Task Supportive and Personalized Human-Large Language Model Interaction:

A User Study

The UNIVERSITY of OKLAHOMA
OUHCIR

Ben Wang *benw@ou.edu*
Jiqun Liu *jiqunliu@ou.edu*
Jamshed Karimnazarov *jamshed.k@ou.edu*

1 Background

- ChatGPT's release has generated significant interest in **human-AI interaction**.
- Various purposes rise: **task assistance, entertainment, education, search engine alternative**.

2 Unsolved Problems

- Users have challenges in **initializing and refining prompts**
- Users have **cognitive barriers** and **biased perceptions** in the usefulness of ChatGPT's output and the task completion

3 This work investigate:

how **background information**

- Task topic and type

and user cognitive aspects

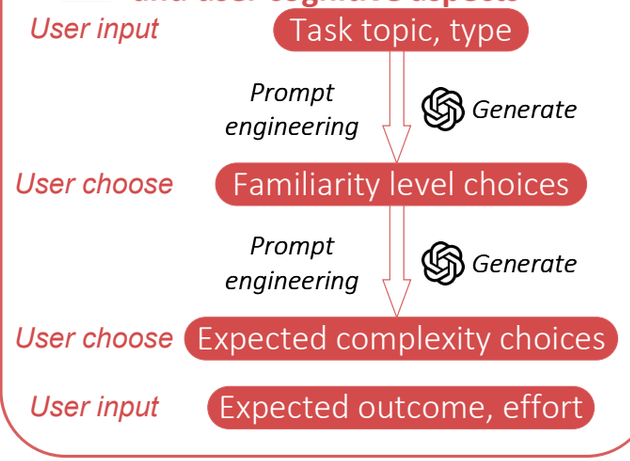
- Task familiarity
- Expectations of task complexity, outcomes, and effort)

affect user interactions in information-seeking and problem-solving tasks with LLMs.

4 Participants

College students in diverse background: Computer Science, Library and Information Science, Education, Psychology, Health ...

5 Collecting task information and user cognitive aspects



6

Users conduct their own tasks using ChatGPT with suggestions:

How can ChatGPT help you perform the task by

initializing the conversation, or predicting your next question, or revising your prompt, or adapting its answer…

Prompt engineering

7 Evaluation

Prompt/output annotation

- Usefulness, credibility, ...

Post-task questionnaire

- Satisfaction, challenges, ...

Interview

8 Future work

- User modeling with behavioral data only
- Dynamic task status at the prompt level
- RLHF involving task information and user cognitive aspects
- How does Auto(Task)GPT help users?

Acknowledgment

IIS-2106152

Microsoft for Startups Founders Hub

DATA INSTITUTE FOR SOCIETAL CHALLENGES
The UNIVERSITY of OKLAHOMA

Task Supportive and Personalized Human-Large Language Model Interaction:

A User Study

The UNIVERSITY of OKLAHOMA
OUHCIR

Ben Wang *benw@ou.edu*
Jiqun Liu *jiqunliu@ou.edu*
Jamshed Karimnazarov *jamshed.k@ou.edu*

Motivation

- IIR: Query reformulation in multi-query search sessions.
  - LLM: User challenges in initializing and refining prompts [1,2]
- IIR: User cognitive barriers, biases in information seeking.
  - LLM: User cognitive barriers and biased perceptions when interpreting ChatGPT's output impede task completion [1,3]

This work as a starting point

Inspired by IIR principles, with context of task background and user cognitive aspects,

ChatGPT can better assist users to perform the task

by

1. enhancing task perception through retrospection,
2. offering early interaction guidance,
3. aligning user needs with output,
4. inspiring prompt refinement ideas,
5. delivering outputs better aligned with user intention and elevating task engagement

Potential task-aware evaluation metrics

System

- Task/topic reliability
- Prompt/output (self)explainability

User

- Task engagement
  - *Prompt refining willingness*
  - *Satisfaction*
- Learning outcome

Next steps…

- **User interface/Plugin:** Task-aware LLM interface/plugin *for data collection in a larger scale study*
- **Fining-tuning and RLHF:** Task/topic specific LLM *involving task information and user cognitive aspects*
- **Context prediction without manual annotations:** Dynamic task state prediction *by semantic/behavioral data*
- **Direction:** Auto(Task)GPT or Copilot?

Acknowledgment

IIS-2106152
Microsoft for Startups Founders Hub
DATA INSTITUTE FOR SOCIETAL CHALLENGES The UNIVERSITY of OKLAHOMA

[1] Zamfirescu-Pereira, et. al. (2023). Why Johnny can't prompt. *CHI 23.*
[2] Skjuve, M., et. al. (2023). The User Experience of ChatGPT. *CUI 23.*
[3] Urban, M., et. al. (2023). Can ChatGPT Improve Creative Problem-Solving Performance in University Students?.