

**Atzeni, Ceri, Paraboschi, Torlone**  
**Basi di dati**

**McGraw-Hill, 1996-2002**

**adattato da Gianluca Amato per il**

**Master in Catalogazione di Beni**  
**Archivistici e Librari**

**28/05/2003**

# Sistemi Informativi

# Ruolo centrale delle Informazioni

- Nello svolgimento di ogni attività, sia a livello individuale sia in grandi organizzazioni, sono essenziali la disponibilità di **informazioni** e la capacità di gestirle in modo efficace.

# Sistema informativo

- **Componente (sottosistema) di una organizzazione che gestisce (acquisisce, elabora, conserva, produce) le informazioni di interesse (cioè utilizzate per il perseguimento degli scopi dell'organizzazione )**

# Sistema informativo, commenti

- **Ogni organizzazione ha un sistema informativo, eventualmente non esplicitato nella struttura**
- **Quasi sempre, il sistema informativo è di supporto ad altri sottosistemi, e va quindi studiato nel contesto in cui è inserito**

# Sistemi informativi e automazione

- Il concetto di “sistema informativo” è indipendente da qualsiasi automatizzazione:
  - esistono organizzazioni la cui ragion d’essere è la gestione di informazioni (p. es. servizi anagrafici e banche) e che operano da secoli

# Sistema Informatico

- **porzione automatizzata del sistema informativo:**

**la parte del sistema informativo che gestisce informazioni con tecnologia informatica**

# Sistema Informatico

**Sistema azienda**

**Sistema informativo**

**Sistema informatico**

# Gestione delle informazioni

- **Raccolta, acquisizione**
- **Archiviazione, conservazione**
- **Elaborazione, trasformazione, produzione**
- **Distribuzione, comunicazione, scambio**

# Gestione delle informazioni

- **Nelle attività umane, le informazioni vengono gestite in forme diverse:**
  - **idee informali**
  - **linguaggio naturale (scritto o parlato, formale o colloquiale, in varie lingue)**
  - **disegni, grafici, schemi**
  - **numeri e codici**
- **e su vari supporti**
  - **mente umana, carta, dispositivi elettronici**

# Informazioni e dati

- Nei sistemi informatici (e non solo), le **informazioni** vengono rappresentate in modo essenziale, spartano: attraverso i **dati**

# Informazioni e dati

(definizioni dal Vocabolario della lingua italiana 1987)

**informazione:** notizia, dato o elemento che consente di avere conoscenza più o meno esatta di fatti, situazioni, modi di essere.

**dato:** ciò che è immediatamente presente alla conoscenza, prima di ogni elaborazione; (in informatica) elementi di informazione costituiti da simboli che debbono essere elaborati

# Dati e informazioni

- Un esempio:

Mario 275

su un foglio di carta sono due **dati** e non significano molto

- Se il foglio di carta viene fornito in risposta alla domanda “A chi mi devo rivolgere per il problema X; qual è il suo numero di telefono?”, allora i dati possono essere interpretati per fornire **informazione** e arricchire la conoscenza

# Importanza dei dati

- **I dati costituiscono spesso una risorsa strategica, perché più stabili nel tempo di altre componenti (processi, tecnologie, ruoli umani):**
  - **ad esempio, i dati delle banche o delle anagrafi**

# Base di dati

(accezione generica, **metodologica**)

- **Insieme organizzato di dati utilizzati per il supporto allo svolgimento delle attività di un ente (azienda, ufficio, persona)**

(accezione specifica, **metodologica e tecnologica**)

- **insieme di dati gestito da un DBMS**

# Sistema di gestione di basi di dati DataBase Management System (DBMS)

- **Sistema** che gestisce collezioni di dati:
  - **grandi**
  - **persistenti**
  - **condivise****garantendo**
  - **privatezza**
  - **affidabilità**
  - **efficienza**
  - **efficacia**

# DBMS

- **Prodotti software (complessi) disponibili sul mercato; esempi:**
  - **DB2**
  - **Oracle**
  - **Informix**
  - **MySQL**
  - **Sybase**
  - **SQLServer**



# Le basi di dati sono ... grandi

- **dimensioni (molto) maggiori della memoria centrale dei sistemi di calcolo utilizzati**
- **il limite deve essere solo quello fisico dei dispositivi di massa, come dischi o nastri magnetici**



# Le basi di dati sono ... persistenti

- hanno un tempo di vita indipendente dalle singole esecuzioni dei programmi che le utilizzano
- quando i programmi che usano una base di dati terminano, la base di dati continua ad esistere e può essere riutilizzata di nuovo successivamente.



# Le basi di dati sono ... condivise

- Ogni organizzazione (specie se grande) è divisa in settori o comunque svolge diverse attività
- Ciascun settore/attività ha un (sotto)sistema informativo (non necessariamente disgiunto)

CCS Ingegneria Informatica Orario - Netscape  
 File Edit View Go Communicator Help

**UNIVERSITA' DEGLI STUDI DI CHISSADOVE**

**Corso di Studi in Ingegneria Informatica**

---

**ORARIO DELLE LEZIONI PER L'ANNO  
 ACCADEMICO 1999-2000**

INSEGNAMENTO	Docente	Aula	Orario
Analisi matematica I	Luigi Neri	N1	8:00-9:30
Basi di dati	Piero Rossi	N2	9:45-11:15
Chimica	Nicola Mori	N1	9:45-11:30
Fisica I	Mario Bruni	N1	11:45-13:00
Fisica II	Mario Bruni	N3	9:45-11:15
Sistemi informativi	Piero Rossi	N3	8:00-9:30

Document: Done

Orari di ricevimento - Netscape

File Edit View Go Communicator Help

# UNIVERSITA' DEGLI STUDI CHISSADOVE

## Corso di Studi in Ingegneria Informatica

---

### Orario di ricevimento dei docenti

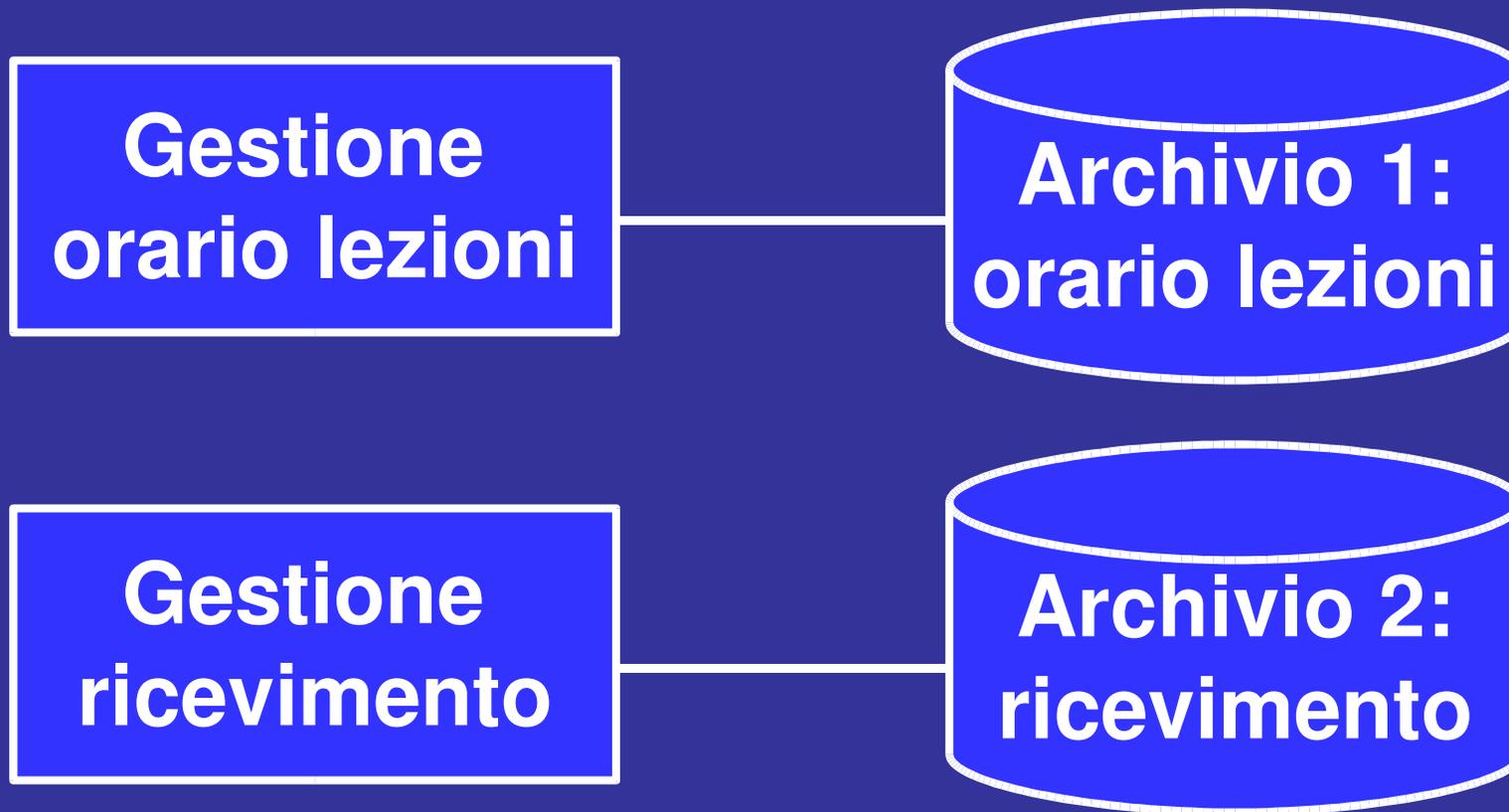
DOCENTE	INSEGNAMENTI	ORARIO
Mario BRUNI	Fisica I Fisica II	Martedì' 10-12
Luigi NERI	Analisi matematica I	Lunedì' 12-13
Piero ROSSI	Basi di dati Sistemi informativi	Giovedì' 11-13
Nicola MORI	Chimica	Martedì' 16-18

Document: Done

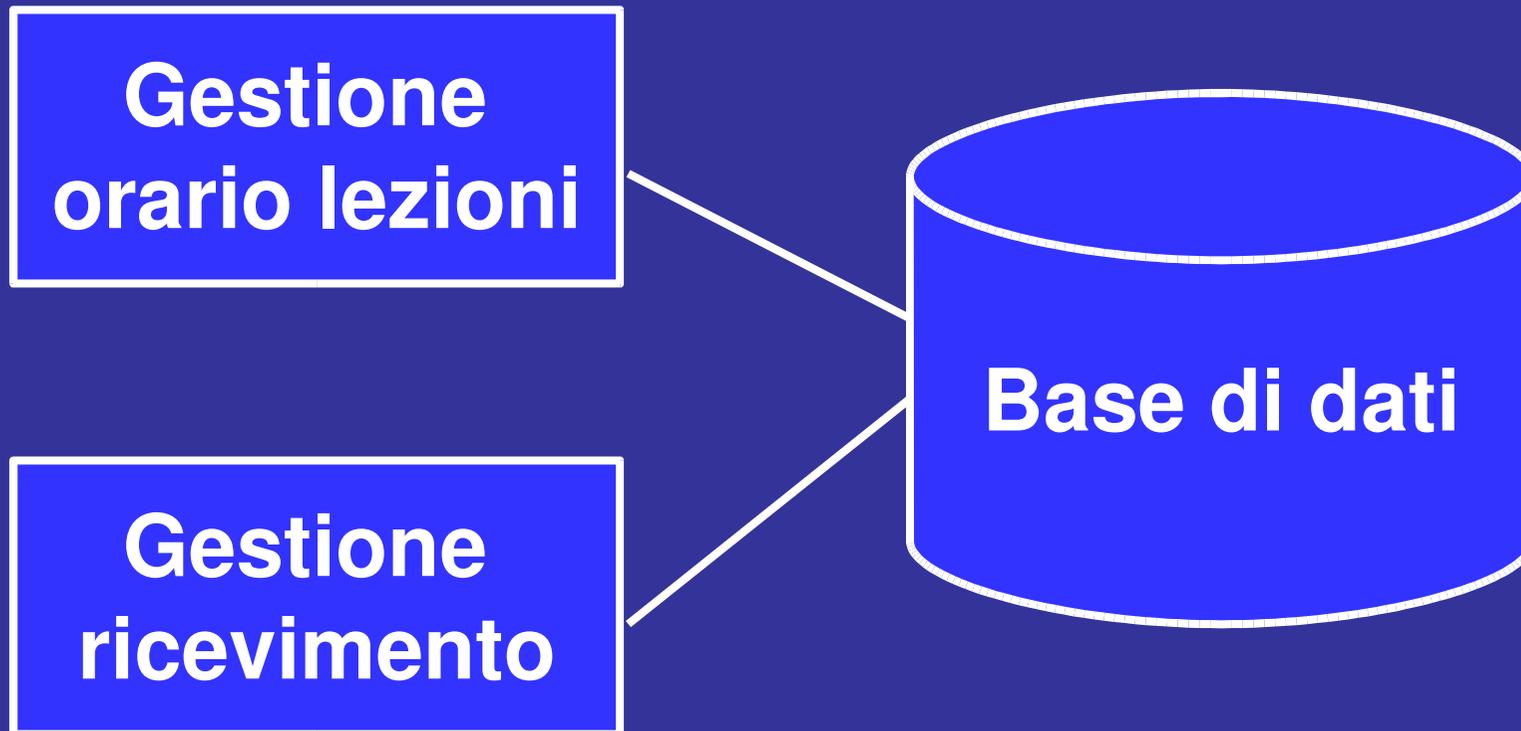
# Problemi

- **Ridondanza:**
  - informazioni ripetute
- **Rischio di incoerenza:**
  - le versioni possono non coincidere

# Archivi e basi di dati



# Archivi e basi di dati



# Le basi di dati sono condivise

- Una base di dati e' una risorsa **integrata, condivisa** fra applicazioni
- conseguenze
  - Attivita' diverse su dati condivisi:
    - meccanismi di **autorizzazione**
  - Accessi di più utenti ai dati condivisi:
    - controllo della **concorrenza**



# I DBMS garantiscono ... privatezza

- **Si possono definire meccanismi di autorizzazione**
  - **l'utente A è autorizzato a leggere tutti i dati e a modificare quelli sul ricevimento**
  - **l'utente B è autorizzato a leggere i X e a modificare Y**



# I DBMS garantiscono... affidabilità

- **Affidabilità** (per le basi di dati):
  - resistenza a malfunzionamenti hardware e software
- Una base di dati è una risorsa pregiata e quindi deve essere conservata a lungo termine

# I DBMS debbono essere...efficienti

- Cercano di utilizzare al meglio le risorse di spazio di memoria (principale e secondaria) e tempo (di esecuzione e di risposta)
- I DBMS, con tante funzioni, rischiano l'inefficienza e per questo ci sono grandi investimenti e competizione
- L'efficienza è anche il risultato della qualità delle applicazioni



# I DBMS debbono essere...efficaci

- Cercano di rendere produttive le attività dei loro utilizzatori, offrendo funzionalità articolate, potenti e flessibili.
- L'efficienza è un concetto molto generico che dipende anche dall'applicazione che si considera.

# DBMS vs file system

- La gestione di insiemi di dati grandi e persistenti è possibile anche attraverso sistemi più semplici — gli ordinari **file system** dei sistemi operativi
- I file system prevedono forme rudimentali di condivisione: "tutto o niente"
- I DBMS estendono le funzionalità dei file system, fornendo più servizi ed in maniera integrata

# Modello Relazionale

# Descrizioni dei dati nei DBMS

- Un DBMS deve poter rappresentare vari tipi di informazione, a seconda dell'applicazione.
- Nasce quindi l'esigenza di un **modello logico dei dati**, ovvero di un insieme di concetti utilizzati per organizzare i dati di interesse e descriverne la struttura.
- Il modello logico più diffuso è quello **relazionale**.

CCS Ingegneria Informatica Orario - Netscape

File Edit View Go Communicator Help

**UNIVERSITA' DEGLI STUDI DI CHISSADOVE**

**Corso di Studi in Ingegneria Informatica**

---

**ORARIO DELLE LEZIONI PER L'ANNO  
ACCADEMICO 1999-2000**

INSEGNAMENTO	Docente	Aula	Orario
Analisi matematica I	Luigi Neri	N1	8:00-9:30
Basi di dati	Piero Rossi	N2	9:45-11:15
Chimica	Nicola Mori	N1	9:45-11:30
Fisica I	Mario Bruni	N1	11:45-13:00
Fisica II	Mario Bruni	N3	9:45-11:15
Sistemi informativi	Piero Rossi	N3	8:00-9:30

Document: Done

# Organizzazione dei dati in una base di dati

## Orario

Insegnamento	Docente	Aula	Ora
Analisi matem. I	Luigi Neri	N1	8:00
Basi di dati	Piero Rossi	N2	9:45
Chimica	Nicola Mori	N1	9:45
Fisica I	Mario Bruni	N1	11:45
Fisica II	Mario Bruni	N3	9:45
Sistemi inform.	Piero Rossi	N3	8:00

# Basi di dati: schema e istanza

## Lo **schema** della base di dati

### Orario

Insegnamento	Docente	Aula	Ora
Analisi matem. I	Luigi Neri	N1	8:00
Basi di dati	Piero Rossi	N2	9:45
Chimica	Nicola Mori	N1	9:45
Fisica I	Mario Bruni	N1	11:45
Fisica II	Mario Bruni	N3	9:45
Sistemi inform.	Piero Rossi	N3	8:00

## L'**istanza** della base di dati

# Schema e istanza

- In ogni base di dati esistono:
  - lo **schema**, sostanzialmente invariante nel tempo, che ne descrive la struttura (aspetto intensionale)
    - es.: le intestazioni delle tabelle
  - l'**istanza**, i valori attuali, che possono cambiare anche molto rapidamente (aspetto estensionale)
    - es.: il “corpo” di ciascuna tabella

Ex

# Il modello è basato su valori

- i riferimenti fra dati in relazioni diverse sono rappresentati per mezzo di valori dei domini che compaiono nelle ennuple

**studenti**

Matricola	Cognome	Nome	Data di nascita
6554	Rossi	Mario	05/12/1978
8765	Neri	Paolo	03/11/1976
9283	Verdi	Luisa	12/11/1979
3456	Rossi	Maria	01/02/1978

**esami**

Studente	Voto	Corso
3456	30	04
3456	24	02
9283	28	01
6554	26	01

**corsi**

Codice	Titolo	Docente
01	Analisi	Mario
02	Chimica	Bruni
04	Chimica	Verdi

**studenti**

Matricola	Cognome	Nome	Data di nascita
6554	Rossi	Mario	05/12/1978
8765	Neri	Paolo	03/11/1976
9283	Verdi	Luisa	12/11/1979
3456	Rossi	Maria	01/02/1978

**esami**

Studente	Voto	Corso
	30	
	24	
	28	
	26	

**corsi**

Codice	Titolo	Docente
01	Analisi	Mario
02	Chimica	Bruni
04	Chimica	Verdi

# Struttura basata su valori: vantaggi

- **indipendenza dalle strutture fisiche (si potrebbe avere anche con puntatori di alto livello) che possono cambiare dinamicamente**
- **si rappresenta solo ciò che è rilevante dal punto di vista dell'applicazione**
- **l'utente finale vede gli stessi dati dei programmatori**
- **i dati sono portabili piu' facilmente da un sistema ad un altro**
- **i puntatori sono direzionali**

# Relazioni su singoli attributi

## studenti

Matricola	Cognome	Nome	Data di nascita
6554	Rossi	Mario	05/12/1978
8765	Neri	Paolo	03/11/1976
9283	Verdi	Luisa	12/11/1979
3456	Rossi	Maria	01/02/1978

## studenti lavoratori

Matricola
6554
3456

# Strutture nidificate

<i>Da Filippo Via Roma 2, Roma</i>		
<i>Ricevuta Fiscale 1235 del 12/10/2000</i>		
<b>3</b>	<b>Coperti</b>	<b>3,00</b>
<b>2</b>	<b>Antipasti</b>	<b>6,20</b>
<b>3</b>	<b>Primi</b>	<b>12,00</b>
<b>2</b>	<b>Bistecche</b>	<b>18,00</b>
<b><i>Totale</i></b>		<b>39,20</b>

<i>Da Filippo Via Roma 2, Roma</i>		
<i>Ricevuta Fiscale 1240 del 13/10/2000</i>		
<b>2</b>	<b>Coperti</b>	<b>2,00</b>
<b>2</b>	<b>Antipasti</b>	<b>7,00</b>
<b>2</b>	<b>Primi</b>	<b>8,00</b>
<b>2</b>	<b>Orate</b>	<b>20,00</b>
<b>2</b>	<b>Caffè</b>	<b>2,00</b>
<b><i>Totale</i></b>		<b>39,00</b>

# Relazioni che rappresentano strutture nidificate

## Ricevute

Numero	Data	Totale
1235	12/10/2000	39,20
1240	13/10/2000	39,00

## Dettaglio

Numero	Qtà	Descrizione	Importo
1235	3	Coperti	3,00
1235	2	Antipasti	6,20
1235	3	Primi	12,00
1235	2	Bistecche	18,00
1240	2	Coperti	2,00
...	...	...	...

# Strutture nidificate, riflessione

- Abbiamo rappresentato veramente tutti gli aspetti delle ricevute?
- Dipende da che cosa ci interessa!
  - l'ordine delle righe e' rilevante?
  - possono esistere linee ripetute in una ricevuta?
- Sono possibili rappresentazioni diverse

# Rappresentazione alternativa per strutture nidificate

## Ricevute

Numero	Data	Totale
1235	12/10/2000	39,20
1240	13/10/2000	39,00

## Dettaglio

Numero	Riga	Qtà	Descrizione	Importo
1235	1	3	Coperti	3,00
1235	2	2	Antipasti	6,20
1235	3	3	Primi	12,00
1235	4	2	Bistecche	18,00
1240	1	2	Coperti	2,00
...	...	...	...	...

# Informazione incompleta

- Il modello relazionale impone ai dati una struttura rigida:
  - le informazioni sono rappresentate per mezzo di ennuple
  - solo alcuni formati di ennuple sono ammessi: quelli che corrispondono agli schemi di relazione
- I dati disponibili possono non corrispondere al formato previsto

# Informazione incompleta: motivazioni

<b>Nome</b>	<b>SecondoNome</b>	<b>Cognome</b>
<b>Franklin</b>	<b>Delano</b>	<b>Roosevelt</b>
<b>Winston</b>		<b>Churchill</b>
<b>Charles</b>		<b>De Gaulle</b>
<b>Josip</b>		<b>Stalin</b>

# Informazione incompleta: soluzioni?

- non conviene (anche se spesso si fa) usare valori del dominio (0, stringa nulla, “99”, ...):
  - potrebbero non esistere valori “non utilizzati”
  - valori “non utilizzati” potrebbero diventare significativi
  - in fase di utilizzo (nei programmi) sarebbe necessario ogni volta tener conto del “significato” di questi valori

# Informazione incompleta nel modello relazionale

- Tecnica rudimentale ma efficace:
  - **valore nullo**: denota l'assenza di un valore del dominio (e non è un valore del dominio)
- Si possono (e debbono) imporre restrizioni sulla presenza di valori nulli

# Troppi valori nulli

**studenti**

Matricola	Cognome	Nome	Data di nascita
6554	Rossi	Mario	05/12/1978
9283	Verdi	Luisa	12/11/1979
NULL	Rossi	Maria	01/02/1978

**esami**

Studente	Voto	Corso
NULL	30	NULL
NULL	24	02
9283	28	01

**corsi**

Codice	Titolo	Docente
01	Analisi	Mario
02	NULL	NULL
04	Chimica	Verdi

# Vincoli di integrità

- **Esistono istanze di basi di dati che, pur sintatticamente corrette, non rappresentano informazioni possibili per l'applicazione di interesse**

# Una base di dati "scorretta"

Esami	Studente	Voto	Lode	Corso
	276545	32		01
	276545	30	e lode	02
	787643	27	e lode	03
	739430	24		04

Studenti	Matricola	Cognome	Nome
	276545	Rossi	Mario
	787643	Neri	Piero
	787643	Bianchi	Luca

# Vincolo di integrità

- Proprietà che deve essere soddisfatta dalle istanze che rappresentano informazioni corrette per l'applicazione
- Un vincolo è una funzione booleana (un **predicato**): associa ad ogni istanza il valore **vero** o **falso**

# Vincoli di integrità, perché?

- descrizione più accurata della realtà
- contributo alla “qualità dei dati”
- usati dai DBMS nella esecuzione delle interrogazioni

# Tipi di vincoli

- vincoli **intrarelazionali**
  - vincoli su valori (o di **dominio**)
  - vincoli di **ennupla**
- vincoli **interrelazionali**

# Una base di dati "scorretta"

Esami	Studente	Voto	Lode	Corso
	276545	32		01
	276545	30	e lode	02
	787643	27	e lode	03
	739430	24		04

Studenti	Matricola	Cognome	Nome
	276545	Rossi	Mario
	787643	Neri	Piero
	787643	Bianchi	Luca

# Vincoli di ennupla

- Esprimono condizioni sui valori di ciascuna ennupla, indipendentemente dalle altre ennuple
- Caso particolare:
  - Vincoli di dominio: coinvolgono un solo attributo

# Sintassi ed esempi

- Una possibile sintassi:
  - espressione booleana di atomi che confrontano valori di attributo o espressioni aritmetiche su di essi

**(Voto  $\geq$  18) AND (Voto  $\leq$  30)**

**(Voto = 30) OR NOT (Lode = "e lode")**

# Vincoli di ennupla, esempio

<b>Stipendi</b>	<b>Impiegato</b>	<b>Lordo</b>	<b>Ritenute</b>	<b>Netto</b>
	<b>Rossi</b>	<b>55.000</b>	<b>12.500</b>	<b>42.500</b>
	<b>Neri</b>	<b>45.000</b>	<b>10.000</b>	<b>35.000</b>
	<b>Bruni</b>	<b>47.000</b>	<b>11.000</b>	<b>36.000</b>

$$\text{Lordo} = (\text{Ritenute} + \text{Netto})$$

# Identificazione delle ennuple

Matricola	Cognome	Nome	Corso	Nascita
27655	Rossi	Mario	Ing Inf	5/12/78
78763	Rossi	Mario	Ing Inf	3/11/76
65432	Neri	Piero	Ing Mecc	10/7/79
87654	Neri	Mario	Ing Inf	3/11/76
67653	Rossi	Piero	Ing Mecc	5/12/78

- non ci sono due ennuple con lo stesso valore sull'attributo **Matricola**
- non ci sono due ennuple uguali su tutti e tre gli attributi **Cognome, Nome e Data di Nascita**

# Chiave

- insieme di attributi che identificano le ennuple di una relazione

Formalmente:

- un insieme  $K$  di attributi è **superchiave** per  $r$  se  $r$  non contiene due ennuple distinte  $t_1$  e  $t_2$  con  $t_1[K] = t_2[K]$
- $K$  è **chiave** per  $r$  se è una superchiave minimale per  $r$  (cioè non contiene un'altra superchiave)

# Una chiave

Matricola	Cognome	Nome	Corso	Nascita
27655	Rossi	Mario	Ing Inf	5/12/78
78763	Rossi	Mario	Ing Inf	3/11/76
65432	Neri	Piero	Ing Mecc	10/7/79
87654	Neri	Mario	Ing Inf	3/11/76
67653	Rossi	Piero	Ing Mecc	5/12/78

- **Matricola è una chiave:**
  - è superchiave
  - contiene un solo attributo e quindi è minimale

# Un'altra chiave

Matricola	Cognome	Nome	Corso	Nascita
27655	Rossi	Mario	Ing Inf	5/12/78
78763	Rossi	Mario	Ing Inf	3/11/76
65432	Neri	Piero	Ing Mecc	10/7/79
87654	Neri	Mario	Ing Inf	3/11/76
67653	Rossi	Piero	Ing Mecc	5/12/78

- **Cognome, Nome, Nascita è un'altra chiave:**
  - è superchiave
  - minimale

## Un'altra chiave??

Matricola	Cognome	Nome	Corso	Nascita
27655	Rossi	Mario	Ing Inf	5/12/78
78763	Rossi	Mario	Ing Civile	3/11/76
65432	Neri	Piero	Ing Mecc	10/7/79
87654	Neri	Mario	Ing Inf	3/11/76
67653	Rossi	Piero	Ing Mecc	5/12/78

- Non ci sono ennuple uguali su Cognome e Corso:
  - Cognome e Corso formano una chiave
- Ma è sempre vero?

# Vincoli, schemi e istanze

- i vincoli corrispondono a proprietà del mondo reale modellato dalla base di dati
- interessano a livello di schema (con riferimento cioè a tutte le istanze)
- ad uno schema associamo un insieme di vincoli e consideriamo **corrette** (valide, ammissibili) le istanze che soddisfano tutti i vincoli
- un'istanza può soddisfare altri vincoli (“per caso”)

# Studenti

**Matricola Cognome Nome Corso Nascita**

- **chiavi:**

**Matricola**  
**Cognome, Nome, Nascita**

Matricola	Cognome	Nome	Corso	Nascita
27655	Rossi	Mario	Ing Inf	5/12/78
78763	Rossi	Mario	Ing Civile	3/11/76
65432	Neri	Piero	Ing Mecc	10/7/79
87654	Neri	Mario	Ing Inf	3/11/76
67653	Rossi	Piero	Ing Mecc	5/12/78

- È corretta: soddisfa i vincoli
- Ne soddisfa anche altri ("per caso"):
  - **Cognome, Corso** è chiave

# Importanza delle chiavi

- l'esistenza delle chiavi garantisce l'accessibilità a ciascun dato della base di dati
- le chiavi permettono di correlare i dati in relazioni diverse:
  - il modello relazionale è basato su valori

# Chiavi e valori nulli

- **In presenza di valori nulli, i valori della chiave non permettono**
  - **di identificare le ennuple**
  - **di realizzare facilmente i riferimenti da altre relazioni**

<b>Matricola</b>	<b>Cognome</b>	<b>Nome</b>	<b>Corso</b>	<b>Nascita</b>
<b>NULL</b>	<b>NULL</b>	<b>Mario</b>	<b>Ing Inf</b>	<b>5/12/78</b>
<b>78763</b>	<b>Rossi</b>	<b>Mario</b>	<b>Ing Civile</b>	<b>3/11/76</b>
<b>65432</b>	<b>Neri</b>	<b>Piero</b>	<b>Ing Mecc</b>	<b>10/7/79</b>
<b>87654</b>	<b>Neri</b>	<b>Mario</b>	<b>Ing Inf</b>	<b>NULL</b>
<b>NULL</b>	<b>Neri</b>	<b>Mario</b>	<b>NULL</b>	<b>5/12/78</b>

- **La presenza di valori nulli nelle chiavi deve essere limitata**

# Chiave primaria

- Chiave su cui non sono ammessi nulli
- Notazione: sottolineatura

<u>Matricola</u>	Cognome	Nome	Corso	Nascita
86765	NULL	Mario	Ing Inf	5/12/78
78763	Rossi	Mario	Ing Civile	3/11/76
65432	Neri	Piero	Ing Mecc	10/7/79
87654	Neri	Mario	Ing Inf	NULL
43289	Neri	Mario	NULL	5/12/78

# Integrità referenziale

- informazioni in relazioni diverse sono correlate attraverso valori comuni
- in particolare, valori delle chiavi (primarie)
- le correlazioni debbono essere "coerenti"

# Infrazioni

<u>Codice</u>	Data	Vigile	Prov	Numero
34321	1/2/95	3987	MI	39548K
53524	4/3/95	3295	TO	E39548
64521	5/4/96	3295	PR	839548
73321	5/2/98	9345	PR	839548

## Vigili

<u>Matricola</u>	Cognome	Nome
3987	Rossi	Luca
3295	Neri	Piero
9345	Neri	Mario
7543	Mori	Gino

# Infrazioni

<u>Codice</u>	<u>Data</u>	<u>Vigile</u>	<u>Prov</u>	<u>Numero</u>
34321	1/2/95	3987	MI	39548K
53524	4/3/95	3295	TO	E39548
64521	5/4/96	3295	PR	839548
73321	5/2/98	9345	PR	839548

# Auto

<u>Prov</u>	<u>Numero</u>	<u>Cognome</u>	<u>Nome</u>
MI	39548K	Rossi	Mario
TO	E39548	Rossi	Mario
PR	839548	Neri	Luca

# Vincolo di integrità referenziale

- Un vincolo di **integrità referenziale** (“**foreign key**”) fra gli attributi  $X$  di una relazione  $R_1$  e un'altra relazione  $R_2$  impone ai valori su  $X$  in  $R_1$  di comparire come valori della chiave primaria di  $R_2$

- **vincoli di integrità referenziale fra:**
  - **l'attributo Vigile della relazione INFRAZIONI e la relazione VIGILI**
  - **gli attributi Prov e Numero di INFRAZIONI e la relazione AUTO**

# Violazione di vincolo di integrità referenziale

## Infrazioni

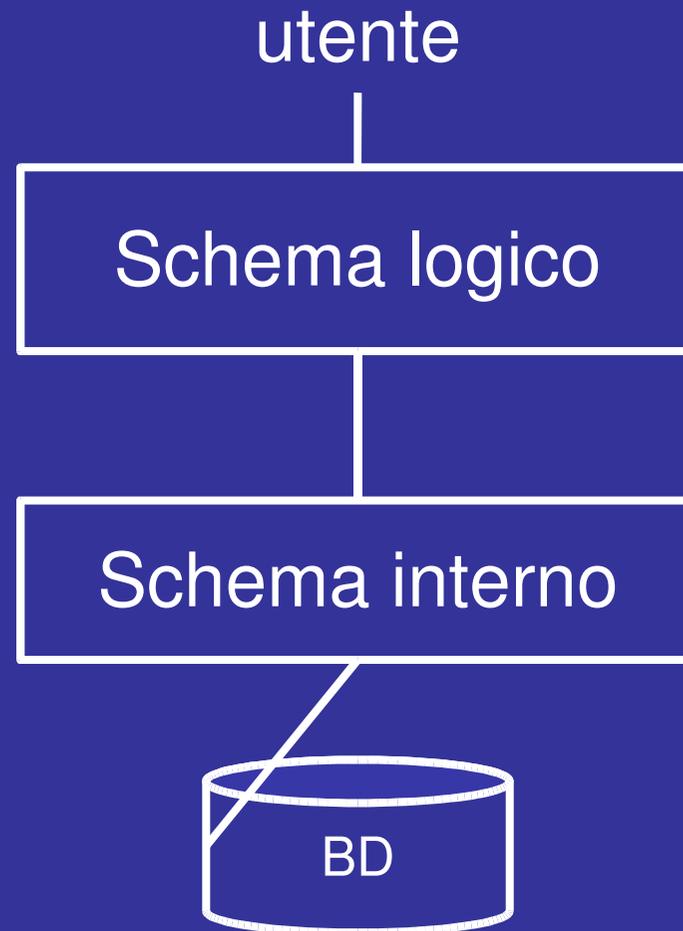
<u>Codice</u>	Data	Vigile	Prov	Numero
34321	1/2/95	3987	MI	39548K
53524	4/3/95	3295	TO	E39548
64521	5/4/96	3295	PR	839548
73321	5/2/98	9345	PR	839548

## Auto

<u>Prov</u>	<u>Numero</u>	Cognome	Nome
MI	E39548	Rossi	Mario
TO	F34268	Rossi	Mario
PR	839548	Neri	Luca

# Indipendenza dei dati

# Architettura (semplificata) di un DBMS



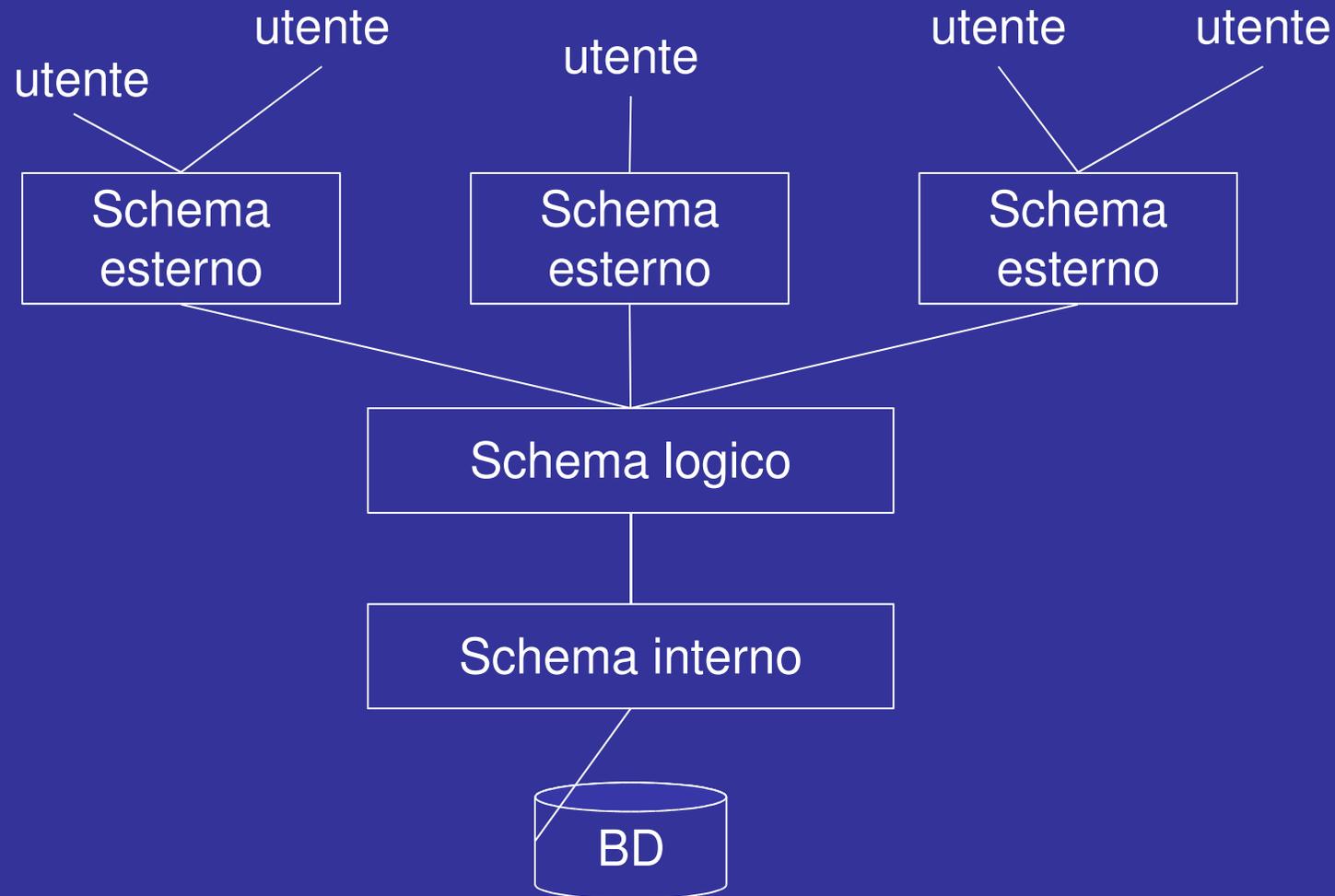
# Architettura semplificata di un DBMS: schemi

- **schema logico**: descrizione della base di dati nel modello logico (ad esempio, la struttura della tabella)
- **schema interno (o fisico)**: rappresentazione dello schema logico per mezzo di strutture memorizzazione (file; ad esempio, record con puntatori, ordinati in un certo modo)

# Indipendenza dei dati

- Il livello logico è indipendente da quello fisico:
  - una tabella è utilizzata nello stesso modo qualunque sia la sua realizzazione fisica (che può anche cambiare nel tempo)
- Perciò in questo corso vedremo solo il livello logico e non quello fisico

# Architettura standard a tre livelli per DBMS



# Architettura a 3 livelli: schemi

**schema logico**: descrizione dell'intera base di dati nel modello logico "principale" del DBMS

**Schema interno (o fisico)**: rappresentazione dello schema logico per mezzo di strutture fisiche di memorizzazione

**schema esterno**: descrizione di parte della base di dati in un modello logico ("viste" parziali, derivate, anche in modelli diversi)

# Una vista

## Corsi

Corso	Docente	Aula
Basi di dati	Rossi	DS3
Sistemi	Neri	N3
Reti	Bruni	N3
Controlli	Bruni	G

## Aule

Nome	Edificio	Piano
DS1	OMI	Terra
N3	OMI	Terra
G	Pincherle	Primo

## CorsiSedi

Corso	Aula	Edificio	Piano
Sistemi	N3	OMI	Terra
Reti	N3	OMI	Terra
Controlli	G	Pincherle	Primo

# Indipendenza dei dati

- **conseguenza della articolazione in livelli**
- **l'accesso avviene solo tramite il livello esterno (che può coincidere con il livello logico)**
- **due forme:**
  - **indipendenza fisica**
  - **indipendenza logica**

# Indipendenza fisica

- **il livello logico e quello esterno sono indipendenti da quello fisico**
  - **una relazione è utilizzata nello stesso modo qualunque sia la sua realizzazione fisica**
  - **la realizzazione fisica può cambiare senza che debbano essere modificati i programmi**



# Indipendenza logica

- **il livello esterno è indipendente da quello logico**
- **aggiunte o modifiche alle viste non richiedono modifiche al livello logico**
- **modifiche allo schema logico che lascino inalterato lo schema esterno sono trasparenti**

# Linguaggi per basi di dati

- **Un altro contributo all'efficacia: disponibilità di vari linguaggi e interfacce**
  - linguaggi testuali interattivi (**SQL**)
  - con interfacce amichevoli (senza linguaggio testuale)

# SQL, un linguaggio interattivo

- "Trovare i corsi tenuti in aule a piano terra"

## Corsi

Corso	Docente	Aula
Basi di dati	Rossi	DS3
Sistemi	Neri	N3
Reti	Bruni	N3
Controlli	Bruni	G

## Aule

Nome	Edificio	Piano
DS1	OMI	Terra
N3	OMI	Terra
G	Pincherle	Primo

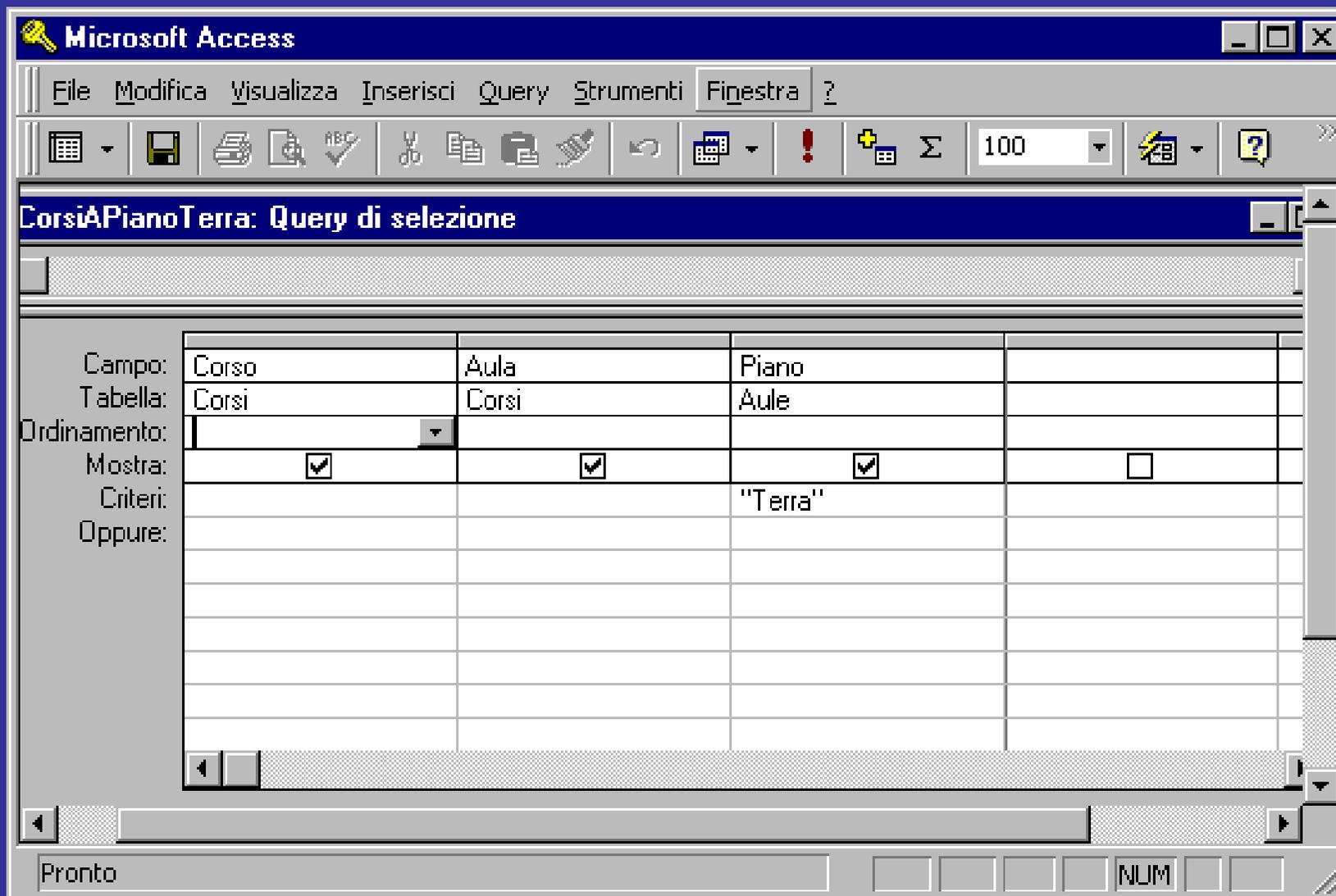
# SQL, un linguaggio interattivo

```
SELECT Corso, Aula, Piano  
FROM Aule, Corsi  
WHERE Nome = Aula  
AND Piano = "Terra"
```

Corso	Aula	Piano
Sistemi	N3	Terra
Reti	N3	Terra



# Interazione non testuale (Access)



**Una distinzione terminologica  
(separazione fra dati e programmi)**

**data manipulation language (DML)**

per l'interrogazione e  
l'aggiornamento di (**istanze** di) basi  
di dati

**data definition language (DDL)**

per la definizione di **schemi** (logici,  
esterni, fisici) e altre operazioni  
generali

# Un'operazione DDL (sullo schema)

```
CREATE TABLE orario (  
  insegnamento CHAR(20) ,  
  docente CHAR(20) ,  
  aula CHAR(4) ,  
  ora CHAR(5) )
```

# Personaggi e interpreti

- **progettisti** e realizzatori di **DBMS**
- **progettisti della base di dati** e amministratori della base di dati (**DBA**)
- **progettisti** e programmatori **di applicazioni**
- **utenti**
  - utenti **finali** (terminalisti): eseguono applicazioni predefinite (**transazioni**)
  - utenti **casuali**: eseguono operazioni non previste a priori, usando linguaggi interattivi



# Database administrator (DBA)

- **Persona o gruppo di persone responsabile del controllo centralizzato e della gestione del sistema, delle prestazioni, dell'affidabilità, delle autorizzazioni**
- **Le funzioni del DBA includono quelle di progettazione, anche se in progetti complessi ci possono essere distinzioni**



# Applicazioni per l'utente

- Programmi che realizzano attività frequenti e predefinite, con poche eccezioni, previste a priori.
- Esempi:
  - versamento presso uno sportello bancario
  - emissione di certificato anagrafico
  - dichiarazione presso l'ufficio di stato civile
  - prenotazione aerea

# Vantaggi e svantaggi dei DBMS, 1

## Pro

- **dati come risorsa comune, base di dati come modello della realtà**
- **gestione centralizzata con possibilità di standardizzazione ed “economia di scala”**
- **disponibilità di servizi integrati**
- **riduzione di ridondanze e inconsistenze**
- **indipendenza dei dati (favorisce lo sviluppo e la manutenzione delle applicazioni)**

# Vantaggi e svantaggi dei DBMS, 2

## Contro

- **costo dei prodotti e della transizione verso di essi**
- **non scorporabilità delle funzionalità (con riduzione di efficienza)**