Few Shot Language Agnoistic Key Word Spotting in Audio

1. The PS is identification of key word of interest from arbitrary duration audio file. The solution must classify (i.e what word is) and localize (i.e start and end time stamps) in given audio file.

2. To develop the solution the participants will be given audio files in the folder structure shown below:
   -Train
     -Language_1
       -key_word_1
           - file1.opus (48 k version)
           - file1.wav (16k version)
           …
       …

   Similarly tests set has also been generated.The tests set kws classes are the same as the test set. There may be intra kws variations in test set as is in the train set. The exactly same kws are not used in the test set.

3. The Data folder consists of 10 folders where the languages are Assamese,Arabic,Chezch, Divehi,Persian,French,Russian,Turkish,Tamil and Chinese

4. For evaluating the solutions we have prepared continuous audio files of arbitrary duration , wherein the keywords from test set (not shared with the participants) have been inserted inside audio files.The audio files used for generating test files are in English . We have inserted the test set key words in those audio files.The audio files will be in "wav" format having 16 k sample rate.The participants needs to provide their output in the following format:

   [{"test_audio_file_name":[{"keyword":<KW_NAME>,
                           "start_time" : <xxxx>,
                            "End_time" : <yyyy>},
                           ... # for each key word found in a single file
                           ],
                            … # for every file
                           }]

5. The Evaluation for the submission will be based on Precsion,Recall, F1 and model size. A time collar of 25 mili secs will be taken into consideration while evaluations.