

Cloud informatica

V anno



MONDADORI EDUCATION

Sistemi informativi e basi di dati



MONDADORI EDUCATION

Introduzione

In ogni **modello organizzativo** (ufficio, biblioteca, scuola ecc.) sono trattate **informazioni**.

Memorizzare tali informazioni permette di:

- **Recuperarle**;
- **Aggiungerne** di nuove;
- **Modificarle**;
- **Cancellarle**.

In questo l'**informatica** ricopre un ruolo di primaria importanza.



MONDADORI EDUCATION

La teoria delle basi di dati

In campo informatico, la **teoria delle basi di dati** permette di gestire grandi quantità di informazioni in modo:

- **semplice**: le applicazioni devono essere trattate da diversi programmi e diversi utenti;
- **efficiente**: l'utilizzo delle risorse deve essere ottimizzato:
 - “in tempo” (efficiente utilizzo del microprocessore)
 - “in spazio” (efficiente utilizzo della memoria);
- **efficace**: le informazioni devono essere rappresentative della realtà che si vuole analizzare;
- **sicuro**: accesso solo su autorizzazione e identificazione.



Ma cosa è una base di dati?

Una **base di dati** (*database* - DB) è una raccolta di dati logicamente correlati, utilizzata per modellare una realtà.

I **dati** sono memorizzati e progettati per una fruizione ottimizzata e sicura.

Per essere definita tale, una base di dati deve essere:

- sicura
- integra
- consistente
- condivisibile
- persistente
- scalabile



L'obiettivo

Per progettare un database occorre sempre avere ben presenti i propri **obiettivi** generali, specifici o di breve durata.

Anche l'**organizzazione** generale della realtà (per es. un'azienda) a cui la base di dati è riferita ricopre un ruolo fondamentale.

Per descrivere meglio l'organizzazione occorre introdurre due utili concetti:

- **risorsa**: è tutto ciò con cui la realtà opera per perseguire i propri obiettivi;
- **processo**: è l'insieme delle attività che l'organizzazione nel suo complesso deve svolgere per gestire il ciclo di vita di una risorsa



Il sistema informativo

Un **sistema informativo** è un insieme organizzato di strumenti automatici, procedure, norme, risorse, orientato alla gestione delle informazioni.

Con **gestione** si intende la raccolta, archiviazione, elaborazione e scambio di informazioni necessarie alle attività operative, di gestione, di preparazione, di controllo e valutazione dell'organizzazione.

Una perfetta gestione delle informazioni permette di ottimizzare l'efficienza e la qualità del servizio offerto.



MONDADORI EDUCATION

Il sistema informatico

La parte del sistema informativo di un'organizzazione che può essere automatizzata è detta sistema informatico.

Il **sistema informatico** è l'insieme degli strumenti informatici utilizzati per il trattamento automatico delle informazioni.

La digitalizzazione delle informazioni agevola il lavoro sul sistema informativo.

In poche parole, **il sistema informatico è un sistema informativo che si avvale di tecnologie informatiche.**



MONDADORI EDUCATION

I dati

In informatica un **dato** è la rappresentazione di una cosa. Lo scopo dei dati è quello di **codificare i fatti rilevanti** per un'organizzazione.

I dati permettono di acquisire informazioni solo se presentano:

- una **chiave di interpretazione** (**schema** o **intensione**) che ne identifica il **significato**.
- l'**istanza di uno schema** (o **estensione**) che comprende l'insieme dei valori assunti da uno schema in un certo istante di tempo.

Lo schema non varia nel tempo, mentre l'istanza di uno schema può essere modificata.

Quando si modifica uno schema si ha un nuovo schema.



Categoria e occorrenza

Un gruppo di dati con uno stesso schema è detto **categoria**.

Una **occorrenza di una base di dati** è l'insieme delle istanze delle categorie in un determinato istante di tempo.



Il modello di dati

Un **modello di dati** è un insieme di concetti utilizzati per organizzare i dati e descriverne la struttura, le associazioni e i vincoli che devono rispettare (dinamica).

Nella teoria delle basi di dati, i modelli si distinguono in:

- **modelli concettuali**: che descrivono i dati e non la struttura con cui possono essere rappresentati (es. **Modello ER** - *Entity Relationship*)
- **modelli logici**: rappresentano in modo astratto l'organizzazione dei dati.

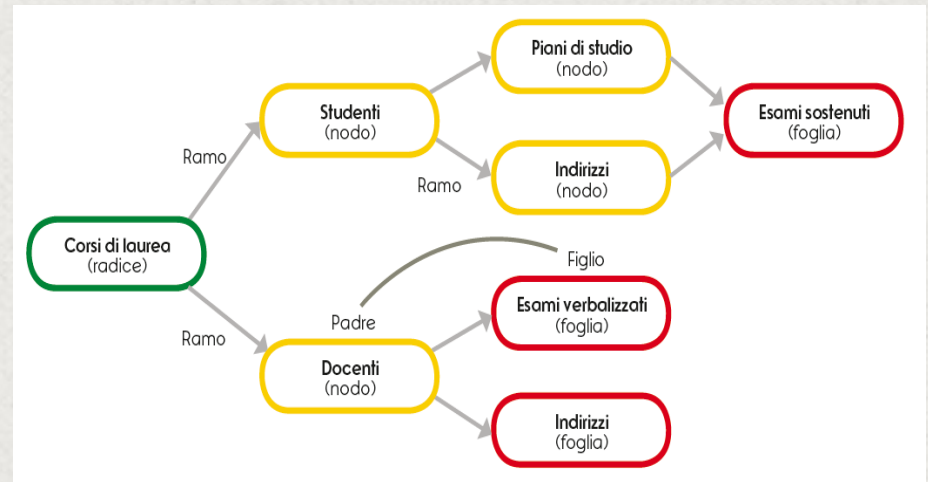
I principali modelli logici sono:

- modello **gerarchico**
- modello **reticolare**
- modello **relazionale**
- modello **a oggetti**



Modelli logici - Il modello gerarchico

Nel **modello gerarchico** i **dati** sono **organizzati in record** connessi secondo **strutture ad albero**.



Ogni record deve avere un **unico padre** ma possono esserci più record, su diversi alberi, che rappresentino la stessa informazione.

Il modello gerarchico comporta la **dipendenza dei programmi dalle strutture**.

Un esempio di modello gerarchico è il **file system**.



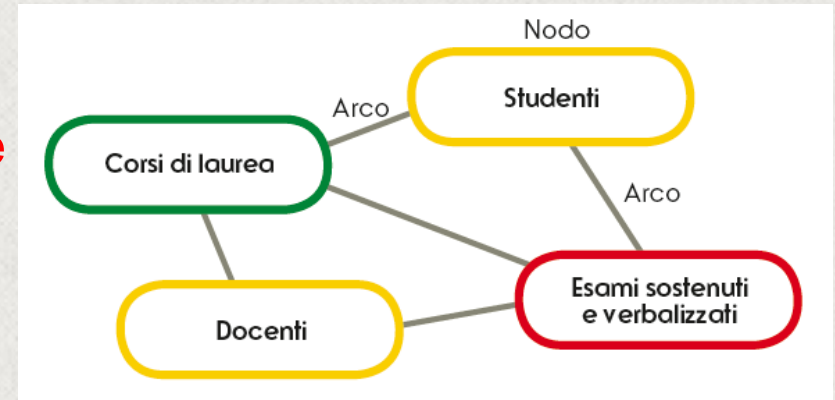
Modelli logici - Il modello reticolare

Detto anche **modello a rete** o **modello Codasyl**, il **modello reticolare** è basato sulle **strutture dati a reticolo**.

Deriva dal modello gerarchico, ma ne elimina la rigidità della struttura ad albero.

I record sono legati tra loro da **strutture ad anello** che permettono all'utente finale di accedere ai dati in maniera semplice, senza vincoli di struttura. Un record può avere **più record padre**.

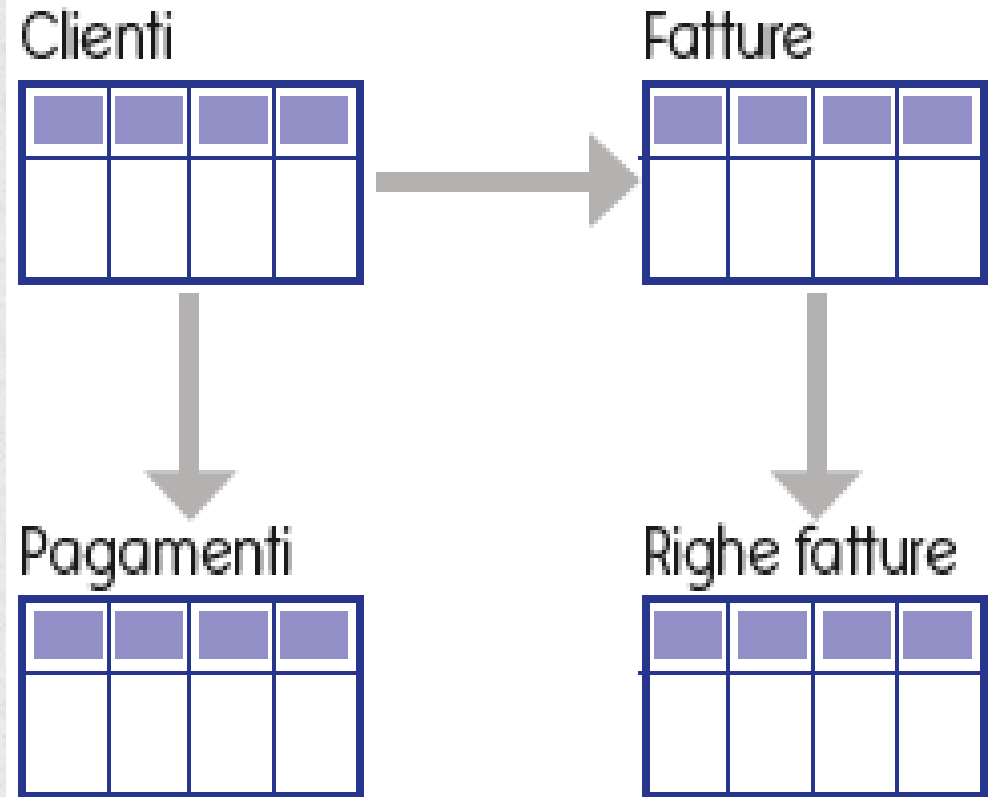
Per modificare un database con questa struttura, occorre cancellarlo e ricrearlo.



Modelli logici - modello relazionale

Il **modello relazionale** è basato sul concetto di **insiemi di record** (**tabelle**).

Le relazioni tra le informazioni derivano dalla corrispondenza di alcuni campi di record appartenenti a tabelle diverse.

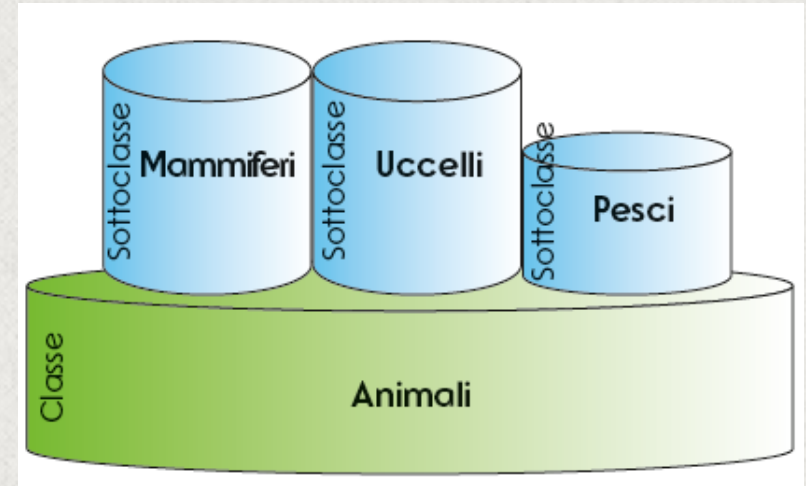


Modelli logici - modello a oggetti

Il **modello a oggetti** (*Object oriented*) è un modello relativamente giovane che nasce come evoluzione del modello relazionale.

Estende alle basi di dati il **concetto di programmazione a oggetti**.

Viene **eliminata la separazione netta tra il software di base e il software di programmazione dell'applicazione** che permette di interrogare il database. Pertanto, chi sviluppa il database ne gestisce tutti gli aspetti direttamente dal software di programmazione della base di dati.



Organizzazione informazioni prima delle basi di dati

Prima dell'introduzione dei database, l'organizzazione delle informazioni risentiva di **alcune problematiche**:

- **ridondanza** dei dati
- **costi addizionali** legati a più copie dello stesso dato
- **aggiornamento** dati **complesso** e **impreciso**
- **incongruenza** dei dati: i diversi archivi non erano aggiornati contemporaneamente
- **inconsistenza** dei dati: in uno stesso archivio potrebbero essere presenti dati aggiornati e dati non aggiornati
- **concorrenza**: con accessi contemporanei è mantenuta solo l'ultima modifica apportata su un dato



I Data Base Management System

I **Data Base Management System (DBMS)** nascono per ovviare a queste problematiche.

Un DBMS è un **insieme di strumenti software** che, partendo dalle specifiche dell'utente, è in grado di gestire dati strutturati.

In poche parole, un DBMS è un'interfaccia tra gli sviluppatori, gli utenti del DB e il sistema di elaborazione.



Caratteristiche di un DBMS

Le **funzioni** a cui un DBMS deve assolvere e garantire sono:

- gestione del database
- persistenza e consistenza dei dati
- privatezza e sicurezza dei dati
- integrità di dati
- supporto alle transazioni (sequenze di operazioni sul database)
- gestione del dizionario dei dati (il dizionario contiene i metadati, ossia le informazioni che descrivono gli oggetti del database).

Un DBMS deve essere, inoltre:

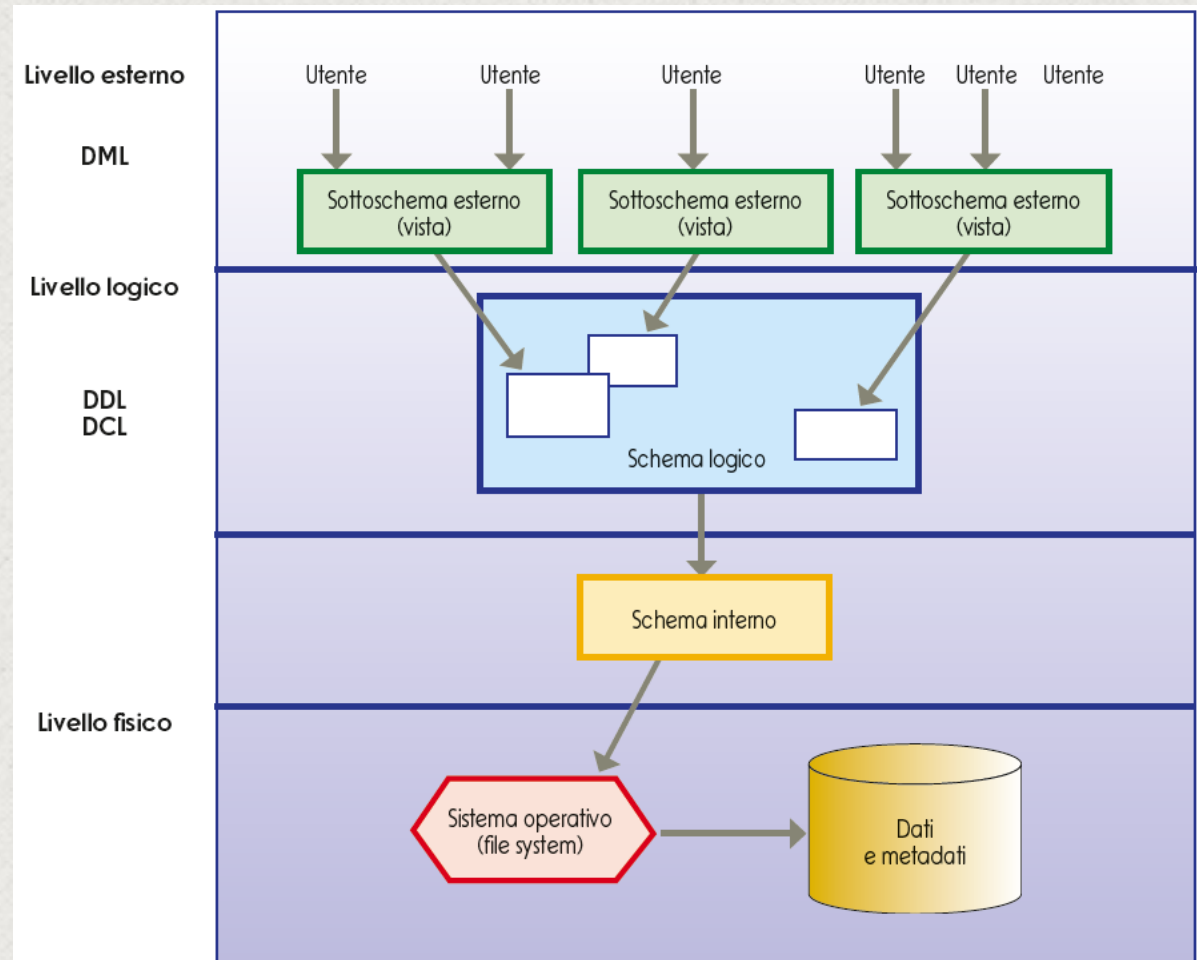
- efficiente, in termini di tempo e spazio
- efficace, rendendo produttive e semplici le attività richieste dagli utenti.



Livelli di astrazione di un DBMS

Il DBMS presenta un **modello astratto dei dati** (macchina DBMS) che permette agli utenti una facile interpretazione e iterazione con i dati Stessi.

L'architettura di tale macchina virtuale è rappresentata dalla seguente figura.



Livello esterno di un DBMS

Il **livello esterno** di un DBMS è quello con cui interagiscono i singoli utenti del database.

Il livello esterno è l'interfaccia che permette agli utenti di compiere le operazioni sui dati attraverso l'utilizzo di uno **specifico linguaggio** detto *Data Manipulation Language* (**DML**).

È possibile realizzare **accessi e navigazioni personalizzati** per utente.

Spetta all'**amministratore di sistema** (*Data Base Administrator* - **DBA**) fornire a ciascun utente un **sottoschema** (**vista**) del database che contenga solo ciò che l'utente desidera o è abilitato a utilizzare.



Livello logico di un DBMS

Il **livello logico** (o **concettuale**) rappresenta **la struttura globale del database**.

In questo livello troviamo:

- il **database logico**: definizione dell'intero schema dei dati;
- **Data Definition Language (DDL)**: linguaggio che descrive le caratteristiche delle categorie di dati presenti nel database e le relative corrispondenze;
- **Data Control Language (DCL)** controlla gli accessi degli utenti.



Livello fisico di un DBMS

Il **livello fisico** (o **interno**) memorizza i dati.

In questo livello troviamo:

- il **database fisico**: rappresenta la forma in cui il database viene memorizzato e usato;
- *Device Media Control Language (DMCL)*: linguaggio per ottimizzare l'organizzazione dei dati nel database.



Livelli di un DBMS, riassumendo...

Riassumendo, i **tre livelli** di un Data Base Management System:

- Il **livello fisico** riguarda la **memorizzazione** dei dati, organizzati in file, record e strutture di accesso;
- Il **livello logico** riguarda la struttura logica assunta dai dati registrati, quindi il loro **schema astratto**;
- Il **livello esterno** si riferisce al modo in cui ciascun utente vede i dati che vengono messi a disposizione secondo il **formato** richiesto.

Questa impostazione permette di ottenere:

- **indipendenza fisica dei dati**: possibilità di modificare l'organizzazione fisica dei dati senza modificare l'organizzazione logica e gli applicativi che operano sul DB.
- **indipendenza logica dei dati**: possibilità di modificare lo schema logico del DB senza dover modificare gli applicativi non interessati dalla modifica.



Fasi di progettazione di un database

La **progettazione di un database** prevede **tre fasi**:

- **progettazione concettuale**: permette di costruire e definire una rappresentazione corretta e completa della realtà di interesse.

Input: documento specifiche formali

Output: schema concettuale della realtà (diagramma ER)

- **progettazione logica**: trasforma il diagramma ER in uno schema logico, efficiente rispetto alle strutture del sistema di gestione che si intende utilizzare.

Input: diagramma ER

Output: schema logico delle relazioni rappresentante tramite tabelle.

- **progettazione fisica**: implementa lo schema logico definendone gli aspetti fisici di memorizzazione e rappresentazione in memoria di massa.

Input: tabelle della progettazione logica

Output: implementazione in memoria di massa delle tabelle.



MONDADORI EDUCATION