

Overview of Database Systems

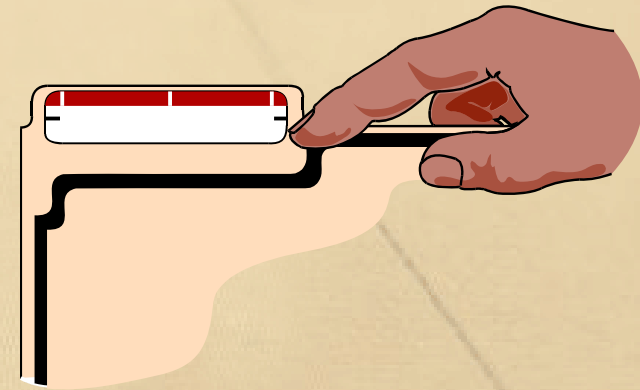
Introduction

- When computers were first used for business applications, they simply performed regular record-keeping functions
- These were called **file processing systems**



File Systems

- Prior to the emergence of database systems, organizational data were kept in numerous computer files
- Information in a folder may be the same as records in a computer
- Files are created and managed independently
- Application programs are used to manage the interrelationships of files when there is need to link data from different files



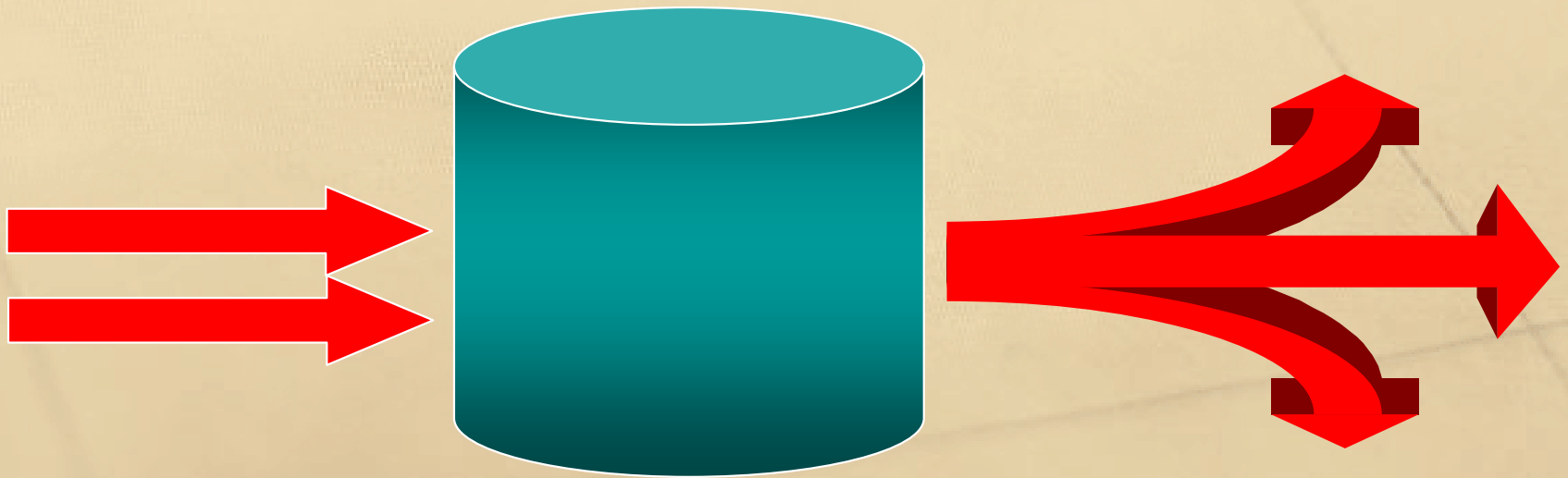
Technology Today

- Organizations today make use of different types of information systems that require well-organized database systems
 - Marketing
 - Credit Cards and Banking
 - Airlines
 - Hospitals
 - Schools
 - Groceries
 - and many more...



Database

- A **database** is an organized collection of logically related data that can be processed by multiple application systems and can be shared by multiple users.



Data

- Traditional Definition

- Facts concerning objects and events that can be recorded and stored on computer media

- Structured Data

- Numeric values, characters, dates
 - Stored in tabular form

- Unstructured Data

- Documents, maps, photos, sound, video

- Data

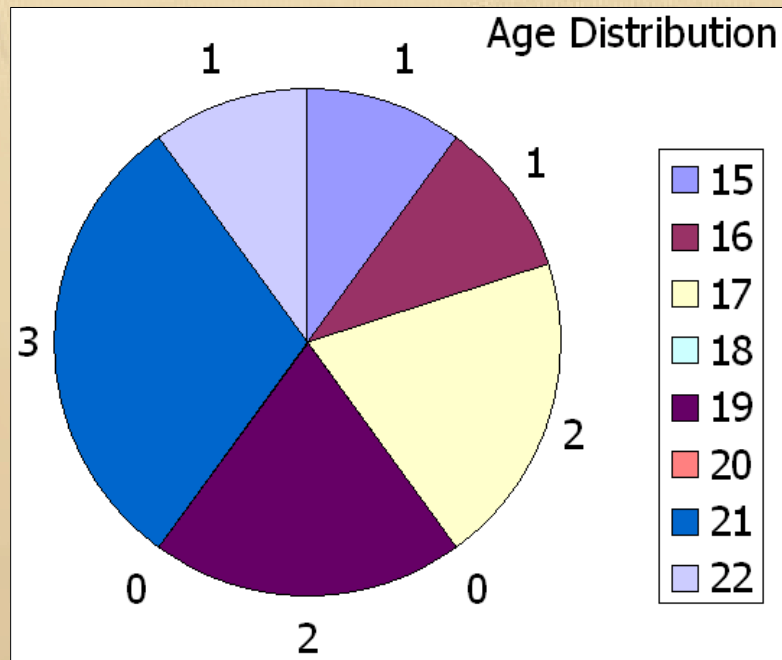
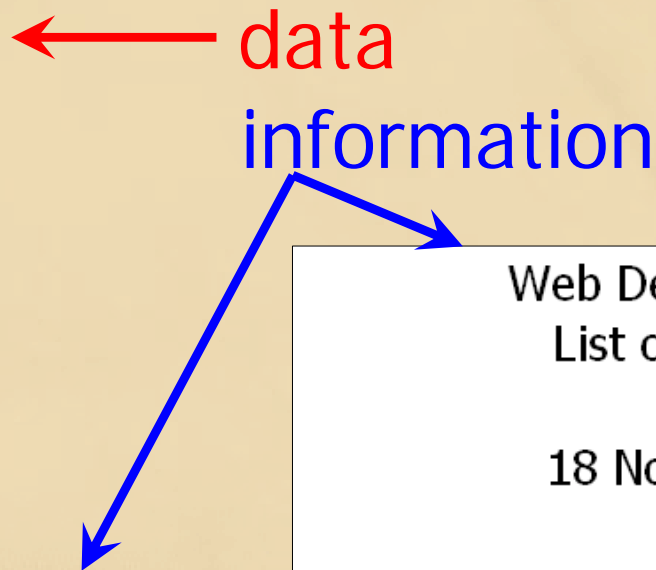
- Stored representations of objects and events that have meaning and importance in the user's environment

Information

- There is a difference between **data** and **information**.
- Information
 - Data that have been processed in such a way as to increase the knowledge of the person who uses the data

Data and Information

Aquino	Zachary	20101375	17
Cordova	Ysabel	20101532	16
Echevarria	Vicente	20101328	21
Galicia	Teresa	20101542	15
Lozano	Pedro	20100280	19
Prieto	Lucas	20100524	21
Robles	Joaquin	20101823	21
Sanchez	Hector	20100342	22
Yamada	Clara	20101802	17
Zabala	Angelina	20101713	19



Web Design Workshop
List of Participants

18 November 2007

<u>Last Name</u>	<u>Given Name</u>	<u>ID Number</u>	<u>Age</u>
Aquino	Zachary	20101375	17
Cordova	Ysabel	20101532	16
Echevarria	Vicente	20101328	21
Galicia	Teresa	20101542	15
Lozano	Pedro	20100280	19
Prieto	Lucas	20100524	21
Robles	Joaquin	20101823	21
Sanchez	Hector	20100342	22
Yamada	Clara	20101802	17
Zabala	Angelina	20101713	19

Metadata

- Data that describe the properties or characteristics of end-user data and the context of that data

Data Item		Value
<u>Name</u>	<u>Type</u>	<u>Length</u>
Workshop	Text	100
Workshop Date	Date	12
Last Name	Text	30
Given Name	Text	30
ID Number	Numeric	8
Age	Numeric	2

Database Management System

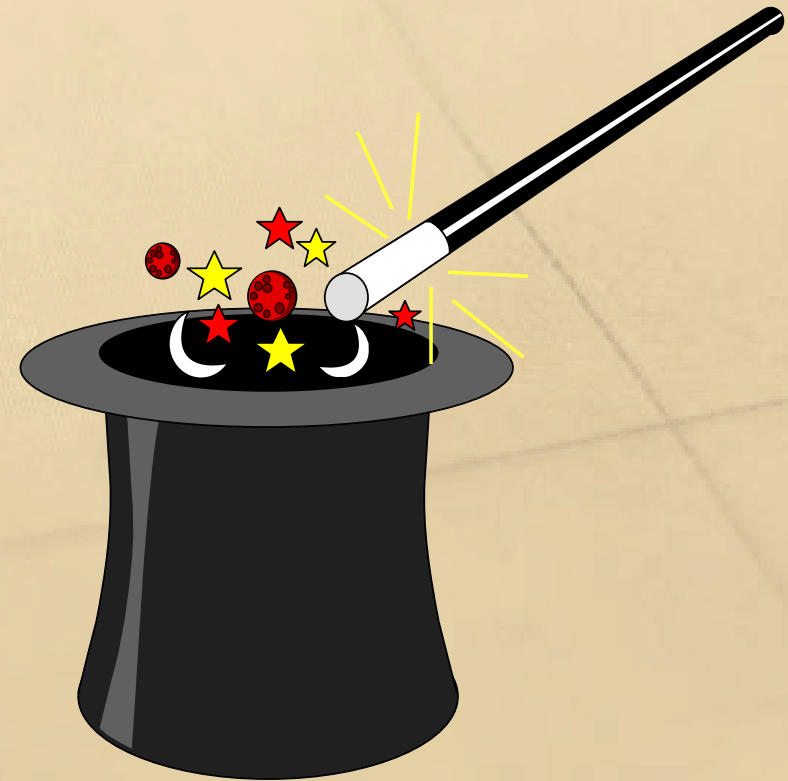
- A **database management system (DBMS)** is a software system that is used to create, maintain, and provide controlled access to user databases.
- The DBMS allows the developer to create the structure of the data in a database.
- The DBMS allows the end-user to manipulate the data in a database.

Database System

- A **database system** is composed of the data stored in the database, the DBMS, the software and applications, hardware, and people
- Software – DBMS, CASE Tools, application software and utilities, report writers, user interface, operating systems
- People – data and database administrators, systems developers, programmers, system administrators

Scenarios

- Changing a student's address
- Retrieving a selected list of student information
- Revising reports (addition of new fields)



Disadvantages of File Systems

- Program-Data Dependence
- Data Redundancy – waste of storage space, multiple entries (duplication) of the same data item
- Data Inconsistency & Poor Data Control – lack of integrity, synonyms & homonyms
- Limited Data Sharing
- Lengthy Development Times
- Excessive Program Maintenance

Advantages of Database Systems

- Program-Data Independence
- Minimal Data Redundancy
- Improved Data Consistency
- Improved Data Sharing
- Increased Productivity of Application Development
- Enforcement of Standards
- Improved Data Quality
- Improved Data Accessibility and Responsiveness
- Reduced Program Maintenance
- Improved Decision Support

Limitations of Database Systems

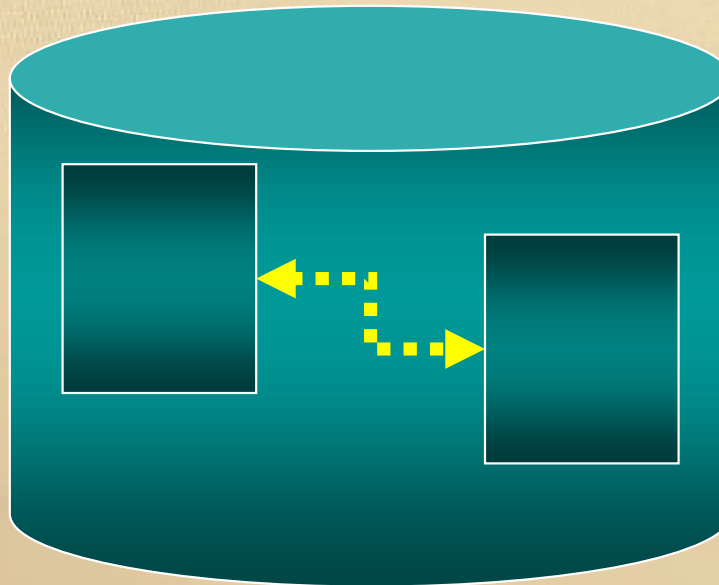
- Specialized Personnel
- Installation and Management Costs and Complexity
- Conversion Costs
- Need for Explicit Backup and Recovery
- Organizational Conflict

Features of DBMS

- Cross-Entity Linking Mechanism
- Data Dictionary
- Query Language
- Mechanisms for Data Access
- Host Language Interface Software
- Data Manipulation Language
- 4GL Application Development Software
- Disaster Recovery & Concurrency Control

Linking Mechanism

- Allows related entities in a database to be linked together
- No need to keep multiple copies of the same data



Data Dictionary

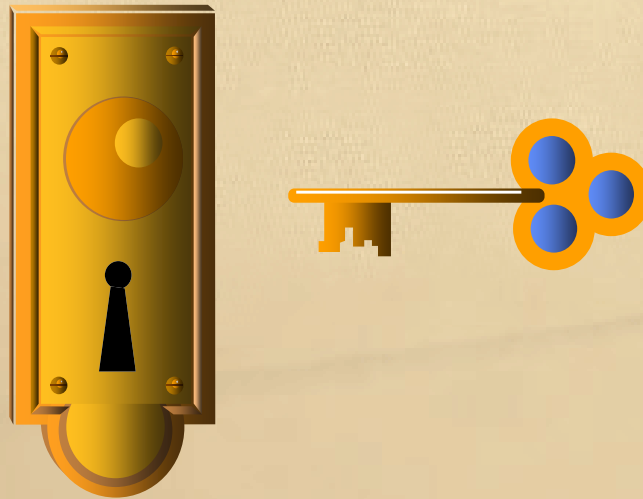
- Describes the structure of data and how they are to be used
- Contains metadata, or data about data
- Also known as **system catalog**
 - a system-created database that contains the characteristics and contents of user-created databases

Query Language

- A language that is used to retrieve selected data from a database
- Used by end-users for ad-hoc queries
- **Structured Query Language (SQL)**
 - used by most relational DBMS as a standard method of issuing database commands

Control for Data Access

- Mechanisms to control access to data
- Protects data integrity and privacy by controlling the actions that users can perform on the data



Host Language Interface

- Software that allows database commands to be combined with standard programming language commands to create applications
- Simplifies programming by allowing the database to handle the data-oriented portions of the application

Data Manipulation Language and 4th Generation Language

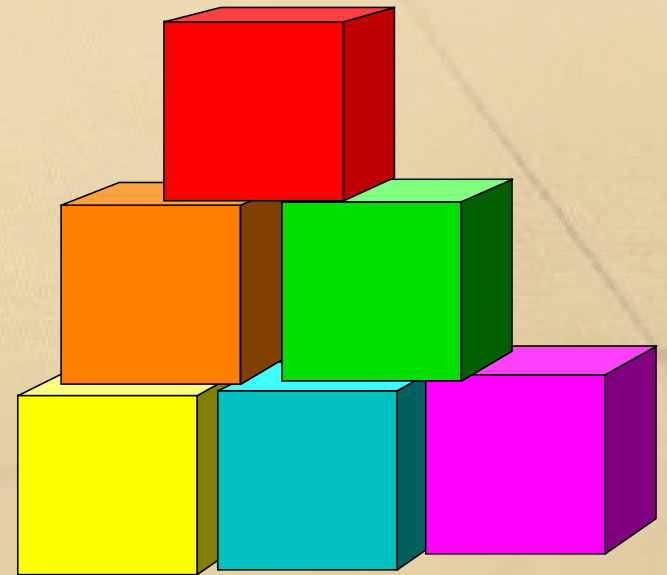
- Data Manipulation Language (DML)
 - enables users to access and manipulate data in a database
- 4GL Application Development Software
 - A visual interface software to develop applications
 - Form and report generators

Disaster Recovery and Concurrency Control

- Software that provides mechanisms to recover damaged data and control concurrent access
- Concurrent access
 - the simultaneous access by multiple users
- Includes maintenance of a transaction log that contains the list of transactions and changes made to the database

Data Models

- **Data models** are conceptual methods of structuring data
- Four (4) data models for implementing database systems:
 - Hierarchical model
 - Network model
 - Relational model
 - Object-oriented model

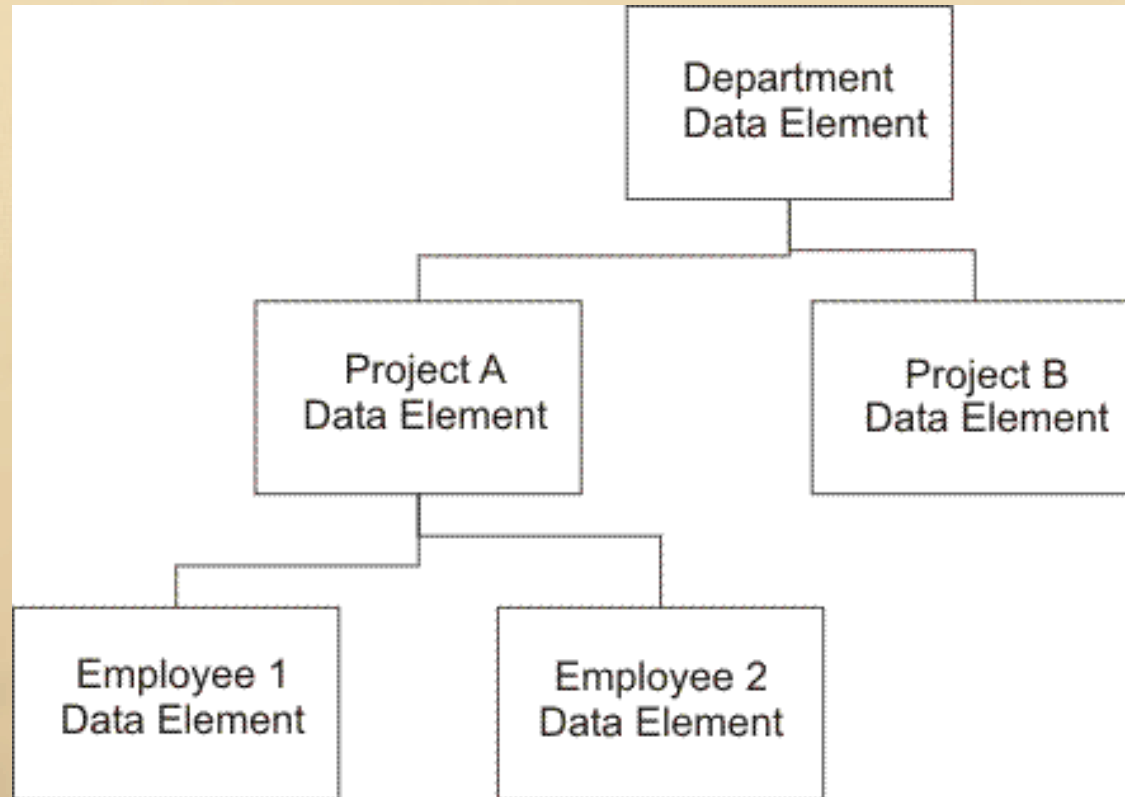


Differences Between Data Models

Model	Description
Hierarchical	Relationships between data are structured as hierarchies; uses pointer-based retrieval
Network	Records are organized as collections of arbitrary graphs; uses pointer-based retrieval
Relational	Uses a collection of tables to represent data and relationships among data; links data by unique columns; uses user-friendly query language
Object-Oriented	Combines data and methods/processes used with the data; features the concept of inheritance of properties

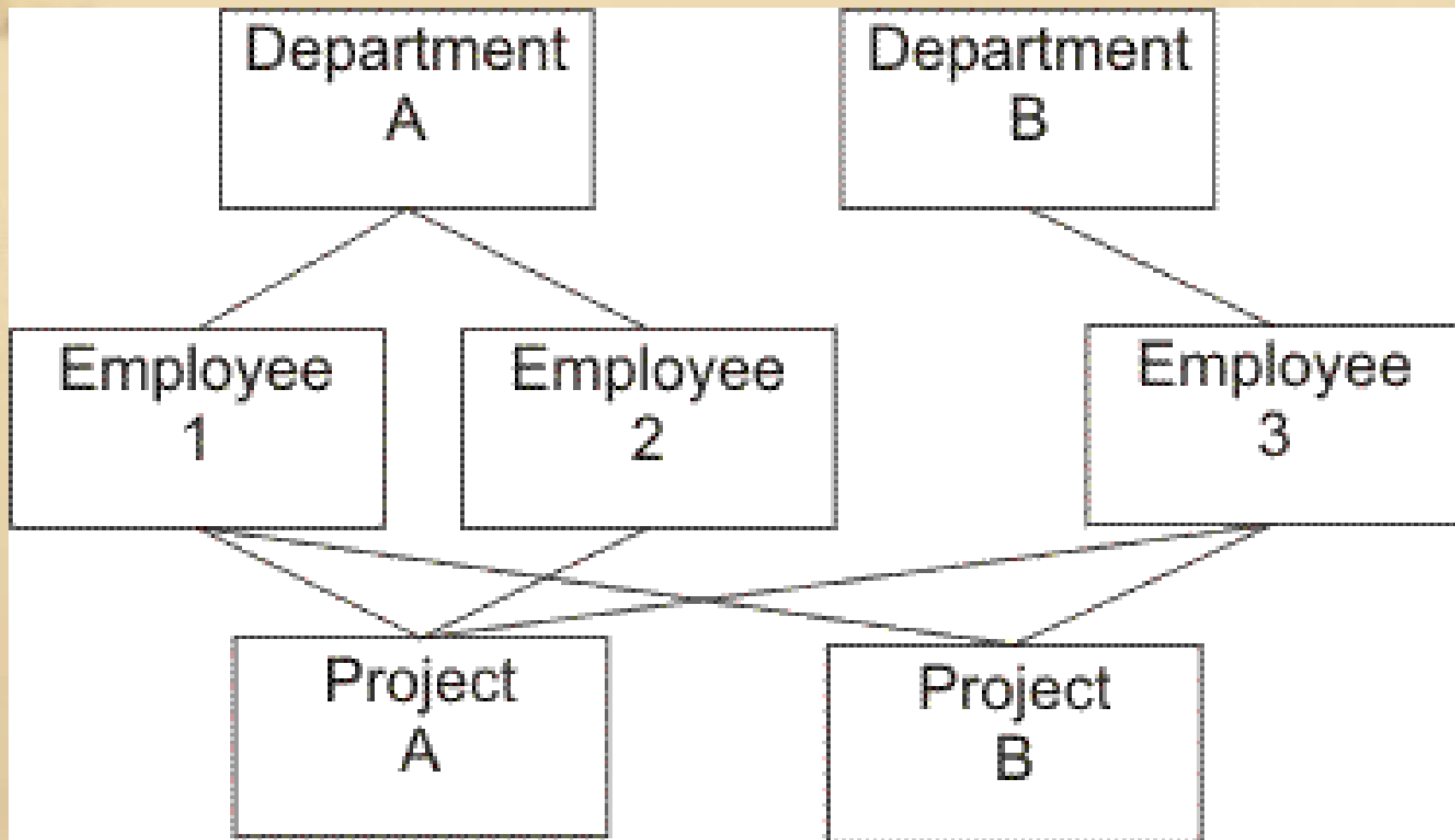
Hierarchical Model

- Treelike structure
- One-to-many relationship
- Highest element is called the root element



Network Model

- Many-to-Many relationships



Hierarchical and Network Models

- Both models use highly efficient pointer-based retrieval methods
- Do not provide a user-friendly mechanism to access data
- Relationships between data have to be determined before they are placed into operation
- Inflexible to changes in the structure of a database

Relational Model

- Most popular
- Data elements are stored in tables
- Relations are illustrated by linking or joining the two tables with a common element

Department Table			
Deptno	Dname	Dloc	Dmgr
Dept A			
Dept B			
Dept C			

Employee Table				
Empno	Ename	Etitle	Esalary	Deptno
Emp1				Dept A
Emp 2				Dept A
Emp 3				Dept B
Emp 4				Dept B
Emp 5				Dept B
Emp 6				Dept C

Relational Model

- Less machine efficient than pointer-based retrieval
- Can accommodate changes in the database structure easily
- Non-technical users can access the database and retrieve selected data through query languages

Relational Model

- Increased the availability of information in corporate databases
- More flexible and easy to use
- Improved the conceptual efficiency of database programming
- Most prevalent model in used today

Object-Oriented Model

- Emerging technology
- Data and the methods that can be performed on the data, are stored in objects
- Objects can inherit the properties of higher level objects
- Have been implemented primarily in engineering rather than business-oriented systems

Types of Databases

- Databases can also be categorized according to the degree of centralization or decentralization of processing that they support:
 - Centralized
 - Personal
 - Client-server
 - Distributed

Centralized Database Systems

- All data are located at a single site and controlled by a single computer
- The central computer does all the data processing activities, data access and retrieval, programming, and conversion of results to the form required by the user
- Remote users can access the database using terminals or data communications links
- Vulnerable to failure

Personal Database Systems

- Developed on a personal computer, normally used and maintained by a single user
- These databases were designed to be easy for non-technical users
- Risk of creating “islands of information” where data cannot readily be shared

Client/Server Database Systems

- Designed to support cooperative processing
- The database is stored on a server, which handles all direct database operations
- Much of the processing can be moved to the client computers (PCs connected to the network)

Client/Server Database Systems

- Reduces workload on database servers
- Allows higher volumes of transactions to be handled with the server's capacity
- Provides greater flexibility to expand systems
- Gives individual users greater control over the system

Distributed Database Systems

- A single logical database that is physically spread across multiple computers
- Data are stored at the location where it is most frequently used
- Minimizes the load on communications network and gives users a sense of ownership over the data they use most frequently
- Adds complexity to database processing