

BASI DI DATI E SISTEMI DI GESTIONE DATI

Contesto, caratteristiche e servizi offerti

MEMORIZZARE INFORMAZIONI

- Risorse di ogni **organizzazione/azienda**
 - persone
 - denaro
 - materiali
 - immobili
 - **informazioni**

— **Alibaba** è il più grande negozio al mondo, ma non ha nemmeno un magazzino.



— **Uber** è la più grande compagnia di noleggio veicoli, ma non possiede nemmeno un'auto.



— **Airbnb** è il più esteso network dedicato alla ricettività, ma è del tutto privo di strutture.



MEMORIZZARE INFORMAZIONI

- Risorse di ogni organizzazione/azienda
 - persone
 - denaro
 - materiali
 - immobili
 - **informazioni**



SISTEMA INFORMATIVO

Compiti

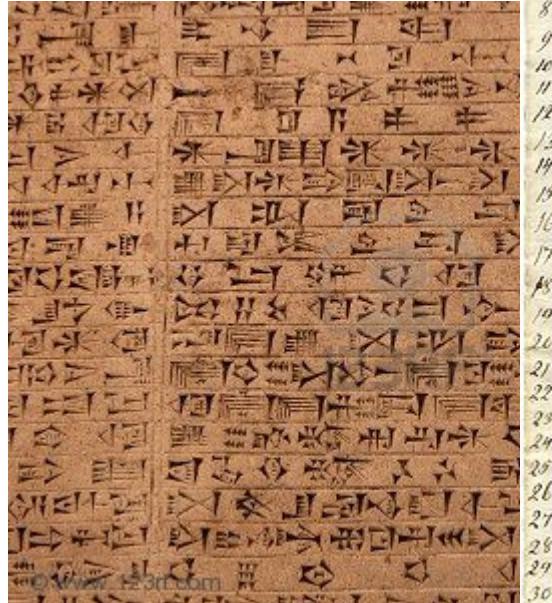
- raccogliere i dati
- conservare i dati raccolti, archiviandoli
- elaborare i dati, trasformandoli in informazioni
- distribuire l'informazione agli utilizzatori

Componenti

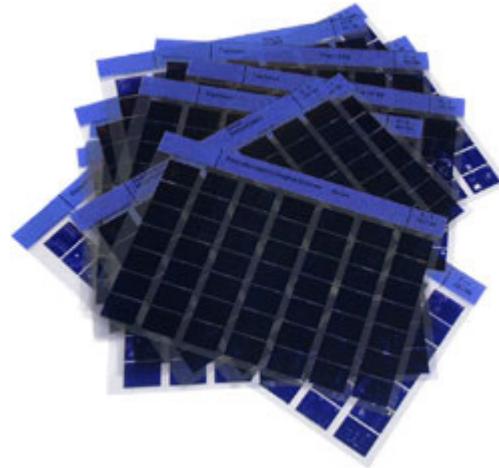
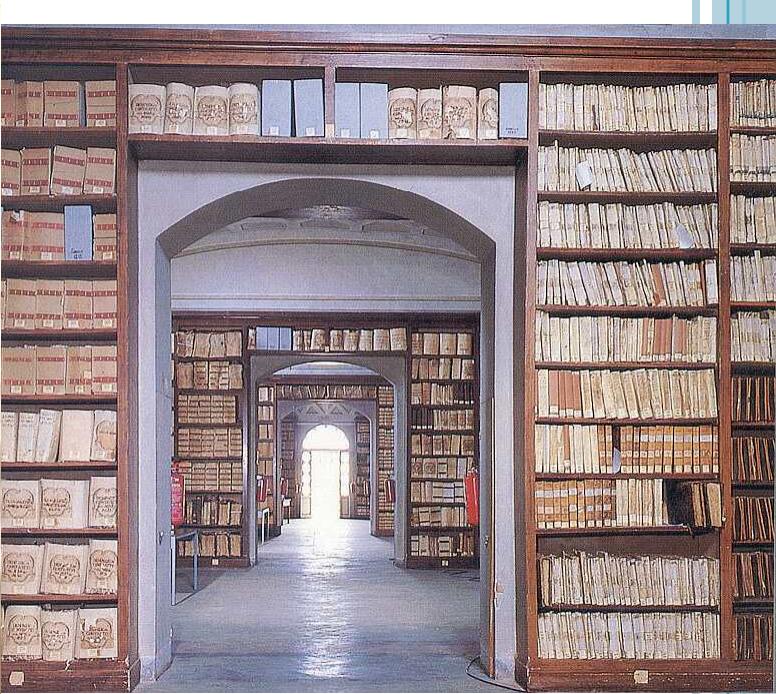
- strumenti
- procedure
- strutture

**Definizione del tutto indipendente
dal grado di automazione**

ESEMPI

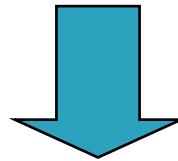


No.	Dødsdagen.	Døgsdørs Døgn.	Den Døde Bon og Til. Navn.	Sted, Hændelæg- og Dødssteds.	Ålder.	Over- førte i her- stadssteds og Døds- stedet.	Namensfølge.
							1822
1	3 ^e Augt	6 ^e heft	Engel Haups	Gjenn.	My	—	St.) Jens Christian
8.	11 ^e dø	15 ^e -2 ^e	Bjergbygård Haups	Støk —	27 ^a	—	—
9.	19 ^e -2 ^e	19 ^e -2 ^e	Hans Haups	Kongsbæk	22 ^a	—	2 ^e
10.	1 ^e Augt	—	Borgmøggs	Egeland —	12 ^a	—	Jacob.
11.	Ge 1	—	Jens Olfert	ibid —	29 ^a	—	Jens.
12.	—	—	Nils Haups	Gjenn. —	21 ^a	—	—
13.	9 ^e	—	Torbjørn Haups	Rosvold —	21 ^a	—	—
14.	9 ^e	—	Lars Petersen	Gjenn. —	46	—	—
15.	2 ^e	—	Hans Haups	Eidsvæn —	11	—	—
16.	8 ^e	—	Hans Larsen	Gjenn. —	24 ^a	—	—
17.	8 ^e	—	Hans Olfert	Hellands —	26 ^a	—	—
18.	8 ^e	—	Hansen Haups	Borge —	21 ^a	—	—
19.	5 ^e	—	Aren Endresen	ibid —	23 ^a	—	—
20.	5 ^e	—	Engel Haups	Schleid —	31 ^a	—	—
21.	5 ^e	—	Fredrik Haups	ibid —	29 ^a	—	—
22.	5 ^e	—	Olen Olfert	Solheim —	21 ^a	—	—
23.	5 ^e	—	Hans Christof	ibid —	19 ^a	—	—
24.	5 ^e	—	Jens Larsen	Kongsvæn —	40 ^a	—	—
25.	8 ^e	—	Daniel Haups	Hestad —	31	—	—
26.	5 ^e	—	Peter Hansen	Hafslund —	27	—	—
27.	5 ^e	—	Peder Asker	Land —	23	—	—
28.	5 ^e	—	Ole Johnson	Fjord —	46	—	—
29.	5 ^e	—	Hans Johnson	ibid —	32	—	—
30.	5 ^e	—	Lars Noddle	Hafslund —	24	—	—
31.	2 ^e	—	Rintje Schouw	Fjord —	25	—	—
32.	2 ^e	—	Elias Christof	Wing —	51	—	—
33.	2 ^e	—	Peter Olfert	Wale —	45	—	—
34.	2 ^e	—	Johannes Haups	ibid —	47	—	—

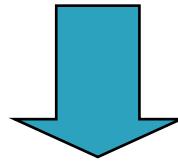


COSA SONO LE INFORMAZIONI?

Tutto ciò che **produce variazioni nel patrimonio cognitivo di un soggetto**, il **percettore dell'informazione**



L'informazione deve essere utile per (comprendibile da) il percettore



Un sistema informativo deve fornire una chiave di lettura mediante cui **interpretare** l'informazione che gestisce

I DATI

- Nei sistemi informatici, le informazioni si rappresentano come **dati** registrati su un supporto che ne garantisce
 - conservazione
 - reperibilità

Esempio: un file di numeri su linee successive

2 2 12

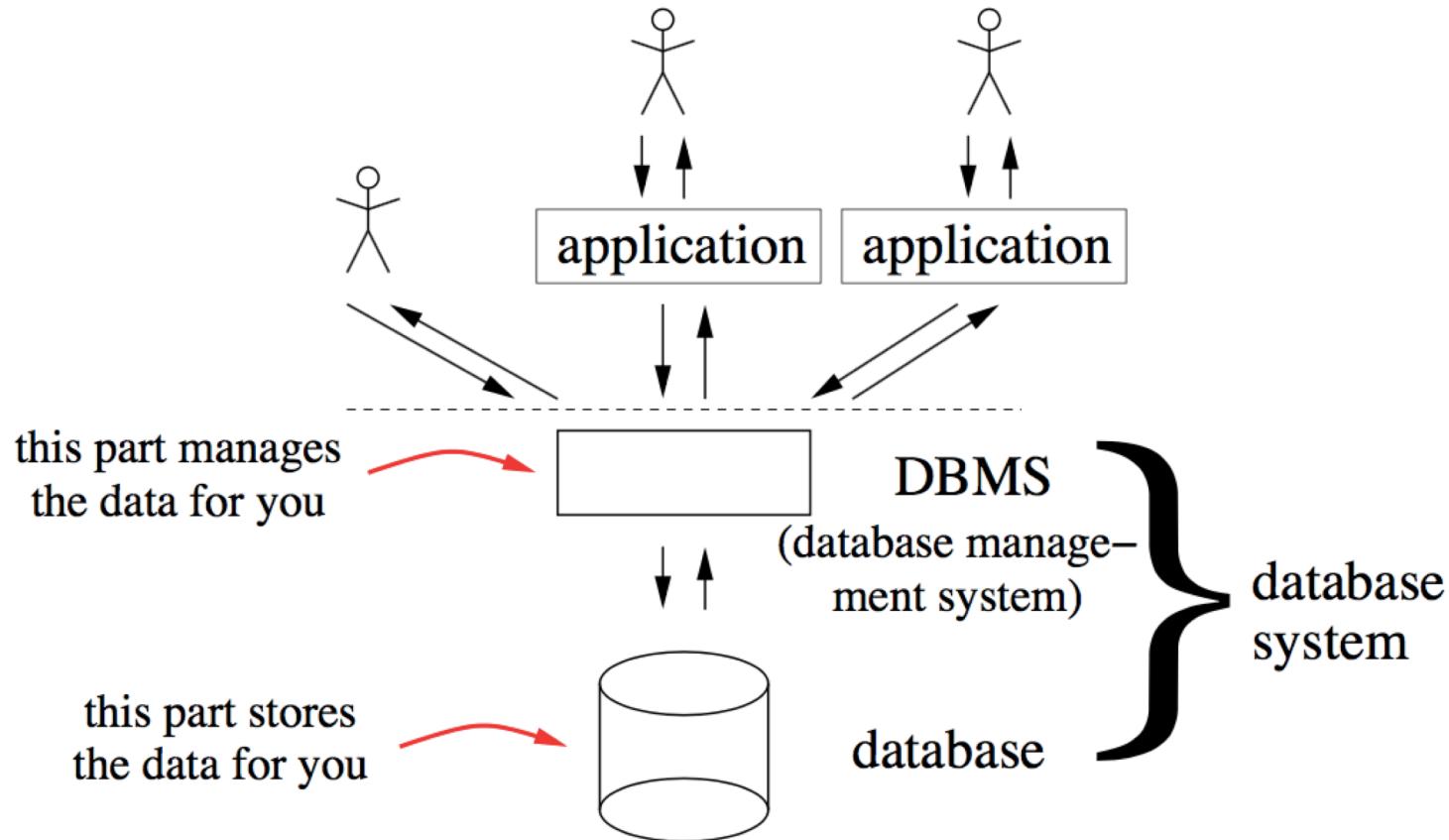
- Il sistema informativo fornisce un **contesto interpretativo** per trasformare i dati in informazioni

Insegnamento	Anno	Semestre	Crediti
Basi di dati (25880)	2	2	12

BASE DI DATI

- **Base di dati**
 - collezione di dati tra loro correlati, che rappresentano le informazioni di interesse in un sistema informativo
 - definizione indipendente dall'automazione
- **Sistema di gestione di basi di dati**
(DBMS – Data Base Management System)
 - sistema software che fornisce gli strumenti necessari a gestire basi di dati
 - se ci focalizziamo sui sistemi informativi supportati da un calcolatore, una base di dati è una collezione di dati gestita da un DBMS

BASI DI DATI, DBMS, APPLICAZIONI?



DAI SISTEMI OPERATIVI AI DBMS

- Ridondanza e inconsistenza
- Difficoltà di accesso ai dati
- Integrità dei dati
- Accesso concorrente ai dati
- Protezione dei dati



```
private String generateName(String[] settings) {
    String name = settings[0];
    for (String singlename : settings) {
        singlename = singlename.replace(" ", "-");
        String[] settings = singlename.split("-");
        if (settings[0].compareTo("s") == 0) {
            if (name.compareTo("") != 0) {
                name += "-";
            }
        } else if (etr.getString(settings[1]) != null) {
            if (name.compareTo("") != 0) {
                name += "-";
            }
            name += etr.getString(settings[1]);
        } else if (settings[0].compareTo("d") == 0) {
            if (name.compareTo("") != 0) {
                name += "-";
            }
            name += DateUtils.format(etr.getDate(settings[1]));
        } else if (settings[0].compareTo("n") == 0) {
            if (name.compareTo("") != 0) {
                name += "-";
            }
            comSysNumber = etr.getDouble(settings[1]);
            f = NumberFormat.getInstance();
            f.setGroupingUsed(false);
        }
    }
}
```

Insegnamento	Anno	Semestre	Crediti
Basi di dati (25880)	4	3	12



12

DAI SISTEMI OPERATIVI AI DBMS

1) Ridondanza e inconsistenza

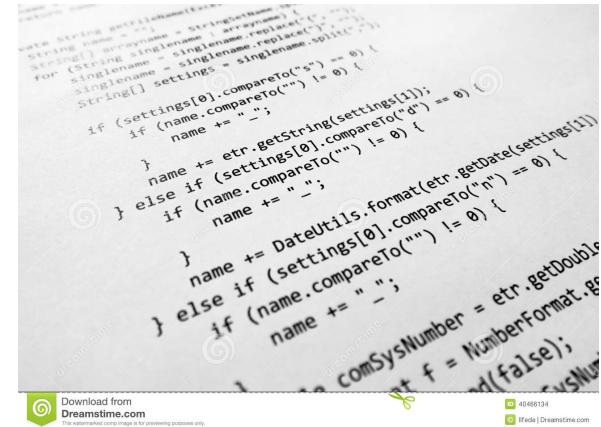
- Piani di studio
- Orario delle lezioni
- Compiti didattici dei prof



DAI SISTEMI OPERATIVI AI DBMS

2) Difficoltà di accesso ai dati

- Lista di studenti che devono sostenere BD
- Lista di studenti che devono sostenere BD e il relativo numero di crediti
- Lista di studenti che devono sostenere BD per 12 CFU e orario degli **altri insegnamenti che seguono**



The watermark shows a spiral-bound notebook with handwritten code. The code is in Java and includes operations like string concatenation, array manipulation, and conditional statements (if, else if). It seems to be part of a larger program, likely related to database management or data retrieval.

DAI SISTEMI OPERATIVI AI DBMS

3) Integrità dei dati

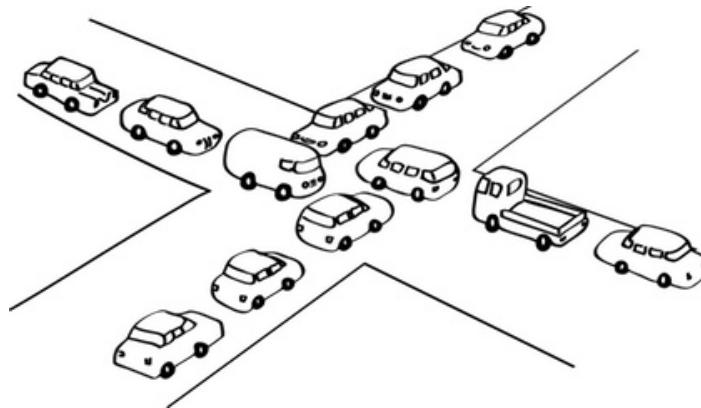
- Se seguo un corso in un anno accademico il corso deve essere erogato in quell'anno

Insegnamento	Anno	Semestre	Crediti
Basi di dati (25880)	4	3	12

DAI SISTEMI OPERATIVI AI DBMS

4) Accesso concorrente

controllo della concorrenza a granularità molto fine
(non quella del file)



DAI SISTEMI OPERATIVI AI DBMS

5) Protezione dei dati

Controllo dell'accesso a granularità molto fine (non quella del file) e anche basata sul contenuto



LA «CARTA VINCENTE» DEI DBMS

- **Schema** (o schema logico) della base di dati
 - descrive la struttura della base di dati
 - astrae dai dettagli dell'implementazione fisica
 - usa un formalismo ad alto livello detto **modello dei dati**
- Unica rappresentazione **centralizzata** e **condivisa**
 - No ridondanze e inconsistenze
 - Permette controllo centralizzato

SCHEMA E ISTANZA

- Indipendentemente dal modello dei dati, in un DBMS distinguiamo:
 - **Schema della base di dati:** descrizione dei dati, specificata tramite il modello dei dati
 - Cambia raramente
 - Descrizione intensionale
 - **Istanza della base di dati:** insieme dei dati presenti in un dato momento in una base di dati
 - Cambia molto spesso nel tempo
 - Descrizione estensionale

Insegnamento	Anno	Semestre	Crediti
Basi di dati (25880)	2	2	12

LA SECONDA «CARTA VINCENTE» DEI DBMS

- Il DBMS mette a disposizione **linguaggi (dichiarativi)** per agire sui dati rappresentati secondo il modello
 - Possibilità di interagire con i dati secondo modalità non anticipate al momento della definizione e realizzazione della base di dati
 - Specifica di cosa si vuole recuperare e non delle specifiche istruzioni per recuperarlo

SELECT ...
FROM ...
WHERE ...

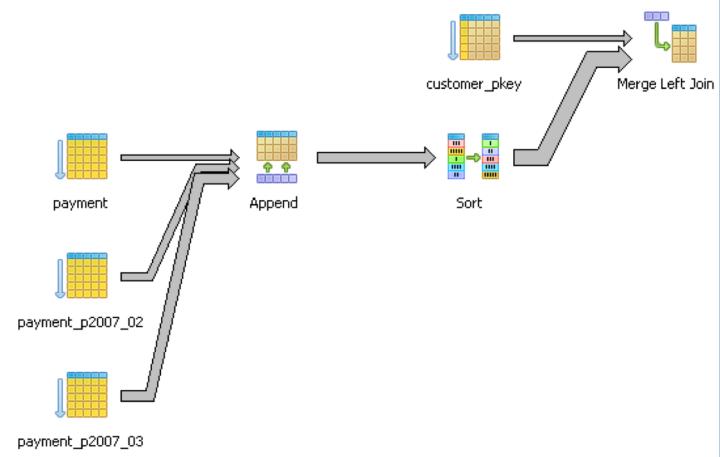
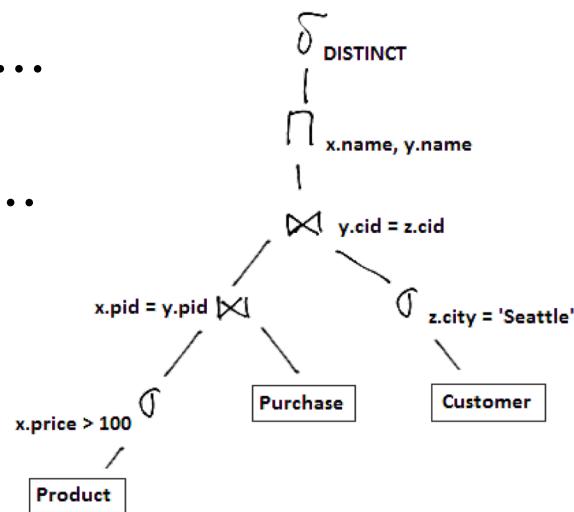
VS

[MV_UI4_C_C 1 0
LOAD_SF8_C 4 1 6
LOAD_SF8_C 5 1 7
LOAD_SF8_C 6 1 8
...]

LA SECONDA «CARTA VINCENTE» DEI DBMS

- Il DBMS mette a disposizione **linguaggi (dichiarativi)** per agire sui dati rappresentati secondo il modello
 - Ottimizza le operazioni e quindi garantisce efficienza

SELECT ...
FROM ...
WHERE ...



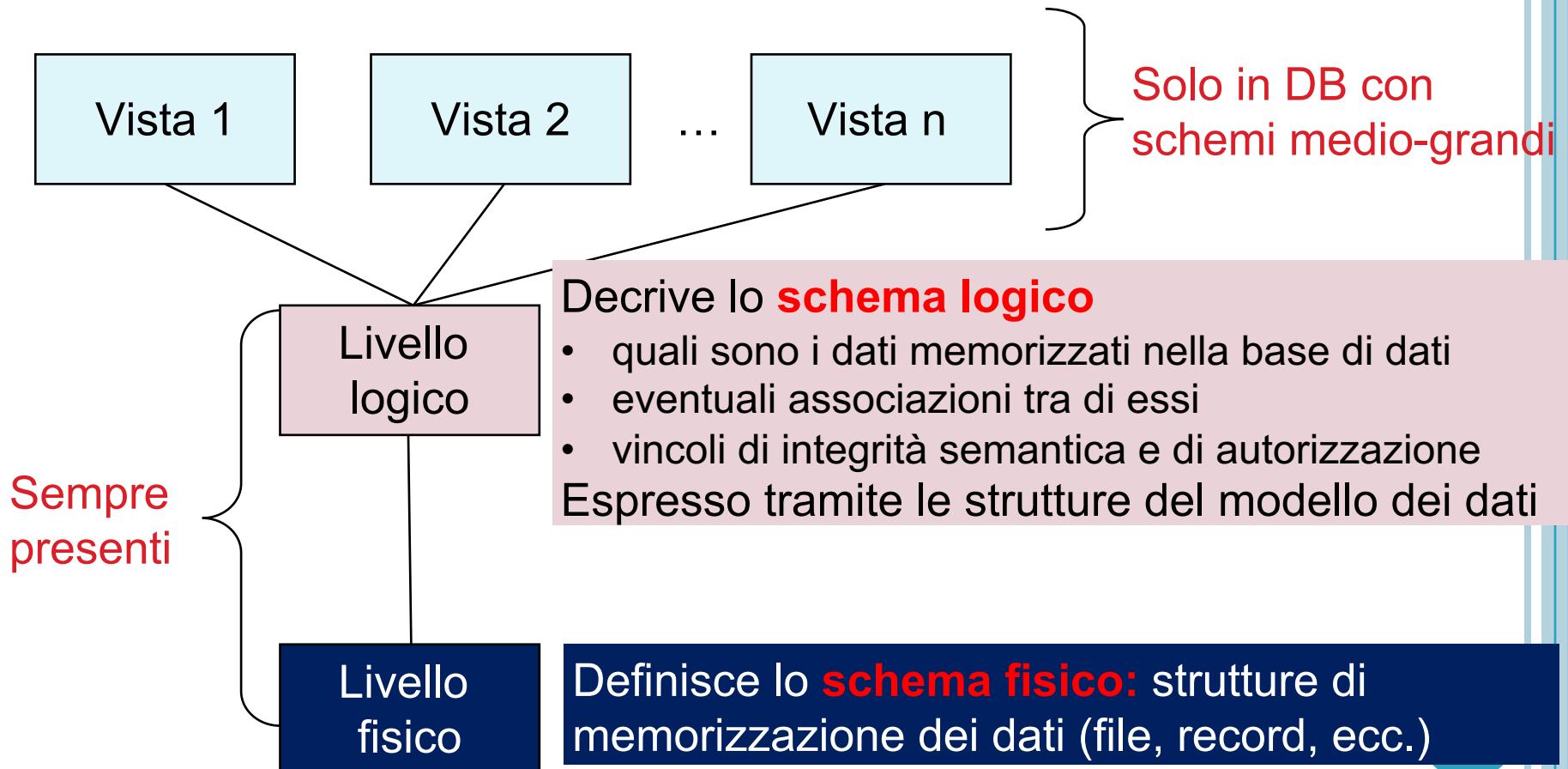
LINGUAGGI DICHIARATIVI

ESEMPIO

- Interrogazione dichiarativa:
 - Voglio conoscere il nome del cliente che l'8 gennaio 2018 ha affittato un video contenente Miss Peregrine e la casa dei ragazzi speciali
- Interrogazione procedurale:
 - Per ogni tupla della tabella Video, guarda se il titolo del film è uguale a “Miss Peregrine e la casa dei ragazzi speciali”
 - Se sì:
 - assegna ad una variabile v la collocazione del video
 - per ogni tupla della tabella Noleggio, guarda se la data di inizio del noleggio è uguale a 8 gennaio 2018 e la collocazione è uguale a v
 - Se sì:
 - assegna ad una variabile c il codice del cliente
 - per ogni tupla della tabella Cliente, guarda se il codice è uguale a c
 - Se sì:
 - Restituisci il nome del cliente
 - Altrimenti, passa ad analizzare la tupla successiva
 - Altrimenti, passa ad analizzare la tupla successiva
 - Altrimenti, passa ad analizzare la tupla successiva

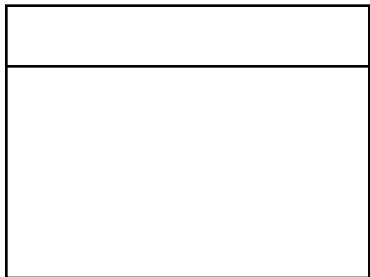
LIVELLI NELLA RAPPRESENTAZIONE DEI DATI

- descrive una porzione dell'intero schema della base di dati (**vista**)
- possono essere definite più viste di una stessa base di dati

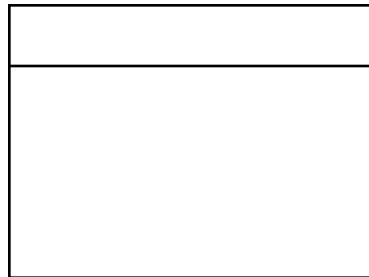


ESEMPIO

Vista_Film_Commedia

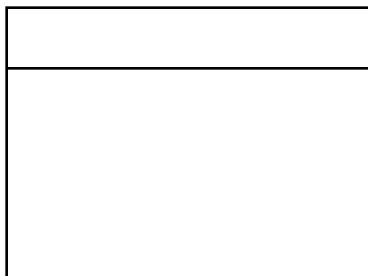


Vista_BluRay

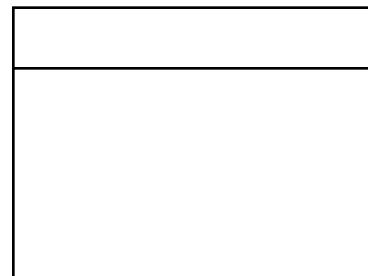


Livello esterno

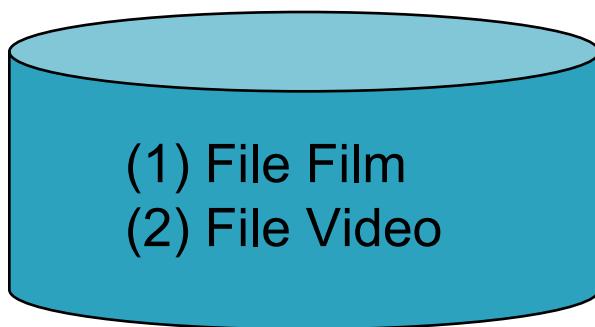
Film



Video



Livello logico



Livello fisico

INDIPENDENZA

- Usando questi tre livelli
Indipendenza fisica

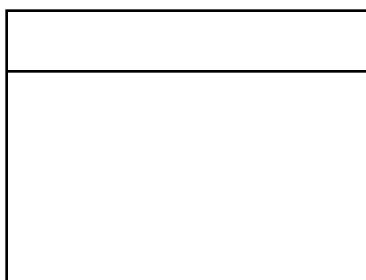
Indipendenza logica

- Si facilitano
 - l'accesso ai dati
 - sviluppo di applicazioni
 - **manutenibilità**

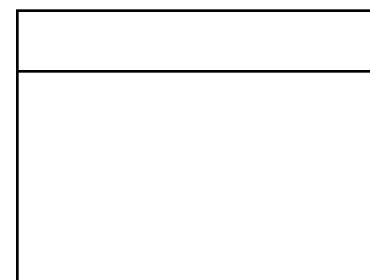
INDIPENDENZA FISICA: ESEMPIO

Tempo T1

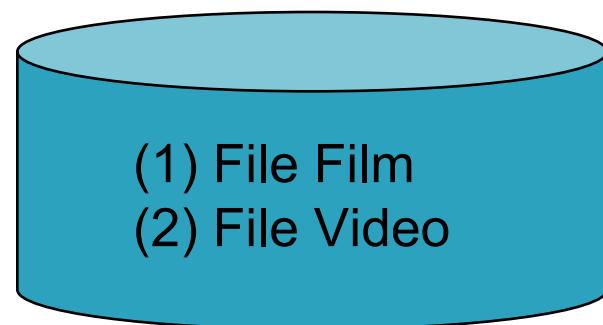
Film



Video



Livello logico



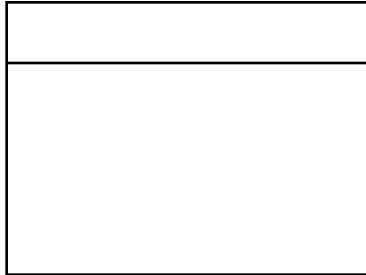
Livello fisico

INDIPENDENZA FISICA: ESEMPIO

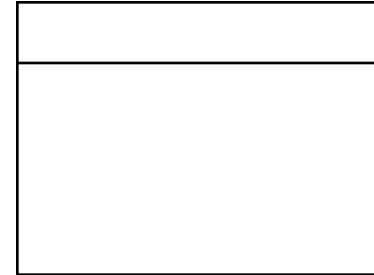
Tempo T2:

**cambiano le strutture di memorizzazione
il livello logico non cambia**

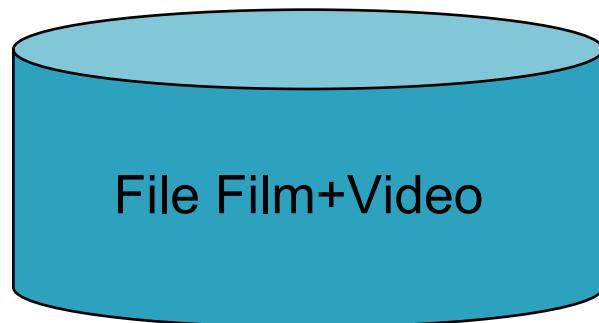
Film



Video



Livello logico

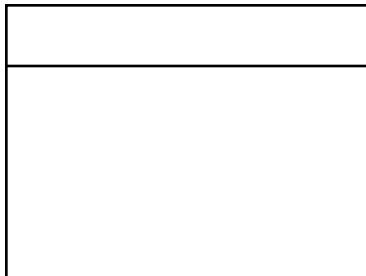


Livello fisico

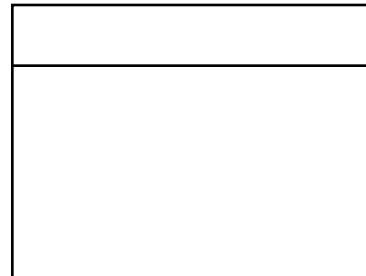
INDIPENDENZA LOGICA: ESEMPIO

Tempo T1

Vista_Film_Commedia

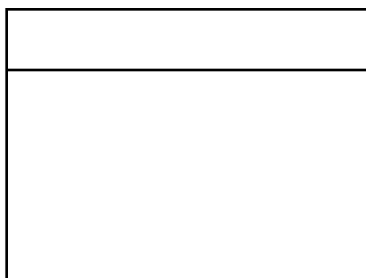


Vista_ClientiVIP

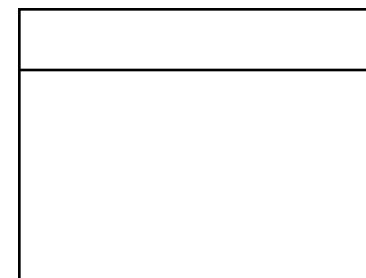


Livello esterno

Film



Video



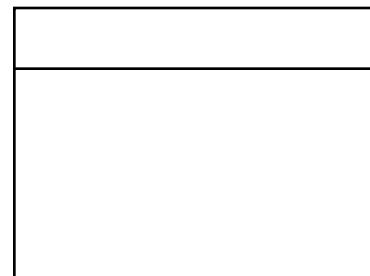
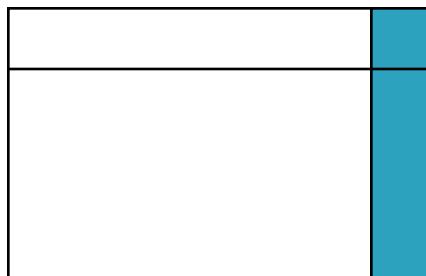
Livello logico

INDIPENDENZA LOGICA: ESEMPIO

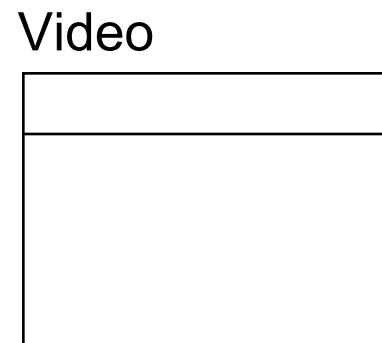
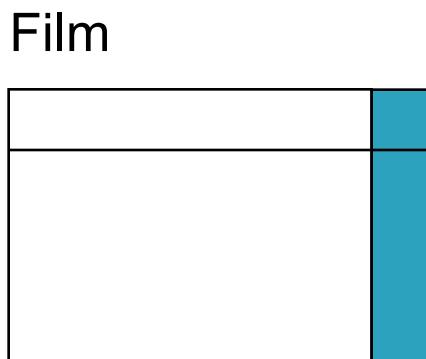
Tempo T2:

**cambia lo schema logico (aggiunta di un attributo)
solo una parte delle viste cambia**

Vista_Film_Commedia Vista_ClientiVIP



Livello esterno



Livello logico

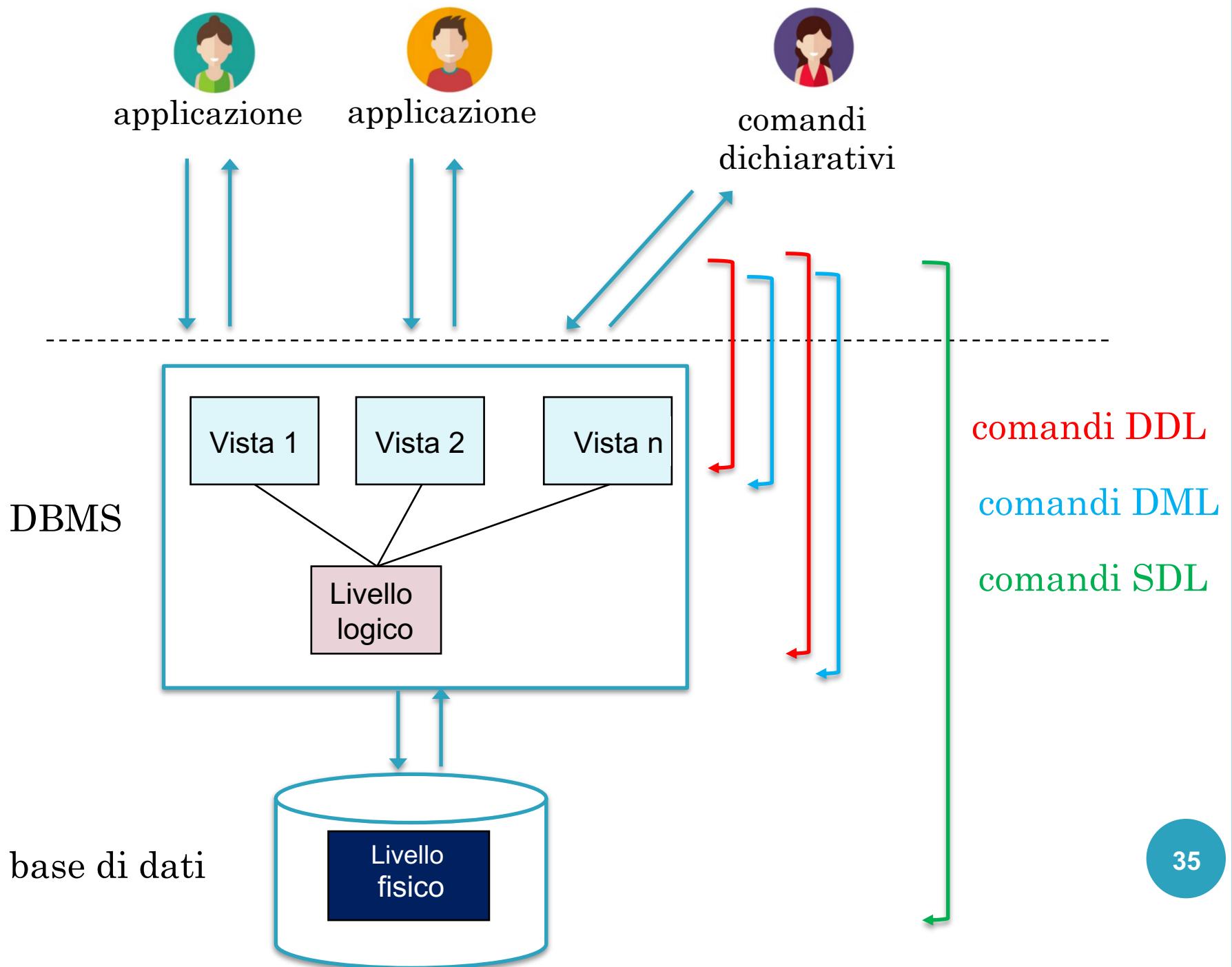
DUE PUNTI FERMI DEI DBMS RELAZIONALI

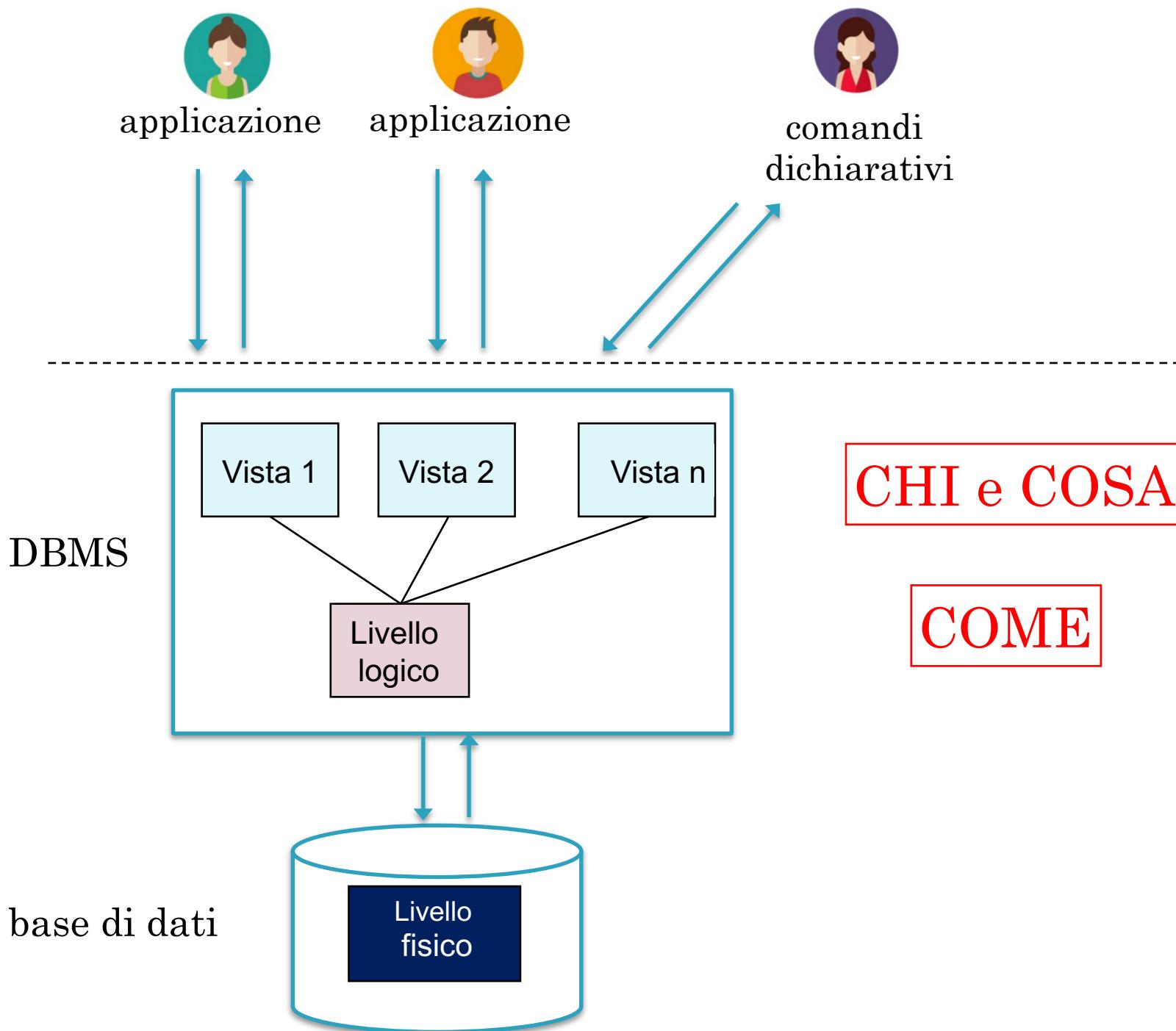
- **Integration DB piuttosto che Application DB**
- **Schema logico della base di dati indipendente dal workload (=come intendo utilizzare i dati)**
- **Punti fermi = Punti di forza?**

... esistono altri DBMS basati su altre scelte
(v. corsi della Laurea Magistrale)

IL MODELLO DEI DATI: LINGUAGGI

- **Data Definition Language (DDL)**
 - Permette di specificare e modificare lo schema della base di dati e lo schema delle viste
 - Livello logico ed esterno
- **Data Manipulation Language (DML)**
 - Permette di creare, modificare e interrogare l'istanza della base di dati
 - Livello logico (ed esterno)
- **Storage Definition Language (SDL)**
 - Definisce lo schema fisico del DB
 - Livello fisico





CHI E COSA?



- Chi può accedere la base di dati e come?
- A quali risorse si può accedere?
- Chi può concedere privilegi e su cosa?
- Chi può revocare privilegi e su cosa?

COME?



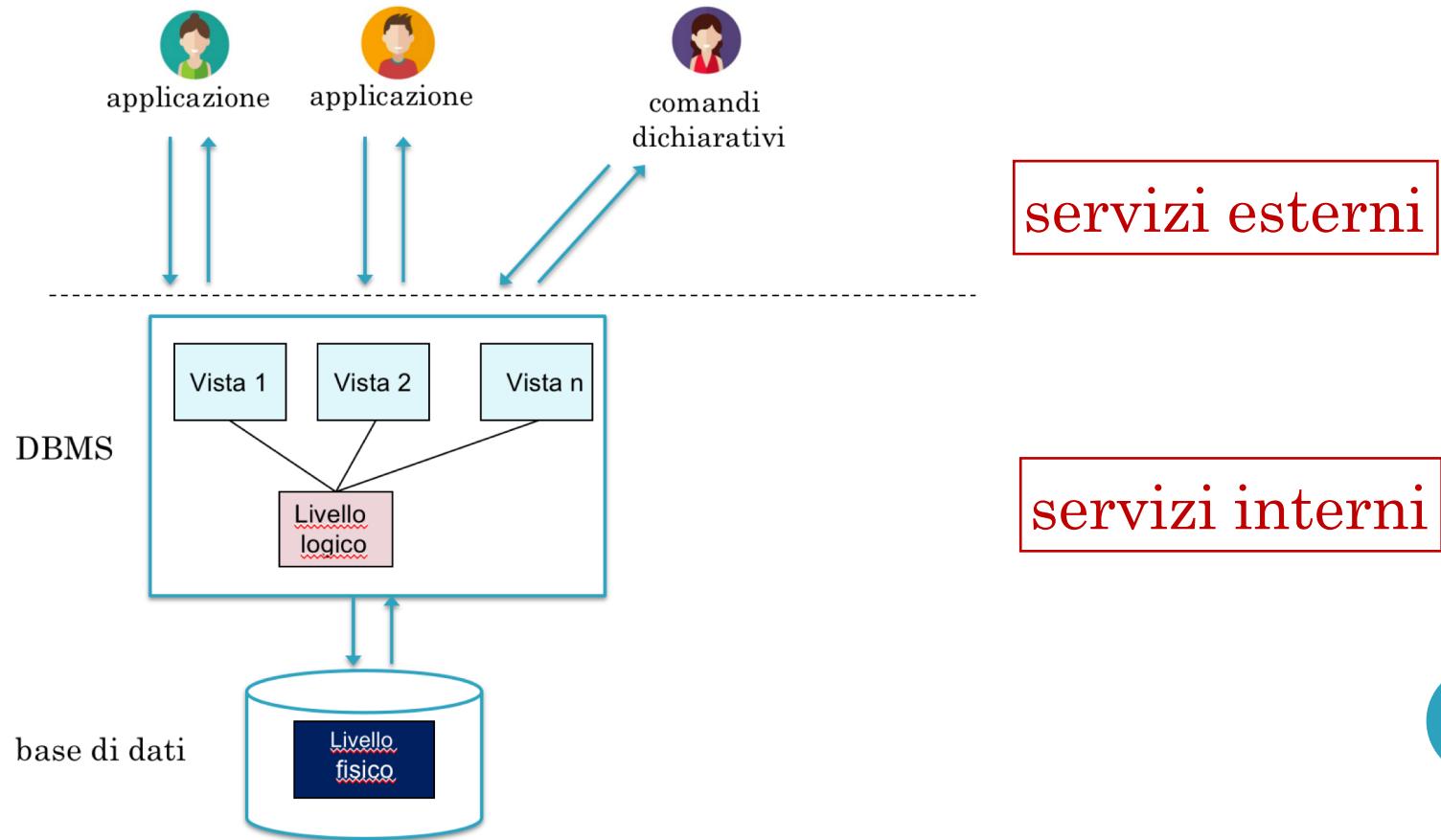
- Come si possono controllare gli accessi alla base di dati?
- Come si possono specificare le autorizzazioni?

- Come si possono eseguire i comandi dichiarativi?
- Come si può rendere più efficiente l'esecuzione dei comandi dichiarativi?
- Come si possono monitorare le prestazioni del sistema e adottare correttivi?

- Come si può eseguire un programma applicativo che richiede l'esecuzione di comandi dichiarativi?
- Come si possono eseguire tanti programmi applicativi che operano sulla stessa base di dati nello stesso momento evitando errori dovuti all'elaborazione concorrente?
- Come si può fare fronte a eventuali guasti che possono portare a perdite di dati?

SERVIZI DI UN DBMS

- Il DBMS risponde alle domande precedenti mediante **servizi**



SERVIZI ESTERNI AL DBMS

- Gli utenti comunicano con il DBMS, tramite comandi in linguaggi dichiarativi, invocando servizi **esterni**

Descrizione dei dati	Specificare i dati da memorizzare
Manipolazione dei dati	<ul style="list-style-type: none">• Inserire nuovi dati• Accedere ai dati esistenti• Modificare dati esistenti• Cancellare dati esistenti

- I servizi esterni possono essere invocati
 - tramite interfacce (testuali o grafiche)
 - tramite API per l'esecuzione dei comandi dichiarativi all'interno di programmi (applicazioni) scritti in linguaggi general purpose

SERVIZI INTERNI AL DBMS

- Il DBMS implementa i servizi esterni mediante molteplici servizi **interni**

Controllo di integrità	Mantenere consistenza e correttezza dei dati
Strutture di memorizzazione	Rappresentare in memoria secondaria i costrutti del modello dei dati
Ottimizzazione di interrogazioni	Determinare strategie efficienti di accesso ai dati
Protezione dei dati	Evitare accessi non autorizzati
Ripristino della base di dati	Evitare che errori e malfunzionamenti <ul style="list-style-type: none">Rendano inconsistente la base di datiProvochino perdite di dati
Controllo della concorrenza	Evitare inconsistenze dovute ad accessi concorrenti

DBMS (DEFINIZIONE RIVISITATA)

- Un DBMS è un sistema software (complesso!) che offre molteplici servizi al fine di gestire

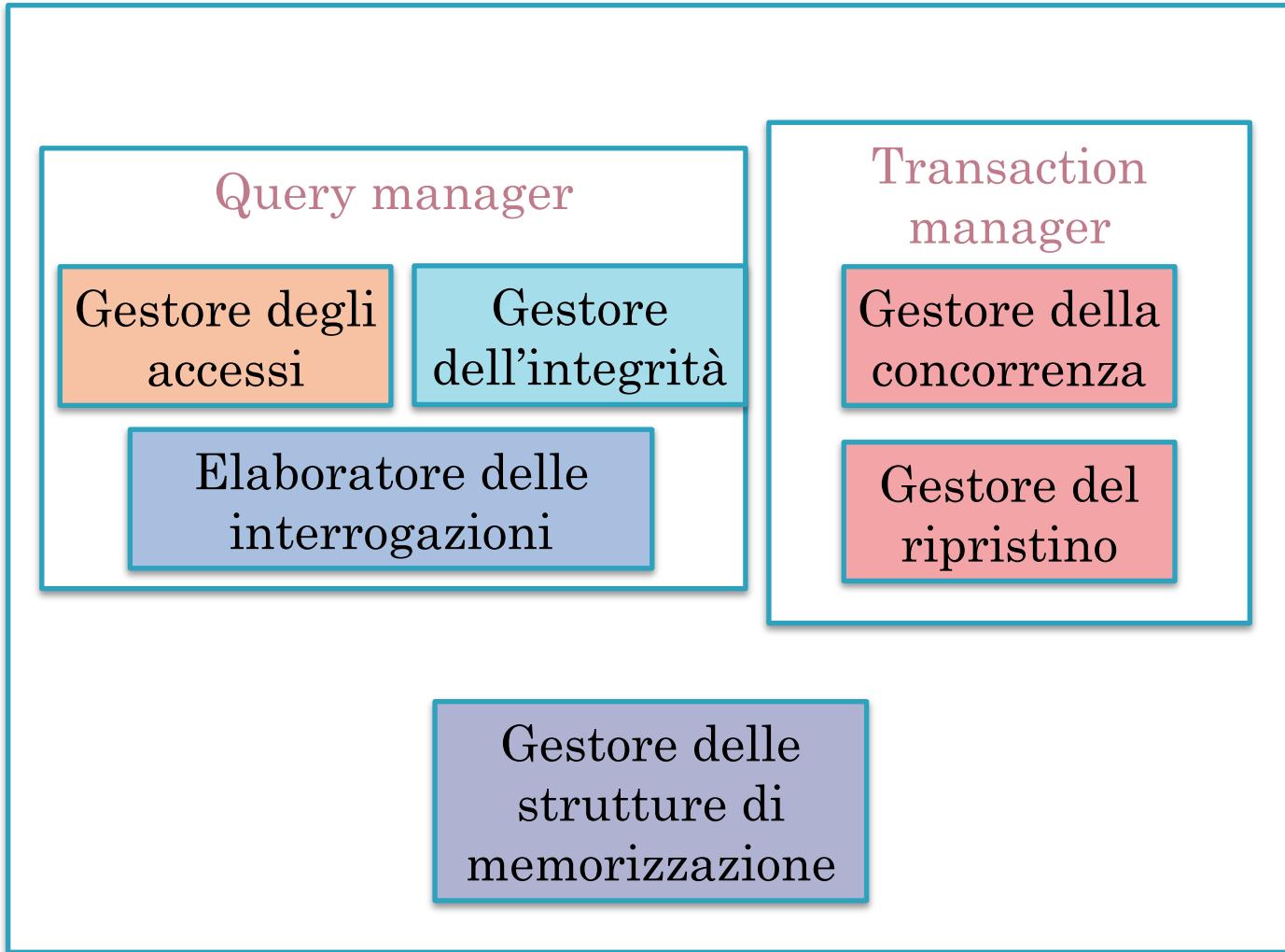
- grandi quantità di dati persistenti e condivisi
- in modo efficiente
- garantendone l'integrità
- seguendo una determinata politica degli accessi
- mediante l'esecuzione coordinata delle richieste e la protezione da malfunzionamenti



- Ogni servizio corrisponde ad uno o più componenti software

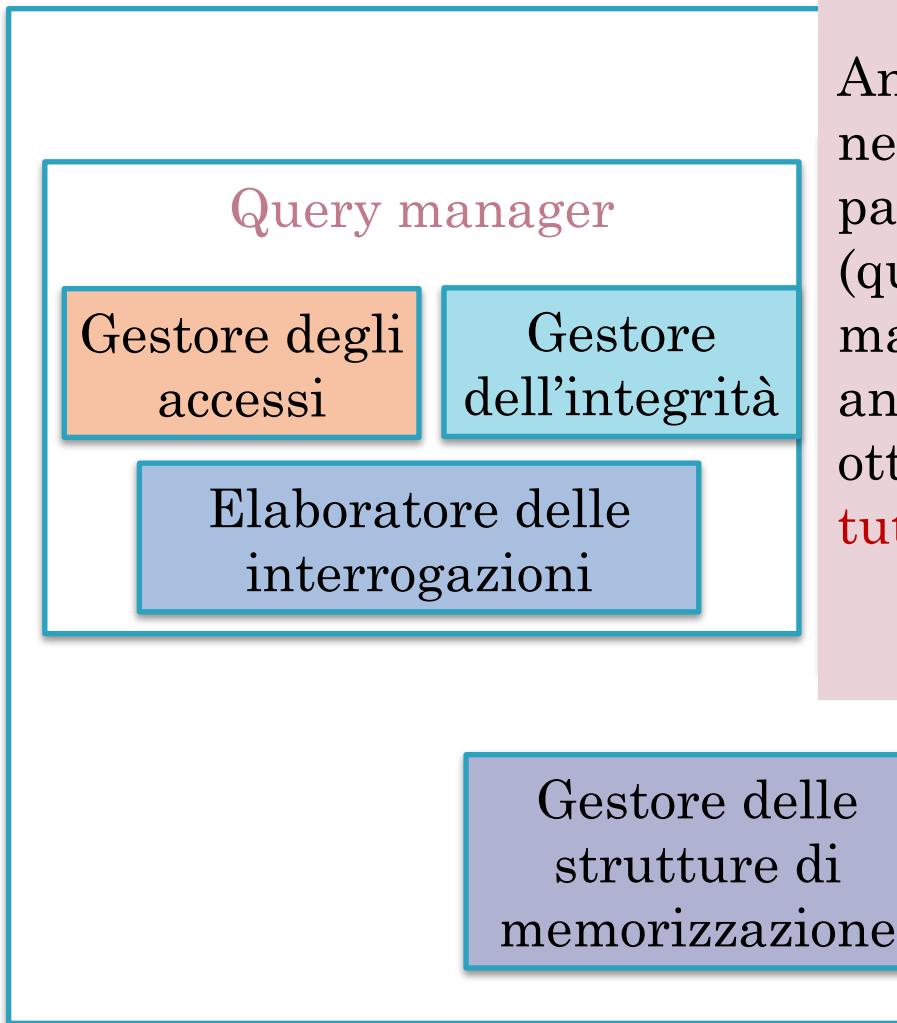
COMPONENTI DI UN DBMS

DBMS



COMPONENTI DI UN DBMS

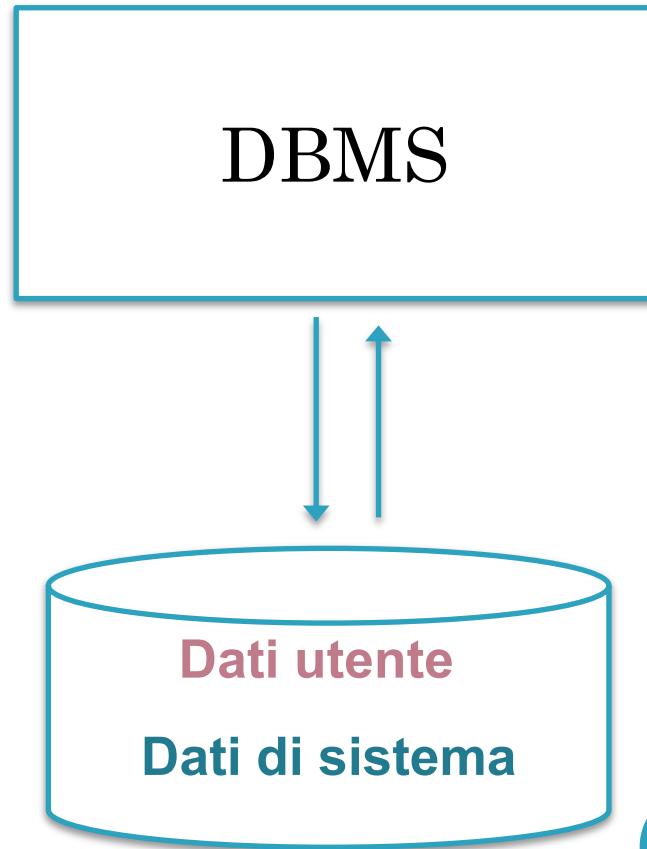
DBMS



Anche se nell'architettura si parla di interrogazioni (query), il query manager si occupa di analizzare, autorizzare, ottimizzare ed eseguire tutti i comandi DML

COMPONENTI DI UN DBMS

- Tutte le diverse componenti, per il corretto funzionamento, hanno bisogno di accedere a **dati utente** e/o **dati di sistema**
- Ad esempio, il gestore delle autorizzazioni deve accedere alle regole di autorizzazione (dati di sistema)
- I dati di sistema vengono memorizzati nella base di dati insieme ai dati utente



UTENTI DI UN DBMS

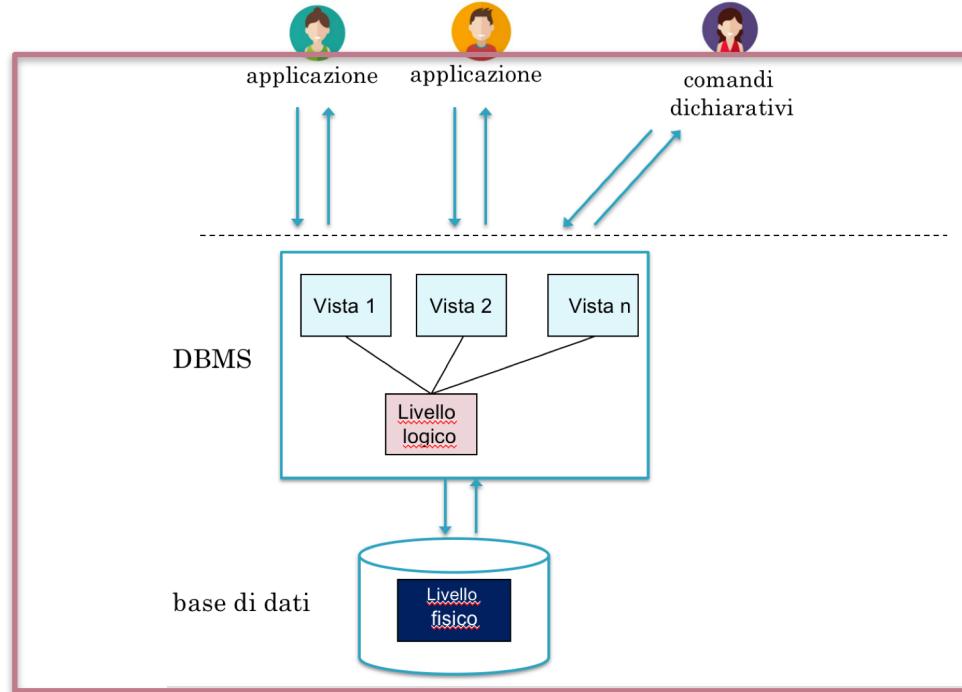
- **Progettista di basi di dati**
- **Programmatore applicativo**
- **Amministratore della base di dati (DBA)**
- Progettista, sviluppatore di DBMS

ARCHITETTURA DI UN DBMS

- Centralizzata
- Client-server
- Distribuita
- Parallelia

ARCHITETTURA CENTRALIZZATA

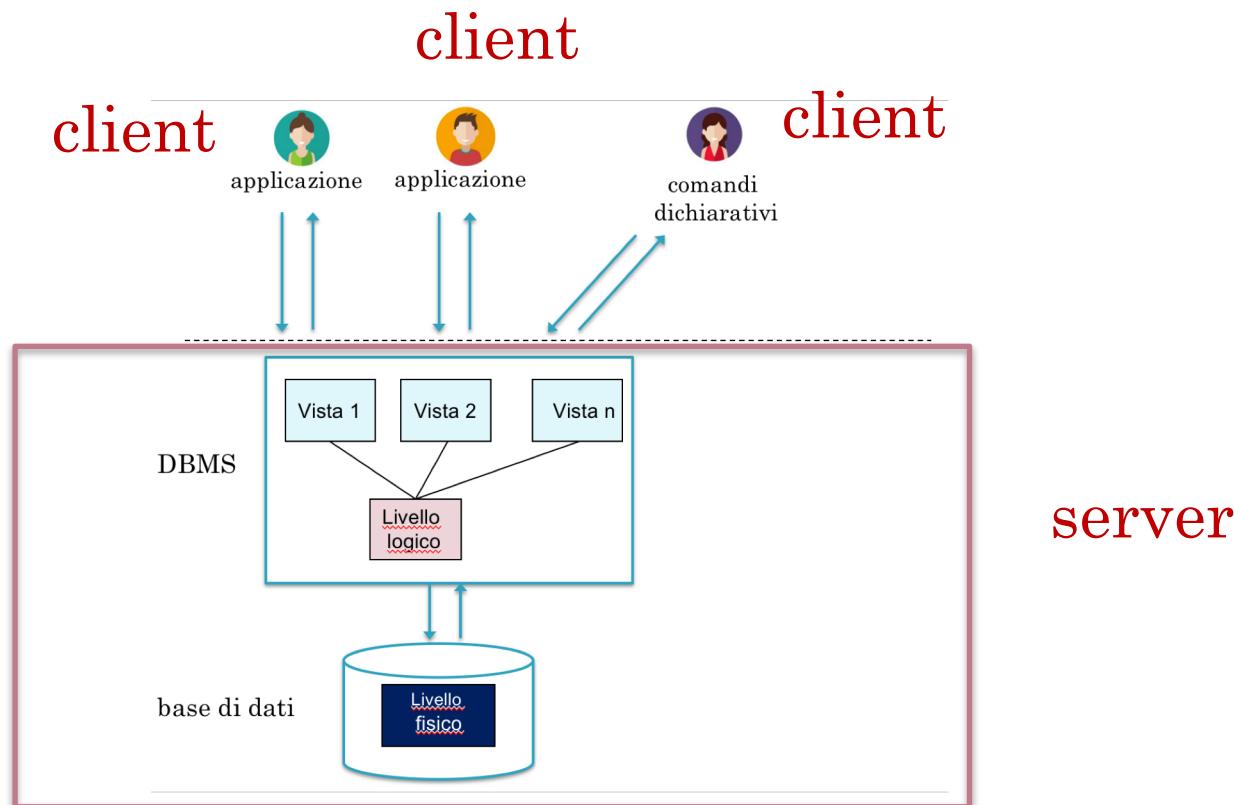
- I servizi interni ed esterni vengono forniti da una singola macchina
- Le applicazioni risiedono e vengono eseguite sulla stessa macchina che offre i servizi
- Le richieste di servizi esterni al DBMS vengono effettuate dalla stessa macchina che offre i servizi



una singola
macchina

ARCHITETTURA CLIENT-SERVER

- I servizi interni ed esterni vengono forniti da una singola macchina, chiamata **server**
- Le applicazioni risiedono e vengono eseguite su macchine **client**, dalle quali vengono richiesti i servizi esterni



ARCHITETTURE DISTRIBUITE

- Un'architettura è distribuita quando si verifica almeno una delle due condizioni
 - Le applicazioni, fra loro cooperanti, risiedono su più macchine (nodi): **elaborazione distribuita**
 - Il patrimonio informativo (dati), unitario, è ospitato su più macchine (nodi): **base di dati distribuita**
- I principi e gli algoritmi su cui si basano i servizi interni ed esterni sono diversi rispetto a sistemi centralizzati o client-server
- L'architettura client-server è un esempio semplice ma molto diffuso di architettura che distribuisce l'elaborazione
- Esempi più complessi di DBMS distribuiti nei corsi della Laurea Magistrale