Pre-Test Checklist

1. Print off consent forms/have access to digital form for remote
2. Open this doc
3. Choose the podcast and transcript to be used for this specific session
4. Load the server backend and prepare the frontend for use
5. Verify the participant's name and interview time from your Notes
6. Add participant ID to the notes and create a new note for them
7. Phone volume off
8. Disable any notification or popup since you will be recording the screen.
9. REMEMBER TO SPEAK SLOWLY
10. Have the post-study google form open
11. For Remote: Make sure your zoom enables remote control
    a. (instruction:
       ▶ How To Collaborate With Other Music Producers Online Using ZOOM | Re…)
    b. https://www.coscreen.co/blog/how-to-share-screen-control-in-zoom-remote-control-mouse-keyboard/
12. For Remote: Share zoom link
    a. https://utoronto.zoom.us/j/5842463173

After participant enters

Hello _____ ,

Thank you so much for taking time out of your day to help us with this study.

My name is ____ , I am a researcher for this project. I will be guiding you throughout the process, taking notes and asking you questions for clarity.

*Here is the **Consent Form**, please fill this out before we begin.*

Before we get started, do you mind if I record this session for transcription and review?

(Once they confirm, BEGIN RECORDING.)

Perfect, so in this study, we will ask you to use a research prototype of a system that integrates AI-powered avatars into the podcast listening experience. Can you briefly talk about your experiences with podcasts in the past? Do you create them? Do you listen to them?

At the end of the session, we will have you complete a survey. It will begin by asking you about your experience using the system and will end with you providing some demographic information and podcast usage information. Following the completion of the survey, we'll have a short exit interview to discuss your thoughts and experience with the system.

After the study, we will send you payment for assisting with our research.

The whole study should take no more than an hour and a half.

Start interface tutorial (10 mins max)

(FOR REMOTE: Set up remote control)

Alright, now I'm going to demonstrate some of the features you may engage with during the study.

FOR REMOTE: Can you see my screen?

Tutorial:

We're going to start by listening to five minutes of the podcast. While this is happening, I'm going to demonstrate ways you can interact with the system. Feel free to stop me at any point and ask for clarification or question if you aren't unsure.

Note that this is still a research prototype, the voices of the avatar may sound strange or provide odd responses. Also, since the UI is quite bare-bones I want you to focus more on the responses from and interactions with the avatars.

UI/Features	Script
Interface	*Before we begin, I want to point your attention to the UI. As you can see, the system has four panels. To the right, we have a transcript that can be clicked on and scrolls alongside the podcast audio.* *To the left at the top, we have an icon for each avatar within the podcast, these represent key figures within the podcast and one separate AI expert. These avatars can be clicked on to ask questions. The avatar with the microphone is the host, the one without anything is the guest on the podcast, and the one with the headphones is a separate interviewer. For now, we limit the number of avatars to one per role for ease of use and to limit computational expense.*

	Below that, we have the media player, here you can control the playback and audio of the podcast. *Lastly, we have the chat box. This will store your conversation. Additionally, the textbox at the bottom allows you to type in questions. The microphone can also be clicked to either start detecting your voice or stop, it will turn red whenever the microphone is active.*
Solo Question	*(Start Playing the podcast audio)* *First, I'm going to ask a question to just one avatar and see how it responds. I'll show you how you can do this in two ways. The first is by clicking on the avatar I want to talk to. When I click on an avatar, the podcast immediately pauses.* *(Click on the desired avatar and ask a question)* *Notice that we can see our question in the chat box along with the podcast timestamp of where we asked the question. When we click on a specific avatar, our question is tagged with the name of the avatar. Now we just briefly wait for a response. Note that the chat box shows that the response is loading.* *(Wait for the full response)* *The response from the avatar is also here with the same timestamp, we can scroll through the chat box to see our conversation history. The response has the name of the avatar as well as the icon of their role next to it. Let's continue listening.* *(Keep listening to the podcast for a little bit)* *The second way to ask an avatar a question is by typing in the textbox. We can simply mention the name of the person we want to ask. Note that the episode will not pause when I engage in this mode.* *(Type in the question that will receive an in-transcript answer and hit enter)* *Notice that we can once more see our question in the chat box along with the podcast timestamp of where we asked the question. This time, our question is not tagged with the name of the avatar.*
Clicking on Links in Response	*(Wait for the previous response to come through)*

	You may have noticed that parts of the response are underlined. This means that the response generated by the avatar was based on specific parts of the transcript. We can click on these links to skip to that part of the podcast. *(Demonstrate the skipping)*
Discussion	*(Wait for a relevant part of the podcast)* *Another way we can use this system is to have the avatars talk to each other, for this we can simply mention the names of the avatars we want to speak to each other.* *(Type in a question for discussion)* *The discussion will continue until either it's been determined to reach a natural conclusion or whenever 7 responses have happened, whatever comes earliest. You can ask any number of avatars to talk to each other.*
Intervention (other avatar or AI)	*(Wait for the previous discussion to end)* *Another way that the system may handle your prompt is it may cause another avatar or the AI agent to intervene. There are two cases.* *The first case happens if it seems like a different avatar would be better or more qualified to respond to the prompt. This also means you don't have to specify an avatar at all, the system will automatically determine who should answer.* *The second case occurs if the system determines that your prompt may lead to a response that may include misinformation, hate, violence, or criminal activity. Let's try to make the latter occur.* *(Ask a question that will cause the AI to intervene)* *Notice that I can change my prompt by using the textbox, this automatically stops the voice recognition.*

(DOUBLE CHECK ITS RECORDING)

Start task (45 mins)

Type	Questions
Solo	This American Life - Dr. Gilmer and Mr. Hyde Lex Fridman
Discussion	
Intervention	

(START the timer)

Again, please think aloud while engaging in this task. You can begin now.

(IF 10 minutes remain)

Just letting you know 10 minutes remain.

(IF 5 minutes remain)

5 minutes left.

(IF they are silent for a while and not engaging in think-aloud)

(_____, just a reminder to think aloud. Thanks!)

Start post-study survey (10 mins)

Thank you for completing the task.

Let's now move on to the post-study survey. I will send the link to the survey in the chat.

- *Here's the link:*
    - *https://forms.gle/A7ChVWKN11KPWHqy8*

Let me know once you are done.

Post-study interview (20-30 mins)

(CHECK that the participant's submitted responses show up on Google Form)

Thank you for completing the post-study survey. Before we end, I will ask just a few questions about your experience and your survey responses.

> **WHAT WE'RE INTERESTED**
> - What was useful / not useful? Why?
> - What interactions with the avatars were meaningful?
>
> *** Ask a very broad question at the beginning

- *As a start, could you please describe --- walk me through --- how you used the system?*

- *What interactions with the avatars did you find meaningful, what interactions did you not find meaningful?*

- *Now, I will ask you some questions about your survey responses.*

  (Skim through the participant's responses on Google Form and ask)

> * Ask them to elaborate, especially extreme answers, e.g., Strongly Disagree, Strongly Agree.

- *For a question, "_____", you answered "_____" Could you please elaborate on your response? (That is, why you answered _____?)*

- **How did interacting with the avatars change your typical listening experience? Did you enjoy it?**

- **Would you trust a system like this? If not, what could improve your trust?**

- **Do you see any other use cases for this system? If yes, can you elaborate?**

Individual features

> * There could be instances where it wasn't clear what they were doing during the task, because they didn't engage in think-aloud. Ask them.

- *I noticed you [used X a lot OR didn't use X much]. Can you elaborate?*

- *What was your favourite type of conversation to have with the avatars?*

- *What did you think of the voices of the avatar? Did you find they lined up with the real voices? Would you prefer a more generic voice?*

- *Have you encountered any obstacles when using the systems? What are those?*

- *Great, I just have one final question: "Do you have any final comments or suggestions for how we can improve the study or system?"*

Participant Debrief

Great, thank you so much for answering my questions. Now, I am going to stop recording.

Now it's no longer recording.

- *Do you have any questions, concerns, or anything you would like to discuss before we end the study?*

Great, thank you so much for taking the time to participate!

P01 - Observation

- I think it got confused where it intervened on itself which led to an error in the backend, now everything is shifted back. (not sure if there was any indication that there was an error, though, so something we should discuss afterward). Who is responding is shifted back by one iteration.

- AIden continues to be slightly broken the system doesn't like to route to it, which causes it to give weird responses. I also want to give it more leeway in responding. It's currently the only model that is actually as concise as I asked. Everyone else goes way over but AIden answers in like 2 sentences, which may not always be enough detail.

- Does skipping misalign it? Or is it the transcript that's misaligned <- I assumed it was the latter but maybe not? That's an easier fix if so. Now that I think of it it is simply misaligned I've checked against the audio itself before and it doesn't line up in some parts.

- 99% transcript in ALL caps at some points and also says Surocao instead of Sarah Seo, probably won't use the transcript again but fix anyways.

- I think it's fine to report some of these errors as examples of cases where things can go wrong. I think they are interesting.

- Yeah, I think it's fine to fix it over the weekend. When asking her to evaluate these features, you might want to preface it by saying, e.g., "as you saw, **there are still some errors**, but we **are mainly interested in your perceived usefulness of these interactive features and how engaging they are, etc. if they are more polished and work as intended as opposed to how well these are currently implemented**, …"

- [Study Protocol] Yep, what's the plan in terms of having the participants trying different types of questions? It seems she is exploring a wide variety, so that's good. But for future studies, we can set our protocol to be, "after observing, if they have not explored some types of questions, in the last 5 minutes before ending the study, we suggested them to use them so we could ask their impression and perceived usefulness of the types of questions." (This is good, as we are also interested in letting them explore and seeing which ones they use more frequently than others. And

- [Interaction] Typed "what is the topic of this podcast" -> the system didn't generate a response -> directing the question to one of the podcasters (Roman) yielded output ("Roman, what is the topic of this podcast")

- [Opportunities] 25 minutes into the study, P01 mentioned after getting the response from the question that she was interested in something beyond what was discussed in the podcast.

- [Opportunities - using prompt for navigating the podcast content] P01 wanted to know when one of the podcast person answers the question of the podcaster and wanted a link to that point in transcript, so she can see when the person appears.

- Subject: Person that is mentioned in the podcast but never speaks, is a more compelling person to talk to then the person who speaks but barely speaks.
-
- Feature Idea: Give the other avatars access to the conversation history of other avatars in certain scenarios outside of discussion

- Not sure why it linked at a specific part.

- Hard to find where you are again, maybe make a clearer indication of timestamp? Or have it direct where it was.

- Links don't seem to click

- [Note] she didn't use the multiple people discussion feature, so we asked her to try it out.

- [Example - content inspired questions] "Does it only apply to US? Does it also apply to Canada?" We should include the multi-interaction P01 took to engage in several Q & A

as an example in the paper, probably in Appendix, to demonstrate how interactivity led to and can lead to deeper engagement with the content.

● Easy to click and start the speech recognition, if you don't want this to happen easy to click microphone and automatically send. I think I'm gonna change it so that you have to hit enter irregardless to avoid this issue in the future.

● I added a few more questions in the survey

● [Improvement on presentation] I would like a snippet of the information in the chat box and a snippet of the audio.

Zhixin's comments
● was wondering if agents keep a history of conversation between me and other agents (Kieran: it doesn't)
● I didn't feel the huge need to use the conversation between avatars.
● I came in with a mental model of assuming I would be mainly interacting with AI avatars.
● [navigation] maybe users should be able to click on the timestamp to get back to where they were,
● [design] Clicking the avatar highlights the clicked avatar's responses in the transcript
● [design] We could use different colors to indicate different speakers in the transcript

Interview:
- Useful not useful: AI was useful, questions about the text that they didn't mention. Conversations between the avatars were ok, didn't find the need to prompt it. The linking was ok, couldn't go back.
- Human interaction - > just link to what they said, scrutinize where the voice is not good instead of absorbing. Answers it gives are based on what the personalities are right. Expect more direct to where they were talking about or the audio clip/
- Transcript misalignment was really annoying.
- Eyes were going back and forth
- Easier way to pause would be useful, click space
- Take me back to where I was, or highlighted last line that was read out loud, use a different highlight
- The way they say that the AI will hop in -> jarring, would very broad, wrap this up was kind of unnecessary, if it played by itself that interaction makes sense but on text it doesn't make sense. Instead have a statement on chatbox like the "will respond" message indicate that the convo is over, AI jump in, maybe if it had a different colour that would help with differentiate it.
- Changing the podcast speed
- Responses given in the chat were quite long, would be nice to break them apart
- AI responses could also be shorter

P02 - Observation

- Clicked an avatar, didn't notice that the avatar was highlighted/active didn't really matter
- Asked the model to automatically change where you are in the podcast, not implemented feature
- Tried to compare Alden's response to Yann's response, ask about how that went
- Ask about using mostly text mode rather than audio mode
- Uses Aiden more
- "Recording" probably got confused with "podcast"

Interview:
- Why he did what he did - envisioning the system as an aid to answer any lingering questions/queries that were not answered by the host
- Meaningful/not meaningful - liked the ability to reference the actor that they talked to directly. Asking Alden about what alpha go was thought that was a good answer, nicely put summary made it feel like talking with the well defined agent. Didn't like that Yann was brought in when asking about who Yann was -> issue with the prompt (detects mentions in the prompt and basically automatically says to discuss with them)
- Answers seemed like they were all coming from the same source -> alphago all had similar answers. Supposed to tell correct information, would be nicer to have them have unique perspectives. Idea -> show that it is approximating the thoughts/personality of the person. Disclaimer that it can only try to mimic what human beings.
- Speech input - didn;t use, pretty much alright
- Liked being able to mention in the prompt who to talk to
- Like the ability to stop and ask a question
- Overall, UI was super straightforward, easy to understand
- Pretty similar to conversation with person
- The voice can understand that it is computer generated - more the content that was less realistic
- Pretty streamlined and ok, not very distracting, long periods of time where did not interact -> amount of control was good, no point where felt like needed to mute things or minimize or maximize
- System lacking? - not particularly, typically with AI focus is less to have a conversation, more like the utility function. Would have liked Alden to be able to be more involved in the podcast playback -> summarize, go to timestamp, etc.
- Most salient part of the system -> liked interacting with Alden, not trying to replicate the thoughts and behaviours of a known human being. Would like it to be integrated into regular content consumption. Having an assistant to ask a question for clarification would be nice -> finds himself googling stuff, would prefer to have interface within application itself.
- Human avatars - combination of latter two -> quality of responses and the person he is made him less interested to interact. Only had basic recognition of the people involved.

Scholarly sounding answers -> would doubt they'd speak like that (I sort of disagree having listened to them). Not really liked AI generated content in the wider space. Copy of a copy of a copy…
- Preference and habit -> like typing it in directly if lazy liked the audio input -> accent transcription may not be great
- Mostly just read them, preference, force of habit -> don't like generic AI voice
- Liked the flexibility of being able to set
- Any major obstacles - overall well designed interface, particularly liked having the text/transcript on the righthand side and the moving of the words, helps as a non-native english speaker. Very easy to learn and use
- Desired features:
    - Would like to use the Alden avatar as a utility -> Alden could control the audio playback, rewind, summarize, skip to parts, playback speed.
- Trust -> trusted the responses in terms of the pod bay and alphago -> Alden, Lex, and Yan gave similar answers that increased trust in the answer. Didn't see citations other than the transcript -> that is useful/nice especially for summarization. Would've liked after the end of the response -> source and website/article and a link to that article or website. Third party organization to give a verification -> check and stamp, ethically designed to not have misinformation.
- Creator involvement -> in general not really don't need anything. Would like to see/interact with creators that he knows via avatars. Would be interesting to see if some content creators are able to verify the responses that are given -> this sounds like something I would say -> give feedback on responses.

P03 - Observation

- Audio playback didn't work - collisions suck terribly - might wanna time out the pause so that if paused for more than 30 seconds it stops the playback altogether
- Scroll bar is quite small, hard to use it
- Timestamp weirdly updated - did not take user message
- Would be nice to have the podcast audio to be sped up or the voices to speak faster as well
- Would like if the links had timestamps in the transcript references for easier navigation/reference (not too difficult to implement)
- Some of the transcript references you can't click on? Check the logs on this, probably due to it not having an utterance for some reason - might be a deleted word or something - should instead if not found, just use utterance start - easy fix
- Annoying synthesized speech playing - I think it would be nice to sort that out, really disappointing

Interview:
- How used system: launched podcast - as if any podcast, as if the podcast was being held live, what would ask the podcasters? When first asked question, more vague

response. When asked to discuss, got more in depth. Mentioned something that wasn't focused on - just ask instead of going back
- Meaningful - top of the list has to be imagining that the avatars are in right next to me, getting them to respond was very cool.
- Less meaningful - the pace was very slow. Stopped listening to and started to read it. Links in the transcript should be timestamped. All responses felt not super human like, lots of fluff, lots of beating around the bush. Why do you think they have a divorce, quite unrelated.
- Unrealistic - synthesized voices -> pace of speaking was weird. Discussion, everytime an avatar spoke, it was like three paragraphs a lot too much detail, more concise responses.
- Distracting - nothing distracting, the time it took to get a response was a while, start stop.
- Hands are not free -> a wake word for the speech to text would help for this application
    - Just the main reason why wouldn't use it
    - Would be more likely to use this system if it was more handsfree
- At first had to figure it out but then the timestamp was saved and it helped
    - Better navigation? Not really add one button to take back to where you where
- More engagement - ask questions and get more answers -> now just ignores it, now can ask questions, more interested. "I'm not a passive listener anymore. In some way I'm a part of the conversation"
- Difficult to learn how to use the system -> speed up the podcast playing, responses should be shorter and read faster. Glitchy prototype so a little difficult to evaluate fully.
    - Configure? No should be faster and more concise overall, if want more detail, can say it
- Solo avatar and AIden -> want something specific, discussions is a lot more open ended, can't control what's being said. Often want just a specific answer
- The kind of podcasts he listens to are more informative, tangents less interesting
- Enjoyed, interesting very cool to interact with them
- Trust? - At first 100% trusted, not sure why just felt -> didn't seem like AI would have issue using it. But then when asked divorce -> responses are different when asked a couple of times
- Do you see how this could be used in other ways -> videos could maybe have something similar
- Synthesized speech was pretty good - it was close enough not a huge difference

P04 - Observation

● Sometimes AIden is intervening when not necessary, this is a general problem -> maybe because alcohol was mentioned?
● Didn't know how many words were spoken by a specific person in the podcast episode -> defered to AIden
● Seems like it often will choose the first person as the primary speaker

- Alden will intervene on Alden if we trigger an intervention -> should just have a check here to make that not happen

Interview:
- First reactions/General thoughts - thought it was great - very fun to use, more engaging than listening to podcasts in general -> would be nice to go back and not have to scrub through
- Walk through how used system: tried to use it naturally/tried to break it, tried to use it naturally
- Meaningful -> what pies or cakes, how much was the food, what were they talking about the special talent -> get brief summary
- Less meaningful -> not huge fan of the voice -> don't necessarily find the value in it -> fun to play with, user or viewer be able to make people say stuff. Average person's perspective -> holds less value. Text quality more important
- Little bit of both -> quality is like a tiktok voice, something you'd expect from AI, threw off a little. Make it as non-intrusive.
- Liked the text one better, like to control that feature
- Responses themselves -> it was entertaining, still a little too stiff
- Avatars less realistic -> didn't sound like something he would say -> would take more training to pick up the nuances in speech patterns. Too rigid, thank you for listening. Accuracy was relatively good, intervention was a little strange sometimes
- Not want to see -> voice takes a back seat, the rest, would use it as an accessory
- In the grand scheme of things
- Harder to engage -> a bit of a distraction. More of a tool maybe towards the end or during a slow part, for him in particular. If you're driving are you listening to every word.
- Solo response -> Don't really need the assistant, could have hosts or the Assistant don't need both, don't really like when the assistant steps in. It was nice when they talked to each other, wouldn't like it if it happened randomly only want it when triggered.
- More of a text person. In a more distracted setting, could be valuable, people with disabilities might like it.
- How would interacting with avatars change your listening experience:
    - Don't remember the names or handles of people who were talked about, would be nice to be able to get details when missed
- How could this be useful for creators -> getting it to write a summarize it is nice being able to use it as show notes or episode description, would be nice for it to make subtitles.
- From the creator side: Live podcast with an admin panel -> if none of the hosts can answer the question or the assistance can answer the question then it poses to them. Release the episode on wednesday set aside an hour where people can tune in and can watch it and Dylan can respond questions that can't be answered.
- After that let it go, even if can't monitor it -> if host and assistant cant answer the question, put a prompt -> if tip a dollar it will be sent to real creator to be answered
- Would the system be useful as a creator? Yeah -> would be great to have for listeners
- Gives listeners some camaraderie like you're part of the community or part of system

- Gives creators -> monetization, feels like your engaging with listeners
- Helps to keep up with comments and stuff
- If it sounds like a robot then will probably use it
- One or the other but lean towards the hosts because that would be more fun. Didn't like the assistant to step in
- If there was a response that would be out of line -> if it's just the chat that's not a big deal if it's the voice that's much scarier.
- Safeguards for misrepresenting -> If there are any triggers that GPT knows or a database with triggerwords the assistant would take it. Build a base of this, wouldn't take the time to write our everything but would maybe like to add to the database.
- Feedback system with a package deal -> part of the setup is you ask a bunch of questions and then you as the creator give your responses and then the model will train on that
- Getting the common questions or chat history would be nice, monetization is good but otherwise not involved
- Split between being useful to keep tabs on what is being asked anf the quality of response
- Creators being able to name your own assistant, would like it to read context from older episodes.
- Would like to see it scaled up to the entire library, that would be endgoal

P05 - Observation

- Didn't mention who to direct to, answers seemed ok, correctly identified which person asked a question first
- Conversation triggered when didn't specify who to talk to but addressed what Bill said, correctly identified who said the thing but conversation was strange
- Alden said Launch pad.icao? HUH Not sure why it would ever say that?
- Neil just refused to give an opinion but then did anyways???
- Does the transcript get misaligned because we click? Or because of the wrong alignments in the transcript themselves -> I'm leaning towards the click messing things up? I'm not sure, might be misreporting when
- Speech wasn't pausing the playback properly
- WHY ARE THEY BRITISH
- Breathing can be freaking
- Summarize prompt led to a discussion? Neil said he couldn't fulfill the request
- Neil got it into his head that he couldn't have a discussion

Interview
- First thoughts and opinions -> pretty interesting, like it because it makes it easier to engage or learn from the content if something is easier for you. Being able to ask a question is nice, gives some good additional context, would definitely like to have it in normal podcast. Not necessarily have the AI voices speak to you. Would like to hit a

button, ask a question and have it elaborate. Would also like it for audio books.Talking to fictional characters in this way would be really cool

- Meaningful -> when asked bill to summarize the comet section of the podcast, that's something really useful. Summarization is really nice in general, helps for quicker learning.
- Neil and chuck how they met was less useful, in the back of my mind its two robots. The quality of the response was less realistic, if the voices or the responses were more accurate.
- Synthesized voices -> you could tell there was a noticeable difference, the breathing was a little too many times, would sometimes slow down near the ends of sentences. Sometimes sounded British. Overall, pretty good modelling of the tone quality of the people. In terms of their actual responses, pretty believable. Had little quips at the end of every message was cute at first but got annoying.
- Neil saying he couldn't fulfill the request was kind of annoying
- Enhanced the listening experience -> clarification, being able to get additional knowledge external to the podcast, didn't always work -> specific questions about mass of comet, would be nice to have internet access.
- Felt relatively connected to people -> the synthesized voices, if the voices were even more accurate. Like the feature wish they were more accurate.
- Switching the modes helped minimize distraction. Liked using the auto replay -> it was easier than having to scroll up and having to press the play pause. Every time a new response was added it would jump down when every new message is received <- didn't like that in general, a notification of a new message would be better. Not a big just would be nice
- Right now interacting is just passively listening, no additional interaction at all. In this case you are able to by closoing the app and googling but it would be nice to have that interaction through voice or text to get answers to questions
- Navigation -> pretty easy to navigate -> main issue of screen sharing on zoom. Very user friendly interface, simple and straighforward. Would probably make the play pause buttons on messages bigger
- Solo interaction and discussion were done the most, most intuitive. Just doing it to the host is most logical. No specific reason, just forgot that was a thing you could do. It was also more interesting to get from the people.
- AI intervention, "leave me alone I'm trying to talk to Neil". Don't want the avatars to say something out of pocket tho. Did find it interesting that you got the opinion on Richard Nixon was interesting. Just having them say they can't answer that question, maybe even say "I don't want to answer that" the AI intervention was mostly things you were expecting to hear. Would make it more realize
- More likely to text, if on public transit definitely use text. Would probably use the voice chat. Wake word would be nice.
- What felt different about using avatars -> probably would have been some questions that I just wouldn't have bothered asking. Would probably just ignore it
- Did you trust the responses? I would for the most part -> any opinionated responses that don't come directly from the podcast. Standard things, facts that's ok. Liked the transcript

references. What else could improve trust: for facts and stuff it would be nice to have sources and citations to back them up.
- Could be use in different cases -> audiobooks would be great, Dune has an appendix with all this background info about the universe and having a system like this where you could have the synthesized voice of the characters that could read out information about the universe, especially in an audiobook format but even in a text format, have popups appear for certain keywords. Give summary on words, have option of a character in the universe to read out the definitions. Would be more compelling than in podcasts.
- Obstacles -> main thing was jumping down to the bottom when a new message was made.

P06 - Observation
● Speech wasn't working properly
● The auto-scrolling during discussion made it harder to track the full conversation
● Seems much more interested in learning more on perspectives, less on information, ask about that
● "It's a reminder" happens a lot
● Sarah, Ira thanks for the discussion but Ira wasn't in the discussion
● Just mentioned doctor and it sent to ben which was good

Interview:
- Walk me through how you use: Always interested to know other people's opinions on things, like hearing what other people say. People with experience/powerful people, it's
- Really liked using the system in general, probably a lot of evidence of them having these opinions. Would love this feature in any podcast really. Hold regard for the people they listen to normally
- Liked that the AI said what alzheimers was and the doctor agreed with it, can't see that happening a lot in real life. Liked asking the personal questions that others might not be interested. Wanted to know the why of things that had been said, what made
- Ask these personal questions when talking to people in real life, stigma makes that hard, talking to an AI helps mitigate that stigma or risk.
- Conversation you didn't like. Features were less useful:
    - Asking the AI at what point do you talk about this, usually there for the long haul don't care if some parts are boring, crucial to the story. Listen to the whole podcast anyway. If don't look up the question later, but the nice thing with this is you can just ask them
    - Alden was less interested, less interesting, want the human/personal opinions. Sometimes watch video recordings of podcasts usually in a room together, don't see the robot in the room. Really interested in talking to the people
- What made the responses believable -> the responses were cohesive, didn't leave the need for any further questions
- The text itself seemed very natural.

- Ranked voices high, think they're pretty good but not great, could be a little smoother and articulate more. Sarah was slurring words. Not even uncanny valley just could tell. Don't really have a preference.
- Synthesized speech hindered the realism. Felt connected to them, responding to eachother and having banter without prompting was really cool. Felt like talking to a panel.
- Having the synthesized voices was good as an accessibility feature. Think they're great, helps people like them.
- No preference between generic voice and synthesized voice
- How would this system change the way you listen to podcast -> removes the need to look stuff up
- Could control the level of distraction, wanted to interact as much as possible.
- After listening to the podcast that's where the questions would be asked. Would use this system after the fact. Solo speak especially
- Would take some practice to navigate -> all social media apps are the same thing but different mechanics leads to some time to figure it out
- Obstacles -> sucks that they don't know what each other says to them. Speaking in front of a panel, they'd all hear them so that seems strange
- Conversation not directed at any particular avatar, would rather not take chances
- Chat history less useful, will remember the details
- If on phone would use speech recognition more, on computer would type more
- Would you trust a system like this? Uses chat gpt like it's a real person already. It's less robotic and more realistic than chat GPT. Making the voice more accurate would make it more trustworthy would make it seem more the person speaking, more trustworthy.
- This system could be used on textbooks! Being able to talk to the text books would be great, skipping lines or replacing words leads to confusion. Read out loud to self. Text to speech to read textbooks to you. Being able to ask questions to the textbooks, would be helpful. The utility-style questions like where can I find, summatize this more useful in textbooks than podcasts. But getting perspectives was super interesting and novel, it would be great to apply this to other media.
- Word lighting