ACS-1803
Introduction to Information Systems

Instructor: Kerry Augustine

Data Management

Lecture Outline 2, Part 1

Overview

- The importance of Data Management
- The file approach vs. the Database approach
- How Data is Categorized
- Database Design
- Relational Databases
- Data Modeling
- Schemas
- Database Management System
- Business Intelligence, Data Warehouses, Data Marts, and Data Mining

Database Technology

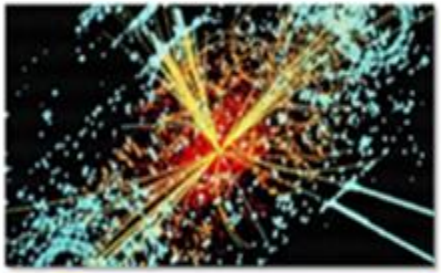
- A collection of related data organized in a way that makes it valuable and useful
- Allows organizations to retrieve, store, and analyze information easily
- Is vital to an organization's success in running operations and making decisions

Definitions

- Database:
  - Organized collection of data
  - For the most part stored in electronic form
  - Data organized to model relevant aspects of reality
- Database management system (DBMS):
  - Group of programs that manipulate the database
  - Provide an interface between the database and its users and other application programs
- Database administrator (DBA):
  - Skilled IS professional who directs all activities related to an organization's database

Databases in Action

- Reservation systems
  - book flights from multiple airlines, hotel rooms etc.
  - e.g. Amadeus systems
    - *Global Distribution System (GDS)* founded by Lufthansa, Air France and other partners



- Scientific databases
  - sensor data, classifications (e.g. human genome) as well as data from simulations
  - e.g. LHC Computing Grid
    - LHC experiments at CERN
    - 15 petabytes of data per year

- Banking and trading
  - customer data, account information, transactions, ...
  - e.g. London Stock Exchange
    - almost 1 million trades per day

- Geographic Information Systems (GIS)
  - store raster (bitmap) or vector data representing real world objects
  - geospatial query language

- Embedded databases in cars, airplanes etc.
  - manage configurations and store sensor data
  - e.g. db4o object database used in BMW's Car IT system

- Many everyday devices contain databases
  - TVs, washing machines, mobile phones, ...
  - e.g. Android phones with SQLite database

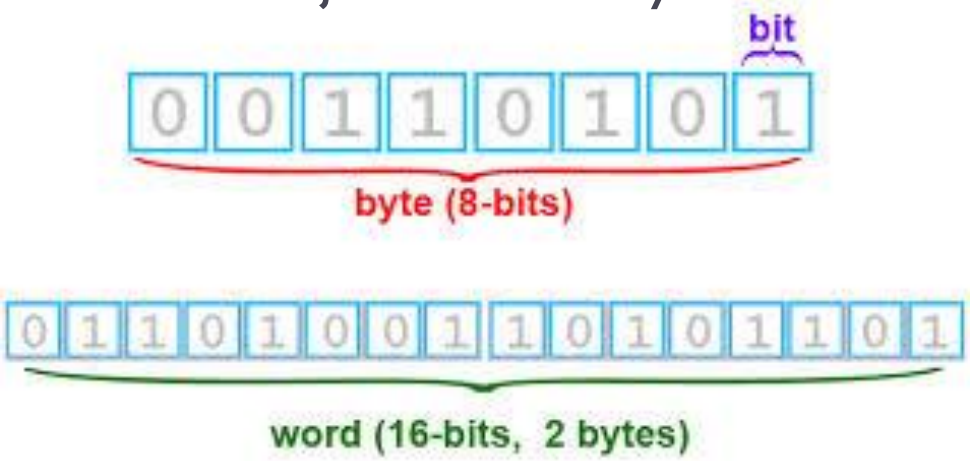
Data Management Concepts

- Why manage data?
  - Without data and the ability to process the data: An organization could not successfully complete most business activities
- Data consists of raw facts
- To transform data into useful information:
  - It must first be organized in a meaningful way i.e. **Database**

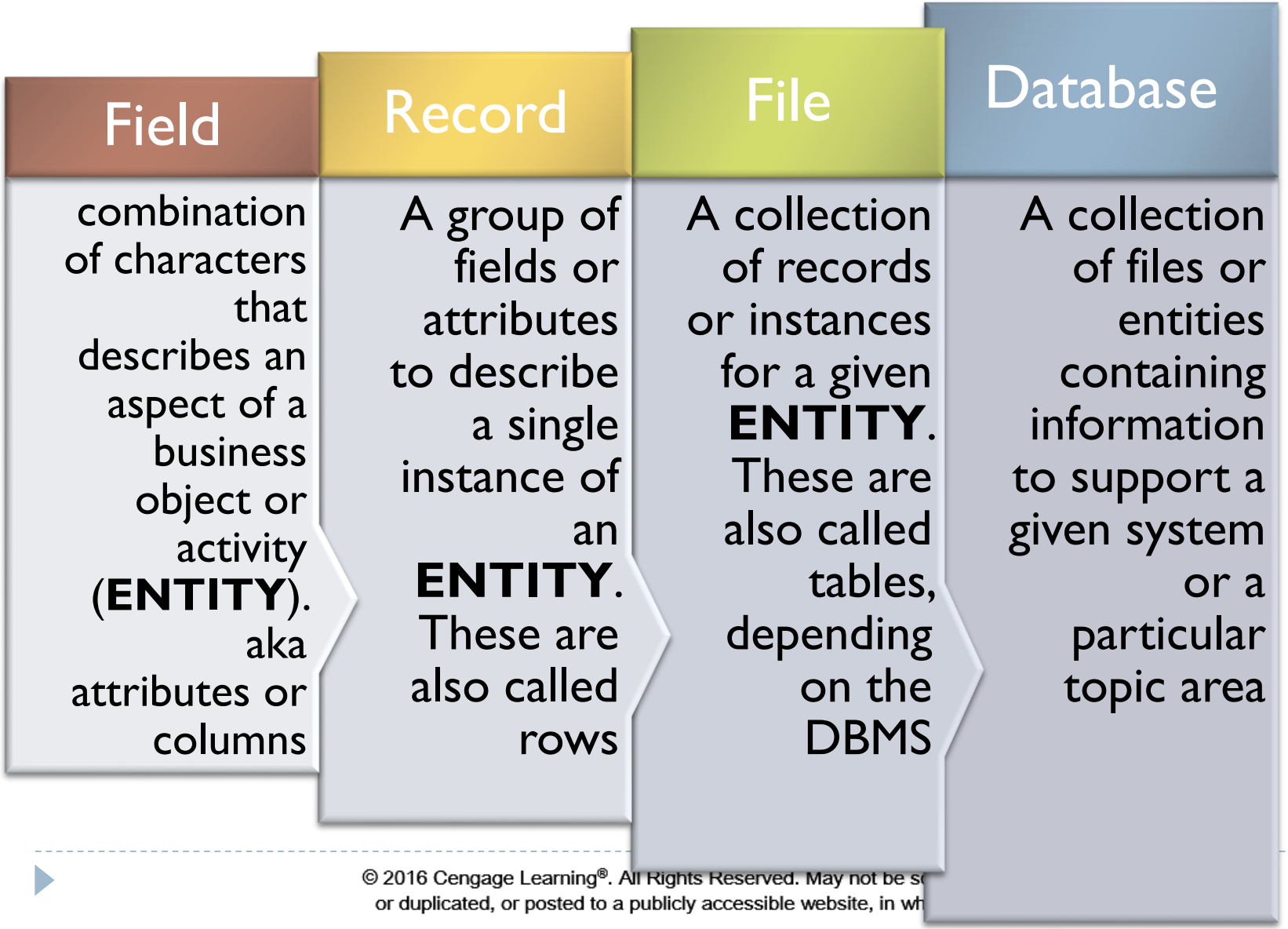
- Database Management System (DBMS)
  - A collection of programs that enables users to store, modify, and extract information from a database

The Hierarchy of Data

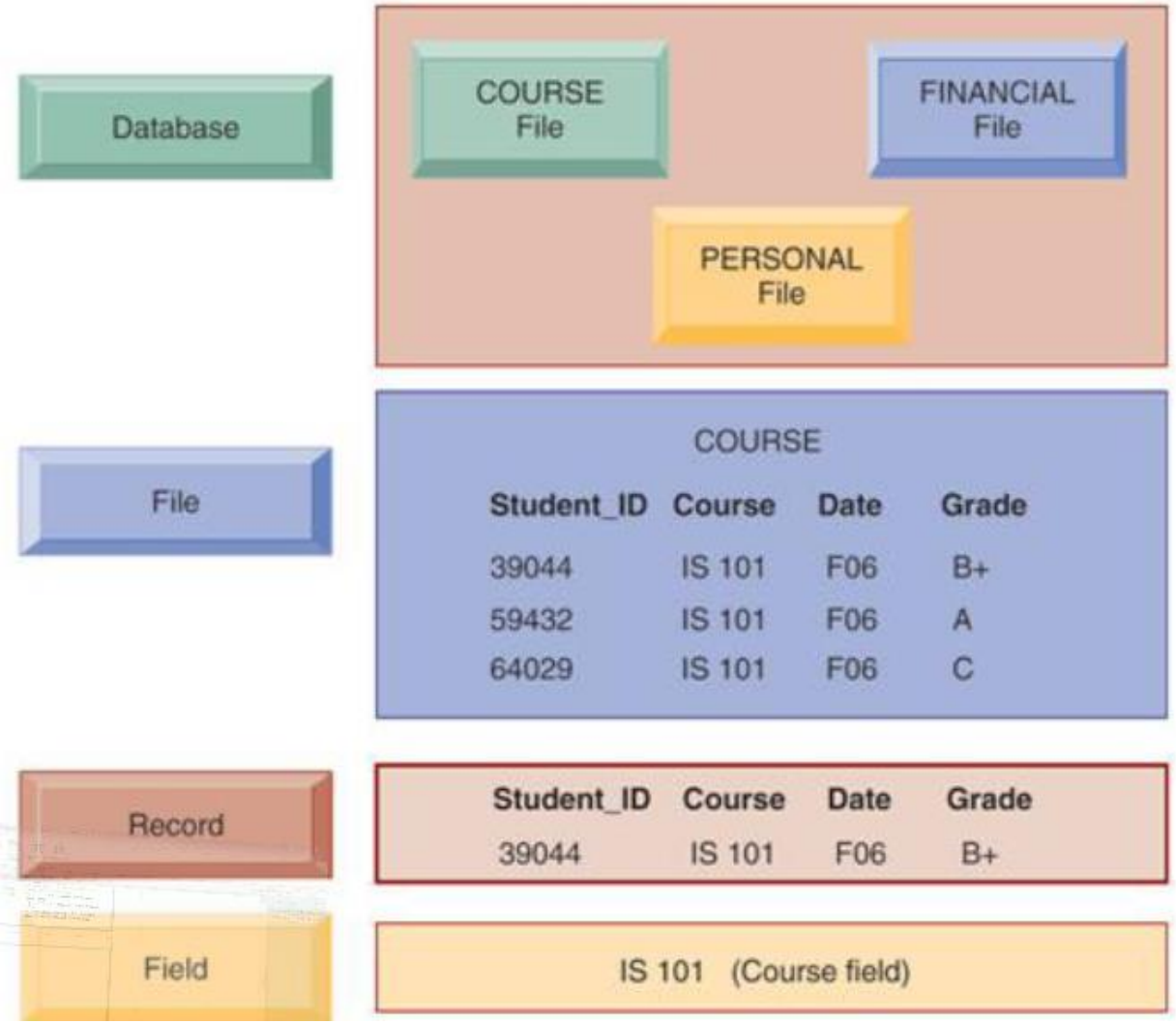
- Bit (a binary digit):
  - Circuit that is either on or off
  - Byte: Typically made up of eight bits
- Character:
  - Basic building block of information
- Field:
  - Name, number, or combination of characters that describes an aspect of a business object or activity



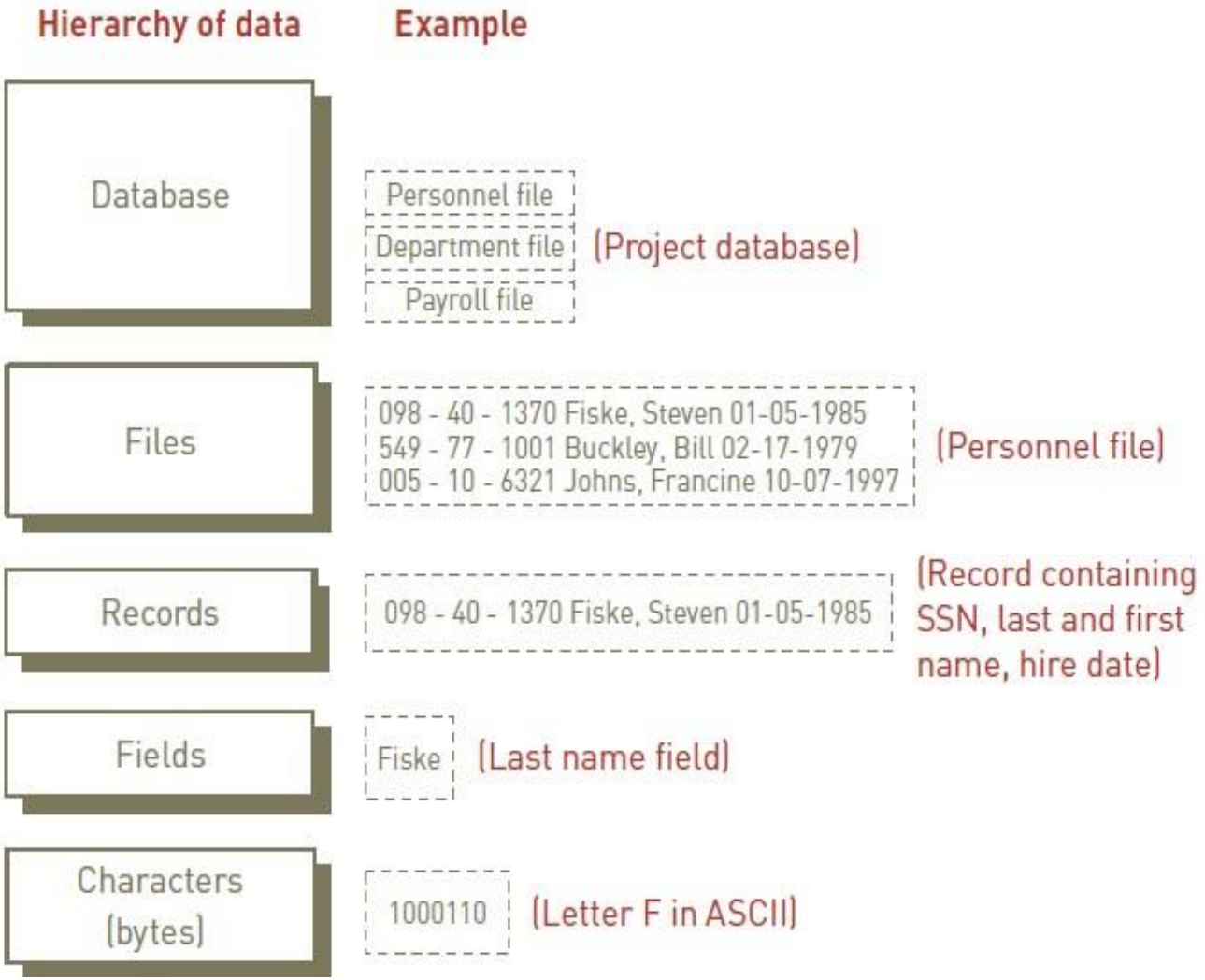
The Hierarchy of Data (continued)



Hierarchy of Data

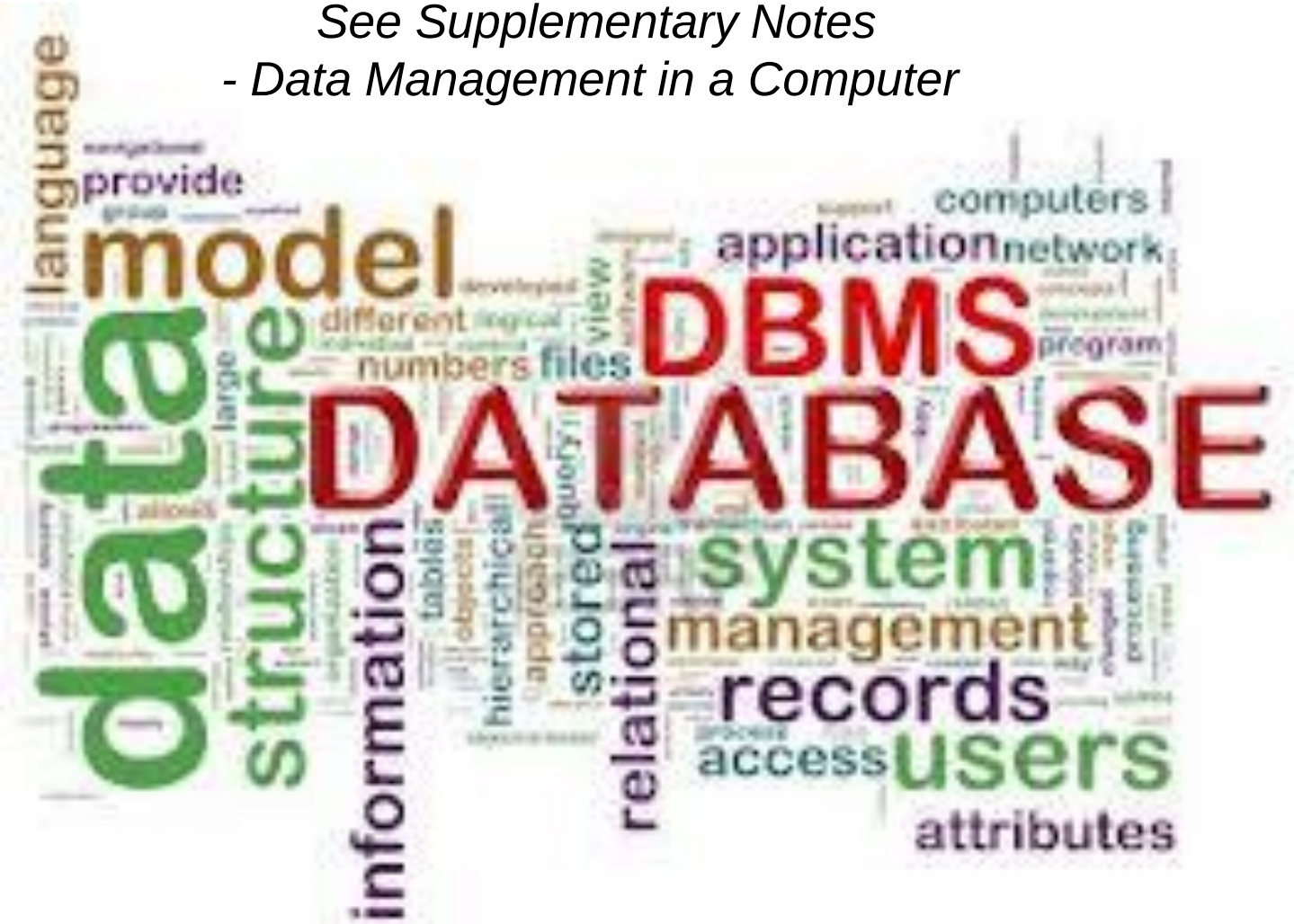


Levels of Data within a Database

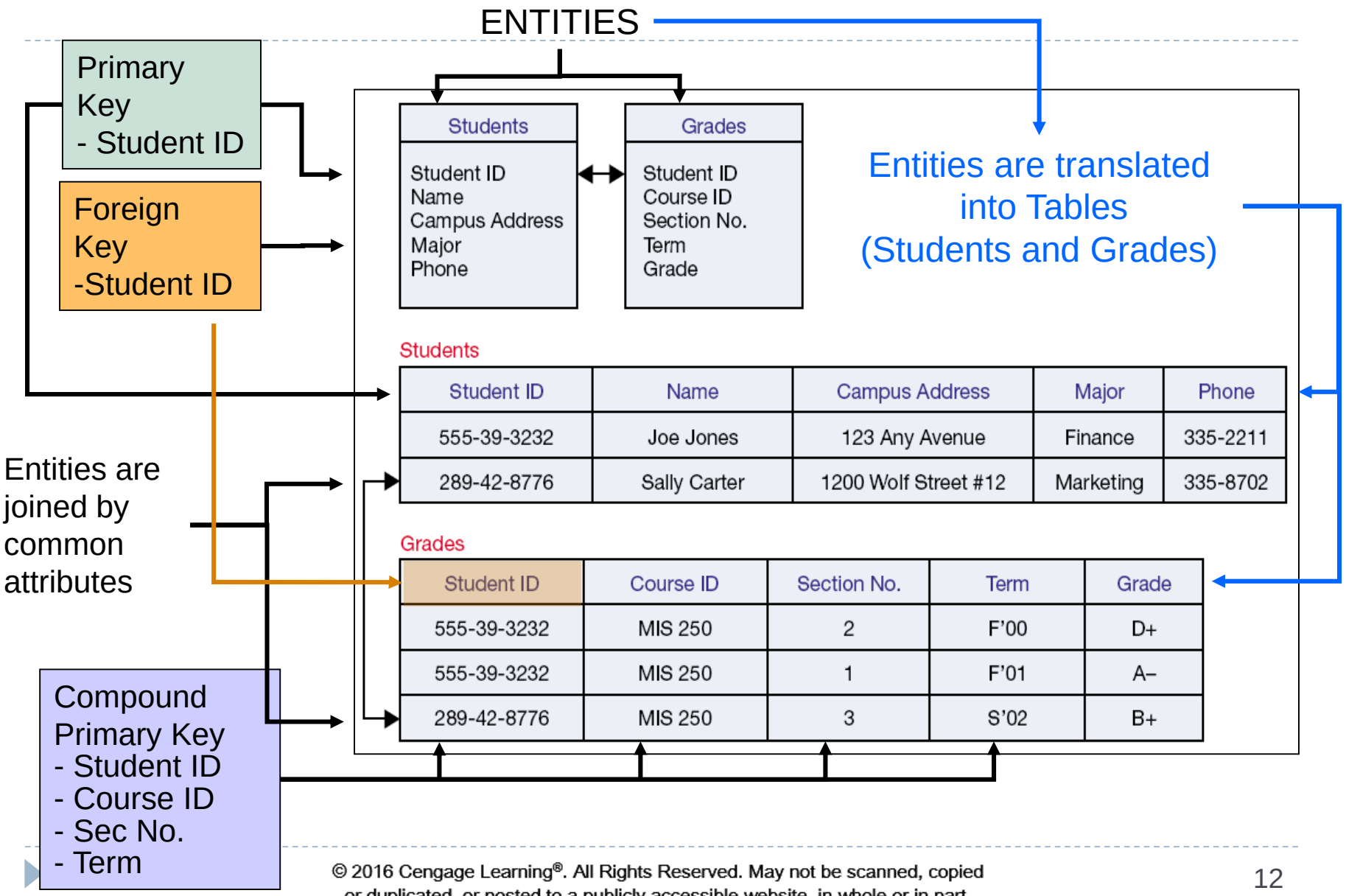


Data Management in a Computer

See Supplementary Notes
- Data Management in a Computer



Data Entities, Attributes, Items, Keys



Students

Student ID	Name	Campus Address	Major	Phone
555-39-3232	Joe Jones	123 Any Avenue	Finance	335-2211
289-42-8776	Sally Carter	1200 Wolf Street #12	Marketing	335-8702

Grades

Student ID	Course ID	Section No.	Term	Grade
555-39-3232	MIS 250	2	F'00	D+
555-39-3232	MIS 250	1	F'01	A–
289-42-8776	MIS 250	3	S'02	B+

Defining Database – Keys

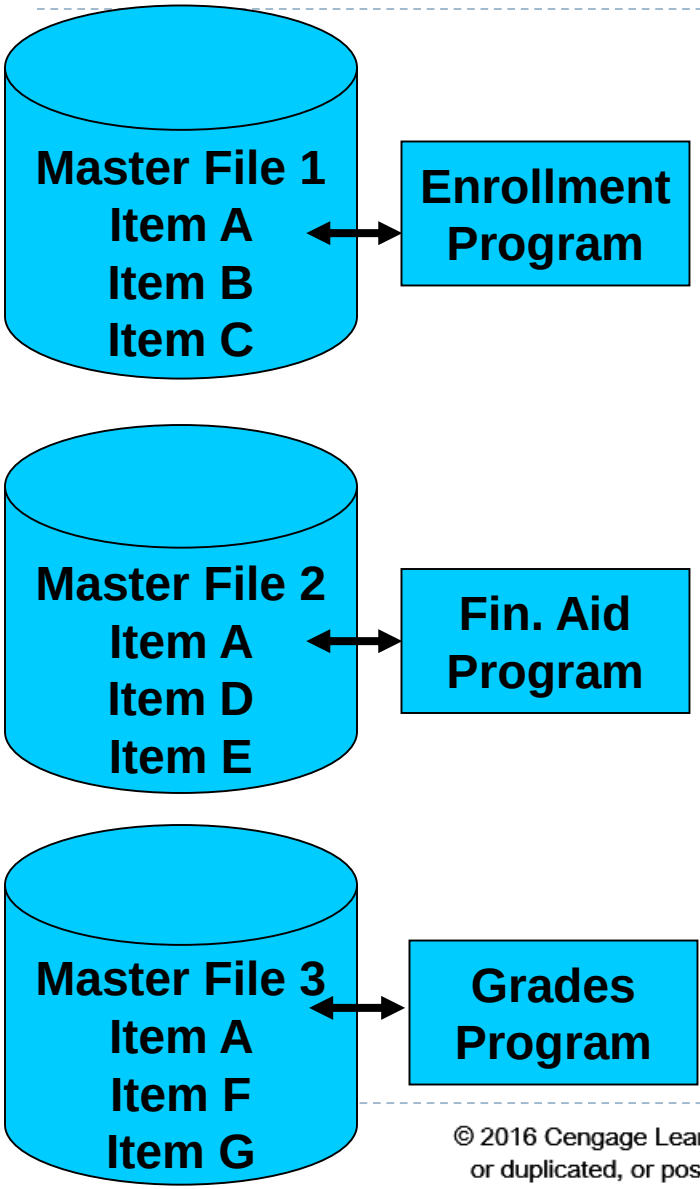
Database Keys
Mechanisms used to identify, select, and maintain one or more records using an application program, query, or report

Primary Key
A unique attribute type used to identify a single instance of an entity

Compound Primary Key
A unique combination of attribute types used to identify a single instance of an entity

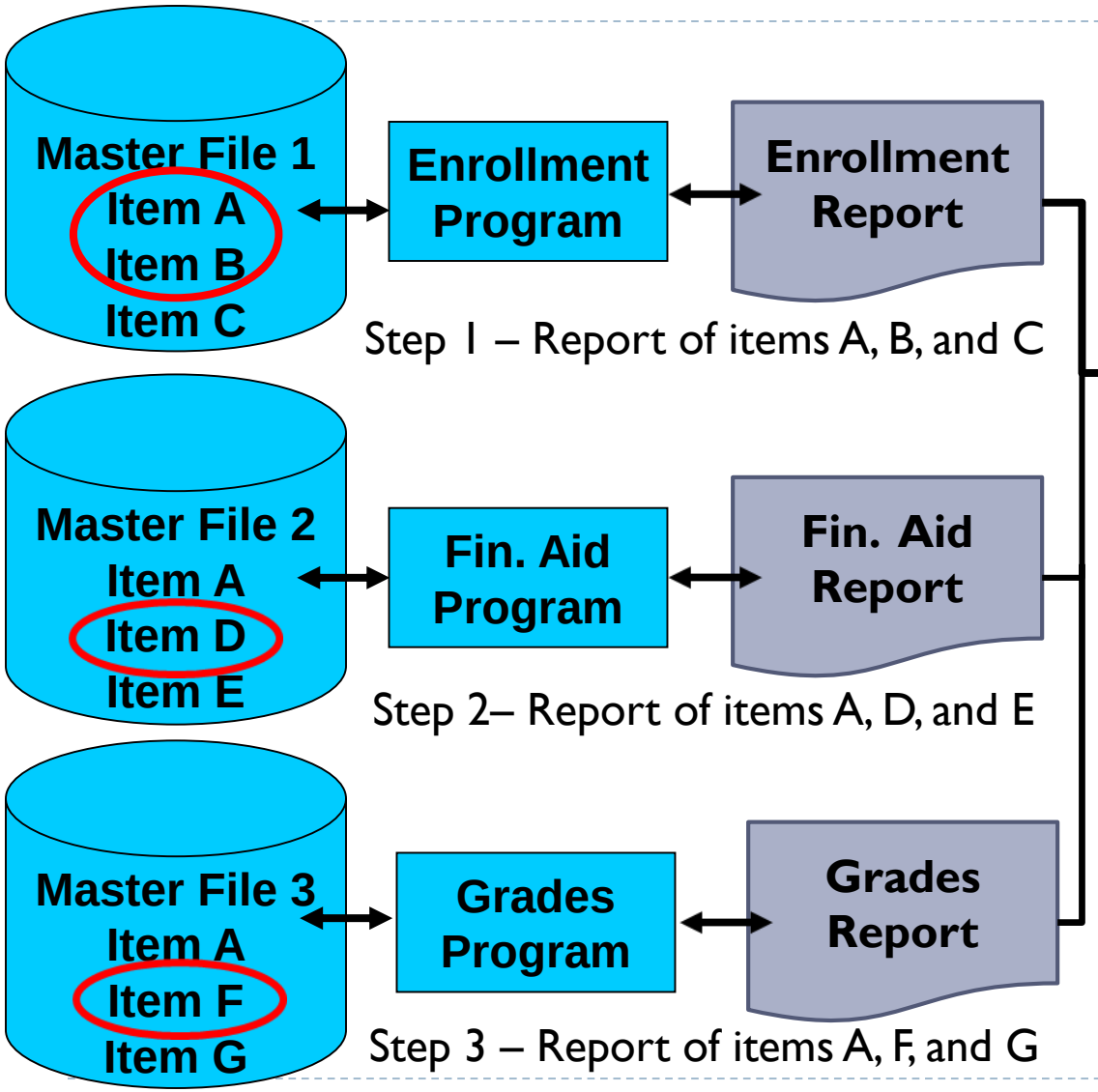
Foreign Key
- **An attribute that appears as a non-primary key in one entity (table) and as a primary key attribute in another entity (table)**

File Approach to Storing Data



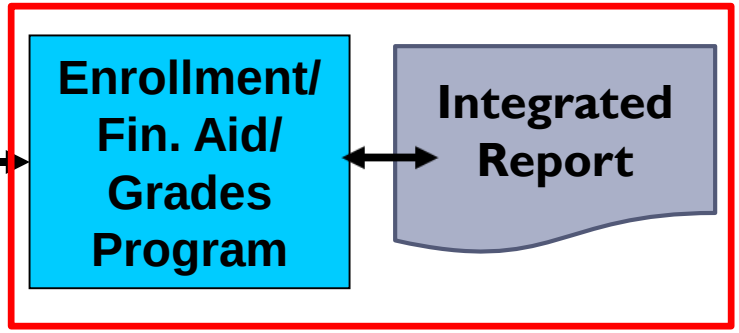
- Often the same information was stored in multiple master files.
- Made it more difficult to effectively integrate data and obtain an organization-wide view of the data.
- Also, the same information may not have been consistent between files.
  - If a student changed his/her phone number, it may have been updated in one master file but not another.

File Approach to Storing Data



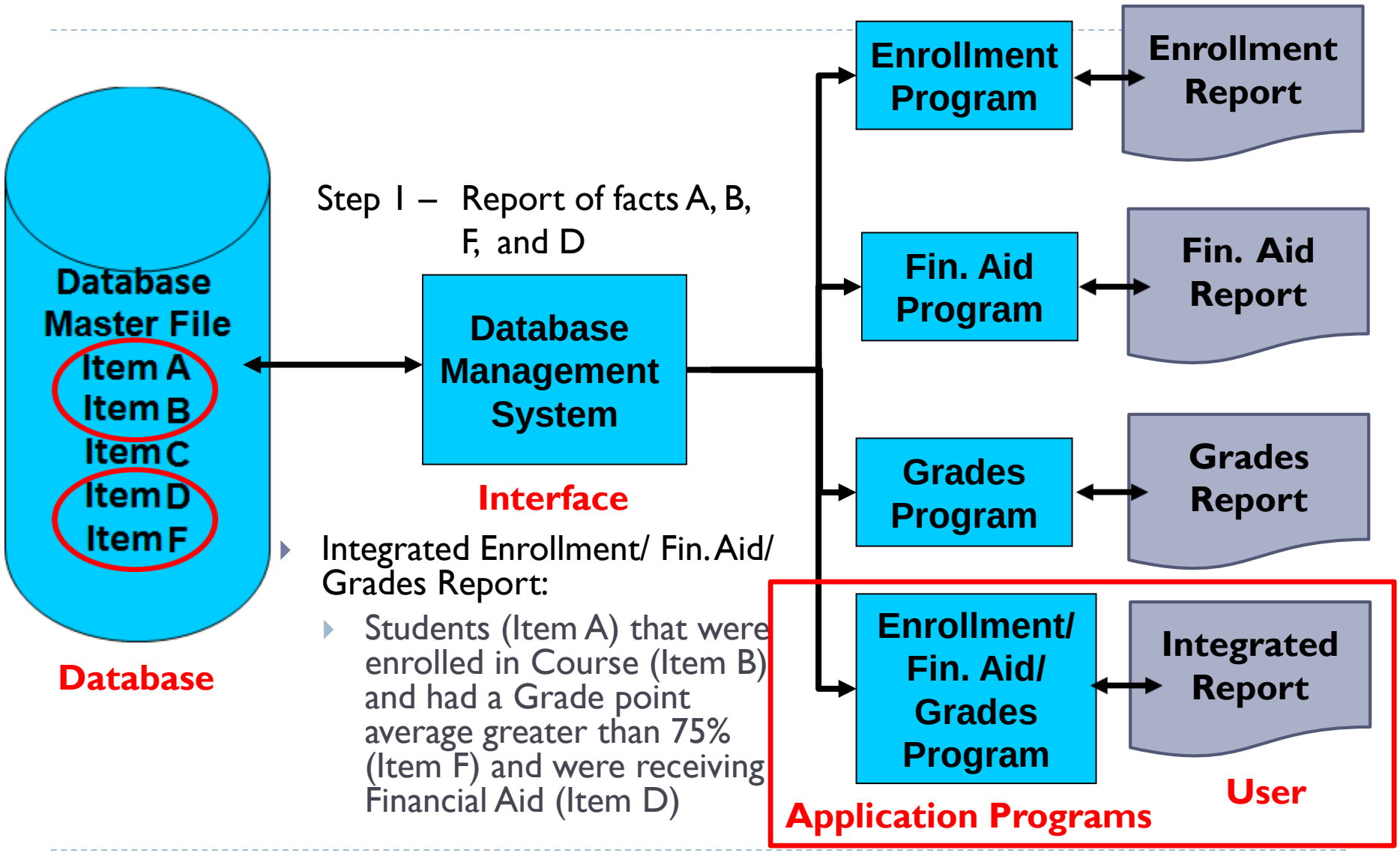
Step 4 – Input items A, B, D, and F

Step 5 – Report of items A, B, F, and D



- Integrated Enrollment/ Fin. Aid/ Grades Report:
  - Students (Item A) that were enrolled in Course (Item B) and had a Grade point average greater than 75% (Item F) and were receiving Financial Aid (Item D)

Database Approach to Storing Data



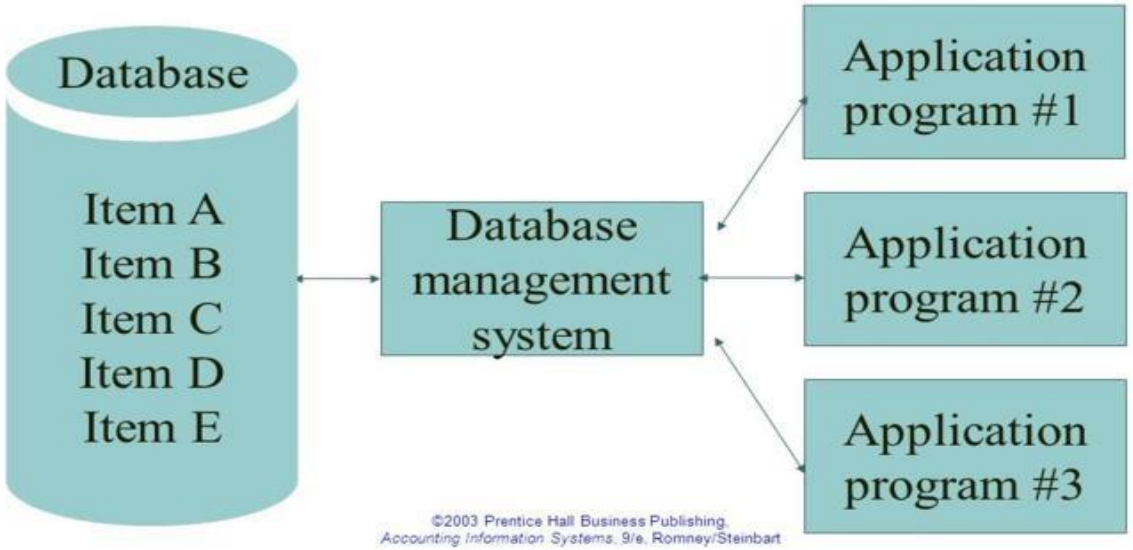
- Integrated Enrollment/ Fin. Aid/ Grades Report:
  - Students (Item A) that were enrolled in Course (Item B) and had a Grade point average greater than 75% (Item F) and were receiving Financial Aid (Item D)

File Approach Characteristics

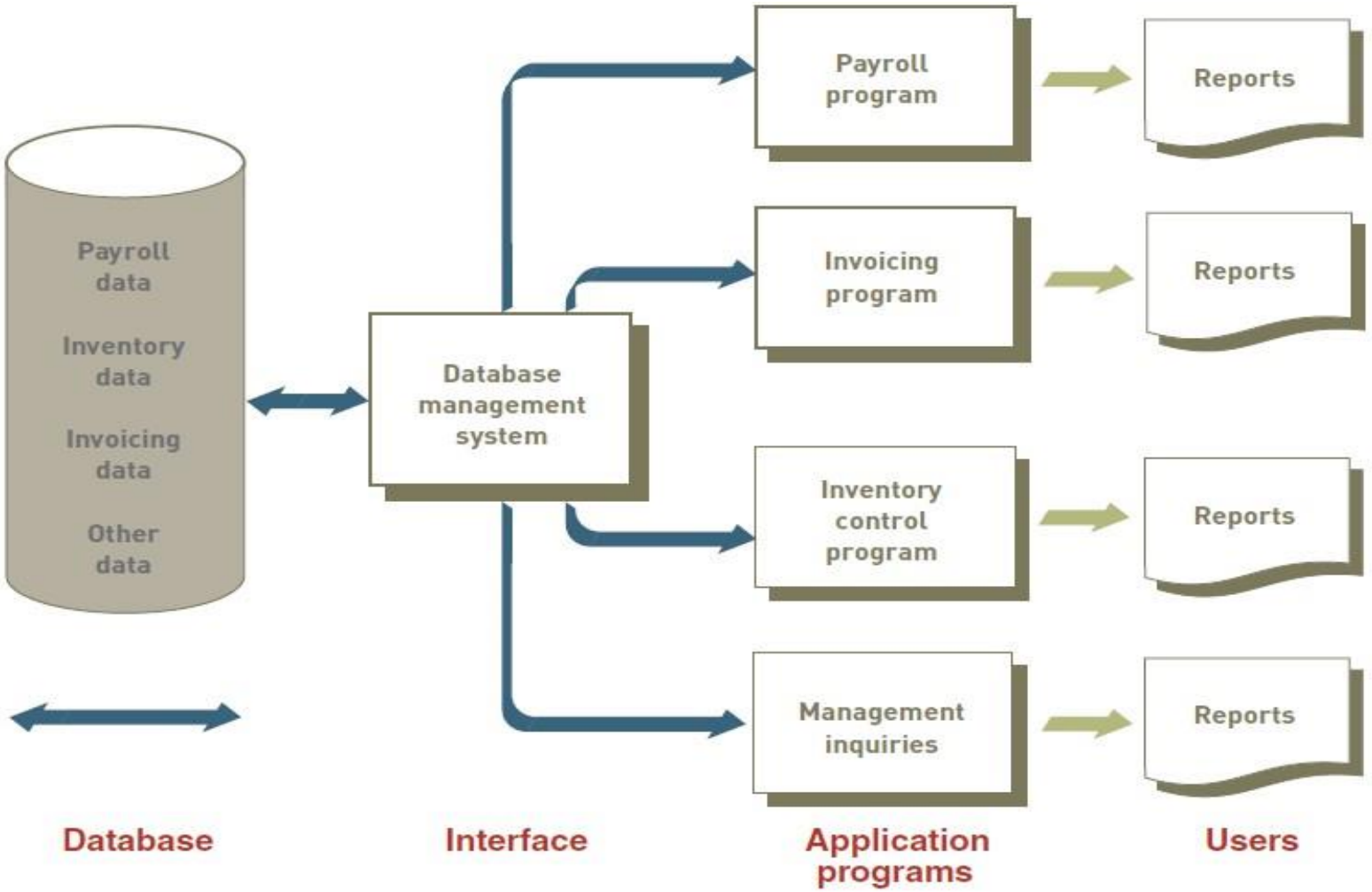
- **Uncontrolled Redundancy**: If these separate applications need to process the same data there are duplicate copies of the data Wastage of valuable storage space.
  - Need to input data to several files.
  - Data inconsistency (one fact may have more than one value – various versions may occur).
- **Poor Enforcement of System Standards**: Data names, formats, access restrictions, … etc. are not standardized across an organization. This makes modifications difficult and hinders sharing of data.
- **Limited Data Sharing**: Each application has its own private file providing little opportunity for users to share existing data. Any new applications would not be able to use existing files leading to low productivity.
- **Program** – Data Dependency: Descriptions of files, records, data items are embedded within application programs. Any modification to a data file requires that the application programs using that file must also be changed. In other words, program maintenance will be excessive

Database Approach Characteristics

- A centralized Database Management System (DBMS) exists, which handles all data management activities.
- The DBMS does not fragment data into separate files but regards data as being stored in a large conceptual repository - *database*. The DBMS handles the addition, storage, update, and retrieval of data.



Database Approach - Overview



Database Approach - Advantages

- Improved strategic use of corporate data
- Reduced Data Redundancy
- Improved Data Integrity
- Easier modification and updating
- Data and program independence
- Better access to data and information
- Standardization of data access
- Improved data safeguarding
- Efficient use of resources

Database Approach - Challenges

Disadvantages	Explanation
More complexity	DBMSs can be difficult to set up and operate. Many decisions must be made correctly for the DBMS to work effectively. In addition, users have to learn new procedures to take full advantage of a DBMS.
More difficult to recover from a failure	With the traditional approach to file management, a failure of a file affects only a single program. With a DBMS, a failure can shut down the entire database.
More expensive	DBMSs can be more expensive to purchase and operate than traditional file management. The expense includes the cost of the database and specialized personnel, such as a database administrator, who is needed to design and operate the database. Additional hardware might also be required.

File Processing vs Database Approach

File Processing Approach	Database Approach
• **Storage Media**: Sequential tapes or files	• Direct Access Storage Device (DASD)
• **Data**: Stored in long sequential files (no relationship with other files)	• Stored in tables with relationships with other files
• **Organization**: Redundant data in multiple files	• Redundant data minimized/ eliminated
• **Updates**: Requires multiple updates in many files	• Requires few or one update for a data field
• **Processing:** Slower query/ faster processing	• Faster query / slower processing