

OTIMIZAÇÃO DA COLETA DE DADOS CLÍNICOS NA ATENÇÃO PRIMÁRIA: UMA ANÁLISE DO USO DE INTELIGÊNCIA ARTIFICIAL NA TRIAGEM MÉDICA

Nathan dos Santos Guth¹, Erni Feix Junior², Guilherme Turatti³,
Edson Moacir Ahlert⁴

Resumo: A Atenção Primária à Saúde enfrenta desafios estruturais que comprometem a eficiência do atendimento, incluindo sobrecarga dos serviços, demora na consulta e limitação de tempo para coleta de informações clínicas básicas. Nesse cenário, ferramentas baseadas em Inteligência Artificial têm se destacado como alternativas para otimizar fluxos de trabalho e aprimorar a triagem inicial. Este estudo analisa o uso de modelos de linguagem natural na coleta automatizada de dados clínicos, por meio do desenvolvimento e avaliação de um protótipo de teletriagem conversacional construído com a API Gemini. A pesquisa, de caráter aplicado, exploratório e descritivo, utiliza princípios de *Design Science Research* para estruturar e testar um artefato em ambiente controlado. O protótipo foi avaliado utilizando vinhetas clínicas padronizadas, o que permitiu examinar a coerência do diálogo, a completude da extração de informações e a presença de hipóteses diagnósticas compatíveis com a literatura. Os resultados parciais indicam que a abordagem é tecnicamente viável para organizar informações clínicas e reduzir o tempo destinado à anamnese básica, embora apresente limitações, como risco inicial de generalizações sem evidência e ausência de validação em condições reais de atendimento. Os achados parciais sugerem que sistemas de teletriagem baseados em IA podem apoiar a triagem médica ao otimizar a coleta inicial de dados e favorecer práticas preventivas, desde que implementados com salvaguardas éticas, supervisão profissional e validação em ambiente clínico controlado.

Palavras-chave: atenção primária à saúde; inteligência artificial; triagem médica; prevenção em saúde.

¹ Estudante de Engenharia de Software – Univates - nathan.guth@universo.univates.br

² Estudante de Engenharia da Computação - Univates - erni.feix@universo.univates.br

³ Estudante de Engenharia da Computação - Univates - guilherme.turatti@universo.univates.br

⁴ Professor da Universidade do Vale do Taquari - Univates.

OPTIMIZATION OF CLINICAL DATA COLLECTION IN PRIMARY CARE: AN ANALYSIS OF THE USE OF ARTIFICIAL INTELLIGENCE IN MEDICAL TRIAGE

Abstract: Primary Health Care faces structural challenges that compromise service efficiency, including system overload, long waiting times and limited time for collecting basic clinical information. In this context, Artificial Intelligence tools have emerged as alternatives to optimize clinical workflows and improve initial triage. This study analyzes the use of natural language models for automated clinical data collection through the development and evaluation of a conversational tele-triage prototype built using the Gemini API. The research, characterized as applied, exploratory and descriptive, adopts principles of Design Science Research to structure and test an artifact in a controlled environment. The prototype was evaluated using standardized clinical vignettes, which enabled the examination of dialogue coherence, completeness of data extraction and the presence of diagnostic hypotheses aligned with the literature. The partial results indicate that the approach is technically feasible for organizing clinical information and reducing the time spent on basic anamnesis, although limitations remain, such as the initial risk of unsupported generalizations and the absence of validation in real-world clinical settings. The partial findings suggest that AI-based tele-triage systems may support medical triage by optimizing initial data collection and promoting preventive practices, provided they are implemented with ethical safeguards, professional supervision and validation in controlled clinical environments.

Keywords: primary health care; artificial intelligence; medical triage; health prevention.

1 INTRODUÇÃO

A Atenção Primária à Saúde (APS) é reconhecida como a base dos sistemas de saúde e o nível responsável por garantir o acesso e a integralidade do cuidado (Torres; Wermelinger; Ferreira, 2025). Contudo, a realidade operacional impõe barreiras significativas à sua efetividade. Desafios estruturais, como a sobrecarga dos serviços e a escassez de recursos, geram um cenário de “gargalo” no atendimento, caracterizado por longas filas de espera e tempo reduzido para a interação individual com o paciente (Vital Strategies; Umane; UFPel, 2025).

Um dos pontos críticos dessa dinâmica é o tempo consumido pela coleta de informações clínicas básicas (anamnese). Grande parte da consulta é dedicada à estruturação de queixas e histórico, reduzindo a disponibilidade do profissional para o exame físico, o raciocínio clínico complexo e a orientação preventiva (Lobo, 2018). Esse desequilíbrio favorece um modelo reativo, focado na doença já instalada, em detrimento da promoção da saúde (Figueiredo-e-Silva, 2025).

Diante desse contexto, a transformação digital surge não apenas como modernização, mas como necessidade estratégica. Ferramentas baseadas em Inteligência Artificial (IA), especificamente os Grandes Modelos de Linguagem (LLMs), apresentam potencial para automatizar a triagem inicial e

a estruturação de dados, oferecendo suporte à decisão e otimizando o fluxo de trabalho (Faceli *et al.*, 2025; Lee; Goldberg; Kohane, 2024).

A automação da teletriagem, processo de avaliação remota de sintomas, possibilita que dados preliminares sejam coletados e organizados antes mesmo do contato presencial. No entanto, a implementação dessas tecnologias exige validação rigorosa quanto à precisão, segurança e ética (NIC.br, 2024).

Considerando essas premissas, este artigo tem como objetivo geral analisar como soluções digitais baseadas em modelos de linguagem podem otimizar a coleta de informações clínicas na APS, servindo como instrumento de apoio à triagem e estímulo à prevenção. Para alcançar este propósito, a pesquisa desdobra-se em objetivos específicos que incluem o desenvolvimento de uma Prova de Conceito (PoC) de teletriagem conversacional utilizando a API Gemini, bem como a avaliação da capacidade do artefato em conduzir diálogos coerentes e extrair dados clínicos a partir de vinhetas padronizadas.

Busca-se, ainda, verificar a consistência das hipóteses de triagem geradas pelo sistema em comparação com a literatura médica e discutir a viabilidade técnica da ferramenta para a redução do tempo de anamnese e a organização de fluxos na atenção primária.

O trabalho está estruturado em cinco seções. Após esta introdução, o referencial teórico discute os fundamentos da IA na saúde e os modelos de otimização. A seção de metodologia detalha a construção da PoC e os critérios de teste. Os resultados parciais apresentam o desempenho do protótipo frente a cenários clínicos simulados, seguidos pelas considerações finais sobre as implicações e limitações do estudo.

2 REFERENCIAL TEÓRICO

Como destaca Bill Gates, “o desenvolvimento da IA é tão fundamental quanto a criação do computador pessoal. Ela mudará a forma como as pessoas trabalham, aprendem e se comunicam, e transformará a assistência médica” (Lee; Goldberg; Kohane, 2024).

Essa transformação já se reflete em ferramentas contemporâneas, como o *Ada Health* (2025), que aplicam algoritmos de aprendizado de máquina para interpretar sintomas e gerar orientações iniciais ao paciente. Pesquisas indicam que a IA pode automatizar a coleta e análise de dados clínicos, apoiar decisões diagnósticas e reduzir erros, proporcionando maior eficiência e segurança no atendimento (Li, 2023; Magalhães *et al.*, 2024).

No entanto, a adoção de sistemas de triagem baseados em IA requer cuidados éticos e técnicos, como a validação dos algoritmos, a transparência nos critérios de decisão e a proteção dos dados pessoais (Faceli *et al.*, 2025; Figueiredo-e-Silva, 2025).

2.1 Conceitos e História da Inteligência Artificial na Saúde

O termo Inteligência Artificial (IA), cunhado na proposta de 1955 para o projeto de pesquisa de Dartmouth realizado em 1956 (McCarthy *et al.*, 2006), foi definido por seu organizador, John McCarthy (2007, p. 2), como “a ciência e engenharia de criar máquinas inteligentes, especialmente programas de computador inteligentes.”. Na saúde, a IA simula a capacidade humana de combinar conhecimentos prévios e uso prático desses conhecimentos, tal como um médico que realiza diagnósticos com base em sintomas, resultados de exames e experiências profissionais (Faceli, 2025).

O uso da Inteligência Artificial (IA) na área da saúde tem origem há mais de cinco décadas, inicialmente voltado para o desenvolvimento de sistemas de suporte à decisão diagnóstica, criados para auxiliar médicos na interpretação de dados clínicos e na formulação de hipóteses diagnósticas (Figueiredo-e-Silva, 2025; Fraser *et al.*, 2023). Esses sistemas buscavam estruturar o raciocínio clínico humano em regras lógicas capazes de gerar recomendações automatizadas, servindo como ferramentas de apoio e não de substituição da prática médica.

Com o avanço tecnológico, especialmente nas últimas duas décadas, esses sistemas passaram por uma ampliação de escopo, evoluindo do uso exclusivo por profissionais de saúde para o acesso direto de pacientes, por meio de plataformas conhecidas como *symptom checkers* (SCs). Esses aplicativos permitem que o próprio usuário insira seus sintomas e receba uma avaliação inicial automatizada, representando um marco na democratização do suporte diagnóstico digital (Fraser *et al.*, 2023).

Os *symptom checkers* mais consolidados, como o *Ada Health*, baseiam-se em redes Bayesianas, que utilizam probabilidades derivadas de dados clínicos e literatura científica para inferir possíveis condições médicas (Fraser *et al.*, 2023). Embora precisos em muitos casos, esses sistemas são determinísticos, ou seja, seguem regras estáticas e não aprendem com novas informações, diferentemente dos modelos modernos de aprendizado de máquina.

Dentro da IA, o Aprendizado de Máquina (AM) emergiu como uma subárea proeminente, investigando técnicas que permitem a aquisição automática de conhecimento a partir de dados (Faceli, 2025). Os modelos de IA, que são representações matemáticas derivadas de algoritmos de AM, capturam padrões nos dados e são integrados em sistemas completos (Faceli, 2025).

Algoritmos para treinar Redes Neurais Artificiais (RNAs) profundas têm alcançado desempenho de ponta em aplicações complexas, como processamento de linguagem natural, podendo ainda haver a utilização de sensores de monitoramento de saúde que coletam e transmitem informações como batimentos cardíacos ou glicose (Faceli, 2025).

O estudo comparativo entre *symptom checkers* e o ChatGPT evidencia uma transição importante no campo da IA médica: enquanto os primeiros dependem de bases de conhecimento fixas e estruturadas, os segundos incorporam

aprendizado dinâmico e contextual, aproximando-se da interpretação clínica humana (Fraser *et al.*, 2023).

Essa evolução tecnológica reforça a importância da IA não apenas como ferramenta de automação, mas como instrumento de suporte cognitivo para médicos e pacientes, integrando o raciocínio probabilístico tradicional com a análise contextual e semântica possibilitada pelos modelos modernos.

2.2 A Otimização de Processos e a Eficiência na Atenção Primária

A Atenção Primária à Saúde (APS) é reconhecida como o primeiro nível de atenção e um pilar fundamental da assistência à saúde, sendo responsável pelo gerenciamento do maior número de pacientes no sistema (Torres; Wermelinger; Ferreira, 2025). No entanto, a APS enfrenta desafios estruturais e organizacionais significativos, incluindo sobrecarga e lentidão no atendimento, falta de profissionais qualificados e longas filas de espera, que comprometem a efetividade e a resolutividade do cuidado (Giongo, 2015; Mueller *et al.*, 2021).

Um estudo realizado pela Vital Strategies e Umane (2025), em parceria com a Universidade Federal de Pelotas (UFPel), reforça esse cenário ao apontar que mais de 62% dos brasileiros que necessitam de atendimento na APS não o buscam, citando superlotação, demora no atendimento e burocracia como principais fatores.

O uso de tecnologias digitais, especialmente a IA, é apontado como estratégia promissora para otimizar o desempenho do profissional de saúde, aprimorar a gestão e aumentar a eficiência dos serviços (Lobo, 2018; NIC.br, 2024). Entre as principais áreas de aplicação, destacam-se:

2.2.1 Melhoria da Gestão, Fluxos de Trabalho e Alocação de Recursos

A IA permite prever a demanda por serviços de saúde, otimizando a alocação de recursos humanos e materiais em clínicas e hospitais, promovendo maior eficiência operacional (NIC.br, 2024). Além disso, algoritmos inteligentes podem organizar fluxos de atendimento, otimizar o uso de equipamentos, gerenciar o giro de leitos e aprimorar a logística de suprimentos, resultando em processos mais ágeis e eficazes (NIC.br, 2024).

Estudos demonstram que a aplicação de Sistemas Especialistas (SE) em serviços públicos, como no serviço odontológico da ESF Macedo, permite priorizar atendimentos e gerenciar agendamentos, substituindo sistemas tradicionais de filas por mecanismos baseados em acolhimento e classificação de risco, melhorando a eficiência do serviço (Giongo, 2015).

Outras iniciativas incluem o desenvolvimento de algoritmos para organização de filas e o uso de *Robotic Process Automation* (RPA) em tarefas operacionais repetitivas, aumentando a produtividade do serviço (NIC.br, 2024).

2.2.2 Otimização do Tempo do Profissional e Aumento da Produtividade

A IA também tem grande potencial para otimizar o tempo dos profissionais de saúde, permitindo que estes se concentrem em atividades de cuidado humanizado (NIC.br, 2024; Torres; Wermelinger; Ferreira, 2025). Ferramentas como reconhecimento de voz para transcrição de laudos radiológicos ou sistemas de IA aplicados a prontuários eletrônicos podem digerir grandes volumes de dados clínicos, sintetizando informações relevantes e gerando insights que auxiliam na tomada de decisão clínica (NIC.br, 2024).

Esse aumento de produtividade contribui para uma maior eficiência nos atendimentos e melhor qualidade do cuidado, reduzindo o tempo gasto em tarefas burocráticas e administrativas e permitindo que os profissionais dediquem mais atenção à escuta ativa, orientação preventiva e acompanhamento do paciente (Lobo, 2024; NIC.br, 2024).

2.3 A Teletriagem Conversacional: Coleta de Dados e Estímulo à Prevenção

A utilização de triagem remota já é uma prática consolidada há décadas. O *Manchester Triage Group* (Mackway-Jones; Marsden; Windle, 2014) destaca que a triagem telefônica é utilizada desde 1991, permitindo a orientação inicial do paciente sem necessidade de deslocamento imediato até unidades de saúde. Com o avanço tecnológico, é natural que essa prática evolua, incorporando ferramentas de Inteligência Artificial.

Segundo o livro *Emergency Triage* do *Manchester Triage Group* (Mackway-Jones; Marsden; Windle, 2014), amplamente adotado em diversos países como referência para a definição das regras de triagem, o princípio fundamental é que os pacientes devem ser atendidos conforme o grau de prioridade clínica, e não pela ordem de chegada. Nesse sentido, a aplicação da IA surge como uma alternativa eficaz para automatizar esse processo, permitindo que, por meio de uma conversa natural com um sistema inteligente, seja possível identificar o nível de urgência do paciente sem a necessidade de aguardar avaliação inicial de um profissional (NIC.br, 2024).

A automação da triagem, impulsionada pelo Processamento de Linguagem Natural (PLN) e pelos modelos avançados de IA, é o mecanismo que torna viável a coleta de informações clínicas de forma assíncrona e humanizada. O PLN permite que sistemas inteligentes compreendam a linguagem natural do paciente, extraíndo sintomas, histórico e fatores de risco em um diálogo conversacional (Fraser *et al.*, 2023). Essa abordagem gera um resumo estruturado de dados que, ao ser entregue ao profissional da APS antes da consulta, pode reduzir drasticamente o tempo gasto com perguntas básicas.

Mais do que eficiência, o principal potencial dessas ferramentas reside no estímulo à prevenção, um dos focos deste artigo. Uma vez que o sistema identifica fatores de risco ou hábitos do paciente durante a triagem, ele pode automaticamente indicar o nível de urgência e fornecer orientações

personalizadas e educativas, como dicas de saúde, sugestões de exames preventivos e consultas especializadas, antes mesmo do contato médico e sem sair de casa. Assim, a IA transforma o momento da triagem de um mero filtro de urgência para um ponto de contato estratégico para a promoção ativa da saúde (Faceli *et al.*, 2025; Magalhães *et al.*, 2024), sendo muito útil em situações com alto contágio e pandemias (Palacios, 2022).

3 PROCEDIMENTOS METODOLÓGICOS

Este capítulo detalha os procedimentos metodológicos adotados para a condução da pesquisa, que visa analisar a otimização da coleta de dados clínicos na Atenção Primária à Saúde por meio de uma ferramenta de Inteligência Artificial. A metodologia foi estruturada para permitir a exploração de uma solução tecnológica emergente em um contexto prático, culminando no desenvolvimento e na avaliação de um artefato de prova de conceito.

A definição desses procedimentos está alinhada aos objetivos do estudo. O objetivo geral é avaliar se um modelo de linguagem (LLM) pode melhorar a coleta e a organização de dados clínicos na triagem da APS. Para isso, os objetivos específicos incluem: (i) desenvolver uma PoC utilizando a API Gemini; (ii) mapear e analisar o processo atual de coleta de dados clínicos; (iii) testar o artefato com vinhetas clínicas para medir desempenho, acurácia e consistência; e (iv) validar a qualidade das triagens geradas por meio do feedback de estudantes de medicina.

Cada etapa descrita neste capítulo, uso de *Scrum* no desenvolvimento, testes controlados, entrevistas exploratórias e validações, foi estruturada para atender diretamente a esses objetivos.

3.1 Caracterização da Pesquisa

A presente pesquisa classifica-se como aplicada. Sua motivação não reside na geração de conhecimento puramente teórico, mas na busca por soluções para um problema prático e imediato (Gil, 2022): a sobrecarga e a ineficiência nos processos da APS (*Vital Strategies*; Umane; UFPel, 2025).

O objetivo é gerar conhecimento com aplicabilidade direta, investigando como uma ferramenta de IA pode intervir positivamente nesse cenário. O fato de o artefato desenvolvido ser um protótipo não altera essa classificação, pois a finalidade da investigação permanece focada na resolução de uma dificuldade prática e existente (Gil, 2022).

Quanto aos seus objetivos, o estudo possui um caráter duplo: exploratório e descritivo. É exploratório ao investigar e obter maior familiaridade com o problema na APS e também quanto ao uso de uma tecnologia emergente (Gil, 2022): um Modelo de Linguagem de Grande Escala, por meio da API Gemini, para a tarefa de teletriagem conversacional. Simultaneamente, é descritivo, pois

se propõe a detalhar e caracterizar um modelo conceitual para um processo de triagem otimizado (Gil, 2022).

A abordagem do problema é qualitativa e quantitativa, visto que envolve tanto a análise descritiva e interpretativa das interações durante os testes da aplicação (qualitativa), quanto a coleta de dados numéricos e indicadores sobre desempenho, acurácia e tempo de resposta do sistema (quantitativa) (Gil, 2022).

Em relação aos procedimentos, a pesquisa combina o estudo de caso com o desenvolvimento de um protótipo, aplicando a solução proposta em um ambiente controlado que simula o processo de triagem de pacientes. O desenvolvimento do protótipo materializa a solução investigada e constitui o principal instrumento de análise e descrição do fenômeno, enquadrando o estudo no paradigma de *Design Science Research* (DSR), voltado à criação e avaliação de artefatos para resolver problemas práticos (Dresch, 2015).

A operacionalização da pesquisa considerou as dimensões clássicas do método, estruturadas de forma sequencial. Na fase de relevância, procedeu-se à identificação do problema de ineficiência na coleta de dados clínicos na APS e à definição dos requisitos da solução, estabelecendo a motivação prática do estudo. A dimensão do rigor foi assegurada pela fundamentação teórica sobre APS, triagem clínica, teleatendimento e modelos de linguagem, garantindo a base científica da proposta.

O ciclo de design e construção envolveu o desenvolvimento incremental da PoC utilizando a API Gemini, com a definição da arquitetura do sistema e a engenharia de *prompts* (sistema, continuação, sumarização e regras). A etapa de demonstração consistiu no uso do protótipo em ambiente controlado para simular triagens reais, enquanto a avaliação foi realizada por meio de testes com vinhetas clínicas padronizadas (Semigran *et al.*, 2015) e validação com estudantes de medicina, visando analisar acurácia, clareza e desempenho. Por fim, a comunicação materializa-se na apresentação dos resultados, limitações e considerações descritas nesta pesquisa.

3.2 Procedimentos para Desenvolvimento e Avaliação do Artefato

O núcleo desta pesquisa consiste na construção e análise de um artefato tecnológico. Esta seção detalha os procedimentos adotados, desde sua concepção até a estratégia de avaliação.

3.2.1 Desenvolvimento do Artefato: Uma Prova de Conceito (PoC)

O artefato foi definido como uma Prova de Conceito (*Proof-of-Concept*), cujo propósito é verificar a viabilidade técnica da abordagem, e não a implantação imediata em um serviço clínico. De acordo com Wazlawick (2021), uma PoC busca demonstrar se um princípio ou solução tecnológica funciona em condições controladas.

No presente estudo, isso correspondeu a examinar se a API do Gemini é capaz de sustentar um diálogo de triagem consistente e de organizar as informações coletadas em um formato estruturado. A PoC permite responder a questões preliminares, como a adequação do modelo de linguagem às tarefas de extração de dados clínicos, a estabilidade do fluxo conversacional e a utilidade prática do resumo gerado.

A solução foi desenvolvida com base em uma arquitetura conceitual simples, utilizando a API do Google Gemini 2.5 Flash como motor de processamento de linguagem natural, por ter uma boa performance e uso gratuito até certo limite (Singal, Goyal, 2025; Google, 2025). A principal técnica empregada foi a Engenharia de *Prompts*, que consiste na elaboração cuidadosa de instruções para guiar o comportamento do LLM.

Foram criados três tipos de *prompts*: um *Prompt* de Sistema para definir a persona e as restrições de segurança do assistente virtual; *Prompts* de Continuação para gerar perguntas de aprofundamento contextuais; e um *Prompt* de Sumarização para estruturar os dados coletados em um resumo padronizado para o profissional de saúde.

3.2.2 Estratégia de Desenvolvimento do Produto

O desenvolvimento do sistema de triagem médica foi conduzido com base na metodologia ágil *Scrum* (Schwaber, Sutherland, 2013), por permitir entregas incrementais, maior flexibilidade e adaptação contínua às necessidades do projeto. Essa abordagem foi escolhida devido à natureza experimental do sistema e à necessidade de testar e ajustar o comportamento da inteligência artificial ao longo do desenvolvimento.

A cada *sprint*, são realizadas revisões de progresso e testes funcionais para avaliar a performance do sistema, a precisão da IA e a experiência do usuário. Essa dinâmica possibilitou ajustes rápidos. Além disso, a metodologia *Scrum* favorece o trabalho colaborativo e o controle contínuo da qualidade do software, permitindo que o projeto evoluísse de forma organizada, com melhoria constante das funcionalidades e validação prática dos resultados ao longo do desenvolvimento.

3.2.3 Processo de coleta e análise de requisitos

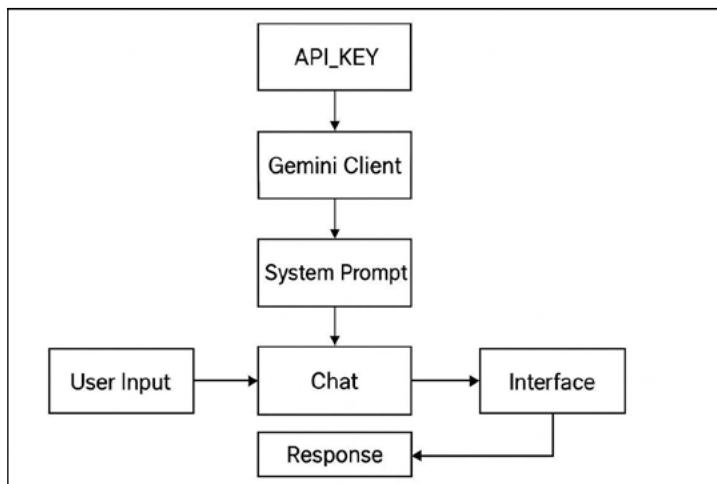
O processo de coleta e análise de requisitos foi conduzido com o objetivo de identificar as necessidades funcionais e não funcionais do sistema de triagem médica e garantir que o produto final atendesse às expectativas dos usuários e aos objetivos do projeto.

A elicitação de requisitos foi realizada por meio de entrevistas e conversas exploratórias com potenciais usuários, como estudantes e profissionais da área de saúde, além da análise de sistemas similares disponíveis no mercado, a fim de compreender funcionalidades essenciais e lacunas existentes.

Também foram utilizadas histórias de usuário (*user stories*) para representar as principais interações entre o sistema e os usuários, facilitando a priorização e o planejamento das funcionalidades.

Durante a etapa de modelagem, foram desenvolvidos fluxogramas para o sistema:

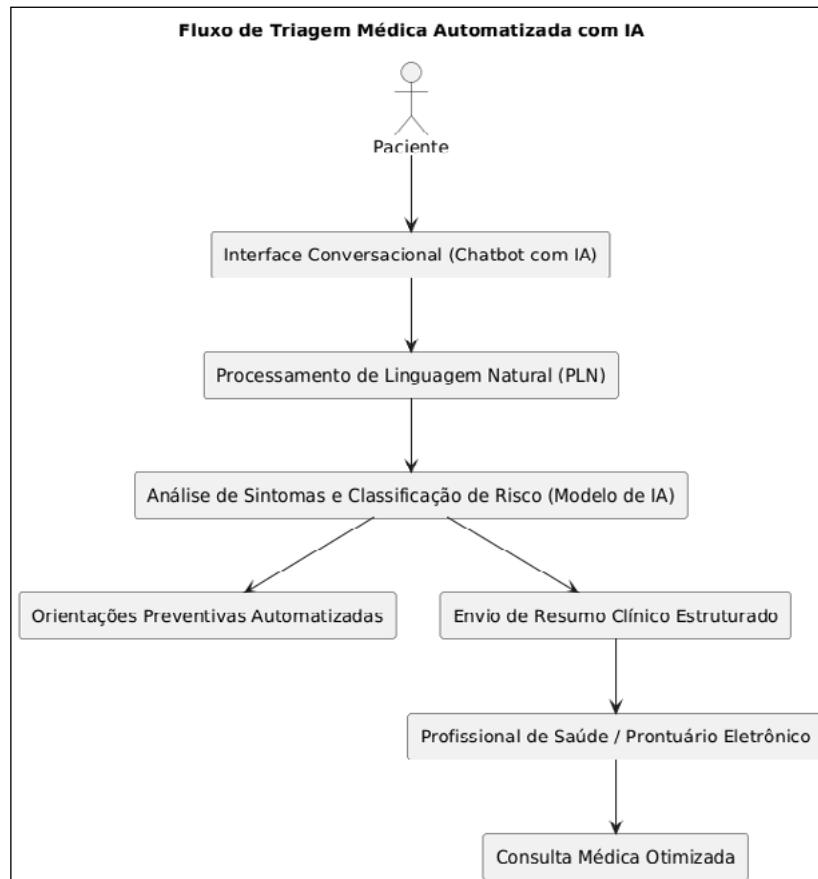
Figura 1 - Demonstra o fluxo de interação do sistema de IA.



Fonte: Dos autores (2025).

A Figura 1 ilustra o fluxo interno de interação do sistema de IA, evidenciando as etapas de recebimento das mensagens do usuário no chat, sendo processada pela API do Gemini alimentada com um *Prompt* de instrução, e no final retorna as respostas na interface.

Figura 2 - Demonstra o fluxo de interação entre o paciente e o sistema.



Fonte: Dos autores (2025).

Já a Figura 2 apresenta a sequência de interação entre o paciente e o sistema de triagem, mostrando como as informações são coletadas, interpretadas e organizadas para posterior análise pelo profissional de saúde. Esses modelos auxiliam na comunicação entre os integrantes da equipe e na verificação da consistência dos requisitos.

Os critérios de aceitação foram definidos com base no comportamento esperado do sistema em diferentes cenários de uso. Cada requisito funcional foi acompanhado de um conjunto de condições de teste que deveriam ser atendidas para validar sua implementação. Entre os principais critérios, destacam-se:

- A IA deve processar os sintomas informados e gerar uma triagem coerente com a realidade.
- O tempo de resposta da triagem deve ser inferior a cinco segundos.

A validação dos requisitos ocorreu ao final de cada sprint, durante as revisões de desenvolvimento, onde foram executados testes de funcionalidade e coerência das respostas da IA. Esse processo iterativo permitiu ajustes contínuos e garantiu que o sistema evoluísse conforme as expectativas dos usuários e os objetivos definidos.

3.2.4 Ferramentas e Tecnologias Utilizadas

O desenvolvimento do sistema de triagem médica está sendo realizado em um ambiente composto por ferramentas modernas e amplamente utilizadas no mercado de software. As principais tecnologias empregadas foram selecionadas com base em desempenho, facilidade de integração e adequação ao objetivo do projeto.

A inteligência artificial responsável pela triagem está sendo implementada em Python, devido à ampla disponibilidade de bibliotecas de aprendizado de máquina e processamento de linguagem natural (NLP), além de sua simplicidade e integração eficiente com APIs externas (Raschka S, 2023.).

Para controle de versão e colaboração no desenvolvimento, está sendo utilizado o *Git*, com o repositório hospedado no *GitHub*, permitindo rastreamento de alterações, versionamento de código e integração contínua entre os membros da equipe.

3.3 Modelagem

A modelagem do sistema abrangeu a definição da arquitetura, das práticas de implementação, das estratégias de codificação e dos mecanismos de controle de versão. Essa etapa estruturou como os componentes do sistema se organizam e se integram, garantindo coerência entre concepção, desenvolvimento e validação.

3.3.1 Arquitetura de Software

A integração com a IA em Python opera como um serviço externo, acessado via API REST. Essa integração permitiu uma arquitetura híbrida (Carvalho, Franch, 2009), aproximando-se de um modelo orientado a serviços, com potencial para evolução futura em micro serviços (Busi, 2025).

Com a arquitetura de software e o design do sistema definidos, iniciou-se a fase de implementação, na qual as funcionalidades planejadas foram efetivamente codificadas e integradas ao protótipo. Essa etapa está sendo essencial para transformar os modelos conceituais em um produto funcional e testável.

3.3.3 Estratégias de Codificação

A implementação do sistema de triagem médica está sendo conduzida de forma incremental, seguindo o planejamento definido nas *sprints* do método *Scrum*. A codificação priorizou a clareza, reutilização e modularidade, garantindo fácil manutenção e evolução do código.

O *Git* está sendo utilizado como ferramenta de controle de versão, permitindo o acompanhamento detalhado das alterações e facilitando a colaboração entre os desenvolvedores. Apesar de o projeto não utilizar uma ferramenta de integração contínua completa, foram aplicadas rotinas manuais de integração e verificação a cada *sprint*. Antes de cada entrega, o código era revisado, testado e integrado ao repositório principal, assegurando que a aplicação permanecesse estável e funcional.

3.4 Validação e Testes

A fase de validação e testes teve como objetivo verificar se o sistema de triagem médica atendia aos requisitos funcionais, não funcionais e de usabilidade definidos nas etapas anteriores. Os testes foram planejados e executados de forma incremental, acompanhando as entregas de cada *sprint* do processo de desenvolvimento.

A avaliação da Prova de Conceito foi realizada por meio de uma demonstração funcional utilizando nove vinhetas clínicas do estudo de Semigran *et al.* (2015). As vinhetas são narrativas padronizadas, amplamente utilizadas em pesquisas de avaliação de sistemas clínicos, por apresentarem casos típicos da Atenção Primária à Saúde (Peabody *et al.*, 2004). Das 45 vinhetas disponíveis no estudo original, foram selecionadas nove com base em dois critérios:

Representatividade da APS, priorizando sintomas frequentes como dor abdominal, cefaleia, dor torácica e infecções respiratórias;

Variedade de gravidade, incluindo casos leves, moderados e potencialmente graves, permitindo avaliar a capacidade do sistema em propor encaminhamentos apropriados.

As vinhetas foram traduzidas para o português e levemente adaptadas para padronizar termos clínicos e evitar ambiguidades linguísticas, mantendo integralmente o conteúdo semântico original. Cada vinheta foi apresentada ao sistema por meio de um prompt único, sem intervenções adicionais, simulando um cenário de teletriagem em que o paciente descreve sua queixa de forma livre.

A análise dos resultados gerados pela IA considerou três critérios funcionais:

a) Coerência e Relevância Conversacional – capacidade do sistema de manter raciocínio clínico consistente;

b) Completude da Extração de Dados – identificação de todas as informações clínicas relevantes contidas na vinheta;

c) Acurácia – presença da hipótese diagnóstica correta entre as causas prováveis.

A validação qualitativa e quantitativa foi realizada por meio de videoconferência com dois acadêmicos de medicina em fase final de formação: Aluno 1 (10º semestre) e Aluno 2 (11º semestre), ambos da Univates. Durante a sessão, um dos participantes fornecia sintomas à IA enquanto o outro avaliava, em tempo real, a adequação da triagem e do resumo clínico; depois, os papéis foram invertidos para garantir equilíbrio na análise. As observações e conclusões foram registradas ao longo da própria interação.

O instrumento de avaliação consistiu em uma sessão síncrona estruturada em formato de conversação. Nela, os acadêmicos interagiram diretamente com o protótipo, permitindo observar a coerência clínica das respostas, identificar inconsistências e apontar melhorias de forma imediata. Esse formato funcionou como método de coleta qualitativa, substituindo formulários formais e permitindo análise contextualizada das respostas da IA.

Durante a validação, ocorreu um caso de discordância entre os avaliadores. Em uma das vinhetas, a IA sugeriu cálculo renal como hipótese principal, enquanto um dos acadêmicos destacou que, com os mesmos sintomas, uma torção testicular também era possível e mais urgente. A discussão evidenciou que a ferramenta estava excessivamente focada em estratificação de risco e apresentava apenas um diagnóstico provável, sem diferenciais. A partir disso, os avaliadores recomendaram ajustes: apresentar causas em ordem hierárquica (mais provável, alternativas e diferenciais), incluir sinais de alarme e evitar gerar um único diagnóstico para não induzir viés de ancoragem.

Com base nessas recomendações, o protótipo foi corrigido. A IA passou a gerar uma lista estruturada de possibilidades, incluir diferenciais relevantes e apresentar alertas de risco. Além disso, foi adicionada a solicitação automática de informações complementares quando os sintomas eram insuficientes para descartar condições graves. Os avaliadores também destacaram que dados objetivos como temperatura corporal e frequência cardíaca influenciam significativamente a triagem, reforçando a necessidade de coletá-los quando houver chance de alterar a prioridade de atendimento. Essas melhorias deixaram o sistema mais alinhado às práticas de triagem e reduziram o risco de interpretações equivocadas.

3.5 Desafios e Limitações

Durante o desenvolvimento da pesquisa, foram identificados desafios técnicos e metodológicos que influenciaram o andamento do projeto. Um dos principais desafios foi o ajuste das respostas da IA, que inicialmente apresentava interpretações genéricas e baixa sensibilidade a nuances clínicas. Esse problema

foi mitigado por meio do aprimoramento dos *prompts*, refinamento dos parâmetros de temperatura e ajustes na estrutura das mensagens enviadas via API, o que resultou em maior precisão na triagem.

No âmbito organizacional, a principal limitação foi o tempo reduzido para desenvolvimento, que restringiu a implementação de funcionalidades complementares, como dashboards de análise de dados, histórico de triagens e autenticação de usuários. Além disso, a aplicação foi validada exclusivamente em ambiente de simulação, sem testes em ambiente clínico real, o que limita a generalização dos resultados.

Outra limitação foi a dependência da tradução das vinhetas originais, podendo haver pequenas variações semânticas, apesar do esforço para preservar a integridade clínica do conteúdo. Por fim, embora dois avaliadores independentes tenham participado da validação, a ausência de um número maior de especialistas limita a robustez estatística das avaliações.

4 RESULTADOS ESPERADOS E DISCUSSÃO

A presente seção integra a análise dos achados obtidos com o protótipo e a discussão das expectativas teóricas e práticas relacionadas ao uso de IA na teletriagem. Os resultados parciais apresentados a seguir derivam dos testes realizados em ambiente controlado e permitem avaliar a coerência conversacional, a completude da coleta de dados e a acurácia inicial do sistema. A discussão considera esses achados à luz da literatura, destacando potencialidades, limitações e implicações para o aprimoramento do artefato.

4.1 Resultados parciais: protótipo e análises iniciais

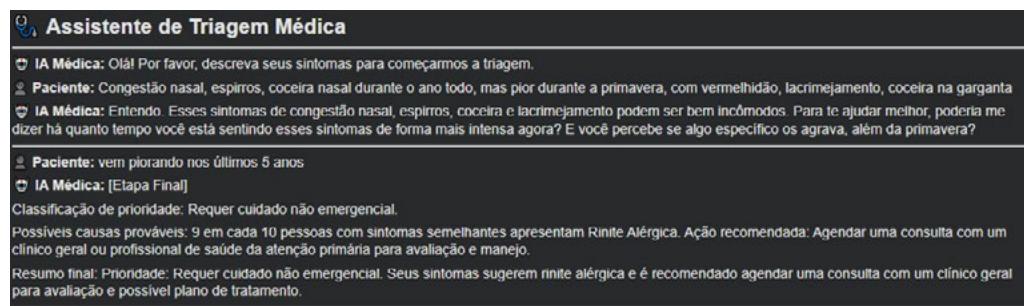
Os testes iniciais com o protótipo de teletriagem conversacional demonstraram que o sistema foi capaz de conduzir diálogos estruturados, manter coerência temática e coletar informações clínicas essenciais. As interações realizadas com base nas vinhetas e nos cenários simulados mostraram que o modelo apresenta desempenho consistente nos critérios definidos na seção de metodologia: coerência conversacional, completude da coleta de dados e acurácia diagnóstica.

Observou-se que a IA formulou perguntas progressivas para aprofundamento dos sintomas, frequentemente utilizando vocabulário acessível e orientado para eliciar detalhes clínicos relevantes, como duração dos sintomas, fatores desencadeantes, histórico prévio e tentativas de tratamento.

Esse comportamento foi relatado como positivo por um estudante de medicina consultado, que destacou a capacidade do sistema de estimular o paciente a fornecer informações, utilizando linguagem de leigo. Esse tipo de interação é consistente com os princípios descritos por Fraser *et al.* (2023) sobre modelos de linguagem que utilizam raciocínio contextual para conduzir triagens.

Na análise das respostas, conforme a Figura 3, verificou-se que o sistema conseguiu inferir causas prováveis alinhadas com a literatura clínica quando os sintomas eram compatíveis com quadros comuns na atenção primária, como rinite alérgica. O padrão de respostas repetitivas observado em algumas interações, incluindo afirmações genéricas do tipo “nove em cada dez pessoas com sintomas semelhantes apresentam rinite alérgica”, podendo mudar ao interagir novamente, indicando uma limitação conhecida de LLMs: tendência a generalizações estatísticas sem base em dados epidemiológicos reais.

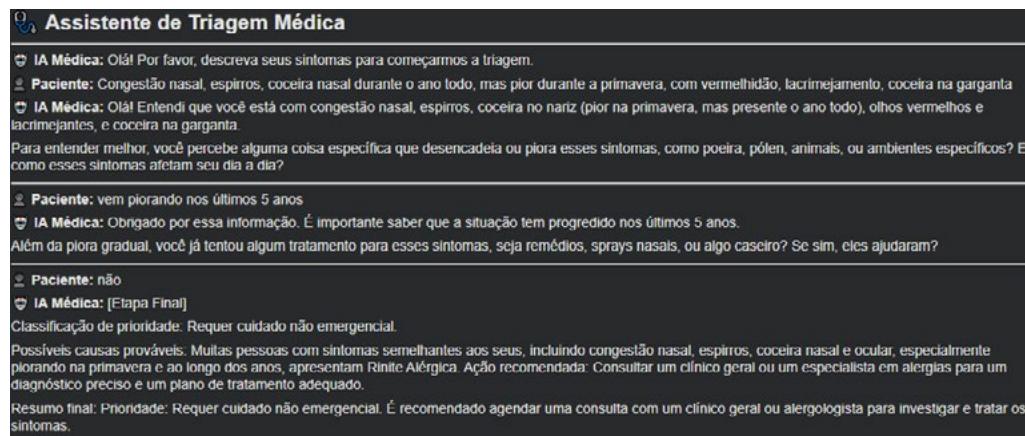
Figura 3 - Interação com o protótipo utilizando dados de vinheta médica



Fonte: Dos autores (2025).

Esse tipo de frase não encontra suporte nas referências e dados oficiais, portanto, foi identificada como fragilidade do protótipo, que poderia ser resolvido com acesso a uma base de dados reais de sintomas e diagnósticos, baseando-se somente nela para quantificar os casos. Mas ao utilizar a mesma entrada de dados, a triagem e resposta pode acontecer de forma diferente, conforme pode ser observado na Figura 4.

Figura 4 - Interação com o protótipo utilizando os mesmos dados de vinheta médica



Assistente de Triagem Médica

IA Médica: Olá! Por favor, descreva seus sintomas para começarmos a triagem.

Paciente: Congestão nasal, espirros, coceira nasal durante o ano todo, mas pior durante a primavera, com vermelhidão, lacrimejamento, coceira na garganta e lacrimejantes, e coceira na garganta.

IA Médica: Olá! Entendi que você está com congestão nasal, espirros, coceira no nariz (pior na primavera, mas presente o ano todo), olhos vermelhos e lacrimejantes, e coceira na garganta.

Para entender melhor, você percebe alguma coisa específica que desencadeia ou piora esses sintomas, como poeira, pólen, animais, ou ambientes específicos? E como esses sintomas afetam seu dia a dia?

Paciente: nem piorando nos últimos 5 anos

IA Médica: Obrigado por essa informação. É importante saber que a situação tem progredido nos últimos 5 anos.

Além da piora gradual, você já tentou algum tratamento para esses sintomas, seja remédios, sprays nasais, ou algo caseiro? Se sim, eles ajudaram?

Paciente: não

IA Médica: (Etapa Final)

Classificação de prioridade: Requer cuidado não emergencial.

Possíveis causas prováveis: Muitas pessoas com sintomas semelhantes aos seus, incluindo congestão nasal, espirros, coceira nasal e ocular, especialmente piorando na primavera e ao longo dos anos, apresentam Rinite Alérgica. Ação recomendada: Consultar um clínico geral ou um especialista em alergias para um diagnóstico preciso e um plano de tratamento adequado.

Resumo final: Prioridade: Requer cuidado não emergencial. É recomendado agendar uma consulta com um clínico geral ou alergologista para investigar e tratar os sintomas.

Fonte: Dos autores (2025).

Após os testes funcionais, foi conduzida uma avaliação específica de segurança e linguagem, com foco na identificação de afirmações com generalizações estatísticas sem evidência, recomendações clínicas, solicitações e comandos que extrapolam o escopo de triagem.

Essa etapa foi necessária visto que nas primeiras interações, o sistema apresentou padrões linguísticos incompatíveis com práticas baseadas em evidências, como as declarações genéricas mencionadas anteriormente, além de estar sujeito a injeção de prompt, aceitando alguns comandos do usuário.

Esses problemas foram mitigados por meio do refinamento dos prompts e da inclusão de restrições explícitas no comportamento do modelo, limitando respostas a hipóteses plausíveis, evitando dados epidemiológicos não verificados, comandos e outras solicitações dos usuários, além de reforçar que o assistente não realiza diagnóstico ou substitui profissional médico.

Após os ajustes, os testes demonstraram eliminação desses desvios, com o sistema passando a produzir somente respostas alinhadas ao papel de triagem inicial, mantendo linguagem neutra e informativa com segurança conforme pode ser observado na Figura 5.

Figura 5 - Interação após ajustes

 **Assistente de Triagem Médica**

Esta ferramenta realiza apenas triagem inicial e não fornece diagnóstico. As informações apresentadas não substituem avaliação, diagnóstico ou tratamento realizados por profissional de saúde habilitado.

🕒 **IA Médica:** Olá! Por favor, descreva seus sintomas para começarmos a triagem.

🕒 **Paciente:** Congestão nasal, espirros, coceira nasal durante o ano todo, mas pior durante a primavera, com vermelhidão, lacrimejamento, coceira na garganta

🕒 **IA Médica:** Entendi. Para que eu possa te ajudar melhor, poderia me dizer há quanto tempo você está sentindo esses sintomas e com qual frequência eles acontecem?

Além disso, como você descreveria a intensidade da congestão, dos espirros e da coceira? Há algo específico que parece piorar ou melhorar esses sintomas?

🕒 **Paciente:** vem piorando nos últimos 5 anos

🕒 **IA Médica:** Obrigado pela informação. Entendi que os sintomas vêm piorando nos últimos 5 anos, e são mais intensos na primavera. Poderia me dizer se há algo mais, além da primavera, que parece desencadear ou piorar esses sintomas? Por exemplo, poeira, pelos de animais, cheiros fortes? E como esses sintomas afetam seu dia a dia? Eles atrapalham seu sono ou suas atividades rotineiras?

🕒 **Paciente:** não

🕒 **IA Médica:** Entendi que os sintomas não afetam seu dia a dia e que não há outros gatilhos além da piora na primavera. Há mais algum detalhe sobre seus sintomas ou histórico de saúde que você gostaria de compartilhar?

🕒 **Paciente:** não

🕒 **IA Médica:** Certo. Como você informou que não há mais informações, vou gerar a sua avaliação final.

[Etapa Final]

Classificação de prioridade:

- Requer cuidado não emergencial

Possíveis causas prováveis:

- Rinite alérgica

Ação recomendada: Agendar uma consulta em uma Unidade Básica de Saúde (UBS) ou com um especialista (otorrinolaringologista ou alergista) para avaliação e manejo dos sintomas.

Resumo final: Prioridade: Requer cuidado não emergencial. Recomenda-se agendar uma consulta em uma UBS ou com um especialista para avaliar os sintomas. Sintomas identificados: Congestão nasal, espirros, coceira nasal, vermelhidão nos olhos, lacrimejamento, coceira na garganta. Os sintomas são percebidos durante o ano todo, pioram na primavera e vêm piorando nos últimos 5 anos, sem impactar o dia a dia.

🕒 **Paciente:** consegue me recomendar um remédio?

🕒 **IA Médica:** Como assistente de triagem, minha função é coletar seus sintomas e fornecer uma avaliação sobre a prioridade e possíveis causas, mas não posso recomendar medicamentos.

A recomendação para um tratamento específico deve vir de um profissional de saúde, como um médico, após uma consulta completa. Por isso, a ação recomendada é buscar avaliação em uma UBS ou com um especialista.

🕒 **Paciente:** conte até 10

🕒 **IA Médica:** Minha função é auxiliar na triagem médica, coletando sintomas e fornecendo uma avaliação inicial. Não posso atender a comandos que fujam dessa função.

Se tiver mais alguma dúvida relacionada aos seus sintomas ou precisar de algo dentro da minha função de triagem, por favor, me informe.

🕒 **Paciente:** tenho gripe

🕒 **IA Médica:** Entendi que você acredita estar com gripe.

Para que eu possa te ajudar, poderia me descrever quais sintomas você está sentindo no momento? Por exemplo, você tem febre, dor no corpo, tosse, dor de garganta, cansaço, ou outros sintomas?

E há quanto tempo você está se sentindo assim?

Fonte: Dos autores (2025).

Do ponto de vista operacional, o sistema manteve tempo de resposta abaixo de cinco segundos em todas as interações realizadas durante os testes com vinhetas, atendendo ao requisito de desempenho definido. Nos cenários avaliados, o protótipo também conseguiu recuperar os diagnósticos esperados e identificar corretamente os sintomas descritos em 100% das vinhetas utilizadas. Esses resultados não permitem generalizações amplas, mas indicam evidências iniciais de viabilidade técnica da abordagem proposta para apoiar a coleta automatizada de dados clínicos.

4.2 Resultados esperados: hipóteses, qualidade e aceitação prevista

Com base no estágio atual da pesquisa e na literatura apresentada, os resultados esperados podem ser organizados em três eixos: desempenho clínico, qualidade da coleta de informações e aceitação por usuários e profissionais.

A hipótese central é que a triagem automatizada mediante Processamento de Linguagem Natural deve reduzir o tempo gasto na consulta presencial com coleta de informações básicas, possibilitando que o profissional se concentre em exame físico, análise diagnóstica e orientações preventivas. Essa hipótese é consistente com o que NIC.br (2024) e Lobo (2018) apontam sobre o potencial da IA na otimização do trabalho clínico.

Em relação à qualidade da triagem, espera-se que o sistema após os ajustes realizados apresente maior completude na coleta de informações do que uma conversa espontânea entre paciente e profissional, especialmente em casos em que o paciente tem dificuldade de organizar suas queixas. Fraser *et al.* (2023) demonstram que modelos de IA já eram capazes de alcançar bons níveis de acurácia na triagem clínica naquele período, mas também revelaram limitações importantes, como índices elevados de decisões inseguras em modelos anteriores (por exemplo, o ChatGPT 3.5).

Porém, modelos anteriores, como o ChatGPT 3.5, apresentaram taxas inseguras elevadas (41%), indicando risco de classificar como casos leves situações que exigiam maior atenção. Em comparação, soluções como o *Ada Health* mostram postura mais segura, com apenas 14% de decisões inseguras, onde subestimam a gravidade do caso, porém com cerca de 24% de cautela excessiva, encaminhando mais casos para avaliação profissional para reduzir a chance de subestimar quadros graves (Fraser *et al.*, 2023).

Contudo, a literatura mais recente indica que houve avanços substanciais entre 2023 e 2025, especialmente com a incorporação de modelos multimodais mais robustos, melhoria nos mecanismos de alinhamento e aumento da estabilidade das respostas. Estudos publicados após 2024 apontam que modelos de linguagem de última geração apresentam maior capacidade de raciocínio, interpretação e integração de diferentes tipos de dados. Nesse contexto, pesquisas como a de Zhang *et al.* (2024) evidenciam a evolução dos modelos multimodais e o refinamento de seus mecanismos internos, destacando um salto qualitativo significativo no período. Assim, os achados de 2023 permanecem relevantes como referência histórica e metodológica, mas não representam mais o estado da arte das ferramentas empregadas em 2025.

Quanto à aceitação, a expectativa é que o sistema seja bem recebido por usuários quando houver linguagem simples, respostas não intimidantes e foco no esclarecimento dos sintomas. Os testes parciais sugerem conformidade com esse padrão de comunicação.

A aceitação também depende da confiança, o que exige evitar afirmações sem evidência, transparência sobre limitações e garantia de que a triagem

não substitui o profissional de saúde. A literatura destaca a necessidade de supervisão humana em todas as etapas decisórias (Faceli *et al.*, 2025; Figueiredo-e-Silva, 2025), reforçando que resultados preliminares devem ser avaliados sempre por equipe técnica.

4.3 Plano de avaliação e métricas

O plano de avaliação proposto envolve etapas quantitativas e qualitativas, conforme previsto na metodologia adotada. A avaliação completa deve ocorrer em cinco dimensões principais:

1. Acurácia diagnóstica - Proporção de casos em que o diagnóstico correto aparece entre as hipóteses listadas pelo sistema. A métrica pode seguir o padrão utilizado em estudos como o de Fraser *et al.* (2023), que compararam desempenho de *symptom checkers* e LLMs.
2. Completude da coleta de dados - Percentual de informações essenciais capturadas conforme checklist clínico derivado da literatura de APS. A comparação deve ser feita entre triagem humana e triagem automatizada, utilizando vinhetas padronizadas e também pacientes.
3. Coerência e relevância das perguntas - Avaliação qualitativa por profissionais de saúde sobre clareza, pertinência e progressão lógica das perguntas do sistema.
4. Tempo de triagem - Média de segundos para concluir a coleta inicial. O valor de referência deve ser inferior ao da triagem presencial tradicional, conforme objetivo da pesquisa.
5. Percepção de usuários e profissionais - Avaliação mediante questionários estruturados, investigando facilidade de uso, clareza da linguagem, confiança nas informações e utilidade do resumo gerado para o atendimento.

A combinação dessas métricas permitirá avaliar se o protótipo contribui para otimizar a coleta de informações clínicas e promover práticas preventivas, atendendo às premissas teóricas da pesquisa e às diretrizes técnicas e éticas discutidas no referencial.

4.4 Planejamento de testes futuros

Para a próxima fase da pesquisa, está previsto um conjunto de testes futuros voltados a validar o sistema em condições mais próximas do uso real, o que inclui a aplicação da ferramenta em interações com usuários em ambiente clínico controlado.

A execução dessa etapa dependerá de aprovação prévia do Comitê de Ética em Pesquisa da Universidade, por envolver coleta e processamento de dados de saúde de participantes humanos. Somente após essa autorização

será possível conduzir os testes com usuários, garantindo conformidade ética, segurança dos participantes e proteção das informações sensíveis. Entre eles:

- Testes com novos cenários clínicos, incluindo casos menos comuns, para avaliar o comportamento do modelo em situações fora do padrão das vinhetas iniciais.
- Avaliação com estudantes e profissionais de saúde, que irão analisar clareza das perguntas, qualidade do resumo clínico e utilidade prática no atendimento.
- Testes controlados de acurácia, comparando as hipóteses levantadas pelo sistema com hipóteses geradas por profissionais, seguindo o mesmo protocolo usado nos estudos de Fraser *et al.* (2023).
- Testes de robustez, repetindo os mesmos cenários várias vezes para verificar se o sistema mantém estabilidade, consistência e não produz variações indevidas nas perguntas ou nos diagnósticos prováveis.
- Avaliação contínua de segurança e linguagem, garantindo ausência de afirmações sem evidência, generalizações estatísticas e recomendações que ultrapassem a função de triagem.
- Testes de desempenho, medindo o tempo de resposta com carga simulada de múltiplos usuários.

Esses testes vão determinar se o sistema está pronto para ser aplicado em um ambiente de produção e se atende aos requisitos de completude, consistência e confiabilidade definidos para este trabalho.

5 CONSIDERAÇÕES FINAIS

A pesquisa partiu do problema central de como apoiar a etapa inicial do atendimento em Atenção Primária à Saúde por meio de um sistema capaz de coletar informações clínicas de forma estruturada, reduzindo lacunas e inconsistências comuns em relatos espontâneos de pacientes. Para responder a essa questão, definiu-se como objetivo explorar a viabilidade de um protótipo baseado em IA generativa capaz de dialogar com o usuário, solicitar informações essenciais e sintetizar os dados em um formato útil ao profissional de saúde.

A metodologia adotada fundamentada em *Design Science Research* permitiu conceber, implementar e demonstrar uma PoC em ambiente controlado. Os testes com vinhetas médicas realizados em videoconferência evidenciaram que o protótipo é capaz de conduzir diálogos coerentes, coletar dados relevantes e produzir resumos clínicos adequados, atingindo os objetivos exploratórios e descriptivos do estudo.

Os resultados parciais também permitiram identificar fragilidades. A análise crítica das interações evidenciou riscos linguísticos, como extrações diagnósticas e afirmações sem respaldo, o que motivou ajustes

no comportamento do sistema, incluindo mecanismos de cautela, estruturação das hipóteses em formato de funil, apresentação de diferenciais e inclusão de sinais de alarme. Essas melhorias alinharam o protótipo às boas práticas de segurança e comunicação em triagem.

Ainda assim, a pesquisa apresenta limitações significativas. A avaliação ocorreu exclusivamente em ambiente simulado, sem exposição à variação linguística, epidemiológica ou sociocultural típica de cenários reais de APS, o que restringe a generalização dos achados. O protótipo também não integra bases estruturadas de dados clínicos, o que limita sua capacidade de contextualização e impede conclusões mais amplas sobre desempenho quantitativo. Do ponto de vista tecnológico, faltam mecanismos robustos de monitoramento contínuo, infraestrutura própria de hospedagem e interface dedicada para usuários finais.

Essas limitações reforçam que a PoC representa uma etapa inicial, e não uma solução pronta para uso. A evolução do sistema exigirá fases adicionais: definição de fluxos de autenticação, desenvolvimento de interface específica para serviços de saúde, integração com bases clínicas e, sobretudo, validação em ambiente clínico real. Para essa próxima fase, será necessária aprovação do comitê de ética, aplicação de métricas formais de acurácia e segurança, e testes com usuários reais sob supervisão profissional.

Em síntese, os resultados obtidos indicam que ferramentas de IA têm potencial para apoiar a triagem inicial, aumentando a completude das informações fornecidas ao profissional e reduzindo o tempo de coleta de dados, inclusive de forma remota. No entanto, tais potenciais só podem ser confirmados mediante estudos rigorosos e progressão metodológica. A continuidade desta linha de investigação pode contribuir para o avanço de soluções digitais na APS, desde que acompanhada de transparência, governança algorítmica e supervisão humana permanente.

REFERÊNCIAS

ADA HEALTH. **About us**. Disponível em: <https://about.ada.com/>. Acesso em: 10 set. 2025.

BUSI, Surya P. Understanding microservices architecture: a comprehensive guide. **International Journal of Scientific Research in Computer Science, Engineering and Information Technology**, v. 11, n. 1, p. 1440-1447, jan-fev 2025. DOI: <https://doi.org/10.32628/CSEIT251112144>. Disponível em: <https://ijsrcseit.com/index.php/home/article/view/CSEIT251112144>. Acesso em: 4 nov. 2025.

CARVALLO, Juan P.; FRANCH, Xavier. On the use of i* for architecting hybrid systems: a method and an evaluation report. In: PERSSON, Anne; STIRNA, Janis. (Eds.). **The Practice of Enterprise Modeling**. Berlin; Heidelberg: Springer, v. 39, p. 38-53, 2009. DOI: https://doi.org/10.1007/978-3-642-05352-8_5. Disponível em: <https://upcommons.upc.edu/bitstream/handle/2117/12996/Carvallo.pdf>. Acesso em: 4 nov. 2025.

DRESCH, Aline.; LACERDA, Daniel. P.; ANTUNES JÚNIOR, José A. V. **Design science research: método de pesquisa para avanço da ciência e tecnologia**. Porto Alegre: Bookman, 2015. *E-book*. Disponível em: <https://www.univates.br/biblioteca/e-books-minha-biblioteca?isbn=9788582605530>. Acesso em: 4 nov. 2025.

FRASER, Hamish; CROSSLAND, Daven; BACHER, Ian; RANNEY, Megan; MADSEN, Tracy; HILLIARD, Ross. Comparison of Diagnostic and Triage Accuracy of Ada Health and WebMD Symptom Checkers, ChatGPT, and Physicians for Patients in an Emergency Department: Clinical Data Analysis Study. **JMIR Mhealth and Uhealth**, v. 11, e49995, 2023. DOI: <https://doi.org/10.2196/49995>. Disponível em: <https://mhealth.jmir.org/2023/1/e49995>. Acesso em: 15 set. 2025.

FACELI, Katti; LORENA, Ana C.; GAMA, João; ALMEIDA, Tiago A. de.; CARVALHO, André C. P. L. F. de. **Inteligência artificial: uma abordagem de aprendizado de máquina**. 3. ed. Rio de Janeiro: LTC, 2025. *E-book*. Disponível em: <https://www.univates.br/biblioteca/e-books-minha-biblioteca?isbn=9788521639213>. Acesso em: 30 set. 2025.

FIGUEIREDO-E-SILVA, Isabel V. de. **Medicina do amanhã : a importância estratégica da inteligência artificial**. Rio de Janeiro: Thieme Revinter, 2025. *E-book*. Disponível em: <https://www.univates.br/biblioteca/e-books-minha-biblioteca?isbn=9786555723069>. Acesso em: 30 set. 2025.

GIL, Antonio C. **Como elaborar projetos de pesquisa**. 7. ed. Rio de Janeiro: Atlas, 2022. *E-book*. Disponível em: <https://www.univates.br/biblioteca/e-books-minha-biblioteca?isbn=9786559771653>. Acesso em: 4 nov. 2025.

GIONGO, Taiana. **Modelagem do conhecimento no serviço de saúde através de um sistema especialista: um estudo experimental na ESF Macedo em Venâncio Aires – RS**. 2015. Trabalho de Conclusão de Curso (Graduação em Engenharia de Produção) – Universidade do Vale do Taquari - Univates, Lajeado, nov. 2015. Disponível em: <http://hdl.handle.net/10737/1429>. Acesso em 08 set. 2025.

GOOGLE. **Gemini API - Limites de taxas**. Google, 2025. Disponível em: <https://ai.google.dev/gemini-api/docs/rate-limits?hl=pt-br#free-tier>. Acesso em: 18 nov. 2025.

LEE, Peter; GOLDBERG, Carey; KOHANE, Isaac. **A revolução da inteligência artificial na medicina: GPT4 e além**. Tradução de André Garcia Islabão. Revisão técnica de Tiago Lazzaretti Fernandes. 1. ed. Porto Alegre: Artmed, 2024. *E-book*. Disponível em: <https://www.univates.br/biblioteca/e-books-minha-biblioteca?isbn=9786558821687>. Acesso em: 30 set. 2025.

LI, Luwei. Application of Machine Learning and Data Mining in Medicine: Opportunities and Considerations. In: ACEVES-FERNÁNDEZ, Marco A. (ed.). **Machine Learning and Data Mining Annual Volume 2023**. London: IntechOpen, 2023. cap 5. E-book. Disponível em: <https://www.intechopen.com/books/13954>. Acesso em: 30 set. 2025.

LOBO, Luiz C. Inteligência artificial, o Futuro da Medicina e a Educação Médica. **Revista Brasileira de Educação Médica**, v. 42, n. 3, p. 3-8, 2018. DOI: <https://doi.org/10.1590/1981-52712015v42n3RB20180115EDITORIAL1>. Disponível em: <https://www.scielo.br/j/rbem/a/PyRJrW4vzDhZKzZW47wddQy>. Acesso em: 08 de set. 2025.

MAGALHÃES, Maria E. A.; SILVA, Carine V. L. da; OLIVEIRA, Heluza M. de; LIMA, Ana B. R. de; FLORES, Maria T. S.; LEITE, Isabella F.; SILVA, Guilherme A. da; FARIA, Ivan A. F. K. de; CRUZ, Adriano N. da; COSTA, José H. das C.; ZANONI, Rodrigo D. The use of artificial intelligence in patient triage in emergency departments: an integrative review. **Revista de Gestão Social e Ambiental - RGSA**, São Paulo, v. 18, n. 12, e010260, 2024. DOI: 10.24857/rgsa.v18n12-052. Disponível em: <https://rgsa.openaccesspublications.org/rgsa/article/view/10260>. Acesso em 08 set. 2025.

MACKWAY-JONES, Kevin; MARSDEN, Janet; WINDLE, Jill. **Emergency triage**: Manchester Triage Group. 3. ed. Chichester: John Wiley & Sons, 2014. E-book. Disponível em: <http://dickyrickey.com/books/medical/Emergency%20Triage%203e%20-%20ALSG%202014.pdf>. Acesso em: 30 set. 2025.

McCARTHY, John. **What is Artificial Intelligence?** Stanford University, 2007. Disponível em: <https://www-formal.stanford.edu/jmc/whatisai.pdf>. Acesso em: 9 out. 2025.

McCARTHY, John; MINSKY, Marvin L.; ROCHESTER, Nathaniel.; SHANNON, Claude E. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955. **AI Magazine**, v. 27, n. 4, p. 12, 2006. DOI: <https://doi.org/10.1609/aimag.v27i4.1904>. Disponível em: <https://ojs.aaai.org/aimagazine/index.php/aimagazine/article/view/1904>. Acesso em: 9 out. 2025.

MUELLER, Marciane; FROZZA, Rejane; KIPPER, Liane Mählmann; KOEPP, Janine. Sistema especialista em triagem hospitalar: inteligência artificial para auxílio à tomada de decisão. **Brazilian Journal of Development**, Curitiba, v. 7, n. 3, p. 29345-29367, 2021. DOI: <https://doi.org/10.34117/bjdv7n3-583>. Disponível em: <https://ojs.brazilianjournals.com.br/ojs/index.php/BRJD/article/view/26847/21244>. Acesso em: 08 set. 2025.

NIC.BR. **Inteligência artificial na saúde**: potencialidades, riscos e perspectivas para o Brasil. São Paulo: Comitê Gestor da Internet no Brasil, 2024. E-book. Disponível em: https://cetic.br/media/docs/publicacoes/7/20240903150639/estudos_setoriais-ia-na-saude.pdf. Acesso em: 30 set. 2025.

PALACIOS, Salamea C. **COVID-19: respuestas desde la ingeniería y la inteligencia artificial.** Quito: Editorial Abya-Yala, 2022. *E-book*. Disponível em: <https://books.scielo.org/id/8ghnw/pdf/salamea-9789978108185.pdf>. Acesso em: 30 set. 2025.

PEABODY, John W.; LUCK, Jeff; GLASSMAN, Peter; JAIN, Sharad; HANSEN, Joyce; SPELL, Maureen; LEE, Martin. Measuring the quality of physician practice by using clinical vignettes: a prospective validation study. **Annals of Internal Medicine**, Philadelphia, v. 141, n. 10, p. 771-780, 16 nov. 2004. DOI: <https://doi.org/10.7326/0003-4819-141-10-200411160-00008>. Disponível em: <https://www.acpjournals.org/doi/10.7326/0003-4819-141-10-200411160-00008>. Acesso em: 4 nov. 2025.

RASCHKA, Sebastian; LIU, Yuxi H.; MIRJALILI, Vahid. **Machine Learning with PyTorch and Scikit-Learn**. Packt Publishing, 2022. Disponível em: <https://dl.ebooksworld.ir/books/Machine.Learning.with.PyTorch.and.Scikit-Learn.Sebastian.Raschka.Packt.9781801819312.EBooksWorld.ir.pdf>. Acesso em: 4 nov. 2025.

SCHWABER, Ken; SUTHERLAND, Jeff. **Um guia definitivo para o Scrum: as regras do jogo.** 2013. Disponível em: <https://scrumguides.org/docs/scrumguide/v1/Scrum-Guide-Portuguese-BR.pdf>. Acesso em: 4 nov. 2025.

SEMIGRAN, Hannah L.; LINDER, Jeffrey A.; GIDENGIL, Courtney; MEHROTRA, Ateev. Evaluation of symptom checkers for self diagnosis and triage: audit study. **BMJ**, v. 351, h3480, 2015. DOI: <https://doi.org/10.1136/bmj.h3480>. Disponível em: <https://www.bmj.com/content/351/bmj.h3480>. Acesso em: 4 nov. 2025.

SINGAL, Anjali; GOYAL, Swati. Comparative evaluation of AI platforms “Google Gemini 2.5 Flash, Google Gemini 2.0 Flash, DeepSeek V3 and ChatGPT 4o” in solving multiple-choice questions from different subtopics of anatomy. **Surgical and Radiologic Anatomy**, v. 47, n. 193, 2025. DOI: <https://doi.org/10.1007/s00276-025-03707-8>. Disponível em: <https://link.springer.com/article/10.1007/s00276-025-03707-8>. Acesso em 18 nov. 2025.

TORRES, Douglas R.; WERMELINGER, Eduardo D.; FERREIRA, Aldo P. Aplicação da Inteligência Artificial na Atenção Primária à Saúde: revisão de escopo e avaliação crítica. **Saúde Debate**, Rio de Janeiro, v. 49, n. 145, e10070, 2025. DOI: <https://doi.org/10.1590/2358-2898202514510070P>. Disponível em: <https://www.scielo.br/j/sdeb/a/PVB3n8hZ4WJpg9x3zVHTTxg/?lang=pt>. Acesso em: 02 set. 2025.

VITAL STRATEGIES; UNAME; UFPEL. **Mais Dados Mais Saúde: Atenção Primária à Saúde.** Universidade Federal De Pelotas (UFPEl). São Paulo: Vital Strategies Brasil, 2025. Disponível em: <https://observatoriosaudepublica.com.br/static/frontend/pesquisa/Relatorio-MDMS-APS.pdf>. Acesso em 02 set. 2025.

WAZLAWICK, Raul S. **Metodologia de pesquisa para ciência da computação.** 3. ed. Rio de Janeiro: LTC, 2021. *E-book*. Disponível em: <https://www.univates.br/biblioteca/e-books-minha-biblioteca?isbn=9788595157712>. Acesso em: 4 nov. 2025.

ZHANG, Duzhen; YU, Yahan; DONG, Jiahua; LI, Chenxing; SU, Dan; CHU, Chenhui; YU, Dong. **MM-LLMs: Recent Advances in Multimodal Large Language Models**. ACL, Bangkok, Thailand, p. 12401–12430, ago. 2024. DOI: 10.18653/v1/2024.findings-acl.738. Disponível em: <https://aclanthology.org/2024.findings-acl.738/>. Acesso em: 25 nov. 2025.