

Introduction to to the course

Data Management for Big Data, Part II:  
Advanced database models, languages, and  
systems

Dario Della Monica

April 1, 2019

# Teacher's information

Dario Della Monica

Dipartimento di Scienze Matematiche, Informatiche e Fisiche (DMIF)  
University of Udine

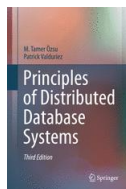
- ▶ **email:** `dario.dellamonica@uniud.it`
- ▶ **web:** `http://users.dimi.uniud.it/~dario.dellamonica/`
  - ▶ [Home] >> Teaching >> Data Management for Big Data
- ▶ **office:** room SN16, 2nd floor, stecca nord  
Polo Scientifico "Rizzi"  
via delle Scienze, 206  
33100 Udine
- ▶ **office hours:** Wednesday, 10:30-12:30 (or by appointment)
- ▶ more info (**cell**, **skype**, ...) on my webpage

# Material

- Database System Concepts  
Silberschatz, Korth, Sudarshan  
6th Edition  
McGraw-Hill, 2011  
<https://www.db-book.com/db6/index.html>  
(extra contents: original slides, solutions, etc)



- Principles of Distributed Database Systems  
Özsu and Valduriez  
3rd Edition  
Springer, 2011  
ISBN: 978-1-4419-8834-8  
<http://extras.springer.com/>  
(original slides)



- other material (e.g., slides) is available on my webpage

# Overview of this part of the course

Data Management for Big Data

Part II: Advanced database models, languages, and systems

3 cfu (24 hours)

---

- ▶ Query processing and optimization
  - ▶ Query processing
  - ▶ Algorithms for the join operation
  - ▶ Cost-based optimization and heuristics
- ▶ Distributed database architectures
  - ▶ An introduction to distributed DBMS
  - ▶ Design of distributed databases (fragmentation and allocation)
  - ▶ Distributed query processing
  - ▶ Transaction processing in distributed databases
    - ▶ the two-phase commit (2PC) protocol
- ▶ Semistructured Data and XML
  - ▶ Definition of semistructured data in XML
  - ▶ Querying XML data (XPath and XQuery)
  - ▶ XML and relational DBMS
  - ▶ Native XML databases
- ▶ Cloud computing and DBMS, Parallel DBMS