# CHAPTER 1. Introductory Database Concepts

## 1.1 Use of Databases

- Databases play a major role in today's computing environment, with applications.
  (ex) online banking, airline reservations, health care providers, credit card, etc.

- Databases are used to satisfy the information needs of many organizations and individuals in a variety of area.

- A poorly designed database may fail to provide the required information or may provide outdated, flawed, or contradictory information.

- It is important to understand the theoretical foundations, internal structure, design, and management of databases.

## 1.2 A Sample Database

- A ***database*** is a collection of related stored data.
  (ex) A simple example of University database by Microsoft Access
  (see Figure 1.1 (a, b, c, d) (pp3-4), Figure 1.2 (p5), and Figure 1.3 (p6))

## 1.3 & 1.4 (Roles in) The Integrated Database Environment

- Having multiple copies of the same data within isolated files or small databases can lead to flawed, outdated, or contradictory information.

- Most organizations can benefit by having their data integrated into a single database.

- An ***integrated database*** is a collection of related data that can be used simultaneously by many departments and users in an enterprise.

- See Figure 1.4 The Integrated Database Environment (p8) and
  Figure 1.5 Roles in the Database Environment (p9)

  - ***Metadata*** : data about data

  - ***Database Administrator (DBA)*** : is responsible for creating and maintaining the database.

- ***Database Management Systems (DBMS)*** : is a s/w package that allows creating and maintaining databases. All access to the database is controlled by a DBMS.
- ***End Users*** : people who use the data to perform their jobs

- The applications programs, which might be written ***in different programming languages*** such as Java, C, C++, C#, Visual Basic or COBOL, go through the DBMS, which can present data in the form each program expects.

## 1.5 Advantages of the Integrated Database Approach

- Sharing of Data
- Control of Redundancy
- Data Consistency
- Improved Data Standards
- Better Data Security
- Improved Data Integrity
- Balancing of Conflicting Requirements
- Faster Development of New Applications
- Better Data Accessibility
- Economy of Scale
- More Control over Concurrency
- Better Backup and Recovery Procedures

## 1.6 Historical Developments in Information Systems

- ***Sequential File Processing System*** using sequential access devices such as
  - punched cards in 1890 US census
  - punched paper tape introduced in 1940s
  - magnetic tape introduced about 1950-used in UNIVAC I

  (ex) payroll, shown in Figure 1.6 (p14)

- ***Early database models***
  - magnetic disk introduced in 1950s - direct access device
  - programming languages COBOL and PL/1 for commercial data processing developed in 1960s
  - ***hierarchical model*** for data base was developed during 1960s
  - ***network model*** for data base was developed during 1960s
  - both were complex and requiring users to understand data structures and access path to data

- ***Relational model*** Proposed by E.F. Codd in 1970
  - strong theoretical (mathematical) foundation

- ***Semantic models***
  - entity-relationship model - P.P. Chen, 1976
  - object-oriented model - Introduced in 1990s
  - object-relational model: object-oriented capabilities in relational databases

- ***Data warehouses*** - developed in 1990s

- ***Internet access*** to vast network of databases
  - e-commerce
  - XML standard for data exchange

- ***Big Data*** refers to capturing, organizing, and analyzing massive amounts of data generated by a wide variety of sources.