

COMPUTER ENGINEERING (LM75)

(Lecce - Università degli Studi)

Teaching BIG DATA MANAGEMENT

GenCod A006801

Owner professor MARCO SALVATORE ZAPPATORE

Reference professors for teaching ANTONELLA LONGO, MARCO SALVATORE ZAPPATORE

Teaching in italian GESTIONE DI BIG DATA

Teaching BIG DATA MANAGEMENT

SSD code ING-INF/05

Reference course COMPUTER ENGINEERING

Course type Laurea Magistrale

Credits 6.0

Teaching hours Front activity hours: 54.0

For enrolled in 2022/2023

Taught in 2022/2023

Course year 1

Language ITALIAN

Curriculum PERCORSO COMUNE

Location Lecce

Semester First Semester

Exam type Oral

Assessment Final grade

Course timetable

<https://easyroom.unisalento.it/Orario>

BRIEF COURSE DESCRIPTION

Il corso ha l'obiettivo di fornire agli studenti la possibilità di apprendere i concetti e i paradigmi principali della generazione, gestione e archiviazione di big data, con particolare riferimento ad architetture, modelli e applicazioni.

REQUIREMENTS

Buona conoscenza di linguaggi orientati agli oggetti (almeno 1), tecniche e strumenti. Elementi di reti informatiche e tecnologie per il Web.

COURSE AIMS

- Comprendere i principi alla base della gestione dei Big Data, della loro governance e della loro generazione
 - Conoscere i modelli dati e le architetture per i Big Data
 - Apprendere i criteri alla base della Big Data Quality
 - Conoscere gli approcci di analisi esplorativa dei Big Data
 - Comprendere definizioni, caratteristiche e differenze di Data Warehouse, Data Lake e Data Lakehouse
 - Conoscere i sistemi e le piattaforme a supporto della gestione dei Big Data
 - Conoscere i DBMS NoSQL e NewSQL

TEACHING METHODOLOGY

Le lezioni frontali mirano a migliorare la conoscenza e la comprensione degli studenti attraverso la presentazione di teorie, modelli e metodi; gli studenti sono invitati a partecipare alla lezione con autonomia di giudizio, ponendo domande e presentando esempi. Le esercitazioni di laboratorio sono finalizzate all'utilizzo di strumenti che supportano i modelli e gli approcci presentati.

ASSESSMENT TYPE

1. Prova di ammissione al progetto (prova scritta obbligatoria)
2. Progetto (prova obbligatoria che consta di implementazione e successiva discussione orale, su argomento concordato con i docenti)
3. Data hackathon (prova opzionale implementazione e successiva discussione orale)

FULL SYLLABUS

- Presentazione del corso
- Database relazionali ed SQL
- Introduzione alla gestione dei Big Data
- Modelli dati, Big Data Quality
- Exploratory Data Analysis (EDA)
- Architetture per Big Data
- Distributed database
- Data Warehouse
- Data Lake, Data Lakehouse
- Introduzione alla virtualizzazione, Map Reduce
- Apache Flask, Apache Hop
- PostgreSQL
- Introduzione alla teoria dei grafi
- NoSQL: MongoDB, Hbase, Neo4j

REFERENCE TEXT BOOKS

- R. Elmasri, S. Navathe, Fundamental of Database Systems, 7th Ed., Pearson
- M. Golfarelli, S. Rizzi, Data Warehouse Design, Mc Graw Hill
- Balamurugan Balusamy, Nandhini Abirami R, Amir H. Gandomi, Big Data: Concepts, Technology, and Architecture, John Wiley & Sons Inc.; 1st Ed.
- Materiale fornito durante il corso