ACS-1803
Introduction to Information Systems

Instructor: Victor Balogun

Data Management

Lecture Outline 2, Part 1

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)

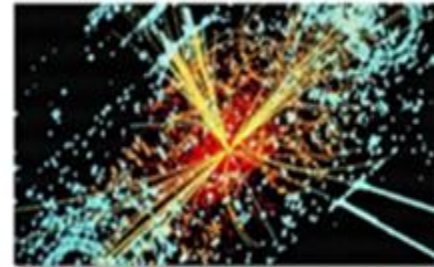 ▸ A collection of programs that enables users to store, modify, and extract information from a database

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

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

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

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

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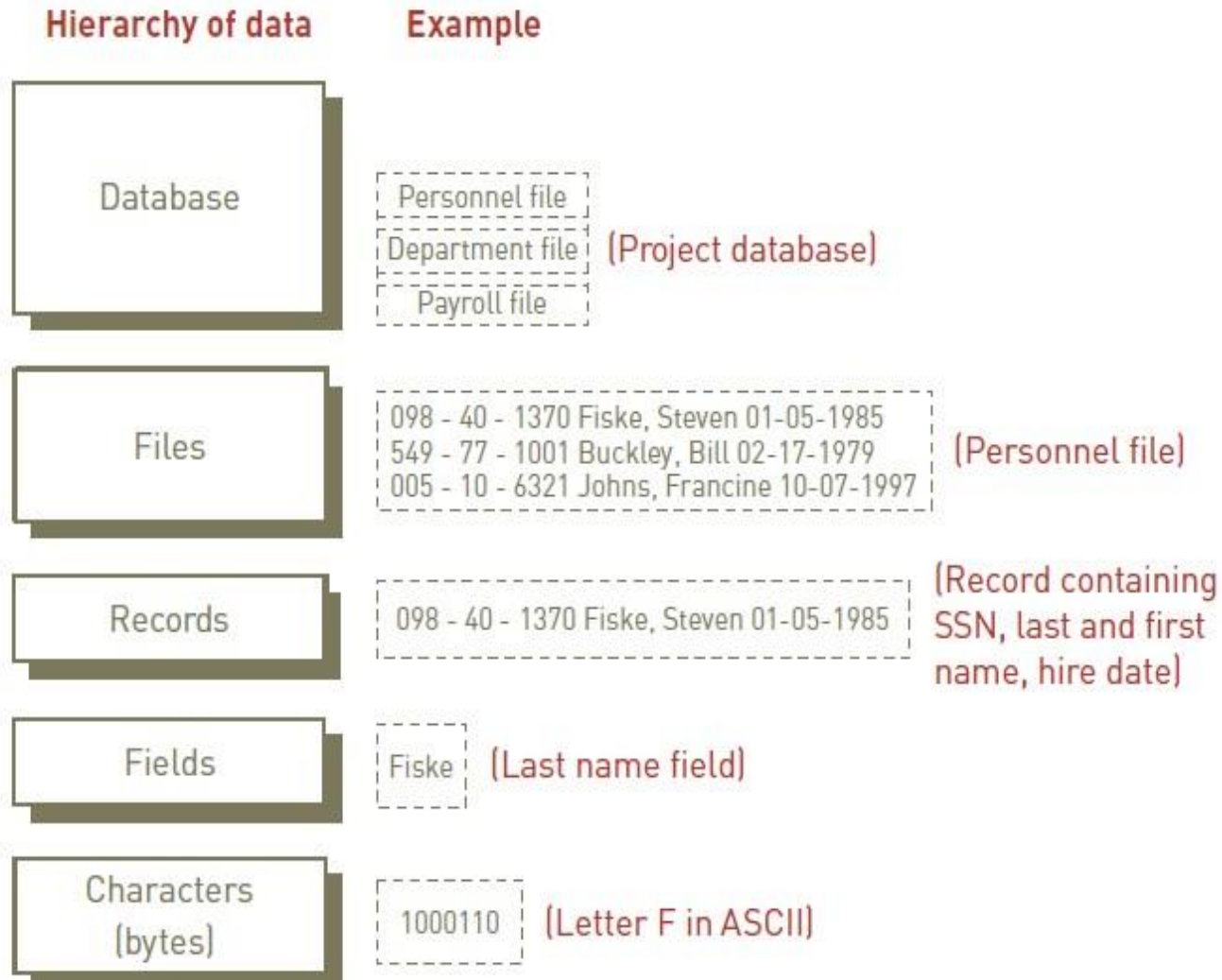- geospatial query language

Many everyday devices contain databases
- TVs, washing machines, mobile phones, ...
- e.g. Android phones with SQLite database
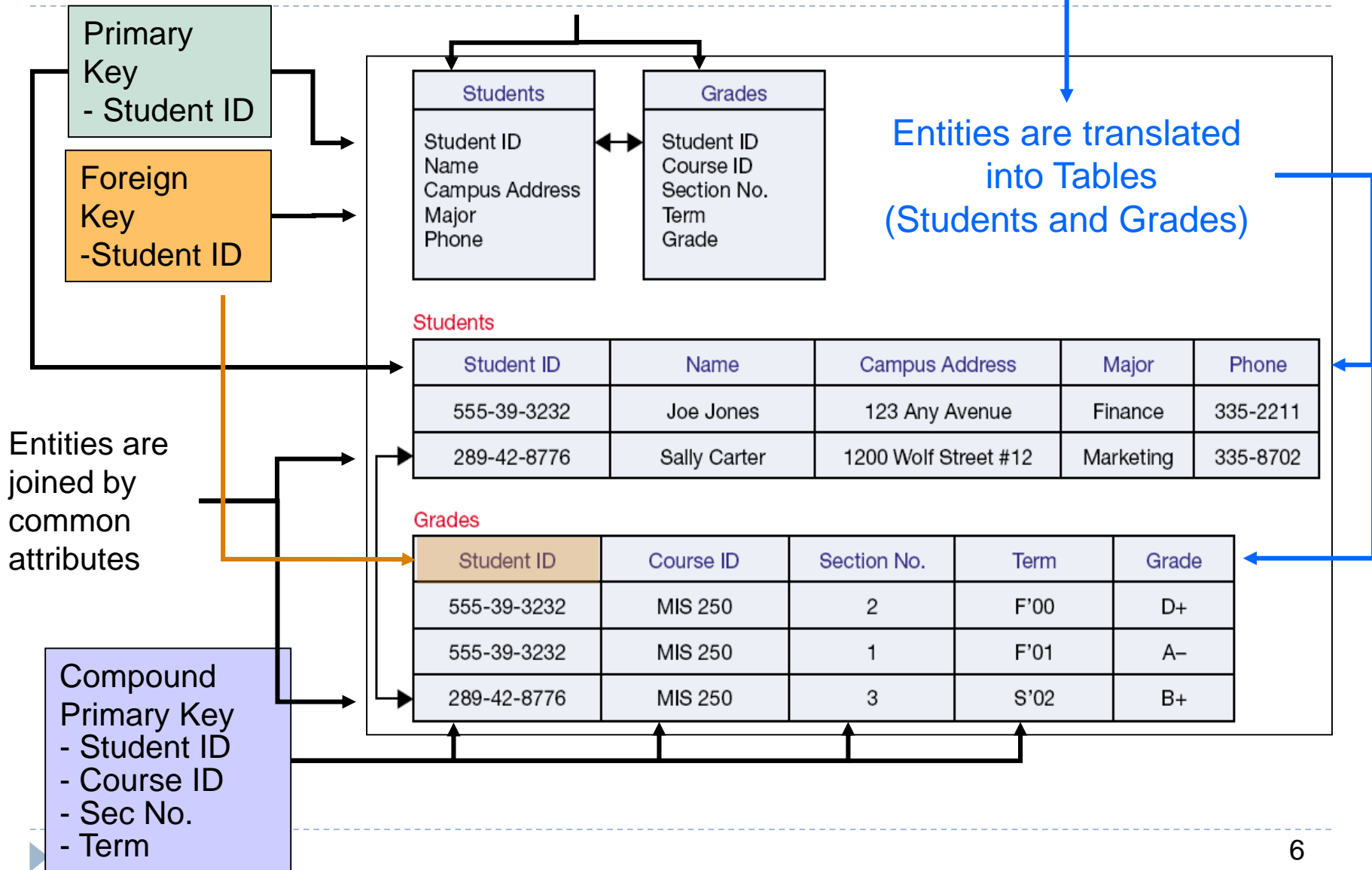
Levels of Data within a Database

Level	Term	Term Definitions
Highest	Database	A collection of files or entities containing information to support a given system or a particular topic area
	File	A collection of records or instances for a given ENTITY. These are also called **tables**, depending on the DBMS
	Record	A group of fields or attributes to describe a single instance of an ENTITY. These are also called **rows** depending on the DBMS
	Field	Individual characteristics about an ENTITY. Fields are also called **attributes or columns** based on the type of DBMS
	Character	Basic building block of information that defines attributes.
Lowest	Bit	Binary digit that defines a character. Communicates with the circuit technology within a computer through on/off code switches.

4

Levels of Data within a Database

Hierarchy of data	Example
Database	Personnel file Department file (Project database) Payroll file
Files	098 - 40 - 1370 Fiske, Steven 01-05-1985 549 - 77 - 1001 Buckley, Bill 02-17-1979 (Personnel file) 005 - 10 - 6321 Johns, Francine 10-07-1997
Records	098 - 40 - 1370 Fiske, Steven 01-05-1985 (Record containing SSN, last and first name, hire date)
Fields	Fiske (Last name field)
Characters (bytes)	1000110 (Letter F in ASCII)

Data Entities, Attributes, Items, Keys

ENTITIES

Primary Key
- Student ID

Foreign Key
-Student ID

Entities are joined by common attributes

Compound Primary Key
- Student ID
- Course ID
- Sec No.
- Term

Entities are translated into Tables
(Students and Grades)

Students

Students
Student ID
Name
Campus Address
Major
Phone

Grades

Grades
Student ID
Course ID
Section No.
Term
Grade

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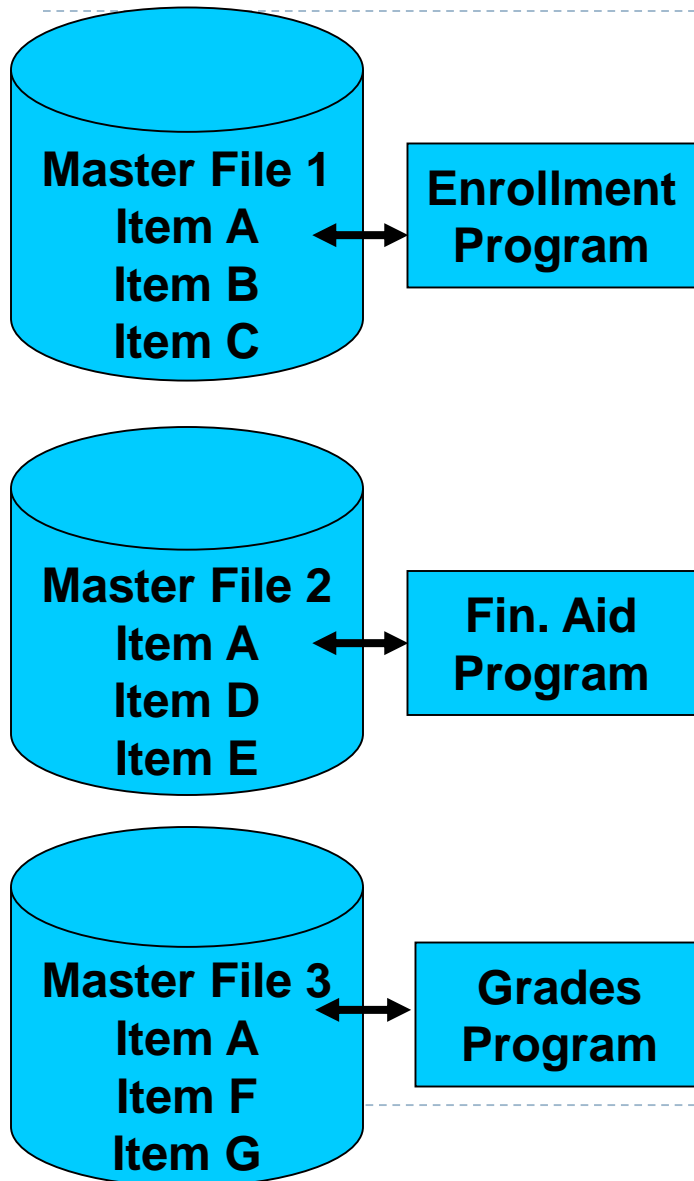A unique attribute type used to identify a single instance of an entity

Compound Primary Key
A unique combination of attribute types used to identify a single instance of an entity
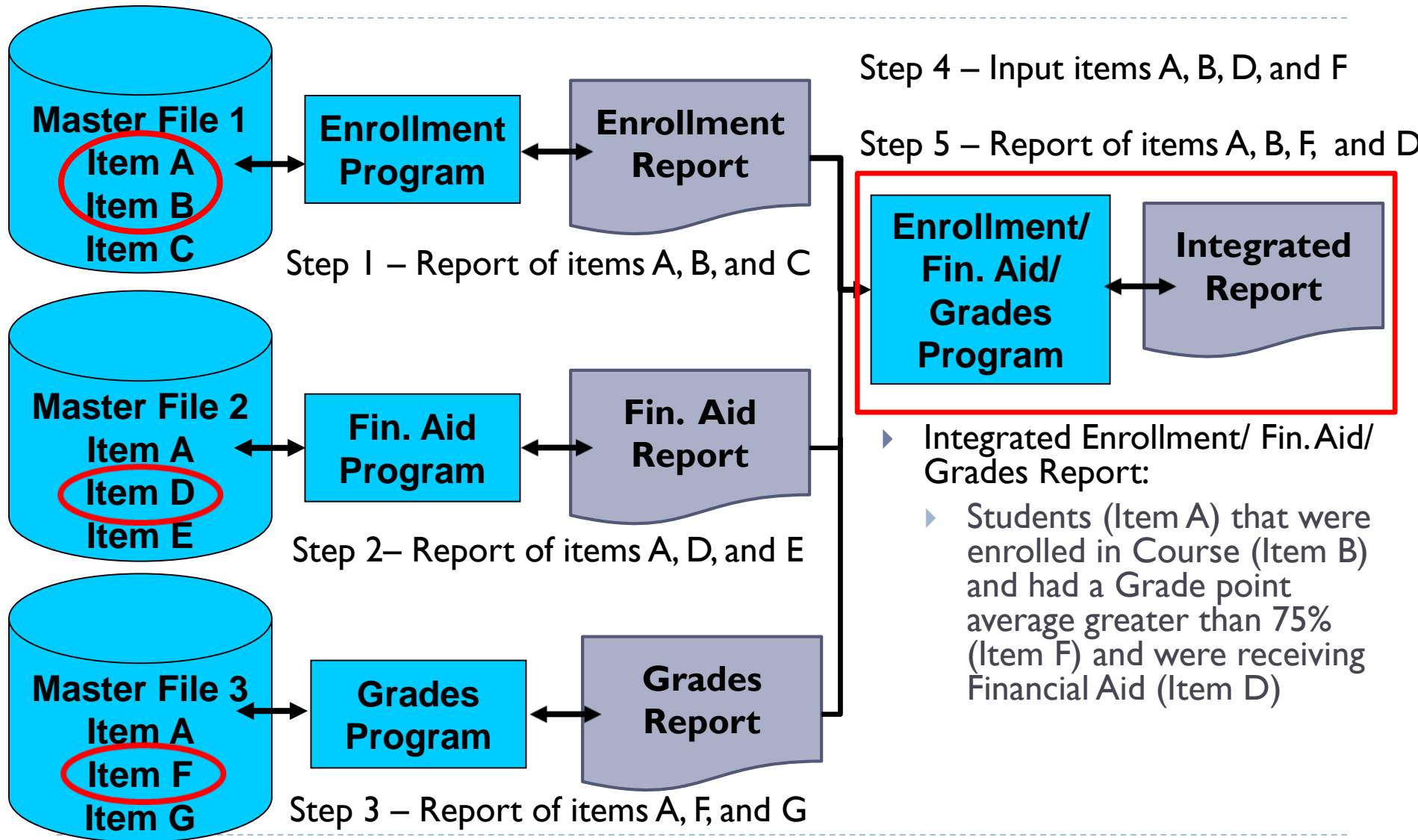
Foreign Key
• An attribute that appears as a non-primary key in one entity (table) and as a primary key attribute in another entity (table)
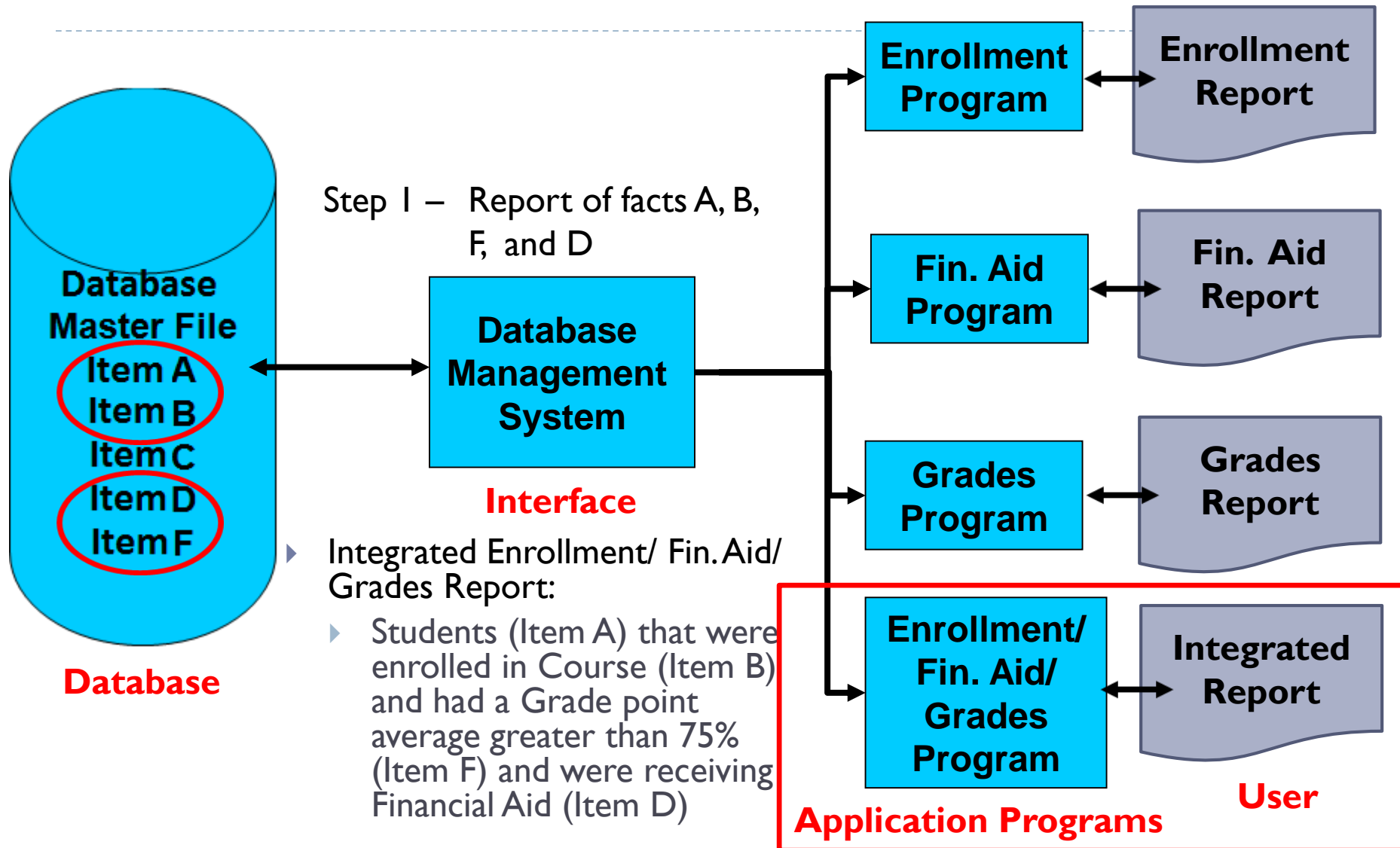
File Approach to Storing Data

Master File 1
  Item A
  Item B
  Item C

Enrollment Program

Master File 2
  Item A
  Item D
  Item E

Fin. Aid Program

Master File 3
  Item A
  Item F
  Item G

Grades Program

▸ 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

Master File 1
Item A
Item B
Item C

Enrollment Program

Enrollment Report

Step 1 – Report of items A, B, and C

Master File 2
Item A
Item D
Item E

Fin. Aid Program

Fin. Aid Report

Step 2– Report of items A, D, and E

Master File 3
Item A
Item F
Item G

Grades Program

Grades Report

Step 3 – Report of items A, F, and G

Step 4 – Input items A, B, D, and F

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

Enrollment/ Fin. Aid/ Grades Program

Integrated Report

▸ 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

Step 1 – Report of facts A, B, F, and D

Interface

▸ 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

Application Programs

User

File Approach Characteristics

▸ Each application has its own set of files.

▸ **Inflexibilit**y: committed to processing particular queries.

▸ **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.
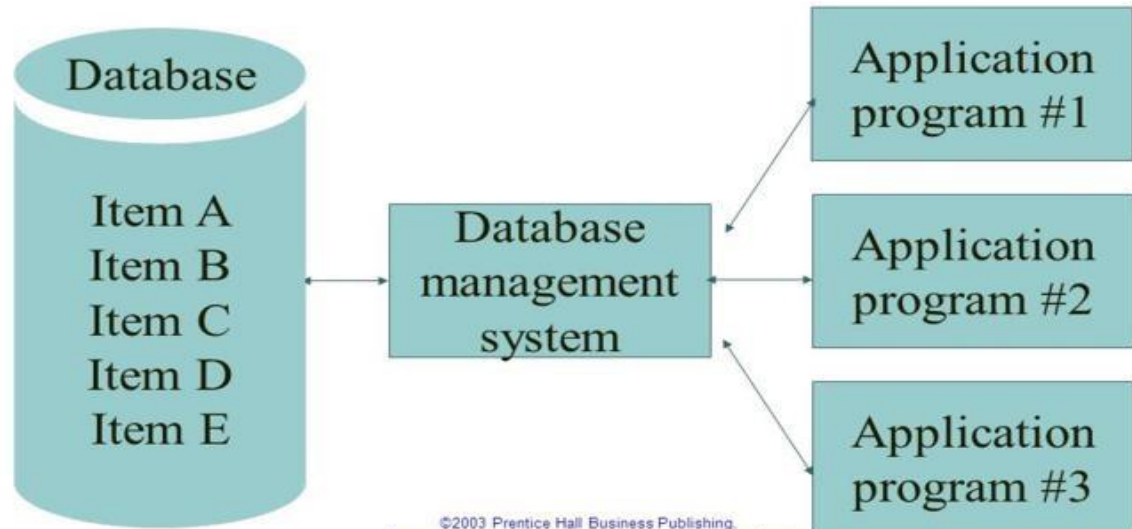
©2003 Prentice Hall Business Publishing,
Accounting Information Systems, 9/e, Romney/Steinbart

Database Approach

| Database | Interface | Application programs | Users |
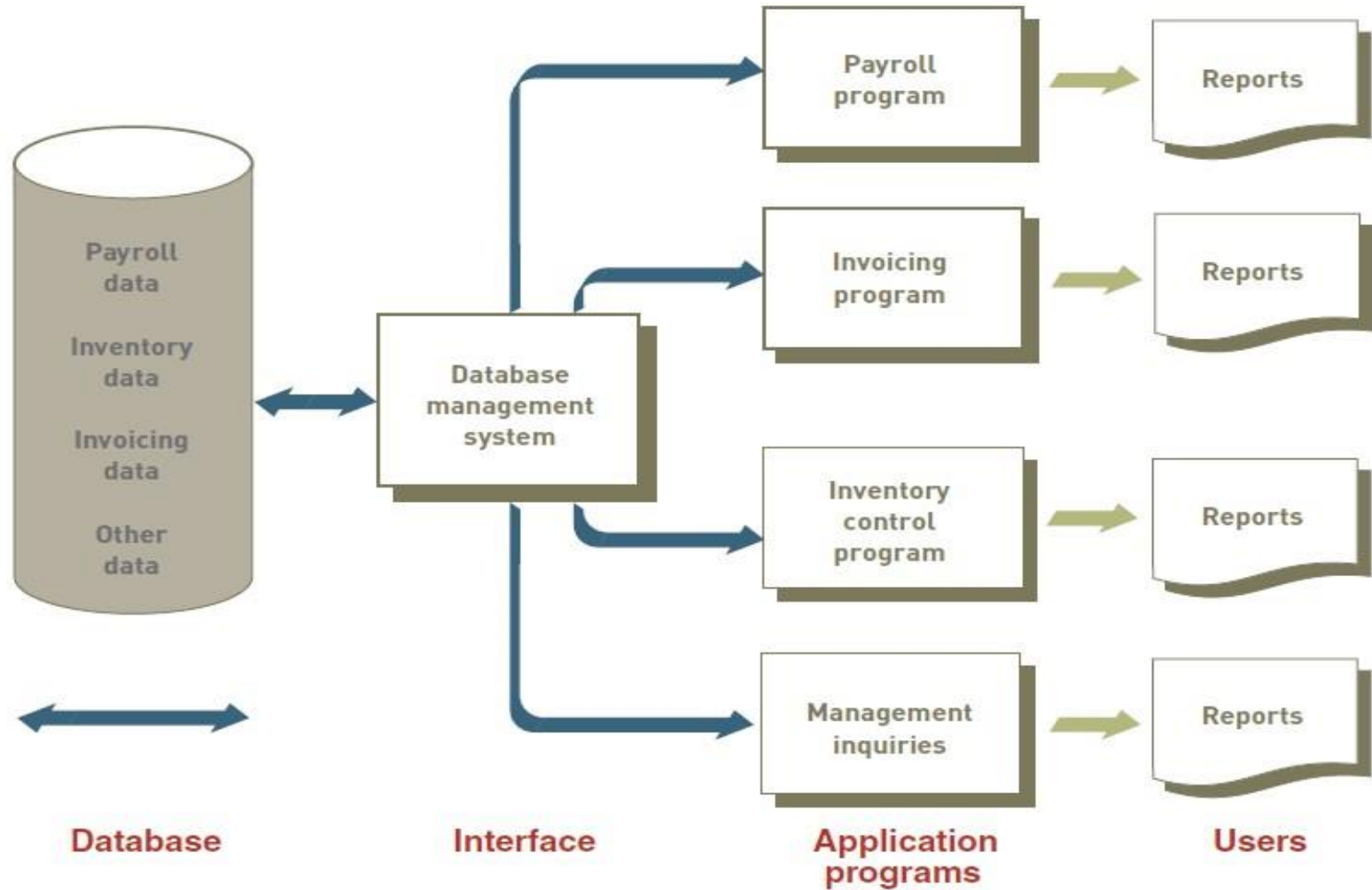
Database Approach - Advantages

- 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 – Disadvantages

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 (Old School)	Data Base Approach (Current)
• **Storage Media**: sequential tapes or files	• <u>D</u>irect <u>A</u>ccess <u>S</u>torage <u>D</u>evice (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