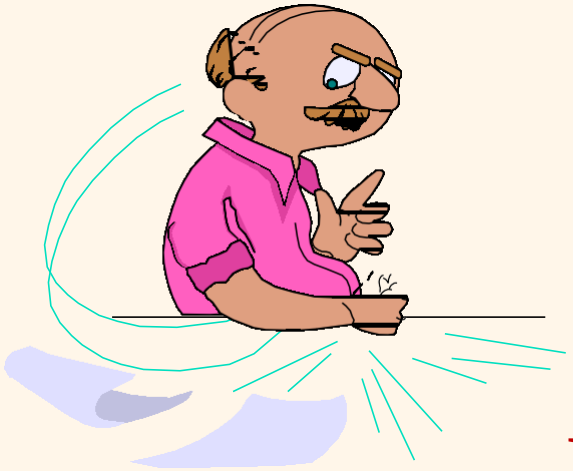


Introduction to Databases

Hafiz Tayyeb Javed

Assistant Professor (CS)

Department of Computer Science

FAST-NU, CFD Campus

Course Particulars

- Credit Hours:
  - Four
- Course Structure:
  - Two Lectures a week (Each of duration 1.5 hour).
  - Lab class each week

Recommended Text/Reference Books

- Fundamentals of Database Systems,
  - 3rd Edition by Elmasri & Navathe
- Database System Concepts
  - 3rd Edition by Abraham Silbershatz, Henery F. Korth, Sudarshan
- An introduction to Database Systems
  - By C.J. Date

Evaluation and Grading (Tentative)

Quizzes	10
Assignments	10
Mid Exam 1&2	30
Class Participation	10
Final Exam	40

Quizzes and Assignments

- Quizzes
  - Frequent quizzes of duration 10-15 minutes will be taken.
  - Students are required to attend the classes regularly and come prepared in each class.
- Assignments
  - No assignments will be accepted after due date.
  - Programming assignments should be well documented.
  - Students are “not” allowed to “copy” each other’s work. Any such work would be marked zero.

Grading Policy

- All deadlines to be strictly followed
- After deadline, any submitted assignments/ Project Deliverable will be marked as zero
- Rechecking can be requested after grade reporting, with in 2 days.

Programming Skills

- Front-End Development
  - ASP.NET using C#/VB.Net, C++, PHP
- Back-End Development
  - MS-Access, Oracle 9i or higher, MS SQL Server
- Documentation
  - ER-Win, Visio, MS-Word

Types of Databases

- Flat Files
- Relational
- Object Relational
- Web Enabled (Online DBs)

Overview

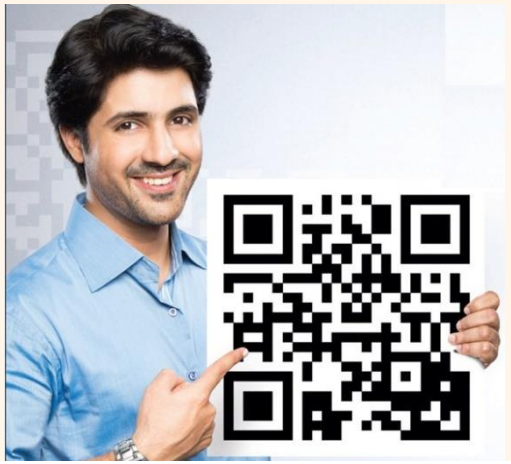
Data: *Known facts that can be recorded*

Database: *Collection of Integrated data*

Typically models a real-world “enterprise”

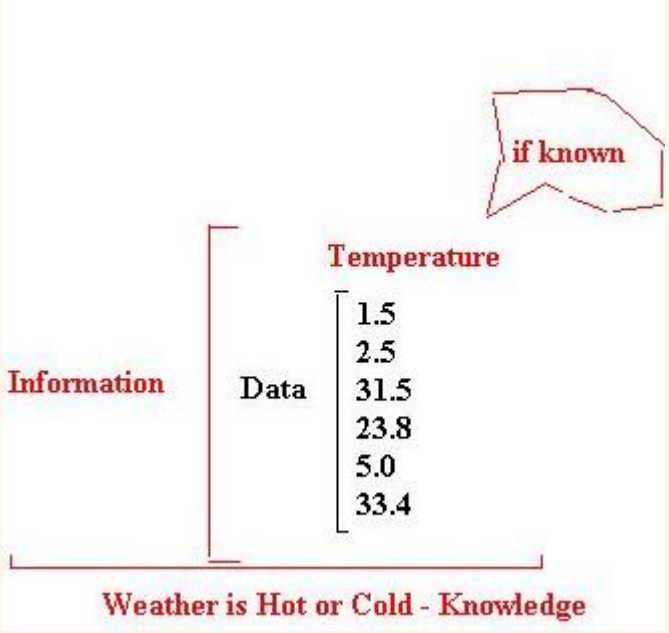
DBMS: *A software system designed to store, manage, and facilitate access to databases.*

Data vs. Information



- Data: Raw facts/Un-processed information
  But they are building blocks for information

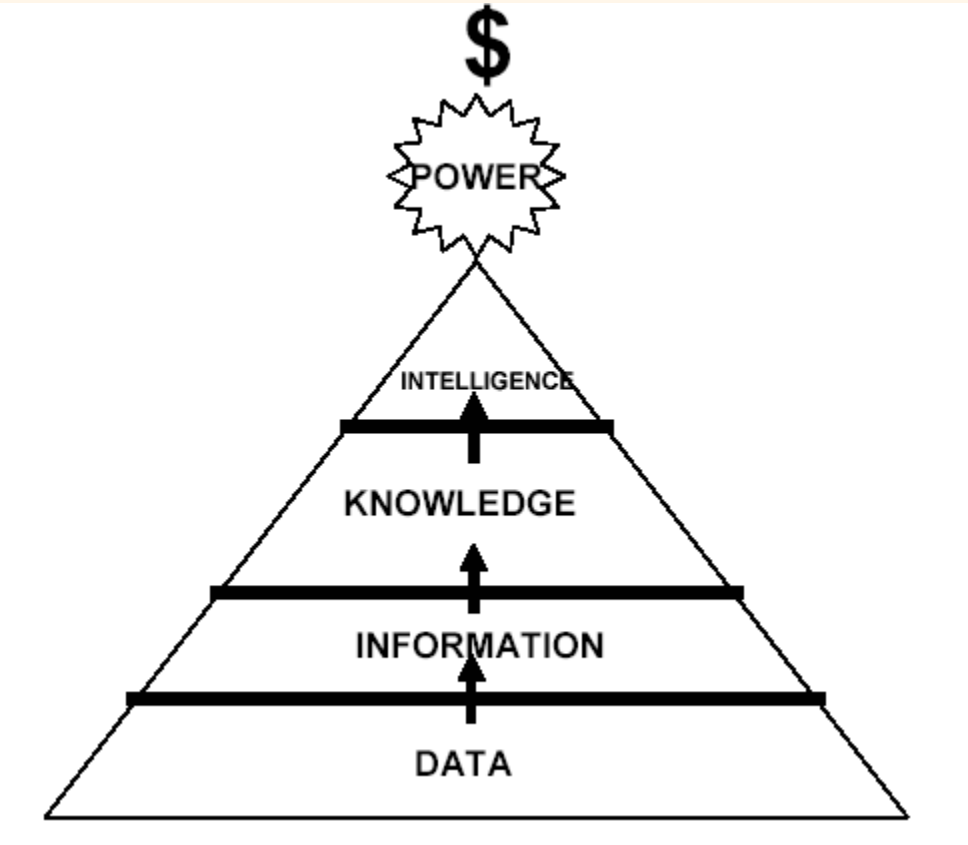
- Information: Data Processed to reveal its meaning
  Information is meaningful
  In today's world, accurate, relevant and timely information is the key to good decision making
  Good decision making is key to survival in today's competitive and global environment



Data and Information …

- Data: is known recorded facts which has specific meanings or interpretation.
  Example: the numbers 10.5, 22, and 119 are data
- Information: is a precise, understandable and specific representation of data.
  Example: The temperature of room 119 in building 22 is 10.5, is an information
- Knowledge: is something which is derived or inferred from available information using some level of intelligence.
  Example: Based on experience, the above information can be used to infer that the room is quite cold and could cause some inconvenience if we work on that room for a long period of time with out wearing warm clothes.

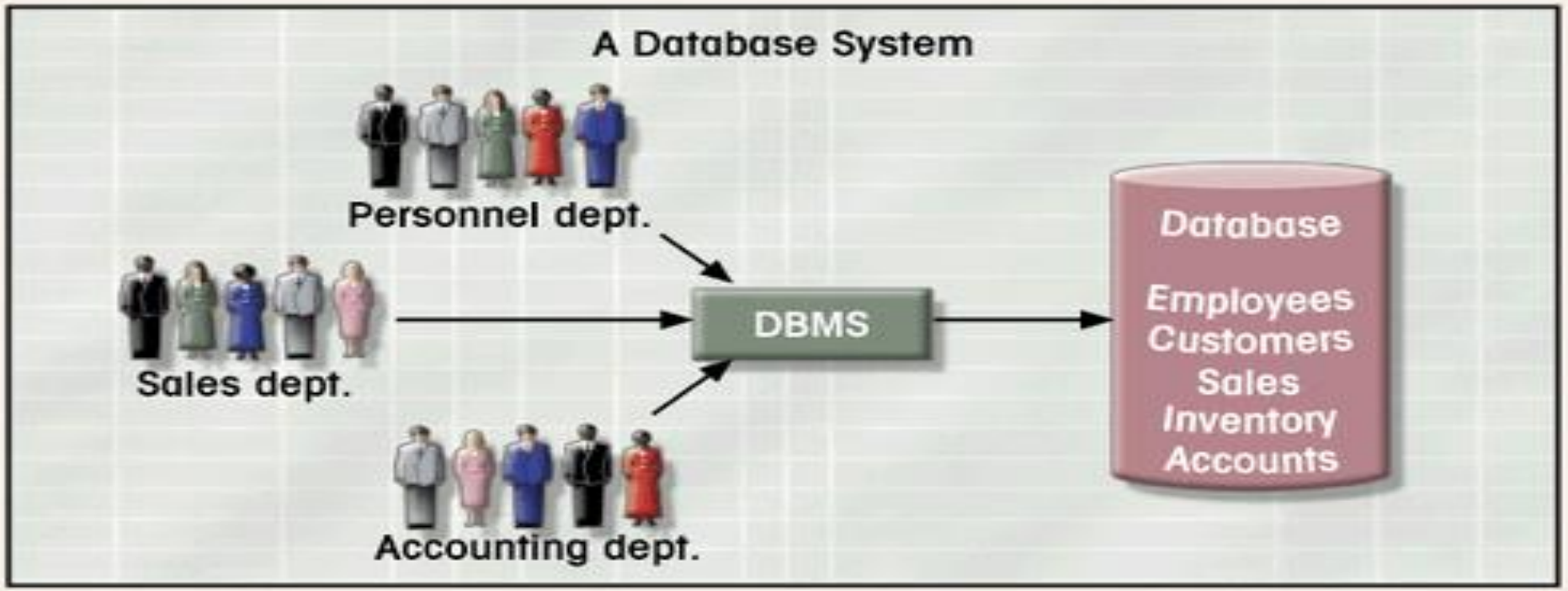
The Need

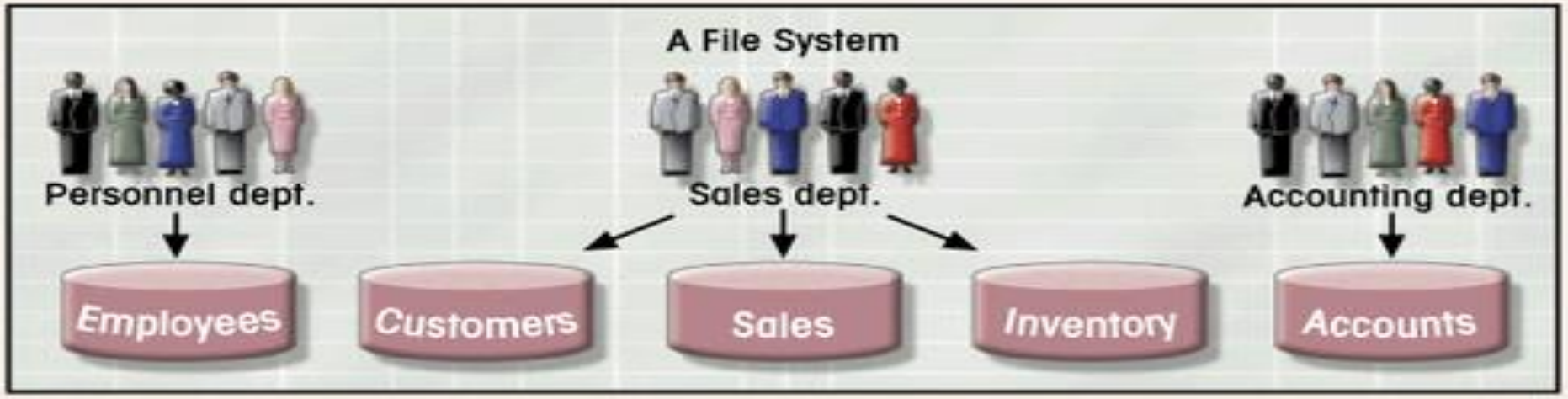


Files and Databases

- Files: A collection of records or documents dealing with one organization, person, area, or subject. It could either be:
  - Computer files
  - Manual files (Paper)
- Database: A collection of similar records with relationships between records.
  - Statistical, Business Data

Database vs. File Systems





Database

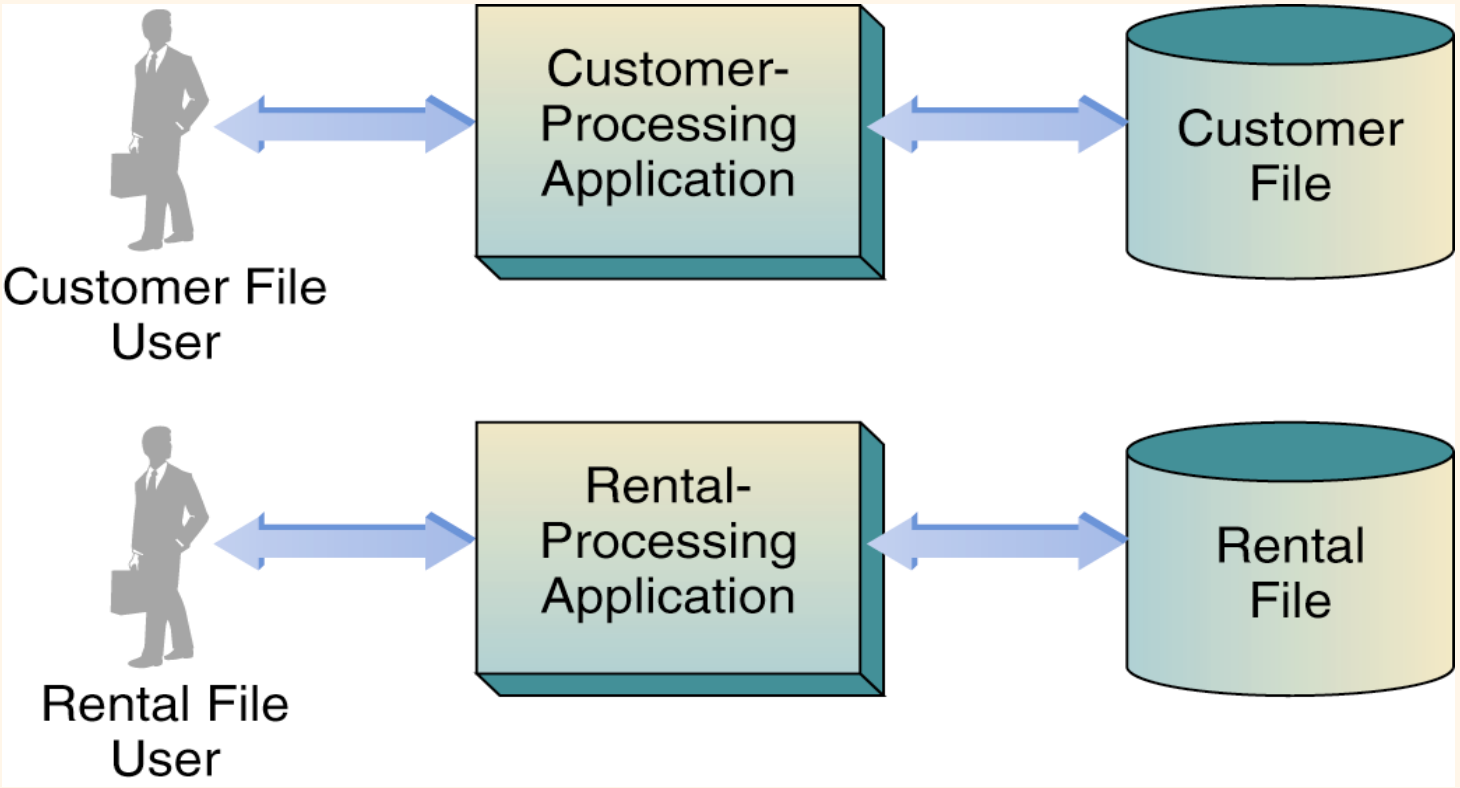
A database is a collection of stored operational data used by application systems of some particular enterprise (C.J. Date)

- Paper Database
  - Still contain a large portion of world's knowledge
- File Processing Systems
  - Early batch processing of business data
- Database Management Systems (DBMS)
  - Will cover in detail

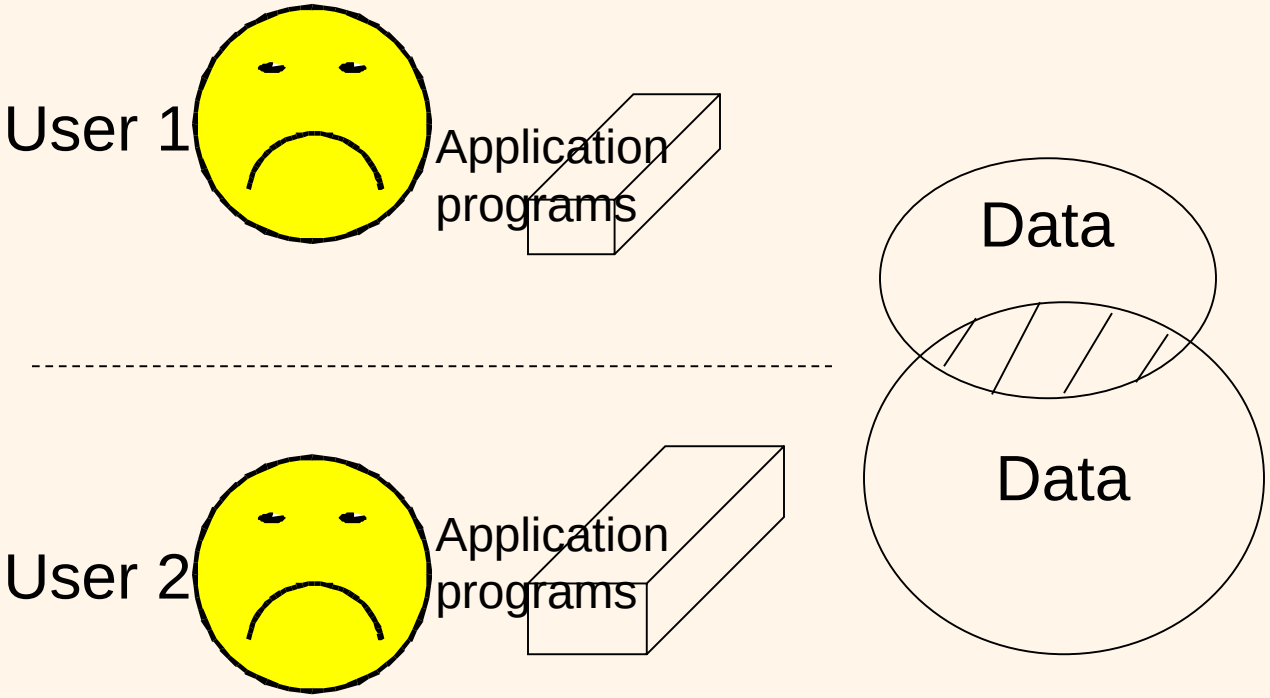
File Processing Systems

Collection of application programs that performs services for the end-users (e.g. Reports)
Each program defines and maintains its own data

File Processing Systems



-- *File-Based Approach*



Limitations of File-based Approach

- **Separation and isolation of data**
  Each program maintains its own set of data. Users of one program may be unaware of potentially useful data held by other programs.

- **Duplication of data**
  Same data is held by different programs. Wasted space and potentially different values and/or different formats for the same item.

- **Atomicity of updates**
  Failure may lead database to an inconsistent state with partial updates carried out
  E.g. Transfer of funds from one account to another should either be complete or incomplete, or may be in in-between state

Limitations of File-based Approach

- **Data dependence**
  File structure is defined in the program code.

- **Incompatible file formats**
  Programs are written in different languages, and so cannot easily access each others files.

- **Integrity Problems**
  Integrity constraints (e.g. account balance > 0) became "buried" in program code rather than being stated explicitly
  Hard to add new constraint or update existing one
  File in the folder cannot keep such conditions

Limitations of File-based Approach

- **Fixed Queries/Proliferation of application programs**
  Programs are written to satisfy particular functions. Any new requirement needs a new program
- **Data Redundancy and Inconsistency**
  Multiple file formats, and duplication of information in different files
- **Difficulty in Accessing Data**
  Need to write a new program for accessing new data

Database Approach

- Overcomes problems associated with file-system based approach
- Central repository of shared data
- The database holds not only the data but also a description of the data.
  - Data dictionary, or metadata
    - A central location where data descriptions are stored
    - Data about data
  - Program-data independence

Advantages of the Database Approach

- Program-data independence
  - The separation of data descriptions from the application programs that use the data
  - Allows the data to change without changing the application programs
- No or Planned data redundancy
- Improved data consistency
- Improved data sharing
- Enforcement of standards