# Logo Detection Project

##### **Alberto Allegri, Aliya Davletshina, Niccolò Diana, Beatrice Guidotti, Kasra Zamanian**

UniversitàBocconi

Data Science & Business Analytics

# Preprocessing

For the purpose of preprocessing, we used both Roboflow and manual preprocessing (you can find all the scripts on the utils folder). However, at some point, we found Roboflow more comfortable to use so we completely shifted to using it for the rest of the journey.

Due to limitations of using free trial on Roboflow, we used different accounts, each time feeding 10k images and applying a different set of augmentations. We chose Yolov5 format for the output.

# The training

We chose three models for our use case:

* Yolov5
* ResNet
* EfficientDet

However, after configuring all the necessary steps, we understood that using ResNet and EfficientDet will require a large amount of time (more than 5 hours per epoch), while even a small Yolov5 was already doing quite good on the data (only 15 minutes per epoch). Thus, we decided not to proceed with the other two models and put our focus on improving the results of Yolov5.

We used five different models in terms of size, datasets, and settings.

1. Yolov5s on all the raw data with freezing 10 layers (only the backbone)
2. Yolov5s on the clean data with freezing 10 layers
3. Yolov5s fine-tuned with extra augmentations on the second model with freezing 18 layers
4. Yolov5s on all the clean data plus extra augmentations with freezing 10 layers
5. Yolov5l on all the clean data plus extra augmentations with freezing 10 layers

Firstly we trained Yolov5s on all the mandatory logos plus some of the optional logos, and we achieved not fair results (first model). After further investigation, we found out the fault is not on the model but the dataset as it included many incorrect annotations and multiple logos per image while there were not annotated. Hence, we manually discarded all the images with wrong annotations or with multiple logos. Also, we excluded **Ralph Lauren** and **Intimissimi** logos completely.

Next step, in order to compensate for multiple detections, we applied Mosaic augmentation (combined with other types of augmentation) to the clean data and reran Yolov5s (second model). The result was great with high precision, recall, and average IOU. But still, the model was suffering from detecting small logos, blur images, and rotated logos. Therefore, we decided to create extra augmentation and fine-tune on the previous model unfreezing the outermost layers (third model) for another 50 epochs. The result got worse since the model was detecting too many false positives, mistaking random objects as logos, which was a result of learning too shallow,

So we decided to run two new models from the beginning again (freezing only the backbone) by combining all the previous clean data with the new augmentations, one on small Yolo (fourth model) and one on a big Yolo (fifth model). The small Yolo took 2 days (20 minutes per epoch), and the big Yolo took 10 days (3.5 hours per epoch) to train. The result was obviously improved on the big Yolo since it was predicting fewer false positives and even managed to predict some images that none of the other models could do.