

ISSN 2178-7107

Volume 13, Number 1, August 2022



# 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

**Maristela Holanda**, Universidade de Brasília, Brazil - Editor in chief

**Carina F. Dorneles** - Universidade Federal de Santa Catarina, Brazil

**Eric Araújo** - Universidade Federal de Lavras, Brazil

## SBBD-DSW 2021 Guest Editorial Board

**André Santanchè** - Universidade Estadual de Campinas, Brazil

**Clodoveu Davis** - Universidade Federal de Minas Gerais, Brazil

**Daniel de Oliveira** - Universidade Federal Fluminense, Brazil

**Daniel Hasan Dalip** - CEFET/MG, Brazil

**Daniel Lichtnow** - Universidade Federal de Santa Maria, Brazil

**Denilson Alves Pereira** - Universidade Federal de Lavras, Brazil

**Eduardo Freire Nakamura** - Universidade Federal do Amazonas, Brazil

**Eduardo Ogasawara** - CEFET/MG, Brazil

**Giseli Rabello Lopes** - Universidade Federal do Rio de Janeiro, Brazil

**Guilherme Domingos Faria Silva** - Universidade de São Paulo, Brazil

**Helena Graziottin Ribeiro** - Universidade de Caxias do Sul, Brazil

**José Palazzo Moreira de Oliveira** - Universidade Federal do Rio Grande do Sul, Brazil

**Luiz Henrique de Campos Merschmann** - Universidade Federal de Lavras, Brazil

**Michele A. Brandão** - Universidade Federal de Minas Gerais, Brazil

**Ronaldo dos Santos Mello** - Universidade Federal de Santa Catarina, Brazil

**Victor Ströele** - Universidade Federal de Juiz de Fora, Brazil

# Journal of Information and Data Management

## Volume 13, Number 1, August 2022

### Contents

Editorial	5
<b>SBBD-DSW 2021 papers</b>	
Covid Data Analytics Repository: An interdisciplinary look into the COVID-19 pandemic in Brazil <i>Ramon A. S. Franco, Pedro Loures Alzamora, Janaina Guiginski, Evandro L. T. P. Cunha, Tereza Bernardes, Juan F. Galindo, Luana Passos, Raquel Schneider, Bruno Chagas, Kícila Ferregetti, Luísa Cardoso, Pedro Moreira, Wallace Pereira, Ana Paula Couto da Silva, Wagner Meira Jr.</i>	6
Collecting, extracting and storing web research survey questionnaires data <i>Carina F. Dorneles, Gilney N. Mathias</i>	20
FeatSet+: Visual Features Extracted from Public Image Datasets <i>Mirela T. Cazzolato, Lucas C. Scabora, Guilherme F. Zobot, Marco A. Gutierrez, Caetano Traina Jr., Agma J. M. Traina</i>	33
Using Musical and Statistical Analysis of the Predominant Melody of the Voice to Create datasets from a Database of Popular Brazilian Hit Songs <i>André A. Bertoni and Rodrigo P. Lemos</i>	49
Essay-BR: a Brazilian Corpus to Automatic Essay Scoring Task <i>Jeziel C. Marinho, Rafael T. Anchiêta, Raimundo S. Moura</i>	65
BovDB: a data set of stock prices of all companies in B3 from 1995 to 2020 <i>Fabian Corrêa Cardoso, Juan Andrey Valverde Malska, Paulo Junior Ramiro, Giancarlo Lucca, Eduardo Nunes Borges, Viviane Leite Dias de Mattos, Rafael Alceste Berri</i>	77
Cross-collection Dataset of Public Domain Portuguese-language Works <i>Mariana O. Silva, Clarisse Scofield, Luiza de Melo-Gomes, Mirella M. Moro</i>	95
Musical Success in the United States and Brazil: Novel Datasets and Temporal Analyses <i>Gabriel P. Oliveira, Gabriel R. G. Barbosa, Bruna C. Melo, Juliana E. Botelho, Mariana O. Silva, Danilo B. Seufitelli, Mirella M. Moro</i>	111
A Comprehensive Dataset of Brazilian Fact-Checking Stories <i>Igor Marques, Isadora Salles, João M. M. Couto, Breno C. Pimenta, Samuel Assis, Julio C. S. Reis, Ana Paula C. da Silva, Jussara M. Almeida, Fabrício Benevenuto</i>	127
PolRoute-DS: a Crime Dataset for Optimization-based Police Patrol Routing <i>Bruno Cunha Sá, Gustavo Muller, Maicon Banni, Wagner Santos, Marcos Lage, Isabel Rosseti, Yuri Frota, Daniel de Oliveira</i>	143

- Curating, Enriching and FAIRifying Datasets from the Brazilian COVID-19 Vaccination 157  
*Marcus Vinicius Ferreira Gonçalves, Jamile Santos dos Santos, Caio Zava Ferreira, Jorge Zavaleta, Sérgio Manuel Serra da Cruz, Jonice Oliveira Sampaio*
- CandiDATA: an enhanced dataset for data analysis of elections in Brazil from 1945 to 2020 172  
*Felipe F. Vasconcelos, João V. S. Tavares, Matheus G. S. Oliveira, Fabio J. Coutinho, João Paulo Clarindo*