Name: Thomas Kolb
Number: s1027332

Research Question:
How do we approach training the wav2vec 2.0 model for automatic end-to-end speech recognition using subtitles?

We still need to figure out how to make this research question more specific and original. The subtitles I refer to in the research question is still not completely clear, we have talked about formats such as teletekst but nothing is final at this point.

Supervisor:
David van Leeuwen

Area:
Automatic speech recognition using deep neural networks, in particular the wav2vec 2.0 model.

Motivation:
Automatic speech recognition is still in a very early stage. Researchers are still struggling with lots of problems with designing and training the models that can be used for ASR. This research could potentially lead to new insights in speech recognition with subtitles using a specific format.

Answer:
The results of this research will be in the form of a trained model that can translate speech to subtitles end to end. Additionally we will also write some scripts (in python probably) that will prepare the available data to something that we can actually use for training.

Strategy:
We will start this research by getting familiar with the model wav2vec 2.0.
We will collect the appropriate data for our research.
We will write scripts to prepare the data for training, depending on the data
We will test the data on the model, and probably keep refining the data/scripts until we get good results
We will compare the results with other automatic speech recognition models or other research