Structure

• Structure
  - Thursday May 5, 1:30-3:20PM
  - 7-10 pages of short answer, essay questions
  - Be precise, be verbose
  - Comprehensive
  - Closed book, notes

• Readings
  - Lecture notes
  - Codd, Spanner, Dremel, NoDB, and Beckman report papers
  - ElMasri and Navathe (relevant chapters are on lecture notes page)
Query Languages

• Domain relational calculus, tuple relational calculus, relational algebra, Datalog, Cypher, Pig Latin, XQuery, OQL

• Be able to write queries in
  - OQL
  - XQuery
  - Pig Latin
  - Cypher
  - Datalog
Transaction Protocols

• Serial vs. serializable

• Protocols
  ■ Two-phase locking
  ■ Timestamp ordering
  ■ Multiversioning
  ■ Optimistic CC

• Strict vs. non-strict

• Cascading aborts/rollback
Recovery

- More practical than theory
- Reasons for failure
- Recovery log
- Recovery strategies
  - Undo/redo
  - Undo/no-redo
- Checkpoint
Object-oriented

- Need for OO in DBMS
  - OODBMS vs. ORDBMS
- Persistence and extents
  - OIDs
- SQL extensions
Data Mining and IR

• Need for

• Association rules
  - Support and confidence

• Information retrieval
  - Precision and recall
  - Search engine techniques
Big Data & OLAP

- Map/Reduce Pig Latin architecture
- Graph Databases
- Column store DBs
- Star vs. snowflake schema