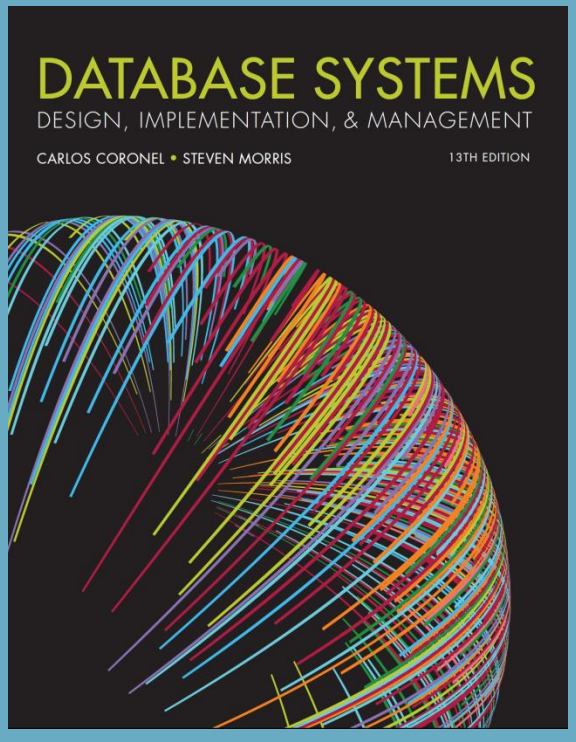


Chapter 1

Database Systems

Learning Objectives

- After completing this chapter, you will be able to:
  - Define the difference between data and information
  - Describe what a database is, various types, and why they are valuable assets for decision making
  - Explain the importance of database design
  - See how modern databases evolved from file systems
  - Understand flaws in file system data management
  - Outline the main components of the database system
  - Describe the main functions of a database management system (DBMS)

Why study databases?

Databases are pervasive in our daily lives. They are everywhere.

When you shop? The store keeps track of purchases, sales, inventory. The bank keeps track of your balance when you pay with your debit or credit card.

When you drive? Road sensors and other traffic detection systems count cars.

When you browse the internet? Every page served is logged.

Can you think of more examples of databases you may have used in the last 24 hours?

What is a database?

A database is a shared, integrated computer structure that stores data and metadata.

- End-user data: raw facts of interest to end user
- Metadata: data about data, through which the end-user data is integrated and managed
  - Describes data characteristics and relationships

"I'd like you to build me a database"

https://www.youtube.com/watch?v=F4OIDszDA9E

Data vs. Information

Databases hold data to be interpreted. Data are the raw facts about people, places, things, and events.

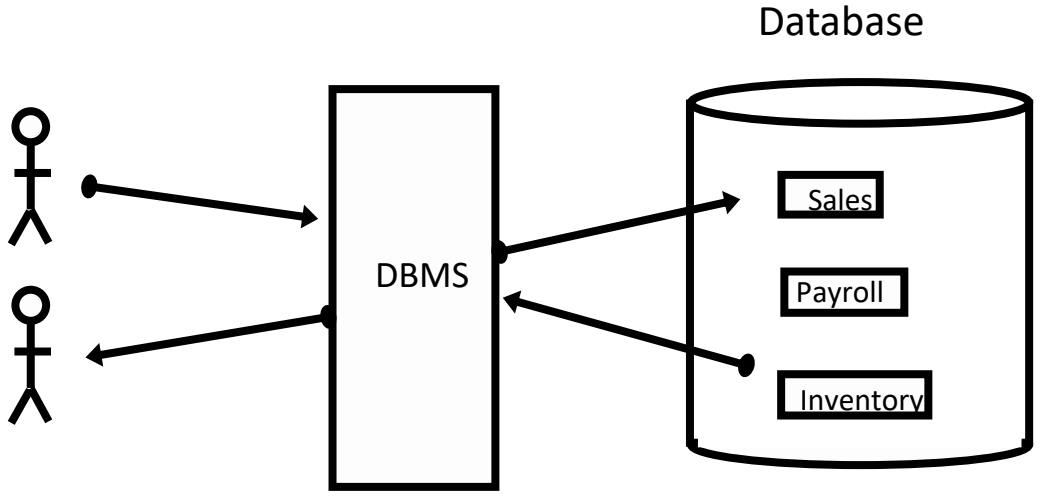
Information is generated as result of interpreting the data.

Data are Raw Facts
Information is interpreted Data

Database Management System

A database management system (DBMS) is a collection of programs that:

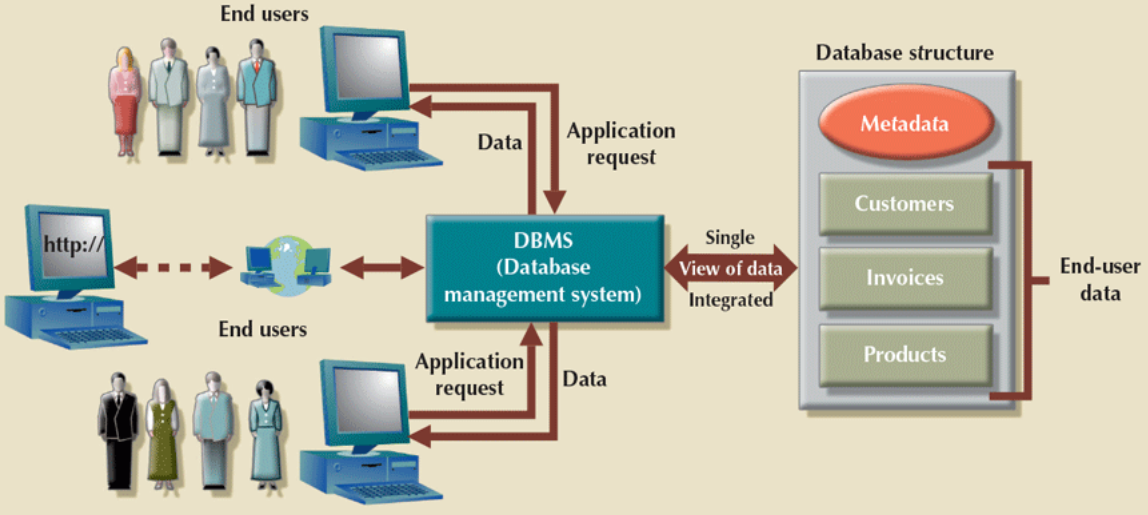
- manages the database structure
- controls access to the structure



Oracle
SAP
SyBase

Database management system (DBMS)

FIGURE 1.4 THE DBMS MANAGES THE INTERACTION BETWEEN THE END USER AND THE DATABASE



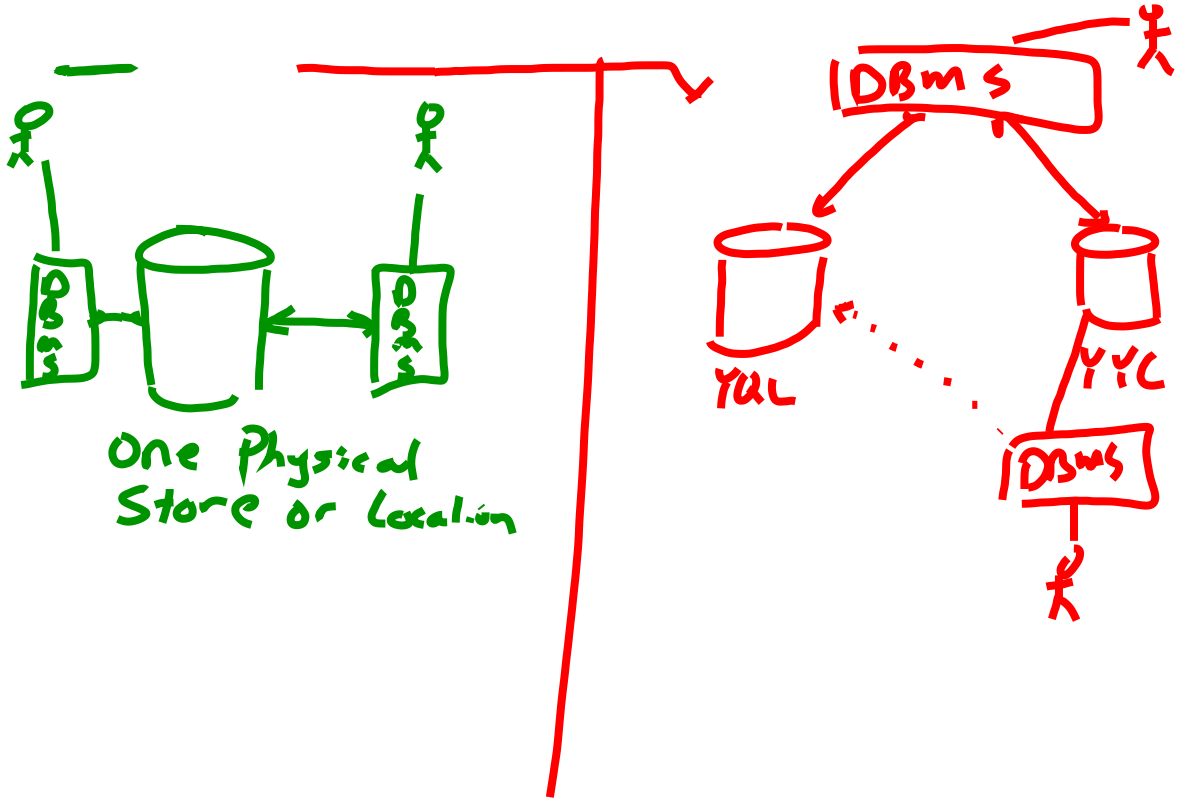
Types of Databases

- Single-user database: supports one user at a time
  - Desktop database: single-user database on a personal computer
- Multiuser database: supports multiple users at the same time
  - Workgroup databases: supports a small number of users or a specific department
  - Enterprise database: supports many users across many departments

Types of Databases

- Classification by location
  - Centralized database: data located at a single site
  - Distributed database: data distributed across different sites
  - Cloud database: created and maintained using cloud data services that provide defined performance measures for the database

centralized vs. distributed



Types of Databases

- Classification by data type
  - General-purpose database: contains a wide variety of data used in multiple disciplines
  - Discipline-specific database: contains data focused on specific subject areas
  - Operational database: designed to support a company's day-to-day operations

Types of Databases

- Databases can be classified to reflect the degree to which the data is structured
  - Unstructured data exists in its original (raw) state
  - Structured data results from formatting
    - Structure is applied based on type of processing to be performed
  - Semistructured data: processed to some extent

Why Database Design Is Important

- Focuses on design of database structure that will be used to store and manage end-user data
  - Well-designed database: facilitates data management and generates accurate and valuable information
  - Poorly designed database: causes difficult-to-trace errors that may lead to poor decision making

Evolution of File System Data Processing

Old manual systems:

- Spreadsheets
- Templates
- Text/Word Processor files
- Stored images (scans, faxes, hand-written, etc.)

File Systems:

- Lineal record system
- Data, but basic storage only
- Difficult to manage, change, query, protect

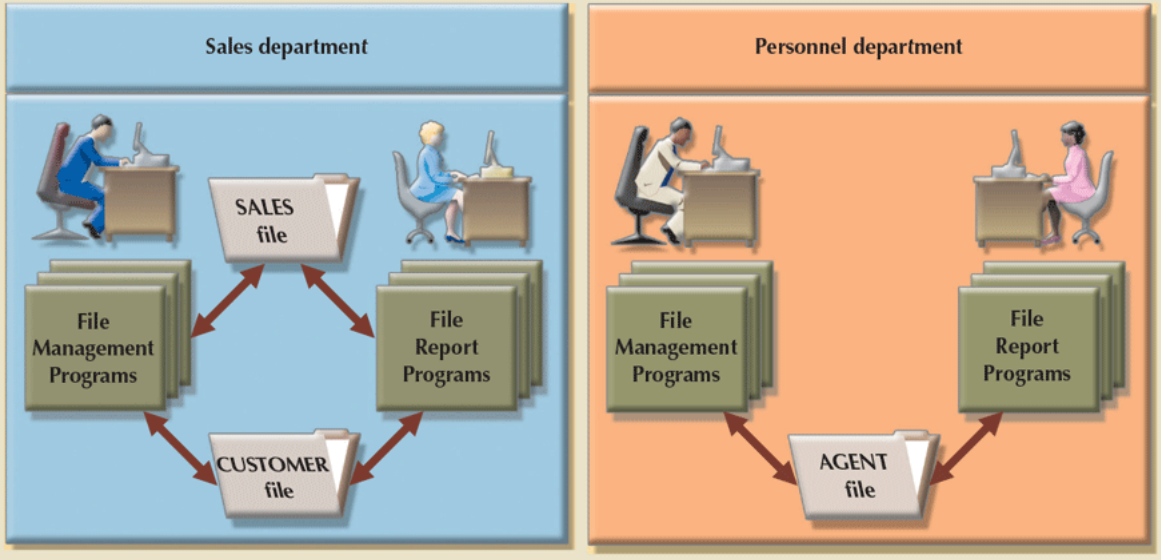
C:\data\YYC\2014\Sales.dat
C:\data\YYC\2015\Sales.dat
C:\data\YYC\2016\Sales.dat

C:\2014\YQL\Sales.dat
C:\2015\YQL\Sales.dat

C:\Sales\2010\YYC.dat

Evolution of File System Data Processing

FIGURE 1.9 A SIMPLE FILE SYSTEM



Evolution of File System Data Processing

Table 1.2	Basic File Terminology
TERM	DEFINITION
Data	Raw facts, such as a telephone number, a birth date, a customer name, and a year-to-date (YTD) sales value. Data has little meaning unless it has been organized in some logical manner.
Field	A character or group of characters (alphabetic or numeric) that has a specific meaning. A field is used to define and store data.
Record	A logically connected set of one or more fields that describes a person, place, or thing. For example, the fields that constitute a record for a customer might consist of the customer's name, address, phone number, date of birth, credit limit, and unpaid balance.
File	A collection of related records. For example, a file might contain data about the students currently enrolled at Gigantic University.

ID	First Name	Last Name	Age	Phone	Email
01	Ricky	Ponting	37	121321	abc@abc
02	Ben	Stokes	35	178214	def@def
03	Eoin	Morgan	36	321324	ghi@ghi

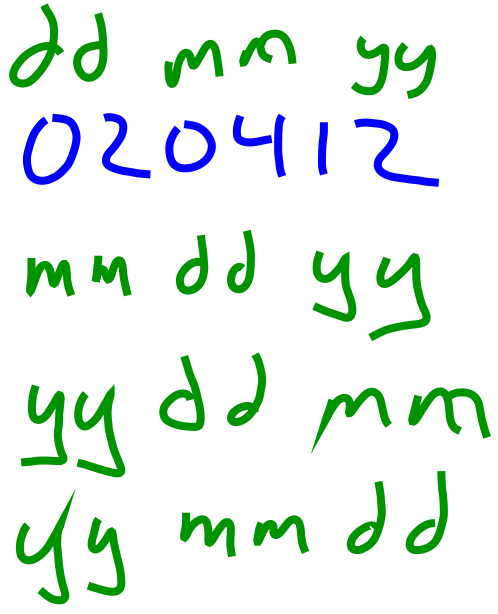
CENGAGE

Structural and Data Dependence (1 of 2)

- Structural dependence
  - Access to a file is dependent on its own structure
  - All file system programs are modified to conform to a new file structure
- Structural independence
  - File structure is changed without affecting the application's ability to access the data

Structural and Data Dependence (2 of 2)

- Data dependence
  - Data access changes when data storage characteristics change
- Data independence
  - Data storage characteristics are changed without affecting the program's ability to access the data



Data Redundancy

- Same data is stored unnecessarily in multiple places, or multiple instances in the same place
- Results in poor security, inconsistency, storage overload

Data Anomalies

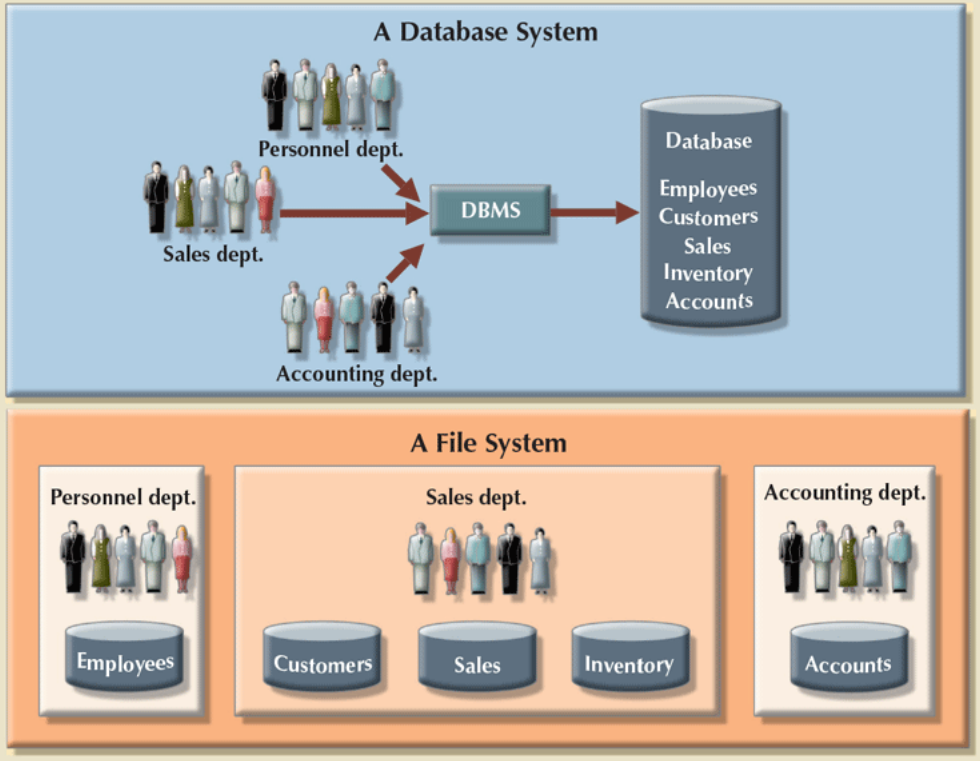
- Develop when not all of the required changes in the redundant data are made successfully
  - Update anomalies
  - Insertion anomalies
  - Deletion anomalies

Database Systems (1 of 2)

- Logically related data stored in a single logical data repository
  - Physically distributed among multiple storage facilities
  - DBMS eliminates most of file system's data inconsistency, data anomaly, data dependence, and structural dependence problems
- Current generation DBMS software
  - Stores data structures, relationships between structures, and access paths
  - Defines, stores, and manages all access paths and components

Database Systems (2 of 2)

FIGURE 1.10 CONTRASTING DATABASE AND FILE SYSTEMS



The Database System Environment (1 of 2)

- Database system: organization of components that define and regulate the collection, storage, management, and use of data within a database environment
  - Hardware
  - Software
  - People
  - Procedures
  - Data

Database Systems

Composed of 5 Major Parts

1. **Hardware**
2. Software
3. People
4. Procedures
5. Data

Hardware (The Physical Devices)

- Servers
- Workstations
    - Desktops
    - Laptops
- Handheld
- Network
- Peripherals

Database Systems

Composed of 5 Major Parts

1. Hardware
2. **Software**
3. People
4. Procedures
5. Data

Software

Three Types:

1. Operating Systems (OS)
    1. Windows
    2. Unix/Linux
    3. MacOS
    4. Android
2. DBMS
    1. Oracle
    2. MS Access
    3. MSSQL
    4. MySQL
3. Applications
    1. Canvas
    2. MS Office
    3. Simply Accounting

Database Systems

Composed of 5 Major Parts

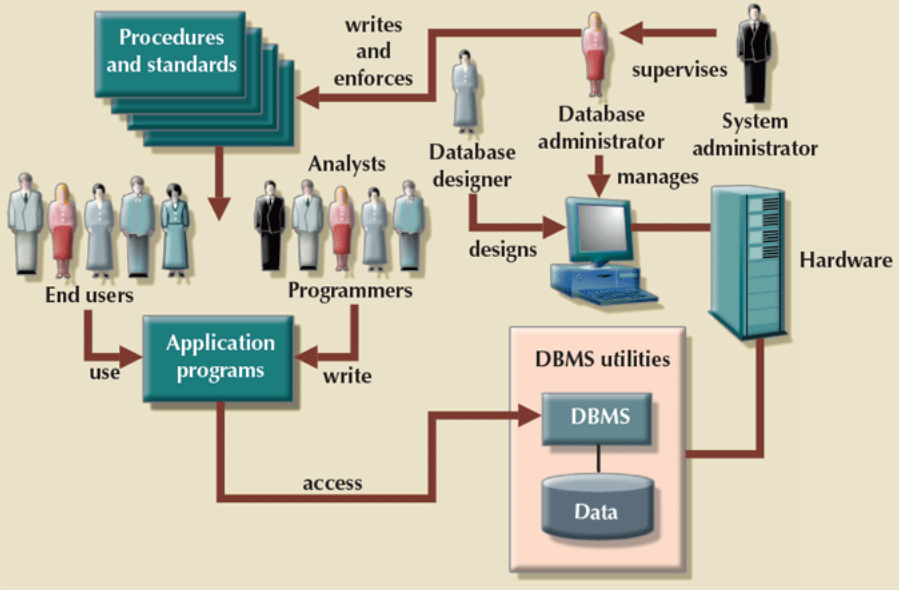
1. Hardware
2. Software
3. **People**
4. Procedures
5. Data

5 Primary User Types

1. System Administrators
2. Database Administrators (DBA)
3. Database Designers
4. System Analysts/Programmers
5. End Users

The Database System Environment (2 of 2)

FIGURE 1.11 THE DATABASE SYSTEM ENVIRONMENT



Database Systems

Composed of 5 Major Parts

1. Hardware
2. Software
3. People
4. **Procedures**
5. Data

The Rules that govern design & use

- Security
- Permissions
- Access
- Design
- NORMALIZATION

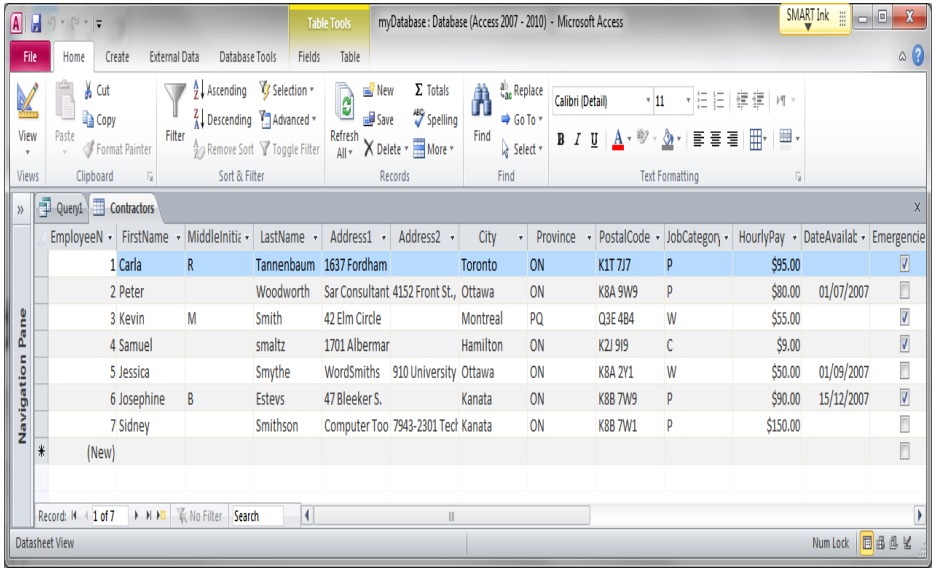
Database Systems

The collection of stored facts

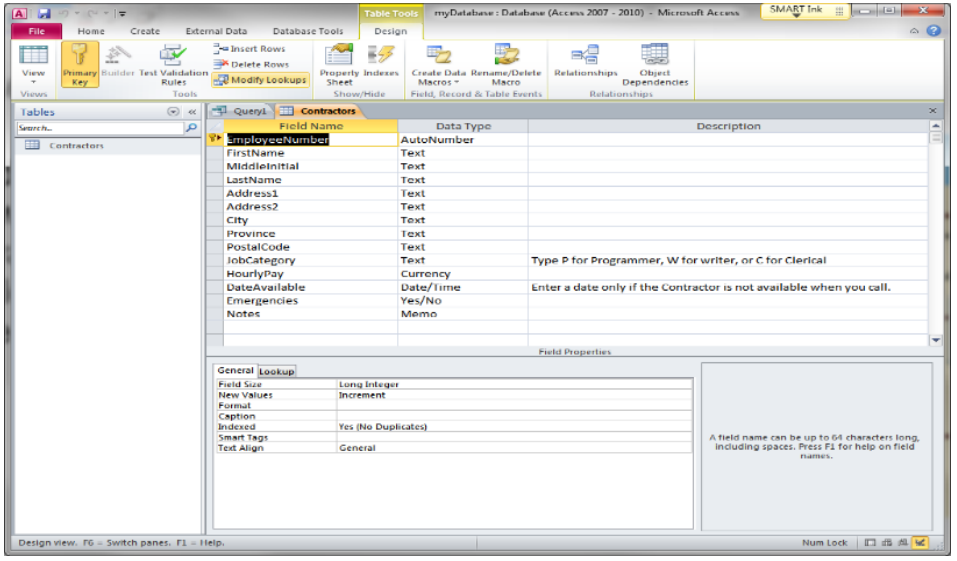
Composed of 5 Major Parts

1. Hardware
2. Software
3. People
4. Procedures
5. **Data**

Data



Metadata



DBMS Functions

Data Dictionary Management
Data definitions and relations

Data Storage Management
Data, forms, reports, validations

Data transformation and presentation
e.g. date formats

Security Management
Who has access and to what?

DBMS Functions

Multi-user access control

Backup and Recovery management

Data integrity management
- Data are "accurate" & "verifiable"
- min. redundancy & max. consistency
- transactional management

Language, API, interfaces
- SQL

```
SELECT lastname, age
  FROM Student
  LEFT JOIN Crs_Reg
  ON Student.id = Crs_Reg.std
  WHERE Crs_Reg.crs = "CIT1163";
```

Managing the Database System: A Shift in Focus

- Disadvantages of database systems
  - Increased costs
  - Maintaining currency
  - Vendor dependence
  - Frequent upgrade/replacement cycles

Preparing for Your Database Professional Career

TABLE 1.3	DATABASE CAREER OPPORTUNITIES	
JOB TITLE	DESCRIPTION	SAMPLE SKILLS REQUIRED
Database Developer	Create and maintain database-based applications	Programming, database fundamentals, SQL
Database Designer	Design and maintain databases	Systems design, database design, SQL
Database Administrator	Manage and maintain DBMS and databases	Database fundamentals, SQL, vendor courses
Database Analyst	Develop databases for decision support reporting	QL, query optimization, data warehouses
Database Architect	Design and implementation of database environments (conceptual, logical, and physical)	DBMS fundamentals, data modeling, SQL, hardware knowledge, etc.
Database Consultant	Help companies leverage database technologies to improve business processes and achieve specific goals	Database fundamentals, data modeling, database design, SQL, DBMS, hardware, vendor-specific technologies, etc.
Database Security Officer	Implement security policies for data administration	DBMS fundamentals, database administration, SQL, data security technologies, etc.
Cloud Computing Data Architect	Design and implement the infrastructure for next-generation cloud database systems	Internet technologies, cloud storage technologies, data security, performance tuning, large databases, etc.
Data Scientist	Analyze large amounts of varied data to generate insights, relationships, and predictable behaviors	Data analysis, statistics, advanced mathematics, SQL, programming, data mining, machine learning, data visualization