

ISSN 2178-7107

Volume 4, Number 3, October 2013

JIDM

Journal of Information & Data Management



Journal of Information and Data Management - JIDM

JIDM is an electronic publication focusing on information and data management in large repositories and document collections. It relates to different areas from Computer Science, including databases, information retrieval, digital libraries, knowledge discovery and data mining, geographic information systems, among others.

Editors

Marco A. Casanova, PUC-Rio, Brazil - Editor in chief
Cristina Dutra de Aguiar Ciferri, USP, Brazil - Editor in chief
Vanessa Braganholo, UFF, Brazil - Associate Editor

Editorial Board

Amr El Abbad - University of California, Santa Barbara, USA
Marcelo Arenas - PUC-Chile, Chile
Paolo Atzeni - Università Roma Tre, Italy
Ricardo Baeza-Yates - Yahoo! Research, Spain
Marco A. Casanova - PUC-Rio, Brazil
Silvana Castano - Università degli Studi di Milano, Italy
Umeshwar Dayal - HP Labs, USA
David W. Embley - Brigham Young University, USA
Alvaro Fernandes - University of Manchester, United Kingdom
Juliana Freire - University of Utah, USA
Theo Härder - University of Kaiserslautern, Germany
Carlos A. Heuser - UFRGS, Brazil
H.V. Jagadish - University of Michigan, USA
Tok-Wang Ling - National University of Singapore, Singapore
Marta Mattoso - UFRJ, Brazil
Claudia Bauzer Medeiros - UNICAMP, Brazil
René J. Miller - University of Toronto, Canada
John Mylopoulos - University of Trento, Italy
Mario Nascimento - University of Alberta, Canada
Tamer Özü - University of Waterloo, Canada
Oscar Pastor - Universidad Politécnica de Valencia, Spain
Alain Pirotte - Université Catholique de Louvain, Belgium
Gustavo Rossi - Universidad de La Plata, Argentina
Ana Carolina Salgado - UFPE, Brazil
Daniel Schwabe - PUC-Rio, Brazil
Altigran S. da Silva - UFAM, Brazil
Mário J. Silva - Universidade de Lisboa, Portugal
Caetano Traina Jr. - USP-SC, Brazil
Patrick Valduriez - INRIA, France
Gerhard Weikum - Max-Planck Institut für Informatik, Germany
Nivio Ziviani - UFMG, Brazil

Copyright ©2013 Permission to copy without fee all or part of the material printed in JIDM is granted provided that the copies are not made or distributed for commercial advantage, and that notice is given that copying is by permission of the Sociedade Brasileira de Computação.

Journal of Information and Data Management
Volume 4, Number 3, October 2013

Contents

Editorial	169
Automatic Web Page Segmentation and Noise Removal for Structured Extraction using Tag Path Sequences <i>Roberto Panerai Velloso, Carina F. Dorneles</i>	173
Evaluating the Diversification of Similarity Query Results <i>Lucio F. D. Santos, Willian D. Oliveira, Mônica R. P. Ferreira, Robson L. F. Cordeiro, Agma J. M. Traina, Caetano Traina Jr.</i>	188
Learning Accurate and Interpretable Classifiers Using Optimal Multi-Criteria Rules <i>Itamar Hata, Adriano Veloso, Nivio Ziviani</i>	204
ObInject: a NoODMG Persistence and Indexing Framework for Object Injection <i>Luiz O. Carvalho, Thatyana F. P. Seraphim, Caetano Traina Júnior, Enzo Seraphim</i>	220
A Method for Building Personalized Ontology Summaries <i>Paulo Orlando Queiroz-Sousa, Ana Carolina Salgado, Carlos Eduardo Pires</i>	236
Finding Missing Cross-Language Links in Wikipedia <i>Carlos Eduardo M. Moreira, Viviane P. Moreira</i>	251
Under Pressure Benchmark for DDBMS Availability <i>Alessandro Gustavo Fior, Jorge Augusto Meira, Eduardo Cunha de Almeida, Ricardo Gonçalves Coelho, Marcos Didonet Del Fabro, Yves Le Traon</i>	266
Stress Testing of Transactional Database Systems <i>Jorge Augusto Meira, Eduardo Cunha de Almeida, Gerson Sunyé, Yves Le Traon, Patrick Valduriez</i>	279
GUARD: A Genetic Unified Approach for Recommendation <i>Adolfo Guimarães, Thales F. Costa, Anísio Lacerda, Gisele L. Pappa, Nivio Ziviani</i>	295
Interaction-Aware Data Management in the Cloud <i>Manoel Siqueira, José Maria Monteiro, Angelo Brayner, Flávio R. C. Sousa, Javam C. Machado</i>	311
Finding Spatio-Temporal Patterns in Multidimensional Data Streams <i>Santiago A. Nunes, Luciana A. S. Romani, Ana M. H. Avila, Priscila P. Coltri, Agma J. M. Traina, Elaine P. M. Sousa</i>	327
A Holistic Hybrid Algorithm for User Recommendation on Twitter <i>Sara Guimarães, Marco Túlio Ribeiro, Renato Assunção, Wagner Meira Jr.</i>	341
Hierarchical Bottom-Up Safe Semi-Supervised Support Vector Machines for Multi-Class Transductive Learning <i>Thiago F. Covões, Rodrigo C. Barros, Tiago S. da Silva, Eduardo R. Hruschka, Andre C. P. L. F. de Carvalho</i>	357
SCM-BP: An Intelligent Buffer Management Mechanism for Database in Storage Class Memory <i>Júlio A. Tavares, José de Aguiar Moraes Filho, Angelo Brayner, Eduardo Lustosa</i>	374

OBAS: An OLAP Benchmark for Analysis Services	390
<i>Bruno Edson Martins de Albuquerque Filho, Thiago Luís Lopes Siqueira, Valéria Cesário Times</i>	
Metamodeling the Enhanced Entity-Relationship Model	406
<i>Robson N. Fidalgo, Edson Alves, Sergio España, Jaelson Castro, Oscar Pastor</i>	
A Relevance Measure for Multivalued Attributes	421
<i>Mariana Tasca, Bianca Zadrozny, Alexandre Plastino</i>	
cx-Sim: A Metric Access Method for Similarity Queries with Additional Conditions	437
<i>Leandro C. Soares, Daniel S. Kaster</i>	
A Comparative Study of Learning-to-Rank Techniques for Tag Recommendation	453
<i>Sérgio D. Canuto, Fabiano M. Belém, Jussara M. Almeida, Marcos A. Gonçalves</i>	
qCube: Efficient Integration of Range Query Operators over a High Dimension Data Cube	469
<i>Rodrigo Rocha Silva, Joubert de Castro Lima, Celso Massaki Hirata</i>	
Towards a Statistical Evaluation of PigLatin Joins	483
<i>Renato Javier M. Mogrovejo, José Maria Monteiro, Javam C. Machado, Carlos Juliano M. Viana, Sérgio Lifschitz</i>	