

# project

## Early Description and Objectives:

The goal of Processing Big Data (PBD) project is to confront students with a typical “data science” problem and apply state-of-the-art “data science” methodologies as well as standard techniques learned in the course. Foundational courses such as linear algebra, programming, calculus or optimization are key so we encourage you to revise those materials.

We will do this training through a specific application of video analytics for cycling races that we will describe below.

The project is structured such that students are required to practice all steps of the data processing “pipeline” : Exploratory Data Analysis, Data Representation, Visualization, Modeling, Algorithm Design and Performance analysis. Furthermore, the project will show the reality of “big data processing”: lack of a clear problem structure, heterogeneity of the data, huge dimension and unreliable data (outliers).

In short, the project does not have a single clear solution or approach. Students are encouraged to explore and autonomously test multiple approaches/solutions though the context is a bit constrained to make it realizable. The application scenario is extracted from the Italian bicycle competition.

## The Giro D'Italia: Analysis of RAI's Media Coverage

Large outdoor events like the bicycle race “Giro D'Italia” follow complex but very well planned “scripts” by the producer/director of the coverage.



Fig1: skeleton extracted by openpose

In particular the director is managing the viewpoints originated from a multitude of video sources such as motorbikes, fixed cameras, drones or helicopters and at each

time instant, he/she selects the best video feed or the most meaningful to the viewer.

## Objectives and data

The original data is comprised of a set of videos taped from RAI's media coverage. These videos were processed and the available data is the following:

- **Visual embeddings:** A high dimensional vector (1024) for each image of the video. This is a representation provided by a Neuronal Network that encodes each image. This is to say that images can be compared in terms of similarity using these vectors. Details will be provided in the practical classes.
- **Skeletons :** Each image was processed and skeletons of visible people were extracted using the "open pose" algorithm. As fig 1 shows, depending on the number of visible persons and the algorithm reliability, a variable number of skeletons are extracted. A skeleton is a set of 25 2D coordinates encoding the image location of each joint. Details will be provided and the original representation is here <https://github.com/CMU-Perceptual-Computing-Lab/openpose>.
- **The original images** in case you want to extract further information from them.

In short, the minimum set of data is represented by a set of matrices as described below.

<https://youtu.be/DEZ8yw78Y94>

Keypoints of detected poses.



There will be three levels of difficulty which involve three different types of data.

- 1 - Complete data: the skeleton data is complete, that is, the whole set of skeleton points is available for each detected person.
- 2 - Incomplete data: skeleton detectors fail and sometimes only part of a body is visible therefore it happens that some points are missing. In fact, the previous dataset (completed) was created from this one using one particular data completion algorithm !
- 3 - Extra data: any data extracted from the images or crawled from the internet or any other source. This is left to each group initiative, students are allowed (and encouraged) to extract further data.