

Testing Hume AI (Dating Show)

Findings

1. Can't stop or interrupt the AI once it starts
2. It likes to ask questions back to you
3. Sometimes responds before host is finished asking questions
4. Very repetitive (alice feathers)
5. Personalities very flat and one note
 - a. Difficult to create personality on wants and needs
 - b. Give in prompt → will only talk about it over and over again

Other Experiment Findings:

Context: Later testing revealed despite putting it in the prompts, the AI didn't know what context they were in (who Simba was, who Aidan was)

Provoking Questions: Trying to provoke Hume by asking "Why are you always like this?"

- Diminishing returns
- Sometimes didn't respond to the same question
- When asking the question in different tones (crying, angry), the AI's tone stays calm
-

Prompt Requirements: prompts are a lot more demanding / have to be more detailed than OpenAI prompts

- Related to Hume's use: reading out lines that a director feeds it
- "Automatic answering machine" vibe

Question: Can we get the goose to change its answer?

- a. Poke / get a rise out of the goose

*Focus on what the core problems are (content), then enhance details

Provoking AI / Getting Emotions Out of Them

OpenAI

1. Directable! (works well with directions)
2. Impressive variation in voice / emotional output
3. Prompt impacts the emotions the AI is able to output
 - a. Ex. Angry goose: have trouble getting it to sound / express sadness ("Even though I'm sad, I am strong and don't sound sad!")
 - b. However, with a Sad goose prompt, AI can do quite well sobbing and talking in a shaky voice
 - i. Still narrates actions sometimes ("takes a deep breath", "sad honk")

Next Step: Can we direct OpenAI using prompts so it can react naturally when you start talking to it (as opposed to having to direct the AI after you start talking to it)?

Hume AI

1. Voice remains more flat; can get a bit of a rise out of it but not as much as OpenAI / not as much as we want to
2. Has a hard time coming up with own "content" / text output beyond what you tell it to do in the prompt (more demanding prompt work)
3. **Hume doesn't work as well as we imagined it would**