

Cơ sở dữ liệu

Tin học · Lớp 11

Giáo viên: Phúc Hảo

- ▶ Hiểu được khái niệm **cơ sở dữ liệu** và **cơ sở dữ liệu quan hệ**.
- ▶ Nhận biết các thành phần: **bảng, bản ghi, thuộc tính, khoá chính, khoá ngoại**.
- ▶ Trình bày được đặc trưng của **mô hình dữ liệu quan hệ**.
- ▶ Đọc và thiết lập được **quan hệ giữa các bảng** qua sơ đồ ERD.

Ý chính: Dữ liệu lộn xộn sẽ khiến việc tìm kiếm trở thành cơn ác mộng.

Hãy tưởng tượng thư viện trường có **5000 quyển sách**, nhưng tất cả thông tin (tên sách, tác giả, ai đang mượn) chỉ ghi lộn xộn vào một cuốn sổ tay.

Tình huống

Cô thủ thư muốn biết: *“Bạn Lan đang mượn những quyển nào?”*
Với cuốn sổ tay, cô phải lật từng trang, dò từng dòng... mất cả buổi!

Câu hỏi: Làm thế nào để lưu trữ dữ liệu một cách **có tổ chức**, giúp tìm kiếm nhanh, tránh trùng lặp và sai sót? $\Rightarrow$ Đó là lí do ta cần **Cơ sở dữ liệu**.

1

Khái niệm cơ sở dữ liệu

Cơ sở dữ liệu là gì?

Định nghĩa

Cơ sở dữ liệu (Database) là một tập hợp các dữ liệu có **liên quan với nhau**, được tổ chức và lưu trữ theo một **cấu trúc nhất định**, nhằm phục vụ cho nhiều người dùng và nhiều mục đích khác nhau.

Liên hệ thực tế

Danh bạ điện thoại, sổ điểm điện tử của trường, danh sách khách hàng của một cửa hàng... đều là các cơ sở dữ liệu quen thuộc.

Ghi nhớ

Dữ liệu trong CSDL không đứng riêng lẻ — chúng **có tổ chức** và **liên quan** nhau. Đây là điểm khác biệt so với việc ghi chép tùy ý.

Định nghĩa

Hệ quản trị CSDL (DBMS) là phần mềm giúp **tạo lập, cập nhật, khai thác** và **quản lí** cơ sở dữ liệu.

Ví dụ: MySQL, SQL Server, Access, PostgreSQL...

Liên hệ thực tế

DBMS giống như **người thủ thư**: ta chỉ cần yêu cầu, thủ thư sẽ tìm, thêm, xoá, sắp xếp sách giúp mà không cần ta tự lục kho.

Ý chính: CSDL là *dữ liệu*; DBMS là *phần mềm* quản lí dữ liệu đó.

2

Cơ sở dữ liệu quan hệ

Định nghĩa

Cơ sở dữ liệu quan hệ là CSDL lưu trữ dữ liệu dưới dạng các **bảng** (quan hệ). Giữa các bảng có thể thiết lập **mối liên kết** với nhau.

Liên hệ thực tế

Mỗi bảng giống một **tờ danh sách kẻ ô**: hàng ngang là từng đối tượng, cột dọc là từng đặc điểm. Đây là mô hình phổ biến nhất hiện nay.

Ý chính: “Quan hệ” = **bảng dữ liệu**. Nhiều bảng liên kết nhau tạo thành CSDL quan hệ.

Các thành phần của một bảng

Một **bảng** gồm:

- ▶ **Thuộc tính** (cột): một đặc điểm của đối tượng, ví dụ *HoTen*, *Lop*.
- ▶ **Bản ghi** (hàng): thông tin đầy đủ của một đối tượng.
- ▶ **Miền giá trị**: tập giá trị hợp lệ của một thuộc tính.

Ghi nhớ

Cột = **thuộc tính**

Hàng = **bản ghi**

Ô = một giá trị cụ thể.

MaHS	HoTen	Lop
HS01	Nguyễn Văn An	11A1
HS02	Trần Thị Bình	11A2

Khoá chính (Primary Key)

Định nghĩa

Khoá chính là một (hoặc một nhóm) thuộc tính dùng để **phân biệt duy nhất** từng bản ghi trong bảng. Giá trị khoá chính **không được trùng** và **không để trống**.

Liên hệ thực tế

Khoá chính giống **số CCCD** hay **mã học sinh**: mỗi người một mã riêng, không ai giống ai.

Ghi nhớ — lỗi thường gặp

Không chọn thuộc tính *có thể trùng* làm khoá chính (ví dụ chọn **HoTen**) — hai người có thể trùng tên!

Định nghĩa

Khoá ngoại là thuộc tính trong một bảng, mà giá trị của nó **tham chiếu đến khoá chính** của một bảng khác. Nhờ khoá ngoại, các bảng được **liên kết** với nhau.

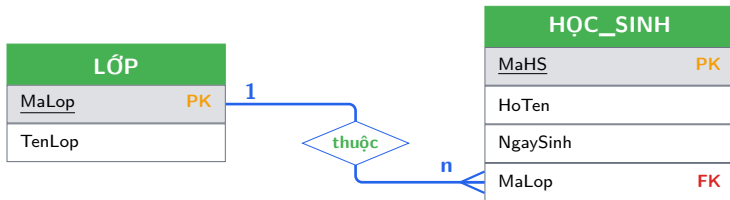
Liên hệ thực tế

Trong phiếu mượn sách, cột *MaHS* chính là “cầu nối” trở về bảng HỌC_SINH để biết ai đang mượn.

Ý chính: Khoá chính *định danh* bản ghi; khoá ngoại *nối* các bảng lại với nhau.

Sơ đồ thực thể - quan hệ (ERD)

Sơ đồ thực thể - quan hệ (ERD)



PK = khoá chính (gạch chân) | **FK** = khoá ngoại | **1-n** = quan hệ một - nhiều

- ▶ Dữ liệu được trình bày dưới dạng các **bảng** (hàng - cột).
- ▶ Thứ tự **bản ghi** và **thuộc tính** là **tuỳ ý**, không ảnh hưởng ý nghĩa dữ liệu.
- ▶ Mỗi bảng có **tên phân biệt**; các thuộc tính trong một bảng **không trùng tên**.
- ▶ Mỗi bảng có một **khoá chính**; giữa các bảng liên kết qua **khoá ngoại**.

Ghi nhớ

Không có **hai bản ghi giống hệt nhau** trong cùng một bảng — nhờ khoá chính bảo đảm điều này.

- ▶ **Giảm dư thừa** và trùng lặp dữ liệu.
- ▶ **Bảo đảm nhất quán**: sửa một nơi, đúng khắp mọi nơi.
- ▶ **Tìm kiếm nhanh**, dễ khai thác.
- ▶ **Chia sẻ** cho nhiều người dùng cùng lúc.

Liên hệ thực tế

Đổi tên một lớp: chỉ cần sửa *một dòng* trong bảng LỚP, tất cả học sinh liên kết đều tự đúng theo — không phải sửa từng người.

Ý chính: Liên kết bảng giúp dữ liệu **gọn, đúng và dễ quản lí**.

3

Ví dụ minh họa

Ví dụ 1: Xác định khoá chính, khoá ngoại

Cho bảng **DIEM** lưu điểm học sinh:

MaHS	MaMH	Diem
HS01	TOAN	8.5
HS01	LI	7.0
HS02	TOAN	9.0

Yêu cầu: Xác định khoá chính và khoá ngoại của bảng DIEM.

Lời giải từng bước

Bước 1. Tìm thuộc tính phân biệt được từng bản ghi. Chỉ *MaHS* hay chỉ *MaMH* đều bị trùng $\Rightarrow$ không dùng riêng lẻ được.

Bước 2. Kết hợp $\{MaHS, MaMH\}$: mỗi cặp là duy nhất $\Rightarrow$ đây là **khóá chính** (khóá gồm nhiều thuộc tính).

Bước 3. *MaHS* trở về bảng HỌC_SINH, *MaMH* trở về bảng MON_HOC $\Rightarrow$ cả hai là **khóá ngoại**.

Ví dụ 2: Thiết lập liên kết giữa hai bảng

Đề bài: Cửa hàng có bảng **KHACH_HANG**(MaKH, HoTen, SDT) và bảng **DON_HANG**(MaDH, NgayDat, MaKH). Hãy chỉ ra cách liên kết hai bảng và lí do.

Lời giải từng bước

Bước 1 — Xác định khoá chính mỗi bảng.

KHACH_HANG có khoá chính **MaKH**; DON_HANG có khoá chính **MaDH**. Vì sao? Vì chúng định danh duy nhất từng khách hàng / đơn hàng.

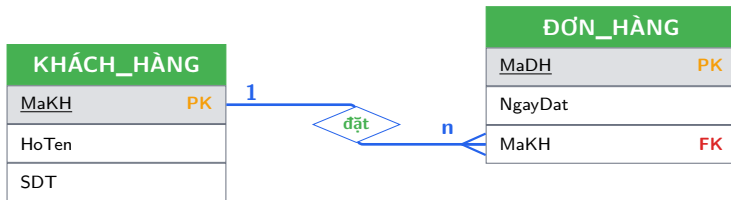
Bước 2 — Tìm cầu nối.

Trong DON_HANG có thuộc tính **MaKH** trùng kiểu với khoá chính của KHACH_HANG $\Rightarrow$ đây là **khoá ngoại**.

Bước 3 — Kết luận.

Liên kết **MaKH (FK)** $\rightarrow$ **MaKH (PK)**. Đây là quan hệ **1 - n**: một khách hàng có thể có nhiều đơn hàng.

Liên kết KHÁCH_HÀNG - ĐƠN_HÀNG



PK = khoá chính (gạch chân) | **FK** = khoá ngoại | **1 - n** = quan hệ một - nhiều

Ghi nhớ tổng hợp

- ▶ **CSDL**: tập dữ liệu có tổ chức, liên quan nhau; **DBMS** là phần mềm quản lí.
- ▶ **CSDL quan hệ**: lưu dữ liệu theo **bảng**; cột = thuộc tính, hàng = bản ghi.
- ▶ **Khoá chính (PK)**: định danh duy nhất mỗi bản ghi.
- ▶ **Khoá ngoại (FK)**: nối bảng này với khoá chính bảng khác.

Ý chính: Hiểu PK và FK là hiểu được “xương sống” của mọi cơ sở dữ liệu quan hệ.

Luyện tập nhanh · Nhận biết

Trong cơ sở dữ liệu quan hệ, dữ liệu được tổ chức và lưu trữ chủ yếu dưới dạng nào?

- A. Các bảng (quan hệ) gồm hàng và cột
- B. Các tệp văn bản rời rạc
- C. Các sơ đồ cây phân cấp
- D. Các đoạn video và hình ảnh

Đáp án đúng: **A. Các bảng (quan hệ) gồm hàng và cột.**

CSDL quan hệ lưu dữ liệu trong các **bảng** (mỗi bảng là một quan hệ). B là kho tệp thường (không có cấu trúc bảng), C nhằm sang mô hình phân cấp, D chỉ là kiểu dữ liệu chứa trong bảng chứ không phải cách tổ chức.

MãHS	Tên	Lớp
HS01	An	
HS02	Bình	
HS03	Chi	

Luyện tập nhanh · Nhận biết

Trong bảng bên, mỗi **hàng** (ví dụ: HS01 - An - ...) được gọi là gì?

- A. Một bản ghi (bộ)
- B. Một thuộc tính
- C. Một khóa chính
- D. Một quan hệ

Đáp án đúng: **A. Một bản ghi (bộ).**

Mỗi **hàng** là một bản ghi (record/bộ) mô tả một đối tượng. B (thuộc tính) là **cột**, dễ nhầm hàng $\leftrightarrow$ cột. C là một cột đặc biệt dùng để phân biệt, D là cả bảng.

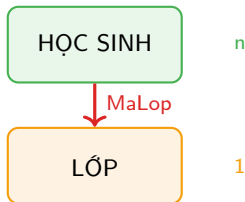
Luyện tập nhanh · Nhận biết

Khóa chính (primary key) của một bảng có đặc điểm nào sau đây?

- A. Giá trị xác định duy nhất mỗi bản ghi và không được trùng lặp
- B. Có thể trùng nhau giữa nhiều bản ghi
- C. Bắt buộc phải là kiểu số nguyên
- D. Là hàng đầu tiên của bảng

Đáp án đúng: **A. Xác định duy nhất mỗi bản ghi, không trùng lặp.**

Khóa chính phải **duy nhất** và **khác rỗng**. B mâu thuẫn tính duy nhất, C sai vì khóa chính có thể là chuỗi (mã HS...), D nhầm khóa chính là một **cột** chứ không phải một hàng.



Luyện tập nhanh · Thông hiểu

Cột MaLop trong bảng HỌC SINH tham chiếu đến khóa chính của bảng LỚP.

MaLop là loại khóa nào?

- A. Khóa ngoại (foreign key)
- B. Khóa chính
- C. Khóa duy nhất riêng của HỌC SINH
- D. Không phải là khóa

Đáp án đúng: **A. Khóa ngoài (foreign key).**

Cột dùng để **liên kết** tới khóa chính của bảng khác là khóa ngoài. B sai vì khóa chính của HỌC SINH là MãHS, C nhầm vì MaLop được phép trùng lặp (nhiều HS cùng lớp), D bỏ sót vai trò liên kết.

Luyện tập nhanh · Thông hiểu

Vì sao nên tách thông tin lớp ra thành bảng LỚP riêng, thay vì ghi tên lớp lặp lại trong mọi bản ghi HỌC SINH?

- A. Tránh trùng lặp dữ liệu, giảm sai sót khi cập nhật
- B. Giúp bảng HỌC SINH có nhiều cột hơn
- C. Làm cho truy vấn luôn chạy chậm hơn
- D. Bắt buộc theo quy định, không có lợi ích thực tế

Đáp án đúng: **A. Tránh trùng lặp dữ liệu, giảm sai sót khi cập nhật.**

Tách bảng (chuẩn hóa) giúp mỗi thông tin lưu **một lần**. B ngược mục tiêu (giảm dư thừa), C sai vì liên kết bằng khóa còn giúp quản lý tốt hơn, D phủ nhận lợi ích thực tế của chuẩn hóa.

MãHS	Điểm
HS01	8
HS02	5
HS03	9
HS04	4

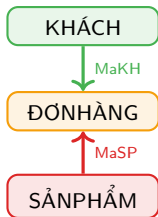
Luyện tập nhanh · Thông hiểu

Truy vấn “chọn các bản ghi có Điểm ≥ 5 ” trả về bao nhiêu bản ghi?

- A. 2
- B. 4
- C. 3
- D. 1

Đáp án đúng: **C. 3 bản ghi** (HS01=8, HS02=5, HS03=9).

Vì $5 \geq 5$ nên HS02 vẫn **thỏa** điều kiện. Các phương án nhiễu: A quên rằng dấu $\geq$ bao gồm giá trị bằng 5, B đếm cả bảng, D bỏ sót HS03.



Luyện tập nhanh · Vận dụng

Để biết “khách hàng nào đã mua sản phẩm nào”, ta cần kết nối các bảng qua đâu?

- A. Qua các khóa ngoài MaKH và MaSP trong ĐƠNHÀNG
- B. Chỉ cần bảng KHÁCH
- C. Chỉ cần bảng SẢN PHẨM
- D. Không thể kết nối được ba bảng

Đáp án đúng: **A. Qua khóa ngoài MaKH và MaSP trong bảng ĐƠNHÀNG.**

Bảng ĐƠNHÀNG đóng vai trò **trung gian** nối KHÁCH với SẢN PHẨM. B, C chỉ dùng một bảng nên thiếu thông tin liên kết, D sai vì đây chính là mục đích của khóa ngoài trong CSDL quan hệ.

Luyện tập nhanh · Vận dụng

Một trường học muốn quản lý MƯỢN SÁCH: một học sinh mượn nhiều sách, một cuốn sách có thể được nhiều học sinh mượn (khác thời điểm). Thiết kế nào hợp lí nhất?

- A. 3 bảng: HỌC SINH, SÁCH và PHIẾU MƯỢN (chứa Mã HS, Mã Sách, ngày)
- B. 1 bảng duy nhất chứa toàn bộ thông tin, lặp lại nhiều lần
- C. 2 bảng HỌC SINH và SÁCH, không cần bảng liên kết
- D. Lưu tên HS trực tiếp vào bảng SÁCH

Đáp án đúng: **A. 3 bảng, có PHIẾUMƯỢN làm bảng liên kết.**

Quan hệ “nhiều - nhiều” cần bảng trung gian. B gây **dư thừa** và sai sót lớn, C thiếu nơi lưu lượt mượn cụ thể, D vi phạm vì một sách nhiều người mượn không thể nhét vào một ô.

1. Cho bảng HỌC SINH (Mã HS, Tên, Lớp, Điểm): xác định khóa chính, đề xuất một truy vấn lọc theo điều kiện điểm.
2. Vẽ sơ đồ liên kết cho hai bảng LỚP và HỌC SINH, chỉ rõ khóa chính và khóa ngoài.
3. Chỉ ra một trường hợp dữ liệu bị trùng lặp và cách tách bảng để khắc phục.

Gợi ý

Khóa chính phải **duy nhất, khác rỗng**. Khóa ngoài là cột ở bảng này **tham chiếu** khóa chính bảng kia. Khi thấy thông tin lặp lại ở nhiều bản ghi $\Rightarrow$ nên tách thành bảng riêng.

1. Thiết kế 3 bảng cho quản lý điểm: HỌCSINH, MÔN HỌC, ĐIỂM; nêu rõ khóa chính, khóa ngoài.
2. Viết bằng lời 3 truy vấn khác nhau trên CSDL vừa thiết kế.
3. Cho ví dụ một bảng có dữ liệu dư thừa và trình bày cách chuẩn hóa.
4. Phân biệt quan hệ “một - nhiều” và “nhiều - nhiều” kèm ví dụ thực tế.
5. Liệt kê 3 lợi ích của việc dùng khóa để liên kết các bảng.

Đọc trước bài sau

Tìm hiểu về **hệ quản trị cơ sở dữ liệu (DBMS)** và ngôn ngữ truy vấn dữ liệu.

Bảng tóm tắt: Khái niệm cốt lõi

Thuật ngữ	Vị trí trong bảng	Ý nghĩa
Quan hệ	Cả bảng	Tập dữ liệu về một đối tượng
Bản ghi (bộ)	Một hàng	Thông tin một đối tượng cụ thể
Thuộc tính	Một cột	Một loại thông tin
Khóa chính	Một/vài cột	Xác định duy nhất bản ghi
Khóa ngoài	Một cột	Liên kết tới bảng khác

Bảng tóm tắt: Loại quan hệ giữa các bảng

Loại	Ví dụ	Cách cài đặt
Một - một	Người $\leftrightarrow$ CCCD	Khóa ngoài + ràng buộc duy nhất
Một - nhiều	Lớp $\rightarrow$ Học sinh	Khóa ngoài ở bảng “nhiều”
Nhiều - nhiều	Học sinh $\leftrightarrow$ Sách	Bảng trung gian (2 khóa ngoài)

Ghi nhớ: quan hệ **nhiều - nhiều** luôn cần một **bảng liên kết**.

- ▶ CSDL quan hệ tổ chức dữ liệu thành các **bảng** gồm hàng (bản ghi) và cột (thuộc tính).
- ▶ **Khóa chính** xác định duy nhất mỗi bản ghi; **khóa ngoài** tạo liên kết giữa các bảng.
- ▶ Tách bảng (chuẩn hóa) giúp **giảm dư thừa**, tránh sai sót khi cập nhật dữ liệu.
- ▶ Truy vấn cho phép **lọc và kết nối** dữ liệu từ nhiều bảng để trả lời câu hỏi.

Liên hệ: Các ứng dụng quen thuộc như quản lý điểm, ngân hàng, đặt vé... đều vận hành trên CSDL quan hệ.

Cảm ơn các em đã lắng nghe!

Bài giảng biên soạn với EduFlow · agentra.io.vn
