



# 43<sup>rd</sup> International Conference on **Very Large Data Bases**

Peter Christen, Bettina Kemme, Erhard Rahm (eds.):

## **Proceedings of the VLDB 2017 PhD Workshop**

August 28, 2017

Munich, Germany

# Preface

This volume contains the proceedings of the VLDB 2017 PhD Workshop, co-located with the 43rd International Conference on Very Large Database (VLDB 2017) and held on August 28, 2016, in Munich, Germany.

The established VLDB PhD Workshop series is a unique opportunity for graduate students to present and discuss their research in the context of a premier international conference. The workshop provides a forum that facilitates interactions among PhD students and stimulates feedback from more experienced researchers—not only in terms of diligent paper reviews but also at the conference site.

Following the call for papers, 27 papers were submitted for presentation at the workshop. Of these, the program committee carefully selected a total of 14 papers, with 10 selected for long and 4 for short presentations, which are contained in this volume.

## Long presentation papers:

- Efficient Migration of Very Large Distributed State for Scalable Stream Processing.  
Bonaventura Del Monte, DFKI Berlin, Germany
- Spatio-temporal Locality in Hash Tables.  
Matthew A Pugh, University of Edinburgh, UK
- Graph Pattern Mining for Business Decision Support.  
André Petermann, University of Leipzig, Germany
- Efficient Clustering for Large-Scale, Sparse, Discrete Data with Low Fundamental Resolution.  
Veronika Strnadova-Neeley, UC Santa Barbara, USA
- Query Processing Based on Compressed Intermediates.  
Patrick Damme, Technical University Dresden, Germany
- A Hardware-Oblivious Optimizer for Data Stream Processing.  
Constantin Pohl, Technical University Ilmenau, Germany
- Generalizing Matching Knowledge using Active Learning.  
Anna Primpeli, University of Mannheim, Germany
- Comparing entities in RDF graphs.  
Alina Petrova, University of Oxford, UK
- Scalable Linkage across Location Enhanced Services.  
Fuat Basik, Bilkent University, Turkey
- Distributed Similarity Joins on Big Textual Data: Toward a Robust Cost-Based Framework.  
Fabian Fier, Humboldt University Berlin, Germany

## Short presentation papers:

- Facilitating User Interaction With Data.  
Zainab Zolaktaf, University of British Columbia Vancouver, Canada
- Processing Moving Object Data Streams with Data Stream Management Systems.  
Tobias L Brandt, University of Oldenburg, Germany
- Symmetric and Asymmetric Aggregate Function in Massively Parallel Computing.  
Chao Zhang, Blaise Pascale University Aubière, France
- Practical Verification of Hierarchical Artifact Systems.  
Yuliang Li, UC San Diego, USA

The workshop program also featured a keynote talk by Prof. Dr. Anastasia Ailamaki, (EPFL, Switzerland) who talked about the past and future of database transaction processing engines. The workshop ended with a panel consisting of leading database experts including Michael Carey (University of California, Irvine, USA), Wolfgang Lehner (Technische Universität Dresden, Germany), and Themis Palpanas (Paris Descartes University, France) who discussed the “life with a PhD” and provided valuable tips and suggestions to the PhDs who attended the workshop.

We like to thank all authors who submitted papers to the VLDB 2017 PhD workshop, to Anastasia for her keynote presentation, and to the panel members for interesting discussions.

Our sincere thanks also go to all members of the program committee for their insightful and valuable reviews. All of you have helped to make the VLDB 2017 PhD Workshop a successful event.

Peter Christen (The Australian National University, Australia)

Bettina Kemme (McGill, Montreal, Canada)

Erhard Rahm (Univ. of Leipzig, Germany)

PC Co-Chairs

August 2017

## PhD Workshop Program Committee

Angela Bonifati (University of Lyon)

Stéphane Bressan (National University of Singapore)

George Fletcher (TU Eindhoven)

Lukasz Golab (University of Waterloo)

Carson Leung (University of Manitoba)

Sebastian Link (University of Auckland)

Rachel Pottinger (University of British Columbia)

Louisa Raschid (University of Maryland)

Donatello Santoro (Università della Basilicata)

Timos Sellis (Swinburne University of Technology)

Andreas Thor (HFT Leipzig)

Wei Wang (University of New South Wales)

Qing Wang (The Australian National University)