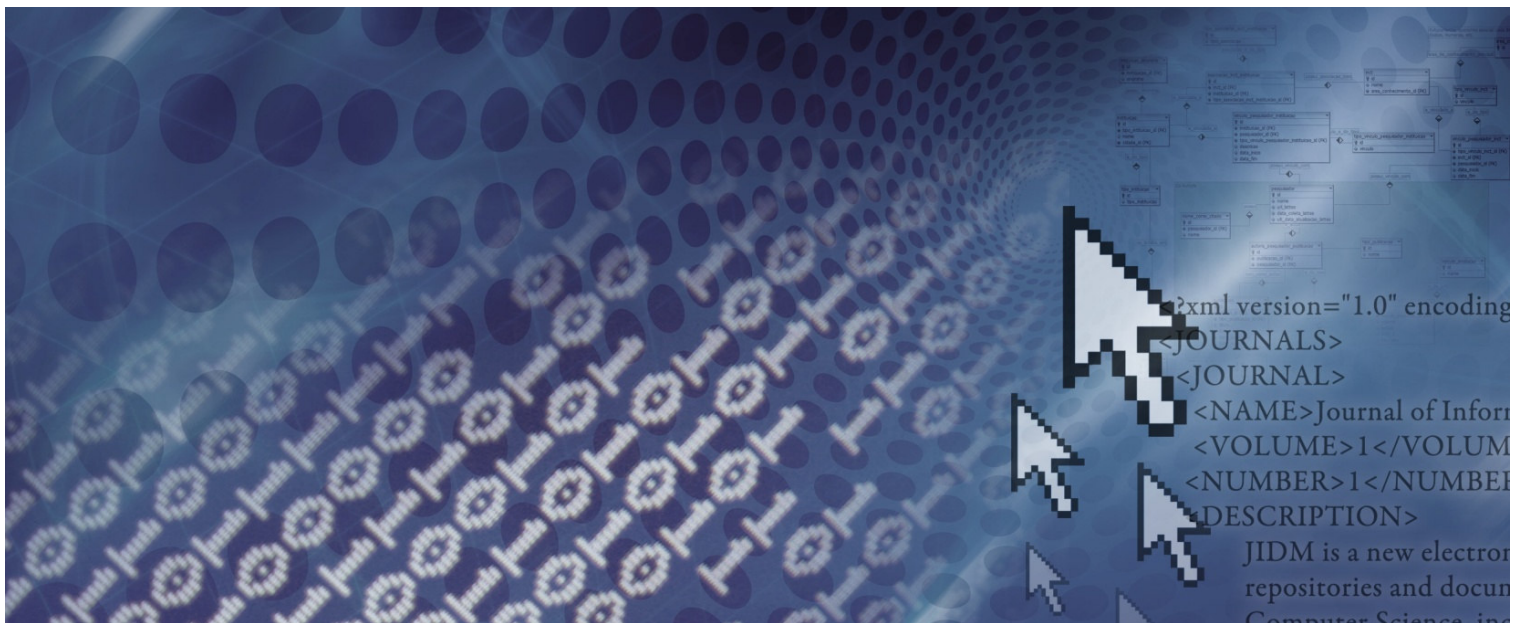


ISSN 2178-7107

Volume 1, Number 3, October 2010

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

Alberto H. F. Laender, UFMG, Brazil - Editor in chief

Mirella M. Moro, UFMG, Brazil - Associate Editor

Editorial Board

Amr El Abbadi - 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

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 Oszü - 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

Nívio Ziviani - UFMG, Brazil

Journal of Information and Data Management
Volume 1, Number 3, October 2010

Contents

Editorial	309
Web Applications	
A Multi-view Approach for Detecting Non-Cooperative Users in Online Video Sharing Systems <i>Hendrickson Reiter Langbehn, Saulo Marques Ribeiro Ricci, Marcos André Gonçalves, Jussara Marques Almeida, Gisele Lobo Pappa, Fabrício Benevenuto</i>	313
A Context-Dependent Supervised Learning Approach to Sentiment Detection in Large Textual Databases <i>Albert Weichselbraun, Stefan Gindl, Arno Scharl</i>	329
Indexing Web Form Constraints <i>Ronaldo dos Santos Mello, Ramesh Pinnamaneni, Juliana Freire</i>	343
Approximate and Similarity Search	
An Adaptive Multi-level Hashing Structure for Fast Approximate Similarity Search <i>Alexander Ocsa, Elaine Sousa</i>	359
DAHC-tree: An Effective Index for Approximate Search in High-Dimensional Metric Spaces <i>Jurandy Almeida, Eduardo Valle, Ricardo da S. Torres, Neucimar J. Leite</i>	375
Adding Knowledge Extracted by Association Rules into Similarity Queries <i>Mônica Ribeiro Porto Ferreira, Marcela Xavier Ribeiro, Agma Juci Machado Traina, Richard Chbeir, Caetano Traina Jr.</i>	391
Data Mining	
RDBMS as an Efficient Tool to Mine Cliques on Complex Networks <i>Ana Paula Appel, Adriano Arantes Paterlini, Caetano Traina Junior</i>	407
HiSP-GC: A Classification Method Based on Probabilistic Analysis of Patterns <i>Luiz Henrique de Campos Merschmann, Alexandre Plastino</i>	423
Estimating the Credibility of Examples in Automatic Document Classification <i>Thiago Salles, João Palotti, Gisele Pappa, Filipe Arcanjo, Marcos Gonçalves, Wagner Meira</i>	439
XML	
Processing Queries over Distributed XML Databases <i>Guilherme Figueiredo, Vanessa Braganholo, Marta Mattoso</i>	455
Supporting Temporal Queries on XML Keyword Search Engines <i>Edimar Manica, Carina F. Dorneles, Renata Galante</i>	471
An Evaluation Study of Search Algorithms for XML Streams <i>Evandrino Gomes Barros, Mirella M. Moro, Alberto H. F. Laender</i>	487

Core Databases

Evaluation of Conditional Preference Queries 503
Fabiola Souza Fernandes Pereira, Sandra de Amo

How does the spatial data redundancy affect query performance in geographic data warehouses? 519
Rodrigo Costa Mateus, Thiago Luís Lopes Siqueira, Valéria Cesário Times, Ricardo Rodrigues Ciferri, Cristina Dutra de Aguiar Ciferri

Towards Automatic Generation of Application Ontologies 535
Eveline Russo Sacramento, Vânia M. P. Vidal, José Antonio F. de Macêdo, Bernadette F. Lóscio, Fernanda Lígia R. Lopes, Marco A. Casanova

Information Retrieval

WCL2R: A Benchmark Collection for Learning to Rank Research with Clickthrough Data 551
Otávio Domingues Alencar Alcântara, Marcos André Gonçalves, Humberto Mossri Almeida, Alvaro Pereira, Ricardo Baeza-Yates, Christian Middleton

Learning to Rank using Query-Level Rules 567
Adriano Veloso, Marcos Gonçalves, Wagner Meira Jr., Humberto Mossri

Using Statistical Features to Find Phrasal Terms in Text Collections 583
André Luiz da Costa Carvalho, Edleno Silva de Moura, Pável Calado