

# CPSC 504: Data Management

---

Rachel Pottinger  
Course Introduction  
2015/01/06

# What is this class about?

---

- Understanding how people design database management systems
- Data management research
- Managing your data

# What is data management research?

---

- Research about managing data including:
  - Traditional (relational) database management systems
    - What they are, how to make them work
  - Other kinds of databases
    - Object-oriented, XML, No SQL
  - Other data management applications
    - OLAP, data mining, etc.

# This class is a seminar

---

A **seminar** is “a small group of advanced students ... under the guidance of a professor who meets regularly with them to discuss...” [dictionary.com]

- 1 or 2 papers to read for most classes. I'll provide:
  - An explanation why we're reading the paper
  - Necessary background beforehand
  - Suggestions on how to read papers where necessary
- Most days students present papers and lead discussions. You'll present once and lead discussion once. I'll provide:
  - The high level goals of reading papers
  - A set of suggested discussion questions
  - Feedback on your plans and answers to questions
  - Possibly a preliminary suggested set of slides
- Sometimes I'll present the papers and lead discussion

# But I haven't taken any database classes!

---

- You do not need to have taken a database class
- Assuming you have a solid basis in CS, I will provide you with all background material you need

# What *are* the prerequisites for this class?

---

- A background in computer science
- Ability to read and respond to 1 – 2 papers a class
- Ability to do a project (not necessarily implementation based) either in a group or on your own
- Ability and willingness to present papers and lead discussions
- Willingness to discuss your own ideas and questions in class

Other handy things: databases, AI, OS

# Where does my grade come from?

---

- Analyzing the readings – 30%
  - Post a summary/analysis on Connect
- Presenting/leading class discussions – 20%
  - One person presents the content
  - One person leads discussion
  - You'll sign up for *different* days for this
- Course project – 30%
  - See website
  - Doesn't have to have an implementation
- One homework – 5%
- In-class participation – 15%

# Any administrative questions?

---



# Introduce your partner

---

Discuss the answers to the following questions with your partner for next 5 minutes:

- What is your name?
- Where are you from?
- What is your database/data management background?
- What do you want to get out of this class?

Grab a card and a marker and write your name on the card

# Some reasons to use a database:

---

- Large amounts of data
- Structured data
- Persistent data
- Valuable data
- Performance requirements
- Concurrent access to data
- Restricted access to data

# What data is stored in databases?

---

This space intentionally left blank

# What data do you have?

---

Here is some data / have:

- Papers I've read
- Addresses
- Job search data
- Experiments I've run
- Grades
- CDs, DVDs, and books I own
- Powerpoint slides
- Research notes
- E-mail
- Drafts of research papers and notes from students

# Class outline

---

- Crash course in databases
- Standard Relational databases
- New Relational Databases
- Other data models
- Management of other data
- Advanced topics/Student request potpourri
- Note that I may have to leave town for another week in the end of March – if this happens, we'll juggle the schedule

# To do:

---

Course website:

<http://www.cs.ubc.ca/~rap/teaching/504/2014>

- Mailing list: mail [majordomo@cs.ubc.ca](mailto:majordomo@cs.ubc.ca) with “subscribe cpsc504” in the body
- Check out Connect (see course webpage)
- Think more about data you have
- Think about which topics you'd like to present/lead discussion on (first come, first served)
- Read the project description, and think about projects
- If there are any topics you'd like to see covered that aren't, let me know