

2016 IEEE 32nd International Conference on Data Engineering Workshops (ICDEW 2016)

**Helsinki, Finland
16 – 20 May 2016**



**IEEE Catalog Number: CFP1645A-POD
ISBN: 978-1-5090-2110-9**

**Copyright © 2016 by the Institute of Electrical and Electronics Engineers, Inc
All Rights Reserved**

Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

******This publication is a representation of what appears in the IEEE Digital Libraries. Some format issues inherent in the e-media version may also appear in this print version.***

IEEE Catalog Number:	CFP1645A-POD
ISBN (Print-On-Demand):	978-1-5090-2110-9
ISBN (Online):	978-1-5090-2109-3

Additional Copies of This Publication Are Available From:

Curran Associates, Inc
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: (845) 758-0400
Fax: (845) 758-2633
E-mail: curran@proceedings.com
Web: www.proceedings.com

CURRAN ASSOCIATES INC.
proceedings
.com

CloudDM 2016 — Workshop on Cloud Data Management

- 1 **CloudDM 2016 Preface Pages**
- 5 **Efficient Skew Handling in Online Aggregation in the Cloud**
(Xiang Ci, Fengming Wang, Yantao Gan, Xiaofeng Meng)
- 12 **High Variety Cloud Databases**
(Shrainik Jain, Dominik Moritz, Bill Howe)
- 20 **A Comparison of Flashcache with IQ-Twemcached**
(Yazeed Alabdulkarim, Marwan Almaymoni, Ziwen Cao, Shahram Ghandeharizadeh, Hieu Nguyen, Lingnan Song)
- 27 **Workload Management for Cloud Databases via Machine Learning**
(Ryan Marcus, Olga Papaemmanouil)
- 31 **Aggressive Buffer Pool Warm-Up After Restart in SQL Server**
(Kwanghyun Park, Jaeyoung Do, Nikhil Teletia, Jignesh M. Patel)
- 39 **IncReStore: Incremental Computation of MapReduce Workflows**
(Ahmed E. Khalifa, Iman Elghandour, Nagwa El-Makky)
-

HDMM 2016 — 1st International Workshop on Health Data Management and Mining

- 47 **HDMM 2016 Preface Pages**
- 50 **A Smartphone App Provides Preventive Care for the Elderly with Dementia**
(Hideyasu Sasaki)
- 54 **Air Quality Assessment from Social Media and Structured Data: Pollutants and Health Impacts in Urban Planning**
(Xu Du, Onyeka Emebo, Aparna Varde, Niket Tandon, Sreyasi Nag Chowdhury, Gerhard Weikum)
- 60 **Data Mining for Better Healthcare: A Path Towards Automated Data Analysis?**
(Tania Cerquitelli, Elena Baralis, Lia Morra, Silvia Chiusano)
- 64 **Timeline and Episode-Structured Clinical Data: Pre-Processing for Data Mining and Analytics**
(Jing Lu, Alan Hales, David Rew, Malcolm Keech)
- 68 **Mining Internet Media for Monitoring Changes of Public Emotions About Infectious Diseases**
(Sungwoon Choi, Jangho Lee, Sangheon Park, Yoon-Seok Chang, Sungroh Yoon)
- 71 **Visual Exploration and Cohort Identification of Acute Patient Histories Aggregated from Heterogeneous Sources**
(Rune Sætre, Øystein Nytrø, Stein Jakob Nordbø, Aslak Steinsbekk)
- 78 **A Sensor Network Query Processing System for Healthcare Data Acquisition in Dr. M**
(Chanhee Lee, Taehun Kim, Soon J. Hyun)
- 84 **Pathology Text Mining — On Norwegian Prostate Cancer Reports**
(Anders Dahl, Atilla Özkan, Hercules Dalianis)
- 88 **A Novel Approach to Estimate Maternal Mortality Rate and its Determinants by Statistical Modelling**
(Arun Sundararaman, Srinivasan Valady Ramanathan)

DESWeb 2016 — 7th International Workshop on Data Engineering meets the Semantic Web

- 94 **DESWeb 2016 Preface Pages**
- 97 **Old Techniques for New Join Algorithms: A Case Study in RDF Processing**
(Christopher R. Aberger, Susan Tu, Kunle Olukotun, Christopher Ré)
- 103 **Profiling Similarity Links in Linked Open Data**
(Blerina Spahiu, Cheng Xie, Anisa Rula, Andrea Maurino, Hongming Cai)
- 109 **Including Hierarchical Navigation in a Graph Database Query Language with an OBDA Approach**
(Nicolle Chaves Cysneiros, Ana Carolina Salgado)
- 115 **Using Analogy Computing for Ontology Development**
(Valentina Nejkovic, Milorad Tosic)
- 121 **From User Graph to Topics Graph: Towards Twitter Followee Recommendation Based on Knowledge Graphs**
(Danae Pla Karidi)
- 124 **Enabling Persistent Identification of Groups of Duplicates in Data Aggregators**
(Giorgos Alexiou, Marios Meimaris, George Papastefanatos)
-
-

HardBD 2016 — International Workshop on Big Data Management on Emerging Hardware

- 127 **HardBD 2016 Preface Pages**
- 130 **HyPer Beyond Software: Exploiting Modern Hardware for Main-Memory Database Systems (Keynote Paper)**
(Alfons Kemper)
- 131 **Towards a Distributed Multi-Tier File System for Cluster Computing**
(Herodotos Herodotou)
- 135 **Towards a GPU Accelerated Spatial Computing Framework**
(Harshada Chavan, Rami Alghamdi, Mohamed F. Mokbel)
- 143 **Exploiting SIMD for Complex Numerical Predicates**
(Dongxiao Song, Shimin Chen)

KEYS 2016 — Workshop on Keyword Search and Data Exploration on Structured Data

150	KEYS 2016 Preface Pages
153	Database Exploration: Problems and Opportunities (Keynote Paper) <i>(Giansalvatore Mecca)</i>
157	ORA-Semantics Based Keyword Search in XML and Relational Databases (Keynote Paper) <i>(Tok Wang Ling, Zhong Zeng, Thuy Ngoc Le, Mong Li Lee)</i>
161	Resourceful Interaction in Information Discovery (Keynote Paper) <i>(Giulio Jacucci)</i>
165	Querying and Mining Geo-Textual Data for Exploration: Challenges and Opportunities (Keynote Paper) <i>(Gao Cong, Kaiyu Feng, Kaiqi Zhao)</i>
169	Parallelizing Multiple Keyword Queries Over XML Streams <i>(Evandrino G. Barros, Fernando G.D.C. Ferreira, Alberto H.F. Laender)</i>
173	Playing LEGO with JSON: Probabilistic Joins Over Attribute-Value Fragments <i>(Manuel Hoffmann, Evica Milchevski, Sebastian Michel)</i>
181	Computing Similar Entity Rankings via Reverse Engineering of Top-k Database Queries <i>(Kiril Panev, Evica Milchevski, Sebastian Michel)</i>
189	A Novel Hierarchical Clustering Framework for Complex Symbolic Data Exploration <i>(Xin Xu)</i>

ICDE PhD Symposium 2016

193	ICDE PhD Symposium 2016 Preface Pages
196	Data Series Management: The Next Challenge (Keynote Paper) <i>(Themis Palpanas)</i>
200	Data Quality Between Promises and Results (Keynote Paper) <i>(Paolo Papotti)</i>
201	On a Self-Tuning Index Recommendation Approach for Databases <i>(Parinaz Ameri)</i>
206	Utilizing User Feedback to Improve Data Integration Systems <i>(Ahmed El-Roby)</i>
	Eventual Causal Consistency for Distributed NoSQL Databases <i>(Mawahib Musa Elbushra)</i>
216	Goal-Aware Data Management for Retrieval and Recommendations <i>(Dimitra Papadimitriou)</i>
221	Why-Query Support in Graph Databases <i>(Elena Vasilyeva)</i>
226	Robust Query Processing <i>(Srinivas Karthik)</i>
231	A Minimally-Intrusive Approach for Query-Driven Data Integration Systems <i>(Andreas M. Wahl)</i>