

- **Bộ tập trung (Multiplexer)** là thiết bị tập trung nhiều luồng thông tin vào một kênh truyền hoặc tách thông tin từ một kênh truyền ra. Chẳng hạn, một cáp quang truyền cùng một lúc được 60000 cuộc điện thoại thì cần phải có thiết bị dồn các cuộc gọi lại rồi truyền đi và ở đầu kia cần phải tách ra dẫn đến các máy đơn lẻ.
- **Giao thức truyền thông (Communications Protocol)** là các quy tắc và các thủ tục quy định thống nhất để thực hiện các nhiệm vụ truyền thông. Các quy trình và thủ tục thường được các phần mềm quản trị truyền thông thực hiện.
- **Phần mềm mạng (WAN software):** mạng WAN cần có các chương trình để điều hành hoạt động và thực hiện các ứng dụng trên mạng. Phần mềm mạng thường bao gồm cả các chương trình quản lý truy cập và truyền thông.

#### 4.3.3. Mạng MAN (Metropolitan Area Network)

Đây là một dạng mạng truyền dữ liệu và thông tin trên phạm vi một thành phố và có tốc độ đường truyền thường lớn hơn mạng LAN (khoảng trên 200 megabits mỗi giây). Mạng MAN cũng thường được thiết kế để truyền nhiều dạng thông tin hơn mạng LAN, như nó có thể truyền được hỗn hợp các âm thanh, dữ liệu, hình ảnh, và phim chuyển động. MAN không được thiết kế với các đường điện thoại, để có thể truyền các dữ liệu kết hợp với tốc độ cao và trên diện toàn thành phố, người ta thường sử dụng các đường cáp quang.

#### 4.3.4. Mạng Internet

Các mạng chuyên mạch gói và chuyên mạch gói diện rộng cho phép người sử dụng truy cập vào các nguồn tài nguyên trong một mạng duy nhất. Tuy nhiên, một mạng đơn lẻ sẽ không còn đáp ứng nhu cầu của người sử dụng, do đó, cần phải liên kết mạng lại với nhau. Nhưng các mạng lẻ lại có cách tổ chức khác nhau, vì vậy, cần phải có một cách thức nào đó để truy nhập liên thông giữa các mạng. Nhu cầu này đã làm nảy sinh và phát triển kỹ thuật mạng Internet.

Có thể hiểu Internet là mạng của các mạng có phạm vi toàn cầu, sử dụng nhiều loại phương tiện truyền thông khác nhau cung cấp nhiều loại dịch vụ trên mạng. Mạng Internet không có chủ nhân riêng mà nhiều chủ nhân, mỗi chủ nhân làm chủ một phần như hệ thống giao thông quốc tế.

Các thành phần của mạng Internet bao gồm:

- **Mạng con (Subnetwork)** là một mạng thành phần của mạng Internet. Để tránh nhầm lẫn với toàn mạng, ta có thể dùng thuật ngữ

mạng riêng lẻ. Mạng con có thể là mạng LAN, mạng WAN, hoặc mạng MAN.

- **Đầu cuối (End System)** là thiết bị được gắn vào mạng con của mạng Internet, được dùng để trợ giúp các ứng dụng của người sử dụng cuối.
- **Hệ thống trung gian (Intermediate System)** là thiết bị được sử dụng để nối tiếp hai mạng con với nhau và cho phép truyền thông giữa hai hệ thống đầu cuối gắn vào hai mạng khác nhau.
- **Cầu nối (Bridge)** là một hệ thống trung gian IS dùng để nối hai mạng LAN có cùng một giao thức. Cầu nối hoạt động như một bộ lọc địa chỉ, nhặt ra những gói tin từ một mạng LAN cần được truyền đến một địa chỉ thuộc mạng LAN khác và chuyển những gói đó sang mạng thứ hai. Cầu nối không làm thay đổi nội dung cũng không thêm dữ liệu vào các gói tin ấy.
- **Bộ định tuyến (Router)** là một hệ thống trung gian IS dùng để nối hai mạng có thể khác giao thức đường truyền. Bộ định tuyến sử dụng giao thức Internet trên các nút và trên các máy chủ mạng.
- **Giao thức Internet (Internet Protocol)** là toàn bộ các quy tắc và thủ tục quy ước được sử dụng để thực hiện việc truyền thông trên mạng Internet. Hiện nay người ta sử dụng chủ yếu là giao thức TCP/IP (Transmission Control Protocol/ Internet Protocol). Ngoài ra có một giao thức chuẩn OSI (Open System Interconnection) dùng để tra cứu khi thiết kế và xây dựng giao thức thực tế.

#### 4.4. Dịch vụ mạng và công nghệ băng thông rộng

##### 4.4.1. Mạng giá trị gia tăng (Value-Added Network)

Mạng giá trị gia tăng (VAN) là mạng riêng, quản lý bởi bên thứ ba cung cấp các dịch vụ chuyển tải thông tin và mạng cho các công ty thuê bao. Thuê bao chỉ phải chi trả cho lượng dữ liệu họ chuyển tải bên cạnh một khoản chi phí thuê bao cố định. Khách hàng không cần phải đầu tư vào phần mềm và trang thiết bị mạng và cả chi phí nối mạng bởi chi phí sử dụng mạng được chia sẻ bởi nhiều khách hàng.

Nhiều công ty hiện đang sử dụng Internet để chuyển tải dữ liệu bởi nó tiết kiệm hơn là sử dụng mạng giá trị gia tăng. Để đối phó, các mạng giá trị gia tăng ngày nay tăng cường cung cấp các dịch vụ phụ trợ bảo vệ quản lý email và chuyển tải dữ liệu, báo cáo quản lý và các dịch vụ trao đổi dữ liệu điện tử.

#### 4.4.2. Các dịch vụ mạng khác

**Chuyển đổi gói tin** (packet switching) là công nghệ chuyển đổi cơ bản dùng để tách các dòng dữ liệu thành những gói nhỏ, gửi đi theo những tuyến đường có hiệu quả nhất rồi thực hiện lắp ráp lại và kiểm tra độ chính xác tại đích đến.

**Mạng chuyển khung** (Frame relay) là công nghệ mạng giúp chia sẻ nhanh và kinh tế hơn so với dịch vụ chuyển đổi gói tin, bởi có thể đạt được tốc độ chuyển tải tối đa là 1.544 megabit/giây. Công nghệ này gộp dữ liệu thành các khung tương tự như gói, nhưng không thực hiện chữa lỗi. Vì vậy, frame relay hoạt động tốt trên các đường truyền ổn định không hay đòi hỏi tái chuyển tải dữ liệu khi gặp lỗi.

**Mạng chuyển dẫn không đồng bộ - ATM** (asynchronous transfer mode) là công nghệ chia thông tin thành các phần tử 8 bit, cho phép dữ liệu được truyền giữa các máy tính khác nhau. ATM hứa hẹn liên kết các LAN và WAN lại với nhau một cách dễ dàng. ATM có thể chuyển tải 10 gigabit/giây.

**Mạng số dịch vụ tích hợp - ISDN** (Integrated Services Digital Network) là dịch vụ truy cập mạng đạt tiêu chuẩn quốc tế sử dụng đường dây điện thoại, có khả năng tích hợp âm thanh, hình ảnh, video và dữ liệu trong một đường truyền đơn. Có hai mức dịch vụ ISDN: mức cơ sở (Basic Rate ISDN) và mức sơ cấp (Primary Rate ISDN). Mức cơ sở có thể chuyển tải dữ liệu với tốc độ 128 kilobit/giây dành cho người sử dụng bình thường. Mức sơ cấp được thiết kế cho các công ty dịch vụ viễn thông với tốc độ chuyển tải lên tới megabit.

**Dịch vụ thuê bao kỹ thuật số** (digital subscriber line - DSL) giống như ISDN cũng thông qua đường dây điện thoại để tải âm thanh, dữ liệu, video nhưng có khả năng chuyển tải cao hơn. Thuê bao kỹ thuật số không đối xứng (asymmetric digital subscriber line - ADSL) cho phép nhận dữ liệu với tốc độ 1,5 tới 9 megabit/giây và gửi dữ liệu với tốc độ 640 kilobit/giây. Thuê bao kỹ thuật số đối xứng (symmetric digital subscriber line - SDSL) có cùng tốc độ gửi với ADSL, còn tốc độ nhận tối đa là 3 megabit/giây.

**Bộ điều giải cáp** (cable modem) được thiết kế riêng cho đường truyền hình cáp, cung cấp khả năng truy cập Internet hoặc mạng nội bộ với tốc độ cao lên tới 4 megabit/giây. Bộ điều giải cáp hiện tại có khả năng nhận dữ liệu cao hơn gửi dữ liệu.

**Băng thông rộng** (broadband) là công nghệ truyền dữ liệu tốc độ cao. Thuật ngữ băng thông rộng còn dùng để chỉ phương tiện chuyển tải nhiều kênh thông tin cùng một lúc thông qua một thiết bị truyền thông đơn.

## 5. Nhân lực

Một nguồn lực quan trọng khác của hệ thống thông tin quản lý đó là nhân lực. Con người (những nhân công kiến thức) thiết lập các mục tiêu, xác định nhiệm vụ, tạo quyết định, phục vụ khách hàng, và trong trường hợp là các chuyên gia công nghệ thông tin, còn có thể cung cấp một môi trường công nghệ đáng tin cậy và ổn định cho một tổ chức. Với sự hỗ trợ của nhân lực, tổ chức sẽ nhận được ưu thế cạnh tranh trong thị trường.

### 5.1. Sự hiểu biết về công nghệ và thông tin

Trong doanh nghiệp, tài sản quý giá nhất không phải là công nghệ mà chính là trí tuệ của nhân công mà nó sở hữu. Công nghệ thông tin đơn giản chỉ là một tập công cụ cho phép chúng ta làm việc và xử lý các thông tin; nó chỉ đóng vai trò là một bộ công cụ hỗ trợ cho quá trình tư duy. Các phần mềm bảng tính cho phép chúng ta nhanh chóng đưa ra được các lược đồ biểu diễn số liệu và thông tin đầy tính thuyết phục và có chất lượng cao. Nhưng nó không thể nói với chúng ta trong trường hợp này thì nên chọn kiểu biểu đồ nào biểu đồ thanh ngang hay là biểu đồ dạng hình tròn. Nó cũng không thể nói với chúng ta xem chúng ta nên chọn biểu diễn doanh số bán hàng theo từng vùng thị trường hay là theo từng người bán hàng. Tất cả những vấn đề đó chính là nhiệm vụ của các nhân công ở trong doanh nghiệp.

Tuy nhiên, công nghệ thông tin vẫn là một tập công cụ rất tốt cho chúng ta. Nó giúp chúng ta làm việc có hiệu quả hơn, xử lý các dữ liệu nhanh hơn. Chính vì vậy, chúng ta cần nắm vững cách thức làm thế nào để sử dụng công nghệ một cách tốt nhất. Và đồng thời, chúng ta cũng cần hiểu được các thông tin mà công nghệ đã giúp chúng ta tạo ra đó. Nói một cách khác, doanh nghiệp cần phải có nguồn nhân lực có kiến thức về công nghệ và thông tin.

Một nhân công có kiến thức về công nghệ là người biết rõ cách thức ứng dụng và khi nào ứng dụng công nghệ thông tin. Cách thức ở đây gồm cả việc chúng ta nên mua những công nghệ nào, làm thế nào để khai thác được hiệu quả các phần mềm ứng dụng, và các cơ sở kỹ thuật nào là cần thiết để giúp cho doanh nghiệp của chúng ta có thể kết nối với các doanh nghiệp khác.

Một nhân công có kiến thức về thông tin là người có thể xác định được loại thông tin nào là cần, biết cách để có thể có được thông tin đó, hiểu rõ về thông tin một khi có được nó, và có thể hành động hợp lý căn cứ vào thông tin đã nhận được để đem lại lợi ích tối đa cho doanh nghiệp.

## **5.2. Trách nhiệm đạo đức đối với xã hội**

Việc có kiến thức về công nghệ và thông tin mới là trách nhiệm của một nhân công đối với doanh nghiệp mà người đó phục vụ. Người đó cũng cần phải có trách nhiệm đối với xã hội: đó là khi đạo đức trở thành một yếu tố quan trọng. Đạo đức là những nguyên tắc và tiêu chuẩn dẫn hướng các hành động của chúng ta đối với những người khác. Đạo đức là một khái niệm khác với luật pháp. Luật pháp đòi hỏi hoặc ngăn cản một số hành động của con người. Đạo đức chỉ là sự diễn giải của riêng một cá nhân về cái gì là đúng và cái gì là sai.

Trong lĩnh vực liên quan tới công nghệ thông tin và hệ thống thông tin, khái niệm đạo đức trở nên khá quan trọng. Do với sự phát triển của công nghệ thông tin, con người có thể nhận được rất nhiều thông tin. Việc xử lý và sử dụng những thông tin đó như thế nào để có thể làm lợi cho doanh nghiệp của mình mà không gây ra những hoạt động phi đạo đức đối với xã hội cũng đòi hỏi lực lượng nhân công phải có hiểu biết rộng hơn và đa dạng hơn, không chỉ là có liên quan tới kỹ thuật và công nghệ mà còn liên quan tới môi trường xã hội xung quanh nữa.

## **5.3. Bộ máy nhân sự công nghệ thông tin trong doanh nghiệp**

Khi thành lập bộ máy nhân sự công nghệ thông tin, doanh nghiệp cần chú ý tới những thành phần sau:

- Quản trị viên hệ thống (System Administrator)
- Lập trình viên (Programmer)
- Nhà thiết kế hệ thống (System Designer)
- Nhà phân tích hệ thống (System Analyst)
- Trưởng phòng CNTT
- GD Dự án
- Phó TGD phụ trách CNTT (Chief Information Officer - CIO)

## **Câu hỏi ôn tập và thảo luận**

1. Phần cứng của hệ thống thông tin bao gồm những thành phần nào?
2. Hãy trình bày những loại phần mềm mà anh/chị biết?
3. Liệt kê tên và tính năng của các loại thiết bị ra của máy tính
4. Liệt kê tên và tính năng của các loại thiết bị vào của máy tính
5. Nhà quản lý cần quan tâm tới những vấn đề gì khi lựa chọn mua phần cứng và phần mềm?
6. Hãy mô tả các phương thức kết nối mạng
7. Mô tả các dạng mạng LAN, WAN, và mạng Internet.
8. Cầu nối khác với bộ định tuyến ở những điểm nào?
9. Thế nào là dịch vụ mạng và công nghệ băng thông rộng?
10. Doanh nghiệp cần chú ý những vấn đề gì khi phát triển các thành phần của HTTT quản lý?

## Chương 3

# THIẾT KẾ CƠ SỞ DỮ LIỆU

**Mục đích:** Chương này cung cấp một số yếu tố kỹ thuật có liên quan tới việc thiết kế một cơ sở dữ liệu trong HTTT.

**Nội dung chính:** Cách thức xây dựng cơ sở dữ liệu, thiết lập mối quan hệ giữa các dạng thực thể trong hệ thống, và các kỹ thuật xây dựng CSDL hiện đang được sử dụng.

## 1. Tập và cơ sở dữ liệu

Chúng ta đã nói nhiều tới việc sử dụng HTTT quản lý như một vũ khí chiến lược của các doanh nghiệp trong thời đại bùng nổ thông tin hiện nay. Vậy, làm thế nào để có thể xây dựng được một HTTT phục vụ được kịp thời những mục đích của doanh nghiệp? Trước khi nói tới vấn đề sử dụng hệ thống tin như thế nào, ta cần nói tới khía cạnh làm thế nào để ta có được một HTTT hiệu quả, nghĩa là nói tới cái gốc căn bản của HTTT - cơ sở dữ liệu. Từ trước tới nay, người ta đã nhiều lần tìm cách định nghĩa và phân loại các cơ sở dữ liệu. Tuy vậy, đứng trên phương diện là người quản lý doanh nghiệp, một hiểu biết đúng đắn về cơ sở dữ liệu sẽ giúp chúng ta rất nhiều trong bước đầu tiên - bước xây dựng cơ sở dữ liệu.

### 1.1. Tập và hệ thống tập dữ liệu

Những ứng dụng ban đầu của HTTT mà chúng ta nhận thấy đầu tiên đó là những ứng dụng cho những người làm công việc thư ký và hạch toán sổ sách trong một doanh nghiệp. Những ứng dụng này chủ yếu tập trung vào việc thực hiện việc đặt hàng, xuất hàng, lập kế hoạch làm việc, lập bảng trả lương hàng tháng, v.v... Tất cả những thông tin phục vụ cho những ứng dụng này, trước đây, được ghi vào sổ sách với thứ tự xác định để giúp những người sử dụng có khả năng lập báo cáo hoặc các bản tổng hợp tình hình một cách nhanh nhất có thể. Ngày nay, khi khả năng công nghệ thông tin với việc sử dụng máy tính đã trở nên phổ biến, tất cả những dữ liệu này bắt đầu được hệ thống lại và ghi dưới dạng các tập dữ liệu trong máy tính. Các dạng tập này cùng với một số các chương trình phần mềm trợ giúp, giúp cho người sử dụng có khả năng ghi dữ liệu không cần theo thứ tự nhưng vẫn có khả năng tạo báo cáo chính xác và nhanh chóng. Hơn nữa, việc tổ chức các tập một cách khoa học sẽ giúp cho việc tạo các cơ sở dữ liệu và các chương trình phân tích và tổng hợp dữ liệu từ những tập chứa dữ liệu có sẵn đó trở

nên dễ dàng hơn và hiệu quả hơn.

Bảng 3.1 ghi lại một ví dụ về tệp dữ liệu chứa thông tin về các khách hàng của một doanh nghiệp.

**Bảng 3.1.** Nội dung của tệp dữ liệu về các khách hàng của một công ty bảo hiểm

Số TT	Tên khách hàng	Số điện thoại	Địa chỉ	Dạng bảo hiểm	Số tiền bảo hiểm	Ngày cập nhật thông tin
1	Nguyễn Văn A	8300154	124, Triệu Việt Vương	T1	200 000	03/12/94
2	Lê Vũ Quỳnh N	8697124	115, Quan Thánh	S1	500 000	15/02/93
3	Hoàng Anh T	8546210	20B, Quốc Tử Giám	S2	450 000	31/12/94

Trong ví dụ về tệp dữ liệu ở trên, người sử dụng sẽ rất dễ dàng thiết lập các báo cáo dựa trên các thông tin này. Họ có thể dễ dàng thống kê được những dạng bảo hiểm nào hay được mua nhất, những người làm ngành nghề nào thường sẽ mua bảo hiểm loại nào. Những báo cáo này nói chung sẽ được lập lại vào cuối mỗi quý, mỗi năm, hoặc thậm chí vào bất cứ lúc nào mà người sử dụng cần những báo cáo dạng này.

Theo cách quản lý trước đây, người ta thường tổ chức các tệp dữ liệu này thành một hệ thống và tạo các mối liên hệ giữa các tệp để dễ dàng truy tìm thông tin và tạo các báo cáo. Việc tổ chức các file theo dạng hệ thống file như vậy mặc dù có vẻ rất logic và dễ hiểu, nhưng thực tế, nó rất phức tạp đặc biệt là khi số tệp lên tới 20 tệp cần kết nối với nhau. Hơn thế nữa, các tệp chương trình giúp người sử dụng làm việc với các tệp này thường chịu ảnh hưởng rất lớn từ cấu trúc của mỗi tệp. Mỗi khi một tệp bị thay đổi cấu trúc, lập tức các chương trình liên quan cũng buộc phải thay đổi theo cho phù hợp với cấu trúc mới này. Đôi khi việc thay đổi này là rất nhỏ, không đáng kể, nhưng nó cũng làm phức tạp, và mất thời gian đối với người sử dụng. Nhất là xét tới sự chậm trễ để thay đổi chương trình điều khiển như vậy làm ảnh hưởng tới các công việc khác của một doanh nghiệp.

Chính vì những hạn chế này, ngày nay, người ta đã tiến tới thiết kế các cơ sở dữ liệu chứ không chỉ còn thiết kế các hệ thống tệp dữ liệu như trước đây nữa.

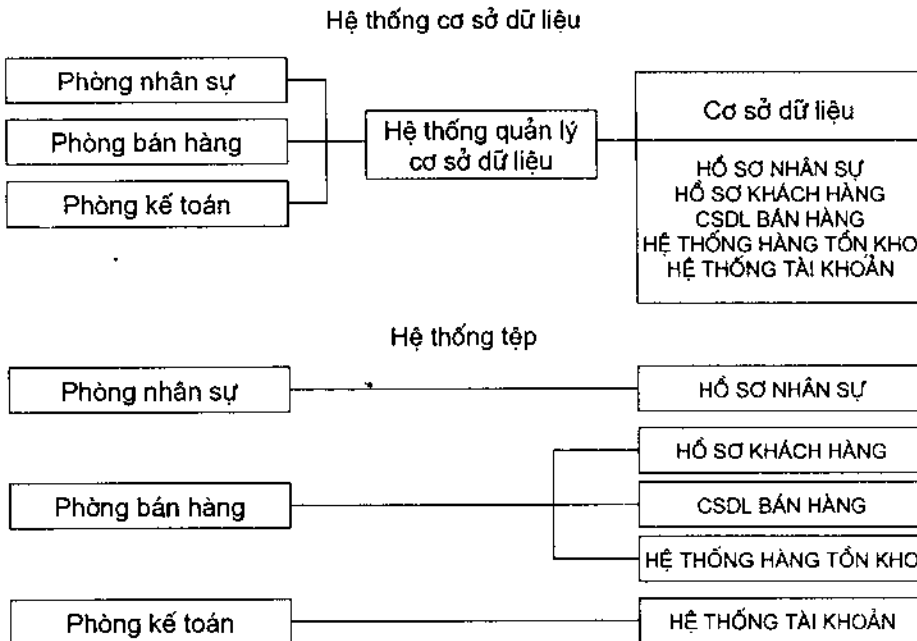


**Bảng 3.2.** Một số thuật ngữ cơ bản

Thuật ngữ	Khái niệm
Dữ liệu	Dữ liệu là những sự kiện thực tế được tổ chức và lưu lại trong máy tính
Trường	Một hoặc một nhóm ký tự được sử dụng để làm tiêu đề cho một nhóm các dữ kiện được lưu lại.
Biểu ghi	Một tập các trường được liên kết lại với nhau liên quan tới một người, một nơi chốn, một vật hoặc một sự kiện nào đó.
Tệp	Một hoặc nhiều các biểu ghi được ghi theo cùng một tiêu chuẩn thống nhất nghĩa là được tổ chức theo cùng một số trường nhất định.

### 1.2. Hệ thống quản lý cơ sở dữ liệu

Khác với hệ thống các tệp dữ liệu, các dữ liệu được lưu trong nhiều tệp khác nhau, cơ sở dữ liệu chứa tất cả các dữ liệu có liên quan trong một kho dữ liệu duy nhất (xem hình 3.1). Hiện nay, không những chỉ cấu trúc chung của dữ liệu được lưu giữ trong một vùng trung tâm mà cả các mối quan hệ



**Hình 3.1.** So sánh giữa hệ thống cơ sở dữ liệu và hệ thống tệp.

giữa các thành phần cơ sở dữ liệu và các đường truy cập tới các cấu tử đó cũng được lưu giữ tại một nơi.

Một hệ thống quản lý dữ liệu có ba thành phần: (1) ngôn ngữ định nghĩa dữ liệu, (2) ngôn ngữ xử lý dữ liệu, và (3) từ điển dữ liệu.

*Ngôn ngữ định nghĩa dữ liệu* là ngôn ngữ chính thức mà lập trình viên sử dụng để chỉ định nội dung và cấu trúc dữ liệu. Nó định nghĩa mỗi phần tử dữ liệu xuất hiện trong cơ sở dữ liệu trước khi phần tử dữ liệu đó được chuyển thành dạng trình ứng dụng đòi hỏi.

*Ngôn ngữ xử lý dữ liệu* là ngôn ngữ chuyên dụng kết hợp chung với các ngôn ngữ lập trình ứng dụng thông thường khác để xử lý dữ liệu trong cơ sở dữ liệu. Ngôn ngữ này chứa các lệnh cho phép người dùng cuối và chuyên viên lập trình lấy ra các dữ liệu thỏa mãn các yêu cầu thông tin và phát triển ứng dụng từ cơ sở dữ liệu. Ngôn ngữ xử lý dữ liệu nổi bật nhất hiện nay là ngôn ngữ truy vấn dữ liệu có cấu trúc SQL.

*Từ điển dữ liệu* là một tệp tin viết tay hoặc tự động chứa định nghĩa của các phần tử dữ liệu và đặc điểm dữ liệu như cách sử dụng, trình bày vật lý, quyền sở hữu, giấy phép và tính bảo mật. Một phần tử dữ liệu thể hiện một trường dữ liệu.

Hệ thống quản lý cơ sở dữ liệu có những chức năng chính như sau:

- **Thiết lập từ điển cơ sở dữ liệu:** Lưu giữ định nghĩa về các mối quan hệ trong một từ điển dữ liệu. Tất cả các chương trình truy cập tới các công việc cần tới dữ liệu sẽ gọi tới hệ thống cơ sở dữ liệu quản lý này để có các thông tin cần thiết. Bất cứ một thay đổi nào trong các tệp dữ liệu, thì lập tức từ điển dữ liệu cũng được tự động thay đổi theo.
- **Quản lý việc lưu trữ dữ liệu:** Hệ thống quản lý cơ sở dữ liệu còn giúp cho người sử dụng tạo các cấu trúc phức tạp cho các dữ liệu lưu trữ. Nhờ đó, ta có thể lập các chương trình liên kết và các chương trình điều khiển phân tích và tổng hợp dữ liệu trực tiếp trên cấu trúc vật lý của dữ liệu.
- **Chuyển đổi và hiển thị thông tin:** Dữ liệu được lưu lại dưới hai hình thức: dữ liệu ghi theo dạng logic và dữ liệu ghi theo dạng vật lý. Do việc ghi dữ liệu đa dạng như vậy, nên hệ thống quản lý cơ sở dữ liệu thực hiện một nhiệm vụ là chuyển những dữ liệu dưới dạng vật lý thành những dữ liệu có tính logic mà người sử dụng cần tới.
- **Bảo đảm độ an toàn của dữ liệu:** hệ thống quản lý cơ sở dữ liệu tạo một hệ thống bảo mật và thiết lập tính an toàn riêng tư cho các dữ

liệu trong cơ sở dữ liệu đó.

- **Khôi phục lại dữ liệu:** hệ thống quản lý cơ sở dữ liệu cung cấp chức năng khôi phục lại những dữ liệu đã mất. Việc này sẽ giúp cho các cơ sở dữ liệu được an toàn và toàn vẹn.
- **Tính toàn vẹn dữ liệu:** Khả năng này cho phép tăng cường tính toàn vẹn và thích hợp của cơ sở dữ liệu.
- **Giao diện trên cơ sở viễn thông của các cơ sở dữ liệu:** hệ thống quản lý cơ sở dữ liệu cho phép người dùng truy cập tới các dữ liệu ở bất cứ đâu trên mạng dữ liệu.

## 2. Mô hình cơ sở dữ liệu

Mô hình cơ sở dữ liệu là một tập hợp các cấu trúc logic được sử dụng để diễn tả cấu trúc dữ liệu và các mối quan hệ dữ liệu được tìm thấy trong một cơ sở dữ liệu. Một cách cơ bản, ta có thể chia các mô hình cơ sở dữ liệu thành hai nhóm: các mô hình khái niệm và các mô hình thực hiện.

### 2.1. Mô hình khái niệm

Mô hình khái niệm tập trung vào bản chất logic của việc biểu diễn dữ liệu. Do đó, mô hình khái niệm liên quan tới vấn đề cái gì được biểu diễn trong cơ sở dữ liệu hơn là làm thế nào để biểu diễn nó. Mô hình khái niệm bao gồm ba dạng quan hệ mô tả sự liên hệ giữa các dữ liệu. Đó là dạng quan hệ một - nhiều, nhiều - nhiều, và quan hệ một - một.

#### 2.1.1. Quan hệ một - một

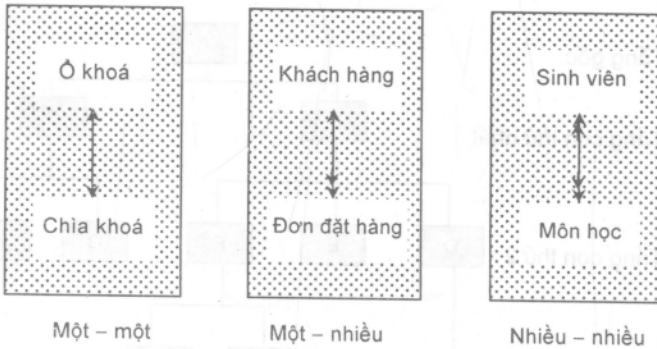
Quan hệ một - một là mối quan hệ duy nhất tồn tại giữa hai thực thể. Từ thực thể này chỉ có thể có duy nhất một đường dẫn tới thực thể kia và ngược lại (xem hình 3.2). Chẳng hạn như một ổ khoá chỉ có thể mở bằng một chìa duy nhất và chìa khoá đó cũng chỉ có thể mở được ổ khoá đó mà không mở được ổ khoá nào khác.

#### 2.1.2. Quan hệ một - nhiều

Đây là mối quan hệ mà từ một gốc có thể chỉ tới nhiều điểm mới, nhưng mỗi điểm chỉ có một gốc duy nhất (xem hình 3.2). Những mối quan hệ dạng này thường có rất nhiều trong các tổ chức doanh nghiệp. Một khách hàng có rất nhiều lần mua hàng với doanh nghiệp, vì vậy, trong hồ sơ lưu trữ tồn tại rất nhiều hoá đơn thanh toán của khách hàng này. Tuy nhiên, ngược lại, mỗi hoá đơn lại chỉ có liên quan tới một khách hàng duy nhất của doanh nghiệp.

### 2.1.3. Quan hệ nhiều - nhiều

Dạng quan hệ này là dạng mà cả gốc và ngọn đều có quan hệ đa phương. Từ một gốc, có thể có nhiều ngọn khác nhau, và ngược lại, từ một ngọn, có thể có nhiều điểm gốc dẫn tới nó. Trong một trường học, một sinh viên có thể học rất nhiều môn học khác nhau. Nhưng ngược lại, mỗi một môn học lại có rất nhiều sinh viên theo học nó (xem hình 3.2).



Hình 3.2. Mô phỏng các dạng quan hệ của các mô hình khái niệm.

## 2.2. Mô hình thực hiện

Khác với mô hình khái niệm, các mô hình thực hiện thì quan tâm tới vấn đề làm thế nào để biểu diễn dữ liệu trong một cơ sở dữ liệu. Mô hình thực hiện bao gồm ba loại: (1) mô hình dữ liệu thứ bậc, (2) mô hình cơ sở dữ liệu dạng mạng lưới, và (3) mô hình cơ sở dữ liệu dạng quan hệ.

### 2.2.1. Mô hình cơ sở dữ liệu thứ bậc

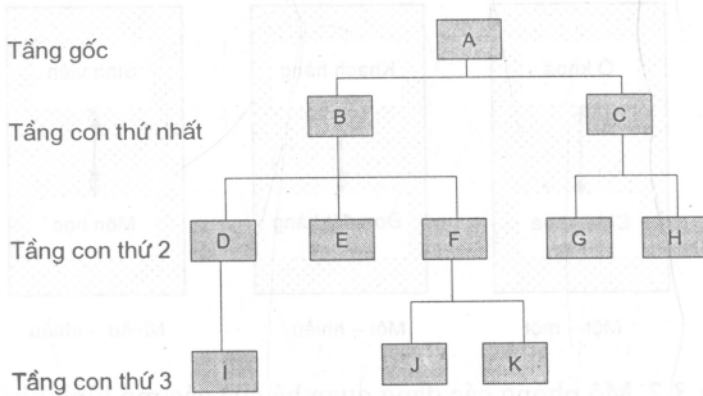
Mô hình cơ sở dữ liệu thứ bậc ra đời vào năm 1969 nhằm giải quyết những khó khăn do sự trùng lặp dữ liệu của hệ thống tệp xảy ra trong quá trình xử lý thông tin trong dự án Apollo của công ty North American Rockwell. Mô hình cơ sở dữ liệu thứ bậc được coi là mô hình chính đầu tiên có tính thương mại dành cho một cơ sở dữ liệu lớn. Những khái niệm cơ sở của nó đã tạo lập nên cơ sở cho sự phát triển cơ sở dữ liệu có thứ tự. Những hạn chế tồn tại trong mô hình cơ sở dữ liệu này dẫn tới hàng loạt các nghiên cứu khác nhau về cách thiết kế cơ sở dữ liệu.

#### Cấu trúc cơ bản

Cơ sở dữ liệu được xây dựng theo dạng thứ bậc có thể hình dung như một cây từ trên xuống dưới với các nút là các dạng báo cáo khác nhau của doanh nghiệp (xem hình 3.3). Trong một dạng thứ bậc như vậy, nút đầu tiên

là nút mẹ. Các nút ở tầng trên là các nút mẹ sinh ra các nút ở tầng dưới. Toàn bộ cây dữ liệu không có bất cứ một sự trùng lặp nào như đối với hệ thống tệp. Thay vào đó, để tìm tới một nút ở dưới nào đó, cây quan hệ sẽ thiết lập một đường dẫn tới vị trí cần thiết đó. Những mối quan hệ quan trọng trong dạng cấu trúc này là:

- Mỗi nút mẹ có thể có nhiều nút con;
- Mỗi nút con chỉ có một nút mẹ và chỉ duy nhất một mà thôi.



**Hình 3.3.** Các phần tử của một cấu trúc thứ bậc.

Mối liên hệ dạng này là mối liên hệ theo kiểu một - nhiều, và thường hay gặp trong các tổ chức doanh nghiệp, như trong một doanh nghiệp có rất nhiều phòng ban, mỗi phòng ban, lại chỉ phụ thuộc vào duy nhất một công ty mà thôi.

### Ưu điểm

Các mô hình dạng cấu trúc thứ bậc như thế này thường có một số ưu điểm cơ bản như sau:

- Do tất cả các dữ liệu đều được giữ trong một cơ sở dữ liệu chung, nên việc phân chia dữ liệu do hệ thống quản lý thông tin điều hành thường phải đòi hỏi khá thực tế và đảm bảo được độ an toàn về dữ liệu.
- Hệ thống quản lý cơ sở dữ liệu tạo ra một môi trường trong đó đảm bảo tính độc lập của các dữ liệu, do đó, làm tăng tính hiệu quả của các chương trình xử lý nó.
- Tạo ra mối liên hệ chặt chẽ giữa các nút mẹ và nút con và nhờ đó đảm bảo tính toàn vẹn của dữ liệu từ trên xuống dưới.

- Mô hình cơ sở dữ liệu thứ bậc rất phù hợp với cơ sở dữ liệu chứa một số lượng lớn dữ liệu có quan hệ một – nhiều và khi người sử dụng cần một số lượng lớn các giao dịch sử dụng những mối quan hệ cố định trong một thời gian dài. Phần lớn các ngân hàng đều sử dụng mô hình quan hệ thứ bậc dạng này.
- Cơ sở dữ liệu được thiết lập từ đầu là rất lớn, và do đó, người lập trình có khả năng thiết lập các chương trình một cách có hiệu quả hơn.
- Các ứng dụng của doanh nghiệp có thể áp dụng được rất nhiều trong môi trường chính của cơ sở dữ liệu này.

### **Hạn chế**

Mặc dù có rất nhiều các ứng dụng có thể áp dụng đối với dạng cơ sở dữ liệu lớn dạng này, nhưng không phải bao giờ người ta cũng sử dụng mô hình này trong việc quản lý cơ sở dữ liệu của các doanh nghiệp do nó còn có một số các nhược điểm sau:

- Mặc dù mô hình dạng thứ bậc này giúp cho các nhà lập trình thoát khỏi các vấn đề phụ thuộc về dữ liệu, nhưng hệ thống quản lý dữ liệu vẫn đòi hỏi phải có kiến thức về mức độ vật lý trên khía cạnh lưu trữ dữ liệu. Bất cứ sự thay đổi trong cấu trúc cơ sở dữ liệu, như thiết lập lại các môđun, đều đòi hỏi thay đổi tất cả các chương trình ứng dụng. Do đó, thực hiện việc thiết kế cơ sở dữ liệu có thể trở nên hết sức phức tạp.
- Nhiều mối quan hệ giữa các dữ liệu thực tế không có mối quan hệ một - nhiều theo tiêu chuẩn mà mô hình thứ bậc cung cấp. Những mối quan hệ dạng nhiều - nhiều thường rất khó sử dụng mô hình thứ bậc này.
- Cơ sở dữ liệu thứ bậc thường phức tạp, khó quản lý, và ít linh hoạt. Khi một khâu nối nào đó bị xoá đi, rất khó có thể xoá những dữ liệu trực tiếp dưới quyền quản lý của nó một cách tự động.
- Các chương trình ứng dụng có vẻ quá bao quát. Các nhà quản lý hay lập trình buộc phải biết rõ về các mã điều khiển để lấy được dữ liệu và phải rất quen thuộc với cấu trúc dữ liệu.

### **2.2.2. Mô hình cơ sở dữ liệu mạng**

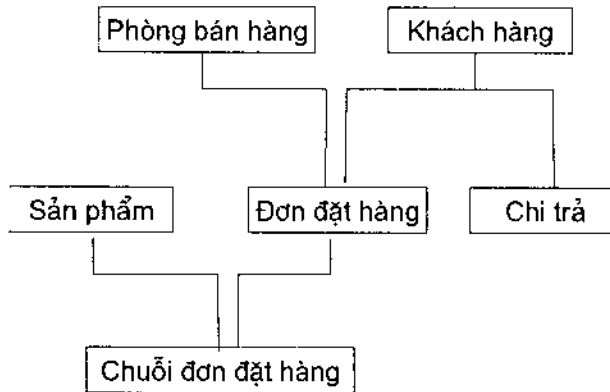
Mô hình cơ sở dữ liệu mạng thường giống như mô hình cơ sở dữ liệu thứ bậc. Tuy vậy, khác biệt lớn nhất để phân biệt hai loại mô hình cơ sở dữ liệu này là trong mô hình cơ sở dữ liệu mạng các báo cáo có thể được thiết

lập từ nhiều nguồn nghĩa là có nhiều nút mẹ tới một nút con.

Mô hình cơ sở dữ liệu mạng được thiết lập một phần để biểu thị những dữ liệu có mối quan hệ phức tạp hơn mô hình cơ sở dữ liệu thứ bậc có thể làm được. Mặt khác mô hình cơ sở dữ liệu mạng còn giúp cho việc thiết lập các chuẩn mực cơ sở dữ liệu, việc này giúp cho việc lập trình và tạo các chương trình quản lý cơ sở dữ liệu rất nhiều.

### Cấu trúc cơ bản

Trong cơ sở dữ liệu dạng mạng, mỗi quan hệ được gọi là một tập. Mỗi tập chứa ít nhất hai dạng: một báo cáo chủ giống như nút mẹ trong cơ sở dữ liệu thứ bậc, và một báo cáo thành phần như các nút con trong mô hình cơ sở dữ liệu thứ bậc. Sự khác biệt của cơ sở dữ liệu mạng với cơ sở dữ liệu thứ bậc là một báo cáo thành phần có thể xuất hiện trong nhiều tập khác nhau, nghĩa là nó có thể có nhiều nút mẹ khác nhau. Hình 3.4 giới thiệu một ví dụ về dạng mô hình CSDL mạng.



**Hình 3.4.** Mô hình cơ sở dữ liệu mạng.

Để xây dựng một cơ sở dữ liệu mạng, các mối quan hệ buộc phải được thiết lập thành các tập khác nhau.

### Ưu điểm

Mô hình cơ sở dữ liệu mạng có một số các ưu điểm giúp nó khắc phục được những nhược điểm của cơ sở dữ liệu thứ bậc như sau:

- Trong mô hình cơ sở dữ liệu mạng, các mối quan hệ nhiều - nhiều có thể dễ dàng biểu diễn hơn trong mô hình cơ sở dữ liệu thứ bậc.
- Việc truy cập dữ liệu và độ linh hoạt của cơ sở dữ liệu cao hơn nhiều so với hệ cơ sở dữ liệu thứ bậc. Một ứng dụng bất kỳ có thể truy cập

tới các báo cáo chủ cũng như các báo cáo thành phần trong bất kỳ một tập nào. Do đó, nếu một báo cáo thành phần có nhiều báo cáo chủ thì chương trình ứng dụng có thể nhanh chóng chuyển từ một báo cáo chủ này sang một báo cáo chủ khác khá nhanh chóng.

- Mô hình cơ sở dữ liệu mạng cho phép nâng cao tính bảo toàn dữ liệu do người sử dụng buộc phải khai báo cáo chủ trước rồi mới tới báo cáo thành phần.
- Mô hình cơ sở dữ liệu mạng đảm bảo tính độc lập của các dữ liệu một cách cao nhất. Chính vì thế, một sự thay đổi tính chất của một dữ liệu này, không ảnh hưởng tới các dữ liệu khác và do đó, các chương trình ứng dụng cũng không phải thay đổi theo.

### **Nhược điểm**

- Mô hình cơ sở dữ liệu mạng khó thiết kế và sử dụng. Người sử dụng buộc phải nắm vững và quen thuộc với cấu trúc dữ liệu mới tận dụng được hết những ưu thế của hệ thống.
- Khó có thể thay đổi trong cơ sở dữ liệu, có một số thay đổi gần như không thể thực hiện được. Mặc dù hệ cơ sở dữ liệu mạng tạo ra sự độc lập về dữ liệu nhưng nó lại không thể tạo ra sự độc lập về cấu trúc. Mỗi khi thay đổi cấu trúc của cơ sở dữ liệu thì các cấu trúc con cũng phải thay đổi theo trước khi thực hiện bất cứ một chương trình truy cập dữ liệu nào.
- Mô hình cơ sở dữ liệu mạng khiến cho việc lập trình trở nên phức tạp hơn. Các nhà lập trình phải nắm rất vững về cấu trúc cơ sở dữ liệu mới có thể làm tốt công việc của họ được.
- Cũng giống như cơ sở dữ liệu dạng thứ bậc, cơ sở dữ liệu mạng tạo ra một môi trường truy cập dữ liệu theo một dòng thông nhất nghĩa là nếu đang từ một biểu ghi A, muốn chuyển sang biểu ghi E, trong dãy A, B, C, D, E, thì người sử dụng buộc phải đi từ biểu ghi A, qua biểu ghi B, C, D, rồi mới tới E.

Nhìn chung, cơ sở dữ liệu mạng không tạo ra được một hệ thống tiện ích cho người sử dụng mà nó có hướng thích hợp với những người lập trình và các nhà quản lý.

### **2.2.3. Mô hình cơ sở dữ liệu quan hệ**

Đòi hỏi phải hiểu biết về các yếu tố vật lý của cấu trúc dữ liệu đã khiến cho những cơ sở dữ liệu được thiết kế tốt trở nên khó khăn hơn. Trong thực



dù cơ sở dữ liệu mạng có rất nhiều điểm mạnh, nhưng cấu trúc phức tạp của nó làm cho rất ít người sử dụng có thể tận dụng được những thế mạnh đó của nó. Khi thông tin cần thiết tăng lên, và do đó, yêu cầu dễ phổ biến trong xã hội của các cơ sở dữ liệu được đòi hỏi cao hơn thì việc thiết kế cơ sở dữ liệu, việc quản lý và sử dụng nó trở nên quá nặng nề và công kềnh.

Mô hình cơ sở dữ liệu quan hệ do E.F.Codd phát minh ra vào năm 1970, đã khắc phục được những nhược điểm nêu trên của cơ sở dữ liệu mạng. Với việc sử dụng một thuật toán tạo ra cơ sở dữ liệu truyền một cách tự động, thay cho cơ sở dữ liệu truyền chuẩn hoá trước đây, mô hình cơ sở dữ liệu quan hệ đã tạo ra một bước ngoặt cho cuộc cách mạng về cơ sở dữ liệu.

Công trình của Codd chưa được sử dụng một cách mạnh mẽ và cũng không có tiếng tăm mấy vào những năm mà nó mới ra đời do công nghệ thông tin còn phát triển chưa đạt mức yêu cầu và chi phí cho máy tính để thực hiện ý tưởng này còn quá cao. Chỉ tới ngày nay, khi mà cuộc cách mạng trong công nghệ cao đã tạo ra những chiếc máy vi tính với khả năng sử dụng khá lớn và tốc độ nhanh tới mức cho phép cũng như đơn giá của những chiếc máy này ngày càng giảm xuống, do đó, tạo cơ hội cho có nhiều người sử dụng chúng thì các cơ sở dữ liệu liên hệ mới có đất để phát triển và hàng loạt các cơ sở dữ liệu quan hệ dạng này đã ra đời và phát huy được thế mạnh tiềm tàng của nó.

### **Cấu trúc cơ bản**

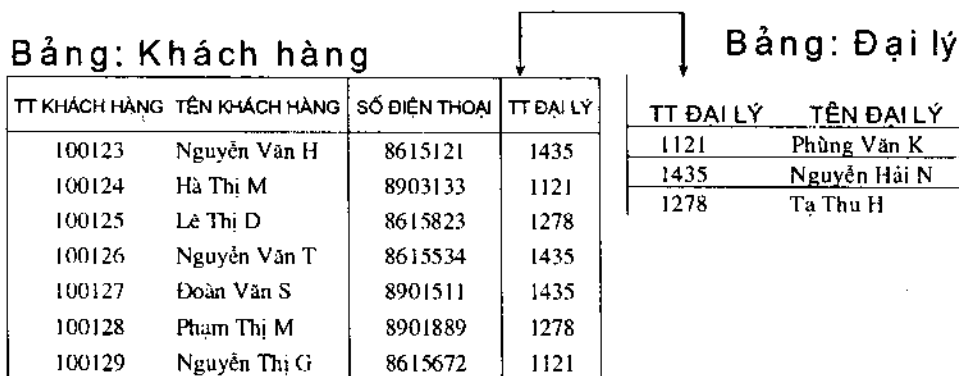
Mô hình cơ sở dữ liệu quan hệ được thực hiện thông qua một hệ thống cơ sở dữ liệu quan hệ. Hệ thống này cũng có những chức năng tương tự như hệ cơ sở dữ liệu mạng và hệ cơ sở dữ liệu thứ bậc và thêm vào đó, nó còn có những chức năng chủ khác cho phép mô hình cơ sở dữ liệu quan hệ trở nên dễ hiểu hơn và dễ thực hiện hơn.

Ưu điểm quan trọng nhất của cơ sở dữ liệu quan hệ là khả năng của nó trong việc thực hiện các mô hình quan hệ giống như trong một môi trường sống thực thụ. Điều này giúp cho người sử dụng và người thiết kế thực hiện được công việc dễ dàng hơn. Cơ sở dữ liệu quan hệ được người sử dụng nhận thức như một tập hợp các bảng biểu lưu trữ dữ liệu.

Mỗi bảng là một ma trận gồm một chuỗi các hàng hoặc cột giao nhau. Các bảng biểu còn được gọi là các mối quan hệ liên kết với nhau bởi một tính chất chung nào đó. Như trong ví dụ biểu diễn trong hình 3.5, bảng Khách hàng và bảng Đại lý bán hàng liên hệ với nhau theo một biến số chung có tên là TT Đại lý. Mặc dù các dữ liệu hoàn toàn độc lập ở mỗi

bảng, ta vẫn có thể dễ dàng liên kết dữ liệu giữa các bảng với nhau. Mô hình cơ sở dữ liệu quan hệ loại bỏ được hầu hết sự trùng lặp dữ liệu giữa các bảng khác nhau mà hệ thống các tệp dữ liệu thường mắc phải.

### Số TT đại lý chung



**Hình 3.5.** Mối liên kết giữa các bảng có quan hệ.

Mô hình cơ sở dữ liệu quan hệ lưu trữ một tập hợp các thực thể có liên quan tới nhau. Trong mô hình cơ sở dữ liệu dạng này, các bảng cơ sở dữ liệu tương tự như một tệp dữ liệu. Nhưng các dữ liệu được lưu trữ là hoàn toàn độc lập về cấu trúc cũng như về dữ liệu. Nó hoàn toàn được tổ chức theo dạng cấu trúc logic. Dù là dữ liệu được lưu trữ vật lý như thế nào, nó cũng hoàn toàn không ảnh hưởng tới công việc của nhà thiết kế cũng như của người sử dụng.

#### Ưu điểm

Cơ sở dữ liệu quan hệ là một kho dữ liệu riêng biệt. Khác với hệ cơ sở dữ liệu thứ bậc và hệ cơ sở dữ liệu mạng, trong hệ cơ sở dữ liệu quan hệ, người sử dụng và người thiết kế hoàn toàn không phải quan tâm tới cấu trúc cơ sở dữ liệu. Do đó, tính độc lập về cấu trúc dữ liệu là ưu điểm nổi bật nhất của cơ sở dữ liệu quan hệ. Hơn thế nữa, do nó giải phóng cho ta về mặt cần phải quan tâm tới khía cạnh vật lý của cơ sở dữ liệu, nên ta có thêm thời gian quan tâm tới khía cạnh logic của cơ sở dữ liệu.

Hệ cơ sở dữ liệu quan hệ có khả năng linh hoạt rất cao. Do đó, nó đòi hỏi ít việc lập trình để truy cập dữ liệu hơn các loại cơ sở dữ liệu khác. Một trong những ưu thế mạnh của cơ sở dữ liệu dạng này là nó dễ tạo ra một giao diện thích hợp với người sử dụng hơn các cơ sở dữ liệu khác.

## Nhược điểm

Hệ cơ sở dữ liệu quan hệ gần như che hết toàn bộ cấu trúc vật lý của cơ sở dữ liệu, do đó, nó đòi hỏi phải có hệ điều hành và phần cứng hoàn hảo đối với người sử dụng. Cũng do đặc tính đó, nó đòi hỏi một hệ thống máy tính mạnh để hỗ trợ cho việc thực hiện những nhiệm vụ truy cập dữ liệu và thông tin. Đồng thời đó cũng là nguyên nhân khiến cho cơ sở dữ liệu dạng này, hoạt động có phần chậm hơn so với các dạng cơ sở dữ liệu khác. Nhưng với điều kiện hiện nay, khi hệ thống máy tính ngày càng được phát triển tốt hơn với sự hỗ trợ của công nghệ cao, thì sự chậm trễ này cũng đã giảm bớt được đáng kể.

## 3. Thiết kế cơ sở dữ liệu

Thiết kế là một bước khá quan trọng trong quá trình xây dựng một cơ sở dữ liệu. Trong hệ thống quản lý dữ liệu, các công việc chủ yếu được thực hiện không chỉ là thu thập dữ liệu, lưu trữ dữ liệu, và khôi phục các dữ liệu, mà quan trọng hơn cả là chuyển đổi dữ liệu đó thành thông tin. Các thông tin thu được từ dữ liệu càng nhanh chóng, và gần gũi với tình trạng thực tiễn bao nhiêu thì việc tạo quyết định càng chính xác và càng có nhiều ý nghĩa.

HTTT không phải đột nhiên xuất hiện mà nó là sản phẩm của một quá trình phát triển từng bước. Để xác định nhu cầu về dữ liệu và thiết kế nó, trước hết ta cần phải phân tích các dữ liệu mà doanh nghiệp thường xuyên liên hệ và làm việc với nó cùng với những thông tin mà doanh nghiệp thu thập được từ những dữ liệu đó. Do bản chất của thông tin là luôn thay đổi và luôn biến động nên HTTT xây dựng được ngày hôm nay, có thể không còn thích hợp trong ngày mai. Trong tương lai, có thể chúng ta sẽ lưu ý tới những thông tin hoàn toàn khác những gì mà ngày nay chúng ta coi là quan trọng nhất. Vì vậy, nói tới thiết kế HTTT, ta cần hình dung đó như là một quá trình không ngừng nghỉ, bước sau coi bước trước là tiền đề để xây dựng và tạo ra một vòng sống của hệ thống.

Do tính chất đó của HTTT, nên hệ cơ sở dữ liệu phục vụ nó cũng cần phải được đánh giá và xây dựng trên cơ sở một quá trình liên tục và cần lưu ý tới chu kỳ sống của một cơ sở dữ liệu, nghĩa là lưu ý tới khi nó không còn đáp ứng được cho nhu cầu đòi hỏi những thông tin cần thiết nữa. Như vậy, vấn đề quan trọng đầu tiên là hệ cơ sở dữ liệu sẽ cung cấp những thông tin gì cho người sử dụng và nó có phù hợp với yêu cầu của người sử dụng không.

### **3.1. Chuyển đổi dữ liệu thành thông tin**

Dữ liệu là tất cả những gì cơ bản nhất xuất hiện một cách tự nhiên trong cuộc sống của chúng ta. Tất cả những sự kiện tự nhiên đó được thu thập và lưu trữ trong một hệ cơ sở dữ liệu. Chẳng hạn như, khi một ngân hàng muốn thu thập thông tin về khách hàng, nó sẽ cần những dữ liệu như tuổi của khách, giới tính, những khoản nợ lâu dài sẵn có của khách, những khoản nợ ngắn hạn mà khách chưa trả cho ngân hàng hoặc cho các chủ nợ khác, mức độ học vấn của khách, quá trình hoạt động của khách, v.v... Những dữ liệu thu thập được càng rõ ràng bao nhiêu thì việc phân tích nó càng dễ dàng bấy nhiêu.

Một vấn đề quan trọng là các dữ liệu thường rất hiếm khi trực tiếp có ích cho những quyết định của người sử dụng chúng. Nói cách khác, người tạo quyết định thường cần dùng thông tin là kết quả của quá trình xử lý và phân tích dữ liệu để đưa chúng về dạng có ý nghĩa hơn.

Quá trình chuyển đổi dữ liệu thành thông tin có thể dựa trên các bảng tổng hợp dữ liệu, hoặc dựa trên các báo cáo chi tiết, hoặc dựa trên các số liệu thống kê phức tạp từ các dữ liệu sẵn có. Bất cứ sử dụng phương pháp nào thì việc tạo quyết định vẫn chính là dựa trên một vài dạng chuyển đổi dữ liệu. Nếu không có dữ liệu, ta sẽ không thể chuyển đổi được thành thông tin cần thiết.

### **3.2. Chu kỳ phát triển cơ sở dữ liệu**

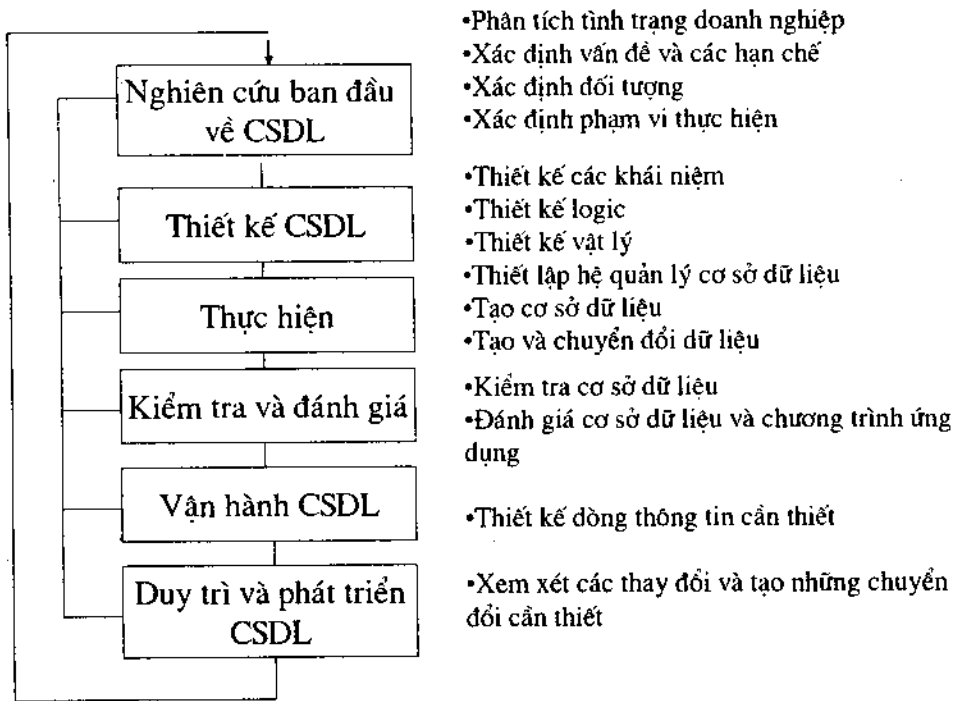
Trong một hệ thống tin lớn thì cơ sở dữ liệu thường được xây dựng thông qua một quá trình liên tục có lặp lại mà người ta thường gọi là vòng đời của cơ sở dữ liệu. Mỗi một quá trình như vậy thường được cấu tạo từ sáu bước cơ bản như minh họa trong hình 3.6.

#### **3.2.1. Nghiên cứu ban đầu về cơ sở dữ liệu**

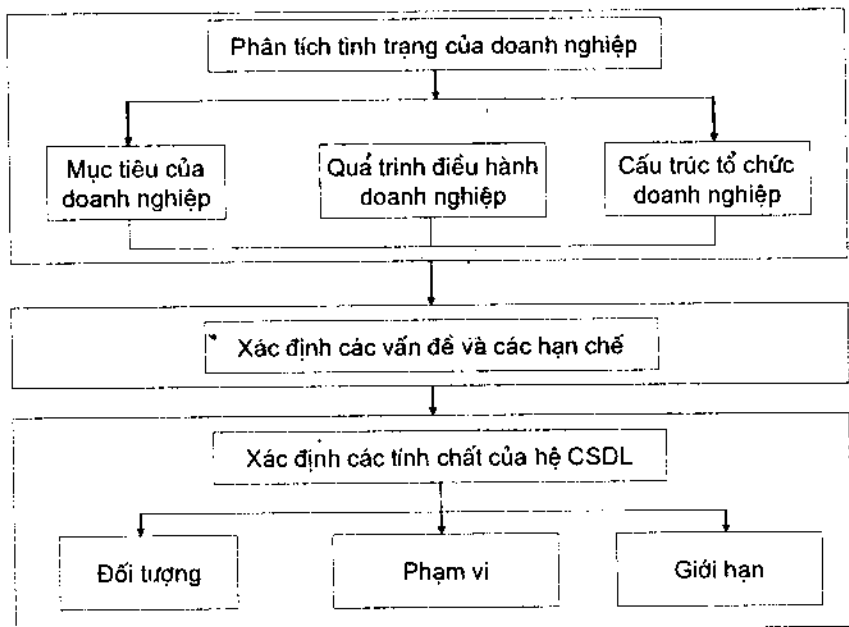
Mục đích cơ bản của quá trình nghiên cứu ban đầu về cơ sở dữ liệu biểu diễn trên hình 3.7 gồm những vấn đề sau:

- Phân tích tình trạng của doanh nghiệp;
- Xác định vấn đề tồn tại và các hạn chế;
- Xác định các mục tiêu cơ bản;
- Xác định phạm vi thực hiện và những ranh giới thực hiện.

Việc xác định chính xác những thông tin đầu tiên này sẽ cho phép chúng ta thiết lập một cơ sở dữ liệu hợp lý và có hiệu quả trong công việc.



**Hình 3.6.** Chu trình thiết kế cơ sở dữ liệu.



**Hình 3.7.** Các hoạt động của bước phân tích ban đầu cơ sở dữ liệu.

## **a) Phân tích tình trạng của doanh nghiệp**

Tình trạng của một doanh nghiệp là tất cả những tình trạng và điều kiện chung ảnh hưởng tới quá trình hoạt động của doanh nghiệp, tới cấu trúc tổ chức của nó, và tới nhiệm vụ chính của nó trong thời gian ngắn cũng như trong một thời gian lâu dài. Để nắm vững được tình trạng của doanh nghiệp, người thiết kế buộc phải nắm vững được những thông tin cơ bản sau:

- Môi trường hoạt động của doanh nghiệp có những đặc điểm gì? Doanh nghiệp có nhiệm vụ chính gì trong môi trường hoạt động đó?
- Cấu trúc tổ chức của doanh nghiệp như thế nào? Ai là người điều hành các hoạt động của doanh nghiệp, của từng bộ phận trong doanh nghiệp? Dòng thông tin được chuyển dịch trong tổ chức theo dạng như thế nào? Từng bộ phận phải tổng hợp những dạng báo cáo nào để phản ánh tình hình hoạt động của bộ phận và của doanh nghiệp?

## **b) Xác định vấn đề và các hạn chế**

Sau khi đã có những hình dung ban đầu, công việc của người thiết kế cơ sở dữ liệu là nhanh chóng tìm ra những vấn đề còn tồn tại trong doanh nghiệp. Thông thường đối với các doanh nghiệp đã có một thời gian hoạt động lâu dài, sẽ tồn tại nhiều nguồn thông tin khác nhau. Đó có thể là những nguồn thông tin chính thức, hoặc không chính thức. Những thông tin này có thể được lưu trữ và thành lập bằng tay, hoặc bằng máy tính. Tìm hiểu được tất cả những nguồn thu thập thông tin sẵn có này trong doanh nghiệp là vấn đề hết sức quan trọng. Đặc biệt là cần tìm hiểu con đường mà các báo cáo di chuyển trong một doanh nghiệp.

Từ tất cả những nguồn đã nắm được đó, quá trình xác định các vấn đề còn tồn tại mới bắt đầu được thực hiện và nó được thực hiện một cách không có bất cứ một tổ chức cấu trúc nào. Một vấn đề cần lưu ý là người quản lý và người trực tiếp thực hiện công việc cụ thể trong một doanh nghiệp có những cách nhìn nhận hoàn toàn khác nhau. Vì vậy, khi tiến hành quá trình tìm hiểu những vấn đề thực sự đang tồn tại trong doanh nghiệp, cần phải thu thập tất cả những ý kiến khác nhau của các đối tượng khác nhau nghĩa là những người nắm những trọng trách khác nhau trong doanh nghiệp.

## **c) Xác định mục tiêu**

Hệ thống cơ sở dữ liệu cần phải được thiết kế sao cho nó có thể giải

quyết những vấn đề cơ bản xác định được trong quá trình hoạt động của doanh nghiệp. Một vấn đề nảy sinh trong một doanh nghiệp có thể sẽ được phát hiện từ nhiều nguồn khác nhau. Chẳng hạn như cả phòng marketing và phòng sản xuất đều phát hiện ra sự dự trữ hàng tồn kho là không hiệu quả trong một doanh nghiệp. Nếu việc thiết kế cơ sở dữ liệu có khả năng tạo một cơ sở dữ liệu một tập các bước có hiệu quả cho việc quản lý, thì tất cả các phòng ban sẽ đều có thể tham gia vào quá trình.

Lưu ý rằng những nghiên cứu đầu tiên làm phát sinh các giải pháp cho các vấn đề. Nhiệm vụ của người thiết kế là phải làm sao cho các đối tượng của hệ cơ sở dữ liệu xuất hiện trước con mắt của người thiết kế cũng phải hoàn toàn phù hợp với con mắt của người sử dụng. Trong bất cứ trường hợp nào, người thiết kế cơ sở dữ liệu cần phải bắt đầu bằng việc trả lời các câu hỏi dạng: đối tượng của hệ cơ sở dữ liệu là ai, liệu hệ giao diện có phù hợp với những hệ thống đang tồn tại hay những hệ thống sẽ ra đời trong tương lai không, liệu hệ cơ sở dữ liệu có chia sẻ được thông tin và dữ liệu với những cơ sở dữ liệu khác không và giữa những người sử dụng khác nhau không.

#### **d) Phạm vi và giới hạn**

Người thiết kế cũng cần phải quan tâm tới những phạm vi và giới hạn hiện đang tồn tại. Phạm vi xác định việc thiết kế cơ sở dữ liệu dựa trên những yêu cầu đòi hỏi của việc điều hành. Người thiết kế phải xác định được rõ ràng, hệ cơ sở dữ liệu sẽ phục vụ những phòng ban nào trong doanh nghiệp, phục vụ những chức năng chủ yếu nào trong những phòng ban đó. Nhờ đó, việc thiết kế cấu trúc của cơ sở dữ liệu, xác định kích cỡ vật lý của cơ sở dữ liệu, dạng và số các trường khác nhau trong cơ sở dữ liệu, cũng như một số các thông số cụ thể khác.

Ngoài ra, những giới hạn bên ngoài của hệ cơ sở dữ liệu cũng cần được xác định cụ thể. Những giới hạn này bao gồm những vấn đề dạng như ngân quỹ phục vụ cho quá trình thiết kế cơ sở dữ liệu, những giới hạn về phần cứng và phần mềm trợ giúp cho quá trình đó. Trong thực tế, giới hạn của các phần cứng và phần mềm hỗ trợ chính nó cũng xác định chu kỳ sống của một hệ cơ sở dữ liệu.

#### **3.2.2. Quá trình thiết kế cơ sở dữ liệu**

Trong quá trình thiết kế cơ sở dữ liệu, điều cơ bản nhất là cần tập trung vào phân tích những tính chất cơ bản của dữ liệu tạo nên cơ sở dữ liệu. Khi thiết kế, ta cần chú ý tới cả hai khía cạnh hay hai cách nhìn nhận khác nhau ảnh hưởng lên cơ sở dữ liệu đó là cách nhìn nhận mang tính kinh tế và cách