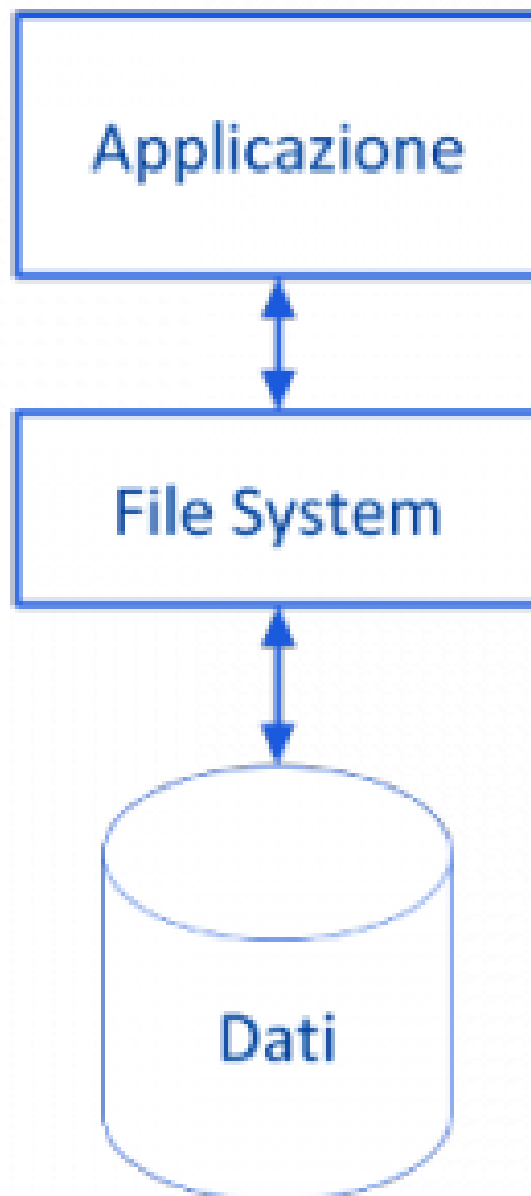


# ORGANIZZAZIONE DEGLI ARCHIVI E BASI DI DATI

1. Archivi e loro organizzazione
2. Limiti della organizzazione tradizionale degli archivi
3. Necessità dei database

Archivi tradizionali



Un file (archivio) è una raccolta di record logici. La gestione dei file è un servizio reso dal Sistema Operativo.

Ogni record logico contiene una sequenza di bit ed è suddiviso in campi.

Ogni campo contiene una differente informazione.

### Esempio:

<b>ROSSINI GIORGIO</b>	<b>Via Nappi 6 - Milano</b>	<b>2372445</b>
<i>campo1: nome</i>	<i>campo2: indirizzo</i>	<i>campo3: telefono</i>

Su questi file esistono meccanismi per l'inserimento, la cancellazione, la modifica e la ricerca dei dati.

### Organizzazioni dei file:

- **Sequenziale:**
  - I record sono letti o scritti uno dopo l'altro. Inserimento, cancellazione e modifica di un file sequenziale presentano problemi, in quanto richiedono di norma lo spostamento di tutti i record che seguono il punto di intervento.
- **Ad accesso diretto:**
  - L'allocazione dei record in memoria di massa è definita da un algoritmo che genera un indirizzo a partire dal valore assunto da un campo del record detto chiave.
  - Il vantaggio delle strutture ad accesso diretto è la rapidità della ricerca: in assenza di conflitti ogni ricerca richiede una sola operazione di ingresso/uscita.
- **A indice:**
  - In queste strutture, come in quelle ad accesso diretto, è il valore di un campo chiave che determina il record a cui accedere.
  - La corrispondenza fra la chiave e l'indirizzo è mantenuta esplicitamente mediante un file ausiliario detto file indice.
  - Su uno stesso archivio è possibile avere diversi file indice relativi a chiavi diverse.

# Archivi tradizionali e basi di dati

- Svantaggi:
  - Ridondanza
  - Inconsistenza
  - Privatezza
  - Condivisione
  - Accesso concorrente



- **Database Management System** è un sistema software progettato per la creazione e la manipolazione efficiente di basi di dati solitamente da parte di più utenti.



Le **basi di dati** sono sorte alla fine degli anni '60. Prima i sistemi informativi si basavano sull'uso di files separati.

I programmi applicativi accedevano ai files individualmente, utilizzando procedure del Sistema Operativo.

L'accesso ai files tradizionali comportava una serie di problemi:

1. **Ridondanza:** informazioni ripetute
2. **Incongruenza:** le modifiche non sempre vengono apportate in tutti gli archivi dove compaiono le stesse informazioni
3. **Inconsistenza:** se ad esempio uno stesso articolo si ritrova con marche diverse quale sarà quella giusta?
4. **Dipendenza dei programmi dai dati:** se cambia il tracciato record o la cartella di un archivio devo cambiare l'applicativo
5. **Difficoltà nel gestire l'integrità dei dati:** va scritto codice ad hoc nell'applicativo
6. **Difficoltà nel gestire la concorrenza:** in un file condiviso se due utenti tentano la modifica, si possono generare informazioni non corrette
7. **Limitata sicurezza e privatezza:** non tutti gli utenti hanno stessi permessi sui dati e la privatezza dei dati è limitata perché gestita solo dal sistema operativo
8. Scarsa protezione dei dati da guasti accidentali

**Questi inconvenienti si possono superare se tutti i dati vengono organizzati in una base di dati e gestiti da un DBMS (Data Base Management System).**

**Confronto tra sistemi di archiviazione tradizionali e DBMS**

<b>ARCHIVI TRADIZIONALI</b>	<b>DBMS</b>	
<i>Sistemi di archiviazione</i>	<i>Sistemi per basi di dati</i>	<i>Benefici</i>
In generale ogni applicazione ha propri archivi, con dati parzialmente duplicati.	I dati sono organizzati per poter essere usati da diverse applicazioni.	Integrazione dei dati.
Gli archivi sono modificati in tempi diversi e copie dello stesso dato possono risultare diverse.	Gli aggiornamenti vengono visti subito da tutte le applicazioni, a causa della condivisione dei dati.	Consistenza dei dati.
Ogni applicazione deve garantire l'integrità dei dati.	I DBMS offrono meccanismi per il controllo centralizzato dell'integrità dei dati.	Integrità dei dati.
Cambiamenti nella definizione dei record devono essere riportati in ogni applicazione.	Con la descrizione centralizzata dei dati solo le viste logiche devono essere modificate.	Indipendenza logica.
Le modalità di accesso ai dati dipendono dalla loro organizzazione fisica che, in generale, va programmata.	Si può accedere ai dati indipendentemente dalla loro organizzazione fisica, che viene scelta fra le organizzazioni previste dal DBMS.	Indipendenza fisica.
I programmi usano nomi diversi per riferirsi agli stessi dati.	I programmi usano i nomi definiti nello schema.	Standardizzazione dei dati.
I dati sono accessibili solo da programmi.	I dati sono accessibili da programmi e interattivamente.	Facilità d'uso dei dati.
La sicurezza può essere assicurata solo scrivendo opportuni programmi.	I DBMS offrono meccanismi per garantire la sicurezza dei dati.	Sicurezza dei dati.
Ogni applicazione deve definire le proprie procedure per proteggere i dati da malfunzionamenti.	I DBMS offrono meccanismi per la protezione dei dati da malfunzionamenti.	Affidabilità dei dati.
I dati sono usati in modo esclusivo da ogni applicazione.	I dati sono accessibili in modo concorrente.	Condivisione dei dati.