
Advances in Database Technology — EDBT 2019

22nd International Conference
on Extending Database Technology
Lisbon, Portugal, March 26–29, 2019
Proceedings

Editors

Melanie Herschel
Helena Galhardas
Berthold Reinwald
Irina Fundulaki
Carsten Binnig
Zoi Kaoudi



Advances in Database Technology – EDBT 2019
Proceedings of the 22nd International Conference
on Extending Database Technology
Lisbon, Portugal, March 26–29, 2019

Series ISSN: 2367-2005

Editors

Melanie Herschel, University of Stuttgart, Germany
Helena Galhardas, INESC-ID and IST, Universidade de Lisboa, Portugal
Berthold Reinwald, IBM Research, USA
Irina Fundulaki, ICS Forth, Greece
Carsten Binnig, TU Darmstadt, Germany
Zoi Kaoudi, QCRI, HBKU, Qatar



OpenProceedings.org
University of Konstanz
University Library
78457 Konstanz, Germany

COPYRIGHT NOTICE: Copyright © 2019 by the authors of the individual papers.

Distribution of all material contained in this volume is permitted under the terms of the Creative Commons license CC-by-nc-nd 4.0

OpenProceedings ISBN: 978-3-89318-081-3

DOI of this front matter: 10.5441/002/edbt.2019.01

In memoriam

Christine Collet

1956–2019

Foreword

The International Conference on Extending Database Technology (EDBT) is an established and renowned forum for the exchange of the latest research results and advances in data management. This year, the 22nd edition of EDBT takes place in Lisbon, Portugal, from March 26 to March 29, 2019. It is jointly organized with the International Conference on Database Theory (ICDT). In a world where increasingly many aspects of our lives and society are data-driven, data management technology continues to broaden its reach and extends its tradition of contributing models, algorithms, and architectures to novel applications adapted to new hardware and software.

As in previous years, EDBT 2019 solicited contributions both on novel research results and on experience and analysis results that focus on a comprehensive and detailed performance evaluation. For the first time, EDBT 2019 further solicited papers that describe innovative systems as part of its main research track. We also continued the recently established short paper track, offering a forum to present research in progress and visionary ideas during plenary poster sessions of the conference. To complement the scientific program, EDBT further solicited demonstrations of research prototypes, descriptions of industrial and application achievements, and proposals for tutorials.

The EDBT 2019 program committee reviewed 157 full research papers, of which 36 were accepted. For short papers, 28 papers out of 122 were selected. Among the 24 submissions to the industry and application track, 8 papers were accepted. The 21 demonstrations presented at the conference were selected among 42 demonstration proposals. Finally, we accepted 3 out of 10 tutorials. All these contributions will be presented at the conference. The program additionally features five workshops, an EDBT/ICDT joint session on research challenges, and four invited EDBT/ICDT joint keynotes.

Shaping the exciting program of EDBT 2019 is the result of a large community effort, and I take this opportunity to thank all persons involved. First, I would like to thank all authors for their high-quality submissions and contributions. I also would like to thank all reviewers who served on the EDBT 2019 program committee and the chairs responsible for our different tracks, namely Berthold Reinwald (IBM, United States) who chaired the industrial and application track, Carsten Binnig (TU Darmstadt, Germany) who served as demonstration chair, our tutorial chair Irini Fundulaki (ICS FORTH, Greece), and Paolo Papotti (EUROCOM, France) who served as workshop chair. The special session on joint EDBT/ICDT research challenges was organized by Julia Stoyanovich (NYU, USA). I also thank Laura Haas (UMass Amherst, USA) and Alon Halevy, who generously accepted to serve on the Test of Time Award Committee. Many thanks also to Paolo Atzeni (Universita' Roma Tre, Italy), Wei Wang (UNSW Sydney, Australia), and Jeffrey Xu Yu (Chinese University of Hong Kong) for serving on the Best Paper Award committee.

The conference would not have been possible without the tireless effort of the general chair Helena Galhardas (INESC-ID and IST, Universidade de Lisboa, Portugal) and the local organization team. Special thanks to the finance chair Manuel J. Fonseca (Universidade de Lisboa, Portugal), the local executive chairs José Borbinha and Luís Rodrigues, the sponsorship chairs João Garcia and Miguel Pardal, the publicity chair Paolo Romano, the student volunteers chair Hugo Nicolau (all from INESC-ID and IST, Universidade de Lisboa, Portugal), and the website chair António Higgs (INESC-ID, Portugal). These proceedings have been produced thanks to our proceedings chair Zoi Kaoudi (QCRI, Qatar). Norman Paton was most helpful in advising and coordinating with the EDBT Executive Board.

I really look forward to an interesting program and exciting conference on March 26–29, 2019 and to meeting you in Lisbon.

Melanie Herschel
EDBT 2019 Program Chair

Program Committee Members

Research Program Committee

Adriane Chapman, U Southampton
Ahmed Elmagarmid, QCRI HBKU
Alekh Jindal, Microsoft
Alessandro Campi, Politecnico di Milano
Alexander Böhm, SAP SE
Alfons Kemper, TU Munich
Alfredo Cuzzocrea, ICAR-CNR and U Calabria
Amelie Marian, Rutgers U
Amr El Abbadi, UC Santa Barbara
Angela Bonifati, U Lyon 1
Anton Dignös, Free U Bozen-Bolzano
Antonios Deligiannakis, TU Crete
Arash Termehchy, Oregon State U
Arnab Bhattacharya, IIT Kanpur
Avigdor Gal, Technion
Bernhard Seeger, U Marburg
Boris Glavic, Illinois IT
Chuan Lei, IBM Research-Almaden
Daisy Zhe Wang, U Florida
Daniele Dell’Aglia, U Zurich
Dario Colazzo, LAMSADE - Universit
Davide Martinenghi, Politecnico di Milano
Demetrios Zeinalipour-Yazti, U Cyprus
Elena Ferrari, U Insubria
Elisa Quintarelli, Politecnico di Milano
Eric Lo, Chinese U Hong Kong
Ernest Teniente, U Politècnica de Catalunya
Esther Pacitti, LIRMM and INRIA U Montpellier 2
Farouk Toumani Limos, B. Pascal U Clermont-Ferrand
Felix Naumann, HPI
Francois Goasdoue, U Rennes 1
George H. L. Fletcher, Eindhoven UT
Georgia Koutrika, Athena Research Center
Goetz Graefe, Google
Gottfried Vossen, ERCIS Muenster
Guoliang Li, Tsinghua U
Haruo Yokota, Tokyo Institute of Technology
Heiko Schuldt, U Basel
Isabel Cruz, U Illinois Chicago
Jean-Marc Petit, U Lyon & INSA Lyon
Jian Pei, Simon Fraser U
Jianwen Su, UC Santa Barbara
Johann Gamper, Free U Bozen-Bolzano
Kai-Uwe Sattler, TU Ilmenau
Karl Aberer, EPF Lausanne
Katja Hose, Aalborg U
Klemens Böhm, Karlsruhe I.T.
Kostas Magoutis, U Ioannina
Kostas Stefanidis, U Tampere
Marta Patino, U Politècnica de Madrid
Martin Theobald, U Luxembourg
Matthias Boehm, IBM
Matthias Renz, George Mason U
Maya Ramanath, IIT Delhi
Michael Grossniklaus, U Konstanz
Minos Garofalakis, TU Crete
Mohammad Sadoghi UC Davis
Nicoleta Preda, U Versailles
Nikos Mamoulis, U Ioannina
Nikos Pelekis, U Piraeus
Panagiotis Karras, Aarhus U
Paolo Missier, Newcastle U
Periklis Andritsos, U Toronto
Peter Triantafillou, U Warwick
Riccardo Torlone, U Roma Tre
Sabrina De Capitani di Vimercati, U Milan
Sherif Sakr, U New South Wales
Stefan Deßloch, TU Kaiserslautern
Tore Risch, Uppsala U
Vassilis Christophides, INRIA
Verena Kantere, U Geneva
Will Wu, Google
Wolfgang Lehner, TU Dresden
Xiaochun Yang, Northeastern U
Xiaofang Zhou, U Queensland
Xin Huang, Hong Kong Baptist U
Yang Cao, U Edinburgh
Yannis Theodoridis, U Piraeus
Yannis Velegarakis, U Trento
Yaron Kanza, AT&T Labs - Research
Yi-Leh Wu, Nat. Taiwan U Sc. & Tech.
Zhifeng Bao, RMIT U
Ziawasch Abedjan, TU Berlin

Industrial/Applications Program Committee

Alkis Simitsis, HP Labs
Amol Ghoting, LinkedIn
Andrey Balmin, Workday
Anil Goel, SAP Canada
Benjamin Schlegel, Oracle Labs
Calisto Zuzarte, IBM Toronto
Cliff Leung, Huawei
David Simmen, Apple
Martin Grund, Amazon AWS
Michael Rys, Microsoft
Nesreen Ahmed, Intel
Niketan Pansare, IBM
Norbert Ritter, U Hamburg
Paul Brown, Teradata Corp
Songting Chen, Tigergraph
Yannis Sismanis, Google

Demonstration Program Committee

Alexander Böhm, SAP SE
Arvind Arasu, Microsoft
Dirk Habich, TU Dresden
Jana Gicheva, Imperial College London
Manos Athanassoulis, Boston U
Martin Theobald, U Luxembourg
Michael Gertz, Heidelberg U
Rick Cole, Tableau Software
Sanjay Krishnan, U Chicago
Stefan Manegold, CWI
Sven Helmer, Free U Bozen-Bolzano
Thibault Sellam, Columbia U
Tilman Rabl, TU Berlin
Uwe Roehm, U Sydney
Viktor Leis, TU Munich
Ziawasch Abedjan, TU Berlin

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 2019 Test-of-Time Award committee was formed by Laura Haas (University of Massachusetts, USA), Alon Halevy, and Melanie Herschel, the EDBT 2019 PC chair. The committee was charged with selecting a paper from the EDBT 2009 Proceedings.

After careful consideration, the Test-of-Time Award committee decided for the following paper from the 2009 EDBT Conference held in Saint Petersburg, Russia to receive the award:

Shore-MT: a scalable storage manager for the multicore era

by Ryan Johnson, Ippokratis Pandis, Nikos Hardavellas, Anastasia Ailamaki, and Babak Falsafi

published in the EDBT 2009 Proceedings, pp. 24–35, DOI: 10.1145/1516360.1516365.

The committee members agreed that this paper clearly stands out in terms of methodology, impact, and influence. It has catalyzed and enabled substantial follow-up research and has demonstrated its high relevance to industry.

Abstract:

Historically, database engines focused on the ability to efficiently overlap many requests over a small number processor cores, with I/O latencies and scalability as the main design driver. However, the advent of increasingly multicore hardware circa 2000 brought new challenges because concurrent transactions begin to stress the limits of the storage manager’s thread scalability by accessing its internal structures simultaneously and in large numbers. This EDBT 2009 paper shows the results of experiments running benchmarks on four (then and still) popular open-source storage managers (Shore, BerkeleyDB, MySQL, and PostgreSQL) on a multi-core machine. The results show that all systems suffer from scalability bottlenecks at the storage engine level. From that research emerged Shore-MT, an open-source multithreaded and highly scalable storage manager, built with Shore as a base. We learned that designers should favor scalability over single-thread performance, and we identified several other key principles for architecting scalable storage engines.

Ten years later, Shore-MT work has concluded, although the system still serves as a research platform in the space. Meanwhile, research on transaction processing scalability continues to mature, the move to main-memory transaction processing and their higher TPS increased the need for scalable storage managers, while the popular open-source systems, such as MySQL and PostgreSQL, significantly improved their scalability. In particular, a significant amount of research and industrial developments in the ten years since the Shore-MT paper focused on improving the scalability of individual components of a storage manager, such as latches, the logging subsystem and access methods. This research was partly carried out by our research group as follow-on work, but other research groups and database vendors have made important contributions as well. Another significant amount of effort has focused on scalable concurrency control protocols, again both within and outside our research group. The knowledge that we have gained from building Shore-MT has been invaluable in maintaining scalability in this new, multi-dimensional ecosystem.

The EDBT Test-of-Time award for 2019 will be presented during the EDBT/ICDT 2019 Conference as part of the Awards session on Wednesday, March 27, 2019, by Anastasia Ailamaki (EPFL, Switzerland).

Table of Contents

Foreword	i
Program Committee Members	ii
Test-of-Time Award	iv
Table of Contents	v
Research Papers	
Umzi: Unified Multi-Zone Indexing for Large-Scale HTAP <i>Chen Luo, Pinar Tozun, Yuanyuan Tian, Ronald Barber, Vijayshankar Raman, Richard Sidle</i>	1
A Highly Scalable Labelling Approach for Exact Distance Queries in Complex Networks <i>Muhammad Farhan, Qing Wang, Yu Lin, Brendan Mckay</i>	13
Stratified Random Sampling over Streaming and Stored Data <i>Trong Nguyen, Ming-Hung Shih, Divesh Srivastava, Srikanta Tirthapura, Bojian Xu</i>	25
Boosting SimRank with Semantics <i>Brit Youngmann, Tova Milo, Amit Somech</i>	37
Leveraging Bitmap Indexing for Subgraph Searching <i>David Luaces, José R.R. Viqueira, Tomás F. Pena, José M. Cotos</i>	49
Spec-QP: Speculative Query Planning for Joins over Knowledge Graphs <i>Madhulika Mohanty, Maya Ramanath, Mohamed Yahya, Gerhard Weikum</i>	61
Iterative Estimation of Mutual Information with Error Bounds <i>Michael Vollmer, Klemens Böhm</i>	73
Functional Geometric Monitoring for Distributed Streams <i>Vasileios Samoladas, Minos Garofalakis</i>	85
Efficient Window Aggregation with General Stream Slicing. <i>Jonas Traub, Philipp M. Grulich, Alejandro Rodriguez Cuellar, Sebastian Breß, Asterios Katsifodimos, Tilmann Rabl, Volker Markl</i>	97
An Efficient Sliding Window Approach for Approximate Entity Extraction with Synonyms <i>Jin Wang, Chunbin Lin, Mingda Li, Carlo Zaniolo</i>	109
Attendance Maximization for Successful Social Event Planning <i>Nikos Bikakis, Vana Kalogeraki, Dimitrios Gunopoulos</i>	121
GroupTravel: Customizing Travel Packages for Groups <i>Sihem Amer-Yahia, Shady Elbassuoni, Behrooz Omidvar-Tehrani, Ria Mae Borromeo, Mehrdad Farokhnejad</i>	133
SEP2P: Secure and Efficient P2P Personal Data Processing <i>Julien Loudet, Iulian Sandu-Popa, Luc Bouganim</i>	145
Indexing Trajectories for Travel-Time Histogram Retrieval <i>Robert Waury, Christian S. Jensen, Satoshi Koide, Yoshiharu Ishikawa, Chuan Xiao</i>	157
BB-Tree: A practical and efficient main-memory index structure for multidimensional workloads <i>Stefan Sprenger, Patrick Schäfer, Ulf Leser</i>	169
Semantic and Influence aware k-Representative Queries over Social Streams <i>Yanhao Wang, Yuchen Li, Kian-Lee Tan</i>	181

Comparative Analysis of Content-based Personalized Microblog Recommendations <i>Efi Karra Taniskidou, George Papadakis, George Giannakopoulos, Manolis Koubarakis</i>	193
Crowdsourced Truth Discovery in the Presence of Hierarchies for Knowledge Fusion <i>Woohwan Jung, Younghoon Kim, Kyuseok Shim</i>	205
A Utility-Preserving and Scalable Technique for Protecting Location Data with Geo-Indistinguishability <i>Ritesh Ahuja, Gabriel Ghinita, Cyrus Shahabi</i>	217
Inves: Incremental Partitioning-Based Verification for Graph Similarity Search <i>Jongik Kim, Dong-Hoon Choi, Chen Li</i>	229
Flow Motifs in Interaction Networks <i>Chrysanthi Kosyfaki, Nikos Mamoulis, Evaggelia Pitoura, Panayiotis Tsaparas</i>	241
DynFD: Functional Dependency Discovery in Dynamic Datasets <i>Philipp Schirmer, Thorsten Papenbrock, Sebastian Kruse, Felix Naumann, Dennis Hempfing, Torben Mayer, Daniel Neuschäfer-Rube</i>	253
CLX: Towards verifiable PBE data transformation <i>Zhongjun Jin, Michael Cafarella, H. V. Jagadish, Sean Kandel, Michael Minar, Joseph M. Hellerstein</i>	265
Reverse-Engineering Conjunctive Queries from Provenance Examples <i>Daniel Deutch, Amir Gilad</i>	277
A Six-dimensional Analysis of In-memory Aggregation <i>Puya Memarzia, Suprio Ray, Virendra C. Bhavsar</i>	289
BionicDB: Fast and Power-Efficient OLTP on FPGA <i>Kangnyeon Kim, Ryan Johnson, Ippokratis Pandis</i>	301
Hyrise Re-engineered: An Extensible Database System for Research in Relational In-Memory Data Management <i>Markus Dreseler, Jan Kossmann, Martin Boissier, Stefan Klauck, Matthias Uflacker, Hasso Plattner</i>	313
Efficient Computation of Probabilistic Core Decomposition at Web-Scale <i>Fatemeh Esfahani, Venkatesh Srinivasan, Alex Thomo, Kui Wu</i>	325
Efficient Network Reliability Computation in Uncertain Graphs <i>Yuya Sasaki, Yasuhiro Fujiwara, Makoto Onizuka</i>	337
Scalable Parallelization of RDF Joins on Multicore Architectures <i>Dimitris Bilidas, Manolis Koubarakis</i>	349
Pivoted Subgraph Isomorphism: The Optimist, the Pessimist and the Realist <i>Ehab Abdelhamid, Ibrahim Abdelaziz, Zuhair Khayyat, Panos Kalnis</i>	361
MinoanER: Schema-Agnostic, Non-Iterative, Massively Parallel Resolution of Web Entities <i>Vasilis Efthymiou, George Papadakis, Kostas Stefanidis, Vassilis Christophides</i>	373
Extending Cross-Domain Knowledge Bases with Long Tail Entities using Web Table Data <i>Yaser Oulabi, Christian Bizer</i>	385
Continuous Deployment of Machine Learning Pipelines <i>Behrouz Derakhshan, Alireza Rezaei Mahdiraji, Tilmann Rabl, Volker Markl</i>	397
Discovering Order Dependencies through Order Compatibility <i>Cristian Consonni, Paolo Sottovia, Alberto Montresor, Yannis Velegarakis</i>	409
Scalable Kernel Density Estimation-based Local Outlier Detection over Large Data Streams <i>Xiao Qin, Lei Cao, Elke Rundensteiner, Samuel Madden</i>	421

Tutorials

RDF graph summarization: principles, techniques and applications <i>Haridimos Kondylakis, Dimitris Kotzinos, Ioana Manolescu</i>	433
Schemas And Types For JSON Data <i>Mohamed-Amine Baazizi, Dario Colazzo, Giorgio Ghelli, Carlo Sartiani</i>	437
Influence Maximization Revisited: The State of the Art and the Gaps that Remain <i>Akhil Arora, Sainyam Galhotra, Sayan Ranu</i>	440
Industry and Applications Papers	
Finding Meaningful Contrast Patterns for Quantitative Data <i>Rohan Khade, Jessica Lin, Nital Patel</i>	444
Predicting "What is Interesting" by Mining Interactive-Data-Analysis Session Logs <i>Amit Somech, Tova Milo, Chai Ozeri</i>	456
Hidden Layer Models for Company Representations and Product Recommendations <i>Katsiaryna Mirylenka, Paolo Scotton, Christoph Miksovic, Jeff Dillon</i>	468
Big POI data integration with Linked Data technologies <i>Spiros Athanasiou, Giorgos Giannopoulos, Damien Graux, Nikos Karagiannakis, Jens Lehmann, Axel-Cyrille Ngonga Ngomo, Kostas Patrourmpas, Mohamed Ahmed Sherif, Dimitrios Skoutas</i>	477
Executing Entity Matching End to End: A Case Study <i>Pradap Konda, Sanjay Seshadri, Elan Segarra, Brent Hueth, Anhai Doan</i>	489
The Copernicus App Lab project: Easy Access to Copernicus Data <i>Konstantina Bereta, Herve Caumont, Ulrike Daniels, Erwin Goor, Manolis Koubarakis, Despina-Athanasia Pantazi, George Stamoulis, Sam Ubels, Valentijn Venus, Firman Wahyudi</i>	501
Modeling and Building IoT Data Platforms with Actor-Oriented Databases <i>Yiwen Wang, Julio Cesar Dos Reis, Kasper Myrtue Borggren, Marcos Antonio Vaz Salles, Claudia Bauzer Medeiros, Yongluan Zhou</i>	512
KSQL: Streaming SQL Engine for Apache Kafka <i>Hojjat Jafarpour, Rohan Desai</i>	524
Demonstrations	
The Power of SQL Lambda Functions <i>Maximilian Schüle, Dimitri Vorona, Linnea Passing, Harald Lang, Alfons Kemper, Stephan Günemann, Thomas Neumann</i>	534
MINARET: A Recommendation Framework for Scientific Reviewers <i>Sherif Sakr, Mohamed Ragab, Mohamed Maher, Ahmed Awad</i>	538
SCube: A Tool for Segregation Discovery <i>Alessandro Baroni, Salvatore Ruggieri</i>	542
SparkTune: tuning Spark SQL through query cost modeling <i>Enrico Gallinucci, Matteo Golfarelli</i>	546
HOTMapper: Historical Open Data Table Mapper <i>Henrique Varella Ehrenfried, Rudolf Eckelberg, Hamer Iboshi, Eduardo Todt, Daniel Weingaertner, Marcos Didonet Del Fabro</i>	550
SmartML: A Meta Learning-Based Framework for Automated Selection and Hyperparameter Tuning for Machine Learning Algorithms <i>Mohamed Maher, Sherif Sakr</i>	554

devUDF: Increasing UDF development efficiency through IDE Integration. It works like a PyCharm! <i>Mark Raasveldt, Pedro Holanda, Stefan Manegold</i>	558
ML2SQL - Compiling a Declarative Machine Learning Language to SQL and Python <i>Maximilian Schüle, Matthias Bungeroth, Dimitri Vorona, Alfons Kemper, Stephan Günemann, Thomas Neumann</i>	562
Incremental structural summarization of RDF graphs <i>Francois Goasdoue, Paweł Guzewicz, Ioana Manolescu</i>	566
VISTA: A visual analytics platform for semantic annotation of trajectories <i>Amilcar Soares, Jordan Rose, Mohammad Etemad, Chiara Renso, Stan Matwin</i>	570
SLIPO: Large-Scale Data Integration for Points of Interest <i>Spiros Athanasiou, Michail Alexakis, Giorgos Giannopoulos, Nikos Karagiannakis, Yannis Kouvaras, Pantelis Mitropoulos, Kostas Patrourmpas, Dimitrios Skoutas</i>	574
A Map Search System based on a Spatial Query Language <i>Yuanyuan Wang, Panote Siritiraya, Haruka Sakara, Yukiko Kawai, Keishi Tajima</i>	578
FaiRank: An Interactive System to Explore Fairness of Ranking in Online Job Marketplaces <i>Ahmad Ghizzawi, Julien Marinescu, Shady Elbassuoni, Sihem Amer-Yahia, Gilles Bisson</i>	582
MM-evolver: A Multi-model Evolution Management Tool <i>Michal Vavrek, Irena Holubova, Stefanie Scherzinger</i>	586
Resense: Transparent Record and Replay of Sensor Data in the Internet of Things <i>Dimitrios Giouroukis, Julius Hülsmann, Janis von Bleichert, Morgan Geldenhuys, Tim Stullich, Felipe Gutierrez, Jonas Traub, Kaustubh Beedkar, Volker Markl</i>	590
Improving Named Entity Recognition using Deep Learning with Human in the Loop <i>Ticiano Coelho Da Silva, Régis Magalhães, José de Macêdo, David Araújo, Natanael Araújo, Vinicius de Melo, Pedro Olímpio, Paulo Rego, Aloisio Neto</i>	594
Demonstrating data collections curation and exploration with CURARE <i>Genoveva Vargas-Solar, Gavin Kemp, Irving Hernandez Gallegos, Javier-Alfonso Espinosa-Oviedo, Catarina Ferrera da Silva, Parisa Ghodous</i>	598
SparkER: Scaling Entity Resolution in Spark <i>Luca Gagliardelli, Giovanni Simonini, Domenico Beneventano, Sonia Bergamaschi</i>	602
Exploring Interpretable Features for Large Time Series with SE4TeC <i>Jingwei Zuo, Karine Zeitouni, Yehia Taher</i>	606
HASQL: a Method of Masking System Failures <i>Mark Hannum, Adi Zaimi, Mike Ponomarenko</i>	610
Query-Driven Data Minimization with the DataEconomist <i>Peter K. Schwab, Julian Matschinske, Andreas M. Wahl, Klaus Meyer-Wegener</i>	614
Short Papers	
ITGC: Information-theoretic grid-based clustering <i>Sahar Behzadi Soheil, Hermann Hinterhauser, Claudia Plant</i>	618
Adaptive Watermarks: A Concept Drift-based Approach for Predicting Event-Time Progress in Data Streams <i>Ahmed Awad, Jonas Traub, Sherif Sakr</i>	622
Identifying Bias in Name Matching Tasks <i>Alexandros Karakasidis, Evaggelia Pitoura</i>	626
Rock - Let the points roam to their clusters themselves <i>Anna Beer, Daniyal Kazempour, Thomas Seidl</i>	630

Fast Trajectory Range Query with Discrete Frechet Distance <i>Jiahao Zhang, Bo Tang, Yiu Man Lung</i>	634
Repairing of Record Linkage: Turning Errors into Insight <i>Quyên Bui-Nguyen, Qing Wang, Jingyu Shao, Dinusha Vatsalan</i>	638
Streaming HyperCube: A Massively Parallel Stream Join Algorithm <i>Yuan Qiu, Serafeim Papadias, Ke Yi</i>	642
Exploring Fairness of Ranking in Online Job Marketplaces <i>Shady Elbassuoni, Sihem Amer-Yahia, Christine El Atie, Ahmad Ghizzawi, Bilel Oualha</i>	646
Snapshot Isolation for Transactional Stream Processing <i>Philipp Götze, Kai-Uwe Sattler</i>	650
Recurrent Neural Networks for Dynamic User Intent Prediction in Human-Database Interaction <i>Venkata Vamsikrishna Meduri, Kanchan Chowdhury, Mohamed Sarwat</i>	654
Optimal Algorithm for Profiling Dynamic Arrays with Finite Values <i>Dingcheng Yang, Wenjian Yu, Junhui Deng, Shenghua Liu</i>	658
Publishing Differentially Private Datasets via Stable Microaggregation <i>Masooma Iftikhar, Qing Wang, Yu Lin</i>	662
Range Query Processing for Monitoring Applications over Untrustworthy Clouds <i>Hoang Van Tran, Tristan Allard, Laurent D’Orazio, Amr El Abbadi</i>	666
Towards Augmented Database Schemes by Discovery of Latent Visual Attributes <i>Tomas Grosup, Ladislav Peska, Tomas Skopal</i>	670
Workload-Driven and Robust Selection of Compression Schemes for Column Stores <i>Martin Boissier, Max Jendruk</i>	674
CLRL: Feature Engineering for Cross-Language Record Linkage <i>Öykü Özlem Çakal, Mohammad Mahdavi, Ziawasch Abedjan</i>	678
Operational Stream Processing: Towards Scalable and Consistent Event-Driven Applications <i>Asterios Katsifodimos, Marios Fragkoulis</i>	682
Metropolis-Hastings Algorithms for Estimating Betweenness Centrality <i>Mostafa Haghir Chehreghani, Talel Abdessalem, Albert Bifet</i>	686
From Copernicus Big Data to Extreme Earth Analytics <i>Manolis Koubarakis, Konstantina Bereta, Dimitris Bilidas, Konstantinos Giannousis, Theofilos Ioannidis, Despina-Athanasia Pantazi, George Stamoulis, Jim Dowling, Seif Haridi, Vladimir Vlassov, Lorenzo Bruzzone, Claudia Paris, Torbjørn Eltoft, Thomas Krämer, Angelos Charalabidis, Vangelis Karkaletsis, Stasinou Konstantopoulos, Theofilos Kakantousis, Mihai Datcu, Corneliu Octavian Dumitru, Florian Appel, Heike Bach, Silke Migdall, Nick Hughes, David Arthurs, Andrew Fleming</i>	690
Neuromorphic Hardware As Database Co-Processors: Potential and Limitations <i>Thomas Heinis</i>	694
Query Driven Data Labeling with Experts: Why Pay Twice? <i>Eyal Dushkin, Shay Gershtein, Tova Milo, Slava Novgorodov</i>	698
A Galaxy of Correlations <i>Daniyal Kazempour, Lisa Krombholz, Peer Kröger, Thomas Seidl</i>	702
Insights into a running clockwork: On interactive process-aware clustering <i>Daniyal Kazempour, Thomas Seidl</i>	706
Interpolation-friendly B-trees: Bridging the Gap Between Algorithmic and Learned Indexes <i>Ali Hadian, Thomas Heinis</i>	710

SynthEdit: Format transformations by example using edit operations <i>Alex Teodor Bogatu, Alvaro A. A. Fernandes, Norman W. Paton, Nikolaos Konstantinou</i>	714
Triad Enumeration at Trillion-Scale Using a Single Commodity Machine <i>Yudi Santoso, Alex Thomo, Venkatesh Srinivasan, Sean Chester</i>	718
Fast Truss Decomposition in Large-scale Probabilistic Graphs <i>Fatemeh Esfahani, Jian Wu, Venkatesh Srinivasan, Alex Thomo, Kui Wu</i>	722
An Experimental Study on Network Immunization <i>Alvis Logins, Panagiotis Karras</i>	726