

ISSN 2178-7107

Volume 5, Number 1, February 2014

JIDM

Journal of Information & Data Management

Official Publication of the Brazilian Computer Society Special Interest Group on Databases

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

Ana Carolina Salgado, UFPE, Brazil - Editor in chief

Vanessa Braganholo, UFF, Brazil - Associate Editor

Editorial Board

Amr El Abbadi - University of California, Santa Barbara, USA

Alvaro Fernandes - University of Manchester, United Kingdom

Alberto H. F. Laender - UFMG, Brazil

Marta Mattoso - UFRJ, Brazil

Mario Nascimento - University of Alberta, Canada

Oscar Pastor - Universidad Politécnica de Valencia, Spain

Alain Pirotte - Université Catholique de Louvain, Belgium

Copyright ©2014 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 5, Number 1, February 2014

Contents

Editorial	1
SBBD 2013 Short Papers	
FPSMining: A Fast Algorithm for Mining User Preferences in Data Streams <i>Jaqueline A. J. Papini, Sandra de Amo, Allan Kardec S. Soares</i>	4
SGProv: Summarization Mechanism for Multiple Provenance Graphs <i>Daniele El-Jaick, Marta Mattoso, Alexandre A. B. Lima</i>	16
ARe-SQL: An Online, Automatic and Non-Intrusive Approach for Rewriting SQL Queries <i>Arlino H. M. de Araújo, José Maria Monteiro, José Antônio F. de Macêdo, Júlio A. Tavares, Angelo Brayner, Sérgio Lifschitz</i>	28
Efficient Entity Matching over Multiple Data Sources with MapReduce <i>Demetrio Gomes Mestre, Carlos Eduardo Santos Pires</i>	40
Experimental Evaluation of Academic Collaboration Recommendation Using Factorial Design <i>Michele A. Brando, Mirella M. Moro, Jussara M. Almeida</i>	52
KDMiLe 2013	
Strategies for Mining User Preferences in a Data Stream Setting <i>Jaqueline A. J. Papini, Sandra de Amo, Allan Kardec S. Soares</i>	64
Similarity Search in Multidimensional Time Series using the Coulomb's Law <i>Claudinei Garcia de Andrade, Marcela Xavier Ribeiro</i>	74
Classifying High-Speed Data Streams Using Statistical Decision Trees <i>Mirela Teixeira Cazzolato, Marcela Xavier Ribeiro</i>	84
Predicting the Learning Rate of Gradient Descent for Accelerating Matrix Factorization <i>Caio B. Nóbrega, Leandro B. Marinho</i>	94
A Fuzzy Decision Tree Model to Support the Task of Bus Reallocation in Public Transport Systems <i>Marcos Evandro Cintra, José Francisco Ferreira Ribeiro, Odacir Almeida Neves</i>	104
Unsupervised Instance Selection from Text Streams <i>Rafael Bonin, Ricardo M Marcacini, Solange O Rezende</i>	114
Exploring Attribute Selection in Hierarchical Classification <i>Bruno C. Paes, Alexandre Plastino, Alex A. Freitas</i>	124
A Constructive Density-Ratio Approach to Mutual Information Estimation: experiments in feature selection <i>Igor Braga</i>	134