

CSCI4333  
Database Design & Implement

Spring 2023

**Lecture One - Introduction**

Instructor: Dr. Yifeng Gao

# Outline

- Course syllabus
- Satisfactory of prerequisites
- Introduction to DB & DBMS

# Basics

**Instructor:** Dr. Yifeng Gao

**Contact Info:**

E-Mail: [yifeng.gao@utrgv.edu](mailto:yifeng.gao@utrgv.edu)

Office location: EIEAB 3.214

Office Hours: TR **3:15 pm-4:15pm**

Phone: (956) 665-3054

**Class Meeting:** **TR 2:00pm – 3:15pm**

**GTA:** [juan.m.perez02@utrgv.edu](mailto:juan.m.perez02@utrgv.edu)

**Prerequisites:** CSCI 3333 or CMPE 3333

# Administration Trivia

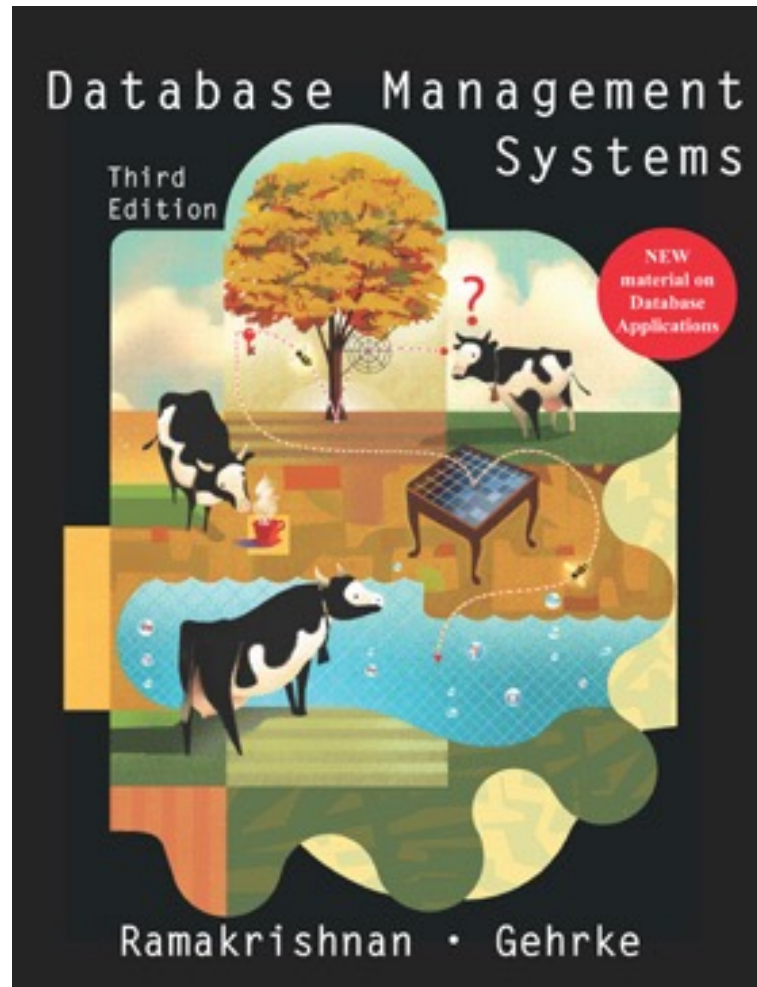
- In most cases, I will put the slides on Blackboard
- You can rely on my slides to review the midterm and final
- You are 100% responsible for any announcements and updates on the class webpage, so visit the page frequently.

# Textbook

- Recommend Book:
  - Database Management Systems, 3rd ed. by Raghu Ramakrishnan & Johannes Gehrke, McGraw-Hill.

[www.cs.wisc.edu/~dbbook](http://www.cs.wisc.edu/~dbbook)

**3<sup>rd</sup> Edition**



# Grading Policies

- Assignment: 40% + Bonus: 10%
  - Project: 20%
  - Midterm exam: 20%
  - Final exam: 20%
- 
- All exams are closed-book, closed-notes.  
The final exam is comprehensive.

# Grading Policies

- Overall grade assignment:
  - A  $\geq 90$ -110
  - B = 80-89
  - C = 70-79
  - D = 60-69
  - F  $< 60$



# Tools we will use for the class

- SQLite
  - More on this later
- Blackboard
  - Assignments, Slides, Project, Announcements
- Whiteboard Notes, HW Answers, Example Questions
  - **In Class Only**

# Contact Policy

- Contact the TA if you
  - Have questions about the course materials
  - Have questions about homework or project
- Contact me if you
  - Have questions about the exams
  - Have general questions/concerns about the course

# Email Policy

- I strongly prefer that you only email me from your official UTRGV email. If you must email me from another account, you must state your full name and your official UTRGV email address.

# Class Schedule

Day	Topic	Assignments
Week 1	Intro to Database	-
Week 2	ER Model - 1	Assignment – 1
Week 3	ER Model - 2	Assignment – 2
Week 4	Relational Model - 1	Assignment – 3
Week 5	Relational Model - 2	Assignment – 4
Week 6	Relational Algebra - 1	Assignment – 5
Week 7	Relational Algebra - 2	Assignment – 6
Week 8	Review & Exam I	-
Week 9	SQL - 1	Assignment – 7
Week 10	SQL - 2	Assignment – 8
Week 11	SQL - 3	Project Start
Week 12	SQL – 4, SQLite	-
Week 13	Normalization	Assignment – 9 and 10
Week 14	Normalization & Review	-
Week 15	Review & Exam II	-
Week 16	-	Project End

# Outcome

Core Objectives	UTRGV Student Learning Outcome Statement	Core Area Requiring this SLO
<b>(1) Design, implement, and evaluate a computing-based solution to meet a given set of computing requirements in the context of the program's discipline.</b>	1. Conceptual & Logical Design 2. Relational Algebra and Calculus 3. Normalization	Assignment 1-6 Assignment 9-10
<b>(2) Apply computer science theory and software development fundamentals to produce computing-based solutions.</b>	1. SQL query 2. Database Design 3. Server Setup and SQLite	Assignment 7 - 8, <u>Project</u>

# Honor Code System

- Vaquero Honor Code
  - <https://www.utrgv.edu/studentlife/about/vaquero-honor-code/index.htm>
- For this class
  - You may work in a team of two for the project.
  - Exams: individual effort, closed books/notes



**Any Questions?**

# About Your Professor



Yifeng Gao



Live in Virginia for 7 years



# Hobby

- Hiking
- Photography



Spring in Rocky  
Mountain National  
Park, CO

\*All photos are taken by me

# Hobby

- Hiking
- Photography



Fall in Shenandoah  
National Park, VA

# Hobby

- Hiking
- Photography



Firefly Summer Night in  
Great Smoky Mountain,  
TN


# Hobby

- Hiking
- Photography



Winter Night (not  
really) in Everglades  
National Park, FL

# Well... not so fun after trip


 **2015** 214.1 GB  
Modified: January 25, 2019 at 1:14 PM

Add Tags...

▼ General:

Kind: Folder

Size: 214,103,703,018 bytes (214.13 GB  
on disk) for 6,762 items

 **2016** 65.49 GB  
Modified: Dec 22, 2017 at 2:52 AM

Add Tags...

▼ General:

Kind: Folder

Size: 65,489,839,730 bytes (65.5 GB  
on disk) for 1,983 items

 **2017** 40.77 GB  
Modified: July 14, 2018 at 12:29 AM

Add Tags...

▼ General:

Kind: Folder

Size: 40,770,759,784 bytes (40.78 GB  
on disk) for 1,312 items

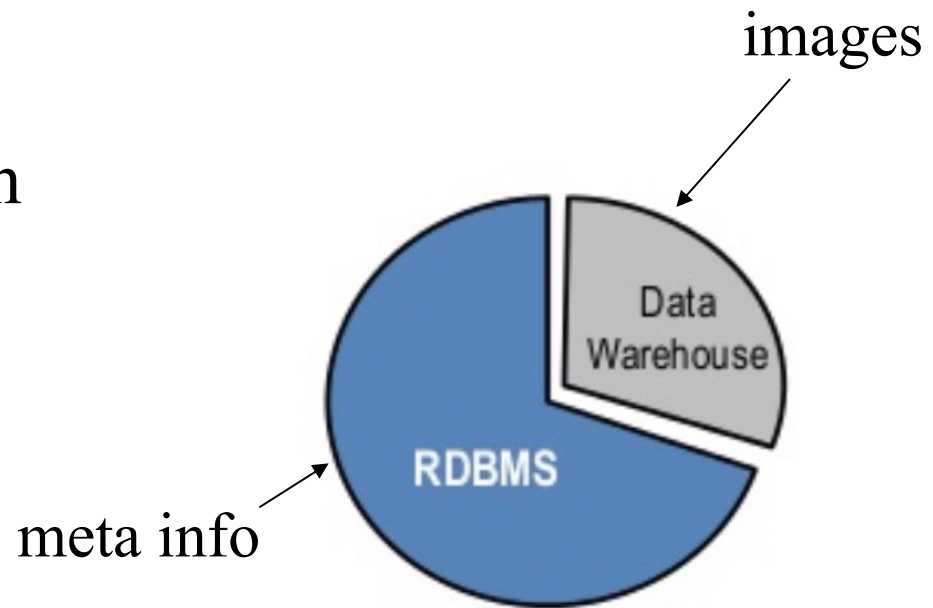
# Given so many images (data)

- How can we organize these images and pick some of them out?

# Given so many images (data)

- How can we organize these images and pick some of them out?

– Database System



# In this example

Title:

Location:

Season:

Time:

State:

meta info



data



# In this example

Title: *Starring Night*

Location: *Everglades  
National Park*

Season: *Winter*

Time: *Night*

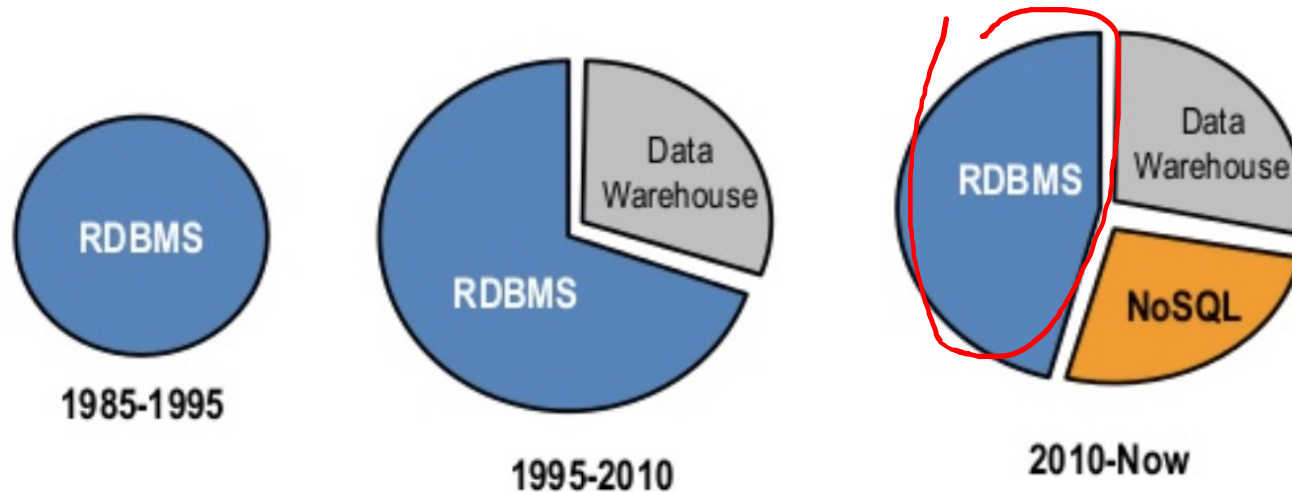
State: *FL*

meta info



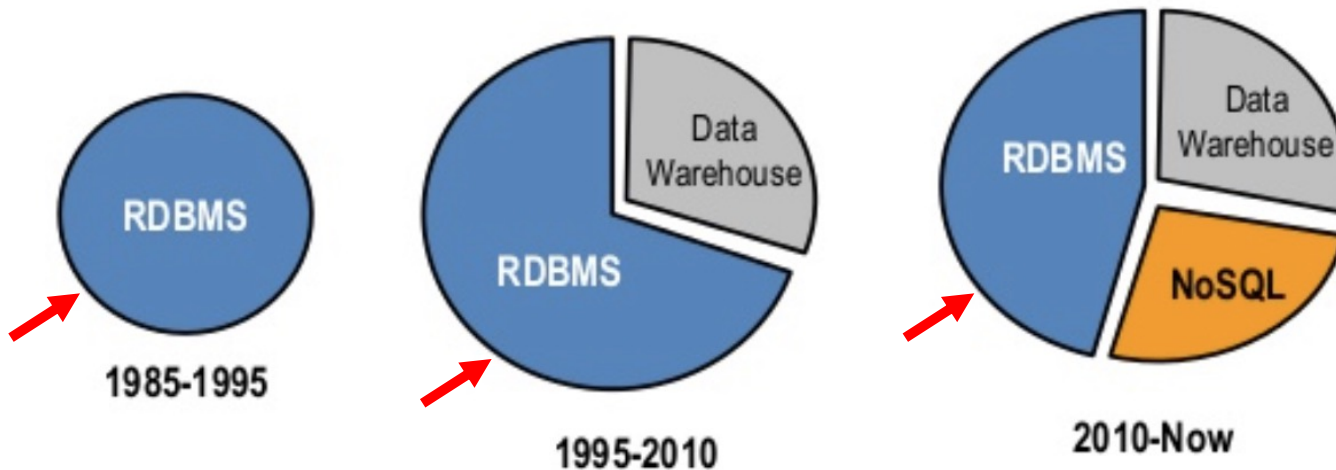
data

# Three Eras of Databases



- RDBMS will continue to be an appropriate solution to many business problems for the foreseeable future

# Three Eras of Databases



- RDBMS will continue to be an appropriate solution to many business problems for the foreseeable future

# Introduction to Databases

# What are Database and DBMS?

- Database:
  - A very large, integrated collection of data.
  - Models real-world enterprise.
    - Entities (e.g., students, courses)
    - Relationships (e.g., Frodo taking INFS 614)
- A Database Management System (DBMS) is a software package designed to store, provide access and manage databases.

# Examples of DBMS Usage

- Airlines: reservations and schedules (expedia.com)
- Universities: student info, grades
- Banking: customer info and accounts (bankofamerica.com)
- Credit Cards: customer info, transactions
- Sales: customer info, inventory (Amazon.com)
- Government: taxes, census

# Top 10 Largest Databases in the World

## 10. Library of Congress

130 million items; 20TB of data; 530 miles of shelves

## 9. CIA

Comprehensive statistics on > 250 countries and entities

## 8. Amazon

59 million customers; > 42 TB of data

## 7. YouTube

45 TB of videos as of 2006

<http://www.comparebusinessproducts.com/fyi/10-largest-databases-in-the-world>

# Top 10 Largest Databases in the World

## 6. ChoicePoint

information on 250 million people; 250TB

## 5. Sprint

2.85 trillion database rows; 365M call records processed/day

## 4. Google

91M searches/day -> 33 trillion DB entries after 1 year

## 3. AT&T

320 TB of info; 1.9 trillion phone call records



# Top 10 Largest Databases in the World

## 2. National Energy Research Scientific Computing Center

2.8 PB of data

## 1. World Data Centre for Climate

220 TB of web data; 6 PB of additional data