

43rd International Conference on Very Large Data Bases



28 August – 1 September 2017 · Munich, Germany



www.vldb.org/2017

Welcome Message

VLDB is a premier annual international forum for data management and database researchers, vendors, practitioners, application developers, and users. The annual conference consists of a mix of research talks, tutorials, demonstrations, and workshops. Its topical coverage includes current issues in data management, database, and information systems research. Data management and databases remain among the main technological cornerstones of the applications of the twenty-first century. With the emergence of Big Data, data-related technologies are becoming more important than ever before.

VLDB 2017 is taking place at the Technical University of Munich (TUM), one of Europe's top universities. It is committed to excellence in research and teaching, interdisciplinary education and the active promotion of promising young scientists. The university also forges strong links with companies and scientific institutions across the world. TUM was one of the first universities in Germany to be named a University of Excellence. Moreover, TUM regularly ranks among the best European universities in international rankings. The VLDB conference takes place at TUM's (original) downtown campus whereas most natural science and technical faculties, such as mechanical engineering, computer science, mathematics, chemistry, and physics are lo-

cated at the modern Garching campus, about 15 km north of Munich. The VLDB 2017 conference is held as part of TUM’s 50 year anniversary of establishing Computer Science as a scientific discipline in Munich. Since TUM was founded in 1868 (almost 150 years ago) it contributed significantly to the conversion of Bavaria from an agricultural to a highly industrialized region and also to the overall technological progress.

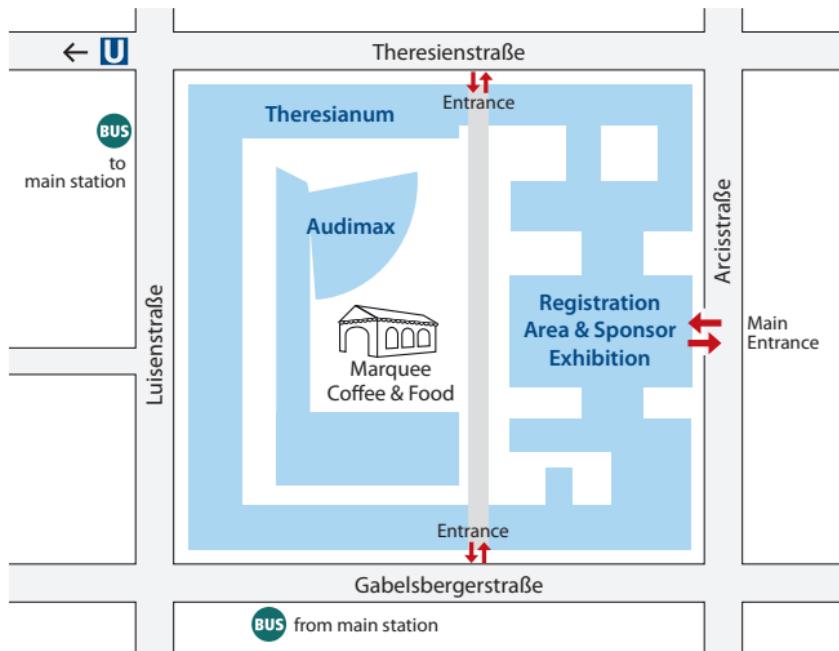
VLDB 2017 received almost 750 research paper submissions, of which 133 have been accepted for presentation at the conference. The conference program also includes 8 “roll-over” papers from VLDB 2016, for a total of 141 research papers. The full conference program also includes 20 papers from the industrial track, 35 research demonstrations, 8 tutorials, and a panel on cross-disciplinary research. Each day of the conference also features a plenary keynote presentation. The main conference is flanked by a variety of workshops on topics of particular interest to the community.

The VLDB 2017 technical program reflects the efforts of hundreds of members of the research community who have prepared papers, demonstrations and presentations for the conference. It is also the result of more than a year of work by the conference officers and the PVLDB Review Board. We’re looking forward to an outstanding program, and we hope that you enjoy it!

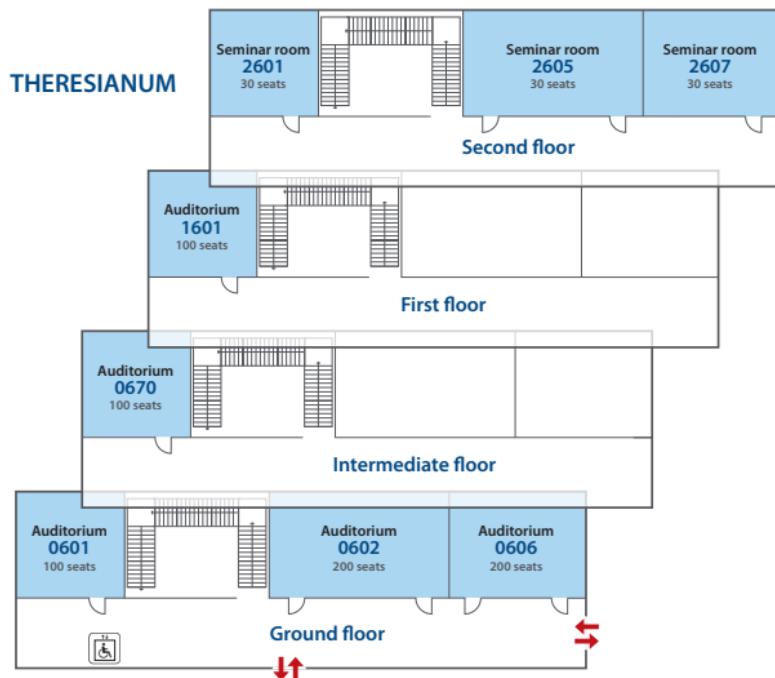
Alfons Kemper, TUM
Thomas Neumann, TUM
VLDB 2017 General Chairs

Peter Boncz, CWI
Ken Salem, University of Waterloo
VLDB 2017 Program Committee Chairs

Location TUM



Lecture Halls in the Theresianum



Street Map

Conference Location (A)

TUM, Arcisstraße 21

Reception (1)

Old Town Hall, Marienplatz 15

Banquet (B)

Hofbräuhaus, Platzl 9



Map data (c) OpenStreetMap contributors

Social Events

Both social events are close by the central place in Munich (Marienplatz). You can either walk there as shown on the map (2.5 km, about 30 min), or take public transport.

Public transport options:

- bus 100 (leaves from Gabelsbergerstraße, south entrance of TUM, direction Ostbahnhof) to Odeonsplatz. There, switch to the subway U3 or U6 towards Marienplatz (1 stop).
- live routing: <https://goo.gl/1T1Sg6>

The **Old Town Hall** is next to the Marienplatz, 50m to the east. It is a white building (<https://goo.gl/XE2OUV>), not to be confused with the large, red New Town Hall immediately next to the Marienplatz.

The **Hofbräuhaus** (<https://goo.gl/SUUQvW>) is a few minutes east and north of the Marienplatz, as shown on the map. Turn left behind the Old Town Hall into Sparkassenstraße, right into Münzstraße, and left into Platzl.

WLAN

Preferably, just use the *EDUROAM* network if you have eduroam access. Otherwise, follow the instructions below:



Wi-Fi-Guide for mwn-events

Wi-Fi name (SSID): mwn-events

Username: VLDB2017

Password: YTn06kdF

Valid from Fri Aug 25 06:00 2017 to Sat Sep 9
23:59 2017

Configuration profiles for wireless network access are available via the QR code or this URL:

<https://www.lrz.de/wlan> (follow the link mwn-events) Access to this site is available via the open Wi-Fi (the SSID) "lrz".

Our Sponsors

Platinum Sponsors



Gold Sponsors



Silver Sponsors



Bronze Sponsors



Exhibitors



Program at a Glance

Monday, August 28							
Time	Audimax	602	606	601	670	1601	260*
08:30-10:00		FADS	BIRTE	TPCTC	VLIoT	PhD	
10:00-10:30				Coffee break (tent)			
10:30-12:00		FADS	BIRTE	TPCTC	VLIoT	PhD	
12:00-13:30				Lunch (tent)			
13:30-15:00		FADS	BIRTE	TPCTC	VLIoT	PhD	
15:00-15:30				Coffee break (tent)			
15:30-17:00		FADS	BIRTE	TPCTC	VLIoT	PhD	
19:00		Welcome Reception with kind support of SAP (Old Townhall Munich)					

Tuesday, August 29							
Time	Audimax	602	606	601	670	1601	260*

Wednesday, August 30							
Time	Audimax	602	606	601	670	1601	260*
Thursday, August 31							
Time	Audimax	602	606	601	670	1601	260*
Friday, September 1							
Time	Audimax	602	606	601	670	1601	260*
08:30-10:00		MATES	ADMS	DMAH	DBPL	BOSS	BOSS
10:00-10:30	Coffee Break (tent)						
10:30-12:00		MATES	ADMS	DMAH	DBPL	BOSS	BOSS
12:00-13:30	Lunch (tent)						
13:30-15:00		MATES	ADMS	DMAH	DBPL	BOSS	BOSS
15:00-15:30	Coffee break (tent)						
15:30-17:00		MATES	ADMS	DMAH	DBPL	BOSS	BOSS

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

Coffee + Light Breakfast

Location: Grote Zaal

Sunday 06/30/2019 08:30-10:30

Tutorial 1: part 1 (starts 09:00)

Location: Effecten

Towards Democratizing Relational Data Visualization

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

DEEM 2019: Session 1

Location: Administratie

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)

GRADES-NDA 2019: Session 1

Location: Veiling

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)*

DSMM 2019: Session 1

Location: Mendes da Costa

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), Louiqa Raschid (University of Maryland)

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

Tutorial 1: part 2

Location: Effecten

Towards Democratizing Relational Data Visualization

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

Tutorial 3: part 2

Location: Berlage

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)*

GRADES-NDA 2019: Session 2

Location: Veiling

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

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

DSMM 2019: Session 2

Location: Mendes da Costa

DSMM 2019: the 5th Workshop on Data Sci- ence for Macro-modeling with Financial and Eco- nomic Datasets

*Douglas Burdick (IBM Almaden Research Center), Ra-
jasekar Krishnamurthy (IBM T. J. Watson Research Cen-
ter), Louiza Raschid (University of Maryland)*

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

Lunch + Workshop Posters

Location: Grote Zaal

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

Tutorial 2: part 1

Location: Effecten

Exploring the Data Wilderness through Examples

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

Phokion Kolaitis Special Event: part 1

Location: Berlage

Phokion Kolaitis Special Event

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

DEEM 2019: Session 3

Location: Administratie

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)

GRADES-NDA 2019: Session 3

Location: Veiling

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)*

DSMM 2019: Session 3

Location: Mendes da Costa

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), Louiqa Raschid (University of Maryland)

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

Coffee + Workshop Posters

Location: Grote Zaal

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

Tutorial 2: part 2 (ends 18:00)

Location: Effecten

Exploring the Data Wilderness through Examples

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

Phokion Kolaitis Special Event: part 2 (starts 16:00)

Location: Berlage

Phokion Kolaitis Special Event

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

DEEM 2019: Session 4

Location: Administratie

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)

GRADES-NDA 2019: Session 4

Location: Veiling

GRADES-NDA 2019: Joint International Work-

shop on Graph Data Management Experiences & Systems and Network Data Analytics

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

DSMM 2019: Session 4

Location: Mendes da Costa

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 18:30-20:30

PODS Reception

Location: Grote Zaal

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

Coffee + Light Breakfast

Location: Grote Zaal

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

PODS Opening + Keynote

Location: Effecten

Differential Privacy and the US Census

Cynthia Dwork (Harvard University)

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

PODS 1

Location: Effecten

DaMoN 2019: Session 1

Location: Berlage

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

Location: Grote Zaal

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

PODS Test-of-Time & Gems

Location: Effecten

Latent Semantic Indexing: A Probabilistic Analysis

Christos Papadimitriou (Columbia University)

Database Repairs and Consistent Query Answering: Origins and Further Developments

Leopoldo Bertossi (RelationalAI & Carleton University)

DaMoN 2019: Session 2

Location: Berlage

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)

Location: Grote Zaal

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

PODS 2

Location: Effecten

DaMoN 2019: Session 3

Location: Berlage

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

Location: Grote Zaal

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

PODS 3

Location: Effecten

DaMoN 2019: Session 4

Location: Berlage

**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

Location: Veiling

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

Location: Berlage

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

Location: Grote Zaal

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

SIGMOD Welcome + Keynote

Location: Effecten

Responsible Data Science

Lise Getoor (University of California, Santa Cruz)

Tuesday 02/07/2018 10:30-11:00

Teaser Talks for all Tuesday SIGMOD Research and Industrial Papers

Location: Effecten

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

Coffee

Location: Grote Zaal

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

SIGMOD Panel on Data Ethics

Location: Effecten

PODS 4

Location: Veiling

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

Lunch

Location: Grote Zaal

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

SIGMOD Research 1

Location: Effecten

SIGMOD Research 2

Location: Graan

SIGMOD Research 3

Location: Berlage

SIGMOD Industry 1: Data Platforms

Location: Administratie

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), Wei-

long 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 Akidau (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), Eu-

gene Koifman (Hortonworks), Owen O'Malley (Hortonworks), Vineet Garg (Hortonworks), Zoltan Haindrich (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 (Hortonworks), 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 Gagねja (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*)

PODS Tutorial 1

Location: Veiling

Making Consistency Protocols Serializable

Alan Fekete (University of Sydney)

Algorithmic Fairness: Measures, Methods and Representations

Suresh Venkatasubramanian (University of Utah)

2019 ACM PODS Alberto O. Mendelzon Test-of-Time Award

*Jianwen Su (University of California, Santa Barbara),
Dirk Van Gucht (Indiana University), Victor Vianu (University of California, San Diego)*

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

Poster & Demo Groups A and B

Location: Grote Zaal

One poster for each SIGMOD and PODS paper presented on Tuesday.

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

Min Xie (Hong Kong University of Science and Technology)

ogy), 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), Phillip 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)

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 Soft-

ware 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), Le-shang 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

Location: Effecten

Student Research Competition (starts 17:20)

Location: Graan

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 Differen-

Potential 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)

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

SIGMOD Reception - sponsored by MonetDB

Location: Van Gogh Museum

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

Coffee + Light Breakfast

Location: Grote Zaal

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

SIGMOD Keynote

Location: Effecten

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

Location: Effecten

PODS 6

Location: Veiling

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

Coffee

Location: Grote Zaal

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

SIGMOD Research 4

Location: Effecten

SIGMOD Research 5

Location: Graan

SIGMOD Research 6

Location: Berlage

SIGMOD Industry 2: Storage and Indexing

Location: Administratie

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)

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)

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)

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)

PODS 7

Location: Veiling

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

Lunch

Location: Grote Zaal

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

SIGMOD Research 7

Location: Effecten

SIGMOD Research 8

Location: Graan

SIGMOD Research 9

Location: Berlage

SIGMOD Research 10

Location: Administratie

PODS Tutorial 2

Location: Veiling

Making Consistency Protocols Serializable

Alan Fekete (University of Sydney)

Algorithmic Fairness: Measures, Methods and Representations

Suresh Venkatasubramanian (University of Utah)

2019 ACM PODS Alberto O. Mendelzon Test-of-Time Award

*Jianwen Su (University of California, Santa Barbara),
Dirk Van Gucht (Indiana University), Victor Vianu (University of California, San Diego)*

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

Poster & Demo Groups B and C

Location: Grote Zaal

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

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)

Pivotal Greenplum[©] for Kubernetes: Demonstra-

tion 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)

*plied 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), Le-
shang Chen (University of Pennsylvania), Paraschos
Koutris (University of Wisconsin, Madison), Arun Ku-
mar (University of California, San Diego)*

Capturing and Querying Structural Provenance in Spark with Pebble

*Ralf Diestelkämper (Universität Stuttgart), Melanie Her-
schel (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), Kon-

stantinos 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 (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 (TUDarmstadt), Carsten Binnig (TUDarmstadt), 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), Jerome 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

Location: 20 boats (Rederij Stromma)

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

SIGMOD Dinner

Location: Noorderlicht Cafe

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

Coffee + Light Breakfast

Location: Grote Zaal

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

Award Talks

Location: Effecten

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

Teaser Talks for all Thursday SIGMOD Research and Industrial Papers

Location: Effecten

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

Coffee

Location: Grote Zaal

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

SIGMOD Research 11

Location: Effecten

SIGMOD Research 12

Location: Graan

SIGMOD Research 13

Location: Berlage

SIGMOD Research 14

Location: Administratie

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

Lunch

Location: Grote Zaal

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

SIGMOD Research 15

Location: Effecten

SIGMOD Research 16

Location: Graan

SIGMOD Research 17

Location: Berlage

SIGMOD Industry 3: Data Applications

Location: Administratie

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)

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

Poster & Demo Groups A and C

Location: 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 Depen-

dencies

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 Model-

ing 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 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), Julianne Freire (New York University), Boris Glavic (Illinois Institute of Technology), Oliver Kennedy (University

of Buffalo), Heiko Meller (New York University), Ruiqi 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 (TUDarmstadt), Carsten Binnig (TUDarmstadt), 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 Sys-

tems)

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

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

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

Coffee + Light Breakfast

Location: Grote Zaal

Friday 07/07/2019 08:30-10:30

Tutorial 4: part 1 (starts 09:00)

Location: Effecten

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 (starts 09:00)

Location: Berlage

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

Location: Administratie

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

Location: Veiling

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

Location: Mendes da Costa

SBD 2019: the Fourth International Workshop on Semantic Big Data

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

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

Tutorial 4: part 2

Location: Effecten

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

Location: Berlage

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

Location: Administratie

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

Location: Veiling

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 2

Location: Mendes da Costa

SBD 2019: the Fourth International Workshop on Semantic Big Data

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

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

Lunch

Location: Grote Zaal

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

Tutorial 5: part 1

Location: Effecten

Data Pipelines for User Group Analytics

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

Tutorial 7

Location: Berlage

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

Location: Administratie

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

Location: Veiling

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

Location: Mendes da Costa

SBD 2019: the Fourth International Workshop on Semantic Big Data

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

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

Coffee + Workshop Posters

Location: Grote Zaal

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

Tutorial 5: part 2 (ends 18:00)

Location: Effecten

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

Location: Administratie

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

Location: Veiling

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

Location: Mendes da Costa

SBD 2019: the Fourth International Workshop on Semantic Big Data

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

Workshops

Monday 08/27/2017 08:30-17:00

FADS

Location: Auditorium 602

Failed Aspirations in Database Systems

Spyros Blanas (Ohio State University), Justin Lewandoski (Microsoft Research), Andy Pavlo (CMU)

BIRTE

Location: Auditorium 606

Eleventh International Workshop on Real-Time Business Intelligence and Analytics

Malu Castellanos (Teradata Aster), Panos K Chrysanthis (University of Pittsburgh)

TPCTC

Location: Auditorium 601

Ninth TPC Technology Conference on Performance Evaluation & Benchmarking

Raghunath Niambar (Cisco), Meikel Poess (Oracle)

VLIoT

Location: Auditorium 670

The International Workshop on Very Large Internet of Things

Sven Groppe (University of Lübeck), Carlo Alberto Boano (Graz University of Technology)

PhD Workshop

Location: Auditorium 1601

VLDB PhD Workshop

Peter Christen (The Australian National University), Bettina Kemme (McGill University), Erhard Rahm (University of Leipzig)

Friday 09/01/2017 08:30-17:00

MATES

Location: Auditorium 602

Workshop on Mobility Analytics for Spatio-temporal and Social Data

Christos Doulkeridis (University of Piraeus), Qiang Qu (Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences)

ADMS

Location: Auditorium 606

Eight International Workshop on Accelerating Analytics and Data Management Systems Using Modern Processor and Storage Architectures

Rajesh Bordawekar (IBM Watson), Tirthankar Lahiri (Oracle)

DMAH

Location: Auditorium 601

The Third International Workshop on Data Management and Analytics for Medicine and Healthcare

Fusheng Wang (Stony Brook University), Gang Luo (University of Washington), Edmon Begoli (Oak Ridge National Laboratory)

DBPL

Location: Auditorium 670

The Sixteenth International Symposium on Database Programming Languages

Tiark Rompf (Purdue University), Alexander Alexandrov (TU Berlin)

BOSS

Location: Auditorium 1601

Third Workshop on Big Data Open Source Systems

Tyson Condie (UCLA), Tilmann Rabl (TU Berlin)

THE WORLD'S FIRST ENGAGEMENT DATABASE

Today's customers aren't just shoppers. They're seekers. That means they spend a lot more time interacting than transacting. **And you have only 8 seconds to grab their attention** before they start looking somewhere else.

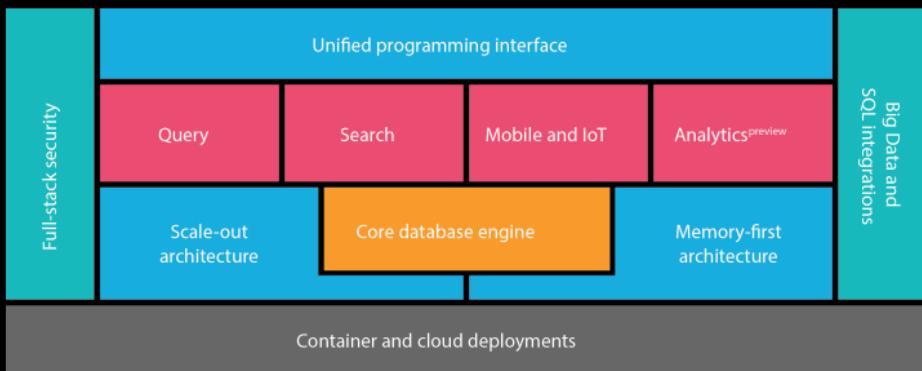
The Couchbase Data Platform makes every interaction count.

To make the most of every moment, you need an **Engagement Database** that delivers an amazing customer experience (CX) every time.

HERE'S WHY

- 1 By 2020, CX will overtake price and product as the key brand differentiator
- 2 Strong CX companies will retain 89% of customers versus 33% for weak CX companies
- 3 70% of Global 500 brands already have dedicated digital experience teams in place

MEET THE COUCHBASE DATA PLATFORM



Start creating amazing customer experiences today.

1-650-417-7500 | www.couchbase.com



Software Engineer, Infrastructure

Google

Software Engineering



Google's software engineers develop the next-generation technologies that change how billions of users connect, explore, and interact with information and one another. Our products need to handle information at massive scale, and extend well beyond web search. We're looking for engineers who bring fresh ideas from all areas, including information retrieval, distributed computing, large-scale system design, networking and data storage, security, artificial intelligence, natural language processing, UI design and mobile; the list goes on and is growing every day. As a software engineer, you will work on a specific project critical to Google's needs with opportunities to switch teams and projects as you and our fast-paced business grow and evolve. We need our engineers to be versatile, display leadership qualities and be enthusiastic to tackle new problems across the full-stack as we continue to push technology forward.

As a Software Engineer working on Google's infrastructure, you have the opportunity to work on everything from the core platform that runs the world's largest distributed network to redefining the systems that allow applications and services to provide useful information to billions of users around the globe. From our Data Center software groups to Google's Cloud Platform, Gmail to YouTube, our infrastructure engineers across departments wrestle with the vast scale of a ubiquitous system, its products, and services and revolutionize industry leading technologies to handle the sheer magnitude at which Google operates.

Google is and always will be an engineering company. We hire people with a broad set of technical skills who are ready to tackle some of technology's greatest challenges and make an impact on millions, if not billions, of users. At Google, engineers not only revolutionize search, they routinely work on massive scalability and storage solutions, large-scale applications and entirely new platforms for developers around the world. From AdWords to Chrome, Android to YouTube, Social to Local, Google engineers are changing the world one technological achievement after another.

Responsibilities

- Build our platforms, systems and infrastructure using your strong background in distributed systems and large scale storage systems.
- Manage individual projects priorities, deadlines and deliverables with your technical expertise.
- Design, develop, test, deploy, maintain, and enhance software solutions.

Qualifications

Minimum qualifications:

- BA/BS degree in Computer Science or related technical field or equivalent practical experience.
- 4 years of relevant work experience, including software development experience, or 1 year of relevant work experience with a PhD in Computer Science or related technical field.
- Professional coding experience in C/C++, Java, Python or Go.
- Experience architecting and developing large scale distributed systems. Experience in concurrency, multithreading and synchronization.

Preferred qualifications:

- MS or PhD in Computer Science.
- Experience with Unix/Linux environments.
- Experience with TCP/IP and network programming.
- Experience with database internals, database language theories, database design, SQL and database programming.
- Understanding of technologies such as virtualization and global infrastructure, load balancing, networking, massive data storage, Hadoop, MapReduce and security.
- Interest or exposure to networking technologies/concepts such as Software Defined Networking (SDN) and OpenFlow.

Microsoft Research

ADVANCE YOUR CAREER



CONNECT WITH US



BOOST YOUR RESEARCH



STAY INFORMED



```
10011001011  
1010010100  
10000110011  
1010110010  
1111011100  
1011011100
```



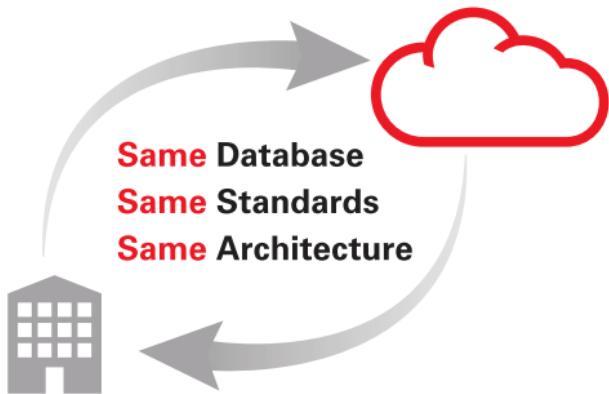
Together, we can achieve more

At Microsoft Research, we're inventing the future of computing. We relentlessly push the boundaries of technology, actively collaborate with world-class researchers, and passionately support the next generation of scientists.

Engage with us: Microsoft.com/research

Push a Button

Move Your Database to the Oracle Cloud



... or Back to Your Data Center

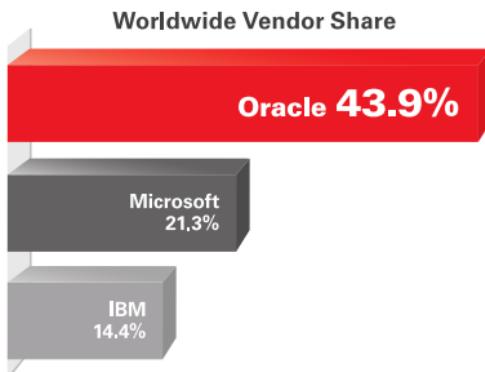
ORACLE®

cloud.oracle.com/database

STILL

#1

Database



Oracle Database

Trusted by 310,000 Customers Worldwide

ORACLE®

oracle.com/database

Source: IDC, "Worldwide Relational Database Management Systems Software Market Shares, 2015: The Year of Transition to the Cloud," IDC #US41484516, June 2016; Table 1 (Worldwide Relational Database Management Systems Revenue by Vendor). Vendor share based on software license and maintenance revenue. Copyright © 2017, Oracle and/or its affiliates. All rights reserved.

Enriching lifestyles with Information Technology

Recruit Institute of Technology (RIT) is the technology hub and research lab for Recruit Holdings, a company that provides over 200 online services in the areas of human resources, travel, housing, education, restaurants and many other areas in which people make daily lifestyle decisions.

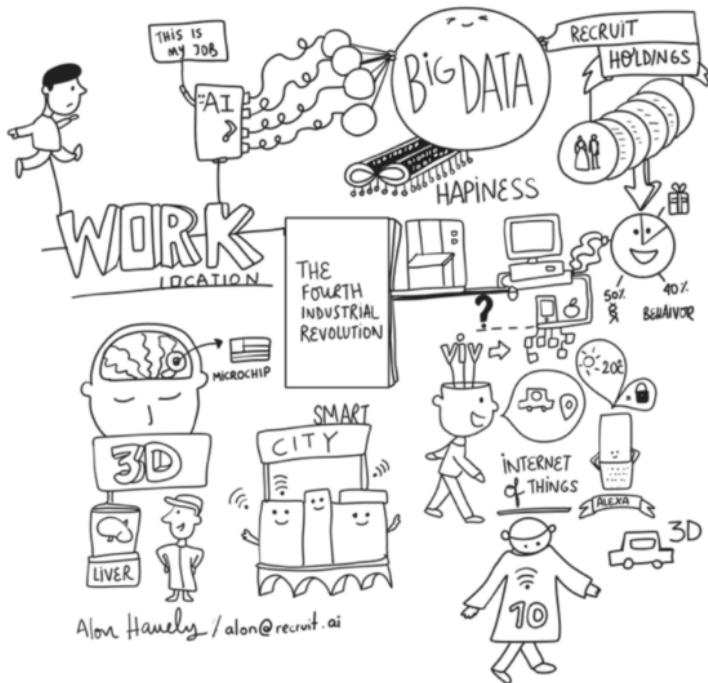
We conduct research in several areas, including data management, data integration, natural language processing, machine learning, and artificial intelligence. We collaborate with universities and publish in top-notch conferences.



Example Project: BigGorilla



BigGorilla is an open-source data integration and data preparation ecosystem (currently in Python) to enable data scientists to perform integration and analysis of data. BigGorilla brings decades of research on data integration into an open-source platform with the goal of accelerating progress in the field and adoption of its techniques. <http://www.biggorilla.org>



*Illustration by Melisa Machurek



At Tableau, innovation is key.

We constantly push the boundaries of Visual Analytics. Tableau is expanding its engineering and R&D power to continue to delight customers with the best analytics platform ever built. You can be part of the team that delivers these breakthroughs.

We are currently hiring:

- Software Engineers in Development
- Software Engineers in Test
- Research Engineers

Visit careers.tableau.com for all open positions.

Haven't had a chance to check out Tableau?

Students & Professors download your free trial today at
www.tableau.com/academic

See yourself at



Jewel Loree

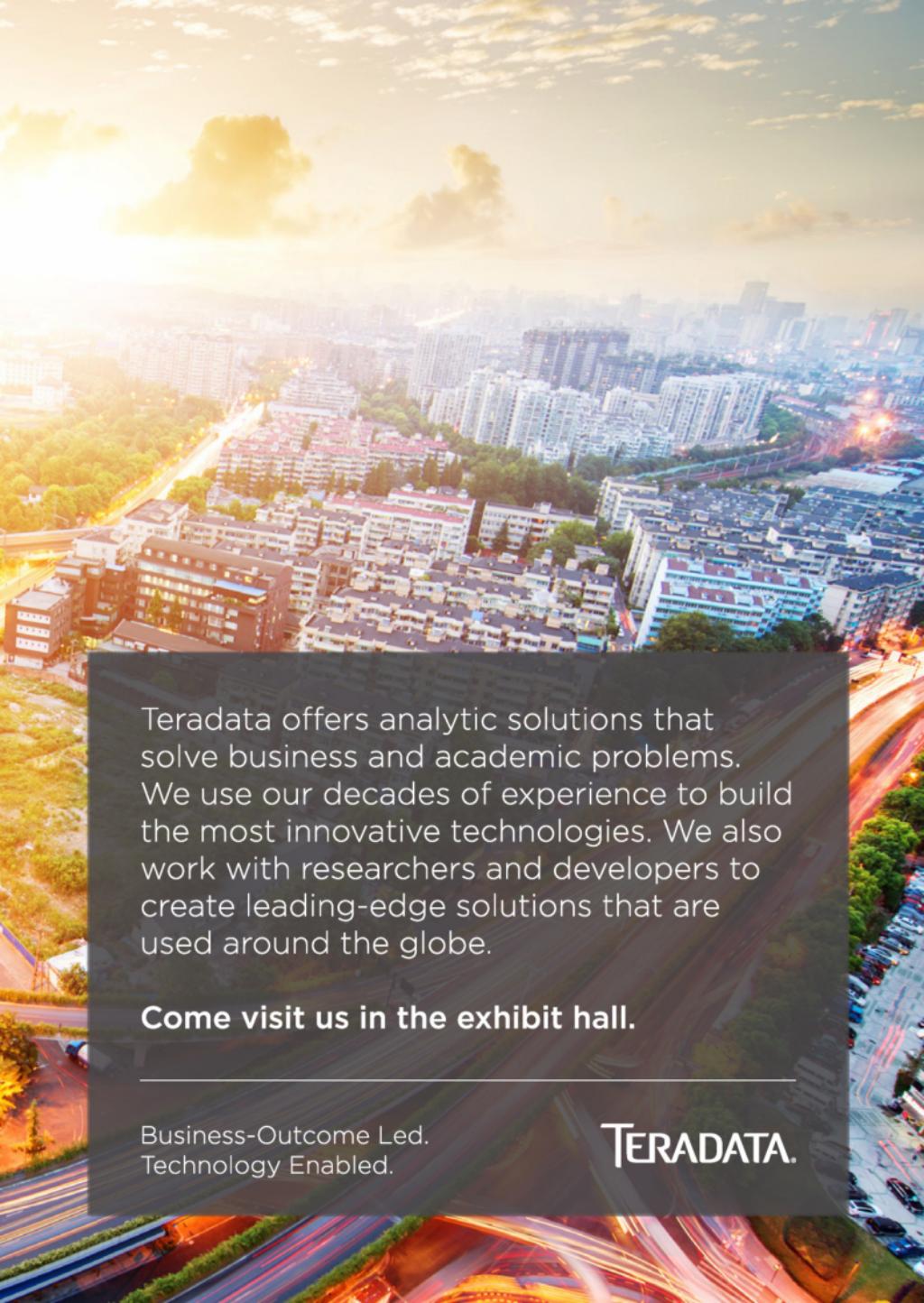
Product Manager, Visual Analytics

Bassist, Indie/Surf band



Advanced Analytics at Enterprise Scale





Teradata offers analytic solutions that solve business and academic problems. We use our decades of experience to build the most innovative technologies. We also work with researchers and developers to create leading-edge solutions that are used around the globe.

Come visit us in the exhibit hall.

Business-Outcome Led.
Technology Enabled.

TERADATA

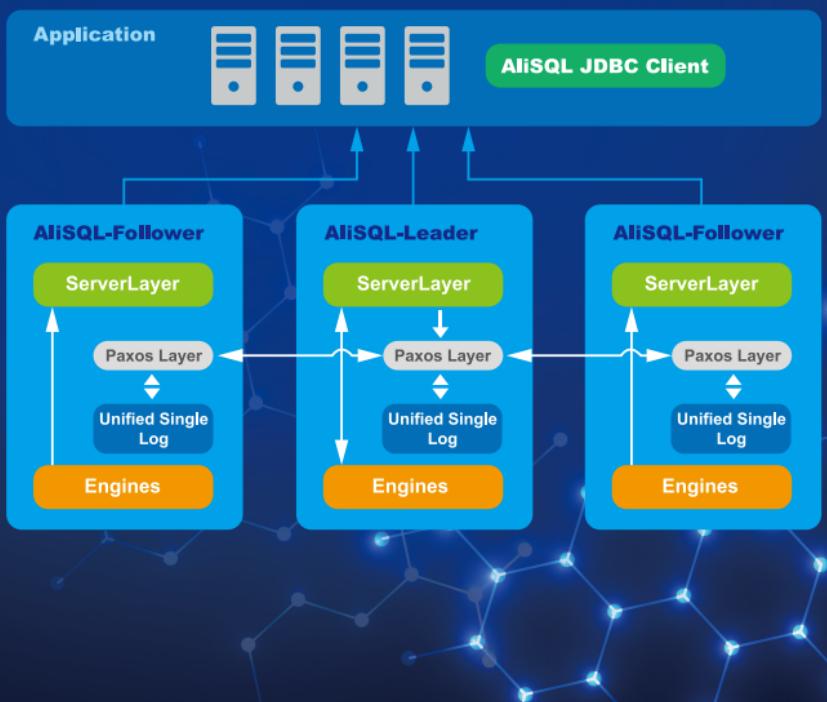
Alibaba Cloud ApsaraDB

ApsaraDB is a Database-as-a-Service platform provided by Alibaba Cloud.

It covers the mainstream database engines including the most popular open-source databases like MySQL (AliSQL), PostgreSQL and Redis, as well as commercial databases like SQL Server and PPAS. It also provides the hybrid analysis database HybridDB and off-line analysis database E-MapReduce, which are compatible with MySQL and PostgreSQL protocols.

AliSQL

AliSQL is a MySQL branch maintained by Alibaba Database team, which supports all lines of business within Alibaba group including Alibaba Cloud, Taobao, Tmall and Ant Financial. We introduced lots of features to AliSQL to support dramatic business growth. Besides providing a fantastic shopping experience in Taobao/Tmall double 11 festival, AliSQL also has more than 100, 000 running RDS instances and 50, 000 users in the cloud. To maximize the outcome we can get from MySQL, we also implemented a high performance Paxos protocol. With the help of Paxos, we can deploy our AliSQL Cluster geo-distributed. Our vision is quite clear that we want to introduce the world the fastest OLTP database at the lowest cost.



Connecting the World

The world's largest social graph

Building the tools & systems to help 1.9 billion people across the world connect, communicate and share requires constant innovation. At Facebook, research permeates everything we do.

The Facebook platform is our lab for research, development, and innovation. Our talented teams of researchers and engineers are constant innovators as they design and build the next generation systems to serve the 1.9 billion people who use our products.

We're hiring!

We're looking for talented industry and academic researchers to join our team!

Visit us at research.fb.com/careers to learn more about a career at Facebook.



Database Systems Expert

Location: Munich Germany

Huawei is a leading global information and communications technology solutions provider.

The European Research Institute (ERI) performs strategic research and cutting-edge development for Huawei.

The Databases group in ERI is advancing database technology for future use in telecommunication and enterprise and is looking for a few technical experts that will take part in that awarding endeavor. It operates out of sites in

Germany and Israel.

What you can expect:

- Research towards conception followed by architecture, design, prototyping, and development.
- Analysis and evaluation of relevant state of the art in the academia and industry, and competitors' products.
- Performing business analysis and technical risk evaluation of proposed designs and plans.

Interested? You are invited to contact the recruiting manager directly

Eliezer Levy, tel. +972 54 2277128
eliezer.levy@huawei.com

HUAWEI TECHNOLOGIES
Duesseldorf GmbH
German Research Center, Munich
Riesstraße 25
80992 Munich, Germany



your future made with IBM

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 Cognitive Computing. 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 and discover what you can make of this moment.

What will you make with IBM?

ibm.com/jobs



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 here:



hana.sap.com

AEROSPIKE

THE HIGH PERFORMANCE, NOSQL
DATABASE FOR REAL-TIME, MISSION
CRITICAL APPLICATIONS.

- * Predictable performance
- * Highest up time & availability
- * Lowest Total Cost of Ownership

www.aerospike.com



EXASOL

The world's fastest*
in-memory database for
large-scale data analytics.

*For over 3 years, we are still unbeaten in the TPC-H
benchmarks. No-one else comes close. Period.

Get started with EXASOL today
wwwexasolcom/vldb



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.



www.MongoDB.com/Careers



@MongoDB



www.Facebook.com/MongoDB



@MongoDB + @MongoDBCareers



Persistent Systems
Delivering the Blueprint
for a Software Driven Business

Join the team that obsessively focuses on the "**how**" of digital.

www.persistent.com