# TABLE OF CONTENTS

## Front Matter
- Copyright Notice .................................................................................................................. ii
- Table of Contents .................................................................................................................. iii
- VLDB 2013 Organization and Review Board ........................................................................ iv

## Letters
- Letter from the Editors-in-Chief ................................................................. *Michael Böhlen, Christoph Koch* xii

## Research Papers
- Probabilistic Query Rewriting for Efficient and Effective Keyword Search on Graph Data........ 1642
  .............................................................................................................................. *Lei Zhang, Thanh Tran, Achim Rettinger*
- QuEval: Beyond high-dimensional indexing a la carte ....................................................... 1654
  .............................................................................................................................. *Martin Schäler, Alexander Grebhahn, Reimar Schröter,*
  .............................................................................................................................. *Sandro Schulze, Veit Köppen, Gunter Saake*
- Discovering Longest-lasting Correlation in Sequence Databases ........................................... 1666
  .............................................................................................................................. *Yuhong Li, Leong Hou U, Man Lung Yiu, Zhiguo Gong*
- PREDIcT: Towards Predicting the Runtime of Large Scale Iterative Analytics ...................... 1678
  .............................................................................................................................. *Adrian Daniel Popescu, Andrey Balmin, Vuk Ercegovac, Anastasia Ailamaki*
- On the Embeddability of Random Walk Distances ............................................................... 1690
  .............................................................................................................................. *Xiaohan Zhao, Adelbert Chang, Atish Das Sarma, Haitao Zheng, Ben Y. Zhao*
- Instant Loading for Main Memory Databases ....................................................................... 1702
  .............................................................................................................................. *Tobias Mühlbauer, Wolf Rödiger, Robert Seilbeck,*
  .............................................................................................................................. *Angelika Reiser, Alfons Kemper, Thomas Neumann*
- Adaptive Range Filters for Cold Data: Avoiding Trips to Siberia ........................................... 1714
  .............................................................................................................................. *Karolina Alexiou, Donald Kossmann, Per-Åke Larson*
- Scalable Progressive Analytics on Big Data in the Cloud ....................................................... 1726
  .............................................................................................................................. *Badrish Chandramouli, Jonathan Goldstein, Abdul Quamar*
- Scalable XML Query Processing using Parallel Pushdown Transducers .................................. 1738
  .............................................................................................................................. *Peter Ogden, David Thomas, Peter Pietzuch*
- Understanding Insights into the Basic Structure and Essential Issues of Table Placement Methods in Clusters ...................................................................................... 1750
  .............................................................................................................................. *Yin Huai, Siyuan Ma, Rubao Lee, Owen O’Malley, Xiaodong Zhang*
- A Probabilistic Optimization Framework for the Empty-Answer Problem ............................. 1762
  .............................................................................................................................. *Davide Mottin, Alice Marascu, Senjuti Basu Roy,*
  .............................................................................................................................. *Gautam Das, Themis Palpanas, Yannis Velegrakis*
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summarizing Answer Graphs Induced by Keyword Queries</td>
<td>1774</td>
</tr>
<tr>
<td>Supporting Keyword Search in Product Database: A Probabilistic Approach</td>
<td>1786</td>
</tr>
<tr>
<td>A Sampling Algebra for Aggregate Estimation</td>
<td>1798</td>
</tr>
<tr>
<td>A Temporal-Probabilistic Database Model for Information Extraction</td>
<td>1810</td>
</tr>
<tr>
<td>Counter Strike: Generic Top-Down Join Enumeration for Hypergraphs</td>
<td>1822</td>
</tr>
<tr>
<td>Efficient Bulk Updates on Multiversion B-trees</td>
<td>1834</td>
</tr>
<tr>
<td>Query-Driven Approach to Entity Resolution</td>
<td>1846</td>
</tr>
<tr>
<td>Expressiveness and Complexity of Order Dependencies</td>
<td>1858</td>
</tr>
<tr>
<td>Counting and Sampling Triangles from a Graph Stream</td>
<td>1870</td>
</tr>
<tr>
<td>An Experimental Analysis of Iterated Spatial Joins in Main Memory</td>
<td>1882</td>
</tr>
<tr>
<td>Scaling Queries over Big RDF Graphs with Semantic Hash Partitioning</td>
<td>1894</td>
</tr>
<tr>
<td>Distributed SociaLite: A Datalog-Based Language for Large-Scale Graph Analysis</td>
<td>1906</td>
</tr>
<tr>
<td>Horton+: A Distributed System for Processing Declarative Reachability Queries over Partitioned Graphs</td>
<td>1918</td>
</tr>
<tr>
<td>Streaming Similarity Search over one Billion Tweets using Parallel Locality-Sensitive Hashing</td>
<td>1930</td>
</tr>
<tr>
<td>Anti-Caching: A New Approach to Database Management System Architecture</td>
<td>1942</td>
</tr>
<tr>
<td>Understanding Hierarchical Methods for Differentially Private Histograms</td>
<td>1954</td>
</tr>
</tbody>
</table>
Towards Social Data Platform: Automatic Topic-focused Monitor for Twitter Stream

Rui Li, Shengjie Wang, Kevin Chen-Chuan Chang

Simple, Fast, and Scalable Reachability Oracle

Ruoming Jin, Guan Wang

Aggregation and Ordering in Factorised Databases

Nurzhan Bakibayev, Tomas Kocisky, Dan Olteanu, Jakub Zavodny

Parallel Computation of Skyline and Reverse Skyline Queries Using MapReduce

Yoonjae Park, Jun-Ki Min, Kyuseok Shim

Fast Iterative Graph Computation with Block Updates

Wenlei Xie, Guozhang Wang, David Bindel, Alan Demers, Johannes Gehrke
VLDB 2013 ORGANIZATION AND REVIEW BOARD

General Chairs
Themis Palpanas, University of Trento
Yannis Velegrakis, University of Trento

Program Chairs
Michael Böhlen, University of Zurich
Christoph Koch, EPFL

Advisory Board
Paolo Atzeni, Universita Roma Tre
Stefano Ceri, Politecnico di Milano
John Mylopoulos, University of Trento

Award Committee
Surajit Chaudhuri, Microsoft (Chair)
Mike Carey, University of California, Irvine
Susan Davidson, University of Pennsylvania
Alon Halevy, Google
Sunita Sarawagi, IIT Bombay

Associate Editors
Ada Wai-Chee Fu, Chinese University of Hong Kong
Alon Halevy, Google
Ashraf Aboulnaga, University of Waterloo
Chee-Yong Chan, National University of Singapore
Dan Olteanu, Oxford University
Divesh Srivastava, AT&T Labs
Jayant Haritsa, Indian Institute of Science Bangalore
Jens Teubner, ETH Zurich
Johannes Gehrke, Cornell University
Nikos Mamoulis, University of Hong Kong
Sihem Amer-Yahia, Qatar Computing Research Institute
Stefan Manegold, CWI
Thomas Neumann, Technische Universität München
Yanlei Diao, University of Massachusetts Amherst
Experiments and Analysis Track Associate Editor
Stefan Manegold, CWI

Industrial and Applications Track Associate Editors
Min Wang, Google Research
Cong Yu, Google Research

Demonstration Chairs
Jun Yang, Duke University
Dimitrios Gunopulos, University of Athens
Letizia Tanca, Politecnico di Milano

Reproducibility Chairs
Philippe Bonnet, IT University of Copenhagen
Juliana Freire, New York University
Dennis Shasha, New York University

Research Track Review Board
Karl Aberer, EPFL, Switzerland
Foto Afrati, NTU Athens
Charu Aggarwal, IBM T. J. Watson Research Center
Yanif Ahmad, JHU
Jose-Luis Ambite, University of Southern California
Walid Aref, Purdue University
Magdalena Balazinska, University of Washington
Srikanta Bedathur, IIIT Delhi
Peter Boncz, CWI
Nico Bruno, Microsoft
Randal Burns, JHU
Andrea Cali, University of London, Birkbeck College
Carlos Castillo, Yahoo!
Gang Chen, Zhejiang University
Lei Chen, Hong Kong University of Science and Technology
Shimin Chen, HP Labs China
James Cheng, CUHK
Reynold Cheng, University of Hong Kong
Gao Cong, Nayang Technological University

Brian Cooper, Google
Bin Cui, Peking University
Carlo Curino, MIT
Sudipto Das, Microsoft Research
Anish Das Sarma, Google Research
Atish Das Sarma, eBay Research Labs
Antonios Deligiannakis, Technical University of Crete
Amol Deshpande, University of Maryland
Xin Luna Dong, AT&T Labs-Research
Sameh Elnikety, Microsoft Research
Mohamed Eltabakh, Worcester Polytechnic Institute
Alan Fekete, University of Sydney
Hakan Ferhatosmanoglu, Bilkent University
Alvaro Fernandez, U. of Manchester
Juliana Freire, New York University
Benjamin C. M. Fung, Concordia University
Fabien Gandon, INRIA
Minos Garofalakis, Technical University of Crete, Greece
Bu̇ra Gedik, Bilkent University
Bordawekar Rajesh, IBM T.J. Watson
Vibhor Rastogi, Yahoo
Christopher Re, University of Wisconsin, Madison
Matthias Renz, Ludwig-Maximilians University Munich, Germany
Marie-Christine Rousset, IMAG
Sourav S. Bhowmick, Nanyang Technological University
Dimitris Sacharidis, IMIS Athena, Greece
Kenneth Salem, University of Waterloo
Maria Sapino, University of Torino
Monica Scannapieco, Istat
Bernhard Seeger, Philipps-Universität Marburg
Pierre Senellart, Télécom ParisTech
Cyrs Shahabi, USC
Lidan Shou, Zhejiang University
Adam Silberstein, Trifacta
Radu Sion, Stony Brook University
Yannis Sismanis, IBM, USA
Mohamed Soliman, University of Waterloo
Julia Stoyanovich, Drexel University and Skoltech
Yufei Tao, Chinese University of Hong Kong
Sandeep Tata, IBM Research
Nesime Tatbul, ETH Zurich
Evimaria Terzi, University of Boston
Martin Theobald, Max Planck Institute, Germany
Anthony Tung, National University of Singapore
Kostas Tzoumas, Technical University of Berlin
Sergei Vassilvitskii, Google
Stratis D. Viglas, University of Edinburgh
Ke Wang, Simon Fraser University
Ingmar Weber, Yahoo!
Raymond Chi-Wing Wong, Hong Kong University of Science and Technology
Xiaokui Xiao, NTU
Dong Xin, Google
Xifeng Yan, University of Santa Barbara
Jiong Yang, Case Western Reserve University
Ke Yi, Hong Kong University of Science and Technology
Man Lung Yu, Hong Kong Polytechnic University
Cong Yu, Google Research
Ge Yu, Northeastern University, China
Jeffrey Yu, Chinese University of Hong Kong
Wenjie Zhang, UNSW Australia
Baihua Zheng, Singapore Management University
Aoying Zhou, East China Normal University
Xiaofang Zhou, University of Queensland

Demonstration Program Committee

Anastasia Ailamaki, EPFL
Sihem Amer-Yahia, Qatar Computing Research Institute
Leopoldo Bertossi, University of Carleton
Francois Bry, University of Munich
Chee-Yong Chan, National University of Singapore
Kevin Chang, UIUC
Chin-Wan Chung, Korea Advanced Institute of SaT
Gautam Das, University of Texas, Arlington
Aris Gkoulalas-Divanis, IBM Research Ireland
Torsten Grust, Universität Tübingen
Herodotos Herodotou, Microsoft Research
Yoshiharu Ishikawa, Nagoya University
Flip Korn, AT&T Labs

Nick Koudas, University of Toronto
Nikos Mamoulis, University of Hong Kong
Giansalvatore Mecca, Università della Basilicata
Alexandra Meliou, University of Washington
Rachel Pottinger, University of British Columbia
Rajeev Rastogi, Yahoo! India
Bernhard Seeger, University of Marburg
Ambuj Singh, University of California, Santa Barbara
Jens Teubner, ETH Zurich
Wei Wang, University of New South Wales
Li Xiong, Emory University
Jia Yuan Yu, IBM Research
Demetris Zeinalipour, University of Cyprus
Shuigeng Zhou, Fudan University
Local Administration
Ufficio Convegni and dbTrento Group, University of Trento

Logo Design
Sakis Palpanas

PVLDB Information Director
Gerald Weber, University of Auckland

PVLDB Advisory Committee
Philip Bernstein, Michael Böhlen, Peter Buneman, Susan Davidson, Z. Meral Ozsoyoglu, S. Sudarshan, Gerhard Weikum
LETTER FROM THE PROCEEDINGS EDITORS

This is the last issue of Vol. 6 of the Proceedings of the VLDB Endowment (PVLDB). The last two issues of Vol. 6 include a total of 47 papers that have been accepted by the VLDB 2013 program committee and will be presented at the VLDB 2014 conference in Hangzhou. Thus, a good number of VLDB 2013 papers will be presented in Hangzhou and we look forward to meet you there.

The 127 research papers published in PVLDB Vol. 6 have been selected from a total of 559 submissions. The decision process consisted of a packed schedule of monthly submissions, reviewing, discussions and revision cycles. The competition was fierce with an acceptance rate of 22% for research, 32% for industrial, and 46% for demonstration papers. There were 97 demonstration and 54 industrial paper submissions. 4% of the research papers were accepted directly without revision, and 23% of the papers were selected for a revision. 29 papers were submitted to the experiments and analysis track, and of these 12 papers got a revision decision.

This is the third year that the papers have been accepted only through journal-style reviewing with revisions and published in PVLDB. Much of this process has stabilized during the last few years and is increasingly being considered as a role model by other conferences and communities. It was a great privilege to further consolidate and advance this model in the context of VLDB 2013.

We would like to use this opportunity to thank the many people who contributed and made VLDB 2013 possible. We are grateful to the general chairs for their unwavering support during the past couple of years. They organized an exciting VLDB 2013 conference in a wonderful location. Our warm thanks go to Min Wang and Cong Yu for handling the industrial program; Jun Yang, Dimitrios Gunopulos, and Letizia Tanca for managing the demonstration program; Serge Abiteboul, Gianni Mecca, and Haixun Wang for choosing the tutorials; and Shivnath Babu, Stavros Harizopoulos, and Ihab Ilyas for selecting the panels. We thank all associate editors, PC members, and reviewers for their dedicated and hard work that made VLDB 2013 a big success. We thank the proceedings co-chairs Stratis Viglas and Peer Kröger, for putting the PVLDB issues together, and year around management of the collection and publication of camera-ready copies of accepted papers. And, finally, many thanks to all the authors for submitting their papers to PVLDB.

Michael Böhlen, University of Zurich
Christoph Koch, EPFL
Program Chairs, VLDB 2013