

TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN
KHOA HỆ THỐNG THÔNG TIN

MÔN CƠ SỞ DỮ LIỆU - IT004

CHƯƠNG 1: TỔNG QUAN VỀ CƠ SỞ DỮ LIỆU

ThS. TẠ VIỆT PHƯƠNG
phuongtv@uit.edu.vn

Nội dung

I. Các khái niệm

II. Các mô hình dữ liệu



CÁC KHÁI NIỆM

1. Đặt vấn đề
2. Dữ liệu (Data)
3. Hệ thống tập tin (File System)
4. Định nghĩa một CSDL
5. Các đối tượng sử dụng CSDL
6. Hệ quản trị CSDL
7. Các mức của một CSDL

1. Đặt vấn đề

Cơ sở dữ liệu là gì?

Tại sao cần có một cơ sở dữ liệu?



1. Đặt vấn đề

- CDSL giữ vai trò quan trọng trong các lĩnh vực có sử dụng máy tính như:
 - Kinh doanh
 - Y học
 - Giáo dục
 - Truyền thông
 - Thư viện
 - ...
- Thuật ngữ cơ sở dữ liệu trở nên phổ biến
- Trong mỗi lĩnh vực các hoạt động của con người

1. Đặt vấn đề

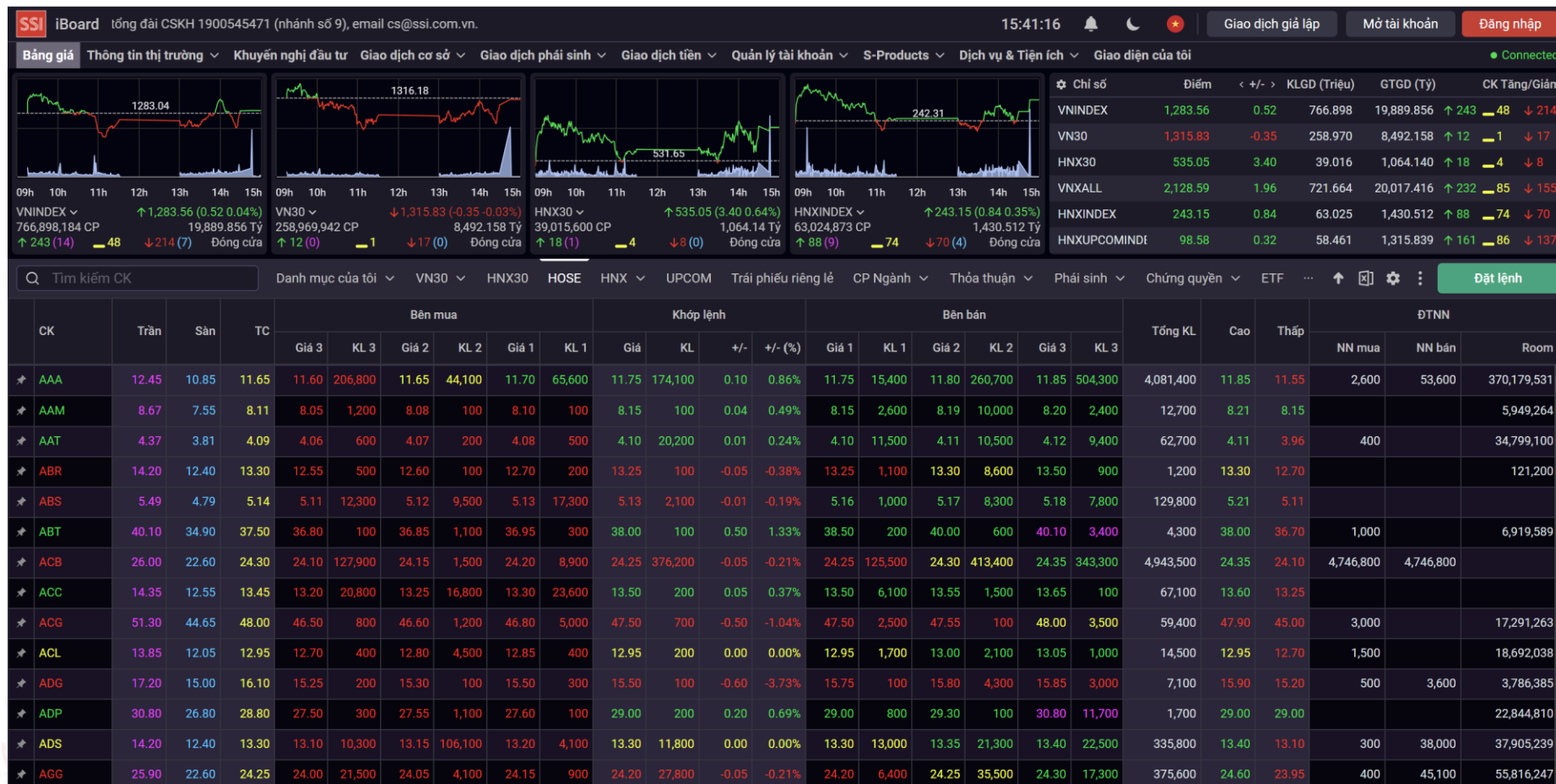
➤ Trong mỗi lĩnh vực, các hoạt động của con người cần có sự giao tiếp với dữ liệu và cơ sở dữ liệu

➤ **Ví dụ:**

- Đến ngân hàng để rút tiền và gửi tiền
- Mua vé máy bay
- Đặt phòng khách sạn
- Truy cập vào thư viện
- Mua hàng trực tuyến
- Đăng ký học phần qua mạng

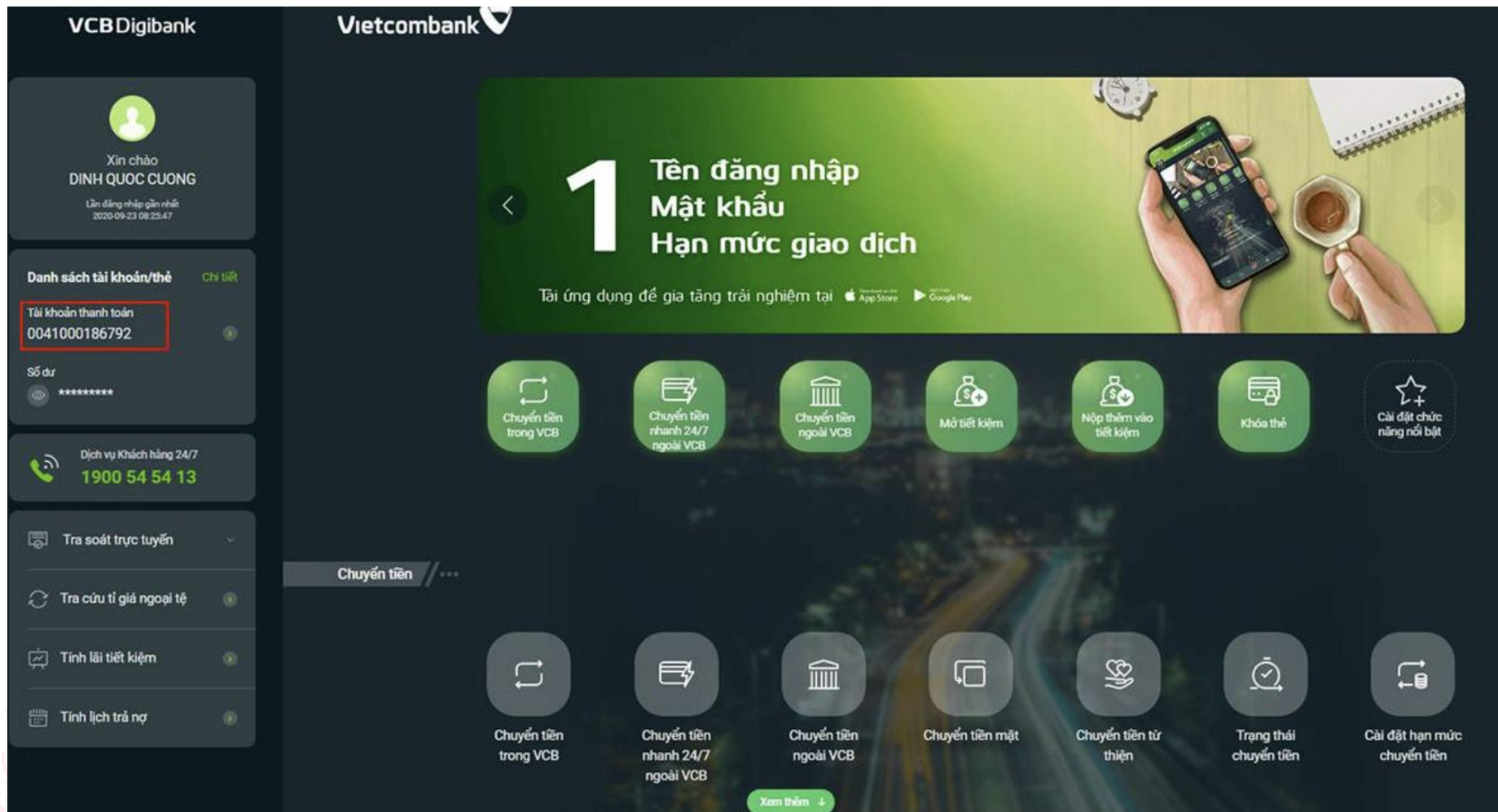
1. Đặt vấn đề

Ví dụ: Theo dõi giá chứng khoán




1. Đặt vấn đề

Ví dụ: Ứng dụng ngân hàng online



1. Đặt vấn đề

Kênh Người Bán | Trở thành Người bán Shopee | Tải ứng dụng | Kết nối  

 Thông Báo  Hỗ Trợ  Tiếng Việt  Đăng Ký | Đăng Nhập



Shopee bao ship 0Đ - Đăng ký ngay!



Sản iPhone 15 Pro Max 1k 1k Áo Giày Nữ Dép HM Vàng Ăn Vật Cặp Da Sạc Sạc Phòng 30000 mAh Quần Jean Ông Loe 1k Body Mist

SALE VẪN CÒN 8 - 9.7

VOUCHER MUA SỚM TỪ 15 TRIỆU ĐỒNG

MUA NGAY TẠI SHOPEE

VINFAST **Shopee Mall**



SALE VẪN CÒN 8 - 9.7

MUỐN ĂN XẾ

SHOPEE FOOD x MONO KHAO

ĐỒNG GIÁ 8.000 ĐỒNG



SALE VẪN CÒN 8 - 9.7

GIẢM ĐẾN 250.000 ĐỒNG

ÁP DỤNG THẺ GHI NỢ VISA TECHCOMBANK

TECHCOMBANK



Voucher Giảm Đến 1 Triệu



Miễn Phí Ship - Có Shopee



Khung Giờ Săn Sale



Mã Giảm Giá



Shopee Siêu Rẻ



Shopee Style Voucher 40%



Hàng Quốc Tế



Nạp Thẻ, Dịch Vụ & Vé Máy Bay

DANH MỤC



Thời Trang Nam



Điện Thoại & Phụ Kiện



Thiết Bị Điện Tử



Máy Tính & Laptop



Máy Ảnh & Máy Quay Phim



Đồng Hồ



Giày Dép Nam



Thiết Bị Điện Gia Dụng



Thể Thao & Du Lịch



Ô Tô & Xe Máy & Xe Đạp

2. Dữ liệu (Data)

- (Trong KHMT) Là các chỉ số của đối tượng (người, vật, một khái niệm, sự việc...) được lưu trữ trên máy tính.
- Là những mảnh ghép thô sơ, rời rạc và chưa được xử lý
- Một mô tả hình thức về những sự kiện, khái niệm
- Dữ liệu là những gì có thể ghi lại được và có ý nghĩa. (By data, we mean known facts that can be recorded and that have implicit meaning - Elmasri, Navathe, 7th)
- Dữ liệu được mô tả dưới nhiều dạng khác nhau (các ký tự, ký số, hình ảnh, ký hiệu, âm thanh...)

Ví dụ:

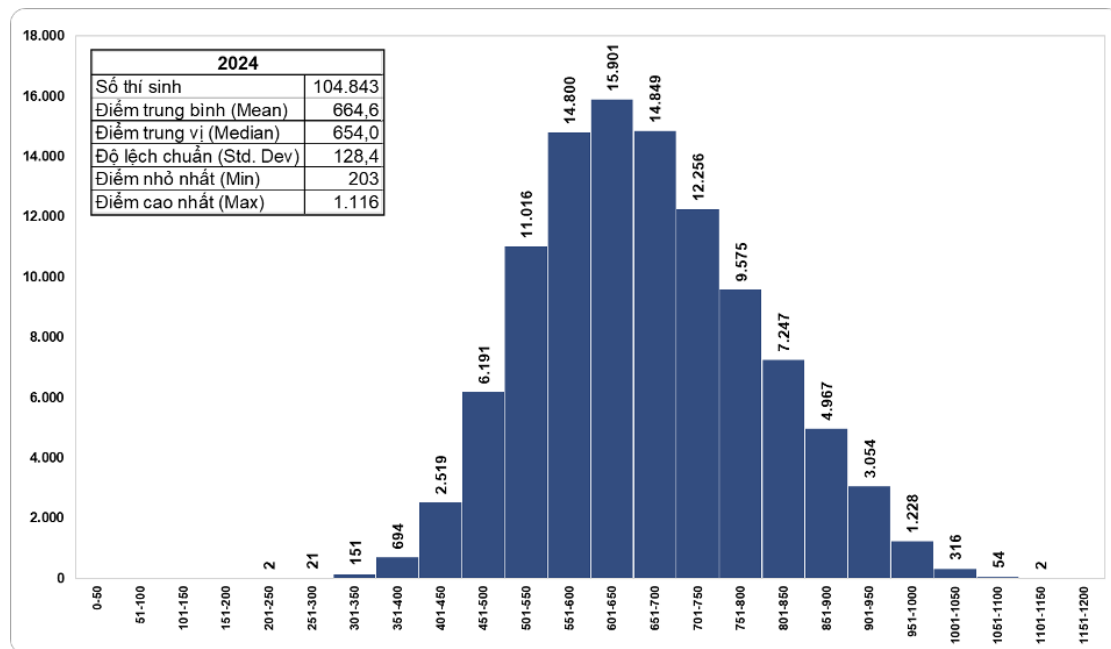
- Thông tin cá nhân của SV
- Bảng điểm của SV

2. Dữ liệu (Data)

➤ Dữ liệu về đối tượng có thể khác nhau, tùy thuộc vào ngữ cảnh

Ví dụ: dữ liệu về SV có thể khác nhau tùy vào mục đích quản lý:

- Quản lý điểm: Tên, mã sinh viên, điểm môn 1, điểm môn 2, điểm môn 3.
- Quản lý nhân thân: Tên, địa chỉ, ngày sinh, quê quán, lớp



2. Dữ liệu (Data)

Ví dụ: trang web của một trung tâm ngoại ngữ



Về chúng tôi

Giảng viên

Học viên

Khóa học

Tư vấn

Đăng ký








Chat trực tiếp với IELTS Fighter





2. Dữ liệu (Data)



Ví dụ: một sản TMĐT



Kênh Người Bán | Trở thành Người bán Shopee | Tải ứng dụng | Kết nối  


Thông Báo  Hỗ Trợ  Tiếng Việt  Đăng Ký | Đăng Nhập

 Shopee 




Điện Thoại 8plus Giá Rẻ | Điện Thoại 1 K | Labubu | Tai Nghe Bluetooth | Ốp Lưng | Capybara Chảy Nước Mũi | Áo Khoác Hoodie Có Dây Kéo | Giày


Shopee > Đồ Chơi > Sở thích & Sưu tầm > Mô hình nhân vật > {  CHÍNH HÃNG POPMART  } LABUBU FLIP WITH ME - HÀNG CÓ SẴN


Yêu Thích {  CHÍNH HÃNG POPMART  } LABUBU FLIP WITH ME - HÀNG CÓ SẴN



5.0  3 Đánh Giá Tổ cáo


₫3.420.000



Chính Sách Trả Hàng  Trả hàng 15 ngày  Đổi ý miễn phí 


Bảo Hiểm **Bảo hiểm Thiệt hại sản phẩm**  [Tìm hiểu thêm](#)






Vận Chuyển  **Miễn phí vận chuyển**

 Vận Chuyển Tới **Huyện Ba Vì** 

Phí Vận Chuyển **₫0** 

Số Lượng  

 **Thêm Vào Giỏ Hàng** **Mua Ngay**



2. Dữ liệu (Data)

Dữ liệu - Thông tin - Kiến thức

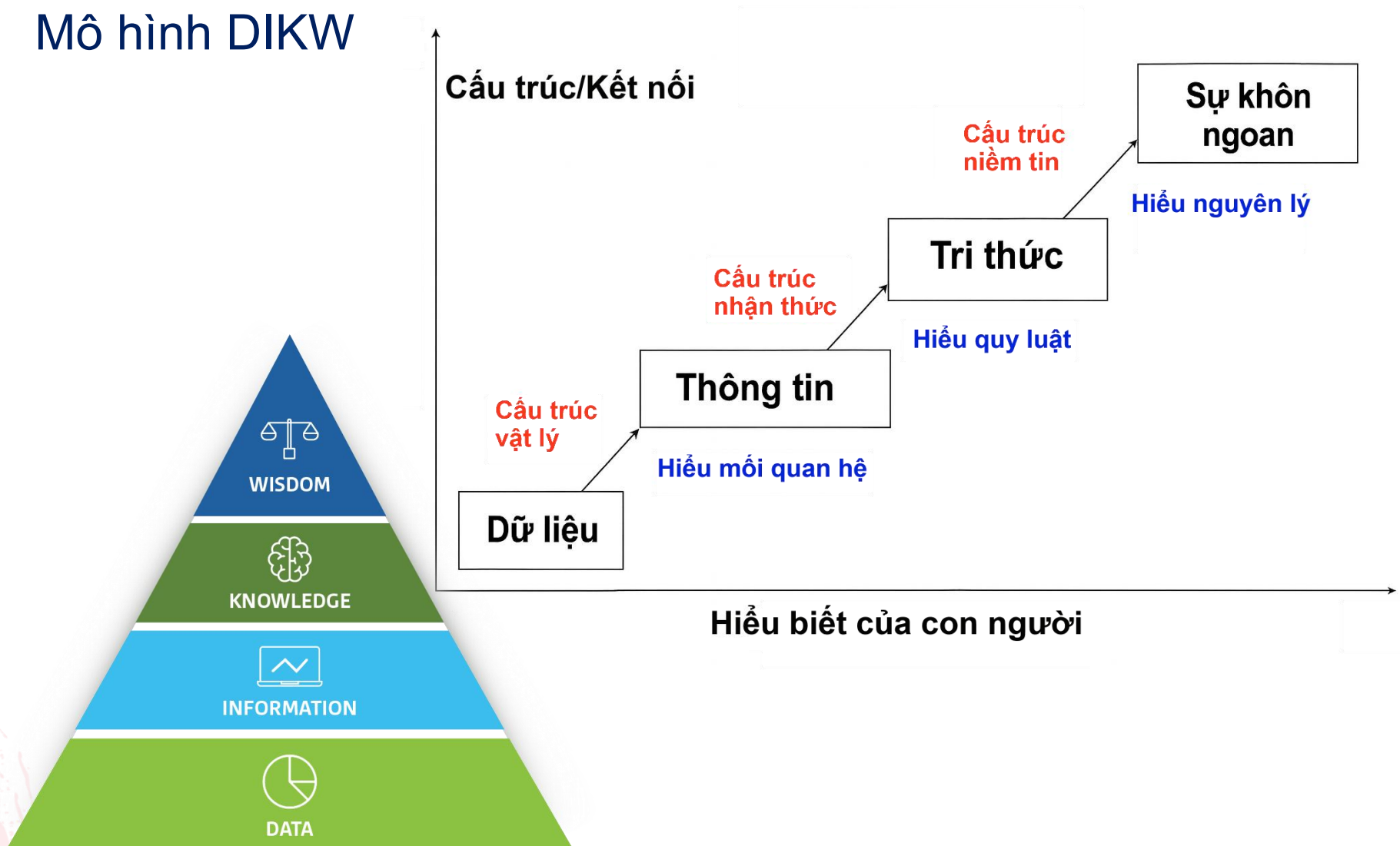
- Bản thân dữ liệu thường không có ý nghĩa trực tiếp đối với con người.
- **Thông tin** (Information) là tất cả những gì mà con người có thể cảm nhận được một cách trực tiếp thông qua các giác quan của mình hoặc gián tiếp thông qua các phương tiện kỹ thuật như tivi, radio, cassette,... Thông tin đối với con người luôn có một ý nghĩa nhất định nào đó. Khi dữ liệu được đưa vào ngữ cảnh, trong tầm hiểu biết của một người quan sát nhận thức, nó sẽ trở thành một thông tin.
- Thông tin là những hiểu biết của con người về một thực thể nào đó, có thể thu thập, lưu trữ, xử lý được. Dữ liệu được dùng để biểu diễn thông tin.

2. Dữ liệu (Data)

Tri thức/kiến thức (Knowledge) là nhận thức về dữ liệu được đưa vào mối quan hệ để hình thành thông tin theo nghĩa rộng hơn. Kiến thức được hỗ trợ bởi kinh nghiệm và các hình thức giáo dục và học tập khác để hiểu được mối quan hệ của dữ liệu với thông tin và cả lý do và ý nghĩa của chúng.

2. Dữ liệu (Data)

Mô hình DIKW



2. Dữ liệu (Data)

- **Quản lý dữ liệu:** Quản lý dữ liệu là quản lý một số lượng lớn dữ liệu, bao gồm cả việc lưu trữ và cung cấp cơ chế cho phép Thao tác (thêm, sửa, xóa dữ liệu) và Truy vấn dữ liệu.
- Hai phương pháp quản lý dữ liệu:
 - Hệ thống quản lý bằng file
 - Hệ thống quản lý bằng CSDL

3. Hệ thống tập tin (File system)

➤ Là **tập hợp các tập tin** riêng lẻ phục vụ cho một mục đích của đơn vị sử dụng.

➤ **Ưu điểm:**

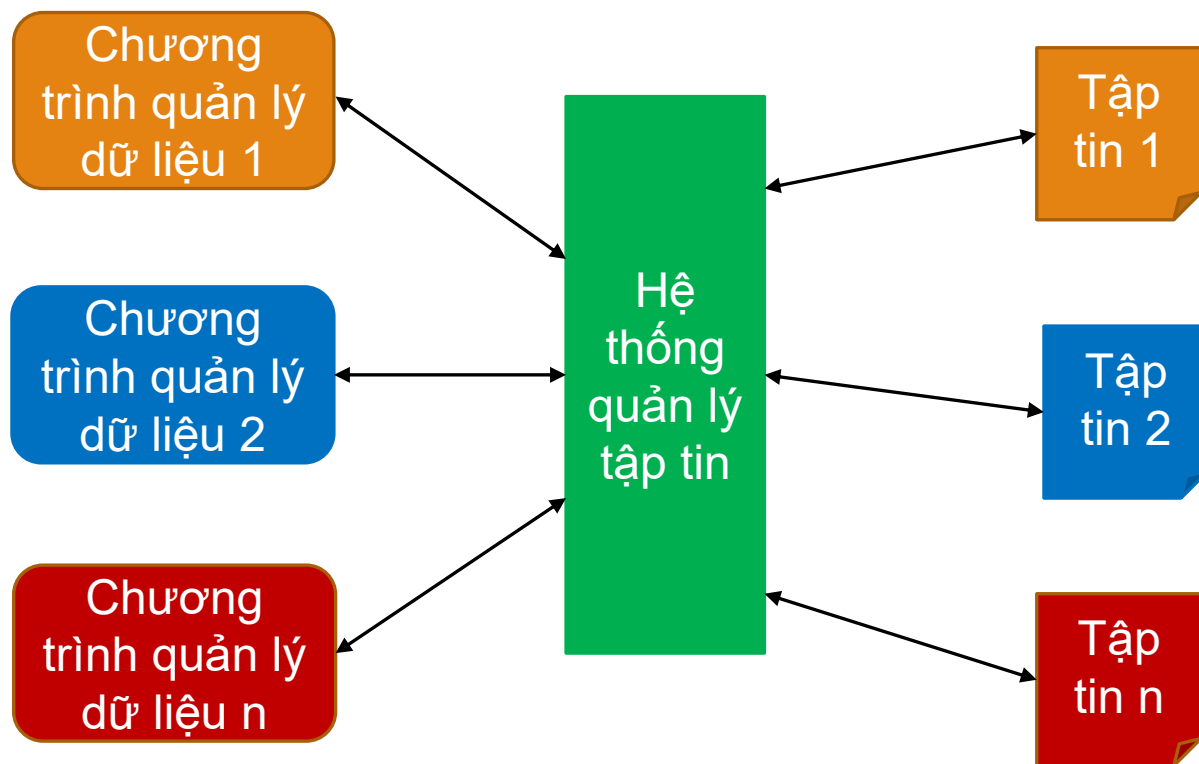
- Triển khai ứng dụng nhanh
- Khả năng đáp ứng nhanh chóng, kịp thời (vì chỉ phục vụ cho mục đích hạn hẹp)

➤ **Khuyết điểm:**

- Trùng lặp dữ liệu → lãng phí, dữ liệu không nhất quán
- Chi phí cao
- Chia sẻ dữ liệu kém
- Khó khăn trong việc truy xuất
- Vấn đề về an toàn và bảo mật

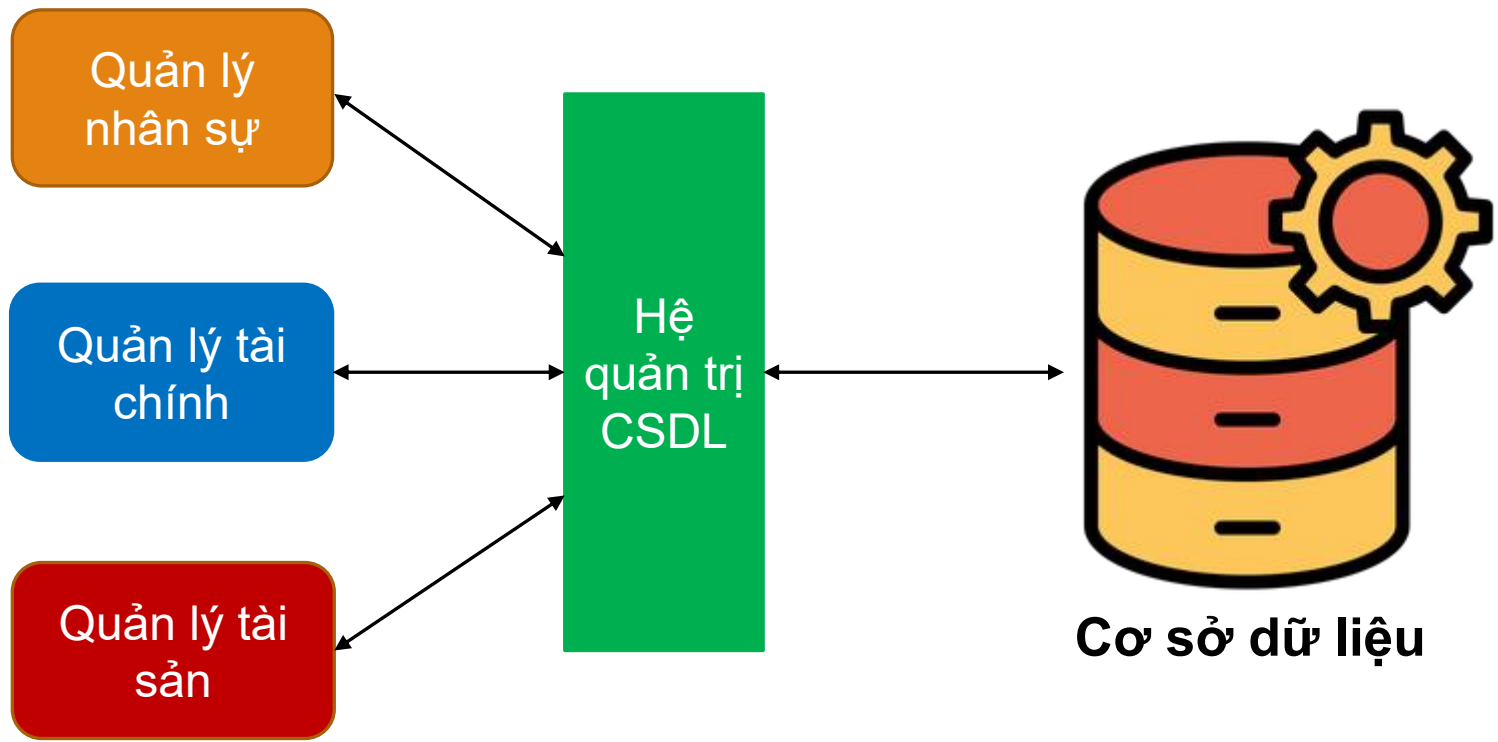
3. Hệ thống tập tin (File system)

- Là tập hợp các tập tin riêng lẻ phục vụ cho một mục đích của đơn vị sử dụng.
- Dữ liệu được lưu trữ trong các file riêng biệt. Ví dụ: các chương trình lưu trữ thông tin bằng hệ thống các file dạng text.



4. Cơ sở dữ liệu (Database)

- Sử dụng cơ sở dữ liệu cho việc quản lý, truy xuất dữ liệu



4. Cơ sở dữ liệu (Database)

- **Cơ sở dữ liệu (CSDL - Database)** là một tập hợp dữ liệu được tổ chức và lưu trữ theo một cấu trúc chặt chẽ, thường mô tả hoạt động của một hay nhiều tổ chức liên quan và được lưu trữ trong máy tính
- CSDL được thiết kế, xây dựng cho phép người dùng lưu trữ dữ liệu, truy xuất thông tin hoặc cập nhật dữ liệu.
- CSDL được tổ chức có cấu trúc: Các dữ liệu được lưu trữ có cấu trúc thành các bản ghi (record), các trường dữ liệu (field). Các dữ liệu lưu trữ có mối quan hệ (relation) với nhau
- Là một hệ thống các thông tin có cấu trúc, được lưu trữ trên các thiết bị lưu trữ nhằm thỏa mãn yêu cầu khai thác thông tin đồng thời của nhiều người sử dụng hay nhiều chương trình ứng dụng với những mục đích khác nhau, được tổ chức theo một mô hình.

4. Cơ sở dữ liệu (Database)

➤ Ưu điểm:

- Giảm trùng lặp thông tin xuống mức thấp nhất, đảm bảo tính nhất quán và toàn vẹn dữ liệu.
- Đảm bảo dữ liệu được truy xuất theo nhiều cách khác nhau.
- Khả năng chia sẻ thông tin cho nhiều người, nhiều ứng dụng khác nhau.

4. Cơ sở dữ liệu (Database)

➤ Những vấn đề cần giải quyết:

- Tính chủ quyền dữ liệu.
- Tính bảo mật và quyền khai thác thông tin của người sử dụng.
- Tranh chấp dữ liệu.
- Đảm bảo dữ liệu khi có sự cố.

4. Cơ sở dữ liệu (Database)

➤ Tính chất của cơ sở dữ liệu

- Một CSDL biểu thị một khía cạnh nào đó của thế giới thực.
- Một cơ sở dữ liệu được thiết kế và được sử dụng cho một mục đích riêng
- Một cơ sở dữ liệu là một tập hợp dữ liệu liên kết với nhau một cách logic và mang một ý nghĩa nào đó.
- Một CSDL không phải là một tập hợp dữ liệu tùy tiện.

4. Cơ sở dữ liệu (Database)

SINH VIÊN

Tên	MSV	Lớp	Ngành
Việt	17	1	CS
Hoài	8	2	IS

KHÓA

Môn học	Mã môn	Số tín chỉ	Khoa
Giới thiệu ngành	CS1310	4	CS
Cấu trúc dữ liệu	CS3320	4	CS
Cơ sở dữ liệu	CS3380	4	IS
Toán cao cấp	MATH2410	3	MATH

4. Cơ sở dữ liệu (Database)

HỌC PHẦN

Mã học phần	Mã khóa	Học kỳ	Năm học	Giảng viên
85	MATH2410	1	2021-2022	A
92	CS1310	2	2021-2022	B
102	CS3320	1	2021-2022	C
135	CS3380	1	2021-2022	D

4. Cơ sở dữ liệu (Database)

ĐIỂM		
MSSV	Mã học phần	ĐIỂM
17	85	A
8	85	A
17	102	B
8	102	B

4. Cơ sở dữ liệu (Database)

➤ Ứng dụng CSDL hiện nay:

- TMĐT (Shopee/TikTok Shop): sản phẩm, đơn hàng, tồn kho, vận chuyển, voucher.
- Mạng xã hội: người dùng, bài đăng, tương tác (like/comment/share), đồ thị quan hệ.
- Fintech (MoMo): ví điện tử, giao dịch, định danh KYC, chống gian lận (fraud).
- Vai trò CSDL: lưu trữ giao dịch → từ đó, phân tích theo thời gian thực → ra quyết định.

5. Các đối tượng sử dụng cơ sở dữ liệu

- **Người sử dụng CSDL (end user):** không chuyên về lĩnh vực tin học và CSDL -> cần công cụ để họ có thể khai thác CSDL khi cần.
- **Chuyên viên tin học:** xây dựng các ứng dụng để phục vụ cho các mục đích quản lý (Người phân tích hệ thống và lập trình ứng dụng)
- **Người quản trị CSDL (Database Administrator):** tổ chức CSDL, bảo mật, cấp quyền, sao lưu, phục hồi dữ liệu, giải quyết các tranh chấp dữ liệu ... (có thể chia ra làm người bảo trì, người vận hành, người thiết kế và cài đặt...)

5. Các đối tượng sử dụng cơ sở dữ liệu

- ❖ Người thiết kế và cài đặt hệ quản trị dữ liệu
- ❖ Người thiết kế cơ sở dữ liệu (Database Designer)
- ❖ Người phân tích hệ thống và lập trình ứng dụng
- ❖ Người vận hành
- ❖ Người bảo trì

6. Hệ quản trị cơ sở dữ liệu

- Hệ quản trị cơ sở dữ liệu (DBMS – DataBase Management System) là hệ thống các phần mềm hỗ trợ tích cực cho người sử dụng phân tích, thiết kế và khai thác CSDL. Giúp người sử dụng định nghĩa, tạo ra, xử lý và duy trì CSDL.
- Các DBMS thông dụng: Visual FoxPro, Microsoft Access, SQL Server, DB2, Oracle ... hầu hết các DBMS hiện nay đều dựa trên mô hình quan hệ.

6. Hệ quản trị cơ sở dữ liệu



6. Hệ quản trị cơ sở dữ liệu

➤ Một DBMS phải có:

- Ngôn ngữ giao tiếp giữa người sử dụng và CSDL
- Từ điển dữ liệu (Data Dictionary)
- Có biện pháp bảo mật khi có yêu cầu
- Cơ chế giải quyết tranh chấp dữ liệu
- Tạo ra các cấu trúc phức tạp cho phép nhiều người sử dụng truy cập đến dữ liệu
- Có cơ chế sao lưu (backup), phục hồi (restore)
- Đảm bảo tính độc lập giữa dữ liệu và chương trình

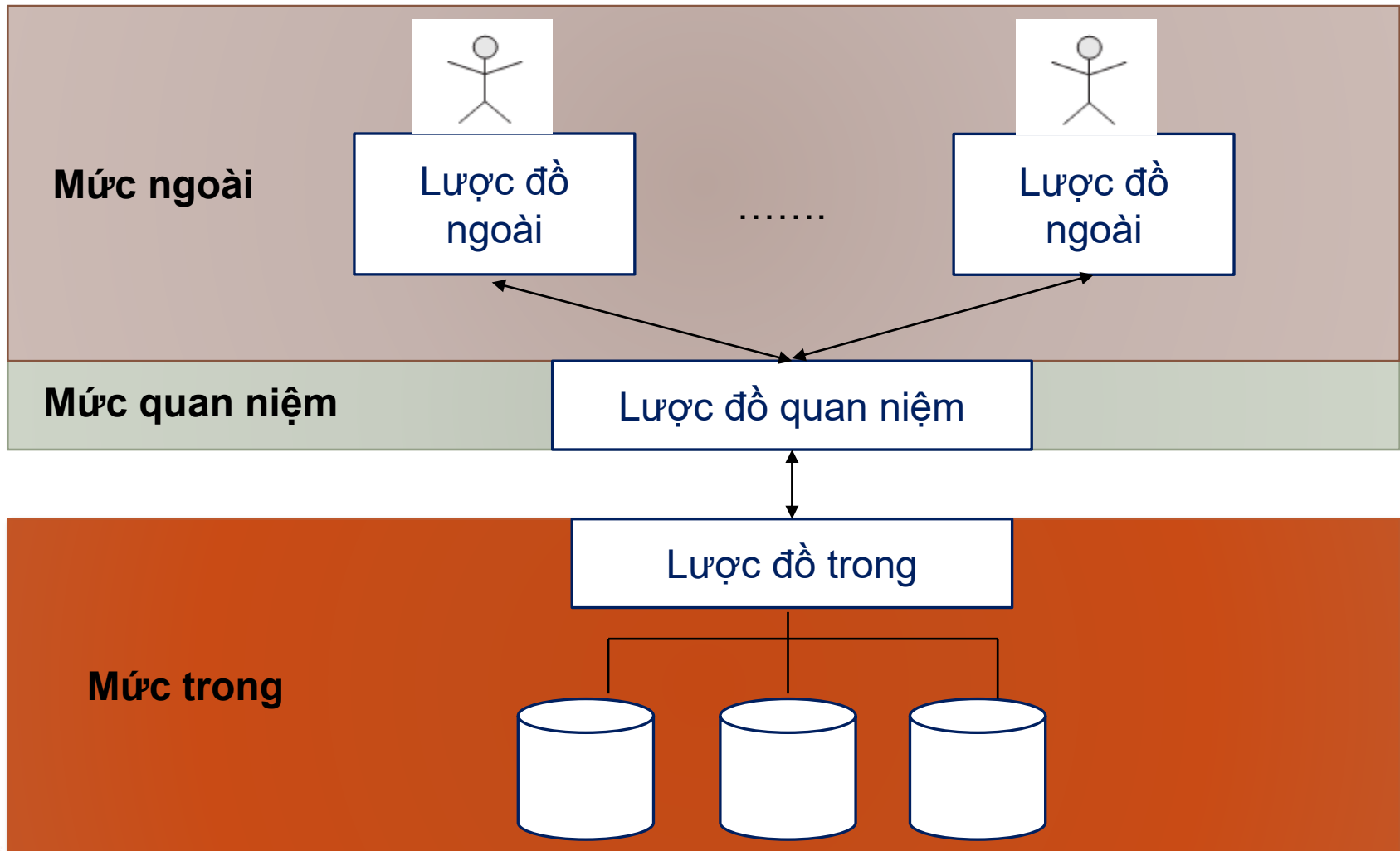
6. Hệ quản trị cơ sở dữ liệu

➤ **Ngôn ngữ giao tiếp:** được chia thành 04 loại

- **Ngôn ngữ mô tả dữ liệu** (DDL – Data Definition Language): cho phép khai báo cấu trúc CSDL, các mối liên hệ của dữ liệu, các quy định, ràng buộc dữ liệu.
- **Ngôn ngữ thao tác dữ liệu** (DML – Data Manipulation Language): cho phép thực hiện thao tác thêm, xóa, sửa dữ liệu.
- **Ngôn ngữ truy vấn có cấu trúc** (DQL – Data Query Language): cho phép người khai thác sử dụng để truy vấn thông tin cần thiết.
- **Ngôn ngữ quản lý dữ liệu** (DCL – Data Control Language) cho phép thay đổi cấu trúc bảng, khai báo bảo mật, cấp quyền cho người sử dụng.

7. Các mức biểu diễn của một CSDL

➤ Kiến trúc 3 lược đồ



7. Các mức biểu diễn của một CSDL

➤ Mức ngoài

- Còn gọi là mức khung nhìn (view)
- Mô tả một phần của CSDL mà 1 nhóm người dùng quan tâm đến và che giấu phần còn lại của CSDL đối với nhóm người dùng đó

➤ Mức quan niệm (mức Conception hay mức Logical)

- Mức nhìn tổng thể về CSDL, dành cho người quản trị CSDL
- Mô tả cấu trúc của toàn thể CSDL gồm thực thể, kiểu dữ liệu, mối liên hệ và ràng buộc

➤ Mức trong (mức vật lý – Physical)

- Mô tả cấu trúc lưu trữ vật lý CSDL: Dữ liệu gì? Lưu trữ thế nào? Ở đâu?...
- Góc nhìn dành cho người phát triển CSDL
- Là mức lưu trữ CSDL

7. Các mức biểu diễn của một CSDL

Ví dụ một ứng dụng TMĐT

External: cách nhìn cho từng nhóm người dùng - View “Báo cáo bán hàng tuần” (chỉ cột cần thiết, dữ liệu đã tổng hợp).
View “CSKH” (ẩn cột nhạy cảm, CCCD mã hóa).

Conceptual: lược đồ tổng thể - Bảng KHACHHANG, DONHANG, CTDH, SANPHAM, KHUYENMAI

Internal: cách lưu trữ vật lý – Index IX_DONHANG_Ngay, partition DONHANG theo Ngay, nonclustered index cho MaKH.

Tách biệt thay đổi. Thể hiện độc lập dữ liệu.

CÁC MÔ HÌNH DỮ LIỆU

1. Giới thiệu
2. Mô hình dữ liệu mạng
3. Mô hình dữ liệu phân cấp
4. Mô hình thực thể mối kết hợp
5. Mô hình dữ liệu quan hệ
6. Mô hình dữ liệu hướng đối tượng
7. Bài tập

1. Giới thiệu

- Mô hình dữ liệu là sự trừu tượng hóa của môi trường thực, biểu diễn dữ liệu ở mức quan niệm.
- Là một tập hợp các khái niệm dùng để biểu diễn các cấu trúc của cơ sở dữ liệu
- Một số mô hình dữ liệu:
 - Mô hình dữ liệu mạng
 - Mô hình dữ liệu phân cấp
 - Mô hình dữ liệu thực thể mối kết hợp
 - Mô hình dữ liệu quan hệ
 - Mô hình dữ liệu hướng đối tượng
- Mô tả của một cơ sở dữ liệu được gọi là lược đồ cơ sở dữ liệu

2. Mô hình dữ liệu mạng

- Mô hình dữ liệu mạng (Network Data Model) còn gọi tắt là mô hình mạng (Network Model) hoặc mô hình lưới là mô hình được biểu diễn bởi một đồ thị có hướng.
- Bao gồm các thành phần:
 - Mẫu tin (record)
 - Loại mẫu tin
 - Loại liên hệ (set type)
 - Bản số

2. Mô hình dữ liệu mạng

- **Mẫu tin:** mô tả 1 đối tượng trong thế giới thực.
(‘NV001’, ‘Nguyen Lam’, ‘Nam’, ‘10/10/1970’, ‘Dong Nai’)
- **Loại mẫu tin:** là 1 tập các mẫu tin có cùng tính chất. Ví dụ:

NHANVIEN

- Ký hiệu:

NHANVIEN

CONGVIEC

- **Loại liên hệ:** mô tả sự liên kết giữa 1 loại mẫu tin chủ và 1 loại mẫu tin thành viên

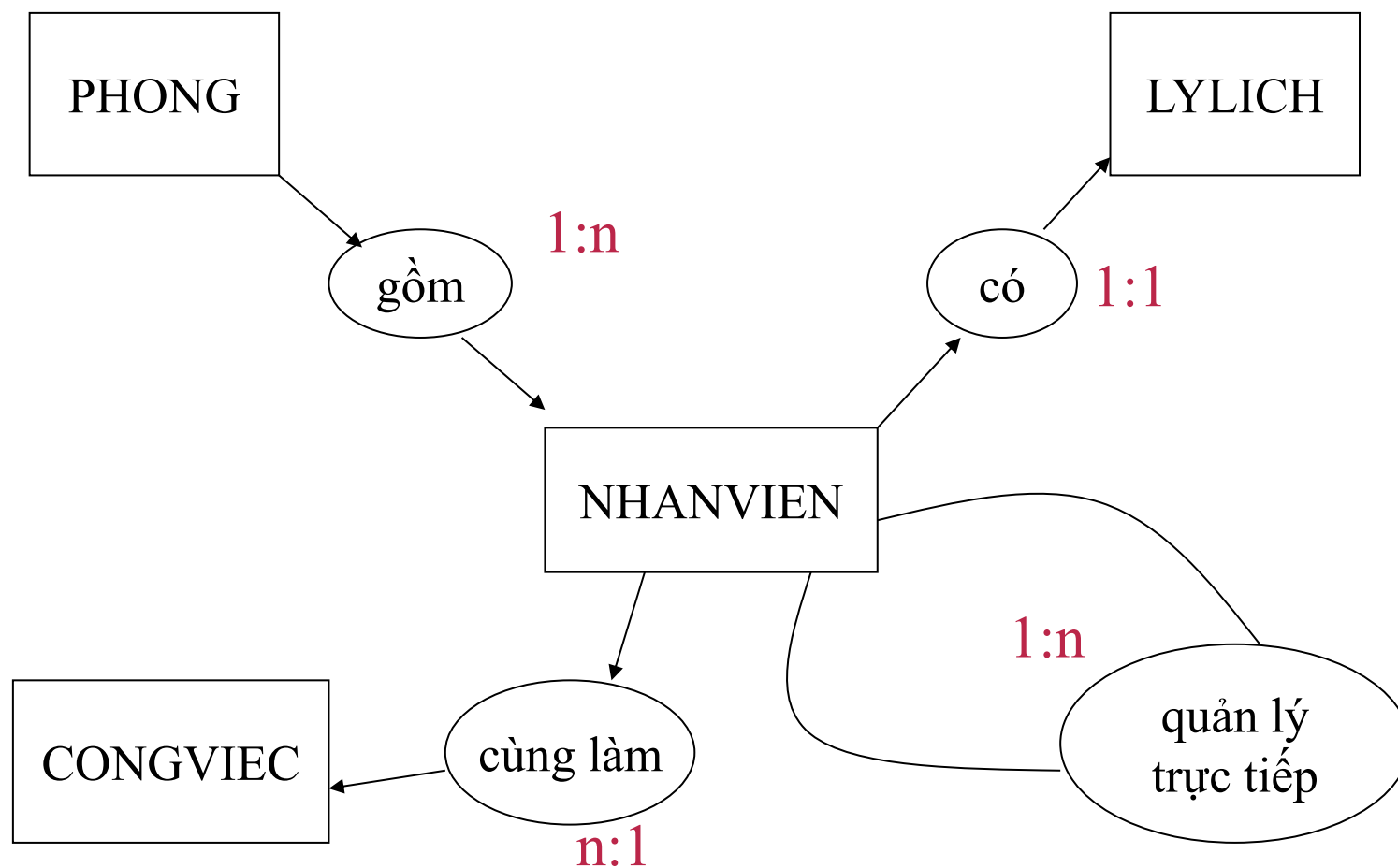
- Ký hiệu:

Tham gia

2. Mô hình dữ liệu mạng

- **(Recursive) - Độ quy:** một loại mẫu tin chủ cũng có thể đồng thời là loại mẫu tin thành viên với chính nó.
- **Bản số:** chỉ ra số lượng các mẫu tin tham gia trong mỗi liên hệ
 - **(1:1)** (one-to-one): mỗi mẫu tin của loại mẫu tin chủ kết hợp với đúng 1 mẫu tin của loại mẫu tin thành viên.
 - **(1:n)** (one-to-many): mỗi mẫu tin của loại mẫu tin chủ kết hợp với 1 hay nhiều mẫu tin thành viên.
 - **(n:1)** (many-to-one): nhiều mẫu tin của loại mẫu tin chủ kết hợp với đúng 1 mẫu tin của loại mẫu tin thành viên.

2. Mô hình dữ liệu mạng



2. Mô hình dữ liệu mạng

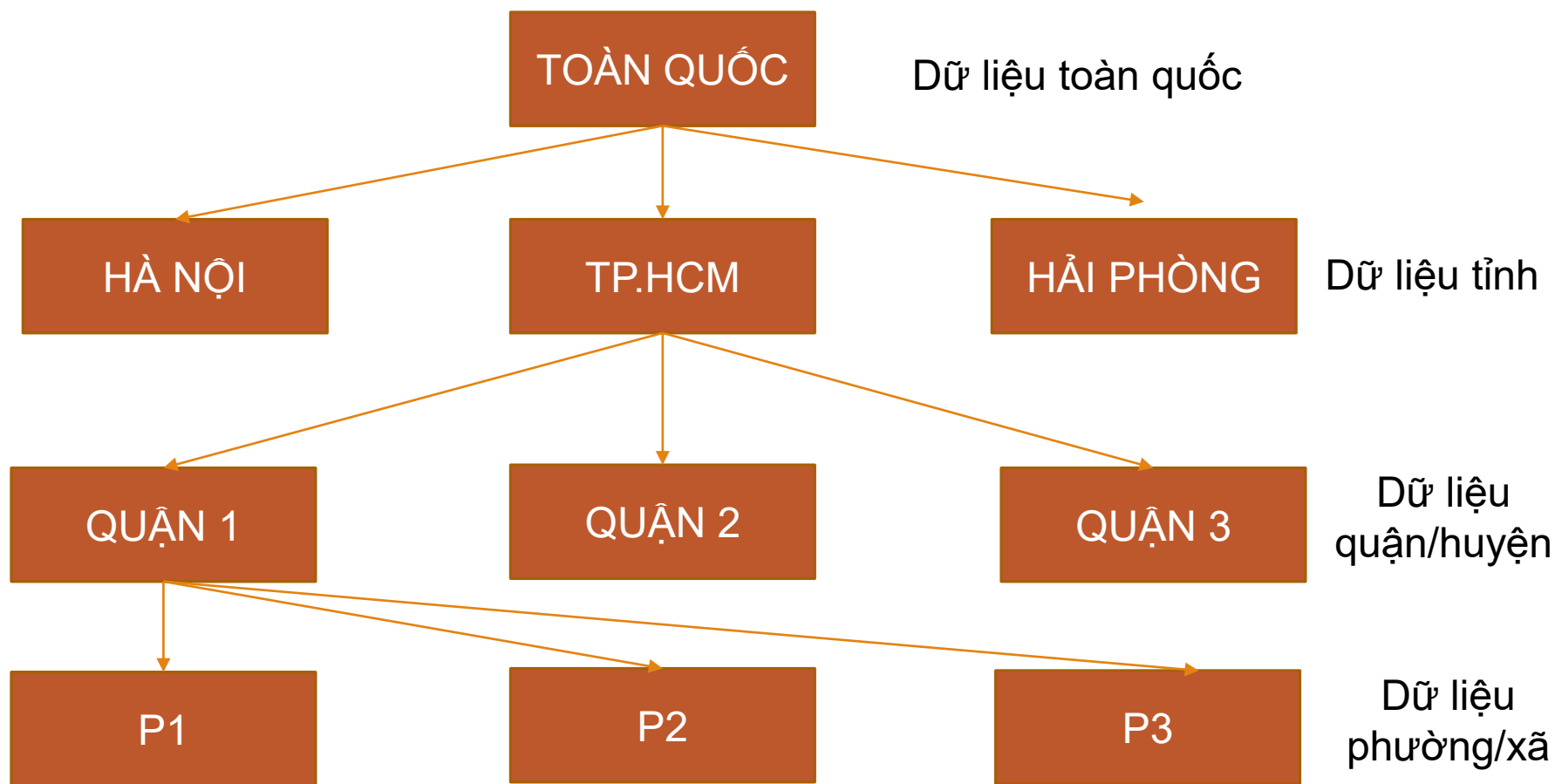
➤ Ưu điểm của mô hình dữ liệu mạng

- Tương đối đơn giản
- Dễ sử dụng
- Không thích hợp biểu diễn CSDL có quy mô lớn
- Khả năng diễn đạt ngữ nghĩa kém

3. Mô hình dữ liệu phân cấp

- Mô hình dữ liệu phân cấp (Hierarchical Data Model) còn gọi tắt là mô hình phân cấp (Hierarchical Model) dạng cây phân cấp.
 - Mỗi nút của cây biểu diễn một thực thể
 - Giữa nút con với nút cha được liên hệ với nhau theo một mối quan hệ xác định
- Một số khái niệm:
 - Loại mẫu tin: giống mô hình dữ liệu mạng
 - Loại mối liên hệ: phân cấp
 - Giữa hai loại mẫu tin **chỉ tồn tại một mối quan hệ duy nhất**.
 - Mối liên hệ từ mẫu tin chủ tới các mẫu tin thành viên là $1 \div n$
 - Mối liên hệ từ mẫu tin thành viên với mẫu tin chủ là $1 \div 1$

3. Mô hình dữ liệu phân cấp



3. Mô hình dữ liệu phân cấp

Ví dụ: hệ thống quản lý thư viện: Mỗi thư viện có thể chứa nhiều kệ sách, mỗi kệ sách có nhiều loại sách khác nhau:

- Thực thể "Library": Gồm các thuộc tính như LibraryID, LibraryName.
- Thực thể "Shelf": Gồm các thuộc tính như ShelfID, Location.
- Thực thể "Book": Gồm các thuộc tính như BookID, Title, Author.
- Mỗi quan hệ phân cấp: Library -> Shelf -> Book.

3. Mô hình dữ liệu phân cấp

➤ Nhận xét

- Đơn giản, tìm kiếm nhanh
- Khả năng diễn đạt ngữ nghĩa kém (các mối quan hệ n-n)
- Không thích hợp biểu diễn CSDL có quy mô lớn

4. Mô hình thực thể mối kết hợp

4.1 Giới thiệu

- Mô hình thực thể mối kết hợp (**Entity-Relationship Model** - viết tắt ER – hoặc gọi là mô hình thực thể mối liên kết) được Peter Pin Shan CHEN giới thiệu năm 1976.
- Mô hình ER được sử dụng nhiều trong thiết kế dữ liệu ở mức **quan niệm**.

4. Mô hình thực thể mối kết hợp

- Nhìn thế giới thực như là một tập các đối tượng căn bản được gọi là **các tập thực thể**, và **các mối quan hệ** (mối liên kết) giữa các đối tượng này.
- Mục đích: **mô tả thế giới thực** bằng các **thực thể – thuộc tính – mối kết hợp**, trước khi chuyển thành bảng.
- Giúp người thiết kế cơ sở dữ liệu mô tả thế giới thực gần gũi với quan niệm và cách nhìn nhận bình thường của con người.
- Là công cụ để phân tích thông tin nghiệp vụ.

4. Mô hình thực thể mối kết hợp

4.2 Loại thực thể (tập thực thể)

- Loại thực thể (Entity Type) là những loại đối tượng hay sự vật của thế giới thực tồn tại cụ thể cần được quản lý. Là tập hợp các thực thể có cùng thuộc tính.
- Ví dụ : HOCVIEN, LOP, MONHOC, ...
- Ký hiệu:

HOCVIEN

LOP

4. Mô hình thực thể mối kết hợp

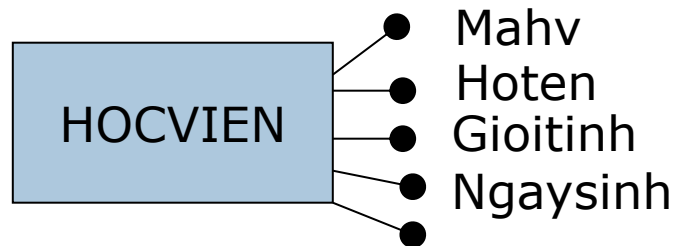
4.3 Thực thể

- Thực thể (Entity) là một thể hiện (hoặc 1 đối tượng) của một loại thực thể.
- Một thực thể (Entity) là một đối tượng của thế giới thực
- Ví dụ: Loại thực thể là HOCVIEN có các thực thể:
 - ('HV001', 'Luc Van Tien', '1/2/1987', 'Nam')
 - ('HV002', 'Thuy Kieu', '13/2/1993', 'Nu')
 - ('HV003', 'Thi Mau', '11/8/1991', 'Nu')

4. Mô hình thực thể mối kết hợp

4.4 Thuộc tính

- Thuộc tính (Attribute) là những tính chất đặc trưng của loại thực thể.
- Ví dụ: Loại thực thể HOCVIEN có các thuộc tính: Mã học viên, họ tên, giới tính, ngày sinh, nơi sinh
- Ký hiệu:



4. Mô hình thực thể mối kết hợp

4.4 Các loại thuộc tính

- **Đơn trị (Simple):** mỗi thực thể chỉ có một giá trị ứng với mỗi thuộc tính.
 - Ví dụ: Mahv,Hoten
- **Đa hợp (Composite):** thuộc tính có thể được tạo thành từ nhiều thành phần
 - Ví dụ: DCHI(SONHA,DUONG,PHUONG,QUAN) hay thuộc tính HOTEN(HO,TENLOT,TEN).
- **Đa trị (Multi-valued):** thuộc tính có thể có nhiều giá trị đối với một thực thể.
 - Ví dụ: BANGCAP ký hiệu {BANGCAP}

4. Mô hình thực thể mối kết hợp

4.4 Các loại thuộc tính

- Các thuộc tính đa hợp và đa trị có thể lồng nhau tùy ý.
 - Ví dụ: thuộc tính BANGCAP của HOCVIEN là một thuộc tính đa hợp được ký hiệu bằng:
{BANGCAP(TRUONGCAP,NAM,KETQUA, CHUYENNGANH)}
- Thuộc tính suy diễn được (thuộc tính dẫn xuất): Thuộc tính có giá trị có thể tính được thông qua giá trị của các thuộc tính khác
Ví dụ: năm sinh và tuổi.
- Trong một số trường hợp, một số thuộc tính của một thực thể cụ thể không xác định được giá trị. Trong trường hợp như vậy, ta phải tạo ra một giá trị đặc biệt gọi là giá trị null.

4. Mô hình thực thể mối kết hợp

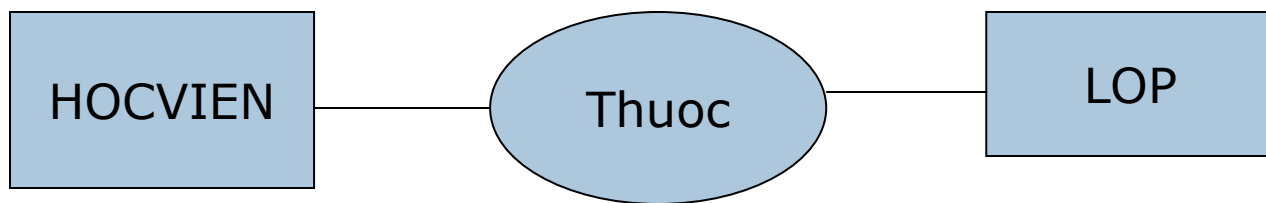
4.5 Khóa của loại thực thể (entity type key)

- Khóa của loại thực thể là **thuộc tính nhận diện thực thể**.
- Căn cứ vào giá trị của khóa có thể xác định **duy nhất** một thực thể.
 - Ví dụ: Mỗi học viên có một mã số duy nhất => Khoá của loại thực thể HOCVIEN là Mã học viên
- Lưu ý
 - Mỗi loại thực thể phải có ít nhất 1 khóa
 - Một khóa có thể có 1 hay nhiều thuộc tính
 - Có thể có nhiều khóa trong 1 loại thực thể, ta sẽ chọn ra 1 khóa làm khóa chính cho tập thực thể đó

4. Mô hình thực thể mối kết hợp

4.6 Loại mối kết hợp

- Loại mối kết hợp (Relationship type) là sự liên kết giữa hai hay nhiều loại thực thể
- Ví dụ: giữa hai loại thực thể HOCVIEN và LOP có loại mối kết hợp THUOC
- Ký hiệu: bằng một hình oval hoặc hình thoi



4. Mô hình thực thể mối kết hợp

Chú ý:

Cung là đường nối giữa

- Tập thực thể và thuộc tính
- Mối quan hệ và tập thực thể

4. Mô hình thực thể mối kết hợp

Ví dụ: một hệ thống quản lý khách hàng (Customer Relationship Management - CRM) cần quản lý các thực thể cơ bản nào?

4. Mô hình thực thể mối kết hợp

Ví dụ: một hệ thống quản lý khách hàng (Customer Relationship Management - CRM) cần quản lý các thực thể cơ bản nào?

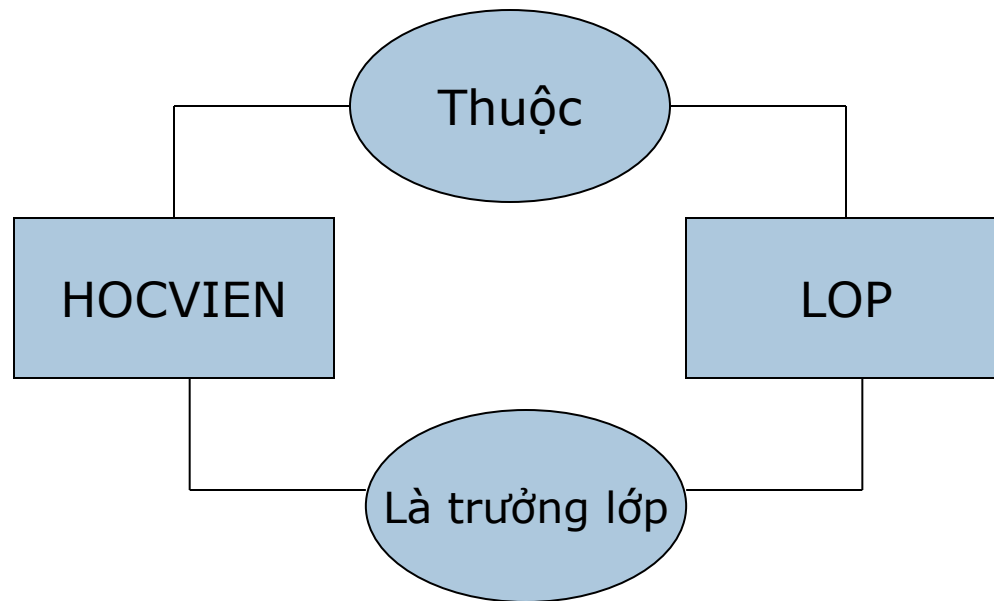
- Thực thể "Customer": Gồm các thuộc tính như CustomerID, Name, Email, Phone.
- Thực thể "Order": Gồm các thuộc tính như OrderID, OrderDate, TotalAmount.
- Mối kết hợp "Places Order": Giữa Customer và Order, trong đó một khách hàng có thể **đặt nhiều** đơn hàng. (?)



4. Mô hình thực thể mối kết hợp

4.6 Loại mối kết hợp

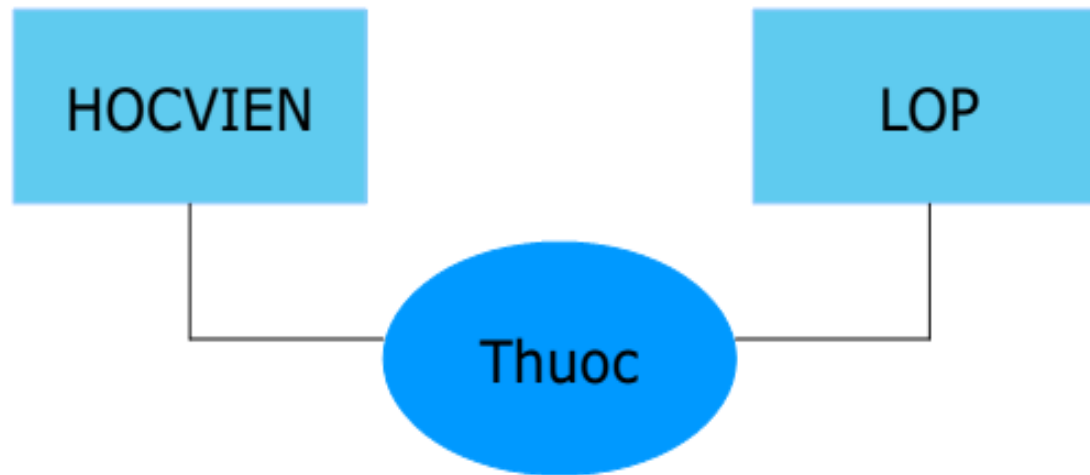
- Giữa hai loại thực thể có thể tồn tại nhiều hơn một loại mối kết hợp.
- Ví dụ:



4. Mô hình thực thể mối kết hợp

4.7 Số ngôi của loại mối kết hợp

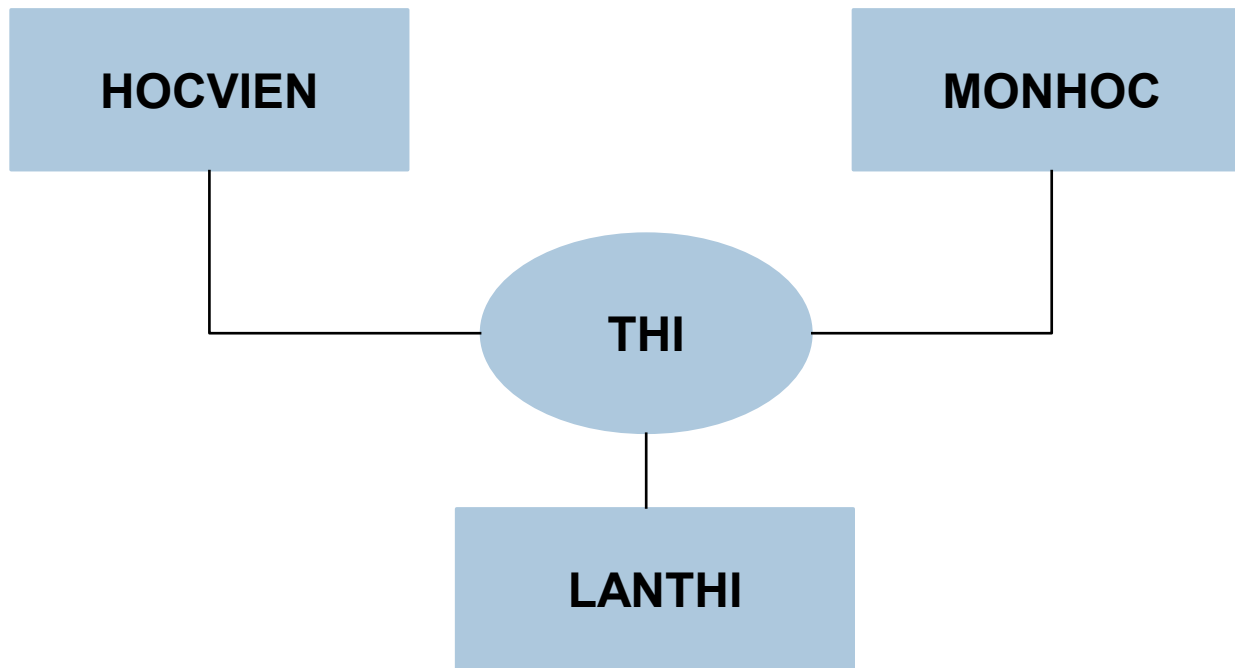
- Số ngôi của loại mối kết hợp (Relationship degree) là **số loại thực thể tham gia vào loại mối kết hợp** đó.
- Ví dụ 1: Loại mối kết hợp Thuộc kết hợp 2 loại thực thể HOCVIEN và LOP nên có số ngôi là 2.



4. Mô hình thực thể mối kết hợp

4.7 Số ngôi của loại mối kết hợp

- Ví dụ 2: Loại mối kết hợp Thi kết hợp 3 loại thực thể LANTHI, HOCVIEN, MONHOC nên có số ngôi là 3.



4. Mô hình thực thể mối kết hợp

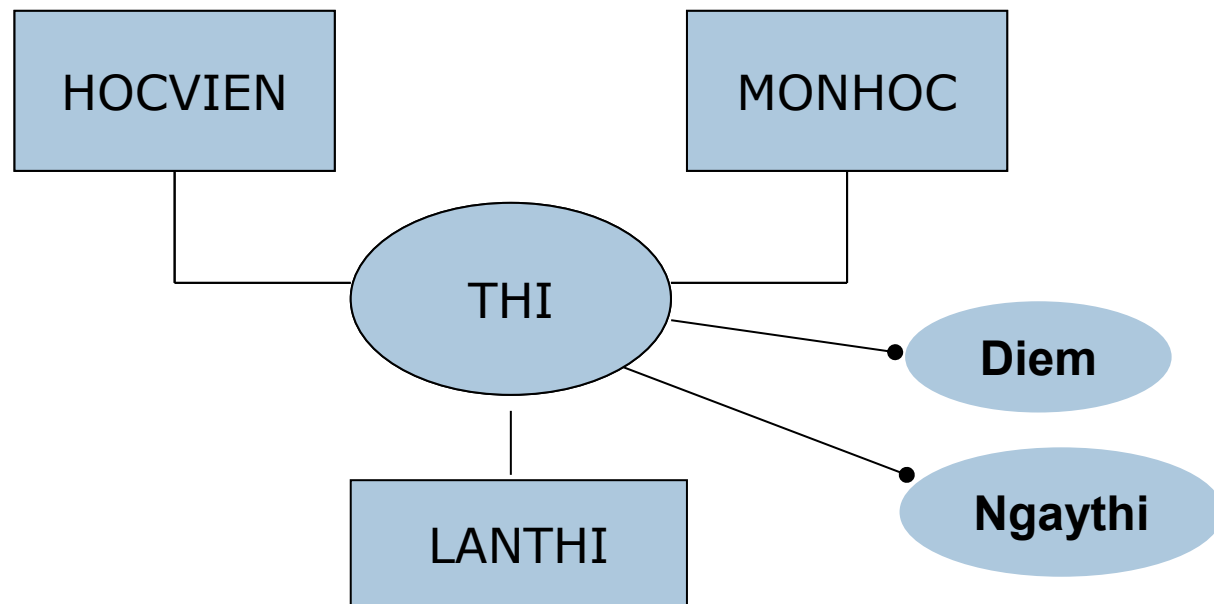
4.8 Thuộc tính của loại mối kết hợp (relationship type attribute)

- Mô tả tính chất của mối kết hợp đó
- Bao gồm các thuộc tính khoá của các loại thực thể tham gia vào loại mối kết hợp đó. Ngoài ra còn có thể có những thuộc tính bổ sung khác.
- Không thể gắn với các thực thể tham gia vào mối kết hợp đó

4. Mô hình thực thể mối kết hợp

4.8 Thuộc tính của loại mối kết hợp (relationship type attribute)

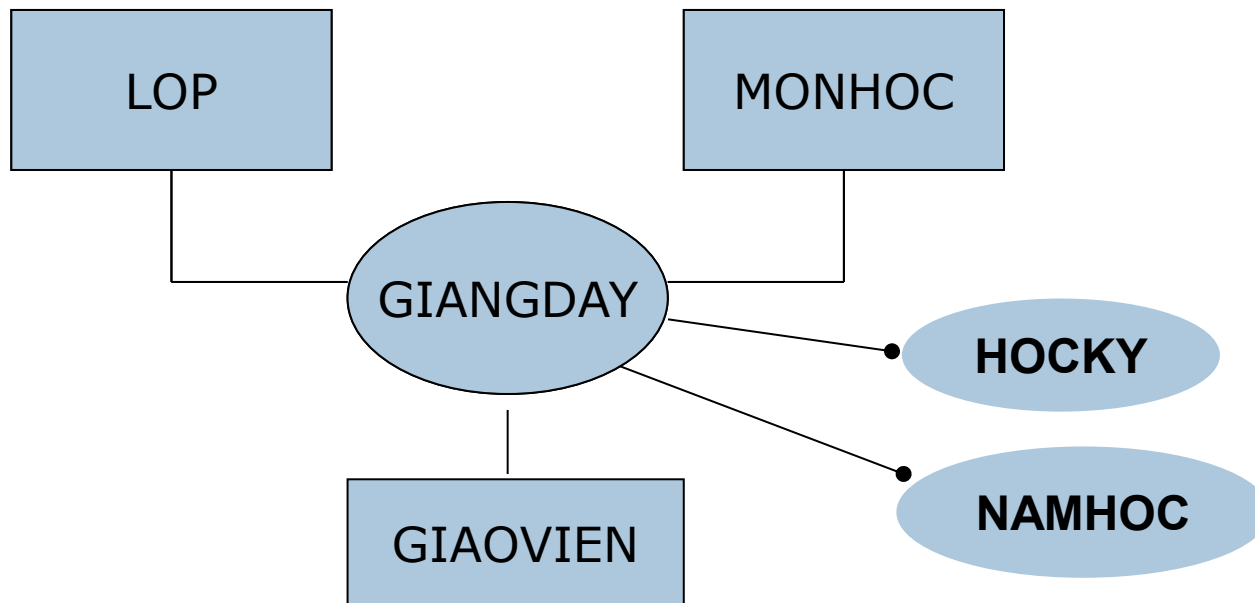
- Ví dụ 1: Loại mối kết hợp Thi giữa ba loại thực thể HOCVIEN, MONHOC và LANTHI có các thuộc tính là Mahv, Mamh, Lanthi, ngoài ra còn có thuộc tính riêng là Diem, Ngaythi



4. Mô hình thực thể mối kết hợp

4.8 Thuộc tính của loại mối kết hợp (relationship type attribute)

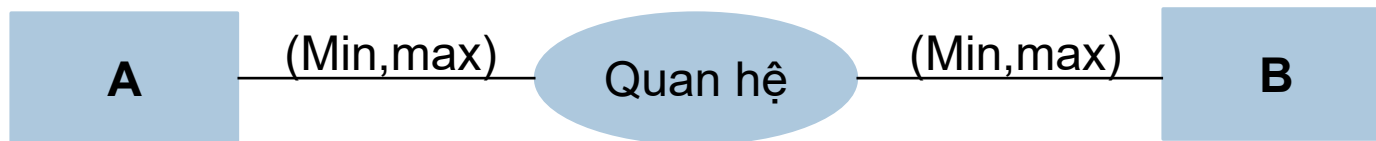
- Ví dụ 2: Loại mối kết hợp Giảng dạy giữa ba loại thực thể GIAOVIEN, MONHOC và LOP có các thuộc tính là Magv, Mamh, Malop. Ngoài ra còn có 2 thuộc tính riêng là HOCKY và NAMHOC



4. Mô hình thực thể mối kết hợp

4.9 Bản số (relationship cardinality)

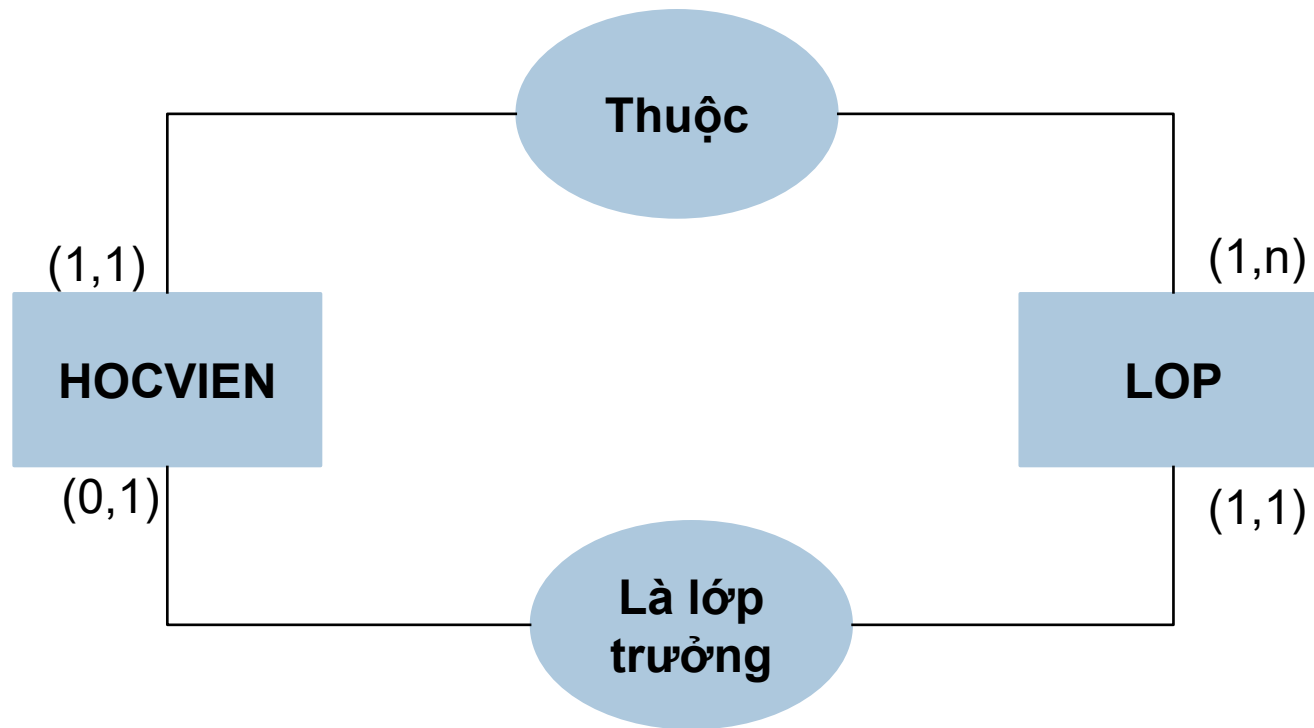
- Loại mối kết hợp thể hiện liên kết giữa các thực thể, mỗi liên kết được gọi là một nhánh
- Bản số của nhánh là **số lượng tối thiểu** và **số lượng tối đa** các thể hiện của loại mối kết hợp mà 1 thực thể thuộc nhánh đó tham gia
- Ký hiệu: (số lượng tối thiểu, số lượng tối đa)
 - $(0,1)$ – không hoặc 1
 - $(1,1)$ – duy nhất 1
 - $(0,n)$ – không hoặc nhiều
 - $(1,n)$ – một hoặc nhiều



4. Mô hình thực thể mối kết hợp

4.9 Bản số (relationship cardinality)

- Ví dụ: Loại thực thể HOCVIEN và LOP có loại mối kết hợp Thuộc và Là lớp trưởng.



4. Mô hình thực thể mối kết hợp

4.10 Mối kết hợp đệ quy

- Là loại mối kết hợp được tạo thành từ cùng một loại thực thể (hay một loại thực thể có loại mối kết hợp với chính nó)
- Ví dụ: Mỗi nhân viên có một người quản lý trực tiếp và người quản lý đó cũng là một nhân viên



4. Mô hình thực thể mối kết hợp

4.11 Loại thực thể yếu

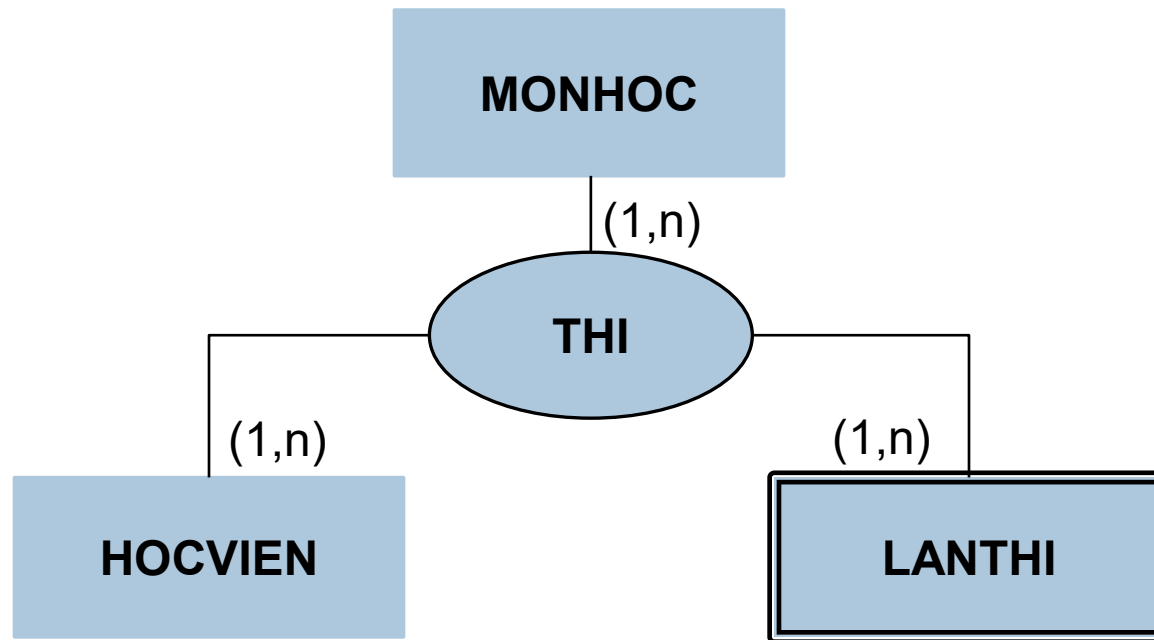
- Là loại thực thể không có thuộc tính khóa
- Khóa có được từ những thuộc tính của tập thực thể khác
- Phải tham gia trong một loại mối kết hợp xác định trong đó có một loại thực thể chủ.
- Ký hiệu

Loại thực thể

4. Mô hình thực thể mối kết hợp

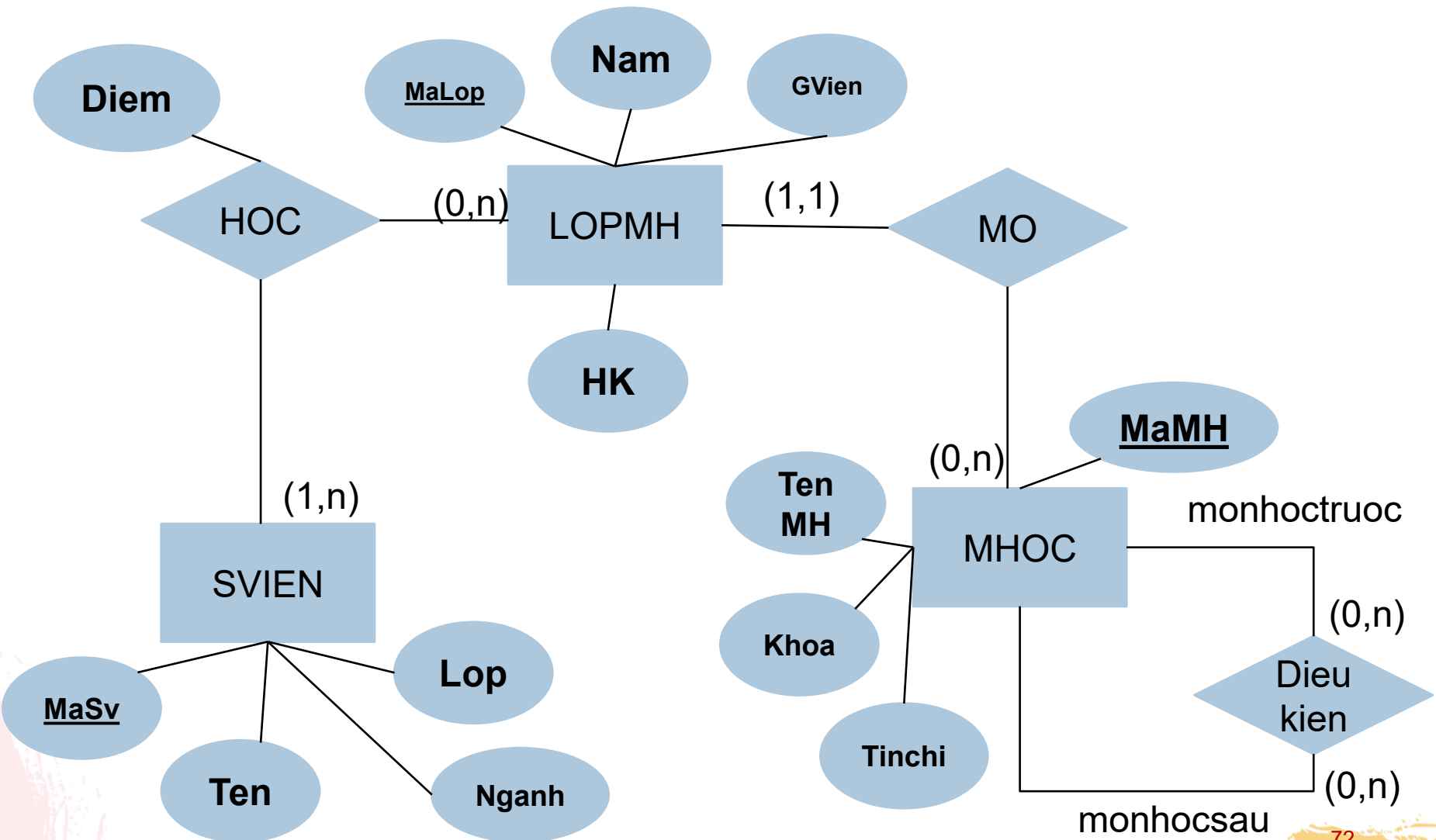
4.11 Loại thực thể yếu

- Ví dụ: loại thực thể LANTHI có thuộc tính Làn và tham gia trong loại mối kết hợp Thi với loại thực thể HOCVIEN và MONHOC là loại thực thể yếu



4. Mô hình thực thể mối kết hợp

Ví dụ mô hình ER



4. Mô hình thực thể mối kết hợp

➤ Các bước thiết kế:

- Xác định loại thực thể
- Xác định loại mối kết hợp
- Xác định Thuộc tính của loại thực thể, loại mối kết hợp
- Xác định bản số (1-1, 1-n, n-n)

5. Mô hình dữ liệu quan hệ

KHOA (MaKhoa, TenKhoa, SoLuongGV)

SINHVIEN(MaSV, TenSV, Lop, Nganh)

MONHOC (MaMH, TenMH, Khoa, TinChi)

DIEUKIEN(MaMH, MaMH_Truoc)

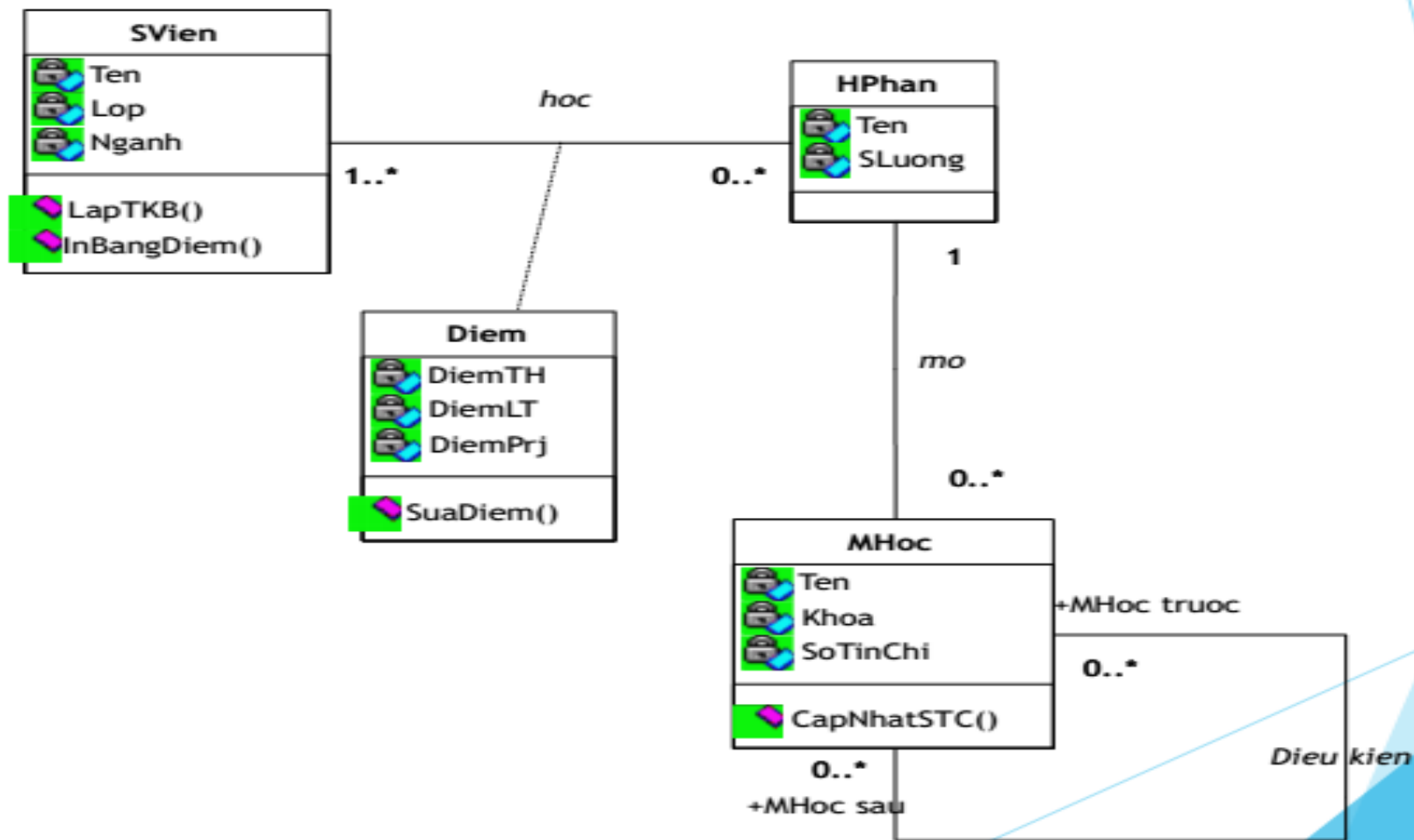
DANGKY (MaSV, MaMH, HocKy, Namhoc)

Qua chương 2 sẽ tìm hiểu rõ hơn về Mô hình dữ liệu quan hệ và cách chuyển từ Mô hình Thực thể-mối kết hợp sang Mô hình dữ liệu quan hệ.

6. Mô hình dữ liệu hướng đối tượng

- Dựa trên cách tiếp cận hướng đối tượng.
- Các khái niệm: Lớp (class), Thuộc tính (Attribute), Phương thức (Operator), ...
- Đặc trưng
 - Tính kế thừa (inheritance)
 - Tính đóng gói (encapsulation)
 - Tính đa hình (polymorphism)
 - Tính tái sử dụng (reusability)

6. Mô hình dữ liệu hướng đối tượng



Bổ sung: Các xu hướng hiện đại

- **NoSQL** là một loại cơ sở dữ liệu phi quan hệ, được thiết kế để xử lý dữ liệu phi cấu trúc hoặc bán cấu trúc, và phù hợp cho các ứng dụng yêu cầu mở rộng quy mô và hiệu suất cao.

Bổ sung: Các xu hướng hiện đại

➤ Các loại cơ sở dữ liệu NoSQL:

- **Document Store** (Cơ sở dữ liệu tài liệu): Ví dụ: MongoDB, nơi dữ liệu được lưu trữ dưới dạng tài liệu JSON hoặc BSON. Các tài liệu này có thể có cấu trúc linh hoạt, cho phép lưu trữ các đối tượng phức tạp.
- **Key-Value Store** (Cơ sở dữ liệu khóa-giá trị): Ví dụ: Redis, nơi dữ liệu được lưu trữ dưới dạng các cặp khóa-giá trị. Loại này phù hợp cho các ứng dụng cần truy xuất nhanh chóng, như bộ nhớ đệm (cache).
- **Column-Family Store** (Cơ sở dữ liệu cột gia đình): Ví dụ: Cassandra, nơi dữ liệu được lưu trữ theo các cột hơn là theo hàng. Điều này cho phép truy vấn dữ liệu theo cột nhanh chóng, phù hợp cho phân tích dữ liệu lớn.
- **Graph Database** (Cơ sở dữ liệu đồ thị): Ví dụ: Neo4j, nơi dữ liệu được lưu trữ dưới dạng các đỉnh và cạnh, phù hợp cho việc xử lý dữ liệu liên kết phức tạp như mạng xã hội, hệ thống gợi ý.

Bổ sung: Các xu hướng hiện đại

- **Ưu điểm:** Mở rộng dễ dàng (cả chiều ngang và chiều dọc), hiệu suất cao, xử lý dữ liệu phi cấu trúc hoặc bán cấu trúc.
- **Nhược điểm:** Thiếu hỗ trợ đầy đủ cho ACID (tùy thuộc vào loại NoSQL), có thể phức tạp hơn trong việc quản lý tính toàn vẹn dữ liệu.
- Ví dụ:
 - Facebook: Sử dụng các hệ thống NoSQL như Apache Cassandra để quản lý và phân tích hàng tỷ dữ liệu người dùng mỗi ngày, từ bài đăng, bình luận đến tin nhắn.
 - Twitter: Sử dụng NoSQL để quản lý hàng triệu tweet mỗi phút, nơi dữ liệu cần được xử lý và truy vấn với tốc độ cao trong thời gian thực.

Bổ sung: Các xu hướng hiện đại

Big Data: khái niệm mô tả việc xử lý và quản lý lượng dữ liệu rất lớn, đa dạng, và tốc độ xử lý nhanh chóng. Big Data thường đi kèm với khái niệm 3V: Volume (Khối lượng), Variety (Đa dạng), và Velocity (Tốc độ).

Bổ sung: Các xu hướng hiện đại

Hệ sinh thái Hadoop:

- HDFS (Hadoop Distributed File System): Hệ thống tệp phân tán cho phép lưu trữ dữ liệu lớn trên nhiều máy chủ khác nhau. HDFS chia nhỏ các tệp tin lớn thành các khối và phân phối chúng trên nhiều nút để đảm bảo khả năng phục hồi và truy cập nhanh chóng.
- MapReduce: Một mô hình lập trình phân tán dùng để xử lý dữ liệu lớn. MapReduce chia dữ liệu thành các phần nhỏ và xử lý chúng song song, sau đó kết hợp kết quả lại với nhau.
- YARN: Hệ điều hành của Hadoop, quản lý tài nguyên và phân phối nhiệm vụ cho các ứng dụng chạy trên hệ sinh thái Hadoop.
- HBase, Hive, Pig: Các công cụ và ngôn ngữ khác trong hệ sinh thái Hadoop hỗ trợ việc quản lý và truy vấn dữ liệu lớn với cú pháp và quy trình gần giống với SQL.

Bổ sung: Các xu hướng hiện đại

Ví dụ :

- **Netflix:** Sử dụng Big Data và Hadoop để thu thập và phân tích dữ liệu từ hàng triệu người dùng trên toàn cầu. Dữ liệu này bao gồm các hành vi xem phim, đánh giá, và lịch sử tìm kiếm, giúp Netflix gợi ý các nội dung phù hợp cho từng người dùng.
- **Amazon:** Sử dụng Big Data để phân tích các giao dịch, hành vi mua sắm của khách hàng, từ đó tối ưu hóa trải nghiệm mua sắm và cung cấp các đề xuất sản phẩm phù hợp.

7. Bài tập

Bài tập 1

❖ **Xây dựng mô hình ER cho CSDL quản lý giáo vụ gồm có các chức năng sau:**

- Lưu trữ thông tin: Học viên (có mã học viên duy nhất, họ tên, giới tính, NTNS) , giáo viên (mã giáo viên duy nhất, họ tên, giới tính), môn học (mã môn học duy nhất, tên môn học, số tín chỉ lý thuyết, số tín chỉ thực hành), lớp học (mã lớp duy nhất, tên lớp, sĩ số dự kiến).
- Xếp lớp cho học viên, chọn lớp trưởng cho lớp
- Phân công giảng dạy: giáo viên dạy lớp nào với môn học gì, ở học kỳ, năm học nào.
- Lưu trữ kết quả thi: học viên thi môn học nào, lần thi thứ mấy, điểm thi bao nhiêu.

7. Bài tập

Bài tập 2

- ❖ CSDL đề án công ty theo dõi các thông tin liên quan đến nhân viên, phòng ban và đề án
- ❖ Công ty có nhiều phòng ban, mỗi phòng ban có tên duy nhất, mã phòng duy nhất, một trưởng phòng và ngày nhận chức. Mỗi phòng ban có thể ở nhiều địa điểm khác nhau.
- ❖ Đề án có tên duy nhất, mã duy nhất, do 1 một phòng ban chủ trì và được triển khai ở 1 địa điểm.
- ❖ Nhân viên có mã số, tên, địa chỉ, ngày sinh, phái và lương. Mỗi nhân viên làm việc ở 1 phòng ban, tham gia vào các đề án với số giờ làm việc khác nhau. Mỗi nhân viên đều có một người quản lý trực tiếp.

7. Bài tập

Bài tập 3: Một ngân hàng có nhu cầu quản lý như sau :

- Mỗi tài khoản tiền gửi có số tài khoản, số dư trong tài khoản, ngày mở tài khoản.
- Mỗi khách hàng có thể mở một hoặc nhiều tài khoản tiền gửi. Có 2 loại khách hàng: khách hàng cá nhân (bao gồm phân loại khách hàng thường và khách hàng VIP) và khách hàng tổ chức (bao gồm loại hình khách hàng doanh nghiệp và khách hàng không phải doanh nghiệp). Mỗi khách hàng cá nhân có mã khách hàng, họ tên khách hàng, ngày tháng năm sinh, CCCD, giới tính, số điện thoại, địa chỉ. Mỗi khách hàng tổ chức có mã khách hàng, tên tổ chức, mã số thuế, người đại diện, CCCD người đại diện, số điện thoại trụ sở, địa chỉ trụ sở.
- Hai tài khoản tiền gửi khác nhau có thể thực hiện giao dịch chuyển khoản cho nhau, mỗi giao dịch sẽ có mã giao dịch, thời gian giao dịch, số tiền giao dịch.
- Mỗi tài khoản tiền gửi sẽ có thể có một (hoặc không) thẻ kèm theo, thẻ có số thẻ (khác số tài khoản), ngày phát hành thẻ, thời gian hiệu lực, loại thẻ. Thẻ phải đi cùng tài khoản tiền gửi.

Tài liệu tham khảo

1. Slides môn Cơ sở dữ liệu, Khoa Hệ thống Thông tin, Trường Đại học Công nghệ Thông tin, ĐHQG Tp. HCM.
2. ThS. Nguyễn Thị Kim Phụng, Slides bài giảng Cơ sở dữ liệu, Khoa HTTT, Trường Đại học CNTT
3. ThS. Nguyễn Hải Châu, Slides bài giảng Cơ sở dữ liệu, Đại học Công nghệ, ĐH Quốc gia Hà Nội
4. Ramez Elmasri, Shamkant B. Navathe, Fundamentals of Database Systems, Seven Edition, 2016



THANK YOU!

Q & A

ThS. TẠ VIỆT PHƯƠNG
phuongtv@uit.edu.vn