***TỔNG QUAN VỀ CSDL***

**1. Thông tin, dữ liệu**

Dữ liệu là thông tin của đối tượng (người, vật, sự việc,…) dưới dạng kí hiệu, chữ viết, hình ảnh, âm thanh,… được lưu trữ trên máy tính.

***Vd:***

* Dữ liệu về nhân viên, quá trình công tác, chuyên môn trong công ty
* Dữ liệu thời tiết, khí hậu,…

Tùy thuộc vào ngữ cảnh mà dữ liệu về đối tượng có thể khác nhau:

***Vd:*** Dữ liệu về một sinh viên có thể khác nhau tùy vào mục đích quản lí:

* Quản lí điểm: Tên, MSSV, Điểm môn 1, Điểm môn 2,…
* Quản lí nhân thân: Tên, Ngày sinh, Địa chỉ, Quê quán,…

Thông tin là dữ liệu đã được xử lí như phân loại, tính toán, chọn lọc,…để phù hợp với ngữ cảnh.

***Vd:*** Điểm trung bình của lớp là thông tin có được từ dữ liệu điểm sinh viên

***Tóm lại:*** Thông tin hợp thành từ các dữ liệu, có được từ việc xử lý dữ liệu, lột tả ý nghĩa của dữ liệu; Thông tin đúng đắn, kịp thời và đầy đủ là mấu chốt để thực hiện những quyết định; Quyết định đúng đắn là then chốt quyết định sự sống còn của tổ chức.

**2. Khái niệm về CSDL**

**Hệ thống tập tin cổ điển:** Là dữ liệu được tổ chức riêng lẻ, phục vụ cho mục đích của một đơn vị cụ thể.

***Ưu điểm:*** Triển khai ứng dụng nhanh, có khả năng đáp ứng nhanh chóng, kịp thời.

***Nhược điểm:***

* Dư thừa dữ liệu (data redundancy), các phiên bản khác nhau của cùng một dữ liệu có thể mâu thuẫn nhau, gây lãng phí công sức nhập liệu và không gian lưu trữ, tạo những dị thường về dữ liệu (data anomalies): khi thêm, sửa, xóa
* Dữ liệu không nhất quán (data inconsistency): thiếu các ràng buộc về dữ liệu (data integrity).
* Thiếu sự chia sẻ thông tin giữa các nơi.
* Truy cập đồng thời bất thường
* Vấn đề bảo mật và toàn vẹn dữ liệu

“***Cơ sở dữ liệu (database)*** *là một tập dữ liệu có liên quan với nhau, được lưu trữ trên máy tính, băng từ, đĩa từ,… có nhiều người sử dụng, nhiều ứng dụng cùng truy cập với nhiều mục đích khác nhau và được tổ chức theo một mô hình cụ thể.”*

***Ví dụ:*** cơ sở dữ liệu SOHUUOTO có các bảng:

* CONGDAN: thông tin về công dân có thể sở hữu xe ô tô
* OTO: thông tin về ô tô
* SOHUU: thông tin về sở hữu ô tô của công dân

A screenshot of a computer

Description automatically generated

**Các tính chất của CSDL:**

* Một CSDL biểu thị một khía cạnh nào đó của thế giới thực. Thông tin được đưa vào trong CSDL tạo thành một không gian CSDL hoặc một “thế giới nhỏ” (miniworld).
* Một cơ sở dữ liệu được thiết kế và được phổ biến cho một mục đích riêng, là một tập hợp dữ liệu liên kết với nhau một cách logic và mang một ý nghĩa nào đó, không phải là một tập dữ liệu tùy tiện.

**Ưu điểm của CSDL:**

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

**Kiến trúc 3 lớp của CSDL:**

A diagram of a cloud

Description automatically generated

**Mức trong:** (mức vật lý – Physical) là mức lưu trữ CSDL

* Lưu trữ như thế nào? ở đâu?
* Dành cho người quản trị CSDL (Database Administrator).

**Mức quan niệm:** (Conception hay Logical)

* Cần phải lưu trữ bao nhiêu loại dữ liệu? là dữ liệu gì? mối quan hệ

**Mức ngoài:** của người sử dụng không chuyên và các chương trình ứng dụng

**Các đối tượng sử dụng CSDL:**

* Người sử dụng CSDL không chuyên về lĩnh vực tin học và CSDL **(End User)**: Cần công cụ để họ có thể khai thác CSDL khi cần.
* Chuyên viên tin học **(Application User)**: Xây dựng các ứng dụng để phục vụ cho các mục đích quản lý.
* Quản trị CSDL **(Database administrator)**: Tổ chức CSDL, bảo mật, cấp quyền, sao lưu, phục hồi dữ liệu, giải quyết các tranh chấp dữ liệu…