Good Afternoon, my name is Francesco Ferraioli. I undertook my final year project with a research group at QUT called SAIVT (Speech Audio Image Video Technologies). As the name suggest, the specialise in speech, audio, image and video related work.

SAIVT has implemented a Django web application which they call the Entity Extraction Webdemo. They use this application to showcase a lot of their work. The main idea of this application is to extract entities from videos and display that information in the webdemo. Entities can be many things, but the webdemo focuses primarily on faces and speakers.

As shown here (pointing to the first box), this is a frame of a video. We can see in the frame, a face. Furthermore, there is also audio that comes with this frame in the video. The job of the webdemo is to capture these entities and identify them as well as try and find linkage between entities in other videos.

The Django application however, needed a lot of work, not only on the front end look and feel, but also on the structure of the application as a whole. Furthermore, a few enhancements and features were wanted for the webdemo. My task was to work on this.

My first task was to implement general enhancements, particularly the look and feel of the page, making it look more modern and alive. A major part of this task was changing the implementation of the tags.

The second task was that of enhancing the entity display and the video page. Previously the faces were being tracked by a square that was encoded in the video. My task was to turn the square into an interactive HTML element. The original video would be shown and a html box would be displayed on top of it, encompassing all the movements of the face. The element is interactive in that hovering over the face would show more information about the entity.

Furthermore, showing the current speaker was a desired functionality for the webdemo. Previously, to see who the current speaker was you would have to look at the right hand side on the speakers tab but now the current speaker is shown underneath the video again as an interactive HTML element.

The final and largest task for the project was implementing functionality to allow users to upload their own videos to the webdemo. The first task was creating a web form in which the user could fill out and choose the video they wish to upload. Validation is in place to ensure only valid data is submitted to the server, this includes file format type.

Once the form has been successfully submitted, the uploading of the video can take sometime, depending on the size of the file, and thus a progress bar has been implemented to track the progress of the upload through the ajax call to the back end.

As soon as it has been uploaded, the video becomes part of the list of the videos, even it if is yet to be processed. A process flag is set to distinguish between processed and unprocessed videos.

I have then implemented a cron job which runs every 30 minutes and its task is to send the uploaded videos to an HPC to undertake the processing and then get back the processed video along with all its information about the entities that were found.

The database is then reloaded with the new information and the video is now flagged as process and it contains all the information found from the entity extraction process.