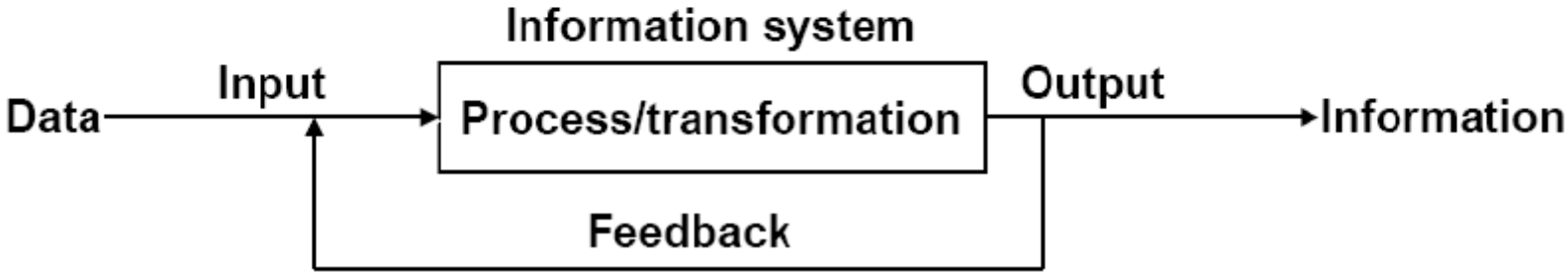
INFORMATION MANAGEMENT (LECTURE)

INTRODUCTION TO DATABASE

LECTURE

MODULE 1

Data vs. Information



- **Data:** Collection of letters, numbers or facts
- **Information:** Processed data that provides value

LECTURE

- ***data is a collection of letters, numbers or facts,*** which by itself derive little or no value. When data is passed as input to an *Information System to be processed and transformed, it becomes* ***Information.***

- Information provides value. During the processing of the data, feedback can be passed back to fine tune the type of data that is collected and how it is collected.

DATABASE SYSTEM

LECTURE

Databases and DBMS

Databases

- A repository of data

DBMS (Database management system)

- Software system that manages databases
- The terms “Database”, “DBMS”, “data server”, “database

• server” often used interchangeably to refer to a DBMS

Why a DBMS?

- Security
- Can handle many users with good performance
- Allows for concurrency while keeping data consistent
- Protects from disaster

- Data can be stored in many ways, for example, text files, spreadsheets, etc. However, using these methods pose problems when trying to access or manipulate the data. For example, if 100 users want to access and make changes to a text file, there will be problems opening the file. This is why databases are needed. Databases are not necessarily cheaper than other solutions, but are better to manage data.
- A **Database** is a repository of data, and a **DBMS (Database management system)** is the software system that manages databases. Often, the terms “Database”, “DBMS”, “data server”, “database server” are used interchangeably to refer to a DBMS
- Why a DBMS?

A DBMS offers security, good performance even when thousands of users are working with it, concurrency while keeping data consistent, and protection from disaster.

WHAT IS A DATABASE SYSTEM

- A database system is an automated system for maintaining records.
- It provides users with several facilities to perform a variety of operations such as:
  - Add new information
  - Delete existing information
  - Update existing information

OVERALL SYSTEM STRUCTURE

- A database system is composed of several specialized modules that perform a particular task or group of tasks.

MODULE	DESCRIPTION
FILE MANAGER	Manages the allocation of disk storage space
DATABASE MANAGER	Provides interface between data stored in database and application programs and queries
QUERY PROCESSOR	Translate statements in query language
DML PRECOMPILER	Converts DML instructions embedded in application program. (data manipulation language)
DDL COMPILER	Converts DDL statement to a set of tables metadata or data about data. (data definition language)

DATABASE SYSTEM COMPONENTS

- DATA
- HARDWARE
- SOFTWARE
- END USER

DATA

- Data stored in a database is both integrated and share.

HARDWARE

- The hardware component is used to store the database itself and the software applications that are used to create and maintain the database.

SOFTWARE

- The software components refers to the interface layer between the users and the physical database.

END USERS

- End users refers to the people whose jobs require access to the database system for querying, updating, and generating useful reports.

- Four categories:
  - APPLICATION PROGRAMMERS
  - NAÏVE OR PARAMETRIC USERS
  - SPECIALIZED USERS
  - SOPHISTICATED USERS

APPLICATION PROGRAMMERS

- Computer professionals who interact with the system through data manipulation language.

NAÏVE OR PARAMETRIC USERS

- Interact with the system by using application programs that were created by the first group of end users.

SPECIALIZED USERS

- Include scientists , engineers, and other specialized professions.

SOPHISTICATED USERS

- Interact with the system without writing programs; they form their request in a database query language such as SQL (Structured Query Language).

DBMS FUNCTIONS

- DATA DEFINITION
- DATA MANIPULATION
- DATA SECURITY AND INTEGRITY
- DATA RECOVERY AND CONCURRENCY
- DATA DICTIONARY

DATA DEFINITION

- The DBMS must be able to accept data definitions(external, internal and conceptual schemas, and all associated mappings) in source form and convert to the appropriate object form.

DATA MANIPULATION

- The DBMS must be able to handle request from the user to retrieve and possibly update existing data in the database, or to add a new data to the database.

DATA SECURITY AND INTEGRITY

- The DBMS must be able to monitor user requests and reject any attempts to violate the security and integrity checks defined by the database administrator.

DATA RECOVERY AND CONCURRENCY

- The DBMS must have the capability of recover from, or minimize the effects of a system crash.

DATA DICTIONARY

- The DBMS must provide a system database called the database dictionary.

ADVANTAGES

- REDUCED DATA REDUNDANCY
- DATA INTEGRITY
- DATA INDEPENDENCE
- DATA SECURITY
- DATA CONSISTENCY
- EASIER USE OF DATA
- LESS STORAGE

REDUCED DATA REDUNDANCY

- In database system, data redundancy is avoided by keeping a single copy of the data.

DATA INTEGRITY

- Since a single copy of a particular data is stored in the database, you can be certain that the changes to the data will be reflected in all future uses of that data.

DATA INDEPENDENCE

- The structure of the database system requires that data be independent of other data in the database and the software used to access the database.

DATA SECURITY

- A database system can have additional security measures as part of the database software product.

DATA CONSISTENCY

- Data consistency is maintained in the database environment.

EASIER USE OF DATA

- Data is easier to use in the database environment.

LESS STORAGE

- Since data redundancy is reduce if not eliminated, the database will occupy less storage space.

DISADVANTAGES

- COMPLEXITY
- EXPENSE
- VULNERABILITY
- SIZE
- TRAINING COSTS
- COMPATIBILITY
- LOCKING-IN OF TECHNOLOGY

COMPLEXITY

- Databases and their associated DBMS are extremely complex.

EXPENSE

- The installation of the DBMS and its databases onto the current system requires careful planning because of the possibility of the need for additional hardware.

VULNERABILITY

- Because data in a database is in a central location, it is vulnerable to partial or complete destruction when a breakdown of hardware components occur or in software mistake.

SIZE

- A typical mainframe database is very large.

TRAINING COSTS

- When a company purchase a database product, it must first train one or more individuals to be the database administrator for the new product.

COMPATIBILITY

- One of the pitfalls of most database systems is their incompatibility with other database systems.

LOCKING-IN TECHNOLOGY

- What if the vendor expresses a desire to abandon the product? There is a possibility that they will have to redesign their system from scratch.