

16AFQPI - Basi di Dati

Introduzione al corso

Obiettivo del corso

- Progettare una base dati relazionale
- Interrogare una base dati relazionale in linguaggio SQL e in algebra relazionale
- Utilizzare i Large Language Model (GenAI) per risolvere interrogazioni in SQL e allenare il pensiero critico
- Perché le basi dati sono così importanti???

Sistemi informativi e basi di dati

- I sistemi informativi

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

struttura

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	14AFQPL 14AFQPI
▶ 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	14AFQPL 14AFQPI
▶ 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

 **Verballi d'esame**

 Procedura da utilizzare per la verbalizzazione delle prove d'esame degli studenti.

[ENTRA](#)

Per problemi di accesso alla procedura contattare help.verballi@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

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

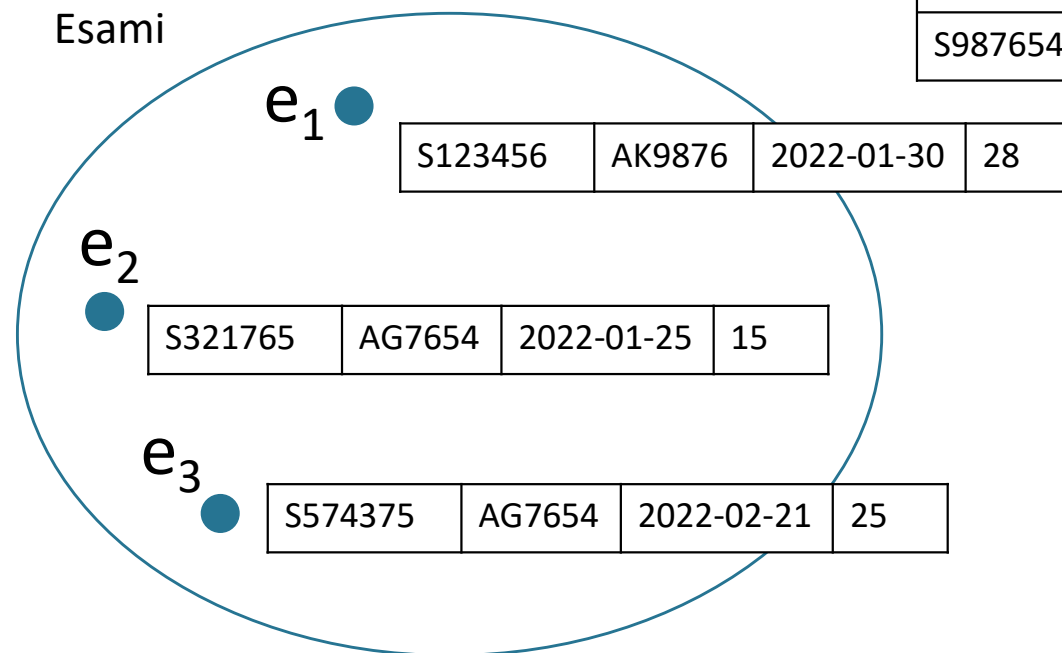
- Serve un modello dei dati
- Noi useremo il modello relazionale
 - Basato sul concetto matematico di «relazione» (informalmente «tabella»)

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

- 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**

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

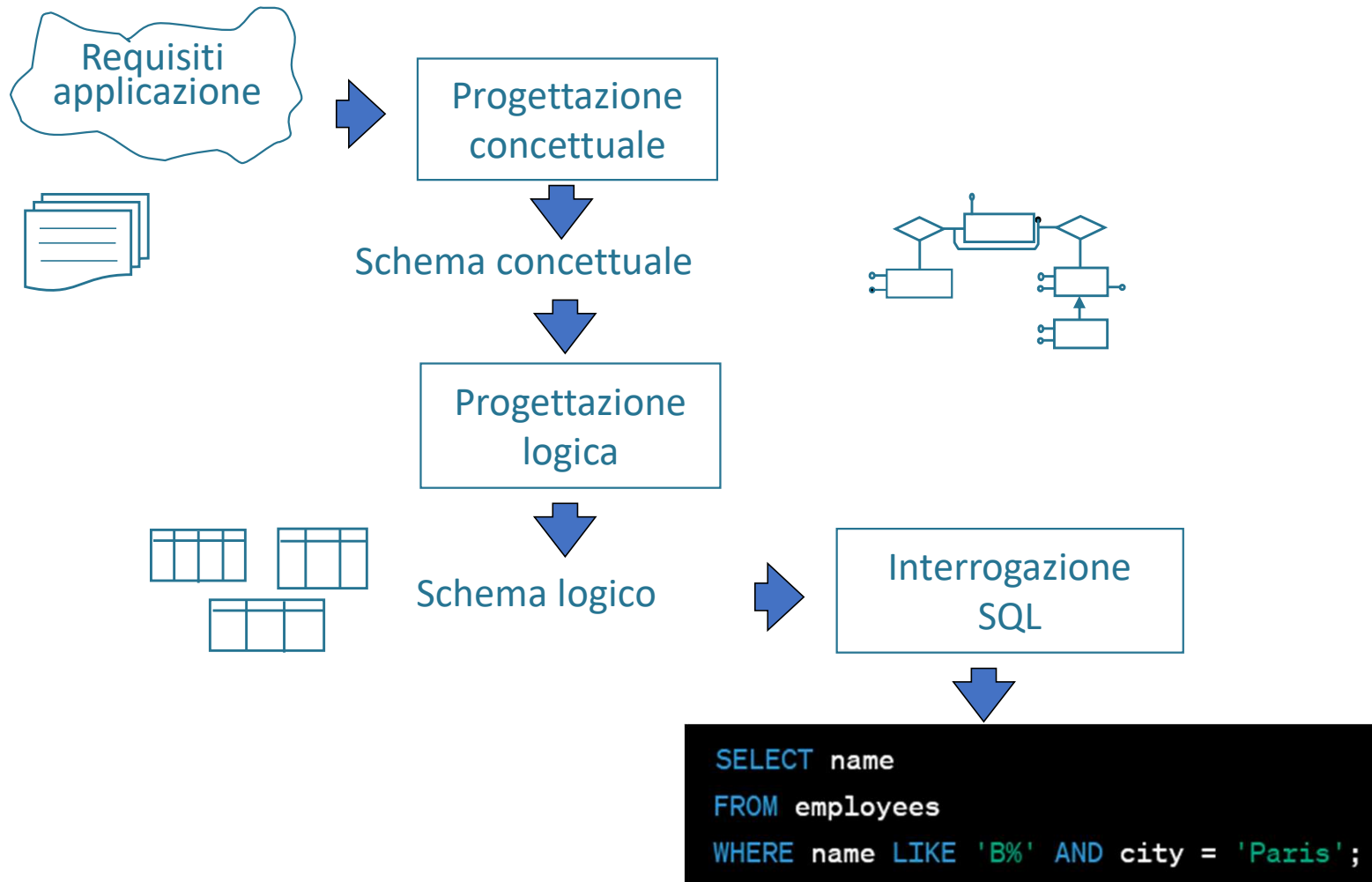
- Pensando alle basi dati in senso più ampio, si può accedere alle informazioni “dialogando” con un servizi interattivi che “parlano” in linguaggio naturale
- Chatbot, or “AI assistants”
- Large Language Models (LLM)
 - Algoritmi AI di deep-learning pre-addestrati su grandi quantità di dati, capaci di riconoscere, generare, riassumere, tradurre e prevedere contenuti



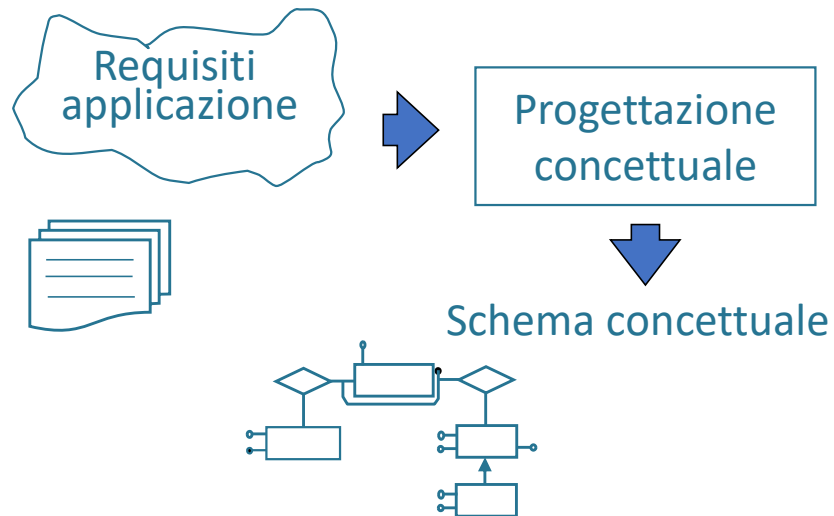
Requisiti didattici

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

Una visita guidata del corso



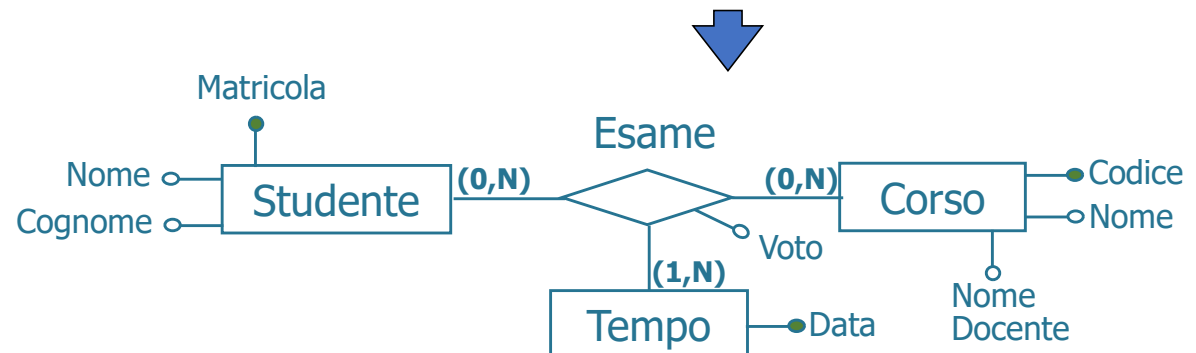
Una visita guidata del corso



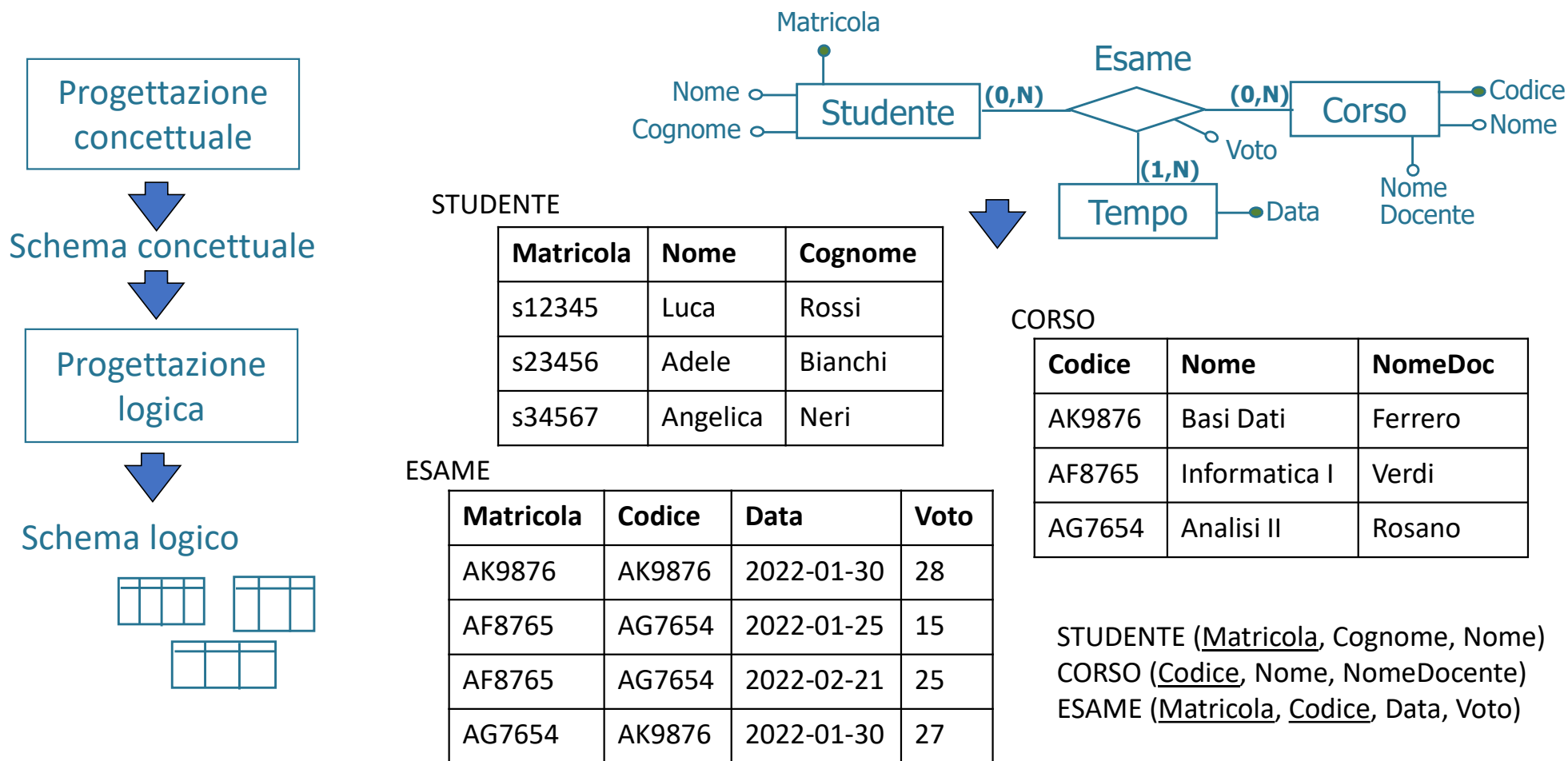
Gli studenti, identificati dalla matricola, e di cui sono noti nome e cognome, sostengono gli esami dei corsi ottenendo un voto.

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

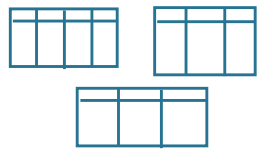
Si noti che gli studenti possono sostenere più volte lo stesso esame, ma non più volte nello stesso giorno.



Una visita guidata del corso



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;
```

Il laboratorio (12 ore)

- 5 esercitazioni su SQL o progettazione
- 2 esercitazioni con l'utilizzo di Large Language Model (GenAI) per risolvere query SQL
 - 1 laboratorio da 1,5 ore + 1 laboratorio da 3 ore (13 maggio)
- I laboratori si svolgono in aula il mercoledì dalle 17.30 alle 19.00, senza suddivisione in squadre
- Occorre portare il proprio PC
- Mercoledì 13 maggio: laboratorio 16.00-19.00 (3 ore)
 - Importante: laboratorio legato all'homework -> presenza + consegna lavoro
- Inizio dei laboratori: 18 marzo

Il laboratorio

- Capacità di scrivere interrogazioni in linguaggio SQL
- Laboratori da 1,5 ore in aula: mercoledì 17.30-19.00
- Occorre portare il proprio PC

- Mercoledì 13 maggio: laboratorio 16.00-19.00 (3 ore)
 - Importante: laboratorio legato all'homework -> presenza + consegna lavoro

- Totale: 12 ore di laboratorio
- A partire dal 18 marzo

Docenti

- Laura Farinetti (lezioni, esercitazioni)
 - Dipartimento di Automatica e Informatica
 - E-mail: laura.farinetti@polito.it
- Luca Cagliero(lezioni, esercitazioni)
 - E-mail: luca.cagliero@polito.it
- Giuseppe Gallipoli, Alberto Verna (laboratorio)
 - E-mail: giuseppe.gallipoli@polito.it
 - E-mail: albero.verna@polito.it

Materiale

- Siti web del corso
 - Sito web del corso: slide, esercizi, materiale per i laboratori, ...
 - https://dbdmg.polito.it/dbdmg_web/basi-di-dati-16afqpl-16afqpi/
- Portale della didattica
 - Avvisi
- Mail ufficiale degli studenti
 - Per comunicazioni urgenti (cambio aula, cancellazione di lezioni o esercitazioni, ...)
 - Scrivete solo dalla mail ufficiale

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 su carta
 - 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
- Può dare l'esame solo chi ha il corso nel carico didattico

Homework

- Gli homework consentono di aggiungere 2 punti al voto finale (se sufficiente)
- Laboratorio del 13 maggio
 - Partecipazione: 0,5 punti
 - Consegna valida: 0,5 punti
- Esercizi da consegnare
 - Algebra + SQL: 0,5 punti
 - Progettazione: 0,5 punti
- Validi fino all'appello di gennaio 2027