



DISTINGUISHED
LECTURE
SERIES 2018

Deep Learning to Learn

SEPT. 10, 2018 (MONDAY)

11:30AM-12:30PM

DAVIS AUDITORIUM
(412 CEPSR)



Pieter Abbeel (UC Berkeley)

CS 4111-Introduction to Databases, EugeneWu

1

W4111
Introduction to Databases
Fall 2018

Computer Science Department
Columbia University

CS 4111-Introduction to Databases, EugeneWu

2

Data

CS 4111-Introduction to Databases, EugeneWu

3

Data

is for serious business

CS 4111-Introduction to Databases, EugeneWu

4

Data

is at the center of most things.

CS 4111-Introduction to Databases, EugeneWu

5

Data

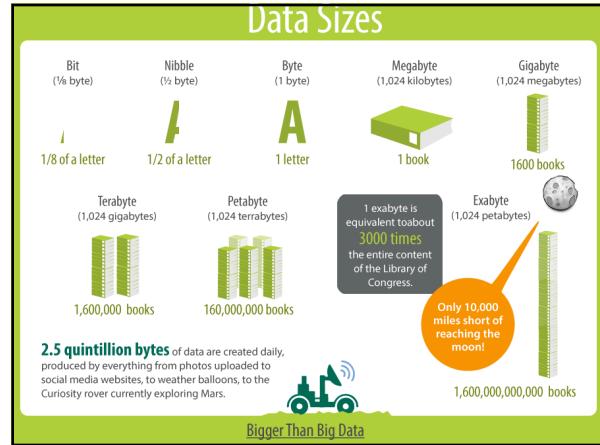
is at the center of *everything*

CS 4111-Introduction to Databases, EugeneWu

6

CS 4111-Introduction to Databases, EugeneWu

7



► THE PAST

Digital storage grew annually by 23% between 1986 and 2007.

Most data was stored on **videotapes** such as VHS cassettes in the pre-digital revolution world of the late 1980s. Vinyl LP records, audio cassette tapes, and photography accounted for significant portions as well.

Paper-based storage represented 33% of all data storage on its own in 1986.

25% of all data stored in the world in 2000 was stored digitally.

2002 is the first year that digital storage capacity overtook analog capacity.

94% of all data was stored in digital format by 2007.

► PRESENT

Today, more than **2.5 exabytes** (2.5 billion gigabytes) of data is generated every single day. This is expected to continue growing at a significant rate with mobile devices accounting for much of this data.

Some experts have estimated that **90%** of all of the data the world today was produced within the last two years.

How did we get here?

CS 4111-Introduction to Databases, EugeneWu

10

Data was Manual

CS 4111-Intro

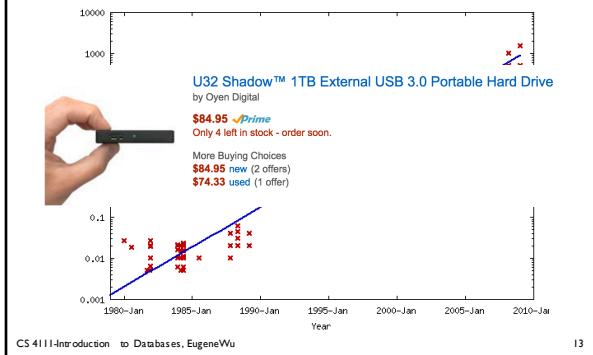
11

Data was Expensive

CS 4111-Introduction to Databases, EugeneWu

12

Data is Cheap



Data is Automated

Physical devices



CS 4111-Introduction to Databases, EugeneWu

14

Data is Automated

Physical devices
Software logs

CS 4111-Introduction to Databases, EugeneWu

15

Data is Ubiquitous

Physical devices
Software logs
Phones



CS 4111-Introduction to Databases, EugeneWu

16

Data is Ubiquitous

Physical devices
Software logs
Phones
GPS/Cars



CS 4111-Introduction to Databases, EugeneWu

17

Data is Everywhere

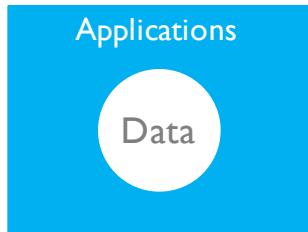
Physical devices
Software logs
Phones
GPS/Cars
Internet of Things



CS 4111-Introduction to Databases, EugeneWu

18

What Applications?



CS 4111-Introduction to Databases, EugeneWu

19

What are we doing with data?

Health



CS 4111-Introduction to Databases, EugeneWu

20

What are we doing with data?

Health



CS 4111-Introduction to Databases, EugeneWu

What are we doing with data?

Health

Investigative Journalism

PROPUBLICA | Patient Safety

Surgeon Scorecard

by Siri Wei, Olga Pierce and Marshall Allen, ProPublica, Updated July 15, 2015
Guided by experts, ProPublica calculated death and complication rates for surgeons performing one of eight elective procedures in Medicare, carefully adjusting for differences in patient health, age and hospital quality. Use this database to know more about a surgeon before your operation.

CS 4111-Introduction to Databases, EugeneWu

22

What are we doing with data?

Health
Investigative Journalism
Recommendations



CS 4111-Introduction to Databases, EugeneWu

24

What are we doing with data?

MACHINE BIAS

Besieged Facebook Says New Ad Limits Aren't Response to Lawsuits

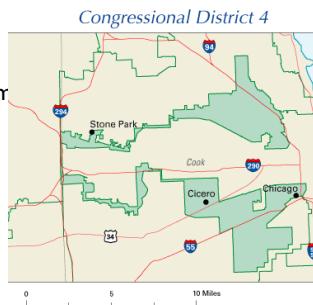
The social network is removing 5,000 options that regulators say enable advertisers to discriminate.

by Ariana Tobin and Jeremy B. Merrill, Aug. 23, 12:48 p.m. EDT



What are we doing with data?

Health
Investigative Journalism
Recommendations
Politics



CS 4111-Introduction to Databases, EugeneWu

26

What are we doing with data?

Health
Investigative Journalism
Recommendations
Politics

A screenshot of the TIME magazine website. The main headline is "Inside the Secret World of the Data Crunchers Who Helped Obama Win". Below it, there's a section for "The Cambridge Analytica Files" with a subtitle "A year-long investigation into Facebook, data, and influencing elections in the digital age". Other menu items visible include News, Opinion, Sport, Culture, Lifestyle, and More.

Key stories

Hide

What are we doing with data?

Health
Investigative Journalism

Every day, the NSA intercepts and stores 1.7 billion emails, phone calls, texts, and other electronic

parallels MANY STORIES, ONE WORLD



Facial Recognition In China Is Big Business As Local Governments Boost Surveillance

April 3, 2018 - 10:40 AM ET

ROB SCHMITZ

QUEUE

DOWNLOAD

EMBED

CS 4111-Introduction to Databases, EugeneWu

29

What are we doing with data?

Health
Investigative Journalism
Recommendations
Politics
Surveillance
Identity

A screenshot of the Harvard Business School website. The main headline is "India's Ambitious National Identification Program". Below it, there's a quote in red: "YOU ARE BASICALLY DENIED ALMOST EVERYTHING IF YOU CAN'T PROVE WHO YOU ARE.". A sidebar on the right provides more details about the Unique Identification Authority of India's program.

Comments: 30 Email Print Download Share Recommended Share (2)

What data?

Applications

Data

CS 4111-Introduction to Databases, EugeneWu

30

What data?

Fake data



CS 4111-Introduction to Databases, EugeneWu

31

What data?

Fake data
Biased data

THE STANFORD OPEN POLICING PROJECT

On a typical day in the United States, police officers make more than 50,000 traffic stops. Our team is gathering, analyzing, and releasing records from millions of traffic stops by law enforcement agencies across the country. Our goal is to help researchers.

What data?

Fake data
Biased data
Mixed data

Reservation

About
Menu
Reviews

SIDES

Fruit Plate \$7
Patatas Bravas, Spicy-Tangy Sauce and Rosemary Aioli \$9

Powered by Owner Verified

SUNLIGHT FOUNDATION

LOG IN

search Follow Us

BLOG TOOLS APIS POLICY ISSUES PRESS ABOUT CONTACT DONATE JOIN

Making government & politics more accountable & transparent .

EU may fine political groups misusing personal data to skew elections

It's hoping to prevent another Cambridge Analytica scandal.

Yale

Jon Fingas, @jonfingas 08.27.18 in Internet 4 Comments 275 Shares

Chris Smith @chris_writes August 17th, 2018 at 12:32 AM

Share Tweet

Google has been dealt two huge blows in Europe in recent years, where antitrust investigations have ruled the company abused its position in search as well as in the mobile market. The company received two record fines as a result, which added up to [more than €6.74 billion \(\\$7.66 billion\)](#). On top of that, a third investigation is in the works and could bring over additional fines.

Computer Science

CS 4111-Introduction to Databases, EugeneWu 35

Computer Science

CS 4111-Introduction to Databases, EugeneWu 36

Computer Science

CS 4111-Introduction to Databases, EugeneWu 37

Data will be crucial to
how we live
as individuals and as a society

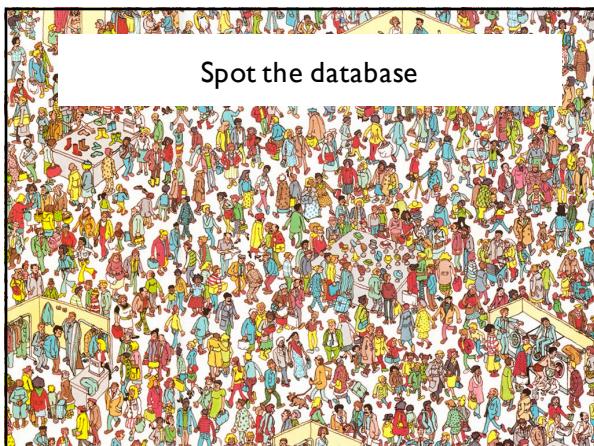
CS 4111-Introduction to Databases, EugeneWu

38

Data ~~will be~~ **is** crucial to
how we live
as individuals and as a society

CS 4111-Introduction to Databases, EugeneWu

39



Spot the database



CS 4111-Introduction to Databases, EugeneWu

41



CS 4111-Introduction to Databases, EugeneWu

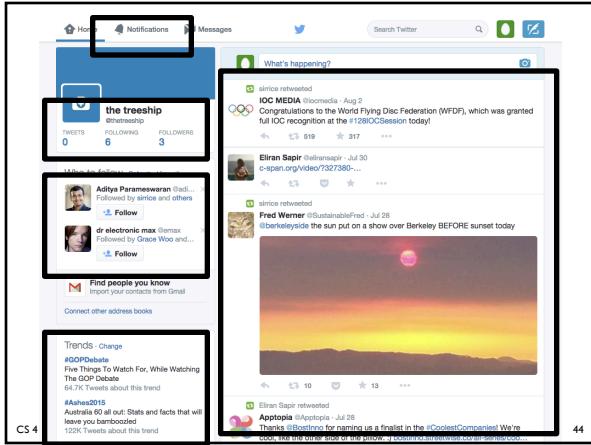
42

Google

Search Google or type URL ↴

CS 4111-Introduction to Databases, EugeneWu

43



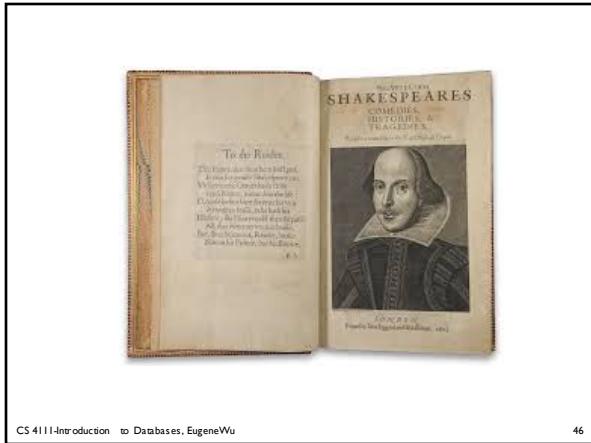
CS 4

44

```

2012-01-04 00:01:23,180 INFO org.apache.hadoop.hdfs.server.datanode.DataNode: Receiving
010
2012-01-04 00:01:23,184 INFO org.apache.hadoop.hdfs.server.datanode.DataNode.clienttrace
cliID: DFSClient_-603743753, offset: 0, srvID: DS-292194659-127.0.1.1-50010-13247633001
2012-01-04 00:01:23,185 INFO org.apache.hadoop.hdfs.server.datanode.DataNode: PacketResp
2012-01-04 00:01:23,291 INFO org.apache.hadoop.hdfs.server.datanode.DataNode: Receiving
10
2012-01-04 00:01:23,293 INFO org.apache.hadoop.hdfs.server.datanode.DataNode.clienttrace
cliID: DFSClient_-603743753, offset: 0, srvID: DS-292194659-127.0.1.1-50010-1324763300176
2012-01-04 00:01:23,293 INFO org.apache.hadoop.hdfs.server.datanode.DataNode: PacketResp
2012-01-04 00:01:23,324 INFO org.apache.hadoop.hdfs.server.datanode.DataNode: Receiving
010
2012-01-04 00:01:23,326 INFO org.apache.hadoop.hdfs.server.datanode.DataNode.clienttrace
cliID: DFSClient_-603743753, offset: 0, srvID: DS-292194659-127.0.1.1-50010-1324763300176
2012-01-04 00:01:23,327 INFO org.apache.hadoop.hdfs.server.datanode.DataNode: PacketResp
2012-01-04 00:01:23,409 INFO org.apache.hadoop.hdfs.server.datanode.DataNode: Receiving
10
2012-01-04 00:01:23,411 INFO org.apache.hadoop.hdfs.server.datanode.DataNode.clienttrace
, cliID: DFSClient_-603743753, offset: 0, srvID: DS-292194659-127.0.1.1-50010-1324763300
2012-01-04 00:01:23,411 INFO org.apache.hadoop.hdfs.server.datanode.DataNode: PacketResp
2012-01-04 00:01:23,433 INFO org.apache.hadoop.hdfs.server.datanode.DataNode.clienttrace
cliID: DFSClient_-2054881899, offset: 0, srvID: DS-292194659-127.0.1.1-50010-1324763300
2012-01-04 00:01:23,494 INFO org.apache.hadoop.hdfs.server.datanode.DataNode: Receiving
10
2012-01-04 00:01:23,498 INFO org.apache.hadoop.hdfs.server.datanode.DataNode.clienttrace
, cliID: DFSClient_-2054881899, offset: 0, srvID: DS-292194659-127.0.1.1-50010-1324763300
2012-01-04 00:01:23,498 INFO org.apache.hadoop.hdfs.server.datanode.DataNode: PacketResp
2012-01-04 00:01:23,523 INFO org.apache.hadoop.hdfs.server.datanode.DataNode: Receiving
000

```



CS 4111-Introduction to Databases, EugeneWu

46

What is a Database?

Structured data

CS 4111-Introduction to Databases, EugeneWu

49

What is a Database?

Lots of
Structured data

CS 4111-Introduction to Databases, EugeneWu

50

Database Management System (DBMS)

A system to **store, manage** and **access** databases

CS 4111-Introduction to Databases, EugeneWu

51

Database Management System (DBMS)

System to **safely** and **reliably** store **lots of persistent** structured data and is **convenient** for **multiple users** to **efficiently** access and modify.

CS 4111-Introduction to Databases, EugeneWu

52

Is a script a DBMS?

Javascript/Python Script

Data stored in variables (RAM)

Very fast access

CS 4111-Introduction to Databases, EugeneWu

53

Is Excel a DBMS?

Microsoft office security

Visually access/modify/compute over data cells

Click save to store persistently

CS 4111-Introduction to Databases, EugeneWu

54

Is the file system a DBMS?

Manages files that are persistently stored on disk

Open/read/seek/write access to files

Access via file names

Access control via permissions

CS 4111-Introduction to Databases, EugeneWu

55

Is the file system a DBMS?

You and a friend edit the same text file

Save at the same time

What happens?

1. Your changes survive
2. Friend's changes survive
3. Both changes survive
4. No changes survive
5. $\exists \cup \cap$

CS 4111-Introduction to Databases, EugeneWu

56

Is the file system a DBMS?

You edit a text file

Computer crashes

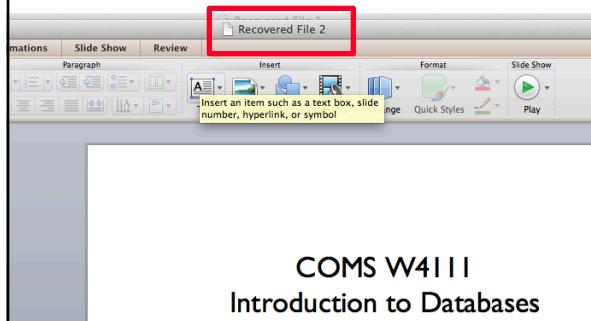
What happens?

1. All changes survive
2. No changes survive
3. Changes from last save survive
4. $\exists \cup \cap$

CS 4111-Introduction to Databases, EugeneWu

57

Is the file system a DBMS?



Want Guarantees from DBMS

You want to write a hot new app on a DBMS.
What do you *not* want to worry about?

Failures disk, machine, human, corruption, deity
Lots of users
Ad-hoc data access
Data formats csv? tsv? custom format?

CS 4111-Introduction to Databases, EugeneWu

59

Database Management System (DBMS)

System to **safely** and **reliably** store **lots** of **persistent** structured data and is **convenient** for **multiple** users to **efficiently** access and modify.

CS 4111-Introduction to Databases, EugeneWu

60

Database Management System (DBMS)

Safe	Consistent and correct data after failures
Reliable	99.99+% Uptime
Lots	>>RAM (terabytes)
Persistent	Lives longer than DBMS application
Convenient	Physical Independence. Declarative.
Multiple Users	Concurrent access. Access control.
Efficient	Fast: 100k+ queries / sec

CS 4111-Introduction to Databases, EugeneWu

61

DBMSes in the Wild

Classic Relational

\$\$: Oracle, IBM, Microsoft, Teradata, EMC, etc
Free: MySQL, PostgreSQL, SQLite

New Relational

In-Memory, Column-store, Streaming

Non-traditional

Search (Google, Bing, Lucene), Scientific, Geo, Graph

NoSQL

Big Data: Hadoop, Spark, etc

Key-value: Mongo, BerkeleyDB, Cassandra, etc

DBMS-as-a-Service

MS Azure, Google BigQuery, Amazon Redshift/RDS ...

CS 4111-Introduction to Databases, EugeneWu

62

Encompasses most of CS

OS	DBMS directly manages hardware
Languages	SQL is a domain specific language
Theory	Algorithms, models, NP-complete
AI/ML	Knowledge Discovery, KDD
Logic	Relational Algebra = 1 st order logic

Scalable Computer Science

CS 4111-Introduction to Databases, EugeneWu

63

Good time to learn!

Cloud programmer

Data science

Data engineer

Machine learning engineer



DATA Data Scientist: The Sexiest Job of the 21st Century

by Thomas H. Davenport and D.J. Patil

FROM THE OCTOBER 2012 ISSUE

64

2 Key Concepts

Data Independence

Declarative Languages

Serve to insulate application programmers
from the system implementation

65

Data Independence

External Schema

Describe how users see data

Conceptual Schema

Describes logical structure

Physical Schema

Describes files and indexes

External Schema



Conceptual Schema



Physical Schema

||

“Data”

CS 4111-Introduction to Databases, EugeneWu

66

Example App: Guuber

Users(**uid int**, name str, age int)

Drivers(**did int**, name str)

Rides(**uid int**, **did int**, distance float, drive_time float)



CS 4111-Introduction to Databases, EugeneWu

67

Data Independence

UID	Name	Age
0	Eugene	17
1	Luis	20
2	Ken	30

0,Eugene,17
1,Luis,20
2,Ken,30
CSV File

What is the number of adults?

CS 4111-Introduction to Databases, EugeneWu

68

Data Independence

UID	Name	Age
0	Eugene	17
1	Luis	20
2	Ken	30

0,Eugene,17
1,Luis,20
2,Ken,30
CSV File

```
n = 0
for line in csv_file:
    attributes = line.split(",")
    if attributes[2] >= 18:
        n += 1
```

CS 4111-Introduction to Databases, EugeneWu

69

Data Independence

UID	Name	Age
0	Eugene	17
1	Luis	20
2	Ken	30

0,1,2
Eugene,Luis,Ken
17,20,30
CSV File

```
n = 0
for line in csv_file:
    attributes = line.split(",")
    if attributes[2] == 18:
        n += 1
```

CS 4111-Introduction to Databases, EugeneWu

70

Data Independence

Conceptual Schema
Describes logical structure

Physical Schema
Describes files and indexes

Users(uid int, name str, age int)

Physical Independence



Conceptual Schema is the API!

CS 4111-Introduction to Databases, EugeneWu

71

Data Independence

Users(uid int, name str, age int)
Drivers(did int, name str)
Rides(uid int, did int, distance float, drive_time float)

“Welcome back Mr. Wu”

CS 4111-Introduction to Databases, EugeneWu

72

Data Independence

Users(uid int, fname str, lname str, age int)
Drivers(did int, name str)
Rides(uid int, did int, distance float, drive_time float)

“Welcome back Mr. Wu”

CS 4111-Introduction to Databases, EugeneWu

73

Data Independence

Conceptual Schema
Describes logical structure

Users(uid int, name str, age int)

Physical Schema
Describes files and indexes



Physical Independence

“Data”

CS 4111-Introduction to Databases, EugeneWu

74

Data Independence

Conceptual Schema
Describes logical structure

Users(uid int, fname str, lname str, age int)

Physical Schema
Describes files and indexes

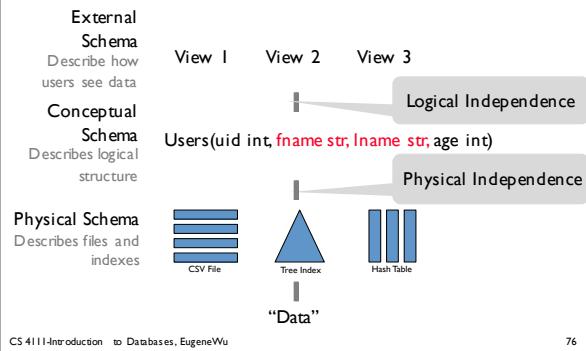


Physical Independence

CS 4111-Introduction to Databases, EugeneWu

75

Data Independence



CS 4111-Introduction to Databases, EugeneWu

76

Data Independence

Physical Independence

Protection from changes in physical structure of data

Logical Independence

Protection from changes in logical structure of data

One of most important properties of a DBMS

CS 4111-Introduction to Databases, EugeneWu

77

Declarative

What you want, not how to do it.

“Make me a sandwich”

“Take two slices of wheat bread out of the 2nd shelf, put them next to each other...”

Buy from pb&j store
Make BLT
½ Tuna
Veggie

What if on 1st shelf?
Out of wheat bread?
No counter space?

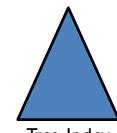
CS 4111-Introduction to Databases, EugeneWu

78

Declarative

“I want all highly rated fast drivers”

————— DBMS ————



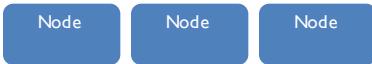
CS 4111-Introduction to Databases, EugeneWu

79

Declarative

“I want all highly rated fast drivers”

————— DBMS ————



CS 4111-Introduction to Databases, EugeneWu

80

Declarative

“I want all highly rated fast drivers”

————— DBMS ————



Classic Components in Databases

Concurrency Control
Transactions
Atomicity
Recovery and Logs

CS 4111-Introduction to Databases, EugeneWu

82

Transaction: Execution of a DB Program

Def. *atomic* sequence of DBMS actions

```
Begin;  
<read beth's account>  
<deduct from beth's account>  
<increase eugene's account>  
Commit; (or Abort;)
```

CS 4111-Introduction to Databases, EugeneWu

83

Transaction: Execution of a DB Program

Def. *atomic* sequence of DBMS actions

Each fully executed transaction must leave DB in *consistent state* if DB is consistent before transaction

- Users specify simple *integrity constraints* on data, and DBMS enforces the constraints.
- DBMS does not understand semantics of its data e.g., doesn't know how bank interest is computed
- User's responsibility to ensure transaction (run alone) preserves consistency

CS 4111-Introduction to Databases, EugeneWu

84

Concurrency Control

Concurrently running multiple user programs needed for good performance

Disk accesses are frequent & slow. Keep CPU working on several user programs while waiting.

Concurrency can cause inconsistencies

- e.g., check cleared while account balance being computed.
- Really hard to program against

DBMS ensures such problems don't arise

- programmers can pretend to use a single-user system.

CS 4111-Introduction to Databases, EugeneWu

85

Scheduling Concurrent Transactions

Transactions T₁, ..., T_n are run concurrently

Equivalent to a *serial ordering* (as if no concurrency)

Locks: T_i requests and waits for lock before read/write.
e.g., T_i locks the database, updates, then releases
e.g., T_i locks the table, updates, then releases
e.g., T_i locks rows, updates, then releases

Will talk about how this works later in course.

CS 4111-Introduction to Databases, EugeneWu

86

Atomicity

Def. Xact fully completes, or never happened even after failures e.g., crashes

Record all actions Xact did during execution in a log

1. **Write ahead logging:** before making any change, ensure the change is safely recorded in log
2. After failure, read log and undo any incomplete Xacts

CS 4111-Introduction to Databases, EugeneWu

87

The Log

A log record contains enough info to undo actions:

Transaction id

T_i writes an object: old and new values

Log record must be safely stored before the changed data

T_i commits/aborts: store commit/abort action

All logging, recovery and concurrency control activities hidden away from user.

Classic Structure of a DBMS

Typical layered architecture
DBMS, not OS, manages
memory and disk

Doesn't show concurrency
control & recovery components



Database Courses at Columbia

COMS W4111 - Intro to Databases

Prerequisites: CS3137 or CS3134; fluency in Python

Intro to DBMSes

Data Models Entity-relation, Relational, ...

Relational Algebra

SQL

Applications + SQL cursors, APIs, embedded ...

Normalization

Peek at DBMS internals:

Storage and indexing

Query optimization

Transaction Processing

COMS W4112-Database Sys. Impl.

Prerequisites: CS3137 or CS3134; fluency in Python

Storage Methods and Indexing

Query Processing and Optimization for 1NF Relations,
including external sorting

Materialized Views and Use in Query Optimization

Query Processing and Optimization for ORDBMSs

Transaction Processing and Recovery

Parallel & Distributed DBMSes: Query Proc. and
Optimization

Parallel and Distributed Databases: Transaction Processing

Performance Considerations Beyond I/Os

COMS E6111-Advanced Databases

Prerequisites: CS4111; fluency in Java or Python

Information Retrieval

Web Search

Distributed Information Retrieval and Web Search

Data Mining

Data Warehousing OLAP, Decision Support

Information Extraction

Scalable Visualization and Interaction

Supporting data analysis

Exploration, explanation and exhibition techniques

COMS E6xxx-DB Research Seminars

Prerequisites: CS4111; fluency in Java or Python

**6998.002 Interactive Data
Exploration Systems**

**6998.005 Database Topics in
Research & Practice**

CS 4111-Introduction to Databases, EugeneWu

94

Administrivia

CS 4111-Introduction to Databases, EugeneWu

95

Next Up

Set up your environment on the cloud

HW0 is out.

Due by Friday 9/7 10AM sharp.

**Must be completed on time to stay in this class
Enrollments will be in order of completion**

CS 4111-Introduction to Databases, EugeneWu

96

Your Instructor: Eugene Wu

B.S. @U.C. Berkeley

Ph.D. @MIT

PostDoc @U.C. Berkeley

Assistant Professor since Fall 2015

Databases, visualization, data analysis
data cleaning crowdsourcing

CS 4111-Introduction to Databases, EugeneWu

97

Your Instructor: Eugene Wu

Contact

www.eugenewu.net
ewu@cs.columbia.edu
421 Mudd
1-212-939-7088

Office hours

Thurs 5:30-6:30PM
By appointment by email

CS 4111-Introduction to Databases, EugeneWu

98

Class Resources

Class web page
w4111.github.io

Discussion board
[piazza](#) (linked from website, public)

Announcements from class staff:
[Website](#)

CS 4111-Introduction to Databases, EugeneWu

99

Your TAs

Felipe Rocha
Amita Shukla
Amit Khat
Ivy Chen
Chih-Chi'an
Mayank Saxena

All TA office hours in CS TA Room (see class web page)
TA office hours will be posted on class web page

CS 4111-Introduction to Databases, EugeneWu

100

Class Information: Prerequisites

COMS W3134 - *Data Structures in Java* or
COMS W3137 - *Data Structures and Algorithms*
(equivalent courses taken elsewhere are acceptable as well)

Fluency in **Python**

CS 4111-Introduction to Databases, EugeneWu

101

Class Information: Lectures

Tuesdays and Thursdays
4 – 5:30PM
501 Northwest Corner Building

(here)

CS 4111-Introduction to Databases, EugeneWu

102

An aside: Success

What does succeeding in this course mean?
Timescales

How to encourage a collaborative environment?
What discourages it?

Assessment

CS 4111-Introduction to Databases, EugeneWu

103

Grading Information

Midterm 1: 25%
Midterm 2: 40% cumulative
HW: 15% (4 HWs equally weighed)
Project 1: 15%
Project 2: 5%
Extra credit: scribe notes + advanced assignments

Median grade: B or slightly higher.
Alternative or make-up exams will not be given.
All homework assignments are equally weighted.
Project 1 has higher weight than Project 2.

CS 4111-Introduction to Databases, EugeneWu

104

Exam Dates

Midterm 1: 10/18, in class
Midterm 2: 12/6 last day of class, in class, cumulative

If you cannot make midterm 2,
do not take this course

CS 4111-Introduction to Databases, EugeneWu

105

Homework

Homeworks usually due at 10AM of due date.

No extensions or exceptions.

Three grace late days for hws throughout the semester.

After using all grace days, 25% grade deduction per late day.

Check full details on web site.

Projects (more details soon)

Two projects.

Teams of two

Run on cloud infrastructure

Get CS account if your team doesn't have a computer

Language is Python

Project 1

Model and build your own database web application
Explore "traditional" relational database features.

Project 2

TBD

Projects (cont.)

No extensions or exceptions for project submission.

3 grace late days total for project.

After using all grace days, 25% grade deduction per late day.

Check full details on web site.

Extra Credit

Added after the curve

Scribe notes: 0-5% extra credit

Advanced Assignments:

- ~same value as HW
- Goes into more depth
- Hack on the DataBass system

Collaboration Policy

Read Syllabus on course site for allowed conduct

CS Dept academic honesty policies
<http://www.cs.columbia.edu/education/honesty>

We will not tolerate any cheating

Collaboration Policy

Discussing lectures and course material strongly encouraged

Homework and exams are individual. No exceptions
Any libraries or code however minor must be disclosed.

Projects are done in teams; no collaboration between teams.

Contact the instructor right away if you have any questions
or are falling behind.

Textbook

Raghu Ramakrishnan, Johannes Gehrke: *Database Management Systems, 3rd edition*, McGraw-Hill, 2002

Available from

[Bookculture](#) bookstore 536 W. 112th St.

Online retailers

Upperclass-persons

On reserve in [Engineering Library](#)

CS 4111-Introduction to Databases, EugeneWu

112

Contests and Rewards

Project I contest

Four best projects chosen as contest winners.

Winners get:

- Option to discuss and demo your project in class.
- 10% boost in Project I grade.

CS 4111-Introduction to Databases, EugeneWu

113

On-going Feedback

Please provide feedback throughout the course.

- What is useful or confusing in lecture
- Thoughts about software stack
- Thoughts about assignments

Email me, come to office hours, talk to staff or:

CS 4111-Introduction to Databases, EugeneWu

114

On-going Feedback

Use form on website

Feedback form
Please share your comments and suggestions for the course
* Required
Feedback *
Share what worked or what was confusing/difficult
Improvements
What change would you suggest to improve things?

CS 4111-Intro

115

Slides borrow material from
Prof. Gravano

Prof. Hellerstein & Franklin@Cal
Prof. Madden & Stonebraker@MIT

(and by transitivity Raghu Ramakrishnan and Johannes Gehrke)

CS 4111-Introduction to Databases, EugeneWu

116



Deep Learning to Learn

SEPT. 10, 2018 (MONDAY)

11:30AM-12:30PM

DAVIS AUDITORIUM

(412 CEPSR)



Pieter Abbeel (UC Berkeley)

CS 4111-Introduction to Databases, EugeneWu

117

w4111.github.io
ewu@cs.columbia.edu

DO HOMEWORK 0!