

# **30th Italian Symposium on Advanced Database Systems (SEBD 2022)**

Pisa, Italy  
19-22 June 2022

## **Editors:**

**Giuseppe Amato  
Valentina Bartalesi  
Riccardo Torlone**

**Devis Bianchini  
Claudio Gennaro**

ISBN: 978-1-7138-6364-9

**Printed from e-media with permission by:**

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571



**Some format issues inherent in the e-media version may also appear in this print version.**

Copyright© for individual papers are retained by the authors and are licensed under a Creative Commons Attribution 4.0 International License.

License details: <http://creativecommons.org/licenses/by/4.0/>.

All rights reserved.

Printed by Curran Associates, Inc. (2023)

Edited by:

Giuseppe Amato  
Valentina Bartalesi  
Devis Bianchini  
Claudio Gennaro  
Riccardo Torlone

**Additional copies of this publication are available from:**

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571 USA  
Phone: 845-758-0400  
Fax: 845-758-2633  
Email: [curran@proceedings.com](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

# Table of Contents

- Preface of the 30th Italian Symposium on Advanced Database Systems  
Summary: There were 81 papers submitted for peer-review to SEBD 2022. Out of these, 70 papers were accepted for this volume, 10 as regular papers and 60 as discussion papers.

## Invited Talks

- Runtime-Optimized Analytics (Abstract) i  
*Anastasia Ailamaki*
- From Database Systems to Global Platforms: Cyberspace and the new Rules of Geopolitics (Abstract) ii  
*Stéphane Grumbach*

## Tutorial

- A Pragmatic Approach to Neural Information Retrieval (Abstract) iii  
*Franco Maria Nardini, Salvatore Trani*

## Taste of Time Award

- Techniques for the Analysis of Conceptual Schema (Extended Abstract) - Winner  
*Silvana Castano, Valeria De Antonellis, Maria Grazia Fugini, Barbara Pernici* iv
- Querying Genome Databases (Extended Abstract) - Honorable Mention vi  
*Giansalvatore Mecca*

## Data Extraction, Integration and Visualization

- OpenTRIAGE: Entity Linkage for Detail Webpages 1-12  
*Roger Voyat, Valter Crescenzi, Paolo Merialdo*
- Extracting a Unified Database from Heterogeneous Bills of Materials via Data Integration Approaches 13-24  
*Luigi De Lucia, Guglielmo Faggioli, Nicola Ferro, Fabio Orobelli*
- Towards an Italian Energy Data Space 25-35  
*Chiara Rucco, Antonella Longo, Marco Zappatore*
- Describing Multidimensional Data Through Highlights (Discussion Paper) 36-43  
*Matteo Francia, Enrico Gallinucci, Matteo Golfarelli, Patrick Marcel, Verónika Peralta, Stefano Rizzi*
- A Multi-Perspective Data Model for Cyber Physical Production Networks (Discussion Paper) 44-51  
*Ada Bagozi, Devis Bianchini, Anisa Rula*
- Progressive Entity Resolution with Node Embeddings (Discussion Paper) 52-60  
*Giovanni Simonini, Luca Gagliardelli, Michele Rinaldi, Luca Zecchini, Giulio De Sabbata, Adeel Aslam, Domenico Beneventano, Sonia Bergamaschi*

- Active E-Learning by Doing with ALDO (Discussion Paper)
61-68  
*Ilaria Bartolini, Andrea Di Luzio*
- Landmark Explanation: a Tool for Entity Matching (Discussion Paper)
69-76  
*Andrea Baraldi, Francesco Del Buono, Matteo Paganelli, Francesco Guerra*
- Automatic Information Extraction from Investment Product Documents (Discussion Paper)
77-84  
*Federico Maria Scafoglieri, Alessandra Monaco, Giulia Neccia, Domenico Lembo, Alessandra Limosani, Francesca Medda*
- Conceptually-grounded Mapping Patterns for Virtual Knowledge Graphs (Discussion Paper)
85-92  
*Diego Calvanese, Avigdor Gal, Davide Lanti, Marco Montali, Alessandro Mosca, Roe Shraga*

## **(Big) Data Governance**

- Understanding RDF Data Representations in Triplestores (Discussion Paper)
93-100  
*Matteo Lissandrini, Tomer Sagi, Torben Bach Pedersen, Katja Hose*
- The Usage of Negation in Real-World JSON Schema Documents (Discussion Paper)
101-108  
*Mohamed-Amine Baazizi, Dario Colazzo, Giorgio Ghelli, Carlo Sartiani, Stefanie Scherzinger*
- Energy and QoE aware Placement of Applications and Data at the Edge (Discussion Paper)
109-116  
*Matteo Mordacchini, Luca Ferrucci, Emanuele Carlini, Hanna Kavalionak, Massimo Coppola, Patrizio Dazzi*
- SoBigData RI: European Integrated Infrastructure for Social Mining and Big Data Analytics (Discussion Paper)
117-124  
*Roberto Trasarti, Valerio Grossi, Michela Natilli, Beatrice Rapisarda*
- Workflows for Bringing Data Science on the Cloud/Edge Computing Continuum (Discussion Paper)
125-132  
*Patrizio Dazzi, Valerio Grossi, Roberto Trasarti*
- Exploiting Curated Databases to Train Relation Extraction Models for Gene-Disease Associations (Discussion Paper)
133-140  
*Stefano Marchesin, Gianmaria Silvello*
- Imputation of Missing Values through Profiling Metadata (Discussion Paper)
141-148  
*Bernardo Breve, Loredana Caruccio, Vincenzo Deufemia, Giuseppe Polese*
- Supporting the Design of Data Preparation Pipelines (Discussion Paper)
149-158  
*Camilla Sancricca, Cinzia Cappiello*
- Counting Database Repairs Entailing a Query: The Case of Functional Dependencies (Discussion Paper)
159-166  
*Marco Calautti, Ester Livshits, Andreas Pieris, Markus Schneider*

## **Data for Smart Society**

- [A Survey of Sentimental Analysis Methods on COVID-19 Research \(Discussion Paper\)](#) 167-174  
*Areeba Umair, Elio Masciari*
- [MAT-Builder: a System to Build Semantically Enriched Trajectories \(Discussion Paper\)](#) 175-182  
*Chiara Pugliese, Francesco Lettich, Chiara Renso, Fabio Pinelli*
- [A Network-based Model and a Related Approach to Represent and Handle the Semantics of Comments in a Social Network \(Discussion Paper\)](#) 183-190  
*Gianluca Bonifazi, Francesco Cauteruccio, Enrico Corradini, Michele Marchetti, Giorgio Terracina, Domenico Ursino, Luca Virgili*
- [Explaining Link Prediction with Kelpie \(Discussion Paper\)](#) 191-198  
*Andrea Rossi, Donatella Firmani, Paolo Merialdo, Tommaso Teofili*
- [Predicting Vehicles Parking Behaviour for EV Recharge Optimization \(Discussion Paper\)](#) 199-206  
*Vinicius Monteiro de Lira, Fabiano Pallonetto, Lorenzo Gabrielli, Chiara Renso*
- [Discovering the Landscape of Decentralized Online Social Networks through Mastodon \(Discussion Paper\)](#) 207-214  
*Lucio La Cava, Sergio Greco, Andrea Tagarelli*
- [Extracting Large Scale Spatio-Temporal Descriptions from Social Media \(Discussion Paper\)](#) 215-222  
*Carlo Bono, Barbara Pernici*
- [Online News Event Extraction for Crime Analysis \(Discussion Paper\)](#) 223-230  
*Federica Rollo, Laura Po, Giovanni Bonisoli*

## Knowledge, Ontologies and Reasoning on Data

- [Semantic Shift Detection in Vatican Publications: a Case Study from Leo XIII to Francis](#) 231-243  
*Silvana Castano, Alfio Ferrara, Stefano Montanelli, Francesco Periti*
- [Shared Metrics of Sustainability: a Knowledge Graph Approach](#) 244-255  
*Claudia Diamantini, Tarique Khan, Domenico Potena, Emanuele Storti*
- [LPG-based Ontologies as Schemas for Graph DBs](#) 256-267  
*Stefano Ferilli, Domenico Redavid, Davide Di Pierro*
- [A Topological Perspective of Port Networks From Three Years \(2017-2019\) of AIS Data \(Discussion Paper\)](#) 268-275  
*Emanuele Carlini, Vinicius Monteiro de Lira, Amilcar Soares, Mohammad Etemad, Bruno Brandoli, Stan Matwin*
- [Expanding the Citation Graph for Data Citations \(Discussion Paper\)](#) 276-283  
*Peter Buneman, Dennis Dosso, Matteo Lissandrini, Gianmaria Silvello*
- [Active Integrity Constraints with Existential Quantification \(Discussion Paper\)](#) 284-291  
*Marco Calautti, Luciano Caroprese, Sergio Greco, Cristian Molinaro, Irina Trubitsyna, Ester Zumpano*
- [Reasoning in Warded Datalog+/- with Harmful Joins \(Discussion Paper\)](#) 292-299  
*Teodoro Baldazzi, Luigi Bellomarini, Emanuel Sallinger, Paolo Atzeni*

- [An Overview of Vadalog: a System for Reasoning over Large Knowledge Graphs \(Discussion Paper\)](#) 300-308  
*Luigi Bellomarini, Davide Benedetto, Emanuel Sallinger*
- [Reasoning about Smart Contracts via LTL Encoding \(Discussion Paper\)](#) 309-316  
*Valeria Fionda, Gianluigi Greco, Marco Antonio Mastratisi*

## Machine Learning and Data Mining

- [Decentralized Federated Learning and Network Topologies: an Empirical Study on Convergence \(Discussion Paper\)](#) 317-324  
*Hanna Kavalionak, Emanuele Carlini, Patrizio Dazzi, Luca Ferrucci, Matteo Mordacchini, Massimo Coppola*
- [Mapping and Compressing a Convolutional Neural Network through a Multilayer Network \(Discussion Paper\)](#) 325-332  
*Alessia Amelio, Gianluca Bonifazi, Enrico Corradini, Michele Marchetti, Domenico Ursino, Luca Virgili*
- [Ambiguity Detection and Textual Claims Generation from Relational Data \(Discussion Paper\)](#) 333-340  
*Enzo Veltri, Donatello Santoro, Gilbert Badaro, Mohammed Saeed, Paolo Papotti*
- [Generating Synthetic Discrete Datasets with Machine Learning \(Discussion Paper\)](#) 341-350  
*Giuseppe Manco, Ettore Ritacco, Antonino Rullo, Domenico Saccà, Edoardo Serra*
- [Distributed Heterogeneous Transfer Learning \(Discussion Paper\)](#) 351-358  
*Paolo Mignone, Gianvito Pio, Michelangelo Ceci*
- [Multiple Instance Learning for viral pneumonia chest X-ray Classification \(Discussion Paper\)](#) 359-366  
*Matteo Avolio, Antonio Fuduli, Eugenio Vocaturo, Ester Zumpano*
- [Towards Extreme Multi-Label Classification of Multimedia Content \(Discussion Paper\)](#) 367-374  
*Marco Minici, Francesco Sergio Pisani, Massimo Guarascio, Giuseppe Manco*
- [Improving Conversational Evaluation via a Dependency-Aware Permutation Strategy \(Discussion Paper\)](#) 375-382  
*Guglielmo Faggioli, Marco Ferrante, Nicola Ferro, Raffaele Perego, Nicola Tonello*

## Data Analysis

- [Attributed Stream-Hypernetwork Analysis: a SocioPatterns Case Study](#) 383-391  
*Andrea Failla, Salvatore Citraro, Giulio Rossetti*
- [Data-Driven, AI-Based Clinical Practice: Experiences, Challenges, and Research Directions](#) 392-403  
*Davide Ferrari, Federica Mandreoli, Federico Motta, Paolo Missier*
- [Supporting Big Moving Objects Tracking and Analysis: An Innovative Scan-](#)

Matching Algorithm 404-414  
*Alfredo Cuzzocrea, Majid Abbasi Sisara, Kristijan Lenac, Enzo Mumolo*

## Data Clustering and Partitioning

- Investigating Binary Partition Power in Metric Query 415-426  
*Richard Connor, Alan Dearle, Lucia Vadicamo*
- Annotating Protein Structures for Understanding SARS-CoV-2 Interactome (Discussion Paper) 427-436  
*Barbara Puccio, Ugo Lomoio, Luisa Di Paola, Pietro Hiram Guzzi, Pierangelo Veltri*
- In-Vehicle Big Data Exploration for Road Maintenance (Discussion Paper) 437-444  
*Devis Bianchini, Valeria De Antonellis, Massimiliano Garda*
- Correlation Clustering: from Local to Global Constraints (Discussion Paper) 445-452  
*Domenico Mandaglio, Andrea Tagarelli, Francesco Gullo*
- SiMBA: Systematic Clustering-Based Methodology to Support Built Environment Analysis (Discussion Paper) 453-460  
*Carlo Andrea Biraghi, Emilia Lenzi, Maristella Matera, Massimo Tadi, Letizia Tanca*
- Clustered Majority Judgement (Discussion Paper) 461-472  
*Arianna Anniciello, Emanuele d'Ajello, Davide Formica, Elio Masciari, Gaia Mattia, Stefano Quintarelli, Davide Zaccarella*

## Recommender Systems, Query Answering and Processing

- Helping Wine Lovers With Taxonomies (Discussion Paper) 473-480  
*Paolo Ciaccia, Davide Martinenghi, Riccardo Torlone*
- Query Answer Explanations under Existential Rules (Discussion Paper) 481-488  
*Ismail Ilkan Ceylan, Thomas Lukasiewicz, Enrico Malizia, Andrius Vaicenavičius*
- Explanations for Inconsistency-Tolerant Query Answering under Existential Rules (Discussion Paper) 489-496  
*Thomas Lukasiewicz, Enrico Malizia, Cristian Molinaro*
- Personalized Context-Aware Recommender System for Travelers (Discussion Paper) 497-504  
*Mahsa Shekari, Alireza Javadian Sabet, Chaofeng Guan, Matteo Rossi, Fabio A. Schreiber, Letizia Tanca*
- Inferring User Decision-Making Processes in Recommender Systems with Knowledge Graphs (Discussion Paper) 505-513  
*Vito Walter Anelli, Tommaso Di Noia, Eugenio Di Sciascio, Antonio Ferrara, Alberto Carlo Maria Mancino*
- The Challenging Reproducibility Task in Recommender Systems Research between Traditional and Deep Learning Models (Discussion Paper) 514-521  
*Vito Walter Anelli, Alejandro Bellogin, Antonio Ferrara, Daniele Malitesta, Felice Antonio Merra, Claudio Pomo, Francesco M. Donini, Eugenio Di Sciascio, Tommaso Di Noia*



- [Investigating Monotone Abstractions \(Discussion Paper\)](#)  
*Gianluca Cima, Marco Console, Maurizio Lenzerini, Antonella Poggi*
522-529
- [Complexity of Inconsistency-Tolerant Query Answering in Datalog+/- under Cardinality-Based Repairs \(Discussion Paper\)](#)  
*Thomas Lukasiewicz, Enrico Malizia, Andrius Vaicenavičius*
530-537
- [SoCRATE: a Framework for Compensating Users Over Time with Limited Availability Recommendations \(Discussion Paper\)](#)  
*Davide Azzalini, Fabio Azzalini, Chiara Criscuolo, Tommaso Dolci, Davide Martinenghi, Letizia Tanca, Sihem Amer-Yahia*
538-545

## Explainable AI, Ethics and Fairness of Data

- [Credit Score Prediction Relying on Machine Learning \(Discussion Paper\)](#)  
*Flora Amato, Antonino Ferraro, Antonio Galli, Francesco Moscato, Vincenzo Moscato, Giancarlo Sperli*
546-553
- [Bias Score: Estimating Gender Bias in Sentence Representations \(Discussion Paper\)](#)  
*Fabio Azzalini, Tommaso Dolci, Mara Tanelli*
554-561
- [Functional Dependencies to Mitigate Data Bias \(Discussion Paper\)](#)  
*Fabio Azzalini, Chiara Criscuolo, Letizia Tanca*
562-569
- [Accounting for Bossy Users in Context-Aware Group Recommendations \(Discussion Paper\)](#)  
*Davide Azzalini, Elisa Quintarelli, Emanuele Rabosio, Letizia Tanca*
570-577
- [Gender Discriminatory Language Identification with an Hybrid Algorithm based on Syntactic Rules and Machine Learning \(Discussion Paper\)](#)  
*Valerio Bellandi, Stefano Siccardi*
578-585
- [A Federated Cloud Solution for Transnational Mobility Data Sharing \(Discussion Paper\)](#)  
*Emanuele Carlini, Thierry Chevalier, Patrizio Dazzi, Francesco Lettich, Raffaele Perego, Chiara Renso, Salvatore Trani*
586-592
- [Cross-Social Network Investigation to Highlight Privacy Violations in Data Sharing Activities \(Discussion Paper\)](#)  
*Francesca Cerruto, Stefano Cirillo, Domenico Desiato, Simone Michele Gambardella, Giuseppe Polese*
593-600
- [Risk Analysis for Unsupervised Privacy-Preserving Tools \(Discussion Paper\)](#)  
*Alfredo Cuzzocrea, Christophe Cerisara, Mojtaba Hajian*
601-609

## Doctoral Consortium

- [Recent Advancements on Bio-Inspired Hebbian Learning for Deep Neural Networks](#)  
*Gabriele Lagani*
610-615
- [A Research on Data Lakes and their Integration Challenges](#)  
*Davide Piantella*
616-621
- [An XAI Framework for Automated Allocation and Exploitation of Resources](#)  
*Diego Marcia*
622-626



