AD HOC TEACHING MODULE Announcement

UNIVERSITÀ DEGLI STUDI DI NAPOLI FEDERICO II

PHD IN INFORMATION AND COMMUNICATION TECHNOLOGY FOR HEALTH

PHD IN INFORMATION TECHNOLOGY AND ELECTRICAL ENGINEERING

PHD IN COMPUTATIONAL AND QUANTITATIVE BIOLOGY

Module Title: Big Data Architecture and Analytics

Lecturer: Prof. Giancarlo Sperlì

University of Naples Federico II

mail: giancarlo.sperli@unina.it

Telephone: +39 081 768 8349

CV: is an assistant 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 main research interests are in the area of Cybersecurity, Multimedia Data and Social Networks Analysis. He served as guest editor of different special issues on International Journals. Finally, he has authored about 104 publications in international journals, conference proceedings.

Dates and Locations

Type	Date	Hours	Lecturer	Room
Lesson 1	**26 June 2023**	**15:00 – 17:00**	**Giancarlo Sperlì**	**CL-I5**
Lesson 2	**29 June 2023**	**15:00 – 17:00**	**Giancarlo Sperlì**	**CL-I5**
Lesson 3	**6 July 2023**	**15:00 – 17:00**	**Giancarlo Sperlì**	**CL-I5**
Lesson 4	**7 July 2023**	**15:00 – 17:00**	**Giancarlo Sperlì**	**CL-I4**
Lesson 5	**10 July 2023**	**15:00 – 17:00**	**Giancarlo Sperlì**	**CL-I5**
Lesson 6	**12 July 2023**	**15:00 – 17:00**	**Giancarlo Sperlì**	**CL-I5**
Lesson 7	**14 July 2023**	**15:00 – 17:00**	**Giancarlo Sperlì**	**CL-I6**
Lesson 8	**19 July 2023**	**10:00 – 13:00**	**Giancarlo Sperlì**	**CL-II2**
Lesson 9	**20 July 2023**	**10:00 – 13:00**	**Giancarlo Sperlì**	**CL-I1**
Final Assessment	**TBD**	**TBD**	**Giancarlo Sperlì**	**TBD**

AD HOC TEACHING MODULE Announcement

Content

The aim of the course is to investigate Big Data methodologies and architectures for supporting analytics in different application domain from different point of views. In particular, the course provides an analysis about Big Data Management and Data Analytics Lifecycle, with reference to the design of large and complex data systems. Furthermore, the course focuses on the processes of ingestion, modelling, analysis and visualization about Big Data. Possible applications of these methodologies and architectures for different case studies (i.e., health, industry 4.0, social media and 5G networks) will be discussed at the end of this course. There will be a final assessment.

Lesson 1

Introduction and fundamentals about Big Data; Design and Definition of Big Data systems (Kappa and Lambda architectural patterns); Data Model for Big Data.

Lesson 2

Introduction of NoSQL databases; Fundamentals of NoSQL databases; Differences between SQL and NoSQL databases; Analysis of NoSQL databases' families: Key-value systems - Column-family storage systems, Graph storage systems, Document Database systems.

Lesson 3

Applications and challenges in using NoSQL databases (Document (MongoDB) and graph (Neo4J) based NoSQL databases) to support different analytics.

Lesson 4

Streaming engines; Main Application framework (i.e., Splunk, Apache Kafka); Use cases; Open challenges and issues.

Lesson 5

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 6

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

Lesson 7

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

Lesson 8

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

Lesson 9

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

AD HOC TEACHING MODULE Announcement

ECTS Credits 5

Notes

The course is in presence.

Remote participation (on Microsoft Teams) is allowed only to students who are spending their research period abroad.

https://teams.microsoft.com/l/team/19%3aFeeXK0sjO2oCrP6BmtpSVMWW5-plI4Q2gPXL2fplVyc1%40thread.tacv2/conversations?groupId=a5101e73-8275-46a0-8ccb-4a30b9e81fa5&tenantId=2fcfe26a-bb62-46b0-b1e3-28f9da0c45fd

(Team Code: axzi6e7)

Participants are requested to send an e-mail to Prof. Giancarlo Sperlì by 27 June, 2023, with the following information:

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

In the email, students abroad need to motivate the request for remote attendance, indicating the place and period they are spending in a foreign institution. For all other students participation is in presence.

Info: **Prof. Giancarlo Sperlì -** tel. (+39) 081-7683849 – mail:giancarlo.sperli@unina.it