

## SYLLABUS DEL CORSO

### Large Scale Data Management

2021-2-F1801Q147

---

#### Obiettivi

Lo studente acquisirà competenze relative alle attività di (i) analisi e riconciliazione di sorgenti operazionali eterogenee, (ii) progettazione concettuale relativa alla realizzazione di schemi di fatto, (iii) utilizzo architetture di big data per la gestione dei dati

#### Contenuti sintetici

Il termine data warehouse si riferisce ai metodi, alle tecniche e strumenti in grado di aggregare ed elaborare i dati di una organizzazione per condurre analisi di tipo strategico. Nella prima parte del corso, erogato in lingua italiana, verranno presentate le metodologie di progettazione e sviluppo di un data warehouse. Nella seconda parte del corso si approfondiranno le architetture di big data e la loro relazione con le soluzioni di datawarehouse esistenti. Durante il corso verranno presentati esempi reali da parte di aziende.

#### Programma esteso

1 "Analisi e riconciliazione di sorgenti operazionali eterogenee:

- a. tecniche per la comparazione e l'allineamento di schemi concettuali;
- b. tecniche per integrazione e ristrutturazione di schemi concettuali.

2 Progettazione concettuale relativa alla realizzazione di schemi di fatto:

a. il Dimensional Fact Model come modello concettuale grafico per la definizione di schemi di fatto;

b. metodologia per la definizione di schemi di fatto a partire da schemi E-R.

3 Big data, architetture di ETL e storage per datawarehouse

a. architetture data lake

b. prodotti disponibili per attività di ETL e storage

c. data processing

## **Prerequisiti**

Conoscenze di base sulla progettazione di basi di dati. Conoscenza del modello entità-relazione (E-R)

## **Modalità didattica**

Lezioni frontali ed esercitazioni in lingua italiana

Durante l'emergenza COVID-19 il corso sarà erogato da remoto asincrono con eventi in videoconferenza sincrona, esercizi di autovalutazione non influenzanti il voto potranno essere proposti

## **Materiale didattico**

Matteo Golfarelli e Stefano Rizzi. Data Warehouse – Teoria e pratica della progettazione (seconda edizione). McGraw-Hill.

## **Periodo di erogazione dell'insegnamento**

primo semestre

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

sviluppo di un progetto di datawarehousing o di big data analysis e relativa discussione

## **Orario di ricevimento**

Giovedì dalle 10 alle 11 edificio U14 stanza 2017

---