

04AFQPC - Basi di Dati

a.a. 2024-2025

INTRODUZIONE AL CORSO



**POLITECNICO
DI TORINO**



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

- 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

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

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		

Verbalì 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 il modello relazionale
 - Basato sul concetto matematico di «relazione» (informalmente «tabella»)

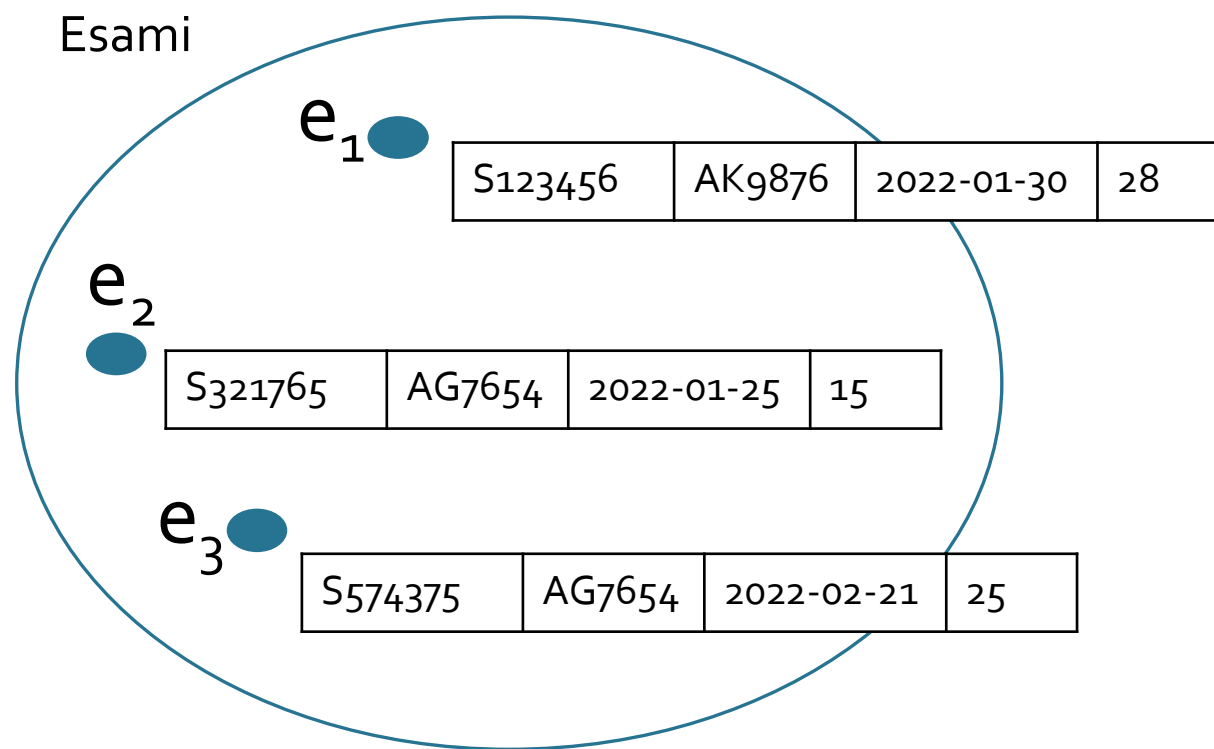
Stato	Continente	Popolazione (M)
Cile	America	18
Australia	Oceania	25
Italia	Europa	59

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



- Pensando alle basi dati in senso più ampio, si può accedere alle informazioni “dialogando” con 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,



Obiettivi del corso

- Progettare una base dati relazionale
- Scrivere interrogazioni in linguaggio SQL
- Soft skills
 - Pensiero critico, creatività, team work, interazione
- Il corso aderisce al progetto OCSE dal titolo «Fostering and assessing creativity and critical thinking in higher education and teacher education»



Argomenti del corso

- I sistemi per la gestione di basi di dati
- Il modello relazionale dei dati
- La metodologia di progettazione concettuale e logica relazionale di una base di dati
- Il linguaggio SQL
- Retrieval Model (RM) e Large Language Model (LLM)

Requisiti di base

- Capacità di
 - Analisi e comprensione di un problema
 - Definizione di una soluzione progettuale
- Conoscenza di Python
 - Utile per il laboratorio

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 privacy

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 ▾

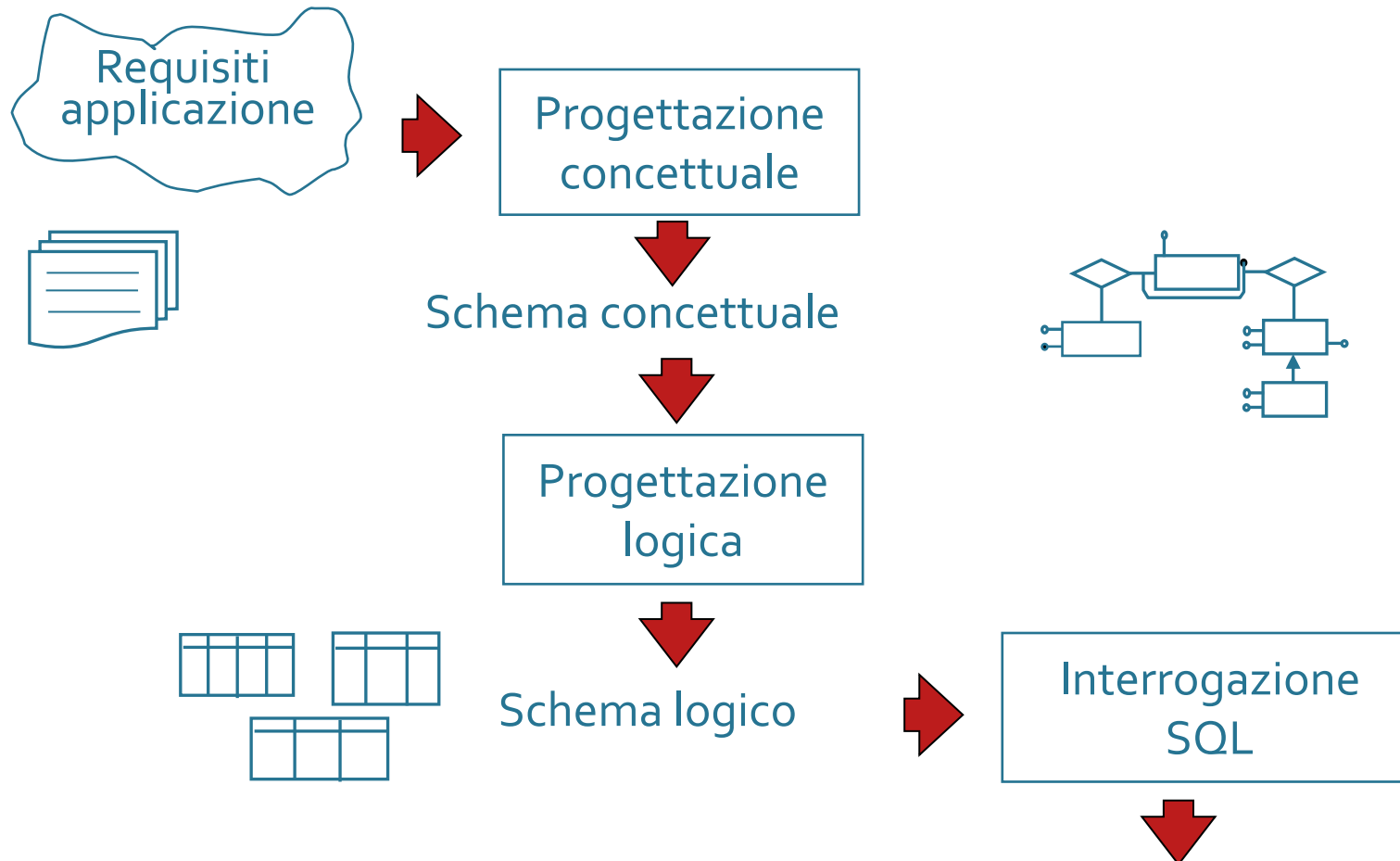
Area: Ingegneria ▾

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 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	

Una visita guidata del corso



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

Una visita guidata del corso

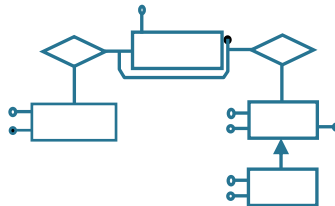
Requisiti
applicazione



Progettazione
concettuale



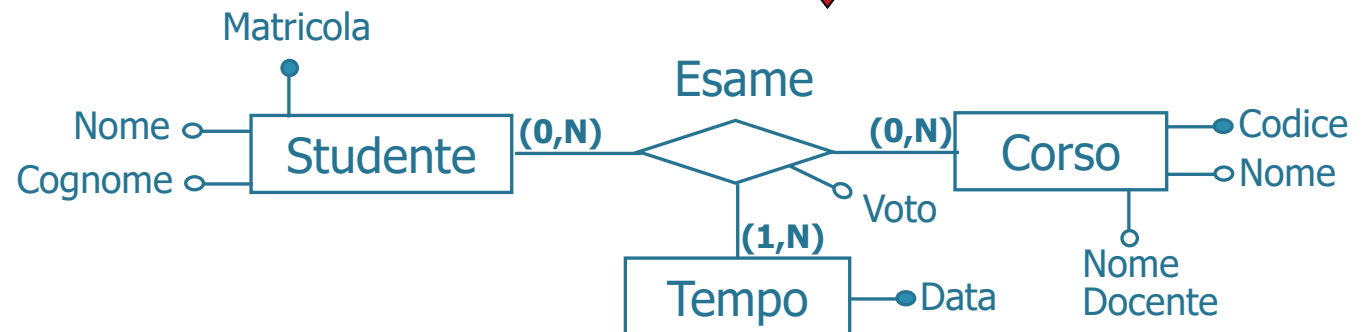
Schema concettuale



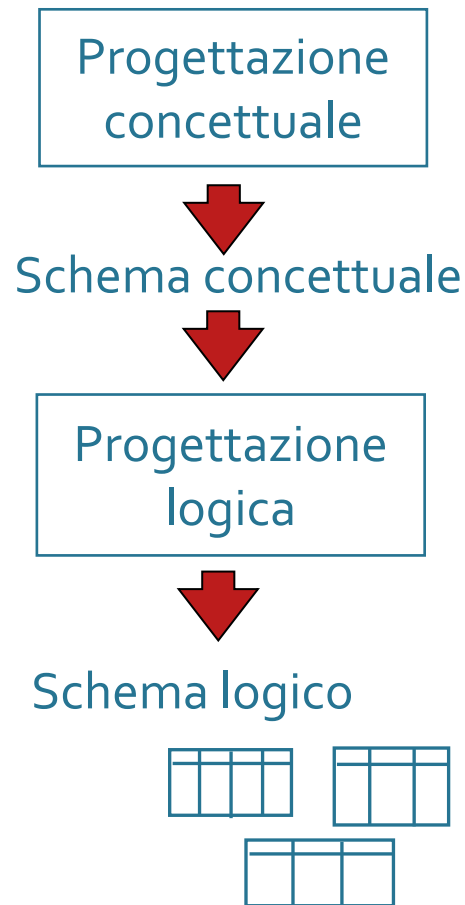
Gli studenti, identificati dalla matricola, e di cui sono noti nome e cognome, sostengono gli esami dei corsi ottenendo un voto.

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

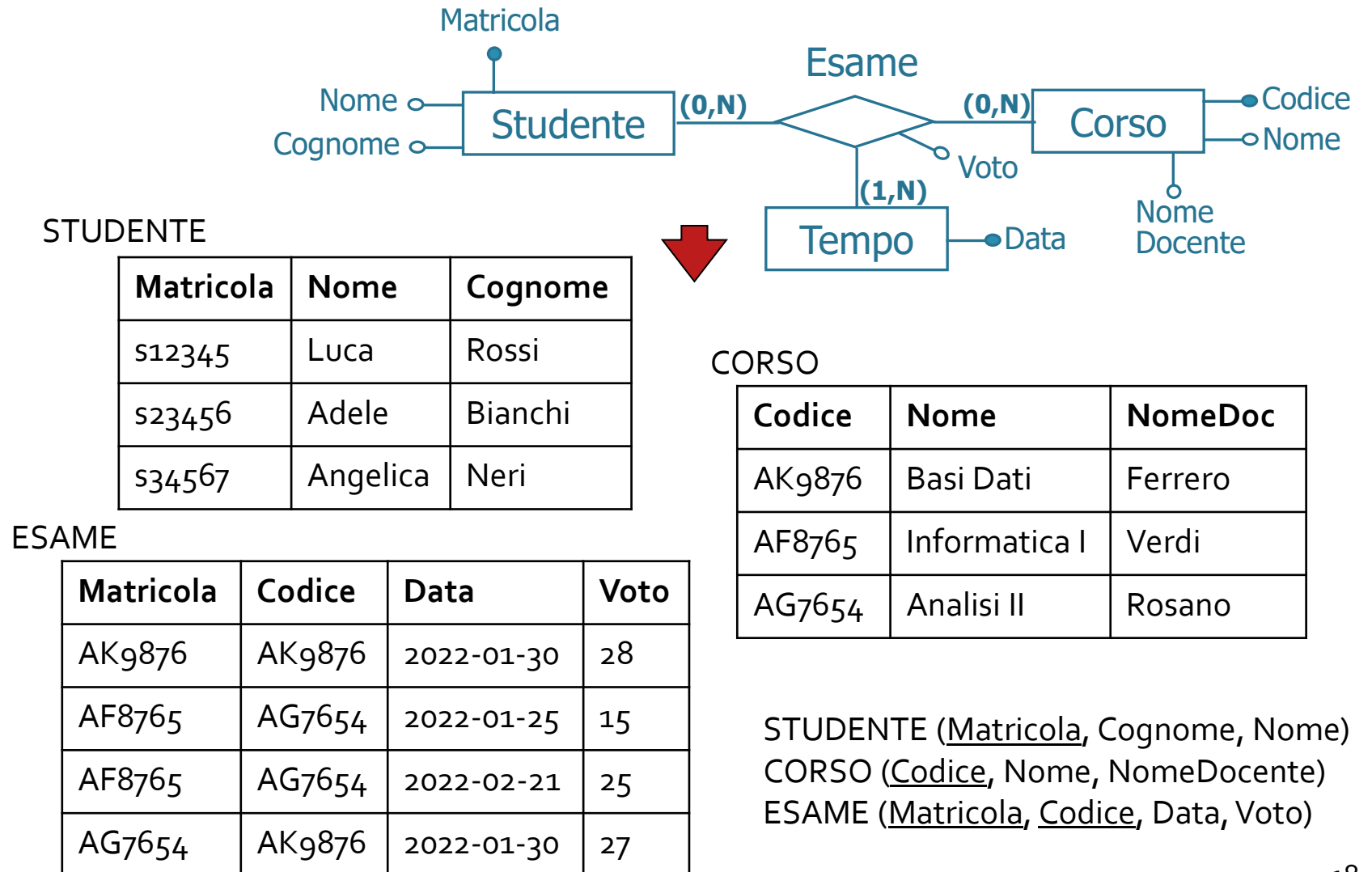
Si noti che gli studenti possono sostenere più volte lo stesso esame, ma non più volte nello stesso giorno.



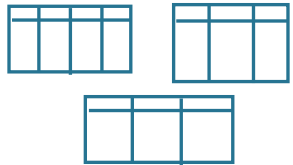
Una visita guidata del corso



Introduzione al corso



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;
```

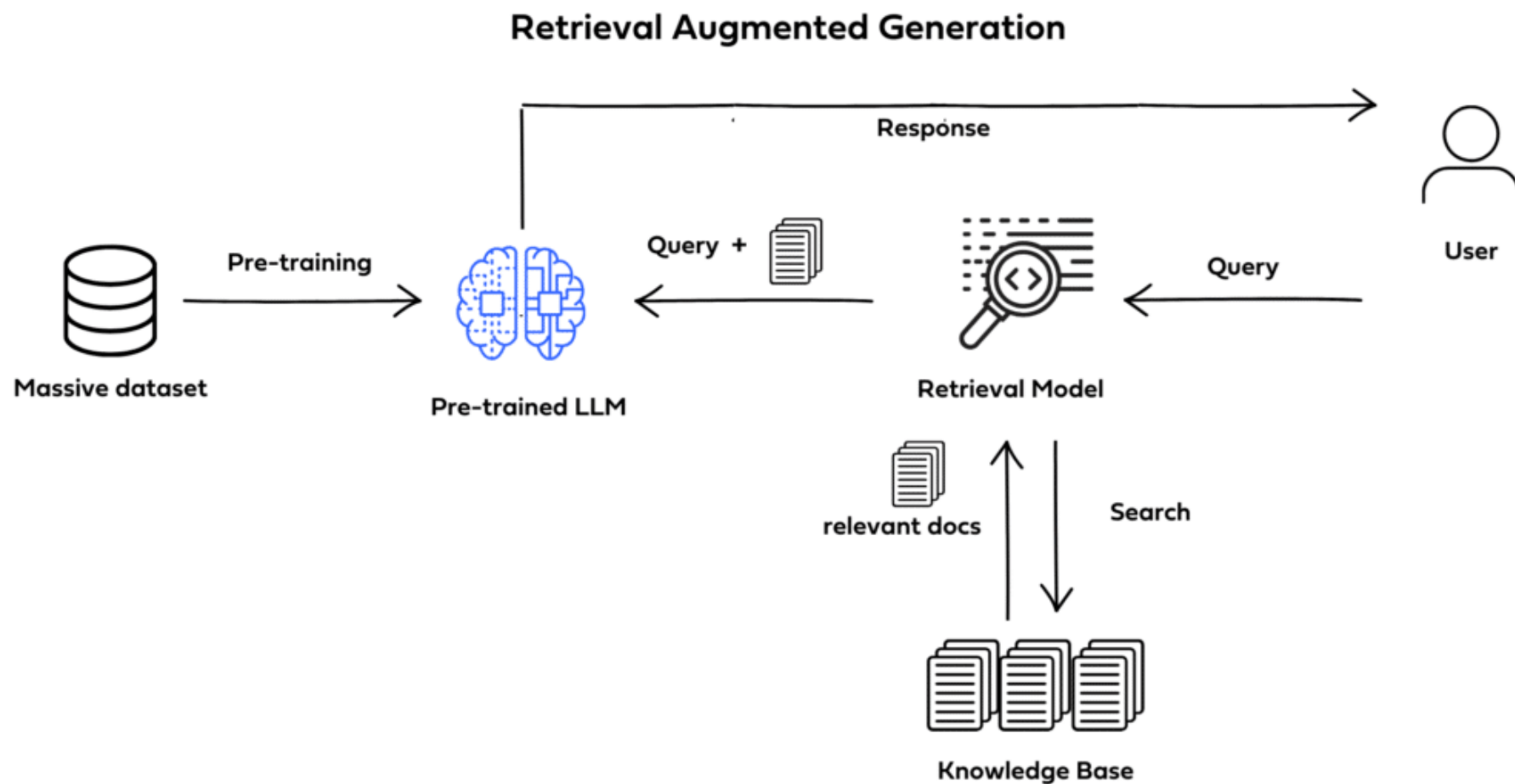
Il laboratorio

- Capacità di scrivere interrogazioni in linguaggio SQL
 - 4 esercitazioni «classiche»
- Capacità di utilizzare ed interrogare dati reali
 - 4 esercitazioni sulla creazione di un chatbot Telegram integrando un Retrieval Model (RM) ed un Large Language Model (LLM) – gruppi da 2 persone
- Due squadre
- Laboratori da 1,5 ore
 - Totale: 12 ore di laboratorio
- A partire dal 23 ottobre

I laboratori 4-7: 4 fasi

- Scelta dei documenti e delle domande
 - Identificare i documenti chiave a cui il chatbot farà riferimento
 - Definire il tipo di domande a cui il chatbot dovrà rispondere basandosi su questi documenti
- Creazione della prima versione del chatbot e valutazione delle risposte
 - Sviluppare la prima versione del chatbot integrando RM e LLM
 - Valutare la qualità delle risposte
- Fine-tuning
 - Perfezionare il chatbot in base alla valutazione, lavorando sui modelli
- Integrazione con Telegram e valutazione fra pari
 - Integrare il chatbot con Telegram per l'interazione con gli utenti in tempo reale
 - Valutazione fra pari: for real-time user interaction. Pair evaluations involve peer-to-peer assessments, where students in the course evaluate each other's chatbot performance.

RM + LMM -> Retrieval Augmented Generation (RAG)



Materiale

- Sito web del corso
 - <https://elite.polito.it/teaching/o4afqpc-bdcin>
 - Tutto il materiale (tranne le videolezioni e i quiz): slide, esercizi proposti, soluzioni, temi d'esame, ...
 - «Piano del corso»: calendario aggiornato delle attività e delle scadenze
- Portale della didattica
 - Videolezioni
 - Quiz (piattaforma Moodle)
- Testo di riferimento
 - Atzeni, Ceri, Paraboschi, Torlone, "Basi di dati - Modelli e linguaggi di interrogazione", 4 ed., McGraw Hill, 2013.

Videolezioni e quiz

- 13 videolezioni pre-registrate da 1 ora
 - Per il ripasso (o chi non riesce a frequentare)
- Quiz (formative assessment) su tutti gli argomenti
 - Feedback significativo
 - Obiettivo: incentivare lo studio progressivo, rendere più efficace la partecipazione ad esercitazioni e laboratori
 - Scadenza per la sottomissione dei quiz (circa 1 settimana dopo l'apertura)
 - Si può fare 1 SOLA VOLTA
 - Utile anche per me, per chiarire i concetti poco chiari
 - Punti extra (vedere regole d'esame)

Comunicazione con i docenti

- Laura Farinetti
 - laura.farinetti@polito.it
- Lorenzo Canale
 - lorenzo.canale@polito.it
- Sito web di riferimento
 - <https://elite.polito.it/teaching/o4afqpc-bdcin>
- Portale della didattica e email per le comunicazioni ufficiali
 - Scrivete solo dalla mail ufficiale

Orario del corso

- **IMPORTANTE:** consultare il «Piano del corso» sul sito web, che verrà costantemente aggiornato con le attività previste

- Mercoledì 8.30 – 11.30, aula 9T
- Giovedì 8.30-10.00, aula 10I
- Martedì 16.00-17.30 (qualche rara volta, quando non ci sono laboratori)
- Laboratori (squadre «a coppie» -> iniziate a formarle)
 - Una squadra: martedì 16.00 – 17.30, aula 5T
 - Una squadra: martedì 17.30 – 19.00, aula 5T
 - Settimane alterne

I quiz

- Su Moodle sono pubblicati 13 quiz (formative assessment) relativi agli argomenti visti a lezione, da svolgere in un intervallo di tempo specifico (circa una settimana)
 - Obiettivo: incentivare lo studio progressivo e rendere più efficace la partecipazione ad esercitazioni e laboratori
 - La sottomissione dei quiz entro la scadenza consente di ottenere alcuni punti extra (vedere le Regole d'esame)
 - Non sottovalutateli (le domande non sono banali), e leggete i feedback!!!

Importante

- Consultare regolarmente il «Piano del Corso» pubblicato sul sito
 - Calendario delle lezioni
 - Esercizi da svolgere prima delle esercitazioni in aula

Calendario delle lezioni

La seguente tabella riporta la pianificazione delle lezioni del corso, per trasparenza di informazione e per facilitare la vostra pianificazione. Ci riserviamo di modificare la pianificazione delle lezioni per esigenze didattiche.

Legenda: la colonna Tipo può essere Lez=Lezione, EsT=Esercitazione "tradizionale" in aula, EsCCT=Esercitazione "CCT" in aula, Lab=Esercitazione di laboratorio, Cons=Consulenza. Le attività in nero sono per tutti, quelle in **viola** o in **verde** sono attività a squadre (**squadra A** e **squadra B** rispettivamente).

Data	Orario	Tipo	Ore	Argomento	Aula	Docente
lun 2/10/23	10.00-13.00	Lez	3	Introduzione al corso. Introduzione alle BD.	LAIB 3B	L. Farinetti
lun 9/10/23	10.00-11.30	EsT	1,5	Esercitazione SQL#1.	LAIB 3B	L. Farinetti
lun 9/10/23	11.30-13.00	Cons	1,5	Consulenza su videolezioni/esercizi.	LAIB 3B	L. Farinetti

L'esame

- Esame scritto in aula: 20/30
 - SQL, progettazione concettuale e logico-relazionale
- Attività di laboratorio (laboratori 4-7): 10/30
 - Consegna di materiale/report (gruppi da due) + domande durante lo scritto (individuale)
 - Alternativa per i non frequentanti: progetto individuale + domande durante lo scritto
- Punti extra per la sottomissione dei quiz (formative assessment): max 3
 - Dipendono dal numero di quiz sottomessi entro la scadenza (minimo la metà) e dalla qualità
- Dettagli nelle regole d'esame (pubblicate a breve sul sito)
- Può dare l'esame solo chi ha il corso nel carico didattico

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/>

