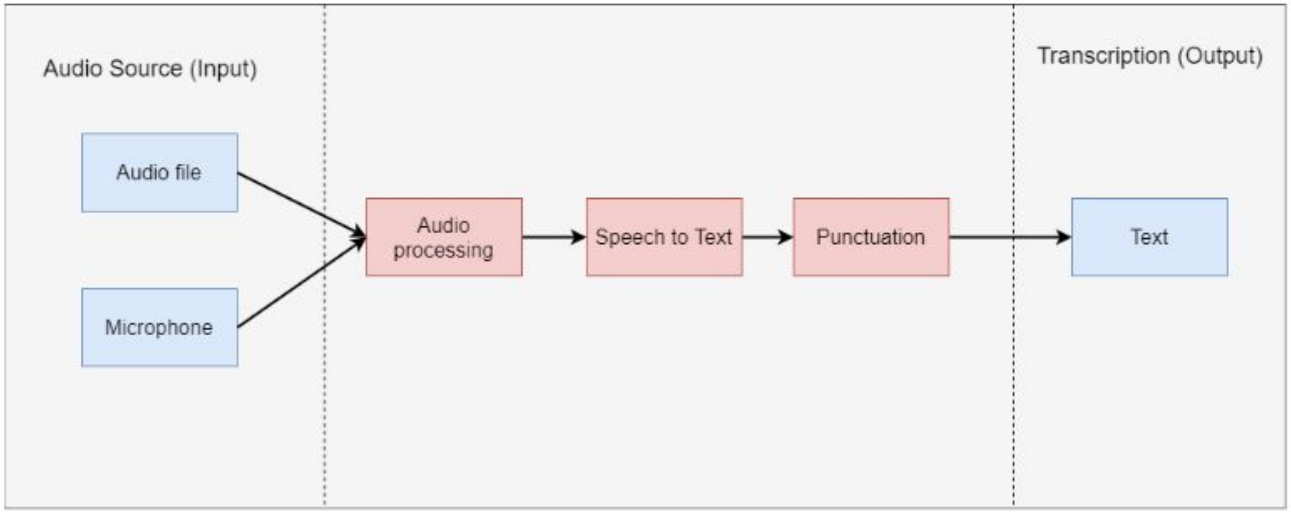
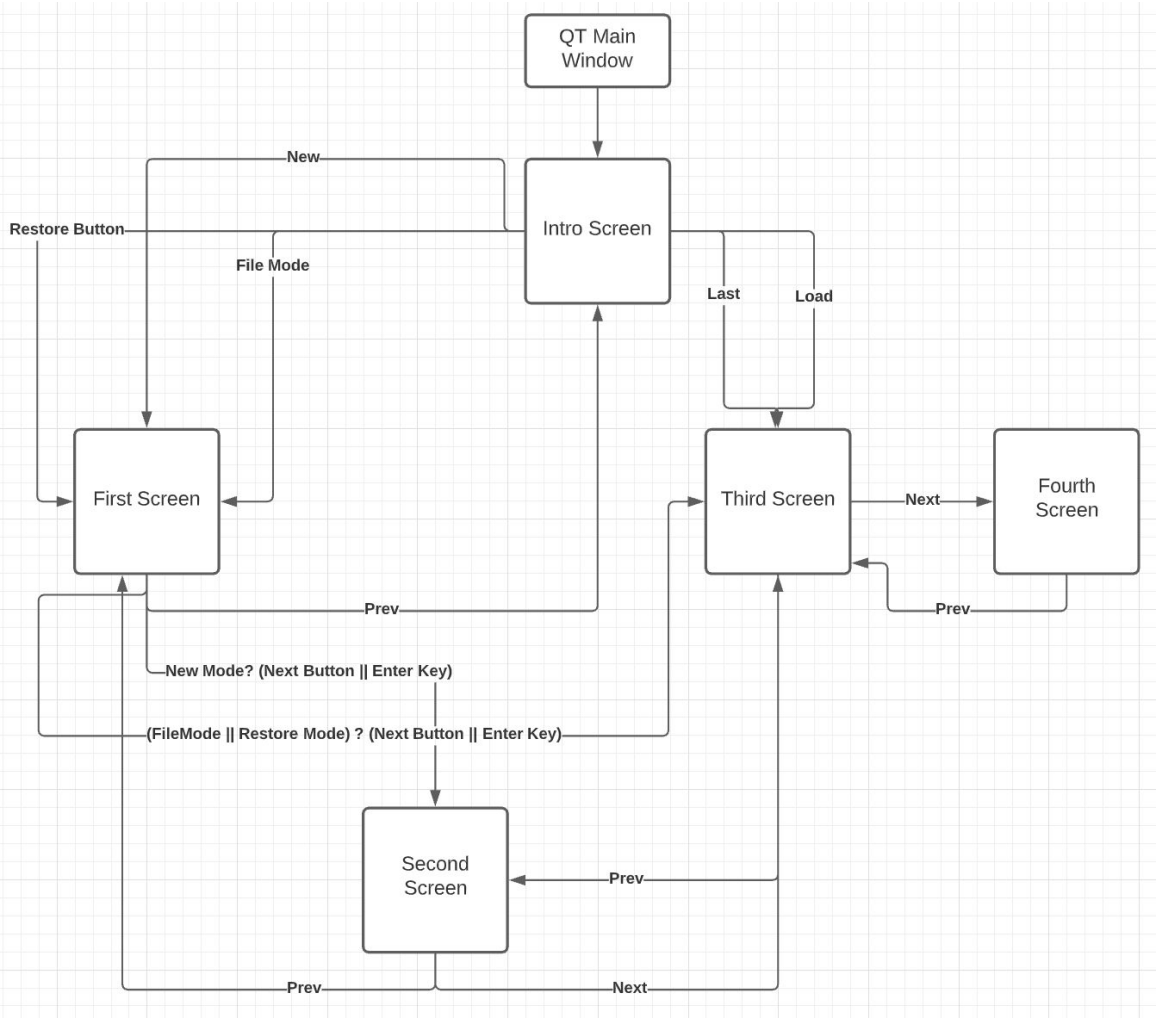
Async ASR

Anand Theertha Nakhate
25 August, 2021

System Architecture

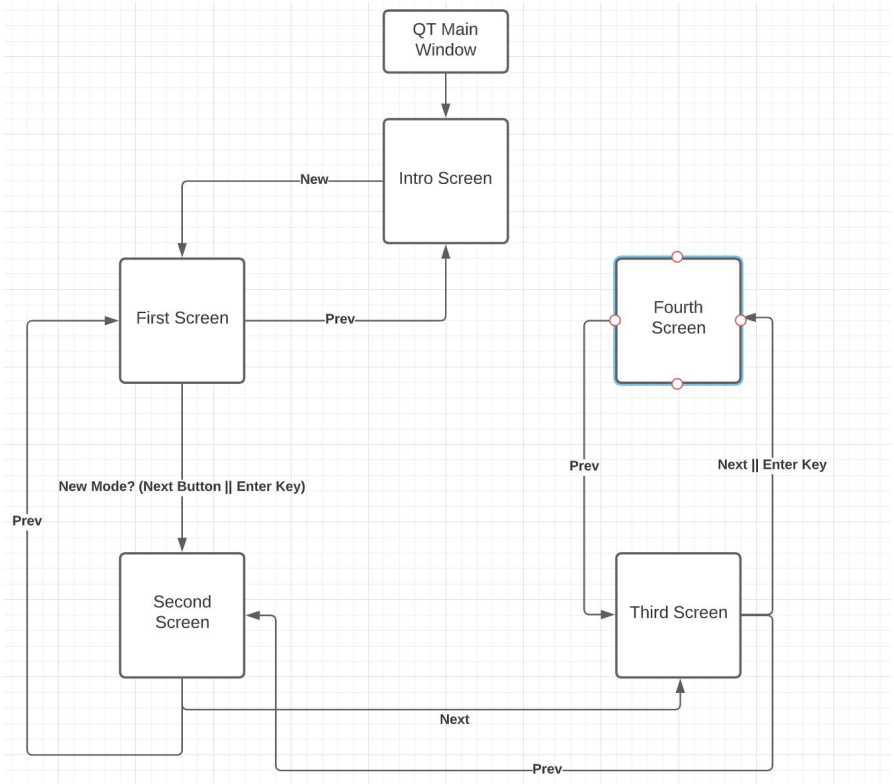


Design Overview


QT Main Window
Intro Screen
New
Restore Button
File Mode
Last
Load
First Screen
Third Screen
Next
Fourth Screen
Prev
Prev
New Mode? (Next Button || Enter Key)
(FileMode || Restore Mode) ? (Next Button || Enter Key)
Second Screen
Prev
Prev
Next

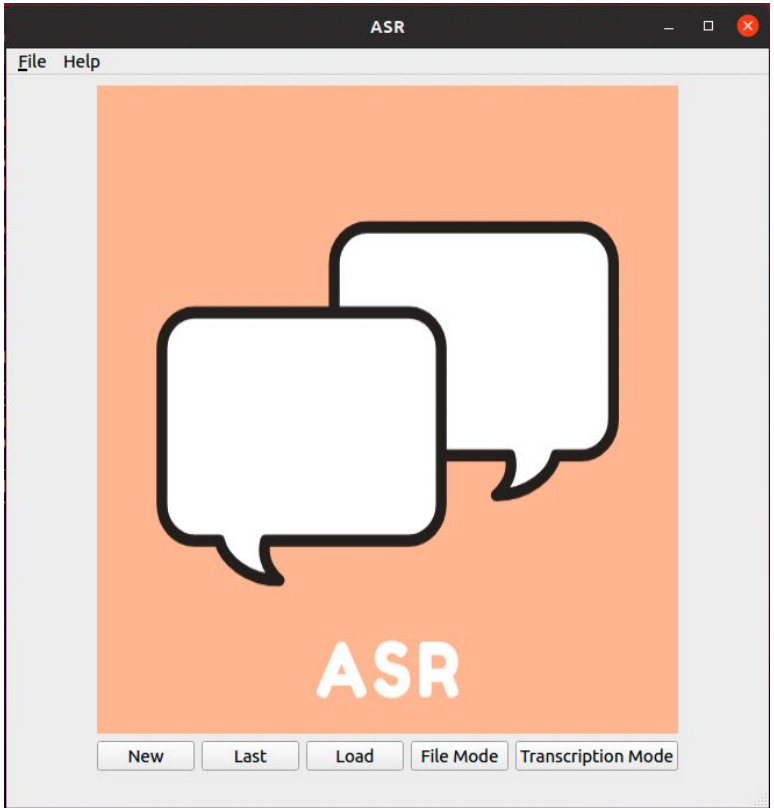
New Mode

Design


QT Main Window
Intro Screen
New
First Screen
Prev
Fourth Screen
Prev
Next || Enter Key
New Mode? (Next Button || Enter Key)
Prev
Second Screen
Third Screen
Next
Prev

New Mode

- Main Window is Declared
- Intro Screen UI is setup
- Intro Screen Attributes:
  - ASR Logo
  - **New Button**
  - Last Button
  - Load Button
  - File Mode Button
  - Transcription Mode Button



First Screen

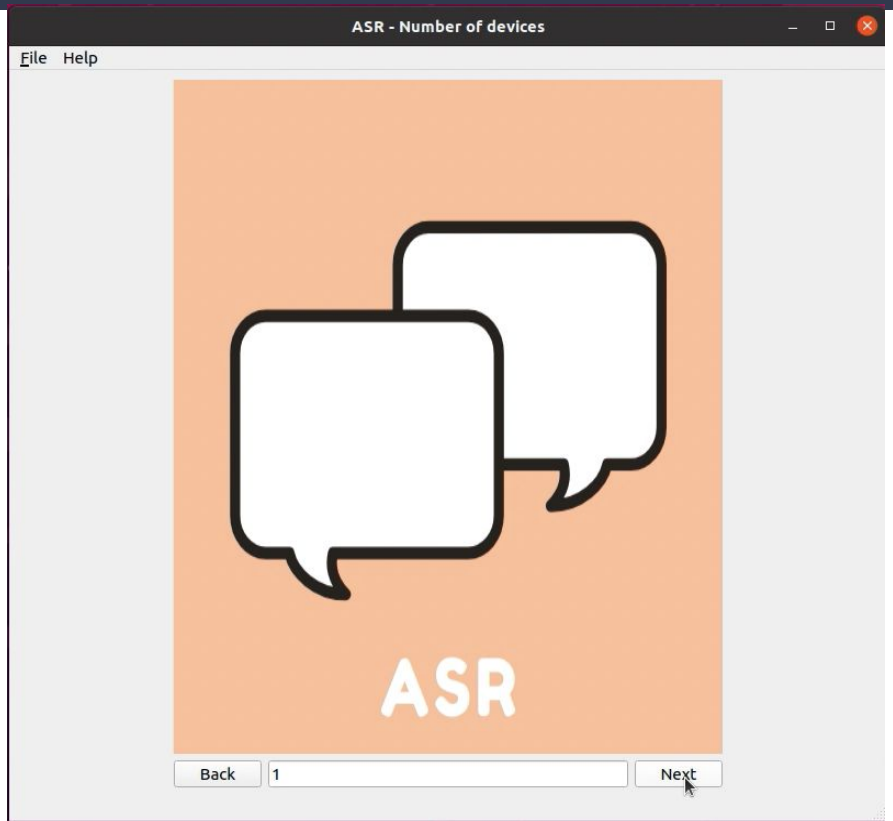
- > Selecting the new Button takes us to the the first screen

- First Screen UI is set up

- Back Button takes us to the previous screen i.e. Intro Screen

- Enter the number of devices (max of 8)

- Next Button handles the number of devices and sets up the second screen by passing the number of speakers.

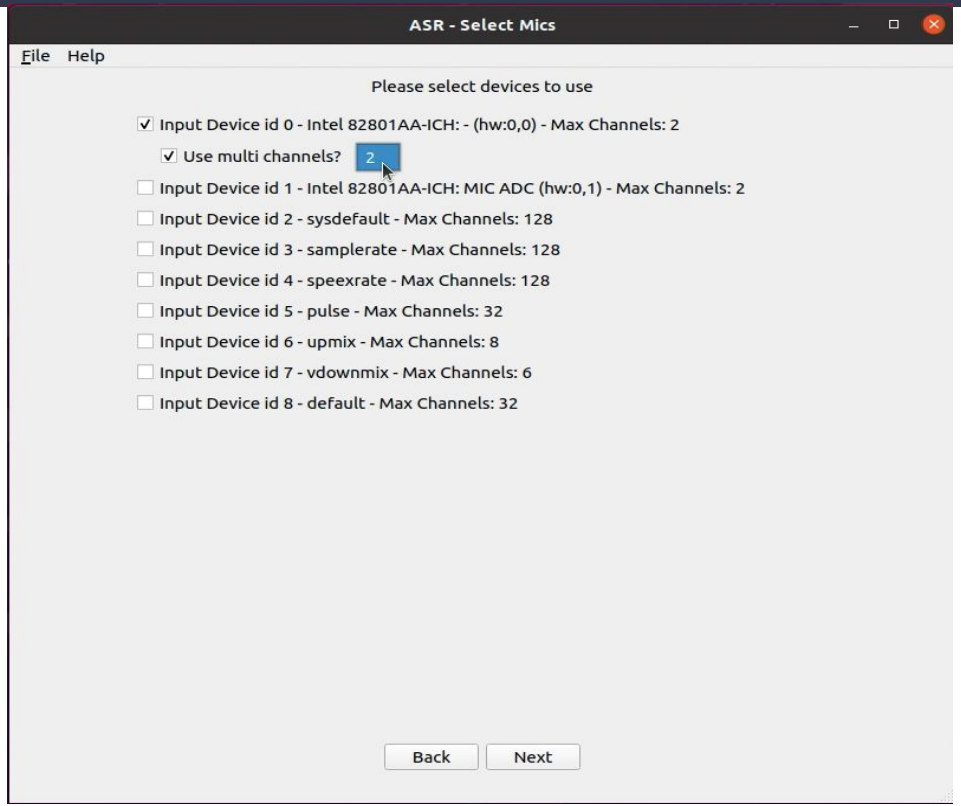


Second Screen

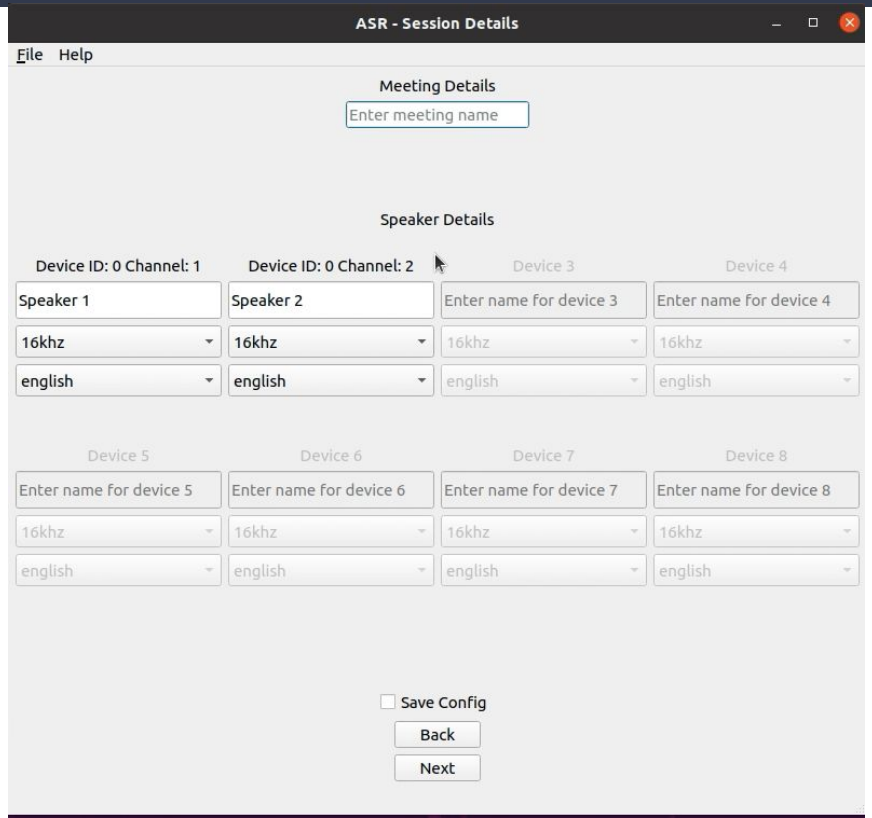
- > Clicking the new Button takes us to the second screen

- We add checkboxes for each of the device that we get from the function deviceInfo().
When a checkbox is selected, a new check box is displayed if we want to use multi channels.
if the checked, a combobox is displayed with the different channels that the device supports.
Select the desired number of channels.

- Next button sets up the UI for the third screen by passing the devices information to the third screen..



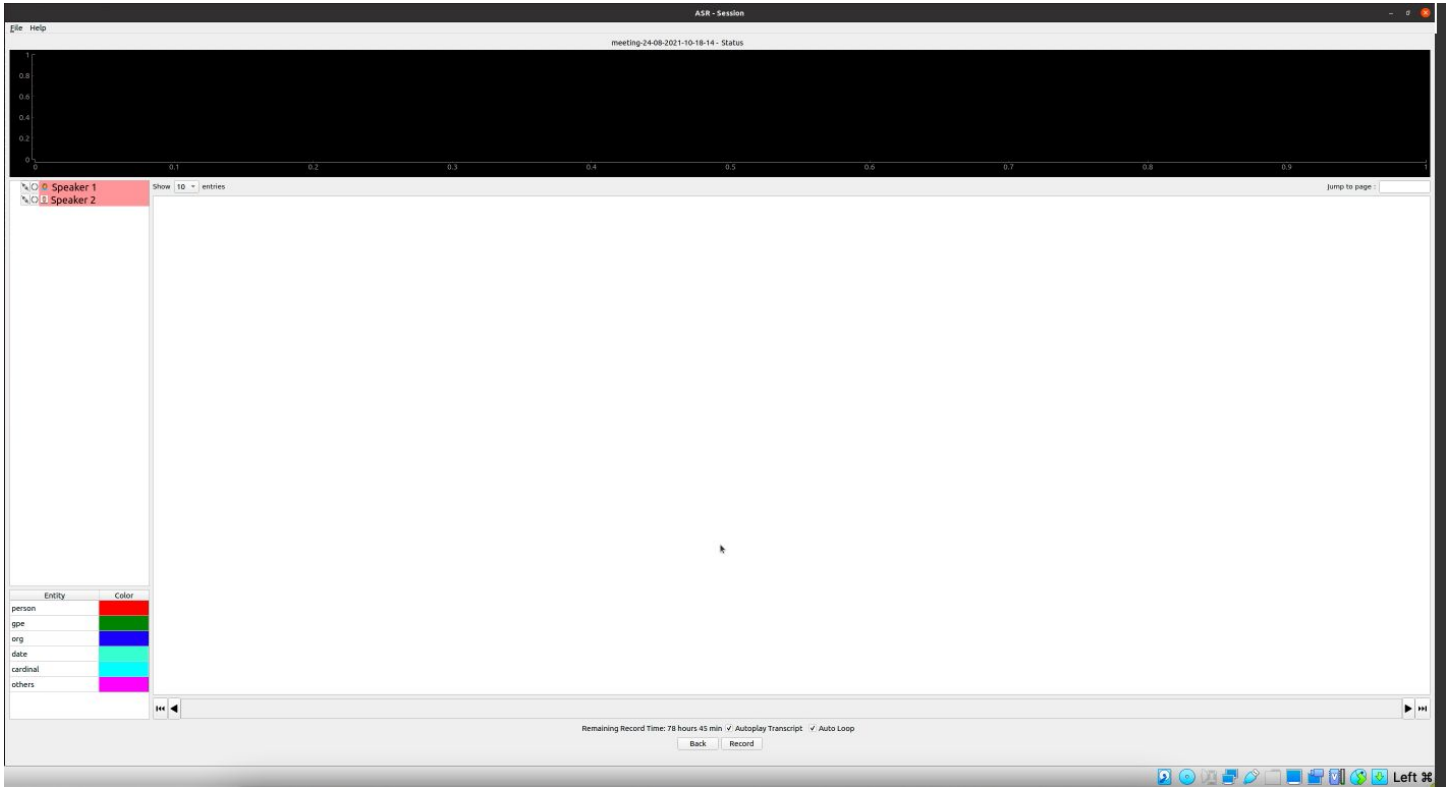
Third Screen



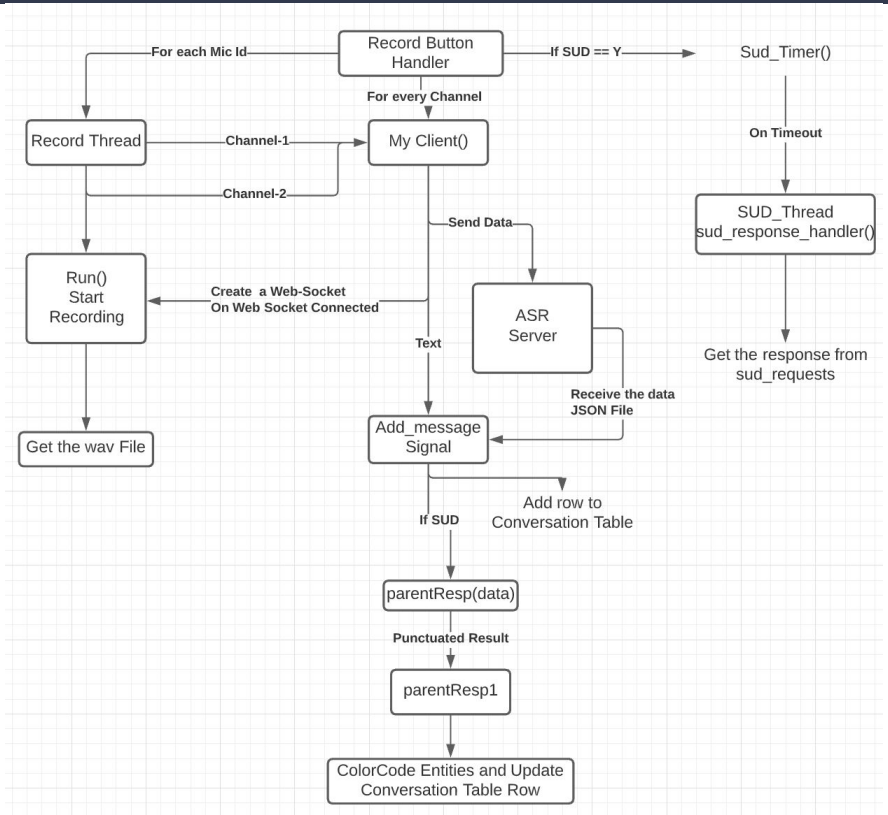
The third Screen UI is loaded

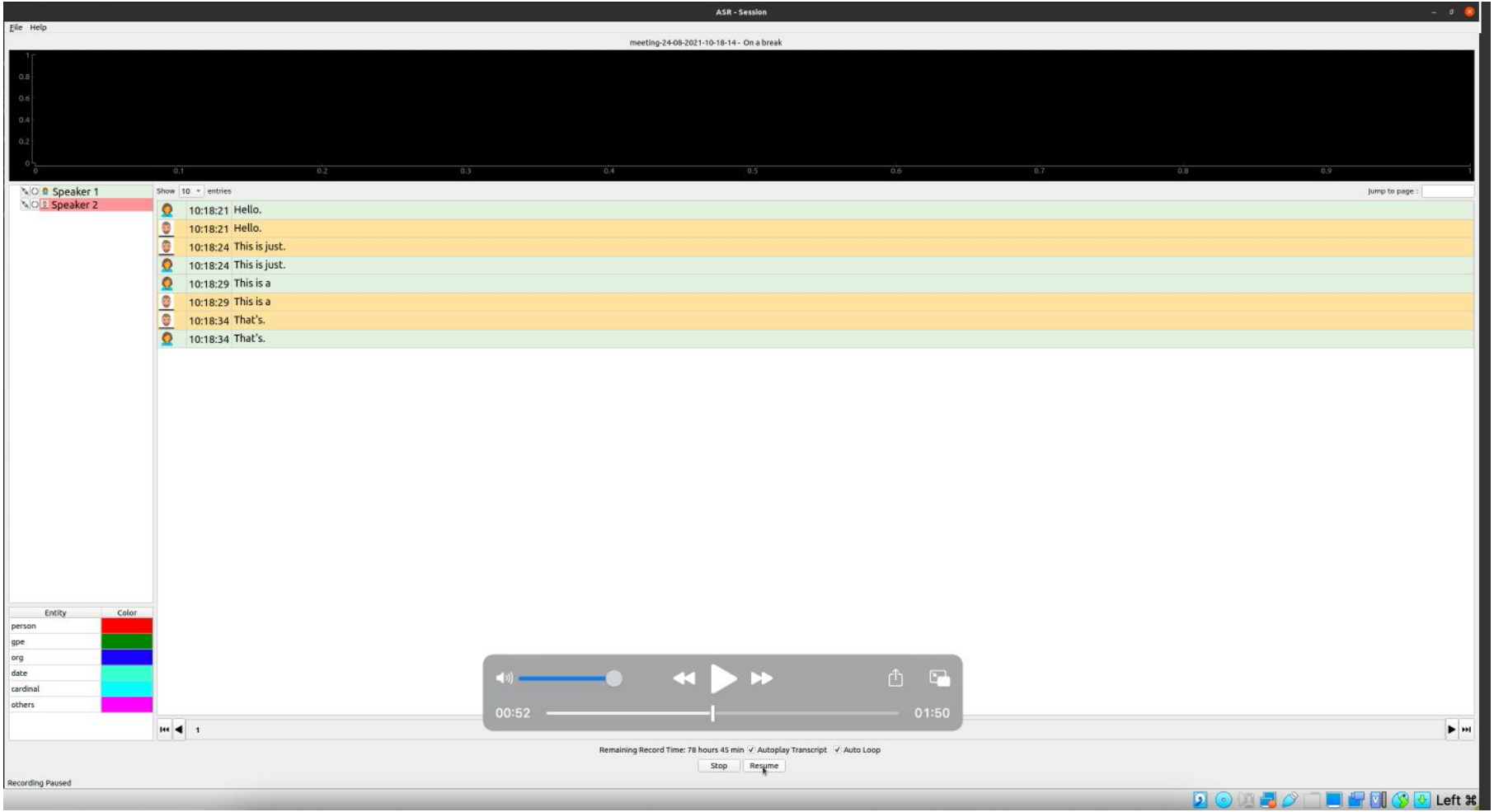
-Enter the name of the meeting (Default: The current Date/Time)
- The Devices/Speakers are setup
-The next button takes us to the fourth screen of the UI

Fourth Screen

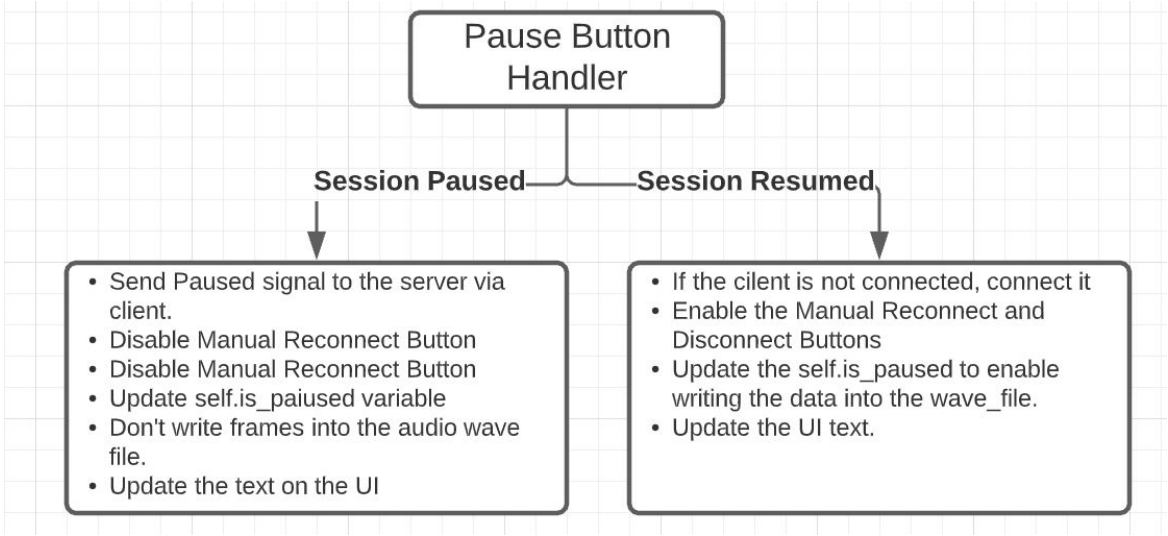


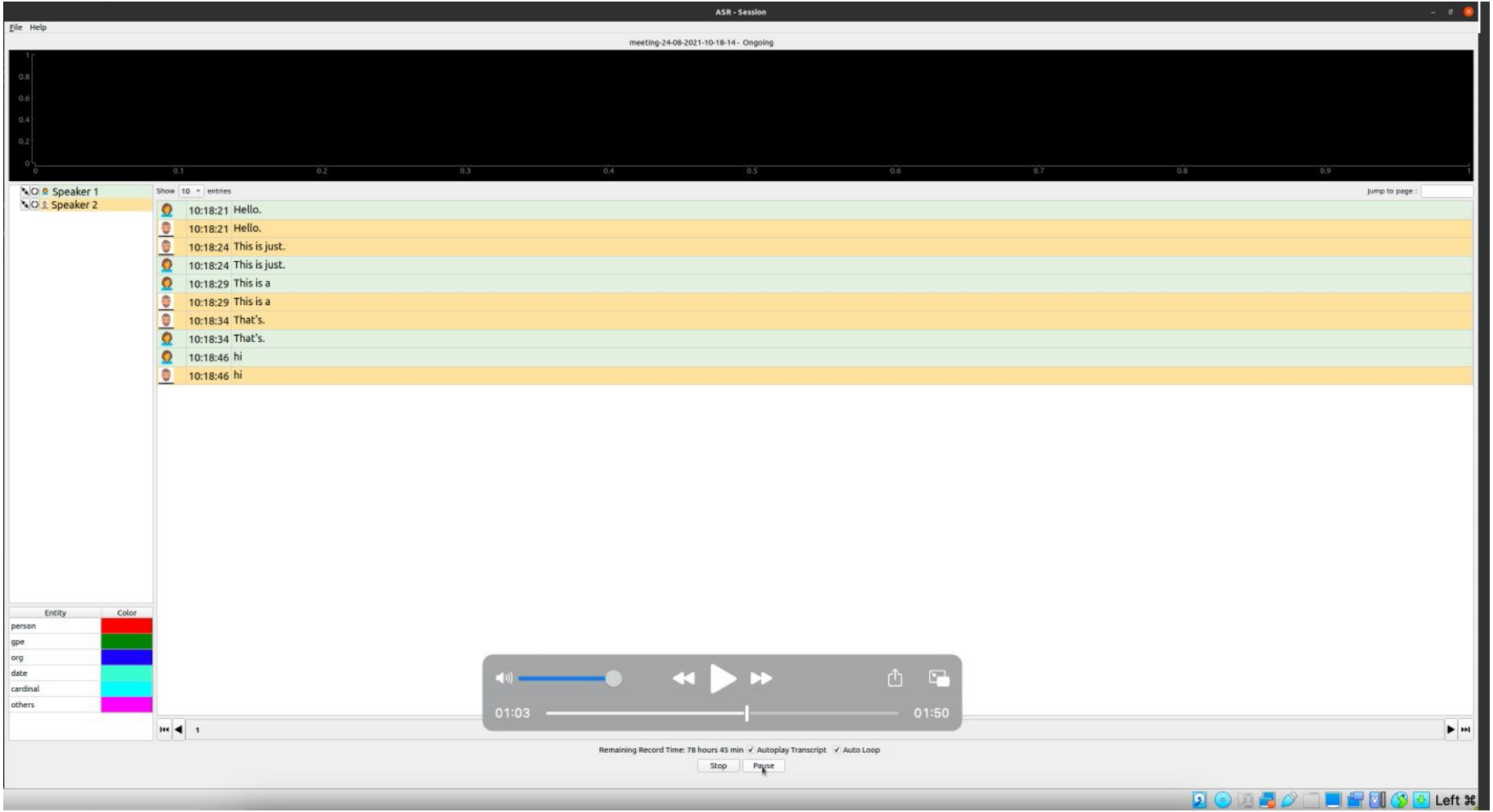
Record Button Handler



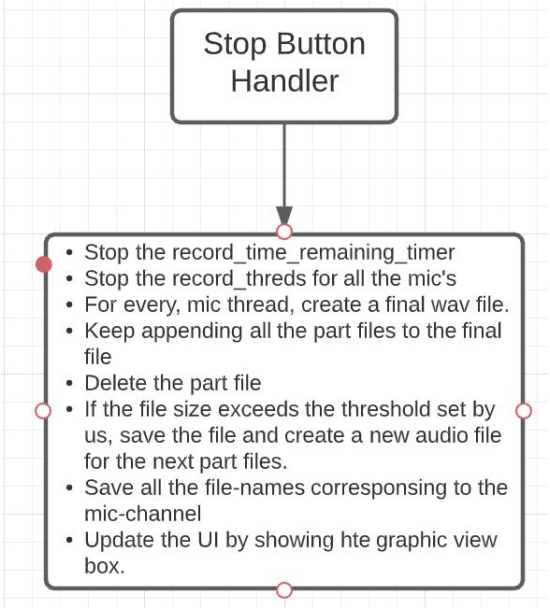

ASR - Session
File Help
meeting-24-08-2021-10-18-14 - On a break
Speaker 1
Speaker 2
Show 10 entries
Jump to page :
10:18:21 Hello.
10:18:21 Hello.
10:18:24 This is just.
10:18:24 This is just.
10:18:29 This is a
10:18:29 This is a
10:18:34 That's.
10:18:34 That's.
Entity
Color
person
gpe
org
date
cardinal
others
00:52
01:50
Remaining Record Time: 78 hours 45 min
Autoplay Transcript
Auto Loop
Stop
Resume
Recording Paused

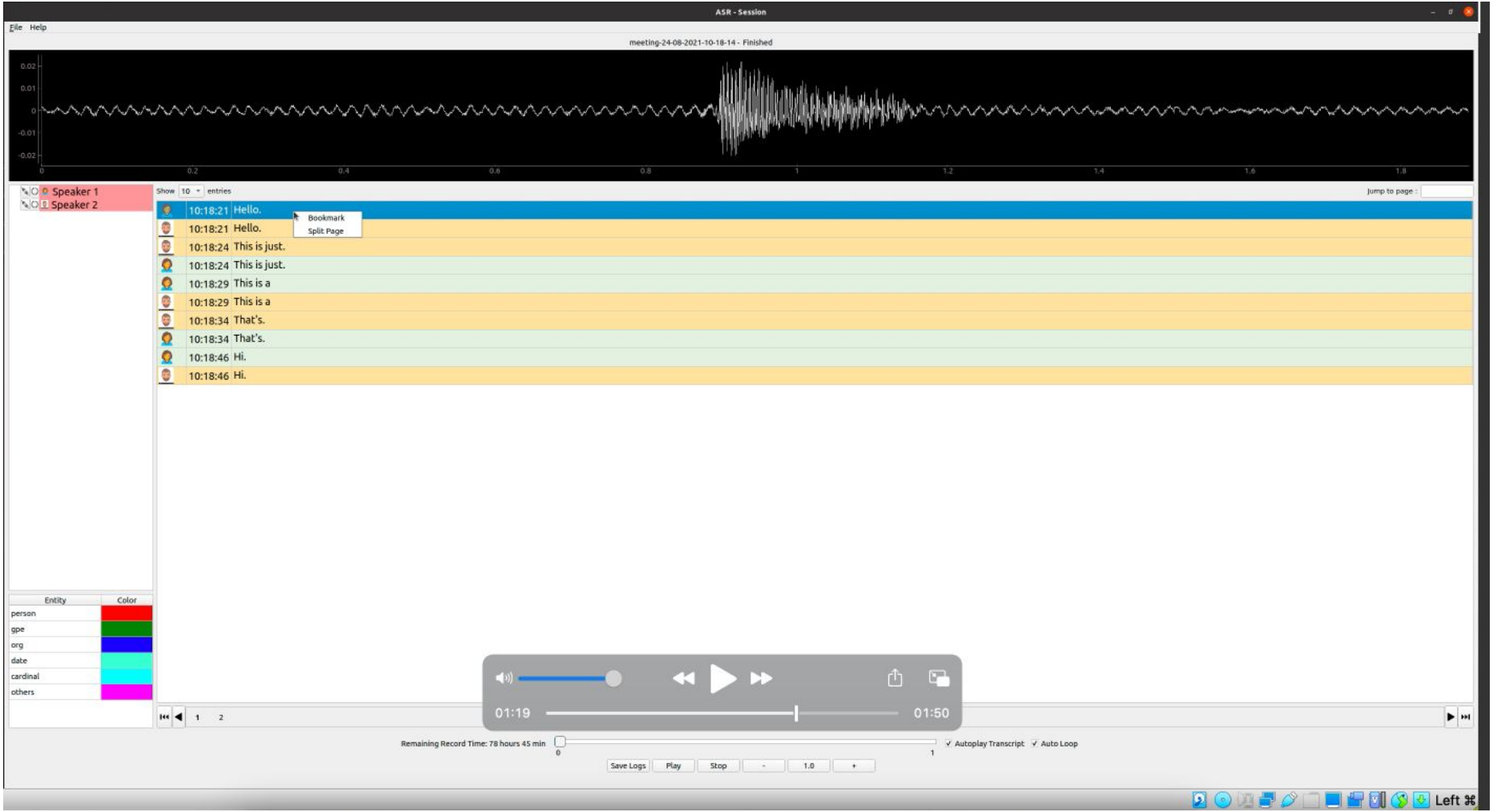
Pause Button Handler



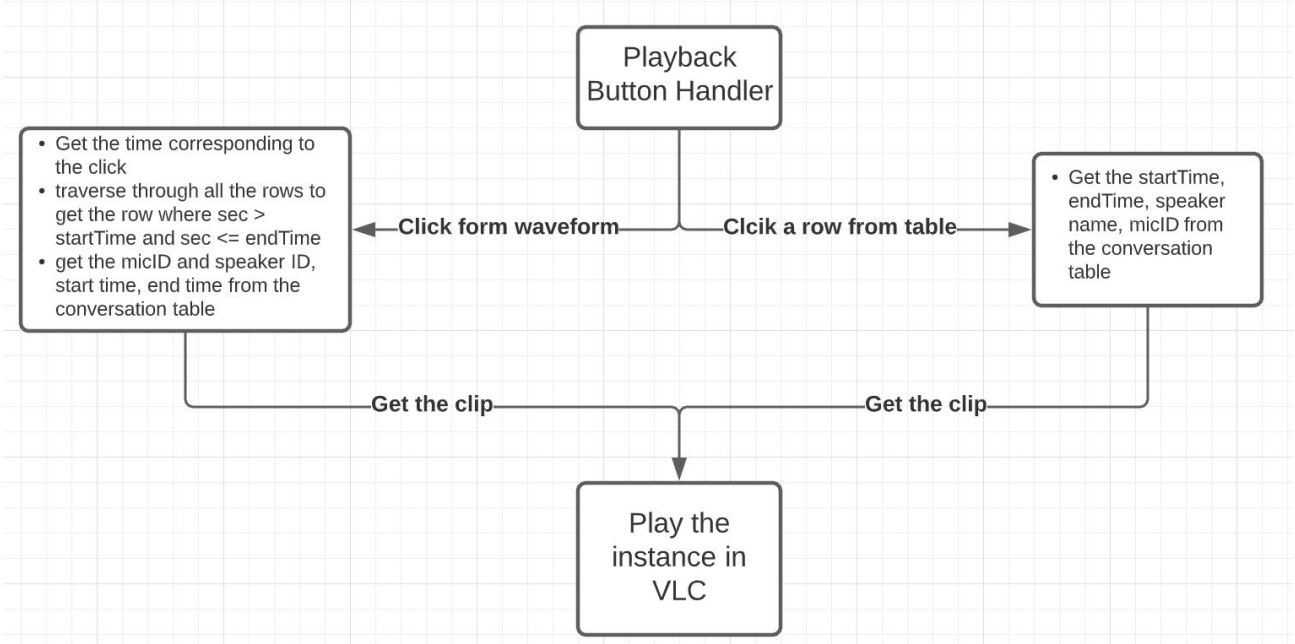

ASR - Session
File Help
meeting-24-08-2021-10-18-14 - Ongoing
Speaker 1
Speaker 2
Show 10 entries
Jump to page :
10:18:21 Hello.
10:18:21 Hello.
10:18:24 This is just.
10:18:24 This is just.
10:18:29 This is a
10:18:29 This is a
10:18:34 That's.
10:18:34 That's.
10:18:46 hi
10:18:46 hi
Entity
Color
person
gpe
org
date
cardinal
others
01:03
01:50
Remaining Record Time: 78 hours 45 min
Autoplay Transcript
Auto Loop
Stop
Pause

Stop Button Handler

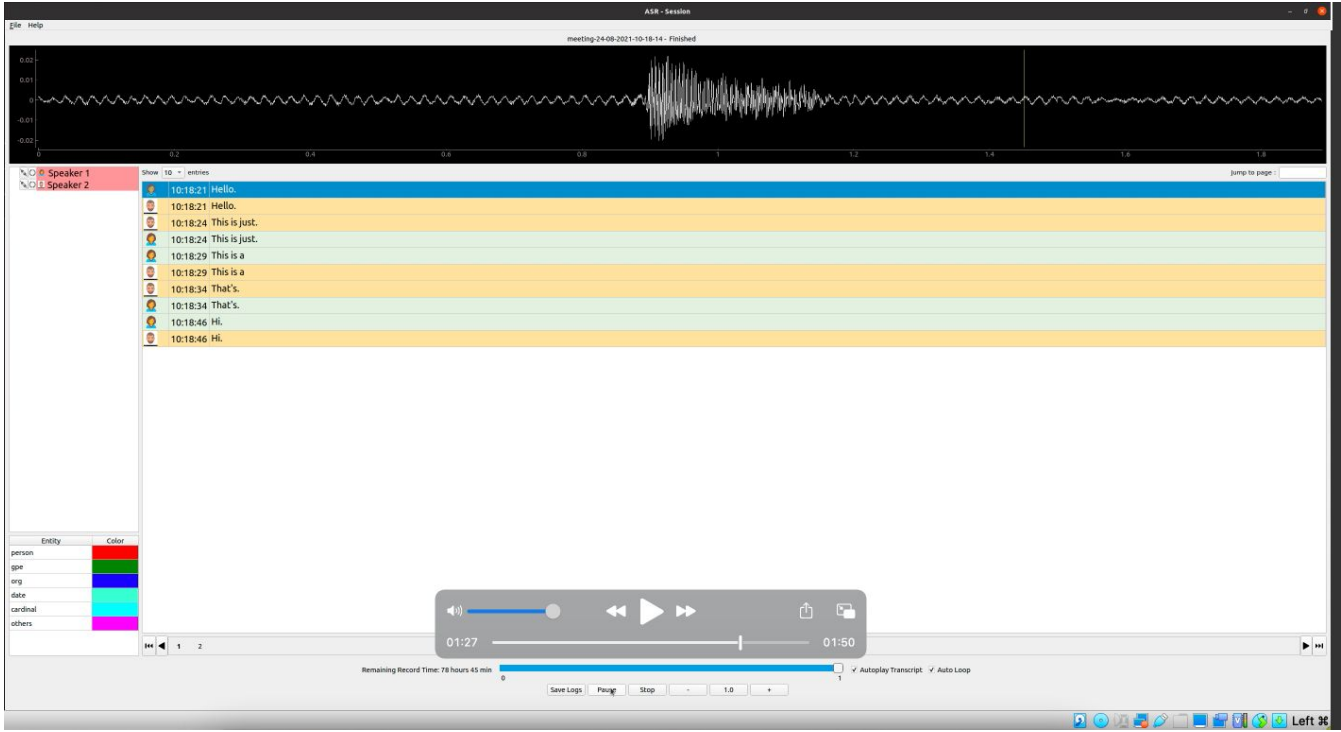



ASR - Session
File Help
meeting-24-08-2021-10-18-14 - Finished
Speaker 1
Speaker 2
Show 10 entries
Jump to page :
10:18:21 Hello.
Bookmark
Split Page
10:18:21 Hello.
10:18:24 This is just.
10:18:24 This is just.
10:18:29 This is a
10:18:29 This is a
10:18:34 That's.
10:18:34 That's.
10:18:46 Hi.
10:18:46 Hi.
Entity
Color
person
gpe
org
date
cardinal
others
01:19
01:50
Remaining Record Time: 78 hours 45 min
Autoplay Transcript
Auto Loop
Save Logs
Play
Stop
1.0
Left ⌘

Playback Button Handler

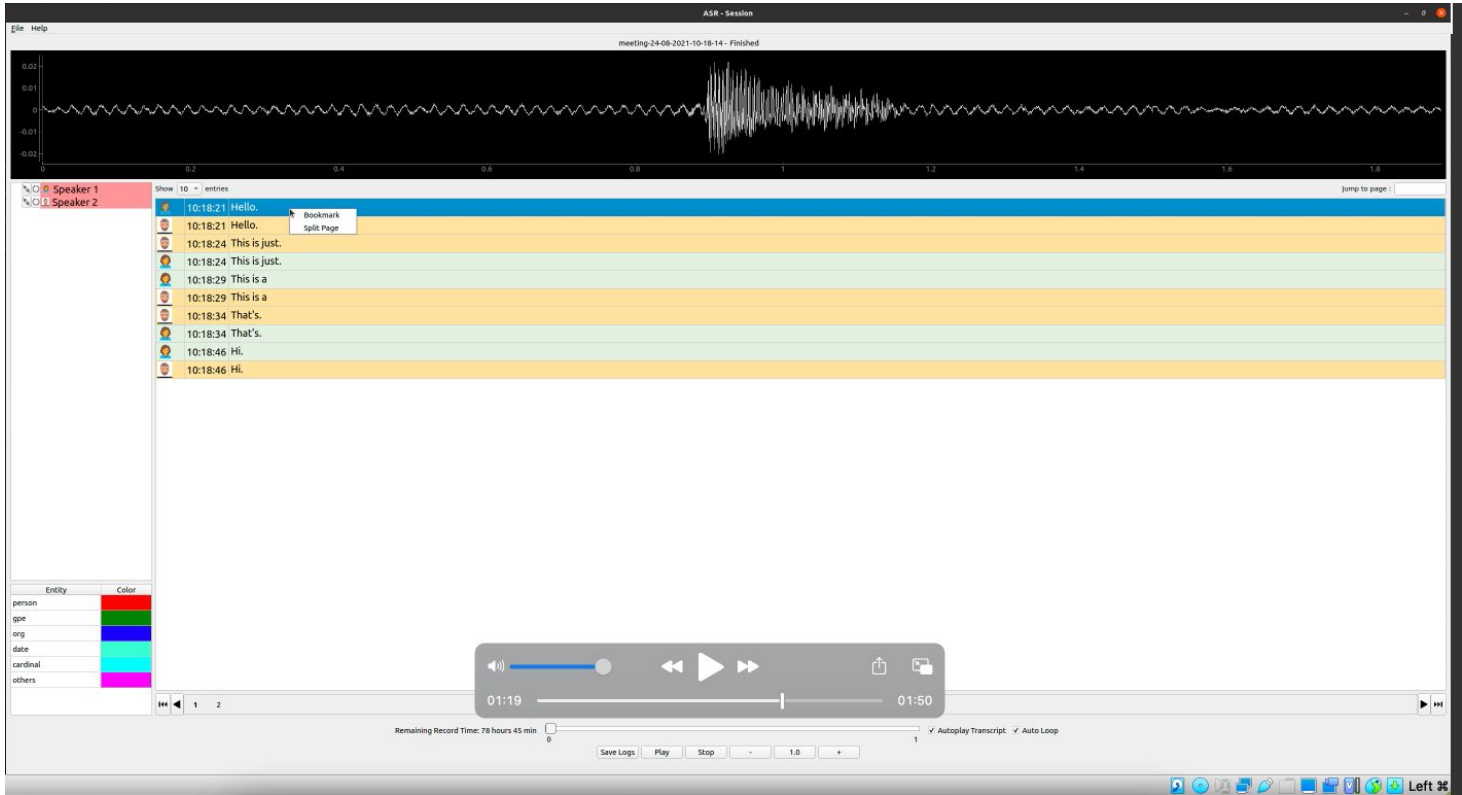


Increase_playrate(): +
Decrease Playrate(): -
Reset Current Rate():



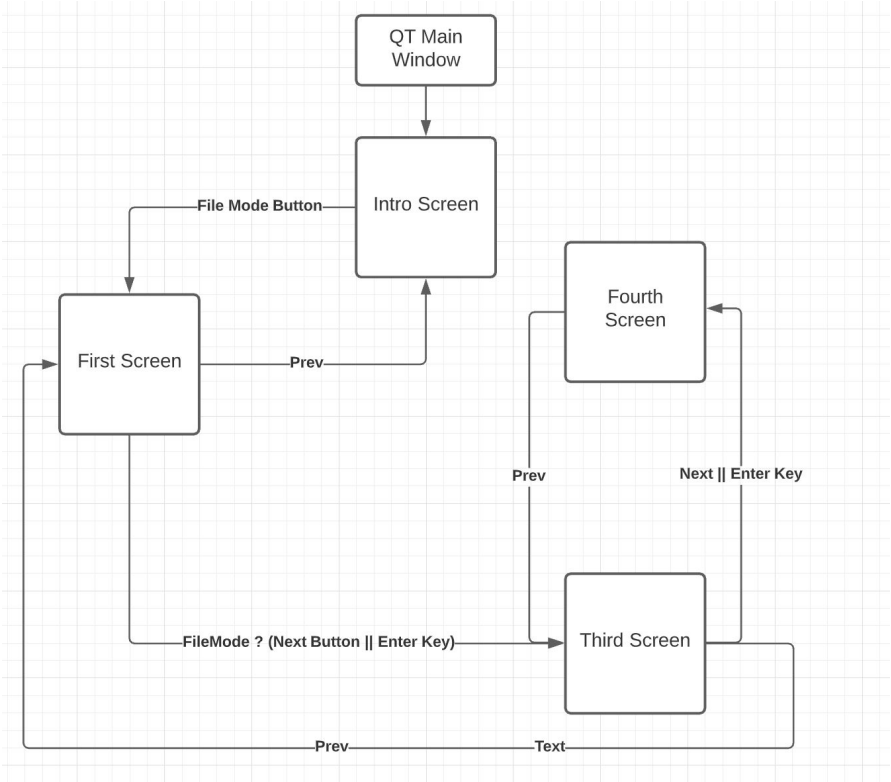
Paging

Next Button Handler():
Prev Button Handler():
Skip First Button Handler():
Skip Last Button():
List Item Clicked():
Update Page Limit():
Page Jump():
Right Click event filter():
BookMark():
Split Page():



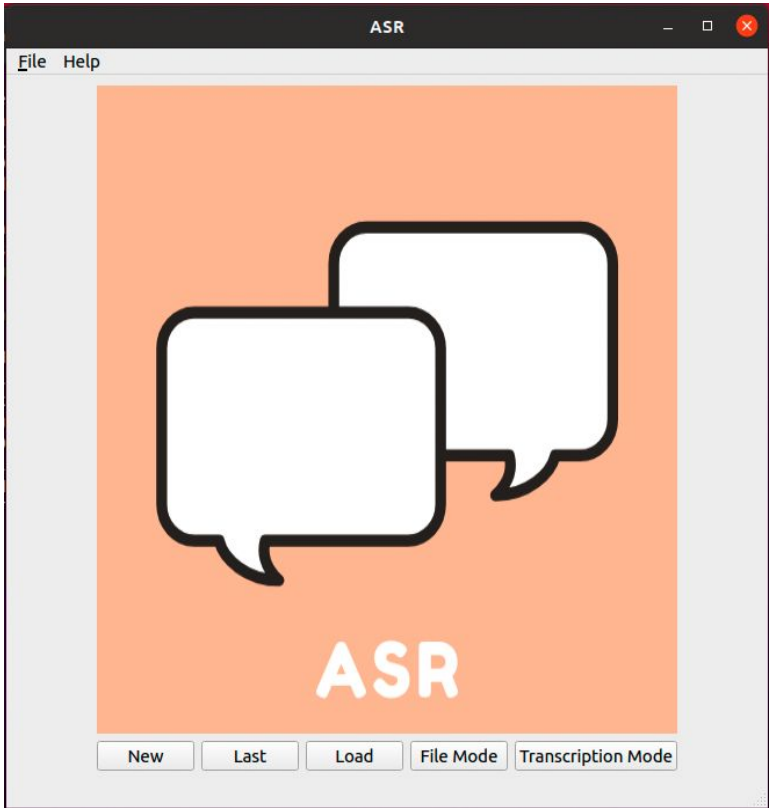
File Mode

Design


QT Main Window
Intro Screen
File Mode Button
First Screen
Prev
Fourth Screen
Prev
Next || Enter Key
FileMode ? (Next Button || Enter Key)
Third Screen
Prev
Text

File Mode

- Main Window is Declared
- Intro Screen UI is setup
- Intro Screen Attributes:
  - ASR Logo
  - New Button
  - Last Button
  - Load Button
  - **File Mode Button**
  - Transcription Mode Button



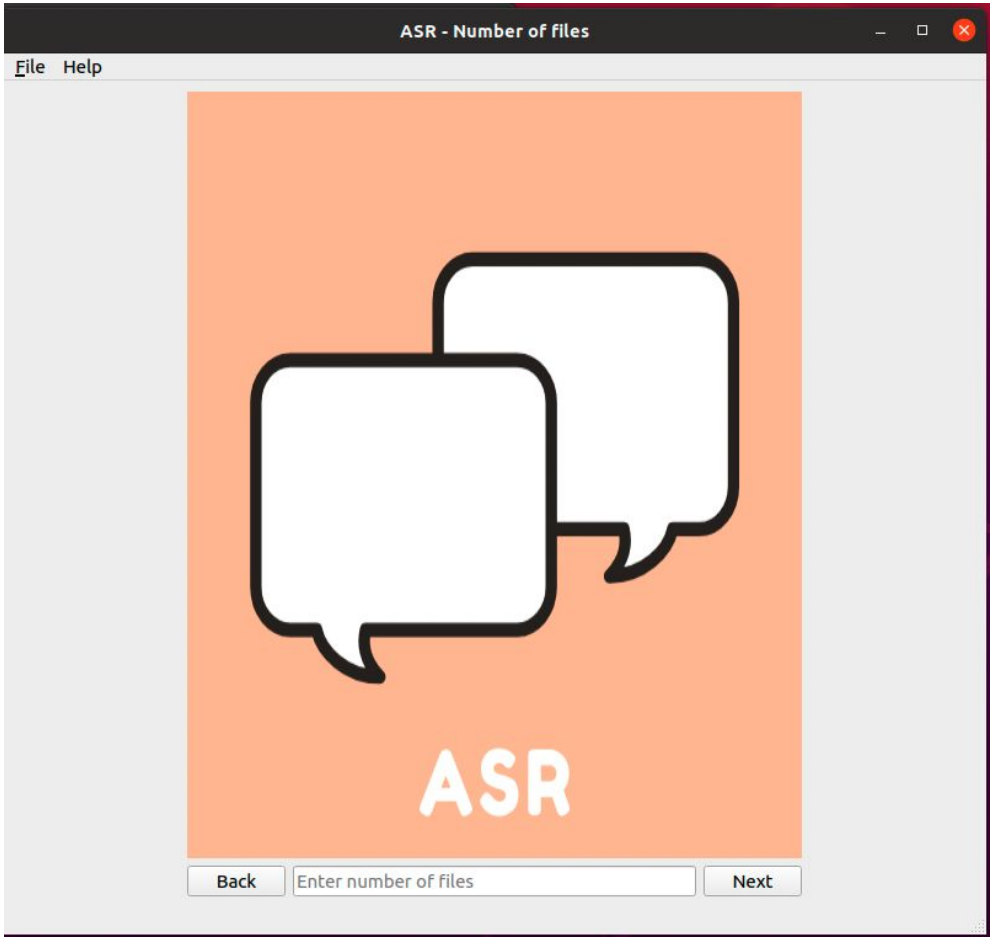
- > Selecting the new Button takes us to the the first screen

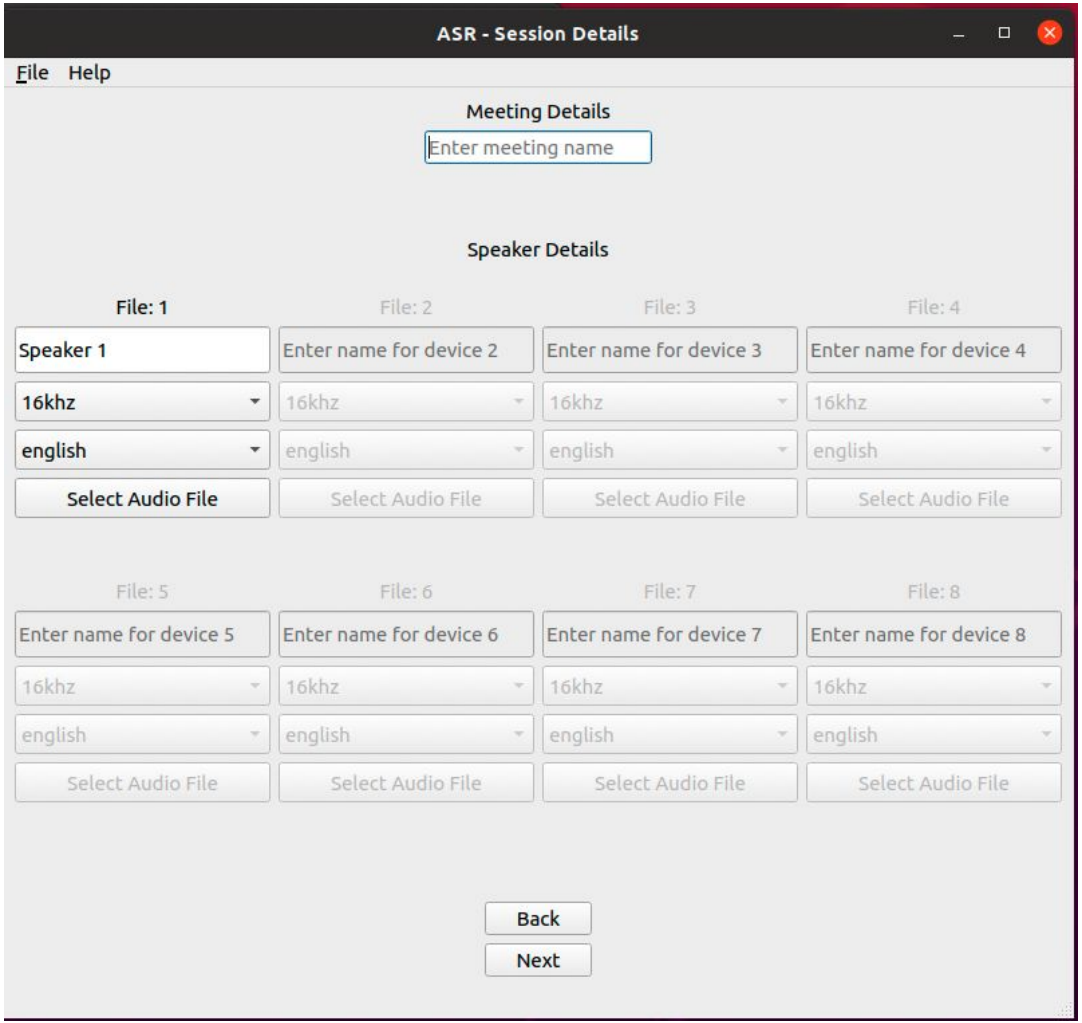
- First Screen UI is set up

- Back Button takes us to the previous screen i.e. Intro Screen

- Enter the number of files (Max 8 Files)

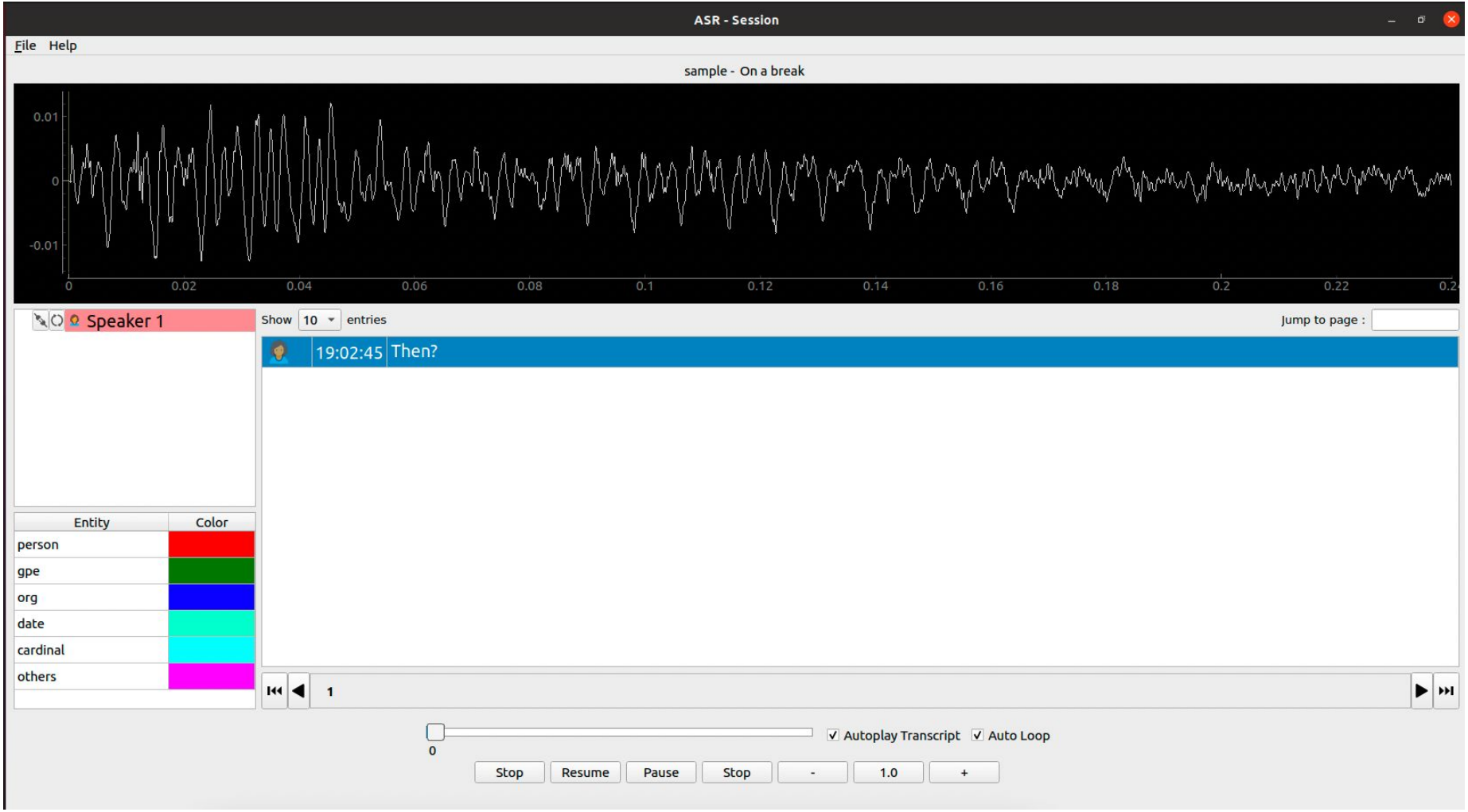
- Next Button handles the number of devices and sets up the second screen by passing the number of files.





The third Screen UI is loaded

-Enter the name of the meeting (Default: The current Date/Time)
- The Devices/Speakers are setup
- The audiofile for that speaker needs to be selected.
-The next button takes us to the fourth screen of the UI


ASR - Session
File Help
sample - On a break
0.01
0
-0.01
0 0.02 0.04 0.06 0.08 0.1 0.12 0.14 0.16 0.18 0.2 0.22
Speaker 1
Show 10 entries
Jump to page :
19:02:45 Then?
Entity Color
person
gpe
org
date
cardinal
others
1
0
Autoplay Transcript
Auto Loop
Stop Resume Pause Stop - 1.0 +

Async ASR

Anand Theertha Nakhate
01 September, 2021

Multi-Channel Audio File Support

- The system currently has a limitation whereby the user is unable to transcribe audio files with multiple channels. The user is required to split the channels into different audio files manually to process the audio file.
- The application can automatically split the audio file with multiple channels into separate audio files and process them.

Library: wave, numpy or spicy.io

- Get the number of channels in the audio file (wav.getnchannels())
- For each channel, store the channel data into a different file with the name given by the user/ speaker name.
- Use the required channel audio files as an input.

Conversation Table Resize

Problems:

- The Row number of the table is used in the payload of the ASR and the SUD services.
- We pass the row number and the data of the row is updated in on receiving the hypothesis.
- Updating the row number will cause errors in the above services.

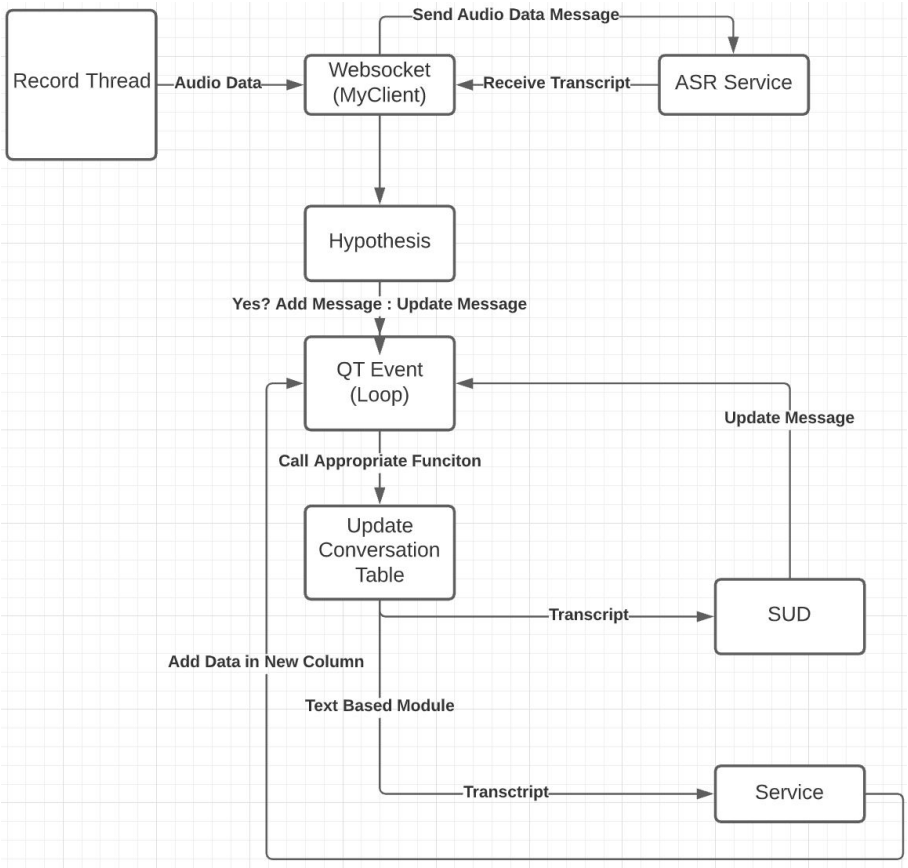
Alternative:

- Add a horizontal scroll bar. (Will need to lookup about this).

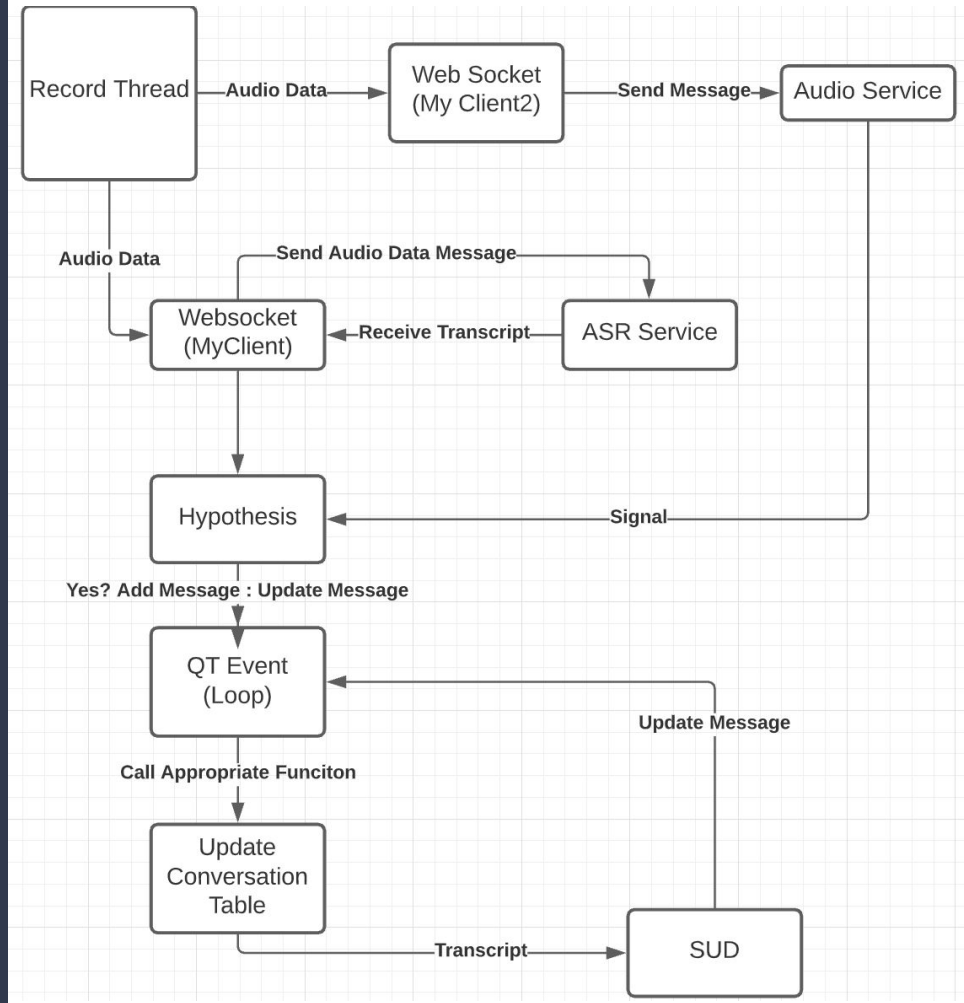
Using a Speech Services

1. HTTP requests
   - Requests Library
   - Futures Add-on to requests Library.
     (Create an instance of Futures Session)
   Helps in asynchronous communication.
   - Whenever we make a http request, we add the request instance to the queue.
   - Every 10 seconds, we check if the request is fulfilled. Else, we resend the request.
2. Server-Client based communication.
   - Create a web-socket for the communication.
   - Send the message through the web-socket
   - Create a signal to process the message when received

Text Based Speech Module



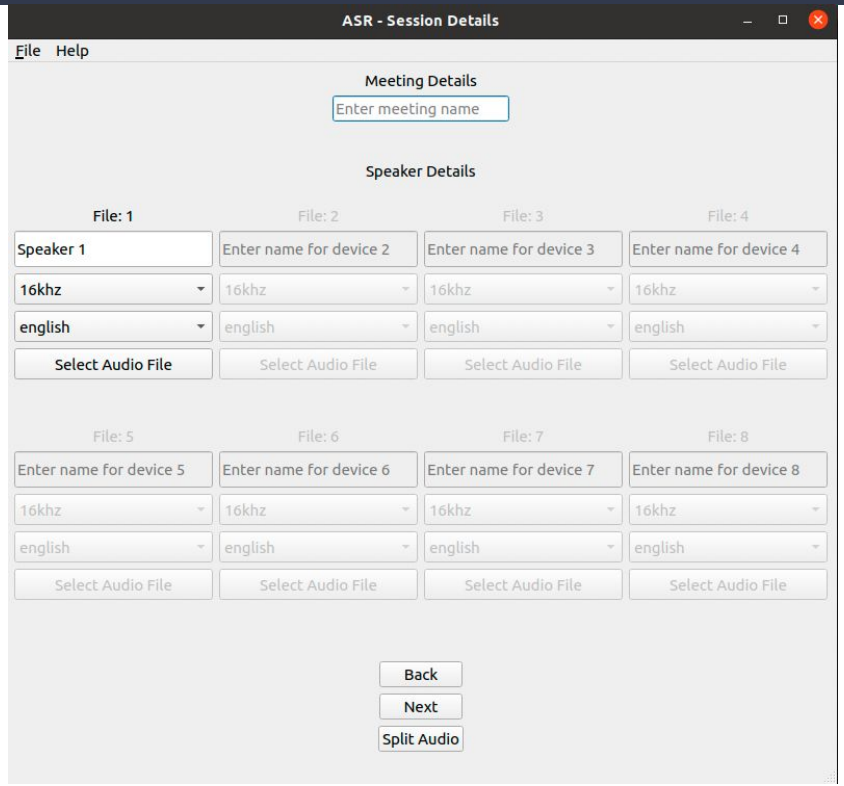
Audio Based Speech Module



Asynchronous Automatic Speech Recognition

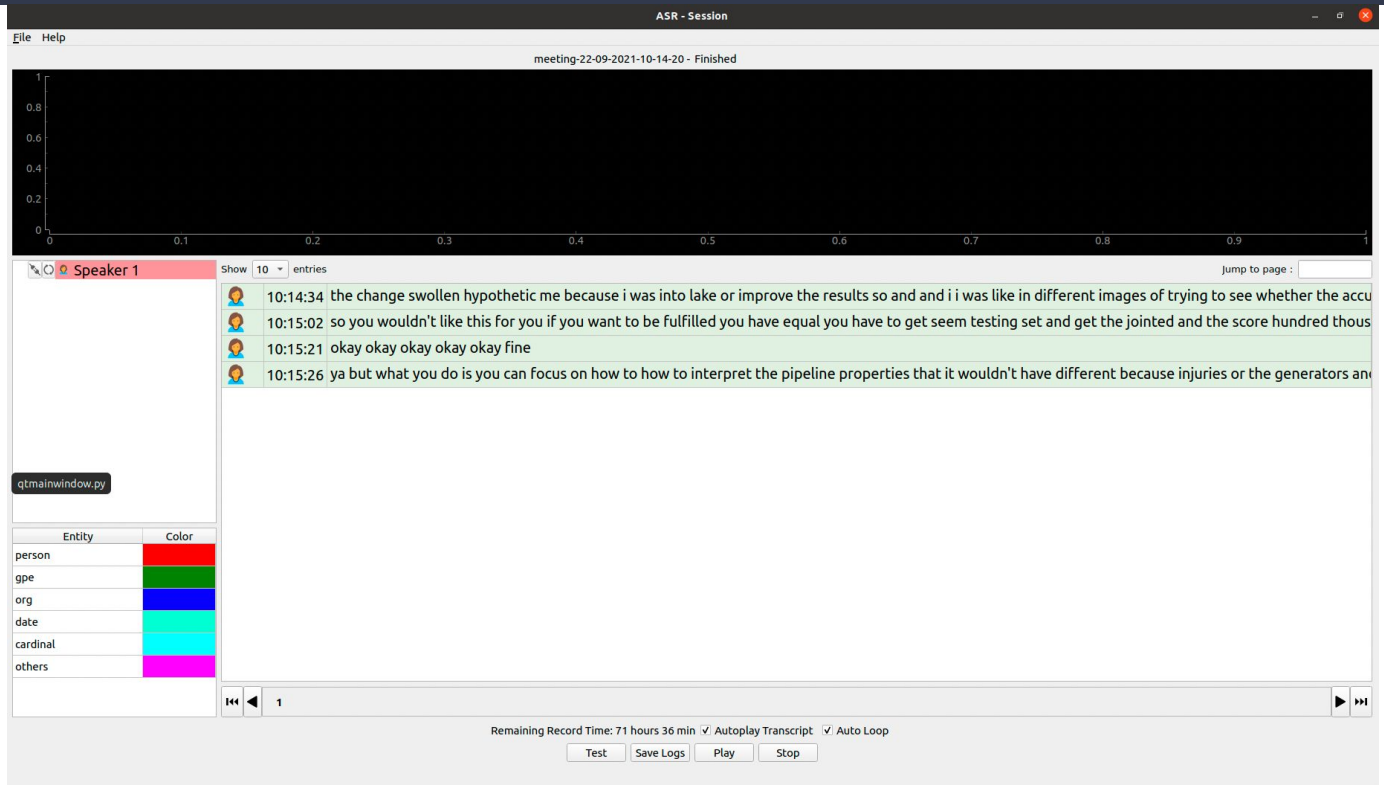
Anand Theertha Nakhate
22 September, 2021

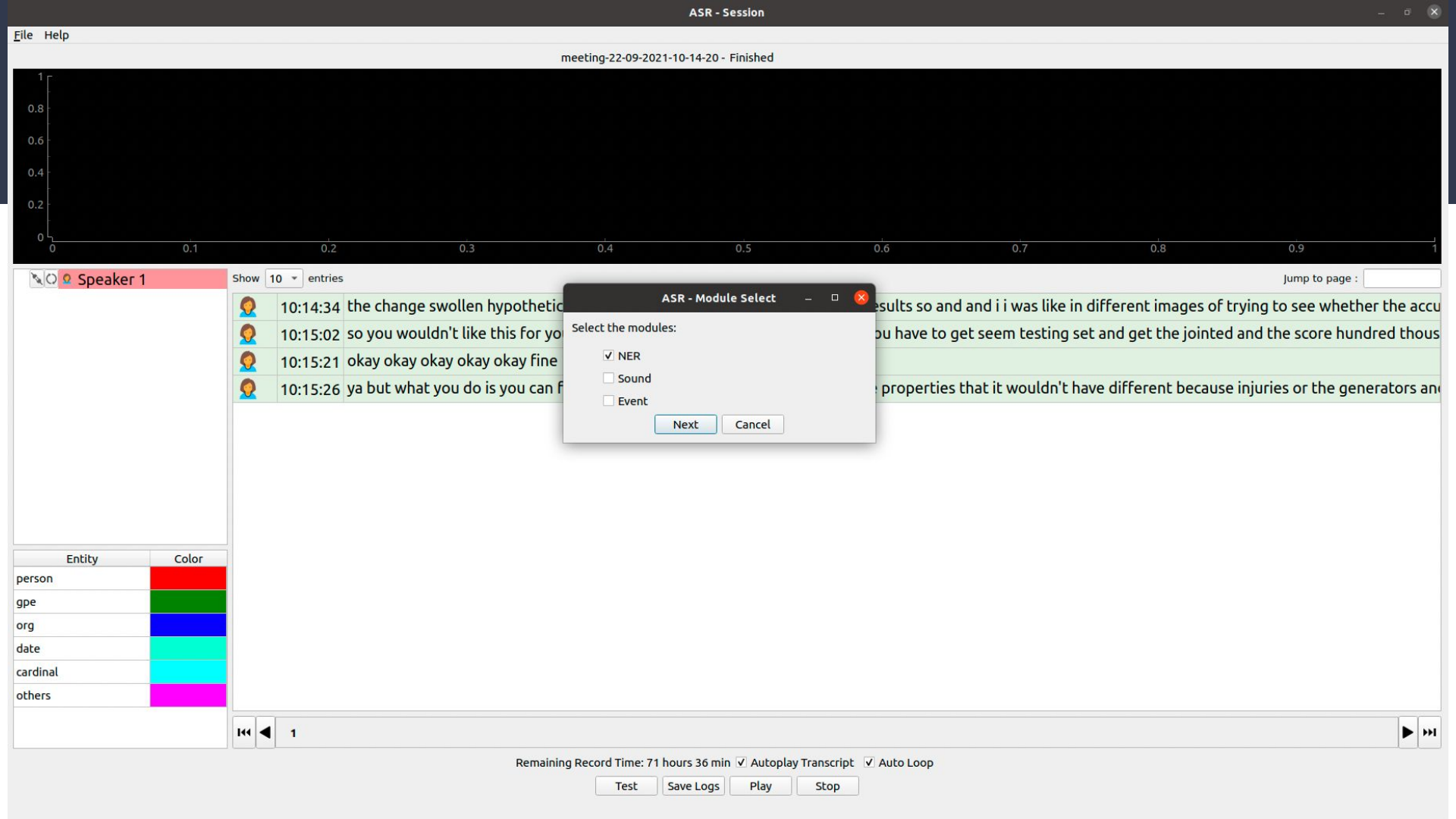
Split Audio with Multiple Channels.



- Split Audio Button Opens a File Selector Window
- Select the Audio with multiple channels to split into individual audio files.
- SOX module splits the audio file into multiple audio files, one file for a channel
- If the audio file is file.wav, the split audio files are saved as file-channel-1.wav, file-channel-2.wav … etc
- The audio files are stored in the temp folder in the current directory.

ASR - Module Analysis




ASR - Session
File Help
meeting-22-09-2021-10-14-20 - Finished
Speaker 1
Show 10 entries
Jump to page :
10:14:34 the change swollen hypothetic esults so and and i i was like in different images of trying to see whether the accu
10:15:02 so you wouldn't like this for yo ou have to get seem testing set and get the jointed and the score hundred thous
10:15:21 okay okay okay okay okay fine
10:15:26 ya but what you do is you can f properties that it wouldn't have different because injuries or the generators an
ASR - Module Select
Select the modules:
NER
Sound
Event
Next
Cancel
Entity
Color
person
gpe
org
date
cardinal
others
1
Remaining Record Time: 71 hours 36 min
Autoplay Transcript
Auto Loop
Test
Save Logs
Play
Stop

👩‍🦰	10:15:02	so you wouldn't like this for you if you want to be fulfilled you have equal you have to get seem testing set and get the jointed and the score hundred thousands of images
👩‍🦰	10:15:26	ya but what you do is you can focus on how to how to interpret the pipeline properties that it wouldn't have different because injuries or the generators and distributorship

Terminal

Cancel

Current Problem

I am not able to use the Analyze button more than once while running the project.

Error:

RuntimeError: wrapped C/C++ object of type QTableWidgetItem has been deleted