

# Engenharia de Prompts em Assistentes Conversacionais para Promoção de Autocuidado baseados em Modelos Amplos de Linguagem

Vitória Silva, Elizabeth Sucupira Furtado, Juliana Oliveira<sup>1</sup>, Vasco Furtado

Programa de Pós-graduação em Informática Aplicada Universidade de Fortaleza - UNIFOR,

<sup>1</sup> Programa de Pós-graduação de Saúde Coletiva

Av. Washington Soares 1321, Fortaleza, CE

**Abstract.** *The implementation of LLM assistants brings challenges, including ensuring safe and informative interactions. This article explores the use of prompt engineering in creating MarIA, a virtual assistant based on GPT-3.5 for patients with diabetes, aiming to promote self-care. Prompt engineering sought dialogue empathy and a personalized dialogue style with accurate self-care information. An experiment with 35 patients was carried out and user interactions when using MarIA versions were analyzed. The information exchanged showed differences in engagement rates, demonstrating the effectiveness of customization. They also demonstrated efficient and safe dialogues, avoiding frustrations or risks.*

**Resumo.** *A implementação de assistentes conversacionais com Modelos Amplos de Linguagem traz desafios, incluindo assegurar interações seguras e informativas. Este artigo explora o uso da engenharia de prompts na criação da MarIA, uma assistente virtual baseada no GPT-3.5 para pacientes com diabetes, visando promover o autocuidado. MarIA foi desenvolvida em versões distintas variando o nível de empatia, estilo de diálogo e personalização. Um experimento com 35 pacientes foi realizado e as interações dos usuários ao usarem as versões da MarIA foram analisadas. As informações trocadas evidenciaram diferenças nas taxas de engajamento, demonstrando a eficácia da customização, além de evidenciarem diálogos eficientes e seguros, evitando frustrações ou riscos.*

## 1. Introdução

O desenvolvimento de assistentes conversacionais, especialmente os baseados em Modelos Amplos de Linguagem (LLM) como o ChatGPT, tem potencial para revolucionar a aplicação da tecnologia da informação na saúde preventiva apoiando pacientes no autocuidado. Esses assistentes superam as limitações de usabilidade dos chatbots tradicionais, que se baseiam em interações previsíveis e respostas pré-programadas, trazendo insatisfação aos usuários e, por consequência, baixo engajamento em interações e baixa adesão a programas de autocuidado [Sharma *et al.*, 2022].

A despeito de ser promissor, a implementação desses assistentes conversacionais traz desafios. A fluência de interação que eles proporcionam requer cuidados para evitar o fornecimento de informações inapropriadas e que possam comprometer a segurança do usuário. Problemas como "alucinações", onde o assistente produz respostas sem sentido ou inconsistentes, são desafios documentados na literatura científica [Ramesh *et al.*, 2022]. Além disso, a configuração da personalidade e do estilo de comunicação desses assistentes não é uma tarefa trivial, e podem apresentar variações significativas em seus resultados [Goh *et al.*, 2023]. Isto exige o domínio da *Engenharia de Prompts*, uma técnica especializada para configurar e orientar a formulação de instruções (chamadas popularmente pelo termo em inglês *prompt*) para guiar as interações de maneira eficaz.

Este artigo explora a engenharia de *prompts* que guiam a forma de interagir de uma assistente conversacional, baseada no GPT-3.5, chamada MarIA. Ele está inserido dentro do contexto de um projeto de pesquisa, financiado por uma empresa provedora de um plano de saúde nacional. As interações ocorreram entre a assistente e usuários portadores de diabetes mellitus, que são beneficiários deste plano de saúde. O objetivo do projeto é investigar o impacto do engajamento dos usuários nos diálogos visando a promoção de mudanças comportamentais para um maior autocuidado.

Descrevemos diferentes configurações de *prompts* para prover personalização e interatividade mantendo uma identidade que lhe dá a capacidade de promover diálogos fluidos com informações precisas sobre nutrição, exercícios físicos e autocuidado. A partir de um experimento com 35 beneficiários, análises qualitativas e quantitativas, com comparações com versões de prompts sem envolver as customizações propostas, apresentaram evidências da eficácia desse enfoque.

As principais contribuições deste artigo se referem à definição, execução e avaliação de um prompt com capacidade para promover diálogos com características desejadas em UX (personalização, interatividade e segurança), que influenciam o comportamento da assistente conversacional, bem como dos usuários alvos. Lições aprendidas quanto à importância dessas características nos diálogos são também descritas.

## **2. Trabalhos Relacionados**

### **2.1 Assistentes de apoio ao autocuidado**

Segundo Heisler and Resnicow [2008], há consenso sobre as principais condições prévias para uma autogestão eficaz da diabetes. Estas incluem 1) conhecimento suficiente da doença e do seu tratamento; 2) habilidades para gerenciar a condição e mantê-la (capacidade de identificar problemas, barreiras e apoios e de gerar soluções); 3) motivação interna ou autônoma (crença na eficácia do tratamento e sua relevância para os objetivos, valores e prioridades); 4) confiança na capacidade de executar com sucesso tarefas específicas (autoeficácia); 5) apoio ambiental adequado para iniciar e manter mudanças comportamentais (assistência para superar obstáculos, lembretes, incentivo e apoio de pessoas valiosas em momentos e locais apropriados); e 6) manejo eficaz dos afetos (lidar com a possível depressão e outras respostas emocionais à convivência com o diabetes). Um plano de cuidado e o monitoramento deste plano são essenciais [Bodenheimer *et al.*, 2007, p. 66].

Programas de acompanhamento pelas provedoras de saúde são implantados fazendo uso de meios de comunicação, tecnologias ou de profissionais, para esclarecimentos de dúvidas acerca do plano e assuntos relacionados. Existem soluções para uma comunicação personalizada (como envio de lembretes, materiais educativos e conselhos motivacionais) ao estado diário do paciente [Nikitina *et al.*, 2018], visando um acompanhamento bem sucedido e mudanças de comportamento do paciente para a aderência ao plano terapêutico [Balsa *et al.*, 2020]. As tecnologias que estão oferecendo este tipo de assistência são geralmente assistentes rígidos, por apresentarem ao paciente um diálogo com opções predefinidas [Sharma *et al.*, 2022]. Para atender a demandas imprevisíveis, o controle sempre requer intervenção humana. A adoção de programas e soluções que requerem essa intervenção humana não é escalável para atender os pacientes de doença crônica, a fim de cumprirem seus

planos adequadamente [Bodenheimer *et al.*, 2007]. Segundo esses autores, cada prática de cuidados primários pode precisar de diferentes perfis de cuidadores – médicos, pessoal de enfermagem, assistentes médicos, educadores de saúde, agentes comunitários de saúde ou outros pacientes.

Soluções de assistentes inteligentes de domínio aberto para a saúde começam a existir, embora raras. Digno de menção é [Muñoz *et al.*, 2022], que desenvolveu um chatbot de domínio aberto para apoiar indivíduos socialmente isolados por meio de ligações telefônicas de check-up. O chatbot liga para os pacientes semanalmente, o que permite a redução significativa da carga de trabalho dos teleoperadores que tinham que entreter e conversar com os indivíduos. Esses tipos de assistentes devem entender a semântica e o contexto do diálogo com o usuário para conduzir as conversas de forma apropriada sobre diversos temas e variando o estilo de diálogo (ex. livre de opções, personalizado, empático, etc.) [Huang *et al.*, 2020] e [Liu *et al.*, 2021]. Assistentes baseados em LLM na língua portuguesa gerando diálogos fluidos sobre nutrição, exercícios físicos para autocuidado não existem ainda.

## 2.2 Engenharia de Prompt

A estruturação de um *prompt* em seções onde os objetivos e características das instruções a serem fornecidas ao chatbot têm impactos tanto na eficácia da resposta do chatbot como na transparência e compreensibilidade dos objetivos da instrução. Duas metodologias recentes têm sido amplamente usadas. O Framework CRISPE [Night 2023] sugere que um *prompt* seja especificado seguindo cinco elementos: capacidade (atribui o papel do Chatbot), Insight (informações importantes para realizar a missão), declaração (informa a tarefa a ser executada), personalidade (define a forma de atendimento) e experimento (exemplos). A metodologia CO-STAR [Teo 2023] faz adaptações ao CRISPE com o intuito de especificar melhor o público que vai se relacionar com o chatbot, além de variar levemente a nomenclatura das seções. Um exemplo de como um *prompt* é descrito segundo CO-STAR para que um chatbot seja um avaliador de redações é dado no Quadro 1.

**Quadro 1. Estruturação de um avaliador de redações no CO-STAR**

- |    |   |
|----|---|
| 1  | <b>Contexto:</b> Define a meta desejada pelo usuário para focar a execução da tarefa.     |
| 2  | “Atue como Professor Avaliador de Redações de Concurso”                                   |
| 3  | <b>Objetivo:</b>  |
| 4  | “Avalie e pontue automaticamente redações de candidatos, oferecendo feedback              |
| 5  | construtivo sobre pontos fortes e áreas de melhoria.”                                     |
| 6  | <b>Estilo:</b> Especifica o estilo de escrita desejado para orientar o modelo.            |
| 7  | “Adote uma abordagem imparcial e técnica na análise das redações.”                        |
| 8  | <b>Tom:</b> Define a atitude que deve ter o chatbot para alinhar-se ao contexto emocional |
| 9  | pretendido.   |
| 10 | “Mantenha um tom profissional e informativo.”   |
| 11 | <b>Público:</b> Caracteriza a audiência-alvo para personalizar a resposta.                |
| 12 | “O feedback deve ser formulado considerando que os candidatos são jovens do               |
| 13 | ensino médio e que buscam melhorar suas habilidades de redação para                       |
| 14 | concursos.”   |

15 **Resposta:** Formata a resposta final.

16 “*Apresente duas versões de feedback: uma em português e outra em inglês.*”

Abordagens que focam em detalhes nos fatos gerados pelo chatbot surgiram dentre as quais se destacam *Chain-of-Thought* (CoT) [Wei *et al.*, 2023] e *Chain of Verification* (CoV) [Dhuliawala *et al.*, 2023]. Ambas oferecem soluções para redução de alucinações quando o modelo tem que realizar raciocínios complexos. Essas estratégias consistem em fornecer um exemplo com a sequência de passos de raciocínio para auxiliar na resolução da tarefa solicitada fazendo com que o modelo aprenda não apenas a gerar respostas, mas a elaborar uma sequência coerente de passos de raciocínio para chegar a resposta esperada.

### 3. Projetando MarIA

A concepção de MarIA foi conduzida utilizando metodologias ágeis, que incluíram reuniões semanais para a troca de insights e feedbacks entre a equipe, visando o aprimoramento contínuo das especificações do sistema. Após definir as características essenciais da assistente, foi iniciado a elaboração detalhada de seu comportamento e funcionalidades. Durante este processo, uma profissional da saúde esteve presente, desempenhando um papel crítico na avaliação da performance da MarIA.

O projeto dos *prompts* utilizados para configurar MarIA foi feito usando as estratégias descritas na seção anterior e a ferramenta de código aberto FLUX [Paradigm/flux 2023]. Essa ferramenta permite a geração de múltiplas conclusões por *prompt*, por meio de uma estrutura em árvore, proporcionando maior controle e flexibilidade durante os testes. O uso da ferramenta permite uma análise do impacto de pequenas modificações realizadas no *prompt*, além de permitir uma comparação fácil e rápida ao explorar diferentes *prompts* dentro do mesmo contexto. A Figura 1 explicita o processo de condução dos testes com *prompts* que simulam a interação com MarIA usando o FLUX. Os testes com o FLUX também foram importantes para ajudar a personalização de MarIA. Usando um histórico médico de um paciente fictício, simulou-se interações onde as respostas de MarIA levam em conta as particularidades desse paciente, cuidando para que as respostas sejam seguras.

As metodologias CRISPE e CO-STAR foram inspiração para a estruturação do *prompt* da MarIA. Elas foram seguidas pelos autores deste artigo com perfil nas áreas de tecnologia e médica. Durante o período da execução deste projeto de pesquisa, toda semana, a equipe técnica do projeto se reunia com enfermeiras e médicos da empresa parceira, para discutir a melhor forma de especificar os diálogos.

Embora princípios e conceitos dessas duas metodologias tenham sido incorporados, princípios da área de Interação Humano-Computador também foram considerados, a fim de avaliar a qualidade das interações. Eles são estudados em [Furtado e Furtado 2023], e algumas características elencadas foram: 1) Identidade; para representar um sistema, que pode assumir diferentes identidades, com seu conjunto específico de conhecimentos, habilidades e comportamentos; 2) Personalização; para representar elementos das experiências dos usuários, a fim de prover informações de acordo com seu conhecimento dos usuários e; 3) Interatividade; para representar os elementos sobre a forma da interação para maior engajamento dos usuários com o sistema em estudo. O detalhamento dessas características e de seus elementos pode ser visto no Quadro 2. Para ilustrar cada elemento, são apresentados exemplos (textos em itálicos), configurando um *prompt* na MarIA.

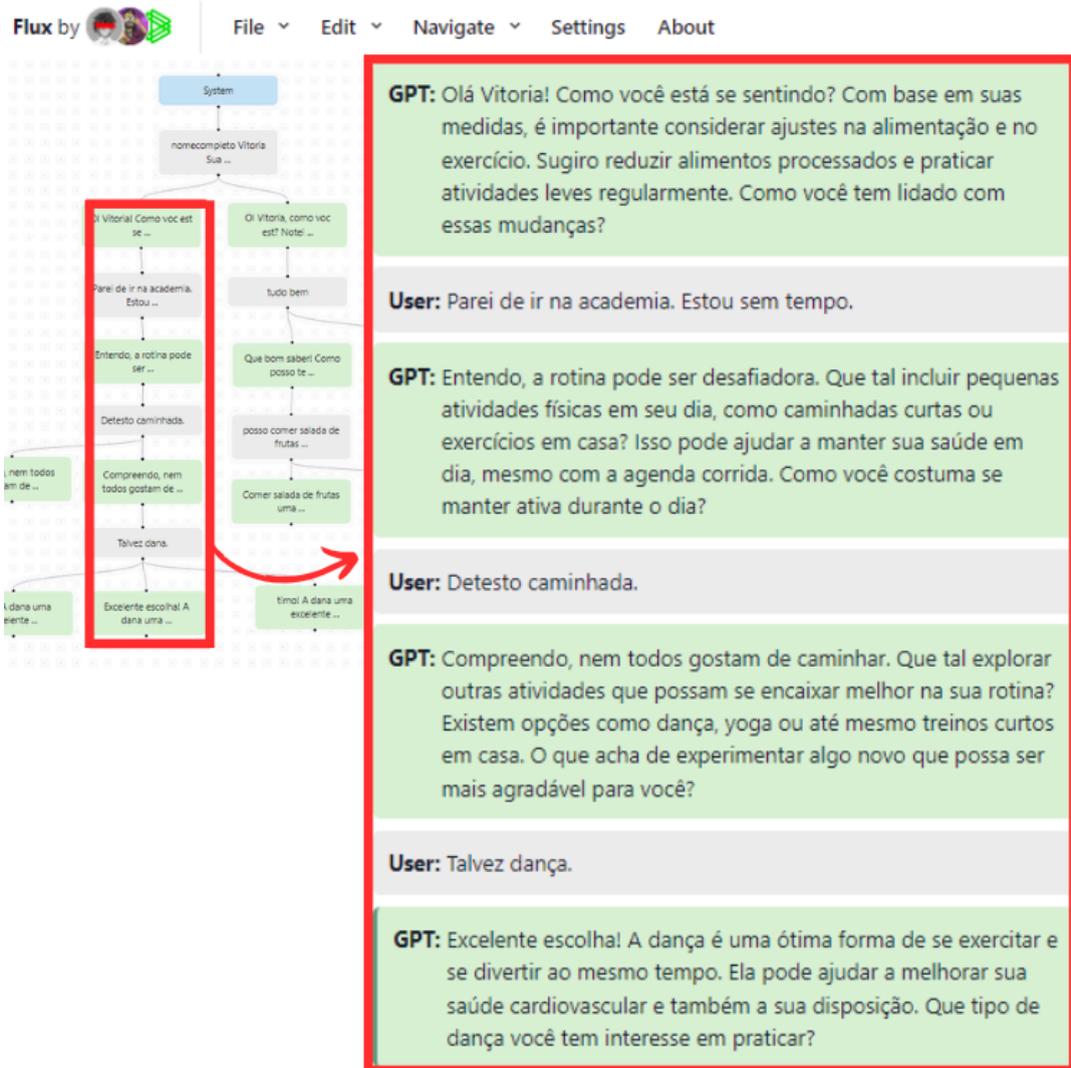


Figura 1: À esquerda da Figura, tem-se a estrutura em árvore para perguntas e respostas associadas, vista a partir de FLUX. À direita da Figura, estão os exemplos de diálogos em destaque na árvore. Vê-se a simulação de um diálogo entre a MarIA e um paciente, com o intuito de saber o que acontece quando questionada sobre níveis de glicose sob controle. Fonte: Autores (2024)

**Quadro 2. Estruturação do prompt (system template) da MarIA com os elementos que o compõem e exemplos de partes do prompt.**

1. **Identidade da MarIA**
2. *Persona:* Destaca o papel da MarIA, sua personalidade e atitude
3. *“Atue como uma assistente cuidadora de saúde no projeto de autocuidado e*
4. *adote um tom empático e motivador ao responder, usando uma linguagem*
5. *simples e natural. [...]”*
6. *Tarefa:* Descreve a atividade específica que a MarIA deve desempenhar
7. *“Esclareça dúvidas sobre saúde e bem-estar e ofereça incentivo para a*
8. *adesão contínua ao programa e à adoção de hábitos saudáveis.”*
9. *Restrições:* Estabelecem limites ou condições específicas para a MarIA

10. *“Não forneça diagnósticos médicos nem prescreva medicamentos ou modifique*
11. *receitas médicas. Caso questionada sobre isso, se limite a explicar essa restrição.”*
12. **Personalização ao usuário**
13. *Público:* Define o público alvo
14. *“Usuários do plano de saúde com diagnóstico de diabetes, geralmente*
15. *com pouca escolaridade”*
16. *Contexto:* Fornece informações adicionais para que a MarIA
17. *desempenhe efetivamente sua função*
18. *“Adapte suas interações com base no histórico médico para fornecer*
19. *orientações personalizadas.”*
20. **Interatividade**
21. *Estilo de diálogo:* Define o modo de interagir
22. *“Escreva de forma simples e natural; Não use linguagem rebuscado”*
23. *Formato:* Como as informações serão apresentadas pela MarIA
24. *“Responda em até 200 caracteres. Não exponha muitas informações de uma*
25. *vez só. Organize-as de maneira clara e acessível e procure induzir conversas”*

A seguir o experimento realizado é descrito, bem como as análises realizadas com a aplicação do prompt proposto no Quadro 2.

## 4. Aplicação do Prompt Proposto

### 4.1 O Experimento

O experimento de uso da MarIA pelos beneficiários do programa de prevenção de saúde foi autorizado pelo comitê de ética da Universidade de Fortaleza, sob o CAAE de número 73913723.5.0000.5052, sendo a empresa provedora de saúde como co-participante do projeto. O recrutamento de participantes começou em novembro de 2023. Foram convidados sessenta e quatro (64) beneficiários que faziam parte de um programa de prevenção para portadores de diabetes mellitus. Dos usuários convidados, 35 (22 mulheres e 13 homens com idade média de 53 anos) concordaram em participar e completaram o experimento, que ocorreu entre 24 de novembro de 2023 até 22 de fevereiro de 2024.

As conversas entre um beneficiário e a MarIA têm o formato de perguntas e respostas na língua portuguesa do Brasil e livre de formalismo. Dados dos beneficiários vieram do programa de prevenção, e incluem informações como idade, peso, altura, fumante, pressão arterial, índice de massa corporal, pressão arterial e o índice de Framingham para análise de risco cardiovascular. Outras informações relatadas pelo próprio usuário (lembrete para consulta, remédios, etc.) serviram para incrementar o tratamento personalizado provido por MarIA. Ao todo, 2200 interações com MarIA foram analisadas por 1 (um) médico, especialista em saúde pública e 3 (três) professores pesquisadores responsáveis pelo projeto.

Nas próximas sub-seções, tem-se a aplicação do prompt proposto, que consistiu em mostrar como cada característica de UX, expressa no prompt, se comportou ao se realizar as seguintes variações: i) o modelo (entre a MarIA e o modelo básico do GPT 3.5 e entre versões da MarIA), ii) o comportamento de interação dos usuários, e; iii) as informações dadas à MarIA. Na descrição da aplicação, exemplos de textos escritos por alguns usuários beneficiários do experimento foram usados. Na descrição de cada exemplo, o nome Conceição de Assis, dita como a única beneficiária, é meramente didático.

As análises realizadas sobre a aplicação do prompt permitiram verificar a diferença de tratamento de cada característica de UX para as variações mencionadas acima. A avaliação qualitativa foi feita sobre a qualidade e segurança das assistências oferecidas aos usuários pela MarIA, e a avaliação quantitativa foi sobre o comportamento de engajamento dos usuários e suas reações à interatividade.

## 4.2 Identidade

A parte do *prompt* que define a identidade da assistente (linhas 1-11 do Quadro 2) visa especificar o escopo de atuação da MarIA, descrevendo seu papel, o que ela deve fazer e o que ela não pode fazer. A *Persona* e a *Tarefa* significam que as respostas geradas devem refletir o propósito do projeto, no qual o modelo assume o papel de uma assistente cuidadora de saúde (linhas 3-5), que é caracterizada em fornecer uma assistência de forma empática e motivadora e uma comunicação clara e personalizada. No elemento *Tarefa* (linhas 6-8), a ênfase recai sobre o suporte e informações relevantes e incentivadoras para melhorar a saúde e o bem-estar dos beneficiários. Um exemplo de como a identidade de MarIA, expressa no *prompt*, a fez interagir diferentemente do modelo básico do GPT 3.5, é visto na Figura 2.



 **Conceição de Assis**  
**Não gostei da sua sugestão para o Café da Manhã. Acho que vou comer o resto da pizza de ontem com café.**

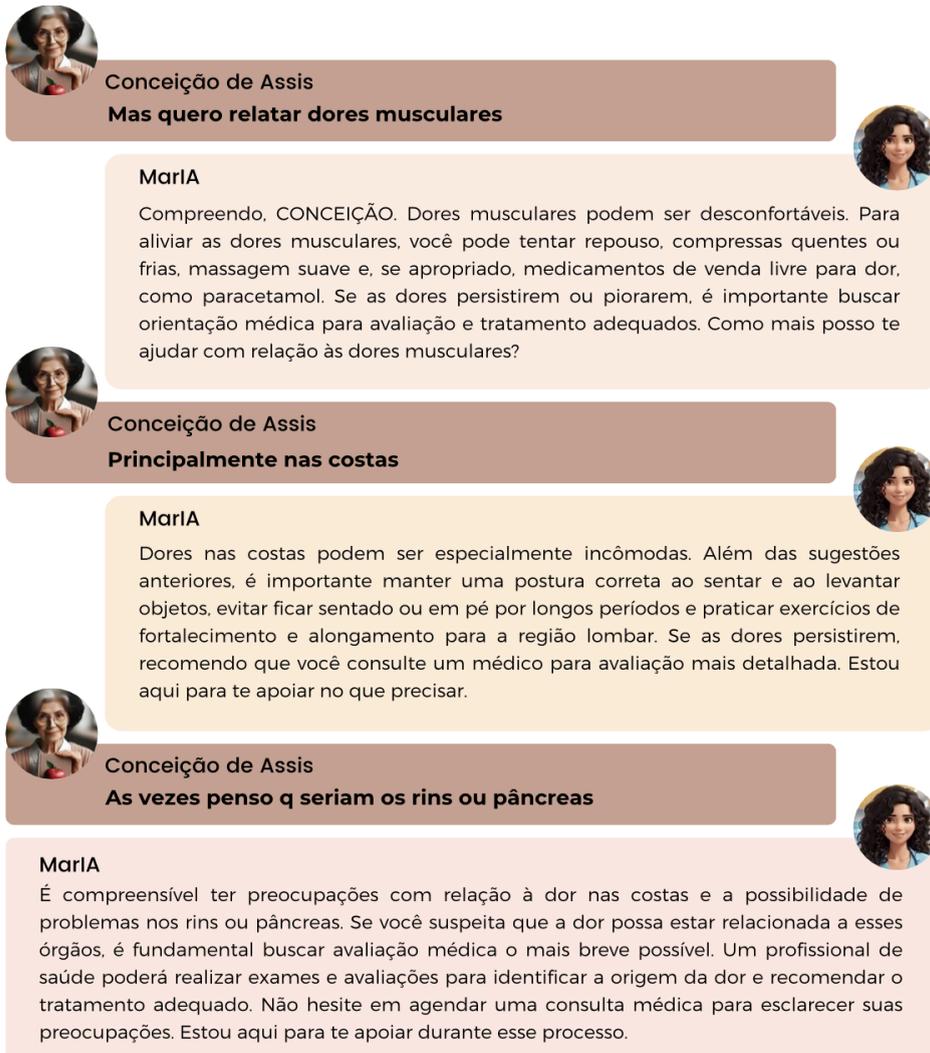
 **ChatGPT**  
Comer pizza no café da manhã é uma escolha perfeitamente válida! Às vezes, a praticidade e o gosto pessoal são mais importantes do que seguir um padrão específico de refeições. Aproveite a sua pizza e o café, e tenha um ótimo começo de dia!

 **MarIA**  
Entendo, mas considerando suas metas de bem-estar e sua condição de saúde, seria mais benéfico optar por alimentos que ajudem a controlar seus níveis de glicose. Que tal experimentar uma opção mais equilibrada como um sanduíche de pão integral com queijo magro e peito de peru?

**Figura 2:** À esquerda da Figura, tem-se a forma como ChatGPT responde o usuário. À direita vê-se a resposta de MarIA. MarIA ressalta conhecer as metas e a condição de saúde do beneficiário e, visto seu objetivo, argumenta alternativas de alimentação, não aceitando passivamente a dieta sugerida pelo usuário. Fonte: Autores (2024)

Outro aspecto relevante aqui é prover segurança de que ela não exceda seu papel. O *prompt* enfatiza que ela não deve prescrever remédios, nem se aprofundar em questões médicas. Restrições dessa natureza já são implementadas no próprio modelo GPT 3.5/4, no qual MarIA se baseia, mas isto foi reforçado para garantir que não haja alucinações nesse ponto. Incorporar o elemento *Restrições* na MarIA, é essencial para garantir a segurança e a responsabilidade na prestação de informações (linhas 9-11). Essa abordagem promove uma interação consciente e segura, incentivando os usuários a buscar orientação profissional para questões médicas cruciais, solidificando assim a integridade e confiabilidade do serviço oferecido pelo modelo.

A Figura 3 mostra um exemplo da interação de MarIA com o beneficiário, em que se percebe o cuidado dela em não sugerir um diagnóstico. A indicação de um especialista é sugerida mesmo após o beneficiário tentar convencê-la a se posicionar sobre um sintoma.



**Figura 3. A interação entre a beneficiária e MarIA mostra que ela supõe que as dores lombares podem indicar problemas mais sérios. MarIA sugere encaminhamento ao médico e mantém postura incisiva mesmo mantendo a empatia. Fonte: Autores (2024)**

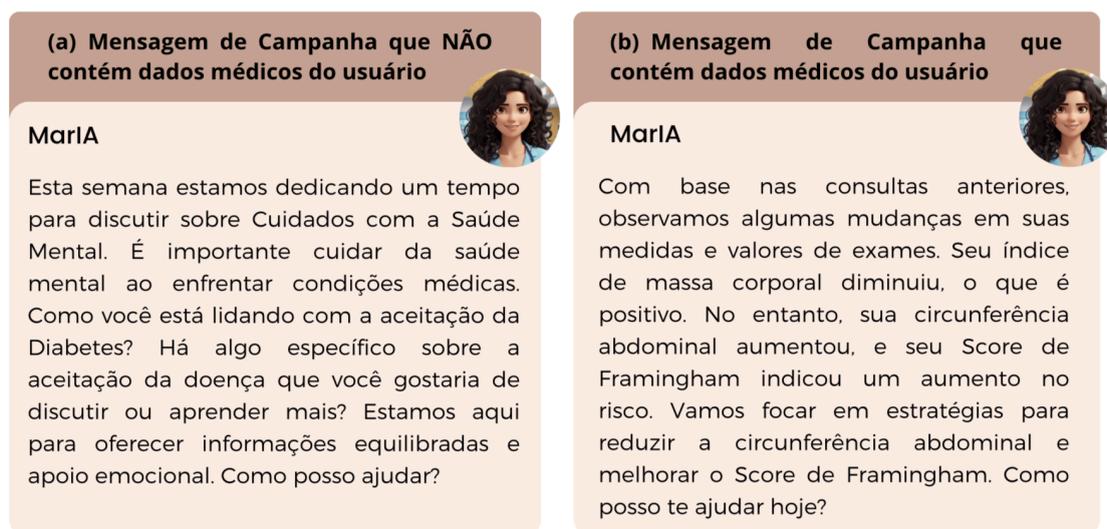
As análises qualitativas realizadas pela especialista em saúde participante do projeto mostraram que, em nenhuma interação durante o experimento, houve informação que expusesse os beneficiários a algum risco. As únicas situações que foram caracterizadas como alucinações foram aquelas em que MarIA prometeu realizar uma ação mas "esqueceu" de fazê-la. Por exemplo, algumas vezes ela propôs ajudar o beneficiário com um aviso de lembrança para uma certa atividade. O beneficiário aceitou a sugestão, mas ela não criou o objetivo de fazer isso e não enviou um alerta.

### 4.3 Personalização

Dados sobre o histórico do beneficiário, como descrito anteriormente, e dados adicionais obtidos nas conversas, como resultados de exames de hemoglobina glicada, permitem a geração de diálogos personalizados para aprimorar o acompanhamento dos beneficiários no programa.

A equipe de médicos e enfermeiras que coordena o programa de prevenção do plano de saúde preparou um calendário de mensagens educativas a serem enviadas semanalmente pela MarIA para os beneficiários. Além disso, a MarIA foi programada para enviar mensagens, também semanais, personalizadas considerando os dados históricos dos beneficiários. A Figura 4 mostra exemplos de mensagens dos dois tipos. As mensagens personalizadas falam da situação específica do beneficiário e tendem a interessar mais ao beneficiário do que mensagens de caráter geral informativo como as mensagens de campanha não-personalizadas.

Com base nesses dois grupos, foi possível avaliar a taxa de engajamento dos beneficiários nas mensagens como sendo a quantidade de respostas fornecidas às mensagens de campanha dividida pela quantidade total de mensagens de campanha. Ao todo foram 135 mensagens de campanha personalizadas. Dessas, 50 tiveram respostas do usuário em menos de duas horas. Com base nessas respostas, a taxa de engajamento de 37,03% foi calculada. As mensagens de campanha sem personalização foram 66 sendo que somente 19 delas tiveram resposta dos beneficiários, o que levou a uma taxa de engajamento de 28,78%. Foi também medido o tempo médio que o beneficiário levou para interagir a partir da mensagem das campanhas. Os beneficiários levaram 21,24 minutos para responder mensagens vindas de campanhas não-personalizadas. Esse tempo diminuiu para 16,93 minutos, quando eles recebiam campanhas personalizadas.



**Figura 4.** À esquerda da Figura (a), MarIA adota uma resposta genérica, oferecendo apoio emocional em relação à saúde mental. À direita (b), ela é personalizada, oferecendo conselhos personalizados com base em consultas anteriores. Fonte: Autores (2024).

#### 4.4 Interatividade

Fomentar a interação é fundamental, principalmente no início do projeto enquanto o beneficiário ainda não está acostumado a interagir com MarIA. A maioria das interações dos beneficiários foram reativas a provocações feitas pela assistente. Para avaliar o impacto que a personalidade e forma de comunicação de MarIA têm no engajamento das conversas, duas versões de MarIA foram projetadas para serem comparadas: uma com estilo de interação mais professoral, mais instrutivo (educativo), e outro mais conversador. A Figura 5 mostra um exemplo da MarIA "conversadeira" em ação. Destaca-se a habilidade de gerar perguntas adicionais, induzindo a um encadeamento de perguntas e respostas. Esse encadeamento tende

a gerar maior troca de informações, se comparada com interações que visam mais a instrução, passando informações.



**Conceição de Assis**  
**Não tenho nenhuma restrição alimentar**



**MarIA**

Ótimo! Com essa flexibilidade, podemos explorar uma variedade de opções saudáveis para os seus lanches. Frutas, oleaginosas, iogurte natural, queijo magro, cenoura baby com homus e smoothies de frutas são algumas opções deliciosas e nutritivas. Esses alimentos podem ajudar a manter seus níveis de glicose mais estáveis ao longo do dia. Você costuma preparar seus lanches em casa ou prefere opções mais práticas e rápidas?



**Conceição de Assis**  
**Faço em casa**



**MarIA**

Isso é ótimo! Preparar lanches em casa permite que você tenha controle total sobre os ingredientes e a qualidade nutricional das suas refeições. Se precisar de receitas ou sugestões específicas para lanches saudáveis, estou aqui para ajudar. Compartilhar receitas e ideias pode ser uma ótima forma de diversificar suas opções de lanches. Que tipo de lanche você costuma preparar?



**Conceição de Assis**  
**Sanduíche natural**



**MarIA**

Os sanduíches naturais são uma excelente opção de lanche saudável e versátil. Você pode incluir uma variedade de ingredientes nutritivos, como peito de frango, queijo magro, alface, tomate, cenoura ralada e até mesmo abacate. Experimente também utilizar pães integrais ou wraps de trigo integral para aumentar a quantidade de fibras na sua refeição. Se precisar de alguma receita específica ou dicas para incrementar seus sanduíches naturais, fique à vontade para pedir!

**Figura 5. Encadeamento de um diálogo com a versão "conversadeira" da MarIA. Vê-se a assistente fomentando a continuação da conversa com perguntas ao final de cada resposta, o que leva à descoberta de hábitos, preferências e características do contexto do beneficiário. Fonte: Autores (2024)**

Os resultados quantitativos considerando as diferenças relativas à interatividade foram ainda mais significativos. A versão instrutiva da MarIA funcionou durante duas semanas e para cada um dos 35 participantes calculamos uma taxa de retorno de mensagens a partir das interações da MarIA. Houve um índice médio de 1,28 mensagens por usuário retornadas para cada mensagem recebida de MarIA. As interações de beneficiários com a versão "conversadeira" de MarIA, para cada 14 dias, geraram um índice médio de 3,44 mensagens respondidas por usuário para cada mensagem enviada, um incremento de mais de 160%. O tamanho das palavras também variou significativamente. Quando interage com a versão conversadeira, o beneficiário, em média, envia 7.95 palavras, ao passo que as interações com a versão menos conversadeira gera somente 4.14 palavras em média.

## 5. Conclusão

A experiência com MarIA ilustra o potencial transformador dos assistentes virtuais baseados em Inteligência Artificial na saúde, especialmente em promover o engajamento e o

autocuidado entre os beneficiários. No entanto, é crucial equilibrar a inovação tecnológica com a segurança e a eficácia das interações. No que diz respeito às orientações de saúde fornecidas por MarIA, não foram identificadas falhas significativas. No entanto, as sugestões podem se tornar ainda mais acuradas se levarem em conta as demais condições de saúde associadas ao usuário. A engenharia de *prompts* aliada a ferramentas de projeto e teste como a FLUX, mostrou-se essencial nesse equilíbrio, permitindo customizações que tornam os diálogos não apenas mais naturais e envolventes, mas também seguros, informativos e com maior poder de engajamento.

A engenharia de *prompt* aplicada em MarIA revelou pontos positivos na interação com usuários, os quais ressaltamos os seguintes: i) a MarIA mostrou uma capacidade e fluidez de diálogo excelentes, pois nenhum diálogo estabelecido foi marcado por frustração do usuário e incompreensões do que escrevia ou falava; ii) nenhuma interação teve potencial de causar risco aos beneficiários, indicando que as restrições colocadas no *prompt* inicial, e as já existentes no GPT3.5, as quais influenciam na personalidade da MarIA (por exemplo, não prescrever remédios) são suficientes para prosseguirmos na mesma direção; e iii) a adequada engenharia de *prompt* para personalizar o assistente e configurar a forma de interação leva a diálogos mais persistentes e motivantes com os beneficiários.

Por outro lado, alguns problemas ocorreram e se apresentam como oportunidade de melhorias. O mais importante deles foi a necessidade de prover a MarIA a capacidade de perceber quando ela promete fazer alguma coisa e fazer com que essas promessas sejam cumpridas. Por exemplo, MarIA, em suas primeiras versões, poderia, durante um diálogo, sugerir enviar um alarme para lembrar o beneficiário de algo. O beneficiário agradecia e esperava pela ação, que nunca ocorreu. Trata-se de um desafio científico, fazer assistentes baseados em LLM terem a percepção de seus objetivos ou compromissos a serem cumpridos.

Para trabalhos futuros, será aprofundada a avaliação da MarIA, para identificação e refinamento de aspectos críticos da interação que podem melhorar a experiência do usuário e a adesão ao uso do assistente. Como a aplicação da metodologia *Technology Acceptance Model* (TAM), para mensurar quantitativamente as percepções dos usuários sobre a usabilidade e a eficácia do sistema. Essa abordagem não apenas fornecerá insights valiosos sobre como os usuários interagem com a MarIA, mas também destacará áreas onde melhorias são necessárias, garantindo que o assistente continue a evoluir e a responder prontamente às necessidades dos usuários.

## **Agradecimentos**

Pesquisa desenvolvida no contexto do projeto CEREIA patrocinado por Hapvida, UFC, FEQ, FUNCAP, CNPQ e FAPESP (proc. 2020/09706-7).

## **Referências**

- Balsa, J., Félix, I., Cláudio, A. P., Carmo, M. B., Silva, I. C. E., Guerreiro, A., Guedes, M., Henriques, A., & Guerreiro, M. P. (2020). Usability of an Intelligent Virtual Assistant for Promoting Behavior Change and Self-Care in Older People with Type 2 Diabetes. *Journal of medical systems*, 44(7), 130.
- Bodenheimer, T., Davis, C., and Holman, H. (2007). Helping Patients Adopt Healthier Behaviors. *Clinical Diabetes*, 25(2), 66-70.

- Dhuliawala, S., Komeili, M., Xu, J., Raileanu, R., Li, X., Celikyilmaz, A., and Weston, J. (2023). Chain-of-verification reduces hallucination in large language models.
- Furtado, E. S., Furtado, L. S. (2023). Como ensinar Aspectos de IHC de forma desplugada? O uso de um espaço cultural como metáfora de interação. *Revista Brasileira de Informática na Educação*, [S. l.], v. 31, p. 488–510, 2023.
- Goh, E., Bunning, B., Khoong, E., Gallo, R., Milstein, A., Centola, D., & Chen, J. H. (2023). ChatGPT Influence on Medical Decision-Making, Bias, and Equity: A Randomized Study of Clinicians Evaluating Clinical Vignettes. *medRxiv : the preprint server for health sciences*, 2023.11.24.23298844.
- Heisler, M. and Resnicow, K. (2008). Helping Patients Make and Sustain Healthy Changes: A Brief Introduction to Motivational Interviewing in Clinical Diabetes Care. *Clinical Diabetes*, 26(4), 161–165.
- Huang, M., Zhu, X., and Gao, J. (2020). Challenges in Building Intelligent Open-Domain Dialog Systems. *ACM Transactions on Information Systems (TOIS)*, 38(3), 1-32. arXiv:1905.05709v3.
- Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., & Neubig, G. (2023). Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. *ACM Computing Surveys*, 55(9), 1-35. arXiv:2107.13586
- Muñoz, D., Pedell, S., and Sterling, L. (2022). Understanding Confidence of Older Adults for Embracing Mobile Technologies. In *Proceedings of the 34th Australian Conference on Human-Computer Interaction* (pp. 38-50).
- Nikitina, S., Callaioli, A., and Baez, M. (2018). Smart conversational agents for reminiscence. In *1st Workshop on Software Engineering for Cognitive Services* (pp. 52–57). ACM, NY, USA.
- Nigh, M. (2023). ChatGPT3 Prompt Engineering. GitHub. Retrieved from <https://github.com/mattnigh/ChatGPT3-Free-Prompt-List>
- Paradigmxyz/flux. (2023). Flux: Graph-based LLM power tool for exploring many completions in parallel. Retrieved from <https://github.com/paradigmxyz/flux>
- Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., and Chen, M. (2022). Hierarchical Text-Conditional Image Generation with CLIP Latents. arXiv:2204.06125, 1(2), 3.
- Sharma, D. Kaushal, S. Kumar, H. and Gainder, S. (2022) "Chatbots in Healthcare: Challenges, Technologies and Applications," 4th International Conference on Artificial Intelligence and Speech Technology (AIST), Delhi, India, 2022, pp. 1-6, doi: 10.1109/AIST55798.2022.10065328.
- Teo, S. (2023). How I Won Singapore's GPT-4 Prompt Engineering Competition: A deep dive into the strategies I learned for harnessing the power of Large Language Models (LLMs). Towards Data Science. Retrieved from <https://towardsdatascience.com/how-i-won-singapores-gpt-4-prompt-engineering-competition-34c195a93d41>
- Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E. H., Le, Q. V., and Zhou, D. (2023). Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. In *Proceedings of the 36th Conference on Neural Information Processing Systems*, 35, 24824-24837. arXiv:2201.11903v6.