

# **29th Italian Symposium on Advanced Database Systems (SEBD 2021)**

Pizzo, Italy  
5 - 9 September 2021

ISBN: 978-1-7138-3890-6

**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.  
All rights reserved.

Printed by Curran Associates, Inc. (2022)

Edited by:  
Sergio Greco  
Maurizio Lenzerini  
Elio Masciari  
Andrea Tagarelli

**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

## **REGULAR PAPERS**

EFFICIENT VALIDATION OF FUNCTIONAL DEPENDENCIES DURING INCREMENTAL DISCOVERY .....	1
<i>Loredana Caruccio, Stefano Cirillo, Vincenzo Deufemia, Giuseppe Polese</i>	
NON-LINEAR ANALYTIC PREDICTION OF IP ADDRESSES FOR SUPPORTING CYBER ATTACK DETECTION AND ANALYSIS .....	13
<i>Alfredo Cuzzocrea, Enzo Mumolo, Edoardo Fadda, Selim Soufargi, Carson K. Leung</i>	
FROM DIGITAL TO COMPUTATIONAL HUMANITIES: THE VAST PROJECT VISION .....	24
<i>Silvana Castano, Alfio Ferrara, Stefano Montanelli, Francesco Periti</i>	
FROM DATA QUALITY TO BIG DATA QUALITY: A DATA INTEGRATION SCENARIO .....	36
<i>Carlo Batini, Anisa Rula</i>	
SNPSTORAGE: A DATABASE TO STORE AND SHARE EXPERIMENTAL SNPS DATA SETS.....	48
<i>Giuseppe Agapito, Mario Cannataro</i>	
THE DOCUMENT CHAIN: A DELTA CRDT FRAMEWORK FOR ARBITRARY JSON DATA.....	59
<i>Amos Brocco</i>	
FEDERATED RECOMMENDER SYSTEMS WITH LEARNING TO RANK.....	71
<i>Vito Walter Anelli, Yashar Deldjoo, Tommaso Di Noia, Antonio Ferrara, Fedelucio Narducci</i>	
A NEURAL NETWORK STRATEGY FOR SUPERVISED CLASSIFICATION VIA THE LEARNING UNDER PRIVILEGED INFORMATION PARADIGM.....	83
<i>Ludovica Sacco, Dino Ienco, Roberto Interdonato</i>	
FROM LOW-LEVEL-EVENT-LOGS TO HIGH-LEVEL-BUSINESS-PROCESS-MODEL-ACTIVITIES: AN ADVANCED FRAMEWORK BASED ON MACHINE LEARNING AND FLEXIBLE BPMN MODEL TRANSLATION .....	94
<i>Alfredo Cuzzocrea, Ernesto Damiani, Hamda Al-Ali, Rabeb Mizouni, Ghalia Tello, Edoardo Fadda</i>	
DEVELOPMENT OF A DECISION TREE MODEL TO IMPROVE CASE DETECTION VIA INFORMATION EXTRACTION FROM VETERINARY ELECTRONIC MEDICAL RECORDS .....	105
<i>Oscar Tamburis, Elio Masciari, Gerardo Fatone</i>	
A WEB TOOL FOR THE SEMANTIC INTEGRATION OF HETEROGENEOUS AND COMPLEX SPREADSHEET TABLES .....	116
<i>Sara Bonfitto, Luca Cappelletti, Elena Casiraghi, Paolo Perlasca, Fabrizio Trovato, Giorgio Valentini, Marco Mesiti</i>	
AMFLP: ADVERSARIAL MATRIX FACTORIZATION-BASED LINK PREDICTOR IN SOCIAL GRAPHS.....	128
<i>Giuseppe De Candia, Tommaso Di Noia, Eugenio Di Sciascio, Felice Antonio Merra</i>	
AN EFFECTIVE AND EFFICIENT GRAPH REPRESENTATION LEARNING APPROACH FOR BIG GRAPHS.....	140
<i>Edoardo Serra, Mikel Joaristi, Alfredo Cuzzocrea, Selim Soufargi, Carson K. Leung</i>	

A KNOWLEDGE-BASED APPROACH TO BUSINESS ANALYSIS FOR PROCESS INNOVATION.....	150
<i>Michele Missikoff</i>	

## **SHORT PAPERS**

HYPERGRAPH DATA ANALYSIS WITH PAOHVIS .....	160
<i>Paolo Buono, Miguel Ceriani, Maria Costabile</i>	
A NOVEL NETWORK-BASED METHODOLOGY FOR ANALYSIS OF COVID-19 DATA .....	168
<i>Marianna Milano, Mario Cannataro</i>	
A COLLABORATIVE TOOL TO SUPPORT JUNIOR DEVELOPERS IN SECURE SOFTWARE DEVELOPMENT .....	176
<i>Maria Teresa Baldassarre, Vita Santa Barletta, Danilo Caivano, Giovanni Dimauro, Antonio Piccinno</i>	
A RESEARCH INFRASTRUCTURE FOR E-HEALTH BIG DATA ANALYTICS.....	185
<i>Alessio Botta, Elio Masciari</i>	
A SHORT ACCOUNT OF FAIR-DB: A SYSTEM TO DISCOVER DATA BIAS (DISCUSSION PAPER) .....	192
<i>Fabio Azzalini, Chiara Criscuolo, Letizia Tanca</i>	
ON DATABASE INCONSISTENCY MEASURES (DISCUSSION PAPER).....	200
<i>Francesco Parisi, John Grant</i>	
BENCHMARKING APPROXIMATE CONSISTENT QUERY ANSWERING (DISCUSSION PAPER) .....	209
<i>Marco Calautti, Marco Console, Andreas Pieris</i>	
PREFERENCE-BASED INCONSISTENCY-TOLERANT QUERY ANSWERING UNDER EXISTENTIAL RULES (DISCUSSION PAPER) .....	217
<i>Marco Calautti, Sergio Greco, Cristian Molinaro, Irina Trubitsyna</i>	
A SYSTEM FOR TRANSLATING NATURAL LANGUAGE QUESTIONS INTO SPARQL QUERIES WITH NEURAL NETWORKS: PRELIMINARY RESULTS (DISCUSSION PAPER).....	226
<i>Manuel Alejandro Borroto, Francesco Ricca, Bernardo Cuteri</i>	
SMARE: AN ENHANCED QUERY PERFORMANCE PREDICTION EVALUATION APPROACH (DISCUSSION PAPER) .....	235
<i>Guglielmo Faggioli, Oleg Zendel, J. Shane Culpepper, Nicola Ferro, Falk Scholer</i>	
A DATA-DRIVEN CONTEXT-BASED APPROACH FOR MODELLING RESILIENT CYBER PHYSICAL PRODUCTION SYSTEMS (DISCUSSION PAPER).....	243
<i>Ada Bagozi, Devis Bianchini, Valeria De Antonellis</i>	
SOME EXPERIMENTS ON ACTIVITY OUTLIER DETECTION (DISCUSSION PAPER) .....	251
<i>Michele Ianni, Elio Masciari</i>	
A TWO-STEP NETWORK INTRUSION DETECTION SYSTEM FOR MULTI-CLASS CLASSIFICATION (DISCUSSION PAPER) .....	259
<i>Giuseppina Andresini, Annalisa Appice, Donato Malerba</i>	

INCREASING PROTECTION AND AUTONOMY IN THE IOT THROUGH A TWO-TIER BLOCKCHAIN FRAMEWORK (DISCUSSION PAPER) .....	267
<i>Enrico Corradini, Serena Nicolazzo, Antonino Nocera, Domenico Ursino, Luca Virgili</i>	
EXTRACTION OF MEDICAL CONCEPTS FROM ITALIAN NATURAL LANGUAGE DESCRIPTIONS (DISCUSSION PAPER) .....	275
<i>Patrizia Agnello, Silvia Maria Ansaldi, Fabio Azzalini, Giovanni Gangemi, Davide Piantella, Emanuele Rabosio, Letizia Tanca</i>	
THE POOR-MAD APPROACH: PREFERRED OBJECTS OVER RICH, MULTI-ATTRIBUTE DATA (DISCUSSION PAPER) .....	283
<i>Paolo Ciaccia, Davide Martinenghi, Riccardo Torlone</i>	
USING DESCRIPTIONS FOR EXPLAINING ENTITY MATCHES (DISCUSSION PAPER) .....	291
<i>Matteo Paganelli, Paolo Sottovia, Antonio Maccioni, Matteo Interlandi, Francesco Guerra</i>	
AUTOMATIC ENTITY LABELING THROUGH EXPLANATION TECHNIQUES (DISCUSSION PAPER) .....	299
<i>Silvana Castano, Alfio Ferrara, Donatella Firmani, Jerin George Mathew, Stefano Montanelli</i>	
A SOFTWARE PIPELINE FOR PRE-PROCESSING AND MINING EEG SIGNALS: APPLICATION IN NEUROLOGY (DISCUSSION PAPER).....	307
<i>Chiara Zucco, Barbara Calabrese, Miriam Sturniolo, Antonio Gambardella, Mario Cannataro</i>	
CATAPLEXY DETECTION: NEUROLOGISTS, YOU ARE NOT ALONE! (DISCUSSION PAPER) .....	315
<i>Ilaria Bartolini, Andrea Di Luzio</i>	
SENTIMENT ANALYSIS OF MEDICAL QUESTIONNAIRES TO IMPROVE ADHERENCE TO TELEMONITORING PROGRAMS (DISCUSSION PAPER) .....	323
<i>Chiara Zucco, Clarissa Paglia, Sonia Graziano, Sergio Bella, Mario Cannataro</i>	
EMBDI: GENERATING EMBEDDINGS FOR RELATIONAL DATA INTEGRATION (DISCUSSION PAPER) .....	331
<i>Riccardo Cappuzzo, Paolo Papotti, Saravanan Thirumuruganathan</i>	
TOWARDS AN ACCESS CONTROL MODEL FOR KNOWLEDGE GRAPHS (DISCUSSION PAPER) .....	339
<i>Marco Valzelli, Andrea Maurino, Matteo Palmonari, Blerina Spahiu</i>	
MULTI-SORTED INVERSE FREQUENT ITEMSETS MINING FOR GENERATING REALISTIC NO-SQL DATASETS (DISCUSSION PAPER) .....	347
<i>Domenico Saccà, Edoardo Serra, Antonino Rullo</i>	
SPICE-UP YOUR CODING LESSONS WITH THE ACME APPROACH (DISCUSSION PAPER) .....	355
<i>Giansalvatore Mecca, Donatello Santoro, Nazzareno Sileno, Enzo Veltri</i>	
THE CASE FOR MULTI-TASK ACTIVE LEARNING ENTITY RESOLUTION (DISCUSSION PAPER) .....	363
<i>Giovanni Simonini, Henrique Saccani, Luca Gagliardelli, Luca Zecchini, Domenico Beneventano, Sonia Bergamaschi</i>	

AN INDUCTIVE FRAMEWORK FOR SEMI-SUPERVISED LEARNING (DISCUSSION PAPER) .....	371
<i>Shuyi Yang, Dino Ienco, Roberto Esposito, Ruggero G. Pensa</i>	
A TWO-STEP METHOD BASED ON EMBEDDING AND CLUSTERING TO IDENTIFY REGULARITIES IN LEGAL CASE JUDGEMENTS (DISCUSSION PAPER).....	379
<i>Graziella De Martino, Gianvito Pio, Michelangelo Ceci</i>	
ACTIVITY DAILY LIVING PREDICTION WITH MARKED TEMPORAL POINT PROCESSES (DISCUSSION PAPER).....	387
<i>Giancarlo Fortino, Antonella Guzzo, Michele Ianni, Francesco Leotta, Massimo Mecella</i>	
A MULTIPLE INSTANCE LEARNING APPROACH FOR THE AUTOMATIC CLASSIFICATION OF SKIN LESIONS (DISCUSSION PAPER).....	395
<i>Eugenio Vocaturo, Ester Zumpano, Giovanni Giallombardo, Giovanna Miglionico</i>	
CONVERSATIONAL OLAP (DISCUSSION PAPER) .....	403
<i>Matteo Francia, Enrico Gallinucci, Matteo Golfarelli</i>	
FINE-GRAINED PROVENANCE FOR HIGH-QUALITY DATA SCIENCE (DISCUSSION PAPER) .....	411
<i>Adriane Chapman, Paolo Missier, Giulia Simonelli, Riccardo Torlone</i>	
AN INVESTIGATION ON NOT SAFE FOR WORK ADULT CONTENT IN REDDIT (DISCUSSION PAPER) .....	419
<i>Francesco Cauteruccio, Enrico Corradini, Giorgio Terracina, Domenico Ursino, Luca Virgili</i>	
A MULTI-RESOLUTION TRAINING FOR EXPRESSION RECOGNITION IN THE WILD (DISCUSSION PAPER) .....	427
<i>Fabio Valerio Massoli, Donato Cafarelli, Giuseppe Amato, Fabrizio Falchi</i>	
APPLYING GENERATIVE ADVERSARIAL NETWORKS TO PERFORM GAZE PATH PREDICTION IN WEBSITES (DISCUSSION PAPER) .....	434
<i>Gianluca Bonifazi, Enrico Corradini, Gianluca Porcino, Alessandro Scopelliti, Domenico Ursino, Luca Virgili</i>	
TRAFFIC DENSITY ESTIMATION VIA UNSUPERVISED DOMAIN ADAPTATION (DISCUSSION PAPER) .....	442
<i>Luca Ciampi, Carlos Santiago, Joao Paulo Costeira, Claudio Gennaro, Giuseppe Amato</i>	
UNBIASING COLLABORATIVE FILTERING FOR POPULARITY-AWARE RECOMMENDATION (DISCUSSION PAPER) .....	450
<i>Luciano Caroprese, Giuseppe Manco, Marco Minici, Francesco Sergio Pisani, Ettore Ritacco</i>	
RECKONITION: A NLP-BASED SYSTEM FOR INDUSTRIAL ACCIDENTS AT WORK PREVENTION (DISCUSSION PAPER) .....	458
<i>Patrizia Agnello, Silvia Maria Ansaldi, Emilia Lenzi, Alessio Mongelluzzo, Manuel Roveri</i>	
RECKON: A REAL-WORLD, CONTEXT-AWARE KNOWLEDGE-BASED LAB (DISCUSSION PAPER) .....	466
<i>Patrizia Agnello, Silvia Maria Ansaldi, Emilia Lenzi, Alessio Mongelluzzo, Davide Piantella, Manuel Roveri, Fabio A. Schreiber, Alessandra Scutti, Mahsa Shekari, Letizia Tanca</i>	
A DEEP LEARNING APPROACH FOR UNSUPERVISED FAILURE DETECTION IN SMART INDUSTRY (DISCUSSION PAPER) .....	474
<i>Angelica Liguori, Giuseppe Manco, Ettore Ritacco, Massimiliano Ruffolo, Salvatore Iiritano</i>	

DP-DILCA: LEARNING DIFFERENTIALLY PRIVATE CONTEXT-BASED DISTANCES FOR CATEGORICAL DATA (DISCUSSION PAPER) .....	482
<i>Elena Battaglia, Ruggero G. Pensa</i>	
PRIVACY PRESERVING QUERY ANSWERING IN DESCRIPTION LOGICS THROUGH INSTANCE INDISTINGUISHABILITY (DISCUSSION PAPER).....	490
<i>Gianluca Cima, Domenico Lembo, Riccardo Rosati, Domenico Fabio Savo</i>	
IMPROVING WORK LIFE CONDITIONS VIA PORTABLE KNOWLEDGE-DRIVEN RECOMMENDER SYSTEM (DISCUSSION PAPER) .....	498
<i>Maurizio Atzori, Alessandro Barengli, Sara Comai, Mariagrazia Fugini, Diego Marcia, Gerardo Pelosi, Manuela Sanguinetti, Vincenzo Scotti</i>	
NANOWEB: SEARCH, ACCESS AND EXPLORE LIFE SCIENCE NANOPUBLICATIONS ON THE WEB (DISCUSSION PAPER) .....	506
<i>Fabio Giachelle, Dennis Dosso, Gianmaria Silvello</i>	
A DATA MANAGEMENT AND ANOMALY DETECTION SOLUTION FOR THE ENTERTAINMENT INDUSTRY (DISCUSSION PAPER).....	514
<i>Michele Berno, Marco Canil, Nicola Chiarello, Luca Piazzon, Fabio Berti, Francesca Ferrari, Alessandro Zaupa, Nicola Ferro, Michele Rossi, Gian Antonio Susto</i>	
SMART CONVERSATIONAL USER INTERFACE FOR RECOMMENDING CULTURAL HERITAGE POINTS OF INTEREST (DISCUSSION PAPER) .....	522
<i>Flora Amato, Francesco Moscato, Vincenzo Moscato, Giancarlo Sperli</i>	
 <b><u>SEMINAR ON REMINISCENCE OF TIDB 1981</u></b>	
PANEL ON "PAST AND FUTURE OF COMPUTER SCIENCE THEORY" (DISCUSSION PAPER) .....	531
<i>Ronald Fagin, Georg Gottlob, Christos Papadimitriou, Moshe Vardi, Giorgio Ausiello, Maurizio Lenzerini, Domenico Saccà, Luigi Palopoli, Francesco Scarcello</i>	

**Author Index**