

Test Plan 3: AI Chatbot Consultation

Goals

The goal of the test is to determine the usability and effectiveness of the AI Chatbot as a health coach, allowing for relevant and helpful responses.

Quantitative Measurement List

- Task Success Rate: How many users receive relevant and useful information based on their queries.
- Single Ease Question Score: A user rating on the difficulty of interacting with the chatbot and understanding the responses.
- AI Latency: How long it takes for the response to be written after the query has been sent.

Scenario

You want to improve your diet to boost your fitness, but you don't know where to start. You ask the Chatbot to provide healthy, nutritious meals based on what you already eat and websites for more information to read-up on afterwards.

Follow-up Scenario: You want the provided information to be saved into your notebook within the app.

Task List

1. Open the Chatbot and ask "What are healthier alternatives for the meals I already eat? *LIST MEALS*".
2. Ask the chatbot "What websites did you get these meals from?".
3. Assess the information given and click on some of the given links.
4. Click the save buttons and write a title for the notes page.

Qualitative Measurement List

- Does the user think that the chatbot feels like a health coach?
- Does the user find the output useful and clear?
- Is the user confident with the AI's advice?

Potential Observations of Users

- Users may submit difficult and long queries, confusing the AI.
- The user may become impatient while waiting for a reply.

Post Scenario Interview

<https://forms.gle/ZWA9yPjVCCeHngau5>

How easy was it to understand the interface for the AI Chatbot? *

1

2

3

4

5

Very Difficult

☐

☐

☐

☐

☐

Very Easy

How easy was it to find the notebook function? *

1

2

3

4

5

Very Difficult

☐

☐

☐

☐

☐

Very Easy

How easy was the notebook interface to understand? *

1

2

3

4

5

Very Difficult

☐

☐

☐

☐

☐

Very Easy

When clicking the save button, did you expect the *entire* conversation to be saved, or just that specific list of meals? *

Short answer text

What information do you feel was missing or hard to understand?
If not, leave it blank.

Short answer text

Any additional comments?

Short answer text

Test Set-up Details

There will be 2 people attempting to use the paper low-fidelity prototype. The participant will have paper copies of the screen in front of them, pretending to go through the different screens to get the required outcome; they will be led through the correct screens at the correct time. A stopwatch will be used to assess the time it takes the user to complete the task. The questionnaire will be completed straight after the scenario is completed.