

6            30  
—            7    5  
2019

**PROGRAM**  
**SIGMxD/**  
A M S P XDS  
T E R  
D A M **2X19**

## Welcome Message

This year, the conference is held in the city center of Amsterdam, capital of The Netherlands. Amsterdam is an internationally oriented city, home to people with origins from all over the world. This used to be already the case even back in the 16th and 17th century, when Amsterdam was the world's biggest trading and financial center; establishing the world's first stock exchange in 1602.

SIGMOD/PODS 2019 is held in the original Amsterdam Stock and Commodities Exchange, constructed by Dutch architect Berlage between 1896 and 1903, which now serves as the well-equipped Amsterdam Conference Center. This architect and his apprentices (the school of Berlage) left an important mark on the city, being responsible for a major expansion of the city in the early 20th century. The sculptures and drawings in the Exchange building refer to the people behind the commodities traded in the various rooms ("Effecten" - stock; "Graan" - grains), e.g., depicting farmers in the grain exchange room; as a reminder that trading affects society.

Amsterdam is a city that offers many cultural activities, including the world-famous classical Concertgebouw Orchestra, as well as many museums (Amsterdam Museum, Rijksmuseum, Rembrandthuis, Anne Frank Huis). In a slight break with SIGMOD tradition, the SIGMOD opening reception will be held one day later, on Tuesday night, when the SIGMOD/PODS attendees will have exclusive access to the Van Gogh museum. The Wednesday conference dinner is organized across the water in Amsterdam North, in Norderlicht Cafe in a festival-like environment. This used to be harbour area and was less-populated and industrial, but in the recent decade has become a hotspot for nightlife ac-

tivities.

Amsterdam is also increasingly a hub for data science companies and services, with multiple universities and CWI in the vicinity; which all participate in the organization of SIGMOD/PODS 2019. On Thursday night, after the SIGMOD program finishes, there will be a meetup of Amsterdam Data Science, where the local data science community will be able to mingle with our data management research community. The SIGMOD 2019 Research Program Committee consists of the Program Chair, two Program Vice Chairs, a core committee with 37 members, and a regular committee with 98 members. During the reviewing period, we solicited additional reviews from 16 external reviewers and occasional input from 10 assistant reviewers. The committee received 430 submissions, out of which 12 were desk-rejected (i.e., without review). There was no bidding; instead, reviewer assignments were made using input from Microsoft’s Conference Management System, the Toronto Paper Matching System, and the reviewers’ background (the detailed assignment procedure is described in a paper which has been submitted for publication to SIGMOD Record). The core committee members had (roughly) double the reviewing load of the regular committee members, and in addition acted as discussion leaders and meta-reviewers for their assigned papers. There were two rounds of submissions, with deadlines in July and November, respectively. Initially, each paper received three reviews. At this point authors could read the reviews and provide feedback about potential factual errors (disclosed to the reviewers) or sensitive issues about potential mishandling (confidentially to the chair). Two additional reviews were solicited for a paper if (a) the reviewers’ expertise level was suboptimal, or (b) if there was significant score discrepancy in the first three reviews, or (c) if it was heading for

rejection but had received a weak accept (or higher) by at least one reviewer. Papers were discussed extensively online; 10 were accepted based on the first round of reviews, while 311 were rejected. The authors of the remaining 97 papers were asked to revise their papers to address reviewers' criticisms; 78 revisions were ultimately accepted for a total of 88 papers which are presented in the research track. Finally, 12 papers were shepherded after acceptance to guarantee that the camera-ready version addresses all of the reviewers' comments.

**Stefan Manegold, Peter Boncz**  
SIGMOD General Chairs  
CWI, Netherlands

**Anastasia Ailamaki**  
SIGMOD Program Chair  
EPFL, Switzerland

# SIGMOD 2019 Sponsors & Supporters

## Supporters



## Diamond Sponsor

facebook

## Platinum Sponsors



Microsoft

ORACLE®

ableau

## Gold Sponsors



IBM Research AI



## Silver Sponsors



## Startup Sponsor



## Platinum Publisher



## Gold Publisher



## Silver Publishers



## Student support



**Sunday 06/30/2019 08:00-09:00**

Coffee + Light Breakfast

(Sun 08:00-09:00)

**Room:** Grote Zaal

**Sunday 06/30/2019 09:00-10:30**

Tutorial 1: part 1

(Sun 09:00-10:30)

**Room:** Veilingzaal

## **Towards Democratizing Relational Data Visualization**

*Nan Tang (Qatar Foundation), Eugene Wu (Columbia University), Guoliang Li (Tsinghua University)*

GRADES-NDA 2019: Session 1 (starts 08:30)

(Sun

09:00-10:30)

**Room:** Administratiezaal

## **GRADES-NDA 2019: Joint International Workshop on Graph Data Management Experiences & Systems and Network Data Analytics**

*Akhil Arora (EPFL), Arnab Bhattacharya (IIT Kanpur), George Fletcher (TU Eindhoven)*

DEEM 2019: Session 1

(Sun 09:00-10:30)

**Room:** Berlage Zaal

## **DEEM 2019: Workshop on Data Management for End-to-End Machine Learning**

*Sebastian Schelter (New York University), Neoklis Polyzotis (Google), Manasi Vartak (Massachusetts Institute of Technology), Stephan Seufert (Amazon Research)*

Tutorial 3: part 1

(Sun 09:00-10:30)

**Room:** Effectenbeurszaal

## **Database and Distributed Computing Foundations of Blockchains**

*Sujaya Maiyya (University of California, Santa Barbara), Victor Zakhary (University of California, Santa Barbara), Mohammad Javad Amiri (University of California, Santa Barbara), Divyakant Agrawal (University of California, Santa Barbara), Amr El Abbadi (University of California, Santa Barbara)*

DSMM 2019: Session 1

(Sun 09:00-10:30)

**Room:** Mendes da Costa Kamer

## **DSMM 2019: the 5th Workshop on Data Science for Macro-modeling with Financial and Economic Datasets**

*Douglas Burdick (IBM Almaden Research Center), Rajasekar Krishnamurthy (IBM T. J. Watson Research Center), Louiga Raschid (University of Maryland)*

**Sunday 06/30/2019 10:30-11:00**

Coffee

(Sun 10:30-11:00)

**Room:** Grote Zaal

**Sunday 06/30/2019 11:00-12:30**

Tutorial 1: part 2

(Sun 11:00-12:30)

**Room:** Veilingzaal

## **Towards Democratizing Relational Data Visualization**

*Nan Tang (Qatar Foundation), Eugene Wu (Columbia University)*

sity), Guoliang Li (Tsinghua University)

GRADES-NDA 2019: Session 2

(Sun 11:00-12:30)

**Room:** Administratiezaal

**GRADES-NDA 2019: Joint International Workshop  
on Graph Data Management Experiences & Systems  
and Network Data Analytics**

*Akhil Arora (EPFL), Arnab Bhattacharya (IIT Kanpur), George Fletcher (TU Eindhoven)*

DEEM 2019: Session 2

(Sun 11:00-12:30)

**Room:** Berlage Zaal

**DEEM 2019: Workshop on Data Management for  
End-to-End Machine Learning**

*Sebastian Schelter (New York University), Neoklis Polyzotis (Google), Manasi Vartak (Massachusetts Institute of Technology), Stephan Seufert (Amazon Research)*

Tutorial 3: part 2

(Sun 11:00-12:30)

**Room:** Effectenbeurszaal

**Database and Distributed Computing Foundations  
of Blockchains**

*Sujaya Maiyya (University of California, Santa Barbara), Victor Zakhary (University of California, Santa Barbara), Mohammad Javad Amiri (University of California, Santa Barbara), Divyakant Agrawal (University of California, Santa Barbara), Amr El Abbadi (University of California, Santa Barbara)*

DSMM 2019: Session 2

(Sun 11:00-12:30)

**Room:** Mendes da Costa Kamer

## **DSMM 2019: the 5th Workshop on Data Science for Macro-modeling with Financial and Economic Datasets**

*Douglas Burdick (IBM Almaden Research Center), Rajasekar Krishnamurthy (IBM T. J. Watson Research Center), Louiza Raschid (University of Maryland)*

**Sunday 06/30/2019 12:30-14:00**

Lunch

(Sun 12:30-14:00)

**Room:** Grote Zaal

**Sunday 06/30/2019 14:00-15:30**

Tutorial 2: part 1

(Sun 14:00-15:30)

**Room:** Veilingzaal

### **Exploring the Data Wilderness through Examples**

*Davide Mottin (Aarhus University), Matteo Lissandrini (Aalborg University), Yannis Velegrakis (Utrecht University), Themis Palpanas (Paris Descartes University)*

**GRADES-NDA 2019: Session 3**

(Sun 14:00-15:30)

**Room:** Administratiezaal

### **GRADES-NDA 2019: Joint International Workshop on Graph Data Management Experiences & Systems and Network Data Analytics**

*Akhil Arora (EPFL), Arnab Bhattacharya (IIT Kanpur), George Fletcher (TU Eindhoven)*

**DEEM 2019: Session 3**

(Sun 14:00-15:30)

**Room:** Berlage Zaal

### **DEEM 2019: Workshop on Data Management for**

## **End-to-End Machine Learning**

*Sebastian Schelter (New York University), Neoklis Polyzotis (Google), Manasi Vartak (Massachusetts Institute of Technology), Stephan Seufert (Amazon Research)*

**Phokion Kolaitis Special Event: part 1** (Sun 14:00-15:30)

**Room:** Effectenbeurszaal

## **Phokion Kolaitis Special Event**

*Georg Gottlob (University of Oxford), Wang-Chiew Tan (Megagon Labs)*

**DSMM 2019: Session 3** (Sun 14:00-15:30)

**Room:** Mendes da Costa Kamer

## **DSMM 2019: the 5th Workshop on Data Science for Macro-modeling with Financial and Economic Datasets**

*Douglas Burdick (IBM Almaden Research Center), Rajasekar Krishnamurthy (IBM T. J. Watson Research Center), Louiza Raschid (University of Maryland)*

**Sunday 06/30/2019 15:30-16:30**

**Coffee + Workshop Posters** (Sun 15:30-16:30)

**Room:** Grote Zaal

**Sunday 06/30/2019 16:30-18:00**

**Tutorial 2: part 2** (Sun 16:30-18:00)

**Room:** Veilingzaal

## **Exploring the Data Wilderness through Examples**

*Davide Mottin (Aarhus University), Matteo Lissandrini (Aalborg University), Yannis Velegrakis (Utrecht University), Themis Pal-*

**GRADES-NDA 2019: Session 4**

(Sun 16:30-18:00)

**Room:** Administratiezaal

**GRADES-NDA 2019: Joint International Workshop on Graph Data Management Experiences & Systems and Network Data Analytics**

*Akhil Arora (EPFL), Arnab Bhattacharya (IIT Kanpur), George Fletcher (TU Eindhoven)*

**DEEM 2019: Session 4**

(Sun 16:30-18:00)

**Room:** Berlage Zaal

**DEEM 2019: Workshop on Data Management for End-to-End Machine Learning**

*Sebastian Schelter (New York University), Neoklis Polyzotis (Google), Manasi Vartak (Massachusetts Institute of Technology), Stephan Seufert (Amazon Research)*

**Phokion Kolaitis Special Event: part 2 (16:00-18:15)**

(Sun 16:30-18:00)

**Room:** Effectenbeurszaal

**Phokion Kolaitis Special Event**

*Georg Gottlob (University of Oxford), Wang-Chiew Tan (Megagon Labs)*

**DSMM 2019: Session 4**

(Sun 16:30-18:00)

**Room:** Mendes da Costa Kamer

**DSMM 2019: the 5th Workshop on Data Science for Macro-modeling with Financial and Economic Datasets**

*Douglas Burdick (IBM Almaden Research Center), Rajasekar Krishnamurthy (IBM T. J. Watson Research Center), Louiza Raschid (University of Maryland)*

**Sunday 06/30/2019 18:15-19:45**

PODS Reception

(Sun 18:15-19:45)

**Room:** Grote Zaal

**Monday 01/07/2019 08:00-08:30**

Coffee + Light Breakfast

(Mon 08:00-08:30)

**Room:** Grote Zaal

**Monday 01/07/2019 08:30-10:00**

PODS Opening & Keynote

(Mon 08:30-10:00)

**Room:** Effectenbeurszaal

**Chair:** Christoph Koch

**Differential Privacy and the US Census**

*Cynthia Dwork (Harvard University)*

**Monday 01/07/2019 10:00-11:00**

PODS 1: Incomplete Information

(Mon 10:00-11:00)

**Room:** Effectenbeurszaal

**Chair:** Pierre Senellart

**Regularizing Conjunctive Features for Classification**

*Pablo Barceló (University of Chile & IMFD Chile), Alexander Baumgartner (University of Chile & RISC, Johannes Kepler University), Victor Dalmau (Universitat Pompeu Fabra), Benny Kimelfeld (Technion)*

## **Probabilistic Databases with an Infinite Open-World Assumption**

*Martin Grohe (RWTH Aachen University), Peter Lindner (RWTH Aachen University)*

### **Query Evaluation in Election Databases**

*Benny Kimelfeld (Technion), Phokion Kolaitis (University of California, Santa Cruz & IBM Research-Almaden), Muhammad Tibi (Technion)*

**DaMoN 2019: Session 1**

(Mon 10:00-11:00)

**Room:** Berlage Zaal

### **DaMoN 2019: the 15th International Workshop on Data Management on New Hardware**

*Thomas Neumann (Technische Universität München), Ken Salem (University of Waterloo)*

**Monday 01/07/2019 11:00-11:30**

**Coffee**

(Mon 11:00-11:30)

**Room:** Grote Zaal

**Monday 01/07/2019 11:30-13:00**

**Gems of PODS and Test-of-Time Award** (Mon 11:30-13:00)

**Room:** Effectenbeurszaal

**Chair:** Benny Kimelfeld

### **A General Datalog-based Framework for Tractable Query Answering**

*Andrea Calì (Birkbeck College), Georg Gottlob (University of Oxford), Thomas Lukasiewicz (University of Oxford)*

### **Database Repairs and Consistent Query Answering:**

## **Origins and Further Developments**

*Leopoldo Bertossi (RelationalAI & Carleton University)*

## **Remembering the Probabilistic Analysis of Latent Semantic Indexing**

*Christos Papadimitriou (Columbia University)*

**DaMoN 2019: Session 2**

(Mon 11:30-13:00)

**Room:** Berlage Zaal

## **DaMoN 2019: the 15th International Workshop on Data Management on New Hardware**

*Thomas Neumann (Technische Universität München), Ken Salem (University of Waterloo)*

**Monday 01/07/2019 13:00-14:30**

Lunch + Posters (PODS, DaMoN, SIGMOD Student Research Competition)

(Mon 13:00-14:30)

**Room:** Grote Zaal

**Monday 01/07/2019 14:30-16:30**

**PODS 2: Enumeration and Counting**

(Mon 14:30-16:30)

**Room:** Effectenbeurszaal

**Chair:** Dirk van Gucht

## **Efficient Logspace Classes for Enumeration, Counting, and Uniform Generation**

*Marcelo Arenas (PUC & IMFD Chile), Luis Alberto Croquevielle (PUC & IMFD Chile), Rajesh Jayaram (Carnegie Mellon University), Cristian Riveros (PUC & IMFD Chile)*

## **Ranked Enumeration of Minimal Triangulations**

*Noam Ravid (Technion), Dori Medini (Technion), Benny*

*Kimelfeld (Technion)*

## **Enumeration on Trees with Tractable Combined Complexity and Efficient Updates**

*Antoine Amarilli (LTCI, CNRS, Télécom ParisTech, Université Paris-Saclay), Pierre Bourhis (CRIStAL, CNRS UMR 9189, Inria Lille), Stefan Mengel (CNRS, CRIL UMR 8188), Matthias Nieuwerth (University of Bayreuth)*

## **Counting Database Repairs under Primary Keys Revisited**

*Marco Calautti (University of Edinburgh), Marco Console (University of Edinburgh), Andreas Pieris (University of Edinburgh)*

## **The Complexity of Counting Cycles in the Adjacency List Streaming Model**

*John Kallaugh (University of Texas at Austin), Andrew McGregor (University of Massachusetts Amherst), Eric Price (University of Texas at Austin), Sofya Vorotnikova (University of Massachusetts Amherst)*

## **On the Enumeration Complexity of Unions of Conjunctive Queries**

*Nofar Carmeli (Technion), Markus Kröll (TU Wien)*

**DaMoN 2019: Session 3**

(Mon 14:30-16:30)

**Room:** Berlage Zaal

## **DaMoN 2019: the 15th International Workshop on Data Management on New Hardware**

*Thomas Neumann (Technische Universität München), Ken Salem (University of Waterloo)*

**Monday 01/07/2019 16:30-17:00**

Coffee

(Mon 16:30-17:00)

**Room:** Grote Zaal

**Monday 01/07/2019 17:00-18:00**

**PODS 3: Information Extraction, Hashing, and Privacy**

(Mon 17:00-18:00)

**Room:** Effectenbeurszaal

**Chair:** Yufei Tao

### **Split-Correctness in Information Extraction**

*Johannes Doleschal (University of Bayreuth & Hasselt University), Benny Kimelfeld (Technion), Wim Martens (University of Bayreuth), Yoav Nahshon (Technion), Frank Neven (Hasselt University & Transnational University of Limburg)*

### **Robust Set Reconciliation via Locality Sensitive Hashing**

*Michael Mitzenmacher (Harvard University), Tom Morgan (Harvard University & Google)*

### **What Storage Access Privacy is Achievable with Small Overhead?**

*Sarvar Patel (Google), Giuseppe Persiano (Google & University of Salerno), Kevin Yeo (Google)*

**DaMoN 2019: Session 4**

(Mon 17:00-18:00)

**Room:** Berlage Zaal

### **DaMoN 2019: the 15th International Workshop on Data Management on New Hardware**

*Thomas Neumann (Technische Universität München), Ken Salem (University of Waterloo)*

**Monday 01/07/2019 18:00-19:00**

PODS Business Meeting

(Mon 18:00-19:00)

**Room:** Effectenbeurszaal

**DaMoN 2019: Session 4 (uninterrupted from 17:00)**

(Mon 18:00-19:00)

**Room:** Berlage Zaal

**DaMoN 2019: the 15th International Workshop on Data Management on New Hardware**

*Thomas Neumann (Technische Universität München), Ken Salem (University of Waterloo)*

**Tuesday 02/07/2018 08:00-08:30**

Coffee + Light Breakfast

(Tue 08:00-08:30)

**Room:** Grote Zaal

**Tuesday 02/07/2018 08:30-10:20**

SIGMOD Welcome + Keynote

(Tue 08:30-10:20)

**Room:** Effectenbeurszaal

**Chair:** Peter Boncz

**Responsible Data Science**

*Lise Getoor (University of California, Santa Cruz)*

**Tuesday 02/07/2018 10:20-11:00**

Teaser Talks for all Tuesday SIGMOD Research, Industrial Papers and TODS Posters  
(Tue 10:20-11:00)

**Room:** Effectenbeurszaal

**Chair:** Peter Boncz

**Tuesday 02/07/2018 11:00-11:30**

Coffee

(Tue 11:00-11:30)

**Room:** Grote Zaal

**Tuesday 02/07/2018 11:30-12:50**

SIGMOD Panel on Data Ethics

(Tue 11:30-12:50)

**Room:** Effectenbeurszaal

**Chair:** H.V. Jagadish

### **The Responsibility Challenge for Data**

*H. V. Jagadish (University of Michigan), Francesco Bonchi (ISI Foundation), Tina Eliassi-Rad (Northeastern University), Lise Getoor (University of California, Santa Cruz), Krishna Gummadi (Max Planck Institute for Software Systems), Julia Stoyanovich (New York University)*

**PODS 4: Streams**

(Tue 11:30-12:50)

**Room:** Veilingzaal

**Chair:** Pablo Barceló

### **Tight Trade-offs for the Maximum k-Coverage Problem in the General Streaming Model**

*Piotr Indyk (MIT), Ali Vakilian (MIT)*

### **Weighted Reservoir Sampling from Distributed Streams**

*Rajesh Jayaram (Carnegie Mellon University), Gokarna Sharma (Kent State University), Srikanta Tirthapura (Iowa State University), David Woodruff (Carnegie Mellon University)*

### **Distributed and Streaming Linear Programming in**

## **Low Dimensions**

*Sepehr Assadi (Princeton University), Nikolai Karpov (Indiana University), Qin Zhang (Indiana University)*

## **Better Sliding Window Algorithms to Maximize Sub-additive and Diversity Objectives**

*Michele Borassi (Google Research), Alessandro Epasto (Google Research), Silvio Lattanzi (Google Research), Sergei Vassilvitskii (Google Research), Morteza Zadimoghaddam (Google Research)*

**Tuesday 02/07/2018 12:50-14:20**

Lunch + SIGMOD Awards

(Tue 12:50-14:20)

**Room:** Grote Zaal

**Tuesday 02/07/2018 14:20-16:20**

SIGMOD Research 1: Query Processing & Optimization  
1 - sponsored by Tableau

(Tue 14:20-16:20)

**Room:** Effectenbeurszaal

**Chair:** Wolfgang Lehner

## **Exact Cardinality Query Optimization with Bounded Execution Cost**

*Immanuel Trummer (Cornell University)*

## **Pessimistic Cardinality Estimation**

*Walter Cai (University of Washington), Magdalena Balazinska (University of Washington), Dan Suciu (University of Washington)*

## **Efficiently Searching In-Memory Sorted Arrays: Revenge of the Interpolation Search?**

*Peter Van Sandt (University of Wisconsin, Madison), Yannis Chronis (University of Wisconsin, Madison), Jignesh Patel (Uni-*

*versity of Wisconsin, Madison)*

## **Iterative Query Processing based on Unified Optimization Techniques**

*Kisung Park (Kyung Hee University), Hojin Seo (Kyung Hee University), Mostofa Rasel (Kyung Hee University), Young-Koo Lee (Kyung Hee University), Chanho Jeong (SAP Labs Korea), Sung Yeol Lee (SAP Labs Korea), Chungmin Lee (SAP Labs Korea), Dong-Hun Lee (SAP Labs Korea)*

## **Approximate Distinct Counts for Billions of Datasets**

*Daniel Ting (Tableau Software)*

## **Cache-oblivious High-performance Similarity Join**

*Martin Perdacher (University of Vienna), Claudia Plant (University of Vienna), Christian Böhm (Ludwig-Maximilians-Universität)*

SIGMOD Research 2: Privacy/Blockchain

(Tue

14:20-16:20)

**Room:** Graanbeurszaal

**Chair:** Raghav Kaushik

## **Blurring the Lines between Blockchains and Database Systems: the Case of Hyperledger Fabric**

*Ankur Sharma (Saarland University), Felix Schuhknecht (Saarland University), Divya Agrawal (Saarland University), Jens Dittrich (Saarland University)*

## **Towards Scaling Blockchain Systems via Sharding**

*Hung Dang (National University of Singapore), Tien Tuan Anh Dinh (National University of Singapore), Dumitrel Loghin (National University of Singapore), Ee-Chien Chang (National University of Singapore), Qian Lin (National University of Singapore), Beng Chin Ooi (National University of Singapore)*

## **vChain: Enabling Verifiable Boolean Range Queries over Blockchain Databases**

*Cheng Xu (Hong Kong Baptist University), Ce Zhang (Hong Kong Baptist University), Jianliang Xu (Hong Kong Baptist University)*

## **Answering Multi-Dimensional Analytical Queries under Local Differential Privacy**

*Tianhao Wang (Purdue University), Bolin Ding (Alibaba Group), Jingren Zhou (Alibaba Group), Cheng Hong (Alibaba Group), Zhicong Huang (Alibaba Group), Ninghui Li (Purdue University), Somesh Jha (University of Wisconsin, Madison)*

## **APEx: Accuracy-Aware Differentially Private Data Exploration**

*Chang Ge (University of Waterloo), Xi He (University of Waterloo), Ihab Ilyas (University of Waterloo), Ashwin Machanavajjhala (Duke University)*

## **Active Sparse Mobile Crowd Sensing Based on Matrix Completion**

*Kun Xie (Hunan University), Xiaocan Li (Hunan University), Xin Wang (Stony Brook University), Gaogang Xie (Institute of Computing Technology & Chinese Academy of Sciences), Jigang Wen (Institute of Computing Technology & Chinese Academy of Sciences), Dafang Zhang (Hunan University)*

SIGMOD Research 3: Information Extraction

(Tue

14:20-16:20)

**Room:** Administratiezaal

**Chair:** Guoliang Li

## **Autocompletion for Prefix-Abbreviated Input**

*Sheng Hu (Nagoya University & Kyoto University), Chuan Xiao (Nagoya University & Osaka University), Jianbin Qin (Shenzhen University), Yoshiharu Ishikawa (Nagoya University), Qiang Ma*

*(Kyoto University)*

## **Progressive Deep Web Crawling Through Keyword Queries For Data Enrichment**

*Pei Wang (Simon Fraser University), Ryan Shea (Simon Fraser University), Jiannan Wang (Simon Fraser University), Eugene Wu (Columbia University)*

## **Visual Segmentation for Information Extraction from Heterogeneous Visually Rich Documents**

*Ritesh Sarkhel (Ohio State University), Arnab Nandi (Ohio State University)*

## **RRR: Rank-Regret Representative**

*Abolfazl Asudeh (University of Michigan), Azade Nazi (Google AI), Nan Zhang (Pennsylvania State University), Gautam Das (University of Texas at Arlington), H. V. Jagadish (University of Michigan)*

## **Strongly Truthful Interactive Regret Minimization**

*Min Xie (Hong Kong University of Science and Technology), Raymond Chi-Wing Wong (Hong Kong University of Science and Technology), Ashwin Lall (Denison University)*

## **Verifying Text Summaries of Relational Data Sets**

*Saehan Jo (Cornell University), Immanuel Trummer (Cornell University), Weicheng Yu (Cornell University), Xuezhi Wang (Google Research), Cong Yu (Google Research), Daniel Liu (Cornell University), Niyati Mehta (Cornell University)*

**SIGMOD Industry 1: Data Applications (Tue 14:20-16:20)**

**Room:** Berlage Zaal

**Chair:** Marco Serafini

**QuickInsights: Quick and Automatic Discovery of Insights from Multi-Dimensional Data**

*Rui Ding (Microsoft Research), Shi Han (Microsoft Research),  
Yong Xu (Microsoft Research), Haidong Zhang (Microsoft Research), Dongmei Zhang (Microsoft Research)*

## **ExplainIt! – A Declarative Root-cause Analysis Engine for Time Series Data**

*Vimalkumar Jeyakumar (Cisco Tetration Analytics), Omid Madani (Cisco Tetration Analytics), Ali Parandeh (Cisco Tetration Analytics), Ashutosh Kulshreshtha (Cisco Tetration Analytics), Weifei Zeng (Cisco Tetration Analytics), Navindra Yadav (Cisco Tetration Analytics)*

## **Automatically Generating Interesting Facts from Wikipedia Tables**

*Flip Korn (Google Research), Xuezhi Wang (Google Research), You Wu (Google Research), Cong Yu (Google Research)*

## **Snorkel DryBell: A Case Study in Deploying Weak Supervision at Industrial Scale**

*Stephen Bach (Brown University), Daniel Rodriguez (Google), Yintao Liu (Google), Chong Luo (Google), Haidong Shao (Google), Cassandra Xia (Google), Souvik Sen (Google), Alex Ratner (Stanford University), Braden Hancock (Stanford University), Houman Alborzi (Google), Rahul Kuchhal (Google), Chris Ré (Stanford University), Rob Malkin (Google)*

## **PS2: Parameter Server on Spark**

*Zhipeng Zhang (Peking University & Tencent Inc.), Bin Cui (Peking University), Yingxia Shao (Beijing University of Posts and Telecommunications), Lele Yu (Tencent Inc.), Jiawei Jiang (Tencent Inc.), Xupeng Miao (Peking University & Tencent Inc.)*

## **Entity Matching Meets Data Science: A Progress Report from the Magellan Project**

*Yash Govind (University of Wisconsin, Madison), Pradap Konda*

(University of Wisconsin, Madison), Paul Suganthan G.C. (Google), Philip Martinkus (University of Wisconsin, Madison), Palaniappan Nagarajan (University of Wisconsin, Madison), Aravind Soundararajan (University of Wisconsin, Madison), Han Li (University of Wisconsin, Madison), Sidharth Mudgal (University of Wisconsin, Madison), Jeff Ballard (University of Wisconsin, Madison), Haojun Zhang (University of Wisconsin, Madison), Adel Ardalan (University of Wisconsin, Madison), Sanjib Das (University of Wisconsin, Madison), Derek Paulsen (University of Wisconsin, Madison), Amanpreet Singh Saini (University of Wisconsin, Madison), Erik Paulson (University of Wisconsin, Madison), Youngchoon Park (Johnson Controls), Marshall Carter (American Family Insurance), Mingju Sun (American Family Insurance), Glenn Fung (American Family Insurance), AnHai Doan (University of Wisconsin, Madison)

## PODS Invited Tutorial 1

(Tue 14:20-16:20)

**Room:** Veilingzaal

**Chair:** Pierre Bourhis

### Making Consistency Protocols Serializable

Alan Fekete (University of Sydney)

**Tuesday 02/07/2018 16:20-17:50**

## Poster & Demo Groups A and B

(Tue 16:20-17:50)

**Room:** Grote Zaal

One poster for each SIGMOD and PODS paper presented on Tuesday. Plus 5 Programming Contest demos.

## Representations and Optimizations for Embedded

## **Parallel Dataflow Languages**

*Alexander Alexandrov (TU Berlin), Georgi Krastev (TU Berlin), Volker Markl (TU Berlin)*

## **A Survey of Spatial Crowdsourcing**

*Srinivasa Raghavendra (Aalborg University), Bhuvan Gummidi (Aalborg University), Xike Xie (University of Science and Technology of China), Torben Bach Pedersen (Aalborg University)*

## **K-Regret Queries Using Multiplicative Utility Functions**

*Jianzhong Qi (The University of Melbourne), Fei Zuo (The University of Melbourne), Hanan Samet (University of Maryland), Jia Cheng Yao (The University of Melbourne)*

## **Historic Moments Discovery in Sequence Data.**

*Ran Bai (The Hong Kong Polytechnic University), Wing-Kai Hon (National Tsing Hua University, Taiwan), Eric Lo (Chinese University of Hong Kong), Zhian He (University of Hong Kong), Kenny Q. Zhu (Shanghai Jiao Tong University)*

## **FindYourFavorite: An Interactive System for Finding the User’s Favorite Tuple in the Database**

*Min Xie (Hong Kong University of Science and Technology), Tianwen Chen (Hong Kong University of Science and Technology), Raymond Chi-Wing Wong (Hong Kong University of Science and Technology)*

## **Large Scale Graph Mining with G-Miner**

*Hongzhi Chen (The Chinese University of Hong Kong), Xiaoxi Wang (The Chinese University of Hong Kong), Chenghuan Huang (The Chinese University of Hong Kong), Juncheng Fang (The Chinese University of Hong Kong), Yifan Hou (The Chinese University of Hong Kong), Changji Li (The Chinese University of Hong Kong), James Cheng (The Chinese University of Hong Kong)*

*Kong)*

## **ANMAT: Automatic Knowledge Discovery and Error Detection through Pattern Functional Dependencies**

*Abdulhakim Qahtan (QCRI, HBKU), Nan Tang (QCRI, HBKU), Mourad Ouzzani (QCRI, HBKU), Yang Cao (University of Edinburgh), Michael Stonebraker (MIT)*

## **Estimating Cardinalities with Deep Sketches**

*Andreas Kipf (Technische Universität München), Dimitri Vorona (Technische Universität München), Jonas Müller (Technische Universität München), Thomas Kipf (University of Amsterdam), Bernhard Radke (Technische Universität München), Viktor Leis (Technische Universität München), Peter Boncz (CWI), Thomas Neumann (Technische Universität München), Alfons Kemper (Technische Universität München)*

## **Unit Testing Data with Deequ**

*Sebastian Schelter (Amazon Research), Felix Biessmann (Amazon Research), Dustin Lange (Amazon Research), Tammo Rukat (Amazon Research), Philipp Schmidt (Amazon Research), Stephan Seufert (Amazon Research), Pierre Brunelle (Amazon Research), Andrey Taptunov (Amazon Research)*

## **DuckDB: an Embeddable Analytical Database**

*Mark Raasveldt (CWI), Hannes Mühlisen (CWI)*

## **CLASH: A High-Level Abstraction for Optimized, Multi-Way Stream Joins over Apache Storm**

*Manuel Dossinger (TU Kaiserslautern), Sebastian Michel (TU Kaiserslautern), Constantin Roudsarabi (TU Kaiserslautern)*

## **PgCuckoo: Laying Plan Eggs in PostgreSQL's Nest**

*Denis Hirn (Universität Tübingen), Torsten Grust (Universität Tübingen)*

## **Demonstration of ModelarDB: Model-Based Management of Dimensional Time Series**

*Søren Kejser Jensen (Aalborg University), Torben Bach Pedersen (Aalborg University), Christian Thomsen (Aalborg University)*

## **NEURON: Query Execution Plan Meets Natural Language Processing For Augmenting DB Education**

*Siyuan Liu (Nanyang Technological University), Sourav Bhowmick (Nanyang Technological University), Wanlu Zhang (Nanyang Technological University), Shu Wang (Nanyang Technological University), Wanyi Huang (Nanyang Technological University), Shafiq Joty (Nanyang Technological University)*

## **PIClean: A Probabilistic and Interactive Data Cleaning System**

*Zhuoran Yu (Georgia Institute of Technology), Xu Chu (Georgia Institute of Technology)*

## **Apollo: A Dataset Profiling and Operator Modeling System**

*Tasos Bakogiannis (National Technical University of Athens), Ioannis Giannakopoulos (National Technical University of Athens), Dimitrios Tsoumakos (Ionian University), Nectarios Koziris (National Technical University of Athens)*

## **Pivotal Greenplum© for Kubernetes: Demonstration of Managing Greenplum Database on Kubernetes**

*Jemish Patel (Pivotal Software Inc), Goutam Tadi (Pivotal Software Inc), Oz Basarir (Pivotal Software Inc), Lawrence Hamel (Pivotal Software Inc), David Sharp (Pivotal Software Inc), Fei Yang (Pivotal Software Inc), Xin Zhang (Pivotal Software Inc)*

## **Demonstration of SpeakQL: Speech-driven Multi-modal Querying of Structured Data**

*Vraj Shah (University of California, San Diego), Side Li (Univer-*

*sity of California, San Diego), Kevin Yang (University of California, San Diego), Arun Kumar (University of California, San Diego), Lawrence Saul (University of California, San Diego)*

### **Ratel: Interactive Analytics for Large Scale Trajectories**

*Haoda Li (Tsinghua University), Guoliang Li (Tsinghua University), Jiayang Liu (Tsinghua University), Haitao Yuan (Tsinghua University), Haiquan Wang (Tsinghua University)*

### **MigCast: Putting a Price Tag on Data Model Evolution in NoSQL Data Stores**

*Andrea Hillenbrand (Darmstadt University of Applied Sciences), Maksym Levchenko (Darmstadt University of Applied Sciences), Uta Störl (Darmstadt University of Applied Sciences), Stefanie Scherzinger (OTH Regensburg), Meike Klettke (University of Rostock)*

### **NeMeSys - A Showcase of Data Oriented Near Memory Graph Processing**

*Alexander Krause (Technische Universität Dresden), Thomas Kissinger (Technische Universität Dresden), Dirk Habich (Technische Universität Dresden), Wolfgang Lehner (Technische Universität Dresden)*

### **Low-latency Spark Queries on Updatable Data**

*Alexandru Uta (Vrije Universiteit Amsterdam), Bogdan Ghit (Databricks), Ankur Dave (University of California, Berkeley), Peter Boncz (CWI)*

### **Demonstration of Nimbus: Model-based Pricing for Machine Learning in a Data Marketplace**

*Lingjiao Chen (University of Wisconsin, Madison), Hongyi Wang (University of Wisconsin, Madison), Leshang Chen (University of Pennsylvania), Paraschos Koutris (University of Wisconsin,*

*Madison), Arun Kumar (University of California, San Diego)*

## **Capturing and Querying Structural Provenance in Spark with Pebble**

*Ralf Diestelkämper (Universität Stuttgart), Melanie Herschel (Universität Stuttgart)*

## **SVQ: Streaming Video Queries**

*Ioannis Xarchakos (University of Toronto), Nick Koudas (University of Toronto)*

## **GraphWrangler: An Interactive Graph View on Relational Data**

*Nafisa Anzum (University of Waterloo), Semih Salihoglu (University of Waterloo), Daniel Vogel (University of Waterloo)*

## **Coconut Palm: Static and Streaming Data Series Exploration Now in your Palm**

*Haridimos Kondylakis (FORTH-ICS), Niv Dayan (Harvard University), Kostas Zoumpatianos (Harvard University), Themis Palpanas (Paris Descartes University)*

## **Natural Language Querying of Complex Business Intelligence Queries**

*Jaydeep Sen (IBM Research AI), Fatma Ozcan (IBM Research AI), Abdul Quamar (IBM Research AI), Greg Stager (IBM Canada), Ashish Mittal (IBM Research AI), Manasa Jammi (IBM Research AI), Chuan Lei (IBM Research AI), Diptikalyan Saha (IBM Research AI), Karthik Sankaranarayanan (IBM Research AI)*

**Tuesday 02/07/2018 17:50-19:50**

New Researcher Symposium

(Tue 17:50-19:50)

**Room:** Effectenbeurszaal

**Room:** Graanbeurszaal

**SpeakQL: Towards Speech-driven Multimodal Querying**

*Vraj Shah (University of California, San Diego)*

**Fingerprints for Compressed Columnar Data Search**

*Carmen Kwan (University of Waterloo)*

**CAvSAT: A System for Query Answering over Inconsistent Databases**

*Akhil Dixit (University of California, Santa Cruz)*

**Scalable Reservoir Sampling on Many-Core CPUs**

*Altan Birler (Technische Universität München)*

**LSM-Trees and B-Trees: The Best of Both Worlds**

*Varun Jain (Harvard University), James Lennon (Harvard University), Harshita Gupta (Harvard University)*

**Generating Selective Filters for Access Method and PhysicalDesign Evaluation**

*Pranav Subramaniam (University of Chicago)*

**Interactive Visualization For Big Spatial Data**

*Saheli Ghosh (University of California, Riverside)*

**Learning to Generate Questions with Adaptive Copying Neural Networks**

*Xinyuan Lu (Carleton University)*

**Query-Driven Learning for Next Generation Predictive Modeling & Analytics**

*Fotis Savva (University of Glasgow)*

**Answering Range Queries Under Local Differential**

## **Privacy**

*Tejas Kulkarni (University of Warwick)*

## **Helios: An Adaptive and Query Workload-driven Partitioning Framework for Distributed Graph Stores**

*Ali Davoudian (Carleton University)*

## **Deep Query Optimization**

*Tin Vu (University of California, Riverside)*

## **Bootstrapping an End-to-End Natural Language Interface for Databases**

*Nathaniel Weir (Brown University), Prasetya Utama (TU Darmstadt)*

## **Recommending Deployment Strategies in Crowd-sourcing Platforms**

*Dong Wei (New Jersey Institute of Technology)*

## **Towards Understanding Data Analysis Workflows using a Large Notebook Corpus**

*Mohammed Suhail Rehman (University of Chicago)*

## **Arachnid: Generalized Visual Data Cleaning**

*Conder Shou (Columbia University), Amita Shukla (Columbia University)*

**PODS 5: Semistructured Data and Knowledge Graphs,  
Logic, and Verification**

(Tue 17:50-19:50)

**Room:** Veilingzaal

**Chair:** Reinhard Pichler

## **The Space-Efficient Core of Vadalog**

*Gerald Berger (TU Wien), Georg Gottlob (University of Oxford & TU Wien), Andreas Pieris (University of Edinburgh),*

*Emanuel Sallinger (University of Oxford &amp; TU Wien)*

## **Decidable XPath Fragments in the Real World**

*David Baelde (ENS Paris-Saclay & CNRS, Université Paris-Saclay), Anthony Lick (ENS Paris-Saclay & CNRS, Université Paris-Saclay), Sylvain Schmitz (ENS Paris-Saclay & CNRS, Université Paris-Saclay)*

## **Containment of Shape Expression Schemas for RDF**

*Slawek Staworko (CNRS & University of Lille), Piotr Wieczorek (University of Wroclaw)*

## **Complexity Bounds for Relational Algebra over Document Spanners**

*Liat Peterfreund (Technion), Dominik Freydenberger (Loughborough University), Benny Kimelfeld (Technion), Markus Kröll (Vienna University of Technology)*

## **Reachability in Database-driven Systems with Numerical Attributes under Recency Bounding**

*Parosh Aziz Abdulla (Uppsala University), C. Aiswarya (Chennai Mathematical Institute), Mohamed Faouzi Atig (Uppsala University), Marco Montali (KRDB Research Centre, Free University of Bozen-Bolzano)*

## **Compiling Existential Positive Queries to Bounded-Variable Fragments**

*Christoph Berkholz (Humboldt-Universität zu Berlin), Hubie Chen (Birkbeck, University of London)*

**Tuesday 02/07/2018 20:30-23:00**

SIGMOD Reception - sponsored by MonetDB (Tue  
20:30-23:00)

**Room:** Van Gogh Museum

**Wednesday 03/07/2018 08:00-08:30**

Coffee + Light Breakfast

(Wed 08:00-08:30)

**Room:** Grote Zaal

**Wednesday 03/07/2018 08:30-10:00**

SIGMOD Keynote

(Wed 08:30-10:00)

**Room:** Effectenbeurszaal

**Chair:** Stefan Manegold

**State of Public and Private Blockchains: Myths and Reality**

*C. Mohan (IBM Almaden Research Center)*

**Wednesday 03/07/2018 10:00-11:00**

Teaser Talks for all Wednesday SIGMOD Research and Industrial Papers

(Wed 10:00-11:00)

**Room:** Effectenbeurszaal

**Chair:** Stefan Manegold

PODS 6: Containment and Homomorphisms

(Wed

10:00-11:00)

**Room:** Veilingzaal

**Chair:** Dan Olteanu

**Testability of Homomorphism Inadmissibility: Property Testing Meets Database Theory**

*Hubie Chen (Birkbeck, University of London), Yuichi Yoshida (National Institute of Informatics)*

**The Selfish Models Property: Bounding the Complexity of Query Containment and Entailment Prob-**

lems

*Hubie Chen (Birkbeck, University of London)*

## **Attacking Diophantus: Solving a Special Case of Bag Containment**

*George Konstantinidis (University of Southampton), Fabio Mogavero (Università degli Studi di Napoli Federico II)*

**Wednesday 03/07/2018 11:00-11:30**

Coffee

(Wed 11:00-11:30)

**Room:** Grote Zaal

**Wednesday 03/07/2018 11:30-12:50**

SIGMOD Research 4: Distributed Data Management

(Wed 11:30-12:50)

**Room:** Effectenbeurszaal

**Chair:** Holger Pirk

## **An End-to-End Automatic Cloud Database Tuning System Using Deep Reinforcement Learning**

*Ji Zhang (Huazhong University of Science and Technology), Yu Liu (Huazhong University of Science and Technology), Ke Zhou (Huazhong University of Science and Technology), Guoliang Li (Tsinghua University), Zhili Xiao (Tencent Inc.), Bin Cheng (Tencent Inc.), Jiashu Xing (Tencent Inc.), Yangtao Wang (Huazhong University of Science and Technology), Tianheng Cheng (Huazhong University of Science and Technology), Li Liu (Huazhong University of Science and Technology), Minwei Ran (Huazhong University of Science and Technology), Zekang Li (Huazhong University of Science and Technology)*

## **Fast General Distributed Transactions with Opacity**

*Alex Shamis (Microsoft Research), Matthew Renzelmann (Mi-*

*crosoft), Stanko Novakovic (VMware), Georgios Chatzopoulos (EPFL), Aleksandar Dragojević (Microsoft Research), Dushyanth Narayanan (Microsoft Research), Miguel Castro (Microsoft Research)*

## **The Log-Structured Merge-Bush & the Wacky Continuum**

*Niv Dayan (Harvard University), Stratos Idreos (Harvard University)*

## **RaSQL: Greater Power and Performance for Big Data Analytics with Recursive-aggregate-SQL on Spark**

*Jiaqi Gu (University of California, Los Angeles), Yugo Watanabe (University of California, Los Angeles), William Mazza (University of Naples Federico II), Alexander Shkapsky (Workday, Inc.), Mohan Yang (Google), Ling Ding (University of California, Los Angeles), Carlo Zaniolo (University of California, Los Angeles)*

### **SIGMOD Research 5: Provenance**

(Wed 11:30-12:50)

**Room:** Graanbeurszaal

**Chair:** Alexandra Meliou

## **Going Beyond Provenance: Explaining Query Answers with Pattern-based Counterbalances**

*Zhengjie Miao (Duke University), Qitian Zeng (Illinois Institute of Technology), Boris Glavic (Illinois Institute of Technology), Sudeepa Roy (Duke University)*

## **Explaining Wrong Queries Using Small Examples**

*Zhengjie Miao (Duke University), Sudeepa Roy (Duke University), Jun Yang (Duke University)*

## **Ariadne: Online Provenance for Big Graph Analytics**

*Vicky Papavasileiou (University of California, San Diego), Ken Yocum (Intuit, Inc. & University of California, San Diego), Alin Deutsch (University of California, San Diego)*

## **Hypothetical Reasoning via Provenance Abstraction**

*Daniel Deutch (Tel Aviv University), Yuval Moskovitch (Tel Aviv University), Noam Rinetzky (Tel Aviv University)*

SIGMOD Research 6: Streams

(Wed 11:30-12:50)

**Room:** Administratiezaal

**Chair:** Jonathan Goldstein

### **Event Trend Aggregation Under Rich Event Matching Semantics**

*Olga Poppe (Microsoft Gray Systems Lab), Chuan Lei (IBM Almaden Research Center), Elke Rundensteiner (Worcester Polytechnic Institute), David Maier (Portland State University)*

### **Elasticutor: Rapid Elasticity for Realtime Stateful Stream Processing**

*Li Wang (Yitu Technology), Tom Z. J. Fu (Advanced Digital Sciences Center), Richard T. B. Ma (National University of Singapore), Marianne Winslett (University of Illinois Urbana-Champaign), Zhenjie Zhang (Yitu Technology)*

### **Real-Time Multi-Pattern Detection over Event Streams**

*Ilya Kolchinsky (Technion), Assaf Schuster (Technion)*

### **ASstream: Ad-hoc Shared Stream Processing**

*Jeyhun Karimov (DFKI GmbH), Tilmann Rabl (DFKI GmbH & TU Berlin), Volker Markl (DFKI GmbH & TU Berlin)*

SIGMOD Industry 2: Storage and Indexing

(Wed

11:30-12:50)

**Room:** Berlage Zaal  
**Chair:** Alexander Shraer

## **Nanosecond Indexing of Graph Data With Hash Maps and VLists**

*Andrew Carter (LinkedIn Corporation), Andrew Rodriguez (LinkedIn Corporation), Yiming Yang (LinkedIn Corporation), Scott Meyer (LinkedIn Corporation)*

## **Implementation of Cluster-wide Logical Clock and Causal Consistency in MongoDB**

*Misha Tyulenev (MongoDB, Inc), Andy Schwerin (MongoDB, Inc), Asya Kamsky (MongoDB, Inc), Randolph Tan (MongoDB, Inc), Alyson Cabral (MongoDB, Inc), Jack Mulrow (MongoDB, Inc)*

## **X-Engine: An Optimized Storage Engine for Large-scale E-commerce Transaction Processing**

*Gui Huang (Alibaba Group), Xuntao Cheng (Alibaba Group), Jianying Wang (Alibaba Group), Yujie Wang (Alibaba Group), Dengcheng He (Alibaba Group), Tieying Zhang (Alibaba Group), Feifei Li (Alibaba Group), Sheng Wang (Alibaba Group), Wei Cao (Alibaba Group), Qiang Li (Alibaba Group)*

## **Automatically Indexing Millions of Databases in Microsoft Azure SQL Database**

*Sudipto Das (Microsoft), Miroslav Grbic (Microsoft), Igor Ilic (Microsoft), Isidora Jovandic (Microsoft), Andrija Jovanovic (Microsoft), Vivek Narasayya (Microsoft), Miodrag Radulovic (Microsoft), Maja Stikic (Microsoft), Gaoxiang Xu (Microsoft), Surajit Chaudhuri (Microsoft)*

**PODS 7: Joins, hypergraphs, and Aggregate Queries**  
(Wed 11:30-12:50)

**Room:** Veilingzaal

Chair: Hubie Chen

# On Functional Aggregate Queries with Additive Inequalities

*Mahmoud Abo Khamis (RelationalAI), Ryan Curtin (RelationalAI), Benjamin Moseley (Carnegie Mellon University), Hung Ngo (RelationalAI), XuanLong Nguyen (University of Michigan), Dan Olteanu (University of Oxford), Maximilian Schleich (University of Oxford)*

# Topology Dependent Bounds For FAQs

*Michael Langberg (University at Buffalo), Shi Li (University at Buffalo), Sai Vikneshwar Mani Jayaraman (University at Buffalo), Atri Rudra (University at Buffalo)*

# Instance and Output Optimal Parallel Algorithms for Acyclic Joins

*Xiao Hu (Hong Kong University of Science and Technology), Ke Yi (Hong Kong University of Science and Technology)*

# HyperBench: A Benchmark and Tool for Hypergraphs and Empirical Findings

*Wolfgang Fischl (Vienna University of Technology), Georg Gottlob (University of Oxford), Davide Mario Longo (Vienna University of Technology), Reinhard Pichler (Vienna University of Technology)*

Wednesday 03/07/2018 12:50-14:20

Lunch + SIGMOD Business Meeting (Wed 12:50-14:20)

## Room: Grote Zaal

Wednesday 03/07/2018 14:20-16:20

SIGMOD Research 7: Modern Hardware (Wed 14:20-16:20)

**Room:** Effectenbeurszaal

**Chair:** Justin Levandoski

## **Concurrent Prefix Recovery: Performing CPR on a Database**

*Guna Prasaad (University of Washington), Badrish Chandramouli (Microsoft Research), Donald Kossmann (Microsoft Research)*

## **BriskStream: Scaling Data Stream Processing on Shared-Memory Multicore Architectures**

*Shuhao Zhang (National University of Singapore), Jiong He (Advanced Digital Sciences Center), Amelie Zhou (Shenzhen University), Bingsheng He (National University of Singapore)*

## **Border-Collie: A Wait-free, Read-optimal Algorithm for Database Logging on Multicore Hardware**

*Jongbin Kim (Hanyang University), Hyeongwon Jang (Hanyang University), Seohui Son (Hanyang University), Hyuck Han (Dongduk Women's University), Sooyong Kang (Hanyang University), Hyungsoo Jung (Hanyang University)*

## **Designing Distributed Tree-based Index Structures for Fast RDMA-capable Networks**

*Tobias Ziegler (TU Darmstadt), Sumukha Tumkur Vani (Brown University), Carsten Binnig (TU Darmstadt), Rodrigo Fonseca (Brown University), Tim Kraska (MIT)*

## **DistME: A Fast and Elastic Distributed Matrix Computation Engine using GPUs**

*Donghyoung Han (Daegu Gyeongbuk Institute of Science & Technology (DGIST)), Yoon-Min Nam (Daegu Gyeongbuk Institute of Science & Technology (DGIST)), Jihye Lee (Daegu Gyeongbuk Institute of Science & Technology (DGIST)), Kyongseok Park (Korea Institute of Science and Technology Information (KISTI)),*

*Hyunwoo Kim (Korea Institute of Science and Technology Information (KISTI)), Min-Soo Kim (Daegu Gyeongbuk Institute of Science & Technology (DGIST))*

## **GPU-based Graph Traversal on Compressed Graphs**

*Mo Sha (National University of Singapore), Yuchen Li (Singapore Management University), Kian-Lee Tan (National University of Singapore)*

**SIGMOD Research 8: Data Integration/Cleaning (Wed  
14:20-16:20)**

**Room:** Graanbeurszaal

**Chair:** Paolo Papotti

## **Interventional Fairness : Causal Database Repair for Algorithmic Fairness**

*Babak Salimi (University of Washington), Luke Rodriguez (University of Washington), Bill Howe (University of Washington), Dan Suciu (University of Washington)*

## **Uni-Detect: A Unified Approach to Automated Error Detection in Tables**

*Pei Wang (Simon Fraser University), Yeye He (Microsoft Research)*

## **HoloDetect: Few-Shot Learning for Error Detection**

*Alireza Heidari (University of Waterloo), Joshua McGrath (University of Wisconsin, Madison), Ihab Ilyas (University of Waterloo), Theodoros Rekatsinas (University of Wisconsin, Madison)*

## **JOSIE: Overlap Set Similarity Search for Finding Joinable Tables in Data Lakes**

*Erkang Zhu (University of Toronto), Dong Deng (Inception Institute of Artificial Intelligence), Fatemeh Nargesian (University of Toronto), Renée Miller (Northeastern University)*

**Raha: A Configuration-Free Error Detection System**  
*Mohammad Mahdavi (TU Berlin), Ziawasch Abedjan (TU Berlin), Raul Castro Fernandez (MIT), Samuel Madden (MIT), Mourad Ouzzani (QCRI, HBKU), Michael Stonebraker (MIT), Nan Tang (QCRI, HBKU)*

## **Speculative Distributed CSV Data Parsing for Big Data Analytics**

*Chang Ge (University of Waterloo), Yinan Li (Microsoft Research), Eric Eilebrecht (Microsoft Research), Badrish Chandramouli (Microsoft Research), Donald Kossmann (Microsoft Research)*

SIGMOD Research 9: Query Processing & Optimization  
2 (Wed 14:20-16:20)

**Room:** Administratiezaal

**Chair:** Jun Yang

**CATAPULT: Data-driven Selection of Canned Patterns for Efficient Visual Graph Query Formulation**  
*Kai Huang (Fudan University), Huey Chua (Nanyang Technological University), Sourav Bhowmick (Nanyang Technological University), Byron Choi (Hong Kong Baptist University), Shuigeng Zhou (Fudan University)*

## **iQCAR: inter-Query Contention Analyzer for Data Analytics Frameworks**

*Prajakta Kalmegh (Duke University), Shivnath Babu (Unravel Data Systems), Sudeepa Roy (Duke University)*

## **A Holistic Approach for Query Evaluation and Result Vocalization in Voice-Based OLAP**

*Immanuel Trummer (Cornell University), Yicheng Wang (Cornell University), Saketh Mahankali (Cornell University)*

## **Top-k Queries over Digital Traces**

*Yifan Li (York University), Xiaohui Yu (York University), Nick Koudas (University of Toronto)*

## **Visual Road: A Video Data Management Benchmark**

*Brandon Haynes (University of Washington), Amrita Mazumdar (University of Washington), Magdalena Balazinska (University of Washington), Luis Ceze (University of Washington), Alvin Cheung (University of Washington)*

## **Mining Precision Interfaces From Query Logs**

*Qianrui Zhang (Tsinghua University), Haoci Zhang (Columbia University), Thibault Sellam (Columbia University), Eugene Wu (Columbia University)*

**SIGMOD Research 10: Graphs 1**

(Wed 14:20-16:20)

**Room:** Berlage Zaal

**Chair:** Angela Bonifati

## **Distance-generalized Core Decomposition**

*Francesco Bonchi (ISI Foundation & Eurecat), Arijit Khan (Nanyang Technological University), Lorenzo Severini (ISI Foundation)*

## **Unboundedness and Efficiency of Truss Maintenance in Evolving Graphs**

*Yikai Zhang (Chinese University of Hong Kong), Jeffrey Yu (Chinese University of Hong Kong)*

## **PRSim: Sublinear Time SimRank Computation on Large Power-Law Graphs**

*Zhewei Wei (Renmin University of China), Xiaodong He (4Paradigm Inc.), Xiaokui Xiao (National University of Singapore), Sibo Wang (The Chinese University of Hong Kong), Yu*

*Liu (Peking University), Xiaoyong Du (Renmin University of China), Ji-Rong Wen (Renmin University of China)*

### **Scaling Distance Labeling on Small-World Networks**

*Wentao Li (University of Technology Sydney), Miao Qiao (University of Auckland), Lu Qin (University of Technology Sydney), Ying Zhang (University of Technology Sydney), Lijun Chang (University of Sydney), Xuemin Lin (University of New South Wales)*

### **Maximizing Welfare in Social Networks under A Utility Driven Influence Diffusion model**

*Prithu Banerjee (University of British Columbia), Wei Chen (Microsoft Research), Laks Lakshmanan (University of British Columbia)*

### **Efficient Approximation Algorithms for Adaptive Seed Minimization**

*Jing Tang (National University of Singapore), Keke Huang (Nanyang Technological University), Xiaokui Xiao (National University of Singapore), Laks Lakshmanan (University of British Columbia), Xueyan Tang (Nanyang Technological University), Aixin Sun (Nanyang Technological University), Andrew Lim (National University of Singapore)*

PODS Invited Tutorial 2

(Wed 14:20-16:20)

**Room:** Veilingzaal

**Chair:** Christoph Koch

### **Algorithmic Fairness: Measures, Methods and Representations**

*Suresh Venkatasubramanian (University of Utah)*

**Wednesday 03/07/2018 16:20-17:50**

Poster & Demo Groups B and C

(Wed 16:20-17:50)

**Room:** Grote Zaal

**One poster for each SIGMOD and PODS paper presented on Wednesday.**

**Pivotal Greenplum© for Kubernetes: Demonstration of Managing Greenplum Database on Kubernetes**

*Jemish Patel (Pivotal Software Inc), Goutam Tadi (Pivotal Software Inc), Oz Basarir (Pivotal Software Inc), Lawrence Hamel (Pivotal Software Inc), David Sharp (Pivotal Software Inc), Fei Yang (Pivotal Software Inc), Xin Zhang (Pivotal Software Inc)*

**Demonstration of SpeakQL: Speech-driven Multi-modal Querying of Structured Data**

*Vraj Shah (University of California, San Diego), Side Li (University of California, San Diego), Kevin Yang (University of California, San Diego), Arun Kumar (University of California, San Diego), Lawrence Saul (University of California, San Diego)*

**Ratel: Interactive Analytics for Large Scale Trajectories**

*Haoda Li (Tsinghua University), Guoliang Li (Tsinghua University), Jiayang Liu (Tsinghua University), Haitao Yuan (Tsinghua University), Haiquan Wang (Tsinghua University)*

**MigCast: Putting a Price Tag on Data Model Evolution in NoSQL Data Stores**

*Andrea Hillenbrand (Darmstadt University of Applied Sciences), Maksym Levchenko (Darmstadt University of Applied Sciences), Uta Störl (Darmstadt University of Applied Sciences), Stefanie Scherzinger (OTH Regensburg), Meike Klettke (University of Rostock)*

**NeMeSys - A Showcase of Data Oriented Near Memory Graph Processing**

*Alexander Krause (Technische Universität Dresden), Thomas Kissinger (Technische Universität Dresden), Dirk Habich (Technische Universität Dresden), Wolfgang Lehner (Technische Universität Dresden)*

## **Low-latency Spark Queries on Updatable Data**

*Alexandru Uta (Vrije Universiteit Amsterdam), Bogdan Ghit (Databricks), Ankur Dave (University of California, Berkeley), Peter Boncz (CWI)*

## **Demonstration of Nimbus: Model-based Pricing for Machine Learning in a Data Marketplace**

*Lingjiao Chen (University of Wisconsin, Madison), Hongyi Wang (University of Wisconsin, Madison), Leshang Chen (University of Pennsylvania), Paraschos Koutris (University of Wisconsin, Madison), Arun Kumar (University of California, San Diego)*

## **Capturing and Querying Structural Provenance in Spark with Pebble**

*Ralf Diestelkämper (Universität Stuttgart), Melanie Herschel (Universität Stuttgart)*

## **SVQ: Streaming Video Queries**

*Ioannis Xarchakos (University of Toronto), Nick Koudas (University of Toronto)*

## **GraphWrangler: An Interactive Graph View on Relational Data**

*Nafisa Anzum (University of Waterloo), Semih Salihoglu (University of Waterloo), Daniel Vogel (University of Waterloo)*

## **Coconut Palm: Static and Streaming Data Series Exploration Now in your Palm**

*Haridimos Kondylakis (FORTH-ICS), Niv Dayan (Harvard University), Kostas Zoumpatianos (Harvard University), Themis Palpanas (Paris Descartes University)*

## **Natural Language Querying of Complex Business Intelligence Queries**

*Jaydeep Sen (IBM Research AI), Fatma Ozcan (IBM Research AI), Abdul Quamar (IBM Research AI), Greg Stager (IBM Canada), Ashish Mittal (IBM Research AI), Manasa Jammi (IBM Research AI), Chuan Lei (IBM Research AI), Diptikalyan Saha (IBM Research AI), Karthik Sankaranarayanan (IBM Research AI)*

## **Peering through the Dark: An Owl’s View of Inter-job Dependencies and Jobs’ Impact in Shared Clusters**

*Andrew Chung (Carnegie Mellon University), Carlo Curino (Microsoft), Subru Krishnan (Microsoft), Konstantinos Karanasos (Microsoft), Panagiotis Garefalakis (Imperial College London), Gregory Ganger (Carnegie Mellon University)*

## **Visual Exploration of Time Series Anomalies with Metro-Viz**

*Philipp Eichmann (Brown University), Franco Solleza (Brown University), Nesime Tatbul (Intel Labs and MIT), Stan Zdonik (Brown University)*

## **Data Debugging and Exploration with Vizier**

*Mike Brachmann (University at Buffalo), Carlos Bautista (New York University), Sonia Castelo (New York University), Su Feng (Illinois Institute of Technology), Juliana Freire (New York University), Boris Glavic (Illinois Institute of Technology), Oliver Kennedy (University of Buffalo), Heiko Meller (New York University), Rumi Rampin (New York University), William Spoth (University at Buffalo), Ying Yang (Oracle)*

## **CrowdGame: A Game-Based Crowdsourcing System for Cost-Effective Data Labeling**

*Tongyu Liu (Renmin University of China), Jingru Yang (Ren-*

*min University of China), Ju Fan (Renmin University of China),  
Zhewei Wei (Renmin University of China), Guoliang Li (Tsinghua University), Xiaoyong Du (Renmin University of China)*

## **Ursprung: Provenance for Large-Scale Analytics Environments**

*Lukas Rupprecht (IBM Almaden Research Center), James Davis (Virginia Tech & IBM Systems), Constantine Arnold (IBM Almaden Research Center), Alexander Lubbock (Vanderbilt University), Darren Tyson (Vanderbilt University), Deepavali Bhagwat (IBM Almaden Research Center)*

## **BlockchainDB - Towards a Shared Database on Blockchains**

*Muhammad El-Hindi (TU Darmstadt), Martin Heyden (TU Darmstadt), Carsten Binnig (TU Darmstadt), Ravi Ramamurthy (Microsoft Research), Arvind Arasu (Microsoft Research), Donald Kossmann (Microsoft Research)*

## **Fluid: A Blockchain based Framework for Crowd-sourcing**

*Siyuan Han (Hong Kong University of Science and Technology), Zihuan Xu (Hong Kong University of Science and Technology), Yuxiang Zeng (Hong Kong University of Science and Technology), Lei Chen (Hong Kong University of Science and Technology)*

## **MorphStore - In-Memory Query Processing based on Morphing Compressed Intermediates LIVE**

*Dirk Habich (Technische Universität Dresden), Patrick Damme (Technische Universität Dresden), Annett Ungethüm (Technische Universität Dresden), Johannes Pietrzyk (Technische Universität Dresden), Alexander Krause (Technische Universität Dresden), Juliana Hildebrandt (Technische Universität Dresden), Wolfgang Lehner (Technische Universität Dresden)*

## **MapRepair: Mapping and Repairing under Policy Views**

*Angela Bonifati (Lyon 1 University & Camp; Liris CNRS), Ugo Comignani (Lyon 1 University & Camp; Liris CNRS), Efthymia Tsamoura (University of Oxford)*

## **RATest: Explaining Wrong Relational Queries Using Small Examples**

*Zhengjie Miao (Duke University), Sudeepa Roy (Duke University), Jun Yang (Duke University)*

## **NAVIGATE: Explainable Visual Graph Exploration by Examples**

*Mohammad Hossein Namaki (Washington State University), Qi Song (Washington State University), Yinghui Wu (Washington State University)*

## **C2Metadata: Automating the Capture of Data Transformations from Statistical Scripts in Data Documentation**

*Jie Song (University of Michigan), George Alter (University of Michigan), H. V. Jagadish (University of Michigan)*

## **MithraRanking: A System for Responsible Ranking Design**

*Yifan Guan (University of Michigan), Abolfazl Asudeh (University of Michigan), Pranav Mayuram (University of Michigan), H. V. Jagadish (University of Michigan), Julia Stoyanovich (New York University), Gerome Miklau (University of Massachusetts Amherst), Gautam Das (University of Texas at Arlington)*

## **NEWS: News Event Walker and Summarizer**

*Radityo Eko Prasojo (Free University of Bozen-Bolzano), Mouna Kacimi (Free University of Bozen-Bolzano), Werner Nutt (Free University of Bozen-Bolzano)*

## **Cost-Effective, Workload-Adaptive Migration of Big Data Applications to the Cloud**

*Victor Giannakouris (Unravel Data Systems), Alejandro Fernandez (Unravel Data Systems), Alkis Simitsis (Unravel Data Systems), Shivnath Babu (Unravel Data Systems)*

## **ChronosDB in Action: Manage, Process, and Visualize Big Geospatial Arrays in the Cloud**

*Ramon Antonio Rodriges Zalipynis (National Research University Higher School of Economics)*

**Wednesday 03/07/2018 17:50-20:30**

Dinner Transfer incl. Canal Cruise

(Wed 17:50-20:30)

**Room:** 20 boats (Rederij Stromma)

**Wednesday 03/07/2018 20:30-23:00**

SIGMOD Dinner - sponsored by Facebook  
(Wed 20:30-23:00)

**Room:** Noorderlicht Cafe

**Thursday 04/07/2018 08:00-08:30**

Coffee + Light Breakfast

(Thu 08:00-08:30)

**Room:** Grote Zaal

**Thursday 04/07/2018 08:30-10:00**

SIGMOD Award Talks

(Thu 08:30-10:00)

**Room:** Effectenbeurszaal

**Chair:** Anastasia Ailamaki

## **Data Management on Non-Volatile Memory**

*Joy Arulraj (Georgia Institute of Technology)*

# Formal Approaches to Querying Big Data in Shared-Nothing Systems

*Bas Ketsman (EPFL)*

**Thursday 04/07/2018 10:00-11:00**

Teaser Talks for all Thursday SIGMOD Research and Industrial Papers

(Thu 10:00-11:00)

**Room:** Effectenbeurszaal

**Chair:** Anastasia Ailamaki

**Thursday 04/07/2018 11:00-11:30**

Coffee

(Thu 11:00-11:30)

**Room:** Grote Zaal

**Thursday 04/07/2018 11:30-12:50**

SIGMOD Research 11: Systems & Machine Learning

(Thu 11:30-12:50)

**Room:** Effectenbeurszaal

**Chair:** Matthias Boehm

## **DeepBase: Deep Inspection of Neural Networks**

*Thibault Sellam (Columbia University), Kevin Lin (Columbia University), Ian Huang (Columbia University), Michelle Yang (University of California, Berkeley), Carl Vondrick (Columbia University), Eugene Wu (Columbia University)*

## **BlinkML: Efficient Maximum Likelihood Estimation with Probabilistic Guarantees**

*Yongjoo Park (University of Michigan), Jingyi Qing (University of Michigan), Xiaoyang Shen (University of Michigan), Barzan Mozafari (University of Michigan)*

## **SkinnerDB: Regret-Bounded Query Evaluation via Reinforcement Learning**

*Immanuel Trummer (Cornell University), Junxiong Wang (Cornell University), Deepak Maram (Cornell University), Samuel Moseley (Cornell University), Saehan Jo (Cornell University), Joseph Antonakakis (Cornell University)*

## **Democratizing Data Science through Interactive Curation of ML Pipelines**

*Zeyuan Shang (MIT), Emanuel Zgraggen (MIT), Benedetto Buratti (Brown University), Ferdinand Kossmann (MIT), Philipp Eichmann (Brown University), Yeounoh Chung (Brown University), Carsten Binnig (Brown University & TU Darmstadt), Eli Upfal (Brown University), Tim Kraska (MIT)*

SIGMOD Research 12: Indexing

(Thu 11:30-12:50)

**Room:** Graanbeurszaal

**Chair:** Stratos Idreos

## **FITing-Tree: A Data-aware Index Structure**

*Alex Galakatos (Brown University), Michael Markovitch (Brown University), Carsten Binnig (TU Darmstadt), Rodrigo Fonseca (Brown University), Tim Kraska (MIT)*

## **Hyperion: Building the Largest In-memory Search Tree**

*Markus Mäsker (Johannes Gutenberg University Mainz), Tim Süß (University of Applied Science Fulda), Lars Nagel (Loughborough University), Lingfang Zeng (Huazhong University of Science and Technology), André Brinkmann (Johannes Gutenberg University Mainz)*

## **Designing Succinct Secondary Indexing Mechanism by Exploiting Column Correlations**

*Yingjun Wu (IBM Almaden Research Center), Jia Yu (Arizona*

*State University), Yuanyuan Tian (IBM Almaden Research Center), Richard Sidle (IBM), Ronald Barber (IBM Almaden Research Center)*

## **AI Meets AI: Leveraging Query Executions to Improve Index Recommendations**

*Bailu Ding (Microsoft Research), Sudipto Das (Microsoft Research), Ryan Marcus (Brandeis University), Wentao Wu (Microsoft Research), Surajit Chaudhuri (Microsoft Research), Vivek Narasayya (Microsoft Research)*

SIGMOD Research 13: Fairness, Uncertainty (Thu  
11:30-12:50)

**Room:** Administratiezaal

**Chair:** Ke Yi

### **Designing Fair Ranking Schemes**

*Abolfazl Asudeh (University of Michigan), H. V. Jagadish (University of Michigan), Julia Stoyanovich (New York University), Gautam Das (University of Texas at Arlington)*

### **Anti-Freeze for Large and Complex Spreadsheets: Asynchronous Formula Computation**

*Mangesh Bendre (University of Illinois Urbana-Champaign), Tana Wattanawaroон (University of Illinois Urbana-Champaign), Kelly Mack (University of Illinois Urbana-Champaign), Kevin Chang (University of Illinois Urbana-Champaign), Aditya Parameswaran (University of Illinois Urbana-Champaign)*

### **Anytime Approximation in Probabilistic Databases via Scaled Dissociations**

*Maarten Van den Heuvel (University of Antwerp), Peter Ivanov (Northeastern University), Wolfgang Gatterbauer (Northeastern University), Floris Geerts (University of Antwerp), Martin Theobald (University of Luxembourg)*

## **Uncertainty Annotated Databases - A Lightweight Approach for Approximating Certain Answers**

*Su Feng (Illinois Institute of Technology), Aaron Huber (University at Buffalo), Boris Glavic (Illinois Institute of Technology), Oliver Kennedy (University at Buffalo)*

**SIGMOD Research 14: Graphs 2**

(Thu 11:30-12:50)

**Room:** Berlage Zaal

**Chair:** Sourav S Bhowmick

## **Efficient Estimation of Heat Kernel PageRank for Local Clustering**

*Renchi Yang (Nanyang Technological University), Xiaokui Xiao (National University of Singapore), Zhewei Wei (Renmin University of China), Sourav Bhowmick (Nanyang Technological University), Jun Zhao (Nanyang Technological University), Rong-Hua Li (Beijing Institute of Technology)*

## **Fractal: A General-Purpose Graph Pattern Mining System**

*Vinicio Dias (Universidade Federal de Minas Gerais), Carlos Teixeira (Universidade Federal de Minas Gerais), Dorival Guedes (Universidade Federal de Minas Gerais), Wagner Meira (Universidade Federal de Minas Gerais), Srinivasan Parthasarathy (Ohio State University)*

## **Experimental Analysis of Streaming Algorithms for Graph Partitioning**

*Anil Pacaci (University of Waterloo), Tamer Özsu (University of Waterloo)*

## **Interactive Graph Search**

*Yufei Tao (Chinese University of Hong Kong), Yuanbing Li (Tsinghua University), Guoliang Li (Tsinghua University)*

**Thursday 04/07/2018 12:50-14:20**

Lunch

(Thu 12:50-14:20)

**Room:** Grote Zaal

**Thursday 04/07/2018 14:20-16:20**

SIGMOD Research 15: Graphs 3

(Thu 14:20-16:20)

**Room:** Effectenbeurszaal

**Chair:** Xuemin Lin

## **Optimizing Declarative Graph Queries at Large Scale**

*Qizhen Zhang (University of Pennsylvania), Akash Acharya (University of Pennsylvania), Hongzhi Chen (The Chinese University of Hong Kong), Simran Arora (University of Pennsylvania), Ang Chen (Rice University), Vincent Liu (University of Pennsylvania), Boon Loo (University of Pennsylvania)*

## **Efficient Subgraph Matching: Harmonizing Dynamic Programming, Adaptive Matching Order, and Failing Set Together**

*Myoungji Han (Seoul National University), Hyunjoon Kim (Seoul National University), Geonmo Gu (Seoul National University), Kunsoo Park (Seoul National University), Wook-Shin Han (Pohang University of Science and Technology (POSTECH))*

## **CECI: Compact Embedding Cluster Index for Scalable Subgraph Matching**

*Bibek Bhattacharai (George Washington University), Hang Liu (University of Massachusetts Lowell), H. Howie Huang (George Washington University)*

## **Efficiently Answering Regular Simple Path Queries on Large Labeled Networks**

*Sarisht Wadhwa (IIT Delhi), Anagh Prasad (IIT Delhi), Sayan Ranu (IIT Delhi), Amitabha Bagchi (IIT Delhi), Srikanta Bedathur (IIT Delhi)*

## **Answering Why-questions by Exemplars in Attributed Graphs**

*Mohammad Hossein Namaki (Washington State University), Qi Song (Washington State University), Yinghui Wu (Washington State University), Shengqi Yang (WeWork Technology)*

## **An Efficient Index for RDF Query Containment**

*Theofilos Mailis (Athena Research Centre & University of Athens), Yannis Kotidis (Athens University of Economics and Business), Vaggelis Nikolopoulos (University of Athens), Evgeny Kharlamov (University of Oslo & Bosch Center for AI), Ian Horrocks (University of Oxford), Yannis Ioannidis (Athena Research Centre & University of Athens)*

**SIGMOD Research 16: Machine Learning (Thu 14:20-16:20)**

**Room:** Graanbeurszaal

**Chair:** Theodoros Rekatsinas

## **Tuple-oriented Compression for Large-scale Mini-batch Stochastic Gradient Descent**

*Fengan Li (University of Wisconsin, Madison), Lingjiao Chen (University of Wisconsin, Madison), Yijing Zeng (University of Wisconsin, Madison), Arun Kumar (University of California, San Diego), Xi Wu (University of Wisconsin, Madison), Jeffrey Naughton (University of Wisconsin, Madison), Jignesh Patel (University of Wisconsin, Madison)*

## **Towards Model-based Pricing for Machine Learning in a Data Marketplace**

*Lingjiao Chen (University of Wisconsin, Madison), Paraschos Koutris (University of Wisconsin, Madison), Arun Kumar (Uni-*

*versity of California, San Diego)*

## **DBEst: Revisiting Approximate Query Processing Engines with Machine Learning Models**

*Qingzhi Ma (University of Warwick), Peter Triantafillou (University of Warwick)*

## **Enabling and Optimizing Non-linear Feature Interactions in Factorized Linear Algebra**

*Side Li (University of California, San Diego), Lingjiao Chen (University of Wisconsin, Madison), Arun Kumar (University of California, San Diego)*

## **Incremental and Approximate Inference for Faster Occlusion-based Deep CNN Explanations**

*Supun Nakandala (University of California, San Diego), Arun Kumar (University of California, San Diego), Yannis Papakonstantinou (University of California, San Diego)*

## **MNC: Structure-Exploiting Sparsity Estimation for Matrix Expressions**

*Johanna Sommer (IBM Germany), Matthias Boehm (Graz University of Technology), Alexandre Evfimievski (IBM Almaden Research Center), Berthold Reinwald (IBM Almaden Research Center), Peter Haas (University of Massachusetts Amherst)*

**SIGMOD Research 17: Scalability**

(Thu 14:20-16:20)

**Room:** Administratiezaal

**Chair:** Norman May

## **A Scalable Index for Top-k Subtree Similarity Queries**

*Daniel Kocher (University of Salzburg), Nikolaus Augsten (University of Salzburg)*

## **A Layered Aggregate Engine for Analytics Work-**

## **loads**

*Maximilian Schleich (University of Oxford), Dan Olteanu (University of Oxford), Mahmoud Abo Khamis (RelationalAI), Hung Ngo (RelationalAI), XuanLong Nguyen (University of Michigan)*

## **Towards Scalable Hybrid Stores: Constraint-Based Rewriting to the Rescue**

*Rana Alotaibi (University of California, San Diego), Damian Bursztyn (Thales), Alin Deutsch (University of California, San Diego), Ioana Manolescu (Inria & Ecole polytechnique), Stamatis Zampetakis (Orchestra Networks)*

## **MIFO: A Query-Semantic Aware Resource Allocation Policy**

*Prajakta Kalmegh (Duke University), Shivnath Babu (Unravel Data Systems)*

## **Dissecting the Performance of Strongly-Consistent Replication Protocols**

*Ailidani Ailijiang (Microsoft), Aleksey Charapko (University at Buffalo, SUNY), Murat Demirbas (University at Buffalo, SUNY)*

## **FishStore: Faster Ingestion with Subset Hashing**

*Dong Xie (University of Utah), Badrish Chandramouli (Microsoft Research), Yinan Li (Microsoft Research), Donald Kossmann (Microsoft Research)*

**SIGMOD Industry 3: Data Platforms**

(Thu 14:20-16:20)

**Room:** Berlage Zaal

**Chair:** Ying Zhang

## **CFS: A Distributed File System for Large Scale Container Platforms**

*Haifeng Liu (University of Science and Technology of China), Wei Ding (JD.com), Yuan Chen (JD.com), Weilong Guo (JD.com),*

*Shuoran Liu (JD.com), Tianpeng Li (JD.com), Mofei Zhang (JD.com), Jianxing Zhao (JD.com), Hongyin Zhu (JD.com), Zhengyi Zhu (JD.com)*

## **Socrates: The New SQL Server in the Cloud**

*Panagiotis Antonopoulos (Microsoft), Alex Budovski (Microsoft), Cristian Diaconu (Microsoft), Alejandro Hernandez (Microsoft), Jack Hu (Microsoft), Hanuma Kodavalla (Microsoft), Donald Kossmann (Microsoft Research), Umar Farooq Minhas (Microsoft Research), Naveen Prakash (Microsoft), Vijendra Purohit (Microsoft), Hugh Qu (Microsoft), Chaitanya Sreenivas Ravella (Microsoft), Krystyna Reisteter (Microsoft), Sheetal Shrotri (Microsoft), Dixin Tang (University of Chicago), Vikram Wakade (Microsoft)*

## **One SQL to Rule Them All - an Efficient and Syntactically Idiomatic Approach to Management of Streams and Tables**

*Edmon Begoli (Oak Ridge National Laboratory), Tyler Aki-dau (Google), Fabian Hueske (Ververica), Julian Hyde (Looker Inc.), Kathryn Knight (Oak Ridge National Laboratory), Kenneth Knowles (Google)*

## **Apache Hive: From MapReduce to Enterprise-grade Big Data Warehousing**

*Jesús Camacho-Rodríguez (Hortonworks), Ashutosh Chauhan (Hortonworks), Alan Gates (Hortonworks), Eugene Koifman (Hortonworks), Owen O'Malley (Hortonworks), Vineet Garg (Hortonworks), Zoltan Hajnrich (Hortonworks), Sergey Shelukhin (Hortonworks), Prasanth Jayachandran (Hortonworks), Siddharth Seth (Hortonworks), Deepak Jaiswal (Hortonworks), Slim Bouguerra (Hortonworks), Nishant Bangarwa (Hortonworks), Sankar Hariappan (Hortonworks), Anishek Agarwal (Hortonworks), Jason Dere (Hortonworks), Daniel Dai (Horton-*

*works), Thejas Nair (Hortonworks), Nita Dembla (Hortonworks), Gopal Vijayaraghavan (Hortonworks), Günther Hagleitner (Hortonworks)*

## **FoundationDB Record Layer: A Multi-Tenant Structured Datastore**

*Christos Chrysafis (Apple), Ben Collins (Apple), Scott Dugas (Apple), Jay Dunkelberger (Apple), Moussa Ehsan (Apple), Scott Gray (Apple), Alec Grieser (Apple), Ori Herrnstadt (Apple), Kfir Lev-Ari (Apple), Tao Lin (Apple), Mike McMahon (Apple), Nicholas Schiefer (Apple), Alexander Shraer (Apple)*

## **Data Platform for Machine Learning**

*Pulkit Agrawal (Apple), Rajat Arya (Apple), Aanchal Bindal (Apple), Sandeep Bhatia (Apple), Anupriya Gagneja (Apple), Joseph Godlewski (Apple), Yucheng Low (Apple), Timothy Muss (Apple), Mudit Manu Paliwal (Apple), Sethu Raman (Apple), Vishruth Shah (Apple), Bochao Shen (Apple), Laura Sugden (Apple), Kaiyu Zhao (Apple), Ming-Chuan Wu (Apple)*

**Thursday 04/07/2018 16:20-17:50**

Poster & Demo Groups A and C

(Thu 16:20-17:50)

Room: Grote Zaal

**One poster for each SIGMOD paper presented on Thursday.**

## **FindYourFavorite: An Interactive System for Finding the User's Favorite Tuple in the Database**

*Min Xie (Hong Kong University of Science and Technology), Tianwen Chen (Hong Kong University of Science and Technology), Raymond Chi-Wing Wong (Hong Kong University of Science and Technology)*

## **Large Scale Graph Mining with G-Miner**

*Hongzhi Chen (The Chinese University of Hong Kong), Xiaoxi Wang (The Chinese University of Hong Kong), Chenghuan Huang (The Chinese University of Hong Kong), Juncheng Fang (The Chinese University of Hong Kong), Yifan Hou (The Chinese University of Hong Kong), Changji Li (The Chinese University of Hong Kong), James Cheng (The Chinese University of Hong Kong)*

## **ANMAT: Automatic Knowledge Discovery and Error Detection through Pattern Functional Dependencies**

*Abdulhakim Qahtan (QCRI, HBKU), Nan Tang (QCRI, HBKU), Mourad Ouzzani (QCRI, HBKU), Yang Cao (University of Edinburgh), Michael Stonebraker (MIT)*

## **Estimating Cardinalities with Deep Sketches**

*Andreas Kipf (Technische Universität München), Dimitri Vorona (Technische Universität München), Jonas Müller (Technische Universität München), Thomas Kipf (University of Amsterdam), Bernhard Radke (Technische Universität München), Viktor Leis (Technische Universität München), Peter Boncz (CWI), Thomas Neumann (Technische Universität München), Alfons Kemper (Technische Universität München)*

## **Unit Testing Data with Deequ**

*Sebastian Schelter (Amazon Research), Felix Biessmann (Amazon Research), Dustin Lange (Amazon Research), Tammo Rukat (Amazon Research), Philipp Schmidt (Amazon Research), Stephan Seufert (Amazon Research), Pierre Brunelle (Amazon Research), Andrey Taptunov (Amazon Research)*

## **DuckDB: an Embeddable Analytical Database**

*Mark Raasveldt (CWI), Hannes Mühleisen (CWI)*

## **CLASH: A High-Level Abstraction for Optimized, Multi-Way Stream Joins over Apache Storm**

*Manuel Dossinger (TU Kaiserslautern), Sebastian Michel (TU Kaiserslautern), Constantin Roudsarabi (TU Kaiserslautern)*

## **PgCuckoo: Laying Plan Eggs in PostgreSQL’s Nest**

*Denis Hirn (Universität Tübingen), Torsten Grust (Universität Tübingen)*

## **Demonstration of ModelarDB: Model-Based Management of Dimensional Time Series**

*Søren Kejser Jensen (Aalborg University), Torben Bach Pedersen (Aalborg University), Christian Thomsen (Aalborg University)*

## **NEURON: Query Execution Plan Meets Natural Language Processing For Augmenting DB Education**

*Siyuan Liu (Nanyang Technological University), Sourav Bhowmick (Nanyang Technological University), Wanlu Zhang (Nanyang Technological University), Shu Wang (Nanyang Technological University), Wanyi Huang (Nanyang Technological University), Shafiq Joty (Nanyang Technological University)*

## **PIClean: A Probabilistic and Interactive Data Cleaning System**

*Zhuoran Yu (Georgia Institute of Technology), Xu Chu (Georgia Institute of Technology)*

## **Apollo: A Dataset Profiling and Operator Modeling System**

*Tasos Bakogiannis (National Technical University of Athens), Ioannis Giannakopoulos (National Technical University of Athens), Dimitrios Tsoumakos (Ionian University), Nectarios Koziris (National Technical University of Athens)*

## **Peering through the Dark: An Owl’s View of Inter-job Dependencies and Jobs’ Impact in Shared Clus-**

ters

*Andrew Chung (Carnegie Mellon University), Carlo Curino (Microsoft), Subru Krishnan (Microsoft), Konstantinos Karanasos (Microsoft), Panagiotis Garefalakis (Imperial College London), Gregory Ganger (Carnegie Mellon University)*

## **Visual Exploration of Time Series Anomalies with Metro-Viz**

*Philipp Eichmann (Brown University), Franco Solleza (Brown University), Nesime Tatbul (Intel Labs and MIT), Stan Zdonik (Brown University)*

## **Data Debugging and Exploration with Vizier**

*Mike Brachmann (University at Buffalo), Carlos Bautista (New York University), Sonia Castelo (New York University), Su Feng (Illinois Institute of Technology), Juliana Freire (New York University), Boris Glavic (Illinois Institute of Technology), Oliver Kennedy (University of Buffalo), Heiko Meller (New York University), Rami Rampin (New York University), William Spoth (University at Buffalo), Ying Yang (Oracle)*

## **CrowdGame: A Game-Based Crowdsourcing System for Cost-Effective Data Labeling**

*Tongyu Liu (Renmin University of China), Jingru Yang (Renmin University of China), Ju Fan (Renmin University of China), Zhewei Wei (Renmin University of China), Guoliang Li (Tsinghua University), Xiaoyong Du (Renmin University of China)*

## **Ursprung: Provenance for Large-Scale Analytics Environments**

*Lukas Rupprecht (IBM Almaden Research Center), James Davis (Virginia Tech & IBM Systems), Constantine Arnold (IBM Almaden Research Center), Alexander Lubbock (Vanderbilt University), Darren Tyson (Vanderbilt University), Deepavali Bhagwat (IBM Almaden Research Center)*

## **BlockchainDB - Towards a Shared Database on Blockchains**

*Muhammad El-Hindi (TU Darmstadt), Martin Heyden (TU Darmstadt), Carsten Binnig (TU Darmstadt), Ravi Ramamurthy (Microsoft Research), Arvind Arasu (Microsoft Research), Donald Kossmann (Microsoft Research)*

## **Fluid: A Blockchain based Framework for Crowd-sourcing**

*Siyuan Han (Hong Kong University of Science and Technology), Zihuan Xu (Hong Kong University of Science and Technology), Yuxiang Zeng (Hong Kong University of Science and Technology), Lei Chen (Hong Kong University of Science and Technology)*

## **MorphStore - In-Memory Query Processing based on Morphing Compressed Intermediates LIVE**

*Dirk Habich (Technische Universität Dresden), Patrick Damme (Technische Universität Dresden), Annett Ungethüm (Technische Universität Dresden), Johannes Pietrzyk (Technische Universität Dresden), Alexander Krause (Technische Universität Dresden), Juliana Hildebrandt (Technische Universität Dresden), Wolfgang Lehner (Technische Universität Dresden)*

## **MapRepair: Mapping and Repairing under Policy Views**

*Angela Bonifati (Lyon 1 University & Liris CNRS), Ugo Comignani (Lyon 1 University & Liris CNRS), Efthymia Tsamoura (University of Oxford)*

## **RATest: Explaining Wrong Relational Queries Using Small Examples**

*Zhengjie Miao (Duke University), Sudeepa Roy (Duke University), Jun Yang (Duke University)*

## **NAVIGATE: Explainable Visual Graph Exploration**

## **by Examples**

*Mohammad Hossein Namaki (Washington State University), Qi Song (Washington State University), Yinghui Wu (Washington State University)*

## **C2Metadata: Automating the Capture of Data Transformations from Statistical Scripts in Data Documentation**

*Jie Song (University of Michigan), George Alter (University of Michigan), H. V. Jagadish (University of Michigan)*

## **MithraRanking: A System for Responsible Ranking Design**

*Yifan Guan (University of Michigan), Abolfazl Asudeh (University of Michigan), Pranav Mayuram (University of Michigan), H. V. Jagadish (University of Michigan), Julia Stoyanovich (New York University), Gerome Miklau (University of Massachusetts Amherst), Gautam Das (University of Texas at Arlington)*

## **NEWS: News Event Walker and Summarizer**

*Radityo Eko Prasojo (Free University of Bozen-Bolzano), Mouna Kacimi (Free University of Bozen-Bolzano), Werner Nutt (Free University of Bozen-Bolzano)*

## **Cost-Effective, Workload-Adaptive Migration of Big Data Applications to the Cloud**

*Victor Giannakouris (Unravel Data Systems), Alejandro Fernandez (Unravel Data Systems), Alkis Simitsis (Unravel Data Systems), Shivnath Babu (Unravel Data Systems)*

## **ChronosDB in Action: Manage, Process, and Visualize Big Geospatial Arrays in the Cloud**

*Ramon Antonio Rodriges Zalipynis (National Research University Higher School of Economics)*

## Thursday 04/07/2018 17:30-18:00

ADS Reception 1 - sponsored by Elsevier (Thu 17:30-18:00)

**Room:** Effectenbeurszaal

## Thursday 04/07/2018 18:00-19:30

ADS Event (Thu 18:00-19:30)

**Room:** Graanbeurszaal

**ADS Event**

*Jeanne Kroeger (ADS)*

## Thursday 04/07/2018 19:30-20:00

ADS Reception 2 - sponsored by Elsevier (Thu 19:30-20:00)

**Room:** Effectenbeurszaal

## Friday 07/07/2019 08:30-09:00

Coffee + Light Breakfast (Fri 08:30-09:00)

**Room:** Grote Zaal

## Friday 07/07/2019 09:00-10:30

Tutorial 4: part 1 (Fri 09:00-10:30)

**Room:** Mendes da Costa Kamer

### Classical and Contemporary Approaches to Big Time Series Forecasting

*Christos Faloutsos (Carnegie Mellon University & Amazon), Jan Gasthaus (AWS AI Labs), Tim Januschowski (AWS AI Labs), Yuyang Wang (AWS AI Labs)*

Tutorial 6: part 1 (Fri 09:00-10:30)

**Room:** Effectenbeurszaal

**From Auto-tuning One Size Fits All to Self-designed and Learned Data-intensive Systems**

*Stratos Idreos (Harvard University), Tim Kraska (MIT)*

HILDA 2019: Session 1

(Fri 09:00-10:30)

**Room:** Berlage Zaal

**HILDA 2019: the International Workshop on Human-In-the-Loop Data Analytics**

*Leilani Battle (University of Maryland), Surajit Chaudhuri (Microsoft), Arnab Nandi (The Ohio State University)*

aiDM 2019: Session 1

(Fri 09:00-10:30)

**Room:** Administratiezaal

**aiDM 2019: the 2nd International Workshop on Exploiting Artificial Intelligence Techniques for Data Management**

*Rajesh Bordawekar (IBM T. J. Watson Research Center), Oded Shmueli (Computer Science Department, Technion)*

SBD 2019: Session 1

(Fri 09:00-10:30)

**Room:** Veilingzaal

**SBD 2019: the Fourth International Workshop on Semantic Big Data**

*Sven Groppe (University of Lübeck), Le Gruenwald (University of Oklahoma)*

**Friday 07/07/2019 10:30-11:00**

Coffee

(Fri 10:30-11:00)

**Room:** Grote Zaal

**Friday 07/07/2019 11:00-12:30**

Tutorial 4: part 2

(Fri 11:00-12:30)

**Room:** Mendes da Costa Kamer

**Classical and Contemporary Approaches to Big Time Series Forecasting**

*Christos Faloutsos (Carnegie Mellon University & Amazon), Jan Gasthaus (AWS AI Labs), Tim Januschowski (AWS AI Labs), Yuyang Wang (AWS AI Labs)*

Tutorial 6: part 2

(Fri 11:00-12:30)

**Room:** Effectenbeurszaal

**From Auto-tuning One Size Fits All to Self-designed and Learned Data-intensive Systems**

*Stratos Idreos (Harvard University), Tim Kraska (MIT)*

HILDA 2019: Session 2

(Fri 11:00-12:30)

**Room:** Berlage Zaal

**HILDA 2019: the International Workshop on Human-In-the-Loop Data Analytics**

*Leilani Battle (University of Maryland), Surajit Chaudhuri (Microsoft), Arnab Nandi (The Ohio State University)*

aiDM 2019: Session 2

(Fri 11:00-12:30)

**Room:** Administratiezaal

**aiDM 2019: the 2nd International Workshop on Exploiting Artificial Intelligence Techniques for Data Management**

*Rajesh Bordawekar (IBM T. J. Watson Research Center), Oded Shmueli (Computer Science Department, Technion)*

**Room:** Veilingzaal

**SBD 2019: the Fourth International Workshop on Semantic Big Data**

*Sven Groppe (University of Lübeck), Le Gruenwald (University of Oklahoma)*

LDBC Technical User Community Meeting: session 1

(Fri 11:00-12:30)

**Room:** Ontvangkamer

**LDBC Technical User Community Meeting**

*Peter Boncz (LDBC (CWI)), Alastair Green (LDBC (Neo4j))*

**Friday 07/07/2019 12:30-14:00**

Lunch

(Fri 12:30-14:00)

**Room:** Grote Zaal

**Friday 07/07/2019 14:00-15:30**

Tutorial 5: part 1

(Fri 14:00-15:30)

**Room:** Mendes da Costa Kamer

**Data Pipelines for User Group Analytics**

*Behrooz Omidvar-Tehrani (University of Grenoble Alpes), Sihem Amer-Yahia (University of Grenoble Alpes and CNRS)*

Tutorial 7

(Fri 14:00-15:30)

**Room:** Effectenbeurszaal

**Schemas and Types for JSON Data: From Theory to Practice**

*Mohamed-Amine Baazizi (Sorbonne Université, LIP6 UMR*

*7606), Dario Colazzo (Université Paris-Dauphine, PSL Research University), Giorgio Ghelli (Università di Pisa), Carlo Sartiani (Università della Basilicata)*

HILDA 2019: Session 3

(Fri 14:00-15:30)

**Room:** Berlage Zaal

## **HILDA 2019: the International Workshop on Human-In-the-Loop Data Analytics**

*Leilani Battle (University of Maryland), Surajit Chaudhuri (Microsoft), Arnab Nandi (The Ohio State University)*

aiDM 2019: Session 3

(Fri 14:00-15:30)

**Room:** Administratiezaal

## **aiDM 2019: the 2nd International Workshop on Exploiting Artificial Intelligence Techniques for Data Management**

*Rajesh Bordawekar (IBM T. J. Watson Research Center), Oded Shmueli (Computer Science Department, Technion)*

SBD 2019: Session 3

(Fri 14:00-15:30)

**Room:** Veilingzaal

## **SBD 2019: the Fourth International Workshop on Semantic Big Data**

*Sven Groppe (University of Lübeck), Le Gruenwald (University of Oklahoma)*

LDBC Technical User Community Meeting: session 2

(Fri 14:00-15:30)

**Room:** Ontvangkamer

## **LDBC Technical User Community Meeting**

Peter Boncz (*LDBC (CWI)*), Alastair Green (*LDBC (Neo4j)*)

**Friday 07/07/2019 15:30-16:30**

Coffee + Workshop Posters

(Fri 15:30-16:30)

**Room:** Grote Zaal

**Friday 07/07/2019 16:30-18:00**

Tutorial 5: part 2

(Fri 16:30-18:00)

**Room:** Mendes da Costa Kamer

### **Data Pipelines for User Group Analytics**

*Behrooz Omidvar-Tehrani (University of Grenoble Alpes), Sihem Amer-Yahia (University of Grenoble Alpes and CNRS)*

HILDA 2019: Session 4

(Fri 16:30-18:00)

**Room:** Berlage Zaal

### **HILDA 2019: the International Workshop on Human-In-the-Loop Data Analytics**

*Leilani Battle (University of Maryland), Surajit Chaudhuri (Microsoft), Arnab Nandi (The Ohio State University)*

aiDM 2019: Session 4

(Fri 16:30-18:00)

**Room:** Administratiezaal

### **aiDM 2019: the 2nd International Workshop on Exploiting Artificial Intelligence Techniques for Data Management**

*Rajesh Bordawekar (IBM T. J. Watson Research Center), Oded Shmueli (Computer Science Department, Technion)*

SBD 2019: Session 4

(Fri 16:30-18:00)

**Room:** Veilingzaal

**SBD 2019: the Fourth International Workshop on Semantic Big Data**

*Sven Groppe (University of Lübeck), Le Gruenwald (University of Oklahoma)*

LDBC Technical User Community Meeting: session 3

(Fri 16:30-18:00)

**Room:** Ontvangkamer

**LDBC Technical User Community Meeting**

*Peter Boncz (LDBC (CWI)), Alastair Green (LDBC (Neo4j))*

# Building the systems to bring the world closer together

Our researchers and engineers are constant innovators as they design and build next generation, scalable, fast, reliable, and efficient systems.

From distributed systems, to data centers, hardware, storage, mobile and beyond, the entire Facebook platform is our lab for research, development, and innovation.

Come visit our booth for more info and learn more at [research.fb.com](http://research.fb.com)

**facebook research**

# Work with us to advance purposeful innovation

At Elsevier, we're combining technology and trusted information to enable giant leaps forward in science and healthcare.

Work with health and scientific data, or join the editorial team working on new Elsevier Computer Science journals like *Array*, *Internet of Things*, and *Software Impact*.

Discover career opportunities at Elsevier

<http://bit.ly/elsevierjobs>



ELSEVIER



Solve real-world  
problems for  
**millions** of users,  
**exabytes** of data,  
and **trillions** of  
transactions

Join **Azure Data** and innovate where the data is

**Explore further**  
[aka.ms/buildingthefuture](https://aka.ms/buildingthefuture)



**Browse jobs**  
[aka.ms/azuredatajobs](https://aka.ms/azuredatajobs)





MORGAN & CLAYPOOL  
PUBLISHERS



Springer

# World's First “Self-Driving” Database



**No Human Labor – Half the Cost**  
**No Human Error – 100x More Reliable**

**ORACLE®**

[oracle.com/selfdrivingdb](http://oracle.com/selfdrivingdb)



## We help people see and understand data.

"A system that allows users to create stunning graphs interactively and easily from large multidimensional datasets."

– Jim Gray, Turing Award winner, about Tableau

### A Home for Innovation: From VizQL to Hyper

VizQL is a visual query language that is the foundation of Tableau's query generation. It can speak to all major database systems as well as Tableau's blazingly fast data engine Hyper.

### Tableau Acquires Hyper

In March 2016, Tableau acquired Hyper, a high-performance database system that started as a research project at Technical University Munich and later spun out into a startup.

Key technical personnel that pioneered code generation for main-memory database systems and achieved breakthroughs in query optimization and hybrid transactional and analytical processing continues to innovate from Tableau's offices in Seattle, Palo Alto, and our European research and development center in Munich.

### We're Hiring!

Calling all technical talent – our Engineering team is hiring Software Engineers on both Development and Test teams as well as Research Engineers.

Visit [careers.tableau.com](http://careers.tableau.com)



# TO MAKE IT EASY TO DO BUSINESS ANYWHERE

We enable businesses to transform the way they market, sell and operate. Our businesses are comprised of core commerce, cloud computing, digital media and entertainment, innovation initiatives and others. Through our subsidiary Cainiao Network and investee affiliate Koubei, respectively, we participate in the logistics and local services sectors. In addition, we have a strategic relationship with Ant Financial Services, the financial services group that operates mainly through Alipay, the leading third-party online payment platform in China.

## • Meet @ Alibaba

We enable hundreds of millions of commercial and social interactions among our users, between consumers and merchants, and among businesses every day.

## • Work @ Alibaba

We empower our customers with the fundamental infrastructure for commerce and data technology, so that they can build businesses and create value that can be shared among our ecosystem participants.



## • Live @ Alibaba

We strive to expand our products and services to become central to the everyday lives of our customers.



**It's only  
impossible  
until it works**

**Ready to pioneer?  
Apply at [Amazon.jobs](#)**

Amazon is an Equal Opportunity Employer



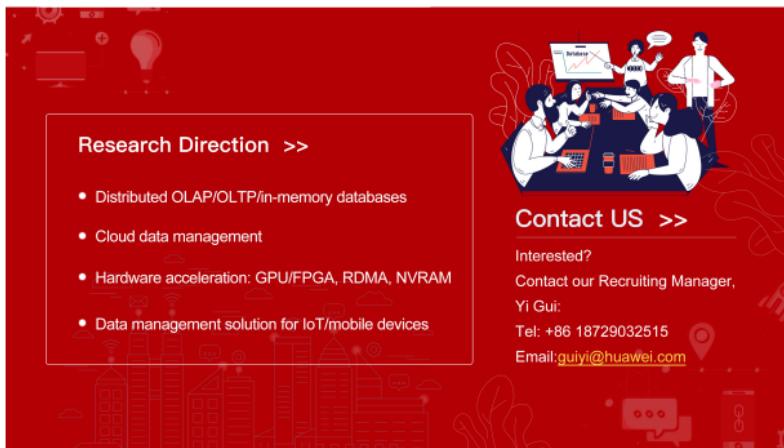




databricks®

Google

# Welcome to join Huawei Database Team



The background is a solid red color with a faint white graphic of a city skyline and various icons like servers, databases, and charts.

### Research Direction >>

- Distributed OLAP/OLTP/in-memory databases
- Cloud data management
- Hardware acceleration: GPU/FPGA, RDMA, NVRAM
- Data management solution for IoT/mobile devices



### Contact US >>

Interested?  
Contact our Recruiting Manager,  
Yi Gui:  
Tel: +86 18729032515  
Email: [guiyi@huawei.com](mailto:guiyi@huawei.com)



A young woman with dark hair and glasses is sitting on a bright red couch in a modern office setting. She is wearing a yellow long-sleeved shirt and blue jeans, and is looking towards the camera while using a silver laptop on a small glass table in front of her.

## IBM Research is hiring

We live in a moment of remarkable change and opportunity. Data and technology are transforming industries and societies, ushering in a new era of Artificial Intelligence. IBM Research is a leader in this worldwide transformation, building on a long history of innovation.

For more than seven decades, IBM Research has defined the future of technology. Our scientists, among them six Nobel Laureates and six Turing Award winners, have produced ten U.S. National Medals of Technology and five U.S. National Medals of Science. Along the way we helped put a man on the moon, defeated Kasparov at chess, and built a Jeopardy! champion named Watson.

At IBM, you can achieve what others think is impossible. And in doing so, you'll play a significant role in shaping the future. Join us.

**Discover what you can do at IBM Research**  
[ibm.com/jobs](http://ibm.com/jobs)

© Copyright IBM Corporation 2018. IBM, the IBM logo and ibm.com are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. See current list at [ibm.com/trademark](http://ibm.com/trademark). Other product and service names might be trademarks of IBM or other companies.



## Megagon Labs

*We empower people with better information  
to make their best decisions.*

Megagon Labs is the research lab for Recruit Holdings Co., Ltd that provides over 200 online services in the areas where people make daily decisions.

We conduct world-class research in the several areas, including data management, natural language processing, machine learning, data integration, and artificial intelligence. We collaborate with research in universities and publish in top-notch conferences.

Megagon Labs is hiring Research Scientists and Software Engineers! Visit our website at [www.megagon.ai](http://www.megagon.ai) for more information.

[www.megagon.ai](http://www.megagon.ai)

## Join the SAP HANA Database Campus!

Are you passionate about Software Development and working on your IT degree? Do you want to start your career by working on the next generation database technologies?

Then join us – The SAP HANA Database Campus team!

The SAP HANA team develops a platform that performs parallel in-memory processing of huge data sets to offer extremely fast real-time responses for analytic and transactional queries. The platform also provides libraries for predictive, planning, text processing, spatial, and business analytics – all on the same architecture.

Learn more about the  
SAP HANA Database Campus





## THE DATA WAREHOUSE BUILT FOR THE CLOUD



Enabling every organization with the power, flexibility and instant elasticity to be data-driven.

© 2018 Snowflake Computing Inc. All Rights Reserved

### AT SNOWFLAKE WE ARE:

Passionate, Innovative, Ambitious

*We challenge ourselves at Snowflake to rethink what's possible for the data warehouse and deliver on that. We're looking for people who share that same passion and ambition.*

Snowflake is hiring top database developers from Europe and around the world to grow our teams in San Mateo, Seattle, and our newly formed European Database R&D Center located in Berlin.

Visit us at :  
[www.snowflake.com/careers](http://www.snowflake.com/careers)



**Avalanche**  
Announcing Actian Avalanche  
The industry's first Gen III  
cloud data warehouse service

You are used to working with the Avalanche. Now you are used to working with data warehouses. Now you can harness the power and scalability of a single, fully managed enterprise tier in the cloud or on-premises designed for the way you work.

**Deliver fast - Delivered Today**

- Cloud Data Warehouse
- Highly Scalable
- Fast Performance
- Easy to Deploy
- Flexible Pricing
- Cloud Native
- Open Standard
- Reliable
- Secure
- Compliant
- Private
- Public
- Hybrid
- Multi-Cloud

**From Start to Insight**  
In as little as 20 minutes  
[www.actian.com/avalanche](http://www.actian.com/avalanche)

**Actian**

# CAMBRIDGE UNIVERSITY PRESS

[www.ebay.com](http://www.ebay.com)    [www.ebayinc.com](http://www.ebayinc.com)



**Empowering People and Creating  
Economic Opportunity for All.**

LEAN  CALE

We are fearless in  
our innovation.

Think what we can do for your career.

We're seeking great minds like you to help us build the next generation in database technology. Join the team that is helping the world's most sophisticated organizations transform their industries by harnessing the power of data.



[mongodb.com/careers](http://mongodb.com/careers)  
[facebook.com/mongodb](https://facebook.com/mongodb)  
[@tongueDB](https://tongueDB) | [@MongoDBCareers](https://MongoDBCareers)



# now

the essence of knowledge



# TigerGraph

 undo

Intermittent software defects  
you can't reproduce?

Eliminate guesswork in software failure diagnosis.  
1. RECORD your program's execution as it fails  
2. REPLAY the recording backwards and forwards  
Get instant visibility into what your program did and why  
and get to the root cause of the issue with 100% certainty.

Learn more at  
<https://undo.io>

**SIGM~~X~~D/**  
A M S **PXDS**  
T E R  
D A M **2X19**

[www.sigmod2019.org](http://www.sigmod2019.org)