

# **HỆ THỐNG THÔNG TIN QUẢN LÝ**

## Management Information System MIS

**TS. ĐỖ QUANG VINH**

**Email: [dqvinh@live.com](mailto:dqvinh@live.com)**

**Hà Nội - 2013**

## ▪ **Tính cấp thiết**

✓ Công nghệ thông tin ngày càng được ứng dụng rộng rãi vào tất cả các ngành kinh tế xã hội nói chung và ngành Xuất bản - Phát hành nói riêng

⇒ Môn học **TIN HỌC QUẢN LÝ** đối với sinh viên ngành Xuất bản - Phát hành là một yêu cầu cấp thiết

## ▶ **Mục đích**

✓ Khái niệm về Hệ thống thông tin IS (HTTT), Hệ thống thông tin quản lý MIS, phân loại HTTT, phương pháp phát triển HTTT

✓ Hệ quản trị cơ sở dữ liệu DBMS, các thành phần cơ bản và các chức năng chính của HQTCSDL, mô hình dữ liệu quan hệ, phương pháp thiết kế CSDL

- **Quản lý trong tổ chức văn hóa thông tin**
- ✓ **Khái niệm quản lý là quá trình tổ chức các nguồn lực nhằm đạt tới một mục tiêu nào đó của một tổ chức, trong đó bao gồm quá trình lập kế hoạch, điều hành, lãnh đạo và giám sát, với nhiều cấp độ và phạm vi quản lý khác nhau phụ thuộc vào mục tiêu, thời gian, tính phức tạp của các quan hệ kinh tế - xã hội và văn hóa trong quản lý**

(Nhập môn quản lý văn hóa thông tin – Quỳ Ford)

- **Quản lý trong tổ chức văn hóa thông tin**

- ✓ **Cấp độ quản lý:**

- **Cấp độ điều hành**

- **Cấp độ quản trị**

- **Cấp độ chiến lược**

# **Cấp độ quản lý của một tổ chức văn hóa thông tin**



**Quản lý cấp cao**  
(Lập kế hoạch chiến lược)

**Quản lý cấp trung**  
(Các kế hoạch điều hành)

**Quản lý cấp thấp**  
(Các kế hoạch hàng ngày)

- ▶ **Định nghĩa 1** (Phan Đình Diệu, Quách Tuấn Ngọc ...[24]):  
Hệ thống là một tập hợp các thành phần có liên hệ với nhau, hoạt động để hướng tới mục đích chung theo cách tiếp nhận các yếu tố vào, sinh ra các yếu tố ra trong một quá trình xử lý có tổ chức
- ▶ **Hệ thống có 3 thành phần cơ bản tương tác với nhau:**
  - + Yếu tố vào (đầu vào)
  - + Xử lý, chế biến
  - + Yếu tố ra (đầu ra)

## ❖ KHÁI NIỆM DỮ LIỆU – THÔNG TIN – TRI THỨC

- **Dữ liệu** (Data) là vật liệu thô để tạo ra thông tin, còn thông tin là dữ liệu đã được thu thập và xử lý chuyển thành dạng có nghĩa. Nói cách khác, dữ liệu là nguồn gốc, là vật mang thông tin, là vật liệu sản xuất ra thông tin.
- ✓ Dữ liệu có thể là:
  - Tín hiệu vật lý (physical signal): tín hiệu điện, tín hiệu sóng điện từ, tín hiệu ánh sáng, tín hiệu âm thanh, nhiệt độ, áp suất, ...
  - Các số liệu (number) là dữ liệu bằng số mà ta đã quen với tên gọi số liệu. Đó là các số liệu trong các biểu bảng tính toán, thống kê, tài chính ...
  - Các ký hiệu (symbol) như các chữ viết (character) và các ký hiệu khắc trên tre, nứa, đá, trên bia, trên vách núi ... của người xưa.

- Văn bản, chữ viết (text, character): sách báo, truyện, thông báo, thông tư, công văn ...
- Âm thanh(sound): tiếng nói, âm nhạc, tiếng ồn
- Hình ảnh (image): phim ảnh, tivi, camera, tranh vẽ ...
- Đồ họa (graphic) ...
- **Thông tin** (INFORMATION) là một khái niệm trừu tượng mô tả tất cả những gì đem lại hiểu biết, nhận thức cho con người về đối tượng mình quan tâm
  - Thông tin tồn tại khách quan. Thông tin có thể tạo ra, phát sinh, truyền đi, lưu trữ, chọn lọc. Thông tin cũng có thể bị méo mó, sai lệch do nhiễu hay do người xuyên tạc
- **Tri thức** (Knowledge) bao gồm các sự kiện và luật dẫn, là một dạng đặc biệt của dữ liệu



## ▪ **Phân loại tri thức**

Khả năng chuyển giao tri thức  
(transferability of knowledge)

### ✓ **Tri thức hiện** (explicit knowledge)

- Tri thức về sự vật:(know-what)

- Tri thức về nguyên nhân hay tri thức giải thích:(know-why)

### ✓ **Tri thức ẩn** (tacit knowledge)

- Tri thức về cách làm(know-how)

- Tri thức về người biết (know-who)

## ▶ **Tri thức hiện**

mô tả bằng văn bản, hình vẽ, âm thanh, có thể truyền bá nhanh chóng

## ▶ **Tri thức ẩn**

- phong phú, đa dạng; chỉ có thể diễn tả trực tiếp bằng hành động và thao tác của chính người sở hữu nó, ví dụ như các nghệ nhân, các thợ có tay nghề cao v.v. . .
- tồn tại trong bộ não của chủ nhân, không thể truyền đi xa được, chỉ có thể truyền cho người đến học bằng kiểu bắt chước hoặc "cầm tay chỉ việc"

# SO SÁNH HAI LOẠI TRI THỨC

## TRI THỨC HIỆN

- Tiếp cận lý thuyết
- Cách giải quyết vấn đề
- Tài liệu
- Cơ sở dữ liệu
- Cơ sở tri thức

## TRI THỨC ẨN

- Nhận thức
  - ✓ Niềm tin
  - ✓ Quan niệm
  - ✓ Trực giác
  - ✓ Mô hình
- Kỹ thuật
  - ✓ Ngón nghề (craft)
  - ✓ Bí quyết (know-how)

## **R.R Nelson và P.Romer (1996)**

- **Phần cứng** (hardware): Tất cả các vật thể vật chất không phải con người
- **Phần mềm** (software): Kiến thức được diễn chế hoá (codified) và lưu trữ bên ngoài não người.
- **Phần ướt** (wetware): Tri thức lưu trữ trong “máy tính ướt” của não người, gồm niềm tin, kỹ năng, tài năng, ý chí ...

# THÁP THÔNG TIN



Mức độ  
xử lý  
thông tin

TRI THỨC ĐÃ ĐƯỢC TRẢI NGHIỆM (WISDOM)

TRI THỨC

THÔNG TIN

DỮ LIỆU

## ▪ Định nghĩa Hệ thống thông tin HTTT

### ✓ Định nghĩa

#### ▶ Định nghĩa 2 (Phan Đình Diệu, Quách Tuấn Ngọc, ... [24]):

HTTT là hệ thống tiếp nhận các nguồn dữ liệu như các yếu tố vào và xử lý chúng thành các sản phẩm thông tin là các yếu tố ra

#### ▶ Định nghĩa 3 (Laudon [14]):

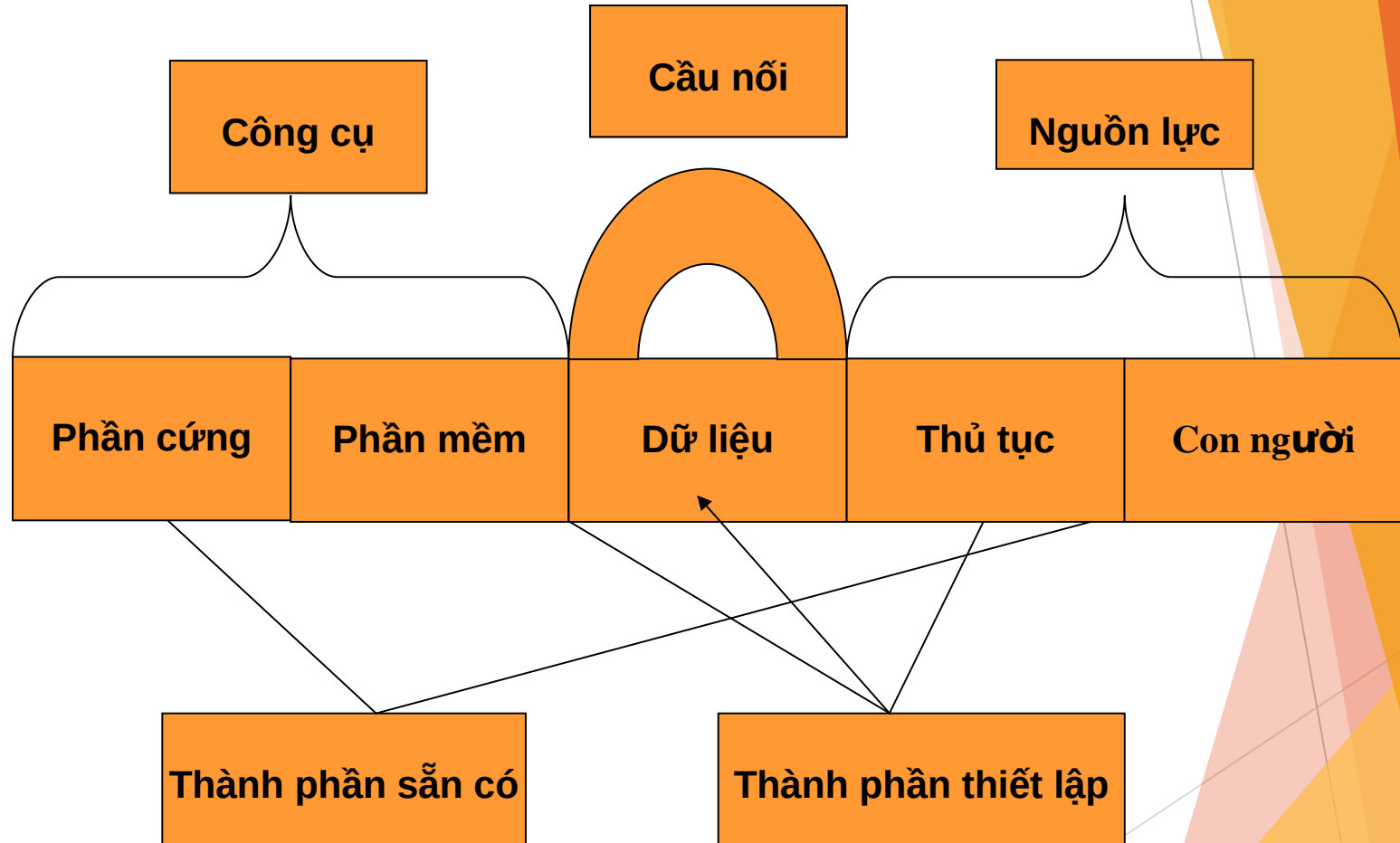
HTTT là tập hợp các thành phần được tổ chức để thu thập, xử lý, lưu trữ, phân phối và biểu diễn thông tin trợ giúp việc ra quyết định và kiểm soát trong một tổ chức

- **Hệ thống thông tin có một ý nghĩa quan trọng trong quá trình phân tích và đánh giá hoạt động của một tổ chức văn hóa thông tin. Không có thông tin, nhà quản lý sẽ không có cơ sở để đánh giá môi trường và đưa ra các giải pháp quản lý**

(Nhập môn Quản lý Văn hóa thông tin – Quỹ Ford )

# Hình 1.3 - Các thành phần của HTTT

(Kroenke D., Hatch R. [12])





▶ Định nghĩa 4 (Keen P.G.W.):

Hệ thống thông tin quản lý MIS (Management Information System) là sự phát triển và sử dụng HTTT có hiệu quả trong một tổ chức. Một HTTT được coi là hiệu quả nếu trợ giúp hoàn thành được các mục tiêu của con người và tổ chức sử dụng nó [49]

## ▪ **VAI TRÒ CỦA HTTT**

Trong bất kỳ một tổ chức, xác định 3 hệ thống:

- Hệ thống điều khiển: có nhiệm vụ ra các quyết định
- Hệ thống thực hiện: hoạt động nhằm thực hiện các quyết định xác định bởi hệ thống điều khiển
- Hệ thống thông tin: thực hiện sự liên hệ giữa hệ thống trên, đảm bảo cho tổ chức hoạt động đạt được các mục tiêu đã đặt ra

## ▪ **Nhiệm vụ của HTTT**

Thu thập thông tin: phân tích, chọn lọc và ghi nhận thông tin cần thiết và có ích cho quản lý

Xử lý thông tin: thực hiện tính toán, cập nhật, lưu trữ dữ liệu

Truyền thông tin: thực hiện truyền thông tin sao cho đảm bảo thời gian và bảo mật

## ▶ **Đặc trưng của HTTT: 4**

- HTTT phải được thiết kế, tổ chức trong ngữ cảnh chung của kinh tế xã hội. HTTT phục vụ nhiều lĩnh vực hoạt động khác nhau, phục vụ nhiệm vụ chung của một tổ chức

- HTTT đạt được mục tiêu là ra các quyết định. Việc xây dựng HTTT nhằm mục đích cuối cùng là hỗ trợ ra quyết định. Để quyết định đúng đắn, cần phải cung cấp cho người ra quyết định đủ thông tin cần thiết

- HTTT phải dựa trên các kỹ thuật tiên tiến về xử lý thông tin, bao gồm các phần mềm.Ứng dụng, các thiết bị CNTT. Một trong những kiến thức cần thiết nhất là tri thức về cơ sở dữ liệu và các hệ quản trị cơ sở dữ liệu DBMS

- HTTT có kiến trúc mềm dẻo, có khả năng phát triển được

## ▪ **HIỆU QUẢ CỦA HTTT**

Một HTTT được coi là có hiệu quả nếu góp phần nâng cao chất lượng hoạt động quản lý tổng thể

Tính hiệu quả của một HTTT được thể hiện trên các mặt sau đây:

- Đạt được các mục tiêu thiết kế đề ra của tổ chức
- Chi phí vận hành là chấp nhận được

- Tin cậy, đáp ứng được các chuẩn mực của một HTTT hiện hành
- Sản phẩm có giá trị xác đáng
- Dễ học và dễ sử dụng
- Mềm dẻo, có thể kiểm tra, mở rộng ứng dụng và phát triển tiếp được

## ▪ **PHÂN LOẠI HTTT** theo chức năng nghiệp vụ

### 1. **HỆ THỐNG XỬ LÝ GIAO DỊCH TPS**

(Transaction Processing System)

- Là một HTTT nghiệp vụ
- phục vụ cho hoạt động của tổ chức ở mức vận hành
- Cung cấp nhiều dữ liệu nhất cho các hệ thống khác trong tổ chức

- Đặc trưng của TPS:

- Khối lượng công việc giao dịch nhiều
- Các quy trình xử lý giao dịch rõ ràng
- Chặt chẽ, có thể mô tả chi tiết
- Ít ngoại lệ

➤ TPS không linh hoạt, không thể điều tiết được việc xử lý dữ liệu hay nhu cầu thông tin khi chưa được xây dựng trước trong hệ thống

## 2. **HỆ THỐNG THÔNG TIN QUẢN LÝ MIS**

(Management Information System)

- Trợ giúp các hoạt động quản lý của tổ chức như lập kế hoạch, kiểm tra thực hiện, tổng hợp và làm báo cáo, ra quyết định quản lý trên cơ sở các quy trình thủ tục cho trước

⇒ MIS sử dụng dữ liệu từ TPS và tạo ra các báo cáo định kỳ hay theo yêu cầu

- MIS ít có khả năng phân tích thông tin

▪ Đặc trưng của MIS:

- Hỗ trợ chức năng xử lý dữ liệu trong giao dịch và lưu trữ

- Sử dụng CSDL thống nhất, có nhiều chức năng xử lý dữ liệu
- Cung cấp đầy đủ thông tin để người quản lý truy cập dữ liệu
- Thích ứng được với những thay đổi của quy trình xử lý thông tin
- Đảm bảo toàn vẹn dữ liệu và an toàn dữ liệu

⇒ So với TPS, MIS mềm dẻo hơn, có nhiều chức năng xử lý dữ liệu hơn

### 3. **HỆ TRỢ GIÚP QUYẾT ĐỊNH DSS**

(Decision Support System)

- Là hệ được sử dụng ở mức quản lý của tổ chức
  - Có nhiệm vụ tổng hợp các dữ liệu và tiến hành phân tích bằng mô hình để trợ giúp cho các nhà quản lý ra quyết định có quy trình rõ ràng (bán cấu trúc) hay hoàn toàn không có quy trình biết trước (không có cấu trúc)
- ✓ Quá trình ra quyết định thường gồm 3 giai đoạn: xác định vấn đề, xây dựng và đánh giá các giải pháp khác nhau và lựa chọn một giải pháp tối ưu, thích hợp



- DSS phải sử dụng nhiều loại dữ liệu khác nhau nên các CSDL phải được tổ chức và liên kết tốt
- DSS có nhiều phương pháp xử lý được tổ chức sao cho có thể sử dụng linh hoạt
- DSS thường được xây dựng cho mỗi tổ chức cụ thể mới đạt hiệu quả cao

## 4. HỆ CHUYÊN GIA ES

(Expert System)

- Là một hệ trợ giúp quyết định ở mức chuyên sâu
  - Ngoài những kiến thức, kinh nghiệm của các chuyên gia và các luật suy diễn, ES còn có thể trang bị các thiết bị cảm nhận để thu các thông tin từ những nguồn khác nhau
  - Có thể xử lý và dựa vào các luật suy diễn để đưa ra các quyết định rất hữu ích và thiết thực

⇒ Sự khác biệt cơ bản của ES với DSS là ES yêu cầu những thông tin xác định đưa vào để ra quyết định có chất lượng cao trong một lĩnh vực hẹp

## 5. **HỆ TRỢ GIÚP ĐIỀU HÀNH ESS**

(Execution Support System)

- ESS được sử dụng ở mức quản lý chiến lược của tổ chức
- ESS được thiết kế hướng sự trợ giúp cho các quyết định không cấu trúc bằng việc làm ra các đồ thị phân tích trực quan và các giao dịch rất thuận tiện với môi trường
- ESS được thiết kế để cung cấp hay trích lọc các thông tin đa dạng lấy từ môi trường hay các hệ MIS, DSS, ...

▪ Đặc trưng của ESS:

- Phục vụ nhu cầu thông tin điều hành
- Đảm bảo giao diện NSD thân thiện để tiện cho việc điều hành của người lãnh đạo
- Sử dụng các mô hình quyết định điều hành của cá nhân
- Đảm bảo theo dõi và điều khiển hiệu quả và theo thời gian
- Truy cập thông tin nhanh
- Trích lọc thông tin

⇒ So với các HTTT khác, ESS có phạm vi ứng dụng rộng rãi hơn, không điều hành các vấn đề chi tiết mà tập trung vào các vấn đề quan trọng

⇒ ESS được sử dụng cho người lãnh đạo của tổ chức

## 6. HỆ THỐNG THÔNG TIN TÍCH HỢP IIS

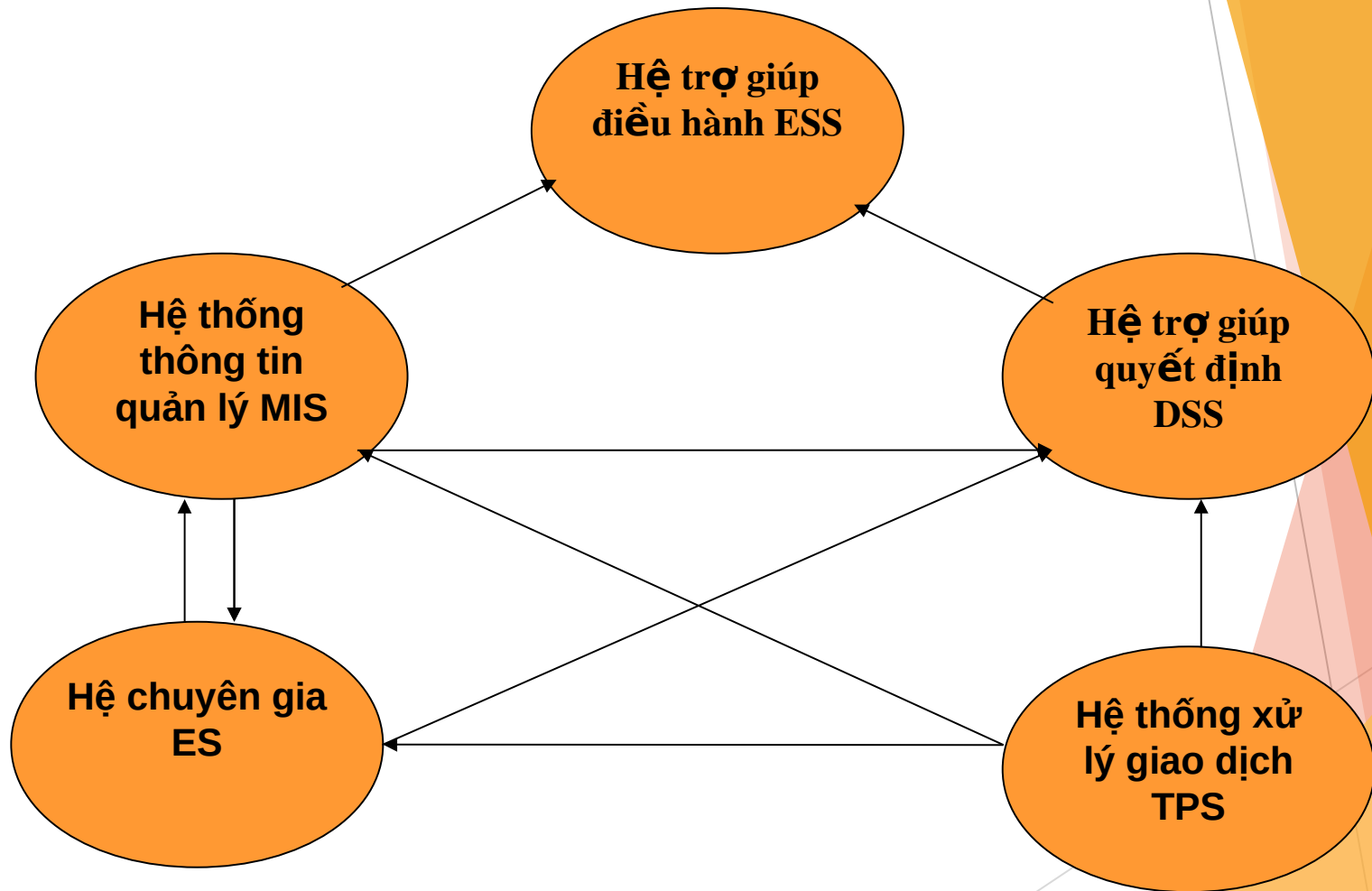
(Integrated Information System)

Một HTTT của tổ chức thường gồm vài loại HTTT cùng được khai thác nhằm đáp ứng được mục tiêu của tổ chức

- ⇒ Vì vậy, cần phải tích hợp nhiều HTTT khác loại để đảm bảo hoạt động hiệu quả của tổ chức
- ⇒ Việc tích hợp nhiều HTTT trong một tổ chức có thể tiến hành theo hai cách: xây dựng một HTTT tích hợp tổng thể hoặc tích hợp các HTTT đã có bằng cách ghép nối chúng nhờ các cầu nối
- ⇒ Việc sử dụng IIS tổng thể thường đưa tổ chức đến một hệ thống tập trung, một sự phối hợp và kiểm tra chặt chẽ

- ⇒ Tuy nhiên, tạo sức ỳ về quản lý và sự quan liêu trong hoạt động và khó thay đổi
- ⇒ Khi sự tập trung của một HTTT đã đạt đến bão hòa, nhiều tổ chức bắt đầu cho các bộ phận của mình tiếp tục phát triển những hệ con với các đặc thù riêng

# Hình 1.4 - Mối quan hệ giữa các loại HTTT (Laudon [14])



## ▪ Phương pháp phát triển HTTT

### ✓ Lý do tổ chức cần phát triển HTTT

1. Tổ chức gặp phải những vấn đề làm cản trở hoặc hạn chế không cho phép tổ chức thực hiện thành công những mục tiêu của họ
2. Tổ chức cần tạo ra các ưu thế mới năng lực mới để có thể vượt qua những thách thức và nắm được cơ hội trong tương lai
3. Do yêu cầu từ bên ngoài có liên quan đến sự phát triển và hợp tác của tổ chức



✓ **Quá trình phát triển HTTT: 6 bước**

- 1. Lập kế hoạch dự án**
- 2. Phân tích hệ thống**
- 3. Thiết kế hệ thống**
- 4. Lập trình và kiểm thử**
- 5. Cài đặt và chuyển đổi hệ thống**
- 6. Vận hành và bảo trì**

## ▪ **Bước 1: Lập kế hoạch dự án**

### ✓ Ý nghĩa:

- Quyết định việc có xây dựng HTTT hay không?
- Là một yêu cầu bắt buộc để tiến hành tiếp theo, không có dự án thì cũng không có xây dựng HTTT

### ✓ Mục tiêu: đưa ra dự án xây dựng HTTT khả thi

### ✓ Nội dung:

- Xác định mục tiêu;
- Xác định các yếu tố quyết định thành công
- Phân tích phạm vi, ràng buộc ảnh hưởng đến thời gian và nguồn lực
- Xác định các vấn đề có ảnh hưởng đến mục tiêu đạt được
- Xác định các nguồn lực: tài chính, nhân lực, vật lực ...
- Lựa chọn các giải pháp hợp lý để đạt được mục tiêu

- ✓ Yêu cầu:
  - Làm rõ HTTT trong tương lai đáp ứng nhu cầu gì?
  - Các nội dung trên có sức thuyết phục: đúng, đủ, tin cậy, khả thi để người lãnh đạo chấp nhận và thông qua
  
- **Bước 2: Phân tích hệ thống**
  - Phân tích HTTT là việc sử dụng các phương pháp và công cụ để nhận thức và hiểu biết được hệ thống, tìm các giải pháp giải quyết các vấn đề phức tạp nảy sinh trong hệ thống đang được nghiên cứu
- ✓ Ý nghĩa:

là giai đoạn chính, trọng tâm khi xây dựng HTTT, đi sâu vào bản chất của HTTT
- ✓ Mục tiêu: xác định nhu cầu thông tin

✓ Nội dung:

- Nghiên cứu hiện trạng: nắm được tình trạng hoạt động của hệ thống cũ

- Xây dựng mô hình hệ thống: dựa vào kết quả điều tra để xây dựng mô hình nghiệp vụ của hệ thống, từ đó làm rõ mô hình thông tin và mô hình tác nghiệp của hệ thống. Đây là công việc quan trọng nhất

- Phân tích tính khả thi: có tầm quan trọng đặc biệt, liên quan đến việc lựa chọn giải pháp mà thực chất là tìm ra điểm cân bằng giữa nhu cầu và khả năng giải quyết vấn đề. Phân tích tính khả thi được tiến hành trên 3 mặt

+ Khả thi về kỹ thuật

+ Khả thi về kinh tế

+ Khả thi hoạt động

- Lập hồ sơ nhiệm vụ

✓ Yêu cầu:

xác định rõ và đầy đủ các chức năng của hệ thống, các kiểu dữ liệu, cấu trúc dữ liệu

## ▪ **Bước 3: Thiết kế hệ thống**

- ✓ Ý nghĩa: là bước chính và đưa ra một phương án tổng thể hay một mô hình đầy đủ về HTTT trong tương lai
- ✓ Mục đích: đạt được các đặc tả về hình thức và cấu trúc HTTT, môi trường hoạt động của HTTT, nhằm hiện thực hóa kết quả phân tích và đưa ra quyết định cài đặt hệ thống
- ✓ Nội dung:
  - Thiết kế logic: bao gồm các thành phần của HTTT và liên kết giữa chúng. Kết quả nhận được là các mô hình dữ liệu và xử lý dữ liệu
  - Thiết kế vật lý: là quá trình chuyển mô hình logic thành thiết kế kỹ thuật của HTTT: hệ thống thiết bị và các chức năng của NSD. Kết quả là tạo ra các đặc tả cụ thể về phần cứng, phần mềm, CSDL, thủ tục xử lý
- ✓ Yêu cầu: đảm bảo HTTT thỏa mãn các yêu cầu đã phân tích

## ▪ **Bước 4: Lập trình và kiểm thử**

- ✓ Ý nghĩa: thể hiện kết quả phân tích và thiết kế. Đây chính là bước thi công
- ✓ Mục tiêu: xây dựng được phần mềm đáp ứng được các yêu cầu đặt ra
- ✓ Nội dung:
  - Lựa chọn phần mềm hạ tầng cơ sở của HTTT, bao gồm: hệ điều hành, hệ quản trị cơ sở dữ liệu DBMS, ngôn ngữ lập trình
  - Lựa chọn các phần mềm đóng gói
  - Chuyển các đặc tả thiết kế thành các phần mềm
  - Kiểm tra và thử nghiệm các module chức năng, hệ thống con và toàn bộ HTTT
- ✓ Yêu cầu:
  - Chuyển mọi kết quả phân tích và thiết kế HTTT trên giấy thành phần mềm chạy được trên máy tính
  - Đưa ra sản phẩm đúng đắn và hợp lệ

## ▪ **Bước 5: Cài đặt và chuyển đổi hệ thống**

✓ Ý nghĩa:

thay đổi và nâng cao chất lượng hoạt động của tổ chức

✓ Mục tiêu:

chuyển đổi toàn bộ hoạt động của tổ chức từ cũ sang mới, nghĩa là đưa HTTT mới vào sử dụng

✓ Nội dung:

chuyển đổi dữ liệu, đào tạo và sắp xếp đội ngũ cán bộ làm việc trên HTTT mới

✓ Yêu cầu:

HTTT mới hoạt động tốt và đem lại hiệu quả cao hơn hệ thống cũ

## ▪ **Bước 6: Vận hành và bảo trì**

✓ Ý nghĩa:

duy trì hoạt động của HTTT

✓ Mục tiêu:

đáp ứng các mục tiêu đặt ra ban đầu

✓ Nội dung:

- Đề xuất những sửa đổi, cải tiến, bổ sung

- Tiến hành những sửa đổi, bổ sung về phần cứng, phần mềm

- Kiểm tra tính đáp ứng được các yêu cầu đã đặt ra

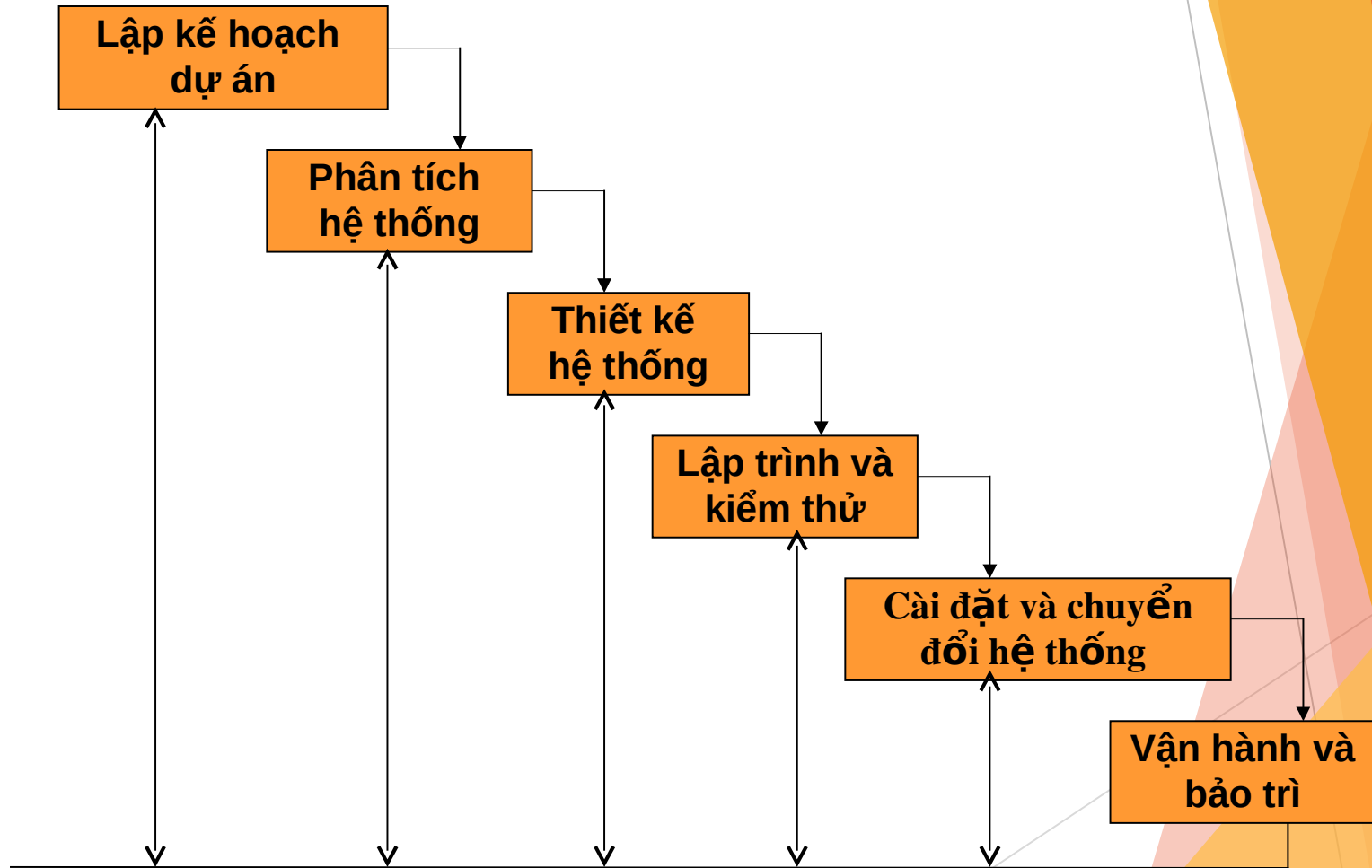
- Bảo trì HTTT

✓ Yêu cầu:

HTTT luôn sẵn sàng hoạt động không ngừng



# Hình 1.5 - Mô hình thác nước của quá trình phát triển HTTT (Pressman [39])



## ▪ **Vai trò của người tham gia phát triển HTTT**

- ✓ Người quản lý HTTT
- ✓ Người phân tích hệ thống
- ✓ Người lập trình
- ✓ Người sử dụng
- ✓ Người quản lý nghiệp vụ
- ✓ Các chuyên viên kỹ thuật

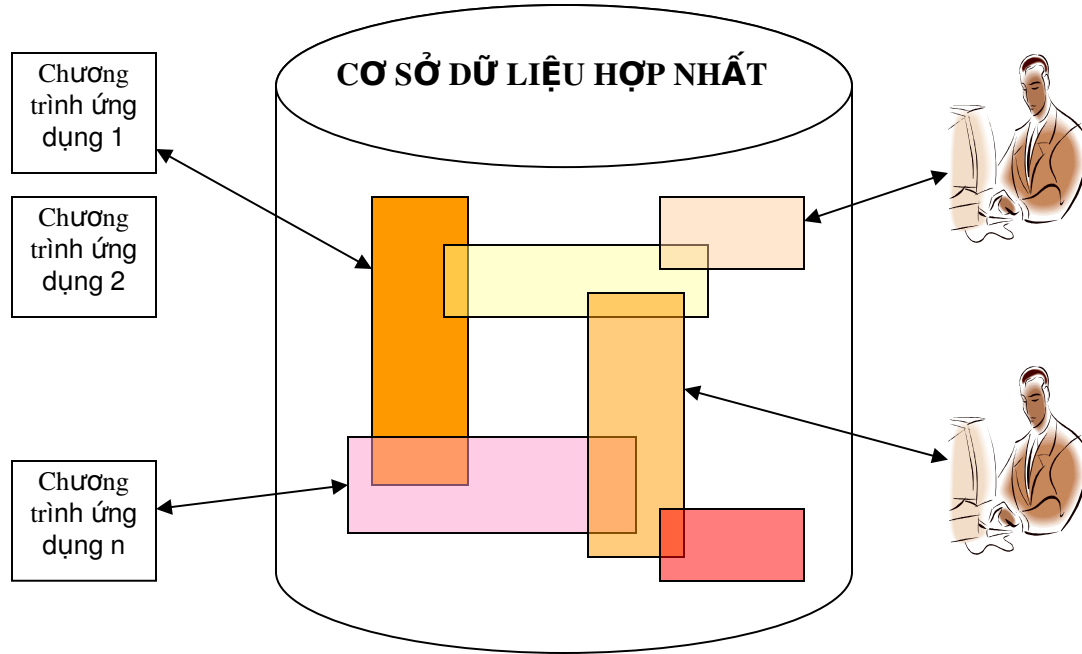
## ▪ **Hệ quản trị cơ sở dữ liệu DBMS**

- ✓ **Định nghĩa cơ sở dữ liệu**
- **Định nghĩa 5:**

**Cơ sở dữ liệu CSDL là một bộ sưu tập những dữ liệu tác nghiệp được lưu trữ lại và được các hệ ứng dụng của một “xí nghiệp” cụ thể nào đó sử dụng (C.J.Date [6])**

“Xí nghiệp” chỉ là một thuật ngữ chung tiện lợi để chỉ những hoạt động thương mại, khoa học, kỹ thuật hoặc các hoạt động khác có quy mô đủ lớn

Ví dụ: trường đại học, công ty, ngân hàng, bệnh viện, cơ quan nhà nước ...



Các hệ  
thống  
chương  
trình ứng  
dụng khai  
thác

Người sử  
dụng NSD  
khai thác cơ  
sở dữ liệu

**Hình 1. Sơ đồ tổng quát về một hệ cơ sở dữ liệu**

Dữ liệu tác nghiệp là các dữ liệu về hoạt động của xí nghiệp được lưu lại. Dữ liệu tác nghiệp của xí nghiệp có thể bao gồm:

- + Dữ liệu về sinh viên
- + Dữ liệu về kế hoạch đào tạo
- + Dữ liệu về sản phẩm
- + Dữ liệu về các tài khoản
- + Dữ liệu về người bệnh ...

• **Đặc trưng: 3**

1. CSDL phải là một tập hợp các thông tin mang tính hệ thống chứ không phải là các thông tin rời rạc, không có mối quan hệ với nhau
2. Thông tin phải có cấu trúc
3. Tập hợp thông tin phải có khả năng đáp ứng các nhu cầu khai thác của nhiều NSD một cách đồng thời

- **Ưu điểm nổi bật: 3**

1. Giảm sự trùng lặp thông tin xuống mức thấp nhất và do đó bảo đảm được tính nhất quán và toàn vẹn dữ liệu
2. Đảm bảo dữ liệu có thể được truy xuất theo nhiều cách khác nhau
3. Khả năng chia sẻ thông tin cho nhiều NSD và nhiều ứng dụng khác nhau

- **4 bài toán:**

1. Tính chủ quyền của dữ liệu
2. Tính bảo mật và quyền khai thác thông tin của NSD
3. Tranh chấp dữ liệu
4. **Đảm bảo dữ liệu khi có sự cố**

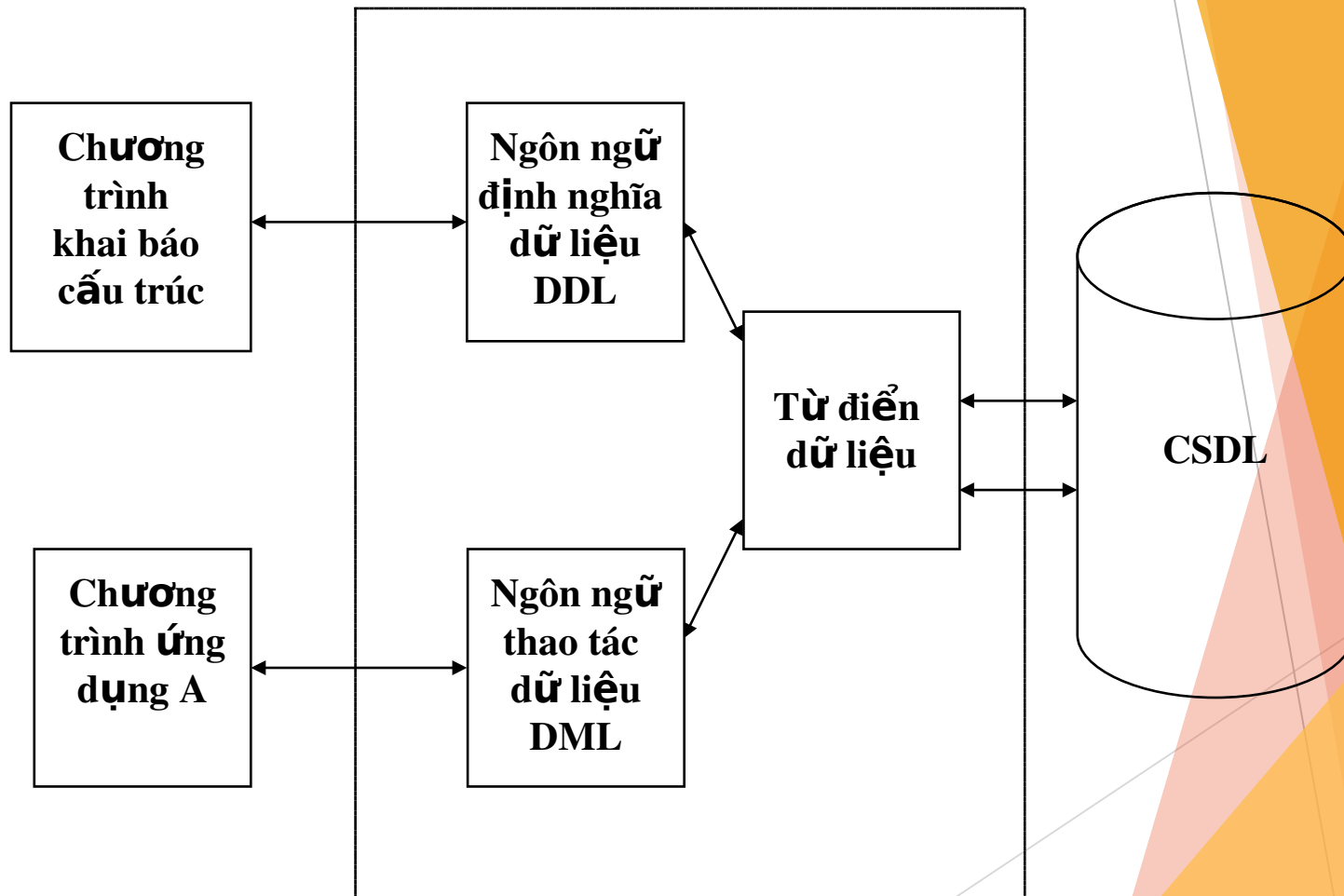
## ✓ **Hệ quản trị cơ sở dữ liệu DBMS**

- **Ngôn ngữ giao tiếp giữa NSD và CSDL:**
  - Ngôn ngữ định nghĩa dữ liệu DDL (Data Definition Language)
  - Ngôn ngữ thao tác dữ liệu DML (Data Manipulation Language)
  - Ngôn ngữ truy vấn có cấu trúc SQL (Structured Query Language)
  - Ngôn ngữ điều khiển dữ liệu DCL (Data Control Language)
  - Từ điển dữ liệu (Data Dictionary)

- Có biện pháp bảo mật tốt khi có yêu cầu bảo mật
- Cơ chế giải quyết vấn đề tranh chấp dữ liệu
- Có cơ chế sao lưu và phục hồi dữ liệu
- Giao diện tốt, dễ sử dụng, dễ hiểu cho những NSD
- Bảo đảm tính độc lập giữa dữ liệu và chương trình



# Hình 1.6 - Sơ đồ tổng quát của DBMS



## ▪ CÁC MỨC BIỂU DIỄN CSDL

Theo kiến trúc **ANSI-PARC** (Standard Planning and Requirements Committee of the American National Standard Institute), một CSDL có 3 mức biểu diễn:

- ✓ mức trong (mức vật lý - *Physical*)
- ✓ mức quan niệm (*Conception / Logical*)
- ✓ mức ngoài

## a. *Mức trong*

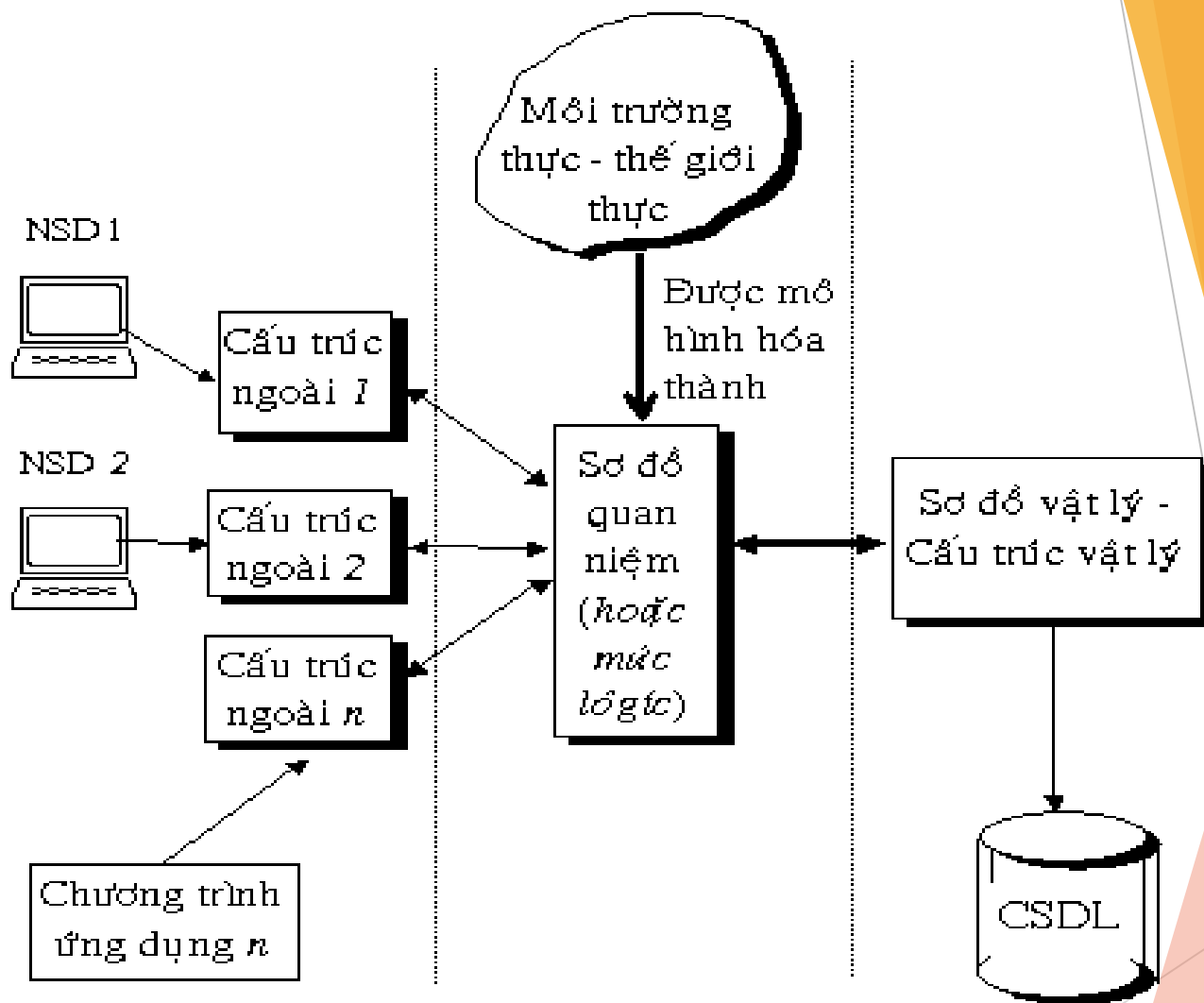
- Là mức quan tâm đến cách tổ chức vật lý của dữ liệu được lưu trữ trên phần cứng như thế nào? Mức trong mô tả cách dùng kỹ thuật của các byte và ở cấp độ máy
- Vấn đề cần giải quyết là dữ liệu gì và được lưu trữ như thế nào? ở đâu (đĩa từ, băng từ, rãnh, cung từ ...)? cần các chỉ mục gì? việc truy xuất là tuần tự (*sequential access*) hay ngẫu nhiên (*random access*) đối với từng loại dữ liệu.
- Những người làm việc với CSDL là người quản trị CSDL, NSD chuyên môn.

## ***b. Mức quan niệm***

- CSDL cần phải lưu trữ bao nhiêu loại dữ liệu? đó là những dữ liệu gì? Mỗi quan hệ giữa các loại dữ liệu này như thế nào?
  - Từ thế giới thực các chuyên viên tin học qua quá trình khảo sát và phân tích, cùng với những người sẽ đảm nhận vai trò quản trị CSDL, sẽ xác định được những loại thông tin gì là cần thiết phải đưa vào CSDL, đồng thời mô tả rõ mối liên hệ giữa các thông tin này
- ⇒ CSDL mức quan niệm là một sự biểu diễn trừu tượng CSDL mức vật lý, ngược lại, CSDL vật lý là sự cài đặt cụ thể của CSDL mức quan niệm

### c. *Mức ngoài*

- Là mức của NSD và các chương trình ứng dụng
  - Làm việc tại mức này có các nhà chuyên môn, các kỹ sư tin học và NSD không chuyên
  - Mỗi NSD hay mỗi chương trình ứng dụng có thể được "nhìn" CSDL theo một góc độ khác nhau. Có thể "nhìn" thấy toàn bộ hay chỉ một phần hoặc chỉ là các thông tin tổng hợp từ CSDL hiện có. NSD hay chương trình ứng dụng có thể hoàn toàn không được biết về cấu trúc tổ chức lưu trữ thông tin trong CSDL, thậm chí ngay cả tên gọi của các loại dữ liệu hay tên gọi của các thuộc tính. Họ chỉ có thể làm việc trên một phần CSDL theo cách "nhìn" do người quản trị hay chương trình ứng dụng quy định, gọi là khung nhìn (*View*)
- ⇒ Cấu trúc CSDL vật lý (mức trong) và mức quan niệm thì chỉ có một, nhưng tại mức ngoài, mức của các chương trình ứng dụng và NSD trực tiếp CSDL thì có thể có nhiều cấu trúc ngoài tương ứng



**Hình - Kiến trúc tổng quát ANSI - PARC của một hệ cơ sở dữ liệu**

- **Mô hình dữ liệu** là sự trừu tượng hóa môi trường thực, là sự biểu diễn dữ liệu ở mức quan niệm

- **Mô hình dữ liệu quan hệ**

**Mô hình dữ liệu quan hệ** (*Relational Data Model*) do E.F.Codd đề xuất.

Nền tảng cơ bản của nó là khái niệm lý thuyết tập hợp trên các quan hệ, tức là tập của các bộ giá trị

## ✓ Các khái niệm cơ bản

### • THUỘC TÍNH (ATTRIBUTE)

- Là một tính chất riêng biệt của một đối tượng cần được lưu trữ trong CSDL để phục vụ cho việc khai thác dữ liệu về đối tượng

- Ví dụ:

+ Loại thực thể SINH\_VIÊN có một số thuộc tính MÃ\_SV, HỌ\_TÊN\_SV, NGÀY\_SINH, QUÊ, MÃ\_KHOA

+ Loại thực thể XUẤT\_BẢN\_PHẨM có một số thuộc tính MÃ\_XBP, TÊN\_XBP, ĐƠN\_GIA, SỐ\_LƯỢNG

+ Loại thực thể NGHỆ\_SỸ có một số thuộc tính MÃ\_NS, HỌ\_TÊN\_NS, NGÀY\_SINH, QUÊ

...



- Đặc trưng bởi một tên gọi, kiểu giá trị và miền giá trị
- Mỗi thuộc tính có thể chỉ chọn lấy những giá trị trong một tập hợp con của kiểu dữ liệu. Tập hợp các giá trị mà một thuộc tính  $A$  có thể nhận được gọi là miền giá trị (*domain*) của thuộc tính  $A$  và được ký hiệu là  $DOM(A)$

- Ví dụ:

Sinh viên đang theo học tại trường đại học HUC thì tuổi nhiều nhất là 60 và tuổi ít nhất là 18

- **QUAN HỆ (RELATION)**

- Một quan hệ  $R$  có  $n$  ngôi được định nghĩa trên tập các thuộc tính  $U = \{A_1, A_2, \dots, A_n\}$  (thứ tự của các thuộc tính là không quan trọng) và kèm theo nó là một tân từ, tức là một quy tắc để xác định mối quan hệ giữa các thuộc tính  $A_i$  và được ký hiệu là  $R(A_1, A_2, \dots, A_n)$ . Tập thuộc tính của quan hệ  $R$  đôi khi còn được ký hiệu là  $R_+$

- Với  $A_i$  là một thuộc tính có miền giá trị là  $DOM(A_i)$ , như vậy,  $R(A_1, A_2, \dots, A_n)$  là tập con của tích Đề các:  $DOM(A_1) \times DOM(A_2) \times \dots \times DOM(A_n)$

- Quan hệ còn được gọi bằng thuật ngữ khác là bảng (*Table*)

- Ví dụ:

XUẤT\_BẢN\_PHẨM(MÃ\_XBP, TÊN\_XBP,  
ĐƠN\_GIA, SỐ\_LƯỢNG) là quan hệ 4 ngôi.

Tên từ: *"Mỗi Xuất bản phẩm có một Tên\_XBP,  
ĐƠN\_GIÁ, ... và được cấp một mã duy nhất để phân  
biệt với những Xuất bản phẩm khác trong nhà sách"*

- **BỘ GIÁ TRỊ (TUPLE)**

- Là các thông tin của một đối tượng thuộc quan hệ. Bộ giá trị thường được gọi là bản ghi (*record*) hoặc hàng của bảng (*row*).

- Về mặt hình thức, một bộ  $q$  là một vectơ gồm  $n$  thành phần thuộc tập hợp con của tích Đề các miền giá trị của các thuộc tính và thỏa mãn tân từ đã cho của quan hệ:

$$q = (a_1, a_2, \dots, a_n) \text{ (DOM}(A_1) \times \text{DOM}(A_2) \times \dots \times \text{DOM}(A_n))$$

- Ví dụ:

4 bộ giá trị dựa trên các thuộc tính của quan hệ SINH\_VIÊN:

$q_1 = (\text{TV42B001}, \text{Lương Văn Sang}, 27/03/1992, \text{Hưng Yên}, \text{TV42B})$

$q_2 = (\text{PHS30A005}, \text{Vũ Bích Nga}, 26/08/1992, \text{Lạng Sơn}, \text{PHS30A})$

$q_3 = (\text{PHS30B014}, \text{Đỗ Xuân Sơn}, 30/04/1991, \text{Khánh Hòa}, \text{PHS30B})$

$q_4 = (\text{VHDL21B015}, \text{Lê Hoài Hà}, 23/10/1994, \text{Hà Nam}, \text{VHDL21B})$

- **THỂ HIỆN CỦA QUAN HỆ**

- Thể hiện TR (hoặc *tình trạng*) của quan hệ R là tập hợp các bộ giá trị của quan hệ R vào một thời điểm.
- Tại những thời điểm khác nhau thì quan hệ sẽ có những thể hiện khác nhau. Thể hiện của các lược đồ quan hệ con  $TR_i$  gọi là tình trạng của lược đồ CSDL **C**

- Ví dụ:

Thể hiện của quan hệ MÔN\_HỌC

<b>Mã_môn_học</b>	<b>Tên_môn_học</b>	
<b>SỐ_Tín_chỉ</b>		
THĐC	Tin học đại cương	3
THQL	Tin học quản lý	2
TKW	Thiết kế Web	2
CBĐT	Chế bản điện tử	2
MMT	Mạng máy tính	2
TOAN	Toán cao cấp	3
TMDT	Thương mại điện tử	3

- **KHÓA (KEY)**

Có nhiều cách khác nhau để định nghĩa khóa:

- ✓ **Định nghĩa 6:**

Khóa của lược đồ quan hệ R định nghĩa trên tập các thuộc tính  $U = \{A_1, A_2, \dots, A_n\}$  là một tập con  $K \subseteq U$  thỏa mãn các tính chất sau: với bộ giá trị  $q_1, q_2$  của R đều một thuộc tính  $A \in K$  sao cho  $q_1.A \neq q_2.A$ .

⇒ ! **2 bộ có giá trị bằng nhau trên thuộc tính của K:**

$$q_1.K = q_2.K$$

⇒ Như vậy, mỗi giá trị của khóa K phải là xác định duy nhất trên quan hệ R.

Theo định nghĩa trên, nếu  $K' \subseteq K \subseteq U$  là khóa của lược đồ quan hệ R thì K cũng là khóa của R

bởi vì  $q_1.K' \neq q_2.K'$  thì có  $q_1.K \neq q_2.K$



- ⇒ Như vậy, trong lược đồ quan hệ có thể có rất nhiều khóa. Việc xác định tất cả các khóa của một lược đồ quan hệ là rất khó khăn
- ⇒ Định nghĩa 6 là chưa chặt chẽ

✓ **Định nghĩa 7:**

Quan hệ R định nghĩa trên tập các thuộc tính

$$U = \{A_1, A_2, \dots, A_n\}$$

$K \subseteq U$  là khóa của quan hệ R nếu thỏa 2 điều kiện sau đây:

- (i)  $K$  xác định được giá trị của  $A_j$  với  $j = 1, 2, \dots, n$
- (ii)  $\nexists K' \subset K$  mà  $K'$  có thể xác định được giá trị của  $A_j$  với  $j = 1, 2, \dots, n$

⇒  **$K$  là tập con nhỏ nhất** mà giá trị của nó có thể xác định duy nhất một bộ giá trị của quan hệ

- Khóa của quan hệ theo định nghĩa 7 được gọi là **khóa dự bị** (*candidate*) và là **khóa nội** của quan hệ
  - $K$  là **siêu khóa** của quan hệ  $R$  nếu  $K'$  là một khóa của quan hệ  
 Một lược đồ quan hệ  $Q$  của quan hệ  $R$  luôn có ít nhất một siêu khóa và có thể có nhiều siêu khóa
  - Ý nghĩa thực tế của khóa là dùng để nhận diện một bộ trong một quan hệ
- ⇒ Khi cần truy tìm một bộ  $q$  chỉ cần biết giá trị của thành phần khóa của  $q$  là đủ để dò tìm và hoàn toàn xác định được nó trong quan hệ

- Trong trường hợp lược đồ quan hệ Q có nhiều khóa chỉ định, khi cài đặt trên một DBMS, NSD có thể chọn một trong số các khóa dự bị để tạo **chỉ mục** (*index*) chỉ phối việc truy cập đến các bộ. Khóa dự bị này được gọi là *khóa chính* (*primary key*)
- Các khóa còn lại gọi là các *khóa tương đương*
- ⇒ Khóa chính chỉ thật sự có ý nghĩa trong quá trình khai thác CSDL và về lý thuyết, khóa chính hoàn toàn không có vai trò gì khác so với các khóa dự bị còn lại
- ✓ Một số DBMS như Microsoft Access, ORACLE, DB2 ... có cài đặt cơ chế tự động kiểm tra tính duy nhất trên khóa chính. Nghĩa là, nếu thêm một bộ mới  $q_2$  có giá trị khóa chính trùng với giá trị khóa chính của một bộ  $q_1$  nào đó đã có trong quan hệ thì hệ thống sẽ báo lỗi và yêu cầu nhập lại một giá trị khác

## 👉 Qui ước:

- Trong một bộ của một quan hệ các thuộc tính khóa không chứa giá trị rỗng.
  - Không được phép sửa đổi giá trị của thuộc tính khóa. Nếu muốn sửa đổi giá trị thuộc tính khóa của một bộ q, NSD phải hủy bỏ bộ q và sau đó, thêm mới một bộ q' với giá trị khóa đã được sửa đổi.
  - Các thuộc tính có tham gia vào một khóa được gọi là thuộc tính khóa. Ngược lại, các thuộc tính không tham gia vào một khóa nào gọi là thuộc tính không khóa
- Ví dụ:
- + KHOA (Mã\_khoa, Tên\_khoa)
  - + MÔN\_HỌC (Mã\_môn\_học, Tên\_môn\_học, Số\_Tín\_chỉ)
  - + SINH\_VIÊN (Mã\_SV, Họ\_Tên\_SV, Ngày\_sinh, Quê, Mã\_khoa)
  - + KẾT\_QUẢ\_THI (Mã\_SV, Mã\_môn\_học, Lần\_thi, Ngày\_thi, Điểm\_thi, Ghi\_chú)

- **Khóa ngoại** (Foreign Key)

Giả sử có hai quan hệ R và S.

Một tập thuộc tính K của quan hệ R được gọi là khóa ngoại của quan hệ R nếu K là khóa nội của quan hệ S

- Ví dụ:

Mã\_khoa trong quan hệ SINH\_VIÊN là khóa ngoại vì nó là khóa nội của quan hệ KHOA

- **PHỤ THUỘC HÀM** (FUNCTIONAL DEPENDENCY)

Quan hệ R được định nghĩa trên tập thuộc tính  $U = \{ A_1, A_2, \dots, A_n \}$

$X, Y \subseteq U$  là 2 tập con của tập thuộc tính U. Nếu tồn tại một ánh xạ  $f : X \rightarrow Y$  thì ta nói rằng X xác định hàm Y, hay Y phụ thuộc hàm vào X và ký hiệu là  $X \twoheadrightarrow Y$

- **RÀNG BUỘC TOÀN VỆN**

(INTEGRITY CONSTRAINT / RULE)

- Là một quy tắc định nghĩa trên một hay nhiều quan hệ do môi trường ứng dụng quy định  $\Rightarrow$  quy tắc để đảm bảo tính nhất quán của dữ liệu trong CSDL.

- Mỗi RBTV được định nghĩa bằng một thuật toán trong CSDL

- Ví dụ:

**SINH\_VIÊN (MA\_SV, HọTên\_SV, Ngày\_sinh, Quê, Mã\_lớp)**

*Quy tắc:* Ngày\_sinh của sinh viên phải  $\geq\{01/01/1953\}$  và  $\leq\{31/12/1995\}$

*Thuật toán:* sv SINH\_VIÊN thì sv.Ngày\_sinh  $\geq\{01/01/1953\}$   
& sv.Ngày\_sinh  $\leq\{31/12/1995\}$

✓ **Các phép toán cơ bản trên các quan hệ: 3**

▶ **PHÉP CHÈN THÊM MỘT BỘ MỚI VÀO QUAN HỆ**

Việc chèn thêm một bộ giá trị mới  $t$  vào quan hệ

$R(A_1, A_2, \dots, A_n)$  làm cho thể hiện TR của nó tăng thêm một phần tử mới:  $TR = TR \cup t$

Dạng hình thức:

$INSERT(R; A_{i_1}=v_1, A_{i_2}=v_2, \dots, A_{i_m}=v_m)$

trong đó:

$A_{i_1}, A_{i_2}, \dots, A_{i_m}$  là các thuộc tính

$v_1, v_2, \dots, v_m$  là các giá trị thuộc  $DOM(A_{i_1}), DOM(A_{i_2}), \dots,$

$DOM(A_{i_m})$  tương ứng



- Ví dụ: Quan hệ

SINH\_VIÊN(Mã\_SV, Họ\_Tên\_SV, Ngày\_sinh, Quê, Mã\_lớp)

Chèn thêm bộ  $q_5=(TV42B002, Hoàng Thu Trang, 17/05/1994, Hà Nội, TV42B)$  vào quan hệ SINH\_VIÊN bởi phép thêm như sau:

INSERT(SINH\_VIÊN; [Mã\_SV] = TV42B002, Họ\_Tên\_SV = Hoàng Thu Trang, [Ngày\_sinh]=17/05/1994, [Quê]=Hà Nội, [Mã\_lớp]=TV42B).

Thể hiện  $T_{SINH\_VIÊN}$  :

$q_1 = (TV42B001, Lương Văn Sang, 27/03/1992, Hưng Yên, TV42B)$

$q_5 = (TV42B002, Hoàng Thu Trang, 17/05/1994, Hà Nội, TV42B)$

$q_2 = (PHS30A005, Vũ Bích Nga, 26/08/1992, Lạng Sơn, PHS30A)$

$q_3 = (PHS30B014, Đỗ Xuân Sơn, 30/04/1991, Khánh Hòa, PHS30B)$

$q_4 = (VHDL21B015, Lê Hoài Hà, 23/10/1994, Hà Nam, VHDL21B)$

## ► PHÉP XOÁ BỘ KHỎI QUAN HỆ

Phép xoá một bộ t của quan hệ sẽ lấy đi bộ t khỏi thể hiện của quan hệ:  $TR = TR \setminus t$

Dạng hình thức:

DELETE (R;  $A_{i_1} = v_1, A_{i_2} = v_2, \dots, A_{i_m} = v_m$ )

trong đó:

$A_{i_j} = v_j$  ( $j = 1, 2, \dots, m$ ) là những điều kiện thỏa một số thuộc tính của bộ t để xoá một bộ ra khỏi quan hệ

- Ví dụ:

Quan hệ SINH\_VIÊN(Mã\_SV, Họ\_Tên\_SV, Ngày\_sinh, Quê, Mã\_lớp)

Với phép xoá như sau:

DELETE (SINH\_VIÊN; [Quê] = Hưng Yên)

Thì bộ  $q_1 = (TV42B001, Lương Văn Sang, 27/03/1992, Hưng Yên, TV42B)$

sẽ bị xoá ra khỏi quan hệ SINH\_VIÊN bởi vì có Quê là Hưng Yên.

Khi đó, thể hiện  $T_{SINH\_VIÊN}$  :

$q_2 = (PHS30A005, Vũ Bích Nga, 26/08/1992, Lạng Sơn, PHS30A)$

$q_3 = (PHS30B014, Đỗ Xuân Sơn, 30/04/1991, Khánh Hòa, PHS30B)$

$q_4 = (VHDL21B015, Lê Hoài Hà, 23/10/1994, Hà Nam, VHDL21B)$

$q_5 = (TV42B002, Hoàng Thu Trang, 17/05/1994, Hà Nội, TV42B)$

## ► PHÉP CẬP NHẬT GIÁ TRỊ CỦA CÁC THUỘC TÍNH

Dữ liệu của CSDL đôi khi cũng cần phải được đổi mới theo thời gian hoặc sửa lại cho đảm bảo tính chính xác hoặc nhất quán của dữ liệu. Do đó, thao tác cập nhật dữ liệu là rất cần thiết. Một số DBMS đưa ra nhiều câu lệnh khác nhau để cập nhật dữ liệu

Dạng hình thức:

UPDATE (R;  $A_{i_1} = c_1, A_{i_2} = c_2, \dots, A_{i_m} = c_m; A_{i_1} = v_1, A_{i_2} = v_2, \dots, A_{i_m} = v_m$ )

trong đó:

R là quan hệ cần thực hiện cập nhật,

$A_{i_j} = c_j$  ( $j = 1, 2, \dots, m$ ) là điều kiện tìm kiếm bộ giá trị để cập nhật và

$A_{i_j} = v_j$  ( $j = 1, 2, \dots, m$ ) là giá trị mới của bộ

- Ví dụ:

Quan hệ SINH\_VIÊN(Mã\_SV, Họ\_Tên\_SV, Ngày\_sinh, Quê, Mã\_lớp)

Với phép cập nhật giá trị như sau:

UPDATE (SINH\_VIÊN;[Mã\_SV]=VHDL21B015, [Quê]=Nam Định)

thì giá trị của bộ q4 được cập nhật thành:

$q_4 = (VHDL21B015, \text{Lê Hoài Hà}, 23/10/1994, \text{Nam Định}, VHDL21B)$

## ▪ Phương pháp Thiết kế cơ sở dữ liệu: 8 bước

1. Phân tích toàn bộ yêu cầu
2. Xác định thực thể
3. Xác định mối tương quan giữa các thực thể
4. Xác định trường khoá chính
5. Xác định trường khoá ngoại
6. Thêm các trường không phải trường khoá vào bảng dữ liệu
7. Xây dựng mạng dữ liệu
8. Khai báo phạm vi của mỗi trường

# NGHIÊN CỨU TRƯỜNG HỢP: Thiết kế cơ sở dữ liệu cho nhà sách/siêu thị sách

# TÀI LIỆU THAM KHẢO

1. Date C.J. (1995), *An Introduction to Database Systems*, 6<sup>th</sup> Edition, Addison-Wesley, Massachusetts.
2. Hawryszkiewicz I. (1998), *Introduction to Systems Analysis and Design*, 4<sup>th</sup> Edition, Prentice Hall, New Jersey.
3. Laudon K.C., Laudon J.P. (2004), *Management Information Systems*, 8<sup>th</sup> Edition, Prentice Hall, New Jersey.
4. A.J. Fabbri, A.R. Schwab (1999), *Quản trị cơ sở dữ liệu*, Trần Đức Quang biên dịch, Nxb Thống kê, TP. Hồ Chí Minh.
5. Phạm Thị Thanh Hồng, Phạm Minh Tuấn (2007), *Hệ thống thông tin quản lý*, Nxb Khoa học và kỹ thuật, Hà Nội.
6. Nguyễn Thanh Hùng (2001), *Hệ thống thông tin quản lý*, Nxb Đại học Quốc gia TP Hồ Chí Minh.
7. Đỗ Quang Vinh (chủ biên) (2009), *Ứng dụng Công nghệ thông tin trong quản lý văn hóa*, Nxb Đại học Quốc gia Hà Nội.
8. Đỗ Quang Vinh (chủ biên) (2010), *Hệ quản trị cơ sở dữ liệu Microsoft Visual FoxPro và Ứng dụng*, xuất bản lần thứ 2, Nxb Đại học Quốc gia Hà Nội.



**TRÂN TRỌNG CẢM ƠN !**