

# **48th International Conference on Very Large Data Bases (VLDB 2022)**

Proceedings of the VLDB Endowment Volume 15

Sydney, Australia  
5-9 September 2022

Part 1 of 6

**Editors:**

**Lijun Chang  
Xin Cao**

ISBN: 978-1-7138-6510-0

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

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



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

Copyright© (2022) by 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-ncnd/4.0/>.

For any use beyond those covered by this license, obtain permission by emailing [info@vldb.org](mailto:info@vldb.org).

Printed with permission by Curran Associates, Inc. (2023)

For permission requests, please contact VLDB Endowment at the address below.

VLDB Endowment  
Attn: Marianne Siroker  
353 Serra Mall – Stanford University  
Stanford CA 94305

Phone: (650) 723-0872

Fax: (650) 725-2588

[info@vldb.org](mailto:info@vldb.org).

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

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

## TABLE OF CONTENTS

### Research Papers

ANN Softmax: Acceleration of Extreme Classification Training.....	1
<i>Kang Zhao, Liuyihan Song, Yingya Zhang, Pan Pan, Xu Yinghui, Rong Jin</i>	
WindTunnel: Towards Differentiable ML Pipelines Beyond a Single Model .....	11
<i>Gyeong-in Yu, Saeed Amizadeh, Sehoon Kim, Artidoro Pagnoni, Ce Zhang, Byung-gon Chun, Markus Weimer, Matteo Interlandi</i>	
DBOS: A DBMS-oriented Operating System.....	21
<i>Athinagoras Skiadopoulos, Qian Li, Peter Kraft, Kostis Kaffes, Daniel Hong, Shana Mathew, David Bestor, Michael Cafarella, Vijay Gadepally, Goetz Graefe, Jeremy Kepner, Christos Kozyrakis, Tim Kraska, Michael Stonebraker, Lalith Suresh, Matei Zaharia</i>	
Deep Indexed Active Learning for Matching Heterogeneous Entity Representations.....	31
<i>Arjit Jain, Sunita Sarawagi, Prithviraj Sen</i>	
A Learned Query Rewrite System using Monte Carlo Tree Search.....	46
<i>Xuanhe Zhou, Guoliang Li, Chengliang Chai, Jianhua Feng</i>	
On Detecting Cherry-picked Generalizations.....	59
<i>Yin Lin, Brit Youngman, Yuval Moskovitch, H. V. Jagadish, Tova Milo</i>	
FACE: A Normalizing Flow based Cardinality Estimator .....	72
<i>Jiayi Wang, Chengliang Chai, Jiabin Liu, Guoliang Li</i>	
Learned Cardinality Estimation: A Design Space Exploration and a Comparative Evaluation .....	85
<i>Ji Sun, Jintao Zhang, Zhaoyan Sun, Guoliang Li, Nan Tang</i>	
DeepEverest: Accelerating Declarative Top-K Queries for Deep Neural Network Interpretation .....	98
<i>Dong He, Maureen Daum, Walter Cai, Magdalena Balazinska</i>	
Cosine: A Cloud-Cost Optimized Self-Designing Key-Value Storage Engine .....	112
<i>Subarna Chatterjee, Meena Jagadeesan, Wilson Qin, Stratos Idreos</i>	
Accelerating Recommendation System Training by Leveraging Popular Choices.....	127
<i>Muhammad Adnan, Yassaman Ebrahimzadeh Maboud, Divya Mahajan, Prashant Nair</i>	

## TABLE OF CONTENTS

### Research Papers

(p,q)-biclique Counting and Enumeration for Large Sparse Bipartite Graphs.....	141
<i>Jianye Yang, Yun Peng, Wenjie Zhang</i>	
Evaluating Query Languages and Systems for High-Energy Physics Data .....	154
<i>Dan Gaur, Ingo Müller, Mason Proffitt, Ghislain Fourny, Gordon T. Watts, Gustavo Alonso</i>	
Distributed Hop-Constrained s-t Simple Path Enumeration at Billion Scale.....	169
<i>Kongzhang Hao, Long Yuan, Wenjie Zhang</i>	
ETO: Accelerating Optimization of DNN Operators by High-Performance Tensor Program Reuse.....	183
<i>Jingzhi Fang, Yanyan Shen, Yue Wang, Lei Chen</i>	
Babelfish: Efficient Execution of Polyglot Queries.....	196
<i>Philipp Marian Grulich, Steffen Zeuch, Volker Markl</i>	
Butterfly Counting on Uncertain Bipartite Networks .....	211
<i>Alexander Zhou, Yue Wang, Lei Chen</i>	
METRO: A Generic Graph Neural Network Framework for Multivariate Time Series Forecasting.....	224
<i>Yue Cui, Kai Zheng, Dingshan Cui, Jiandong Xie, Liwei Deng, Feiteng Huang, Xiaofang Zhou</i>	
LargeEA: Aligning Entities for Large-scale Knowledge Graphs .....	237
<i>Congcong Ge, Xiaoze Liu, Lu Chen, Baihua Zheng, Yunjun Gao</i>	
HVS: Hierarchical Graph Structure Based on Voronoi Diagrams for Solving Approximate Nearest Neighbor Search.....	246
<i>Kejing Lu, Mineichi Kudo, Chuan Xiao, Yoshiharu Ishikawa</i>	
Origami: A High-Performance Mergesort Framework.....	259
<i>Arif Arman, Dmitri Loguinov</i>	
Learning to be a Statistician: Learned Estimator for Number of Distinct Values .....	272
<i>Renzhi Wu, Bolin Ding, Xu Chu, Zhewei Wei, Xiening Dai, Tao Guan, Jingren Zhou</i>	
ParChain: A Framework for Parallel Hierarchical Agglomerative Clustering using Nearest-Neighbor Chain .....	285
<i>Shangdi Yu, Yiqiu Wang, Yan Gu, Laxman Dhulipala, Julian Shun</i>	
Answering Regular Path Queries through Exemplars.....	299
<i>Komal Chauhan, Kartik Jain, Sayan Ranu, Srikanta Bedathur, Amitabha Bagchi</i>	
HET: Scaling out Huge Embedding Model Training via Cache-enabled Distributed Framework .....	312
<i>Xupeng Miao, Hailin Zhang, Yining Shi, Xiaonan Nie, Zhi Yang, Yangyu Tao, Bin Cui</i>	

FINEdex: A Fine-grained Learned Index Scheme for Scalable and Concurrent Memory Systems.....	321
<i>Pengfei Li, Yu Hua, Jingnan Jia, Pengfei Zuo</i>	
TaGSim: Type-aware Graph Similarity Learning and Computation .....	335
<i>Jiyang Bai, Peixiang Zhao</i>	
Analysis of Influence Contribution in Social Advertising .....	348
<i>Yuqing Zhu, Jing Tang, Xueyan Tang, Lei Chen</i>	
Scabbard: Single-Node Fault-Tolerant Stream Processing .....	361
<i>Georgios R Theodorakis, Fotios Kounelis, Peter Pietzuch, Holger Pirk</i>	
Enabling Personal Consent in Databases.....	375
<i>George Konstantinidis, Jet Holt, Adriane Chapman</i>	

## TABLE OF CONTENTS

### Research Papers

Enabling SQL-based Training Data Debugging for Federated Learning.....	388
<i>Yejia Liu, Weiyuan Wu, Lampros Flokas, Jiannan Wang, Eugene Wu</i>	
Leveraging Query Logs and Machine Learning for Parametric Query Optimization.....	401
<i>Kapil Vaidya, Anshuman Dutt, Vivek Narasayya, Surajit Chaudhuri</i>	
Pre-training Summarization Models of Structured Datasets for Cardinality Estimation.....	414
<i>Yao Lu, Srikanth Kandula, Arnd Christian König, Surajit Chaudhuri</i>	
xFraud: Explainable Fraud Transaction Detection.....	427
<i>Susie Xi Rao, Shuai Zhang, Zhichao Han, Zitao Zhang, Wei Min, Zhiyao Chen, Yinan Shan, Yang Zhao, Ce Zhang</i>	
Subgraph Matching over Graph Federation .....	437
<i>Ye Yuan, DeLong Ma, Zhenyu Wen, Zhiwei Zhang, Guoren Wang</i>	
Provenance-based Data Skipping .....	451
<i>Xing Niu, Boris Glavic, Ziyu Liu, Pengyuan Li, Dieter Gawlick, Vasudha Krishnaswamy, Zhen Hua Liu, Danica Porobic</i>	
Deep Transfer Learning for Multi-source Entity Linkage via Domain Adaptation .....	465
<i>Di Jin, Bunyamin Sisman, Hao Wei, Xin Luna Dong, Danaï Koutra</i>	
An Experimental Evaluation and Investigation of Waves of Misery in R-trees.....	478
<i>Lu Xing, Eric Lee, Tong An, Bo-cheng Chu, Ahmed Mahmood, Ahmed M. Aly, Jianguo Wang, Walid G. Aref</i>	
PRUC : P-Regions with User-Defined Constraint .....	491
<i>Yongyi Liu, Ahmed R. Mahmood, Amr Magdy, Sergio Rey</i>	
Points-of-Interest Relationship Inference with Spatial-enriched Graph Neural Networks .....	504
<i>Yile Chen, Xiucheng Li, Gao Cong, Cheng Long, Zhifeng Bao, Shang Liu, Wanli Gu, Fuzheng Zhang</i>	
SAFE: A Share-and-Aggregate Bandwidth Exploration Framework for Kernel Density Visualization ...	513
<i>Tsz Nam Chan, Pak Lon Ip, Leong Hou U, Byron Choi, Jianliang Xu</i>	
The next 50 Years in Database Indexing or: The Case for Automatically Generated Index Structures .....	527
<i>Jens Dittrich, Joris Nix, Christian Schön</i>	
DARLING: Data-Aware Load Shedding in Complex Event Processing Systems .....	541
<i>Koral Chapnik, Ilya Kolchinsky, Assaf Schuster</i>	

Rearchitecting In-Memory Object Stores for Low Latency.....	555
<i>Danyang Zhuo, Kaiyuan Zhang, Zhuohan Li, Siyuan Zhuang, Stephanie Wang, Ang Chen, Ion Stoica</i>	
MT-Teql: Evaluating and Augmenting Neural NLIDB on Real-world Linguistic and Schema Variations .....	569
<i>Pingchuan Ma, Shuai Wang</i>	
Theoretically and Practically Efficient Parallel Nucleus Decomposition .....	583
<i>Jessica Shi, Laxman Dhulipala, Julian Shun</i>	
APEX: A High-Performance Learned Index on Persistent Memory.....	597
<i>Baotong Lu, Jialin Ding, Eric Lo, Umar Farooq Minhas, Tianzheng Wang</i>	
Unsupervised Time Series Outlier Detection with Diversity-Driven Convolutional Ensembles .....	611
<i>David Campos, Tung Kieu, Chenjuan Guo, Feiteng Huang, Kai Zheng, Bin Yang, Christian S. Jensen</i>	
Efficient and Effective Data Imputation with Influence Functions .....	624
<i>Xiaoye Miao, Yangyang Wu, Lu Chen, Yunjun Gao, Jun Wang, Jianwei Yin</i>	
Parallel Training of Knowledge Graph Embedding Models: A Comparison of Techniques.....	633
<i>Adrian Kochsiek, Rainer Gemulla</i>	
Detecting Layout Templates in Complex Multiregion Files.....	646
<i>Gerardo Vitagliano, Lan Jiang, Felix Naumann</i>	
What Is the Price for Joining Securely? Benchmarking Equi-Joins in Trusted Execution Environments .....	659
<i>Kajetan Maliszewski, Jorge-Arnulfo Quiane-Ruiz, Jonas Traub, Volker Markl</i>	
Efficient Temporal Pattern Mining in Big Time Series Using Mutual Information.....	673
<i>Van Long Ho, Nguyen Ho, Torben Bach Pedersen</i>	
Efficient Label-Constrained Shortest Path Queries on Road Networks: A Tree Decomposition Approach .....	686
<i>Junhua Zhang, Long Yuan, Wentao Li, Lu Qin, Ying Zhang</i>	
Ember: No-Code Context Enrichment via Similarity-Based Keyless Joins .....	699
<i>Sahaana Suri, Ihab F Ilyas, Christopher Re, Theodoros Rekatsinas</i>	
Incremental Partitioning for Efficient Spatial Data Analytics .....	713
<i>Tin Vu, Ahmed Eldawy, Vagelis Hristidis, Vassilis J. Tsotras</i>	
Lux: Always-on Visualization Recommendations for Exploratory Dataframe Workflows.....	727
<i>Doris Lee, Dixin Tang, Kunal Agarwal, Thyne Boonmark, Caitlyn Chen, Jake Kang, Ujjaini Mukhopadhyay, Jerry Song, Micah Yong, Marti A. Hearst, Aditya Parameswaran</i>	
Flexible Rule-Based Decomposition and Metadata Independence in Modin: A Parallel Dataframe System .....	739
<i>Devin Petersohn, Dixin Tang, Rehan S Durrani, Areg Melik-adamyam, Joseph Gonzalez, Anthony Joseph, Aditya Parameswaran</i>	

## TABLE OF CONTENTS

### Research Papers

Cardinality Estimation in DBMS: A Comprehensive Benchmark Evaluation .....	752
<i>Yuxing Han, Ziniu Wu, Peizhi Wu, Rong Zhu, Jingyi Yang, Liang Wei Tan, Kai Zeng, Gao Cong, Yanzhao Qin, Andreas Pfadler, Zhengping Qian, Jingren Zhou, Jiangneng Li, Bin Cui</i>	
Redy: Remote Dynamic Memory Cache .....	766
<i>Qizhen Zhang, Philip A Bernstein, Daniel S Berger, Badrish Chandramouli</i>	
Robust and Budget-Constrained Encoding Configurations for In-Memory Database Systems .....	780
<i>Martin Boissier</i>	
Fast Neural Ranking on Bipartite Graph Indices.....	794
<i>Shulong Tan, Weijie Zhao, Ping Li</i>	
BAGUA: Scaling up Distributed Learning with System Relaxations .....	804
<i>Shaoduo Gan, Xiangru Lian, Rui Wang, Jianbin Chang, Chengjun Liu, Hongmei Shi, Shengzhuo Zhang, Xianghong Li, Tengxu Sun, Jiawei Jiang, Binhang Yuan, Sen Yang, Ji Liu, Ce Zhang</i>	
SWS: A Complexity-Optimized Solution for Spatial-Temporal Kernel Density Visualization .....	814
<i>Tsz Nam Chan, Pak Lon Ip, Leong Hou U, Byron Choi, Jianliang Xu</i>	
Projected Federated Averaging with Heterogeneous Differential Privacy .....	828
<i>Junxu Liu, Jian Lou, Li Xiong, Jinfei Liu, Xiaofeng Meng</i>	
Popularity Prediction for Social Media over Arbitrary Time Horizons.....	841
<i>Daniel Haimovich, Dmytro Karamshuk, Thomas J. Leeper, Evgeniy Riabenko, Milan Vojnovic</i>	
LANNS: A Web-Scale Approximate Nearest Neighbor Lookup System .....	850
<i>Ishita Doshi, Dhritiman Das, Ashish Bhutani, Rajeev Kumar, Rushi Bhatt, Niranjana Balasubramanian</i>	
Fast Detection of Denial Constraint Violations .....	859
<i>Eduardo H. M. Pena, Eduardo Cunha De Almeida, Felix Naumann</i>	
Chukonu: A Fully-Featured Big Data Processing System by Efficiently Integrating a Native Compute Engine into Spark.....	872
<i>Bowen Yu, Guanyu Feng, Huanqi Cao, Xiaohan Li, Zhenbo Sun, Haojie Wang, Xiaowei Zhu, Weimin Zheng, Wenguang Chen</i>	
COMET: A Novel Memory-Efficient Deep Learning Training Framework by Using Error-Bounded Lossy Compression.....	886
<i>Sian Jin, Chengming Zhang, Xintong Jiang, Yunhe Feng, Hui Guan, Guanpeng Li, Shuaiwen Song, Dingwen Tao</i>	
Federated Matrix Factorization with Privacy Guarantee .....	900



*Zitao Li, Bolin Ding, Ce Zhang, Ninghui Li, Jingren Zhou*

Scalable Robust Graph Embedding with Spark ..... 914  
*Chi Thang Duong, Dung Trung Hoang, Hongzhi Yin, Matthias Weidlich, Quoc Viet Hung Nguyen, Karl Aberer*

Database Workload Characterization with Query Plan Encoders ..... 923  
*Debjyoti Paul, Jie Cao, Feifei Li, Vivek Srikumar*

New Query Optimization Techniques in the Spark Engine of Azure Synapse ..... 936  
*Abhishek Modi, Kaushik Rajan, Srinivas Thimmaiah, Prakhar Jain, Swinky Mann, Ayushi Agarwal, Ajith Shetty, Shahid K I, Ashit Gosalia, Partho Sarthi*

DQDF: Data-Quality-Aware Dataframes ..... 949  
*Phanwadee Sinthong, Dhaval Patel, Nianjun Zhou, Shrey Shrivastava, Arun Iyengar, Anuradha Bhamidipaty*

Retrofitting GDPR Compliance onto Legacy Databases ..... 958  
*Archita Agarwal, Marilyn George, Aaron Jeyaraj, Malte Schwarzkopf*

AutoCTS: Automated Correlated Time Series Forecasting ..... 971  
*Xinle Wu, Dalin Zhang, Chenjuan Guo, Chaoyang He, Bin Yang, Christian S. Jensen*

Replicated Layout for In-Memory Database Systems ..... 984  
*Sivaprasad Sudhir, Michael Cafarella, Samuel Madden*

## TABLE OF CONTENTS

### Research Papers

Projection-Compliant Database Generation .....	998
<i>Anupam Sanghi, Shadab Ahmed, Jayant R Haritsa</i>	
Making RDBMSs Efficient on Graph Workloads Through Predefined Joins.....	1011
<i>Guodong Jin, Semih Salihoglu</i>	
Ranked Enumeration of Join Queries with Projections .....	1024
<i>Shaleen Deep, Xiao Hu, Paraschos Koutris</i>	
Hippo: Sharing Computations in Hyper-Parameter Optimization .....	1038
<i>Ahnjae Shin, Joo Seong Jeong, Do Yoon Kim, Soyoung Jung, Byung-gon Chun</i>	
DSON: JSON CRDT Using Delta-Mutations For Document Stores.....	1053
<i>Arik Rinberg, Tomer Solomon, Roei Shlomo, Guy Khazma, Gal Lushi, Idit Keidar, Paula Ta-shma</i>	
A Neural Database for Differentially Private Spatial Range Queries.....	1066
<i>Sepanta Zeighami, Ritesh Ahuja, Gabriel Ghinita, Cyrus Shahabi</i>	
A Critical Analysis of Recursive Model Indexes.....	1079
<i>Marcel Maltry, Jens Dittrich</i>	
Hybrid Blockchain Database Systems: Design and Performance .....	1092
<i>Zerui Ge, Dumitrel Loghin, Beng Chin Ooi, Pingcheng Ruan, Tianwen Wang</i>	
Threshold Queries in Theory and in the Wild.....	1105
<i>Angela Bonifati, Stefania Dumbrava, George Fletcher, Jan Hidders, Matthias Hofer, Wim Martens, Filip Murlak, Joshua Shinavier, Sławek Staworko, Dominik Tomaszuk</i>	
User-Defined Operators: Efficiently Integrating Custom Algorithms into Modern Databases.....	1119
<i>Moritz Sichert, Thomas Neumann</i>	

## TABLE OF CONTENTS

### Research Papers

PACK: An Efficient Partition-based Distributed Agglomerative Hierarchical Clustering Algorithm for Deduplication .....	1132
<i>Yue Wang, Vivek Narasayya, Yeye He, Surajit Chaudhuri</i>	
A Near-Optimal Approach to Edge Connectivity-Based Hierarchical Graph Decomposition.....	1146
<i>Lijun Chang, Zhiyi Wang</i>	
Hu-Fu: Efficient and Secure Spatial Queries over Data Federation .....	1159
<i>Yongxin Tong, Xuchen Pan, Yuxiang Zeng, Yexuan Shi, Chunbo Xue, Zimu Zhou, Xiaofei Zhang, Lei Chen, Yi Xu, Ke Xu, Weifeng Lv</i>	
Sortledton: a Universal, Transactional Graph Data Structure .....	1173
<i>Per Fuchs, Jana Giceva, Domagoj Margan</i>	
NBTree: a Lock-free PM-friendly Persistent B+-Tree for eADR-enabled PM Systems.....	1187
<i>Bowen Zhang, Shengan Zheng, Zhenlin Qi, Linpeng Huang</i>	
TranAD: Deep Transformer Networks for Anomaly Detection in Multivariate Time Series Data.....	1201
<i>Shreshth Tuli, Giuliano Casale, Nicholas R Jennings</i>	
SpaceSaving± An Optimal Algorithm for Frequency Estimation and Frequent items in the Bounded Deletion Model .....	1215
<i>Fuheng Zhao, Divy Agrawal, Amr El Abbadi, Ahmed Metwally</i>	
ByteGNN: Efficient Graph Neural Network Training at Large Scale .....	1228
<i>Chenguang Zheng, Hongzhi Chen, Yuxuan Cheng, Zhezheng Song, Yifan Wu, Changji Li, James Cheng, Hao Yang, Shuai Zhang</i>	
Query Driven-Graph Neural Networks for Community Search: From Non-Attributed, Attributed, to Interactive Attributed.....	1243
<i>Yuli Jiang, Yu Rong, Hong Cheng, Xin Huang, Kangfei Zhao, Junzhou Huang</i>	
Hyper-Tune: Towards Efficient Hyper-parameter Tuning at Scale .....	1256
<i>Yang Li, Yu Shen, Huaijun Jiang, Wentao Zhang, Jixiang Li, Ji Liu, Ce Zhang, Bin Cui</i>	
Multivariate Correlations Discovery in Static and Streaming Data .....	1266
<i>Koen Minartz, Jens D'hondt, Odysseas Papapetrou</i>	
Moneyball: Proactive Auto-Scaling in Microsoft Azure SQL Database Serverless .....	1279
<i>Olga Poppe, Qun Guo, Willis Lang, Pankaj Arora, Morgan Oslake, Shize Xu, Ajay Kalhan</i>	
PGE: Robust Product Graph Embedding Learning for Error Detection .....	1288
<i>Kewei Cheng, Xian Li, Yifan Xu, Xin Dong, Yizhou Sun</i>	

CHEX: Multiversion Replay with Ordered Checkpoints ..... 1297  
*Naga Nithin Manne, Shilvi Satpati, Tanu Malik, Amitabha Bagchi, Ashish Gehani, Amitabh Chaudhary*

## TABLE OF CONTENTS

### Research Papers

Prefix Filter: Practically and Theoretically Better Than Bloom .....	1311
<i>Tomer Even, Guy Even, Adam Morrison</i>	
Scalar DL: Scalable and Practical Byzantine Fault Detection for Transactional Database Systems ...	1324
<i>Hiroyuki Yamada, Jun Nemoto</i>	
In-Network Leaderless Replication for Distributed Data Stores .....	1337
<i>Gyuyeong Kim, Wonjun Lee</i>	
Fast Algorithms for Core Maximization on Large Graphs .....	1350
<i>Xin Sun, Xin Huang, Di Jin</i>	
NLC: Search Correlated Window Pairs on Long Time Series .....	1363
<i>Shuye Pan, Peng Wang, Chen Wang, Wei Wang, Jianmin Wang</i>	
Edge-based Local Push for Personalized PageRank .....	1376
<i>Hanzhi Wang, Zhewei Wei, Junhao Gan, Ye Yuan, Xiaoyong Du, Ji-rong Wen</i>	
Continuous Social Distance Monitoring in Indoor Space .....	1390
<i>Harry Kai-ho Chan, Huan Li, Xiao Li, Hua Lu</i>	
An In-Depth Study of Continuous Subgraph Matching .....	1403
<i>Xibo Sun, Shixuan Sun, Qiong Luo, Bingsheng He</i>	
OnlineSTL: Scaling Time Series Decomposition by 100x .....	1417
<i>Abhinav Mishra, Ram Sriharsha, Sichen Zhong</i>	
Stingy Sketch: A Sketch Framework for Accurate and Fast Frequency Estimation .....	1426
<i>Haoyu Li, Qizhi Chen, Yixin Zhang, Tong Yang, Bin Cui</i>	
A Study of Database Performance Sensitivity to Experiment Settings .....	1439
<i>Yang Wang, Miao Yu, Yujie Hui, Fang Zhou, Yuyang Huang, Rui Zhu, Xueyuan Ren, Tianxi Li, Xiaoyi Lu</i>	
The Inherent Time Complexity and An Efficient Algorithm for Subsequence Matching Problem .....	1453
<i>Zemin Chao, Hong Gao, Yinan An, Jianzhong Li</i>	
Selective Data Acquisition in the Wild for Model Charging .....	1466
<i>Chengliang Chai, Jiabin Liu, Nan Tang, Guoliang Li, Yuyu Luo</i>	
Discovering Association Rules from Big Graphs .....	1479
<i>Wenfei Fan, Wenzhi Fu, Ruochun Jin, Ping Lu, Chao Tian</i>	

DeepTEA: Effective and Efficient Online Time-dependent Trajectory Outlier Detection .....	1493
<i>Xiaolin Han, Reynold Cheng, Chenhao Ma, Tobias Grubenman</i>	
Entity Resolution On-Demand .....	1506
<i>Giovanni Simonini, Luca Zecchini, Sonia Bergamaschi, Felix Naumann</i>	

## TABLE OF CONTENTS

### Research Papers

ForBackBench: A Benchmark for Chasing vs. Query-Rewriting .....	1519
<i>Afnan G Alhazmi, Tom Blount, George Konstantinidis</i>	
Accurate Summary-based Cardinality Estimation Through the Lens of Cardinality Estimation Graphs .....	1533
<i>Jeremy Chen, Yuqing Huang, Mushi Wang, Semih Salihoglu, Kenneth Salem</i>	
Distributed D-core Decomposition over Large Directed Graphs .....	1546
<i>Xuankun Liao, Qing Liu, Jiaxin Jiang, Xin Huang, Jianliang Xu, Byron Choi</i>	
Efficient Maximal Biclique Enumeration for Large Sparse Bipartite Graphs .....	1559
<i>Lu Chen, Chengfei Liu, Rui Zhou, Jiajie Xu, Jianxin Li</i>	
TGL: A General Framework for Temporal GNN Training on Billion-Scale Graphs .....	1572
<i>Hongkuan Zhou, Da Zheng, Israt Nisa, Vassilis N. Ioannidis, Xiang Song, George Karypis</i>	
Distributed Learning of Fully Connected Neural Networks using Independent Subnet Training .....	1581
<i>Binhang Yuan, Cameron Wolfe, Chen Dun, Yuxin Tang, Anastasios Kyrillidis, Chris Jermaine</i>	
Netherite: Efficient Execution of Serverless Workflows .....	1591
<i>Sebastian C Burckhardt, Badrish Chandramouli, Chris Gillum, David A Justo, Konstantinos Kallas, Connor McMahon, Christopher Meiklejohn, Xiangfeng Zhu</i>	
Endure: A Robust Tuning Paradigm for LSM Trees Under Workload Uncertainty .....	1605
<i>Andy Huynh, Harshal Chaudhari, Evimaria Terzi, Manos Athanassoulis</i>	
An I/O-Efficient Disk-based Graph System for Scalable Second-Order Random Walk of Large Graphs .....	1619
<i>Hongzheng Li, Yingxia Shao, Junping Du, Bin Cui, Lei Chen</i>	
SNARF: A Learning-Enhanced Range Filter .....	1632
<i>Kapil Vaidya, Tim Kraska, Subarna Chatterjee, Eric R Knorr, Michael Mitzenmacher, Stratos Idreos</i>	
DLCR: Efficient Indexing for Label-Constrained Reachability Queries on Large Dynamic Graphs .....	1645
<i>Xin Chen, You Peng, Sibor Wang, Jeffrey Xu Yu</i>	
QueryFormer: A Tree Transformer Model for Query Plan Representation .....	1658
<i>Yue Zhao, Gao Cong, Jiachen Shi, Chunyan Miao</i>	
Index Checkpoints for Instant Recovery in In-Memory Database Systems .....	1671
<i>Leon Lee, Siphrey Xie, Yunus Ma, Shimin Chen</i>	

MATE: Multi-Attribute Table Extraction.....	1684
<i>Mahdi Esmailoghli, Jorge Arnulfo Quiane Ruiz, Ziawasch Abedjan</i>	
TSB-UAD: An End-to-End Benchmark Suite for Univariate Time-Series Anomaly Detection .....	1697
<i>John Paparrizos, Yuhao Kang, Paul Boniol, Ruey Tsay, Themis Palpanas, Michael Franklin</i>	
A Critical Re-evaluation of Neural Methods for Entity Alignment.....	1712
<i>Manuel Leone, Stefano Huber, Akhil Arora, Alberto Garcia-duran, Robert West</i>	
Analyzing How BERT Performs Entity Matching.....	1726
<i>Matteo Paganelli, Francesco Del Buono, Andrea Baraldi, Francesco Guerra</i>	



## TABLE OF CONTENTS

### Research Papers

Scalable Byzantine Fault Tolerance via Partial Decentralization .....	1739
<i>Balaji Arun, Binoy Ravindran</i>	
Efficient and Error-bounded Spatiotemporal Quantile Monitoring in Edge Computing Environments	1753
<i>Huan Li, Lanjing Yi, Bo Tang, Hua Lu, Christian S Jensen</i>	
HDPView: Differentially Private Materialized View for Exploring High Dimensional Relational Data..	1766
<i>Fumiyuki Kato, Tsubasa Takahashi, Shun Takagi, Yang Cao, Seng Pei Liew, Masatoshi Yoshikawa</i>	
Anomaly Detection in Time Series: A Comprehensive Evaluation.....	1779
<i>Sebastian Schmidl, Phillip Wenig, Thorsten Papenbrock</i>	
Guided Exploration of Data Summaries.....	1798
<i>Brit Youngmann, Sihem Amer-yahia, Aurélien Personnaz</i>	
Facilitating Database Tuning with Hyper-Parameter Optimization: A Comprehensive Experimental Evaluation.....	1808
<i>Xinyi Zhang, Zhuo Chang, Yang Li, Hong Wu, Jian Tan, Feifei Li, Bin Cui</i>	
Efficient Secure and Verifiable Location-Based Skyline Queries over Encrypted Data .....	1822
<i>Zuan Wang, Xiaofeng Ding, Hai Jin, Pan Zhou</i>	
AB-tree: Index for Concurrent Random Sampling and Updates .....	1835
<i>Zhuoyue Zhao, Dong Xie, Feifei Li</i>	
On Repairing Timestamps for Regular Interval Time Series .....	1848
<i>Chenguang Fang, Shaoxu Song, Yinan Mei</i>	
Towards Event Prediction in Temporal Graphs.....	1861
<i>Wenfei Fan, Ruochun Jin, Ping Lu, Chao Tian, Ruiqi Xu</i>	
Decentralized Crowdsourcing for Human Intelligence Tasks with Efficient On-Chain Cost .....	1875
<i>Yihuai Liang, Yan Li, Byeong-seok Shin</i>	
Towards Distributed Bitruss Decomposition on Bipartite Graphs.....	1889
<i>Yue Wang, Ruiqi Xu, Xun Jian, Alexander Zhou, Lei Chen</i>	
Generalized Supervised Meta-blocking .....	1902
<i>Luca Gagliardelli, George Papadakis, Giovanni Simonini, Sonia Bergamaschi, Themis Palpanas</i>	
Your Read is Our Priority in Flash Storage.....	1911
<i>Mijin An, Soojun Im, Dawoon Jung, Sang Won Lee</i>	

New Wine in an Old Bottle: Data-aware Hash Functions for Bloom Filters .....	1924
<i>Arindam Bhattacharya, Chathur Gudesu, Amitabha Bagchi, Srikanta Bedathur</i>	
SANCUS: Staleness-Aware Communication-Avoiding Full-Graph Decentralized Training in Large-Scale Graph Neural Networks .....	1937
<i>Jingshu Peng, Zhao Chen, Yingxia Shao, Yanyan Shen, Lei Chen, Jiannong Cao</i>	
CORE: a Complex event Recognition Engine .....	1951
<i>Marco Bucchi, Alejandro Grez, Andres F Quintana, Cristian Riveros, Stijn Vansummeren</i>	
TAOBench: An End-to-End Benchmark for Social Networking Workloads.....	1965
<i>Audrey Cheng, Xiao Shi, Aaron N Kabcenell, Shilpa Lawande, Hamza Qadeer, Jason Chan, Harrison Tin, Ryan Zhao, Peter Bailis, Mahesh Balakrishnan, Nathan Bronson, Natacha Crooks, Ion Stoica</i>	

## TABLE OF CONTENTS

### Research Papers

VIP Hashing - Adapting to Skew in Popularity of Data on the Fly .....	1978
<i>Aarati Kakaraparthi, Jignesh Patel, Brian Kroth, Kwanghyun Park</i>	
Near-Data Processing in Database Systems on Native Computational Storage under HTAP Workloads .....	1991
<i>Tobias Vincon, Christian Knoedler, Leonardo Solis-vasquez, Arthur Bernhardt, Sajjad Tamimi, Lukas Weber, Florian Stock, Andreas Koch, Ilija Petrov</i>	
Hercules Against Data Series Similarity Search .....	2005
<i>Karima Echihabi, Panagiota Fatourou, Kostas Zoumpatianos, Themis Palpanas, Houda Benbrahim</i>	
DISTILL: Low-Overhead Data-Driven Techniques for Filtering and Costing Indexes for Scalable Index Tuning .....	2019
<i>Tarique Siddiqui, Wentao Wu, Vivek Narasayya, Surajit Chaudhuri</i>	
Optimizing Machine Learning Inference Queries with Correlative Proxy Models .....	2032
<i>Zhihui Yang, Zuozhi Wang, Yicong Huang, Yao Lu, Chen Li, X. Sean Wang</i>	
Banyan: A Scoped Dataflow Engine for Graph Query Service .....	2045
<i>Li Su, Xiaoming Qin, Zichao Zhang, Rui Yang, Le Xu, Indranil Gupta, Wenyuan Yu, Zeng Kai, Jingren Zhou</i>	
Frequency Estimation Under Multiparty Differential Privacy: One-shot and Streaming .....	2058
<i>Ziyue Huang, Yuan Qiu, Ke Yi, Graham Cormode</i>	
Optimizing Inference Serving on Serverless Platforms .....	2071
<i>Ahsan Ali, Riccardo Pincioli, Feng Yan, Evgenia Smirni</i>	
Columnar Formats for Schemaless LSM-based Document Stores .....	2085
<i>Wail Y Alkowaleet, Michael Carey</i>	
Efficient Shortest Path Counting on Large Road Networks .....	2098
<i>Yu-xuan Qiu, Dong Wen, Lu Qin, Wentao Li, Ronghua Li, Ying Zhang</i>	
Towards Communication-efficient Vertical Federated Learning Training via Cache-enabled Local Update .....	2111
<i>Fangcheng Fu, Xupeng Miao, Jiawei Jiang, Huanran Xue, Bin Cui</i>	
DESIRE: An Efficient Dynamic Cluster-based Forest Indexing for Similarity Search in Multi-Metric Spaces .....	2121
<i>Yifan Zhu, Lu Chen, Yunjun Gao, Baihua Zheng, Pengfei Wang</i>	
ABC: Attributed Bipartite Co-clustering .....	2134
<i>Junghoon Kim, Kaiyu Feng, Gao Cong, Diwen Zhu, Wenyuan Yu, Chunyan Miao</i>	

Time Series Data Encoding for Efficient Storage: A Comparative Analysis in Apache IoTDB .....	2148
<i>Jinzhao Xiao, Yuxiang Huang, Changyu Hu, Shaoxu Song, Xiangdong Huang, Jianmin Wang</i>	
SA-LSM: Optimize Data Layout for LSM-tree Based Storage using Survival Analysis .....	2161
<i>Teng Zhang, Jian Tan, Xin Cai, Jiaying Wang, Feifei Li, Jianling Sun</i>	
Improving Matrix-vector Multiplication via Lossless Grammar-Compressed Matrices .....	2175
<i>Paolo Ferragina, Giovanni Manzini, Travis Gagie, Dominik Köppl, Gonzalo Navarro, Manuel Striani, Francesco Tosoni</i>	
NFL: Robust Learned Index via Distribution Transformation .....	2188
<i>Shangyu Wu, Yufei Cui, Jinghuan Yu, Xuan Sun, Tei-wei Kuo, Chun Jason Xue</i>	
LEGOStore: A Linearizable Geo-Distributed Store Combining Replication and Erasure Coding .....	2201
<i>Hamidreza Zare, Viveck Cadambe, Bhuvan Uргаonkar, Nader Alfares, Praneet Soni, Chetan Sharma, Arif Merchant</i>	
Misinformation Mitigation under Differential Propagation Rates and Temporal Penalties .....	2216
<i>Michael Simpson, Laks V.s. Lakshmanan, Farnoosh Hashemi</i>	
Serving Deep Learning Models with Deduplication from Relational Databases.....	2230
<i>Lixi Zhou, Jiaqing Chen, Amitabh Das, Hong Min, Lei Yu, Ming Zhao, Jia Zou</i>	
Density-optimized Intersection-free Mapping and Matrix Multiplication for Join-Project Operations	2244
<i>Zichun Huang, Shimin Chen</i>	
Design Trade-offs for a Robust Dynamic Hybrid Hash Join .....	2257
<i>Shiva Jahangiri, Michael Carey, Johann-christoph Freytag</i>	
YeSQL: “You extend SQL” with Rich and Highly Performant User-Defined Functions in Relational Databases.....	2270
<i>Yannis E Fofoulas, Alkis Simitsis, Eleftherios Stamatogiannakis, Yannis Ioannidis</i>	
Magic Shapes for SHACL Validation.....	2284
<i>Shqiporja Ahmetaj, Bianca Löhnert, Magdalena Ortiz, Mantas Simkus</i>	

## TABLE OF CONTENTS

### Research Papers

Succinct Graph Representations as Distance Oracles: An Experimental Evaluation .....	2297
<i>Arpit Merchant, Aristides Gionis, Michael Mathioudakis</i>	
Effective Community Search over Large Star-Schema Heterogeneous Information Networks .....	2307
<i>Yangqin Jiang, Yixiang Fang, Chenhao Ma, Xin Cao, Chunshan Li</i>	
A New Distributional Treatment for Time Series and An Anomaly Detection Investigation .....	2321
<i>Kai Ming Ting, Zongyou Liu, Hang Zhang, Ye Zhu</i>	
Witan: Unsupervised Labelling Function Generation for Assisted Data Programming .....	2334
<i>Benjamin Denham, Edmund M K Lai, Roopak Sinha, M. Asif Naeem</i>	
Skellam Mixture Mechanism: A Novel Approach to Federated Learning with Differential Privacy .....	2348
<i>Ergute Bao, Yizheng Zhu, Xiaokui Xiao, Yin Yang, Beng Chin Ooi, Benjamin Tan, Khin Mi Mi Aung</i>	
Zero-Shot Cost Models for Out-of-the-box Learned Cost Prediction .....	2361
<i>Benjamin Hilprecht, Carsten Binnig</i>	
Waffle: In-memory Grid Index for Moving Objects with Reinforcement Learning-based Configuration Tuning System .....	2375
<i>Dalsu Choi, Hyunsik Yoon, Hyubjin Lee, Yon Dohn Chung</i>	
Designing an Open Framework for Query Optimization and Compilation .....	2389
<i>Michael Jungmair, André Kohn, Jana Giceva</i>	
In-Page Shadowing and Two-Version Timestamp Ordering for Mobile DBMSs .....	2402
<i>Duy Lam Nguyen, Sang Won Lee, Beomseok Nam</i>	
RapidFlow: An Efficient Approach to Continuous Subgraph Matching .....	2415
<i>Shixuan Sun, Xibo Sun, Bingsheng He, Qiong Luo</i>	
A Scalable AutoML Approach Based on Graph Neural Networks.....	2428
<i>Mossad Helali, Essam Mansour, Ibrahim Abdelaziz, Julian Dolby, Kavitha Srinivas</i>	
Don't Be a Tattle-Tale: Preventing Leakages through Data Dependencies on Access Control Protected Data .....	2437
<i>Primal Pappachan, Shufan Zhang, Xi He, Sharad Mehrotra</i>	
Efficient Load-Balanced Butterfly Counting on GPU .....	2450
<i>Qingyu Xu, Feng Zhang, Zhiming Yao, Lv Lu, Xiaoyong Du, Dong Deng, Bingsheng He</i>	
PerMA-Bench: Benchmarking Persistent Memory Access .....	2463
<i>Lawrence Benson, Leon Papke, Tilmann Rabl</i>	

Evaluating Persistent Memory Range Indexes: Part Two .....	2477
<i>Yuliang He, Duo Lu, Kaisong Huang, Tianzheng Wang</i>	
Orchestrating Data Placement and Query Execution in Heterogeneous CPU-GPU DBMS .....	2491
<i>Bobbi W Yogatama, Weiwei Gong, Xiangyao Yu</i>	
Interactive Mining with Ordered and Unordered Attributes .....	2504
<i>Weicheng Wang, Raymond Chi-wing Wong</i>	
Fast Dataset Search with Earth Mover's Distance.....	2517
<i>Wenzhe Yang, Sheng Wang, Yuan Sun, Zhiyong Peng</i>	
AcX: System, Techniques, and Experiments for Acronym Expansion.....	2530
<i>João L. M. Pereira, João Casanova, Helena Galhardas, Dennis Shasha</i>	
G-Tran: A High Performance Distributed Graph Database with a Decentralized Architecture .....	2545
<i>Hongzhi Chen, Changji Li, Chenguang Zheng, Chenghuan Huang, Juncheng Fang, James Cheng, Jian Zhang</i>	
Tenant Placement in Over-subscribed Database-as-a-Service Clusters .....	2559
<i>Arnd Christian König, Yi Shan, Tobias Ziegler, Aarati Kakaraparthi, Willis Lang, Justin Moeller, Ajay Kalhan, Vivek Narasayya</i>	
Example-based Spatial Pattern Matching .....	2572
<i>Yue Chen, Kaiyu Feng, Gao Cong, Han Mao Kiah</i>	
NeuChain: A Fast Permissioned Blockchain System with Deterministic Ordering .....	2585
<i>Zeshun Peng, Yanfeng Zhang, Qian Xu, Haixu Liu, Yuxiao Gao, Xiaohua Li, Ge Yu</i>	
AIM: An Adaptive and Iterative Mechanism for Differentially Private Synthetic Data.....	2599
<i>Ryan Mckenna, Brett Mullins, Daniel Sheldon, Gerome Miklau</i>	
Troubles with Nulls, Views from the Users .....	2613
<i>Etienne Jr Toussaint, Paolo Guagliardo, Leonid Libkin, Juan Sequeda</i>	
Ginex: SSD-enabled Billion-scale Graph Neural Network Training on a Single Machine via Provably Optimal In-memory Caching .....	2626
<i>Yeonhong Park, Sunhong Min, Jae W. Lee</i>	
Shortest-Path Queries on Complex Networks: Experiments, Analyses, and Improvement .....	2640
<i>Junhua Zhang, Wentao Li, Long Yuan, Lu Qin, Ying Zhang, Lijun Chang</i>	
MIDE: Accuracy Aware Minimally Invasive Data Exploration For Decision Support .....	2653
<i>Sameera Ghayyur, Dhruvajyoti Ghosh, Xi He, Sharad Mehrotra</i>	
JENNER: Just-in-time Enrichment in Query Processing .....	2666
<i>Dhruvajyoti Ghosh, Peeyush Gupta, Sharad Mehrotra, Roberto Yus, Yasser Altowim</i>	
CARMI: A Cache-Aware Learned Index with a Cost-based Construction Algorithm .....	2679
<i>Jiaoyi Zhang, Yihan Gao</i>	
Maximizing Fair Content Spread via Edge Suggestion in Social Networks .....	2692
<i>Ian Swift, Sana Ebrahimi, Azade Nova, Abolfazl Asudeh</i>	

Turbo-Charging SPJ Query Plans with Learned Physical Join Operator Selections.....	2706
<i>Axel Hertzschuch, Claudio Hartmann, Dirk Habich, Wolfgang Lehner</i>	
Finding Locally Densest Subgraphs: A Convex Programming Approach.....	2719
<i>Chenhao Ma, Reynold Cheng, Laks V.s. Lakshmanan, Xiaolin Han</i>	
Decoupled Dynamic Spatial-Temporal Graph Neural Network for Traffic Forecasting.....	2733
<i>ZeZhi Shao, Zhao Zhang, Wei Wei, Fei Wang, Yongjun Xu, Xin Cao, Christian S Jensen</i>	
Harmony: Overcoming the hurdles of GPU memory capacity to train massive DNN models on commodity servers .....	2747
<i>Youjie Li, Amar Phanishayee, Derek Murray, Jakub Tarnawski, Nam Sung Kim</i>	
On Shapley Value in Data Assemblage Under Independent Utility .....	2761
<i>Xuan Luo, Jian Pei, Zicun Cong, Cheng Xu</i>	
Volume Under the Surface: A New Accuracy Evaluation Measure for Time-Series Anomaly Detection .....	2774
<i>John Paparrizos, Paul Boniol, Themis Palpanas, Ruey Tsay, Aaron J Elmore, Michael Franklin</i>	
Algorithm and System Co-design for Efficient Subgraph-based Graph Representation Learning.....	2788
<i>Haoteng Yin, Muhan Zhang, Yanbang Wang, Jianguo Wang, Pan Li</i>	
Memory-Optimized Multi-Version Concurrency Control for Disk-Based Database Systems.....	2797
<i>Michael Freitag, Alfons Kemper, Thomas Neumann</i>	
Query Processing on Tensor Computation Runtimes .....	2811
<i>Dong He, Supun C Nakandala, Dalitso Banda, Rathijit Sen, Karla Saur, Kwanghyun Park, Carlo Curino, Jesús Camacho-rod�ríguez, Konstantinos Karanasos, Matteo Interlandi</i>	
Reliable Community Search in Dynamic Networks.....	2826
<i>Yifu Tang, Jianxin Li, Nur Al Hasan Haldar, Ziyu Guan, Jiajie Xu, Chengfei Liu</i>	
Qanaat: A Scalable Multi-Enterprise Permissioned Blockchain System with Confidentiality Guarantees .....	2839
<i>Mohammad Javad Amiri, Boon Thau Loo, Divy Agrawal, Amr El Abbadi</i>	
Fast Network K-function-based Spatial Analysis.....	2853
<i>Tsz Nam Chan, Leong Hou U, Yun Peng, Byron Choi, Jianliang Xu</i>	
Cost Modelling for Optimal Data Placement in Heterogeneous Main Memory.....	2867
<i>Robert Lasch, Thomas Legler, Norman May, Bernhard Scheirle, Kai-uwe Sattler</i>	
SwitchTx: Scalable In-Network Coordination for Distributed Transaction Processing .....	2881
<i>Junru Li, Youyou Lu, Yiming Zhang, Qing Wang, Zhuo Cheng, Keji Huang, Jiwu Shu</i>	
Plush: A Write-Optimized Persistent Log-Structured Hash-Table.....	2895
<i>Lukas Vogel, Alexander Van Renen, Satoshi Imamura, Jana Giceva, Thomas Neumann, Alfons Kemper</i>	
Effective Indexing for Dynamic Structural Graph Clustering .....	2908
<i>Fangyuan Zhang, Sib0 Wang</i>	

CodexDB: Synthesizing Code for Query Processing from Natural Language Instructions using GPT-3 Codex.....	2921
<i>Immanuel Trummer</i>	
UPLIFT: Parallelization Strategies for Feature Transformations in Machine Learning Workloads .....	2929
<i>Arnab Phani, Lukas Erlbacher, Matthias Boehm</i>	
Lotus: Scalable Multi-Partition Transactions on Single-Threaded Partitioned Databases.....	2939
<i>Xinjing Zhou, Xiangyao Yu, Goetz Graefe, Michael Stonebraker</i>	
LlamaTune: Sample-Efficient DBMS Configuration Tuning .....	2953
<i>Konstantinos Kanellis, Cong Ding, Brian Kroth, Andreas C Mueller, Carlo Curino, Shivaram Venkataraman</i>	
On-Demand State Separation for Cloud Data Warehousing.....	2966
<i>Christian Winter, Jana Giceva, Thomas Neumann, Alfons Kemper</i>	
BABOONS: Black-Box Optimization of Data Summaries in Natural Language .....	2980
<i>Immanuel Trummer</i>	
ConnectorX: Accelerating Data Loading From Databases to Dataframes .....	2994
<i>Xiaoying Wang, Weiyuan Wu, Jinze Wu, Yizhou Chen, Nick Zrymiak, Changbo Qu, Lampros Flokas, George Chow, Jiannan Wang, Tianzheng Wang, Eugene Wu, Qingqing Zhou</i>	
Are Updatable Learned Indexes Ready?.....	3004
<i>Chaichon Wongkham, Baotong Lu, Chris Liu, Zhicong Zhong, Eric Lo, Tianzheng Wang</i>	
A Scalable and Generic Approach to Range Joins.....	3018
<i>Maximilian Reif, Thomas Neumann</i>	
SCAR - Spectral Clustering Accelerated and Robustified .....	3031
<i>Ellen Hohma, Christian M.m. Frey, Anna Beer, Thomas Seidl</i>	
Rewriting the Infinite Chase .....	3045
<i>Michael Benedikt, Maxime Buron, Stefano Germano, Kevin Kappelmann, Boris Motik</i>	
Chimp: Efficient Lossless Floating Point Compression for Time Series Databases .....	3058
<i>Panagiotis Liakos, Katia Papakonstantinou, Yannis Kotidis</i>	
Spooky: Granulating LSM-Tree Compactions Correctly .....	3071
<i>Niv Dayan, Tamar Weiss, Shmuel Dashevsky, Michael Pan, Edward Bortnikov, Moshe Twitto</i>	
Identifying Similar-Bicliques in Bipartite Graphs.....	3085
<i>Kai Yao, Lijun Chang, Jeffrey Xu Yu</i>	
Fine-Grained Modeling and Optimization for Intelligent Resource Management in Big Data Processing .....	3098
<i>Chenghao Lyu, Qi Fan, Fei Song, Arnab Sinha, Yanlei Diao, Wei Chen, Li Ma, Yihui Feng, Yaliang Li, Kai Zeng, Jingren Zhou</i>	
FHL-Cube: Multi-Constraint Shortest Path Querying with Flexible Combination of Constraints .....	3112
<i>Ziyi Liu, Lei Li, Mengxuan Zhang, Wen Hua, Xiaofang Zhou</i>	



Tiresias: Enabling Predictive Autonomous Storage and Indexing .....	3126
<i>Michael Abebe, Horatiu Lazu, Khuzaima Daudjee</i>	
Towards Distribution-aware Query Answering in Data Markets .....	3137
<i>Abolfazl Asudeh, Fatemeh Nargesian</i>	
Cardinality Estimation of Approximate Substring Queries using Deep Learning .....	3145
<i>Suyong Kwon, Woohwan Jung, Kyuseok Shim</i>	
Containerized Execution of UDFs: An Experimental Evaluation .....	3158
<i>Karla Saur, Tara Mirmira, Konstantinos Karanasos, Jesús Camacho-rodríguez</i>	
Data Station: Delegated, Trustworthy, and Auditable Computation to Enable Data-Sharing Consortia with a Data Escrow .....	3172
<i>Siyuan Xia, Zhiru Zhu, Christopher Zhu, Jinjin Zhao, Kyle Chard, Aaron J Elmore, Ian Foster, Michael Franklin, Sanjay Krishnan, Raul Castro Fernandez</i>	
Optimizing Differentially-Maintained Recursive Queries on Dynamic Graphs .....	3186
<i>Khaled Ammar, Siddhartha Sahu, Semih Salihoglu, Tamer Özsu</i>	
Diversified Top-k Route Planning in Road Network .....	3199
<i>Zihan Luo, Lei Li, Mengxuan Zhang, Wen Hua, Yehong Xu, Xiaofang Zhou</i>	
Migrating Social Event Recommendation Over Microblogs .....	3213
<i>Xiangmin Zhou, Lei Chen</i>	
Spatial and Temporal Constrained Ranked Retrieval over Videos.....	3226
<i>Yueting Chen, Nick Koudas, Xiaohui Yu, Ziqiang Yu</i>	
SCARA: Scalable Graph Neural Networks with Feature-Oriented Optimization.....	3240
<i>Ningyi Liao, Dingheng Mo, Siqiang Luo, Xiang Li, Pengcheng Yin</i>	
Enabling Efficient and General Subpopulation Analytics in Multidimensional Data Streams.....	3249
<i>Antonis Manousis, Zhuo Cheng, Ran Ben Basat, Zaoxing Liu, Vyas Sekar</i>	
Dynamic Spanning Trees for Connectivity Queries on Fully-dynamic Undirected Graphs.....	3263
<i>Qing Chen, Oded Lachish, Sven Helmer, Michael H Böhlen</i>	

## TABLE OF CONTENTS

### Industrial Papers

Hardware Acceleration of Compression and Encryption in SAP HANA.....	3277
<i>Monica Chiosa, Fabio Maschi, Ingo Müller, Gustavo Alonso, Norman May</i>	
Frost: A Platform for Benchmarking and Exploring Data Matching Results.....	3292
<i>Martin Graf, Lukas Laskowski, Florian Papsdorf, Florian Sold, Roland Gremmelspacher, Felix Naumann, Fabian Panse</i>	
ByteGraph: A High-Performance Distributed Graph Database in ByteDance .....	3306
<i>Changji Li, Hongzhi Chen, Shuai Zhang, Yingqian Hu, Chao Chen, Zhenjie Zhang, Meng Li, Xiangchen Li, Dongqing Han, Xiaohui Chen, Xudong Wang, Huiming Zhu, Xuwei Fu, Tingwei Wu, Hongfei Tan, Hengtian Ding, Mengjin Liu, Kangcheng Wang, Ting Ye, Lei Li, Xin Li, Yu Wang, Chenguang Zheng, Hao Yang, James Cheng</i>	
CDI-E: An Elastic Cloud Service for Data Engineering.....	3319
<i>Prakash C Das, Shivangi Srivastava, Valentin Moskovich, Anmol Chaturvedi, Anant Mittal, Yongqin Xiao, Mosharaf Chowdhury</i>	
Operon: An Encrypted Database for Ownership-Preserving Data Management .....	3332
<i>Sheng Wang, Yiran Li, Huorong Li, Feifei Li, Chengjin Tian, Le Su, Yanshan Zhang, Yubing Ma, Lie Yan, Yuanyuan Sun, Xuntao Cheng, Xiaolong Xie, Yu Zou</i>	
Tair-PMem: a Fully Durable Non-Volatile Memory Database.....	3346
<i>Caixin Gong, Chengjin Tian, Zhengheng Wang, Sheng Wang, Xiyu Wang, Qiulei Fu, Wu Qin, Qian Long, Rui Chen, Jiang Qi, Ruo Wang, Guoyun Zhu, Chenghu Yang, Wei Zhang, Feifei Li</i>	
Trie memtables in Cassandra.....	3359
<i>Branimir Lambov</i>	
Velox: Meta’s Unified Execution Engine.....	3372
<i>Pedro Pedreira, Orri Erling, Maria Basmanova, Kevin Wilfong, Laith S Sakka, Krishna Pai, Wei He, Biswapesh Chattopadhyay</i>	
OceanBase: A 707 Million tpmC Distributed Relational Database System.....	3385
<i>Zhenkun Yang, Chuanhui Yang, Fusheng Han, Mingqiang Zhuang, Bing Yang, Zhifeng Yang, Xiaojun Cheng, Yuzhong Zhao, Wenhui Shi, Huafeng Xi, Huang Yu, Bin Liu, Yi Pan, Boxue Yin, Junquan Chen, Quanqing Xu</i>	
VRE: A Versatile, Robust, and Economical Trajectory Data System.....	3398
<i>Hai Lan, Jiong Xie, Zhifeng Bao, Feifei Li, Wei Tian, Fang Wang, Sheng Wang, Ailin Zhang</i>	

ByteHTAP: ByteDance’s HTAP System with High Data Freshness and Strong Data Consistency.....	3411
<i>Jianjun Chen, Yonghua Ding, Ye Liu, Fangshi Li, Li Zhang, Mingyi Zhang, Kui Wei, Lixun Cao, Dan Zou, Yang Liu, Lei Zhang, Rui Shi, Wei Ding, Kai Wu, Shangyu Luo, Jason Sun, Yuming Liang</i>	
Beaconnect: Continuous Web Performance A/B Testing at Scale.....	3425
<i>Wolfram Wingerath, Benjamin Wollmer, Markus Bestehorn, Stephan Succo, Sophie Ferrlein, Florian Bücklers, Jörn Domnik, Fabian Panse, Erik Witt, Anil Sener, Felix Gessert, Norbert Ritter</i>	
CloudJump: Optimizing Cloud Databases for Cloud Storages .....	3432
<i>Zongzhi Chen, Xinjun Yang, Feifei Li, Xuntao Cheng, Qingda Hu, Zheyu Miao, Rongbiao Xie, Xiaofei Wu, Kang Wang, Zhao Song, Haiqing Sun, Zechao Zhuang, Yuming Yang, Jie Xu, Liang Yin, Wenchao Zhou, Sheng Wang</i>	
DyHealth: Making Neural Networks Dynamic for Effective Healthcare Analytics .....	3445
<i>Kaiping Zheng, Shaofeng Cai, Horng-ruey Chua, Melanie Herschel, Meihui Zhang, Beng Chin Ooi</i>	
Blueprint: a constraint-solving approach for document extraction .....	3459
<i>Andrey Mishchenko, Dominique Danco, Abhilash Jindal, Adrian Blue</i>	
TencentCLS: The Cloud Log Service with High Query Performances .....	3472
<i>Muzhi Yu, Zhaoxiang Lin, Jinan Sun, Runyun Zhou, Guoqiang Jiang, Hua Huang, Shikun Zhang</i>	
Ganos: A Multidimensional, Dynamic, and Scene-Oriented Cloud-Native Spatial Database Engine ..	3483
<i>Jiong Xie, Zhen Chen, Jianwei Liu, Fang Wang, Feifei Li, Zhida Chen, Yinpei Liu, Songlu Cai, Zhenhua Fan, Fei Xiao, Yue Chen</i>	
Magma: A high data density storage engine used in Couchbase.....	3496
<i>Sarath Lakshman, Apaar Gupta, Rohan Suri, Scott D Lashley, John Liang, Srinath Duvuru, Ravi Mayuram</i>	
Doppler: Automated SKU Recommendation in Migrating SQL Workloads to the Cloud.....	3509
<i>Joyce Cahoon, Wenjing Wang, Yiwen Zhu, Katherine Lin, Sean Liu, Raymond Truong, Neetu Singh, Chengcheng Wan, Alexandra M Ciortea, Sreraman Narasimhan, Subru Krishnan</i>	
Meta's Next-generation Realtime Monitoring and Analytics Platform .....	3522
<i>Stavros Harizopoulos, Taylor Hopper, Morton Mo, Shyam Sundar Chandrasekaran, Tongguang Chen, Yan Cui, Nandini Ganesh, Gary Helmling, Hieu Pham, Sebastian Wong</i>	
SQLite: Past, Present, and Future .....	3535
<i>Kevin P Gaffney, Martin Prammer, Laurence C Brasfield, Richard Hipp, Dan R Kennedy, Jignesh Patel</i>	
Manu: A Cloud Native Vector Database Management System .....	3548
<i>Rentong Guo, Xiaofan Luan, Long Xiang, Xiao Yan, Xiaomeng Yi, Jigao Luo, Qianya Cheng, Weizhi Xu, Jiarui Luo, Frank Liu, Zhenshan Cao, Yanliang Qiao, Ting Wang, Bo Tang, Charles Xie</i>	

## Demonstrations

Automated Relational Data Explanation using External Semantic Knowledge.....	3562
<i>Sainyam Galhotra, Udayan Khurana</i>	
Kelpie: an Explainability Framework for Embedding-based Link Prediction Models .....	3566
<i>Andrea Rossi, Donatella Firmani, Paolo Merialdo, Tommaso Teofili</i>	

OREO: Detection of Cherry-picked Generalizations .....	3570
<i>Yin Lin, Brit Youngmann, Yuval Moskovitch, H. V. Jagadish, Tova Milo</i>	
DuckDB-Wasm: Fast Analytical Processing for the Web .....	3574
<i>André Kohn, Dominik Moritz, Mark Raasveldt, Hannes Mühleisen, Thomas Neumann</i>	
EasyDR: A Human-in-the-loop Error Detection&Repair Platform for Holistic Table Cleaning .....	3578
<i>Yihai Xi, Ning Wang, Xinyu Chen, Yiyi Zhang, Zilong Wang, Zhihong Xu, Yue Wang</i>	
Hu-Fu: A Data Federation System for Secure Spatial Queries.....	3582
<i>Xuchen Pan, Yongxin Tong, Chunbo Xue, Zimu Zhou, Junping Du, Yuxiang Zeng, Yexuan Shi, Xiaofei Zhang, Lei Chen, Yi Xu, Ke Xu, Weifeng Lv</i>	
Demonstrating CAT: Synthesizing Data-Aware Conversational Agents for Transactional Databases	3586
<i>Marius Gassen, Benjamin Hättasch, Benjamin Hilprecht, Nadja Geisler, Alexander Fraser, Carsten Binnig</i>	
EDA4SUM: Guided Exploration of Data Summaries .....	3590
<i>Aurélien Personnaz, Brit Youngmann, Sihem Amer-yahia</i>	
CaJaDE: Explaining Query Results by Augmenting Provenance with Context .....	3594
<i>Chenjie Li, Juseung Lee, Zhengjie Miao, Boris Glavic, Sudeepa Roy</i>	
Share the Tensor Tea: How Databases can Leverage the Machine Learning Ecosystem .....	3598
<i>Yuki Asada, Victor Fu, Apurva Gandhi, Advitya Gemawat, Lihao Zhang, Vivek Gupta, Ehi Nosakhare, Dalitso Banda, Rathijit Sen, Matteo Interlandi</i>	
MOCHA: A Tool for Visualizing Impact of Operator Choices in Query Execution Plans for Database Education .....	3602
<i>Jess Tan, Desmond Yeoh, Rachael Neoh, Huey Eng Chua, Sourav S Bhowmick</i>	
LIBKDV: A Versatile Kernel Density Visualization Library for Geospatial Analytics.....	3606
<i>Tsz Nam Chan, Pak Lon Ip, Kaiyan Zhao, Leong Hou U, Byron Choi, Jianliang Xu</i>	
A Demonstration of Multi-Region CockroachDB .....	3610
<i>Arul Ajmani, Aayush Shah, Alexander Shraer, Adam Storm, Rebecca Taft, Oliver Tan, Nathan Vanbenschoten</i>	
DPDS: Assisting Data Science with Data Provenance.....	3614
<i>Adriane Chapman, Luca Lauro, Paolo Missier, Riccardo Torlone</i>	
POEM: Pattern-Oriented Explanations of CNN Models .....	3618
<i>Vargha Dadvar, Lukasz Golab, Divesh Srivastava</i>	
WebArrayDB: A Geospatial Array DBMS in Your Web Browser.....	3622
<i>Ramon Antonio Rodrigues Zalipynis, Nikita A Terlych</i>	
AutoDI: Towards an Automatic Plan Regression Analysis.....	3626
<i>Hai Lan, Yuanjia Zhang, Zhifeng Bao, Yu Dong, Dongxu Huang, Liu Tang, Jian Zhang</i>	
PHoCus: Efficiently Archiving Photos .....	3630
<i>Susan B Davidson, Shay Gershtein, Tova Milo, Slava Novgorodov, May Shoshan</i>	

VINCENT: Towards Efficient Exploratory Subgraph Search in Graph Databases .....	3634
<i>Kai Huang, Qingqing Ye, Jing Zhao, Xi Zhao, Haibo Hu, Xiaofang Zhou</i>	
ActivePDB: Active Probabilistic Databases .....	3638
<i>Osnat Drien, Matanya Freiman, Yael Amsterdamer</i>	
CERTEM: Explaining and Debugging Black-box Entity Resolution Systems with CERTA.....	3642
<i>Tommaso Teofili, Donatella Firmani, Nick Koudas, Paolo Merialdo, Divesh Srivastava</i>	
Satellite Image Search in AgoraEO .....	3646
<i>Ahmet Kerem Aksoy, Pavel Dushev, Eleni Tzirita Zacharitou, Holmer Hensen, Marcela Charfuelan, Jorge Arnulfo Quiane Ruiz, Begum Demir, Volker Markl</i>	
SENSOR: Data-driven Construction of Sketch-based Visual Query Interfaces for Time Series Data .	3650
<i>Li Yan, Nerissa Xu, Guozhong Li, Sourav S Bhowmick, Byron Choi, Jianliang Xu</i>	
DiscoPG: Property Graph Schema Discovery and Exploration .....	3654
<i>Angela Bonifati, Stefania G. Dumbrava, Emile Martinez, Fatemeh Ghasemi, Malo Jaffré, Pacome Luton, Thomas Pickles</i>	
SA-Q: Observing, Evaluating, and Enhancing the Quality of the Results of Sentiment Analysis Tools .....	3658
<i>Wissam Mammam Kouadri, Salima Benbernou, Mourad Ouziri, Themis Palpanas, Iheb Benamor</i>	
SmartBench: Demonstrating Automatic Generation of Comprehensive Benchmarks for Question Answering Over Knowledge Graphs.....	3662
<i>Abdelghny Orogat, Ahmed El-robby</i>	
DADER: Hands-Off Entity Resolution with Domain Adaptation.....	3666
<i>Jianhong Tu, Xiaoyue Han, Ju Fan, Nan Tang, Chengliang Chai, Guoliang Li, Xiaoyong Du</i>	
Sigma Workbook: A Spreadsheet for Cloud Data Warehouses.....	3670
<i>James L Gale, Max Seiden, Deepanshu Utkarsh, Jason Frantz, Rob Woollen, Cagatay Demiralp</i>	
ReMac: A Matrix Computation System with Redundancy Elimination.....	3674
<i>Zihao Chen, Zhizhen Xu, Baokun Han, Chen Xu, Weining Qian, Aoying Zhou</i>	
TimeEval: A Benchmarking Toolkit for Time Series Anomaly Detection Algorithms.....	3678
<i>Phillip Wenig, Sebastian Schmidl, Thorsten Papenbrock</i>	
DBMS Annihilator: A High-Performance Database Workload Generator in Action.....	3682
<i>Alberto Lerner, Matthias Jasný, Theo Jepsen, Carsten Binnig, Philippe Cudre-mauroux</i>	
FedTSC: A Secure Federated Learning System for Interpretable Time Series Classification .....	3686
<i>Zhiyu Liang, Hongzhi Wang</i>	
AMRAS: A Visual Analysis System for Spatial Crowdsourcing.....	3690
<i>Qingshun Wu, Yafei Li, Huiling Li, Di Zhang, Guanglei Zhu</i>	
SparkCAD: Caching Anomalies Detector for Spark Applications .....	3694
<i>Hani Al-sayeh, Muhammad Attahir Jibril, Muhammad Waleed Bin Saeed, Kai-uwe Sattler</i>	

AvantGraph Query Processing Engine.....	3698
<i>Wilco Van Leeuwen, Thomas Mulder, Bram Van De Wall, George Fletcher, Nikolay Yakovets</i>	
Theseus: Navigating the Labyrinth of Time-Series Anomaly Detection .....	3702
<i>Paul Boniol, John Paparrizos, Yuhao Kang, Themis Palpanas, Ruey S. Tsay, Aaron J Elmore, Michael Franklin</i>	
A Demonstration of AutoOD: A Self-tuning Anomaly Detection System.....	3706
<i>Dennis M Hofmann, Peter Vannostrand, Huayi Zhang, Yizhou Yan, Lei Cao, Samuel Madden, Elke A Rundensteiner</i>	
Pipemizer: An Optimizer for Analytics Data Pipelines.....	3710
<i>Sunny Gakhar, Joyce Cahoon, Wangchao Le, Xiangnan Li, Kaushik Ravichandran, Hiren Patel, Marc Friedman, Brandon Haynes, Shi Qiao, Alekh Jindal, Jyoti Leeka</i>	
DORIAN in action: Assisted Design of Data Science Pipelines.....	3714
<i>Sergey Redyuk, Zoi Kaoudi, Sebastian Schelter, Volker Markl</i>	
WebMILE: Democratizing Network Representation Learning at Scale.....	3718
<i>Yuntian He, Yue Zhang, Saket Gurukar, Srinivasan Parthasarathy</i>	
Demonstrating Quest: A Query-Driven Framework to Explain Classification Models on Tabular Data .....	3722
<i>Nadja Geisler, Benjamin Hättasch, Carsten Binnig</i>	
IsoBugView: Interactively Debugging Isolation Bugs in Database Applications .....	3726
<i>Drew Ripberger, Yifan Gan, Xueyuan Ren, Spyros Blanas, Yang Wang</i>	
YeSQL: Rich User-Defined Functions without the Overhead .....	3730
<i>Yannis E Fofoulas, Alkis Simitsis, Yannis Ioannidis</i>	
Demonstration of Accelerating Machine Learning Inference Queries with Correlative Proxy Models	3734
<i>Zhihui Yang, Yicong Huang, Zuozhi Wang, Feng Gao, Yao Lu, Chen Li, X. Sean Wang</i>	
Demonstration of Collaborative and Interactive Workflow-Based Data Analytics in Texera.....	3738
<i>Xiaozhen Liu, Zuozhi Wang, Shengquan Ni, Sadeem Alsudais, Yicong Huang, Avinash Kumar, Chen Li</i>	
SimDB in Action: Road Traffic Simulations Completely Inside Array DBMS.....	3742
<i>Ramon Antonio Rodrigues Zalipynis</i>	

## Tutorials

Transformers for Tabular Data Representation: A Tutorial on Models and Applications.....	3746
<i>Gilbert Badaro, Paolo Papotti</i>	
Polyglot Data Management: State of the Art & Open Challenges.....	3750
<i>Felix Kiehn, Mareike Schmidt, Daniel Glake, Fabian Panse, Wolfram Wingerath, Benjamin Wollmer, Martin Poppinga, Norbert Ritter</i>	
Machine Programming: Turning Data into Programmer Productivity .....	3754
<i>Abdul Wasay, Nesime Tatbul, Justin Gottschlich</i>	

Cloud Databases: New Techniques, Challenges, and Opportunities .....	3758
<i>Guoliang Li, Haowen Dong, Chao Zhang</i>	
Modern Techniques for Querying Graph-Structured Relations: Foundations, System Implementations, and Open Challenges .....	3762
<i>Amine Mhedhbi, Semih Salihoglu</i>	
Densest Subgraph Discovery on Large Graphs: Applications, Challenges, and Techniques .....	3766
<i>Yixiang Fang, Wensheng Luo, Chenhao Ma</i>	
From BERT to GPT-3 Codex: Harnessing the Potential of Very Large Language Models for Data Management.....	3770
<i>Immanuel Trummer</i>	
The Past, Present and Future of Indexing on Persistent Memory.....	3774
<i>Kaisong Huang, Yuliang He, Tianzheng Wang</i>	
Unified Data Analytics: State-of-the-art and Open Problems .....	3778
<i>Zoi Kaoudi, Jorge Arnulfo Quiane Ruiz</i>	

## Keynotes

Big Graphs: Challenges and Opportunities .....	3782
<i>Wenfei Fan</i>	
Towards AI-Powered Data-Driven Education .....	3798
<i>Sihem Amer-yahia</i>	

## Endowment Awards

Heterogeneous Information Networks: the Past, the Present, and the Future.....	3807
<i>Yizhou Sun, Jiawei Han, Xifeng Yan, Philip S. Yu, Tianyi Wu</i>	
Toward Interpretable and Actionable Data Analysis with Explanations and Causality.....	3812
<i>Sudeepa Roy</i>	
Reflections On My Data Management Research Journey (VLDB Women in Database Research Award Talk) .....	N/A
<i>Fatma Özcan</i>	

## Panel

Panel: Startups Founded by Database Researchers.....	3823
<i>C. Mohan</i>	
Cloud Data Systems: What are the Opportunities for the Database Research Community? .....	3826
<i>Magdalena Balazinska, Surajit Chaudhuri, Anhai Doan, Joseph M. Hellerstein, Hanuma Kodavalla, Ippokratis Pandis, Matei Zaharia</i>	

## TABLE OF CONTENTS

High-dimensional Data Cubes.....	3828
<i>Sachin Basil John, Christoph Koch</i>	
Fast and Scalable Mining of Time Series Motifs with Probabilistic Guarantees.....	3841
<i>Matteo Ceccarello, Johann Gamper</i>	
FEDEX: An Explainability Framework for Data Exploration Steps .....	3854
<i>Daniel Deutch, Amir Gilad, Tova Milo, Amit Muaem, Amit Somech</i>	
Enabling Transparent Acceleration of Big Data Frameworks using Heterogeneous Hardware .....	3869
<i>Maria N Xekalaki, Juan Fumero, Athanasios Stratikopoulos, Katerina Doka, Christos Katsakioris, Constantinos Bitsakos, Nectarios Koziris, Christos Kotselidis</i>	
Discovering Polarization Niches via Dense Subgraphs with Attractors and Repulsers.....	3883
<i>Adriano Fazzino, Tommaso Lanciano, Riccardo Denni, Charalampos Tsourakakis, Francesco Bonchi</i>	
Sage: A System for Uncertain Network Analysis .....	3897
<i>Eunjae Lee, Sam H. Noh, Jiwon Seo</i>	
Mining Bursting Core in Large Temporal Graph.....	3911
<i>Hongchao Qin, Rong-Hua Li, Ye Yuan, Guoren Wang, Lu Qin, Zhiwei Zhang</i>	
Cost-based or Learning-based? A Hybrid Query Optimizer for Query Plan Selection .....	3924
<i>Xiang Yu, Chengliang Chai, Guoliang Li, Jiabin Liu</i>	
One Index for All Kernels (ONIAK): A Zero Re-Indexing LSH Solution to ANNS-ALT .....	3937
<i>Jingfan Meng, Huayi Wang, Jun Xu, Mitsunori Ogihara</i>	
Learned Index Benefits: Machine Learning Based Index Performance Estimation .....	3950
<i>Jiachen Shi, Gao Cong, Xiaoli Li</i>	
Online Ridesharing with Meeting Points.....	3963
<i>Jiachuan Wang, Peng Cheng, Libin Zheng, Lei Chen, Wenjie Zhang</i>	
Exploiting the Power of Equality-generating Dependencies in Ontological Reasoning .....	3976
<i>Luigi Bellomarini, Davide Benedetto, Matteo Brandetti, Emanuel Sallinger</i>	
No Repetition: Fast and Reliable Sampling with Highly Concentrated Hashing .....	3989
<i>Anders Aamand, Debarati Das, Evangelos Kipouridis, Jakob B. T. Knudsen, Peter M. R. Rasmussen, Mikkel Thorup</i>	
Witness Generation for JSON Schema.....	4002
<i>Lyes Attouche, Mohamed-Amine Baazizi, Dario Colazzo, Giorgio Ghelli, Carlo Sartiani, Stefanie Scherzinger</i>	
Towards Observability for Production Machine Learning Pipelines [Vision].....	4015
<i>Shreya Shankar, Aditya Parameswaran</i>	
DINOMO: An Elastic, Scalable, High-Performance Key-Value Store for Disaggregated Persistent Memory .....	4023
<i>Sekwon Lee, Soujanya Ponnappalli, Sharad Singhal, Marcos Aguilera, Kimberly Keeton, Vijay Chidambaram</i>	



Bolt-on, Compact, and Rapid Program Slicing for Notebooks [Scalable Data Science].....	4038
<i>Shreya Shankar, Stephen Macke, Sarah Chasins, Andrew Head, Aditya Parameswaran</i>	
Fairness Matters: A Tit-For-Tat Strategy Against Selfish Mining .....	4048
<i>Weijie Sun, Zihuan Xu, Lei Chen</i>	
SageDB: An Instance-Optimized Data Analytics System .....	4062
<i>Jialin Ding, Ryan C Marcus, Andreas Kipf, Vikram Nathan, Aniruddha Nrusimha, Kapil Vaidya, Alexander Van Renen, Tim Kraska</i>	
Budget-Conscious Fine-Grained Configuration Optimization for Spatio-Temporal Applications .....	4079
<i>Keven Richly, Rainer Schlosser, Martin Boissier</i>	
Nemo: Guiding and Contextualizing Weak Supervision for Interactive Data Programming .....	4093
<i>Cheng-Yu Hsieh, Jieyu Zhang, Alexander J Ratner</i>	

**Author Index**