

# Basi di dati

# Basi di dati per bioinformatica



**DOCENTI**  
**PROF. ALBERTO BELUSSI**  
**PROF CARLO COMBI**

**Anno accademico 2013/14**

# Organizzazione degli insegnamenti

3

Basi di dati

Basi di dati per  
Bioinformatica

**Teoria (Belussi)**  
(modulo comune)

**Tecnologie**  
(Belussi)

**Tecnologie**  
(Combi)

**Laboratorio**  
(Belussi)

**Laboratorio**  
(Combi)

# Organizzazione didattica nei semestri

4

Gli insegnamenti si svolgono su due semestri

Moduli	CFU	I Sem	II Sem
Teoria	6	48 ore	-
Tecnologie delle basi di dati	2 +1ese	4 ore	24 ore
Laboratorio	3	-	36 ore

# Informazioni generali sull'organizzazione

5

- **Orario I Semestre**
  - Lunedì 11.30 – 13.30 (aula A)
  - Martedì 11.30 – 13.30 (aula A)
- **Ricevimento (Belussi)**
  - Lunedì 14.30 – 16.30
- **Ricevimento (Combi)**
  - Venerdì 14.30 – 16.30

# Informazioni generali sull'organizzazione

6

- **Testi e eserciziari**
  - P. Atzeni, S. Ceri, S. Paraboschi, R. Torlone  
Basi di dati: modelli e linguaggi di interrogazione  
(IV edizione) McGraw-Hill
  - E. Baralis, A. Belussi, G. Psaila  
Basi di dati: temi d'esame svolti  
Progetto Leonardo – Editrice Esculapio

# Programma

7

Insegnamento: BASI DI DATI

Sul sito di dipartimento:

<http://www.di.univr.it>

Pagina dell'insegnamento nel  
[corso di laurea in Informatica](#)

Pagina dell'insegnamento nel  
[corso di laurea in Bioinformatica](#)

# Modalità d'esame (Basi di dati)

8

## *Modulo Teoria (9 crediti)*

### ESAME SCRITTO:

- Progettazione concettuale/logica di una base di dati
- Algebra relazionale
- SQL
- Domande di teoria (indici, concorrenza, transazioni, ecc..)
- Dati multimediali
- XML

## *Modulo Laboratorio (3 crediti)*

### ESAME SCRITTO:

- Tecnologie per applicazioni web data-centric: Servlet, JSP, JDBC, JPA
- Progettazione logica di un applicazioni web
- Implementazione di applicazioni web secondo l'approccio MVC-2
- SQL

### oppure PROGETTO di una applicazione WEB

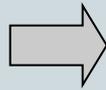
- Progettazione di un'applicazione
- Implementazione MVC-2 (servlet/JSP, PHP, Hybernate)



# Relazione con Algoritmi (e Strutture Dati)

10

Algoritmi e  
strutture dati



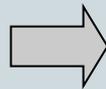
**Memoria  
Centrale**

(RAM, ...)

MEMORIA VOLATILE

Accesso veloce e  
casuale

Basi di dati



**Memoria  
secondaria**

(nastri, dischi,  
memorie a stato  
solido)

MEMORIA  
PERSISTENTE

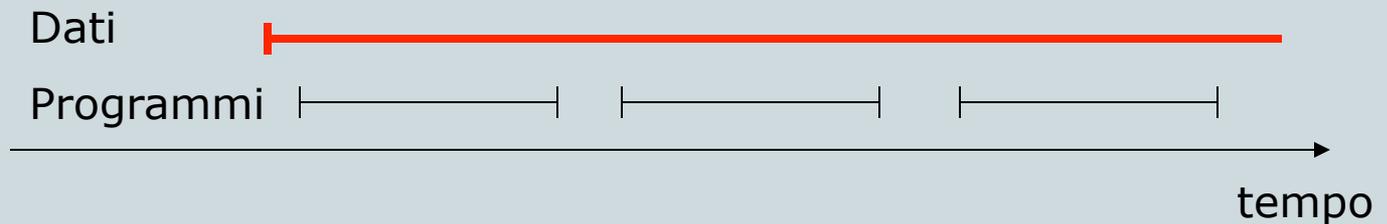
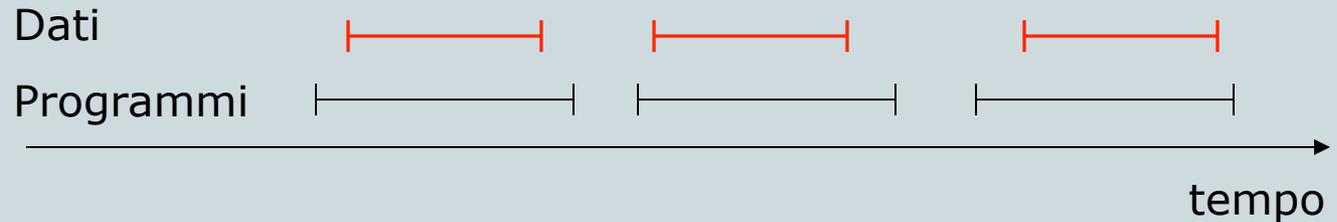
Accesso lento e  
sequenziale (file)

# Relazione con Algoritmi (e Strutture Dati)

11

## Tempo di vita dei dati

**Memoria  
Centrale**



# Evoluzione dell'informatica applicata

12

- anni '60: applicazioni negli ambienti di **ricerca scientifica** (laboratori) con enfasi sui programmi (e quindi sugli algoritmi).
- anni '70: nascono le applicazioni informatiche in **ambito gestionale**

# Applicazioni gestionali

13

## Caratteristiche

- algoritmi di elaborazione semplici.
- grande quantità di dati CONDIVISI da PIU' applicazioni.

Tali caratteristiche derivano dall'ambiente in cui vennero introdotte le applicazioni informatiche (anni '70):

## IL SISTEMA INFORMATIVO

# Sistema Informativo

14

## Definizione

*“è l’insieme delle attività umane e dei dispositivi di memorizzazione ed elaborazione che organizza e gestisce l’informazione di interesse per un’organizzazione di dimensioni qualsiasi”*

N.B.: un sistema informativo NON contiene necessariamente tecnologia informatica.

# Sistema Informativo

15

## **Dato**

Elemento di conoscenza di base costituito da simboli che devono essere elaborati.

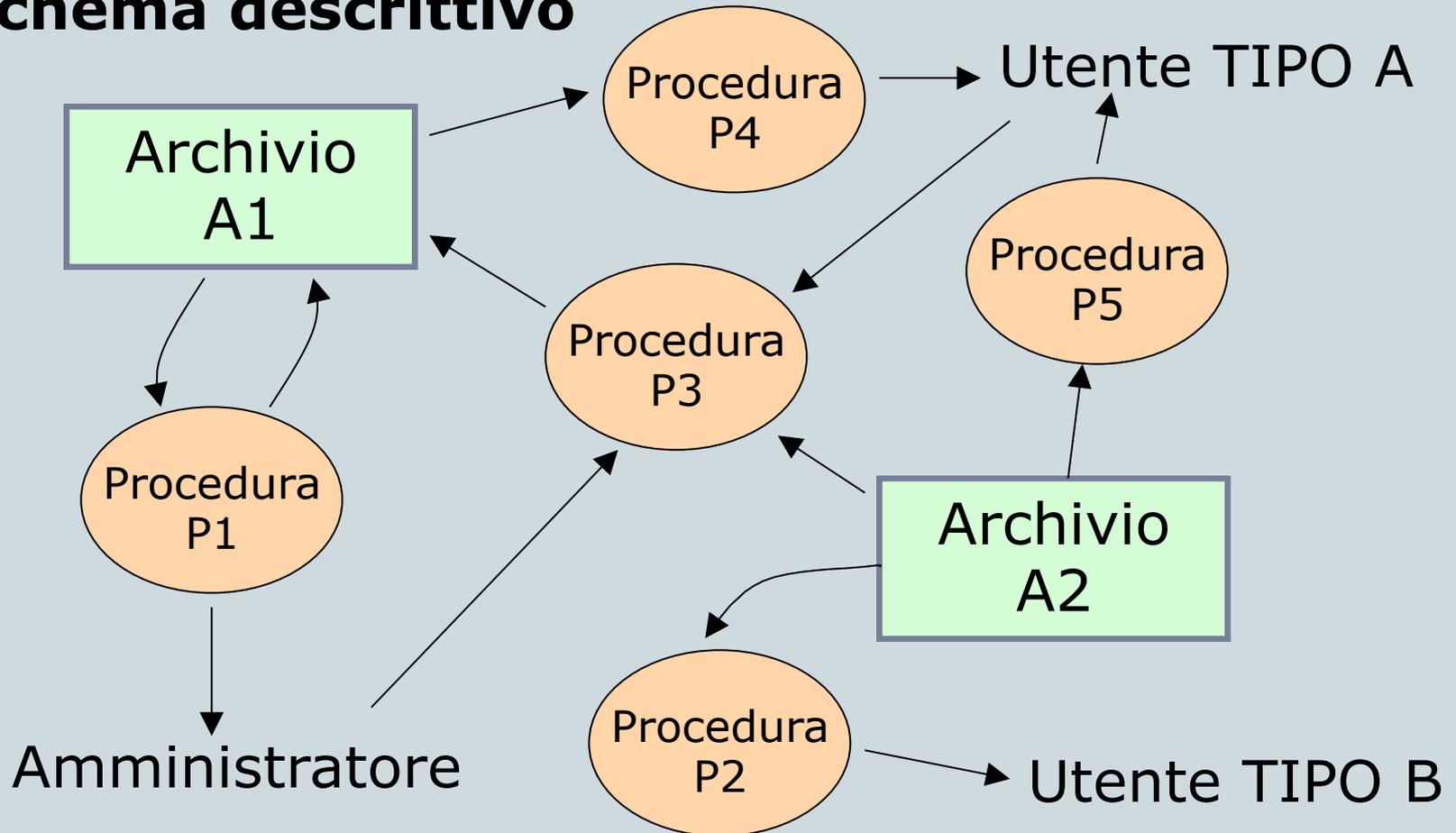
## **Informazione**

Interpretazione dei dati che permette di ottenere conoscenza più o meno esatta di fatti e situazioni.

# Sistema Informativo

16

## Schema descrittivo



## **BASE DI DATI**

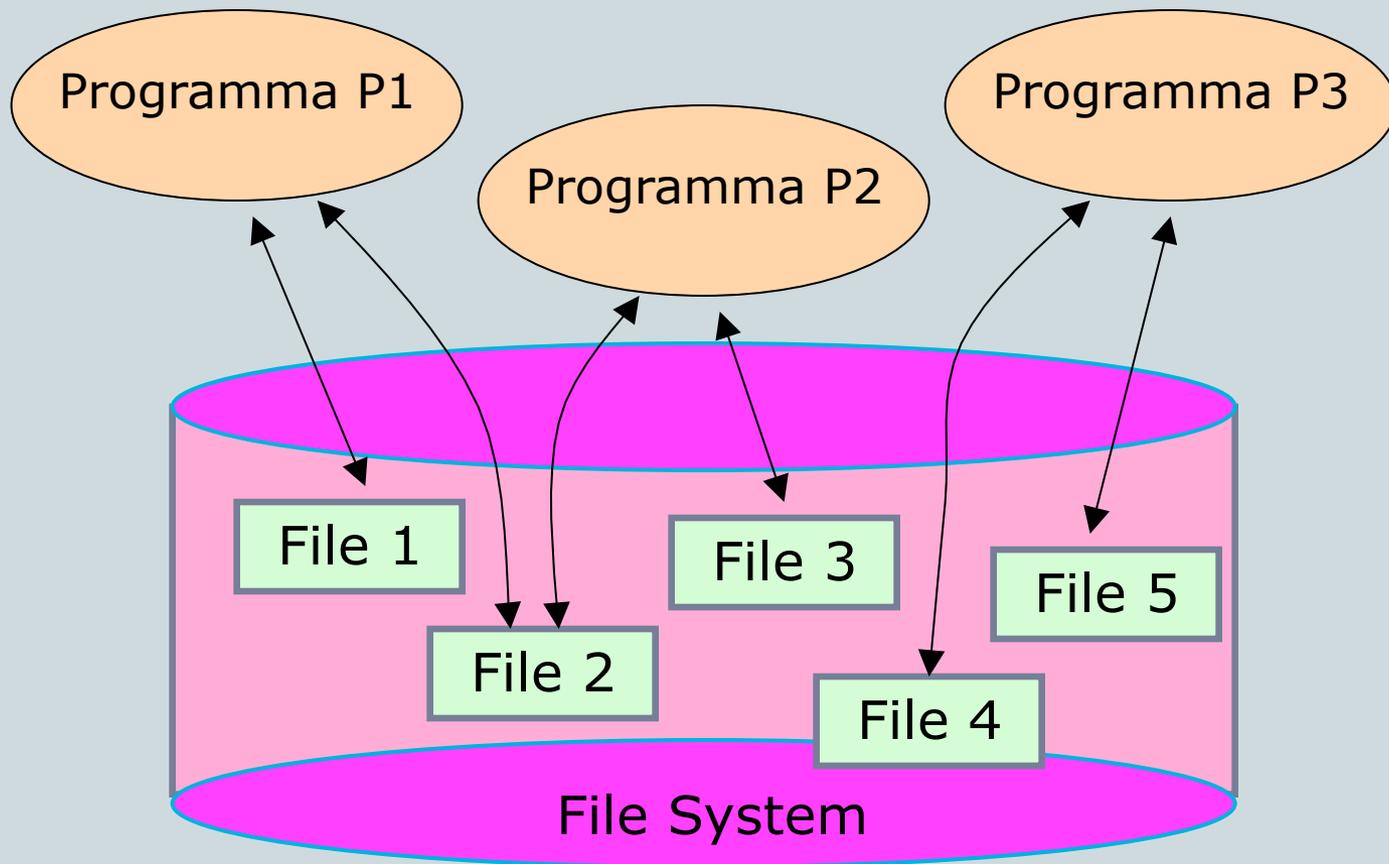
### **Definizione**

“è una collezione di dati utilizzati per rappresentare con tecnologia informatica le informazioni di interesse per un sistema informativo”

# Applicazioni gestionali

18

## Soluzione convenzionale



# Applicazioni gestionali

19

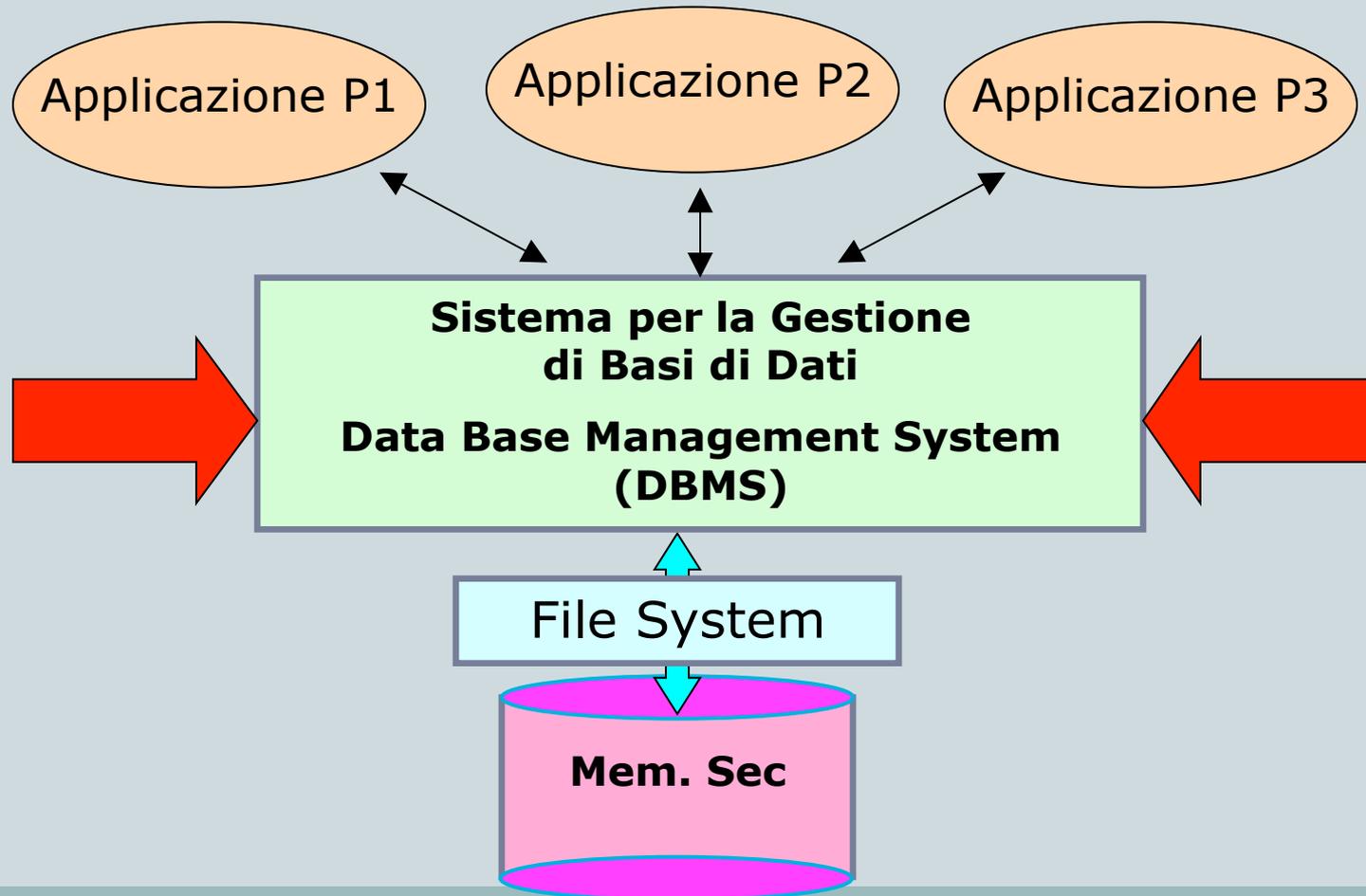
## **Problemi della soluzione convenzionale**

- Scarsa efficienza nell'accesso ai dati su file (struttura ad accesso sequenziale)
- Ridondanza nei dati (duplicazioni dello stesso dato su più file)
- Inconsistenza (aggiornamenti parziali)
- Progettazione dei dati replicata per ogni programma

# Applicazioni gestionali

20

## Soluzione innovativa (anni '70 - '80)



# DBMS (Data Base Management System)

21

## **DBMS**

### **Definizione**

È un sistema che gestisce su memoria secondaria collezioni di dati (chiamate “Basi di Dati”):

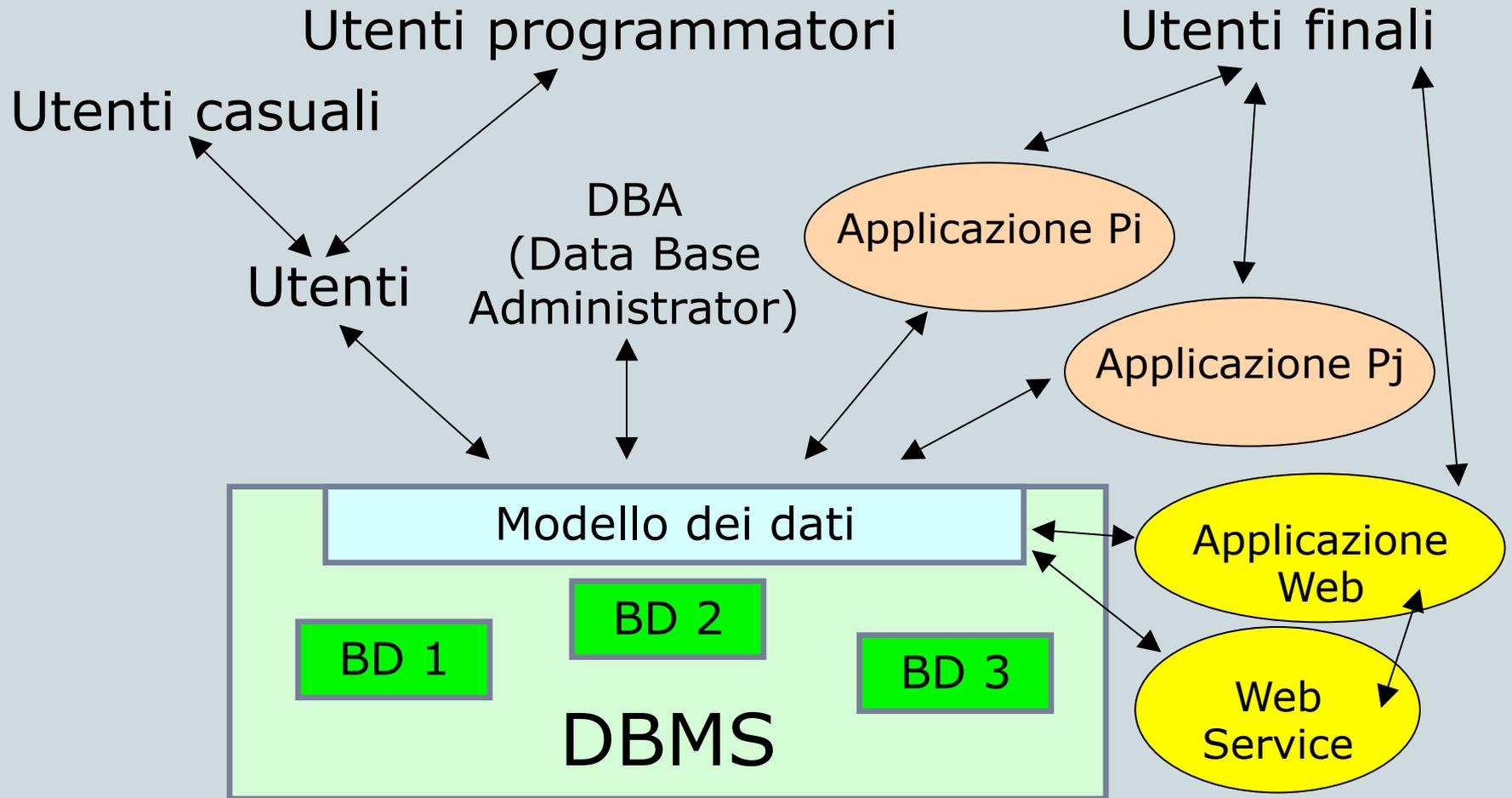
- GRANDI, CONDIVISE e PERSISTENTI

assicurando:

- AFFIDABILITA', PRIVATEZZA e ACCESSO EFFICIENTE

# DBMS: interazione

22



# DBMS: interazione

23

## Linguaggi di interazione:

- Linguaggio per la definizione dei dati (Data Definition Language - DDL)
- Linguaggio per l'interrogazione e aggiornamento dei dati (Data Manipulation Language – DML):
  - Linguaggio di interrogazione: estrae informazioni da una base di dati (esempio: SQL, algebra relazionale)
  - Linguaggio di manipolazione: popola la base di dati, modifica il suo contenuto con aggiunte, cancellazioni e variazioni sui dati (esempio: SQL)

# DBMS: modello dei dati

24

## Definizione

È l'insieme dei **costrutti** forniti dal DBMS per descrivere la **struttura** e le **proprietà** dell'informazione contenuta in una base di dati.

Costrutti: permettono

- di definire le strutture dati che conterranno le informazioni della base di dati (analogia con i costruttori di tipo di un linguaggio di programmazione) e
- di specificare le proprietà che dovranno soddisfare le istanze di informazione che saranno contenute nelle strutture dati

# DBMS: modello dei dati

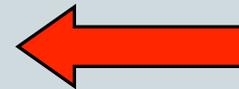
25

## Modelli dei dati del passato

- Modello reticolare
- Modello gerarchico

## Modelli dei dati attuali

- Modello relazionale (E.F. Codd 1970)
- Modello ad oggetti
- Modello object-relational



# Modello, schema e istanza

26

**È fondamentale nel contesto delle basi di dati distinguere correttamente questi tre concetti:**

- **Modello dei dati:** visto nel lucido precedente
- **Schema di una base di dati:** è la descrizione della struttura e delle proprietà di una specifica base di dati fatta utilizzando i costrutti del modello dei dati (lo schema di una base di dati è invariante nel tempo)
- **Istanza di una base di dati:** è costituita dai valori effettivi che in un certo istante popolano le strutture dati della base di dati (l'istanza di una base di dati varia nel tempo)

# Modello, schema e istanza

27

Esempio

Modello dei dati

Basi di dati

Tabella  
(o relazione)

Schema

```
P(cognome: VARCHAR(40),  
nome: VARCHAR(30))
```

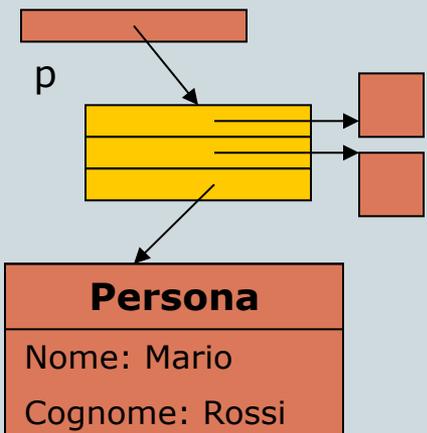
Istanza

cognome	nome
Rossi	Mario
Bianchi	Lia

Linguaggi di  
progr.

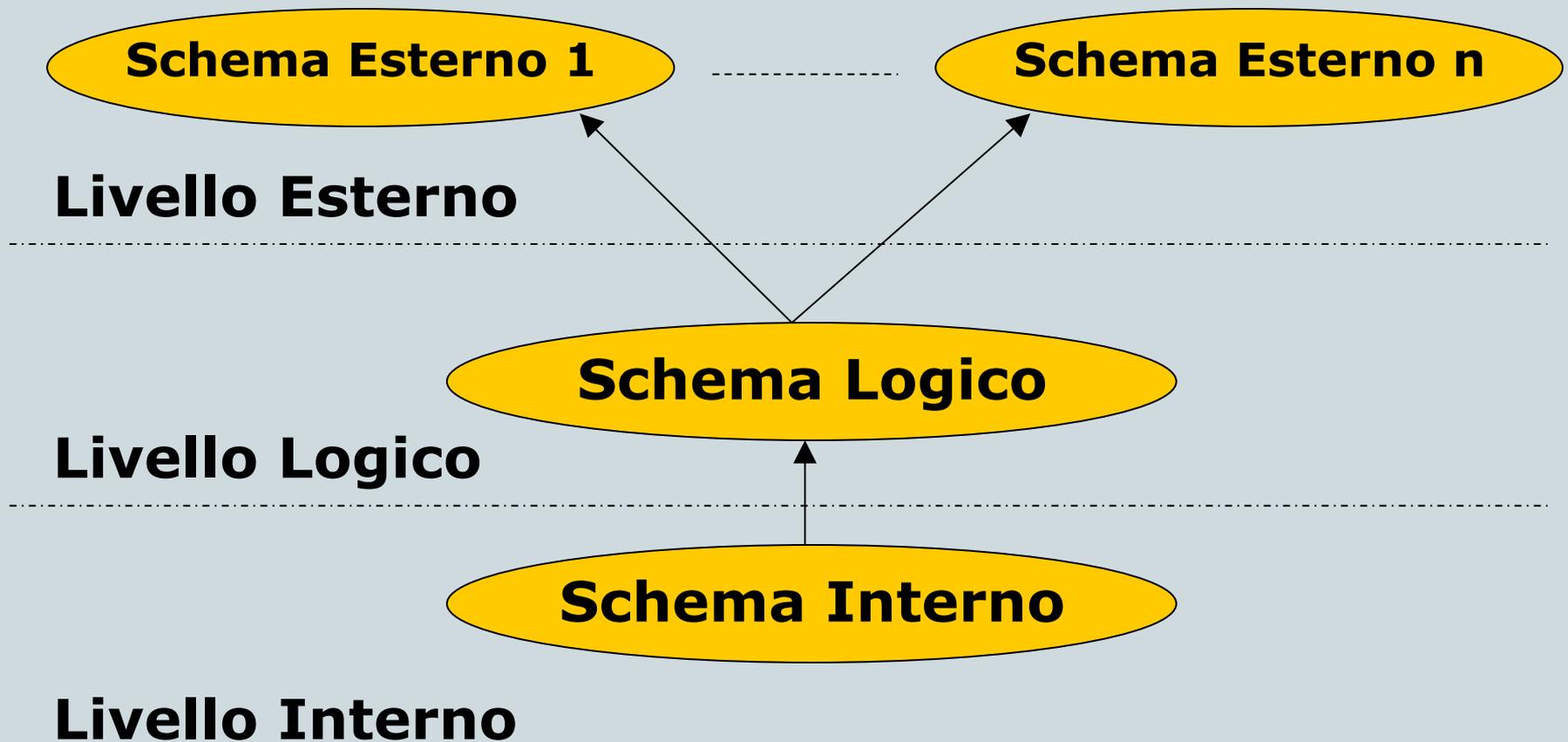
Array

```
Class Persona {  
    String cognome;  
    String nome; }  
  
Class X {  
    ...  
    Persona[] p;  
    p = new Persona[100];  
}
```



# Architettura di un DBMS (schema based)

28



# Architettura di un DBMS

29

- **Schema Logico:** è la rappresentazione della struttura e delle proprietà della base di dati attraverso i costrutti del modello dei dati del DBMS
- **Schema Interno:** è la rappresentazione della base di dati per mezzo delle strutture fisiche di memorizzazione (file dati, file indice, ecc...)
- **Schema Esterno:** descrive una porzione dello schema logico di interesse per uno specifico utente o applicazione (attraverso viste sullo schema logico)

# Indipendenza dei dati

30

**Indipendenza FISICA:** lo schema logico della base di dati è completamente indipendente dallo schema interno.

## *Conseguenza*

variazioni delle strutture fisiche non impattano sullo schema logico e quindi sulle applicazioni

# Indipendenza dei dati

31

**Indipendenza LOGICA:** gli schemi esterni della base di dati sono indipendenti dallo schema logico.

## *Conseguenza*

variazioni dello schema logico (purché non tolgano dati) non impattano sugli schemi esterni e quindi sulle applicazioni (va eventualmente ridefinita solo l'espressione di derivazione degli schemi esterni dallo schema logico "viste")