

How Far Are We? The Triumphs and Trials of Generative AI in Learning Software Engineering

Choudhuri, Rudrajit, et al. "How Far Are We? The Triumphs and Trials of Generative AI in Learning Software Engineering." *Proceedings of the IEEE/ACM 46th International Conference on Software Engineering*. 2024.

PROBLEM BEING SOLVED

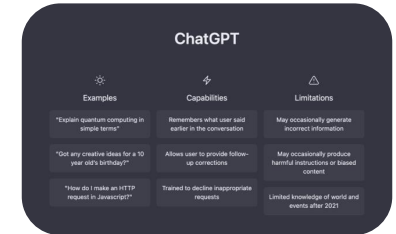
Prior work studies GenAI in the context of Introductory Computer Science

Prior work focuses on using genAI to solve algorithmic problems or improving genAI

Knowledge GAP : how can conversational genAI help in learning advanced CS topics like SE?

RQ1: How effective is convo-genAI in helping students in software engineering tasks?

RQ2: What are the current pitfalls in convo-genAI?



SOLUTION

RQ1: How effective is convo-genAI in helping students in software engineering tasks?



Tasks 1 Debugging: Logical programming, API usage, Web Scrapping

Tasks 2: Removing Code Smells

Tasks 3: Commit code to the remote branch and create a pull request to the base branch

Task performance and Self-efficacy

RQ2: What are the current pitfalls in convo-genAI?

After Action Review for AI (AAR/AI)

Quantitative Analysis of AAR/AI reveals 5 fault categories + 7 consequence categories

Microsoft's design guidelines for Human AI Interaction

Participant perception of violations reveal pitfalls

Study shows increased frustration, uncertainty, and induced self-doubt for Treatment Group

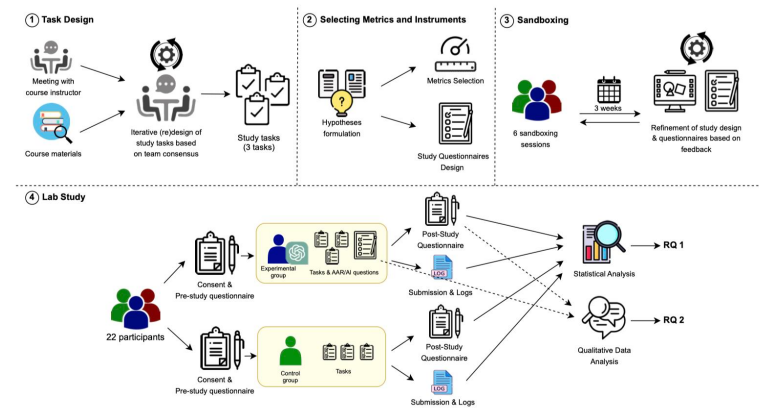


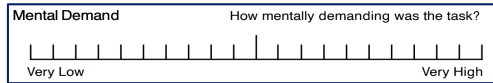
Figure 1: Overview of the research design

POSITIVE POINT: Well Structured and Detailed Methodology

RQ1: How effective is convo-genAI in helping students in software engineering tasks?

H1: Students using ChatGPT for the study tasks perceive lower cognitive load than students using alternate resources

Use Original NASA Task Load Index (TLX): mental, physical, and temporal demand, performance, effort, and frustration



H2: ChatGPT positively impacts students' productivity

Time boxed task, evaluate correctness via blind grading

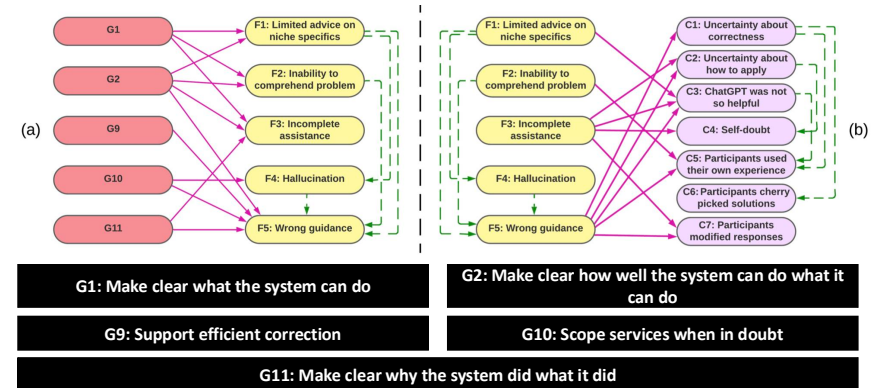
H3: ChatGPT promotes students' self-efficacy

Self-efficacy questionnaire + Users' continuance intention

RQ2: What are the current pitfalls in convo-genAI?

AAR/AI + Microsoft HAI guidelines

POSITIVE POINT: Analysis of ChatGPT's faults and Consequences



POSITIVE POINT: Sound Recommendations for future genAI

Use of a scaffolding learning agent to ensure that answers are not directly given

Incorporating templates, heuristics or human intervention to clarify AI's problem-solving process

Incorporate adherence of AI to HAI guidelines through an iterative approach

Different learning styles when it comes to specific genders – make AI inclusive of both learning styles

NEGATIVE POINTS

NEGATIVE POINT: Python tasks used as a representative tasks for Software Engineering

Negative POINT: Sample size of 22 students may limit generalizability + students from different courses may have different expertise

Negative POINT: Evaluation in a single session

FUTURE WORK

Develop a genAI model that captures past interactions, mistakes and problem-solving strategies of the student
-> use these to gauge expertise level and understand learning strategy of the student -> Treatment Group

Conduct the experiment in a co-op setting, partnered with diverse tech companies over a whole co-op term -> gather statistical and qualitative insights

Conduct the experiment with students with 2 sections of the same SE course. Course could contain weekly lab components

DISCUSSION POINTS

Which pitfalls (limited advice, inability to comprehend problems, incomplete assistance, hallucination, wrong guidance) have you personally experienced while using ChatGPT for SE related tasks?

Which particular SE tasks does ChatGPT help with better than traditional sources (in your experience)?

Does using genAI in SE adversely impact a student's independent debugging capabilities?

If an experiment is conducted with a group of student that can use a mixture of ChatGPT + traditional resources -> how do you expect this group to perform?

RATING

4.5 / 5

A very well detailed investigation!

