# INFORMATION SYSTEM DEPARTMENT INFORMATION TECHNOLOGY FACULTY- HCM UNIVERSITY SCIENCE

# INTRODUCTION TO DATABASE

# Chapter 01 Overview of Database System

Lecturer- PhD. NGUYEN TRAN MINH THU





- Why Database
- Data & Information
- Database
- Type of Databases
- Evolution of File System Data Processing
- Database Management System
- Information System
- Database Design
- Systems Development Life Cycle (SDLC)
- Database Life Cycle (DBLC)
- Database Career Opportunities



- Why Database
- Data & Information
- Database
- Type of Databases
- Evolution of File System Data Processing
- Database Management System
- Information System
- Database Design
- Systems Development Life Cycle (SDLC)
- Database Life Cycle (DBLC)
- Database Career Opportunities



#### A Day In Susan's Life

See how many databases she interacts with each day

Before leaving for work, Susan checks her Facebook and Twitter accounts On her lunch break, she picks up her prescription at the pharmacy After work, Susan goes to the grocery store At night, she plans for a trip and buys airline tickets and hotel reservations online Then she makes a few online purchases











Where is the data about the friends and groups stored?
Where are the "likes" stored

and what would they be

used for?

Where is the pharmacy inventory data stored?

What data about each product will be in the inventory data?

What data is kept about each customer and where is it stored?

Where is the product data stored?

Is the product quantity in stock updated at checkout?

Does she pay with a credit card?

Where does the online travel website get the airline and hotel data from?

What customer data would be kept by the website?

Where would the customer data be stored?

Where are the product and stock data stored?

Where does the system get the data to generate product "recommendations" to the customer?

Where would credit card information be stored?



#### A Day In Susan's Life

See how many databases she interacts with each day

Before leaving for work, Susan checks her Facebook and Twitter accounts On her lunch break, she picks up her prescription at the pharmacy After work, Susan goes to the grocery store At night, she plans for a trip and buys airline tickets and hotel reservations online

Then she makes a few online purchases











Where is the data about the friends and groups stored?

Where are the "likes" stored and what would they be used for?

Where is the pharmacy inventory data stored?

What data about each product will be in the inventory data?

What data is kept about each customer and where is it stored?

Where is the product data stored?

Is the product quantity in stock updated at checkout?

Does she pay with a credit card?

Where does the online travel website get the airline and hotel data from?

What customer data would be kept by the website?

Where would the customer data be stored?

Where are the product and stock data stored?

Where does the system get the data to generate product "recommendations" to the customer?

Where would credit card information be stored?



Products
Sales
Customers









- Characteristics of data in today's world
  - Ubiquitous (i.e., abundant, global, and everywhere)
  - Pervasive (i.e., unescapable, prevalent, and persistent)
- Databases make data persistent and shareable in a secure way
  - Specialized structures that allow computer-based systems to store, manage, and retrieve data very quickly



- Why Database
- Data & Information
- Database
- Type of Databases
- Evolution of File System Data Processing
- Database Management System
- Information System
- Database Design
- Systems Development Life Cycle (SDLC)
- Database Life Cycle (DBLC)
- Database Career Opportunities



### **Data versus Information**

Baker, Kenneth D.	324917628
Doyle, Joan E.	476193248
Finkle, Clive R.	548429344
Lewis, John C.	551742186
McFerran, Debra R.	409723145
	/ V

(a)

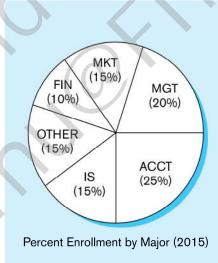
-	Lancas and a second	<b>n</b>
-	200	Roster

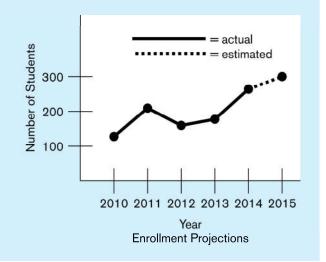
Course: MGT 500 Semester: Spring 2015

**Business Policy** 

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





(b)

**(c)** 



#### Data

#### Raw facts

- Have not yet been processed to reveal their meaning to the end user
- Building blocks of information

### Data management

 Generation, storage, and retrieval of data

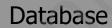
#### Information

- Produced by processing raw data to reveal its meaning
- Requires context
- Bedrock of knowledge
- Should be accurate, relevant, and timely to enable good decision making



- Why Database
- Data & Information
- Database
- Type of Databases
- Evolution of File System Data Processing
- Database Management System
- Information System
- Database Design
- Systems Development Life Cycle (SDLC)
- Database Life Cycle (DBLC)
- Database Career Opportunities





Contains employee, order, inventory, pricing, and customer data

Database: organized collection of logically related data



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

Central Database

Contains employee, order, inventory, pricing, and customer data Requires a Database Management System (DBMS)



- Why Database
- Data & Information
- Database
- Type of Databases
- Evolution of File System Data Processing
- Database Management System
- Information System
- Database Design
- Systems Development Life Cycle (SDLC)
- Database Life Cycle (DBLC)
- Database Career Opportunities

# Types of Databases

- Single-user database: Supports one user at a time
  - Desktop database: Runs on PC
- Multiuser database: Supports multiple users at the same time
  - Workgroup databases: Supports a small number of users or a specific department
  - Enterprise database: Supports many users across many departments



### Classification by location

- Centralized database: Data is located at a single site
- Distributed database: Data is distributed across different sites
- Cloud database: Created and maintained using cloud data services that provide defined performance measures for the database



### Classification by data type

- General-purpose databases: Contains a wide variety of data used in multiple disciplines
- Discipline-specific databases: Contains data focused on specific subject areas
- Operational database: Designed to support a company's day-to-day operations



- Analytical database: Stores historical data and business metrics used exclusively for tactical or strategic decision making
  - Data warehouse: Stores data in a format optimized for decision support
  - Online analytical processing (OLAP)
    - Tools for retrieving, processing, and modeling data from the data warehouse
- **Business intelligence**: Captures and processes business data to generate information that support decision making



- Databases can be classified to reflect the degree to which the data is structured:
  - Unstructured data: It exists in their original state
  - Structured data: It results from formatting
    - Structure is applied based on type of processing to be performed
  - **Semistructured data**: Processed to some extent
- Extensible Markup Language (XML)
  - Represents data elements in textual format



TYPES OF DATABASES								
PRODUCT	NUMBER OF USERS			DATA LOCATION		DATA USAGE		XML
	SINGLE MULTIUSE		USER					
	USER	WORKGROUP	ENTERPRISE	CENTRALIZED	DISTRIBUTED	OPERATIONAL	ANALYTICAL	
MS Access	X	X		X		X		
MS SQL Server	X <sup>3</sup>	X	X	X	X	X	X	Х
IBM DB2	X <sup>3</sup>	X	X	X	X	X	X	X
MySQL	Х	X	X	X	X	X	X	Х
Oracle RDBMS	X <sup>3</sup>	X	X	X	X	Х	X	Х



- Why Database
- Data & Information
- Database
- Type of Databases
- Evolution of File System Data Processing
- Database Management System
- Information System
- Database Design
- Systems Development Life Cycle (SDLC)
- Database Life Cycle (DBLC)
- Database Career Opportunities



**Duplicate Data** Orders Department Accounting Department Payroll Department Program A Program B Program A Program B Program C Program A Program B Order Filling Invoicing Payroll System System System Customer Employee Customer Inventory Back Inventory Order Pricing Master Master Master Master File File File File File File

File Systems Processing



Lengthy development times

Difficulty of getting quick answers

Complex system administration

Lack of security and limited data sharing

Extensive programming

**Problems with File System Data Processing** 



Table 1.2	Basic File Terminology
TERM	DEFINITION
Data	Raw facts, such as a telephone number, a birth date, a customer name, and a year-to-date (YTD) sales value. Data has little meaning unless it has been organized in some logical manner.
Field	A character or group of characters (alphabetic or numeric) that has a specific meaning. A field is used to define and store data.
Record	A logically connected set of one or more fields that describes a person, place, or thing. For example, the fields that constitute a record for a customer might consist of the customer's name, address, phone number, date of birth, credit limit, and unpaid balance.
File	A collection of related records. For example, a file might contain data about the students currently enrolled at Gigantic University.

#### FIGURE 1.8 CONTENTS OF THE AGENT FILE

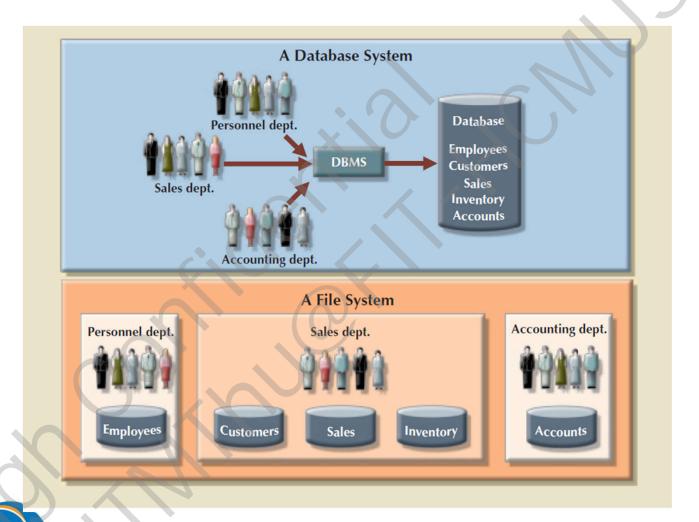
Database name: Ch01\_Text

A_NAME	A_PHONE	A_ADDRESS	ZIP	HIRED	YTD_PAY	YTD_FIT	YTD_FICA	YTD_SLS	DEP
Alex B. Alby	713-228-1249	123 Toll, Nash, TN	37119	01-Nov-2000	26566.24	6641.56	2125.30	132737.75	3
Leah F. Hahn	615-882-1244	334 Main, Fox, KY	25246	23-May-1986	32213.78	8053.44	2577.10	138967.35	0
John T. Okon	615-123-5589	452 Elm, New, TN	36155	15-Jun-2005	23198.29	5799.57	1855.86	127093.45	2

A\_NAME = Agent name YTD\_PAY = Year-to-date pay A\_PHONE = Agent phone YTD\_FIT = Year-to-date federal income tax paid A\_ADDRESS = Agent address YTD\_FICA = Year-to-date Social Security taxes paid

YTD\_SLS = Year-to-date sales ZIP = Agent zip code HIRED = Agent date of hire DEP = Number of dependents





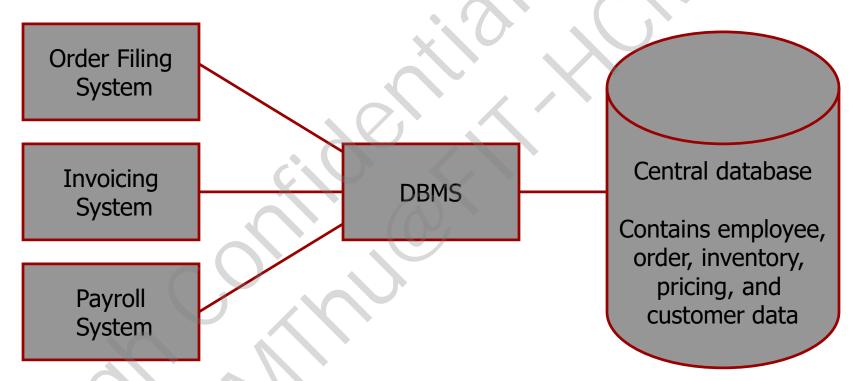


- Why Database
- Data & Information
- Database
- Type of Databases
- Evolution of File System Data Processing
- Database Management System
- Information System
- Database Design
- Systems Development Life Cycle (SDLC)
- Database Life Cycle (DBLC)
- Database Career Opportunities



## **Database Management System**

A software system that is used to create, maintain, and provide controlled access to user databases

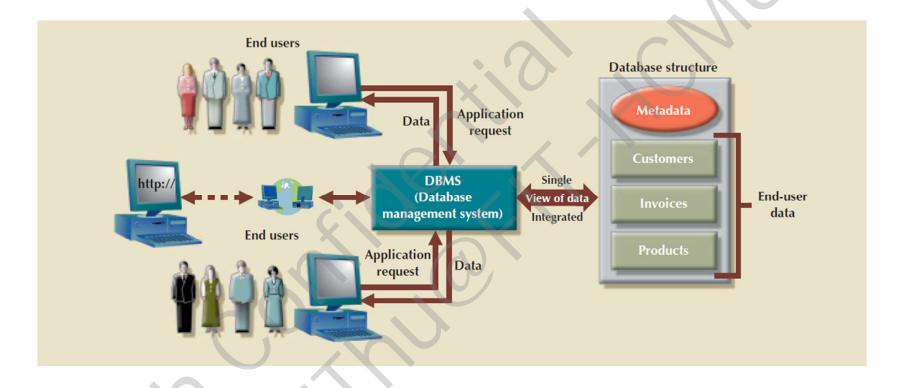


DBMS manages data resources like an operating system manages hardware resources



- Intermediary between the user and the database
- Enables data to be shared
- Presents the end user with an integrated view of the data
- Receives and translates application requests into operations required to fulfill the requests
- Hides database's internal complexity from the application programs and users





The DBMS Manages the Interaction between the End User and the Database

# Metadata Metadata

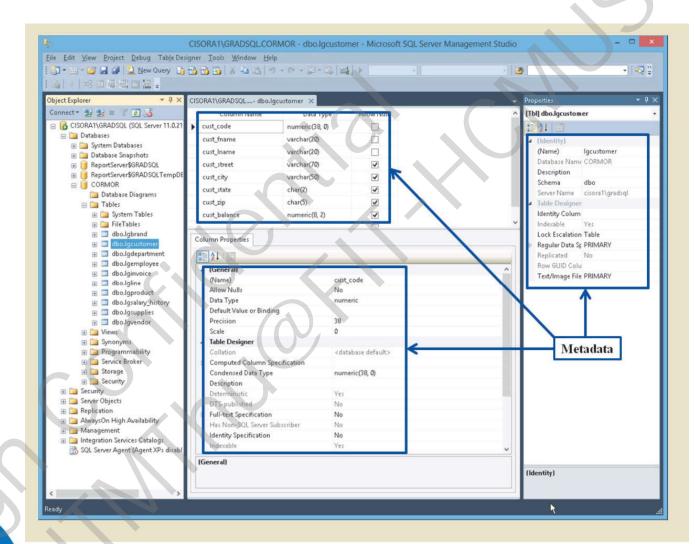
Metadata are data that describe the properties or characteristics of end user data and the context of that data

TABLE 1-1	Example Metadata fo	or Class Roster	· ·	0 10		
Data Item Metadata						
Name	Туре	Length I	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	

Data dictionary: Stores definitions of the data elements and their relationships

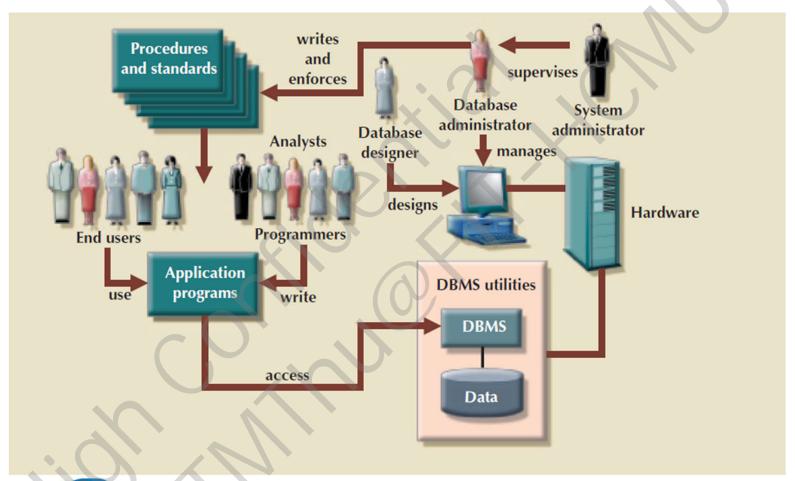


### Metadata





### **The Database System Environment**





### Data dictionary management

• **Data dictionary**: Stores definitions of the data elements and their relationships

#### Data storage management

• **Performance tuning**: Ensures efficient performance of the database in terms of storage and access speed

### Data transformation and presentation

• Transforms entered data to conform to required data structures

### Security management

• Enforces user security and data privacy



### Multiuser access control

• Sophisticated algorithms ensure that multiple users can access the database concurrently without compromising its integrity

### Backup and recovery management

• Enables recovery of the database after a failure

### Data integrity management

• Minimizes redundancy and maximizes consistency



### Database access languages and application programming interfaces

- Query language: Lets the user specify what must be done without having to specify how
- Structured Query Language (SQL): De facto query language and data access standard supported by the majority of DBMS vendors

#### Database communication interfaces

• Accept end-user requests via multiple, different network environments



## **Disadvantages of DBMS**

Increased costs

Management complexity

Maintaining currency

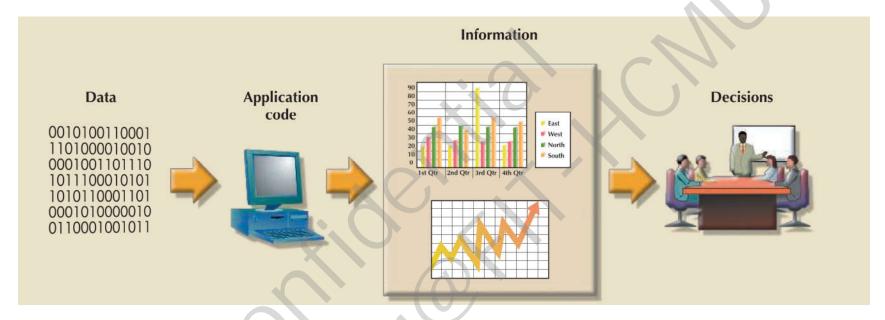
Vendor dependence

Frequent upgrade/replacement cycles



- Why Database
- Data & Information
- Database
- Type of Databases
- Evolution of File System Data Processing
- Database Management System
- Information System
- Database Design
- Systems Development Life Cycle (SDLC)
- Database Life Cycle (DBLC)
- Database Career Opportunities





- Provides for data collection, storage, and retrieval
- Composed of:
  - People, hardware, software
  - Database(s), application programs, procedures

# **Information System**

- Systems analysis: Process that establishes need for and extent of information system
- Systems development: Process of creating information system
- The performance of an information system depends on three factors:
  - Database Design & Implementation
  - Application Design & Implementation
  - Administrative procedures
- Database development: Process of database design and its implementation



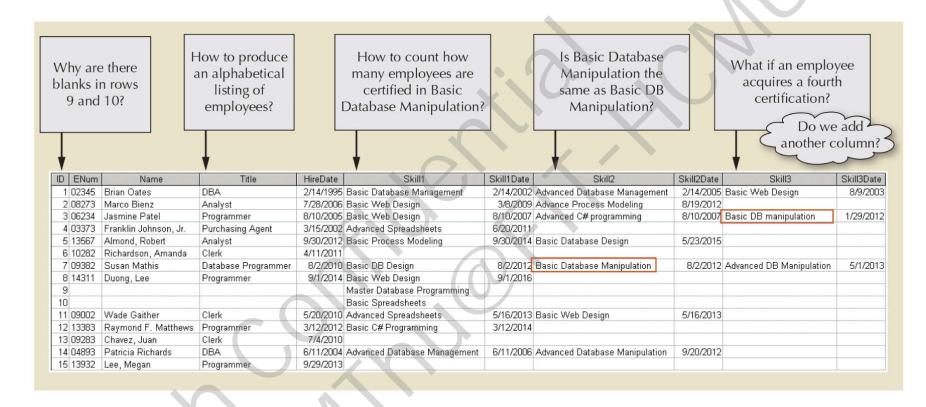
- Why Database
- Data & Information
- Database
- Type of Databases
- Evolution of File System Data Processing
- Database Management System
- Information System
- Database Design
- Systems Development Life Cycle (SDLC)
- Database Life Cycle (DBLC)
- Database Career Opportunities



- Focuses on the design of the database structure that will be used to store and manage end-user data
- Well-designed database
  - Facilitates data management
  - Generates accurate and valuable information
- Poorly designed database causes difficult-to-trace errors



### Why Database Design is Important?



#### **EMPLOYEE SKILLS CERTIFICATION IN A POOR DESIGN**



## Why Database Design is Important?

#### Table name: EMPLOYEE

Employee_ID	Employee_FName	Employee_LName	Employee_HireDate	Employee_Title
02345	Johnny	Jones	2/14/1995	DBA
03373	Franklin	Johnson	3/15/2002	Purchasing Agent
04893	Patricia	Richards	6/11/2004	DBA
06234	Jasmine	Patel	8/10/2005	Programmer
08273	Marco	Bienz	7/28/2006	Analyst
09002	Ben	Joiner	5/20/2010	Clerk
09283	Juan	Chavez	7/4/2010	Clerk
09382	Jessica	Johnson	8/2/2010	Database Programmer
10282	Amanda	Richardson	4/11/2011	Clerk
13383	Raymond	Matthews	3/12/2012	Programmer
13567	Robert	Almond	9/30/2012	Analyst
13932	Megan	Lee	9/29/2013	Programmer
14311	Lee	Duong	9/1/2014	Programmer

#### Table name: CERTIFIED

Employee_ID	Skill_ID	Certified_Date	
02345	100	2/14/2002	
02345	110	8/9/2003	
02345	180	2/14/2005	
03373	120	6/20/2011	
04893	180	6/11/2008	
04893	220	9/20/2012	
06234	110	8/10/2007	
06234	200	8/10/2007	
06234	210	1/29/2012	
08273	110	3/8/2009	
08273	190	8/19/2012	
09002	110	5/16/2013	
09002	120	5/16/2013	
09382	140	8/2/2012	
09382	210	8/2/2012	
09382	220	5/1/2013	
13383	170	3/12/2014	
13567	130	9/30/2014	
13567	140	5/23/2015	
14311	110	9/1/2018	

#### Table name: SKILL

Skill_ID	Skill_Name	Skill_Description		
100	Basic Database Management	Create and manage database user accounts.		
110	Basic Web Design	Create and maintain HTML and CSS documents.		
120	Advanced Spreadsheets	Use of advanced functions, user-defined functions, and macroing.		
130 Basic Process Modeling Create core business process models using standard libr		Create core business process models using standard libraries.		
140	Basic Database Design	Create simple data models.		
150	Master Database Programming	se Programming Create integrated trigger and procedure packages for a distributed environment.		
160	Basic Spreadsheets	Create single tab worksheets with basic formulas		
170	Basic C# Programming	Create single-tier data aware modules.		
180	Advanced Database Management	Manage Database Server Clusters.		
190	Advance Process Modeling	Evaluate and Redesign cross-functional internal and external business processes.		
200	Advanced C# Programming	ed C# Programming Create multi-tier applications using multi-threading		
210	Basic Database Manipulation	Create simple data retrieval and manipulation statements in SQL.		
220	Advanced Database Manipulation	Use of advanced data manipulation methods for multi-table inserts, set operations, and correlated subqueries.		

#### **EMPLOYEE SKILLS CERTIFICATION IN A GOOD DESIGN**



#### Existing Database

- Old data: worksheet, files.
- Extracting data from another databases
- Design using Normal Form

#### Design a new database

- Database analysis based on user requirements.
- Database design for data extracting.

#### Re-design Database

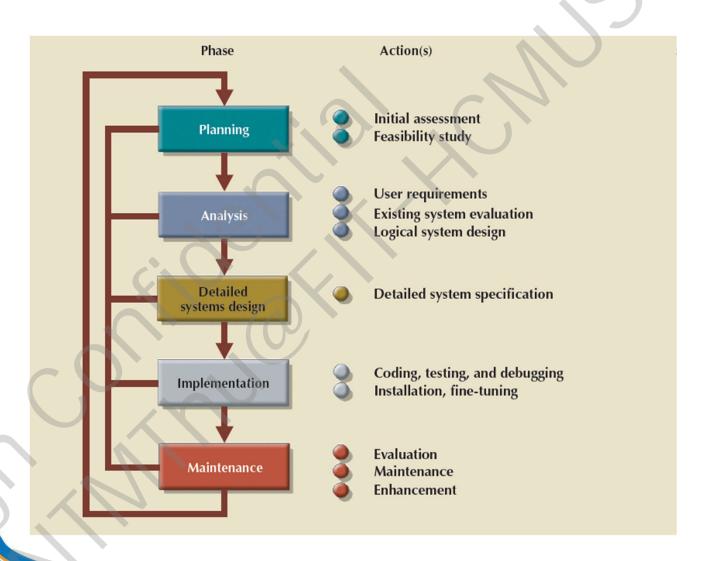
- Transform existing database to a new one.
- Integrated two or more databased together.
- Giving techniques & design methods for databases.
- Applying normal form principle and mapping rules between models



- Why Database
- Data & Information
- Database
- Type of Databases
- Evolution of File System Data Processing
- Database Management System
- Information System
- Database Design
- Systems Development Life Cycle (SDLC)
- Database Life Cycle (DBLC)
- Database Career Opportunities



## Systems Development Life Cycle (SDLC)





### **Systems Development Life Cycle (SDLC)**

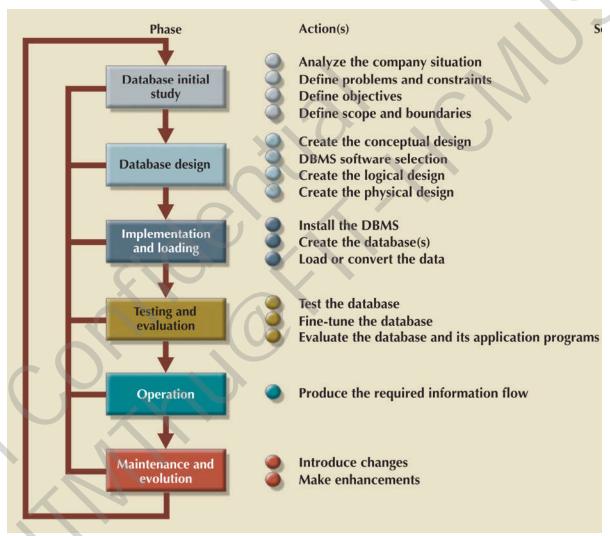
- Traces history of an information system
- Provides a picture within which database design and application development are mapped out and evaluated
- Iterative rather than sequential process



- Why Database
- Data & Information
- Database
- Type of Databases
- Evolution of File System Data Processing
- Database Management System
- Information System
- Database Design
- Systems Development Life Cycle (SDLC)
- Database Life Cycle (DBLC)
- Database Career Opportunities



### The Database Life Cycle (DBLC)





### The Database Life Cycle (DBLC)

Analyze company situation

Define problems and constraints

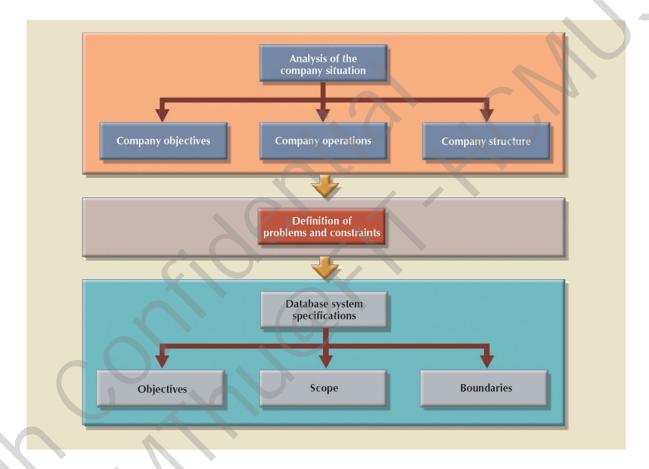
Define objectives

Define scope and boundaries

**Purpose of Database Initial Study** 



## Purpose of Database Initial Study



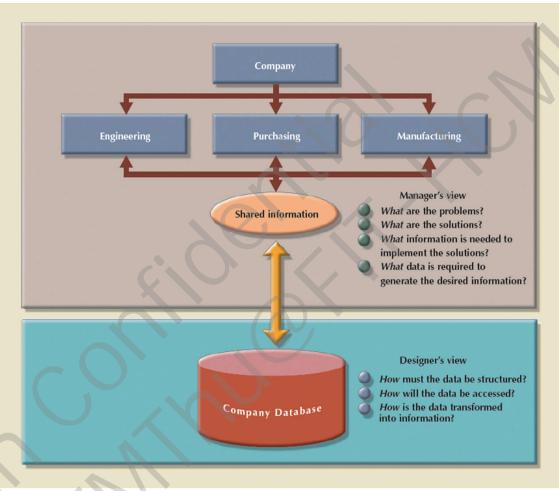
A SUMMARY OF ACTIVITIES IN THE DATABASE INITIAL STUDY



- Supports company's operations and objectives
- Most critical phase
  - Ensures final product meets user and system requirements
- Points for examining completion procedures
  - Data component is an element of whole system
  - System analysts/programmers design procedures to convert data into information
  - Database design is an iterative process



## 14.0 Database Design



Two Views of Data: Business Manager and **Database Designer** 

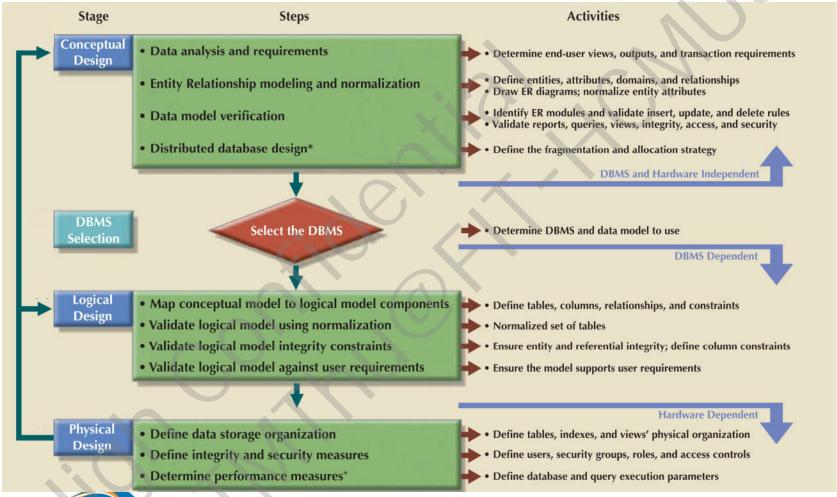


- Install the DBMS
  - Virtualization: Creates logical representations of computing resources independent of underlying physical computing resources
- Create the databases
  - Requires the creation of special storage-related constructs to house the end-user tables
- Load or convert the data
  - Requires aggregating data from multiple sources

**Implementation and Loading** 



### **Database Design Process**



©2020 Information System Department



- Physical security
- Password security
- Access rights
- Audit trails
- Data encryption
- Diskless workstations
- Optimization

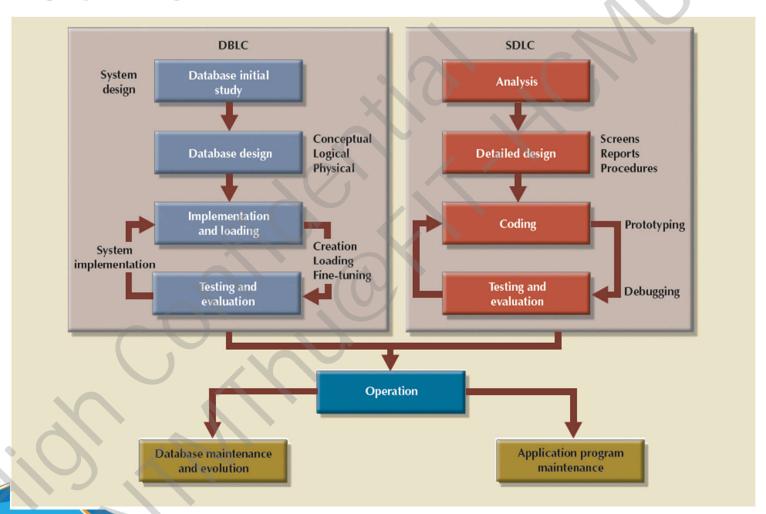
**Testing and Evaluation** 



- Preventive maintenance (backup)
- Corrective maintenance (recovery)
- Adaptive maintenance
- Assignment of access permissions and their maintenance for new and old users
- Generation of database access statistics
- Periodic security audits
- Periodic system-usage summaries

### **Maintenance and Evolution**

# Parallel Activities in the DBLC and the SDLC





- Why Database
- Data & Information
- Database
- Type of Databases
- Evolution of File System Data Processing
- Database Management System
- Information System
- Database Design
- Systems Development Life Cycle (SDLC)
- Database Life Cycle (DBLC)
- Database Career Opportunities



### **Database Career Opportunities**

#### **DATABASE CAREER OPPORTUNITIES**

JOB TITLE	DESCRIPTION	SAMPLE SKILLS REQUIRED
Database Developer	Create and maintain database-based applications	Programming, database fundamentals, SQL
Database Designer	Design and maintain databases	Systems design, database design, SQL
Database Administrator	Manage and maintain DBMS and databases	Database fundamentals, SQL, vendor courses
Database Analyst	Develop databases for decision support reporting	SQL, query optimization, data warehouses
Database Architect	Design and implementation of database environments (conceptual, logical, and physical)	DBMS fundamentals, data modeling, SQL, hardware knowledge, etc.
Database Consultant	Help companies leverage database technologies to improve business processes and achieve specific goals	Database fundamentals, data modeling, database design, SQL, DBMS, hardware, vendor-specific technologies, etc.
Database Security Officer	Implement security policies for data administration	DBMS fundamentals, database administration, SQL, data security technologies, etc.
Cloud Computing Data Architect	Design and implement the infrastructure for next-generation cloud database systems	Internet technologies, cloud storage technologies, data security, performance tuning, large databases, etc.



