

ISSN 2178-7107

Volume 7, Number 2, August 2016

# 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**

**Caetano Traina Junior**, Universidade de São Paulo, Brazil - Editor in chief  
**Robson L. F. Cordeiro**, Universidade de São Paulo, Brazil - Associate Editor

## **Editorial Board**

**Alberto Laender** - Universidade Federal de Minas Gerais, Brazil  
**Altigran Silva** - Universidade Federal do Amazonas, Brazil  
**Alvaro Fernandes** - University of Manchester, United Kingdom  
**Ana Carolina Salgado** - Universidade Federal de Pernambuco, Brazil  
**Angelo Brayner** - Universidade de Fortaleza, Brazil  
**Daniel Keim** - Konstanz University, Germany  
**Divesh Srivastava** - AT&T LabsResearch, USA  
**Eamonn Keogh** - University of California, USA  
**Edgar Leonel Chávez González** - Centro de Investigación Científica y de Educación Superior de Ensenada, Mexico  
**Gerhard Weikum** - Max Planck Institute for InformaticsSaarbruecken, Germany  
**Javam de Castro Machado** - Universidade Federal do Ceará, Brazil  
**João Eduardo Ferreira** - Universidade de São Paulo, Brazil  
**Juliana Freire** - New York University, USA  
**Marco Antonio Casanova** - Pontifícia Universidade Católica do Rio de Janeiro, Brazil  
**Marcos André Gonçalves** - Universidade Federal de Minas Gerais, Brazil  
**Maria Luisa Sapino** - Università degli Studi di Torino, Italy  
**Marta Lima de Queiros Mattoso** - Universidade Federal do Rio de Janeiro, Brazil  
**Mirella Moura Moro** - Universidade Federal de Minas Gerais, Brazil  
**Oscar Pastor** - Universidad Politécnica de Valencia, Spain  
**Renata de Matos Galante** - Universidade Federal do Rio Grande do Sul, Brazil  
**Renato Fileto** - Universidade Federal de Santa Catarina, Brazil  
**Ricardo da Silva Torres** - Universidade Estadual de Campinas, Brazil  
**Sandra de Amo** - Universidade Federal de Uberlândia, Brazil  
**Sang-Wook Kim** - Hanyang University, Korea  
**Selçuk Candan** - Arizona State University, USA  
**Sergio Lifschitz** - Pontifícia Universidade Católica do Rio de Janeiro, Brazil  
**Spiros Papadimitriou** - Rutgers University, USA  
**Tok Wang Ling** - National University of Singapore, Singapore  
**Vassilis J. Tsotras** - University of California, USA  
**Vanessa Braganholo** - Universidade Federal Fluminense, Brazil  
**Vincent Oriá** - New Jersey Institute of Technology, USA  
**Wagner Meira Júnior** - Universidade Federal de Minas Gerais, Brazil  
**Yannis Manolopoulos** - Aristotle University of Thessaloniki, Greece  
**Yasin N. Silva** - Arizona State University, USA  
**Zhongfei (Mark) Zhang** - Binghamton University, USA

Copyright ©2016 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 7, Number 2, August 2016

Contents

Editorial	68
ETL4LinkedProv: Managing Multigranular Linked Data Provenance <i>Rogers Reiche de Mendonça, Sérgio Manuel Serra da Cruz, Maria Luiza Machado Campos</i>	70
Improving Pairwise Preference Mining Algorithms Using Preference Degrees <i>Juliete A. Ramos Costa, Sandra de Amo</i>	86
<b>KDMiLe 2015</b>	
Exploiting Social Information in Pairwise Preference Recommender System <i>Crícia Z. Felício, Klérisson V. R. Paixão, Guilherme Alves, Sandra de Amo, Philippe Preux</i>	99
Data Preparation for Longitudinal Data Mining: a case study on human ageing <i>Caio Eduardo Ribeiro, Luis Enrique Zárate</i>	116
From Sensor Data Streams to Linked Streaming Data: a survey of main approaches <i>Kathrin Rodriguez Llanes, Marco Antonio Casanova, Noel Moreno Lemus</i>	130
Aspect-based Sentiment Analysis using Semi-supervised Learning in Bipartite Heterogeneous Networks <i>Ivone P. Matsuno, Rafael G. Rossi, Ricardo M. Marcacini, Solange O. Rezende</i>	141
A Fast and Effective Strategy for Feature Selection in High-dimensional Datasets <i>Mariana Tasca, Alexandre Plastino, Celso Ribeiro, Bianca Zadrozny</i>	155
Exploiting Academic Records for Predicting Student Drop Out: a case study in Brazilian higher education <i>Allan Sales, Leandro Balby, Adalberto Cajueiro</i>	166
Better Initialization Heuristics for Order-based Bayesian Network Structure Learning <i>Walter Perez Urcia, Denis Deratani Mauá</i>	181