

Introduction to Database Systems

CISC637, Lecture #1

Ben Carterette

Copyright © Ben Carterette

1

The screenshot shows a web browser with several tabs and windows open. The main window displays a university's website with sections for Faculty Center, Faculty Center, and My Schedule & Rosters. A sidebar on the right lists courses like Database Systems, Special Problem, and Research. Below this is a 'Teaching Schedule' for Spring Semester 2013, listing courses such as Introduction to Database Systems, Database Systems, Special Problem, and Research. A 'My Teaching Schedule' table provides detailed information for each course, including room numbers and times.

A second tab titled 'Reddit' is open, showing a post from /r/awesomeness about a software company that offers their products for free. The post has received 14 comments and 253 upvotes. Another post from /r/awesomeness discusses Mugabe's policies, with 11 upvotes and 442 comments. A third post from /r/awesomeness discusses the book 'The Science of Hand, Heart and Mind' by David Linden, with 3,111 upvotes and 442 comments. A fourth post from /r/awesomeness discusses the movie 'Spider-Man: Homecoming', with 4,445 upvotes and 10 comments. The browser also shows a sidebar for Amtrak with travel options and a rider's guide.

The screenshot shows a mobile application interface. At the top, there's a banner for "UNIVERSITY of DELAWARE" with a blue background and white text. Below the banner, there's a "FlightTrack" section showing flight details from Indianapolis to Denver, Honolulu to San Diego, and Denver to San Diego. The time is 7:10 AM, and the battery level is at 23%. To the right of the flight info is a faculty profile for "Ben Carterette". The profile includes his name, title ("Assistant Professor"), department ("Department of Computer & Information Sciences"), and contact information ("email: carterette@udel.edu, phone: (302) 831-3185"). Below the profile is a sidebar with links like "Recent and Upcoming Projects", "Information Retrieval Lab", and "Project Web Pages". The bottom of the screen shows a footer with navigation icons and the text "Updated 2/4/13 7:10 AM".

Database Programs (DBMSs)

The slide displays logos for several database management systems:

- Microsoft SQL Server**: Features a red and white logo with a stylized globe.
- MySQL**: Features a blue and orange logo with a stylized dolphin.
- ORACLE**: Features a red and white logo with the word "ORACLE" in large, bold letters.
- PostgreSQL**: Features a blue and white logo with a stylized elephant.
- Microsoft Access**: Features a purple and white logo with a stylized microphone icon.
- SQLite**: Features a blue and white logo with a stylized feather quill pen.

Copyright © Ben Carterette

4

Not a Database

Total Sales by Store

| Store | Sales |
|----------|------------|
| Store 1 | \$ 313,765 |
| Store 2 | \$ 107,160 |
| Store 3 | \$ 351,793 |
| Store 4 | \$ 110,747 |
| Store 5 | \$ 232,136 |
| Store 6 | \$ 167,462 |
| Store 7 | \$ 210,073 |
| Store 8 | \$ 308,092 |
| Store 9 | \$ 97,492 |
| Store 10 | \$ 393,484 |
| Store 11 | \$ 396,891 |
| Store 12 | \$ 151,168 |
| Store 13 | \$ 251,390 |
| Store 14 | \$ 397,776 |
| Store 15 | \$ 259,644 |
| Store 16 | \$ 225,184 |
| Store 17 | \$ 335,785 |

Total Sales by Region

| Region | Sales |
|---------|--------------|
| West | \$ 1,718,258 |
| South | \$ 534,389 |
| Midwest | \$ 1,009,248 |
| East | \$ 900,431 |
| Total | \$ 4,162,346 |

TOTAL SALES BY CATEGORY

| Category | Sales |
|-------------|-----------|
| Automotive | \$ 62,858 |
| Gardening | \$ 52,048 |
| Electronics | \$ 33,026 |
| Jewelry | \$ 23,035 |
| Sporting | \$ 16,036 |
| Housewares | \$ 10,149 |
| Books | \$ 12,247 |
| Games | \$ 18,420 |

About Me

- Professor here at UD for 7 years
- PhD from UMass Amherst
- Before that, I worked as a DB admin/web admin/web programmer for 3 years
- Research search systems/information retrieval

Copyright © Ben Carterette

About You

Copyright © Ben Carterette

7

What is a Database?

- A **database** is a collection of data
 - Often representing something in the real world
 - E.g. a university, a bookstore, a social network, ...
 - Interrelated data
 - students, faculty, courses, classrooms, payrolls
 - books, authors, customers, mailing addresses
 - users, friends, updates, likes, comments
 - Nearly always *structured* in a way that supports computing answers to arbitrary questions
- A **database management system** (DBMS) is a software system designed to store, manage, and query that data

Copyright © Ben Carterette

8

Why Use a Database?

- You're building an application. Why consider using a database?
- Some good reasons:
 - lots of interrelations in data
 - data needs to be persistent across sessions
 - too much data to fit in main memory
 - large files too big for efficient random access
 - complex queries must be answered quickly
 - many users access data at the same time
 - different users need different permissions
 - many many more...

Copyright © Ben Carterette

9

Why Use a DBMS?

- Data independence
- Efficient access
- Integrity and security
- Access administration
- Concurrent access
- Application development time

Copyright © Ben Carterette

10

Database Users

- **DBMS programmers** actually implement the DBMS software
- **Database administrators** design schema and storage requirements, handle security, ensure graceful recovery, tune database performance
- **Applications programmers** write software that interacts with a database
- **End users** use the software written by applications programmers



Copyright © Ben Carterette

12

Architecture of a Database-Driven Application

“views”/external schema



stuff the user cares about

conceptual schema



stuff the DBA and application programmers care about

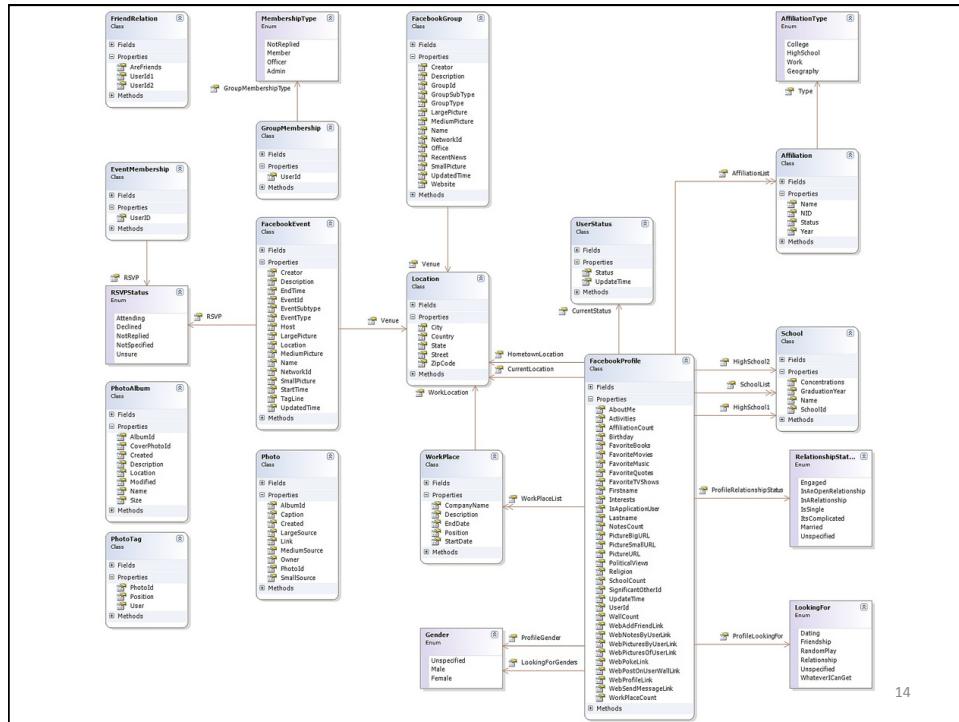
physical schema



stuff the DBMS and system admins care about

Copyright © Ben Carterette

13



14

Why Study Databases?

- Multibillion dollar industry, second only to operating systems
- Databases form backbone of many information-centric applications
 - Using computation to create and understand information
- Understanding databases well incorporates knowledge from every area of CS
 - Systems, theory, artificial intelligence

When and Where

- **Charles Bachman** designed the Integrated Data Store at General Electric in the 1960s
- The *network data model*, a tree-based representation designed for exploration rather than querying
 - Similar to a modern filesystem
- First Turing Award winner in 1973

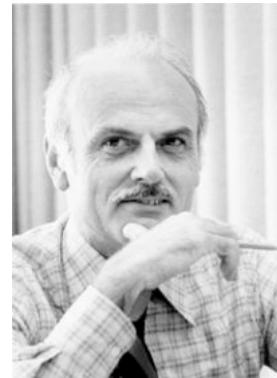


Copyright © Ben Carterette

16

When and Where

- **Edgar Codd** proposed relational data model in 1970 at IBM
- Quickly became the basis of commercial systems; strong theoretical foundation developed
- Turing Award 1981

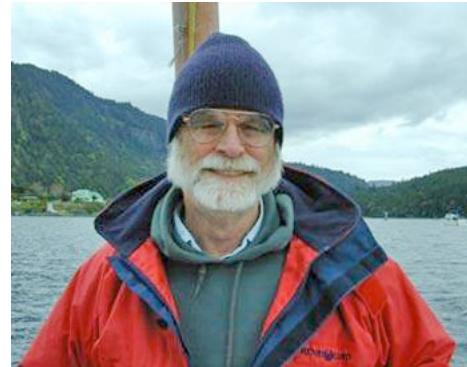


Copyright © Ben Carterette

17

When and Where

- Jim Gray made fundamental contributions to transaction management in the 80s and 90s
- Allowed DBMSs to scale to huge applications with thousands or millions of users
- Turing Award 1999



Copyright © Ben Carterette

18

Summary

- DBMS used to maintain and query large amounts of data
- They allow concurrent access, recovery from failure, fast application development, security
- Levels of abstraction mean that one can work on one subproblem without knowing about others
- Huge industry and huge research area in CS

Copyright © Ben Carterette

19

Activity

- Form small groups (2-4 people)
- Discuss a problem (next slide)
- Write down a solution
 - Be sure to put everyone's name on it

Copyright © Ben Carterette

20

Problem: Reserve a Shelter at a Public Park



A large, two-story stone building with a prominent tower and a gabled roof, situated on a grassy hill overlooking a body of water. The sky is filled with dark, heavy clouds. To the left of the main image, there is a vertical strip containing a map or diagram.

| Accessible? |
|-------------|
| no |
| yes |
| yes |
| no |
| yes |

Activity

- Design a simple phone app to help someone reserve a shelter
 - Just the user interface (the “view” in database terminology)
 - Draw a few screens with buttons, use arrows to show transitions between screens
- Goal: think about the data—what is represented, what information is kept track of—and relationships among them

Copyright © Ben Carterette

22