

# **2013 IEEE 29th International Conference on Data Engineering Workshops**

**(ICDEW 2013)**

**Brisbane, Australia  
8 – 12 April 2013**



**IEEE Catalog Number: CFP1345A-PRT  
ISBN: 978-1-4673-5303-8**

---

## Workshop 1: Data Engineering Meets the Semantic Web (DESWEB)

---

(i)  **DESWEB 2013 Preface Pages**


### REGULAR PAPERS

1  **WARP: Workload-Aware Replication and Partitioning for RDF**

*(Katja Hose, Ralf Schenkel)*

7  **SESM: Semantic Enrichment of Schema Mappings**

*(Yoones A. Sekhavat, Jeffrey Parsons)*

13  **Introducing Shadows: Flexible Document Representation and Annotation on the Web**

*(Matheus Silva Mota, Claudia Bauzer Medeiros)*


### LATE-BREAKING RESULTS, VISIONS AND CHALLENGES

19  **Automated Educated Guessing**


*(Aleksandar Stupar, Sebastian Michel)*

21  **Eight Fallacies When Querying the Web of Data**


*(Jürgen Umbrich, Claudio Gutierrez, Aidan Hogan, Marcel Karnstedt, Josiane Xavier Parreira)*

23  **Hybrid Graph and Relational Query Processing in Main Memory**

*(Martin Grund, Philippe Cudre-Mauroux, Jens Krueger, Hasso Plattner)*

25  **A Vision for SPARQL Multi-Query Optimization on MapReduce**

*(Kemafor Anyanwu)*

27  **Recommending Environmental Knowledge as Linked Open Data Cloud Using Semantic Machine Learning**

*(Ahsan Morshed, Ritaban Dutta, Jagannath Aryal)*

---

## Workshop 2: Self-Managing Database Systems (SMDB)

---

(i)  **SMDB 2013 Preface Pages**

### APPLICATIONS OF SELF-MANAGEMENT

29  **Realistic Tenant Traces for Enterprise DBaaS**

*(Jan Schaffner, Tim Januschowski)*

36  **Model Ensemble Tools for Self-Management in Data Centers**

*(Jin Chen, Gokul Soundararajan, Saeed Ghanbari, Cristiana Amza)*

44  **Total Operator State Recall — Cost-Effective Reuse of Results in Greenplum Database**

*(George C. Caragea, Carlos Garcia-Alvarado, Michalis Petropoulos, Florian M. Waas)*

### FOUNDATIONS OF SELF-MANAGEMENT

50  **INUM<sup>+</sup>: A Leaner, More Accurate and More Efficient Fast What-If Optimizer**


*(Rui Wang, Quoc Trung Tran, Ivo Jimenez, Neoklis Polyzotis)*

56  **Automatic Schema Design for Co-Clustered Tables**

*(Stephan Baumann, Peter Boncz, Kai-Uwe Sattler)*

62  **Performance Optimization for Distributed Intra-Node-Parallel Streaming Systems**

*(Matthias J. Sax, Malu Castellanos, Qiming Chen, Meichun Hsu)*

70  **Self-Managing Load Shedding for Data Stream Management Systems**

*(Thao N. Pham, Panos K. Chrysanthis, Alexandros Labrinidis)*




---

## Workshop 3: Privacy-Preserving Data Publication and Analysis (PrivDB)



---

(i)  **PrivDB 2013 Preface Pages**

### RESEARCH SESSION 1

- 77  **Empirical Privacy and Empirical Utility of Anonymized Data**  
*(Graham Cormode, Cecilia M. Procopiuc, Entong Shen, Divesh Srivastava, Ting Yu)*
- 83  **Privacy-Protecting Index for Outsourced Databases**  
*(Chung-Min Chen, Andrzej Cichocki, Allen McIntosh, Euthimios Panagos)*
- 88  **On Syntactic Anonymity and Differential Privacy**  
*(Chris Clifton, Tamir Tassa)*

### RESEARCH SESSION 2

- 94  **On Information Leakage by Indexes Over Data Fragments**  
*(Sabrina De Capitani di Vimercati, Sara Foresti, Sushil Jajodia, Stefano Paraboschi, Pierangela Samarati)*
- 99  **Privacy Against Aggregate Knowledge Attacks**  
*(Olga Gkountouna, Katerina Lepenioti, Manolis Terrovitis)*




---

## Workshop 4: Mobile Data Analytics (MoDA)




---

(i)  **MoDA 2013 Preface Pages**

### RESEARCH PAPERS

- 104  **Client-Centric OLAP on Mobile Devices**  
*(Zheng Xu, Wo-Shun Luk, Stephen Petchulat)*
- 112  **Signature Generation for Sensitive Information Leakage in Android Applications**  
*(Hiroki Kuzuno, Satoshi Tonami)*
- 120  **RFID Based Vehicular Networks for Smart Cities**  
*(Joydeep Paul, Baljeet Malhotra, Simon Dale, Meng Qiang)*









### INVITED PAPERS

- 128  **ShareLikesCrowd: Mobile Analytics for Participatory Sensing and Crowd-Sourcing Applications**  
*(Arkady Zaslavsky, Prem Prakash Jayaraman, Shonali Krishnaswamy)*
- 136  **Strong Location Privacy: A Case Study on Shortest Path Queries**  
*(Kyriakos Mouratidis)*
- 144  **On the Link(s) Between “D” and “A” in Mobile Data Analytics**  
*(Goce Trajcevski)*

---

## Workshop 5: Data-Driven Decision Guidance and Support Systems (DGSS)







---

- (i)  **DGSS 2013 Preface Pages**
- 152  **Will Internet of Things Flood DGSS with Data?**  
*(Arkady Zaslavsky)*
- 153  **Using Military Operational Planning System Data to Drive Reserve Stocking Decisions**  
*(Rajesh Thiagarajan, Mirza Arif Mekhtiev, Greg Calbert, Nikifor Jeremic, Don Gossink)*
- 163  **SmartCart: A Consolidated Shopping Cart for Pareto-Optimal Sourcing and Fair Discount Distribution**  
*(Brian Goodhart, Venkata Yerneni, Alex Brodsky, Venkata Rudraraju, Nathan Egge)*
- 173  **Water Desalination Supply Chain Modelling and Optimization**  
*(Malak T. Al-Nory, Stephen C. Graves)*
- 181  **Using Graphical Models and Multi-Attribute Utility Theory for Probabilistic Uncertainty Handling in Large Systems, with Application to the Nuclear Emergency Management**  
*(Manuele Leonelli, James Q. Smith)*
- 193  **Multivariate Data-Driven Decision Guidance for Clinical Scientists**  
*(Frada Burstein, Daswin De Silva, Herbert F. Jelinek, Andrew Stranieri)*
- 200  **ODSS: A Decision Support System for Ocean Exploration**  
*(Kevin Gomes, Danelle Cline, Duane Edgington, Michael Godin, Thom Maughan, Mike McCann, Tom O'Reilly, Fred Bahr, Francisco Chavez, Monique Messié, Jnaneshwar Das, Kanna Rajan)*

---

## Workshop 6: Graph Data Management: Techniques and Applications (GDM)

---

- (i)  **GDM 2013 Preface Pages**
- RESEARCH SESSION 1
- 212  **SuReQL: A Subgraph Match Based Relational Model for Large Graphs**  
*(Shijie Zhang, Jiong Yang, Boya Sun)*
- 216  **Ranking Outlier Nodes in Subspaces of Attributed Graphs**  
*(Emmanuel Müller, Patricia Iglesias Sánchez, Yvonne Mülle, Klemens Böhm)*
- 223  **Chordless Cycles in Networks**  
*(John L. Pfaltz)*
- RESEARCH SESSION 2
- 229  **PSOGD: A New Method for Graph Drawing**  
*(Jianhua Qu, Yi Song, Stéphane Bressan)*
- 236  **Clustering Remote RDF Data Using SPARQL Update Queries**  
*(Letao Qi, Harris T. Lin, Vasant Honavar)*

---

## Workshop 7: Data Management in the Cloud (DMC)

---

(i)  **DMC 2013 Preface Pages**

### PAPER SESSION 1

243  **HotROD: Managing Grid Storage with On-Demand Replication**

*(Sriram Rao, Benjamin Reed, Adam Silberstein)*


250  **Materialized Views for Eventually Consistent Record Stores**

*(Changjiu Jin, Rui Liu, Kenneth Salem)*

258  **Packing Light: Portable Workload Performance Prediction for the Cloud**

*(Jennie Duggan, Yun Chi, Hakan Hacigümüs, Shenghuo Zhu, Ugur Çetintemel)*

### PAPER SESSION 2

266  **P-Mine: Parallel Itemset Mining on Large Datasets**

*(Elena Baralis, Tania Cerquitelli, Silvia Chiusano, Alberto Grand)*

272  **Towards Dynamic Pricing-Based Collaborative Optimizations for Green Data Centers**

*(Yang Li, David Chiu, Changbin Liu, Linh T.X. Phan, Tanveer Gill, Sanchit Aggarwal, Zhuoyao Zhang, Boon Thau Loo, David Maier, Bart McManus)*

279  **ISP Business Models in Caching**

*(Jörn Künsemöller, Nan Zhang, João Soares)*

---

## ICDE-13 PhD Symposium

---

(i)  **PhD Symposium Preface Pages**

### IT'S ALL ABOUT DATA

286  **Taming the Metadata Mess**

*(V.M. Megler)*


290  **High Quality Information Provisioning and Data Pricing**

*(Florian Stahl)*


294  **A Framework of Ontology Guided Data Linkage for Evidence Based Knowledge Extraction and Information Sharing**

*(Mohammed Gollapalli)*


### IT'S NOW ABOUT DATABASE (AND OTHER) SYSTEMS

298  **On Answering Why and Why-Not Questions in Databases**

*(Md. Saiful Islam)*

302  **Towards Elastic Key-Value Stores on IaaS**

*(Han Li)*


306  **User-Oriented Modelling of Scientific Workflows for High Frequency Event Data Analysis**

*(Aarthi Natarajan)*

### SHORT PRESENTATIONS





310  **Self-Organizing Structured RDF in MonetDB**

*(Minh-Duc Pham)*

314  **E-Research Event Data Quality**

*(Weisi Chen)*

*ICDE-13 PhD Symposium Short Presentations continued ...*

- 318  **Indexing and Querying Moving Objects in Indoor Spaces**  
*(Sultan Alamri)*
- 322  **Stock Prediction by Searching Similar Candlestick Charts**  
*(Zen-Yu Quan)*
- 326  **News Recommendation Based on Web Usage and Web Content Mining**  
*(Husna Sarirah Husin)*
- 330  **Making the H-Index More Relevant: A Step Towards Standard Classes for Citation Classification**  
*(Mohammad Abdullatif)*