

Indice

1	Introduzione	2
1.1	Sistemi informativi, informazioni e dati	2
1.2	Basi di dati e sistemi di gestione di basi di dati	3
1.3	Linguaggi per basi di dati	4
1.4	Modelli dei dati	5

1 Introduzione

1.1 Sistemi informativi, informazioni e dati

Ogni organizzazione è dotata di un *sistema informativo*, che organizza e gestisce le informazioni necessarie per perseguire gli scopi dell'organizzazione stessa. Per indicare la **porzione automatizzata del sistema informativo** viene di solito utilizzato il termine *sistema informatico*.

Nei sistemi informatici le informazioni vengono rappresentate per mezzo di *dati*, che hanno bisogno di essere interpretati per fornire informazioni.

Esiste una differenza sottile tra dato e informazioni. Solitamente i primi, se presi da soli, non hanno significato, ma, una volta interpretati e correlati opportunamente, essi forniscono informazioni, che consentono di arricchire la conoscenza:

Informazione: notizia, dato o elemento che consente di avere conoscenza più o meno esatta di fatti, situazioni, modi di essere;

Dato: ciò che è immediatamente presente alla conoscenza, prima di ogni elaborazione. In informatica, sono elementi di informazione costituiti da simboli che devono essere elaborati.

[ESAME] Definizione base di dati: Una *base di dati* è una collezione di dati, utilizzati per rappresentare con tecnologia informatica le informazioni di interesse per un sistema informativo.

1.2 Basi di dati e sistemi di gestione di basi di dati

Inizialmente, venne adottato un “approccio convenzionale” alla gestione dei dati. Esso **sfruttava** la presenza di archivi o **file per memorizzare e per ricercare dati**. Tuttavia, i metodi di accesso e condivisione erano semplici e banali. Infatti, erano presenti numerosi **problemi**:

- ✗ **Accesso sequenziale**: la scarsa efficienza nell’accesso ai dati su file rendeva lento l’accesso a tali informazioni;
- ✗ **Ridondanza**: i dati di interesse per più programmi sono replicati tante volte quanti sono i programmi che li utilizzano, con evidente ridondanza e possibilità di incoerenza;
- ✗ **Inconsistenza**: una diretta conseguenza della ridondanza. Con la presenza di più copie di un determinato dato, l’eventuale cambiamento di uno solo potrebbe portare a questo effetto;
- ✗ **Progettazione duplicata**: per ogni programma viene replicata la progettazione.

La **soluzione** è arrivata negli anni ’80 con l’avvento delle **basi di dati**. Quest’ultime gestiscono in modo integrato e flessibile le informazioni di interesse per diversi soggetti.

[ESAME] Definizione DBMS: Un *sistema di gestione di basi di dati* (in inglese *Data Base Management System*, **DBMS**) è un sistema software in grado di gestire collezioni di dati che siano:

- ✓ **Grandi**;
- ✓ **Condivise**;
- ✓ **Persistenti**.

assicurando allo stesso tempo:

- ★ **Affidabilità**;
- ★ **Privatezza**;
- ★ **Accesso efficiente**.

Il **vantaggio** di utilizzare un DBMS è stato evidenziato nella definizione. Quindi:

- ✓ **Maggiore astrazione** poiché le sue funzioni estendono il *file system*, fornendo la possibilità di accesso condiviso agli stessi dati da parte di più utenti e applicazioni;
- ✓ **Maggiore efficacia** poiché le operazioni di accesso ai dati si basano su un linguaggio di interrogazione.

1.3 Linguaggi per basi di dati

Su un DBMS è possibile specificare operazioni di vario tipo, ma principalmente si distinguono in due categorie:

- **Linguaggi di definizione dei dati** (*Data Definition Language*, abbreviato con **DDL**) utilizzati per definire gli schemi logici, esterni e fisici e le autorizzazioni per l'accesso;
- **Linguaggi di manipolazione dei dati** (*Data Manipulation Language*, abbreviato con **DML**) utilizzati per l'interrogazione e l'aggiornamento delle istanze di basi di dati:
 - *Linguaggio di interrogazione*, estrae informazioni da una base di dati (SQL, algebra relazionale);
 - *Linguaggio di manipolazione*, popola la base di dati, modifica il suo contenuto con aggiunte, cancellazioni e variazioni sui dati (SQL).

1.4 Modelli dei dati

Definizione modello dei dati: Un *modello dei dati* è un insieme di concetti utilizzati per organizzare i dati di interesse e descriverne la struttura in modo che essa risulti comprensibile ad un elaboratore.

Ogni modello dei dati fornisce **meccanismi di strutturazione**, analoghi ai **costruttori** di tipo dei linguaggi di programmazione (es: Java), che permettono di definire nuovi tipi sulla base di tipi predefiniti (elementari) e costruttori di tipo.

Quindi, i *costruttori* consentono di:

- ☞ *Definire* le **strutture dati che conterranno le informazioni** della base di dati;
- ☞ *Specificare* le **proprietà che dovranno soddisfare le istanze** di informazione che saranno contenuto nelle strutture dati.

Definizione schemi e istanze: È molto importante distinguere gli **schemi** e le **istanze** dal concetto di modello dei dati:

- **Schema:** parte invariante nel tempo, è costituita dalle caratteristiche dei dati. In altre parole, è la descrizione della struttura e delle proprietà di una specifica base di dati fatta utilizzando i costrutti del modello dei dati;
- **Istanza o stato:** parte variabile nel tempo, è costituita dai valori effettivi. Quest'ultimi, in un certo istante, popolano le strutture dati della base di dati.