

1. Record Nr.	UNISA996485665303316
Titolo	New trends in database and information systems : ADBIS 2022 short papers, doctoral consortium and workshops: DOING, K-GALS, MADEISD, MegaData, SWODCH, Turin, Italy, September 5-8, 2022, proceedings / / Silvia Chiusano, [and six others], editors
Pubbl/distr/stampa	Cham, Switzerland : , : Springer, , [2022] ©2022
ISBN	3-031-15743-5
Descrizione fisica	1 online resource (675 pages)
Collana	Communications in computer and information science ; ; Volume 1652
Disciplina	006.3
Soggetti	Artificial intelligence Databases
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di contenuto	Intro -- Preface -- Organization -- Contents -- ADBIS Short Papers: Data Understanding, Modeling and Visualization -- On an Extensible Conceptual Data Model by a Non-formal Example -- 1 Introduction -- 2 Formal Bases -- 3 An Extensible Conceptual Data Model -- 3.1 Formalization of Hierarchical Relations -- 3.2 Conceptual Schema -- 4 Formal Semantics of Hierarchical Relations Algebra -- 5 The Conceptual Data Model Extensibility Concept -- 5.1 Non-formal Example of Extending the Conceptual Data Model -- 6 Related Work -- 7 Conclusions -- References -- Understanding Misinformation About COVID-19 in WhatsApp Messages -- 1 Introduction -- 2 Related Work -- 3 COVID-19.BR Data Set -- 4 Understanding Misinformation About COVID-19 Using LIME and LIWC -- 4.1 Applying LIME -- 4.2 Applying LIWC -- 5 Conclusion -- References -- Automatic Inference of Taxonomy Relationships Among Legal Documents -- 1 Introduction -- 2 Related Work -- 3 The Classification System -- 3.1 Preliminaries -- 3.2 Problem Statement -- 3.3 Text Representations -- 3.4 Classifiers -- 4 Experimental Results -- 4.1 Experimental Design -- 4.2 Results Discussion -- 5 Takeaways and Future Directions -- References -- Visualisation of Numerical Query Results on Industrial Data Streams -- 1 Introduction -- 2 State of the Art -- 3 The DISPLAY Operator -- 4

Experimental Validation -- 5 Conclusions and Contributions --
 References -- ADBIS Short Papers: Fairness in Data processing --
 Clustering-Based Subgroup Detection for Automated Fairness Analysis -- 1 Introduction -- 2 Related Work -- 3 AI Fairness -- 3.1 Subgroup Fairness Metrics -- 4 Automatic Subgroup Detection -- 5 Results -- 6 Conclusion and Future Work -- References -- Open Science in the Cloud: The CloudFAIR Architecture for FAIR-compliant Repositories -- 1 Introduction -- 2 The BigFAIR Architecture -- 3 The CloudFAIR Architecture. 4 Performance Evaluation -- 4.1 Queries Involving Different Storage Components -- 4.2 Effect of the Number of Data Providers -- 5 Conclusions and Future Work -- References -- Detecting Simpson's Paradox: A Step Towards Fairness in Machine Learning -- 1 Introduction -- 2 Yule-Simpson's Paradox -- 3 Detecting Simpson's Paradox -- 3.1 Algorithm -- 4 Experiments -- 4.1 Datasets -- 5 Discussion and Future Work -- 6 Conclusion -- References -- ADBIS Short Papers: Data Management Pipeline, Information and Process Retrieval -- FusionFlow: An Integrated System Workflow for Gene Fusion Detection in Genomic Samples -- 1 Introduction -- 2 The Workflow -- 2.1 Architecture -- 3 Workflow Test and Discussion -- 4 Conclusions -- References -- ProSA Pipeline: Provenance Conquers the Chase -- 1 Introduction -- 2 Basics and Related Work -- 3 The ProSA-Pipeline -- 3.1 The Theory Behind ProSA -- 3.2 The Steps of the ProSA Pipeline -- 3.3 Implementing ProSA -- 4 Conclusion -- References -- A Tabular Open Data Search Engine Based on Word Embeddings for Data Integration -- 1 Introduction -- 2 Related Work -- 3 Tabular Open Data Search and Integration with Word Embeddings -- 3.1 Crawler -- 3.2 Storage -- 3.3 Microservice -- 3.4 Web Application -- 4 Using Search! -- 5 Conclusions and Future Work -- References -- The Use of M2P in Business Process Improvement and Optimization -- 1 Introduction -- 2 Related Works -- 3 Preliminaries -- 3.1 Business Process Management (BPM) -- 3.2 Business Process Improvement -- 3.3 Business Process Modelling -- 3.4 Unified Modelling Language -- 3.5 Model Transformation -- 4 Contribution -- 4.1 Transformation Table -- 4.2 The Algorithm of Transformation -- 4.3 Insurance Claim Case Study -- 5 Conclusion and Future Works -- References -- Embedding Process Structure in Activities for Process Mapping and Comparison -- 1 Introduction -- 2 Related Work. 3 Methods -- 4 Case Study -- 4.1 Settings -- 4.2 Results -- 5 Conclusions -- References -- ADBIS Short Papers: Data Access Optimization -- Bulk Loading of the Secondary Index in LSM-Based Stores for Flash Memory -- 1 Introduction -- 2 Bulk Loading of the Secondary Index -- 2.1 LSM Tree -- 2.2 Bulk Loading Outline -- 2.3 Bulk Loading for Flash Memory -- 3 Proposed Solution -- 4 Experiments -- 5 Conclusions -- References -- Blink: Lightweight Sample Runs for Cost Optimization of Big Data Applications -- 1 Introduction -- 2 Related Work -- 3 Background -- 4 Efficient Sample Runs -- 4.1 Size of Sample Runs -- 4.2 Parallelism -- 4.3 Cluster Configuration -- 4.4 Number of Sample Runs -- 5 Blink -- 5.1 Sample Runs Manager -- 5.2 Data Size Predictor -- 5.3 Execution Memory Predictor -- 5.4 Cluster Size Selector -- 6 Evaluation -- 6.1 Selected Cluster Size -- 6.2 Overhead of Sample Runs -- 7 Conclusion -- References -- ADBIS Short Papers: Data Pre-processing and Cleaning -- A RoBERTa Based Approach for Address Validation -- 1 Introduction -- 2 Related Work -- 2.1 Address Parsing -- 2.2 Address Classification -- 3 Address Validation Approach -- 3.1 Address Standardization -- 3.2 Address Classification -- 4 Evaluation -- 4.1 Experimental Setting -- 4.2 Address Parsing Evaluation -- 4.3 Address Classification

Evaluation -- 5 Conclusion -- References -- Empirical Comparison of Semantic Similarity Measures for Technical Question Answering -- 1 Introduction -- 1.1 Document Retrieval -- 1.2 Technical Question Answering -- 1.3 Semantic Similarity Measures -- 2 Compared Methods -- 2.1 Keyword Extraction Based Methods -- 2.2 Word Vector Based Methods -- 2.3 Deep Learning Based Methods -- 3 Computational Comparison -- 3.1 Evaluation Data -- 3.2 Evaluation Measures -- 3.3 Results -- 4 Conclusion and Further Works -- References.

Denoising Architecture for Unsupervised Anomaly Detection in Time-Series -- 1 Introduction -- 2 Related Work -- 3 Unsupervised Time-Series Anomaly Detection -- 4 Anomaly Detection Using LSTM Autoencoder -- 5 The Denoising Architecture -- 6 Experimental Results -- 6.1 Datasets -- 6.2 Experimental Setup -- 6.3

Benchmarking Results -- 7 Conclusion -- References -- Discretizing

Numerical Attributes: An Analysis of Human Perceptions -- 1 Introduction -- 2 Related Work -- 3 Motivation -- 4 The LSQM Method -- 5 Experimental Design -- 6 Analysis and Result -- 7 Conclusion --

References -- ADBIS Short Papers: Modeling and Querying of Graph Databases -- Interactive Knowledge Graph Querying Through Examples and Facets -- 1 Introduction -- 2 Model -- 2.1 Knowledge Graphs --

2.2 Interaction Model -- 2.3 Multiple Semantics for User Interactions --

3 Optimizing the Interaction -- 3.1 Assuming Known Probabilities -- 3.2 Estimating the Probabilities -- 4 Related Work -- References -- It's Too Noisy in Here: Using Projection to Improve Differential Privacy on

RDF Graphs -- 1 Introduction -- 2 Background: Differential Privacy -- 2.1 Definition of Differential Privacy -- 2.2 Noise Calibration -- 2.3 Sensitivity and Privacy on (Bounded QL-Out-Degree) Graphs -- 3

Related Work -- 3.1 Models and Distances for DP on Graphs -- 3.2

Applying DP on Graphs -- 3.3 Privacy over RDF -- 4 Proposed

Approach: From a Subspace with Low-Sensitivity Queries to G -- 4.1

Proposed Projection Method -- 4.2 Privacy and Projections -- 4.3

Privacy on Unbounded Graphs Through Projection -- 5 Analytical and Experimental Evaluation -- 5.1 Considered Query and Interest of the Approach -- 5.2 Overall Utility of the Approach -- 6 Conclusion --

References -- Modeling and Querying Sensor Networks Using Temporal Graph Databases -- 1 Introduction and Related Work -- 2 Background and Preliminary Definitions.

3 Temporal Graphs for Sensor Networks -- 4 Temporal Paths in Sensor

Networks -- 5 Computing the Paths -- 6 Conclusion and Future Work

-- References -- Indexing Continuous Paths in Temporal Graphs -- 1

Introduction -- 2 Background -- 3 An Index for Continuous Paths:

TGIndex -- 4 Using the Index -- 5 Experimental Evaluation -- 6

Conclusion -- References -- ADBIS Short Papers: Data Science and Machine Learning -- Unsupervised Extractive Text Summarization

Using Frequency-Based Sentence Clustering -- 1 Introduction -- 2

Related Works -- 3 Proposal -- 3.1 Linguistic Pre-processing -- 3.2

Sentence Feature Vector Representation -- 3.3 Sentence Clustering --

3.4 Summary Building -- 4 Experimental Evaluation -- 4.1 Prototype

Implementation -- 4.2 Experimental Metrics -- 4.3 Experimental Data

-- 4.4 Experimental Results -- 5 Conclusion -- References --

Experimental Comparison of Theory-Guided Deep Learning Algorithms

-- 1 Introduction -- 2 Related Works -- 2.1 Knowledge Source -- 2.2

Knowledge Representation -- 2.3 Knowledge Integration -- 3 Datasets

and Methods -- 3.1 Lake Temperature -- 3.2 Convective Movements in

Climate Modeling -- 4 Experimental Results -- 5 Conclusions --

References -- Hyper-parameters Tuning of Artificial Neural Networks:

An Application in the Field of Recommender Systems -- 1 Introduction

-- 2 Related Work -- 3 Background -- 3.1 Activation Functions -- 3.2 Weight Initialization -- 4 Experiments -- 4.1 Datasets -- 4.2 Evaluation Metrics -- 4.3 CiteULike Experiments in Recommendation of Scientific Publications -- 4.4 AMiner Citation Network Experiments -- 5 Discussion of the Experimental Results -- 6 Conclusion and Future Work -- References -- Exploring Latent Space Using a Non-linear Dimensionality Reduction Algorithm for Style Transfer Application -- 1 Introduction -- 2 Related Work -- 3 Methodology -- 4 Experimental Validation.
4.1 The Dataset.
