Task Description

The aim of this Task is to create a good validation framework for analysing people who appeared in News Video Segments. The Task should collect all the instances where a person appears, and this data should include the timestamps, name, and flags marking whether that person is a presenter or interviewee.

An important keyword before the task description, an 'Instance' refers to the part of a person when a person appears.

Structure of task

1. Open a video from the provided folder.
2. Copy the template csv and rename the csv to match the mp4 name.
3. Watch the video and list down all instances where a person appears.
4. These steps should be repeated for all videos provided.

Video and CSV structure

The videos will be provided in a folder where each video should have a respective CSV file with the filled data. The names of the CSV files should match the names of the videos.

All created CSVs should be placed in the folder named 'Validation'.

Which instances/people to note down

Instances that should be noted are people that have their names appear on screen or is speaking to the audience (the people who are watching the video).

Below is a list of examples that should be noted down.

Figure 1 Person speaking with name (ADRIAN AXIAK) shown.

Figure 2 Person speaking without name shown.

Figure 3 Presenter outside studio

Figure 4 Presenter inside studio

CSV format

Each CSV file should contain the following columns:

1. The starting time of the instance
2. The ending time of the instance
3. The name of the person (leave empty is not shown)
4. A Boolean (write either 'TRUE' or 'FALSE') noting is name was shown.
5. A Boolean (write either 'TRUE' or 'FALSE') noting whether the person is a presenter.

Time start	Time end	Name	Is name shown?	Is presenter?

Figure 5 Template CSV

It is important to note that the CSV files should be comma separated values as shown below.

```
Time start,Time end,Name,Is name shown?,Is presenter?
```

Figure 6 Template CSV showing comma separated values

Starting and ending time of the instance

The start and ending time of the instance should be noted according to when the person appears up till the last consecutive appearance. This means that a person can have multiple appearances which should include different timestamps.

Name of the person

When it comes to identifying the name it is important to do so, since this is required for my validation. For better clarification I have also included some samples of names in the 'Samples/Images' folder.

It is **important** to update a person's name in the case where initially the person's name did not appear but appeared later. All instances where a person is shown, it is important to note that each instance of the same person should have the same name.

Name shown Boolean

In this column, only a 'TRUE' or 'FALSE' is required. TRUE meaning that the person's name was shown in this instance, and FALSE meaning that the person's name was not shown in this instance, but in anther instance, or not shown at all throughout the video.

Presenter Boolean

Like the previous section, in this column, only a 'TRUE' or 'FALSE' is required. TRUE meaning that a presenter person was shown in this instance, and FALSE meaning that an interviewee was shown.

It is important to distinguish between presenter and interviewee, and this can be shown in the examples of Figure 3 and 4.

Samples

For better clarification for this Task, I have provided a 'Samples' folder which includes 3 video examples for how this Task should be carried out. It is highly suggested to go through the examples so to limit errors and miscommunications.

In the same folder, I also provided an 'Images' folder which contains some sample examples of how the names of people appears during the video.