# Problem-Solving with an AI Assistant: A Usability Test

Carson Rivera, Benhur Ravuri, and Dr. Jonathan Adams

Florida State University College of Communication and Information

## Background

As artificial intelligence becomes more prevalent with improvements in software and hardware, humans will need to take on an editor role as opposed to being a master of a skill. The domain of higher education is the pivotal stepping-stone in implementing this shift. This study aims to assess people's current abilities in problem-solving with AI and provide a basis for how we can teach future generations to be better engineers of large language models.

## Methods

- 7 volunteer-participants were asked to complete a set of tasks using the Monica AI Chatbot
- The participants were equipped with a prompting guide to produce better answers from the Chatbot
- The participants were asked to "think-aloud" while working with the Chatbot to gain an understanding of their live thought processes
- Live notes and chat logs were analyzed drawing trends and themes on how users generally used the Chatbot to solve problems

## Task
### Write A Haiku



## Definitions

Large Language Models - Artificial intelligence systems that can process and generate text with coherent communication and generalize to multiple tasks

## Themes

### Disregard of Prompting Guide
- Most of the participants opted not to use the prompting guide given to them. There appears to be a disconnect with how these unorthodox strategies can be used to optimize large language models.

### Disconnect with Initial Approach
- Many participants didn't know how to approach questions or how to get started. Most opted to simply copy the tasks word-for-word.

### Exploratory Behavior
- Participants often explored capabilities of the Chatbot outside of their designated tasks.

### Intuitive and Human-like Behavior
- Rather than some of the more rigid prompting strategies offered in the guide (e.g., role, task, feature), participants talked their way through problems as if they had a human assistant they could negotiate, correct, and clarify with.

### Task Satisfaction
- Task satisfaction varied greatly throughout testing. Participants generally fell into 3 groups:
  - Satisfied by Monica's answer with little verification
  - Not satisfied by answer and continued prompting
  - Satisfied with answer but continued prompting to test further capabilities

## Prompting Guide

1. **Ask directly.** Some requests given to an AI are relatively straightforward if you do not have any special preferences

**Example:**

*Please create a 3-day travel itinerary for a weekend trip to the Bahamas. Please include flights that leave for the Bahamas in the morning, and depart the Bahamas for home in the afternoon.*

2. **Sometimes, a Chat GPT needs directions or one-line helpful hints.**

**Example:**

Please write a Haiku
vs.
Please write a Haiku (*about going to college*)

**Accuracy Example:**

A juggler can juggle 16 balls. Half of the balls are Golf balls, half of the Golf balls are blue. How many blue golf balls are there?

Let's think step by step, be sure to check your answer

3. **More complicated examples require structured prompts. Use this format**

Role: (a profession, for example Role: teacher)
Task: (what you want the agent to do)
Features: (the qualities you want in the response)

**Example:**

*Role:* You are an experienced travel agent
*Task:* Please create a 3 day travel itinerary for a weekend trip to the Bahamas
*Features:*
Depart home in the morning
Arrive home in the evening
Please plan one day at the beach
Please plan a snorkel diving expedition

## References

Adams, J., Roberts, A., Obermaier, L, et al. (2023). Uses of Artificial Intelligence in Higher Education. EdMedia + Innovate Learning, Jul 10, 2023 in Vienna, Austria ISBN 978-1-939797-71-1 Publisher: Association for the Advancement of Computing in Education (AACE), Waynesville, NC

Dumas, J. S., Redish, J.C. (1999). A Practical Guide to Usability Testing Revised, Subsequent Edition. Intellect Ltd; Revised, Subsequent edition. ISBN-13 : 978-1841500201

Naveed, H., Khan, A. U., Qiu, S., Saqib, M., Anwar, S., Usman, M., Akhtar, N., Barnes, N., & Mian, A. (2024, February 20). A Comprehensive Overview of Large Language Models. arXiv.org. https://arxiv.org/abs/2307.06435