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).

Lecture Objectives

- Typical functions of a DBMS.
- Major components of the DBMS environment.
- Personnel involved in the DBMS environment.
- History of the development of DBMSs.
- Advantages and disadvantages of DBMSs.

Examples of Database Applications

- Purchases from the supermarket
- Purchases using your credit card
- Booking a holiday at the travel agents
- Using the local library
- Taking out insurance
- Renting a video
- Using the Internet
- Studying at university

Definitions of Database

- ***Def 1:*** Database is an *organized collection* of *logically related data*
- ***Def 2:*** A database is a *shared* collection of *logically related data* that is stored to meet the requirements of *different users* of an organization
- ***Def 3:*** A database is a *self-describing* collection of integrated records
- ***Def 4:*** A database models a *particular real world system* in the computer in the form of data

Definitions

- ***Data***: stored representations of meaningful objects and events or
- Referred to facts concerning objects and events that could be recorded and stored on computer media
  - Structured: numbers, text, dates
  - Unstructured: images, video, documents
- ***Information***: data processed to increase knowledge in the person using the data
- ***Metadata:*** data that describes the properties and context of user data

What is a Database

- Shared collection of logically related data (and a description of this data), designed to meet the information needs of an organization.
- System catalog (metadata) provides description of data to enable program–data independence.
- Logically related data comprises entities, attributes, and relationships of an organization's information.

Figure 1-1a Data in Context

Class Roster

Course: MGT 500 Business Policy

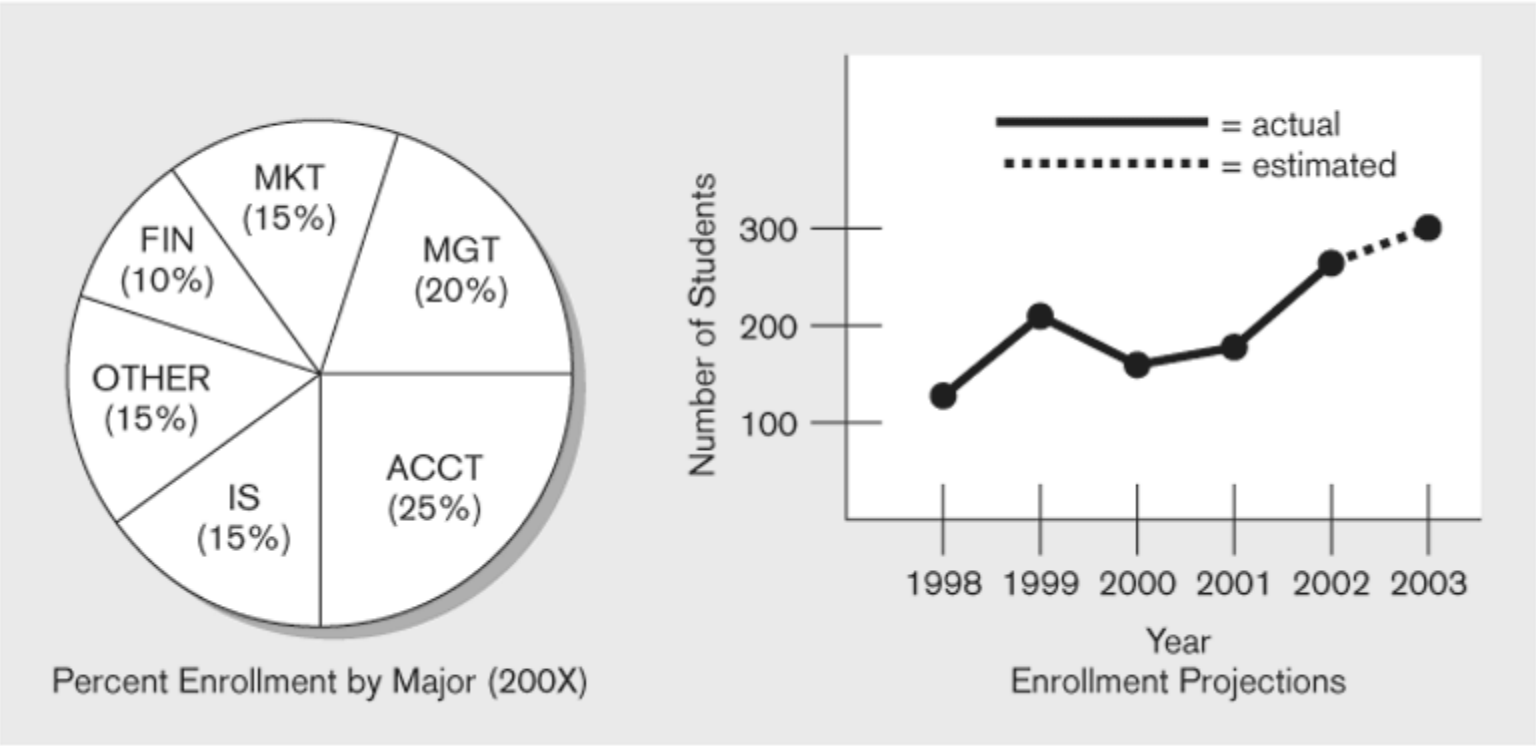
Semester: Spring 200X

Section: 2

Name	ID	Major	GPA
Baker, Kenneth D.	324917628	MGT	2.9
Doyle, Joan E.	476193248	MKT	3.4
Finkle, Clive R.	548429344	PRM	2.8
Lewis, John C.	551742186	MGT	3.7
McFerran, Debra R.	409723145	IS	2.9
Sisneros, Michael	392416582	ACCT	3.3

Context helps users understand data

Figure 1-1b Converting data to information - Summarized data



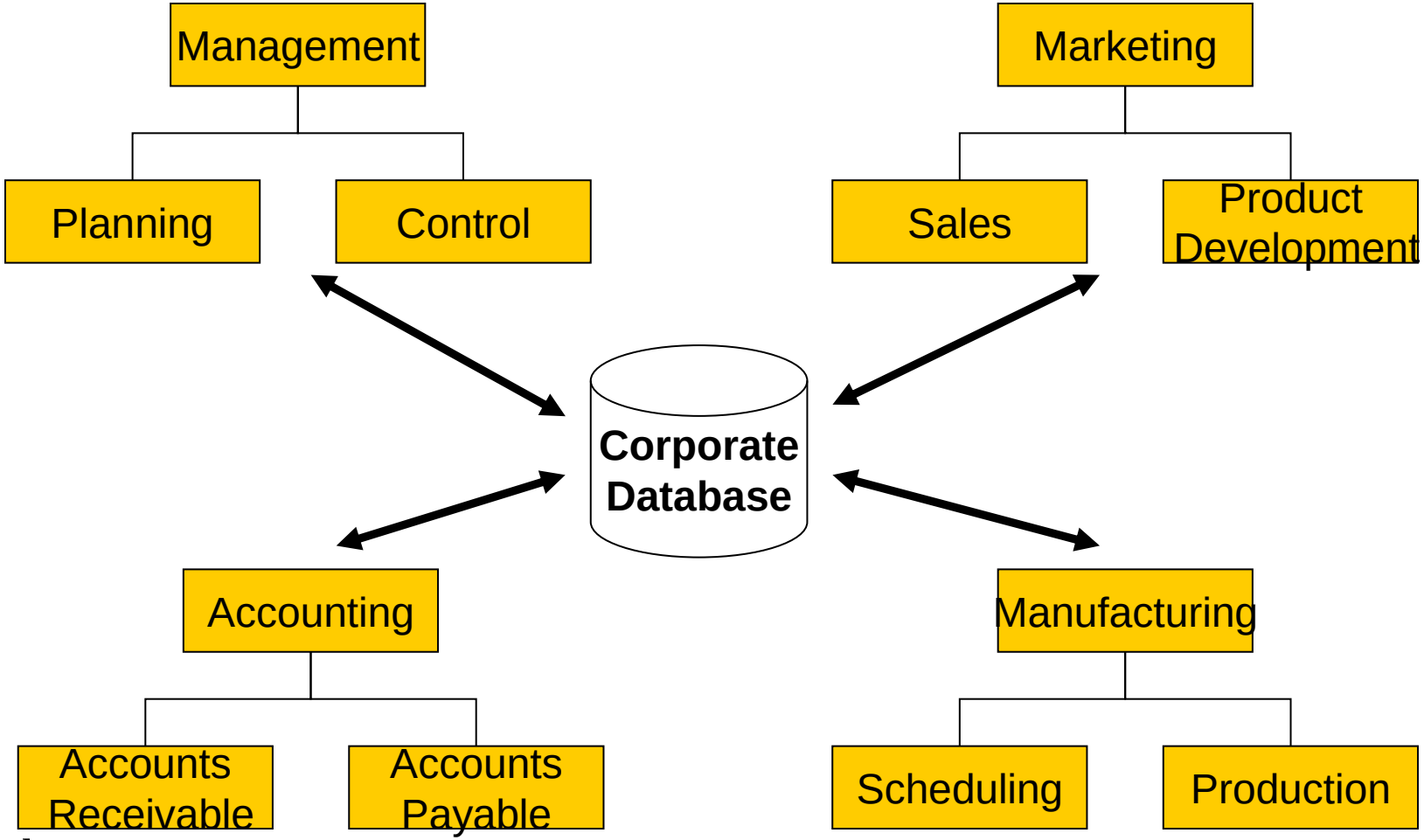
Graphical displays turn data into useful information that managers can use for decision making and interpretation

Table 1-1 Example Metadata for Class Roster

Data Item		*Value*				
Name	**Type**	**Length**	**Min**	**Max**	**Description**	**Source**
Course	Alphanumeric	30			Course ID and name	Academic Unit
Section	Integer	1	1	9	Section number	Registrar
Semester	Alphanumeric	10			Semester and year	Registrar
Name	Alphanumeric	30			Student name	Student IS
ID	Integer	9			Student ID (SSN)	Student IS
Major	Alphanumeric	4			Student major	Student IS
GPA	Decimal	3	0.0	4.0	Student grade point average	Academic Unit

Descriptions of the properties or characteristics of the data, including data types, field sizes, allowable values, and data context

The concept of a shared organizational database



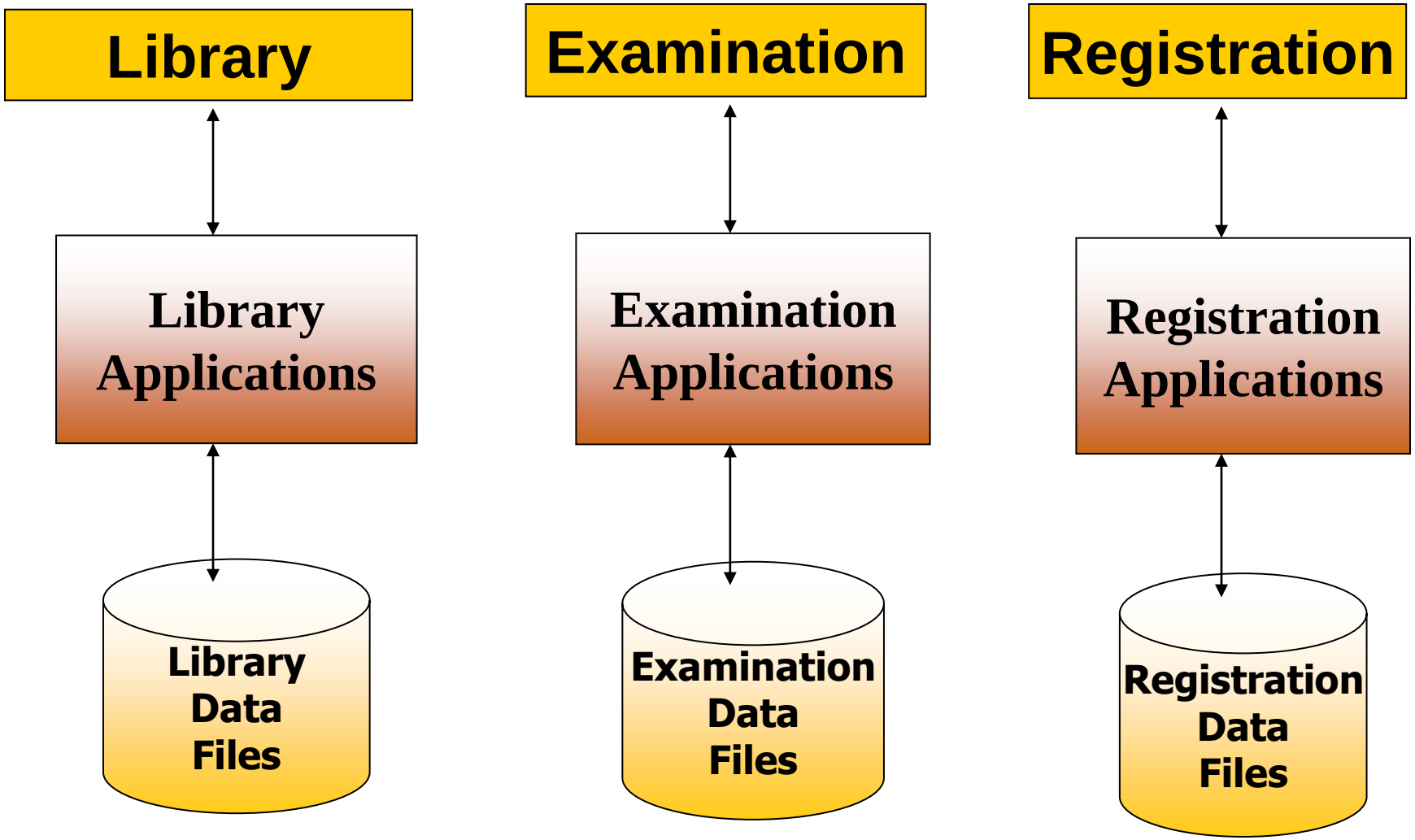
A bit of History

- Computer initially used for computational/ engineering purposes
- Commercial applications introduced File Processing System

File Processing System

- A collection of application programs that perform services for the end-users such as production of reports
- Each program defines and manages its own data

File Processing Systems



File Processing Systems

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

Files Based Processing

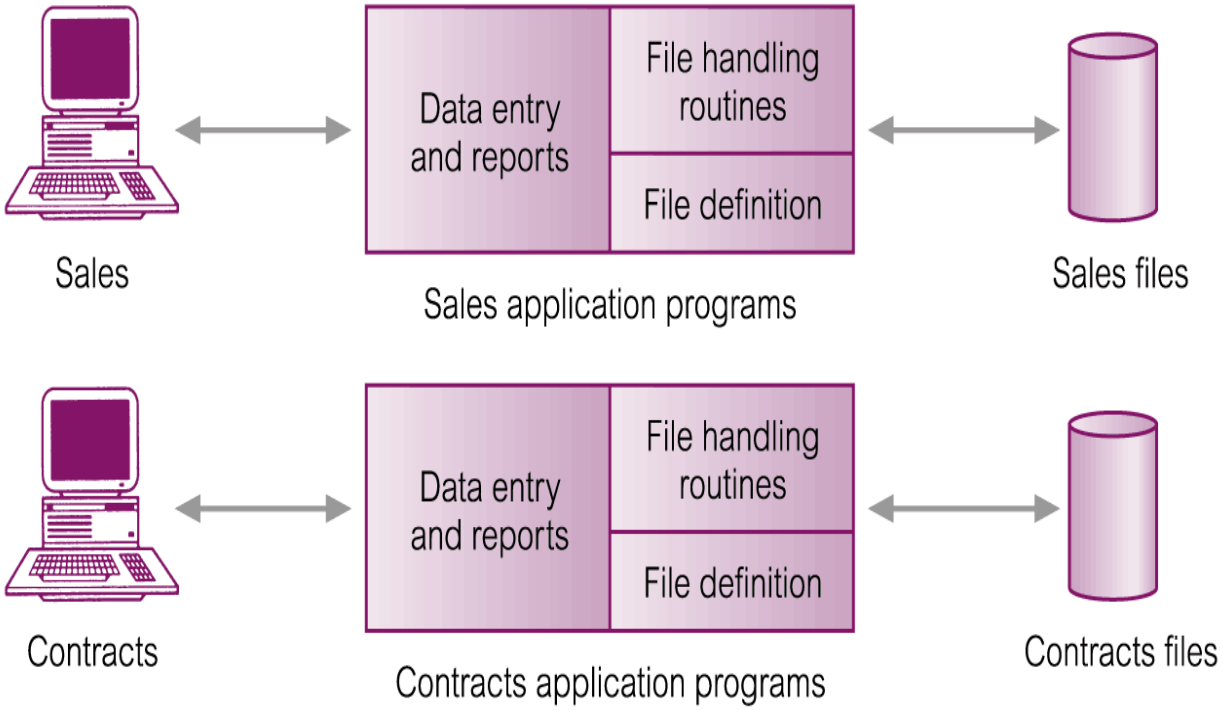


Figure 1.5
File-based processing.

Disadvantages of File Processing

- **Program-Data Dependence**
  - File structure is defined in the program code.
  - All programs maintain metadata for each file they use
- **Duplication of Data (Data Redundancy)**
  - Different systems/programs have separate copies of the same data
  - Same data is held by different programs.
  - Wasted space and potentially different values and/or different formats for the same item.
- **Limited Data Sharing**
  - No centralized control of data
  - Programs are written in different languages, and so cannot easily access each other's files.

Disadvantages of File Processing

- **Lengthy Development Times**
  - Programmers must design their own file formats
- **Excessive Program Maintenance**
  - 80% of of information systems budget
- **Vulnerable to Inconsistency**
  - Change in one table need changes in corresponding tables as well otherwise data will be inconsistent

Problems with Data Dependency

- Each application programmer must maintain their own data
- Each application program needs to include code for the metadata of each file
- Each application program must have its own processing routines for reading, inserting, updating and deleting data
- Lack of coordination and central control
- Non-standard file formats

Problems with Data Redundancy

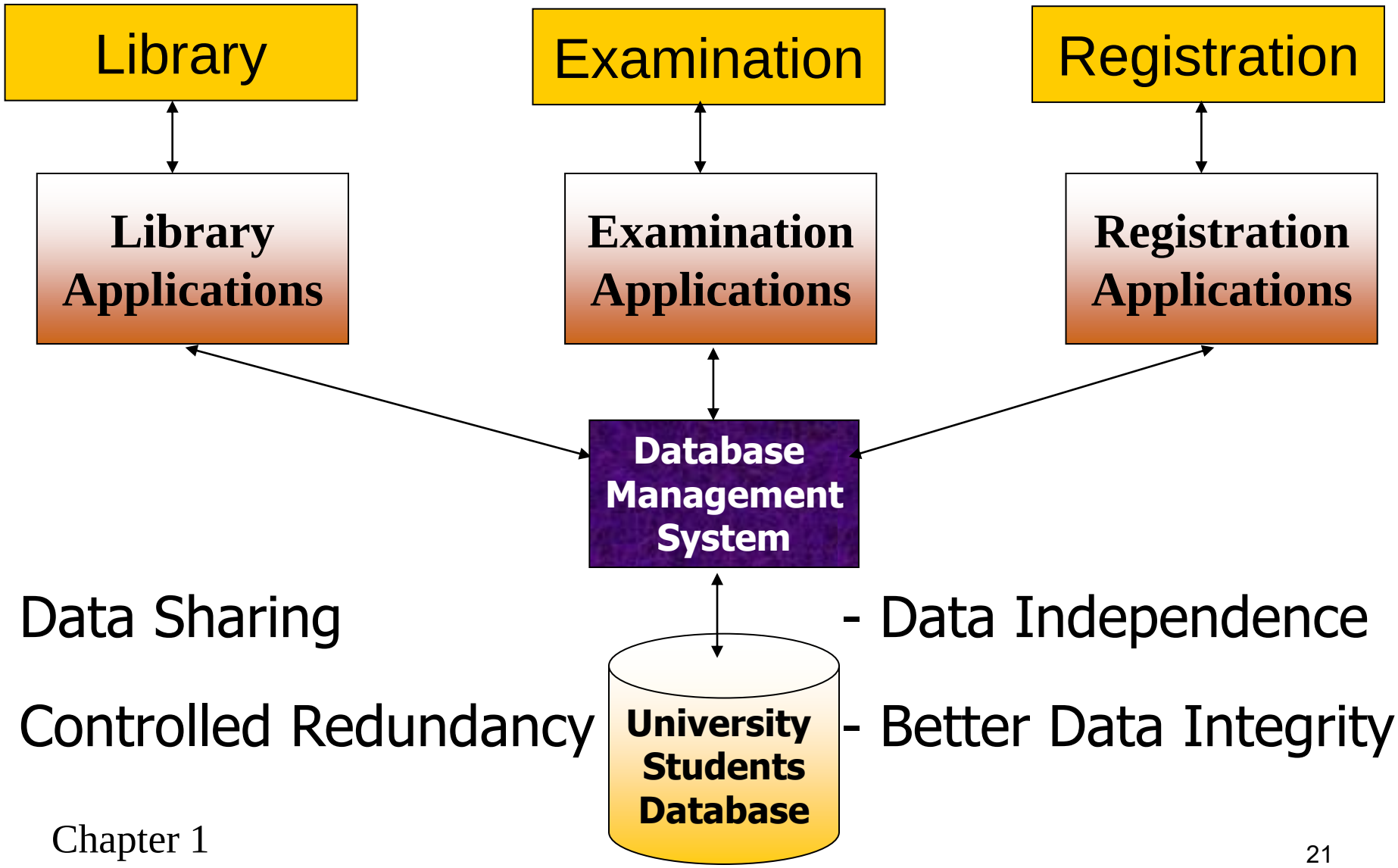
- Waste of space to have duplicate data
- Causes more maintenance headaches
- The biggest problem:
  - **When data changes in one file, could cause inconsistencies (Vulnerable to Inconsistency)**
  - Compromises ***data integrity (data reliability)***

SOLUTION:
The DATABASE Approach

- Central repository of shared data
- Data is managed by a controlling agent
- Stored in a standardized, convenient form

This requires a
Database and Database Management System (DBMS)

Advantages of Database Approach



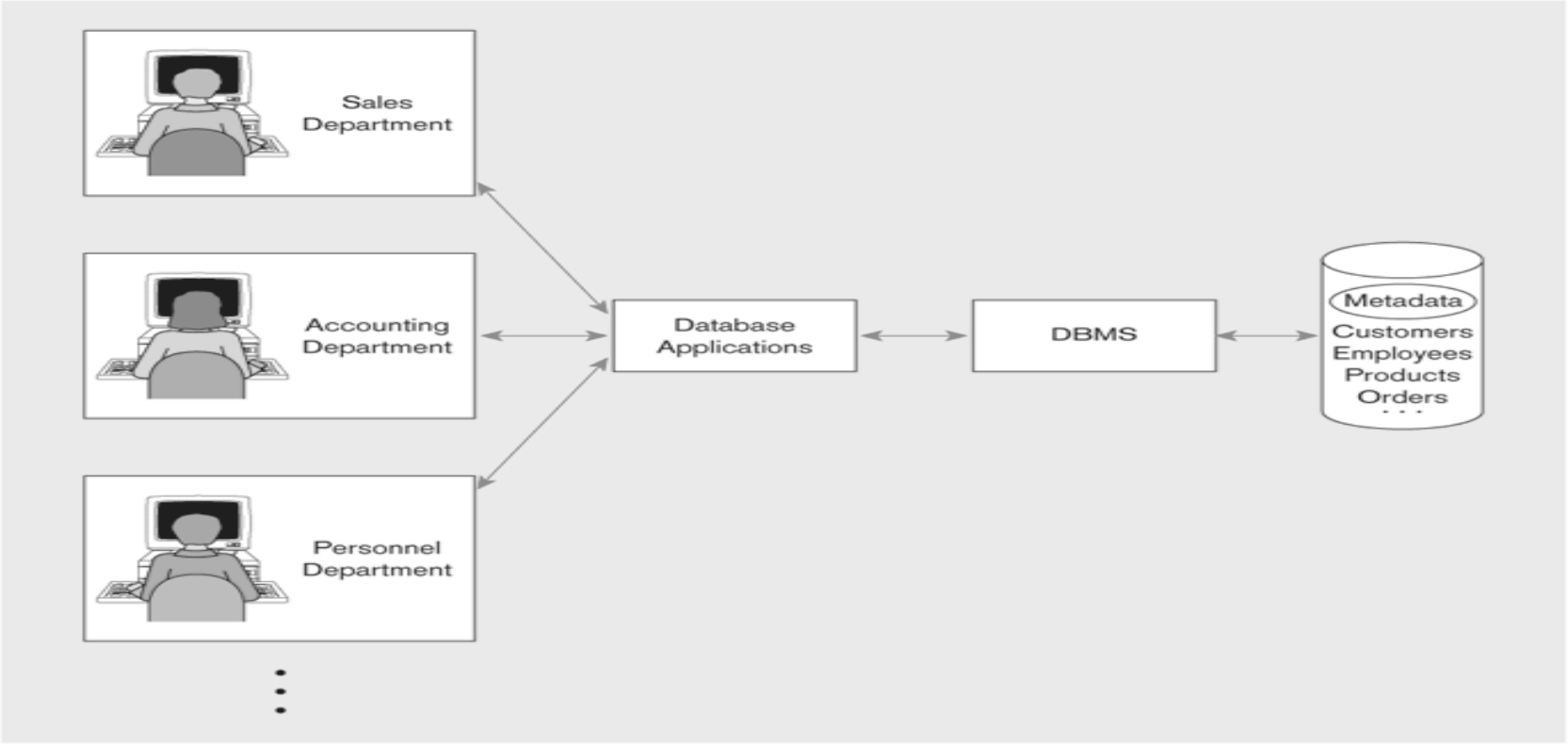
- Data Sharing
- Data Independence
- Controlled Redundancy
- Better Data Integrity

Database Management System

- A software system that is used to create, maintain, and provide controlled access to users of a database
- (Database) application program: A computer program that interacts with database by issuing an appropriate request (SQL statement) to the DBMS

Database Management System

Figure 1-3 Database approach at Pine Valley Furniture Company



DBMS manages data resources like an operating system manages hardware resources

Table 1-5

Advantages of the Database Approach

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

Table 1-6

Costs and Risks of the Database Approach

New, specialized personnel
Installation and management cost and complexity
Conversion costs
Need for explicit backup and recovery
Organizational conflict