

INFORMATICA

BASE DI DATI

Nella pratica quotidiana spesso si usano in modo interscambiabile i termini «dato» e «informazione»: in effetti non è immediato distinguerli. Un **dato** è la misura di un fenomeno che siamo interessati a osservare. L'**informazione** è ciò che si ottiene dall'elaborazione di un insieme di dati e che accresce lo stato di conoscenza relativo a un fenomeno. Da un punto di vista puramente intuitivo è possibile affermare che maggiore è la quantità di dati di cui si dispone rispetto a un fenomeno da analizzare migliore sarà la qualità dell'informazione. Una base di dati (in inglese database) è una collezione di dati organizzati immagazzinata e accessibile per via elettronica. Piccole banche dati possono essere immagazzinate su un file system, mentre banche dati più grandi sono ospitati su computer cluster o su cloud storage.

Un **sistema di gestione delle banche dati**, noto anche come database management system (DBMS), è il software che interagisce con gli utenti finali, le applicazioni e la banca dati stessa per prendere e analizzare i dati. Inoltre, il software di gestione comprende le funzionalità di base necessarie per amministrare la banca dati. L'insieme della banca dati, il sistema di gestione e le applicazioni associate può essere chiamato un database system. Le basi di dati relazionali divennero predominanti negli anni '80. Queste modellano i dati come righe e colonne in una serie di tabelle, e la stragrande maggioranza utilizza SQL come linguaggio per scrivere e interrogare i dati. Negli anni 2000, presero piede anche modelli non relazionali, collettivamente chiamati NoSQL, perché usano linguaggi di interrogazione diversi.

Banca Dati può anche indicare contemporaneamente: l'archivio a livello fisico (hardware) cioè il sistema con i supporti di memorizzazione (storage, es. dischi rigidi) che contengono i dati stessi e il processore per l'elaborazione di questi (database server); l'archivio a livello logico, cioè i dati strutturati, e la parte software, cioè il database management system (DBMS). Nelle banche dati più moderne, ovvero quelle basate sul modello relazionale, i dati vengono suddivisi in apposite tabelle per argomenti e questi argomenti vengono suddivisi per categorie. Questa suddivisione rendono le basi di dati notevolmente più efficienti, almeno per la gestione di dati complessi.

I principali punti che qualificano il ricorso all'approccio fondato su DBMS sono i seguenti:

1. **Integrazione:** una base di dati è un insieme integrato di dati strutturati e permanenti memorizzati senza ridondanze superflue e organizzati in modo tale da poter essere usati da applicazioni diverse senza dipendere da alcuna di esse.
2. **Indipendenza logica:** i dati sono definiti indipendentemente dalle procedure che li

gestiscono, in questo modo la struttura logica della base di dati può essere ampliata senza la necessità di modificare i programmi applicativi.

3. **Indipendenza fisica:** descrive la struttura dei dati astraendo da quella che è la loro implementazione fisica (organizzazione della memorizzazione, modalità di accesso, ecc.) in modo tale che si possa modificare quest'ultima senza modificare la struttura logica dei dati ed di conseguenza i programmi applicativi.
4. **Integrità:** è il sistema di gestione delle basi di dati e non le procedure di gestione a dover prevedere meccanismi per controllare che i dati inseriti o modificati soddisfino ai vincoli di integrità specificati.

È possibile classificare l'utenza di un sistema di gestione della basi di dati nel seguente modo:

1. **Programmatori di applicazioni**, ossia utenti professionali che hanno il compito di sviluppare applicazioni software.
2. **Utenti finali**, cioè utenti non professionali che interagiscono con la base di dati esclusivamente mediante programmi applicativi (per es: addetti a sportelli postali o bancari, magazzinieri, utenti web, ecc.) senza conoscere la struttura dei dati con cui interagiscono.
3. **Utenti avanzati**, ovvero utenti che conoscono la struttura dei dati e sono in grado di operare attività di indagine utilizzando un linguaggio di interrogazione del database senza alterarlo.
4. **Utenti amministratori (DBA, *Data Base Administrator*)**, cioè utenti professionali a cui è demandata la manutenzione della base di dati nel tempo.