In this task, you are going to build an inference engine that is capable of predicting whether a person is wearing a **face mask** and a **safety helmet**.

1. You will be using the dataset provided by us. The dataset folder consists of an image folder and a xml annotation file. The annotations have human heads and several other objects. You only need human head objects for this task. Please refer to the meta element of the xml file for more details.

   The ‘id’ attribute of the image element refers to the name of the image file. xtl, ytl, xbr, ybr represents the location of the bounding box of the object. x, y refers to the coordinates and tl, br are abbreviations of top left and bottom right respectively.

2. The engine does not have to be one model

3. You are free to use extra training data.

4. You need to write your code in python. You are free to use any framework to train your model/models(pytorch, tensorflow, mxnet, keras, chainer etc)

5. Propose an evaluation metric and justify your choice. It does not have to be mAP.

6. Include the dependencies in read.me

The goal of the test is to see whether the candidate is familiar with the deep learning workflow. Since only limited data is provided, we don’t expect an accurate model