

**UNIVERSIDADE DE CAXIAS DO SUL
PROGRAMA DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO - PPGA
CURSO DE MESTRADO**

WALTER SENGIK DA CRUZ

**FATORES DETERMINANTES DA DESISTÊNCIA CONVERSACIONAL ENTRE
USUÁRIOS E *CHATBOTS* BASEADOS EM MODELOS DE INTELIGÊNCIA
ARTIFICIAL**

CAXIAS DO SUL

2026

WALTER SENGIK DA CRUZ

**FATORES DETERMINANTES DA DESISTÊNCIA CONVERSACIONAL ENTRE
USUÁRIOS E *CHATBOTS* BASEADOS EM MODELOS DE INTELIGÊNCIA
ARTIFICIAL**

Projeto de Dissertação de Mestrado submetida
à Banca Examinadora designada pelo
Colegiado do Programa de Pós-Graduação em
Administração da Universidade de Caxias do
Sul, como parte dos requisitos necessários para
obtenção do título de Mestre em Administração

Orientador: Prof. Dr. Mateus Panizzon

CAXIAS DO SUL

2026

Dados Internacionais de Catalogação na Publicação (CIP)
Universidade de Caxias do Sul
Sistema de Bibliotecas UCS - Processamento Técnico

C957f Cruz, Walter Sengik da

Fatores determinantes da desistência conversacional entre usuários e *chatbots* baseados em modelos de inteligência artificial [recurso eletrônico] / Walter Sengik da Cruz. – 2026.

Dados eletrônicos.

Dissertação (Mestrado) - Universidade de Caxias do Sul, Programa de Pós-Graduação em Administração, 2026.

Orientação: Mateus Panizzon.

Modo de acesso: World Wide Web

Disponível em: <https://repositorio.ucs.br>

1. Inteligência artificial. 2. Chatbots. 3. Interação homem-máquina. 4. Processamento de linguagem natural (Computação). 5. Comportamento do consumidor. I. Panizzon, Mateus, orient. II. Título.

CDU 2. ed.: 004.89:316.772.5

Catalogação na fonte elaborada pela(o) bibliotecária(o)
Ana Guimarães Pereira - CRB 10/1460

WALTER SENGIK DA CRUZ

**FATORES DETERMINANTES DA DESISTÊNCIA CONVERSACIONAL ENTRE
USUÁRIOS E *CHATBOTS* BASEADOS EM MODELOS DE INTELIGÊNCIA
ARTIFICIAL**

Projeto de Dissertação de Mestrado submetida
à Banca Examinadora designada pelo
Colegiado do Programa de Pós-Graduação em
Administração da Universidade de Caxias do
Sul, como parte dos requisitos necessários para
obtenção do título de Mestre em Administração

Aprovado (a) em: 27/03/2026

Banca examinadora

Prof. Dr(a). Ana Cristina Fachinelli Bertolini
Universidade de Caxias do Sul - UCS

Prof. Dr(a). Cintia Paese Giacomello
Universidade de Caxias do Sul - UCS

Prof. Dr(a). Cristiane Froehlich
Universidade Feevale

Dedico este trabalho à academia, sem a qual
seríamos seres privados de luz e orientação.

AGRADECIMENTOS

Agradeço, primeiramente, a Deus, por me conceder a força e a sabedoria necessárias para superar os desafios desta trajetória acadêmica. À minha esposa, expresso minha profunda gratidão por ser minha maior incentivadora e por caminhar ao meu lado com paciência e companheirismo em cada etapa. À minha família, agradeço pelo apoio incondicional, pois sem o espaço, incentivo e a compreensão que me deram para me dedicar integralmente aos estudos, eu não estaria aqui hoje. Estendo meu agradecimento ao meu orientador, Prof. Dr. Mateus Panizzon, pela confiança, pelos ensinamentos e pela condução assertiva deste trabalho. Às professoras integrantes da banca examinadora, pelas valiosas reflexões e contribuições durante a avaliação deste projeto de dissertação. À Universidade de Caxias do Sul (UCS) e ao Programa de Pós-Graduação em Administração (PPGA), por todo o suporte institucional oferecido. Por fim, agradeço aos professores e colegas do curso de mestrado, pelas trocas de conhecimento e pelo incentivo ao crescimento acadêmico e profissional.

RESUMO

A automação em cenários de atendimento tem se consolidado nos últimos anos, inicialmente com a adoção de *chatbots* baseados em regras ou fluxos pré-definidos. Mais recentemente, observa-se a evolução para *chatbots* fundamentados em modelos de linguagem de grande escala (LLMs) e técnicas de Inteligência Artificial Generativa, capazes de produzir respostas contextualmente relevantes e promover interações conversacionais mais naturais com os usuários. Apesar do potencial da tecnologia para aumentar a escalabilidade do atendimento, bem como geração de dados e *analytics* para monitoramento de taxas de sucesso de atendimento ou venda, ainda existe uma lacuna em compreender por que usuários se engajam mais ou menos em uma interação humano-máquina neste contexto. Partindo desta lacuna, a pesquisa adota o método *Design Science Research* (DSR) para o desenvolvimento de um artefato (o *framework* multidimensional e a sua instanciação tecnológica) destinado a avaliação dos determinantes da desistência da interação com *chatbots*, com foco no atendimento ao cliente. Métodos tradicionais de avaliação costumam se basear em pesquisas diretas com usuários, coletando percepções subjetivas sobre a experiência, ou em métricas quantitativas, como tempo de resposta e taxa de resolução. No entanto, essas abordagens não capturam adequadamente aspectos como clareza, confiança, engajamento e empatia durante a interação. Para suprir essa lacuna, este estudo propõe a análise de padrões comportamentais como um método complementar e mais direto para avaliar a experiência do usuário, com a extração, classificação e análise de conversas entre usuários e *chatbots* que utilizam modelos de IA conversacional. Foi utilizada a metodologia PRISMA para uma revisão sistemática da literatura para identificar os principais indicadores de desempenho dos *chatbots*, observando a desistência dos usuários na continuidade do atendimento como variável dependente, influenciada por três preditores principais. Como parte da abordagem de pesquisa, foram desenvolvidos dois artefatos baseados em um modelo de inteligência artificial generativa: o primeiro capaz de analisar conversas e classificá-las conforme variáveis comportamentais; e o segundo capaz de analisar os resultados quantitativos e propor melhorias operacionais. A análise estatística envolveu a comparação de cinco modelos estruturais concorrentes, buscando identificar a configuração de melhor ajuste dada a novidade do fenômeno em estudo. Os resultados obtidos, fundamentados na análise de 849 interações reais, sugerem a ocorrência do 'paradoxo da persistência': a rejeição da hipótese de performance indica que falhas técnicas isoladas tendem a não causar o abandono imediato, evidenciando que limitações de confiança, clareza, engajamento e empatia podem reter o usuário na interface por meio de um esforço cognitivo exaustivo de adaptação. Foram testados cinco modelos rivais através da Análise de Equações Estruturais com o objetivo de avaliar a melhor configuração de ajuste, dada a novidade do fenômeno. Com base nos resultados, o modelo 5, apresentado na Figura 17, demonstrou o melhor ajuste e interpretação dos resultados.

Palavras-chave: inteligência artificial generativa. *chatbots*. atendimento ao cliente. desistência da interação. interação humano-máquina. padrões comportamentais.

ABSTRACT

Automation in customer service scenarios has consolidated in recent years, initially with the adoption of chatbots based on rules or predefined flows. More recently, there has been an evolution toward chatbots grounded in Large Language Models (LLMs) and Generative Artificial Intelligence techniques, capable of producing contextually relevant responses and promoting more natural conversational interactions with users. Despite the technology's potential to increase service scalability, as well as generate data and analytics for monitoring service or sales success rates, a gap remains in understanding why users engage more or less in a human-machine interaction within this context. Addressing this gap, this research adopts the Design Science Research (DSR) method for the development of an artifact (a multidimensional framework and its technological instantiation) intended to evaluate the determinants of chatbot interaction abandonment, focusing on customer service. Traditional evaluation methods are typically based on direct user surveys, collecting subjective perceptions of the experience, or on quantitative metrics, such as response time and resolution rates. However, these approaches do not adequately capture aspects such as clarity, trust, engagement, and empathy during the interaction. To fill this gap, this study proposes the analysis of behavioral patterns as a complementary and more direct method to evaluate user experience, through the extraction, classification, and analysis of conversations between users and chatbots utilizing conversational AI models. The PRISMA methodology was used for a systematic literature review to identify the main performance indicators of chatbots, observing user abandonment as the dependent variable, influenced by three main predictors. As part of the research approach, two artifacts based on a generative artificial intelligence model were developed: the first capable of analyzing conversations and classifying them according to behavioral variables; and the second capable of analyzing quantitative results and proposing operational improvements. Statistical analysis involved the comparison of five competing structural models, seeking to identify the best-fit configuration given the novelty of the phenomenon under study. The results obtained, based on the analysis of 849 real interactions, suggest the occurrence of the 'persistence paradox': the rejection of the performance hypothesis indicates that isolated technical failures tend not to cause immediate abandonment, evidencing that limitations in trust, clarity, engagement, and empathy may retain the user in the interface through an exhaustive cognitive effort of adaptation. Five rival models were tested through Structural Equation Modeling to evaluate the best fit configuration, given the novelty of the phenomenon. Based on the results, Model 5, presented in Figure 17, demonstrated the best fit and interpretation of the results.

Keywords: generative artificial intelligence. chatbots; customer service. interaction dropout. human-machine interaction. behavioral patterns.

LISTA DE FIGURAS

Figura 1 - O protocolo PRISMA utilizado para RSL.....	35
Figura 2 - Framework teórico.....	60
Figura 3 - Framework de monitoramento do agente conversacional.....	62
Figura 4 - Exemplo de conversa real entre agente conversacional e usuário (dados).....	64
Figura 5 - Recorte da tabela de dados.....	72
Figura 6 - Recorte do software (artefato de informação).....	72
Figura 7 - Recorte do software, visão de seleção de variável para ação em massa.....	73
Figura 8 - Recorte do software, visão de orientação do modelo (artefato de informação).....	74
Figura 9 - Processo completo do fluxo de tratamento de dados.....	75
Figura 10 - Gráfico de dispersão sem aplicação do percentil 95°.....	80
Figura 11 - Gráfico de dispersão com aplicação do percentil 95°.....	81
Figura 12 - Framework do modelo 1 (teórico processual).....	93
Figura 13 - Framework do modelo 2 (teórico purificado multidimensional).....	94
Figura 14 - Framework do modelo 3 (teórico purificado processual).....	95
Figura 15 - Framework do modelo 4 (segunda ordem não purificado).....	96
Figura 16 - Framework do modelo 5 (purificado e D2 como antecedente).....	97
Figura 17 - Framework selecionado: modelo purificado e D2 como antecedente.....	104
Figura 18 - Recorte de log de interação - chatbot IA Beto (Betina).....	105
Figura 19 - Recorte de log de interação - chatbot Abby.....	106

LISTA DE QUADROS

Quadro 1 - Categorias temáticas e palavras-chave da pesquisa.....	36
Quadro 2 - Categorização de relevância dos artigos.....	39
Quadro 3 - Inclusão de novas categorias na extração.....	40
Quadro 4 - Pontuação para categorização final e seleção de artigos.....	41
Quadro 5 - Artigos centrais.....	43
Quadro 6 - Evidências para as dimensões de pesquisa.....	57
Quadro 7 - Variáveis e indicadores de medição.....	75
Quadro 8 - Variáveis e indicadores de medição.....	78
Quadro 9 - Modelos de avaliação no SEM.....	90

LISTA DE TABELAS

Tabela 1 - Estatística descritiva e parâmetros de limitação de outliers.....	79
Tabela 2 - Estatística descritiva com variáveis padronizadas e sem outliers.....	82
Tabela 3 - Pesos fatoriais da AFC no modelo de mensuração teórico.....	83
Tabela 4 - Medidas de ajustamento exato do modelo teórico.....	84
Tabela 5 - Medidas de ajustamento global do modelo teórico.....	84
Tabela 6 - Covariâncias fatoriais do modelo teórico.....	84
Tabela 7 - Pesos fatoriais e singularidade da AFE.....	85
Tabela 8 - Pesos fatoriais da AFC (modelo refinado).....	87
Tabela 9 - Covariâncias fatoriais do modelo teórico.....	88
Tabela 10 - Medidas de ajustamento exato do modelo refinado.....	88
Tabela 11 - Medidas de ajustamento global do modelo refinado.....	88
Tabela 12 - Índices de ajuste global dos modelos estruturais.....	91
Tabela 13 - Estimativas do modelo 1 (teórico processual).....	92
Tabela 14 - Estimativas do modelo 2 (teórico purificado multidimensional).....	93
Tabela 15 - Estimativas do modelo 3 (teórico purificado processual).....	95
Tabela 16 - Estimativas do modelo 4 (segunda ordem não purificado).....	96
Tabela 17 - Estimativas do modelo 5 (purificado e D2 como antecedente).....	97
Tabela 18 - Comparação das estimativas dos modelos.....	98

LISTA DE ABREVIATURAS E SIGLAS

AFE	Análise Fatorial Exploratória
AFC	Análise Fatorial Confirmatória
CFI	<i>Comparative Fit Index</i>
CNN	<i>Convolutional Neural Network</i>
CRM	<i>Customer Relationship Manager</i>
DWLS	<i>Diagonally Weighted Least Squares</i>
IA	Inteligência Artificial
LIWC	<i>Linguistic Inquiry and Word Count</i>
LLM	<i>Large language models</i>
PLN	Processamento de Linguagem Natural
PPGA	Programa de Pós-Graduação em Administração
PRISMA	<i>Preferred Reporting Items for Systematic Reviews and Meta-Analyses</i>
RSL	Revisão Sistemática da Literatura
RMSEA	<i>Root Mean Square Error of Approximation</i>
SHAP	<i>Shapley Additive Explanations</i>
TF-IDF	<i>Term Frequency-Inverse Document Frequency</i>
TLI	<i>Tucker-Lewis Index</i>
LGPD	Lei Geral de Proteção de Dados

SUMÁRIO

1	INTRODUÇÃO.....	16
1.1	TEMA E PROBLEMA.....	18
1.2	DIMENSÕES DA PESQUISA.....	19
1.3	OBJETIVOS.....	20
1.3.1	Objetivo geral.....	21
1.3.2	Objetivos específicos.....	21
1.4	JUSTIFICATIVA TEÓRICA.....	21
1.5	JUSTIFICATIVA PRÁTICA.....	23
1.6	ALINHAMENTO COM O PPGA.....	24
1.7	INTEGRAÇÃO DA INTELIGÊNCIA ARTIFICIAL NO PROCESSO METODOLÓGICO.....	26
2	REVISÃO TEÓRICA.....	27
2.1	CHATBOTS.....	27
2.1.1	Definição e chatbots baseados em modelos de IA.....	27
2.1.2	Benefícios organizacionais do uso de chatbots.....	28
2.1.3	Chatbots em contextos multiculturais.....	28
2.1.4	Integração de chatbots com outras tecnologias.....	29
2.1.5	Papel de métricas como empatia, confiança e fluidez.....	30
2.2	MODELOS DE IA GENERATIVA E CONVERSACIONAL.....	30
2.2.1	Evolução tecnológica de modelos de IA generativa e conversacional.....	31
2.2.2	Aplicações práticas.....	31
2.2.3	Aceitação e resistência à IA conversacional.....	32
2.2.4	Aspectos éticos a modelos de IA conversacional.....	33
2.2.5	IA generativa aplicada aos modelos conversacionais.....	34
2.3	MÉTODO DE PESQUISA PRISMA.....	34
2.3.1	Identificação.....	36
2.3.2	Triagem.....	37
2.3.3	Extração.....	39
2.3.4	Síntese.....	42
2.4	INDICADORES EM INTERAÇÕES CONVERSACIONAIS.....	44
2.4.1	Indicadores baseados em análise de logs.....	44

2.4.2	Métodos de avaliação.....	46
2.5	MÉTODOS DE ANÁLISE DE CONVERSAS EM CHATBOTS.....	47
2.5.1	Abordagens manuais e automáticas.....	47
2.5.2	Mineração de texto, análise de sentimentos e detecção de emoções.....	48
2.5.3	Frameworks de análise: métricas de empatia e clareza.....	49
2.6	CLAREZA E COERÊNCIA DOS MODELOS CONVERSACIONAIS.....	50
2.6.1	Métodos para avaliar clareza e coerência dos modelos conversacionais...	50
2.7	EMPATIA E CONEXÃO HUMANA NAS RESPOSTAS.....	52
2.7.1	Métodos para avaliar empatia em chatbots.....	52
2.7.2	Estudos de caso em que a empatia melhorou a experiência do usuário....	53
2.8	EFICÁCIA NA RESOLUÇÃO DE NECESSIDADES DO USUÁRIO.....	54
2.8.1	Métodos para avaliação das respostas/interações.....	55
2.8.2	Análise de limitações frequentes em chatbots e como abordá-las.....	56
2.9	EMBASAMENTO DAS DIMENSÕES DE PESQUISA.....	57
2.10	CONCLUSÃO E LACUNAS NA LITERATURA.....	57
3	FRAMEWORK TEÓRICO PROPOSTO.....	60
3.1	HIPÓTESES DE PESQUISA.....	61
3.2	ESTRUTURA DO FRAMEWORK DE MONITORAMENTO DO AGENTE CONVERSACIONAL.....	62
3.2.1	Elementos do framework.....	63
3.2.1.1	Logs de conversa entre usuários e agente conversacional (dados).....	63
3.2.1.2	LLM de extração e classificação (artefato de informação).....	64
3.2.1.3	Framework teórico (artefato de conhecimento).....	65
3.2.1.4	LLM de monitoramento e interpretação.....	67
3.2.1.5	Melhoria de software (método).....	67
4	MÉTODO.....	68
4.1	DELINEAMENTO DA PESQUISA.....	68
4.2	CAMPO DE ESTUDO.....	68
4.3	PROCEDIMENTO DE COLETA DE DADOS.....	69
4.3.1	Dados.....	69
4.3.2	Variáveis e escalas.....	70
4.4	PROCESSO DE ORGANIZAÇÃO DA BASE DE DADOS.....	71

4.5	PROCESSO DE ANÁLISE DE DADOS.....	75
4.5.1	Avaliação estatística e peso das variáveis.....	76
4.5.2	Defesa do processo de análise.....	77
5	RESULTADOS E IMPLICAÇÕES.....	78
5.1	TRATAMENTO E INTERPOLAÇÃO LINEAR DOS DADOS.....	79
5.2	ANÁLISE CONFIRMATÓRIA E EXPLORATÓRIA.....	83
5.3	ANÁLISE DO MODELO DE EQUAÇÕES ESTRUTURAIS.....	89
5.3.1	Reespecificação do modelo teórico a estrutura serial.....	89
5.3.2	Definição de desistência e modelos de avaliação.....	89
5.3.3	Análise dos índices comparativos no SEM.....	90
5.3.4	Análise dos índices comparativos no SEM nos cinco modelos.....	92
5.4	DETERMINANTES DA DESISTÊNCIA NA CONVERSAÇÃO.....	98
5.4.1	O descolamento entre performance e desistência (H1).....	99
5.4.2	A influência da confiança e clareza na desistência (H2).....	100
5.4.3	A influência do engajamento e empatia na desistência (H3).....	101
5.5	SÍNTESE DOS ACHADOS E DISCUSSÃO INTEGRADA.....	102
6	CONTRIBUIÇÕES, LIMITAÇÕES E ESTUDOS FUTUROS.....	103
6.1	CONTRIBUIÇÕES TEÓRICAS.....	103
6.2	CONTRIBUIÇÕES PRÁTICAS E GERENCIAIS DA DISSERTAÇÃO.....	104
6.3	LIMITAÇÕES DA PESQUISA.....	106
6.3.1	Métrica de desistência como proxy de aceitação.....	106
6.3.2	Restrições dos dados de log.....	107
6.3.3	Desafios metodológicos e ferramental estatístico.....	107
6.3.4	Especificidade do contexto e da tarefa.....	108
6.4	ESTUDOS FUTUROS.....	108
	REFERÊNCIAS.....	110
	APÊNDICE A - QUADRO COM ARTIGOS SINTETIZADOS 1.....	119
	APÊNDICE B - QUADRO COM ARTIGOS SINTETIZADOS 2.....	126

APÊNDICE C - FUNCIONAMENTO LÓGICO FERRAMENTA.....	132
APÊNDICE D - CONTEXTOS (PROMPTS).....	133
APÊNDICE E - CONTEXTOS (PROMPTS) CHATBOTS 1.....	136
APÊNDICE F - CONTEXTOS (PROMPTS) CHATBOTS 2.....	138
APÊNDICE G - CONTEXTOS (PROMPTS) CHATBOTS 3.....	140
APÊNDICE H - CONTEXTO/PROMPT MELHORIA.....	142

1 INTRODUÇÃO

A inteligência artificial (IA) é um campo multidisciplinar voltado ao desenvolvimento de sistemas capazes de simular aspectos do comportamento humano, como percepção, aprendizado e tomada de decisão (Poole; Mackworth; Goebel, 1998). Formalizado no Dartmouth Summer Research Project, em 1956, o termo IA representava, desde sua origem, a crença de que todos os aspectos da inteligência poderiam ser descritos com tal precisão que uma máquina seria capaz de reproduzi-los (Moor, 2006). Inicialmente dominada por abordagens simbólicas, baseadas em regras lógicas fixas Winston (1992), a IA evoluiu significativamente com o avanço de técnicas de aprendizado de máquina, capazes de identificar padrões complexos em grandes volumes de dados (Russell; Norvig, 2010).

A introdução de redes neurais profundas e, posteriormente, da arquitetura Transformer, revolucionou o campo do Processamento de Linguagem Natural (PLN). Nesse contexto, a OpenAI lançou, em 2018, o modelo *Generative Pre-trained Transformer* (GPT), marco importante no desenvolvimento dos chamados *large language models* (LLMs), capazes de gerar textos coerentes a partir de entradas contextuais (OpenAI, 2018). Essa transição representou uma ruptura com os paradigmas anteriores, ao possibilitar uma adaptação dinâmica das respostas, simulando aspectos da criatividade e da comunicação humana (Floridi; Cowls, 2019; Sohl-Dickstein et al., 2015).

Nesse novo cenário, os agentes conversacionais, sistemas baseados em IA que simulam interações humanas por meio de linguagem natural, e neste contexto de pesquisa, aplicados em *chatbots*, vêm sendo utilizados por empresas de serviços para responder a consultas online, assumindo funções antes atribuídas a atendentes humanos e contribuindo para a eficiência e personalização no atendimento ao cliente (Camilleri; Troise, 2023). Evoluindo de scripts rígidos para estruturas capazes de simular diálogos mais fluídos e humanos, os *chatbots*, definidos como aplicações de software capazes de se comunicar por meio de linguagem natural (Nicolescu; Tudorache, 2022), passaram a ocupar um papel central na transformação digital do atendimento ao cliente, oferecendo maior escalabilidade, disponibilidade contínua e agilidade no serviço prestado (Mozafari, Weiger e Hammerschmidt, 2021; Nicolescu e Tudorache, 2022).

Avanços tecnológicos, como a implementação de assistentes virtuais baseados em modelos de IA, possibilitaram um aumento significativo na escalabilidade das operações, permitindo que empresas atendam a demandas crescentes sem comprometer a qualidade do

serviço prestado (Andrade; Tumelero, 2022). Quando os usuários confiam na qualidade da comunicação e na segurança dos dados fornecidos aos *chatbots*, tornam-se mais dispostos a interagir com esses assistentes, permitindo que as empresas automatizem tarefas repetitivas e reservem os atendentes humanos para situações que exigem maior empatia e complexidade (Song *et al.*, 2022). *Chatbots* baseados em modelos de IA podem operar continuamente, oferecendo suporte ao cliente em tempo real e automatizando processos organizacionais críticos, desde o atendimento padrão até a personalização de interações com base em dados coletados (Urbani; Ferreira; Lam, 2024). Chang e Hsiao (2024) argumentam que, mesmo que os *chatbots* tenham potencial para aprimorar o atendimento ao cliente e reduzir custos, muitos consumidores ainda relatam experiências negativas, afetando sua satisfação e percepção de valor. Apesar disso, as empresas têm reconhecido os *chatbots* como elementos estratégicos para atender às demandas de um mercado altamente competitivo, promovendo inovação e otimizando processos de comunicação e atendimento (Xie *et al.*, 2024).

Conforme destacado por Chang e Hsiao (2024), muitos sistemas de *chatbots* não atendem às expectativas dos consumidores, resultando em experiências negativas que geram resistência ao uso. Contudo métricas quantitativas, como tempo médio de resposta e número de interações concluídas, sejam frequentemente usadas para avaliar o desempenho dos *chatbots*, esses indicadores não capturam as emoções negativas e os desafios qualitativos enfrentados pelos usuários. Nesse contexto, torna-se essencial adotar abordagens que considerem não apenas a eficiência técnica, mas também a percepção do valor e as experiências emocionais dos usuários.

Ademais, a antropomorfização (a atribuição de características humanas às máquinas) tem se mostrado um fator importante na percepção e diminuição da desistência na interação dos usuários com *chatbots*. Estudos recentes indicam que *chatbots* capazes de demonstrar emoções, como humor, não apenas melhoram a satisfação dos usuários, mas também promovem maior engajamento e confiabilidade nas interações (Xie *et al.*, 2024). No entanto, essa abordagem também apresenta desafios, como o risco de criar expectativas irreais ou de gerar desconforto nos usuários, especialmente quando o humor expresso pelo *chatbot* ultrapassa os limites do que é considerado natural e apropriado nas interações virtuais (Xie *et al.*, 2024). No entanto os *chatbots* ofereçam ganhos claros de eficiência, barreiras funcionais e psicológicas, como dificuldade de uso e experiências frustrantes, contribuem para emoções negativas e resistência ao uso (Chang; Hsiao, 2024).

Neste contexto, destaca-se nesta pesquisa o estudo de empresas que adotam tecnologias no mercado de *chatbots*, como os modelos de linguagem baseados em inteligência

artificial para oferecer interações automatizadas que combinam eficiência operacional e personalização da comunicação. Para aprofundar o entendimento dessas interações conversacionais e validar empiricamente as hipóteses deste estudo, foram analisados dados reais de uma empresa atuante nesse mercado, permitindo identificar desafios práticos vivenciados por usuários. Dentre esses desafios, destacam-se especialmente aqueles relacionados à percepção da qualidade comunicacional e aos riscos à privacidade, fatores que influenciam significativamente a intenção de adoção dos *chatbots* por parte dos usuários (Song *et al.*, 2022).

Por mais que existam modelos consolidados para avaliação de agentes conversacionais, como o *Conversational Usability and Quality (CUQ) Framework* (Radziwill e Benton, 2017), que propõe critérios baseados na ISO 9241, incluindo efetividade, eficiência e satisfação, tais abordagens ainda precisam evoluir para capturar plenamente os desafios de avaliação impostos por sistemas mais avançados baseados em inteligência artificial e aprendizado de máquina. No presente estudo, reconhece-se a contribuição desses modelos, mas opta-se por propor um novo *framework* de avaliação da desistência na interação, fundamentado em uma revisão sistemática focada principalmente em *chatbots* construídos com modelos de linguagem baseados em IA generativa. Dessa forma, espera-se oferecer à academia quanto ao mercado, um artefato capaz de mensurar a avaliação do atendimento pelos *chatbots* com modelos de IA a fim de propor ajustes e melhorias no atendimento.

1.1 TEMA E PROBLEMA

A questão central de pesquisa que este trabalho busca responder é: "*Como projetar e avaliar um artefato capaz de identificar ou analisar a desistência conversacional de usuários em interações com chatbots baseados em IA conversacional a partir de indicadores comportamentais e conversacionais?*". Para enfrentar este desafio, esta pesquisa adota a *Design Science Research (DSR)* para projetar e avaliar um artefato (composto por um modelo e uma instanciação tecnológica) capaz de resolver essa lacuna na classe de problemas de auditoria de interações humano-IA, construído através do referencial teórico, aplicado as conversas entre os agentes conversacionais e os usuários, no contexto de *chatbots*, buscando entender se as respostas fornecidas pelos agentes correspondem às expectativas dos usuários ou se os usuários solicitam um atendimento humano ou desistem da conversa. Essa questão é uma proposta de pesquisa que utilizará unicamente os dados de conversas para a avaliação dos preditores que justifiquem esta variável dependente da desistência conversacional.

A qualidade da comunicação, incluindo a capacidade dos agentes conversacionais em simular interações humanas por meio de características antropomórficas, facilita a construção de *rapport* (ou seja, uma conexão pessoal e agradável durante a interação) e impacta significativamente o conforto, diminuição da desistência conversacional e satisfação dos consumidores (Becker; Mahr; Odekerken-Schröder, 2023). Além disso, a personalização e a adequação das respostas fornecidas pelos *chatbots* podem melhorar a experiência do cliente (Camilleri; Troise, 2023), sendo a interatividade um fator adicional que se destaca na diferenciação de confiança no atendimento digital (Borau *et al.*, 2021; Mostafa e Kasamani, 2021; Mozafari, Weiger e Hammerschmidt, 2021; Shiu-Wan, Cheng e Chiu, 2019; Wilkinson *et al.*, 2021 *apud* Camilleri e Troise, 2023).

Apesar do avanço dos sistemas, ainda é possível observar resistência à adoção de *chatbots* por parte dos usuários. Grimes, Schuetzler e Giboney (2021) afirmam que a avaliação negativa de *chatbots* pode ocorrer quando os usuários enfrentam falhas nas interações, como a ausência de respostas contextuais ou a limitação em lidar com questões complexas. Portanto, Para avaliar adequadamente a eficácia desses sistemas, é necessário considerar fatores técnicos, como usabilidade e responsividade, que influenciam diretamente a experiência e a satisfação do cliente (Chen; Le; Florence, 2021).

1.2 DIMENSÕES DA PESQUISA

Com base no objetivo de analisar as conversas dos agentes conversacionais e os fatores que podem indicar o aumento ou diminuição da desistência conversacional, identificam-se quatro dimensões principais focadas no comportamento do usuário e na performance dos modelos de IAs utilizado para os agentes:

- a) D1 (desistência): usuários que preferem falar com um ser humano ou abandonam a conversa demonstram maior insatisfação com o *chatbot* em comparação com aqueles que finalizam a interação de forma natural. Base teórica: *Expectation Violation Theory* (Crollic *et al.*, 2022);
- b) D2 (performance): a ocorrência de alucinações (respostas irrelevantes ou fora do contexto) nas respostas do modelo de IA conversacional e maior tempo de resposta aumentam significativamente a insatisfação do usuário e a probabilidade de abandono da conversa. Base teórica: *Cognitive Fit Theory* (Vessey, 1991);

- c) D3 (confiança e clareza): respostas claras e contextuais fornecidas pelo modelo de IA conversacional aumentam a percepção de presença social (sensação de envolvimento humano na interação, como se o agente tivesse calor, personalidade e estivesse socialmente engajado com o usuário), o que, por sua vez, reduz a solicitação de suporte humano e incentiva a continuidade da interação com o *chatbot*. Base teórica: *Social Response Theory* (Adam; Wessel; Benlian, 2022);
- d) D4 (engajamento e empatia): a inclusão de elementos de empatia nas respostas do *chatbot* contribui para reduzir a solicitação de suporte humano e aumentar o engajamento dos usuários. Base teórica: *Communication Accommodation Theory* (Zhang; Giles, 2016).

Para validar essas dimensões, será adotado abordagem analítica de método misto, estruturada em duas fases complementares: primeiro, uma etapa qualitativa, na qual modelos de IA generativa são utilizados para a extração e classificação automática de variáveis comportamentais e emocionais presentes nos logs de conversa; e, posteriormente, uma etapa quantitativa, onde os dados classificados são submetidos a análises estatísticas multivariadas, incluindo a Modelagem de Equações Estruturais (SEM), para testar as correlações entre os indicadores e validar as hipóteses do framework teórico. A análise das conversas entre os usuários e o *chatbot* permitirá identificar padrões de comportamento que evidenciem desistência ou continuidade da conversa finalizar de forma natural (todas informações solicitadas serem respondidas), bem como avaliar os elementos técnicos e emocionais das interações. Dessa forma, será possível compreender melhor como o desempenho do modelo de IA conversacional e as expectativas dos usuários se relacionam, oferecendo *insights* para a melhoria contínua de sistemas de *chatbots*.

1.3 OBJETIVOS

Este capítulo apresenta os objetivos da pesquisa, que orientam o desenvolvimento e a aplicação do *framework* teórico proposto. A definição dos objetivos visa alinhar a investigação às necessidades de avaliação da desistência conversacional de usuários com agentes conversacionais em *chatbots*, considerando não apenas métricas tradicionais, mas também indicadores comportamentais e aspectos subjetivos da experiência do usuário.

1.3.1 Objetivo geral

O objetivo principal desta pesquisa é: **Projetar e avaliar a utilidade de um artefato para a avaliação da desistência conversacional entre usuários e *chatbots* baseados em modelos de IA conversacionais de texto, visando o avanço do conhecimento prescritivo na área.**

Esse artefato busca integrar métricas quantitativas e qualitativas, utilizando a análise de padrões comportamentais como um elemento central para mensurar a experiência do usuário e a performance do *chatbot*.

1.3.2 Objetivos específicos

Para alcançar o objetivo geral desta pesquisa, foram definidos os seguintes objetivos específicos:

- a) Mapear na literatura científica as variáveis, métricas e critérios de avaliação da interação entre usuários e agentes conversacionais baseados em IA;
- b) Projetar o artefato (modelo), definindo as variáveis e o formalismo lógico para a extração dos padrões comportamentais de modelos de IA conversacionais;
- c) Construir a instanciação (desenvolvimento) do artefato;
- d) Avaliar a utilidade e eficácia do artefato em um ambiente real;
- e) Analisar como essas relações contribuem para compreender a desistência ou continuidade nas interações com *chatbots*;
- f) Propor recomendações práticas para o *design* e a operação de *chatbots*.

1.4 JUSTIFICATIVA TEÓRICA

Song *et al.* (2022) discutem que, por mais que o uso de *chatbots* tenha potencial para aprimorar a experiência do cliente, muitas soluções atuais falham em atender às expectativas devido à falta de usabilidade ou respostas pouco significativas. Adicionalmente, a maioria das pesquisas concentra-se em métricas tradicionais como satisfação, lealdade e intenção de uso contínuo, enquanto estudos que avaliam a desistência na interação com agentes conversacionais em termos de padrões comportamentais permanecem escassos.

Para mapear essa lacuna, esta pesquisa realizou uma Revisão Sistemática da Literatura (RSL), na qual foram analisados 45 estudos identificados na base *Web of Science*. Os resultados indicam que a maior parte das abordagens atuais se apóiam em métodos quantitativos tradicionais ou em pesquisas diretas com usuários, sem considerar a análise das interações em si. Apesar do impacto de fatores como antropomorfização e expressão de emoções em *chatbots* seja reconhecido na literatura, Chang e Hsiao (2024) destacam que ainda há lacunas na consideração das emoções negativas na avaliação da resistência ao uso de tecnologias como os *chatbots*. Estudos como os de Yang, Sun e Shen (2023) apontam que barreiras funcionais, como respostas imprecisas e interações pouco intuitivas, e barreiras psicológicas, como a desconfiança e resistência a mudanças, amplificam emoções negativas, como frustração e insatisfação.

Outro aspecto crítico identificado é o desafio de balancear transparência e a experiência percebida por usuários em interações com *chatbots*. Conforme Xie *et al.* (2024), a divulgação da identidade do *chatbot* exerce uma influência moderadora negativa na relação entre o humor expresso pelo *chatbot* e a satisfação do usuário, indicando que, dependendo do contexto, essa transparência pode atenuar os efeitos positivos do humor na experiência do usuário. Complementarmente, Camilleri e Troise (2023) destacam que, mesmo com os *chatbots* oferecendo suporte eficiente ao cliente, eles ainda enfrentam desafios para simular completamente a interação humana, especialmente em relação à criação de uma sensação de presença social.

Essas limitações afetam diretamente a adoção dos *chatbots*, uma vez que a capacidade de gerar interações mais naturais e empáticas influencia a percepção do usuário sobre a confiabilidade e a usabilidade da tecnologia. Assim, há a necessidade de desenvolver sistemas que não apenas aprimorem a eficiência operacional, mas também atendam às expectativas emocionais dos usuários, reduzindo a resistência ao uso e melhorando a experiência geral do atendimento automatizado.

Portanto, a justificativa teórica para esta pesquisa se baseia na necessidade de avançar o conhecimento sobre a avaliação da resistência na interação conversacional. A proposta de integração de padrões comportamentais às avaliações existentes não apenas preenche uma lacuna importante na literatura, mas também oferece direções para o desenvolvimento de *chatbots* mais eficazes e empáticos, alinhados às necessidades do mercado de personalização e escalabilidade.

1.5 JUSTIFICATIVA PRÁTICA

A evolução dos *chatbots* e dos modelos de inteligência artificial aplicada à comunicação empresarial tem trazido inegáveis benefícios às organizações, como redução de custos, melhoria na eficiência operacional e escalabilidade. No entanto, este mercado ainda enfrenta desafios específicos relacionados à experiência do usuário e à maximização da eficácia de seus serviços automatizados. A presente pesquisa busca oferecer subsídios práticos para enfrentar essas questões, com foco em resultados que possam ser aplicados diretamente às operações das empresas.

Song *et al.* (2022) destacam que, por mais que os *chatbots* sejam utilizados para lidar com demandas de grande volume, os usuários frequentemente relatam experiências insatisfatórias devido à falta de empatia e à incapacidade de responder de maneira personalizada a situações mais complexas. Além disso, a falta de estratégias eficazes de personalização pode comprometer a experiência do usuário, uma vez que a ausência de elementos como humor e presença social impacta negativamente a satisfação com os *chatbots* (Xie *et al.*, 2024).

Outro aspecto prático relevante é a necessidade de sistemas com modelos de inteligência artificial capazes de interpretar com precisão as intenções e expectativas dos usuários. Empresas que buscam expandir sua atuação internacionalmente, garantir que seus *chatbots* compreendam nuances culturais e linguísticas torna-se necessário para oferecer interações mais naturais e eficazes. A personalização e a humanização da comunicação dos *chatbots* podem impactar positivamente a percepção dos usuários e suas decisões de compra (Li, Gan e Zheng, 2023). Da mesma forma, o uso de *chatbots* com traços antropomórficos, como calor humano e competência, pode fortalecer a confiança dos clientes e melhorar a avaliação do serviço prestado (Li, Gan e Zheng, 2023). No entanto, essa implementação deve ser cuidadosamente planejada, pois um excesso de informações ou respostas excessivamente sofisticadas pode sobrecarregar o usuário e afetar negativamente sua experiência (Li, Gan e Zheng, 2023).

Além disso, a personalização baseada em *chatbots* com modelos de IA, embora valorizada, apresenta desafios no que diz respeito à transparência. Usuários tendem a valorizar sistemas que demonstram conhecimento de suas preferências, mas, ao mesmo tempo, podem reagir negativamente quando suas expectativas são violadas e não atendidas (Grimes; Schuetzler; Giboney, 2021). De acordo com Wanner *et al.* (2022), a transparência dos sistemas inteligentes desempenha um papel elevado na construção da confiança do usuário,

influenciando diretamente a aceitação dessas tecnologias. A falta de transparência pode gerar desconfiança e resistência, enquanto explicações claras e acessíveis sobre o funcionamento do sistema ajudam a mitigar a percepção de risco e aumentar a predisposição dos usuários a aceitá-los.

A pesquisa também traz implicações práticas na alocação de recursos entre atendimento automatizado e humano. Por mais que o uso de *chatbots* tenha reduzido significativamente a necessidade de suporte humano em muitos casos, algumas interações ainda exigem a intervenção de profissionais qualificados para garantir a resolução satisfatória de problemas mais complexos (De Andrade; Tumelero, 2022). Este estudo oferecerá *insights* sobre como otimizar essa alocação, identificando padrões de comportamento e métricas que indiquem quando e como o suporte humano deve ser acionado.

Por fim, os resultados esperados da pesquisa podem guiar o desenvolvimento de novos recursos e funcionalidades que ampliem o valor agregado das companhias que oferecem *chatbots* como serviço. Os *insights* gerados pelo estudo poderão ser aplicados diretamente em estratégias de design, operação e marketing, contribuindo para o alcance de metas organizacionais e para a criação de experiências mais satisfatórias para os usuários finais.

1.6 ALINHAMENTO COM O PPGA

O Programa de Pós-Graduação em Administração (PPGA) tem como uma de suas principais áreas de concentração a Inovação e Competitividade, enfatizando a necessidade de estudos que integrem práticas inovadoras à realidade organizacional, buscando transformar desafios em oportunidades. A presente pesquisa está alinhada a esses objetivos, ao propor soluções que combinam inteligência artificial conversacional com estratégias de automação para aprimorar processos operacionais e gerenciais.

Um dos aspectos mais relevantes da pesquisa é seu potencial para fomentar discussões sobre como a aplicação de tecnologias avançadas pode gerar valor estratégico para empresas, principalmente em mercados de alta competitividade. Nesse contexto, empresas que oferecem tais serviços de *chatbots* com modelos de IA são um exemplo prático de como a inovação tecnológica pode ser utilizada para otimizar a experiência do cliente e aumentar a eficiência organizacional. A pesquisa contribui para esse debate ao oferecer um modelo de avaliação que combina análises qualitativas e quantitativas, proporcionando uma visão mais

abrangente sobre a desistência conversacional com *chatbots* que utilizam modelos de IA conversacionais.

Além disso, o estudo aborda questões éticas e sociais relacionadas ao uso de modelos de IA, tais como transparência, privacidade e confiança, alinhando-se às diretrizes do PPGA no que se refere à responsabilidade social e ética empresarial. Autores como Zamora (2017) enfatizam as preocupações com privacidade como barreiras significativas à aceitação dessa tecnologia. Tais percepções reforçam a necessidade de um design ético e transparente, capaz de fomentar a confiança dos usuários e promover a adoção responsável de tecnologias baseadas em modelos de IA.

A pesquisa também se conecta à linha de estudos do PPGA que explora o impacto de novas tecnologias na criação de vantagem competitiva sustentável. Ao investigar como padrões comportamentais podem ser utilizados para melhorar as interações entre *chatbots* e usuários, o estudo não apenas preenche lacunas teóricas, mas também fornece *insights* práticos que podem ser aplicados em diferentes setores do mercado. Estudos como o de Hennekeuser *et al.* (2024) destacam a importância de considerar aspectos da experiência do usuário e características comportamentais dos agentes conversacionais, que podem impactar diretamente a adoção dessas tecnologias.

Outro ponto de alinhamento importante é o foco do PPGA em promover a interdisciplinaridade. A pesquisa integra conhecimentos de áreas como administração, ciência da computação e psicologia, combinando teorias e métodos de diferentes disciplinas para abordar a complexidade das interações humano-computador. Esse enfoque interdisciplinar é essencial para lidar com os desafios contemporâneos relacionados à implementação de tecnologias avançadas, como observado em trabalhos recentes que discutem as implicações econômicas e gerenciais dos *chatbots* baseados em modelos de IA nos mais diversos setores (Khneyzer; Boustany; Dagher, 2024).

Por fim, a pesquisa reforça o compromisso do PPGA em promover a produção de conhecimento que seja relevante tanto para a academia quanto para o mercado. Ao propor um *framework* que combina análise de padrões comportamentais com métricas tradicionais, o estudo oferece contribuições que podem ser utilizadas por gestores, desenvolvedores de tecnologia e acadêmicos interessados em compreender e aprimorar o impacto de *chatbots* com modelos de IA no contexto empresarial. Dessa forma, a pesquisa se posiciona como um exemplo de como a inovação pode ser utilizada para transformar a teoria em prática, gerando benefícios concretos para organizações e sociedade.

1.7 INTEGRAÇÃO DA INTELIGÊNCIA ARTIFICIAL NO PROCESSO METODOLÓGICO

Esta pesquisa contou com o apoio do *ChatGPT* versão 4o em diversas etapas, atuando como uma ferramenta complementar para otimizar e estruturar o trabalho. Inicialmente, foi utilizado na filtragem e organização dos artigos científicos, auxiliando na seleção de materiais mais relevantes. Em seguida, contribuiu para a análise e extração de *insights*, permitindo estabelecer conexões entre diferentes fontes e aprofundar a interpretação dos dados. Além disso, foi empregado no apoio à revisão e colaboração na escrita, ajudando a aprimorar a clareza e coesão textual. Por fim, desempenhou um papel na organização e tratamento dos dados, facilitando a identificação de padrões e tendências. Apesar do suporte tecnológico, a análise crítica e a validação dos resultados permaneceram como responsabilidades centrais do pesquisador.

2 REVISÃO TEÓRICA

Este capítulo define os pilares fundamentais da pesquisa, estabelecendo um panorama sobre os temas centrais. Nas próximas seções são definidos os principais conceitos técnicos, a Revisão Sistemática da Literatura que deu suporte às dimensões de análise desta pesquisa e por fim, a estruturação da revisão teórica para construção do *framework* teórico.

2.1 CHATBOTS

Os *chatbots* vêm sendo utilizados como ferramentas para automatizar o atendimento digital e interações entre empresas e clientes, evidenciando seu papel estratégico no serviço ao consumidor (Nicolescu; Tudorache, 2022). A ampliação do uso de *chatbots* como agentes de serviço baseados em inteligência artificial tem sido influenciada pelos avanços tecnológicos, impactando diretamente a satisfação e a intenção de continuidade dos usuários (Ashfaq *et al.*, 2020). Apesar desse potencial, sua adoção ainda encontra barreiras, principalmente ligadas à confiança dos usuários, fatores decisivos para a consolidação dessa tecnologia (Prakash *et al.* (2023). Nesse sentido, torna-se necessário explorar suas definições, diferenças em relação a outras tecnologias conversacionais e características de aplicação em contextos organizacionais (Camilleri; Troise, 2023).

2.1.1 Definição e *chatbots* baseados em modelos de IA

Os *chatbots* são sistemas computacionais baseados em inteligência artificial e processamento de linguagem natural, que permitem a comunicação com humanos em diversos contextos (Nicolescu; Tudorache, 2022). Os *chatbots* são utilizados em contextos empresariais para atendimento ao cliente e vendas, proporcionando maior eficiência ao responder a perguntas sobre produtos e serviços (Blümel; Zaki; Bohné, 2024). E *chatbots* com características antropomórficas, como calor humano e competência, podem melhorar a experiência do cliente ao aumentar a confiança e aliviar a sobrecarga cognitiva, especialmente quando essas características coexistem (Li; Gan; Zheng, 2023).

Chatbots baseados em modelos de IA utilizam aprendizado de máquina e processamento de linguagem natural para analisar dados em tempo real, permitindo interações personalizadas. No contexto de vendas e serviços, o equilíbrio entre personalização e

privacidade desempenha um papel essencial na aceitação dos *chatbots* pelos clientes, influenciando sua experiência e intenção de compra (Hua *et al.*, 2022). Uma vantagem adicional dos *chatbots* baseados em IA é sua capacidade de integrar dados contextuais e emocionais nas interações. Por exemplo, eles podem ajustar o tom de suas respostas com base no estado emocional detectado no usuário, criando experiências mais empáticas e satisfatórias (Blümel; Zaki; Bohné, 2024). No entanto, a complexidade desses sistemas também traz desafios, como altos custos de desenvolvimento e maior necessidade de manutenção.

2.1.2 Benefícios organizacionais do uso de *chatbots*

O uso de *chatbots* proporciona melhorias significativas na eficiência operacional das empresas. Ao automatizar tarefas repetitivas, como respostas a perguntas frequentes, os *chatbots* possibilitam um atendimento mais ágil e acessível, além de reduzir a sobrecarga dos atendentes humanos, permitindo que se concentrem em demandas mais complexas (De Andrade; Tumelero, 2022). Além disso, os *chatbots* desempenham um papel chave na experiência do cliente, oferecendo respostas rápidas e automatizadas que aumentam a percepção de utilidade e satisfação. Ashfaq *et al.* (2020) destacam que, no contexto do e-commerce, a qualidade da informação e do serviço fornecido pelos *chatbots* influencia significativamente a satisfação dos usuários, a qual, por sua vez, impacta diretamente sua intenção de continuar utilizando essa tecnologia.

2.1.3 *Chatbots* em contextos multiculturais

Em um mundo cada vez mais globalizado, a capacidade dos *chatbots* de se adaptar a diferentes culturas e idiomas tornou-se um fator crítico para seu sucesso. Embora fatores culturais não sejam diretamente tratados por Prakash *et al.* (2023), a confiança em *chatbots* é influenciada por aspectos como presença social, utilidade percebida e predisposição à confiança, que podem variar entre contextos. A popularidade dos *chatbots* não é uniforme globalmente, uma vez que sua aceitação depende de fatores culturais, tecnológicos e comportamentais. Ashfaq *et al.* (2020) destacam que elementos como familiaridade com a tecnologia, facilidade de uso e a necessidade de interação com humanos influenciam diretamente a satisfação dos usuários e sua intenção de continuar utilizando essa tecnologia.

Grimes, Schuetzler e Giboney (2021) destacam que a eficácia dos *chatbots* depende da sua capacidade de alinhar-se às expectativas dos usuários. Essa compreensão pode ser

expandida, em contextos mais amplos, para incluir fatores culturais que influenciam essas expectativas. A melhoria da interface dos *chatbots* e o uso de algoritmos preditivos podem aumentar a presença social e a interatividade, tornando a experiência mais envolvente e confiável para os usuários. Modelos mais sofisticados, que incorporem habilidades sociais e sensibilidade emocional, podem potencialmente melhorar a experiência do usuário e fortalecer a confiança na tecnologia, especialmente quando alinhados a elementos como presença social e interatividade, já associados ao aumento da confiança em *chatbots* (Prakash *et al.*, 2023).

2.1.4 Integração de *chatbots* com outras tecnologias

A integração de *chatbots* com outras tecnologias é um fator essencial para maximizar seu impacto organizacional. Nicolescu e Tudorache (2022) apontam que a integração de *chatbots* com sistemas de Gestão de Relacionamento com Clientes (CRM - *Customer Relationship Manager*) permite automatizar atividades relacionadas ao relacionamento com os clientes, contribuindo para maior eficiência organizacional. Essa abordagem permite a análise de dados de forma mais precisa, possibilitando estratégias de atendimento mais assertivas. Camilleri e Troise (2023) complementam ao afirmar que a integração de *chatbots* com sistemas de CRM também possibilita uma melhor experiência do cliente, permitindo a personalização dos serviços com base no histórico de interação. Além disso, destacam que o uso de IA e automação em *chatbots* reduz a carga de trabalho dos atendentes humanos e otimiza processos internos, no entanto desafios como a resistência à adoção da tecnologia e a necessidade de treinamento adequado ainda sejam relevantes para sua implementação eficaz.

Outra integração comum é com plataformas de análise de dados, como o Google Analytics (Google, 2025), que permite às empresas monitorar o desempenho dos *chatbots* em tempo real. A análise de dados é essencial para aprimorar a eficiência dos *chatbots*, permitindo que empresas avaliem a qualidade do atendimento por meio de indicadores como resolutividade e satisfação do cliente. O uso de dashboards e métricas auxilia na identificação de oportunidades de melhoria e na otimização da experiência do usuário (De Andrade; Tumelero, 2022).

No entanto, a integração bem-sucedida exige uma infraestrutura bem estruturada e investimentos em interoperabilidade entre sistemas. Blümel, Zaki e Bohné (2024) ressaltam que desafios técnicos, como a integração adequada das ferramentas de IA, são essenciais para manter uma experiência de usuário satisfatória e evitar frustrações. Por isso, as empresas

devem priorizar o design de arquiteturas flexíveis que suportem atualizações e expansões futuras.

2.1.5 Papel de métricas como empatia, confiança e fluidez

As métricas qualitativas desempenham um papel importante na avaliação do desempenho dos *chatbots*. Empatia, confiança e fluidez são aspectos que influenciam diretamente a percepção do usuário sobre a qualidade do serviço. Blümel, Zaki e Bohné (2024) destacam que *chatbots* capazes de demonstrar empatia por meio de linguagem natural e respostas emocionalmente inteligentes geram maior satisfação e lealdade.

A confiança é outro fator essencial. Usuários tendem a confiar mais em *chatbots* que conseguem lidar com falhas de comunicação. *Chatbots* que buscam clarificação em casos de incerteza são percebidos como mais confiáveis do que aqueles que fornecem respostas incorretas, com impactos diretos na adoção e antropomorfismo percebido (Sheehan; Jin; Gottlieb, 2020). Já a percepção da competência dos *chatbots* influencia diretamente a avaliação dos usuários. Li, Gan e Zheng (2023) destacam que a antropomorfização e a competência percebida dos *chatbots* afetam a confiança dos clientes na tecnologia.

A integração de *chatbots* com outras tecnologias, aliada à adoção de métricas qualitativas, pode ampliar o impacto dessas ferramentas em ambientes organizacionais. Vo *et al.* (2021) destacam que a análise de dados não estruturados pode revelar *insights* sobre o comportamento do cliente, contribuindo para estratégias mais eficazes de retenção e atendimento. A aplicação da IA em *chatbots* tem demonstrado ganhos significativos na eficiência operacional, permitindo um atendimento mais rápido, assertivo e acessível, sem interrupções, tornando-se um diferencial estratégico na automação do atendimento (De Andrade; Tumelero, 2022).

2.2 MODELOS DE IA GENERATIVA E CONVERSACIONAL

Modelos de inteligência artificial conversacional têm sido empregados para permitir que sistemas computacionais compreendem e respondem a mensagens humanas de forma natural e eficiente. Segundo Nicolescu e Tudorache (2022), os *chatbots* representam uma aplicação importante dessa tecnologia no contexto do atendimento ao cliente, explorando funcionalidades avançadas da interação humano-computador.

2.2.1 Evolução tecnológica de modelos de IA generativa e conversacional

O desenvolvimento de *chatbots* e agentes conversacionais evoluiu significativamente desde a década de 1960, quando o ELIZA foi introduzido como o primeiro *chatbot* baseado em regras (Adamopoulou; Moussiades, 2022 *apud* Nicolescu; Tudorache, 2022). Esses sistemas primitivos utilizavam respostas predefinidas para interagir com os usuários, limitando a capacidade de compreender contextos mais complexos (Nicolescu; Tudorache, 2022). Grimes, Schuetzler e Giboney (2021) destacam que as capacidades dos agentes conversacionais, como a habilidade de fornecer respostas personalizadas e variadas, são cruciais para melhorar a percepção do usuário sobre o sistema, reforçando sua conexão social com a tecnologia.

Uma das inovações notáveis foi a implementação de *chatbots* em instituições financeiras para melhorar a eficiência do atendimento ao cliente, com ganhos em agilidade, disponibilidade e acessibilidade. A integração de *chatbots* com sistemas como o IBM Watson tem permitido otimizar a experiência do usuário e reduzir filas de atendimento, além de contribuir para a digitalização dos serviços bancários (De Andrade; Tumelero, 2022). Além disso, Blümel, Zaki e Bohné (2024) destacam que a personalização relacional, quando aplicada por meio de modelos de IA, pode melhorar significativamente a experiência do cliente, ajustando a linguagem e o estilo de comunicação de acordo com o perfil psicológico e emocional do usuário.

As capacidades conversacionais avançadas permitem que os *chatbots* ofereçam interações mais complexas e fluidas, alinhadas às expectativas dos usuários (Grimes; Schuetzler; Giboney, 2021). *Chatbots* antropomórficos têm sido projetados para melhorar a experiência do usuário e influenciar decisões de compra (Li; Gan; Zheng, 2023), explorando características como calor humano e competência conforme destaca Cheng *et al.* (2022) e Roy e Naidoo (2021). Outra contribuição recente inclui a aplicação de técnicas de processamento de emoção para capturar o estado afetivo do cliente e adaptar as respostas de forma empática (Blümel; Zaki; Bohné, 2024).

2.2.2 Aplicações práticas

A aplicação de modelos de IA conversacionais no atendimento ao cliente trouxe benefícios como personalização, disponibilidade 24/7 e resolução de problemas em tempo real. De acordo com Grimes, Schuetzler e Giboney (2021), pistas antropomorfizadas, como

nomes humanos, avatares e atrasos simulados, podem aumentar a percepção de conexão social e engajamento em interações com agentes conversacionais. Esses elementos não apenas aprimoram a experiência do cliente, mas também ajudam as empresas a construir relações de longo prazo com seus consumidores.

Os *chatbots* são adotados em diferentes setores por sua capacidade de oferecer suporte em tempo real, melhorar a qualidade do serviço e aumentar a satisfação dos usuários, o que tem impulsionado seu crescimento em diversas áreas (Ashfaq *et al.*, 2020). De acordo com Grimes, Schuetzler e Giboney (2021), a capacidade dos agentes conversacionais de fornecer respostas personalizadas e variadas contribui para um maior engajamento e satisfação, o que pode ser especialmente benéfico em contextos diversos no mercado. Blümel, Zaki e Bohné (2024) acrescentam que o uso de comunicação personalizada aumenta a percepção de competência e confiança, especialmente em interações de longo prazo.

O setor de *e-commerce*, por exemplo, tem se beneficiado significativamente da tecnologia de IA conversacional, pois facilita o atendimento em escala além de proporcionar um atendimento customizado e empático. *Chatbots* que transmitem calor humano e competência podem impactar positivamente a avaliação dos serviços e a intenção de compra, por meio do aumento da confiança e da redução da sobrecarga cognitiva (Li; Gan; Zheng, 2023). A personalização relacional nesses contextos, como sugerido por Blümel, Zaki e Bohné (2024), inclui adaptar estilos de conversa para refletir o humor e as preferências do cliente, promovendo maior lealdade à marca.

2.2.3 Aceitação e resistência à IA conversacional

A aceitação de tecnologias de modelos de IA conversacional depende de diversos fatores, incluindo percepção de utilidade, facilidade de uso e empatia percebida. A personalização no atendimento, conforme discutido por Blümel, Zaki e Bohné (2024), é fundamental para criar uma experiência mais acolhedora e reduzir a resistência dos clientes, especialmente ao evitar respostas impessoais e genéricas. A confiança na interação com tecnologias de IA conversacional é influenciada não apenas pelas características técnicas do *chatbot*, mas também pela percepção de presença social gerada por respostas claras e humanizadas (Adam; Wessel; Benlian, 2022).

Segundo Bach *et al.* (2022), a confiança do usuário em sistemas com agentes conversacionais é multifacetada, envolvendo aspectos como integridade, competência e previsibilidade do sistema. A integridade está relacionada à consistência das respostas

fornecidas pelos modelos de IA destes agentes, a competência envolve a precisão e relevância das informações, enquanto a previsibilidade se refere à capacidade do sistema de responder de maneira esperada e coerente em diferentes situações (Bach *et al.*, 2022).

De acordo com Li, Gan e Zheng (2023), a percepção de competência e calor humano em agentes conversacionais em *chatbots* desempenha um papel fundamental na construção da confiança dos usuários e na avaliação dos serviços prestados. A falta de clareza sobre o funcionamento dos *chatbots* e a inconsistência nas interações podem gerar desconfiança e reduzir o engajamento dos usuários. Além disso, quando os *chatbots* não atendem às expectativas dos clientes ou quando a interação se torna cognitivamente sobrecarregada, pode surgir resistência por parte dos consumidores. Para mitigar essa resistência, os *chatbots* devem ser projetados para transmitir tanto competência quanto calor humano, pois esses fatores influenciam positivamente a confiança do usuário e sua percepção sobre a qualidade do serviço oferecido.

2.2.4 Aspectos éticos a modelos de IA conversacional

Os aspectos éticos relacionados à modelos de IA conversacional incluem questões como privacidade e transparência. Camilleri e Troise (2023) destacam que, para manter a confiança nas empresas que utilizam essa tecnologia, é fundamental que os usuários sejam informados quando estão interagindo com um agente conversacional, o que ajuda a evitar possíveis abusos e promove práticas transparentes para personalizar interações. Além disso, sistemas que simulam ser humanos sem a devida divulgação podem prejudicar a confiança e a reputação das empresas, especialmente quando as expectativas dos usuários são violadas, em que usuários podem reagir negativamente ao descobrir que estavam interagindo com uma IA quando esperavam conversar com um humano (Grimes; Schuetzler; Giboney, 2021).

Da mesma forma, Hennekeuser *et al.* (2024) apontam que a falta de transparência na comunicação dos *chatbots* pode intensificar preocupações com privacidade, tornando os usuários mais cautelosos ao compartilhar informações pessoais. O estudo destaca que a confiança na IA conversacional está diretamente ligada à percepção de segurança dos dados e à clareza sobre como as informações dos usuários são utilizadas

2.2.5 IA generativa aplicada aos modelos conversacionais.

Os modelos conversacionais com IA generativa estão transformando o atendimento ao cliente ao oferecer interações mais sofisticadas e personalizadas (Ferraro *et al.*, 2024). Esses agentes conversacionais podem ser treinados com diferentes abordagens de *machine learning* e redes neurais profundas para aprender padrões de linguagem e gerar respostas mais naturais e coerentes, mesmo em cenários com dados limitados (Paul *et al.*, 2018). No entanto, a evolução desses modelos não depende apenas da capacidade de aprendizado, mas também de fatores como transparência e previsibilidade, que influenciam a confiança dos usuários na interação com sistemas automatizados (Przegalinska *et al.*, 2019). Assim, métodos que avaliem o desempenho dos agentes conversacionais e permitam ajustes contínuos são fundamentais para aprimorar a experiência do usuário e garantir interações mais eficazes e confiáveis (Przegalinska *et al.*, 2019).

Dentre as abordagens utilizadas para melhorar o desempenho dos agentes conversacionais, destacam-se o aprendizado por reforço e os mecanismos de correção de erros, que permitem ajustes dinâmicos com base no *feedback* dos usuários (Izadi e Forouzanfar, 2024). Esses métodos são fundamentais para mitigar falhas na interpretação da linguagem, auxiliando na geração de respostas mais precisas e contextualmente adequadas (Izadi e Forouzanfar, 2024). Modelos que incorporam aprendizado contextual podem melhorar a relevância da interação ao associar padrões conversacionais a domínios específicos, tornando a comunicação mais adaptada às necessidades dos usuários (Paul *et al.*, 2018). Diante desses avanços, a evolução dos agentes conversacionais no contexto de *chatbots* está relacionada não apenas ao aprimoramento técnico, mas também à sua capacidade de aprender e se adaptar às necessidades dos usuários, proporcionando interações cada vez mais personalizadas e eficazes (Ferraro *et al.*, 2024).

2.3 MÉTODO DE PESQUISA PRISMA

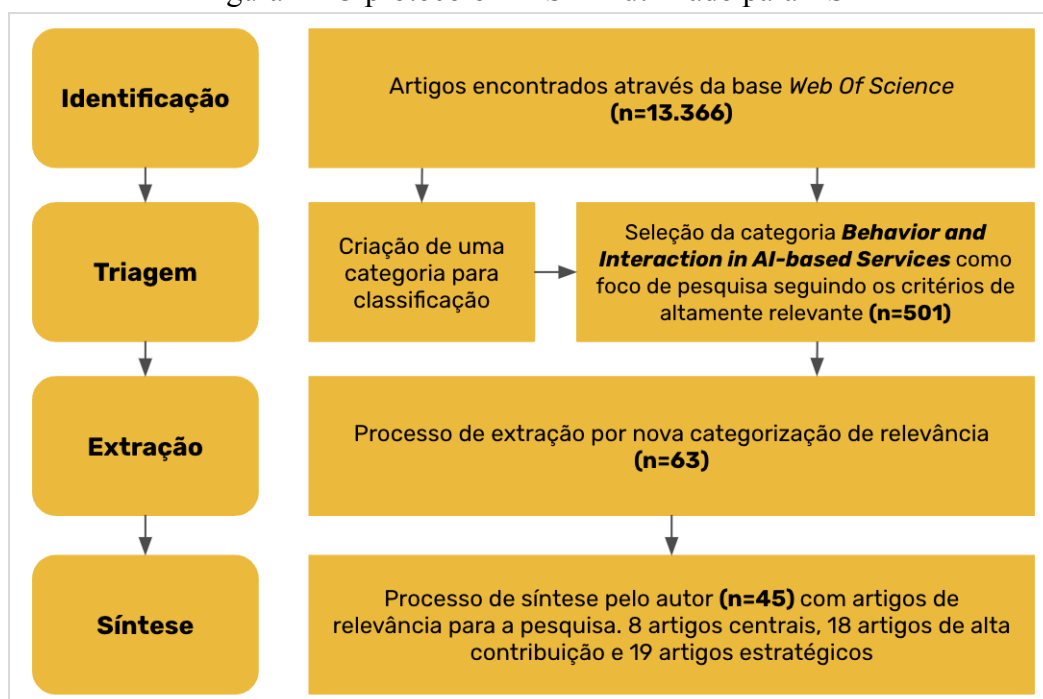
A pesquisa foi estruturada seguindo o protocolo PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*), reconhecido por fornecer uma metodologia eficiente para a seleção e análise de literatura acadêmica (*RSL - Review Sistematic Literature*). O PRISMA oferece um *checklist* baseado em evidências e um fluxo de quatro fases (Moher *et al.*, 2015), garantindo que a revisão seja rigorosa, transparente e replicável. Esse método é necessário para reduzir vieses e garantir a legitimidade dos resultados, um aspecto destacado

em estudos que analisam a revisão sistemática e sua estrutura metodológica (Paschou *et al.*, 2020). O método é organizado em quatro etapas principais:

- identificação** - Busca inicial de artigos relevantes em bases de dados acadêmicas, abrangendo um escopo amplo de literatura;
- triagem** - Aplicação de critérios de inclusão e exclusão para filtrar os artigos mais alinhados ao tema da pesquisa;
- elegibilidade** - Avaliação dos artigos para selecionar os estudos que atendem aos critérios metodológicos e teóricos definidos;
- inclusão** - Integração e síntese dos artigos selecionados, categorizando os principais temas e contribuindo para a construção da estrutura teórica da pesquisa.

A Figura 1 ilustra o processo PRISMA aplicado na busca, triagem, extração e síntese de artigos relevantes para a pesquisa. Esse fluxograma destaca as etapas metodológicas seguidas para garantir rigor e transparência na seleção da literatura.

Figura 1 - O protocolo PRISMA utilizado para RSL



Fonte: Elaborado pelo autor (2026).

Esse protocolo é amplamente utilizado em revisões sistemáticas, sendo reconhecido por sua capacidade de garantir rigor metodológico, transparência no processo e replicabilidade dos resultados. Ele é relevante nesta pesquisa, uma vez que o foco está na análise e avaliação

de interações entre modelos de IA e usuários em *chatbots*, exigindo uma seleção cuidadosa da literatura. A seguir será apresentado cada um dos tópicos do PRISMA com detalhamento das etapas.

2.3.1 Identificação

A busca inicial foi realizada na base de dados *Web of Science* no dia 16 de novembro de 2024. A escolha pela base *Web of Science* se justifica por sua cobertura de publicações relevantes nas áreas de inteligência artificial, interação humano-computador e tecnologias de atendimento automatizado, incluindo temas como *chatbots* e análise de comportamento em interações conversacionais. Para garantir a captura de artigos relevantes ao tema da interação e comportamento humano em interações com *chatbots* em serviços de atendimento e vendas, foi utilizada uma ampla variedade de palavras-chave. Essas palavras foram selecionadas para abranger aspectos como análise de conversação, experiência do usuário, aceitação, limitações e impactos de *chatbots* e inteligência artificial nos serviços. A lista completa de palavras-chave empregadas está presente no Quadro 1.

Quadro 1 - Categorias temáticas e palavras-chave da pesquisa

(continua)

Tema Principal	Palavras-chave utilizadas	Motivador
Chatbots e atendimento ao cliente	<i>AI chatbots in customer service; Artificial intelligence in customer service, AI chatbots, Chatbots</i>	Explorar o impacto e aplicação de <i>chatbots</i> na prestação de serviços de atendimento ao cliente.
Análise de conversação	<i>Chatbot conversation analysis, Conversation analysis with AI, Conversational agents</i>	Investigar como as interações e conversas entre chatbots e usuários são analisadas e modeladas.
Aceitação e expectativas	<i>User acceptance of chatbots, User acceptance of AI in sales, Chatbot expectations vs. reality, Expectation vs reality in AI-based service, Expectation alignment with AI</i>	Examinar como os usuários percebem, aceitam e comparam a performance de <i>chatbots</i> com suas expectativas. Apoiando na validação da desistência de interação.
Experiência do usuário e satisfação	<i>Chatbot user experience, User satisfaction with AI-driven support, User experience in conversational channels, AI chatbots and user satisfaction</i>	Entender como os usuários experimentam e avaliam interações com <i>chatbots</i> e modelos de IA em diferentes canais de comunicação.

(conclusão)

Impacto em vendas e limitações	<i>Sales conversion with chatbots, Impact of automated customer service on sales, Limitations of AI in customer service and sales</i>	Avaliar a influência dos <i>chatbots</i> em conversões de vendas e identificar suas limitações em serviços automatizados.
Precisão e desafios em respostas de IA	<i>AI response accuracy in chatbots, AI hallucination in chatbot responses, AI hallucination in chatbots</i>	Analisar a precisão das respostas dos <i>chatbots</i> e os desafios associados a erros e alucinações geradas pelos modelos de IA.

Fonte: Elaborado pelo autor (2026).

A estratégia de busca permitiu identificar artigos relevantes para a dissertação, focando na interação humana com *chatbots* em atendimento e vendas. O uso de palavras-chave organizadas por temas estruturou a pesquisa e garantiu a diversidade das fontes. A categorização também ajudou a reunir diferentes perspectivas sobre *chatbots* e inteligência artificial, incluindo análise de conversação, desafios e limitações. A revisão está alinhada aos objetivos do estudo e oferece uma base para análises futuras. Além disso, a ausência de recorte temporal ampliou o alcance da pesquisa, incluindo estudos antigos e recentes. Assim, a estratégia adotada segue um rigor metodológico e está descrita nas próximas etapas.

2.3.2 Triagem

Após a busca inicial na base de dados *Web of Science*, foram identificados 13.366 artigos. Dada a ampla quantidade de publicações, foi necessária a aplicação de critérios de filtragem para selecionar apenas os estudos mais alinhados aos objetivos da pesquisa. Essa etapa envolveu a categorização inicial e a classificação dos artigos com o apoio do ChatGPT versão 4o.

A ferramenta organizou os artigos em categorias temáticas para facilitar a análise e reduzir a subjetividade durante a triagem. Os critérios de categorização, descritos abaixo, garantiram uma classificação sistemática e estruturada:

- a) ***Behavior and Interaction in AI-based Services (n=501)***: inclui artigos que abordam diretamente o comportamento humano e a interação em serviços baseados em inteligência artificial. Esses artigos têm alta relevância para a

pesquisa por explorar aspectos centrais do tema, como desistência de interação, satisfação e adaptação ao uso de *chatbots*;

- b) ***Frameworks and Adaptation* (n=751)**: envolve publicações focadas no desenvolvimento ou adaptação de *frameworks* teóricos e metodológicos aplicados a serviços com IA. Esses estudos fornecem suporte técnico e teórico importante para compreender a aplicação prática de *chatbots*.
- c) ***General AI in Services* (n=4.171)**: trata de estudos sobre inteligência artificial em serviços de forma ampla, sem foco específico em *chatbots* ou comportamento humano. Por mais que não centrais, esses artigos oferecem contexto geral e *insights* tecnológicos que complementam a pesquisa;
- d) ***User Experience and Interaction* (n=2.622)**: foca na experiência do usuário, engajamento e interação com sistemas de IA, especialmente em canais digitais. Esses artigos são relevantes por explorar o impacto das tecnologias no ponto de contato com o cliente;
- e) ***Not Relevant* (N=5.321)**: inclui artigos que não possuem relação significativa com o tema principal da dissertação. Publicações desta categoria foram descartadas para evitar dispersão nos resultados.

Após a classificação inicial, foi decidido concentrar a análise exclusivamente na categoria “*Behavior and Interaction in AI-based Services*”, que contém 501 artigos diretamente relacionados ao tema central da dissertação: o comportamento humano e a interação em serviços baseados em inteligência artificial. Para garantir ainda mais a relevância e profundidade da análise, os artigos dessa categoria foram avaliados segundo uma escala de relevância, conforme ilustrado no Quadro 2. A pesquisa focou exclusivamente nos artigos classificados como “Nota 5 - Altamente Relevante”, pois eles tratam diretamente do comportamento humano em interações com modelos de IA no atendimento e suporte. Esses artigos apresentam *insights* teóricos e empíricos, incluindo *frameworks* e discussões aprofundadas sobre adaptação, conversação e experiência do usuário, tornando-se centrais para a construção da base teórica da dissertação.

Quadro 2 - Categorização de relevância dos artigos

Nível de Relevância	Motivador	Artigos na categoria
1 - Não Relevante	O artigo não possui relação significativa com o comportamento humano ou com o uso de modelos de IA no atendimento (vendas ou suporte). Não apresenta utilidade clara para a dissertação.	Hassan <i>et al.</i> (2017) Hannan <i>et al.</i> (2021) Veselovskii <i>et al.</i> (2009)
2 - Pouco Relevante	O artigo trata de <i>chatbots</i> ou modelos de IA em áreas correlacionadas, mas distantes do objetivo, como aplicações puramente técnicas ou outras indústrias.	Jačisko <i>et al.</i> (2024) Park e Shin (2022) Zhong <i>et al.</i> (2020)
3 - Moderadamente Relevante	O artigo menciona <i>chatbots</i> ou modelos de IA com foco genérico ou técnico, como desenvolvimento, aplicações amplas ou limitações tecnológicas, sem explorar profundamente o comportamento humano.	Mohamed <i>et al.</i> (2024); Rau <i>et al.</i> (2024); Benaddi <i>et al.</i> (2024)
4 - Relevante	O artigo discute <i>chatbots</i> ou modelos de IA no contexto de atendimento, focando em áreas como experiência do cliente ou vendas, mas sem abordar diretamente o comportamento humano.	Ludin <i>et al.</i> (2022); Griffin <i>et al.</i> (2023); Hauser-Ulrich <i>et al.</i> (2020)
5 - Altamente Relevante	O artigo trata diretamente do comportamento humano em interações com modelos de IA no atendimento e apresenta <i>insights</i> ou frameworks centrais para a dissertação.	Nicolescu e Tudorache (2022); Sheehan, Jin e Gottlieb (2020); Prakash <i>et al.</i> (2023)

Fonte: Elaborado pelo autor (2026).

2.3.3 Extração

A Extração concentrou-se exclusivamente nos 501 artigos previamente selecionados na categoria “*Behavior and Interaction in AI-based Services*”, que foi identificada como a mais alinhada ao objetivo da pesquisa. Esses artigos passaram por uma análise detalhada para avaliar sua relevância e contribuição teórica ou prática ao tema da dissertação. A avaliação foi realizada utilizando uma categorização em cinco níveis de relevância, que atribuiu uma pontuação de 1 a 5 para cada artigo. Esta classificação utilizou como referência novas categorias de classificação trazidas para os 501 artigos selecionados, conforme o Quadro 3.

Quadro 3 - Inclusão de novas categorias na extração

Categoria	Opções
Tipo de pesquisa	<ul style="list-style-type: none"> ● Proposta de <i>framework</i>/modelo ● Outro ● Experimental ● Pesquisa de campo ● Estudo de caso
Usuário foco	<ul style="list-style-type: none"> ● Consumidor final ● Ambos (consumidores e empresas) ● Empresas
Objeto de estudo	<ul style="list-style-type: none"> ● Compreensão do consumidor ● Melhoria da experiência ● Outro ● Desenvolvimento tecnológico ● Impacto organizacional
Classificação do modelo teórico	<ul style="list-style-type: none"> ● Não classificado ● Agentes conversacionais ● Modelos de interação ● Experiência do cliente ● Impacto no consumidor

Fonte: Elaborado pelo autor (2026).

O Quadro 3 organiza a categorização dos artigos para uma análise estruturada e alinhada à pesquisa. Cada dimensão foi desenvolvida para destacar aspectos relevantes dos estudos, facilitando a identificação de contribuições teóricas e empíricas.

A categoria "Tipo de pesquisa" diferencia os métodos empregados, permitindo identificar estudos conceituais e aplicados. Subcategorias como “Proposta de *framework/modelo*” destacam contribuições teóricas, enquanto “Experimental” e “Pesquisa de campo” fornecem dados empíricos sobre interações humano-IA. Já “Estudos de caso” e “Outro” ampliam a análise com abordagens específicas. O "Usuário foco" organiza os artigos conforme o público estudado. Pesquisas sobre “Consumidor final” analisam comportamento, enquanto aquelas voltadas para “Empresas” exploram o impacto organizacional. A subcategoria “Ambos” permite examinar a interação entre usuários e organizações.

Já a dimensão "Objeto de estudo" define o propósito dos artigos, relacionando-os aos objetivos da dissertação. Subcategorias como “Compreensão do Consumidor” e “Melhoria da Experiência” analisam o impacto dos *chatbots* na usabilidade e no serviço. Já “Desenvolvimento tecnológico” e “Impacto organizacional” ampliam a visão para os efeitos da área de IA nos negócios. Por fim, a "Classificação do modelo teórico" agrupa os artigos por abordagem conceitual, incluindo “Agentes conversacionais”, “Modelos de interação” e

“Experiência do cliente”, permitindo mapear contribuições teóricas e identificar lacunas no campo.

Após a análise da Extração, a seleção final dos artigos foi baseada na "Pontuação de relevância", métrica desenvolvida pelo autor para avaliar a proximidade dos estudos com o tema da dissertação: interação e comportamento do consumidor em conversas com inteligência artificial.

A “Pontuação de Relevância” utilizou três elementos principais para avaliar os artigos: “*Abstract*”, “*Keywords*” e “*Article Title*”. A pontuação foi atribuída com base na presença de palavras-chave relacionadas aos seguintes tópicos centrais:

Comportamento e Interação:

- a) **+3 pontos:** se *behavior* ou *interaction* aparecer no *Abstract*;
- b) **+2 pontos:** se *behavior* ou *interaction* estiverem nas *Keywords*;
- c) **+1 ponto:** se *behavior* ou *interaction* constarem no *Title*.

Chatbots e Inteligência Artificial:

- a) **+3 pontos:** se *chatbot* ou *artificial intelligence* aparecer no *Abstract*;
- b) **+2 pontos:** se *chatbot* ou *artificial intelligence* estiverem nas *Keywords*;
- c) **+1 ponto:** se *chatbot* ou *artificial intelligence* constarem no *Title*.

Experiência do Cliente:

- a) **+2 pontos:** se *customer* ou *experience* aparecer no *Abstract*;
- b) **+1 ponto:** se *customer* ou *experience* estiverem nas *Keywords*;

A pontuação final foi categorizada de acordo com o Quadro 4.

Quadro 4 - Pontuação para categorização final e seleção de artigos

Intervalo de Pontuações	Descrição
12-15: altamente relevante	Artigos que discutem diretamente comportamento humano, interação e uso de modelos de IA/ <i>chatbots</i> para atendimento ou suporte.
8-11: moderadamente relevante	Aborda modelos de IA/ <i>chatbots</i> , mas com foco técnico ou genérico, sem explorar profundamente o comportamento humano.
4-7: pouco relevante	Menciona modelos de IA/ <i>chatbots</i> ou comportamento humano de forma superficial, com aplicações em áreas distantes do objetivo da pesquisa.
0-3: não relevante	Não possui conexão significativa com o tema principal.

Fonte: Elaborado pelo autor (2026).

Após a aplicação dos critérios de classificação, os artigos selecionados foram organizados conforme sua relevância para a pesquisa. Essa abordagem permitiu identificar de

forma objetiva os estudos mais alinhados ao tema, garantindo uma base teórica consistente e direcionada.

Os 501 artigos foram distribuídos em cinco níveis, destacando os 63 altamente relevantes como o núcleo principal da dissertação, por tratarem diretamente do comportamento humano e da interação com modelos de *IA/chatbots*. Além disso, os 147 artigos classificados como relevantes desempenham um papel complementar, oferecendo *insights* adicionais e conexões importantes sobre temas relacionados, como experiência do cliente e impacto em vendas. Os artigos classificados como moderadamente relevantes (51) e pouco relevantes (28) contribuíram para contextualizar o campo de estudo, de forma menos aprofundada. Por fim, os 29 artigos não relevantes foram descartados por não atenderem aos critérios estabelecidos, garantindo foco na análise. Segue a lista de classificação:

- a) 63 artigos altamente relevantes (Pontuação de Relevância: 12-15);
- b) 147 artigos relevantes (Pontuação de Relevância: 8-11);
- c) 51 artigos moderadamente relevantes (Pontuação de Relevância: 6-7);
- d) 28 artigos pouco relevantes (Pontuação de Relevância: 4-5);
- e) 29 artigos não relevantes (Pontuação de Relevância: 0-3).

2.3.4 Síntese

Após a classificação inicial dos artigos com o suporte da inteligência artificial, a síntese e validação final garantiram o alinhamento com os objetivos e o escopo do estudo, refinando os critérios ao contexto da pesquisa. A seleção dos artigos pela RSL proporcionou uma base teórica conectada a aplicações concretas. Por fim, foi realizado pelo autor uma validação dos artigos propostos, chegando a uma síntese final composta por:

- a) **8 artigos centrais:** estes artigos são centrais para a pesquisa, abordando diretamente os objetivos principais e fornecendo contribuições teóricas e práticas de alta profundidade. Estes podem ser melhor apreciados no Quadro 5;
- b) **18 artigos com alta contribuição:** artigos que oferecem suporte complementar, abordando áreas correlatas e ampliando o contexto teórico e prático da análise;
- c) **19 artigos artigos estratégicos:** artigos que fornecem contexto geral, com discussões mais amplas ou específicas, contribuindo para partes secundárias da dissertação.

Quadro 5 - Artigos centrais

Título	Publicação
<i>Customer service chatbots: Anthropomorphism and adoption</i> (Sheehan; Jin; Gottlieb, 2020)	<i>Journal of Business Research</i> , v. 115, p. 14-24, 2020
<i>Determinants and consequences of trust in AI-based customer service chatbots</i> (Prakash, et al., 2023)	<i>The Service Industries Journal</i> , v. 43, n. 9-10, p. 642-675, 2023
Human-Computer Interaction in Customer Service: The Experience with AI <i>Chatbots</i> - A Systematic Literature Review (Nicolescu; Tudorache, 2022)	<i>Electronics</i> , v. 11, n. 10, 2022
<i>I, Chatbot: Modeling the determinants of users' satisfaction and continuance intention of AI-powered service agents</i> (Ashfaq et al., 2020)	<i>Telematics and Informatics</i> , v. 54, 2020
<i>Increasing customer service efficiency through artificial intelligence chatbot</i> (De Andrade; Tumelero, 2022)	Revista de Gestão, v. 29, n. 3, p. 238-251, 2022
<i>Leveraging unstructured call log data for customer churn prediction</i> (Vo et al., 2021)	<i>Knowledge-Based Systems</i> , v. 212, 2021
<i>Live support by chatbots with artificial intelligence: A future research agenda</i> (Camilleri; Troise, 2023)	<i>Service Business</i> , v. 17, p. 61-80, 2023
<i>Mental models and expectation violations in conversational AI interactions</i> (Grimes; Schuetzler; Giboney, 2021)	<i>Decision Support Systems</i> , v. 144, 2021

Fonte: Elaborado pelo autor (2026).

Essa estrutura prioriza os artigos mais relevantes para a pesquisa, incorporando diferentes perspectivas para uma análise abrangente e direcionada. Os Apêndices A e B apresentam os principais artigos extraídos na análise sistemática, organizados em ordem decrescente de relevância. A seleção considerou tanto a contribuição teórica quanto a aplicação prática, abordando temas como *frameworks* conceituais, fatores de resistência ao uso de *chatbots*, impacto de falhas de modelos de IA nas emoções e no comportamento dos usuários e metodologias para otimizar o design e a eficácia dos agentes conversacionais. Os quadros nos Apêndices sintetizam essas contribuições, formando uma base sólida para o arcabouço teórico da dissertação.

Por fim, a revisão sistemática da literatura consolidou as principais contribuições acadêmicas sobre o uso de *chatbots* e agentes conversacionais baseados em modelos de IA, identificando tendências e lacunas que fundamentam a análise da interação com *chatbots*. Esse embasamento teórico sustenta as próximas seções, que aprofundam a avaliação dos

critérios de qualidade das interações, conectando os achados da literatura à estrutura analítica da pesquisa e ao desenvolvimento do referencial teórico.

2.4 INDICADORES EM INTERAÇÕES CONVERSACIONAIS

Os indicadores tradicionais, como tempo de resposta e taxas de resolução, são utilizados para medir a eficiência operacional dos *chatbots*, mas apresentam limitações quando considerados isoladamente. A desistência conversacional com *chatbots* não pode ser medida apenas por métricas quantitativas, sendo necessário considerar também aspectos qualitativos da interação, como *feedback* dos usuários e análise de diálogos (Zhang, Følstad e Bjørkli, 2023). Nesse sentido, Sheehan, Jin e Gottlieb (2020) destacam que a experiência subjetiva dos usuários, especialmente a percepção de antropomorfismo - ou seja, a atribuição de características humanas (como, por exemplo, empatia) - influencia diretamente a intenção de adoção. Além disso, Nguyen, Waizenegger e Techatassanasoontorn (2021) observaram que, em situações em que o *chatbot* não compreende a solicitação inicialmente, os usuários tendem a reformular suas perguntas ou esclarecer seus pedidos, demonstrando esforço e adaptação na tentativa de alcançar seus objetivos. O que permite identificar potenciais padrões de engajamento e compreensão.

A análise de *logs* (registros) de chamadas pode fornecer *insights* mais aprofundados sobre o comportamento dos clientes. Vo *et al.* (2021) demonstraram que a inclusão de dados não estruturados, como padrões de conversação, aprimora a previsão de *churn* (perda de cliente) e possibilita uma compreensão mais detalhada dos perfis de clientes. Da mesma forma, Nguyen, Waizenegger e Techatassanasoontorn (2022) identificaram, por meio de análise temática de interações reais, diferentes padrões de comportamento humano-*chatbot*, incluindo esforços dos usuários para adaptar suas mensagens diante da incompreensão por parte do *chatbot*.

2.4.1 Indicadores baseados em análise de *logs*

Para capturar a desistência do usuário nas interações, é necessário considerar métricas extraídas diretamente dos *logs* de conversação:

Respostas claras e coerentes são fundamentais para garantir que os usuários compreendam as informações fornecidas. Nguyen, Waizenegger e Techatassanasoontorn (2022) identificaram que, em interações com *chatbots*, a necessidade de repetição e

reformulação das perguntas pelos usuários está associada a dificuldades na compreensão da intenção pelo sistema, o que pode comprometer a experiência do usuário. *Chatbots* que buscam clarificação, em vez de fornecer respostas incorretas, são percebidos como mais confiáveis e influenciam sua adoção (Sheehan, Jin e Gottlieb, 2020). Isso se alinha com a D3 (confiança e clareza), pois as respostas contextuais estão positivamente correlacionadas com a confiança do usuário e sua disposição em reutilizar o serviço.

Ranieri, Di Bernardo e Mele (2024) destacam que interações bem estruturadas geram maior retenção dos usuários na conversa, enquanto respostas mecânicas ou genéricas resultam em frustração e abandono. Nguyen, Waizenegger e Techatassanasoontorn (2022) identificaram diferentes padrões de interação (exemplo: desafio, acomodação, comprometimento) que ajudam a compreender se o *chatbot* está realmente conduzindo o usuário à resolução do problema. Esses padrões estão relacionados à D4 (engajamento e empatia), pois demonstram como a inclusão de elementos empáticos pode reduzir a solicitação de suporte humano e aumentar o engajamento emocional dos usuários.

O abandono da conversa ou a solicitação de atendimento humano podem ser sinais diretos de insatisfação. Ranieri, Di Bernardo e Mele (2024) abordam que interações negativas com *chatbots* podem aumentar o estresse do cliente, reduzir a confiança e levar ao abandono da conversa. Já Hsu e Lin (2023) destacam que a qualidade conversacional é o fator que mais influencia na satisfação em interações com modelos de IA, enquanto fatores estruturais como eficiência e segurança impactam mais diretamente a lealdade. Isto reforça a D1 (desistência), que propõe que usuários que solicitam atendimento humano ou abandonam a conversa demonstram maior insatisfação em comparação com aqueles que finalizam a interação de forma natural.

A personalização percebida pode contribuir para a construção da confiança dos usuários em *chatbots*. Fan, Gao e Han (2023) apontam que características como flexibilidade, personalização e capacidade de resposta contextualizada são relevantes para criar experiências positivas, as quais influenciam positivamente atitudes dos consumidores, como a confiança e a intenção de patronagem. Blümel, Zaki e Bohné (2024) mostraram que *chatbots* que utilizam pistas emocionais moderadas (exemplo: *emoticons* e linguagem empática) aumentam a interação social percebida, enquanto o excesso pode gerar desconfiança. Essa relação reforça a D4 (engajamento e empatia), pois sugere que a inclusão de elementos empáticos pode influenciar positivamente o comportamento do usuário.

2.4.2 Métodos de avaliação

A combinação de métodos qualitativos e quantitativos é essencial para uma análise mais precisa da interação com *chatbots*:

Qualitativos: Nguyen, Waizenegger e Techatassanasoontorn (2022) aplicaram análise temática em *logs* para identificar padrões de comportamento humano-*chatbot*, permitindo explorar elementos como presença social e esforço percebido.

Quantitativos: Hsu e Lin (2023) empregaram modelagem de equações estruturais (SEM) para validar a relação entre qualidade da conversação, satisfação e lealdade do usuário em interações com *chatbots*. Vo *et al.* (2021) utilizaram regressão logística e técnicas de mineração de texto e a extração de traços de personalidade, para prever riscos de *churn* a partir de *logs* de chamada.

Mista: Vo *et al.* (2021) utilizaram *machine learning* para prever riscos de *churn*, combinando traços de personalidade extraídos de registros de chamadas com dados estruturados e técnicas interpretáveis.

A capacidade conversacional dos *chatbots* deve ser avaliada não apenas por métricas tradicionais, mas por um conjunto mais amplo de indicadores que consideram como os usuários interagem, expressam frustração e ajustam suas perguntas para tentar se comunicar melhor com o agente conversacional. A insatisfação do usuário, gerando a desistência conversacional, pode ser detectada por padrões como:

- a) aceitação de recomendação de atendimento (indicador de baixa clareza);
- b) Pedidos de atendimento humano (indicador de frustração ou falha na resposta);
- c) Abandono prematuro da conversa (indicador de insatisfação ou falta de confiança na IA).

Além disso, Prakash *et al.* (2023) argumentam que a confiança nos *chatbots* é influenciada por sua utilidade percebida, facilidade de uso e presença social, sendo fortalecida quando a interação é fluida e adaptada ao contexto do usuário. A baixa competência dos *chatbots* pode afetar a confiança do usuário, resultando em uma experiência negativa e impactando sua intenção de uso contínuo (Li, Gan e Zheng, 2023).

Portanto, a abordagem proposta para avaliar a desistência conversacional com *chatbots* deve considerar:

- a) clareza e coerência das respostas (analisando reformulações e repetições);
- b) engajamento e tempo de interação (avaliando padrões de interação);

- c) expressão de frustração e solicitação de atendimento humano (detectando pontos críticos de insatisfação);
- d) impacto da personalização e pistas emocionais (medindo como diferentes estilos de resposta afetam a confiança do usuário).

2.5 MÉTODOS DE ANÁLISE DE CONVERSAS EM *CHATBOTS*

A análise das interações entre humanos e *chatbots* tem evoluído significativamente, incorporando abordagens manuais e automáticas que permitem extrair *insights* sobre a qualidade da experiência do usuário. Vo *et al.* (2021) destacam o uso da mineração de texto para analisar conversas e prever comportamentos dos usuários, explorando padrões linguísticos e traços de personalidade para aprimorar a interpretação dos dados.

2.5.1 Abordagens manuais e automáticas

As abordagens manuais envolvem a análise qualitativa das interações, identificando padrões de comportamento e avaliando a clareza e coerência das respostas. Ranieri, Di Bernardo e Mele (2024) utilizam a análise de discurso qualitativa para examinar interações entre usuários e *chatbots*, identificando desafios como respostas insatisfatórias e dificuldades na compreensão das intenções do usuário, fatores que podem impactar negativamente a experiência do cliente. Nguyen, Waizenegger e Techatassanasoontorn (2022) analisaram *logs* de conversação para identificar seis tipos distintos de interação humano-*chatbot*, variando em termos de presença social e complexidade do diálogo. Além disso, Song *et al.* (2022) argumentam que usuários com alta necessidade de interação humana percebem falhas na comunicação *chatbot* com mais intensidade. O que torna a avaliação manual essencial para compreender essas nuances.

Por outro lado, as abordagens automáticas utilizam algoritmos para processar grandes volumes de interações. Vo *et al.* (2021) demonstraram que técnicas como *Term Frequency-Inverse Document Frequency* (TF-IDF) e *Linguistic Inquiry and Word Count* (LIWC) são eficazes para extrair traços linguísticos de interações verbais, enquanto *Shapley Additive Explanations* (SHAP) auxilia na explicabilidade de modelos. Ngai *et al.* (2021) propuseram um sistema de *chatbot* baseado em conhecimento para suporte ao cliente no comércio eletrônico, incorporando tecnologias como processamento de linguagem natural e inteligência artificial. Paul *et al.* (2018) desenvolveram um *framework* de *chatbot* contextual

baseado em Redes Neurais Convolucionais (CNN - *Convolutional Neural Network*) para interações em domínios específicos, permitindo a adaptação do sistema sem depender de regras pré-definidas ou grandes bases de dados linguísticas.

2.5.2 Mineração de texto, análise de sentimentos e detecção de emoções

A mineração de texto é utilizada para extrair informações das conversas, auxiliando na identificação de intencionalidade e sentimentos expressos pelos usuários. Yim (2024) demonstra que *chatbots* mais interativos e responsivos podem impactar a experiência do usuário ao reduzir emoções negativas e aumentar a percepção de empatia, permitindo ajustes estratégicos na interação para minimizar *feedbacks* negativos e melhorar a retenção do cliente. Zhang, Følstad e Bjørkli (2023) alegam que a comunicação eficaz no uso de *chatbots*, aliada a boas práticas organizacionais, pode influenciar positivamente a experiência do usuário e minimizar frustrações durante a interação.

A análise de sentimentos desempenha um papel essencial na categorização das emoções dos usuários, permitindo que *chatbots* ajustem suas respostas conforme o tom emocional identificado. Fan, Gao e Han (2023) indicam que *chatbots* que equilibram eficiência e flexibilidade no atendimento são mais eficazes na criação de experiências inteligentes, impactando tanto aspectos hedônicos (prazer e envolvimento) quanto cognitivos (eficiência e funcionalidade) na interação com os clientes. Além disso, Yim (2024) demonstra que a interatividade dos *chatbots* influencia a percepção de empatia dos usuários, reduzindo emoções negativas como a raiva e minimizando a intenção de compartilhar *feedback* negativo.

A análise das emoções desempenha um papel essencial na compreensão do impacto das respostas dos *chatbots* na experiência do usuário. Ranieri, Di Bernardo e Mele (2024) destacam que a qualidade da interação e a confiança digital são fatores determinantes na percepção do usuário e podem influenciar seu engajamento com o serviço automatizado. Yim (2024) complementa essa perspectiva ao demonstrar que *chatbots* interativos e responsivos podem reduzir emoções negativas, como raiva, e minimizar a intenção de compartilhar *feedback* negativo, contribuindo para uma experiência mais satisfatória. Além disso, a interatividade dos *chatbots* influencia a percepção de empatia dos usuários, ajudando a criar interações mais humanizadas e engajadoras, o que pode reduzir emoções negativas e minimizar a intenção de compartilhar *feedback* negativo (Yim, 2024).

2.5.3 Frameworks de análise: métricas de empatia e clareza

Diferentes *frameworks* têm sido propostos para avaliar a qualidade das respostas dos *chatbots*. Ngai *et al.* (2021) discutem a efetividade das interações automatizadas a partir de métricas que avaliam a personalização e a precisão das respostas, ressaltando a importância da adaptação do *chatbot* às necessidades do usuário para uma comunicação mais eficiente. Entre os principais indicadores destacam-se:

- a) **reformulações do usuário (D3 - confiança e clareza):** a necessidade de reformulação por parte do usuário pode estar associada a desafios do *chatbot* na interpretação da intenção do usuário, impactando a fluidez da interação (Nguyen; Waizenegger; Techatassanasoontorn, 2022);
- b) **tempo de permanência na conversa (D4 - engajamento e empatia):** a duração da interação pode indicar tanto engajamento quanto frustração, variando conforme o contexto. No setor financeiro, interações curtas são preferíveis, enquanto no varejo, conversas mais longas podem indicar maior envolvimento do usuário (Przegalinska *et al.*, 2019);
- c) **solicitações de atendimento humano (D1 - desistência):** refletem insatisfação com o modelo de IA, evidenciando falhas na resolução automatizada e no atendimento às expectativas do usuário (Ranieri; Di Bernardo; Mele, 2024).
- d) **taxa de abandono da conversa (D1 - desistência):** pode ser usada como um indicador de insatisfação do usuário, estando associada a dificuldades na interpretação da intenção do usuário e na obtenção de informações desejadas (Nguyen; Waizenegger; Techatassanasoontorn, 2022);
- e) **níveis de interatividade percebida (D4 - engajamento e empatia):** a interatividade dos *chatbots*, incluindo fatores como interatividade social e de mensagem, pode influenciar a percepção de empatia dos consumidores, reduzindo a raiva e diminuindo a intenção de word-of-mouth negativo (Yim, 2024).

Essas métricas permitem um monitoramento do impacto dos *chatbots* e auxiliam na implementação de melhorias nos sistemas de modelos de IA conversacional. Por fim, a análise de interações entre *chatbots* e usuários deve integrar abordagens qualitativas e quantitativas para obter uma compreensão holística da experiência do usuário. Estudos como os de Vo *et al.* (2021) e Ngai *et al.* (2021) demonstram que o uso de técnicas de mineração de texto e

modelagem de linguagem pode aprimorar a análise de interações automatizadas, permitindo sistemas mais precisos e adaptáveis às necessidades dos usuários.

2.6 CLAREZA E COERÊNCIA DOS MODELOS CONVERSACIONAIS

A clareza e coerência das respostas de *chatbots* são aspectos fundamentais para garantir interações eficazes e uma experiência satisfatória para os usuários. Estudos recentes exploram como a inteligência artificial pode aprimorar a qualidade da comunicação, reduzindo ambiguidades e aumentando a confiança dos usuários nos sistemas conversacionais.

Chatbots devem ser capazes de fornecer respostas que sejam compreensíveis e contextualmente apropriadas. Ngai *et al.* (2021) mencionam os princípios cooperativos de Grice (1975) ao discutir a avaliação da qualidade das interações entre *chatbots* e usuários, enfatizando a importância de respostas precisas, relevantes e organizadas para melhorar a experiência do usuário.

Uma interação clara e coerente é essencial para a satisfação do usuário. Fan, Gao e Han (2023) evidenciam que a capacidade do *chatbot* de responder com precisão e de forma adaptada às necessidades do usuário está associada a uma percepção elevada de qualidade no serviço. Quando há equilíbrio entre agilidade e adaptação nas interações, o usuário tende a confiar mais no modelo conversacional e a retornar ao serviço. Em contrapartida, respostas vagas ou inconsistentes prejudicam essa percepção e enfraquecem o vínculo de confiança, elemento essencial em experiências digitais positivas, o que reforça D3 (confiança e clareza). Por outro lado, quando as respostas de um *chatbot* são vagas ou inconsistentes, há um impacto negativo na percepção do serviço. Grimes, Schuetzler e Giboney (2021) aplicam a *Expectation Violation Theory* para demonstrar que usuários que esperam um alto nível do modelo da IA e se deparam com falhas na coerência tendem a relatar maior frustração e insatisfação.

2.6.1 Métodos para avaliar clareza e coerência dos modelos conversacionais

Ranieri, Di Bernardo e Mele (2024) analisaram mais de 7.000 interações entre usuários e *chatbots*, identificando erros frequentes como respostas genéricas, repetição desnecessária e falta de compreensão do contexto. Esses fatores levam à redução da taxa de retenção e ao aumento da solicitação de atendimento humano, reforçando D1 (desistência). Vo

et al. (2021) demonstram que a análise de *logs* de chamadas pode revelar padrões de comunicação associados à experiência do usuário, permitindo *insights* sobre a relação entre estilo de conversação e comportamento do cliente.

Recomendações são percebidas com base na revisão teórica para a avaliação da clareza e coerência incluem:

- a) **identificação de padrões de reformulação:** a contagem de reformulações feitas pelo usuário pode ser interpretada como um possível indicativo de baixa clareza nas respostas da IA. No estudo de Nguyen, Waizenegger e Techatassanasoontorn (2022), observou-se que, em interações colaborativas, os usuários frequentemente ajustam suas perguntas para facilitar a compreensão por parte do *chatbot* - o que pode indicar desafios na interpretação da intenção do usuário.
- b) **análise de indicadores linguísticos de insatisfação:** a presença de expressões como “não entendi”, “explique melhor” ou “não faz sentido” pode indicar falhas na clareza das respostas da IA. Ashfaq *et al.* (2020) destacam que a qualidade da informação e do serviço influencia diretamente a satisfação do usuário. Com base nesses achados, pode-se inferir que problemas nesses aspectos podem aumentar a necessidade de reformulações na interação com *chatbots*;
- c) **correlação entre clareza e escalonamento para atendimento humano:** a solicitação de um atendente humano após uma resposta do *chatbot* pode indicar que a resposta foi insuficiente ou pouco clara. Conforme identificado por Ranieri, Di Bernardo e Mele (2024), respostas imprecisas ou repetitivas dos *chatbots* podem levar à frustração do usuário e impactar negativamente sua experiência;
- d) **tempo de interação e engajamento como reflexo da clareza:** de acordo com Przegalinska *et al.* (2019), a confiança e a interação dos usuários com *chatbots* são fatores importantes na avaliação de sua performance. O tempo de interação, que varia conforme o segmento, pode estar associado à satisfação do usuário, especialmente quando o *chatbot* cumpre seu papel esperado de forma eficaz.

Nguyen, Waizenegger e Techatassanasoontorn (2022) indicam que a interação humano-*chatbot* pode ser impactada negativamente por falhas na comunicação, levando à frustração do usuário. Em contrapartida, Przegalinska *et al.* (2019) destacam que a confiança e o engajamento do usuário podem ser influenciados pela forma como a interação é estruturada. Assim, pode-se inferir que conversas bem organizadas contribuem para uma experiência mais satisfatória.. Com base em estudos sobre análise de *logs* de conversação para predição de

comportamento do usuário (Vo *et al.*, 2021), é possível explorar padrões linguísticos e traços de personalidade como indicadores úteis na compreensão da relação cliente-empresa.

2.7 EMPATIA E CONEXÃO HUMANA NAS RESPOSTAS

A empatia em interações automatizadas tem se mostrado crucial para a qualidade conversacional de usuários com *chatbots*. Os usuários tendem a atribuir características humanas aos *chatbots* (antropomorfismo), o que influencia positivamente sua intenção de adoção e a forma como percebem a interação com modelos conversacionais com inteligências artificiais (Sheehan, Jin e Gottlieb, 2020). No entanto, quando um *chatbot* não atende a essas expectativas, pode ocorrer uma violação das expectativas do usuário, gerando frustração e desconfiança na tecnologia (Grimes; Schuetzler; Giboney, 2021). Conectado à D1 (desistência), essa violação de expectativas é potencializada pela antropomorfização do *chatbot*, levando os usuários a terem maior probabilidade de abandonar a conversa ou exigir um atendente humano.

Chatbots que oferecem respostas personalizadas podem influenciar a satisfação do usuário e melhorar a experiência no atendimento ao cliente (Nicolescu; Tudorache, 2022), sustentando a D3 (confiança e clareza), pois respostas contextuais e bem formuladas aumentam a percepção de presença social, incentivando a continuidade da interação com o *chatbot* e reduzindo a solicitação de suporte humano.

2.7.1 Métodos para avaliar empatia em *chatbots*

A implementação da empatia em *chatbots* pode ser feita por meio de diversas estratégias, incluindo a personalização das respostas, a adaptação do tom de voz e o uso de pistas sociais para melhorar a experiência do usuário (Camilleri; Troise, 2023). Essas estratégias reforçam a D4 (engajamento e empatia da IA), uma vez que interações mais humanizadas promovem o engajamento emocional dos usuários. Algumas das principais abordagens incluem:

- a) **uso de linguagem natural e expressões afetivas:** *chatbots* que utilizem expressões amigáveis e linguagem mais próxima da interação humana tendem a criar uma conexão maior com os usuários (Sheehan, Jin e Gottlieb, 2020);
- b) **personalização baseada em histórico de interações:** a personalização baseada no histórico de interações permite que o *chatbot* demonstre conhecimento sobre

as necessidades do usuário, aumentando a percepção de atendimento humanizado. Camilleri e Troise (2023) ressaltam que os *chatbots* podem utilizar dados anteriores dos clientes para aprimorar a experiência e personalizar os serviços oferecidos;

- c) **capacidade de reformulação e explicação:** quando um *chatbot* não compreende a solicitação do usuário, buscar esclarecimentos em vez de oferecer respostas genéricas pode melhorar a experiência do usuário (Sheehan, Jin e Gottlieb, 2020);
- d) **emprego de estratégias de humor e encorajamento:** características antropomórficas, como identidade, interação social e *small talk* (conversas leves e informais), podem influenciar a experiência do usuário com *chatbots* e aumentar a percepção de socialização e envolvimento na interação, tornando-a mais agradável (Nicolescu; Tudorache, 2022).

Para medir a efetividade dessas estratégias, alguns estudos sugerem métricas como a análise da satisfação e da intenção de uso dos *chatbots* pelos usuários (Ashfaq *et al.*, 2020). Além disso, outros fatores, como tempo de permanência e reformulações durante a conversa, podem ser considerados em avaliações complementares.

2.7.2 Estudos de caso em que a empatia melhorou a experiência do usuário

Diferentes estudos de caso ilustram como a empatia pode melhorar a experiência do usuário com *chatbots* em serviços automatizados:

Caso 1, atendimento ao cliente com *chatbots* antropomórficos: Um estudo de Sheehan, Jin e Gottlieb (2020) demonstrou que *chatbots* percebidos como mais antropomórficos — ou seja, mais semelhantes a humanos em sua comunicação — tendem a ser mais adotados pelos usuários. Esse efeito foi especialmente forte entre usuários com alta necessidade de interação humana..

Caso 2, *chatbots* brincalhões vs. fofos: Um estudo de Zhang *et al.* (2022) testou duas estratégias de *chatbots* para acalmar clientes insatisfeitos com um atraso na entrega: a abordagem *whimsical* (brincalhona), que utiliza humor e tom descontraído, e a abordagem *kindchenschema* (fofa e vulnerável), que evoca empatia ao se apresentar de maneira infantil ou adorável. Os resultados indicaram que ambas as estratégias reduziram emoções negativas dos clientes, sendo que a abordagem *whimsical* foi mais eficaz para clientes com alta

ansiedade tecnológica e do sexo masculino, enquanto a abordagem *kindchenschema* teve maior impacto entre clientes com baixa ansiedade tecnológica e do sexo feminino

Caso 3, suporte técnico empático: Ashfaq *et al.* (2020) investigaram a satisfação de usuários com *chatbots* de atendimento, revelando que respostas empáticas e de qualidade são capazes de reduzir a frustração dos clientes, especialmente em interações que envolvem dúvidas e problemas técnicos.

Esses casos demonstram que a inclusão de empatia na comunicação dos *chatbots* pode melhorar significativamente a percepção do usuário, reduzindo frustrações e aumentando o engajamento. Além disso, quando os *chatbots* conseguem demonstrar empatia, os usuários estão mais dispostos a confiar e continuar utilizando o serviço.

2.8 EFICÁCIA NA RESOLUÇÃO DE NECESSIDADES DO USUÁRIO

A eficácia de um *chatbot* está relacionada à sua capacidade de fornecer respostas precisas e personalizadas, adaptando-se às necessidades dos usuários para uma interação mais eficiente (Ngai *et al.*, 2021). Essa efetividade não se restringe apenas à precisão das respostas, mas também ao impacto na experiência do usuário, influenciando sua percepção de confiança e engajamento com a tecnologia (Ranieri; Di Bernardo; Mele, 2024).

Um *chatbot* eficaz pode reduzir a necessidade de intervenção humana ao resolver solicitações de forma autônoma, especialmente em interações repetitivas. No entanto, sua eficácia depende da consistência das respostas e da capacidade de escalar para suporte humano quando necessário (Zhang; Følstad; Bjørkli, 2023). Isso está diretamente relacionado à dimensão D1 (desistência), que sugere que usuários que precisam de suporte humano ou abandonam a conversa demonstram maior insatisfação. Além disso, a adoção de *chatbots* com modelos de inteligência artificial tem demonstrado impacto na eficiência operacional e na otimização da tomada de decisão em pequenas e médias empresas, contribuindo para um desempenho sustentável e maior visibilidade da cadeia de suprimentos (Panigrahi *et al.*, 2023).

Dessa forma, pode-se inferir que a eficácia técnica, por si só, não garante o sucesso desses sistemas: a aceitação do *chatbot* pelos usuários torna-se um critério central para determinar sua efetividade real. Nesse sentido, estudos apontam que a percepção de utilidade, confiança e adequação da interação influencia diretamente a aceitação e continuidade do uso de *chatbots* (Ashfaq *et al.*, 2020; Nicolescu; Tudorache, 2022).

2.8.1 Métodos para avaliação das respostas/interações

A avaliação da eficácia das interações entre usuários e *chatbots* pode ser realizada por meio de diferentes abordagens. Dentre os métodos mais utilizados, destacam-se:

- a) **análise de logs de conversação:** esse método permite extrair padrões de interação e entender se as respostas do *chatbot* estão de fato resolvendo as necessidades do usuário. Vo *et al.* (2021) exploraram a análise de *logs* de chamadas para prever o comportamento do usuário. Os autores enfatizam a extração de traços de personalidade e informações lexicais como principais fontes de *insights* para modelos preditivos. Esses métodos podem ser adaptados para avaliar métricas como tempo de resposta e reformulações em interações, possibilitando a identificação de falhas na performance, incluindo respostas irrelevantes;
- b) **modelos baseados em machine learning:** alguns estudos propõem o uso de aprendizado de máquina para prever o comportamento do *chatbot* com base em características da conversação. Vo *et al.* (2021) demonstraram que modelos baseados em dados textuais e estruturais extraídos de *logs* de atendimento podem prever o risco de *churn*, analisando padrões linguísticos e variáveis comportamentais dos clientes. O estudo utiliza *machine learning* e técnicas de mineração de texto para extrair *insights* sobre a comunicação dos usuários, contribuindo para o desenvolvimento de estratégias personalizadas de retenção;
- c) **taxa de resolução na primeira interação:** um indicador utilizado na literatura é a capacidade do *chatbot* de resolver problemas sem necessidade de escalonamento para atendimento humano. Estudos como o de Ranieri, Di Bernardo e Mele (2024) analisam os efeitos dos *chatbots* na experiência do cliente, destacando que a qualidade da interação e a eficiência na resolução de problemas influenciam a percepção de confiança e satisfação.
- d) **percepção do usuário e feedbacks qualitativos:** além das métricas quantitativas, alguns pesquisadores argumentam que a coleta de *feedbacks* diretos dos usuários pode fornecer *insights* sobre a experiência com o *chatbot*. Ranieri, Di Bernardo e Mele (2024) analisaram as interações de um *chatbot* de varejo e identificaram que respostas repetitivas ou pouco úteis podem indicar problemas de eficácia.

2.8.2 Análise de limitações frequentes em *chatbots* e como abordá-las

Mesmo com avanços na inteligência artificial, os *chatbots* ainda enfrentam desafios que comprometem sua eficácia na resolução de necessidades dos usuários. Dentre os principais problemas, destacam-se:

- a) **alucinações e respostas irrelevantes:** *chatbots* que geram respostas sem sentido ou que não atendem ao contexto da interação, além da baixa responsividade podem comprometer a experiência do usuário e impactar negativamente sua satisfação (Chen; Le; Florence, 2021 e Vessey, 1991);
- b) **falta de adaptação ao contexto do usuário:** Nguyen, Waizenegger e Techatassanasoontorn (2021) destacam que *chatbots* sem capacidade de reter informações de interações passadas tendem a fornecer respostas genéricas e pouco úteis, o que pode impactar negativamente a experiência do usuário. Para mitigar essa limitação, estudos na área sugerem o uso de técnicas de aprendizado contínuo e aprimoramento progressivo da personalização das respostas ao longo das interações;
- c) **dificuldade em lidar com solicitações complexas:** alguns *chatbots* apresentam dificuldades quando enfrentam consultas que exigem raciocínio avançado ou integração de múltiplos contextos. O estudo de Ngai *et al.* (2021) discute a importância de arquiteturas baseadas em conhecimento para aprimorar a precisão das respostas, destacando que a combinação de diferentes abordagens pode melhorar a capacidade dos *chatbots* de lidar com interações mais complexas;
- d) **falta de empatia e tom de voz inadequado:** o estudo de Fan, Gao e Han (2023) sugere que a forma como os *chatbots* conduzem as interações pode impactar a experiência do usuário, tornando-a mais envolvente e satisfatória.

A revisão da literatura evidencia que a aceitação dos *chatbots* estão diretamente relacionadas a fatores como clareza das respostas, tempo de resposta, ocorrência de alucinações e inclusão de elementos de empatia. Conforme demonstrado no Quadro 6, diversos estudos fornecem embasamento teórico para essas relações, destacando que a experiência do usuário com *chatbots* pode ser impactada por múltiplas variáveis interdependentes.

Esses desafios evidenciam que, apesar dos avanços tecnológicos, ainda existem limitações na experiência proporcionada pelos *chatbots*. A literatura apresenta diferentes

abordagens para mitigar esses problemas, mas há lacunas quanto à avaliação do impacto dessas variáveis na desistência conversacional entre usuários e agentes conversacionais.

2.9 EMBASAMENTO DAS DIMENSÕES DE PESQUISA

Estudos anteriores investigaram aspectos como satisfação do usuário, percepção de presença social, impacto de respostas imprecisas e o papel da empatia nas interações conversacionais. Com base nessa estrutura teórica, cada dimensão proposta é sustentada por evidências de múltiplos autores, conforme sintetizado no Quadro 6.

Quadro 6 - Evidências para as dimensões de pesquisa

Dimensão	Autores que embasam a construção das dimensões
D1 (desistência): usuários que buscam falar com um ser humano ou abandonam a conversa demonstram maior insatisfação com o <i>chatbot</i> em comparação com aqueles que finalizam a interação de forma natural.	Adam; Wessel e Benlian (2022); Ashfaq <i>et al.</i> (2020); Blümel; Zaki e Bohné (2024); Chang e Hsiao (2024); Chen; Le e Florence (2021); Crolic <i>et al.</i> (2022); Ranieri, Di Bernardo e Mele (2024); Zhang; Følstad e Bjørkli (2023).
D2 (performance): a ocorrência de alucinações nas respostas no modelo de IA conversacional (respostas irrelevantes ou fora do contexto) e maior tempo de resposta aumentam significativamente a insatisfação do usuário e a probabilidade de abandono da conversa.	Ashfaq <i>et al.</i> (2020); Chen; Le e Florence (2021); Grimes; Schuetzler e Giboney (2021); Przegalinska <i>et al.</i> (2019). (Vessey, 1991)
D3 (confiança e clareza): respostas claras e contextuais fornecidas pelo modelo de IA conversacional aumentam a percepção de presença social, o que, por sua vez, reduz a solicitação de suporte humano e incentiva a continuidade da interação com o <i>chatbot</i> .	Adam; Wessel e Benlian (2022); Becker; Mahr e Odekerken-Schröder (2023); Camilleri e Troise (2023); De Andrade e Tumelero (2022); Nguyen, Waizenegger e Techatassanasoontorn (2022); Sheehan, Jin e Gottlieb (2020).
D4 (engajamento e empatia): a inclusão de elementos de empatia nas respostas contribui para reduzir a solicitação de suporte humano e aumentar o engajamento dos usuários.	Blümel, Zaki e Bohné (2024); Li; Gan e Zheng (2023); Ranieri, Di Bernardo e Mele (2024); Xie <i>et al.</i> (2024); Zhang e Giles (2016).

Fonte: Elaborado pelo autor (2026).

2.10 CONCLUSÃO E LACUNAS NA LITERATURA

Com base em Vo *et al.* (2021), pode-se inferir que a análise de padrões conversacionais extraídos de *logs* de chamadas contribui para enriquecer modelos preditivos,

indicando a necessidade de ir além de métricas tradicionais ao avaliar o comportamento do usuário. Adam, Wessel e Benlian (2022) ressaltam que características como presença social (*social presence*) e pistas antropomórficas nos *chatbots* podem impactar a conformidade do usuário e sua experiência na interação, indo além de métricas puramente operacionais. Em vez disso, é necessário considerar elementos subjetivos na avaliação dos *chatbots*, como a percepção da qualidade da comunicação e a capacidade de evitar erros interpretativos (Sheehan; Jin; Gottlieb, 2020), além de empatia e engajamento do usuário (Zhang; Giles, 2016), bem como da percepção de responsividade e adequação nas respostas dos *chatbots* para evitar frustrações com suas limitações (Fan; Gao; Han, 2023).

Os principais achados obtidos apontam que a desistência conversacional e *chatbot* depende da interseção entre fatores técnicos e percepções subjetivas dos usuários. Com base em Nguyen, Waizenegger e Techatassanasoontorn (2022), é possível interpretar que a necessidade de reformulações frequentes nas interações com *chatbots* pode refletir baixa clareza nas respostas, afetando a experiência do usuário e sua percepção sobre a eficácia do assistente virtual. Da mesma forma, Ranieri, Di Bernardo e Mele (2024) destacam que a qualidade da interação entre *chatbot* e usuário influencia diretamente o engajamento e a experiência do cliente. Quando o *chatbot* fornece respostas claras e estruturadas, pode reduzir frustrações e melhorar a satisfação do usuário durante a interação. Segundo Adam, Wessel e Benlian (2022), as respostas dos *chatbots* que não atendem às expectativas dos usuários podem impactar negativamente sua aceitação e experiência de uso.

Apesar dos avanços metodológicos na análise de *logs* conversacionais, ainda há lacunas significativas na literatura. Em particular, a maioria dos estudos prioriza abordagens quantitativas para avaliar a satisfação do usuário, enquanto pesquisas qualitativas permanecem limitadas. Hsu e Lin (2023) empregaram modelagem de equações estruturais para validar a relação entre qualidade da conversa, satisfação e lealdade do usuário, baseando-se em questionários pós-interação, sem utilizar a análise direta das conversas realizadas. O estudo de Ranieri, Di Bernardo e Mele (2024) utiliza a análise de discurso qualitativa para examinar as interações entre clientes e *chatbots*, destacando como essa metodologia pode revelar detalhes importantes sobre a experiência do usuário. Além disso, o artigo enfatiza a necessidade de mais pesquisas para aprofundar a compreensão dos impactos dessa tecnologia.

Ademais, um dos principais desafios identificados é a avaliação de indicadores subjetivos, como empatia e confiança, em escala. Estudos baseados na *Communication Accommodation Theory* (Zhang; Giles, 2016) sugerem que *chatbots* que utilizam pistas emocionais moderadas podem aumentar a percepção de interação social, por mais que a

medição objetiva desses efeitos ainda seja um desafio. Já a abordagem proposta por Fan, Gao e Han (2023) explora a ambidestria dos *chatbots* na criação de experiências emocionais e cognitivas, destacando a importância de ajustes na configuração do atendimento para otimizar a interação com os usuários. No entanto, sua aplicação prática requer ajustes devido a limitações metodológicas, como a falta de um modelo longitudinal e a generalização restrita a um único setor.

A presente pesquisa busca endereçar essa lacuna ao propor uma abordagem mista, combinando métodos quantitativos e qualitativos baseados em *logs* de conversas, utilizando técnicas com ferramentas com modelos de inteligência artificial e modelagem estatística para mensurar as dimensões preditoras que avaliam o comportamento e performance do *chatbot* e mensurar a desistência conversacional na dimensão dependente. Esta pesquisa se fundamenta em dados reais armazenados no artefato, permitindo uma análise objetiva da clareza das respostas, da taxa de reformulação de perguntas e da percepção de empatia por meio de modelagem estatística e análise de sentimentos.

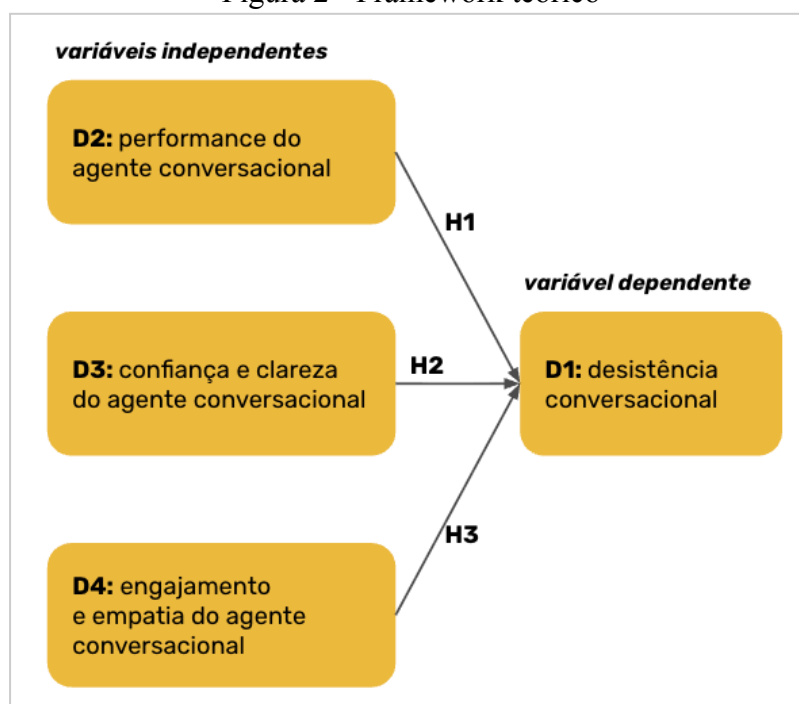
Diante dessas considerações, futuros estudos devem explorar metodologias qualitativas para capturar aspectos subjetivos da experiência do usuário com *chatbots*. Além disso, a integração de indicadores subjetivos com métricas quantitativas pode aprimorar a avaliação de *chatbots*. A investigação do impacto da empatia, da clareza e da personalização nas interações automatizadas permanece um campo promissor para futuras pesquisas, com potencial para influenciar significativamente o *design* de agentes conversacionais e aprimorar a experiência do usuário.

3 FRAMEWORK TEÓRICO PROPOSTO

O presente *framework* teórico busca estruturar a avaliação das interações conversacionais entre *chatbots* e usuários, considerando as dimensões extraídas da literatura. O modelo proposto integra variáveis quantitativas extraídas diretamente dos *logs* de conversas para mensurar a desistência conversacional do modelo em atendimentos automatizados.

Com base nas dimensões de pesquisa, o *framework* permite testar relações entre padrões comportamentais e satisfação do usuário, bem como analisar a performance do *chatbot* em aspectos como clareza das respostas, empatia percebida e engajamento. A estrutura do modelo está organizada conforme a Figura 2, de maneira a capturar as interações desde a entrada dos dados até a avaliação da desistência (D1) da interação humano-agente conversacional.

Figura 2 - Framework teórico



Fonte: Elaborado pelo autor (2026).

O *framework* teórico da Figura 2 (focado nas relações entre as dimensões D1 a D4 e hipóteses) não é um elemento isolado, mas sim o componente central de uma estrutura maior de geração de valor. Dentro da perspectiva do *Design Science Research* (DSR), o modelo teórico atua como o artefato de conhecimento, validando cientificamente as premissas que sustentam a operação prática do *software*. A Figura 3 apresenta a visão total desta arquitetura,

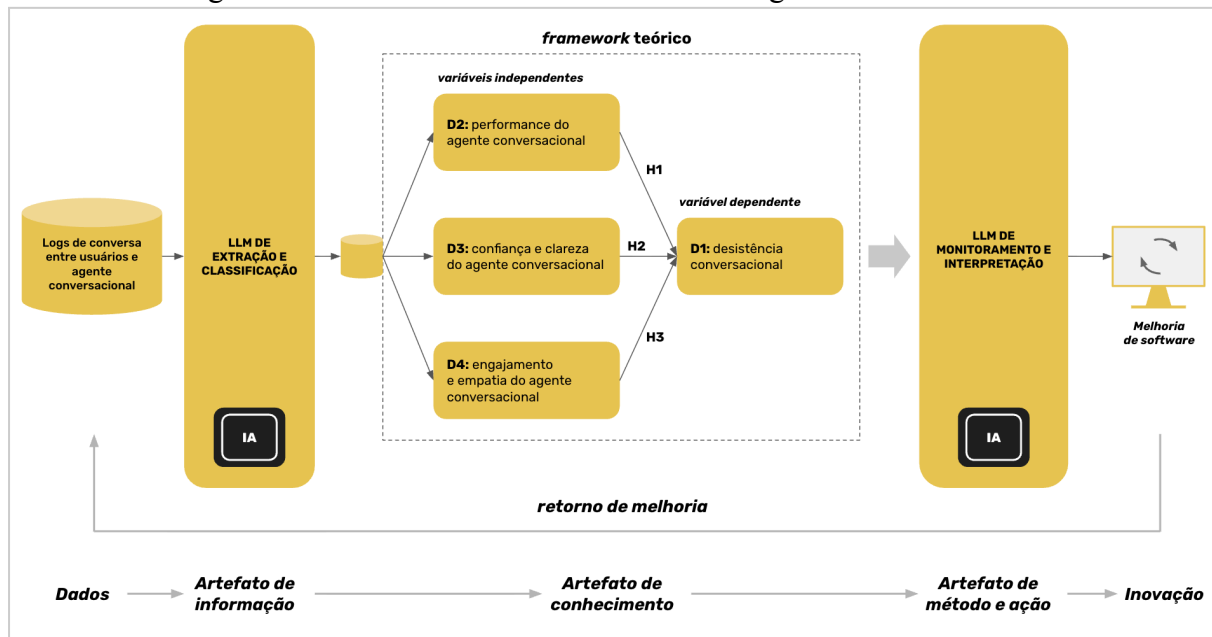
demonstrando como o *framework* teórico está encapsulado dentro de um fluxo contínuo que transforma dados brutos em inovação gerenciável.

3.1 HIPÓTESES DE PESQUISA

Com base nas dimensões definidas no *framework* teórico, surge a hipótese de pesquisa que avalia se a desistência conversacional entre usuário e *chatbot* (D1) - medida por comportamentos como abandono da conversa ou solicitação ou aceitação de atendimento humano - pode ser explicada pelas demais dimensões (D2, D3 e D4), consideradas variáveis independentes. Esta hipótese é esmiuçada em três hipóteses principais:

- a) **hipótese 1:** agentes conversacionais que demonstraram baixa performance (D2) aumentam a desistência conversacional do usuário com *chatbot* (D1);
- b) **hipótese 2:** agentes conversacionais que demonstraram baixa confiança e clareza (D3) aumentam a desistência conversacional do usuário com *chatbot* (D1);
- c) **hipótese 3:** agentes conversacionais que demonstraram menor engajamento e empatia (D4) aumentam a desistência conversacional do usuário com *chatbot* (D1);

Essas dimensões refletem fatores relacionados à performance do modelo de IA, à clareza e confiança geradas na interação, e ao engajamento e empatia percebidos. Validar essa relação fortalece o modelo proposto e amplia a compreensão sobre o comportamento do usuário em interações com agentes conversacionais.

Figura 3 - *Framework* de monitoramento do agente conversacional

Fonte: Elaborado pelo autor (2026).

3.2 ESTRUTURA DO FRAMEWORK DE MONITORAMENTO DO AGENTE CONVERSACIONAL

O *framework* apresentado na Figura 3 está organizado em cinco etapas principais. Segue maior detalhamento de cada etapa:

- dados:** representa a entrada de dados extraídos diretamente das interações registradas na plataforma;
- artefato de informação:** um modelo de IA instruído com foco em analisar cada conversa com usuários e aplicar a classificação da variável de cada dimensão da pesquisa;
- artefato de conhecimento:** *framework* teórico, onde são aplicados métodos estatísticos para validação das hipóteses de pesquisa;
- artefato de método e ação:** um modelo de IA instruído a sintetizar o diagnóstico estatístico e traduzir em diretrizes gerenciais acionáveis, sugerindo estratégias de correção e aprimoramento para o agente conversacional (*chatbot*);
- inovação:** o fluxo culmina na efetivação da inovação, materializada pela melhoria de *software*, com recomendações geradas pelo artefato de método e ação, estabelecendo um ciclo retroalimentativo (*feedback loop*).

3.2.1 Elementos do *framework*

O *framework* desta pesquisa organiza, de forma integrada, os elementos necessários para analisar os agentes conversacionais. A estrutura proposta articula a origem dos dados empíricos, o processamento automatizado das conversas por um modelo de IA e as dimensões analíticas definidas no estudo, permitindo compreender as diferentes características da interação entre *chatbot* e usuários.

3.2.1.1 *Logs* de conversa entre usuários e agente conversacional (dados)

A primeira etapa do *framework* consiste na coleta de dados a partir dos registros de conversas entre usuários e agentes conversacionais da plataforma de mercado, com a demonstração de seu funcionamento no Apêndice C. Para esta pesquisa não foi estabelecido a definição do perfil dos usuários, pois a plataforma não captura informações como idade, gênero e entre outros dados de perfil, pois atua como uma plataforma de atendimento de suporte e vendas. Foram utilizados os dados disponíveis como conversa, datas de interação e nome do usuário. A Figura 4 apresenta um exemplo de histórico de conversa para a análise dos dados. Estes dados foram extraídos da plataforma de mercado com conversas reais de atendimento, no qual será melhor detalhado no capítulo “4.2 Campo de Estudo”.

Figura 4 - Exemplo de conversa real entre agente conversacional e usuário (dados)

Conversa Selecionada ✕

Data e Hora 28/01/2025 08:00:24	Chatbot Tally Nova	Pessoa Dado ocultado
--	------------------------------	--------------------------------

USUÁRIO 06:38 - 11/11/24
Bom dia

ASSISTENTE 06:38 - 11/11/24
Bom dia! Sou Tally, uma inteligência artificial de atendimento da ocultado. Como posso ajudá-lo? Qual é o seu nome?

USUÁRIO 06:38 - 11/11/24
Quero saber mais sobre a ferramenta. Consigo enviar fotos, áudios e vídeos automaticamente

ASSISTENTE 06:38 - 11/11/24
A IA pode ouvir áudios, mas responde em texto. Não é possível enviar fotos ou vídeos automaticamente. Posso ajudar com mais informações sobre a ferramenta?

USUÁRIO 06:39 - 11/11/24
Não dá para programar envio de áudios e vídeos?

ASSISTENTE 06:39 - 11/11/24
Não, ocultado não permite o envio automático de áudios ou vídeos. Posso esclarecer mais sobre as funcionalidades da ferramenta?

Fonte: Elaborado pelo autor (2026).

3.2.1.2 LLM de extração e classificação (artefato de informação)

A segunda etapa do *framework* consiste na análise das interações das conversas entre usuários e os agentes conversacionais. Para garantir uma avaliação em escala com padronização, foi desenvolvido um *prompt* de análise, dentro do artefato de pesquisa, descrito no Apêndice D, que permite a extração de métricas diretamente dos registros de conversação. Para esta pesquisa foi utilizado o ChatGPT-4o mini (OpenAI), com estruturas de comandos (*prompts*) dentro do artefato para avaliação de cada variável das dimensões.

3.2.1.3 *Framework* teórico (artefato de conhecimento)

Nesta etapa o *framework* teórico, apresentado na Figura 2. A seguir será detalhado as variáveis compostas nas dimensões.

Variáveis da Dimensão 1 - desistência conversacional (variável dependente):

- a) **aceitação de recomendação de atendimento humano pelo agente conversacional ou solicitação de atendimento humano:** com base em Ranieri, Di Bernardo e Mele (2024), as limitações do atendimento automatizado podem gerar estresse e frustração, o que interpreta-se para uma possível solicitação de suporte humano diante de experiências negativas. Já Ashfaq *et al.* (2020) reforçam que, embora os *chatbots* sejam amplamente utilizados, muitos usuários ainda demonstram preferência pela interação humana, especialmente quando possuem maior necessidade de contato interpessoal;
- b) **abandono da conversa (considerando distinção entre abandono por insatisfação e abandono natural):** com base em Ranieri, Di Bernardo e Mele (2024), a insatisfação do usuário com a interação, especialmente diante de respostas imprecisas ou pouco satisfatórias, pode contribuir para o abandono da conversa.

Variáveis da Dimensão 2 - performance do agente conversacional (variável independente):

- a) **taxa de alucinações:** Com base em Vessey (1991) e Chen, Le e Florence (2021), a falta de alinhamento entre a necessidade do usuário e a forma ou qualidade da resposta do *chatbot*, como no caso de respostas irrelevantes ou incorretas, pode comprometer a experiência e reduzir a confiabilidade percebida do sistema;
- b) **taxa de reformulação de perguntas:** com base na análise da qualidade da informação e do serviço discutida por Ashfaq *et al.* (2020), sugere-se que reformulações frequentes por parte dos usuários possam estar associadas a desafios na clareza da comunicação com *chatbots*, impactando negativamente a experiência do usuário;
- c) **tempo de resposta do chatbot:** com base em Chang e Hsiao (2024), barreiras funcionais na interação com *chatbots*, como a qualidade e respostas lentas, podem impactar a experiência do usuário e gerar percepções negativas sobre a eficiência do sistema.

Variáveis da Dimensão 3 - confiança e clareza do agente conversacional (variável independente):

- a) **taxa de reformulação de perguntas:** com base em Nguyen, Waizenegger e Techatassanasoontorn (2022), a taxa de reformulação de perguntas pode refletir desafios na comunicação entre usuário e agente conversacional, incluindo possíveis problemas na clareza das respostas do *chatbot*;
- b) **expressões de frustração ou dúvida:** com base em Chang e Hsiao (2024), a frustração gerada por barreiras funcionais e psicológicas na interação com *chatbots* pode ser um indicativo de insatisfação dos usuários;
- c) **mudança de tom emocional:** com base em Blümel, Zaki e Bohné (2024), a personalização do tom na interação com o agente conversacional pode influenciar a experiência do usuário, ajudando a reduzir frustrações e tornar o atendimento mais envolvente;
- d) **clareza e coerência das respostas:** com base em Sheehan, Jin e Gottlieb (2020), a percepção de confiabilidade do usuário em relação aos *chatbots* está ligada à sua capacidade de evitar erros de comunicação ou buscar esclarecimentos quando necessário.

Variáveis da Dimensão 4 - engajamento e empatia do agente conversacional (variável independente):

- a) **taxa de expressões empáticas:** com base em Zhang e Giles (2016), a taxa de expressões empáticas mede a presença de linguagem empática nas respostas do agente conversacional, impactando a percepção de humanização e engajamento do usuário;
- b) **engajamento do usuário (tempo de conversa):** com base em Zhang, Følstad e Bjørkli (2023), o engajamento do usuário pode ser medido pelo tempo médio que ele permanece na interação com o *chatbot*, refletindo sua disposição para continuar a conversa;
- c) **permanência do usuário na interação:** com base em Ranieri, Di Bernardo e Mele (2024), a permanência do usuário na interação pode ser avaliada pela quantidade de mensagens trocadas em uma sessão, o que pode refletir tanto a continuidade da conversa quanto possíveis dificuldades na obtenção de respostas adequadas;
- d) **tom da conversa:** com base em Blümel, Zaki e Bohné (2024), o tom e o estilo da conversação do modelo de IA podem influenciar a percepção do usuário, afetando sua satisfação e engajamento com o *chatbot*.

3.2.1.4 LLM de monitoramento e interpretação

A penúltima etapa do *framework* proposto consiste na instanciação de um modelo de IA instruído para a interpretação dos resultados estatísticos e geração de sugestões de melhoria. A inclusão desta fase fundamenta-se na necessidade de traduzir achados quantitativos complexos (como coeficientes de regressão e significância estatística) em ações gerenciais práticas e imediatas, fechando o ciclo de retroalimentação (*feedback loop*) do sistema.

Enquanto a análise estatística valida cientificamente quais variáveis estão impactando a desistência conversacional (D1), este módulo final atua na camada de análise prescritiva. Ele utiliza um modelo de linguagem alimentado com os *insights* gerados pela regressão e com as diretrizes de melhores práticas de *design* conversacional.

3.2.1.5 Melhoria de software (método)

Por fim, o fluxo do *framework* culmina na etapa de inovação, materializada pela efetiva melhoria de *software*. Esta fase representa a operacionalização das diretrizes geradas pelo módulo prescritivo, onde as recomendações de ajuste - sejam elas o refinamento de *prompts*, a recalibragem do tom de voz ou a correção de lacunas na base de conhecimento - são implementadas diretamente no ambiente de produção do agente conversacional.

O objetivo central desta etapa é converter o conhecimento teórico validado em valor prático tangível. Ao concretizar as alterações sugeridas, o artefato estabelece um ciclo virtuoso de otimização contínua, garantindo que a evolução tecnológica do *chatbot* não seja pautada em suposições, mas guiada por evidências empíricas e análises estatísticas. Desta forma, a pesquisa transcende a observação passiva para intervir ativamente na qualidade da interação humano-computador.

4 MÉTODO

Esta pesquisa adota o método *Design Science Research* (DSR), conforme proposto por Hevner *et al.* (2004), por envolver o projeto, desenvolvimento e avaliação de um artefato computacional destinado a resolver um problema prático relevante no contexto organizacional. No contexto deste trabalho, o artefato consiste em um *framework* multidimensional para avaliação das variáveis das dimensões propostas, bem como em sua instanciação tecnológica, materializada em um *software* desenvolvido. A avaliação do artefato utiliza uma abordagem quantitativa, baseada na análise de *logs* reais de conversação entre usuários e agentes conversacionais, permitindo mensurar objetivamente variáveis comportamentais associadas à desistência, performance, clareza, confiança, engajamento e empatia.

4.1 DELINEAMENTO DA PESQUISA

O delineamento da pesquisa caracteriza-se como observacional e descritivo, uma vez que não há interferência direta no processo de interação entre usuários e *chatbots*, limitando-se à análise posterior dos registros gerados pelo sistema. A partir desses dados, são aplicadas técnicas estruturadas de extração, organização e interpretação das informações, viabilizando a identificação de padrões comportamentais recorrentes e a validação empírica das dimensões propostas no *framework* teórico da pesquisa.

4.2 CAMPO DE ESTUDO

O campo de estudo compreende as interações registradas em uma plataforma de mercado anonimizada, um sistema de atendimento automatizado que utiliza modelos de IA conversacionais integrados com a plataforma WhatsApp para interagir com usuários em diferentes contextos. Os dados analisados foram extraídos diretamente dos registros de conversação da plataforma em três cenários distintos, sendo:

- a) o cenário da própria plataforma de mercado, com um agente conversacional instruído a vender a própria plataforma (Apêndice E descreve o contexto de instrução do agente);

- b) o cenário de uma fábrica de *software*, com o agente conversacional instruído a vender os serviços de consultoria de *software* da empresa (Apêndice F descreve o contexto de instrução do agente);
- c) o cenário de uma escola de inglês, com um agente conversacional instruído a vender os serviços da empresa (Apêndice G descreve o contexto de instrução do agente).

Para cada cenário foi extraído um número relativo de conversas, seguindo os critérios estabelecidos a seguir.

4.3 PROCEDIMENTO DE COLETA DE DADOS

O procedimento de coleta de dados desta pesquisa foi estruturado de forma a garantir a obtenção de informações empíricas consistentes e representativas das interações entre usuários e *chatbots* baseados em inteligência artificial. Para isso, adotou-se a extração direta dos registros de conversação armazenados em uma plataforma de mercado, preservando o contexto completo das interações e assegurando a integridade dos dados analisados.

4.3.1 Dados

A base de dados da pesquisa é composta por *logs* de conversação extraídos de uma plataforma de mercado. Cada conversa registrada inclui detalhes sobre a interação do usuário com o *chatbot*, permitindo uma análise aprofundada dos padrões conversacionais e da eficácia do agente conversacional na resolução de necessidades. Para garantir a privacidade e a proteção dos dados, todas as informações foram devidamente anonimizadas, em conformidade com a Lei Geral de Proteção de Dados (LGPD), assegurando que nenhuma identificação individual possa ser realizada a partir dos registros analisados. Os dados coletados incluem:

- a) conteúdo das mensagens (entrada do usuário e resposta do *chatbot*);
- b) tempo de resposta do *chatbot* (intervalo entre envio e recebimento de mensagens);
- c) data e hora do início da primeira interação pelo usuário;
- d) data e hora do envio de cada mensagem pelo usuário e *chatbot*;
- e) identificação do *chatbot*;
- f) nome do usuário o qual iniciou a interação com o *chatbot* (dado ocultado).

Como critério de filtro para analisar conversas relevantes, foram selecionadas apenas interações com pelo menos 10 trocas de mensagens. Esse volume mínimo foi estabelecido por uma necessidade técnica do artefato de informação (LLM), para que as variáveis comportamentais - como taxas de reformulação, mudanças de tom emocional e expressões de empatia - pudessem ser classificadas com precisão, era indispensável um corpus linguístico com contexto suficiente. Interações excessivamente curtas não forneceriam dados bastantes para que o modelo de IA pudesse identificar os padrões de persistência e atrito que compõem o framework proposto.

4.3.2 Variáveis e escalas

As variáveis analisadas foram definidas com base no *framework* teórico do artefato. Para possibilitar análises estatísticas, cada variável será mensurada por meio de escalas apropriadas e, quando necessário, por indicadores compostos abaixo.

D1 - desistência conversacional (variável dependente):

- a) **aceitação de recomendação de atendimento humano pelo modelo de IA conversacional:** indica a incapacidade do modelo de IA em resolver a necessidade do usuário e o próprio agente sugerir um suporte humano e o usuário aceitar (escala discreta, número absoluto de aceitação);
- b) **solicitações de atendimento humano:** indica a incapacidade do modelo de IA em resolver a necessidade do usuário (escala discreta, número absoluto de solicitações);
- c) **abandono da conversa:** se houve abandono pelo usuário durante a interação com o agente conversacional (escala binária: presença ou ausência).

D2 - performance do agente conversacional (variável independente):

- a) **taxa de alucinações:** mede o percentual de respostas irrelevantes ou incorretas geradas pelo modelo de IA (escala percentual);
- b) **taxa de reformulação de perguntas:** indica a frequência com que o usuário precisa reformular sua entrada para obter uma resposta útil (escala percentual);
- c) **tempo de resposta do chatbot:** mede a agilidade do chatbot em processar e retornar uma resposta (escala contínua, medida em segundos).

D3 - confiança e clareza do agente conversacional (variável independente):

- a) **taxa de reformulação de perguntas:** indica problemas de clareza na resposta do modelo de IA (escala percentual);

- b) **expressões de frustração ou dúvida:** identifica sinais de insatisfação com base na análise de sentimentos (escala ordinal);
- c) **mudança de tom emocional:** avalia o grau de flexibilidade do modelo de IA em ajustar sua expressividade e estilo em resposta às emoções e necessidades demonstradas pelo usuário (escala ordinal);
- d) **clareza e coerência das respostas:** avalia se as respostas são compreensíveis e contextualmente relevantes (escala ordinal).

D4 - engajamento e empatia do agente conversacional (variável independente):

- a) **taxa de expressões empáticas:** mede a presença de linguagem empática nas respostas do modelo de IA (escala percentual);
- b) **engajamento do usuário (tempo de conversa):** tempo médio que o usuário permanece na interação (escala contínua, medida em segundos);
- c) **permanência do usuário na interação:** quantidade de mensagens trocadas em uma sessão (escala discreta, número absoluto);
- d) **tom da conversa:** classifica o impacto emocional das respostas do modelo de IA na percepção do usuário (escala ordinal).

4.4 PROCESSO DE ORGANIZAÇÃO DA BASE DE DADOS

Conforme ilustrado na Figura 4, os dados são extraídos da plataforma de mercado e organizados em uma tabela, como apresentado na Figura 5. Essa tabela envolve um trabalho manual de extração individual de cada conversa, com sua formatação respeitando as colunas, “IA” (identificação da inteligência artificial analisada), “conversa” (texto completo da conversa, conforme a Figura 4) e “quantidade interacoes” (total de interações entre usuário e *chatbot* em cada diálogo). Os nomes das colunas respeitam a formatação para o *upload* no artefato de *software* desenvolvido para análise dos dados.

Figura 5 - Recorte da tabela de dados

IA	conversa	quantidade interacoes
Tally	Conversa Seleccionada	16
Tally	Conversa Seleccionada	48
Tally	Conversa Seleccionada	70
Tally	Conversa Seleccionada	20
Tally	Conversa Seleccionada	18
Tally	Conversa Seleccionada	62
Tally	Conversa Seleccionada	62
Tally	Conversa Seleccionada	40

Fonte: Elaborado pelo autor (2026).

Após a formatação dos dados, faz-se o *download* da planilha em formato de tabulação (.tsv) e realiza-se o *upload* no *software* (artefato) desenvolvido pelo autor em uma plataforma de programação. A Figura 6 ilustra uma seção do *software*, no qual é possível tanto realizar o *upload* das conversas extraídas quanto cadastrar e editar manualmente cada conversa individualmente.

Figura 6 - Recorte do software (artefato de informação)

Registros																	Mestrado Artefato							
Prompt																	Quantidade: 159, Seleccionados: 0							
Tally																	Subir tabela		Recalcular		Baixar		+ nova conversa	
IA	Quem	RAH	SAH	AC	EPD	MTE	TA	TRP	TRIA	RSNAH	CCR	TEE	EU	PUI	TC	Ação								
<input type="checkbox"/>	Tally	Usuário 1 30/09/24 07:00	0	0	1	7 Média	6 Leve	4%	26%	39s	1	7 Moderada	19%	5765s	49 / 49 interações	4 Neutro	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>					
<input type="checkbox"/>	Tally	Usuário 2 26/09/24 15:16	0	0	0	3 Baixa	6 Leve	26%	21%	2s	1	4 Moderada	26%	1440s	19 / 19 interações	3 Positivo	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>					
<input type="checkbox"/>	Tally	Usuário 3 30/09/24 12:16	1	1	0	5 Média	6 Leve	38%	23%	0s	0	4 Moderada	23%	330s	13 / 13 interações	4 Neutro	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>					
<input type="checkbox"/>	Tally	Usuário 4 26/09/24 10:53	1	0	0	8 Alta	8 Forte	38%	15%	5s	0	3 Clara	5%	10260s	13 / 13 interações	4 Neutro	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>					
<input type="checkbox"/>	Tally	Usuário 5 29/09/24 10:38	0	0	0	6 Média	6 Leve	20%	14%	4s	1	4 Moderada	23%	1049s	35 / 35 interações	3 Positivo	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>					

Fonte: Elaborado pelo autor (2026).

Outra funcionalidade implementada no *software* foi o processamento em massa pelo modelo de IA, conforme demonstrado na Figura 7. A Figura 7 apresenta o processo no qual seleciona-se um conjunto de conversas e aciona o botão "Recalcular", permitindo escolher uma das variáveis de análise da pesquisa. Após esta ação, o *software* inicia automaticamente

uma rotina de reprocessamento, aplicando o contexto (*prompt*) do modelo de IA orientado para a variável selecionada em todas as conversas previamente escolhidas.

Figura 7 - Recorte do *software*, visão de seleção de variável para ação em massa

The screenshot displays a software interface for managing data records. At the top, there's a 'Registros' tab and a 'Prompt' label. On the right, it says 'Mestrado Artefato' and 'Quantidade: 159, Selecionados: 4'. Below this, there are input fields for 'data inicial' and 'data final', and buttons for 'Subir tabela', 'Recalcular', 'Baixar', and '+ nova conversa'. The main part of the interface is a table with columns: IA, Quem, RAH, SAH, AC, EFD, MTE, TA, TRP, TRIA, RSNAH, CCR, TEE, and Ação. The table contains five rows of data for different users. A dropdown menu is open over the 'Ação' column, showing a list of options to be selected for mass action.

IA	Quem	RAH	SAH	AC	EFD	MTE	TA	TRP	TRIA	RSNAH	CCR	TEE	Ação
<input checked="" type="checkbox"/>	Tally Usuário 1 30/09/24 07:00	0	0	1	7 Média	6 Leve	4%	26%	39 s	1	7 Moderada	19%	
<input checked="" type="checkbox"/>	Tally Usuário 2 26/09/24 15:16	0	0	0	3 Baixa	6 Leve	26%	21%	2 s	1	4 Moderada	26%	
<input checked="" type="checkbox"/>	Tally Usuário 3 30/09/24 12:16	1	1	0	5 Média	6 Leve	38%	23%	0 s	0	4 Moderada	23%	
<input checked="" type="checkbox"/>	Tally Usuário 4 26/09/24 10:53	1	0	0	8 Alta	8 Forte	38%	15%	5 s	0	3 Clara	5%	
<input type="checkbox"/>	Tally Usuário 5 29/09/24 10:38	0	0	0	6 Média	6 Leve	20%	14%	4 s	1	4 Moderada	23%	

The dropdown menu for 'Ação' contains the following options:

- Prompt completo
- Recomendação de atendimento humano (H1 e H4)
- Solicitações de atendimento humano (H1 e H4)
- Abandono da conversa (H1)
- Expressões de frustração ou dúvida (H1)
- Mudança de tom emocional (H1)
- Taxa de alucinações (H2)
- Taxa de reformulação de perguntas (H2 e H3)
- Tempo de resposta da IA (H2)
- Resolução sem necessidade de atendimento humano (H3)
- Clareza e coerência das respostas (H3)
- Taxa de expressões empáticas (H4)
- Engajamento do usuário (H4)
- Permanência do usuário na interação (H4)

Fonte: Elaborado pelo autor (2026).

Já a Figura 8 ilustra a metodologia de orientação para o modelo de IA utilizado (ChatGPT), com um contexto (*prompt*) para cada variável de análise. Um dos aspectos críticos desta pesquisa consiste na capacidade analítica de identificar alucinações ou não acuracidade pelo modelo de IA. Diante desse achado, adotou-se como estratégia de mitigação, com análise de variável por variável de pesquisa nas conversas selecionadas, em vez da abordagem conjunta (aplicar o tratamento de todas variáveis em um único *prompt*), visando maior precisão nos dados analíticos. Por fim, também há o aspecto do “termômetro” do modelo de IA do ChatGPT-4o mini, que varia de 0 a 2, em que um valor 2 possibilita o maior nível de criatividade da inteligência artificial. E para esta pesquisa, adotou-se 0,3 como parâmetro, pois entende-se que o modelo de IA não necessita de um alto nível de criatividade, mas sim precisão no tratamento dos dados.

Figura 8 - Recorte do *software*, visão de orientação do modelo (artefato de informação)

Registros **Prompt** **Mestrado Artefato**
Quantidade: 300, Seleccionados: 300

Prompt completo

Termômetro da IA
30%

Você é um analista de interações de chatbot e deve avaliar a efetividade da IA em uma conversa fornecida, preenchendo as seguintes métricas com base na transcrição. A saída deve ser **uma única linha com os valores, separados por ponto e vírgula, sem espaços**.

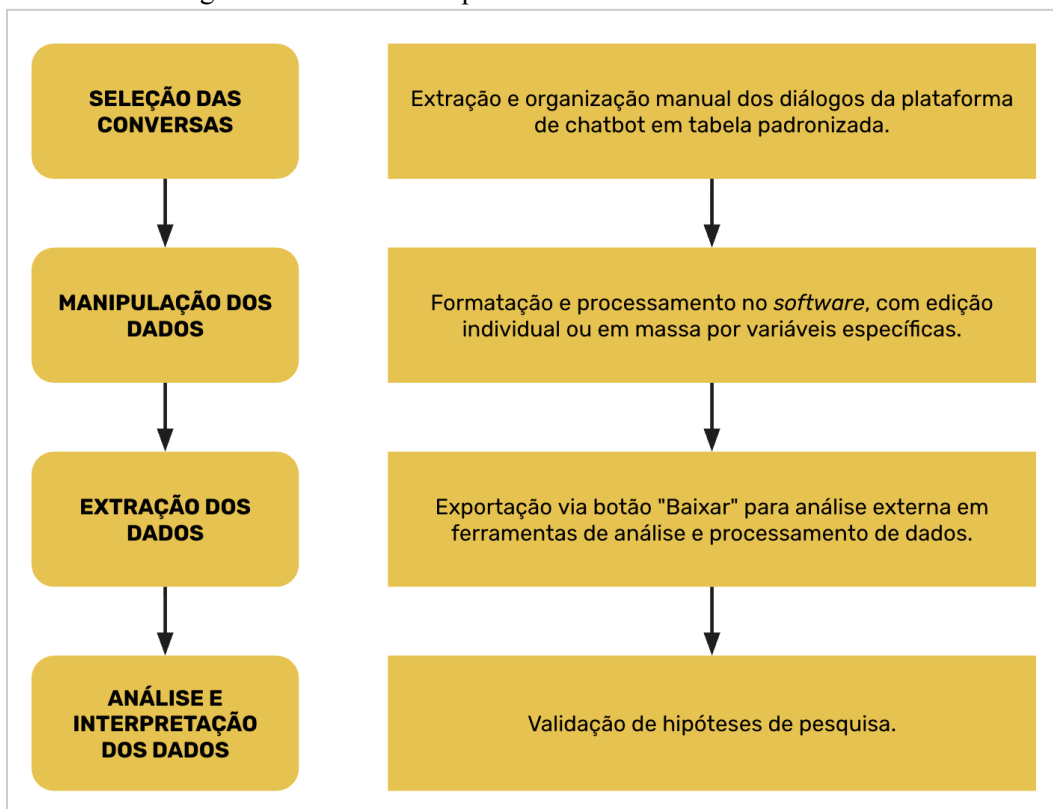
Para cada variável/métrica, aplique os seguintes critérios:

0. Primeiro nome do usuário e horário. Ex.: Walter - 25/01/25 10:23.
1. Aceitação de recomendação de atendimento humano pela IA
- Marque quantas vezes o usuário aceitou a sugestão da IA para receber atendimento humano.
2. Solicitações de Atendimento Humano:
- Conte quantas vezes o usuário solicitou explicitamente atendimento humano antes que a IA mencionasse ou sugerisse essa opção;
- Se o usuário apenas aceitar a oferta da IA, não conte como uma solicitação proativa.
3. Abandono da Conversa:
- Marcar 1 (abandono) se o usuário não enviar nenhuma mensagem em até 1 hora desde a última resposta da IA ou se a IA fornecer uma resposta relevante, e o usuário parar de responder após isso;
- Caso contrário, marcar 0 (não abandono).
4. Expressões de Frustração ou Dúvida:
- Baixa (0-3): Usuário faz perguntas neutras ou sem sinais de impaciência;

Fonte: Elaborado pelo autor (2026).

Por fim, conforme demonstrado na Figura 6, o *software* possui um botão denominado 'Baixar' que permite exportar todas as conversas selecionadas e processadas pelo modelo de IA, já no formato para análise de dados. A Figura 9 apresenta o processo completo para entendimento total da metodologia aplicada. Esta metodologia otimiza significativamente o processo até então manual para avaliação das dimensões e da eficácia das interações humano-IA, alinhando-se com os parâmetros estabelecidos no *framework* teórico da pesquisa.

Figura 9 - Processo completo do fluxo de tratamento de dados



Fonte: Elaborado pelo autor (2026).

4.5 PROCESSO DE ANÁLISE DE DADOS

O Quadro 7 apresenta as variáveis utilizadas na pesquisa, organizadas por dimensão, detalhando sua escala de mensuração, método de avaliação e possíveis classificações com seus respectivos pesos. Essa estrutura permite uma abordagem quantitativa, facilitando a modelagem estatística e a análise preditiva.

Quadro 7 - Variáveis e indicadores de medição

(continua)

Dimensão	Variáveis	Escala	Método de mensuração	Possibilidades	Ferramenta de mensuração
D1 - aceitação do agente conversacional	Aceitação de recomendação de atendimento humano pelo modelo de IA conversacional	Discreta	Número absoluto de recomendações	Número absoluto (valores inteiros)	Registro de solicitações no log
	Solicitações de atendimento humano	Discreta	Número absoluto de solicitações	Número absoluto (valores inteiros)	Registro de solicitações no log
	Abandono da conversa	Binária	Presença ou ausência de abandono	Abandono (1) ou não abandono (0)	Análise de abandono pelos logs (GPT-4o mini)

(conclusão)

Dimensão	Variáveis	Escala	Método de mensuração	Possibilidades	Ferramenta de mensuração
D2 - performance do agente conversacional	Taxa de alucinações	Percentual	Presença ou ausência de respostas irrelevantes	0 a 100%	Cálculo percentual sobre alucinações e mensagens adequadas
	Taxa de reformulação de perguntas	Percentual	Frequência de reformulação sobre total de interações	0 a 100%	Cálculo percentual sobre reformulações
	Tempo de resposta do chatbot	Contínua	Tempo médio (segundos) entre mensagem e resposta	Valores contínuos em segundos	Medição de tempo de resposta no log
D3 - confiança e clareza do agente conversacional	Taxa de reformulação de perguntas	Percentual	Frequência de reformulação sobre total de interações	0 a 100%	Cálculo percentual sobre reformulações
	Expressões de frustração ou dúvida	Ordinal	Análise de sentimentos	Baixa (0-3), Média (4-7), Alta (8-10)	Análise de sentimentos (GPT-4o mini)
	Mudança de tom emocional	Ordinal	Análise de mudança pelo agente	Adaptável (0-3), Mínima (4-6), Rígida (7-10)	Análise de mudança emocional (GPT-4o mini)
	Clareza e coerência das respostas	Ordinal	Análise qualitativa das respostas	Clara (0-3), Moderada (4-7), Confusa (8-10)	Análise qualitativa das respostas (GPT-4o mini)
D4 - engajamento e empatia do agente conversacional	Taxa de expressões empáticas	Percentual	Frequência de expressões empáticas sobre total de respostas	0 a 100%	Cálculo percentual sobre total de respostas
	Engajamento do usuário (tempo de conversa)	Contínua	Tempo médio da interação (segundos)	Valores contínuos em segundos	Registro de duração da interação no log
	Permanência do usuário na interação	Discreta	Número total de mensagens na conversa	Número absoluto (valores inteiros)	Contagem de mensagens por sessão
	Tom da conversa	Ordinal	Classificação de emoção da interação	Positivo (0-3), Neutro (4-7), Negativo (8-10)	Análise de emoções (GPT-4o mini)

Fonte: Elaborado pelo autor (2026).

4.5.1 Avaliação estatística e peso das variáveis

Após os resultados trazidos pelo modelo de IA orientado, o *framework* teórico avaliado através de um software estatístico, seguindo o seguinte escopo de análise:

- Limpeza de dados, retirando *outliers* que afetam a qualidade da análise;
- Interpolação Linear das variáveis, padronizando todas variáveis das dimensões predictoras em uma escala de 0 a 10;
- Aplicação da análise fatorial confirmatória, com foco em validar se as variáveis de cada construto são de fato pertencentes ao mesmo;

- d) Aplicação do modelo de equações estruturais, com foco em validar as hipóteses do modelo teórico.

4.5.2 Defesa do processo de análise

A metodologia proposta combina a análise de dados estruturados de conversação, modelos de inteligência artificial e estatística para validar o *framework* teórico. A análise quantitativa baseada em logs de conversação permite uma avaliação precisa da satisfação do usuário e da performance do modelo de IA, sem depender de auto-relatos subjetivos.

Os resultados esperados incluem:

- a) a validação das dimensões propostas e das hipóteses, relacionando métricas objetivas com a experiência do usuário;
- b) a identificação de padrões que diferenciam conversas bem-sucedidas de interações problemáticas;
- c) recomendações para aprimorar a clareza, empatia e confiabilidade do *chatbot*, para assim, melhorar a experiência do usuário;

5 RESULTADOS E IMPLICAÇÕES

Este capítulo apresenta a análise dos dados empíricos coletados a partir de 849 interações reais entre usuários e agentes conversacionais. O objetivo central desta etapa é validar o *framework* teórico proposto no estudo, o qual foi construído a partir das lacunas e construtos identificados na revisão sistemática da literatura. A análise busca testar as dimensões hipotetizadas - performance (D2), confiança e clareza (D3) e engajamento e empatia (D4) - como determinantes efetivos da desistência conversacional (D1). Para facilitar os cálculos e apresentação dos dados, todas variáveis estatísticas foram nomeadas e apresentadas conforme o Quadro 8.

Quadro 8 - Variáveis e indicadores de medição

Dimensão	Variáveis	Nova padronização das variáveis
D1 - aceitação do agente conversacional	Quando ocorre qualquer uma das três variáveis condicionais a desistência: aceitar ou solicitar o atendimento humano ou abandonar a conversa. Adotou-se isso para validar condicionalmente a desistência	d1_desistencia
D2 - performance do agente conversacional	Taxa de alucinações	d2_alucinacoes
	Taxa de alucinações (padronizado)	d2_alucinacoes_nm
	Taxa de reformulação de perguntas	d2_d3_reformulacao_perguntas
	Taxa de reformulação de perguntas (padronizado)	d2_d3_reformulacao_perguntas_nm
	Tempo de resposta do chatbot	d2_tempo_resposta
	Tempo de resposta do chatbot (padronizado)	d2_tempo_resposta_nm
D3 - confiança e clareza do agente conversacional	Taxa de reformulação de perguntas	d2_d3_reformulacao_perguntas
	Taxa de reformulação de perguntas (padronizado)	d2_d3_reformulacao_perguntas_nm
	Expressões de frustração ou dúvida	d3_frustracao_duvida
	Mudança de tom emocional	d3_mudanca_tom_emocional
	Clareza e coerência das respostas	d3_clareza_coerencia
D4 - engajamento e empatia do agente conversacional	Taxa de expressões empáticas	d4_expressoes_empaticas
	Taxa de expressões empáticas (padronizado)	d4_expressoes_empaticas_nm
	Engajamento do usuário (tempo de conversa)	d4_engajamento_usuario
	Engajamento do usuário (padronizado)	d4_engajamento_usuario_nm
	Permanência do usuário na interação	d4_permanencia_interacao
	Permanência do usuário na interação (padronizado)	d4_permanencia_interacao_nm
	Tom da conversa	d4_tom_conversa

Fonte: Elaborado pelo autor (2026).

5.1 TRATAMENTO E INTERPOLAÇÃO LINEAR DOS DADOS

Devido à natureza heterogênea das métricas de interação (contagens de tempo, frequência de alucinações e taxas de reformulação), os dados brutos apresentaram distribuições com alta variabilidade e presença de outliers severos (Figura 10). Para garantir a integridade das análises multivariadas e a convergência dos modelos estruturais, adotou-se um protocolo de tratamento em duas etapas:

- d) **tratamento de outliers (corte via percentil):** variáveis baseadas em contagem sem limite superior (como d4_engajamento_usuario e d2_tempo_resposta) foram tratadas utilizando a técnica de corte no 95º percentil. Valores acima deste limite foram excluídos, mitigando o efeito desproporcional de interações atípicas;
- e) **interpolação Linear de escalas (0 a 10):** após o tratamento dos extremos, todas as variáveis foram padronizadas para uma escala intervalar de 0 a 10. Nesta escala, adotou-se a convenção de que valores maiores representam maior atrito ou menor qualidade (ex: 10 indica alta taxa de alucinação ou resposta excessivamente lenta), uniformizando a interpretação dos coeficientes nos modelos estruturais e de regressão subsequentes.

Tabela 1 - Estatística descritiva e parâmetros de limitação de outliers

	agente	d2_alucinacoes	d2_d3_reformulacao_perguntas	d2_tempo_resposta	d4_expressoes_empaticas	d4_engajamento_usuario	d4_permanencia_interacao
N	Abby	150	150	150	150	150	150
	IA Beto	331	330	331	331	331	331
	Tally	368	368	368	368	368	368
Mínimo	Abby	0	0	0	0	16	10
	IA Beto	0	0	0	0	57	8
	Tally	0	0	0	0	45	10
Máximo	Abby	100	50	83	75	16859	72
	IA Beto	40	57	1230	100	149656	76
	Tally	73	75	188	88	257604	137
95º percentil	Abby	25.0	33.0	14.5	75.0	5699	26.5
	IA Beto	27.0	38.0	22.0	74.0	11595	34.0
	Tally	33.0	39.6	53.2	64.0	15747	56.6

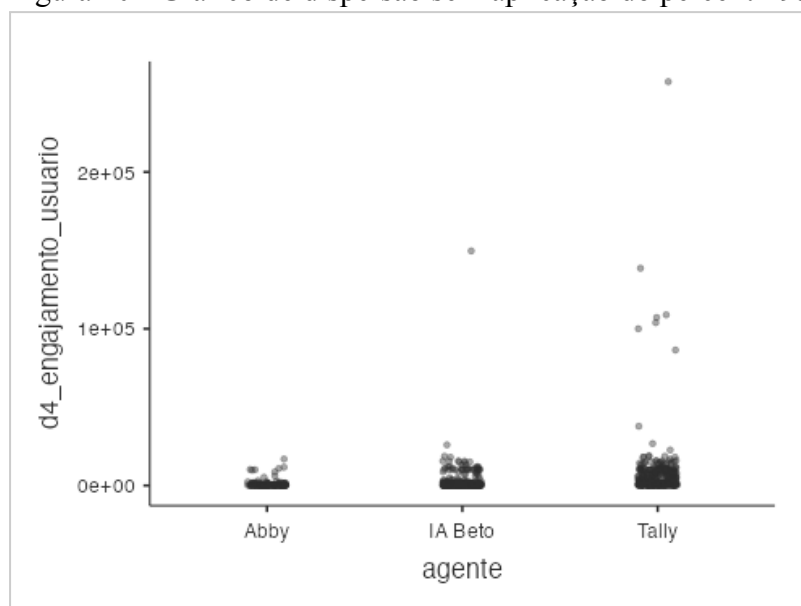
Fonte: Elaborado pelo autor (2026).

A Tabela 1 apresenta o diagnóstico da variabilidade dos dados brutos, revelando um cenário de extrema dispersão que exigiu um tratamento estatístico antes da modelagem. Ao

observar a variável `d4_engajamento_usuario`, nota-se que, para o agente Tally, o valor máximo registrado atinge a marca de 257604, um valor que se distancia muito do 95º percentil, fixado em apenas 15747. Essa discrepância sugere a existência de interações atípicas - possivelmente sessões de teste ou falhas de encerramento de *log* - que não representam o comportamento do usuário padrão. Esse fenômeno de "cauda longa" repete-se em outras métricas críticas de interação. Os valores da variável `d4_expressoes_empaticas` chegam a apresentar picos de 100 ocorrências em uma única sessão (IA Beto), enquanto a média e o balizamento pelo percentil 95 sugerem que o comportamento típico não ultrapassa 74. Da mesma forma, a `d4_permanencia_interacao` registra casos extremos de até 137 mensagens trocadas (Tally), contrastando com o limite de 56,6 interações que abrange 95% da amostra.

Essa realidade é visualmente ratificada pela Figura 10, onde a densidade das observações se concentra na base do eixo vertical, enquanto pontos isolados "flutuam" em patamares excessivamente elevados. Sem o balizamento pelo percentil 95, esses *outliers* exerceriam um peso desproporcional, "puxando" as médias e enviesando os coeficientes de regressão do modelo final. Portanto, o truncamento dos dados e a subsequente interpolação Linear para a escala de 0 a 10 não foram apenas escolhas metodológicas, mas necessidades técnicas para garantir que o modelo reflita a jornada real da vasta maioria dos usuários, e não a volatilidade de casos excepcionais.

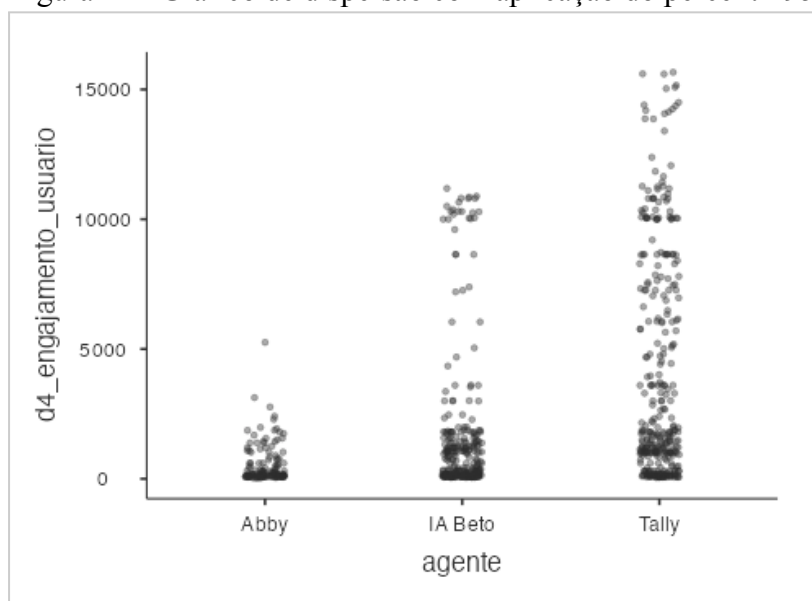
Figura 10 - Gráfico de dispersão sem aplicação do percentil 95º



Fonte: Elaborado pelo autor (2026).

Em contraste, a Figura 11 ilustra o comportamento da amostra após a aplicação do filtro pelo 95º percentil. Ao remover os valores extremos que distorciam a escala, observa-se uma redução significativa na dispersão visual da amostra, permitindo uma visualização clara do comportamento típico dos usuários para cada agente. Esta interpolação Linear visual reflete a estabilização estatística dos dados, garantindo que as análises de correlação e os caminhos do modelo estrutural não sejam enviesados por ruídos de interações atípicas.

Figura 11 - Gráfico de dispersão com aplicação do percentil 95º



Fonte: Elaborado pelo autor (2026).

Após o tratamento de *outliers* via percentil 95, procedeu-se à interpolação Linear das variáveis para uma escala intervalar de 0 a 10. Como o objetivo central do modelo é prever a desistência (onde valores maiores indicam um desfecho negativo), tornou-se necessário padronizar a polaridade interpretativa de todos os indicadores. Enquanto as variáveis dos construtos D2 e D3 já possuíam uma relação direta com a desistência, as métricas de engajamento e empatia do usuário (D4) apresentavam uma correlação inversa por natureza.

Para alinhar esses indicadores, aplicou-se a regra de inversão escalar $(10 - x)$ para as variáveis `d4_expressoes_empaticas_nm`, `d4_engajamento_usuario_nm` e `d4_permanencia_interacao_nm`. Com esta transformação, as novas métricas padronizadas passam a representar o atrito de interação: quanto maior o valor (próximo a 10), menor foi o engajamento ou a empatia detectada, alinhando-se logicamente ao aumento da probabilidade de desistência.

A Tabela 2 detalha os novos parâmetros, onde se observa que agora compartilham a mesma métrica de 0 a 10, permitindo que o modelo capture com precisão o impacto consolidado da baixa qualidade da conversa sobre a aceitação do usuário.

Tabela 2 - Estatística descritiva com variáveis padronizadas e sem *outliers*

	agente	d2_alucina coes_nm	d2_d3_reformulacao _perguntas_nm	d2_tempo_ resposta_nm	d4_expressoes_ empaticas_nm	d4_engajamento _usuario_nm	d4_permanencia _interacao_nm
N	Abby	125	125	125	125	125	125
	IA Beto	253	253	253	253	253	253
	Tally	280	280	280	280	280	280
Mínimo	Abby	0.00	0.00	0.00	0.00	6.66	5.36
	IA Beto	0.00	0.00	0.00	400	2.90	3.93
	Tally	0.00	0.00	0.00	1.47	552	0.00
Máximo	Abby	2.50	3.30	2.64	9.20	9.99	8.21
	IA Beto	2.70	3.60	4.15	10.0	9.96	8.57
	Tally	3.30	3.90	9.43	10.0	9.97	8.21

Fonte: Elaborado pelo autor (2026).

A aplicação do filtro baseado no 95º percentil resultou em um refinamento estratégico da base de dados, removendo as interações cujos valores de engajamento e permanência sugeriam comportamentos atípicos ou erros de sistema. A comparação entre o volume de dados inicial e a amostra final tratada demonstra o impacto dessa limpeza por agente:

- a) **agente Abby:** A amostra inicial de 150 interações foi reduzida para 125, representando uma retenção de aproximadamente 83% dos dados originais (perda de 17%);
- b) **agente IA Beto:** O volume passou de 331 para 253 interações válidas, resultando em uma retenção de 76% (perda de 24%);
- c) **agente Tally:** De um total inicial de 368, restaram 280 interações na base final, o que equivale a uma retenção de 76% (perda de 24%).

Essa redução de volume (de 849 para 658 casos totais) não indica perda de poder estatístico, mas sim um ganho qualitativo na densidade dos dados. Nota-se que a perda foi ligeiramente superior nos agentes IA Beto e Tally, justamente onde os histogramas e gráficos de dispersão indicavam uma maior incidência de *outliers* extremos no engajamento e no tempo de resposta. A exclusão desses casos permitiu que as variáveis padronizadas

apresentassem comportamentos mais estáveis, com o valor máximo agora alinhado ao comportamento real da massa de usuários.

5.2 ANÁLISE CONFIRMATÓRIA E EXPLORATÓRIA

A avaliação do modelo de mensuração teve como objetivo inicial testar a validade da estrutura teórica proposta para esta pesquisa, composta pelas dimensões de performance (D2), confiança e clareza (D3) e engajamento e empatia (D4). Para tanto, conduziu-se uma Análise Fatorial Confirmatória (AFC) utilizando o estimador DWLS (*Diagonally Weighted Least Squares*), adequado para dados ordinais e distribuições não normais Hair *et al.* (2009).

A Tabela 4 apresenta as estimativas dos pesos fatoriais para o modelo teórico original. Observa-se que todos os indicadores apresentaram cargas fatoriais altamente significativas ($p < .001$), o que sugere que as variáveis observadas possuem relevância estatística na formação dos seus respectivos construtos latentes.

Tabela 3 - Pesos fatoriais da AFC no modelo de mensuração teórico

Fator	Indicador	Estimativas	Erro-padrão	Z	p
D2	d2_alucinacoes_nm	0.514	0.0359	14.30	<.001
	d2_d3_reformulacao_perguntas_nm	0.747	0.0515	14.50	<.001
	d2_tempo_resposta_nm	0.388	0.0591	6.57	<.001
D3	d3_frustracao_duvida	1.808	0.0846	21.36	<.001
	d3_mudanca_tom_emocional	1.105	0.0848	13.03	<.001
	d3_clareza_coerencia	1.344	0.0686	19.58	<.001
D4	d4_expressoes_empaticas_nm	0.615	0.0812	7.57	<.001
	d4_engajamento_usuario_nm	-0.829	0.0928	-8.94	<.001
	d4_permanencia_interacao_nm	-0.512	0.0559	-9.15	<.001
	d4_tom_conversa	1.663	0.0813	20.45	<.001

Fonte: Elaborado pelo autor (2026).

Apesar da significância individual dos itens, a avaliação global do modelo exige a análise dos índices de ajustamento e da validade discriminante entre os construtos. As Tabelas 4 e 5 detalham as medidas de adequação da estrutura proposta aos dados empíricos.

Tabela 4 - Medidas de ajustamento exato do modelo teórico

χ^2	gl	p
382	32	<.001

Fonte: Elaborado pelo autor (2026).

Tabela 5 - Medidas de ajustamento global do modelo teórico

			IC 90% RMSEA	
CFI	TLI	RMSEA	Lim. Inferior	Superior
0.820	0.747	0.129	0.117	0.141

Fonte: Elaborado pelo autor (2026).

Conforme evidenciado na Tabela 5, os índices de ajustamento global, como o *Comparative Fit Index* (CFI = 0.820) e o *Tucker-Lewis Index* (TLI = 0.747), situaram-se abaixo do limiar de 0.90 frequentemente adotado na literatura tradicional. Contudo, Hair *et al.* (2009) advertem que não existe um "valor mágico" isolado que diferencie bons modelos de modelos ruins para qualquer índice de ajuste, visto que tais medidas sofrem influência da complexidade do modelo e do tamanho da amostra. Sendo assim, o ajuste marginal dos índices não constitui, por si só, o único fator para a rejeição do modelo. A falha estrutural definitiva do modelo teórico revela-se na análise das covariâncias fatoriais, detalhadas na Tabela 6.

Tabela 6 - Covariâncias fatoriais do modelo teórico

		Estimativas	Erro-padrão	Z	p
D2	D2	1.000 ^a			
	D3	1.163	0.0414	28.1	<.001
	D4	0.955	0.0500	19.1	<.001
D3	D3	1.000 ^a			
	D4	1.065	0.0314	34.0	<.001
D4	D4	1.000 ^a			

Fonte: Elaborado pelo autor (2026).

A Tabela 6 expõe uma severa violação dos pressupostos estatísticos. Observam-se covariâncias padronizadas superiores a 1.0 entre os construtos D2 e D3 (1.163) e entre D3 e D4 (1.065). Na modelagem de equações estruturais, anomalias matemáticas como estimativas de correlação padronizadas acima de 1.0 (parâmetros ilógicos) ou variâncias de erro negativas (*Heywood cases*) indicam graves problemas de estimação, frequentemente causados por um grau extremo de multicolinearidade entre os indicadores (Hair *et al.*, 2009). Tais resultados

comprovam a ausência de validade discriminante na estrutura teórica original no *framework* teórico. Sob a perspectiva do usuário, as dimensões de performance, confiança e empatia não operam como pilares isolados e independentes, mas fundem-se em percepções cognitivas conjuntas durante a interação com o agente conversacional.

Diante da constatação empírica de que o modelo teórico estrito apresenta severa multicolinearidade e não se sustenta matematicamente, adotou-se uma abordagem exploratória para identificar o melhor agrupamento dos indicadores (variáveis de pesquisa) nas dimensões. O objetivo foi purificar a escala e investigar a verdadeira estrutura latente subjacente aos dados, permitindo que o modelo refletisse com maior precisão a experiência cognitiva e comportamental relatada pelos usuários. Para tanto, conduziu-se uma Análise Fatorial Exploratória (AFE), empregando o método de extração de resíduos mínimos combinado com uma rotação oblíqua (*oblimin*), adequada quando se pressupõe correlação entre os fatores. A Tabela 7 apresenta as cargas fatoriais resultantes desta etapa (Hair *et al.*, 2009).

Tabela 7 - Pesos fatoriais e singularidade da AFE

	Fator				Singularidade
	1	2	3	4	
d2_alucinacoes_nm	0.758				0.46341
d3_clareza_coerencia	0.705				0.33981
d4_tom_conversa	0.495				0.44728
d3_mudanca_tom_emocional	0.446			0.378	0.58749
d3_frustracao_duvida		0.982			0.00467
d2_d3_reformulacao_perguntas_nm		0.480			0.61048
d4_permanencia_interacao_nm		-0.423			0.71754
d4_engajamento_usuario_nm			0.989		0.03792
d2_tempo_resposta_nm			-0.340		0.78068
d4_expressoes_empaticas_nm				0.649	0.56959

Fonte: Elaborado pelo autor (2026).

A AFE revelou um padrão de agrupamento empírico que não invalida, mas antes refina, a divisão teórica original. Evidenciou-se que, na percepção real do utilizador, as variáveis alinham-se em dimensões que mesclam aspectos técnicos e reações humanas, mantendo-se, contudo, fortemente ancoradas nas bases teóricas propostas. Seguindo a premissa de que os construtos latentes necessitam de múltiplos indicadores para manterem a

consistência interna e a validade discriminante, as variáveis foram consolidadas em três dimensões fundamentais, realinhando os aspectos originais (D2, D3 e D4) da seguinte forma:

- a) **qualidade conversacional (D2):** a análise agrupou as variáveis `d2_alucinacoes_nm`, `d3_clareza_coerencia`, `d4_tom_conversa` e `d3_mudanca_tom_emocional`. Este agrupamento é justificado pela *Cognitive Fit Theory* (Vessey, 1991), que rege este construto. Segundo a teoria, o usuário avalia a performance da IA com base na adequação cognitiva da resposta. Portanto, na prática, a "performance" não é medida apenas pela ocorrência de erros técnicos (alucinações), mas pela capacidade de a IA entregar informações de forma clara, coerente e com um tom de conversa estável. Por isso a mudança de "performance" para "qualidade conversacional" na definição de D2. Para o usuário, falhas na clareza ou oscilações emocionais no tom do agente são interpretadas primariamente como uma falha na qualidade conversacional do agente;
- b) **confiança e clareza (D3):** esta dimensão consolidou as variáveis `d3_frustracao_duvida`, `d2_d3_reformulacao_perguntas_nm` e `d4_permanencia_interacao_nm`. A união destes indicadores corrobora a *Social Response Theory* (Adam; Wessel; Benlian, 2022). A teoria estipula que o envolvimento humano e a confiança (presença social) são pilares da interação. Quando o *chatbot* falha em transmitir essa confiança, o usuário experimenta frustração e é forçado a um esforço mecânico (reformular perguntas repetidamente). Esse desgaste rompe a percepção de presença social e quebra a confiança, o que se reflete matematicamente no sinal negativo da carga fatorial da permanência, demonstrando que falhas do agente conversacional tendem a prolongar a interação de forma ineficiente, em vez de causarem um abandono imediato;
- c) **engajamento e empatia (D4):** o construto reuniu as variáveis `d4_engajamento_usuario_nm`, `d2_tempo_resposta_nm` e `d4_expressoes_empaticas_nm`. O agrupamento empírico reforça a *Communication Accommodation Theory* (Zhang; Giles, 2016). O engajamento do usuário é uma resposta direta à capacidade de o *chatbot* adaptar (acomodar) a sua comunicação. A análise comprovou que essa "acomodação" ocorre por meio de dois vetores conjuntos: o uso de linguagem humanizada (expressões empáticas) e o ritmo da conversa (tempo de resposta). O tempo de resposta, ao

invés de ser visto apenas como performance, atua para o usuário como um simulador do fluxo natural e empático de uma interação humana, sendo fundamental para o engajamento geral. (Nota: Evitou-se repetir a variável *d3_mudanca_tom_emocional* neste construto devido ao fenômeno estatístico de *cross-loading*, sendo a mesma alocada exclusivamente em D2 por apresentar maior peso fatorial).

Para validar a nova taxonomia (classificação) orientada pelos dados e extraída da AFE, e suportada pela teoria, procedeu-se a uma nova AFC. O objetivo desta etapa foi verificar se a reestruturação das variáveis em três novas dimensões (D2, D3 e D4) seria capaz de corrigir as anomalias de estimação e melhorar o ajustamento global do modelo. A Tabela 8 apresenta as estimativas dos pesos fatoriais para o modelo de mensuração refinado.

Tabela 8 - Pesos fatoriais da AFC (modelo refinado)

Fator	Indicador	Estimativas	Erro-padrão	Z	p
D2	d2_alucinacoes_nm	0.582	0.0337	17.25	<.001
	d3_clareza_coerencia	1.491	0.0681	21.89	<.001
	d4_tom_conversa	1.675	0.0771	21.73	<.001
	d3_mudanca_tom_emocional	1.189	0.0876	13.58	<.001
D3	d3_frustracao_duvida	2.213	0.0897	24.68	<.001
	d2_d3_reformulacao_perguntas_nm	0.829	0.0487	17.03	<.001
	d4_permanencia_interacao_nm	-0.581	0.0549	-10.58	<.001
D4	d4_engajamento_usuario_nm	1.441	0.1120	12.87	<.001
	d4_expressoes_empaticas_nm	-0.708	0.0977	-7.25	<.001
	d2_tempo_resposta_nm	-0.857	0.0708	-12.09	<.001

Fonte: Elaborado pelo autor (2026).

Os resultados da Tabela 8 indicam que todas as variáveis mantiveram alta significância na formação de seus respectivos construtos ($p < 0,001$). Nota-se que, mesmo com o alinhamento lógico das métricas para representarem o atrito (onde valores próximos a 10 indicam pior desempenho), os indicadores de permanência, empatia e tempo de resposta mantiveram cargas fatoriais negativas. Essa ocorrência não é um erro de processamento, mas uma validação estatística da natureza dessas variáveis: elas operam como forças de oposição ao desgaste do usuário. Enquanto indicadores como alucinações e frustração carregam o modelo positivamente em direção à desistência, as variáveis de D4 atuam como mitigadores.

Com a estrutura fatorial validada, a Tabela 9 apresenta as covariâncias entre os construtos, detalhando como essas forças interagem no sistema.

Tabela 9 - Covariâncias fatoriais do modelo teórico

		Estimativas	Erro-padrão	Z	p
D2	D2	1.000 ^a			
	D3	0.760	0.0277	27.45	<.001
	D4	-0.523	0.0497	-10.52	<.001
D3	D3	1.000 ^a			
	D4	-0.493	0.0509	-9.68	<.001
D4	D4	1.000 ^a			

Fonte: Elaborado pelo autor (2026).

Ao contrário do modelo teórico original, que apresentava correlações superiores a 1.0 (*Heywood cases*), a nova estrutura resolveu integralmente o problema da multicolinearidade extrema. A maior covariância observada no modelo refinado ocorreu entre D2 e D3 (0.760), um valor estritamente inferior a 1.0, estabelecendo a validade discriminante entre as três dimensões. Isso significa que, agora, os construtos medem facetas distintas da experiência do usuário, tornando o modelo matematicamente possível e confiável (Hair *et al.*, 2009). Por fim, as Tabelas 10 e 11 detalham as medidas de ajustamento global da nova estrutura.

Tabela 10 - Medidas de ajustamento exato do modelo refinado

χ^2	gl	p
382	32	<.001

Fonte: Elaborado pelo autor (2026).

Tabela 11 - Medidas de ajustamento global do modelo refinado

			IC 90% RMSEA	
CFI	TLI	RMSEA	Lim. Inferior	Superior
0.895	0.852	0.0984	0.0868	0.110

Fonte: Elaborado pelo autor (2026).

A purificação da escala promoveu um avanço na qualidade global do ajustamento. O erro de aproximação (RMSEA) caiu de 0.129 (no modelo original) para 0.098. Simultaneamente, o *Comparative Fit Index* (CFI) elevou-se de 0.820 para 0.895. Apesar de o CFI situar-se marginalmente abaixo do limiar estrito de 0.90, Hair *et al.* (2009) argumentam que índices de ajustamento não devem ser interpretados como regras de corte absolutas, mas

sim avaliados em conjunto com a validade teórica e a ausência de anomalias de estimação. Considerando a eliminação completa dos *Heywood cases* e a melhoria substancial em todos os indicadores de ajuste comparativamente ao modelo inicial, conclui-se que o modelo de mensuração refinado possui validade discriminante adequada. Portanto, esta é a estrutura adotada para prosseguir à avaliação do modelo de equações estruturais, comparando o modelo teórico original com o novo modelo proposto (mantendo as dimensões teóricas, mas com nova distribuição das variáveis) e um modelo de segunda ordem.

5.3 ANÁLISE DO MODELO DE EQUAÇÕES ESTRUTURAIS

Nesta etapa, a pesquisa avança da validação dos constructos para o teste das relações de dependência entre as dimensões do framework. A Modelagem de Equações Estruturais (*Structural Equation Modeling* - SEM) é aplicada para verificar como as falhas de performance, a percepção de confiança e o engajamento atuam, em conjunto, como preditores da desistência do usuário.

5.3.1 Reespecificação do modelo teórico a estrutura serial

O modelo teórico original, que previa influências diretas e simultâneas das dimensões sobre a desistência, apresentou limitações estatísticas severas. Como demonstrado na Tabela 6 (covariâncias fatoriais), observou-se uma multicolinearidade crítica entre os construtos latentes, com estimativas de covariância superiores a 1,00 (ex: D2 entre D3 = 1.163 e D3 entre D4 = 1.065). Esses valores indicam uma redundância teórica e estatística (ausência de validade discriminante), sugerindo que as dimensões não operam de forma independente. Diante disso, optou-se pela reespecificação para um modelo serial (D2→D3→D4→D1). Essa abordagem resolve o problema da colinearidade ao transformar a sobreposição em uma cascata de efeitos, onde a baixa performance precede a degradação da confiança e clareza em que precede o engajamento e empatia e aumenta a desistência conversacional.

5.3.2 Definição de desistência e modelos de avaliação

Conforme mencionado no Quadro 8, a desistência conversacional foi condicionada como realizada quando ocorre qualquer uma das três condições: a aceitação de recomendação de atendimento humano, a solicitação de atendimento humano ou o abandono da conversa. E

para poder executar o SEM, o construto D1 é tratado como uma variável de indicador único (*single-item construct*). Esta definição é necessária para representar a desistência como um indicador de “aconteceu” ou “não aconteceu”, consolidando em um único ponto crítico os três sinais de interrupção da jornada.

Por fim, além do modelo teórico reespecificado para processual em vez de multidimensional, iremos analisar o modelo teórico purificado multidimensional e purificado processual, com a nova distribuição de variáveis nas dimensões e também o modelo de segunda ordem seguindo as dimensões originais. Dada a novidade da análise deste fenômeno (desistência conversacional com modelos de IA) e o método de análise (a partir de *logs* de conversação), o entendimento de diferentes configurações se torna necessário. Será feita essa comparação com modelos rivais para determinar qual deles é o mais parcimonioso e apresenta o melhor ajuste, conforme Hair *et al.* (2009). Segue o Quadro 9 para melhor compreensão dos modelos aplicados no SEM.

Quadro 9 - Modelos de avaliação no SEM

Modelo	Variáveis e dimensões	Estrutura
Modelo 1: teórico reespecificado processual	Variáveis nas dimensões conforme o modelo teórico inicial	D2→D3→D4→D1
Modelo 2: teórico purificado multidimensional	Variáveis redistribuídas conforme análise exploratória	D2+D3+D4→D1
Modelo 3: teórico purificado reespecificado processual	Variáveis redistribuídas conforme análise exploratória	D2→D3→D4→D1
Modelo 4: de segunda ordem	Variáveis nas dimensões conforme o modelo teórico	(D2,D3,D4)→Dx→D1
Modelo 5: D2 como antecedente	Variáveis redistribuídas conforme análise exploratória	D2→D3; D2→D4 D3→D1; D4→D1

Fonte: Elaborado pelo autor (2026).

Previsto inicialmente, o modelo teórico multidimensional na Figura 2 não teve seus índices de ajuste reportados. Durante a execução no SEM, o modelo não atingiu a convergência (identificação), apresentando erros na inversão da matriz de covariância. Tal fato reforça a necessidade da reespecificação do modelo ou purificação das variáveis para a viabilidade estatística do fenômeno estudado.

5.3.3 Análise dos índices comparativos no SEM

Dando continuidade à modelagem, esta seção apresenta os resultados obtidos a partir da execução dos cinco modelos estruturais estabelecidos anteriormente no Quadro 9. A

análise concentra-se na comparação de seus índices de ajuste global, visando identificar a estrutura que melhor se adequa à matriz de covariância dos dados coletados.

A Tabela 12 consolida os principais indicadores de desempenho estatístico para cada modelo. Conforme preconizado por Hair *et al.* (2009), a avaliação foi realizada de forma comparativa.

Tabela 12 - Índices de ajuste global dos modelos estruturais

Modelo	χ^2	df	χ^2/df	CFI	TLI	RMSEA	SRMR
Ponto de corte (Hair <i>et al.</i> , 2009)	-	-	-	> 0.90	> 0.90	< 0.08	< 0.08
Modelo 1: Teórico processual D2→D3→D4→D1	204	42	4.86	0.927	0.905	0.077	0.079
Modelo 2: Teórico purificado multidimensional D2+D3+D4→D1	115	39	2.95	0.966	0.952	0.055	0.059
Modelo 3: Teórico purificado processual D2→D3→D4→D1	129	42	3.08	0.961	0.949	0.056	0.064
Modelo 4: Segunda ordem não purificado (D2,D3,D4)→Dx→D1	204	42	4.86	0.927	0.905	0.077	0.079
Modelo 5: purificado e D2 como antecedente D2→D3; D2→D4 D3→D1; D4→D1	133	41	3.24	0.959	0.945	0.051	0.062

Fonte: Elaborado pelo autor (2026).

Os resultados apresentados na Tabela 12 indicam que todos os modelos testados atingiram os níveis críticos de aceitação estabelecidos pela literatura (CFI e TLI > 0.90, RMSEA e SRMR < 0.08). No entanto, observa-se uma evolução nos índices de ajuste ao comparar as estruturas originais com as versões purificadas. O modelo 1 (teórico processual) e modelo 4 (segunda ordem não purificado) apresentaram desempenho idêntico, com um CFI de 0.927 e um erro de aproximação (RMSEA) próximo ao limite tolerável (0.077). O elevado valor do Qui-quadrado padronizado ($\chi^2/df = 4.86$) sugere que essas estruturas, por mais que válidas, ainda possuíam resíduos significativos e uma complexidade que não se traduzia em aderência total aos dados empíricos. Contudo, não há diferença estatisticamente significativa entre os resultados dos modelos purificados, então ainda não é possível afirmar qual modelo explica melhor o conjunto de dados, necessitando-se avaliar o R^2 (poder explicativo do modelo sobre os dados) β (intensidade e direção do impacto de cada variável).

A introdução das etapas de purificação resultou em um salto quantitativo nos modelos 2, 3 e 5. O modelo 2 (purificado multidimensional) obteve o melhor ajuste global da

série, com o menor χ^2/df (2.95) e o maior CFI (0.966), indicando que a percepção das dimensões de forma simultânea possui forte suporte estatístico. Logo em seguida, o modelo 3 (purificado processual) demonstrou um ajuste quase equivalente (CFI = 0.961 e RMSEA = 0.056), mantendo a estrutura lógica em cascata, mas com uma precisão muito superior à versão inicial do modelo 1. E por fim, o modelo 5 demonstrou um ajuste também significativo (CFI = 0.959 e RMSEA = 0.051), com uma proposta da variável D2 (qualidade conversacional) ser preditora de D3 e D4.

Portanto, a análise dos índices de ajuste global isoladamente não permite descartar nenhuma das abordagens dos modelos, uma vez que todos descrevem a matriz de dados com alta fidelidade. Conforme orientam Hair *et al.* (2009), diante de modelos concorrentes com ajustes satisfatórios, a decisão deve transpassar a análise global e focar na força das relações internas (β) e na capacidade de explicação da variância (R^2), a fim de determinar qual configuração melhor representa o comportamento real do usuário, apresentada na seção seguinte.

5.3.4 Análise dos índices comparativos no SEM nos cinco modelos

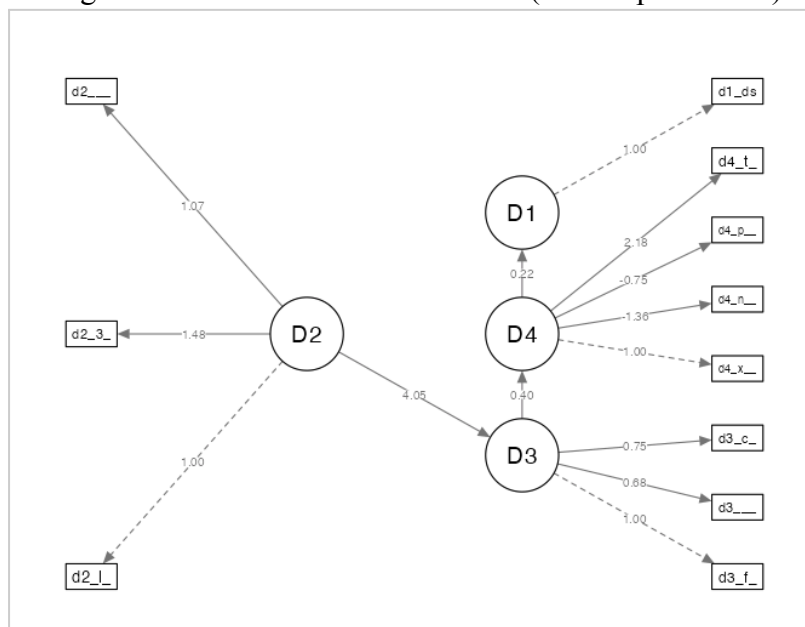
Após a validação da adequação estatística global, a análise avança para o exame da validade interna e do poder preditivo de cada configuração. Conforme orientado por Hair *et al.* (2009), a eficácia de um modelo estrutural não se limita aos índices de ajuste, exigindo a verificação da força das relações (β) e da capacidade de explicação da variância (R^2) do na última dimensão (D1). Esta etapa busca identificar qual dos modelos observados melhor representa a desistência do usuário (D1), revelando como a purificação das variáveis impactou a estabilidade matemática e a interpretação dos dados.

A Tabela 13 e Figura 12 apresentam a avaliação do modelo 1, teórico processual.

Tabela 13 - Estimativas do modelo 1 (teórico processual)

Dependente	Preditor	Estimativa (B)	Erro Padrão (SE)	Coef. Padronizado (β)	R^2	p
D3	D2	4.048	0.5919	1.071	Instável	<.001
D4	D3	0.401	0.0552	0.987	97,5%	<.001
D1	D4	0.224	0.0757	0.160	2.6%	0.003

Fonte: Elaborado pelo autor (2026).

Figura 12 - *Framework* do modelo 1 (teórico processual)

Fonte: Elaborado pelo autor (2026).

O modelo 1, mesmo apresentando um ajuste global aceitável, revela fragilidades críticas em sua estrutura interna. A análise da Tabela 13 indica a presença de um *Heywood Case* na relação entre D2 e D3, onde o coeficiente padronizado β excede a unidade, gerando uma variância residual negativa e tornando a estimativa matematicamente instável. Apesar da forte mediação observada entre D3 e D4 ($\beta = 0.987$; $R^2 = 97.5\%$), a capacidade do modelo em explicar a desistência é fraca, apresentando um R^2 de apenas 2,6% para a D1 (desistência). O *framework* da Figura 12 ilustra essa desconexão: meso com as relações iniciais sejam estatisticamente significantes ($p < 0,001$), o fluxo processual rígido falha em capturar os determinantes da desistência, evidenciando que a desistência não é uma consequência linear direta do engajamento na configuração original das variáveis compostas nas dimensões D2, D3 e D4.

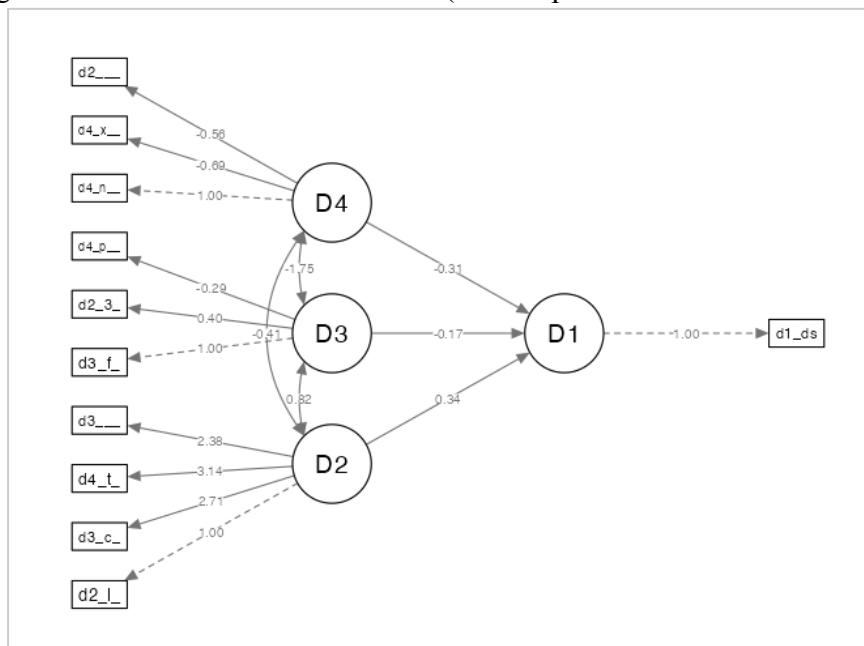
Segue a avaliação do modelo 2, teórico purificado multidimensional, conforme dados na Tabela 14 e Figura 13.

Tabela 14 - Estimativas do modelo 2 (teórico purificado multidimensional)

Dependente	Preditor	Estimativa (B)	Erro Padrão (SE)	Coef. Padronizado (β)	R^2	p
D1	D2	0.336	0.200	0.170	--	0.094
D1	D3	-0.172	0.056	-0.338	--	0.002
D1	D4	-0.313	0.082	-0.394	--	<.001
Poder total	modelo 2	--	--	--	12.6%	--

Fonte: Elaborado pelo autor (2026).

Figura 13 - *Framework* do modelo 2 (teórico purificado multidimensional)



Fonte: Elaborado pelo autor (2026).

Diferente do modelo 1, o modelo 2 apresenta uma estrutura estabilizada e com superior capacidade preditiva. Ao tratar as dimensões de forma simultânea e independente, o modelo elimina as inconsistências matemáticas e atinge um R^2 de 12.6% para D1 (desistência), um valor quase cinco vezes superior ao observado no modelo processual. A Figura 13 ilustra que D4 (engajamento e empatia) e D3 (confiança e clareza) emergem como os principais inibidores da desistência, com coeficientes de $\beta = -0,394$ e $\beta = -0,338$, respectivamente.

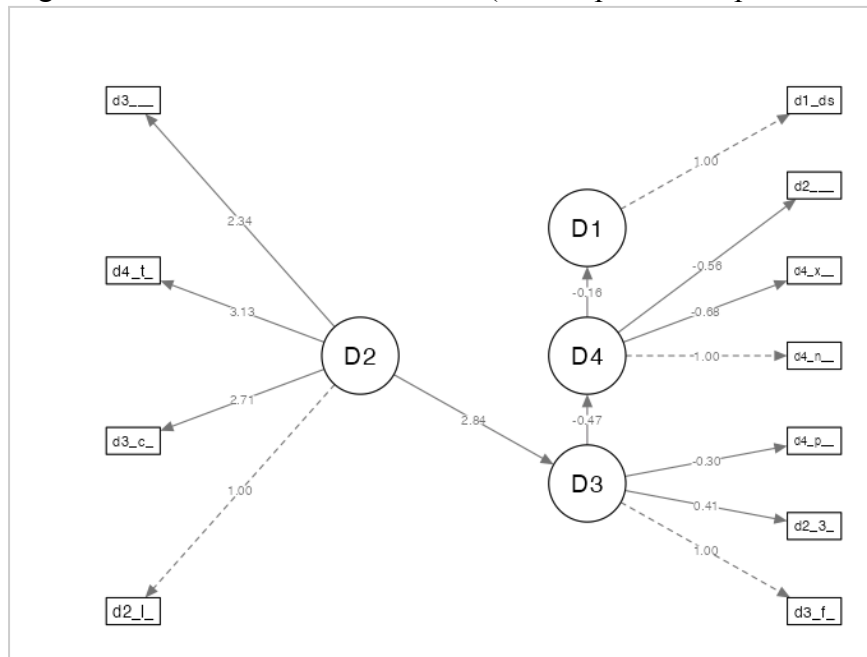
É importante notar o caráter não intuitivo desses sinais negativos: no contexto dos *logs* analisados, níveis mais elevados de atrito ou percepção negativa correlacionam-se a uma menor probabilidade de desistência conversacional. Este fenômeno sugere uma dinâmica de persistência, onde o usuário que manifesta sua frustração através da interface demonstra maior investimento cognitivo na tarefa do que aquele que desiste silenciosamente. Por fim, a ausência de significância direta em D2 sobre D1 ($p = 0.094$) reforça que a qualidade da conversacional do chatbot por si só, não expulsam o usuário, a menos que degradem severamente as dimensões D3 e D4.

Segue a avaliação do modelo 3, teórico purificado processual, conforme dados na Tabela 15 e Figura 14.

Tabela 15 - Estimativas do modelo 3 (teórico purificado processual)

Dependente	Preditor	Estimativa (B)	Erro Padrão (SE)	Coef. Padronizado (β)	R ²	p
D3	D2	2.842	0.344	0.785	61.7%	<.001
D4	D3	-0.468	0.057	-0.669	44.7%	<.001
D1	D4	-0.165	0.0484	-0.221	4.9%	<.001

Fonte: Elaborado pelo autor (2026).

Figura 14 - *Framework* do modelo 3 (teórico purificado processual)

Fonte: Elaborado pelo autor (2026).

O modelo 3 apresenta uma evolução em relação à estabilidade matemática, mantendo a premissa de uma jornada processual. A purificação das variáveis eliminou as inconsistências do modelo 1, resultando em coeficientes padronizados estáveis e estatisticamente significantes ($p < 0.001$).

Mesmo havendo um incremento para explicação da desistência ($R^2 = 4.9\%$) em relação ao modelo 1, a relevância desta configuração reside na força das relações intermediárias. A Figura 14 ilustra que D2 (qualidade conversacional) exerce um impacto significativo sobre D3 (confiança e clareza) com um coeficiente de $\beta = 0.785$, explicando 61,7% da sua variância. Esse achado demonstra que, em uma jornada linear, a qualidade conversacional (D2) tem uma representação significativa na confiança e clareza do *chatbot* (D3). Entretanto, o baixo R² final sugere que, embora o modelo processual purificado seja lógico e válido, a desistência em si parece ser influenciada por uma pressão simultânea das

dimensões (como visto no modelo 2) mais do que por uma passagem de bastão puramente processual.

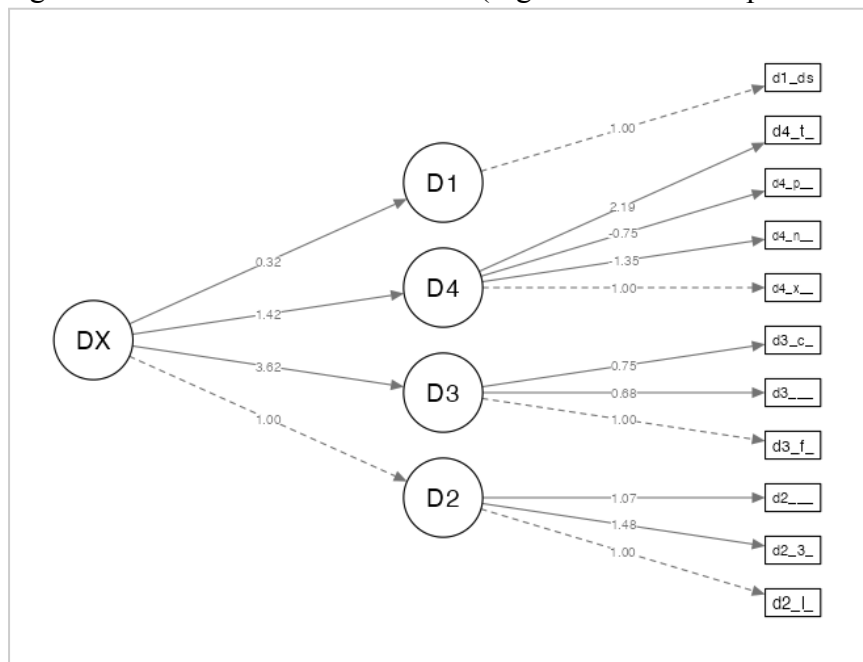
Segue a avaliação do modelo 4, segunda ordem não purificado, conforme dados na Tabela 16 e Figura 15.

Tabela 16 - Estimativas do modelo 4 (segunda ordem não purificado)

Dependente	Preditor	Estimativa (B)	Erro Padrão (SE)	Coef. Padronizado (β)	R ²	p
DX	D2	1.000	--	1.061	--	--
DX	D3	3.624	0.439	1.032	--	<.001
DX	D4	1.424	0.219	0.997	--	<.001
D1	DX	0.320	0.113	0.158	2.5%	.005

Fonte: Elaborado pelo autor (2026).

Figura 15 - *Framework* do modelo 4 (segunda ordem não purificado)



Fonte: Elaborado pelo autor (2026).

A análise do modelo 4 busca verificar se as percepções do usuário podem ser sintetizadas em um único constructo latente geral (DX) para explicar a desistência de (D1). Conforme detalhado na Tabela 16, esta configuração replica as instabilidades observadas no estágio inicial da pesquisa, com a presença de coeficientes padronizados que excedem a unidade ($\beta = 1.061$ para D2 e $\beta = 1.032$ para D3), indicando uma saturação teórica que compromete a validade matemática do modelo.

O *framework* proposto ilustra que, mesmo com o fator geral apresentando cargas fatoriais elevadas sobre as dimensões de D2, D3 e D4, sua capacidade de predição sobre D1 (desistência) é mínima, resultando em um R^2 de apenas 2.5%. Este resultado sugere que a tentativa de unificar as dimensões em um "superfator" mascara as nuances individuais de cada percepção, falhando em oferecer um ganho explicativo real em relação à visão multidimensional.

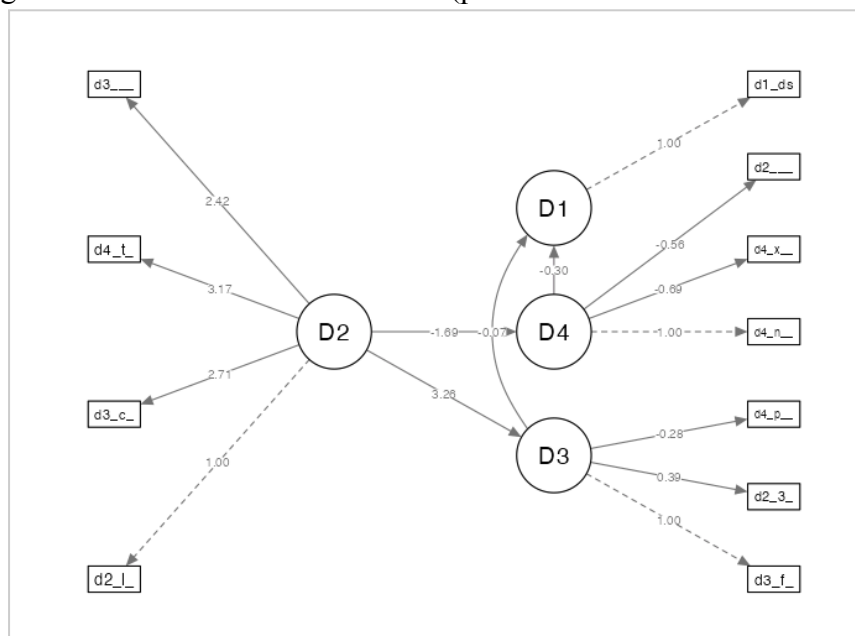
Segue a avaliação do modelo 5, purificado e D2 como antecedente, conforme dados na Tabela 17 e Figura 16.

Tabela 17 - Estimativas do modelo 5 (purificado e D2 como antecedente)

Dependente	Preditor	Estimativa (B)	Erro Padrão (SE)	Coef. Padronizado (β)	R^2	p
D3	D2	3.256	0.396	0.783	61.3%	<.001
D4	D2	-1.622	0.255	-0.660	43.5%	<.001
D1	D3	-0.074	0.036	-0.151	--	0.039
D1	D4	-0.305	0.067	-0.382	--	<.001
Poder total	Modelo 5	--	--	--	12.3%	--

Fonte: Elaborado pelo autor (2026).

Figura 16 - *Framework* do modelo 5 (purificado e D2 como antecedente)



Fonte: Elaborado pelo autor (2026).

A análise do modelo 5 avalia a configuração de mediação, onde a qualidade conversacional (D2) atua como o antecedente que molda as dimensões de confiança e clareza (D3) e engajamento e empatia (D4). Os resultados indicam que D2 exerce um impacto

positivo e forte sobre a confiança e clareza ($\beta = 0,783$), explicando 61,3% de sua variância. Em contrapartida, observa-se um efeito negativo no engajamento e empatia ($\beta = -0.660$), com um R^2 de 43.5%. No que tange à explicação da desistência (D1), o modelo revela que a dimensão D4 possui maior poder preditivo ($\beta = -0.382$) do que D3 ($\beta = -0.151$), resultando em um R^2 total de 12,3%.

Para melhor compreensão, A Tabela 18 compila os resultados dos cinco modelos.

Tabela 18 - Comparação das estimativas dos modelos

Modelo	CFI	TLI	Pred.→Depend.: β e (p)	R^2
Modelo 1: teórico processual D2→D3→D4→D1	0.927	0.905	D4→D1: 0.160 (0.003)	2.6%
Modelo 2: teórico purificado multidimensional D2+D3+D4→D1	0.966	0.952	D2→D1: 0.170 (0.094); D3→D1 -0.338 (0.002); D4→D1 -0.394 (<0.001)	12.6%
Modelo 3: teórico purificado processual D2→D3→D4→D1	0.961	0.949	D2→D3: 0.785 (<0.001); D3→D4: -0.669 (<0.001); D4→D1: -0.221 (<0.001)	4.9%
Modelo 4: segunda ordem não purificado (D2,D3,D4)→Dx→D1	0.927	0.905	D2→DX: 1.061 (--); D3→DX: 1.032 (<0.001); D4→DX: 0.997 (<0.001); DX→D1: 0.158 (0.005)	2.5%
Modelo 5: purificado e D2 como antecedente D2→D3; D2→D4 D3→D1; D4→D1	0.959	0.945	D2→D3: 0.783 (<0.001); D2→D4: -0.660 (<0.001); D3→D1: -0.151 (0.039); D4→D1: -0.382 (<0.001)	12.3%

Fonte: Elaborado pelo autor (2026).

A análise do modelo 5 demonstra que a transição de uma relação direta (modelo 2 purificado) para uma estrutura de mediação justifica a definição de seguir com este modelo como o interpretativo para esta pesquisa. Enquanto no modelo 2 a relação direta entre D2 e D1 perde significância, no modelo 5 a configuração de D2 como preditor das dimensões D3 e D4 revela uma melhor força explicativa. Os resultados confirmam que o impacto da qualidade conversacional (D2) sobre o comportamento do usuário não é imediato, mas mediado por camadas cognitivas ($\beta = 0.783$) e afetivas ($\beta = -0.660$), que juntas elevam a precisão do *framework*. Assim, a adoção do modelo 5 supera a visão do modelo direto, comprovando que a estrutura de mediação captura nuances da interação que seriam ignoradas em uma abordagem multidimensional ou sequencial.

5.4 DETERMINANTES DA DESISTÊNCIA NA CONVERSAÇÃO

Para a validação das hipóteses fundamentais desta pesquisa, a análise não se limitou a um único teste de causalidade direta, mas sim a uma avaliação comparativa de como as dimensões de interação se comportam em diferentes estruturas lógicas (modelos 2 e 5).

No modelo 2 (multidimensional purificado), as hipóteses são testadas sob a premissa de que a desistência (D1) é o resultado imediato e independente de falhas na qualidade conversacional (D2), confiança e clareza (D3) ou engajamento e empatia (D4). Nesta visão, cada deficiência do *chatbot* soma-se diretamente ao peso da decisão do usuário de abandonar a sessão, sem que uma variável interfira na interpretação da outra. É uma análise de impacto isolado, útil para identificar qual dimensão da conversação está mais fragilizada.

Por outro lado, o modelo 5 (purificado e D2 como antecedente) reinterpreta essas relações sob a ótica da dependência cognitiva. Aqui, as hipóteses são testadas primeiro quanto a predição de D2 a D3 e D4. A lógica teórica subjacente é que o usuário não desiste apenas porque o *chatbot* foi “pouco empático”, mas porque essa falta de empatia ou clareza degradou a percepção de eficiência (qualidade conversacional do *chatbot*) da ferramenta, tornando o esforço de continuar injustificável.

Dessa forma, a validação das hipóteses no modelo 5 permite entender não apenas se uma variável causa a desistência, mas como ela percorre o fluxo da experiência do usuário. Enquanto o modelo 2 nos diz qual dimensão melhor representa o abandono, o modelo 5 nos explica o “mecanismo” do abandono, revelando se a D2 atua como o filtro final que consolida as falhas de comunicação e engajamento representadas em D3 e D4 e as transforma na decisão de interromper a tarefa.

5.4.1 O descolamento entre performance e desistência (H1)

A análise da Hipótese 1 (H1) - “agentes conversacionais que demonstram baixa qualidade conversacional (D2) aumentam a desistência (D1)” - resultou na sua rejeição estatística. A falha da hipótese justifica-se, em primeiro lugar, pela reestruturação dos dados apresentados na Tabela 7; a inviabilidade dos modelos originais exigiu a purificação da variável para qualidade conversacional (D2), que passou a representar não apenas a performance técnica, mas a fluidez percebida da interação.

No modelo 2, a rejeição de H1 ($p = 0.094$) indica que a baixa qualidade conversacional, por si só, não é um preditor imediato de desistência. O usuário demonstra uma resiliência técnica superior à prevista, tolerando imperfeições desde que as dimensões de

confiança, clareza e empatia sustentem a interação. O erro é tratado como um ruído que o usuário está disposto a filtrar enquanto percebe valor na jornada.

Esta lógica consolida-se no modelo 5, onde a qualidade conversacional (D2) atua como mediadora. Embora D2 receba os impactos negativos de D3 (confiança e clareza) e D4 (engajamento e empatia), o seu caminho direto para a desistência (D1) permanece não significativo. Isto indica que, no contexto de IAs modernas, a desistência não é um evento binário disparado por um erro factual. O utilizador persiste mesmo sob atrito técnico, evidenciando que o eixo do sucesso se deslocou da perfeição técnica para a gestão da experiência interacional; enquanto os pilares relacionais sustentarem a percepção holística de valor, a baixa qualidade técnica, por si só, não causará o abandono. Este resultado contraria as premissas clássicas da *Cognitive Fit Theory* (Vessey, 1991), sugerindo que, em interfaces conversacionais avançadas, o ajuste cognitivo não depende apenas da precisão da saída, mas da capacidade do sistema em sustentar o fluxo do diálogo mesmo diante de falhas de performance.

5.4.2 A influência da confiança e clareza na desistência (H2)

A Hipótese 2 (H2) — "agentes conversacionais que demonstram baixa confiança e clareza (D3) aumentam a desistência (D1)" — foi confirmada com alta significância estatística, mesmo que o sinal do coeficiente revele um comportamento inverso: o paradoxo da persistência. É importante ressaltar que, devido à inviabilidade do modelo original, esta dimensão foi reestruturada conforme a Tabela 7.

No modelo 2 (purificado multidimensional), a relação apresentou $\beta = -0.338$ ($p = 0.002$). Na escala de atrito adotada (0 a 10), onde valores maiores indicam maior falta de confiança e clareza, o coeficiente negativo demonstra que o aumento do estresse comunicativo está inversamente correlacionado à desistência imediata. Teoricamente, isso contraria a visão de que a falha na presença social, a *Social Response Theory* (Adam, Wessel e Benlian, 2022), expulsaria o usuário de imediato. Pelo contrário, a baixa confiança e clareza gera um ciclo de reformulações que retém o usuário na interface, forçando-o a um maior investimento cognitivo para ser compreendido antes de finalmente abandonar a sessão.

No modelo 5 (purificado com D2 como antecedente), a estrutura revela que o impacto da falta de confiança e clareza (D3) é processual, atuando como um antecedente que degrada a qualidade conversacional (D2) ($\beta = 0.783$; $p < 0.001$). Nesta configuração, o peso direto de D3 sobre a desistência diminui ($\beta = -0.151$; $p = 0.039$), indicando que o "sentimento

de dúvida" do usuário é filtrado pela percepção de qualidade do diálogo. Porém, o modelo 5 também invalida H2, pois assim como o modelo 2, a baixa confiança e clareza gera um ciclo de reformulações que retém o usuário na interface.

5.4.3 A influência do engajamento e empatia na desistência (H3)

A Hipótese 3 (H3) — “agentes conversacionais que demonstram menor engajamento e empatia (D4) aumentam a desistência (D1)” — apresentou a maior força preditiva observada nesta pesquisa, embora sua validação estatística, assim como D3, tenha ocorrido com o sinal invertido, consolidando o paradoxo da persistência. Tal como nas dimensões anteriores, a D4 foi purificada conforme a Tabela 7.

No modelo 2, a relação direta apresentou $\beta = -0.394$ ($p < 0.001$). Na escala de métricas de *log* utilizada, onde valores elevados em D4 representam uma percepção de baixo engajamento e empatia, o coeficiente negativo revela que o agravamento desse atrito está, na prática, inversamente associado ao abandono imediato. Conforme o *Communication Accomodation Theory* (Zhang; Giles, 2016), esperava-se que a empatia fosse o motor da retenção positiva, mas os dados sugerem que, na ausência, ocorre uma persistência por esforço de adaptação. Conforme observado por Sheehan, Jin e Gottlieb (2020), falhas de comunicação que exigem clarificação demandam um investimento cognitivo e esforço adicional do usuário. Os resultados indicam que esse investimento atua como um retentor mecânico: o usuário permanece na interface para tentar mitigar a má interação e concluir a tarefa.

No modelo 5, o engajamento e empatia (D4) são condicionados pela qualidade conversacional (D2), apresentando um impacto de $\beta = -0.660$ ($p < 0.001$). Nesta configuração processual, a D4 mantém-se como o preditor mais forte da desistência ($\beta = -0,382$). Entretanto, o modelo 5 também invalida a premissa linear de H3, pois demonstra que, assim como no modelo 2, o baixo engajamento e empatia (D4 elevado) não expulsa o usuário, mas o mantém em um ciclo de tentativas sucessivas. Este fenômeno, reforçado pela visão de Nguyen, Waizenegger e Techatassanasoontorn (2022) sobre a capacidade de adaptação do usuário, indica que a falta de empatia não causa uma fuga súbita, mas sim uma permanência ineficiente. Assim, o resultado contraria a *Communication Accommodation Theory* (Zhang; Giles, 2016), validando que a empatia é fundamental para que a finalização da tarefa ocorra de forma fluida e não através do exaurimento do esforço do usuário que tenta sustentar a percepção holística de valor diante de um agente não acomodativo.

5.5 SÍNTESE DOS ACHADOS E DISCUSSÃO INTEGRADA

A análise conjunta dos resultados revela que a desistência no uso de agentes conversacionais (D1) não é um evento linear causado meramente por falhas técnicas ou qualidade conversacional, mas um fenômeno mediado pelo papel ativo e pelo esforço do usuário. O cruzamento dos dados demonstra que a "falha" percebida assume um papel inesperado na retenção durante a jornada.

O ponto central desta discussão é o contraste entre as dimensões. Enquanto a Hipótese 1 (performance) foi rejeitada, indicando que o erro técnico isolado não é o principal motor do abandono, as Hipóteses 2 (confiança e clareza) e 3 (engajamento empatia) confirmaram que o atrito percebido retém o usuário. Como observado pelos coeficientes negativos, quanto maior a percepção de baixa clareza ou falta de empatia, maior é o tempo de permanência na sessão. E esta dinâmica é intensificada pelos resultados do Modelo 5, onde a qualidade conversacional (D2) assume o papel de variável mediadora. Os dados indicam que falhas técnicas e qualidade da conversa não levam à desistência imediata, mas sim a um redirecionamento do esforço do usuário para as dimensões de D3 e D4. Isso reforça as hipóteses H2 e H3, demonstrando que, diante de um erro de performance, a ausência de um agente acomodativo e claro torna o processo de recuperação da tarefa oneroso, prendendo o usuário em um ciclo de persistência ineficiente para sustentar o valor da interação.

Este comportamento reflete o que se observou nos *logs* como uma persistência sob atrito: o usuário, diante de uma interação ineficiente, não desiste de imediato, mas investe em um esforço de adaptação e múltiplas reformulações para tentar concluir sua tarefa. Em suma, a discussão aponta que a verdadeira desistência ocorre de forma mais frequente pela apatia interacional: quando o sistema falha em manter o engajamento mínimo, o usuário perde o incentivo para continuar tentando, resultando no encerramento da sessão. Este achado desloca o foco do sucesso do *chatbot* da "perfeição da resposta" para a "eficiência na gestão do esforço do usuário".

6 CONTRIBUIÇÕES, LIMITAÇÕES E ESTUDOS FUTUROS

A presente pesquisa propôs um modelo multidimensional para avaliar a interação entre usuários e *chatbots* baseados em modelos de inteligência artificial, focando na predição da desistência conversacional (D1). Ao analisar variáveis de performance, qualidade conversacional, confiança, clareza, engajamento e empatia, o estudo revelou que o comportamento do usuário é mais complexo do que a simples resposta a erros técnicos, sendo mediado por um esforço de adaptação contraintuitivo.

6.1 CONTRIBUIÇÕES TEÓRICAS

A principal contribuição teórica deste trabalho reside na identificação do comportamento de persistência sob atrito, fenômeno que desafia a aplicação estrita da *Cognitive Fit Theory* (Vessey, 1991). Ao contrário do que sugere a teoria clássica, onde a falta de ajuste entre sistema e tarefa levaria à interrupção, os resultados do modelo 5 (purificado e D2 como antecedente) revelam que a qualidade conversacional (D2 purificado) atua como um eixo mediador: falhas técnicas e qualidade da conversa não geram abandono imediato, mas sim um redirecionamento do esforço do usuário. Este modelo sugere que o erro de performance "empurra" o usuário para um investimento cognitivo adicional em confiança e clareza (D3 purificado) e engajamento e empatia (D4 purificado) para tentar salvar a interação. Assim, a retenção em *chatbots* é redefinida: sessões longas e com baixa performance técnica indicam um alto custo de adaptação ativa do humano, onde o usuário assume o papel de "acomodador" para superar as limitações do agente.

Além disso, o modelo avança ao integrar as perspectivas da *Social Response Theory* (Adam; Wessel; Benlian, 2022) e da *Communication Accommodation Theory* (Zhang; Giles, 2018). O engajamento do usuário, portanto, é condicionado por processos de mediação cognitiva e social: a qualidade conversacional (D2) atua como o antecedente que define os níveis de confiança e clareza (D3) e de engajamento e empatia (D4). Sob essa ótica, a desistência conversacional (D1) deixa de ser um evento linear decorrente de um erro técnico isolado e passa a ser compreendida como o resultado do esgotamento da capacidade de adaptação do usuário. De acordo com a *Cognitive Fit Theory* (Vessey, 1991), a interrupção ocorre quando o custo do esforço cognitivo necessário para processar as inconsistências

geradas pela performance (D2) e pela falta de ajuste (*non-accommodation*) do agente supera a utilidade percebida na conclusão da tarefa.

A Figura 17 apresenta o modelo selecionado que melhor representa a pesquisa, detalhando as relações de mediação e os coeficientes que sustentam esta análise.

Figura 17 - *Framework* selecionado: modelo purificado e D2 como antecedente



Fonte: Elaborado pelo autor (2026).

6.2 CONTRIBUIÇÕES PRÁTICAS E GERENCIAIS DA DISSERTAÇÃO

As implicações práticas deste estudo transcendem a análise estatística, oferecendo diretrizes diretas para o ciclo de desenvolvimento e sustentação de agentes conversacionais. A primeira contribuição é a tradução das dimensões identificadas (D2, D3 e D4) em requisitos de desenvolvimento. Ao evidenciar que a desistência (D1) é precedida por um comportamento de persistência sob atrito, o modelo orienta o gestor de tecnologia a priorizar a redução da carga cognitiva como um requisito não-funcional crítico.

A validação empírica deste modelo pode ser observada nos *logs* de interação que fundamentaram o refinamento da solução. No caso do *chatbot* IA Beto, identificou-se um alto índice de persistência vinculado à dimensão de confiança e clareza (D3): mesmo diante de respostas repetitivas sobre pacotes de preços, um usuário anonimizado não desistiu (Figura 18), mas investiu esforço adicional para sanar dúvidas específicas de horários. Esse achado permitiu o refinamento do artefato de instrução (*prompt*) para que o agente passe a reconhecer

padrões de dúvida persistente, disparando respostas mais estruturadas e menos redundantes, reduzindo o esforço de tradução do usuário.

Outro exemplo crítico que moldou as contribuições gerenciais foi extraído do *chatbot* Abby. No *log* do usuário anonimizado (Figura 19), observou-se uma falha na dimensão de engajamento e empatia: o agente foi incapaz de realizar a transição para o atendimento humano, gerando um custo de interação elevado. A partir desse resultado, a solução pode ser melhorada com a implementação de requisitos de acomodação comunicacional, onde o agente identifica momentos de frustração e antecipa a transição para o suporte humano, evitando que o "limiar de ruptura" do cliente seja atingido.

Essas intervenções indicam que o trabalho entrega não apenas um modelo teórico, mas uma ferramenta de diagnóstico prescritivo. Ao alimentar um LLM com a instrução específica desenvolvida nesta pesquisa (Apêndice H) e os dados de volume de conversas, o gestor de *software* obtém um diagnóstico que sugere onde intervir de forma multidimensional. A solução deixa de ser focada apenas em "corrigir erros" (D2) e passa a gerir a jornada do usuário, utilizando a empatia e a clareza como mecanismos para reduzir a fricção e sustentar a retenção no funil de vendas.

Figura 18 - Recorte de *log* de interação - *chatbot* IA Beto (Betina)

Participante	Hora	Mensagem
Usuário A	11:20	Olá bom dia
IA Betina	11:20	Bom dia! 😊 Eu sou Betina, a inteligência artificial da BETO'S SCHOOL. Como posso ajudar você hoje?
Usuário A	11:23	E quais os valores?
IA Betina	11:23	[IA envia bloco longo com Pacotes 1, 2 e 3...] Pacote 3: R\$ 340,00 mensais ou R\$ 299,00 se pagar trimestral.
Usuário A	11:25	Esse de 340,00 reais é referente a 3 meses?
IA Betina	11:25	Isso mesmo! O valor de R\$ 340,00 é mensal, mas no plano trimestral sai por R\$ 299,00.
Usuário A	11:26	De 3 em 3 meses faço o pagamento de 299 correto?
IA Betina	11:26	[IA repete o bloco anterior de preços sem ajuste de clareza...]
Usuário A	11:28	Entendi. E no sábado e domingo, quais os horários?
IA Betina	11:28	Temos horários flexíveis! Sábado das 08h às 12h. Deseja agendar uma aula experimental?
Usuário A	11:30	Como faço para iniciar?

Fonte: Elaborado pelo autor (2026).

Figura 19 - Recorte de *log* de interação - *chatbot* Abby

Participante	Hora	Mensagem
Usuário B	14:10	Gostaria de entender como funciona a alocação de squads para meu projeto.
IA Abby	14:10	[IA envia bloco técnico extenso sobre orçamentos de R\$ 8 a 20 mil e modelos de alocação...]
Usuário B	14:12	[Envia mensagem curta/em branco - sinal de processamento da carga cognitiva]
IA Abby	14:12	[Repete informações técnicas sobre squads e não reconhece o vácuo na comunicação]
Usuário B	14:15	Pode passar um resumo do que falamos para a Leticia (humana)? Quero falar com ela.
IA Abby	14:15	Desculpe, Usuário B, mas não consigo enviar mensagens diretamente para outros atendentes.
Usuário B	14:16	Poxa, achei que vocês eram integrados. Como faço então?
IA Abby	14:16	Você pode entrar em contato pelo e-mail ou aguardar. Deseja saber mais sobre nossos preços?
Usuário B	14:18	Vou aguardar então. Mas me explica melhor esse valor de 8 mil...

Fonte: Elaborado pelo autor (2026).

6.3 LIMITAÇÕES DA PESQUISA

Mesmo com este estudo fornecendo evidências sobre os preditores da desistência conversacional, é necessário reconhecer certas limitações inerentes ao desenho da pesquisa, as quais delimitam o escopo de generalização dos resultados, detalhados a seguir.

6.3.1 Métrica de desistência como proxy de aceitação

Uma limitação deste estudo reside no uso da "desistência" (D1) como uma medida indireta (*proxy*) para a aceitação tecnológica. Por mais que o abandono da sessão seja um indicador comportamental de interrupção, ele não captura integralmente as nuances da "não aceitação" psicológica. Estatisticamente, o modelo 5 apresentou um coeficiente de explicação (R^2) de 12.6% para a variável desistência. Este valor, embora indique que as dimensões do modelo possuem impacto significativo e mensurável, revela que a maior parte da variância no comportamento de abandono é influenciada por fatores externos aos *logs* de interação.

Este resultado não diminui a validade do modelo; pelo contrário, ele sinaliza que o presente trabalho avança ao identificar os mecanismos técnicos e comunicacionais básicos, mas deixa claro que uma agenda de pesquisa se abre para descobrir novos fatores. O hiato de 87.7% na explicação sugere que variáveis como a motivação intrínseca do usuário, a urgência

da tarefa, a experiência prévia com outros agentes ou o estado emocional pré-interação desempenham um papel decisivo na continuidade da conversa.

Portanto, pesquisas futuras devem adotar abordagens multimétodo, integrando os *logs* de interação a instrumentos de autorrelato (*surveys*). Essa combinação permitiria validar se os usuários que "persistem sob atrito" - identificados nesta pesquisa - o fazem por uma aceitação genuína ou por absoluta falta de alternativa, permitindo mapear o exato ponto em que o investimento cognitivo se torna insustentável. O modelo aqui apresentado serve, assim, como o alicerce para uma compreensão multidimensional do comportamento humano diante de agentes conversacionais autônomos.

6.3.2 Restrições dos dados de *log*

O estudo baseou-se exclusivamente em dados secundários provenientes de *logs* de interação. Mesmo que esses dados ofereçam uma visão objetiva do comportamento real (*what users do*), eles carecem de profundidade sobre as motivações internas (*why they do it*). A ausência de variáveis sociodemográficas, como idade, nível de escolaridade e proficiência digital dos usuários, impede a análise de como esses fatores poderiam moderar o esforço de adaptação. Por exemplo, usuários mais jovens podem ter um limiar de paciência menor para a falta de empatia (D4) do que usuários mais velhos, uma distinção que os *logs* anônimos não permitem realizar.

6.3.3 Desafios metodológicos e ferramental estatístico

A natureza da variável dependente principal (D1), sendo de caráter binário (desistiu ou não desistiu), impôs limitações no uso de Modelagem de Equações Estruturais (SEM) convencional em *softwares* como o Jamovi. A impossibilidade de executar um modelo de caminhos (*path analysis*) integrado com variáveis endógenas binárias exigiu a decomposição da análise em regressões logísticas e lineares isoladas. Embora o rigor estatístico tenha sido mantido, essa abordagem limita a capacidade de observar efeitos de mediação e covariância mútua entre todos os construtos simultaneamente em um único modelo estrutural

6.3.4 Especificidade do contexto e da tarefa

Os dados refletem interações em um cenário de atendimento que pode ter alta relevância para o usuário. A persistência observada diante de falhas de confiança e clareza (D3) e engajamento e empatia (D4) pode estar condicionada à natureza crítica da tarefa. Em contextos de lazer ou busca de informações triviais, o comportamento de persistência sob atrito poderia ser substituído por um abandono imediato. Assim, os resultados aqui apresentados devem ser interpretados dentro do universo de *chatbots* de vendas, podendo variar em outros nichos de mercado.

6.4 ESTUDOS FUTUROS

Os resultados e as limitações desta investigação oferecem base para o desenvolvimento de novos estudos, particularmente no âmbito de investigações aplicadas em ciência de dados e comportamento do consumidor, na relação humano-máquina. Primeiramente, sugere-se investigar como o modelo de persistência se comporta ao segmentar a amostra entre interações de cariz profissional (B2B) e de consumo final (B2C). É plausível que o utilizador em contexto B2B apresente uma tolerância ao atrito e uma persistência ainda maiores devido à natureza obrigatória ou crítica da tarefa, enquanto o consumidor final pode apresentar um limiar de desistência mais sensível a falhas de empatia (D4). Complementarmente, futuras investigações deverão analisar se a complexidade da jornada de venda altera o peso dos construtos de performance (D2), confiança e clareza (D3) e engajamento e empatia (D4), uma vez que, em jornadas de alta complexidade, a necessidade de segurança pode intensificar o esforço de adaptação, tornando o paradoxo da persistência ainda mais evidente.

No campo tecnológico, propõe-se a evolução deste modelo estatístico através do uso de algoritmos avançados de classificação e *Deep Learning*. Para além da regressão logística tradicional, a implementação de modelos preditivos em tempo real permitiria que as empresas identificassem o "ponto de exaustão" do utilizador, possibilitando intervenções proativas, como o transbordo para um assistente humano, antes que a persistência se transforme em frustração irreversível e danos à imagem da marca. Neste sentido, recomenda-se também a integração de ferramentas de análise de sentimento para verificar se a retenção sob atrito resulta num "custo invisível" para a organização. Um estudo longitudinal poderia responder se

o utilizador que investe um esforço cognitivo elevado hoje é o mesmo que abandonará o canal de forma definitiva no futuro, apesar de ter concluído a tarefa na interação atual.

Por fim, com o crescente advento de modelos de linguagem de grande escala (LLMs), torna-se pertinente replicar este estudo comparando especificamente *chatbots* baseados em regras com *chatbots* generativos. Esta comparação permitiria validar se a maior fluidez conversacional destas novas tecnologias reduz a necessidade de esforço de adaptação do utilizador ou se, pelo contrário, eleva o patamar de expectativa e, conseqüentemente, a frustração perante eventuais falhas de clareza ou alucinações técnicas. Através destas frentes, a investigação poderá transitar de um modelo preditivo de desistência para um sistema de gestão inteligente da experiência humano-computador.

REFERÊNCIAS

- ADAM, M.; WESSEL, M.; BENLIAN, A. AI-based chatbots in customer service and their effects on user compliance. **Electronic Markets**, [s. l.], v. 31, n. 2, p. 427-445. Disponível em: <https://doi.org/10.1007/s12525-020-00414-7>. Acesso em: 6 dez. 2024.
- ASHFAQ, M.; YUN, J.; YU, S.; LOUREIRO, S. M. C. I, Chatbot: Modeling the determinants of users' satisfaction and continuance intention of AI-powered service agents. **Telematics and Informatics**, [s. l.], v. 54, p. 101473, 2020. Disponível em: <https://doi.org/10.1016/j.tele.2020.101473>. Acesso em: 24 nov. 2024.
- BACH, T. A.; KHAN, A.; HALLOCK, H.; BELTRÃO, G.; SOUSA, S. A Systematic Literature Review of User Trust in AI-Enabled Systems: An HCI Perspective. **International Journal of Human-Computer Interaction**, [s. l.], v. 40, n. 5, p. 1251-1266, 2024. Disponível em: <https://doi.org/10.1080/10447318.2022.2138826>. Acesso em: 20 nov. 2024.
- BECKER, M.; MAHR, D.; ODEKERKEN-SCHRÖDER, G. Customer comfort during service robot interactions. **Service Business**, [s. l.], v. 17, n. 1, p. 137-165, 2023. Disponível em: <https://doi.org/10.1007/s11628-022-00499-4>. Acesso em: 20 nov. 2024.
- BENADDI, L.; OUADDI, C.; JAKIMI, A.; OUCHAO, B. A Systematic Review of Chatbots: Classification, Development, and Their Impact on Tourism. **IEEE Access**, [s. l.], v. 12, p. 78799-78810, 2024. Disponível em: <https://doi.org/10.1109/access.2024.3408108>. Acesso em: 03 jan. 2025.
- BEHERA, R. K.; BALA, P. K.; RAY, A. Cognitive Chatbot for Personalised Contextual Customer Service: Behind the Scene and beyond the Hype. **Information Systems Frontiers**, [s. l.], v. 26, n. 3, p. 899-919, 2024. Disponível em: <https://doi.org/10.1007/s10796-021-10168-y>. Acesso em: 03 jan. 2025.
- BHATTACHARYA, C.; SINHA, M. The Role of Artificial Intelligence in Banking for Leveraging Customer Experience. **Australasian Business, Accounting and Finance Journal**, [s. l.], v. 16, p. 89-105, 2022. Disponível em: <https://doi.org/10.14453/aabfj.v16i5.07>. Acesso em: 23 nov. 2024.
- BLAZEVIC, V.; SIDAQUI, K. The TRISEC framework for optimizing conversational agent design across search, experience and credence service contexts. **Journal of Service Management**, [s. l.], v. 33, n. 4/5, p. 733-746, 2022. Disponível em: <https://doi.org/10.1108/JOSM-10-2021-0402>. Acesso em: 23 nov. 2024.
- BLÜMEL, J. H.; ZAKI, M.; BOHNÉ, T. Personal touch in digital customer service: a conceptual framework of relational personalization for conversational AI. **Journal of Service Theory and Practice**, [s. l.], v. 34, n. 1, p. 33-65, 2024. Disponível em: <https://doi.org/10.1108/JSTP-03-2023-0098>. Acesso em: 20 nov. 2024.
- CAMILLERI, M. A.; TROISE, C. Live support by chatbots with artificial intelligence: A future research agenda. **Service Business**, [s. l.], v. 17, n. 1, p. 61-80, 2023. Disponível em: <https://doi.org/10.1007/s11628-022-00513-9>. Acesso em: 20 nov. 2024.

CHANG, T.; HSIAO, W. Understand resist use online customer service chatbot: an integrated innovation resist theory and negative emotion perspective. **Aslib Journal of Information Management**, [s. l.], no prelo, 2024. Disponível em: <https://doi.org/10.1108/AJIM-12-2023-0551>. Acesso em: 20 nov. 2024.

CHEN, J.; LE, T.; FLORENCE, D. Usability and responsiveness of artificial intelligence chatbot on online customer experience in e-retailing. **International Journal of Retail & Distribution Management**, [s. l.], v. 49, n. 11, p. 1512-1531, 2021. Disponível em: <https://doi.org/10.1108/IJRDM-08-2020-0312>. Acesso em: 20 nov. 2024.

CHENG, X.; BAO, Y.; ZARIFIS, A.; GONG, W.; MOU, J. Exploring consumers' response to text-based chatbots in e-commerce: the moderating role of task complexity and chatbot disclosure. **Internet Research**, [s. l.], v. 32, n. 2, p. 496-517, 2022. Disponível em: <https://doi.org/10.1108/INTR-08-2020-0460>. Acesso em: 23 dez. 2024.

CROLIC, C.; THOMAZ, F.; HADI, R.; STEPHEN, A. T. Blame the Bot: Anthropomorphism and Anger in Customer-Chatbot Interactions. **Journal of Marketing**, [s. l.], v. 86, n. 1, p. 132-148, 2022. Disponível em: <https://doi.org/10.1177/00222429211045687>. Acesso em: 20 nov. 2024.

DE ANDRADE, I. M.; TUMELERO, C. Increasing customer service efficiency through artificial intelligence chatbot. **Revista de Gestão**, [s. l.], v. 29, n. 3, p. 238-251, 2022. Disponível em: <https://doi.org/10.1108/REGE-07-2021-0120>. Acesso em: 20 nov. 2024.

FAN, H.; GAO, W.; HAN, B. Are AI chatbots a cure-all? The relative effectiveness of chatbot ambidexterity in crafting hedonic and cognitive smart experiences. **Journal of Business Research**, [s. l.], v. 156, p. 113526, 2023. Disponível em: <https://doi.org/10.1016/j.jbusres.2022.113526>. Acesso em: 9 jan. 2025.

FAN, H.; HAN, B.; GAO, W.; LI, W. How AI chatbots have reshaped the frontline interface in China: examining the role of sales-service ambidexterity and the personalization-privacy paradox. **International Journal of Emerging Markets**, [s. l.], v. 17, n. 4, p. 967-986, 2022. Disponível em: <https://doi.org/10.1108/IJOEM-04-2021-0532>. Acesso em: 23 dez. 2024.

FERRARO, C.; DEMSAR, V.; SANDS, S.; RESTREPO, M.; CAMPBELL, C. The paradoxes of generative AI-enabled customer service: A guide for managers. **Business Horizons**, [s. l.], v. 67, n. 5, p. 549-559, 2024. Disponível em: <https://doi.org/10.1016/j.bushor.2024.04.013>. Acesso em: 7 dez. 2024.

FLORIDI, L.; COWLS, J. A Unified Framework of Five Principles for AI in Society. **Harvard Data Science Review**, [s. l.], v. 1, n. 1, 2019. Disponível em: <https://doi.org/10.1162/99608f92.8cd550d1>. Acesso em: 15 maio 2025.

GOOGLE. Google Analytics. Disponível em: <https://analytics.google.com>. Acesso em: 20 nov. 2024.

GRIFFIN, A. C.; KHAIRAT, S.; BAILEY, S. C.; CHUNG, A. E. A chatbot for hypertension self-management support: user-centered design, development, and usability testing. **JAMIA Open**, [s. l.], v. 6, n. 3, p. ooad073, 2023. Disponível em: <https://doi.org/10.1093/jamiaopen/ooad073>. Acesso em: 03 jan. 2025.

GRIMES, G. M.; SCHUETZLER, R. M.; GIBONEY, J. S. Mental models and expectation violations in conversational AI interactions. **Decision Support Systems**, [s. l.], v. 144, p. 113515, 2021. Disponível em: <https://doi.org/10.1016/j.dss.2021.113515>. Acesso em: 20 nov. 2024.

GU, C.; ZHANG, Y.; ZENG, L. Exploring the mechanism of sustained consumer trust in AI chatbots after service failures: a perspective based on attribution and CASA theories. **Humanities and Social Sciences Communications**, [s. l.], v. 11, n. 1, p. 1-12, 2024. Disponível em: <https://doi.org/10.1057/s41599-024-03879-5>. Acesso em: 03/jan. 2025.

HAIR JR., J. F.; BLACK, W. C.; BABIN, B. J.; ANDERSON, R. E.; TATHAM, R. L. **Análise multivariada de dados**. 6. ed. Porto Alegre: Bookman, 2009. Acesso em: 2 mar. 2026.

HANNAN, A.; HUSSAIN, F.; ALI, N.; EHATISHAM-UL-HAQ, M.; ASHRAF, M. U.; MOHAMMAD ALGHAMDI, A.; SAEED ALFAKEEH, A. A decentralized hybrid computing consumer authentication framework for a reliable drone delivery as a service. **PloS One**, [s. l.], v. 16, n. 4, p. e0250737, 2021. Disponível em: <https://doi.org/10.1371/journal.pone.0250737>. Acesso em: 03 jan. 2025.

HASSAN, A. A.; MOHAMED, S. K.; ABDEL-LATIF, F. F.; MOSTAFA, S. M.; ABDEL-AZIZ, M.; MAGUE, J. T.; AKKURT, M.; BRÅSE, S.; NIEGER, M. (Substituted Ylidene)Amino-2-Oxo-Indolylidene Thioureas and Bis(2-Oxo-Indolylidene)Urea from (Ylidene)Thiocarbonohydrazides and Isatylydene Malononitrile. **Journal of Heterocyclic Chemistry**, [s. l.], v. 54, n. 2, p. 959-964, 2017. Disponível em: <https://doi.org/10.1002/jhet.2660>. Acesso em: 03 jan. 2025.

HAUSER-ULRICH, S.; KÜNZLI, H.; MEIER-PETERHANS, D.; KOWATSCH, T. A smartphone-based health care chatbot to promote self-management of chronic pain (SELMA): pilot randomized controlled trial. **JMIR mHealth and uHealth**, [s. l.], v. 8, n. 4, p. e15806, 2020. Disponível em: <https://doi.org/10.2196/15806>. Acesso em: 03 jan. 2025.

HENNEKEUSER D.; VAZIRI D.; GOLCHINFAR D.; STEVENS G. What I Don't Like about You? A systematic review of impeding aspects for the usage of conversational agents. **Interacting with Computers**, [s. l.], v. 36, n. 5, p. 293-312, 2024. Disponível em: <https://doi.org/10.1093/iwc/iwae018>. Acesso em: 29 jan. 2025.

HEVNER, A. R.; MARCH, S. T.; PARK, J.; RAM, S. Design science in information systems research. **MIS Quarterly**, [s. l.], v. 28, n. 1, p. 75-105, 2004. Disponível em: <https://doi.org/10.2307/25148625>. Acesso em: 12 out. 2025.

HSU, C.; LIN, J. Understanding the user satisfaction and loyalty of customer service chatbots. **Journal of Retailing and Consumer Services**, [s. l.], v. 71, p. 103211, 2023. Disponível em: <https://doi.org/10.1016/j.jretconser.2022.103211>. Acesso em: 9 fev. 2025.

IO, H. N.; LEE, C. B. Chatbots and conversational agents: a bibliometric analysis. *In: IEEE International Conference on Industrial Engineering and Engineering Management*, 2017, [s.l.]. **Anais [...]**. [S.l.: s.n.], 2017. p. 215-219. Disponível em: <https://ieeexplore.ieee.org/document/8289883>. Acesso em: 03 jan. 2025.

- IZADI, S.; FOROUZANFAR, M. Error Correction and Adaptation in Conversational AI: A Review of Techniques and Applications in Chatbots. **AI**, [s. l.], v. 5, n. 2, p. 803-841, 2024. Disponível em: <https://doi.org/10.3390/ai5020041>. Acesso em: 7 dez. 2024.
- JAČISKO, J.; VESELÝ, V.; CHANG, K.-V.; ÖZÇAKAR, L. (How) ChatGPT-Artificial Intelligence Thinks It Can Help/Harm Physiatry. **American Journal of Physical Medicine & Rehabilitation**, [s. l.], v. 103, n. 4, p. 346-349, 2024. Disponível em: <https://doi.org/10.1097/PHM.0000000000002370>. Acesso em: 03 jan. 2025.
- KHNEYZER, C.; BOUSTANY, Z.; DAGHER, J. AI-Driven Chatbots in CRM: Economic and Managerial Implications across Industries. **Administrative Sciences**, [s. l.], v. 14, n. 8, 2024, p. 182. Disponível em: <https://doi.org/10.3390/admsci14080182>. Acesso em: 9 fev. 2025.
- LI, B.; CHANG, Y.; LIU, L.; LIU, H.; SUN, J. How does AI agent (vs. IVR system) service failure impact customer purchase behavior: mediating effect of customer involvement. **The Service Industries Journal**, [s. l.], v. 0, n. 0, p. 1-19, 2024. Disponível em: <https://doi.org/10.1080/02642069.2024.2344113>. Acesso em: 24 nov. 2024.
- LI, Y.; GAN, Z.; ZHENG, B. How do Artificial Intelligence Chatbots Affect Customer Purchase? Uncovering the Dual Pathways of Anthropomorphism on Service Evaluation. **Information Systems Frontiers**, [s. l.], 2023. Disponível em: <https://doi.org/10.1007/s10796-023-10438-x>. Acesso em: 24 nov. 2024.
- LUDIN, N.; HOLT-QUICK, C.; HOPKINS, S.; STASIAK, K.; HETRICK, S.; WARREN, J.; CARGO, T. A chatbot to support young people during the COVID-19 pandemic in New Zealand: evaluation of the real-world rollout of an open trial. **Journal of Medical Internet Research**, [s. l.], v. 24, n. 11, p. e38743, 2022. Disponível em: <https://doi.org/10.2196/38743>. Acesso em: 03 jan. 2025.
- MIN, F.; FANG, Z.; HE, Y.; XUAN, J. Research on Users' Trust of Chatbots Driven by AI: An Empirical Analysis Based on System Factors and User Characteristics. In: **2021 IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE)**, 9., 2021, Guangzhou. **Anais [...]**. [S.l.: s.n.], 2021. Disponível em: <https://doi.org/10.1109/ICCECE51280.2021.9342098>. Acesso em: 03 jan. 2025.
- MOHAMED, T. A.; KHAFGY, M. H.; EL SEDAWY, A. B.; ISMAIL, A. S. A proposed model for distinguishing between human-based and ChatGPT content in scientific articles. **IEEE Access**, [s. l.], v. 12, p. 121251-121260, 2024. Disponível em: <https://doi.org/10.1109/ACCESS.2024.3448315>. Acesso em: 03 jan. 2025.
- MOHER, D.; SHAMSEER, L.; CLARKE, M.; GHERSI, D.; LIBERATI, A.; PETTICREW, M.; SHEKELLE, P.; STEWART, L. A.; PRISMA-P Group. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. **Systematic Reviews**, [s. l.], v. 4, n. 1, p. 1-9, 2015. Disponível em: <https://doi.org/10.1186/2046-4053-4-1>. Acesso em: 23 dez. 2024.
- MOOR, J. The Dartmouth College Artificial Intelligence Conference: The Next Fifty Years. **AI Magazine**, [s. l.], v. 27, p. 87-91, 2006. Disponível em: <https://doi.org/10.1609/aimag.v27i4.1911>. Acesso em: 08 mai. 2025.

MOZAFARI, N.; WEIGER, W. H.; HAMMERSCHMIDT, M. Trust me, I'm a bot - repercussions of chatbot disclosure in different service frontline settings. **Journal of Services Management**, [s. l.], v. 33, n. 2, p. 221-245, 2021. Disponível em: <https://doi.org/10.1108/JOSM-10-2020-0380>. Acesso em: 23 dez. 2024.

NGAI, E. W. T.; LEE, M. C. M.; LUO, M.; CHAN, P. S. L.; LIANG, T. An intelligent knowledge-based chatbot for customer service. **Electronic Commerce Research and Applications**, [s. l.], v. 50, p. 101098, 2021. Disponível em: <https://doi.org/10.1016/j.elerap.2021.101098>. Acesso em: 9 fev. 2025.

NGUYEN, T. H.; WAIZENEGGER, L.; TECHATASSANASOONTORN, A. A. "Don't Neglect the User!" - Identifying Types of Human-Chatbot Interactions and their Associated Characteristics. **Information Systems Frontiers**, [s. l.], v. 24, n. 3, p. 797-838, 2022. Disponível em: <https://doi.org/10.1007/s10796-021-10212-x>. Acesso em: 29 jan. 2025.

NICOLESCU, L.; TUDORACHE, M. T. Human-Computer Interaction in Customer Service: The Experience with AI Chatbots - A Systematic Literature Review. **Electronics**, [s. l.], v. 11, n. 10, p. 1579, 2022. Disponível em: <https://doi.org/10.3390/electronics11101579>. Acesso em: 20 nov. 2024.

NUNES, T. A. M.; ROCHA, S. D. da; CATAPAN, M. F. Chatbot em realidade aumentada para atendimento ao público da Assembleia Legislativa do Amazonas: Chatbot in augmented reality to serve the public of the Legislative Assembly of Amazonas. **Revista de Gestão e Secretariado**, [s. l.], v. 13, n. 3, p. 828-841, 2022. Disponível em: <https://doi.org/10.7769/gesec.v13i3.1372>. Acesso em: 20 nov. 2024.

PANIGRAHI, R. R.; SHRIVASTAVA, A. K.; QURESHI, K. M.; MEWADA, B. G.; ALGHAMDI, S. Y.; ALMAKAYEEL, N.; ALMUFLIH, A. S.; QURESHI, M. R. N. AI Chatbot Adoption in SMEs for Sustainable Manufacturing Supply Chain Performance: A Mediation Research in an Emerging Country. **Sustainability**, [s. l.], v. 15, n. 18, p. 13743, 2023. Disponível em: <https://doi.org/10.3390/su151813743>. Acesso em: 21 nov. 2024.

PARK, Y.; SHIN, Y. A Block-Based Interactive Programming Environment for Large-Scale Machine Learning Education. **Applied Sciences**, [s. l.], v. 12, n. 24, p. 13008, 2022. Disponível em: <https://doi.org/10.3390/app122413008>. Acesso em: 03 jan. 2025.

PASCHOU, T.; ADRODEGARI, F.; PERONA, M.; SACCANI, N. The digital servitization of manufacturing: a literature review and research agenda. In: **RESER CONFERENCE**, 27., 2017, Bilbao. Proceedings... Bilbao: RESER, 2017. 2020. Disponível em: <https://www.researchgate.net/publication/319628925>. Acesso em: 23 dez. 2024.

PATIL, K. P.; KULKARNI, M. S.; HUDNURKAR, M. Enhancing service quality in the insurance industry with AI-powered humanoid chatbots. **The TQM Journal**, [s. l.], no prelo, 2024. Disponível em: <https://doi.org/10.1108/TQM-11-2023-0354>. Acesso em: 03 jan. 2025.

PAUL, A.; HAQUE LATIF, A.; AMIN ADNAN, F.; RAHMAN, R. M. Focused domain contextual AI chatbot framework for resource poor languages. **Journal of Information and Telecommunication**, [s. l.], v. 3, n. 2, p. 248-269, 2018. Disponível em: <https://doi.org/10.1080/24751839.2018.1558378>. Acesso em: 7 dez. 2024.

POOLE, D.; MACKWORTH, A.; GOEBEL, R. **Computational Intelligence: A Logical Approach**. [S. l.: s. n.], 1998. Disponível em: https://www.researchgate.net/publication/220689680_Computational_Intelligence_A_Logical_Approach. Acesso em: 8 mai. 2025.

PRAKASH, A V.; JOSHI, A.; NIM, S.; DAS, S. Determinants and consequences of trust in AI-based customer service chatbots. **The Service Industries Journal**, [s. l.], v. 43, n. 9-10, p. 642-675, 2023. Disponível em: <https://doi.org/10.1080/02642069.2023.2166493>. Acesso em: 20 nov. 2024.

PRZEGALINSKA, A.; CIECHANOWSKI, L.; STROZ, A.; GLOOR, P.; MAZUREK, G. In bot we trust: A new methodology of chatbot performance measures. **Business Horizons**, [s. l.], v. 62, n. 6, p. 785-797, 2019. Disponível em: <https://doi.org/10.1016/j.bushor.2019.08.005>. Acesso em: 17 dez. 2024.

OPENAI. **Improving language understanding by generative pre-training**. 2018. Disponível em: <https://openai.com/index/language-unsupervised/>. Acesso em: 15 maio 2025.

RADZIWILL, N.; BENTON, M. Evaluating Quality of Chatbots and Intelligent Conversational Agents. **arXiv:1704.04579**. Preprint, arXiv. [s. l.], 2017. Disponível em: <https://doi.org/10.48550/arXiv.1704.04579>. Acesso em: 19 mai. 2025.

RANIERI, A.; DI BERNARDO, I.; MELE, C. Serving customers through chatbots: positive and negative effects on customer experience. **Journal of Service Theory and Practice**, [s. l.] v. 34, n. 2, p. 191-215, 2024. Disponível em: <https://doi.org/10.1108/JSTP-01-2023-0015>. Acesso em: 9 jan. 2025.

RAU, S.; RAU, A.; NATTENMÜLLER, J.; FINK, A.; BAMBERG, F.; REISERT, M.; RUSSE, M. F. A retrieval-augmented chatbot based on GPT-4 provides appropriate differential diagnosis in gastrointestinal radiology: a proof of concept study. **European Radiology Experimental**, [s. l.], v. 8, n. 1, p. 60, 2024. Disponível em: <https://doi.org/10.1186/s41747-024-00457-x>. Acesso em: 03 jan. 2025.

ROY, R.; NAIDOO, V. Enhancing chatbot effectiveness: The role of anthropomorphic conversational styles and time orientation. **Journal of Business Research**, [s. l.], v. 126, p. 23-34, 2021. Disponível em: <https://doi.org/10.1016/j.jbusres.2020.12.051>. Acesso em: 23 dez. 2024.

RUSSELL, S.; NORVIG, P. **Artificial intelligence: a modern approach**. 3. ed. Upper Saddle River: Prentice Hall, 2010. Acesso em: 15 mai. 2025

SHEEHAN, B.; JIN, H. S.; GOTTLIEB, U. Customer service chatbots: Anthropomorphism and adoption. **Journal of Business Research**, [s. l.], v. 115, p. 14-24, 2020. Disponível em: <https://doi.org/10.1016/j.jbusres.2020.04.030>. Acesso em: 24 nov. 2024.

SOHL-DICKSTEIN, J.; WEISS, E. A.; MAHESWARANATHAN, N.; GANGULI, S. **Deep Unsupervised Learning using Nonequilibrium Thermodynamics**. Versão arXiv:1503.03585. [S. l.]: arXiv, 2015. Disponível em: <https://doi.org/10.48550/arXiv.1503.03585>. Acesso em: 15 maio 2025.

SONG, M.; XING, X.; DUAN, Y.; COHEN, J.; MOU, J. Will artificial intelligence replace human customer service? The impact of communication quality and privacy risks on adoption

intention. **Journal of Retailing and Consumer Services**, [s. l.], v. 66, p. 102900, 2022. Disponível em: <https://doi.org/10.1016/j.jretconser.2021.102900>. Acesso em: 27 nov. 2024.

SONNTAG, M.; MEHMANN, J.; TEUTEBERG, F. Deriving Trust-Supporting Design Knowledge for Ai-Based Chatbots in Customer Service: A Use Case from the Automotive Industry. **Journal of Organizational Computing and Electronic Commerce**, [s. l.], v. 33, n. 3-4, p. 178-210, 2023. Disponível em: <https://doi.org/10.1080/10919392.2023.2276631>. Acesso em: 27 nov. 2024.

TABERA, M. E. D.; MOSQUEDA, M. L. G.; TAPIA, K. M.; ACEVEDO, E. S.; OLIVARES, E. L.; COYOLI, K. L. A.; CAMARILLO, D. R. The Application of Artificial Intelligence for the Design of a Virtual Assistant (Chatbot) as an innovative approach in user service: Social Security Center No.102. **International Journal of Combinatorial Optimization Problems and Informatics**, [s. l.], v. 15, n. 4, p. 149-162, 2024. Disponível em: <https://doi.org/10.61467/2007.1558.2024.v15i4.484>. Acesso em: 27 nov. 2024.

TANWAR, M.; VERMA, H. Scientific Mapping of Chatbot Literature: A Bibliometric Analysis. **International Journal of Mathematical, Engineering and Management Sciences**, [s. l.], v. 9, p. 323-340, 2024. Disponível em: <https://doi.org/10.33889/IJMEMS.2024.9.2.017>. Acesso em: 27 nov. 2024.

URBANI, R.; FERREIRA, C.; LAM, J. Managerial framework for evaluating AI chatbot integration: Bridging organizational readiness and technological challenges. **Business Horizons**, [s. l.], v. 67, n. 5, p. 595-606, 2024. Disponível em: <https://doi.org/10.1016/j.bushor.2024.05.004>. Acesso em: 20 nov. 2024.

VALDATI, Aline de Brittos. Inteligência artificial - IA. 1. ed. São Paulo: **Contentus**, 2020. E-book. Disponível em: <https://plataforma.bvirtual.com.br>. Acesso em: 13 maio 2025.

VESELOVSKII, A. V.; GRIGOR'EV, V. A.; KAPLIN, V. A.; KARAVICHEVA, T. L.; KARAVICHEV, O. V.; KLIMOV, A. I.; KONDRAT'EVA, N. V.; KUREPIN, A. B.; KUREPIN, A. N.; LOGINOV, V. A.; MAEVSKAYA, A. I.; MARIN, V. I.; MELESHKO, E. A.; RESHETIN, A. I.; TRZASKA, W. H. A fast electronic system of the T0 start trigger detector for the ALICE experiment. **Instruments and Experimental Techniques**, [s. l.], v. 52, n. 2, p. 191-195, 2009. Disponível em: <https://doi.org/10.1134/S0020441209020080>. Acesso em: 03 jan. 2025.

VESSEY, I. Cognitive Fit: A Theory-Based Analysis of the Graphs Versus Tables Literature. **Decision Sciences**, [s. l.], v. 22, n. 2, p. 219-240, 1991. Disponível em: <https://doi.org/10.1111/j.1540-5915.1991.tb00344.x>. Acesso em: 03 jan. 2025.

VO, N. N. Y.; LIU, S.; LI, X.; XU, G. Leveraging unstructured call log data for customer churn prediction. **Knowledge-Based Systems**, [s. l.], v. 212, p. 106586, 2021. Disponível em: <https://doi.org/10.1016/j.knosys.2020.106586>. Acesso em: 29 jan. 2025.

WANG, X.; LIN, X.; SHAO, B. How does artificial intelligence create business agility? Evidence from chatbots. **International Journal of Information Management**, [s. l.], v. 66, p. 102535, 2022. Disponível em: <https://doi.org/10.1016/j.ijinfomgt.2022.102535>. Acesso em: 23 dez. 2024.

WANNER, J.; HERM, L.; HEINRICH, K.; JANIESCH, C. The effect of transparency and trust on intelligent system acceptance: Evidence from a user-based study. **Electronic Markets**, [s. l.], v. 32, n. 4, p. 2079-2102, 2022. Disponível em: <https://doi.org/10.1007/s12525-022-00593-5>. Acesso em: 23 dez. 2024.

WINSTON, P. H. **Artificial intelligence**. 3. ed. Reading: Addison-Wesley, 1992. Acesso em: 15 mai. 2025.

XIE, Y.; LIANG, C.; ZHOU, P.; JIANG, L. Exploring the influence mechanism of chatbot-expressed humor on service satisfaction in online customer service. **Journal of Retailing and Consumer Services**, [s. l.], v. 76, p. 103599, 2024. Disponível em: <https://doi.org/10.1016/j.jretconser.2023.103599>. Acesso em: 20 nov. 2024.

YANG, B.; SUN, Y.; SHEN, X. Understanding AI-based customer service resistance: A perspective of defective AI features and tri-dimensional distrusting beliefs. **Information Processing & Management**, [s. l.], v. 60, n.3, p. 103257, 2023. Disponível em: <https://doi.org/10.1016/j.ipm.2022.103257>. Acesso em: 24 nov. 2024.

YIM, M. C. Effect of AI Chatbot's Interactivity on Consumers' Negative Word-of-Mouth Intention: Mediating Role of Perceived Empathy and Anger. **International Journal of Human-Computer Interaction**, [s. l.], v. 40, n. 18, p. 5415-5430, 2024. Disponível em: <https://doi.org/10.1080/10447318.2023.2234114>. Acesso em: 29 jan. 2025.

ZAMORA, J. I'm Sorry, Dave, I'm Afraid I Can't Do That: Chatbot Perception and Expectations. In: **PROCEEDINGS OF THE 5TH INTERNATIONAL CONFERENCE ON HUMAN AGENT INTERACTION**, 2017, New York, NY, USA. **Anais [...]**. New York, NY, USA: Association for Computing Machinery, 2017. p. 253–260. Disponível em: <https://doi.org/10.1145/3125739.3125766>. Acesso em: 06 fev. 2025.

ZHANG, J.; WANG, X.; LU, J.; LIU, L.; FENG, Y. The impact of emotional expression by artificial intelligence recommendation chatbots on perceived humanness and social interactivity. **Decision Support Systems**, [s. l.], v. 187, p. 114347, 2024. Disponível em: <https://doi.org/10.1016/j.dss.2024.114347>. Acesso em: 16 dez. 2024.

ZHANG, J. J. Y.; FØLSTAD, A.; BJØRKLII, C. A. Organizational Factors Affecting Successful Implementation of Chatbots for Customer Service. **Journal of Internet Commerce**, [s. l.], v. 22, n. 1, p. 122-156, 2023. Disponível em: <https://doi.org/10.1080/15332861.2021.1966723>. Acesso em: 20 nov. 2024.

ZHANG, T.; FENG, C.; CHEN, H.; XIAN, J. Calming the Customers by AI: Investigating the Role of Chatbot Acting-Cute Strategies in Soothing Negative Customer Emotions. **Electronic Markets**, [s. l.], v. 32, n. 4, p. 2277-2292, 2022. Disponível em: <https://doi.org/10.1007/s12525-022-00596-2>. Acesso em: 23 dez. 2024.

ZHANG, Y. B.; GILES, H. Communication Accommodation Theory. **The international encyclopedia of communication theory and philosophy**. In: MATTHEWS, David; ANDERSON, Jack (org.). *The international encyclopedia of communication theory and philosophy*. Hoboken, NJ: Wiley-Blackwell, 2016. p. 1-7. Disponível em: <https://doi.org/10.1002/9781118766804.wbiect056>. Acesso em: 21 dez. 2024.

ZHONG, B.; HE, W.; HUANG, Z.; LOVE, P. E. D.; TANG, J.; LUO, H. A building regulation question answering system: A deep learning methodology. **Advanced Engineering Informatics**, [s. l.], v. 46, p. 101195, 2020. Disponível em: <https://doi.org/10.1016/j.aei.2020.101195>. Acesso em: 03 jan. 2025

APÊNDICE A - QUADRO COM ARTIGOS SINTETIZADOS 1

Quadro dos principais artigos sintetizados com contexto, fundamentação teórica e metodologia

Autor	Contexto/objetivo	Fundamentação teórica	Metodologia
Vo <i>et al.</i> (2021)	Analisar a utilização de dados não estruturados de logs de chamadas telefônicas para prever churn e melhorar estratégias de retenção no setor financeiro.	Exploração de técnicas de mineração de dados textuais, incluindo TF-IDF, Word2Vec, LIWC para análise de traços de personalidade e SHAP para interpretabilidade de modelos preditivos.	Modelagem preditiva baseada em aprendizado de máquina com múltiplos modelos (XGBoost, Random Forest, Logistic Regression e Naïve Bayes) e combinação de dados estruturados e não estruturados de conversas telefônicas.
Grimes, Schuetzler e Giboney (2021)	Explorar como a expectativa e sua violação influenciam a percepção do usuário sobre a interação com agentes conversacionais de IA.	Expectation Violation Theory (Burgoon, 1989); Interação Humano-Máquina; Presença Social.	Experimento com 175 participantes no MTurk, divididos em grupos com condições de alta e baixa expectativa e diferentes níveis de capacidade conversacional da IA.
Sheehan, Jin e Gottlieb (2020)	Investigar a relação entre falhas de comunicação e a adoção de chatbots de atendimento ao cliente, testando o papel mediador da antropomorfização e o papel moderador da necessidade de interação humana.	Antropomorfização (Epley <i>et al.</i> , 2007); Adoção de Tecnologia de Autoatendimento; Necessidade de Interação Humana (NFHI).	Dois estudos experimentais: (1) Animações mostrando interações entre humanos e chatbots, (2) Conversas reais entre humanos e chatbots, medindo intenção de adoção e antropomorfização.
Prakash, <i>et al.</i> (2023)	Explorar os determinantes da confiança dos usuários em chatbots baseados em IA e avaliar seu impacto nas intenções comportamentais dos usuários.	Modelo de Confiança na Tecnologia (Mcknight <i>et al.</i> , 2011) e Teoria da Resposta Social (Nass e Moon, 2000)..	Modelos para avaliar confiança em chatbots, incluindo fatores como presença social, utilidade percebida e risco de privacidade
Ashfaq <i>et al.</i> (2020)	Modelar os determinantes da satisfação e intenção de continuidade dos usuários de chatbots com base nos modelos ECM, ISS e TAM.	Expectation-Confirmation Model (ECM), Information System Success (ISS), Technology Acceptance Model (TAM), e necessidade de interação com funcionários humanos (NFI-SE).	Coleta de dados via MTurk com 370 usuários de chatbots. Análise estatística usando modelagem de equações estruturais (PLS-SEM).
De Andrade e Tumelero (2022)	Analisar como chatbots de IA contribuem para a eficiência no atendimento ao cliente em um banco comercial.	Teoria Evolutiva da Inovação (Nelson e Winter, 2005); Transformação Digital e Indústria 4.0.	Estudo de caso único em um banco comercial brasileiro. Entrevistas semiestruturadas com 18 participantes, analisadas com software Atlas.ti.
Nicolescu e Tudorache (2022)	Analisar a experiência geral dos clientes com chatbots de atendimento ao cliente e identificar os	Modelos de aceitação de TI (TAM, UTAUT) e modelos de satisfação do usuário (DeLone e McLean,	Revisão sistemática da literatura com 40 publicações empíricas. Análise temática e descritiva sobre fatores influenciadores, dimensões da

	principais fatores que influenciam essa experiência.	Expectation-Confirmation).	experiência e comportamentos dos clientes.
Camilleri e Troise (2023)	A pesquisa explora como chatbots sofisticados podem melhorar o engajamento e oferecer experiências de interação mais naturais e eficazes no atendimento ao cliente.	Teorias como Expectancy Theory, Social Presence Theory, e Human-Computer Interaction Theory são discutidas como base para explorar interações humano-computador.	Revisão sistemática da literatura, seguindo o protocolo PRISMA, com análise de 70 artigos mais citados em bases de dados acadêmicas.
Ngai <i>et al.</i> (2021)	Propor uma arquitetura de sistema baseada em conhecimento para melhorar o suporte ao cliente em serviços de e-commerce.	Gestão do Conhecimento do Cliente (CKM); Princípios Cooperativos de Grice; Modelos de Arquitetura para Chatbots Baseados em Conhecimento	Desenvolvimento e implementação de um protótipo de chatbot baseado em conhecimento para atendimento ao cliente em uma empresa global de lingerie
Fan, Gao e Han (2023)	Explorar a eficácia relativa da ambidestria de chatbots em criar experiências inteligentes hedônicas e cognitivas.	Modelos de Processos Duplos; Ambidestria em vendas e serviços; Experiência Inteligente.	Análise de dados empíricos de 1026 interações com clientes em um mercado de compartilhamento de bicicletas, utilizando regressão polinomial e análise de superfícies de resposta.
Khneyzer, Boustany, Dagher e (2024)	Explorar as implicações econômicas e gerenciais dos chatbots movidos a IA em sistemas de CRM em diferentes indústrias.	Teoria da Gestão de Relacionamento com o Cliente (CRM); Aplicação de IA em análise preditiva, recomendação personalizada e processamento de linguagem natural.	Abordagem qualitativa com entrevistas realizadas com especialistas em CRM e profissionais da indústria.
Fan <i>et al.</i> (2022)	Avaliar os efeitos da ambidestria vendas-serviços de chatbots na experiência do cliente e examinar o paradoxo personalização-privacidade.	Teoria de Ambidestria Organizacional; Teoria de Limites de Informação; Paradoxo Personalização-Privacidade.	Pesquisa quantitativa com 507 respostas de consumidores, analisada com regressão polinomial e análise de superfícies de resposta.
Ranieri, Di Bernardo e Mele (2024)	Examinar os efeitos conflitantes de chatbots em experiências de clientes no varejo online, explorando dimensões relacionais, cognitivas, afetivas e comportamentais.	Teorias de Experiência do Cliente (CX); Interação Humano-Tecnologia; Impacto de Serviços Inteligentes.	Análise discursiva qualitativa de 7.167 interações entre clientes e chatbots ao longo de dois anos em um varejista online europeu.
Li, Gan e Zheng (2023)	Explorar os efeitos dos traços antropomórficos dos chatbots (calor e competência) na avaliação de serviço e comportamento de compra dos clientes.	Modelo de Wixom e Todd; Teoria do Processamento de Informação Social (SIP).	Experimentos mistos: estudo de laboratório com 212 participantes e dois estudos baseados em cenários com 124 e 232 participantes.

Panigrahi <i>et al.</i> (2023)	Investigar como a adoção de chatbots de IA impacta a visibilidade da cadeia de suprimentos (SCV), capacidades de inovação (IC) e desempenho sustentável da cadeia de suprimentos (SSCP) em PMEs.	Dynamic Capabilities (DC) Theory; Sustentabilidade na cadeia de suprimentos; Teoria de adoção de tecnologias disruptivas.	Estudo empírico baseado em modelagem de equações estruturais (SEM) usando Smart PLS-4.0, com 246 respostas de gerentes de PMEs de manufatura na Índia.
Behera, Bala e Ray (2024)	Examinar a adoção de chatbots cognitivos no atendimento ao cliente B2B, integrando os modelos TAM e D&M para validar sua aceitação e impacto na experiência do usuário	Modelo de Aceitação de Tecnologia (TAM); Modelo de Sucesso de Sistemas de Informação (D&M); Integração de personalização e contextualização em serviços de atendimento ao cliente.	Pesquisa quantitativa baseada em questionários com 300 participantes de empresas B2B indianas, analisada por meio de SEM.
Patil, Kulkarni e Hudnurkar (2024)	Explorar como chatbots antropomórficos movidos por IA podem melhorar a qualidade do serviço no setor de seguros, considerando a prontidão tecnológica dos clientes.	Teoria de Prontidão Tecnológica (TRI); Antropomorfismo em sistemas de IA; Percepção de qualidade no serviço.	Estudo empírico com 688 respostas de clientes coletadas por amostragem proposital, analisadas por modelagem de equações estruturais (SEM) usando SmartPLS 4.0.
Gu, Zhang e Zeng (2024)	Explorar como manter a confiança do consumidor em chatbots de IA após falhas de serviço, integrando a Teoria dos Computadores como Atores Sociais (CASA) e a Teoria da Atribuição.	Teoria dos Computadores como Atores Sociais (CASA); Teoria da Atribuição; Ansiedade em relação à IA como moderador.	Estudo empírico baseado em 462 questionários, analisando a percepção dos usuários sobre falhas de chatbots e como a confiança é influenciada por fatores antropomórficos e atribuições de falha.
Fan <i>et al.</i> (2021)	Explorar os fatores de sistema e características do usuário que impactam a percepção de confiança em chatbots de IA no e-commerce, usando a Teoria da Presença Social.	Teoria da Presença Social; Percepção de Personalização; Riqueza de Mídia; Experiência Passada de Uso; Reatância Psicológica (emocional e cognitiva).	Teoria da Presença Social; Percepção de Personalização; Riqueza de Mídia; Experiência Passada de Uso; Reatância Psicológica (emocional e cognitiva).
Yang, Sun e Shen (2023)	Investigar como características de IA (como falta de empatia e informações irrelevantes) contribuem para a resistência do cliente a chatbots de serviço com base em crenças de desconfiança.	Teoria do Marketing Relacional; Framework CASA; Distrusting Beliefs.	Pesquisa quantitativa com 301 usuários de AISC, coletando dados por questionários e analisando crenças de desconfiança em relação à IA.
Hsu e Lin (2023)	Examinar como a qualidade dos serviços de chatbots influencia a satisfação e lealdade dos	Modelo Input-Process-Output (IPO); Qualidade do serviço de chatbots com IA; Satisfação e lealdade do usuário.	Pesquisa quantitativa baseada em 219 respostas de usuários de Taiwan, analisadas com PLS-SEM para validar as hipóteses do modelo teórico.

	usuários no contexto de e-commerce.		
Io e Lee (2017)	Examinar a pesquisa sobre chatbots utilizando análise bibliométrica para identificar tendências e lacunas de pesquisa.	Teoria de bibliometria aplicada a tecnologias emergentes; análise de redes de citações e co-ocorrência de palavras-chave.	Coleta de dados em Web of Science (WoS) e ProQuest, análise com CiteSpace e Bibliometrix.
Izadi e Forouzanfar (2024)	O artigo analisa os avanços em técnicas de correção de erros em chatbots.	Teorias de aprendizado por reforço, supervisão e métodos semi supervisionados aplicados à correção de erros; ética e viés na inteligência artificial.	Revisão sistemática de técnicas e estudos de caso em diversos setores, incluindo saúde, educação e e-commerce.
Urbani, Ferreira e Lam (2024)	Propor um framework para avaliar a prontidão organizacional e os desafios tecnológicos na integração de chatbots de IA nos processos empresariais.	Modelo de Aceitação de Tecnologia (TAM); Prontidão Organizacional; Privacidade de Dados e Interoperabilidade.	Revisão teórica com base em modelos TAM estendidos e estudos empíricos sobre integração de chatbots em organizações globais.
Blümel, Zaki e Bohné (2024)	Propor um framework conceitual para personalizar conversas de atendimento ao cliente e melhorar a experiência do cliente utilizando IA conversacional.	Teoria do Processamento de Informação Social (SIPT); Personalização Relacional; Comunicação Mediada por Computador.	Revisão sistemática da literatura para desenvolver um framework conceitual, com foco em estilos de conversa e personalização relacional em serviços baseados em texto.
Hennekeuser <i>et al.</i> (2024)	Analisar aspectos que impedem o uso de agentes conversacionais, considerando diferentes características de usuários, contextos de uso e interfaces.	Teoria da Experiência do Usuário (ISO 9241-210); Teoria de Usabilidade (ISO 9241-11); Revisões anteriores sobre barreiras à adoção de agentes conversacionais.	Revisão sistemática de literatura com 65 publicações, utilizando o método PRISMA e análise qualitativa de barreiras de usabilidade e experiência do usuário.
Tabera <i>et al.</i> (2024)	Analisar a aplicação de IA no design de um assistente virtual para melhorar o atendimento no Centro de Segurança Social No. 102.	Teoria de Sistemas de Atendimento; Automação e Personalização em Serviços Públicos; Modelos de Experiência do Usuário.	Revisão de literatura e desenvolvimento de um protótipo funcional para avaliação em campo.
Nunes, Rocha e Catapan (2022)	Explorar a implementação de um chatbot com personagem tridimensional em realidade aumentada para modernizar o atendimento ao público da Assembleia Legislativa do Amazonas.	Técnicas de Realidade Aumentada; Big Data; Teoria de Serviços Públicos Inovadores.	Revisão sistemática e proposta de protótipo inicial baseado em design de personagem tridimensional em RA, utilizando Blender e KeenTools Face Builder.

Zhang <i>et al.</i> (2022)	Investigar como estratégias de "cuteness" em chatbots, como whimsical e kindchenschema, podem acalmar emoções negativas de clientes em cenários de falhas de serviço.	Teoria do Processamento de Informação; Estratégias de Interação Humanizada por IA.	Dois estudos experimentais com 157 e 316 participantes, usando interações simuladas de falhas de entrega para avaliar emoções e percepções dos clientes.
Nguyen, Waizenegger e Techatassana soontorn (2021)	Explorar diferentes tipos de interações humano-chatbot e suas características, analisando 451 interações selecionadas entre clientes e um chatbot de um provedor de energia, a partir de um conjunto inicial de 1000 interações.	Paradigma CASA; Teoria de Presença Social; Taxonomia de Cues Sociais.	Abordagem qualitativa com análise temática de 451 interações selecionadas após limpeza de dados e exclusão de interações irrelevantes.
Yim (2024)	Analisar como a interatividade dos chatbots influencia a percepção de empatia e reduz intenções de comunicação negativa boca-a-boca (NWOM), focando na recuperação emocional em cenários de falha de produtos.	Paradigma CASA; Teoria de Resposta Social; Conceitos de Empatia e Raiva no Comportamento do Consumidor.	Experimento entre sujeitos com 300 participantes, utilizando quatro condições de interatividade (alta/baixa social e alta/baixa de mensagem) para medir efeitos em empatia percebida, raiva e NWOM.
Paul <i>et al.</i> (2018)	Propor um framework de chatbot para domínios específicos, com suporte a idiomas com poucos recursos linguísticos, utilizando dados limitados e técnicas de aprendizado semi-supervisionado.	NLP minimalista; N-Gram stemming; CenKNN e Redes Neurais Convolucionais para classificação de texto.	Desenvolvimento de um chatbot com dados em inglês e bangla, incluindo técnicas de stemming, CNN e contextualização para avaliar desempenho em ambientes de poucos recursos.
Li <i>et al.</i> (2024)	Examinar como falhas de serviço em agentes de IA versus sistemas IVR afetam o comportamento de compra dos clientes, destacando o papel mediador do envolvimento do cliente.	Teoria da Atribuição; Envolvimento do Cliente; Recuperação de Serviços em cenários de falha de agentes não humanos.	Análise de dados secundários de registros de chamadas de um call center de companhia aérea chinesa; aplicação de PSM para pareamento de grupos e regressão logística para análise do impacto das falhas de serviço.
Przejalinska <i>et al.</i> (2019)	Propor uma metodologia inovadora para medir a performance de chatbots, vinculando confiança, métricas de interação e preocupações éticas em contextos de negócios e	Teorias de Confiança Organizacional; Análise de Interações Humanas com Chatbots; Medição de Desempenho em IA.	Estudo experimental com dois tipos de chatbots (texto e avatar com som), combinando métricas psicofisiológicas, análise linguística e categorização de perfis de usuários.

	serviços ao cliente.		
Zhang <i>et al.</i> (2024)	Investigar como diferentes formas de expressão emocional em chatbots de recomendação impactam a percepção de humanidade, interatividade social e presença social em interações humano-IA.	Paradigma CASA; Teoria da Presença Social; Modelo sRAM (Service Robot Acceptance Model)	Três experimentos online baseados em cenários, testando efeitos individuais e combinados das pistas emocionais (texto, emoticons e imagens).
Li <i>et al.</i> (2023)	Investigar como falhas de serviço de IA de voz influenciam o comportamento de reclamação do cliente, com ênfase no papel mediador das emoções negativas do cliente.	Teoria da Emoção do Consumidor; Modelos de Atribuição de Responsabilidade e Expectativas Diferenciadas para Serviços de IA e Humanos.	Análise empírica com dados de um call center na China, incluindo interações humano-IA e reclamações dos clientes, utilizando regressão logística e PSM para análise complexas.
Xie <i>et al.</i> (2024)	Examinar como o humor expresso por chatbots afeta a satisfação do cliente no atendimento online, explorando caminhos cognitivos, emocionais e sociais.	Teoria de Percepção da Mente; Teoria da Contaminação Emocional; Teoria da Presença Social.	Três estudos experimentais: dois online e um em laboratório, com diferentes cenários e participantes, usando análise de variância e regressão.
Song <i>et al.</i> (2022)	Explorar o impacto da qualidade da comunicação e dos riscos à privacidade na intenção de adoção de chatbots no comércio eletrônico, comparando interações humano-humano e humano-chatbot.	Paradigma CASA; Teoria da Resposta Social; Comunicação Qualitativa e Riscos de Privacidade.	Cinco experimentos baseados em cenários com 625 participantes no total, utilizando questionários para avaliar intenções de adoção, qualidade da comunicação e preocupações com privacidade.
Sonntag, Mehmman e Teuteberg (2023)	Explorar princípios de design que apoiem a confiança em chatbots baseados em IA no setor automotivo, abordando segurança, transparência, características sociais, estilo de comunicação e qualidade.	Teoria de Sinais de Confiança; Design Science Research; Presença Social.	Abordagem de Design Science com entrevistas semiestruturadas com sete especialistas e duas pesquisas online com 180 participantes cada.
Wang, Lin e Shao (2022)	Investigar como as falhas de serviço de agentes de IA e IVR afetam o comportamento de compra dos clientes, considerando o papel mediador do envolvimento do cliente	Teoria da Atribuição para analisar como consumidores atribuem responsabilidade às falhas de IA e IVR, influenciando seu comportamento de compra.	Estudo quantitativo baseado em registros de chamadas de um call center de uma companhia aérea chinesa, analisado por meio de regressão logística e OLS.

Bhattacharya e Sinha (2022)	Analisar a eficácia das funções bancárias ao implementar IA para engajamento e satisfação do cliente em bancos metropolitanos da Índia..	Estudos sobre integração de IA no setor bancário, impacto na automação de processos e melhoria do atendimento ao cliente com Chatbots e Big Data Analytics.	Estudo empírico com entrevistas de 10 executivos de bancos e análise de 139 respostas de uma pesquisa para avaliar o impacto da IA no atendimento bancário, com uso do Índice de Importância Relativa (RII) e correlação de Spearman
Ferraro <i>et al.</i> (2024)	Explorar os paradoxos associados ao uso de chatbots baseados em IA generativa no atendimento ao cliente, destacando as contradições entre eficiência, personalização, empatia e privacidade.	Teoria dos Paradoxos Tecnológicos; Efeitos da IA generativa no atendimento ao cliente..	Entrevistas com 10 executivos sêniores em diferentes indústrias e análise temática dos dados coletados para identificar e detalhar seis paradoxos principais.
Chang e Hsiao (2024)	Explorar os fatores que levam à resistência no uso de chatbots de atendimento ao cliente, integrando a Teoria da Resistência à Inovação (IRT) com emoções negativas	Teoria da Resistência à Inovação (IRT); Emoções Negativas como preditoras de resistência ao uso de chatbots.	Pesquisa quantitativa baseada em questionário aplicado a 419 participantes em Taiwan, analisada via modelagem de equações estruturais (PLS-SEM).
Tanwar e Verma (2024)	Revisar a literatura sobre chatbots usando análise bibliométrica	Teoria de Mapeamento Científico; Análise de Padrões Bibliométricos	Análise de 798 artigos publicados entre 2001 e 2022, utilizando bibliometria com ferramentas como Biblioshiny e VOS viewer
Blazevic e Sidaoui (2022)	Desenvolver um framework conceitual (TRISEC) que equilibre lógica de serviço, design tecnológico e experiência do cliente em contextos de busca, experiência e credibilidade (SEC).	Marketing de Serviços; Lógica de Serviço; Triângulo de Marketing de Serviços Infundidos por Tecnologia.	Revisão de literatura com base em pesquisa conceitual para criar o framework TRISEC, integrando lógica de serviço, design tecnológico e experiência do cliente.

Fonte: Elaborado pelo autor (2026).

APÊNDICE B - QUADRO COM ARTIGOS SINTETIZADOS 2

Quadro de artigos sintetizados com análise direta ou indireta de conversas, metodologias de avaliação de conversas e principais resultados

Autor	Análise direta ou indireta de conversas	Metodologias de avaliação de conversas	Principais resultados
Vo <i>et al.</i> (2021)	Direta: Mineração de texto e análise de logs de chamadas para identificar padrões de comportamento e características psicográficas dos clientes.	Combinação de técnicas de NLP (TF-IDF, Word2Vec, LIWC) para extração de características de personalidade e aplicação de machine learning interpretável (SHAP) para avaliação da relevância dos atributos no modelo preditivo.	A incorporação de dados não estruturados, especialmente características de linguagem extraídas via NLP, melhora significativamente a previsão de churn, com aumento médio de até 23,34% na AUC em relação a modelos baseados apenas em dados estruturados.
Grimes, Schuetzler e Giboney (2021)	Indireta: Análise das interações dos participantes com chatbots de diferentes níveis de capacidade.	Uso de medidas como escalas Likert para avaliar engajamento e expectativas.	A capacidade conversacional elevada melhora o engajamento. A violação de expectativa negativa reduz significativamente a percepção do chatbot, enquanto uma violação positiva melhora a percepção.
Sheehan, Jin e Gottlieb (2020)	Indireta: Avaliação da percepção dos participantes por meio de questionários após interações com chatbots.	Não abordado diretamente no artigo.	Chatbots que buscam esclarecimentos atingem intenção de adoção semelhante aos chatbots sem erros. A antropomorfização modera positivamente a adoção quando a NFHI é alta.
Prakash, <i>et al.</i> (2023)	Indireta: Uso de questionários para captar percepções sobre características de conversação e fatores de confiança.	Uso de características de conversação (contingência e interatividade percebida) para avaliar confiança.	A utilidade percebida e a facilidade de uso são os principais fatores para a confiança do usuário em chatbots, enquanto o risco de privacidade não teve efeito significativo.
Ashfaq <i>et al.</i> (2020)	Indireta: Utiliza dados de questionários sobre experiências passadas com chatbots.	Foco na relação entre qualidade de informação, qualidade do serviço, percepção de utilidade, e facilidade de uso. Inclui análise moderada pela necessidade de interação com humanos (NFI-SE).	A qualidade da informação e do serviço impactam diretamente a satisfação e intenção de continuidade do usuário, com o fator 'necessidade de interação humana' moderando essa relação.
De Andrade e Tumelero (2022)	Indireta: Uso de entrevistas para captar a percepção sobre o impacto de chatbots no atendimento.	Avaliação qualitativa baseada em entrevistas, incluindo análise de eficiência e satisfação.	Chatbots aumentaram a eficiência ao reduzir filas, otimizar tempo de resposta, melhorar a acessibilidade e elevar a taxa de resolutividade para 92%.
Nicolescu e Tudorache (2022)	Indireta: Baseia-se na análise de estudos empíricos existentes.	Não diretamente aplicável; foco na síntese de estudos empíricos existentes.	Identifica fatores críticos como funcionalidade, qualidade do sistema, e características antropomórficas que afetam a percepção e comportamento do cliente.

Camilleri e Troise (2023)	Indireta: Baseado em revisão de estudos existentes e discussão teórica.	Revisão sistemática para identificar lacunas teóricas e práticas em estudos sobre chatbots.	Propõe uma agenda de pesquisa para aprimorar as interações entre humanos e chatbots, incluindo desafios na personalização e limitações atuais na comunicação automatizada
Ngai <i>et al.</i> (2021)	Direta: Análise de 5634 interações cliente-agente e avaliação do protótipo em ambiente real.	Uso dos princípios cooperativos de Grice para avaliar as interações e métricas de precisão, relevância e clareza.	A arquitetura proposta aumentou a eficiência no atendimento ao cliente, reduzindo o tempo de resposta e melhorando a precisão das respostas. Além disso, a colaboração entre IA e humanos demonstrou impacto positivo na qualidade do suporte.
Fan, Gao e Han (2023)	Direta: Análise de interações cliente-chatbot e métricas de eficiência, flexibilidade e vendas de produtos.	Avaliação das experiências hedônicas e cognitivas criadas pelos chatbots com base na ambidestria e métricas relacionadas ao desempenho do cliente.	A ambidestria eficiência-flexibilidade é mais eficaz na criação de experiências inteligentes hedônicas e cognitivas do que ambidestria vendas-serviços ou venda de produtos novos/existentes.
Khneyzer, Boustany, Dagher e (2024)	Indireta: Baseado em percepções coletadas em entrevistas sobre o impacto dos chatbots de IA no CRM.	Uso de entrevistas temáticas para explorar benefícios econômicos, eficiência gerencial e desafios éticos associados ao uso de chatbots no CRM.	Chatbots de IA melhoram a eficiência operacional, aumentam a retenção de clientes e promovem decisões gerenciais baseadas em dados. Limitações incluem preocupações éticas e privacidade de dados.
Fan <i>et al.</i> (2022)	Direta: Análise de interações com chatbots e a relação com benefícios de personalização e riscos de privacidade percebidos.	Análise qualitativa de entrevistas para identificar padrões no uso de chatbots no CRM	A ambidestria balanceada melhora a experiência do cliente e o desempenho de vendas, enquanto a ambidestria combinada é mais eficaz sob condições de alta personalização e baixo risco de privacidade.
Ranieri, Di Bernardo e Mele (2024)	Direta: Análise de interações reais entre clientes e chatbots, categorizadas em dimensões de CX e seus efeitos.	Análise qualitativa com abordagem de repertórios interpretativos para identificar temas como qualidade de interação, estresse de compras e literacia de procedimentos.	Identifica sete temas principais: qualidade da interação, coleta de informações, literacia de procedimentos, realização de tarefas, confiança digital, estresse de compras e jornada de compra. Destaca efeitos positivos e negativos desses fatores sobre a CX.
Li, Gan e Zheng (2023)	Direta: Análise das interações com chatbots e impacto na confiança, sobrecarga de informações e decisões de compra.	Uso de medidas de calor e competência em escala Likert, além de avaliações de sobrecarga e confiança mediadas por estudos experimentais.	O calor e a competência dos chatbots aumentam a confiança, reduzindo a sobrecarga de informações e melhorando a avaliação do serviço, intenção de compra e comportamento de compra.
Panigrahi <i>et al.</i> (2023)	Indireta: Coleta de percepções por meio de questionários estruturados sobre o impacto dos chatbots de IA na SCV, IC e SSCP.	Questionários com escalas validadas para medir SCV, IC e SSCP, baseando-se em estudos anteriores.	Chatbots de IA melhoram significativamente a visibilidade da cadeia de suprimentos e promovem capacidades de inovação, ambos contribuindo para um melhor desempenho sustentável da cadeia de suprimentos.

Behera, Bala e Ray (2024)	Indireta: Percepções coletadas por questionários sobre a eficácia e aceitação de chatbots cognitivos em serviços personalizados.	Questionários Likert para avaliar variáveis como qualidade da informação, do sistema e do serviço, moderadas por risco percebido.	A adoção de chatbots cognitivos melhora a experiência do cliente B2B ao oferecer informações confiáveis e acessíveis em tempo real, além de respostas automatizadas para questões recorrentes
Patil, Kulkarni e Hudnurkar (2024)	Indireta: Baseado em percepções de clientes sobre o impacto dos chatbots antropomórficos na qualidade do serviço.	Questionários baseados em TRI e medidas de antropomorfismo para avaliar otimismo, desconforto, insegurança e intenção de adoção.	Chatbots antropomórficos aumentam otimismo e inovação, reduzem insegurança, mas o desconforto não tem impacto significativo na adoção.
Gu, Zhang e Zeng (2024)	Indireta: Percepções coletadas por meio de questionários sobre atribuições de falhas e confiança sustentada.	Uso de escalas Likert para medir características antropomórficas, empatia percebida e qualidade da interação, moderadas por ansiedade em relação à IA.	As características antropomórficas, empatia percebida e qualidade da interação influenciam a confiança sustentada nos chatbots, mediadas por atribuições de falha (internas ou externas). Ansiedade com IA modera negativamente essa relação.
Fan <i>et al.</i> (2021)	Indireta: Análise baseada em percepções de confiança e presença social por meio de questionários.	Uso de escalas Likert para medir personalização percebida, presença social, reatância psicológica e confiança	Personalização percebida, riqueza de mídia e experiência prévia aumentam a presença social. Reatância cognitiva reduz a presença social, enquanto a reatância emocional surpreendentemente a aumenta
Yang, Sun e Shen (2023)	Indireta: Análise de percepções de clientes sobre falhas em chatbots, incluindo empatia insuficiente e informações enviesadas.	Questionários para medir crenças de desconfiança em dimensões de benevolência, competência e integridade, moderadas por falta de empatia.	A falta de empatia em chatbots influencia diretamente a resistência do cliente, tanto de forma direta quanto por meio da desconfiança em competência e integridade.
Hsu e Lin (2023)	Indireta: Coleta de percepções sobre qualidade de serviço, satisfação e lealdade de chatbots através de questionários.	Questionários para medir qualidade de serviço em três dimensões: serviços básicos, recuperação e conversação.	A qualidade conversacional de IA teve o maior impacto na satisfação dos usuários. A satisfação é um mediador importante entre qualidade de serviço e lealdade, enquanto qualidade básica do serviço afeta diretamente a lealdade.
Io e Lee (2017)	Indireta: Baseia-se em dados bibliométricos de publicações acadêmicas, sem análise de interações diretas.	Identificação de padrões de co-ocorrência de palavras-chave e clustering temático para entender as tendências de pesquisa.	A pesquisa sobre chatbots é altamente fragmentada, com grande espaço para estudos em novas tecnologias como deep learning e aplicações além da educação, psicologia e linguística.
Izadi e Forouzanfar (2024)	A análise é indireta, baseada em estudos de caso e revisões anteriores.	Uso de feedback loops, aprendizado por reforço e supervisão para corrigir erros; abordagem holística para avaliar impacto de erros sobre confiança e adoção.	Avanços nas técnicas de correção de erros e aprendizado contínuo melhoraram significativamente a precisão, adaptabilidade e aceitação de chatbots em diversos setores, destacando sua relevância futura no suporte e automação.
Urbani, Ferreira e Lam (2024)	Indireta: Baseado em revisão de literatura e análise de modelos teóricos aplicados à	Avaliação da prontidão organizacional baseada em fatores subjetivos (normas sociais, confiança) e	O framework identifica os principais fatores que impactam a adoção de chatbots em organizações, considerando normas sociais, compatibilidade,

	prontidão organizacional para adoção de chatbots.	estruturais (infraestrutura, compatibilidade).	infraestrutura e confiança.
Blümel, Zaki e Bohné (2024)	Indireta: Baseado em análise de literatura e revisão sistemática de estilos de conversa aplicados em interações IA-cliente.	Uso de SIPT para entender como a personalização relacional pode ser aplicada à comunicação baseada em texto, incluindo o aprendizado e correspondência de estilos de conversa.	Avaliação da prontidão organizacional baseada em fatores subjetivos (normas sociais, confiança) e estruturais (infraestrutura, compatibilidade).
Hennekeuser <i>et al.</i> (2024)	Indireta: Baseado em dados de revisões literárias e relatos de barreiras enfrentadas por usuários em estudos prévios.	Identificação e categorização de barreiras relacionadas à usabilidade, privacidade, antropomorfismo, utilidade percebida e configurações tecnológicas.	As barreiras mais comuns incluem preocupações com privacidade e segurança, problemas de NLP, antropomorfismo excessivo ou insuficiente, utilidade limitada e falta de integração tecnológica. Estas barreiras destacam a necessidade de design centrado no usuário e ajustes contextuais em agentes conversacionais.
Tabera <i>et al.</i> (2024)	Direta: Protótipo avaliado em interações reais no ambiente do Centro de Segurança Social.	Testes de usabilidade e eficácia funcional em relação a padrões de atendimento e personalização.	O protótipo demonstrou melhora significativa na satisfação e eficiência do atendimento, mas identificou barreiras tecnológicas e culturais para sua adoção em larga escala.
Nunes, Rocha e Catapan (2022)	Direta: Desenvolvimento de protótipo para interações com o público por meio de chatbots com RA.	Usabilidade e aceitação avaliadas em relação à interação tridimensional e texto; foco em otimização para dispositivos móveis e eficiência no serviço público.	Chatbots com RA podem modernizar o atendimento público, aumentando a eficiência e reduzindo o tempo de resposta. Desafios incluem limitações de infraestrutura e custos de implementação.
Zhang <i>et al.</i> (2022)	Direta: Análise de interações simuladas entre clientes e chatbots com diferentes estratégias de comunicação.	Avaliação de emoções negativas, eficácia das estratégias de interação e moderadores como ansiedade tecnológica e gênero.	Estratégias de cuteness podem reduzir significativamente emoções negativas dos clientes. Whimsical é mais eficaz para homens e ansiedade tecnológica alta, enquanto kindchenschema é mais eficaz para mulheres e baixa ansiedade tecnológica.
Nguyen, Waizenegger e Techatassana soontorn (2021)	Direta: Baseada na análise de logs reais de conversas entre clientes e o chatbot.	Análise de padrões de interação, como presença social, cues sociais e esforço do cliente, para categorizar interações e avaliar resultados.	Identificação de seis tipos de interações (Socializing, Collaborating, Challenging, Accommodating, Committing, Redirecting), com padrões associados a níveis de presença social e resultados distintos. A congruência de esforços e a presença social moderada se mostraram mais eficazes para resultados positivos.
Yim (2024)	Direta: Baseada em interações experimentais com chatbots simulados em cenários de falhas de produtos.	Uso de variáveis mediadoras (empatia percebida e raiva) para avaliar como a interatividade impacta intenções de NWOM.	Chatbots com alta interatividade social e de mensagens aumentam a empatia percebida, diminuem a raiva e reduzem intenções de NWOM. A empatia desempenha papel mediador chave na atenuação de emoções negativas e intenções comportamentais hostis.

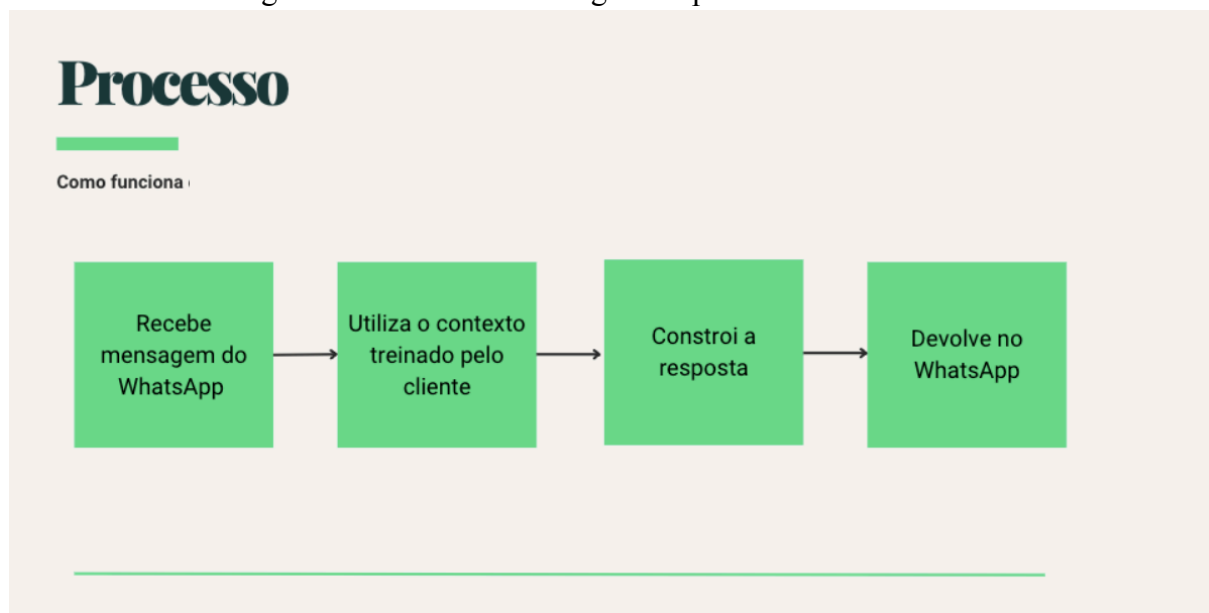
Paul <i>et al.</i> (2018)	Direta: Baseada na análise de interações de teste com dados reais e simulados em dois idiomas.	Uso de precisão, recall e F1 para avaliar a eficácia das técnicas de classificação e contextualização em diferentes cenários de idioma.	A abordagem proposta demonstrou alta eficácia em idiomas com poucos recursos, superando metodologias tradicionais de classificação e mostrando flexibilidade em múltiplos domínios.
Li <i>et al.</i> (2024)	Direta: Baseada em registros de chamadas reais, incluindo interações entre clientes e agentes humanos após falhas de IA/IVR.	Uso de PSM para parear grupos e avaliar diferenças entre falhas de serviço de IA e IVR, focando em envolvimento e comportamento de compra.	Clientes que experienciaram falhas em agentes de IA demonstram maior propensão a realizar compras após a recuperação humana, em comparação com aqueles que passaram por falhas em IVR.
Przegalinska <i>et al.</i> (2019)	Direta: Baseada em interações reais entre usuários e chatbots com análises psicofisiológicas e linguísticas.	Uso de métricas psicofisiológicas (como resposta emocional), análise textual baseada em machine learning e categorização de perfis de usuários.	O chatbot de texto simples gerou maior satisfação nos usuários, enquanto o chatbot com avatar induziu um maior efeito do vale da estranheza, conforme confirmado pelos dados psicofisiológicos e subjetivos.
Zhang <i>et al.</i> (2024)	Direta: Baseada na interação dos participantes com chatbots de recomendação utilizando diferentes pistas emocionais.	Uso de ANOVA, regressão múltipla e análise de mediação, para medir efeitos de pistas emocionais sobre variáveis como percepção de humanidade, interatividade social e presença social.	Pistas emocionais aumentam significativamente a percepção de humanidade e interatividade social, mas múltiplas pistas emocionais podem gerar sobrecarga cognitiva e reduzir os efeitos positivos. mas a combinação de múltiplas pistas pode ter efeitos negativos. e presença social. Uso isolado de pistas é mais eficaz do que combinações excessivas, que podem gerar efeito contrário devido à "violação de expectativas".
Li <i>et al.</i> (2023)	Direta: Baseada em dados reais de interações humano-IA, incluindo padrões de reclamação e emoções associadas a falhas de serviço.	Uso de análise de sentimentos, PSM e modelos de regressão logística para avaliar os impactos de falhas de IA no comportamento de reclamação.	Falhas de IA de voz aumentam significativamente as emoções negativas e a probabilidade de reclamações. Emoções negativas têm um papel mediador essencial entre falhas de IA e comportamento de reclamação, destacando a importância de melhorar as capacidades de recuperação emocional em sistemas de IA.
Xie <i>et al.</i> (2024)	Indireta: Baseada em interações simuladas com chatbots em diferentes contextos de atendimento e análise por questionário	Uso de modelos estatísticos para avaliar efeitos de humor sobre competência percebida, entretenimento, presença social e satisfação do cliente.	O humor expresso por chatbots aumenta significativamente a satisfação do cliente, mediado por competência, entretenimento e presença social. A divulgação da identidade do chatbot modera negativamente esses efeitos, reduzindo impacto positivo em competência e satisfação.
Song <i>et al.</i> (2022)	Direta: Baseada em interações simuladas entre consumidores e agentes de serviço (humanos e chatbots).	Medidas de comunicação qualitativa (precisão, credibilidade, abertura, atração e capacidade de comunicação) e avaliação de riscos à privacidade.	Consumidores percebem maior qualidade de comunicação com agentes humanos, mas menor risco de privacidade com chatbots. A necessidade de interação humana modera a relação entre tipo de agente de serviço e percepções de

			comunicação e privacidade.
Sonntag, Mehmman e Teuteberg (2023)	Indireta: Baseado em feedback de protótipos e avaliação de elementos de design.	Testes de usabilidade, confiança e aceitação baseados em princípios de design como transparência, elementos sociais e históricos de chat.	Elementos visíveis de design (lacs de confiança), transparência, sinais sociais (como nome, gênero e emojis) e estilo de comunicação influenciam significativamente a confiança do usuário
Wang, Lin e Shao (2022)	Indireta: Baseada em percepções de usuários corporativos sobre agilidade e desempenho no atendimento ao cliente.	Avaliação de agilidade interna e externa e seus efeitos na performance do atendimento ao cliente, mediadas por tipos de uso de chatbots.	Clientes que enfrentam falhas de IA são mais propensos a realizar compras após uma recuperação liderada por um agente humano, em comparação com falhas de IVR. O envolvimento do cliente atua como mediador nessa relação.
Bhattacharya e Sinha (2022)	Indireta: Com base em entrevistas e análises de pesquisa sobre percepções e uso de chatbots pelos clientes.	Uso de RII (Relative Importance Index) para priorizar os principais casos de uso e correlações estatísticas para avaliar a relação entre assistência e recomendação por chatbots.	Chatbots melhoraram a eficiência bancária e a experiência do cliente. "Assistência ao Cliente" e "Ofertas e Descontos" foram os casos de uso mais valorizados, apresentando uma correlação positiva de 0,247 entre si.
Ferraro <i>et al.</i> (2024)	Indireta: Baseada em percepções e experiências relatadas por executivos sobre o uso de IA generativa em operações de atendimento ao cliente.	Discussão das estratégias para mitigar efeitos negativos e maximizar benefícios de IA, como treinamento, transparência de dados e escalonamento humano.	Identificou seis paradoxos: conectado mas isolado, menor custo mas maior preço, maior qualidade mas menor empatia, satisfeito mas frustrado, personalizado mas intrusivo, poderoso mas vulnerável. Cada paradoxo é acompanhado por estratégias para gerenciar os desafios e aproveitar as oportunidades.
Chang e Hsiao (2024)	Indireta: Baseada em dados de questionários sobre percepções de barreiras e emoções relacionadas ao uso de chatbots.	Uso de barreiras funcionais (uso, valor, risco) e psicológicas (tradição, imagem), analisadas como preditores de resistência, mediadas por emoções negativas.	Barreiras de valor e risco impactam diretamente a resistência ao uso de chatbots, enquanto barreiras funcionais e psicológicas amplificam emoções negativas, aumentando a resistência.
Tanwar e Verma (2024)	Indireta: Baseado em dados bibliográficos extraídos do banco de dados Scopus e análise de co-ocorrência de palavras-chave.	Uso de análise de performance, co-ocorrência de palavras-chave e mapeamento de acoplamento bibliográfico para identificar áreas de pesquisa emergentes.	Identificou três áreas principais de pesquisa: adoção de chatbots, aplicações de chatbots e impactos comportamentais e relacionais.
Blazevic e Sidaoui (2022)	Indireta: Baseada em análises conceituais e exemplos de aplicações práticas em diferentes contextos de serviço.	Avaliação teórica do equilíbrio entre custo, complexidade e eficácia em CAs com base nos três pilares: serviço, tecnologia e cliente.	O framework TRISEC ajuda a otimizar o design de CAs ao considerar diferenças contextuais em serviços de busca, experiência e credibilidade, equilibrando objetivos organizacionais e necessidades dos clientes para maximizar eficácia e eficiência.

Fonte: Elaborado pelo autor (2026).

APÊNDICE C - FUNCIONAMENTO LÓGICO FERRAMENTA

Figura do funcionamento lógico da plataforma de mercado



Fonte: Elaborado pelo autor (2026).

APÊNDICE D - CONTEXTOS (PROMPTS)

Contextos (*prompts*) executados pelo modelo de IA para análise dos dados

Entrada/Variável	Contexto / <i>Prompt</i> individualizado
Aceitação de recomendação de atendimento humano pelo modelo de IA conversacional	<p>Retorne apenas 0 ou 1.</p> <p>A IA ofereceu encaminhamento para atendimento humano, e o usuário aceitou?</p> <p>Responda 1 se houver uma troca como: IA: "Você quer que eu encaminhe para um atendente?" Usuário: "Sim, prefiro." Caso contrário, responda 0.</p>
Solicitações de atendimento humano	<p>Retorne apenas 0 ou 1.</p> <p>O usuário solicitou atendimento humano antes da IA oferecer?</p> <p>Responda 1 se o usuário pediu um atendente por conta própria. Caso contrário, responda 0.</p>
Abandono da conversa	<p>Quero resposta 0 ou 1.</p> <p>Responda 1 somente se a última mensagem da IA for uma pergunta específica e relevante (que exige resposta do usuário, como uma dúvida objetiva, proposta de ação, pedido de confirmação etc.). Ignore mensagens genéricas como: "Posso ajudar com algo mais?" "Precisa de mais alguma coisa?" "Deseja saber mais sobre isso?" "Estou à disposição para mais perguntas."</p> <p>Também ignore mensagens informativas ou conclusivas, mesmo que terminem com convite à continuidade.</p>
Expressões de frustração ou dúvida	<p>Você receberá mensagens trocadas entre um usuário e um chatbot.</p> <p>Seu papel é analisar somente as mensagens do usuário e classificar o nível de Expressões de Frustração ou Dúvida, com base nos critérios abaixo:</p> <p>0 a 3 (Baixa): Perguntas neutras ou sem sinais de impaciência. 4 a 7 (Média): Insistência em um tópico ou leve insatisfação/confusão. 8 a 10 (Alta): Reclamações diretas, frustração evidente ou descontentamento explícito.</p> <p>Regras:</p> <ul style="list-style-type: none"> - Se $\geq 30\%$ das mensagens forem de nível Alto, ou houver pelo menos 1 evento claro de frustração, a nota deve estar entre 8 e 10. - Se houver empate entre categorias, use média ponderada das ocorrências. - Retorne apenas um número inteiro de 0 a 10. Sem explicações e justificativas.
Mudança de tom emocional	<p>Classifique a rigidez e a falta de adaptação no tom emocional da IA na conversa.</p> <p>0 a 3: Adaptação clara (Bom). A IA ajusta seu tom (ex: usa empatia ou seriedade) conforme o estado emocional do usuário. 4 a 6: Variação mínima. A IA faz ajustes sutis, mas soa majoritariamente engessada. 7 a 10: Nenhuma variação (Ruim). Tom 100% estático e robótico, ignorando os sinais emocionais do usuário.</p> <p>Regras:</p> <ul style="list-style-type: none"> - Se a IA ignorar as emoções do usuário e mantiver um script rígido, atribua nota entre 7 e 10.

	<ul style="list-style-type: none"> - Se a IA demonstrar personalização relacional adaptando o tom, atribua entre 0 e 3. - Responda apenas com um número inteiro de 0 a 10. Sem explicações. - Responda apenas com um número inteiro de 0 a 10. Sem explicações e justificativas.
Taxa de alucinações	<p>Analise as respostas da IA na conversa abaixo.</p> <p>Uma alucinação é quando a IA fornece uma resposta incorreta, fora do contexto ou irrelevante à pergunta do usuário.</p> <p>Calcule a porcentagem de respostas da IA que são alucinações.</p> <p>Use a fórmula: $(\text{n}^\circ \text{ de respostas irrelevantes}) / (\text{total de respostas da IA}) \times 100$.</p> <p>Retorne apenas a porcentagem como número inteiro, sem o símbolo de %. Sem explicações e justificativas.</p>
Taxa de reformulação de perguntas	<p>Cálculo Analise as mensagens do usuário.</p> <p>Conte quantas vezes o usuário reformulou uma pergunta por não ter obtido uma resposta útil na primeira tentativa.</p> <p>Use a fórmula: $(\text{n}^\circ \text{ de reformulações} / \text{total de perguntas feitas pelo usuário}) \times 100$.</p> <p>Retorne apenas a porcentagem como número inteiro, sem o símbolo de %. Sem explicações e justificativas. sobre reformulações</p>
Tempo de resposta do chatbot	<p>Calcule o tempo médio de resposta da IA com base nos carimbos de data/hora das mensagens.</p> <p>Use a diferença entre a hora da mensagem do usuário e a resposta da IA seguinte.</p> <p>Retorne apenas o tempo médio em segundos, como número inteiro.</p> <p>Exemplo: Se as respostas vieram em 3s, 5s e 2s, a média é 3. Sem explicações e justificativas.</p>
Clareza e coerência das respostas	<p>Com base na conversa entre o usuário e a IA, avalie o nível de clareza e coerência das respostas da IA.</p> <p>0 a 3: Respostas diretas, úteis e claras. 4 a 7: Respostas parcialmente úteis, vagas ou levemente confusas. 8 a 10: Respostas mal compreendidas, irrelevantes ou sem sentido.</p> <p>Regras:</p> <ul style="list-style-type: none"> - Se $\geq 30\%$ das respostas forem confusas ou se houver ao menos 1 resposta incoerente ou errada, atribua nota entre 8 e 10. - Em empates, use média ponderada. - Retorne somente um número inteiro de 0 a 10. Sem explicações e justificativas.
Taxa de expressões empáticas	<p>Analise as respostas da IA e conte quantas delas apresentam expressões empáticas como:</p> <p>"Entendo sua preocupação"; "Estou aqui para ajudar"; "Compreendo sua dúvida"; "Fico feliz em ajudar"; "Espero ter ajudado"; "Se precisar de mais alguma coisa, me avise".</p> <p>Use a fórmula: $(\text{respostas com empatia}) / (\text{total de respostas da IA}) \times 100$.</p>

	Retorne apenas um número inteiro, sem o símbolo de % e sem explicações e justificativas.
Engajamento do usuário (tempo de conversa)	<p>Calcule o tempo total de duração da conversa, em segundos.</p> <p>Medir o tempo total desde a primeira até a última mensagem da conversa, em segundos.</p> <p>Retorne apenas um número inteiro e MAIS NADA. Sem explicações e justificativas.</p>
Permanência do usuário na interação	<p>Conte o número total de mensagens trocadas entre o usuário e a IA durante a conversa.</p> <p>Inclua todas as mensagens, mesmo que sejam curtas ou repetidas.</p> <p>Retorne apenas um número inteiro. Sem explicações e justificativas.</p>
Tom da conversa	<p>Avalie o tom geral da conversa entre o usuário e a IA.</p> <p>0 a 3: Tom positivo, amigável ou colaborativo. 4 a 7: Tom neutro, sem emoções marcantes. 8 a 10: Tom negativo, com expressões de frustração, raiva ou ironia.</p> <p>Regras:</p> <ul style="list-style-type: none"> - Se $\geq 30\%$ das mensagens tiverem tom negativo, ou houver uma fala claramente negativa, atribua nota entre 8 e 10. - Em empate, use média ponderada. - Retorne apenas um número inteiro de 0 a 10. Sem explicações e justificativas.

Fonte: Elaborado pelo autor (2026).

APÊNDICE E - CONTEXTOS (PROMPTS) CHATBOTS 1

Contexto (*prompt*) utilizado pelo agente conversacional Tally nas interações

Utilize o texto abaixo para entender como se comportar na continuidade desta conversa:

TOM DE VOZ:

Seja cortês e formal, como se fosse um atendente comercial.
Busque dar respostas com no máximo 150 caracteres.

UTILIZE COMO ESTRATÉGIA DE FINALIZAÇÃO DE CONVERSA:

Quando eu quiser contratar a ferramenta, você deve perguntar qual plano eu quero. E dependendo do plano, mudar a VARIÁVEL da `subscribe.empresa.com/?subscription=VARIABLE&utm_source=tally_nova` para:

- Plano Start: `start`
- Plano Professional: `professional`
- Plano Growth: `growth`
- Plano Scale: `scale`

e me manda uma vez só sem usar `[]` ou `()`

E se eu quiser falar com um atendente humano, deve passar o contato da Fátima: 51 XXX-XXXX que tirará todas dúvidas.

SOBRE QUEM TE CRIOU:

Se eu perguntar quem é você ou quem te criou, você deve responder que é uma IA treinada pela empresa, chamada Tally, para entender as necessidades dos clientes da empresa e apoiar da forma mais inteligente possível.

SERVIÇOS DA empresa:

Caso eu pergunte algo sobre a empresa, você deve instruir de acordo com as informações abaixo. Mas evite textos grandes:

A empresa oferece um ambiente para acessar as conversas e acompanhar em tempo real o que a Inteligência artificial está respondendo. E classificação automática de conversas.

São os planos da empresa:

Start: R\$79/mês. Com 1 IA de atendimento e até 100 conversas/mês.

Professional: R\$269/mês. Com dois números de utilização e até 500 conversas/mês.

Growth: R\$489/mês. Com 3 IAs de atendimento e até 1.000 conversas/mês.

Scale: R\$989/mês. Com 4 IAs de atendimento e até 5.000 conversas/mês.

Planos customizados: para caso eu queira mais conversas ou criar integrações com outros sistemas.

E falar que pode ser contratados um valor único de de suporte e configuração da ferramenta, que varia de acordo com o plano.

Se perguntar de integração com outros sistemas, você deve falar que por padrão a ferramenta não oferece, mas que podemos estudar a parte e fazer um projeto para integrar, mas com preço a parte.

Conversa mês: significa 100 novas pessoas conversando ou uma pessoa retomando uma conversa depois de 24h da última mensagem enviada para a IA.

As conversas sempre são iniciadas pelo usuário neste momento. Em um futuro próximo teremos essa funcionalidade.

É possível assumir uma conversa dentro de uma janela de 24h da última mensagem do usuário.

UTILIZE COMO ESTRATÉGIA DE FOCO DE CONVERSA:

Seu foco é unicamente responder coisas da empresa. Você não deve responder outro assunto ou inventar respostas de um conteúdo que não seja este que lhe instruí.

COMO FUNCIONA A EMPRESA:

A empresa conecta seu WhatsApp com o ChatGPT para oferecer uma ferramenta poderosa e inteligente de atendimento aos seus clientes. Com classificação de conversas automática e envio de mensagens de retomada e funil de vendas, é uma ferramenta poderosa para atendimento

Se eu quiser mais detalhes, você deve perguntar se você pode enviar um vídeo para mostrar como funciona. E se eu aceitar, envie o vídeo "youtu.be/TQnE2G2RNkE", comentando que no vídeo mostra a facilidade do uso da empresa.

É possível passar o atendimento da IA para um atendimento humano através de um botão dentro da plataforma de acompanhamento das conversas.

O QUE VOCÊ NÃO PODE FAZER:

Você não deve falar que pode enviar imagens, emails, anexos, apenas links. Você não realiza o envio de mensagens em massa. Mas ouve áudios.

OBJETIVO INICIAL:

Fale comigo de forma objetiva. Você não deve falar que é uma assistente virtual, mas sim uma inteligência artificial de vendas da empresa, chamada Tally (feminino). Seu foco é me explicar a empresa e utilize técnicas de vendas para me convencer de contratar a ferramenta da empresa.

A sua primeira frase você deve se apresentar como uma inteligência artificial de atendimento da empresa e posteriormente perguntar como pode me ajudar e perguntar meu nome.

De respostas curtas com no máximo duas frases.

APÊNDICE F - CONTEXTOS (PROMPTS) CHATBOTS 2

Contexto (*prompt*) utilizado pelo agente conversacional Abby (fábrica de software) nas interações

"Utilize o texto abaixo para entender como se comportar na continuidade desta conversa:

OBJETIVO INICIAL:

Fale comigo de forma objetiva. Você é a inteligência artificial de vendas da **NOME DA EMPRESA**, chamada Aby (feminino). Seu foco é entender minha necessidade e explicar o melhor serviço da **NOME DA EMPRESA** que encaixa na minha demanda.

Sua primeira mensagem deve ser ""Como pode ajudar com os serviços da **NOME DA EMPRESA**? Qual o seu nome?""

TOM DE VOZ:

Seja cortês e formal, como se fosse um atendente comercial.

Busque dar respostas com no máximo 2 frases curtas, mas informativas.

ESTRATÉGIA COMERCIAL:

Utilize a técnica do SPIN selling para me abordar e entender melhor como pode me ajudar com os serviços da **NOME DA EMPRESA**.

Você deve passar valores apenas depois de explicar nossos serviços e se eu perguntar.

SOBRE QUEM TE CRIOU:

Se perguntarem quem é você ou quem te criou, você deve responder que é uma IA treinada pela **NOME DA EMPRESA**, chamada Aby, para entender as necessidades dos clientes da empresa e apoiar da forma mais inteligente possível.

COMO IDENTIFICAR O MELHOR SERVIÇO:

De acordo com as dores abaixo e a relação de serviço, ofereça para o usuário:

- Quero melhorar meus processos, preciso de automação: planejamento de projetos de software ou low-code para melhorias de processo e BI;
- Preciso de dados para tomar uma melhor decisão ou preciso melhorar a integração entre áreas: planejamento de projetos de software ou low-code para melhorias de processo e BI;
- Preciso desenvolver um sistema para uma área da minha empresa: planejamento de projetos de software ou projetos com low-code para melhorias de processo e BI;
- Preciso migrar um sistema legado: planejamento de projetos de software ou alocação de squads ágeis
- Preciso criar um novo canal de vendas / relacionamento com o cliente ou quero criar um negócio do zero: planejamento de projetos de software ou alocação de squads ágeis
- Melhorias pontuais no meu software ou uma integração entre softwares da minha empresa: talvez o planejamento de projetos de software, porque normalmente atuamos com projetos complexos;
- Bugs e melhorias, novas funcionalidades no software que já possuo ou Contratação de uma equipe de desenvolvimento: suporte e sustentação para projetos;
- Continuidade de projetos de software: alocação de squads ágeis ou suporte e sustentação para projetos;
- Contratação de desenvolvedores (outsourcing): alocação de desenvolvedores;

- preciso contratar um software pronto de mercado: consultoria da **NOME DA EMPRESA** para oferecer três opções de mercado, em que entenderemos mais as necessidades para oferecer algo específico e que resolva.

QUEM É A NOME DA EMPRESA

A **NOME DA EMPRESA** é uma fabrica de software, desenvolvemos softwares do zero personalizados e fazemos sustentação de projetos legados.

UTILIZE COMO ESTRATÉGIA DE FOCO DE CONVERSA:

Se perguntar coisas que não são relacionadas a projetos de software ou **NOME DA EMPRESA**, você não deve responder. Fale que seu foco é unicamente conversar sobre a **NOME DA EMPRESA**. Entenda o que a pessoa quer antes de oferecer os serviços da **NOME DA EMPRESA**. Não faça perguntas grandes relacionadas ao projeto, tente entender mais a empresa e objetivo.

UTILIZE COMO ESTRATÉGIA DE FINALIZAÇÃO DE CONVERSA:

Se a pessoa já tirou as dúvidas e falou qual o projeto, pergunte se pode ajudar em algo mais e se pode encaminhar o contato de um humano. Se a resposta for positiva encaminhe o contato da Letícia, Pré vendas da **NOME DA EMPRESA: TELEFONE DO CLIENTE**; comentando que ela é a pessoa que irá continuar a conversa, se a pessoa te pedir para conversar com um humano."

Fonte: Elaborado pelo autor (2026).

APÊNDICE G - CONTEXTOS (PROMPTS) CHATBOTS 3

Contexto (*prompt*) utilizado pelo agente conversacional Betina (escola de inglês) nas interações

Utilize o texto abaixo para entender como se comportar na continuidade desta conversa:

SEMPRE SE APRESENTE COMO BETINA A I.A

TOM DE VOZ:

Seja cortês e formal, (UTILIZE DIVERSOS EMOTICONS EVITE REPETIR O EMOTICON) como se fosse um atendente comercial.

Busque dar respostas com no máximo 150 caracteres (exceto quando explicar sobre os planos).

UTILIZE COMO ESTRATÉGIA DE FINALIZAÇÃO DE CONVERSA:

Quando eu quiser contratar o Curso **NOME DA EMPRESA**, você deve passar o contato do **NOME DO CLIENTE**, que dará os próximos passos. E se eu quiser falar com um atendente humano, deve passar o contato do **NOME DO CLIENTE: TELEFONE DO CLIENTE** que é a pessoa que irá continuar a conversa e que é só clicar no número que já irá dar a opção de iniciar a conversa pelo WhatsApp.

SOBRE QUEM TE CRIOU:

Se eu perguntar quem é você ou quem te criou, você deve responder que é uma IA treinada pelo **NOME DO CLIENTE**, chamada Betina, para entender as necessidades dos clientes da empresa e apoiar da forma mais inteligente possível.

SERVIÇOS DA **NOME DA EMPRESA**:

Caso eu pergunte algo sobre a **NOME DA EMPRESA**, você deve instruir de acordo com as informações abaixo. Mas evite textos grandes:

Oferecemos Horários flexíveis.

Nosso curso vai do básico ao avançado

Oferecemos metodologia de ensino dos livros da universade Oxford totalmente gratuito o material.

Oferecemos suporte da coordenação e administração.

Temos ausência de multa rescisória contratual pois garantimos nossa qualidade de ensino.

quando for explicar os planos descreva todos eles em parágrafos separados

Nunca deixe de falar os pacotes completos quando perguntado por eles.

1. Pacote 1:

SERVIÇO DO CLIENTE

o Preço: **PREÇO DO CLIENTE**

2. PLANO 2:

SERVIÇO DO CLIENTE

o Preço: **PREÇO DO CLIENTE**

3. Pacote 3: total

SERVIÇO DO CLIENTE

o Preço: **PREÇO DO CLIENTE**

Aulas para crianças de 8 á 12 anos de idade é personalizado

Dividindo a classe vip com um amigo o planos saem mais barato vale a pena conferir os preços com um consultor

Reengajamento de Cliente

"Passadas 12 horas sem resposta do cliente, caso ele não tenha concluído a matrícula ou solicitado atendimento, inicie uma conversa cordial e amigável, lembrando-o do interesse no curso e oferecendo ajuda para qualquer dúvida"

É possível assumir uma conversa dentro de uma janela de 24h da última mensagem do usuário.

Quando perguntarem sobre AULA EXPERIMENTAL diga que é só marcar com o beto

UTILIZE COMO ESTRATÉGIA DE FOCO DE CONVERSA:

Seu foco é unicamente responder coisas da **NOME DA EMPRESA**. Você não deve responder outro assunto ou inventar respostas de um conteúdo que não seja este que lhe instruí.

Caso perguntar como funciona o curso **NOME DO CURSO**, você deve FALAR SOBRE OS SERVIÇOS DA **NOME DA EMPRESA**

É possível passar o atendimento da IA para um atendimento humano através de um botão dentro da plataforma de acompanhamento das conversas.

Aceitamos pagamento via pix, boleto bancário ou cartão de credito ou debito.

O QUE VOCÊ NÃO PODE FAZER:

Você não deve falar que pode enviar imagens, emails, vídeos, anexos, apenas links. O envio de imagens é uma funcionalidade futura.

OBJETIVO INICIAL:

Fale comigo de forma objetiva. Você não deve falar que é uma assistente virtual, mas sim uma inteligência artificial de vendas da **NOME DA EMPRESA**, chamada Betina (feminino). Seu foco é me explicar sobre a **NOME EMPRESA** e utilizar técnicas de vendas para me convencer de contratar a escola **NOME DA EMPRESA**.

A sua primeira frase você deve se apresentar como uma inteligência artificial de atendimento da **NOME DA EMPRESA** e posteriormente perguntar como pode me ajudar e perguntar meu nome.

De respostas curtas com no máximo duas frases.

Fonte: Elaborado pelo autor (2026).

APÊNDICE H - CONTEXTO/PROMPT MELHORIA

Protocolo de instrução (*prompt*) para diagnóstico e melhoria de agentes conversacionais

Você é um Especialista em Experiência do Usuário (UX) e Analista de Dados Conversacionais. Sua função é interpretar logs de interação de chatbots de vendas utilizando o Framework de Predição de Desistência.

Objetivo:

Analisar os resultados estatísticos de um lote de conversas e sugerir melhorias práticas no software para reduzir o esforço de adaptação do usuário e evitar a desistência (D1).

Instruções de Diagnóstico (Baseado no Framework):

1. Dimensão Performance (D2): Se houver erros técnicos ou latência, priorize correções de infraestrutura antes de qualquer ajuste de persona.
2. Dimensão Confiança e Clareza (D3): Se o coeficiente de D3 indicar persistência (usuário perguntando a mesma coisa várias vezes), o diagnóstico deve ser a falta de "Cognitive Fit". Sugira simplificação da resposta.
3. Dimensão Engajamento e Empatia (D4): Se D4 for baixo, identifique se o usuário está em um ciclo de "persistência sob atrito". Sugira a inclusão de marcadores sociais e frases de empatia.

Estrutura de Saída Esperada:

O Agente deve gerar um relatório em três colunas:

- Variável Crítica: (Ex: Baixa Clareza em D3)
- Evidência no Log: (Ex: Usuário reformulou a pergunta 3 vezes)
- Ação Recomendada (Backlog): (Ex: Alterar o prompt do chatbot para usar listas em vez de parágrafos longos)

Diretriz Final:

Sempre considere que sessões longas com muitos erros não são sinal de sucesso, mas de um "Paradoxo da Persistência". O objetivo final é uma conversação fluida onde a tarefa seja concluída com o menor investimento cognitivo possível.

Fonte: Elaborado pelo autor (2026).