**Seminar in Databases 2019:** Cloud-based systems, Machine Learning Applications to Databases, and current topics: Provenance, Graphs, Vectors, Word Embedding, Consensus

**TOPICS**

**By lecturer:**
1. Concurrency Control - Chapter 2 BHG: Serializability Theory
2. Concurrency Control - Chapter 3 BHG: Two-Phase Locking
3. Recovery in Distributed Databases - Chapter 19 WV: Distributed Recovery

**By guest lecturers**
1. Blockchain technologies
2. The modern cloud

**Main Topics**

**Cloud**
(2 presenters) Amazon Aurora: On Avoiding Distributed Consensus for I/Os, Commits, and Membership Changes.
Alexandre Verbitski, Anurag Gupta, Debanjan Saha, James Corey, Kamal Gupta, Murali Brahmadessam, Raman Mittal, Sailesh Krishnamurthy, Sandor Maurice, Tengiz Kharatishvili, Xiaofeng Bao.
*SIGMOD Conference 2018*: 789-796.

Serverless Computing: One Step Forward, Two Steps Back.
Joseph M Hellerstein (UC Berkeley); Jose Faleiro (UC Berkeley); Joseph Gonzalez (UC Berkeley); Johann Schleier-Smith (UC Berkeley); Vikram Sreekanti (UC Berkeley); Alexey Tumanov (UC Berkeley); Chenggang Wu (UC Berkeley).
*CIDR 2019*.

Veritas: Shared Verifiable Databases and Tables in the Cloud.
Johannes Gehrke (Microsoft); Lindsay Allen (Microsoft); Panagiotis Antonopoulos (Microsoft); Arvind Arasu (Microsoft); Joachim Hammer (Microsoft); James Hunter (Microsoft Research); Raghav Kaushik (Microsoft); Donald Kossmann (Microsoft Research); Ravi Ramamurthy (MICROSOFT); Srinath Setty (Microsoft); Jakub Szynaszek (Microsoft); Alexander van Renen (Technical University of Munich); Jonathan Lee (Microsoft Research); Ramarathnam Venkatesan (Microsoft).
*CIDR 2019*.

The Kubernetes book, Nigel Poulton, 2 talks:
   a. Chapters 1,2,3,4 (including a demo on MiniKube).
   b. Chapters 5,6,7 (including a demo on MiniKube).

The Case for Network-Accelerated Query Processing.
Alberto Lerner, Rana Hussein, Philippe Cudre-Mauroux.
*CIDR 2019*.
**Provenance**

Provenance in Databases: Why, How, and Where

Two talks:
- Where-Provenance, Comparing Models of provenance.

Provenance semirings.
Todd J. Green, Grigoris Karvounarakis, Val Tannen.

Circuits for Datalog Provenance.
Daniel Deutch, Tova Milo, Sudeepa Roy, Val Tannen.
*ICDT 2014: 201-212.*

Provenance for Natural Language Queries.
Daniel Deutch, Nave Frost, Amir Gilad.

Two papers (one presenter):
1. Provenance and Probabilities in Relational Databases.
   Pierre Senellart.
   *SIGMOD Record 46(4): 5-15 (2017).*
2. ProvSQL: Provenance and Probability Management in PostgreSQL.
   Pierre Senellart, Louis Jachiet, Silviu Maniu, Yann Ramusat.
   *PVLDB 11(12): 2034-2037 (2018).*

**Machine Learning Applications to Databases**

Deep Learning for Entity Matching: A Design Space Exploration.
Sidharth Mudgal, Han Li, Theodoros Rekatsinas, AnHai Doan, Youngchoon Park, Ganesh Krishnan, Rohit Deep, Esteban Arcaute, Vijay Raghavendra.
*SIGMOD Conference 2018: 19-34.*

Fine-grained Concept Linking using Neural Networks in Healthcare. 51-66.
Jian Dai, Meihui Zhang, Gang Chen, Ju Fan, Kee Yuan Ngiam, Beng Chin Ooi.
*SIGMOD Conference 2018: 51-66.*

Two papers (2 presenters):
1. The Case for Learned Index Structures.
   Tim Kraska, Alex Beutel, Ed H. Chi, Jeffrey Dean, Neoklis Polyzotis.
   *SIGMOD Conference 2018: 489-504.*

2. SageDB: A Learned Database System.
   Tim Kraska, Mohammad Alizadeh, Alex Beutel, Ed H. Chi, Jialin Ding,
Ani Kristo, Guillaume Leclerc, Samuel Madden, Hongzi Mao, Vikram Nathan.  
**CIDR 2019.**

Query-based Workload Forecasting for Self-Driving Database Management Systems. 
Lin Ma, Dana Van Aken, Ahmed Hefny, Gustavo Mezerhane, Andrew Pavlo, Geoffrey J. Gordon.  
**SIGMOD Conference 2018:** 631-645.

Sketching Linear Classifiers over Data Streams. 
Kai Sheng Tai, Vatsal Sharan, Peter Bailis, Gregory Valiant.  
**SIGMOD Conference 2018:** 757-772.

ZigZag: Supporting Similarity Queries on Vector Space Models. 
Wenhai Li, Lingfeng Deng, Yang Li, Chen Li.  
**SIGMOD Conference 2018:** 873-888.

**Two papers (one presenter):**

Deep Reinforcement Learning for Join Order Enumeration. aiDM@SIGMOD 2018: 3:1-3:4
Both by Ryan Marcus, Olga Papaemmanouil.

**Two papers (two presenters):**
1. End-to-End Machine Learning with Apache AsterixDB. Wail Alkowailleet et. al.  
DEEM@SIGMOD 2018: 6:1-6:10.

**CIDR 2019.**
Graphs, Vectors, Word Embedding
Maverick: Discovering Exceptional Facts from Knowledge Graphs. 1317-1332.
Gensheng Zhang, Damian Jimenez, Chengkai Li. SIGMOD Conference 2018: 1317-1332.

Embedding Logical Queries on Knowledge Graphs.

Cypher: An Evolving Query Language for Property Graphs. 1433-1445.

AND
Formal Semantics of the Language Cypher.

Quantum-inspired Complex Word Embedding.
Qiuchi Li, Sagar Uprety, Benyou Wang, Dawei Song. Rep4NLP@ACL 2018: 50-57.

Word Vectors and Quantum Logic, Experiments with negation and disjunction. (DEMO?)

Articles covered by THE MORNING PAPER https://blog.acolyer.org/2018/07/
Each day an article is reviewed on this site (you may also subscribe to receive notifications).
Choose an article covered by THE MORNING PAPER that deal with this seminar’s topics, that can be the basis of a one hour presentation, and that was reviewed on or after Jan. 1, 2017.

Here are interesting ones:

