Preface

DASFAA is an annual international database conference, located in the Asia-Pacific region, which showcases state-of-the-art R & D activities in database systems and their applications. It provides a forum for technical presentations and discussions among database researchers, developers and users from academia, business and industry. DASFAA 2009, the 14th in the series, was held during April 20-23, 2009 in Brisbane, Australia.

In this year, we carefully selected six workshops, each focusing on specific research issues that contribute to the main themes of the DASFAA conference. This volume contains the final versions of papers accepted for these six workshops that were held in conjunction with DASFAA 2009. They are:

- First International Workshop on Benchmarking of XML and Semantic Web Applications (BenchmarX 2009)
- First International Workshop on Data and Process Provenance (WDPP 2009)
- First International Workshop on Privacy-Preserving Data Analysis (PPDA 2009)
- First International Workshop on Mobile Business Collaboration (MBC 2009)
- DASFAA 2009 PhD Workshop

All the workshops were selected via a public call-for-proposals process. The workshop organizers put a tremendous amount of effort into soliciting and selecting papers with a balance of high quality, new ideas and new applications. We asked all workshops to follow a rigid paper selection process, including the procedure to ensure that any Program Committee members are excluded from the paper review process of any paper they are involved with. A requirement about the overall paper acceptance rate of no more than 50% was also imposed on all the workshops.

The conference and the workshops received generous financial support from The University of Melbourne, The University of New South Wales, The University of Sydney, The University of Queensland, National ICT Australia (NICTA), Australian Research Council (ARC) Research Network in Enterprise Information Infrastructure (EII), ARC Research Network on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), and ARC Research Network for a Secure Australia. We also received extensive help and logistic support from the DASFAA Steering Committee, The University of Queensland, Tokyo Institute of Technology and the Conference Management Toolkit Support Team at Microsoft.

We are very grateful to Xiaofang Zhou, Haruo Yokota, Qing Liu, Ke Deng, Shazia Sadiq, Gabriel Fung, Kathleen Williamson and many other people for
their effort in supporting the workshop organization. We would like to take this opportunity to thank all workshop organizers and Program Committee members for their effort to put together the workshop program of DASFAA 2009.

April 2009

Lei Chen
Chengfei Liu
DASFAA 2009 Workshop Organization

Workshop Chairs

Lei Chen  
Hong Kong University of Science and Technology, China

Chengfei Liu  
Swinburne University of Technology, Australia

Publication Chairs

Qing Liu  
CSIRO, Australia

Ke Deng  
University of Queensland, Australia

BenchmarkX 2009 Workshop

Program Committee Chairs

Michal Kratky  
Technical University of Ostrava, Czech Republic

Irena Mlynkova  
Charles University in Prague, Czech Republic

Eric Pardede  
La Trobe University, Bundoora, Australia

Program Committee

Radim Baca  
Technical University of Ostrava, Czech Republic

Martine Collard  
INRIA Sophia Antipolis, France

Jiri Dokulil  
Charles University in Prague, Czech Republic

Peter Gursky  
Pavol Jozef Safarik University in Kosice, Slovakia

Tomas Horvath  
Pavol Jozef Safarik University in Kosice, Slovakia

Jana Katreniakova  
Comenius University in Bratislava, Slovakia

Markus Kirchberg  
Institute for Infocomm Research, A*STAR, Singapore

Agnes Koschmider  
Institute AIFB, Universität Karlsruhe, Germany

Michal Kratky  
Technical University of Ostrava, Czech Republic

Sebastian Link  
Victoria University of Wellington, New Zealand
VIII Organization

Pavel Loupal  
Czech Technical University in Prague, Czech Republic

Mary Ann Malloy  
The MITRE Corporation, USA

Marco Mevius  
Institute AIFB, Universität Karlsruhe, Germany

Irena Mlynkova  
Charles University in Prague, Czech Republic

Martin Necasky  
Charles University in Prague, Czech Republic

Alexander Paar  
Universität Karlsruhe, Germany

Incheon Paik  
The University of Aizu, Japan

Eric Pardede  
La Trobe University, Bundoora, Australia

Jorge Perez  
Pontificia Universidad Catolica, Chile

Dmitry Shaporenkov  
University of Saint Petersburg, Russia

Michal Valenta  
Czech Technical University in Prague, Czech Republic

---

**MCIS 2009 and WDPP 2009 Workshops**

**Program Committee Chairs**

**MCIS**

Shazia Sadiq  
The University of Queensland, Australia

Ke Deng  
The University of Queensland, Australia

Xiaofang Zhou  
The University of Queensland, Australia

Xiaochun Yang  
Northeastern University, China

**WDPP**

Walid G. Aref  
Purdue University, USA

Alex Delis  
University of Athens, Greece

Qing Liu  
CSIRO ICT Centre, Australia

**Publicity Chair (WDPP)**

Kai Xu  
CSIRO ICT Centre, Australia

**Program Committee**

**MCIS**

Yi Chen  
Arizona State University, USA

Markus Helfert  
Dublin City University, UK

Ruoming Jin  
Kent State University, USA

Chen Li  
UC Irvine, USA

Jiaheng Lu  
Renmin University, China

Graeme Shanks  
Monash University, Australia

Can Turker  
FGCZ Zurich, Switzerland

Haixun Wang  
IBM, USA

Xuemin Lin  
UNSW, Australia
WDPP
Mohamed S. Abougabal  University of Alexandria, Egypt
Ilkay Altintas  San Diego Supercomputer Centre, USA
Athan Bouguettaya  CSIRO, Australia
Susan B. Davidson  University of Pennsylvania, USA
Antonios Deligiannakis  Technical University of Crete, Greece
Juliana Freire  University of Utah, USA
James Frew  University of California, Santa Barbara, USA
Paul Groth  University of Southern California, USA
Georgia Koutrika  Stanford University, USA
Bertram Lodascher  University of California, Davis, USA
Simon McBride  The Australian E-Health Research Centre, Australia
Simon Miles  King’s College London, UK
Brahim Medjahed  University of Michigan, Dearborn, USA
Khaled Nagi  Alexandria University, Egypt
Anne Ngu  Texas State University, San Marcos, USA
Mourad Ouazzani  Purdue University, USA
Thomas Risse  L3S Lab, Germany
Satya S. Sahoo  Kno.e.sis Center, Wright State University, USA
Zahir Tari  RMIT, Australia
Qi Yu  Rochester Institute of Technology, USA
Jun Zhao  University of Oxford, UK

PPDA 2009 Workshop
Raymond Chi-Wing Wong  The Hong Kong University of Science and Technology, China
Ada Wai-Chee Fu  The Chinese University of Hong Kong, China

Program Committee
Claudio Bettini  University of Milan, Italy
Chris Clifton  Purdue University, USA
Claudia Diaz  K.U.Leuven, Belgium
Josep Domingo-Ferrer  Rovira i Virgili University, Spain
Elena Ferrari  University of Insubria, Italy
Sara Foresti  University of Milan, Italy
Benjamin C.M. Fung  Concordia University, Canada
Christopher Andrew Leckie  The University of Melbourne, Australia
Jiuyong Li  University of South Australia, Australia
Jun-Lin Lin  Yuan Ze University, Taiwan
Kun Liu IBM Almaden Research Center, USA
Ashwin Machanavajjhala Cornell University, USA
Bradley Malin Vanderbilt University, USA
Nikos Mamoulis Hong Kong University, China
Wee Keong Ng Nanyang Technological University, Singapore
Jian Pei Simon Fraser University, Canada
Yucel Saygin Sabanci University, Turkey
Jianhua Shao Cardiff University, UK
Yufei Tao The Chinese University of Hong Kong, China
Vicenc Torra Universitat Autonoma de Barcelona, Spain
Carmela Troncoso K.U. Leuven, Belgium
Hua Wang University of Southern Queensland, Australia
Ke Wang Simon Fraser University, Canada
Sean Wang University of Vermont, USA
Duminda Wijesekera George Mason University, USA
Xintao Wu University of North Carolina at Charlotte, USA
Jeffrey Yu The Chinese University of Hong Kong, China
Philip S. Yu University of Illinois at Chicago, USA

MBC 2009 Workshop

General Chairs
Qing Li City University of Hong Kong, China
Hua Hu Zhejiang Gongshang University, China

Program Committee Chairs
Dickson K.W. Chiu Dickson Computer Systems, Hong Kong, China
Yi Zhuang Zhejiang Gongshang University, China

Program Committee
Patrick C.K. Hung University of Ontario Institute of Technology, Canada
Samuel P.M. Choi The Open University of Hong Kong, China
Eleanna Kafeza Athens University of Economics and Commerce, Greece
Baihua Zheng Singapore Management University, Singapore
Edward Hung Hong Kong Polytechnic University, China
Ho-fung Leung Chinese University of Hong, China
Zakaria Maamar Zayed University, UAE
Stefan Voss University of Hamburg, Germany
Cuiping Li Renmin University, China
Chi-hung Chi  National Tsing Hua University, Taiwan, China
Stephen Yang  National Central University, Taiwan, China
Ibrahim Kushchu  Mobile Government Consortium International, UK
Haiyang Hu  Zhejiang Gongshang University, China
Huiye Ma  CWI, The Netherlands
Pirkko Walden  Abo Akademi University, Finland
Raymond Wong  National ICT, Australia
Lidan Shou  Zhejiang University, China
Matti Rossi  Helsinki School of Economics, Finland
Achim Karduck  Furtwangen University, Germany

DASFAA 2009 PhD Workshop

Program Committee Chairs

Wei Wang  University of New South Wales, Australia
Baihua Zheng  Singapore Management University, Singapore

Program Committee

Bin Cui  Peking University, China
Jianlin Feng  Zhongshan University, China
Haifeng Jiang  Google, USA
Takahiro Hara  Osaka University, Japan
Wang-chien Lee  Pennsylvania State University, USA
Jiaheng Lu  Renming University, China
Lidan Shou  Zhejiang University, China
Bill Shui  NICTA, Australia
Xueyan Tang  Nanyang Technological University, Singapore
Jianliang Xu  Hong Kong Baptist University, China
# Table of Contents

## First International Workshop on Benchmarking of XML and Semantic Web Applications (BenchmarX’09))

- Workshop Organizers’ Message .......................................................... 3  
  Michal Krátký, Irena Mlynkova, and Eric Pardede
- Current Approaches to XML Benchmarking (Invited Talk) .................... 4  
  Stéphane Bressan
- TJDewey – On the Efficient Path Labeling Scheme Holistic Approach ....................................................... 6  
  Radim Bača and Michal Krátký
- The XMLBench Project: Comparison of Fast, Multi-Platform XML Libraries .......................................................... 21  
  Suren Chilingaryan
- A Synthetic, Trend-Based Benchmark for XPath .................................. 35  
  Curtis Dyreson and Hao Jin
- An Empirical Evaluation of XML Compression Tools ............................ 49  
  Sherif Sakr
- Benchmarking Performance-Critical Components in a Native XML Database System .................................................. 64  
  Karsten Schmidt, Sebastian Bächle, and Theo Härder
- On Benchmarking Transaction Managers ............................................... 79  
  Pavel Strnad and Michal Valenta

## Second International Workshop on Managing Data Quality in Collaborative Information Systems and First International Workshop on Data and Process Provenance (MCIS’09 & WDPP’09)

- Workshop Organizers’ Message .......................................................... 95  
  Shazia Sadiq, Ke Deng, Xiaofang Zhou, Xiaochun Yang,  
  Walid G. Aref, Alex Delis, Qing Liu, and Kai Xu
- Data Provenance Support in Relational Databases for Stored Procedures .......................................................... 97  
  Winly Jurnawan and Uwe Röhm
A Vision and Agenda for Theory Provenance in Scientific Publishing ... 112
  *Ian Wood, J. Walter Larson, and Henry Gardner*

Probabilistic Ranking in Uncertain Vector Spaces ......................... 122
  *Thomas Bernecker, Hans-Peter Kriegel, Matthias Renz, and Andreas Zuefle*

Logical Foundations for Similarity-Based Databases ..................... 137
  *Radim Belohlavek and Vilem Vychodil*

Tailoring Data Quality Models Using Social Network Preferences ...... 152
  *Ismael Caballero, Eugenio Verbo, Manuel Serrano, Coral Calero, and Mario Piattini*

The Effect of Data Quality Tag Values and Usable Data Quality Tags on Decision-Making .................................................. 167
  *Rosanne Price and Graeme Shanks*

Predicting Timing Failures in Web Services .............................. 182
  *Nuno Laranjeiro, Marco Vieira, and Henrique Madeira*

A Two-Tire Index Structure for Approximate String Matching with Block Moves ......................................................... 197
  *Bin Wang, Long Xie, and Guoren Wang*

**First International Workshop on Privacy-Preserving Data Analysis (PPDA’09)**

Workshop Organizers’ Message ..................................................... 215
  *Raymond Chi-Wing Wong and Ada Wai-Chee Fu*

Privacy Risk Diagnosis: Mining l-Diversity .................................. 216
  *Mohammad-Reza Zare-Mirakabad, Aman Jantan, and Stéphane Bressan*

Towards Preference-Constrained k-Anonymisation ......................... 231
  *Grigoris Loukides, Achilles Tziatzios, and Jianhua Shao*

Privacy FP-Tree ............................................................................. 246
  *Sampson Pun and Ken Barker*

Classification with Meta-learning in Privacy Preserving Data Mining ... 261
  *Piotr Andruszkiewicz*

Importance of Data Standardization in Privacy-Preserving K-Means Clustering ................................................................. 276
  *Chunhua Su, Justin Zhan, and Kouichi Sakurai*
# First International Workshop on Mobile Business Collaboration (MBC’09)

## Workshop Organizers’ Message

Dickson K.W. Chiu and Yi Zhuang

A Decomposition Approach with Invariant Analysis for Workflow Coordination

Jidong Ge and Haiyang Hu

An Efficient P2P Range Query Processing Approach for Multi-dimensional Uncertain Data

Ye Yuan, Guoren Wang, Yongjiao Sun, Bin Wang, Xiaochun Yang, and Ge Yu

Flexibility as a Service

W.M.P. van der Aalst, M. Adams, A.H.M. ter Hofstede, M. Pesic, and H. Schonenberg

Concept Shift Detection for Frequent Itemsets from Sliding Windows over Data Streams

Jia-Ling Koh and Ching-Yi Lin

A Framework for Mining Stochastic Model of Business Process in Mobile Environments

Haiyang Hu, Bo Xie, JiDong Ge, Yi Zhuang, and Hua Hu

## DASFAA 2009 PhD Workshop

## Workshop Organizers’ Message

Wei Wang and Baihua Zheng

Encryption over Semi-trusted Database

Hasan Kadhem, Toshiyuki Amagasa, and Hiroyuki Kitagawa

Integration of Domain Knowledge for Outlier Detection in High Dimensional Space

Sakshi Babbar

Towards a Spreadsheet-Based Service Composition Framework

Dat Dac Hoang, Boualem Benatallah, and Hye-young Paik

## Author Index

375