



**UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS QUIXADÁ
CURSO DE GRADUAÇÃO EM SISTEMAS DE INFORMAÇÃO**

ANA BEATRIZ SILVA LINHARES

AVALIAÇÃO DE USABILIDADE DO CHATGPT

QUIXADÁ

2025

ANA BEATRIZ SILVA LINHARES

AVALIAÇÃO DE USABILIDADE DO CHATGPT

Trabalho de Conclusão de Curso apresentado ao
Curso de Graduação em Sistemas de Informação
do Campus Quixadá da Universidade Federal
do Ceará, como requisito parcial à obtenção do
grau de bacharel em Sistemas de Informação.

Orientadora: Profa. Dra. Andréia Libório Sampaio

QUIXADÁ

2025

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Sistema de Bibliotecas

Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

L728a Linhares, Ana Beatriz Silva.

Avaliação de Usabilidade do ChatGPT / Ana Beatriz Silva Linhares. – 2025.

98 f. : il. color.

Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Campus de Quixadá,
Curso de Sistemas de Informação, Quixadá, 2025.

Orientação: Profa. Dra. Andréia Libório Sampáio.

1. Usabilidade. 2. ChatGPT. 3. Interação usuário-IA. 4. Avaliação Heurística. 5. Teste de Usabilidade. I.
Título.

CDD 005

ANA BEATRIZ SILVA LINHARES

AVALIAÇÃO DE USABILIDADE DO CHATGPT

Trabalho de Conclusão de Curso apresentado ao
Curso de Graduação em Sistemas de Informação
do Campus Quixadá da Universidade Federal
do Ceará, como requisito parcial à obtenção do
grau de bacharel em Sistemas de Informação.

Aprovada em: 06 de Março de 2025

BANCA EXAMINADORA

Profa. Dra. Andréia Libório Sampaio (Orientadora)
Universidade Federal do Ceará (UFC)

Profa. Dra. Ingrid Texeira Monteiro
Universidade Federal do Ceará (UFC)

Profa. Dra. Rainara Maia Carvalho
Universidade Federal do Ceará (UFC)

A Deus, minha família e todas as pessoas que
caminharam comigo até aqui.

AGRADECIMENTOS

A Deus, por me conceder o dom da vida e a força necessária para enfrentar cada desafio desta jornada.

À minha mãe, Solonja, e à minha querida avó, a quem carinhosamente chamo de "Manhê", dedico este trabalho e toda a minha trajetória. Obrigada por cuidarem de mim com tanto amor e dedicação. Tudo o que sou e tudo o que ainda serei é reflexo do carinho e dos valores que vocês me ensinaram.

Mãe, obrigada por ser meu porto seguro, por me oferecer amor e conforto incondicionais, independentemente de como tenha sido os nossos dias. Sua força, apoio e fé em mim foram fundamentais para que eu chegassem até aqui. Caminhar ao seu lado e segurar sua mão, sempre tornou qualquer desafio leve.

Ao meu grande amor, Hugo Patrício, por confiar em mim em todos os momentos e por estar ao meu lado com paciência e cumplicidade. Encontrar você no "meio do caminho" me deu ainda mais força para continuar. Obrigada por ser você e por tornar essa caminhada mais especial. Que para sempre seja você.

Aos meus grandes amigos, Sávila Amâncio, Yuri Lima, Elano Nunes, Marcos Vinícius, Roberto Coutinho, Edoardo Nogueira, Karen Oliveira, Mariana Oliveira, Talita Maciel, Thayná Cavalcante, Alysson Alexandre, Arislayne Souza, Renan Almeida e todos os meus colegas que, com amizade, apoio e muito humor, tornaram esses momentos mais leves. Em especial, aos amigos que me acompanharam de perto durante esse último período na faculdade, obrigada por acreditarem em mim. Obrigada por estarem comigo nas matérias difíceis, nas dúvidas e nas conquistas. Nada teria sido o mesmo sem vocês.

Aos professores que marcaram minha trajetória acadêmica, transmitindo conhecimentos valiosos e oferecendo conselhos que impulsionaram meu crescimento. Em especial, à minha orientadora, Profa. Dra. Andréia Libório, sou imensamente grata por sua paciência, incentivo e dedicação. Apesar do breve contato, sua orientação foi essencial para que este trabalho fosse concluído.

A todos que, de alguma forma, contribuíram para esta jornada, meu mais sincero agradecimento.

“Lembre-se, quando você estiver perdido na escuridão... Procure a luz.”

(The Last of Us)

RESUMO

A usabilidade desempenha um papel essencial na interação humano-computador, sendo um dos pilares para o desenvolvimento de sistemas digitais eficazes, acessíveis e satisfatórios. Este trabalho realiza uma análise detalhada da usabilidade do Chat Generative Pre-trained Transformer (GPT), uma inteligência artificial de conversação desenvolvida para interagir com usuários de maneira natural e eficiente. Utilizando metodologias como avaliação heurística e testes de usabilidade, foram identificadas as principais falhas e pontos fortes da interface, propondo melhorias baseadas nas percepções dos usuários. Os resultados destacam a importância de interfaces claras e intuitivas para a aceitação e o sucesso de sistemas digitais, além de evidenciar a necessidade de um processo contínuo de validação e refinamento. A inclusão de participantes de diferentes idades e áreas de estudo e atuação nas avaliações, mostrou-se essencial para o desenvolvimento dessa pesquisa. Este trabalho conclui que priorizar a usabilidade é fundamental para criar experiências digitais mais satisfatórias, engajadoras e eficientes, servindo como base para futuras pesquisas e aprimoramentos no campo da inteligência artificial conversacional.

Palavras-chave: usabilidade; chatGPT; interação usuário-IA; avaliação heurística; teste de usabilidade.

ABSTRACT

Usability plays an essential role in human-computer interaction, being one of the pillars for the development of effective, accessible and satisfactory digital systems. This work performs a detailed analysis of the usability of Chat GPT, a conversational artificial intelligence developed to interact with users in a natural and efficient way. Using methodologies such as heuristic evaluation and usability testing, the main flaws and strengths of the interface were identified, proposing improvements based on user perceptions. The results highlight the importance of clear and intuitive interfaces for the acceptance and success of digital systems, in addition to highlighting the need for a continuous process of validation and refinement. The inclusion of participants of different ages and areas of study and activity in the evaluations proved to be essential for the development of this research. This work concludes that prioritizing usability is fundamental to create more satisfactory, engaging and efficient digital experiences, serving as a basis for future research and improvements in the field of conversational artificial intelligence.

Keywords: usability; chatGPT; user-AI interaction; heuristic evaluation; usability testing.

LISTA DE FIGURAS

| | |
|--|----|
| Figura 1 – Arquitetura de funcionamento de um chatbot | 20 |
| Figura 2 – Captura de tela do antigo ChatGPT | 22 |
| Figura 3 – Captura de tela do ChatGPT | 23 |
| Figura 4 – Captura de tela do canal Alanzoka | 26 |
| Figura 5 – Símbolos | 27 |
| Figura 6 – Captura de tela Google Docs | 27 |
| Figura 7 – Captura de Tela Gmail | 28 |
| Figura 8 – Captura de tela | 28 |
| Figura 9 – Captura de tela do Canva | 29 |
| Figura 10 – Atalhos do teclado | 29 |
| Figura 11 – Captura de tela Sigaa | 30 |
| Figura 12 – Passo-a-passo da metodologia | 36 |
| Figura 13 – Divisão ChatGPT | 37 |
| Figura 14 – Gráfico de violações por nível de problema | 45 |
| Figura 15 – Número de violações por heurística | 46 |
| Figura 16 – Botão de enviar mensagem | 47 |
| Figura 17 – Prompts pré-escritos | 47 |
| Figura 18 – Botões de criar novo chat | 48 |
| Figura 19 – Botão de dúvidas | 49 |
| Figura 20 – Modal de confirmação de exclusão de chats | 49 |
| Figura 21 – Links de domínios | 50 |
| Figura 22 – Chats arquivados | 50 |
| Figura 23 – Aba de funções do ChatGPT | 51 |
| Figura 24 – Configurações do ChatGPT | 51 |
| Figura 25 – Limpar chat temporário | 52 |
| Figura 26 – Opções de upgrades e chat temporário | 53 |
| Figura 27 – Botões de respostas | 54 |
| Figura 28 – Mensagens ambíguas | 54 |
| Figura 29 – Respostas invisíveis | 55 |
| Figura 30 – Mensagem de dúvida sobre o ChatGPT | 55 |
| Figura 31 – Pontos brancos | 56 |

| | |
|--------------------------------------|----|
| Figura 32 – Mensagem em Inglês GPT | 57 |
| Figura 33 – Visão Global - Meus GPTs | 58 |
| Figura 34 – Nuvem de Palavras | 70 |

LISTA DE TABELAS

| | |
|---|----|
| Tabela 1 – Descrição das Atividades | 41 |
| Tabela 2 – Participantes do Teste de Usabilidade | 59 |
| Tabela 3 – Tarefas e Métricas | 59 |
| Tabela 4 – Taxa de sucesso por tarefa | 61 |
| Tabela 5 – Tempo médio de execução por tarefa | 62 |
| Tabela 6 – Solicitação de ajuda por tarefa | 63 |
| Tabela 7 – Perguntas e Respostas - Entrevista | 64 |
| Tabela 8 – Percepções Qualitativas e Métricas Quantitativas | 71 |
| Tabela 9 – Análise por Usuário – Dificuldades e Desempenho | 72 |
| Tabela 10 – Impacto das Violações de Heurísticas na Satisfação do Usuário | 74 |
| Tabela 11 – Problemas consolidados | 91 |

LISTA DE QUADROS

| | |
|---|----|
| Quadro 1 – Comparação dos trabalhos | 34 |
| Quadro 2 – Severidade do problema | 39 |
| Quadro 3 – Ferramentas usadas | 43 |

LISTA DE ABREVIATURAS E SIGLAS

| | |
|-------|---|
| GPT | Generative Pre-trained Transformer |
| TICs | Tecnologias da Comunicação e Informação |
| IA | Inteligência Artificial |
| IHC | Interação Humano-Computador |
| NLP | Processamento de Linguagem Natural |
| ISO | Organization for Standardization |
| FAQs | Frequently Asked Questions |
| SIDRA | Sistema IBGE de Recuperação Automática |
| IBGE | Instituto Brasileiro de Geografia e Estatística |

SUMÁRIO

| | | |
|--------------|---|-----------|
| 1 | INTRODUÇÃO | 16 |
| 1.1 | Objetivos | 17 |
| 1.1.1 | <i>Objetivo Geral</i> | 17 |
| 1.1.2 | <i>Objetivos Específicos</i> | 17 |
| 2 | FUNDAMENTAÇÃO TEÓRICA | 18 |
| 2.1 | Inteligência Artificial | 18 |
| 2.1.1 | <i>Chatbots</i> | 19 |
| 2.2 | Conhecendo o ChatGPT | 21 |
| 2.2.1 | <i>Interface</i> | 22 |
| 2.2.2 | <i>Funcionalidades do ChatGPT</i> | 23 |
| 2.3 | Por que avaliar, o ChatGPT? | 24 |
| 2.4 | Teste de Usabilidade | 25 |
| 2.5 | Avaliação Heurística | 26 |
| 3 | TRABALHOS RELACIONADOS | 31 |
| 3.1 | Usabilidade do ChatGPT como ferramenta de apoio no ensino de português como língua estrangeira | 31 |
| 3.2 | Usabilidade de um chatbot para pessoas com hipertensão e diabetes | 32 |
| 3.3 | CHARTBOT: UMA INTERFACE CONVERSACIONAL PARA VISUALIZAÇÃO DE DADOS ABERTOS DO IBGE | 33 |
| 3.4 | Comparação entre os trabalhos relacionados e a pesquisa proposta | 34 |
| 4 | METODOLOGIA | 36 |
| 4.1 | Avaliação Heurística | 36 |
| 4.1.1 | <i>Preparação</i> | 36 |
| 4.1.2 | <i>Coleta e Interpretação dos Dados</i> | 39 |
| 4.1.3 | <i>Consolidação dos Resultados</i> | 40 |
| 4.1.4 | <i>Discussão dos Resultados</i> | 40 |
| 4.2 | Teste de Usabilidade | 40 |
| 4.2.1 | <i>Preparação</i> | 40 |
| 4.2.2 | <i>Coleta de Dados</i> | 43 |
| 4.2.3 | <i>Consolidação e Interpretação dos Resultados</i> | 43 |

| | | |
|----------------|---|----|
| 4.2.4 | <i>Análise Coomparativa dos Resultados</i> | 44 |
| 5 | RESULTADOS | 45 |
| 5.1 | Avaliação Heurística | 45 |
| 5.1.1 | <i>Consolidação dos Resultados</i> | 45 |
| 5.1.1.1 | <i>Porção 1 - Prompt Inicial com Sugestões de Prompts</i> | 46 |
| 5.1.1.2 | <i>Porção 2 - Aba Lateral com Histórico e Nomes dos Chats</i> | 47 |
| 5.1.1.3 | <i>Porção 3 - Menu de Ajuda na Lateral Direita</i> | 48 |
| 5.1.1.4 | <i>Porção 4 - Menu Principal com Configurações de Conta, Informações e Logout</i> | 49 |
| 5.1.1.5 | <i>Porção 5 - Seção para Alteração de Plano e Ativação de Chat Temporário</i> | 52 |
| 5.1.1.6 | <i>Porção 6 - Interface de Conversação</i> | 53 |
| 5.1.1.7 | <i>Porção 7 - Seção Meus GPTs</i> | 57 |
| 5.2 | Teste de Usabilidade | 58 |
| 5.2.1 | <i>Coleta e Interpretação de Dados</i> | 58 |
| 5.2.1.1 | <i>Taxa de Sucesso por Tarefa</i> | 61 |
| 5.2.1.2 | <i>Tempo Médio de Execução por Tarefa</i> | 62 |
| 5.2.1.3 | <i>Solicitação de Ajuda por Tarefa</i> | 63 |
| 5.2.1.4 | <i>Entrevistas</i> | 63 |
| 5.2.2 | <i>Discussão dos dados</i> | 70 |
| 6 | CONCLUSÕES E TRABALHOS FUTUROS | 79 |
| 6.1 | Trabalhos Futuros | 80 |
| 6.1.1 | <i>Dicas</i> | 81 |
| | REFERÊNCIAS | 82 |
| | APÊNDICES | 85 |
| | APÊNDICE A-INFORMAÇÕES PARA AVALIAÇÃO HEURÍSTICA | 85 |
| | APÊNDICE B-FORMULÁRIO DE RECRUTAMENTO | 87 |
| | APÊNDICE C-ENTREVISTA PÓS-USO | 88 |
| | APÊNDICE D-TERMO DE CONSENTIMENTO LIVRE E ESCLARE- CIDO (TCLE) | 89 |
| | APÊNDICE E-PROBLEMAS HEURÍSTICOS CONSOLIDADOS | 91 |

1 INTRODUÇÃO

Não é segredo que desde a Terceira Revolução Industrial e a Revolução Informacional, que ocorreram em meados dos anos 90, algumas tecnologias começaram a fazer parte do cotidiano das pessoas, com pouco tempo depois essas tecnologias foram nomeadas de Tecnologias da Comunicação e Informação (TICs), que ganharam cada vez mais espaço na sociedade atual por trazerem facilidade em tarefas do dia a dia (Santos, 2018). Com a inclusão exponencial desses novos artifícios, o âmbito da Inteligência Artificial (IA) e os *chatbots* ganharam um destaque notável no auxílio de seus usuários, sendo alguns deles muito conhecidos como a assistente de voz do Google, a Siri da Apple, a Cortana da Microsoft e a Alexa da Amazon. Porém, em 2022 uma empresa até então pouco famosa traria consigo uma novidade que em breve mudaria o rumo da inteligência artificial e o modo como ela é tratada.

O ChatGPT, lançado pela OpenAI em novembro de 2022, em essência, é um programa de computador projetado para entender e gerar linguagem natural. Isso significa que ele pode conversar, responder a perguntas, escrever textos e até mesmo gerar diferentes tipos de conteúdo criativo, como poemas, código, scripts, peças musicais, e-mails, cartas, etc (OpenAI, 2025). Sua sigla GPT, traduzindo para o português “transformador pré-treinado gerativo”, informa que a plataforma é em uma espécie de robô que pode conversar com um usuário qualquer a partir de treinamentos e simulações de diálogos pré-existentes, partindo do entendimento do funcionamento de uma IA.

Com a implantação dessas tecnologias, apesar das mudanças positivas, também foi possível notar dificuldades em abranger todos os públicos, visto que pessoas com pouca ou nenhuma experiência na área da Tecnologia da Informação não possuem um conhecimento prévio para usufruir da ferramenta com êxito. Em concordância com o (Forum, 2020), após a transformação digital ocorrida durante a pandemia de Covid-19, vários trabalhos que eram realizados manualmente passaram a ser digitalizados e com isso uma grande porcentagem dos trabalhadores precisaram mudar e se adequar ao novo formato do trabalho e aos processos digitais, incluindo o trabalho remoto. Porém, nem todas as ferramentas que poderiam ser utilizadas dentro desse âmbito trabalhista possuem uma “recepção” adequada para novos usuários, levando em consideração que estes não possuem algum tipo de formação ou conhecimento prévio necessário para usar a ferramenta sem impasses. O ChatGPT é um exemplo amplamente utilizado de inteligência artificial. No entanto, em sua versão gratuita, pode apresentar desafios de usabilidade para indivíduos com baixo letramento digital ou em diferentes contextos de uso. A falta de uma

abordagem direcionada a esses perfis pode comprometer tanto a acessibilidade quanto a eficácia da interação. Isso ressalta a importância de avaliar a usabilidade considerando a diversidade de usuários. Como resultado, a ferramenta pode não atingir plenamente aqueles que poderiam se beneficiar dela, caso sua usabilidade fosse melhor adaptada a diferentes necessidades e níveis de familiaridade com a tecnologia.

Este trabalho busca avaliar a usabilidade do ChatGPT-4o mini¹, sua versão gratuita, que é um modelo de linguagem desenvolvido pela OpenAI, derivado do modelo maior GPT-4o, sua versão paga. Para isso, iremos utilizar dois métodos da literatura: Avaliação Heurística e o Teste de Usabilidade (Prates; Barbosa, 2003) seguido de uma entrevista pós-teste, visando analisar a sua usabilidade, utilizando usuários totalmente iniciantes no ChatGPT, gerando possíveis cenários de uso da ferramenta a fim de explorar a curva de aprendizado e pontuar possíveis melhorias na plataforma.

Com a conclusão deste trabalho, profissionais e estudantes da área de Interação Humano-Computador (IHC) poderão ter acesso a dados e informações relevantes para melhoria do ChatGPT, ou outros *chatbots* semelhantes.

1.1 Objetivos

1.1.1 *Objetivo Geral*

Este trabalho tem como objetivo realizar uma Avaliação de Usabilidade do ChatGPT, visando os usuários iniciantes no ChatGPT que não possuem conhecimento do uso da plataforma.

1.1.2 *Objetivos Específicos*

1. Explorar contextos de uso da plataforma;
2. Investigar a usabilidade da interface;
3. Compreender como o usuário inciante interage com a plataforma;
4. Perceber a eficácia, eficiência e satisfação durante o uso do ChatGPT.

¹ <https://chat.openai.com/>

2 FUNDAMENTAÇÃO TEÓRICA

O objetivo dessa seção é apresentar os conceitos necessários para entender as informações expostas nesse trabalho, de forma clara e objetiva. Essa fundamentação iniciará falando sobre IA, apresentando um pouco das suas possibilidades, e evolução, logo em seguida será apresentado o conceito de *chatbot*, mostrando de forma sucinta, seu funcionamento e características. No tópico seguinte será abordado a visão da IHC, onde também é falado sobre Teste de Usabilidade e Avaliação Heurística, mostrando como funcionaram e sua importância dentro da tecnologia.

2.1 Inteligência Artificial

De acordo com Schwab (2019), alguns especialistas consideram que a era tecnológica atual é uma consequência direta da terceira revolução industrial. No entanto, o autor destaca que estamos à beira da quarta revolução, caracterizada por três aspectos fundamentais: a velocidade da evolução tecnológica, sua amplitude e profundidade, e seu impacto sistêmico.

Entre os avanços mais conhecidos da inteligência artificial, destacam-se sistemas capazes de interagir por meio da linguagem natural, como os assistentes pessoais presentes em smartphones. Além disso, essas tecnologias podem interpretar e compreender o mundo ao seu redor, reconhecendo imagens, documentos e áudio. Algumas abordagens de inteligência artificial permitem que as máquinas planejem sequências de ações para alcançar objetivos específicos. Já existem sistemas capazes de jogar xadrez, resolver problemas complexos e demonstrar habilidades notáveis em diferentes domínios (Garcia, 2020).

A ideia de Inteligência Artificial está sustentada na reprodução da inteligência humana mediante software, ou seja, na fundação de sistemas especializados de apoio à tomada de decisão, bem como no esforço de replicar atividades tipicamente humanas como a fala, audição, visão, deslocamento e entre outros. Segundo Demertzis e Iliadis (2014), a Inteligência Artificial desempenha um papel fundamental ao impulsionar o desenvolvimento e ampliar a disseminação de aplicações inteligentes em diversas áreas, como medicina, finanças e engenharia. Essa evolução tecnológica pode ser considerada um avanço significativo rumo à quarta revolução industrial mencionada por Schwab (2019), dado seu impacto na transformação dos setores e na redefinição da interação entre humanos e máquinas.

Nos últimos anos, a inteligência artificial tem aparecido cada vez mais em vários

campos, como mecanismos inteligentes de busca na Internet, reconhecimento de fala, robótica, mineração de dados e casas inteligentes. Essas tecnologias trabalham juntas para produzir resultados surpreendentes. Por exemplo, eles podem ajudar as empresas a controlar o estoque, ser usados em carros inteligentes, prever tendências de aplicativos usando dados de GPS em tempo real, reconhecer rostos, detectar fraudes e fornecer soluções de atendimento ao cliente usando *chatbots* e assistentes (Coutinho, 2021).

2.1.1 Chatbots

Os *chatbots* são sistemas de IA desenvolvidos para simular interações humanas, proporcionando conversas fluidas com usuários reais. Seu impacto abrange diversas áreas, como educação, negócios e disseminação de informações (Shawar; Atwell, 2007). Segundo Guichard *et al.* (2019), um *chatbot*, também conhecido como *chatterbot*, é um software baseado em texto e voltado para um domínio específico, projetado para auxiliar usuários humanos por meio de serviços especializados.

Nos últimos anos, os *chatbots* se tornaram mais populares à medida que a demanda por comércio eletrônico aumentou. Com isso também houve a diminuição de barreiras para a criação de chatbots. Hoje as pessoas podem criar seus próprios chatbots em plataformas de mídia social como Facebook Messenger e Twitter sem conhecer programação complexa e outras habilidades técnicas altamente especializadas (Ren *et al.*, 2022).

No mundo dos *chatbots* eles são divididos em duas categorias principais, que são domínio aberto e domínio fechado, os que possuem domínio aberto conseguem falar sobre vários assuntos gerais e responder de maneira adequada e natural para o usuário, já os *chatbots* que estão dentro do domínio fechado são focados em uma área de conhecimento específica e não conseguem responder adequadamente nenhuma outra pergunta que não esteja dentro de sua área. De acordo com Nimavat e Champaneria (2017), os *chatbots* podem ser classificados em quatro categorias principais:

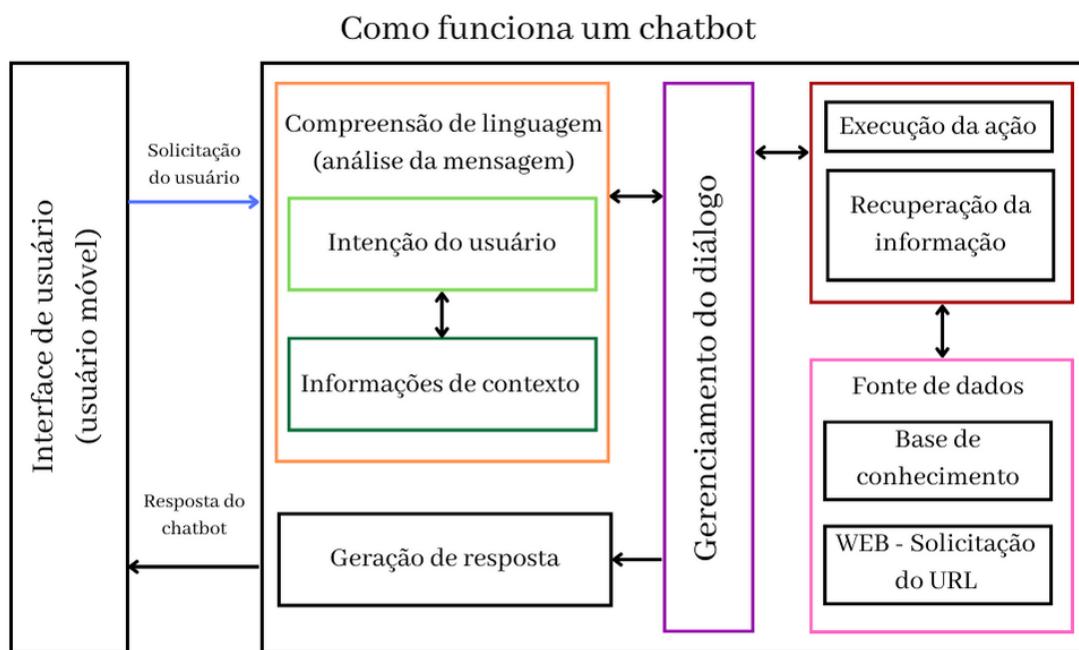
- *Chatbots* de serviço prestado: focados em interações frequentes e diretas com os usuários, oferecendo serviços específicos e personalizados.
- *Chatbots* interpessoais: destinados a facilitar comunicações diretas, como reservas em hotéis e restaurantes, ou fornecer respostas a perguntas frequentes, entregando informações pontuais sem acompanhamento contínuo.
- *Chatbots* intrapessoais: integrados em plataformas já utilizadas pelos usuários, como

aplicativos de mensagens, oferecendo assistência contínua e personalizada; um exemplo é o Google Assistente.

- *Chatbots* interagentes: projetados para interconectar diferentes assistentes virtuais, permitindo a comunicação entre eles; a interação entre Alexa e Cortana exemplifica essa categoria.

Essas classificações destacam a versatilidade dos *chatbots* e sua capacidade de se adaptar a diversas necessidades e contextos de uso.

Figura 1 – Arquitetura de funcionamento de um chatbot



Fonte: (Adamopoulou; Moussiades, 2020). Traduzido pelo autor

Na Figura 1, é possível observar como funciona a arquitetura de um *chatbot*. O processo tem início na seta azul, onde o usuário pergunta algo para o *chatbot*, em seguida o processo segue para dentro do quadrado laranja “compreensão da linguagem”, começando dentro da primeira caixa verde clara, onde o que foi pedido é armazenado e segue para a caixa verde escuro para ser colocado em um contexto, em seguida, dentro da caixa roxa a solicitação é compreendida completamente e logo é passada para dentro da caixa vermelha, onde os processos de “execução da ação” ou “recuperação da informação” ocorrem dentro do seu banco de dados conhecido como Knowledge Base ou “base de conhecimento”, localizada dentro da caixa rosa, após todo esse processo a resposta do *chatbot* é colocada em uma linguagem natural através do Processamento de Linguagem Natural (NLP) e a informação é devolvida para o usuário de maneira compreensível (Adamopoulou; Moussiades, 2020).

2.2 Conhecendo o ChatGPT

O GPT, desenvolvido pela OpenAI, é um modelo de linguagem baseado na arquitetura Transformer (Vaswani *et al.*, 2017), que se destaca pela eficiência no processamento de linguagem natural. O GPT foi pré-treinado em um extenso conjunto de dados textuais da internet, permitindo-lhe compreender e gerar texto de forma eficaz. Durante o pré-treinamento, ChatGPT aprendeu linguística complexa e desenvolveu uma compreensão profunda do idioma. A arquitetura do Transformer permite que o ChatGPT processe determinadas sequências, como palavras em uma frase, enquanto mantém o relacionamento entre elas. Foi submetido a processos de aprendizagem supervisionados e não supervisionados envolvendo várias tarefas de linguagem natural, incluindo prever palavras em uma frase, preencher um texto e responder a perguntas (Brown *et al.*, 2020). Assim como outros modelos, o ChatGPT é capaz de gerar respostas relevantes e baseadas no contexto fornecido pelos usuários.

Desenvolvido em 2022, o ChatGPT rapidamente se consolidou como uma ferramenta amplamente utilizada em diversos contextos desde seu lançamento. Dados recentes indicam que a plataforma atinge aproximadamente 300 milhões de usuários ativos semanalmente, o que evidencia seu rápido crescimento e adoção global (Turley, 2024). Essa expressiva base de usuários reflete a versatilidade do ChatGPT, que é empregado em áreas tão variadas quanto suporte ao cliente, educação, saúde e programação.

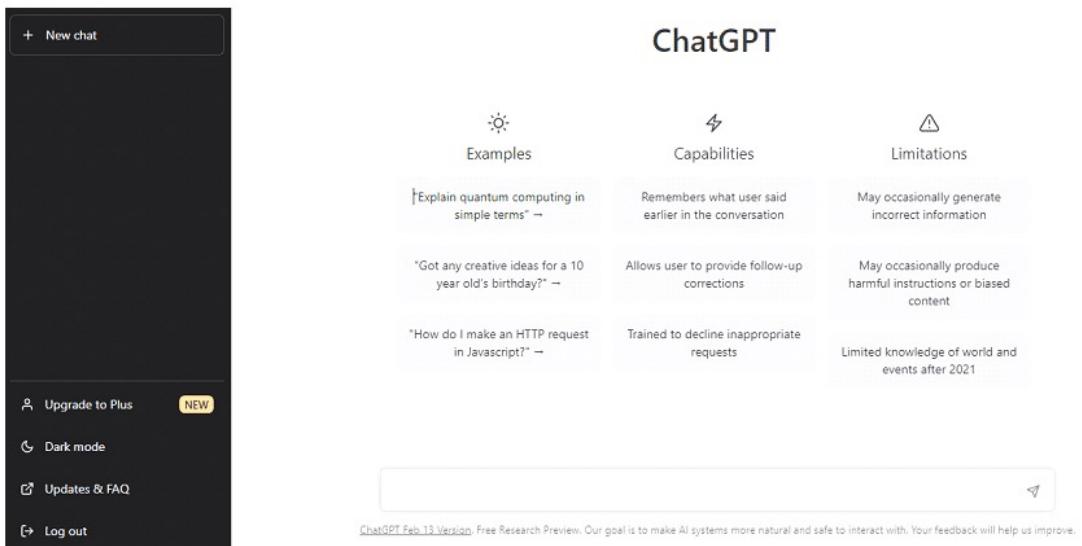
Na educação, o ChatGPT tem sido explorado como uma ferramenta de apoio à aprendizagem da língua portuguesa. Por exemplo, Junior (2024) avaliou as possibilidades e contribuições que o ChatGPT pode oferecer a estudantes e professores durante o processo de ensino-aprendizagem do português. O estudo revelou que o ChatGPT é útil na correção de erros, sugestão de vocabulário e fornecimento de explicações claras, embora apresente limitações culturais e emocionais.

Esses dados não apenas destacam a amplitude de usos do ChatGPT, mas também ressaltam a necessidade de avaliações contínuas tanto da usabilidade quanto da segurança da plataforma. Uma análise detalhada dos diferentes contextos de aplicação permite compreender as nuances de cada ambiente, garantindo que as melhorias implementadas sejam alinhadas às necessidades específicas dos usuários. Dessa forma, a investigação de métricas de satisfação, eficiência na execução de tarefas e a confiabilidade das respostas é crucial para orientar o desenvolvimento futuro da ferramenta.

2.2.1 Interface

Em meados de 2022, ano de lançamento do ChatGPT, sua interface era simples, possuindo uma proposta de simular uma conversação em linguagem natural. Sendo inicialmente projetada para facilitar a interação textual com os usuários, proporcionando respostas claras e detalhadas em uma ampla gama de tópicos. No entanto, algumas limitações e áreas de melhoria foram identificadas ao longo do tempo, o que levou ao desenvolvimento de versões mais avançadas (OpenAI, 2023). Sua antiga interface possuía poucas funcionalidades e sua estrutura e conseguia simular uma conversação natural, mas com menos sofisticação em termos de adaptação ao tom e estilo preferidos pelo usuário., assim como mostrado na Figura 2:

Figura 2 – Captura de tela do antigo ChatGPT



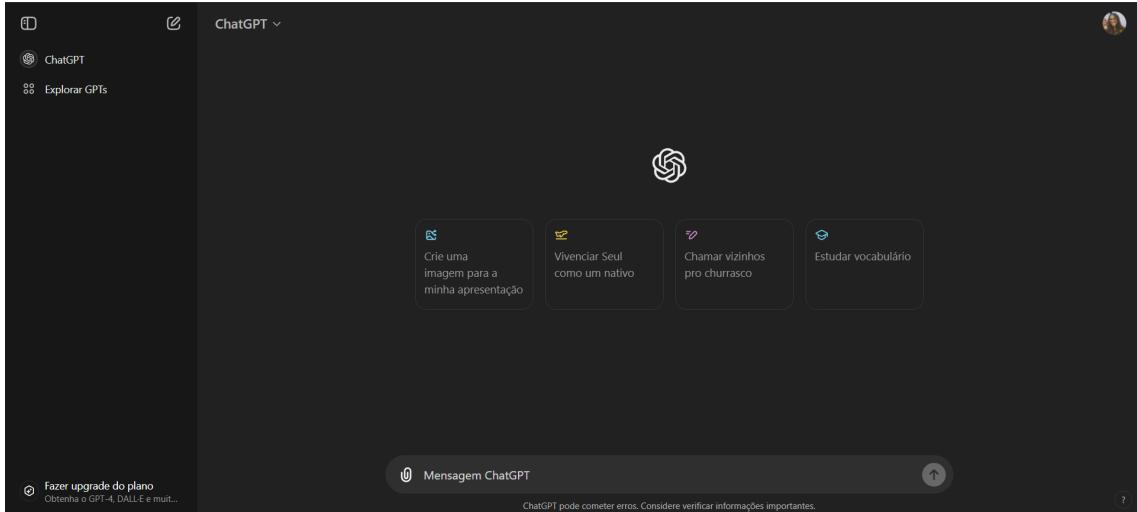
Fonte: <https://chatgpt.com/>. Retirado do Google.

A interface de interação do ChatGPT atual (2025), ainda segue a linha de simular uma conversação natural, permitindo que o usuário se comunique da mesma forma que faria com outra pessoa. Isso inclui a capacidade de lidar com linguagem coloquial, gírias, e variações de tom e estilo, o que torna a interação fluida e acessível. A interface possibilita que o usuário faça perguntas ou solicitações usando linguagem natural, e é adaptável a uma ampla gama de contextos, como suporte técnico, educação, escrita criativa e programação.

A comunicação na interface é puramente textual, semelhante a um aplicativo de mensagens, o que facilita a familiaridade e o uso intuitivo. Além disso, a interface integra ferramentas avançadas que ampliam a interatividade como as funções de excluir, arquivar e desarquivar conversas, permitindo uma experiência mais rica e dinâmica.

Uma das características mais notáveis dessa interface é sua capacidade de manter o contexto ao longo da conversa. Isso significa que o ChatGPT pode "lembrar" de informações fornecidas anteriormente, utilizando-as para gerar respostas mais relevantes e precisas, aprimorando a coerência e a continuidade do diálogo. Sua interface atual pode ser vista na Figura 3:

Figura 3 – Captura de tela do ChatGPT



Fonte: <https://chatgpt.com/>. Acesso em 09 out 2024.

2.2.2 Funcionalidades do ChatGPT

O modelo funciona por meio de inferência, em que o usuário insere uma entrada textual (prompt) e o ChatGPT processa essa entrada para gerar uma resposta, considerando o contexto da conversa. A interface do ChatGPT se assemelha a um chat convencional, onde o usuário digita mensagens e recebe respostas em tempo real. A interação é fluida e contínua, permitindo que o usuário explore diferentes tópicos dentro da mesma conversa.

O ChatGPT possui diversas funcionalidades que o tornam uma ferramenta versátil para diferentes aplicações. Entre as principais funcionalidades, destacam-se:

1. Geração de Texto: Produz respostas textuais coerentes para perguntas, explicações, resumos e diálogos naturais.
2. Auxílio na Produção de Conteúdo: Pode ser utilizado para criação de artigos, redação de e-mails, desenvolvimento de roteiros e escrita criativa.
3. Suporte Técnico e Educacional: Atua como assistente no aprendizado, fornecendo explicações detalhadas sobre diversos temas, auxiliando estudantes e pesquisadores.
4. Automação de Atendimento ao Cliente: Empresas utilizam o ChatGPT para responder

dúvidas frequentes e interagir com usuários em tempo real.

5. Programação e Codificação: Ajuda desenvolvedores fornecendo exemplos de código, correção de erros e explicações sobre algoritmos e estruturas de dados.
6. Tradução e Revisão Textual: Suporte na tradução de idiomas e na correção gramatical e estilística de textos.
7. Criação e Gerenciamento de Chats: O ChatGPT permite a criação de conversas individuais e múltiplas. O usuário pode iniciar um novo chat a qualquer momento e personalizar o contexto da interação.
8. Armazenamento e Continuidade: Algumas versões da tecnologia possibilitam a continuidade da conversa, permitindo que o usuário retome tópicos discutidos anteriormente.
9. Exclusão de Chats: O usuário pode excluir conversas anteriores para remover dados e iniciar novos diálogos com um novo contexto.
10. Personalização das Respostas: O usuário pode refinar as respostas pedindo reformulações, ajustes no tom da conversa ou maior nível de detalhe.
11. Integração com Outras Plataformas: O ChatGPT pode ser incorporado a sistemas de atendimento, assistentes virtuais e aplicações empresariais.
12. Gerenciamento de Contexto: Algumas versões avançadas conseguem manter um histórico de contexto mais longo, tornando as respostas mais precisas e coerentes ao longo da conversa.

2.3 Por que avaliar, o ChatGPT?

Para Moraes (2002) é importante minimizar o tempo demandado para a aprendizagem em aplicativos, a irritação dos usuários, a incapacidade de navegação em aplicativos, a subutilização de recursos, as possibilidades de erros e o baixo rendimento do trabalho. Atualmente as pessoas vivem conectadas e dependem dos seus smartphones para as atividades do dia a dia. Cybis (2003) relata que usuários expostos a interfaces deficientes frequentemente com uma alta taxa de experiências negativas, podem ficar frustrados e desenvolver ansiedade e estresse, então ofereça aplicativos, sites de qualidade e até o próprio aparelho estudos e testes.

A usabilidade de um produto pode ser aferida por meio de inspeções de usabilidade ou Testes de Usabilidade (Quiñones; Rusu, 2017). A adaptabilidade da interface permitirá que os usuários executem diferentes tarefas e em diferentes ambientes físicos, técnicos e organizacionais. Eles podem ter seus objetivos alcançados com mais facilidade para atingir os objetivos, podendo

também trazer retorno para a negócio, ao evitar perdas, atender os clientes e evitar custos de manutenção (Nielsen, 1994).

É relevante levar em consideração que mesmo seguindo critérios-se de qualidade e os passos para a elaboração de produtos adequados, isso nem sempre trará o resultado esperado, afinal erros existem em qualquer lugar, mesmo que sejam poucos é normal, mas é fato que isso afeta a qualidade do produto final. Igualmente a um produto qualquer, um sistema também pode e deve ser avaliado antes de ser entregue para seu público, mesmo que seja difícil assegurar que ele será perfeito, essa avaliação não pode ser descartada (Barbosa *et al.*, 2021).

2.4 Teste de Usabilidade

Para Carroll (2006), a usabilidade refere-se à facilidade de uso de um produto, enquanto a Organization for Standardization (ISO) a define como a capacidade de um sistema em permitir que usuários específicos alcancem objetivos com eficiência e satisfação (ABNT, 2002). Nesse contexto, o Teste de Usabilidade é uma metodologia que avalia a interação entre usuário e sistema, observando aspectos como eficácia, eficiência e satisfação (Nielsen, 1994).

O processo se inicia com o planejamento, etapa na qual se definem objetivos, público-alvo e tarefas a serem avaliadas, além dos métodos de coleta de dados e recursos necessários (Nielsen, 1993). Em seguida, ocorre o recrutamento de participantes representativos do público-alvo, geralmente entre cinco e dez pessoas, número suficiente para identificar problemas relevantes (Rubin; Chisnell, 2008; Nielsen, 2000).

As tarefas propostas devem simular cenários reais de uso e ser testadas previamente em um piloto para evitar problemas metodológicos (Dumas; Redish, 1999). Durante o teste, os participantes interagem com o sistema em um ambiente controlado, enquanto o moderador observa e coleta dados. A técnica *think aloud*, onde os usuários verbalizam seus pensamentos, é amplamente utilizada para compreender os processos cognitivos envolvidos (Nielsen, 1993).

Após a execução, os dados qualitativos e quantitativos são analisados, permitindo identificar problemas de usabilidade e propor melhorias. A etapa final consiste na validação das alterações realizadas por meio de novos testes, seguindo um processo iterativo para refinar continuamente a interface (Nielsen, 2000; Dumas; Redish, 1999).

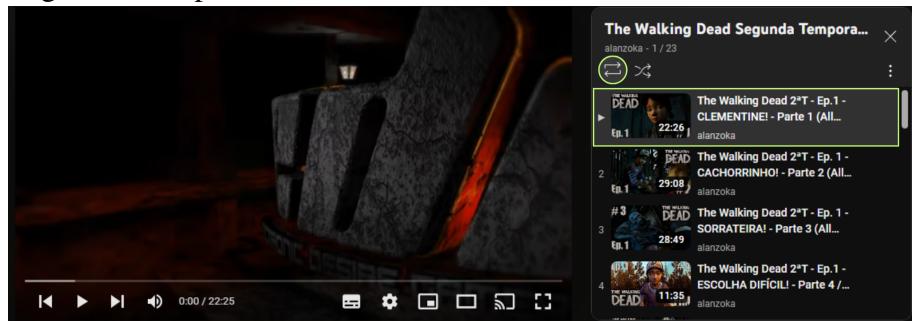
2.5 Avaliação Heurística

A Avaliação Heurística, desenvolvida por Jakob Nielsen em 1994, é um método utilizado para identificar problemas de usabilidade em sistemas e interfaces. Consiste em um conjunto de 10 princípios, conhecidos como heurísticas, que orientam designers e desenvolvedores na criação de interfaces mais intuitivas e eficientes. Essas diretrizes são aplicáveis tanto em ambientes digitais quanto físicos, com foco na melhoria da experiência do usuário. O principal objetivo é detectar falhas de usabilidade de forma rápida e prática, permitindo ajustes que atendam melhor às necessidades dos usuários. As heurísticas de Nielsen são reconhecidas como fundamentais no design de interfaces centradas no usuário.

1. Visibilidade do Status do Sistema: Segundo Nielsen (1994), esta heurística destaca a importância de manter o usuário informado sobre o estado atual do sistema, oferecendo feedback claro e em tempo adequado. Essa prática garante que o usuário tenha uma compreensão contínua das ações realizadas e das opções disponíveis.

Um exemplo prático dessa aplicação pode ser observado no YouTube, onde a plataforma fornece indicadores visuais sobre o status de reprodução de uma playlist. Como ilustrado na Figura 4, o vídeo atualmente em exibição é destacado dentro do quadrado verde, enquanto os vídeos seguintes na fila de reprodução são apresentados logo abaixo. Além disso, no círculo verde, observa-se a seleção da opção de repetir a playlist, demonstrando como o sistema informa o usuário sobre as configurações escolhidas. Esses elementos visuais asseguram que o usuário esteja ciente das suas ações e do estado atual do sistema, promovendo uma experiência mais intuitiva e satisfatória.

Figura 4 – Captura de tela do canal Alanzoka



Fonte: YouTube, 2023

2. Correspondência entre o sistema e o mundo real: Essa heurística explica que é importante que o sistema use a mesma linguagem do usuário – palavras, termos, rótulos e

ícones. É relevante que o sistema e o usuário tenham uma comunicação fluida, sem ruídos, a fim de facilitar a linha de aprendizado do usuário. Assim como mostra a Figura 5, os símbolos de telefone, calendário, relógio e mensagem, são facilmente reconhecidos pois já se tornaram familiares para seus usuários (Nielsen, 1994).

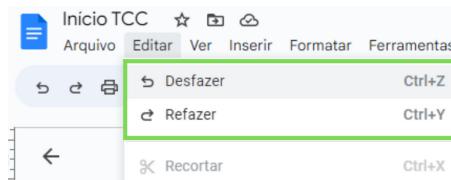
Figura 5 – Símbolos



Fonte: Elaborado pelo autor

3. Controle e liberdade do usuário: Dentro dessa heurística o autor explica que, de vez em quando, o usuário toma algumas decisões erradas, e por isso é necessário que existam algumas formas de reverter essas ações, uma espécie de “saída de emergência”, proporcionando liberdade para o indivíduo que usa aquele sistema (Nielsen, 1994). Um bom exemplo são as teclas de desfazer e refazer apresentadas no Google Docs, elas dão oportunidade para o usuário ter mais fluidez na hora de usar a ferramenta, como mostra dentro do quadrado verde, na Figura 6:

Figura 6 – Captura de tela Google Docs



Fonte: Elaborado pelo autor

4. Consistência e padrões: A quarta heurística de Nielsen, "Consistência e Padrões", destaca a importância de seguir convenções de design e comportamento estabelecidas tanto dentro de um sistema quanto em relação a outros sistemas similares. Essa heurística visa promover uma experiência de usuário mais intuitiva e familiar, permitindo que os usuários se sintam confortáveis e confiantes ao interagir com um sistema.

A consistência em um sistema significa que elementos similares devem funcionar da mesma maneira em diferentes partes do sistema. Isso inclui aspectos como layout, terminologia, design de elementos interativos e fluxos de interação. Quando os elementos de interface de usuário são consistentes, os usuários não precisam aprender novos conceitos ou comportamentos toda vez que encontram uma nova tela ou funcionalidade, o que facilita a aprendizagem e a navegação (Nielsen, 1994). A Figura 7 mostra um exemplo dessa heurística:

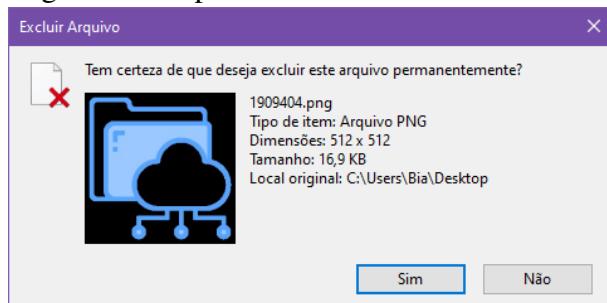
Figura 7 – Captura de Tela
Gmail



Fonte: Gmail, 2023

5. Prevenção de erros: Nessa heurística é explicado que o sistema deve ser projetado de forma a prevenir erros sempre que possível. Isso pode incluir a adição de confirmações de ações irreversíveis, restrições de entrada de dados para evitar formatos incorretos, entre outros mecanismos (Nielsen, 1994). Por esse motivo, é importante que exista algum tipo de confirmação para a realização de ações mais importantes, assim como mostra a Figura 8.

Figura 8 – Captura de tela

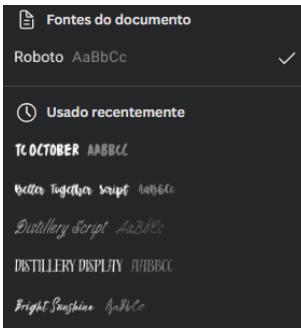


Fonte: Elaborado pelo autor

6. Reconhecimento em vez de memória: Como objetivo de facilitar a vida do usuário, fazendo com que o indivíduo reconheça ao invés de lembrar, ou seja é melhor dar um sentimento de familiaridade do que fazer com que o usuário tenha que memorizar cada processo no caminho percorrido pela aplicação. Também é importante deixar “rastros” de informações que podem vir a ser úteis. A Figura 9, mostra de uma forma simples como a sexta heurística funciona, pode-se observar que dentro do menu de fontes tem o tópico “usado recentemente”, que serve como um rastro para o usuário lembrar o que ele já fez anteriormente. Isso inclui a apresentação de menus, botões e opções sempre que necessário (Nielsen, 1994).

7. Flexibilidade e eficiência de uso: Essa regra afirma que o sistema ideal é útil tanto para usuários experientes quanto inexperientes. Os usuários leigos precisam das informações

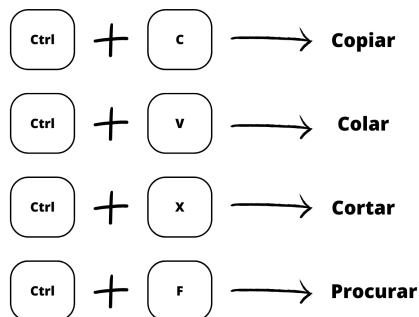
Figura 9 – Captura de tela do Canva



Fonte: Canva, 2023

necessárias para aprenderem a usar a interface, mas com o passar do tempo surge a necessidade de usar essas mesmas ferramentas por meio de atalhos. Dessa forma, é necessário que o sistema possua a opção de personalização, a fim de promover flexibilidade e adaptabilidade da interface (Nielsen, 1994). Um exemplo bem conhecido são os atalhos de teclado mostrados na Figura 10:

Figura 10 – Atalhos do teclado



Fonte: Elaborado pelo autor

8. Estética e design minimalista: Destaca a importância de criar interfaces de usuário limpas, simples e visualmente atraentes. Essa heurística reconhece que o design estético desempenha um papel significativo na experiência do usuário, influenciando a percepção da usabilidade e da qualidade de um sistema.

O design minimalista se concentra em eliminar elementos desnecessários e simplificar a apresentação de informações, priorizando a clareza e a funcionalidade. Ele visa criar uma experiência visualmente agradável, onde cada elemento na interface tem um propósito claro e contribui para a compreensão e eficácia do sistema (Nielsen, 1994).

9. Recuperação diante de erros: Essa heurística tem o intuito de ajudar o usuário a se recuperar de algum erro caso os outros métodos como o “Ctrl+Z” não funcione, o erro pode ser mostrado para o usuário através de cores e avisos escritos próximos ao erro (Nielsen, 1994). Um exemplo disso é o Sigaa, que exibe uma mensagem de erro, como mostra dentro do quadrado

verde na Figura 11:

Figura 11 – Captura de tela Sigaa

Fonte: Sigaa, 2023.

10. Ajuda e documentação: Mesmo que essa heurística fique um pouco "escondida", comparada às outras, essa área é de suma importância, pois algumas vezes o usuário precisa de ajuda em âmbitos mais profundos do sistema. Mesmo que as outras heurísticas ajudem a evitar erros, é sempre importante oferecer algo mais completo para ajudar o usuário. Um bom exemplo disso são os Frequently Asked Questions (FAQs) ou Perguntas Feitas Frequentemente, documentos mais completos oferecidos para problemas mais encorpados (Nielsen, 1994).

3 TRABALHOS RELACIONADOS

Nesta seção são apresentados alguns trabalhos que se relacionam com a pesquisa proposta, expondo um breve resumo de cada um deles. De maneira geral todos os trabalhos abordam sobre Avaliação de Usabilidade de chatbots, sendo alguns mais voltados para usuários específicos.

3.1 Usabilidade do ChatGPT como ferramenta de apoio no ensino de português como língua estrangeira

O estudo de Junior (2024) avaliou as possibilidades e contribuições do ChatGPT na aprendizagem do português como língua estrangeira, demonstrando o potencial da inteligência artificial para apoiar tanto professores quanto estudantes. Com base em um referencial teórico que abrange teorias de aprendizagem de línguas, inteligência artificial, processamento de linguagem natural (PLN) e o conceito de memória metálica, a pesquisa combinou uma análise de usabilidade da ferramenta com uma revisão da literatura para oferecer uma visão ampla de seu impacto pedagógico.

A metodologia adotada permitiu identificar que o ChatGPT é eficaz na correção de erros, na sugestão de vocabulário e na oferta de explicações claras e contextualizadas, contribuindo para uma experiência de aprendizagem mais dinâmica. Contudo, o estudo também apontou limitações, sobretudo em relação a aspectos culturais e emocionais, evidenciando que, apesar de suas capacidades técnicas avançadas, a ferramenta ainda enfrenta desafios para replicar a complexidade das interações humanas.

Além disso, a análise ressaltou a importância de atualizar os métodos tradicionais de ensino de línguas e de desenvolver novos recursos didáticos que integrem a inteligência artificial de maneira a tornar o processo de aprendizagem mais flexível e personalizado. A proposta é que, ao complementar as abordagens convencionais, o ChatGPT possa oferecer uma experiência de ensino mais interativa e adaptada às necessidades individuais dos usuários, promovendo maior autonomia e engajamento no estudo da língua.

Em síntese, o ChatGPT se apresenta como uma solução promissora para o apoio ao ensino do português como língua estrangeira, demonstrando potencial para inovar as práticas pedagógicas e ampliar o acesso a métodos de aprendizagem mais eficientes. A integração dessa ferramenta com plataformas digitais consolida seu papel como um recurso de apoio que, aliado à

supervisão e à adaptação dos educadores, pode contribuir significativamente para a evolução do ensino e da aprendizagem de idiomas.

3.2 Usabilidade de um chatbot para pessoas com hipertensão e diabetes

O estudo de Feitosa *et al.* (2024) buscou avaliar a usabilidade do GISSA Intelligent Bot, um chatbot desenvolvido para interagir com pessoas com Hipertensão Arterial Sistêmica (HAS) e Diabetes Mellitus (DM), oferecendo diálogos textuais sobre autocuidado e a Rede de Atenção Primária à Saúde. A pesquisa foi realizada com 119 participantes distribuídos entre os municípios de Sobral e Crateús, no Ceará, e Cruz Alta, no Rio Grande do Sul. A maioria dos participantes era do sexo feminino, com idades entre 40 e 60 anos, e utilizava a internet por mais de 30 minutos diários.

Para avaliar a usabilidade do chatbot, foi utilizado um Teste de Usabilidade e a System Usability Scale (SUS), que apresentou uma pontuação média de 72,3%, indicando boa usabilidade. Os resultados mostraram que 77% dos participantes afirmaram que gostariam de utilizar o chatbot frequentemente, enquanto 72,7% o consideraram fácil de usar. Esses dados reforçam que o GISSA Bot é uma ferramenta intuitiva e acessível, com potencial para contribuir significativamente para o autocuidado e suporte a pessoas com doenças crônicas não transmissíveis (DCNTs).

A metodologia utilizada incluiu um estudo transversal e quantitativo. Os participantes interagiram com o GISSA Bot em seus próprios dispositivos móveis ou em equipamentos fornecidos pela equipe de pesquisa. Após essa interação, responderam a um questionário baseado na escala SUS. A análise dos dados revelou que o sistema foi bem avaliado, com poucos relatos de inconsistências (87,2% dos participantes notaram poucas ou nenhuma inconsistência). Além disso, apenas 16,2% indicaram necessidade de ajuda técnica para manusear o chatbot, o que evidencia sua simplicidade e eficiência.

Apesar dos resultados positivos, os autores sugerem aprimoramentos no chatbot, como a inclusão de comandos de voz, mídias interativas (como vídeos e animações) e lembretes para consultas e medicamentos. Tais funcionalidades poderiam ampliar o alcance e a utilidade da ferramenta, especialmente para pessoas com dificuldades de leitura ou escrita. O estudo também destacou a importância de considerar a inclusão digital, já que 30,3% dos participantes utilizavam a internet por menos de 30 minutos ao dia, indicando a necessidade de soluções acessíveis para esse público.

O GISSA Intelligent Bot demonstrou ser uma solução promissora para promover o autocuidado e a educação em saúde. Sua integração com plataformas populares, como o Facebook Messenger, facilita a adoção pelo público-alvo e reforça seu papel como uma ferramenta eficaz no suporte à atenção primária à saúde. Além disso, o estudo enfatiza a importância de adaptar tecnologias às necessidades e limitações dos usuários, considerando fatores como idade, gênero e familiaridade com recursos digitais.

3.3 CHARTBOT: UMA INTERFACE CONVERSACIONAL PARA VISUALIZAÇÃO DE DADOS ABERTOS DO IBGE

Segundo Sousa (2022), o estudo apresenta uma avaliação heurística, a estruturação da modelagem de interação e um teste de usabilidade na plataforma Sistema IBGE de Recuperação Automática (SIDRA) do Instituto Brasileiro de Geografia e Estatística (IBGE). Além disso, o autor propõe a criação de um chatbot voltado à visualização de dados, visando aprimorar a experiência do usuário no acesso a informações estatísticas. Foram escolhidos 12 usuários de 18 a 36 anos com pelo menos o ensino médio completo, após a seleção, foram divididos em dois grupos, onde o primeiro grupo iria usar o método tradicional (SIDRA) e o segundo grupo iria usar a *chatbot* proposto para pesquisar as seguintes questões:

1. Pesquisar o PIB da cidade de Quixadá-CE;
2. Buscar pelo PIB de Quixadá-CE em 2015;
3. Verificar, em um período de dez anos, o PIB da cidade de Quixadá-CE.

Ao fim da Avaliação Heurística a plataforma SIDRA violou as seguintes heurísticas: H04: Consistência e padrões, H06: Reconhecimento ao invés de memória, H07:Flexibilidade e eficiência de uso e H09: Recuperação diante de erros.

No Teste de Usabilidade os usuários foram orientados a selecionar o grau de dificuldade dentro de cada atividade proposta por Sousa (2022), o grupo que ficou com o método tradicional não conseguiu concluir a atividade número 3: Verificar, em um período de dez anos, o PIB da cidade de Quixadá-CE, e alguns tiveram dificuldades nas outras atividades, já o grupo do *chatbot*, apenas 2 participantes não concluíram a atividade 3.

Assim, Sousa (2022) concluiu que a implantação de *chatbots* na Exploração de dados abertos do governo trariam benefícios para os seus usuários. O ChartBot proposto, consegue atuar como um complemento do portal SIDRA, além de poder pesquisar, acessar informações rapidamente e também visualizar mais detalhes sobre essas informações ao interagir junto com

um link fornecido em conversas, gráficos, tabelas e entre outros.

3.4 Comparação entre os trabalhos relacionados e a pesquisa proposta

Nesta seção é realizada uma análise comparativa de forma ampla entre os trabalhos relacionados e o trabalho proposto.

O estudo de Junior (2024), avaliou o uso do ChatGPT no ensino de português como língua estrangeira, destacando sua eficácia na correção de erros, sugestão de vocabulário e fornecimento de explicações claras. Contudo, foram identificadas limitações em aspectos culturais e emocionais, sugerindo a necessidade de atualizar métodos de ensino e desenvolver recursos que integrem a inteligência artificial para uma aprendizagem mais flexível e personalizada.

O estudo de Feitosa *et al.* (2024) avaliou a usabilidade do GISSA Intelligent Bot, um chatbot voltado para pessoas com hipertensão e diabetes, utilizando Teste de Usabilidade e a System Usability Scale (SUS). Com uma pontuação média de 72,3%, o sistema foi considerado fácil de usar por 72,7% dos participantes, e 77% gostariam de utilizá-lo frequentemente. A pesquisa apontou o potencial do chatbot para promover o autocuidado e sugeriu aprimoramentos como comandos de voz e mídias interativas. A integração com o Facebook Messenger demonstrou facilitar o acesso e a inclusão digital.

Sousa (2022) propõe uma Avaliação Heurística e um Teste de usabilidade na plataforma SIDRA do IBGE, e após identificar lacunas, Sousa (2022) vê a oportunidade de criar uma ferramenta chamada Chartbot que poderá facilitar a interação dos usuários, além disso, é feito uma Avaliação Heurística e um Teste de usabilidade do aplicativo. Mesmo de forma intrínseca, os trabalhos se assemelham quando eles buscam entender algo por meio de Avaliação Heurística e Testes de Usabilidade.

Quadro 1 – Comparação dos trabalhos

| Trabalhos Relacionados | Contém chatbot? | Contém ChatGPT? | Faz Avaliação Heurística? | Faz Teste de Usabilidade? |
|--------------------------------|-----------------|-----------------|---------------------------|---------------------------|
| (Junior, 2024) | Sim | Sim | Não | Não |
| (Feitosa <i>et al.</i> , 2024) | Sim | Não | Não | Sim |
| (Sousa, 2022) | Sim | Não | Sim | Sim |
| Trabalho Proposto | Sim | Sim | Sim | Sim |

Fonte: Elaborado pelo autor

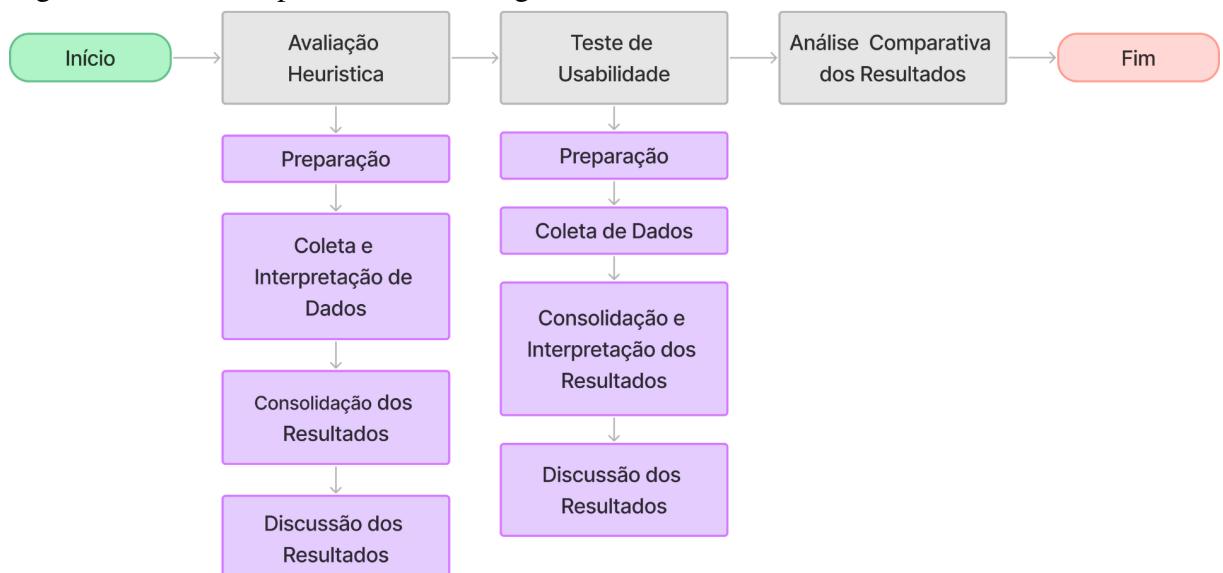
É importante salientar que alguns dos trabalhos citados não se concentram diretamente no ChatGPT. Além disso, os estudos de Junior (2024) e Feitosa *et al.* (2024) avaliam

aspectos da conversação, focando na fluidez, coerência e adequação das interações do sistema. Em contrapartida, o trabalho de Sousa (2022) – assim como o trabalho proposto – concentra-se exclusivamente na avaliação da interface, analisando fatores como a disposição dos elementos visuais, a naveabilidade e a responsividade do sistema. Essa distinção é crucial para entender a real proposta de cada estudo e possibilitar uma comparação mais precisa entre as abordagens adotadas.

4 METODOLOGIA

Nesta seção são apresentados os procedimentos metodológicos para a execução deste trabalho, em que serão usados métodos de avaliação de IHC por inspeção e observação, além de métodos qualitativos de coleta de dados, a fim de buscar um melhor entendimento do caso. A Figura 12 apresenta os passos a serem seguidos no trabalho:

Figura 12 – Passo-a-passo da metodologia



Fonte: Elaborado pelo autor

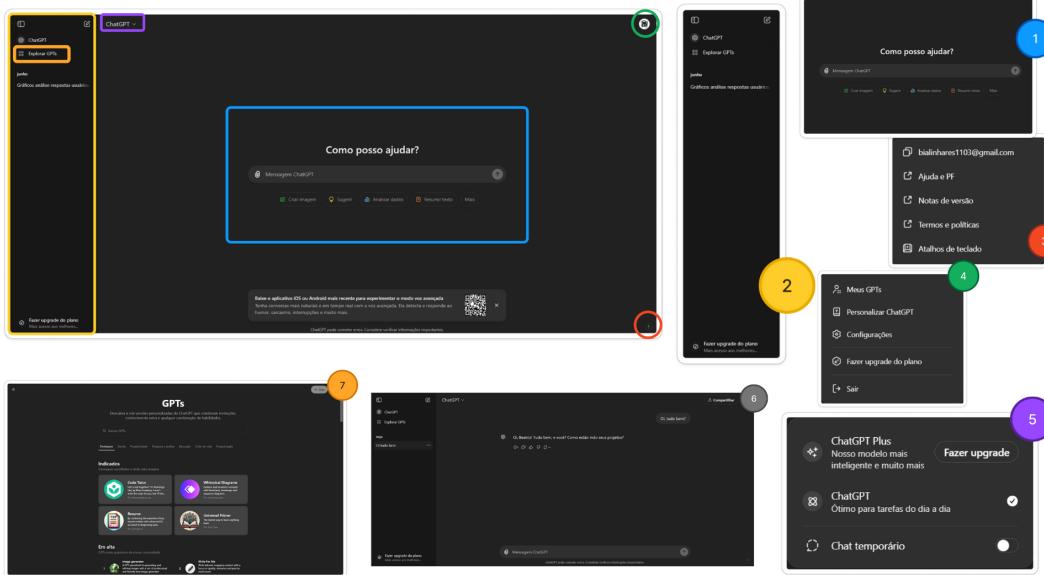
4.1 Avaliação Heurística

4.1.1 Preparação

Nesta etapa, foram convidados dois estudantes de Design Digital com experiência na execução de avaliações heurísticas. O recrutamento dos avaliadores ocorreu na Universidade Federal do Ceará (UFC), Campus Quixadá. O pré-requisito para o convite foi verificar se esses estudantes já haviam cursado a disciplina de Avaliação da Interação Humano-Computador, oferecida pela faculdade e já haviam feito uma avaliação heurística antes.

O primeiro passo foi identificar os principais fluxos de uso que um usuário poderia executar gratuitamente no ChatGPT, com o objetivo de realizar uma análise completa da experiência de interação. Com base nessa análise exploratória, a interface foi dividida em diferentes seções, conforme ilustrado na Figura 13, para facilitar a identificação de possíveis problemas de usabilidade.

Figura 13 – Divisão ChatGPT



Fonte: Elaborado pelo autor

Essa divisão foi tomada com o objetivo de tornar viável uma análise detalhada da interface, assegurando que cada aspecto seja cuidadosamente considerado. Essa abordagem sistemática não apenas aprimora a compreensão global do sistema, mas também facilita futuras revisões e análises, permitindo uma evolução contínua do design. A seguir, apresenta-se a divisão detalhada dos componentes da interface:

- Prompt Inicial com Sugestões de Prompts: Ao iniciar uma conversa, o ChatGPT apresenta um campo de entrada de texto onde os usuários podem digitar suas perguntas ou comandos. Além disso, são fornecidas sugestões de prompts predefinidos que auxiliam na formulação de consultas, facilitando a interação e garantindo respostas mais precisas.
- Aba Lateral com Histórico e Nomes dos Chats: Na interface, há uma aba lateral que exibe o histórico completo das conversas anteriores. Cada interação é listada com seu respectivo título ou nome atribuído pelo usuário, permitindo fácil navegação e referência a diálogos passados.
- Menu de Ajuda na Lateral Direita: Localizado na parte direita da interface, o menu de ajuda oferece recursos adicionais, como tutoriais, FAQs e suporte ao vivo. Essa seção visa auxiliar os usuários a tirar dúvidas sobre funcionalidades, solucionar problemas comuns e obter orientações sobre o uso eficaz do ChatGPT.
- Menu Principal com Configurações de Conta, Informações e Logout: O menu principal, acessível geralmente no canto superior esquerdo, contém opções para:
 - Configurações de Conta: Permite ao usuário gerenciar preferências pessoais, alterar

dados cadastrais e ajustar configurações de privacidade. Opções: Fornece detalhes sobre a versão atual do ChatGPT, atualizações disponíveis e informações de contato para suporte.

- Logout: Opção para encerrar a sessão atual, garantindo a segurança e privacidade do usuário.
- Seção para Alteração de Plano e Ativação de Chat Temporário: Nesta área, os usuários podem:
 - Alterar Plano: Visualizar e selecionar entre diferentes planos de assinatura, cada um oferecendo benefícios específicos, como acesso prioritário, recursos adicionais ou limites expandidos de uso.
 - Ativar Chat Temporário: Permite iniciar conversas com duração limitada ou específicas, úteis para consultas rápidas ou sessões de brainstorming.
- Interface de Conversação: A área central da interface é dedicada à interação direta com o ChatGPT. Nela, as mensagens são exibidas em formato de balões de chat, diferenciando claramente entre as entradas do usuário e as respostas do modelo. Recursos adicionais, como opções de formatação de texto, envio de anexos ou integração com outras ferramentas, podem estar disponíveis para enriquecer a experiência de conversação.
- Seção "Meus GPTs" com Funcionalidades Exclusivas: A seção "Meus GPTs" oferece aos usuários acesso a funcionalidades avançadas e exclusivas, disponíveis mediante assinatura. Entre os recursos destacados estão:
 - Criação de Imagens: Utiliza modelos de geração de imagens para criar visualizações baseadas em descrições textuais fornecidas pelo usuário.
 - Edição de Código: Permite ao ChatGPT auxiliar na escrita, depuração e otimização de código em diversas linguagens de programação.
 - Análise de Dados: Oferece ferramentas para processar e interpretar conjuntos de dados, gerando insights e visualizações relevantes.
 - Aprimoramento de Texto: Auxilia na revisão, edição e aprimoramento de textos, garantindo clareza, coerência e correção gramatical.

A seguir foi realizada uma reunião via Google Meet, com ambos os avaliadores em um horário e dia de disponibilidade deles. Posteriormente, como indicado por Barbosa e Silva (2010), foram apresentadas todas as informações sobre funções, objetivo e público-alvo (usuários iniciantes), que são informações pertinentes para a avaliação no geral. Ainda nessa reunião foi

repassado, um link do FigJam, incluindo o documento, disponível no Apêndice A que continha o escopo da avaliação, mostrando as atividades, telas que seriam avaliadas e uma breve explicação das heurísticas de Nielsen (1993).

A escolha das heurísticas de Nielsen se justifica por serem amplamente reconhecidas na área de usabilidade e por sua abrangência na inspeção dos diferentes aspectos dos sistemas.

4.1.2 *Coleta e Interpretação dos Dados*

A seguir, após o alinhamento do escopo da avaliação junto aos avaliadores, a análise individual do ChatGPT foi iniciada, utilizando as heurísticas mencionadas na etapa anterior. O objetivo era verificar se as diretrizes foram seguidas ou infringidas em cada tela. A inspeção foi conduzida utilizando o próprio computador dos avaliadores, realizando duas etapas:

1. Visão geral: Explorar as funções ChatGPT, para obter uma visão geral.
2. Avaliação: Análise detalhada de cada elemento proposto pelas atividades.

A avaliação foi feita por partes, ou seja, foi selecionado uma das porções mostradas anteriormente na Figura 13 e com base nela, os avaliadores examinaram todas as funções e heurísticas em busca de possíveis problemas. Posteriormente, o processo é repetido para cada uma das outras porções.

Segundo Barbosa *et al.* (2021), cada problema identificado precisa estar atrelado a uma heurística violada, o local do problema, a gravidade desse problema e o motivo de ser considerado um problema e, por fim, as sugestões de possíveis soluções. Visando propor o julgamento da gravidade dos problemas pelos avaliadores, foi apresentada a escala de severidade proposta por Nielsen (1994), exposta no Quadro 2:

Quadro 2 – Severidade do problema

| Caracterização do Problema | Descrição |
|----------------------------|---|
| 1. Problema cosmético | Detalhe mínimo, corrigir apenas se tiver tempo sobrando no projeto; |
| 2. Problema pequeno | Problema com conserto de baixa prioridade; |
| 3. Problema grande | Necessidade de atenção elevada, deve receber alta prioridade; |
| 4. Problema catastrófico | É extremamente importante, não é possível entregar um produto com um erro desses, prioridade total. |

Fonte: (Nielsen, 1994), Adaptado pelo autor.

4.1.3 Consolidação dos Resultados

De acordo com Barbosa e Silva (2010), ao fim da inspeção, os avaliadores se reuniram, para analisar e comparar os problemas encontrados no ChatGPT. O objetivo era proporcionar a todos uma visão abrangente dos problemas encontrados na interface por cada avaliador.

Durante essa consolidação, foi necessário que o grupo reavaliasse os problemas em conjunto, atribuindo um novo grau de severidade - se necessário - e decidindo quais problemas e soluções seriam incluídos nos resultados finais. As novas informações resultantes dessa discussão foram registradas coletivamente no FigJam, que representou a consolidação dos resultados da avaliação.

4.1.4 Discussão dos Resultados

Por fim, após a reunião de consolidação mencionada anteriormente, foi registrado o número de heurísticas violadas. Para facilitar a visualização dos problemas, foram inseridos prints das interfaces, destacando os pontos exatos onde ocorreram as violações.

Além disso, foram gerados gráficos utilizando o FigJam que demonstram as heurísticas mais violadas, quantidade de ocorrências e apresentando o grau de severidade dos problemas identificados. Junto com esses gráficos, foi desenvolvida uma tabela detalhada contendo todos os problemas detectados, as heurísticas violadas, o grau de severidade e possíveis sugestões de correções. Esses recursos visuais proporcionam uma compreensão mais clara dos elementos que podiam ser melhorados.

4.2 Teste de Usabilidade

4.2.1 Preparação

Antes de qualquer preparativo, foram criadas cinco contas para evitar qualquer influência nos resultados. Além disso, foi escolhido um ambiente calmo e controlado para garantir que não haja interferências ou imprevistos durante o teste.

Assim como recomenda Nielsen (1993), foram recrutados cinco participantes para a realização do teste de usabilidade. O recrutamento dos perfis de usuários foi realizado através de um Formulário Google disponibilizado no Apêndice B, incluindo perguntas básicas de

identificação, nível de interação com Inteligências Artificiais e familiaridade com ChatGPT. Esse formulário foi repassado em grupos do WhatsApp, convidando os usuários para a pesquisa de maneira voluntária.

Baseando-se nas respostas, foram escolhidos um grupo de usuários que representavam pessoas novatas em relação ao uso do ChatGPT, com níveis distintos de conhecimento sobre IA Assistentes. Essa diversidade foi intencional para garantir uma coleta de dados rica e variada, refletindo diferentes perspectivas dos participantes.

O teste foi realizado na versão desktop do ChatGPT 4o-mini, em um notebook Asus Vivobook 15. Nessa etapa, foram utilizadas as mesmas tarefas da avaliação heurística, focando em atividades que abrangem os aspectos mais relevantes do ChatGPT. De acordo com a orientação de Barbosa e Silva (2010), as tarefas foram organizadas em formato de cenários, conforme apresentado na Tabela 1, permitindo que os usuários tivessem as próprias percepções das atividades propostas

Tabela 1 – Descrição das Atividades

| Nome da Atividade | Descrição |
|--|---|
| Cenário 1: Iniciar uma Conversa | Você acabou de conversar com um amigo(a) sobre assuntos atuais do Brasil, porém em um momento da conversa ela(ola) falou de um assunto no qual você desconhecia, chegando em casa você resolveu pesquisar mais sobre as atualidades do Brasil. Seu objetivo é obter essa informação usando o ChatGPT. |
| Cenário 02: Continuando a exploração simples | Vamos supor que você está conversando sobre um assunto do seu interesse, porém não lembra ou não sabe sobre um tópico que seria importante durante a conversa. Seu objetivo é pesquisar e entender o tópico (escolhido pelo usuário) usando o ChatGPT. |
| Cenário 03: Conhecendo funções | Você pretende perguntar algo mais específico e quer que esse tópico fique separado dos demais, a fim de manter as perguntas separadas e organizadas para melhor análise. Seu objetivo é criar uma nova conversa para iniciar outra pergunta. |

Continua na próxima página

Tabela 1 Continuado da página anterior

| Nome da Atividade | Descrição |
|--------------------------------------|--|
| Cenário 04: Excluir uma conversa | Alguns chats antigos já não são mais usados e precisam ser deletados de alguma forma. Seu objetivo é excluir uma conversa criada anteriormente. |
| Cenário 05: Arquivar uma conversa | Um chat específico precisa ser ocultado da interface por algum motivo. Seu objetivo é arquivar uma conversa específica. |
| Cenário 06: Desarquivar uma conversa | Você arquivou uma conversa há algum tempo, mas agora precisa acessá-la novamente. Seu objetivo é recuperar essa conversa arquivada, trazendo-a de volta para o menu principal para facilitar o acesso. |

Fonte: Elaborado pelo autor

Segundo o padrão ISO 9241-11 a usabilidade é moldada conforme um produto pode ser usado por seus usuários a fim de atingir objetivos próprios e obtendo três características dentro do seu contexto de uso (ABNT, 2002), portanto também foram utilizadas as seguintes medidas:

1. Eficácia: A eficácia refere-se à precisão e à completude com que os participantes executam as tarefas propostas. Em outras palavras, avalia se os usuários conseguem atingir os objetivos estabelecidos de forma correta e completa. No Teste de Usabilidade, um dos principais indicadores analisados foi:
 - a) Taxa de sucesso: A porcentagem de tarefas que os participantes conseguiram completar com êxito
2. Eficiência: A eficiência é avaliada com base na quantidade de recursos utilizados pelos usuários para concluir uma tarefa. Quando os usuários conseguem atingir seus objetivos utilizando menos tempo, esforço ou ações, isso indica que o design é intuitivo e de fácil aprendizado. No Teste de Usabilidade, foram analisados dois aspectos principais:
 - a) Tempo de conclusão das tarefas: O tempo que cada usuário levou para finalizar cada tarefa proposta, considerando se o resultado foi alcançado de maneira correta e sem erros.
 - b) Solicitações de ajuda: O número de vezes que os usuários precisaram recorrer a

assistência externa ou buscar orientações para finalizar uma tarefa.

3. Satisfação: Métrica que avalia se os participantes conseguem concluir as tarefas sem apresentar reclamações, sendo complementada por entrevistas que exploram diretamente suas percepções e pontos de melhoria. Dessa forma, é possível obter uma visão abrangente da satisfação dos usuários com o sistema.

A métrica de satisfação foi medida através de um questionário pós-uso disponível no Apêndice C. Para validar o material utilizado, foi realizado um teste piloto com um usuário que possui o mesmo perfil do estudo. O objetivo foi identificar possíveis problemas que pudessem impactar os resultados do experimento. Com base nesse teste, foi estabelecido um tempo de aproximadamente 10 minutos para a conclusão do teste por cada participante.

4.2.2 Coleta de Dados

Na fase de observação, cada usuário foi individualmente convidado para uma sala, onde foram apresentados o objetivo da avaliação, a plataforma em estudo e os tipos de participantes do teste. Além disso, os usuários receberam orientações das tarefas a serem realizadas, essas informações (Apêndice C) foram fornecidas em uma folha impressa, proporcionando aos participantes um contexto mais claro para o estudo. Em seguida, foi pedido que o usuário assinasse o termo de consentimento também impresso, confirmando sua participação voluntária, disponível para visualização no Apêndice D.

Para observar o fluxo seguido pelo usuário, foi utilizado um aplicativo de gravação de tela e voz para desktop, disponível no Windows, dessa forma, foi possível visualizar se o usuário seguiu o caminho com sucesso, como também ouvir comentários durante a execução das tarefas. O Quadro 3 mostra as ferramentas utilizadas para a coleta.

Quadro 3 – Ferramentas usadas

| Ferramentas | |
|--------------------|------------------------|
| Dispositivo | Gravador de voz |
| Asus Vivobook 15 | Xbox Game Bar |

Fonte: Elaborado pelo autor

4.2.3 Consolidação e Interpretação dos Resultados

Nesta etapa, foram consolidados os dados individuais de cada usuário com o objetivo de avaliar se eles concluíram com sucesso as tarefas propostas. Para isso, foram analisadas as

gravações da tela do notebook utilizados durante o teste, permitindo interpretar e extrair todos os dados relevantes para o estudo.

Com base nessas descobertas, foi possível obter uma compreensão detalhada sobre as métricas de eficácia, eficiência e satisfação discutidas na fase de preparação. No que diz respeito à medição de satisfação, foram analisadas as respostas ao questionário pós-uso, que buscou avaliar o grau de satisfação dos usuários ao utilizar o aplicativo, bem como entender suas impressões sobre essa experiência inicial com o ChatGPT.

4.2.4 Análise Coomparativa dos Resultados

Nesta seção, foi feito um comparativo das descobertas relatadas tanto no método de avaliação heurística como no teste de usabilidade. Essa abordagem permitiu estabelecer conexões entre os problemas identificados durante a inspeção e aqueles enfrentados pelos usuários em situações reais de utilização.

5 RESULTADOS

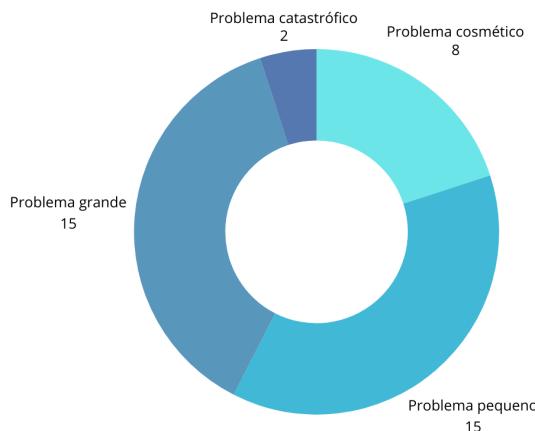
Neste capítulo, são apresentados os resultados que foram alcançados através da execução da metodologia mencionada no Capítulo 4. Primeiramente, foi executada a avaliação heurística que trouxe informações valiosas sobre problemas encontrados no ChatGPT. Posteriormente, o teste de usabilidade foi realizado, contando com participação de usuários reais na utilização do aplicativo. Logo, esses dois métodos foram extremamente relevantes para ter o conhecimento necessário sobre a usabilidade do aplicativo em estudo.

5.1 Avaliação Heurística

5.1.1 *Consolidação dos Resultados*

A Avaliação Heurística foi realizada entre 14 e 16 de outubro de 2024, utilizando notebooks pessoais e um link do FigJam. Foram identificadas 40 violações de heurísticas no ChatGPT: 8 problemas cosméticos, 15 pequenos, 15 grandes e 2 catastróficos, conforme ilustrado na Figura 14.

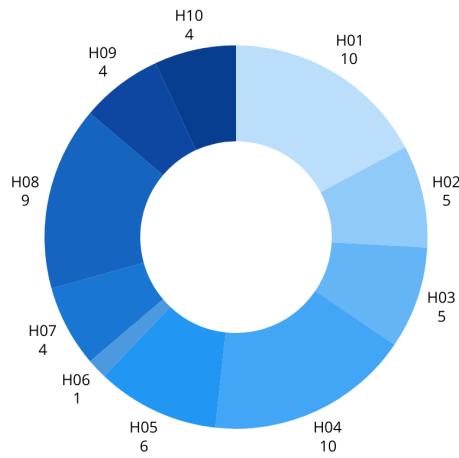
Figura 14 – Gráfico de violações por nível de problema



Fonte: Elaborado pelo autor

A quantidade de violações por heurística é ilustrada pela Figura 15, que apresenta de maneira gráfica as diretrizes de Nielsen violadas juntamente com a quantidade. Dessa forma, esses resultados proporcionam uma compreensão mais clara e concisa das áreas específicas que demandam melhorias, permitindo uma abordagem mais eficaz na otimização da experiência do usuário.

Figura 15 – Número de violações por heurística



Fonte: Elaborado pelo autor

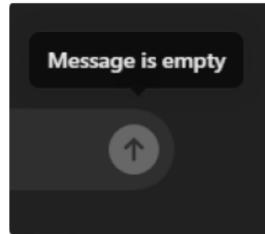
Após a consolidação das análises realizadas pelos três avaliadores, foram identificados e detalhados os problemas de usabilidade, os quais são descritos a seguir e apresentados na Tabela 11, que pode ser consultada no Apêndice E. Essas informações foram repartidas em porções, assim como mostra a Figura 13.

5.1.1.1 Porção 1 - Prompt Inicial com Sugestões de Prompts

A exibição da mensagem "message is empty" ao passar o cursor sobre o botão "enviar" viola as heurísticas H02 (correspondência entre o sistema e o mundo real), H04 (consistência e padrões) e H09 (ajuda os usuários a reconhecer, diagnosticar e se recuperar de erros), conforme descrito por Jakob Nielsen. A H02 enfatiza que o sistema deve falar a linguagem do usuário, utilizando palavras, frases e conceitos familiares, ao invés de termos técnicos ou estrangeiros. A presença de uma mensagem em inglês em uma interface destinada a usuários que falam português pode causar confusão e dificultar a compreensão, especialmente para aqueles com conhecimento limitado do idioma. Além disso, a H04 destaca a importância da consistência e da aderência a padrões na interface. A inconsistência linguística, como a mistura de idiomas, pode comprometer a experiência do usuário, levando a erros e frustrações. Por fim, a H09 ressalta que as mensagens de erro devem ser expressas em linguagem simples, indicando claramente o problema e sugerindo soluções. Uma mensagem de erro em inglês pode não ser eficaz para todos os usuários, dificultando a identificação e correção do problema. Portanto, é essencial que todas as mensagens do sistema estejam no idioma principal da interface, garantindo uma comunicação clara e consistente com os usuários.

Outro problema encontrado refere-se ao baixo contraste nos botões que oferecem

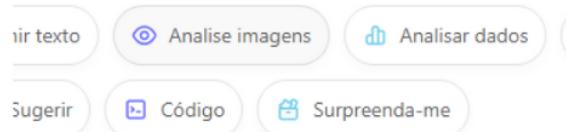
Figura 16 – Botão de enviar mensagem



Fonte: ChatGPT. out 2024.

"prompts pré-escritos" no modo claro, durante o efeito hover. Isso está em desacordo com a heurística H08 (Estética e design minimalista), que sugere que o design deve ser claro e visível, apresentado na Figura 17. O baixo contraste dificulta a identificação dos botões interativos, o que pode ser confuso para usuários novatos ou com pouca familiaridade com interfaces digitais. Eles podem não perceber que esses botões são clicáveis, o que gera incerteza sobre as ações possíveis, prejudicando a eficiência e gerando frustração.

Figura 17 – Prompts pré-escritos



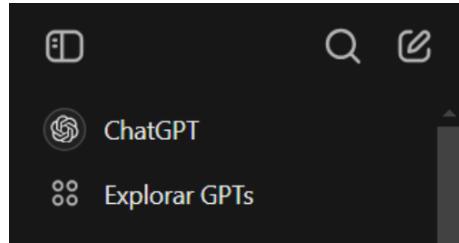
Fonte: ChatGPT. out 2024.

5.1.1.2 Porção 2 - Aba Lateral com Histórico e Nomes dos Chats

Um dos problemas identificados na análise refere-se aos botões, como o de lápis no menu "ChatGPT", que não apresentam feedback visual ao serem clicados, mostrado na Figura 18. A falta de um *feedback* claro após a interação compromete a comunicação com o usuário, tornando incerta a ação realizada. Para os usuários novatos, que podem não estar familiarizados com os efeitos de interação, isso pode gerar confusão sobre se o comando foi ou não executado.

Outro problema identificado envolve o botão de "arquivar", que exibe um modal apenas uma vez, e não repete a mensagem caso o usuário clique novamente. Esse problema está relacionado às heurísticas H01 (Visibilidade do status do sistema), H03 (Controle e liberdade do usuário) e H06 (Prevenção de erros). A falta de consistência no comportamento do modal

Figura 18 – Botões de criar novo chat



Fonte: ChatGPT. out 2024.

pode gerar incertezas sobre o estado da ação, especialmente para usuários que podem não perceber que a conversa foi arquivada. Para usuários com pouca experiência em tecnologia, esse comportamento pode ser frustrante.

Outro ponto observado é a falta de *feedback* visual ao editar o nome de uma conversa. Embora essa ação aconteça, o usuário não recebe uma confirmação visual, como ocorre em outras ações (como arquivar ou excluir conversas). De acordo com a heurística H01 (Visibilidade do status do sistema) e H04 (Consistência e padrões), a ausência de um feedback claro pode deixar o usuário inseguro se a ação foi realizada com sucesso.

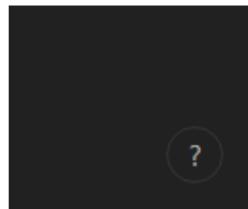
Por fim, um problema crítico foi identificado na edição do nome de uma conversa antiga, onde o histórico da conversa, incluindo a data de início, é perdido. Esse erro está relacionado às heurísticas H03 (Controle e liberdade do usuário) e H05 (Erro de prevenção), já que a perda de dados importantes compromete completamente a utilidade da função de renomeação. Para usuários que dependem dessas informações, como datas e contextos de conversas, a perda de dados pode resultar em frustração e até mesmo perda de confiança no sistema.

5.1.1.3 Porção 3 - Menu de Ajuda na Lateral Direita

O problema identificado refere-se ao ícone de Interrogação no canto inferior direito, que é descrito como muito pequeno, claro e discreto, mostrado na Figura 19. De acordo com a heurística H08 (Estética e design minimalista), um ícone tão pequeno e discreto pode dificultar a identificação e interação do usuário, especialmente em contextos onde o acesso a informações de ajuda é necessário. Para usuários novatos, que podem precisar de assistência frequente, esse ícone pode ser facilmente ignorado, o que prejudica a usabilidade do sistema. Além disso, essa questão também viola a heurística H10 (Ajuda e documentação), uma vez que a ajuda deve ser facilmente localizável e acessível quando necessária. Se os usuários não identificam o

ícone de suporte devido ao seu tamanho reduzido e baixa visibilidade, a experiência pode ser comprometida, dificultando o acesso a informações essenciais e impactando negativamente a aprendizagem e a eficiência no uso do sistema.

Figura 19 – Botão de dúvidas

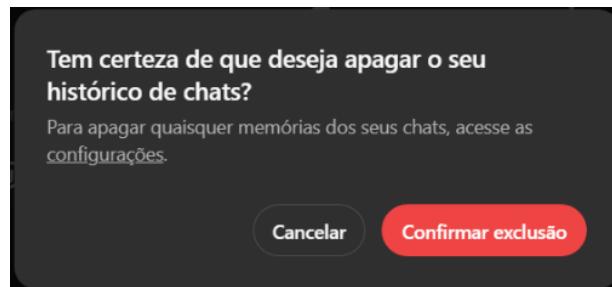


Fonte: ChatGPT. out 2024.

5.1.1.4 Porção 4 - Menu Principal com Configurações de Conta, Informações e Logout

Um dos problemas identificados diz respeito à funcionalidade de excluir todos os chats, onde não é informado que os chats arquivados também serão excluídos, mostrado na figura Figura 20. Esse problema está relacionado à heurística H05 (Prevenção de erros), pois a falta de aviso prévio pode resultar em ações não intencionais, gerando frustração. Para os usuários novatos, a exclusão inesperada de chats arquivados pode causar perda de dados importantes. A recomendação é informar claramente que os chats arquivados também serão excluídos, evitando surpresas e proporcionando maior controle sobre a ação.

Figura 20 – Modal de confirmação de exclusão de chats



Fonte: ChatGPT. out 2024.

Outro problema refere-se à falta de descrição ao lado do ícone de planeta na seção "Configurações → Perfil de Construtor → Links", apresentado na Figura 21. De acordo com a heurística H04 (Consistência e padrões), todos os ícones deveriam seguir o mesmo padrão

visual, o que inclui a presença de descrições ao lado de ícones, como ocorre com outros ícones na interface. A recomendação seria adicionar uma descrição ao lado do ícone de planeta, garantindo consistência e facilitando a compreensão do usuário.

Figura 21 – Links de domínios

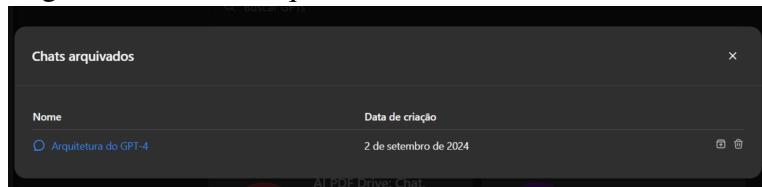


Fonte: ChatGPT. out 2024.

Na opção "Configurações → Perfil de Construtor → Selecionar um domínio", é possível adicionar qualquer termo, mas a mensagem de erro só aparece após a tentativa de salvar. Esse comportamento viola a heurística H05 (Prevenção de erros), pois permite que o usuário insira dados inválidos sem uma restrição imediata, o que pode gerar frustração. A recomendação seria restringir o input para aceitar apenas domínios válidos, evitando erros e otimizando a experiência do usuário.

Outro ponto importante é a funcionalidade de excluir ou desarquivar chats arquivados individualmente, mostrado na Figura 22, o que pode ser inconveniente para usuários que desejam realizar essas ações em múltiplos chats. De acordo com a heurística H07 (Flexibilidade e eficiência de uso), seria interessante adicionar uma opção para selecionar e desarquivar todos os chats de uma vez. Isso proporcionaria mais eficiência na interação com a interface, especialmente para usuários com maior volume de chats.

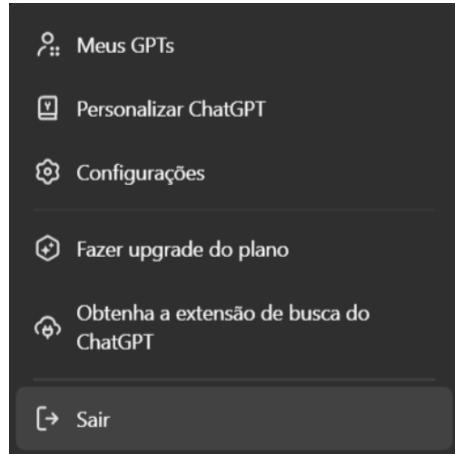
Figura 22 – Chats arquivados



Fonte: ChatGPT. out 2024.

O botão “sair” localizado em “Configurações”, mostrado na Figura 23 também apresenta um problema: ao ser clicado, o logout ocorre imediatamente sem um diálogo de confirmação. Esse comportamento está em desacordo com a heurística H05 (Prevenção de erros), pois a ausência de confirmação pode resultar em saídas não intencionais. A recomendação seria adicionar um modal de confirmação, garantindo que o usuário esteja ciente da ação que está prestes a realizar.

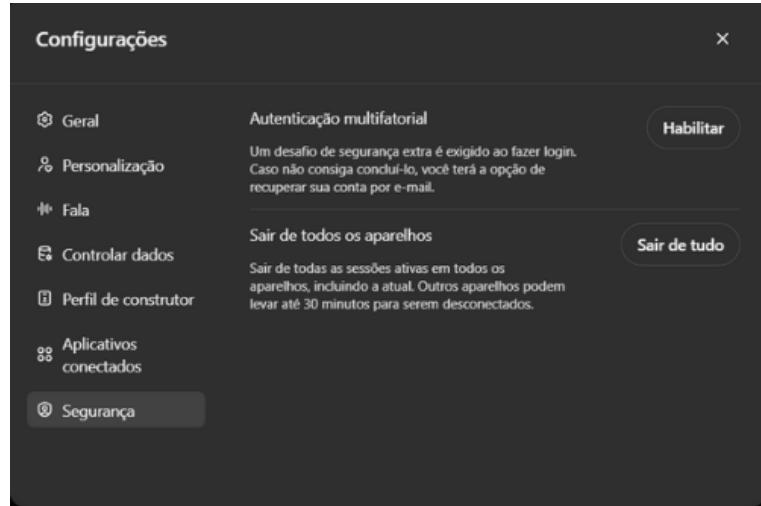
Figura 23 – Aba de funções do ChatGPT



Fonte: ChatGPT. out 2024.

Da mesma forma, a funcionalidade "Sair de todos os aparelhos" nas configurações, assim como é apresentado na figura Figura 24 de segurança não possui um diálogo de confirmação antes de executar a ação, o que pode ser problemático para o usuário. Isso também viola a heurística H05 (Prevenção de erros), pois a falta de um aviso pode levar a ações irreversíveis. A recomendação é adicionar um modal de confirmação antes de realizar o logout em todos os aparelhos.

Figura 24 – Configurações do ChatGPT



Fonte: ChatGPT. out 2024.

No caso do "Editar ChatGPT personalizado", o botão de salvar permanece desabilitado sem uma explicação clara, mesmo após a atualização das informações. De acordo com as heurísticas H04 (Consistência e padrões) e H09 (Ajuda e documentação), o usuário pode não entender o motivo pelo qual o botão continua desabilitado, o que pode gerar incerteza.

A recomendação seria adicionar um *tooltip* explicando o motivo do botão estar desabilitado, proporcionando clareza sobre o estado da interação.

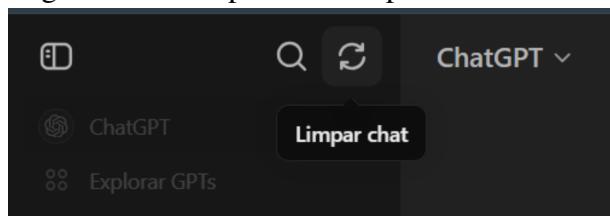
A falta da opção "Home" no menu de perfil também foi identificada. A heurística H03 (Controle e liberdade do usuário) sugere que os usuários devem ter fácil acesso a uma tela inicial. A recomendação é adicionar a opção "Home" no menu, proporcionando uma navegação mais fluida e eficiente.

Por fim, ao criar um ChatGPT personalizado, não há uma descrição sobre as mudanças que ocorreram após a criação. Isso está em desacordo com a heurística H01 (Visibilidade do status do sistema), pois o usuário não recebe feedback sobre as mudanças realizadas. A recomendação é apresentar um modal explicando as alterações após a criação, garantindo transparência sobre o que foi modificado no sistema.

5.1.1.5 Porção 5 - Seção para Alteração de Plano e Ativação de Chat Temporário

Um problema identificado refere-se ao botão de apagar o chat na seção "Chat temporário", apresentado na Figura 25. Quando clicado e sem conversas para apagar, o sistema não retorna nenhuma informação ao usuário. Isso está em desacordo com a heurística H01 (Visibilidade do status do sistema), pois o usuário não recebe um feedback claro sobre o estado da ação. A recomendação seria utilizar modais de informação para informar o estado do sistema, como "Nenhuma conversa para apagar", garantindo que o usuário saiba o que está acontecendo ou desabilitar o botão.

Figura 25 – Limpar chat temporário



Fonte: ChatGPT. out 2024.

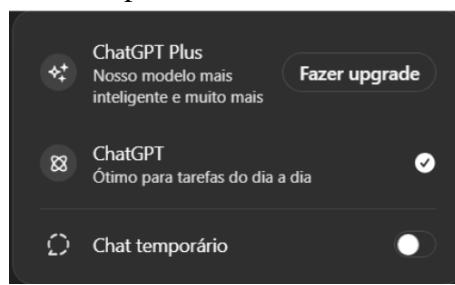
Outro ponto, ainda mostrado na Figura 25, é a localização do botão de apagar o chat temporário, que está distante da área onde ele deveria estar, de acordo com a heurística H08 (Estética e design minimalista). O afastamento do botão da sua função principal pode dificultar a interação, especialmente para usuários novatos. A recomendação é reposicionar o botão para uma localização mais intuitiva e próxima ao conteúdo relevante, tornando a navegação mais eficiente e clara.

O ícone utilizado para o botão de apagar o chat temporário, mostrado anteriormente na Figura 25, também foi mencionado, pois não deixa claro que a ação se refere à exclusão de um chat temporário. Isso viola a heurística H02 (Visibilidade do status do sistema), pois a falta de clareza no ícone pode causar confusão no usuário. A recomendação é utilizar um ícone mais representativo, que indique claramente a ação de excluir um chat temporário, facilitando a compreensão do usuário sobre a função do botão.

Além disso, a análise aponta que o ícone utilizado no menu "ChatGPT", mostrado anteriormente na Figura 18 é diferente do que aparece em outras seções do sistema, o que vai contra a heurística H04 (Consistência e padrões). A discrepância nos ícones pode gerar uma sensação de incoerência na interface, afetando a fluidez da navegação. A solução seria padronizar os ícones para que todos sigam o mesmo estilo, proporcionando uma experiência mais consistente e previsível para o usuário.

Por fim, ao clicar no botão "ChatGPT" no menu "ChatGPT", apresentado na Figura 26, o sistema não retorna nenhum feedback para o usuário caso ele esteja utilizando a versão gratuita do ChatGPT, nem informa sobre a indisponibilidade da opção de selecionar o plano pago. Isso está em desacordo com as heurísticas H01 (Visibilidade do status do sistema) e H08 (Estética e design minimalista), pois o usuário não recebe nenhuma resposta sobre sua interação, o que pode causar confusão. A recomendação seria retornar um feedback visual que explique a indisponibilidade dessa funcionalidade na versão gratuita, além de modificar o layout para tornar essa interação mais clara e evitar frustrações.

Figura 26 – Opções de upgrades e chat temporário



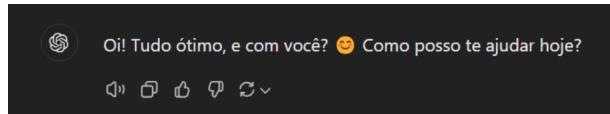
Fonte: ChatGPT. out 2024.

5.1.1.6 Porção 6 - Interface de Conversação

O primeiro problema identificado refere-se ao botão de "curtir" ou "descurtir" apresentado ao final de cada resposta do ChatGPT, assim como é apresentado na Figura 27. Ao

clicar nesses botões, a ação não pode ser desfeita, o que pode ser problemático para o usuário. De acordo com a heurística H03 (Prevenção de erros), seria mais adequado que um modal de confirmação seja exibido antes da execução da ação, evitando frustrações no caso de uma escolha acidental.

Figura 27 – Botões de respostas



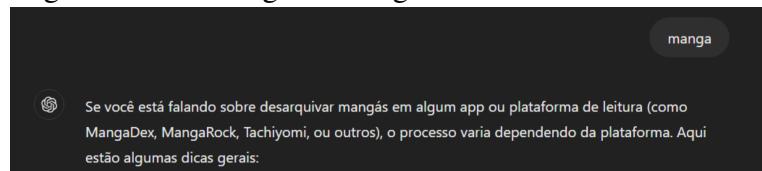
Fonte: ChatGPT. out 2024.

Na interface conversacional, foi observado que não há opção para excluir uma pergunta, apenas para editá-la. Isso representa uma limitação no controle do usuário sobre suas interações com o sistema, violando a heurística H03 (Controle e liberdade do usuário). A recomendação seria adicionar um botão de excluir pergunta, permitindo ao usuário maior controle sobre suas interações.

Outro problema é a rolagem na interface conversacional, que não acompanha automaticamente a resposta quando o ChatGPT oferece respostas longas. Esse comportamento é inconsistente com a heurística H07 (Flexibilidade e eficiência de uso), pois a rolagem deveria ser automática para garantir que o usuário possa acompanhar o fluxo de conversa sem interrupções. A solução seria ajustar a rolagem para acompanhar a resposta.

A interface conversacional também tem dificuldades ao lidar com perguntas ambíguas, o que pode levar a respostas imprecisas ou insatisfatórias, assim como mostra a Figura 28. Isso está relacionado à heurística H05 (Prevenção de erros) e H07 (Flexibilidade e eficiência de uso), pois o chatbot deveria ser capaz de identificar quando uma pergunta é ambígua e solicitar informações adicionais antes de enviar uma resposta.

Figura 28 – Mensagens ambíguas

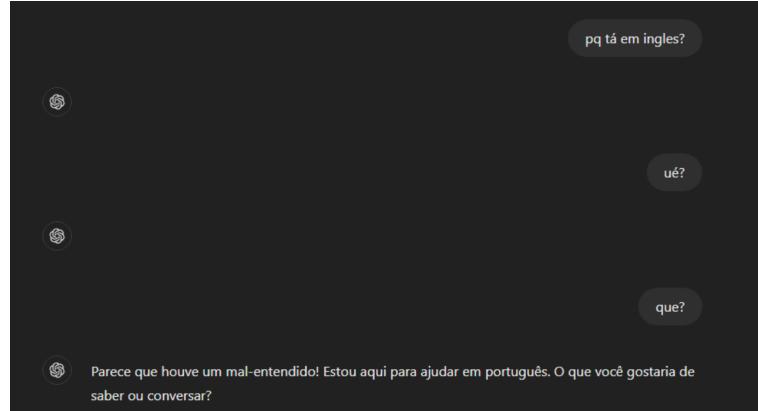


Fonte: ChatGPT. out 2024.

Além disso, em algumas situações, a interface não gera nenhuma resposta ou não retorna nenhuma informação para o usuário, como é mostrado na Figura 29, o que pode gerar confusão e frustração. Isso viola as heurísticas H01 (Visibilidade do status do sistema) e H09

(Ajuda e documentação). A solução seria exibir uma mensagem de erro ou algum outro tipo de feedback para indicar que algo deu errado.

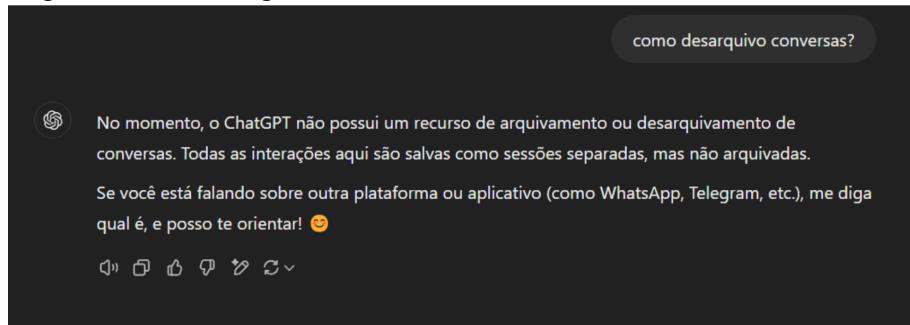
Figura 29 – Respostas invisíveis



Fonte: ChatGPT. out 2024.

O ChatGPT também não tem consciência de sua própria interface, como é mostrado na Figura 30, o que significa que, quando o usuário pergunta sobre seu funcionamento, ele não recebe uma resposta adequada. Isso está em desacordo com a heurística H10 (Ajuda e documentação), já que o chatbot deveria ser capaz de fornecer informações sobre suas funções e localização na interface. A recomendação é que o chatbot explique suas funcionalidades quando solicitado.

Figura 30 – Mensagem de dúvida sobre o ChatGPT



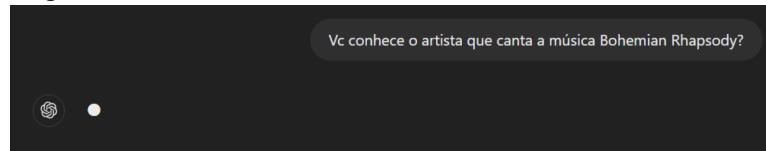
Fonte: ChatGPT. out 2024.

Outro ponto levantado foi a proximidade dos botões de "curtir" e "descurtir", que são visualmente semelhantes. De acordo com a heurística H08 (Estética e design minimalista), esses botões poderiam ser reposicionados ou ter seu layout alterado para evitar confusão entre as duas opções. Isso garantiria uma interação mais intuitiva para o usuário.

Quanto ao feedback visual durante o carregamento das respostas, foi identificado que o único indicativo de carregamento é um círculo, apresentado na Figura 31, que pode travar

em alguns momentos. Isso prejudica a visibilidade do status do sistema (H01) e pode gerar insegurança no usuário. A recomendação seria adicionar outros signos de carregamento ou indicadores mais claros para que o usuário saiba que a resposta está sendo processada.

Figura 31 – Pontos brancos



Fonte: ChatGPT. out 2024.

A interação com o ChatGPT também possui um problema quando o usuário tenta enviar um novo texto enquanto o sistema está gerando uma resposta. A heurística H01 (Visibilidade do status do sistema) sugere que o usuário seja avisado sobre essa restrição. A recomendação é adicionar um modal de aviso, informando que o envio de novos prompts é proibido enquanto uma resposta está sendo gerada, bloquear o campo ou permitir o cancelamento do processamento.

Outro problema refere-se à edição de mensagens, onde uma nova subconversa é criada, mas essa alteração não é claramente indicada ao usuário. Isso pode gerar confusão, violando a heurística H01 (Visibilidade do status do sistema). A solução seria adicionar um modal de aviso, informando ao usuário sobre a criação da nova subconversa.

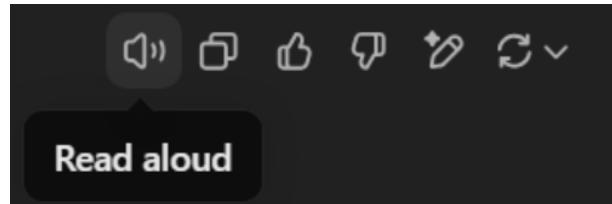
O ChatGPT possibilita que o usuário responda a trechos específicos de uma resposta anterior. No entanto, quando a mensagem é editada, a referência à resposta original é perdida, o que compromete a continuidade e a integridade da conversa. Esse problema está relacionado à heurística H04 (Consistência e padrões), que enfatiza a importância de manter a coerência ao longo da interação. A recomendação seria aprimorar o sistema para garantir que, ao editar uma mensagem, a referência à resposta original seja preservada, promovendo uma experiência mais coesa e fluida.

Além disso, ao editar mensagens, a interface tende a criar muitas ramificações na conversa, o que pode gerar uma interface confusa, especialmente em conversas longas. Isso está em desacordo com a heurística H08 (Estética e design minimalista) e H07 (Flexibilidade e eficiência de uso). A sugestão é remodelar a forma como as edições e conversas são apresentadas, criando uma estrutura mais simples e acessível.

Por fim, foi identificado que ao passar o cursor sobre o ícone para "ler o conteúdo" do chat, aparece uma legenda em inglês, assim como mostrado na Figura 32, enquanto outros ícones apresentam as legendas em português. Isso representa uma inconsistência, violando

a heurística H02 (Visibilidade do status do sistema) e H04 (Consistência e padrões), pois a interface não mantém um padrão linguístico claro, podendo confundir os usuários. Além disso, a heurística H10 (Ajuda e documentação) também é comprometida, uma vez que a legenda em um idioma diferente pode dificultar o entendimento e o acesso à informação, especialmente para usuários que não dominam o inglês. A recomendação seria traduzir a mensagem do **tooltip** para garantir consistência no idioma da interface e assegurar que todos os usuários compreendam corretamente as instruções fornecidas pelo

Figura 32 – Mensagem em Inglês GPT



Fonte: ChatGPT. out 2024.

5.1.1.7 Porção 7 - Seção Meus GPTs

O primeiro problema identificado nesta porção refere-se ao modal exibido quando o usuário seleciona um GPT em "Meus GPTs" e tenta obter mais informações por meio de redes sociais, como LinkedIn ou X (Twitter). A dificuldade de acesso ao conteúdo compromete a experiência do usuário, tornando a navegação menos intuitiva e eficiente. Para aprimorar a usabilidade, recomenda-se a correção do modal de informações extras, garantindo que ele seja mais acessível, claro e funcional, permitindo que os usuários obtenham as informações desejadas de forma fluida e sem obstáculos.

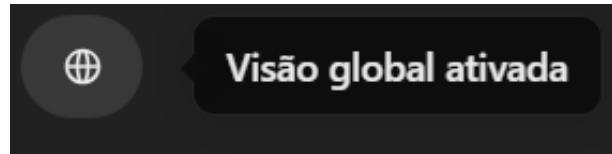
Em relação à seleção de um GPT, no plano gratuito, o usuário só pode interagir com o GPT caso não tenha excedido seus usos diários gratuitos. Contudo, essa informação é apresentada apenas no último passo de interação, o que pode causar frustração. Isso viola a heurística H07 (Flexibilidade e eficiência de uso), já que o usuário deveria ser informado sobre essa restrição logo no início do fluxo. A solução seria indicar essa limitação desde o início, proporcionando uma navegação mais transparente.

Outro problema está relacionado à rolagem das informações sobre os GPTs criados pela organização, que não pode ser visualizada através do scroll do mouse. Esse problema prejudica a flexibilidade de uso da interface, violando a heurística H07 (Flexibilidade e eficiência de uso), que enfatiza a adaptação da interface às diferentes necessidades e preferências dos

usuários. A recomendação é corrigir a funcionalidade de rolagem, permitindo que os usuários naveguem de maneira intuitiva e eficiente pelas opções.

Foi identificado também um botão no canto superior esquerdo, mostrado na Figura 33, cuja função não é explicada, o que gera incerteza para o usuário. Isso contraria a heurística H10 (Ajuda e documentação), que sugere que a interface deve fornecer clareza sobre as ações que o usuário pode realizar. A solução seria adicionar um tooltip ou substituir o ícone por um texto explicativo, tornando a função do botão mais evidente.

Figura 33 – Visão Global - Meus GPTs



Fonte: ChatGPT. out 2024.

Além disso, o botão no canto superior esquerdo utiliza um ícone do planeta Terra, que já é utilizado para outros propósitos, o que pode gerar confusão e violar a heurística H04 (Consistência e padrões). A recomendação é trocar o ícone, garantindo que cada ícone tenha um único significado.

Por fim, algumas categorias e itens estão em inglês enquanto outros estão em português, o que gera inconsistência na interface e pode dificultar a navegação dos usuários que não são fluentes em inglês. Esse problema está em desacordo com a heurística H02 (Visibilidade do status do sistema) e H04 (Consistência e padrões). A solução é traduzir todas as categorias e itens para o mesmo idioma, garantindo uma experiência mais coesa e acessível.

5.2 Teste de Usabilidade

5.2.1 Coleta e Interpretação de Dados

O Teste de Usabilidade foi realizado entre os dias 21 e 31 de outubro, logo após a consolidação dos dados da Avaliação Heurística. Nessa etapa, foram organizadas e analisadas as informações coletadas dos participantes da pesquisa, com o objetivo de traçar um perfil detalhado dos usuários. A partir dessa análise, foi possível avaliar variáveis comportamentais dos participantes, considerando métricas de eficácia e eficiência, conforme proposto por Ramos et al. (2017). A Tabela 2 apresenta os dados coletados de cada usuário.

Tabela 2 – Participantes do Teste de Usabilidade

| Identificador | Faixa etária | Área que atua |
|---------------|--------------|-------------------------------------|
| Usuário 1 | 50 e 60 anos | Educação Básica |
| Usuário 2 | 20 e 30 anos | Saúde |
| Usuário 3 | 20 a 30 anos | Saúde mental e comportamento humano |
| Usuário 4 | 40 a 50 anos | Educação Especial |
| Usuário 5 | 50 a 60 anos | Gestão Administrativa |

Fonte: Elaborado pelo autor

Essas métricas proporcionaram uma visão clara do desempenho dos participantes em relação às tarefas propostas, permitindo uma avaliação detalhada de sua interação com o sistema. A Tabela 3 apresenta os resultados de cada uma das métricas, com foco na compreensão das necessidades e dificuldades enfrentadas pelos usuários.

Tabela 3 – Tarefas e Métricas

| Usuário | Conclusão da Tarefa (Sucesso) | Tempo (segundos) | Ajuda |
|-----------------|-------------------------------|------------------|-------|
| Tarefa 1 | | | |
| Usuário 1 | Concluiu | 122 | 2 |
| Usuário 2 | Concluiu | 215 | 0 |
| Usuário 3 | Concluiu | 51 | 0 |
| Usuário 4 | Concluiu | 120 | 1 |
| Usuário 5 | Concluiu | 87 | 0 |
| Tarefa 2 | | | |
| Usuário 1 | Concluiu | 147 | 0 |
| Usuário 2 | Concluiu | 102 | 1 |
| Usuário 3 | Concluiu | 68 | 0 |
| Usuário 4 | Concluiu | 158 | 0 |
| Usuário 5 | Concluiu | 129 | 1 |
| Tarefa 3 | | | |
| Usuário 1 | Concluiu | 204 | 1 |
| Usuário 2 | Concluiu | 69 | 0 |
| Usuário 3 | Concluiu | 57 | 0 |

| | | | |
|-----------|----------|----|---|
| Usuário 4 | Concluiu | 87 | 0 |
| Usuário 5 | Concluiu | 82 | 0 |

Tarefa 4

| | | | |
|-----------|----------|-----|---|
| Usuário 1 | Concluiu | 119 | 2 |
| Usuário 2 | Concluiu | 215 | 1 |
| Usuário 3 | Concluiu | 127 | 0 |
| Usuário 4 | Concluiu | 185 | 1 |
| Usuário 5 | Concluiu | 157 | 0 |

Tarefa 5

| | | | |
|-----------|----------|----|---|
| Usuário 1 | Concluiu | 69 | 0 |
| Usuário 2 | Concluiu | 47 | 2 |
| Usuário 3 | Concluiu | 58 | 0 |
| Usuário 4 | Concluiu | 67 | 1 |
| Usuário 5 | Concluiu | 52 | 0 |

Tarefa 6

| | | | |
|-----------|--------------|-----|---|
| Usuário 1 | Concluiu | 206 | 4 |
| Usuário 2 | Não Concluiu | 618 | 5 |
| Usuário 3 | Concluiu | 127 | 0 |
| Usuário 4 | Não Concluiu | 489 | 3 |
| Usuário 5 | Concluiu | 204 | 0 |

Fonte: Elaborado pelo autor

Nas primeiras cinco tarefas, todos os usuários conseguiram completá-las com sucesso, o que demonstra que as funcionalidades básicas do sistema são compreensíveis e podem ser executadas sem grandes obstáculos. No entanto, houve uma variação significativa tanto nos tempos de execução quanto na necessidade de ajuda entre os participantes. Na Tarefa 1, por exemplo, o tempo de conclusão variou de 51 a 215 segundos. O Usuário 2 foi o que levou mais tempo para concluir a tarefa, principalmente porque demonstrou dúvidas sobre como formular a pergunta inicial e mostrou-se excessivamente preocupado com possíveis erros ortográficos. Além disso, o número de solicitações de ajuda variou de 0 a 2 vezes, dependendo do participante. Essas discrepâncias podem ser explicadas por fatores como o nível de familiaridade de cada usuário com o uso de computador, sendo que os usuários 1 e 4, por terem menos familiaridade

com a tecnologia, podem ter enfrentado maiores dificuldades ao interagir com o sistema. A clareza (ou falta de clareza) das instruções fornecidas e até mesmo a confiança individual de cada um em sua capacidade de interagir com a interface também desempenham um papel importante. Já os outros usuários, que já utilizam tecnologia no dia a dia, demonstraram maior facilidade e confiança ao navegar pela interface.

Já a Tarefa 6, que envolvia desarquivar uma conversa, se mostrou mais desafiadora. Dois usuários não conseguiram concluir-la, e os tempos de execução dos que finalizaram a tarefa foram significativamente mais altos, variando de 127 a 618 segundos. Além disso, a necessidade de ajuda foi mais frequente, com solicitações que variaram de 0 a 5 vezes. Esses resultados sugerem que o processo de desarquivamento pode ser complexo ou pouco intuitivo, indicando a necessidade de ajustes para tornar a experiência do usuário mais fluida e acessível.

Ao analisar as métricas de usabilidade, como a taxa de sucesso nas tarefas e o tempo de conclusão, fica evidente que esses são indicadores essenciais para avaliar a eficácia do sistema. A variação nos tempos de execução e no número de solicitações de ajuda reforça a importância de considerar as diferenças individuais dos usuários ao avaliar a usabilidade.

5.2.1.1 *Taxa de Sucesso por Tarefa*

A taxa de sucesso representa a proporção de usuários que completaram a tarefa com êxito, refletindo a eficácia da interface em permitir a realização das atividades de forma satisfatória (Nielsen, 1994). Essa métrica é essencial para avaliar a usabilidade de um sistema, pois indica se os usuários conseguem atingir seus objetivos sem dificuldades significativas. Foi calculado a métrica para cada tarefa com a fórmula:

$$\text{Taxa de Sucesso (\%)} = \frac{\text{Nº de Sucessos}}{\text{Total de Tentativas}} \times 100 \quad (5.1)$$

Tabela 4 – Taxa de sucesso por tarefa

| Tarefa | Total de Usuários | Concluíram com Sucesso | Taxa de Sucesso (%) |
|----------|-------------------|------------------------|---------------------|
| Tarefa 1 | 5 | 5 | 100% |
| Tarefa 2 | 5 | 5 | 100% |
| Tarefa 3 | 5 | 5 | 100% |
| Tarefa 4 | 5 | 5 | 100% |
| Tarefa 5 | 5 | 5 | 100% |
| Tarefa 6 | 5 | 3 | 60% |

Fonte: Elaborado pelo autor

Ao analisar, podemos concluir que todas as tarefas, com exceção da Tarefa 6, atingiram uma taxa de sucesso de 100%. Esse desempenho sugere que a maioria dos participantes, apesar de solicitarem ajuda, conseguiram concluir o que foi pedido durante o Teste de Usabilidade. Entretanto, a Tarefa 6 apresentou uma performance significativamente inferior, com uma taxa de sucesso de apenas 60%, o que também implicou no tempo e na quantidade de "ajuda" que os usuários solicitaram.

5.2.1.2 *Tempo Médio de Execução por Tarefa*

O tempo médio de execução reflete a eficiência dos usuários ao concluir as tarefas, indicando a rapidez com que eles conseguem realizar as ações necessárias dentro da interface (Nielsen; Landauer, 1993). Ele foi calculado com a fórmula:

$$\text{Tempo Médio (segundos)} = \frac{\text{Soma do Tempo de Execução}}{\text{Total de Usuários que Concluíram}} \quad (5.2)$$

Tabela 5 – Tempo médio de execução por tarefa

| Tarefa | Tempo Total (s) | Usuários que Concluíram | Tempo Médio (s) |
|-------------|-----------------|-------------------------|-----------------|
| Tarefa 1 | 575 | 5 | 119 |
| Tarefa 2 | 249 | 2 | 120,8 |
| Tarefa 3 | 640 | 5 | 99,8 |
| Tarefa 4 | 839 | 5 | 160,6 |
| Tarefa 5 | 669 | 5 | 58,6 |
| Tarefa 6 | 734 | 2 | 238,8 |
| Média geral | | | 152,6 |

Fonte: Elaborado pelo autor

A Tarefa 3 apresentou o menor tempo médio de execução, com 99,8 segundos. Como as Tarefas 1 e 2 eram similares à Tarefa 3, é possível que essa melhoria no tempo seja reflexo de maior aprendizado ou familiaridade adquirida durante a execução das tarefas anteriores. Já Tarefa 6, por outro lado, teve o maior tempo médio de execução, com 238,8 segundos, quase o dobro da média geral (152,6 segundos). Esse dado reforça a ideia de que essa tarefa apresenta desafios significativos para os usuários. As demais tarefas (2, 3, 4 e 5) apresentaram tempos médios relativamente próximos, variando entre 124,5 e 167,8 segundos. Esses valores estão dentro de uma faixa aceitável e indicam que essas tarefas são equilibradas em termos de tempo de execução.

5.2.1.3 *Solicitação de Ajuda por Tarefa*

O número de vezes em que os usuários solicitam ajuda reflete diretamente a complexidade de cada tarefa, indicando a dificuldade percebida durante a execução das ações e a clareza das funcionalidades da interface (Nielsen; Landauer, 1993). A ajuda média foi calculada com a fórmula:

$$\text{Ajuda Média} = \frac{\text{Nº Total de Pedidos de Ajuda}}{\text{Total de Usuários}} \quad (5.3)$$

Tabela 6 – Solicitação de ajuda por tarefa

| Tarefa | Total de Ajuda | Usuários | Ajuda Média |
|----------|----------------|----------|-------------|
| Tarefa 1 | 3 | 5 | 0,6 |
| Tarefa 2 | 2 | 2 | 0,4 |
| Tarefa 3 | 1 | 5 | 0,2 |
| Tarefa 4 | 4 | 5 | 0,8 |
| Tarefa 5 | 5 | 5 | 0,6 |
| Tarefa 6 | 12 | 5 | 2,4 |

Fonte: Elaborado pelo autor

A Tarefa 6 evidenciou-se como a mais desafiadora, com uma média de 2,4 solicitações de ajuda por usuário, indicando elevada complexidade ou menor intuitividade em sua execução. Por outro lado, a Tarefa 3 apresentou a menor média, com apenas 0,2 solicitações de ajuda por usuário, o que pode ser atribuído à maior simplicidade da tarefa ou à familiaridade progressivamente adquirida pelos participantes ao longo da realização das atividades anteriores.

5.2.1.4 *Entrevistas*

A Tabela 7 apresenta um resumo das respostas coletadas durante as entrevistas realizadas com os participantes do teste de usabilidade do ChatGPT. O objetivo principal do teste foi avaliar a experiência dos usuários ao interagir com o sistema, identificando pontos positivos, dificuldades enfrentadas, bem como possíveis melhorias.

Os participantes, provenientes de diferentes faixas etárias e áreas de atuação, forneceram informações detalhadas sobre aspectos como facilidade de uso, clareza nas respostas, e comparação com outras ferramentas similares. Além disso, foi explorada a percepção sobre a relevância e confiabilidade das respostas oferecidas pelo ChatGPT, bem como sugestões para aprimorar a usabilidade da ferramenta.

A tabela organiza as informações em categorias específicas, como:

1. Experiência Geral: Primeiras impressões e desafios ao utilizar o ChatGPT.
2. Facilidade de Uso: Avaliação da interface e facilidade de navegação.
3. Clareza e Relevância nas Respostas: Qualidade e adequação das informações fornecidas.
4. Comparação com Outras Ferramentas: Experiência dos usuários com ferramentas semelhantes.
5. Sugestões de Melhoria: Recomendações dos participantes para aprimorar a ferramenta.
6. Comentários Finais: Observações adicionais sobre a experiência geral.

Esse levantamento fornece uma visão ampla e detalhada sobre as percepções e necessidades dos usuários, servindo como base para futuros aprimoramentos no sistema e na experiência oferecida pelo ChatGPT.

Tabela 7 – Perguntas e Respostas - Entrevista

| Usuários | Respostas |
|---|--|
| 1. Experiência Geral: | |
| Você usa algum sistema ou aplicativo no seu trabalho? Como ele é? | |
| Usuário 1 | "Sim, é um sistema para colocar notas de alunos, provas, atividades e também cadastrar presença (Aluno Online)." |
| Usuário 2 | "Sim, um sistema de estoque e pedido de remédios para serem distribuídos para a população, esse sistema também armazena receitas, e verifica usuários frequentes." |
| Usuário 3 | "Sim, uso para arquivar alguns documentos da empresa." |
| Usuário 4 | "Sim, preciso informar sobre problemas e escrever relatórios sobre o avanço dos meus alunos para colocar no sistema da escola." |
| Usuário 5 | "Sim, é um sistema bem completo e complexo, eu uso para cadastrar, checar estoque, emitir notas fiscais e vender os produtos da loja." |
| Como você descreveria sua experiência geral ao utilizar o ChatGPT? | |
| Houve algo que se destacou para você? | |
| Usuário 1 | "Gostei da rapidez das respostas, com certeza iria otimizar um tempo gigantesco na hora de fazer algumas atividades do meu dia." |

| | |
|-----------|--|
| Usuário 2 | "No geral eu gostei da rapidez das respostas, mas algumas coisas nele ainda parecem um pouco confusas, não sei o que posso perguntar, o que não posso... Ele não oferece um "tutorial" nem informação nenhuma nesse início." |
| Usuário 3 | "Gostei, pelo fato dele ser bem rápido, não é didático, porque preciso voltar várias vezes, e tem processos longos, mas é bem rápido nas respostas." |
| Usuário 4 | "É realmente bem novo e diferente do Google, não gostei muito porque não é fácil e me deixa um pouco aflita." |
| Usuário 5 | "Gostei do fato de ser uma espécie de "conversa", mas ainda não me passa muita confiança, me sinto estranha em conversar com um robô. Me senti receosa em informar minhas dúvidas." |

2. Facilidade de Uso:

O quanto fácil foi para você interagir com o ChatGPT? Quais dificuldades, se houve alguma, você enfrentou?

| | |
|-----------|--|
| Usuário 1 | "Foi difícil, não tenho experiência com essas coisas, pareceu algo de "outro mundo", mas com ajuda parece menos difícil." |
| Usuário 2 | "Achei "ok", já tinha ouvido falar do ChatGPT, mas nunca tinha usado, essa primeira experiência me deu um pouco de receio, mas espero conseguir aprender mais depois." |
| Usuário 3 | "Não gostei muito, tive dificuldade de formular as perguntas, porque quando eu entrei não entendi nada, sabe? Ai não sabia como começar a perguntar o que eu queria." |
| Usuário 4 | "Não achei tão fácil, tive bastante dificuldade, especialmente na hora de arquivar e desarquivar os chats." |
| Usuário 5 | "Acho que foi uma experiência um pouco "chata", pois me deixou um pouco confusa, as funções são diferentes de como eu esperava e a parte de desarquivar é praticamente um labirinto" |

Você achou fácil encontrar as funcionalidades desejadas no ChatGPT?

| | |
|-----------|---|
| Usuário 1 | "Algumas sim, mas não gostei da função de criar uma conversa nova, é escondido, não tem um nome que informe o que vai acontecer,, Também não gostei de desarquivar, isso não é para mim." |
| Usuário 2 | "Não muito, fiquei nervoso na hora de desarquivar, não apareceu nada informando se tinha arquivado, não sabia onde achar, demorei bastante, o que parecia fácil ficou muito confuso." |
| Usuário 3 | "A de excluir é bem familiar, mas a de criar um novo chat e de desarquivar estranhas, principalmente a de desarquivar, é difícil de achar." |
| Usuário 4 | "Achei dificultoso, poderia ter informações mais claras e simples de entender." |
| Usuário 5 | "Algumas são bem escondidas, principalmente quando fui desarquivar a conversa, nunca imaginei que ia ser ali." |

3. Clareza e Relevância nas Respostas:

As respostas do ChatGPT foram claras e compreensíveis? Houve

algum momento em que você achou as respostas confusas ou vagas?

| | |
|-----------|---|
| Usuário 1 | "Sim, foram respostas boas e coesas." |
| Usuário 2 | "Sim, consegui entender o que ele quis passar." |
| Usuário 3 | "Achei as respostas bem objetivas, gostei de como ele respondeu, mas em alguns pontos poderiam ser mais extensas, dependendo do conhecimento de quem pergunta." |
| Usuário 4 | "Gostei das respostas, achei que elas responderam bem as minhas dúvidas." |
| Usuário 5 | "Sim, sinto que foram bem claras, rasas mas ainda sim claras." |

Como você avalia a relevância das respostas fornecidas pelo

ChatGPT em relação às suas perguntas? Houve alguma resposta

que não atendeu às suas expectativas?

| | |
|-----------|--|
| Usuário 1 | "Achei que me ajudaram bastante, facilitariam muito se usasse no dia-a-dia." |
|-----------|--|

| | |
|-----------|---|
| Usuário 2 | "Em relação a isso, achei que as respostas, apesar de objetivas, foram fracas, elas poderiam ter mais contexto, explicar melhor alguns pontos, porque dependendo do conhecimento da pessoa que está perguntando podem existir algum mal entendido." |
| Usuário 3 | "Achei que faltou fundamento, tipo, de onde ele tirou isso? Eu queria saber para poder confiar mais." |
| Usuário 4 | "Perguntei sobre as atualidades do nosso país, porem ele me retornou 2 notícias, isso me deixou um pouco descontente, porque eu esperava um grande resumo do ultimo mês." |
| Usuário 5 | "As respostas supriram minhas expectativas, mas elas poderiam ter algum "fundamento", pois ele não fala de onde tirou a informação." |

5. Comparação com Outras Ferramentas:

Se você já utilizou outras ferramentas semelhantes ao ChatGPT, como você compara a experiência? O que o ChatGPT faz melhor ou pior?

| | |
|-----------|---|
| Usuário 1 | "Só a Cortana. O ChatGPT, é mais rápido, tem uma forma diferente." |
| Usuário 2 | "Já usei a Siri no IOS, mas acho o ChatGPT, rápido, mesmo sendo mais difícil de acessar, ele também é menos "simpático"." |
| Usuário 3 | "Só a Cortana e a assistente de voz do Google, no celular, são bem diferentes do Chat porque são mais simples, porém gosto mais delas." |
| Usuário 4 | "Nunca tive contato com esse tipo de tecnologia, mas gostei de como essa funciona." |
| Usuário 5 | "Já usei a assistente da Google no celular, é mais simples que o ChatGPT, apesar de não dar informações tão completas de forma rápida." |

6. Expectativas Atendidas:

Você consideraria usar o ChatGPT regularmente com base na experiência que teve durante este teste de usabilidade?

| | |
|-----------|---|
| Usuário 1 | "Não, apesar de ser bom, eu acho que não conseguiria fazer tudo sozinha, tive muita dificuldade." |
|-----------|---|

| | |
|-----------|---|
| Usuário 2 | "Sim, talvez em alguma situação mais simples do dia-a-dia, mas não usaria no trabalho por ser da área da saúde, seria totalmente irresponsável da minha parte." |
| Usuário 3 | "Não, por ele não ser didático, prefiro pesquisar no Google mesmo, assim tenho mais segurança." |
| Usuário 4 | "Talvez, ele parece bom, mas ainda sinto muito receio de usar sozinha." |
| Usuário 5 | "Não, tenho mais segurança pesquisando nas fontes convencionais." |

7. Sugestões de Melhoria:

Você tem alguma sugestão ou recomendação para melhorar a usabilidade e eficácia do ChatGPT?

| | |
|-----------|--|
| Usuário 1 | "Tudo poderia ser mais fluido. Me pareceu que estou falando com um sistema rígido que não se adapta bem a mudanças no fluxo da conversa." |
| Usuário 2 | "Ele deveria ser mais transparente sobre suas limitações. Por exemplo, informar claramente quando não pode responder algo em vez de tentar gerar uma resposta que pode ser imprecisa." |
| Usuário 3 | "O ChatGPT às vezes não é muito assertivo em suas respostas. Uma sugestão seria priorizar respostas mais focadas, evitando rodeios ou informações irrelevantes que podem confundir o usuário." |
| Usuário 4 | "Ele me deu algumas respostas vagas, que complicou meu entendimento. Ser mais direto e oferecer respostas objetivas seria uma grande melhoria na experiência." |
| Usuário 5 | "Poderia variar mais nas respostas, parece algo engessado, com informações muito simples e sem detalhes, e também poderia ensinar algo para usuários novos." |

8. Comentários Finais:

Há mais algum comentário que você gostaria de fazer sobre sua experiência com o ChatGPT?

| | |
|-----------|--|
| Usuário 1 | "Gostei do ChatGPT, mas entendi que ele tem limites, quando se trata de informações muito recentes ou de opiniões mais subjetivas pode ser que não seja a melhor opção." |
|-----------|--|

| | |
|-----------|---|
| Usuário 2 | "Por mais que o ChatGPT seja útil em várias situações, às vezes ele não consegue responder com profundidade, e até oferecer informações atuais. Senti que falta precisão em temas mais complexos." |
| Usuário 3 | "O ChatGPT é interessante, mas senti que ele não consegue captar nuances ou interpretar contextos mais subjetivos. Isso pode tornar algumas respostas muito mecânicas, o que limita sua utilidade em conversas mais profundas." |
| Usuário 4 | "Usar o ChatGPT é como ter um assistente sempre à disposição. Ele ajuda a economizar tempo com pesquisas e esclarece dúvidas rapidamente. O único ponto a melhorar talvez seja a facilidade de algumas funções." |
| Usuário 5 | "De início eu gostei, mas às vezes ele parece superficial, e as vezes difícil demais." |

Fonte: Elaborado pelo autor

Com base nas respostas apresentadas na Tabela 7, a maioria destacou a rapidez das respostas como um ponto positivo, mas também apontou dificuldades, como a falta de um tutorial inicial e dúvidas sobre o que perguntar.

A facilidade de uso foi um desafio para alguns, especialmente na navegação e no arquivamento de chats, fato que é citado várias vezes durante a avaliação heurística, na subseção 5.1.1.2. Alguns usuários tiveram dificuldades em formular perguntas e sentiram que faltava clareza nas funcionalidades. Quanto à clareza e relevância das respostas, a maioria achou as respostas compreensíveis, mas alguns mencionaram que poderiam ser mais completas ou ter mais embasamento.

Comparado a outras ferramentas, como assistentes de voz, o ChatGPT foi visto como mais rápido, mas menos intuitivo e menos "simpático". Sobre expectativas, poucos usariam a ferramenta regularmente, principalmente por receio ou por não se sentirem confiantes no uso.

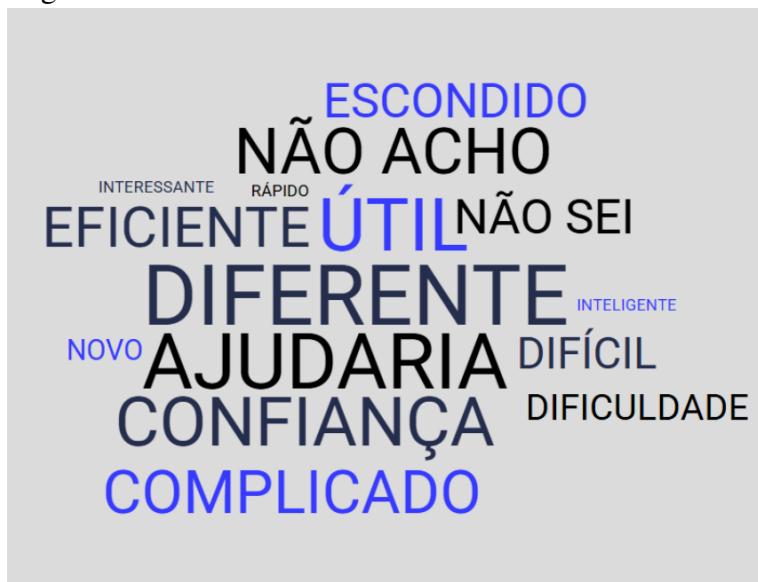
Entre as sugestões de melhoria, destacam-se tornar as respostas mais adaptáveis ao contexto, melhorar a transparência sobre limitações e tornar a navegação mais intuitiva. No geral, a experiência foi mista, com pontos positivos na rapidez e pontos negativos na usabilidade e profundidade das respostas.

Analizando as entrevistas e comentários realizados durante o teste, foi criada uma

nuvem de palavras, mostrada na Figura 34, que reflete a experiência do usuário durante o primeiro uso do ChatGPT. Esse conjunto de palavras visa ilustrar de forma concisa e visual as impressões mais comuns, sentimentos e reações observadas, oferecendo uma representação clara e acessível da interação inicial com a ferramenta.

As palavras refletidas na nuvem abrangem tanto desafios enfrentados pelos usuários, como "difícil" e "complicado", quanto aspectos positivos, como "rápido", "útil" e "inteligente". A combinação desses termos oferece uma visão holística sobre a experiência, destacando tanto áreas que podem exigir melhorias quanto os pontos fortes que foram mais apreciados. Essa análise proporciona percepções valiosas para aprimorar a interface e a usabilidade do ChatGPT, visando sempre uma experiência mais intuitiva e eficaz para os usuários.

Figura 34 – Nuvem de Palavras



Fonte: Elaborado pelo autor

5.2.2 Discussão dos dados

Nessa parte, foi realizada a triangulação de dados, combinando as abordagens de análise heurística, teste de usabilidade e entrevistas com os usuários. Essa combinação permitirá integrar dados qualitativos e quantitativos. Ao comparar os resultados obtidos por meio desses diferentes métodos, será possível validar as conclusões e obter uma visão mais clara sobre a eficácia da interface, suas dificuldades e oportunidades de melhoria, além de proporcionar uma compreensão mais rica sobre a experiência do usuário.

A Tabela 8, apresenta as percepções qualitativas e as métricas quantitativas rela-

cionadas ao teste de usabilidade. As percepções qualitativas são baseadas nas observações e comentários dos usuários durante as tarefas, enquanto as métricas quantitativas, como a taxa de sucesso, o tempo médio de execução e o número de solicitações de ajuda, oferecem dados numéricos que refletem a eficiência e a satisfação dos usuários com a interface. Juntas, essas informações permitem uma análise detalhada da experiência do usuário e ajudam a identificar áreas para melhorias.

Tabela 8 – Percepções Qualitativas e Métricas Quantitativas

| Questões Citadas (dores citadas pelos usuários durante o Teste de Usabilidade e a entrevista) | Usuários com Dificuldade | Tarefa Correspondente | Taxa de Sucesso (%) | Auxílio Média |
|--|---------------------------------|------------------------------|----------------------------|----------------------|
| Dificuldade em entender a interface | 5 | Tarefa 6 | 60% | 2,4 |
| Falta de clareza nas respostas | 2 | Tarefa 3 | 100% | 0,2 |
| Problemas com botões escondidos | 5 | Tarefas 4 e 6 | 100% | 0,8 |
| Navegação confusa | 4 | Tarefa 5 | 100% | 0,6 |

Fonte: Elaborado pelo autor

A principal dificuldade identificada foi a "dificuldade em compreender a interface", relatada por 5 usuários. Esse problema também foi evidenciado durante a avaliação heurística, que apontou questões como botões pouco visíveis, funções pouco intuitivas e ícones de difícil interpretação, aspectos que podem ser revisitados no Capítulo 5. Essas informações indicam que a clareza e a intuitividade da interface do sistema são fatores críticos para a usabilidade, podendo impactar diretamente a eficácia dos usuários na execução das tarefas. A taxa de sucesso de 60% associada a essa questão sugere que, embora os usuários tenham conseguido concluir as tarefas, a experiência foi significativamente prejudicada pela falta de compreensão da interface, reforçando a necessidade de melhorias nesse aspecto para garantir uma interação mais fluida e satisfatória.

As tarefas relacionadas à "Falta de clareza nas respostas" e aos "Problemas com botões escondidos"—também identificados na avaliação heurística—apresentaram uma taxa de sucesso de 100%. Isso sugere que, apesar das dificuldades enfrentadas, os usuários conseguiram concluir as tarefas, possivelmente devido a fatores como treinamento prévio ou familiaridade com o sistema. No entanto, a taxa de sucesso mais baixa observada (60%) indica que a interface pode exigir melhorias significativas para reduzir a carga cognitiva dos usuários e proporcionar uma experiência mais intuitiva.

A média de ajuda solicitada para a dificuldade de "Entender a interface" foi 2,4, um

valor relativamente alto, o que sugere que os usuários tiveram que buscar ajuda frequentemente para completar a tarefa. Isso destaca a importância de um design mais intuitivo, com melhor sinalização e feedback visual, a fim de reduzir a dependência de assistência externa.

A Tabela 9, oferece uma análise detalhada sobre as dificuldades e o desempenho de diferentes usuários ao interagirem com a interface de uma ferramenta. O objetivo é identificar os principais desafios encontrados durante o uso, o desempenho de cada usuário nas tarefas propostas, o tempo médio gasto para realizar essas tarefas e a quantidade de ajuda solicitada em média. As colunas incluem as dificuldades relatadas, destacando aspectos como a falta de experiência com tecnologia, a confusão na navegação e problemas de confiança nas respostas, entre outros.

Tabela 9 – Análise por Usuário – Dificuldades e Desempenho

| Usuário | Principais Dificuldades Relatadas | Tarefa com Maior Dificuldade | Taxa de Sucesso Geral (%) | Tempo Médio Geral (s) | Ajuda Média Geral |
|----------------|---|-------------------------------------|----------------------------------|------------------------------|--------------------------|
| Usuário 1 | Falta de experiência com tecnologia, interface confusa | Tarefa 6 | 83,33% | 144,5 | 1,67 |
| Usuário 2 | Interface rígida, dificuldades para desarquivar conversas | Tarefa 6 | 83,33% | 211,0 | 1,67 |
| Usuário 3 | Respostas pouco confiáveis, dificuldade em formular perguntas | Tarefa 2 e 3 | 100% | 81,33 | 0 |
| Usuário 4 | Navegação pouco intuitiva, falta de informações claras | Tarefa 6 | 100% | 184,33 | 1,17 |

| | | | | | |
|-----------|---|----------|------|-------|------|
| Usuário 5 | Perda de confiança na interface, funções escondidas | Tarefa 6 | 100% | 118,5 | 1,17 |
|-----------|---|----------|------|-------|------|

Fonte: Elaborado pelo autor

A análise da tabela revela uma diversidade de dificuldades enfrentadas pelos usuários, que variam de falta de experiência com tecnologia até problemas relacionados à clareza e navegação da interface. O Usuário 1 destacou limitações relacionadas à inexperiência com tecnologia e uma interface confusa, alcançando uma taxa de sucesso de 83,33%, com um tempo médio de 144,5 segundos e uma dependência de ajuda média de 1,67, o que sugere que este perfil se beneficiaria de um design mais intuitivo.

O Usuário 2 também apresentou uma taxa de sucesso de 83,33%, mas com um tempo médio de 211 segundos, o mais elevado entre os usuários, indicando que o esforço cognitivo necessário para completar as tarefas foi maior. Sua ajuda média de 1,5 reforça a necessidade de simplificar a interface e melhorar as instruções.

O Usuário 3, por outro lado, obteve 100% de sucesso e registrou o menor tempo médio geral (81,33 segundos) e ajuda média (0). Isso sugere que, apesar de relatar dificuldades em formular perguntas e lidar com respostas inconsistentes, sua interação foi menos dependente de suporte externo, possivelmente devido à maior familiaridade com a interface.

Os Usuários 4 e 5 também alcançaram 100% de sucesso, demonstrando que foram capazes de realizar todas as tarefas com êxito. O Usuário 4 registrou um tempo médio de 184,33 segundos e uma ajuda média de 1,17, enquanto o Usuário 5 completou as tarefas com um tempo médio menor (118,5 segundos) e uma ajuda média similar (1,17). Apesar do sucesso, ambos relataram desafios relacionados à navegação pouco intuitiva e funções escondidas, indicando que ajustes no design poderiam beneficiar até mesmo usuários mais eficientes.

A Tarefa 6 foi apontada como a mais desafiadora para os Usuários 1, 2 e 3, enquanto os Usuários 4 e 5 conseguiram completá-la sem dificuldades aparentes. Esse padrão ressalta a importância de ajustar áreas específicas da interface para melhorar a acessibilidade e a experiência de todos os perfis de usuários.

A análise geral reforça a necessidade de um design mais claro, intuitivo e inclusivo, especialmente para usuários com menos experiência tecnológica ou que enfrentam barreiras

cognitivas ao utilizar a interface. Focar em melhorias direcionadas, como reduzir o esforço necessário para completar tarefas e aumentar a visibilidade de funcionalidades, pode gerar ganhos significativos em usabilidade.

A Tabela 10 apresenta as violações de heurísticas de usabilidade identificadas durante a avaliação da interface, com foco no impacto dessas falhas na experiência do usuário. As heurísticas, como estabelecido por Nielsen, são princípios fundamentais para criar interfaces eficazes e intuitivas. Ao identificar a frequência dessas violações e o impacto que causam na interação do usuário, é possível propor melhorias para otimizar a experiência de uso.

Tabela 10 – Impacto das Violações de Heurísticas na Satisfação do Usuário

| Heurística Violada | Frequência (%) | Exemplos de Problemas | Impacto na Experiência do Usuário |
|---------------------------------------|-----------------------|---|---|
| H01: Visibilidade do Status | 19,61% | Ausência de feedback visual ao arquivar/desarquivar; rolagem não automática em respostas longas | Gera confusão e insegurança; dificulta a compreensão do estado atual da interface |
| H02: Compatibilidade com o Mundo Real | 9,80% | Ícones pouco intuitivos, como o ícone de planeta utilizado para múltiplas funções | Reduz a familiaridade e aumenta a curva de aprendizado |
| H03: Controle e Liberdade do Usuário | 9,80% | Falta de confirmação para ações críticas (excluir todos os chats, sair de sessão); impossibilidade de desfazer curtidas ou respostas editadas | Diminui a confiança do usuário e aumenta o risco de erros irreversíveis |
| H04: Consistência e Padrões | 11,76% | Ícones e mensagens inconsistentes (uso de idiomas mistos: português e inglês) | Compromete a fluidez e aumenta o esforço cognitivo para navegar pela interface |

| | | | |
|---|--------|--|---|
| H05: Prevenção de Erros | 11,76% | Input inválido permitido (ex.: domínios incorretos); exclusão não confirmada de chats arquivados | Provoca frustração e reduz a confiabilidade da interface |
| H06: Reconhecimento ao invés de Memorização | 1,96% | Opções importantes escondidas em menus não intuitivos (ex.: botão de desarquivar); ausência de descrições em ícones como o "planeta" | Aumenta a dificuldade de localizar funções, prejudicando a navegação fluida |
| H07: Flexibilidade e Eficiência de Uso | 1,96% | Ausência de ações em massa (ex.: desarquivar todos os chats de uma vez); rolagem impossível para usuários de mouse | Reduz a eficiência, especialmente para usuários avançados |
| H08: Estética e Design Minimalista | 17,65% | Botões pouco contrastantes no modo claro; ícones pequenos e discretos como o de "ajuda" | Prejudica a visibilidade e a acessibilidade, dificultando a interação |
| H09: Ajuda e Documentação | 7,84% | Falta de informações claras sobre limitações da interface; ausência de mensagens explicativas em erros | Gera frustração para usuários novatos que precisam de suporte adicional |

| | | | |
|--|-------|---|---|
| H10: Ajuda ao Reconhecer e Recuperar Erros | 7,84% | Mensagens de erro pouco claras ou ausentes (ex.: ao editar uma mensagem, não há explicação sobre a criação de subconversas) | Complica a recuperação de erros, deixando o usuário inseguro e desorientado |
|--|-------|---|---|

Fonte: Elaborado pelo autor

A heurística "Visibilidade do Status" (H01) aparece como a mais violada, com 19,61% dos problemas. A ausência de feedback visual em ações como arquivar e desarquivar, além da rolagem não automática em respostas longas, gera confusão e insegurança, dificultando a compreensão do estado atual da interface. Durante o Teste de Usabilidade, mais da metade dos usuários sentiram dificuldade e pediram ajuda pois o sistema não mostrou informações de onde encontrar os chats arquivados. Esse problema é crítico, pois interfere diretamente na confiança do usuário no sistema.

"Compatibilidade com o Mundo Real" (H02), embora menos frequente (9,80%), apresenta problemas significativos, como o uso de ícones pouco intuitivos, especialmente o ícone de planeta com múltiplas funções. Essa violação dificulta o aprendizado, pois quebra a familiaridade entre elementos da interface e o contexto real, aumentando o esforço cognitivo do usuário.

A heurística "Controle e Liberdade do Usuário" (H03), com a mesma frequência de 9,80%, é violada por ações que não possuem confirmação, como exclusão de chats ou saída de sessão, além da impossibilidade de desfazer curtidas ou edições. Isso diminui a confiança do usuário, gerando frustração e risco de danos irreversíveis.

"Consistência e Padrões" (H04) e "Prevenção de Erros" (H05) compartilham a mesma taxa de violação, 11,76%. A inconsistência em ícones e mensagens, incluindo o uso de idiomas mistos, compromete a fluidez da navegação e aumenta o esforço necessário para compreender a interface. Já a aceitação de inputs inválidos e a exclusão de chats sem confirmação comprometem a confiabilidade, gerando frustração ao permitir erros que poderiam ser evitados.

A heurística "Reconhecimento ao invés de Memorização" (H06), com apenas 1,96%, apresenta um impacto menor, mas ainda assim compromete a navegação do sistema. Botões importantes estão escondidos em menus pouco intuitivos, e ícones como o "planeta" carecem de

descrições claras, dificultando a localização das funções, especialmente para usuários menos experientes.

"Flexibilidade e Eficiência de Uso" (H07), também com 1,96%, apresenta problemas como a ausência de ações em massa e rolagem ineficiente para usuários de mouse, reduzindo a eficiência do sistema, principalmente para usuários avançados. Durante o teste de usabilidade, na primeira tarefa, o usuário 1 teve dificuldade para visualizar a informação completa. A rolagem ineficiente afetou a experiência, tornando a navegação confusa e dificultando o acesso ao conteúdo necessário.

"Estética e Design Minimalista" (H08), com 17,65%, é a segunda heurística mais violada. Problemas como botões pouco contrastantes e ícones discretos dificultam a visibilidade e acessibilidade da interface, prejudicando tanto iniciantes quanto pessoas com limitações visuais.

As heurísticas "Ajuda e Documentação" (H09) e "Ajuda ao Reconhecer e Recuperar Erros" (H10), ambas com 7,84%, refletem a falta de suporte para usuários em situações de erro ou incerteza. Mensagens de erro pouco claras ou inexistentes dificultam a recuperação de falhas e a compreensão das limitações da interface, deixando os usuários desorientados e frustrados.

De modo geral, as violações mais frequentes estão relacionadas à clareza do feedback da interface e à acessibilidade visual, sendo H01 e H08 os principais pontos críticos. Problemas como falta de consistência, controle e suporte tornam a experiência frustrante, especialmente para usuários com conhecimento limitado. Melhorar a visibilidade do estado do sistema, oferecer feedback claro e garantir uma interface mais intuitiva e acessível devem ser prioridades para uma usabilidade mais satisfatória.

A Avaliação Heurística revelou problemas que não foram identificados no Teste de Usabilidade, como ícones pouco intuitivos, ausência de confirmação em ações críticas, inconsistências visuais e textuais, permissão de entradas inválidas, falta de funcionalidades para ações em massa, botões com baixa visibilidade e mensagens de erro pouco informativas ou inexistentes. Essa diferença ocorreu porque o Teste de Usabilidade focou em tarefas específicas e básicas, como conversação, exclusão, arquivamento e desarquivamento de chats, enquanto a Avaliação Heurística analisou o sistema de forma mais abrangente e detalhada.

No entanto, os principais problemas apontados no Teste de Usabilidade — como a falta de intuitividade, a ausência de um tutorial e a navegação confusa — também foram detectados na Avaliação Heurística. Isso confirma que essas falhas têm um impacto direto e significativo na experiência do usuário.

Além disso, a Avaliação Heurística destacou falhas específicas no design da interface, como botões com baixo contraste, falta de feedback visual, inconsistências linguísticas - com elementos misturando inglês e português - e problemas na iconografia. Esses fatores comprometem a usabilidade e tornam a interação menos intuitiva. Durante o Teste de Usabilidade, os participantes relataram dificuldades na navegação, mas não detalharam essas falhas de forma técnica, como foi feito na Avaliação Heurística.

Um problema comum identificado nos dois métodos foi a ausência de feedback do sistema. Por exemplo, na funcionalidade de desarquivar conversas, não havia nenhuma indicação visual ou mensagem confirmando que a ação havia sido concluída. No Teste de Usabilidade, os usuários tiveram dificuldade para encontrar essa opção e não perceberam que a falta de feedback era a principal causa da confusão.

6 CONCLUSÕES E TRABALHOS FUTUROS

Este estudo buscou compreender como os princípios de usabilidade podem influenciar a interação de novos usuários com o ChatGPT, destacando sua importância no desenvolvimento de soluções mais eficazes e satisfatórias. A aplicação de metodologias consolidadas, como a Avaliação Heurística e o Teste de Usabilidade, mostrou-se essencial para identificar falhas, propor melhorias e entender as percepções dos usuários.

Esses métodos clássicos, originados em períodos anteriores à popularização de tecnologias avançadas, demonstram que os princípios fundamentais de design e usabilidade transcendem as inovações tecnológicas. A continuidade de sua aplicação não apenas valida sua eficácia, mas também destaca a importância de aspectos como consistência, feedback informativo e prevenção de erros no desenvolvimento de interfaces.

Os resultados obtidos demonstraram que uma boa usabilidade é decisiva para a aceitação e o sucesso de chatbots. Usuários valorizam interfaces claras, intuitivas e que atendam às suas necessidades de forma eficiente. No entanto, pequenas inadequações podem comprometer a experiência, evidenciando a necessidade de um processo contínuo de validação e refinamento.

É importante ressaltar que, apesar deste estudo estar focado em analisar a interface do ChatGPT, alguns resultados envolveram análises da conversação por parte dos usuários. Embora esse não fosse o objetivo principal, essas análises podem enriquecer significativamente o estudo, abrindo espaço para novas perspectivas e possibilidades para futuros trabalhos relacionados ao tema.

Entre os principais desafios identificados, destaca-se a necessidade de maior inclusão e diversidade nos estudos de usabilidade. A presença de participantes com diferentes níveis de experiência, idades e contextos culturais pode trazer possibilidades mais abrangentes e contribuir para o desenvolvimento de sistemas mais democráticos.

Ademais, constatou-se que o feedback constante dos usuários desempenha um papel crucial na evolução das interfaces. Desenvolvedores e designers devem investir em ciclos iterativos de design, onde as sugestões dos usuários sejam incorporadas como parte essencial do processo de melhoria.

Em suma, este estudo não apenas reforça a importância da usabilidade no contexto digital, mas também aponta que ela deve ser vista como uma prática estratégica. Ao priorizar a usabilidade, é possível criar sistemas que não apenas atendam às expectativas dos usuários, mas que também promovam maior engajamento e eficiência. Espera-se que as contribuições

deste trabalho inspirem futuros estudos e práticas, ajudando a consolidar a usabilidade como um elemento-chave na construção de experiências digitais mais acessíveis e impactantes.

6.1 Trabalhos Futuros

Para aprimorar a análise de usabilidade do ChatGPT em futuros trabalhos, é fundamental adotar uma abordagem abrangente que considere tanto os aspectos de interface quanto os de conversação. Ao separar problemas de interface (como layout, navegação e responsividade) dos problemas de conversação (como coerência, relevância e naturalidade das respostas), é possível identificar com mais precisão as áreas que necessitam de melhorias específicas.

Explorar a usabilidade do ChatGPT em diversos contextos práticos, como suporte ao cliente, educação e saúde, permite compreender como as interações variam conforme o uso, conforme mencionado anteriormente na seção 3.2. Além disso, investigar a interação do ChatGPT com diferentes grupos de usuários, incluindo idosos, crianças e pessoas com deficiências, é essencial para garantir a inclusão e acessibilidade do sistema. Desenvolver novas métricas e métodos de avaliação específicos para a interação com IA conversacional pode capturar de maneira mais precisa a experiência do usuário, considerando tanto os aspectos de interface quanto os de conversação.

Analizar a usabilidade do ChatGPT em diferentes idiomas e culturas é crucial para identificar desafios e adaptações necessárias que melhorem a interação global. Examinar o impacto das atualizações contínuas no ChatGPT sobre a usabilidade ajuda a entender como as mudanças no sistema influenciam a experiência do usuário a longo prazo. Investigar o papel do feedback dos usuários no desenvolvimento do ChatGPT é vital para incorporar sugestões e resolver reclamações de forma eficaz no processo de design e melhoria do sistema. Por fim, estudar o impacto da usabilidade do ChatGPT na satisfação e engajamento dos usuários permite identificar quais características do sistema contribuem para uma experiência mais envolvente e satisfatória.

Ao abordar essas áreas e distinguir entre problemas de interface e de conversação, futuros trabalhos podem avançar na compreensão e aplicação dos princípios de usabilidade, contribuindo para o desenvolvimento de melhorias que sejam funcionais, acessíveis, inclusivas e satisfatórias para todos os usuários.

6.1.1 Dicas

Para aprimorar a avaliação heurística e os testes de usabilidade, é fundamental adotar uma abordagem abrangente que considere a complexidade do sistema e as diversas perspectivas dos avaliadores. Dividir sistemas complexos em módulos menores facilita a identificação de problemas específicos e evita a sobrecarga de informações durante a análise. Esse fracionamento permite que cada parte seja avaliada detalhadamente por uma equipe diversificada de pesquisadores, minimizando vieses individuais. Além disso, distribuir as sessões de avaliação ao longo de vários dias possibilita captar variações no comportamento dos usuários, contribuindo para insights mais profundos e precisos durante o teste de usabilidade.

Durante os testes, é crucial que cada participante utilize uma conta única, permitindo o rastreamento preciso das interações individuais e a correlação dos resultados às características específicas de cada usuário. A criação de ambientes de teste controlados e padronizados é igualmente importante, pois elimina interferências externas que possam distorcer os dados coletados. A seleção imparcial dos participantes deve refletir o perfil real dos usuários do sistema, e todo o processo deve ser minuciosamente documentado, registrando o que foi dito, feito e observado. A definição de métricas claras, como o tempo de realização de tarefas, a taxa de erros e o nível de satisfação, possibilita uma análise quantitativa robusta, que, aliada a métodos qualitativos – como sessões de feedback pós-teste –, propicia uma compreensão abrangente da experiência do usuário. A constante revisão dos critérios de avaliação e a capacitação contínua da equipe são práticas fundamentais para aprimorar os processos e obter resultados que possam, de forma efetiva, orientar melhorias futuras no sistema.

Ao integrar essas práticas e princípios, é possível conduzir avaliações heurísticas e testes de usabilidade mais eficazes, resultando em sistemas que oferecem uma experiência de usuário superior e atendem aos objetivos de negócio.

REFERÊNCIAS

- ABNT, A. B. de N. T. Requisitos ergonômicos para trabalho de escritórios com computadores. parte 11: orientações sobre usabilidade. **ABNT - Associação Brasileira de Normas Técnicas**, 2002.
- ADAMOPOULOU, E.; MOUSSIADES, L. An overview of chatbot technology. In: **Artificial Intelligence Applications and Innovations: 16th IFIP WG 12.5 International Conference, AIAI 2020, Neos Marmaras, Greece, June 5–7, 2020, Proceedings, Part II**. [S. l.]: Springer, 2020. p. 373–383.
- BARBOSA, S.; SILVA, B. **Interação humano-computador**. São Paulo: Elsevier Brasil, 2010.
- BARBOSA, S. D. J.; SILVA, B. S. d.; SILVEIRA, M. S.; GASPARINI, I.; DARIN, T.; BARBOSA, G. D. J. **Interação Humano-Computador e Experiência do Usuário**. [S. l.]: Autopublicação, 2021. ISBN 978-65-00-19677-1.
- BROWN, T.; MANN, B.; RYDER, N.; SUBBIAH, M.; KAPLAN, J. D.; DHARIWAL, P.; NEELAKANTAN, A.; SHYAM, P.; SASTRY, G.; ASKELL, A. *et al.* Language models are few-shot learners. **Advances in Neural Information Processing Systems**, Curran Associates, Red Hook, NY, v. 33, p. 1877–1901, 2020.
- CARROLL, J. M. Human–computer interaction. **Encyclopedia of Cognitive Science**, Wiley Online Library, 2006.
- COUTINHO, G. F. Avaliação de interfaces de usuário em chatbots de serviço público e o uncanny valley. **Sem nome de periódico**, 2021. Não publicado.
- CYBIS, W. d. A. Engenharia de usabilidade: uma abordagem ergonômica. **Labiutil**, Florianópolis, 2003.
- DEMERTZIS, K.; ILIADIS, L. Evolving computational intelligence system for malware detection. In: SPRINGER. **Advanced Information Systems Engineering Workshops: CAiSE 2014 International Workshops, Thessaloniki, Greece, June 16-20, 2014. Proceedings 26**. Berlin, 2014. p. 322–334.
- DUMAS, J. S.; REDISH, J. C. **A Practical Guide to Usability Testing**. Bristol: Intellect Books, 1999.
- FEITOSA, A. L. M.; BARRETO, I. C. d. H. C.; PERNENCAR, C. A. d. C.; OUTROS. Usabilidade de um chatbot para pessoas com hipertensão e diabetes. **Journal of Health Informatics**, v. 16, n. Especial, p. 1–14, 2024.
- FORUM, W. E. **The Future of Jobs Report 2020**. Genebra, 2020. Disponível em: <https://www.weforum.org/reports/the-future-of-jobs-report-2020> Acesso em: 08 dez. 2024.
- GARCIA, A. C. Ética e inteligência artificial. **Computação Brasil**, Computação Brasil, São Paulo, n. 43, p. 14–22, 2020.
- GUICHARD, J.; RUANE, E.; SMITH, R.; BEAN, D.; VENTRESQUE, A. Assessing the robustness of conversational agents using paraphrases. In: IEEE. **2019 IEEE International Conference On Artificial Intelligence Testing (AITest)**. New York, 2019. p. 55–62.

- JUNIOR, M. A. R. P. A aprendizagem da língua portuguesa como língua estrangeira através do uso da inteligência artificial: análise da usabilidade do chatgpt. **Revista Contemporânea**, v. 4, n. 6, p. e4446–e4446, 2024.
- MORAES, A. Ergodesign e interação humano-computador: Design e avaliação de interface. **User**, Rio de Janeiro, 2002.
- NIELSEN, J. **Usability Engineering**. San Francisco: Morgan Kaufmann, 1993.
- NIELSEN, J. **Usability Engineering**. [S. l.]: Morgan Kaufmann, 1994.
- NIELSEN, J. **Why You Only Need to Test with Five Users**. Fremont: Nielsen Norman Group, 2000.
- NIELSEN, J.; LANDAUER, T. K. A mathematical model of the finding of usability problems. In: **Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems**. New York: [S. n.], 1993. p. 206–213.
- NIMAVAT, K.; CHAMPANERIA, T. Chatbots: An overview of types, architecture, tools, and future possibilities. **International Journal of Scientific Research and Development (IJSRD)**, v. 5, n. 7, p. 1019–1024, 2017.
- OPENAI. **Introducing GPT-4: A new generation of language models**. 2023. Disponível em: <https://www.openai.com/research/gpt-4>. Acesso em: 02 set. 2024.
- OPENAI. **ChatGPT Overview**. 2025. Disponível em: <https://openai.com/chatgpt/overview/>. Acesso em: 14 feb. 2025.
- PRATES, R. O.; BARBOSA, S. D. J. Avaliação de interfaces de usuário: conceitos e métodos. In: **Jornada de Atualização em Informática do Congresso da Sociedade Brasileira de Computação, Capítulo**. [S. l.: s. n.], 2003. v. 6, p. 13–36.
- QUIÑONES, D.; RUSU, C. How to develop usability heuristics: A systematic literature review. **Computer Standards & Interfaces**, Elsevier, Amsterdam, v. 53, p. 89–122, 2017.
- REN, R.; ZAPATA, M.; CASTRO, J. W.; DIESTE, O.; ACUÑA, S. T. Experimentation for chatbot usability evaluation: A secondary study. **IEEE Access**, IEEE, New York, v. 10, p. 12430–12464, 2022.
- RUBIN, J.; CHISNELL, D. **Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests**. [S. l.]: John Wiley & Sons, 2008.
- SANTOS, W. P. Tecnologias da informação e comunicação (tics) e suas possibilidades de uso no ensino de língua portuguesa. **Revista Desempenho**, n. 28, 2018.
- SCHWAB, K. **A quarta revolução industrial**. 1. ed. São Paulo: Edipro, 2019.
- SHAWAR, B. A.; ATWELL, E. Chatbots: Are they really useful? **Journal for Language Technology and Computational Linguistics**, v. 22, n. 1, p. 29–49, 2007.
- SOUZA, F. J. d. Chartbot: uma interface conversacional para visualização de dados abertos do ibge. **Sem nome de periódico**, 2022. Não publicado.

TURLEY, N. Inside the launch – and future – of chatgpt. **The Verge**, 2024. Disponível em: <https://www.theverge.com/2024/12/12/24318650/chatgpt-openai-history-two-year-anniversary>. Acesso em: 06 mar. 2025.

VASWANI, A.; SHAZER, N.; PARMAR, N.; USZKOREIT, J.; JONES, L.; GOMEZ, A. N.; KAISER, ; POLOSUKHIN, I. Attention is all you need. **Advances in Neural Information Processing Systems**, Curran Associates, Red Hook, NY, v. 30, 2017.

APÊNDICE A – INFORMAÇÕES PARA AVALIAÇÃO HEURÍSTICA

O que é o ChatGPT: O ChatGPT é uma plataforma de inteligência artificial desenvolvida pela OpenAI, acessível tanto via dispositivos móveis quanto pela web. Focado em interação conversacional, o ChatGPT é projetado para responder perguntas, auxiliar na resolução de problemas, criar conteúdos e oferecer suporte em diversas áreas, como programação, escrita criativa, estudos e muito mais.

Público-alvo: Pessoas de fora da área da tecnologia que não tiveram nenhuma interação com o ChatGPT.

Concorrentes: Entre as principais plataformas semelhantes ao ChatGPT, destacam-se Bard (Google), Claude (Anthropic) e ferramentas específicas de IA em sistemas como Bing Chat e Jasper AI.

Objetivo: Identificar possíveis problemas de usabilidade que possam dificultar a interação entre usuários e o ChatGPT, com foco na experiência de aprendizado e assistência oferecida pela ferramenta.

Heurísticas: As diretrizes escolhidas para apoiar na avaliação, são as 10 Heurísticas de Nielsen que estão listadas abaixo:

1. **Visibilidade do estado do sistema:** O sistema deve sempre manter os usuários informados sobre o que está acontecendo, por meio de feedback apropriado em tempo hábil.
2. **Correspondência entre o sistema e o mundo real:** O sistema deve usar uma linguagem familiar ao usuário, com palavras, frases e conceitos que sejam conhecidos, e seguir convenções do mundo real.
3. **Controle e liberdade do usuário:** Usuários frequentemente escolhem ações por engano. Deve-se oferecer opções fáceis de desfazer ou refazer essas ações.
4. **Consistência e padrões:** Os usuários não devem precisar se perguntar se diferentes palavras, situações ou ações significam a mesma coisa. Siga as convenções da plataforma.
5. **Prevenção de erros:** Melhor do que boas mensagens de erro é um design cuidadoso que previna a ocorrência de problemas.
6. **Reconhecimento em vez de memorização:** Minimize a carga de memória do usuário, tornando objetos, ações e opções visíveis. O usuário não deve precisar lembrar informações de uma parte para outra.
7. **Flexibilidade e eficiência de uso:** O sistema deve atender tanto usuários iniciantes quanto experientes, permitindo atalhos para ações frequentes.

8. **Estética e design minimalista:** Diálogos e telas não devem conter informações irrelevantes ou raramente necessárias, pois isso pode competir com as informações relevantes.
9. **Ajuda os usuários a reconhecer, diagnosticar e corrigir erros:** As mensagens de erro devem ser expressas em linguagem clara (sem códigos), indicar o problema e sugerir uma solução.
10. **Ajuda e documentação:** Embora seja melhor que o sistema possa ser usado sem documentação, pode ser necessário fornecer ajuda e documentação acessíveis e focadas na tarefa do usuário.

APÊNDICE B – FORMULÁRIO DE RECRUTAMENTO

1. Idade

- Menos de 18 anos
- 18-24 anos
- 25-34 anos
- 35-44 anos
- 45-54 anos
- 55-64 anos
- Mais de 65 anos

2. Gênero

- Masculino
- Feminino
- Prefiro não responder

3. Qual é a sua profissão?

4. Você já usou algum tipo de assistente de inteligência artificial (IA) antes?

- Sim
- Não

5. Se sim, por favor, descreva brevemente sua experiência anterior com assistentes de IA:

6. Com que frequência você usa assistentes de IA em sua vida cotidiana?

- Diariamente
- Semanalmente
- Mensalmente
- Raramente
- Nunca usei

7. Experiência com o ChatGPT:

- Sim
- Não

APÊNDICE C – ENTREVISTA PÓS-USO

1. Você usa algum sistema ou aplicativo no seu trabalho? Como ele é?
2. Como você descreveria sua experiência geral ao utilizar o ChatGPT? Houve algo que se destacou para você?
3. O quanto fácil foi para você interagir com o ChatGPT? Quais dificuldades, se houve alguma, você enfrentou?
4. Você achou fácil encontrar as funcionalidades desejadas no ChatGPT?
5. As respostas do ChatGPT foram claras e compreensíveis? Houve algum momento em que você achou as respostas confusas ou vagas?
6. Como você avalia a relevância das respostas fornecidas pelo ChatGPT em relação às suas perguntas? Houve alguma resposta que não atendeu às suas expectativas?
7. Se você já utilizou outras ferramentas semelhantes ao ChatGPT, como você compara a experiência? O que o ChatGPT faz melhor ou pior?
8. Você consideraria usar o ChatGPT regularmente com base na experiência que teve durante este teste de usabilidade?
9. Você tem alguma sugestão ou recomendação para melhorar a usabilidade e eficácia do ChatGPT?
10. Há mais algum comentário que você gostaria de fazer sobre sua experiência com o ChatGPT?

APÊNDICE D – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO (TCLE)

Você está sendo convidado por Ana Beatriz Silva Linhares e Andréia Libório Sampaio como participante do questionário intitulado “Avaliação de Usabilidade do Chat GPT”. Você não deve participar contra a sua vontade. Leia atentamente as informações abaixo e faça qualquer pergunta que desejar, para que todos os procedimentos desta pesquisa sejam esclarecidos.

1. PARTICIPANTES DA PESQUISA: Pode participar da pesquisa qualquer estudante e profissional que não seja da área de tecnologia, com idade igual ou superior a 18 anos, que deseje colaborar e esteja de acordo com o presente termo.

2. ENVOLVIMENTO NA PESQUISA: Nesta pesquisa, as seguintes atividades serão realizadas: Iniciaremos com um questionário para coletar dados simples (nome, idade, profissão, experiências posteriores com inteligências artificiais), após a coleta de dados será solicitado a realização diversas tarefas de interação com o ChatGPT em um ambiente controlado.

1. Iniciar uma conversa sobre a situação atual do país;
2. Iniciar uma conversa sobre algo de seu interesse;
3. Criar uma nova conversa e perguntar algo relacionado a sua profissão;
4. Excluir uma conversa;
5. Arquivar uma conversa;
6. Desarquivar uma conversa.

As interações serão observadas por um moderador, que registrará suas ações e coletará seu feedback verbal. O teste pode incluir o uso de gravações de áudio e vídeo para fins de análise e documentação. Após o teste, faremos um breve questionário sobre sua experiência durante a pesquisa. Todas as informações coletadas durante o teste serão mantidas confidenciais e utilizadas apenas para fins de pesquisa. Mesmo que você queira participar agora, você pode voltar atrás ou parar de participar a qualquer momento. A sua participação é voluntária e o fato de não querer participar não vai trazer qualquer penalidade.

3. RISCOS E DESCONFORTOS: Por tratar-se de dois questionários e uma pesquisa de usabilidade existe a possibilidade de ocorrerem falhas técnicas, ou ainda, cansaço ou desconforto ao responder algumas perguntas. No entanto, de forma geral, considera-se que a pesquisa envolve riscos mínimos. Mesmo assim, no caso de cansaço ou desconforto, faça intervalos ou deixe a questão em branco quando for possível.

4. CONFIDENCIALIDADE DA PESQUISA: Você não será identificado (a) em nenhuma publicação que esta pesquisa possa resultar. Os pesquisadores tratarão a sua identidade

com padrões profissionais de sigilo, atendendo a legislação brasileira (Resolução Nº 466/12 do Conselho Nacional de Saúde), utilizando as informações fornecidas apenas para fins científicos.

5. BENEFÍCIOS E CUSTOS: Os benefícios incluem a oportunidade de contribuir para a pesquisa de usabilidade do ChatGPT, bem como a possibilidade de influenciar futuras melhorias na usabilidade e na experiência do usuário. Para participar deste estudo você não vai ter nenhum custo.

6. DECLARAÇÃO DE RESPONSABILIDADE: Os pesquisadores responsáveis, Ana Beatriz Silva Linhares (graduanda em Sistemas de Informação pela Universidade Federal do Ceará, campus de Quixadá) e Andréia Libório Sampaio, (diretora da Universidade Federal do Ceará, campus de Quixadá) declaram cumprir as exigências da RESOLUÇÃO Nº 466, DE 12 DE DEZEMBRO DE 2012, e suas complementares, do Conselho Nacional de Saúde, que trata dos cuidados éticos de pesquisas científicas envolvendo pessoas.

7. CONTATO: Em caso de dúvidas, solicitações ou notificação de acontecimentos não previstos, você poderá contatar os pesquisadores responsáveis, Ana Beatriz Silva Linhares, pelo e-mail: analisnq@alu.ufc.br ou Andréia Libório Sampaio, pelo e-mail: andreia.ufc@gmail.com. CONSENTIMENTO DA PARTICIPAÇÃO Tendo compreendido perfeitamente tudo o que me foi informado sobre a minha participação no mencionado estudo, ao assinar abaixo, confirmo que li e entendi as informações fornecidas neste termo de consentimento e concordo em participar voluntariamente do teste de usabilidade do ChatGPT.

Assinatura do pesquisador(a):

Assinatura do Participante:

APÊNDICE E – PROBLEMAS HEURÍSTICOS CONSOLIDADOS

Tabela 11 – Problemas consolidados

| Heurísticas | Local e Descrição | Severidade | Possível sugestão |
|------------------|---|-----------------------|--|
| <i>Porção 01</i> | | | |
| H02, H04, H09 | Botão de “enviar”. Ao colocar o cursor em cima do ícone, aparece o termo “message is empty”, em inglês. | Problema pequeno | Traduzir a mensagem. |
| H08 | No modo claro, os botões que auxiliam com “prompts pré-escritos” têm baixo contraste no efeito hover. | Problema pequeno | Ajustar a cor do overlay. |
| <i>Porção 02</i> | | | |
| H01, H08 | Existem, nessa posição, alguns botões que apresentam problemas de interação, como o botão de lápis no menu “ChatGPT”. Quando clicados, não apresentam feedback. | Problema pequeno | Postar alguma mensagem de feedback ou ajustar o botão. |
| H01, H03, H06 | O botão de “arquivar” exibe um modal apenas uma vez. Se o usuário clicar novamente, o modal não se repete. | Problema grande | Mostrar a mensagem sempre que o chat for arquivado. |
| H01, H04 | Ao editar o nome da conversa, o usuário não recebe feedback visual como em outras ações, como arquivar ou excluir conversas. | Problema cosmético | Mostrar um modal informando a conclusão da mudança. |
| H03, H05 | Ao editar o nome de uma conversa antiga, o histórico da conversa, incluindo a data de início, é perdido. | Problema catastrófico | Preservar o histórico da conversa ao renomear. |
| <i>Porção 03</i> | | | |

| Heurísticas | Local e Descrição | Severidade | Possível sugestão |
|--------------------|--|--------------------|---|
| H08, H10 | O ícone de Interrogação no canto inferior direito é muito pequeno, claro e discreto. | Problema cosmético | Aumentar o tamanho do ícone. |
| <i>Porção 04</i> | | | |
| H05 | Configurações → Geral → Excluir todos os chats → Excluir tudo. Não é informado que os chats arquivados também serão excluídos. | Problema grande | Informar que os chats arquivados também serão excluídos. |
| H04 | Configurações → Perfil de Construtor → Links. Ao lado do ícone de planeta, falta uma descrição, ao contrário dos outros ícones. | Problema cosmético | Adicionar uma descrição ao lado do ícone, como os outros ícones. |
| H05 | Configurações → Perfil de Construtor → Adicionar novo domínio. É possível adicionar qualquer termo e só depois aparece uma mensagem de erro. | Problema pequeno | Restringir o input para aceitar apenas domínios válidos. |
| H07 | Configurações → Chats Arquivados. A exclusão ou desarquivamento dos chats arquivados precisa ser feita um por vez. | Problema pequeno | Adicionar uma opção para selecionar e desarquivar todos os chats. |
| H05 | Botão “sair”. Ao clicar, não é apresentado um diálogo de confirmação, e o logout ocorre imediatamente. | Problema pequeno | Adicionar um modal de confirmação. |
| H05 | Configurações → Segurança → Sair de todos os aparelhos. Não há diálogo de confirmação antes de executar a ação. | Problema grande | Adicionar um modal de confirmação. |

| Heurísticas | Local e Descrição | Severidade | Possível sugestão |
|-------------|--|------------------|---|
| H04, H09 | Editar ChatGPT personalizado. O botão de salvar permanece desabilitado sem explicar o motivo, mesmo após atualizar as informações. | Problema pequeno | Adicionar um tooltip explicando o motivo do botão desabilitado. |
| H03 | Menu de perfil. Falta a opção de tela inicial no menu. | Problema pequeno | Adicionar a opção "Home" no menu. |
| H01 | Criar ChatGPT personalizado. Não há uma descrição sobre as mudanças que ocorreram após a criação. | Problema grande | Apresentar um modal explicando as mudanças após a criação. |

Porção 05

| | | | |
|-----|---|--------------------|---|
| H01 | Chat temporário → Botão de apagar Chat. Ao clicar, quando não tem nenhuma conversa, nada é retornado ao usuário. | Problema pequeno. | Utilizar modais de informação do estado do sistema. |
| H08 | Chat temporário → Botão de apagar Chat. A localização do botão (considerando sua função) fica distante da área em que se relaciona. | Problema pequeno | Reposicionar o botão. |
| H02 | Chat temporário → Botão de apagar Chat. O ícone utilizado não deixa claro que a ação refere-se a excluir chat temporário. | Problema cosmético | Utilizar outro ícone. |
| H04 | Menu ChatGPT → ChatGPT. O ícone utilizado para retratar o chat gpt é diferente do utilizado em outros lugares. | Problema cosmético | Padronizar os ícones. |

| Heurísticas | Local e Descrição | Severidade | Possível sugestão |
|------------------|---|---------------------|---|
| H01, H08 | Menu ChatGPT → ChatGPT. Na versão gratuita do ChatGPT, essa opção é ativada automaticamente. Caso o usuário tenha um plano pago, ele pode selecionar entre ChatGPT gratuito ou Plus. Ao clicar no botão “ChatGPT” nada acontece, pois o usuário não tem acesso ao plano pago, e nada é dito ao usuário. | Problema pequeno | Retornar o feedback e modificar o layout para que essa interação seja mais clara. |
| <i>Porção 06</i> | | | |
| H03 | Ao final de cada resposta, são apresentados botões de “curtir” ou “descurtir”, porém ao clicar, a ação não pode mais ser desfeita. | Problema grande. | Modal da confirmação de ação ou fazer com que a ação possa ser desfeita também por um clique. |
| H03, H08 | Interface conversacional não possui opção de excluir pergunta, apenas de editar. | Problema pequeno | Adicionar um botão de excluir pergunta. |
| H07 | Interface conversacional. Quando o ChatGPT oferece uma resposta longa, a rolagem não acompanha a resposta. | Problema cosmético. | Seria importante a rolagem acompanhar a resposta. |
| H05, H07, H09 | Interface conversacional não sabe lidar bem com perguntas ambíguas. | Problema grande | O chatbot deveria perguntar qual o contexto antes de enviar a resposta para o usuário ou após a pergunta ambígua, pedir mais informações. |

| Heurísticas | Local e Descrição | Severidade | Possível sugestão |
|-------------|---|-----------------------|---|
| H01, H09 | Interface conversacional as vezes não gera nenhuma resposta e não retorna nenhuma informação para o usuário. | Problema catastrófico | O chatbot deveria exibir alguma mensasgem de erro. |
| H10 | Interface conversacional. O ChatGPT não tem consciência da sua própria interface. Assim, quando o usuário realiza questões sobre seu funcionamento, não é alcançado nenhum auxílio. | Problema grande | O chatbot deveria explicar para o usuário sobre suas funções e localização na interface. |
| H08 | Curtir ou “descurtir” mensagens do chatGPT. Os botões estão bem próximos e semelhantes. | Problema pequeno | Alterar o layout dos ícones. |
| H01, H02 | Quando o ChatGPT inicia sua resposta, a única informação que é recebida quanto ao carregamento da mensagem, é um signo de círculo, que às vezes trava. | Problema grande | Adicionar outros signos de carregamento da informação. |
| H01, H10 | É sabido que não é possível enviar textos enquanto o ChatGPT está gerando outro texto. No sistema, a única indicação disso é de um botão que só aparece no momento de geração do texto. | Problema pequeno | Adicionar algum modal de aviso, com: “Proibido enviar prompts enquanto uma resposta está sendo gerada. Caso não goste da resposta gerada até o momento, experimente cancelar o processo no botão da barra de pesquisa”. |

| Heurísticas | Local e Descrição | Severidade | Possível sugestão |
|--------------------|---|---------------------|--|
| H01 | Quando o usuário realiza a edição de uma mensagem enviada em uma conversa, é criado uma outra subconversa, com o usuário. Essa alteração não é indicada claramente ao usuário. | Problema cosmético. | Adicionar algum modal de aviso sobre a alteração. |
| H04 | Quando o usuário responde um trecho específico da mensagem resposta do ChatGPT e edita essa mensagem, o novo conteúdo perde a referência da resposta. | Problema grande | Consertar a referência de resposta. |
| H08 | Quando o usuário realiza a edição de uma mensagem enviada em uma conversa, é criado uma outra subconversa, com o usuário. Em larga escala, a interface fica com muitas ramificações. | Problema grande | Remodelar conversas e edições, por outra forma de interação. |
| H02, H04, H10 | Interface conversacional. Ao colocar o cursor em cima do ícone para “ler o conteúdo” que está no chat, aparece a legenda em inglês, diferentemente dos outros ícones, como o “copiar” e o “curtir”, que estão em português. | Problema pequeno. | Traduzir a mensagem do tooltip. |

Porção 07

| | | | |
|-----|--|-----------------|--|
| H04 | Quando o usuário seleciona um GPT disponível e busca ver mais informações em outras redes sociais, como sites, páginas no LinkedIn e X (Twitter), aparece um modal com acesso dificultoso. | Problema grande | Consertar o modelo de informações extras do GPT. |
|-----|--|-----------------|--|

| Heurísticas | Local e Descrição | Severidade | Possível sugestão |
|-------------|---|--------------------|--|
| H07 | Quando o usuário seleciona um GPT, ele só consegue interagir, no caso do plano gratuito, caso ele não tenha excedido seus usos gratuitos diários de ChatGPT. Essa informação é dita no último passo de escolha e interação. | Problema grande | Indicar, logo no início do fluxo, que os GPTs só estão disponíveis para usuários do plano "premium" da plataforma. |
| H07 | A rolagem de outros gpts criados pela empresa / organização dos GPTs destacado no card, é impossibilitada em usuários de mouse. | Problema grande | Corrigir a rolagem de informações. |
| H10 | Existe um botão, no canto superior esquerdo, que não tem explicação sobre sua função. | Problema grande | Usar outros termos ou adicionar um tooltip. |
| H04, H10 | Existe um botão, no canto superior esquerdo, que utiliza um ícone que já é utilizado para outros propósitos. (Planeta Terra é utilizado tanto para essa visão global quanto para sites). | Problema cosmético | Trocar o ícone. |
| H02, H04 | Algumas categorias e itens estão em inglês enquanto outras estão em português. | Problema grande | Traduzir. |