The Landscape of Databases or What We Did **Not** Cover

Lecture 15

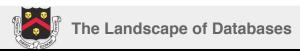
3 April 2015

Broad Areas

- Using an RDBMS
- Conceptual/Logical design
- Physical design
- How an RDBMS works
- Non-Relational models
- Scaling
- Data analytics

Using an RDBMS

- DBMS diversity
 - DB2, Oracle
- Advanced database programming
 - Views
 - Stored procedures, triggers
 - Cursors, ORM
 - Isolation levels
 - Data formats: CSV, XML, JSON
 - [Web] APIs
- Reporting
 - Analytics



Conceptual/Logical Design

- ERDs
 - Multi-way relationships
 - Other dialects

- Normal forms
 - BCNF, 4NF, 5NF, 6NF

Physical Design

- Storage
 - Cost models
 - File organization
 - Buffering

- Indexing, details of ...
 - Hashing
 - Trees
 - Specialized



How an RDBMS Works

- Relational algebra & calculus
- Query evaluation/optimization
 - Join processing, sorting
- Transaction management
 - Concurrent schedules
 - Locking
- Logging & recovery
 - Write-Ahead (WAL), checkpointing

3 April 2015

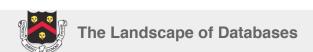
Non-Relational Models

- XML, Object-Relational
- Spatial
- Graphs/networks
- Events, streams
- Documents
- Key-value stores
- Probabilistic, deductive

3 April 2015 7

Scaling

- Distributed Databases
 - OLAP
- Cloud computing
 - MapReduce
 - Horizontal vs. vertical scaling
- Crowdsourcing
 - Mechanical Turk
 - Human computation (e.g. reCAPTCHA, Duolingo)
- Security
 - Encryption vs. functionality, monetization
 - Differential privacy
 - Regulation (e.g. HIPAA), provenance



Data Analytics

- Data Mining/Science
- Ranking (e.g. PageRank)
- Data Warehouses
 - Data cube
- Natural Language Processing (NLP)
 - Entity matching
- Human-Computer Interaction (HCI)
 - Visualization
 - [Facilitated] exploration
 - Explainable queries



3 April 2015