

# BÀI GIẢNG CƠ SỞ DỮ LIỆU



## Chương 1 TỔNG QUAN CƠ SỞ DỮ LIỆU

Giáo viên: TS. Nguyễn Thị Uyên Nhi  
Email: nhintu@due.edu.vn

TIN HỌC QUẢN LÝ

# MỤC TIÊU

2

- Thế nào là dữ liệu – thông tin?
- Các dạng của dữ liệu
- Cơ sở dữ liệu và quản lý dữ liệu
- Tiêu chuẩn của CSDL là gì?
- Đặc tính của CSDL
- Hệ quản trị CSDL: lợi ích và kiến trúc
- Người dùng CSDL
- Ngôn ngữ CSDL
- Vòng đời và phát triển CSDL

# Dữ liệu

3

- Dữ liệu là các giá trị phản ánh về sự vật, hiện tượng trong thế giới khách quan.
- Dữ liệu là các giá trị thô, chưa có ý nghĩa với người sử dụng. Có thể là một tập hợp các giá trị mà không biết được sự liên hệ giữa chúng
  - Ví dụ: *Nguyễn Văn Nam, 845102, 14 / 10 / 02, 18*
- Dữ liệu có thể biểu diễn dưới nhiều dạng khác nhau (*âm thanh, văn bản, hình ảnh, v.v...*)
- 2 loại dữ liệu:
  - Tĩnh: Tên, ngày sinh, địa chỉ...
  - Động: Doanh thu, số lượng kho...

# Dữ liệu

4

- Dữ liệu về đối tượng có thể khác nhau, tùy thuộc vào ngữ cảnh.
- Ví dụ: dữ liệu về sinh viên có thể khác nhau tùy thuộc vào mục đích quản lý
  - Quản lý điểm: tên, mã sinh viên, điểm môn A, điểm môn B,...
  - Quản lý nhân thân: tên, địa chỉ, ngày sinh, quê quán,...

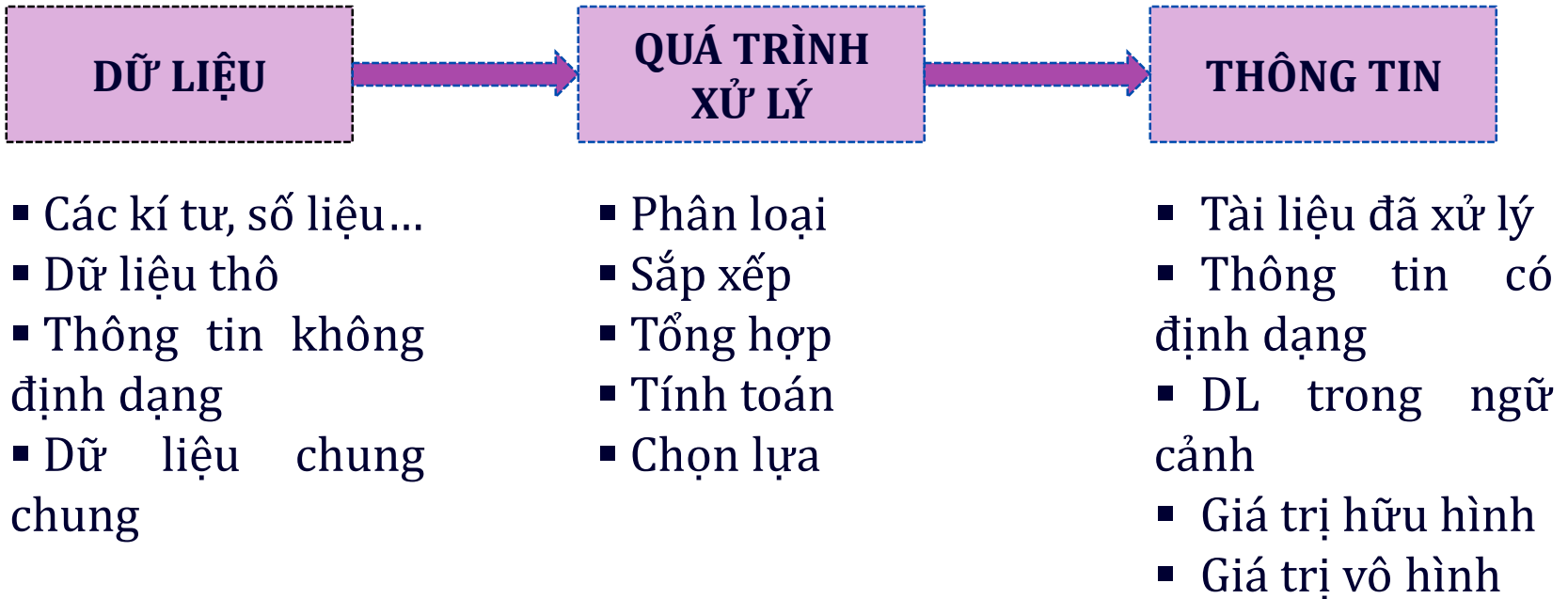
# Thông tin

5

- Thông tin là **ý nghĩa** được rút ra từ dữ liệu thông qua quá trình **xử lý** (phân tích, tổng hợp, v.v.), phù hợp với **mục đích cụ thể** của người sử dụng.
- Thông tin có thể gồm nhiều giá trị dữ liệu được tổ chức sao cho nó mang lại một **ý nghĩa cho một đối tượng cụ thể, trong một ngữ cảnh cụ thể**.
  - Ví dụ với dữ liệu trên có thông tin như sau: Thủ kho Nguyễn Văn Nam xuất mặt hàng có danh mục là 845102 vào ngày 14/10/ 02 với số lượng 18.

# Quá trình tạo thông tin

6



- Thông tin = Dữ liệu + Xử lý
- Xử lý thông tin ~ Xử lý dữ liệu

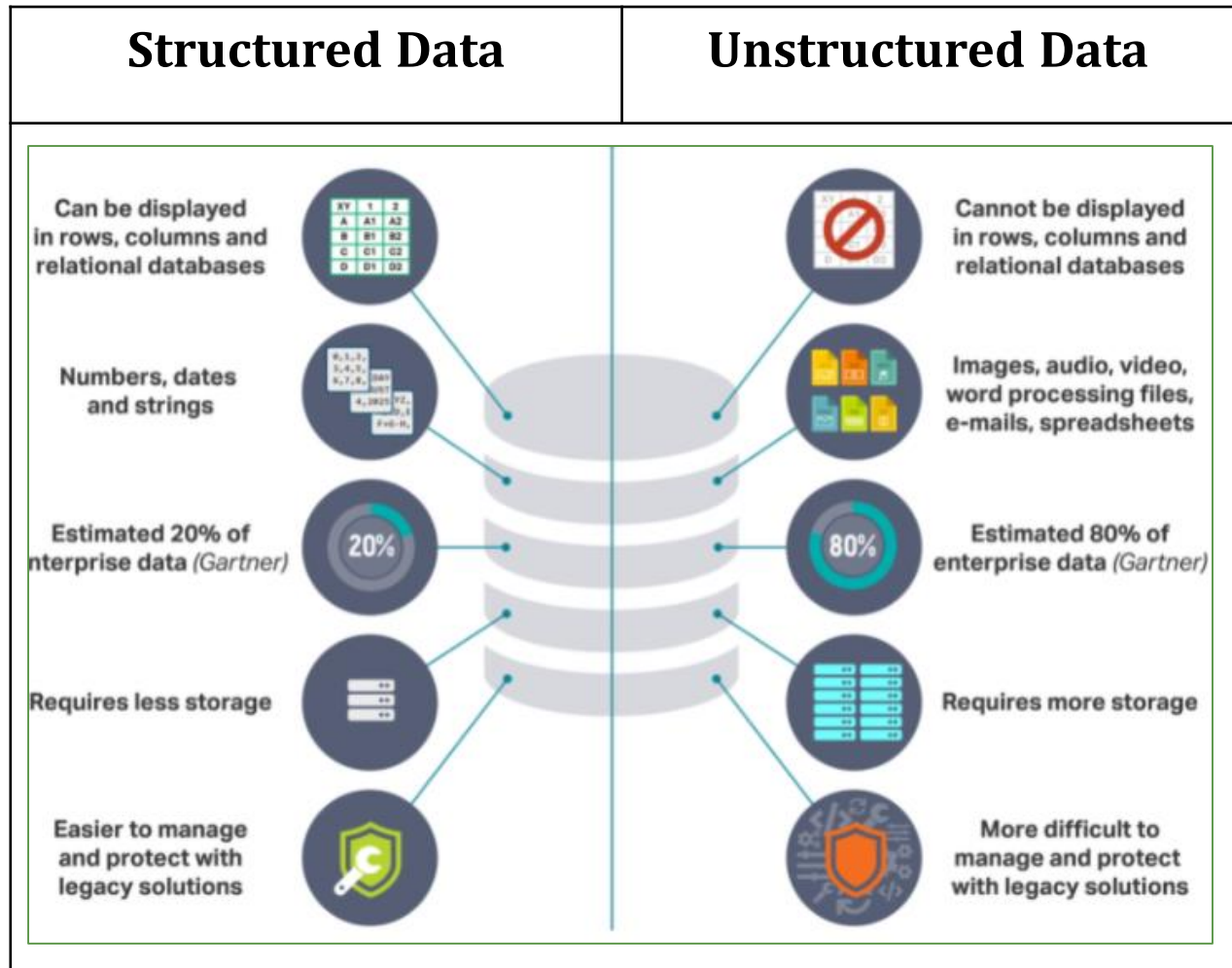
# Các dạng của dữ liệu

7

## ➤ Có cấu trúc



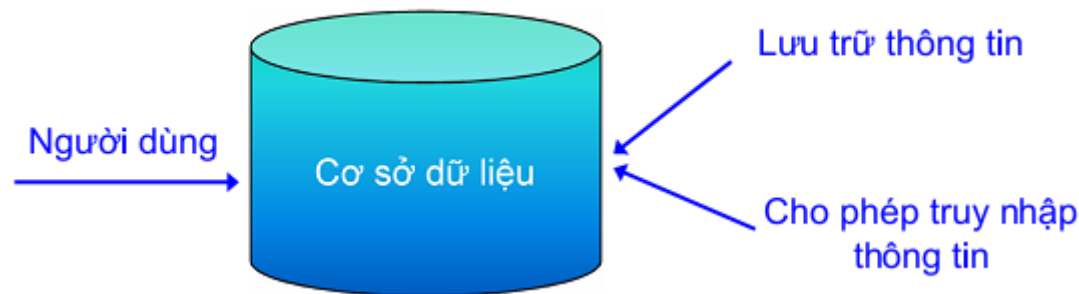
## ➤ Phi cấu trúc



# Cơ sở dữ liệu

8

- ❖ **Định nghĩa:** **CSDL (database)** = Tập hợp dữ liệu được tổ chức có cấu trúc liên quan với nhau và được lưu trữ trong máy tính.
- ❖ Được thiết kế, xây dựng cho phép người dùng lưu trữ dữ liệu, truy xuất thông tin hoặc cập nhật dữ liệu.
- ❖ **Mục đích:** Nhằm thoả mãn nhu cầu khai thác thông tin của một hay nhiều người sử dụng phục vụ cho nhiều mục đích khác nhau.





# Cơ sở dữ liệu

9

## ❖ CSDL được tổ chức có cấu trúc:

- Dữ liệu được lưu trữ thành các bản ghi (record), các trường (field)
- Dữ liệu có mối quan hệ với nhau

## ❖ Khả năng truy xuất thông tin từ CSDL:

- CSDL được cấu trúc để dễ dàng truy cập, quản lý và cập nhật dữ liệu

→ Cần phải quản trị CSDL

# Cơ sở dữ liệu

10

## ❖ Ví dụ CSDL:

SINHVIEN	TENSV	MASV	NAMSINH
	Hoang	1002013105	1995
	Van	1002013303	2004

MONHOC	TENMH	MAHP	TC
	Cơ sở dữ liệu	841107	4
	HTTT DN	841048	3

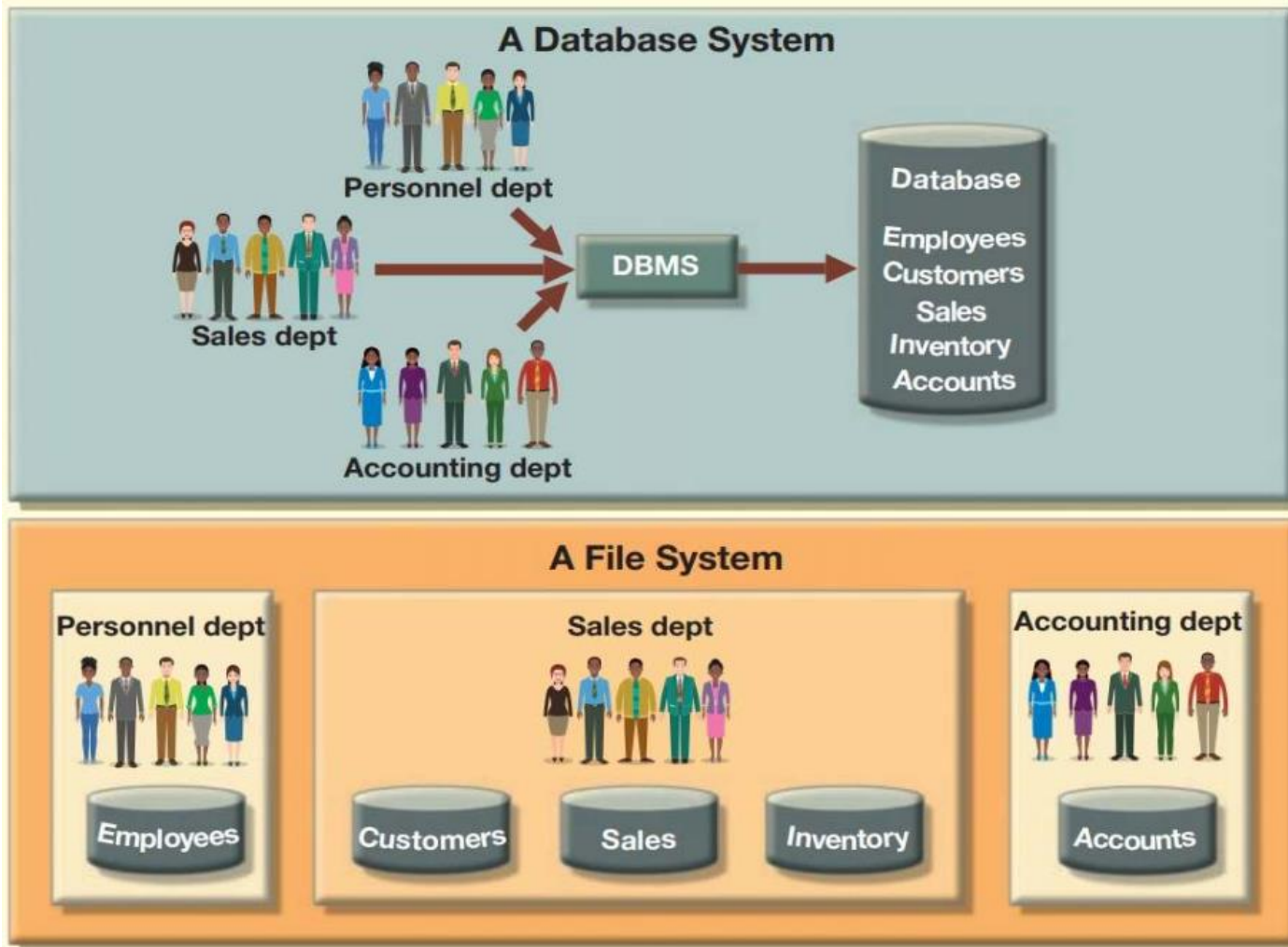
# Quản lý dữ liệu

11

- **Quản lý dữ liệu**: là quản lý số lượng lớn dữ liệu, bao gồm các việc lưu trữ và cung cấp cơ chế cho phép **thao tác** (thêm, sửa, xóa) và **truy vấn** dữ liệu.
- Các phương pháp quản lý dữ liệu:
  - ✦ Hệ thống quản lý bằng file
  - ✦ Hệ thống quản lý bằng cơ sở dữ liệu

# Quản lý dữ liệu

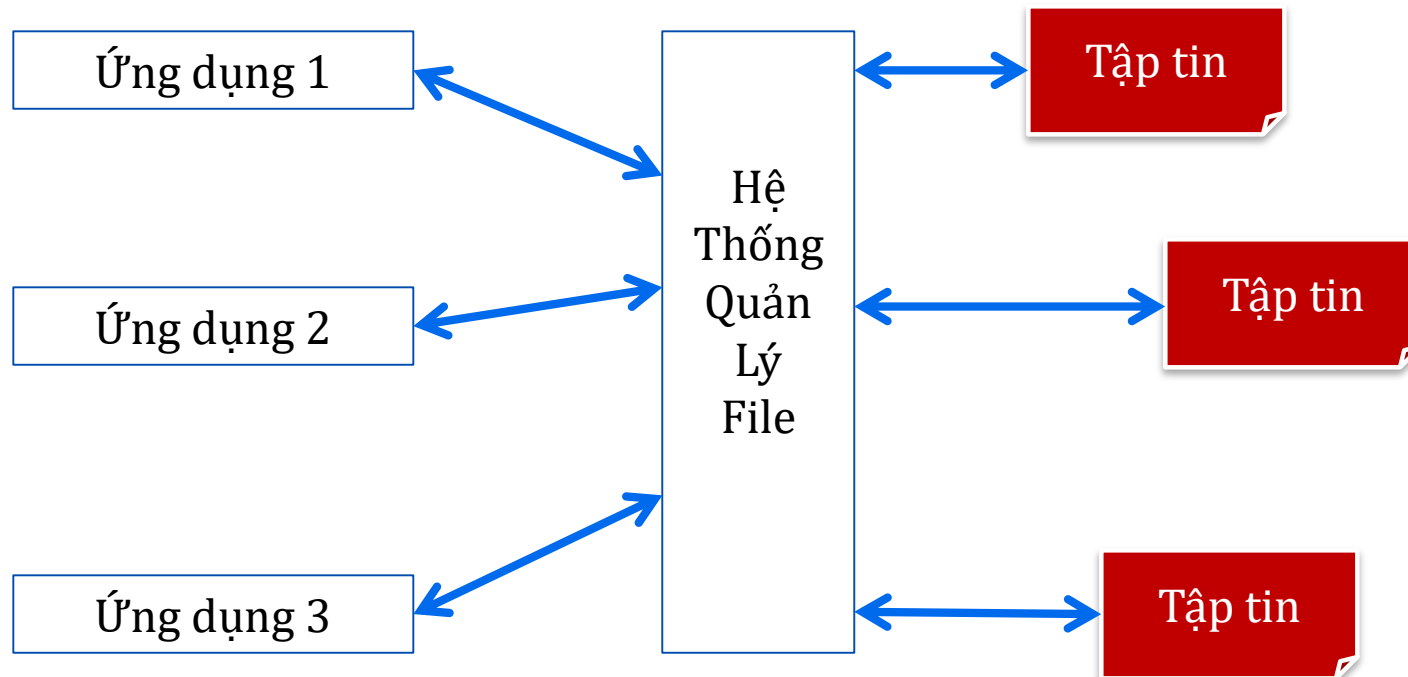
12



# Quản lý dữ liệu bằng file (tập tin)

13

- Dữ liệu lưu trữ trong những file riêng biệt



# Quản lý dữ liệu bằng file (tập tin)

14

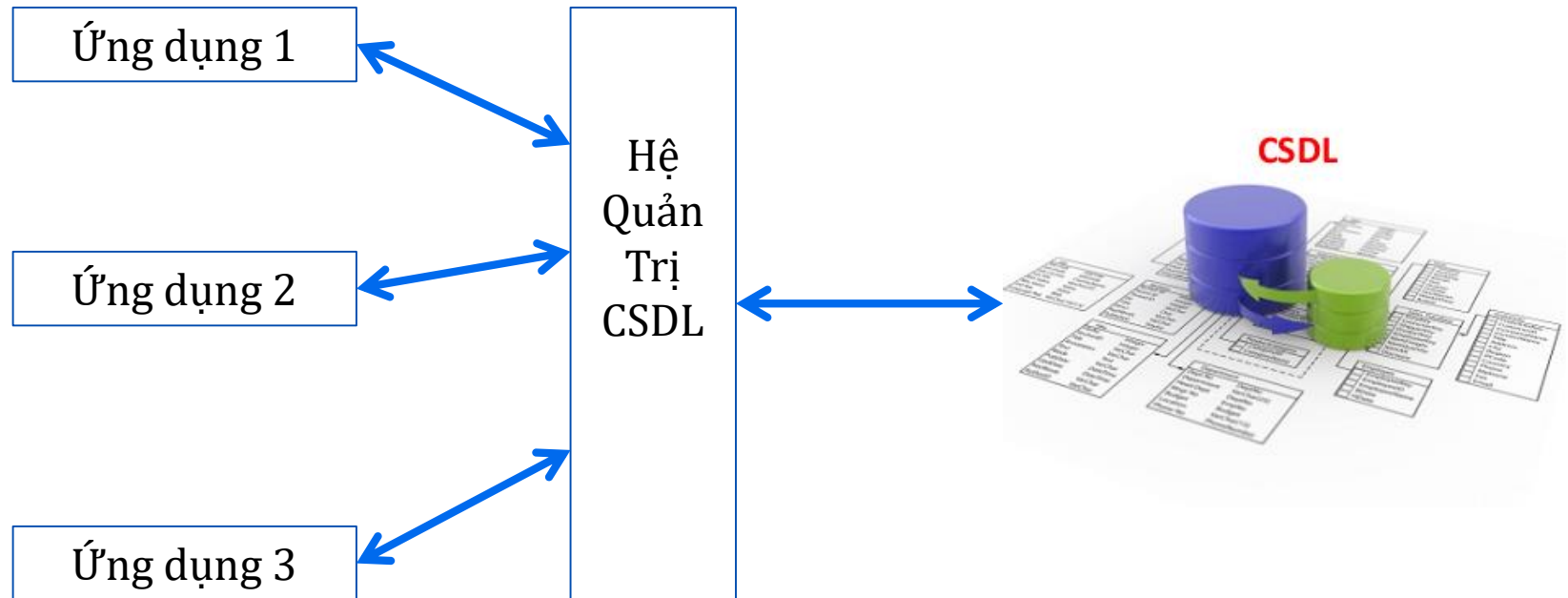
## ❖ Nhược điểm

- Mỗi người dùng khác nhau sẽ lưu trữ tập tin với thông tin theo nhu cầu riêng → Dữ liệu lưu trữ rời rạc
- Dữ liệu dễ bị trùng lặp.
- Dễ bị thiếu nhất quán, mâu thuẫn
- Chia sẻ dữ liệu rất khó khăn.
- Kém hiệu quả trong truy xuất ngẫu nhiên hoặc truy cập đồng thời. Truy xuất khó.
- Khó khôi phục dữ liệu khi có sự cố
- Có vấn đề về an toàn và bảo mật dữ liệu

# Quản lý dữ liệu bằng CSDL

15

- Giúp dữ liệu được lưu trữ một cách hiệu quả và có tổ chức, cho phép quản lý dữ liệu nhanh chóng và hiệu quả.



# Quản lý dữ liệu bằng CSDL

16

## ❖ Ưu điểm:

- Tối thiểu hoá dư thừa thông tin
- Nâng cao tính nhất quán dữ liệu
- Nâng cao tính chia sẻ dữ liệu
- Nâng cao chất lượng dữ liệu
- Nâng cao khả năng truy cập và trả lại dữ liệu
- Giảm chi phí bảo trì chương trình.
- Bảo mật dữ liệu



# Tiêu chuẩn của một CSDL

17

- Tính bảo mật
- Tính độc lập
- Tránh sự dư thừa thông tin
- Có khả năng lưu trữ, phục hồi khi gặp sự cố
- Hiệu suất sử dụng tốt, đảm bảo chế độ ưu tiên khi truy nhập dữ liệu.

# Đặc tính của Cơ sở dữ liệu

18

Tính tự mô tả

1

Tính độc lập

2

Tính trừu tượng

3

Tính nhất quán

4

Đa khung nhìn

5

# Đặc tính của Cơ sở dữ liệu

19

## ➤ Tính tự mô tả

- ✓ Hệ CSDL không chỉ chứa bản thân CSDL mà còn chứa định nghĩa đầy đủ mô tả (cấu trúc+ràng buộc) của CSDL
- ✓ Các định nghĩa được lưu trữ trong catalog: Chứa các thông tin về cấu trúc tập tin, kiểu và dạng thức lưu trữ của mỗi thành phần dữ liệu và những ràng buộc dữ liệu
- ✓ Dữ liệu trong catalog gọi là meta-data
- ✓ Thông tin lưu trong catalog được sử dụng bởi:
  - ✓ Hệ QTCSDL
  - ✓ Người lập trình CSDL

# Đặc tính của Cơ sở dữ liệu

20

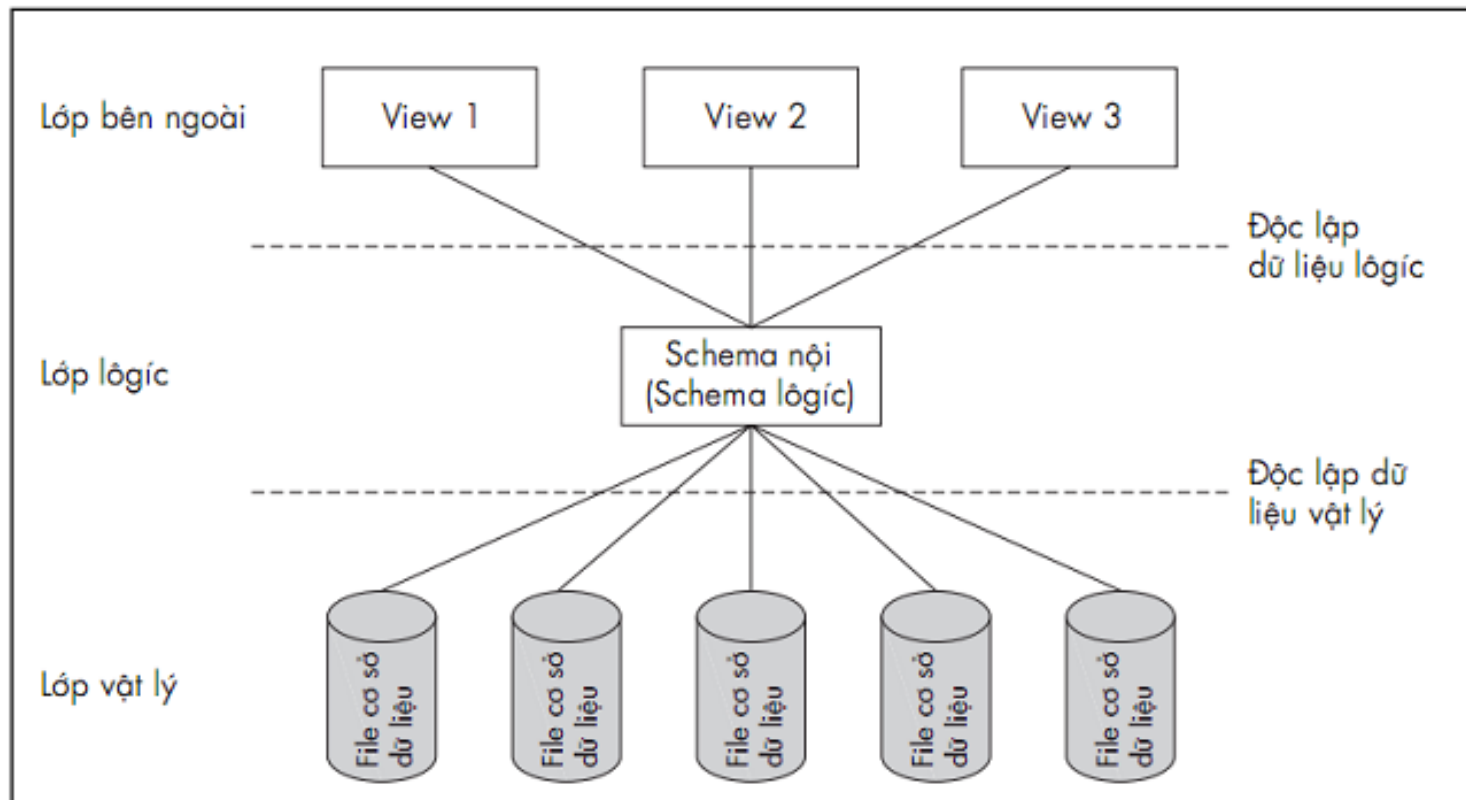
## ➤ Tính độc lập giữa chương trình và dữ liệu

- Cung cấp khả năng trừu tượng hóa thông qua các lớp: vật lý, logic, bên ngoài.
- Sự phân biệt giữa các lớp tạo nên 2 tầng độc lập: độc lập dữ liệu vật lý và độc lập dữ liệu logic
- Một CSDL là độc lập dữ liệu nếu nó có hoặc có thể có khả năng phát triển mà không ảnh hưởng tới các hệ ứng dụng.

# Đặc tính của Cơ sở dữ liệu

21

## ➤ Tính độc lập giữa chương trình và dữ liệu



# Đặc tính của Cơ sở dữ liệu

22

- **Lớp bên ngoài** : là tập tất cả các dữ liệu mà người sử dụng cụ thể có thể nhìn thấy và được phép truy cập (khung nhìn), là mức của người sử dụng và các chương trình ứng dụng.
- **Lớp logic (quan niệm)**: giải quyết vấn đề lưu trữ dữ liệu loại gì, lưu trữ bao nhiêu, mối quan hệ trong CSDL và mối quan hệ giữa các loại dữ liệu này như thế nào?
- **Lớp vật lý**: Đây là mức lưu trữ dữ liệu. Mục đích của mức này giải quyết vấn đề dữ liệu là gì, dữ liệu được lưu trữ như thế nào và nhằm mục đích gì. Người dùng CSDL không nhất thiết phải nắm được cấu trúc tổ chức của các file vật lý khi sử dụng một CSDL.

# Đặc tính của Cơ sở dữ liệu

23

## ➤ Tính độc lập giữa chương trình và dữ liệu

### ➤ Độc lập dữ liệu mức vật lý :

- ✦ Khả năng làm thay đổi cấu trúc file vật lý của một CSDL mà không làm gián đoạn người dùng đang truy cập vào các quá trình đang diễn ra
- ✦ Việc phân biệt lớp vật lý và lớp logic tạo ra sự độc lập vật lý

### ➤ Độc lập dữ liệu mức logic:

- ✦ Khả năng tạo ra các thay đổi tới lớp logic mà không làm gián đoạn người dùng hiện tại và các quá trình đang diễn ra
- ✦ Việc phân biệt giữa lớp logic và lớp bên ngoài gọi là độc lập dữ liệu logic

# Đặc tính của Cơ sở dữ liệu

24

## ➤ Tính trừu tượng

- ✓ HQT CSDL chỉ cung cấp biểu diễn về dữ liệu ở mức khái niệm và che dấu nhiều chi tiết về cách thức lưu trữ
- ✓ Sự trừu tượng hóa giúp cho tính độc lập chương trình và dữ liệu
- ✓ Sự trừu tượng hóa được thể hiện qua mô hình dữ liệu

## ➤ Tính nhất quán

- ✓ Lưu trữ dữ liệu thống nhất
  - Tránh được tình trạng trùng lặp thông tin
- ✓ Có cơ chế điều khiển truy xuất dữ liệu hợp lý
  - Tránh được việc tranh chấp dữ liệu
  - Bảo đảm dữ liệu luôn đúng tại mọi thời điểm



# Đặc tính của Cơ sở dữ liệu

25

- **Đa khung nhìn:** Hỗ trợ nhiều cách nhìn dữ liệu
  - ✓ Một CSDL có nhiều người sử dụng
  - ✓ Mỗi người đòi hỏi cách nhìn (view) khác nhau về CSDL
  - ✓ Một khung nhìn (view) là một phần của CSDL hoặc các dữ liệu được tổng hợp từ CSDL

# Hệ quản trị CSDL

26

## ❖ Hệ quản trị CSDL (database management system – DBMS)

- ✓ Hệ thống **phần mềm**, cung cấp công cụ:
  - Xây dựng CSDL
  - Thao tác trên dữ liệu của CSDL
  - Kiểm soát việc truy xuất dữ liệu
- ✓ Cho phép **định nghĩa, xây dựng và xử lý** dữ liệu
  - Định nghĩa – khai báo khung, các mô tả chi tiết về dữ liệu
  - Xây dựng – lưu trữ dữ liệu trên bộ nhớ phụ
  - Xử lý – truy vấn, cập nhật và phát sinh báo cáo


# Hệ quản trị CSDL

27

## ❖ Một số ví dụ về hệ quản trị CSDL

**ORACLE®**  
DATABASE

**MySQL®**

 **mongoDB**

  
Microsoft®  
**SQL Server®**

  
**MariaDB**

  
**PostgreSQL**

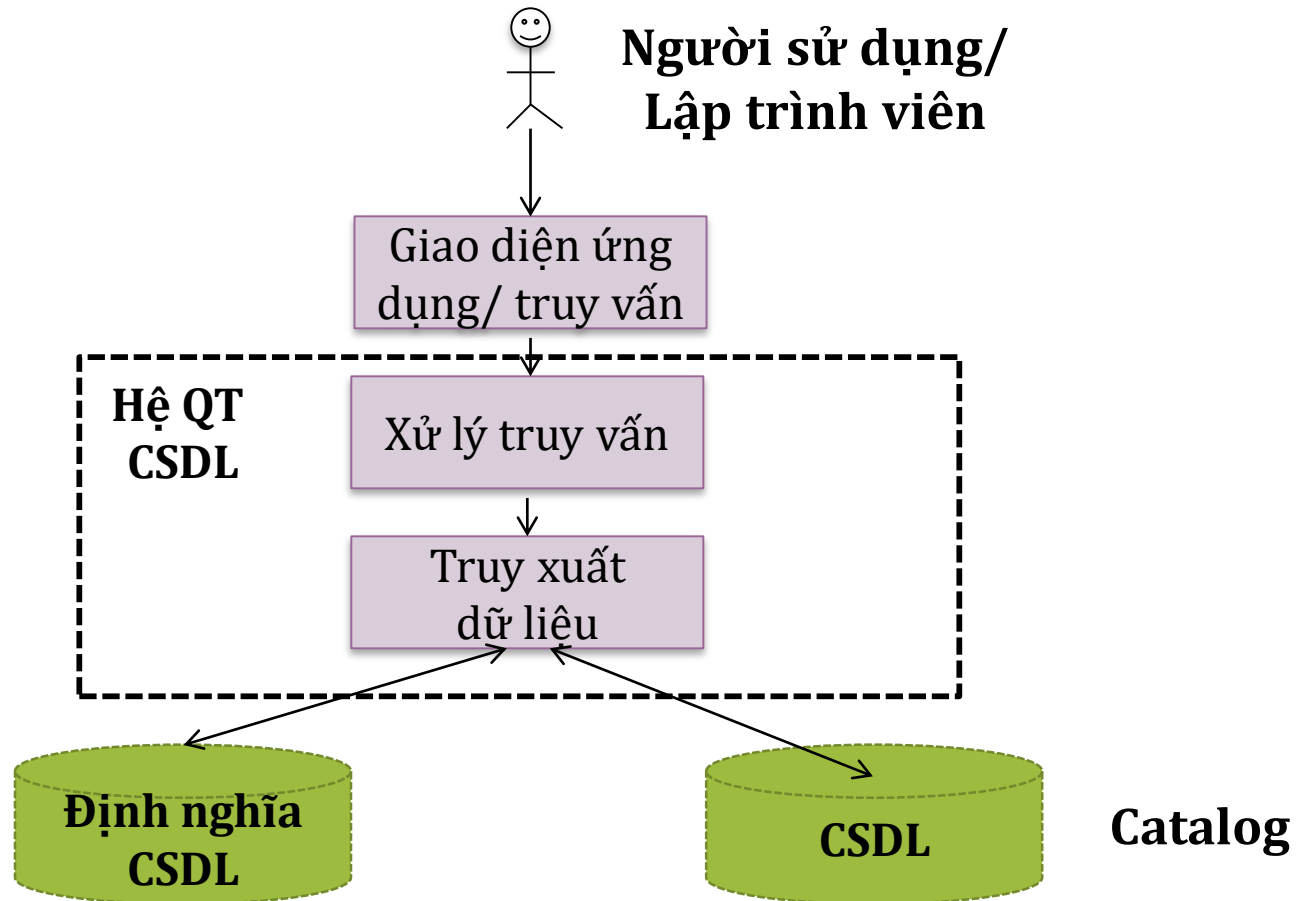
 Microsoft®  
**Access**

**IBM** **DB2**

# Hệ quản trị CSDL

28

❖ Hệ CSDL = CSDL + Hệ quản trị CSDL + Người dùng



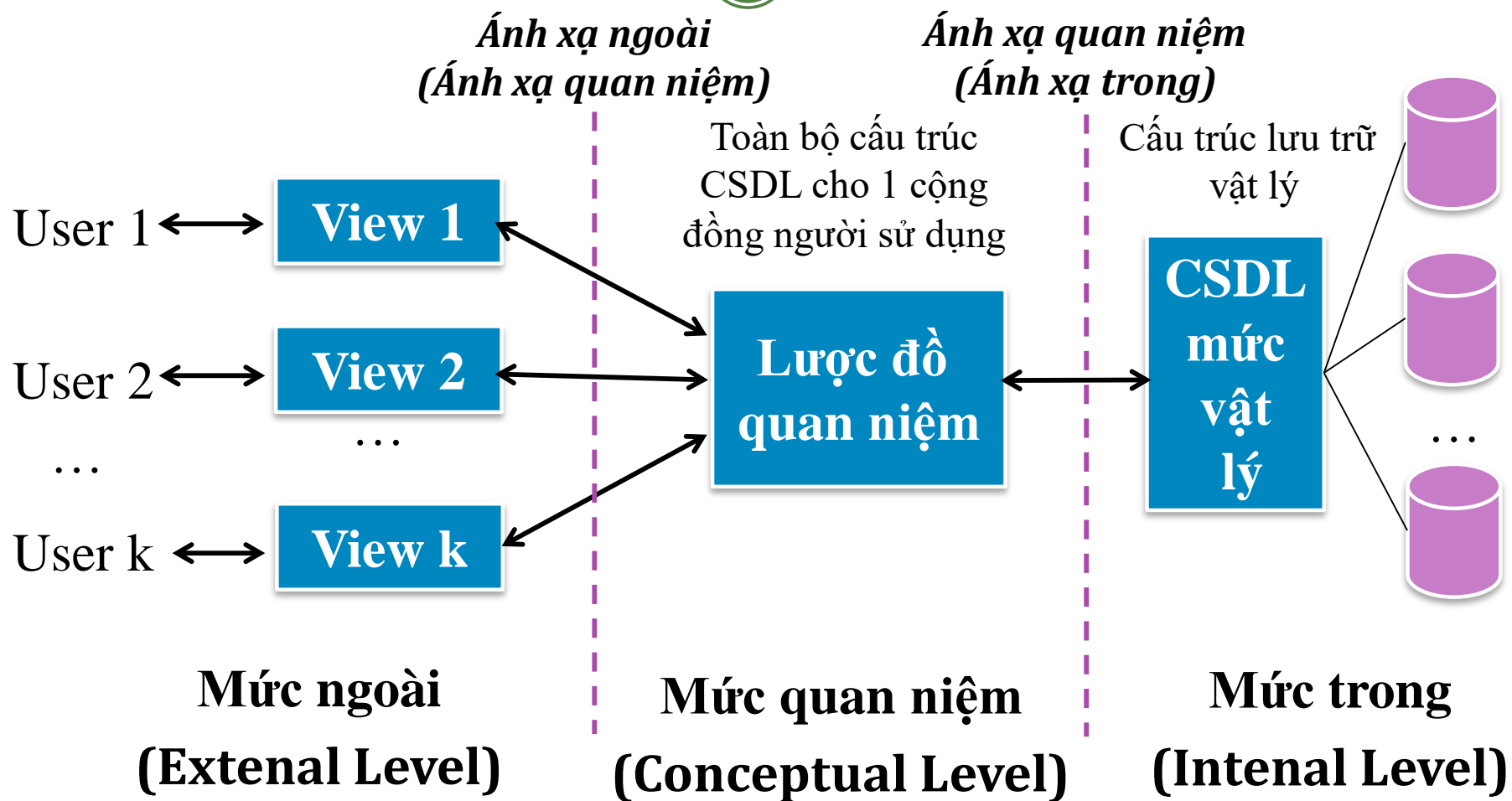
# Lợi ích của hệ quản trị CSDL

29

- ❖ Quản trị CSDL: hạn chế trùng lặp, chia sẻ dữ liệu, Phân quyền, quản lý tài khoản đăng nhập, đảm bảo các ràng buộc giữa CSDL.
- ❖ Cung cấp giao diện làm việc để che dấu các đặc tính phức tạp về mặt cấu trúc tổ chức dữ liệu vật lý.
- ❖ Hỗ trợ ngôn ngữ giao tiếp CSDL.
- ❖ Có cơ chế bảo mật an toàn, khả năng sao lưu, backup.

# Kiến trúc của hệ quản trị CSDL

30



# Người dùng Cơ sở dữ liệu

31

- Người quản trị CSDL (Database Administrator\_DBA)
  - ✓ Có trách nhiệm quản lý hệ CSDL
    - Cấp quyền truy cập
    - Điều phối việc sử dụng
    - Giám sát việc sử dụng
    - Tìm kiếm & đảm bảo phần cứng, phần mềm liên quan
  - ✓ Chịu trách nhiệm cho các vấn đề liên quan đến
    - Tính an toàn của hệ CSDL
    - Khả năng đáp ứng của hệ CSDL

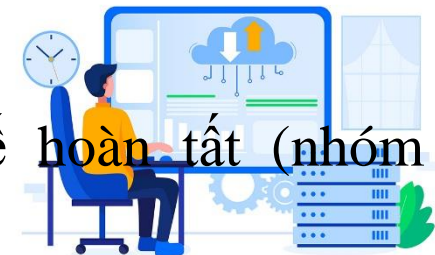


# Người dùng Cơ sở dữ liệu

32

## ➤ Thiết kế viên (DBD – Database Designer)

- ✓ Chịu trách nhiệm về vấn đề :
  - Quyết định những dữ liệu nào cần được lưu trữ
  - Lựa chọn cấu trúc phù hợp để lưu trữ dữ liệu
- ✓ Trao đổi với người dùng để nắm bắt được những yêu cầu và đưa ra một thiết kế CSDL thỏa yêu cầu này
- ✓ Có thể là thành viên của
  - Nhóm DBA quản lý các CSDL
  - Các nhóm khác sau khi việc thiết kế hoàn tất (nhóm lập trình,...)



DATABASE DEVELOPER  
VECTOR ILLUSTRATION



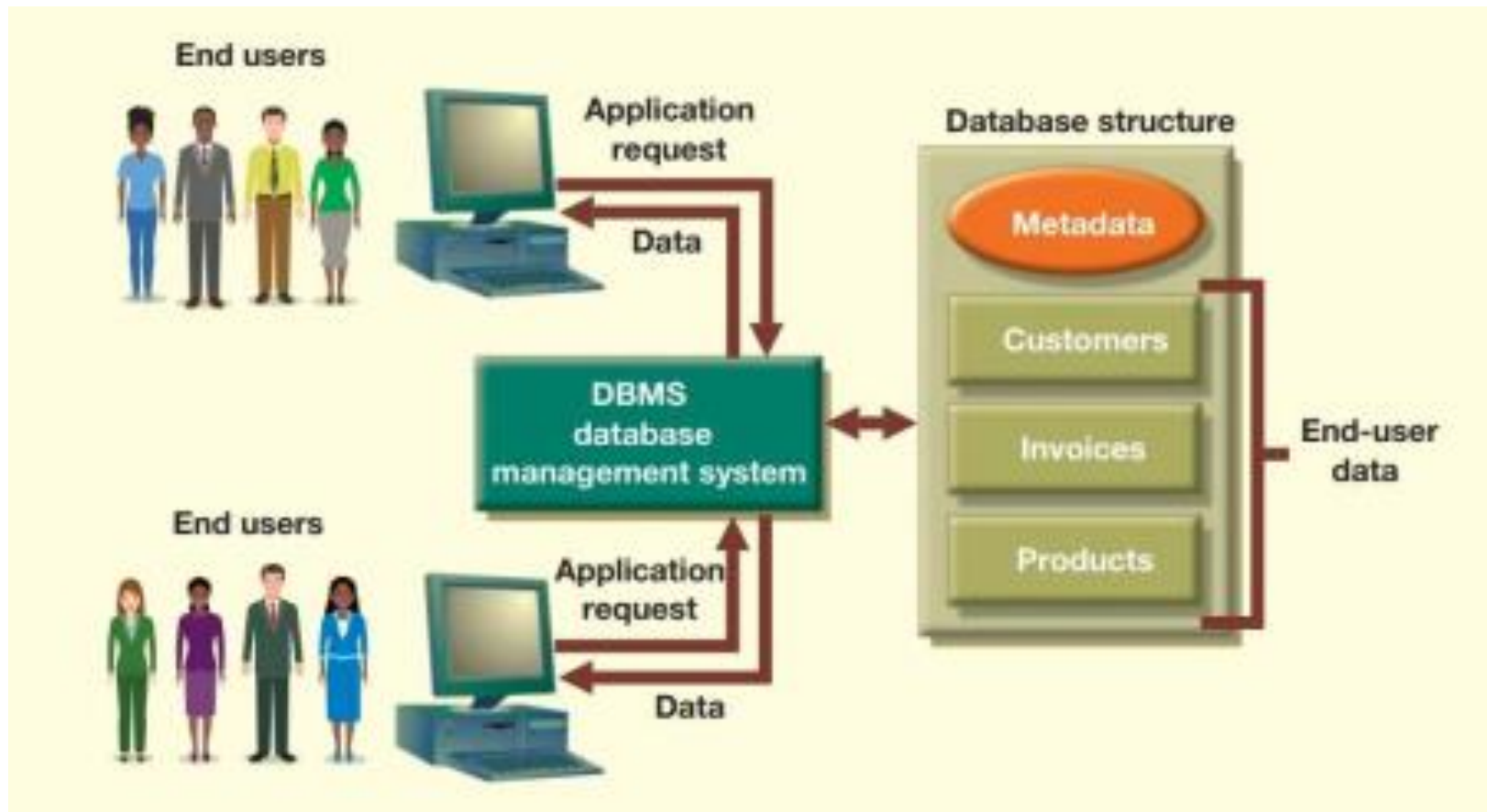
# Người dùng Cơ sở dữ liệu

33

- Người dùng cuối (End user)
  - ✓ Người ít sử dụng (Casual users)
    - Ít khi truy cập CSDL, nhưng cần những thông tin khác nhau trong mỗi lần truy cập và dùng những câu truy vấn phức tạp
    - Người quản lý
  - ✓ Người sử dụng thường xuyên (Naïve users)
    - Thường xuyên truy vấn và cập nhật CSDL nhờ vào một số các chức năng đã được xây dựng sẵn và kiểm thử
    - Nhân viên
  - ✓ Người sử dụng đặc biệt (Sophisticated users)
    - Thông thạo về HQT CSDL, tự xây dựng những truy vấn phức tạp cho công việc
    - Kỹ sư, nhà khoa học, người phân tích,...
  - ✓ Người sử dụng cá nhân (stand-alone users)
    - Sử dụng các ứng dụng truy cập HQT CSDL để phục vụ công việc cá nhân
    - Người sử dụng chương trình quản lý thuế cá nhân, ...

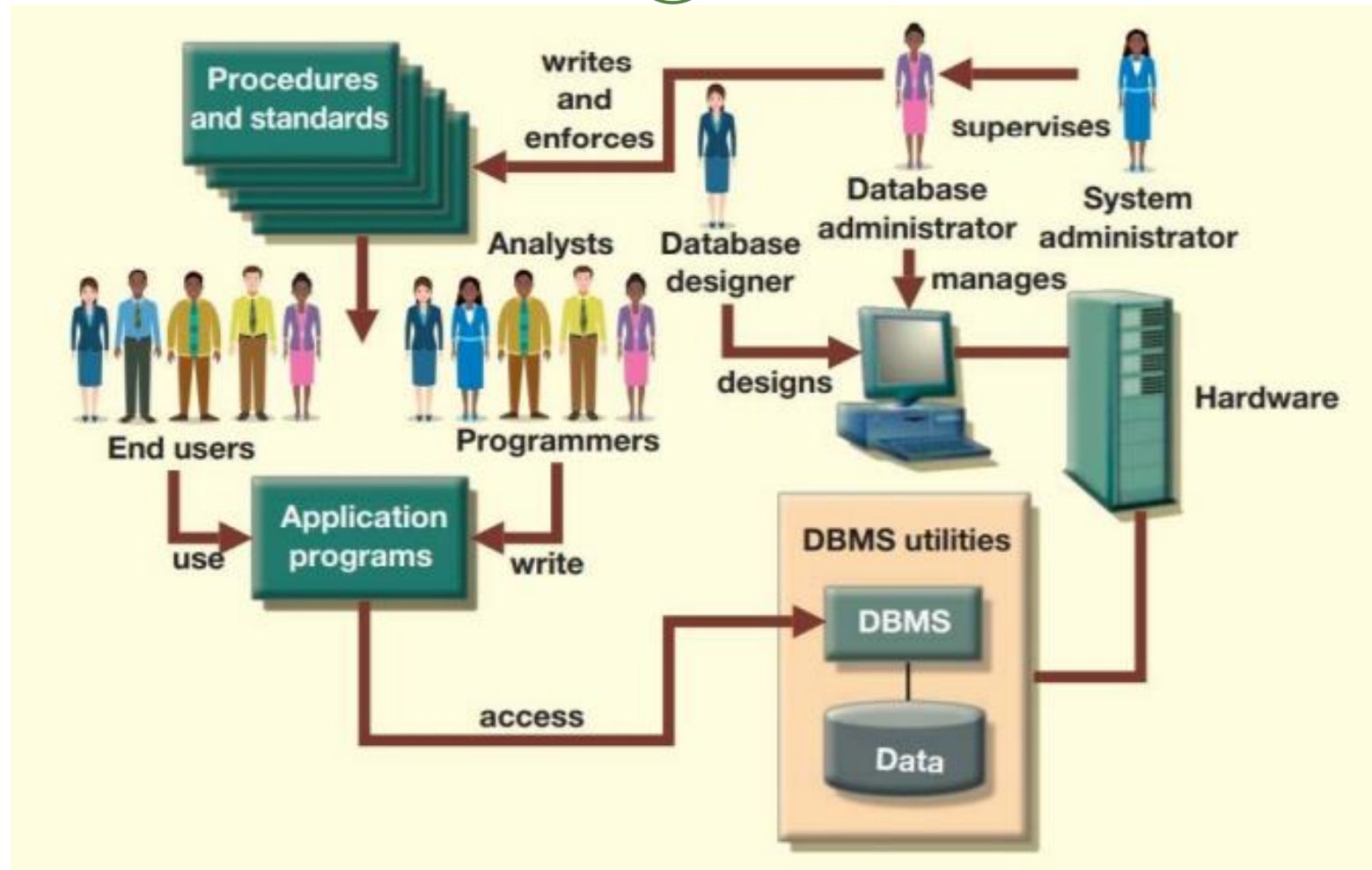
# Người dùng Cơ sở dữ liệu

34



# Người dùng Cơ sở dữ liệu

35



# Hệ quản trị CSDL quan hệ

36

- Hệ quản trị CSDL quan hệ (Relational Database Management System – RDBMS).
- Là một dạng của DBMS được sử dụng rộng rãi nhất, trong đó tất cả dữ liệu được tổ chức chặt chẽ dưới dạng các **bảng** dữ liệu.
- Tất cả các thao tác trên CSDL đều diễn ra trên bảng.

# Hệ quản trị CSDL quan hệ

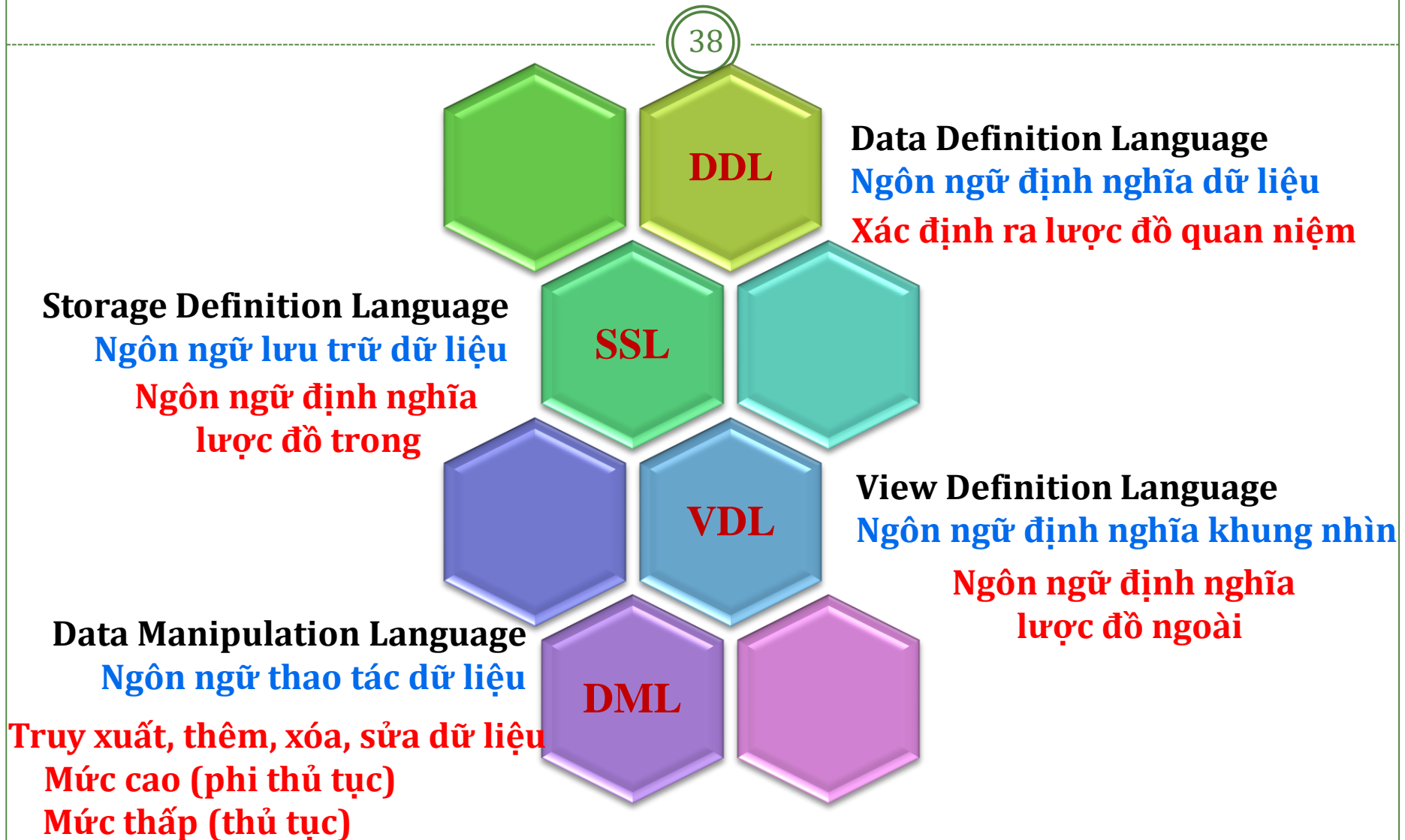
37

- Dễ dàng định nghĩa, duy trì và thao tác dữ liệu lưu trữ
- Trích xuất dữ liệu dễ dàng
- Dữ liệu được chuẩn hóa và bảo vệ tốt
- Nhiều nhà cung cấp phần mềm
- Dễ dàng chuyển đổi giữa nh cung cấp và nhà triển khai
- Là sản phẩm trưởng thành và ổn định



# Ngôn ngữ cơ sở dữ liệu

38



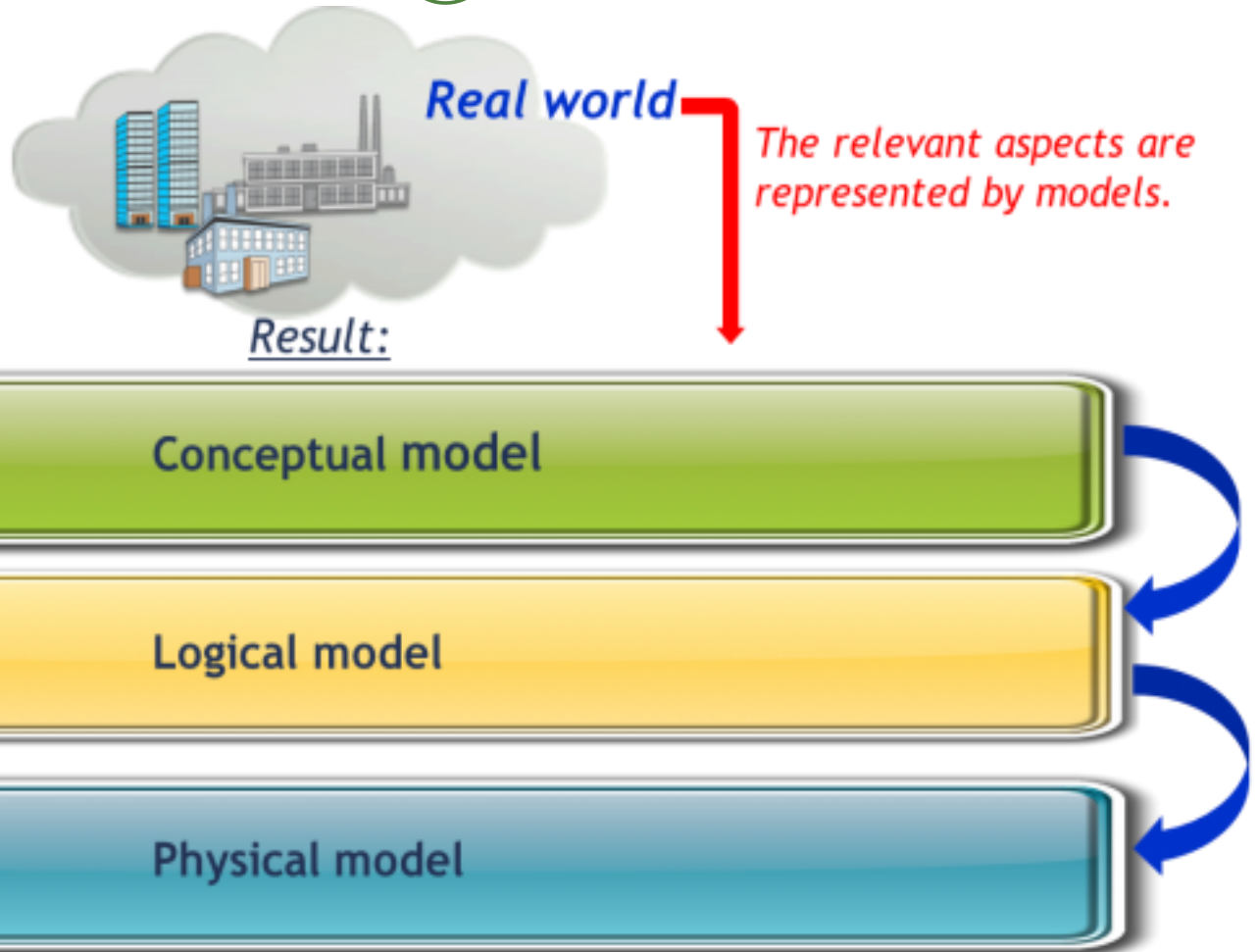
# Vòng đời phát triển CSDL

39



# Thiết kế cơ sở dữ liệu

40





# Thiết kế cơ sở dữ liệu

41

