

14AFQPL - Basi di Dati

a.a. 2024-2025

INTRODUZIONE AL CORSO



POLITECNICO
DI TORINO



Obiettivo del corso

- Progettare una base dati relazionale
- Scrivere interrogazioni in linguaggio SQL e algebra relazionale
- Conoscere alternative alle basi dati relazionali
 - I sistemi NoSQL
 - Large Language Models (AI)
- Perché le basi dati sono così importanti???

Sistemi informativi e basi di dati

struttura

- I sistemi informativi
 - Scopo: gestire la risorsa «informazione»
 - Strumento: tecnologie informatiche e della comunicazione (ICT)
 - Vincolo: rispettare regole ed obiettivi aziendali

Matricola	Codice	Data	Voto
S123456	AK9876	2022-01-30	28
S321765	AG7654	2022-01-25	15
S574375	AG7654	2022-02-21	25
S987654	AK9876	2022-01-30	27

contenuto

- Le basi di dati sono il CUORE dei sistemi informativi
 - Le applicazioni (programmi) che creano i servizi sono costruite sui dati
 - Le applicazioni vengono create/modificate/cancellate in modo dinamico e frequente
 - Il contenuto della base dati cambia nel tempo
 - La struttura della base dati dovrebbe cambiare il meno possibile

Un esempio familiare

Appelli		
Basi di dati (AA-LK)	Prof. Cagliero Luca	14AFQPL14AFQPI
▶ 08/09/2023 ore 14:00 - 16:30; Scritto 📅 Scadenza: 04/09/2023 14:00 - 29 studenti prenotati Aula/e: 3 - AULA Corso Duca		
Basi di dati (LL-ZZ)	Prof. Farinetti Laura	14AFQPL14AFQPI
▶ 08/09/2023 ore 14:00 - 16:30; Scritto 📅 Scadenza: 04/09/2023 14:00 - 10 studenti prenotati Aula/e: 27B - AULA Corso Duca		
Basi di dati	Prof. Farinetti Laura	04AFQPC
▶ 18/09/2023 ore 17:00 - 20:00; Scritto 📅 Scadenza: 12/09/2023 14:00 - 1 studenti prenotati Aula/e: 13S		
▶ Visualizza gli appelli degli ultimi 3 mesi		

Verbali d'esame
① Procedura da utilizzare per la verbalizzazione delle prove d'esame degli studenti.
ENTRA
Per problemi di accesso alla procedura contattare help.verbali@polito.it Informazioni per richiedere il certificato di firma remota 📄

Dati e informazioni

- In informatica non sono sinonimi
- Un dato è una rappresentazione oggettiva e non interpretata della realtà, ed è solitamente costituito da simboli che devono essere elaborati e contestualizzati
- Un'informazione è una visione della realtà derivante dall'elaborazione e interpretazione dei dati, cioè dipende dal significato che associamo ai dati



Dati e informazioni

- La memorizzazione di un dato deve quindi essere valutata e attribuita a un suo contesto, cioè «organizzata»

Matricola	Codice	Data	Voto
S123456	AK9876	2022-01-30	28
S321765	AG7654	2022-01-25	15
S574375	AG7654	2022-02-21	25
S987654	AK9876	2022-01-30	27

- Serve un modello dei dati
- Noi useremo principalmente il modello relazionale
 - Basato sul concetto matematico di «relazione» (informalmente «tabella»)

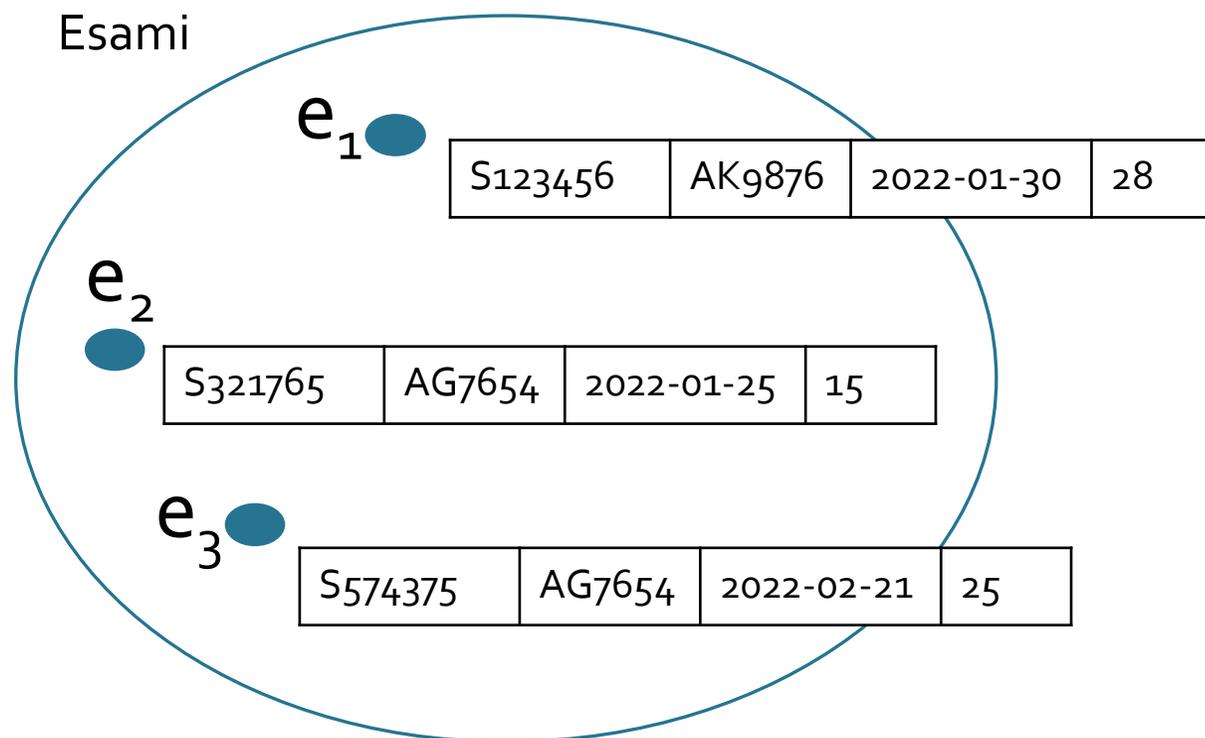
Stato	Continente	Popolazione (M)
Cile	America	18
Australia	Oceania	25
Italia	Europa	59

Modello	Colore	Numero	Prezzo (€)
SG-237	Blu	25	62
SG-239	Verde	39	57
RZ-135	Nero	37	102
RZ-431	Bianco	37	25
KT-762	Blu	44	82

Il modello relazionale

- Nasce dalla teoria degli insiemi

Matricola	Codice	Data	Voto
S123456	AK9876	2022-01-30	28
S321765	AG7654	2022-01-25	15
S574375	AG7654	2022-02-21	25
S987654	AK9876	2022-01-30	27



Obiettivi (informali) del corso

Matricola	Codice	Data	Voto
S123456	AK9876	2022-01-30	28
S321765	AG7654	2022-01-25	15
S574375	AG7654	2022-02-21	25
S987654	AK9876	2022-01-30	27

- Come creare le tabelle, in modo da memorizzare in modo efficiente tutti i dati di interesse per il sistema informativo?
- Principi chiave: organizzazione dei contenuti, no ridondanza dei dati

Progettazione delle basi dati

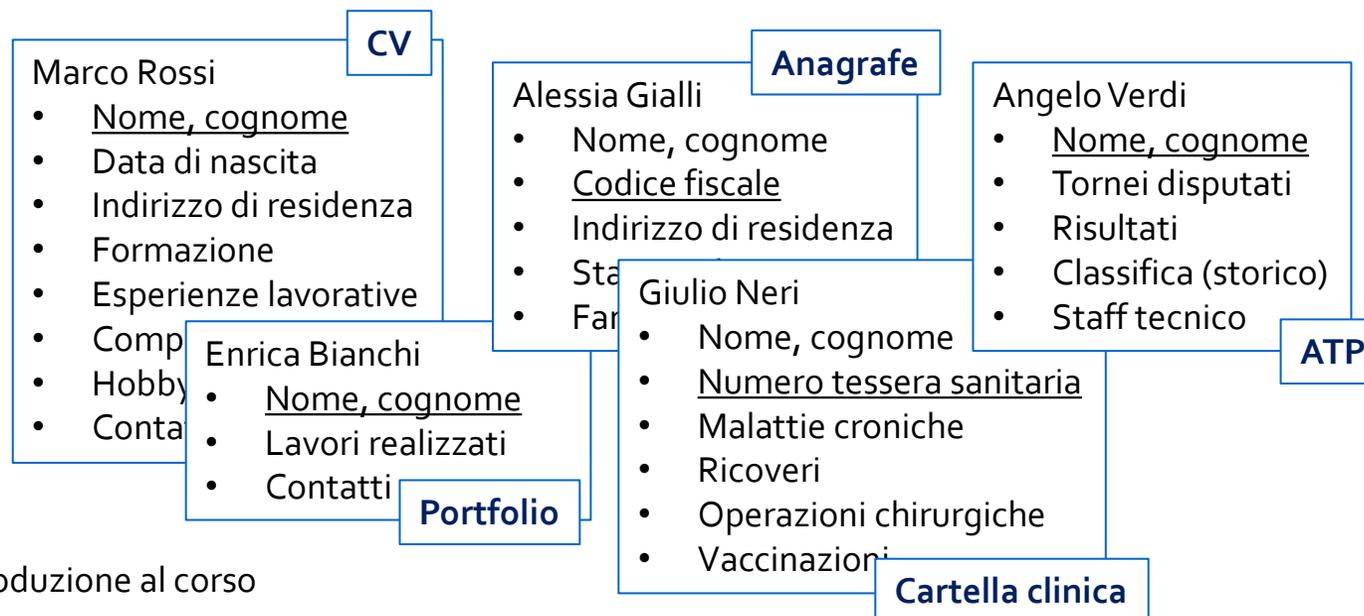
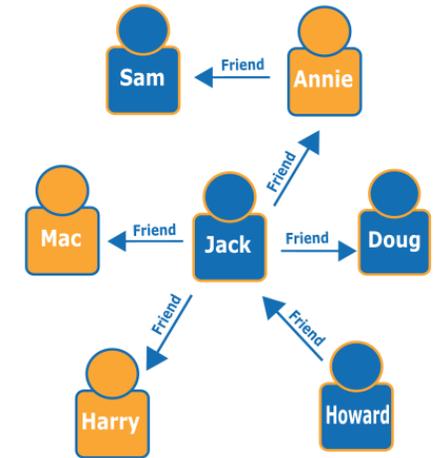
- Come leggere le tabelle, elaborando i dati contenuti in modo da rispondere a domande «interessanti»?
- Estrazione di righe/colonne/celle, calcolo di dati aggregati

Interrogazione delle basi dati: linguaggio SQL

Obiettivi (informali) del corso

- E se è difficile «costringere» i dati in una tabella, perchè non sono sufficientemente strutturati? Esistono alternative?
- Organizzazione flessibile dei contenuti (basata su documenti, grafi, ...) linguaggi di interrogazione diversi da SQL

Basi di dati non relazionali o NoSQL



Obiettivi (informali) del corso



- Pensando alle basi dati in senso più ampio, si può accedere alle informazioni “dialogando” con un servizi interattivi che “parlano” in linguaggio naturale
- Chatbot, or “AI assistants”

Large Language Models (LLM)



- Algoritmi AI di deep-learning pre-addestrati su grandi quantità di dati, capaci di riconoscere, generare, riassumere, tradurre e prevedere contenuti,



Base di dati e DBMS

- Definizione generale
 - Una base di dati è una collezione di dati che rappresenta le informazioni di interesse per un sistema informativo
- Definizione «tecnica»
 - Una base di dati è una collezione di dati gestita da un DBMS
- Un DBMS (Data Base Management System) o Sistema per la Gestione di Basi di Dati è un sistema software in grado di gestire collezioni di dati che siano grandi, condivise, persistenti, assicurando la loro affidabilità e privatezza

Linguaggi di accesso ai dati

- Linguaggi testuali interattivi (SQL)
- Interfacce amichevoli
 - Comandi simili a quelli interattivi introdotti in linguaggi di programmazione (Python, Java, ...), detti linguaggi ospite

Filtri di Ricerca

Anno Accademico: 2022/2023 ▾

Sede: TORINO ▾

Tipo Laurea: Laurea ▾

Area: Ingegneria ▾

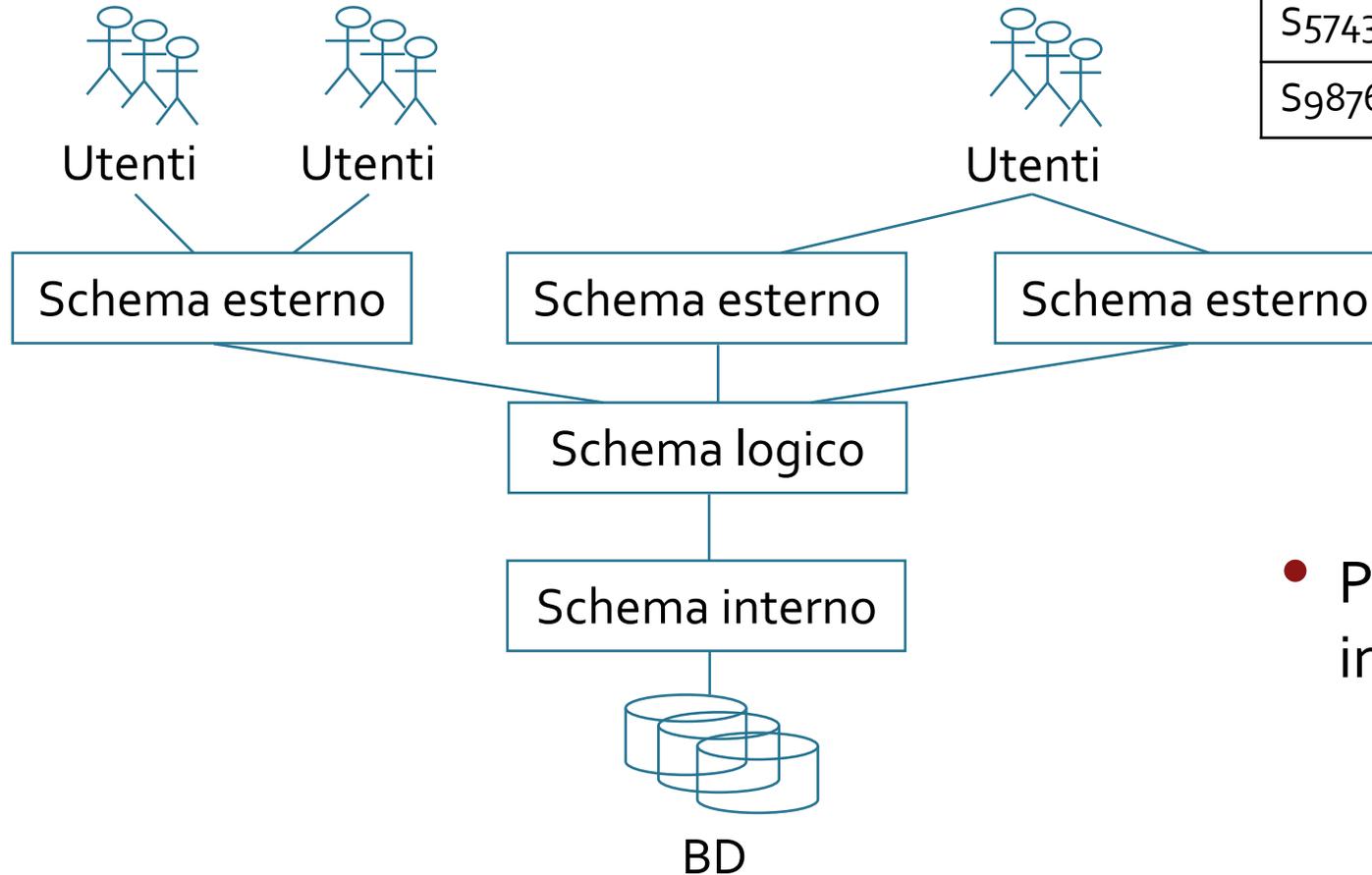
Cdl: INGEGNERIA DEL CINEMA E DEI MEZZI DI COMUNICAZIONE (ingegneria dell'informazione ▾)

Cerca

Visualizzazione Giornaliera Stampa

	lunedì 05/06/2023	martedì 06/06/2023	mercoledì 07/06/2023	giovedì 08/06/2023	venerdì 09/06/2023
8 ⁰⁰					
9 ⁰⁰	Comunicare il patrimonio cu... MOREZZI EMANUELE AA - ZZ - 0 4VM Lezione/Esercitazione	Diritto ed etica della comu... CIURCINA MARCO AA - ZZ - 0 7B	Transmedia MORREALE DOMENICO AA - ZZ - 0 LAIB5 Lezione/Esercitazione SQUADRA 2	Diritto ed etica della comu... CIURCINA MARCO AA - ZZ - 0 Transmedia MORREALE DOMENICO AA - ZZ - 0 8I Lezione/Esercitazione	
10 ⁰⁰					
11 ⁰⁰		Computer grafica BOTTINO ANDREA	Transmedia MORREALE DOMENICO AA - ZZ - 0 LAIB5 Lezione/Esercitazione SQUADRA 1		
12 ⁰⁰				Computer grafica BOTTINO ANDREA AA - ZZ - 0 8I Lezione/Esercitazione	
13 ⁰⁰					Elaborazione di immagine e ... BIANCHI TIZIANO

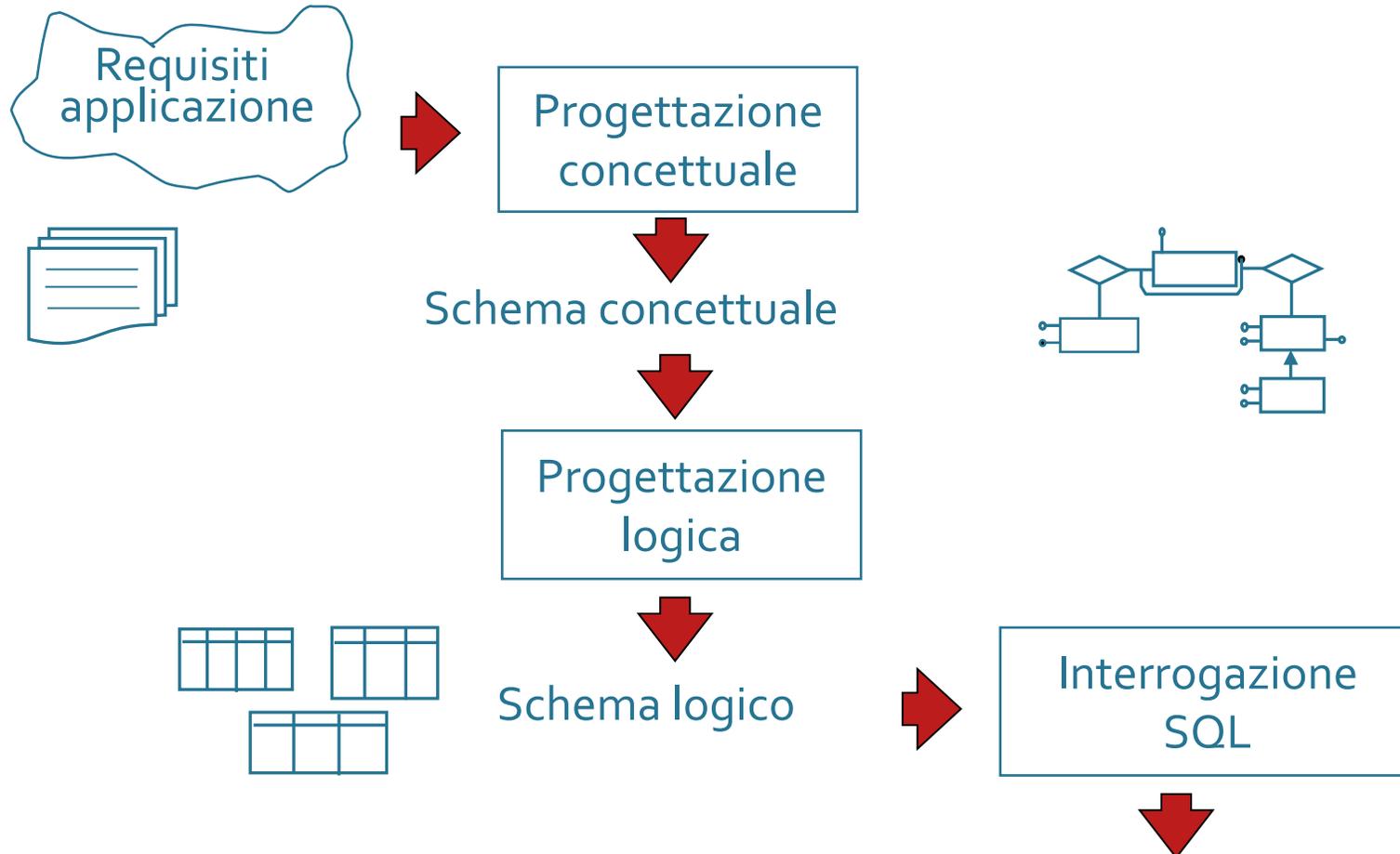
I livelli di astrazione di un DBMS



Matricola	Codice	Data	Voto
S123456	AK9876	2022-01-30	28
S321765	AG7654	2022-01-25	15
S574375	AG7654	2022-02-21	25
S987654	AK9876	2022-01-30	27

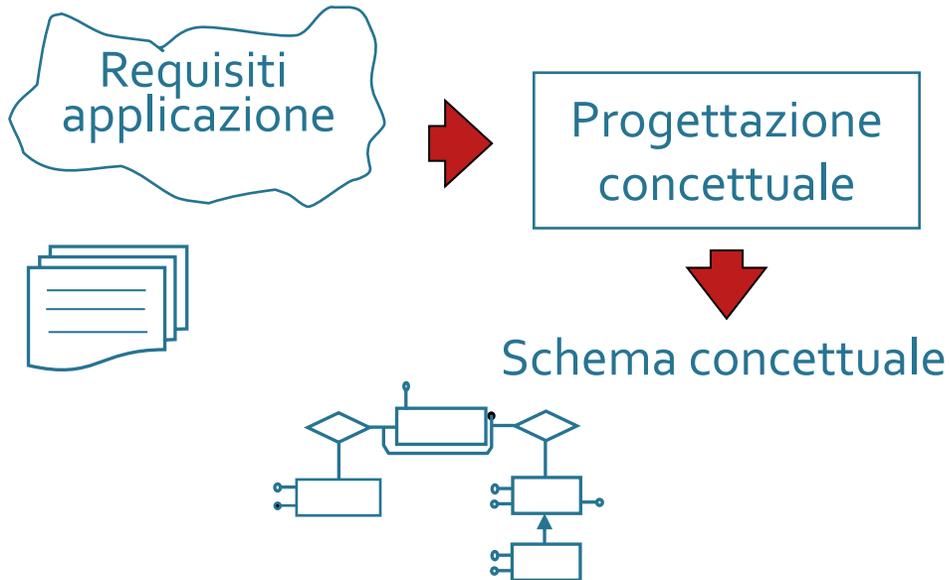
- Proprietà fondamentale: indipendenza dei dati

Una visita guidata del corso (per l'85%)



```
SELECT name
FROM employees
WHERE name LIKE 'B%' AND city = 'Paris';
```

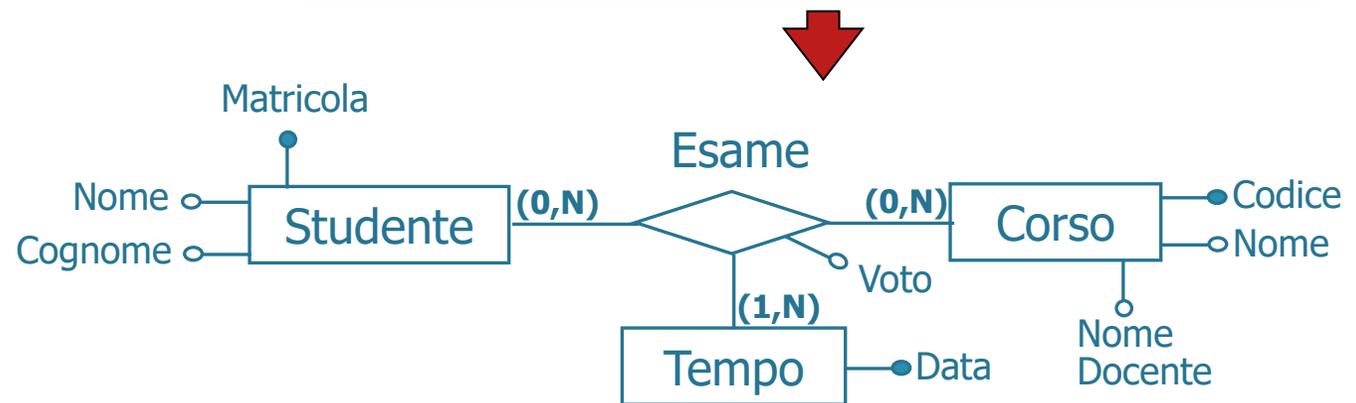
Una visita guidata del corso (per l'85%)



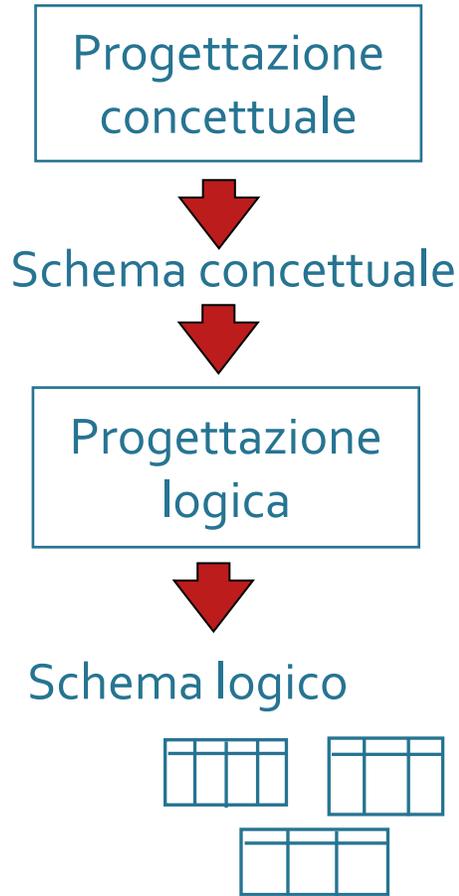
Gli studenti, identificati dalla matricola, e di cui sono noti nome e cognome, sostengono gli esami dei corsi ottenendo un voto.

I corsi sono identificati da un codice e sono caratterizzati dal nome del corso e dal nome del docente.

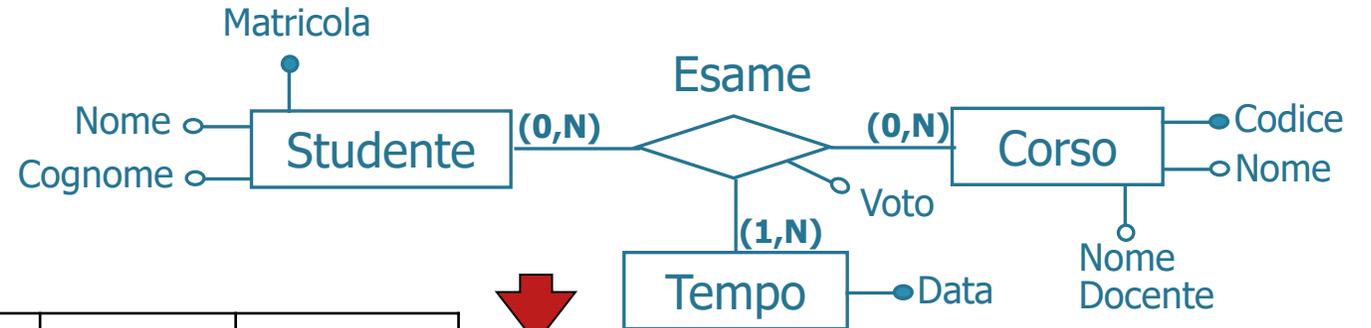
Si noti che gli studenti possono sostenere più volte lo stesso esame, ma non più volte nello stesso giorno.



Una visita guidata del corso



Introduzione al corso



STUDENTE

Matricola	Nome	Cognome
s12345	Luca	Rossi
s23456	Adele	Bianchi
s34567	Angelica	Neri

ESAME

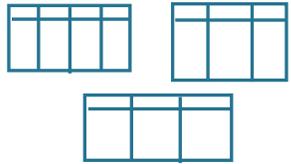
Matricola	Codice	Data	Voto
S123456	AK9876	2022-01-30	28
S321765	AG7654	2022-01-25	15
S574375	AG7654	2022-02-21	25
S987654	AK9876	2022-01-30	27

CORSO

Codice	Nome	NomeDoc
AK9876	Basi Dati	Ferrero
AF8765	Informatica I	Verdi
AG7654	Analisi II	Rosano

STUDENTE (Matricola, Cognome, Nome)
 CORSO (Codice, Nome, NomeDocente)
 ESAME (Matricola, Codice, Data, Voto)

Una visita guidata del corso



Schema logico



Interrogazione
SQL

STUDENTE (Matricola, Cognome, Nome)
CORSO (Codice, Nome, NomeDocente)
ESAME (Matricola, Codice, Data, Voto)

Trovare cognome e nome di tutti gli studenti che nel 2021 hanno superato almeno un esame con voto maggiore di 25.



```
SELECT Cognome, Nome
FROM STUDENTE S, ESAME E
WHERE S.Matricola = E.Matricola AND Voto > 25
AND Data >= '2021-01-01'
AND Data <= '2021-12-31';
```

Per ciascun corso, visualizzarne il nome e la media dei voti degli studenti.



```
SELECT Nome, AVG (Voto) AS MediaVoti
FROM CORSO C, ESAME E
WHERE C.Codice = E.Codice
GROUP BY C.Codice, Nome;
```

Requisiti didattici

- Capacità di
 - Analisi e comprensione di un problema
 - Definizione di una soluzione progettuale
- Conoscenze di base di programmazione

Materiale

- Sito web del corso
 - https://dbdmg.polito.it/dbdmg_web/2024/base-di-dati-gestionale/
 - Slide, esercizi, materiale per i laboratori, ...
- Portale della didattica
 - Avvisi
 - Consegna/valutazione homework (via piattaforma Moodle)
- Mail ufficiale degli studenti / SMS
 - Per comunicazioni urgenti (cambio aula, cancellazione di lezioni o esercitazioni, ...)
 - Scrivete solo dalla mail ufficiale
- Le lezioni saranno registrate

Orario del corso

- Lunedì 14.30 – 17.30, aula 27
- Giovedì 11.30 – 14.30 (o 11.30 – 13.00), aula 27
- Martedì 11.30 – 14.30, LAIB 1 (a partire dal 15/10)
 - Squadra 1 dalle 11.30 alle 13.00
 - Squadra 2 dalle 13.00 alle 14.30
- Lo slot del giovedì 13.00 – 14.30 sarà a settimane alterne (è in comune con il corso parallelo)
 - Consultate l'orario ufficiale sul sito di Ateneo
- Eventuali lezioni di recupero saranno in Virtual Conference

Il laboratorio

- Capacità di scrivere interrogazioni in linguaggio SQL e MongoDB
- Capacità di utilizzare ed interrogare dati reali
- Due squadre
- Laboratori da 1,5 ore
- Totale: 12 ore di laboratorio
- A partire dal 15 ottobre

Docenti

- Laura Farinetti (lezioni, esercitazioni)
 - Dipartimento di Automatica e Informatica
 - E-mail: laura.farinetti@polito.it
- Alessandro Fiori (lezioni, esercitazioni)
 - E-mail: alessandro.fiori@polito.it
- Giuseppe Gallipoli (laboratorio)
 - E-mail: giuseppe.gallipoli@polito.it
- Davide Napolitano (laboratorio)
 - E-mail: davide.napolitano@polito.it

Materiale di supporto

- Testo di riferimento
 - Atzeni, Ceri, Paraboschi, Torlone, "Basi di dati - Modelli e linguaggi di interrogazione", 4 ed., McGraw Hill, 2013.
- Esercizi svolti
 - Baralis, Belussi, Psaila, "Basi di dati: temi d'esame svolti", Esculapio, 2000
- Altri testi
 - Raghu Ramakrishnan, Johannes Gehrke, "Sistemi di basi di dati", McGraw-Hill, 2004.
 - Albano, Ghelli, Orsini, "Fondamenti di basi di dati, Zanichelli, 2005

Esame

- Esame scritto
 - Abilità nel problem solving (+++): interrogazioni e progettazione
 - Fondamenti teorici (+)
 - NON si può consultare materiale né cartaceo né elettronico
- Criteri, regole e procedure di valutazione disponibili sulla scheda d'insegnamento
 - I dettagli delle regole d'esame saranno pubblicati sul sito
- Può dare l'esame solo chi ha il corso nel carico didattico
- **IMPORTANTE:** è l'ultima edizione di questo corso
 - Il corso passa al I anno
 - Appelli fino a settembre 2026

Homework

- Gli homework svolti durante l'anno consentono di aggiungere alcuni punti (2) al voto finale (solo se sufficiente)
- Attività da svolgere in parte in laboratorio e in parte a casa, e da consegnare rispettando la scadenza
 - Attività in laboratorio -> 0,5 punti (partecipazione in presenza + consegna)
 - Attività a casa -> 0,5 (analisi + consegna)
 - Consegna su piattaforma Moodle
- Sono due
 - SQL (laboratorio a metà novembre)
 - Progettazione concettuale e logico-relazionale (laboratorio a metà dicembre)
- Utilizzo di Large Language Models (generative AI)
- Validi fino all'appello di settembre 2025

License

- This work is licensed under the Creative Commons “Attribution-NonCommercial-ShareAlike Unported (CC BY-NC-SA 3,0)” License.
- You are free:
 - to Share - to copy, distribute and transmit the work
 - to Remix - to adapt the work
- Under the following conditions:
 - Attribution - You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).
 - Noncommercial - You may not use this work for commercial purposes.
 - Share Alike - If you alter, transform, or build upon this work, you may distribute the resulting work only under the same or similar license to this one.
- To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/3.0/>

