Labwork on **image classification**

Nicolas Sutton-Charani

Solutions will be given on this online notebook.

Chest CT-Scan images Dataset

(a) Download the data here.

(b) Install and load the '**OpenImageR**' package.

(c) Load and plot any image with the 'readImage' and 'imageShow' functions.

(d) Compute the dimensions of the image, convert it to a black and white image with the 'rgb_2gray' function and recompute its dimensions.

(e) Compute the ***HOG*** descriptors of the image with the 'HOG' function.

(f) Write an R script that scraps and loads all the downloaded images, computes their HOG descriptors and their labels (adenocarcinoma, large.cell.carcinoma, squamous.cell.carcinom or normal) and saves them into a descriptors matrix and a labels vector.

(g) In the *scraping* part of script, add a test that avoids duplicates.

(h) Evaluate the predictive power of the processed dataset with tree and forest (and other) classifiers.

(i) Try to increase the accuracy of your best models.

https://www.kaggle.com/mohamedhanyyy/chest-ctscan-images