

**Seminar:  
Database Systems  
Algorithms for Database Systems**

**Topic 2019:  
Algorithms and Systems for Data Science**

Michael Böhlen  
Sven Helmer  
Paolo Penna

19. Februar 2019

# Agenda

- ▶ Goals of the Seminar
- ▶ Logistics
- ▶ Topics
- ▶ Next Steps
- ▶ Grading

# Goals of the Seminar

- ▶ Get a deep understanding of an advanced topic/paper in algorithms for database systems
- ▶ Focus on one technical paper/topic (**not** an overview of ten papers!)
- ▶ Give a talk on this paper, write a manuscript, lead the discussion
  - ▶ Present a result
  - ▶ Do not present a paper
- ▶ Act as a buddy to another student (read paper, read report, give feedback)

# Topics

1. Architectures and Systems
2. Column Stores
3. Streams
4. Spark
5. Query Processing
6. Clustering

- ▶ VLDB
- ▶ SIGMOD
- ▶ ICDE
- ▶ KDD
- ▶ NIPS
- ▶ OSDI
- ▶ SOSP

# Logistics

- ▶ Prerequisites:
  - ▶ Good understanding of database systems and of algorithms
  - ▶ Enjoy algorithms and database systems
  - ▶ Available on seminar days (SA April 13 and SA May 11)
  
- ▶ Maximum number of participants:
  - ▶ The total number of seminar participants is 18
  - ▶ Please decide now if you want to participate or not
  
- ▶ Paper selection:
  - ▶ We use RSD to assign 1) topics and 2) buddies

## Next steps

- ▶ Read your paper (and related papers)
- ▶ Write manuscript (4 - 8 pages) on how would you present it: motivation, formal model, interesting proofs, main results
- ▶ Send your manuscript to buddy + the advisor 3 weeks before your talk
- ▶ Meet with the advisor and buddy 2 weeks before the talk
- ▶ Submit final version of manuscript 1 week before the talk
- ▶ Prepare talk, rehearse with buddy
- ▶ Talks: blackboard talk (30min, no slides) + lead discussion (10min)

# Grading

- ▶ In order of importance:
  - ▶ Presentation
  - ▶ Manuscript
  - ▶ Buddy
  - ▶ Active Participation