

# 1 L'organizzazione dei dati negli archivi

**EUCIP**  
item  
B.1.2.3

Gli archivi assolvono alla funzione di conservare nel tempo dati e informazioni perché potranno essere utili in momenti successivi: questo vale per dati di tipo personale, come quelli contenuti in una normale rubrica di contatti, ma anche per i documenti utili alla vita di uno Stato, di un'azienda o di un ente.

→ Che cos'è un archivio?

In generale, un **archivio** è un insieme organizzato di informazioni caratterizzate da alcune proprietà fondamentali:

- tra esse esiste un **nesso logico** (cioè sono inerenti a un medesimo argomento);
- sono rappresentate secondo un **formato** che ne rende possibile l'interpretazione;
- sono registrate su **supporti** che ne permettono la rilettura nel tempo;
- sono **organizzate** in modo che siano facilmente consultabili.

225	Bianchi	Giovanni	...
236	Carminati	Alfredo	...
325	Gialli	Giovanna	...
425	Verdi	Enrico	...
456	Negri	Francesca	...
538	Viola	Annalisa	...
585	Rossi	Giuliano	...
624	Magnani	Pietro	...
788	Bruni	Alessandra	...

file con le informazioni anagrafiche degli studenti

il campo con il nome degli studenti presenti in anagrafica

il record con l'insieme dei campi che descrivono la studentessa **Bruni Alessandra**

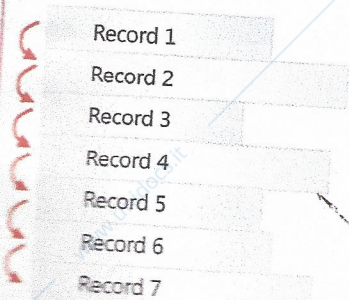
Un archivio (o **file**) è una collezione di **record** cioè di informazioni logicamente omogenee che descrivono i singoli elementi di una realtà considerata.

Ogni **record** è composto da un insieme di **campi** che contengono i valori assunti dalle caratteristiche scelte per descrivere la realtà.

## I tipi di organizzazione

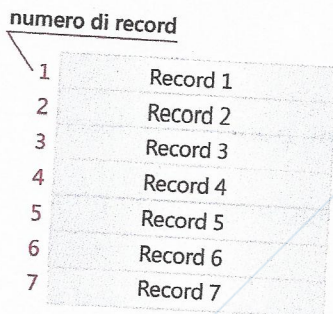
L'organizzazione dei dati negli archivi deve essere fatta in modo da ottimizzare l'uso dello spazio disponibile sul supporto e minimizzare i tempi di ritrovamento dei dati memorizzati.

### Accesso sequenziale



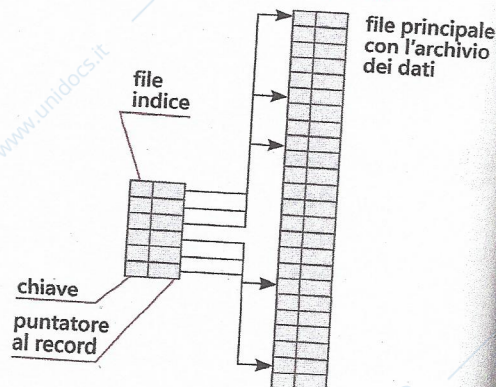
Registrazione dei record uno di seguito all'altro, in modo sequenziale.

### Ad accesso diretto (o relativo o random)



I record sono identificati dalla **posizione** che occupano nel file. Per accedere a un record occorre specificare, con un valore numerico, la sua posizione.

### Indexed sequential



Ciascun record è identificato univocamente da un campo **chiave** (matricola, codice ecc.). Accanto all'archivio con i record, è gestita una tabella ordinata delle chiavi (**file indice**) con il puntatore al record corrispondente.

→ In quali modi è possibile organizzare gli archivi di dati?

I sistemi informativi costruiti con **archivi non integrati** consistevano di diverse applicazioni indipendenti, ognuna delle quali operava su propri archivi (approccio basato su file *file-based*) e presentavano diversi problemi.

- **Ridondanza**: duplicazione degli stessi dati in archivi diversi.
- **Incongruenza**: valori diversi per lo stesso dato in archivi diversi.
- **Inconsistenza**: i dati sono inutilizzabili perché non si sa quale sia il valore corretto.

**File e memorie di massa**

# Le basi di dati

Per risolvere i problemi e superare i limiti dell'organizzazione tradizionale degli archivi, gestiti in modo non integrato, si sono affermate le tecniche di gestione delle **basi di dati**.

database 

Il **database** è una collezione di archivi logicamente correlati e condivisi, che ha lo scopo di soddisfare i fabbisogni informativi di una specifica organizzazione.

Il sistema software che gestisce i dati, congiuntamente con la loro descrizione, si chiama **DBMS (DataBase Management System)**: esso amministra e controlla gli accessi, offrendo misure di **sicurezza** e di **integrità** dei dati, per garantire la consistenza del database.

## La consistenza del database

**SICUREZZA**

impedire che il database venga danneggiato da interventi accidentali o non autorizzati

**INTEGRITÀ**

garantire che le operazioni effettuate sul database da utenti autorizzati non provochino una perdita di consistenza ai dati

→ Che cos'è il DBMS e che cosa garantisce?

I servizi di un DBMS sono attivati con appositi comandi, che possono essere classificati nelle seguenti categorie di linguaggi:

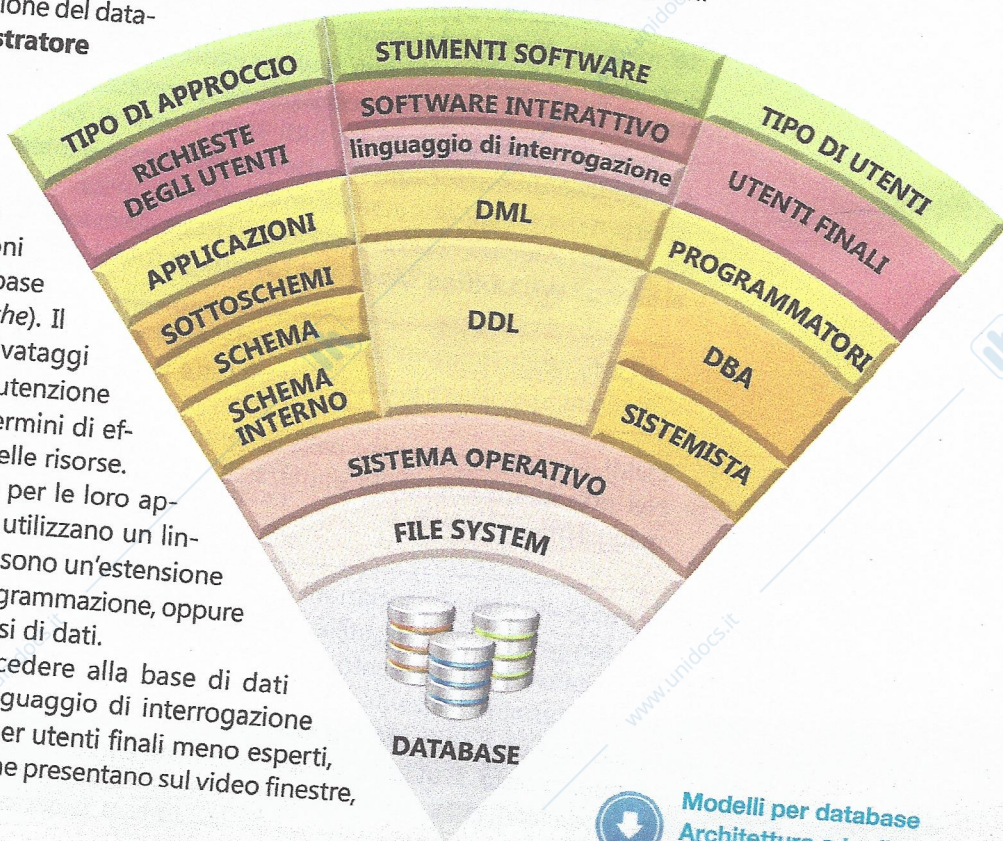
- **DDL (Data Definition Language)**: linguaggio per la descrizione dei dati e dei controlli sulla sicurezza, per la creazione o la modifica della struttura logica e fisica del database.
- **DML (Data Manipulation Language)**: linguaggio per il trattamento (inserimenti, modifiche e cancellazioni) dei dati contenuti nel database.
- **QL (Query Language)**: linguaggio per le interrogazioni alla base di dati, che consente il ritrovamento dei dati che interessano, sulla base dei criteri di ricerca richiesti dall'utente.

Un database viene utilizzato da diverse **figure professionali**, per funzioni e applicazioni diverse, usando differenti strumenti software.

La responsabilità della gestione del database è affidata all'**Amministratore della Base di Dati (DBA, DataBase Administrator)**,

che crea la base di dati e la sua struttura logica (*schemi e sottoschemi*), gestisce gli utenti con le autorizzazioni degli accessi all'intero database o a parti di esso (*viste logiche*). Il DBA provvede anche ai salvataggi periodici dei dati e alla manutenzione del database nel tempo, in termini di efficienza e di ottimizzazione delle risorse.

I **programmatori**, che usano per le loro applicazioni i dati del database, utilizzano un linguaggio **DML**, o comandi che sono un'estensione dei tradizionali linguaggi di programmazione, oppure un linguaggio specifico per basi di dati. Gli **utenti finali** possono accedere alla base di dati attraverso i comandi di un linguaggio di interrogazione (**QL, query language**) oppure, per utenti finali meno esperti, attraverso interfacce grafiche, che presentano sul video finestre, menu e icone.



Modelli per database  
Architettura a tre livelli  
e indipendenza dei dati