



DBKDA 2012

The Fourth International Conference on Advances in Databases, Knowledge,
and Data Applications

February 29 - March 5, 2012

Saint Gilles, Reunion Island

DBKDA 2012 Editors

Friedrich Laux, Reutlingen University, Germany

Gledson Elias, Federal University of Paraíba, Brazil

Chris Ireland, Open University, UK

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2012) by International Academy, Research, and Industry Association (IARIA)
Please refer to the Copyright Information page.

Printed by Curran Associates, Inc. (2012)

International Academy, Research, and Industry Association (IARIA)
412 Derby Way
Wilmington, DE 19810

Phone: (408) 893-6407
Fax: (408) 527-6351

petre@iaria.org

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2634
Email: curran@proceedings.com
Web: www.proceedings.com

TABLE OF CONTENTS

DBKDA 1: ADVANCES IN FUNDAMENTALS ON DATABASES

Reasoning about Domain Semantics over Relations, Bags, Partial Relations and Partial Bags	1
<i>Sebastian Link</i>	
Distinguishing Soft-Goals and Quality Requirements in Software Requirements Modeling	9
<i>Thi-Thuy-Hang Hoang, Alain Pirotte</i>	
A Case for Inverted Indices in In-Memory Databases	15
<i>Jens Krueger, Johannes Wust, Martin Faust, Tim Berning, Hasso Plattner</i>	
Sparse Dictionaries for In-Memory Column Stores	25
<i>Johannes Wust, Jens Krueger, Martin Grund, Uwe Hartmann, Hasso Plattner</i>	

DBKDA 2: KNOWLEDGE AND DECISION BASES

BI: Lean Manufacturing Indicators Applied to HR Applications - An Implementation Study	34
<i>Robson Poffo, Mehran Misaghi</i>	
Improving the Quality of Knowledge Discovery Process by Information Gain Computing	38
<i>Dumitru Dan Burdescu, Marian Cristian Mihaescu, Bogdan Logofatu, Costel Marian Ionascu</i>	
Education Cluster for Intelligent Provision of eLearning Services	45
<i>Stanimir Stoyanov, Emil Doychev, Veselina Valkanova, Georgi Cholakov</i>	
Towards Accurate Electricity Load Forecasting in Smart Grids	51
<i>Zeyar Aung, Mohamed Toukhy, John Williams, Abel Sanchez, Sergio Herrero</i>	

DBKDA 3: DATA QUERY, ACCESS, MINING, AND CORRELATION I

Operating on Hierarchical Enterprise Data in an In-Memory Column Store	58
<i>Christian Tinnefeld, Bjoern Wagner, Hasso Plattner</i>	
About Top-k Flexible Queries in Large Databases	64
<i>Khaoula Mabrouki, Amel Grissa Touzi, Habib Ounalli</i>	
Association Rule Mining from Large and Heterogeneous Databases with Uncertain Data Using Genetic Network Programming	74
<i>Eloy Gonzales, Koji Zettsu</i>	
A Multidimensional Data Modeling of the SEER Database from the USA National Cancer Institute	81
<i>Heidy M. Marin-Castro, Jose Torres-Jimenez, Diana I. Escalona-Vargas</i>	

DBKDA 4: DATA QUERY, ACCESS, MINING, AND CORRELATION II

Constructing a Synthetic Longitudinal Health Dataset for Data Mining	86
<i>Shima Ghassem Pour, Anthony Maeder, Louisa Jorm</i>	
Business Lead Generation for Online Real Estate Services: A Case Study	91
<i>Md. Rahman, Xinghui Zhao, Maria Mosquera, Qigang Gao, Vlado Keselj, Hai Wang</i>	
On Parallel Evaluation of SPARQL Queries	97
<i>Zbynek Falt, David Bednarek, Miroslav Cermak, Filip Zavoral</i>	
A New Approach For Top-k Flexible Queries In Large Database Using The Knowledge Discovered	103
<i>Amel Grissa Touzi, Habib Ounalli</i>	

DBKDA 5: DATA QUERY, ACCESS, MINING, AND CORRELATION III

A Knowledge Management Methodology for Studying Health Science Students' Development of Misconceptions	112
<i>Meaghen Regts, Arturo Fernandez Espinosa, Miguel Vargas Martin, Jayshiro Tashiro</i>	
Prediction Model Based on User Profile and Partial Course Progress for a Digital Media Learning Environment	120
<i>Arturo Fernandez Espinosa, Meaghen Regts, Jayshiro Tashiro, Miguel Vargas Martin</i>	

Clustering Large-Scale, Distributed Software Component Repositories	124
<i>Marcos Paulo Paixão, Leila Silva, Gledson Elias</i>	
OLAP Authentication and Authorization via Query Re-writing	130
<i>Todd Eavis, Ahmad Altamimi</i>	

DBKDA 6: DATABASES TECHNOLOGIES

Basic Components for Building Column Store-based Applications	140
<i>Andreas Schmidt, Daniel Kimmig</i>	
Leveraging Compression in In-Memory Databases	147
<i>Jens Krueger, Johannes Wust, Martin Linkhorst, Hasso Plattner</i>	
A System Overview for netCDF-FastBit Integration	154
<i>David Marks, Mahdi Abdelguerfi, Elias Ioup, John Sample, Kevin Shaw</i>	

DBKDA 7: DATA MANAGEMENT

The Multi-Tenant Data Placement Problem	157
<i>Jan Schaffner, Dean Jacobs, Tim Kraska, Hasso Plattner</i>	
Evaluation of Data Anonymization Tools	163
<i>Sergey Vinogradov, Alexander Pastsyak</i>	
Dynamic Stream Allocation with the Discrepancy between Data Access Time and CPU Usage Time	169
<i>Sayaka Akioka, Yoichi Muraoka, Hayato Yamana</i>	

DBKDA 8: DATA QUALITY

Relevance of Quality Criteria According to the Type of Information Systems	175
<i>Maria Del Pilar Angeles, Francisco Javier Garcia-Ugalde</i>	
Cloud-based Medical Image Collection Database with Automated Annotation	182
<i>Anthony Maeder, Birgit Planitz</i>	
Algebraic Constructs for Querying Provenance	187
<i>Murali Mani, Mohamad Alawa, Arunlal Kalyanasundaram</i>	
Author Index	