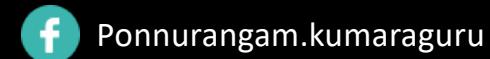
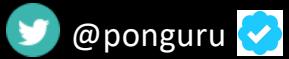
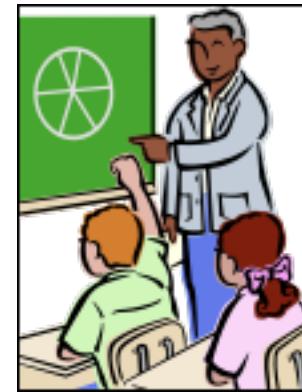


CS4.301 Data & Applications

Ponnurangam Kumaraguru ("PK")
#ProfGiri @ IIIT Hyderabad



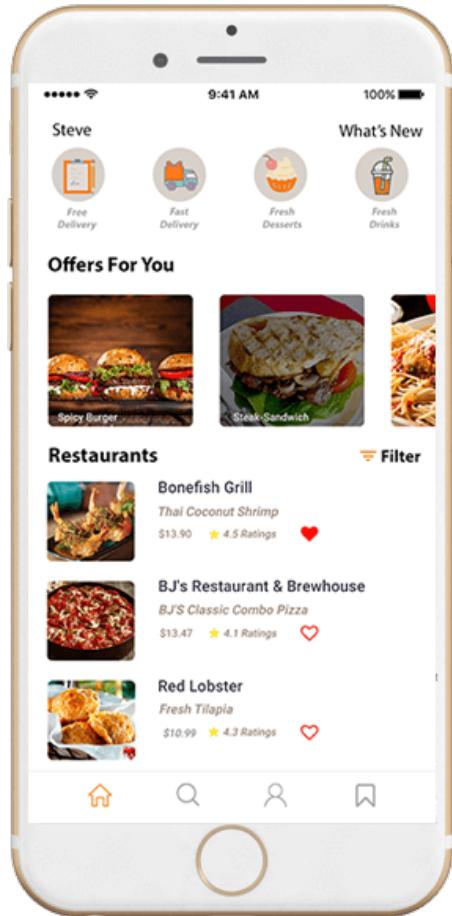
Protocol



Who am I?

- ~~Assistant~~ Associate Professor of Computer Science
- Ph.D. from School of Computer Science, Carnegie Mellon University (CMU)
- Research interests
 - Computational Social Science
 - Social (Societal) Computing
 - Privacy & Security in Social Media
- Courses I teach
 - Online Privacy (1)
 - Privacy and Security in Online Social Media (8), 4+
 - Designing Human Centered Systems (5), 4+
 - Research methods / Advanced research methods (2), 4+
 - Foundations of Computer Security (5), 4+
 - Big Data & Policing (1), 4+

A red square frame containing the number 27 in a large, bold, black font.



make my trip

Download the App
Get exciting welcome offer

myBiz Introducing myBiz Business Travel Solution

My Trips Manage your bookings

Login or Create Account

IN | ENG | INR

Flights Hotels Homestays Holiday Packages Trains Buses Cabs Forex Charter Flights Activities

ONEWAY ROUND TRIP MULTI CITY

FROM TO DEPARTURE RETURN TRAVELLERS & CLASS

Hyderabad **Mumbai** 4 Nov'22 4 Nov'22 1 Traveller

HYD, Rajiv Gandhi International Airport I... BOM, Chhatrapati Shivaji Internation... Friday Friday Economy/Premium Economy Group Booking Now Available!

Select A FARE Type: Regular Fares Armed Forces Fares NEW Student Fares Senior Citizen Fares Doctors & Nurses Fares Double Seat Fares

Recent Searches: Hyderabad → Chennai 4 Nov'22 - 4 Nov'22 Chandigarh → Hyderabad 16 Oct'22

SEARCH

What kind of data are we generating here?



What kind of data are we generating here?

Grading, Relative

Type of Evaluation	Weightage (in %)
Quizzes & Lab Exam (3)	32
Assignments (4) + Activities	20
Project	32
End Sem Exam	16

TAs

11 for now

Students will be rotated among Tas for all evaluations

Plagiarism

What is it?

Copying HWs

Any content taken from another source without citation

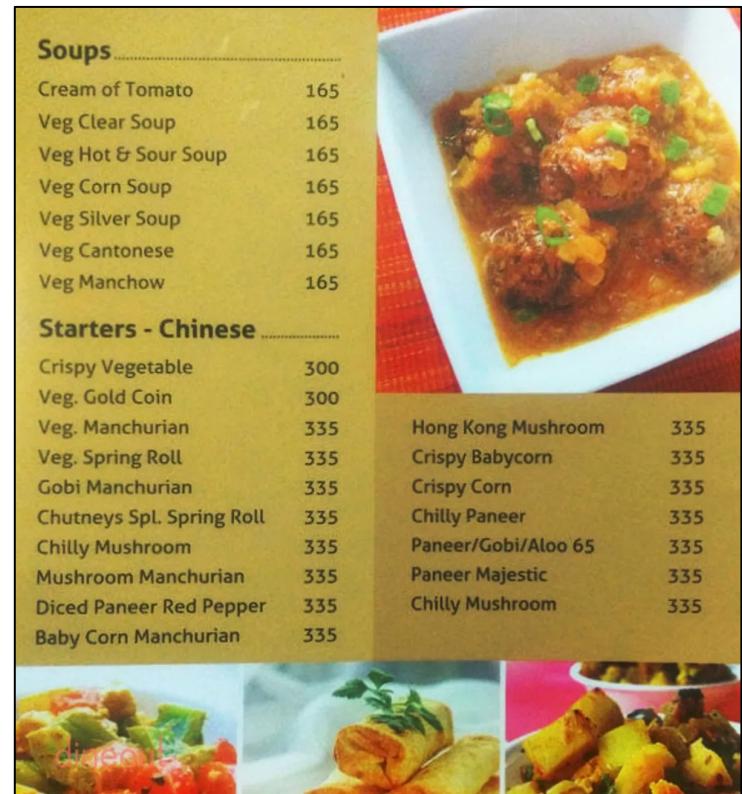
Whatever policy from IIITH

Moodle

We will use Moodle for all content sharing – slides, HWs, etc.

Topics that we will cover

Relational Database Systems
SQL
Database design process
Data Models, Normalization



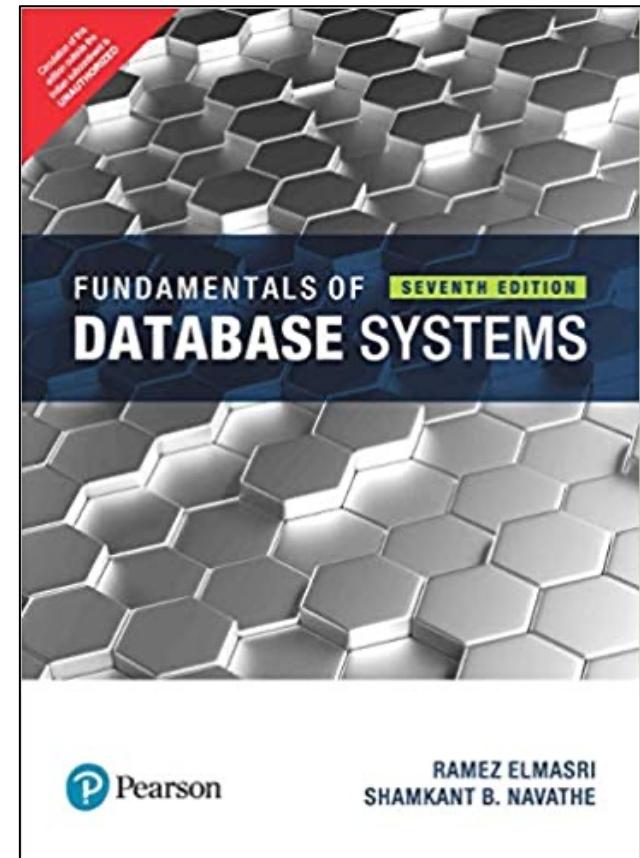
<https://www.dineout.co.in/hyderabad/chutneys-madhapur-west-hyderabad-11747/menu>

Tentative schedule

			03-Nov	Th	23:59hrs		HW3 submission	
5	10		03-Nov	Th	0830 - 1000	Normalization		
		4	05-Nov	S	1130 - 1300	Normalization		
			05-Nov	S	23:59hrs		HW4 publish	Relational Database design
6	11		07-Nov	M	0830 - 1000	SQL		
6	12		10-Nov	Th	0830 - 1000	SQL		Quiz 3
		5	12-Nov	S	1130 - 1300	SQL		
			14-Nov	S	23:59hrs		HW4 submission	
7	13		14-Nov	M	0830 - 1000	SQL		
			15-Nov			All marks check		
			16-Nov			Final project demo		Application
			19 - 26 Nov			End Sem		
			02-Dec			End Sem paper check		
			05-Dec			Grades to be submitted		

Week	Class	Tutorial	Date	Day	Time	Topic	HW	Quiz	Project
1	1		26-Sep	M	0830 - 1000	Intro & data models			
1	2		29-Sep	Th	0830 - 1000	Intro & data models			
			30-Sep		23:59hrs		HW1 publish		
		1	01-Oct	S	1130 - 1300	Data models & project			
2	3		06-Oct	Th	0830 - 1000	DB design & ER models		Quiz 1	
			08-Oct	S	1130 - 1300				
			10-Oct		23:59hrs		HW1 submission		
2	4		10-Oct	M	0830 - 1000	DB design & ER models	HW2 publish		
			12-Oct	W	23:59hrs				Data requirements
3	5		13-Oct	Th	0830 - 1000	DB design & ER models			
		2	15-Oct	S	1130 - 1300	ER models			3
			17 - 19 Oct			Institute Quiz			
3	6		20-Oct	Th	0830 - 1000	Relational DB			
		3	22-Oct	S	1130 - 1300	Relational DB	HW2 publish		
			22-Oct	S	23:59hrs		HW3 submission		
			24-Oct	M	23:59hrs				ER Model
4	7		24-Oct	M	0830 - 1000	Relational DB			
4	8		27-Oct	Th	0830 - 1000	Normalization			
5	9		31-Oct	M	0830 - 1000	Normalization		Quiz 2	

Book I will follow



Any questions / clarifications?

Basic Definitions

Database:

A collection of related data.

Data:

Known facts that can be recorded and have an implicit meaning.

Mini-world:

Some part of the real world about which data is stored in a database. For example, student grades and transcripts at a university.

Database Management System (DBMS):

A software package/ system to facilitate the creation and maintenance of a computerized database.

Database System:

The DBMS software together with the data itself. Sometimes, the applications are also included.

What is a Database?

Data: factual (undoubted) information that can be recorded and have implicit meaning

A database is a collection of related data

What is a Database?

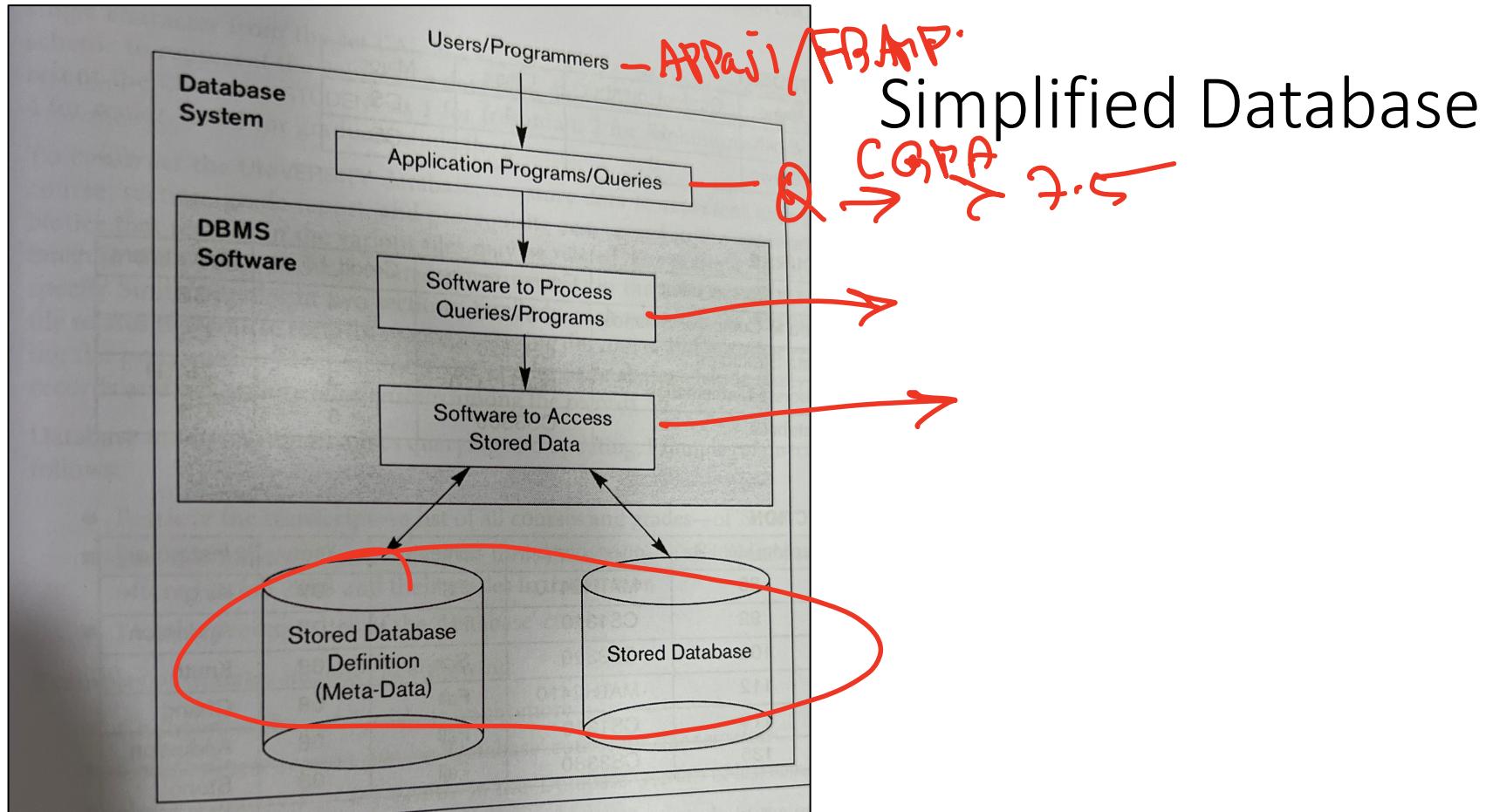
A database has the following implicit properties:

A database represents some aspect of the real world (mini-world or Universe of Discourse (UoD))

A database is a logically coherent (associated, related) collection of data with some inherent meaning

A database is designed, built and populated with data for a specific purpose

It has an intended group of users and some preconceived (already thought of) applications in which these users are interested



Example of Database: University

Data:

STUDENTs

COURSEs

SECTIONs (of COURSEs)

(academic) DEPARTMENTs

INSTRUCTORs

Relation:

SECTIONs are of specific COURSEs

STUDENTs take SECTIONs

COURSEs have prerequisite COURSEs

INSTRUCTORs teach SECTIONs

COURSEs are offered by DEPARTMENTs

STUDENTs major in DEPARTMENTs

Database

STUDENT			
Name	Student Number	Class	Major
Smith	17	1	COSC
Brown	8	2	COSC

GRADE REPORT		
Student Number	Section-Identifier	Grade
17	85	A
18	102	B+

PREREQUISITE	
Course Number	Prerequisite Number
COSC3380	COSC3320
COSC3320	COSC1310

COURSE			
Course Name	Course Number	Credit Hours	Department
Intro to CS	COSC1310	4	COSC
Data Structures	COSC3320	4	COSC
Discrete Mathematics	MATH2410	3	MATH
Data Base	COSC3380	3	COSC

SECTION				
Section-Identifier	Course Number	Semester	Year	Instructor
85	MATH2410	Fall	91	King
92	COSC1310	Fall	91	Anderson
102	COSC3320	Spring	92	Knuth
135	COSC3380	Fall	92	Stone

table

DnA

ADB

CFSO

Database catalogue

RELATIONS		
Column_name	Data_type	Belongs_to_relation
Name	Character (30)	STUDENT
Student_number	Character (4)	STUDENT
Class	Integer (1)	STUDENT
Major	Major_type	STUDENT
Course_name	Character (10)	COURSE
Course_number	XXXXNNNN	COURSE
....
....
....
Prerequisite_number	XXXXNNNN	PREREQUISITE

*Note: Major_type is defined as an enumerated type with all known majors.
XXXXNNNN is used to define a type with four alphabetic characters followed by four digits.*

Book

Views

Many users to DB

Each user may require a different view

View may be a subset or virtual data derived

14 Chapter 1 Databases and Database Views

(a) TRANSCRIPT

Student_name	Course_number	Grade	Semester	Year	Section_id
Smith	CS1310	C	Fall	08	119
	MATH2410	B	Fall	08	112
	MATH2410	A	Fall	07	85
	CS1310	A	Fall	07	92
Brown	CS3320	B	Spring	08	102
	CS3380	A	Fall	08	135

(b) COURSE_PREREQUISITES

Course_name	Course_number	Prerequisites
Database	CS3380	CS3320
Data Structures	CS3320	MATH2410
		CS1310

Figure 1.5

Two views derived from the database in Figure 1.2. (a) The TRANSCRIPT view.
(b) The COURSE_PREREQUISITES view.

Online Transaction Processing (OLTP)

Multiuser DB

Concurrency control

Flight ticket booking, seats available

Transaction

Executing program or process that includes one or more database accesses, reading or updating of database records

Properties [ACID]

Atomicity: either all are executed or none are executed [A/c A \rightarrow A/c B]

Consistency: any data written to a DB must be valid according to the defined rules [telephone number]

Isolation: each transaction appears to execute in isolation, even though 100s may be executing at the same time [updating the seat preference]

Durability: guarantees that once a transaction has been committed, it will remain committed even in the case of a system failure

Actors on the Scene: Day-to-Day use of DB

Database administrators

authorizing access to DB, coordinating & monitoring its use, accountable for security breaches & response time

Database designers

responsible for identifying the data to be stored in the DB, interact with potential group of users and develop *views* of the DB

End Users: Casual, naïve / parametric, sophisticated, stand-alone users

Casual: occasional users, typically middle or high-level managers

Naïve / parametric: constantly updating the db using *canned transaction*, done using mobile apps

bank tellers checking balances post withdrawals & deposits

reservation agents checking for availability

social media users post and read items on platforms

Actors on the Scene: Day-to-Day use of DB

End Users: Casual, naïve, sophisticated, stand-alone users

sophisticated: thoroughly familiarize themselves with all facilities of DBMS, implement their own, complex requirements

stand-alone: maintain personal DB using ready-made programs; TALLY

System analysts & application programmers

determine the requirements of end-users, including naïve, develop specifications for canned transactions

ap implement above specifications as programs, they test – debug – maintain these canned transactions

software developers / engineers play these roles sometimes

Workers Behind the Scene: Maintain the DB

DBMS designers & implementers

a

design and implement the DBMS modules; complex modules like query language processing, interface processing, controlling concurrency, handling data recovery & security

Tool developers

b

design & implement tools; optional packages that are often purchased separately; facilitate DB modeling & design, system design, and improved performance

Operators & maintenance personnel

c

responsible for running & maintenance of the hardware & software environment for DB

Advantages of using DBMS approach

Controlling redundancy

redundancy in storing the same data multiple times

e.g. student details in university maintained by acad & finance office separately

duplication of efforts, storage space, inconsistent data [Jan-19-1998 vs Jan-29-1998]

ideally student details in only one place, *data normalization*

keeping all needed data together, *denormalization*

Restricting unauthorized access

Any questions?

Bibliography / Acknowledgements

Instructor materials from Elmasri & Navathe 7e



pk.profgiri



Ponnurangam.kumaraguru



/in/ponguru



ponguru



pk.guru@iiit.ac.in

Thank you
for attending
the class!!!