

TRƯỜNG ĐÀO TẠO LẬP TRÌNH VIÊN VÀ QUẢN TRỊ MẠNG QUỐC TẾ BACHKHOA-APTECH

Bài 01

Giới thiệu RDBMS

Tóm tắt

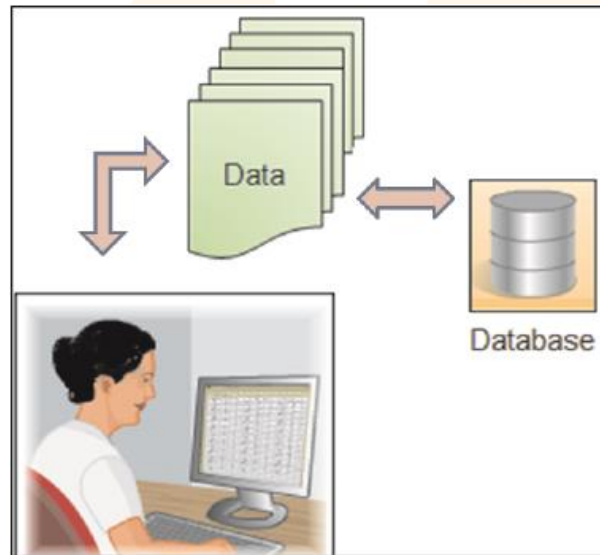
- Giải thích khái niệm dữ liệu và CSDL
- Các mô hình CSDL
- Hệ Quản trị CSDL (DBMS)
- Hệ Quản trị CSDL Quan hệ (RDBMS)
- Mô tả thực thể và bảng



Khái niệm dữ liệu và CSDL

Dữ liệu

- Các tổ chức/cơ quan/doanh nghiệp phải **duy trì** một lượng lớn dữ liệu **phát sinh hàng ngày** trong quá trình hoạt động.
- Dữ liệu là **thông tin có ý nghĩa**, có thể chứa nhiều giá trị.



Khái niệm dữ liệu và CSDL

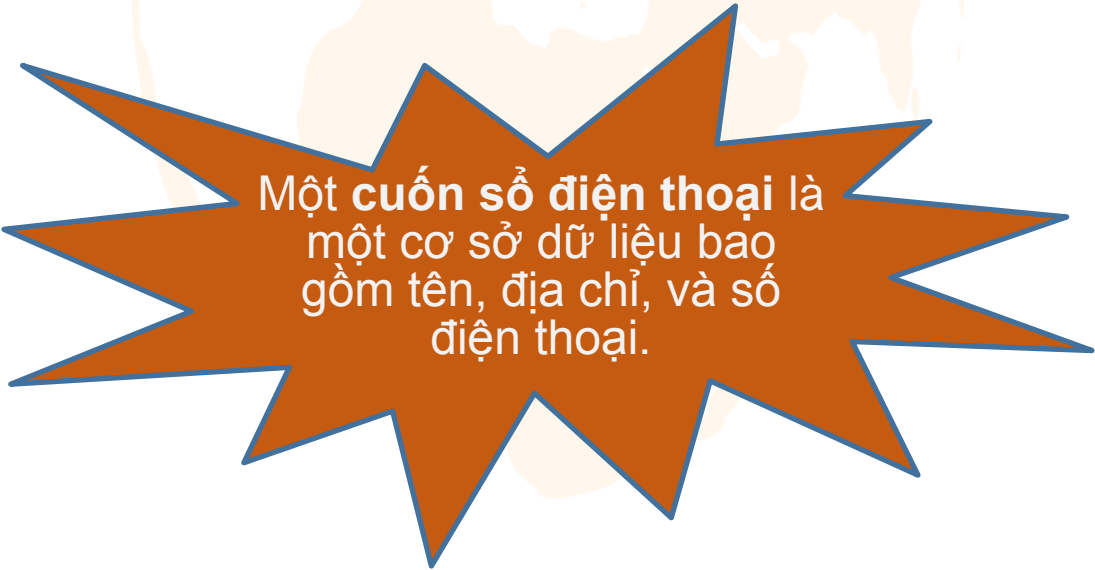
Ví dụ:

- ✓ **Thông tin về mỗi trận đấu** bóng đá trong một mùa giải WorldCup được lưu lại gồm có: **tên đội** tham gia, **tỷ số**, danh sách **cầu thủ**, **địa điểm** thi đấu, **thời gian**...
- ✓ Dựa vào các dữ liệu (thông tin của mỗi trận đấu đã được lưu lại) chúng ta có thể **biết được thông tin** về 10 **trận đấu hay nhất**, các **đội chơi hay**, **cầu thủ xuất sắc**,... trong kỳ WorldCup

Khái niệm dữ liệu và CSDL

Cơ sở dữ liệu (CSDL)

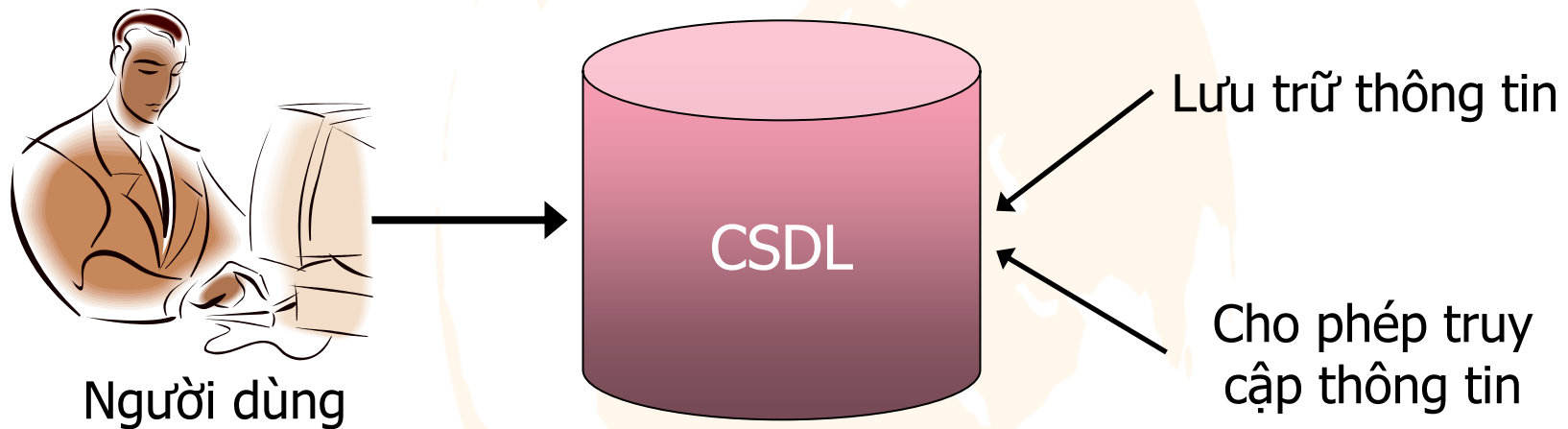
- Là tập hợp dữ liệu được **tổ chức có cấu trúc** liên quan với nhau và được **lưu trữ** trong máy tính.
- Được **thiết kế**, xây dựng cho phép người dùng lưu trữ dữ liệu, **truy xuất** thông tin hoặc **cập nhật** dữ liệu...



Một cuốn sổ điện thoại là một cơ sở dữ liệu bao gồm tên, địa chỉ, và số điện thoại.

Khái niệm dữ liệu và CSDL

Minh họa CSDL đơn giản



Khái niệm dữ liệu và CSDL

Ví dụ




Khái niệm dữ liệu và CSDL

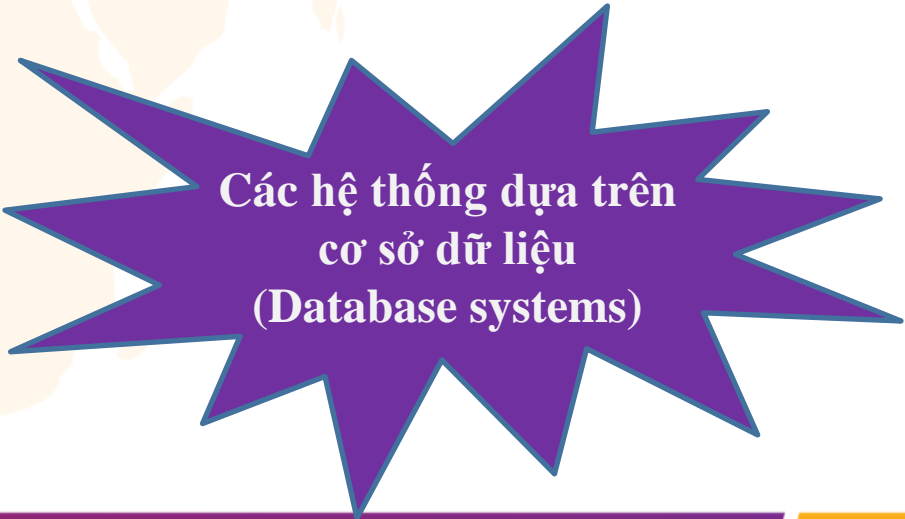
Quản trị Dữ liệu: là đề cập đến việc quản lý một khối lượng lớn thông tin bao gồm:

- Lưu trữ.
- Các kỹ thuật để thao tác với thông tin
- Đảm bảo thông tin được an toàn, chính xác.

Có 2 cách tiếp cận:

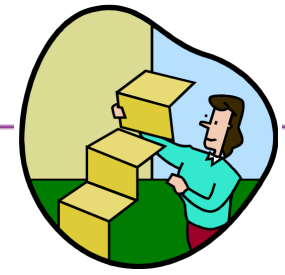


Các hệ thống dựa trên
file (File-based systems)



Các hệ thống dựa trên
cơ sở dữ liệu
(Database systems)

Khái niệm dữ liệu và CSDL



Hệ thống File:

- Dữ liệu được lưu trữ trong các file riêng biệt (có thể là dạng tập tin truyền thống trên máy tính).
- Nhược điểm của việc quản lý bằng file:
 - ✓ **Dư thừa** và không nhất quán
 - ✓ **Kém hiệu quả** trong truy xuất ngẫu nhiên, hoặc xử lý đồng thời
 - ✓ Dữ liệu lưu trữ **rời rạc và không toàn vẹn**
 - ✓ Gặp vấn đề về **an toàn và bảo mật**



Khái niệm dữ liệu và CSDL

Hệ thống CSDL:

- Quản lý dữ liệu bằng CSDL giúp dữ liệu được lưu trữ một cách **hiệu quả** và có **tổ chức**, cho phép quản lý dữ liệu **nhANH chóng và hiệu quả**.
- Lợi ích của hệ thống quản lý bằng CSDL:
 - ✓ **Tránh dư thừa**, trùng lặp dữ liệu
 - ✓ **Đảm bảo sự nhất quán** trong CSDL
 - ✓ Các dữ liệu lưu trữ có thể được **chia sẻ**
 - ✓ Có thể **thiết lập các chuẩn** trên dữ liệu
 - ✓ Duy trì **tính toàn vẹn** dữ liệu
 - ✓ Đảm bảo **bảo mật** dữ liệu



Các mô hình CSDL

- Các CSDL có thể khác nhau về **chức năng** và **mô hình dữ liệu** (data model).
- Mô hình dữ liệu sẽ quyết định **cách thức lưu trữ** và **truy cập dữ liệu**.
- Tùy từng ngữ cảnh quan hệ giữa các thành phần dữ liệu trong CSDL, một mô hình được áp dụng để việc **lưu trữ và truy xuất** dữ liệu đạt **hiệu quả cao nhất**.

Mô hình dữ liệu flat-file

Mô hình dữ liệu phân cấp

Mô hình dữ liệu mạng

Mô hình dữ liệu quan hệ

Mô hình dữ liệu flat-file

- Mô hình này chỉ dùng cho các CSDL đơn giản.
- CSDL dạng file phẳng thường là file kiểu văn bản chứa dữ liệu dạng bảng.

Ví dụ: một file phẳng thể hiện thông tin một Customer (Khách hàng) dưới dạng bảng của một công ty Northwind Traders chuyên cung cấp các mặt hàng thực phẩm

Customer ID	Company Name	Contact First Name	Contact Last Name	Job Title	City	State
6	Company F	Francisco	Pérez-Olaeta	Purchasing Manager	Milwaukee	WI
26	Company Z	Run	Liu	Accounting Assistant	Miami	FL

Mô hình dữ liệu flat-file

id	name	team
1	Amy	Blues
2	Bob	Reds
3	Chuck	Blues
4	Dick	Blues
5	Ethel	Reds
6	Fred	Blues
7	Gilly	Blues
8	Hank	Reds
9	Hank	Blues

Bob|123 street|California|\$200.00
Nathan|800 Street|Utah|\$10.00

Flat File Model

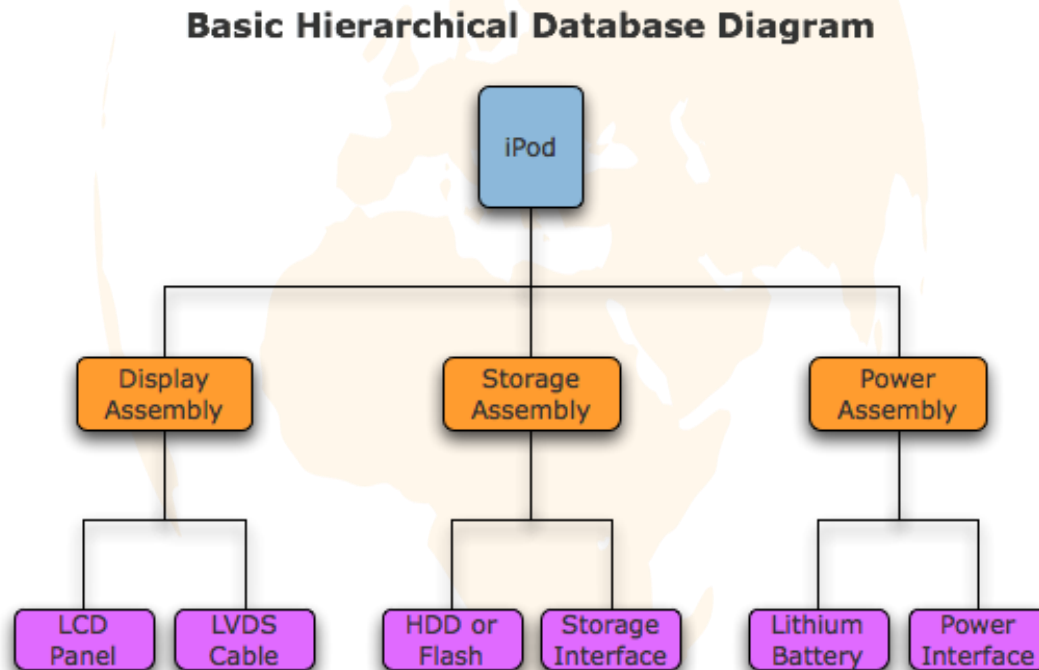
	Route No.	Miles	Activity
Record 1	I-95	12	Overlay
Record 2	I-495	05	Patching
Record 3	SR-301	33	Crack seal

Mô hình dữ liệu phân cấp

- Tổ chức theo **hình cây**, mỗi nút biểu diễn một thực thể dữ liệu.
- Liên hệ dữ liệu thể hiện trên liên hệ giữa **nút cha và nút con**. Mỗi nút cha có thể có một hoặc nhiều nút con, nhưng mỗi nút con chỉ có thể có một nút cha.
- Do đó mô hình phân cấp thể hiện các kiểu quan hệ:
 - ✓ 1-1
 - ✓ 1-N

Mô hình dữ liệu phân cấp

- Một nút con không thể có quá một nút cha -> Không biểu diễn được các quan hệ dữ liệu phức tạp
- Ví dụ



©2006 Scott Macdonald

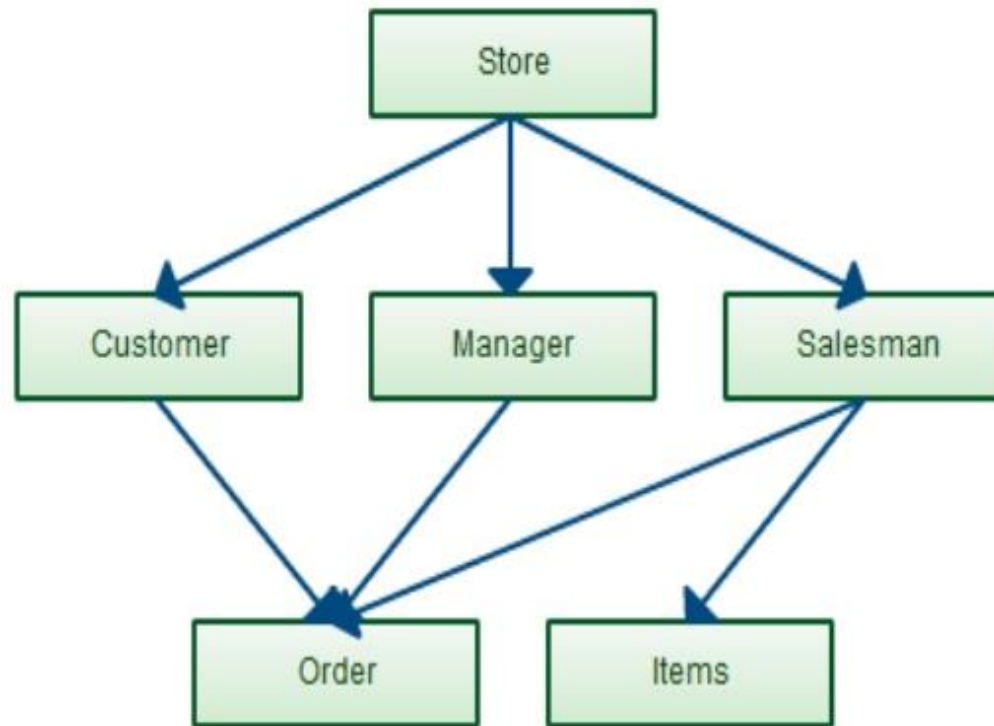
Mô hình dữ liệu mạng

- Mô hình này tương tự như mô hình dữ liệu phân cấp, thực sự nó là tập con của mô hình phân cấp.
- Trong mô hình mạng, dữ liệu được lưu trữ trong các bộ (sets), thay cho dạng cấu trúc phân cấp. Điều này nhằm để giải quyết vấn đề dư thừa dữ liệu.
- **Ưu điểm:**
 - ✓ Dễ biểu diễn mô hình
 - ✓ Diễn đạt được các liên hệ dữ liệu phức tạp
- **Nhược điểm:**
 - ✓ Truy xuất chậm
 - ✓ Không thích hợp với các CSDL có quy mô lớn.

Mô hình dữ liệu mạng

- Cấu trúc mô hình mạng

Network DBMS



Mô hình dữ liệu quan hệ

- Mô hình cơ sở dữ liệu quan hệ đáp ứng yêu cầu về **sự tăng trưởng và phức tạp** của dữ liệu trong nhiều doanh nghiệp.
- Thuật ngữ "**quan hệ**" có nguồn gốc từ lý thuyết tập hợp của toán học. Trong mô hình quan hệ, không giống như các mô hình phân cấp và mạng, **không có liên kết vật lý**.
- Tất cả các dữ liệu được **duy trì ở dạng bảng** bao gồm các hàng và cột. Dữ liệu trong hai bảng có quan hệ thông qua các cột chung chứ không liên kết vật lý.

Mô hình dữ liệu quan hệ

- Trong mô hình dữ liệu quan hệ, không có các liên kết vật lý. Dữ liệu được biểu diễn **dưới dạng bảng** với các hàng và các cột:
 - ✓ CSDL là **tập hợp các bảng** (còn gọi là quan hệ)
 - ✓ Mỗi hàng là một bản ghi (**record**), còn được gọi là bộ (**tuple**)
 - ✓ Mỗi cột là một **thuộc tính**, còn được gọi là trường (field)
- Dữ liệu trong hai bảng liên hệ với nhau thông qua các **cột chung**.
- Có các toán tử để thao tác trên các hàng của bảng.

Mô hình dữ liệu quan hệ

- Ví dụ:

Students Table

Student	ID*
John Smith	084
Jane Bloggs	100
John Smith	182
Mark Antony	219

Participants Table

ID*	Activity*
084	Tennis
084	Swimming
100	Squash
100	Swimming
182	Tennis
219	Golf
219	Swimming
219	Squash

Activities Table

Activity*	Cost
Golf	\$47
Sailing	\$50
Squash	\$40
Swimming	\$15
Tennis	\$36

Roll Number	Student Name
1	Sam Reiner
2	John Parkinson
3	Jenny Smith
4	Lisa Hayes
5	Penny Walker
6	Peter Jordan
7	Joe Wong

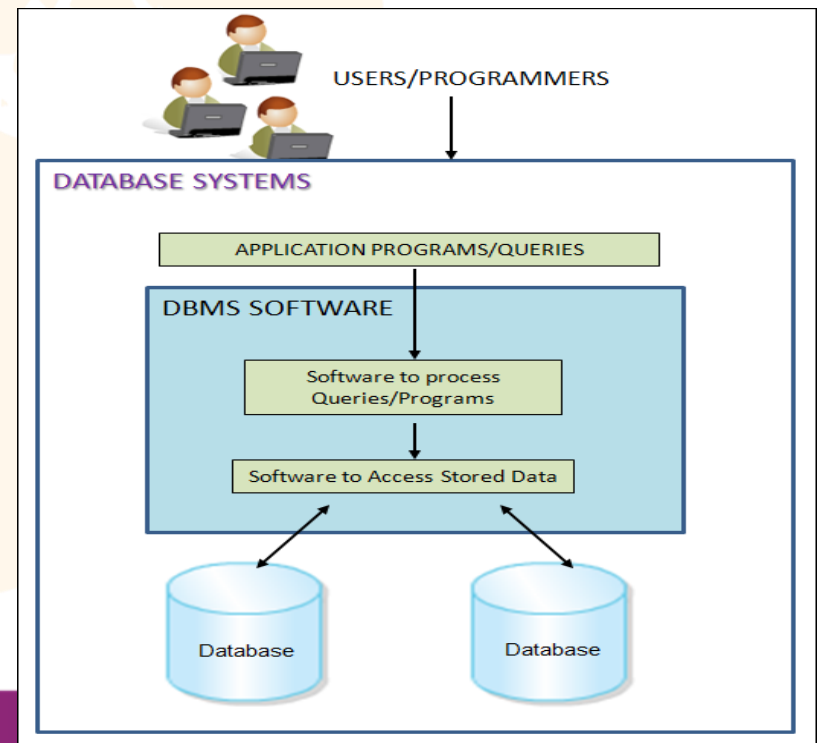
Students Table

Roll Number	Marks Obtained
1	34
2	87
3	45
4	90
5	36
6	65
7	89

Marks Table

Hệ quản trị CSDL (DBMS)

- Hệ quản trị CSDL (Database Management System - DBMS) là **phần mềm** giúp tạo các csdl và cung cấp cơ chế **lưu trữ**, **truy cập**.
- VD: SQL Server, MySQL, Microsoft Access, Oracle, PostgreSQL, DB2, ..v..v....



Hệ quản trị CSDL (DBMS)

Những lợi ích DBMS mang lại:

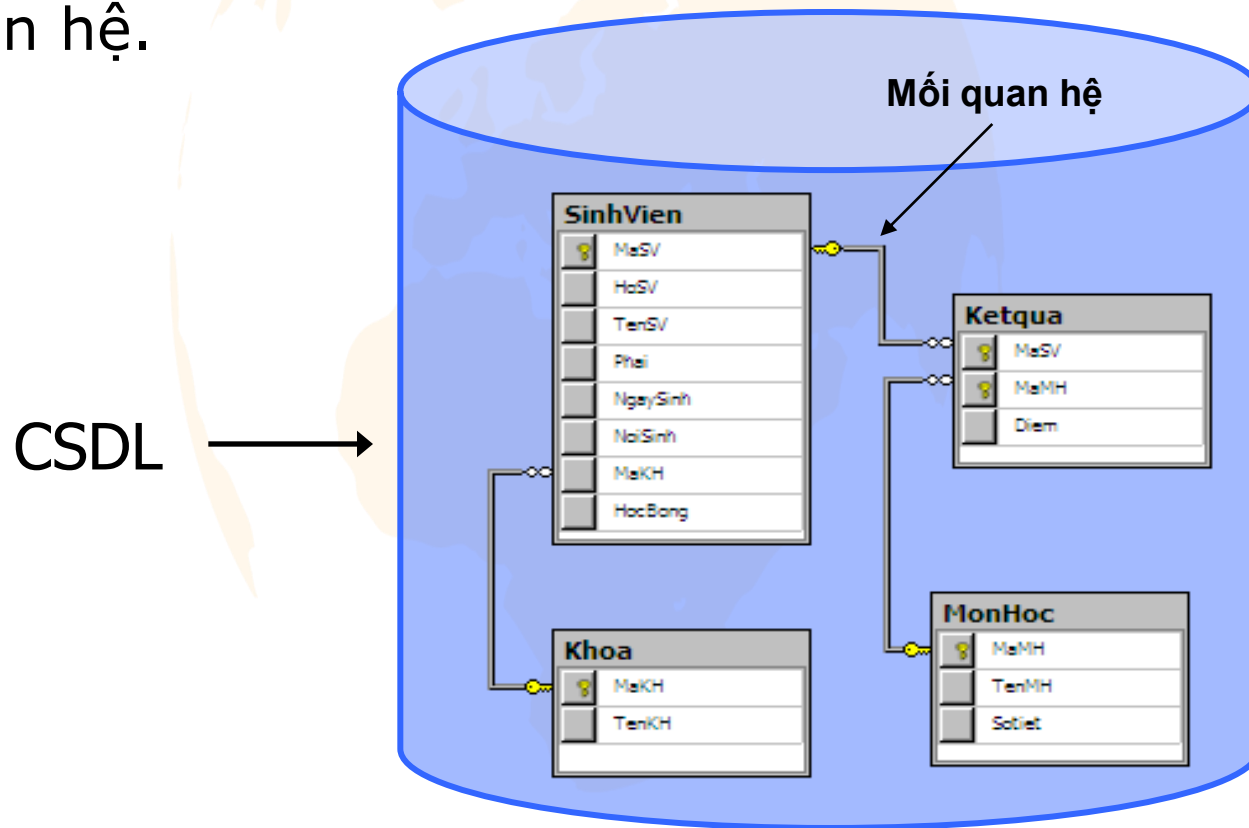
- **Quản trị** các CSDL
- **Cung cấp giao diện truy cập** để che dấu các đặc tính phức tạp về mặt cấu trúc tổ chức dữ liệu vật lý
- Hỗ trợ các **ngôn ngữ giao tiếp**. Ví dụ:
 - ✓ Ngôn ngữ mô tả, định nghĩa dữ liệu – **DDL**
 - ✓ Ngôn ngữ thao tác dữ liệu – **DML**
 - ✓ Ngôn ngữ truy vấn dữ liệu có cấu trúc – **SQL**
- Có cơ chế **an toàn, bảo mật** cao

Hệ quản trị CSDL quan hệ (RDBMS)

- Hệ quản trị CSDL quan hệ (**Relational DataBase Management System** = RDBMS)
- Là một hệ quản trị CSDL được xây dựng trên **mô hình quan hệ**.
- Một CSDL quan hệ là một CSDL được chia thành các đơn vị logic gọi là **bảng**, và chúng có mỗi **quan hệ** với nhau (khóa chung, trường chung).

Hệ quản trị CSDL quan hệ (RDBMS)

Hình dưới đây cho thấy các bảng quan hệ với nhau thông qua các **khóa chung** (giá trị dữ liệu) trong một csdl quan hệ.



Hệ quản trị CSDL quan hệ (RDBMS)

Khái niệm	Nghĩa
Relation	Một bảng
Tuple	Một dòng hoặc một bản ghi trong quan hệ
Attribute	Một trường hoặc một cột trong quan hệ
Cardinality of a relation	Số dòng trong một quan hệ
Degree of a relation	Số thuộc tính trong một quan hệ
Domain of an attribute	Tập các giá trị có thể lưu trữ trong thuộc tính
Primary Key of a relation	Một thuộc tính hoặc một tập các thuộc tính mà giá trị của nó xác định duy nhất các bản ghi trong quan hệ
Foreign key	Một thuộc tính hoặc một tập các thuộc tính trong quan hệ R1 mà nó chỉ ra mối quan hệ của R1 với quan hệ R2 Các thuộc tính khóa ngoại trong R1 phải chứa các giá trị so khớp với các giá trị tương ứng trong quan hệ R2

Hệ quản trị CSDL quan hệ (RDBMS)

Rất nhiều người dùng tham gia vào **hệ thống RDBMS** như:

- ✓ Người quản trị CSDL (**DataBase Administrator**)
- ✓ Người thiết kế CSDL (**DataBase Designer**)
- ✓ Người phân tích hệ thống (**System Analysts**)
- ✓ Người lập trình ứng dụng (**Application Programmers**)
- ✓ Người thiết kế và triển khai CSDL (**DBMS Designers and Implementers**)
- ✓ Người dùng cuối (**End User**)

Thực thể

- Một thực thể là một người, địa điểm, điều, đối tượng, sự kiện, hoặc thậm chí là một khái niệm, có thể nhận diện được rõ ràng..
- Ví dụ, các thực thể trong một trường đại học là sinh viên, giảng viên, và các khóa học.

Sinh Viên

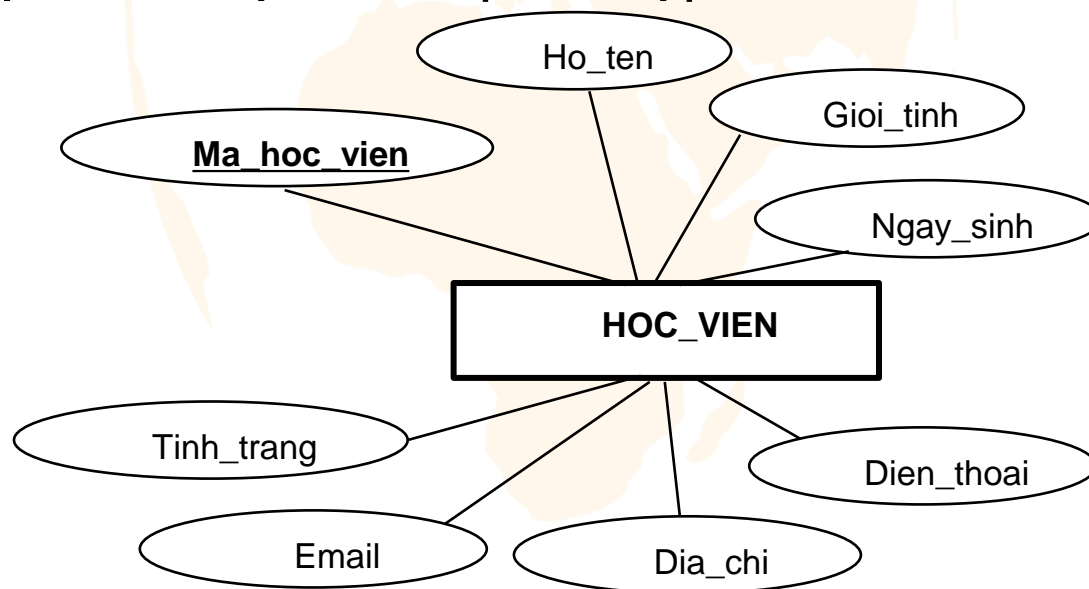
Giảng Viên

Khóa học

Điểm

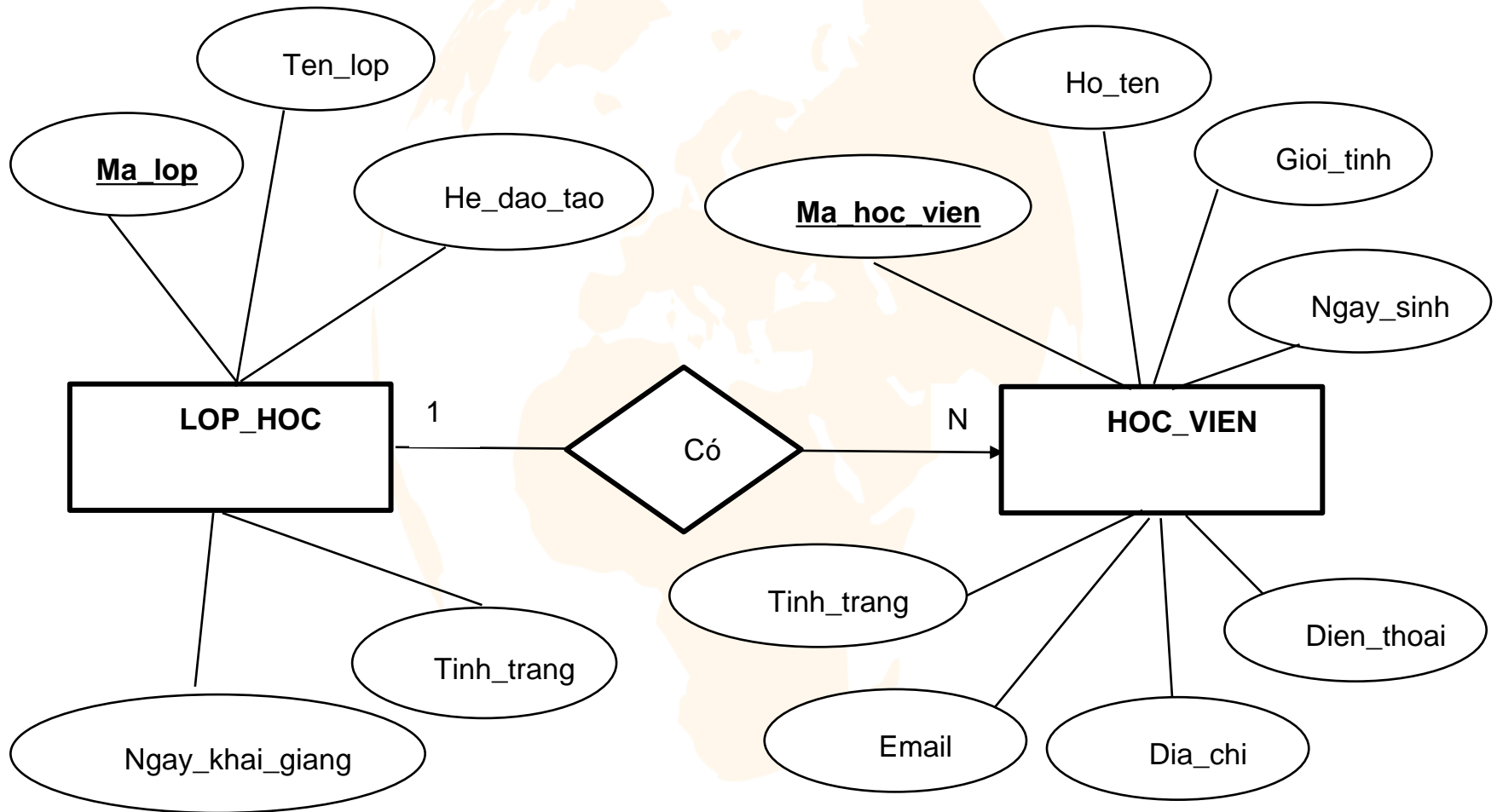
Thực thể

- Mỗi thực thể có một số đặc điểm được biết đến như là **thuộc tính**.
- Ví dụ, các thực thể sinh viên có thể gồm có các thuộc tính như số học sinh, tên, và lớp. Mỗi thuộc tính nên được đặt tên một cách phù hợp.



Thực thể

- Quan hệ giữa các thực thể.




Bảng

- Một nhóm các thực thể có liên quan sẽ trở thành một tập thực thể. Mỗi tập thực thể được đặt một tên. Tên của tập thực thể phản ánh các nội dung.
- Do đó, các thuộc tính của tất cả các sinh viên của các trường đại học sẽ được lưu trữ trong một tập thực thể gọi là **SinhVien**.
- Tập thực thể chính là bảng trong csdl quan hệ.
- Trong bảng, các hàng gọi là bộ dữ liệu còn cột gọi là thuộc tính.

Bảng

Đặc tính của bảng:

Attributes



Emp_No	Emp_Name	Emp_DOB	Emp_DOJ
345	James McElroy	24-Sep-1968	30-May-1990
873	Pamela Smith	10-Feb-1960	19-Nov-1993
693	Allan Howard	14-May-1975	24-Aug-2012
305	Geoff Bridges	12-Feb-1973	05-Jan-2013

Tuple

Table: EMPLOYEE

Tóm tắt bài học

- Một CSDL là một tập dữ liệu có quan hệ được lưu trữ trong các bảng.
- Mô hình dữ liệu mô tả về container lưu trữ dữ liệu, và quá trình lưu trữ và truy xuất dữ liệu từ container đó.
- Một hệ quản trị CSDL có thể được định nghĩa như một tập các bản ghi quan hệ và các chương trình có thể truy cập và thao tác trên các bản ghi.
- DBMS là một tập các chương trình mà nó cho phép người dùng lưu trữ, cập nhật và trích rút thông tin từ một CSDL.
- Một hệ quản trị CSDL quan hệ (RDBMS) là một phần mềm cho phép tạo, thao tác, cập nhật trên một CSDL quan hệ.
- Một CSDL quan hệ được chia nhỏ thành các đơn vị logic gọi là bảng. Các đơn vị logic này có quan hệ với nhau trong CSDL.
- Các thành phần chính của RDBMS là các thực thể và các bảng.

TRƯỜNG ĐÀO TẠO LẬP TRÌNH VIÊN VÀ QUẢN TRỊ MẠNG QUỐC TẾ BACHKHOA-APTECH

Thank for watching!

