UNIVERSITÀ DEGLI STUDI DI NAPOLI FEDERICO II

PHD PROGRAM IN INFORMATION TECHNOLOGY AND ELECTRICAL ENGINEERING

PHD PROGRAM IN INFORMATION AND COMMUNICATION TECHNOLOGY FOR HEALTH

PhD Course announcement

Title: Big Data Architecture and Analytics

Lecturer: Prof. Giancarlo Sperlì

University: University of Naples Federico II
Email: Giancarlo.sperli@unina.it

Credits: 5

Short bio notes

Giancarlo Sperlì is an Associate Professor at the Department of Electrical Engineering and Information Technology of the University of Naples Federico II. He obtained his PhD in Information Technology and Electrical Engineering at the same University, defending his thesis: ``Multimedia Social Networks". He is a member of the Pattern analysis and Intelligent Computation for mUltimedia Systems (PICUS) departmental research groups. His research interests focus on multimedia big data, social network analysis, and design of data-driven approaches for different domains.
Email: giancarlo.sperli@unina.it

Overview

The course aims to investigate Big Data methodologies and architectures for supporting analytics in several application domains from different points of view. Specifically, the course provides an analysis of Big Data Management and Data Analytics Lifecycle, concerning the design of large and complex data systems. Furthermore, the course focuses on the processes of ingestion, modeling, analysis, and visualization of Big Data. We further investigate both batch and streaming processing architectures for supporting different case studies (i.e., social network analysis, health, and Industry 4.0), also discussing their deployment on cloud infrastructures and the integration of recent Artificial Intelligence models. There will be a final assessment.

Schedule

Lecture	Date	Time	Room	Topics	Lecturer
1	7 May 2025	14:30 – 17:30	**Ex SOFTEL**	**Big Data Lifecycle**	Giancarlo Sperlì
2	12 May 2025	10:00 – 13:00	**Ex SOFTEL**	**NoSQL databases: Theory and Use Cases**	Giancarlo Sperlì
3	19 May 2025	10:00 – 13:00	**Ex SOFTEL**	**Distributed Systems**	Giancarlo Sperlì
4	22 May 2025	14:30 – 17:30	**Ex SOFTEL**	**Lambda Architecture**	Giancarlo Sperlì
5	26 May 2025	10:00 – 13:00	**Ex SOFTEL**	**Kappa Architecture**	Giancarlo Sperlì
6	27 May 2025	14:30 – 17:30	**Ex SOFTEL**	**Open Challenges in Big Data Analytics**	Giancarlo Sperlì
7	28 May 2025	10:00 – 12:00	**Ex SOFTEL**	**Open Challenges in Big Data Analytics**	Giancarlo Sperlì
				Assessment test	Giancarlo Sperlì

Content details

Lesson 1 – Big Data Lifecycle

Introduction and fundamentals about Big Data; Design and Definition of Big Data systems (Kappa and Lambda architectural patterns); Data Model for Big Data, Introduction of NoSQL databases; Fundamentals of NoSQL databases; Differences between SQL and NoSQL databases.

Lesson 2 – NoSQL databases: Theory and Use Cases

Analysis of NoSQL databases' families: Key-value systems - Column-family storage systems, Graph storage systems, Document Database systems, Applications and challenges in using NoSQL databases (Document (MongoDB) and graph (Neo4J) based NoSQL databases) to support different analytics.

Lesson 3 – Distributed Systems

Streaming engines; Main Application framework (i.e., Splunk, Apache Kafka); Use cases; Open challenges and issues. Introduction and fundamentals of Big Data Analytics (BDA); Big Data Analytics Lifecycle; Introduction Apache Hadoop distributed File System (HDFS); Architecture of HDFS; Map Reduce; Optimization strategy in Apache Hadoop.

Lesson 4 – Lambda Architecture

Introduction to Distributed Streaming Processing, Fundamentals of Apache Kafka; Introduction to Lambda Architecture, Fundamentals of Apache Spark; Use cases; Issues and challenges.

Lesson 5 – Kappa Architecture

Introduction to Kappa Architecture, Fundamentals of Apache Storm; Use cases; Issues and challenges.

Lesson 6 – Open Challenges in Big Data Analytics

Challenges of the Big data methodologies and architecture in different domains.

Lesson 7 – Open Challenges in Big Data Analytics

Challenges of the Big data methodologies and architecture in different domains.

Participants are requested to send an e-mail to giancarlo.sperli@unina.it by May 5, with the following information:

Student name and surname, name of the PhD course, PhD cycle.

For information: Prof. Giancarlo Sperlì (DIETI, UniNA) – giancarlo.sperli@unina.it *(organizer)*