The preliminary build focuses on the most popular strategies tried in combination and then evaluated by an LLM by taking into account the user's context. The best prompt, as well as the best response, is explored.

➔ Prompt strategies:

1. Write clear instructions
   Reducing ambiguity by adding context to questions. Add a system prompt of the demographics of the user to eliminate some ambiguity. Add instructions like "answer consistently" and include a word limit.

2. Provide reference text
   Instruct the model to answer using the reference material only. Ask for citations in the answer.

3. Split complex tasks into simpler subtasks
   Add a summary of the conversation until now for improved context. Process the prompt recursively for long texts. (Use intent classification - a future exploration idea)

4. Give the model time to think
   Add an inner monologue as system dialogue. Ask to solve it itself first before answering yes/no questions when presented with a solution. (Ask if it missed anything on its previous passes - a future exploration idea)

➔ The evaluation of answers is based on a fuzzy criterion. OpenAI Evals provides some existing specifications and is customizable to evaluate for large samples.

DIFFERENCE TO DETECT	SAMPLE SIZE NEEDED FOR 95% CONFIDENCE
30%	~10
10%	~100
3%	~1,000
1%	~10,000

➔ For this current build, we can use a simple LLM prompt with the 2^4 combinations possible to get the best strategy - both for prompts and answers. The prompt is based on user information and demographics and is given to the LLM to create its evaluation criteria and then give the best combination of prompt engineering techniques used.