

# Lecture 08 – Database



06016402 : Information Technology Fundamentals  
Asst.Prof.Dr. Pornsuree Jamsri

# 1. Databases, Data, and Information

## Database

- Collection of data organized in a manner that allows access, retrieval, and use of that data

ធនាគារកំណត់អនុវត្តន៍  
ទូរសព្ទ

## Data

- Collection of unprocessed items
  - Text
  - Numbers
  - Images
  - Audio
  - Video

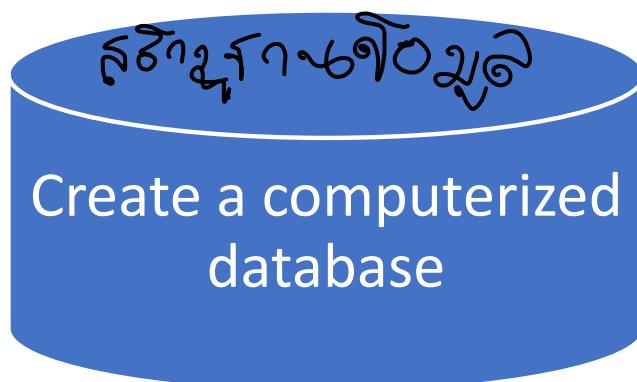
## Information

- Processed data
  - Organized
  - Meaningful
  - Useful

Computers process data in a database to generate information for users.

# 1. Databases, Data, and Information

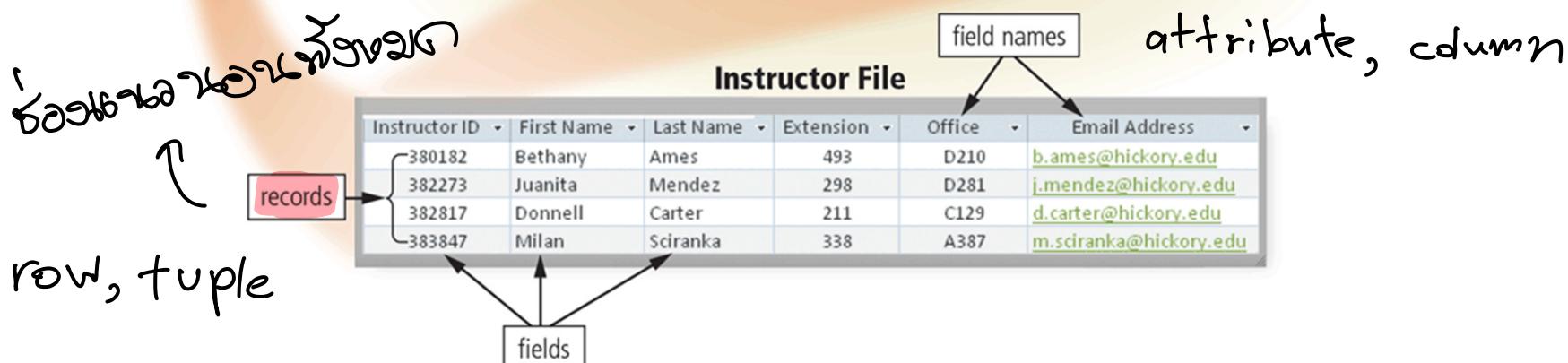
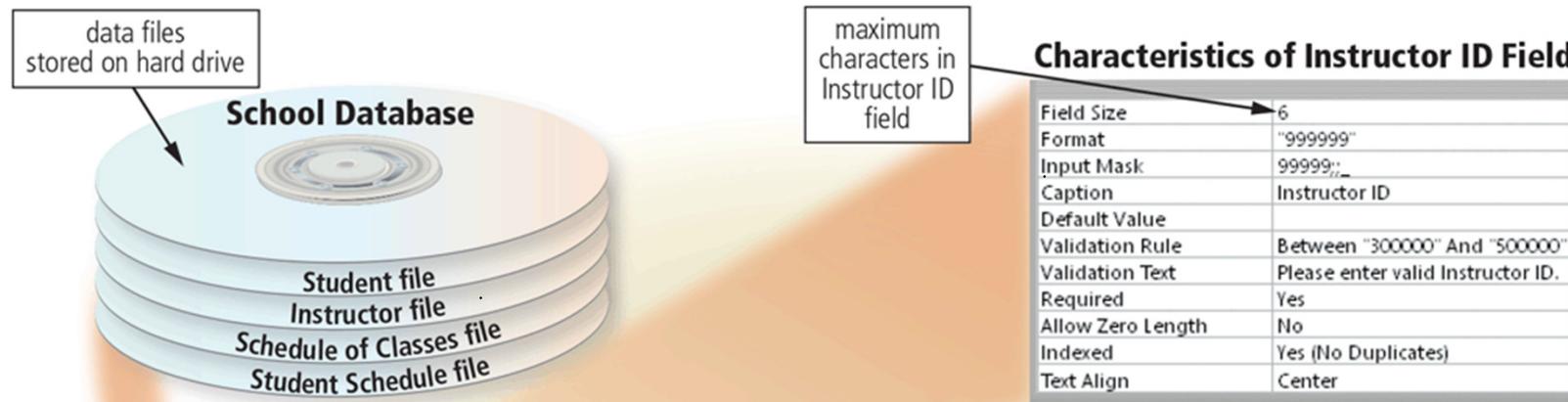
- **Database software**, often called a **database management system (DBMS)**, allows users to:



## 2. The Hierarchy of Data

ពិនិត្យកម្មវិធាន / ០១២៧៨៩

- Data is organized in levels. Each higher level of data consists of one or more items from the lower level.
- A **database** contains a group of related **data files**. A data file contains **records**, a record contains **fields**, and a field is composed of one or more **characters**.



1 single character ex. R, 1, &, etc.



1 byte = 8 bit

## 2. The Hierarchy of Data

**Characters**, a **bit** is the smallest unit of data the computer can process.

Eight bits grouped together in a unit constitute a **byte**. In the ASCII coding scheme, each byte represents a single **character**, which can be a number (4), letter (R), blank space (spacebar), punctuation mark (?), or other symbol (&).

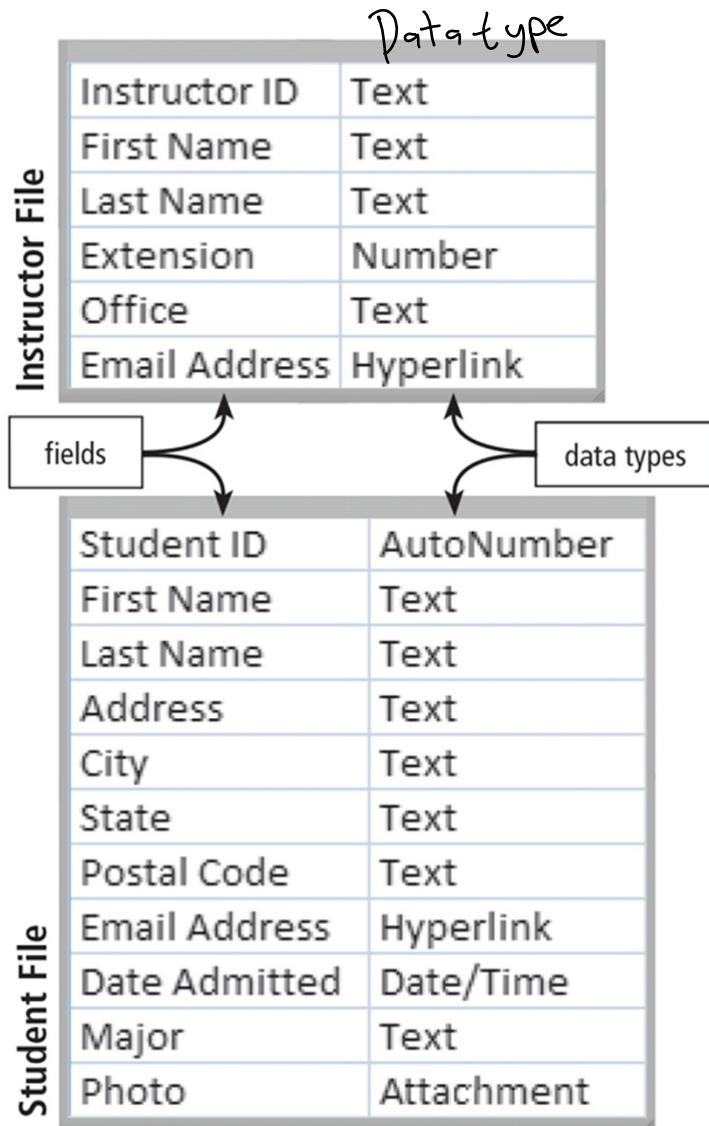
**Fields** A field is a combination of one or more related characters or bytes and is the smallest unit of data a user accesses. A **field name** uniquely identifies each field.

A database uses a variety of characteristics, such as **field size** and **data type**, to define each field.

The **field size** defines the maximum number of characters a field can contain

The **data type** specifies the kind of data a field can contain and how the field is used.

# 2. The Hierarchy of Data



**Common data types** include the following:

- **Text**: Letters, numeric characters, or special characters *String*
- **Number** (also called numeric values): Positive or negative numbers, and the number zero, with or without decimal points *int*
- **AutoNumber**: Unique number automatically assigned by the DBMS to each added record, which provides a value that identifies the record (such as a student ID) គម្រោងលេខ *e.g. 1, 2, 3, 7, ..., 1000*
- **Currency**: Dollar and cent amounts or numbers containing decimal values
- **Date** (also called date/time): Month, day, year, and sometimes time
- **Memo** (also called long text): Lengthy text entries, which may or may not include separate paragraphs
- **Hyperlink**: Email address or web address that links to a webpage on the Internet or document on a network

តើក្នុងទូរសព្ទ, web, email

# 2. The Hierarchy of Data

**Records** A record is a group of related fields.

A primary key is a field that uniquely identifies each record in a file.

The data in a primary key is unique to a specific record. → key ឯកតាត់ទីនូវប្រព័ន្ធដែលមិនមែន

**Data Files** A data file, often simply called a file, is a collection of related records stored on a storage medium, such as a hard drive, or on cloud storage.

↳ google drive, dropbox

Sample Student File											
Student ID	First Name	Last Name	Address	City	State	Postal Code	Email Address	Date Admitted	Major	Photo	
2295	Milton	Brewer	54 Lucy Court	Charlestown	IN	46176		6/10/2016	EE	mbrewer.jpg	
3876	Louella	Drake	33 Timmons Place	Bonner	IN	45208	<u>lou@world.com</u>	8/9/2016	BIO	ldrake.jpg	
3928	Adelbert	Ruiz	99 Tenth Street	Sheldon	IN	46033		10/8/2016	CT	aruiz.jpg	
2872	Benjamin	Tu	2204 Elm Court	Rowley	IN	46167	<u>tu@indi.net</u>	9/14/2016	GEN	btu.jpg	

records → student ID  
key field → student ID

fields → all other columns

# 3. File Maintenance

កំណត់ទៅ  
វិធាននៃការបញ្ចូលព័ត៌មាន  
សំគាល់ការ update

**File maintenance** refers to the procedures that keep data current.

File maintenance includes **adding records to**, **modifying records in**, and **deleting records from a file**.

ផ្តល់, កែតាំង, លក់

Adding ឲ្យកើឡេ

Users add new records to a file when they obtain additional data that should be stored, such as data about a new student admitted to a school.

Modifying នឹងចុចិត្តិភាព

Generally, users modify a record in a file for two reasons:

- (1) to correct inaccurate data or
- (2) to update old data with new data, such as replacing a student's address when she moves to a new address.

Deleting ស្វែងរក ឬលើក

When a record no longer is needed, a user deletes it from a file.

DBMSs use a variety of techniques to manage deleted or obsolete records. Sometimes, the DBMS removes the record from the file immediately, which means the deleted record cannot be restored. Other times, the record is **flagged, or marked**, so that the DBMS will not process it again.

# 4. Validating Data

**Validation** is the process of comparing data with a set of rules or values to determine if the data meets certain criteria. Many programs perform a validity check that analyzes data, either as you enter the data or after you enter it, to help ensure that it is valid.

\* \* ດົນເລີດຈິ່ງຂູ້ຂໍ້ມູນ ຂະໜາງ

**Alphabetic/Numeric Check** An alphabetic check ensures that users enter only alphabetic data into a field.

ຮ່ວງທັງຈັນຂູ້ລູກຕົ້ນມື້ ex. ສັນຕະຫຼຸງ ເວລາລົກ Facebook ລຳເນັ້ນມັນຍຸ

**Range Check** A range check determines whether a number is within a specified range.

ຄວາມສອດຄລ້ອງ

✓ ex. nson Password 2 ຮອບ

**Consistency Check** A consistency check tests the data in two or more associated fields to ensure that the relationship is logical and their data is in the correct format.

**Completeness Check** A completeness check verifies that a required field contains data. For example, some fields cannot be left blank; others require a minimum number of characters. Check ລຳກວດຄວບປັບຍຸ ອາກະນີ \* ບົດເປົ້າ ລຳຕັ້ງການການໃໝ່ຄວນ

**Check Digit** A check digit is a number(s) or character(s) that is appended to or inserted in a primary key value. ຖົດເປົ້າ Password ກຽມສູ່ສິຂະກົດ/ກົດ special char. ພັດ

# 4. Validating Data

**Table 11-1 Sample Valid and Invalid Data**

<b>Validity Check</b>	<b>Field(s) Being Checked</b>	<b>Valid Data</b>	<b>Invalid Data</b>
Alphabetic Check	First Name	Karen	Ka24n
Numeric Check	Current Enrollment	24	s8q
Range Check	Per Credit Hour Fee	\$220.25	\$2,120.00
Consistency Check	Date Admitted, Birth Date	9/19/2016 8/27/1998	9/19/2016 8/27/2017
Completeness Check	Last Name	Gupta	
Other Check	Email Address	eg@earth.net	egearth.net

# 5. File Processing Systems and Databases

## File Processing Systems

ឯកសារឈ្មោះ DB

In the past, many organizations exclusively used file processing systems to store and manage data.

In a typical file processing system, each department or area within an organization has its own set of files. The records in one file may not relate to the records in any other file. Many of these systems have two major weaknesses: redundant data and isolated data.

ចាំងក្រុង

ឧបករណ៍ទីលាក់ និងកំណត់បញ្ជាផលរបាយការណ៍

- **Redundant data:** Because each department or area in an organization has its own files in a file processing system, the same fields are stored in multiple files. Duplicating data in this manner can increase the chance of errors.

ិត្តិយោ

- **Isolated data:** It often is difficult to access data that is stored in separate files in different departments. ឧបករណ៍ទីលាក់ និងកំណត់បញ្ជាផលរបាយការណ៍

# 5. File Processing Systems and Databases

## The Database Approach

When an organization uses a database approach, many programs and users share the data in the database.

### Advantages of a Database Approach

- Reduced data redundancy
- Improved data integrity  
    ความสมบูรณ์
- Shared data
- Easier access
- Reduced development time

### Disadvantages of a Database Approach

- Can be more complex than a file processing system
- Require more memory and processing power
- Data can be more vulnerable  
    เสี่ยง, เป็นภัย

- ធ្វើការជាមួយបច្ចេកទេស  
- ស្ថាបន្ទាន់

# 6. Types of Databases

A **data model** defines **how users view the organization of the data**. It does not define how the operating system actually arranges the data on the storage media. A database typically is based on one data model. Three popular data models in use today are relational, object-oriented, and multidimensional.

Oracle

ເກີນແບ່ງວິທີ row/column (ນີ້ຈະໃຊ້ກົວ stock, warehouse, inventory)

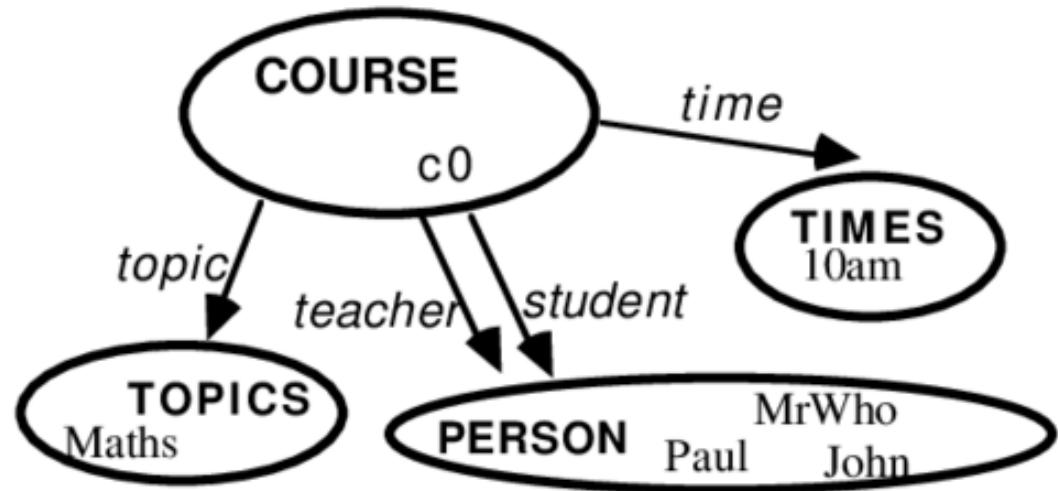
**Relational Database** A relational database is a database that **stores data in tables that consist of rows and columns**. Many organizations use relational databases for payroll, accounts receivable, accounts payable, general ledger, inventory, order entry, invoicing, and other business-related functions.

ຈົດເກີນ data ແລະ ອົບ

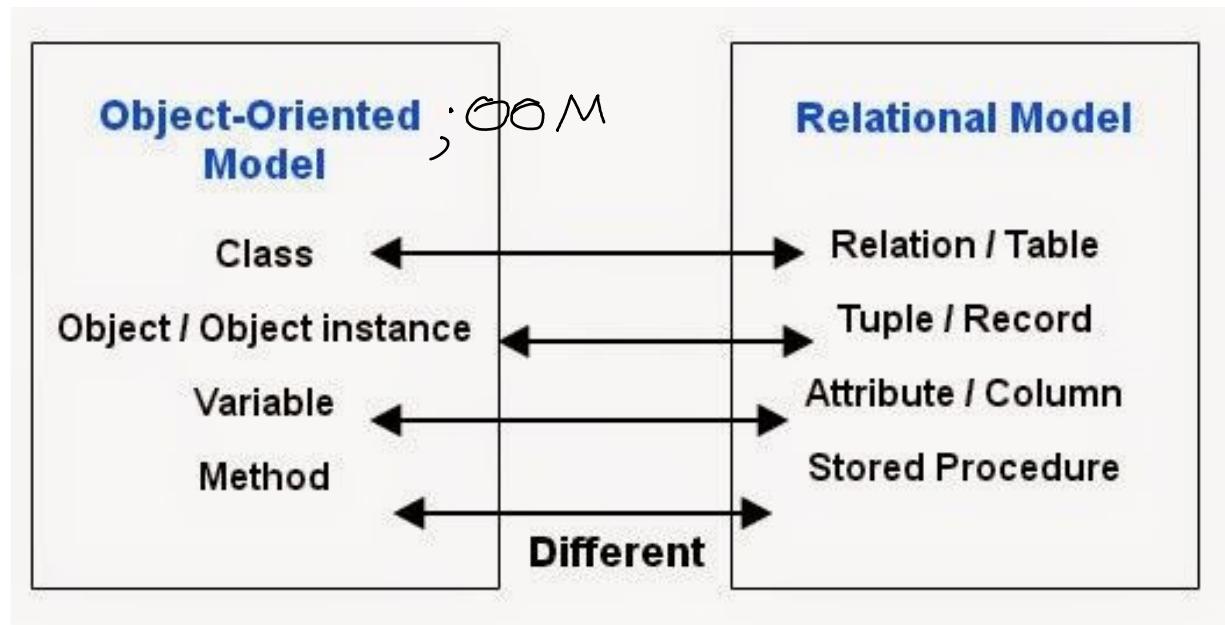
**Object-Oriented Database** An object-oriented database (OODB) **stores data in objects**. Examples of applications appropriate for an object-oriented database include media databases **that store images, audio clips, and/or video clips**.

COURSE			
topic	teacher	student	time
Maths	MrWho	John	10am
Maths	Mr Who	Paul	10am
...			

relational database approach



object-oriented database approach

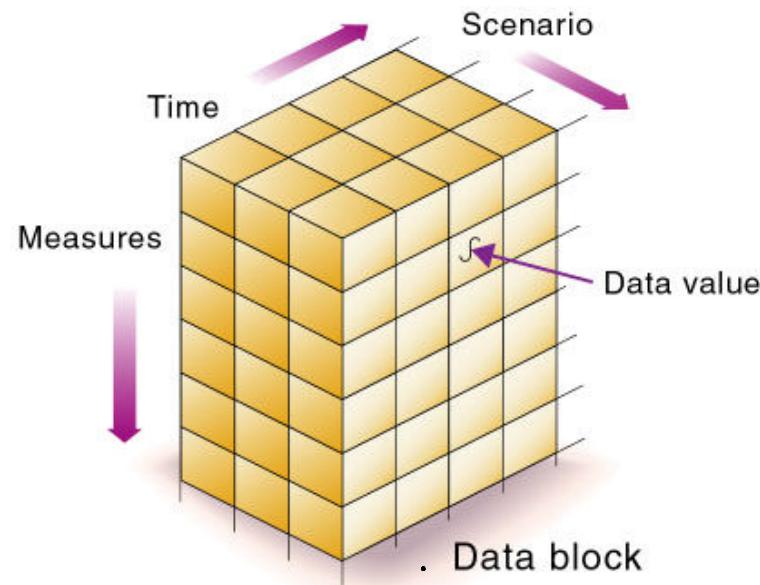


# 6. Types of Databases

**Multidimensional Database** A multidimensional database stores data in dimensions. Whereas a relational database is a two-dimensional table, a multidimensional database can store more than two dimensions of data. These multiple dimensions allow users to access and analyze any view of the database data.

One application that uses multidimensional databases is a **data warehouse**. A data warehouse is a huge database that stores and manages the data required to analyze historical and current transactions.

புது பார்வைகள் மற்றும் தொழில்



	Product A	Product B	Product C	Product D	Product E
Chicago	20M	2M	12M	2M	21M
Cincinnati	30M	4M	10M	8M	26M
Dallas	14M	3M	14M	9M	24M
Louisville	16M	5M	11M	4M	23M
				2002	2003
				2004	

# 7. Database Management Systems

A **database management system** (DBMS), or database program, is software that allows you to create, access, and manage a database. Managing a company's databases requires a great deal of coordination.

ବ୍ୟବସ୍ଥାକ ଡେଟାବେସ୍

The **database administrator (DBA)** is the person in the organization who is responsible for managing and coordinating all database activities, including development, maintenance, and permissions.

ex. ଜୀବନରେ କୌଣସିଲେ  
ଠାର୍ମିକ୍ ପାଇଁ କାହାରେ ବେଳେ

**Data Dictionary** A data dictionary, sometimes called a repository, contains data about each file in the database and each field in those files

Field Name	Data Type	Description (Optional)
Student ID	AutoNumber	Student's ID Number
First Name	Short Text	Student's First Name
Last Name	Short Text	Student's Last Name
Address	Short Text	Student's Address
City	Short Text	
State	Short Text	Lives
Postal Code	Short Text	Lives
Email Address	Hyperlink	Student's Email Address
Date Admitted	Date/Time	Date Student Admitted to School
Major	Short Text	Student's Major Code
Photo	OLE Object	Digital Photo of Student

primary key

fields in Student file

field name

data type for State field

default value

General Lookup

Field Size: 2

Format:

Input Mask:

Caption: State

Default Value: "IN"

Validation Rule:

Validation Text:

Required: Yes

Allow Zero Length: No

Indexed: No

Unicode Compression: No

IME Mode: No Control

IME Sentence Mode: None

Text Align: General

metadata about State field

One can be up to 64 characters long, spaces. Press F1 for help on field names.

# 7. Database Management Systems

**File Retrieval and Maintenance** A DBMS provides several tools that allow users and programs to retrieve and maintain data in the database. To retrieve or select data in a database, you **query** it. A **query** is a request for specific data from the database.

The four more commonly used are

query languages,  
query by example,  
forms,  
and report writers.

డेटाबैज़ / डेटाबैज़

# 7. Database Management Systems

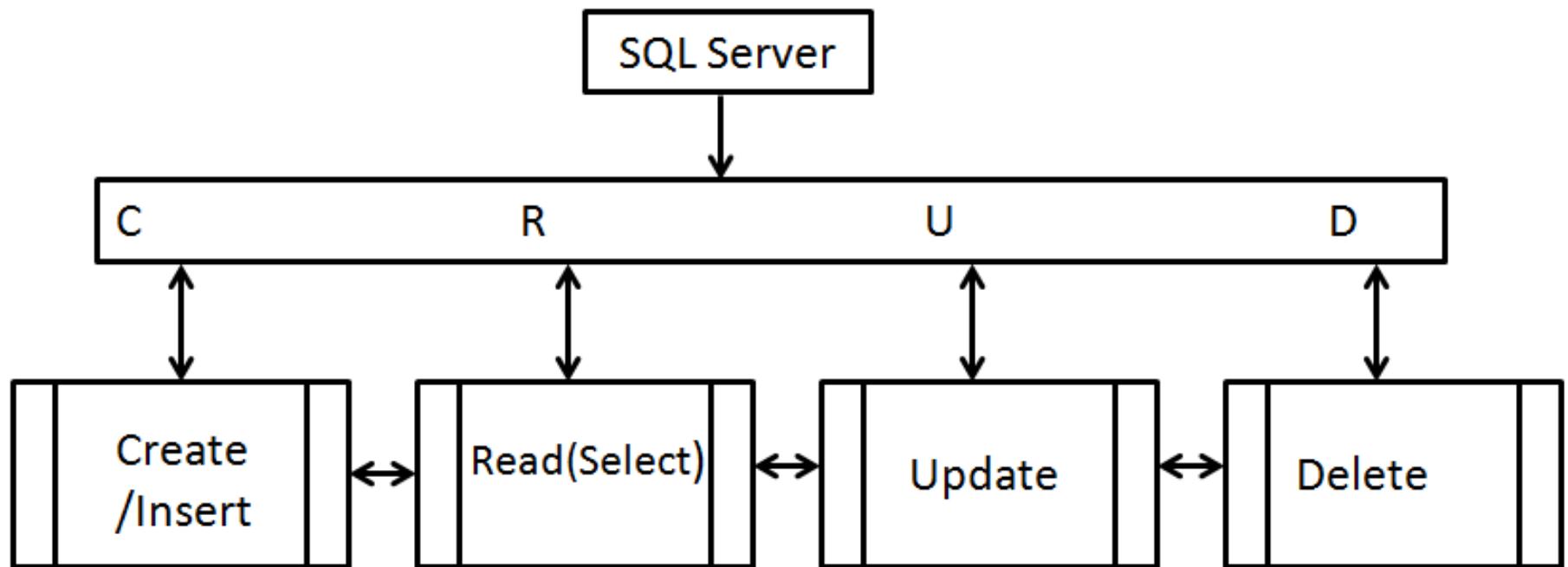
- A **query** is a request for specific data from the database
- A **query language** consists of simple, English-like statements that allow users to specify the data to display, print, store, update, or delete
- **Structured Query Language (SQL)** is a popular query language that allows users to manage, update, and retrieve data

See -Q uq

```
SELECT CLASS_TITLE, CLASS_SECTION,
       MAXIMUM_ENROLLMENT - CURRENT_ENROLLMENT AS SEATS_REMAINING
    FROM SCHEDULE_OF_CLASSES, CLASS_CATALOG
   WHERE SCHEDULE_OF_CLASSES.CLASS_CODE = CLASS_CATALOG.CLASS_CODE
  ORDER BY CLASS_TITLE
```

Class Title	Class Section	Seats Remaining
Algebra 1	51	14
Art Appreciation	52	19
English Composition 1	02	5
Introduction to Sociology	01	14

# Structured Query Language (SQL)



- SELECT ... FROM ... WHERE ...
- INSERT INTO ... VALUES ...
- UPDATE ... SET ... WHERE ...
- DELETE FROM ... WHERE ...

Primary Key

3 fields

	ID	Fruit_Name	Fruit_Color
1	1	Banana	Yellow
2	2	Apple	Red
3	3	Lemon	Yellow
4	4	Strawberry	Red
5	5	Watermelon	Green
6	6	Lime	Green

TABLE ชื่อ Fruits

ຄົວກຸມຕົວ

condition

SELECT \* FROM Fruits WHERE Fruit\_Color='Red'

	ID	Fruit_Name	Fruit_Color
1	2	Apple	Red
2	4	Strawberry	Red

# 7. Database Management Systems

- Most DBMSs include **query by example (QBE)**, a feature that has a graphical user interface to assist users with retrieving data

The diagram illustrates the QBE process across three windows:

- Top Window (Student Table):** Shows a list of student records with columns: Student ID, First Name, Last Name, Address, City, State, Postal Code, Email Address, Date Admitted, and Major. A record for Milton Brewer is selected.
- Middle Window (Filter by Form):** Shows a "Look for" field containing "SOC". An annotation "OrganizaPon" = "SOC" points to the "Major" column header. A box labeled "Major field" points to the "Major" column header, and another box labeled "criteria" points to the "Look for" field.
- Bottom Window (Filtered Student Table):** Shows the same student table with only the two rows for Milton Brewer and Elena Gupta selected, as indicated by the red border around their rows.

# 7. Database Management Systems

- A **form** is a window on the screen that provides areas for entering or modifying data in a database
- A **report writer** allows users to design a report on the screen, retrieve data into the report design, and then display or print the report

Student List by Major						
Major	Last Name	Student ID	First Name	Address	City	Date Admitted
<b>BIO</b>						
	Drake	3876	Louella	33 Timmons Place	Borner	8/9/2016
<b>CT</b>						
	Ruiz	3928	Adelbert	99 Tenth Street	Sheldon	10/8/2016
<b>GEN</b>						
	Tu	2928	Benjamin	2204 Elm Court	Rowley	9/4/2017
<b>SOC</b>						
	Brewer	2295	Milton	54 Lucy Court	Charlestown	6/10/2016
	Gupta	4872	Elena	2 East Penn Drive	Rowley	9/3/2016

# 7. Database Management Systems

## Data Security

A DBMS provides means to ensure that only authorized users access data

อนุญาต

สิทธิ

- Access privileges
- Principle of least privilege policy

e.g. อาจารย์ได้กำหนดเจ้าหน้าที่ Data ไว้ 4 ชื่อ คือ

\* ผู้ดูแลระบบ  
รักษาข้อมูล

การอนุญาตให้ผู้ใช้เข้าระบบหรือเข้าถึงข้อมูลแตกต่างกันตามภาระหน้าที่ความรับผิดชอบ

# 7. Database Management Systems

## Backup and Recovery

PBMS

- A ~~DMS~~ provides a variety of techniques to restore the database to a usable form in case it is damaged or destroyed

សំរាប់ចិត្ត (Duplicate វត្ថុ)

Backup

Monitor ទូទាត់ការបញ្ចូនអ្នកប្រើប្រាស់ ឬឱ្យការណ៍  
ex. CPU ការងារ ផែលកម្ម

Log រួចឱ្យការ recovery

ex. ពាណិជ្ជកម្ម ឬការកែតាំង log  
នៃការ recovery ។

Recovery  
utility

ការបង្ហាញ និងការបង្ហាញ

Continuous  
backup

# Big Data

ទំនុលានាថី ឡើ

នានា Freelance  
stocks, etc.

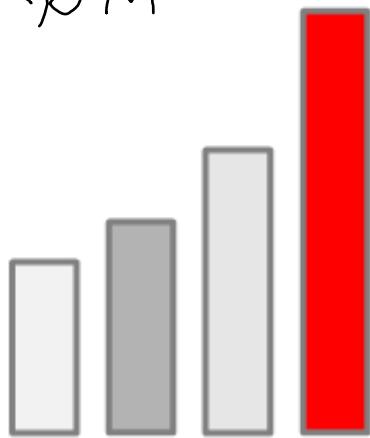
នានាដែវ យកច័ងតាមការ

Data និភាពនៃការលាស់  
អំពីគ្មានមួយចំណាំ

Data ដឹងថ្លែងរបស់ខ្លួន  
ex. អ្នកអ្នកបានដឹងថ្លែង តើខ្លួនតើអ្នកបាន  
ធ្វើដំណឹងនៅក្នុងការណិត  
(ផែនទេសយកឯងជាអំពីការណិត)

## What Makes it Big Data?

សំណុះសំណុះរបស់ពួកយើង  
និង ៧០ M



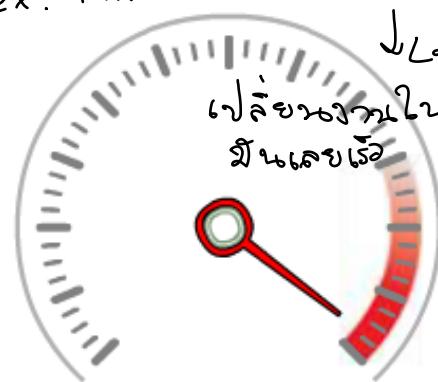
VOLUME

សំណុះសំណុះរបស់ពួកយើង

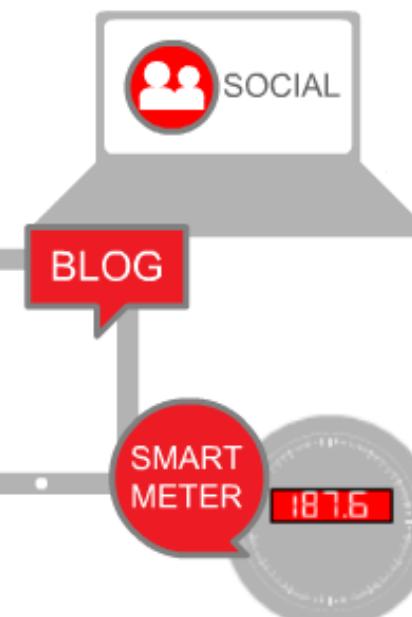
ex. Tax គ្រប់គ្រង ឱ្យរាជការការងារ

Lay-off

សំណុះសំណុះរបស់ពួកយើង



VELOCITY



VARIETY

Data និភាពនៃការលាស់  
អំពីគ្មានមួយចំណាំ

101100101001  
001001101010  
101011100101  
010100100101

VALUE

ទីតាំងផ្ទាល់  
personal life

ORACLE

Goal of Big data : Find useful pattern or models in data



Netflix ก็จะแนะนำสิ่งที่ลูกค้าต้องการด้วย

*Big Data* ที่วิเคราะห์จากการเลือกดูหนังที่ผ่านมาว่า ผู้ชมน่าจะอยากรู้เรื่องใดเป็นเรื่องต่อไป และนำเสนอได้ตรงหรือใกล้เคียงกับความชอบและสนใจของผู้บริโภคมากขึ้นเรื่อยๆ ส่งผลต่อประสบการณ์ และความพึงพอใจ ทำให้มีผู้สมัครสมาชิกมากกว่า 213 ล้านราย (ข้อมูล ณ สิ้นปี 2021)