

ChatGPT for Learning HCI Techniques: A Case Study on Interviews for Personas

Apresentado por: Simone de Oliveira Santos
Programa de Pós-Graduação em Ciência da Computação



- Publicado no IEEE Transactions on Learning Technologies
- Qualis A1
- Ano de publicação 2024

ChatGPT for Learning HCI Techniques: A Case Study on Interviews for Personas

Jose Barambones , Cristian Moral , Angélica de Antonio , Ricardo Imbert , Loïc Martínez- Normand ,
and Elena Villalba-Mora 

Motivação

- Antes de interagir com usuários reais, desenvolvedores devem entender de IHC para **não desperdiçar a paciência e disponibilidade dos usuários**.
- Especialistas no campos de IHC passam por **treinamento** em várias técnicas.
- Necessidade de analisar usuários em potencial para os sistemas que são produzidos.

Motivação

- Conduzir entrevistas **é desafiador**.
- Recrutar participantes, definir questões efetivas para obter respostas úteis é imperativo.
- Entrevistas de baixa qualidade comprometem a análise.
- Técnicas como a criação de **Personas** somente funciona através de uma **boa coleta de dados e análise do contexto de uso**.

Motivação

- O surgimento das IA generativas e o seu uso como chatbots treinados (ChatGPT) oferece uma oportunidade uso para aprendizado.
- Estudos de caso recentes:
 - aquisição de habilidades pedagógicas
 - efeitos na motivação de estudantes e desempenho na aprendizagem
 - sala de aula invertida

Lacuna identificada

Uso de ChatGPT para o aprendizado de técnicas para o design centrado no usuário.

Técnica alvo: Criação de **Personas**

Personas

No contexto de IHC, a criação de Personas é uma técnica usada para representar usuários reais e potenciais usuários para aplicações.

Uma persona é:

- fictícia
- uma representação realística de um usuário em potencial
- exemplifica atitudes associadas ao contexto dos usuários em relação a um produto específico.

Problema

No contexto educacional é difícil treinar estudantes para conduzir atividades de IHC em contexto real.

Fatores limitantes:

- acesso limitado à usuários reais
- limitações de tempo
- **falta de prática e habilidades em entrevistas**

Problema

Para melhor entender a percepção dos problemas enfrentados, os pesquisadores conduziram um survey com estudantes de graduação.

Achados:

- Dificuldades de recrutamento
- Falta de variabilidade de usuários
- Definir questões de qualidade

Objetivo

O objetivo da pesquisa é estudar e validar se o ChatGPT, propriamente configurado e parametrizado, pode gerar informações fictícias mas de boa qualidade para ser usado como material de treinamento em IHC.

Proposta

A questão central da pesquisa é se um chatbot baseado em LLM pode servir como uma **ferramenta educacional de treinamento** para atividades de IHC.

Hipótese levantada

ChatGPT pode melhorar a experiência educacional pela simulação de interações de conversa, e conseqüentemente, ajudar estudantes a aprender.

Questões de pesquisa

- 1 O ChatGPT pode representar um usuário fictício e crível através de uma configuração apropriada e um prompt único?
- 2 É possível usar esses usuários fictícios para gerar um conjunto de entrevistas de alta qualidade que sirva como entrada válida para criar personas realísticas?

Proposta

Principal requisito: que os resultados das entrevistas pudessem ser comparáveis, assim, foi usada a entrevista estruturada.

Principal desafio: operacionalizar o conceitos de “usuário crível” e “entrevistas de qualidade”.

Proposta

Como cenário de experimentação: foi usado o contexto e os dados de um projeto real que objetivava melhorar a intranet da universidade.

Este projeto foi desenvolvido de acordo com o processo centrado no usuário com criação de 8 personas.

Metodologia

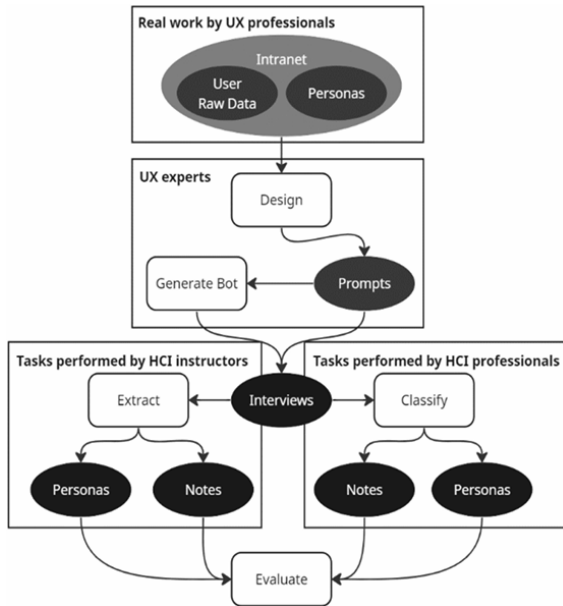
A pesquisa foi estruturada com 4 grupos distintos:

Supervisor Coordenador dos times

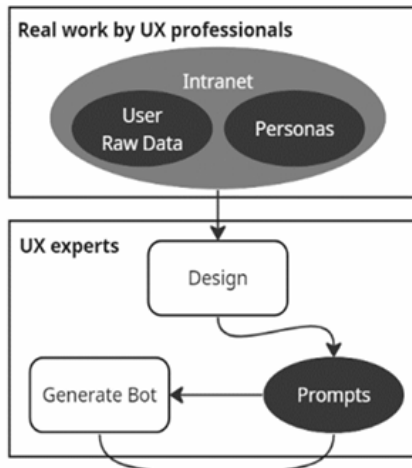
Especialistas UX Criam os prompts e entrevistas baseado no projeto original. Foram criados 40 usuários fictícios.

Profissionais UX Conhecedores do projeto original.

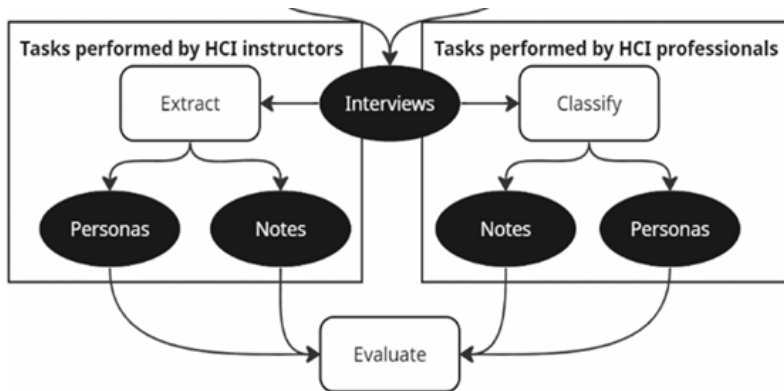
Professores IHC Avaliam a qualidade do material e criam personas baseado nas entrevistas geradas.



Metodologia



Metodologia



O objetivo da análise foi determinar se as respostas geradas pelo ChatGPT poderia ser **corretamente associadas com as personas originais** e se foram **informativas o suficiente** para apoiar a criação de novas personas.

Medida de sucesso

Recapitulando as questões de pesquisa:

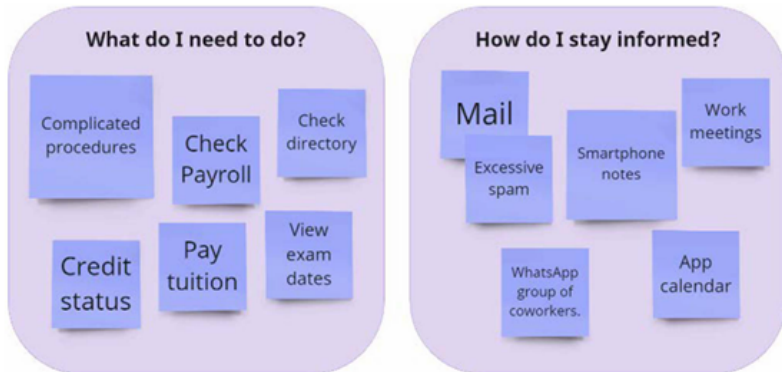
RQ1 O ChatGPT pode representar um usuário fictício e crível através de uma configuração apropriada e um prompt único?

RQ2 É possível usar esses usuários fictícios para gerar um conjunto de entrevistas de alta qualidade que sirva como entrada válida para criar personas realísticas?

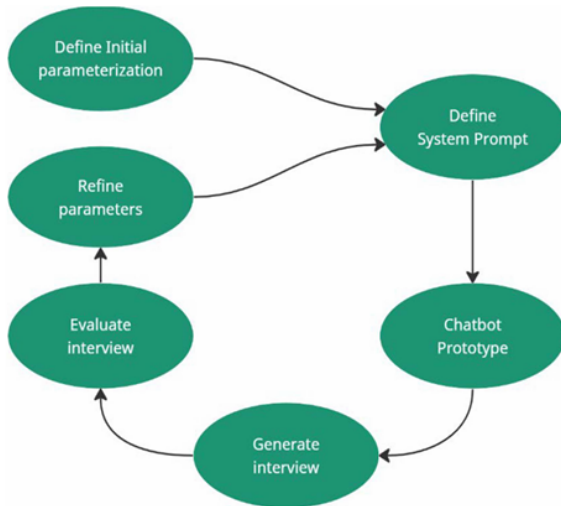
Medida de sucesso

- RQ1** Os profissionais classificam as respostas geradas com relação às 8 personas de referência, se os usuários fictícios são críveis.
- RQ2** A qualidade do material foi medida de duas formas:
- Profissionais** a qualidade do material do ponto de vista técnico, se representa o domínio do projeto original.
 - Professores** a qualidade do material do ponto de vista educacional, se é útil para criar personas.

Template para criação de personas



Teste de operação do ChatGPT



Proposta de prompt

- 1) Tell ChatGPT who the target person to be simulated is (using a general definition like a university intranet user or a football team coach).
- 2) Ask ChatGPT to answer a series of questions for a study, applying a specific configuration of the model (in ChatGPT 4, this is specified using brackets).
- 3) Ask ChatGPT to act as a person with specific characteristics (name, age, gender, educational level, and type of language spoken).
- 4) Ask ChatGPT not to reveal that it is a virtual assistant during the conversation.
- 5) Ask ChatGPT not to use enumerations in its responses.

Medida de sucesso

- Ao final do procedimento, os profissionais e professores devem responder a um questionário sobre a qualidade das entrevistas.
- Questões sobre formato da entrevista, informações relevantes, coerência das respostas, utilidade educacional, grau de simulação e ética.

Resultados - RQ1 “believability”

- De forma geral, os itens da entrevista **foram melhor avaliados pelos profissionais** do que pelos professores.
- Os profissionais já tinham conhecimento das personas do projeto original e apenas indicariam se a entrevista representava bem.
- Os professores precisavam das entrevistas para criar personas e foram mais críticos.

Resultados - RQ1 “believability”

Principais achados da avaliação qualitativa:

- Respostas repetitivas, superficiais, e não elaboravam as respostas.
- Inconsistência na linguagem: muito boa para um estudante internacional, por exemplo, e inconsistência de gênero (Espanhol).
- Informações incompletas para criação de personas.

Resultados - RQ1 “believability”

- Apesar dos problemas, **a resposta à RQ1 foi positiva.**
- Considerando cada entrevista separadamente, ambos **consideraram os usuários fictícios críveis** para o contexto do sistema.

Resultados - RQ2 “quality”

- Os profissionais **relacionaram todas as entrevistas** com uma persona do projeto original, validando o processo de criação de prompts para gerar entrevistas com informações úteis.
- Os professores **foram capazes de gerar personas**, com exceção de uma, que casavam com as originais.

Resultados - RQ2 “quality”

- A resposta à pergunta foi considerada **positiva**.
- A maior limitação considerada foi a repetição do conteúdo devido à limitação de respostas do chat para uma melhor consistência e falta de elaboração das respostas.

Discussões - Limitações

- O prompt fornecia algumas alternativas para o chat assumir na criação de usuário, e o chat não elaborava na escolha e não randomizada nas escolhas, repetindo escolhas.
- O chat tende a escolher o gênero masculino (39 de 40).
- Possível viés dos professores que conhecem o tópico do contexto e conhecem os papéis do projeto.

Discussões - Recomendações

- Atenção ao trabalhar com mais de uma **versão do ChatGPT** ou similares pois podem apresentar comportamentos diferentes.
- Em entrevistas, **simplificar e reduzir o número de perguntas**, pois ele pode juntar todas as respostas em uma só e depois repetia as respostas nas demais perguntas.
- **Não há garantia de diversidade** quando são fornecidas escolhas. É preciso explicitamente indicar variedade.

Conclusões - Lições aprendidas

- Entendimento do funcionamento de um chatbot baseado em LLM para fornecer boas instruções.
- Fornecer flexibilidade suficiente para evitar respostas repetidas. Tentativa e erro foi a forma encontrada para resolver a questão.

Passo interessante para integrar tecnologias baseadas em LLM no **treinamento de técnicas de IHC**, sem intenção de substituir os usuários reais.

ChatGPT for Learning HCI Techniques: A Case Study on Interviews for Personas

Apresentado por: Simone de Oliveira Santos
Programa de Pós-Graduação em Ciência da Computação

