UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

Data Architecture

2122-1-F1801Q159

Aims

The course is provided in Italian.

First of all the concept of data life cycle and of data base management system is discussed in traditioal environements and information systems, and the so called ACID properties are examined,

The evolution of data base management systems from centralized architectures to distributed and federated architectures is examined in detail, providing also general concepts of data architecture design.

As to data governance, two relevant phases are examined in more depth, data linkage (also called object identification) and fusion, and data quality assessment, that are experimented using tools.

The recent evolution of management systerms from "small data" to big data is investigated, with specific reference to the most important phenomena, related to the increase in the volume of data and the increase in the heterogeneity of data types, from relational data bases to document bases, linked open data and graph data.

Contents

Data life cycle, management and analysis phases

Traditiona Data Base Management systems (DBMS)

Evolution of DBMSs from centralized to distributed, replicated, federated.

Data governance

NoSQL database models and architetture

Big data principles and architectures

Detailed program

Traditional Data Base Management Systems (DBMS)

- Centralized data base management systems (DBMS)
- Distributed DBMSs
- Replicated systems
- Data warehouses
- Virtual data integration systems

Data governance

- Record linkage and fusion
- Data Quality

Evolution of data management systems from small data to big data

- NoSql Architectures and models
- Key value databases
- Document based DBMSs
- Graph Databases
- Linked Open Data
- Big Data

Prerequisites

Knowledge on data models, relational model and Entity Relationship model, and database design methodologies is useful, but non mandatory

Teaching form

Lectures and exercises in Italian, company presentation

Self-assessment tests, that will not be considered for the final evaluation will be provided

Textbook and teaching resource

slides will be published on the Elearning platform

Text book

Next Generation Databases: NoSQL, NewSQL, and Big Data

by Guy Harrison Publisher: Apress

Release Date: January 2016

ISBN 9781484213292

Semester

second semester

Assessment method

Written examination related to the topics of the course and exercises. Teacher could ask and oral exam

Alternative assessment method: Project work, individual or in group related to the technologies shown during the course

Office hours

Thursday from 10 to 11 building U14 room 2017 prof Maurino please send an email message before