Lecture1 – Introducing Database Concepts

ICDT 1202Y - DATABASE SYSTEMS

Learning Outcomes

After this session you will be able to :

- Discuss on the importance of information
- Differentiate between data and information
- Explain the importance of database
- Explain the role of a DBMS
- Describe the database system architecture
- Explain how to go about the various steps for developing a database
- Identify the various people involved in the field of database

THE NEED FOR INFORMATION

- Vast variety of information is present in organizations today
- If information has value then it is an asset for the organization depending on strategy adopted
- Investment in IT contributes to proper information management
- IT delivers economic value to businesses which can be see through an increase in competitiveness, high productivity, increase revenue

DATA, INFORMATION & METADATA

Data - facts that are recorded and can be accessed

- Data formats – text, numbers, figures, graphics, images, audio/video recordings and more
- Data is recorded and kept because it is considered to be of use to an intended user

Information - refers to the data that is accessed by a user for some particular purpose

- Typically, getting the needed information from a collection of data requires performing an activity, such as searching through, processing, or manipulating the data in some form or fashion

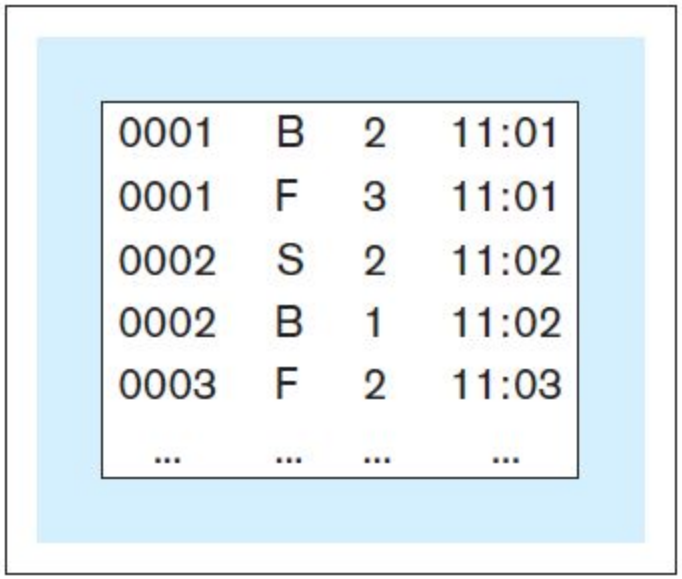
DATA, INFORMATION & METADATA(2)

Metadata - data that describes the structure and the properties of the data

- Metadata is essential for the proper understanding and use of the data

DATA, INFORMATION & METADATA (3)

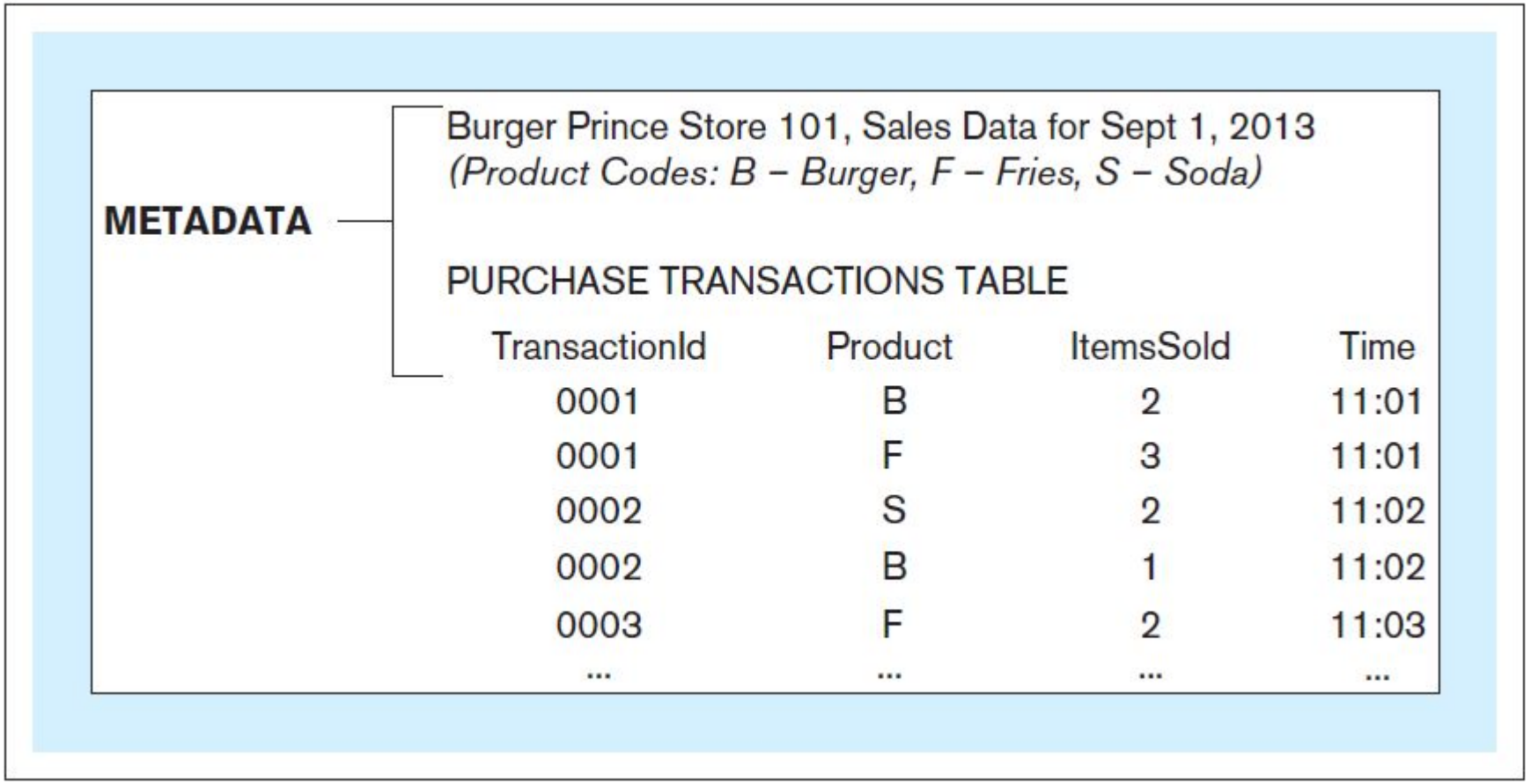
Data without metadata - example



0001	B	2	11:01
0001	F	3	11:01
0002	S	2	11:02
0002	B	1	11:02
0003	F	2	11:03
...	...	...	...

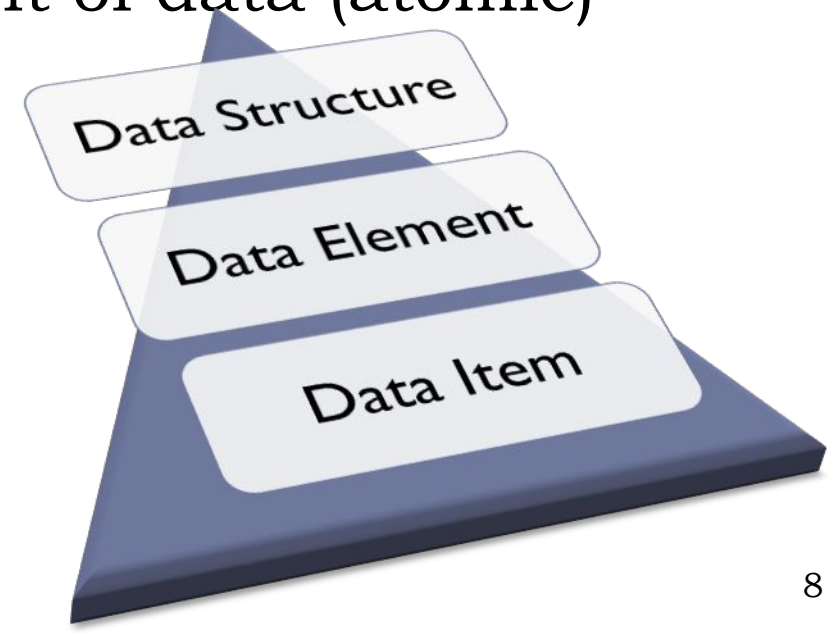
DATA, INFORMATION & METADATA(4)

Data with metadata - example



ORGANISING DATA

- Data model : provides set of principles for organising data
- Data structure : logical collection of data element
- Data element : group of data item
- Data item: simplest element of data (atomic)



Organising data _ File Based

- File based : Most popular data model for decades
- Logically organised constructs of fields (data items), records (data elements) and files (data structures)

File Based Systems

- Collection of application programs that perform services for end-users
- Each program defines and manages its own data
- File based systems:
  - Can lack efficient access
  - Have no direct support for queries
  - Limit organisation to directory creation and hierarchical organization
  - Have no sophisticated support for concurrency
  - Do not ensure durability

File Based Systems - Example

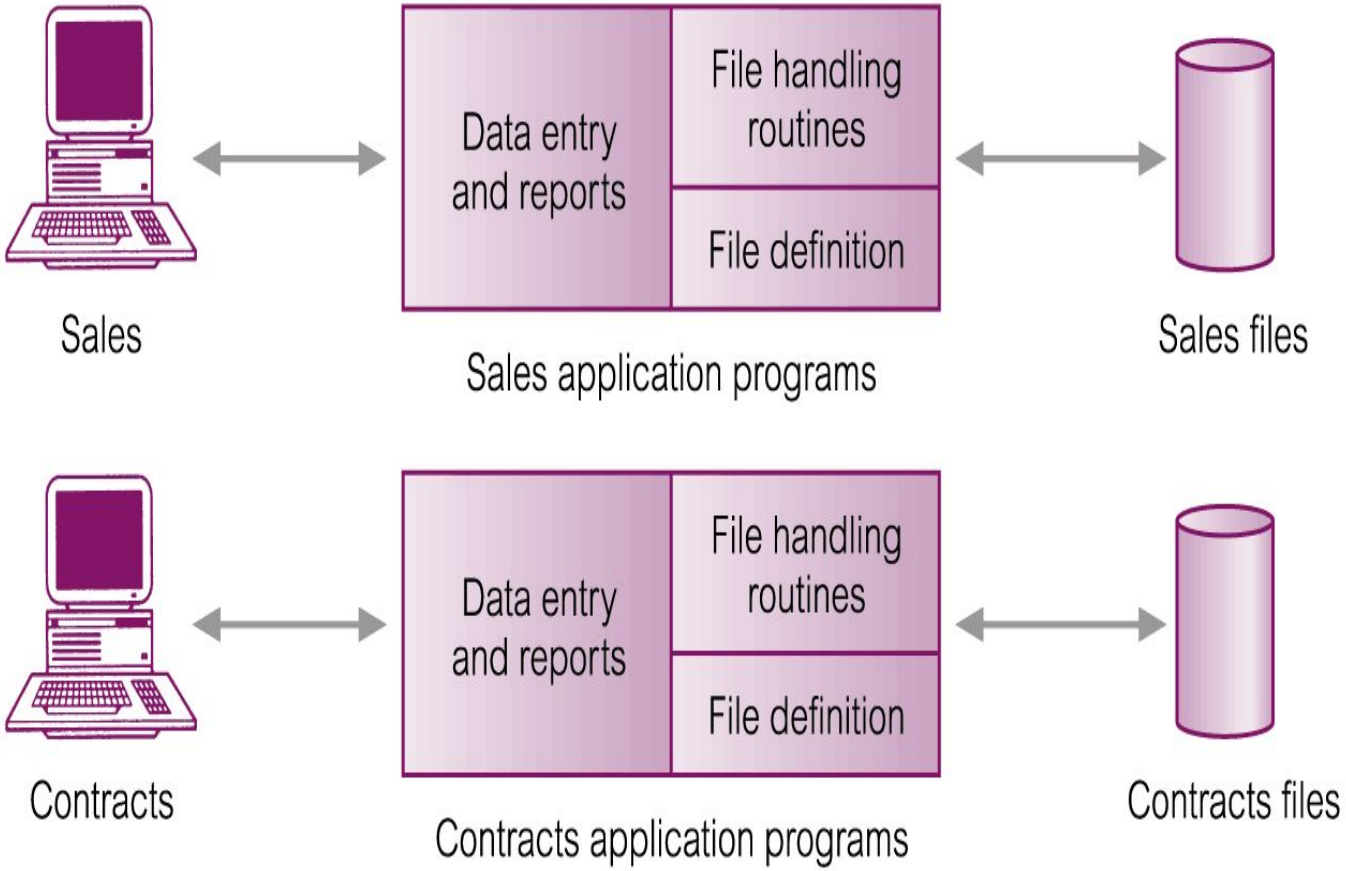


Figure 1

File-base
processir

Drawbacks of Using File Systems to Store Data

- Data Redundancy and Inconsistency
  - Multiple file formats, duplication of information in different files
- Difficulty in accessing data
  - Need to write a new program to carry out a new task
- Data isolation
  - Multiple files may be in different formats , difficult to write new programs to retrieve appropriate data
- Integrity Problems
  - Integrity constraints (balance > 0) become part of program code
  - Hard to add new constraints or add existing ones

Drawbacks of Using File Systems to Store Data

- Atomicity of updates
  - Failures may leave database in an inconsistent state with partial updates carried out
- Concurrent access by multiple users
  - Concurrent accesses needed for performance
  - Uncontrolled concurrent accesses can lead to inconsistencies
    - E.g. two people reading a balance and updating it at the same time
- Security Problems

DATABASE APPROACH

- Controlled access to database may include:
  - a security system
  - an integrity system
  - a concurrency control system
  - a recovery control system
  - a user-accessible catalog.

WHAT IS A DATABASE?

- A database is a **shared** collection of **logically related** data (and a description of this data), designed to **meet the information needs of an organization**
- It is a collection of information that is organized so that it can easily be **accessed**, **managed** and **updated**
- The **system catalog** (metadata) provides the description of data to enable program-data independence

WHAT IS A DATABASE?

- Store large quantities of information
- Enable information sharing
- Allow for flexible use of data

DATABASE APPLICATIONS

- Supermarket Stock Control
- University Registration System
- Video/Car Rental
- Travel Agency
- Library Information System
- Reservation Systems
- Banking Systems
- Bioinformatics, e.g. gene databases
- Criminal Justice, e.g. fingerprint matching
- Multimedia Systems, e.g. image/audio/video retrieval
- Satellite imaging
- The web

In sum, databases are everywhere!!!

DATABASE TERMINOLOGIES

- Table: a group of related records

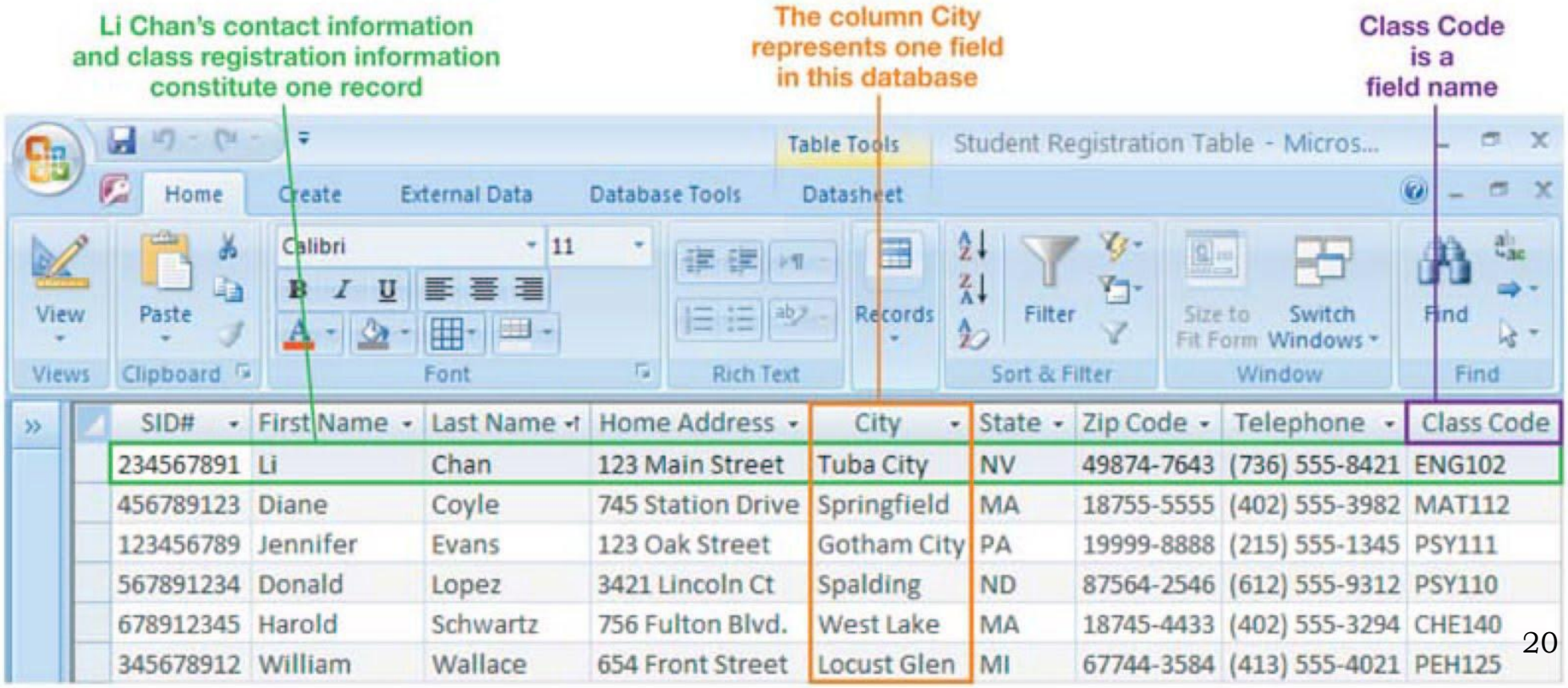
SS#	Last Name	First Name	Home Address	City	State	Zip Code	Residence Hall	Room #
234567891	Chan	Li	123 Main Street	Tuba City	NV	49874-7643	Wilson Hall	218
456789123	Coyle	Diane	745 Station Drive	Springfield	MA	18755-5555	Montgomery Hall	231
987654321	Evans	Jennifer	312 Brookside Road	Springfield	MA	18755-4444	Forsman Quad	127
123456789	Evans	Jennifer	123 Oak Street	Gotham City	PA	19999-8888	Montgomery Hall	312
567891234	Lopez	Donald	3421 Lincoln Court	Spalding	ND	87564-2546	Stone House	102
345678912	Wallace	William	654 Front Street	Locust Glen	MI	67744-3584	Forsman Quad	124

DATABASE TERMINOLOGIES

- **Database** - structured collection of related data stored on a computer medium
  - Organizes the data in a way that facilitates efficient access to the information captured in the data
- **Database metadata** – represents the structure of the database
  - Database content that is not the data itself (data about the data)
  - Contains:
    - Names of data structures
    - Data types
    - Data descriptions
    - Other information describing the characteristics of the data

DATABASE TERMINOLOGIES

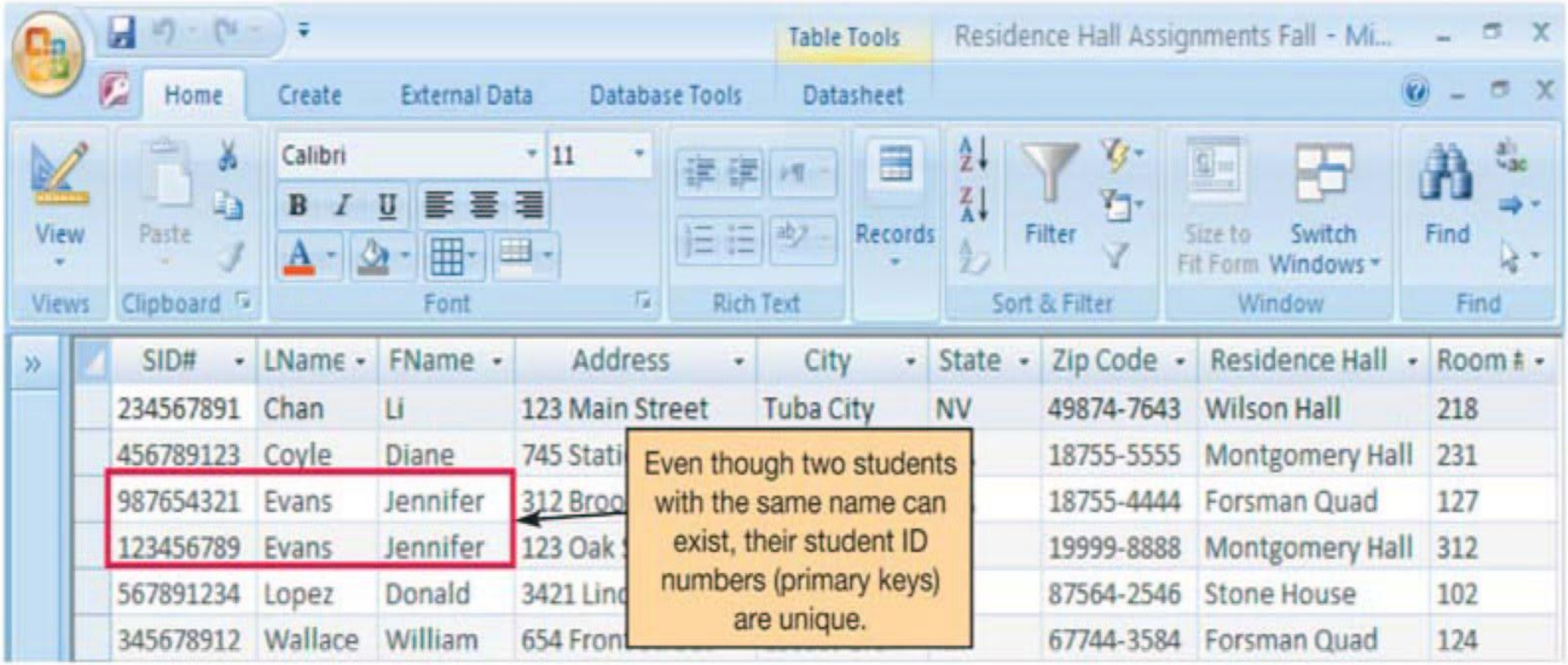
- Field: category of information displayed in columns
- Record: a group of related fields



SID#	First Name	Last Name	Home Address	City	State	Zip Code	Telephone	Class Code
234567891	Li	Chan	123 Main Street	Tuba City	NV	49874-7643	(736) 555-8421	ENG102
456789123	Diane	Coyle	745 Station Drive	Springfield	MA	18755-5555	(402) 555-3982	MAT112
123456789	Jennifer	Evans	123 Oak Street	Gotham City	PA	19999-8888	(215) 555-1345	PSY111
567891234	Donald	Lopez	3421 Lincoln Ct	Spalding	ND	87564-2546	(612) 555-9312	PSY110
678912345	Harold	Schwartz	756 Fulton Blvd.	West Lake	MA	18745-4433	(402) 555-3294	CHE140
345678912	William	Wallace	654 Front Street	Locust Glen	MI	67744-3584	(413) 555-4021	PEH125

DATABASE TERMINOLOGIES

- Primary Key: a field value unique to a record



DATABASE TERMINOLOGIES

- Data Types: type of data that can be stored in the field

FIGURE 11.5 Common Data Types and Examples of the Types of Information They Can Contain

Data Type	Used to Store	Example of Data Stored in the Field
Text	Alphabetic or alphanumeric data	Mary, CIS110
Numeric	Numbers	256, 1.347, $5600
Computational	Computational formulas	Credit hours x per-credit tuition charges
Date	Dates in standard date notation	4/15/2008
Memo	Long blocks of text	Four score and seven years ago our fathers brought forth on this continent, a new nation, conceived in liberty, and dedicated to the proposition that all men are created equal.
Object	Multimedia files or entire documents	MP3 file, AVI file
Hyperlink	A hyperlink to a Web page on the Internet	**www.prenhall.com/techinaction**

DATABASE MANAGEMENT SYSTEM

- A Database Management System (DBMS) is a set of computer programs that controls the creation, maintenance and the use of the database of an organization and its end users
- It is a powerful tool for creating, managing and manipulating large amounts of data efficiently and allowing it to persist over long periods of time, safely
- The DBMS software enables users to define, create, maintain and control access to the database
- On the other hand, a database application program is a computer program that interacts with database by issuing an appropriate request (SQL statement) to the DBMS

DATABASE MANAGEMENT SYSTEM

- **Database management system (DBMS)** - software used for:
  - Creation of databases
  - Insertion, storage, retrieval, update, and deletion of the data in the database
  - Maintenance of databases
- **Database system** - computer-based system whose purpose is to enable an efficient interaction between the users and the information captured in a database

24

TYPICAL DBMS FUNCTIONALITY

- Define a database: in terms of data types, structures and constraints
- Construct or Load the database on a secondary medium
- Manipulate 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
- Other Features:
  - Protection or Security measures to prevent unauthorized access
  - "Active" processing, to take internal actions on data
  - Presentation or Visualization of data

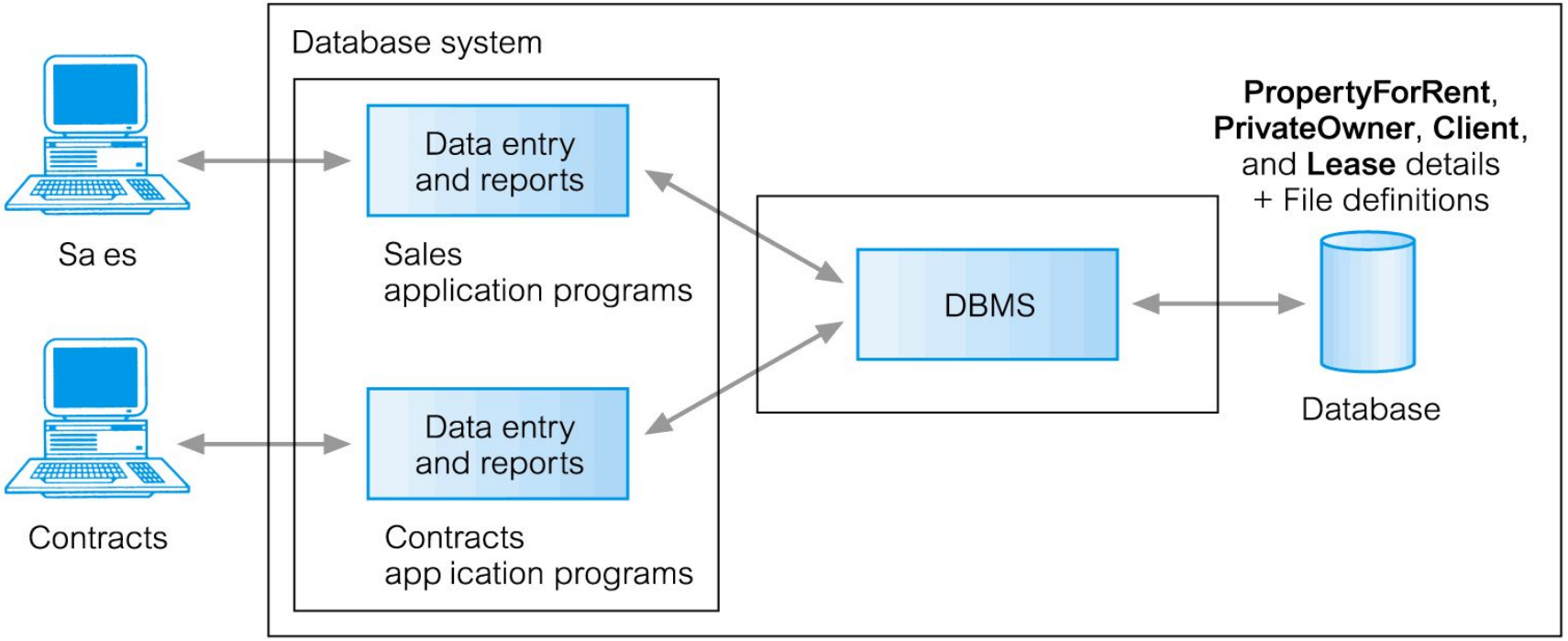
BENEFITS OF DBMS

- Data Independence
  - Application programs should be as independent as possible from details of data representation and storage
- Efficient data access
  - Use if sophisticated techniques to store and retrieve data efficiently
- Reduces application development time
- Data integrity and security
  - Enforce integrity constraints
- Data centralization and administration
  - When several users share data, centralizing the administration of data is important
- Concurrent access and crash recovery

BENEFITS OF DBMS

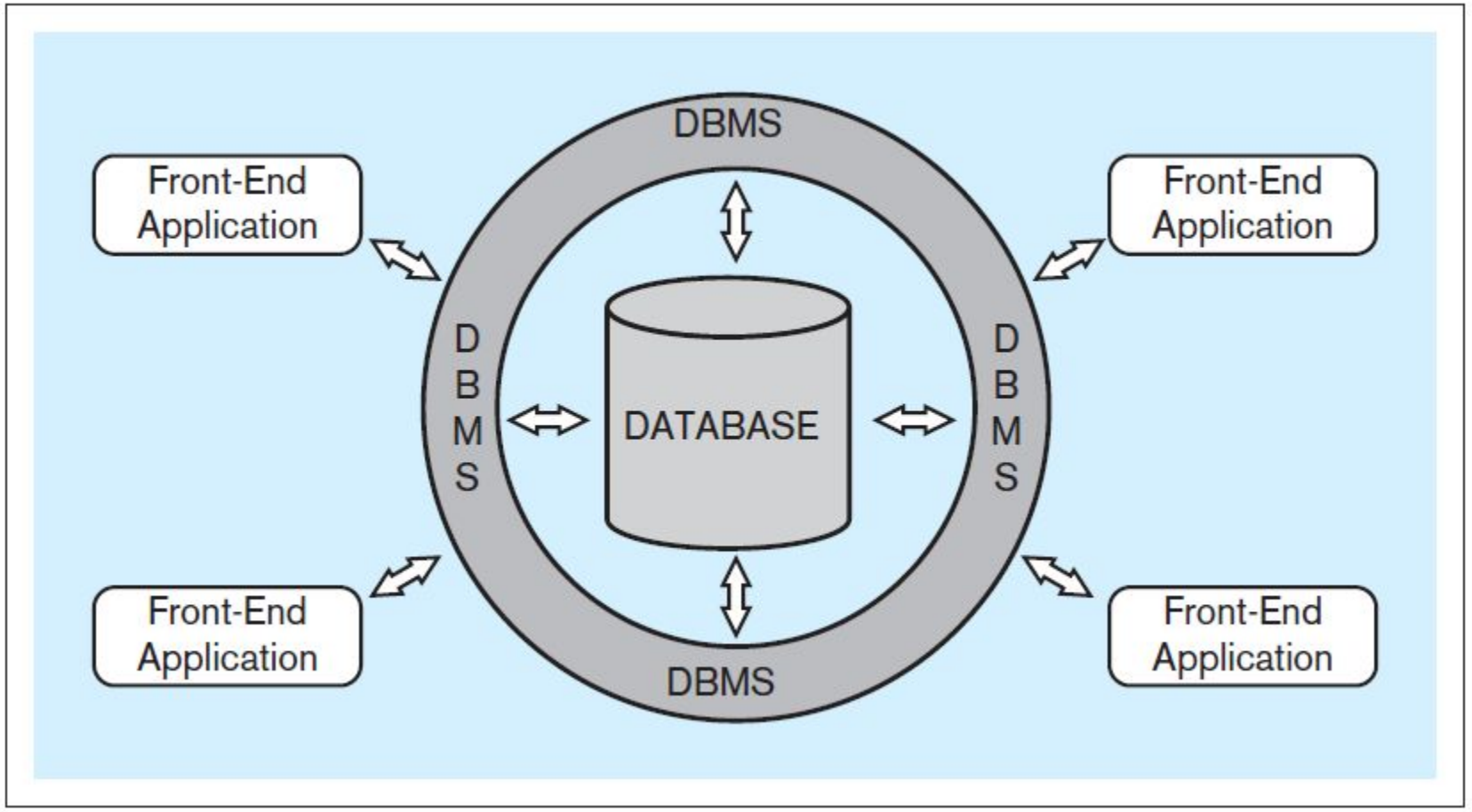
- Control of data redundancy
- Data consistency
- More information from the same amount of data
- Sharing of data
- Improved data integrity
- Improved security
- Enforcement of standards
- Economy of scale
- Balance conflicting requirements
- Improved data accessibility and responsiveness
- Increased productivity
- Improved maintenance through data independence
- Increased concurrency
- Improved backup and recovery services

AN EXAMPLE OF A DATABASE MANAGEMENT SYSTEM



PropertyForRent (propertyNo, street, city, postcode, type, rooms, rent, ownerNo)
PrivateOwner (ownerNo, fName, lName, address, telNo)
Client (clientNo, fName, lName, address, telNo, prefType, maxRent)
Lease (leaseNo, propertyNo, clientNo, paymentMethod, deposit, paid, rentStart, rentF nish)

TYPICAL DATABASE SYSTEM ARCHITECTURE

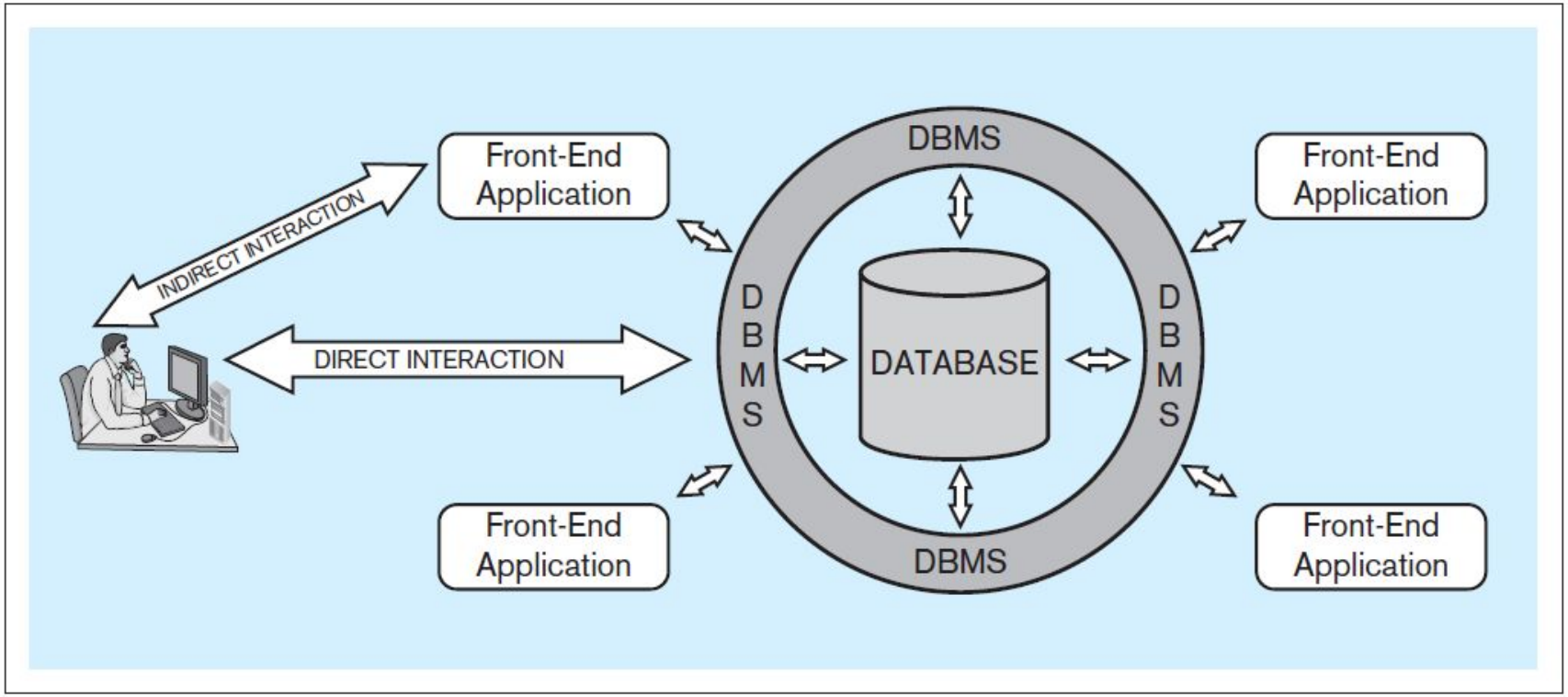


TYPICAL DATABASE SYSTEM ARCHITECTURE

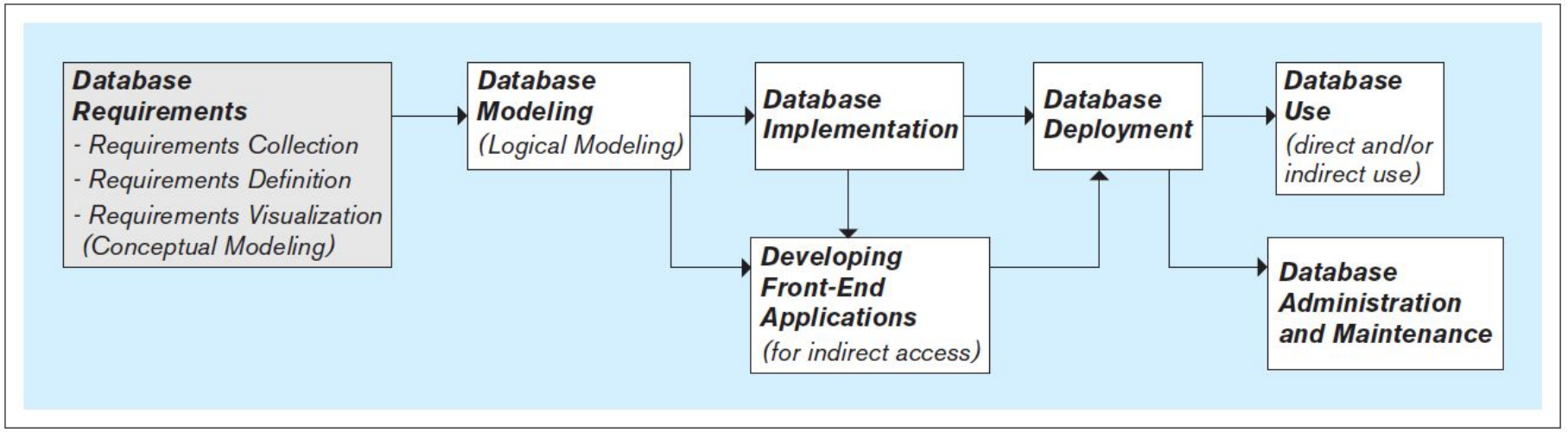
- **Front-end applications** - provide a mechanism for easy interaction between the users and the DBMS
- **End-users** (**business-users**) - users using a database system to support their tasks and processes
- **Indirect interaction** - end-user communicating with the database through front-end applications
- **Direct interaction** - end-user communicating with the database directly through DBMS

TYPICAL DATABASE SYSTEM ARCHITECTURE

Typical database system architecture



STEPS IN THE DEVELOPMENT OF DATABASE SYSTEMS



STEPS IN THE DEVELOPMENT OF DATABASE SYSTEMS

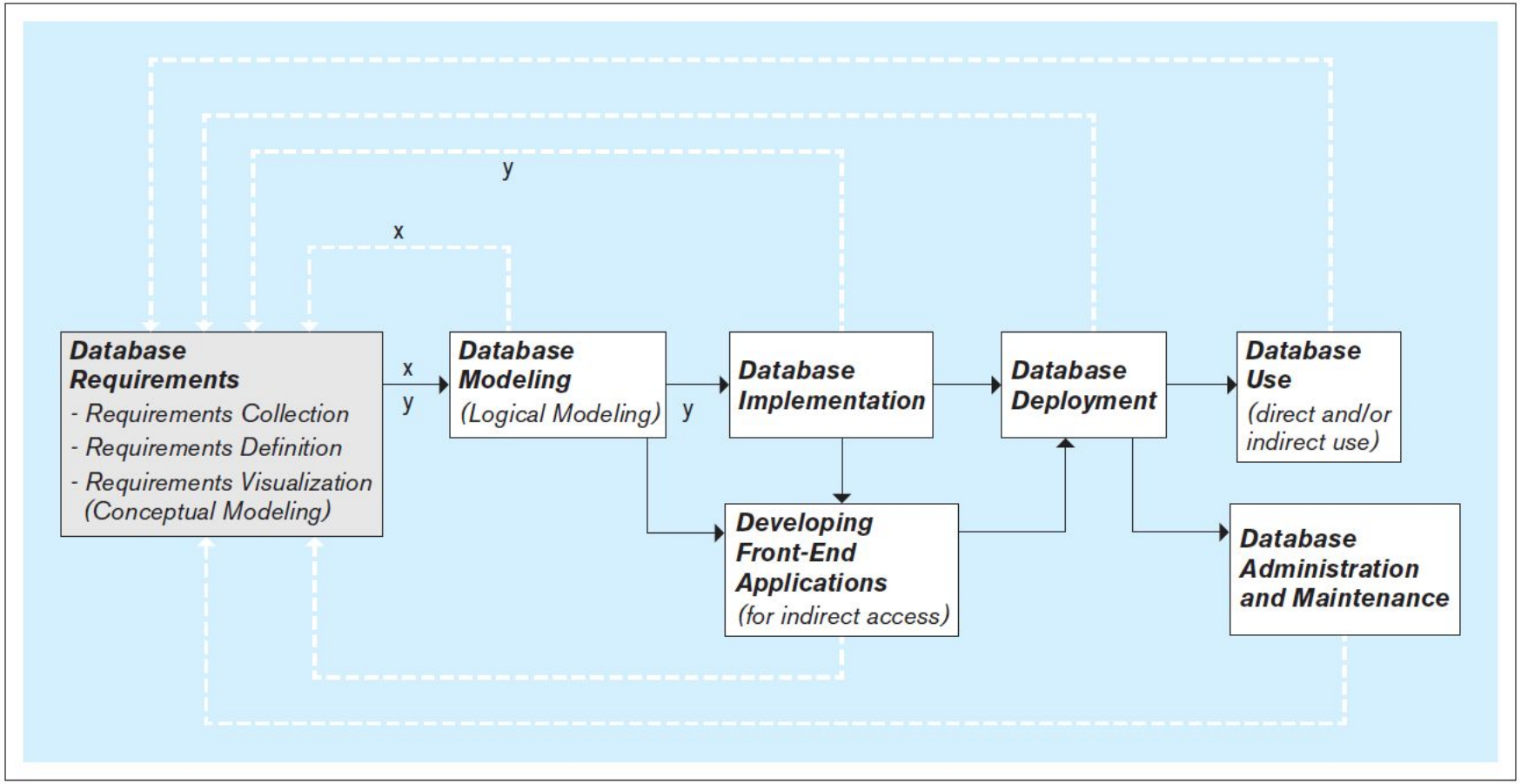
- **Requirements collection, definition, and visualization** - results in the requirements specifying which data the future database system will hold and in what fashion, and what the capabilities and functionalities of the database system will be
  - The **collected** requirements should be clearly **defined** and stated in a written document, and then **visualized**
  -
  - **The first and most critical step in the development of the database.**

STEPS IN THE DEVELOPMENT OF DATABASE SYSTEMS

- **Requirements collection, definition, and visualization**
  - **Conceptual database model** – a visualization of requirements by using a conceptual data modeling technique (such as entity-relationship [ER] modeling)

STEPS IN THE DEVELOPMENT OF DATABASE SYSTEMS

Iterative nature of the database requirements collection, definition, and visualization process



STEPS IN THE DEVELOPMENT OF DATABASE SYSTEMS

- **Database modeling** (**logical database modeling**) - creation of the database model that is implementable by the DBMS software
  - *Logical database modeling* follows *conceptual database modeling*
  - Conceptual model - created as a visualization of requirements during the requirements collection, definition, and visualization step and serves as a as a blueprint for the actual (logical) database model
  - Logical model - actual database model, created during the database modeling step to be used in the subsequent step of database implementation using the DBMS

STEPS IN THE DEVELOPMENT OF DATABASE SYSTEMS

- **Database implementation** - using a DBMS to implement the database model as an actual database
  - Most modern databases are implemented using a relational DBMS (RDBMS) software

- SQL is a language used by most relational DBMS software packages.
- SQL includes commands for creating, modifying and deleting database structures (these commands are used during database implementation).

STEPS IN THE DEVELOPMENT OF DATABASE SYSTEMS

- **Developing front-end applications** - designing and creating applications for indirect use by the end-users
  - Front-end applications are based on the database model and the requirements specifying the front-end functionalities
  - Front-end applications contain interfaces (such as forms and reports) accessible via a navigation mechanism (such as a menu)

STEPS IN THE DEVELOPMENT OF DATABASE SYSTEMS

- **Database deployment** - releasing the database system for use by the end users
- Typically, database deployment also involves populating the implemented database with the initial set of data.

STEPS IN THE DEVELOPMENT OF DATABASE SYSTEMS

- **Database use** - the insertion, modification, deletion and retrieval of the data in the database system

STEPS IN THE DEVELOPMENT OF DATABASE SYSTEMS

- **Database administration and maintenance -** performing activities that support the database end user, including dealing with technical issues, such as:
  - Providing security for the information contained in the database
  - Ensuring sufficient hard-drive space for the database content
  - Implementing the backup and recovery procedures

DATABASE SCOPE

- Databases can vary in their scope from small single-user (personal) databases to large enterprise databases that can be used by thousands of end-users
- Regardless of their scope, all databases go through the same fundamental development steps (*requirements*, *modeling*, *implementation*, *deployment*, *use*, etc.)

PEOPLE INVOLVED WITH DATABASE SYSTEMS

- **Database analysts, designers, and developers**
  - **Database analysts** - involved in the requirements collection, definition, and visualization stage
  - **Database designers** (a.k.a. **database modelers** or **architects**) - involved in the database modeling stage
  - **Database developers** – in charge of implementing the database model as a functioning database using the DBMS software

PEOPLE INVOLVED WITH DATABASE SYSTEMS

- **Front-end applications analysts and developers**
  - **Front-end application analysts** - in charge of collecting and defining requirements for front-end applications
  - **Front-end applications developers** - in charge of creating the front-end applications

PEOPLE INVOLVED WITH DATABASE SYSTEMS

- **Database administrators (DBAs) -** perform the tasks related to the maintenance and administration of a database system

PEOPLE INVOLVED WITH DATABASE SYSTEMS

- **Database end users -** use a database system to support their work- or life-related tasks and processes
  - Users differ in:
    - Level of technical sophistication
    - Amount of data that they need
    - Frequency with which they access the database system

DRAWBACK OF DBMSs

- Complexity
- Size
- Cost of DBMS
- Additional hardware costs
- Cost of conversion
- Performance
- Higher impact of a failure

MAJOR DBMS

- Oracle
- IBM DB2
- Microsoft SQL Server
- Sybase
- Informix
- PostgreSQL
- MySQL
- Microsoft Access