Foreword

The International Conference on Extending Database Technology (EDBT) is a leading international forum for database researchers, developers, and users to present and discuss novel and cutting-edge ideas and techniques, and to showcase tools and experiences related to data management. Data management is an essential enabling technology that has applications in several scientific, business and social communities, and runs on diverse technical platforms associated with the web, enterprises, clouds and mobile devices. The database community has a continuing tradition of contributing with models, algorithms and architectures to the set of tools and applications that enable day-to-day functioning of our societies. Faced with the broad challenges of today’s applications, data management technology constantly broadens its reach, exploiting new hardware and software to achieve innovative results and embracing new challenges in the years to come.

EDBT 2020 solicited submissions of original research contributions, descriptions of industrial solutions and applications, and proposals for tutorials and software demonstrations. We encouraged submissions of research papers related to all aspects of data management. We also encouraged submissions of visionary papers as well as innovative system papers and experimental analyses papers. In addition to long research paper submissions, EDBT 2020 again encouraged the submission of short research papers, which provide an excellent opportunity to describe significant work or research in progress that can foster the discussions at the conference. Short papers are presented as posters at plenary poster sessions of the conference. This year, they will also be communicated at a plenary lightning talks session.

The program committees of EDBT accepted 30 out of 151 submitted regular research papers, resulting in an acceptance rate of 20% for the research track; 26 out of 85 submitted short research papers, resulting in an acceptance rate of 31% for short research papers; 16 out of 37 demos, resulting in an acceptance rate of 43% for the demonstration track; and 10 out of 37 industrial and application papers, resulting in an acceptance rate of 27% for industrial and application papers. The papers will be presented in eight research paper sessions, three industrial and application sessions, as well as two plenary poster and demonstration sessions.

The program additionally features five workshops, one of which is the well-established DOLAP workshop that has successfully been co-located with EDBT since many years. Finally, the conference program includes four tutorials and an EDBT and ICDT joint session on climate change.

I would like to thank all authors for their contributions: a successful conference crucially depends on high-quality submissions. I also would like to thank all senior reviewers and reviewers for serving on the EDBT 2020 program committee, in particular for the high quality and timely handling of all reviews and discussions. This community service requires a lot of work on a tight schedule, and is what makes our research community function and ensures the sustained impact of our research. Thanks to their valuable effort we can look forward to an exciting program and attractive EDBT conference in Copenhagen from March 30–April 2, 2020.

A warm thanks to Anastasia Alaimaki and Tamer Özsu for serving on the Test-of-Time Award committee to select the paper from EDBT 2010 that has had the most lasting influence. Wook-shin Han, Erhard Rahm and Nesime Tatbul generously accepted to serve on the Best Paper committee. The EDBT 2020 program is the result of the joint effort of many people who shared their experience and time to contribute to the EDBT 2020 program and make the conference a great success. Alexander Boehm served as PC chair for industrial and application papers; George Fletcher as PC chair for the demonstration track; Dan Olteanu as tutorial chair; Alexandra Poulouvasilis as workshop chair. My warmest thanks to all these people. The general chairs, Yongluan Zhou and Marcos Vaz Salles and the other local organizers worked hard to make all necessary arrangements for a successful event. Special thanks to Arijit Khan, the EDBT proceedings chair; Davide Mottin, the publicity chair; Bin Yang, the local executive chair; Boris Düdder, the sponsorship chair; Nanna Hejbohl, the finance chair and Yiwen Wang, the website chair, for tirelessly finding solutions for all our requests. Norman Paton was most helpful in advising and coordinating with the EDBT Executive Board. Last but not least, I would like to thank Marc H. Scholl for assembling the EDBT proceedings on openproceedings.org I hope that you find EDBT 2020 inspiring, enriching, and enjoyable and look forward to meeting you in Copenhagen.

Angela Bonifati
EDBT 2020 Program Chair
Program Committee Members

Research Program Committee

Angela Bonifati (Lyon 1 U, France) – Chair

Senior Program Committee Members

Karl Aberer (EPFL Lausanne, Switzerland)
Walid Aref (Purdue U, USA)
Michael Benedikt (U Oxford, UK)
Michael Böhlen (U Zurich, Switzerland)
K. Selcuk Candan (Arizona State U, USA)
Kevin Chang (U Illinois at Urbana-Champaign, USA)
Vassilis Christophides (INRIA, France)
Daniel Deutch (Tel Aviv U, Israel)
Floris Geerts (U Antwerp, Belgium)
Jan Hidders (VU Brussel, Belgium)
Katja Hose (Aalborg U, Denmark)
Christoph Koch (EPFL Lausanne, Switzerland)
Georgia Koutrika (Athena Research Center, Greece)

Program Committee Members

Bernd Amann (Sorbonne U – LIP6, France)
Akhil Arora (EPFL Lausanne, Switzerland)
Elena Baralis (Politecnico di Torino, Italy)
Denilson Barbosa (U Alberta, Canada)
Senjuti Basu Roy (New Jersey Inst. of Techn., USA)
Luigi Bellomarini (Banca d’Italia, Italy)
Sonia Bergamaschi (U Modena Reggio Emilia, Italy)
Laure Berti-Equille (IRD, France)
Arnab Bhattacharya (IIT Kanpur, India)
Luc Bouganim (INRIA-UVSQ, France)
Andrea Cali (U London, Birkbeck College, UK)
Bogdan Cautis (Paris Sud U, France)
Lei Chen (Hong Kong U Sc. & Techn., China)
Dario Colazzo (Paris Dauphine U, France)
Bin Cui (Peking U, China)
Alfredo Cuzzocrea (U Calabria, Italy)
Sabrina De Capitani di Vimercati (U Milan, Italy)
Stefania Dumbрова (ENSIEE Rennes, France)
Donatella Firmani (Roma Tre U, Italy)
Rainer Gemulla (U Mannheim, Germany)
Paolo Guagliardo (U Edinburgh, UK)
Xi He (U Waterloo, Canada)
Xin Huang (Hong Kong Baptist U, China)
Zsolt Istvan (IMDEA Software Institute, Spain)
Panos Kalnis (King Abdullah UST, Saudi Arabia)
Vana Kalogeraki (Athens U Ec. & Busin., Greece)
Verena Kantere (U Ottawa, Canada)

Ulf Leser (Humboldt-U Berlin, Germany)
Guoliang Li (Tsinghua U, China)
Chengkai Li (U Texas at Arlington, USA)
Eric Lo (Chinese U Hong Kong, China)
Evaggelia Pitoura (U Ioannina, Greece)
Louisa Raschid (U Maryland, USA)
Sherif Sakr (U Tartu, Estonia)
Semih Salihoglu (U Waterlo, Canada)
Kai-Uwe Sattler (TU Ilmenau, Germany)
Arash Termehchy (Oregon State U, USA)
Riccardo Torlone (Roma Tre U, Italy)
Peter Triantafillou (U Warwick, UK)
Yannis Velegrakis (Utrecht U, The Netherlands)

Yaron Kanza (AT&T Labs – Research, USA)
Asterios Katsifodimos (Delft UT, The Netherlands)
Xiang Lian (Kent State U, USA)
Ping Lu (Beihang U, China)
Paolo Missier (Newcastle U, USA)
Davide Mottin (Aarhus U, Denmark)
Behrooz Omidvar-Tehrani (Grenoble Alpes, France)
Eric Peukert (Leipzig U, Germany)
Holger Pirk (Imperial College London, UK)
Dimitris Plexousakis (Institute of CS, FORTH, Greece)
Giuseppe Polese (U Salerno, Italy)
Arnau Prat (U Politècn. Catalunya, Spain)
Mammad Sadoghi (UC Davis)
Carlo Sartiani (U della Basilicata, Italy)
Stefanie Scherzinger (OTH Regensburg, Germany)
Petra Selmer (Neo4j, UK)
Juan F. Sequeda (Capsenta Labs, USA)
Hala Skaf-Molli (U Nantes, France)
Kostas Stefanidis (U Tampere, Finland)
Gábor Szárnyas (Budapest U Tech. & Eco., Hungary)
Ernest Teniente (U Politècn. Catalunya, Spain)
Jens Teubner (TU Dortmund, Germany)
Farouk Toumani (U Clermont Auvergne, France)
Anthony K. H. Tung (National U Singapore)
Wendy Hui Wang (Stevens Inst. of Techn., USA)
Nikolay Yakovets (TU Eindhoven, The Netherlands)
Demetrios Zeinalipour (U Cyprus, Greece)
**Industrial/Applications Program Committee**

Roy (Sudipto) Chowdhuri (Salesforce)  
Colin Florendo (Google)  
Prasanta Ghosh (Microsoft)  
Fisnik Kastrati (Huawei)  
Martin Kaufmann (Teradata)  
Justin Levandoski (Amazon)  
Stefan Mandl (Exasol)  
Norman May (SAP SE)  
Ismail Oukid (TU Dresden)  
Fatma Ozcan (IBM)  
Orestis Polychroniou (Amazon)  
Danica Porobic (Oracle)  
Harald Schöning (Software AG)  
Peter Thawley (Amazon)  
Adrian Vogelsgesang (Tableau)  
Hannes Voigt (Neo4j)  
Till Westmann (Couchbase)  
Fanfeng Zhou (Alibaba)

**Demonstration Program Committee**

Ashvin Agrawal (Microsoft Research, USA)  
Toshiyuki Amagasa (U Tsukuba, Japan)  
Siheun Amer-Yahia (CNRS LIG, France)  
Sourav Bhowmick (Nanyang Techn. U, Singapore)  
Carsten Binnig (TU Darmstadt, Germany)  
Chee-Yong Chan (National U Singapore)  
Adriane Chapman (U Southampton, UK)  
Dong-Wan Choi (Inha U, South Korea)  
Rick Cole (Tableau Software, Inc., USA)  
Vasiliki Kalavri (ETH Zurich, Switzerland)  
Romans Kasperovics (SAP, Germany)  
Wolfgang Lehner (TU Dresden, Germany)  
Hannes Mühleisen (CWI, Netherlands)  
Hubert Naacke (Sorbonne U, France)  
Keisuke Nakano (U Tohoku, Japan)  
Jorge Pérez (U Chile)  
Antonella Poggi (Sapienza U Roma, Italy)  
Cristian Riveros (Pontificia U Catolica Chile)  
Jagan Sankaranarayanan (Google, USA)  
Yuya Sasaki (Osaka U, Japan)  
Michael Schmidt (Amazon, USA)  
Slawek Staworko (U Lille 3, France)  
Dimitri Surinx (Hasselt U, Belgium)  
Alex Thomo (U Victoria, Canada)  
Xiaolan Wang (Megagon Labs, USA)  
Raymond Chi-Wing Wong (Hong Kong UST)  
Meihui Zhang (Beijing Inst. Techn., China)
Conference Organization

General Chairs
Yongluan Zhou, University of Copenhagen, Denmark
Marcos Antonio Vaz Salles, University of Copenhagen, Denmark

EDBT Program Chair
Angela Bonifati, Lyon 1 University, France

ICDT Program Chair
Carsten Lutz, University of Bremen, Germany

EDBT Industrial/Application Chair
Alexander Böhm, SAP Research, Germany

EDBT Demonstrations Chair
George Fletcher, Eindhoven University of Technology, The Netherlands

Tutorial Chair
Dan Olteanu, University of Oxford, United Kingdom

Workshops Chair
Alex Poulovassilis, Birbeck University of London, United Kingdom

EDBT Proceedings Chair
Arijit Khan, Nanyang Technological University, Singapore

ICDT Proceedings Chair
Jean Christoph Jung, University of Bremen, Germany

Local Executive Chair
Bin Yang, Aalborg University, Denmark

Sponsorship Chairs
Boris Düdder, University of Copenhagen, Denmark

Publicity Chair
Davide Mottin, Aarhus University, Denmark

Finance Chair
Nanna Højholt, University of Copenhagen, Denmark

Website Chair
Yiwen Wang, University of Copenhagen, Denmark
Test-of-Time Award

Established in 2014, the Test-of-Time Award awarded by the Extended Database Technology (EDBT) Conference recognizes papers presented at EDBT Conferences that have had the most impact in terms of research, methodology, conceptual contribution, or transfer to practice over the past ten years. The EDBT 2020 Test of Time Award committee was formed by Anastasia Ailamaki (Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland), Tamer Oszu (University of Waterloo, Canada), and Angela Bonifati (Lyon 1 University, France).

After careful consideration, the committee has decided to select the following paper from the EDBT 2010 conference as the EDBT ToT Award winner for 2020:

**Optimizing joins in a map-reduce environment**

by Foto Afrati and Jeff Ullman


This paper presented optimization strategies for executing multi-way joins in a map-reduce environment. It focused on large-scale data and provided algorithms to choose the number of map-keys and shares in order to minimize the communication cost among the map and reduce processes. The committee members agreed that this paper clearly pioneered the field of join processing in map-reduce environments. It has triggered substantial follow-up research and impact on big data processing in parallel and distributed architectures.

The EDBT Test-of-Time award for 2020 will be presented during the EDBT/ICDT 2020 Conference in Copenhagen, Denmark, as part of the Awards session on March 31, 2020.
Best Paper Award

The EDBT 2020 Best Paper Award committee was formed by Wook-shin Han (Postech, Korea), Erhard Rahm (University of Leipzig, Germany), Nesime Tatbul (Intel & MIT, USA), and Angela Bonifati (Lyon 1 University, France). After careful consideration, the committee has decided to select the following paper as the EDBT Best Paper for 2020:

Provenance for Probabilistic Logic Programs
by Shaobo Wang, Hui Lyu, Jaichi Zhang, Chenyuan Wu, Xinyi Chen, Wenchao Zhou, Boon Thau Loo, Susan B. Davidson, Chen Chen.
DOI: 10.5441/002/edbt.2020.14

Abstract: Despite the emergence of probabilistic logic programming (PLP) languages for data driven applications, there are currently no debugging tools based on provenance for PLP programs. In this paper, we propose a novel provenance model and system, called P3 (Provenance for Probabilistic logic Programs) for analyzing PLP programs. P3 enables four types of provenance queries: traditional explanation queries, queries for finding the set of most important derivations within an approximate error, top-K most influential queries, and modification queries that enable us to modify tuple probabilities with fewest modifications to program or input data. We apply these queries into real-world scenarios and present theoretical analysis and practical algorithms for such queries. We have developed a prototype of P3, and our evaluation on real-world data demonstrates that the system can support a wide-range of provenance queries with explainable results. Moreover, the system maintains provenance and executes queries efficiently with low overhead.

The EDBT Best Paper Awards for 2020 will be presented during the EDBT/ICDT 2020 Conference in Copenhagen, Denmark, as part of the Awards session on March 31, 2020.
Table of Contents

Foreword ......................................................... i
Program Committee Members ................................... ii
Conference Organization ....................................... iv
Test-of-Time Award ............................................ v
Best Paper Award ............................................. vi
Table of Contents ............................................. vii

Research Papers

Automatic Canonical Utterance Generation for Task-Oriented Bots from API Specifications
Mohammad-Ali Yaghoub-Zadeh-Fard, Boualem Benatallah, Shayan Zamanirad ........................................... 1

Efficient Continuous Multi-Query Processing over Graph Streams
Lefteris Zervakis, Vinay Setty, Christos Tryfonopoulos, Katja Hose ................................................................. 13

Zooming Out on an Evolving Graph
Amir Aghasadeghi, Vera Moffitt, Sebastian Schelter, Julia Stoyanovich ............................................................ 25

Efficient Search for Multi-Scale Time Delay Correlations in Big Time Series Data
Nguyen Ho, Torben Bach Pedersen, Van Long Ho, Mai Vu ................................................................. 37

Dynamic Query Refinement for Interactive Data Exploration
Alexander Kalinin, Ugur Cetintemel, Zheguang Zhao, Stanley Zdonik ............................................................ 49

Micro Analysis to Enable Energy-Efficient Database Systems
Chen Yang, Yongjie Du, Zhihui Du, Meng Xiaofeng ................................................................. 61

Q-Store: Distributed, Multi-partition Transactions via Queue-oriented Execution and Communication
Thamir Qadah, Suyash Gupta, Mohammad Sadoghi ......................................................................................... 73

Ensemble Grammar Induction For Detecting Anomalies in Time Series
Yifeng Gao, Jessica Lin, Constantin Brif ............................................................................................................ 85

Incremental Based Framework for Efficient Top-K Similarity Search in Interactive Data Analysis Sessions
Amit Somech, Tova Milo, Oded Elbaz ................................................................. 97

pgFMU: Integrating Data Management with Physical System Modelling
Olga Rybnytska, Laurynas Šikšnys, Torben Bach Pedersen, Bijay Neupane ......................................................... 109

Differentially-Private Next-Location Prediction with Neural Networks
Ritesh Ahuja, Gabriel Ghinita, Cyrus Shahabi .................................................................................................... 121

Explaining Differences Between Unaligned Table Snapshots
Manuel Fink, Christian Meilicke, Heiner Stuckenschmidt .................................................................................. 133

Provenance for Probabilistic Logic Programs
Shaobo Wang, Hai Lyu, Jiachi Zhang, Cheryuan Wu, Xinyi Chen, Wencao Zhou, Boon Thau Loo, Susan B. Davidson, Chen Chen ........................................................................................................ 145

Improved Cardinality Estimation by Learning Queries Containment Rates
Rojeh Hayek, Oded Shmueli ................................................................................................................................. 157

Automated Management of Indexes for Dataflow Processing Engines in IaaS Clouds
Herald Kllapi, Ilia Pietri, Verena Kantere, Yannis Ioannidis .................................................................................. 169
<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimal Histograms with Outliers</td>
<td>Rachel Behar, Sara Cohen</td>
<td>181</td>
</tr>
<tr>
<td>BalanSiNG: Fast and Scalable Generation of Realistic Signed Networks</td>
<td>Jinhong Jung, Ha-Myung Park, U Kang</td>
<td>193</td>
</tr>
<tr>
<td>Distributed Similarity Joins over Top-K Rankings</td>
<td>Evica Milchevski, Sebastian Michel</td>
<td>205</td>
</tr>
<tr>
<td>MV-PBT: Multi-Version Indexing for Large Datasets and HTAP Workloads</td>
<td>Christian Riegger, Tobias Vincon, Robert Gottstein, Iliia Petrov</td>
<td>217</td>
</tr>
<tr>
<td>Lineage-Preserving Anonymization of the Provenance of Collection-Based Workflows</td>
<td>Khalid Belhajjame</td>
<td>229</td>
</tr>
<tr>
<td>Sharing Computations for User-Defined Aggregate Functions</td>
<td>Chao Zhang, Farouk Toumani</td>
<td>241</td>
</tr>
<tr>
<td>Tracing nested data with structural provenance for big data analytics</td>
<td>Ralf Diestelkämper, Melanie Herschel</td>
<td>253</td>
</tr>
<tr>
<td>A Parallel and Distributed Approach for Diversified Top-k Best Region Search</td>
<td>Hamid Shahrivari, Matthais Olma, Odysseas Papapetrou, Dimitrios Skoutas, Anastasia Ailamaki</td>
<td>265</td>
</tr>
<tr>
<td>Data Curation with Deep Learning</td>
<td>Saravanan Thirumuruganathan, Nan Tang, Mourad Ouzzani, Anhai Doan</td>
<td>277</td>
</tr>
<tr>
<td>Fairness in Clustering with Multiple Sensitive Attributes</td>
<td>Savitha Abraham, Deepak P, Sowmya S Sundaram</td>
<td>287</td>
</tr>
<tr>
<td>Ontology-Based RDF Integration of Heterogeneous Data</td>
<td>Maxime Buron, Francois Goadsoue, Ioana Manolescu, Marie-Laure Mugnier</td>
<td>299</td>
</tr>
<tr>
<td>Manually Detecting Errors for Data Cleaning Using Adaptive Crowdsourcing Strategies</td>
<td>Haojun Zhang, Chengliang Chai, Anhai Doan, Paris Koutris, Esteban Arcaute</td>
<td>311</td>
</tr>
<tr>
<td>Publishing Video Data with Indistinguishable Objects</td>
<td>Han Wang, Yuan Hong, Yu Kong, Jaideep Vaidya</td>
<td>323</td>
</tr>
<tr>
<td>Efficient PrefDiv Algorithms for Effective Top-k Result Diversification</td>
<td>Xiaoyu Ge, Panos Chrysanthis</td>
<td>335</td>
</tr>
<tr>
<td>Adaptive Main-Memory Indexing for High-Performance Point-Polygon Joins</td>
<td>Andreas Kipf, Harald Lang, Varun Pandey, Raul Alexandru Persa, Christoph Annesser, Eleni Tzirita Zacharatu, Harish Doraiswamy, Peter Boncz, Thomas Neumann, Alfons Kemper</td>
<td>347</td>
</tr>
</tbody>
</table>

**Short papers**

<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>DISGD: A Distributed Shared-nothing Matrix Factorization for Large Scale Online Recommender Systems</td>
<td>Heidy Hazem, Ahmed Awad, Ahmed Hassan, Sherif Sakr</td>
<td>359</td>
</tr>
<tr>
<td>Forming Compatible Teams in Signed Networks</td>
<td>Ioannis Kouvatis, Konstantinos Semertzidis, Maria Zerva, Evaggelia Pitoura, Panayiotis Tsaparas</td>
<td>363</td>
</tr>
<tr>
<td>A Learning Based Approach to Predict Shortest-Path Distances</td>
<td>Jianzhong Qi, Wei Wang, Rui Zhang, Zhuowei Zhao</td>
<td>367</td>
</tr>
<tr>
<td>Optimizing Data Movement with Near-Memory Acceleration of In-memory DBMS</td>
<td>Donghun Lee, Andrew Chang, Minseon Ahn, Jongmin Gim, Jungmin Kim, Jeumin Jung, Kangwoo Choi, Vincent Pham, Oliver Rebholz, Krishna Malladi, Yang-Seok Ki</td>
<td>371</td>
</tr>
<tr>
<td>ODSA: Open Database Storage Access</td>
<td>James Wagner, Alexander Rasin, Dai Hai Ton That, Tanu Malik, Jonathan Grier</td>
<td>375</td>
</tr>
</tbody>
</table>
An Integrated Graph Neural Network for Supervised Non-obvious Relationship Detection in Knowledge Graphs
Phillipp Müller, Xiao Qin, Balaji Ganesan, Nasrullah Sheikh, Berthold Reinwald .......................... 379

Outlier detection in multivariate functional data based on a geometric aggregation
Clément Lejeune, Josiane Mothe, Olivier Teste ............................................................................. 383

REMI: Mining Intuitive Referring Expressions on Knowledge Bases
Luis Galárraga, Julien Delaunay, Jean-Louis Dessalles .................................................................. 387

The Case for Hybrid Succinct Data Structures
Christoph Anneser, Andreas Kipf, Harald Lang, Thomas Neumann, Alfons Kemper .......................... 391

FairPrep: Promoting Data to a First-Class Citizen in Studies on Fairness-Enhancing Interventions
Sebastian Schelter, Yuxuan He, Jatin Khilnani, Julia Stoyanovich .................................................... 395

Partially Materializable Delta Trees for Efficient Data Wrangling of Semi-Structured Contents
Nico Schäfer, Sebastian Michel ......................................................................................................... 399

Towards Fine-Grained Data Access Control Through Active Peer Probing
Yael Amsterdamer, Osnat Drien ....................................................................................................... 403

The ML-Index: A Multidimensional, Learned Index for Point, Range, and Nearest-Neighbor Queries
Angjela Davitkova, Evica Milchevski, Sebastian Michel .................................................................. 407

Retro: Relation Retrofitting For In-Database Machine Learning on Textual Data
Michael Günther, Maik Thiele, Wolfgang Lehner .............................................................................. 411

Revisiting the Theory and Practice of Database Cracking
Fatemeh Zardbani, Peyman Afshani, Panagiotis Karras .................................................................. 415

Boosting Blocking Performance in Entity Resolution Pipelines: Comparison Cleaning using Bloom Filters
Leonardo Gazzarri, Melanie Herschel ............................................................................................... 419

Disco: Efficient Distributed Window Aggregation
Lawrence Benson, Philipp M. Grulich, Steffen Zeuch, Volker Markl, Tilmann Rabl ......................... 423

Explaining Missing Query Results in Natural Language
Daniel Deutch, Nave Frost, Amir Gilad, Tomer Haimovich ............................................................... 427

A Context-based Approach for Partitioning Big Data
Sara Migliorini, Alberto Belussi, Elisa Quintarelli, Damiano Carra .................................................. 431

SlideSide: A fast Incremental Stream Processing Algorithm for Multiple Queries
Georgios Theodorakis, Peter Pietzuch, Holger Pirk ........................................................................... 435

Efficient Enumeration of Four Node Graphlets at Trillion-Scale
Yudi Santoso, Venkatesh Srinivasan, Alex Thomo .......................................................................... 439

Band Joins for Interval Data
Panagiotis Bouros, Konstantinos Lampropoulos, Dimitrios Tsitsigkos, Nikos Mamoulis, Manolis Terrovitis .............................................................. 443

Personalized Page Rank on Knowledge Graphs: Particle Filtering is all you need!
Denis Gallo, Matteo Lissandrini, Yannis Velegrakis ........................................................................ 447

Towards Planning of Regular Queries with Memory
Thomas Mulder, Nikolay Yakovets, George Fletcher ........................................................................ 451

Accurate Demand Forecasting for Retail with Deep Neural Networks
Shanhe Liao, Weixiong Rao ............................................................................................................ 455

Efficient Skyline Computation in High-Dimensionality Domains
Rui Liu, Dominique Li ...................................................................................................................... 459
Industry and Applications Papers

Entity Matching with Transformer Architectures - A Step Forward in Data Integration
Ursin Brunner, Kurt Stockinger ......................................................... 463

Gallery: A Machine Learning Model Management System at Uber
Chong Sun, Nader Azari, Chintan Turakhia ....................................... 474

Diverse User Selection for Opinion Procurement
Yael Amsterdamer, Oded Goldreich ................................................. 486

Elastic Scaling in VectorH
Steffen Kläbe, Kai-Uwe Sattler, Stephan Baumann, Michael Rink ........ 498

Fairness in Online Jobs: A Case Study on TaskRabbit and Google
Sihem Amer-Yahia, Shady Elbassuoni, Ahmad Ghizzawi, Ria Mae Borromeo, Emilie Hoareau, Philippe Mulhem .... 510

Inventory Reduction via Maximal Coverage in E-Commerce
Shay Gershtein, Tova Milo, Slava Novgorodov .................................. 522

Cost Estimation Across Heterogeneous SQL-Based Big Data Infrastructures in Teradata IntelliSphere
Kassem Awada, Mohammed Eltabakh, Conrad Tang, Mohammed Al-Kateb, Sanjay Nair, Grace Au ................... 534

Fast Entropy Maximization for Selectivity Estimation of Conjunctive Predicates on CPUs and GPUs
Diego Havenstein, Peter Lysakovski, Norman May, Guido Moerkotte, Gabrielle Steidl .................................. 546

Weaving Enterprise Knowledge Graphs: The Case of Company Ownership Graphs
Paolo Atzeni, Luigi Bellomarini, Michela Iezzi, Emanuel Sallinger, Adriano Vlad ........................................... 555

Expanding Query Answers on Medical Knowledge Bases
Chuan Lei, Vasilis Efthymiou, Rebecca Geis, Fatma Ozcan .................. 567

Demonstrations

VAP: A Visual Analysis Tool for Energy Consumption Spatio-temporal Pattern Discovery
Xiufeng Liu, Zhibin Niu, Yanyan Yang, Junqi Wu, Dawei Cheng, Xin Wang ......................................................... 579

Chronos: The Swiss Army Knife for Database Evaluations
Marco Vogt, Alexander Stiemer, Sein Coray, Heiko Schuldte ................ 583

skyex: an R Package for Entity Linkage
Suela Isaj, Torben Bach Pedersen .................................................. 587

Data Quality Checking for Machine Learning with MeSQuaL
Ugo Comignani, Noël Novelli, Laure Berti-Equille .............................. 591

MALOS: A Movement-Aware Location Selection System
Di Yang, Hui Li, Meng Wang, Dan Li, Jiangtao Cui .............................. 595

RRAMEN: An Interactive Tool for Evaluating Choices and Changes in Transportation Networks
Camila Ferreira Costa, Theodoros Chondrogiannis, Mario A. Nascimento, Panagiotis Bouros ............................. 599

JedAI3̂: beyond batch, blocking-based Entity Resolution
George Papadakis, Leonidas Tsekouras, Emmanuel Thanos, Nikiforos Pittaras, Giovanni Simonini, Dimitrios Skoutas, Paul Isaris, George Giannakopoulos, Themis Palpanas, Manolis Koubarakis ................................. 603

RetroLive: Analysis of Relational Retrofitted Word Embeddings
Michael Günther, Maik Thiele, Erik Nikulski, Wolfgang Lehner ................ 607

RulER: Scaling Up Record-level Matching Rules
Luca Gagliardelli, Giovanni Simonini, Sonia Bergamaschi ...................... 611
<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schema Mapping Generation in the Wild: A Demonstration with Open Government Data</td>
<td>Lacramioara Mazilu, Nikolaos Konstantinou, Norman Paton, Alvaro A. A. Fernandes</td>
<td>615</td>
</tr>
<tr>
<td>EPIQUE: Extracting Meaningful Science Evolution Patterns from Large Document Archives</td>
<td>Ke Li, Hubert Naacke, Bernd Amann</td>
<td>619</td>
</tr>
<tr>
<td>Task-Tuning in Privacy-Preserving Crowdsourcing Platforms</td>
<td>Joris Duguépéroux, Antonin Voyez, Tristan Allard</td>
<td>623</td>
</tr>
<tr>
<td>Scaling a Public Transport Monitoring System to Internet of Things Infrastructures</td>
<td>Haralambos Gavriilidis, Adrian Michalke, Laura Mons, Steffen Zeuch, Volker Markl</td>
<td>627</td>
</tr>
<tr>
<td>Governor: Operator Placement for a Unified Fog-Cloud Environment</td>
<td>Ankit Chaudhary, Steffen Zeuch, Volker Markl</td>
<td>631</td>
</tr>
<tr>
<td>Human-in-the-Loop Schema Inference for Massive JSON Datasets</td>
<td>Mohamed-Amine Baazizi, Clément Berti, Dario Colazzo, Giorgio Ghelli, Carlo Sartiani</td>
<td>635</td>
</tr>
<tr>
<td>SQL Query Processing Using an Integrated FPGA-based Near-Data Accelerator in ReProVide</td>
<td>Lekshmi B.G., Andreas Becher, Klaus Meyer-Wegener, Stefan Wildermann, Jürgen Teich</td>
<td>639</td>
</tr>
<tr>
<td>Tutorials</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Declarative Languages for Big Streaming Data</td>
<td>Riccardo Tommasini, Sherif Sakr, Emanuele Della Valle, Hojjat Jafarpour</td>
<td>643</td>
</tr>
<tr>
<td>Entity Resolution: Past, Present and Yet-to-Come</td>
<td>George Papadakis, Ekaterini Ioannou, Themis Palpanas</td>
<td>647</td>
</tr>
<tr>
<td>Fairness in Rankings and Recommenders</td>
<td>Evaggelia Pitoura, Georgia Koutrika, Kostas Stefanidis</td>
<td>651</td>
</tr>
<tr>
<td>NoSQL Schema Evolution and Data Migration: State-of-the-Art and Opportunities</td>
<td>Uta Störl, Meike Klettke, Stefanie Scherzinger</td>
<td>655</td>
</tr>
</tbody>
</table>