DATABASE SYSTEMS

WEEK 1 LECTURE 1 & 2

Course Objectives

Introduction to very basics

Guides through different design stages

Familiarize with tools

Emphasis on design stages of database

Lecture Objectives

Some common uses of database systems.

Characteristics of file-based systems.

Problems with file-based approach.

Meaning of the term database.

Meaning of the term Database Management System (DBMS).

Typical functions of a DBMS.

Major components of the DBMS environment.

Personnel involved in the DBMS environment.

Textbook

- Recommended textbooks:
  - **Database Systems: A practical approach to design, implementation and management** by Connolly and Begg
  - *Fundamentals of Database Systems (7th Edition)* by Ramez Elmasri, Shamkant B. Navathe

- Other textbooks:
  - There are *lots* of database texts
  - Most of them would be fine also

- For example:
  - 'Database Systems' by CJ Date

Data vs. Information
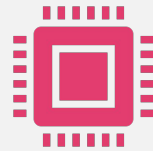
Data are simply facts or figures - bits of information, but not information itself.

When **data** are processed, interpreted, organized, structured or presented so as to make them meaningful or useful, they are called **information**. **Information** provides context for **data**.

Thing
Cricket Player
Scholars
Movies
Food
Vehicle

Data vs. Information

Data vs. Information

Thing	Data (Facts or figures)
Cricket Player	Country, name, date of birth, specialty, matches played, runs etc.
Scholars	Name, data of birth, age, country, field, books published etc.
Movies	Name, director, language (Punjabi is default in case of Pakistan) etc.
Food	Name, ingredients, taste, preferred time, origin, etc.
Vehicle	Registration number, make, owner, type, price, etc.

Traditional File Systems

- File processing systems was an early attempt to computerize the manual filing system that we are all familiar with.

- A file system is a method for storing and organizing computer files and the data they contain to make it easy to find and access them.

- File systems may use a storage device such as a hard disk or CD-ROM and involve maintaining the physical location of the files.

Traditional File Systems

- The manual filing system works well when the number of items to be stored is small.

- It even works quite adequately when there are large numbers of items and we have only to store and retrieve them.

- However, the manual filing system breaks down when we must cross-reference or process the information in the files.

Characteristics of File Processing System

It is a group of files storing data of an organization.

Each file is independent from one another.

Each file is called a flat file.

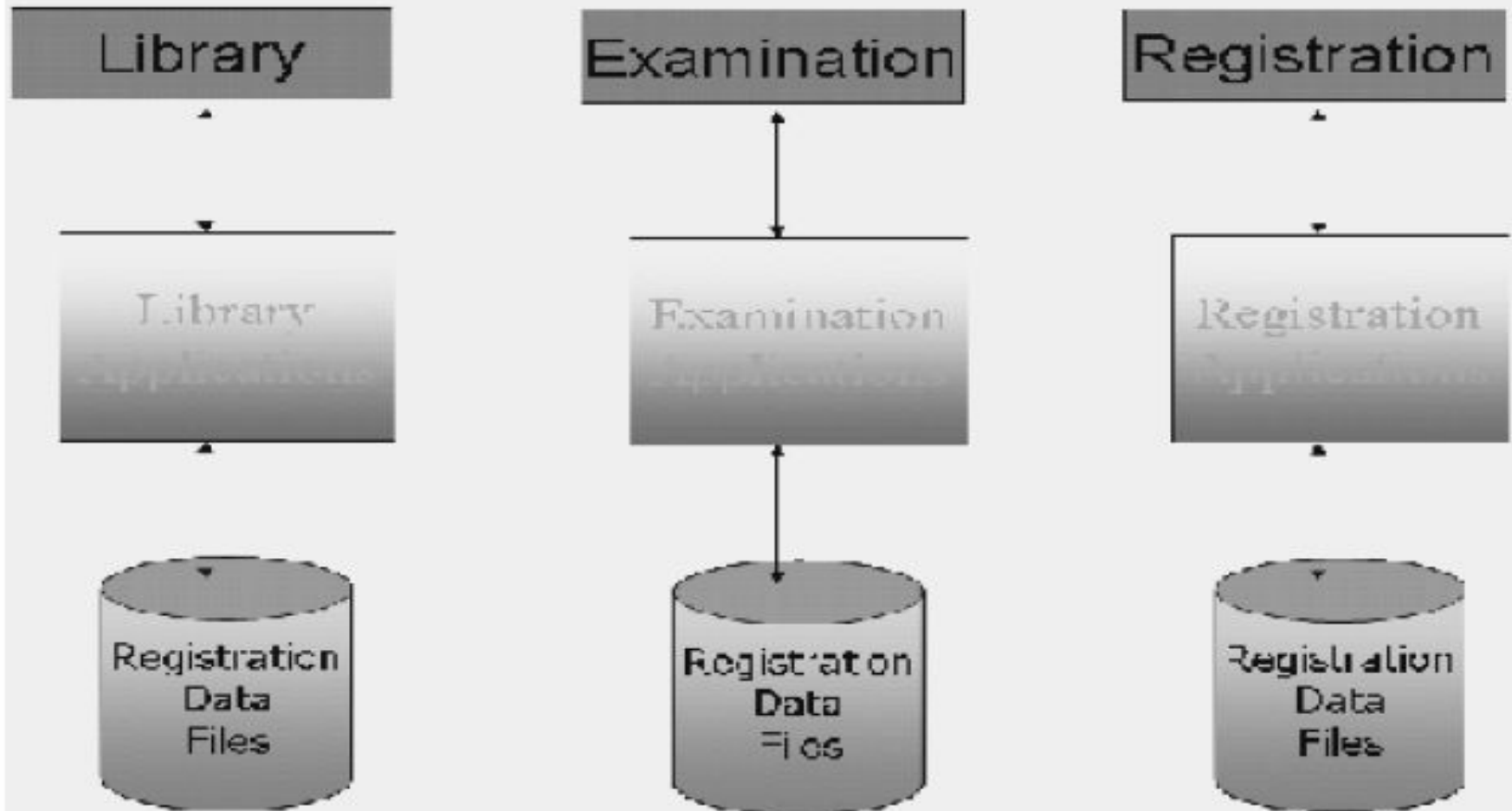Each file contained and processed information for one specific function, such as accounting or inventory.

Files are designed by using programs written in programming languages such as C, C++.

As systems became more complex, file processing systems offered little flexibility, presented many limitations, and were difficult to maintain.

File Processing Systems

| Library | Examination | Registration |

Library Applications | Examination Applications | Registration Applications

Registration Data Files | Registration Data Files | Registration Data Files

Program and Data Interdependence

File Processing System

Library
Reg_Number
Name
Father Name
Books Issued
Fine

Examination
Reg_Number
Name
Address
Class
Semester
Grade

Registration
Reg_Number
Name
Father Name
Phone
Address
Class

Duplication of Data
Vulnerable to Inconsistency

Problems With the File System

➤ Separated and Isolated Data: To decide, a user might need data from two separate files. First, the files were evaluated by analysts and programmers to determine the specific data required from each file and the relationships between the data and then applications could be written in a programming language to process and extract the needed data. Imagine the work involved if data from several files was needed.

Problems With the File System

Duplication of data: Often the same information is stored in more than one file. Uncontrolled duplication of data is not required for several reasons, such as:

Duplication is wasteful. It costs time and money to enter the data more than once

It takes up additional storage space, again with associated costs.

File Based Systems (Summary)

- 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

- Problems:
  - No standards
  - Data duplication
  - Data dependence
  - No way to generate ad hoc queries
  - No provision for security, recovery, concurrency, etc.
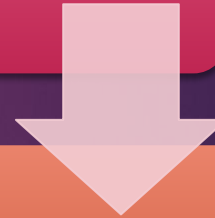
Why Study Databases?

➤ Databases are useful

  ➤ Many computing applications deal with large amounts of information

  ➤ Database systems give a set of tools for storing, searching and managing this information

What is a Database?

"A set of information held in a computer"
- Oxford English Dictionary

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

Databases – Introduction

- ► **Database Def-1**

- ► A database is a _shared_ collection of _logically related_ _data_ that is stored to meet the requirements of _different users_ of an organization

- ► **Database Def-2**

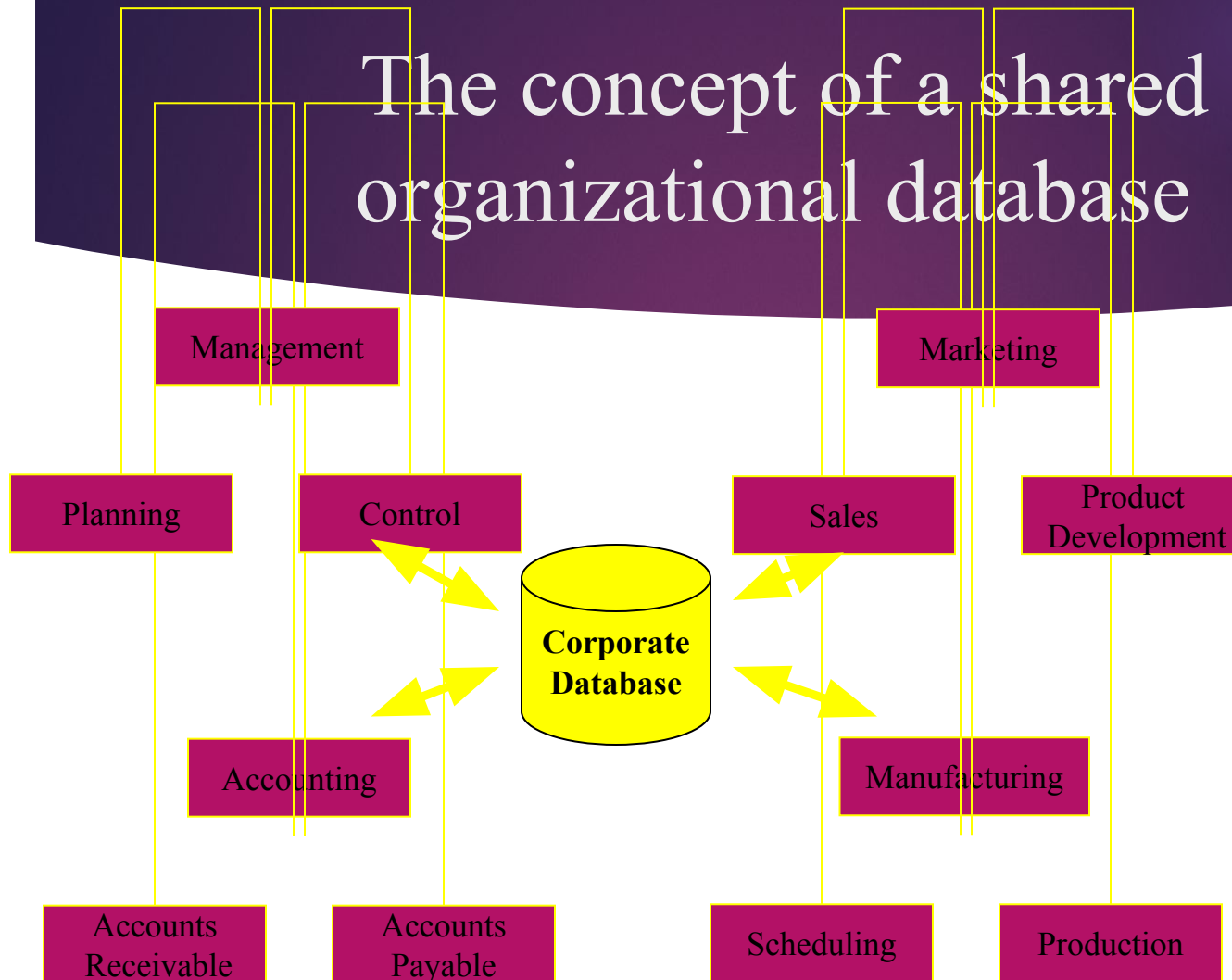- ► A database is a _self-describing_ collection of integrated records

- ► **Database Def-3**

- ► A database models a _particular real world system_ in the computer in the form of data

Databases - Introduction

- ► Database approach

  - ► **Database Management System (DBMS)**: a general-purpose software system that facilitates the processes of defining, constructing, manipulating, and sharing databases among various users and applications.

  - ► Database: a collection of related data managed by a DBMS

  - ► Data: known facts that can be recorded and that have implicit meaning

  - ► Database system = the database + DBMS software

  - ► DBMS provides facilities for querying, data security, and integrity and concurrent control

  - ► Database application - set of programs that use DBMS to perform a particular business function

The concept of a shared organizational database

Management

Marketing

Planning

Control

Sales

Product Development

Corporate Database

Accounting

Manufacturing

Accounts Receivable

Accounts Payable

Scheduling

Production

Databases

- 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
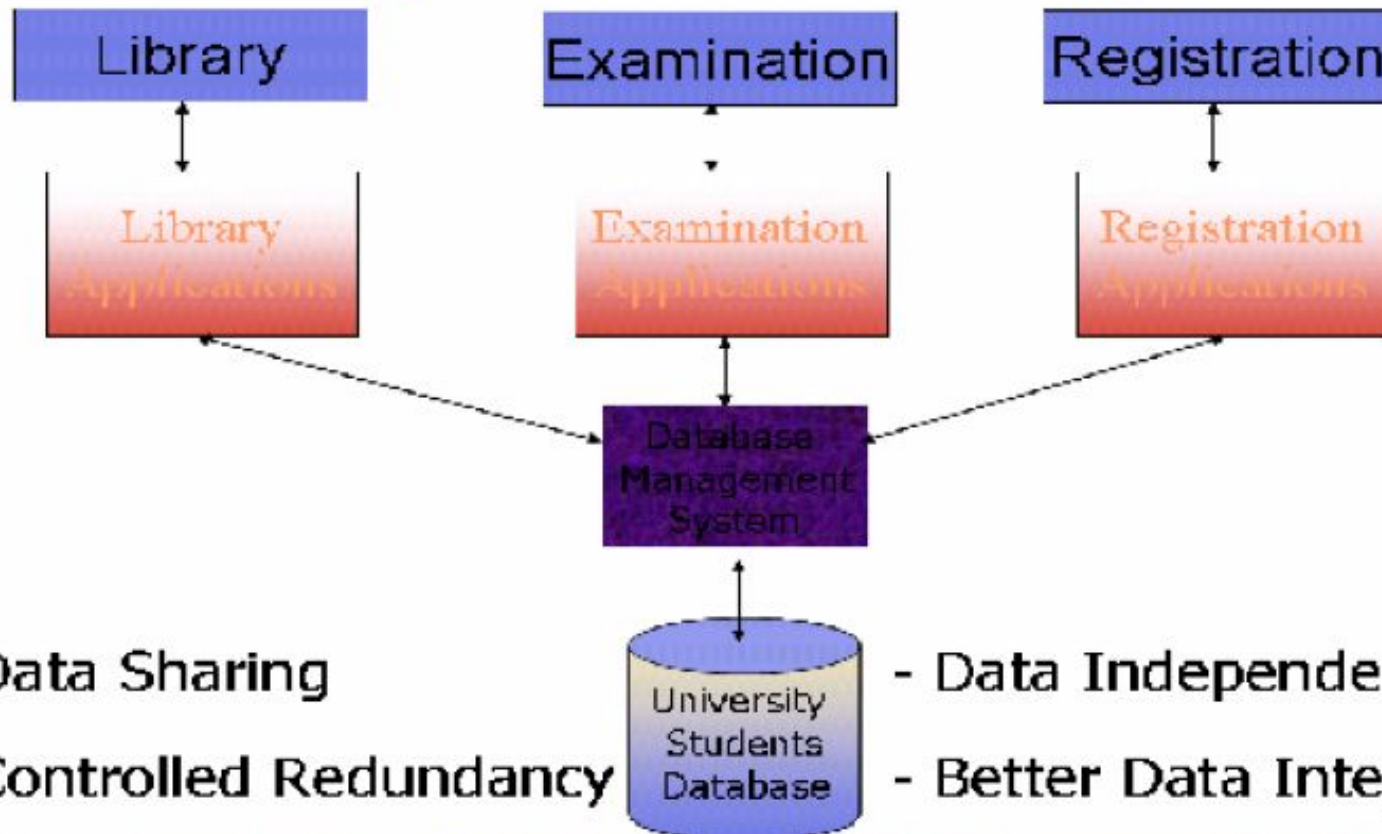- Stock market prices
- Discussion boards
- and so on…

Database Systems

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

- Database systems allow users to
  - Store
  - Update
  - Retrieve
  - Organise
  - Protect

their data.

Advantages of Databases

It will be helpful to reiterate our database definition here, that is, database is a shared collection of logically related data, designed to meet the information needs of multiple users in an organization. A typical database system environment is shown in the figure 3 below:

Advantages of Database Approach

| Library | Examination | Registration |

Library Applications | Examination Applications | Registration Applications

Database Management System

University Students Database

- Data Sharing
- Controlled Redundancy
- Data Independence
- Better Data Integrity

Database Users

- End users
  - Use the database system to achieve some goal
- Application developers
  - Write software to allow end users to interface with the database system
- Database Administrator (DBA)
  - Designs & manages the database system
- Database systems programmer
  - Writes the database software itself

Database Management Systems

- ► A database is a collection of information

- ► A database management system (DBMS) is the software which controls that information

- ► Examples:
  - ► Oracle
  - ► DB2 (IBM)
  - ► MS SQL Server
  - ► MS Access
  - ► Ingres
  - ► PostgreSQL
  - ► MySQL

What the DBMS does

- ► Provides users with
  - ► Data definition language (DDL)
  - ► Data manipulation language (DML)
  - ► Data control language (DCL)
- ► Often these are all the same language

- ► **DBMS provides**
  - ► **Persistence**
  - ► **Concurrency**
  - ► **Integrity**
  - ► **Security**
  - ► **Data independence**
- ► **Data Dictionary**
  - ► **Describes the database itself**

Data Dictionary - Metadata

- The dictionary or catalogue stores information about the database itself

- This is data about data or 'metadata'

- Almost every aspect of the DBMS uses the dictionary

- The dictionary holds
  - Descriptions of database objects (tables, users, rules, views, indexes,…)
  - Information about who is using which data (locks)
  - Schemas and mappings

Relational Systems

- Information is stored as *tuples* or *records* in *relations* or *tables*

- There is a sound mathematical theory of relations

- Most modern DBMS are based on the relational model

- The relational model covers 3 areas:
  - Data structure
  - Data integrity
  - Data manipulation

- More details in the next lecture…

Next Lecture

- ► Database Management System in detail

- ► Importance of database design.

- ► Database system basic structures.