



BÀI 10

CƠ SỞ DỮ LIỆU VÀ

HỆ THỐNG THÔNG TIN

Trường Công nghệ thông tin và Truyền thông
2023

Nội dung

1. Dữ liệu, tổ chức dữ liệu
2. Cơ sở dữ liệu
3. Các mô hình cơ sở dữ liệu
4. Sử dụng CSDL
5. Giới thiệu chung về hệ thống thông tin
6. Các hệ thống thông tin của tổ chức hay doanh nghiệp
7. Phân tích và Thiết kế Hệ thống thông tin

1. Dữ liệu và tổ chức dữ liệu

- Dữ liệu (data) bao gồm các yếu tố hay các quan sát về con người, địa điểm, vạn vật và các sự kiện.
- Các kiểu dữ liệu:
 - Số
 - Ký tự
 - Âm thanh
 - Âm nhạc
 - Hình ảnh
 - Video
 - ...
- → đều được mã hóa theo nhị phân

Cách nhìn về dữ liệu

- Cách nhìn về vật lý: Tập trung vào dạng thức thực tế và nơi đặt dữ liệu.
- Cách nhìn về logic: Tập trung vào ý nghĩa, nội dung và ngũ cảnh của dữ liệu.

Tổ chức dữ liệu

- Ký tự (Character): là phần tử dữ liệu cơ bản nhất.
- Trường (Field) : là nhóm các ký tự có quan hệ với nhau để mô tả một thuộc tính của thực thể (entity) (một người, vị trí, sự vật hay một đối tượng).
- Bản ghi (Record): tập hợp các trường có quan hệ với nhau để mô tả các thuộc tính của thực thể.
- Bảng (Table): tập hợp các bản ghi có quan hệ với nhau, được lưu trữ dưới dạng file.
- Cơ sở dữ liệu (Database): Tập hợp các bảng có quan hệ logic với nhau

Trường khóa (Key Field)

- Còn gọi là Khóa chính (Primary Key)
- Định danh duy nhất cho bản ghi
- Ví dụ
 - Mã số công dân
 - Mã số sinh viên
 - Mã số cán bộ
 - Mã số phụ tùng (Part Number)
 - Số seri (Serial number)

Ví dụ

	Họ đệm	Tên	Mã số SV	Ngày sinh	Nơi sinh
Bản ghi 1	Nguyễn	Hoàng Long	20172168	02/06/1999	Hà Nội
Bản ghi 2	Trần	Hoài Nam	20172912	21/10/1999	Hải Dương
Bản ghi 3	Lê	Thu Trang	20173265	30/05/1999	Hải Phòng

- Bảng dữ liệu sinh viên KT22 K62
- Cơ sở dữ liệu sinh viên K62-ĐHBKHN

Xử lý theo lô và xử lý thời gian thực

- Xử lý theo lô (Batch processing): Dữ liệu được thu thập và tập hợp sau một khoảng thời gian xác định, khi đó mới xử lý tất cả cùng một lần.
 - VD: Xử lý điểm sinh viên sau mỗi học kỳ
- Xử lý thời gian thực (Real-time processing): dữ liệu được xử lý cùng lúc với các giao tác dữ liệu.
 - VD: xử lý dữ liệu ngân hàng sau mỗi giao dịch

2. Cơ sở dữ liệu (CSDL)

- CSDL là tập các bảng dữ liệu có quan hệ logic được tích hợp với nhau
- CSDL cần có các đặc điểm sau:
 - Chia sẻ thông tin: thông tin từ bộ phận này có thể sẵn sàng được chia sẻ với bộ phận khác.
 - An toàn: người dùng chỉ có thể truy cập được thông tin cần thiết cho họ.
 - Ít dư thừa dữ liệu (less data redundancy): Nếu không có CSDL chung, khi đó các bộ phận riêng lẻ cần phải tạo ra và duy trì dữ liệu của mình và dẫn đến dư thừa dữ liệu.
 - Tính toàn vẹn dữ liệu (data integrity): khi có nhiều nguồn dữ liệu, thông tin trong các nguồn có thể khác nhau.

Hệ quản trị cơ sở dữ liệu

- Hệ quản trị CSDL (DBMS - Database Management System): là phần mềm để tạo, biến đổi và truy cập cơ sở dữ liệu.
- Ví dụ:
 - Microsoft Access
 - SQL Server
 - MySQL
 - Oracle

Các thành phần của DBMS

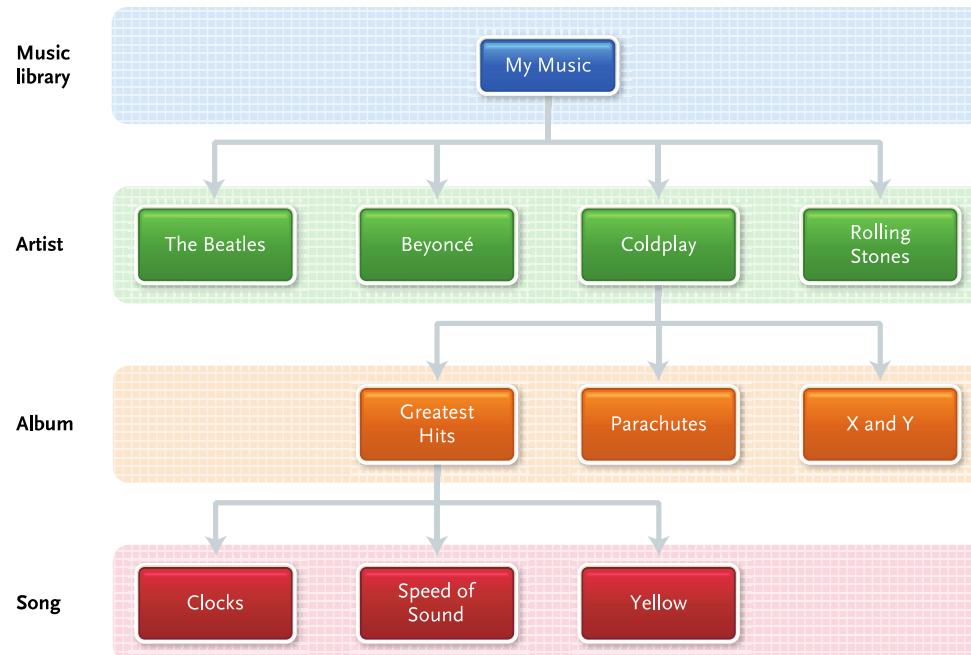
- DBMS engine: cung cấp cầu nối giữa dữ liệu logic và dữ liệu vật lý.
- Hệ thống con định nghĩa dữ liệu: định nghĩa cấu trúc logic của CSDL bằng cách sử dụng từ điển dữ liệu hay lược đồ dữ liệu.
- Hệ thống con thao tác dữ liệu: cung cấp các công cụ để duy trì và phân tích dữ liệu (thêm, xóa, sửa đổi, truy vấn).
- Hệ thống con tạo ứng dụng: cung cấp các công cụ để tạo các dạng vào dữ liệu và các ngôn ngữ lập trình chuyên biệt.
- Hệ thống con quản trị dữ liệu: để quản lý CSDL, nó bao gồm: duy trì an toàn, cung cấp sự hỗ trợ khôi phục thảm họa và giám sát hiệu năng tổng thể của các thao tác CSDL.

3. Các mô hình cơ sở dữ liệu

- Cơ sở dữ liệu phân cấp (Hierarchical Database)
- Cơ sở dữ liệu mạng (Network Database)
- Cơ sở dữ liệu quan hệ (Relational Database)
- Cơ sở dữ liệu đa chiều (Multidimensional Database)
- Cơ sở dữ liệu hướng đối tượng (Object-Oriented Database)

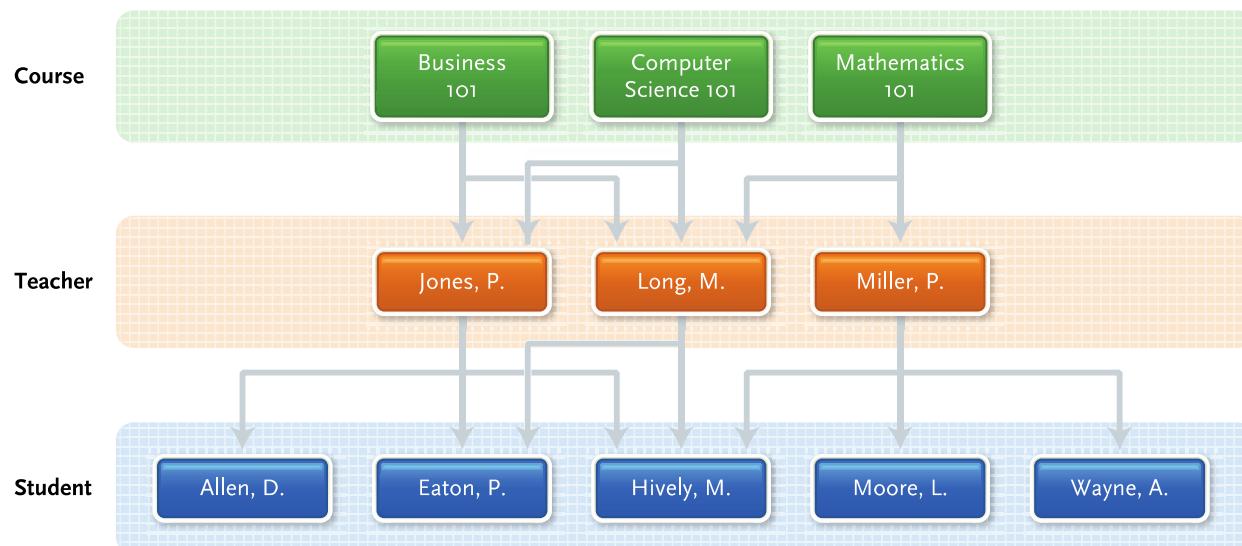
Cơ sở dữ liệu phân cấp

- Các trường hoặc các bản ghi được cấu trúc theo node
- Mỗi điểm vào có thể có một node cha với một vài node con theo quan hệ một nhiều
- Mỗi node con chỉ có một node cha
- Không có quan hệ giữa các node con



Cơ sở dữ liệu mạng

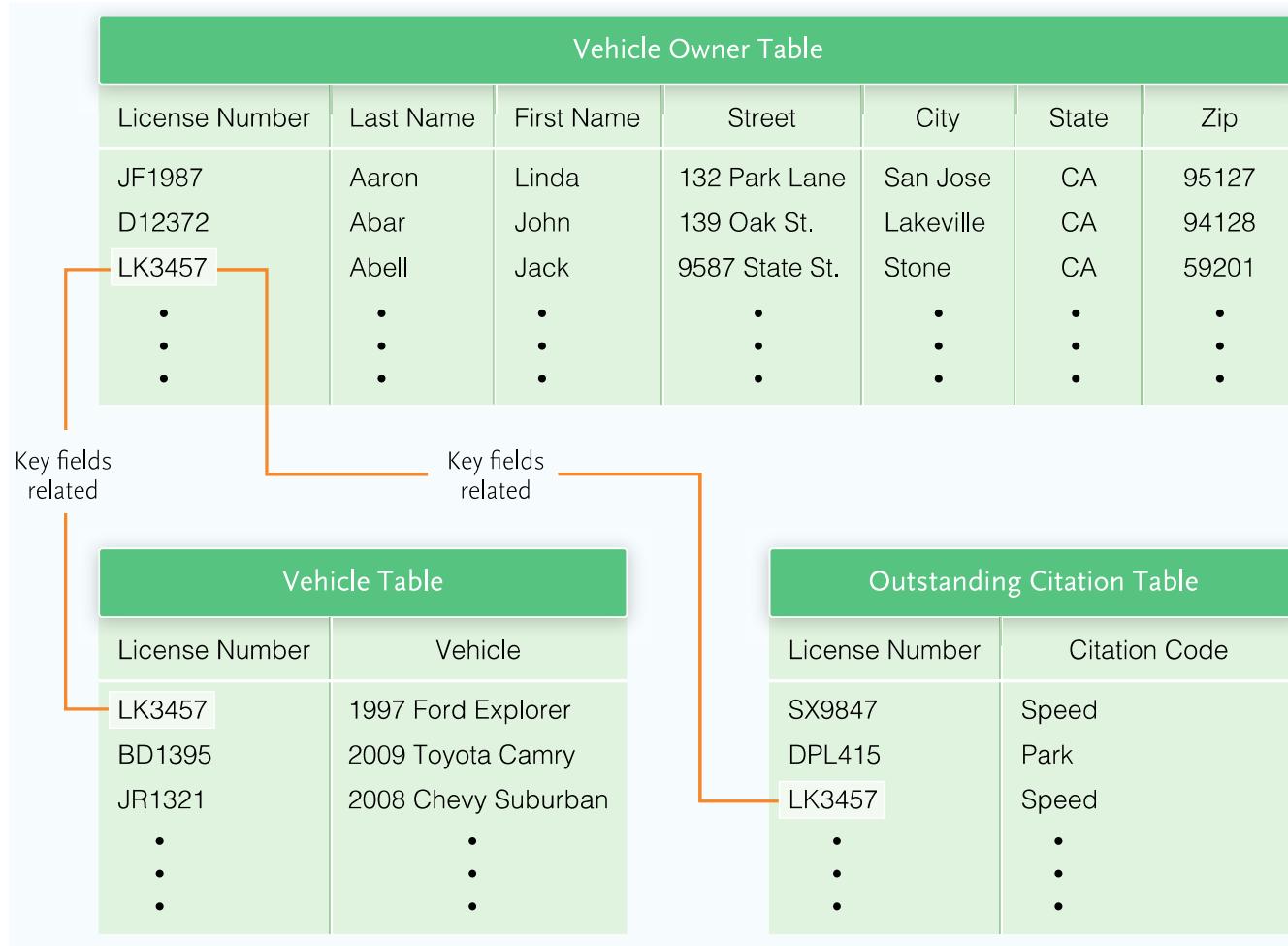
- Tổ chức phân cấp cho các node
- Mỗi node con có thể có nhiều hơn một node cha
- Có thể có quan hệ nhiều-nhiều
- Để truy cập vào một node có thể có các đường khác nhau.



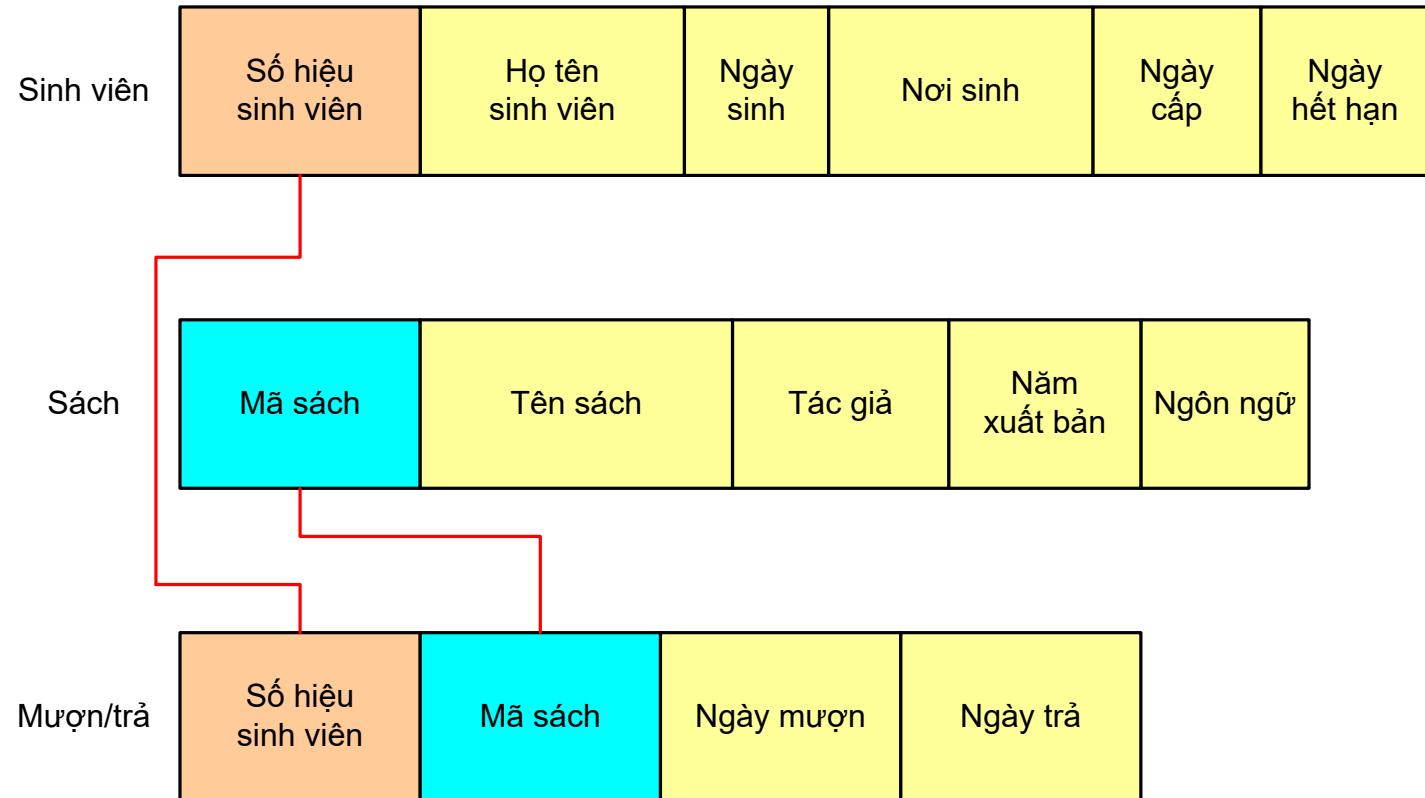
Cơ sở dữ liệu quan hệ

- Các phần tử dữ liệu được lưu trữ trong các bảng khác nhau
- Các bảng bao gồm các hàng và các cột
- Các bảng cần có mục dữ liệu chung (trường khóa) cho phép thông tin được lưu trong bảng này có thể được liên kết với thông tin trong bảng khác
- Tổ chức đơn giản và dễ hiểu
- Các thực thể dễ dàng thêm, xóa và thay đổi
- Được sử dụng phổ biến nhất

Ví dụ cơ sở dữ liệu quan hệ



Ví dụ cơ sở dữ liệu quan hệ



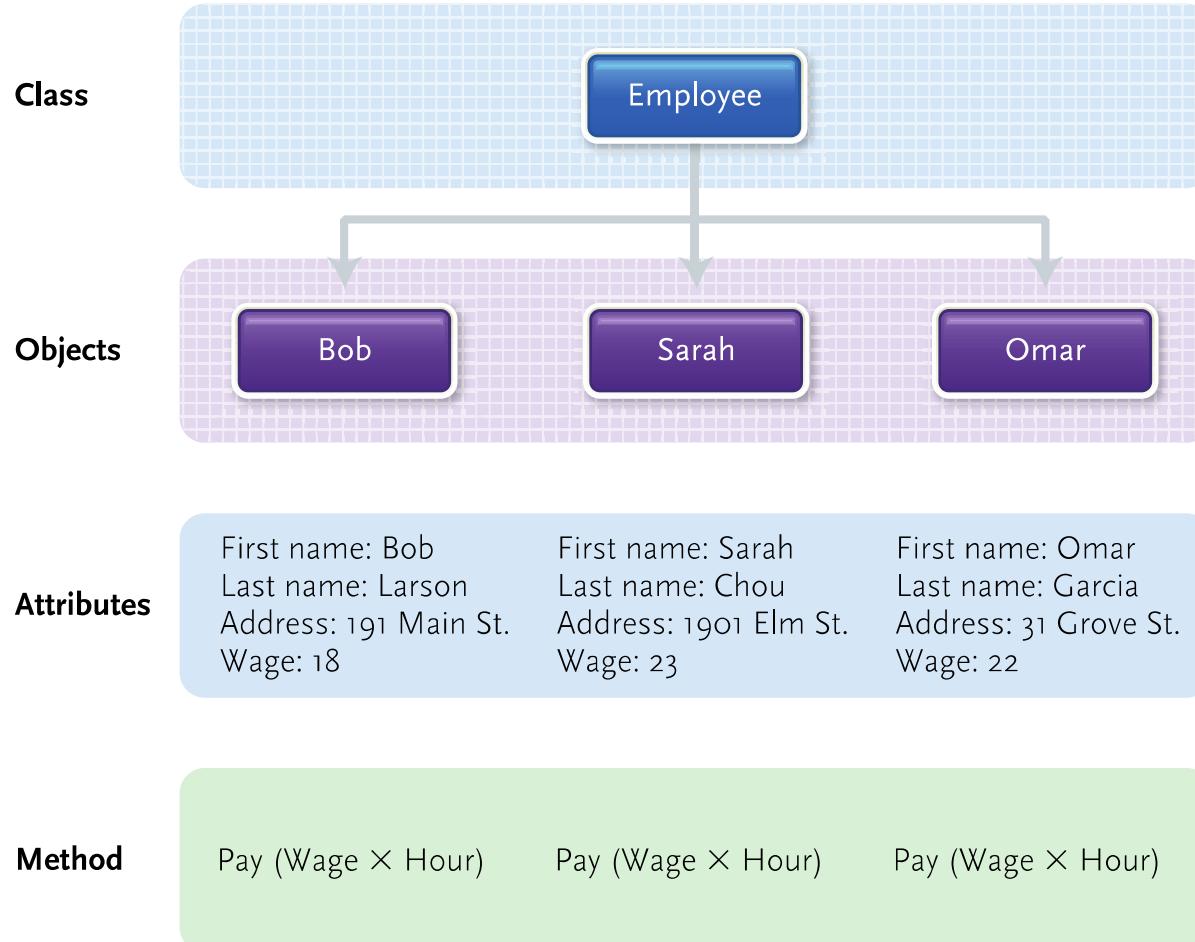
Cơ sở dữ liệu đa chiều

- Mở rộng của cơ sở dữ liệu quan hệ từ hai chiều thành nhiều chiều hơn
- Sử dụng cho các quan hệ dữ liệu phức tạp
- Tăng tốc độ xử lý

Cơ sở dữ liệu hướng đối tượng

- Cơ sở dữ liệu hướng đối tượng lưu giữ dữ liệu và cả lệnh để thao tác với dữ liệu đó.
- Tổ chức của CSDL hướng đối tượng bao gồm:
 - Lớp (Classes): Định nghĩa khuôn mẫu chung cho các đối tượng
 - Đối tượng (Objects): là các trường hợp cụ thể của lớp chứa cả dữ liệu và lệnh để thao tác dữ liệu
 - Thuộc tính (Attributes): là các trường dữ liệu của đối tượng sở hữu
 - Phương thức (Methods): là các lệnh để tìm kiếm hoặc thao tác các giá trị của thuộc tính.

Cơ sở dữ liệu hướng đối tượng



Các kiểu cơ sở dữ liệu

- CSDL cá nhân: dùng cho một người
- CSDL công ty: được dùng chung chia sẻ bởi nhiều người trong công ty
- CSDL phân tán: phân bố ở các vị trí địa lý khác nhau, được truy nhập thông qua các đường truyền thông
- CSDL thương mại: sẵn dùng với các thuê bao (ngân hàng dữ liệu)

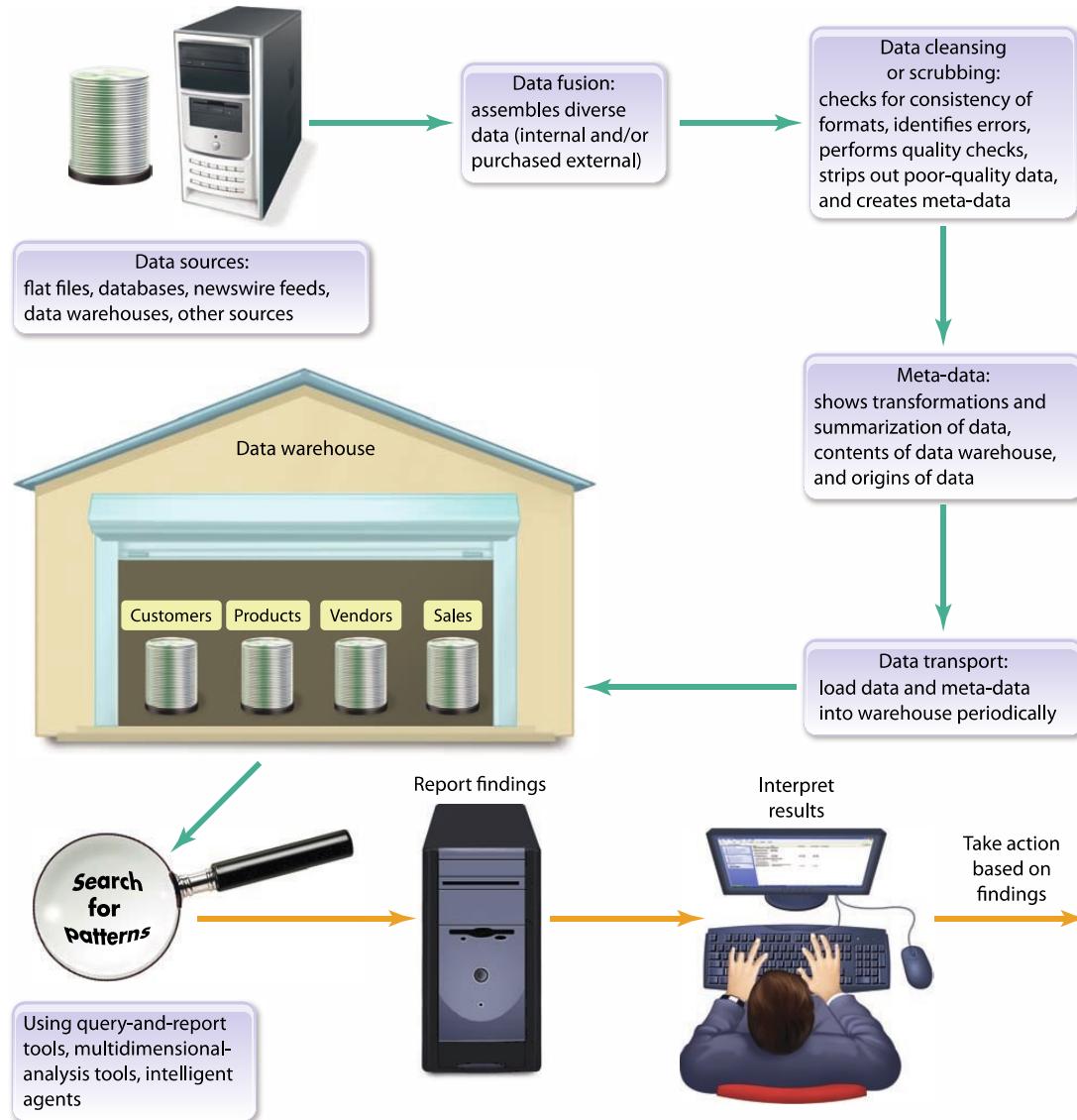
4. Sử dụng CSDL

- CSDL đưa ra cơ hội lớn để tăng năng suất lao động
- Việc duy trì CSDL cần cố gắng để không bị xáo trộn hoặc bị lạm dụng
- Chiến lược sử dụng:
 - Thường xuyên cập nhật từ các nguồn thông tin bên trong hoặc bên ngoài
 - Tổ chức thành kho dữ liệu (data warehouse)
 - Khai phá dữ liệu (data mining): là kỹ thuật dùng để tìm kiếm và khai phá dữ liệu để lấy ra các thông tin và các mẫu liên quan.
- An toàn cơ sở dữ liệu:
 - Chống sử dụng bất hợp pháp → sử dụng firewalls
 - Chống truy cập không đúng thẩm quyền → cần phân quyền truy cập cho người dùng

Khai phá dữ liệu và Kho dữ liệu

- Khai phá dữ liệu (Data Mining) là quá trình sàng lọc và phân tích lượng lớn dữ liệu để trích xuất ra các mẫu và các nghĩa ẩn và để phát hiện tri thức mới.
- Kho dữ liệu (Data Warehouse) là cơ sở dữ liệu đặc biệt của dữ liệu và siêu dữ liệu (metadata) sạch
- Dữ liệu được đưa vào một kho dữ liệu thông qua các bước sau:
 - Xác định và kết nối với các nguồn dữ liệu
 - Thực hiện tổng hợp dữ liệu và làm sạch dữ liệu
 - Nhận được cả dữ liệu và siêu dữ liệu (dữ liệu về dữ liệu)
 - Chuyển dữ liệu và siêu dữ liệu cho kho dữ liệu

Khai phá dữ liệu và Kho dữ liệu



5. Giới thiệu chung về hệ thống thông tin

- Hệ thống thông tin (HTTT) dựa trên máy tính bao gồm: Con người, Các qui trình, Phần cứng, Phần mềm, Dữ liệu, Kết nối mạng.
- Các thành phần trên làm việc cùng nhau để cung cấp các thông tin cần thiết cho tổ chức hay doanh nghiệp hoạt động.

Các bộ phận chính của tổ chức/doanh nghiệp

- Kế toán và tài chính (Accounting and Finance)
- Tiếp thị và bán hàng (Marketing and Sales)
- Nhân sự (Human Resources)
- Sản xuất (Production)
- Nghiên cứu và phát triển (Research and Development – R&D)

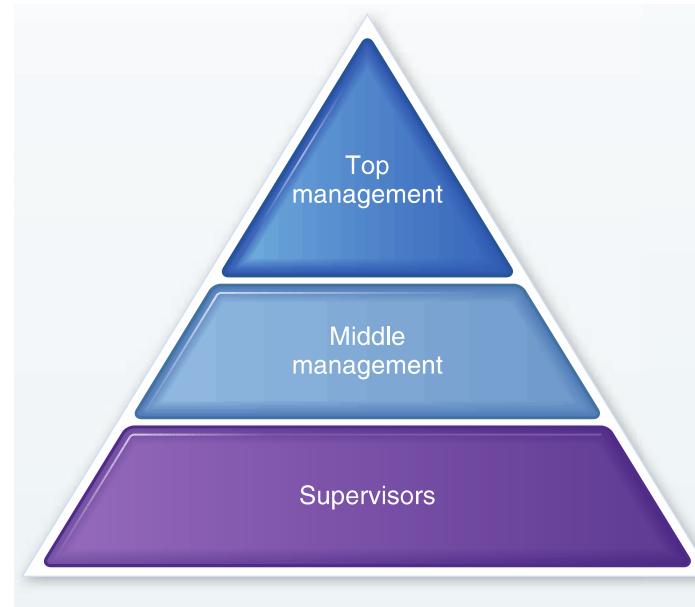
Các mức quản lý trong tổ chức/doanh nghiệp

- Quản lý mức đỉnh:
 - Liên quan với kế hoạch dài hạn
 - Cần thông tin để lập kế hoạch phát triển tương lai và định hướng của tổ chức/doanh nghiệp.
- Quản lý mức trung gian:
 - Kiểm soát, lập kế hoạch ngắn hạn
 - Thực hiện các mục đích ngắn hạn
- Quản lý mức giám sát:
 - Giám sát công việc hàng ngày của nhân viên và công nhân.



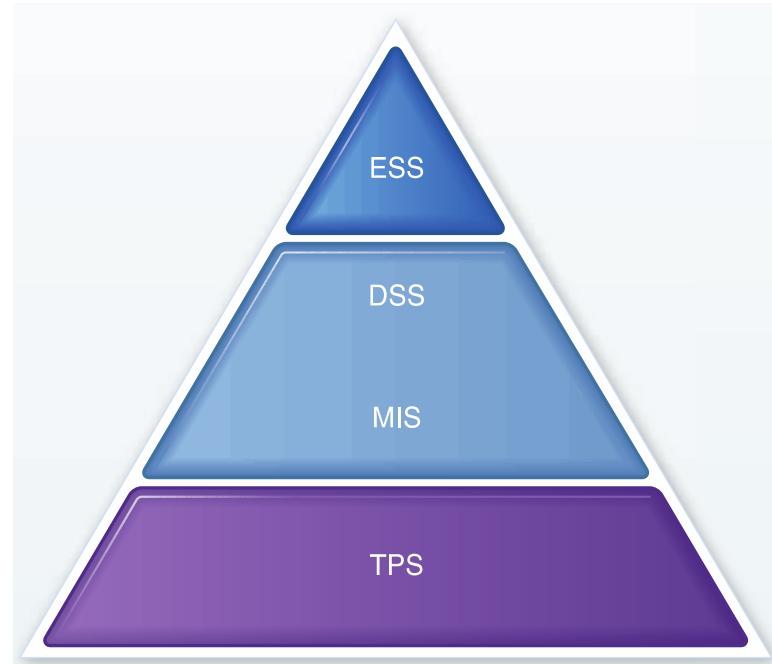
Luồng thông tin trong tổ chức/doanh nghiệp

- Quản lý mức đỉnh:
 - Luồng thông tin theo chiều ngang và dọc
 - CEO cần thông tin từ dưới các bộ phận và từ bên ngoài tổ chức/doanh nghiệp
 - Lãnh đạo từng bộ phận cần thông tin từ cấp dưới (dọc) và từ bộ phận khác (ngang)
- Quản lý mức trung gian:
 - Luồng thông tin theo chiều ngang và dọc
- Quản lý mức giám sát:
 - Luồng thông tin chủ yếu theo chiều dọc
 - Người giám sát trao đổi thông tin với quản lý mức trung gian và với nhân viên cấp dưới



6. Các HTTT của tổ chức/doanh nghiệp

- Hệ thống xử lý giao dịch TPS
 - Transaction Processing Systems
 - cho quản lý giám sát
- Hệ thống thông tin quản lý MIS
 - Management Information Systems
 - cho quản lý mức trung gian
- Hệ thống hỗ trợ quyết định DSS
 - Decision Support Systems
 - cho quản lý mức trung gian
- Hệ thống hỗ trợ điều hành ESS
 - Executive Support Systems
 - cho quản lý mức đỉnh



Hệ thống xử lý giao dịch (TPS)

- Ghi lại các giao dịch hằng ngày vào cơ sở dữ liệu
- Còn được gọi là hệ thống xử lý dữ liệu
- Một trong những sử dụng chủ yếu nhất của TPS là trong lĩnh vực kế toán:
 - Xử lý đơn đặt hàng
 - Các khoản thu
 - Hàng tồn kho và mua nguyên vật liệu
 - Các khoản chi trả
 - Lương

Hệ thống thông tin quản lý (MIS)

- Đưa ra các báo cáo chuẩn để hỗ trợ quyết định cho người quản lý mức trung gian.
- Tích hợp dữ liệu và tổng hợp chi tiết từ cơ sở dữ liệu theo dạng có cấu trúc
- Đưa ra các báo cáo:
 - Báo cáo định kỳ
 - Báo cáo ngoại lệ
 - Báo cáo theo yêu cầu

Hệ thống hỗ trợ quyết định (DSS)

- Công cụ linh hoạt để phân tích dữ liệu cho mục đích quyết định
- Cho phép người quản lý có câu trả lời với các vấn đề không mong muốn và hiếm xảy ra
 - Đưa ra các báo cáo không có dạng cố định
- Các thành phần của DSS:
 - Người sử dụng: người quyết định
 - Phần mềm của hệ thống
 - Dữ liệu: bên trong, bên ngoài
 - Mô hình quyết định: chiến lược, chiến thuật và hoạt động

Hệ thống hỗ trợ điều hành (ESS)

- Được thiết kế cho quản lý mức đỉnh
- Phần mềm để trình bày, tóm tắt và phân tích dữ liệu, nhưng phải dễ sử dụng
- Cung cấp truy cập tức thì vào những chỉ số hiệu suất chính của tổ chức/doanh nghiệp.

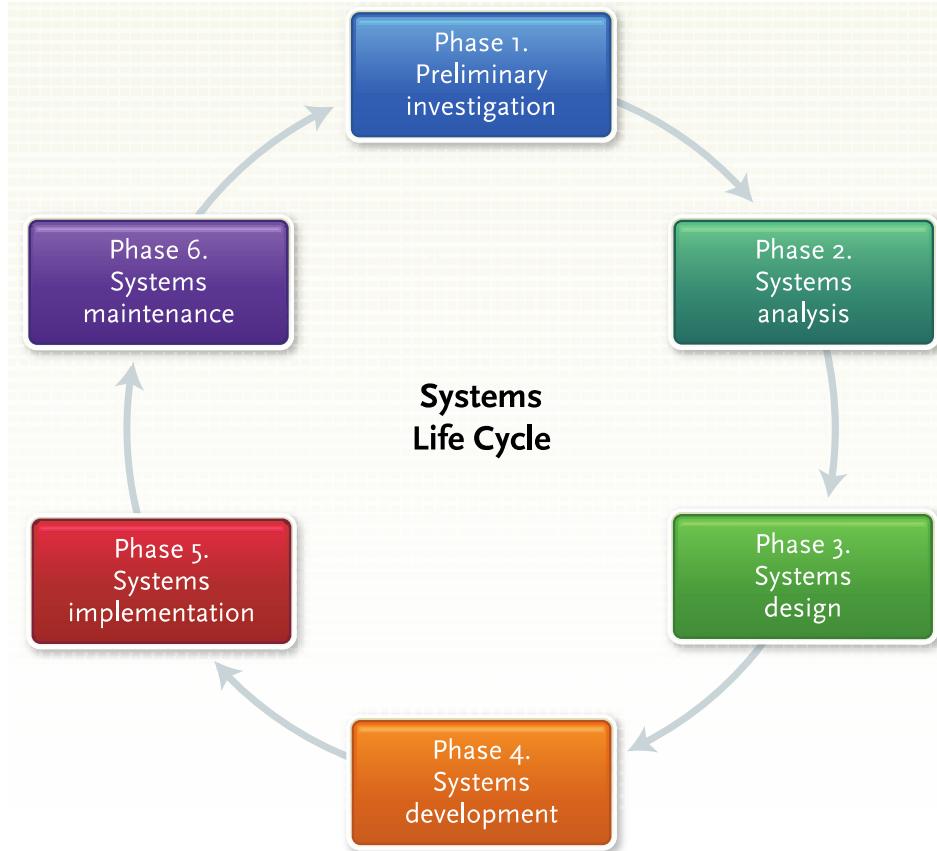
Các hệ thống khác

- Nhân viên thông tin
 - Nhân viên dữ liệu
 - Nhân viên tri thức
- Hệ thống tự động hóa văn phòng
 - Office Automation Systems (OAS)
 - Hỗ trợ nhân viên dữ liệu
 - Quản lý tài liệu, lập kế hoạch, lập lịch, trao đổi thông tin
 - Ví dụ: Chương trình quản lý dự án
- Hệ thống công việc tri thức
 - Knowledge Work Systems (KWS)
 - Các hệ thống thông tin chuyên dụng
 - Ví dụ: Hệ thống Computer-Aided Design/Computer-Aided Manufacturing Systems (CAD/CAM)

7. Phân tích và thiết kế hệ thống thông tin

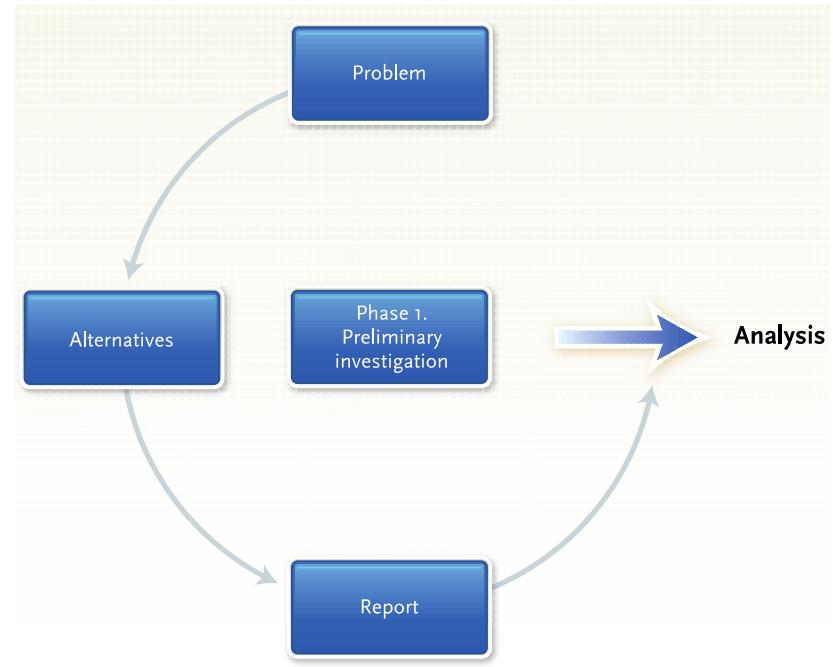
- Pha 1: Khảo sát sơ bộ (Preliminary investigation)
- Pha 2: Phân tích hệ thống (Systems analysis)
- Pha 3: Thiết kế hệ thống (Systems design)
- Pha 4: Phát triển hệ thống (Systems development)
- Pha 5: Thực hiện hệ thống (Systems implementation)
- Pha 6: Bảo trì hệ thống (Systems maintenance)

Vòng đời phát triển hệ thống



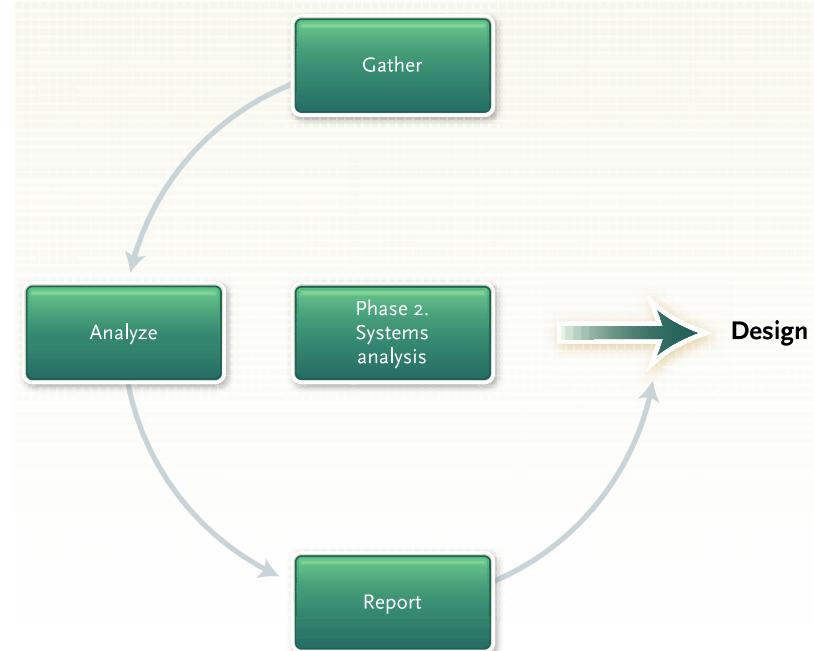
Pha 1: Khảo sát sơ bộ

- Xác định vấn đề:
 - Khảo sát hệ thống hiện tại
 - Xác định ai cần thông tin gì, khi nào thì cần thông tin, tại sao?
- Đề xuất các hệ thống thay thế
- Chuẩn bị báo cáo ngắn



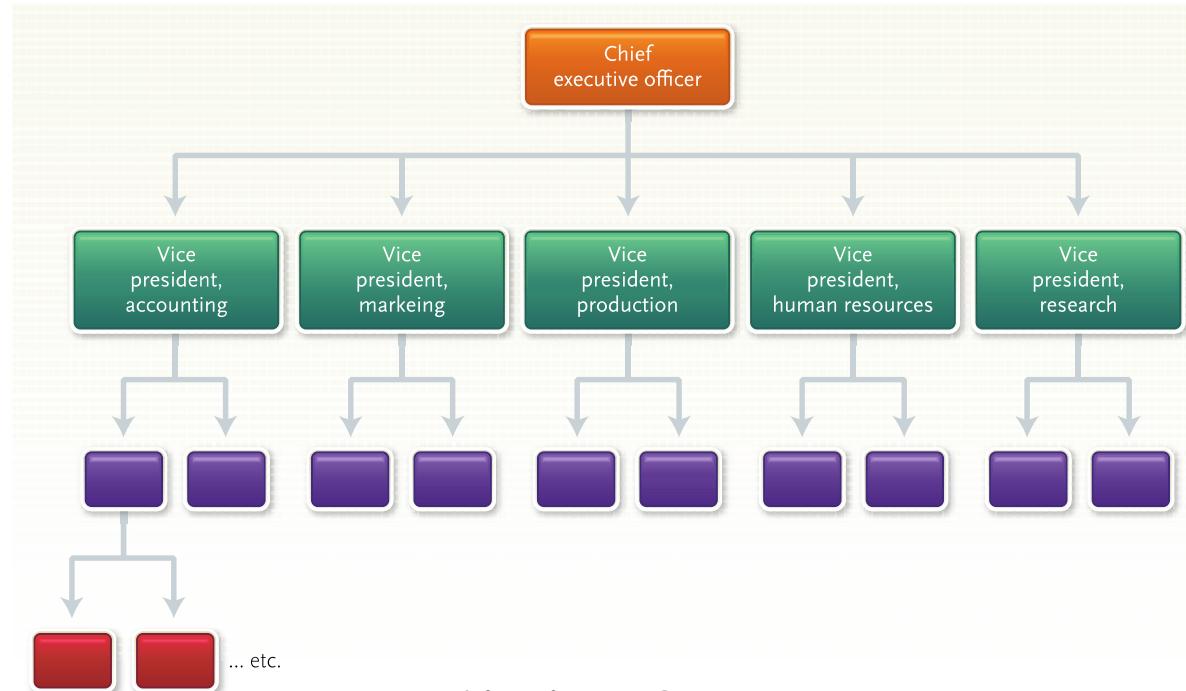
Pha 2: Phân tích hệ thống

- Thu thập dữ liệu
- Phân tích dữ liệu
- Lập báo cáo



Thu thập dữ liệu

- Quan sát
- Phỏng vấn
- Bảng câu hỏi
- Xem các tài liệu
 - Sơ đồ tổ chức của doanh nghiệp

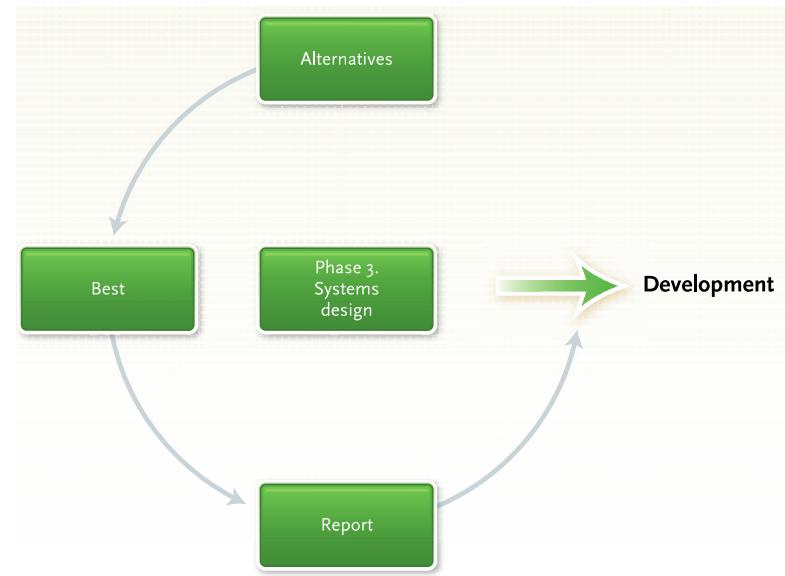


Các công cụ phân tích dữ liệu

- Checklist
- Phân tích top-down
- Sơ đồ dạng lưới
- Các bảng quyết định
- Lưu đồ hệ thống
- Biểu đồ luồng dữ liệu
- Công cụ thiết kế tự động (CASE)
 - Computer-Aided Software Engineering tools

Pha 3: Thiết kế hệ thống

- Thiết kế các hệ thống thay thế
 - Đánh giá hệ thống theo:
 - Khả thi về kinh tế
 - Khả thi về kỹ thuật
 - Khả thi về vận hành
- Lựa chọn hệ thống tốt nhất
- Viết báo cáo

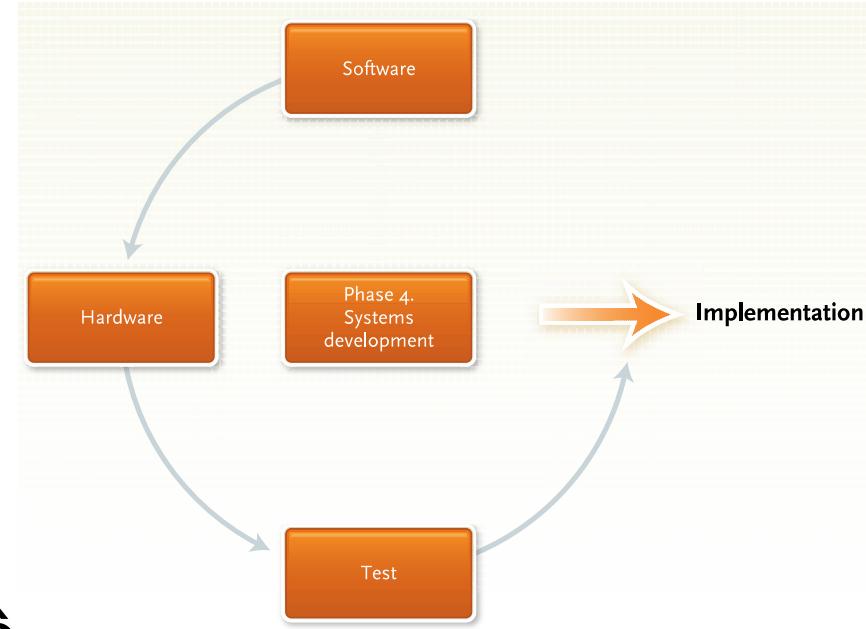


Tiêu chí lựa chọn thiết kế

- (1) Hệ thống có phù hợp với HTTT tổng thể của tổ chức/doanh nghiệp ?
- (2) Hệ thống có đủ mềm dẻo để nó có thể được thay đổi trong tương lai ?
- (3) Nó có thể được thực hiện an toàn chống lại sử dụng trái phép ?
- (4) Những lợi ích đáng giá trị với chi phí ?

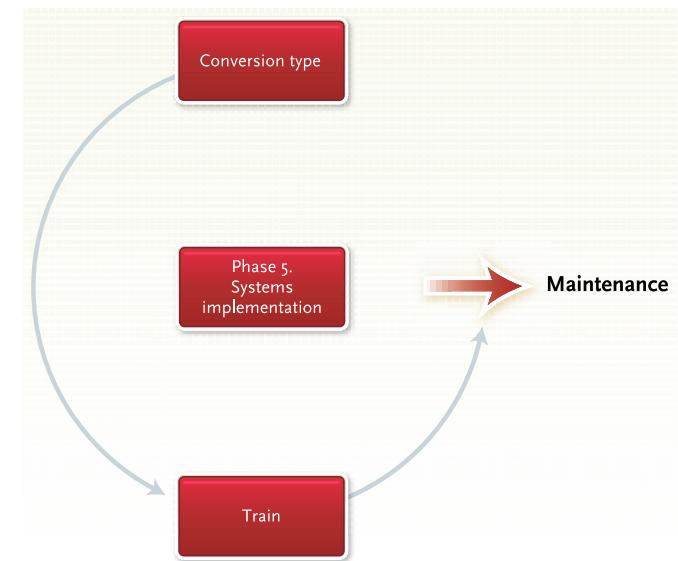
Pha 4: Phát triển hệ thống

- Phần mềm
 - Mua gói phần mềm thích hợp có khả năng tùy biến để phù hợp với hệ thống mới
 - Thiết kế và xây dựng mới
- Phần cứng:
 - Sử dụng thiết bị sẵn có
 - Mua thêm hoặc mua mới
 - Kết nối mạng
- Tích hợp và Kiểm thử hệ thống



Pha 5: Thực hiện hệ thống

- Các kiểu chuyển đổi sang hệ thống mới
 - Giải pháp trực tiếp
 - Giải pháp song song
 - Giải pháp thí điểm
 - Giải pháp theo từng giai đoạn
- Đào tạo người sử dụng



Các kiểu chuyển đổi

Kiểu chuyển đổi	Mô tả	Đặc điểm
Trực tiếp	Bỏ hệ thống cũ, xây hệ thống mới từ đầu	Rủi ro cao, không khuyên dùng
Song song	Chạy cả hệ thống cũ và hệ thống mới	Rủi ro rất thấp, tốn kém, không khuyên dùng
Thí điểm (pilot)	Chuyển đổi một phần sang hệ thống mới	Rủi ro thấp, đỡ tốn kém
Theo pha (phase)	Triển khai từ từ từng giai đoạn	Rủi ro thấp, tốn kém, hoàn thành chậm

Pha 6: Bảo trì hệ thống

- Kiểm nghiệm hệ thống
- Đánh giá định kỳ

