require reliable information on dataset characteristics, including general metadata, quality features, statistical information, dynamics, licensing and provenance. The lack of a thorough understanding of the nature, scope and characteristics of data from particular sources limits their take-up and reuse, such that applications are often limited and focused on well-known reference datasets. The PROFILES workshop series aims at gathering approaches to analyse, describe and discover data sources – including but not limited to SPARQL endpoints – as a facilitator for applications and tasks such as query distribution, semantic search, entity retrieval and recommendation. PROFILES offers a highly interactive forum for researchers and practitioners bringing together experts in the fields of Semantic Web, Web Data, Semantic Search, Databases, NLP, IR and various application domains.

The PROFILES '17 workshop is a continuation of the workshop series successfully started as PROFILES '14 - PROFILES '16 at ESWC 2014 - 2016. These workshops aim to gather innovative query and search approaches for large-scale, distributed and heterogeneous linked datasets in line with dedicated approaches to analyse, describe and discover endpoints, as an inherent task of query distribution and dataset recommendation. This year we have accepted four full papers and two short papers.

We would like to take this opportunity to sincerely thank the authors for their invaluable and inspiring contributions to the workshops. Our sincere thanks are due

to the program committee members for reviewing the submissions and ensuring the high quality of our workshop program.

We are also very grateful to the organisers of the ISWC 2017 conference, and in particular to the Workshops and Tutorials Chairs, Aidan Hogan and Valentina Presutti, for their support in the workshop organisation.

Sincere appreciation for his time and expertise goes to our keynote speaker Prof. Dr. Heiko Paulheim from the Data and Web Science Group, University of Mannheim, Germany for his talk entitled “Towards Knowledge Graph Profiling”.

September 2017

Elena Demidova  
Stefan Dietze  
Julian Szymański  
John Breslin

(The PROFILES '17 workshop was co-organised by members of the KEYSTONE COST Action IC1302 on semantic KEYword-based Search on sTructured data sOurcEs and supported by the Data4UrbanMobility project funded by the Federal Ministry of Education and Research, Germany (BMBF, 02K15A040).)

## Organisation

### Organising Committee:

Elena Demidova, L3S Research Center, Leibniz Universität Hannover, Germany  
Stefan Dietze, L3S Research Center, Leibniz Universität Hannover, Germany  
Julian Szymański, Gdańsk University of Technology, Poland  
John Breslin, National University of Ireland Galway, Ireland

### Program Committee:

Axel Ngonga Ngomo, InfAI, University of Leipzig  
Muhammad Saleem, InfAI, University of Leipzig  
Andreas Nürnberger, Otto-von-Guericke University of Magdeburg  
Alessandro Adamou, Knowledge Media Institute, The Open University  
Enrico Daga, The Open University  
Dmitry Mouromtsev, ITMO University, St. Petersburg  
Bernardo Pereira Nunes, PUC-Rio  
Thanassis Tiropanis University of Southampton  
Konstantin Todorov LIRMM / University of Montpellier  
Tom Heath, Arup  
Charlie Abela, University of Malta  
Marco Antonio Casanova, PUC-Rio  
Raquel Trillo-Lado, Universidad de Zaragoza  
Simon Gottschalk, Leibniz Universität Hannover  
Ran Yu, Leibniz Universität Hannover



