953212 DB SYS & DB SYS DESIGN (2, 1, 4)

Prerequisite SE953102 ADT

Agenda

- Course outline
- Introduction to Database system

Course information

- **Prerequisites:** 953102 Abstract data type and problem solving
- **Credit:** 3(2-2-4)
- **Lecture:** Pree Thiengburanathum, PhD.

  Email: pree.t@cmu.ac.th Office room: 415-1
- **TA: 3 TAs (Mr. Elf, Gui, Wang)**

Course information

- Lecture
  - Date: Monday and Thursday
  - Room: online via zoom
  - Time: 13:00-14:30pm
- Lab
  - Date: Tuesday and Friday
  - Room: online via zoom
  - Time: 9:00-11:00am

Course objectives

- **Course objectives:**
  - Student are able to
    - Design logical data model and physical data model.
    - Design and implement physical database.
    - Organize, maintain and retrieve information form a Database Management System.
    - Apply the course knowledge to the following courses:

      953233 Programming methodology (Adv. Programming)

      953261 Interactive web development

Course content

Week no.	Lecture (1 hour)	Lab (2 hours)	Instructor
1	- Course policy announcement - Introduction to Database system	Pre-test on data structure (2-D array and basic sorting, searching)	Pree
2	- Database approach - Component of the DB environment - Range of DB applications	- Lab sheet involves terms, definitions, explain applications.	Pree
3	- Modeling entities and ttributes	- Lab sheet involves basic entity and attributes.	Pree
4	- Modeling relationships ER	- Lab sheet involve basic ER	Pree
5	- EER modeling	- Workbench ER modeling	Pree
6	- EER to relations	- Workbench EER modeling	Pree
7	- Logical database design	- Normalization 1, 2, 3	Pree
8	- Physical database design	Lab sheet file organization	Pree
9	- Intro to SQL	- Introduction to XAMPP Introduction to simple SQL and DB Programming, processing single table	Pree
10	- Advanced SQL	Processing multiple tables, workbench on Advanced SQL	Pree
11	- Database application development	Develop connection (e.g., rest service)	Pree
12	- Database application development	Create connection (e.g., phpMyAdmin)	Pree
13	- Data Warehousing/ Project idea	Integrate with other course projects	Pree
14	- Project implement	Integrate with other course projects	Pree/Toey/Kong

- Sid, sname, lname, gpa
- 1, a, doh
- ..
- ..
- ..

Course requirements

- Lectures in class (3 hour per week)
- Lab Practice (4 hours per week)
- Unannounced Quizzes
- Assignments
- Lab examination
- Midterm and final exams
- Project assignment

Grading system

- Lecture attendance and quizzes
  - Lecture Attendance 5 %
  - Quiz 5 %
- Lab attendance
  - Attendance 5%
  - Assignment 5 %
- Midterm Examination 30%
- Final Examination 30%
- Project Assignment 20 %
- Total 100%

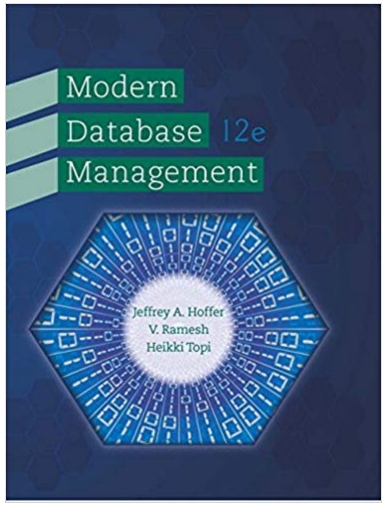
- The semester grade is computed by group reference

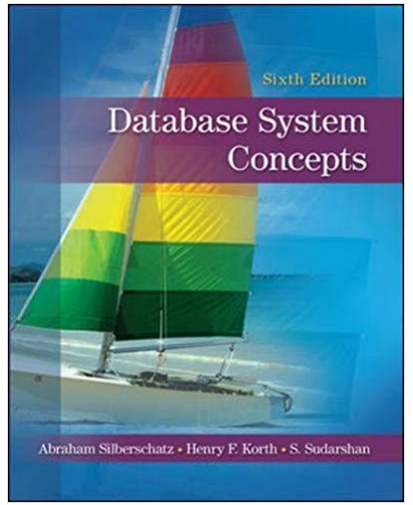
Grade policy:

- Any late assignment submissions will either be penalized (at least 50% reduction) or NOT be accepted.
- If a student is late more **than 15 minutes** in either lab or lecture, you will be regarded as absence.
- If a student needs to be absent with legitimate causes, please notify the lecturer or TA before the date of absence.
- The student who has come to class **less than 80% will NOT** allow to take the FINAL EXAM.
- The student who does not take the final exam gets “F” for this course.
- The work that does not strictly follow the instruction is not accepted.

Course Texts

- Jeff Hoffer, Ramesh Venkataraman, Heikki Topi - Modern Database Management (2016, Pearson Education Limited)
- Silberschatz A., Korth H., Sudarshan S. - Database system concepts (2010, McGraw-Hill Higher Education)
- Google classroom (for the lecture slides)





Course communication

- Facebook group: 953212 63/1
- https://www.facebook.com/groups/824516401285911
- **pree.t@cmu.ac.th**
- **Tas contacts**
- liqian_w@elearning.cmu.ac.th
- aoxue_gui@elearning.cmu.ac.th
- supavas_s@elearning.cmu.ac.th

Google classroom

What is data?

Unit	Value	Size
bit (b)	0 or 1	1/8 of a byte
byte (B)	8 bits	1 byte
kilobyte (KB)	1000^1 bytes	1,000 bytes
megabyte (MB)	1000^2 bytes	1,000,000 bytes
gigabyte (GB)	1000^3 bytes	1,000,000.000 bytes
terabyte (TB)	1000^4 bytes	1,000,000,000,000 bytes
petabyte (PB)	1000^5 bytes	1,000,000,000,000,000 bytes
exabyte (EB)	1000^6 bytes	1,000,000,000,000,000,000 bytes
zettabyte (ZB)	1000^7 bytes	1,000,000,000,000,000,000,000 bytes
yottabyte (YB)	1000^8 bytes	1,000,000,000,000,000,000,000,000 bytes

Hierarchy of data (cont.)

- **Bit** – a bit is the smallest unit of data representation (value of a bit may be 0 or 1)
- **Byte** – a unit of digital information that most commonly consists of eight bits to encode a single character.
- **Field** – a filed consists of a grouping of characters. A data field represents an attribute (a characteristic or quality) of some entity (object, person, place, or event)

Hierarchy of data (cont.)

- **Record** – a record represents a collection of attributes that describe an entity
- **File** – a group of related records.
- **Database** – is an integrated collection of logically related records or files.
- Data structure vs Database?
  - Temporary vs permeant

History of Database Management

Paper-based

- Paper
- Box card index
- Filling

File-based System (Computerized)

- Magnetic tap file, Magnetic disk, Hard disk
- Server

History of Database Systems

1950's and early 1960's:

- Data processing using magnetic tapes for storage
  - Tapes provided only sequential access
- Punched cards for input

Late 1960's and 1970's:

- Hard disks allowed direct access to data
- Network and hierarchical data models in widespread use
- Ted Codd defines the relational data model
  - Would win the ACM Turing Award for this work
  - IBM Research begins System R prototype
  - UC Berkeley begins Ingres prototype
- High-performance (for the era) transaction processing

History of Database Systems

1980's:

- Research relational prototypes evolve into commercial systems
- SQL becomes industrial standard
- Parallel and distributed database systems
- Object-oriented database systems

1990's:

- Large decision support and data-mining applications

Early 2000's:

- XML and XQuery standards
- Automated database administration

Later 2000's: **NoSQL and NewSQL (RDMS enhanced)**

- Giant data storage systems
- Google BigTable, Yahoo PNuts, Amazon,etc.

Database Matter!

- Continental Airlines aka. United Airlines
- Lowest rank to the most admired global company

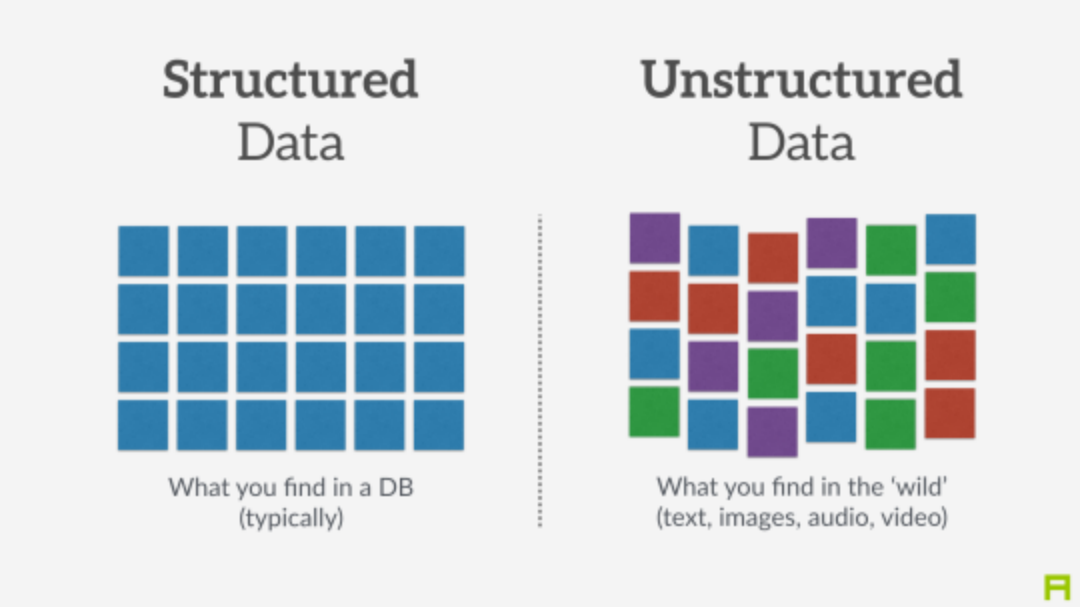
Source: https://www.ausbt.com.au/fyi-continental-airlines-is-no-more-flights-now-united (Accessed August 2018)

More Definitions

- **Database**: organized collection of logically related data
- **Data**: stored representations of meaningful objects and events
  - Structured: numbers, string, dates
  - Unstructured: images, video, documents, textual
- **Information**: data processed to increase knowledge in the person using the data
- **Metadata**: data that describes the properties and context of user data

Structure vs non-structure data

- Structure data
  - Spread sheet consists of numeric, character and dates
- Non-structure data
  - Photo, sound, video segments



Source: http://bigdata.black/infrastructure/storage/unstructured-data/

What is Database?

- A database is any collection of data.
- A DBMS is a software system
  designed to maintain a database.
- We use a DBMS when
  - there is a large amount of data
  - security and integrity of the data are important
  - many users access the data concurrently

Metadata

- Data that describe the properties or characteristics of end-use data and the context of those data.

TABLE 1-1 Example Metadata for Class Roster

Data Item	Metadata					
Name	**Type**	**Length**	**Min**	**Max**	**Description**	**Source**
Course	Alphanumeric	30			Course ID and name	Academic Unit
Section	Integer	1	1	9	Section number	Registrar
Semester	Alphanumeric	10			Semester and year	Registrar
Name	Alphanumeric	30			Student name	Student IS
ID	Integer	9			Student ID (SSN)	Student IS
Major	Alphanumeric	4			Student major	Student IS
GPA	Decimal	3	0.0	4.0	Student grade point average	Academic Unit

2. Hierarchy of Data

Hierarchy	Example
Database	**Employee Database** Employee Details File Training Records File Salary File
File	**Employee Details File** EMP_NAME / JOB TITLE / DATE EMPLOYED Alice Carter / Lecturer / 31 Mar 2002 Faridah bte Hassan / Sales Manager / 9 Aug 2013 Jeffrey Tan / Lecturer / 19 Sep 2004 Steve Willis / HR Manager / 23 Dec 2005
Record	**Employee Record** EMP_NAME / JOB TITLE / DATE EMPLOYED Jeffrey Tan / Lecturer / 19 Sep 2004
Field	**Employee Name Field** EMP_NAME Jeffrey Tan
Byte	01001010 (Letter J in ASCII)
Bit	0

Note: EMP = employee

Source: Jeffrey TL Tan Wikipedia original contributor for Data Hierarchy. 9 Aug 2013
Permission is given to freely use this diagram in its entirety & unedited.

Data Hierarchy Diagram – with Employee Database example

https://en.wikipedia.org/wiki/Data_hierarchy

Example Database Application

- Consider a Internet or Phone Company in Thailand, such as AT&T
- Kinds of information they deal with:

customer records

employee records

billing information

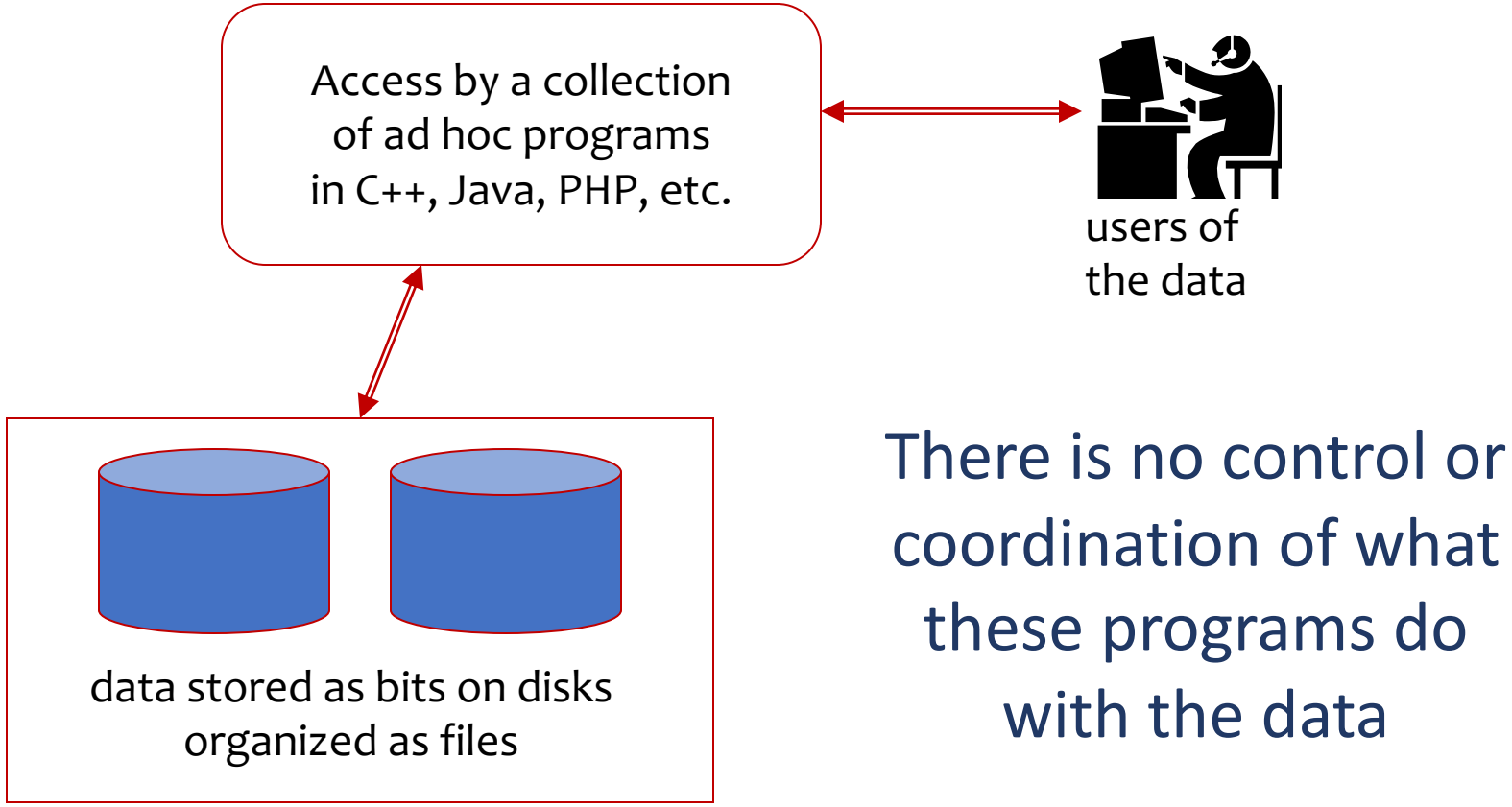
management records

switching and wiring diagrams

customer service orders

Why Use a DBMS?

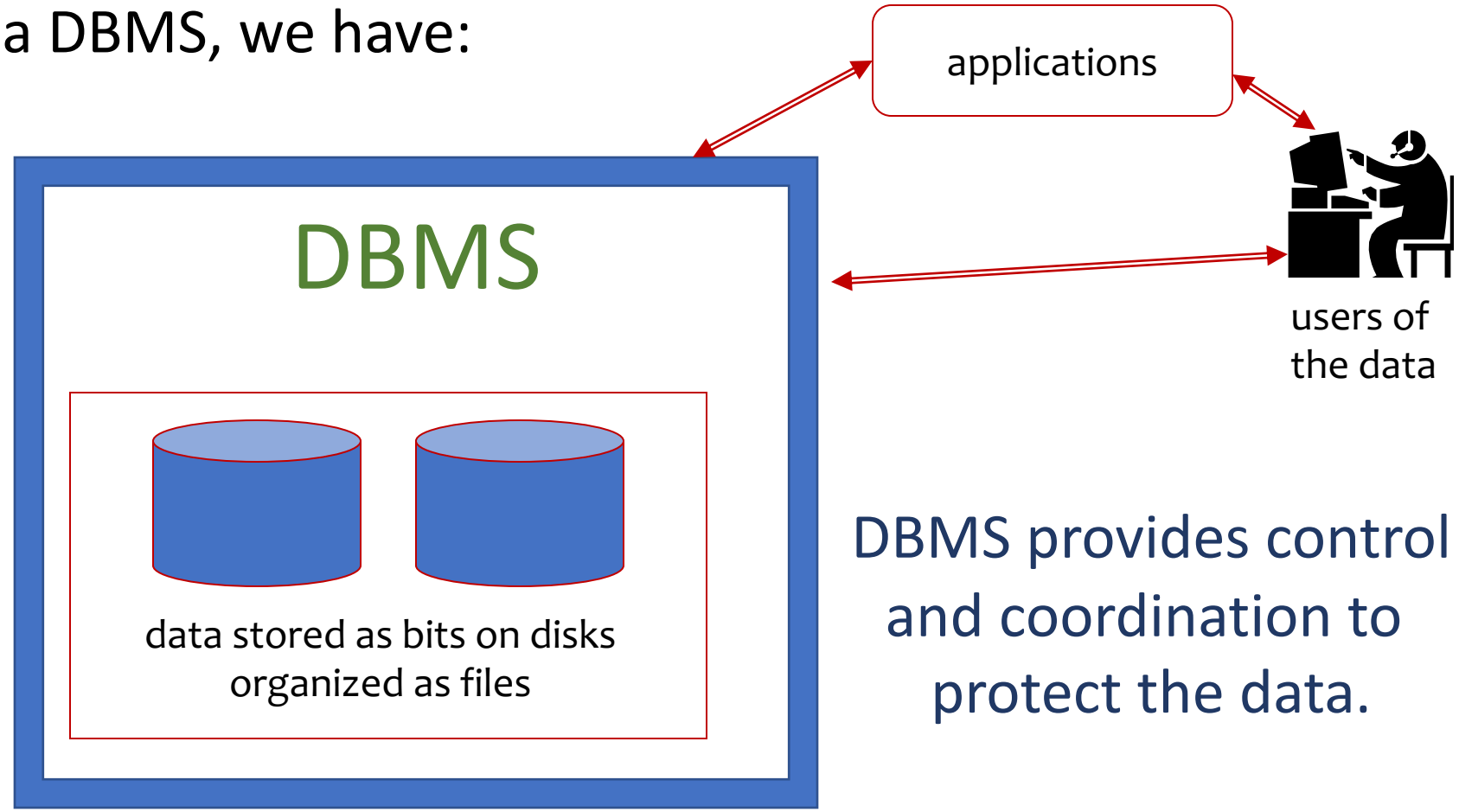
- Without a DBMS, we'd have:



There is no control or coordination of what these programs do with the data

Why Use a DBMS?

- With a DBMS, we have:



DBMS provides control and coordination to protect the data.

DBMS Roles

- **Actors On the Scene**
  (people interested in the actual data):

  - Database administrators
  - Database designers
  - Systems analysts and application programmers
  - End users

1. The Need for Databases

- Database Applications:
  - Banking: transactions
  - Airlines: reservations, schedules
  - Universities: registration, grades
  - Sales: customers, products, purchases
  - Online retailers: order tracking, customized recommendations
  - Manufacturing: production, inventory, orders, supply chain
  - Human resources: employee records, salaries, tax deductions
- Databases can be very large.
- Database are accessed by customer and other remote users via various tech
  - (ATM, Web-browser, mobile phones, etc.)
- Databases touch all aspects of our lives.

University Database Example

- Application program examples
  - Add new students, instructors and courses
  - Register students for courses and generate class rosters
  - Assign grades to students, compute grade point averages (GPA) and generate transcripts

Actors on the Scene

- Database Administrators
    - Acquiring a DBMS
    - Managing the system
    - Acquiring HW and SW to support the DBMS
    - Authorizing access (security policies)
    - Managing staff, including DB designers

Actors on the Scene

- Database Designers
  - Identifying the information of interested
  - Designing the database conceptual schema
  - Designing views for particular users
  - Designing the physical data layout and logical schema
  - Adjusting data parameters for performance

Actors Behind the Scene

- Tools Developers
  - Design and implement DBMS add-ons or plug-ins
  - May work for DBMS supplier or be independent
  - kinds of tools: database design aids, performance monitoring tools, user and designer interfaces

Actors Behind the Scene

- Operators and maintenance personnel
  - Run and maintain the computer environment in which a DBMS operates
  - Probably work for the database administrator (DBA)

Actors Behind the Scene

- Database Researchers
  - academic or industrial researchers
  - develop new theory, new designs, new data models and new algorithms to improve future database management systems

Lab 1

- Tomorrow at 9am