VLDB2011
37th International Conference on Very Large Data Bases, Seattle, WA

Proceedings of the VLDB Endowment

Volume 4, No. 11 – Research Track Papers and Keynotes
Proceedings of the 37th International Conference on Very Large Data Bases, Seattle, WA

Editor-in-Chief:
H. V. Jagadish

Guest Editors:
José Blakeley, Joseph M. Hellerstein, Nick Koudas, Wolfgang Lehner, Sunita Sarawagi, Uwe Röhm

PVLDB – Proceedings of the VLDB Endowment

Volume 4, No. 11, August 2011.

The 37th International Conference on Very Large Data Bases, Seattle, WA.

Copyright 2011 VLDB Endowment

Permission to make digital or hard copies of portions of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyright for components of this work owned by others than VLDB Endowment must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers or to redistribute to lists requires prior specific permission and/or a fee. Request permission to republish from PVLDB under email: info@vldb.org.

Volume 4, Number 11:
VLDB 2011 Keynotes and Final Research Track Papers

Pages ii – xv and 693 – 1212

ISSN 2150-8097, August 2011.

Additional copies only online at: portal.acm.org, arxiv.org/corr, and www.vldb.org

TABLE OF CONTENTS

Front Matter

Copyright Notice ii
Table of Contents iii
VLDB 2011 Conference Officers vi
PVLDB Review Board viii
List of External Reviewers xi
VLDB Endowment Board of Trustees xii

Letters

Welcome Message from the VLDB 2011 General Chair ..... *Per-Åke (Paul) Larson* xiii
Message from the VLDB 2011 General Program Chairs ..... *Joseph M. Hellerstein, José Blakeley* xiv

Keynotes

Towards a Global Brain ..... *Tim O'Reilly* 693
Is It Still "Big Data" If It Fits In My Pocket ..... *David Campbell* 694

10-Year Best Paper Award

Generic Schema Matching, Ten Years Later 695
..... *Philip A. Bernstein, Jayant Madhavan, Erhard Rahm*

Research Papers

A Framework for Supporting DBMS-like Indexes in the Cloud 702
..... *Gang Chen, Hoang Tam Vo, Sai Wu, Beng Chin Ooi, M. Tamer Özsu*

On Link-based Similarity Join 714
..... *Liwen Sun, Reynold Cheng, Xiang Li, David W. Cheung, Jiawei Han*

On Querying Historical Evolving Graph Sequences 726
..... *Chenghui Ren, Eric Lo, Ben Kao, Xinjie Zhu, Reynold Cheng*

RemusDB: Transparent High Availability for Database Systems 738
Umar Minhas, Shriram Rajagopalan, Brendan Cully, Ashraf Aboulnaga, Ken Salem, Andrew Warfield

Completeness of Queries over Incomplete Databases 749
..... *Simon Razniewski, Werner Nutt*

A Subsequence Matching with Gaps-Range-Tolerances Framework: Query-By-Humming Application 761
..... *Alexios Kotsifakos, Panagiotis Papapetrou, Jaakko Hollmén, Dimitris Gunopulos*

Approximate Substring Matching over Uncertain Strings 772
..... *Tingjian Ge, Zheng Li*

Serializable Snapshot Isolation for Replicated Databases in High-Update Scenarios..... 783
..... *Hyungsoo Jung, Hyuck Han, Alan Fekete, Uwe Röhm*

PALM: Parallel Architecture-Friendly Latch-Free Modifications to B+ Trees on Many-Core Processors 795
Jason Sewall, Jatin Chhugani, Changkyu Kim, Nadathur Satish, Pradeep Dubey

Mining Top-K Large Structural Patterns in a Massive Network 807
Feida Zhu, Qiang Qu, David Lo, Xifeng Yan, Jiawei Han, Philip S. Yu

Structure-Aware Sampling: Flexible and Accurate Summarization 819
Edith Cohen, Graham Cormode, Nick Duffield

Data Coordination: Supporting Contingent Updates 831
Michael Lawrence, Rachel Pottinger, Sheryl Staub-French

Accelerating Queries with Group-By and Join by Groupjoin 843
Guido Moerkotte, Thomas Neumann

Lightweight Graphical Models for Selectivity Estimation Without Independence Assumptions 852
Kostas Tzoumas, Amol Deshpande, Christian S. Jensen

Dissemination of Models over Time-Varying Data 864
Yongluan Zhou, Zografoula Vagena, Jonas Haustad

Efficient Subgraph Search over Large Uncertain Graphs 876
Ye Yuan, Guoren Wang, Haixun Wang, Lei Chen

Entangled Transactions 887
Nitin Gupta, Milos Nikolic, Sudip Roy, Gabriel Bender, Lucja Kot, Johannes Gehrke, Christoph Koch

Summary Graphs for Relational Database Schemas 899
Xiaoyan Yang, Cecilia M. Procopiuc, Divesh Srivastava

RecBench: Benchmarks for Evaluating Performance of Recommender System Architectures 911
Justin Levandoski, Michael Ekstrand, Michael Ludwig, Ahmed Eldawy, Mohamed Mokbel, John Riedl

Business Policy Modeling and Enforcement in Databases 921
Ahmed A. Ataullah, Frank W. Tompa

Online Data Fusion 932
Xuan Liu, Xin Luna Dong, Beng Chin Ooi, Divesh Srivastava

Optimistic Concurrency Control by Melding Trees 944
Philip A. Bernstein, Colin W. Reid, Ming Wu, Xinhao Yuan

Linking Temporal Records 956
Pei Li, Xin Luna Dong, Andrea Maurino, Divesh Srivastava

Efficient Algorithms for Finding Optimal Meeting Point on Road Networks 968
Da Yan, Zhou Zhao, Wilfred Ng

Optimal Schemes for Robust Web Extraction 980
Aditya Parameswaran, Nilesh Dalvi, Hector Garcia-Molina, Rajeev Rastogi

PathSim: Meta Path-Based Top-K Similarity Search in Heterogeneous Information Networks 992
Yizhou Sun, Jiawei Han, Xifeng Yan, Philip Yu, Tianyi Wu

Optimizing Query Answering under Ontological Constraints 1004
Giorgio Orsi, Andreas Pieris

OXPath: A Language for Scalable, Memory-efficient Data Extraction from Web Applications 1016
Tim Furche, Georg Gottlob, Giovanni Grasso, Christian Schallhart, Andrew Sellers

Optimizing and Parallelizing Ranked Enumeration 1028
Konstantin Golenberg, Benny Kimelfeld, Yehoshua Sagiv

Where in the World is My Data? ..... Sudarshan Kadambi, Jianjun Chen, Brian Cooper, 1040
David Lomax, Raghu Ramakrishnan, Adam Silberstein, Erwin Tam, Hector Garcia-Molina

Queries with Difference on Probabilistic Databases 1051
Sanjeev Khanna, Sudeepa Roy, Val Tannen

MRI: Meaningful Interpretations of Collaborative Ratings 1063
Mahashweta Das, Sihem Amer-Yahia, Gautam Das, Cong Yu

Storing Matrices on Disk: Theory and Practice Revisited 1075
Yi Zhang, Kamesh Munagala, Jun Yang

Publishing Set-Valued Data via Differential Privacy 1087
Rui Chen, Noman Mohammed, Benjamin C. M. Fung, Bipin C. Desai, Li Xiong

Randomized Generalization for Aggregate Suppression Over Hidden Web Databases 1099
Xin Jin, Nan Zhang, Aditya Mone, Gautam Das

Profiling, What-if Analysis, and Cost-based Optimization of MapReduce Programs 1111
Herodotos Herodotou, Shivnath Babu

Scalable SPARQL Querying of Large RDF Graphs 1123
Jiewen Huang, Daniel J. Abadi, Kun Ren

Online Aggregation for Large MapReduce Jobs 1135
Niketan Pansare, Vinayak Borkar, Chris Jermaine, Tyson Condie

Private Analysis of Graph Structure 1146
Vishesh Karwa, Sofya Raskhodnikov, Adam Smith, Grigory Yaroslavtsev

Stratification Criteria and Rewriting Techniques for Checking Chase Termination 1158
Sergio Greco, Francesca Spezzano, Irina Trubitsyna

Optimizing Probabilistic Query Processing on Continuous Uncertain Data 1169
Liping Peng, Yanlei Diao, Anna Liu

Massive Scale-out of Expensive Continuous Queries 1181
Erik Zeitler, Tore Risch

Keyword Search on Form Results 1189
Aditya Ramesh, S. Sudarshan, Purva Joshi

Efficient Rank Join with Aggregation Constraints 1201
Min Xie, Laks V.S. Lakshmanan, Peter T. Wood

VLDB 2011 CONFERENCE OFFICERS

General Chair

Per-Åke (Paul) Larson, Microsoft Research

Program Chairs

General PC Chairs:	José Blakeley, Microsoft
	Joseph M. Hellerstein, University of California - Berkeley
Research Track Chairs:	Nick Koudas, University of Toronto
	Wolfgang Lehner, Dresden University of Technology
	Sunita Sarawagi, IIT Bombay
Industrial Track Chairs:	Berthold Reinwald, IBM Almaden Research Center
	Phil Bohannon, Yahoo! Research
Challenges and Vision Track Chair:	Gerhard Weikum, Max-Planck-Institut für Informatik

Demo Program

Demo PC Chairs:	Jignesh Patel, University of Wisconsin – Madison
	Masatoshi Yoshikawa, University of Kyoto

Tutorial Program

Tutorial PC Chairs:	Qiong Luo, Hong Kong University of Science & Technology
	Gerome Miklau, University of Massachusetts – Amherst

Workshop Program

Workshop Chairs:	Luis Gravano, Columbia University
	Chris Jermaine, Rice University

PhD Program

PhD Workshop Chair:	Philippe Bonnet, IT University of Copenhagen

Proceedings

Proceedings Chair:	Uwe Röhm, University of Sydney

Ten-year Awards Committee

Bettina Kemme, McGill University, Canada

Tova Milo, Tel Aviv University, Israel

Divesh Srivastava, AT&T Labs (Chair)

Best Paper Awards Committee

Umesh Dayal, HP Labs

Donald Kossmann, ETH Zurich, Switzerland

Stan Zdonik, Brown University (Chair)

Sponsorship

David J. DeWitt, Jim Gray Systems Labs, Microsoft

Local Organization

Chair:	Arnd Christian König, Microsoft Research
Treasurer:	Magdalena Balazinska, University of Washington
Program Booklet Coordinator:	Yasin Silva, Arizona State University
Webmaster:	Justin Levandoski, University of Minnesota
Conference Management:	University of Washington Conference Management Team

PVLDB REVIEW BOARD

VLDB 2011 General PC Co-Chairs

José Blakeley, Microsoft

Joseph M. Hellerstein, University of California – Berkeley

VLDB 2011 Research Track Co-Chairs

Nick Koudas, University of Toronto and Sysomos Inc.

Wolfgang Lehner, Dresden University of Technology

Sunita Sarawagi, IIT Bombay

Reviewer

Ashraf Aboulnaga (University of Waterloo)
Sibel Adali (Rensselaer Polytechnic Institute)
Charu Aggarwal (IBM Watson Research Center)
Divyakant Agrawal (Univ. California, Santa Barbara)
Anastasia Ailamaki (EPFL Lausanne)
Gustavo Alonso (ETH Zurich)
Walid Aref (Purdue University)
Shivnath Babu (Duke University)
Roberto Bayardo (Google)
Elisa Bertino (Purdue University)
Peter Boncz (CWI, Netherlands)
Angela Bonifati (Icar-CNR)
Christof Bornhoevd (SAP Palo Alto)
Randal Burns (JHU)
Michael J. Cafarella (University of Michigan)
K. Selcuk Candan (Arizona State University)
Malu Castellanos (HP Labs)
Tiziana Catarci (University of Rome)
Chee-Yong Chan (National University of Singapore)
Kevin Chang (University of Illinois, Urbana-Champaign)
Surajit Chaudhuri (Microsoft Research)
Rada Chirkova (North Carolina State University)
Junghoo Cho (UCLA)
Jan Chomicki (University at Buffalo)
Chin-Wan Chung (Korea Advanced Institute of SaT)
Chris Clifton (Purdue University)
Christine Collet (Grenoble Institute of Technology)
Graham Cormode (AT&T Labs)
Nilesh Dalvi (Yahoo! Research)
Gautam Das (University of Texas, Arlington)
Anish Das Sarma (Yahoo! Research)
Umeshwar Dayal (HP Labs)
Amol Deshpande (University of Maryland)
AnHai Doan (University of Wisconsin)
Xin Dong (AT&T Labs)
Vuk Ercegovac (IBM Research – Almaden)
Alexandre Evfimievski (IBM Research)
Wenfei Fan (University of Edinburgh & Bell Labs)
Alan Fekete (University of Sydney)
Johann-Christoph Freytag (Humboldt-Universität Berlin)
Wolfgang Gatterbauer (University of Washington)
Johannes Gehrke (Cornell University)
Rainer Gemulla (Max-Planck-Institut für Informatik)
Dimitris Georgakopoulos (CSIRO)
Aristides Gionis (Yahoo! Research)
Goetz Graefe (HP Labs)
Torsten Grust (Universität Tübingen, Germany)
Giovanna Guerrini (University of Genova)

Dimitris Gunopulos (University of Athens, Greece)
Theo Haerder (University of Kaiserslautern)
Alon Halevy (Google)
Vagelis Hristidis (Florida International University)
Meichun Hsu (HP Labs, Palo Alto)
Svein-Olaf Hvasshovd (NTU, Norway)
Ihab Ilyas (University of Waterloo)
Zachary Ives (University of Pennsylvania)
Dean Jacobs (SAP)
Christian Jensen (Aalborg University)
Chris Jermaine (University of Florida)
Carl-Christian Kanne (University of Mannheim)
Raghav Kaushik (Microsoft Research)
Bettina Kemme (McGill University)
Alfons Kemper (Technische Universität München)
Eamonn Keogh (University of California, Riverside)
Martin Kersten (CWI)
Christoph Koch (Cornell University)
Flip Korn (AT&T Labs)
Donald Kossmann (ETH Zurich)
Rajasekar Krishnamurthy (IBM Research – Almaden)
Alberto Laender (Federal University of Minas Gerais)
Dongwon Lee (Penn State University)
Kristen Lefevre (University of Michigan)
Chen Li (University of California, Irvine)
Bin Liu (University of Michigan)
Fred Lochovsky (HKUST)
David Lomet (Microsoft Research)
Samuel Madden (MIT)
Nikos Mamoulis (University of Hong Kong)
Ioana Manolescu (INRIA)
Claudia Medeiros (University of Campinas, Brazil)
Alexandra Meliou (University of Washington)
Sergey Melnik (Google)
Marco Mesiti (Universita degli Studi di Milano)
Gerome Miklau (U. of Massachusetts – Amherst)
Chaitanya Mishra (Facebook Inc.)
Raymond Ng (University of British Columbia)
Siliva Nittel (University of Maine)
Christopher Olston (Yahoo! Research)
Fatma Ozcan (IBM Research – Almaden)
Sriram Padmanabhan (IBM)
Themis Palpanas (University of Trento)
Dimitris Papadias (Hong Kong University of SaT)
Stavros Papadopoulos (Chinese University of Hong Kong)
Yannis Papakonstantinou (UC San Diego)
Stefano Paraboschi (University of Bergamo)
Jian Pei (Simon Fraser University)
Rachel Pottinger (University of British Columbia)
Sunil Prabhaka (Purdue University)
Vijayshankar Raman (IBM Almaden Research Center)
Prakash Ramanan (Wichita State University)
Louiqa Raschid (University of Maryland)
Frank Renkes (SAP)
Kenneth Ross (Columbia University)
Elke Rundensteiner (Worcester Polytechnic Institute)
Yehoshua Sagiv (Hebrew University, Jerusalem)
Ken Salem (University of Waterloo)
Nikos Sarkas (University of Toronto)
Kai-Uwe Sattler (Ilmenau University of Technology)
Bernhard Seeger (University of Marburg)
Jayavel Shanmugasundaram (Yahoo! Research)
Eugene Shekita (IBM Research – Almaden)
Kyuseok Shim (Seoul National University)
Ambuj Singh (UC Santa Barbara)
Divesh Srivastava (AT&T Labs)
Dan Suciu (University of Washington)
S. Sudarshan (IIT Bombay)
Kian-Lee Tan (National University of Singapore)
Val Tannen (University of Pennsylvania)
Jens Teubner (ETH Zurich)
Martin Theobald (Max-Planck-Institut für Informatik)

Frank Tompa (University of Waterloo)
Panayiotis Tsaparas (Microsoft Research)
Anthony Tung (National University of Singapore)
Patrick Valduriez (INRIA)
Wei Wang (University of North Carolina)
Gerhard Weikum (Max Planck Institute, Germany)
Yuqing Wu (Indiana University)
Fei Xu (Microsoft Search)
Sihem Amer-Yahia (Yahoo! Research)
Jun Yang (Duke University)
Cong Yu (Yahoo! Research)
Jeffrey Xu Yu (Chinese University of Hong Kong)
Ting Yu (North Carolina State University)
Xiaohui Yu (York University)
Jingren Zhu (Microsoft Research)
Justin Zobel (University of Melbourne)

Steering Committee

Serge Abiteboul, Peter Apers, Philip Bernstein, Elisa Bertino, Peter Buneman, Martin Kersten, Z. Meral Ozsoyuglu

LIST OF EXTERNAL REVIEWERS

Sattam Alsubaiee (UC Irvine)
Hugo Alves (University of Campinas, Brazil)
Yael Amsterdamer (Tel Aviv University)
Paschalis Antoniou (University of Athens)
Arvind Arasu (Microsoft Research)
Nikos Armenantzoglou (HKUST, China)
Artin Avanes (Humboldt University Berlin)
Zhuowei Bao (University of Pennsylvania)
Sumita Barahmand (U Southern California)
Senjuti Basu Roy (UT Arlington)
Alex Behm (UC Irvine)
Manish Bhide (IBM Research - India)
Petko Bogdanov (UC Santa Barbara)
Xin Cao (NTU Singapore)
Federico Cavalieri (University of Genoa, Italy)
Inci Cetindil (UC Irvine)
Lijun Chang (Chinese University of Hong Kong)
Altigran da Silva (UFAM, Brazil)
Michele Dallachiesa (University of Trento)
Mahashweta Das (UT Arlington)
Sudipto Das (UC Santa Barbara)
Camil Demetrescu (University of Rome)
Prasad Deshpande (IBM Research - India)
Daniel Deutch (Ben Gurion University)
Alin Deutsch (UC San Diego)
Iman Elghandour (University of Waterloo)
Shel Finkelstein (SAP)
Joana G. Malaverri (University of Campinas)
Sandro Gatto (CPqD Research Center, Brazil)
Gabriel Ghinita (Purdue University)
George Giannakopoulos (NCSR Demokritos)
Lukasz Golab (AT&T Labs Research)
Marcos Gonçalves (UFMG, Brazil)
Torsten Grabs (Microsoft)
Jin Guang (University of Maine)
Lin Guo (Harbin English University)
Chetan Gupta (HP Labs)
Bingsheng He (NTU Singapore)
Volker Höfner (University of Kaiserslautern)
Jian Huang (Pennsylvania State University)
Shunping Huang (University of North Carolina)
Shengyue Ji (UC Irvine)
Ricardo Jimenez-Peris (University of Madrid)
Ning Jin (University of North Carolina)
Alexis Joly (INRIA)
Konstantinos Karanasos (INRIA)
George Kellaris (HKUST, China)
Kraig King (University of Maine)
Dimitrios Kotsakos (University of Athens)
Isaac Kunen (Microsoft)
Jaewoo Lee (Purdue University)
Domenico Lembo (University of Rome)
Maurizio Lenzerini (University of Rome)
Jian Li (University of Maryland)
Qinghan Liang (University of Maine)
Jessica Lin (George Mason University)
Kathy Macropol (UC Santa Barbara)
Alice Marascu (University of Trento)
Hossein Maserrat (Simon Fraser University)
Florent Masseglia (INRIA)
Katsiaryna Mirylenka (University of Trento)
Misael Mongiovi (UC Santa Barbara)
Mirella Moro (UFMG, Brazil)
Matheus Motta (University of Campinas, Brazil)
Davide Mottin (University of Trento)
Victor Muntés i Mulero (CA Labs, Barcelona)
Mohamed Nabeel (Purdue University)
Alan Nakai (University of Campinas, Brazil)
Yi Ou (University of Kaiserslautern)
Stavros Papadopoulos (Chinese University HK)
Panagiotis Papapetrou (Aalto University)
Jignesh Patel (University of Wisconsin)
Gerardo Pelosi (Politecnico di Milano)
Pouria Pirzadeh (UC Irvine)
Lucian Popa (IBM Almaden Research Center)
Jon Purnell (Rensselaer Polytechnic Institute)
Yan Qi (Teradata)
Miao Qiao (Chinese University of Hong Kong)
Lu Qin (Chinese University of Hong Kong)
Sayan Ranu (UC Santa Barbara)
Brian Ruttenberg (UC Santa Barbara)
André Santanchè (UNICAMP, Brazil)
Ismael Sanz (University Jaume I, Spain)
Monica Scannapieco (Istat, Italy)
James Schaffer (UC Santa Barbara)
Daniel Schall (University of Kaiserslautern)
Karsten Schmidt (University of Kaiserslautern)
Jie Shao (University of Melbourne)
Entong Shen (North Carolina State University)
Vishwakarma Singh (UC Santa Barbara)
Emad Soroush (University of Washington)
Weifeng Su (BNU-HKBU United Int'l College)
Christine Task (Purdue University)
Saravanan Thirumuruganathan (UT Arlington)
Andreas Thor (University of Maryland)
George Trimponias (HKUST, China)
Kostas Tzoumas (Aalborg University)
Georgios Valkanas (University of Athens)
Adriano Veloso (UFMG, Brazil)
Rares Vernica (UC Irvine)
Bruno Vilar (University of Campinas, Brazil)
Akrivi Vlachou (Athens University of EaB)
Hannes Voigt (TU Dresden)
Di Wang (Worcester Polytechnic Institute)
Jiying Wang (City University of Hong Kong)
Lixing Wang (HKUST, China)
Rui Wang (Microsoft Research)
Weibo Wang (University of North Carolina)
Yafang Wang (Max-Planck-Institut f.Informatik)
Wei Wei (North Carolina State University)
Andreas Weiner (University of Kaiserslautern)
Karen Works (Worcester Polytechnic Institute)
Dingming Wu (Aalborg University)
Fei Wu (Google)
Xiaoheng Xie (HKUST, China)
Di Yang (Worcester Polytechnic Institute)
Yin Yang (HKUST, China)
Nan Zhang (George Washington University)
Xiang Zhang (Case Western Reserve University)
Zhaojun Zhang (University of North Carolina)

VLDB ENDOWMENT BOARD OF TRUSTEES

The VLDB Endowment now has a board of 21 elected trustees, who are the legal guardians of the Endowment's charter and activities. Trustees are elected for a six-year period; the election procedure is documented in the code of regulations available on the Endowment's Web site. The board is continuously renewed, with one third of its members being replaced every two years. The trustees are elected among internationally distinguished researchers and professionals in the field of database and information systems who have contributed to the objectives of the Endowment with dedication and distinction. All trustees provide voluntary service.

Executive

President: Renée J. Miller

Vice-President: Paolo Atzeni

Treasurer: Michael J. Franklin

Secretary: Beng Chin Ooi

Current trustees (as of January 2010):

Gustavo Alonso

Paolo Atzeni

Susan B. Davidson

Michael J. Franklin

Donald Kossmann

David B. Lomet

Renée J. Miller

Tova Milo

Beng Chin Ooi

Z. Meral Özsoyoğlu

Sunita Sarawagi

Anastasia Ailamaki

Sihem Amer-Yahia

Michael H. Böhlen

Michael J. Carey

Surajit Chaudhuri

Alon Y. Halevy

Volker Markl

Divesh Srivastava

S. Sudarshan

Kyu-Young Whang

WELCOME MESSAGE FROM THE VLDB 2011 GENERAL CHAIR

On behalf of the conference committee for VLDB 2011, it is my pleasure to welcome you to Seattle and the state of Washington for the 37th International Conference on Very Large Data Bases.

VLDB is a premier international forum for data management and database researchers, vendors, practitioners, application developers, and users. The conference will feature two keynotes, research talks, industry presentations, panels, tutorials, demonstrations, and eleven workshops. The program covers a wide range of topics in data management, database and information systems research. Data management and databases remain one of the key technical underpinnings of a broad variety of applications.

Seattle is the largest city in Washington State with over 3.3 million residents in the Seattle metropolitan area, around half of the state's population. The city is situated on the shores of Puget Sound, a long narrow inlet from the Pacific Ocean. It has a vibrant economy with many well-known companies in aerospace, software, and retailing. University of Washington is one of the top research universities in the country and is particularly renowned for its medical education, care and research. Many companies and research centers in the medical and biological fields are clustered near the university. Washington State is also a major fruit and wine producer with vast orchards and vineyards east of the Cascade Mountains. Don't miss the opportunity to sample some excellent Washington wines!

The conference would not have been possible without the efforts of many people. Thanks are due to the Technical Program Chairs - Dr. Jose Blakeley and Dr. Joe Hellerstein, and their team of track program chairs: Research track – Dr. Nick Koudas, Dr. Wolfgang Lehner, and Dr. Sunita Sarawagi; Industrial and Applications track – Dr. Berthold Reinwald and Dr. Phil Bohannon; Demo track – Dr. Jignesh Patel and Dr. Masatoshi Yoshikawa; and Tutorial track – Dr. Qiang Lou, and Dr. Gerome Miklau. Thanks also for the efforts of the workshop co-chairs Dr. Luis Gravano and Dr. Chris Jermaine and the Ph.D. Workshop chair Dr. Philippe Bonnet. Many, many thanks go to Dr. Uwe Roehm (Proceedings chair), Dr. Arnd Christian König (Local Organization chair), Dr. Magda Balazinska (Treasurer), Dr. David Dewitt (Sponsorship chair), Dr. Justin Levandoski (Webmaster), Dr. Yasin Silva (Program Booklet organizer), and the student volunteers. The contributions of the University of Washington Conference Management Team, ably led by Ms. Jan Kvamme, were invaluable and saved the organizers many sleepless nights.

We are grateful for the generous support of our sponsors: Platinum Sponsors – Facebook, Google, Microsoft and Oracle; Gold Sponsors - EMC, IBM, SAP, and Sybase; Silver Sponsors - HP, NEC, TPC, and Yahoo!, and Workshop Sponsors – Intelius (QDB) and Microsoft (LADIS).

Last but not least, our sincere thanks go to the authors of papers and demos, the many reviewers, the speakers, and everybody who contributed to the program. Without your contributions there would be no conference.

Welcome and enjoy the conference and your time in Seattle!

Per-Åke (Paul) Larson, Microsoft Research
VLDB 2011 General Chair

MESSAGE FROM THE VLDB 2011 GENERAL PROGRAM CHAIRS

Welcome to VLDB 2011 in Seattle! After a year of hard work from authors, reviewers, conference and journal organizers, we are excited to join in the community discussion at the conference.

VLDB kicks off this year with keynote talks from two speakers who provide a broad view of the landscape of data-driven technologies and their applications: Tim O'Reilly of O'Reilly Media, and David Campbell of Microsoft. We're delighted they were able to set aside time from their schedules to join us.

VLDB 2011 marks the end of the first year of a new experiment in paper reviewing and conference presentations designed by the VLDB Endowment. The Research Track of this year's conference consists of the full set of papers that appeared in the 2011 volume of the Proceedings of the VLDB Endowment (PVLDB). PVLDB is a journal consisting of short "conference-style" papers, with a monthly process designed to promote timely submission, review, and revision of scholarly results. We are grateful to the community of authors, reviewers and conference officers who worked year-round to make this experiment a reality. We also thank VLDB Proceedings Chair Uwe Röhm and PVLDB Editor-in-Chief H. V. Jagadish for their coordination and guidance in handling the rolling publication of PVLDB papers from the Research Track, as well as the final conference proceedings.

As in previous years, VLDB 2011 includes technical papers in research and industrial tracks, as well as a strong set of workshops, tutorials, demos and panels. In addition, as is traditional, selection committees singled out two research papers for recognition. The 10-year Best PaperAward for the paper from VLDB 2001 that has had the largest impact is "Generic Schema Matching with Cupid" by Jayant Madhavan, Philip A. Bernstein, and Erhard Rahm. The VDLB 2011 Best Paper Award is "RemusDB: Transparent High-Availability for Database Systems" by Umar Farooq Minha, Shriram Rajagopalan, Brendan Cully, Ashraf Aboulnaga, Ken Salem, and Andrew Warfield. We are grateful to the 10-year award selection committee consisting of Divesh Srivastava, Bettina Kemme, and Tova Milo; and the best paper award selection committee consisting of Umesh Dayal, Donald Kossman, and Stan Zdonik for their contribution.

This year's conference includes a Challenges and Vision Track held in coordination with the Computing Community Consortium (CCC). The track emphasizes visionary ideas, long term challenges, and opportunities in data-centric research that are outside of the current mainstream topics of the field. Submissions were judged on the extent to which they expand the possibilities and horizons of the field. The CCC sponsored three awards for best papers in the Challenges and Vision Track, which will be announced at the conference.

This year's edition of VLDB received 553 submissions to the research track, of which 100 papers (18.1%) were accepted. We received 298 submissions (54%) spread from April 1, 2010 to February 1, 2011 for an average of 27 papers per month plus 255 submissions (46%) for the March 1, 2011 deadline. We would like to encourage the community to submit their work more evenly throughout the year.

The industrial and applications track received 41 submissions, of which 12 papers (27%) were accepted. The demonstrations track received 89 submissions, of which 28 (31%) were accepted.

Our sincere thanks to all our colleagues who volunteered to serve as conference officers, program committee members and reviewers on this year's VLDB and PVLDB. We want to especially acknowledge the Research Track Co-Chairs – Wolfgang Lehner, Sunita Sarawagi and Nick Koudas – who took on the relentless task of driving the PVLDB review process to a conclusion every month. In addition, we would like to thank Berthold Reinwald and Phil Bohannon for handling the industry and applications track; Jignesh Patel and Masatoshi Yoshikawa for handling the demonstrations program; Qiong Luo and Gerome Miklau for handling the tutorials program.

We are also extremely grateful to General Chair Paul Larson and PVLDB Editor-in-Chief H.V. Jagadish, who calmly guided us all through this year's challenging new format. The staff at Microsoft's Conference Management Toolkit deserves recognition for their help in customizing CMT to accommodate the new structure of PVLDB and the VLDB conference.

Joseph M. Hellerstein, University of California – Berkeley
José Blakeley, Microsoft
VLDB 2011 General Program Chairs