

Basi di Dati

Corso di Laurea in Informatica
Anno Accademico 2021/2022

Alessandra Raffaetà
raffaeta@unive.it

-
- **Docenti:** Alessandra Raffaetà, Stefano Calzavara
 - **Periodo:** annuale
 - **Lezioni primo semestre:**
 - Mer 15:45 - 17:15 Aula 1, 2 + zoom
 - Gio 14:00 - 15:30 Aula 1, 2 + zoom
 - Esercitazioni in laboratorio a fine Novembre/Dicembre
 - **Ricevimento**

Venerdì 11.00 - 13:00 (avvisare tramite mail)
 - Corso successivo: Advanced Data Management (Laurea Magistrale)

- **Modalità d'esame:**

- **Scritto: 5 esercizi.**

- la capacità di progettare lo schema concettuale di una base di dati e l'acquisizione della metodologia per trasformare lo schema concettuale in uno schema relazionale,
- l'abilità di effettuare query in SQL,
- la conoscenza delle nozioni e degli algoritmi relativi alla teoria della normalizzazione,
- la capacità di amministrare in maniera adeguata una base di dati tramite la definizione di trigger, procedure ed opportune pratiche di sicurezza,
- la capacità di scrivere codice applicativo che si interfacci con una base di dati e di rappresentare conoscenza tramite modelli non relazionali.

- **Progetto: base di dati + interfaccia web**

- **Orale** [discussione scritto e progetto]

● **Compitini:**

- gennaio 2022 (Modellazione concettuale e SQL)
- Fine maggio/Giugno 2022 - (Normalizzazione e amministrazione DB, codice applicativo per interfacciarsi con DB e modelli non relazionali)

● **Appelli:**

- Fine maggio/giugno 2022 (due appelli)
- settembre 2022
- gennaio 2023

- Il progetto deve essere svolto in gruppi di 2 o 3 persone
- Il progetto può essere consegnato in qualunque appello, per tutto l'anno accademico (fino all'esame di gennaio 2023). Dopo tale data dovrà essere fatto l'intero esame (prova scritta + progetto).
- Il progetto si consegna su appuntamento in uno dei giorni previsti durante le sessioni di esame.
- Alla presentazione devono essere presenti tutti i membri del gruppo.

-
- Scritto del modulo 1 vale 50 punti
 - Scritto del modulo 2 vale 30 punti
 - Progetto vale 20 punti

● Testi adottati

- A. Albano, G. Ghelli, R. Orsini. *Fondamenti di Basi di Dati* scaricabile dal sito <http://fondamentidibasididati.it>
- H. Garcia-Molina, J.D. Ullman, J. Widom - *Database Systems: The Complete Book (Second Edition)*, Pearson, 2008.
- M. Grinberg. *Flask Web Development: Developing Web Applications with Python*. O'Reilly Media, 2018.

● Materiali aggiuntivi, avvisi e comunicazioni

- Corso su moodle: [CT0006-1 \(CT3\) 2021-22](#)
- [Lucidi delle lezioni](#)
- Testi di esame
- Esercitazioni di laboratorio
- Risultati esami
- ...

- Se sei uno studente o una studentessa con una disabilità o con un disturbo specifico dell'apprendimento (DSA) contatta il Servizio Disabilità e DSA: disabilita@unive.it.
- Vai al sito www.unive.it/disabilita per informazioni su:
 - a cosa hai diritto per legge;
 - servizi di supporto alla frequenza delle lezioni e dello studio individuale;
 - esami in modalità accessibile.

- La Commissione Paritetica (CPDS – composta in egual numero da docenti e studenti di tutti i CdS del Dipartimento) raccoglie le segnalazioni degli studenti sia direttamente sia attraverso l'analisi degli esiti dei questionari.
- Componenti studenti di Informatica:
 - Thomas Babato (CT3)
 - Diletta Olliaro (CM9)
- Rappresentanti studenti nel Consiglio di Dipartimento DAIS:
 - Thomas Babato (CT3)
 - Sant Xuan Mai (CT5)
- **Elezioni rappresentanti DAIS: 9, 10 Novembre**

Consultare la pagina **Organi ed elezioni** Referente: Dott.ssa Equizi 041 2347506

Basi di Dati: Perché?

- Interagiamo con una base di dati per
 - Acquisto al supermercato
 - Acquisto con carta di credito
 - Prelievo con il Bancomat
 - Prenotazione di un viaggio aereo
 - Biblioteca
 - Iscrizione ad un esame tramite WEB
 - Cellulare
 - ...

-
- Sistemi informativi aziendali
 - Sistemi informativi territoriali (GIS)
 - Applicazioni Internet
 - Basi di dati distribuite
 - Sistemi di supporto alle decisioni
 - Data mining
 - ...

- Area di sintesi di competenze
 - linguaggi
 - ingegneria del software
 - algoritmi e strutture dati
 - reti
 - intelligenza artificiale
- Presenta aspetti modellistici, ingegneristici, teorici
- Pone interessanti problemi di ricerca

Metodi e strumenti per la gestione di grandi quantità di dati

memorizzati in modo persistente,

per permetterne l'archiviazione, la consultazione e la modifica

da parte di utenti (uso interattivo) o programmi

in modo flessibile, efficiente e sicuro

- Una banca è interessata a gestire informazioni sui mutui dei propri clienti per produrre rapporti del tipo:

RAPPORTO MUTUO			
NUMERO MUTUO: 250		DATA: 7/2/02	
DATA SCADENZA : 1/1/10			
AMMONTARE : 70.000,00			
INTERESSE: 5%		CODICE CLIENTE : 2000	
VERSAMENTI: 4		NOME CLIENTE : Mario Rossi	
SALDO: 14.000		INDIRIZZO CLIENTE: Via Roma, 13 -Pisa	
N. RATA	DATA SCADENZA	AMMONTARE	DATA PAGAMENTO
1	1/7/00	3.500	29/06/00
2	1/1/01	3.500	30/12/00
3	1/7/01	3.500	30/06/01
4	1/1/02	3.500	30/12/01

- Quali funzionalità sono richieste ai sistemi HW/SW che gestiscono BD?
- Quali metodologie usare per progettare BD?
- Come, a quali livelli di astrazione, con quali strumenti concettuali e matematici si descrive l'organizzazione dei dati?
- Quali linguaggi usare per caricare, modificare, ricercare, aggiungere, riorganizzare i dati?
- Come realizzare applicazioni/interfacce con basi di dati?

-
- Sistemi per Basi di dati
 - Modello dei dati
 - Progettazione di Basi di dati
 - Modello relazionale
 - SQL per la definizione, manipolazione e consultazione di BD
 - Teoria della normalizzazione
 - Amministrazione di una base di dati
 - Programmazione di applicazioni che utilizzano basi di dati
 - Applicazione: Flask + SQLAlchemy
 - NoSQL: Modelli di rappresentazione dell'informazione diversi da quello relazionale

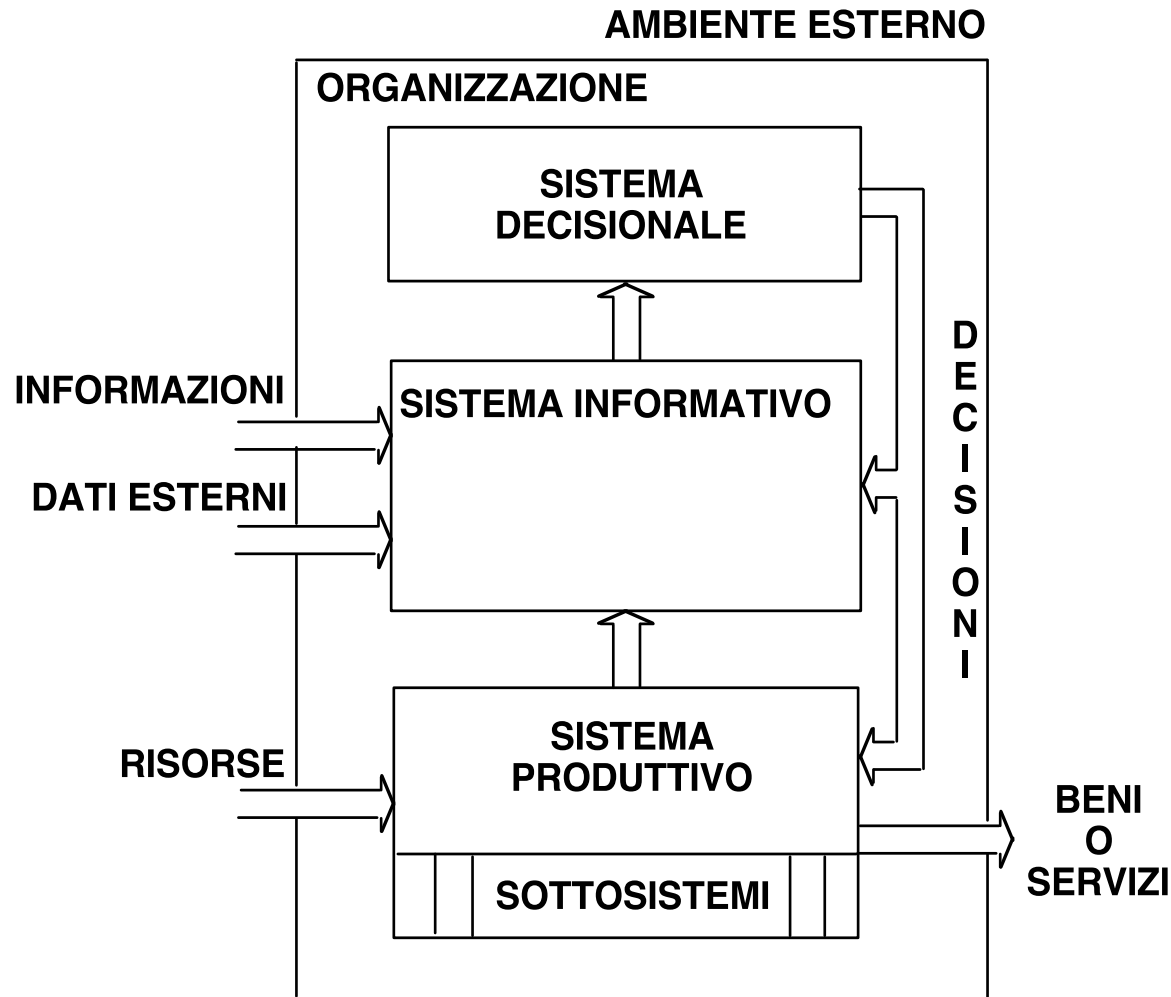
Sistema Informativo e Sistema Informatico

- L'attenzione sarà sulle basi di dati (BD) di supporto ai sistemi informativi di organizzazioni
- Le organizzazioni (viste come insiemi strutturati di uomini, risorse, strumenti e procedure finalizzati al conseguimento di determinati obiettivi) hanno in genere bisogno di gestire grandi quantità di informazioni per realizzare le proprie attività

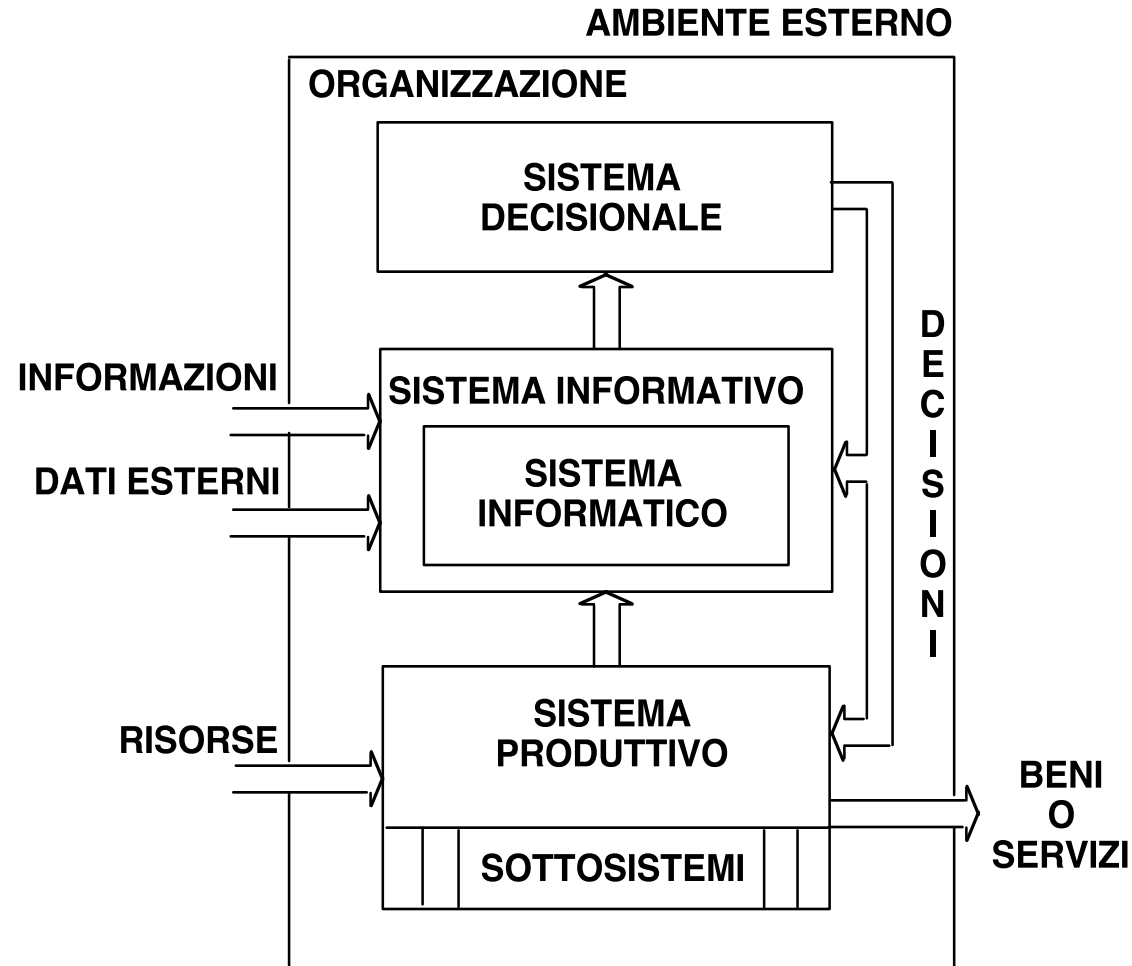
Un **sistema informativo** di un'organizzazione è una combinazione di risorse, umane e materiali, e di procedure per:

- la raccolta e acquisizione,
- l'archiviazione, conservazione
- l'elaborazione, trasformazione, produzione
- la distribuzione, comunicazione e lo scambio

delle informazioni necessarie alle attività dell'organizzazione.



- Il **sistema informatico** è l'insieme delle tecnologie informatiche e della comunicazione (Information and Communication Technologies, ICT) a supporto delle attività di un'organizzazione.
- Il **sistema informativo automatizzato** è quella parte del sistema informativo in cui le informazioni sono raccolte, elaborate, archiviate e scambiate usando un sistema informatico.
- Terminologia
 - sistema informativo \approx sistema informativo automatizzato
 - sistema informativo automatizzato \approx sistema informatico



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

(Definizioni dal Vocabolario della lingua italiana 1987)

● Anni '60: Sistemi informatici settoriali

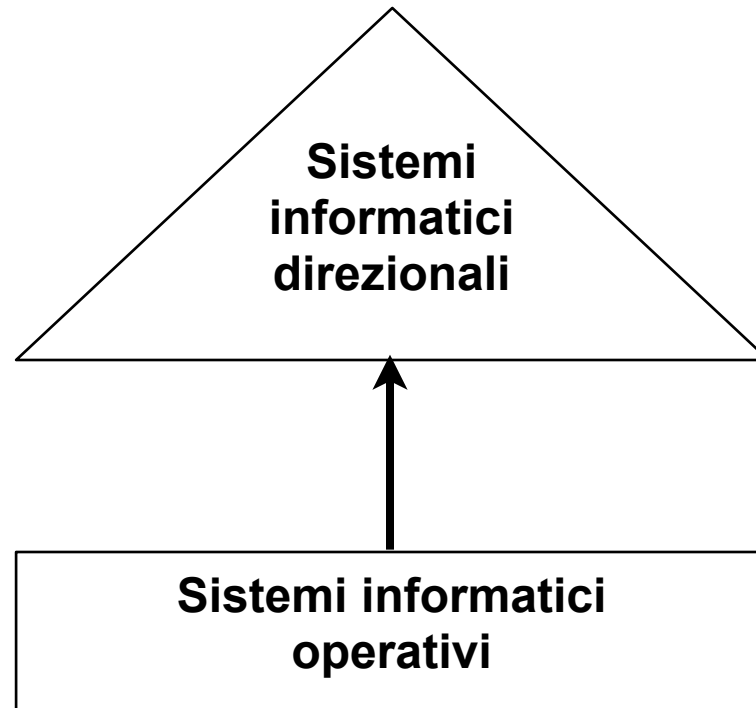
- Automatizzazione di operazioni ripetitive: elaborazione sistematica e ripetitiva di grandi quantità di dati (calcolo delle paghe, emissione di fatture)
- Supporto alle attività di gestione: contabilità generale, controllo di gestione, valutazione dell'andamento dell'azienda (informazioni sintetiche per ...)
- Tecnologia dei sistemi di gestione di archivi (file processing systems - FPSs)
 - nessuna integrazione dei dati
 - duplicazione e incoerenze
 - difficoltà nel correlare dati settoriali (anche per semplici problemi di naming)

- Inizi anni '70 : **Sistemi informatici per l'organizzazione**
 - Tecnologia dei **sistemi di gestione di basi di dati** (DBMSs), che consentono il supporto ad ogni livello dell'organizzazione.
 - Caratteristiche
 - **integrazione dei dati**: un'unica raccolta di dati comuni (informazioni di base) e diverse applicazioni che operano accedendo ai dati di loro interesse
 - Disponibilità
 - Limitazione delle ridondanze
 - Efficienza

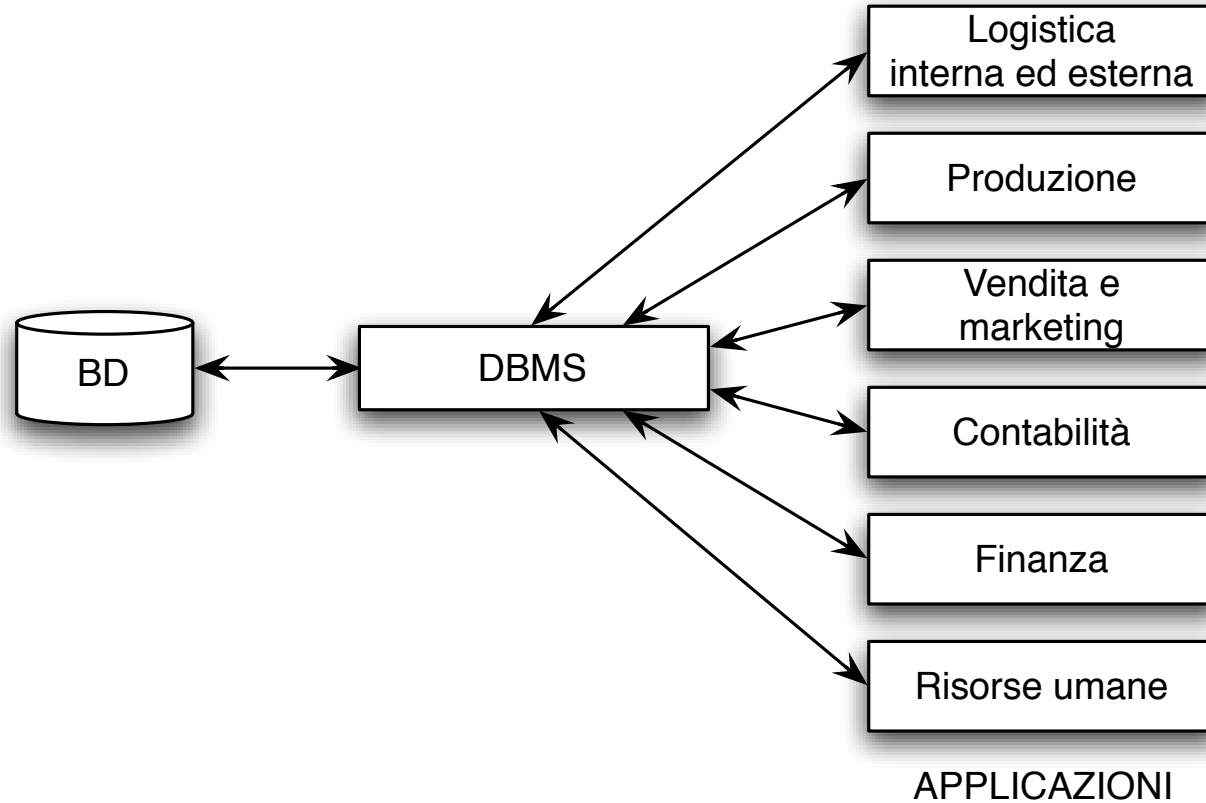
- **Flessibilità**

- **Schema**: descrive la struttura della base informativa ed è parte della base di dati stessa,
- DB **auto-descrittiva**: i programmi applicativi non sono tenuti a “conoscere” la struttura della base informativa, in quanto essa è esplicitamente rappresentata, e a tale rappresentazione essi possono accedere

-
- Anni '80: Servizi informatici per la pianificazione strategica
 - Supporto alla direzione intermedia e alta delle organizzazioni
 - Data Warehouse e Data Mining
 - ...
 - Anni 90: Web Services
 - Big Data:
 - Volume
 - Velocità
 - Varietà
 - Veridicità
 - Valore
-



- I dati sono organizzati in BD
- Le applicazioni si usano per svolgere le classiche **attività strutturate e ripetitive** dell'azienda nelle aree amministrativa e finanziaria, vendite, produzione, risorse umane ecc. (calcolo paghe, emissione fatture, magazzino, ...)
- Terminologia anglosassone
 - Data processing (DP), Electronic Data processing (EDP)
 - Transaction Processing Systems (TPS)
 - **Enterprise Resource Planning (ERP)**



- **OLTP**: Acronimo di **On-Line Transaction Processing**
- Tradizionale elaborazione di transazioni, che realizzano i processi operativi per il funzionamento di organizzazioni:
 - Operazioni predefinite e relativamente semplici
 - Ogni operazione coinvolge “pochi” dati
 - Dati di dettaglio, aggiornati
- Uso principale dei DBMS

Scontrino

Grandi Magazzini Rossetti

CodArt	Des	Prezzo	Qta	Importo
#190	Penna	3.50	2	7.00
#69	Matita	1.50	10	15.00
#90	Libro	16.00	1	16.00
TOTALE				38.00

Pagamento
Data

Carta
11-06-09

BD

Anagrafica Prodotto

#Articolo
Descrizione
Prezzo
Scorta corrente
Scorta a inizio giornata

Testata Scontrini

#Punto di vendita
#Scontrino
Importo
Mezzo pagamento
Data

Dettaglio Scontrini

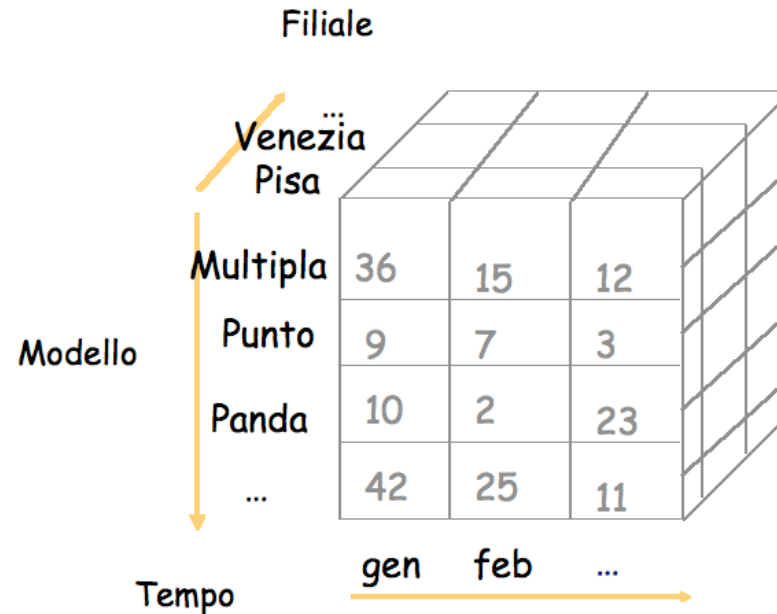
#Scontrino
#Articolo
Importo
Quantità

- La direzione intermedia e alta necessitano di:
 - **analisi storiche** dell'andamento degli affari
 - **produzione interattiva (non programmata) di rapporti di sintesi**, da punti di vista ed a livelli di dettaglio diversi [situazione anomale, tendenze interessanti]
- Le **basi di dati operative** risultano **inadeguate**:
 - contengono solo **dati recenti** (situazione corrente)
 - le operazioni coinvolgono grandi quantità di dati o sono molto complesse e quindi rallenterebbero in modo inaccettabile le funzioni operative.

- **Sistemi informatici direzionali**

- Dati sono organizzati in **Data Warehouse** (DW)
 - Gestiti da un opportuno **sistema per analisi interattive dei dati** [dati storici, aggregati, di sintesi]
 - **Applicazioni di Business intelligence** come strumenti di supporto ai processi di controllo delle prestazioni aziendali e di decisione manageriale
- Terminologia anglosassone:
 - Management Information Systems (MIS)
 - Decision support systems (DSS), data or model based
 - Executive Information System (EIS)

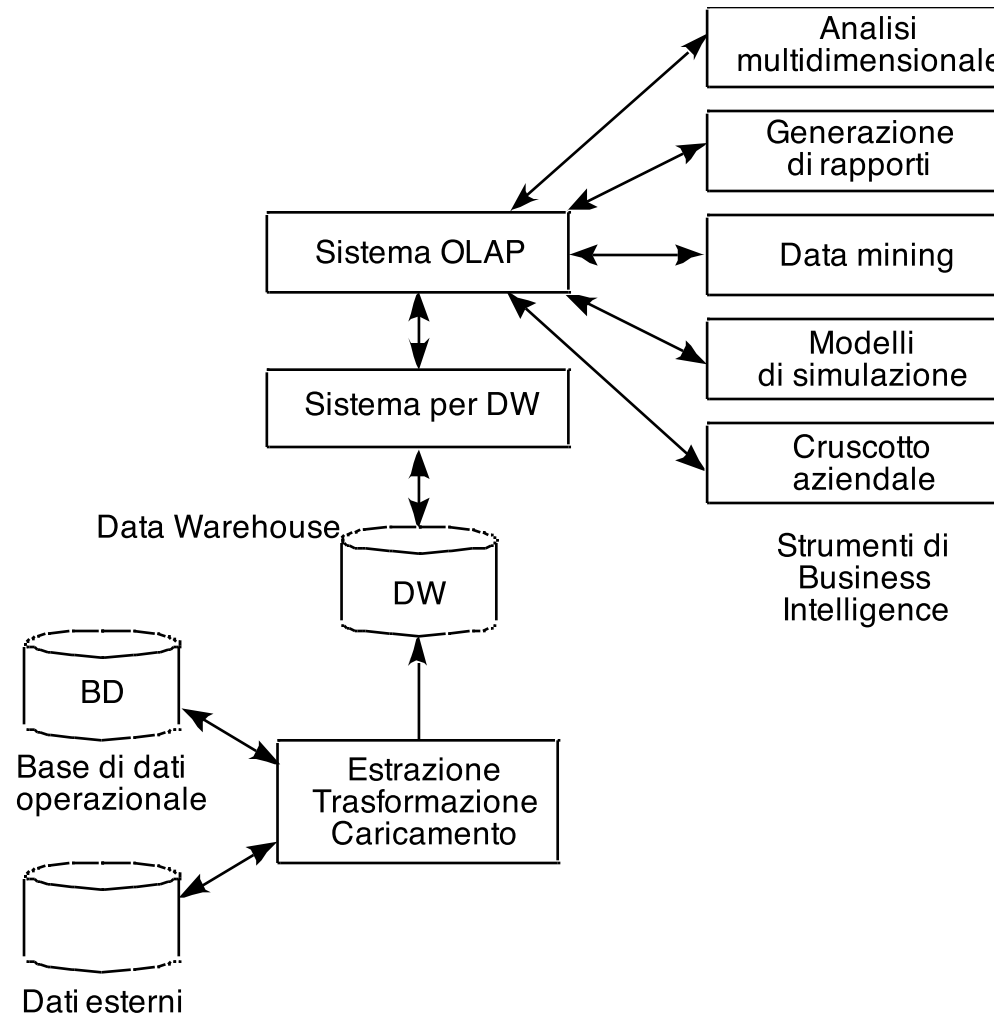
Le informazioni sono:
tempificate,
aggregate,
multidimensionali



Aggrega i dati
eliminando la
dimensione “filiale”
(roll up)

Vendite totali auto										
Modello	gennaio			febbraio			...	anno		
	bdg	eff	Δ	bdg	eff	Δ		bdg	eff	Δ
Multipla										
Panda										
...										

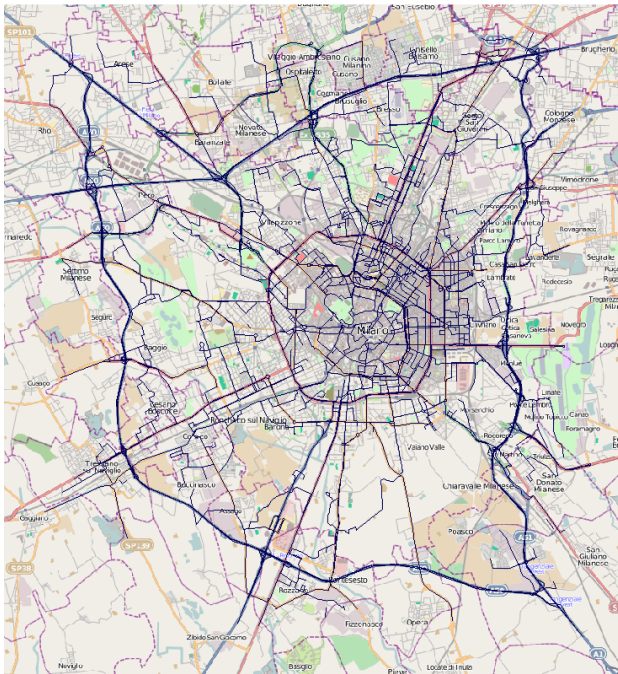
- **OLAP**: Acronimo di **On-Line Analytical Processing**
- Uso principale dei **data warehouse**
- Caratteristiche
 - Operazioni complesse e casuali
 - Ogni operazione può coinvolgere moltissimi dati
 - I dati sono letti, ma non modificati
 - Dati aggregati, storici, anche non attualissimi



Traffic in Milan: How does traffic vary along the day?

39

0.00am - 3.00am



3.00am - 6.00am



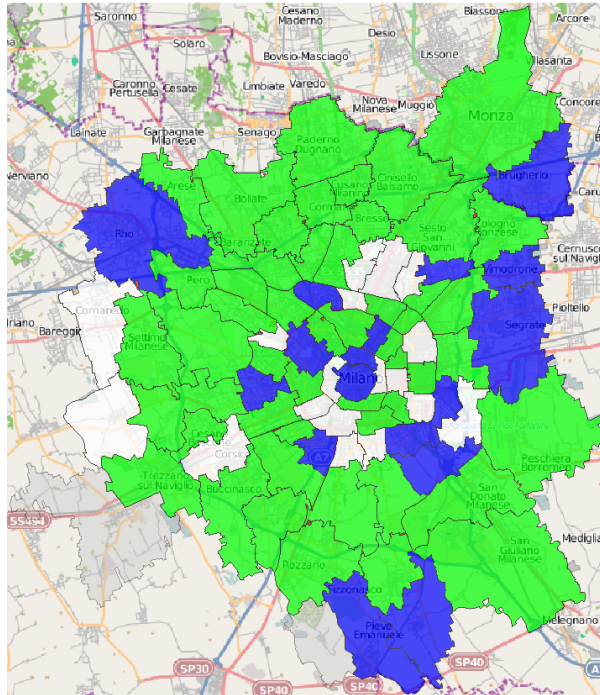
6.00am - 9.00am



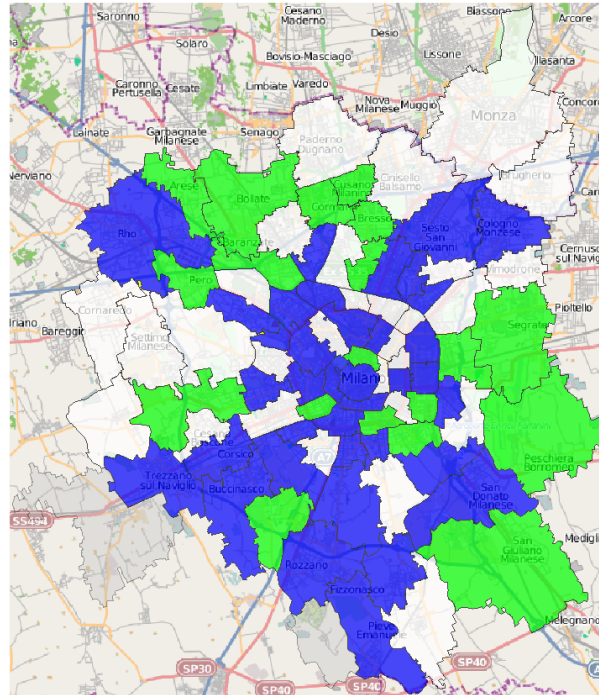
Which are possible residential areas?

40

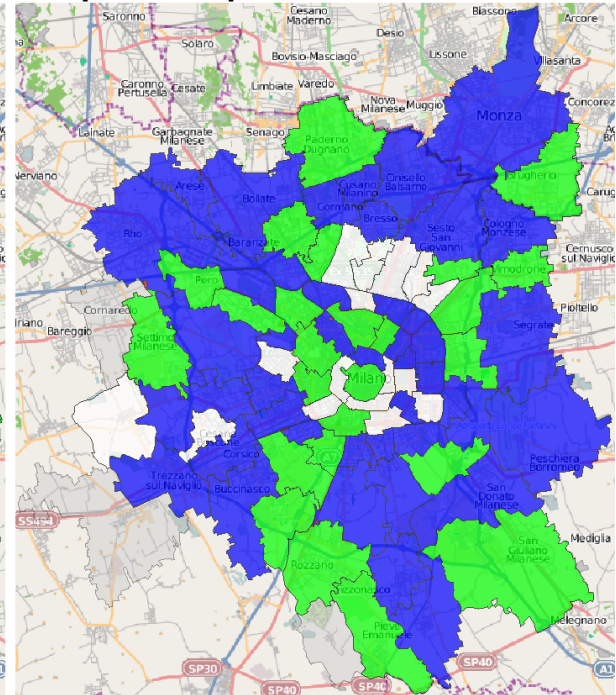
5.00am - 6.00am



7.00am - 8.00am



5.00pm - 6.00pm



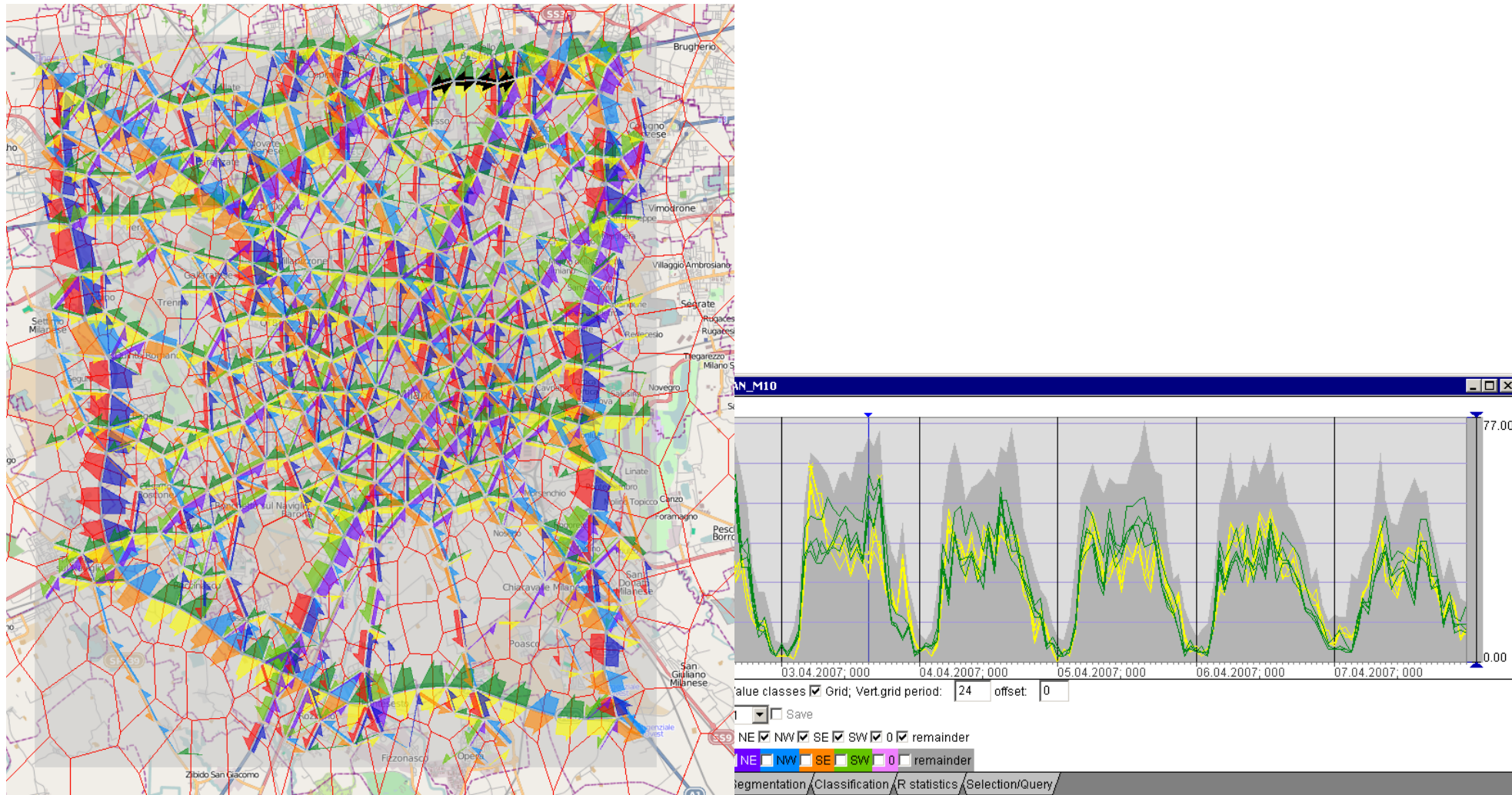
Green starts > ends

Blue starts < ends

White starts ≈ ends

Which are the main traffic flows?

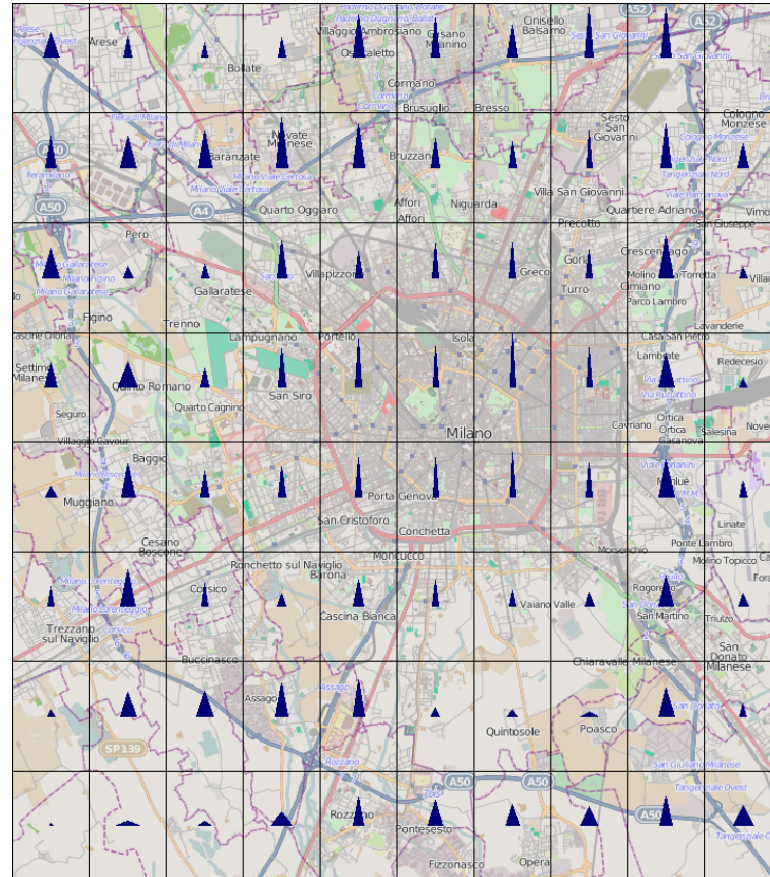
41



Where do traffic jams appear?

42

Monday 6am - 9 am



Sistemi per la gestione di Basi di Dati (DBMS)

- In generale, una qualsiasi raccolta di dati permanenti gestiti tramite un elaboratore elettronico, ma per noi ...
- **Definizione:** Una **base di dati** è una raccolta di dati permanenti, gestiti da un elaboratore elettronico, suddivisi in due categorie:
 - **Metadati:** definiscono lo **schema** della BD, che descrive:
 - struttura dei dati [e operazioni eseguibili]
 - restrizioni sui valori ammissibili (**vincoli di integrità**)
 - utenti autorizzati, ecc.

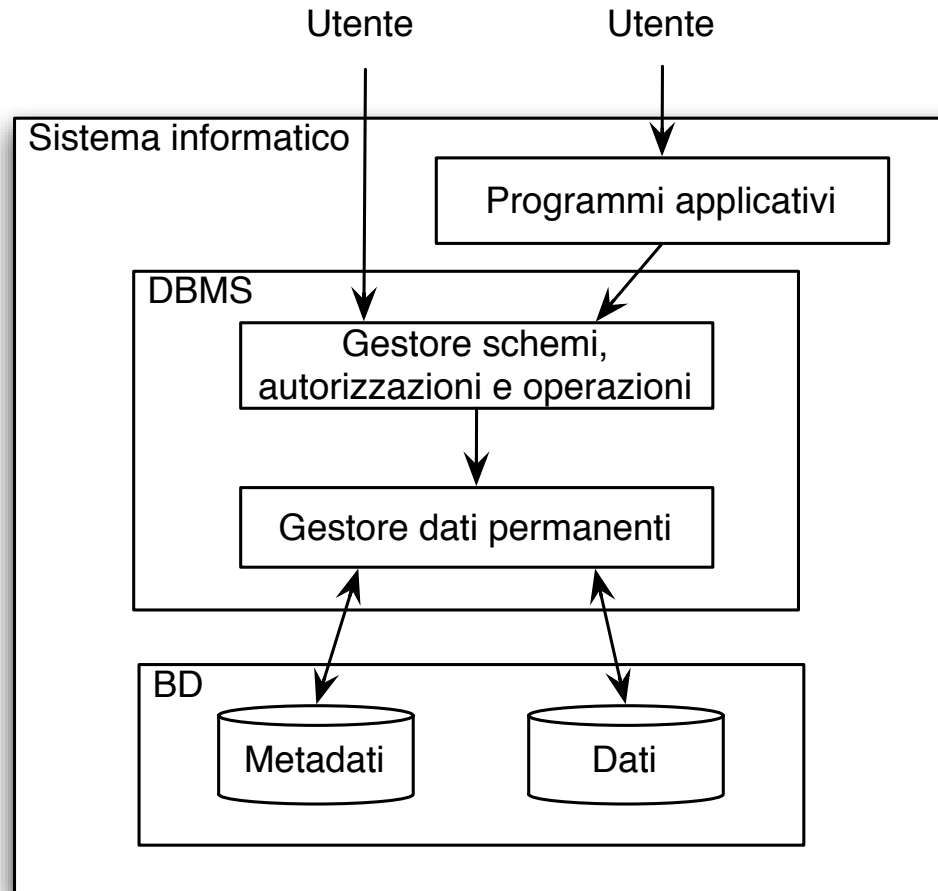
Definito prima di creare i dati e indipendente dalle applicazioni

- **Dati:** le rappresentazioni di certi fatti conformi alle definizioni dello schema, con le seguenti caratteristiche ...

- Sono organizzati in **insiemi omogenei**, fra i quali sono definite delle **relazioni**. La struttura dei dati e le relazioni sono descritte nello schema usando i meccanismi di astrazione del **modello dei dati** adottato;
- Sono **molti**, in assoluto e rispetto ai metadati, e non possono essere gestiti in memoria temporanea;
- Sono **permanenti**, continuano ad esistere finché non sono rimossi esplicitamente (indipendenti dalle applicazioni)
- Sono **utilizzabili contemporaneamente** da utenti diversi.

- Sono **protetti** sia da accesso da parte di utenti non autorizzati, sia da corruzione dovuta a malfunzionamenti hardware e software;
- Sono accessibili mediante **transazioni**, unità di lavoro atomiche che non possono avere effetti parziali;

- **Definizione:** Un **DBMS (Data Base Management System)** è un sistema centralizzato o distribuito che offre opportuni linguaggi/strumenti per:
 - **definire** lo **schema** della BD (va definito prima di creare dati)
 - definito usando il modello dei dati adottato dal DBMS
 - interrogabile con le stesse modalità previste per i dati;
 - scegliere le **strutture dati** per la memorizzazione dei dati,
 - **memorizzare** i dati rispettando i vincoli definiti nello schema;
 - **recuperare e modificare** i dati interattivamente (linguaggio di interrogazione o **query language**) o da programmi.



- Il modello relazionale dei dati è il più diffuso fra i DBMS commerciali.
- Il meccanismo di astrazione fondamentale è la **relazione (tabella)**
~ insieme di record con campi di tipo elementare;

Nome	<u>Matricola</u>	Citta	AnnoNascita
Verdi	71523	Padova	1987
Rossi	76366	Dolo	1988
Zeri	71347	Venezia	1988

Studenti

- Lo schema specifica le tabelle
 - nome
 - struttura degli elementi (nome e tipo degli attributi).
- Definizione base di dati (schema vuoto)
 - **CREATE DATABASE** EsempioEsami;

- Definizione schema:

- CREATE TABLE** Studenti (
Nome char(8),
Matricola int **NOT NULL**,
Citta char(10),
AnnoNascita int,
PRIMARY KEY (Matricola));
- CREATE TABLE** ProveEsami (
Materia char(5),
Matricola int,
Data char(6),
Voto int,
Lode char(1),
PRIMARY KEY (Materia,Matricola));
- ...

Nome	<u>Matricola</u>	Citta	AnnoNascita
Verdi	71523	Padova	1987
Rossi	76366	Dolo	1988
Zeri	71347	Venezia	1988

Studenti

ProveEsami

<u>Materia</u>	<u>Matricola</u>	Data	Voto	Lode
CN	71523	08.07.06	27	N
FIS	76366	08.07.07	26	N
BD	71523	28.12.06	30	S

● Inserzione dati:

- **INSERT INTO** ProveEsami

VALUES ('BD' , 71523 , '28.12.06' , 30 , 'S');

- Interrogazione:

```
SELECT Matricola  
FROM ProveEsami  
WHERE Materia = 'BD' AND Voto = 30;
```

Matricola

71523

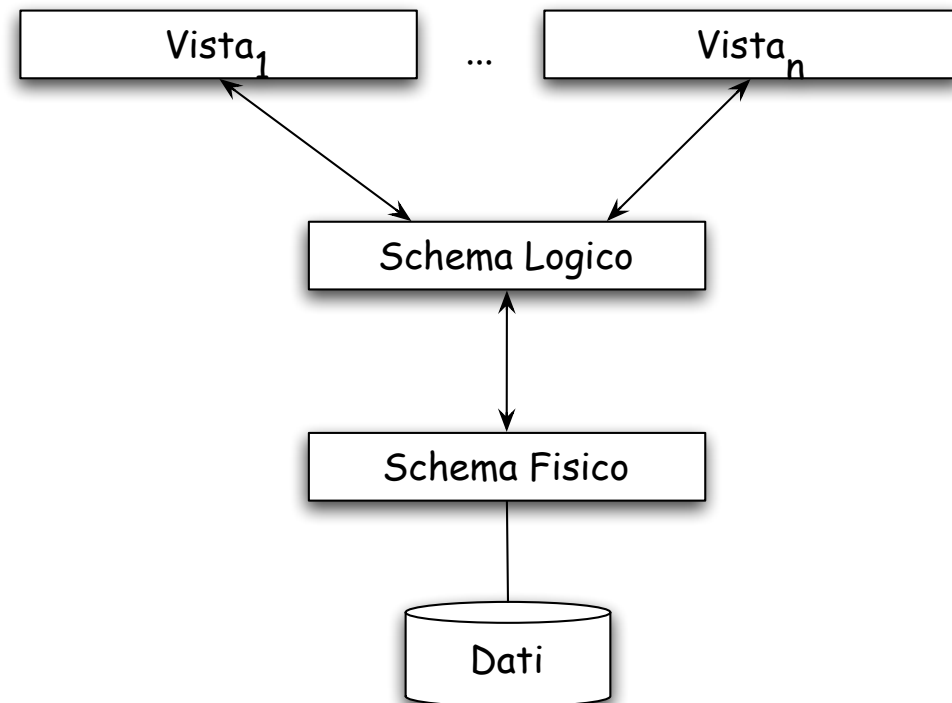
Funzionalità dei DBMS

- Linguaggio per la **definizione della base di dati (DDL)**;
- Linguaggi per l'**uso dei dati (DML)**;
- Meccanismi per il **controllo dei dati**;
- Strumenti per il **responsabile della base di dati**;
- Strumenti per lo **sviluppo delle applicazioni**.

DDL: Definizione della BD

- La descrizione della BD è indipendente dalle applicazioni che la usano
- Tre diversi livelli di descrizione dei dati (schemi):

- **livello di vista logica,**
- **livello logico**
- **livello fisico**



- **Schema logico:** Descrive la **struttura degli insiemi di dati e delle relazioni** fra questi, secondo un certo **modello dei dati**, senza nessun riferimento alla loro organizzazione fisica nella memoria permanente.
- Esempio:

```
Studenti(Matricola int, Nome char(20), Login char(8),  
         AnnoNascita int, Reddito real )
```

```
Corsi(IdeC char(8), Titolo char(20), Credito int )
```

```
Esami(Matricola int, IdeC char(8), Voto int )
```

... realizzata in SQL con **CREATE TABLE**.

- Descrive lo **schema fisico** o **interno**:
 - come vanno organizzati fisicamente i dati nelle memorie permanenti
 - strutture dati ausiliarie per l'uso (es. indici)
- Esempio:
 - Corsi e Esami organizzate in modo seriale
 - Studenti organizzata in modo sequenziale con indice `Indice su Matricola`

```
CREATE INDEX Indice ON Studenti(Matricola);
```

- Descrive come deve apparire la struttura della base di dati ad una certa applicazione (**schema esterno** o **vista**).
- Esempio:
 - InfCorsi(IdeC char(8), Titolo char(20), NumEsami int)

```
CREATE VIEW InfCorsi (IdeC, Titolo, NumEsami) AS  
    SELECT IdeC,  
           Titolo,  
           COUNT(*)  
FROM Corsi NATURAL JOIN Esami  
GROUP BY IdeC, Titolo;
```

- **Indipendenza fisica**: i programmi applicativi non devono essere modificati in seguito a modifiche dell'organizzazione fisica dei dati. Es.
 - strutture dati ausiliarie
 - modifica della distribuzione

- **Esempio**: Se si deve risalire spesso agli studenti che hanno sostenuto un particolare esame:

```
CREATE INDEX IndiceIdeC ON Esami(IdeC);
```

- **Indipendenza logica**: i programmi applicativi non devono essere modificati in seguito a modifiche dello schema logico
 - difficile da ottenere
 - richiederà la ridefinizione dello schema esterno
- Esempio: per suddividere la collezione degli studenti in part-time e full-time:

```
CREATE TABLE StudentiFull (...);
```

```
CREATE TABLE StudentiPart (...);
```

```
CREATE VIEW Studenti AS
```

```
SELECT * FROM StudentiFull
```

```
UNION
```

```
SELECT * FROM StudentiPart;
```

DML: Linguaggi per l'uso dei dati

- Un DBMS prevede varie modalità d'uso per soddisfare le esigenze delle diverse categorie di utenti che possono accedere alla base di dati:
 - Utenti delle applicazioni
 - Utenti non programmatori
 - Programmatori delle applicazioni
- **Utenti non programmatori**
 - Interfaccia grafica per accedere ai dati
 - Linguaggio di interrogazione

● Utenti programmatori

- Linguaggio integrato (dati e DML)
 - Linguaggio disegnato ad-hoc per usare SQL. I comandi SQL sono controllati staticamente dal traduttore ed eseguiti dal DBMS.
- Linguaggio convenzionale + funzioni di libreria predefinita
 - Linguaggio convenzionale che usa delle funzioni di una libreria predefinita per usare SQL. I comandi SQL sono **stringhe** passate come parametri alle funzioni che poi vengono controllate dinamicamente dal DBMS prima di eseguirle.
- Linguaggio che ospita l'SQL (SQL embedded)
 - Linguaggio convenzionale esteso con un nuovo costrutto per marcare i comandi SQL. Occorre un **pre-compilatore** che controlla i comandi SQL, li sostituisce con chiamate a funzioni predefinite e genera un programma nel linguaggio convenzionale + funzioni di libreria

Controllo dei Dati

- Meccanismi offerti per garantire le seguenti proprietà:
 - **Integrità**: mantenimento delle proprietà specificate in modo dichiarativo nello schema (vincoli d'integrità)
 - **Sicurezza**: protezione dei dati da usi non autorizzati
 - restrizione dell'**accesso ai soli utenti autorizzati**
 - **limitazione delle operazioni eseguibili**
Es: banca dati sui cittadini (CF, dati anagrafici, reddito)
 - alcuni utenti accedono, non modificano i dati
 - alcuni utenti possono accedere solo ai propri dati
 - alcuni utenti accedono solo a dati statistici (es. media), non singoli
 - **Nota**: nel caso di uso statistico, per proteggere i dati riservati non basta limitare l'utilizzo delle funzioni su insiemi.

-
- **Affidabilità**: protezione dei dati da
 - **interferenze** indesiderate dovute all'**accesso concorrente** ai dati da parte di più utenti.
 - **malfunzionamenti hardware o software** (fallimenti di transazione, fallimenti di sistema, disastri)

- **Definizione:** Una **transazione** è una sequenza di azioni di lettura e scrittura in memoria permanente e di elaborazioni di dati in memoria temporanea, con le seguenti proprietà:
 - **Atomicità:** Le transazioni che terminano prematuramente (aborted transactions) sono trattate dal sistema come se non fossero mai iniziate; pertanto eventuali loro effetti sulla base di dati sono annullati.
 - **Serializzabilità:** Nel caso di esecuzioni concorrenti di più transazioni, l'effetto complessivo è quello di una esecuzione seriale.
 - **Persistenza:** Le modifiche sulla base di dati di una transazione terminata normalmente sono permanenti, cioè non sono alterabili da eventuali malfunzionamenti.

- Protezione da interferenze indesiderate tra accessi concorrenti ai dati
 - Basterebbe impedire l'inizio di una transazione prima che un'altra finisca, però ...
 - scheduling dei singoli passi di ciascuna transazione in $\{T1, \dots, Tn\}$ che garantisca che l'effetto complessivo sarebbe ottenibile eseguendo le transazioni isolatamente in qualche ordine

● Es:

T1	T2
leggi(SALDO)	leggi(SALDO)
SALDO := SALDO + 100	SALDO := SALDO - 100
scrivi(SALDO)	scrivi(SALDO)

- Protezione da malfunzionamenti hardware o software:
 - **fallimenti di transazione**: dovuta a una situazione prevista dall'applicazione o a eventi imprevisti, come la violazione di vincoli di integrità o accessi non autorizzati
 - **fallimenti di sistema**: dovuti ad un'anomalia HW o SW dell'unità centrale o di una periferica, che determina l'interruzione di tutte le transazioni attive e la perdita del contenuto della memoria temporanea
 - **disastri**: danni alla memoria permanente
- Ripristino basato su journal e copie di sicurezza.

Strumenti per l'amministrazione

- Strumenti per l'amministratore della base di dati
 - Linguaggio per la definizione e la modifica degli schemi della base di dati
 - logico, esterno, e fisico
 - Strumenti per il controllo e messa a punto del funzionamento del sistema.
 - Strumenti per stabilire i diritti di accesso ai dati
 - Strumenti per ripristinare la base di dati in caso di malfunzionamenti di sistemi o disastri

- Indipendenza fisica e logica
- Gestione efficiente dei dati
- Integrità e sicurezza dei dati
- Accessi interattivi, concorrenti e protetti dai malfunzionamenti
- Amministrazione dei dati
- Riduzione dei tempi di sviluppo delle applicazioni

- Possono essere costosi e complessi da installare e mantenere in esercizio.
- Richiedono personale qualificato (se si tratta di personale esterno, aumenta la dipendenza da ditte esterne)
- Le applicazioni sviluppate possono essere trasferite con difficoltà su sistemi diversi se vengono usati linguaggi troppo “legati” al DBMS usato

MA

- La riduzione dei costi della tecnologia e i possibili tipi di DBMS disponibili sul mercato facilitano la loro diffusione.