



INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO
TOCANTINS – CAMPUS ARAGUATINS
COORDENAÇÃO DE COMPUTAÇÃO
CURSO DE LICENCIATURA EM COMPUTAÇÃO

JOAQUIM GABRIEL DE FARIAS LIMA NETO

**EFICÁCIA DE PROGRAMAS DE INTELIGÊNCIA ARTIFICIAL COM APLICAÇÃO
DE CHATBOTS NO AUXÍLIO AOS CUIDADOS COM A SAÚDE MENTAL**

ARAGUATINS – TO
2026

JOAQUIM GABRIEL DE FARIAS LIMA NETO

**EFICÁCIA DE PROGRAMAS DE INTELIGÊNCIA ARTIFICIAL COM APLICAÇÃO
DE CHATBOTS NO AUXÍLIO AOS CUIDADOS COM A SAÚDE MENTAL**

Trabalho de Conclusão de Curso apresentado à
Coordenação do Curso Superior de Licenciatura em
Computação do Instituto Federal do Tocantins (IFTO) –
Campus Araguatins, como requisito parcial para obtenção
do título de Licenciado em Computação.

Orientador: Prof. Ancelmo Frank Coelho Castro.

Dados Internacionais de Catalogação na Publicação (CIP)
Bibliotecas do Instituto Federal do Tocantins

L732e Lima Neto, Joaquim Gabriel de Farias
 Eficácia De Programas De Inteligência Artificial Com Aplicação De
 Chatbots No Auxílio Aos Cuidados Com A Saúde Mental / Joaquim
 Gabriel de Farias Lima Neto. – Araguatins, TO, 2026.
 58 p.

Trabalho de Conclusão de Curso (Licenciatura em Computação) –
Instituto Federal de Educação, Ciência e Tecnologia do Tocantins - IFTO,
Araguatins, TO, 2026.

Orientador: Me. Ancelmo Frank Coelho Castro.

1. Chatbots. 2. Inteligência Artificial. 3. Saúde mental. I. Castro,
Ancelmo Frank Coelho. II. Título.

CDD 004

A reprodução total ou parcial, de qualquer forma ou por qualquer meio, deste documento é autorizada para fins
de estudo e pesquisa, desde que citada a fonte.

**Elaborado pelo sistema de geração automática de ficha catalográfica do IFTO com os dados fornecidos
pelo(a) autor(a).**



Ministério da Educação
Secretaria de Educação Profissional e Tecnológica
Instituto Federal de Educação, Ciência e Tecnologia do Tocantins
Campus Araguatins

FOLHA DE APROVAÇÃO

TÍTULO DO TRABALHO: EFICÁCIA DE PROGRAMAS DE INTELIGÊNCIA ARTIFICIAL COM APLICAÇÃO DE CHATBOTS NO AUXÍLIO AOS CUIDADOS COM A SAÚDE MENTAL

AUTOR: Joaquim Gabriel de Farias Lima Neto

ORIENTADOR: Prof. Me. Ancelmo Frank Coêlho Castro

Trabalho de Conclusão de Curso apresentado ao Instituto Federal de Educação, Ciência e Tecnologia do Tocantins, *Campus Araguatins*, como parte das exigências para a conclusão do Curso de Licenciatura em Computação.

Aprovado em 18 de junho de 2026.

Prof. Me. Ancelmo Frank Coêlho Castro
Presidente da Banca

Prof. Me. Rogério Pereira de Sousa
Primeiro avaliador

Prof. Dr. Ramásio Ferreira de Melo
Segundo avaliador



Documento assinado eletronicamente por **Ancelmo Frank Coelho Castro, Servidor**, em 19/06/2026, às 17:05, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Rogério Pereira de Sousa, Servidor**, em 22/06/2026, às 16:15, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Ramasio Ferreira de Melo, Servidor**, em 22/06/2026, às 16:41, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site http://sei.iftto.edu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **3223813** e o código CRC **17DF6B24**.

Referência: Processo nº 23233.013212/2026-66

SEI nº 3223813

Criado por 1899549, versão 5 por 1899549 em 19/06/2026 17:05:12.

Dedico este trabalho a **Stela O. S. Lima**, pelo apoio constante, pela paciência nos dias mais corridos e pelo carinho que tornou o caminho mais leve. Dedico também a **Deus**, em quem encontro direção e força para seguir com seriedade e propósito, consagrando a Ele meus esforços e planos.

AGRADECIMENTOS

Agradeço, antes de tudo, a **Deus**, por ter sustentado minha caminhada na graduação quando o cansaço bateu, quando a rotina apertou e quando a motivação oscilou. Em um tempo onde se questiona o valor de um diploma, eu aprendi que a formação não se resume ao papel: ela marca a disciplina, a persistência e a coragem de concluir o que se começa.

Agradeço de forma especial ao professor **Antonio Clerton Santana de Araujo**, que, ao longo da carreira dele no **IFTO**, foi uma referência de incentivo, inspiração e competência na área de tecnologia. Seu modo de ensinar, com firmeza, prática e clareza e conhecimento, ajudou a fortalecer meu interesse pela área de tecnologia em específico a programação.

Registro minha gratidão ao professor **Ancelmo Frank Coelho Castro**, que acolheu este trabalho desde a fase de artigo e, posteriormente, como TCC. Agradeço pela orientação, pela atenção sempre que possível e pelas contribuições que ajudaram a dar forma, organização e direção ao texto.

Aos **docentes do curso de Licenciatura em Computação**, deixo meu agradecimento pelo compromisso com a formação acadêmica, pelas cobranças necessárias e pelas oportunidades de aprendizagem ao longo das disciplinas. Cada aula, atividade e orientação, mesmo as mais simples das aulas contribuiu para o profissional que estou construindo. Também agradeço ao **IFTO – Campus Araguatins** pelo espaço de formação e por tudo o que a instituição representa para a comunidade.

Por fim, agradeço a todos que torceram por mim, direta ou indiretamente.

RESUMO

Os distúrbios relacionados à saúde mental, como ansiedade, depressão e estresse, impactam diretamente a qualidade de vida e demandam estratégias de cuidado acessíveis, seguras e escaláveis. Com a popularização de smartphones e da internet móvel, soluções de mHealth passaram a incorporar chatbots baseados em Inteligência Artificial (IA) como recursos de apoio conversacional, psicoeducação e incentivo a práticas de autocuidado. Entretanto, apesar do interesse crescente e de relatos de resultados promissores em estudos publicados, a adoção dessas ferramentas levanta questões relevantes sobre como avaliar, de modo consistente, sua efetividade reportada, seus limites de atuação, a segurança em situações de crise, a privacidade e os cuidados com os dados, a transparência do funcionamento, a mitigação de vieses e a qualidade da experiência do usuário.

Nesse contexto, este Trabalho de Conclusão de Curso tem como objetivo propor um framework e um checklist de avaliação para programas de IA com aplicação de chatbots voltados ao apoio à saúde mental, organizando critérios técnicos, éticos e de qualidade com base em literatura científica e documentos correlatos. Trata-se de um estudo qualitativo, exploratório e descritivo, fundamentado em pesquisa bibliográfica e na organização temática dos critérios identificados. Como resultado, apresenta-se um checklist estruturado em dimensões como a finalidade e escopo do sistema, evidências e validação reportadas, segurança e manejo de crise, privacidade e governança de dados, transparência e comunicação de limitações, vieses, equidade e acessibilidade, usabilidade e engajamento, e por fim, operação, atualização e monitoramento. Conclui-se que a sistematização desses critérios pode apoiar avaliações mais comparáveis e incentivar o desenvolvimento e a adoção responsável de chatbots em saúde mental, reduzindo riscos e ampliando a confiabilidade do uso em contextos reais.

Palavras-chave: Chatbots. Inteligência Artificial. Saúde mental. Avaliação. Segurança e privacidade.

ABSTRACT

Mental health disorders, such as anxiety, depression, and stress, directly impact quality of life and require accessible, safe, and scalable care strategies. With the widespread adoption of smartphones and mobile internet, mHealth solutions have begun to incorporate chatbots based on Artificial Intelligence (AI) as resources for conversational support, psychoeducation, and encouragement of self-care practices. However, despite growing interest and reports of promising results in published studies, the adoption of these tools raises important questions about how to consistently evaluate their reported effectiveness, their operational limitations, safety in crisis situations, privacy and data protection, transparency of operation, mitigation of biases, and the quality of the user experience.

In this context, this Final Course Project aims to propose a framework and an evaluation checklist for AI programs that utilize chatbots for mental health support, organizing technical, ethical, and quality criteria based on scientific literature and related documents. This is a qualitative, exploratory, and descriptive study grounded in a literature review and the thematic organization of the identified criteria. As a result, a checklist is presented, structured around dimensions such as the system's purpose and scope, reported evidence and validation, safety and crisis management, privacy and data governance, transparency and communication of limitations, biases, equity and accessibility, usability and engagement, and finally, operation, updating, and monitoring. It is concluded that the systematization of these criteria can support more comparable evaluations and encourage the responsible development and adoption of mental health chatbots, reducing risks and increasing the reliability of their use in real-world contexts.

Keywords: Chatbots. Artificial Intelligence. Mental health. Evaluation. Safety and privacy.

Sumário

1 INTRODUÇÃO	9
1.1 Problema de pesquisa	10
1.2 Questão norteadora	10
1.3 Objetivos	11
1.3.1 Objetivo geral	11
1.3.2 Objetivos específicos	11
1.3.3 Justificativa	11
1.4 Organização do trabalho	12
2 REFERENCIAL TEÓRICO	13
2.1 Saúde mental: contexto, demanda e barreiras de cuidado	13
2.2 mHealth e intervenções digitais em saúde mental	13
2.3 Chatbots, agentes conversacionais e fundamentos de IA (PLN)	14
2.4 Chatbots em saúde mental: panorama de evidências e desenhos de estudo	15
2.5 Avaliação de qualidade em soluções digitais de saúde mental e necessidade de checklists	16
2.6 Dimensões críticas: segurança, privacidade, transparência, vieses e ciclo de vida do software	17
3 METODOLOGIA	19
3.1 Natureza e abordagem da pesquisa	19
3.2 Procedimentos de levantamento bibliográfico/documental	19
3.3 Estratégia de análise: extração e organização temática de critérios	19
3.4 Produto do trabalho: framework e checklist de avaliação	20
3.5 Aplicação demonstrativa do checklist (escopo e limites)	21
4 RESULTADOS	22
4.1 Panorama dos estudos e soluções citados na literatura	22
4.2 Caracterização descritiva das intervenções e desfechos reportados	24
4.3 Síntese dos tipos de estudo reportados (descrição)	25
4.4 Framework proposto	25
4.5 DISCUSSÃO	28
5 CONSIDERAÇÕES FINAIS	36
5.1 Como trabalhos futuros, recomenda-se:	37
REFERÊNCIAS	39
APÊNDICE A	41

Estratégias de busca.....	41
APÊNDICE B	45
Checklist de avaliação.....	45
1. Finalidade e escopo do sistema.....	45
2. Evidência e validação reportada	46
3. Segurança e manejo de crise	48
4. Privacidade e governança de dados.....	50
5. Transparência e comunicação de limitações	52
6. Mitigação de vieses, equidade e acessibilidade	53
7. Usabilidade e engajamento.....	54
8. Operação, atualização e monitoramento (ciclo de vida).....	55

1 INTRODUÇÃO

Em nível global, transtornos e sintomas relacionados à saúde mental como ansiedade, depressão e estresse, um impacto direto na qualidade de vida. A Organização Mundial da Saúde (OMS) estimou, em 2020, que mais de 264 milhões de pessoas são afetadas pela depressão, condição associada a prejuízos funcionais e sociais relevantes. Esses quadros podem repercutir em diversos domínios, incluindo desempenho acadêmico e profissional, aumento de afastamentos por motivos de saúde e sobrecarga dos sistemas de cuidado.

Embora abordagens psicológicas e farmacológicas sejam utilizadas no tratamento de condições moderadas a graves, ainda existem desafios para ampliar o acesso a intervenções oportunas, sustentáveis e compatíveis com diferentes contextos. Estudos recentes destacam que intervenções que combinam Terapia Cognitivo-Comportamental (TCC) e estratégias de relaxamento (por exemplo, meditação e ioga) podem apresentar bons resultados em determinados cenários (AHMED et al., 2021). Contudo, na prática, barreiras como disponibilidade de profissionais, custos, deslocamento, estigma e baixa adesão ao acompanhamento contínuo contribuem para lacunas no cuidado.

Logo, o aumento do acesso à internet móvel e a ampla adoção de smartphones impulsionaram o crescimento de aplicações digitais voltadas à saúde mental, frequentemente referidas como mobile health (mHealth). A literatura aponta a existência de muitos aplicativos dessa categoria, com parcela significativa direcionada a diagnóstico, tratamento ou suporte em saúde mental. Esse cenário amplia possibilidades de autocuidado e acompanhamento, mas também evidencia um problema central onde, nem toda solução digital apresenta qualidade equivalente, e o uso de aplicativos para saúde mental requer critérios claros para avaliação de segurança, confiabilidade e adequação ao usuário.

Para apoiar decisões mais responsáveis, pesquisadores vêm propondo e discutindo instrumentos de avaliação, como listas de verificação (checklists), escalas de classificação e modelos de análise de qualidade, os quais buscam orientar profissionais e usuários na escolha e recomendação de aplicativos (DRISSI et al., 2020). Essa necessidade torna-se ainda mais relevante quando se trata de

tecnologias com interação conversacional, que podem influenciar percepções, emoções e comportamentos.

Nesse contexto, destacam-se os chatbots (ou softwares de conversação), sistemas capazes de simular diálogo humano por meio de mecanismos como processamento de linguagem natural (PLN). Em geral, esses sistemas interpretam a entrada do usuário e selecionam respostas com base em fluxos conversacionais, bancos de respostas, classificação de intenções e outros componentes de IA (HUNGERBUEHLER et al., 2021). Os chatbots podem ser acessados por sites, aplicativos móveis ou integrados a assistentes virtuais, e, em alguns casos, são representados por avatares ou agentes incorporados com elementos verbais e não verbais (NOBLE et al., 2022).

A adoção de chatbots para apoio em saúde mental é frequentemente justificada por características como disponibilidade contínua, escalabilidade, baixo custo marginal e possibilidade de reduzir barreiras de acesso. Ao mesmo tempo, trata-se de um domínio sensível, que exige cautela quanto a limites de atuação, manejo de situações de crise, privacidade e governança de dados, transparência sobre capacidades e limitações, mitigação de vieses, além de fatores de usabilidade e engajamento. Assim, ainda que a literatura reporte resultados promissores em diferentes aplicações e populações, permanece a necessidade de uma forma mais compreensível de avaliar essas soluções sob múltiplas dimensões, especialmente quando se pretende orientar desenvolvimento, seleção e uso responsável.

1.1 Problema de pesquisa

Apesar do crescimento do uso de chatbots com IA em aplicações de saúde mental, observa-se heterogeneidade nas propostas, no nível de evidência reportado e nas medidas adotadas para segurança, privacidade, transparência e qualidade de interação. Essa diversidade dificulta comparar soluções e estabelecer padrões mínimos para adoção responsável.

1.2 Questão norteadora

Diante desse cenário, este trabalho busca responder à seguinte questão:

Quais critérios técnicos, éticos e de qualidade podem orientar a avaliação de programas de Inteligência Artificial com aplicação de chatbots no auxílio ao cuidado à saúde mental?

1.3 Objetivos

1.3.1 Objetivo geral

Propor um framework e um checklist de avaliação para programas de Inteligência Artificial com aplicação de chatbots voltados ao auxílio no cuidado à saúde mental, a partir de pesquisa documental e análise crítica da literatura.

1.3.2 Objetivos específicos

- a) Apresentar fundamentos sobre saúde mental, mHealth, chatbots e IA aplicados ao apoio emocional;
- b) Identificar, na literatura, dimensões e critérios recorrentes para avaliação de soluções digitais em saúde mental (por exemplo: escopo, evidência reportada, segurança, privacidade, transparência, vieses, usabilidade e manutenção);
- c) Organizar os critérios identificados em um checklist estruturado, com itens verificáveis;
- d) Demonstrar o uso do checklist por meio de uma aplicação ilustrativa baseada em informações reportadas em estudos e descrições públicas, registrando “não informado” quando não houver dados disponíveis;
- e) Discutir limitações e implicações para o desenvolvimento e a adoção responsável de chatbots em contextos de saúde mental.

1.3.3 Justificativa

A relevância deste trabalho se apresenta em três dimensões. No âmbito acadêmico, contribui para organizar e sistematizar critérios de avaliação em uma área interdisciplinar que envolve Computação, saúde e ética. No plano social, dialoga com demandas contemporâneas por soluções acessíveis de apoio, ao mesmo tempo em que destaca riscos e responsabilidades associados a tecnologias conversacionais. Por fim, do ponto de vista técnico e prático, a proposição de um checklist pode apoiar

desenvolvedores, pesquisadores e avaliadores na identificação de requisitos mínimos e pontos críticos, favorecendo maior transparência e segurança no uso dessas ferramentas.

1.4 Organização do trabalho

Este Trabalho de Conclusão de Curso está organizado da seguinte forma: o Capítulo 2 apresenta o referencial teórico sobre saúde mental, mHealth, chatbots, fundamentos de IA (PLN, Processamento de Linguagem Natural) e aspectos éticos relevantes. O Capítulo 3 descreve a metodologia adotada, baseada em pesquisa bibliográfica e análise temática para construção do framework ou checklist. No Capítulo 4, são apresentados o panorama dos estudos e soluções, Caracterização descritiva das intervenções e desfechos reportados e discussão. Por fim, o Capítulo 5 reúne as considerações finais e limitações do estudo.

2 REFERENCIAL TEÓRICO

2.1 Saúde mental: contexto, demanda e barreiras de cuidado

Transtornos e sintomas relacionados à saúde mental como ansiedade, depressão e estresse, constituem um dos principais desafios contemporâneos em saúde pública, com impacto direto na qualidade de vida no desempenho acadêmico e profissional e na organização dos serviços de saúde. A depressão, em especial, é descrita como uma condição associada a sofrimento significativo e prejuízos funcionais, exigindo estratégias de cuidado que sejam oportunas, acessíveis e sustentáveis (WHO, 2023). Ao mesmo tempo, a procura por apoio em saúde mental enfrenta barreiras práticas e sociais: escassez de profissionais em determinadas regiões, custos, dificuldades de deslocamento, estigma, baixa disponibilidade de horários e, em muitos casos, descontinuidade do acompanhamento.

Do ponto de vista do cuidado, as intervenções psicológicas e, quando indicado, farmacológicas são abordagens consolidadas em diferentes graus de gravidade. Contudo, o desafio contemporâneo não se limita a “existirem tratamentos”, mas a como **ampliar acesso, melhorar adesão e oferecer suporte contínuo** de forma segura. Nesse cenário, recursos digitais são frequentemente discutidos como complemento ao cuidado tradicional, especialmente por favorecerem escalabilidade e disponibilidade, ainda que não substituam, por si só, o acompanhamento profissional em situações de maior complexidade.

2.2 mHealth e intervenções digitais em saúde mental

O crescimento do uso de smartphones e internet móvel impulsionou o desenvolvimento de soluções digitais em saúde, frequentemente agrupadas sob o termo *mobile health* (mHealth). No campo da saúde mental, tais soluções variam desde aplicativos de psicoeducação e autocuidado até plataformas com registro de sintomas, lembretes, exercícios estruturados e ferramentas de suporte conversacional. A literatura aponta que a popularização dessas aplicações ampliou possibilidades de acesso, mas também trouxe um problema relevante: a heterogeneidade de qualidade e de critérios de avaliação. Em outras palavras, a existência de um aplicativo não implica, necessariamente, segurança, adequação ou efetividade reportada.

A discussão sobre mHealth em saúde mental costuma enfatizar a utilidade de recursos de autoajuda e de componentes inspirados em intervenções psicológicas (por exemplo, técnicas associadas à terapia cognitivo-comportamental), relaxamento e construção de hábitos de autocuidado. Uma revisão sobre aplicativos com chatbots voltados a ansiedade e depressão descreve funcionalidades recorrentes de autocuidado, além de apontar variações importantes no tipo e no nível de suporte oferecido (AHMED et al., 2021). Esses achados reforçam que, para além da tecnologia, é necessário observar **o que o aplicativo entrega, como entrega e com quais limites**.

2.3 Chatbots, agentes conversacionais e fundamentos de IA (PLN)

Chatbots (ou agentes conversacionais) são sistemas capazes de simular diálogo com usuários por meio de texto (e, em alguns casos, voz), variando desde abordagens baseadas em regras e fluxos predefinidos até soluções que empregam técnicas de Inteligência Artificial (IA) e Processamento de Linguagem Natural (PLN). Em termos gerais, tais sistemas operam interpretando a entrada do usuário (por exemplo, intenção, entidades e contexto), selecionando ou gerando respostas e, frequentemente, conduzindo o usuário por um fluxo conversacional com objetivos específicos.

Na prática, chatbots podem ser integrados a aplicativos móveis, sites e plataformas de mensageria, e também podem aparecer como parte de assistentes virtuais mais amplos. Há propostas que incluem avatares (personagens humanoides), buscando aumentar engajamento e a sensação de presença social. Em um protocolo de chatbot voltado a navegação de recursos em saúde mental, por exemplo, observa-se a preocupação em descrever desenvolvimento, implementação e avaliação, sinalizando que sistemas conversacionais em contextos sensíveis exigem planejamento e critérios de validação (NOBLE et al., 2022). Esse tipo de literatura é importante porque aproxima o tema do ponto de vista de engenharia e avaliação de sistemas, não apenas do resultado clínico.

Além disso, chatbots aplicados a saúde mental lidam com um desafio técnico particular, onde os usuários expressam sofrimento com ambiguidades, metáforas, gírias, ironia e mensagens incompletas. Mesmo em sistemas com PLN, a

interpretação de linguagem natural pode falhar, e falhas de interpretação, em um domínio sensível, não são meramente bugs, podem gerar respostas inadequadas, frustrar o usuário ou produzir orientações imprudentes. Essa característica faz com que chatbots em saúde mental exijam avaliação que considere não apenas funcionalidade, mas também segurança, limites comunicados e tratamento de situações críticas.

2.4 Chatbots em saúde mental: panorama de evidências e desenhos de estudo

A literatura recente reúne diferentes tentativas de avaliar chatbots em saúde mental, estudos pilotos e abordagens experimentais, para melhoria de saúde mental, sugerindo potencial benefício em determinados contextos, ao mesmo tempo em que destaca questões metodológicas e necessidade de análises cuidadosas (ALRAZAQ et al., 2020). Esse tipo de síntese é relevante para este trabalho por indicar que “eficácia” é um tema em debate, porém dependente do **nível de evidência**, do **contexto de aplicação** e do **tipo de intervenção** que o chatbot realiza.

Nos estudos empíricos citados na pesquisa, observa-se diversidade de públicos e objetivos. Em estudantes universitários, por exemplo, um ensaio piloto randomizado controlado avaliou um chatbot (Tess) com foco em viabilidade, aceitabilidade e sintomas de ansiedade e depressão, reportando achados promissores em ansiedade e destacando relação entre engajamento e percepção de utilidade (KLOS et al., 2021). Durante a pandemia de COVID-19, um ensaio randomizado comparou um chatbot baseado em TCC (XiaoE) com alternativas como chatbot geral e e-book, utilizando instrumentos como PHQ-9 e medidas de usabilidade, e reportando melhora em sintomas depressivos no grupo do chatbot específico (HE et al., 2022). Outro estudo, também com universitários, comparou chatbot e biblioterapia, indicando vantagens para redução de depressão e ansiedade em determinados resultados reportados (LIU et al., 2022).

Há, ainda, propostas que investigam chatbots como complemento à terapia ou como suporte em intervenções para manejo de estresse, com grupos e arranjos distintos (DANIELI et al., 2022). Além do formato “chatbot textual”, existem intervenções com humano digital voltadas a solidão e estresse, frequentemente com

objetivos de viabilidade e aceitação, além de desfechos de bem-estar (LOVEYS et al., 2021). Em paralelo, uma linha de pesquisa discute chatbots em saúde mental também em perspectiva histórica e sociotécnica, indicando aproximações e distanciamentos ao longo do tempo (COSTA; ALMEIDA, 2021).

Em síntese, o panorama de evidências aponta que chatbots podem ser associados a melhorias em alguns desfechos em contextos específicos, mas a heterogeneidade de métodos, instrumentos e intervenções reforça a necessidade de cuidado ao interpretar resultados e de critérios claros para avaliação. Para um trabalho como este, o ponto central não é “provar” eficácia, mas estruturar **como avaliar** e **o que verificar** para interpretar, com responsabilidade, a efetividade reportada e os requisitos mínimos de segurança e qualidade.

2.5 Avaliação de qualidade em soluções digitais de saúde mental e necessidade de checklists

A avaliação de aplicativos e sistemas digitais em saúde mental envolve dimensões que extrapolam, “funciona ou não funciona”. É necessário considerar a qualidade do conteúdo, adequação ao público, consistência de recomendações, usabilidade, engajamento, privacidade, transparência e governança de dados. Nesse sentido, a literatura aponta o uso de instrumentos de avaliação e modelos para orientar análise e recomendação de aplicativos, como listas de verificação e escalas de classificação (DRISSI et al., 2020). A relevância desses instrumentos está em oferecer um caminho mais estruturado para comparar soluções e reduzir decisões baseadas apenas em popularidade, marketing ou percepção subjetiva.

No contexto de chatbots, essa necessidade se intensifica porque a interação conversacional cria uma dinâmica semelhante a relacionamento, com o sistema o chatbot “responde”, “acolhe”, “lembra”, “orienta”. Essa forma de interação pode aumentar adesão e sensação de suporte, mas também pode induzir confiança indevida se o sistema não comunica claramente limites e se não há evidência suficiente para suas promessas. Assim, a avaliação precisa observar tanto o **desempenho reportado** (quando há estudos) quanto os **requisitos de responsabilidade** que permitem uso seguro.

Por essa razão, checklists são particularmente úteis no tema do presente trabalho acadêmico de conclusão de curso, pois eles transformam dimensões abstratas (por exemplo, “segurança” e “transparência”) em itens verificáveis, permitindo que diferentes avaliadores cheguem a conclusões mais comparáveis. Além disso, o checklist oferece um meio de evidenciar lacunas de reporte: quando um estudo ou uma descrição pública não informa política de dados ou manejo de crise, isso não pode ser presumido; deve ser registrado como “não informado”.

2.6 Dimensões críticas: segurança, privacidade, transparência, vieses e ciclo de vida do software

A adoção de chatbots em saúde mental exige uma discussão sobre critérios críticos, especialmente em **cinco eixos**, que seria a segurança e manejo de crise, privacidade e governança de dados, transparência e comunicação de limitações, vieses, equidade e acessibilidade e operação, manutenção e monitoramento.

Segurança e manejo de crise são aspectos fundamentais porque usuários podem expressar risco agudo (ideação suicida, automutilação, pânico, abuso de substâncias) durante interações. Mesmo sistemas voltados a bem-estar geral podem receber mensagens de crise. Assim, critérios mínimos incluem orientar busca por serviços formais, comunicar que não se trata de atendimento emergencial e evitar respostas iatrogênicas. A literatura empírica citada na pesquisa, ao descrever intervenções e métodos de avaliação, reforça que diferentes estudos se concentram em sintomas (PHQ-9, GAD-7) e aceitação, mas nem sempre detalham mecanismos de crise com a mesma clareza, o que reforça a necessidade de padronização de reporte (HE et al., 2022; KLOS et al., 2021).

Privacidade e governança de dados são centrais porque conversas em saúde mental frequentemente carregam informações sensíveis como sintomas, histórico, gatilhos emocionais, contexto familiar e social. Logo, uma avaliação responsável deve verificar políticas de privacidade, consentimento e finalidades de uso. Mesmo quando a pesquisa não realiza auditoria jurídica, do ponto de vista técnico e ético é essencial observar se há documentação acessível e se há consistência entre o que o sistema promete e o que efetivamente coleta e armazena.

Transparência e comunicação de limitações dizem respeito à clareza sobre o que o sistema é (IA/chatbot), o que pode fazer e o que não pode fazer, e quando o usuário deve procurar ajuda profissional. Quanto mais “humanizada” a interface (por exemplo, com avatares e humanos digitais), maior a necessidade de comunicar limites para evitar antropomorfização e expectativas irreais (LOVEYS et al., 2021). Transparência também se relaciona à evidência, se um aplicativo afirma reduzir depressão, precisa indicar a base de validação ou, pelo menos, evitar promessas absolutas.

Vieses, equidade e acessibilidade são dimensões essenciais porque modelos de linguagem e fluxos conversacionais podem funcionar melhor para determinados perfis linguísticos e culturais e pior para outros. Além disso, a acessibilidade (inclusive para baixa conectividade e baixo letramento digital) influencia diretamente quem consegue se beneficiar de uma intervenção digital. Logo, é necessário avaliar linguagem, inclusão, possibilidades de personalização responsável e riscos de discriminação.

Por fim, **operação, manutenção e monitoramento** devem compor qualquer framework de avaliação, porque chatbots não são sistemas estáticos. Eles dependem de atualizações, correções, curadoria de conteúdo e monitoramento de falhas. Mudanças em plataformas (por exemplo, mensageria), atualização de modelos e ajustes em políticas de dados podem impactar o comportamento do sistema. Estudos que descrevem design, implementação e pilotos reforçam que o ciclo de vida e as decisões de engenharia influenciam tanto a experiência do usuário quanto a segurança do uso em saúde mental (HUNGERBUEHLER et al., 2021; NOBLE et al., 2022).

Diante do panorama apresentado, observa-se que a literatura oferece evidências e descrições relevantes sobre o uso de chatbots em saúde mental, mas também evidencia a necessidade de critérios estruturados para avaliação comparável e responsável. Assim, este trabalho propõe, na sequência, uma metodologia de extração e organização temática de critérios e apresenta um checklist estruturado por dimensões, visando apoiar análises acadêmicas e técnicas de programas de IA com aplicação de chatbots no auxílio ao cuidado à saúde mental.

3 METODOLOGIA

3.1 Natureza e abordagem da pesquisa

Este trabalho adota uma abordagem qualitativa, de caráter exploratório e descritivo, fundamentada em pesquisa bibliográfica e documental. O foco do estudo não é testar clinicamente uma solução específica, mas analisar criticamente o uso de chatbots com Inteligência Artificial (IA) no contexto de saúde mental e, a partir da literatura, **propor um framework e um checklist** de avaliação que apoiem o desenvolvimento e a adoção responsável dessas tecnologias.

3.2 Procedimentos de levantamento bibliográfico/documental

O levantamento bibliográfico foi realizado a partir de literatura científica e técnica relacionada a chatbots e agentes conversacionais, IA aplicada ao apoio emocional e à saúde mental, avaliação de aplicações de *mHealth* e aspectos de segurança, privacidade e ética em tecnologias digitais aplicadas à saúde.

As buscas e consultas foram conduzidas por meio de combinações de descritores e palavras-chave em língua portuguesa e inglesa, incluindo termos relacionados a saúde mental (por exemplo, *mental health*, *depression*, *anxiety*, *stress*) e a agentes conversacionais (por exemplo, *conversational agent*, *virtual agent*), com uso de operadores booleanos (AND/OR) quando aplicável. Como parte do processo de transparência metodológica, exemplos de estratégias de busca (strings) utilizadas em bases acadêmicas podem ser apresentados em apêndice, permitindo que o leitor compreenda o tipo de consulta realizada.

3.3 Estratégia de análise: extração e organização temática de critérios

Após a seleção das fontes bibliográficas e documentais, realizou-se uma leitura analítica voltada a identificar **como a literatura descreve, avalia e problematiza** o uso de chatbots com Inteligência Artificial (IA) no apoio à saúde mental. A leitura teve foco em trechos que apresentavam recomendações explícitas (boas práticas, requisitos e cuidados), formas de avaliação reportadas (indicadores, instrumentos, métricas de usabilidade/aceitabilidade e resultados) e riscos e limitações (privacidade, segurança, transparência, vieses e manejo de situações sensíveis).

A partir dessa leitura, os elementos identificados foram registrados e organizados como **critérios preliminares**, descritos inicialmente em linguagem livre e, em seguida, consolidados por aproximação de significado. Nessa etapa, buscou-se reduzir repetições e evitar sobreposição de itens, reunindo critérios semelhantes sob categorias mais abrangentes. Ao final, essas categorias foram tratadas como **dimensões** do framework proposto, de forma a permitir uma visão estruturada do que deve ser observado na avaliação de soluções conversacionais em saúde mental.

Com o framework delineado, procedeu-se à redação dos itens do checklist em formato **verificável**, preferencialmente como perguntas objetivas (por exemplo: “há política de privacidade acessível?”; “o sistema comunica limites e recomenda busca de ajuda em crise?”). Para diferenças significativas tornar o instrumento aplicável mesmo quando a fonte consultada não apresenta todas as informações, adotou-se uma escala simples de classificação por item, contemplando a possibilidade de ausência de reporte, como, **Atende**, **Parcialmente atende**, **Não atende** e **Não informado**. Por fim, realizou-se uma **aplicação demonstrativa** do checklist em exemplos descritos na literatura, registrando-se “Não informado” nos casos em que as publicações não traziam elementos suficientes para julgamento, o que também contribui para evidenciar lacunas recorrentes de documentação e transparência.

3.4 Produto do trabalho: framework e checklist de avaliação

O produto central deste Trabalho de Conclusão de Curso é um **checklist estruturado** para avaliação de programas de IA com aplicação de chatbots voltados ao apoio à saúde mental. O instrumento foi elaborado com a intenção de ser útil em contexto acadêmico, servindo tanto para orientar análises críticas quanto para apoiar discussões sobre requisitos mínimos de qualidade e responsabilidade no desenvolvimento e adoção dessas soluções.

O checklist foi organizado em **dimensões**, cada uma representando um conjunto de aspectos que se mostraram recorrentes e relevantes na literatura consultada. Em linhas gerais, o instrumento contempla:

- (a) finalidade e escopo do sistema (o que promete e quais limites declara);

- (b) evidências e validação reportadas (existência e tipo de avaliação descrita, população-alvo e desfechos);
- (c) segurança e manejo de crise (orientações, encaminhamentos e limites em situações críticas);
- (d) privacidade e governança de dados (coleta, finalidade, armazenamento e consentimento);
- (e) transparência e comunicação de limitações (clareza de que se trata de IA, responsabilidades e advertências);
- (f) mitigação de vieses, equidade e acessibilidade (inclusão, linguagem e adequação a diferentes perfis de usuários);
- (g) usabilidade e engajamento (experiência do usuário, aderência e personalização responsável); e
- (h) operação, atualização e monitoramento (manutenção, auditoria e melhoria contínua).

O checklist, com itens verificáveis e escala de avaliação, é apresentado no **Apêndice (pagina 41)**.

3.5 Aplicação demonstrativa do checklist (escopo e limites)

A aplicação do checklist possui caráter **demonstrativo**, com o objetivo de mostrar como o instrumento pode ser utilizado para analisar chatbots descritos em estudos e publicações (o checklist se encontra no APÊNDICE B). Essa aplicação não pretende substituir avaliações clínicas, auditorias técnicas formais ou certificações, mas sim **evidenciar lacunas de reporte** (quando informações essenciais não são apresentadas) e apoiar uma discussão crítica sobre requisitos mínimos para uso responsável.

4 RESULTADOS

4.1 Panorama dos estudos e soluções citados na literatura

A partir do levantamento descrito na Metodologia, foram identificados estudos recentes que descrevem o uso de chatbots e agentes conversacionais baseados em Inteligência Artificial (IA) no contexto de apoio à saúde mental. Nesta seção, apresenta-se um panorama descritivo de exemplos recorrentes na literatura, destacando o objetivo de cada estudo, o tipo de solução utilizada e os focos de intervenção reportados (depressão, ansiedade, estresse, solidão e bem-estar).

Ressalta-se que os estudos apresentados a seguir são utilizados como exemplos representativos para contextualizar funcionalidades e formas de avaliação descritas na literatura, servindo de base para a proposição do framework e do checklist apresentados posteriormente.

Quadro 1 — Estudos citados na literatura e seus objetivos (exemplos)

Autor (ano)	Título	Principais objetivos
Klos et al. (2021)	<i>Artificial intelligence-based chatbot for anxiety and depression in university students: pilot randomized controlled trial</i>	Avaliar viabilidade, aceitabilidade e potencial impacto do chatbot Tess sobre sintomas de depressão e ansiedade em estudantes universitários.
He et al. (2022)	<i>Mental Health Chatbot for Young Adults With Depressive Symptoms During the COVID-19 Pandemic: Single-Blind, Three-Arm Randomized Controlled Trial</i>	Testar eficácia clínica e desempenho não clínico de um chatbot (ex.: XiaoE) baseado em TCC para jovens adultos com sintomas depressivos durante a pandemia.

Autor (ano)	Título	Principais objetivos
Liu et al. (2022)	<i>Using AI chatbots to provide self-help depression interventions for university students: A randomized trial of effectiveness</i>	Comparar intervenção por chatbot com biblioterapia como estratégia de autoajuda em contexto universitário.
Danieli et al. (2022)	<i>Assessing the Impact of Conversational Artificial Intelligence in the Treatment of Stress and Anxiety in Aging Adults: Randomized Controlled Trial</i>	Avaliar contribuição do agente conversacional TeO como apoio em intervenção voltada ao manejo de estresse e ansiedade em adultos/idosos.
Loveys et al. (2021)	<i>A digital human for delivering a remote loneliness and stress intervention....: Randomized pilot trial</i>	Avaliar viabilidade e aceitabilidade de intervenção remota mediada por humano digital (Dh) voltada a solidão, estresse e bem-estar.
Stephanie G. S. et al. (2022)	<i>The Effect of Mental Health App Customization on Depressive Symptoms in College Students: Randomized Controlled Trial</i>	Investigar o efeito da personalização (ex.: avatar) em aplicativo conversacional (AirHeart) sobre sintomas depressivos e adesão.

4.2 Caracterização descritiva das intervenções e desfechos reportados

Os estudos citados apresentam diversidade de propostas, incluindo chatbots de texto, agentes conversacionais com suporte a intervenções inspiradas em TCC (Terapia Cognitivo-Comportamental) e “humanos digitais” voltados a interação e engajamento. Em comum, observa-se a finalidade de oferecer suporte emocional, psicoeducação, orientação de autocuidado ou acompanhamento de sintomas, com desfechos predominantemente relacionados a **depressão** e **ansiedade**, além de **estresse**, **solidão** e **bem-estar psicológico** em parte das publicações.

Para sistematizar as soluções descritas, o Quadro 2 apresenta a relação entre os trabalhos citados, o nome da IA/chatbot e o tipo de condição-alvo abordada.

Quadro 2 — Soluções de IA citadas e foco de intervenção (exemplos)

Autor (ano)	Nome da IA	Tipo de condição-alvo
Klos et al. (2021)	Tess	Ansiedade e depressão
He et al. (2022)	XiaoE e Xiaoi	Sintomas depressivos
Liu et al. (2022)	XiaoNan	Ansiedade e depressão
Danieli et al. (2022)	TeO	Ansiedade, depressão, estresse e bem-estar geral
Loveys et al. (2021)	Dh (humano digital)	Solidão, estresse e bem-estar
Stephanie G. S. et al. (2022)	AirHeart	Estresse e depressão

4.3 Síntese dos tipos de estudo reportados (descrição)

Quanto aos desenhos de pesquisa descritos nas publicações citadas, observam-se estudos experimentais e ensaios randomizados (incluindo ensaios piloto). Essa predominância de pesquisas com intervenção é relevante para o tema da pesquisa por indicar que parte da literatura busca mensurar efeitos ou avaliar viabilidade e aceitabilidade de chatbots no apoio à saúde mental. Ao mesmo tempo, a diversidade de populações, abordagens e formas de reporte reforça a necessidade de critérios padronizados para avaliar qualidade, segurança, evidência reportada e adequação de uso — aspecto que fundamenta a proposta do checklist apresentada nos próximos tópicos.

4.4 Framework proposto

O framework proposto neste Trabalho de Conclusão de Curso organiza, de forma estruturada, os principais critérios identificados na literatura para avaliar programas de Inteligência Artificial (IA) com aplicação de chatbots voltados ao apoio à saúde mental. Neste trabalho, entende-se por framework um modelo conceitual de avaliação, isto é, um conjunto de dimensões que sintetizam aspectos técnicos, éticos e de qualidade necessários para analisar essas soluções de maneira comparável e orientada por evidências reportadas. A motivação para a construção do framework decorre do fato de que a literatura apresenta intervenções heterogêneas e níveis de reporte distintos, o que torna insuficiente avaliar chatbots apenas por percepções de utilidade ou por afirmações gerais de efetividade.

A partir da análise temática descrita na Metodologia, os critérios recorrentes foram agrupados em oito dimensões, que funcionam como camadas de avaliação. Em conjunto, tais dimensões permitem observar desde o posicionamento do sistema (o que ele promete e quais limites declara) até aspectos críticos de segurança, privacidade, transparência, equidade e manutenção do software. O Quadro 3 sintetiza essas dimensões e sua finalidade no processo avaliativo.

Quadro 3 — Dimensões e finalidade de avaliação

Dimensão	Finalidade (o que observar)
----------	-----------------------------

Finalidade e escopo do sistema	verifica se o chatbot explicita o objetivo de uso, público-alvo e limites de atuação, evitando promessas incompatíveis com suas capacidades.
Evidências e validação reportadas	observa se há estudos, relatórios ou descrições metodológicas que sustentem resultados e afirmações do sistema, bem como quais desfechos são reportados.
Segurança e manejo de crise	identifica se o sistema prevê orientação adequada em situações críticas, reconhece limites em emergências e oferece encaminhamentos apropriados.
Privacidade e governança de dados	avalia se há clareza sobre coleta, finalidade, armazenamento e compartilhamento de dados, considerando a sensibilidade das informações em saúde mental.
Transparência e comunicação de limitações	verifica se o usuário é informado de que interage com IA, quais são as limitações do sistema e quais responsabilidades cabem ao usuário e ao serviço.
Mitigação de vieses, equidade e acessibilidade	considera se a solução atende a diferentes perfis de usuários, reduzindo riscos de exclusão, discriminação e inadequação cultural/linguística.
Usabilidade e engajamento	observa se o uso é intuitivo, se há aderência e se estratégias de engajamento são aplicadas de forma responsável, sem indução indevida.
Operação, atualização e monitoramento	verifica se existem práticas mínimas de manutenção, registro de incidentes,

	melhoria contínua e responsabilidade institucional pelo funcionamento.
--	--

O checklist apresentado no Apêndice B é a operacionalização desse framework. Enquanto o framework descreve o que avaliar (dimensões e seus focos), o checklist define como avaliar (itens verificáveis), permitindo registrar de forma objetiva se determinada informação está presente e em que nível. Por isso, cada dimensão do framework é desdobrada em perguntas de verificação, acompanhadas de criticidade e de uma escala simples de classificação (Atende, Parcialmente atende, Não atende e Não informado). A inclusão do marcador **não informado** é especialmente relevante neste contexto, pois permite evidenciar lacunas de reporte, quando um estudo ou descrição pública não apresenta informação suficiente (por exemplo, política de privacidade ou manejo de crise), o instrumento não assume conformidade por inferência, registrando a ausência de dados.

Dessa forma, o framework fornece a estrutura conceitual para interpretação crítica, e o checklist funciona como instrumento prático para registrar evidências disponíveis e sustentar comparações. Com base nessa organização, a seção seguinte discute os estudos e soluções citados à luz das dimensões propostas, destacando implicações para adoção responsável e para desenvolvimento de chatbots em saúde mental.

4.5 DISCUSSÃO

A literatura recente sobre o uso de chatbots e agentes conversacionais em saúde mental revela um campo em franca expansão, marcado por propostas tecnológicas heterogêneas, públicos distintos e objetivos que variam desde suporte emocional básico até intervenções inspiradas em terapia cognitivo-comportamental (TCC). Nesse cenário, apesar de diferentes estudos relatarem resultados promissores para sintomas de depressão, ansiedade, estresse, solidão e bem-estar psicológico, observa-se que a variabilidade de desenhos de estudo, métricas utilizadas e características dos sistemas dificulta comparações diretas e conclusões generalizáveis. Por isso, uma discussão madura sobre “eficácia” precisa ir além da pergunta “funciona?” e incorporar questões como, **em quais condições, para quais usuários, com qual transparência, com quais salvaguardas e com qual nível de evidência reportada.**

Um primeiro ponto que se destaca é que chatbots podem assumir papéis diferentes dentro de um aplicativo de saúde mental. Alguns sistemas funcionam como “companheiros” de conversa, oferecendo mensagens de apoio e psicoeducação; outros são projetados como “coaches” estruturados, conduzindo exercícios e tarefas; e há ainda agentes mais elaborados, como humanos digitais, que tentam aumentar o engajamento e a sensação de presença social. Essa diferença de intenção e arquitetura importa porque altera expectativas e riscos. Um chatbot que apenas oferece dicas e registro de humor demanda critérios distintos de um sistema que se apresenta como ferramenta terapêutica estruturada, sobretudo quando atua em um domínio sensível, em que respostas inadequadas podem amplificar sofrimento ou retardar a busca por ajuda profissional.

No estudo de Klos et al. (2021), por exemplo, o chatbot Tess é descrito como um sistema de conversas por texto que oferece lembretes, conteúdo psicoeducativo e respostas de apoio emocional com base no que os usuários relatam. O desenho do Tess evidencia uma característica central de muitos chatbots de saúde mental, que seria a tentativa de produzir uma interação “humanizada”, por meio de linguagem amigável e elementos expressivos (incluindo emojis). Do ponto de vista de experiência do usuário, isso pode reduzir barreiras iniciais e favorecer adesão, especialmente em públicos como estudantes universitários, que frequentemente buscam soluções

rápidas, discretas e acessíveis. Ao mesmo tempo, a “humanização” comunicacional exige um cuidado adicional com **transparência**. Quanto mais humano o sistema parece, maior a chance de o usuário atribuir ao chatbot competências que ele não possui, elevando o risco de dependência, de confiança indevida e de frustração quando a conversa falha em momentos críticos.

Ainda no trabalho de Klos et al. (2021), a avaliação de sintomas foi realizada com instrumentos amplamente utilizados em saúde mental, como o PHQ-9 (para depressão) e o GAD-7 (para ansiedade). A presença de métricas padronizadas é positiva porque permite diálogo com a prática clínica e com a literatura comparável, mas também sinaliza uma limitação frequente, trata-se, em grande parte, de medidas de autorrelato, influenciadas por fatores como motivação, expectativa, compreensão das perguntas e contexto emocional do respondente. No mesmo estudo, o engajamento medido pelo volume de mensagens trocadas e o feedback subjetivo dos participantes oferecem pistas sobre aceitabilidade e percepção de utilidade. É relatado que, após oito semanas, houve média elevada de mensagens trocadas e predominância de feedback positivo, associada ao nível de interação. Esse tipo de achado é relevante porque, em intervenções digitais, o efeito do sistema pode estar intimamente ligado ao engajamento, não basta o chatbot existir, é necessário que o usuário o utilize de modo suficientemente consistente para que qualquer intervenção psicoeducativa ou exercício produza resultados observáveis.

Contudo, mesmo quando se relata melhora em ansiedade ou percepção de apoio, a literatura sugere cautela para evitar extrapolações. No caso de Klos et al. (2021), há indicação de redução moderada em ansiedade no grupo de intervenção em comparação ao controle, mas não foram observadas diferenças expressivas para depressão naquele contexto. Essa discrepância é útil para a discussão porque revela que chatbot em saúde mental não é uma categoria uniforme. Diferentes sintomas podem responder de modos distintos, e intervenções mais leves podem ter efeito perceptível para ansiedade situacional, mas menor impacto para quadros depressivos mais persistentes, que frequentemente exigem acompanhamento clínico estruturado, suporte social e, em alguns casos, farmacoterapia.

Essas observações ajudam a identificar um primeiro conjunto de critérios que um checklist precisa contemplar qual é o escopo do chatbot (apoio geral, triagem,

intervenção estruturada), quais sintomas pretende impactar, quais evidências reporta para essas promessas e quais são seus limites declarados. Em outras palavras, discutir eficácia exige avaliar se o sistema se posiciona com responsabilidade e se a narrativa do produto não ultrapassa o que a literatura sustenta.

O estudo de He et al. (2022) amplia esse debate ao investigar um chatbot de saúde mental (XiaoE) baseado em TCC (Terapia Cognitivo-Comportamental), comparando-o a um chatbot geral (Xiaoai) e a um e-book. Esse tipo de comparação é particularmente interessante porque separa o efeito, **conversacional** (interagir com um agente) do efeito **ter conteúdo** (ler material psicoeducativo) e do efeito **ser uma intervenção estruturada** (com princípios de Terapia Cognitivo-Comportamental). Os autores utilizam novamente o PHQ-9, o que mantém consistência com a literatura, e incluem instrumentos adicionais voltados à relação terapêutica e à experiência do usuário, como o Working Alliance Questionnaire (WAQ), além de medidas de usabilidade (UMUX-LITE) e aceitabilidade (AS). A presença dessas métricas reforça um ponto importante: em saúde mental, resultados não se resumem a redução de sintomas; fatores como aliança de trabalho, conforto, usabilidade e aceitação podem ser determinantes para adesão e manutenção do uso.

Os resultados reportados por He et al. (2022) indicam redução maior de sintomas depressivos no grupo que utilizou o XiaoE, em comparação ao controle, com variações de tamanho de efeito em diferentes momentos. Ainda que esse achado seja promissor, ele também chama atenção para um desafio típico de intervenções digitais: a estabilidade do efeito ao longo do tempo. Melhoras iniciais podem se reduzir quando a novidade passa, quando o usuário perde motivação ou quando situações externas (como estressores socioeconômicos e eventos de crise) se intensificam. Portanto, ao avaliar um chatbot, um critério relevante é a existência de estratégia de manutenção de engajamento (por exemplo, lembretes, metas, progressão de conteúdo), bem como clareza sobre para quem a ferramenta é adequada e em que intensidade ela deve ser utilizada.

Além disso, o próprio desenho de He et al. (2022) evidencia uma dimensão que deve estar explícita em qualquer framework de avaliação: **o chatbot realiza intervenção, suporte ou navegação de recursos?** Há uma diferença prática entre um sistema que conduz exercícios terapêuticos estruturados e outro que apenas

orienta o usuário a procurar serviços e materiais. Essa distinção afeta a responsabilidade do sistema, a necessidade de supervisão clínica, e, sobretudo, o conjunto de salvaguardas em caso de piora de sintomas. Em sistemas voltados a saúde mental, um ponto crítico é o **manejo de crise**: mesmo que o objetivo primário seja psicoeducação ou autocuidado, o usuário pode revelar ideação suicida, automutilação ou risco agudo em algum momento. Nesse caso, critérios mínimos incluem detecção de linguagem de crise (ainda que simples), respostas não iatrogênicas, encaminhamento para recursos formais e comunicação clara de que o chatbot não substitui atendimento emergencial.

Já o estudo de Liu et al. (2022) introduz um aspecto relevante para a Licenciatura em Computação: o uso de infraestrutura e tecnologias de conversação como o framework aberto RASA e plataformas amplamente utilizadas (no caso, WeChat). A adoção de ferramentas abertas pode facilitar replicação e evolução do sistema, mas também impõe desafios de governança como versão do modelo, atualização de intents, curadoria de respostas, monitoramento de falhas e controle de qualidade. Liu et al. (2022) comparam a intervenção mediada por chatbot com biblioterapia, indicando vantagens para redução de depressão e ansiedade, com efeitos reportados em instrumentos como PHQ-9 e GAD-7. Ainda que o resultado seja promissor, ele reforça a necessidade de avaliar o **conteúdo** que o chatbot entrega. Muitas vezes, o efeito não vem do chat, mas do conteúdo psicoeducativo ou dos exercícios estruturados (por exemplo, técnicas de reestruturação cognitiva, diário de pensamento, exercícios de respiração). Assim, um item importante do checklist é verificar se o chatbot explicita **quais técnicas baseadas em evidências** orientam suas interações e se o conteúdo é consistente com boas práticas (evitando recomendações genéricas, moralizantes ou potencialmente culpabilizadoras).

O estudo de Danieli et al. (2022) oferece um exemplo ainda mais complexo ao combinar psicoterapia tradicional com apoio de um agente conversacional (TeO), incluindo múltiplos grupos (terapia remota, terapia + agente, somente agente, e controle). Esse tipo de desenho é valioso porque explora uma hipótese realista, onde chatbots podem ser mais úteis como **complemento** do cuidado do que como substituto. Em serviços de saúde e em contextos com pouca disponibilidade de profissionais, um agente conversacional pode ampliar contato entre sessões, reforçar

tarefas e registrar progresso, desde que esteja bem integrado ao plano terapêutico. Ao mesmo tempo, os resultados reportados ao longo do tempo parecem variar entre grupos e momentos, e são descritas tendências de oscilação de estresse e bem-estar. Esse comportamento não linear ao longo do tempo é comum em saúde mental e deve ser entendido como parte do fenômeno, não como falha isolada do estudo. Para a avaliação de chatbots, isso sugere que o checklist deve considerar critérios de **monitoramento longitudinal**, não apenas efeito final. O sistema registra evolução? Identifica piora? Sugere procura de ajuda quando há agravamento? Há mecanismos para reduzir risco de abandono?

Outro ponto importante em Danieli et al. (2022) é a noção de satisfação com o tratamento e percepção do usuário. Mesmo quando uma intervenção não mostra superioridade clara em todas as medidas, pode haver ganho em engajamento, sensação de suporte e adesão a práticas de autocuidado. Em tecnologia aplicada à saúde, esses desfechos secundários podem ser a diferença entre o usuário abandonar uma intervenção em duas semanas ou sustentar um hábito de monitoramento e reflexão por meses. Portanto, além de medir sintomas, é crucial que uma avaliação considere **aceitabilidade, usabilidade, aderência e qualidade da interação**.

No estudo de Loveys et al. (2021), a intervenção é mediada por um humano digital (um agente com aparência, ou seja, uma representação humana) e envolve exercícios comportamentais e de psicologia positiva por um período curto. Os autores relatam que os efeitos podem variar conforme o desfecho, a solidão percebida pode não mudar significativamente, enquanto estresse e bem-estar psicológico podem apresentar tendências em favor do grupo de intervenção. Esses achados lembram que diferentes constructos psicológicos respondem a ritmos diferentes. Solidão, por exemplo, tem forte componente social e contextual, um agente digital pode ajudar a reduzir sofrimento momentâneo, mas dificilmente substitui vínculos sociais reais. Assim, um item de checklist relevante é verificar se o chatbot evita promessas irreais, reconhece limites do autocuidado digital e orienta o usuário a estratégias complementares (contato com rede de apoio, hábitos sociais, busca de serviços).

Além disso, a adoção de humanos digitais ou avatares levanta questões próprias: que tipo de persona é construída? Ela se adapta ao usuário? Há risco de

reforçar estereótipos? Existe acessibilidade para pessoas com deficiências (visual, auditiva, cognitiva), baixa conectividade ou baixa familiaridade tecnológica? Em um país com desigualdades marcantes de acesso e letramento digital, essas dimensões são particularmente relevantes. Assim, equidade e acessibilidade não devem ser tratadas como “detalhe”, mas como requisitos centrais para avaliação responsável.

O trabalho de Stephanie G. S. et al. (2022) acrescenta outro ângulo, onde o efeito da personalização (por exemplo, do avatar) sobre sintomas e adesão em um aplicativo baseado em TCC (Terapia Cognitivo-Comportamental). Personalização pode aumentar identificação e engajamento, mas também pode operar apenas como camada estética sem melhorar, de fato, a qualidade terapêutica do conteúdo. Portanto, ao avaliar soluções desse tipo, torna-se necessário diferenciar personalização superficial (cores, avatar) de personalização clinicamente relevante (adaptação de conteúdo, metas e intervenções ao perfil do usuário). Esse ponto se traduz em critérios onde o sistema explica quais dados usa para personalizar? Há consentimento? O usuário pode desligar personalização? Como a personalização afeta privacidade e risco de inferência de dados sensíveis?

A partir do conjunto desses estudos, é possível consolidar uma leitura crítica que fundamenta diretamente o checklist proposto neste Trabalho de conclusão de curso. Em termos de **benefícios potenciais**, a literatura sugere que chatbots podem:

- a) Ampliar acesso ao suporte emocional em horários e locais variados;
- b) Reduzir barreiras relacionadas a estigma e medo de julgamento;
- c) Oferecer conteúdo psicoeducativo e exercícios estruturados;
- d) Apoiar automonitoramento de sintomas e hábitos;
- e) Contribuir para engajamento em práticas de autocuidado.

Tais benefícios, porém, dependem fortemente de fatores como qualidade do design conversacional, adequação cultural e linguística, clareza de escopo e presença de mecanismos de segurança.

Em termos de **limitações e riscos**, destacam-se ao menos seis pontos que precisam estar explícitos em qualquer avaliação séria de chatbots para saúde mental. **Primeiro**, a dependência de medidas de autorrelato e períodos de acompanhamento

relativamente curtos em parte da literatura pode limitar inferências sobre sustentabilidade do efeito. **Segundo**, a heterogeneidade de intervenções (conteúdos, frequências, duração) dificulta atribuir melhorias exclusivamente ao chatbot, pois muitas vezes o efeito está no conteúdo terapêutico, na estrutura do programa ou no suporte adicional. **Terceiro**, há desafios técnicos inerentes ao Processamento de Linguagem Natural (PLN): usuários expressam sofrimento de forma ambígua, com ironia, gírias, variações culturais e mensagens incompletas. Um sistema que entende errado pode responder de modo inadequado, invalidante ou perigoso. **Quarto**, privacidade e governança de dados são críticas, onde conversas sobre saúde mental são altamente sensíveis, e o usuário frequentemente não tem clareza sobre o que é coletado, por quanto tempo, com que finalidade e com quais terceiros. **Quinto**, transparência é indispensável, por que o usuário deve saber que está conversando com IA, quais são as limitações e quando procurar ajuda profissional. **Sexto**, manejo de crise não é opcional, mesmo sistemas voltados ao bem-estar geral podem receber mensagens indicando risco agudo.

Além disso, do ponto de vista de engenharia de software, um chatbot em saúde mental não pode ser tratado como aplicativo comum “lançou e esqueceu”. Há necessidade de manutenção contínua, auditoria de respostas, monitoramento de falhas e atualização de conteúdo. Mudanças em plataformas (APIs), políticas de privacidade, comportamento do usuário e evolução de modelos de IA podem impactar diretamente o funcionamento. Logo, um framework de avaliação coerente deve incluir critérios sobre **operação e ciclo de vida**: existe monitoramento? Há canal de suporte? Há registro de incidentes? Como o sistema lida com respostas incorretas? Há mecanismo de escalonamento para humano em casos críticos?

Diante dessas considerações, esta discussão sustenta a proposta do checklist como uma forma de organizar, em linguagem verificável, o que frequentemente aparece disperso nos estudos, evidência reportada (quando existe), características do sistema, métricas de experiência do usuário, e salvaguardas éticas e de segurança. Em síntese, a contribuição prática deste TCC está em transformar achados e preocupações recorrentes da literatura — exemplificados por estudos como Klos et al. (2021), He et al. (2022), Liu et al. (2022), Danieli et al. (2022), Loveys et al. (2021) e Stephanie G. S. et al. (2022) em um conjunto de critérios que pode orientar tanto a

análise acadêmica quanto decisões de desenvolvimento e adoção responsável de chatbots no apoio ao cuidado em saúde mental.

5 CONSIDERAÇÕES FINAIS

Este Trabalho de Conclusão de Curso teve como propósito discutir criticamente o uso de chatbots baseados em Inteligência Artificial (IA) no apoio ao cuidado em saúde mental e, principalmente, **propor um framework e um checklist de avaliação** que organizem critérios técnicos, éticos e de qualidade para análise dessas soluções. Ao longo do trabalho, observou-se que a literatura recente tem atribuído aos chatbots potencial para ampliar acesso a suporte emocional, psicoeducação e estratégias de autocuidado, especialmente em cenários nos quais há barreiras de acesso a serviços presenciais, limitações de tempo, estigma e necessidade de disponibilidade contínua. Ainda assim, a discussão mostrou que a adoção de tecnologias conversacionais em saúde mental exige cuidado redobrado, porque a promessa de conveniência e escala convive com riscos e lacunas de reporte que, se ignorados, podem comprometer a segurança do usuário e a credibilidade das intervenções digitais.

No que se refere aos **achados centrais** da análise bibliográfica, a literatura citada aponta que determinadas soluções conversacionais são associadas a melhora de indicadores como ansiedade, estresse, sintomas depressivos e bem-estar em contextos específicos, bem como a medidas de aceitabilidade e engajamento. Entretanto, também se evidenciou que tais resultados são apresentados em condições variadas (públicos diferentes, durações distintas, instrumentos diversos e níveis de detalhamento metodológico heterogêneos). Assim, do ponto de vista acadêmico e prático, não é suficiente afirmar que “chatbots funcionam” ou “chatbots não funcionam”, o que se impõe é a necessidade de perguntar como essas soluções foram desenhadas, quais limites declaram, que evidências reportam, quais salvaguardas oferecem, e como são mantidas e monitoradas ao longo do tempo.

A contribuição mais objetiva deste TCC consiste na **transformação dessas preocupações recorrentes da literatura em um instrumento de avaliação**. O framework (ou checklist) proposto organiza dimensões que tendem a determinar a qualidade e a adequação de chatbots em saúde mental, tais como a finalidade e escopo (o que o sistema promete e o que não promete), evidências e validação reportadas, segurança e manejo de crise, privacidade e governança de dados, transparência e comunicação de limitações, vieses, equidade e acessibilidade, usabilidade e engajamento, e também operação, atualização e monitoramento. Ao

estruturar critérios verificáveis, o checklist visa apoiar análises mais consistentes e comparáveis, reduzindo avaliações baseadas apenas em percepção subjetiva, em marketing do aplicativo ou em descrições incompletas.

Do ponto de vista **prático**, o checklist pode ser útil em pelo menos três situações como guia de requisitos para desenvolvimento de chatbots (engenharia de software com responsabilidade), como roteiro para análise de soluções existentes em trabalhos acadêmicos, projetos institucionais e iniciativas de extensão e como mecanismo de transparência, incentivando que propostas de chatbots explicitem limitações, políticas de dados, estratégias de segurança e evidências que sustentam suas recomendações. Em um domínio sensível como saúde mental, essa transparência não é um “detalhe de documentação”, ela é parte da segurança.

Como limitações, reconhece-se que este trabalho adotou abordagem bibliográfica e documental e não realizou validação empírica própria do checklist com usuários finais ou equipes clínicas. Além disso, a aplicação apresentada tem caráter demonstrativo, dependente do nível de informação disponibilizado nas fontes consultadas, o que, por si só, evidencia um problema recorrente. Muitas soluções e estudos não reportam de forma completa aspectos como governança de dados, políticas de privacidade, mecanismos de manejo de crise e procedimentos de monitoramento pós-implantação. Dessa forma, parte dos itens do checklist pode resultar em marcações do tipo “não informado”, não por falha do instrumento, mas por ausência de reporte nas fontes.

5.1 Como trabalhos futuros, recomenda-se:

(a) validar o checklist com especialistas de áreas complementares (Computação, Psicologia, Saúde Coletiva e Direito/LGPD), buscando melhorar clareza, abrangência e prioridade dos critérios;

(b) aplicar o checklist a um conjunto maior de soluções, construindo um panorama comparativo;

(c) estabelecer pesos ou níveis de criticidade (por exemplo, requisitos mínimos obrigatórios versus critérios desejáveis);

(d) realizar estudos empíricos de usabilidade e segurança, incluindo cenários de crise e análise de respostas inadequadas;

(e) quando houver um corpus de ensaios clínicos devidamente levantado, empregar instrumentos formais de avaliação metodológica (como ferramentas de risco de viés aplicáveis a ensaios) para qualificar ainda mais o nível de confiança em resultados de efetividade reportados.

Conclui-se, portanto, que chatbots baseados em IA apresentam **potencial** para compor estratégias de apoio em saúde mental, mas seu uso responsável depende de critérios claros e verificáveis que considerem, simultaneamente, evidência reportada, qualidade técnica, ética, privacidade, segurança e manutenção. Nesse sentido, o checklist proposto busca contribuir para que a discussão sobre eficácia deixe de ser apenas uma expectativa e passe a ser tratada como uma **propriedade avaliável**, sustentada por requisitos e transparência compatíveis com a sensibilidade do contexto em que essas tecnologias operam.

REFERÊNCIAS

AHMED, Arfan *et al.* A review of mobile chatbot apps for anxiety and depression and their self-care features. **Computer Methods and Programs in Biomedicine Update**, v. 1, p. 100012, 2021.

ALRAZAQ, Alaa Ali *et al.* Effectiveness and safety of using chatbots to improve mental health: systematic review and meta-analysis. **Journal of medical Internet research**, v. 22, n. 7, p. e16021, 2020.

DANIELI, Morena *et al.* Assessing the Impact of Conversational Artificial Intelligence in the Treatment of Stress and Anxiety in Aging Adults: Randomized Controlled Trial. **JMIR Mental Health**, v. 9, n. 9, p. e38067, 2022.

DRISSI, Nidal *et al.* An analysis on self-management and treatment-related functionality and characteristics of highly rated anxiety apps. **International journal of medical informatics**, v. 141, p. 104243, 2020.

HE, Yuhao *et al.* Mental Health Chatbot for Young Adults With Depressive Symptoms During the COVID-19 Pandemic: Single-Blind, Three-Arm Randomized Controlled Trial. **Journal of Medical Internet Research**, v. 24, n. 11, p. e40719, 2022.

HUNGERBUEHLER, Ines *et al.* Chatbot-based assessment of employees' mental health: Design process and pilot implementation. **JMIR Formative Research**, v. 5, n. 4, p. e21678, 2021.

KLOS, Maria Carolina *et al.* Artificial intelligence–based chatbot for anxiety and depression in university students: pilot randomized controlled trial. **JMIR formative research**, v. 5, n. 8, p. e20678, 2021.

LEE, Sang Min *et al.* Psychological impact of the 2015 MERS outbreak on hospital workers and quarantined hemodialysis patients. **Comprehensive psychiatry**, v. 87, p. 123-127, 2018.

LIU, Hao *et al.* Using AI chatbots to provide self-help depression interventions for university students: A randomized trial of effectiveness. **Internet Interventions**, v. 27, p. 100495, 2022.

LOVEYS, Kate *et al.* A digital human for delivering a remote loneliness and stress intervention to at-risk younger and older adults during the COVID-19 pandemic: Randomized pilot trial. **JMIR mental health**, v. 8, n. 11, p. e31586, 2021.

NOBLE, Jasmine M. *et al.* Developing, Implementing, and Evaluating an Artificial Intelligence–Guided Mental Health Resource Navigation Chatbot for Health Care Workers and Their Families During and Following the COVID-19 Pandemic: Protocol for a Cross-sectional Study. **JMIR Research Protocols**, v. 11, n. 7, p. e33717, 2022.

Stephanie G. S *et al.* The Effect of Mental Health App Customization on Depressive Symptoms in College Students: Randomized Controlled Trial. **JMIR Mental Health**, v. 9, n. 8, p. e39516, 2022.

WHO, World Health Organization. **Depressive disorder (depression)**. 2023. Disponível em: <https://www.who.int/en/news-room/fact-sheets/detail/depression>. Acesso em: 12 jun. 2023.

COSTA, Roger Vinícius; ALMEIDA, Iara Carnevale. CHATBOTS E SAÚDE MENTAL: uma linha do tempo com aproximações e distanciamentos de 1966 a 2021. **XI Mostra Interna de Trabalhos de Iniciação Científica**. 2021.

APÊNDICE A

Estratégias de busca

Figura I

Base de dados	Estratégias de busca
Pubmed	("mental disorder*" OR "mental health" OR "mood disorder*" OR "psychotic disorder*" OR "obsessive-compulsive disorder*" OR autism OR "panic disorder*" OR "phobic disorder*" OR "post-traumatic stress disorder*" OR "substance-related disorder*" OR depression OR depressed OR melancholia OR "anxiety disorder*" OR anxious OR phobia OR bipolar OR schizophrenia OR psychosis OR eating disorder* OR anorexia OR binge-eating OR bulimia OR "drug dependence" OR "substance dependence" OR addiction OR "drug abuse" OR "substance abuse") AND ("conversational agent*" OR "conversational bot*" OR "conversational system*" OR "conversational interface*" OR "chatbot*" OR "chat bot*" OR "chatterbot*" OR "chatter bot*" OR "chat-bot*" OR "smartbot*" OR "smart bot*" OR "smart-bot*" OR "virtual coach*" OR "virtual agent*" OR "embodied agent*" OR "relational agent*" OR avatar OR "virtual

	character*" OR "animated character*" OR "virtual human")
Cochrane	(("mental disorder*" OR "mental health" OR "mood disorder*" OR "psychotic disorder*" OR "obsessive-compulsive disorder*" OR autism OR "panic disorder*" OR "phobic disorder*" OR "post-traumatic stress disorder*" OR "substance-related disorder*" OR depression OR depressed OR melancholia OR "anxiety disorder*" OR anxious OR phobia OR bipolar OR schizophrenia OR psychosis OR eating disorder* OR anorexia OR binge-eating OR bulimia OR "drug dependence" OR "substance dependence" OR addiction OR "drug abuse" OR "substance abuse") AND ("conversational agent*" OR "conversational bot*" OR "conversational system*" OR "conversational interface*" OR "chatbot*" OR "chat bot*" OR "chatterbot*" OR "chatter bot*" OR "chat-bot*" OR "smartbot*" OR "smart bot*" OR "smart-bot*" OR "virtual coach*" OR "virtual agent*" OR "embodied agent*" OR "relational agent*" OR avatar OR "virtual character*" OR "animated character*"

	OR "virtual human")):ti,ab,kw (Word variations have been searched)
Scopus	TITLE-ABS-KEY "mental disorder*" OR "mental health" OR "mood disorder*" OR "psychotic disorder*" OR "obsessive-compulsive disorder*" OR autism OR "panic disorder*" OR "phobic disorder*" OR "post-traumatic stress disorder*" OR "substance-related disorder*" OR depression OR depressed OR melancholia OR "anxiety disorder*" OR anxious OR phobia OR bipolar OR schizophrenia OR psychosis OR eating AND TITLE-ABS-KEY disorder* OR anorexia OR binge-eating OR bulimia OR "drug dependence" OR "substance dependence" OR addiction OR "drug abuse" OR "substance abuse" AND TITLE-ABS-KEY "conversational agent*" OR "conversational bot*" OR "conversational system*" OR "conversational interface*" OR "chatbot*" OR "chat bot*" OR "chatterbot*" OR "chatter bot*" OR "chat-bot*" OR "smartbot*" OR "smart bot*" OR "smart-bot*" OR "virtual

	coach*" OR "virtual agent*" OR "embodied agent*" OR "relational agent*" OR avatar OR "virtual character*" OR "animated character*" OR "virtual human"
Embase	("mental disorder*" OR "mental health" OR "mood disorder*" OR autism OR depression OR anxiety OR phobia OR bipolar OR schizophrenia) AND ("conversational agent*" OR "chatbot*" OR "chat bot*" OR "chatterbot*")

Fonte: Modificado de Alrazaq *et al.*, (2020)

APÊNDICE B

Checklist de avaliação

Finalidade do checklist: apoiar a avaliação acadêmica e técnica de chatbots em saúde mental, considerando critérios de escopo, evidência reportada, segurança, privacidade, transparência, vieses, usabilidade e manutenção.

Como aplicar: para cada chatbot analisado, marcar (A/P/N/NI) e registrar, em “Evidência/observação”, de onde a informação foi obtida (artigo, site, manual, termo de privacidade, descrição do app). Quando não houver informação, marcar **NI**.

Legenda: **A** (Atende) | **P** (Parcial) | **N** (Não atende) | **NI** (Não informado)

1. Finalidade e escopo do sistema

ID	Critério	Pergunta de verificação	Criticidade	A/P/N/NI	Evidência/observação
1.1	Objetivo declarado	O sistema declara claramente o objetivo (apoio, psicoeducação, exercícios, triagem, navegação de recursos)?	Média		
1.2	Público-alvo	O público-alvo é definido (idade, contexto, necessidade, restrições)?	Média		

ID	Critério	Pergunta de verificação	Criticidade	A/P/N/NI	Evidência/observação
1.3	Limites do que faz	O sistema explicita limites (não substitui profissional; não é emergência; limitações de interpretação)?	Alta		
1.4	Limites do que não faz	O sistema informa o que não oferece (diagnóstico, prescrição, terapia completa, etc.)?	Alta		
1.5	Contexto de uso	Há orientação de contexto de uso (frequência, duração, quando interromper e buscar ajuda)?	Alta		

2. Evidência e validação reportada

ID	Critério	Pergunta de verificação	Criticidade	A/P/N/NI	Evidência/observação
2.1	Evidência citada	Existem estudos, relatórios ou validações citadas	Média		

ID	Critério	Pergunta de verificação	Criticidade	A/P/N/NI	Evidência/observação
		(mesmo que preliminares)?			
2.2	Tipo de estudo reportado	A fonte informa o tipo de avaliação (ensaio, piloto, estudo observacional, usabilidade, etc.)?	Média		
2.3	Instrumentos de medida	Foram usados instrumentos reconhecidos (ex.: PHQ-9, GAD-7) ou métricas justificadas?	Média		
2.4	Desfechos claros	O que foi medido está claro (depressão, ansiedade, estresse, bem-estar, engajamento)?	Média		
2.5	Transparência de resultados	A fonte reporta resultados	Média		

ID	Critério	Pergunta de verificação	Criticidade	A/P/N/NI	Evidência/observação
		com clareza (não só “funciona”, mas “como foi avaliado”)?			
2.6	Generalização prudente	O sistema evita promessas absolutas (“cura”, “garante”), usando linguagem compatível com evidência?	Alta		

3. Segurança e manejo de crise

ID	Critério	Pergunta de verificação	Criticidade	A/P/N/Nl	Evidência/observação
3.1	Aviso de emergência	O sistema orienta que não é serviço de emergência e indica o que fazer em crise?	Alta		

ID	Critério	Pergunta de verificação	Criticidade	A/P/N/N	Evidência/observação
3.2	Encaminhamento	Há encaminhamento para serviços/recursos (ex.: emergência, apoio local) quando necessário?	Alta		
3.3	Respostas seguras	Há evidência de cuidado para não fornecer conselhos perigosos, julgamentos ou invalidar sofrimento?	Alta		
3.4	Sinais de risco	O sistema possui alguma estratégia para lidar com mensagens de risco (mesmo que seja orientar busca imediata de ajuda)?	Alta		

ID	Cr�terio	Pergunta de verifica��o	Criticidad e	A/P/N/N l	Evid�ncia/observa��o
3.5	Escalonamento humano	Existe op���o de contato humano (suporte, equipe, servi��o) quando o caso ultrapassa o chatbot?	Alta		
3.6	Mensagens de conten���o	Em situa���es sens���veis, o chatbot usa linguagem de acolhimento e orienta���es prudentes?	Alta		

4. Privacidade e governan  a de dados

ID	Cr�terio	Pergunta de verifica��o	Criticidad e	A/P/N/N l	Evid�ncia/observa��o
4.1	Pol�tica de privacidade	Existe pol�tica de privacidade clara e acess��vel?	Alta		
4.2	Consentimento	H� consentimento expl�cito para coleta/uso de dados sens��veis?	Alta		

ID	Critério	Pergunta de verificação	Criticidade	A/P/N/N	Evidência/observação
4.3	Minimização de dados	O sistema coleta apenas o necessário para a finalidade declarada?	Alta		
4.4	Finalidade definida	A finalidade do uso de dados é informada (suporte, melhoria, pesquisa, publicidade)?	Alta		
4.5	Compartilhamento com terceiros	O compartilhamento com terceiros é informado e justificado?	Alta		
4.6	Retenção e exclusão	O usuário consegue solicitar exclusão/retificação e há prazo de retenção informado?	Alta		
4.7	Segurança da informação	Há menção a medidas de segurança (criptografia,	Alta		

ID	Critério	Pergunta de verificação	Criticidade	A/P/N/N	Evidência/observação
		controle de acesso, etc.) ou pelo menos compromisso explícito?			

5. Transparência e comunicação de limitações

ID	Critério	Pergunta de verificação	Criticidade	A/P/N/N	Evidência/observação
5.1	Identidade do sistema	Está explícito que o usuário conversa com IA/chatbot?	Alta		
5.2	Limitações comunicadas	O sistema informa limitações (erros, incompreensão, não diagnóstico)?	Alta		
5.3	Linguagem apropriada	A linguagem é clara, respeitosa e evita termos que culpabilizem o usuário?	Alta		
5.4	Referências e fontes	Quando oferece orientações, há indicação de base (diretrizes, técnicas	Média		

ID	CrITÉrio	Pergunta de verificação	Criticidad e	A/P/N/N/I	Evidência/observação
		reconhecidas) ou cautela ao orientar?			
5.5	Rastreabilidade básica	Existe registro mínimo do que foi recomendado (para auditoria/seguimento), preservando privacidade?	Média		

6. Mitigação de vieses, equidade e acessibilidade

ID	CrITÉrio	Pergunta de verificação	Criticidad e	A/P/N/NI	Evidência/observação
6.1	Inclusão linguística	O sistema considera idioma/variação linguística e evita interpretações perigosas por ambiguidades?	Média		
6.2	Público diverso	Há preocupação com diferentes públicos (idade, escolaridade, contexto cultural)?	Média		
6.3	Acessibilidade	Há recursos/compatibilidade básica com	Média		

ID	CrITÉrio	Pergunta de verificação	Críticidad e	A/P/N/NI	Evidência/observação
		acessibilidade (leitura, contraste, navegação simples)?			
6.4	Declaração de vieses	O sistema reconhece limitações e risco de vieses (quando aplicável)?	Média		
6.5	Não discriminação	Não há indícios de respostas discriminatórias (gênero, etnia, orientação, religião)?	Alta		

7. Usabilidade e engajamento

ID	CrITÉrio	Pergunta de verificação	Críticidad e	A/P/N/NI	Evidência/observação
7.1	Facilidade de uso	O fluxo é intuitivo, com instruções claras e baixa carga cognitiva?	Média		
7.2	Feedback do usuário	O sistema permite feedback (útil/inútil; relatório de problema)?	Média		

ID	Critério	Pergunta de verificação	Criticidade	A/P/N/NI	Evidência/observação
7.3	Personalização responsável	Há personalização (se existir) com controle do usuário e explicação do que muda?	Média		
7.4	Engajamento sem manipulação	Evita técnicas de engajamento que pressionem/culpabilizem (ex.: vergonha, ameaça)?	Alta		
7.5	Tom empático	A conversa busca acolhimento sem simular capacidade clínica que não possui?	Alta		

8. Operação, atualização e monitoramento (ciclo de vida)

ID	Critério	Pergunta de verificação	Criticidade	A/P/N/NI	Evidência/observação
8.1	Responsável/contato	Existe canal de contato e responsabilidade institucional definida?	Alta		

ID	Critério	Pergunta de verificação	Criticidade	A/P/N/NI	Evidência/observação
8.2	Atualizações	Há política de atualização/correção de falhas e melhoria contínua?	Média		
8.3	Monitoramento de incidentes	Existe mecanismo para registrar incidentes/erros críticos e corrigi-los?	Alta		
8.4	Auditoria	Existe possibilidade de auditoria (mesmo que interna), principalmente em casos críticos?	Alta		
8.5	Dependência de plataforma	Há clareza sobre dependências (app, mensageria, web) e limitações técnicas?	Baixa		