

**Laurea in Informatica
A.A. 2021-2022**

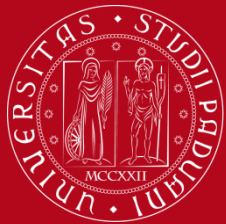
Corso "Base di Dati"

Introduzione alle Basi di Dati

Prof. Massimiliano de Leoni



**UNIVERSITÀ
DEGLI STUDI
DI PADOVA**



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Aziende e Sistemi Informativi

Organizzazione



**gestisce le informazioni di
interesse per il
perseguimento degli scopi
dell'organizzazione**

Sistema informativo

deliveroo

Sistema informatico

**Porzione automatizzata del
sistema informativo con
tecnologie informatiche**



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Sistemi Informativi e Basi di Dati

Organizzazione



Sistema informativo

Sistema informatico



**insieme organizzato,
persistente e condiviso di
dati utilizzati per lo
svolgimento delle attività
automatizzate
dell'organizzazione**

(automatizzate = con il supporto del sistema informatico)

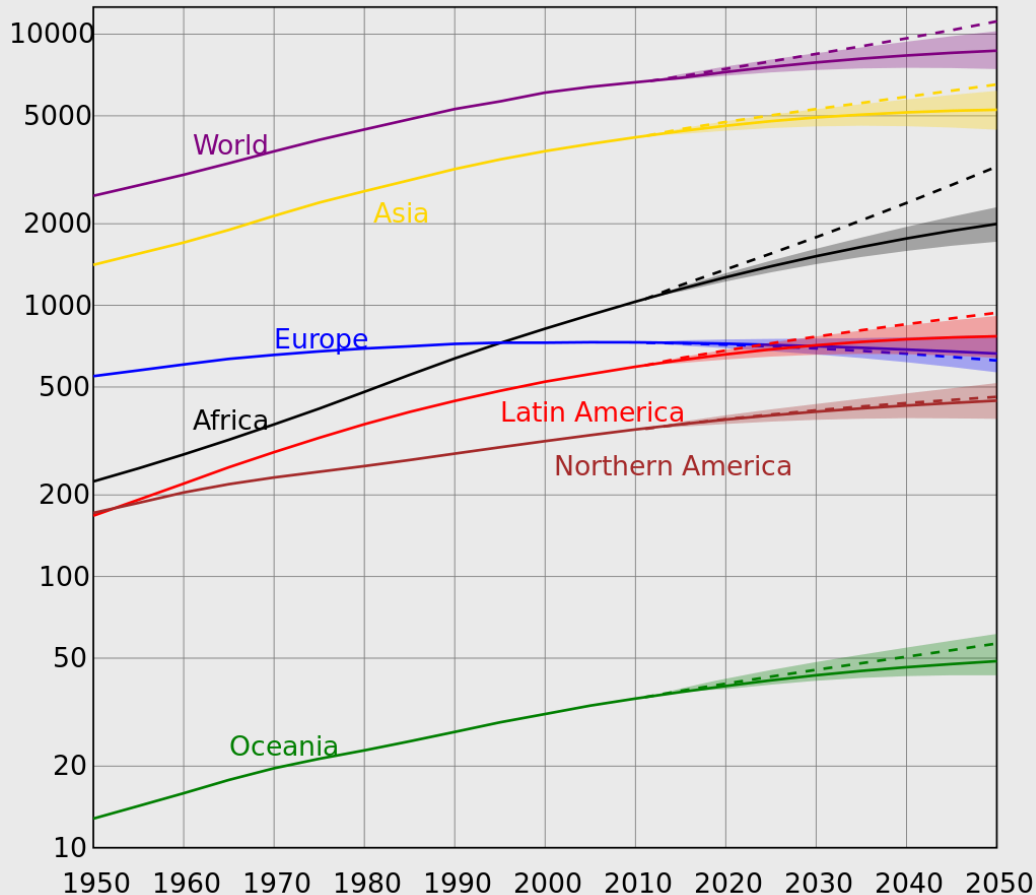
Informazioni e dati

(definizioni dal Vocabolario della lingua italiana 1987)

dato: ciò che è immediatamente presente alla conoscenza, **prima di ogni elaborazione**;

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

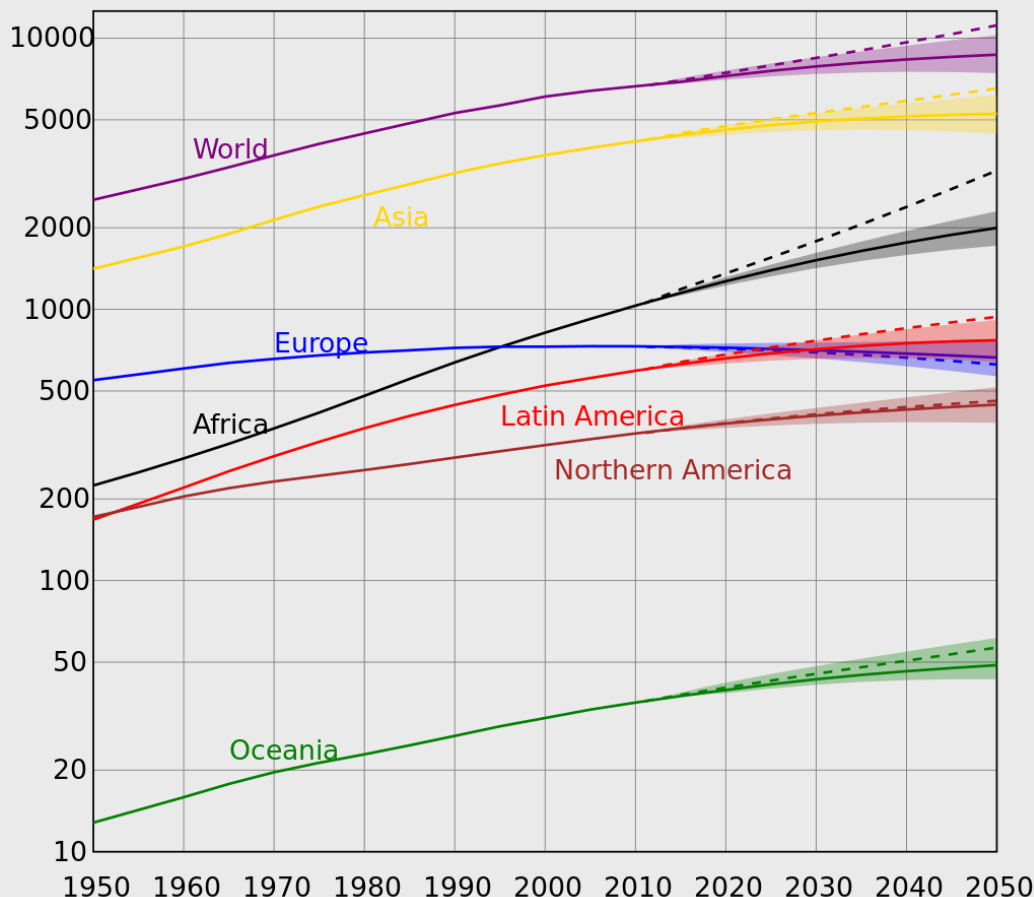
Dati



Dati:

- Misurazione delle numero di abitanti nei diversi continenti
- Stime della popolazione fino al 2050 per continente.

Dati + Elaborazione = Informazione



Dati:

- Misurazione delle numero di abitanti nei diversi continenti
- Stime della popolazione fino al 2050 per continente.

Informazione:

- Nel 2050, la popolazione mondiale sarà 4 volte
- La polazione in Europa tenderà a scendere

I dati sono inutili senza elaborazione.

I sistemi informativi generano informazione

https://it.wikipedia.org/wiki/Popolazione_mondiale

Gestione delle informazioni

- **I dati codificano le informazioni**

Ad esempio, nei servizi anagrafici e nel riferimento a persone

- descrizioni discorsive
- nome e cognome
- estremi anagrafici
- codice fiscale

- **La rappresentazione precisa di forme più ricche di informazione e conoscenza è difficile**

Base di dati

Insieme organizzato di dati utilizzati per il supporto allo svolgimento delle attività di un ente (azienda, ufficio, comuni, ecc...)

DataBase Management System (DBMS)

- Le basi di dati sono gestiti da un DBMS (Database Management System)
- Una basi di dati è:
 1. Grande
 2. Persistente
 3. Condivisa
- Un DBMS garantisce alla base di dati:
 1. Privacy
 2. Affidabilità
 3. Efficienza
 4. Efficacia

Le basi di dati sono ...

- **Grandi**: Il limite deve essere solo quello fisico dei dispositivi (>100 miliardi di record, >10 Terabyte di dati);
- **Persistenti**: indipendente dalle singole esecuzioni dei programmi;
- **Condivise**: Applicazioni e utenti diversi accedono ad una porzione (sovrapponibile) delle basi di dati

Problemi delle condivisioni

- **Ridondanza**: informazioni ripetute;
- Rischio di **incoerenza**: le versioni possono non coincidere;
- Problemi di **privacy**: i fruitori delle basi di dati devono accedere ai dati per cui sono autorizzati;

Problemi delle condivisioni: La Soluzione!

- Attività diverse su dati condivisi:
 - meccanismi di **autorizzazione**
- Accessi di più utenti ai dati condivisi:
 - controllo della **concorrenza**
 - uso di lock e semafori

I DBMS garantiscono ... privacy

- Si possono definire meccanismi di autorizzazione
- Esempio:
 - l'utente A è autorizzato a leggere tutti i dati e a modificare X
 - l'utente B è autorizzato a leggere dati X e a modificare Y

I DBMS garantiscono... affidabilità

- **Affidabilità:** resistenza a malfunzionamenti hardware e software
- Affidabilità → Gestione delle **transazioni**
 - **Transazioni** = Insieme di operazioni da considerare indivisibile corretto anche in presenza di concorrenza e con effetti definitivi

Affidabilità → Gestione delle transazioni: Un esempio

- Trasferimento di fondi da un conto A ad un conto B:
 1. Prelevamento da A
 2. Versamento su B
- Si fanno il prelevamento da A e il versamento su B o nessuno dei due

Le transazioni sono ... concorrenti

- Esempi:
 1. se due assegni emessi sullo stesso conto corrente vengono incassati contemporaneamente
... si deve evitare di trascurarne uno
 2. se due agenzie rischiedono lo stesso posto (libero) su un treno
... si deve evitare di assegnarlo due volte

Quindi:

Transazioni T_1, \dots, T_n

Fare T_1, \dots, T_n concorrentemente =
– Fare T_1 , poi T_2 , ..., poi T_n

Affidabilità: i risultati delle transazioni sono permanenti

- La conclusione positiva di una transazione corrisponde ad un impegno (in inglese **commit**) a mantenere traccia del risultato in modo definitivo, anche in presenza di guasti e di esecuzione concorrente

I DBMS debbono essere... efficienti

- Utilizzare al meglio le risorse memoria, disco e tempo
- Se un DBMS non è efficiente, il sistema informativo non è efficiente!

I DBMS debbono essere... efficaci

- Cercano di rendere produttive le attività dei loro utilizzatori, offrendo funzionalità articolate, potenti e flessibili:
- **Il corso è in buona parte dedicato ad illustrare come i DBMS perseguono l'efficacia**

DBMS vs file system

- DBMS si costruiscono sopra i file system
- I DBMS estendono le funzionalità dei file system, garantendo alle base di dati:
 1. Privatezza
 2. Affidabilità
 3. Efficienza
 4. Efficacia

Modello dei dati

- Insieme di costrutti utilizzati per organizzare i dati di interesse e descriverne la dinamica
- Questo corso usa il **modello relazionale**
- Il **modello relazionale** prevede il costruttore **relazione**, che permette di definire insiemi di record omogenei

Esempio di (porzione di) 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

Esempio di (porzione di) base di dati

Orario

Domini



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

Esempio di (porzione di) 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

Insegnamento = String

Docente = String

Aula = {N1, N2, N3}

Ora = {0:01,0:02,...,23:58,23:59}

Orario \subseteq Insegnamento X Docente X Aula X Ora

Schema e istanza

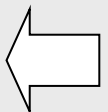
- In ogni base di dati esistono:
 - lo **schema**, sostanzialmente invariante nel tempo, che ne descrive la struttura (aspetto intensionale)
 - es.: le intestazioni delle tabelle
 - l'**istanza**, i valori attuali, che possono cambiare anche molto rapidamente (aspetto estensionale)
 - es.: il “corpo”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

Due tipi (principali) di modelli

- modelli logici
- modelli concettuali

Modelli logici

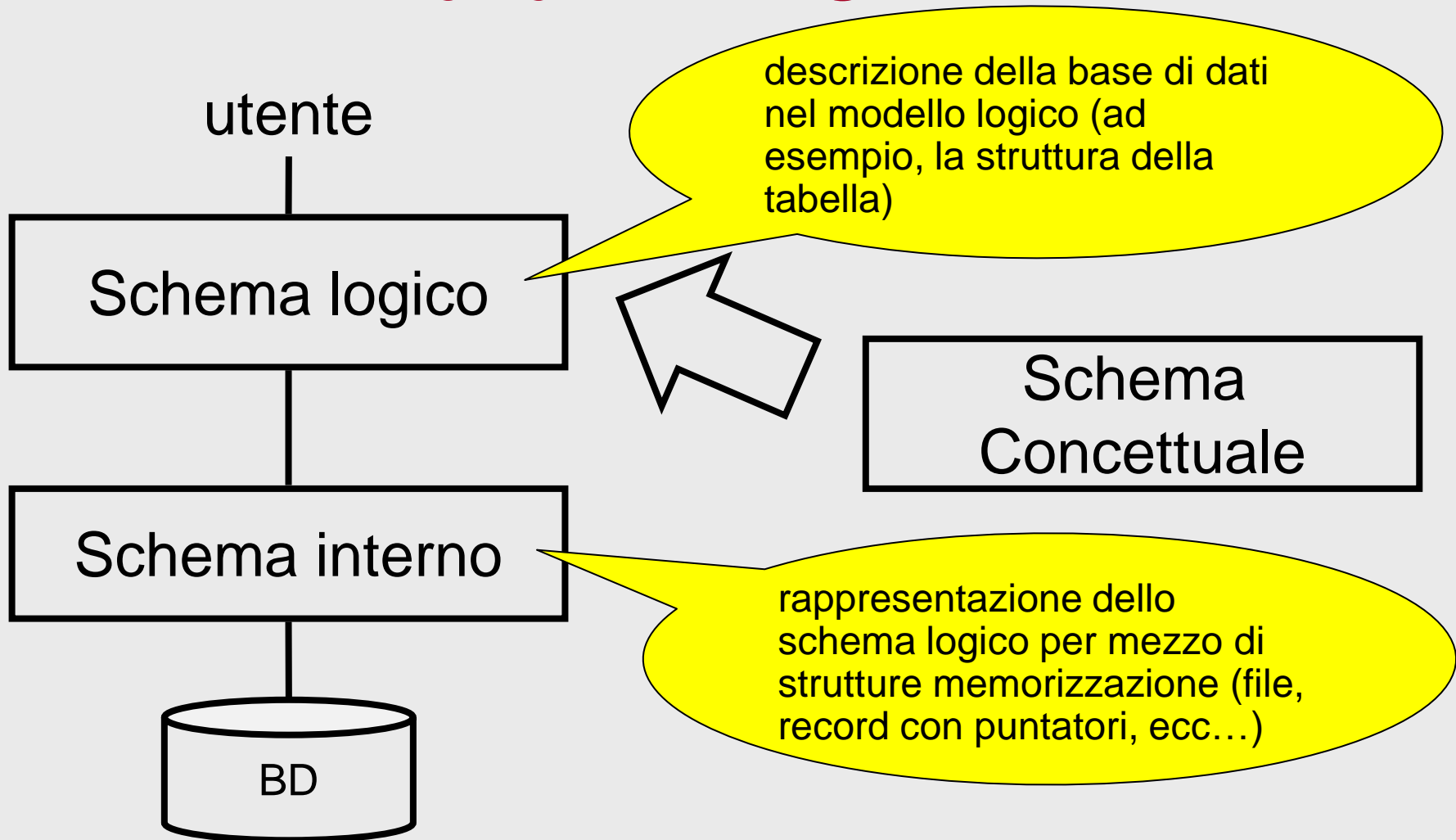
- Adottati nei DBMS esistenti per l'organizzazione dei dati
 - utilizzati dai programmi
 - indipendenti dalle strutture fisiche
- Esempi: **relazionale**, reticolare, gerarchico, a oggetti, basato su XML



Modelli concettuali

- Permettono di rappresentare i dati in modo indipendente da ogni sistema
 - cercano di descrivere i concetti del mondo reale
 - sono utilizzati nelle fasi preliminari di progettazione
- Il più diffuso è il modello **Entity-Relationship**

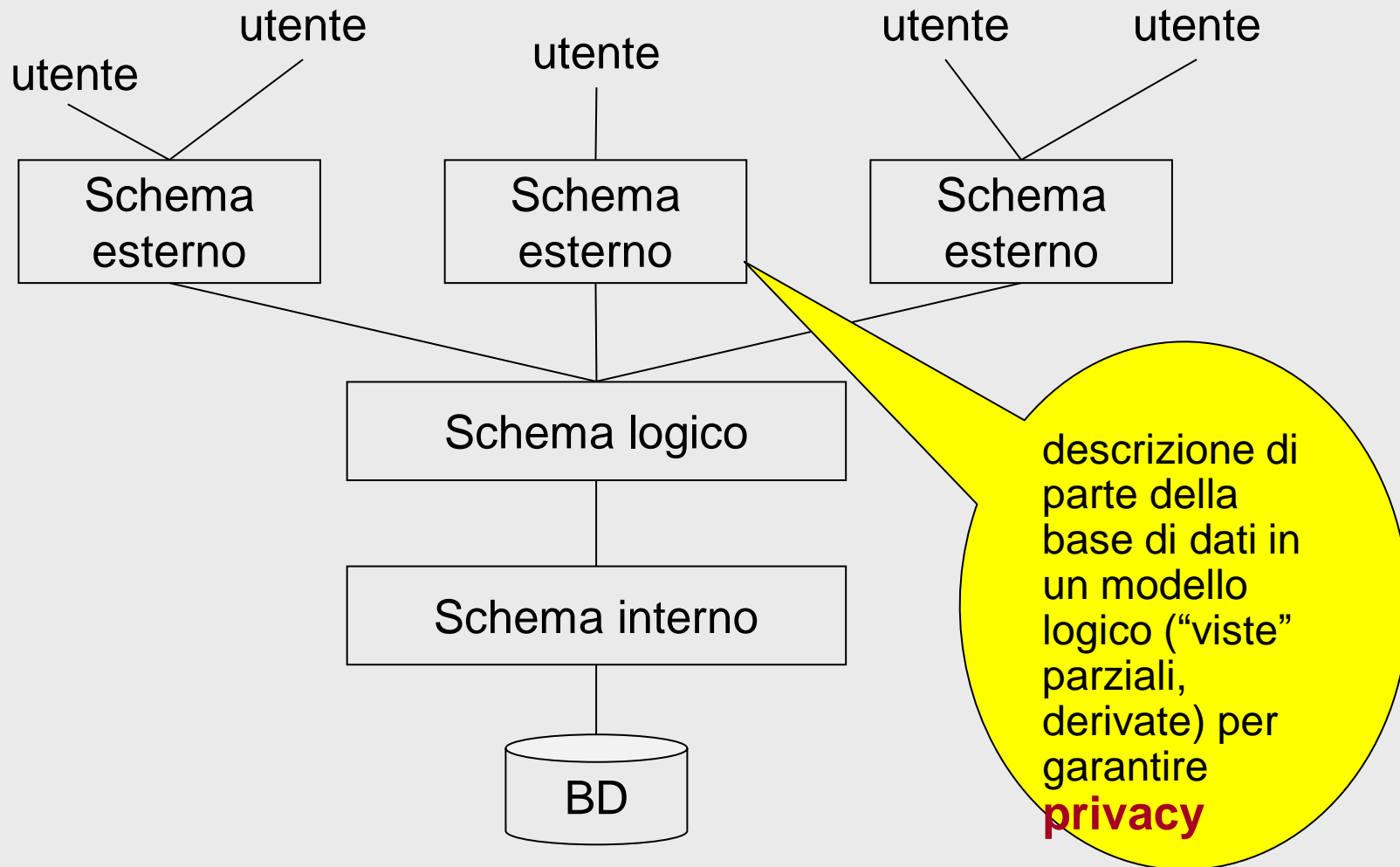
Architettura (semplificata) di un DBMS



Indipendenza dei dati

- Il livello logico è indipendente da quello fisico.
- In questo corso vedremo solo il livello logico e non quello fisico

Architettura standard (ANSI/SPARC) a tre livelli per DBMS



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

Corsi Sedi

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

Linguaggi per basi di dati

- Un altro contributo all'efficacia: disponibilità di vari linguaggi e interfacce
 - ⇒ linguaggi testuali interattivi (**SQL**)
 - ⇒ comandi (SQL) immersi in un linguaggio **ospite** (Pascal, Java, C ...)
 - ⇒ con interfacce amichevoli (senza linguaggio testuale)

SQL, un linguaggio interattivo

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

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

SQL, un linguaggio interattivo

"Trovare i corsi
tenuti in aule a
piano terra"

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

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

"Trovare i corsi
tenuti in aule a
piano terra"

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

Corso	Docente	Aula	Nome	Edificio	Piano
Sistemi	Neri	N3	N3	OMI	Terra
Reti	Bruni	N3	N3	OMI	Terra
Controlli	Bruni	G	G	Pincherle	Primo

SQL, un linguaggio interattivo

"Trovare i corsi
tenuti in aule a
piano terra"

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

Corso	Docente	Aula	Nome	Edificio	Piano
Sistemi	Neri	N3	N3	OMI	Terra
Reti	Brani	N3	N3	OMI	Terra
Controllo	Brani	G	G	Finchione	Primo

SQL immerso in linguaggio ospite (es. Java)

```
import java.sql.*;

public class FirstExample {
    public static void main(String[] args) {
        Class.forName("com.mysql.jdbc.Driver");
        Connection conn=
        DriverManager.getConnection("jdbc:mysql://localhost/EMP",
            "user", "pass");
        Statement stmt= conn.createStatement();
        ResultSet rs =
            stmt.executeQuery("SELECT id, first, last, age FROM Employees");
        while(rs.next()){
            int id  = rs.getInt("id");
            int age = rs.getInt("age");
            String first = rs.getString("first");
            String last = rs.getString("last");
            System.out.print("ID: " + id);
            System.out.print(", Age: " + age);
            System.out.print(", First: " + first);
            System.out.println(", Last: " + last);
        }
    }
}
```

SQL in linguaggio ad hoc (Oracle PL/SQL)

```
declare Stip number;
begin
    SELECT STIPENDIO INTO STIP FROM IMPIEGATO
    WHERE MATRICOLA = '575488' FOR UPDATE OF STIPENDIO;
    if Stip > 30 then
        UPDATE IMPIEGATO SET STIPENDIO = STIPENDIO * 1.1
        WHERE MATRICOLA = '575488';
    else
        UPDATE IMPIEGATO SET STIPENDIO = STIPENDIO * 1.15
        WHERE MATRICOLA = '575488';
    end if;
    commit;
exception
    when no_data_found then
        INSERT INTO ERRORI
        VALUES('MATRICOLA INESISTENTE',SYSDATE);
end;
```


Interazione Grafica (Postgres)

The screenshot displays the pgAdmin 4 web interface. The left sidebar shows the database structure, with the 'customers' table selected under the 'public' schema. The main panel shows the 'Query Editor' with the following SQL query:

```
1 SELECT customerName
2 FROM customers;
```

Below the query editor, the 'Data Output' tab is active, showing the results of the query. The results are displayed in a table with 12 rows and 1 column, 'customername'.

	customername
1	Atelier graphique
2	Signal Gift Stores
3	Australian Collectors, Co.
4	La Rochelle Gifts
5	Baane Mini Imports
6	Mini Gifts Distributors Ltd.
7	Havel & Zbyszek Co
8	Blauer See Auto, Co.
9	Mini Wheels Co.
10	Land of Toys Inc.
11	Euro+ Shopping Channel
12	Value Model Baglines Co

Linguaggio SQL: Non solo per Query

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) )
```

Vantaggi dei DBMS

- DBMS permettono ai dati di essere una risorsa comune, condivisa nell'organizzazione
- 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

Svantaggi dei DBMS

- I DBMS possono essere complessi e costosi
 - Ne vale la pena solo per grandi organizzazioni
 - Non utilizzare senza:
 - tanti utenti;
 - accessi concorrenti e stabili;
- A volte: meglio file semplici!

Riferimenti

- Capitolo 1 del libro di riferimento



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Organizzazione



Obiettivi del corso

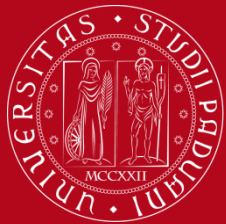
- Conoscenze di base per la progettazione, l'implementazione e l'utilizzo di sistemi di gestione di basi di dati (DBMSs = Database Management Systems)
 1. Creazione di una BD, Base di Dati (inglese database)
 2. Interrogazione di una BD per l'estrazione di dati
 3. Progettazione di una BD
 4. Ottimizzazione di una BD (tramite «normalizzazione»)
 5. Gestione degli accessi concorrenti ad una BD
 6. Accesso ad una BD da software C/C++
- Laboratori per «sporcarsi le mani con le BD»



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Organizzazione del corso

- Crediti: 9 CFU di cui:
 - 6 CFU Introduzione Concetti (~48 ore)
 - 2 CFU Didattica Laboratorio (16 ore)
 - 1 CFU Esercitazioni (~8 ore)
- **Nessuna registrazione delle lezioni per massimizzare partecipazione sincrona**
- **Esercitazioni e Didattica Laboratorio solo in presenza**
(<https://www.unipd.it/node/86476>)
- Supporto alle attività di laboratorio:
 - Dott. Samuel Congolato
 - Dott. Alessandro Padella



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Materiale

- Libro di testo:
P. Atzeni, S. Ceri, P. Fraternali,
S. Paraboschi, R. Torlone. *Basi di dati*.
McGraw-Hill Education; 5a edizione
- Sulla pagina Moodle del corso:
 1. Annunci
 2. Slides del Corso
 3. Materiale Esercitazioni e Laboratorio
 4. Esempi di Esame
 5. ...
- Registratevi su Moodle se non lo siete già!





UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Come fare pratica?

- Numerosi esercizi alla fine di ogni capitolo del libro
- Soluzioni disponibili su:
https://highered.mheducation.com/sites/8838694451/student_view0/soluzioni_per_gli_esercizi_.html
- Numerosi esercizi «per casa» alla fine dei laboratori



Come contattare i docenti?

- **Non utilizzare la email!**
 - Circa 150 studenti
 - Se ogni studente invia una email al mese ed ogni risposta richiede 20 minuti
 - ➔ $20 \times 150 = 50$ ore $>$ 1 settimana lavorativa (40h)
 - Discussioni per email sono meno efficienti di quelle verbali
- Ricevimento dopo lezione del mercoledì: in LUM250 [Via Luzzatti]



Orari ed aula

- Giorni/Orario (pausa di 15 minuti dopo ca. 45 minuti) :
 - Martedì: 12.30 - 14.15
 - Mercoledì: 14.30 - 16.15
 - Giovedì: 12.30 - 14.15
- Lezioni ed Esercitazioni: Aula LUM250
- Laboratorio: P140 + P036
- **Non sempre lezione:** Consultare il Calendario delle lezioni sulla pagina Moodle del corso per sapere se e quando lezione, esercitazione o laboratorio



Due parti:

- **Progetto (25%)**
 - Gruppi da due persone (salvo casi eccezionali e documentabili)
 - Definizione di un Dominio Applicativo e dei Requisiti
 - Progettazione e Realizzazione di una Base di Dati
 - Interrogazione ed Accesso da software
- **Prova Scritta (75%)**
 - Esercizi sulla Esecuzione di Interrogazioni in SQL e/o Algebra Relazionale
 - Esercizi sulla Progettazione e Realizzazione di una Base di Dati
 - Esercizi sulla Normalizzazione o Gestione delle Transazioni
 - Quiz Scelta Multipla



Criteri di Valutazione / 2

- Occorre raggiungere la sufficienza ($\geq 18/30$) in ciascuna prova
 - Nessuna possibilità di fare l'orale
 - Rifiutare il voto del progetto e/o della prova scritta equivale a ripetere la prova
 - Esame \rightarrow Nuovo Esame in un appello successivo
 - Progetto \rightarrow Nuovo Progetto su un nuovo dominio
- Non è possibile fare lo scritto senza aver consegnato il Progetto
- **Dopo due insufficienze allo scritto, occorre fare un nuovo progetto!**
- È possibile comunque consegnare il progetto e fare lo scritto in appelli successivi



- Progetto:
 - Scadenza per il Primo Scritto: 6 Gennaio alle 22.00
 - Scadenze dal Secondo Scritto: a partire dal secondo scritto, 10 giorni prima di ogni scritto alle 22.00 anche se weekend (per esempio se lo scritto fosse il 12/2, la scadenza sarebbe 2/2 alle 22.00)
- **Non farti sorprendere (negativamente) dal voto del progetto!**
 - (Porzioni di) Laboratori dedicati al feedback sul progetto
- Quando chiedi feedback:
 - A domanda precisa, risposta precisa
 - A domanda generica (per es. «Va bene il progetto?»), risposta generica



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Considerazioni Conclusive

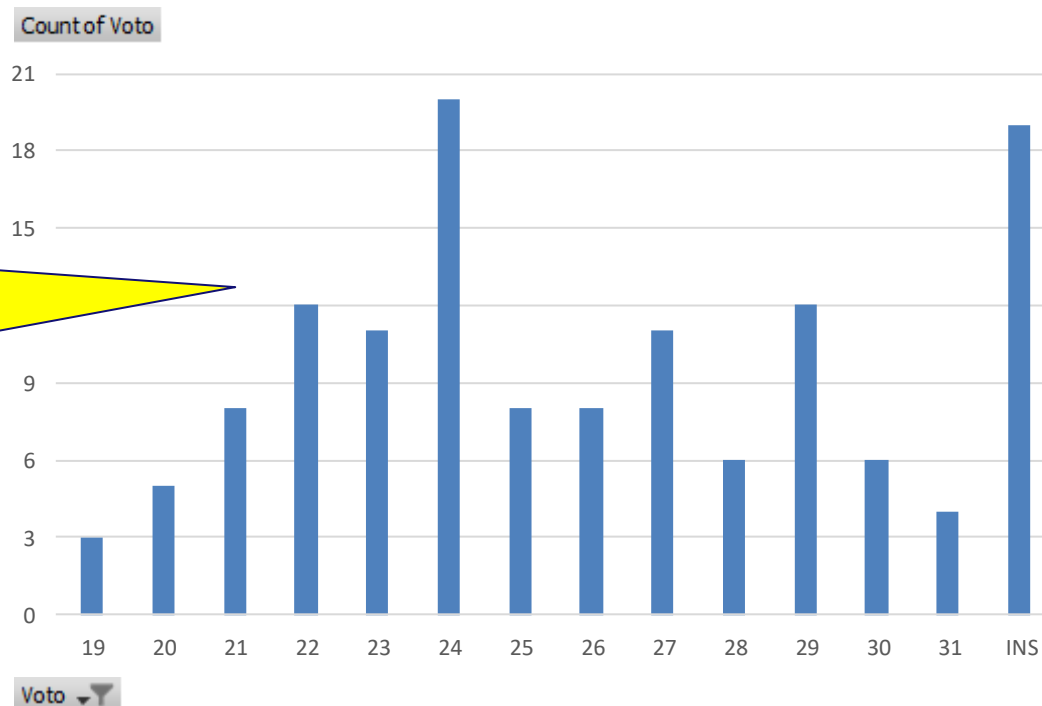


Capire un Modello \neq Essere in grado di modellare

**Modellare ed Interrogare una Base di Dati
non è banale!**

**Distribuzione
voti A.A 20-21**

**(ca. 20% non è
passato dopo 5
appelli)**





UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Anno Lezioni in Presenza vs Anno Lezioni Duale

Numero Studenti Passati



Frazione Studenti Passati per Appello

