Project idea:
Online Collaborative annotation for communication-based multimodal data

Proposed by Slim Ouni (LORIA/CNRS)

When dealing with multimodal data, the annotation of this temporal data is very often needed. An example of such data is acoustic speech, audiovisual speech (video), gestures; and an example of such a scenario is the annotation of dialog (speaker turn, speech transcription, gestures description, etc.), annotation of sign language data, etc.

In several cases, the first annotations need to be performed by a human before using the data by an AI-based system for training. This data is valuable not only for training but also for testing and validation. The human annotation is time consuming and very lengthy.

We propose developing an online tool to do the annotation of multimodal data online. The online application will run in a web browser and can be used by several human annotators simultaneously working on the corpus and portion of the corpus. This should accelerate the annotation process. This web application should be available to all researchers interested in annotating multimodal data. The web application interface display files to annotate and the progress of the overall annotation. An identified user (who should have an account provided by the researcher who is responsible of the corpus) will have access to a set of files for annotation. The files can be audio, video and eventually any temporal 2D/3D data (to be defined according to the needs of researchers). It is possible to add several annotation streams (for example, words, actions, gestures, speaker turn, etc.)

The researcher can distribute the annotation task of the corpus to several people. It is also possible to have more than one annotator working on the same portion of the corpus, but in this case, we need to evaluate and measure the inter-evaluator agreement.

The web application should be visual and interactive, where the user can play, replay, forward, rewind, etc.