

## COURSE SYLLABUS

### Data Architecture

2526-1-F1802Q100

---

#### Obiettivi

Il corso, erogato in lingua Italiana, intende fornire conoscenze e competenze nell'ambito delle moderne architetture dati. Veranno presentati sia gli aspetti teorici e di ricerca ancora aperti sia come tali tecnologie possono essere utilizzate per risolvere specifici contesti applicativi. Particolare attenzione sarà posta sullo studio e analisi di come le architetture e le soluzioni presentate consentano l'esecuzione di scritture, anche concorrenti in ambito distribuito e letture in ambiente distribuito e in presenza di possibili malfunzionamenti sulla rete. Saranno anche presentate le modalità di uso dei sistemi di ai generativa applicati alla gestione dei dati

Al termine del corso lo studente sarà in grado di scegliere la più opportuna soluzione architetturale per la gestione di dati

#### Contenuti sintetici

Attraverso la presentazione di use case diversi verranno affrontati gli aspetti teorici, metodologici e applicativi previsti dal corso. Ogni use case intende presentare dei problemi per i quali sono previste delle nuove soluzioni in termini architetturali rispetto al quanto visto fino a quel momento

1. sistemi relazionali distribuiti
2. sistemi non relazionali e poliglotti
3. data centric ai (data management per machine learning)
4. Al generativa per la gestione dei dati

#### Programma esteso

User cases 1

Richiami di architetture relazionali centralizzate: transazioni, ottimizzazione query  
Architetture dbms distribuiti: protocollo two phase commit, deadlock distribuito  
Use case 2  
Modelli non relazionali  
Architetture distribuite di sistemi non relazionali  
Sistemi poliglotti  
Use case 3  
Data management for machine learning  
Data understanding  
Data validation  
Data preparation ( integration, quality, fusion)  
Future engineering  
MLOps  
Use case 4  
Modelli di AI generativa  
Fine tuning  
Architetture RAG

## Prerequisiti

E' utile una conoscenza delle tematiche dei modelli dei dati, relazionale e Entità Relazione, del linguaggio SQL e delle metodologie di progettazione di basi di dati

## Modalità didattica

Sono previste in totale 48 ore per il corso; così organizzate  
32 ore di lezioni in aula in presenza  
16 ore di esercitazioni in aula in presenza

sia le ore di lezioni che quelle di esercitazioni saranno di tipo interattivo. Il docente porrà delle domande per sviluppare la capacità di ragionamento critico degli studenti

## Materiale didattico

slide pubblicate sulla piattaforma elearning

Libro di testo adottato

Text book Next Generation Databases: NoSQL, NewSQL, and Big Data by Guy Harrison Publisher: Apress  
Release Date: January 2016 ISBN 9781484213292

## Periodo di erogazione dell'insegnamento

secondo semestre

## **Modalità di verifica del profitto e valutazione**

Prove scritta con domande a risposte aperte sui contenuti del corso, esercizi numerici eventualmente seguite da un orale a richiesta del docente

In alternativa progetto relativo a uno degli use cases presentato durante il corso. il progetto può essere esteso a tesi e può essere fatto anche a gruppi di studenti. l'argomento del progetto deve essere concordato con il docente

## **Orario di ricevimento**

Si prega mandare una email in anticipo indicando il nome del corso. il ricevimento può essere effettuato anche online o al termine delle lezioni

## **Sustainable Development Goals**

ISTRUZIONE DI QUALITÀ

---