

# Mô Hình Hóa

# Dữ Liệu

## Data Modeling – DM



Người trình bày:



Đoàn Thị Kim Nhung



ITBA Team

# Mục Lục

01

Tổng quan  
Overview

02

Mô hình hóa dữ liệu  
Data Modeling

03

Chuẩn hóa CSDL quan hệ  
Relational Database Normalization



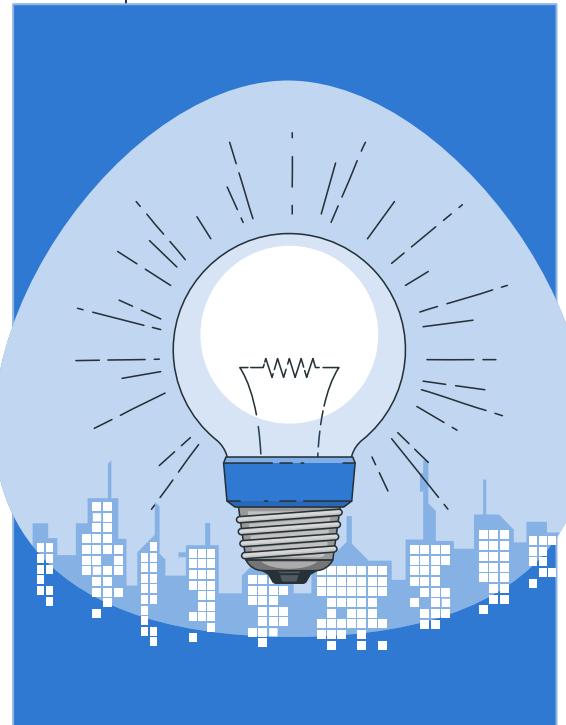
BATIZENS



# Từ ngữ viết tắt và định nghĩa

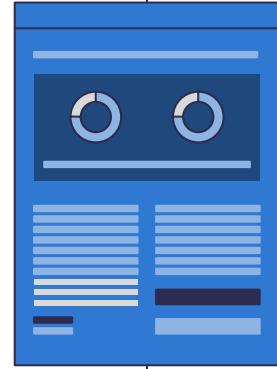
#	Từ ngữ viết tắt	Định nghĩa
1	PM	Phần mềm
2	CLPM	Chất lượng phần mềm
3		
4		
5		
6		
7		
8		





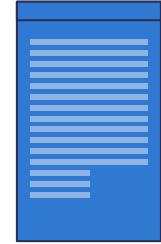
01

# Tổng quan





# Tổng quan



**1. Dữ liệu**

Các mục  
tổng quan

**3. Siêu dữ liệu**

**2. Thông tin**

**4. CSDL**



TQ

# 1. Dữ liệu (Data)

Dữ liệu

Thông tin

Siêu dữ liệu

Cơ sở dữ liệu



## 1. Khái niệm:

Dữ liệu là các dạng của **dấu hiệu, sự kiện, hoặc tác động mà không có ý nghĩa** tức thì.

## 2. Ví dụ:



Chữ cái



Hình ảnh



Nhiệt độ



Chữ số



Bit



Âm thanh





TQ

# 1. Dữ liệu (Data)



Dữ liệu

Thông tin

Siêu dữ liệu

Cơ sở dữ liệu

## 3. Dữ liệu trong máy tính:

Máy tính có thể lưu trữ các dữ liệu như video, hình ảnh, âm thanh, văn bản, tập tin,... và được lưu dưới dạng nhị phân bằng số 1 và 0.





## Dữ liệu

Thông tin

Siêu dữ liệu

Cơ sở dữ liệu

## 4. Dữ liệu cấu trúc (Structured Data)

**Khái niệm:** Dữ liệu có cấu trúc là **dữ liệu được tổ chức và định dạng** một cách **cụ thể** giúp **dễ đọc** và **dễ hiểu** cho cả **người và máy**. Điều này thường được thực hiện thông qua việc sử dụng một lược đồ hoặc mô hình dữ liệu (như SQL) để định nghĩa rõ ràng, cung cấp cấu trúc cho dữ liệu.

**Ví dụ:** Bảng dữ liệu Khách hàng gồm các 5 cột ID, Họ tên, Địa chỉ, Ngày sinh, Số điện thoại

#	ID	Họ tên	Địa chỉ	Ngày sinh	Số điện thoại
1	001	Nguyễn Văn A	123 Đường ABC, Quận 1, HCM	01-01-1990	0912345678
2	002	Trần Thị B	456 Đường DEF, Quận 2, HCM	12-05-1985	0934567890
3	003	Lê Hoàng C	789 Đường GHI, Quận 3, HCM	20-11-1992	0901234567
4	004	Phạm Văn D	321 Đường JKL, Quận 4, HCM	15-07-1988	0923456789
5	005	Hoàng Thị E	654 Đường MNO, Quận 5, HCM	22-09-1995	0945678901
6	006	Võ Thanh F	987 Đường PQR, Quận 6, HCM	30-03-1983	0956789012

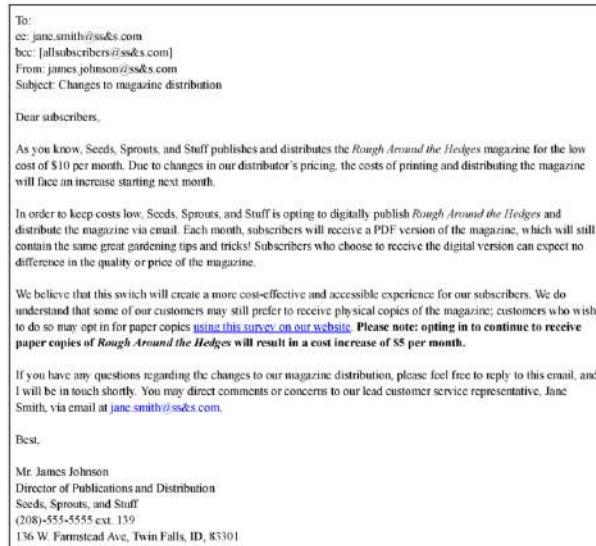




## 5. Dữ liệu phi cấu trúc (Un-Structured Data)

**Khái niệm:** Dữ liệu phi cấu trúc là **dữ liệu không có định dạng cố định** hoặc **không tuân theo một cấu trúc rõ ràng** như dữ liệu cấu trúc. Dữ liệu phi cấu trúc có thể tồn tại dưới nhiều dạng như văn bản, hình ảnh, âm thanh, video, email, tin nhắn, mạng xã hội, và các tập tin khác mà không có cấu trúc rõ ràng để máy tính dễ dàng xử lý.

## Ví dụ:



## Nội dung Email



## Bài đăng MXH



## Hình ảnh/Video/Audio



## 6. Dữ liệu bán cấu trúc (Semi-Structured Data)

**Khái niệm:** Dữ liệu bán cấu trúc là loại **dữ liệu có một số yếu tố cấu trúc**, nhưng **không hoàn toàn theo cấu trúc chặt chẽ như dữ liệu cấu trúc**. Dữ liệu bán cấu trúc có thể có một số dạng tổ chức, nhưng không nhất thiết phải tuân theo một khuôn mẫu hoặc bảng cố định. Điều này giúp dữ liệu bán cấu trúc linh hoạt hơn, nhưng vẫn có một số mức độ tổ chức giúp dễ dàng hơn trong việc xử lý và phân tích so với dữ liệu phi cấu trúc.

**Ví dụ:**

Dữ liệu trong định dạng XML	Dữ liệu trong định dạng JSON	Email/Lịch sử hệ thống
<pre>&lt;khachhang&gt; &lt;id&gt;001&lt;/id&gt; &lt;ten&gt;Nguyen Van A&lt;/ten&gt; &lt;diachi&gt;123 Duong A&lt;/diachi&gt; &lt;sodienthoai&gt;0909123456&lt;/sodienthoai&gt; &lt;/khachhang&gt;</pre>	<pre>{   "id": "001",   "ten": "Nguyen Van A",   "diachi": "123 Duong A",   "sodienthoai": "0909123456" }</pre>	<p><b>Email:</b></p> <ul style="list-style-type: none"> <li>Có cấu trúc: Người gửi, người nhận, thời gian gửi, chủ đề</li> <li>Phi cấu trúc: Nội dung email như văn bản/hình ảnh/video/audio,...</li> </ul> <p><b>Lịch sử hệ thống</b></p> <ul style="list-style-type: none"> <li>Có cấu trúc: Giá trị khóa như thời gian, sự kiện,...</li> <li>Phi cấu trúc: Mô tả sự kiện</li> </ul>





Dữ liệu

Thông tin

Siêu dữ liệu

Cơ sở dữ liệu

## 7. So sánh 3 loại dữ liệu

#	Khía cạnh	Dữ liệu cấu trúc	Dữ liệu phi cấu trúc	Dữ liệu bán cấu trúc
1	Ý nghĩa	Dữ liệu được tổ chức theo các hàng và cột với định dạng cố định. VD: Bảng CSDL quan hệ	Không có cấu trúc cố định hoặc tổ chức, khó phân tích trực tiếp. VD: Hình ảnh, video, audio	Có cấu trúc nhất định nhưng không hoàn toàn cứng nhắc, linh hoạt trong tổ chức. VD: Tệp XML, JSON
2	Định dạng lưu trữ	Bảng biểu, CSDL quan hệ với hàng, cột rõ ràng. VD: Bảng khách hàng	Văn bản, hình ảnh, video, âm thanh, tài liệu không theo định dạng cố định. VD: Hình ảnh	Các tệp có định dạng linh hoạt như XML, JSON, CSV, log file. VD: XML, JSON, CSV
3	Đơn vị đo lường	Số nguyên, chuỗi, ngày tháng, tiền tệ, kiểu dữ liệu đơn giản. VD: Số liệu bán hàng	Không có đơn vị đo lường cụ thể hoặc có thể là bất kỳ loại dữ liệu nào. VD: Hình ảnh, âm thanh	Cặp giá trị khóa, thẻ đánh dấu hoặc các yếu tố tổ chức linh hoạt. VD: Dữ liệu từ log file





Dữ liệu

Thông tin

Siêu dữ liệu

Cơ sở dữ liệu

## 7. So sánh 3 loại dữ liệu

#	Khía cạnh	Dữ liệu cấu trúc	Dữ liệu phi cấu trúc	Dữ liệu bán cấu trúc
4	Tổ chức quản lý	<p>Dễ quản lý nhờ vào định dạng rõ ràng, thường được lưu trong các hệ quản trị CSDL (DBMS).</p> <p>VD: Bảng khách hàng gồm có ID, Họ tên, Địa chỉ, Ngày sinh, Số điện thoại</p>	<p>Không tuân theo cấu trúc nhất quán, khó quản lý và tổ chức.</p> <p>VD: Một tập hợp các email không được tổ chức rõ ràng</p>	<p>Có cấu trúc cơ bản để tổ chức dữ liệu, nhưng không bị giới hạn bởi bảng và cột cố định.</p> <p>VD: Tập XML chứa thông tin khách hàng</p>
5	Xử lý tính toán	<p>Dễ xử lý với các công cụ cơ bản như SQL, bảng tính, CSDL quan hệ.</p> <p>VD: Truy vấn SQL với bảng dữ liệu khách hàng</p>	<p>Đòi hỏi các công cụ phân tích đặc biệt như AI, machine learning, phân tích văn bản hoặc hình ảnh.</p> <p>VD: Xử lý hình ảnh AI</p>	<p>Cần các công cụ chuyên biệt để xử lý (XML parser, JSON parser), nhưng dễ hơn so với dữ liệu phi cấu trúc.</p> <p>VD: Phân tích tệp JSON bằng công cụ chuyên biệt như Postman</p>





## 8. Tệp dữ liệu (Data File)

**Khái niệm:** Tệp dữ liệu là một **tệp chứa dữ liệu**, được lưu trữ trên máy tính hoặc các thiết bị lưu trữ khác. Tệp dữ liệu có thể **chứa các loại thông tin khác nhau**, từ **văn bản**, **số liệu**, **hình ảnh** đến các **tệp nhị phân** hoặc **tệp chuyên dụng** cho các phần mềm cụ thể. Các tệp này được sử dụng để **lưu trữ**, **trao đổi** và **xử lý dữ liệu** giữa các hệ thống hoặc chương trình.

**Ví dụ:**

Tệp cấu trúc JSON	Tệp Log
<p>Mô tả: Tệp JSON chứa thông tin cấu trúc về một đơn hàng trong một trang web thương mại điện tử.</p> <p>VD:</p> <pre>{   "order_id": "12345",   "customer": {     "name": "Nguyen Van A",     "address": "123 Duong A, Hanoi"   },   "items": [     {"product": "Laptop", "quantity": 1, "price": 15000000},     {"product": "Mouse", "quantity": 2, "price": 200000}   ] }</pre>	<p>Mô tả: Tệp log chứa các bản ghi về hoạt động của một ứng dụng web trong ngày, bao gồm thời gian, loại yêu cầu và trạng thái.</p> <p>VD:</p> <pre>[2023-08-17 12:34:56] INFO: User 'admin' logged in. [2023-08-17 12:36:01] ERROR: Failed to connect to database.</pre>





## Dữ liệu

## Thông tin

## Siêu dữ liệu

## Cơ sở dữ liệu

### 8. Tệp dữ liệu (Data File)

**Các loại Data File chính thường gặp:**

#	Loại Data File	Định dạng	Đặc điểm	Ví dụ
1	Tệp văn bản (Text Files)	TXT, CSV, JSON, XML, HTML,...	Chứa các ký tự văn bản và có thể đọc được bằng phần mềm soạn thảo văn bản đơn giản.	<ul style="list-style-type: none"> <li>- Tệp TXT chứa ghi chú cá nhân.</li> <li>- Tệp CSV chứa dữ liệu bảng biểu, như danh sách học sinh với các cột tên, tuổi, lớp.</li> </ul>
2	Tệp nhị phân (Binary Files)	BIN, EXE, DLL, OBJ, JPG, MP4, MP3,...	Chứa dữ liệu dưới dạng nhị phân (binary) và thường không thể đọc được trực tiếp bằng phần mềm soạn thảo văn bản. Thường được sử dụng bởi các phần mềm hoặc hệ thống cụ thể để thực thi hoặc hiển thị nội dung.	<ul style="list-style-type: none"> <li>- Tệp hình ảnh JPG.</li> <li>- Tệp nhạc MP3.</li> <li>- Tệp thực thi EXE của chương trình phần mềm.</li> </ul>





Dữ liệu

Thông tin

Siêu dữ liệu

Cơ sở dữ liệu

## 8. Tệp dữ liệu (Data File)

Các loại Data File chính thường gặp:

#	Loại Data File	Định dạng	Đặc điểm	Ví dụ
3	Tệp CSDL (Database Files)	MDB, ACCDB, SQL, DBF, SQLite,...	Chứa dữ liệu được tổ chức và quản lý theo một cấu trúc nhất định, thường được sử dụng bởi các hệ quản trị CSDL.	<ul style="list-style-type: none"> <li>- Tệp MDB chứa CSDL Microsoft Access.</li> <li>- Tệp SQL chứa các câu lệnh và dữ liệu của một CSDL QH.</li> </ul>
4	Tệp cấu trúc (Structured Files)	JSON, XML, YAML,...	Chứa dữ liệu được tổ chức theo định dạng có cấu trúc như cặp giá trị khóa, thẻ đánh dấu và được sử dụng để trao đổi dữ liệu giữa các hệ thống.	<ul style="list-style-type: none"> <li>- Tệp JSON chứa dữ liệu cấu trúc dưới dạng cặp khóa-giá trị (key-value).</li> <li>- Tệp XML chứa dữ liệu được đánh dấu bởi các thẻ (tags).</li> </ul>
5	Tệp log (Log Files)	LOG, TXT,...	Chứa các bản ghi (logs) của một hệ thống hoặc chương trình, được sử dụng để theo dõi hoạt động hoặc phát hiện lỗi.	<ul style="list-style-type: none"> <li>- Tệp LOG ghi lại các sự kiện xảy ra trong hệ thống máy tính.</li> <li>- Tệp nhật ký của ứng dụng web ghi lại các yêu cầu HTTP từ người dùng.</li> </ul>





TQ

## 2. Thông tin (Information)



Dữ liệu

Thông tin

Siêu dữ liệu

Cơ sở dữ liệu

### 1. Khái niệm:

Thông tin là **dữ liệu đã được xử lý và có ý nghĩa**. Nó mang lại sự hiểu biết hoặc giải thích về một vấn đề cụ thể.

### 2. Ví dụ:

Dữ liệu giao dịch của khách hàng			
#	Mã giao dịch	Mã khách hàng	Tháng giao dịch
1	GD1	KH1	03/2024
2	GD2	KH1	04/2024
3	GD3	KH2	05/2024
4	GD4	KH2	08/2024
5	GD5	KH2	06/2024
6	GD6	KH3	08/2024



Thông tin số lượng giao dịch theo tháng		
#	Tháng giao dịch	Số lượng giao dịch
1	03/2024	1
2	04/2024	1
3	05/2024	1
4	06/2024	1
5	08/2024	2





TQ

# 2. Thông tin (Information)



Dữ liệu

Thông tin

Siêu dữ liệu

Cơ sở dữ liệu

## 3. So sánh dữ liệu và thông tin

#	Khía cạnh	Dữ liệu	Thông tin
1	Ý nghĩa	Dữ liệu là các dạng của dấu hiệu, sự kiện, hoặc tác động mà không có ý nghĩa tức thì.	Thông tin là dữ liệu đã được xử lý và có ý nghĩa.
2	Định dạng	Dạng số, văn bản, hình ảnh, âm thanh, ký tự, các sự kiện thô, bảng dữ liệu, cảm biến,...	Dạng biểu đồ, bảng thống kê, báo cáo, đồ thị, văn bản diễn giải.
3	Đơn vị đo lường	Số nguyên, số thực, ký tự, điểm ảnh, các dữ kiện thô.	Không có đơn vị đo lường cụ thể vì nó phụ thuộc vào dạng thông tin được truyền tải.
4	Thể hiện	Thường không có ý nghĩa rõ ràng nếu đứng riêng lẻ và cần phải được tổ chức lại.	Có ý nghĩa trong ngữ cảnh cụ thể và có thể hiểu dễ dàng hơn.
5	Hỗ trợ ra quyết định	Dữ liệu không trực tiếp hỗ trợ ra quyết định vì chưa có ngữ cảnh rõ ràng.	Thông tin hỗ trợ ra quyết định vì nó cung cấp hiểu biết từ dữ liệu.
6	Ví dụ minh họa	Nhiệt độ hàng ngày: 30°C, 31°C, 29°C, 33°C,...	Trung bình nhiệt độ trong tuần là 31°C và có xu hướng tăng lên, cho thấy khả năng sẽ có đợt nắng nóng vào cuối tuần.



## 1. Khái niệm:

Siêu dữ liệu (Metadata) là **dữ liệu mô tả** các **thuộc tính** và **thông tin** về **dữ liệu khác**. Nói cách khác, Metadata cung cấp thông tin về cấu trúc, nội dung, nguồn gốc, quyền truy cập, và nhiều yếu tố khác liên quan đến tập dữ liệu chính. Metadata là một **yếu tố** quan trọng **trong việc quản lý** và **sử dụng dữ liệu**, **cung cấp thông tin bổ sung** giúp **hiểu rõ** hơn về bản chất, nguồn gốc, và cách sử dụng **dữ liệu** đó mà không cần trực tiếp xử lý dữ liệu gốc.

## 2. Ví dụ:

Metadata của một bức ảnh	Metadata của một Email
Dữ liệu chính: Hình ảnh chụp cảnh thiên nhiên. Metadata: <ul style="list-style-type: none"><li>- Định dạng ảnh: JPEG</li><li>- Kích thước: 1920x1080 pixels</li><li>- Kích thước file: 2 MB</li><li>- Ngày chụp: 15/08/2023</li><li>- Địa điểm: Hồ Tây, Hà Nội</li><li>- Máy ảnh: Canon EOS 5D Mark IV</li><li>- Thông số máy ảnh: ISO 200, F/2.8, 1/250s</li></ul>	Dữ liệu chính: Nội dung email. Metadata: <ul style="list-style-type: none"><li>- Người gửi: abc@example.com</li><li>- Người nhận: xyz@example.com</li><li>- Chủ đề: Cuộc họp ngày 20/08</li><li>- Ngày gửi: 19/08/2023</li><li>- Định dạng: Văn bản</li></ul>





TQ

# 4. CSDL (Database)



Dữ liệu

Thông tin

Siêu dữ liệu

Cơ sở dữ liệu

## 1. Khái niệm:

CSDL (CSDL) là **một tập hợp có tổ chức của dữ liệu được lưu trữ và quản lý** sao cho có thể dễ dàng truy cập, thao tác và cập nhật. CSDL thường chứa thông tin được **sắp xếp theo một cấu trúc cụ thể**, giúp việc tìm kiếm và xử lý dữ liệu trở nên hiệu quả hơn. CSDL có thể chứa các dạng dữ liệu khác nhau như văn bản, số liệu, hình ảnh, và thường được quản lý bởi các hệ quản trị CSDL (DBMS - Database Management System).

## 2. Ví dụ:

CSDL khách hàng của một cửa hàng với 3 bảng: KhachHang, SanPham, HoaDon

**TABLE: KhachHang – Lưu thông tin Khách hàng**

MaKhachHang	TenKhachHang	SoDienThoai	Email	DiaChi
KH001	Nguyễn Văn A	0987654321	nguyenvana@gmail.com	123 Đường ABC
KH002	Trần Thị B	0912345678	tranthib@gmail.com	456 Đường XYZ





TQ

# 4. CSDL (Database)



Dữ liệu

Thông tin

Siêu dữ liệu

Cơ sở dữ liệu

## 2. Ví dụ:

CSDL khách hàng của một cửa hàng với 3 bảng: KhachHang, SanPham, HoaDon

**TABLE: SanPham – Lưu thông tin Sản phẩm**

MaSanPham	TenSanPham	LoaiSanPham	GiaSanPham
SP001	TV	Điện tử	10,000,000
SP002	Tủ lạnh	Điện lạnh	12,000,000

**TABLE: HoaDon – Lưu thông tin Hóa đơn**

MaHoaDon	MaKhachHang	NgayMua	MaSanPham	SoLuong	TongTien
HD001	KH001	2023-08-01	SP001	1	10,000,000
HD002	KH002	2023-08-15	SP002	2	24,000,000





## 3. Tính chất CSDL – ACID

ACID là viết tắt của 4 tính chất quan trọng đảm bảo tính toàn vẹn và độ tin cậy của giao dịch trong CSDL hay các hoạt động trên CSDL diễn ra một cách an toàn và có thể khôi phục trong TH có lỗi.

### 3.1. Tính nguyên tố (Atomicity)

Mô tả: Tính nguyên tố đảm bảo rằng 1 giao dịch phải được thực hiện hoàn toàn hoặc không thực hiện gì cả. Giao dịch được coi là **1 đơn vị không thể chia nhỏ**. Nếu 1 phần của giao dịch không thể hoàn thành vì lỗi, toàn bộ giao dịch sẽ bị hủy bỏ và CSDL sẽ quay lại trạng thái trước khi giao dịch diễn ra.

VD: Một giao dịch chuyển tiền từ tài khoản A sang tài khoản B. Giao dịch này bao gồm hai bước:

- Trừ số tiền từ tài khoản A.
- Cộng số tiền vào tài khoản B.

Nếu có sự cố xảy ra sau khi bước 1 hoàn thành nhưng trước khi bước 2 diễn ra (ví dụ: mất điện), giao dịch phải được hủy bỏ hoàn toàn để số tiền không bị trừ khỏi tài khoản A mà không cộng vào tài khoản B. Kết quả là, không có sự thay đổi nào trên CSDL nếu cả hai bước không thực hiện thành công.

**Kết luận: Giao dịch được coi là 1 đơn vị không thể chia nhỏ**





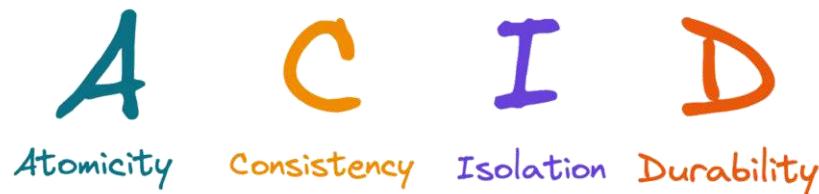
## 3. Tính chất CSDL – ACID

### 3.2. Tính nhất quán (Consistency)

Mô tả: Tính nhất quán đảm bảo rằng CSDL luôn chuyển từ 1 trạng thái hợp lệ sang 1 trạng thái hợp lệ khác sau khi một giao dịch hoàn tất. Điều này có nghĩa là các quy tắc về toàn vẹn dữ liệu được duy trì trước và sau giao dịch.

VD: Trong CSDL ngân hàng, 1 ràng buộc nhất quán có thể là không tài khoản nào được phép có số dư âm. Khi thực hiện giao dịch rút tiền, nếu sau khi trừ tiền mà số dư của tài khoản trở nên âm, giao dịch sẽ không được chấp nhận để đảm bảo tính nhất quán của CSDL.

**Kết luận: Một quy tắc phải được duy trì trước và sau giao dịch.**





## 3. Tính chất CSDL – ACID

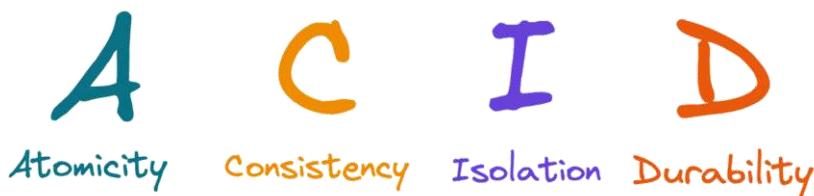
### 3.3. Tính tách biệt (Isolation)

Mô tả: Tính tách biệt đảm bảo rằng các giao dịch đang diễn ra đồng thời không được ảnh hưởng lẫn nhau. Mỗi giao dịch phải được thực thi trong trạng thái như thể nó là giao dịch duy nhất trong hệ thống. Sau khi một giao dịch hoàn tất, các thay đổi của nó mới có thể được thấy bởi các giao dịch khác.

VD: Giả sử 2 giao dịch diễn ra đồng thời trên cùng 1 tài khoản ngân hàng: 1 giao dịch rút tiền và 1 giao dịch nạp tiền.

- Tính tách biệt đảm bảo rằng nếu giao dịch rút tiền xảy ra trước, giao dịch nạp tiền sẽ chỉ thấy số dư tài khoản sau khi giao dịch rút tiền đã hoàn tất.
- Ngược lại, nếu giao dịch nạp tiền xảy ra trước, giao dịch rút tiền sẽ chỉ thấy số dư đã được nạp vào. Điều này đảm bảo rằng các giao dịch không can thiệp vào nhau và không gây ra lỗi dữ liệu.

**Kết luận: Các giao dịch đang diễn ra đồng thời không được ảnh hưởng lẫn nhau.**





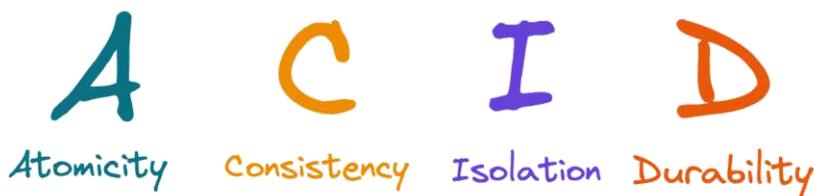
## 3. Tính chất CSDL – ACID

### 3.4. Tính bền vững (Durability)

Mô tả: Tính bền vững đảm bảo rằng khi một giao dịch đã hoàn tất và được xác nhận, mọi thay đổi mà giao dịch đó tạo ra sẽ được lưu trữ vĩnh viễn trong CSDL, ngay cả khi hệ thống gặp sự cố (chẳng hạn như mất điện hoặc lỗi hệ thống). Điều này có nghĩa là dữ liệu sẽ không bị mất sau khi giao dịch đã được cam kết (commit).

VD: Sau khi giao dịch chuyển tiền giữa hai tài khoản đã được hoàn tất và cam kết, dù hệ thống gặp sự cố (mất điện, treo máy) thì số tiền đã chuyển sẽ không bị mất, và dữ liệu sẽ không bị đảo ngược về trạng thái trước giao dịch. Khi hệ thống phục hồi, trạng thái mới của dữ liệu vẫn được duy trì chính xác như sau khi giao dịch kết thúc.

**Kết luận: Giao dịch sau khi đã được cam kết (commit) thì sẽ không bị mất đi.**





Dữ liệu

Thông tin

Siêu dữ liệu

Cở sở dữ liệu

## 3. Tính chất CSDL – ACID

### 3.5. Ví dụ tổng quát

Một người mua hàng trên trang thương mại điện tử và thanh toán trực tuyến.

#	ACID	Nội dung
1	Nguyên tố (A – Atomicity)	Nếu tiền đã bị trừ khỏi tài khoản nhưng không thể đặt đơn hàng, giao dịch sẽ bị hủy để tránh mất tiền.
2	Nhất quán (C – Consistency)	Số dư tài khoản phải phản ánh đúng giao dịch đã hoàn thành, không được âm bất hợp lý.
3	Tách biệt (I – Isolation)	Nếu có nhiều người dùng đang mua hàng và thực hiện giao dịch đồng thời, các giao dịch này sẽ không ảnh hưởng đến nhau.
4	Bền vững (D – Durability)	Sau khi thanh toán hoàn tất và đơn hàng được xác nhận, dữ liệu của giao dịch này sẽ không bị mất ngay cả khi hệ thống gặp sự cố.

A      C      I      D  
 Atomicity    Consistency    Isolation    Durability





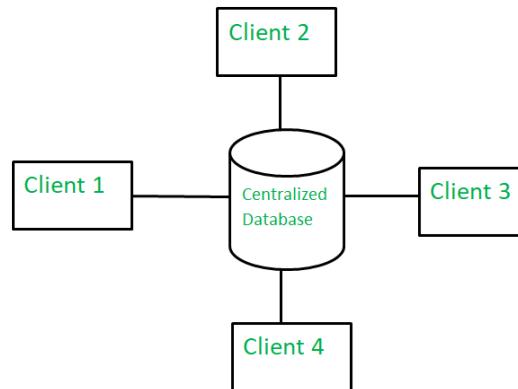
## 4. CSDL tập trung (Centralized Database)

**Khái niệm:** CSDL tập trung (Centralized Database) là loại CSDL trong đó tất cả dữ liệu được lưu trữ và quản lý tại một vị trí trung tâm, thường là trên một máy chủ đơn hoặc một hệ thống máy chủ đặt tại một trung tâm dữ liệu duy nhất. Người dùng từ các vị trí khác nhau truy cập CSDL tập trung này thông qua mạng, nhưng tất cả các dữ liệu đều được quản lý tại vị trí duy nhất này.

**Ví dụ:** Một ngân hàng với nhiều chi nhánh khác nhau trên cả nước có thể sử dụng một CSDL tập trung để lưu trữ toàn bộ thông tin về khách hàng, tài khoản, giao dịch,... Các chi nhánh sẽ truy cập vào máy chủ trung tâm này để tra cứu và xử lý dữ liệu.

Khách hàng thực hiện giao dịch tại một chi nhánh, nhưng dữ liệu về giao dịch này sẽ được cập nhật ngay lập tức vào CSDL tập trung, và các chi nhánh khác có thể truy cập thông tin mới nhất của khách hàng đó.

**Kết luận: Một máy chủ.**





Dữ liệu

Thông tin

Siêu dữ liệu

Cơ sở dữ liệu

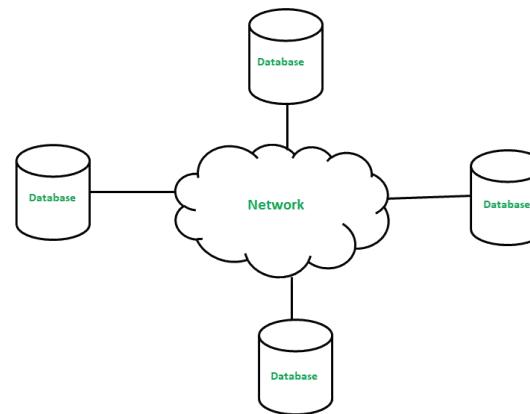
## 5. CSDL phân tán (Distributed Database)

**Khái niệm:** CSDL phân tán (Distributed Database) là loại CSDL trong đó dữ liệu được phân bố tại nhiều địa điểm khác nhau, có thể là các máy chủ khác nhau trong cùng một hệ thống hoặc trên các khu vực địa lý khác nhau. Mỗi địa điểm có thể lưu trữ một phần của dữ liệu hoặc một bản sao của toàn bộ dữ liệu. Các máy chủ này được kết nối qua mạng, và người dùng có thể truy cập và thao tác với dữ liệu.

**Ví dụ:** Các ngân hàng có nhiều chi nhánh trên toàn thế giới thường sử dụng cơ sở dữ liệu phân tán. Mỗi chi nhánh của ngân hàng có thể có máy chủ lưu trữ dữ liệu khách hàng của mình, nhưng vẫn đồng bộ với các chi nhánh khác để chia sẻ và cập nhật thông tin.

Khách hàng có thể thực hiện giao dịch tại một chi nhánh ở New York, và giao dịch này sẽ được cập nhật ngay lập tức vào cơ sở dữ liệu phân tán để chi nhánh khác ở London có thể xem thông tin cập nhật về tài khoản đó.

**Kết luận: Nhiều máy chủ kết nối với nhau qua mạng.**





Dữ liệu

Thông tin

Siêu dữ liệu

Cơ sở dữ liệu

## 6. So sánh CSDL tập trung và CSDL phân tán

#	Khía cạnh	CSDL tập trung	CSDL phân tán
1	Khái niệm	Một cơ sở dữ liệu được lưu trữ, định vị và duy trì ở một vị trí duy nhất.	Một cơ sở dữ liệu bao gồm nhiều cơ sở dữ liệu được kết nối với nhau và trải rộng trên các vị trí vật lý khác nhau.
2	Ưu điểm	<ul style="list-style-type: none"> <li>- Dễ quản lý và bảo trì vì toàn bộ dữ liệu tập trung tại một nơi.</li> <li>- Bảo mật dễ kiểm soát hơn do chỉ có một vị trí lưu trữ dữ liệu.</li> <li>- Dễ dàng sao lưu và khôi phục dữ liệu từ một vị trí trung tâm.</li> </ul>	<ul style="list-style-type: none"> <li>- Tăng tính sẵn sàng và độ tin cậy vì hệ thống vẫn hoạt động ngay cả khi một số máy chủ bị hỏng.</li> <li>- Giảm thiểu thời gian truy cập đối với người dùng ở các địa điểm khác nhau.</li> <li>- Khả năng mở rộng cao: Dữ liệu có thể được phân phối tới nhiều máy chủ khi hệ thống phát triển.</li> </ul>
3	Nhược điểm	<ul style="list-style-type: none"> <li>- Dễ gặp phải sự cố hệ thống nếu máy chủ trung tâm bị lỗi, dẫn đến gián đoạn toàn bộ hệ thống.</li> <li>- Khó mở rộng hệ thống khi khối lượng dữ liệu và số lượng người dùng tăng cao.</li> <li>- Hiệu suất truy xuất dữ liệu có thể giảm đối với người dùng ở xa máy chủ trung tâm.</li> </ul>	<ul style="list-style-type: none"> <li>- Phức tạp trong việc quản lý, đồng bộ và duy trì tính nhất quán của dữ liệu trên nhiều máy chủ.</li> <li>- Chi phí thiết lập, bảo trì và bảo mật cao hơn.</li> <li>- Khó khăn trong việc đảm bảo tính toàn vẹn dữ liệu khi nhiều giao dịch xảy ra đồng thời trên các máy chủ khác nhau.</li> </ul>





# TQ 4. CSDL (Database)



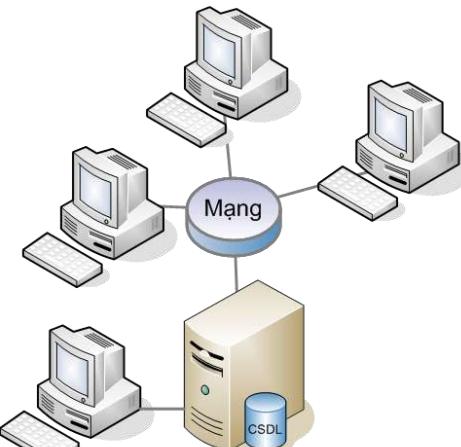
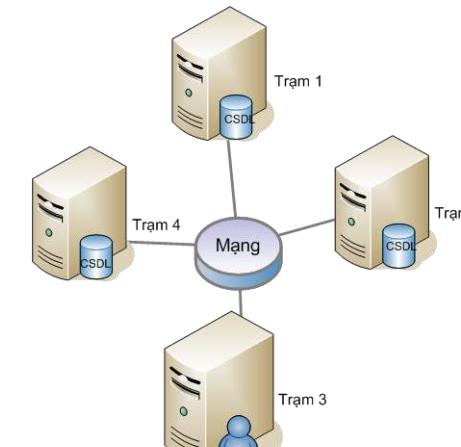
Dữ liệu

Thông tin

Siêu dữ liệu

Cơ sở dữ liệu

## 6. So sánh CSDL tập trung và CSDL phân tán

#	Khía cạnh	CSDL tập trung	CSDL phân tán
4	Ứng dụng	<ul style="list-style-type: none"> <li>Các hệ thống quản lý nhỏ gọn hoặc vừa phải, không yêu cầu tính sẵn sàng cao, ví dụ như các hệ thống quản lý trong các tổ chức cỡ nhỏ.</li> <li>Hệ thống cơ sở dữ liệu nội bộ của một công ty nhỏ hoặc tổ chức có phạm vi địa lý giới hạn.</li> </ul>	<ul style="list-style-type: none"> <li>Các tổ chức lớn có chi nhánh toàn cầu như ngân hàng, chuỗi cung ứng, và các công ty đa quốc gia.</li> <li>Mạng xã hội, hệ thống thương mại điện tử quốc tế yêu cầu tốc độ truy xuất dữ liệu nhanh từ nhiều địa điểm khác nhau.</li> <li>Hệ thống y tế quốc gia hoặc các tổ chức chăm sóc sức khỏe có nhiều bệnh viện, phòng khám ở nhiều khu vực.</li> </ul>
5	Hình ảnh minh họa		





TQ

# 4. CSDL (Database)



Dữ liệu

Thông tin

Siêu dữ liệu

Cơ sở dữ liệu

## 7. CSDL OLTP

**Khái niệm:** CSDL OLTP (Online Transaction Processing) là một loại cơ sở dữ liệu được thiết kế để hỗ trợ các giao dịch trực tuyến trong thời gian thực. OLTP chủ yếu được sử dụng cho các ứng dụng xử lý giao dịch với số lượng lớn, liên quan đến việc thêm, sửa, xóa dữ liệu nhanh chóng. Các giao dịch này thường là ngắn, yêu cầu độ chính xác và tính nhất quán cao. Hệ thống OLTP thường được tối ưu hóa cho việc xử lý số lượng lớn các giao dịch đồng thời.

**Kết luận: Xử lý các giao dịch trực tuyến liên quan đến thêm-sửa-xóa trong thời gian thực.**

**Ví dụ:** Trong một ngân hàng trực tuyến, người dùng có thể thực hiện các giao dịch như chuyển tiền, thanh toán hóa đơn, hoặc kiểm tra số dư tài khoản. Hệ thống OLTP sẽ xử lý các giao dịch này trong thời gian thực, cập nhật ngay lập tức dữ liệu về tài khoản của người dùng.

KH A thực hiện một giao dịch chuyển tiền từ tài khoản của mình sang tài khoản của KH B, hệ thống OLTP sẽ cập nhật dữ liệu ngay lập tức để ghi nhận số tiền đã bị trừ từ tài khoản của KH A và cộng vào tài khoản của KH B.





## 8. CSDL OLAP

**Khái niệm:** CSDL OLAP (Online Analytical Processing) là loại CSDL được thiết kế để hỗ trợ việc phân tích và truy vấn dữ liệu phức tạp, thường dùng cho các mục đích phân tích kinh doanh. OLAP cho phép người dùng thực hiện các phân tích dữ liệu đa chiều, từ đó tạo ra các báo cáo chi tiết và trực quan như tổng hợp, phân loại, và so sánh dữ liệu.

**Kết luận:** Hỗ trợ việc phân tích dữ liệu đa chiều để cung cấp thông tin ra quyết định.

**Ví dụ:** Giả sử một chuỗi siêu thị lớn sử dụng hệ thống OLAP để phân tích dữ liệu bán hàng. Dữ liệu được lưu trữ dưới dạng các khối dữ liệu đa chiều (cubes) với các chiều chính như:

- Chiều Thời gian: Quý, Năm
- Chiều Sản phẩm: Danh mục sản phẩm (Thực phẩm, Đồ uống, Đồ gia dụng).
- Chiều Địa điểm: Vị trí các cửa hàng (Miền Bắc, Miền Trung, Miền Nam).
- Chiều Doanh thu: Doanh thu của sản phẩm tại từng địa điểm theo thời gian quý năm.





Dữ liệu

Thông tin

Siêu dữ liệu

Cơ sở dữ liệu

## 8. CSDL OLAP

**Ví dụ:** Giả sử một chuỗi siêu thị lớn sử dụng hệ thống OLAP để phân tích dữ liệu bán hàng. Dữ liệu được lưu trữ dưới dạng các khối dữ liệu đa chiều (cubes) với các chiều chính như:

	#	Địa điểm	Sản phẩm	Thời gian	Doanh thu (triệu đồng)
Dữ liệu	1	Miền Bắc	Thực phẩm	Quý 1 – 2024	500
Thông tin	2	Miền Bắc	Đồ uống	Quý 1 – 2024	200
Siêu dữ liệu	3	Miền Bắc	Đồ gia dụng	Quý 1 – 2024	300
Cơ sở dữ liệu	4	Miền Nam	Thực phẩm	Quý 1 – 2024	600
	5	Miền Nam	Đồ uống	Quý 1 – 2024	450
	6	Miền Nam	Đồ gia dụng	Quý 1 – 2024	240
	7	Miền Bắc	Thực phẩm	Quý 2 – 2024	450
	8	Miền Nam	Đồ gia dụng	Quý 2 – 2024	610





TQ

# 4. CSDL (Database)



Dữ liệu

Thông tin

Siêu dữ liệu

Cơ sở dữ liệu

## 8. CSDL OLAP

**Ví dụ:** Giả sử một chuỗi siêu thị lớn sử dụng hệ thống OLAP để phân tích dữ liệu bán hàng. Dữ liệu được lưu trữ dưới dạng các khối dữ liệu đa chiều (cubes) với các chiều chính như:

- Chiều Thời gian: Quý, Năm
- Chiều Sản phẩm: Danh mục sản phẩm (Thực phẩm, Đồ uống, Đồ gia dụng).
- Chiều Địa điểm: Vị trí các cửa hàng (Miền Bắc, Miền Trung, Miền Nam).
- Chiều Doanh thu: Doanh thu của sản phẩm tại từng địa điểm theo thời gian quý năm.

### Truy vấn:

- Doanh thu bán hàng của "Thực phẩm" tại khu vực "Miền Bắc" trong "Quý 1 năm 2024" là bao nhiêu?  
Kết quả: 500 triệu VND.
- Tổng doanh thu bán hàng tại "Miền Nam" trong "Quý 2 năm 2024" là bao nhiêu?  
Kết quả: 610 triệu VND.
- So sánh doanh thu giữa các mặt hàng "Đồ uống" ở "Miền Bắc" và "Miền Nam" trong "Quý 1 và Quý 2 năm 2024"? Kết quả: Miền Bắc (Quý 1): 200 triệu VND; Miền Nam (Quý 1): 450 triệu VND.





TQ

# 4. CSDL (Database)



Dữ liệu

Thông tin

Siêu dữ liệu

Cở sở dữ liệu

## 9. So sánh CSDL OLTP và CSDL OLAP

#	<b>Khía cạnh</b>	<b>CSDL tập trung</b>	<b>CSDL phân tán</b>
1	Khái niệm	CSDL xử lý các giao dịch trực tuyến theo thời gian thực.	CSDL hỗ trợ phân tích dữ liệu và truy vấn phức tạp, thường dùng cho mục tiêu phân tích kinh doanh.
2	Đối tượng sử dụng	Người dùng cuối như nhân viên bán hàng, khách hàng, người vận hành,...	Nhà quản lý, nhà phân tích dữ liệu, bộ phận tài chính,...
3	Lưu trữ dữ liệu	Lưu trữ dữ liệu hiện tại, ngắn hạn (thường theo sự kiện giao dịch).	Lưu trữ dữ liệu lịch sử, dài hạn để phân tích và báo cáo.
4	Tổ chức dữ liệu	Dữ liệu được tổ chức dạng bảng với các hàng và cột đơn giản (quan hệ).	Dữ liệu được tổ chức dưới dạng khối (cubes) đa chiều.
5	Tần suất truy suất	Cao, xử lý hàng ngàn giao dịch mỗi giây (real-time).	Thấp hơn, chủ yếu phục vụ các truy vấn phức tạp và tổng hợp dữ liệu.





TQ

# 4. CSDL (Database)



Dữ liệu

Thông tin

Siêu dữ liệu

Cở sở dữ liệu

## 9. So sánh CSDL OLTP và CSDL OLAP

#	Khía cạnh	CSDL tập trung	CSDL phân tán
6	Dư thừa dữ liệu	Ít dư thừa, tập trung vào tính toàn vẹn dữ liệu.	Có thể có dư thừa, vì dữ liệu thường được lưu trữ dưới dạng tổng hợp, nhóm lại theo các chiều.
7	Mức độ xử lý song song	Thường xử lý từng giao dịch riêng lẻ.	Có khả năng xử lý nhiều truy vấn đồng thời trên các chiều dữ liệu khác nhau.
8	Giới hạn dữ liệu	Thường nhỏ và cụ thể, tập trung vào dữ liệu hiện tại.	Lớn, bao gồm dữ liệu lịch sử từ nhiều năm và nhiều nguồn khác nhau.
	Kết luận	OLTP phù hợp với các hệ thống giao dịch thời gian thực như bán hàng, ngân hàng, quản lý đặt hàng, tập trung vào hiệu quả và tốc độ xử lý các giao dịch nhỏ lẻ.	OLAP phù hợp cho phân tích dữ liệu, cung cấp các báo cáo và insights từ dữ liệu lịch sử để hỗ trợ ra quyết định kinh doanh dài hạn.





## 10. CSDL quan hệ (Relational Database)

**Khái niệm:** Cơ sở dữ liệu quan hệ (Relational Database) là loại cơ sở dữ liệu lưu trữ dữ liệu dưới dạng các bảng (table), trong đó các bảng này có mối liên hệ với nhau thông qua các khóa (keys). Dữ liệu trong các bảng được tổ chức theo các hàng (row) và cột (column), với mỗi cột đại diện cho một thuộc tính của dữ liệu, và mỗi hàng là một bản ghi dữ liệu.

**Kết luận:** CSDL quan hệ lưu trữ DL dưới dạng bảng có mối liên hệ với nhau thông qua các khóa.

**Ví dụ:** Giả sử một công ty quản lý việc bán hàng bằng cơ sở dữ liệu quan hệ với các bảng sau:

- Bảng **KhachHang** liên kết với bảng **HoaDon** qua MaKhachHang (khóa ngoại trong bảng **HoaDon**).
- Bảng **SanPham** liên kết với bảng **HoaDon** qua MaSanPham (khóa ngoại trong bảng **HoaDon**).

**TABLE: KhachHang – Lưu thông tin Khách hàng**

MaKhachHang	TenKhachHang	SoDienThoai	Email	DiaChi
KH001	Nguyễn Văn A	0987654321	nguyenvana@gmail.com	123 Đường ABC
KH002	Trần Thị B	0912345678	tranthib@gmail.com	456 Đường XYZ





TQ

# 4. CSDL (Database)



Dữ liệu

Thông tin

Siêu dữ liệu

Cơ sở dữ liệu

## 10. CSDL quan hệ (Relational Database)

**Ví dụ:** Giả sử một công ty quản lý việc bán hàng bằng cơ sở dữ liệu quan hệ với các bảng sau:

**TABLE: SanPham – Lưu thông tin Sản phẩm**

<b>MaSanPham</b>	<b>TenSanPham</b>	<b>GiaSanPham</b>
SP001	Laptop	20,000,000
SP002	Điện thoại	15,000,000

**TABLE: HoaDon – Lưu thông tin Hóa đơn**

<b>MaHoaDon</b>	<b>MaKhachHang</b>	<b>NgayMua</b>	<b>MaSanPham</b>	<b>SoLuong</b>	<b>TongTien</b>
HD001	KH001	2023-08-01	SP001	1	10,000,000
HD002	KH002	2023-08-15	SP002	2	24,000,000





## 11. CSDL phi quan hệ (NoSQL Database)

**Khái niệm:** CSDL phi quan hệ (NoSQL Database) được thiết kế để xử lý dữ liệu phi cấu trúc hoặc bán cấu trúc với hiệu suất cao, khả năng mở rộng linh hoạt, và khả năng xử lý các khối lượng lớn dữ liệu. NoSQL không yêu cầu việc thiết kế sơ đồ (schema) cứng nhắc như CSDL quan hệ, cho phép lưu trữ dữ liệu với cấu trúc linh hoạt hơn.

**Kết luận:** **CSDL quan hệ lưu trữ DL phi cấu trúc hoặc bán cấu trúc.**

**Ví dụ:**

MongoDB (CSDL dạng tài liệu)	Redis (CSDL dạng khóa-giá trị):
<p>Một trang web thương mại điện tử lớn lưu trữ thông tin sản phẩm. Thay vì sử dụng các bảng như trong CSDL quan hệ, hệ thống sử dụng MongoDB để lưu trữ dữ liệu sản phẩm dưới dạng tài liệu JSON, trong đó mỗi sản phẩm có thể có các thuộc tính khác nhau. Dữ liệu có thể bao gồm tên sản phẩm, mô tả, hình ảnh, và các đánh giá từ người dùng dưới dạng một cấu trúc linh hoạt.</p> <p>Lợi ích: Giúp trang web dễ dàng thêm các thuộc tính mới cho sản phẩm mà không cần thay đổi cấu trúc CSDL.</p>	<p>Redis được sử dụng để lưu trữ các phiên đăng nhập của người dùng trên một trang web. Mỗi phiên đăng nhập có một khóa là ID người dùng và giá trị là thông tin phiên như token xác thực, thời gian đăng nhập. Khi người dùng đăng nhập, hệ thống có thể nhanh chóng truy vấn và cập nhật phiên đăng nhập trong Redis.</p> <p>Lợi ích: Hiệu suất cực kỳ cao khi xử lý truy xuất dữ liệu theo cặp khóa-giá trị, thích hợp cho các ứng dụng thời gian thực.</p>





TQ

# 4. CSDL (Database)



## 11. CSDL phi quan hệ (NoSQL Database)

Ví dụ:

Dữ liệu

Thông tin

Siêu dữ liệu

Cở sở dữ liệu

### Neo4j (CSDL đồ thị)

Một mạng xã hội sử dụng Neo4j để lưu trữ và phân tích mối quan hệ giữa người dùng. Mỗi người dùng là một nút trong đồ thị, và mối quan hệ bạn bè, theo dõi, thích, hoặc bình luận giữa các người dùng là các cạnh trong đồ thị. Neo4j giúp mạng xã hội dễ dàng truy xuất và phân tích mối quan hệ phức tạp giữa hàng triệu người dùng.

Lợi ích: Xử lý tốt các bài toán về phân tích mối quan hệ phức tạp, thích hợp cho các ứng dụng như mạng xã hội hoặc hệ thống khuyến nghị.

### Cassandra (CSDL dạng cột)

Một hệ thống quản lý dữ liệu lớn như Facebook sử dụng Cassandra để lưu trữ và truy xuất dữ liệu người dùng trên toàn cầu. Mỗi người dùng có thể có hàng triệu bài đăng và tương tác, và Cassandra giúp xử lý các truy vấn nhanh chóng qua việc lưu trữ dữ liệu dạng cột.

Lợi ích: Khả năng mở rộng linh hoạt, hiệu quả khi xử lý dữ liệu lớn phân tán trên nhiều máy chủ.

**Tổng kết:** CSDL phi quan hệ (NoSQL) là giải pháp tốt cho các ứng dụng yêu cầu tính linh hoạt trong lưu trữ dữ liệu, khả năng mở rộng cao và hiệu suất vượt trội khi xử lý khối lượng lớn dữ liệu phi cấu trúc hoặc bán cấu trúc. Nó được sử dụng rộng rãi trong các hệ thống hiện đại như mạng xã hội, thương mại điện tử, và các ứng dụng thời gian thực.





Dữ liệu

Thông tin

Siêu dữ liệu

Cở sở dữ liệu

## 12. So sánh CSDL quan hệ và CSDL phi quan hệ

#	Khía cạnh	CSDL tập trung	CSDL phân tán
1	Khái niệm	<ul style="list-style-type: none"> <li>Lưu trữ dữ liệu theo cấu trúc bảng với các hàng và cột.</li> <li>Các bảng có mối quan hệ với nhau thông qua khóa ngoại (Foreign Key).</li> </ul>	<ul style="list-style-type: none"> <li>Lưu trữ dữ liệu theo nhiều định dạng linh hoạt như tài liệu, cặp khóa-giá trị, đồ thị, cột.</li> <li>Không tuân theo mô hình quan hệ cố định.</li> </ul>
	Ưu điểm	<ul style="list-style-type: none"> <li>Đảm bảo các thuộc tính ACID (Atomicity, Consistency, Isolation, Durability) cho các giao dịch, giúp dữ liệu luôn được xử lý an toàn.</li> <li>Ngôn ngữ truy vấn chuẩn: SQL là ngôn ngữ truy vấn phổ biến, mạnh mẽ, được chuẩn hóa trên nhiều hệ thống.</li> </ul>	<ul style="list-style-type: none"> <li>Linh hoạt về cấu trúc dữ liệu: Không cần tuân theo một sơ đồ cứng nhắc, phù hợp cho dữ liệu phi cấu trúc và bán cấu trúc.</li> <li>Khả năng mở rộng cao: Mở rộng dễ dàng theo chiều ngang (thêm máy chủ), lý tưởng cho các ứng dụng quy mô lớn và phân tán.</li> <li>Hiệu suất cao: Xử lý nhanh chóng các khối lượng lớn dữ liệu phi cấu trúc với tốc độ vượt trội trong các trường hợp không cần tính nhất quán chặt chẽ như SQL.</li> </ul>





Dữ liệu

Thông tin

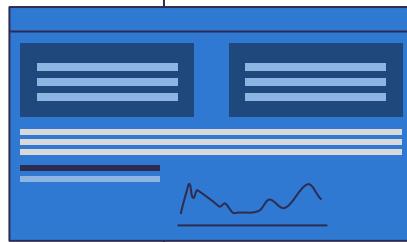
Siêu dữ liệu

Cở sở dữ liệu

## 12. So sánh CSDL quan hệ và CSDL phi quan hệ

#	Khía cạnh	CSDL tập trung	CSDL phân tán
3	Nhược điểm	<ul style="list-style-type: none"> <li>- Hạn chế về khả năng mở rộng: Khó mở rộng theo chiều ngang (thêm máy chủ) so với NoSQL.</li> <li>- Khả năng xử lý dữ liệu phi cấu trúc hạn chế: Dữ liệu không phải lúc nào cũng dễ dàng mô hình hóa trong cấu trúc bảng cột, dẫn đến sự phức tạp trong lưu trữ và xử lý dữ liệu phi cấu trúc hoặc bán cấu trúc.</li> </ul>	<ul style="list-style-type: none"> <li>- Không hỗ trợ ACID toàn diện: Trong nhiều hệ thống NoSQL, tính nhất quán và tính toàn vẹn dữ liệu không được đảm bảo nghiêm ngặt như SQL (tuân theo CAP Theorem).</li> <li>- Ngôn ngữ truy vấn không chuẩn hóa: Không có một ngôn ngữ truy vấn chung như SQL, mỗi hệ thống NoSQL có thể sử dụng các ngôn ngữ khác nhau.</li> </ul>
4	Ứng dụng	<ul style="list-style-type: none"> <li>- Ứng dụng truyền thống: Hệ thống kế toán, quản lý khách hàng (CRM), hệ thống thanh toán ngân hàng, quản lý nguồn nhân lực, ERP (Enterprise Resource Planning).</li> <li>- Dữ liệu có cấu trúc cao: Thích hợp cho các ứng dụng yêu cầu dữ liệu có cấu trúc chặt chẽ và tính toàn vẹn cao.</li> </ul>	<ul style="list-style-type: none"> <li>- Ứng dụng quy mô lớn: Mạng xã hội, hệ thống khuyến nghị, hệ thống dữ liệu lớn (Big Data), ứng dụng IoT.</li> <li>- Dữ liệu phi cấu trúc hoặc bán cấu trúc: Phù hợp với các hệ thống lưu trữ hình ảnh, video, nội dung web, phân tích dữ liệu, và các ứng dụng thời gian thực.</li> </ul>





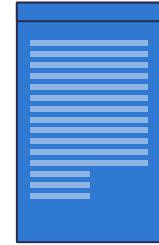
02

# Mô hình hóa dữ liệu





# Mô hình hóa dữ liệu



Thiết kế CSDL

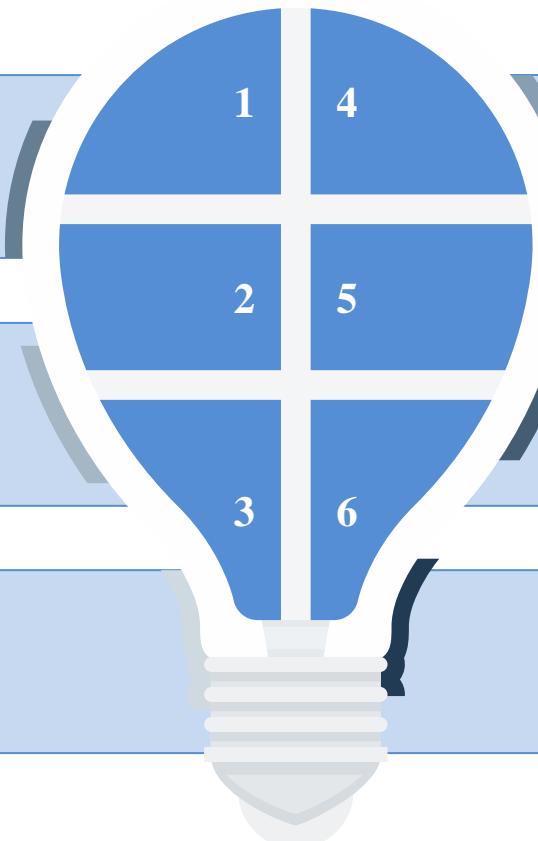
Mô hình hóa dữ liệu

KTXĐ thực thể

KTXĐ thuộc tính

KTXĐ quan hệ

KTXĐ khóa chính & ngoại



# 1. Thiết kế CSDL



## Thiết kế CSDL

Mô hình hóa  
dữ liệu

KTXĐ  
thực thể

KTXĐ  
thuộc tính

KTXĐ  
quan hệ

KTXĐ khóa  
chính & ngoại

### Khái niệm:

Thiết kế cơ sở dữ liệu (CSDL) là quá trình xác định cấu trúc và tổ chức của dữ liệu nhằm đảm bảo dữ liệu được lưu trữ, quản lý và truy xuất hiệu quả nhất trong một hệ quản trị cơ sở dữ liệu (DBMS). Quá trình này bao gồm việc xác định các thực thể (entities), các thuộc tính (attributes), các mối quan hệ (relationships) giữa các thực thể, cũng như các quy tắc, ràng buộc về tính toàn vẹn dữ liệu.

**Quá trình thiết kế CSDL:** Có thể được chia thành các giai đoạn chính:

- Thiết kế khái niệm: Xác định các thực thể, thuộc tính và mối quan hệ giữa các thực thể, thường được biểu diễn bằng các mô hình khái niệm như sơ đồ ERD (Entity-Relationship Diagram).
- Thiết kế logic: Chuyển đổi mô hình khái niệm thành mô hình logic, thường là mô hình quan hệ với các bảng, cột, và khóa.
- Thiết kế vật lý: Xác định cách lưu trữ dữ liệu trên đĩa, như việc phân vùng dữ liệu, lựa chọn chỉ mục, và các cấu trúc lưu trữ vật lý khác.

**Kết luận:** Thiết kế CSDL là quá trình tạo cấu trúc và tổ chức dữ liệu trong hệ thống cơ sở dữ liệu.



## 2. Mô hình hóa dữ liệu

Thiết kế  
CSDLMô hình hóa  
dữ liệuKTXĐ  
thực thểKTXĐ  
thuộc tínhKTXĐ  
quan hệKTXĐ khóa  
chính & ngoại

### 1. Khái niệm:

Mô hình hóa dữ liệu là quá trình xây dựng các mô hình biểu diễn cấu trúc và tổ chức của dữ liệu trong một hệ thống. Mô hình hóa dữ liệu giúp tạo ra một cái nhìn trừu tượng và có cấu trúc về dữ liệu, bao gồm các thực thể, thuộc tính của chúng, và mối quan hệ giữa các thực thể trong hệ thống.

**Kết luận:** **Mô hình hóa dữ liệu là quá trình tạo ra các mô hình biểu diễn cấu trúc và mối quan hệ của dữ liệu.**





## 2. Mô hình hóa dữ liệu

### 2. Mục tiêu:

- Tạo ra sự hiểu biết chung về dữ liệu:** Mô hình hóa dữ liệu giúp các bên liên quan (nhà phân tích, lập trình viên, người dùng) có cái nhìn rõ ràng về cấu trúc và tổ chức của dữ liệu trong hệ thống.
- Tối ưu hóa lưu trữ và truy xuất dữ liệu:** Đảm bảo rằng dữ liệu được lưu trữ và truy xuất một cách hiệu quả, giảm thiểu sự trùng lặp và tăng tốc độ truy vấn.
- Đảm bảo tính nhất quán và toàn vẹn dữ liệu:** Thiết kế các quy tắc và ràng buộc để đảm bảo dữ liệu luôn chính xác, đồng nhất và đáp ứng các yêu cầu nghiệp vụ.
- Dễ dàng bảo trì và mở rộng:** Một mô hình dữ liệu tốt sẽ dễ dàng được bảo trì, cập nhật khi có yêu cầu thay đổi về mặt nghiệp vụ hoặc công nghệ, và dễ dàng mở rộng khi có nhu cầu thêm dữ liệu mới.
- Hỗ trợ quá trình phát triển phần mềm:** Giúp nhà phát triển hiểu rõ cách thức tổ chức dữ liệu, qua đó thiết kế và phát triển các ứng dụng dựa trên dữ liệu một cách hiệu quả hơn.

**Kết luận:** Mục tiêu của mô hình hóa dữ liệu là tạo sự hiểu biết chung về dữ liệu, tối ưu hóa lưu trữ và truy xuất, đảm bảo tính nhất quán và toàn vẹn, và hỗ trợ quá trình phát triển và bảo trì hệ thống.



## 2. Mô hình hóa dữ liệu

	
Thiết kế CSDL	
Mô hình hóa dữ liệu	
KTXĐ thực thể	
KTXĐ thuộc tính	
KTXĐ quan hệ	
KTXĐ khóa chính & ngoại	

### 3. Giai đoạn 1: Mô hình dữ liệu mức khái niệm

Làm gì?	Thực hiện ở bước nào trong SDLC?	Làm việc với đối tượng nào?
<p>Bước đầu tiên trong quá trình thiết kế CSDL. Xây dựng một cái nhìn tổng quan về cấu trúc dữ liệu của hệ thống ở mức trừu tượng cao. Các nhiệm vụ bao gồm:</p> <ul style="list-style-type: none"><li>- Xác định các thực thể (entities) chính trong hệ thống.</li><li>- Xác định các thuộc tính (attributes) của các thực thể.</li><li>- Xác định các mối quan hệ (relationships) giữa các thực thể.</li><li>- Thiết lập các quy tắc và ràng buộc nghiệp vụ cơ bản.</li></ul>	Bước 3: Thiết kế kiến trúc (Designing Architecture)	<ol style="list-style-type: none"><li>1. Phân tích nghiệp vụ (BA)</li><li>2. Tham vấn các chuyên gia về thiết kế CSDL</li></ol>



## 2. Mô hình hóa dữ liệu

Thiết kế  
CSDLMô hình hóa  
dữ liệuKTXĐ  
thực thểKTXĐ  
thuộc tínhKTXĐ  
quan hệKTXĐ khóa  
chính & ngoại

### 3. Giai đoạn 1: Mô hình dữ liệu mức khái niệm

Các bước thực hiện:

Đầu vào	Các bước thực hiện	Kỹ thuật áp dụng	Kết quả đầu ra
<ul style="list-style-type: none"><li>- Các tài liệu nghiệp vụ từ BA: BRD/User Story/URD/SRS/FSD</li><li>- Buổi họp với chuyên gia CSDL</li></ul>	<p>B1: Xác định thực thể (Entity Identification)</p> <p>B2: Xác định các thuộc tính (Attribute Identification)</p> <p>B3: Xác định mối quan hệ (Relationship Identification)</p> <p>B4: Sơ đồ ER (Entity-Relationship Diagram) – biểu diễn trực quan cấu trúc dữ liệu trên</p> <p>B5: Đánh giá và điều chỉnh lại với các bên liên quan</p>	<ul style="list-style-type: none"><li>- Phân tích nghiệp vụ</li><li>- Tạo sơ đồ ER</li><li>- Làm việc trực tiếp với các bên liên quan</li></ul>	<ul style="list-style-type: none"><li>- Sơ đồ ER</li><li>- Tài liệu mô tả sơ đồ ER</li></ul>



## 2. Mô hình hóa dữ liệu



Thiết kế  
CSDL

**Mô hình hóa  
dữ liệu**

KTXĐ  
thực thể

KTXĐ  
thuộc tính

KTXĐ  
quan hệ

KTXĐ khóa  
chính & ngoại

### 4. Giai đoạn 2: Mô hình dữ liệu mức logic

Làm gì?	Thực hiện ở bước nào trong SDLC?	Làm việc với đối tượng nào?
<p>Mô hình hóa dữ liệu mức logic là quá trình chuyển đổi mô hình dữ liệu khái niệm thành một mô hình chi tiết hơn. Mô hình này mô tả cách dữ liệu sẽ được tổ chức, lưu trữ và truy xuất trong hệ thống dưới dạng các bảng, cột, khóa chính, khóa ngoại, và các quy tắc liên quan.</p> <ul style="list-style-type: none"> <li>- Xác định và mô hình hóa các bảng (tables) từ thực thể.</li> <li>- Xác định các thuộc tính (attributes) sẽ trở thành các cột trong bảng.</li> <li>- Xác định các khóa chính (primary key) và khóa ngoại (foreign key).</li> <li>- Xác định các quan hệ giữa các bảng (ví dụ: một-nhiều, nhiều-nhiều).</li> <li>- Định nghĩa các ràng buộc toàn vẹn dữ liệu và quy tắc kinh doanh chi tiết.</li> </ul>	Bước 3: Thiết kế kiến trúc (Designing Architecture)	<ol style="list-style-type: none"> <li>1. Phát triển CSDL (Database Developer)</li> <li>2. Kiến trúc sư dữ liệu (Data Architect)</li> <li>3. Phân tích nghiệp vụ (Business Analyst)</li> </ol>



Thiết kế  
CSDLMô hình hóa  
dữ liệuKTXĐ  
thực thểKTXĐ  
thuộc tínhKTXĐ  
quan hệKTXĐ khóa  
chính & ngoại

## 2. Mô hình hóa dữ liệu

### 4. Giai đoạn 2: Mô hình dữ liệu mức logic

Các bước thực hiện:

Đầu vào	Các bước thực hiện	Kỹ thuật áp dụng	Kết quả đầu ra
<ul style="list-style-type: none"><li>- Mô hình dữ liệu mức khái niệm ER</li><li>- Nghiệp vụ chi tiết</li><li>- Yêu cầu hệ thống: Thông tin về các yêu cầu hệ thống liên quan đến bảo mật, tính toàn vẹn, và hiệu suất</li></ul>	<p>B1: Chuyển đổi thực thể thành bảng (Table Conversion)</p> <p>B2: Xác định khóa chính và khóa ngoại (Primary and Foreign Keys)</p> <p>B3: Xác định các ràng buộc toàn vẹn (Integrity Constraints)</p> <p>B4: Chuẩn hóa dữ liệu (Normalization)</p> <p>B5: Định nghĩa quan hệ giữa các bảng (Define Relationships)</p> <p>B6: Xem xét hiệu suất và tối ưu hóa (Performance Considerations)</p>	<ul style="list-style-type: none"><li>- Chuẩn hóa dữ liệu</li><li>- Sơ đồ ER chi tiết</li><li>- Xác định các quy tắc ràng buộc dữ liệu</li></ul>	<ul style="list-style-type: none"><li>- Mô hình dữ liệu mức logic</li><li>- Tài liệu mô tả mô hình mức logic</li><li>- Bảng dữ liệu chuẩn hóa</li></ul>



## 2. Mô hình hóa dữ liệu

	
Thiết kế CSDL	
Mô hình hóa dữ liệu	
KTXĐ thực thể	
KTXĐ thuộc tính	
KTXĐ quan hệ	
KTXĐ khóa chính & ngoại	

### 5. Giai đoạn 3: Mô hình dữ liệu mức vật lý

Làm gì?	Thực hiện ở bước nào trong SDLC?	Làm việc với đối tượng nào?
<p>Mô hình hóa dữ liệu mức vật lý là quá trình chuyển đổi mô hình dữ liệu logic thành cấu trúc dữ liệu thực tế trên hệ quản trị cơ sở dữ liệu (DBMS). Trong giai đoạn này, các bảng, cột, chỉ mục, ràng buộc và các yếu tố khác sẽ được triển khai trên một nền tảng cụ thể.</p> <ul style="list-style-type: none"><li>- Tạo bảng (Create Tables)</li><li>- Tạo chỉ mục (Indexing)</li><li>- Quản lý lưu trữ (Storage Management)</li><li>- Thiết lập bảo mật (Security Setup)</li></ul>	<p>Bước 4: Phát triển sản phẩm (Developing Product) Bước 5: Kiểm thử và tích hợp sản phẩm (Product Testing and Integration)</p>	<ol style="list-style-type: none"><li>1. Chuyên viên CSDL (Database Administrator - DBA)</li><li>2. Phát triển phần mềm (Software Developer)</li><li>3. Kiến trúc sư hệ thống (System Architect)</li></ol>



Thiết kế  
CSDLMô hình hóa  
dữ liệuKTXĐ  
thực thểKTXĐ  
thuộc tínhKTXĐ  
quan hệKTXĐ khóa  
chính & ngoại

## 2. Mô hình hóa dữ liệu

### 5. Giai đoạn 3: Mô hình dữ liệu mức vật lý

Các bước thực hiện:

Đầu vào	Các bước thực hiện	Kỹ thuật áp dụng	Kết quả đầu ra
<ul style="list-style-type: none"> <li>- Mô hình dữ liệu mức logic</li> <li>- Yêu cầu về hiệu suất: Thông tin về số lượng truy vấn, yêu cầu hiệu suất và tốc độ truy cập dữ liệu.</li> <li>- Thông tin về hệ quản trị cơ sở dữ liệu (DBMS): Các đặc điểm kỹ thuật của hệ quản trị cơ sở dữ liệu được chọn (ví dụ: MySQL, SQL Server, Oracle).</li> </ul>	<p>B1: Xác định kiểu dữ liệu vật lý (Physical Data Types Identification)</p> <p>B2: Tạo bảng và chỉ mục (Create Tables and Indexes)</p> <p>B3: Tối ưu hóa lưu trữ và hiệu suất (Storage and Performance Optimization)</p> <p>B4: Thiết lập bảo mật và kiểm soát truy cập (Security and Access Control)</p> <p>B5: Thiết lập sao lưu và khôi phục (Backup and Recovery Setup)</p> <p>B6: Kiểm thử và điều chỉnh (Testing and Tuning)</p>	<ul style="list-style-type: none"> <li>- Tối ưu hóa chỉ mục (Index Optimization)</li> <li>- Phân vùng dữ liệu (Data Partitioning)</li> <li>- Quản lý lưu trữ (Storage Management)</li> <li>- Bảo mật dữ liệu (Data Security)</li> </ul>	<ul style="list-style-type: none"> <li>- CSDL vật lý</li> <li>- Chiến lược sao lưu và khôi phục</li> <li>- Tài liệu cấu trúc CSDL vật lý</li> </ul>



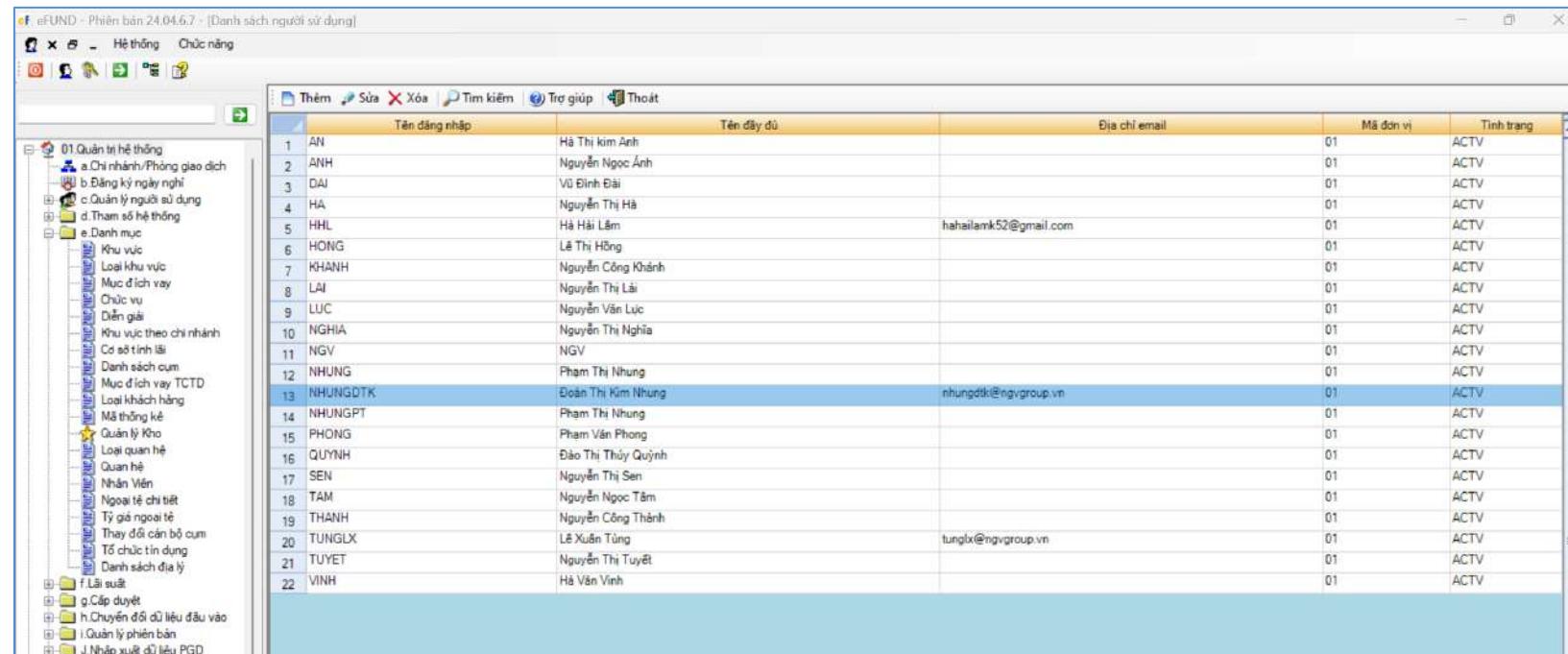


# 3. Kỹ thuật xác định Thực thể

## 1. Đối tượng là gì?

Đối tượng (Object) là các thực thể (entities) trong mô hình dữ liệu, thường được mô tả bằng các bảng trong CSDL quan hệ. Đối tượng thường đại diện cho các bản ghi trong bảng và có các thuộc tính tương ứng với các cột trong bảng.

**Ví dụ đối tượng trong CSDL:** Đối tượng trong tính năng “Danh sách người sử dụng”



	Tên đăng nhập	Tên đầy đủ	Địa chỉ email	Mã đơn vị	Tình trạng
1	AN	Hà Thị Kim Anh		01	ACTV
2	ANH	Nguyễn Ngọc Ánh		01	ACTV
3	DAI	Vũ Đình Đài		01	ACTV
4	HA	Nguyễn Thị Hà		01	ACTV
5	HHL	Hà Hải Lâm		01	ACTV
6	HONG	Lê Thị Hồng		01	ACTV
7	KHANH	Nguyễn Công Khanh		01	ACTV
8	LAI	Nguyễn Thị Lai		01	ACTV
9	LUC	Nguyễn Văn Lực		01	ACTV
10	NGHIA	Nguyễn Thị Nghĩa		01	ACTV
11	NGV	NGV		01	ACTV
12	NHUNG	Phạm Thị Nhung		01	ACTV
13	NHUNGDTK	Đoàn Thị Kim Nhung	nhungdtk@ngvgroup.vn	01	ACTV
14	NHUNGPT	Phạm Thị Nhung		01	ACTV
15	PHONG	Phạm Văn Phong		01	ACTV
16	QUYNH	Đào Thị Thúy Quynh		01	ACTV
17	SEN	Nguyễn Thị Sen		01	ACTV
18	TAM	Nguyễn Ngọc Tâm		01	ACTV
19	THANH	Nguyễn Công Thành		01	ACTV
20	TUNGLX	Lê Xuân Tùng	tunglx@ngvgroup.vn	01	ACTV
21	TUYET	Nguyễn Thị Tuyết		01	ACTV
22	VINH	Hà Văn Vinh		01	ACTV



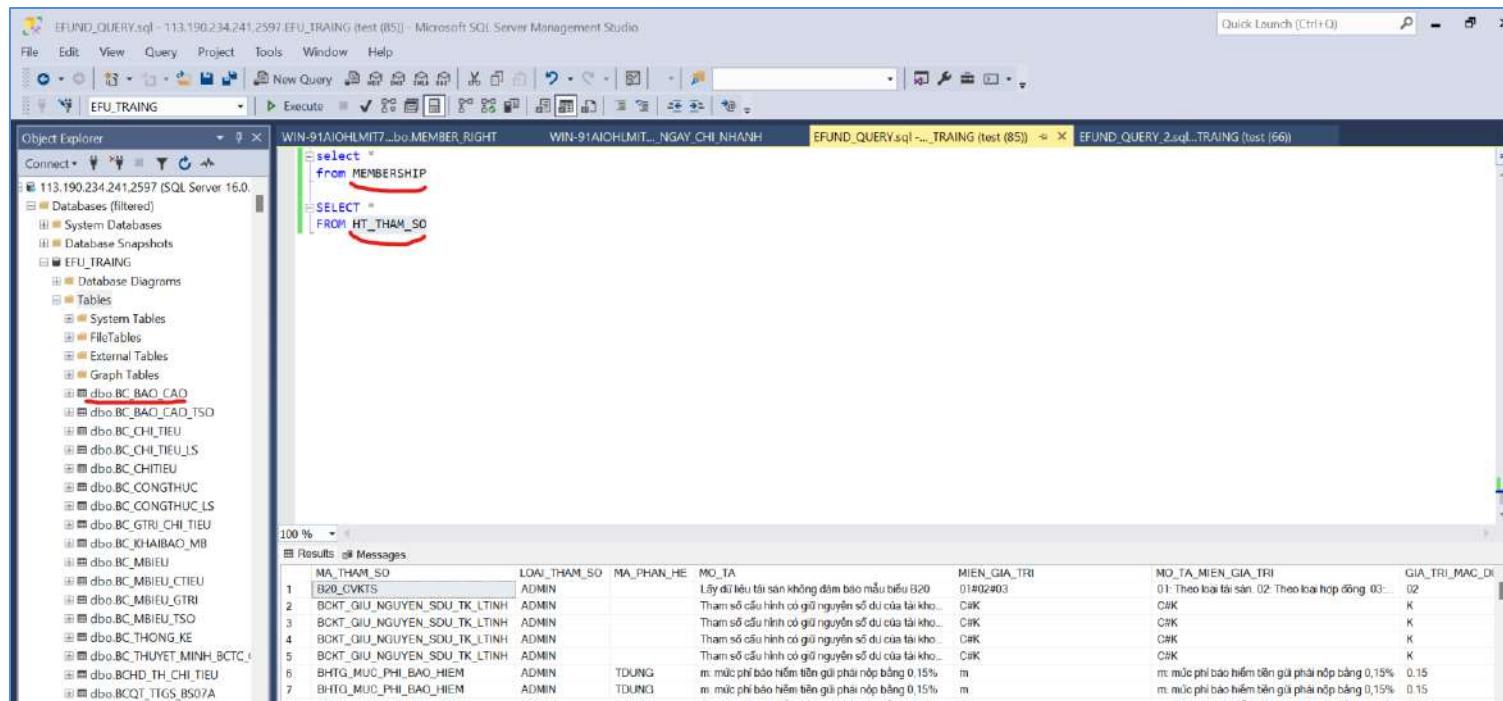


# 3. Kỹ thuật xác định Thực thể

## 2. Lớp là gì?

Lớp (class) là một bản thiết kế hoặc khuôn mẫu cho các đối tượng (objects). Lớp định nghĩa các thuộc tính – biến (attributes) và phương thức – hàm (methods) mà các đối tượng thuộc lớp đó sẽ có. Trong ngữ cảnh CSDL, lớp có thể tương đương với một bảng trong cơ sở dữ liệu.

**Ví dụ lớp trong CSDL:** Danh sách bảng trong eFUND



The screenshot shows the Microsoft SQL Server Management Studio interface. In the Object Explorer on the left, a tree view shows the database structure, including databases like 'EFU\_TRAING' and tables such as 'HT\_THAM\_SO'. In the center pane, a query window displays two SELECT statements:

```
select * from MEMBERSHIP
SELECT * FROM HT_THAM_SO
```

The results pane at the bottom shows the data for the 'HT\_THAM\_SO' table. The columns are: MA\_THAM\_SO, LOAI\_THAM\_SO, MA\_PHAN\_HE, MO\_TA, MIEN\_GIA\_TRI, MO\_TA\_MIEN\_GIA\_TRI, and GIA\_TRI\_MAC\_DI. The data rows are:

MA_THAM_SO	LOAI_THAM_SO	MA_PHAN_HE	MO_TA	MIEN_GIA_TRI	MO_TA_MIEN_GIA_TRI	GIA_TRI_MAC_DI
B20_CVKT	ADMIN		Lấy dữ liệu tài sản không đảm bảo mẫu biểu B20	01802403	01: Theo loại tài sản: 02: Theo loại hợp đồng: 03:	
BCKT_GIU_NGUYEN_SO_U_TK_LTINH	ADMIN		Tham số cấu hình có giá nguyên số dù của tài kho...	C4K	C4K	K
BCKT_GIU_NGUYEN_SO_U_TK_LTINH	ADMIN		Tham số cấu hình có giá nguyên số dù của tài kho...	C4K	C4K	K
BCKT_GIU_NGUYEN_SO_U_TK_LTINH	ADMIN		Tham số cấu hình có giá nguyên số dù của tài kho...	C4K	C4K	K
BHTG_MUC_PHI_BAO_HIEM	ADMIN	TDUNG	m: mức phí bảo hiểm tiền gửi phải nộp bằng 0,15%	m	m: mức phí bảo hiểm tiền gửi phải nộp bằng 0,15%	
BHTG_MUC_PHI_BAO_HIEM	ADMIN	TDUNCI	m: mức phí bảo hiểm tiền gửi phải nộp bằng 0,15%	m	m: mức phí bảo hiểm tiền gửi phải nộp bằng 0,15%	0,15



# 3. Kỹ thuật xác định Thực thể

## 3. Thực thể là gì?

Thực thể (Entity) là một đối tượng hoặc khái niệm độc lập mà ta muốn lưu trữ thông tin trong CSDL. Thực thể có một tập các thuộc tính mô tả nó, và các thực thể cụ thể sẽ có các giá trị khác nhau cho các thuộc tính này.

Trong cơ sở dữ liệu, một thực thể thường được biểu diễn dưới dạng một bảng, trong đó:

- Thực thể tương ứng với bảng.
- Thuộc tính của thực thể tương ứng với các cột của bảng.
- Thực thể cụ thể tương ứng với một dòng dữ liệu trong bảng.

**Ví dụ thực thể trong CSDL:** Thực thể “Hệ thống tham số” trên eFUND

The screenshot shows a SQL query window with the following details:

```
SELECT *
FROM HT_THAM_SO
```

The results table has the following columns and data:

	MA_THAM_SO	LOAI_THAM_SO	MA_PHAN_HE	MO_TA	MIEN_GIA_TRI	MO_TA_MIEN_GIA_TRI	GIA_TRI_MAC_DI
1	B20_CVKT\$	ADMIN		Lấy dữ liệu tài sản không đảm bảo mẫu biểu B20	01#02#03	01: Theo loại tài sản. 02: Theo loại hợp đồng. 03:...	02
2	BCKT_GIU_NGUYEN_SDU_TK_LTINH	ADMIN		Tham số cấu hình có giữ nguyên số dư của tài kho...	C#K	C#K	K
3	BCKT_GIU_NGUYEN_SDU_TK_LTINH	ADMIN		Tham số cấu hình có giữ nguyên số dư của tài kho...	C#K	C#K	K
4	BCKT_GIU_NGUYEN_SDU_TK_LTINH	ADMIN		Tham số cấu hình có giữ nguyên số dư của tài kho...	C#K	C#K	K
5	BCKT_GIU_NGUYEN_SDU_TK_LTINH	ADMIN		Tham số cấu hình có giữ nguyên số dư của tài kho...	C#K	C#K	K
6	BHTG_MUC_PHI_BAO_HIEM	ADMIN	TDUNG	m: mức phí bảo hiểm tiền gửi phải nộp bằng 0,15%	m	m: mức phí bảo hiểm tiền gửi phải nộp bằng 0,15%	0.15





# 3. Kỹ thuật xác định Thực thể

## 4. Bảng trong CSDL

Bảng trong cơ sở dữ liệu (CSDL) là một cấu trúc cơ bản dùng để lưu trữ dữ liệu. Mỗi bảng bao gồm các hàng (row) và cột (column), trong đó mỗi hàng đại diện cho một bản ghi (record) và mỗi cột đại diện cho một thuộc tính (attribute) của bản ghi đó. Bảng thường được sử dụng để tổ chức và quản lý dữ liệu một cách có cấu trúc.

### Ví dụ bảng trong CSDL:

Column Name	Data Type	Allow Nulls
MA_LOAI_TK	varchar(100)	<input type="checkbox"/>
MA_TK	varchar(20)	<input type="checkbox"/>
TEN_MA_THONG_KE	nvarchar(255)	<input type="checkbox"/>
GHI_CHU	nvarchar(255)	<input checked="" type="checkbox"/>
MA_CHI_NHANH	varchar(20)	<input type="checkbox"/>
BAN_GHI_THU	varchar(4)	<input type="checkbox"/>
TRANG_THAI	varchar(4)	<input type="checkbox"/>
NGUOI_NHAP	varchar(30)	<input type="checkbox"/>
NGAY_NHAP	varchar(14)	<input type="checkbox"/>
NGUOI_DUYET	varchar(30)	<input type="checkbox"/>
NGAY_DUYET	varchar(14)	<input type="checkbox"/>

DC\_LOAI\_TKE

DC\_LOAI\_TK

Column Name	Data Type	Allow Nulls
MA_LOAI_TK	varchar(20)	<input type="checkbox"/>
MO_TA	nvarchar(255)	<input type="checkbox"/>
MA_CHI_NHANH	varchar(20)	<input type="checkbox"/>
BAN_GHI_THU	varchar(4)	<input type="checkbox"/>
TRANG_THAI	varchar(4)	<input type="checkbox"/>
NGUOI_NHAP	varchar(30)	<input type="checkbox"/>
NGAY_NHAP	varchar(14)	<input type="checkbox"/>
NGUOI_DUYET	varchar(30)	<input type="checkbox"/>
NGAY_DUYET	varchar(14)	<input type="checkbox"/>
NGUON_MAPPING	nvarchar(1000)	<input checked="" type="checkbox"/>

# 3. Kỹ thuật xác định Thực thể



Thiết kế  
CSDL

Mô hình hóa  
dữ liệu

KTXĐ  
thực thể

KTXĐ  
thuộc tính

KTXĐ  
quan hệ

KTXĐ khóa  
chính & ngoại

## 5. Phân loại thực thể

Thực thể chính (Primary Entity)	Thực thể phụ thuộc (Dependent Entity)	Thực thể liên kết (Associative Entity)
<p><b>Mô tả:</b> Đây là những thực thể cơ bản và độc lập, không phụ thuộc vào thực thể khác.</p> <p><b>Ví dụ:</b> Trong một hệ thống quản lý sinh viên, <b>SinhVien</b> là một thực thể chính vì sinh viên là đối tượng chính cần được quản lý.</p> <p><b>Kết luận:</b> Thực thể chính độc lập.</p>	<p><b>Mô tả:</b> Đây là những thực thể không tồn tại độc lập mà phụ thuộc vào một thực thể chính nào đó.</p> <p><b>Ví dụ:</b> Thực thể <b>DangKyHocPhan</b> (Đăng ký học phần) phụ thuộc vào thực thể <b>SinhVien</b> và <b>HocPhan</b>. Mỗi bản ghi trong <b>DangKyHocPhan</b> sẽ cần có một sinh viên cụ thể và một học phần cụ thể.</p> <p><b>Kết luận:</b> Thực thể phụ thuộc được sinh ra từ một thực thể chính</p>	<p><b>Mô tả:</b> Thực thể này dùng để mô tả mối quan hệ giữa hai hay nhiều thực thể khác.</p> <p><b>Ví dụ:</b> Trong hệ thống quản lý trường học, thực thể <b>GiangDay</b> có thể liên kết giữa thực thể <b>GiangVien</b> (giảng viên) và <b>MonHoc</b> (môn học) để chỉ ra giảng viên nào dạy môn học nào.</p> <p><b>Kết luận:</b> Thực thể liên kết được sinh ra để liên kết các thực thể chính có mối quan hệ N-N với nhau.</p>



# 3. Kỹ thuật xác định Thực thể



Thiết kế  
CSDL

Mô hình hóa  
dữ liệu

KTXĐ  
thực thể

KTXĐ  
thuộc tính

KTXĐ  
quan hệ

KTXĐ khóa  
chính & ngoại

## 6. Cách xác định thực thể

Xác định đối tượng chính	Xác định hành động/sự kiện	Xác định mối quan hệ thực thể
Hãy xem xét các đối tượng chính trong hệ thống cần quản lý. Thông thường, đây sẽ là những thực thể chính.  <b>Ví dụ:</b> Trong hệ thống quản lý thư viện, các thực thể chính có thể bao gồm <b>Sach</b> (sách), <b>DocGia</b> (độc giả), và <b>MuonTra</b> (mượn trả sách).	Các hành động hoặc sự kiện có thể là các thực thể phụ thuộc hoặc thực thể liên kết.  <b>Ví dụ:</b> Trong hệ thống quản lý nhân viên, hành động <b>DiLam</b> (đi làm) có thể là một thực thể lưu trữ thông tin về thời gian làm việc của nhân viên.	Khi đã xác định được các thực thể chính, ta có thể tiếp tục xác định các thực thể liên kết dựa trên mối quan hệ giữa các thực thể chính.  <b>Ví dụ:</b> Nếu đã có <b>SinhVien</b> và <b>MonHoc</b> , thì <b>DangKyHocPhan</b> là một thực thể liên kết giữa hai thực thể này.
Trong một hệ thống quản lý bán hàng: <b>Thực thể chính:</b>  KhachHang (Khách hàng) SanPham (Sản phẩm)	Trong một hệ thống quản lý bán hàng: <b>Thực thể phụ thuộc:</b>  HoaDon (Hóa đơn) vì hóa đơn cần có thông tin về khách hàng.	Trong một hệ thống quản lý bán hàng: <b>Thực thể liên kết:</b>  ChiTietHoaDon (Chi tiết hóa đơn) liên kết giữa HoaDon và SanPham để lưu trữ thông tin về những sản phẩm nào được bán trong một hóa đơn cụ thể.



Thiết kế  
CSDLMô hình hóa  
dữ liệuKTXĐ  
thực thểKTXĐ  
thuộc tínhKTXĐ  
quan hệKTXĐ khóa  
chính & ngoại

### 3. Kỹ thuật xác định Thực thể

#### 7. Phân tích mối quan hệ giữa Đối tượng-Lớp-Thực thể-Bảng-Bản ghi

**Lớp với Thực thể:** Cả hai đều là các mô tả tổng quát cho các đối tượng hoặc bản ghi. Lớp định nghĩa các thuộc tính và hành vi của các đối tượng trong một chương trình, trong khi thực thể định nghĩa các thuộc tính của các bản ghi trong cơ sở dữ liệu.

**Đối tượng với Bản ghi:** Đối tượng là một thể hiện cụ thể của một lớp trong lập trình, tương tự như bản ghi là một thể hiện cụ thể của một thực thể trong cơ sở dữ liệu. Cả hai đều đại diện cho một đơn vị dữ liệu cụ thể với các giá trị thuộc tính được xác định.

**Thực thể với Bảng:** Mỗi thực thể trong mô hình dữ liệu thường thành một bảng trong CSDL. Bảng là hiện thực hóa của thực thể, với mỗi bản ghi trong bảng là một thể hiện của thực thể đó.

##### Kết luận:

- Đối tượng là một đơn vị dữ liệu trong Thực thể/Lớp.
- Bản ghi là một đơn vị dữ liệu trong Bảng.
- Thực thể được hiện thực hóa thành Bảng.





# 4. Kỹ thuật xác định Thuộc tính

## 1. Thuộc tính là gì?

Thuộc tính (Attribute) là một đặc điểm hoặc tính chất của một thực thể trong mô hình dữ liệu. Nó mô tả các yếu tố, đặc điểm hoặc thông tin liên quan đến thực thể đó. Trong CSDL, thuộc tính thường được biểu diễn dưới dạng các cột (columns) trong bảng.

**Ví dụ thuộc tính:** Thông tin cá nhân của một Người chính là các thuộc tính:

- Họ tên
- Ngày sinh
- Nghề nghiệp
- Chức danh
- Địa chỉ liên hệ
- Địa chỉ thư điện tử
- Số điện thoại
- Số CMND/CCCD/Hộ chiếu
- ...

# 4. Kỹ thuật xác định Thuộc tính



Thiết kế  
CSDL

Mô hình hóa  
dữ liệu

KTXĐ  
thực thể

KTXĐ  
thuộc tính

KTXĐ  
quan hệ

KTXĐ khóa  
chính & ngoại

## 2. Phân loại thuộc tính

#	Tên	Định nghĩa	Ví dụ
1	Thuộc tính đơn trị (Simple Attribute)	Thuộc tính đơn trị chỉ có thể nhận một giá trị duy nhất cho một đối tượng.	Trong thực thể <b>SinhVien</b> , thuộc tính HO_TEN chỉ chứa một tên duy nhất cho mỗi sinh viên.
2	Thuộc tính đa trị (Multivalued Attribute)	Thuộc tính đa trị có thể nhận một hoặc một vài giá trị cho một đối tượng trong thực thể.	Trong thực thể <b>SinhVien</b> , thuộc tính SO_THICH có thể cho rất nhiều sở thích của mỗi sinh viên.
3	Thuộc tính nguyên tố (Atomic Attribute)	Thuộc tính nguyên tố là thuộc tính không thể chia nhỏ thành các phần riêng biệt nhỏ hơn và có ý nghĩa.	CHIEU_CAO (chiều cao), CAN_NANG (cân nặng)
4	Thuộc tính phức hợp (Composite Attribute)	Thuộc tính phức hợp là thuộc tính có thể phân chia được thành các phần nhỏ hơn, để biểu diễn các thuộc tính cơ bản hơn với các ý nghĩa độc lập.	HO_TEN (họ tên) có thể chia thành: HO (họ), TEN_DEM (tên đệm), TEN (tên)
5	Thuộc tính suy dẫn (Derived Attribute)	Thuộc tính suy dẫn là thuộc tính không phải nhập, được tính thông qua các thuộc tính khác.	TUOI (tuổi) sẽ được tính toán từ NAM_SINH (năm sinh)





# 4. Kỹ thuật xác định Thuộc tính

## 2. Phân loại thuộc tính

#	Tên	Định nghĩa	Ví dụ
6	Thuộc tính lưu trữ (Stored Attribute)	Thuộc tính lưu trữ là thuộc tính mà giá trị của nó phải được nhập vào.	NAM_SINH (năm sinh), HO_TEN (họ tên)
7	Thuộc tính khóa (Key Attribute)	Thuộc tính khóa là thuộc tính mà giá trị của nó là duy nhất cho mỗi đối tượng trong thực thể, giúp phân biệt đối tượng này và đối tượng khác trong cùng một thực thể.	CCCD (căn cước công dân), MA_SINH_VIEN (mã sinh viên), MA_SO_THUE (mã số thuế)





# 4. Kỹ thuật xác định Thuộc tính

## 2. Trường/cột của Bảng?

Trường (Field)/Cột (Column): Trường, hay còn gọi là cột trong bảng cơ sở dữ liệu, là đơn vị dữ liệu nhỏ nhất lưu trữ giá trị của một thuộc tính cụ thể cho mỗi bản ghi trong bảng. Mỗi cột trong bảng đại diện cho một thuộc tính của thực thể mà bảng đó biểu diễn.

### Ví dụ trường/cột:

ThongTinCaNhan – Thực thể							
	Cột 1	Cột 2	Cột 3	Cột 4	Cột 5	Cột 6	Cột 7
	ID	HoTen	NgaySinh	SoDienThoai	GioiTinh	Email	DiaChi
Bản ghi 1	CN1	Nguyen Van A	1990-05-12	0905123456	Nữ	nguyenvana@example.com	123 Đường ABC, Hà Nội
Bản ghi 2	CN2	Tran Thi B	1992-07-20	0912345678	Nam	tranthib@example.com	456 Đường XYZ, TP HCM
Bản ghi 3	CN3	Le Van C	1988-11-15	0987654321	Nữ	levancc@example.com	789 Đường QWE, Đà Nẵng



# 4. Kỹ thuật xác định Thuộc tính

Thiết kế  
CSDLMô hình hóa  
dữ liệuKTXĐ  
thực thểKTXĐ  
thuộc tínhKTXĐ  
quan hệKTXĐ khóa  
chính & ngoại

## 3. Cách xác định thuộc tính:

- Xác định thực thể.
- Xác định các thông tin cần thiết cho thực thể đó.
- Xác định những thông tin có phải là thông tin lưu trữ hay không, ngược lại sẽ là suy diễn.
- Xác định những thông tin lưu trữ xem có phải là thông tin đơn trị hay không, ngược lại sẽ là đa trị.
- Xác định những thông tin đơn trị xem có phải là thông tin nguyên tố hay không, ngược lại sẽ là phức hợp.
- Xác định những thông tin nguyên tố xem có phải là thông tin khóa hay không, ngược lại sẽ là nguyên tố.

## 4. Mối quan hệ giữa Thuộc tính-Trường/Cột?

Các thuộc tính thường trực tiếp trở thành các trường/cột trong bảng cơ sở dữ liệu.

# 5. Kỹ thuật xác định Quan hệ

Thiết kế  
CSDLMô hình hóa  
dữ liệuKTXĐ  
thực thểKTXĐ  
thuộc tínhKTXĐ  
quan hệKTXĐ khóa  
chính & ngoại

## 1. Khái niệm:

Quan hệ thực thể (Entity Relationship) trong mô hình dữ liệu là mối liên kết giữa hai hoặc nhiều thực thể trong hệ thống. Quan hệ thể hiện cách các thực thể liên kết với nhau và tương tác với nhau.

**Ví dụ về quan hệ:** Mỗi quan hệ giữa sinh viên và khóa học, hoặc giữa khách hàng và đơn hàng.

## 2. Phân loại quan hệ:

- Quan hệ 1-1: 1 thực thể kiểu A liên kết với 1 thực thể kiểu B và ngược lại.
  - Quan hệ 1-N: 1 thực thể kiểu A liên kết với nhiều thực thể kiểu B; 1 thực thể kiểu B chỉ liên kết duy nhất với 1 thực thể kiểu A.
  - Quan hệ N-N: 1 thực thể kiểu A liên kết với nhiều thực thể kiểu B và ngược lại
- ⇒ Cách xác định ĐÚNG quan hệ thực thể là tùy thuộc vào NGỮ CẢNH của Doanh nghiệp.
- ⇒ Cách xác định ĐÚNG quan hệ thực thể là dựa vào LUẬT.





# 5. Kỹ thuật xác định Quan hệ

## 2. Xác định quan hệ sai?

- **Dữ liệu bị trùng lặp hoặc thiếu**

Ví dụ: Nếu thiết kế hệ thống với giả định mỗi sinh viên chỉ có thể đăng ký một khóa học (1:1), trong khi thực tế sinh viên có thể đăng ký nhiều khóa học (1:N), thì sẽ không thể lưu trữ đầy đủ dữ liệu về các khóa học mà sinh viên đã đăng ký.

- **Khó khăn trong việc truy vấn và báo cáo**

Ví dụ: Nếu mối quan hệ giữa Giáo viên và Lớp học bị hiểu nhầm là 1:1 thay vì 1:N, việc truy vấn dữ liệu để tìm tất cả các lớp học do một giáo viên giảng dạy sẽ trở nên khó khăn hoặc không thể thực hiện đúng.

- **Khó khăn trong mở rộng và bảo trì hệ thống**

Ví dụ: Nếu ban đầu hệ thống được thiết kế với quan hệ 1:1 giữa SinhVien và LopHoc, nhưng sau đó cần mở rộng để hỗ trợ phát triển hệ thống, việc thay đổi sẽ đòi hỏi sửa đổi toàn bộ cấu trúc bảng và có thể ảnh hưởng đến các chức năng hiện có.





# 6. Kỹ thuật xác định Khóa chính và khóa ngoại

## 1. Khóa chính (Primary Key)

**Khái niệm:** Khóa chính (Primary Key) là một thuộc tính của một bảng trong CSDL, được sử dụng để xác định duy nhất mỗi bản ghi trong bảng đó. Mỗi bảng trong cơ sở dữ liệu chỉ có một khóa chính và giá trị của khóa chính phải là duy nhất và không thể rỗng (NULL) để đảm bảo tính toàn vẹn của dữ liệu.

**Ví dụ minh họa:** Bảng SinhVien

Cột ID được sử dụng làm khóa chính. Mỗi giá trị của ID là duy nhất cho mỗi sinh viên.

ID	HoTen	NgaySinh
1	Nguyen Văn A	1990-05-12
2	Trần Thị B	1992-07-20
3	Lê Văn C	1988-11-15





# 6. Kỹ thuật xác định Khóa chính và khóa ngoại

## 2. So sánh khóa chính (Primary Key) và khóa duy nhất (Unique Key)

**Khái niệm:** Khóa duy nhất là một thuộc tính của một bảng mà giá trị của nó phải là duy nhất cho mỗi bản ghi trong bảng. Khóa duy nhất có thể chấp nhận giá trị NULL, nhưng mỗi giá trị không được trùng lặp với các giá trị khác trong cùng cột.

#	Khía cạnh	Khóa chính (Primary Key)	Khóa duy nhất (Unique Key)
1	Số lượng trường khóa chính	Một trường khóa chính trong 1 bảng.	Nhiều trường khóa duy nhất trong 1 bảng.
2	Giá trị NULL	Không có	Có (tùy hệ quản trị CSDL)
3	Giá trị trong trường	Mỗi giá trị phải là duy nhất	Mỗi giá trị phải là duy nhất
4	Tính toàn vẹn	Đảm bảo tính duy nhất và toàn vẹn của mỗi bản ghi trong bảng.	Đảm bảo tính duy nhất và toàn vẹn của mỗi bản ghi trong bảng.
5	Chỉ mục (Index)	Tự động tạo chỉ mục duy nhất.	Tự động tạo chỉ mục duy nhất.





# 6. Kỹ thuật xác định Khóa chính và khóa ngoại

## 3. Khóa ngoại (Foreign Key)

**Khái niệm:** Khóa ngoại (Foreign Key) là một thuộc tính trong một bảng, dùng để tạo mối liên kết với khóa chính của một bảng khác. Khóa ngoại giúp duy trì tính toàn vẹn của dữ liệu bằng cách đảm bảo rằng giá trị của khóa ngoại phải tồn tại trong bảng mà nó liên kết.

### Ví dụ minh họa:

Trong bảng **LopHoc**, cột **GiangVienID** là khóa ngoại tham chiếu đến cột **ID** trong bảng **GiangVien**. Điều này đảm bảo rằng mỗi lớp học phải có một giáo viên hợp lệ từ bảng **GiangVien**.

Table: GiangVien		Table: LopHoc		
ID	HoTen	ID	TenLop	GiangVienID
GV1	Nguyễn Văn A	LOP1	Toán A	GV1
GV2	Trần Thị B	LOP2	Hóa B	GV2
GV3	Lê Văn C	LOP3	Toán C	GV1



# 6. Kỹ thuật xác định Khóa chính và khóa ngoại

Thiết kế  
CSDLMô hình hóa  
dữ liệuKTXĐ  
thực thểKTXĐ  
thuộc tínhKTXĐ  
quan hệKTXĐ khóa  
chính & ngoại

## 4. Xác định khóa chính (Primary Key)

Xác định thuộc tính duy nhất: Chọn thuộc tính hoặc tập hợp thuộc tính có thể duy nhất xác định mỗi bản ghi trong bảng. Thuộc tính này không được trùng lặp và không thể có giá trị NULL.

**Ví dụ:** Trong bảng **SinhVien**, **ID** là khóa chính vì mỗi sinh viên có một **ID** duy nhất và không có giá trị NULL.

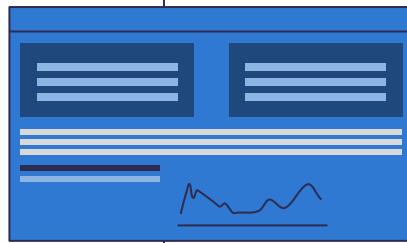
## 5. Xác định khóa ngoại (Foreign Key)

Xác định mối quan hệ giữa các bảng: Xác định thuộc tính trong một bảng cần liên kết với khóa chính của bảng khác. Đây thường là thuộc tính dùng để thiết lập mối liên kết giữa các bảng.

Chắc chắn rằng giá trị của khóa ngoại tồn tại trong bảng mà nó tham chiếu: Điều này đảm bảo tính toàn vẹn của dữ liệu và tránh các giá trị không hợp lệ.

**Ví dụ:** Trong bảng **LopHoc**, **GiangVienID** là khóa ngoại tham chiếu đến **ID** trong bảng **GiangVien**. Điều này tạo mối liên kết giữa lớp học và giáo viên.





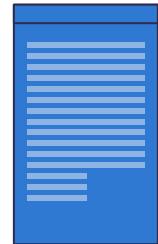
03

# Chuẩn hóa CSDL quan hệ





# Chuẩn hóa CSDL quan hệ



1. Tổng quan

3. Chuẩn hóa 2NF

2. Chuẩn hóa 1NF

4. Chuẩn hóa 3NF

Các mục  
chuẩn hóa CSDL  
quan hệ



# 1. Tổng quan



Tổng quan

Chuẩn hóa 1NF

Chuẩn hóa 2NF

Chuẩn hóa 3NF

## 1. Chuẩn hóa lược đồ CSDL quan hệ là gì?

**Khái niệm:** Chuẩn hóa (Normalization) lược đồ CSDL quan hệ là quá trình tổ chức các bảng trong cơ sở dữ liệu để giảm thiểu sự dư thừa và tránh những vấn đề bất thường trong dữ liệu (data anomalies) như thêm, xóa, hay cập nhật. Quá trình này bao gồm việc phân tách các bảng lớn thành các bảng nhỏ hơn và thiết lập mối quan hệ giữa chúng theo một cách hợp lý.

### Mục tiêu:

1. Giảm thiểu sự dư thừa dữ liệu (Data Redundancy)
2. Đảm bảo tính toàn vẹn và nhất quán của dữ liệu (Data Integrity and Consistency)
3. Loại bỏ các vấn đề bất thường trong cơ sở dữ liệu (Eliminate Data Anomalies)
4. Tối ưu hóa việc truy vấn và bảo trì cơ sở dữ liệu



## Tổng quan



## 1. Tổng quan

## 1. Chuẩn hóa lược đồ CSDL quan hệ là gì?

Ví dụ:

MaSV	TenSV	LopHoc	GiangVien
SV001	Nguyễn Văn A	Toán A	Nguyễn Văn C
SV002	Trần Thị B	Hóa B	Lê Thị D
SV003	Nguyễn Văn A	Lý C	Nguyễn Văn C

Vấn đề trong bảng này:

- Dữ thừa dữ liệu: Tên giảng viên Nguyen Van C được lặp lại nhiều lần.
- Khó bảo trì: Nếu tên giảng viên Nguyen Van C cần thay đổi, cần cập nhật tất cả các dòng liên quan.
- Bất thường khi xóa: Nếu sinh viên Nguyen Van A rời khỏi lớp Toan A, thông tin về giảng viên Nguyen Van C có thể bị mất.

## Chuẩn hóa 2NF

## Chuẩn hóa 3NF

## 2. Chuẩn hóa 1NF



Tổng quan

Chuẩn hóa 1NF

Chuẩn hóa 2NF

Chuẩn hóa 3NF

1NF (First Normal Form) là cấp độ chuẩn hóa đầu tiên trong quá trình chuẩn hóa cơ sở dữ liệu quan hệ.

Một bảng được coi là đạt 1NF nếu nó thỏa mãn các điều kiện sau:

- **Mỗi cột chỉ chứa các giá trị đơn trị:** Tức là mỗi cột trong bảng chỉ chứa một giá trị duy nhất và không phải là một danh sách, mảng hoặc tập hợp các giá trị.
- **Mỗi cột chỉ chứa một loại dữ liệu:** Tất cả các giá trị trong cùng một cột phải thuộc cùng một kiểu dữ liệu. Ví dụ: Nếu một cột là kiểu số, tất cả các giá trị trong cột đó phải là số.
- **Mỗi bản ghi là duy nhất:** Không có hai bản ghi nào trong bảng có thể giống hệt nhau về tất cả các giá trị trong các cột. Thông thường, để đảm bảo điều này, bảng sẽ có một khóa chính.
- **Mỗi cột phải có một tên duy nhất:** Không có cột nào trong bảng có thể có tên trùng lặp với cột khác.



## 2. Chuẩn hóa 1NF

Tổng quan



Ví dụ:

Bảng **KhachHang** không đạt 1NF

<b>MaKH</b>	<b>TenKH</b>	<b>SoDienThoai</b>
001	Nguyen Văn A	0123456789, 0987654321
002	Trần Thị B	0912345678

Chuẩn hóa 1NF

Bảng **KhachHang** đạt 1NF

<b>MaKH</b>	<b>TenKH</b>	<b>SoDienThoai</b>
001	Nguyen Văn A	0123456789
001	Nguyen Văn A	0987654321
002	Trần Thị B	0912345678

Chuẩn hóa 2NF

Chuẩn hóa 3NF

### 3. Chuẩn hóa 2NF



Tổng quan

Chuẩn hóa 1NF

Chuẩn hóa 2NF

Chuẩn hóa 3NF

2NF (Second Normal Form) là cấp độ thứ hai trong quá trình chuẩn hóa cơ sở dữ liệu quan hệ.

Một bảng được coi là đạt 2NF nếu nó thỏa mãn các điều kiện sau:

- **Bảng đã đạt 1NF:** Tức là bảng phải ở dạng chuẩn thứ nhất (1NF), nghĩa là tất cả các cột chứa giá trị nguyên tố và mỗi bản ghi là duy nhất.
- **Mọi thuộc tính không khóa đều phụ thuộc hoàn toàn vào toàn bộ khóa chính:** Điều này có nghĩa là tất cả các thuộc tính không thuộc về khóa chính (non-prime attributes) phải phụ thuộc đầy đủ vào khóa chính. Không được phép có thuộc tính không khóa nào phụ thuộc vào một phần của khóa chính (partial dependency).



### 3. Chuẩn hóa 2NF

Tổng quan



Ví dụ:

Bảng **DangKyHocPhan** không đạt 2NF

<b>MaSV</b>	<b>MaHP</b>	<b>TenHP</b>	<b>Diem</b>
SV001	HP01	Toán Cao Cấp	8.5
SV002	HP02	Cơ sở dữ liệu 1	7.0
SV003	HP02	Cơ sở dữ liệu 1	9.0

Bảng **DangKyHocPhan** đạt 2NF

<b>MaSV</b>	<b>MaHP</b>	<b>Diem</b>
SV001	HP01	8.5
SV002	HP02	7.0
SV003	HP02	9.0

## 4. Chuẩn hóa 3NF



Tổng quan

Chuẩn hóa 1NF

Chuẩn hóa 2NF

Chuẩn hóa 3NF

3NF (Third Normal Form) là cấp độ thứ ba trong quá trình chuẩn hóa cơ sở dữ liệu quan hệ.

Một bảng được coi là đạt 3NF nếu nó thỏa mãn các điều kiện sau:

- **Bảng đã đạt 2NF:** Bảng phải thỏa mãn tất cả các điều kiện của 1NF và 2NF.
- **Không có phụ thuộc bắc cầu (transitive dependency):** Không có thuộc tính không khóa nào phụ thuộc vào một thuộc tính không khóa khác. Nói cách khác, mọi thuộc tính không khóa chỉ phụ thuộc trực tiếp vào khóa chính.

# 4. Chuẩn hóa 3NF

Tổng quan

Chuẩn hóa 1NF

Chuẩn hóa 2NF

Chuẩn hóa 3NF



Ví dụ:

Bảng **NhanVien** không đạt 3NF

<b>MaNV</b>	<b>TenNV</b>	<b>MaPhongBan</b>	<b>TenPhongBan</b>	<b>DiaChiPhongBan</b>
NV001	Nguyen Van A	PB01	Phòng Kế toán	123 Đinh Tiên Hoàng
NV002	Tran Thi B	PB02	Phòng Nhân sự	456 Lê Lợi
NV003	Le Van C	PB01	Phòng Kế toán	123 Đinh Tiên Hoàng

Bảng **NhanVien** đạt 3NF

<b>MaNV</b>	<b>TenNV</b>	<b>MaPhongBan</b>
NV001	Nguyen Van A	PB01
NV002	Tran Thi B	PB02
NV003	Le Van C	PB01



BATIZENS



*Trân trọng cảm ơn!*

