

# SI440: Database Systems

---

Dr. Adina Crainiceanu

[www.cs.usna.edu/~adina/teaching/si440/fall2007](http://www.cs.usna.edu/~adina/teaching/si440/fall2007)

Database Management Systems, R. Ramakrishnan and Johannes Gehrke

## Outline

---

- Class Survey
- Why Databases (DB)?
  - A Problem
  - DB Benefits
- This Course?
- Admin

Database Management Systems, R. Ramakrishnan and Johannes Gehrke

## SI440: Database Systems

---

- **How does Wal-Mart manage its 200 TB data warehouse?**
- **What is the database technology behind ebay's website?**
- **How do you build an Oracle 9i, MySQL or Microsoft SQL Server database?**
- **How do you build a search engine?**

Database Management Systems, R. Ramakrishnan and Johannes Gehrke

## Database Management Systems

---

- Information is one of the most valuable resources in this information age
- How do we effectively and efficiently manage this information?
  - Relational database management systems
    - Dominant data management paradigm today
  - Search engines
  - 10+ billion dollar a year industry
    - You will see this in the job market!

Database Management Systems, R. Ramakrishnan and Johannes Gehrke

## ICE: The Mid Store

---

- Create a system to keep track of inventory

## Problems

---

- Changes to data - Data model
- “on the fly” queries
- Data inconsistencies
- Security of information (views)
- Performance
- Concurrency
- Computer crash

## What is a Database?

---

- A very large, integrated collection of data
- Models real-world *enterprise*.
  - Entities (e.g., students, courses)
  - Relationships
- A *Database Management System (DBMS)* is a software package designed to store and manage databases.

## Why Use a DBMS?

---

- Data independence and efficient access
- Reduced application development time
- Data integrity and security
- Performance and scalability
- Concurrent data access
- Recovery from system crashes

## Why Study Databases?

- Used everywhere
  - Universities (MIDS), military, enterprises
- Datasets increasing in diversity and volume.
  - Digital libraries, interactive video, Facebook, YouTube, Google
  - ... need for DBMS exploding
- DBMS encompasses most of CS
  - OS, languages, theory, data mining, multimedia, logic

Database Management Systems, R. Ramakrishnan and Johannes Gehrke

## Best Jobs!

**CNNMoney.com™**

HOME NEWS MARKETS TECHNOLOGY JOBS & ECONOMY PERSONAL FINANCE AUTO

**BEST JOBS IN AMERICA** Money salary.com™

MONEY Magazine and Salary.com rate careers on salary and job prospects.  
By Tara Kalwarzki, Daphne Mosher, Janet Paskin and Donna Rosato

**Top 10 best jobs**  
MONEY Magazine and Salary.com researched hundreds of jobs, considering their growth, pay, stress-levels and other factors. These careers ranked highest. (more)

1. Software Engineer
2. College professor
3. Financial adviser
4. Human Resources Manager
5. Physician assistant
6. Market research analyst
7. Computer IT analyst
8. Real Estate Appraiser
9. Pharmacist
10. Psychologist

Database Management Systems, R. Ramakrishnan and Johannes Gehrke

## IT Analyst

http://money.cnn.com - MONEY Magazine: Best Jobs in America - Microsoft Internet Explorer

**Money.com** BEST JOBS IN AMERICA salary.com™

PREVIOUS NEXT Back to feature

**No. 7: COMPUTER/IT ANALYST**

AVERAGE SALARY: **\$83,500**

10-YEAR GROWTH: **36%**

AVERAGE ANNUAL JOB OPENINGS: **67,300**

**C C D B**  
STRESS FLEXIBILITY CREATIVITY EASE OF ENTRY

**7. Computer IT analyst**

**Why it's great** Seems like the entire world is at the mercy of information technology folks, thanks to the rapid spread of computers and swell of the Internet. And all of these jobs pay well, from desktop support technician to Webmaster to database work.

Entry-level analysts make \$60,000 and above. Senior database specialists and IT managers command six-figure salaries and decent bonuses. A Bachelor's degree is enough to get started.

**What's cool** Telecommuting and freelance gigs abound. Plus: e-mail snooping!

Database Management Systems, R. Ramakrishnan and Johannes Gehrke

## Outline

- Class Survey
- Why Databases (DB)?
  - A Problem
  - DB Benefits
- **This Course?**
- Admin

Database Management Systems, R. Ramakrishnan and Johannes Gehrke

## Prerequisites

---

- SI321 Advanced Data Structures,
- Strong programming skills (C++)

Database Management Systems, R. Ramakrishnan and Johannes Gehrke

## Course Grading

---

- Three components
  - Assignments, quizzes (20%)
  - Projects (25%)
  - Exams (55%)

Database Management Systems, R. Ramakrishnan and Johannes Gehrke

## Class Lectures

---

- Textbook: “Database Management Systems”
  - By Raghu Ramakrishnan and Johannes Gehrke
  - Required textbook
- Syllabus
  - Defined by class lectures
  - Not defined by textbook

Database Management Systems, R. Ramakrishnan and Johannes Gehrke

## Course Topics

---

- Database design
- Relational model
- Relational algebra
- SQL
- Indexing
- Query evaluation and optimization
- Transaction processing: concurrency, crash recovery
- Normalization
- Database security

Database Management Systems, R. Ramakrishnan and Johannes Gehrke

## Course Goals

---

- Query relational databases using SQL;
- Design, and create relational databases to satisfy user requirements;
- Build parts of a real database management system;
- Explain the main functionality provided by modern database management systems: query optimization, concurrency control, crash recovery;
- Analyze the ethical issues and responsibilities related to records management.

How to BUILD a Database Management System

Database Management Systems, R. Ramakrishnan and Johannes Gehrke

## Things We Will NOT Cover

---

- Database applications
- Web interface
- PHP

Create applications that USE a Database Management System

Database Management Systems, R. Ramakrishnan and Johannes Gehrke

## Success in SI440

---

- Lecture – stay engaged
  - Take notes – provided slides are not enough!
  - Exams closed-book – but open-note!
  - Ask & answer questions
- Complete assignments / projects
  - Think before you start writing /typing
  - Don't stay stuck!
- Don't fall behind
  - Finish assignments early and leave time for reading
  - See me for help and/or talk to friends
  - Course material builds on itself and gets more complex

Database Management Systems, R. Ramakrishnan and Johannes Gehrke

## Academic Integrity - Honor

---

- Collaboration on hws is allowed, but submitted work should be your own
  - Cite any assistance, from any sources
- Collaboration on exams/quizzes is prohibited
- Projects will be done in groups. Collaboration between groups is prohibited.
- <http://www.cs.usna.edu/academics/honor.htm>

Database Management Systems, R. Ramakrishnan and Johannes Gehrke

## Resources

---

- Lecture slides / your notes
- Textbook: Database Management Systems by R. Ramakrishnan and J. Gehrke

Database Management Systems, R. Ramakrishnan and Johannes Gehrke

## Is SI440 a lot of work?

---

- It depends!
  - Much of the material is probably new to you
  - There is at least one substantial programming project
- Then why on earth should I take this course?
  - Intellectual argument
    - Big conceptual ideas
    - Meeting of theory and practice
  - Utilitarian argument
    - Many, many real applications
    - Job market!

Database Management Systems, R. Ramakrishnan and Johannes Gehrke

## Instructor

---

- Asst. Prof. Adina Crainiceanu
- Ph.D. Cornell University
- Area of Specialization: Databases
- Research: Search in Peer-to-Peer Systems
  - Lots of military applications
  - Internship possibilities for students

Database Management Systems, R. Ramakrishnan and Johannes Gehrke