



UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS QUIXADÁ
CURSO DE GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

MARIA SOFIA JERONIMO DE LIMA

APLICANDO MÉTODOS DE AVALIAÇÃO DE IHC NO CONTEXTO DE UM
***CHATBOT* DE ENSINO E APRENDIZAGEM DE MATEMÁTICA**

QUIXADÁ

2026

MARIA SOFIA JERONIMO DE LIMA

APLICANDO MÉTODOS DE AVALIAÇÃO DE IHC NO CONTEXTO DE UM *CHATBOT*
DE ENSINO E APRENDIZAGEM DE MATEMÁTICA

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

Orientador: Prof. Me. Marcelo Martins da Silva.

QUIXADÁ

2026

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)

L699a Lima, Maria Sofia Jeronimo de.
Aplicando métodos de avaliação de IHC no contexto de um Chatbot de ensino e aprendizagem de Matemática / Maria Sofia Jeronimo de Lima. – 2026.
105 f. : il. color.

Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Campus de Quixadá, Curso de Ciência da Computação, Quixadá, 2026.
Orientação: Prof. Me. Marcelo Martins da Silva.

1. Chatbots. 2. Comunicabilidade. 3. Usabilidade. 4. Agentes Conversacionais. 5. Mathos AI. I. Título.
CDD 004

MARIA SOFIA JERONIMO DE LIMA

APLICANDO MÉTODOS DE AVALIAÇÃO DE IHC NO CONTEXTO DE UM *CHATBOT*
DE ENSINO E APRENDIZAGEM DE MATEMÁTICA

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

Aprovada em: 21/01/2026.

BANCA EXAMINADORA

Prof. Me. Marcelo Martins da Silva (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Me. Adson Roberto Pontes Damasceno
Universidade Estadual do Ceará (UECE)

Prof. Dr. Marcos Devaner do Nascimento
Universidade Federal do Ceará (UFC)

À todos que acreditaram em mim quando eu mais duvidava, a sua fé e força me tornaram quem sou hoje.

AGRADECIMENTOS

À minha família, por mostrarem desde sempre o valor do estudo, por sempre lutarem para que eu consiga atingir meus objetivos e por me ensinarem o que nenhuma instituição poderia ensinar.

Aos meus pais, por serem minha base nessa jornada. À minha mãe, pelos aconselhamentos, pela ajuda nos momentos difíceis e pelo apoio, e ao meu pai, por nunca ter deixado nada faltar, por todas as vezes que acordou cedo para que eu pudesse ir tranquila para as aulas e por toda a força que me deu.

Aos meus irmãos, Clara e Davi, por me darem um motivo para ser exemplo, por me acompanharem e por estarem presentes sempre que eu precisei. Dedico o meu sucesso a vocês e a tudo o que vocês fizeram por mim.

Aos meus avós, que, apesar de não entenderem exatamente o que eu estudo e apenas disserem que “é difícil”, sempre estiveram presentes, sempre perguntaram como eu estava e sempre me apoiaram nessa escolha. Sem vocês, eu não teria toda a resiliência necessária para chegar aqui.

Ao meu companheiro, Pedro, por ter me acompanhado durante praticamente toda a graduação, por toda ajuda que você me deu quando eu precisava, e por continuar acreditando em mim, mesmo que eu mesma duvide. Muito obrigado por sempre estar presente, por me acolher e me permitir compartilhar a vida com você.

Ao Prof. Me. Marcelo Martins, pela excelente orientação e por me desafiar durante o processo de construção desse trabalho, me provando que eu sou capaz de fazer coisas incríveis.

Aos professores do campus da UFC Quixadá, por contribuírem diretamente para o meu aprendizado e ao meu crescimento enquanto profissional. Obrigado por todos os seus ensinamentos, por me mostrarem todos os diferentes caminhos que eu posso escolher trilhar durante a minha jornada.

Aos meus amigos e colegas, por me mostrarem que este percurso é tudo menos solitário. Agradeço toda a ajuda durante as disciplinas difíceis, por me acolherem quando precisei, e por me mostrarem que a experiência acadêmica não é apenas estudo.

E aos que participaram deste estudo, muito obrigado! O seu feedback e a sua contribuição foram essenciais para esta pesquisa.

"O mundo é tão grande quanto você quiser que seja. Para onde você vai agora?" (TOBY FOX, 2025, UNDERTALE 10th Anniversary Stream. Disponível em <https://youtu.be/FS903HbnfWg>)

RESUMO

O aprendizado online tornou-se uma realidade cada vez mais presente para muitos estudantes. Nesse contexto, diversas soluções inovadoras foram desenvolvidas para o ensino em ambientes virtuais, entre elas, o uso de *Chatbots* tutores. *Chatbots* são softwares criados para conduzir uma conversa com o usuário, apresentando aplicações variadas no cotidiano. Por se tratarem de interfaces que interagem diretamente com o usuário, é necessário garantir a qualidade e eficiência desses sistemas por meio de avaliações de Usabilidade e Comunicabilidade. O objetivo deste trabalho é realizar diferentes avaliações dentro de um *Chatbot* chamado *Mathos AI*, com a finalidade de obter tanto uma visão geral sobre o estado atual do software quanto compreender as estratégias de avaliação em *Chatbots* por meio de um exemplo prático. Para realizar a avaliação, serão utilizados os métodos Modelo para Avaliação da interação em Sistemas Sociais a partir da Linguagem Textual do Usuário (MALTU), Método de Inspeção Semiótica (MIS), Método de Avaliação de Comunicabilidade (MAC) e Teste de Usabilidade. Espera-se, com isso, reunir perspectivas complementares que permitam identificar potenciais melhorias na interface e na experiência de uso do sistema.

Palavras-chave: chatbots; comunicabilidade; usabilidade; agentes conversacionais; Mathos AI; ensino virtual; MALTU; MIS; MAC.

ABSTRACT

Online learning became a reality even more present for a lot of students. In this context, various innovative solutions were developed for learning in virtual environments, including the use of tutoring *Chatbots*. *Chatbots* are software developed to conduct a conversation with the user, having various applications in our daily life. Since they are interfaces that interact directly with the user, it's necessary to ensure the quality and efficiency of the system with the use of communicability and usability evaluations. The objective of this research is to make various evaluations inside a *Chatbot* called *Mathos AI*, to obtain an overview about the current state of the software, as well as to understand the evaluation strategies inside *Chatbots* by a practical example. To execute the evaluation, there will be used the methods Model for Evaluation of the Interaction in Social Systems by the Textual Language of the User (MALTU), Semiotic Inspection Method (MIS), Communicability Evaluation Method (MAC). It's expected that, with the evaluations, we get complementary perspectives that allow us to identify potential increments on the interface and in the system's use experience.

Keywords: chatbots; communicability; usability; conversational agents; Mathos AI; online learning; MALTU; MIS; MAC.

LISTA DE FIGURAS

Figura 1 – Versão web do <i>Mathos AI</i>	24
Figura 2 – Capturas de tela da versão mobile	24
Figura 3 – Representação de um caso de uso básico do sistema	25
Figura 4 – Método de avaliação MIS	30
Figura 5 – Método MALTU	32
Figura 6 – Heurísticas de Langevin.	36
Figura 7 – Funcionamento do FeedBot.	38
Figura 8 – Funcionamento do LitBot.	39
Figura 9 – Execução do MALTU	42
Figura 10 – Execução do MIS	44
Figura 11 – Execução do MAC	46
Figura 12 – Triangulação dos resultados	48
Figura 13 – Visão geral dos resultados do MALTU.	50
Figura 14 – Signos Metalinguísticos encontrados na interface correspondente ao cenário 1.	53
Figura 15 – Signos Metalinguísticos encontrados na interface correspondente ao cenário 2.	54
Figura 16 – Signos Metalinguísticos encontrados na interface correspondente ao cenário 3.	54
Figura 17 – Signos Estáticos encontrados na interface correspondente ao cenário 1.	56
Figura 18 – Signos Estáticos encontrados na interface correspondente ao cenário 2.	56
Figura 19 – Signos Estáticos encontrados na interface correspondente ao cenário 3.	57
Figura 20 – Signos Dinâmicos encontrados na interface correspondente ao cenário 1.	58
Figura 21 – Signos Dinâmicos encontrados na interface correspondente ao cenário 2.	59
Figura 22 – Signos Dinâmicos encontrados na interface correspondente ao cenário 3.	60
Figura 23 – Frequência total por etiqueta	63
Figura 24 – <i>Paywall</i> exibido durante as interações	64
Figura 25 – Menu com símbolos e fórmulas \LaTeX	64
Figura 26 – Tela de flashcards	65

LISTA DE TABELAS

Tabela 1 – Participantes do experimento MAC	62
Tabela 2 – Frequência de etiquetas por tarefa	62
Tabela 3 – Frequência de rupturas por tarefa e usuário	62

LISTA DE QUADROS

Quadro 1 – Comparação entre os trabalhos relacionados	40
Quadro 2 – Quadro de cenários para a avaliação	45
Quadro 3 – Quadro de tarefas para a avaliação	47

LISTA DE ABREVIATURAS E SIGLAS

AVA	Ambientes Virtuais de Aprendizagem
IA	Inteligência Artificial
IHC	Interação Humano-Computador
MAC	Método de Avaliação de Comunicabilidade
MALTU	Modelo para Avaliação da interação em Sistemas Sociais a partir da Linguagem Textual do Usuário
MIS	Método de Inspeção Semiótica
MRC	<i>Machine Reading Comprehension</i>
PLN	Processamento de Linguagem Natural
PRU	Postagem Relacionada ao uso do Sistema
SBF	<i>Social Bot Framework</i>
SUS	<i>System Usability Scale</i>
TCC	Trabalho de Conclusão de Curso
UX	Experiência do Usuário

SUMÁRIO

1	INTRODUÇÃO	15
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	<i>Chatbots</i>	18
2.1.1	<i>Classificação de Chatbots</i>	19
2.1.2	<i>Chatbots na educação</i>	20
2.2	Interfaces Conversacionais	20
2.2.1	<i>Design de interfaces conversacionais</i>	21
2.2.1.1	<i>Práticas propostas para o design conversacional</i>	21
2.2.1.2	<i>Práticas a serem evitadas</i>	23
2.3	<i>Mathos AI</i>	23
2.4	Engenharia Semiótica	25
2.4.1	<i>Signo, Significação e Semiose</i>	26
2.4.2	<i>O Sistema como Artefato de Comunicação</i>	26
2.5	Avaliação de Interação	27
2.5.1	<i>Métodos de Avaliação</i>	28
2.5.1.1	<i>Métodos de Inspeção</i>	28
2.5.1.2	<i>Métodos de Observação</i>	29
2.5.1.3	<i>Métodos de Investigação</i>	31
3	TRABALHOS RELACIONADOS	33
3.1	STUART: An Intelligent Tutoring System for Increasing Scalability of Distance Education Courses	33
3.2	Avaliação de Usabilidade do CoderBot como Recurso Pedagógico no Ensino de Programação	34
3.3	Investigando chatbots governamentais: um panorama sobre a usabilidade dentro e fora do Brasil	35
3.4	Chatbots Como Uma Ferramenta Para Escalar Processos de Mentoria: Individualmente Apoiando o Auto-Estudo na Educação Superior	37

3.5	Quadro comparativo	39
4	METODOLOGIA	41
4.1	Definição dos objetivos	41
4.2	Aplicação do MALTU	42
4.3	Aplicação do MIS	43
4.4	Aplicação do MAC	46
4.5	Consolidação dos dados	47
5	RESULTADOS	49
5.1	Resultados do MALTU	49
5.2	Resultados do MIS	52
5.2.1	<i>Signos Metalinguísticos</i>	52
5.2.1.1	<i>Reconstrução da Metamensagem</i>	55
5.2.2	<i>Signos Estáticos</i>	55
5.2.2.1	<i>Reconstrução da Metamensagem</i>	57
5.2.3	<i>Signos Dinâmicos</i>	57
5.2.3.1	<i>Reconstrução da Metamensagem</i>	59
5.2.4	<i>Metamensagem final do designer</i>	60
5.3	Resultados do MAC	61
5.3.1	<i>Execução</i>	61
5.3.1.1	<i>Enviar uma pergunta ao Chatbot</i>	63
5.3.1.2	<i>Enviar uma questão para o Chatbot resolver</i>	64
5.3.1.3	<i>Enviar um erro conceitual / erro de cálculo</i>	65
5.3.1.4	<i>Utilizar as outras funcionalidades dentro do Chat</i>	65
5.3.1.5	<i>Utilizar a funcionalidade Teste-me</i>	66
5.3.1.6	<i>Resolver questões utilizando a funcionalidade de leitura de PDF</i>	66
5.3.2	<i>Perfil Semiótico</i>	66
5.3.2.1	<i>Quem o designer pensa ser o usuário do produto por ele projetado?</i>	67
5.3.2.2	<i>Quais os desejos e necessidades dos usuários, na visão do designer?</i>	67
5.3.2.3	<i>Na visão do designer, de que maneiras os usuários preferem fazer o que desejam e precisam, onde, quando e por quê?</i>	67
5.3.2.4	<i>Qual foi o sistema que o designer projetou para os usuários e como eles devem utilizá-lo?</i>	67

5.3.2.5	<i>Qual é a visão de design e quão bem a lógica é compreendida pelos usuários?</i>	68
5.4	Triangulação dos Resultados	68
5.4.1	Análise e Interpretação dos Resultados	68
5.4.1.1	<i>Principais Convergências</i>	68
5.4.1.2	<i>Principais Complementações</i>	69
5.4.1.3	<i>Principais Divergências</i>	69
5.4.2	O Processo de Construção-Síntese	70
5.5	Respondendo às questões de pesquisa	70
5.5.1	Como os métodos de avaliação de Interação Humano-Computador (IHC) podem ser utilizados dentro do contexto de Chatbots tutores/educacionais?	70
5.5.2	Quais métodos de avaliação seriam mais adequados no contexto de um voltado para a educação?	71
5.5.3	Quais são possíveis requisitos para um Chatbot voltado para a educação?	71
5.5.4	Como os Chatbots podem auxiliar dentro do processo de aprendizagem de matemática?	71
6	CONCLUSÕES E TRABALHOS FUTUROS	72
6.1	Limitações e Ameaças à Validade	73
6.2	Trabalhos futuros	74
6.3	Considerações finais	74
	REFERÊNCIAS	76
	APÊNDICE A –CÓDIGO-FONTE DO SCRIPT UTILIZADO PARA A EXTRAÇÃO DE PRUS	80
	APÊNDICE B –TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO	81
	APÊNDICE C –LISTA DE QUESTÕES DE APOIO PARA O TESTE	85
	APÊNDICE D –ROTEIRO ENTREVISTA SEMIESTRUTURADA	88
	APÊNDICE E –ROTEIRO PARA O MÉTODO DE AVALIAÇÃO DE COMUNICABILIDADE	91
	APÊNDICE F –QUESTIONÁRIO SOBRE A COMUNICABILIDADE DEO APLICATIVO MATHOS AI	94

1 INTRODUÇÃO

Com a pandemia de COVID-19 durante o ano de 2020, diversos setores como empresas e instituições passaram por mudanças drásticas. Dentro do campo da educação, essas alterações puderam ser sentidas com uma rápida transição do ensino presencial para o modelo remoto, o que levou à necessidade de adaptações na forma de ensinar, aprender e avaliar os estudantes (Flores *et al.*, 2021; Behar, 2020).

Diante desse cenário, o uso do aprendizado online se tornou uma realidade frequente para muitos estudantes. Com a necessidade de adaptação ao novo contexto, houveram diversas soluções de sucesso nas diferentes áreas (Marinoni *et al.*, 2020). Dentre as soluções desenvolvidas, tem-se o uso de Ambientes Virtuais de Aprendizagem (AVA), ambientes que se caracterizam como espaço onde educandos e educadores podem construir as relações de aprendizagem (Sousa *et al.*, 2023).

Dessa forma, surge a problemática do ensino da matemática no contexto de aprendizado remoto. Com o avanço tecnológico, passaram a adotar o uso de recursos de informática educativa no ensino e aprendizagem de matemática, conforme Corrêa e Brandemberg (2021), visto que os métodos tradicionais têm se mostrado insuficientes diante das demandas educacionais contemporâneas (Castro *et al.*, 2025).

Atualmente, uma ferramenta que possui grande potencial de aplicação dentro do ensino à distância e da aprendizagem é a Inteligência Artificial (IA). Segundo Boulay (2023), “a IA pode proporcionar um ensino personalizado e adaptado às necessidades individuais dos alunos, permitindo um aprendizado eficiente e engajante”. Essa área pode ser definida como o estudo dos métodos para fazer computadores se comportarem de forma inteligente (Russell; Norwig, 2022). Segundo Silva *et al.* (2023), a presença da IA na educação almeja criar uma abordagem dinâmica, adaptativa e inclusiva para o ensino, afetando alunos, educadores e administradores.

No momento, uma das principais aplicações da IA é o Processamento de Linguagem Natural (PLN). O PLN é “um campo de pesquisa que tem como objetivo investigar e propor métodos e sistemas de processamento computacional da linguagem humana.” (Caseli; Nunes, 2024). Dentro dessa área, são estudadas soluções para problemas computacionais que utilizam a comunicação humana, como sons, palavras, sentenças e discursos (Gonzalez; Lima, 2003).

Existem diferentes aplicações de PLN no cotidiano, entre elas, o uso dentro de *Chatbots*. Os *Chatbots* são softwares criados para conduzir uma conversa com o usuário, podendo

ser baseados em voz (Siri¹, Alexa²), baseados em texto (ChatGPT³, DeepSeek⁴, Gemini⁵), ou multimodais. Essa área de estudo está mostrando um grande potencial de desenvolvimento, pois é aplicada em diversos contextos, como entretenimento, saúde, educação, atendimento ao cliente, negócios (Casas *et al.*, 2020).

Ademais, o uso de *Chatbots* na educação tem se expandido como uma estratégia para potencializar o processo de ensino-aprendizagem, como pode ser exemplificado pelos trabalhos de Hobert (2019) e Carreira *et al.* (2022). O uso de *Chatbots* na educação possui diferentes objetivos, como desenvolvimento de habilidades, o aumento da eficiência do processo de ensino-aprendizagem, a elevação da motivação dos estudantes e a ampliação da disponibilidade do acesso à educação (Wollny *et al.*, 2021). No contexto da matemática, *Chatbots* podem ajudar os alunos a compreender conceitos matemáticos com exercícios e explicações detalhadas, adaptando-se às necessidades individuais dos alunos, tornando-a mais interessante para eles (Silva, 2024).

Com o crescimento na popularidade e na demanda dos *Chatbots*, é necessário um nível de qualidade alto para que estes sejam realmente úteis a seus usuários (Guerino; Valentim, 2020). Diante disso, a área de IHC se torna essencial, pois fornece princípios e métodos para atingir este nível de qualidade. A área de IHC está interessada no projeto, implementação e avaliação de sistemas computacionais interativos para o uso humano, juntamente com os fenômenos relacionados a esse uso (Hewett *et al.*, 1992). Através da IHC, é possível analisar aspectos como usabilidade, acessibilidade, eficiência e satisfação nas interações, garantindo uma boa experiência para os usuários.

Nesse cenário, torna-se importante o uso de avaliações de IHC. De acordo com Barbosa e Silva (2010), “A avaliação de IHC é uma atividade fundamental em qualquer processo de desenvolvimento que busque produzir um sistema interativo com alta qualidade de uso”, pois, mesmo que seja feito o uso de critérios e modelos de projetos, ainda é possível obter um produto com uma qualidade inadequada. Essa realidade também se aplica aos *Chatbots*, que, ao atuarem como agentes conversacionais que simulam interações humanas, aspectos como a experiência do usuário e a qualidade da comunicação tornam-se elementos centrais para a eficácia da interação.

Portanto, neste contexto, este trabalho visa responder às seguintes questões de pesquisa:

¹ <https://www.apple.com/siri/>

² https://www.amazon.com/dp/B0DCCNHV5?ref=aucc_web_red_xaa_evgn_tx_0001

³ <https://chatgpt.com/>

⁴ <https://www.deepseek.com/>

⁵ <https://gemini.google.com/>

- Como os métodos de avaliação de IHC podem ser utilizados dentro do contexto de *Chatbot* tutores/educacionais?
- Quais métodos de avaliação seriam mais adequados no contexto de um *Chatbot* voltado para a educação?
- Quais são possíveis requisitos para um *Chatbot* projetado nesse contexto?
- Como os *Chatbots* podem auxiliar dentro do processo de aprendizagem de matemática?

Entretanto, há uma escassez de métodos que avaliem a experiência do usuário dentro do contexto de *Chatbots* (Guerino; Valentim, 2020). Atualmente, existem estudos que focam na usabilidade e na Experiência do Usuário (UX) por meio de métodos de avaliação adaptados, como Damasceno *et al.* (2020) e Miranda *et al.* (2024). De forma similar, o objetivo deste trabalho é analisar a qualidade da interação entre usuários e o aplicativo *Mathos AI*, um *Chatbot* que visa auxiliar estudantes com o aprendizado de matemática. Para isso, foram empregados diferentes métodos de avaliação, que permitiram identificar aspectos sobre a usabilidade e a comunicabilidade do sistema. A partir disso, propôs-se melhorias que contribuam para a eficácia pedagógica da ferramenta, além de explorar seu potencial de uso em contextos educacionais.

1.1 Objetivos

1.1.1 Objetivo geral

Realizar a avaliação da interação de um *Chatbot* de ensino e aprendizagem voltado para a matemática, identificando possíveis lacunas e barreiras no uso, propondo melhorias e sugestões.

1.1.2 Objetivos específicos

- Realizar uma avaliação utilizando métodos de investigação, inspeção e observação;
- Comparar os resultados obtidos a partir de cada avaliação, categorizando os problemas encontrados e suas possíveis causas;
- Consolidar um diagnóstico abrangente do estado atual do *Chatbot*, incluindo recomendações para aprimorar sua eficácia pedagógica e a satisfação dos usuários.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo, serão apresentados conceitos necessários para a compreensão deste trabalho. Na Seção 2.1, é apresentada a definição de *Chatbots*. Em seguida, a Seção 2.5 se aprofunda em sub-tópicos da área de IHC focados em métodos, como os métodos de avaliação, métodos de inspeção, métodos de observação e métodos de investigação.

2.1 *Chatbots*

A IA destaca-se como um dos campos mais proeminentes da computação, sendo uma das suas aplicações mais recorrentes os *Chatbots*. *Chatbots* estão presentes no imaginário há muitos anos, como por exemplo, o personagem *AM*, do conto americano “Eu Não Tenho Boca e Preciso Gritar”¹, publicado em 1967. Um dos primeiros registros do conceito de *Chatbot* na literatura técnica remonta ao ELIZA, desenvolvido em 1966, que é “um programa que torna a conversação em linguagem natural com um computador possível” (Weizenbaum, 1966).

Chatbots são agentes de software utilizados para a interação entre um computador e um humano em linguagem natural, conforme define Sharma *et al.* (2017), recebendo entradas e enviando respostas baseadas no que foi recebido. De acordo com Shawar e Atwell (2003), *Chatbots* são sistemas conversacionais que utilizam IA, cujas funcionalidades variam de questões e tarefas simples até interações complexas. O atual interesse em *Chatbots* está fortemente relacionado aos avanços na computação e na adoção de aplicativos de troca de mensagens (Molnár; Szüts, 2018).

Atualmente, as motivações mais frequentes para o uso de *Chatbots* são a produtividade e o entretenimento, como diz Brandtzaeg e Følstad (2017), porém também estão presentes em diversos contextos, sendo amplamente utilizados para preencher a demanda por interações mais rápidas, eficientes e personalizadas. Um exemplo de aplicação popular é dentro da área de atendimento ao cliente, devido à redução aos custos com o serviço e a possibilidade de lidar com muitos usuários ao mesmo tempo (Adamopoulou; Moussiades, 2020). Segundo estimativas, o mercado global de *Chatbots* deverá alcançar a marca de US\$ 15,5 bilhões de dólares até o ano de 2028 (MarketsAndMarkets, 2023).

¹ Harlan Ellison, *I Have No Mouth, And I Must Scream* (IF: Worlds of Science Fiction, 1967)

2.1.1 *Classificação de Chatbots*

Os *Chatbots* podem ser classificados utilizando diferentes parâmetros, como o domínio de conhecimento, serviço providenciado, metas, processamento de entradas e produção de saídas, contribuição humana e o método de desenvolvimento, além de poderem ser baseados em voz, texto ou um formato híbrido (Adamopoulou; Moussiades, 2020). Além disso, as implementações de um *Chatbot* podem utilizar desde padrões de palavras-chave dentro de repositórios, até PLN e Aprendizado de Máquina (Brandtzaeg; Følstad, 2017; Molnár; Szüts, 2018).

Em relação ao domínio de conhecimento, *Chatbots* podem ser classificados como domínio aberto, podendo falar sobre tópicos gerais, ou como domínio fechado, sendo focados em conhecimentos específicos. A principal diferença entre estes tipos de *Chatbots* está no conteúdo que o sistema pode acessar e a quantidade de dados em que é treinado (Adamopoulou; Moussiades, 2020).

Quanto ao serviço providenciado, eles se diferenciam em interpessoais, intrapessoais e interagentes (Adamopoulou; Moussiades, 2020). Os *Chatbots* interpessoais são sistemas que possuem características que os aproximam do comportamento humano, podendo ter uma personalidade ou lembrar de informações do usuário, porém possuem o objetivo de providenciar serviços para o usuário, sem demonstrar proximidade com este. Já os intrapessoais são aqueles que serão companheiros do usuário, podendo realizar tarefas e compreender o usuário como um ser humano. Além disso, os *Chatbots* também podem ser denominados interagentes, ou seja, dispositivos que podem se comunicar entre si.

Sobre metas, os *Chatbots* podem ser divididos entre informativos, conversacionais ou aqueles com foco em tarefas (Adamopoulou; Moussiades, 2020; Kucherbaev *et al.*, 2018). *Chatbots* informativos são projetados para disponibilizar dados previamente armazenados em bancos de dados ou em fontes estruturadas. Já os conversacionais falam com o usuário respondendo de acordo com as mensagens que recebe. Por fim, os *Chatbots* orientados a tarefas são capazes de obter informações adicionais e executar ações específicas ou auxiliar na realização de atividades.

No que se refere ao processamento de entradas e geração de respostas, existem diversas abordagens de desenvolvimento. A abordagem baseada em regras consiste em ter um sistema que possui um conjunto fixo de regras predefinidas, com objetivos específicos, sem a flexibilidade de respostas ou habilidades. Já a abordagem generativa utiliza de modelos de IA e

Machine Learning para realizar o processamento de dados, pois geram respostas mais parecidas com a forma como um humano responderia, independente da mensagem enviada (Ramesh *et al.*, 2017).

Além disso, os *Chatbots* podem ser classificados quanto à quantidade de ajuda humana em seu funcionamento, ou seja, *Chatbots* que, em algum elemento, podem utilizar da mediação humana para concluir solicitações (Kucherbaev *et al.*, 2018). O uso de computação humana dentro do contexto de *Chatbots* provê mais flexibilidade e robustez sobre as respostas, embora não alcance a mesma velocidade de processamento das soluções inteiramente automatizadas.

2.1.2 *Chatbots na educação*

No contexto da educação, o uso da IA se mostrou muito proeminente, principalmente nos domínios da Matemática Discreta e Álgebra, segundo Hwang e Tu (2021), por melhorar o desempenho dos estudantes e oferecer um suporte pessoal, reconhecendo as habilidades individuais dos alunos (Brandtzaeg; Følstad, 2017; Moraes; Souza, 2015). Tradicionalmente, a Matemática é ensinada como uma matéria de certo e errado, desconsiderando que os estudantes possuem diferentes habilidades, estilo de aprendizagem e velocidade o que é uma fonte frequente de ansiedade (Stuart, 2000; Boaler, 2002). Nesse contexto, os *Chatbots* podem auxiliar os estudantes a superarem suas dificuldades por meio de um ensino personalizado, capaz de se adaptar às necessidades individuais dos alunos.

2.2 Interfaces Conversacionais

Interfaces conversacionais são um tipo de interface baseado no diálogo, onde a interação se dá por meio de uma conversação entre o usuário e o sistema. Candello e Pinhanez (2016) definem 4 tipos de interfaces conversacionais:

- **Diálogos baseados em fala:** são interfaces onde a entrada é baseada em comandos de voz e a saída em áudio. Exemplos desse tipo de interface são a Alexa e a SIRI.
- **Diálogos baseados em texto:** agentes conversacionais que interagem com usuários por meio de linguagem natural em texto. Aqui entram os *Chatbots*, como por exemplo o ChatGPT e o *Mathos AI*, objeto de estudo deste trabalho.
- **Agentes virtuais interativos:** agentes conversacionais que possuem propriedades similares

aos humanos, porém de forma animada, como por exemplo, os Sims do jogo The Sims.

- **Robôs conversacionais:** máquinas que possuem corpos físicos e que podem ter comportamentos humanos durante a interação, como a personagem Ava do filme *Ex_Machina*² ou o robô NAO³.

Atualmente, a maioria das interfaces conversacionais é construída com uma abordagem intenção-ação, segundo Candello e Pinhanez (2016), ou seja, o sistema é criado a partir de um conjunto de expressões tanto do usuário como do sistema, e regras sobre como essas expressões são correspondidas entre si.

2.2.1 *Design de interfaces conversacionais*

De acordo com Candello e Pinhanez (2016), “a experiência de conversação entre seres humanos e não-humanos é complexa”, pois os humanos aplicam contextos e expectativas de uma interação humana sobre os computadores. Nesse contexto, ao desenvolver uma interface conversacional, é importante considerar a naturalidade da interação para a garantia de uma boa experiência do usuário. Existem diversos fatores a serem considerados para a criação de um diálogo de sucesso. Silva e Canedo (2024) propõem um guia para auxiliar no design de conversas para agentes que se comunicam via texto, incluindo práticas a serem evitadas, boas práticas, estratégias adequadas e um caso de uso.

2.2.1.1 *Práticas propostas para o design conversacional*

Uma das boas práticas propostas por Silva e Canedo (2024) é a **naturalidade**. Segundo o autor, *Chatbots* devem participar de conversas utilizando de práticas que podem não fazer parte do seu objetivo principal, porém fazem parte de qualquer interação entre humanos. A seguir, apresentaremos boas práticas que o *Chatbot* deve incluir:

- **Apresentação:** *Chatbots* devem iniciar a conversa se apresentando como agentes virtuais, dizendo seu nome e seu objetivo para o usuário;
- **Referir ao usuário pelo seu nome:** A depender do contexto da aplicação, é possível que o usuário possua alguma identificação, como o nome de usuário. Neste caso, o *Chatbot* deve sempre se referir ao usuário pelo seu nome;
- **Uso de conversa fiada:** Apesar de não estar de acordo com a máxima conversacional

² <https://www.imdb.com/pt/title/tt0470752/>

³ <https://aldebaran.com/en/nao6/>

de falar apenas o necessário, diálogos que não possuem objetivo, como saudações e agradecimentos, fazem parte de uma conversa natural entre seres humanos e, portanto, devem ser incluídos na conversa com o *Chatbot*;

- **Ecoar respostas:** Evitar respostas curtas, como um simples “Ok”, e, em vez disso, enfatizar a mensagem que o usuário mandou por meio de paráfrases;
- **Linguagem casual:** O uso de uma linguagem casual e informal durante a interação faz com que o sistema pareça menos como uma fonte de informação e mais como um agente conversacional. Entretanto, o nível de casualidade deve estar de acordo com o objetivo e público-alvo do *Chatbot*.

Também, é importante considerar a **emocionalidade** do *Chatbot*, ou seja, a sua capacidade de expressar emoções e compreender os sentimentos do usuário. A seguir, discutiremos melhor essas características:

- **Feedback exclamatório:** Utilizar da pontuação para expressar emoções dentro de uma simples mensagem, como o uso de pontos de exclamação para expressar entusiasmo;
- **Mídia gráfica:** A depender do contexto da aplicação, é possível que o *Chatbot* faça o uso de mídias como imagens e emojis para tornar a conversa mais leve;
- **Mensagens enfáticas:** Adaptar as mensagens de acordo com os sentimentos do usuário, de modo a demonstrar empatia. Por exemplo, o *Chatbot* pode expressar felicidade quando o usuário realiza alguma meta mencionada;
- **Humor:** Para o desenvolvimento de uma conversa leve e agradável, é possível que o *Chatbot* faça o uso de piadas ou histórias para melhorar o humor da conversa.

Um outro fator importante para uma conversa agradável com um *Chatbot* é a **transparência**. O sistema deve ser claro sobre si mesmo e suas habilidades, para alinhar as expectativas do usuário. A seguir, apresentaremos algumas características que podem auxiliar com a transparência do sistema:

- **Habilidades presentes:** *Chatbots* devem iniciar a conversa apresentando o objetivo e as habilidades do sistema;
- **Reconhecer limitações:** No início de uma conversa ou após uma falha, o sistema deve apresentar suas limitações, para evitar perguntas indesejadas dos usuários;
- **Realizar sugestões:** *Chatbots* que possuem bases de conhecimento variadas podem sugerir tópicos de conversa para auxiliar o usuário a permanecer de acordo com as habilidades do usuário;

- **Pedir clarificações:** Se o *Chatbot* não entender a requisição do usuário, o sistema deve pedir ao usuário para refazer sua pergunta.

2.2.1.2 *Práticas a serem evitadas*

Além de boas práticas, existem algumas características que podem prejudicar a experiência do usuário ou invalidar as boas práticas descritas na seção 2.2.1.1. Uma das práticas descritas por Silva e Canedo (2024) são as mensagens repetitivas, pois fazem com que o *Chatbot* pareça robótico, o que impede a naturalidade da conversação. Exageros também devem ser evitados, pois podem trazer efeitos opostos àqueles esperados. Também, o *Chatbot* não deve "fingir" ser um humano, ou seja, é importante que o sistema deixe claro que se trata de uma máquina falando com o usuário. Além disso, por mais que erros façam parte de uma conversa natural entre humanos, a inserção proposital de erros compromete a confiabilidade do *Chatbot*. Por fim, o sistema deve evitar fontes que pareçam mecânicas ou computadorizadas, optando por fontes mais legíveis e amigáveis para o usuário.

2.3 *Mathos AI*

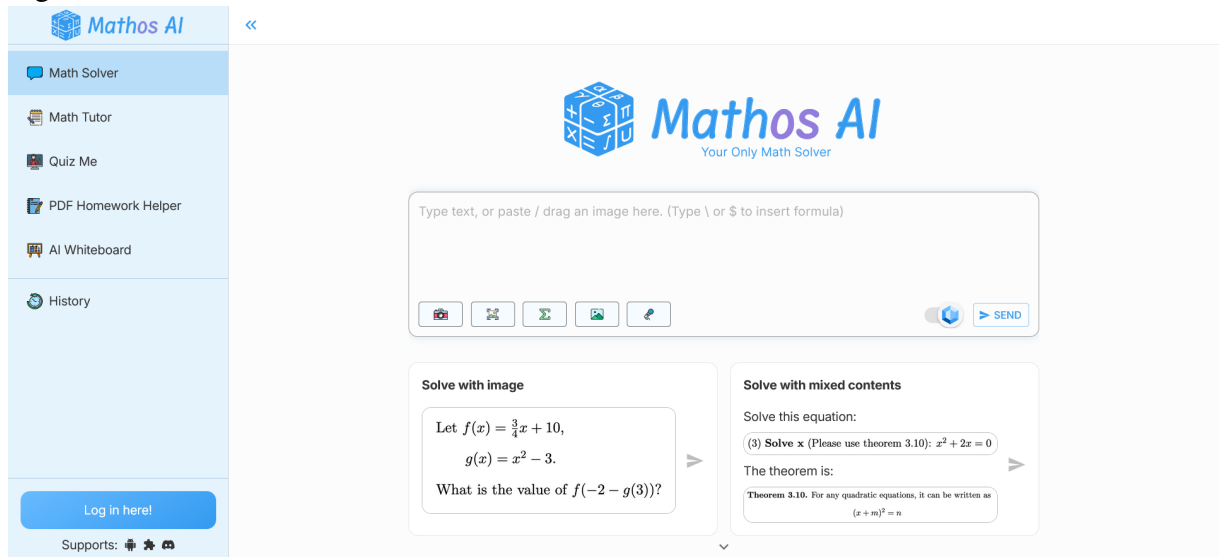
Sob o contexto de estudo apoiado por tecnologia, surgiram diversas soluções para auxiliar os estudantes no seu processo de aprendizado, entre elas, o uso de *Chatbots* tutores. Neste trabalho, foi utilizado o *Chatbot Mathos AI*, um tutor de matemática baseado em IA⁴.

O aplicativo foi desenvolvido pela Upstage, uma startup sul-coreana, como um “GPT de matemática” utilizando dados da QANDA⁵ (LeeSang-Gu *et al.*, 2024). O aplicativo está disponível em uma versão web, mas também como um aplicativo mobile. É possível ver capturas de tela nas figuras 1 e 2.

⁴ <https://info.mathos.ai/>

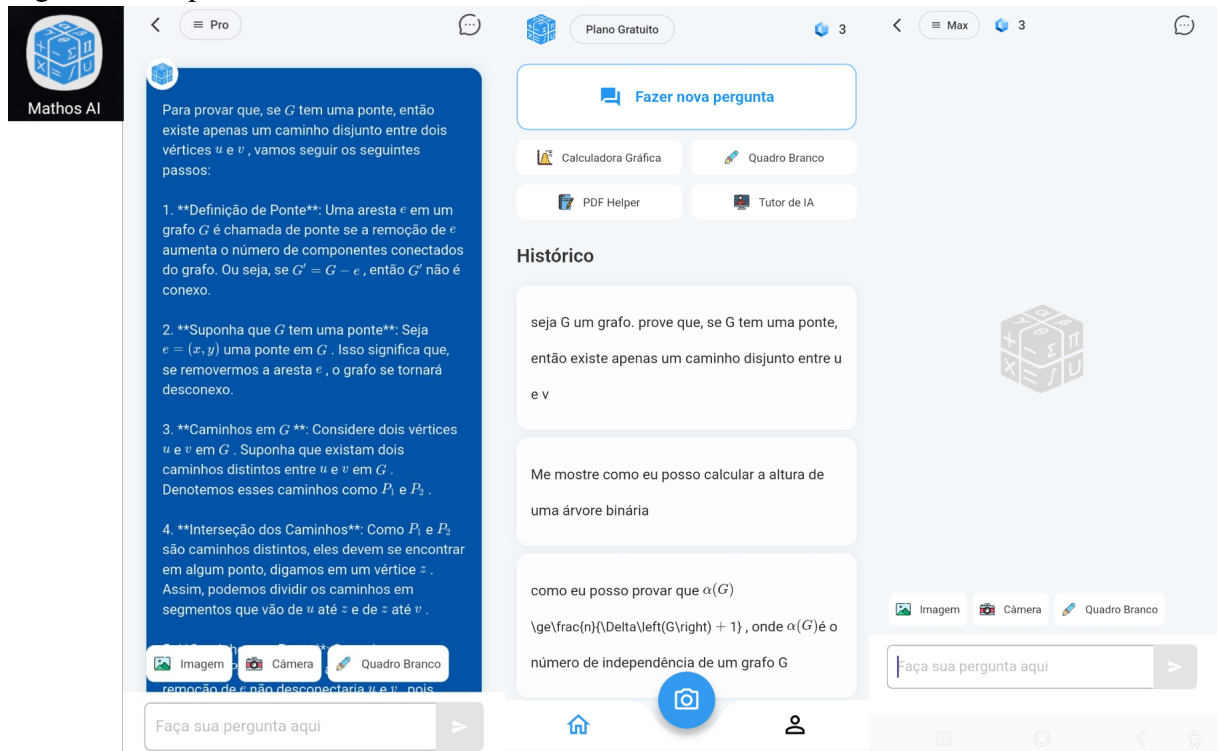
⁵ <https://qanda.ai/pt>

Figura 1 – Versão web do *Mathos AI*



Fonte: Adaptado de www.mathgptpro.com

Figura 2 – Capturas de tela da versão mobile



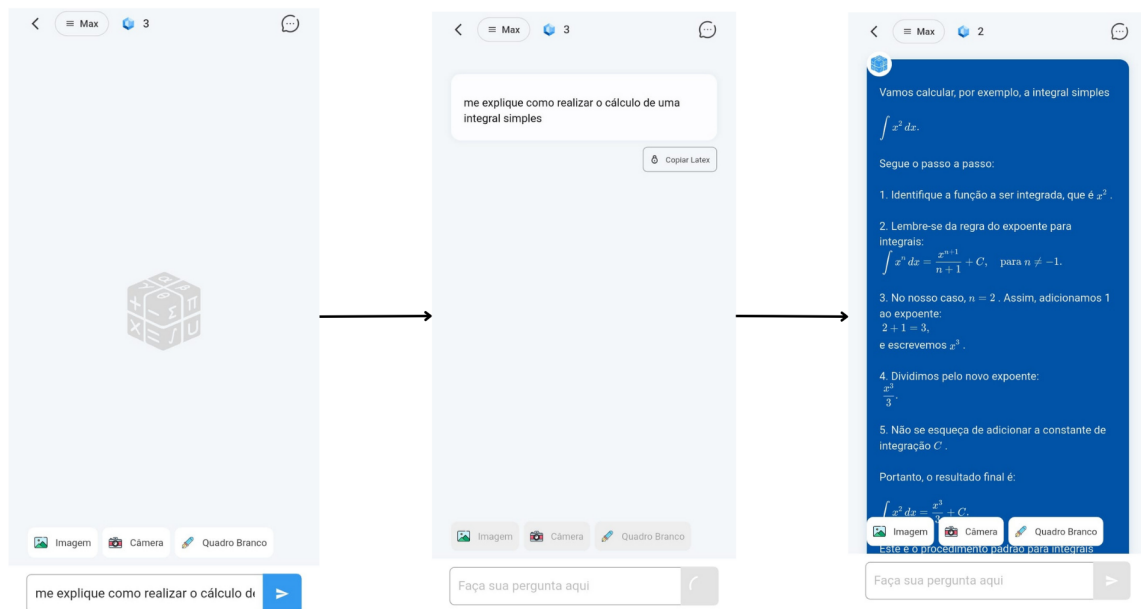
Fonte: Elaborado pela autora. Adaptado de www.mathgptpro.com

Segundo Seto *et al.* (2025), a aplicação teve um impacto positivo e significativo na melhoria das habilidades na resolução de problemas matemáticos de estudantes, auxiliando na compreensão do problema, no planejamento e execução da solução e na verificação das respostas.

O funcionamento do aplicativo funciona como uma aplicação de *Chatbot* qualquer: o usuário envia uma requisição de texto para o sistema, e, com o uso de PLN e *Deep Learning*,

o aplicativo retorna a resposta considerada mais adequada para a requisição. Na Figura 3, é possível ver de forma gráfica um exemplo de uso do sistema.

Figura 3 – Representação de um caso de uso básico do sistema



Fonte: Elaborado pela autora. Adaptado de www.mathgptpro.com

2.4 Engenharia Semiótica

A Engenharia Semiótica é uma teoria de IHC centrada na comunicação, ou seja, uma teoria que investiga a comunicação entre designers, usuários e sistemas (Barbosa *et al.*, 2021). Segundo Souza (2005), a Semiótica é o “estudo dos signos, processos de significações e como os signos e a significação atuam na comunicação”, e, dentro da área de IHC, a investigação é realizada tanto dentro da comunicação direta usuário-sistema como pela metacomunicação através da interface.

Ao contrário de outras teorias de IHC, a Engenharia Semiótica tem como foco a comunicabilidade do sistema, estudando as intenções comunicativas do designer, a lógica e os princípios de interação subjacentes. Dentro desta teoria, as aplicações computacionais são caracterizadas como artefatos de metacomunicação, ou seja, “artefatos que comunicam uma mensagem do designer para os usuários sobre a comunicação usuário-sistema, sobre como eles podem e devem utilizar o sistema, por que e com que efeitos” (Barbosa *et al.*, 2021; Souza, 2005).

Durante o processo de design, o designer realiza um estudo sobre o contexto de uso

do sistema e seus usuários e, a partir deste estudo, expressa sua visão por meio de palavras, gráficos, ícones e explicações, para que os usuários interpretem adequadamente, gostem e se beneficiem do produto (Souza, 2005). Aqui, o sistema se torna o preposto do designer, ou seja, o meio responsável para comunicar ao usuário a metamensagem do designer.

2.4.1 Signo, Significação e Semiose

Segundo Peirce (1934), o signo é “uma coisa que serve para veicular conhecimento de outra coisa que ele representa”. Para que uma representação seja considerada um signo, esta deve possuir um objeto, um *representamen* e um interpretante, para que possa ser realizada a associação de significado a este signo. O processo de significação de um signo, interpretação ou semiose, é um processo infinito, pois, quanto mais pensamos sobre algo significativo, mais significados são associados a ele.

Dentro da área de IHC, os signos são utilizados como forma de comunicar ao usuário a mensagem que o designer deseja entregar por meio do sistema, fazendo uso da metacomunicação. A metacomunicação é um tipo de comunicação que fala sobre a comunicação em si. Dentro de um sistema, a mensagem pode ser considerada única e unidirecional, pois representa um conteúdo único e imutável, desenhado para informar aos usuários o que devem fazer para alcançar um objetivo. Podemos utilizar o seguinte modelo para representar a metamensagem do designer:

Este é o meu entendimento, como designer, de quem você, usuário, é, do que aprendi que você quer ou precisa fazer, de que maneiras prefere fazer, e por quê. Este, portanto, é o sistema que projetei para você, e esta é a forma como você pode ou deve utilizá-lo para alcançar uma gama de objetivos que se encaixam nesta visão. (Souza, 2005, p. 84)

Em tempo de interação, os usuários decodificam e interpretam gradualmente a metamensagem do designer, buscando atribuir sentido aos significados nela codificados e respondendo de forma apropriada.

2.4.2 O Sistema como Artefato de Comunicação

Um artefato é “algo criado pelo ser humano, geralmente para um propósito prático, algo característico ou resultante de uma instituição, período, tendência ou indivíduo particular”. No contexto computacional, um sistema pode ser definido como um artefato, pois resulta de

atividades de análise sobre um problema e sintetiza um conjunto de soluções para dada situação. Souza (2005) define os sistemas como um artefato intelectual, pois:

- A codificação do problema e das soluções são fundamentalmente linguísticas;
- O propósito final só pode ser alcançado por seus usuários se eles conseguem formulá-lo dentro do sistema linguístico.

Ou seja, designers, usuários e seus prepostos devem utilizar uma mesma linguagem e um sistema de signos para que se comuniquem através do sistema.

A Engenharia Semiótica utiliza o modelo de espaço de comunicação proposto por Jakobson e Sebeok (1960) para estruturar o espaço de design de IHC, onde “Um emissor transmite uma mensagem a um receptor através de um canal. A mensagem é expressa em um código e se refere a um contexto. Na comunicação, os interlocutores exercem alternadamente os papéis de emissor e receptor”.

Para a realização da comunicação, os sistemas utilizam uma linguagem interativa única, definida pelo designer, contendo os códigos expressivos que os usuários deverão utilizar para se comunicarem com o sistema. A Engenharia semiótica classifica os signos utilizados dentro do design de interfaces em três tipos (Barbosa *et al.*, 2021; Souza, 2005):

- Signos estáticos, que expressam o estado do sistema e cujo significado é interpretado independentemente de relações causais e temporais da interface;
- Signos dinâmicos, que expressam o comportamento do sistema, envolvendo aspectos temporais e causais da interface, fazendo referência à própria interação;
- Signos metalinguísticos, que se referem a outros signos da interface.

Apesar de o designer possuir um papel ativo dentro do espaço de designer, como um interlocutor que deve auxiliar os usuários a entenderem a metagemagem, este não pode determinar de que maneira os usuários interpretarão os signos ou garantir que a mensagem será recebida de forma correta. Portanto, é necessário que o designer tenha como objetivo introduzir aos usuários um sistema computacional interativo, considerando aspectos estratégicos e operacionais da tecnologia, para que os usuários possam fazer um bom uso do sistema.

2.5 Avaliação de Interação

O processo de avaliação em IHC é uma atividade fundamental pois garante a qualidade do produto final do usuário. Esse processo tem por objetivo verificar se o sistema possui quaisquer problemas na interação e na interface que podem prejudicar a UX, além de analisar se

o sistema apoia adequadamente os usuários a atingirem seus objetivos de uso (Barbosa *et al.*, 2021).

Um aspecto importante da avaliação é definir o *quê* avaliar, ou seja, quais os objetivos, qual o público alvo e por que avaliar. Essa etapa é essencial pois direciona toda a investigação, determinando quais interações, funcionalidades ou aspectos da interface devem ser avaliados. Além disso, ao definir o propósito da avaliação, podemos definir quais métodos e técnicas se adequam melhor às características do objeto de estudo, além de garantir coerência com os objetivos definidos. A partir disso, obtém-se uma base para uma análise consistente, orientando a coleta de evidências e interpretação dos resultados obtidos.

Ao realizar uma avaliação de interação devemos conhecer os principais critérios de qualidade de uso presentes na literatura, sendo eles usabilidade, UX, acessibilidade e comunicabilidade (Barbosa *et al.*, 2021). A usabilidade está relacionada com a facilidade de uso e aprendizado da interface, além como a satisfação de uso do sistema (Nielsen, 1994). De forma similar, temos a UX, que envolve os sentimentos e emoções do usuário. O critério de acessibilidade está relacionado com a remoção de barreiras que impedem que alguns usuários sejam capazes de acessar e interagir com a interface. Já a comunicabilidade fala sobre a responsabilidade do designer de comunicar ao usuário suas intenções de design e a lógica que rege o comportamento da interface.

2.5.1 Métodos de Avaliação

Os métodos de avaliação de IHC são ferramentas para a realização da avaliação em diferentes aspectos do sistema (Barbosa *et al.*, 2021). Os métodos são divididos em três principais tipos, sendo eles métodos de Inspeção, Observação e Investigação.

2.5.1.1 Métodos de Inspeção

Segundo Barbosa *et al.* (2021) os “[...]métodos de inspeção permitem ao avaliador examinar (ou inspecionar) uma solução de IHC para tentar prever as possíveis consequências de certas decisões de design.”. Esses métodos não envolvem o usuário diretamente, portanto, cabe ao avaliador se colocar no papel do usuário identificar potenciais problemas na interface e dificuldades enfrentadas pelo usuário ao interagir com o sistema.

Os métodos de inspeção são geralmente empregados durante o desenvolvimento para examinar uma solução de IHC para tentar prever as possíveis consequências de certas decisões

de design dentro do sistema, ou seja, tentar identificar problemas que os usuários podem ter ao interagir com os sistema.

Um exemplo de método de Inspeção é a Avaliação Heurística. A Avaliação Heurística consiste em avaliadores analisarem a interface de forma sistemática, visando encontrar problemas que prejudiquem a usabilidade. Para avaliar a usabilidade do sistema, esse método usa como base as heurísticas de Nielsen (1994), onde o avaliador pode medir a usabilidade a partir de quantas vezes cada heurística não é seguida adequadamente. Abaixo, serão listadas as heurísticas de Nielsen:

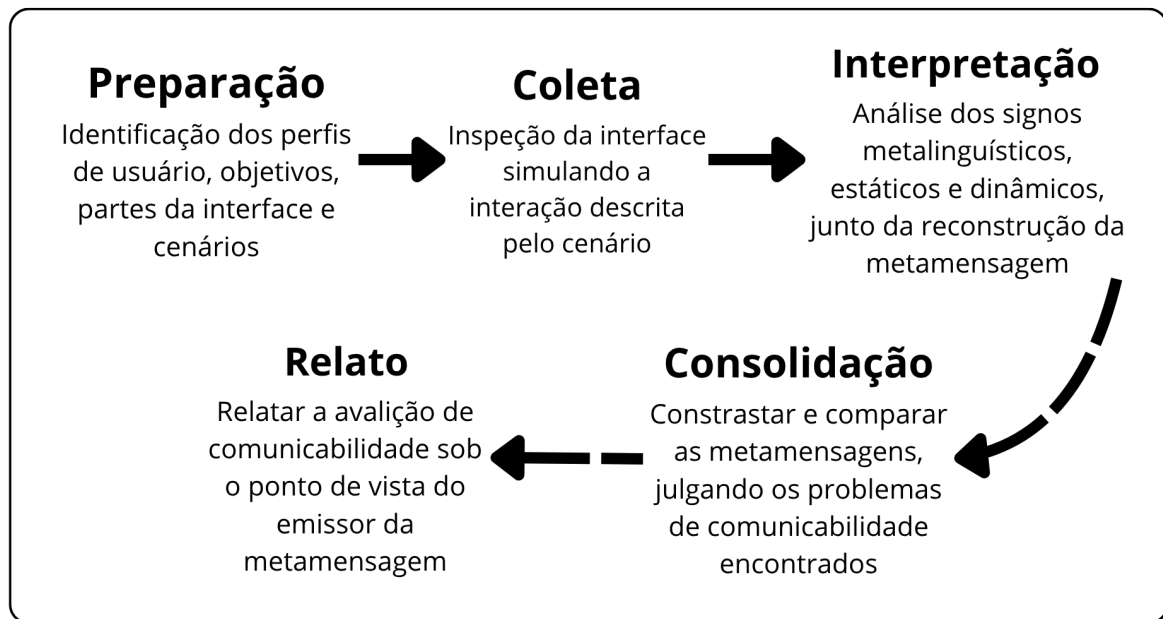
- Visibilidade do status do sistema;
- Correspondência entre o sistema e o mundo real;
- Controle e liberdade para o usuário;
- Consistência e padronização;
- Reconhecimento em vez de recordação;
- Flexibilidade e eficiência de uso;
- Projeto estético e minimalista;
- Prevenção de erros;
- Ajude os usuários a reconhecerem, diagnosticarem e se recuperarem de erros;
- Ajuda e documentação.

Outro método de Inspeção é o MIS. O MIS, segundo Barbosa *et al.* (2021), “[...] *Java-lia a comunicabilidade de uma solução de IHC por meio de inspeção[...]*”, tendo como principal objetivo avaliar a qualidade da metacomunicação entre usuário e designer (Barbosa *et al.*, 2021). A principal métrica deste método são os signos codificados na interface, sendo classificados como estáticos, dinâmicos e metalinguísticos. A partir destes signos, o avaliador deve realizar a reconstrução da metamensagem do designer de acordo com cada tipo de signo, para que, ao final, possa realizar a comparação das metamensagens e realizar o julgamento sobre a comunicabilidade do sistema. O método de avaliação pode ser exemplificado na Figura 4.

2.5.1.2 Métodos de Observação

Segundo Barbosa *et al.* (2021), “[...] *os métodos de observação permitem ao avaliador coletar dados sobre situações em que os participantes realizam suas atividades[...]*”, assim, os métodos classificados como de observação consistem em observar usuários interagindo com o sistema em tempo real, podendo então o avaliador tirar conclusões sobre problemas reais

Figura 4 – Método de avaliação MIS



Fonte: Elaborado pela autora, adaptado de Barbosa *et al.* (2021)

durante a interação entre o usuário e a interface.

Um exemplo de Método de Observação é o Teste de Usabilidade. Segundo Rubin e Chisnell (2008), o teste de usabilidade visa avaliar o sistema a partir da interação e da experiência de uso de seus usuários-alvo. Neste teste, o avaliador deverá selecionar os participantes e as tarefas a serem realizadas dentro de um ambiente controlado, e, a partir da observação dos usuários na realização das tarefas, coletar diferentes dados e métricas sobre a usabilidade do sistema a partir de critérios de usabilidade.

O Teste de Usabilidade é conduzido por meio de diferentes atividades. Durante a **preparação**, serão definidas as tarefas a serem executadas pelos participantes, além da preparação de material e a execução de um teste piloto. Depois, na fase de **coleta**, o avaliador deverá observar e registrar a performance e a opinião dos participantes durante a sessão; aqui, é possível medir o grau de sucesso, o total de erros cometidos, o grau de satisfação do usuário, entre outras métricas. Então, na fase de **consolidação**, os dados dos participantes serão organizados de modo a evidenciar as relações entre eles, como pelo o cálculo de médias, índices e contagens (Barbosa *et al.*, 2021).

Outro Método de Observação amplamente utilizado é o Método de Avaliação de Comunicabilidade (MAC). De acordo com Barbosa *et al.* (2021), “O método de avaliação de comunicabilidade visa apreciar a qualidade da comunicação da metamensagem do designer

para os usuário”. Sendo assim, o modelo consiste em fazer uso dos signos definidos pela engenharia semiótica para analisar a qualidade da recepção da metamensagem ao observar o usuário.

De forma similar ao Teste de Usabilidade, o MAC é realizado por meio da observação de usuários em um ambiente controlado. Na fase de **Preparação**, o avaliador deve realizar uma inspeção prévia dos signos da interface, para auxiliar na definição de tarefas e preparação do material; além disso, o avaliador deverá recrutar os participantes e realizar um teste-piloto. Durante a **Coleta**, o avaliador observa as sessões de uso, além de realizar registros por meio de anotações e gravações de cada participante. Na fase de **interpretação**, cada vídeo é etiquetado de acordo com as rupturas de comunicação encontradas, categorizadas por expressões de comunicabilidade. Por fim, na fase de **consolidação**, o avaliador deve interpretar o significado do conjunto de todas as etiquetas nos vídeos de interação, por meio de fatores como frequência e contexto, sequências e nível, além de poder utilizar os dados dentro de outras ontologias e classes. (Barbosa *et al.*, 2021).

2.5.1.3 Métodos de Investigação

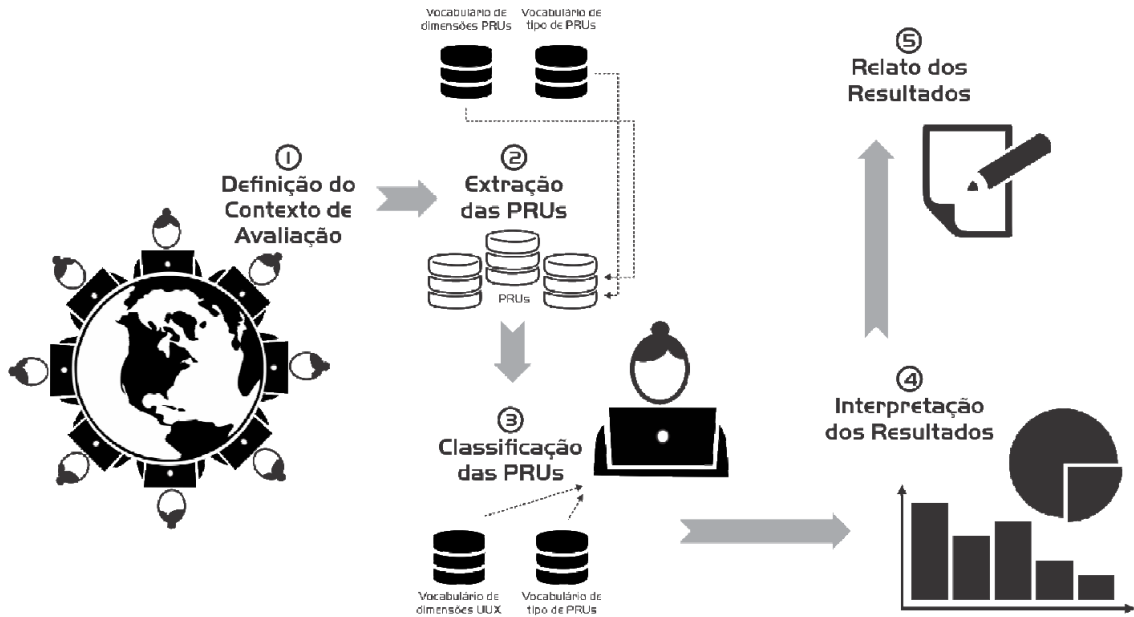
Os Métodos de Investigação são métodos utilizados para identificar a situação atual do sistema, coletando dados a partir de uma variedade de fontes para identificar quais funcionalidades devem ser incluídas ou aprimoradas no sistema (Barbosa *et al.*, 2021). Dentro desses métodos são coletados dados sobre os usuários, como dados demográficos e conhecimento do domínio, e quais os principais objetivos e requisitos do usuário.

Um dos métodos de Investigação mais utilizados frequentemente são os questionários. Segundo Barbosa *et al.* (2021), “*Um questionário é um formulário impresso ou on-line com perguntas que os usuários e demais participantes devem responder, a fim de fornecer os dados necessários em uma pesquisa, análise ou avaliação.*”, que permite a coleta de dados de um grande volume de pessoas, além de poder obter informações sobre diferentes aspectos por meio de perguntas que podem ser abertas ou fechadas.

Outro método, proposto por Mendes (2015) é o MALTU. O MALTU é um método que busca apresentar uma visão geral de um sistema a partir de um conjunto de Postagem Relacionada ao uso do Sistema (PRU). A extração de PRUs pode ser realizada tanto manualmente como de forma automática com o apoio de ferramentas. Após a extração, as PRUs são classificadas de acordo com o seu **tipo, intenção**, pela análise dos sentimentos por **polaridade e intensidade**,

pela **funcionalidade** e pelos **critérios de qualidade**, além de poderem incluir o dispositivo que o usuário está utilizando. Por fim, os dados obtidos devem ser organizados de modo a evidenciar relações entre eles, como definido por Mendes (2015), como por exemplo, pela frequência de cada tipo de PRU e pela frequência em que as funcionalidades são citadas. A metodologia pode ser demonstrada na Figura 5.

Figura 5 – Método MALTU



Fonte: Adaptado de Mendes (2015).

3 TRABALHOS RELACIONADOS

Nesta seção, serão apresentados e discutidos os trabalhos relacionados que serviram como base e referência para o planejamento e execução da metodologia utilizada. Esses estudos, ao abordarem diferentes perspectivas e metodologias de avaliação aplicadas a sistemas interativos, constituem referências relevantes que contribuíram para o delineamento deste trabalho.

3.1 STUART: An Intelligent Tutoring System for Increasing Scalability of Distance Education Courses

No contexto de aprendizado à distância, Damasceno *et al.* (2020) identificaram o desequilíbrio entre o número elevado de estudantes e a quantidade de tutores disponíveis. Como solução apresentada, eles desenvolveram um Sistema de Tutor Inteligente para reduzir a carga dos tutores em cursos online dentro da plataforma Dell Accessible Learning¹, que visa otimizar o suporte pedagógico no aprendizado.

O STUART é um *Chatbot* que utiliza de PLN, *Machine Learning*, e integração com as ferramentas da plataforma para mandar recomendações pedagógicas de acordo com o perfil do estudante e responder às demandas técnicas, pedagógicas e estudantis comumente designadas a um tutor. Seu funcionamento baseia-se em uma requisição feita pelo cliente, que é encaminhada ao agente conversacional Dialogflow². Se o sistema não o encontrar, então o *Chatbot* utiliza um modelo de *Machine Reading Comprehension* (MRC) para produzir uma resposta. Adicionalmente, o *Chatbot* avalia respostas de atividades, oferecendo feedback automático aos estudantes.

Dentro da sua avaliação, Damasceno *et al.* (2020) realizaram um estudo piloto para verificar se o *Chatbot* auxiliava a reduzir as atividades humanas no sistema para identificar o impacto que o *Chatbot* oferece na escalabilidade. A partir disso, os autores contataram estudantes de dentro da plataforma, explicando o objetivo do teste e o contexto da pesquisa.

Os participantes realizaram seis tarefas dentro de um curso sobre Atendimento ao Cliente descritas por meio de um cenário desenvolvido a partir de uma pesquisa prévia, onde os estudantes deveriam realizar diferentes atividades com e sem o apoio da ferramenta. Dentro do teste, os autores realizaram uma análise por meio de diferentes perspectivas. Primeiramente, realizaram uma análise de performance com dados sobre o tempo para completar as tarefas, a

¹ Dell Accessible Learning

² Dialogflow

quantidade de tarefas concluídas e o quanto de ajuda foi necessário para cada tarefa. Depois, realizaram um questionário de usabilidade baseado no *System Usability Scale* (SUS) e, por fim, realizaram uma análise sobre sistemas que atuam de forma similar (Damasceno *et al.*, 2020).

Como resultados, foi identificado uma redução no uso de assistências humanas nas tarefas após o uso do *Chatbot*, além de observarem que mais estudantes obtiveram respostas utilizando o *Chatbot*. Além disso, o teste de usabilidade classificou a satisfação dos usuários do sistema como excelente, além de identificar que o sistema é fácil de usar, intuitivo e transmite confiança.

O trabalho oferece uma solução de implementação de um *Chatbot* seguido de uma avaliação. Ao contrário de Damasceno *et al.* (2020), este trabalho busca apenas realizar uma avaliação. Entretanto, de forma similar, o trabalho de Damasceno *et al.* (2020) realiza uma combinação de diferentes métodos, com o objetivo de identificar uma visão geral de usabilidade e qualidade do sistema.

3.2 Avaliação de Usabilidade do CoderBot como Recurso Pedagógico no Ensino de Programação

No trabalho de Miranda *et al.* (2024), os autores desenvolveram um *Chatbot* voltado para o ensino de programação. O *Chatbot* visa “proporcionar um ambiente de aprendizagem em que os estudantes possam praticar exercícios de programação.” (Miranda *et al.*, 2024).

O *Chatbot* desenvolvido utiliza da Aprendizagem Baseada em Exemplos, oferece aos estudantes exemplos práticos que reduzem a carga cognitiva, facilitando a compreensão e o desenvolvimento de esquemas mais generalizados e aplicáveis a diferentes contextos de ensino.

O sistema foi integrado a um ambiente web e estruturado com base em regras e interações por botões, onde serão apresentados desafios de programação de acordo com determinado conteúdo, incluindo sua descrição, o resultado esperado e instruções do teste. Aqui, o usuário pode escolher entre visualizar um exemplo correto ou incorreto de solução, e, ao selecionar a visualização de um exemplo incorreto, o usuário é convidado a identificar o erro, recebendo feedback automatizado do *Chatbot* de acordo com sua resposta.

Para a avaliação, foram selecionados 103 estudantes de graduação em disciplinas introdutórias de programação que deveriam resolver problemas utilizando o *Chatbot*. Inicialmente, um estudo piloto foi conduzido para validar o roteiro de testes. Em seguida, os autores instruíram os professores responsáveis para orientarem os alunos no uso da ferramenta. Após

a interação com o sistema, os participantes responderam ao questionário SUS para mensurar a capacidade de aprendizagem, eficiência, capacidade de memorização, minimização de erros e satisfação de aplicações.

Após a avaliação, o *Chatbot* obteve um *score* de 79,2 pontos, podendo ser classificado como média usabilidade (Miranda *et al.*, 2024). Os participantes concordaram que o sistema era fácil de utilizar, porém possui uma falta de exemplos práticos e funcionalidades limitadas.

O trabalho de Miranda *et al.* (2024) se relaciona com este Trabalho de Conclusão de Curso (TCC) por apresentar uma avaliação de um *Chatbot* que atua dentro do contexto de educação, mesmo que voltado para o ensino de programação. O trabalho emprega um teste seguido de um questionário para a obtenção de métricas quantitativas sobre a usabilidade do sistema. Entretanto, ao contrário do trabalho apresentado, este TCC busca apenas fazer a avaliação de um sistema de *Chatbot*, sem propor uma nova solução. Além disso, neste trabalho serão utilizados outros métodos de avaliação, de modo a obter uma visão mais geral sobre o estado do sistema.

3.3 Investigando chatbots governamentais: um panorama sobre a usabilidade dentro e fora do Brasil

No contexto do uso de *Chatbots* no ambiente governamental, Batista *et al.* (2022) identificaram diversos desafios no uso da tecnologia, como a complexidade em atender às expectativas dos usuários, mantê-los engajados com o uso, além de problemas relacionados à usabilidade de *Chatbots*. Diante desse cenário, os autores ressaltaram a necessidade de medir a qualidade da entrega de acordo com a qualidade da interação dos usuários com a tecnologia.

O objetivo principal do trabalho é identificar os impactos das violações de padrões de usabilidade dentro dos *Chatbots* governamentais nacionais e estrangeiros, buscando identificar a qualidade das soluções apresentadas. Para a realização desse objetivo, os autores realizaram uma avaliação utilizando das heurísticas de Nielsen (1994), adaptadas por Langevin *et al.* (2021) para a avaliação de *Chatbots*. Na Figura 6, é possível visualizar as heurísticas utilizadas, juntamente de suas descrições.

O estudo foi conduzido em quatro etapas por três avaliadores que possuíam experiência variada na área de IHC. Na fase de preparação, aconteceu o mapeamento e busca dos *Chatbots*, onde foram selecionados diferentes sistemas tanto brasileiros como estrangeiros. Após uma filtragem dos resultados, os pesquisadores avaliaram 10 *Chatbots*, sendo 5 brasileiros e 5

Figura 6 – Heurísticas de Langevin.

Heurística	Descrição
Visibilidade do estado do sistema	O chatbot sempre deve informar ao usuário sobre o status do sistema, dentro de um tempo razoável.
Correspondência entre o Sistema e o mundo real	O sistema deve entender e falar a linguagem dos usuários. Sem termos orientados ao sistema ou terminologias confusas.
Controle e Liberdade do usuário	O sistema deve dar suporte para o usuário desfazer e refazer comandos.
Consistência e Padrão	Os usuários não devem se perguntar se palavras, opções ou ações diferentes significam a mesma coisa. O chatbot deve ser capaz de enviar respostas consistentes, mesmo que comuniquem a mesma função de várias maneiras (e modalidades).
Prevenção de erros	O chatbot deve prevenir ou eliminar prontamente possíveis condições propensas a erros, verificando e confirmando com os usuários antes que eles executem uma ação.
Ajuda e orientação	O sistema deve orientar o usuário ao longo do diálogo, esclarecendo suas capacidades e ter os recursos de ajuda fáceis de recuperar e pesquisar.
Flexibilidade e eficiência de uso	O sistema deve ser eficiente tanto para especialista quanto para novatos, fornecendo aos usuários a modalidade e o hardware de entrada e saída apropriados (ou preferidos).
Design estético, minimalista e envolvente	O chatbot deve fornecer elementos de interação necessários para envolver o usuário e se adequar ao objetivo do sistema. Suportando interações curtas e expandindo a conversa, se o usuário escolher.
Ajude o usuário a reconhecer, diagnosticar e se recuperar de erros	O sistema deve indicar com precisão e clareza o problema (mensagens de erros), e sugerir uma solução de forma construtiva.
Preservação de contexto	O chatbot deve permitir ao usuário fazer referência a mensagens anteriores, fazendo relação entre os tópicos das sessões.
Confiabilidade	O sistema deve ser transparente quanto ao uso dos dados e de como o chatbot irá ser apresentado ao usuário.

Fonte: Adaptado de (Batista *et al.*, 2022; Langevin *et al.*, 2021; Nielsen, 1994).

estrangeiros.

Para a realização da coleta de dados, os avaliadores interagem com cada *Chatbot* para a realização de uma determinada função presente no sistema, com a funcionalidade variando de acordo com o objetivo do *Chatbot*. Durante esta fase, todas as violações foram registradas juntamente de capturas da interface, além de serem analisadas com base no grau de severidade das violações, sendo divididos em Cosméticos, Pequeno, Grande e Catastrófico (Batista *et al.*,

2022).

Durante a consolidação dos resultados, foi observado que a heurística de Controle e Liberdade do Usuário foi a mais violada, estando presente em 6 dos 10 *Chatbots*, sendo seguida pela heurística Visibilidade do Status do Sistema. Dentro dos *Chatbots* brasileiros, foi identificado um alto índice de problemas catastróficos, ou seja, violações que provocam grandes problemas de usabilidade; além das violações relacionadas à heurística de Controle e Liberdade do Usuário, foram observadas falhas com relação à heurística de recuperação de erros. Em relação aos *Chatbots* estrangeiros, houve um grande número de violações sobre a Visibilidade do Status do Sistema, além de problemas com a heurística de Prevenção de Erros.

Ao comparar os *Chatbots* selecionados, foram identificados problemas comuns em ambos os cenários de *Chatbots* estrangeiros e nacionais, como por exemplo, o sistema não permitir que o usuário retorne ao menu anterior (Batista *et al.*, 2022). Além disso, também foram identificados problemas sobre a confiabilidade dos sistemas avaliados, em que foram identificados casos em que o sistema não assegurava a privacidade dos dados solicitados ou sobre o sistema ser ou não automatizado.

O trabalho de Batista *et al.* (2022) se relaciona com este trabalho por se tratar de uma avaliação de *Chatbots* com o uso de métodos de Avaliação de IHC. Apesar de possuir uma metodologia diferente da proposta neste trabalho, a avaliação de Batista *et al.* (2022) permite identificar aspectos da usabilidade em *Chatbots* que não se restringem ao contexto governamental, mas que podem ser identificados em outras situações.

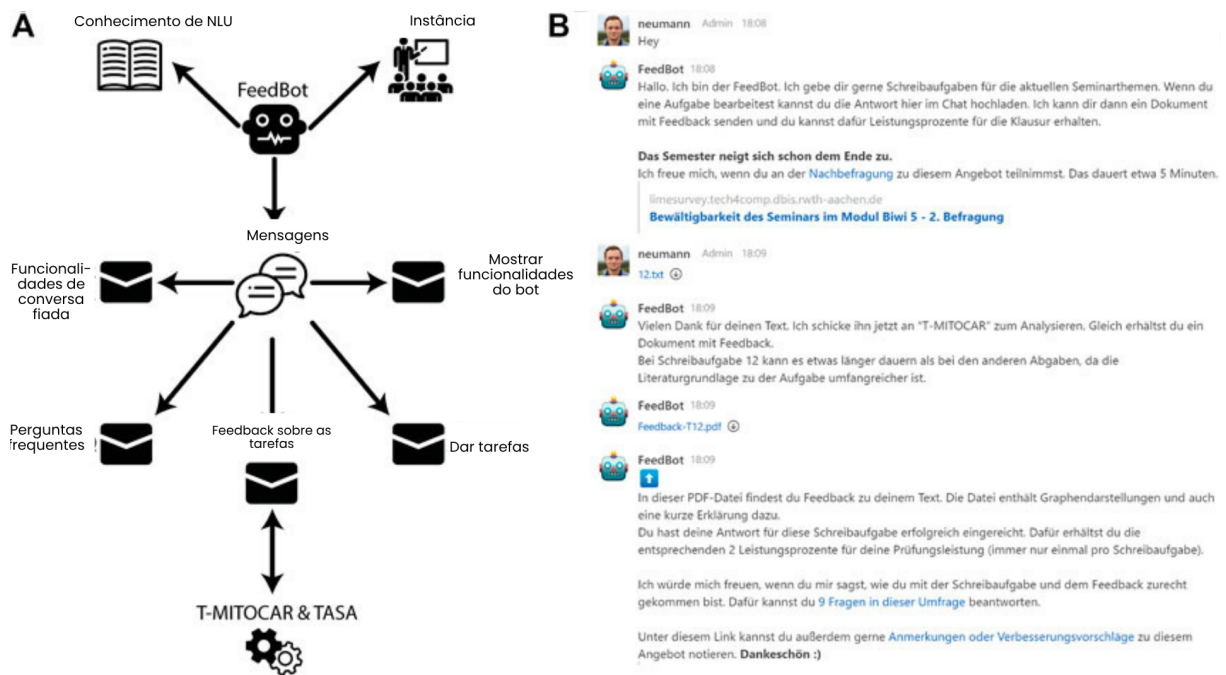
3.4 Chatbots Como Uma Ferramenta Para Escalar Processos de Mentoria: Individualmente Apoiando o Auto-Estudo na Educação Superior

O trabalho de Neumann *et al.* (2021) explora o uso de *Chatbots* para escalar processos de mentoria no ensino superior, especificamente em cursos que possuem uma grande carga de leitura e escrita. Sob esta perspectiva, os autores desenvolveram dois *Chatbots* para oferecer suporte e feedback para os alunos de um curso de ciências da educação, como ferramenta para auxiliar no autoestudo.

O trabalho traz como problemática principal a necessidade de mentorias individuais para os estudantes, pois a demanda é maior que os recursos disponíveis. Segundo Neumann *et al.* (2021), “os estudantes gostariam de mais suporte para lidar com a quantidade de trabalhos de leitura, de modo a entendê-los e se prepararem para provas.”

Nesse contexto, os autores desenvolveram os seguintes *Chatbots*: **FeedBot** (Figura 7), que auxilia o autoestudo a partir de tarefas de escrita sobre os trabalhos de leitura, e o **LitBot** (Figura 8), cuja abordagem visa ativar o conhecimento prévio dos alunos antes da leitura e incentiva a reflexão após a tarefa. Como bases técnicas, Neumann *et al.* (2021) utilizaram o *Social Bot Framework* (SBF), RASA, T-MITOCAR e grafos de conhecimento para a construção dos *Chatbots*.

Figura 7 – Funcionamento do FeedBot.



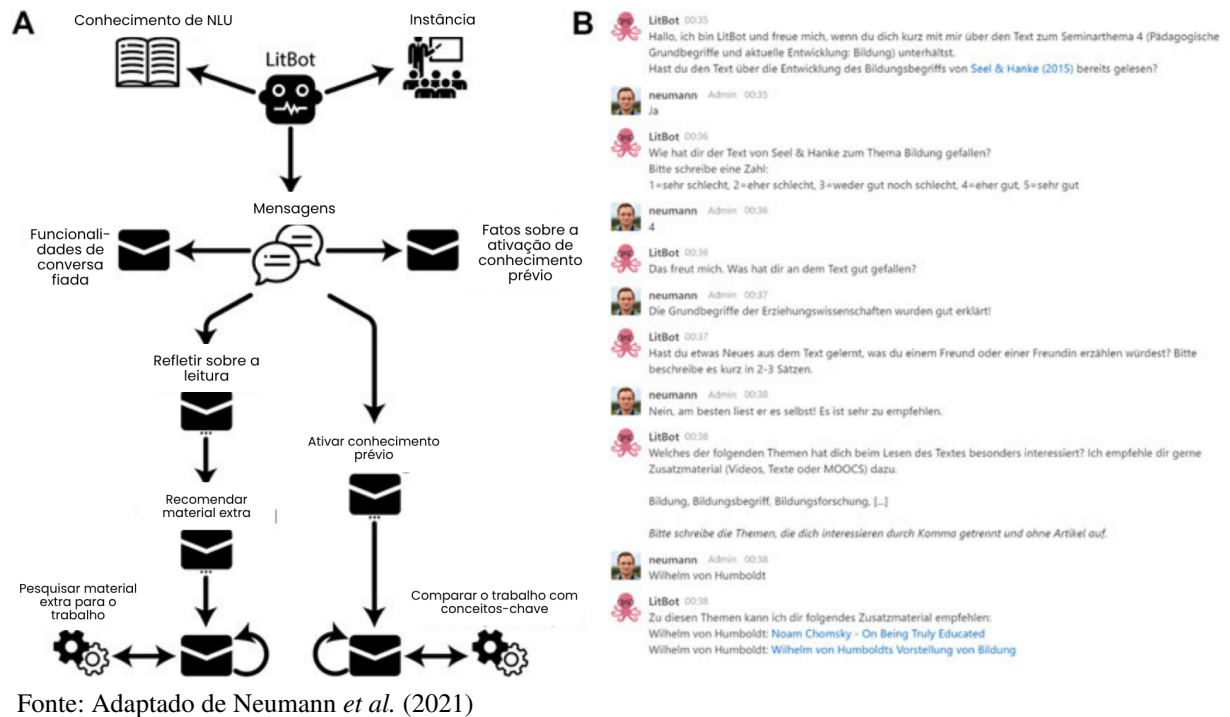
Fonte: Adaptado de Neumann *et al.* (2021)

A avaliação dos *Chatbots* se deu em 3 etapas. Primeiramente, os autores avaliaram a aplicabilidade do SBF para a criação de *bots* por meio do questionário SUS. Em seguida, o FeedBot foi avaliado utilizando uma pesquisa qualitativa com 6 alunos e, por fim, a avaliação do LitBot foi feita com uma pesquisa qualitativa com 13 alunos após uma sessão de seminários.

Após a aplicação do questionário, Neumann *et al.* (2021) concluiu que a plataforma SBF era adequada para a integração de participantes não técnicos no desenvolvimento de bots. Segundo os autores, “a maioria dos participantes demonstraram interesse real no framework, expressando que agora eles seriam capazes de modelar um bot sozinhos ou com pouca ajuda.”

Durante a avaliação dos *Chatbots*, os usuários demonstraram reações positivas, afirmando que os sistemas os auxiliaram no processo de autoestudo. Entretanto, foi expressado o desejo de poder falar com uma pessoa real para oferecer suporte. Também, os usuários demonstraram a necessidade de respostas personalizadas e o desejo por mais exercícios.

Figura 8 – Funcionamento do LitBot.



Apesar de este trabalho não apresentar uma solução nova, o trabalho de Neumann *et al.* (2021) se relaciona com este TCC por avaliar uma ferramenta que auxilia os estudantes no seu processo de estudo. O trabalho foca em métricas qualitativas sobre os sistemas, utilizando questionários e pesquisas com usuários, enquanto este trabalho busca obter métricas tanto quantitativas quanto qualitativas por meio de diferentes métodos de avaliação.

3.5 Quadro comparativo

Para compreender o estado da arte e situar este trabalho no cenário atual de pesquisas sobre a avaliação de *Chatbots* utilizando diferentes métodos, foi realizada uma análise de quatro trabalhos relacionados. A seguir, apresentamos um quadro comparativo que sintetiza os principais pontos de convergência e divergência entre as abordagens, objetivos e metodologias empregadas pelos autores.

Cada um dos trabalhos analisados contribui de forma distinta para a compreensão dos critérios de avaliação de interfaces conversacionais, com foco em *Chatbots*. Ao mesmo tempo em que alguns apresentam novas soluções em determinados contextos, outros buscam compreender a percepção dos usuários sobre esses sistemas por meio da avaliação. Este TCC se diferencia por aplicar múltiplos métodos avaliativos — incluindo MALTU, MIS, MAC e Teste de Usabilidade — sobre um sistema existente voltado à matemática, realizando uma análise de

Quadro 1 – Comparação entre os trabalhos relacionados

	(Damasceno <i>et al.</i> , 2020)	(Miranda <i>et al.</i> , 2024)	(Batista <i>et al.</i> , 2022)	(Neumann <i>et al.</i> , 2021)	Este trabalho
Contexto	Educacional para cursos	Educacional para programação	Governamental	Educacional	Educacional para matemática
Desenvolvimento de Chatbot	Sim	Sim	Não	Sim	Não
Métodos de avaliação	Observação + Teste de usabilidade	Teste de usabilidade	Avaliação Heurística	Questionários	MALTU, MIS, MAC
Tipo de análise	Quantitativa e Qualitativa	Quantitativa	Qualitativa	Qualitativa	Quantitativa e Qualitativa

Fonte: Elaborado pela autora.

diferentes aspectos do sistema avaliado.

4 METODOLOGIA

Este capítulo descreve os procedimentos metodológicos aplicados nesta avaliação sobre o *Chatbot Mathos AI*. A escolha do aplicativo se deu graças à emergência do uso de *Chatbots* dentro do contexto educacional, principalmente na forma de tutores e para a pesquisa. Trabalhos como Damasceno *et al.* (2020) e Miranda *et al.* (2024) possuem uma abordagem similar, empregando diferentes métodos para a avaliação de *Chatbots* educacionais, entretanto, as pesquisas são comumente centradas em aspectos únicos da interface conversacional.

Este trabalho busca a aplicação de diferentes métodos de avaliação, com a intenção de obter uma visão geral sobre o estado do aplicativo, mas também apresentar evidências sobre como métodos de Avaliação de IHC podem ser aplicados dentro do contexto de *Chatbots*. Na Seção 4.1, serão detalhados os objetivos a serem alcançados com esta pesquisa, além de definir o público-alvo buscado para esta avaliação. Depois, nas Seções 4.2, 4.3, 4.4, serão detalhados o processo de aplicação dos métodos selecionados para esta avaliação.

4.1 Definição dos objetivos

O objetivo principal desta pesquisa é obter uma visão geral do estado do aplicativo selecionado, investigando como os usuários percebem e interagem com o sistema. Para esta pesquisa, foram estabelecidos os seguintes objetivos:

- Aplicar o MALTU com o intuito de analisar as formas pelas quais os usuários descrevem e interpretam o sistema, identificando PRUs e quais os principais pontos a serem analisados nas próximas avaliações.
- Realizar o MIS com o propósito de reconstruir a metamensagem projetada pelo designer, a partir da interpretação dos signos presentes na interface.
- Aplicar o MAC para examinar a eficácia da emissão da metamensagem, com base na forma como os usuários interpretam e compreendem os elementos da interface.
- Realizar um teste de usabilidade, para identificar aspectos que possam comprometer a qualidade do uso a partir da observação dos usuários.
- Integrar os resultados obtidos nas diferentes avaliações, buscando convergências, complementaridades e padrões recorrentes, com o objetivo de construir uma visão abrangente e crítica do desempenho comunicacional do sistema.

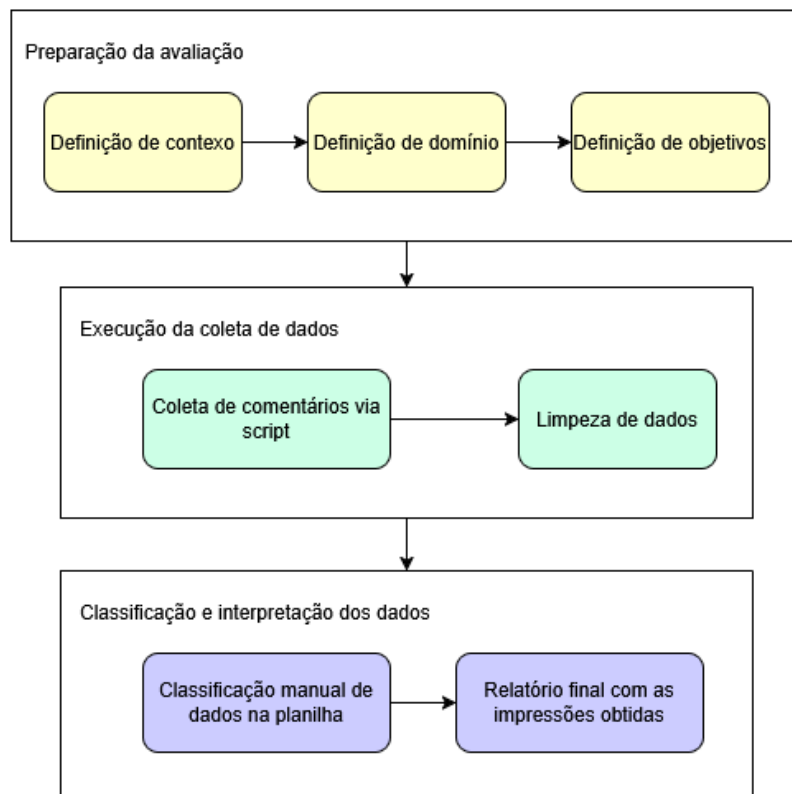
Para definir o escopo da avaliação, foi definido um público-alvo para os estudos serem

realizados. Observou-se que o usuário deve ser alguém que possui a necessidade de ferramentas que auxiliem com o estudo de Matemática, oferecendo resultados rápidos e confiáveis. Além disso, também foi observado que o usuário deve ter um nível de escolaridade adequado, para que possa ler e interpretar a interface e as respostas dadas pelo *Chatbot*. Definiu-se, portanto, como público-alvo alunos universitários de diferentes cursos que estão cursando as disciplinas de Cálculo *I* e Cálculo *III*. A partir dessa definição, foi possível desenvolver os cenários de execução do MIS e o recrutamento de usuários para a realização do MAC

4.2 Aplicação do MALTU

A próxima etapa da metodologia foi a aplicação do MALTU, proposto por Mendes (2015), cuja metodologia possui 5 etapas básicas. O fluxo da avaliação pode ser ilustrado na Figura 9.

Figura 9 – Execução do MALTU



Fonte: Elaborado pela autora.

A primeira etapa da aplicação do método consiste em definir o contexto de avaliação, que consiste no contexto de uso do sistema, o domínio e os objetivos de avaliação. O **contexto de uso** do sistema escolhido é o contexto educacional, sendo empregado por estudantes que buscam

apoio no aprendizado de matemática por meio de interações com um agente conversacional. O **domínio do sistema** são resoluções matemáticas, abrangendo diferentes conteúdos como álgebra, geometria e aritmética, compreendendo dúvidas dos estudantes por meio de PLN e oferecendo orientações pedagógicas de acordo com a necessidade do usuário. Os **objetivos** da avaliação serão a identificação de problemas na interação e interface, avaliar a satisfação dos usuários com o uso, identificar as sugestões e como os usuários comentam sobre o sistema.

Após a definição do contexto de avaliação, foi realizada a extração de PRUs, que pode ser realizada de forma manual ou automática (Mendes, 2015). Neste trabalho, a coleta das PRUs foi realizada de forma automática a partir da plataforma Google Play Store¹, utilizando um script Python e a biblioteca Google Play Scapper², que pode ser consultado no Apêndice A.

Depois de realizar a extração das PRUs, foi realizada a classificação dos comentários recolhidos. Segundo Mendes (2015), as categorias sugeridas são:

- Tipo
- Intenção
- Análise de Sentimentos
- Funcionalidade Relacionada
- Critérios de qualidade de uso
- Artefato

Durante esta fase, foi realizada a classificação manual de 49 comentários, organizados dentro de uma planilha.

Por fim, foi conduzida a interpretação dos resultados, considerando tanto a frequência dos dados quanto as relações evidenciadas entre eles. A análise inicial pode ser vista de forma mais profunda na seção 5.1.

4.3 Aplicação do MIS

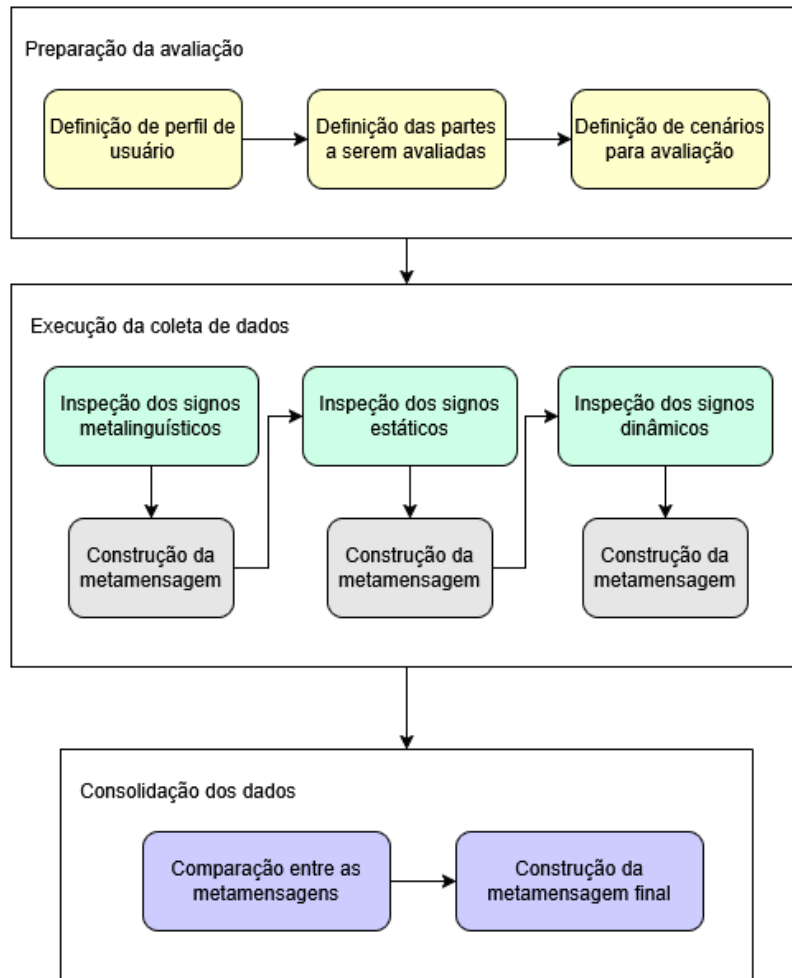
O método de inspeção semiótica é um método voltado para a análise da emissão da metamensagem do designer, ou seja, um método para “*avaliar a qualidade da emissão da metacomunicação do designer codificada na interface*” (Barbosa *et al.*, 2021). O fluxo da avaliação pode ser ilustrado na Figura 10.

A primeira etapa do método consiste na preparação da avaliação, onde foi definido o

¹ Google Play Store

² Google Play Scrapper

Figura 10 – Execução do MIS



Fonte: Elaborado pela autora.

público-alvo e as partes da interface a serem analisadas. O perfil dos usuários desta avaliação foi descrito no início deste capítulo e as partes do sistema a serem analisadas serão definidas à seguir. Com base nessas informações, foi elaborado um cenário-base com o objetivo de contextualizar o avaliador em uma situação de uso próxima da realidade. Um cenário é uma narrativa fictícia que possui detalhes sobre uma situação de interação, envolvendo usuários, processos e dados reais ou potenciais (Barbosa *et al.*, 2021). O cenário-base para a avaliação está definido a seguir:

“Um estudante está se preparando para uma avaliação de Cálculo e necessita de apoio especializado para auxiliar no seu aprendizado. Em virtude da escassez de professores disponíveis para atender às suas necessidades específicas, o estudante decide explorar abordagens alternativas. Durante suas buscas, toma conhecimento da existência de agentes conversacionais aplicados ao ensino e, em particular, do Mathos AI — um Chatbot voltado para o apoio no aprendizado de matemática, e decide utilizá-lo dentro dos seus estudos. Portanto, o estudante irá utilizar a ferramenta para alcançar 3 objetivos principais:

- *Obter auxílio na resolução de questões complexas;*
- *Compreender conceitos por meio de explicações e exemplos;*
- *Contar com apoio na execução de atividades escolares.”*

Após a preparação da avaliação, prosseguiu-se para a preparação e coleta de dados. Com base no cenário inicial, foram obtidos 3 novos cenários complementares e específicos a cada funcionalidade, definidos no Quadro 2.

Quadro 2 – Quadro de cenários para a avaliação

1	Você está com dificuldade em um conteúdo de cálculo, e decidiu testar uma nova ferramenta para lhe auxiliar com a resolução de questões. Para iniciar, você pediu para que o <i>Chatbot</i> lhe explique o conteúdo em termos simples.
2	Após receber uma resposta agradável, você decidiu incluir o aplicativo como ferramenta de estudo, e agora você irá utilizá-lo para lhe praticar problemas que você tem dificuldade de entender
3	Por fim, depois de muito explorar, você notou que você esqueceu de fazer uma lista de exercícios que seu professor enviou e que deve ser entregue amanhã. Como o tempo é curto e as questões são extensas, você decidiu pedir ajuda ao <i>Chatbot</i> para a resolução dos problemas.

Fonte: Elaborado pela autora.

Após a definição dos cenários, o avaliador iniciou a análise dos signos por tarefa. Aqui, foi feita uma coleta dos signos metalinguísticos, estáticos e dinâmicos da aplicação, e, para cada um deles, foi feita uma interpretação para identificar a intenção do designer dentro da interface. À medida que cada tipo de signo foi interpretado, a metamensagem do designer foi reconstruída para explicitar a intenção por trás da interface. Para a reconstrução da metamensagem, o modelo de Souza (2005) será seguido:

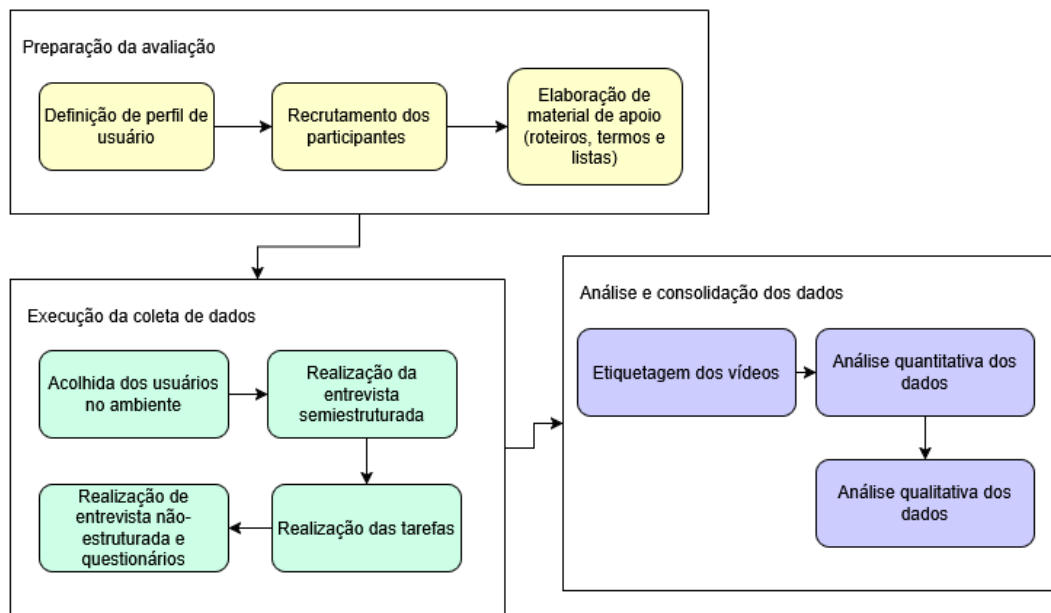
"Este é o meu entendimento, como designer, de quem você, usuário, é, do que aprendi que você quer ou precisa fazer, de que maneiras prefere fazer, e por quê. Este, portanto, é o sistema que projetei para você, e esta é a forma como você pode ou deve utilizá-lo para alcançar uma gama de objetivos que se encaixam nesta visão."

Por fim, ao finalizar as três metamensagens, foi feita a consolidação dos resultados, onde cada mensagem obtida foi comparada com o objetivo de encontrar redundâncias, ambiguidades, inconsistências ou possibilidades de ruptura. Os resultados desta avaliação podem ser vistos na Seção 5.2.

4.4 Aplicação do MAC

O Método de Avaliação de Comunicabilidade tem como foco principal a análise da recepção da metamensagem pelo usuário, ao contrário do MIS, que se concentra na produção dessa mensagem pelo designer. Esse método busca identificar rupturas na comunicação entre usuário e sistema, revelando obstáculos à compreensão e à interação eficaz (Souza, 2005). O fluxo da avaliação pode ser ilustrado na Figura 11.

Figura 11 – Execução do MAC



Fonte: Elaborado pela autora.

Para a preparação do MAC, foram recrutados diferentes perfis de estudantes universitários, familiares com tecnologia e diferentes experiências relacionadas à IA e ao aprendizado de matemática. Após isso, foi elaborado um roteiro para a realização da avaliação, que pode ser acessado no Apêndice E, especificando os passos a serem seguidos e como deverão ser realizados. Além disso, também foi elaborado um termo de consentimento que esclarece o objetivo e as etapas do teste, garantindo a transparência do processo e a privacidade dos usuários, que pode ser visto no Apêndice B.

Logo em seguida, foi realizada a coleta de dados, onde, em ambiente controlado, os usuários foram instruídos a realizarem as tarefas de acordo com os cenários descritos (Barbosa *et al.*, 2021). As tarefas realizadas podem ser vistas no Quadro 3, e a lista de exercícios utilizada como apoio para o teste pode ser vista no Apêndice C. Nesta etapa, foi feita uma entrevista semi-estruturada prévia para a coleta de dados sobre o perfil dos usuários, cujo roteiro pode

ser visto no Apêndice D. Para o registro das tarefas, tanto a tela como a imagem dos usuários foi gravada. Após a realização das tarefas, cada participante passou por uma entrevista não estruturada e, posteriormente, por um questionário, que pode ser visto no Apêndice F, para a coleta de possíveis informações extras.

Quadro 3 – Quadro de tarefas para a avaliação

Tarefa 1	Pedir para o <i>Chatbot</i> explicar um conteúdo que você tem dificuldade
Tarefa 2	Pedir para o <i>Chatbot</i> resolver um exercício
Tarefa 3	Pedir para o <i>Chatbot</i> corrigir um erro de cálculo ou conceitual
Tarefa 4	Exercitar o conteúdo utilizando diferentes funcionalidades
Tarefa 5	Exercitar o conteúdo utilizando a funcionalidade de quiz
Tarefa 6	Resolver questões utilizando a funcionalidade de leitura de PDF

Fonte: Elaborado pela autora.

Para a etapa de interpretação, os vídeos gravados serão analisados e etiquetados de acordo com Souza (2005), para a identificação de rupturas de comunicabilidade. Durante a consolidação dos resultados, as etiquetas serão interpretadas de acordo com os fatores listados (Souza, 2005):

- Frequência de cada etiqueta
- Sequências e repetições
- Problemas identificados durante o processo de interpretação
- Outros problemas que não necessariamente dizem respeito à comunicabilidade do sistema.

Por fim, na etapa de relato, será construído o perfil semiótico do sistema avaliado, elaborado com base na metamensagem do designer, construída de acordo com a maneira como ela foi interpretada pelo usuário durante as avaliações (Souza, 2005). Os resultados desta avaliação podem ser vistos na Seção 5.3.

4.5 Consolidação dos dados

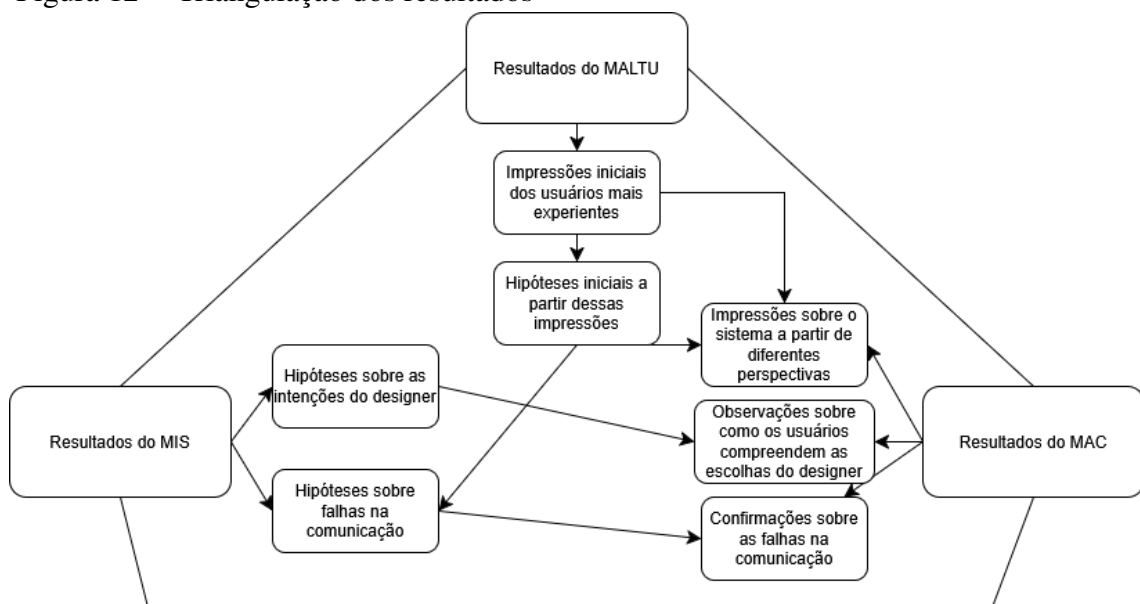
Para assegurar a validade deste trabalho, além de obter resultados confiáveis e de qualidade, será realizada uma triangulação dos dados, visando compreender o sistema em diferentes níveis. Este método foi escolhido por ser uma estratégia de validação da pesquisa, pois permite relacionar dados qualitativos distintos em uma única síntese, sendo convergentes, complementares ou divergentes (Denzin, 2017; Oelerich; Otto, 2011).

A escolha da triangulação de dados se deu pela decisão de utilizar diferentes métodos para a avaliação do *Chatbot*, que produzem resultados distintos. Com o MALTU, foi possível identificar as percepções dos usuários a partir dos comentários encontrados na loja de aplicativos.

Por meio da aplicação do MIS e do MAC, será investigada a comunicabilidade do sistema, além de possíveis aspectos sobre a usabilidade e a experiência geral do usuário ao utilizar o sistema.

Para realizar a integração dos dados, será realizada uma análise cruzada dos resultados, identificando pontos convergentes, divergentes e complementares entre cada método, como recomenda Gomes *et al.* (2005), como pode ser exemplificado na Figura 12. A partir disso, será possível fortalecer a validade da pesquisa, ampliar a compreensão sobre o desempenho do *Chatbot* e fornecer recomendações mais consistentes e fundamentadas tanto sobre o sistema, como sobre a metodologia adotada.

Figura 12 – Triangulação dos resultados



Fonte: Elaborado pela autora.

5 RESULTADOS

5.1 Resultados do MALTU

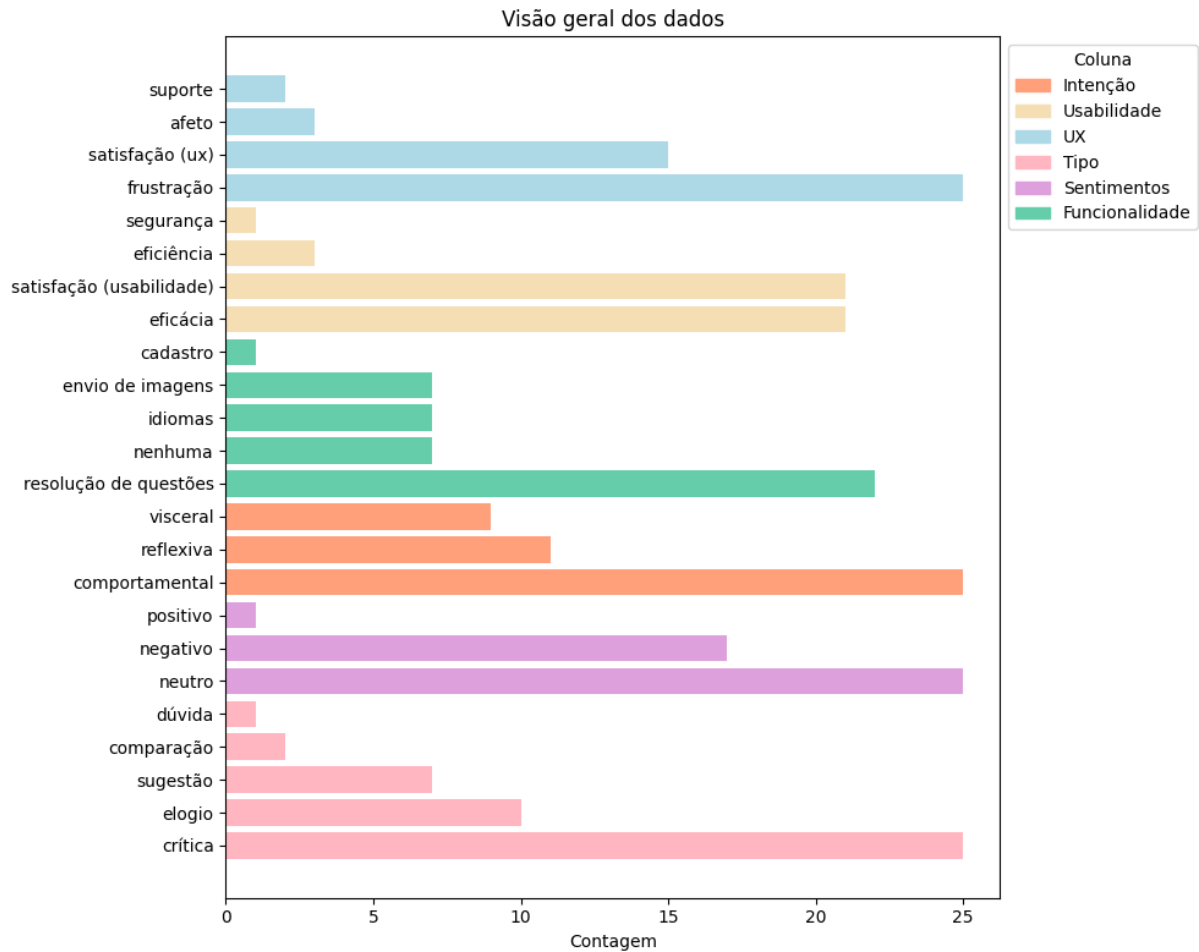
Durante a aplicação da metodologia descrita na seção 4.2, foi realizada a extração dos comentários via *web scrapping*, sendo considerados comentários em português, inglês e espanhol. Esta extração retornou aproximadamente 196 comentários, entretanto, nem todos poderiam ser considerados como PRUs, por consistirem de apenas *emojis* ou não conterem conteúdo relevante para a pesquisa dentro do texto (i.e. um comentário falando "pior app de todos"), o que trouxe a necessidade da realização de uma filtragem nos dados obtidos, resultando em 45 comentários a serem analisados. Na Figura 13, é possível ver a distribuição geral dos resultados obtidos, classificados em Tipo, Intenção, Sentimentos, Funcionalidade, Usabilidade e UX Mendes (2015).

O gráfico evidencia algumas tendências encontradas durante a classificação dos comentários, pois mostra que, apesar de um número alto de críticas, os comentários foram em sua maioria neutros e de intenção comportamental. Tais comentários, por muitas vezes, se mostraram relacionados à problemas técnicos, como problemas para conseguir atualizar o sistema.

As avaliações do aplicativo indicam um alto nível de aprovação por parte dos usuários, refletido em uma média de 4,6 estrelas na Play Store, com cerca de 84% dos comentários sendo notas 5. No entanto, embora em menor quantidade, também foram identificados comentários negativos que apontam insatisfações específicas com o serviço.

Ao analisar os comentários, foi notada uma certa frequência nas menções sobre problemas de atualizações, ocorrendo em diferentes períodos do ano de 2024. Os comentários relacionados a este problema técnico em específico possuem diferentes pontos de vista, como por exemplo nas seguintes PRUs: **PRU 1:** "*[o aplicativo] continua me pedindo para atualizar e me redireciona para a Google Play, então eu clico para abrir e novamente vejo a tela de atualização. Estou em um ciclo triste.*", que demonstra frustração e uma baixa satisfação com o sistema, além de possuir uma polaridade negativa; **PRU 2:** "*Ele me diz para atualizar o app, mas já está atualizado e não me deixa [utilizar/atualizar].*", que também mostra frustração, porém em um tom neutro e com menos detalhes sobre o uso. **PRU 3:** "*Acabei de instalá-lo e ele me diz que há uma nova atualização e não permite que eu faça login e use o aplicativo. Spoiler: não há atualização.*"

Figura 13 – Visão geral dos resultados do MALTU.



Fonte: Elaborado pela autora.

Nota: Neste gráfico, é possível ver a contagem de PRUs, classificadas a partir de diferentes parâmetros, descritos pelas cores utilizadas para cada barra.

Nota: As linhas que vão de suporte à frustração são classificações relacionadas à UX, coloridas de azul; as linhas que vão de segurança a eficácia estão relacionadas à Usabilidade, coloridas de amarelo; as linhas que vão de cadastro a resolução de questões estão relacionadas a funcionalidades, coloridas de verde; as linhas que vão de visceral a comportamental estão relacionadas à Intenção, coloridas de laranja; as linhas de positivo a neutro são os Sentimentos identificados, coloridas de roxo; as linhas de dúvida a crítica estão relacionadas ao Tipo, coloridas de rosa.

Também, diversos usuários expressaram insatisfação quanto ao modelo de assinatura adotado pela plataforma. As críticas mais recorrentes dizem respeito à cobrança automática sem aviso prévio, à dificuldade em cancelar o plano e à ausência de suporte eficiente, que podem ser evidenciadas a seguir: **PRU 4:** “*Esse aplicativo é bom, mas o problema é que ele me obriga a comprar a assinatura premium. É por isso que vou excluir este aplicativo.*”; **PRU 5:** “*Decepcionante, não resolve bem as coisas e, além disso, eles me cobraram \$629 pela assinatura sem me avisar antes e não tenho mais o aplicativo baixado e nenhuma possibilidade de reembolso.*”; **PRU 6:** “*Inteligência artificial muito boa, resolve problemas com números complexos e aplicações de engenharia. O que eu não gostei é que eu só queria pagar por um mês, e no mês seguinte eles me cobraram automaticamente, e eu tentei contatá-los pelos meios*

que eles dizem, e é falso que eles têm suporte 24/7. Até agora eles nunca me responderam.”

Além disso, vários comentários falaram sobre a eficácia do aplicativo, possuindo tanto elogios sobre as respostas corretas (como pode ser visto na **PRU 6**), como críticas. A seguir, serão mostrados diferentes comentários sobre a eficácia do *Chatbot*: **PRU 7**: *“O pior aplicativo de matemática de todos os tempos, ele mostra a solução errada com confiança e, quando peço para corrigi-la, ele diz que é a única solução possível. Além disso, comete erros até mesmo em questões básicas de geometria [...]”*; **PRU 8**: *“Este aplicativo não vale nada: tive um problema e queria que me dissesse os valores de a para que seja verdadeiro que $a^x = x$. É claro que não é um problema fácil (por isso mesmo foi colocado) e que tem como solução um intervalo definido de valores de a . Eu disse que só existe um valor de a para que se cumpra essa condição. Depois disso, não me atrevi a dizer que ele estava equivocado. O que eu disse: você fez errado e depois não me deixou corrigir.”*; **PRU 9**: *“Eu sou um estudante francês e acho que este é um aplicativo bem útil. Pode responder em outras línguas, no meu caso em francês, porém não tem representações gráficas, então darei 4 estrelas.”*; **PRU 10**: *“Não é tão confiável. Preciso que me dê resultados que nem uma calculadora. Funcionou bem em 90% das vezes que o usei.”*

Também, foram encontrados alguns comentários relacionados à sugestão, sendo mais comum a adição de suporte à diferentes línguas. Porém, além disso, houveram sugestões sobre funcionalidades a serem adicionadas ou melhoradas, como podem ser vistos nas seguintes PRUs: **PRU 11**: *“10/10 mas eu gostaria que ele permitisse que você enviasse várias imagens para resolver vários exercícios, por isso eu dei um 4, mas o resto é super bom.”*; **PRU 12**: *“É bom, mas se tivesse ilustrações seria melhor, por exemplo, diagramas de fenômenos matemáticos como círculos, contos, objetos tridimensionais, etc.”*; **PRU 13**: *“A aplicação só tem o idioma inglês e só resolve alguns problemas, não recomendo mas cada um faz o que quiser.”*

A análise dos comentários sobre o aplicativo *Mathos AI* mostrou certas divergências de opiniões. Apesar de muitos comentários positivos, foram percebidas diversas críticas e comentários negativos. Entretanto, é possível perceber que os desenvolvedores percebem e valorizam os comentários dos usuários, pois, após sugestões sobre a adição de novas línguas (PRU 13), o aplicativo adicionou suporte a diferentes línguas, como português, espanhol, francês (PRU 9), italiano, entre outros.

5.2 Resultados do MIS

Nesta seção, serão discutidos os resultados obtidos com a realização do Método de Inspeção Semmiótica (MIS) sobre a interface do *Chatbot Mathos AI*. O método foi aplicado por um avaliador especialista, que executou os cenários descritos na seção 4.3.

A seguir, os signos serão analisados e interpretados e, em seguida, será apresentada a metamensagem obtida a partir de cada análise. Na subseção 5.2.1, serão analisados os signos metalinguísticos, seguida pela subseção 5.2.2 onde serão discutidos os signos estáticos e, por fim, na subseção 5.2.3, serão abordados os signos dinâmicos, conforme orienta Souza (2005).

5.2.1 Signos Metalinguísticos

Para iniciar a aplicação do método, foi realizada a inspeção e análise dos signos metalinguísticos, como proposto por Souza (2005). Em sua maioria, os signos metalinguísticos foram apresentados dentro da interface como descrições, instruções e *tooltips* para que o usuário possa compreender melhor como utilizar a interface.

Cenário 1: “Você está com dificuldade em um conteúdo de cálculo, e decidiu testar uma nova ferramenta para lhe auxiliar com a resolução de questões. Para iniciar, você pediu para que o *Chatbot* lhe explique o conteúdo em termos simples.”

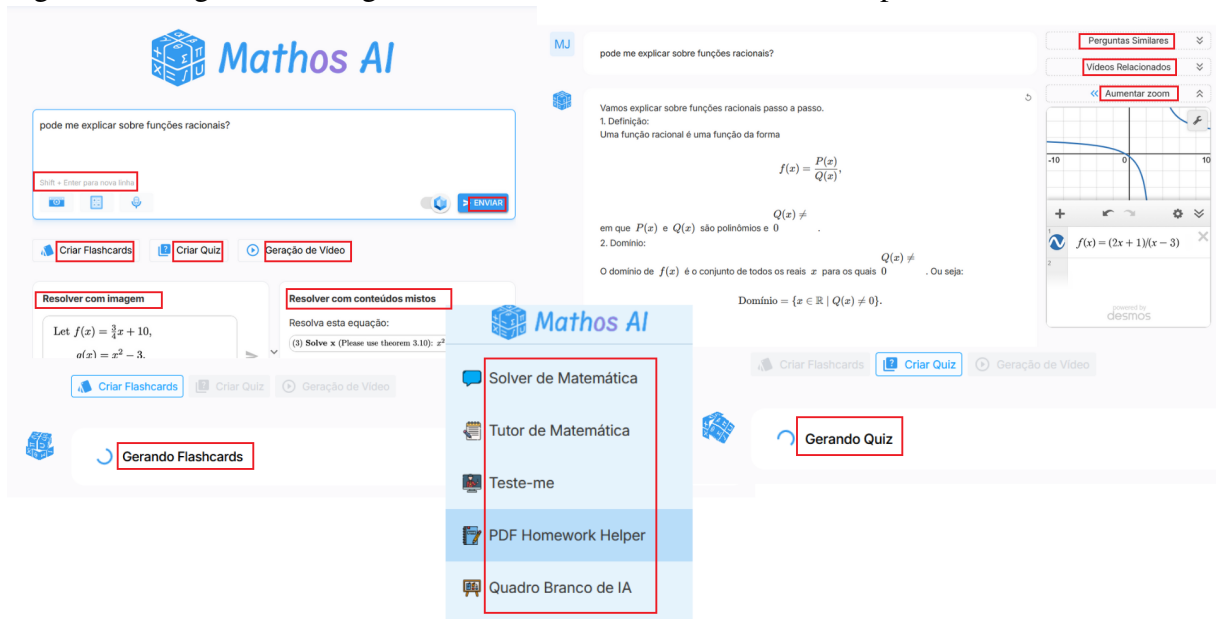
Neste cenário, o usuário deve enviar uma mensagem para o *Chatbot* pedindo para que este explique os conceitos básicos do conteúdo de forma simples de se entender.

Os signos metalinguísticos encontrados foram:

- Instruções dentro da caixa de mensagens;
- Rótulos de botões;
- Tooltips;
- Mensagens de erro;
- Descrições sobre a aplicação;

Na Figura 14 é possível ver alguns dos signos identificados durante o processo de análise. Aqui, é possível notar que o designer da interface buscou torná-la bastante descritiva, para que facilite o uso e a navegação por usuários iniciantes, entretanto, determinados signos não comunicam claramente o que o usuário deve fazer. Além disso, a resposta do *Chatbot* é curta e simples, o que torna a interação menos parecida com uma conversa e mais como uma resolução de tarefas.

Figura 14 – Signos Metalinguísticos encontrados na interface correspondente ao cenário 1.



Fonte: Elaborado pela autora.

Cenário 2: “Após receber uma resposta agradável, você decidiu incluir o aplicativo como ferramenta de estudo, e agora você irá utilizá-lo para lhe praticar problemas que você tem dificuldade de entender.”

Neste cenário, o usuário irá utilizar a funcionalidade de quiz do Chatbot para praticar a resolução de problemas mais complexos. Nessa atividade, o usuário envia o enunciado de um problema, e o sistema o decompõe em etapas, promovendo uma resolução mais guiada e estruturada.

Os signos metalinguísticos encontrados foram:

- Descrição da funcionalidade “Teste-me”;
- Indicadores de estado e progresso;

Na Figura 15 é possível ver alguns dos signos identificados durante o processo de análise. Durante a execução desta tarefa, foi encontrada uma falha de comunicação, pois a interface traz a ideia de que o *Chatbot* irá criar uma lista de questões práticas baseado em um conteúdo. No entanto, ao realizar essa ação, o sistema indica que essa funcionalidade não está disponível da forma inicialmente sugerida.

Cenário 3: “Por fim, depois de muito explorar, você notou que você esqueceu de fazer uma lista de exercícios que seu professor enviou e que deve ser entregue amanhã. Como o tempo é curto e as questões são extensas, você decidiu pedir ajuda ao *Chatbot* para a resolução dos problemas.”

Neste cenário, o usuário deve realizar o upload de uma lista de exercícios para

Figura 15 – Signos Metalinguísticos encontrados na interface correspondente ao cenário 2.

O que é Tutor de Matemática?

1. Extraia insights chave

2. Divida em etapas

3. Domine o problema

Pergunta
determine o domínio, raízes, assíntotas verticais e estudo do sinal da função
 $f(x) = \frac{-x+9}{x+8}$

Progresso
Entendendo sua pergunta
Reunindo insights chave
Vamos começar!
Tudo está pronto! Vamos mergulhar juntos e conquistar este desafio.

Domínio de uma função
Conjunto de todos os valores possíveis de x que podem ser usados na função.
Raízes de uma função
Valores de x para os quais $f(x) = 0$.
Assíntotas verticais
 $x =$
Linhas verticais $x =$ onde a função tende ao infinito ou menos infinito.
Estudo do sinal
Análise dos intervalos onde a função é positiva ou negativa.

Etapa 1
Considere a função $f(x) = \frac{-x+9}{x+8}$. Qual das alternativas abaixo representa corretamente o domínio desta função?
O domínio é \mathbb{R} .
O domínio é $\mathbb{R} \setminus \{8\}$.
O denominador 8 não pode ser zero, ou seja, 8 .
O domínio é $\mathbb{R} \setminus \{9\}$.

Considere a função $f(x) = \frac{-x+9}{x+8}$. Qual procedimento é usado para determinar a raiz (zero da função)?
Dividimos o numerador pelo denominador e procuramos o valor que torna a fração igual a 1.
Igualamos o numerador a zero: $0 = -x + 9$, resultando em 9 .
Igualamos o denominador a zero: $0 = x + 8$, o que implicaria 8 .
Igualamos tanto o numerador quanto o denominador a zero simultaneamente.

Próximo

Incorreto

Atualmente uma experiência de pré-visualização, tutor completo disponível mediante assinatura.

Fonte: Elaborado pela autora.

realizar a resolução destes. O sistema provê o usuário com um tutorial de como utilizar a funcionalidade, e os signos identificados demonstram diferentes instruções e indicadores de estado. Os signos identificados podem ser vistos na Figura 16.

Figura 16 – Signos Metalinguísticos encontrados na interface correspondente ao cenário 3.

Mathos AI

Solte ou clique para enviar seu PDF aqui

Oii! Como posso ajudar você hoje?

Selecione uma região para começar

Efetue as seguintes divisões de polinômio a seguir,

1. $\frac{x^2 - 15x + 56}{x - 7}$

2. $\frac{-x + 2}{-x + 2}$

Resolver

Exercícios

1. ainda que o grau do polinômio numere o do polinômio denominador...
2. $\frac{x^4 - 2x^3 - 7x^2 - 8x - 65}{x^2 - 5x - 5}$
3. $\frac{x^4 - 4x^2 + 12x - 9}{x^2 - 2x + 3}$
4. $\frac{x^4 - 4x^3 - 7x^2 + 23x + 14}{x^2 - 13x^2 + 3x + 2}$
5. Determine os valores de a , nos casos a seguir, para que a divisão polinomial seja exata.
6. $\frac{x^3 + (2a^2 - 7)x - (13a^2 + 1)}{x - 7}$
7. $\frac{x^2 + (3 - a)x + (a^2 - 10a)}{x - 4}$
8. Para as funções racionais a seguir determine o domínio, raízes, assíntotas verticais, buracos, estudo de sinal da imagem e esboço do gráfico.
9. $f(x) = \frac{1}{x}$
10. $f(x) = \frac{4}{x^2}$
11. $-x + 9$

Fonte: Elaborado pela autora.

5.2.1.1 *Reconstrução da Metamensagem*

Você é alguém que valoriza instruções claras na tela, de modo que possa identificar facilmente os campos e elementos da interface. Aprendi que o que você mais precisa é resolver problemas no domínio da matemática, de forma rápida e simples e com ferramentas intuitivas e de fácil acesso, pois seu principal objetivo é encontrar soluções de maneira eficiente. Dessa forma, projetei um Chatbot baseado em IA generativa, capaz de compreender e resolver problemas matemáticos, oferecendo suporte direto durante o processo de aprendizagem. O sistema deve ser utilizado tanto com o envio de enunciados de problemas, assim como o uso de descrições e perguntas, para que dessa forma, o Chatbot possa fornecer explicações e soluções, auxiliando-o a alcançar seus objetivos de forma prática e orientada.

5.2.2 *Signos Estáticos*

Para iniciar a aplicação do método, foi realizada a inspeção e análise dos signos estáticos, como proposto por Souza (2005). Em sua maioria, os signos estáticos foram apresentados dentro da interface como signos de aplicação, ou seja, signos que o designer da interface atribuiu significado a este, como diz Souza (2005), como por exemplo os ícones relacionados às funcionalidades do *Chatbot*.

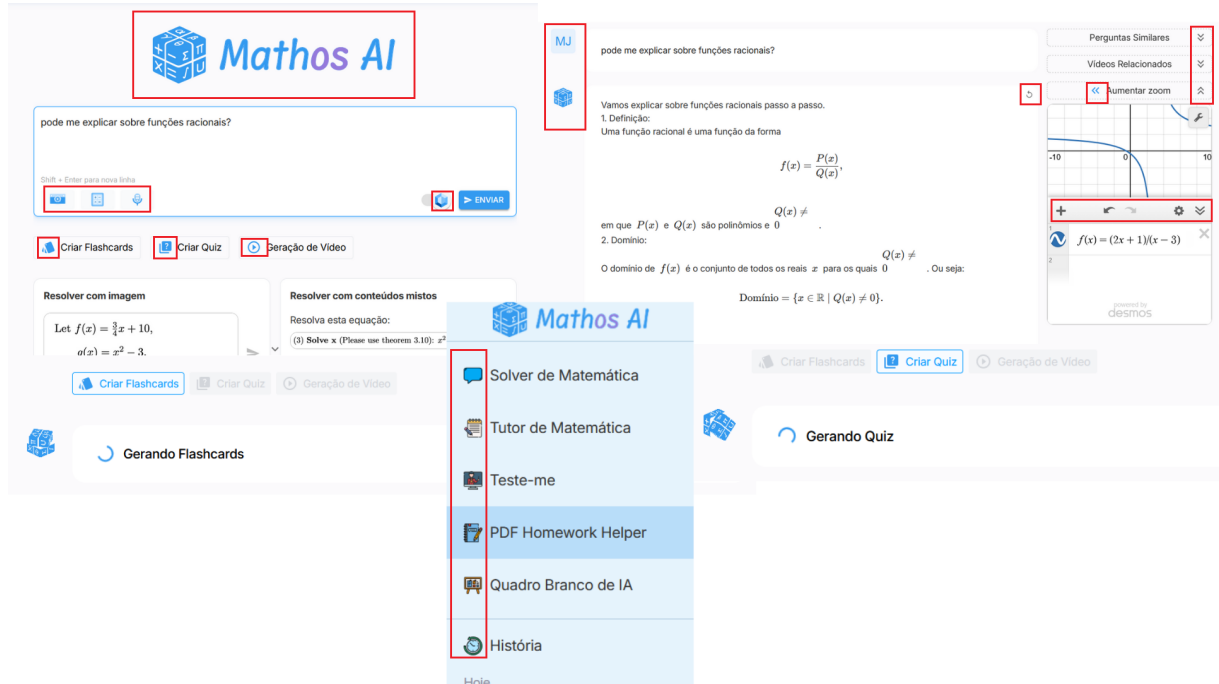
Cenário 1: “Você está com dificuldade em um conteúdo de cálculo, e decidiu testar uma nova ferramenta para lhe auxiliar com a resolução de questões. Para iniciar, você pediu para que o *Chatbot* lhe explique o conteúdo em termos simples.”

Os signos identificados foram, em sua maioria, ícones presentes na interface, como botões e indicativos de estado, como o “diamante” para indicar o modelo pro e a “moeda” para o modelo básico. Além disso, o layout do *Chatbot* após o envio de uma mensagem também pode ser interpretado como um signo sobre a intenção do designer com esta escolha, como pode ser visto na Figura 17.

Cenário 2: “Após receber uma resposta agradável, você decidiu incluir o aplicativo como ferramenta de estudo, e agora você irá utilizá-lo para lhe praticar problemas que você tem dificuldade de entender.”

Os signos identificados foram ícones que indicam funcionalidades e imagens para descrever a funcionalidade envolvida no cenário. Entretanto, apesar do uso de texto e imagens, a interface ainda não comunica de forma clara como a funcionalidade deve ser utilizada, como

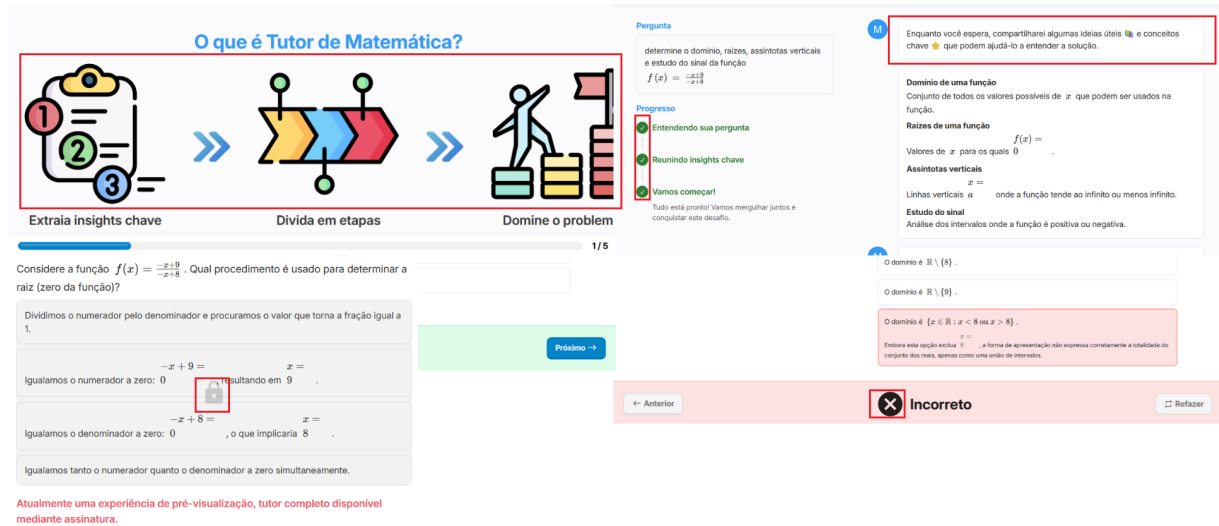
Figura 17 – Signos Estáticos encontrados na interface correspondente ao cenário 1.



Fonte: Elaborado pela autora.

pode ser ilustrado na Figura 18.

Figura 18 – Signos Estáticos encontrados na interface correspondente ao cenário 2.



Fonte: Elaborado pela autora.

Cenário 3: “Por fim, depois de muito explorar, você notou que você esqueceu de fazer uma lista de exercícios que seu professor enviou e que deve ser entregue amanhã. Como o tempo é curto e as questões são extensas, você decidiu pedir ajuda ao Chatbot para a resolução dos problemas.”

Assim como nos outros cenários, os signos estáticos identificados foram ícones que indicam outras ferramentas, além de uma imagem de um robô, para personificar o Chatbot

durante a resolução da tarefa, como pode ser visto na Figura 19.

Figura 19 – Signos Estáticos encontrados na interface correspondente ao cenário 3.

The image shows the Mathos AI interface. At the top left is the logo 'Mathos AI' with a cube icon. Below it is a large light blue box with the text 'Solte ou clique para enviar seu PDF aqui'. To the right is a chatbot character with a blue and purple body, enclosed in a red box. Below the character is the text 'Oi! Como posso ajudar você hoje?'. At the bottom left, there is a math problem: 'Efetue as seguintes divisões de polinômio a seguir, 1. $\frac{x^2 - 15x + 56}{x - 7}$ '. Below the problem are icons for chat, voice, and a 'Resolver' button. At the bottom right, there is a list of math problems, with the first one highlighted. The first problem is: '8. $\frac{x^4 - 2x^3 - 7x^2 - 8x - 65}{x^2 - 5x - 5}$ '. Below the list is a graph of a function and a 'Praticios' section with a list of problems.

Fonte: Elaborado pela autora.

5.2.2.1 Reconstrução da Metamensagem

Você é alguém que gosta de uma interface bonita e original, com as informações bem dispostas na tela, mas que também pode se aventurar e descobrir como utilizar as funcionalidades presentes no sistema. Aprendi que você precisa acessar as funcionalidades de cálculos de forma rápida e eficiente, e que para isso você precisa de botões que indiquem claramente para que servem, além de ícones únicos que identifiquem cada funcionalidade. Projetei para você um Chatbot com IA generativa, que possui uma interface amigável e descritiva, que você pode utilizá-lo enviando perguntas e problemas a serem decompostos e resolvidos com o auxílio da ferramenta.

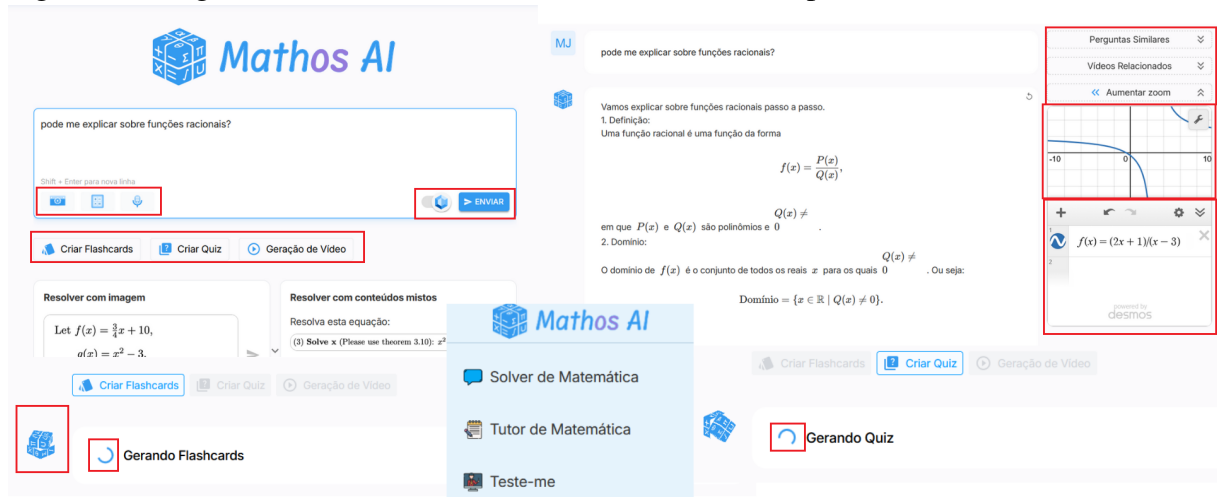
5.2.3 Signos Dinâmicos

Cenário 1: “Você está com dificuldade em um conteúdo de cálculo, e decidiu testar uma nova ferramenta para lhe auxiliar com a resolução de questões. Para iniciar, você pediu para que o Chatbot lhe explique o conteúdo em termos simples.”

Os signos dinâmicos que foram identificados foram botões que indicam outras

funcionalidades, como flashcards e quiz, e também há um switch que permite alterar o modelo generativo do *Chatbot*. Por padrão, este switch está ativado no modo “Max”, que possui um limite de uso para usuários comuns que não possuem assinatura. Além disso, a interface de resposta do *Chatbot* possui uma seção interativa para gráficos, que o usuário pode utilizar para visualizar gráficos de funções. Os signos identificados podem ser vistos na Figura 20.

Figura 20 – Signos Dinâmicos encontrados na interface correspondente ao cenário 1.



Fonte: Elaborado pela autora.

Cenário 2: “Após receber uma resposta agradável, você decidiu incluir o aplicativo como ferramenta de estudo, e agora você irá utilizá-lo para lhe praticar problemas que você tem dificuldade de entender.”

Os signos identificados foram botões e indicadores de progresso, como mostra na Figura 21. Além disso, a interface também indica dinamicamente o seu estado com a questão, se ela foi respondida correta ou incorretamente.

Cenário 3: “Por fim, depois de muito explorar, você notou que você esqueceu de fazer uma lista de exercícios que seu professor enviou e que deve ser entregue amanhã. Como o tempo é curto e as questões são extensas, você decidiu pedir ajuda ao *Chatbot* para a resolução dos problemas.”

O *Chatbot* contém um tutorial que explica como utilizar a funcionalidade de resolução de questões a partir de um PDF. Para resolver uma equação, o usuário deverá selecionar a parte que deseja ser resolvida que ficará marcada por um quadrado azul, e é possível refinar esta seleção. Após isso, o usuário deverá apertar para “enviar”, e o *Chatbot* irá resolver a questão. Além disso, também há signos dinâmicos de botões para outras funcionalidades e para personalizar a experiência nesta funcionalidade, como pode ser visto na Figura 22.

Figura 21 – Signos Dinâmicos encontrados na interface correspondente ao cenário 2.

Pergunta

determine o domínio, raízes, assíntotas verticais e estudo do sinal da função

$$f(x) = \frac{-x+9}{-x+8}$$

Progresso

- ✓ Entendendo sua pergunta
- ✓ Reunindo insights chave
- ✓ Vamos começar!

Tudo está pronto! Vamos mergulhar juntos e conquistar este desafio.

M Enquanto você espera, compartilharei algumas ideias úteis e conceitos chave que podem ajudá-lo a entender a solução.

Domínio de uma função
Conjunto de todos os valores possíveis de x que podem ser usados na função.

Raízes de uma função
 $f(x) =$
Valores de x para os quais 0 .

Assíntotas verticais
 $x =$
Linhas verticais α onde a função tende ao infinito ou menos infinito.

Estudo do sinal
Análise dos intervalos onde a função é positiva ou negativa.

M Agora estamos todos prontos! Pronto para mergulhar na solução do seu problema?

1 / 5

Pergunta

determine o domínio, raízes, assíntotas verticais e estudo do sinal da função

$$f(x) = \frac{-x+9}{-x+8}$$

Etapa 1

Considere a função $f(x) = \frac{-x+9}{-x+8}$. Qual das alternativas abaixo representa corretamente o domínio desta função?

- O domínio é \mathbb{R} .
- O domínio é $\mathbb{R} \setminus \{8\}$.
- O domínio é $\mathbb{R} \setminus \{9\}$.
- O domínio é $\{x \in \mathbb{R} : x < 8 \text{ ou } x > 8\}$.

Embora esta opção exclua 8 , a forma de apresentação não expressa corretamente a totalidade do conjunto dos reais, apenas como uma união de intervalos.

← Anterior

X Incorreto

Refazer

Fonte: Elaborado pela autora.

5.2.3.1 Reconstrução da Metamensagem

Você é alguém que gosta de uma interface interativa, que permite você acessar as funcionalidades de forma rápida e fácil, que permite que você visualize funções matemáticas e resolva questões de forma interativa. Aprendi que você precisa acessar as funcionalidades de forma rápida e fácil, que permita a interação rápida. Projetei para você um Chatbot com IA generativa, que possui uma interface interativa, que você pode utilizá-lo enviando perguntas e explorando suas funcionalidades.

Figura 22 – Signos Dinâmicos encontrados na interface correspondente ao cenário 3.

The image shows the Mathos AI interface. At the top left is the logo 'Mathos AI' with a cube icon. Below it is a large light blue box with the text 'Solte ou clique para enviar seu PDF aqui'. To the right is a cartoon robot character with the text 'Oi! Como posso ajudar você hoje?'. Below the robot is a search bar with the text 'Selecione uma região para começar' and a magnifying glass icon. The main content area is divided into two columns. The left column contains the text 'Efetue as seguintes divisões de polinômio a seguir,' followed by a math problem:
$$\frac{x^2 - 15x + 56}{x - 7}$$
 and a 'Resolver' button. The right column contains a list of exercises, including:

- $\frac{x^2 - 2x^3 - 7x^2 - 80x - 65}{x^2 - 5x - 5}$
- $\frac{x^2 - 4x^2 + 12x - 9}{x^2 - 2x + 3}$
- $\frac{x^4 - 4x^3 - 74x^2 + 23x + 14}{x^3 - 11x^2 + 3x + 2}$
- $\frac{x^2 + (2x^2 - 7)x - (13x^2 + 1)}{x - 7}$
- $\frac{x^2 + (3 - a)x + (a^2 - 10a)}{x - a}$

 Below the exercises is a graph of a function and a section titled 'exercícios' with a list of numbers.

Fonte: Elaborado pela autora.

5.2.4 Metamensagem final do designer

A partir dos signos metalinguísticos, estáticos e dinâmicos, observa-se que o sistema busca trazer uma comunicação clara, para que ele possa formular seus problemas matemáticos de maneira clara e completa para o uso do *Chatbot*. A interface traz diversos campos de texto e instruções diretas, fornecendo informações para os usuários para que a interação seja feita de forma eficaz.

Os elementos visuais e organizacionais trazem um ambiente simples e objetivo, que, em combinação com o uso de instruções e descrições, traz a ideia de que o processo de resolução será rápido e direto. Entretanto, algumas rupturas foram encontradas que comunicam ao usuário que nem todas as funcionalidades operam exatamente como sugerido, como por exemplo a falha com a funcionalidade "Teste-me".

O sistema busca facilitar a resolução de problemas matemáticos de forma acessível e imediata, entretanto, possui algumas limitações relacionados às capacidades reais do *Chatbot*. Ao mesmo tempo que é possível identificar a intenção de simplicidade na interação, resultando em uma experiência que nem sempre corresponde à visão projetada pelo designer.

A partir das metamensagens identificadas e das conclusões trazidas acima, podemos construir uma metamensagem final sobre a intenção do designer sobre o sistema do *Chatbot* Mathos AI. *Você é alguém que valoriza instruções claras, visualmente organizadas e acessíveis,*

que permitam identificar facilmente os elementos na interface, além de apreciar uma estética original e agradável. Você aprecia uma estética original e agradável, com ícones e botões que comuniquem claramente sua função, além de poder explorar e descobrir novas funcionalidades por conta própria, além de preferir experiências rápidas e interativas, que agilizem seu processo de resolução de problemas no domínio da matemática.

Você precisa resolver problemas matemáticos de maneira eficiente, com suporte direto, explicações acessíveis e recursos que favoreçam tanto a compreensão como a prática. Para isso, projetei um Chatbot baseado em IA Generativa, que possui uma interface amigável, informativa e interativa. Com isso, você pode utilizar o sistema enviando os seus problemas matemáticos, dúvidas e explorações conceituais, para que o Chatbot possa decompor, explicar e resolver questões, para que você possa atingir seus objetivos de forma prática, clara e interativa.

5.3 Resultados do MAC

Esta seção irá trazer com detalhes os resultados obtidos com o Método de Avaliação de Comunicabilidade MAC.

5.3.1 Execução

Para a execução do MAC, foram chamados 6 participantes (U1-U6), alunos das disciplinas de Cálculo *I* e Cálculo *III*, estudantes do curso de Engenharia da Computação, sendo 4 homens e 4 mulheres, com diferentes experiências acerca das ferramentas utilizadas para apoiar os seus estudos. Os detalhes sobre cada participante podem ser vistos na Tabela 1. Em média, cada participante levou entre 2 e 6 minutos para fazer cada tarefa, com a T1 levando mais tempo pois os usuários foram incentivados a explorar a interface e suas funcionalidades.

Conforme descrito na metodologia, o teste foi gravado e posteriormente analisado pela avaliadora, de acordo com as etapas do MAC. Durante a análise, foram identificadas rupturas de comunicabilidade, que foram associadas às etiquetas adequadas. A tabela 2 mostra a frequência de etiquetas por tarefa e o total geral, e a Figura 23 mostra a frequência total de cada etiqueta.

Na tabela 3, é possível ver o desempenho dos usuários ao realizar cada tarefa. Todas as tarefas tiveram ao menos uma ruptura, com a com o maior número de rupturas sendo T6, referente à funcionalidade "Teste-me". Em geral, os usuários que tinham mais experiência

Tabela 1 – Participantes do experimento MAC

Usuário	Turma de Cálculo	Principais dificuldades com a disciplina	Método de estudo que usa normalmente	Experiência com ferramentas de IA
U1	Cálculo I	Dificuldade em aplicar os conteúdos	Estudo por vídeoaulas	Já utilizou mas não com frequência
U2	Cálculo I	Dificuldade em fazer os exercícios	Estudo por vídeoaulas e IA	Utiliza com frequência
U3	Cálculo I	Dificuldade média	Estudo por aulas e exercícios	Já utilizou mas não com frequência
U4	Cálculo I	Dificuldade em entender os conceitos	Estudo por vídeoaulas e IA	Utiliza não com frequência
U5	Cálculo III	Dificuldade média	Estudo por exercícios	Já utilizou mas não com frequência
U6	Cálculo III	Dificuldade em fazer os exercícios	Estudo por vídeoaulas, exercícios e IA	Já utilizou mas não com frequência

Fonte: Elaborada pela autora.

Tabela 2 – Frequência de etiquetas por tarefa

Etiqueta	T1	T2	T3	T4	T5	T6	T7	Total
Ué, o que houve?	3	2	2	2	5	3	3	20
Cadê?	0	0	1	1	2	1	3	8
Socorro!	0	0	0	0	1	0	2	3
Epa!	1	2	2	2	0	1	0	8
Por que não funciona?	0	2	2	1	2	5	2	14
O que é isso?	2	1	2	1	0	1	1	8
E agora?	0	1	1	0	0	1	0	3
Não, obrigado	0	0	1	0	0	0	0	1
Pra mim está bom	0	0	0	0	0	1	0	1
Onde estou?	0	0	0	1	0	1	0	2
Assim não dá	0	0	0	0	0	0	0	0
Vai de outro jeito	0	0	0	0	0	0	0	0
Desisto	0	0	0	0	0	0	0	0
Total	6	8	11	8	10	14	11	68

Fonte: Elaborada pela autora.

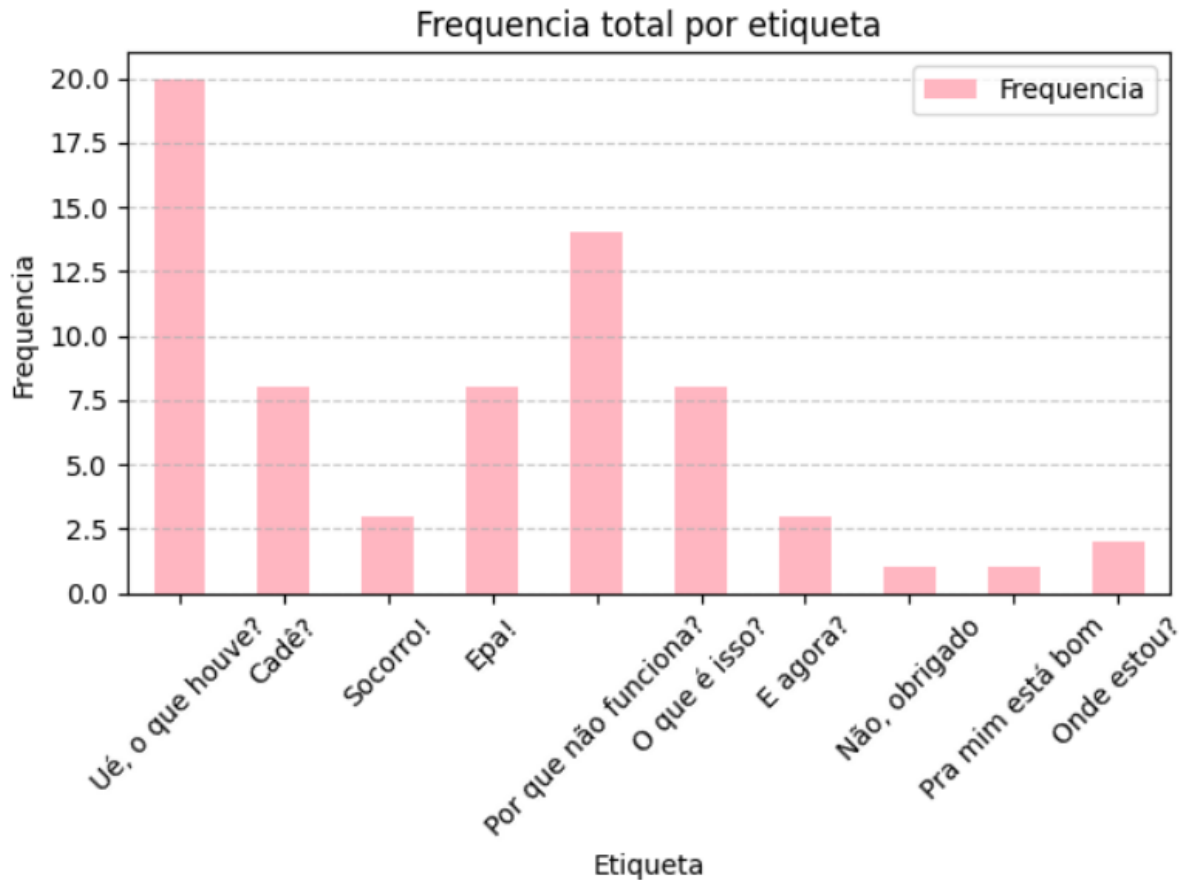
demonstraram mais conforto e confiança, como U1, U2 e U6, explorando mais a funcionalidade de Chat, reformulando *prompts* e gerando novas respostas durante a execução das tarefas T1 e T2.

Tabela 3 – Frequência de rupturas por tarefa e usuário

Tarefa	U1	U2	U3	U4	U5	U6	Total
T1	6	0	0	0	0	0	6
T2	0	0	4	1	2	1	8
T3	4	3	2	0	2	0	11
T4	0	0	0	2	0	6	8
T5	0	1	1	3	0	5	10
T6	1	2	0	4	2	5	14
T7	0	4	1	5	0	1	11
Total	11	10	8	15	6	18	68

Fonte: Elaborada pela autora.

Figura 23 – Frequência total por etiqueta

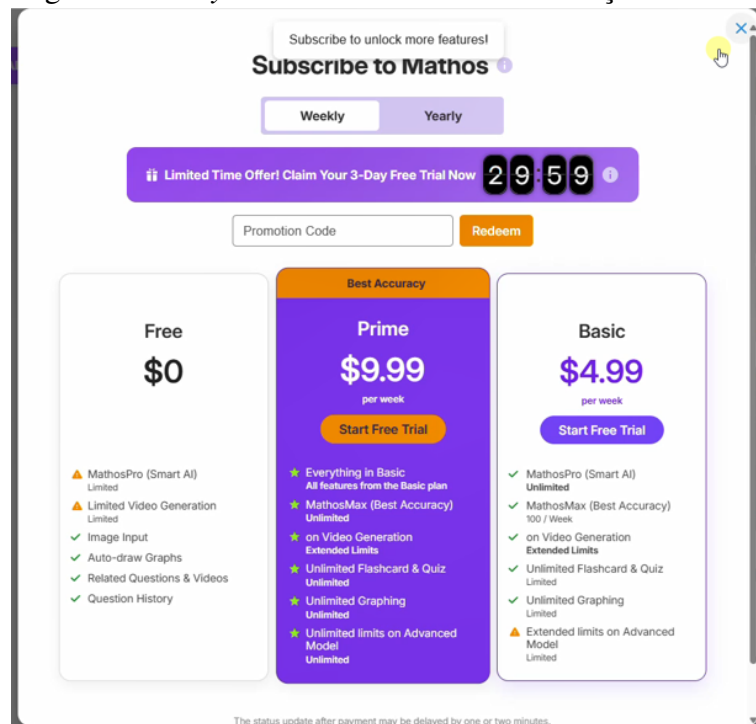


Fonte: Elaborado pela autora.

5.3.1.1 Enviar uma pergunta ao Chatbot

Todos os usuários, por já terem experiência com interfaces de chat, conseguiram executar a tarefa sem problemas. Entretanto, durante a interação, U1 acabou enviando a pergunta com um erro de digitação, o que fez com que o chat enviasse uma resposta inadequada para a pergunta feita. Além disso, U1 permaneceu na interação, pedindo que o *Chatbot* enviasse mais exemplos, mesmo que na intenção incorreta. Para além disso, após o envio de três mensagens, a tarefa foi interrompida por um *paywall*, como mostra a captura na Figura 24. Segundo CAMBRIDGE Dictionary (2026), um *paywall* é “um programa que impede as pessoas que não pagaram uma inscrição de usar um website.” Esta interrupção também ocorreu nas tarefas T2, T3, T4 e T7. Durante a ocorrência dessa ruptura em todas as ocorrências, a interface não deu nenhum indicativo de limite de interação, apenas mostrou a opção de um plano pago.

Figura 24 – *Paywall* exibido durante as interações

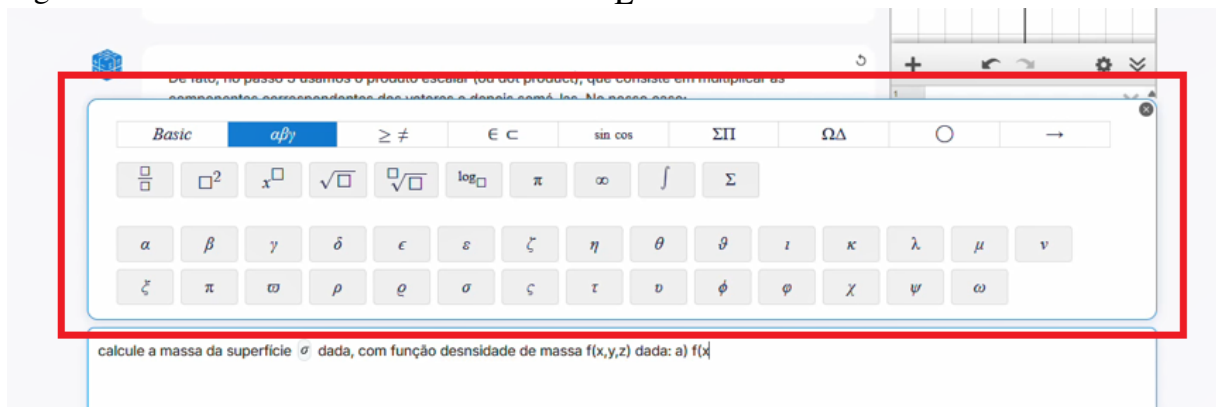


Fonte: Adaptado de www.mathgptpro.com.

5.3.1.2 Enviar uma questão para o Chatbot resolver

Esta tarefa também pôde ocorrer sem problemas para a maioria dos usuários, com a maior confusão sendo para formatar o texto utilizando o teclado \LaTeX embutido na ferramenta, como pode ser visto na Figura 25. Porém, para os usuários U3 e U5, a atividade foi interrompida por um *paywall*, sem a opção de tentar novamente, gerando etiquetas como "Ué, o que houve?", "O que é isso?", "Por que não funciona?", devido a interrupção imprevista.

Figura 25 – Menu com símbolos e fórmulas \LaTeX



Fonte: Elaborado pela autora. Adaptado de www.mathgptpro.com

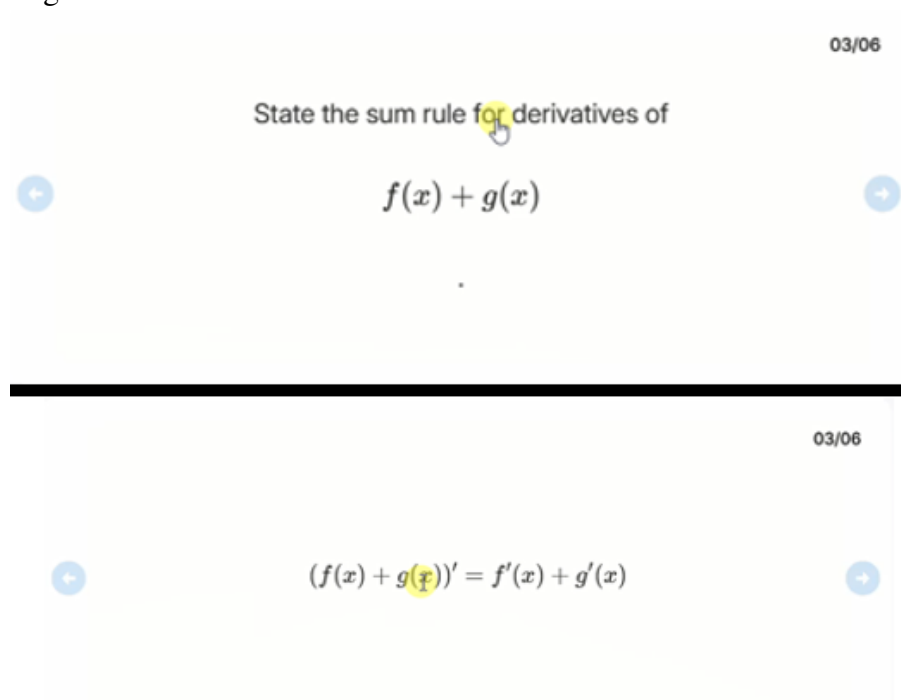
5.3.1.3 Enviar um erro conceitual / erro de cálculo

Durante os testes, os usuários tiveram experiências variadas com as tarefas T3 e T4. Os usuários U2, U3, U4 e U6 receberam *paywalls*, entretanto, durante a interação, U4 encontrou uma forma de fazer com que o *Chatbot* gere novamente a resposta, apesar de estar sem créditos, o que foi marcado pelo avaliador como "Não, obrigado". Também, durante a interação, U6 não entendeu a resposta dada pelo *Chatbot*, pois este trouxe um conceito que não havia sido mencionado em nenhuma das outras mensagens durante a interação, o que, apesar de estar mais relacionado com o modelo generativo, também causou uma ruptura na comunicação entre usuário e sistema.

5.3.1.4 Utilizar as outras funcionalidades dentro do Chat

Durante a execução desta atividade, os usuários U4 e U6 não identificaram as funcionalidades de imediato, precisando de ajuda para encontrá-las e executá-las. Ao gerar flashcards, o usuário U1 não conseguiu entender como prosseguir, pois a interface não comunica que, ao clicar, é possível ver a resposta para a flashcard, como mostra a Figura 26. Além disso, para a maioria dos usuários, o material gerado pelo *Chatbot* foi em inglês, o que causou confusão em alguns participantes pois a interação toda estava ocorrendo em português.

Figura 26 – Tela de flashcards



Fonte: Adaptado de www.mathgptpro.com.

5.3.1.5 Utilizar a funcionalidade Teste-me

Durante a execução do MIS, ao examinar esta funcionalidade foi identificado que esta não comunica o que deve ser enviado como entrada, podendo causar confusões nos usuários. Esta hipótese foi confirmada durante os testes, com o usuário U4 tendo que repetir a tarefa após obter instruções com o avaliador. Além disso, o usuário U6, ao enviar uma questão para a realização do teste, não enviou mais detalhes sobre o enunciado, o que fez com que este não entendesse o progresso na tarefa e na resolução da questão; por causa do ocorrido, o usuário tentou iniciar a tarefa novamente, porém o sistema não permitiu cancelar o teste que já estava em andamento.

5.3.1.6 Resolver questões utilizando a funcionalidade de leitura de PDF

Durante a execução desta tarefa, dois usuários tiveram problemas após trocar de ferramenta dentro da funcionalidade, o que gerou rupturas de “Socorro!”. Também, ao realizar a inspeção dentro do MIS, o avaliador interpretou esta funcionalidade como sendo fácil de utilizar, com o ícone de “varinha mágica” sendo utilizado para representar seleção não causando nenhum estranhamento, entretanto, durante a execução do teste, foi visto que esta escolha causou confusão nos usuários.

5.3.2 Perfil Semiótico

Segundo Prates *et al.* (2000), a elaboração de perfil consiste em interpretar os dados em termos semióticos, com o objetivo de obter a metacomunicação original do designer, ou seja, o significado da mensagem entre o designer e o usuário. Para a construção do perfil semiótico, serão respondidas as seguintes perguntas propostas por Souza e Leitão (2009):

- Quem o designer pensa ser o usuário do produto por ele projetado?
- Quais são os desejos e as necessidades dos usuários, na visão do designer?
- Na visão do designer, de que maneiras os usuários preferem fazer o que desejam e precisam, onde, quando e por quê?
- Qual foi o sistema que o designer projetou para os usuários e como eles devem utilizá-lo?
- Qual é a visão de design e quão bem a lógica é compreendida pelos usuários?

5.3.2.1 *Quem o designer pensa ser o usuário do produto por ele projetado?*

O aplicativo parece ter sido desenvolvido para usuários que possuem a necessidade de resolver problemas matemáticos e compreender conceitos de forma prática e simples, natural como uma conversa. A função principal de Chat da interface é acessível até para usuários menos experientes, entretanto, funcionalidades como o quiz e a leitura de PDF requerem uma certa carga cognitiva maior para usuários que não possuem experiência com ferramentas de estudo digital interativas, além de eventualmente gerar conteúdo em inglês. A partir disso, é possível concluir que o perfil de usuário buscado são pessoas que têm uma certa afinidade com ferramentas de estudo digital e não apenas *Chatbots* de IA.

5.3.2.2 *Quais os desejos e necessidades dos usuários, na visão do designer?*

Os desejos e necessidades dos usuários, como foi percebido durante a execução do teste, é a resolução de perguntas relacionadas à matemática, além de praticar conceitos e questões. Entretanto, o sistema também espera que o usuário explore mais a interface para que possa identificar as outras funcionalidades, testando e descobrindo como utilizar o sistema para alcançar os seus objetivos.

5.3.2.3 *Na visão do designer, de que maneiras os usuários preferem fazer o que desejam e precisam, onde, quando e por quê?*

Ao que parece, os usuários preferem realizar as suas tarefas como se fosse uma interação com um colega ou tutor, porém em um ambiente diferente da sala de aula ou trabalho. O sistema oferece, por exemplo, diferentes maneiras de testar o conhecimento do usuário sobre determinado tópico, com flashcards, quizzes ou apenas resolvendo questões durante a conversa dentro do Chat.

5.3.2.4 *Qual foi o sistema que o designer projetou para os usuários e como eles devem utilizá-lo?*

O sistema projetado é uma mistura de interface conversacional com ferramentas comumente utilizadas para a prática de conceitos e questões de forma interativa, e o usuário pode tanto utilizar o sistema a partir de uma conversa (via Chat) ou utilizar as funcionalidades para praticar os conceitos de forma ativa.

5.3.2.5 *Qual é a visão de design e quão bem a lógica é compreendida pelos usuários?*

A interface busca incluir todas as informações necessárias dentro da interface para facilitar a identificação e a encontrabilidade, entretanto, isso acaba causando algumas confusões que fazem com que os usuários tenham dificuldade de encontrar determinadas funcionalidades. Mesmo assim, a interface ainda é "agradável", oferecendo múltiplas funções ao alcance imediato do usuário.

5.4 **Triangulação dos Resultados**

A partir do uso e combinação de métodos de avaliação a partir de tanto investigação, observação e inspeção, foi possível obter noções sobre múltiplos aspectos da interface e do sistema. Esta triangulação de dados possui foco em aspectos qualitativos da interação, entretanto, os dados quantitativos encontrados durante a aplicação dos testes também permitiu a obtenção de perspectivas complementares sobre o comportamento do sistema, a interpretação dos usuários e as rupturas observadas durante a interação.

O uso da triangulação de métodos se deu para a obtenção de conclusões que são confirmadas por diferentes perspectivas, minimizando possíveis vieses que poderiam surgir ao restringir a experimentação a um único método. Além disso, como cada método utiliza diferentes grupos de usuários, também obtemos a triangulação de dados, ou seja, dados que foram produzidos em diferentes momentos e com diferentes pessoas.

5.4.1 *Análise e Interpretação dos Resultados*

O processo de triangulação foi conduzido por meio do cruzamento dos resultados obtidos em cada método. Inicialmente, os dados foram analisados de forma independente, e então foram comparados em aspectos sobre divergências, convergências e informações complementares. Aqui, serão considerados os resultados que puderam ser encontrados em mais de um método.

5.4.1.1 *Principais Convergências*

Por exemplo, a partir do MALTU, foi possível ver que, eventualmente, o sistema comete erros durante a interação, que podem passar despercebidos pelo usuário, o que foi confirmado dentro do MAC, onde, em uma instância, o usuário fez uma correção no enunciado

enviado porém a resposta enviada era extremamente similar à resposta errada, dando a impressão de que o *Chatbot* está persistindo no erro.

Dentro do MIS, foi notado que a interface da funcionalidade “Teste-me” não possui uma comunicação clara sobre o seu funcionamento, podendo causar confusão no usuário durante o uso, o que foi confirmado durante o MAC, com usuários precisando repetir a tarefa por não enviar a entrada adequada durante a interação.

Além disso, durante o MIS, foi suposto que o uso de ícones próprios poderia causar confusão durante o uso com o usuário, o que foi provado dentro do MAC, com os usuários não conseguindo identificar como trocar o modelo generativo por causa do ícone no switch.

5.4.1.2 Principais Complementações

A partir dos resultados do MALTU, foram pontuadas reclamações acerca das cobranças dentro do aplicativo e da falta de comunicação entre o sistema e o usuário sobre renovações de planos e limitações da versão gratuita. Por limitações da avaliação, não foi possível testar como a cobrança é realizada atualmente ao assinar o sistema, entretanto, durante a realização do MAC, as interações foram interrompidas múltiplas vezes por paywalls e cobranças, portanto, surge a hipótese de que não houveram melhoras sobre o método de cobrança dos usuários.

Além disso, também no MALTU, foram feitos comentários sobre a linguagem padrão do sistema ser o inglês, o que, apesar de melhorado e traduzido para diferentes línguas, durante algumas interações o conteúdo gerado pelo *Chatbot* é em inglês, ao contrário da língua em que a interação está sendo realizada.

5.4.1.3 Principais Divergências

Durante o MIS, a interface foi dada como descritiva e fácil de navegar, entretanto, ao realizar o MAC, foi notado que, com a quantidade de informações presentes na tela, alguns usuários tiveram dificuldade em se encontrar dentro da interface.

Também, na funcionalidade de leitura de PDF, o MIS a classificou como simples de utilizar, principalmente por causa dos ícones utilizados e do tutorial, porém, durante o MAC, esta funcionalidade obteve algumas rupturas durante a sua execução.

5.4.2 O Processo de Construção-Síntese

A partir do uso da triangulação, foi possível obter um conjunto de dados e conclusões que são confirmadas por diferentes métodos. Por exemplo, a partir dos métodos utilizados, foi possível identificar falhas na comunicação acerca de cobranças e limitações, e, também, foi concluído que a interface, apesar de amigável, pode ser confusa em certos aspectos.

Com o uso de diferentes métodos, foi possível conhecer quem são os usuários que utilizam o sistema atualmente, quais suas satisfações e insatisfações e quem estaria interessado em utilizar o sistema futuramente, além de obter conhecimentos sobre quem concebeu e projetou o sistema a partir das decisões identificadas na interface (Barbosa; Silva, 2010). Com o uso do MALTU, foi possível compreender as impressões já existentes sobre o sistema e direcionar os testes posteriores; com o uso do MIS, foi possível conhecer o designer e suas intenções com o sistema; e, com o MAC, foi possível validar este conhecimento prévio a partir de uma interação real com usuários que, apesar de diferentes perfis, possuem o perfil de estudantes que necessitam de ferramentas para auxiliar nos seus estudos.

5.5 Respondendo às questões de pesquisa

5.5.1 Como os métodos de avaliação de IHC podem ser utilizados dentro do contexto de Chatbots tutores/educacionais?

A partir dos experimentos realizados e os resultados obtidos, concluiu-se que os métodos de avaliação de IHC podem ser utilizados para avaliar em diferentes contextos, seja com a participação de usuários ou não. Neste trabalho, métodos sem participação direta de usuários, como o MIS, possibilitaram investigar as intenções do designer e os significados embutidos nos signos da interface, permitindo antecipar potenciais rupturas de interpretação. Já métodos com participação de usuários, como o MAC, evidenciaram como essas intenções são efetivamente interpretadas na prática, revelando discrepâncias entre o que o sistema “diz” e o que o usuário compreende durante o uso real. Complementarmente, o MALTU contribuiu ao trazer indícios prévios de dificuldades e expectativas dos usuários, ajudando a contextualizar os achados das demais avaliações. Os métodos de avaliação auxiliam para o desenvolvimento de uma interface que seja útil, agradável e que proporcione uma boa experiência de usuário.

5.5.2 *Quais métodos de avaliação seriam mais adequados no contexto de um voltado para a educação?*

Os resultados indicam que, no contexto de *Chatbots* educacionais, os métodos mais adequados são aqueles capazes de investigar não apenas aspectos tradicionais de usabilidade, mas principalmente a comunicabilidade e a construção de sentido na interação. Por se tratar de uma interface baseada em linguagem natural, em que instruções, feedbacks e respostas são apresentados de forma textual e conversacional, torna-se essencial utilizar métodos que analisem como o sistema comunica suas intenções e como os usuários interpretam essa comunicação.

5.5.3 *Quais são possíveis requisitos para um Chatbot voltado para a educação?*

Com os resultados obtidos, foi observado que é essencial para um *Chatbot* educacional uma boa comunicação com seu usuário, dando explicações claras e permitindo que este possa esclarecer suas dúvidas livremente. Além disso, uma boa comunicabilidade de interface também é essencial, indicando ao usuário o que ele pode fazer, como ele pode fazer e oferecendo feedback após ser feito. Não existe um método único que possa avaliar todos os aspectos importantes da interface, mas a realização de uma análise multimétodo que envolva tanto a opinião de especialistas quanto a de usuários permite uma análise sobre a perspectiva do designer e a sua percepção pelos usuários.

5.5.4 *Como os Chatbots podem auxiliar dentro do processo de aprendizagem de matemática?*

Segundo Stuart (2000) e Boaler (2002), o ensino tradicional de Matemática não abrange todas as necessidades dos estudantes, o que gera dificuldades no aprendizado. A partir desse contexto, os *Chatbots* podem auxiliar dentro desse processo oferecendo suporte personalizado, adequado às necessidades individuais dos alunos. No contexto observado neste trabalho, o *Chatbot* avaliado mostrou potencial para atuar como mediador do raciocínio matemático, auxiliando na decomposição de problemas complexos, na explicitação de etapas intermediárias e na reformulação de explicações quando solicitado. Dessa forma, *Chatbots* podem contribuir para reduzir barreiras comuns no aprendizado de Matemática, funcionando como um apoio complementar ao professor e às práticas pedagógicas formais.

6 CONCLUSÕES E TRABALHOS FUTUROS

Este trabalho buscou avaliar diferentes aspectos sobre o *Chatbot Mathos AI*, buscando compreender as intenções do designer durante o projeto da interface e como essas intenções são interpretadas pelos usuários. A avaliação analisou possíveis falhas de interação que pudessem comprometer a experiência geral do usuário, identificando pontos de ruptura de comunicação, satisfações e insatisfações de quem usa o sistema. Para isso, foram usados os seguintes métodos:

MALTU: Para realizar uma investigação a partir dos comentários e opiniões já existentes dos usuários, permitindo compreender previamente possíveis rupturas e dificuldades enfrentadas pelos usuários;

MIS: Para inspecionar a interface, buscando compreender as intenções do designer sobre os signos utilizados e o que ele busca comunicar ao usuário;

MAC: Para acompanhar o uso do sistema em um cenário de uso controlado, porém representativo, possibilitando a análise da interpretação dos usuários ao longo da interação.

A aplicação do MALTU permitiu investigar o estado do sistema a partir dos comentários obtidos na Play Store, obtendo uma visão prévia do sistema que orientou os experimentos seguintes. Durante a avaliação, foi possível identificar algumas questões que deveriam ser identificadas durante os próximos métodos, como por exemplo a falta de comunicação acerca de cobranças e as limitações do *Chatbot*.

Já na execução do MIS, foi possível inspecionar a interface, investigando as decisões do designer e como elas são aplicadas dentro do sistema, além de antecipar possíveis rupturas de comunicabilidade que poderiam ser encontradas durante a execução do MAC. Nessa etapa, algumas hipóteses levantadas na avaliação prévia foram confirmadas, mas também foram obtidas novas hipóteses a serem testadas, como a percepção dos usuários em relação à funcionalidade “Teste-me”.

Em seguida, na execução do MAC, foi possível observar os usuários diretamente, entendendo as suas decisões e os seus processos de pensamento ao utilizar a interface. Durante a interação, foram encontradas algumas dificuldades relacionadas à navegação e à comunicação entre usuário e designer, além de encontrar novas rupturas causadas pelas falhas de comunicação.

Por fim, durante a consolidação dos dados, foi realizada uma triangulação sobre os diferentes resultados, permitindo identificar convergências entre os métodos e aspectos destacados de forma isolada em cada avaliação, fortalecendo a análise e fornecendo uma visão mais abrangente sobre os pontos fortes e fragilidades do sistema.

Como contribuição, este trabalho reforça o uso combinado de métodos de avaliação de IHC, aplicados no contexto de *Chatbots* educacionais, além de prover subsídios para o aprimoramento da interface e comunicação do sistema avaliado. A partir desta análise, foi possível compreender padrões presentes em interfaces conversacionais, e como estas podem ser adaptadas para uma experiência de aprendizado enriquecida. Ademais, a pesquisa contribui para a discussão sobre a avaliação de fatores que influenciam na experiência do usuário em sistemas baseados em inteligência artificial, ainda pouco explorada no contexto de interfaces educacionais.

6.1 Limitações e Ameaças à Validade

É importante reconhecer que este estudo apresenta limitações que podem influenciar a generalização dos resultados e o entendimento dos problemas de comunicabilidade do *Chatbot*. Uma das limitações diz respeito à amostra utilizada, que, apesar de ser um número adequado para o tipo de avaliação, foi considerada limitada para as expectativas do avaliador. Juntamente disso, a limitação da análise sobre as turmas de Cálculo restringiu a perspectiva dos resultados, não contemplando outras áreas que poderiam oferecer perspectivas complementares sobre o uso do sistema.

Para além disso, a seleção de tarefas e cenários adotada neste estudo esteve alinhada a um recorte específico de uso do sistema, o que implica que outras situações e problemas podem não ter sido contemplados neste estudo. Ademais, a falta de respostas ao questionário reduziu a obtenção de dados mais detalhados; entretanto, as evidências obtidas foram suficientes para sustentar a análise e fornecer insights relevantes sobre a interação com o sistema.

Apesar das limitações identificadas, o presente trabalho evidenciou a relevância e a efetividade da investigação de interfaces conversacionais no contexto do aprendizado, neste caso, utilizando o sistema *Mathos AI*, por meio da combinação de métodos e triangulação de dados. A análise permitiu uma compreensão mais aprofundada sobre diversos aspectos da interface e como os usuários interpretam e respondem a essa comunicação, indicando caminhos de melhorias que podem ser aplicados em diferentes contextos e sistemas. Assim, este trabalho busca contribuir para o fortalecimento das pesquisas em IHC e para o desenvolvimento de soluções mais eficazes e centradas no usuário.

6.2 Trabalhos futuros

Como continuidade deste trabalho, recomenda-se a ampliação do escopo das avaliações realizadas, especialmente por meio do aumento do número de participantes e a diversificação do perfil dos usuários envolvidos. A inclusão de estudantes em diferentes níveis de ensino, presentes em diferentes contextos, pode contribuir para uma compreensão mais abrangente dos aspectos de comunicabilidade e usabilidade do sistema em distintos contextos educacionais.

Além disso, sugere-se a aplicação dos métodos utilizados neste estudo em outros contextos, como outros *Chatbots* educacionais ou interfaces com objetivos similares. Dessa forma, será possível identificar padrões recorrentes de problemas e boas práticas de design de interfaces conversacionais voltadas ao ensino.

Em conjunto, trabalhos futuros podem explorar a condução de estudos longitudinais, acompanhando a interação ao longo do tempo, podendo observar a evolução da compreensão do usuário, além do impacto do uso do *Chatbot* no processo de aprendizagem individual.

Por fim, recomenda-se que os resultados obtidos neste trabalho sejam utilizados como subsídio para o aprimoramento do *Mathos AI*, especialmente no que diz respeito à clareza das funcionalidades, à comunicação de limitações do sistema e à consistência da interface, mas que também possam servir como indicações e recomendações dentro da implementação de diferentes sistemas.

6.3 Considerações finais

O crescimento do uso de *Chatbots* no contexto educacional, principalmente como ferramenta de apoio ao ensino e aprendizado de Matemática, evidencia a necessidade de investigações e avaliações sobre a qualidade da interação entre usuários e sistemas. Nesse cenário, este trabalho buscou contribuir para a compreensão de como interfaces conversacionais podem impactar a experiência do usuário a partir da avaliação do *Chatbot* educacional *Mathos AI*.

Ao longo da pesquisa, foi conduzida uma avaliação fundamentada em diferentes métodos da área de IHC, como o Modelo para Avaliação da Interação em Sistemas Sociais a partir da Linguagem Textual do Usuário MALTU, o Método de Inspeção Semiótica MIS e o Método de Avaliação de Comunicabilidade MAC, com o apoio de entrevistas e questionários. Com a combinação desses métodos, foi possível analisar o sistema sob diferentes perspectivas, ampliando a compreensão sobre aspectos relacionados à usabilidade, comunicabilidade e à

percepção dos usuários.

Apesar das limitações inerentes ao escopo do estudo, como o recorte do público avaliado e das tarefas analisadas, os achados fornecem subsídios relevantes tanto para o aprimoramento do Mathos AI quanto para futuras pesquisas na área de IHC. As reflexões apresentadas reforçam o papel da avaliação como etapa fundamental dentro do desenvolvimento de interfaces mais transparentes, eficientes e alinhadas às necessidades reais dos usuários.

Dessa forma, espera-se que este trabalho contribua para o avanço das discussões sobre a avaliação de *Chatbots* educacionais e incentive a adoção de abordagens metodológicas integradas na análise de interfaces conversacionais, desenvolvendo soluções tecnológicas centradas no ser humano.

REFERÊNCIAS

- ADAMOPOULOU, E.; MOUSSIADES, L. An overview of chatbot technology. In: MAGLOGIANNIS, I.; ILIADIS, L.; PIMENIDIS, E. (Ed.). **Artificial Intelligence Applications and Innovations**. Cham: Springer International Publishing, 2020. p. 373–383. ISBN 978-3-030-49186-4.
- BARBOSA, S.; SILVA, B. **Interação humano-computador**. [S. l.]: 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.
- BATISTA, G. O. da S.; MONTEIRO, M. de S.; SALGADO, L. C. de C. Investigando chatbots governamentais: um panorama sobre a usabilidade dentro e fora do brasil. In: SBC. **Workshop de Computação Aplicada em Governo Eletrônico (WCGE)**. [S. l.], 2022. p. 73–84.
- BEHAR, P. A. O ensino remoto emergencial e a educação a distância. **Jornal da Universidade**, 2 jul. 2020, ed. 15, 2020.
- BOALER, J. Paying the price for "sugar and spice": Shifting the analytical lens in equity research. **Mathematical thinking and learning**, Taylor & Francis, v. 4, n. 2-3, p. 127–144, 2002.
- BOULAY, B. du. Inteligência artificial na educação e ética. **RE@ D-Revista de Educação a Distância e Elearning**, v. 6, n. 1, p. e202303–e202303, 2023.
- BRANDTZAEG, P. B.; FØLSTAD, A. Why people use chatbots. In: SPRINGER. **Internet Science 4th International Conference, INSCI 2017, Thessaloniki, Greece, November 22-24, 2017, Proceedings 4**. [S. l.], 2017. p. 377–392.
- CAMBRIDGE Dictionary. **Paywall**. 2026. Disponível em: <https://dictionary.cambridge.org/dictionary/english/paywall>. Acesso em: 10 jan. 2026.
- CANDELLO, H.; PINHANEZ, C. Designing conversational interfaces. **Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais-IHC**, v. 100, 2016.
- CARREIRA, G.; SILVA, L.; MENDES, A. J.; OLIVEIRA, H. G. Pyo, a chatbot assistant for introductory programming students. In: IEEE. **2022 International Symposium on Computers in Education (SIIE)**. [S. l.], 2022. p. 1–6.
- CASAS, J.; TRICOT, M.-O.; KHALED, O. A.; MUGELLINI, E.; CUDRE-MAUROUX, P. Trends and methods in chatbot evaluation. In: **Proceedings of ICMI 2020 22nd ACM International Conference on Multimodal Interaction**. [S. l.: s. n.], 2020. p. 280–286.
- CASELI, H. M.; NUNES, M. G. V. (Ed.). **Processamento de Linguagem Natural: Conceitos, técnicas e aplicações em português**. 3. ed. BPLN, 2024. ISBN 978-65-01-20581-6. Disponível em: <https://brasileiraspln.com/livro-pln/3a-edicao/>.
- CASTRO, U. S.; SILVA, A. C. de C.; PRESTES, E. T. P.; CORREIA, F. da C.; TEIXEIRA, M. L. L. D.; NETO, R. C. da S.; OLIVEIRA, R.; SILVA, W. C. C. N. O ensino de matemática no século XXI: Desafios e tecnologias emergentes. **REVISTA FOCO**, v. 18, n. 2, p. e7811–e7811, 2025.

CORRÊA, J. N. P.; BRANDEMBERG, J. C. Tecnologias digitais da informação e comunicação no ensino de matemática em tempos de pandemia: desafios e possibilidades. **Boletim Cearense de Educação e História da Matemática**, v. 8, n. 22, p. 34–54, 2021.

DAMASCENO, A. R.; MARTINS, A. R.; CHAGAS, M. L.; BARROS, E. M.; MAIA, P. H. M.; OLIVEIRA, F. C. Stuart: an intelligent tutoring system for increasing scalability of distance education courses. In: **Proceedings of the 19th Brazilian Symposium on Human Factors in Computing Systems**. [S. l.: s. n.], 2020. p. 1–10.

DENZIN, N. K. **The research act: A theoretical introduction to sociological methods**. [S. l.]: Routledge, 2017.

FLORES, M. A.; SIMÃO, A. M. V.; BARROS, A.; FLORES, P.; PEREIRA, D.; FERNANDES, E. L.; FERREIRA, P. C.; COSTA, L. Ensino e aprendizagem à distância em tempos de covid-19: um estudo com alunos do ensino superior. **Revista Portuguesa de pedagogia**, v. 55, p. e055001–e055001, 2021.

GOMES, R.; SOUZA, E. R. d.; MINAYO, M. d. S.; MALAQUIAS, J. V.; SILVA, C. F. R. d. Organização, processamento, análise e interpretação de dados: o desafio da triangulação. **Avaliação por triangulação de métodos abordagem de programas sociais. Rio de Janeiro Fiocruz**, p. 185–221, 2005.

GONZALEZ, M.; LIMA, V. L. S. Recuperação de informação e processamento da linguagem natural. In: SN. **XXIII Congresso da Sociedade Brasileira de Computação**. [S. l.], 2003. v. 3, p. 347–395.

GUERINO, G. C.; VALENTIM, N. M. C. Usability and user experience evaluation of conversational systems: A systematic mapping study. In: **Proceedings of the XXXIV Brazilian Symposium on Software Engineering**. [S. l.: s. n.], 2020. p. 427–436.

HEWETT, T.; BAECKER, R.; CARD, S.; CAREY, T.; GASEN, J.; TREMAINE, M.; PERLMAN, G.; STRONG, G.; VERPLANK, W. Acm sigchi curricula for human-computer interaction. 01 1992.

HOBERT, S. Say hello to 'coding tutor'! design and evaluation of a chatbot-based learning system supporting students to learn to program. In: MUNICH. **Proceedings of the International Conference on Information Systems (ICIS)**. [S. l.], 2019.

HWANG, G.-J.; TU, Y.-F. Roles and research trends of artificial intelligence in mathematics education: A bibliometric mapping analysis and systematic review. **Mathematics**, MDPI, v. 9, n. 6, p. 584, 2021.

JAKOBSON, R.; SEBEOK, T. A. Closing statement: Linguistics and poetics. **Semiotics An introductory anthology**, Indiana University Press Bloomington, p. 147–175, 1960.

KUCHERBAEV, P.; BOZZON, A.; HOUBEN, G.-J. Human-aided bots. **IEEE Internet Computing**, v. 22, n. 6, p. 36–43, 2018.

LANGVIN, R.; LORDON, R. J.; AVRAHAMI, T.; COWAN, B. R.; HIRSCH, T.; HSIEH, G. Heuristic evaluation of conversational agents. In: **Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems**. [S. l.: s. n.], 2021. p. 1–15.

LEESANG-GU; HWA, L.; SUN, L.; PARKDOYOUNG. Teaching mathematics for ai with chatgpt and gpt-4 omni. **Research in Mathematical Education**, v. 27, n. 4, p. 449–466, 12 2024.

MARINONI, G.; LAND, H. V.; JENSEN, T. *et al.* The impact of covid-19 on higher education around the world. **IAU global survey report**, v. 23, n. 1, p. 1–17, 2020.

MARKETSANDMARKETS. **Chatbot Market by Offering, Bot Communication (Text, Audio, Video), Type, Business Function (Sales Marketing, Contact Centers), Channel Integration, Vertical (Retail eCommerce, Healthcare Life Sciences) and Region - Global Forecast to 2028**. MarketsAndMarkets, 2023. Disponível em: <https://www.marketsandmarkets.com/Market-Reports/chatbot-market-72302363.html>. Acesso em: 27 mai. 2025.

MENDES, M. S. **MALTU – Um modelo para avaliação da interação em sistemas sociais a partir da linguagem textual do usuário**. 199 p. Dissertação (Tese (Doutorado em Ciência da Computação)) – Universidade Federal do Ceará, Fortaleza, 2015.

MIRANDA, A. L.; GARCIA, R.; ORAN, A. C.; GUEDES, G. T.; SANTANA, B. S.; SILVA, D. G.; VALLE, P. H.; SILVA, W. Avaliação de usabilidade do coderbot como recurso pedagógico no ensino de programação. In: SBC. **Workshop sobre Bots na Engenharia de Software**. [S. l.], 2024. p. 11–20.

MOLNÁR, G.; SZÜTS, Z. The role of chatbots in formal education. In: IEEE. **2018 IEEE 16th international symposium on intelligent systems and informatics (SISY)**. [S. l.], 2018. p. 000197–000202.

MORAES, S. M.; SOUZA, L. S. de. Uma abordagem semiautomática para expansão e enriquecimento linguístico de bases aiml para chatbots. In: **Congresso Internacional de Informática Educativa**. [S. l.: s. n.], 2015. v. 20, p. 600–605.

NEUMANN, A. T.; ARNDT, T.; KÖBIS, L.; MEISSNER, R.; MARTIN, A.; LANGE, P. de; PENGEL, N.; KLAMMA, R.; WOLLERSHEIM, H.-W. Chatbots as a tool to scale mentoring processes: Individually supporting self-study in higher education. **Frontiers in artificial intelligence**, Frontiers Media SA, v. 4, p. 668220, 2021.

NIELSEN, J. **Usability engineering**. [S. l.]: Morgan Kaufmann, 1994.

OELERICH, G.; OTTO, H.-U. **Empirische Forschung und Soziale Arbeit**. [S. l.]: Springer, 2011.

PEIRCE, C. S. **Collected papers of charles sanders peirce**. [S. l.]: Harvard University Press, 1934. v. 5.

PRATES, R. O.; SOUZA, C. S. D.; BARBOSA, S. D. Methods and tools: a method for evaluating the communicability of user interfaces. **interactions**, ACM New York, NY, USA, v. 7, n. 1, p. 31–38, 2000.

RAMESH, K.; RAVISHANKARAN, S.; JOSHI, A.; CHANDRASEKARAN, K. A survey of design techniques for conversational agents. In: KAUSHIK, S.; GUPTA, D.; KHARB, L.; CHAHAL, D. (Ed.). **Information, Communication and Computing Technology**. Singapore: Springer Singapore, 2017. p. 336–350. ISBN 978-981-10-6544-6.

RUBIN, J.; CHISNELL, D. **Handbook of usability testing: How to plan, design and conduct effective tests.** [S. l.]: John Wiley & Sons, 2008.

RUSSELL, S. J.; NORWIG, P. **Artificial Intelligence: A modern approach.** [S. l.]: Gen-LTC, 2022. v. 4.

SETO, S. B.; KADE, G.; LESTARI, M. I.; SENE, M. T.; MUGA, K. V.; TENGGONG, Y. Pengaruh penggunaan aplikasi mathos ai dalam meningkatkan kemampuan pemecahan masalah matematika siswa kelas x sman 1 nangapanda. **De Fermat Jurnal Pendidikan Matematika**, v. 8, n. 1, p. 214–222, 2025.

SHARMA, V.; GOYAL, M.; MALIK, D. An intelligent behaviour shown by chatbot system. **International Journal of New Technology and Research**, v. 3, n. 4, 4 2017.

SHAWAR, B. A.; ATWELL, E. Using dialogue corpora to train a chatbot. In: **Proceedings of the Corpus Linguistics 2003 conference.** [S. l.: s. n.], 2003. p. 681–690.

SILVA, A. S. Uso do chatbot como suporte educacional no ensino e aprendizagem da matemática: um estudo de caso da disciplina de funções de variáveis reais. Instituto Federal de Educação, Ciência e Tecnologia da Bahia, 2024.

SILVA, G. R. S.; CANEDO, E. D. Human factors in the design of chatbot interactions: Conversational design practices. In: **Proceedings of the XXIII Brazilian Symposium on Human Factors in Computing Systems.** [S. l.: s. n.], 2024. p. 1–12.

SILVA, K. R. da; BARBOSA, L. S. de O.; BOTELHO, W. L.; PINHEIRO, J. M. B.; PEIXOTO, I. dos S.; MENEZES, I. V. C. B. de. Inteligência artificial e seus impactos na educação: uma revisão sistemática. **RECIMA21-Revista Científica Multidisciplinar-ISSN 2675-6218**, v. 4, n. 11, p. e4114353–e4114353, 2023.

SOUSA, R. d. *et al.* **AVA na EaD: linguagem total e comunicação.** 123 p. Dissertação (Dissertação (Programa de Pós-Graduação em Educação)) – Universidade Nove de Julho, São Paulo, 2023.

SOUZA, C. S. D. **The semiotic engineering of human-computer interaction.** [S. l.]: MIT press, 2005.

SOUZA, C. S. de; LEITÃO, C. F. **Semiotic engineering methods for scientific research in HCI.** [S. l.]: Morgan & Claypool Publishers, 2009.

STUART, V. Math curse or math anxiety? **Teaching children mathematics**, National Council of Teachers of Mathematics, v. 6, n. 5, p. 330–335, 2000.

WEIZENBAUM, J. Eliza — a computer program for the study of natural language communication between man and machine. **Communications of the ACM**, ACM New York, NY, USA, v. 9, n. 1, p. 36–45, 1966.

WOLLNY, S.; SCHNEIDER, J.; MITRI, D. D.; WEIDLICH, J.; RITTBERGER, M.; DRACHSLER, H. Are we there yet? - a systematic literature review on chatbots in education. **Frontiers in artificial intelligence**, Frontiers Media SA, v. 4, p. 654924, 2021.

APÊNDICE A – CÓDIGO-FONTE DO SCRIPT UTILIZADO PARA A EXTRAÇÃO DE PRUS

Código-fonte 1 – Script para a extração de PRUs

```
1 pip install google-play-scraper
2 import pandas as pd
3 from google_play_scraper import Sort, reviews, reviews_all
4
5 app_id = "com.mathgptpro.mclient"
6 langs = ['pt','en', 'es', 'fr']
7 scores = [1, 2, 3, 4]
8 all_reviews = []
9
10 for lang in langs:
11     for score in scores:
12         all_reviews_lang = reviews_all(
13             app_id,
14             sleep_milliseconds=0,
15             lang=lang,
16             country='us',
17             sort=Sort.NEWEST,
18             filter_score_with=score,
19         )
20         all_reviews.extend(all_reviews_lang)
21         print(f"Reviews in {lang} with {score} stars: {len(
22             all_reviews_lang)}")
23
24 df = pd.DataFrame(all_reviews)
25 df = df[['content', 'score', 'at']]
26 df.to_csv('todas_reviews.csv', index=False)
```

APÊNDICE B – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Termo de Consentimento que foi entregue aos participantes do Teste MAC

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO (TCLE)

Maria Sofia Jeronimo de Lima

December 1, 2025

Você está sendo convidado(a) como voluntário(a) a participar da pesquisa Aplicando métodos de avaliação de IHC no contexto de um *Chatbot* de ensino e aprendizagem de Matemática, que tem como objetivo identificar rupturas de comunicabilidade entre o sistema *Mathos AI* e seus usuários utilizando métodos de avaliação de comunicabilidade. Este estudo busca compreender os padrões utilizados no desenvolvimento do sistema estudado, avaliando se a ferramenta facilita a interação e a compreensão dos usuários, além de sugerir melhorias com base nos testes.

Sua participação na pesquisa é totalmente voluntária, ou seja, não é obrigatória. Caso você decida não participar, ou ainda, desistir de participar e retirar seu consentimento durante a pesquisa, não haverá nenhum prejuízo à sua integridade enquanto estudante.

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 pessoa com idade igual ou superior a 18 anos, que deseje colaborar e esteja de acordo com o presente termo.

2 ENVOLVIMENTO NA PESQUISA

A sua participação no estudo será voluntária e consistirá em participar de um teste dividido em três etapas: uma entrevista pré-teste, onde serão feitas perguntas sobre o seu perfil, a sua experiência dentro das disciplinas de Cálculo I ou III e a sua experiência com ferramentas de apoio ao ensino e aprendizagem; uma avaliação utilizando o Método de Avaliação de Comunicabilidade, onde você irá executar 7 tarefas na ferramenta *Mathos AI*, em plataforma Web, conforme os cenários designados; um questionário final para avaliar a comunicabilidade do sistema. Todo o processo deverá durar cerca de 20 minutos.

Durante a execução da avaliação, a sua interação com a ferramenta será gravada para análise futura. Mesmo que você aceite participar agora, poderá desistir ou interromper sua participação a qualquer momento, sem qualquer penalidade.

3 RISCOS E DESCONFORTOS:

Por se tratar da realização de cenários em uma ferramenta de plataforma web, existe a possibilidade de ocorrerem falhas técnicas, ou, ainda, cansaço ou desconforto ao realaizar alguma tarefa. De forma geral, considera-se que a pesquisa envolve riscos mínimos.

No caso de cansaço ou desconforto, faça intervalos ou pule a tarefa quando for possível.

4 CONFIDENCIALIDADE DA PESQUISA:

A sua privacidade é prioridade e será respeitada. Durante a realização desta pesquisa, você não será identificado(a) em nenhuma publicação resultante. Além disso, os pesquisadores se responsabilizam pela guarda e confidencialidade de todos os dados coletados, que serão mantidos pelo período necessário para a análise e publicação dos resultados. Posteriormente, os dados serão descartados de acordo com os meios legalmente autorizados, atendendo à legislação brasileira (Resolução Nº 466/12 do Conselho Nacional de Saúde). As informações fornecidas serão utilizadas exclusivamente para fins científicos.

5 BENEFÍCIOS E CUSTOS:

A sua participação ajudará a identificar possíveis falhas na comunicação entre o sistema e seus usuários, colaborando para a sugestão de melhorias que tornem a ferramenta mais intuitiva e eficiente. Para participar deste estudo você não vai ter nenhum custo nem receberá qualquer vantagem financeira.

6 DECLARAÇÃO DE RESPONSABILIDADE:

A pesquisadora responsável, Maria Sofia Jeronimo de Lima, aluna da Universidade Federal do Ceará - Campus Quixadá, sob orientação do professor Marcelo Martins da Silva, declara 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 o pesquisador responsável, Maria Sofia Jeronimo de Lima, através do email msofiajmolima@gmail.com, ou Marcelo Martins da Silva, através do email martins2016eng@gmail.com.

ATENÇÃO: Se você tiver alguma consideração ou dúvida, sobre a sua participação na pesquisa, entre em contato com o Comitê de Ética em Pesquisa da UFC/PROPESQ – Rua Coronel Nunes de Melo, 1000 - Rodolfo Teófilo, fone: 3366-8344/46. (Horário: 08:00-12:00 horas de segunda a sexta-feira). O CEP/UFC/PROPESQ é a instância da Universidade Federal do Ceará responsável pela avaliação e acompanhamento dos aspectos éticos de todas as pesquisas envolvendo seres humanos.

8 CONSENTIMENTO DA PARTICIPAÇÃO

Eu, _____, _____ anos, declaro que li cuidadosamente este Termo de Consentimento Livre e Esclarecido (TCLE) e que, após sua leitura, tive a oportunidade de fazer perguntas sobre o seu conteúdo, como também sobre a pesquisa, e recebi explicações que responderam por completo minhas dúvidas. Portanto, concordo em dele participar e dou meu consentimento sem que para isso tenha sido forçado(a) ou obrigado(a).

_____, _____ de _____ de 2025
Assinatura do participante

Pesquisador (Orientando): Maria Sofia Jeronimo de Lima

Professor (Orientador): M^e Marcelo Martins da Silva

APÊNDICE C – LISTA DE QUESTÕES DE APOIO PARA O TESTE

Lista de exercícios que foram utilizados como exemplo durante a execução do teste

MAC.

Exercícios de exemplo para a avaliação de comunicabilidade

Maria Sofia Jeronimo de Lima

December 2, 2025

1 Cálculo I

1.1 Derivadas

1) Para cada função $f(x)$, determine a derivada $f'(x_0)$ no ponto x_0 indicado:

- a) $f(x) = x^2$ $x_0 = 4$ c) $f(x) = x^2 - 3x$ $x_0 = 2$
b) $f(x) = 2x + 3$ $x_0 = 3$ d) $f(x) = \frac{1}{x}$ $x_0 = 2$

2) Dada a função:

$$f(x) = \begin{cases} x & \text{se } x \leq 1 \\ 2 & \text{se } x > 1 \end{cases}$$

Mostre que não existe $f'(1)$.

4) Obtenha a derivada de cada função a seguir:

- a) $f(x) = 10$ i) $f(x) = (2x^2 - 3x + 5)(2x - 1)$
b) $f(x) = x^5$ j) $f(x) = \frac{\sin x}{x^2}$
c) $f(x) = 10x^5$ k) $f(x) = \frac{2}{x^3} + \frac{5}{x^2}$
d) $f(x) = \frac{1}{2}x^2$ l) $f(x) = x^{\frac{2}{3}}$
e) $f(u) = 5u^3 - 2u^2 + 6u + 7$ m) $f(x) = 3\sqrt{x} + 5\sqrt[3]{x}$
f) $f(x) = 3 \ln x + 5$ n) $f(x) = \frac{\ln x}{\sqrt{x}}$
g) $5 \sin x + 2 \cos x - 4$
h) $f(x)x^2 \cdot \ln x$

1.2 Integrais

1) Calcule as integrais.

APÊNDICE D – ROTEIRO ENTREVISTA SEMIESTRUTURADA

Roteiro para a entrevista semiestruturada que foi realizada antes do teste MAC.

Roteiro Entrevista Pré-teste

Maria Sofia Jeronimo de Lima

December 1, 2025

1 ENTREVISTA PRÉ-TESTE

1.1 Introdução e apresentação

- Se apresentar e conhecer o entrevistado
- Entregar termo de consentimento
- Mostrar o objetivo da entrevista
 - ”Antes de começarmos as tarefas, gostaria de entender um pouco sobre seu perfil, sua experiência com tecnologia e suas expectativas sobre um chatbot para aprender matemática. Não há respostas certas ou erradas — estamos avaliando o sistema, não você.”

1.2 Experiência com as disciplinas de Cálculo

- Identificar o nível de familiaridade com a disciplina
- Identificar as principais dificuldades do estudante
- Identificar as principais formas que o estudante busca apoio durante o estudo

1.3 Experiência com ferramentas que auxiliam com o estudo e/ou ferramentas de IA

- Identificar o uso de ferramentas de apoio, quais são e sua frequência;
- Identificar o uso de Chatbots, mesmo que não diretamente com a Matemática;

1.4 Expectativas sobre o uso

- Entender o que o usuário espera que o Chatbot faça a partir do seu conceito inicial, um ”Chatbot voltado para a resolução de problemas matemáticos”
- Quais as principais funcionalidades que um aplicativo desse deve ter
- O que você espera obter como resposta ao enviar um problema matemático para um chatbot
 - Resultado, explicação, outras formas de resolver, analogias e sugestões
- Entender o que é uma explicação boa e uma explicação ruim para o usuário

1.5 Finalização

- Agradecer a participação do usuário
- Repassar o segmento do teste, com entrega de termo de consentimento específico para este.

**APÊNDICE E – ROTEIRO PARA O MÉTODO DE AVALIAÇÃO DE
COMUNICABILIDADE**

Roteiro utilizado para guiar o teste MAC.

Roteiro Método de Avaliação de Comunicabilidade

Maria Sofia Jeronimo de Lima

December 1, 2025

1 ENTREVISTA PRÉ TESTE

- Coletar dados iniciais sobre o usuário e sua experiência com as disciplinas de Cálculo e tecnologias para estudo

2 SEGMENTO DO TESTE

2.1 Tarefa 1: Pedir para o *Chatbot* explicar um conteúdo que você tem dificuldade

Objetivos

- Avaliar a interpretação do Chatbot sobre prompts não muito detalhadas;
- Avaliar a compreensão da resposta pelo usuário;

2.2 Tarefa 2: Enviar um conceito errado para o *Chatbot*

Objetivos

- Avaliar a confiabilidade do Chatbot, ou seja, ver se o sistema corrige, questiona, hesita ou confirma erroneamente.

2.3 Tarefa 3: Pedir para o *Chatbot* resolver um exercício

Objetivos

- Avaliar a resposta enviada pelo Chatbot e a sua compreensão pelo usuário

2.4 Tarefa 4: Pedir para o *Chatbot* corrigir um erro de cálculo

Objetivos

- Avaliar a correção enviada pelo Chatbot e a sua compreensão pelo usuário

2.5 Tarefa 5: Exercitar o conteúdo utilizando diferentes funcionalidades

Objetivos

- Identificar se o sistema deixa claro as outras funcionalidades que estão presentes dentro do chat
- Identificar se o usuário consegue compreender o que o sistema pode fazer com as suas outras funcionalidades

2.6 Tarefa 6: Exercitar o conteúdo utilizando a funcionalidade de quiz

Objetivos

- Identificar se o sistema comunica adequadamente sobre a funcionalidade e o que esta faz
- Identificar se o funcionamento da funcionalidade é claro para o usuário

2.7 Tarefa 7: Resolver questões utilizando a funcionalidade de leitura de PDF

Objetivos

- Identificar se a interface comunica adequadamente ao usuário como a funcionalidade deve ser utilizada;

3 QUESTIONÁRIO PÓS-TESTE E FINALIZAÇÃO

- Entregar ao usuário o questionário pós-teste;
- Identificar possíveis perguntas e/ou lacunas após o uso;
- Novamente, agradecer a participação, repassando os objetivos do teste e reforçando a privacidade sobre os dados coletados durante o processo.

**APÊNDICE F – QUESTIONÁRIO SOBRE A COMUNICABILIDADE DEO
APLICATIVO *MATHOS AI***

Questionário aplicado via Google Forms após a realização do teste MAC.

Questionário sobre a Comunicabilidade do aplicativo *Mathos AI*

Agradeço a sua participação dentro do teste de Avaliação de Comunicabilidade do aplicativo *Mathos AI*, conduzido pela estudante Maria Sofia Jeronimo de Lima, como parte da sua pesquisa para o Trabalho de Conclusão de Curso.

O objetivo deste formulário é **coletar outras impressões sobre a comunicabilidade da interface**, obtendo dados sobre suas opiniões durante o uso da aplicação.

Declaro que estou ciente de que:

- A minha participação é **voluntária**;
- Os dados fornecidos neste formulário serão utilizados **exclusivamente para fins acadêmicos**;
- Posso **desistir a qualquer momento**, sem qualquer prejuízo;
- Meus dados serão tratados de forma **confidencial** e utilizados apenas para fins de **pesquisa e desenvolvimento acadêmico**;

* Indica uma pergunta obrigatória

1. Termo de Consentimento Livre e Esclarecido (TCLE) *

Marcar apenas uma oval.

- Eu concordo com os termos aqui descritos e esclarecidos
- Não concordo com os termos

Clareza de Comunicação do Sistema

2. As mensagens exibidas pelo sistema são fáceis de entender *

Marcar apenas uma oval.

- Discordo totalmente
- Discordo parcialmente
- Nem concordo nem discordo
- Concordo parcialmente
- Concordo totalmente

3. A interface transmite claramente o que eu posso fazer em cada tela *

Marcar apenas uma oval.

- Discordo totalmente
- Discordo parcialmente
- Nem concordo nem discordo
- Concordo parcialmente
- Concordo totalmente

4. As instruções do sistema são suficientes para que eu realize as tarefas sem dúvidas *

Marcar apenas uma oval.

- Discordo totalmente
- Discordo parcialmente
- Nem concordo nem discordo
- Concordo parcialmente
- Concordo totalmente

5. A linguagem usada pelo sistema é consistente ao longo de toda a experiência *

Marcar apenas uma oval.

- Discordo totalmente
- Discordo parcialmente
- Nem concordo nem discordo
- Concordo parcialmente
- Concordo totalmente

Intenções de Design

6. Fica claro qual é a intenção do sistema em cada funcionalidade *

Marcar apenas uma oval.

- Discordo totalmente
- Discordo parcialmente
- Nem concordo nem discordo
- Concordo parcialmente
- Concordo totalmente

7. Eu entendo facilmente o que eu devo colocar como entrada dentro das funcionalidades do sistema *

Marcar apenas uma oval.

- Discordo totalmente
- Discordo parcialmente
- Nem concordo nem discordo
- Concordo parcialmente
- Concordo totalmente

8. O sistema comunica adequadamente os resultados das minhas ações *

Marcar apenas uma oval.

- Discordo totalmente
- Discordo parcialmente
- Nem concordo nem discordo
- Concordo parcialmente
- Concordo totalmente

9. As pistas de design (ícones, cores, labels) ajudam a entender como usar a interface *

Marcar apenas uma oval.

- Discordo totalmente
- Discordo parcialmente
- Nem concordo nem discordo
- Concordo parcialmente
- Concordo totalmente

Navegação

10. É fácil encontrar onde estão as principais funcionalidades *

Marcar apenas uma oval.

- Discordo totalmente
- Discordo parcialmente
- Nem concordo nem discordo
- Concordo parcialmente
- Concordo totalmente

11. Foi fácil navegar pela interface *

Marcar apenas uma oval.

- Discordo totalmente
- Discordo parcialmente
- Nem concordo nem discordo
- Concordo parcialmente
- Concordo totalmente

12. Foi muito fácil ficar confuso enquanto uso o chatbot *

Marcar apenas uma oval.

- Discordo totalmente
- Discordo parcialmente
- Nem concordo nem discordo
- Concordo parcialmente
- Concordo totalmente

13. É claro pra mim onde estou dentro do sistema a cada momento *

Marcar apenas uma oval.

- Discordo totalmente
- Discordo parcialmente
- Nem concordo nem discordo
- Concordo parcialmente
- Concordo totalmente

14. A interface ajuda a prever o que vai acontecer ao clicar em um botão ou ação *

Marcar apenas uma oval.

- Discordo totalmente
- Discordo parcialmente
- Nem concordo nem discordo
- Concordo parcialmente
- Concordo totalmente

Feedback e Coerência de Linguagem

15. O sistema oferece feedback imediato quando realizo ações *

Marcar apenas uma oval.

- Discordo totalmente
- Discordo parcialmente
- Nem concordo nem discordo
- Concordo parcialmente
- Concordo totalmente

16. Consigo entender o feedback que o sistema fornece (mensagens, alertas, destaques) *

Marcar apenas uma oval.

- Discordo totalmente
- Discordo parcialmente
- Nem concordo nem discordo
- Concordo parcialmente
- Concordo totalmente

17. Quando ocorre um erro, a interface deixa claro o que aconteceu e como resolver *

Marcar apenas uma oval.

- Discordo totalmente
- Discordo parcialmente
- Nem concordo nem discordo
- Concordo parcialmente
- Concordo totalmente

18. A linguagem usada pelo sistema é adequada e não causa confusão *

Marcar apenas uma oval.

- Discordo totalmente
- Discordo parcialmente
- Nem concordo nem discordo
- Concordo parcialmente
- Concordo totalmente

19. O sistema mantém um estilo comunicativo consistente (formal, informal, técnico, etc.) *

Marcar apenas uma oval.

- Discordo totalmente
- Discordo parcialmente
- Nem concordo nem discordo
- Concordo parcialmente
- Concordo totalmente

Impressões gerais

20. O sistema facilita a realização das tarefas, sem exigir esforços desnecessários *

Marcar apenas uma oval.

- Discordo totalmente
- Discordo parcialmente
- Nem concordo nem discordo
- Concordo parcialmente
- Concordo totalmente

21. Os elementos na tela ajudam a compreender o fluxo da atividade *

Marcar apenas uma oval.

- Discordo totalmente
- Discordo parcialmente
- Nem concordo nem discordo
- Concordo parcialmente
- Concordo totalmente

22. A interface comunica claramente como eu devo interagir com ela *

Marcar apenas uma oval.

- Discordo totalmente
- Discordo parcialmente
- Nem concordo nem discordo
- Concordo parcialmente
- Concordo totalmente

23. Senti que a comunicação do sistema reduziu a necessidade de ajuda externa *

Marcar apenas uma oval.

- Discordo totalmente
- Discordo parcialmente
- Nem concordo nem discordo
- Concordo parcialmente
- Concordo totalmente

24. No geral, a comunicabilidade do sistema é eficiente *

Marcar apenas uma oval.

- Discordo totalmente
- Discordo parcialmente
- Nem concordo nem discordo
- Concordo parcialmente
- Concordo totalmente

Este conteúdo não foi criado nem aprovado pelo Google.

Google Formulários