

Cơ Sở Dữ Liệu

# Chương 1. Tổng quan về Cơ sở dữ liệu

Trần Hoài Thuận

Ngày 06 tháng 08 năm 2023



**GIA DINH**  
UNIVERSITY



DEPARTMENT OF  
INFORMATION TECHNOLOGY  
GIA DINH UNIVERSITY



# Nội Dung

1. Định nghĩa Dữ liệu, CSDL, hệ quản trị CSDL, hệ CSDL
2. Các chức năng của hệ quản trị CSDL
3. Các đặt trưng của hướng tiếp cận CSDL
4. Kiến trúc của hệ quản trị CSDL
5. Các mô hình cơ sở dữ liệu
6. Ví dụ ứng dụng cơ sở dữ liệu



Dữ liệu

# Dữ liệu

- **Dữ liệu** (data) là toàn bộ những gì chúng ta biết, có thể **ghi lại** **được** và **có ý nghĩa rõ ràng**. Ví dụ: các sự kiện, văn bản, hình ảnh, âm thanh,...
- Dữ liệu được chia làm 2 loại: **dữ liệu tĩnh** và **dữ liệu động**.
  - Dữ liệu tĩnh: Tên, địa chỉ, số điện thoại của khách hàng,...
  - Dữ liệu động: Báo cáo doanh thu, dữ liệu về thời tiết, khí hậu,...
- **Thông tin** là một khái niệm phản ánh tri thức, **sự hiểu biết** của con người về một đối tượng nào đó.
- Thông tin là nội dung của dữ liệu sau khi được xử lý.





# Cơ sở dữ liệu

# Cơ sở dữ liệu

- Cơ sở dữ liệu (database) là một **tập hợp dữ liệu** được tổ chức theo một **cấu trúc chặt chẽ**, được lưu trữ trên máy, nhằm **phục vụ (chia sẻ)** cho **nhiều mục tiêu** khác nhau một cách **có chọn lọc**.
- Một CSDL thường bao gồm một hoặc nhiều bảng (table). Mỗi bảng được xác định thông qua một tên. Bảng chứa các cột (column), dòng (record - row) là dữ liệu của bảng.
- Ví dụ: Cơ sở dữ liệu SOHUUOTO có các bảng sau:
  - CONGDAN: Thông tin về công dân có thẻ sở hữu xe ôtô
  - OTO: Thông tin về xe ôtô
  - SOHUU: Thông tin về công dân sở hữu xe ôtô



# Ví dụ 1: Cơ sở dữ liệu sở hữu xe ôtô

Số CMND	Họ tên	Ngày sinh	Giới tính
1	Đỗ Nam Trung	1990-06-07	Nam
2	Nguyễn Đông Anh	1991-02-10	Nam
3	Trần Thu Ngân	1995-10-17	Nữ
4	Hoàng Vũ	1980-01-01	Nam
5	Lê Quỳnh Chi	2000-10-10	Nữ

Biển số	Số máy	Số khung	Số chỗ	Nhãn hiệu	Màu sơn
BS1	1	3	5	Honda	Đen
BS2	5	9	5	Mazda	Đỏ
BS3	3	2	2	Hyundai	Xám
BS4	10	7	7	Toyota	Vàng

Số CMND	Biển số	Ngày sở hữu
1	BS1	2020-01-01
1	BS4	2019-06-09
3	BS3	2016-07-09
2	BS2	2018-01-20
5	BS3	2016-07-09



## Ví dụ 2: Cơ sở dữ liệu phân công đề án

NHANVIEN	HONV	TENLOT	TENNVL	MANV	NGSINH	MA_NQL	PHG
	Tran	Hong	Quang	987987987	03/09/1969	987654321	4
	Nguyen	Thanh	Tung	333445555	12/08/1955	888665555	5
	Nguyen	Manh	Hung	666884444	09/15/1962	333445555	5
	Tran	Thanh	Tam	453453453	07/31/1972	333445555	5

DEAN	TENDA	MADA	DDIEM_DA	PHONG
	San pham X	1	VUNG TAU	5
	San pham Y	2	NHA TRANG	5
	San pham Z	3	TP HCM	5
	Tin hoc hoa	10	HANOI	4

PHANCONG	MA_NVIENT	SODA	THOIGIAN
	123456789	1	32.5
	123456789	2	7.5
	666884444	3	40.0
	453453453	1	20.0



# Các tính chất của một CSDL

- Một CSDL phải biểu thị một khía cạnh nào đó của thế giới thực. Thông tin được đưa vào trong CSDL tạo thành một không gian cơ sở dữ liệu.
- Một CSDL được thiết kế, xây dựng và lưu trữ để phục vụ cho mục đích cụ thể.
- Một CSDL là một tập hợp có cấu trúc của những phần dữ liệu có liên quan với nhau được lưu trữ trong máy tính. Một tập các dữ liệu ngẫu nhiên không được xem là CSDL.



# Hệ quản trị CSDL

# Hệ quản trị cơ sở dữ liệu

- Một **hệ quản trị cơ sở dữ liệu** (Database management system - DBMS) là một **tập hợp các chương trình** cho phép người dùng tạo ra, duy trì và khai thác một cơ sở dữ liệu. Ngoài ra nó còn làm dễ quá trình định nghĩa, xây dựng, xử lý và chia sẻ CSDL giữa những người dùng khác nhau và ứng dụng.
  - Định nghĩa CSDL: Đặt tả các kiểu dữ liệu, các cấu trúc và các ràng buộc của dữ liệu được lưu trữ trong CSDL.
  - Xây dựng CSDL: là quá trình lưu trữ dữ liệu trên các phương tiện lưu trữ được điều khiển bởi một hệ quản trị CSDL.
  - Xử lý CSDL: Bao gồm các thao tác như truy vấn CSDL, cập nhật dữ liệu và tạo ra các báo cáo.
  - Chia sẻ CSDL: Cho phép nhiều người dùng và nhiều ứng dụng truy xuất đến CSDL.
- Ví dụ: Microsoft Access, MySQL, SQL Server, Oracle,...



# Các chức năng của hệ quản trị CSDL

- Lưu trữ dữ liệu, tìm kiếm, cập nhật.
- Quản trị siêu cơ sở dữ liệu (Catalog).
- Hỗ trợ các giao tác
- Chia sẻ dữ liệu và cho phép truy xuất dữ liệu đồng thời.
- Đảm bảo an toàn dữ liệu, hạn chế truy cập không được phép.
- Có khả năng sao lưu và khôi phục dữ liệu khi có sự cố về phần cứng hoặc phần mềm.
- Cung cấp giao diện giao tiếp giữa cơ sở dữ liệu và người dùng.
- Quản lý lưu trữ vật lý



# Các chức năng của hệ quản trị CSDL

- Có thể kiểm tra và đảm bảo ràng buộc toàn vẹn của dữ liệu.
- Hạn chế sự trùng lặp và kiểm soát tính dư thừa của dữ liệu.
- Một số tính năng khác: Chuẩn hoá, uyển chuyển, khả dụng, giảm thời gian phát triển ứng dụng,...



# Các vai trò trong cơ sở dữ liệu

- Quản trị dữ liệu (Data Administrator)
- Quản trị cơ sở dữ liệu (Database Administrator)
- Thiết kế cơ sở dữ liệu (Database Designer)
- Lập trình viên ứng dụng (Programmer)
- Người dùng cuối (End User)



# Thành phần của hệ quản trị cơ sở dữ liệu

- Các ngôn ngữ của hệ quản trị CSDL:
  - Ngôn ngữ định nghĩa dữ liệu (data definition language – DDL):  
Dùng để định nghĩa các lược đồ
  - Ngôn ngữ thao tác dữ liệu (data manipulation language – DML): Dùng để thao tác cơ sở dữ liệu



# Thành phần của hệ quản trị cơ sở dữ liệu

- Các ngôn ngữ của hệ quản trị CSDL:
  - Ngôn ngữ định nghĩa dữ liệu (data definition language – DDL):  
Dùng để định nghĩa các lược đồ
  - Ngôn ngữ thao tác dữ liệu (data manipulation language – DML): Dùng để thao tác cơ sở dữ liệu
- Các loại giao tiếp của hệ quản trị CSDL:
  - Menu: Thể hiện cho người dùng một danh sách các tùy chọn
  - Form: Thể hiện các hình thức cụ thể cho người dùng
  - GUI: Hiển thị các lượt đồ đến người dùng bằng biểu đồ
  - Ngôn ngữ tự nhiên: truy vấn dữ liệu bằng ngôn ngữ tiếng Anh hoặc bằng một ngôn ngữ cụ thể được quy định
  - Kết hợp các loại giao tiếp trên



# Ưu điểm của hệ quản trị CSDL

- Kiểm soát sự dư thừa dữ liệu
- Dữ liệu nhất quán
- Có nhiều thông tin hơn từ cùng một khối lượng dữ liệu
- Chia sẻ dữ liệu
- Cải thiện tính nhất quán của dữ liệu
- Cải thiện độ an toàn
- Yêu cầu về chuẩn
- Kinh tế



# Ưu điểm của hệ quản trị CSDL

- Cân bằng các yêu cầu có tranh chấp
- Khả năng truy xuất và đáp ứng được cải thiện
- Khả năng sản xuất được cải thiện
- Dễ bảo trì dữ liệu do nó độc lập
- Tăng truy xuất đồng thời
- Cải thiện dịch vụ sao lưu và khôi phục



# Khuyết điểm của hệ quản trị CSDL

- Phức tạp
- Kích thước
- Chi phí mua DBMS
- Chi phí phần cứng phải thêm
- Chi phí chuyển đổi
- Hiệu năng
- Khả năng sự cố tăng





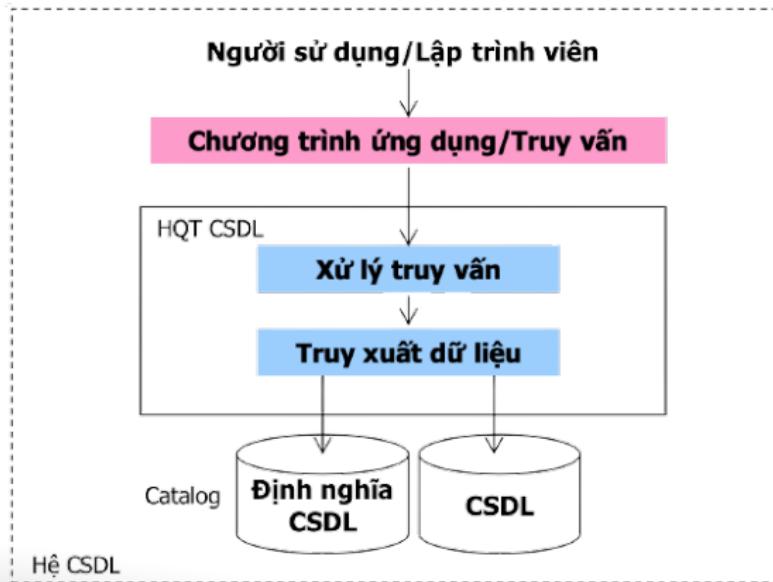
# Hệ cơ sở dữ liệu

# Hệ cơ sở dữ liệu

- Hệ cơ sở dữ liệu (Database system) là sự tích hợp giữa cơ sở dữ liệu và hệ quản trị cơ sở dữ liệu.
- Hệ cơ sở dữ liệu = Cơ sở dữ liệu + hệ quản trị cơ sở dữ liệu

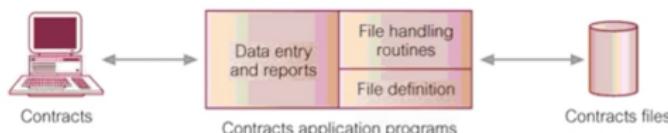
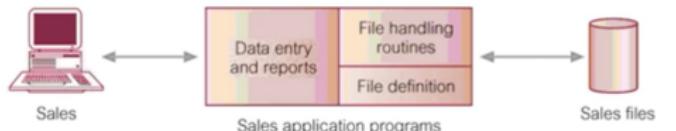


# Hệ cơ sở dữ liệu



# Quá trình phát triển hệ CSDL

- Tiếp cận theo tập tin



Sales Files

**PropertyForRent** (propertyNo, street, city, postcode, type, rooms, rent, ownerNo)

**PrivateOwner** (ownerNo, fName, lName, address, telNo)

**Client** (clientNo, fName, lName, address, telNo, prefType, maxRent)

Contracts Files

**Lease** (leaseNo, propertyNo, clientNo, rent, paymentMethod, deposit, paid, rentStart, rentFinish, duration)

**PropertyForRent** (propertyNo, street, city, postcode, rent)

**Client** (clientNo, fName, lName, address, telNo)



# Quá trình phát triển hệ CSDL

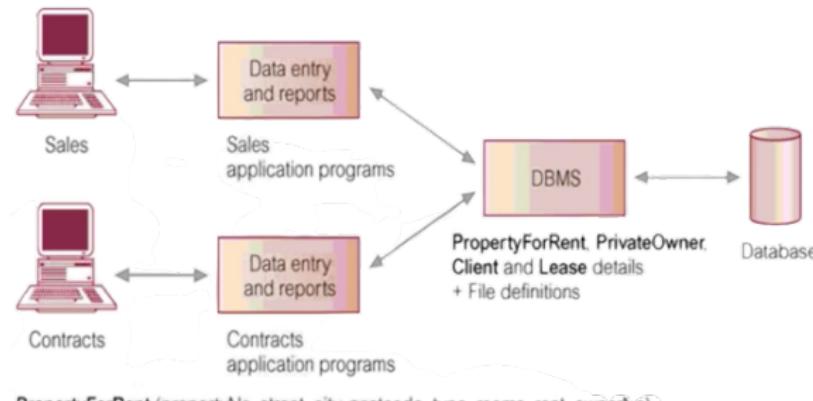
Các hạn chế của cách tiếp cận theo tập tin:

- Dữ liệu tách biệt và không chia sẻ
- Dữ liệu dễ bị trùng lặp và thiếu nhất quán
- Phụ thuộc dữ liệu
- Truy xuất khó khăn do định dạng tập tin không tương thích
- Khó khôi phục dữ liệu khi có sự cố



# Quá trình phát triển hệ CSDL

- Tiếp cận theo cơ sở dữ liệu



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, rentFinish)



# Đặt trưng của hướng tiếp cận CSDL

- Dữ liệu được lưu trữ độc lập với chương trình
- Dữ liệu được quản lý tập trung
- Cho phép chia sẻ cho nhiều ứng dụng sử dụng đồng thời trên một CSDL
- Dữ liệu được đảm bảo an toàn
- Lưu trữ dữ liệu thống nhất, tránh dư thừa, trùng lặp
- Tính tự mô tả của hệ cơ sở dữ liệu
- Tính trừu tượng
- Hỗ trợ nhiều khung nhìn khác nhau về cơ sở dữ liệu



# Lưu trữ dữ liệu dưới dạng CSDL

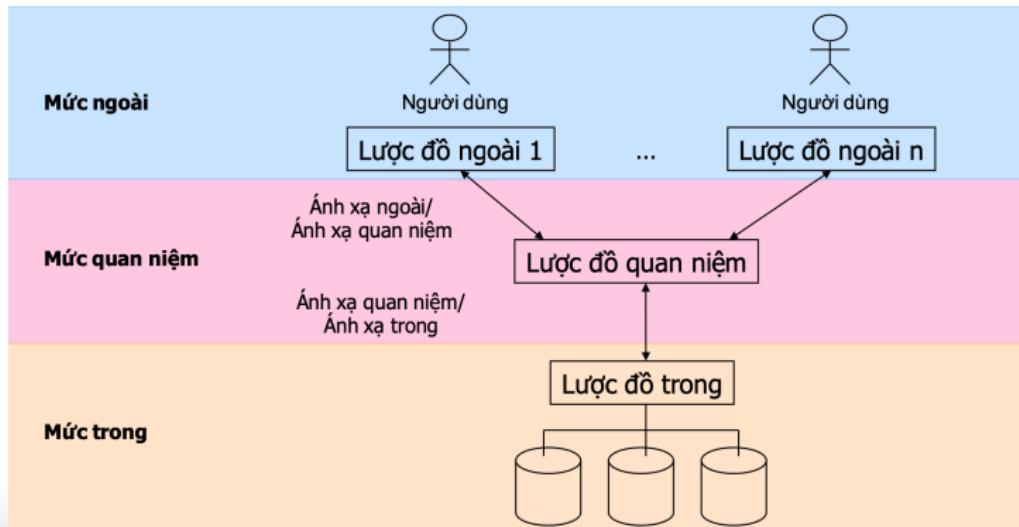
- Cách tiếp cận theo hướng CSDL là một phương pháp quan trọng để xử lý và truy vấn dữ liệu, bao gồm việc sử dụng các câu lệnh truy vấn và các công cụ để truy cập và xử lý dữ liệu.
- Ví dụ, để truy vấn tất cả các hàng trong một bảng có tên là SinhVien trong một CSDL, chúng ta có thể dùng câu truy vấn sau:

```
SELECT * FROM SinhVien;
```



# Kiến trúc của hệ quản trị CSDL

Kiến trúc 3 lượt đồ:



# Ví dụ

Khung nhìn 1

MaNV	Ten	Tuoi	Luong
------	-----	------	-------

Khung nhìn 2

MaNV	Ten	ChiNhanh
------	-----	----------

Mức quan niệm

MaNV	Ten	Tuoi	NgaySinh	Luong	Chinhanh
------	-----	------	----------	-------	----------

Mức trong

```
CREATE TABLE NHANVIEN{
    MaNV CHAR(10),
    Ten NVARCHAR(50),
    Tuoi INT,
    NgaySinh DATETIME,
    Luong DECIMAL(10,3),
    ChiNhanh NVARCHAR(100)};
```



# Các mức nhìn một cơ sở dữ liệu

- Mức ngoài (external level): hay còn gọi là mức khung nhìn. Mức này mô tả các góc nhìn khác nhau của các ứng dụng, người sử dụng cơ sở dữ liệu lên cơ sở dữ liệu.
- Mức quan niệm (conceptual level): hay còn gọi là lượt đồ quan niệm. Mô tả tổng thể những gì có trên cơ sở dữ liệu, đây là góc nhìn của người có trách nhiệm quản trị cơ sở dữ liệu.
- Mức trong (internal level): Mô tả cấu trúc lưu trữ vật lý dữ liệu, đây là góc nhìn của người phát triển cơ sở dữ liệu.



# Độc lập dữ liệu

- Độc lập logic: Các thay đổi của lược đồ quan niệm (ví dụ thêm/xoá thực thể) thì không làm ảnh hưởng đến lược đồ ngoài hay phải viết lại chương trình ứng dụng.



# Độc lập dữ liệu

- Độc lập logic: Các thay đổi của lược đồ quan niệm (ví dụ thêm/xoá thực thể) thì không làm ảnh hưởng đến lược đồ ngoài hay phải viết lại chương trình ứng dụng.
- Độc lập vật lý: Các thay đổi của lược đồ vật lý (ví dụ thay đổi cách thức tổ chức dữ liệu) thì không làm thay đổi lược đồ quan niệm.





# Mô hình dữ liệu

# Định nghĩa mô hình dữ liệu

- Mô hình dữ liệu (data model) là một tập hợp các khái niệm để mô tả cấu trúc CSDL, cũng như các ràng buộc trên CSDL đó.



# Định nghĩa mô hình dữ liệu

- Mô hình dữ liệu (data model) là một tập hợp các khái niệm để **mô tả cấu trúc CSDL**, cũng như các **ràng buộc** trên CSDL đó.
- Một số mô hình CSDL còn có thêm một tập hợp các phép toán cơ bản để đặc tả các thao tác trên CSDL



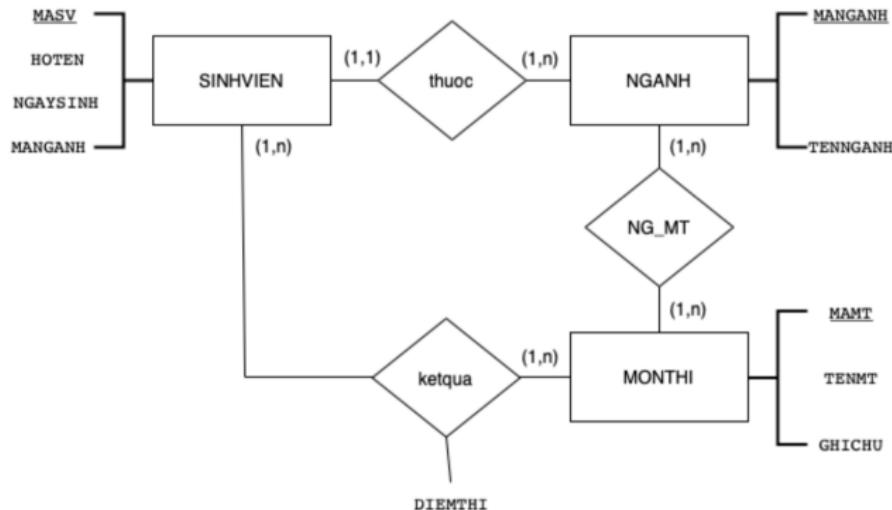
# Các loại mô hình dữ liệu

Mô hình dữ liệu được phân chia thành các loại mô hình chính sau:

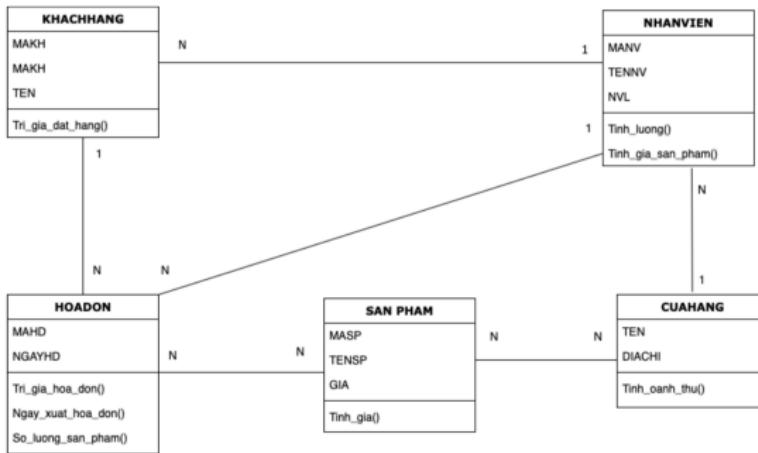
- **Mô hình dữ liệu bậc cao** (mô hình dữ liệu mức quan niệm) cung cấp các khái niệm gần gũi với người dùng.
- **Mô hình dữ liệu bậc thấp** (mô hình dữ liệu vật lý) cung cấp các khái niệm mô tả chi tiết về việc dữ liệu được lưu trữ trong máy tính như thế nào.
- **Mô hình dữ liệu thể hiện** (mô hình dữ liệu mức logic) cung cấp những khái niệm mà người sử dụng có thể hiểu được, đồng thời không xa với cách tổ chức dữ liệu bên trong máy tính.
- **Mô hình dữ liệu quan hệ** là một trong các mô hình dữ liệu mức logic được sử dụng nhiều nhất hiện nay.
- Một số mô hình dữ liệu khác: mô hình dữ liệu đồ thị (graph database), mô hình dữ liệu địa lý (geographical database).



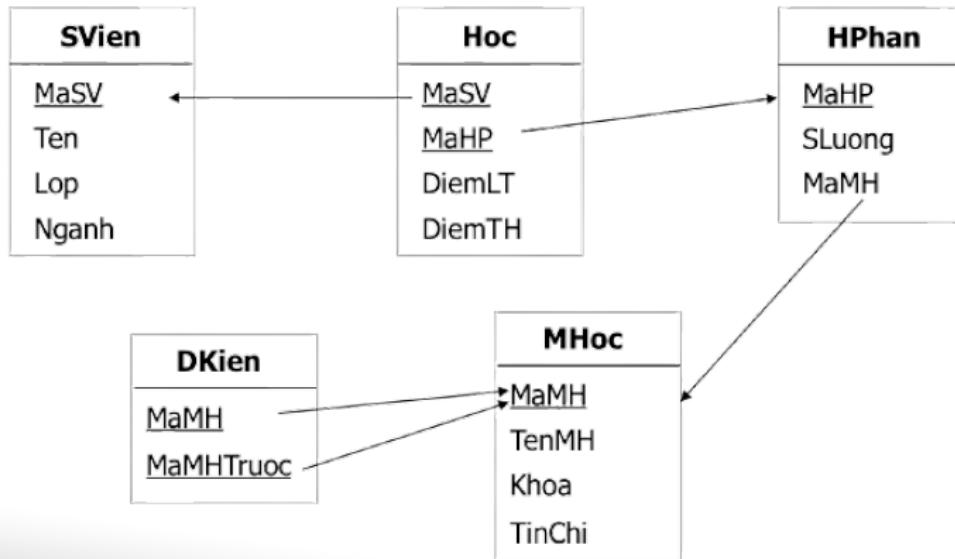
# Ví dụ: Mô hình thực thể kết hợp ER



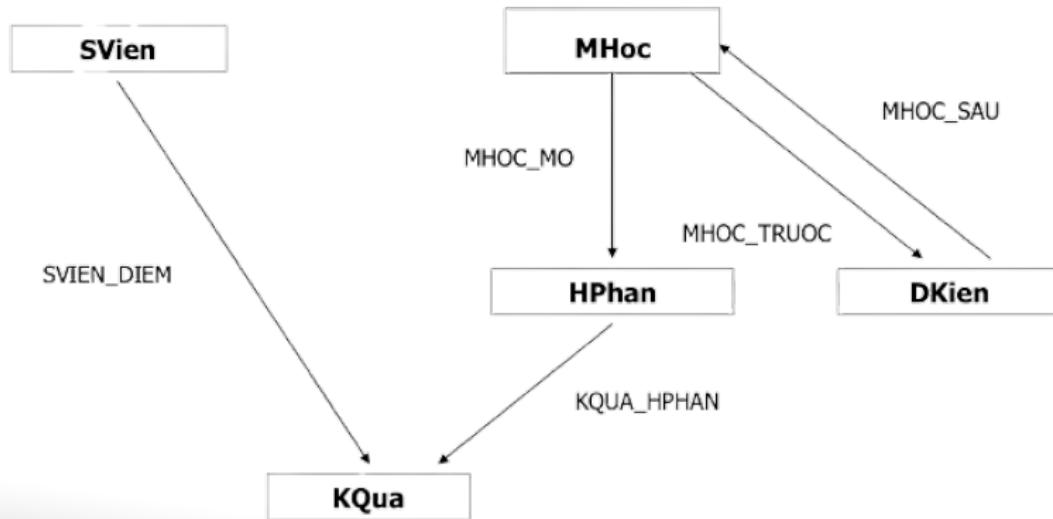
# Ví dụ: Mô hình hướng đối tượng



# Ví dụ: Mô hình dữ liệu quan hệ



# Ví dụ: Mô hình mạng



# Ví dụ: Mô hình phân cấp

Mức 1:

KQua	
DiemTH	
	DiemLT

Mức 2:

HPhan	
TenHP	SLuong

SVien		
TenSV	Lop	Nganh

Mức 3:

MHoc		
TenMH	Khoa	TinChi



# Lược đồ và trạng thái cơ sở dữ liệu

- Mô tả của một cơ sở dữ liệu được gọi là **lược đồ cơ sở dữ liệu**, nó được xác định rõ trong quá trình thiết kế cơ sở dữ liệu và không bị thay đổi thường xuyên.



# Lược đồ và trạng thái cơ sở dữ liệu

- Mô tả của một cơ sở dữ liệu được gọi là **lược đồ cơ sở dữ liệu**, nó được xác định rõ trong quá trình thiết kế cơ sở dữ liệu và không bị thay đổi thường xuyên.
- Các dữ liệu trong một cơ sở dữ liệu tại một thời điểm cụ thể được gọi là một **trạng thái cơ sở dữ liệu** hoặc là ảnh (snapshot) của cơ sở dữ liệu.



# Lược đồ và trạng thái cơ sở dữ liệu

- Mô tả của một cơ sở dữ liệu được gọi là **lược đồ cơ sở dữ liệu**, nó được xác định rõ trong quá trình thiết kế cơ sở dữ liệu và không bị thay đổi thường xuyên.
- Các dữ liệu trong một cơ sở dữ liệu tại một thời điểm cụ thể được gọi là một **trạng thái cơ sở dữ liệu** hoặc là ảnh (snapshot) của cơ sở dữ liệu.
- Trạng thái cơ sở dữ liệu thường xuyên thay đổi trong quá trình sử dụng cơ sở dữ liệu.



# Một vài ứng dụng cơ sở dữ liệu

Ứng dụng Quản lý kết quả học tập của sinh viên.

Mỗi sinh viên cần lưu trữ các thông tin gồm mã số sinh viên, họ tên, địa chỉ, số điện thoại, ngày sinh, giới tính của sinh viên đó. Mỗi sinh viên có thể có nhiều số điện thoại và địa chỉ gồm có các thông tin số nhà, tên đường, quận, thành phố. Mỗi sinh viên phải thuộc về một lớp học nào đó.

Mỗi lớp lưu trữ thông tin mã lớp, số lượng lớp. Lớp phải có một sinh viên làm lớp trưởng và mỗi lớp phải thuộc một chuyên ngành nào đó. Mỗi chuyên ngành cần lưu trữ các thông tin về mã chuyên ngành, tên chuyên ngành. Một chuyên ngành phải thuộc về sự quản lý của một khoa nào đó cụ thể.

Mỗi khoa lưu trữ thông tin về mã khoa, tên khoa, năm thành lập, phòng làm việc, số điện thoại của khoa.



# Ví dụ ứng dụng cơ sở dữ liệu

Mỗi môn thi gồm có thông tin mã môn thi, tên môn thi. Mỗi chuyên ngành sẽ thi nhiều môn, và ngược lại, một môn thi cũng có thể có nhiều chuyên ngành khác nhau cùng thi. Mỗi sinh viên ứng với mỗi môn thi sẽ có một kết quả điểm thi, nếu sinh viên vắng thi môn nào thì bị điểm 0 ở môn thi đó và cần có thông tin ghi chú là “Vắng thi” để phân biệt với một bài thi bị điểm 0.



# Câu hỏi ôn tập

1. Hãy định nghĩa các khái niệm: dữ liệu, cơ sở dữ liệu, hệ quản trị cơ sở dữ liệu, hệ cơ sở dữ liệu.
2. Hãy nêu một cách ngắn gọn ba hoạt động chính liên quan đến cơ sở dữ liệu.
3. Hãy trình bày những thuận lợi trong việc quản lý dữ liệu theo hướng tiếp cận cơ sở dữ liệu mang lại so với việc quản lý dữ liệu theo hướng tiếp cận tập tin.
4. Hãy nêu một số mô hình dữ liệu mà bạn biết.
5. Người sử dụng cơ sở dữ liệu gồm có những loại nào? Hãy cho biết vai trò của từng người dùng.



# Bài tập

1. Hãy tìm một vài ứng dụng thực tế có thể áp dụng cơ sở dữ liệu để quản lý dữ liệu. Với mỗi ứng dụng, hãy trình bày ngắn gọn mô tả về ứng dụng, xác định sơ lược một vài bảng dữ liệu và thuộc tính cho các bảng dữ liệu đó.



Cơ Sở Dữ Liệu

# Chương 1. Tổng quan về Cơ sở dữ liệu

Trần Hoài Thuận

Ngày 06 tháng 08 năm 2023



DEPARTMENT OF  
INFORMATION TECHNOLOGY  
GIA DINH UNIVERSITY



GIA DINH  
UNIVERSITY