Fundamentals of Databases
CSU 07203

The Institute of Finance Management
FCIM

BCS & BIT

Year I -2015/2016

BY

A. S. Siphy

Logistics

Instructor: Siphy, A. S(Mr.)

email: dullextz@gmail.com

Office: Block D, 020

Consultation Time

Mondays

10:00 am -12:00

Or

By appointment

Course Overview

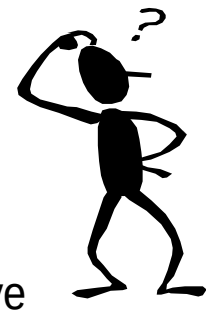
- Description of the Course (fact sheet)
- Assignments & projects (fact sheet)
- Readings (fact sheet)
- Grading (fact sheet)
- Schedule -as per your timetable….

Course Description

- This course is concerned with the design of the database itself -- not with the design of database system software.!!!
  - We will discuss DBMS internals only as they relate to the database and its design and structure
- Developing and managing efficient and effective database applications – requires understanding the fundamentals of databases, database management systems, techniques for the design of database apps, database systems , understanding new development trends etc.

Assessment

- Two kinds of assignments and projects
  - Using a database modeling approach, designing entity relationship diagrams, conceptual and logical designs
  - Designing, populating, and running queries against your *own personal* database in MS Access
    - Types of database projects
      - Individual
        - » Work related….
        - » Course only….
      - **Group**
        - » Course related// collaborative
        - » Projects from around campus that need to be done……!
        - » Min projects – Case studies and scenarios (based on ERM) about database design -to be done in MS Access Database and other listed apps

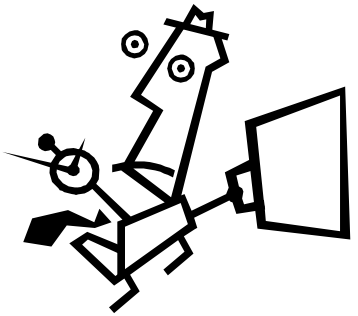


4/06/2016 5

Grading

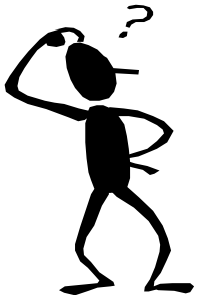
- You will be assessed through continuous assessments (Coursework) that comprise of two compulsory Tests, Lab test and assignments, min projects etc.

- CW carries **40%**
- Final Exam carries **60%**



Readings

- Ramez Elmasri and Shamkant B. Navathe (2006), *Fundamentals of Databases Systems*, 5th Edition, Pearson / Addison Wesley
- Hector Garcia-molina, Jeffrey D. Ullman and Jennifer Widom(2008), *Database Systems*: The Complete Book, 2nd Edition, Prentice Hall
- David Kroenke (2002), *Database Processing: Fundamentals, Design and Implementation*, 8th Edition, Prentice Hall, Upper Saddle River NJ



Today's buzzwords

- Organizational data cycle
- Database concepts
- Database Management Systems
- Data Independence
- Database systems
  - Centralized Database Systems
  - Client-Server Database Systems
  - Distributed Database Systems

Objectives of today's lecture

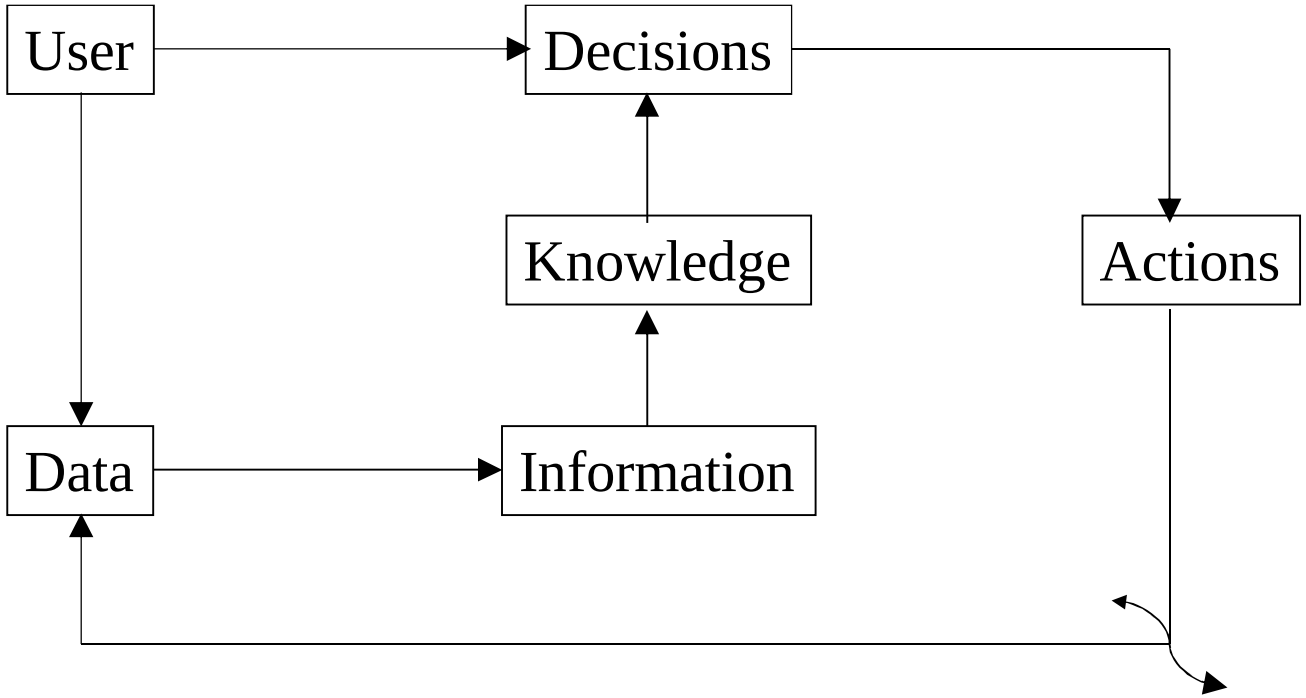
- Know common database terminologies and basic concepts
- Understand requirements and uses of data
- Know the differences between databases and flat files (legacy –file-based system)
- Realize the importance and need for databases in problem-solving
- Understand the different types of databases and their differences

Data: A Resource

- The Success of an organization depends on efficient use of its resources:
  - Buildings, factories, equipment
  - Technical know-how
  - Human resources
  - *Data*

- *Data*: An important organizational resource
- **Data** :Known facts that can be recorded and have an implicit meaning

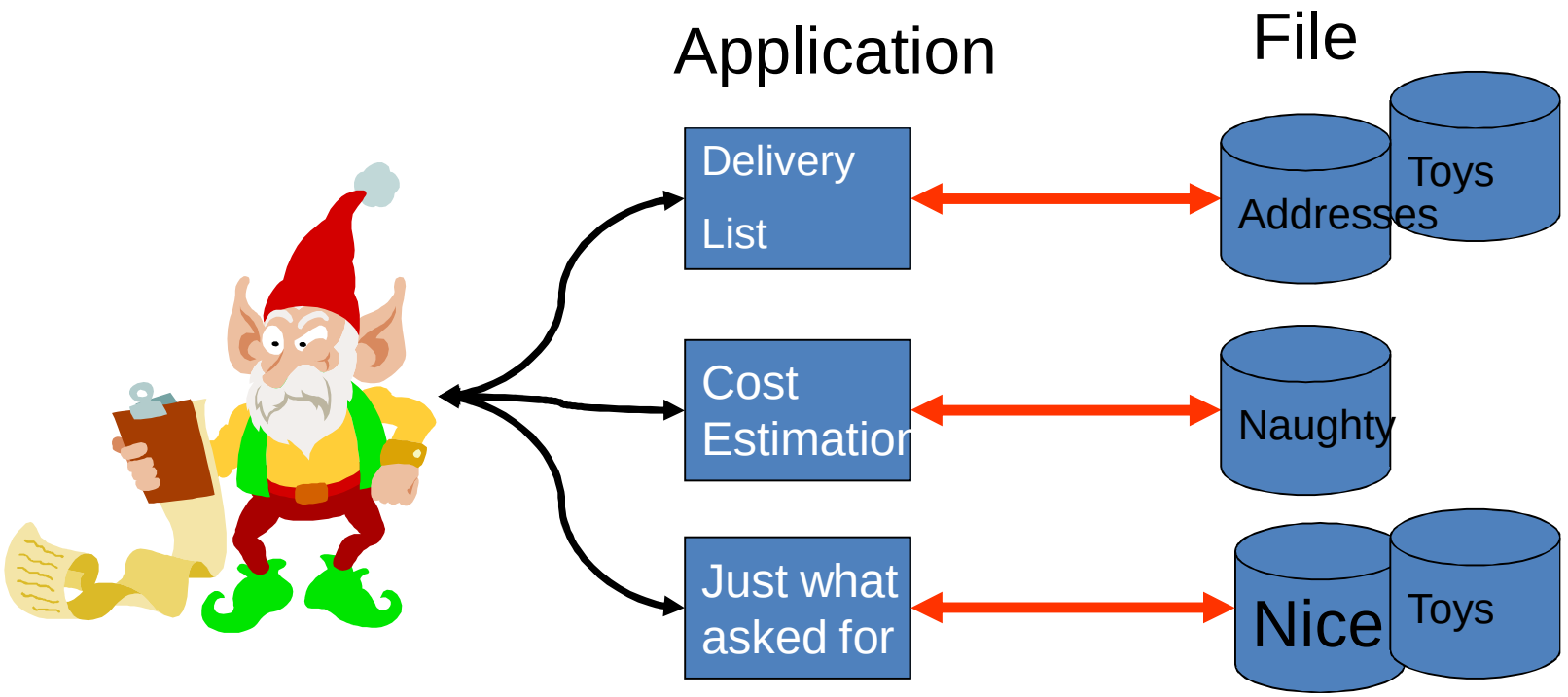
The Organizational Data Cycle



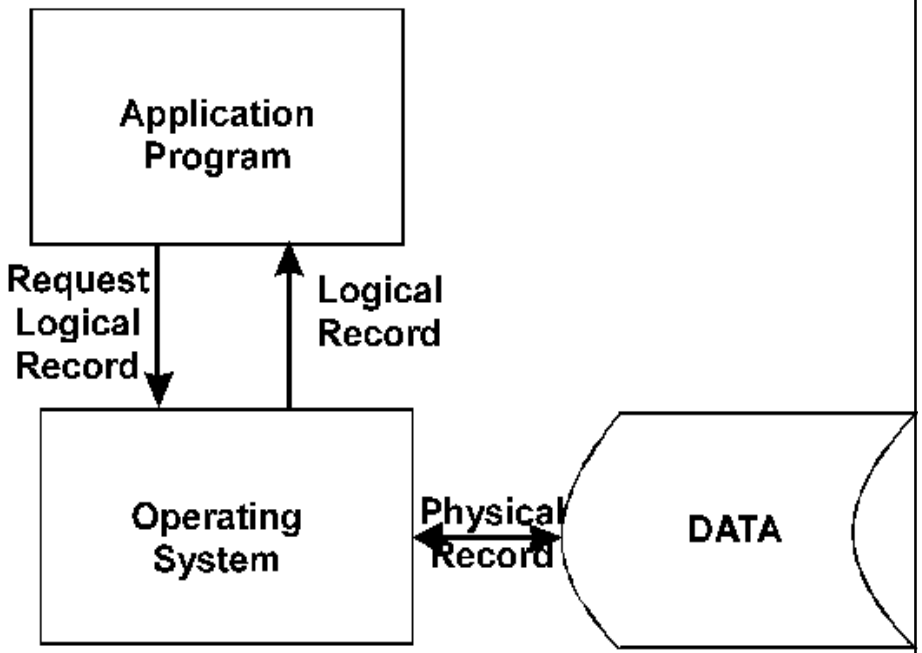
Electronic Data

- Why?
  - Large volume in a small space
  - Ease of sharing
  - Ease of use
  - Data analysis
- How?
  - File-based system versus databases

Legacy (File-Based) Systems



FILE-BASED SYSTEM

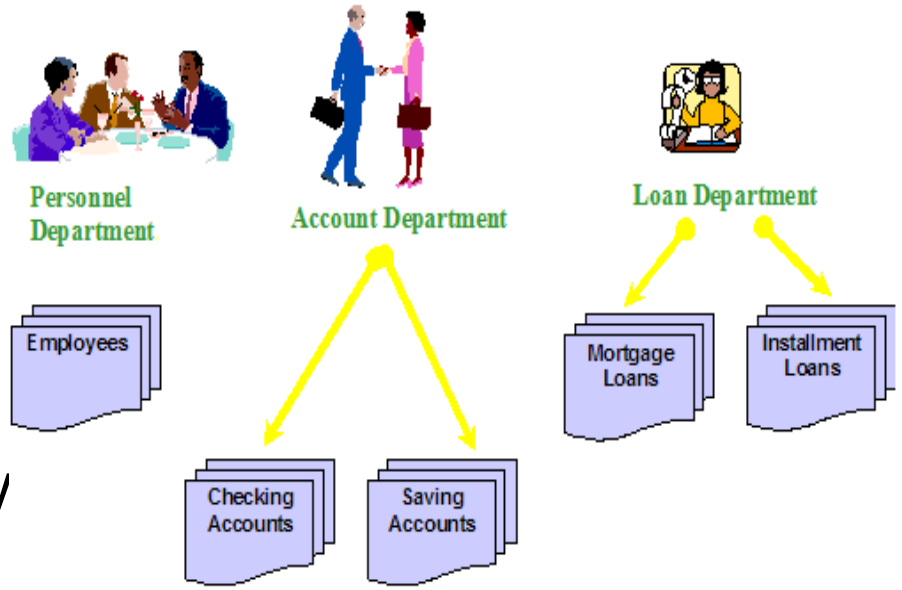


File-based systems

- Data is stored in files
- Each file has a specific format
- Programs that use these files depend on knowledge about that format

File-Based Systems..

- Problems:
  - No standards
  - Data duplication
  - Poor data sharing
  - Data inconsistency
  - Data dependence
  - No way to generate ad hoc queries
  - No provision for security, recovery, concurrency, etc.



Legacy (File-based) Systems

- Problems …..
- Uncontrolled data redundancy,
- Difficult to keep up with changes
- Record format Vs. user requirements
- Programs Vs. record format
- Low productivity
- High maintenance cost

From File Systems to DBMS

- Why shifted to Database?
  - Inconsistent Data
  - Inflexibility
  - Limited Data Sharing
  - Poor enforcement of standards
  - Excessive program maintenance

So what is a database?

- According to Oxford English Dictionary:

  "A structured collection of data held in computer storage; esp. one that incorporates software to make it accessible in a variety of ways"

- So does it make pretty much every collection of data is a "database"?

So What is a Database?..

- "One or more large structured sets of persistent data, usually associated with software to update and query the data"

  Free On-Line Dictionary of Computing

- "A collection of data arranged for ease and speed of search and retrieval"

  Dictionary.com

Files and Databases

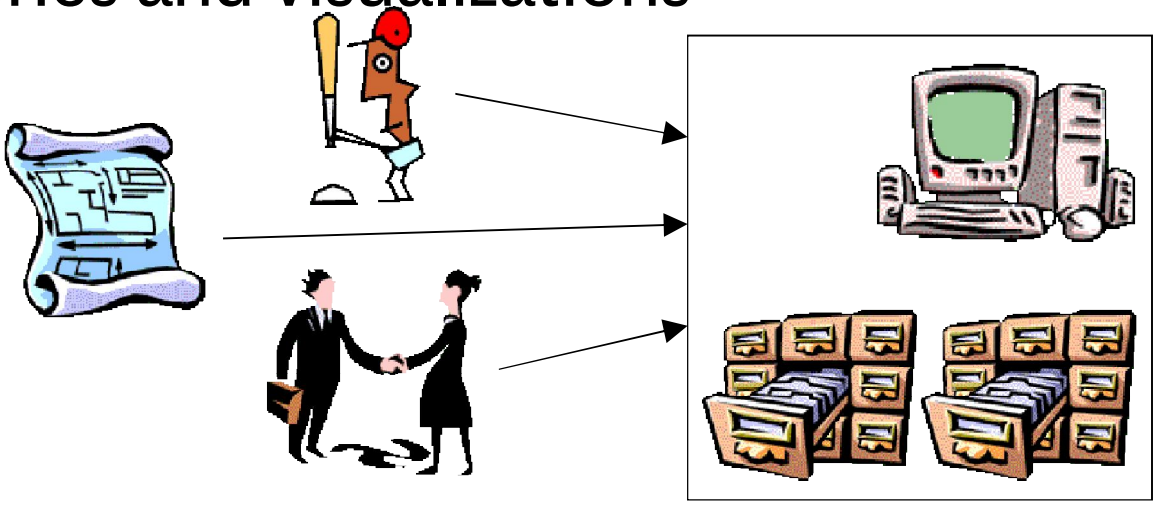
- File: A collection of records or documents dealing with one organization, person, area or subject. (Rowley)
  - Manual (paper) files
  - Computer files
- Database: “A collection of similar records with relationships between the records” (Rowley)
  - bibliographic, statistical, business data, images, etc.

Databases examples

- Web indexes
- Library catalogues
- Medical records
- Bank accounts
- Stock control
- Personnel systems
- Product catalogues
- Telephone directories
- Train timetables
- Airline bookings
- Credit card details
- Student records
- Customer histories
- Stock market prices
- Discussion boards
- and so on…

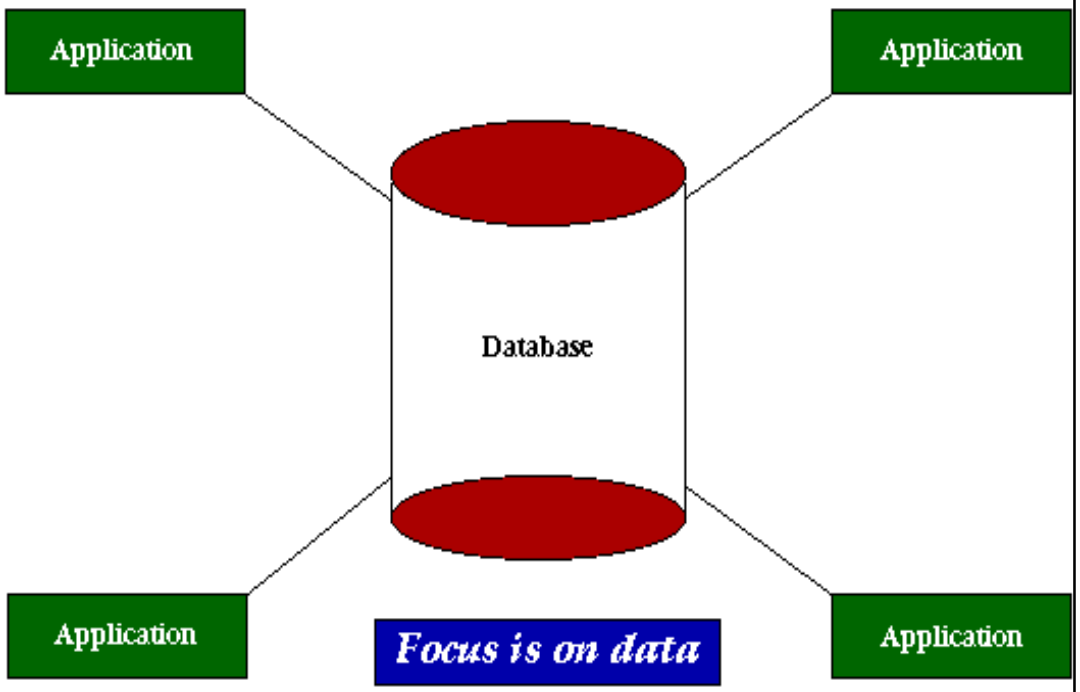
The Database

- The data itself PLUS The data definitions (metadata), applications, queries and visualizations



Database Approach

- Permits sharing
- Allows different user views



- Database Concepts
  - File processing is replaced by an integrated DBMS (Database Management System)
  - Data treated as a resource and is independent
  - Redundancy avoided for consistency (logical) and efficiency (physical)
  - Flexible access through content
  - Interaction between separate tasks on the same data

Database approach…

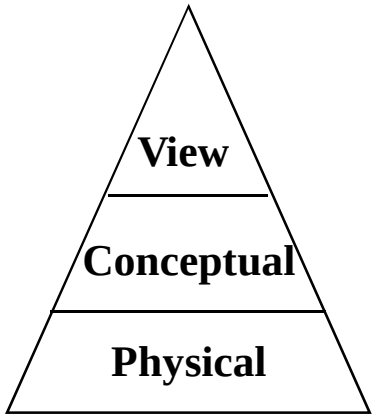
- Non-redundant collection of logically related facts
  - ***representing*** some aspect of the real world
  - the data itself ***plus*** the data definitions
- Consistent representation for each piece of data
- Avoids (minimizes) redundancy
- Users are isolated from most changes

Why Databases?

- Independence from representation formats
- Control redundancy and consistency
- Ensure integrity/security
- Better scalability
- Allow ad hoc access
- Better maintenance
- Better concurrency

Data Independence

- Does data have to be part of programs?
- Do we need to change one if the other changes?
- Three-tier architecture of databases



What the users see

How we model data

How data is stored

Data Independence….

- Physical representation and location of data and the use of that data are separated
  - The application doesn't need to know how or where the database has stored the data, but just how to ask for it.
  - Moving a database from one DBMS to another should not have a material effect on application program
  - Recoding, adding fields, etc. in the database should not affect applications

Access Flexibility

- Easy to ask *ad-hoc* questions
- No need for separate codes
- User-friendly interface
- Command-based (e.g., SQL)
- Graphical (e.g., QBE)

Data Integrity

- Ensures that the stored data are consistent and correct
- Easy to define global rules
  - customer_age > 21 years
  - number_of_credits < 18
- Can allow multiple users to access data without compromising on data integrity

Data Security

- Access definition
  - Global
  - Local
- Uniform access authorization

Data Redundancy

- Data need not be replicated
- Less wastage of storage space
- Less data anomaly
- Reduced and controlled redundancy
- Tighter control of replicated data

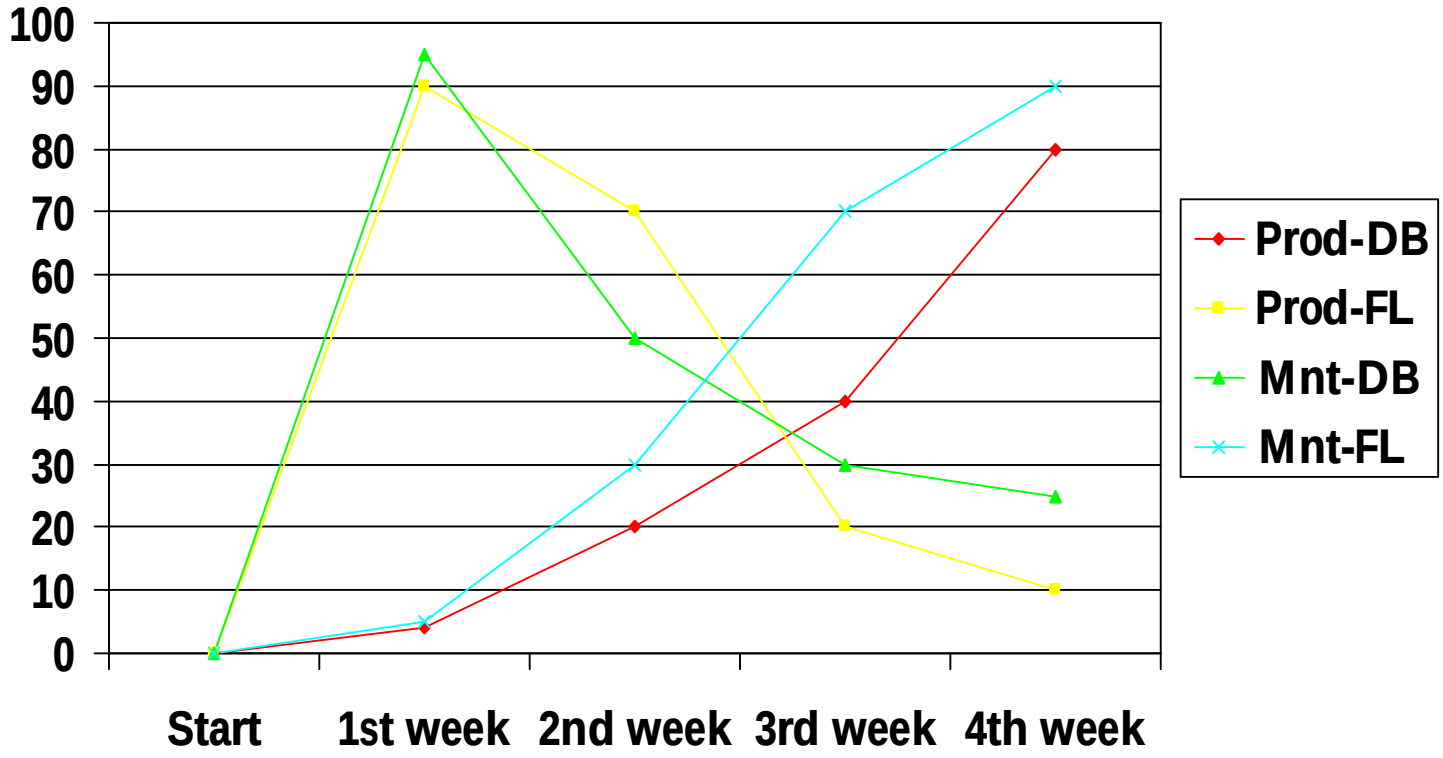
Standardization

- Everybody talks the same talk.
  - Less chance of misunderstanding
  - Easier to interpret other's data
- Easier to merge
  - Useful when several organizations combine to form one.

Productivity and Maintenance

- Increase in productivity
  - User-friendly interface
  - Independence from specific data structure
- Easier maintenance
  - Less code to maintain
- The DBMS is the bulk of the code.
- Ad-hoc queries make it possible to make do with much less code.
  - The vendor makes revisions of the DBMS.
- Economy of scale

Productivity vs. Maintenance



Disadvantages of Databases

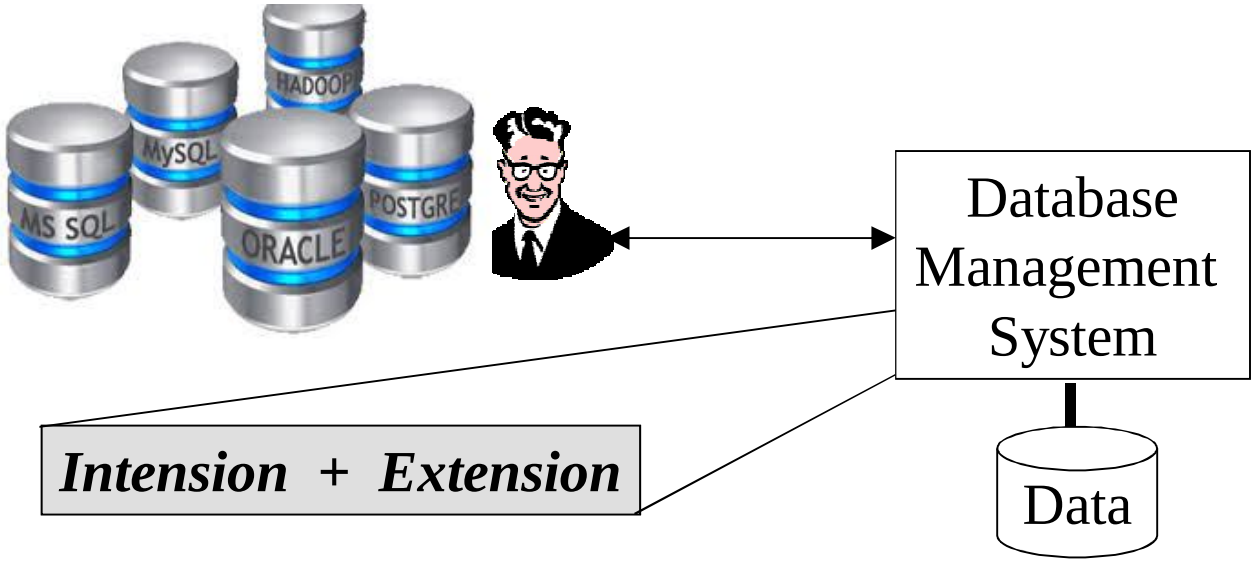
- Software complexity
- Processing inefficiency
- Need for co-ordination
- Organizational impact
- Risk

Database Management System(DBMS)

- **DBMS**
- Software system used to define, create, maintain and provide controlled access to the database and repository.
  - **Examples:**
    - Oracle
    - DB2 (IBM), MS SQL Server, MS Access, Ingres, PostgreSQL, MySQL

DBMS..

- A Specialized piece of software that sits between the data and its users.



What the DBMS does?

- Provides users with
  - Data definition language (DDL)
  - Data manipulation language (DML)
  - Data control language (DCL)
- DBMS provides
  - Persistence
  - Concurrency
  - Integrity
  - Security
  - Data independence
- Data Dictionary
  - Describes the database itself

Typical DBMS Functionality

- Define a database : in terms of data types, structures and constraints
- Construct or Load the Database on a secondary storage medium
- Manipulating the database : querying, generating reports, insertions, deletions and modifications to its content
- Concurrent Processing and Sharing: by a set of users and programs – yet, keeping all data valid and consistent

Typical DBMS Functionality..

Other features:

- Protection or Security measures to prevent unauthorized access
- Presentation and Visualization of data

DBMS Benefits

- Minimal Data Redundancy
- Consistency of Data
- Integration of Data
- Sharing of Data
- Ease of Application Development
- Uniform Security, Privacy, and Integrity Controls
- Data Accessibility and Responsiveness
- Reduced Program Maintenance
- Data Independence

Advantages of DBMS

- Data independence
- Efficient data access
- Data integrity & security
- Data administration
- Concurrent access, crash recovery
- Reduced application development time

When not to use a DBMS?

- **Main inhibitors (costs) of using a DBMS**:
  - High initial investment and possible need for additional hardware.
  - Overhead for providing generality, security, concurrency control, recovery, and integrity functions.
- **When a DBMS may be unnecessary:**
  - If the database and applications are simple, well defined, and not expected to change.
  - If there are stringent real-time requirements that may not be met because of DBMS overhead.
  - If access to data by multiple users is not required

When not to use a DBMS?..

- **When no DBMS may suffice:**
  - If the database system is not able to handle the complexity of data because of modeling limitations
  - If the database users need special operations not supported by the DBMS.

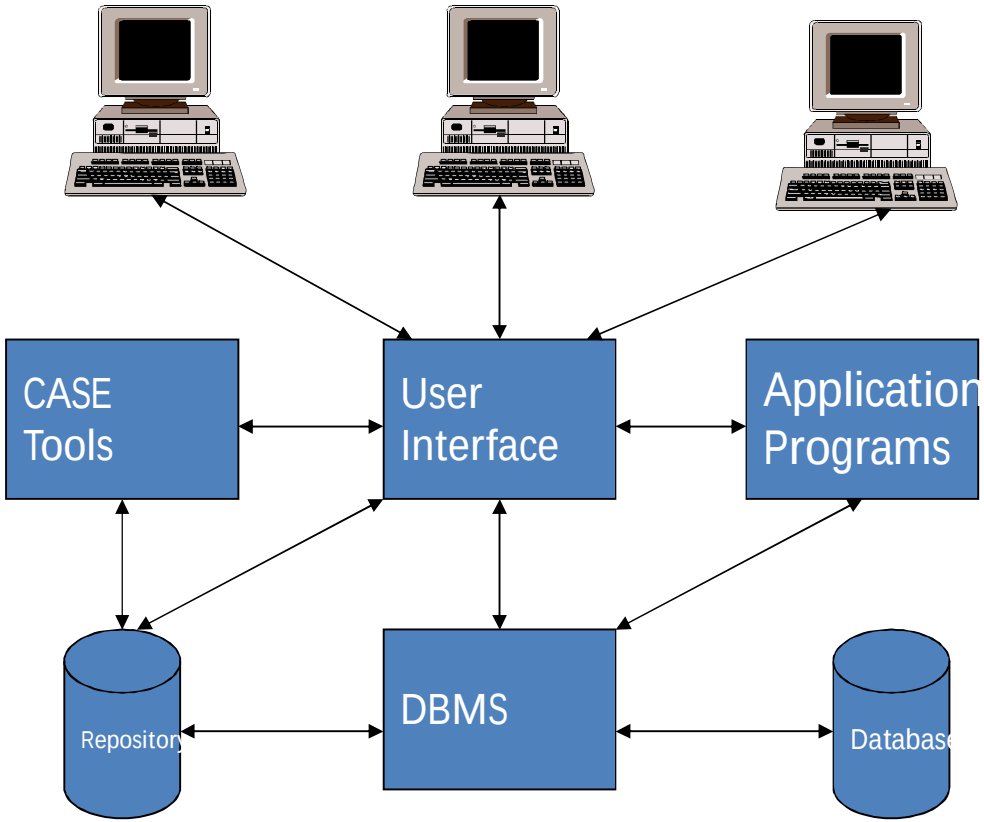
Terms and Concepts

- Repository
  - AKA Data Dictionary
  - The place where all metadata for a particular database is stored
  - may also include information on relationships between files or tables in a particular database

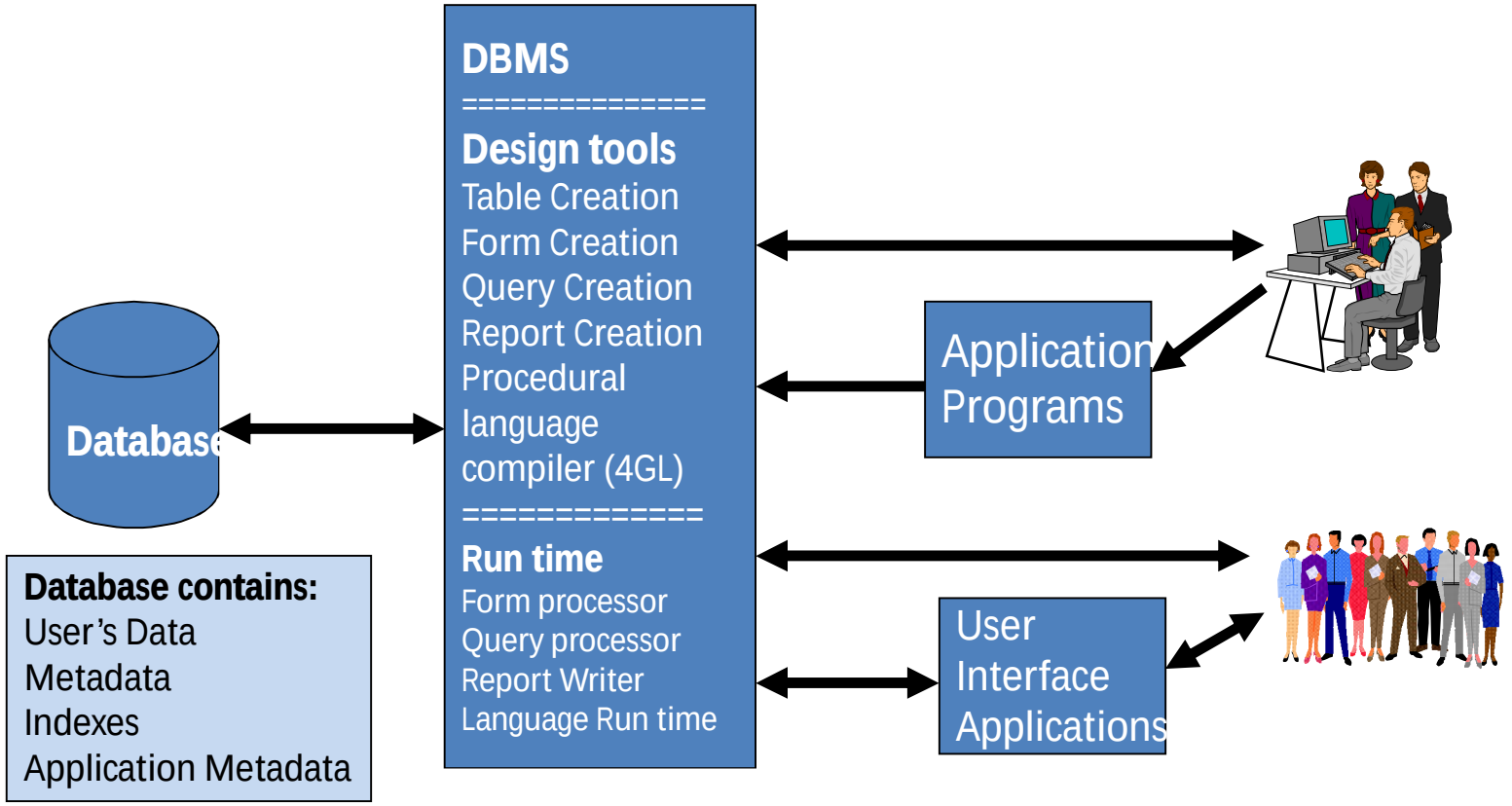
Terms and Concepts

- Metadata
  - Data about data
    - In DBMS means all of the characteristics describing the attributes of an entity, E.G.:
      - name of attribute
      - data type of attribute
      - size of the attribute
      - format or special characteristics
  - Characteristics of files or relations
    - name, content, notes, etc.

Database Environment



Database Components

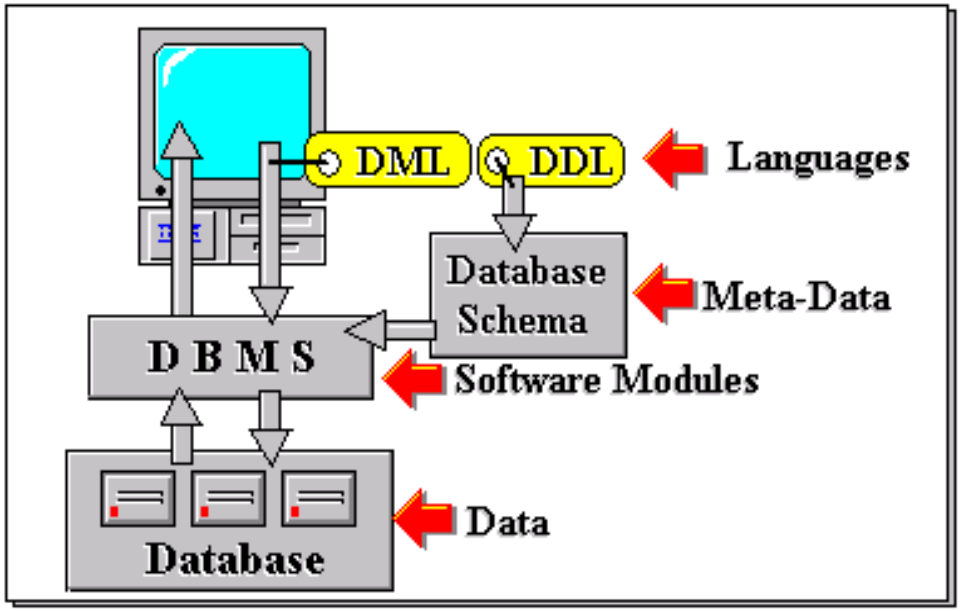


Components of the Database Environment

- **CASE Tools** – computer-aided software engineering
- **Repository** – centralized storehouse of metadata
- **Database Management System (DBMS)** – software for managing the database
- **Database** – storehouse of the data
- **Application Programs** – software using the data
- **User Interface** – text and graphical displays to users
- **Data Administrators** – personnel responsible for maintaining the database
- **System Developers** – personnel responsible for designing databases and software
- **End Users** – people who use the applications and databases

Database Systems

- **Database System**:
  - The DBMS software together with the data itself. Sometimes, the applications are also included.



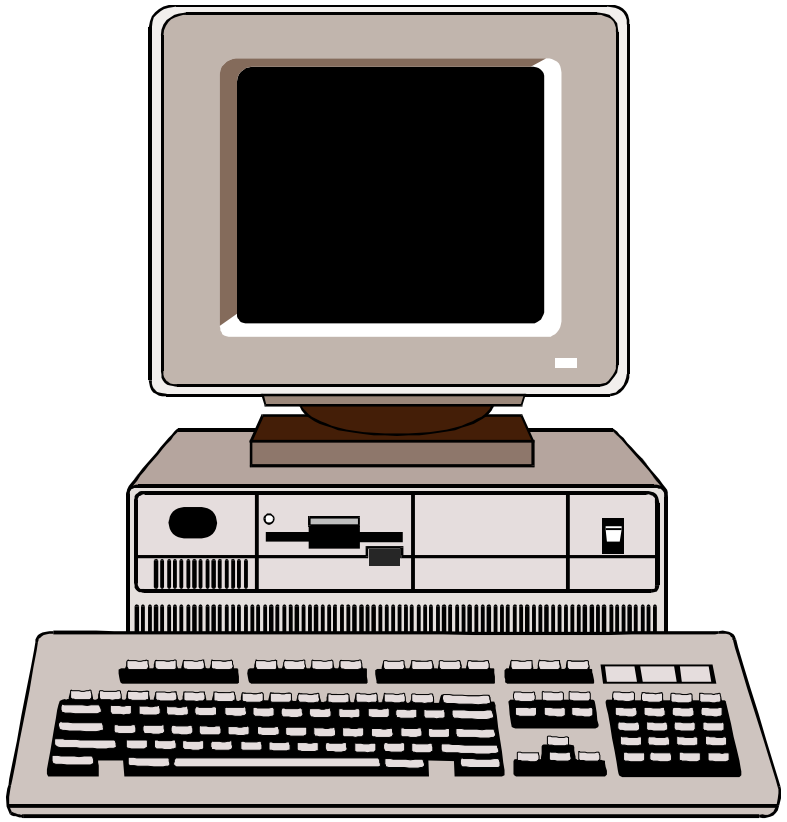
Database Systems..

- A database system consists of
  - Data (the database)
  - Software
  - Hardware
  - Users
- We focus mainly on the software

Types of Database Systems

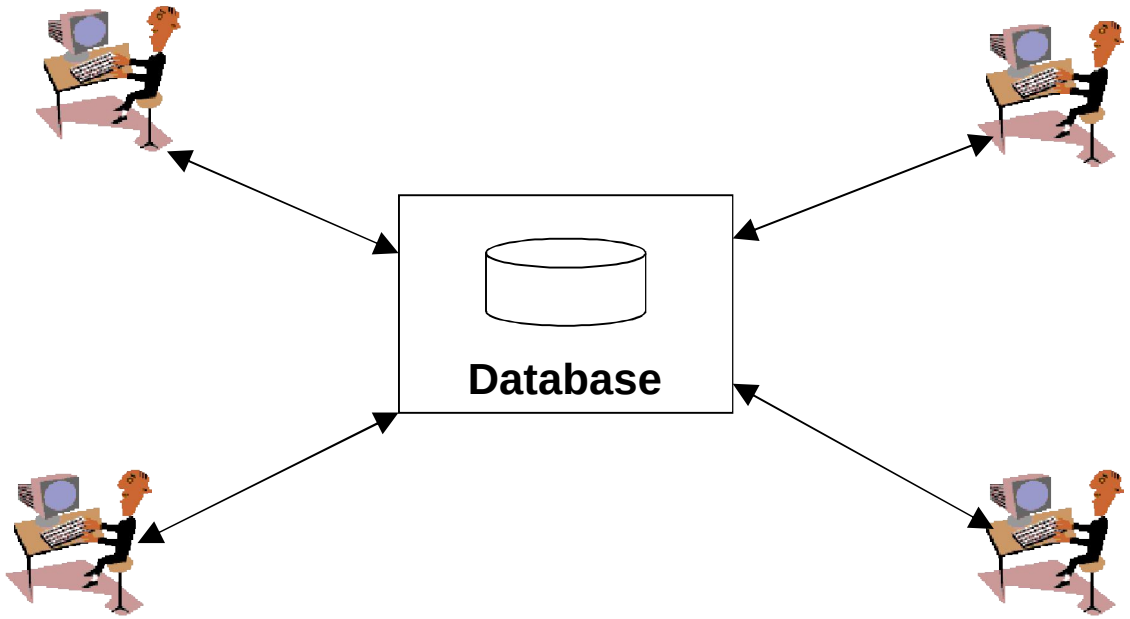
- PC Databases
- Centralized Database
- Client/Server Databases
- Distributed Databases
- Database Models

PC Databases

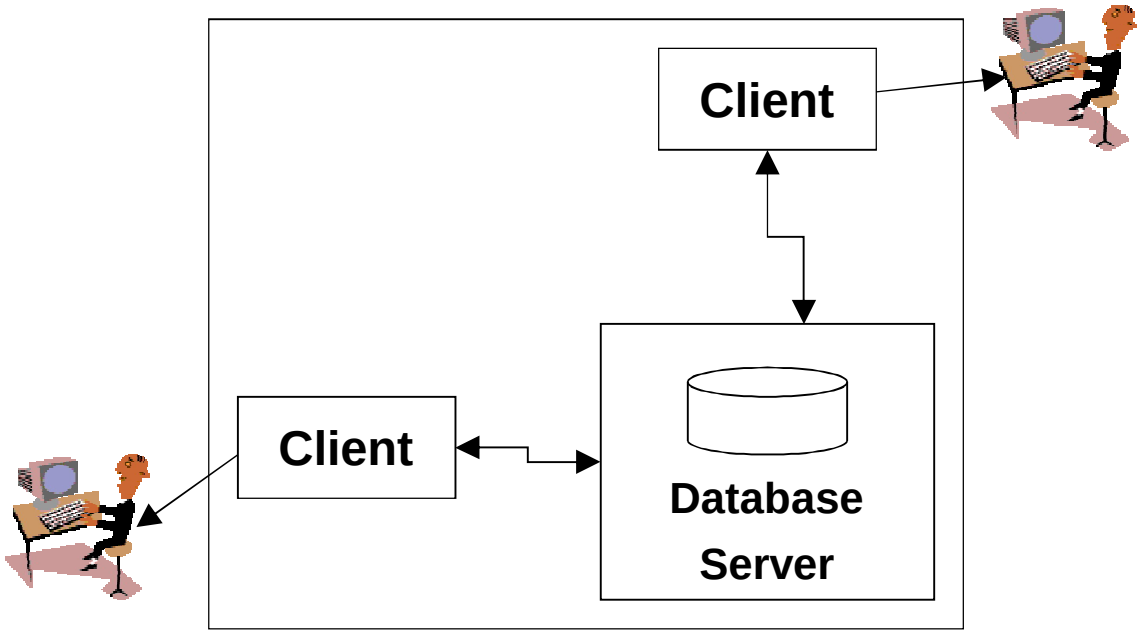


E.G.
Access
FoxPro
Dbase
Etc.

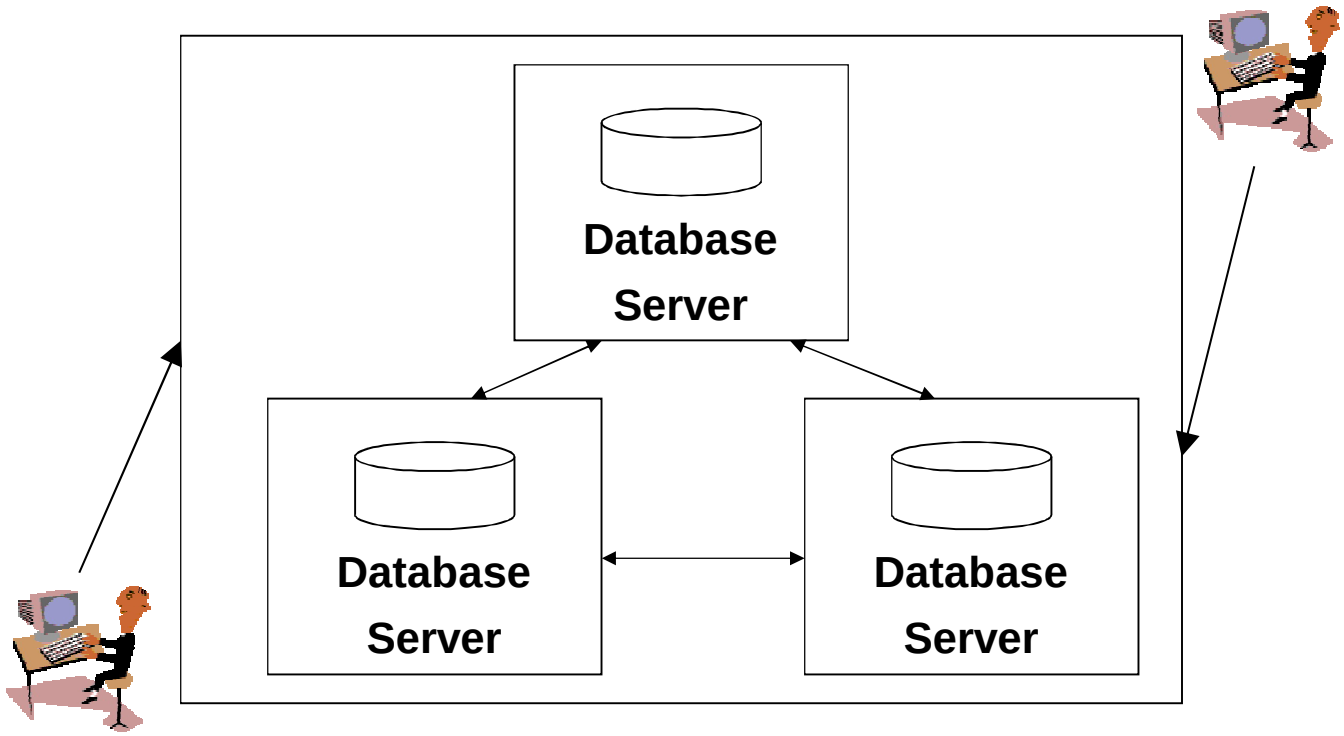
Centralized Database Systems



Client/Server Database Systems



Distributed Database Systems



Range of Database Applications

- PC databases
  - Usually for individual
- WorkGroup databases
  - Small group use where everyone has access to the database over a LAN
- Departmental databases
  - Larger than a workgroup – but similar
- Enterprises databases
  - For the entire organization over an intranet (or sometimes the internet)

Terms and Concepts

- ***Database Application***
  - An application program (or set of related programs) that is used to perform a series of database activities:
    - ***Create :*** Add new data to the database
    - ***Read:*** Read current data from the database
    - ***Update:*** Update or modify current database data
    - ***Delete:*** Remove current data from the database
    - On behalf of database users

Terms and Concepts

- ***Enterprise***
  - Organization
- ***Entity***
  - Person, Place, Thing, Event, Concept…
- ***Attributes***
  - Data elements (facts) about some entity
  - Also sometimes called fields or items or domains
- ***Data values***
  - instances of a particular attribute for a particular entity

Terms and Concepts

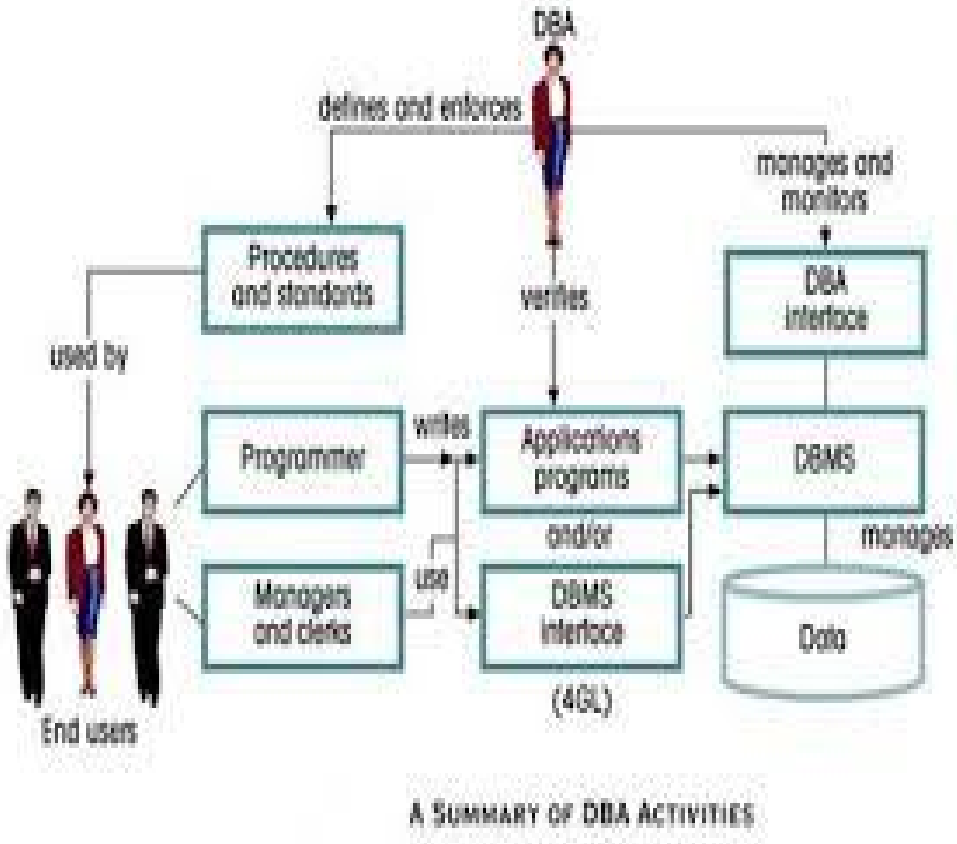
- ***Records***
  - The set of values for all attributes of a particular entity
  - AKA “tuples” or “rows” in relational DBMS
- ***File***
  - Collection of records
  - AKA “Relation” or “Table” in relational DBMS

Terms and Concepts

- ***Key***
  - an attribute or set of attributes used to identify or locate records in a file
- ***Primary Key***
  - an attribute or set of attributes that *uniquely* identifies each record in a file

Terms and Concepts

- ***DA***
- Responsibility for the overall management of data resources within an organization
- ***DBA***
- Responsibility for physical database design and technical issues in database management



A Summary of DBA Activities

Database Users

- Database administrators (DBAs)
- Database designer
- End Users

Database Users….

Actors on the scene

- **Database administrators:** responsible for authorizing access to the database, for co-ordinating and monitoring its use, acquiring software, and hardware resources, controlling its use and monitoring efficiency of operations.
- **Database Designers:** responsible to define the content, the structure, the constraints, and functions or transactions against the database. They must communicate with the end-users and understand their needs.
- **End-users:** they use the data for queries, reports and some of them actually update the database content.

Categories of End-users

- **Casual** : access database occasionally when needed
- **Naïve or Parametric** : they make up a

large section of the end-user population.

- **Sophisticated** : these include business analysts, scientists, engineers, others thoroughly familiar with the system capabilities.
- **Stand-alone** : mostly maintain personal databases using ready-to-use packaged applications.

Summary

- DBMS used to maintain, query large datasets.
  - can manipulate data and exploit *semantics*
- Other benefits include:
  - recovery from system crashes,
  - concurrent access,
  - quick application development,
  - data integrity and security.

- DBAs, DB developers the bedrock of the information economy

Summary..

- Data is essential for an organization
- A Database is usually the most effective way of storing and organizing data
- File-based Vs. database systems
- Database system properties
- Types of database systems

Match terms and definitions

1. Data
2. Database application
3. Repository
4. Metadata
5. Information
6. DBMS
7. Database
8. ERP
9. Enterprise data model

a. Data placed in context or summarized
b. Application program
c. Facts, text, graphics, etc.
d. A graphical model about an organization
e. Organized collection of related data
f. Data definitions and constraints
g. Centralized storehouse for all data definitions
h. An integrated enterprise level MIS
i. A software application to create, maintain and control a database