Meeting 2 - Oct. 6th

We need transcription built into the web app. As well as capture audio for a chat box. Using those we need to be able to direct a user how to dictate their speech.

Google was used for facial landmarks. Looks for the nose and creates 468 points, used for monitoring facial expressions. Movement disorder is often deadpan, and its hard to tell if they are talking or not due to not being able to properly expressing themselves. The challenge with using this approach is that since everyone is different, its hard to distinctly recognize an individual’s range of motion. So we need to find a way to map an individual’s expression with how motivated they are to continue their speech therapy. One of the goals is to bridge the gap between actual physical therapy and speech therapy.

The therapy app is able to sort of skirt the regulatory bodies for healthcare due to the fact that its software, rather than a full on medical device. So the app can assist with therapy without needing to provide an actual diagnosis.

The app goal is two fold, one to collect data for training ML models on how to transcribe speech troubles better, and second to be able to assist in therapy at the same time.

Go over the document he will send to get a decent overview of some of the problems we are trying to solve/assist with in developing the software.

At a baseline, we want to change speech therapy into something that takes their typical speech during conversations to develop techniques for therapy. Rather than the typical ‘boring’ or not intuitively effective therapy i.e not repeating a word 10 times to measure your effectiveness at a baseline. Our software would be able to gauge a patient’s wants for a therapy sessions in real-time and based off of that tailor their speech therapy.

Ask the patient the length of time they have for the session, how difficult of a therapy session they are wanting that day, and then what they want to do during the session. All based on the context of their previous therapy sessions.

He will give us access to the FTP server, and then we can play around in the system to figure out how to work in an external server environment. He doesn’t care if we break anything, he wants us to learn and part of that is learning is breaking things.

He will have the current project migrated to the new server and then be able to give us example functionality currently present.

We need to research(maybe not necessarily us) a bit on the menu system for the app, whether or not a menu-less route will be beneficial or not.

There is a speech pathologist as a co-investigator, which dave meets with on Mondays every other week. As well as him meeting with UBC on Friday afternoons.

We may join a few survey calls he has with actual speech therapy patients, so we can get more information and context directly from the users.

The real-time chat transcription is to gauge their speech, and then the therapy portion is still whatever the traditional method is to improve.

Sort of an idea of where we may start:

Get react installed on the server, as well as our directories set up. Potentially get maybe whisper, googles tool, and/or deepgram on the server as well.

Start with getting transcription up and running (?). They were using openAI Whisper, but the limitation ended up being a struggle with real time transcription and it doesn’t give you a word error rate i.e the tool doesn’t understand a word. If there is a word error with Whisper, it would automatically use context to provide a replacement. Which is not what we want for therapy.

Google has a better tool for real time transcription, it will give “word error rate” at a confidence rating. Then it will give likely word replacements, but it isn’t like a traditional word error rate.

Deep Gram(?) is a company that can give real time transcriptions, which may be an option.

We will develop multiple designs for the app, at a miniature level, to give to users and get feedback to gauge whether the design is effective or not.

Get react installed on the server, as well as maybe whisper, googles tool, and/or deepgram.

The real-time chat transcription is to gauge their speech, and then the therapy portion is still whatever the traditional method is to improve.