

Introduzione alle basi di dati

Un database è **un'insieme organizzato di dati** che esistono e si sviluppano nel tempo, disponibili in una certa struttura (impresa, banca, ospedale ecc..) per lo svolgimento della propria attività.

Esempi di utilizzo di un database:

- mantenimento di "*record*" interni ad una struttura;
- offrire servizi attraverso il *World-Wide-Web*;
- alla base di ricerche scientifiche e usate per memorizzare e rappresentare dati.

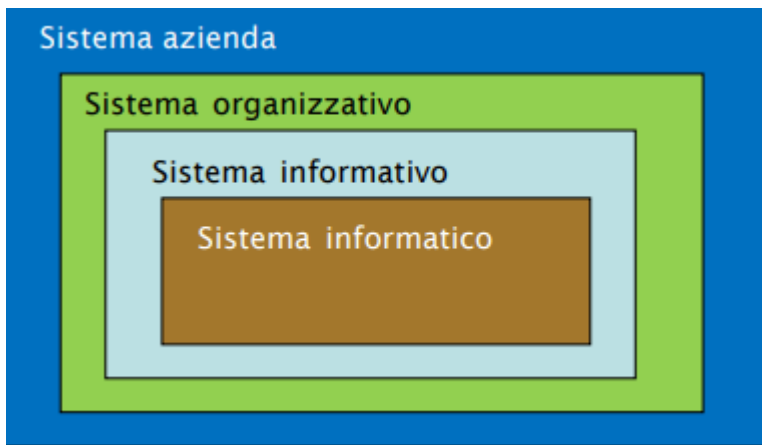
Sistema organizzativo e sistema informativo

Il **sistema organizzativo** di un'azienda è formato da:

- insieme di **risorse** (persone, denaro, **informazioni**) dove le informazioni possono essere, per esempio, quelle personali riguardo i propri interessi online affinché si ricevano le cosiddette "*pubblicità targhettate*";
- **regole** che servono per lo svolgimento coordinato delle attività al fine del perseguimento degli scopi.

Il **sistema informativo** è parte del sistema organizzativo che *esegue/gestisce* processi informativi (cioè i processi che coinvolgono le informazioni).

Non vuol dire, però, che un sistema informativo debba necessariamente fare ricorso a strumenti automatici; esistono organizzazioni la cui ragion d'essere è la gestione di informazioni e che operano da secoli, fin da prima dell'invenzione degli automatismi (servizi anagrafici e banche) che usavano, appunto, carta e penna per la registrazione delle informazioni.



Gestione delle informazioni

In qualsiasi attività le informazioni vengono gestite (*registrate e scambiate*) in diversi **modi**:

- idee informali;
- *linguaggio naturale*(scritto o parlato);
- *disegni*, grafici, schemi;
- *numeri* e codici.
..e su vari **supporti**:
- *memoria umana*, carta, dispositivi elettronici.

Informazioni e dati

Le informazioni vengono rappresentate in modo **essenziale** attraverso i dati. Si distinguono:

- **Informazione** = notizia, dato che consente di avere conoscenza più o meno esatta dei fatti;
- **Dato** = ciò che è immediatamente presente alla coscienza, prima di ogni elaborazione.

I dati devono essere interpretati, per esempio:

"Mario" "2075"

Se questi dati vengono forniti in risposta alla domanda "A chi mi devo rivolgere per il problema X" "Qual è il suo **interno?**"

Allora i dati **possono essere interpretati** per fornire informazione e *arricchire la conoscenza*.

I dati costituiscono spesso una risorsa strategica, perché più stabili nel tempo di altre componenti (processi, tecnologie, ruoli umani).

Basi di dati DBMS

La potenza delle basi di dati deriva da un bagaglio di conoscenze e tecnologie che sono state sviluppate nel tempo che hanno dato luogo a software specializzati a questo compito chiamati **DataBase Management System** (Sistema di gestione di basi di dati) **DBMS**.

Un DBMS è uno strumento potente per la creazione e gestione efficiente ed efficace di grandi quantità di dati.

Caratteristiche DBMS

Questo sistema (prodotto software) è in grado di gestire collezioni di dati che siano anche:

- **grandi**, di dimensioni maggiori della memoria centrale dei sistemi di calcolo usati;
- **persistenti**, con un periodo di vita indipendente dalle singole esecuzioni del programmi che la utilizzano;
- **condivise**, cioè usate da applicazioni diverse.

Un DBMS consente all'utente di accedere e modificare i dati attraverso un potente linguaggio di interrogazione garantendo **affidabilità** (cioè resistente a malfunzionamenti software o hardware) e **privatezza** (cioè con una disciplina di controllo degli accessi ai dati stessi).

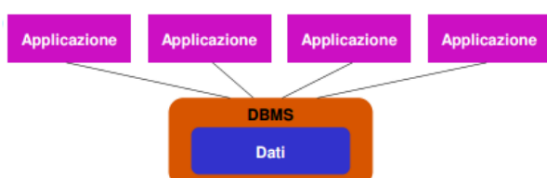
Ogni DBMS deve essere **efficiente**, cioè deve usare al meglio le risorse spazio-temporali del sistema del sistema operativo, ed **efficace**, cioè rendere produttive le attività dei suoi utilizzatori.

Esempi di DBMS sul mercato

Essi sono prodotti software (complessi):

- DB2;
- Oracle;
- SQLServer;
- MySQL;
- PostgreSQL.

Condivisione delle informazioni



Una base di dati è una **risorsa integrata, condivisa** fra le varie applicazioni. Si distinguono **conseguenze** e **problemi** di tale procedura:

- **Conseguenze**: attività diverse su dati in parte condivisi: meccanismi di **autorizzazione** e attività multi-utente su dati condivisi: controllo della **concorrenza**
- **Problemi** relativi alla condivisione: **ridondanza**, cioè informazioni ripetute; e rischio di **incoerenza**, cioè le versioni non possono coincidere.

Modello dei dati

Il modello dei dati è un formalismo (matematico) che è formato da due parti:

- **notazione** che serve per descrivere i dati;
- **operatori** che servono per manipolare i dati.

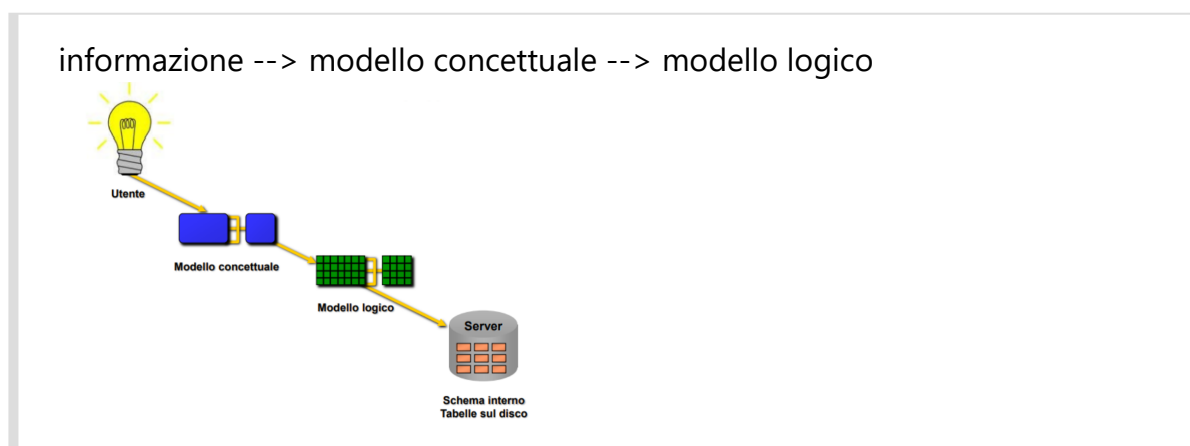
Ogni modello di dati fornisce meccanismi di strutturazione, cioè come nei linguaggi di programmazione esistono meccanismi che permettono di definire nuovi tipi, così ogni modello dei dati prevede alcuni costruttori.

I principali due tipi di modelli sono:

1. Modello **logico**, usato nei DBMS esistenti per l'organizzazione di dati. Esso è usato dai programmi e sono indipendenti dalle strutture fisiche;
Per esempio: relazionale (che usa una notazione semplice, cioè la tabella), reticolare, gerarchico, a oggetti.
2. Modello **concettuale** che permette di rappresentare i dati indipendentemente da ogni sistema. Esso cerca di descrivere i concetti del mondo reale piuttosto che i dati utili a rappresentarli e sono usati nelle fasi preliminari di progettazione.
Per esempio: Entità-Relazione

Architettura di un DBMS

Un DBMS segue una particolare architettura, del tipo:



Per schema logico si intende la descrizione della base di dati (ad esempio, la struttura della tabella).

Una vista

Corsi

Corso	Docente	Aula
Basi di dati	Rossi	4
Algoritmi	Neri	2
Reti	Bruni	1
Analisi dati	Rossi	G

Aule

Nome	Edificio	Piano
4	DMI	Terra
1	DMI	Terra
G	DAU	Primo

CorsiSedi

Corso	Aula	Edificio	Piano
Basi di dati	4	DMI	Terra
Reti	1	DMI	Terra
Analisi dati	G	DAU	Primo

Una vista è un "prodotto" fornito dall'**unione di 2 o più tabelle**. E' una *tabella derivata* da altre tabelle precedenti esistenti e si dice "*virtuale*". Essa contiene parti di righe/colonne di alcune tabelle che la generano, appunto.

Se si aggiorna una tabella, in automatico, si aggiorna anche la vista.

Schermo interno: il metodo di memorizzazione all'interno di un DBMS a livello proprio di segnali (che è una parte che a noi non interessa);

Schema logico: interazione con l'utente per l'avviamento della procedura di memorizzazione dei dati (la parte a noi interessante).

Linguaggi per le basi di dati

Nei DBMS si distinguono:

- **Data-Definition Language (DDL):** consente all'utente di creare nuovi database e di specificarne i loro schemi e la costruzione logica;
- **Data-Manipulation Language (DML):** da agli utenti la possibilità di interrogare e modificare le istanze di basi di dati.
- **Data Query Language (DQL):** consente di avviare interrogazioni, sia semplici che complesse, dei dati memorizzati

Query è un termine che serve per indicare che il DBMS risponde ad una precisa domanda rispetto a uno o più tabelle. Per esempio:

- Vorrei conoscere l'indirizzo e il telefono di Giovanni Verdi

Nome	Matricola	Indirizzo	Telefono
Mario Rossi	123456	Via Etna 18	777777
Maria Bianchi	234567	Via Roma 2	888888
Giovanni Verdi	345678	Via Etna 18	999999
Enzo Gialli	456789	Via Catania 3	444444

Indirizzo	Telefono
Via Etna 18	999999