

Le Basi di dati: generalità



1 Cosa è una base di dati

- In ogni modello di organizzazione della vita dell'uomo vengono trattate **informazioni**
- Una volta individuate e raccolte devono essere necessariamente memorizzate in modo che si possano:
 - Recuperare
 - Aggiungerne di nuove
 - Modificare
 - Cancellare
- Tale necessità è nata ancor prima dell'avvento dell'informatica

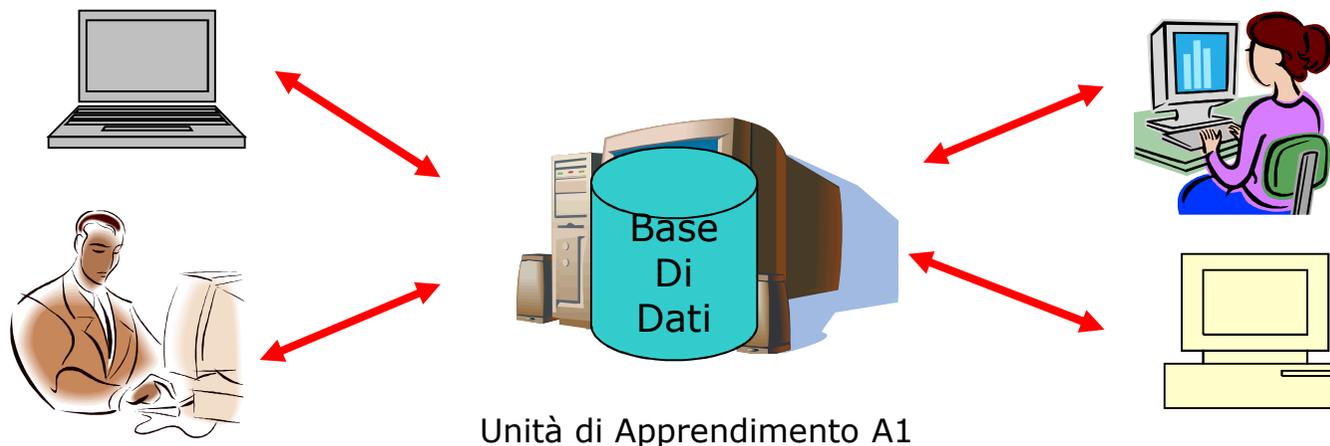


1 Cosa è una base di dati

- La **Teoria della Base di Dati** studia come organizzare al meglio grandi quantità di informazioni, per gestirle in modo:
 - **Semplice** → fruibili da applicazioni e/o da utenti diversi
 - **Efficiente** → ottimizzato in tempo, in spazio
 - **Efficace** → rappresentative della realtà di interesse, rendere produttive le attività richieste
 - **Sicure** → i soggetti e/o applicazioni che vi operano devono essere appositamente autorizzati

1 Cosa è una base di dati

- Definizione:
“Base di Dati (DataBase) è una raccolta di dati **logicamente correlati** e progettati per essere **fruiti in maniera ottimizzata** da differenti applicazioni e/o utenti”





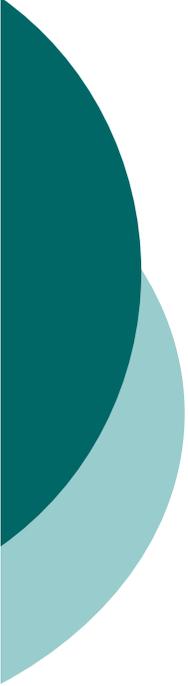
1 Cosa è una base di dati

- Una base di dati per essere definita tale deve essere:
 1. **Sicura** → contro eventi accidentali o interventi non autorizzati
 2. **Integra** → gli interventi autorizzati non devono provocare inconsistenza dei dati (dati irreali)
 3. **Consistente** → I dati devono sempre essere significativi e quindi utilizzabili (i dati devono essere sempre consistenti anche in seguito ad accesso concorrente)
 4. **Condivisibili** → Più utenti e/o applicazioni diverse devono poter accedere ai dati
 5. **Persistente** → Tempo di vita delle informazioni non limitato a quello di esecuzione delle applicazioni



2 Sistemi informativi e Sistemi informatici

- Le risorse di una azienda:
 - Persone
 - Denaro
 - Materiali
 - **Sistema Informativo**
- Funzioni di un Sistema Informativo:
 - Raccolta, acquisizione delle informazioni
 - Archiviazione, conservazione delle informazioni
 - Elaborazione delle informazioni
 - Distribuzione, scambio di informazioni

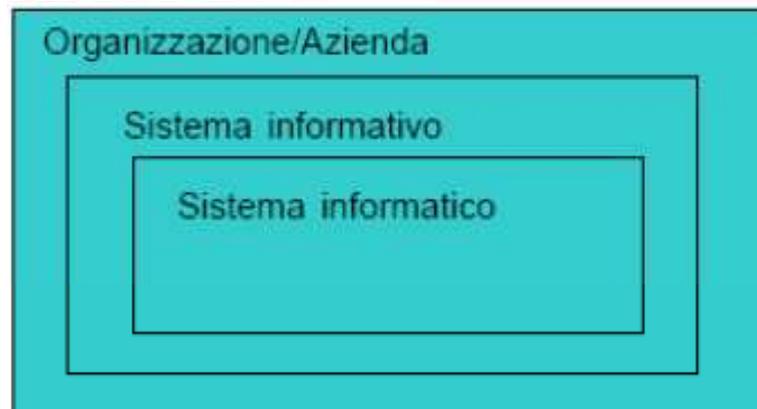


2 Sistemi informativi e Sistemi informatici

- Il concetto di “**sistema informativo**” è indipendente da qualsiasi automatizzazione
 - Esistono organizzazioni che gestiscono esclusivamente informazioni (es: servizi anagrafici e banche) e che operano da secoli
 - Vi sono infine organizzazioni che implicitamente da sempre gestiscono informazioni (Elenco dei clienti di un'azienda, Raccolta dei biglietti da visita, Gestione dei biglietti di una compagnia aerea ecc..)

2 Sistemi informativi e Sistemi informatici

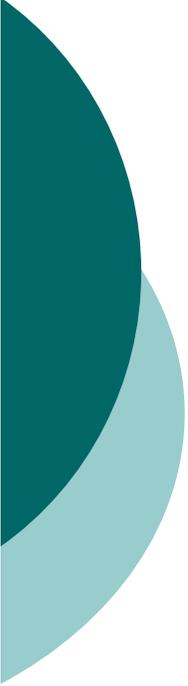
- **“Sistema Informatico”** è l’insieme degli strumenti informatici utilizzati per il **trattamento automatico delle informazioni, al fine di agevolare le funzioni del sistema informativo.**





3 Dati e Informazioni: schemi e istanze

- **Dato:**
Letteralmente fatto. Codifica dei fatti ritenuti importanti nell'ambito di una organizzazione.
- **Informazione:**
E' l'incremento di conoscenze che può essere acquisita (inferita) dai dati

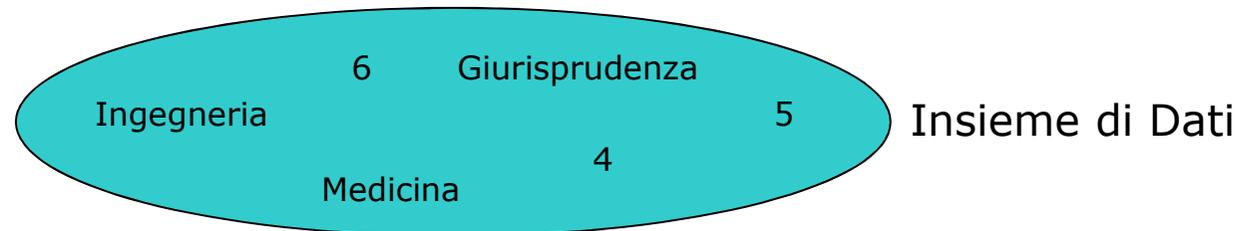


3 Dati e Informazioni: schemi e istanze

- I dati sono fatti elementari che devono essere **interpretati** per fornire informazioni
- Esempio:
 - Si consideri un dato rappresentato dal numero 5: tale dato non fornisce alcuna informazione
 - Viceversa, Interpretando 5 come il numero degli anni del corso di Laurea di Giurisprudenza questo dato fornisce informazione
- I dati sono fruibili (cioè danno informazioni) se viene fornita la **“Chiave di Interpretazione dei dati”**

3 Dati e Informazioni: schemi e istanze

- Una possibile Chiave di interpretazione di dati è lo **“Schema tabellare o Significato Intensionale”**
- Esempio:



Università Italiana	
Corso di Laurea	Anni di Corso
Medicina	6
Giurisprudenza	4
Ingegneria	5

Schema o Intensione

3 Dati e Informazioni: schemi e istanze

- Terminologia:

- **Schema o Intensione:**

Chiave di interpretazione dei dati (significato semantico che si attribuisce ai dati). Dallo schema scaturisce l'interpretazione dei dati. Non varia nel tempo

Corso di Laurea	Anni di Corso

3 Dati e Informazioni: schemi e istanze

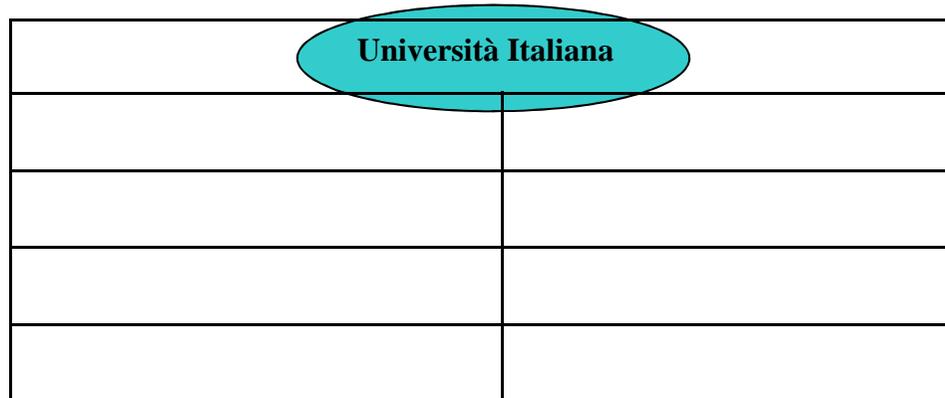
- **Attributo:**

Uno schema o istensione è costituito da più attributi che raggruppano caratteristiche comuni ad insieme di dati

	Anni di Corso

3 Dati e Informazioni: schemi e istanze

- **Categoria:**
Un gruppo di dati avente la stessa chiave di interpretazione, cioè lo stesso schema



A diagram illustrating a data table. The table has a header row and four data rows. The header cell contains the text "Università Italiana" and is highlighted with a light blue oval. The table is structured as follows:

Università Italiana	

3 Dati e Informazioni: schemi e istanze

- **Istanza di uno schema o Estensione:**
Insieme dei valori assunti da uno schema in un certo istante di tempo

Medicina	6
Giurisprudenza	4
Ingegneria	5

3 Dati e Informazioni: schemi e istanze

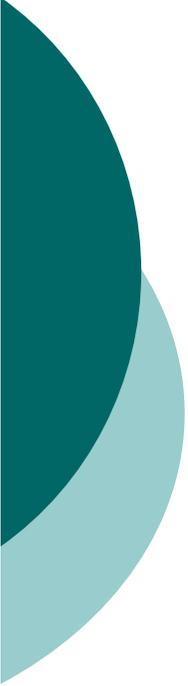
- **Istanza:**
Valori assunti da un **singola entità** della
Categoria

Ingegneria	5

3 Dati e Informazioni: schemi e istanze

- **Valore:**
Singolo dato all'interno di uno schema

Ingegneria	5



3 Dati e Informazioni: schemi e istanze

- **Occorrenza di una Base di Dati:**
Insieme delle Istanze **di tutte le categorie** (tabelle) presenti all'interno di un database in un **determinato istante di tempo**



4 Il Modello di dati

- Un modello di dati è un'insieme di concetti e di costrutti per organizzare i dati di interesse e descriverne la struttura e la dinamica (associazioni e vincoli che devono rispettare)
- I modelli astratti si distinguono in due categorie:
 - **Modello Concettuale** → permette di **rappresentare i concetti** (dati) in modo indipendente da ogni sistema. Rappresenta una descrizione del mondo reale attraverso una opportuna simbologia. Si mettono in luce i concetti che legano i vari dati **a prescindere da come verranno eventualmente rappresentati su di un calcolatore.** Sono modelli preliminari di progettazione. Uno dei più noti e usati: **Modello ER (Entità Relazione)**



4 Il Modello di dati

- **Modello Logico** → Pur essendo ancora una rappresentazione astratta dei dati riflette già di una particolare organizzazione. Descrive l'organizzazione dei dati con l'influenza del tipo di implementazione fisica definitiva.

Vi sono vari modelli logici:

- Modello Gerarchico
- Modello Reticolare
- **Modello Relazionale** → Approfondiremo questo modello
Inventato nel 1976 dal ricercatore Peter P. Chen basato sul concetto di insieme di Tabelle costituite da record. Le relazioni fra le informazioni derivano dalla corrispondenza di alcuni campi fra tabelle diverse.
- Modello a Oggetti → moderno e molto promettente

5 La Progettazione di una base di dati

- La progettazione di un Data Base, analogamente a quanto avviene per la progettazione di un software prevede due momenti importanti:

- **Analisi Astratta** →
Indipendente dalle tecniche di implementazioni che si useranno (Progettazione **concettuale** e Progettazione **Logica**)

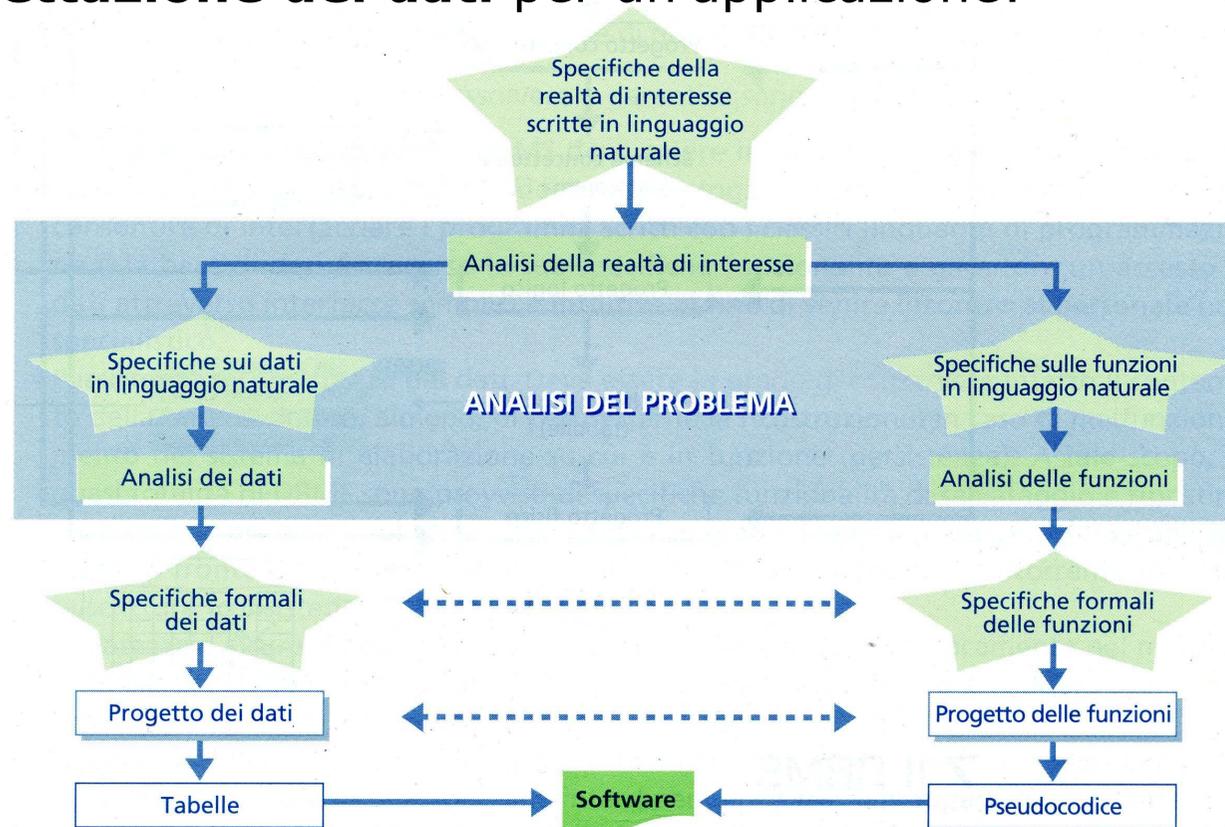
- **Progetto dei dati** →
Determinazione di come il programma dovrà fare quanto stabilito (Progettazione **Fisica**)

Unità di Apprendimento A1



5 La Progettazione di una base di dati

- Possiamo immaginare una analogia nelle fasi di **progettazione di funzioni** di un'applicazione e nelle fasi di **progettazione dei dati** per un'applicazione:



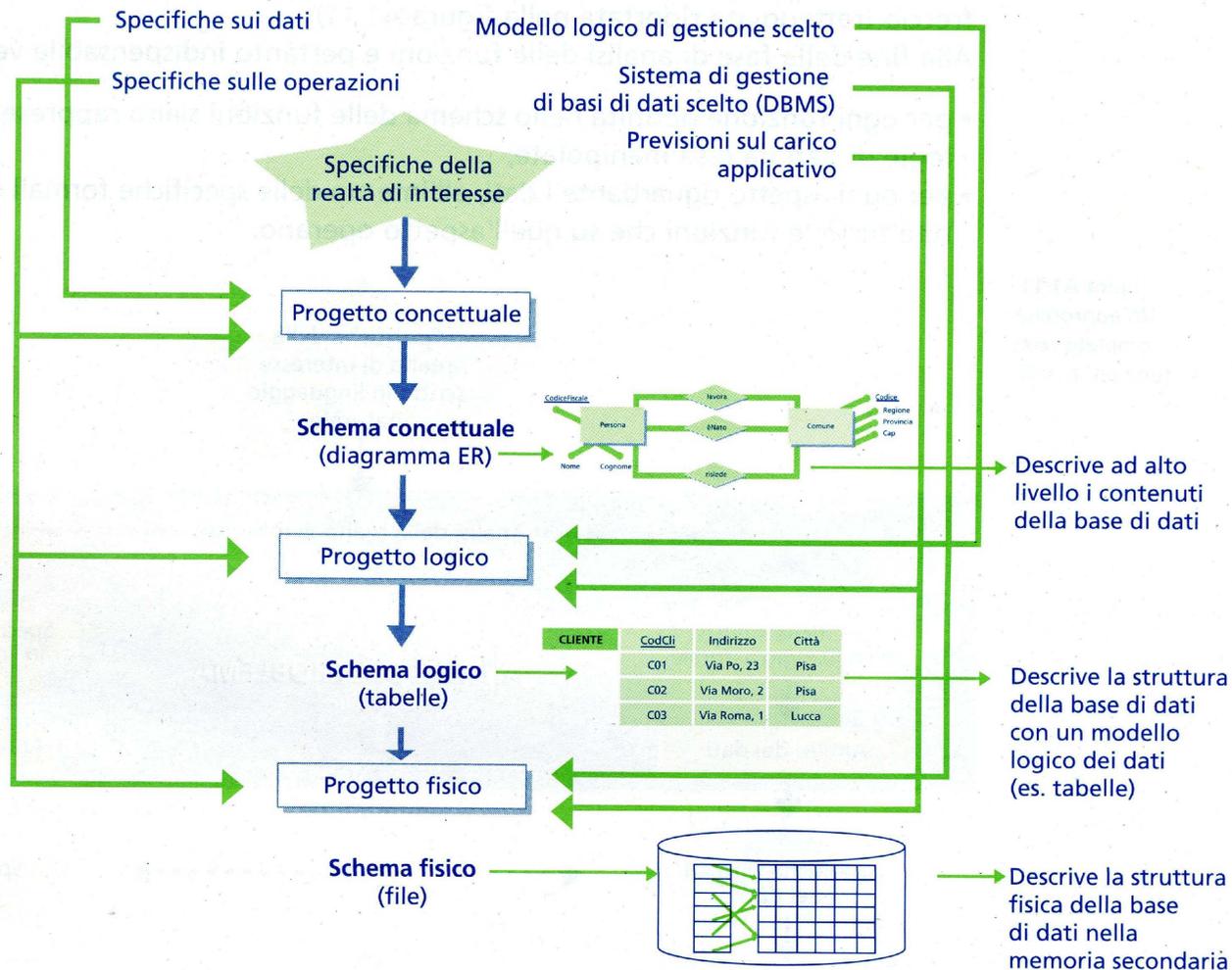


6 Le fasi della Progettazione

- Le attività da svolgere sono sostanzialmente tre:
 1. **Progettazione Concettuale** → Costruire e definire una rappresentazione corretta e completa della realtà.
 - Input → Documento delle specifiche formali
 - Output → Schema concettuale (**Diagramma ER**)
 2. **Progettazione Logica** → Rappresentare in modo efficiente la realtà di interesse, in funzione delle strutture che si intende utilizzare
 - Input → Diagramma ER
 - Output → Schema logico riassuntivo attraverso relazioni rappresentate tramite Tabelle (**Schema Relazionale**)
 3. **Progettazione Fisica** → Implementare lo schema logico all'interno di un DataBase specifico
 - Input → (Schema Relazionale) Tabelle della fase di progettazione Logica
 - Output → **Implementazione** in memoria di massa delle tabelle

6 Le fasi della Progettazione

○ Schema Riassuntivo delle fasi di Progettazione





7 II DBMS

- Prima dell'introduzione dei database **gli archivi** avevano svantaggi notevoli:
 - **Ridondanza** → dati ripetuti in più file
 - **Costi aggiuntivi** → per generare e memorizzare più copie dei singoli dati in più file
 - **Complessità della gestione** → ricerche complesse per la ricerca di dati
 - **Livello di Aggiornamento** → Impossibilità oggettiva di avere lo stesso livello di aggiornamento dei dati collocati in più file che chiaramente conduce a:
 - **Incongruenza** dei dati → attraverso un'interrogazione si potevano ricevere versioni di dati non aggiornati ed incongruente con i restanti dati dell'archivio
 - **Inconsistenza** dei dati → nella stessa serie di dati vi potevano essere alcuni aggiornati ed altri no, pertanto le interrogazioni sui dati potevano condurre ad informazioni inconsistenti (non reali e quindi inutilizzabili)
 - Impossibilità di gestire la **concorrenza** → più utenti/programmi che tentavano di accedere agli stessi dati per modifiche potevano ritrovarsi con dati non fisicamente modificati



7 II DBMS

- I DBMS nascono per ovviare alle difficoltà emerse nella gestione attraverso archivi:
- Un DBMS (Data Base Management System) è un'insieme di strumenti software che sulla base delle specifiche del programmatore, provvede alla **generazione** dello schema progettato, alla **ricerca** e all'**aggiornamento** dei dati.
Nel caso specifico di un database relazionale, di cui parleremo, si parla di RDBMS (Relational Data Management System).
- Il DBMS rappresenta quindi un'interfaccia tra gli sviluppatori, gli utenti del database e il sistema di elaborazione.
- Il DBMS è un sistema attivo, mentre la base di dati è l'oggetto passivo su cui opera il DBMS.



7 II DBMS

- Le 6 funzioni di un DBMS sono:
 1. **Gestione della base di dati** → permettere la creazione, l'inserimento, l'aggiornamento e l'interrogazione della base di dati.
Deve inoltre fornire un'interfaccia semplice ed intuitiva per l'utente/programmatore.
 2. **Persistenza e consistenza** dei dati → conservare intatto il contenuto della base di dati (o almeno permettere la ricostruzione automatica, backup e restore).
Gestire l'accesso concorrente per preservare la consistenza dei dati
 3. **Privatezza e sicurezza** dei dati → deve gestire accessi autorizzati con diritti prefissati



7 II DBMS

- Le 6 funzioni di un DBMS sono:
 4. **Integrità** dei dati → In merito ai valori che i dati possono assumere anche in relazione alle interdipendenze fra tabelle diverse o vincoli imposti dal programmatore.
 5. **Supporto alle transazioni** → Per transazione si intende un'operazione su di un database che può concludersi positivamente o negativamente ed in tal caso il DBMS deve preservare per lo meno lo stato precedente dei dati.
 6. **Gestione del dizionario dei dati** → informazioni che descrivono gli oggetti della base di dati attraverso **metadati** (tali informazioni riguardano i nomi delle tabelle, dei campi, le associazioni ecc..)



7 II DBMS

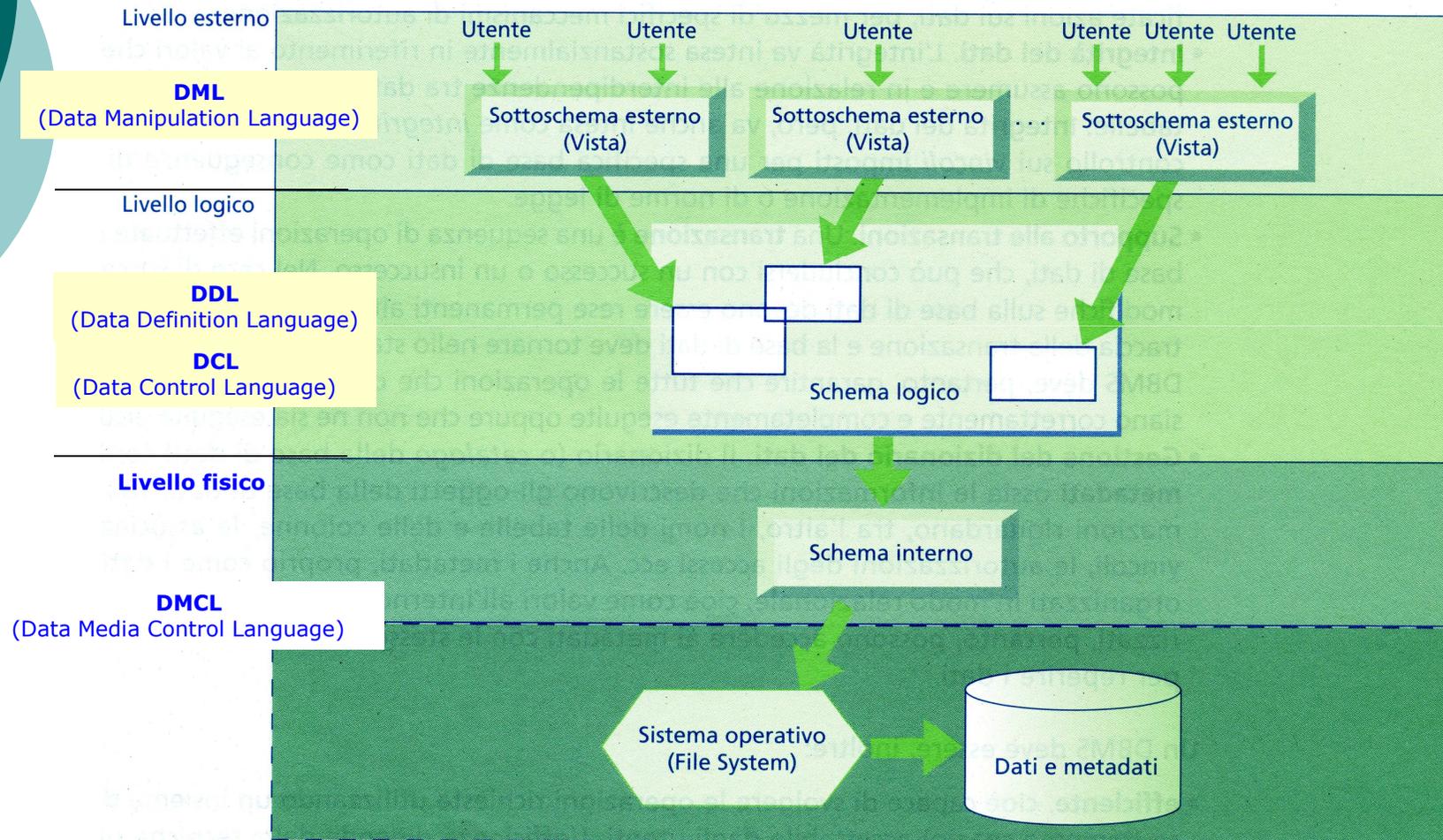
- Un DBMS deve essere inoltre:
 - **Efficiente** → ottimizzato in tempo, in spazio.
L'efficienza risente:
 - delle tecniche utilizzate nell'implementazione del DBMS
 - della bontà della base di dati progettata dal programmatore.
 - del dimensionamento del sistema informatico complessivo su cui il DBMS è installato
 - **Efficace** → deve essere in grado di rendere produttive e semplici le attività richieste dagli utenti



8 Architettura di un DBMS

- L'architettura di un software DBMS è fondamentalmente costituita da tre livelli:
 1. **Livello esterno** → Livello nel quale operano gli utenti attraverso la propria personale "vista" esterna del database. (predisposta dal **DBA** Data Base Administrator)
Il linguaggio **DML** (Data Manipulation Language) consente agli utenti di **interrogare, inserire, cancellare e modificare** la base di dati.
 2. **Livello logico** → Livello all'interno del quale viene definito l'intero schema dei dati (schema relazionale) indipendente dalla realizzazione fisica.
Il linguaggio **DDL** (Data Definition Language) viene utilizzato per definire le caratteristiche e le corrispondenze delle varie categorie presenti.
Il linguaggio **DCL** (Data Control Language) permette operazioni relative al controllo degli accessi
 3. **Livello Fisico** → Implementazione del database logico.
Rappresenta la forma in cui il database viene memorizzato sui supporti fisici.
Il linguaggio **DMCL** (Data Media Control Language) utilizzato per operare a questo livello.

8 Architettura di un DBMS





8 Architettura di un DBMS

- Lo standard SQL che studieremo si colloca all'interno dei linguaggi per la gestione dei DBMS assolvendo alle funzioni:
 - **DDL** (Data Definition Language) → istruzioni che ci permettono di definire la struttura del data-base (creare tabelle, impostare vincoli, creare viste, ecc..)
 - **DML** (Data Manipulation Language) → istruzioni che ci permettono di manipolare i dati collocati nelle varie tabelle (inserire dati, cancellare dati, modifiche delle righe delle tabelle, interrogare la base di dati, ecc..)
 - **DCL** (Data Control Language) → istruzioni che ci permettono di controllare gli accessi ed i diritti (permessi) dei vari utenti