**-IDEAS FOR STEPS2**

**INSTALLATION**

1. Create a Github account and generate a github token
2. Install Matlab following instructions [here](https://www.mathworks.com/help/install/install-products.html). Log in with your Mathworks account.
   1. Install the following toolboxes during initial setup. To add toolboxes later, just click “Add-ons” at the top of the screen.
      1. Audio toolbox
      2. DSP System toolbox
      3. Signal Processing toolbox
      4. Statistics and Machine Learning toolbox
      5. Text analytics toolbox
3. Install [anaconda](https://www.anaconda.com/products/distribution).
   1. In mac, if the installation fails, open a terminal window and type the following:

source /opt/anaconda3/bin/activate (replace with path to anaconda)

conda init zsh

* 1. In Windows, make sure you select the “Add to Path” option

1. In a terminal window, enter the following commands to install the required Python packages

conda create -n ideas python=3.8

conda activate ideas

pip install pyannote.audio

conda install -c conda-forge ffmpeg

pip install -U openai-whisper

1. Clone IDEAS\_STEPS from the repository by typing in the terminal window:

git clone <https://github.com/hugonvilla/IDEAS_STEPS.git>

if prompted, log in with your username and token

1. To download IDEAS\_STEPS updates, type:

git stash (only if you made local changes to IDEAS)

git pull

**USAGE**

1. Open a terminal window and activate the IDEAS environment by typing

conda activate ideas

1. In the same terminal window, start matlab
   1. In Windows: type matlab
   2. In Mac: type open matlabroot

Replace “matlabroot” with the path to the matlab executable (.app). This should be present in the “Applications” folder. You can copy the path by selecting the file and typing the keys option+command+c, and later paste it with command+v.

* 1. In Linux: type matlab

1. Create a .txt file with the paths to the media you want to process. The folder “data” inside IDEAS has an example.
2. Retrieve the path the .txt file and save as a Matlab variable.
3. Run IDEAS by typing in the Matlab prompt

run\_IDEAS(X)

where X is the variable containing the path

1. Retrieve the output in the output folder
2. If using the option “oracle” in parameters.m, the ELAN output (in .txt format) should be located in the “data” folder and named as the video/audio name. For example, for the “xyz.mp4” video, the transcript should be saved as “xyz.txt”