Validate AI

More transparent, accurate AI responses

Sprinternship Project - Lexie Leong

The Problem

Univ. of Michigan study reports...

92%

older adults want to know whether content comes from an AI or a human

46%

older adults have little to no trust in AI-generated information

Our seniors expressed concerns with

- the accuracy of AI-generated material
- the “yes-man” attitude most AI are programmed to be
- the authenticity and credibility of source of AI responses

My project:

A Chrome Extension to validate AI

Three layers of verification

1 Source validation

- The URLs provided by chatbots will be automatically checked for existence via a HTTP fetch request
- URLs will be rated in security, trustworthiness and the kind of information it is (e.g. governmental,

2 Prompt Engineering

- Inject a system prompt for AI responses that ...
  - provides sources and citations
  - is rigorous and impartial
  - is short and concise
  - is simple and friendly for older adults
  - is accurate

3 Google Fact Check API

- Look into Google Fact Check database for similar/identical claims that have been proven/debunked by a fact-check or news organization

Problem Solved

Biased, false and agreeing AI responses

Authenticity and credibility of AI responses

- Google Fact Check API cross checks responses with actual humans.
- Prompt engineering encourages ChatGPT to be rigorous and impartial in its response.
- Source verification gives the power back to the user.
  - Transparency: Users can see who is speaking (Reuters vs. a random blog).
  - Bias detection: Users can see bias from sources or quotes used.
  - Error checking: Check if AI is hallucinating or misinterpreting sources, and allow users to

Problem Solved

The accuracy of AI-generated material

- Prompt engineering forces ChatGPT to use sources and quotes when responding. as well as double checking.
- Source verification validates the existence of sources and verifies the credibility of them.
- Fact-check API cross checks AI claims with third-party expert evaluation.

How It’s Built

01 Tech Stack

JavaScript
HTML
CSS
Figma
JSON
Chrome Extension API
React
CANVA

02 Process

1. Understand what seniors need by talking to them
2. Creating a proposal
3. Create the two features and debug
4. User test, receiving feedback and make changes
5. TADA!

How It’s Built

03 Collaboration

- I presented my project to Lois, Janelle, Leann, Alice, Rich and Mary, and Charles.
- They all liked the project, found it useful, and were super supportive
- Gave me new insights and perspectives for the project
- Thanks so much for breaking it with your amazing prompts!

04 Feedback

- It is useful
- The popup is simple, but the extension is hard to install or find on the browser without help
- Different sources have different internal credibility for people (e.g. government)
- Would be good if depth of a source is evaluated as well
- The font size can be bigger

Salford

Presented by :
Ketut Susilo, CEO

Demo

Next Steps

- Expand it to multiple platforms and chatbots
- Publish it in the Chrome Extensions Store
- Implement the third feature
- Evaluate the depth/bias of the sources