Proceedings of the VLDB Endowment

Volume 13, No. 12 – August 2020

Editors in Chief:

Magdalena Balazinska and Xiaofang Zhou

Associate Editors:

Azza Abouzied, Amr El Abbadi, Phil Bernstein, Xin Luna Dong, Zi (Helen) Huang, Nick Koudas, Georgia Koutrika, Guoliang Li, Alexandra Meliou, Felix Naumann, Dan Olteanu, M. Tamer Özsu, Aditya Parameswaran, Andy Pavlo, Xiaokui Xiao, Jeffrey Xu Yu, Meihui Zhang, Jingren Zhou

Publication Editors:

Hiroaki Shiokawa and Sen Wang

PVLDB – Proceedings of the VLDB Endowment

Volume 13, No. 12, August 2020.

PVLDB is indexed in Scopus (Elsevier) as well as covered by the following Clarivate Analytics services:

- Science Citation Index Expanded (also known as SciSearch®)
- Journal Citation Reports/Science Edition, and
- Current Contents®/Engineering Computing and Technology

All papers published in this issue will be presented at the 46th International Conference on Very Large Data Bases, Tokyo, Japan, 2020.

Copyright 2020 VLDB Endowment

This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-nd/4.0/. For any use beyond those covered by this license, obtain permission by emailing info@vldb.org.

Volume 13, Number 12, August 2020

Pages i – xvi and 2801 - 3516

ISSN 2150-8097

Available at: http://www.pvldb.org and https://dl.acm.org.

TABLE OF CONTENTS

Front Matter

Copyright Notice i
Table of Contents ii
PVLDB Organization and Review Board – Vol. 13 ix
Demonstration Track Chairs and Reviewers – Vol. 13 xii
Industrial Track Chairs and Reviewers – Vol. 13 xiii
Tutorial Track Chairs – Vol. 13 xiv

Demonstrations

Demand-based Sensor Data Gathering with Multi-Query Optimization 2801
Julius Hülsmann, Jonas Traub, Volker Markl

CheetahVIS: A Visual Analytical System for Large Urban Bus Data 2805
Wentao Ning, Qiandong Tang, Yi Zhao, Chuan Yang, Xiaofeng Wang, Teng Wang, Haotian Liu, Chaozu Zhang, Zhiyuan Zhou, Qiaomu Shen, Bo Tang

UNMASQUE: A Hidden SQL Query Extractor 2809
Kapil Khurana, Jayant Haritsa

G3: When Graph Neural Networks Meet Parallel Graph Processing Systems on GPUs 2813
Husong Liu, Shengliang Lu, Xinyu Chen, Bingsheng He

PiBench Online: Interactive Benchmarking of Persistent Memory Indexes 2817
Xiangpeng Hao, Lucas Lersch, Tianzheng Wang, Ismail Oukid

VisClean: Interactive Cleaning for Progressive Visualization 2821
Yuyu Luo, Chengliang Chai, Xuedi Qin, Nan Tang, Guoliang Li

IMO: A Toolbox for Simulating and Querying "Infected" Moving Objects 2825
Jianqiu Xu, Hua Lu, Zhifeng Bao

COUNTATA: Dataset Labeling Using Pattern Counts 2829
Yuval Moskovitch, H. Jagadish

A Demonstration of Willump: A Statistically-Aware End-to-end Optimizer for Machine Learning *Inference* 2833
Peter Kraft, Daniel Kang, Deepak Narayanan, Shoumik Palkar, Peter Bailis, Matei Zaharia

Ease.ml/snoopy in Action: Towards Automatic Feasibility Analysis for Machine Learning Application Development 2837
Cedric Renggli, Luka Rimanic, Luka Kolar, Wentao Wu, Ce Zhang

DeepTrack: Monitoring and Exploring Spatio-Temporal Data – A Case of Tracking COVID-19 2841
Yuyu Luo, Wenbo Li, Tianyu Zhao, Xiang Yu, Lixi Zhang, Guoliang Li, Nan Tang

DF-Toolkit: Interacting with Low-Level Database Storage 2845
James Wagner, Alexander Rasin, Karen Heart, Tanu Malik, Jonathan Grier

Like Water and Oil: With a Proper Emulsifier, Query Compilation and Data Parallelism Will Mix Well 2849
Henning Funke, Jens Teubner

SQUARES : A SQL Synthesizer Using Query Reverse Engineering 2853
Pedro Orvalho, Miguel Terra-Neves, Miguel Ventura, Ruben Martins, Vasco Manquinho

SuccinctEdge: A Succinct RDF Store for Edge Computing 2857
Weiqin Xu, Olivier CURE, Philippe Calvez

SuDocu: Summarizing Documents by Example 2861
Anna Fariha, Matteo Brucato, Peter Haas, Alexandra Meliou

CONCIERGE: Improving Constrained Search Results by Data Melioration 2865
Ido Guy, Tova Milo, Slava Novgorodov, Brit Youngmann

Demonstrating the Voice-Based Exploration of Large Data Sets with CiceroDB-Zero 2869
Immanuel Trummer

FASTS: A Satisfaction-Boosting Bus Scheduling Assistant 2873
Songsong Mo, Zhifeng Bao, Baihua Zheng, Zhiyong Peng

Vaas: Video Analytics At Scale 2877
Favyen Bastani, Oscar Moll, Samuel Madden

sPaQLTooLs: A Stochastic Package Query Interface for Scalable Constrained Optimization 2881
Matteo Brucato, Miro Mannino, Azza Abouzied, Peter Haas, Alexandra Meliou

ActiveDeeper: A Model-based Active Data Enrichment System 2885
Liang Zhao, Qingcan Li, Pei Wang, Jiannan Wang, Eugene Wu

RDFFrames: Knowledge Graph Access for Machine Learning Tools 2889
Aisha Mohamed, Ghadeer Abuoda, Abdurrahman Ghanem, Zoi Kaoudi, Ashraf Aboulnaga

Scalable, Resilient and Configurable Permissioned Blockchain Fabric 2893
Sajjad Rahnama, Suyash Gupta, Thamir Qadah, Jelle Hellings, Mohammad Sadoghi

BIRDS: Programming view update strategies in Datalog 2897
Van-Dang Tran, Hiroyuki Kato, Zhenjiang Hu

Apache IoTDB: Time-series database for Internet of Things 2901
Chen Wang, Huang Xiangdong, Jialin Qiao, Jianmin Wang, Jiaguang Sun, Kevin McGrail, Julian Feinauer, Jinrui Zhang, Peng Wang, Jinrui Zhang, Rong Kang, Tian Jiang, Lei Rui, Jun Yuan

Evaluating Ridesharing Algorithms using the Jargo Real-Time Stochastic Simulator 2905
James Pan, Guoliang Li, Yong Wang

BitFun: Fast Answers to Queries with Tunable Functions in Geospatial Array DBMS 2909
Ramon Antonio Rodriges Zalipynis

SPHINX: A System for Metapath-based Entity Exploration in Heterogeneous Information Networks 2913

Serafeim Chatzopoulos, Kostas Patroumpas, Alexandros Zeakis, Thanasis Vergoulis, Dimitrios Skoutas
ExplainED: Explanations for EDA Notebooks 2917
Daniel Deutch, Amir Gilad, Tova Milo, Amit Somech

MuSe: Multiple Deletion Semantics for Data Repair 2921
Amir Gilad, Yihao Hu, Daniel Deutch, Sudeepa Roy

Tabula in Action: A Sampling Middleware for Interactive Geospatial Visualization Dashboards 2925
Jia Yu, Kanchan Chowdhury, Mohamed Sarwat

X^2R^2: a Tool for Explainable and Explorative Reidentification Risk Analysis 2929
Tom Rolandus Hagedoorn, Rohit Kumar, Francesco Bonchi

Obi-Wan: Ontology-Based RDF Integration of Heterogeneous Data 2933
Maxime Buron, François Goasdoué, Ioana Manolescu, Marie-Laure Mugnier

CrocodileDB in Action: Resource-Efficient Query Execution by Exploiting Time Slackness 2937
Dixin Tang, Zechao Shang, Aaron Elmore, Sanjay Krishnan, Michael Franklin

GraphAn: Graph-based Subsequence Anomaly Detection 2941
Paul Boniol, Themis Palpanas, Mohammed Meftah, Emmanuel Remy

LMFAO: An Engine for Batches of Group-By Aggregates 2945
Maximilian Schleich, Dan Olteanu

ESTOCADA: Towards Scalable Polystore Systems 2949
Rana Alotaibi, Bogdan Cautis, Alin Deutsch, Moustafa Latrache, Ioana Manolescu, Yifei Yang

Demonstration of Interactive Runtime Debugging of Distributed Dataflows in Texera 2953
Zuozhi Wang, Avinash Kumar, Shengquan Ni, Chen Li

DeepTRANS: A Deep Learning System for Public Bus Travel Time Estimation using Traffic Forecasting 2957
Luan Tran, Min Young Mun, Matthew Lim, Jonah Yamato, Nathan Huh, Cyrus Shahabi

Demonstration of ScroogeDB: Getting More Bang For the Buck with Deterministic Approximation in the Cloud 2961
Saehan Jo, Jialing Pei, Immanuel Trummer

Scrutinizer: Fact Checking Statistical Claims 2965
Georgios Karagiannis, Mohammed Saeed, Paolo Papotti, Immanuel Trummer

SciLens News Platform: A System for Real-Time Evaluation of News Articles 2969
Angelika Romanou, Panayiotis Smeros, Carlos Castillo, Karl Aberer

HDAG-Explorer: A System for Hierarchical DAG Summarization and Exploration 2973
Xuliang Zhu, Xin Huang, Jinbin Huang, Byron Choi, Jianliang Xu

Orca-SR: A Real-Time Traffic Engineering Framework leveraging Similarity Joins 2977
Jees Augustine, Suraj Shetiya, Abolfazl Asudeh, Saravanan Thirumuruganathan, Azade Nazi, Nan Zhang, Gautam Das, Divesh Srivastava

nKV in Action: Accelerating KV-Stores on NativeComputational Storage with Near-Data Processing 2981
Tobias Vincon, Lukas Weber, Arthur Bernhard, Andreas Koch, Ilia Petrov, Christian Knoedler, Sergey Hardock, Sajjad Tamimi, Christian Riegger

Demonstration of Inferring Causality from Relational Databases with CaRL 2985
Moe Kayali , Babak Salimi, Dan Suciu

SQL for Data Scientists: Designing SQL Tutorials for Scalable Online Teaching 2989
Uwe Roehm, Lexi Brent, Tim Dawborn, Bryn Jeffries

Debugging Large-Scale Data Science Pipelines using Dagger..... 2993
El Kindi Rezig, Ashrita Brahmaroutu, Nesime Tatbul, Mourad Ouzzani, Nan Tang, Timothy Mattson, Samuel Madden, Michael Stonebraker

I-Rex: An Interactive Relational Query Explainer for SQL 2997
Zhengjie Miao, Tiangang Chen, Alexander Bendeck, Kevin Day, Sudeepa Roy, Jun Yang

PANDA: Policy-aware Location Privacy for Epidemic Surveillance 3001
Yang Cao, Shun Takagi, Yonghui Xiao, Li Xiong, Masatoshi Yoshikawa

Industrial and Applications

PyTorch Distributed: Experiences on Accelerating Data Parallel Training 3005
Shen Li, Yanli Zhao, Rohan Varma, Omkar Salpekar, Pieter Noordhuis, Teng Li, Adam Paszke, Jeff Smith, Brian Vaughan, Pritam Damania, Soumith Chintala

Towards Multi-way Join Aware Optimizer in SAP HANA 3019
Sungheun Wi, Wook-Shin Han, Chuho Chang, Kihong Kim

InvaliDB: Scalable Push-Based Real-Time Queries on Top of Pull-Based Databases 3032
Wolfram Wingerath, Felix Gessert, Norbert Ritter

Automated Generation of Materialized Views in Oracle..... 3046
Rafi Ahmed, Randall Bello, Andrew Witkowski, Praveen Kumar

Native JSON Datatype Support: Maturing SQL and NoSQL convergence in Oracle Database..... 3059
Zhen Hua Liu, Beda Hammerschmidt, Douglas McMahon, Hui Chang, Ying Lu, Josh Spiegel, Alfonso Colunga Sosa, srikrishnan suresh, Geeta Arora, vikas Arora

TiDB: A Raft-based HTAP Database..... 3072
Dongxu Huang, Qi Liu, Qiu Cui, Zhuhe Fang, Xiaoyu Ma, Fei Xu, Li Shen, Liu Tang, Yuxing Zhou, Menglong Huang, Wan Wei, Cong Liu, Jian Zhang, Jianjun Li, Xuelian Wu, Lingyu Song, Ruoxi Sun, Shuaipeng Yu, Lei Zhao, Nicholas Cameron, Liquan Pei, Xin Tang

A system design for elastically scaling transaction processing engines in virtualized servers 3085
Angelos Christos Anadiotis, Raja Appuswamy, Anastasia Ailamaki, Ilan Bronshtein, Hillel Avni, David Dominguez-Sal, Shay Goikhman, Eliezer Levy

Industrial Strength OLTP Using Main Memory and Many Cores 3099
Hillel Avni, Alisher Aliev, Oren Amor, Aharon Avitzur, Ilan Bronshtein, Eli Ginot, Shay Goikhman, Eliezer Levy, Idan Levy, Fuyang Lu, Liran Mishali, Yeqin Mo, Nir Pachter, Dima Sivov, Vinoth Veeraraghavan, Vladi Vexler, Lei Wang, Peng Wang

Asymmetric-Partition Replication for Highly Scalable Distributed Transaction Processing in Practice 3112
Juchang Lee, Hyejeong Lee, Seongyun Ko, Kyu Hwan Kim, Mihnea ANDREI, Friedrich Keller, Wook-Shin Han

AGL: A Scalable System for Industrial-purpose Graph Machine Learning 3125
Dalong Zhang, Xin Huang, ziqi liu, Jun Zhou, Zhiyang Hu, Xianzheng Song, Zhibang Ge, Lin Wang, Zhiqiang Zhang, Yuan Qi

LedgerDB: A Centralized Ledger Database for Universal Audit and Verification 3138
Xinying Yang, Yuan Zhang, Sheng Wang, Benquan Yu, Feifei Li, Yize Li, Wenyuan Yan

AnalyticDB-V: A Hybrid Analytical Engine Towards Query Fusion for Structured and Unstructured Data 3152
Chuangxian Wei, Bin Wu, Sheng Wang, Renjie Lou, Chaoqun Zhan, Feifei Li, Yuanzhe Cai

Oracle AutoML: A Fast and Predictive AutoML Pipeline 3166
Anatoly Yakovlev, Hesam Fathi Moghadam, Ali Moharrer, Jingxiao Cai, Nikan Chavoshi, Venkatanathan Varadarajan, Sandeep Agrawal, Tomas Karnagel, Sam Idicula, Sanjay Jinturkar, Nipun Agarwal)

Monarch: Google's Planet-Scale In-Memory Time Series Database..... 3181
Colin Adams, Luis Alonso, Benjamin Atkin, John Banning, Sumeer Bhola, Rick Buskens, Ming Chen, Xi Chen, Yoo Chung, Qin Jia, Nick Sakharov, George Talbot, Nick Taylor, Adam Tart

Concurrent Updates to Pages with Fixed-Size rows Using Lock-Free Algorithms..... 3195
Raghavendra Thallam Kodandaramaih, Hanuma Kodavalla, Girish Venkataramanappa

POLARIS: The Distributed SQL Engine in Azure Synapse 3204
Josep Aguilar Saborit, Raghu Ramakrishnan

MyRocks: LSM-Tree Database Storage Engine Serving Facebook's Social Graph 3217
Yoshinori Matsunobu, Siying Dong, Herman Lee

Helios: Hyperscale Indexing for the Cloud & Edge 3231
Rahul Potharaju, Terry Kim, Wentao Wu, Vidip Acharya, Steve Suh, Andrew Fogarty, Apoorve Dave, Sinduja Ramanujam, Tomas Talius, Lev Novik, Raghu Ramakrishnan

Replication at the Speed of Change – a Fast, Scalable Replication Solution for Near Real-Time HTAP Processing 3245
Dennis Butterstein, Daniel Martin, Knut Stolze, Felix Beier, Jia Zhong, Lingyun Wang

Exploiting Domain Knowledge to address Multi-Class Imbalance and a Heterogeneous Feature Space in Classification Tasks for Manufacturing Data 3258
Vitali Hirsch, Peter Reimann, Bernhard Mitschang

Alibaba Hologres: A Cloud-Native Service for Hybrid Serving/Analytical Processing 3272
Xiaowei Jiang, Yuejun Hu, Yu Xiang, Guangran Jiang, Xiaojun Jin, Chen Xia, Weihua Jiang, Jun Yu, Haitao Wang, Yuan Jiang, Jihong Ma, Li Su, Kai Zeng

DIAMetrics: Benchmarking Query Engines at Scale 3285
Anja Gruenheid, Shaleen Deep, Kruthi Nagaraj, Hiro Naito, Jeff Naughton, Stratis Viglas

Db2 Event Store: A Purpose-Built IoT Database Engine 3299
Christian Garcia-Arellano, Adam Storm, David Kalmuk, Hamdi Roumani, Ronald Barber, Yuanyuan Tian, Richard Sidle, Fatma Ozcan, Matthew Spilchen, Josh Tiefenbach, Daniel Zilio, Lan Pham, Kostas Rakopoulos, Alexander Cheung, Darren Pepper, Imran Sayyid, Gidon Gershinsky, Gal Lushi, Hamid Pirahesh

F1 Lightning: HTAP as a Service 3313
Jiacheng Yang, Ian Rae, Jun Xu, Jeff Shute, Zhan Yuan, Kelvin Lau, Qiang Zeng, Xi Zhao, Jun Ma, Ziyang Chen, Yuan Gao, Qilin Dong, Junxiong Zhou, Jeremy Wood, Goetz Graefe, Jeff Naughton, John Cieslewicz

AutoToken: Predicting Peak Parallelism for Big Data Analytics at Microsoft 3326
Rathijit Sen, Alekh Jindal, Hiren Patel, Shi Qiao

A Drop-in Middleware for Serializable DB Clustering across Geo-distributed Sites 3340
Enrique Saurez, Bharath Balasubramanian, Richard Schlichting, Brendan Tschaen, Shankaranarayanan Puzhavakath Narayanan, Zhe Huang, Umakishore Ramachandran

Improving Reproducibility of Data Science Pipelines through Transparent Provenance Capture ... 3354
Lukas Rupprecht , James Davis, Constantine Arnold, Yaniv Gur, Deepavali Bhagwat

Conversational BI: An Ontology-Driven ConversationSystem for Business Intelligence Applications 3369
Abdul Quamar, Fatma Ozcan, Dorian Miller, Robert Moore, Rebecca Niehus, Jeffrey Kreulen

Magnet: Push-based Shuffle Service for Large-scale Data Processing 3382
Min Shen, Ye Zhou, Chandni Singh

Leveraging Organizational Resources to Adapt Models to New Data Modalities 3396
Sahaana Suri, Abishek Sethi, Girija Narlikar, Neslihan Bulut, Raghuveer Chanda, Sugato Basu, Pradyumna Narayana, Peter Bailis, Christopher Re, Yemao Zeng

Delta Lake: High-Performance ACID Table Storage over Cloud Object Stores 3411
Michael Armbrust, Tathagata Das, Sameer Paranjpye, Reynold Xin, Shixiong Zhu, Ali Ghodsi, Burak Yavuz, Mukul Murthy, Joseph Torres, Liwen Sun, Peter Boncz, Mostafa Mokhtar, Herman van Hovell, Adrian Ionescu, Alicja Luszczak, Michal Switakowski, Takuya Ueshin, Xiao Li, Michal Szafranski, Pieter Senster, Matei Zaharia

Tutorials

Robust Query Processing: Mission Possible 3425
Jayant Haritsa

Data Collection and Quality Challenges for Deep Learning 3429
Steven Whang, Jae-Gil Lee

Table Extraction and Understanding for Scientific and Enterprise Applications 3433
Doug Burdick, Marina Danilevsky, Alexandre Evfimievski, Yannis Katsis, Nancy X.R. Wang

Similarity Query Processing for High-Dimensional Data 3437
Jianbin Qin, Wei Wang, Chuan Xiao, Ying Zhang

Building High Throughput Permissioned Blockchain Fabrics: Challenges and Opportunities 3441
Suyash Gupta, Jelle Hellings, Sajjad Rahnama, Mohammad Sadoghi

Fairly Evaluating and Scoring Items in a Data Set 3445
Abolfazl Asudeh, H. Jagadish

Award Talks, Keynotes, and Panel

Spitz: A Verifiable Database System 3449
Meihui Zhang, Zhongle Xie, Cong Yue, Ziyue Zhong

Dremel: A Decade of Interactive SQL Analysis at Web Scale 3461
Sergey Melnik, Andrey Gubarev, Jing Jing Long, Geoffrey Romer, Shiva Shivakumar, Matt Tolton, Theo Vassilakis, Hossein Ahmadi, Dan Delorey, Slava Min, Mosha Pasumansky, Jeff Shute

JIT works: decide when all data is known 3473
Anastasia Ailamaki

Responsible Data Management 3474
Julia Stoyanovich, Bill Howe, H.V. Jagadish

Out-of-order Execution of Database Queries 3489
Kazuo Goda, Yuto Hayamizu, Hiroyuki Yamada, Masaru Kitsuregawa

The Relational Data Borg is Learning 3502
Dan Olteanu

Winds from Seattle: Database Research Directions 3516
Peter Bailis, Magda Balazinska, Xin Luna Dong, Juliana Freire, Raghu Ramakrishnan, Michael Stonebraker, Joseph M. Hellerstein

PVLDB ORGANIZATION AND REVIEW BOARD - Vol. 13

Editors in Chief of PVLDB
Magdalena Balazinska (University of Washington, USA)
Xiaofang Zhou (University of Queensland, Australia)

Associate Editors of PVLDB
Azza Abouzied (New York University Abu Dhabi, UAE)
Amr El Abbadi (University of California, Santa Barbara, USA)
Phil Bernstein (Microsoft Research, USA)
Xin Luna Dong (Amazon, USA)
Zi (Helen) Huang (University of Queensland, Australia)
Nick Koudas (University of Toronto, Canada)
Georgia Koutrika (Athena Research Center, Greece)
Guoliang Li (Tsinghua University, China)
Alexandra Meliou (University of Massachusetts, Amherst, USA)
Felix Naumann (Hasso Plattner Institute, University of Potsdam, Germany)
Dan Olteanu (University of Oxford, United Kingdom)
M. Tamer Özsu (University of Waterloo, Canada)
Aditya Parameswaran (University of California, Berkeley, USA)
Andy Pavlo (Carnegie Mellon University, USA)
Xiaokui Xiao (National University of Singapore, Singapore)
Jeffrey Xu Yu (The Chinese University of Hong Kong, China)
Meihui Zhang (Beijing Institute of Technology, China)
Jingren Zhou (Alibaba Group, China)

Publication Editors
Hiroaki Shiokawa (University of Tsukuba, Japan)
Sen Wang (University of Queensland, Australia)

PVLDB Managing Editor
Wolfgang Lehner (TU Dresden, Germany)

PVLDB Advisory Committee
Divesh Srivastava (AT&T Labs-Research, USA)
M. Tamer Özsu (University of Waterloo, Canada)
Juliana Freire (New York University, USA)
Xin Luna Dong (Amazon, USA)
Peter Boncz (CWI, Netherlands)
Xiaofang Zhou (University of Queensland, Australia)
Magdalena Balazinska (University of Washington, USA)
Lei Chen (Hong Kong University of Science and Technology, China)
Fatma Ozcan (IBM Almaden, USA)
Graham Cormode (University of Warwick, United Kingdom)
Felix Naumann (HPI, Germany)

Review Board

Ziawasch Abedjan (TU Berlin, Germany)
Ashraf Aboulnaga (Qatar Computing Research Institute, Qatar)
Pelin Angın (Middle East Technical University, Turkey)
Arvind Arasu (Microsoft Research, USA)
Joy Arulraj (Georgia Tech, USA)
Manos Athanassoulis (Boston University, USA)
Zhifeng Bao (RMIT University, Australia)
Ilaria Bartolini (University of Bologna, Italy)
Leilani Battle (University of Maryland, USA)
Kaustubh Beedkar (TU Berlin, Germany)
Arnab Bhattacharya (IIT Kanpur, India)
Sourav S Bhowmick (Nanyang Technological University, Singapore)
Carsten Binnig (TU Darmstadt, Germany)
Spyros Blanas (The Ohio State University, USA)
Matthias Boehm (Graz University of Technology, Austria)
Alexander Böhm (SAP SE, Germany)
Michael Böhlen (University of Zürich, Switzerland)
Peter Boncz (Centrum Wiskunde & Informatica, Netherlands)
Angela Bonifati (Lyon 1 University, France)
Philippe Bonnet (IT University of Copenhagen, Denmark)
Renata Borovica-Gajic (University of Melbourne, Australia)
Huiping Cao (New Mexico State University, USA)
Lei Cao (MIT, USA)
Lijun Chang (The University of Sydney, Australia)
Surajit Chaudhuri (Microsoft Research, USA)
Lei Chen (Hong Kong University of Science and Technology, China)
Hong Cheng (The Chinese University of Hong Kong, China)
Reynold Cheng (The University of Hong Kong, China)
Fei Chiang (McMaster University, Canada)
Xu Chu (Georgia Tech, USA)
Bobbie Cochrane (IBM, USA)
Gao Cong (Nanyang Technological University, Singapore)
Brian Cooper (Google, USA)
Natacha Crooks (University of Texas at Austin, USA)
Andrew Crotty (Brown University, USA)
Bin Cui (Peking University, China)
Sudipto Das (Amazon Web Service, USA)
Akash Das Sarma (Facebook, USA)
Khuzaima Daudjee (University of Waterloo, Canada)
Niv Dayan (Harvard University, USA)
Dong Deng (Rutgers University, USA)
Bailu Ding (Microsoft Research, USA)
Bolin Ding (Alibaba Group, China)
Jens Dittrich (Saarland University, Germany)
Harish Doraiswamy (New York University, USA)
Eduard C. Dragut (Temple University, USA)
Curtis Dyreson (Utah State University, USA)
Mohamed Y. Eltabakh (Teradata Labs, USA)
Jose M. Faleiro (Microsoft Research, USA)
Ju Fan (Renmin University of China, China)
Raul Castro Fernandez (The University of Chicago, USA)
Avrilia Floratou (Microsoft Research, USA)
Avigdor Gal (Technion, Israel)
Alex Galakatos (Brown University, USA)
Johann Gamper (Free University of Bozen-Bolzano, Italy)
Jing Gao (University at Buffalo, USA)
Yunjun Gao (Zhejiang University, China)
Tingjian Ge (University of Massachusetts, Lowell, USA)
Floris Geerts (University of Antwerp, Belgium)
Johannes Gehrke (Microsoft Research, USA)
Jonathan Goldstein (Microsoft Research, USA)
Torsten Grust (University of Tübingen, Germany)
Wook-Shin Han (POSTECH, South Korea)
Takahiro Hara (Osaka University, Japan)
Oktie Hassanzadeh (IBM Research, USA)
Michael Hay (Colgate University, USA)
Xi He (University of Waterloo, Canada)
Melanie Herschel (University of Stuttgart, Germany)
Katja Hose (Aalborg University, Denmark)
Wen Hua (The University of Queensland, Australia)
Xin Huang (Hong Kong Baptist University, China)
Yan Huang (University of North Texas, USA)
Seung-won Hwang (Yonsei University, South Korea)
Christopher Jermaine (Rice University, USA)
Ruoming Jin (Kent State University, USA)
Eser Kandogan (Megagon Labs, USA)
Murat Kantarcioglu (University of Texas at Dallas, USA)
Verena Kantere (University of Ottawa, Canada)
Pinar Karagoz (Middle East Technical University, Turkey)
Manos Karpathiotakis (Facebook, United Kingdom)
Batya Kenig (University of Washington, USA)
Oliver Kennedy (University at Buffalo, USA)
Arijit Khan (Nanyang Technological University, Singapore)
Daniel Kifer (Pennsylvania State University, USA)
Hideaki Kimura (Oracle, USA)
Sanjay Krishnan (University of Chicago, USA)
Arun Kumar (University of California, San Diego, USA)
Chuan Lei (IBM Research - Almaden, USA)
Viktor Leis (Technical University of Munich, Germany)
Ulf Leser (Humboldt-Universität zu Berlin, Germany)
Chengkai Li (The University of Texas at Arlington, USA)
Feifei Li (University of Utah, USA)
Rong-Hua Li (Beijing Institute of Technology, China)
Sebastian Link (The University of Auckland, New Zealand)
Chengfei Liu (Swinburne University of Technology, Australia)
Hua Lu (Aalborg University, Denmark)
Jiaheng Lu (University of Helsinki, Finland)
Wei Lu (Renmin University of China, China)
Shuai Ma (Beihang University, China)
Nikos Mamoulis (University of Ioannina, Greece)

Ioana Manolescu (INRIA, France)
Essam Mansour (Concordia University, Canada)
Ryan Marcus (MIT, USA)
Sergey Melnik (Google, USA)
Mohamed Mokbel (Qatar Computing Research Institute, Qatar)
Mirella Moura Moro (Universidade Federal de Minas Gerais, Brazil)
Davide Mottin (Aarhus University, Denmark)
Parth Nagarkar (New Mexico State University, USA)
Faisal Nawab (University of California, Santa Cruz, USA)
Thomas Neumann (Technical University of Munich, Germany)
Milos Nikolic (The University of Edinburgh, United Kingdom)
Beng Chin Ooi (National University of Singapore, Singapore)
Ismail Oukid (SAP SE, USA)
Mourad Ouzzani (Qatar Computing Research Institute, Qatar)
Themis Palpanas (Paris Descartes University, France)
George Papadakis (University of Athens, Greece)
Olga Papaemmanouil (Brandeis University, USA)
Thorsten Papenbrock (Hasso Plattner Institute, Germany)
Paolo Papotti (EURECOM, France)
Stefano Paraboschi (Universita' degli Studi di Bergamo, Italy)
Yongjoo Park (University of Michigan, USA)
Jignesh M. Patel (University of Wisconsin-Madison, USA)
Peter Pietzuch (Imperial College London, United Kingdom)
Holger Pirk (Imperial College London, United Kingdom)
Fábio Porto (National Laboratory for Scientific Computing (LNCC), Brazil)
Dan R. K. Ports (Microsoft Research, USA)
Lu Qin (University of Technology Sydney, Australia)
Abdul H. Quamar (IBM Research – Almaden, USA)
Tilmann Rabl (TU Berlin, Germany)
Karthik Ramachandra (Microsoft Research, USA)
Maya Ramanath (IIT Delhi, India)
Berthold Reinwald (IBM Research, USA)
Theodoros Rekatsinas (University of Wisconsin-Madison, USA)
Uwe Roehm (The University of Sydney, Australia)
Jennie Rogers (Northwestern University, USA)
Florin Rusu (University of California, Merced, USA)
Diptikalyan Saha (IBM Research AI India, India)
Ken Salem (University of Waterloo, Canada)
Semih Salihoglu (University of Waterloo, Canada)
Maria Luisa Sapino (University of Torino, Italy)
A. Erdem Sariyuce (University at Buffalo, USA)
Mohamed Sarwat (Arizona State University, USA)
Maximilian Schleich (University of Oxford, United Kingdom)
Mohamed Sharaf (University of Queensland, Australia)
Yanyan Shen (Shanghai Jiao Tong University, China)
Kyuseok Shim (Seoul National University, South Korea)
Prashant Shiralkar (Amazon, USA)
Alkis Simitsis (Hewlett Packard Enterprise, USA)
Kostas Stefanidis (Tampere University, Finland)
Rebecca Taft (Cockroach Labs, USA)
Nan Tang (Qatar Computing Research Institute, Qatar)
Yufei Tao (The Chinese University of Hong Kong, China)
Jens Teubner (TU Dortmund, Germany)
Andreas Thor (University of Applied Sciences for Telecommunications Leipzig, Germany)
Yongxin Tong (Beihang University, China)
Anthony K. H. Tung (National University of Singapore, Singapore)
Yannis Velegrakis (Utrecht University, Netherlands)
Stratis Viglas (University of Edinburgh, United Kingdom)
Daisy Zhe Wang (University of Florida, USA)
Guoren Wang (Beijing Institute of Technology, China)
Jiannan Wang (Simon Fraser University, USA)
Junhu Wang (Griffith University, Australia)
Sibo Wang (The Chinese University of Hong Kong, China)
Eugene Wu (Columbia University, USA)
Yingjun Wu (Amazon Web Service, USA)
Yinglong Xia (Huawei Research America, USA)
Chuan Xiao (Osaka University, Japan)
Yanghua Xiao (Fudan University, China)
Li Xiong (Emory University, USA)
Jianliang Xu (Hong Kong Baptist University, China)
Xiaochun Yang (Northeastern University, China)
Junjie Yao (East China Normal University, China)
Hongzhi Yin (The University of Queensland, Australia)
Man Lung Yiu (Hong Kong Polytechnic University, China)
Haruo Yokota (Tokyo Institute of Technology, Japan)
Masatoshi Yoshikawa (Kyoto University, Japan)
Xiangyao Yu (University of Wisconsin-Madison, USA)
Demetrios Zeinalipour-Yazti (University of Cyprus, Cyprus)
Baihua Zheng (Singapore Management University, Singapore)
Rui Zhang (University of Melbourne, Australia)
Wenjie Zhang (University of New South Wales, Australia)
Xiaofei Zhang (The University of Memphis, USA)
Ying Zhang (University of Technology Sydney, Australia)
Yuanyuan Zhu (Wuhan University, China)
Lei Zou (Peking University, China)
Kostas Zoumpatianos (Harvard University, USA)

DEMOSTRATION TRACK CHAIRS AND REVIEWERS - Vol. 13

Demonstration Track PC Chairs

Yasuhiro Fujiwara (NTT Communication Science Laboratories, Japan)
Ioana Manolescu (Inria and Ecole Polytechnique, France)
Ying Zhang (The University of Technology Sydney, Australia)

Demonstration Track Reviewers

Ahmed Eldawy (University of California Riverside, USA)
Alekh Jindal (Microsoft, USA)
Alvin Cheung (University of California Berkeley, USA)
Amélie Marian (Rutgers, USA)
Angelos-Christos Anadiotis (École Polytechnique, France)
Annika Hinze (The University of Waikato, USA)
Benoît Groz (Université Paris-Saclay, France)
Boris Glavic (Illinois Institute of Technology, USA)
Chaokun Wang (Tsinghua University, China)
Daniel Deutch (Tel Aviv University, Israel)
David Eyers (University of Otago, USA)
Dong Wen (University of Technology Sydney, Australia)
Fabian Suchanek (Télécom Paris University, France)
Fan Zhang (Guangzhou University, China)
Guanfeng Liu (Macquarie University, Australia)
Helena Galhardas (University of Lisbon, Portugal)
Ioana Giurgiu (IBM Research, USA)
Jens Teubner (TU Dortmund University, Germany)
Jiang Du (Facebook, USA)
Katerina Tzompanaki (CY Cergy Paris University, France)
Katja Hose (Aalborg University, Denmark)
Kazuo Goda (The University of Tokyo, Japan)
Keishi Tajima (Kyoto University, Japan)
Letizia Tanca (Politecnico di Milano, Italy)
Lu Chen (Aalborg University, Denmark)
Luc Bouganim (INRIA, France)
Lucian Popa (IBM Almaden Research Center, USA)
Michael Grossniklaus (University of Konstanz, Germany)
Michael Gubanov (Florida State University, USA)
Nesreen Ahmed (Intel Labs, USA)
Oana Balalau (INRIA and École Polytechnique, France)
Oscar Romero (Universitat Politècnica de Catalunya, Spain)
Pinar Karagoz (METU, Turkey)
Qiong Luo (Hong Kong University of Science and Technology, China)
Saravanan Thirumuruganathan (QCRI, Hamad Bin Khalifa University, Qatar)
Slava Novgorodov (eBay Research, USA)
Vijay Gadepally (MIT Lincoln Laboratory, USA)
Wentao Wu (Microsoft Research, USA)
Xiaoyang Wang (Zhejiang Gongshang University, China)
Xin Cao (University of New South Wales, Australia)
Yannis Katsis (IBM Research, USA)
Yannis Velegrakis (Utrecht University, Netherlands)
Yash Govind (University of Wisconsin Madison, USA)
Yasuhiro Fujiwara (NTT Communication Science Laboratories, Japan)
Yuanyuan Tian (IBM Almaden, USA)
Yuchen Li (Singapore Management University, Singapore)
Zoi Kaoudi (TU Berlin, Germany)

INDUSTRIAL TRACK CHAIRS AND REVIEWERS - Vol. 13

Industrial Track PC Chairs
Tiziana Catarci (University of Rome La Sapienza, Italy)
Carlo Curino (Microsoft Research, USA)
Divesh Srivastava (AT&T Labs, USA)

Industrial Track Reviewers
Allison Holloway (Oracle, USA)
Altigran da Silva (Universidade Federal do Amazonas, Brazil)
Anja Gruenheid (Google Inc., USA)
Azza Abouzied (New York University Abu Dhabi, United Arab Emirates)
Carlo Curino (Microsoft, USA)
Danica Porobic (Oracle, USA)
Divesh Srivastava (AT&T Labs Research, USA)
Entong Shen (Amazon Web Services, USA)
Fabrizio Silvestri (Facebook, United Kingdom)
Fatma Ozcan (IBM Research - Almaden, USA)
Feifei Li (Alibaba Group, China)
Guoliang Li (Tsinghua University, China)
Heiko Schuldt (University of Basel, Switzerland)
Irini Fundulaki (ICS FORTH, Greece)
Jana Giceva (TU Munich, Germany)
Jianguo Wang (Amazon Web Services, USA)
Johannes Gehrke (Microsoft, USA)
Meihui Zhang (Beijing Institute of Technology, China)
Mónica Scannapieco (ISTAT, Italy)
Norman May (SAP SE, Germany)
Olga Poppe (Microsoft, USA)
Roberto Zicari (Goethe University Frankfurt, Germany)
Russell Sears (Apple, USA)
Sameer Agarwal (Facebook, USA)
Sang Cha (Seoul National University, South Korea)
Sihem Amer-Yahia (CNRS, France)
Songtao Guo (LinkedIn, USA)
Sonia Bergamaschi (Università di Modena e Reggio Emilia, Italy)
Tiziana Catarci (University of Rome "La Sapienza", Italy)
Wolfram Wingerath (Baqend, Germany)
Yaron Kanza (AT&T Labs-Research, USA)
Yasushi Sakurai (Osaka University, Japan)

TUTORIAL TRACK CHAIRS - Vol. 13

Tutorial Track Chairs:
Michael J. Carey (University of California, Irvine, USA)
Renée J. Miller (Northeastern University, USA)

LETTER FROM THE EDITORS IN CHIEF

The Proceedings of the VLDB Endowment (PVLDB) provides a high-quality journal publication service to the data management research community. Each volume offers twelve monthly submission deadlines on the first day of each month and a quick, six week reviewing cycle. This publication model was pioneered by PVLDB and combines a journal-style reviewing process, which includes a three-month revision cycle, with the agility and visibility provided by rapid on-line publication, and presentation at the annual VLDB conference.

PVLDB attracts many submissions spanning diverse data management topics, and the PVLDB reviewing process is implemented by a large team of dedicated researchers. The Review Board of PVLDB Volume 13 consists of 186 expert researchers, and reviewing is coordinated by 18 Associate Editors. Review Board members provide timely (within a 3-week deadline) high-quality reviews, and participate actively in online discussions led by the Associate Editors for each paper. When needed, the Associate Editors together with the Editors-in-Chief solicit additional reviews from external experts. We give special thanks here to those additional reviewers who in most cases need to complete their expert reviews on a very short notice.

Most of the accepted papers go through a revision process that requires a second round of reviews after the authors have addressed an initial set of issues and concerns raised by the reviewers during the first round. Some papers are further accepted with shepherding, which means that one of the reviewers works with the authors to address a final set of comments.

This is the twelfth issue of the thirteenth volume of PVLDB. This issue features 95 papers, all associated with the VLDB'20 conference: 51 demonstrations, 31 industrial track papers, six tutorials, three invited papers for the three VLDB Endowment awards, three invited papers for the three keynote talks, and one panel abstract.

The 51 accepted demonstrations cover topics such as query optimization at all levels, advanced applications such as explanation and causality on relational databases, and a variety of works studying the interplay between machine learning and scalable data management. Knowledge base management, news verification and fact-checking, as well as three tools for monitoring an epidemic such as the Covid one are also part of this year's selection.

The industrial track at VLDB 2020 accepted 31 papers from a total of 69 submissions; 4 of these papers were accepted directly and 27 papers were accepted after a revision phase. These works, reported by large and small companies from around the world, showcase the diversity of industrial research on analytical query processing, benchmarking, blockchain, cloud systems, data storage, graph systems, HTAP, indexing, machine learning systems, natural language processing, performance tuning, provenance, query optimization, semi-structured data, time series, transaction processing and views.

This issue also includes descriptions of the six tutorials presented at the VLDB'20 conference. Those tutorials covered six timely and exciting topics: permissioned blockchain technology, robust query processing techniques, data collection and quality for deep learning, similarity query processing (for high-dimensional data), table extraction/understanding techniques, and fairness in evaluating and scoring data.

This issue further includes three invited papers for the three VLDB Endowment awards: The 2020 VLDB Early Career Research Contribution Award. The award this year went to Prof. Meihui Zhang for contributions to cleaning, integration and analysis techniques for healthcare data, and deployment of those techniques in hospitals. The second award is the 2020 VLDB Women in Database Research Award, which went to Prof. Anastasia Ailamaki for pioneering research on the interaction between hardware micro-architecture and database engine performance. Finally, the 2020 VLDB Test of Time Award went to the PVLDB Vol 3 paper entitled "Dremel: interactive analysis of web-scale datasets".

Three keynote speakers were invited to deliver talks at the VLDB'20 conference and invited papers associated with those talks are included in this issue. Prof. Dan Olteanu spoke on "The Relational Data Borg is Learning". Prof. Julia Stoyanovich talked about "Responsible Data Management" and Prof. Masaru Kitsuregawa presented "Out-of-order Execution of Query Processing and New Advances in COVID-19 Information".

Finally, the VLDB'20 conference also included a panel entitled "Panel - Winds from Seattle: Database Research Directions", and an abstract for the panel is included in this issue.

These papers will be presented at the 46th International Conference on Very Large Data Bases (PVLDB 2020), originally planned to be held in Tokyo, Japan during August 31 to September 4, 2020 but now to be held online due to the coronavirus pandemic.

We hope that the readers will find the selected papers engaging, and thought provoking. We also hope that the selected papers will provide valuable insights and inspire novel systems contributions and follow-up research.

Magdalena Balazinska and Xiaofang Zhou

PVLDB Volume 13 Editors in Chief