



UNIVERSITÀ  
DEGLI STUDI DI MILANO-BICOCCA

## SYLLABUS DEL CORSO

### Technological Infrastructures For Data Science

2021-2-F9101Q013

---

#### Aims

The course aims at providing a solid understanding of the technological platforms (sensors and networks) which allow the collection of data in an IoT environment, as well as of the computing platforms (architectures, algorithms and infrastructures) which can be used to analyse those data.

The exercises will provide the student with the basic capabilities necessary to interact with such platforms.

#### Contents

Sources of sensor data, quality of sensor data, data collection networks, sensor management platforms. Data processing architectures, Infrastructure Management, Big Data Platforms for data Science, platform examples.

#### Detailed program

- Introduction to the IoT paradigm (2)
- Sources of sensor data (4)
  - o Environment sensors
  - o Home and appliance-based sensors
  - o Wearables

- o Sensor embedded in mobiles
  - Quality of sensor data (Precision, Timeliness and Availability) (2)
  - Data collection networks (8)
- o Home networks and broadband access
- o Wireless networks: LPWAN and Industry-specific solutions
- o Cellular networks: the eSIM concept, LTE-M, NB-IoT and 5G
  - Data collection and sensor management platforms (8)
  - Data processing architectures (4)
  - Infrastructure Management (6)
- o Virtualization vs containerization
- o Cloud environments, models and pricing
  - Big Data Platforms for data Science
- o Workload types
- o MapReduce - the Hadoop ecosystem
- o Event processing and Stream processing
- o Lambda and Kappa architectures
  - Platform examples (8)
- o Apache Spark
- o Apache Kafka
- o Google Cloud Platform

## **Prerequisites**

No prerequisites

## **Teaching form**

Classroom lectures, classroom exercises

The course will be held in english

During the Covid-19 emergency period, the teaching form is changed. Lectures and exercises will be mostly delivered remotely with video recordings and with live videoconferencing events.

## **Textbook and teaching resource**

Lecture notes and slides provided by the lecturers

## **Semester**

second year, first semester

## **Assessment method**

The exam will be in written form and made up of two parts, relating respectively to the processing platforms and to the technological platforms (see the course description in the previous points). The two parts can be carried out independently even in different sessions.

Each part will consist of a set of closed and open questions to be carried out in an hour (approximately 6 multiple choice questions and one open question for the part relating to the processing platforms and three open questions for the part relating to the technological platforms, however small variations in the structure of the exam will be possible). The exam will be held in paper form and the weight of each question for the formation of the vote will be explicitly indicated in the upper part of the sheet containing the questions. The answers to the questions will be written on the sheet itself, but it is up to the student to use additional sheets to provide a more extensive answer to the open questions.

Once the student has done both tests, the exam will be considered passed if both these conditions are met:

- for both parts the student will have obtained more than half of the 15 points available
- the sum of the points of the two parts will be greater than or equal to 18

in this case the student can register a grade consisting of the sum of the points.

In addition to the standard sessions, there will be an ongoing test relative to the part of the program carried out in the first half of the course, which will correspond to one of the two parts of the exam. Consequently, a student who has completed (with a positive result, more than half of the points available) the on-going test may perform only the missing part of the exam during the sessions.

## **Office hours**

tuesday 12:30-14:30 ask for email confirmation

---