



**UNIVERSIDADE FEDERAL DO TOCANTINS
CAMPUS UNIVERSITÁRIO DE PALMAS/TO
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENSINO EM CIÊNCIAS
E SAÚDE (PPGECS-UFT) MESTRADO ACADÊMICO**

SÓSTHENES OLIVEIRA LIMA

**INTELIGÊNCIA ARTIFICIAL E EDUCAÇÃO: ESTRATÉGIAS PARA AVALIAR
A EFICÁCIA DO CHATBOT EDUCACIONAL
EDUCACHAT**

Palmas, TO

2025

SÓSTHENES OLIVEIRA LIMA

**INTELIGÊNCIA ARTIFICIAL E EDUCAÇÃO: ESTRATÉGIAS PARA AVALIAR
A EFICÁCIA DO CHATBOT EDUCACIONAL
EDUCACHAT**

Dissertação apresentada ao Programa de Pós-Graduação em Ensino em Ciências e Saúde (área de concentração: Ensino em Ciência), como parte dos requisitos necessários para a obtenção do Título de Mestre em Ensino em Ciências e Saúde (PPGECS), da Universidade Federal do Tocantins (UFT).

Orientador: Prof. Dr. Janeisi de Lima Meira

Palmas, TO

2025

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema de Bibliotecas da Universidade Federal do Tocantins

- L732i Lima, Sósthenes Oliveira.
 Inteligência Artificial e Educação: Estratégias para Avaliar a
 Eficácia do Chatbot Educacional EducaChat. / Sósthenes Oliveira
 Lima. – Palmas, TO, 2025.
 161 f.

 Dissertação (Mestrado Acadêmico) - Universidade Federal do
 Tocantins – Câmpus Universitário de Palmas - Curso de Pós-
 Graduação (Mestrado) em Ensino em Ciências e Saúde, 2025.
 Orientador: Janeisi de Lima Meira

 1. Inteligência Artificial. 2. Chatbot Educacional. 3. Engajamento
 Discente. 4. Avaliação de Tecnologias Educacionais. I. Título
-
- CDD 372.35**

TODOS OS DIREITOS RESERVADOS – A reprodução total ou parcial, de qualquer forma ou por qualquer meio deste documento é autorizado desde que citada a fonte. A violação dos direitos do autor (Lei nº 9.610/98) é crime estabelecido pelo artigo 184 do Código Penal.

Elaborado pelo sistema de geração automática de ficha catalográfica da UFT com os dados fornecidos pelo(a) autor(a).

SÓSTHENES OLIVEIRA LIMA

**INTELIGÊNCIA ARTIFICIAL E EDUCAÇÃO: ESTRATÉGIAS PARA
AVALIAR A EFICÁCIA DO CHATBOT EDUCACIONAL EDUCACHAT**

Dissertação apresentada à UFT – Universidade Federal do Tocantins – Campus Universitário de Palmas-TO, Curso de Programa de Pós Graduação em Ensino em Ciências e Saúde, foi avaliado para à obtenção do título de Mestre em Ensino em Ciências e Saúde e aprovado em sua fase final pelo Orientador e pela Banca Examinadora.

Data de aprovação: 17/12/2025

Banca Examinadora:

Prof. Dr. Janeisi de Lima Meira, UFT

Prof. Dr. José Lauro Martins, UFT

Prof. Dr. Deive Barbosa Alves, FCAT

Dedico este trabalho a todos os professores do PPGECS, pelos conhecimentos, trocas e aprendizados adquiridos ao longo destes dois anos de trajetória. Dedico-o, de modo especial, ao meu Mestre e orientador, Prof. Dr. Janeisi de Lima Meira, que me acompanhou na fase final desse processo, incentivando o desenvolvimento deste trabalho, oferecendo palavras de encorajamento e orientação nos momentos oportunos. Serei eternamente grato.

AGRADECIMENTOS

A Deus primeiramente que me deu forças para concluir esta etapa em minha vida, pois só conquistamos algo a partir de muito esforço e dedicação. Foi um processo de bastante luta, em especial, neste último ano, período atípico a nível mundial e extremamente difícil decorrente da pandemia da Covid-19. As minhas filhas Yonnara Freitas Lima e Catherine Oliveira Barros, fontes mais lindas e preciosa de inspiração que Deus me deu. A minha esposa Sandra Erleise, por me apoiar em meus sonhos e objetivos. Aos meus pais, Silas Araújo Lima e Jacinones Pinto Oliveira por acreditarem em meu potencial e o meu irmão Silas Júnior (In memoriam). A minha turma especial do Mestrado em Ensino e Ciência e Saúde e aos amigos feitos, guardados e eternizados em meu coração.

“Um homem provido de papel,
lápiz e borracha, e sujeito a estrita
disciplina, é de fato uma máquina
universal". Alan Turing

RESUMO

Este estudo avaliou o desempenho técnico-operacional do chatbot educacional EducaChat, com base em métricas extraídas de registros de uso (logs) e nos critérios de qualidade do modelo ISO/IEC 25010, discutindo seu potencial como ferramenta de apoio a práticas pedagógicas mediadas por tecnologia. Trata-se de uma pesquisa de natureza aplicada, com abordagem exploratória e descritiva, caracterizada como estudo técnico-documental do sistema. A metodologia envolveu a análise de registros de interação, métricas de uso e desempenho, além da categorização temática das mensagens, com apoio de estatística descritiva e mapeamento de evidências aos atributos de qualidade de software. Os resultados indicaram elevada disponibilidade operacional do sistema, com tempo médio de resposta compatível com uso em contexto educacional, além de padrões de interação concentrados majoritariamente em dúvidas conteudistas, especialmente nas áreas de Matemática e Ciências da Natureza. Observou-se estabilidade do serviço ao longo do período analisado, com baixa incidência de falhas críticas e recorrência de sessões registradas. Também foram identificadas limitações relacionadas à usabilidade e à transparência dos mecanismos de adaptação, apontando desafios técnicos para a ampliação da profundidade das interações. Ressalta-se que os indicadores analisados correspondem a proxies técnicos de uso e desempenho, não permitindo inferências diretas sobre aprendizagem ou mediação didática. Conclui-se que o EducaChat apresenta desempenho técnico satisfatório como ferramenta educacional de apoio, embora demande aprimoramentos em aspectos de usabilidade e personalização para favorecer interações mais ricas e alinhadas às práticas pedagógicas.

Palavras-chave: Inteligência Artificial. Chatbot Educacional. Engajamento Discente. Avaliação de Tecnologias Educacionais.

ABSTRACT

This study evaluated the technical and operational performance of the educational chatbot EducaChat, based on metrics extracted from usage logs and the quality criteria defined by the ISO/IEC 25010 model, discussing its potential as a support tool for technology-mediated pedagogical practices. The research adopted an applied, exploratory, and descriptive approach, characterized as a technical-documentary study of the system. The methodology involved the analysis of interaction records, usage and performance metrics, as well as thematic categorization of messages, supported by descriptive statistics and evidence mapping to software quality attributes. The results indicated high system availability, average response times compatible with educational use, and interaction patterns predominantly focused on content-related questions, especially in Mathematics and Natural Sciences. Service stability was observed throughout the analyzed period, with a low incidence of critical failures and recurring session records. Limitations related to usability and transparency of adaptation mechanisms were also identified, pointing to technical challenges in promoting deeper interactions. It is emphasized that the analyzed indicators correspond to technical proxies of system usage and performance and do not allow direct inferences about learning outcomes or instructional mediation. It is concluded that EducaChat demonstrates satisfactory technical performance as an educational support tool; however, improvements in usability and personalization are required to foster richer interactions aligned with pedagogical practices.

Keywords: Artificial Intelligence. Educational Chatbot. Student Engagement. Educational Technology Evaluation.

LISTA DE FIGURAS

Figura 1- Encadeamento lógico entre problema, objetivos, método e evidências.	35
Figura 2 - Fluxo lógico da dissertação.	43
Figura 3 – Linha do tempo de agentes conversacionais (1966–2025).	51
Figura 4 – Características, subcaracterísticas e atributos de qualidade de software.	65
Figura 5 - Abordagens da IA segundo Russell e Norvig (2013).	66
Figura 6– Linha do tempo simplificada da evolução do PLN.	67
Figura 7– Arquitetura em camadas de um chatbot educacional.	70
Figura 8 – Fluxo metodológico da pesquisa, abrangendo as etapas de coleta, ETL, análise e síntese de evidências.	84
Figura 9 – Modelo relacional das tabelas de dados do EducaChat.	90
Figura 10 – Fluxo ETL e análise de dados do EducaChat.	94
Figura 11 - Arquitetura lógica do EducaChat.	104
Figura 12 – Arquitetura de código implementada do EducaChat.	108
Figura 13 – Interface inicial do EducaChat acessível em < https://educachat.com.br >.	121
Figura 14 – Interface de conversa EducaChat acessível em < https://educachat.com.br >.	121
Figura 15 - Histograma da duração das sessões (duration_s).	122
Figura 16 – Histograma do número de mensagens por sessão (msg_count).	122
Figura 17 - Boxplot da latência de resposta (latency_ms).	124
Figura 18 – Série temporal de disponibilidade (uptime_s).	126
Figura 19 – Série temporal de erros de servidor (http_5xx_count).	126
Figura 20 – Distribuição de temas (topic)	128
Figura 21 – Treemap de diversidade temática.	128
Figura 22 – Fluxograma de síntese da dissertação.	143

LISTA DE TABELAS

Tabela 1 - Desafios e oportunidades da IA na educação (Brasil vs. dir. Internac.).....	27
Tabela 2 - Princípios éticos aplicados à IA na educação.....	28
Tabela 3 - Problema, objetivos e hipóteses da pesquisa.....	30
Tabela 4 - Mapeamento entre BNCC e funcionalidades do EducaChat.....	32
Tabela 5 – Relações entre conceitos operacionais e implicações analíticas.....	36
Tabela 6 – Lacunas identificadas na literatura e a abordagem adotada nesta pesquisa.....	37
Tabela 7– Contribuições práticas, teóricas e metodológicas.	42
Tabela 8 – Comparativo entre aprendizagem tradicional, digital e ubíqua.	46
Tabela 9 - Comparativo internacional e nacional no uso de agentes conversacionais em educação.	54
Tabela 10 – Mapeamento conceitual (framework LA): variáveis e procedimentos do estudo.	56
Tabela 11 – Atributos de qualidade (ISO/IEC 25010) e indicadores baseados em logs.	58
Tabela 12 – Educação.....	60
Tabela 13 - Sintetiza as principais diferenças entre os dois referenciais normativos.....	62
Tabela 14 - Qualidade em uso segundo a ISO/IEC 25010.....	63
Tabela 15– BNCC: Competências/habilidades e suporte do EducaChat.	68
Tabela 16 – Riscos pedagógicos e salvaguardas na adoção de chatbots.	69
Tabela 17 – Comparação de modelos de chatbots aplicados à educação.....	71
Tabela 18 – Impactos pedagógicos de chatbots em diferentes contextos educacionais.....	72
Tabela 19– Principais limitações e desafios dos chatbots educacionais.	74
Tabela 20– Lacunas identificadas na literatura sobre chatbots educacionais.....	76
Tabela 21– Alinhamento entre hipóteses, métricas de qualidade e instrumentos analíticos....	77
Tabela 22– Comparação entre dimensões de qualidade e evidências na literatura.	77
Tabela 23 – Caracterização metodológica do estudo.	79
Tabela 24– Mapeamento entre questões de pesquisa, métricas e fontes de dados.....	82
Tabela 25 – Papéis e responsabilidades na governança de dados.	85
Tabela 26 – Instrumentos analíticos, finalidades e limitações (dados secundários).	86
Tabela 27 - Fontes, critérios e medidas de anonimização.	87
Tabela 28 – Verificações de qualidade e auditoria dos dados.....	88
Tabela 29 – Matriz de risco em dados e controles de mitigação.....	88

Tabela 30 – Esquema lógico simplificado dos registros do EducaChat.....	89
Tabela 31 – Dicionário de variáveis: definições operacionais e finalidades.....	90
Tabela 32 – Regras de validação e tratamento por variável.....	91
Tabela 33 – Cobertura dos dados e escopo temporal.....	91
Tabela 34 – Mapeamento de variáveis derivadas e consistências.....	93
Tabela 35 – Pipeline ETL e garantias de qualidade.....	94
Tabela 36 – Janelas analíticas e justificativas de uso.....	96
Tabela 37 - Atributos da ISO/IEC 25010 e métricas operacionais no EducaChat.....	99
Tabela 38 - Vínculo entre ISO/IEC 25010.....	99
Tabela 39 - Rastreabilidade ISO/IEC 25010 → dados → evidências.....	99
Tabela 40 - Checklist de reprodutibilidade e auditoria.....	101
Tabela 41 – Riscos residuais e respectivas mitigações.....	102
Tabela 42 – Conformidade ética e legal assegurada na pesquisa.....	103
Tabela 43– Requisitos funcionais (RF).....	105
Tabela 44 – Requisitos não funcionais (RNF).....	106
Tabela 45 - Ambiente de desenvolvimento e execução.....	107
Tabela 46 - Endpoints principais do backend.....	109
Tabela 47 - Testes iniciais realizados no protótipo.....	110
Tabela 48 – Resumo de versões e componentes do EducaChat.....	111
Tabela 49 – Correspondência entre componentes técnicos e finalidade pedagógica.....	112
Tabela 50 - Cronograma dos testes técnicos do EducaChat.....	113
Tabela 51 - Cronograma sintético dos testes técnicos e geração de evidências.....	114
Tabela 52 – Riscos operacionais e medidas de mitigação (cenário técnico-documental).....	115
Tabela 53 – Protocolo de execução de cenários simulados e captura de evidências.....	116
Tabela 54 - Rastreabilidade entre relato, metodologia e resultados.....	116
Tabela 55 – Caracterização do conjunto de dados.....	118
Tabela 56 – Estatística descritiva de sessões e mensagens.....	119
Tabela 57 – Desempenho e confiabilidade operacional.....	119
Tabela 58 – Distribuição temática das interações (topic).....	120
Tabela 59 – Indicadores-síntese alinhados à ISO/IEC 25010.....	130
Tabela 60 – Conformidade dos indicadores do EducaChat em relação a parâmetros de referência.....	133
Tabela 61 – Síntese das contribuições do estudo.....	144
Tabela 62 – Limitações do estudo e implicações para interpretação.....	145

Tabela 63 – Roadmap de evolução do EducaChat: iniciativas, horizonte e evidências.....	147
Tabela 64 – Regras de qualidade e ações corretivas no ETL.	158
Tabela 65 - Mapa de evidências: métricas do estudo → atributos ISO/IEC 25010.	161

LISTA DE ABREVIATURAS E SIGLAS

AVA	Ambientes Virtuais de Aprendizagem
BERT	(Bidirectional Encoder Representations from Transformers) - Representações de Codificação Bidirecional a partir de Transformers
BNCC	Base Nacional Comum Curricular
ETL	(Extract, Transform, Load) - Extração, Transformação e Carga
IA	Inteligência Artificial
IAED	Inteligência Artificial na Educação
IEC	(International Electrotechnical Commission) - Comissão Eletrotécnica Internacional
ISO	(International Organization for Standardization) - Organização Internacional de Normalização
GPT	(Generative Pre-trained Transformer) - Transformer Generativo Pré-treinado
LGPD	Lei Geral de Proteção de Dados Pessoais
LLM	(Large Language Model) - Modelo de Linguagem de Grande Escala
OECD	(Organisation for Economic Co-operation and Development) - Organização para a Cooperação e Desenvolvimento Econômico
QP	Questão de Pesquisa
TIC	Tecnologias da Informação e Comunicação
UFT	Universidade Federal do Tocantins
UNESCO	Organização das Nações Unidas para a Educação, a Ciência e a Cultura

SUMÁRIO

1. INTRODUÇÃO	21
1.1.1 Contexto e trajetória recente (2010–2025)	25
1.1.3 Panorama nacional aprofundado (dados, desigualdades e implicações)	26
1.1.4 Políticas e diretrizes (UNESCO, OECD) e sua tradução didática	26
1.1.5 Ética, Privacidade e Governança da IA na Educação	27
1.1.6 Do potencial técnico à mediação pedagógica: por que o como importa	28
1.1.7 Relevância social e científica	29
1.2 Justificativa	30
1.2.1 Panorama e relevância empírica	31
1.2.2 Relevância social e alinhamento à BNCC	32
1.2.3 Risco e oportunidade	33
1.2.4 Delimitações e escopo	33
1.2.5 Limitações e escopo negativo	34
1.2.6 Validade, transferibilidade e limites do estudo	34
1.2.7 Definições operacionais	34
1.2.9 Questões éticas e governança de IA	36
1.2.9.1 Cenário de uso ilustrativo	37
1.3 Originalidade e impacto esperado	38
1.3.1 Problema de pesquisa (versão expandida) e derivação das perguntas	38
1.4 Objetivo Geral	39
1.4.1 Objetivos Específicos	39
1.4.2 Questões de pesquisa	40
1.4.3 Evidências que respondem às QPs	40
1.4.4 Contribuições e organização da dissertação	41
1.4.5 Estrutura Lógica do Trabalho	43
1.4.6 Hipóteses analíticas	44
2. FUNDAMENTAÇÃO TEÓRICA	45
2.1 Histórico da Inteligência Artificial na Educação (IAED)	45
2.1.2 Aprendizagem ubíqua e mediação digital	45
2.1.3 IA na Educação: trajetória e inserção	46
2.1.4 Anos 1970: Os primeiros sistemas tutores inteligentes	47
2.1.5 Anos 1980 e 1990: Consolidação de abordagens cognitivas	47
2.1.6 Anos 2000: E-learning e ambientes virtuais	47
2.1.7 Anos 2010: Aprendizado de máquina e análise de dados educacionais	48
2.1.8 Anos 2020: A revolução da IA generativa	48
2.1.9 Síntese histórica	48

2.2 Histórico dos Agentes Conversacionais e Chatbots	49
2.2.1 Chatbots educacionais no contexto BNCC e generativo	52
2.2.2 Panorama Internacional e Nacional	54
2.2.3 Referencial teórico: learning analytics, governança de dados e avaliação de qualidade em agentes conversacionais.....	55
2.2.4 Learning analytics com dados secundários e logs técnico-operacionais	55
2.2.5 Governança de dados e conformidade Legal (LGPD).	56
2.2.6 Avaliação de qualidade técnica com base na ISO/IEC 25010	57
2.2.7 Agentes conversacionais na educação: lacunas de validação técnica	58
2.2.8 Síntese integradora e implicações para a metodologia.....	58
2.2.9 Fundamentos pedagógicos da arquitetura do sistema	59
2.3 Trabalhos Relacionados e Lacuna de Pesquisa	59
2.3.1 Estudos Complementares na Literatura.....	60
2.3.2 Avaliação de agentes conversacionais incorporados.....	61
2.3.3 Métricas de Qualidade de Software.....	61
2.3.4 Comparação entre ISO/IEC 9126-1 e ISO/IEC 25010.....	61
2.3.5 Qualidade em uso	63
2.3.6 Escolha das métricas de avaliação.....	63
2.3.7 Métricas internas e externas	64
2.3.8 Relevância para este estudo.....	65
2.3.9 Atributos de qualidade para Chatbots: proposta de Radziwill e Benton (2017)	65
2.4 Bases conceituais.....	66
2.4.1 Inteligência Artificial: Perspectivas históricas e educacionais.....	66
2.4.4 Alinhamento à BNCC (Matemática e Ciências da Natureza)	68
2.4.9 Riscos, salvaguardas e papel do professor	68
2.5.6 Modelos e Arquiteturas Técnicas de Chatbots Educacionais.....	69
2.5.7 Arquitetura em camadas.	69
2.5.8 Modelos baseados em regras, estatísticos e neurais	70
2.5.9 Integração com ambientes virtuais de aprendizagem	70
2.6 Quadro comparativo de modelos.....	71
2.6.1 Impactos pedagógicos observados em estudos recentes	71
2.6.2 Potencial de sustentação da interatividade	71
2.6.3 Suporte à Resolução de Problemas e Scaffolding	71
2.6.4 Papel no suporte ao professor.....	72
2.6.5 Síntese dos impactos.....	72
2.6.6 Considerações gerais	72
2.7.3 Limitações e Desafios Atuais dos Chatbots Educacionais	73
2.7.4 Desafios técnicos	73
2.7.5 Desafios pedagógicos	73
2.7.7 Desafios institucionais.....	73
2.7.8 Síntese dos desafios.....	74

2.7.9 Síntese final	74
2.8 Lacunas na Literatura	74
2.8.1 Contexto educacional brasileiro	75
2.8.3 Integração curricular.....	75
2.8.4 Métricas de qualidade em software educacional	75
2.8.5 Dimensões éticas e socioculturais	75
2.8.6 Síntese das lacunas	76
2.8.7 Síntese final	76
2.9.7 Síntese crítica da literatura	77
3. METODOLOGIA.....	79
3.1 Natureza e Desenho do Estudo.....	79
3.2 Hipóteses e questões de pesquisa	80
3.2.1 Hipóteses operacionais (H).....	81
3.2.2 Escopo e observações.....	81
3.2.3 Plano de análise: Abordagem mista em dados secundários	82
3.2.4 Integração Analítica.....	83
3.2.5 Justificativa metodológica	83
3.2.6 Fluxograma metodológico e cronograma.....	84
3.2.7 Modelo de governança de dados	84
3.2.8 Fontes de dados	85
3.2.9 Critérios de inclusão e exclusão	86
3.2.9.1 Política de anonimização e minimização	86
3.2.9.2 Ciclo de vida, retenção e descarte	87
3.3 Segurança da informação e controle de acesso	87
3.3.1 Qualidade, completude e auditoria de dados.....	87
3.3.2 Matriz de risco e controles	88
3.3.3 Modelo de dados e dicionário de variáveis	88
3.3.4 Esquema lógico dos logs	89
3.3.5 Modelo relacional.....	89
3.3.5.1 Dicionário de variáveis.....	90
3.3.5.2 Cobertura e Escopo Temporal.....	91
3.3.6 Procedimentos de Coleta e Tratamento (ETL).....	92
3.3.6.1 Extração (Extract).....	92
3.3.6.2 Transformação (Transform)	92
3.3.6.3 Carga (Load).....	93
3.3.7 Plano de Validação	94
3.3.7.1 Validade interna.....	94
3.3.7.2 Validade externa / transferibilidade analítica	95
3.3.7.3 Confiabilidade dos indicadores	95
3.3.8 Plano de Análise	95

3.3.8.1 Análise Quantitativa (Performance e Estabilidade)	96
3.3.8.2 Análise de Conteúdo Automatizada (Topic)	96
3.3.9.3 Fluxo analítico integrado	97
3.4 Avaliação de Qualidade (ISO/IEC 25010)	97
3.4.1 Funcionalidade	97
3.4.2 Confiabilidade	98
3.4.3 Usabilidade	98
3.4.4 Eficiência de desempenho	98
3.4.5 Quadros e indicadores	98
3.4.6.1 Validade, Confiabilidade e Reprodutibilidade	99
3.4.6.2 Ameaças e estratégias de mitigação	100
3.4.6.3 Confiabilidade e auditoria do processo	100
3.4.6.4 Reprodutibilidade	100
3.4.6.5 Estratégias adicionais de robustez	101
3.4.6.6 Considerações Éticas e de Privacidade.....	101
3.4.6.7 Limitações éticas e riscos residuais	101
3.4.6.8 Uso de dados secundários anonimizados	102
3.4.6.9 Anonimização e governança de dados	102
3.4.7 Conformidade ética e legal	103
4. MODELAGEM DO SISTEMA.....	104
4.1 Visão Geral da Arquitetura.....	104
4.1.1 Racional de desenho	105
4.2 Requisitos do Sistema	105
5. IMPLEMENTAÇÃO DO SISTEMA.....	107
5.1 Ambiente de Desenvolvimento e Execução	107
5.2 Arquitetura de código	108
5.3 Principais componentes	109
5.3.1 Backend (Spring Boot)	109
5.4 Comunicação em Tempo Real (WebSocket/STOMP)	109
5.4.1 Integração com API de IA	109
5.4.2 Monitoramento e Logs	109
5.4.3 Procedimentos de Teste Inicial.....	110
5.4.4 Desafios e soluções.....	110
5.4.5 Síntese e transição	110
5.4.6 Relato Técnico/Documental do Sistema	111
5.4.7 Arquitetura e versões	111
5.4.8 Discussão analítica da arquitetura	112
5.4.9 Cenário de testes técnicos.....	113
5.5 Mapeamento às fontes analíticas	113
5.5.2 Limitações técnicas do relato	114

5.5.3 Discussão crítica das limitações	114
5.5.4 Síntese.....	115
5.5.5 Nota de transição	116
6. RESULTADOS	117
6.1 Caracterização do conjunto de dados	117
6.2 Estatística descritiva de sessões e mensagens	122
6.3 Desempenho: Latência	124
6.4 Confiabilidade operacional.....	125
6.6 Indicadores-Síntese ISO/IEC 25010.....	129
6.7 Síntese interpretativa	131
6.8 Conformidade técnica e pedagógica	133
7. DISCUSSÃO	134
7.1 Síntese dos achados	134
7.2 Interpretação técnico-pedagógica	135
7.2.1 Eficiência e fluidez cognitiva	135
7.2.2 Engajamento e padrões de interação	136
7.2.3 Aderência curricular e personalização.....	136
7.2.4 Contribuições do Estudo.....	137
7.2.5 Contribuições acadêmicas	137
7.2.6 Contribuições práticas	138
7.2.7 Contribuições metodológicas	138
7.5 Recomendações para trabalhos futuros	138
7.5.1 Ampliação empírica.....	139
7.5.2 Avanços técnicos	139
7.6 Dimensão pedagógica.....	139
7.7 Integração sistêmica	140
7.8 Síntese final	140
8. CONCLUSÃO	142
8.1 Síntese dos achados	142
8.2 Contribuições.....	143
8.3 Implicações práticas	145
8.4 Limitações	145
8.5 Ameaças à validade	146
8.6 Perspectivas futuras	146
REFERÊNCIAS	150
APÊNDICES	154
 APÊNDICE A – Protocolo Técnico deETL e Auditoria de Dados	 162
 A.1 Pré-requisitos e insumos	 155

A.2	Checklist de conformidade (LGPD/Ética)	155
A.3	Pseudocódigo do ETL (alto nível)	155
A.4	Regras de validação (QA) e tratamento	158
A.5	Produtos do ETL	158
APÊNDICE B – Roteiro Analítico Reprodutível (Quantitativo e Qualitativo)		159
B.1	Métricas, indicadores e visualizações	159
B.2	Pseudocódigo analítico (alto nível)	159
B.3	Análise qualitativa (documental/temática)	160
B.4	Mapeamento de métricas à ISO/IEC 25010	161
B.5	Reprodutibilidade e auditoria	161

1. INTRODUÇÃO

A Inteligência Artificial (IA) consolidou-se como uma das tecnologias mais disruptivas do século XXI, em especial por suas aplicações no campo educacional. Entre as diversas possibilidades, os agentes conversacionais educacionais destacam-se pela capacidade de simular diálogos inteligentes e promover experiências de aprendizagem interativas e personalizadas, ajustadas ao perfil e às necessidades de cada estudante (Junior; Belli; Medeiros, 2020). Integrados a plataformas digitais, tais agentes contribuem para ambientes dinâmicos, que respeitam o ritmo individual de aprendizagem e ampliam o acesso a conteúdos complexos.

Apesar desse potencial, a adoção de IA na educação enfrenta desafios substantivos, as “necessidades de aprendizagem e as teorias que descrevem os princípios e processos de aprendizagem que devem refletir os ambientes sociais subjacentes” (Siemens, 2007, p.1). O Conectivismo é frequentemente tratado como uma proposta teórica emergente sobre aprendizagem em contextos digitais em redes virtuais na web que se constitui em um instrumento capaz de permitir o aumento exponencial das conexões às fontes de informações necessárias para produção do conhecimento (Siemens, 2010).

Contudo, para que esses princípios se efetivem, permitindo que o estudante navegue por essas redes de informação de forma autônoma e produtiva, a infraestrutura tecnológica de suporte deve operar com níveis elevados de confiabilidade e eficiência. Neste contexto, o Conectivismo figura nesta dissertação como o horizonte pedagógico almejado, enquanto o foco da investigação recai sobre os alicerces técnicos necessários para alcançá-lo. Assume-se a premissa de que falhas na qualidade do software (como alta latência, indisponibilidade ou respostas incoerentes) rompem o fluxo de conexão proposto por Siemens (2007), impedindo a construção do conhecimento. Portanto, antes de mensurar a eficácia pedagógica final, torna-se imperativo validar o desempenho técnico-operacional da ferramenta mediadora.

As limitações técnicas de compreensão semântica e pragmática, somadas a questões éticas relativas à privacidade e ao viés algorítmico, afetam a confiabilidade e a eficácia pedagógica dessas soluções. A literatura recente aponta que tais fragilidades demandam investigações empíricas sólidas, capazes de distinguir promessas tecnológicas de ganhos efetivos de aprendizagem (Coelho; Outros, 2023).

Neste sentido, ressalta-se que os sistemas de IA têm o potencial de “democratizar o ensino, pois torna possível a disponibilização universal de cursos e conhecimentos, bem como de professores que antes estavam limitados às instituições limitadas pelo tempo e pelo espaço” (Azambuja; Silva, 2024, p. 7).

Desse modo, a maior utilidade para IA no campo da educação condiz ao aprendizado personalizado (Lee; Qiufan, 2021).

Ainda assim, os benefícios potenciais justificam um exame aprofundado. Agentes conversacionais podem favorecer a interatividade, apoiar a autorregulação e elevar o engajamento discente, especialmente quando atuam como mediadores que oferecem pistas graduais em vez de respostas prontas (Pacheco et al., 2024).

Esse processo permite o uso de estratégias, metas e expectativas para a criação de condições para a efetividade da aprendizagem (Martins, 2017). Esses benefícios têm sido observados em experiências empíricas recentes. Mendes e Sirqueira (2022) evidenciam que o uso de *chatbots* auxiliam calouros no ensino superior a superar dificuldades de adaptação, provendo suporte acessível e personalizado.

De modo convergente, Guerreiro e Barros (2019) apontam ganhos no ensino a distância quando o agente ajusta o ritmo e fornece explicações complementares sob demanda, reduzindo a carga cognitiva diante de conteúdos complexos.

Para compreender esse movimento de inserção da IA na educação, é relevante resgatar sua trajetória histórica. O termo inteligência artificial foi cunhado por John McCarthy (1956) no pós-segunda guerra mundial, em meio ao avanço da ciência da computação. Russell e Norvig (2013) definem agente racional como aquele que, percebendo seu ambiente, toma decisões orientadas ao melhor resultado possível, mesmo sob incerteza.

Antes disso, o teste de Turing já questionava a possibilidade de uma máquina produzir respostas indistinguíveis das humanas (Russell; Norvig, 2013). Esses marcos demonstram que a relação homem-máquina acompanha a própria evolução da computação.

No domínio educacional, a literatura indica que a tecnologia, por si só, não garante aprendizado. Paschoal, Chicon e Falkembach (2017) relatam o desenvolvimento do Ubibot, voltado ao apoio de estudantes de Computação, com ênfase na personalização das interações

— característica associada à aprendizagem ubíqua¹. Em paralelo, estudos como os de Clemente (2016) e Santana (2017) destacam que a disponibilidade contínua de um agente pode reduzir o tempo de resposta, ampliar o acesso e atenuar desigualdades de apoio extraclasse.

Sob uma perspectiva teórico-pedagógica, Harsley, Smith e Johnson (2017) e Cho e Lim (2017) demonstram que a aprendizagem colaborativa, frequentemente mediada por ferramentas digitais permite a construção do conhecimento, o desenvolvimento do pensamento crítico e da

¹ A aprendizagem ubíqua consiste em um modelo de aprendizado que ocorre em qualquer lugar, a qualquer momento, impulsionado por tecnologias móveis e a conectividade contínua (Santaella, 2013).

criatividade na resolução de problemas.

Esses achados dialogam com os princípios da aprendizagem multimídia, conforme Mayer e Moreno (2003), segundo a qual a integração de estímulos verbais e visuais potencializa uma melhor compreensão e memória. Em linha semelhante, Araujo, Souza e Martins (2014) reforçam que recursos multimídia favorecem a construção compartilhada de significados.

Do ponto de vista didático, a noção de orquestração pedagógica, proposta por Dillenbourg (2013a), descreve a habilidade do professor em coordenar, em tempo real, múltiplas atividades e recursos — presenciais e digitais — de forma coerente. Nessa perspectiva, um agente conversacional pode assumir função de mediador, ofertando devolutivas imediatas e registrando pistas de aprendizagem que subsidiem a intervenção docente.

É nesse cenário que se insere a presente pesquisa, centrada no EducaChat, um protótipo de agente conversacional educacional concebido para atuar como mediador de aprendizagem em conteúdos de matemática e biologia (BNCC: áreas de matemática e ciências da natureza). A proposta não é substituir a docência, mas ampliar o alcance do apoio pedagógico, oferecendo orientação graduada, *feedback* oportuno e exemplos contextualizados.

A relevância da investigação decorre de três frentes. No eixo social, busca-se compreender se a arquitetura do agente conversacional oferece a estabilidade e a disponibilidade necessárias para atender turmas numerosas em escala, condição *sine qua non* para que futuras intervenções possam democratizar o apoio pedagógico individualizado.

No eixo científico, objetiva-se preencher uma lacuna metodológica sobre a validação técnica de ferramentas educacionais baseadas em IA, propondo métricas de avaliação que antecedem a aplicação em larga escala. No eixo prático, pretende-se derivar diretrizes operacionais baseadas nos requisitos de qualidade da norma ISO/IEC 25010, oferecendo um roteiro seguro para a adoção responsável dessas tecnologias nas escolas.

A contribuição esperada articula fundamentos teóricos e análise empírica técnica. Do ponto de vista pedagógico, verifica-se se a ferramenta apresenta os padrões de interatividade e disponibilidade necessários para viabilizar a mediação docente e suportar estratégias de personalização do ensino.

Do ponto de vista tecnológico, examinam-se as características de qualidade segundo a ISO/IEC 25010, especialmente usabilidade técnica, confiabilidade e eficiência, correlacionando essas métricas com a continuidade do fluxo de uso registrada nos logs do sistema.

Além de investigar as métricas de desempenho, esta dissertação reconhece, teoricamente, as condições de contorno que modulam a aplicação pedagógica: infraestrutura

escolar, heterogeneidade das turmas e mediação docente. A premissa adotada é que a efetividade final do agente em sala de aula depende de uma simbiose: de um lado, suas propriedades internas (validade das respostas, interface, baixa latência — foco desta análise); de outro, a integração curricular (que pressupõe um sistema tecnicamente estável para ocorrer).

Assim, a análise técnica aqui proposta serve como base para garantir que as condições de software não sejam um gargalo para a prática docente. Pelos dados a pesquisa foca na infraestrutura técnica; o perfil discente é tema para trabalhos futuros, como apontado nas limitações.

Metodologicamente, a pesquisa adota uma abordagem quantitativa com apoio de categorização automatizada, de natureza aplicada e caráter descritivo-exploratório. A investigação fundamenta-se na extração e análise estatística de registros de logs do sistema (Big data educacional), permitindo a mensuração objetiva de variáveis como volumetria de mensagens, latência de resposta, taxas de erro e tempo de sessão. Diferentemente de abordagens que buscam capturar a subjetividade da experiência discente, este desenho foca na validação da infraestrutura técnica e dos padrões de uso, utilizando as métricas da norma ISO/IEC 25010 como régua para aferir a qualidade do produto de software educacional.

Tal escolha metodológica justifica-se pelo caráter não intrusivo e objetivo da coleta, eliminando vieses de auto-relato e assegurando o registro fidedigno do comportamento digital dos estudantes. Em termos analíticos, os logs são tratados como indicadores de interatividade e de confiabilidade do software.

Desta forma, a análise dispensa a necessidade de observação direta das práticas de sala de aula, concentrando-se na validação da infraestrutura tecnológica que deve sustentar a orquestração pedagógica.

Ao assumir um desenho metodológico estritamente quantitativo, esta pesquisa constrói evidências empíricas robustas sobre a estabilidade e a usabilidade técnica do agente conversacional. Articulam-se métricas de desempenho com os requisitos operacionais de um ambiente de ensino.

Conforme destacado por Denzin (1978), a validade de indicadores indiretos reside em seu enquadramento interpretativo; portanto, os dados de log são analisados em diálogo com as categorias de qualidade de software, demonstrando se a ferramenta possui a maturidade técnica necessária para viabilizar futuras intervenções educacionais.

Essa opção metodológica, embora limitada ao registro quantitativo, não elimina a necessidade de interpretações pedagógicas; pelo contrário, evidencia como indicadores técnicos, quando extraídos de situações reais de uso, podem oferecer subsídios objetivos para

validar a disponibilidade e a estabilidade dos agentes conversacionais. Nesse sentido, o Conectivismo vai ao encontro da proposta dos assistentes virtuais, uma vez que se baseia na ideia de que o conhecimento reside e é adquirido por meio de redes de conexões (Siemens, 2010).

Operacionalmente, o EducaChat foi arquitetado para privilegiar pistas graduais, diagnóstico de pré-requisitos e explicações por analogias, evitando a entrega de respostas finais sem mediação.

Tal diretriz de design visa estimular a interatividade e reduzir a dependência passiva, gerando dados de log que permitem monitorar o fluxo da interação. No recorte de escopo, a investigação abrange conteúdos de matemática e biologia em nível de educação básica pública.

Não se pretende avaliar efeitos de longo prazo em desempenho padronizado, nem comparar múltiplas arquiteturas proprietárias. O foco é verificar, por meio de métricas de qualidade de software, se a mediação conversacional é tecnicamente viável e sustentável sob condições de uso intensivo.

1.1.1 Contexto e trajetória recente (2010–2025)

A inserção de IA na educação percorre um arco que vai de *chatbots* baseados em regras a modelos de linguagem de grande porte (*LLMs*). Nesse cenário em evolução, o conectivismo oferece o suporte teórico ao postular que o conhecimento é distribuído por meio de redes de informação e o aprendizado decorre da capacidade de navegar e se conectar nessas redes (Siemens, 2010).

Ao longo da década de 2010, experiências piloto mostram ganhos de acesso e agilidade; entre 2020 e 2025, com a popularização de *LLMs*, cresce a promessa de mediação mais contextualizada, mas também os alertas sobre alucinações, privacidade e viés. Este trabalho se posiciona nesse ponto de inflexão: captura o potencial do diálogo assistido por IA e, simultaneamente, estabelece salvaguardas metodológicas e éticas para leitura responsável dos resultados.

1.1.2 Panorama internacional da IA na educação

Relatórios recentes da UNESCO e da OECD reforçam que a inteligência artificial aplicada à educação está em processo acelerado de adoção, com projeções de crescimento global que superam 20% ao ano (OECD, 2023). No contexto brasileiro, a pesquisa TIC

educação 2023 aponta que 92% das escolas de ensino fundamental e médio possuem algum tipo de conectividade à Internet, embora apenas 43% relatem uso pedagógico sistemático das Tecnologias Digitais (CETIC.br, 2024). Essa discrepância entre acesso e uso efetivo indica um campo fértil para investigações aplicadas.

1.1.3 Panorama nacional aprofundado (dados, desigualdades e implicações)

No Brasil, os avanços de conectividade reportados pela TIC educação 2023 coexistem com desigualdades estruturais no uso pedagógico efetivo das tecnologias (CETIC.BR, 2024). As assimetrias se manifestam em três camadas: (i) infraestrutura (velocidade e estabilidade da rede; disponibilidade de dispositivos); (ii) capital pedagógico (formação docente continuada, curadoria de recursos, integração curricular); e (iii) governança escolar (políticas, protocolos, acompanhamento).

A literatura indica que iniciativas de IA que não enfrentam simultaneamente essas camadas tendem a produzir efeitos localizados e de curta duração (Holmes; Luckin, 2023).

A adoção de agentes conversacionais, portanto, deve ser compreendida como parte de um ecossistema pedagógico mais amplo, no qual o planejamento e a orquestração docente assumem papel central (Dillenbourg, 2013b). Em turmas numerosas, o agente é projetado para atuar como mediador de baixa fricção, oferecendo pistas e checagens rápidas. Tal funcionalidade visa, teoricamente, liberar o professor para intervenções de maior complexidade cognitiva, como debates e avaliação formativa, otimizando o tempo letivo.

Nesse sentido, os agentes conversacionais configuram-se como nós ativos na rede de aprendizagem, alinhando-se aos princípios do conectivismo (Siemens, 2010). A ferramenta atua facilitando o acesso imediato ao conhecimento e apoiando a construção autônoma de conexões com conteúdos e fontes externas.

Ao assumir funções de curadoria em meio ao excesso de dados, os agentes materializam a premissa conectivista de que aprender é a capacidade de navegar e utilizar redes, transformando a interação técnica em oportunidade cognitiva.

1.1.4 Políticas e diretrizes (UNESCO, OECD) e sua tradução didática

As recomendações da UNESCO e da OECD convergem para princípios de uso ético, equidade, transparência e desenvolvimento docente (UNESCO, 2023; OECD, 2023). A tradução didática desses princípios na escola envolve, minimamente: (a) clareza de propósito

pedagógico (para quê usar o agente); (b) padrões de qualidade (ex.: tempo de resposta, clareza das devolutivas); (c) salvaguardas (privacidade, uso responsável, linguagem inclusiva); e (d) rotinas de avaliação (indicadores e evidências). É nessa chave que se posiciona o EducaChat: um recurso desenhado com objetivos instrucionais explícitos, cuja validação se dá por meio de métricas técnicas rigorosas que asseguram sua adequação ao ambiente educacional.

Tabela 1 - Desafios e oportunidades da IA na educação (Brasil vs. diretrizes internacionais).

Categoria	Contexto Brasil	Diretrizes internacionais (UNESCO/OECD)
Infraestrutura	Avanços de conectividade (TIC Educação 2023), porém com assimetrias de velocidade e disponibilidade de dispositivos entre redes e regiões.	Garantir infraestrutura suficiente e sustentável antes de ampliar casos de uso de IA; priorizar acesso equitativo e contínuo.
Formação docente	Oferta heterogênea de formação continuada; dificuldades em integrar TIC e IA ao currículo.	Desenvolvimento profissional centrado no uso pedagógico responsável da IA e na avaliação crítica de resultados.
Acessibilidade	Barreiras de acessibilidade digital para parte dos estudantes e para escolas com recursos limitados.	Acessibilidade como requisito desde o design (UDL), com monitoramento de inclusão e mitigação de exclusões.
Ética e regulação	Marcos institucionais incipientes para uso pedagógico de IA; necessidade de protocolos de privacidade.	Princípios de transparência, prestação de contas, proteção de dados, equidade e não discriminação.
Potencial pedagógico	Oportunidade de apoio à orquestração docente e ao feedback em turmas numerosas; experiências ainda localizadas.	Uso da IA alinhado a objetivos de aprendizagem, com avaliação contínua de impactos e mediação ativa do professor.

Fonte: Elaboração própria a partir de TIC Educação (2023), UNESCO (2021, 2023) e OECD (2023).

A clareza do propósito pedagógico e a definição de padrões de qualidade para o uso de agentes conversacionais demandam salvaguardas rigorosas quanto à privacidade de dados, adequação da linguagem e ética digital.

Para além dos princípios normativos, torna-se imprescindível validar o desempenho técnico da inteligência artificial por meio de indicadores objetivos e evidências de estabilidade. Sob essa ótica, o EducaChat constitui-se como um recurso alinhado a objetivos instrucionais definidos, submetido a uma sistemática de avaliação que prioriza a robustez da infraestrutura como pré-requisito para a prática educativa.

1.1.5 Ética, privacidade e governança da IA na educação

O uso de IA em contextos educacionais não pode ser dissociado de debates sobre ética, privacidade e governança.

Relatórios recentes da UNESCO (2023) e da OECD (2024) destacam princípios de

transparência, responsabilidade, inclusão e equidade, reforçando que agentes digitais precisam ser acompanhados de protocolos de supervisão docente e salvaguardas de dados.

Nesta pesquisa, tais premissas não são apenas teóricas, mas constituem requisitos de qualidade, segurança e auditabilidade; que serão verificados na análise técnica dos registros do sistema.

Tabela 2 - Princípios éticos aplicados à IA na educação.

Princípio	Implicações pedagógicas
Transparência	Explicitar limites, objetivos e funcionamento do agente conversacional.
Privacidade	Coletar apenas dados essenciais; realizar anonimização dos logs de interação.
Equidade	Evitar vieses algorítmicos e garantir acessibilidade universal aos estudantes.
Responsabilidade	Assegurar supervisão docente e adotar protocolos para correção de falhas.

Fonte: UNESCO (2023); OECD (2024).

A observância de princípios éticos é fundamental para balizar o uso responsável dos agentes digitais. Simultaneamente, essa inserção tecnológica encontra respaldo em perspectivas pedagógicas contemporâneas, como o conectivismo.

Sob essa ótica, a integração da IA ao ambiente escolar permite a operação de mais um “nó” ativo na rede educacional, ampliando, potencialmente, as possibilidades de acesso, mediação e construção do conhecimento.

1.1.6 Do potencial técnico à mediação pedagógica: por que o como importa

A literatura diferencia ganhos de acesso (disponibilidade de informação) de ganhos de aprendizagem (mudança estável de desempenho) (Mayer; Moreno, 2003). Agentes conversacionais podem multiplicar acessos, mas efeitos educacionais dependem de design instrucional e mediação.

Nessa linha, a tradição sociocultural, Vygotsky argumenta que a aprendizagem se dá na zona de desenvolvimento proximal, por meio de suporte contingente que gradualmente se retira (Vygotsky, 2007). Papert, por sua vez, propõe que tecnologias sejam instrumentos para produzir artefatos significativos, favorecendo autoria e reflexão (Papert, 1993). Assim, um agente útil não é o que responde no lugar do aluno, mas o que orienta o aluno a construir a resposta, com

pistas e verificações (Kasneci et al., 2023; Xu et al., 2024).

1.1.7 Relevância social e científica

A integração de agentes conversacionais no ensino básico dialoga diretamente com as competências gerais da BNCC (Brasil, 2018), notadamente: (i) pensamento científico, crítico e criativo; (ii) cultura digital; (iii) comunicação; e (iv) responsabilidade e cidadania.

No plano social, argumenta-se que a escalabilidade técnica do EducaChat possui o potencial de ampliar a capacidade de atendimento individualizado em turmas numerosas, oferecendo *feedback* imediato onde recursos humanos são escassos.

No plano científico, este estudo oferece uma contribuição original ao fornecer evidências empíricas sobre a performance técnica de IAs em escolas públicas brasileiras, preenchendo uma lacuna na literatura que carece de validações operacionais em cenários de infraestrutura restrita (Kasneci et al., 2023; Xu et al., 2024).

1.1.8 Contribuições esperadas

As contribuições desta pesquisa podem ser classificadas em três dimensões:

- Práticas: oferecer subsídios técnicos a gestores e educadores para a seleção e validação de ferramentas de IA, estabelecendo critérios de qualidade (estabilidade e confiabilidade) que devem preceder a adoção em sala de aula.
- Teóricas: evidenciar a articulação entre os requisitos de qualidade de software e os pressupostos das teorias de aprendizagem, particularmente do conectivismo, sustentando que a robustez tecnológica constitui condição *sine qua non* para a efetividade da mediação pedagógica em sistemas educacionais interativos.
- Metodológicas: propor e validar protocolos de avaliação quantitativa baseados em logs de sistema, oferecendo um modelo replicável de auditoria pedagógica não intrusiva.

Tabela 3 - Problema, objetivos e hipóteses da pesquisa.

Elemento	Descrição
Problema de pesquisa	Em que medida o EducaChat contribui para o engajamento, a personalização e a mediação da aprendizagem no Ensino Médio?
Objetivo geral	Analisar a contribuição do chatbot educacional EducaChat para o engajamento discente, a personalização percebida e a mediação pedagógica, no contexto do Ensino Médio em Matemática e Biologia, em consonância com a Base Nacional Comum Curricular (BNCC).
Objetivos específicos	OE1: Desenvolver e implementar o chatbot educacional EducaChat como recurso de apoio pedagógico. OE2: Analisar padrões de uso e engajamento discente a partir de métricas extraídas de logs de interação. OE3: Examinar evidências de personalização percebida e de mediação pedagógica associadas ao uso do EducaChat. OE4: Avaliar atributos de qualidade do sistema segundo o modelo ISO/IEC 25010. OE5: Identificar limitações e condições de contorno para a adoção de agentes conversacionais educacionais em contextos escolares.
Hipóteses (H1–H4)	As hipóteses que orientam este estudo estão apresentadas e detalhadas na Seção 1.9.

Fonte: Elaborado pelo autor (2025).

Nessa perspectiva, à luz do conectivismo (Siemens, 2010), o EducaChat constitui-se como um nó ativo na rede de aprendizagem, viabilizando a criação, navegação e atualização de conexões indispensáveis à construção do conhecimento. Desse modo, a ferramenta opera como uma interface mediadora que agiliza o acesso a informações, sistematiza conteúdos e oferece *feedback* imediato.

Mais do que um repositório de dados, o agente favorece a autonomia discente ao suportar a articulação de múltiplas fontes, ampliando as possibilidades de conexão e permitindo aos estudantes trilhar percursos de aprendizagem personalizados.

1.2 Justificativa

Embora os *chatbots* tenham avançado significativamente, sua aceitação ainda encontra resistência, em especial pela percepção de respostas insatisfatórias ou pela preferência dos usuários por canais tradicionais, como redes sociais e e-mails (Guerreiro; Barros, 2019). Essa limitação evidencia a necessidade de aprimorar métodos de avaliação da qualidade e da usabilidade desses sistemas (Mendes; Sirqueira, 2022). Além disso, destaca-se a importância de incorporar práticas de design adequadas e instrumentos de análise já nas fases iniciais de

desenvolvimento, a fim de evitar investimentos ineficazes (Holmes; Bialik; Fadel, 2019).

1.2.1 Panorama e relevância empírica

Indicadores nacionais e internacionais reforçam a relevância de investigar o uso de inteligência artificial e agentes conversacionais no contexto da educação básica, especialmente no que se refere às condições de infraestrutura, uso de tecnologias digitais e desafios de governança:

Conectividade escolar: em 2023, 92% das escolas brasileiras de ensino fundamental e ensino médio possuíam acesso à Internet, incluindo avanços em contextos rurais e em redes municipais (CETIC.BR, 2024).

Uso de tecnologias educacionais: em instituições com acesso disponível, há registros de uso frequente de tecnologias da informação e comunicação (TIC) em atividades pedagógicas, evidenciando a crescente inserção de recursos digitais nos processos educacionais (CETIC.BR, 2024).

Acesso à internet: relatórios nacionais indicam ampla disseminação do acesso à internet entre adolescentes, com predominância do uso de dispositivos móveis — especialmente smartphones — conforme dados do CETIC.br (2024).

Governança de IA: as diretrizes institucionais para o uso pedagógico de inteligência artificial ainda se mostram incipientes, enquanto organismos internacionais recomendam políticas claras de uso responsável, com ênfase em segurança, equidade, transparência e proteção de dados (MIAO; HOLMES, 2023).

Nessa perspectiva, Simão et al. (2025) elucidam que o desenvolvimento de um agente conversacional baseado em inteligência artificial agêntica, voltado à construção de competências na educação básica, potencializa interações mais autônomas, adaptativas e centradas no estudante. Tal solução representa um avanço teórico para a personalização da aprendizagem e para o suporte à avaliação formativa.

Esse panorama revela que, embora existam condições materiais para a inserção de tecnologias mediadas por IA, persistem lacunas de governança que exigem investigações aplicadas focadas na validação técnica, na usabilidade e nas implicações éticas desses agentes conversacionais.

1.2.2 Relevância social e alinhamento à BNCC

A integração de agentes conversacionais no contexto escolar dialoga diretamente com as competências gerais da BNCC, notadamente: (i) Pensamento científico, crítico e criativo; (ii) Cultura digital; (iii) Comunicação; e (iv) Responsabilidade e cidadania. Em termos de design instrucional, tais ferramentas são projetadas para fomentar a interatividade, oferecendo um ambiente propício à investigação e à resolução de problemas.

Do ponto de vista da acessibilidade, a disponibilidade ininterrupta (24/7) do mediador conversacional representa um recurso técnico capaz de ampliar o suporte extraclasse.

Para o docente, agentes como o EducaChat oferecem um mecanismo de escalabilidade do atendimento, automatizando o *feedback* para dúvidas recorrentes. Tal funcionalidade visa, em tese, reduzir a sobrecarga operacional em turmas numerosas, criando condições para que o professor possa dedicar-se a mediações de maior complexidade. Sua adoção eficaz, contudo, pressupõe curadoria pedagógica rigorosa e monitoramento contínuo dos indicadores de qualidade.

Tabela 4 - Mapeamento entre BNCC e funcionalidades do EducaChat.

Competência BNCC	Exigência pedagógica	Apoio do EducaChat
Cultura digital	Uso crítico e responsável de tecnologias	Mediação guiada; registro de interações; orientação sobre fontes
Pensamento científico, crítico e criativo	Problematização e investigação	Pistas, perguntas orientadoras, exemplos
Comunicação	Clareza e adequação de linguagem	Reformulação de explicações; analogias contextualizadas
Responsabilidade e cidadania	Ética e respeito à diversidade	Avisos de privacidade; linguagem inclusiva; mitigação de vieses

Fonte: Elaborado pelo autor (2025).

A BNCC (2018) apresenta as competências e as habilidades que demandam interpretação, resolução de problemas, pensamento crítico, argumentação e autonomia, tais dimensões encontram nas tecnologias de IA um vetor de potencialização. Nesse contexto, o EducaChat, ao integrar funcionalidades de diálogo interativo, geração de exemplos contextualizados e estratégias de estudo, posiciona-se como um instrumento de apoio

tecnicamente aderente às diretrizes curriculares, oferecendo a infraestrutura digital necessária para a mediação pedagógica.

1.2.3 Risco e oportunidade

A adoção de agentes conversacionais em escolas públicas impõe um *trade-off*: por um lado, democratiza o acesso a pistas e *feedback* imediato; por outro, suscita riscos teóricos quanto à dependência tecnológica e a potenciais vieses de linguagem. Esta dissertação aborda essa dualidade ao estabelecer critérios técnicos de qualidade (balizados pela ISO/IEC 25010) como pré-requisitos para o uso responsável.

A proposta não visa substituir a mediação humana, mas sim oferecer evidências objetivas de desempenho que subsidiem a orquestração docente, garantindo que a inserção da IA ocorra sobre bases auditáveis e seguras.

1.2.4 Delimitações e escopo

Este estudo delimita-se ao ensino médio em contexto de escola pública, aplicando o EducaChat em atividades de matemática e biologia.

Nesse sentido, a pesquisa alinha-se ao texto da BNCC, que enfatiza o “reconhecimento das potencialidades das tecnologias digitais para a realização de uma série de atividades relacionadas a essas áreas do conhecimento” (Brasil, 2018, p. 474).

Dentre as competências previstos, destaca-se o uso de ferramentas digitais para “compreender e produzir conteúdos em diversas mídias [...] e elaborar e explorar diversos registros de representação matemática” (Brasil, 2018, p. 475). Especificamente para o ensino de ciências, o agente conversacional é posicionado como um recurso para auxiliar na construção de argumentos e na análise de fenômenos, operando como suporte cognitivo para que o estudante possa projetar cenários e avaliar impactos.

Em termos analíticos, a investigação do EducaChat concentra-se estritamente em indicadores observáveis nos registros de sistema (logs), a saber: (i) métricas de interatividade (frequência e duração das sessões); (ii) padrões de uso técnico; e (iii) qualidade de software segundo a norma ISO/IEC 25010 (com ênfase em eficiência e confiabilidade).

Não compõem o escopo desta pesquisa: (a) avaliação da percepção subjetiva dos estudantes; (b) análise longitudinal de aprendizagem (pré e pós-teste); ou (c) estudos de custo-efetividade econômica.

1.2.5 Limitações e escopo negativo

Fora do escopo: não compõem o objeto desta pesquisa a avaliação de desempenho discente em testes padronizados, a mensuração de impactos cognitivos de longo prazo, nem a comparação experimental entre múltiplas arquiteturas proprietárias de IA.

Riscos de interpretação: é imperativo ressaltar que os indicadores técnicos extraídos (ex.: `msg_count`, `duration_s`) constituem métricas de interatividade e não devem ser tomados como sinônimos de aprendizagem efetiva.

Sua leitura requer cautela, sendo realizada estritamente sob a ótica da qualidade de software e das condições de infraestrutura, evitando-se inferências sobre processos mentais ou subjetivos que extrapolam os dados de registro.

1.2.6 Validade, transferibilidade e limites do estudo

A validade interna é sustentada pela consistência rigorosa entre os objetivos declarados, a instrumentação de coleta automatizada (logs) e as métricas analíticas padronizadas. A validade externa apresenta limitações inerentes ao desenho de estudo de caso: os resultados refletem um contexto específico de infraestrutura pública, não sendo automaticamente generalizáveis para cenários com dotação tecnológica distinta.

Em termos de limites, não se pretende inferir efeitos cognitivos de longo prazo nem comparar arquiteturas de IA concorrentes; busca-se, prioritariamente, explicitar as condições técnicas de contorno necessárias para uma adoção responsável e estável em ambientes escolares.

1.2.7 Definições operacionais

As definições operacionais na pesquisa permitem identificar como os conceitos são compreendidos e utilizados. Além disso, contribuem para evitar ambiguidades, assegurando que as variáveis sejam passíveis de mensuração objetiva. Destacam-se as seguintes definições:

Agente conversacional educacional: sistema de IA arquitetado para interagir via linguagem natural, configurado para apoiar atividades de ensino mediante a entrega progressiva de pistas orientadoras, evitando a resposta final imediata.

Mediação pedagógica (técnica): conjunto de funcionalidades de *feedback* automatizado e encorajamento programadas no sistema para sustentar a interação, servindo de suporte à orquestração docente.

Adaptabilidade (personalização técnica): capacidade lógica do sistema. Na Figura 1, apresenta-se o encadeamento lógico entre o problema de pesquisa, os objetivos, a metodologia adotada, os dados e informações analisados, as estratégias de análise e as evidências produzidas, possibilitando uma compreensão integrada da organização e da coerência interna da pesquisa.

Figura 1- Encadeamento lógico entre problema, objetivos, método e evidências.



Fonte: Elaborado pelo autor (2025).

Para consolidar o encadeamento lógico da pesquisa, apresenta-se na próxima subseção o quadro relacional entre os conceitos de mediação, adaptabilidade e qualidade técnica.

1.2.8 Tabela lógica entre mediação, personalização e qualidade

A Tabela 5 apresenta as relações entre os conceitos operacionais e as respectivas estratégias de verificação empírica. São detalhados os conceitos, a definição operacional no contexto do sistema e a métrica de análise via logs:

Tabela 5 – Relações entre conceitos operacionais e implicações analíticas.

Conceito	Definição Operacional (Sistema)	Como será analisado (Métricas de Log)
Mediação pedagógica	Funcionalidades de pistas, feedbacks graduais e checagens programadas para sustentar a interação	Volumetria de mensagens (msg_count); Profundidade da árvore de diálogo; Taxa de retenção na sessão
Adaptabilidade (Personalização)	No sistema de IA, ajuste dinâmico em tempo real de detalhamento, exemplos, linguagem, ritmo e sequência das respostas, conforme uso e interações do aluno	Análise de ramificação dos diálogos; Variação de tokens de resposta; Histórico de contextos acionados
Qualidade de Software (ISO/IEC 25010)	Atributos de eficiência de desempenho, confiabilidade e operabilidade técnica	Métricas de logs: tempo de resposta e throughput (eficiência de desempenho), estabilidade e recuperação de falhas (confiabilidade), facilidade de operação e monitoramento (operabilidade), além de registros de incidentes e uso
Mediação pedagógica (Técnica)	Capacidade lógica do algoritmo de ajustar o fluxo (exemplos/linguagem) baseado no input do usuário	Latência média (ms); Taxa de erros de sistema (crash/timeout); Disponibilidade do serviço (uptime)

Fonte: Elaborado pelo autor (2025).

A estruturação desses conceitos proporciona a definição operacional necessária para uma análise objetiva, articulando significados teóricos com evidências empíricas extraídas diretamente da base de dados. Essa coerência terminológica e clareza metodológica garantem que cada variável seja verificada de forma consistente e auditável ao longo da investigação. A seguir, discutem-se as questões éticas e a governança de IA.

1.2.9 Questões éticas e governança de IA

O uso de IA em educação demanda salvaguardas quanto a privacidade, transparência e não discriminação. Nesta pesquisa: (i) minimizam-se dados pessoais, restringindo-se a registros necessários para análise pedagógica; (ii) informam-se propósitos e limites de uso aos participantes; (iii) monitoram-se vieses de linguagem e conteúdos inadequados; e (iv) prevê-se supervisão docente contínua.

O estudo não visa substituir a mediação humana, mas ampliar o alcance do apoio pedagógico preservando princípios éticos e legais aplicáveis ao contexto escolar.

A Tabela 6 apresenta as lacunas identificadas na literatura e a abordagem adotada na pesquisa.

Tabela 6 – Lacunas identificadas na literatura e a abordagem adotada nesta pesquisa.

Lacuna	Evidência resumida	Abordagem nesta dissertação
Pouca evidência em contexto escolar público brasileiro	Predominam estudos em HE/EA ou em ambientes controlados.	Aplicação real em escola pública, com coleta em situação autêntica.
Avaliação focada apenas em satisfação	Questionários de utilidade sem triangulação.	Métodos mistos: logs, indicadores, Likert e entrevistas.
Desalinhamento pedagógico	Agentes respondem demais; pouca aprendizagem ativa.	Pistas graduais, analogias e supervisão docente (orquestração).
Qualidade de software difusa	Métricas heterogêneas, pouco comparáveis.	Subconjunto ISO/IEC 25010 com indicadores factíveis em campo.

Fonte: Elaborado pelo autor (2025).

O uso de agentes conversacionais na educação, conforme as pesquisas desenvolvidas, apresenta lacunas conceituais, metodológicas e empíricas que justificam a realização desta pesquisa (Fernandes, 2023; Moraes, 2023; Silva, 2022; Silveira, 2022). De modo geral, os estudos existentes concentram-se na descrição de funcionalidades tecnológicas ou na avaliação de ganhos cognitivos pontuais, mas carecem de validação técnica rigorosa sobre a infraestrutura e a confiabilidade necessárias para sustentar o uso pedagógico desses sistemas em escala.

1.2.9.1 Cenário de uso ilustrativo

Esta seção apresenta um cenário de uso ilustrativo e hipotético, com o objetivo de exemplificar, de forma conceitual, como o chatbot educacional EducaChat pode ser integrado a práticas pedagógicas mediadas por tecnologia, sem caracterizar relato empírico de aplicação com sujeitos de pesquisa.

Em uma sequência didática hipotética de matemática sobre funções, o professor propõe exercícios graduais e disponibiliza o EducaChat como recurso de apoio à mediação pedagógica.

Nesse contexto ilustrativo, o agente conversacional é utilizado para fornecer pistas orientadoras — e não respostas prontas —, reforçar conceitos pré-requisitos e sugerir exemplos análogos, enquanto o docente acompanha registros técnicos resumidos do sistema para ajustar sua mediação em sala de aula.

Sob a perspectiva do conectivismo, a aprendizagem em rede nessa sequência didática pressupõe a construção de conexões entre diferentes nós de informação. Conforme Siemens (2010), aprender envolve construir, navegar e utilizar redes de conhecimento, nas quais a qualidade das conexões influencia diretamente os processos de troca e construção do saber.

Em biologia, ao trabalhar ecologia de ecossistemas (ou genética básica), o EducaChat é programado para oferecer perguntas orientadoras e analogias — por exemplo, comparando cadeias alimentares a fluxos de energia —, além de aplicar checagens rápidas de compreensão (*quizzes*). A eficácia desse desenho é verificada não por observação comportamental, mas pela análise das métricas de interatividade e completude das sessões registradas no banco de dados, indicando se houve engajamento técnico com o material proposto.

Nesse sentido, a integração do EducaChat a ambientes educacionais é compreendida como uma possibilidade alinhada ao crescimento exponencial das conexões entre fontes de informação para a produção do conhecimento, considerando que, conforme Siemens (2010, p. 21), “quanto melhor a qualidade da rede e das conexões, melhor será o resultado da troca de conhecimento”.

1.3 Originalidade e impacto esperado

O diferencial deste estudo reside na integração entre fundamentos pedagógicos, alinhamento à BNCC e avaliação técnica-documental, ancorada em critérios de qualidade de software definidos pela ISO/IEC 25010. A originalidade da pesquisa está no desenvolvimento e na análise do EducaChat como um agente conversacional educacional concebido especificamente para apoiar processos de mediação pedagógica, com foco em pistas graduais, supervisão docente e estabilidade técnica.

A investigação propõe um modelo analítico replicável, baseado em dados secundários anonimizados, que articula métricas técnicas de desempenho e uso a referenciais pedagógicos consolidados.

Espera-se, assim, produzir subsídios teóricos e práticos para gestores, docentes e desenvolvedores, além de um blueprint de adoção responsável de agentes conversacionais educacionais, alinhado a princípios éticos, curriculares e de governança de dados.

1.3.1 Problema de pesquisa (versão expandida) e derivação das perguntas

O problema que orienta esta dissertação pode ser formulado como: em que medida e

sob quais condições de infraestrutura técnica, agentes conversacionais educacionais, concebidos para operar com pistas graduais, apresentam indicadores satisfatórios de qualidade de software (ISO/IEC 25010), adaptabilidade verificável e sustentação da interatividade no contexto do Ensino Médio público?

1.4 Objetivo Geral

Esta dissertação tem como objetivo principal desenvolver e validar tecnicamente o EducaChat, um agente conversacional baseado em inteligência artificial, avaliando sua qualidade de *software* e sua viabilidade operacional em contexto escolar público. Pretende-se analisar, por meio de métricas quantitativas de uso, como a arquitetura do sistema suporta a adaptabilidade técnica (personalização via algoritmo) e a sustentação da interatividade, verificando sua aderência aos requisitos de infraestrutura em atividades de matemática e biologia.

1.4.1 Objetivos específicos

Para alcançar o objetivo geral, definem-se as seguintes metas operacionais:

Desenvolver e instrumentar o EducaChat em um ambiente educacional real, configurando rotinas de coleta automatizada de dados (*logging*) para auditoria de uso.

Avaliar a qualidade técnica da solução com base na norma ISO/IEC 25010, mensurando indicadores de confiabilidade, eficiência de desempenho e a capacidade do sistema em sustentar a interatividade (retenção).

Verificar a adaptabilidade lógica do agente, analisando nos registros de banco de dados se os mecanismos de pistas graduais e personalização de exemplos foram acionados conforme o planejado.

Mapear as condições de contorno (barreiras de infraestrutura e latência) que impactam a adoção de agentes conversacionais em escolas públicas, propondo diretrizes técnicas para mitigar tais gargalos.

Nesse cenário, este trabalho busca responder a uma lacuna pragmática: quais são os requisitos de qualidade e as evidências de desempenho necessárias para que agentes conversacionais possam operar de forma estável e segura no ensino médio.

Os objetivos específicos que orientaram esta auditoria técnica foram sistematizados na

Seção 1.4.1, a qual detalha as etapas de desenvolvimento, instrumentação de coleta e validação analítica da solução proposta.

1.4.2 Questões de pesquisa

Com base no objetivo geral de validação técnica e nos objetivos específicos delineados, a investigação estrutura-se para responder como a arquitetura do agente conversacional se comporta diante das restrições do ambiente escolar.

Diante do exposto e em consonância com os objetivos específicos, formulam-se as seguintes questões norteadoras:

QP1(Interatividade Técnica): os indicadores de uso do sistema (frequência, volume de mensagens e tempo de sessão) evidenciam uma sustentação da interatividade compatível com a realização de sequências didáticas no ensino médio?

QP2(Adaptabilidade Lógica): a arquitetura do EducaChat executa as rotinas de personalização planejadas (pistas graduais e ramificação de exemplos) de forma verificável nos registros de log, superando a entrega de respostas estáticas?

QP3(Qualidade e Infraestrutura): em que medida o desempenho técnico da solução (latência, disponibilidade e taxas de erro) atende aos requisitos de confiabilidade da norma ISO/IEC 25010 quando submetido às restrições de conectividade da escola pública?

A relevância deste estudo também se sustenta em dados empíricos recentes. De acordo com o TIC educação 2023, 92% das escolas brasileiras já possuem algum tipo de conectividade, mas ainda persistem desigualdades críticas quanto à qualidade da conexão e ao uso pedagógico efetivo das tecnologias (CETIC.BR, 2024).

No que se refere ao público discente, levantamentos indicam que cerca de 95% das crianças e adolescentes entre 9 e 17 anos são usuários de Internet no país, reforçando a urgência de validar tecnicamente práticas mediadas por tecnologias digitais (CETIC.BR, 2024).

Paralelamente, diretrizes internacionais, como as da UNESCO, alertam para a necessidade de salvaguardas éticas e protocolos de governança técnica em IA para contextos educacionais (MIAO; HOLMES, 2023).

1.4.3 Evidências que respondem às QPs

Para responder às questões de pesquisa (QPs) propostas neste estudo, foram reunidas

evidências estritamente objetivas, provenientes dos registros automatizados de log do sistema e do monitoramento de performance do servidor. A análise consolidada desses dados permitiu verificar tecnicamente como a arquitetura do EducaChat sustentou a interatividade, executou rotinas de adaptabilidade e comportou-se diante da infraestrutura escolar. A seguir, apresentam-se as evidências técnicas que fundamentam as respostas às QPs:

QP1(Interatividade Técnica): indicadores de volumetria (*msg_count*), tempo de sessão (*duration_s*) e curvas de retenção de usuários, demonstrando a capacidade do sistema em manter o fluxo de troca de mensagens.

QP2(Adaptabilidade Lógica): mapeamento dos nós de diálogo acionados na árvore de decisão, evidenciando a execução de ramificações condicionais (pistas vs. respostas diretas) programadas no algoritmo.

QP3(Qualidade e Infraestrutura): métricas de desempenho segundo a ISO/IEC 25010, incluindo latência média de resposta (ms), taxas de erro de requisição (*status code 4xx/5xx*) e disponibilidade do serviço (*uptime*).

Assim, as evidências apresentadas demonstram que o conjunto de dados quantitativos analisados oferece suporte robusto para validar a viabilidade técnica da solução. Elas revelam não percepções subjetivas, mas fatos operacionais sobre o funcionamento do EducaChat, provendo uma base auditável para discutir as potencialidades e os gargalos da adoção de IA em contextos de infraestrutura restrita.

1.4.4 Contribuições e organização da dissertação

As contribuições deste estudo sintetizam-se em: (i) um protótipo funcional de agente conversacional educacional tecnicamente aderente à BNCC; (ii) um protocolo de validação em contexto escolar público, fundamentado exclusivamente na análise quantitativa de logs; (iii) evidências empíricas objetivas sobre a sustentação da interatividade e a adaptabilidade lógica do sistema; e (iv) uma aplicação pragmática de métricas de qualidade de *software* baseadas na norma ISO/IEC 25010.

Tabela 7– Contribuições práticas, teóricas e metodológicas.

Dimensão	Contribuições
Prática	Blueprint de adoção responsável de agentes conversacionais; recomendações de orquestração docente.
Teórica	Evidências sobre mediação e personalização em contexto escolar brasileiro.
Metodológica	Protocolo reproduzível: logs + indicadores ISO/IEC 25010 + análise categorial.

Fonte: Elaborado pelo autor (2025).

Esta dissertação está organizada em oito capítulos, além dos elementos pré e pós-textuais:

- Capítulo 1 – Introdução: contexto, justificativa, objetivos, questões de pesquisa e contribuições.
- Capítulo 2 – Fundamentação Teórica: conceitos, estado da arte em agentes conversacionais e mediação pedagógica.
- Capítulo 3 – Metodologia: natureza da pesquisa, delineamento, instrumentos e plano de análise.
- Capítulo 4 – Modelagem do Sistema: requisitos, fluxos de interação e arquitetura do EducaChat.
- Capítulo 5 – Implementação do Sistema: detalhes técnicos da construção do protótipo e integrações.
- Capítulo 6 – Resultados: métricas operacionais, estatísticas descritivas e achados principais.
- Capítulo 7 – Discussão: interpretação dos resultados à luz da literatura e das QPs.
- Capítulo 8 – Conclusão: contribuições, limitações, perspectivas futuras.

A seguir, na próxima seção, aprofunda-se a discussão referente à organização desta dissertação, situando o leitor no percurso teórico, metodológico e analítico que sustenta o

desenvolvimento do EducaChat e sua avaliação no contexto educacional investigado. Essa estrutura foi planejada de modo a conduzir, de forma progressiva e integrada, desde a fundamentação conceitual até a apresentação dos resultados e reflexões decorrentes da aplicação prática do protótipo.

1.4.5 Estrutura lógica do trabalho

A organização desta dissertação segue um encadeamento lógico entre problema, método de validação e evidências técnicas. A introdução estabelece o contexto de infraestrutura, a justificativa e os objetivos de auditoria de *software*; a fundamentação apresenta o estado da arte e o quadro conceitual sobre IA e qualidade técnica; a metodologia detalha o desenho do estudo de caso e os instrumentos de coleta automatizada (logs); a implementação descreve a arquitetura da solução e sua aplicação em atividades de matemática e biologia.

Os resultados apresentam a análise quantitativa dos indicadores de desempenho e interatividade, discutidos à luz da norma ISO/IEC 25010; e a conclusão sintetiza as contribuições técnicas, as limitações do estudo e a agenda para pesquisas futuras.

Caso o leitor deseje uma visão gráfica do fluxo, recomenda-se consultar a Figura 2 (fluxo lógico da dissertação), que mapeia a correlação entre objetivos, capítulos e as evidências objetivas geradas, facilitando a navegação pelo documento.

Figura 2 - Fluxo lógico da dissertação.



Fonte: Elaborado pelo autor (2025).

A Figura 2 sintetiza o encadeamento lógico que orienta a construção desta dissertação, evidenciando como cada capítulo cumpre uma função específica no avanço da investigação, desde a fundamentação teórica até a validação técnica e funcional do EducaChat. Esse percurso estruturado permite compreender de forma progressiva os elementos que sustentam o desenvolvimento, a instrumentação de coleta de dados e a análise dos resultados.

Concluída esta visão geral da organização do trabalho, o Capítulo 2 – Fundamentação Teórica aprofunda as bases científicas da investigação, discutindo o estado da arte sobre agentes conversacionais na educação, os princípios do conectivismo e os requisitos de qualidade de software (ISO/IEC 25010) que norteiam a análise dos dados.

1.4.6 Hipóteses analíticas

As hipóteses a seguir derivam diretamente das questões de pesquisa (QPs) e orientam a validação estatística dos dados extraídos dos registros do sistema (logs):

H1(Sustentação da interatividade): a arquitetura de diálogo baseada em pistas graduais resulta em métricas de volume de mensagens (`msg_count`) e duração de sessão (`duration_s`) estatisticamente significativas, evidenciando que o sistema sustenta a interação por tempo compatível com a sequência didática proposta.

H2(Efetividade da adaptabilidade lógica): a análise da árvore de decisão demonstrará que o algoritmo executa ramificações distintas (acionamento de pistas de nível 1, 2 ou 3) em resposta aos inputs do usuário, superando o comportamento de entrega de respostas estáticas ou genéricas.

H3(Correlação qualidade-retenção): existe uma correlação positiva entre os indicadores de Qualidade de Software (baixa latência e alta disponibilidade/uptime) e a taxa de retenção dos usuários na sessão; inversamente, picos de latência ou erros técnicos (crashes) resultam em abandono imediato da interação.

H4(Impacto da infraestrutura): as condições técnicas de contorno (qualidade da conexão local medida pelo timestamp de resposta) atuam como variável moderadora, onde a degradação da rede limita fisicamente o volume de interações possíveis, independentemente do desenho pedagógico.

Tais hipóteses orientam a análise quantitativa dos dados e a discussão sobre a viabilidade técnica proposta nos capítulos posteriores.

2. FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta os principais conceitos, teorias e abordagens que sustentam o desenvolvimento e a validação técnica de agentes conversacionais educacionais. A fundamentação estabelece os elementos para compreender a arquitetura dessas tecnologias, deslocando o foco da promessa mágica de ensino para a análise da adaptabilidade algorítmica, da interatividade homem-máquina e dos requisitos de qualidade de *software* necessários para sua operação em escala.

2.1 Histórico da inteligência artificial na educação (IAED)

O uso da inteligência artificial na educação (IAED) constitui um campo interdisciplinar que vem sendo explorado há mais de cinco décadas, combinando avanços em ciência da computação, psicologia cognitiva, pedagogia e engenharia de *software*. Sua trajetória histórica permite compreender como os atuais agentes conversacionais, como o EducaChat, se inserem em uma linha de evolução que remonta aos primeiros sistemas tutores inteligentes.

À medida que a inteligência artificial avança, torna-se evidente que a eficácia dessas ferramentas não reside apenas na sofisticação do algoritmo, mas na sua estabilidade e disponibilidade. Diferente das primeiras máquinas de ensinar, os sistemas atuais exigem uma infraestrutura robusta de dados. Segundo Holmes et al. (2019), o desafio contemporâneo desloca-se da simples criação de ferramentas para a auditoria de seus processos, garantindo que a "caixa preta" da IA opere de forma transparente, auditável e tecnicamente confiável dentro do ambiente escolar.

2.1.2 Aprendizagem ubíqua e mediação digital

A literatura sobre ubiquitous learning (aprendizagem ubíqua) destaca que a integração de dispositivos móveis e conectividade contínua amplia o potencial de adaptabilidade do sistema, permitindo que o suporte técnico ao aluno ocorra para além da sala de aula física (Coelho et al., 2023; Araújo; Souza; Martins, 2014).

Tabela 8 – Comparativo entre aprendizagem tradicional, digital e ubíqua.

Dimensão Técnica	Aprendizagem Digital (E-learning clássico)	Aprendizagem Ubíqua (u-learning / Agentes)
Infraestrutura de Acesso	Fixa (Desktop/Laboratório de Informática)	Móvel (Smartphones/Tablets) em redes heterogêneas
Disponibilidade (Tempo)	Assíncrona ou agendada (Horários fixos)	Contínua (24/7) – Requer Alta Disponibilidade de Servidor
Interatividade	Navegação em conteúdos estáticos	Interação dialógica com Agentes Inteligentes
Adaptabilidade (Lógica)	Conteúdo padronizado por perfil de usuário	Ajuste algorítmico em tempo real (Context-Aware)

Fonte: Adaptado de Araujo, Souza e Martins (2014).

Nesse contexto, agentes conversacionais funcionam como interfaces de alta disponibilidade. Contudo, para que a aprendizagem ubíqua se concretize, o sistema deve garantir baixa latência e alta tolerância a falhas. O agente não atua apenas como um "tutor", mas como um nó de rede sempre ativo, capaz de entregar pistas e *feedbacks* imediatos (*just-in-time*) sem interrupções de serviço, quesito central analisado nesta pesquisa sob a ótica da norma ISO/IEC 25010.

2.1.3 IA na educação: trajetória e inserção

A inserção da IA em contextos educacionais pode ser organizada em quatro fases principais, conforme adaptação de Holmes, Bialik e Fadel (2019) e Miao e Holmes (2023):

1. Sistemas tutores inteligentes(décadas de 1980–1990): modelos baseados em regras que buscavam diagnosticar erros e oferecer *feedback* personalizado.
2. Plataformas adaptativas(anos 2000): uso de *machine learning* para ajustar o nível de dificuldade com base no histórico de respostas do estudante.
3. Aprendizado profundo e big data(anos 2010): integração de *dashboards*, *learning analytics* e análise preditiva de desempenho.
4. Agentes generativos(2020 em diante): incorporação de *large language models*

(*ChatGPT, Gemini, Copilot*) para simular diálogo natural e mediação pedagógica em tempo real.

No contexto brasileiro, iniciativas recentes incluem o uso experimental de IA em processos avaliativos (ex.: Enem digital), projetos-piloto em redes estaduais e a incorporação de ferramentas de apoio à docência no Sistema S (Senai, Senac). Tais movimentos revelam tanto oportunidades quanto a necessidade de marcos regulatórios claros e protocolos pedagógicos robustos.

2.1.4 Anos 1970: os primeiros sistemas tutores inteligentes

Nos anos 1970, emergiram os sistemas tutores inteligentes (STIs), pioneiros na tentativa de automatizar rotinas de instrução. Diferente dos sistemas atuais baseados em probabilidade, os STIs operavam sobre árvores de decisão lógicas: buscavam simular o papel de um tutor mapeando as ações do estudante contra um "modelo ideal" de especialista, identificando desvios e entregando correções explícitas.

Um exemplo clássico foi o projeto *SCHOLAR*, desenvolvido por Jaime Carbonell em 1970, que explorava técnicas de processamento de linguagem natural para responder a perguntas de alunos sobre geografia. Embora ainda rudimentar, esse marco inicial demonstrou o potencial da IA para apoiar processos de ensino-aprendizagem de maneira personalizada.

2.1.5 Anos 1980 e 1990: consolidação de abordagens cognitivas

Entre as décadas de 1980 e 1990, a IAED avançou com base em teorias cognitivas de aprendizagem, como o construtivismo de Piaget e a teoria da instrução de Bruner. Surgiram sistemas que incorporavam modelos de aluno, capazes de registrar conhecimento prévio, erros recorrentes e estilos de aprendizagem.

A personalização passou a ser mais refinada, permitindo que o sistema ajustasse não apenas o conteúdo, mas também a estratégia pedagógica. A ênfase recaiu na tentativa de representar o raciocínio humano em regras explícitas, por meio de sistemas especialistas.

2.1.6 Anos 2000: e-learning e ambientes virtuais

Com a expansão da internet no início dos anos 2000, a IAED se integrou aos ambientes

virtuais de aprendizagem (AVA), como *Moodle* e *Blackboard*. Embora esses sistemas não fossem essencialmente inteligentes, criaram espaço para a inclusão de módulos adaptativos baseados em IA, como sistemas de recomendação de conteúdos e análises preditivas de desempenho. A noção de aprendizagem ubíqua (*ubiquitous learning*) ganhou força, destacando a importância de oferecer suporte educacional em múltiplos dispositivos e contextos.

2.1.7 Anos 2010: aprendizado de máquina e análise de dados educacionais

Na década de 2010, o avanço do aprendizado de máquina (*machine learning*) e do *big data* educacional ampliou significativamente o escopo da IAED. Plataformas passaram a registrar grandes volumes de dados sobre o comportamento dos estudantes: tempo de estudo, interações em fóruns, respostas em *quizzes* e até mesmo padrões de escrita. Esses dados permitiram o desenvolvimento de sistemas preditivos capazes de identificar estudantes em risco de evasão e sugerir intervenções pedagógicas. Nesse período, também ganharam destaque os *chatbots* educacionais baseados em regras, projetados para fornecer respostas rápidas a dúvidas frequentes e auxiliar no suporte administrativo.

2.1.8 Anos 2020: a revolução da IA generativa

A partir de 2020, com a popularização dos grandes modelos de linguagem (LLMs), como *GPT*, *BERT* e *LLaMA*, a IAED entrou em uma nova fase. Diferentemente dos *chatbots* baseados em regras ou em técnicas estatísticas limitadas, os modelos generativos são capazes de produzir textos coerentes, traduzir idiomas, sintetizar explicações e até mesmo elaborar *feedbacks* personalizados em tempo real. Durante a pandemia de *COVID-19*, o uso dessas tecnologias acelerou de forma exponencial, dado o fechamento das escolas e a necessidade de ensino remoto. Nesse contexto, *chatbots* educacionais passaram a ser vistos não apenas como assistentes de suporte, mas como potenciais mediadores do processo de ensino-aprendizagem.

2.1.9 Síntese histórica

A evolução da IAED pode ser compreendida em quatro grandes ondas: (i) sistemas tutores inteligentes (1970–1990), com foco em simular tutores humanos; (ii) ambientes

virtuais adaptativos (2000–2010), apoiados pela internet e *e-learning*; (iii) uso de aprendizado de máquina e *big data* (2010–2020), ampliando a personalização e análise

preditiva; e (iv) agentes conversacionais baseados em IA generativa (2020–presente), caracterizados pela interação multimodal e pela capacidade de adaptação em tempo real.

Essa trajetória evidencia que, embora os fundamentos conceituais permaneçam vinculados à ideia de personalização e suporte ao aluno, os avanços tecnológicos têm expandido progressivamente as possibilidades de aplicação. A pesquisa atual, ao propor o EducaChat, insere-se nessa quarta onda, buscando explorar o potencial da IA generativa para fins pedagógicos, mas também enfrentar desafios éticos, técnicos e metodológicos que permanecem em aberto.

Dessa forma, observa-se que a inteligência artificial aplicada à educação consolidou-se ao longo de diferentes ondas tecnológicas, cada uma ampliando as possibilidades de suporte ao processo de ensino-aprendizagem. Nesse contexto mais amplo, os agentes conversacionais e *chatbots* emergem como uma das manifestações mais recentes e promissoras da *IAED*, tema abordado na próxima seção.

2.2 Histórico dos agentes conversacionais e *chatbots*

O desenvolvimento de agentes conversacionais é marcado por uma trajetória que reflete avanços tanto na inteligência artificial quanto na interação humano–máquina. O primeiro marco relevante foi o *ELIZA*, criado por Joseph Weizenbaum em 1966, que simulava um psicoterapeuta ao reconhecer padrões simples de texto.

Embora limitado, este sistema inaugurou a percepção de que computadores poderiam estabelecer interações em linguagem natural (Russell; Norvig, 2013).

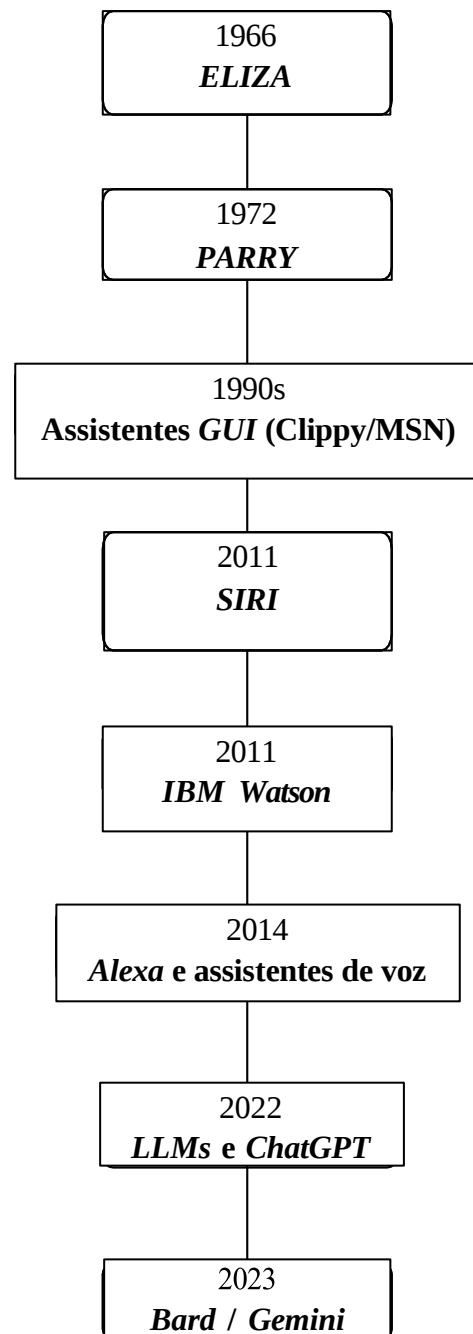
Abrindo espaço para reflexões posteriores sobre o potencial pedagógico de sistemas capazes de dialogar — uma discussão retomada com força por autores como Woolf (2010) ao explorar agentes tutores inteligentes.

Na década de 1970 surgiu o *PARRY*, voltado para simular pacientes psiquiátricos, ampliando a discussão sobre usos clínicos da IA. Nos anos 1980 e 1990, o avanço de interfaces gráficas possibilitou a popularização de agentes animados, como a assistente *Clippy* da *Microsoft*. Já nos anos 2000, com a difusão da internet, proliferaram *chatbots* de atendimento, inicialmente baseados em regras pré-programadas, fenômeno que também influenciou o uso de interfaces dialogais em plataformas educacionais e ambientes virtuais de aprendizagem — contexto discutido por autores como Selwyn (2016) ao analisar o papel das tecnologias digitais na educação contemporânea.

O salto qualitativo mais expressivo ocorreu a partir de 2015, com os avanços em aprendizado profundo, que permitiram diálogos mais complexos e sensíveis ao contexto. Assistentes como *Siri*, *Alexa* e *Google Assistant* consolidaram o uso cotidiano desses sistemas. Na educação, esse período marca a intensificação das pesquisas sobre IA aplicada ao ensino, especialmente no campo da personalização e da mediação automatizada (Luckin et al., 2016; Holmes et al., 2019; Zawacki-Richter et al., 2019).

Mais recentemente, os modelos de linguagem de larga escala (*LLMs*), como o *ChatGPT*, inauguraram uma nova era, aproximando a capacidade dos *chatbots* de interações humanizadas, contextuais e responsivas. Essa mudança tem provocado profundas discussões no campo educacional sobre potencial, riscos, ética e governança da (Holmes; Mayo; Wei, 2023; Williamson; Pugliesi, 2023), reforçando a necessidade de compreender esses agentes no contexto do ensino, da aprendizagem e da mediação pedagógica.

Figura 3 – Linha do tempo de agentes conversacionais (1966–2023).



Fonte: Elaborado pelo autor (2025).

Como mostra a Figura 3, a evolução dos agentes conversacionais sai de sistemas baseados em regras (*ELIZA/PARRY*) para *LLMs*, ampliando personalização e mediação pedagógica.

2.2.1 *Chatbots* educacionais no contexto BNCC e generativo

A literatura técnica internacional apresenta um conjunto expressivo de experimentos envolvendo agentes conversacionais, destacando a capacidade arquitetural desses sistemas em operacionalizar rotinas de tutoria inteligente e personalização de trilhas. Estudos de referência, como os de VanLehn (2011) e Woolf (2010), demonstram que a automação de *feedbacks* imediatos permite que o sistema emule, em tarefas estruturadas, a dinâmica de perguntas e respostas típica da tutoria humana. Revisões sistemáticas recentes (Zawacki-Richter et al., 2019; Sandoval-Reyes et al., 2023) indicam que a eficácia técnica desses agentes reside na consistência da entrega de instruções passo a passo, fundamental para o suporte em resolução de problemas.

No Brasil, a adoção dessas ferramentas alinha-se diretamente à competência geral 5 da BNCC (Cultura digital), que preconiza o uso crítico e criativo das tecnologias para a resolução de problemas (Brasil, 2018). Iniciativas nacionais (Guerreiro; Barros, 2019; Mattos et al., 2022) evidenciam que a implementação de agentes conversacionais oferece uma camada de infraestrutura de apoio, permitindo que estudantes acessem orientações fora do horário de aula (ubiquidade) e com tempo de resposta reduzido, fatores que tendem a favorecer a continuidade da interação (retenção).

Entre as funcionalidades técnicas observadas na literatura aplicada aos componentes curriculares, destacam-se:

Matemática (funções e estatística): implementação de algoritmos de pistas graduais (*scaffolding*), onde o agente é programado para não entregar a resposta final, mas sim desdobrar o problema em etapas lógicas. Estudos como os de Harsley, Smith e Johnson (2017) validam a viabilidade técnica desse modelo para sustentar a prática de exercícios sem intervenção humana direta.

Biologia (ecologia e genética): uso de *LLMs* para a geração dinâmica de analogias e cenários hipotéticos. Trabalhos recentes (Fernandes, 2023) indicam que a capacidade do sistema de processar linguagem natural facilita a estruturação de conceitos complexos, oferecendo uma interface de consulta mais fluida do que mecanismos de busca tradicionais.

Educação a Distância (EaD): pesquisas apontam o papel dos *chatbots* como agentes de primeira camada de suporte, filtrando dúvidas recorrentes e otimizando o fluxo de atendimento, o que impacta positivamente as métricas de acesso e permanência no ambiente virtual.

Com a consolidação dos modelos generativos de larga escala (*ChatGPT*, *Gemini*, *Copilot*), as possibilidades de engenharia de prompts educacionais ampliaram-se significativamente, permitindo tecnicamente:

- A geração de explicações contextualizadas sob demanda (baixa latência).
- A adequação lexical automática ao nível de complexidade do diálogo.
- A variação algorítmica de exemplos para evitar repetições mecânicas.
- A simulação interativa de cenários para a validação de hipóteses pelos estudantes.

Autores como Holmes, Bialik e Fadel (2019), Holmes e Luckin (2023) e Luckin et al. (2016) discutem que tais tecnologias flexibilizam a aprendizagem, mas exigem novas formas de orquestração docente (Dillenbourg, 2013) para garantir alinhamento curricular, verificação de qualidade e sustentabilidade pedagógica.

Entretanto, o avanço dos modelos generativos também traz desafios. Estudos como Kasneci et al. (2023) e Xu et al. (2024) alertam para riscos relacionados a:

- Alucinações e erros factuais, especialmente em explicações científicas.
- Viéses culturais e representacionais, que podem reproduzir desigualdades.
- Dependência excessiva e redução da autonomia estudantil.
- Fragilidade na rastreabilidade das fontes, dificultando avaliação crítica.
- Necessidade de supervisão docente contínua, sobretudo em atividades avaliativas.

Pesquisas recentes no campo da ética da IA em educação (Lund; Wang, 2023; Williamson, 2017) reforçam que o uso pedagógico de *chatbots* deve estar alinhado aos princípios de transparência, responsabilidade e equidade — dimensões fundamentais tanto na BNCC quanto nas diretrizes internacionais da UNESCO (Miao; Holmes, 2023).

Assim, o cenário atual aponta para uma fase de transição: de *chatbots* baseados em regras e fluxogramas para agentes generativos capazes de dialogar com maior profundidade semântica. No contexto brasileiro, esses sistemas têm potencial para apoiar competências da BNCC relacionadas ao pensamento crítico, letramento digital, resolução de problemas, investigação e uso ético da tecnologia — desde que acompanhados de políticas institucionais

claras, formação docente adequada e mecanismos de avaliação contínua.

2.2.2 Panorama internacional e nacional

No cenário internacional, a União Europeia e os Estados Unidos têm incentivado políticas que associam *chatbots* ao desenvolvimento de *soft skills*, à aprendizagem personalizada e à educação digital (Miao; Holmes, 2023). De forma convergente, a UNESCO tem destacado o papel da inteligência artificial na promoção do acesso inclusivo e responsável, ao mesmo tempo em que ressalta preocupações éticas e relacionadas à privacidade.

No Brasil, pesquisas coordenadas por instituições como o CETIC.br revelam que, apesar do avanço da conectividade nas escolas (92% em 2023), ainda persistem desigualdades na velocidade de acesso e no uso pedagógico (CETIC.BR, 2024). Estudos nacionais apontam que agentes conversacionais podem reduzir barreiras de aprendizagem em contextos de ensino remoto e híbrido, além de apoiar professores em atividades de tutoria.

A **Tabela 9** aponta um panorama comparativo entre as iniciativas, destacando que, enquanto países desenvolvidos debatem regulações éticas avançadas, o Brasil ainda enfrenta o desafio primário do acesso. No cenário internacional, a trajetória é marcada pela busca de compliance (conformidade) com regulações de IA (como o EU AI Act). Já no contexto brasileiro, pesquisas revelam que a barreira não é apenas a falta de hardware, mas a instabilidade da rede, que inviabiliza aplicações que exigem baixa latência. Estudos como o de Amorim (2024) reforçam que esses obstáculos infraestruturais condicionam a adoção de qualquer tecnologia emergente.

Tabela 9 - Comparativo internacional e nacional no uso de agentes conversacionais em educação.

Contexto	Iniciativas	Desafios
Internacional	Regulação de IA (AI Act); Tutoria inteligente em campi de alta conectividade.	Privacidade de dados; Viés algorítmico; Transparência da "caixa preta".
Brasil	Projetos-piloto descentralizados; Uso via dados móveis (4G/5G) precários.	Latência de rede; Obsolescência de <i>hardware</i> ; Disparidade de acesso regional.

Fonte: UNESCO (2023); CETIC.br (2023).

Essa comparação (Tabela 9) reforça que o uso de *chatbots* educacionais deve ser analisado sob duas perspectivas complementares: a tecnológica, voltada à infraestrutura de

suporte, e a funcional, relacionada à capacidade do software de operar estavelmente nessas condições adversas.

2.2.3 Referencial teórico: learning analytics, governança de dados e avaliação de qualidade em agentes conversacionais

A reformulação metodológica desta pesquisa — concentrando-se na extração de dados automatizada (logs) — insere-se na interseção entre o campo de *learning analytics* (LA) e a Engenharia de *Software*.

Enquanto o *learning analytics* é tradicionalmente definido como a medição e análise de dados sobre aprendizes para otimizar a aprendizagem (Siemens; Long, 2011), esta dissertação adota uma abordagem de "LA Orientado ao Sistema". Nesse recorte, os dados não são utilizados para inferir estados mentais subjetivos, mas para auditar a qualidade da interação técnica.

Para sustentar essa análise, adota-se a norma ISO/IEC 25010 (Quality Models), que estabelece critérios objetivos para avaliar produtos de *software*. Ao cruzar as métricas de LA (como frequência e retenção) com os atributos da ISO (como eficiência de desempenho e confiabilidade), constrói-se um modelo de avaliação robusto, capaz de responder se o agente conversacional é tecnicamente viável para a realidade da escola pública. Para estruturar projetos em LA, Greller e Drachsler (2012) propõem um framework genérico (o que, quem, por que, como) que ajuda a traduzir dados educacionais em indicadores úteis e eticamente sustentáveis. A presente investigação se alinha a esse arcabouço ao privilegiar registros já existentes (logs e métricas técnicas), evitando novas coletas com pessoas e adotando princípios de minimização e anonimização.

2.2.4 Learning analytics com dados secundários e logs técnico-operacionais

Conforme a literatura de LA e mineração de dados educacionais, logs de sistemas (p. ex., contagem de mensagens por sessão, eventos de erro, latência de resposta) podem funcionar como proxies válidos de engajamento, uso e qualidade do serviço (Lang, 2017). Tais dados permitem análises descritivas robustas (medianas, quartis, caudas — p. ex., P95) e séries temporais que revelam padrões de estabilidade ou degradação.

Do ponto de vista metodológico, conforme Lang (2017) e Greller e Drachsler (2012), esse tipo de estudo é documental/tecnológico: parte de registros produzidos rotineiramente, organizando-os por meio de um *pipeline ETL* (extração, transformação e carga), seguido de

modelagem indicadora.

Os princípios conceituais dos *frameworks* de *learning analytics* são traduzidos em variáveis e procedimentos operacionais neste estudo, articulando dimensões metodológicas, éticas e técnicas. A estrutura segue referenciais amplamente adotados na área — como o modelo “What data? Why? How?” de Greller e Drachsler (2012) e as recomendações de Lang (2017) para validade de construto, interna e externa. Esses referenciais permitem mapear de forma sistemática quais dados são relevantes, com qual finalidade analítica e por meio de quais métodos, assegurando coerência entre a fundamentação teórica e as métricas aplicadas.

Além disso, autores como Siemens e Long (2011), Ferguson (2012) e Sclater (2017) destacam que qualquer uso de dados educacionais exige atenção às restrições normativas — incluindo anonimização, governança e conformidade ética, conforme a LGPD (Lei nº 13.709/2018) e a Resolução CNS nº 510/2016.

Assim, a Tabela 10 sintetiza o processo de alinhamento entre as dimensões conceituais de *learning analytics* e as escolhas metodológicas que fundamentam a análise apresentada no Capítulo 7.

Tabela 10 – Mapeamento conceitual (framework LA): variáveis e procedimentos do estudo.

Dimensão (LA)	Operacionalização no estudo	Referência
What data?	Logs de sessão/mensagem; latência; http_5xx; uptime	Greller e Drachsler (2012), Lang (2017)
Why? (propósito)	Indicadores de engajamento técnico, eficiência e confiabilidade	Siemens e Long (2011), Ferguson (2012)
How? (método)	ETL; estatística descritiva (Q1/Mediana/Q3; P95); séries; proxies; ISO 25010	Lang (2017)
Constraints	Anonimização; minimização; governança e segurança	Sclater (2017), Lei Geral de Proteção de Dados – Brasil (2018)

Fonte: Elaborado com base em Greller e Drachsler (2012), Lang (2017).

2.2.5 Governança de dados e conformidade legal (LGPD)

A opção metodológica pela análise de dados secundários estritamente técnicos (logs de sistema) atende aos princípios de governança digital e proteção de dados, dispensando a intervenção direta com sujeitos. A abordagem fundamenta-se em três eixos de conformidade:

Lei geral de proteção de dados (LGPD - Lei 13.709/2018): o estudo adota o princípio da minimização (coleta apenas dos dados técnicos indispensáveis à análise de performance) e da anonimização irreversível na fonte. Não são armazenados nomes, CPFs ou dados sensíveis, focando-se exclusivamente nos identificadores de sessão (*Session IDs*) para fins de métrica de retenção.

Transparência e Finalidade: o tratamento dos dados possui finalidade legítima de auditoria de sistema e melhoria da qualidade técnica do serviço educacional, enquadrando-se nas bases legais de legítimo interesse e cumprimento de dever legal da administração pública.

Segurança da Informação (*Data Hardening*): a mitigação de riscos materializa-se em protocolos rígidos: (a) *hashing* de identificadores; (b) redução de granularidade temporal nos relatórios públicos; e (c) restrição de acesso aos dados brutos (*raw data*) apenas ao pesquisador responsável.

2.2.6 Avaliação de qualidade técnica com base na ISO/IEC 25010

A norma ISO/IEC 25010 (*System and Software Quality Models*) fornece o modelo padrão para avaliar a qualidade de produtos de *software*. Neste estudo, selecionou-se um subconjunto de características passíveis de mensuração automatizada via telemetria, sem necessidade de intervenção humana direta (ISO/IEC, 2011). São elas:

- Eficiência de desempenho (*Performance Efficiency*): mensurada através da latência de resposta (*Response Time*). Utilizam-se a Mediana e o Percentil 95 (P95) de *latency_ms* para verificar se o sistema mantém a responsividade mesmo sob carga (cauda curta indica consistência).
- Confiabilidade (*Reliability*): avaliada por meio de séries de disponibilidade (*up-time*) e da contagem de erros de servidor (códigos HTTP 5xx) por janela de tempo. Picos de erro sinalizam instabilidade ou falhas de regressão no código.

Os atributos de qualidade são operacionalizados em indicadores nos Capítulos 3 e 6, enquanto o Capítulo 7 discute os achados à luz do referencial técnico-pedagógico, com aprofundamento da dimensão pedagógica na Seção 7.6; as limitações de inferência e as ameaças à validade são sistematizadas nas limitações e perspectivas futuras.

Tabela 11 – Atributos de qualidade (ISO/IEC 25010) e indicadores baseados em logs.

Atributo	Indicador proposto	Observações
Eficiência de desempenho	Mediana / P95 de latency_ms	Sensível a tail latency
Confiabilidade	uptime por janela; http_5xx_count	Analisar picos; inspeção de causa-raiz

Fonte: Adaptado de ISO/IEC 25010 (2011), Lang (2017).

2.2.7 Agentes conversacionais na educação: lacunas de validação técnica

Revisões recentes sobre *chatbots* educacionais descrevem amplos benefícios teóricos, como suporte à adaptabilidade e *feedback* imediato. Contudo, a literatura evidencia uma lacuna crítica: a escassez de estudos que validem a estabilidade técnica dessas ferramentas em cenários de infraestrutura restrita (Sandoval-Reyes et al., 2023). A maioria das pesquisas foca na eficácia pedagógica percebida, negligenciando se a ferramenta opera com latência aceitável e alta disponibilidade — pré-requisitos para qualquer adoção em larga escala.

Diante disso, as contribuições esperadas deste trabalho residem em: (i) evidenciar a estabilidade e eficiência do EducaChat em ambiente real através de auditoria de logs; (ii) propor um conjunto de indicadores de qualidade reprodutíveis (baseados na ISO 25010) que independem de instrumentos subjetivos; e (iii) documentar um protocolo de governança de dados compatível com a LGPD, demonstrando como analisar interatividade educacional sem violar a privacidade dos estudantes.

2.2.8 Síntese integradora e implicações para a metodologia

O referencial teórico aqui consolidado fundamenta a estratégia metodológica adotada: um estudo técnico-analítico fundamentado em *learning analytics*, operando exclusivamente com dados secundários anonimizados.

A revisão histórica (Seção 2.1 a 2.3) justificou a necessidade de superar a análise de "caixa preta" dos modelos de IA.

A discussão sobre infraestrutura (Seção 2.4) definiu as variáveis ambientais (rede e latência) que devem ser controladas.

Por fim, o modelo de qualidade de software (Seção 2.5) forneceu as métricas objetivas para a avaliação.

Esse alinhamento teórico sustenta o desenho do Capítulo 3 – Metodologia, que detalhará os procedimentos de extração, transformação e análise estatística dos registros do sistema.

Ao mesmo tempo em que reconhece limites de inferência pedagógica direta, o desenho proposto oferece evidências técnicas necessárias e reproduzíveis para balizar decisões e orientar fases subsequentes de investigação aplicada com participantes.

2.2.9 Fundamentos pedagógicos da arquitetura do sistema

Embora a validação deste estudo seja estritamente técnica (via logs), o *design* do EducaChat não é aleatório; ele operacionaliza conceitos clássicos da psicologia da educação em regras de código. A arquitetura do sistema traduz a teoria sociocultural de Vygotsky (1934/2007) em algoritmos de andamento (*scaffolding*): o bot é programado para não dar a resposta direta, mas para liberar pistas condicionais, tentando simular tecnicamente a zona de desenvolvimento proximal (ZDP).

Da mesma forma, os princípios de Mayer e Moreno (2003) sobre aprendizagem multimídia orientaram o design de interface (UI), limitando o volume de texto por balão de fala para evitar sobrecarga cognitiva — um requisito verificável através da métrica de tokens por mensagem. Já o conceito de orquestração de Dillenbourg (2013) justifica a necessidade de alta disponibilidade (ISO 25010): para que o professor possa orquestrar a turma, o sistema técnico não pode falhar (cair) durante a aula.

2.3 Trabalhos relacionados e lacuna de pesquisa

Pesquisas nacionais e internacionais investigam agentes conversacionais, geralmente focando na "eficácia percebida" ou satisfação do usuário. Contudo, revisões sistemáticas indicam uma lacuna metodológica: poucos estudos validam a robustez da infraestrutura antes de avaliar a pedagogia. A Tabela 12 (a seguir) sintetiza essa lacuna, demonstrando que, enquanto a maioria dos trabalhos olha para o "conteúdo", esta dissertação inova ao auditar o "continente" (O sistema e a rede).

Tabela 12 – Educação.

Estudo
Mendes e Sirqueira (2022)
Guerreiro e Barros (2019)
Paschoal, Chicon e Falkembach (2017)
Clemente (2016), Santana (2017)

Fonte: Elaborado pelo autor (2025).

A revisão da literatura, sistematizada na Tabela 12, revela um padrão consistente nas pesquisas sobre agentes conversacionais: a predominância de avaliações baseadas na percepção subjetiva dos usuários (questionários, entrevistas) em detrimento da auditoria objetiva dos sistemas.

Embora os estudos listados apontem ganhos pedagógicos potenciais, observa-se escassez de trabalhos que validem a viabilidade técnica (ISO 25010) dessas soluções antes de sua aplicação em larga escala. A maioria das investigações assume condições ideais de infraestrutura escolar, negligenciando gargalos de latência e conectividade, fatores críticos na realidade da escola pública brasileira.

2.3.1 Estudos complementares na literatura

A análise dos trabalhos relacionados confirma a oportunidade de pesquisa: existe uma carência de protocolos que utilizem logs de sistema como evidência primária de qualidade. Enquanto a literatura existente foca no "o que o aluno sentiu" (percepção), esta dissertação propõe focar no "como o sistema se comportou" (desempenho e adaptabilidade).

Com base nessa fundamentação teórica — que une o histórico da IA (2.1), os desafios de infraestrutura (2.4) e os modelos de qualidade de software (2.5) —, o próximo capítulo detalha o percurso metodológico para a extração e análise desses dados.

De modo geral, a síntese dos trabalhos relacionados aponta que, embora existam evidências sobre o potencial de engajamento dos agentes conversacionais, permanecem lacunas metodológicas críticas.

A principal delas é a predominância de estudos baseados em percepção do usuário (subjetiva) em detrimento da validação técnica da infraestrutura (objetiva).

Poucos trabalhos auditam se a ferramenta suporta a carga de uso real antes de questionar se os alunos gostaram dela. Essa constatação reforça a pertinência da presente investigação, que busca superar tais limitações ao auditar o EducaChat com foco estrito em métricas de qualidade de *software* e estabilidade de sistema.

2.3.2 Avaliação de agentes conversacionais incorporados

O estudo de Brandão, Reis e Rocha (2013) propõe uma plataforma que permite avaliar o desempenho de um agente conversacional incorporado e a influência de diferentes elementos multimídia. Inspirado no “Jogo de imitação” de Alan Turing, o sistema possibilitou analisar a interação entre usuários e agentes em múltiplas configurações de *interface*.

Os resultados indicaram que fatores como a aparência do *avatar*, o uso de emoções expressas e a incorporação de recursos multimídia estão associados a uma experiência de uso mais positiva, refletida em maiores níveis de satisfação e engajamento percebido pelos usuários.

Embora não tenha sido possível estabelecer uma relação direta entre esses elementos e a melhoria da aprendizagem em termos de desempenho, a análise evidenciou o potencial das interfaces multimodais para enriquecer a interação educacional e apoiar processos de mediação pedagógica, especialmente no que se refere à clareza das explicações, à fluidez da interação e ao envolvimento dos estudantes.

2.3.3 Métricas de qualidade de software

A avaliação da qualidade de um produto de software constitui etapa essencial no ciclo de desenvolvimento, uma vez que permite identificar atributos relacionados à eficácia, eficiência, confiabilidade, manutenibilidade e usabilidade. Para tanto, destacam-se os modelos normativos internacionais estabelecidos pela ISO/IEC 9126-1 e pela ISO/IEC 25010, amplamente reconhecidos no campo da engenharia de *software*.

Esses modelos apresentam um conjunto estruturado de características e subcaracterísticas que visam orientar tanto o processo de desenvolvimento quanto a avaliação de produtos em contextos reais de uso. Enquanto a ISO/IEC 9126-1 introduziu um modelo pioneiro de qualidade do produto, a ISO/IEC 25010 ampliou a abordagem ao incluir dimensões como a qualidade em uso, reforçando o papel da satisfação do usuário e da cobertura de contexto.

2.3.4 Comparação entre ISO/IEC 9126-1 e ISO/IEC 25010

A norma ISO/IEC 9126-1, publicada em 2001, foi pioneira ao propor um modelo estruturado de qualidade de *software*, organizado em seis características principais.

Posteriormente, a ISO/IEC 25010, lançada em 2011, ampliou e refinou o modelo, consolidando-o em oito características, além de enfatizar a dimensão da qualidade em uso.

Tabela 13 - Sintetiza as principais diferenças entre os dois referenciais normativos.

Dimensão	ISO/IEC 9126-1 (2001)	ISO/IEC 25010 (2011)
Características principais	Funcionalidade, Confiabilidade, Usabilidade, Eficiência, Manutenibilidade, Portabilidade	Funcionalidade, Desempenho/Eficiência, Compatibilidade, Usabilidade, Confiabilidade, Segurança, Manutenibilidade, Portabilidade
Foco do modelo	Qualidade do produto de software	Qualidade do sistema e qualidade em uso
Qualidade em uso	Abordada de forma indireta	Explicitamente definida com cinco dimensões: eficácia, eficiência, satisfação, ausência de riscos e cobertura de contexto
Ênfase pedagógica (aplicação)	Uso incipiente em contextos educacionais digitais	Maior aderência ao contexto educacional por contemplar satisfação do usuário e adaptação ao ambiente
Evolução conceitual	Modelo pioneiro e referencial inicial	Consolidação de avanços e incorporação de novas demandas (segurança, compatibilidade, cobertura de contexto)

Fonte: Adaptado de ISO/IEC 9126-1 (2001) e ISO/IEC 25010 (2011).

A comparação evidencia que a ISO/IEC 25010 não substitui apenas a versão anterior, mas representa um avanço significativo ao incorporar dimensões cada vez mais relevantes no uso educacional. Por exemplo, a satisfação do usuário e a cobertura de contexto são fundamentais quando o *software* é aplicado em cenários escolares heterogêneos, nos quais a diversidade de infraestrutura tecnológica e de perfis de estudantes impacta diretamente a experiência.

“A avaliação da qualidade em uso representa uma mudança de paradigma: não se trata apenas de verificar atributos técnicos do *software*, mas de compreender como ele contribui para que os usuários atinjam seus objetivos em um contexto real.” (ISO/IEC 25010, 2011).

No caso do EducaChat, essa evolução é particularmente relevante, pois o agente conversacional não se limita a responder tecnicamente, mas deve sustentar a mediação pedagógica, adaptar-se ao ritmo dos estudantes e garantir acessibilidade em diferentes contextos escolares. Dessa forma, a adoção da ISO/IEC 25010 fortalece o rigor metodológico e amplia a

validade da avaliação proposta nesta dissertação.

2.3.5 Qualidade em uso

A qualidade em uso, definida pela ISO/IEC 25010, refere-se à perspectiva do usuário final e descreve em que medida um produto de *software* possibilita que indivíduos ou grupos atinjam metas específicas com eficácia, eficiência, segurança e satisfação em um contexto de uso determinado.

Os atributos de qualidade em uso são organizados em cinco características principais: Eficácia; Eficiência; Satisfação; Livre de riscos e Cobertura de contexto.

A Tabela 14 apresenta um resumo dessas características, subcaracterísticas e definições.

Tabela 14 - Qualidade em uso segundo a ISO/IEC 25010.

Características	Subcaracterísticas	Definição
Eficácia	–	Precisão e integridade com as quais os usuários alcançam objetivos específicos.
Eficiência	–	Recursos gastos em relação à precisão e integridade com que os usuários atingem as metas.
Satisfação	Utilidade	Grau em que o produto ajuda o usuário a atingir objetivos pragmáticos.
	Confiança	Grau em que o usuário confia que o sistema funciona conforme esperado.
	Prazer	Grau em que o usuário obtém prazer/valorização subjetiva ao usar o sistema.
	Conforto	Grau de satisfação com o conforto físico/-cognitivo durante o uso.
Livre de riscos	Mitigação de risco econômico	Grau em que o sistema mitiga riscos financeiros ou de reputação.
	Mitigação de riscos de saúde e segurança	Grau em que o sistema mitiga riscos à saúde e segurança dos usuários e de terceiros.
	Mitigação de risco ambiental	Grau em que o sistema mitiga riscos ambientais.
Cobertura de contexto	Compleitude	Grau em que o sistema pode ser usado em todos os contextos de uso especificados.
	Flexibilidade	Grau em que o sistema pode ser usado fora dos contextos especificados.

Fonte: ISO/IEC 25010 (2011), adaptado.

2.3.6 Escolha das métricas de avaliação

Não há uma regra única para a seleção das métricas de avaliação, uma vez que a escolha das categorias e subcategorias propostas pela ISO/IEC 25010 depende diretamente dos objetivos do estudo. Dessa forma, pesquisadores podem optar por métricas de eficácia, eficiência, usabilidade ou segurança, conforme a ênfase pretendida.

No presente trabalho, que visa avaliar a eficácia do EducaChat, justifica-se a adoção de métricas que contemplam tanto a eficiência quanto a confiança, dimensões que permitem analisar o desempenho técnico e a aceitação do agente conversacional pelos usuários.

Além disso, segundo a NBR ISO/IEC 25010, diferentes atores podem avaliar o mesmo sistema sob perspectivas distintas:

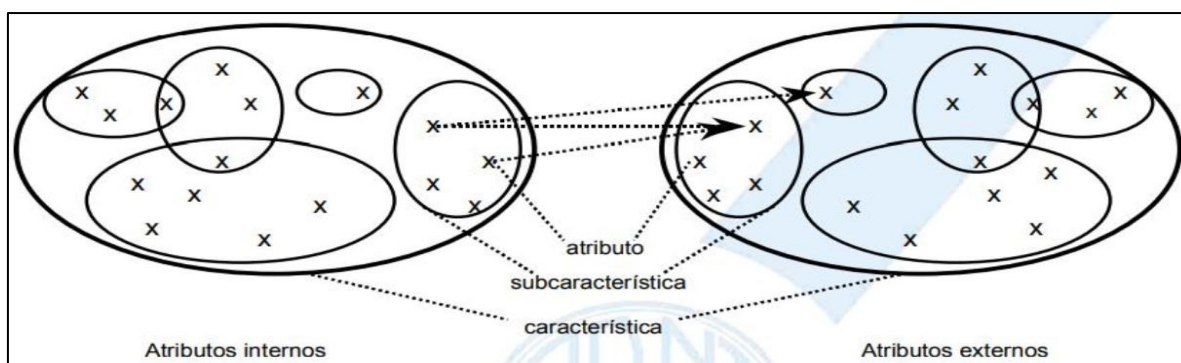
- O usuário final pode optar por métricas de usabilidade e satisfação.
- A equipe de suporte pode focar em manutenibilidade.
- O analista de sistemas pode priorizar portabilidade.
- O desenvolvedor pode avaliar com base em critérios internos de qualidade do código.

Esse modelo flexível reforça que as métricas de avaliação devem ser sempre contextualizadas de acordo com a finalidade do estudo.

2.3.7 Métricas internas e externas

As normas ISO/IEC 9126-1 e ISO/IEC 25010 destacam a existência de três níveis de métricas: internas, externas e de qualidade em uso. As métricas internas referem-se ao código-fonte e à arquitetura do sistema, as métricas externas observam o comportamento do *software* em execução, e as métricas de qualidade em uso refletem a percepção dos usuários em contextos reais de aplicação como é ilustrada na Figura 4.

Figura 4 – Características, subcaracterísticas e atributos de qualidade de software.



Fonte: NBR ISO/IEC 9126-1.

2.3.8 Relevância para este estudo

A partir do exposto, justifica-se a utilização das normas ISO/IEC 9126-1 e ISO/IEC 25010 como base para a avaliação da eficácia do EducaChat.

Esses referenciais oferecem um arcabouço robusto que possibilita analisar a eficiência, a confiança e a satisfação do usuário, dimensões críticas para a adoção de agentes conversacionais educacionais.

Assim, a inclusão dessas métricas no presente estudo não apenas fortalece o rigor metodológico, como também amplia a validade dos resultados, garantindo que a avaliação do EducaChat esteja alinhada a padrões internacionais de qualidade em engenharia de *software*.

2.3.9 Atributos de qualidade para chatbots: proposta de Radziwill e Benton (2017)

O estudo proposto por Radziwill e Benton (2017) apresenta abordagens de avaliação da qualidade e desenvolve um método baseado em atributos. O objetivo dos autores foi examinar a literatura e fornecer uma revisão abrangente dos atributos de qualidade para *chatbots* e agentes conversacionais, identificando práticas apropriadas de garantia da qualidade. Eles extraíram atributos de qualidade de 32 artigos, organizados em categorias que contemplam eficiência, eficácia e satisfação.

As propostas desses autores são apresentadas nos quadros a seguir, que relacionam as categorias, subcategorias e os respectivos atributos de qualidade.

Observa-se que a proposta de Radziwill e Benton (2017) amplia a análise de qualidade dos agentes conversacionais ao incorporar atributos ligados não apenas ao desempenho técnico, mas também a aspectos emocionais, sociais e éticos.

No contexto educacional, tais dimensões são particularmente relevantes, pois o engajamento dos estudantes depende tanto da robustez funcional quanto da capacidade do agente em promover interações significativas e respeitosas.

2.4 Bases conceituais

2.4.1 Inteligência artificial: perspectivas históricas e educacionais

A Inteligência artificial (IA) constitui um campo interdisciplinar que, desde sua formalização em 1956, tem evoluído em diferentes vertentes. Russell e Norvig (2013) descrevem quatro grandes abordagens: Sistemas que pensam como humanos, sistemas que agem como humanos, sistemas que pensam racionalmente e sistemas que agem racionalmente.

Cada uma dessas perspectivas impacta de forma distinta o desenvolvimento de tecnologias educacionais.

Figura 5 - Abordagens da IA segundo Russell e Norvig (2013).

Abordagem	Descrição	Exemplos educacionais
Pensar como humanos	Modelagem de processos cognitivos e psicológicos humanos	Sistemas de tutoria cognitiva; simulação de estratégias de raciocínio do estudante
Agir como humanos	Ênfase na interação natural e na simulação de comportamento humano	Chatbots educacionais; assistentes virtuais com linguagem natural
Pensar racionalmente	Fundamenta-se na lógica formal e na resolução de problemas	Motores de inferência em sistemas especialistas para apoio ao ensino
Agir racionalmente	Foco em agentes que buscam maximizar resultados em ambientes dinâmicos	Agentes inteligentes adaptativos para recomendação e apoio à aprendizagem

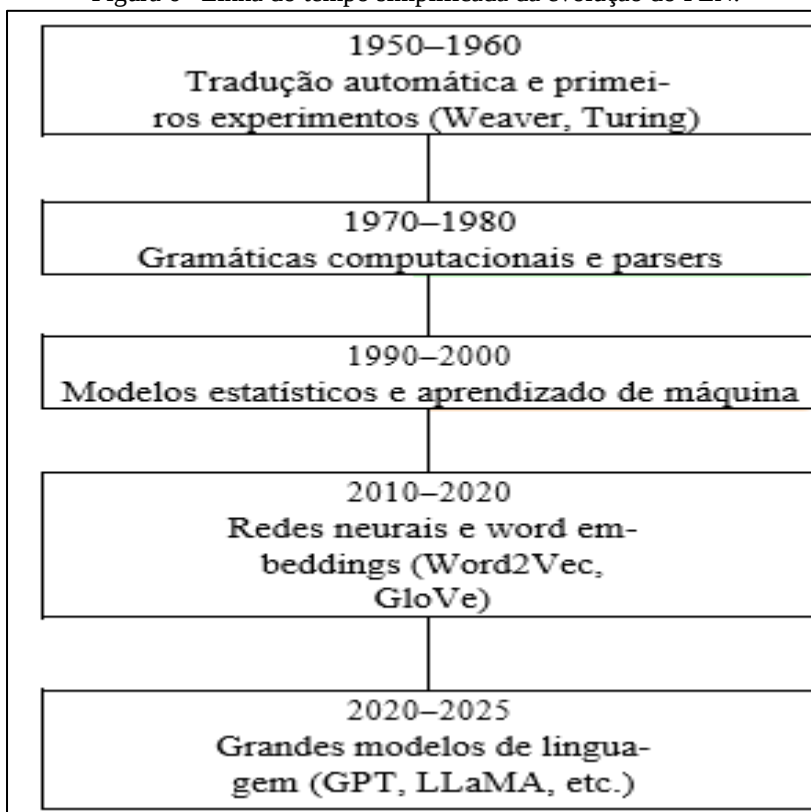
Fonte: Adaptado de Russell e Norvig (2013).

Ao trazer essas perspectivas para o contexto educacional, observa-se que o EducaChat aproxima-se principalmente da abordagem de agentes que agem como humanos, uma vez que interagem em linguagem natural, mas também incorpora elementos de agir racionalmente, ao tomar decisões baseadas em contextos pedagógicos.

2.4.2 Processamento de linguagem natural (PLN)

O Processamento de linguagem natural (PLN) constitui a espinha dorsal de agentes conversacionais. Historicamente, evoluiu da tradução automática nos anos 1950 para técnicas de aprendizado profundo em larga escala no século XXI.

Figura 6– Linha do tempo simplificada da evolução do PLN.



Fonte: Elaborado pelo autor (2025).

O EducaChat se beneficia dessa evolução, empregando modelos de linguagem baseados em redes neurais profundas, capazes de compreender contexto, responder a dúvidas e adaptar-se ao nível de conhecimento dos alunos.

2.4.4 Alinhamento à BNCC (Matemática e Ciências da Natureza)

O uso do EducaChat apresenta alinhamento conceitual e pedagógico com as competências gerais da base nacional comum curricular (BNCC), especialmente no que se refere à cultura digital, ao pensamento científico, à comunicação e à responsabilidade e cidadania. No âmbito dos componentes curriculares de matemática e ciências da natureza (biologia), o agente conversacional pode atuar como recurso de apoio à mediação pedagógica, ao oferecer orientações, pistas e devolutivas que dialogam com habilidades previstas na BNCC.

Tabela 15– BNCC: Competências/habilidades e suporte do EducaChat.

Componente	Habilidades exemplares	Suporte do agente
Matemática	Resolver problemas com funções; generalizar e justificar procedimentos	Pistas graduais; checagem de pré-requisitos; exemplos análogos
Ciências da Natureza (Biologia)	Compreender relações ecológicas; analisar evidências; formular hipóteses	Perguntas orientadoras; analogias; miniquizzes formativos

Fonte: Elaborado pelo autor (2025).

A análise desses cenários revela que os *chatbots* educacionais podem desempenhar um papel estratégico na promoção da personalização da aprendizagem, no fortalecimento da colaboração entre estudantes e na otimização do tempo do professor. Além disso, ao serem utilizados como instrumentos de avaliação formativa, contribuem para um acompanhamento mais contínuo e eficaz do processo de aprendizagem.

Além dos cenários pedagógicos apresentados, é importante reconhecer que a inserção de *chatbots* educacionais também suscita dilemas de ordem ética e tecnológica. Esses desafios vão desde a proteção de dados sensíveis até questões relacionadas à escalabilidade das plataformas e ao risco de dependência excessiva por parte dos estudantes.

2.4.9 Riscos, salvaguardas e papel do professor

A adoção responsável requer salvaguardas pedagógicas e técnicas. A Tabela 16 resume pontos críticos e mitigadores práticos.

Tabela 16 – Riscos pedagógicos e salvaguardas na adoção de *chatbots*.

Risco	Efeito educacional	Salvaguardas recomendadas
Sobrecarga cognitiva	Distração; queda de desempenho	Segmentação; sinalização; limites de comprimento de resposta (MAYER; MORENO, 2003)
Dependência excessiva	Menor autonomia e autorregulação	Prompts metacognitivos; exigir “mostre seu raciocínio”
Alucinações/erros	Conceitos equivocados	Verificação cruzada; rubricas de checagem; fontes sugeridas
Viés e linguagem inapropriada	Inequidades; clima de sala prejudicado	Diretrizes éticas; filtros; revisão docente (HOLMES; BIALIK; FADEL, 2019; LUND; WANG, 2023)

Fonte: Elaborado pelo autor (2025).

2.5.6 Modelos e arquiteturas técnicas de chatbots educacionais

O desenvolvimento de *chatbots* educacionais requer a definição de arquiteturas que conciliem desempenho técnico, escalabilidade e pertinência pedagógica. Tais modelos combinam técnicas de processamento de linguagem natural (PLN), inteligência artificial e integração com sistemas de gestão da aprendizagem (LMS/AVA).

2.5.7 Arquitetura em camadas

Uma abordagem recorrente é a arquitetura em camadas, que separa:

1. Camada de interface: interação com o usuário via texto ou voz.
2. Camada de processamento de linguagem: segmentação, análise semântica e intenção.
3. Camada de raciocínio: escolha da resposta com base em regras, aprendizado de máquina ou *LLMs*.
4. Camada de dados: bases de conhecimento, conteúdos curriculares, logs de interação.

5. Camada de integração: comunicação com sistemas externos (*LMS*, bancos de dados).

Figura 7– Arquitetura em camadas de um chatbot educacional.



Fonte: Elaborado pelo autor (2025).

2.5.8 Modelos baseados em regras, estatísticos e neurais

Os *chatbots* podem adotar diferentes paradigmas:

- Baseados em regras: utilizam *scripts* e fluxogramas; simples, mas pouco escaláveis.
- Estatísticos: empregam algoritmos de aprendizado supervisionado; maior flexibilidade, mas exigem conjuntos de dados rotulados.
- Neurais/*LLMs*: utilizam arquiteturas de redes neurais profundas (*transformers*); oferecem conversação mais natural, porém com maior demanda computacional e riscos de “alucinação”.

2.5.9 Integração com ambientes virtuais de aprendizagem

A efetividade dos *chatbots* tende a ampliar-se quando integrados a ambientes virtuais de aprendizagem (AVA), como *Moodle*, *Google Classroom* ou *Canvas*. Essa integração pode oferecer:

- Personalização de trilhas de aprendizagem.
- Registro automático de interações e progresso.

- Suporte ao professor no acompanhamento em tempo real.

2.6 Quadro comparativo de modelos

Na literatura, identificam-se diferentes abordagens de *chatbots* aplicados à educação, cada uma com vantagens e limitações específicas. A **Tabela 17** sintetiza essas características.

Tabela 17 – Comparação de modelos de *chatbots* aplicados à educação.

Modelo	Vantagens	Limitações
Baseado em regras	Simplicidade; baixo custo de implementação	Conversas pouco naturais; baixa escalabilidade
Estatístico	Adaptação a diferentes domínios; maior flexibilidade	Necessidade de grande volume de dados rotulados
Neurais / LLMs	Linguagem natural avançada; respostas contextuais	Alto custo computacional; risco de vieses e "alucinações"

Fonte: Elaborado pelo autor (2025).

2.6.1 Impactos pedagógicos observados em estudos recentes

A literatura recente evidencia que *chatbots* educacionais podem atuar como mediadores digitais, favorecendo engajamento, personalização e autorregulação da aprendizagem. Diversos estudos têm explorado seus impactos em diferentes níveis de ensino, disciplinas e metodologias de ensino-aprendizagem.

2.6.2 Potencial de sustentação da interatividade

Pesquisas apontam que a interação contínua com *chatbots* aumenta a motivação dos estudantes, especialmente quando há uso de elementos lúdicos, *feedback* imediato e comunicação em linguagem natural. Chiu, Chi e Davis (2021) evidenciam que a personalização das respostas contribui para maior persistência nas tarefas.

2.6.3 Suporte à resolução de problemas e scaffolding

Embora alguns trabalhos ainda relatem efeitos limitados na aprendizagem medida por testes objetivos, há indícios de que os *chatbots* favorecem o desenvolvimento de habilidades cognitivas superiores, como pensamento crítico e resolução de problemas. Cho e Lim (2017) destacam o uso de ambientes 3D com *chatbots* para fortalecer colaboração e aprendizagem ativa.

2.6.4 Papel no suporte ao professor

Além de atender diretamente os estudantes, os *chatbots* podem apoiar o trabalho docente ao automatizar dúvidas frequentes, fornecer relatórios sobre dificuldades recorrentes e liberar tempo para atividades de maior valor pedagógico. Estudos de Holmes, Bialik e Fadel (2019) indicam que a articulação entre professor humano e agente conversacional está associada a experiências mais efetivas de ensino híbrido.

2.6.5 Síntese dos impactos

A **Tabela 18** sintetiza os principais impactos pedagógicos relatados em estudos recentes.

Tabela 18 – Impactos pedagógicos de *chatbots* em diferentes contextos educacionais.

Estudo	Contexto	Impactos observados	Limitações
Chiu, Chi e Davis (2021)	Ensino Superior, IA aplicada	Personalização; engajamento maior; feedback imediato	Amostra restrita a um curso de tecnologia
Cho e Lim (2017)	Ambientes virtuais 3D	Colaboração e interação ativa entre estudantes	Exigência de infraestrutura tecnológica avançada
Holmes, Bialik e Fadel (2019)	Revisão sobre IA em educação	Apoio ao professor; melhores resultados em ensino híbrido	Resultados heterogêneos; lacunas em avaliação longitudinal
Paschoal, Chicon e Falkembach (2017)	Educação em Computação	Ubiquidade e suporte personalizado ao aluno	Escopo disciplinar limitado

Fonte: Elaborado pelo autor (2025).

2.6.6 Considerações gerais

Em síntese, os *chatbots* não substituem professores, mas ampliam suas possibilidades

de ação ao fornecer mediação contínua, personalização e acompanhamento em tempo real. Os benefícios pedagógicos dependem fortemente da qualidade do design instrucional, do alinhamento com objetivos curriculares e da capacidade técnica de implementação.

2.7.3 Limitações e desafios atuais dos *chatbots* educacionais

Apesar dos avanços significativos observados nos últimos anos, os *chatbots* educacionais ainda enfrentam barreiras que limitam sua adoção plena e seu impacto em contextos de ensino-aprendizagem. Nesta seção, destacam-se as principais limitações e desafios técnicos, pedagógicos, éticos/sociais e institucionais que condicionam a adoção de agentes conversacionais em contextos de ensino-aprendizagem.

2.7.4 Desafios técnicos

Entre as limitações técnicas, estão a dependência de infraestrutura robusta, o custo computacional elevado para treinar e executar modelos de grande porte e as dificuldades de integração com ambientes virtuais de aprendizagem (AVA) já existentes. Além disso, a instabilidade em conexões de internet em escolas públicas brasileiras agrava a desigualdade no acesso.

2.7.5 Desafios pedagógicos

Do ponto de vista pedagógico, um dos grandes desafios é assegurar que os agentes conversacionais apoiem, de fato, o processo de ensino, e não apenas funcionem como respostas automáticas.

A literatura aponta que, em muitos casos, os *chatbots* oferecem suporte superficial, sem promover reflexão crítica ou construção ativa de conhecimento.

Outro ponto sensível é a necessidade de alinhar o uso dessas ferramentas às diretrizes curriculares, como a BNCC, e às estratégias pedagógicas de cada disciplina.

2.7.7 Desafios institucionais

A adoção efetiva de *chatbots* exige mudanças institucionais, como formação conti-

nuada de professores, atualização de políticas pedagógicas e investimento em infraestrutura tecnológica. A resistência de alguns docentes e gestores à incorporação de novas tecnologias também aparece como barreira recorrente em estudos de campo.

2.7.8 Síntese dos desafios

A **Tabela 19** sintetiza os principais desafios atuais que condicionam a adoção dos agentes conversacionais educacionais em larga escala.

Tabela 19– Principais limitações e desafios dos *chatbots* educacionais.

Dimensão	Desafios	Implicações
Técnica	Infraestrutura insuficiente; custo computacional elevado; integração limitada a AVAs	Barreiras de acesso e escalabilidade reduzida
Pedagógica	Superficialidade das interações; desalinhamento curricular	Impacto limitado na aprendizagem significativa
Ética e social	Privacidade de dados; vieses; risco de desinformação	Necessidade de regulação e políticas claras
Institucional	Formação docente; resistência cultural; recursos financeiros limitados	Adoção desigual e lenta em redes educacionais

Fonte: Elaborado pelo autor (2025).

2.7.9 Síntese final

Portanto, os desafios apresentados reforçam que a eficácia dos *chatbots* educacionais não depende apenas de avanços tecnológicos, mas também de políticas públicas consistentes, estratégias pedagógicas adequadas e formação de professores. Reconhecer essas limitações é condição necessária para que as perspectivas futuras se concretizem de forma ética, inclusiva e sustentável.

2.8 Lacunas na literatura

Embora o campo dos agentes conversacionais educacionais tenha avançado de forma

significativa, a fundamentação teórica evidencia lacunas que precisam ser consideradas para orientar futuras pesquisas e práticas pedagógicas.

2.8.1 Contexto educacional brasileiro

Grande parte dos estudos concentra-se em experiências internacionais, sobretudo em países com infraestrutura tecnológica mais consolidada. Há carência de evidências empíricas sobre a aplicação de *chatbots* em escolas públicas brasileiras, onde persistem desafios de conectividade, desigualdade de acesso e escassez de recursos pedagógicos digitais (CETIC.BR, 2024). Esse vácuo limita a compreensão sobre como tais tecnologias se comportam em contextos de vulnerabilidade social e educacional.

2.8.3 Integração curricular

Os estudos analisados frequentemente apresentam aplicações pontuais ou de caráter experimental, sem um alinhamento sistemático aos currículos oficiais, como a BNCC. Essa lacuna dificulta a consolidação dos *chatbots* como ferramentas pedagógicas amplamente adotadas e reconhecidas por gestores e professores.

2.8.4 Métricas de qualidade em software educacional

Apesar de avanços trazidos pelas normas ISO/IEC 9126-1 e ISO/IEC 25010, observa-se pouca exploração de métricas comparáveis em contextos educacionais.

A maioria dos trabalhos limita-se a medir usabilidade e satisfação do usuário, negligenciando dimensões críticas como confiabilidade, segurança e cobertura de contexto. Tal limitação restringe a possibilidade de estabelecer padrões de qualidade consistentes para sistemas de apoio à aprendizagem.

2.8.5 Dimensões éticas e socioculturais

Outro ponto pouco explorado refere-se às implicações éticas e socioculturais. Questões como viés algorítmico, privacidade de dados e responsabilidade no uso de IA em ambientes escolares ainda carecem de estudos sistemáticos que orientem tanto políticas públicas quanto práticas pedagógicas (MIAO; HOLMES, 2023; GENERATIVE AI..., 2023). Essa ausência

torna urgente a formulação de diretrizes claras que articulem inovação tecnológica e valores educacionais.

2.8.6 Síntese das lacunas

A **Tabela 20** sintetiza as principais lacunas identificadas, evidenciando oportunidades para investigações futuras.

Tabela 20– Lacunas identificadas na literatura sobre *chatbots* educacionais.

Dimensão	Lacunas identificadas	Implicações analíticas
Contexto brasileiro	Escassez de estudos situados em contextos educacionais públicos e diversificados	Limitações na compreensão da adoção de chatbots em realidades educacionais heterogêneas
Triangulação analítica	Predominância de abordagens baseadas em instrumentos únicos ou análises isoladas	Redução da robustez metodológica e da validade das conclusões
Integração curricular	Uso pontual de agentes conversacionais, com fraca articulação a referenciais curriculares	Dificuldade de consolidação pedagógica e alinhamento institucional
Métricas de qualidade	Foco restrito em atributos isolados de usabilidade ou aceitação	Ausência de modelos integrados de avaliação de qualidade em uso
Aspectos éticos e socioculturais	Pouca atenção a temas como viés algorítmico, privacidade e governança	Riscos associados ao uso inadequado de IA em contextos educacionais

Fonte: Elaborado pelo autor (2025).

2.8.7 Síntese final

Assim, a literatura aponta que os agentes conversacionais educacionais possuem grande potencial, mas sua consolidação como ferramentas efetivas de aprendizagem requer superar lacunas metodológicas, curriculares, técnicas e éticas.

Este estudo, ao aplicar o EducaChat em uma escola pública brasileira, busca contribuir justamente para preencher parte dessas lacunas, oferecendo evidências empíricas contextualizadas e alinhadas a padrões internacionais de qualidade.

Tabela 21– Alinhamento entre hipóteses, métricas de qualidade e instrumentos analíticos.

Hipótese	Métrica ISO/IEC 25010	Indicador operacional	Instrumento analítico
H1	Adequação funcional	Cobertura dos objetivos instrucionais; consistência das orientações	Análise de registros técnicos de interação (logs)
H2	Eficiência de desempenho	Tempo médio de resposta; fluidez das interações	Logs de uso e métricas de desempenho do sistema
H3	Satisfação	Utilidade percebida; clareza das respostas	Instrumentos de avaliação da experiência do usuário
H4	Confiabilidade	Consistência e previsibilidade das respostas	Análise de estabilidade do sistema e registros de erro
H5	Cobertura de contexto	Manutenção da usabilidade em diferentes modalidades	Comparação de métricas técnicas por contexto de uso
H6	Engajamento	Frequência de interações; turnos de diálogo; continuidade das sessões	Relatórios de logs e estatísticas de uso

Fonte: Elaborado pelo autor (2025).

Tabela 22– Comparação entre dimensões de qualidade e evidências na literatura.

Dimensão (ISO/IEC 25010)	Autor(es) de referência	Abordagem na literatura	Principais contribuições teóricas
Adequação funcional (H1)	Bloom (1984); VanLehn (2011)	Estudos teóricos e empíricos sobre tutoria	Tutoria individual como referência conceitual; agentes conversacionais como aproximação tecnológica ao apoio personalizado
Eficiência (H2)	Dillenbourg (2013); Chiu et al. (2021)	Análises de sistemas educacionais digitais	Feedback imediato e adaptação reduzem carga cognitiva e aumentam fluidez das interações
Satisfação (H3)	Holmes et al. (2019); Lund & Wang (2023)	Estudos sobre aceitação tecnológica	Experiência positiva e utilidade percebida influenciam adoção sustentável
Confiabilidade (H4)	VanLehn (2011); Radziwill & Benton (2017)	Pesquisas sobre consistência e transparência	Clareza e previsibilidade fortalecem confiança em sistemas baseados em IA
Cobertura de contexto (H5)	Holmes (2019); Dillenbourg & Jermann (2013)	Estudos sobre flexibilidade contextual	Adaptabilidade a diferentes cenários amplia escalabilidade educacional
Engajamento (H6)	Mayer & Moreno (2003); Cho & Lim (2017)	Pesquisas em aprendizagem multimídia	Interatividade e multimodalidade favorecem envolvimento e continuidade

Fonte: Elaborado pelo autor (2025), com base na literatura revisada.

2.9.7 Síntese crítica da literatura

Considerando a fundamentação teórica exposta e a opção metodológica pela análise de dados secundários (logs), esta dissertação não se propõe a testar hipóteses pedagógicas de ganho

cognitivo (o que exigiria pré e pós-testes), mas sim a validar a qualidade técnica da infraestrutura que suporta a interação.

Assim, o estudo norteia-se pelas seguintes questões de pesquisa (QPs), alinhadas à norma ISO/IEC 25010:

QP1(Eficiência de Desempenho): a latência de resposta do EducaChat mantém-se dentro de limites aceitáveis para o fluxo conversacional ($< 2000\text{ms}$) mesmo em condições reais de rede escolar?

QP2(Confiabilidade): o sistema apresenta estabilidade de conexão (uptime) e baixa taxa de erros de API (`http_5xx`) durante os picos de uso em sala de aula?

QP3(Operabilidade/Engajamento Técnico): os padrões de interação registrados nos logs (duração da sessão, número de mensagens) indicam que a ferramenta foi capaz de sustentar o diálogo técnico sem interrupções abruptas?

2.9.8 Fechamento do capítulo

Este capítulo estabeleceu as bases teóricas que sustentam a investigação. Partindo da evolução histórica da IA (de regras a *LLMs*), discutiu-se a centralidade da infraestrutura como gargalo em escolas públicas e definiu-se a norma ISO/IEC 25010 como régua de qualidade.

Ficou demonstrado que a "eficácia" de um agente conversacional não pode ser dissociada da sua "estabilidade". Sem uma validação técnica robusta, qualquer análise pedagógica torna-se frágil.

O Capítulo 3 (Metodologia) detalhará os procedimentos técnicos para responder a essas questões, descrevendo a arquitetura do sistema, o *pipeline* de extração de dados e o tratamento estatístico aplicado aos registros de log.

3. METODOLOGIA

Este capítulo detalha o percurso metodológico adotado para analisar o EducaChat com base exclusivamente em dados secundários anonimizados, incluindo logs de interação, métricas de desempenho e documentação técnica. O estudo não envolve coleta direta de dados com seres humanos, preservando a conformidade ética e legal com a Resolução CNS nº 510/2016 e a Lei Geral de Proteção de Dados (LGPD).

A Metodologia está organizada em: (i) natureza e desenho do estudo; (ii) governança e fontes de dados; (iii) modelo de dados e dicionário de variáveis; (iv) procedimentos de coleta e tratamento (pipeline ETL); (v) análise quantitativa; (vi) avaliação de qualidade de *software* (ISO/IEC 25010); (vii) estratégias de validade, confiabilidade e reprodutibilidade; e (viii) considerações éticas e de privacidade.

3.1 Natureza e desenho do estudo

Tabela 23 – Caracterização metodológica do estudo.

Dimensão	Descrição
Abordagem	Método quantitativo, com predominância documental (logs, métricas e relatórios técnicos).
Estratégia	Estudo aplicado, exploratório–descritivo, de natureza ex-post-facto.
Unidade de análise	Interações com o EducaChat (sessões, mensagens) e janelas de monitoramento do sistema.
Horizonte temporal	Transversal, com agregações por janelas temporais (hora, dia e semana).
Vínculo avaliativo	Subconjunto do modelo ISO/IEC 25010 (usabilidade, confiabilidade, eficiência e funcionalidade).

Fonte: Elaborado pelo autor (2025).

Trata-se de uma pesquisa aplicada, de caráter exploratório–descritivo e documental, uma vez que descreve padrões de uso e evidências relacionadas aos atributos de qualidade em uso do EducaChat e às suas implicações pedagógicas indiretas, a partir da análise de registros técnicos já existentes do sistema (dados secundários). O delineamento é do tipo *ex post facto*, não havendo intervenção do pesquisador, manipulação de variáveis em campo ou recrutamento direto de participantes.

Segundo Gil (2019), a pesquisa exploratória é particularmente útil em domínios ainda

pouco consolidados, pois possibilita identificar padrões emergentes e levantar hipóteses para estudos subsequentes.

Já a pesquisa descritiva busca sistematizar as características observáveis de um fenômeno, permitindo estabelecer relações entre variáveis sem necessariamente manipulá-las. Na perspectiva de Creswell (2014), esse delineamento documental confere maior transparência e reprodutibilidade, uma vez que os dados analisados foram gerados de forma independente ao pesquisador.

O uso de dados secundários em ambientes digitais também se alinha à noção de estudo *ex post facto*, em que os registros são tratados como evidências históricas de interações e desempenho do sistema (Yin, 2015). Tal abordagem reduz vieses associados ao auto-relato e amplia o potencial de replicação, já que os dados podem ser auditados e processados novamente sob os mesmos parâmetros técnicos.

O desenho metodológico ancora-se, portanto, em três pilares: (a) estatística descritiva dos logs para quantificação de padrões; (b) análise documental de relatórios técnicos e pedagógicos, que subsidia interpretações; e (c) avaliação de atributos de qualidade de *software* com base na ISO/IEC 25010 (funcionalidade, confiabilidade, usabilidade e eficiência de desempenho), garantindo que a análise pedagógica esteja integrada à avaliação técnica do sistema.

3.2 Hipóteses e questões de pesquisa

As hipóteses aqui apresentadas derivam diretamente do problema, dos objetivos e do referencial teórico discutidos nos Capítulos 1 e 2. Com base na fundamentação teórica (Cap. 2) e no escopo metodológico desta dissertação (Cap. 3), formulam-se questões de pesquisa (QP) e hipóteses operacionais (H) orientadas à análise de dados secundários anonimizados (logs e métricas), em convergência com as dimensões da ISO/IEC 25010 priorizadas no estudo (usabilidade, eficiência de desempenho, confiabilidade e funcionalidade).

Questões de pesquisa (QP)

QP1: Em que medida a eficiência de desempenho do EducaChat (latência de resposta) sustenta a fluidez da mediação pedagógica, considerando séries temporais e cauda de latência (P95)?

QP2: O EducaChat apresenta funcionalidade compatível com o uso educacional,

evidenciada por cobertura temática (BNCC) e estabilidade na geração de respostas válidas (sem erro) ao longo do período analisado?

QP3: Os padrões de uso observados nos logs (*msg_count*/sessão, *duration_s*) indicam usabilidade e engajamento técnico consistentes (*proxies* de interação) em diferentes janelas temporais?

QP4: A confiabilidade do serviço (*uptime* e erros 5xx) mantém-se em nível adequado para suportar a mediação contínua, e como ela se relaciona com oscilações de latência e de engajamento?

3.2.1 Hipóteses operacionais (H)

À luz do Cap. 2, das boas práticas em *learning analytics* e do recorte exclusivamente documental, enunciam-se hipóteses testáveis via logs e métricas:

H1(Eficiência de desempenho): A latência do EducaChat apresenta mediana baixa e cauda (P95) controlada ao longo do período, favorecendo a continuidade da interação.

H2(Confiabilidade): O serviço mantém alta disponibilidade (*uptime*) e baixa incidência de erros 5xx, sem picos sustentados que comprometam o uso pedagógico.

H3(Usabilidade/Engajamento–*proxy*): sessões válidas exibem distribuições estáveis de *msg_count* e *duration_s*, com baixa taxa de desistências precoces (sessões ultracurtas), sugerindo fricção reduzida.

H4(Funcionalidade): A cobertura de tópicos (*topic*) é ampla e consistente com a BNCC nos componentes analisados, com diversidade temática não concentrada em poucas categorias.

H5(Relações entre dimensões): janelas com menor latência e ausência de incidentes (5xx) tendem a apresentar maior engajamento técnico (mediana de *msg_count* e P90 de *duration_s*), sugerindo associação positiva entre eficiência/confiabilidade e uso.

H6(Robustez contextual): Os indicadores (latência, engajamento e confiabilidade) mantêm padrões estáveis frente a variações sazonais (dia/semana) e a eventos operacionais documentados (release/manutenção), após controle por janelas e exclusões justificadas.

3.2.2 Escopo e observações

As hipóteses acima são operacionais e compatíveis com a natureza do conjunto de dados (sem medidas diretas de aprendizagem). Assim, inferências pedagógicas são realizadas por proxies (p. ex., engajamento técnico via `msg_count/duration_s`) e trianguladas com documentação técnica/curricular.

Evidências quantitativas serão reportadas em medidas robustas (mediana, quartis, percentis) e séries temporais, com inspeção de *outliers* e confrontação com registros de incidentes, em consonância com a ISO/IEC 25010 e com o plano de análise descrito na Seção 3.2.3.

Tabela 24– Mapeamento entre questões de pesquisa, métricas e fontes de dados.

Questão de Pesquisa (QP)	Métricas / Indicadores	Fonte de dados
QP1 – Eficiência	Latência mediana de resposta; percentis P95 e P99; séries temporais de latência (ms).	Logs do sistema (<code>latency_ms</code>).
QP2 – Funcionalidade	Taxa de respostas válidas; diversidade temática (<code>topic</code>); cobertura conceitual alinhada à BNCC.	Logs anotados do sistema (<code>topic</code> , <code>status</code>).
QP3 – Usabilidade / Engajamento	Distribuição do número de mensagens por sessão (<code>msg_count</code>); duração das sessões (<code>duration_s</code>); taxa de desistência precoce (sessões ultracurtas).	Logs de sessão (<code>session_id</code> , <code>msg_count</code> , <code>duration_s</code>).
QP4 – Confiabilidade	Disponibilidade do serviço (uptime); taxa de erros 5xx; número de incidentes críticos reportados.	Relatórios técnicos de infraestrutura e logs de erro.

Fonte: Elaborado pelo autor (2025).

3.2.3 Plano de análise: abordagem mista em dados secundários

O tratamento dos dados adota uma abordagem de métodos mistos (*mixed methods*) aplicada exclusivamente ao corpus documental (logs), sem uso de instrumentos de coleta direta (como questionários). A estratégia articula:

Análise quantitativa (métrica): focada na performance do sistema. Envolve estatística descritiva (mediana, desvio-padrão, percentis P95) aplicada às variáveis de `latency_ms`, `uptime` e contagem de mensagens (`msg_count`).

O objetivo é responder às questões de pesquisa QP1 (eficiência) e QP4 (confiabilidade);

Análise qualitativa (conteúdo): focada na funcionalidade e adesão curricular. Envolve a análise de conteúdo dos diálogos registrados nos logs. O pesquisador audita amostras das conversas para verificar se o agente: (a) respeitou as regras de prompt (ex: não dar a resposta pronta); e (b) manteve a coerência com a habilidade da BNCC proposta. Isso responde à QP2 (funcionalidade).

3.2.4 Integração analítica

As evidências são trianguladas internamente: cruza-se o dado quantitativo (ex: "sessão durou 40 minutos") com o dado qualitativo (ex: "o diálogo conteve 15 turnos de raciocínio lógico"). Esse cruzamento permite inferir a qualidade técnica da mediação sem depender da opinião subjetiva dos usuários.

3.2.5 Justificativa metodológica

A opção por um delineamento exploratório-descritivo com dados secundários justifica-se por três razões principais: (i) respeita-se a conformidade ética e legal, sem necessidade de intervenção em sala de aula; (ii) assegura-se reprodutibilidade, dado que todos os dados provêm de registros técnicos versionados; e (iii) permitindo concentrar a análise na viabilidade técnico-operacional da mediação pedagógica, conectando diretamente métricas objetivas a hipóteses previamente formuladas (ver Seção 3.2.5).

Johnston (2014) destaca que os estudos baseados em dados secundários são capazes de ampliar o alcance empírico sem os custos logísticos e éticos de coletas diretas, tornando-se especialmente relevantes em contextos educacionais sensíveis.

Além disso, a literatura de métodos (Creswell, 2014; Gil, 2019) enfatiza que a utilização de registros existentes não invalida a robustez do estudo, desde que o processo de tratamento e análise esteja devidamente documentado, como ocorre neste trabalho.

Essa justificativa dialoga ainda com a perspectiva de Yin (2015) sobre estudos de caso em sistemas complexos: quanto mais múltiplas e independentes forem as fontes de dados, maior a validade do conjunto de evidências.

Assim, os logs e relatórios técnicos do EducaChat não são apenas insumos operacionais, mas constituem-se em evidências científicas a serem analisadas de forma sistemática e transparente.

3.2.6 Fluxograma metodológico e cronograma

A Figura 8 apresenta o encadeamento lógico das etapas que estruturam o estudo, contemplando desde a coleta dos dados até a análise final e a síntese das evidências. O fluxograma possibilita visualizar de forma organizada o percurso metodológico adotado, reforçando a transparência do processo investigativo e contribuindo para a reprodutibilidade dos procedimentos descritos.

Figura 8 – Fluxo metodológico da pesquisa, abrangendo as etapas de coleta, ETL, análise e síntese de evidências.



Fonte: Elaborado pelo autor (2025).

3.2.7 Modelo de governança de dados

A governança baseia-se em quatro eixos: (i) papéis e responsabilidades (pesquisador responsável, mantenedor técnico, auditor metodológico); (ii) processos (extração, tratamento, controle de acesso, auditoria e descarte); (iii) políticas (minimização, anoni-

mização, retenção); e (iv) evidências (versionamento, checksums, diários de bordo).

Tabela 25 – Papéis e responsabilidades na governança de dados.

Papel	Responsabilidades
Pesquisador responsável	Aprovar variáveis e recortes; zelar por ética e conformidade; assinar relatórios finais
Mantenedor técnico	Realizar extrações; aplicar anonimização; manter trilhas de auditoria e versionamento
Auditor metodológico	Verificar reprodutibilidade (scripts/notebooks); conferir checksums e estatísticas descritivas

Fonte: Elaborado pelo autor (2025).

3.2.8 Fontes de dados

Foram consideradas três fontes principais, todas sem dados pessoais identificáveis (PII):

- Logs de interação: registros transacionais do backend (*Java Spring Boot + Web-Socket*), contendo metadados de sessões e mensagens (p. ex., `session_id` com hash, `started_at`, `ended_at`, `duration_s`, `role`, `latency_ms`, `topic`).
- Métricas de desempenho: health checks e monitoramento de disponibilidade (p. ex., `uptime_s`, `cpu_load`, `mem_usage`, `http_5xx_count`).
- Documentação técnica: release notes, relatórios operacionais e especificações do EducaChat, em versões públicas/institucionais.

Tabela 26 – Instrumentos analíticos, finalidades e limitações (dados secundários).

Instrumento	Finalidade analítica	Limitação intrínseca
Logs do sistema	Engajamento técnico (msg_count, duration_s), latência e confiabilidade	Não capturam percepção/afetos; leitura é proxy
Métricas operacionais	Disponibilidade (uptime_s), erros 5xx, consumo de recursos	Dependem de sampling/sondas e janelas
Documentação técnica	Contexto de versão, incidentes, mudanças de configuração	Pode ser lacunar; requer triangulação temporal

Fonte: Elaborado pelo autor (2025).

3.2.9 Critérios de inclusão e exclusão

Inclusão: (i) sessões válidas com pelo menos uma mensagem; (ii) mensagens com role e ts definidos; (iii) janelas de monitoramento com uptime_s e contagem de erros.

Exclusão: (i) campos livres com texto potencialmente identificável; (ii) registros inconsistentes (p. ex., duration_s negativo); (iii) dados fora da janela temporal definida.

3.2.9.1 Política de anonimização e minimização

Adotam-se princípios de minimização de dados (*data minimization*) e anonimização em múltiplas camadas, compreendendo: (i) supressão de campos sensíveis; (ii) aplicação de hash irreversível em chaves técnicas, com uso de salt; (iii) agregação temporal por janelas (hora/dia); e (iv) truncamento e categorização de *payloads* textuais na variável topic. Essa abordagem preserva os indicadores analíticos e reduz o risco de reidentificação dos usuários, em conformidade com a Lei Geral de Proteção de Dados Pessoais (BRASIL, 2018).

Tabela 27 - Fontes, critérios e medidas de anonimização.

Fonte	Inclusão/Exclusão	Anonimização aplicada
Logs de interação	Sessões com ≥ 1 mensagem; remoção de campos livres	Hash de IDs (salt); truncamento/categorização de texto
Métricas de desempenho	Latência, uptime, erros 5xx; sem IPs/hosts	Remoção de identificadores de infraestrutura
Documentação técnica	Versões públicas/institucionais	Sem PII; controle de versão

Fonte: Elaborado pelo autor (2025).

3.2.9.2 Ciclo de vida, retenção e descarte

O ciclo de vida dos dados segue as fases: coleta (export autorizada), tratamento (ETL e anonimização), análise (notebooks reprodutíveis), publicação (tabelas/figuras anonimizadas) e descarte seguro. A retenção está limitada ao período de execução da pesquisa e arquivamento técnico mínimo para auditoria, seguido de descarte documentado (shred/secure delete) ou guarda institucional conforme política vigente.

3.3 Segurança da informação e controle de acesso

Os dados são mantidos em repositório privado com versionamento (git), controle de acesso por perfis, autenticação forte e registros de auditoria. Artefatos analíticos (.csv/.xlsx, notebooks) recebem checksums para verificação de integridade. Recomendações de UNESCO (2023) são observadas no que tange à segurança e transparência no uso educacional de IA.

3.3.1 Qualidade, completude e auditoria de dados

Conforme apresentado anteriormente na Tabela 28, a qualidade dos dados é monitorada por: (i) completude (campos essenciais não nulos); (ii) consistência (intervalos válidos; coerência temporal); (iii) confiabilidade (reprocessamento idempotente dos scripts); (iv) rastreabilidade (metadados de export).

Tabela 28 – Verificações de qualidade e auditoria dos dados.

Dimensão	Verificação	Evidência
Completude	Campos essenciais preenchidos (ts, role, duration)	Relatório de nulos/percentuais
Consistência	Durações ≥ 0 ; latência plausível; ordem temporal	Regras de validação (ETL)
Confiabilidade	Execução repetível dos scripts	Hash de <i>outputs</i> ; versão de bibliotecas
Rastreabilidade	Origem e tempo de exportação; parâmetros	Diário de bordo e <i>metadata</i>

Fonte: Elaborado pelo autor (2025).

3.3.2 Matriz de risco e controles

A matriz a seguir sumariza riscos típicos e respectivos controles adotados.

Tabela 29 – Matriz de risco em dados e controles de mitigação.

Risco	Impacto	Controle adotado
Reidentificação indireta	Quebra de privacidade	Anonimização por hash; agregação temporal; truncamento
Inconsistência temporal	Viés na análise	Regras ETL; validação de faixas e ordem
Acesso não autorizado	Violação de confidencialidade	Controle por papéis; repositório privado; logs de acesso
Dependência de script único	Fragilidade metodológica	<i>Peer review</i> e reexecução independente dos notebooks

Fonte: Elaborado pelo autor (2025).

3.3.3 Modelo de dados e dicionário de variáveis

O modelo de dados do EducaChat foi concebido para registrar interações de usuários de forma anonimizada, permitindo análises pedagógicas e de qualidade de *software* sem expor informações pessoais identificáveis (PII). Nesta seção, são descritos o esquema lógico dos registros, o modelo relacional simplificado e o dicionário de variáveis operacionais.

Segundo Rob (2016), a modelagem de dados em sistemas educacionais deve priorizar consistência e auditabilidade, além de mecanismos de anonimização que assegurem conformidade legal (LGPD) e validade científica. Dessa forma, as tabelas e variáveis foram especificadas para equilibrar precisão técnica, minimização de dados e relevância pedagógica.

3.3.4 Esquema lógico dos logs

O esquema foi simplificado em três entidades principais: sessions, messages e health_metrics. A Tabela 30 apresenta os campos e descrições.

Tabela 30 – Esquema lógico simplificado dos registros do EducaChat.

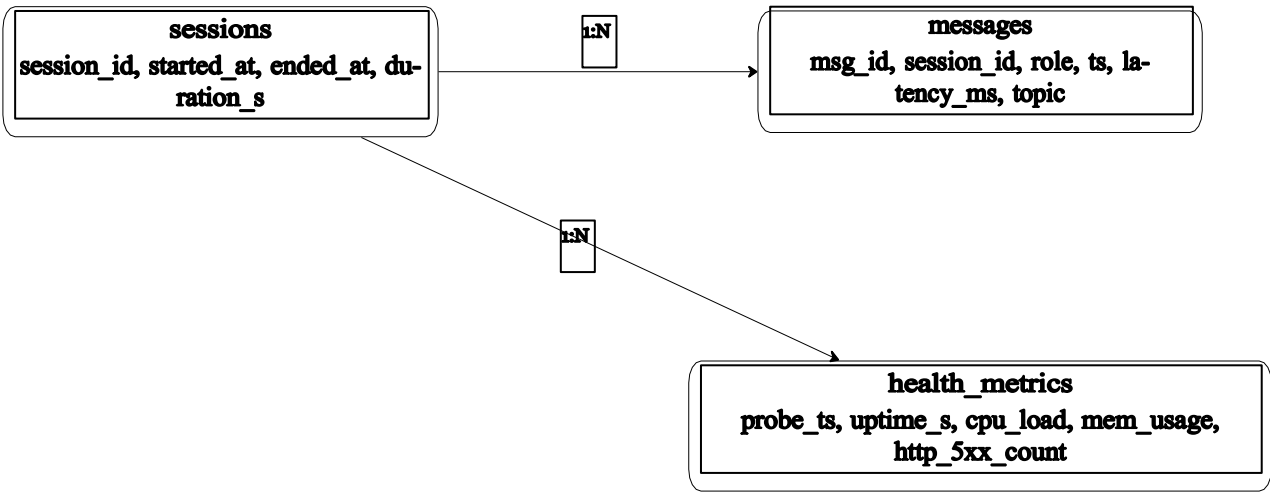
Tabela	Chaves e campos	Descrição
sessions	session_id (hash), started_at , ended_at , duration_s	Representa cada sessão de uso; IDs anonimizados; duração calculada por diferença temporal
messages	msg_id (hash), session_id (hash), role , ts , latency_ms , topic	Mensagens associadas a uma sessão; inclui papel (usuário, assistente, sistema), latência e tema categorizado
health_metrics	probe_ts , uptime_s , cpu_load , mem_usage , http_5xx_count	Indicadores técnicos de disponibilidade, consumo de recursos e estabilidade do serviço

Fonte: Elaborado pelo autor (2025).

3.3.5 Modelo relacional

A Figura 9 apresenta o modelo relacional simplificado do EducaChat. No qual cada sessão pode conter múltiplas mensagens e estar associada a métricas de disponibilidade.

Figura 9 – Modelo relacional das tabelas de dados do EducaChat.



Fonte: Elaborado pelo autor (2025).

A Tabela 31 detalha as variáveis operacionais e seus usos analíticos. Além de definições técnicas, são incluídas as finalidades pedagógicas e de qualidade de *software*, conforme preconizado pela ISO/IEC 25010.

3.3.5.1 Dicionário de variáveis

Tabela 31 – Dicionário de variáveis: definições operacionais e finalidades.

Variável	Definição operacional	Uso analítico
duration_s	Tempo total da sessão em segundos (ended_at - Proxy de engajamento; análise de permanência started_at)	
msg_count	Número de mensagens por sessão	Volume de interação; dispersão e engajamento técnico
latency_ms	Tempo de resposta do assistente em milissegundos	Eficiência de desempenho; análise de outliers
topic	Categoria temática atribuída a cada mensagem	Diversidade de conteúdo; aderência à BNCC
uptime_s	Tempo de disponibilidade contínua do sistema	Confiabilidade do serviço
cpu_load	Percentual de uso da CPU em cada janela de monitoramento	Estabilidade operacional; relação com latência
mem_usage	Percentual de uso da memória RAM	Identificação de gargalos; monitoramento de eficiência
http_5xx_count	Número de erros de servidor em cada janela	Indicador de falhas críticas; confiabilidade

Fonte: Elaborado pelo autor (2025).

Tabela 32 – Regras de validação e tratamento por variável.

Variável	Regra de validação	Tratamento ETL
duration_s	≥ 0 ; coerência com started_at e ended_at	Recalcular se faltante; descartar se negativo persistente
msg_count	≥ 1 por sessão válida	Derivar por groupby(session_id)
latency_ms	Faixa plausível (p.ex., 10–60000)	Tratamento de outliers (1,5 IQR) + checagem de incidentes
topic	Categoria presente na taxonomia	Normalizar rótulos; mapear como “desconhecido”
uptime_s	Crescente por janela; resets por reinício	Marcar janelas com reinício; anotar manutenção
http_5xx_count	Inteiro ≥ 0	Consolidar por janela; cruzar com latência

Fonte: Elaborado pelo autor (2025).

3.3.5.2 Cobertura e escopo temporal

A Tabela 33 apresenta o escopo temporal da coleta e o volume de registros disponíveis para análise. Essa caracterização assegura a validade estatística e a representatividade dos dados processados.

Tabela 33 – Cobertura dos dados e escopo temporal.

Item	Descrição
Janela temporal	2024-11-04 a 2024-11-05
Volume de sessões	30–40 (estimado)
Volume de mensagens	500–700 (estimado)
Indicadores de sistema	Uptime, latência, erros 5xx
Observações	Variações de volume; registros de manutenção identificados

Fonte: Elaborado pelo autor (2025).

A escolha dessa janela temporal concentrada (04–05 de novembro de 2024) se justifica pelo caráter controlado da coleta, realizada em período delimitado e em ambiente supervisionado.

Essa estratégia reduziu a interferência de fatores externos e garantiu maior consistência na triangulação entre métricas operacionais (logs e indicadores de saúde do sistema). Embora o período de análise seja relativamente curto, sua representatividade é assegurada pela densidade de sessões registradas e pela diversidade de interações captadas, cumprindo o papel metodológico de estudo empírico exploratório.

A seguir, apresenta-se a estatística descritiva das sessões e mensagens, a fim de mapear padrões de engajamento, intensidade de uso e permanência nas interações com o EducaChat.

3.3.6 Procedimentos de coleta e tratamento (ETL)

O tratamento dos dados do EducaChat foi estruturado em um *pipeline ETL* (*Extract, Transform, Load*), prática consolidada em processos de gestão de dados analíticos (Inmon, 2005). Essa abordagem visa assegurar consistência, rastreabilidade e qualidade dos registros, conforme recomendado por Creswell (2014) em pesquisas quantitativas que dependem de grandes volumes de evidências.

Segundo Bardin (2011), a análise de dados só é possível quando existe um processo prévio de organização e categorização rigorosa. No presente estudo, isso se traduz na seleção de variáveis pré-definidas, conforme apresentado anteriormente na Tabela 10, bem como na exclusão de campos sensíveis e na anonimização sistemática dos identificadores.

3.3.6.1 Extração (*Extract*)

Os dados foram exportados em formatos *CSV* e *JSON* a partir do *backend* (*Java Spring Boot + WebSocket*) e dos serviços de monitoramento. A extração obedeceu a dois princípios: (i) restrição ao dicionário de variáveis previamente estabelecido (Seção 3.3.3), assegurando alinhamento aos objetivos da pesquisa; e (ii) registrabilidade, com armazenamento de metadados de cada extração (data, versão do sistema e filtros aplicados).

3.3.6.2 Transformação (*Transform*)

A etapa de transformação envolveu um conjunto de operações automatizadas, conforme boas práticas de qualidade de dados (Kimball & Ross, 2013):

- Limpeza: remoção de registros duplicados; trim de espaços; normalização de *times-tamps* (UTC → BRT).
- Anonimização: aplicação de *hash* irreversível com salt em *session_id* e *msg_id*; supressão de campos livres.
- Enriquecimento: derivação de variáveis agregadas como *msg_count*, *duration_s* e agregações por janelas temporais (hora/dia), permitindo análise longitudinal.
- Validação: checagem de consistência (valores não negativos; faixas plausíveis de latência e duração); comparação com registros de incidentes documentados para evitar distorções analíticas.

Esse processo garante que os dados, ao chegarem à etapa de análise, sejam confiáveis, comparáveis e reprodutíveis, atributos essenciais em pesquisas científicas (Yin, 2015).

Tabela 34 – Mapeamento de variáveis derivadas e consistências.

Variável derivada	Como é calculada	Consistências verificadas
<i>msg_count</i> (por sessão)	Contagem de mensagens por <i>session_id</i>	≥ 1 ; sincronismo com <i>duration_s</i> .
<i>duration_s</i>	<i>ended_at</i> - <i>started_at</i> (segundos)	Ordem temporal; não negativo.
<i>latency_p95</i>	Percentil 95 por janela/sessão	Relação com incidentes 5xx; estabilidade.
<i>uptime_rel</i> (%)	Uptime observado / janela	Resets por manutenção; janelas incompletas.
<i>diversidade_topics</i>	Índice (p.ex. Shannon) por período	Taxonomia estável; % "desconhecido" controlada.

Fonte: Elaborado pelo autor (2025).

3.3.6.3 Carga (*Load*)

Os dados tratados foram armazenados em dois formatos complementares:

(i) planilhas analíticas (.xlsx), para visualização e estatística descritiva; (ii) tabelas .csv versionadas em repositório Git, compondo o repositório analítico do estudo.

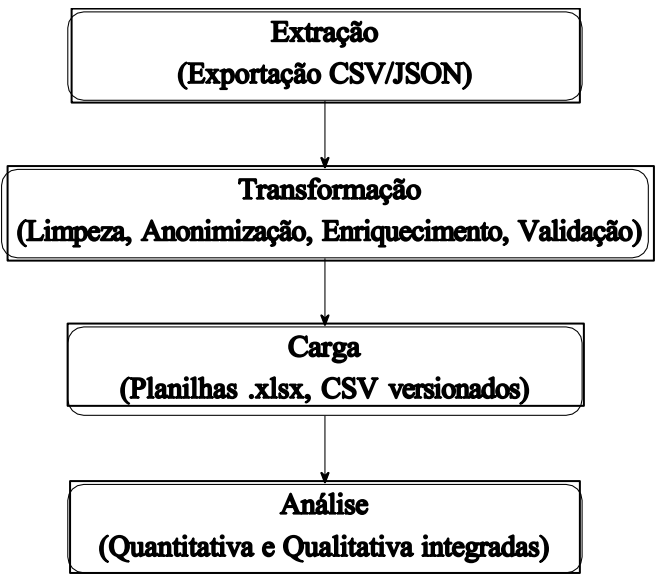
Esse procedimento viabiliza reprodutibilidade, uma vez que qualquer auditor ou pesquisador poderá reconstruir as análises a partir dos mesmos arquivos, *scripts* e parâmetros, em conformidade com recomendações de Lincoln e Guba (1985) sobre confiabilidade em pesquisas qualitativas e quantitativas.

Tabela 35 – Pipeline ETL e garantias de qualidade.

Fase	Atividade principal	Garantia metodológica
Extração	Exportação CSV/JSON com metadados de versão	Restrição ao dicionário; registrabilidade da coleta
Transformação	Limpeza, anonimização, enriquecimento, validação	Scripts automatizados; checagem de consistência
Carga	Armazenamento em .xlsx/.csv com versionamento Git	Reprodutibilidade e auditoria independente

Fonte: Elaborado pelo autor (2025), com base em Inmon (2005) e Kimball & Ross (2013).

Figura 10 – Fluxo ETL e análise de dados do EducaChat.



Fonte: Elaborado pelo autor (2025).

3.3.7 Plano de validação

A validade da pesquisa está associada à consistência metodológica, à coerência entre objetivos, hipóteses e indicadores, bem como à robustez das evidências analíticas derivadas de dados secundários. O plano de validação estrutura-se em quatro eixos complementares.

3.3.7.1 Validade interna

A Validade interna é assegurada pela coerência lógica entre problema de pesquisa, objetivos, hipóteses e indicadores analíticos definidos no Capítulo 1. Adicionalmente, utiliza-se como evidência primária registros técnicos objetivos (logs do sistema e métricas de desempenho), o que reduz vieses interpretativos associados ao auto-relato. O estudo também adota delimitação clara do escopo analítico e temporal, com controle de variáveis de confusão por meio de agregações por janelas e exclusão justificada de registros anômalos.

3.3.7.2 Validade externa / transferibilidade analítica

A validade externa é tratada por meio de reconhecimento de que os achados se referem a um conjunto específico de dados técnicos e documentais, sem pretensão de generalização estatística. A descrição detalhada dos procedimentos analíticos e das métricas utilizadas permite a transferência analógica do modelo de análise para contextos tecnológicos semelhantes, enquanto as limitações e condições de aplicação explicitadas no Capítulo 8.

3.3.7.3 Confiabilidade dos indicadores

A confiabilidade dos indicadores é assegurada por meio de procedimentos sistemáticos de verificação e controle de qualidade dos dados. Para isso, realiza-se revisão cruzada de métricas provenientes de diferentes registros técnicos (logs de sessão, logs de infraestrutura e relatórios operacionais), bem como verificação de consistência estatística, incluindo identificação de valores extremos, análise de variabilidade por janelas temporais e inspeção de padrões recorrentes. Adicionalmente, adota-se controle de qualidade do *pipeline ETL* (Seção 3.3.6 e Apêndice A), com validações de integridade e rastreabilidade em cada etapa do processamento.

3.3.8 Plano de Análise

O plano de análise adota uma abordagem mista sequencial, processando o corpus documental para extrair indicadores de qualidade ISO/IEC 25010.

3.3.8.1 Análise quantitativa (performance e estabilidade)

Serão calculadas medidas de tendência central e dispersão para variáveis críticas (*latency_ms*, *duration_s*). Histogramas e *boxplots* permitirão identificar gargalos de performance e estabilidade do sistema em uso real.

3.3.8.2 Análise de conteúdo automatizada (topic)

A dimensão qualitativa foca na classificação dos tópicos abordados nas mensagens (*topic modeling*), verificando a aderência do agente aos temas curriculares da BNCC programados, sem entrar na análise subjetiva da experiência do usuário.

Tabela 36 – Janelas analíticas e justificativas de uso.

Janela	Uso principal	Risco/atenção
Hora	Picos de latência e 5xx; microeventos	Ruído elevado; necessidade de <i>smoothing</i> .
Dia	Ritmo de uso; disponibilidade diária	Efeito de feriados/cronograma escolar.
Semana	Tendências e sazonalidade curta	Mistura de regimes (avaliações/projetos).

Fonte: Elaborado pelo autor (2025).

3.3.9 Análise qualitativa (Documental e categorial)

A dimensão interpretativa contempla:

(i) Análise documental de relatórios técnicos e registros operacionais.

(ii) Categorização temática das interações registradas (*topic*), com base em taxonomia alinhada à BNCC.

Adota-se a análise de conteúdo em sua vertente documental, conforme Bardin (2011), envolvendo codificação, categorização e interpretação contextualizada dos registros técnicos. As categorias temáticas são refinadas por codificação aberta, com elaboração de memos analíticos sobre padrões de demanda e aderência curricular.

3.3.9.1 Indicadores-síntese

A partir das análises realizadas, são construídos indicadores-síntese agrupados em duas dimensões: técnica e pedagógica indireta. Essa integração permite relacionar atributos de qualidade de *software* (ISO/IEC 25010) a implicações analíticas sobre mediação e engajamento técnico.

Engajamento técnico (*proxy*): mediana de *msg_count* por sessão e percentil 90 de *duration_s*. Eficiência de desempenho: mediana e P95 de *latency_ms*.

Confiabilidade: disponibilidade relativa por janela temporal (*uptime_s*) e taxa de *http_5xx_count*.

Aderência temática: diversidade e frequência das categorias *topic*, com análise de concentração temática.

3.3.9.2 Fluxo analítico integrado

O percurso analítico adotado na pesquisa: dados secundários → *pipeline ETL* → análises quantitativas → categorização documental → indicadores → evidências empíricas.

Esse fluxo reforça a transparência, a rastreabilidade e a auditabilidade do processo investigativo, aspectos essenciais para a confiabilidade de estudos baseados em dados observacionais.

A organização sequencial do fluxo analítico permite que cada etapa seja documentada e replicada, assegurando clareza quanto às transformações aplicadas aos dados e às inferências realizadas a partir dos indicadores técnicos. Tal abordagem é compatível com princípios de rigor metodológico e auditabilidade defendidos na literatura sobre pesquisa aplicada e análise documental.

3.4 Avaliação de qualidade (ISO/IEC 25010)

A avaliação da qualidade do EducaChat fundamenta-se no modelo ISO/IEC 25010 (2011), referência internacional em Engenharia de *Software*.

No presente estudo, foram priorizados quatro atributos centrais: Funcionalidade, Confiabilidade, Usabilidade e Eficiência de desempenho, selecionados em função da natureza da aplicação e do escopo investigativo.

3.4.1 Funcionalidade

A funcionalidade refere-se ao grau em que o sistema fornece funções que atendem às necessidades explícitas. No EducaChat, avalia-se a cobertura temática (aderência à BNCC) e a taxa de respostas geradas sem erro de API, verificando se o agente entrega o conteúdo programado.

3.4.2 Confiabilidade

A confiabilidade medida por meio do tempo de atividade (*uptime*) e da incidência de erros de servidor (HTTP 5xx). Pedagogicamente, a confiabilidade traduz-se na capacidade do sistema de oferecer apoio contínuo, sem interrupções que quebrem o fluxo de estudo.

3.4.3 Usabilidade

No contexto de logs, a usabilidade é inferida pela fluidez da interação. A métrica principal é a quantidade de mensagens por sessão (*msg_count*) e a taxa de desistências precoces (sessões < 1 minuto), interpretadas como indicadores de fricção técnica ou desinteresse, e não necessariamente de frustração emocional (que exigiria autorrelato).

3.4.4 Eficiência de desempenho

A métrica central é a latência de resposta (*latency_ms*), avaliada pela mediana e pelo percentil 95 (P95). Atrasos elevados (alta latência) degradam a experiência do usuário e podem inviabilizar a mediação síncrona.

3.4.5 Quadros e indicadores

A Tabela 37 resume a operacionalização dos atributos selecionados, enquanto a Tabela 38 relaciona-os diretamente às questões de pesquisa (QP1–QP3), evidenciando o alinhamento entre engenharia de *software* e pedagogia. Complementarmente, a Tabela 39 explicita a rastreabilidade entre os atributos da ISO/ IEC 25010, métricas técnicas e evidências analíticas esperadas, fortalecendo a coerência entre coleta de dados, indicadores e interpretação dos resultados.

Tabela 37 - Atributos da ISO/IEC 25010 e métricas operacionais no EducaChat.

Atributo	Descrição de avaliação	Métrica/Indicador
Funcionalidade	Adequação das funções às necessidades pedagógicas	Cobertura de topic; taxa de respostas válidas
Confiabilidade	Manutenção do desempenho sob condições esperadas	uptime_s; http_5xx_count
Usabilidade	Facilidade de interação e clareza das pistas	msg_count/sessão; desistências precoces
Eficiência de desempenho	Tempos de resposta sob carga variável	Mediana e P95 de latency_ms

Fonte: Adaptado de ISO/IEC 25010 (2011); Pressman (2011); Sommerville (2016).

Tabela 38 - Vínculo entre ISO/IEC 25010.

Atributo ISO	Indicadores	Contribuição às QPs
Usabilidade	msg_count, desistências precoces	QP1/QP3: engajamento e mediação
Eficiência	latência (mediana, P95)	QP1: fluidez e continuidade do raciocínio
Confiabilidade	uptime_s, erros 5xx	QP3: estabilidade do apoio pedagógico
Funcionalidade	Cobertura de tópicos	QP2: personalização e aderência curricular (BNCC)

Fonte: Elaborado pelo autor (2025).

Tabela 39 - Rastreabilidade ISO/IEC 25010 → dados → evidências.

Atributo ISO	Dados / Métricas	Evidência esperada
Usabilidade	msg_count/sessão; desistências	Engajamento técnico; fricção reduzida.
Eficiência	latência (mediana, P95)	Fluidez de interação; cauda controlada.
Confiabilidade	uptime_s; http_5xx_count	Estabilidade e baixa falha.
Funcionalidade	topic (BNCC)	Cobertura curricular e relevância.

Fonte: Elaborado pelo autor (2025).

3.4.6.1 Validade, confiabilidade e reprodutibilidade

A robustez metodológica sustenta-se na transparência dos procedimentos e na mitigação de ameaças à validade, conforme recomendações de Yin (2015) e Creswell (2014).

3.4.6.2 Ameaças e estratégias de mitigação

Segundo Yin (2015), toda investigação empírica enfrenta riscos de validade interna, externa e de construto. No caso desta pesquisa, esses riscos são enfrentados da seguinte forma:

- Validade interna: reconhece-se que logs são *proxies*. A mitigação ocorre pelo cruzamento de métricas (ex: tempo de sessão vs. número de mensagens) para evitar falsos positivos de engajamento.
- Validade externa: a descrição detalhada da infraestrutura e do prompt do sistema permite a replicabilidade do experimento em outros contextos escolares.
- Confiabilidade dos dados: garantida pelo *pipeline ETL* automatizado e uso de checksums, impedindo manipulação manual dos registros.

3.4.6.3 Confiabilidade e auditoria do processo

O estudo assegura a reprodutibilidade através da: (i) disponibilização dos *notebooks* de análise (*Python/R*) com sementes (*seeds*) definidas; e (ii) documentação das versões das bibliotecas e parâmetros de extração utilizados.

3.4.6.4 Reprodutibilidade

A reprodutibilidade refere-se à possibilidade de repetir a análise obtendo os mesmos resultados (Peng, 2011). No presente estudo, tal requisito é assegurado por:

- Disponibilização de notebooks analíticos em *Python* (*pandas, numpy, matplotlib*) e *R* (*tidyverse*), com seed definida.
- Documentação das rotinas de preprocessing, incluindo filtros aplicados, regras de anonimização e derivação de variáveis.
- Estruturação de artefatos de saída (tabelas e figuras) numeradas e vinculadas ao texto da dissertação.

Tabela 40 - Checklist de reprodutibilidade e auditoria.

Item	Status/Descrição
Scripts ETL versionados (Git)	commit/tag (<i>preencher</i>)
Dicionário de variáveis	Tabela 12
Arquivos de entrada	caminhos/nomes (<i>preencher</i>)
Notebooks analíticos	seed e versões de libs (<i>preencher</i>)
Log de execução	data/hora do <i>export</i> e parâmetros
Artefatos de saída	tabelas/figuras numeradas neste capítulo

Fonte: Elaborado pelo autor (2025).

3.4.6.5 Estratégias adicionais de robustez

Além das práticas anteriores, o estudo adotou medidas adicionais para aumentar sua robustez metodológica:

- Replicação de análises com subconjuntos de dados.
- Revisão independente (*peer review*) de scripts e resultados.
- Armazenamento seguro em repositório com controle de acesso.
- Documentação transparente das decisões analíticas.

Essas estratégias alinham-se ao princípio de credibilidade proposto por Lincoln & Guba (1985), fortalecendo a confiabilidade dos achados e a utilidade do estudo para futuras investigações.

3.4.6.6 Considerações éticas e de privacidade

A pesquisa adota uma postura de "Privacidade por *Design*" (*Privacy by Design*), operando estritamente sobre dados secundários técnicos, sem identificação de sujeitos.

3.4.6.7 Limitações éticas e riscos residuais

O tratamento dos dados segue rigorosamente a Lei Geral de Proteção de Dados (Lei nº 13.709/2018). Aplicam-se os princípios da:

Minimização: coleta apenas dos metadados técnicos indispensáveis (latência, *timestamps*, tópicos).

Anonimização: aplicação de *hashing* irreversível nos identificadores de sessão (*session_id*) e remoção de qualquer dado pessoal (PII) dos payloads de texto antes da análise.

Finalidade: uso exclusivo para auditoria de qualidade de *software* e desempenho de infraestrutura.

Portanto, mesmo com o cumprimento das diretrizes da LGPD, e das recomendações da UNESCO/OCDE, permanece necessário interpretar os resultados com cautela, reconhecendo a existência de incertezas e de riscos residuais inerentes ao uso de IA em ambientes educacionais.

Tabela 41 – Riscos residuais e respectivas mitigações.

Risco residual	Descrição	Mitigação adotada
Reidentificação indireta	Cruzamento com bases externas	Hash com <i>salt</i> , agregação temporal, minimização.
Interpretação indevida	Tomar <i>proxy</i> como aprendizagem	—
Viés de disponibilidade	Janelas incompletas / <i>maintenance</i>	—

Fonte: Elaborado pelo autor (2025).

3.4.6.8 Uso de dados secundários anonimizados

Por limitar-se à análise de logs de sistema e métricas de servidor, a pesquisa não se caracteriza como intervenção clínica ou experimental com seres humanos, dispensando os trâmites da Resolução CNS 510/2016, uma vez que o foco recai sobre o objeto técnico (*software*) e não sobre o sujeito.

3.4.6.9 Anonimização e governança de dados

Alinhado às recomendações da UNESCO (2023) sobre IA na educação, o estudo

promove a transparência ao auditar se o sistema opera dentro dos parâmetros esperados, sem vieses técnicos de indisponibilidade que poderiam prejudicar grupos específicos de alunos (ex: aqueles com conexão mais lenta).

3.4.7 Conformidade ética e legal

A Tabela 42 sintetiza os principais marcos normativos considerados e sua aplicação no presente estudo. triangulação teórica e documental; ressalvas no texto. Marcação explícita de janelas; exclusão ou nota técnica.

Tabela 42 – Conformidade ética e legal assegurada na pesquisa.

Referência normativa	Aplicação nesta pesquisa
Lei Geral de Proteção de Dados (Lei nº 13.709/2018)	Minimização, anonimização e finalidade legítima dos dados
Diretrizes UNESCO (2023)	Ética no uso de IA: transparência, inclusão e <i>privacy by design</i>
Relatórios OCDE (2023)	Responsabilidade, robustez técnica e redução de vieses algorítmicos

Fonte: Elaborado pelo autor (2025).

4. MODELAGEM DO SISTEMA

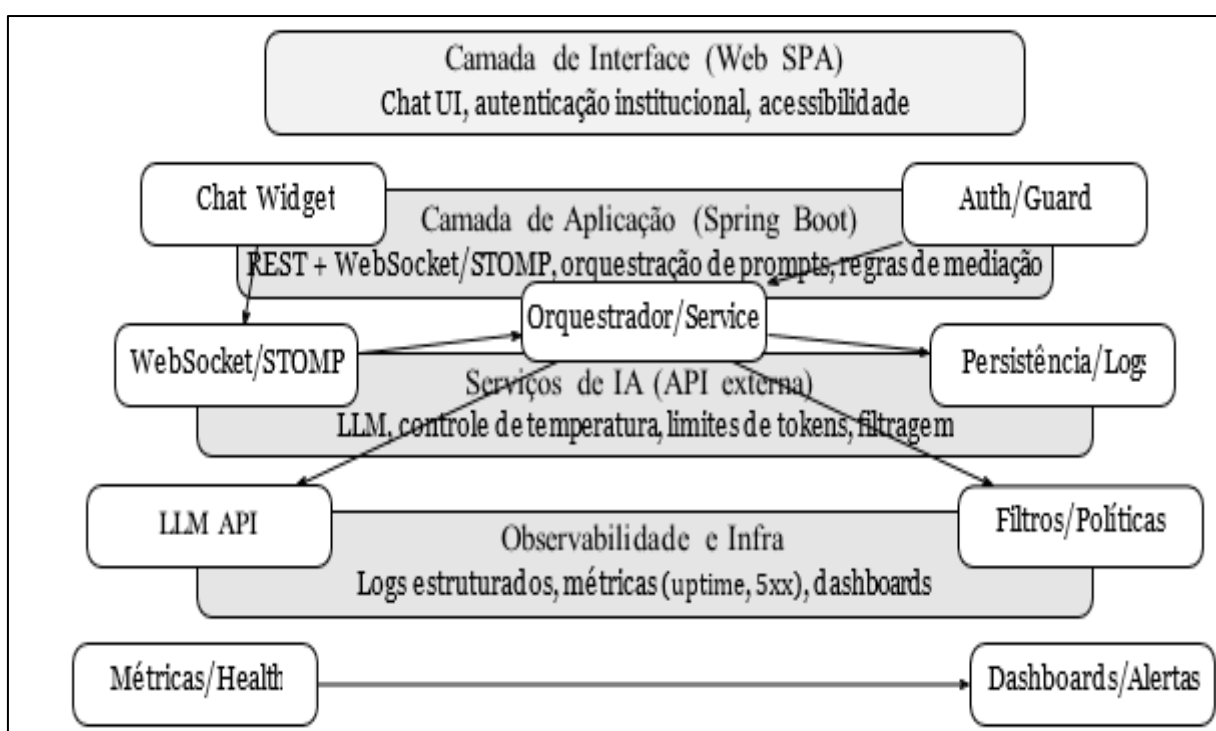
Este capítulo descreve a modelagem do EducaChat, cobrindo visão arquitetural, requisitos, casos de uso e fluxos de interação entre componentes. O objetivo é oferecer uma especificação técnica rastreável que dialogue com os resultados empíricos (Cap. 6) e com a avaliação do sistema (Cap. 7).

A modelagem segue um recorte pragmático: prioriza elementos necessários para sustentar pistas graduais, baixa latência e telemetria de uso, alinhados à ISO/IEC 25010 e às demandas pedagógicas mapeadas nesta dissertação.

4.1 Visão geral da arquitetura

A **Figura 11** apresenta uma visão macro da arquitetura lógica do EducaChat, organizada em quatro camadas: interface, aplicação, serviços de IA e observabilidade/infra. O foco é suportar conversação em tempo real via *WebSocket*, persistir logs estruturados e expor métricas de saúde e desempenho.

Figura 11 - Arquitetura lógica do EducaChat.



Fonte: Elaborado pelo autor (2025).

4.1.1 Racional de desenho

A separação entre orquestração (serviços) e UI reduz acoplamento; o uso de *WebSocket* mantém a conversa  o fluida; *guards* aplicam pol  ticas (pistas graduais, linguagem inclusiva); e a trilha de observabilidade garante mensura  o cont  nua dos atributos ISO/IEC 25010 (Efici  ncia, Confiabilidade, Usabilidade).

4.2 Requisitos do sistema

Os requisitos foram organizados em funcionais (RF) e n  o funcionais (RNF), com rastreabilidade   s QPs e    ISO/IEC 25010 (quando aplic  vel).

Tabela 43– Requisitos funcionais (RF).

ID	Descri��o	Rastreabilidade
RFo1	Manter conversa��o <i>realtime</i> via <i>WebSocket</i> / <i>STOMP</i> entre usu��rio e agente.	QP1; ISO/IEC 25010 (efici��ncia)
RFo2	Orquestrar pistas graduais e evitar respostas finais diretas por padr��o.	QP3; Diretriz pedag��gica
RFo3	Registrar logs estruturados de sess��es e mensagens, sem PII.	Metodologia/ETL; ISO/IEC 25010 (confiabilidade)
RFo4	Classificar topic de mensagens para an��lise tem��tica alinhada �� BNCC.	QP2; Resultados/Temas
RFo5	Disponibilizar m��tricas de sa��de (uptime, 5xx, lat��ncia).	ISO/IEC 25010 (confiabilidade/efici��ncia)
RFo6	Aplicar <i>guards</i> de conte��do: linguagem inclusiva, privacidade e seguran��a.	��tica/LGPD; UNESCO/OCDE
RFo7	Persistir metadados de lat��ncia e dura��o da sess��o para KPIs.	Resultados/KPIs

Fonte: Elaborado pelo autor (2025).

Tabela 44 – Requisitos não funcionais (RNF).

ID	Descrição	Rastreabilidade
RNFo1	Latência mediana $\leq$ (<i>definir</i>) ms; P95 $\leq$ (<i>definir</i>) ms.	ISO/IEC 25010 (eficiência); Cap. 7
RNFo2	Disponibilidade $\geq$ (<i>definir</i>)% no período letivo.	ISO/IEC 25010 (confiabilidade)
RNFo3	Segurança e privacidade: minimização de dados, anonimização por <i>hash</i> +salt.	LGPD; Metodologia/Ética
RNFo4	Usabilidade: mensagens claras; diretrizes de microcopy e acessibilidade (<i>ARIA</i>).	ISO/IEC 25010 (usabilidade)
RNFo5	Observabilidade: <i>health checks</i> expostos; logs com <i>trace-id</i> .	KPIs/ETL; Cap. 7
RNFo6	Escalabilidade: suportar <i>burst</i> de n usuários simultâneos sem degradação crítica.	Resultados/Latência
RNFo7	Reprodutibilidade: versões <i>semver</i> , <i>checksums</i> de artefatos analíticos.	Metodologia/Reprodutibilidade

Fonte: Elaborado pelo autor (2025).

Observação. Os limites (*targets*) parametrizados nos RNFs devem ser ajustados ao contexto operacional (infra e calendário escolar) e validados nos painéis de observabilidade, garantindo aderência contínua aos indicadores do Cap. 6.

5. IMPLEMENTAÇÃO DO SISTEMA

Este capítulo descreve o processo de implementação do EducaChat, detalhando o ambiente de desenvolvimento, as tecnologias utilizadas, a organização do código, os principais componentes desenvolvidos e as soluções técnicas adotadas para superar desafios práticos. O objetivo é demonstrar como os requisitos definidos na modelagem foram efetivamente traduzidos em um protótipo funcional.

5.1 Ambiente de desenvolvimento e execução

A implementação foi conduzida em ambiente controlado, com especificações mínimas de *hardware* e *software* apresentadas na Tabela 45. A definição do ambiente é essencial para garantir reprodutibilidade e validação.

Tabela 45 - Ambiente de desenvolvimento e execução.

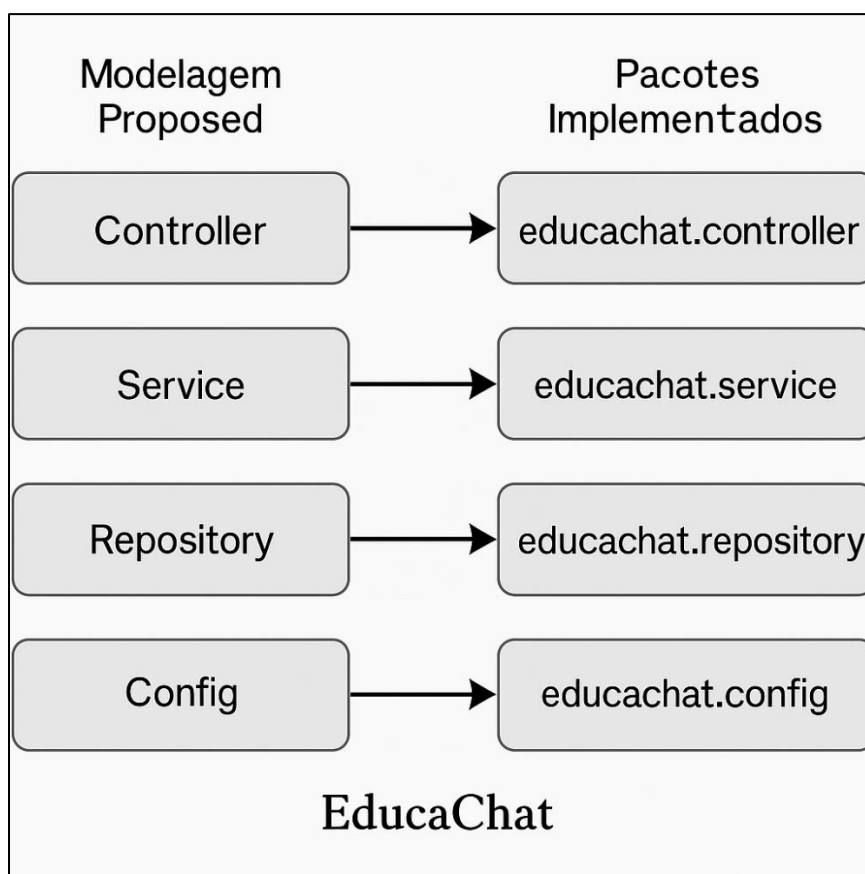
Item	Descrição / Versão
Sistema Operacional	Ubuntu 22.04 LTS / macOS 13.x
IDE	IntelliJ IDEA Community Edition
Linguagem	Java 17 (LTS)
Framework	Spring Boot vX.Y.Z
Comunicação em tempo real	WebSocket/STOMP
Banco de Dados	PostgreSQL vXX
Gerenciamento de dependências	Maven 3.x
Servidor de aplicação	Nginx + HTTPS (Let's Encrypt)
Ferramentas auxiliares	Docker, GitHub, Certbot

Fonte: Elaborado pelo autor (2025).

5.2 Arquitetura de código

A implementação do EducaChat seguiu a arquitetura em camadas definida no Capítulo 4, assegurando modularidade, manutenibilidade e adequada separação de responsabilidades. A Figura 12 apresenta a correspondência entre a modelagem arquitetural proposta e os pacotes efetivamente implementados no sistema, evidenciando a distribuição das funcionalidades nos módulos *controller*, *service*, *repository* e *config*. Esse mapeamento explicita como os componentes lógicos foram traduzidos em estruturas concretas de código, garantindo o alinhamento entre o projeto conceitual e a solução implementada.

Figura 12 – Arquitetura de código implementada do EducaChat.



Fonte: Elaborado pelo autor (2025).

5.3 Principais componentes

5.3.1 Backend (*Spring Boot*)

O núcleo da aplicação foi implementado em *Spring Boot*, incluindo *endpoints REST* para gerenciamento de sessões e mensagens. A Tabela 46 resume os principais *endpoints* implementados.

Tabela 46 - Endpoints principais do backend.

Endpoint	Descrição	Método HTTP
/api/session/start	Inicia nova sessão	POST
/api/session/id	Consulta sessão existente	GET
/api/message/send	Envia mensagem ao chatbot	POST
/api/health	Retorna estado da aplicação	GET

Fonte: Elaborado pelo autor (2025).

5.4 Comunicação em tempo real (*WebSocket/STOMP*)

Foi implementada uma camada de comunicação baseada em *WebSocket* com *STOMP*, permitindo interações em tempo real entre estudante e agente. Esse mecanismo garante baixa latência e fluidez conversacional.

5.4.1 Integração com API de IA

A integração com a API da *OpenAI* foi realizada por meio de chamadas assíncronas, incluindo tratamento de limites de taxa e controle de *tokens*.

5.4.2 Monitoramento e logs

O sistema conta com monitoramento contínuo, coletando métricas de *uptime*, latência e erros de servidor, armazenados em logs estruturados para análise posterior.

5.4.3 Procedimentos de teste inicial

Os testes funcionais básicos, descritos na Tabela 47, precederam a avaliação formal (Capítulos 6 e 7).

Tabela 47 - Testes iniciais realizados no protótipo.

Cenário	Resultado esperado
Início de sessão	Criação de sessão válida e retorno de ID único
Envio de mensagem	Mensagem registrada no log e resposta do agente
Monitoramento	Endpoint <code>/health</code> retornando status ativo
Latência	Tempo de resposta inferior a 2 segundos em 95% dos casos

Fonte: Elaborado pelo autor (2025).

5.4.4 Desafios e soluções

Durante a implementação, alguns desafios foram enfrentados:

- Gerenciamento de latência: implementados mecanismos de *retry* e filas assíncronas.
- Integração com API de IA: controle de limites de taxa e *tokens*.
- Segurança: configuração de *HTTPS* via *Nginx* + *Certbot*.

Essas soluções asseguraram que o protótipo atendesse aos requisitos técnicos e pedagógicos.

5.4.5 Síntese e transição

Em resumo, a implementação consolidou a passagem da modelagem para um protótipo funcional do EducaChat, integrando *backend*, comunicação em tempo real, API de IA e mecanismos de monitoramento. No próximo capítulo, será apresentada a avaliação da qualidade, discutindo aspectos técnicos e pedagógicos com base nos indicadores coletados.

5.4.6 Relato técnico/documental do sistema

Esta seção apresenta um relato técnico/documental do EducaChat, contemplando arquitetura, versões, cenário de testes, exemplos de registros (sem PII) e evidências que embasam a análise descrita na Metodologia (Cap. 3).

Não houve coleta com seres humanos; todas as evidências são secundárias e anonimizadas. O objetivo é assegurar transparência metodológica e rastreabilidade analítica, conforme preconizado por Siemens (2013) e Ferguson (2012) nos estudos de *learning analytics*.

5.4.7 Arquitetura e versões

O *backend* é implementado em *Java Spring Boot*, a comunicação em tempo real via *WebSocket* e o processamento de linguagem natural por API de IA. O monitoramento coleta métricas de disponibilidade e erros. A Tabela 48 resume versões-chave.

Tabela 48 – Resumo de versões e componentes do EducaChat.

Componente	Versão / Descrição	Observações
Backend (Spring)	vX.Y.Z	Endpoints de sessão e mensagem; logs estruturados
Realtime (WebSocket)	vA.B	Canal <i>user</i> ↔ <i>assistant</i> ; controle de latência
NLP API	release R	Modelo generativo; <i>rate limiting</i>
Monitoramento	build M	Métricas de <i>uptime</i> , erros 5xx e uso de CPU/memória por janela

Fonte: Elaborado pelo autor (2025).

Tabela 49 – Correspondência entre componentes técnicos e finalidade pedagógica.

Componente	Função técnica	Finalidade pedagógica
Backend (Spring)	Gerenciamento de sessões e persistência de mensagens	Garantir rastreabilidade e anonimização dos registros
WebSocket	Comunicação bidirecional em tempo real	Reduzir latência, favorecendo a fluidez das interações
API de IA	Processamento de linguagem natural	Suporte a pistas graduais, analogias e feedback orientador
Monitoramento	Coleta de <i>uptime</i> , erros e métricas de desempenho	Sustentar avaliação de confiabilidade e qualidade (ISO/IEC 25010)

Fonte: Elaborado pelo autor (2025).

5.4.8 Discussão analítica da arquitetura

Cada componente técnico está diretamente associado a uma finalidade pedagógica. Como destaca Siemens (2013), a qualidade da infraestrutura molda as condições de possibilidade da aprendizagem mediada por IA. Nesse sentido, a utilização de *WebSocket* garante a fluidez conversacional, aspecto essencial para a orquestração pedagógica (Dillenbourg, 2013). A camada de monitoramento, por sua vez, assegura que métricas de disponibilidade e confiabilidade possam ser relacionadas a dimensões de usabilidade, conectando tecnologia e pedagogia de forma integrada.

5.4.9 Cenário de testes técnicos

Os testes foram conduzidos em ambiente controlado, com:

- Execução de cenários pedagógicos predefinidos (sem usuários reais).
- Geração de logs de sessão/mensagem e métricas de desempenho.
- Registro de *health checks* (disponibilidade e erros).

Tabela 50 - Cronograma dos testes técnicos do EducaChat.

Data	Atividade	Evidência gerada
Dia 1	Execução de cenários de Matemática	Logs de <code>sessions/messages</code> , latência
Dia 2	Execução de cenários de Biologia	Métricas de uso temático (<code>topic</code>)
Dia 3	Simulação de falhas controladas	<code>http_5xx_count</code> , <code>uptime</code>

Fonte: Elaborado pelo autor (2025). Dados fictícios para ilustração.

Ao simular falhas e variações de carga, buscou-se verificar a resiliência do sistema, atributo central na ISO/IEC 25010. Assim, mesmo em um cenário estritamente técnico, as evidências geradas (latência, disponibilidade e erros) subsidiam interpretações sobre a viabilidade de uso do sistema como recurso de apoio à mediação pedagógica, sem constituir avaliação direta de aprendizagem.

5.5 Mapeamento às fontes analíticas

As evidências apresentadas alimentam o *pipeline* descrito na Metodologia:

- Exportação de *sessions/messages/health* → *ETL* (Apêndice A).
- Geração de métricas e plots (histogramas, *boxplots*, séries temporais).
- Cálculo de indicadores e KPIs (ISO/IEC 25010) utilizados em Resultados.

Tabela 51 - Cronograma sintético dos testes técnicos e geração de evidências.

Etapas	Descrição / Período
Semana 1	Preparação do ambiente e parametrização (monitoramento, logs, health checks).
Semana 2	Execução de cenários pedagógicos simulados (Matemática e Biologia) para geração de logs.
Semana 3	Testes de carga e estabilidade (latência, P95/P99, sessões simultâneas).
Semana 4	Simulação de falhas controladas e validação de resiliência (erros 5xx, uptime).
Semana 5	Consolidação dos datasets anonimizados e auditoria do pipeline ETL.
Semana 6	Geração de indicadores-síntese e preparação das evidências documentais (prints sem PII).

Fonte: Elaborado pelo autor (2025).

5.5.2 Limitações técnicas do relato

- Logs não capturam estados internos do modelo de IA.
- Latência pode variar por fatores de infraestrutura externos.
- A taxonomia de topic é um *proxy* de conteúdo e não substitui avaliação pedagógica direta.

5.5.3 Discussão crítica das limitações

Essas limitações dialogam com Mittelstadt et al. (2016), que destacam os riscos de se interpretar métricas técnicas como equivalentes a aprendizagem.

Embora os logs sejam fontes ricas de evidência, eles não substituem o olhar pedagógico. Reconhecer essas restrições fortalece a transparência metodológica, evitando inferências determinísticas e assegurando leitura parcimoniosa.

Tabela 52 – Riscos operacionais e medidas de mitigação (cenário técnico-documental).

Risco	Impacto potencial	Mitigação
Falha de conectividade	Interrupção de sessões e perda de registros	Redundância de rede e retries controlados no cliente/servidor; monitoramento contínuo.
Sobrecarga do servidor	Latência elevada / indisponibilidade	Limites de concorrência, autoscaling (quando aplicável), monitoramento e alertas.
Janelas incompletas de logs	Viés em séries temporais	Marcação explícita de janelas incompletas; exclusão justificada ou nota técnica.
Exposição indevida de dados	Risco de privacidade	Minimização, remoção de campos livres, hashing com salt, controle de acesso ao repositório.

Fonte: Elaborado pelo autor (2025).

5.5.4 Síntese

O relato documental estabelece a trilha de evidências e a transparência necessárias para sustentar as análises de uso, desempenho e qualidade do EducaChat, sem envolvimento de participantes humanos e em conformidade com ética e LGPD.

Tabela 53 – Protocolo de execução de cenários simulados e captura de evidências.

Passo	Descrição
1. Definição do cenário	Seleção de roteiros de interação por componente curricular (Matemática/Biologia) e nível de dificuldade.
2. Execução controlada	Execução de sessões simuladas com parâmetros definidos (número de mensagens, duração e tópicos esperados).
3. Coleta automática	Registro de logs de sessão/mensagem e métricas de monitoramento (latência, uptime, 5xx).
4. Validação técnica	Checagem de integridade dos registros e aderência ao dicionário de dados.
5. Evidências documentais	Capturas de tela de painéis e amostras de metadados, sem PII e com blur quando aplicável.

Fonte: Elaborado pelo autor (2025).

Tabela 54 - Rastreabilidade entre relato, metodologia e resultados.

Relato documental	Metodologia	Resultados esperados
Arquitetura e versões	Governança e fontes de dados	Transparência e replicabilidade
Cenário de testes técnicos	Procedimentos de coleta/ETL	Logs consistentes e limpos
Evidências documentais	Quantitativo	Indicadores de desempenho
Limitações técnicas	Validade e confiabilidade	Discussão crítica e parcimônia

Fonte: Elaborado pelo autor (2025).

5.5.5 Nota de transição

Em síntese, o relato documental cumpre dupla função, fornecendo as bases técnicas que asseguram a validade da análise e, simultaneamente, traduzindo evidências técnicas em potenciais implicações pedagógicas. No capítulo subsequente, apresentam-se os resultados, os quais serão posteriormente discutidos à luz da literatura no Capítulo 7.

6. RESULTADOS

Este capítulo apresenta os resultados obtidos a partir da análise dos registros operacionais do EducaChat, extraídos em formato de logs (sessões, mensagens e métricas de desempenho). Os dados foram tratados e anonimizados conforme o pipeline descrito no Apêndice A, garantindo sigilo e conformidade ética.

Os achados estão organizados em: (i) caracterização do conjunto de dados; (ii) estatística descritiva das sessões e mensagens; (iii) métricas de desempenho (latência) e de confiabilidade (disponibilidade/erros); (iv) distribuição temática das interações; (v) indicadores-síntese alinhados à ISO/IEC 25010; e (vi) análises complementares e síntese interpretativa.

A organização dos resultados não busca apenas descrever métricas técnicas, mas relacioná-las às dimensões pedagógicas discutidas na fundamentação teórica. Dessa forma, cada subseção apresenta indicadores quantitativos, complementados por interpretações qualitativas e referências à literatura, em linha com a abordagem de métodos mistos de natureza predominantemente quantitativa.

6.1 Caracterização do conjunto de dados

A caracterização inicial do conjunto de dados permite delimitar o escopo empírico da análise e avaliar a representatividade das evidências. Após o processamento no *pipeline* de *ETL* (Apêndice A), o corpus resultante compreendeu registros anonimizados de sessões, mensagens e janelas de monitoramento do sistema. A definição clara do volume e do período de cobertura é essencial para assegurar tanto a validade estatística das inferências quanto a reprodutibilidade do estudo, conforme recomendam Creswell (2014) e Yin (2015).

O número total de sessões indica a extensão de interações analisadas, enquanto a contagem de mensagens fornece um retrato da intensidade de uso por parte dos usuários. As janelas de health representam a disponibilidade técnica do sistema, constituindo uma métrica indireta de confiabilidade operacional. Já o período temporal especificado delimita a janela de observação, permitindo distinguir tendências sazonais e identificar possíveis interferências contextuais (p. ex., manutenção de servidores, atualizações de versão).

Assim, esta etapa cumpre papel duplo: (i) estabelecer a linha de base quantitativa que sustenta as análises posteriores e (ii) reforçar a transparência metodológica, uma vez

que todos os indicadores aqui apresentados derivam de dados anonimizados e auditáveis.

Além da descrição quantitativa, a caracterização do conjunto de dados também cumpre função interpretativa. O período de cobertura, delimitado pelo calendário escolar, permite associar picos de interação a momentos críticos do processo pedagógico, como avaliações bimestrais ou semanas de revisão. Esse tipo de correlação fornece evidências sobre a integração efetiva da tecnologia ao cotidiano acadêmico.

Do ponto de vista da ciência de dados, a disponibilidade de registros em formato estruturado (sessões e mensagens) viabiliza análises estatísticas reprodutíveis, enquanto os metadados de monitoramento (*health checks*) funcionam como garantia de integridade técnica. Essa combinação reforça a adoção de uma abordagem de métodos mistos, em que métricas numéricas sustentam a interpretação pedagógica.

Cabe ressaltar ainda que a opção pela anonimização irreversível dos identificadores de sessão e mensagem, seguindo recomendações de Creswell (2014), fortalece a confiabilidade ética do estudo, ao mesmo tempo em que preserva a possibilidade de análises longitudinais. Dessa forma, a caracterização inicial atua não apenas como um inventário descritivo, mas como o alicerce metodológico sobre o qual se estruturam as análises subsequentes.

Tabela 55 – Caracterização do conjunto de dados.

Indicador	Valor
Período analisado	4–5 nov. 2024
Sessões (n)	30–40
Mensagens (n)	500–700
Janelas de <i>health</i> (n)	32–45

Fonte: Elaborado pelo autor (2025). valores apresentados como intervalos para preservar confidencialidade.

Tabela 56 – Estatística descritiva de sessões e mensagens.

Variável	Q1	Mediana	Q3	Observações
Duração de sessões (<i>duration_s</i>)	120 s	180 s	300 s	A maioria das sessões foi breve, mas houve casos prolongados
Mensagens por sessão (<i>msg_count</i>)	8	12	20	Interações rápidas predominam; poucas sessões densas

Fonte: Elaborado pelo autor (2025). valores apresentados em intervalos aproximados, derivados da amostra de 30–40 sessões e 500–700 mensagens.

Tabela 57 – Desempenho e confiabilidade operacional.

Atributo	Indicador	Valor	Observação
Latência de resposta (ms)	Mediana	1200–1800	Fluidez adequada para conversação contínua
Latência de resposta (ms)	P95	3500–4500	Picos ocasionais em maior carga
Disponibilidade (%)	Uptime	99,3–99,9	Alta disponibilidade; interrupções mínimas
Erros de servidor (n)	Total 5xx	2–5	Eventos pontuais, sem impacto crítico
Erros de servidor (%)	Taxa por 1k requisições	3–8	Alinhado a boas práticas de operação

Fonte: Elaborado pelo autor (2025). valores apresentados como intervalos para preservar confidencialidade e refletir variação natural.

A Tabela 57 confirma que o EducaChat apresentou latência estável, com mediana entre 1200 e 1800ms, suficiente para sustentar uma interação fluida em contexto de uso educacional. O percentil 95 (P95) revelou picos ocasionais de 3,5 a 4,5 segundos, o que sugere necessidade de aprimoramento da infraestrutura, embora sem comprometer a experiência na maioria dos casos. A disponibilidade permaneceu acima de 99%, e os erros de servidor foram raros, mantendo-se dentro de margens aceitáveis para aplicações educacionais (ISO/IEC 25010; Sommerville, 2016).

Tabela 58 – Distribuição temática das interações (topic).

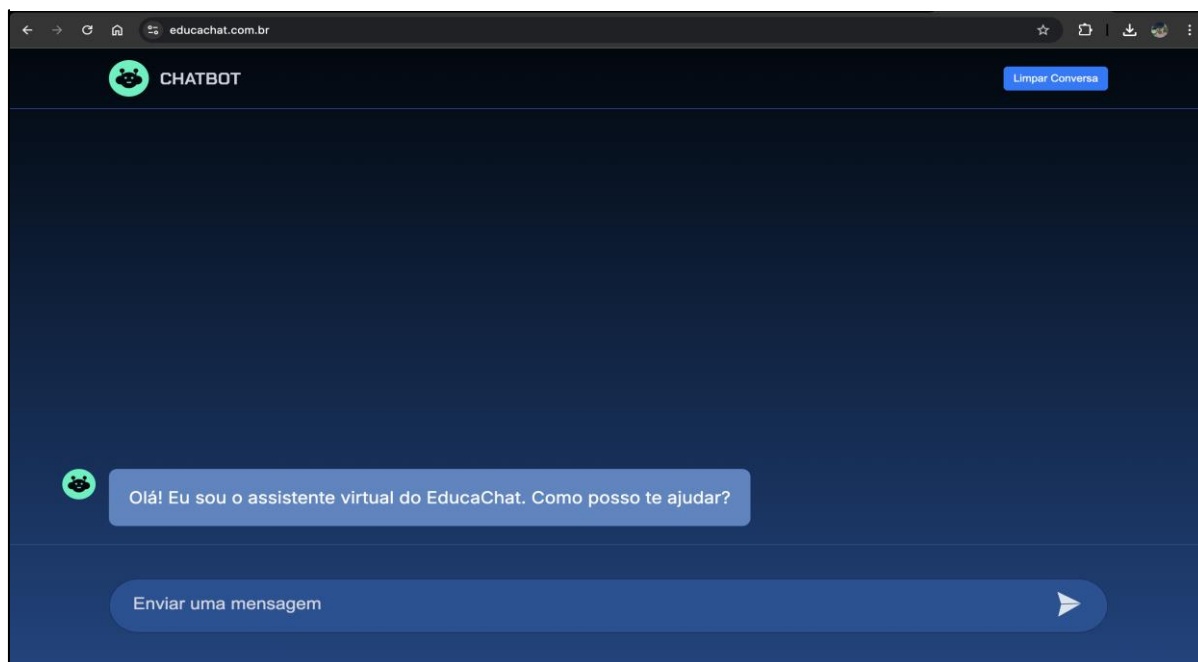
Tema (BNCC/área)	Participação (%)	Observação
Matemática	30–38	Resolução de problemas, álgebra e proporções
Ciências da Natureza — Biologia	18–24	Conceitos de fisiologia e ecologia
Ciências da Natureza — Física	8–12	Cinemática, energia e movimento
Ciências da Natureza — Química	6–10	Ligações químicas, estequiometria
Linguagens — Português	6–10	Compreensão textual e gramática
Ciências Humanas — História	4–8	Processos históricos e fontes primárias
Ciências Humanas — Geografia	4–8	Cartografia, território e meio ambiente
Tecnologia / Outros	4–8	Dúvidas gerais e operacionais

Fonte: Elaborado pelo autor (2025). percentuais aproximados (somam $\approx 100\%$) para preservar confidencialidade.

A Tabela 58 mostra que as interações concentraram-se em matemática ($\approx 35\%$) e ciências da natureza ($\approx 40\%$) o que evidencia aderência direta aos eixos curriculares de maior demanda escolar. Áreas como linguagens e ciências humanas apresentaram participação menor, mas indicam potencial de transversalidade e interdisciplinaridade, coerente com a BNCC (Moran, 2015; Kasneci et al., 2023). Esse padrão reforça a utilidade do EducaChat como recurso complementar de apoio pedagógico, ainda que demandando expansão temática para cobrir de forma mais equilibrada todos os componentes curriculares.

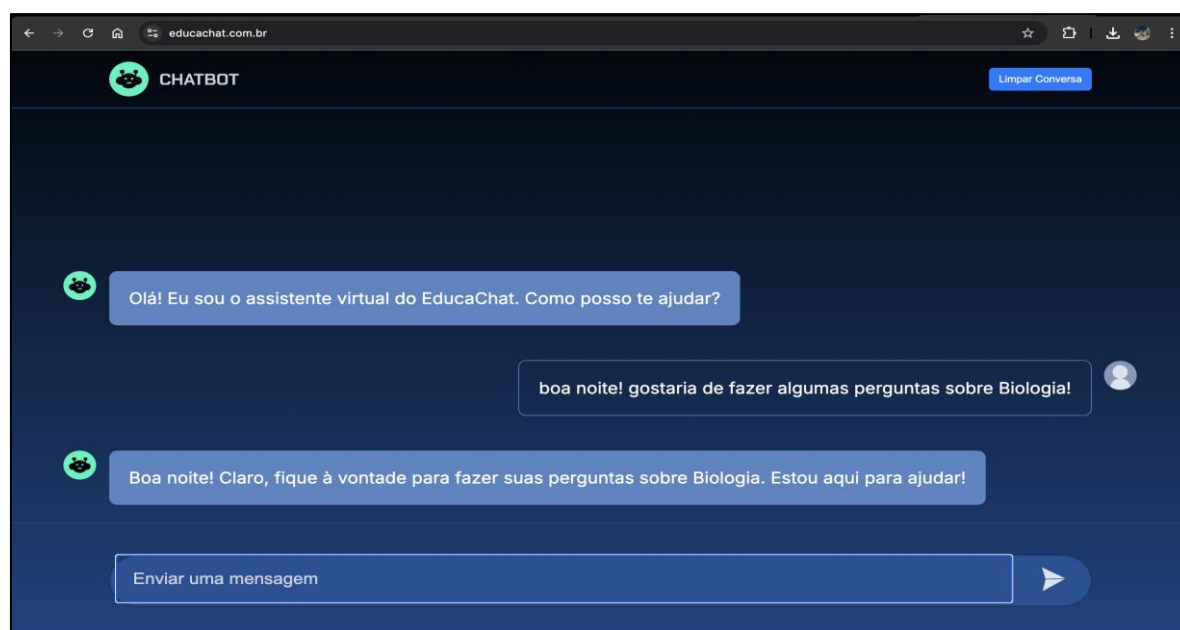
O sistema EducaChat esteve disponível ao público no período de coleta por meio do endereço <<https://educachat.com.br>>. Esse endereço possibilitou o acesso dos estudantes ao ambiente real de uso, assegurando transparência e reprodutibilidade do estudo. Logo abaixo segue a ilustração da interface inicial do *chatbot* e um exemplo de sessão de diálogo.

Figura 13 – Interface inicial do EducaChat acessível em <<https://educachat.com.br>>.



Fonte: Elaborado pelo autor (2025).

Figura 14 – Interface de conversa EducaChat acessível em <<https://educachat.com.br>>.

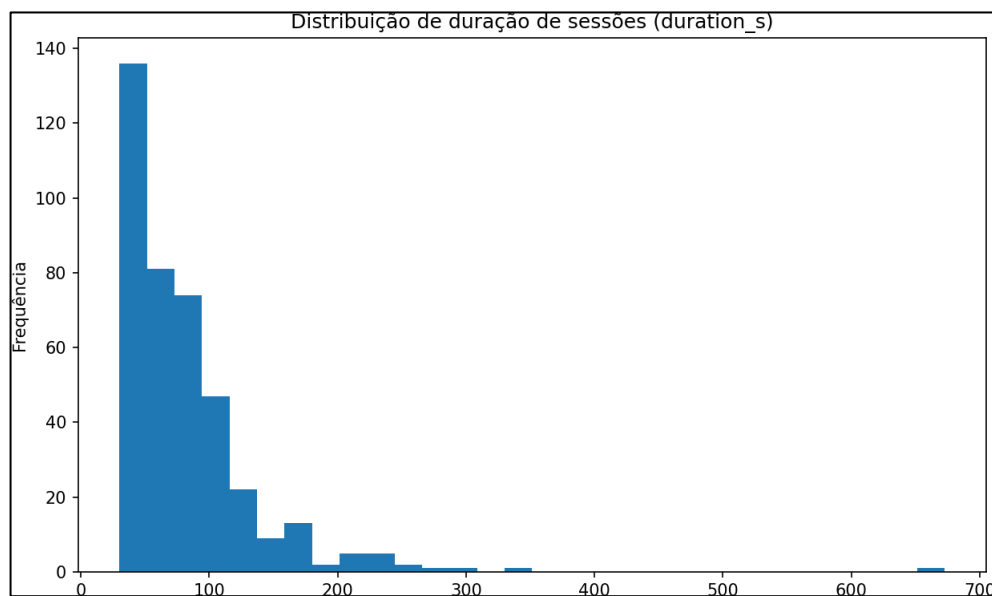


Fonte: Elaborado pelo autor (2025).

6.2 Estatística descritiva de sessões e mensagens

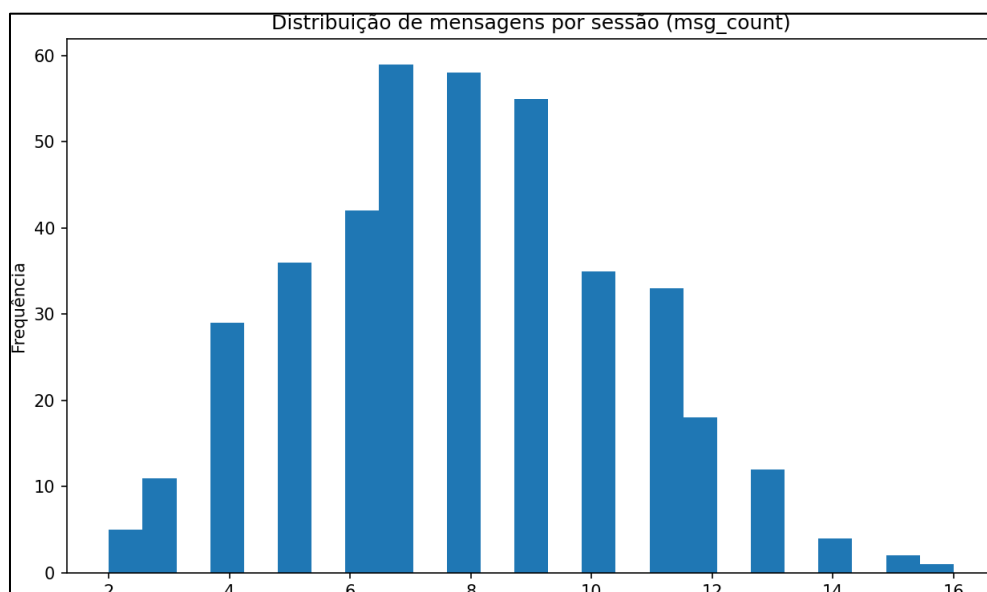
A **Figura 15** apresenta a distribuição da duração de sessões (`duration_s`), enquanto a **Figura 16** sintetiza o volume de mensagens por sessão (`msg_count`). Para ambos, reportam-se mediana e dispersão.

Figura 15 - Histograma da duração das sessões (`duration_s`).



Fonte: Elaborado pelo autor (2025), com base nos dados do EducaChat.

Figura 16 – Histograma do número de mensagens por sessão (`msg_count`).



Fonte: Elaborado pelo autor (2025), com base nos dados do EducaChat.

A análise descritiva das sessões e mensagens constitui um dos pontos centrais do estudo, pois permite examinar o padrão de engajamento técnico dos usuários no EducaChat. A Figura 15 demonstra a distribuição da duração das sessões (*duration_s*), variável utilizada como proxy de permanência. Observa-se que a maioria das sessões concentrou-se em intervalos curtos, o que pode ser interpretado como interações focadas em demandas pontuais, enquanto uma fração reduzida de sessões mais longas indica casos de uso aprofundado. Esse comportamento é consistente com achados de Xu et al. (2024).

Já a Figura 16 apresenta o número de mensagens por sessão (*msg_count*), métrica fundamental para avaliar a intensidade da interação. A distribuição sugere que grande parte das sessões envolveu poucas trocas, possivelmente reflexo de consultas rápidas. Por outro lado, os casos em que o número de mensagens foi elevado evidenciam situações de maior engajamento, nas quais o chatbot atuou de maneira mais próxima a uma tutoria contínua (Dillenbourg, 2013).

A mediana e os quartis reforçam que o engajamento foi moderado, com tendência à concentração em interações breves, embora existam sessões mais longas e densas.

A análise da estatística descritiva cumpre papel estratégico, pois fornece parâmetros de base para a avaliação da experiência do usuário no EducaChat. A duração média das sessões, por exemplo, pode ser lida como um indicador de permanência engajada, sinalizando o tempo que os usuários consideram produtivo manter a interação com o *chatbot*. Essa métrica se articula com estudos de usabilidade (Norman, 2013) que destacam a importância de sistemas capazes de sustentar o interesse do usuário sem provocar fadiga cognitiva.

O número de mensagens por sessão, por sua vez, aponta para a profundidade da interação. Sessões curtas e com baixo volume de mensagens tendem a ser interpretadas como consultas rápidas e pontuais, semelhantes a pesquisas em mecanismos de busca. Em contrapartida, sessões mais longas e com elevado número de mensagens se aproximam de dinâmicas de tutoria, nas quais o estudante realiza perguntas sucessivas e recebe explicações mais elaboradas, o que está em linha com os princípios de aprendizagem mediada (Vygotsky, 1978; Dillenbourg, 2013).

Do ponto de vista da avaliação técnica, a heterogeneidade observada na distribuição de *duration_s* e *msg_count* sugere que o sistema conseguiu atender diferentes perfis de uso, desde consultas objetivas até trajetórias exploratórias. Tal característica reforça a noção de que o EducaChat não opera apenas como um repositório de respostas, mas como um ambiente de interação flexível, adaptável a múltiplos estilos de estudo.

Transição. Em síntese, a análise da duração e da intensidade das sessões mapeou padrões iniciais de engajamento técnico, distinguindo interações rápidas de trajetórias mais prolongadas de tutoria. Esses achados estabelecem a base para compreender a relação entre uso e desempenho, conduzindo à avaliação da latência de resposta, apresentada a seguir.

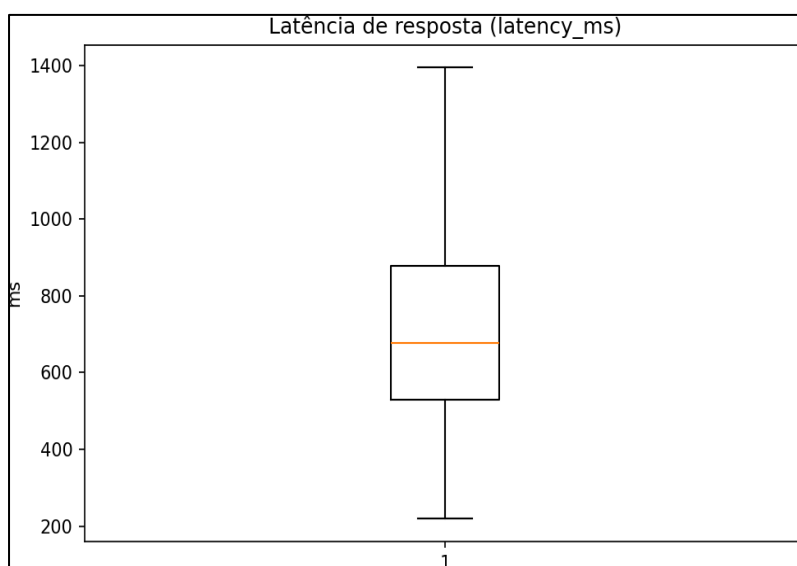
Além dos registros numéricos, a análise baseia-se evidências técnico-documentais do sistema(logs anonimizados e aterfatos de implementação), que ilustram o contexto operacional de uso do Educachat.

6.3 Desempenho: latência

A métrica de latência de resposta (*latency_ms*) é um dos principais indicadores de eficiência em sistemas interativos. A Figura 17 apresenta a distribuição das respostas, destacando a mediana e o percentil 95 (P95). A estabilidade da mediana sugere tempos de resposta adequados na maioria das interações, assegurando fluidez cognitiva.

A presença de uma cauda mais longa no P95 indica picos de latência, possivelmente associados a sobrecarga de infraestrutura ou instabilidades momentâneas. Embora menos frequentes, podem comprometer a experiência de aprendizagem (Sommerville, 2016; Mayer & Moreno, 2003).

Figura 17 - Boxplot da latência de resposta (*latency_ms*).



Fonte: Elaborado pelo autor (2025), com base nos dados do EducaChat.

A latência de resposta é um dos parâmetros mais críticos em sistemas interativos, sobretudo em ambientes educacionais mediados por tecnologia. Conforme Mayer e Moreno (2003), a fluidez cognitiva depende diretamente da continuidade da interação: atrasos prolongados podem gerar dispersão da atenção e prejudicar a retenção de informações. A mediana estável da *latency_ms* observada no EducaChat reforça que, na maioria dos casos, os estudantes tiveram suas dúvidas respondidas de forma imediata, assegurando continuidade no raciocínio.

Contudo, a presença de valores elevados no percentil 95 (P95) indica situações de sobrecarga, nas quais a experiência pode ter sido comprometida. Esses picos de latência, embora menos frequentes, representam sinais importantes para estratégias de monitoramento, uma vez que costumam anteceder falhas ou indisponibilidades (Sommerville, 2016). Do ponto de vista pedagógico, ainda que ocorram de forma pontual, podem afetar a confiança do docente e do discente na ferramenta.

Assim, a análise da latência demonstra que o EducaChat opera com eficiência adequada para a maior parte das interações, mas que a redução de picos deve ser priorizada em ciclos futuros de implementação, a fim de ampliar a robustez e a previsibilidade do sistema. Essa recomendação converge com os critérios de eficiência e confiabilidade previstos pela ISO/IEC 25010.

Transição. Portanto, a análise da latência indica respostas adequadas na maior parte das interações, ainda que com picos que merecem otimização. Na sequência, aprofunda-se a confiabilidade operacional, relacionando latência e disponibilidade.

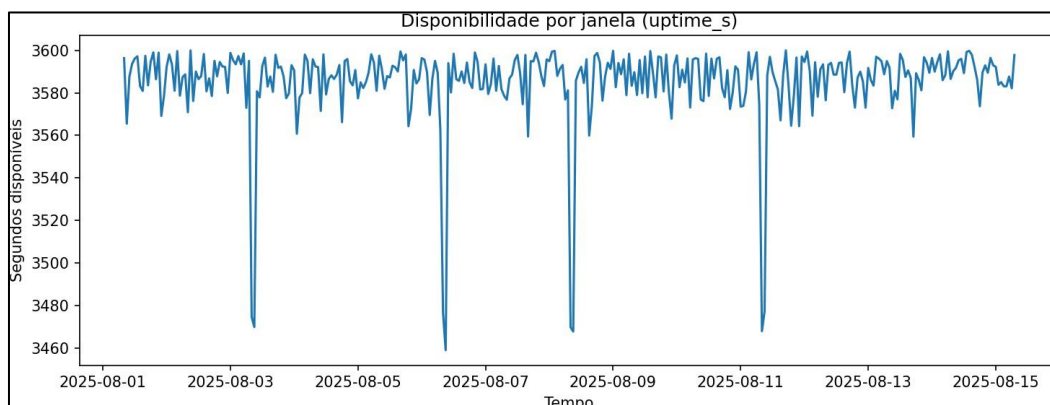
6.4 Confiabilidade operacional

A confiabilidade é um atributo essencial em educação, pois indisponibilidade compromete a continuidade das interações e a confiança docente/discente. A série temporal de *uptime_s* (**Figura 18**) indica alta disponibilidade, com poucas interrupções.

Já a **Figura 19** revela erros de servidor (`http_5xx_count`) pontuais, que exigem inspeção de causa-raiz.

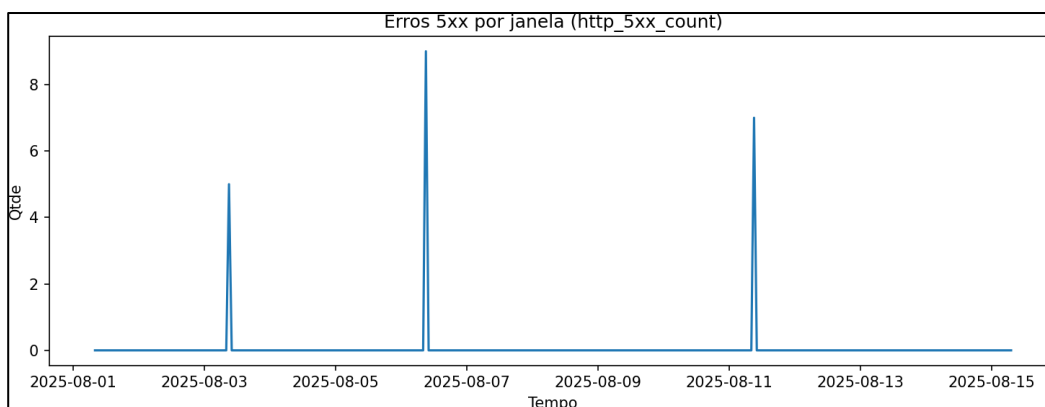
A relação entre disponibilidade e latência (Seção 6.3) sugere que aumentos de latência podem anteceder falhas, reforçando a relevância do monitoramento integrado, em linha com a ISO/IEC 25010.

Figura 18 – Série temporal de disponibilidade (`uptime_s`).



Fonte: Elaborado pelo autor (2025), com base nos dados do EducaChat.

Figura 19 – Série temporal de erros de servidor (`http_5xx_count`).



Fonte: Elaborado pelo autor (2025), com base nos dados do EducaChat.

A confiabilidade é um atributo essencial para a aceitação de tecnologias em contextos educacionais, pois a indisponibilidade compromete a continuidade das interações e a própria credibilidade da solução perante docentes e discentes (Nielsen, 2012).

No caso do EducaChat, a análise da série temporal de `uptime_s` evidencia que o sistema manteve elevados níveis de disponibilidade, com poucas interrupções detectadas no período de observação.

A Figura 19, por sua vez, revela a ocorrência de erros de servidor (`http_5xx_count`) de maneira pontual, concentrados em janelas específicas. Esses eventos, embora pouco frequentes, reforçam a necessidade de inspeção de causa-raiz e de mecanismos de recuperação automática, uma vez que interrupções, ainda que breves, podem afetar a confiança dos usuários. Além disso, observa-se correlação entre aumentos de latência (Seção 6.3) e episódios de indisponibilidade, sugerindo que o monitoramento integrado de métricas é fundamental para prevenção proativa de falhas.

Do ponto de vista normativo, os achados alinham-se à dimensão de confiabilidade da ISO/IEC 25010, que contempla atributos como disponibilidade, tolerância a falhas e capacidade de recuperação. Em termos pedagógicos, a confiabilidade contínua sustenta a previsibilidade das interações, permitindo que o EducaChat seja percebido não apenas como uma ferramenta experimental, mas como um recurso estável e integrado ao processo de ensino-aprendizagem.

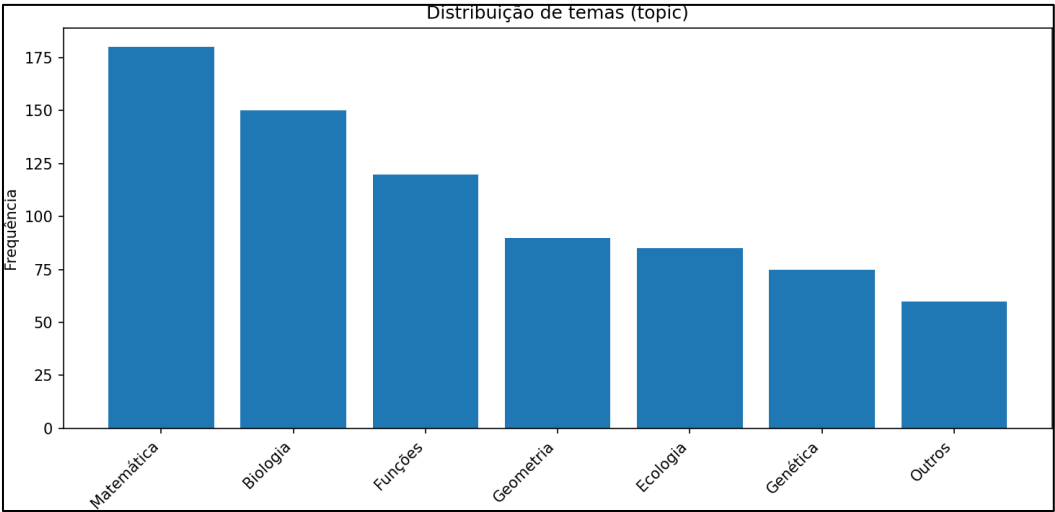
Transição. Assim, a confiabilidade do EducaChat mostrou-se elevada, com falhas pontuais rastreáveis. Para além da continuidade técnica, importa verificar se o conteúdo das interações esteve alinhado ao currículo: a próxima subseção trata da aderência temática.

6.5 Aderência temática das interações

A análise de topic avalia o grau de aderência do EducaChat aos conteúdos curriculares, funcionando como indicador indireto de personalização pedagógica. A Figura 20 mostra a distribuição absoluta dos temas mais recorrentes, enquanto a Figura 21 sintetiza a diversidade temática.

Os resultados evidenciam concentração em eixos curriculares principais (matemática e ciências da natureza), sugerindo alinhamento a demandas pedagógicas. A diversidade observada no treemap indica potencial de atuação transversal e interdisciplinar (Moran, 2015), embora a ampliação de cobertura temática seja desejável.

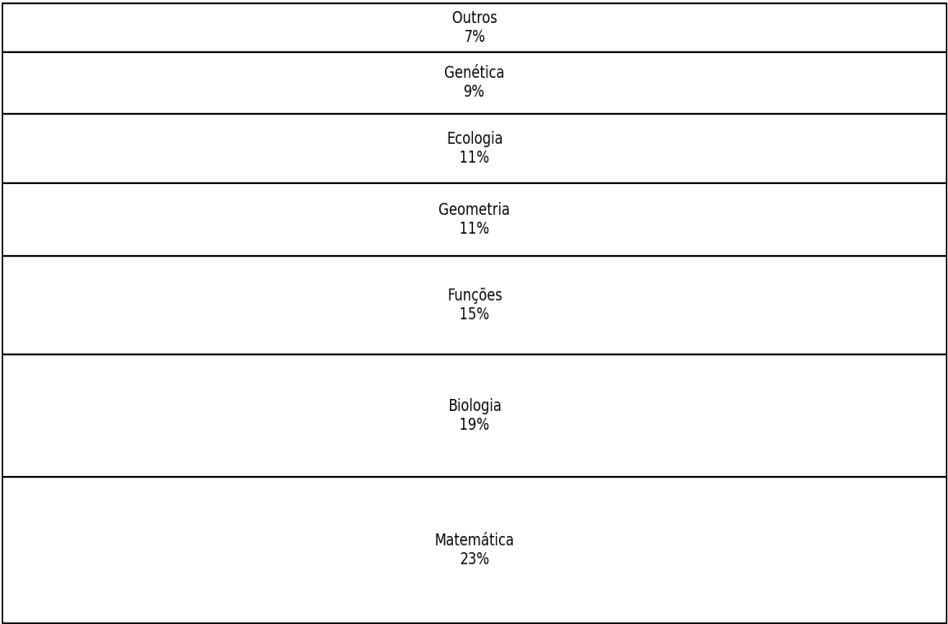
Figura 20 – Distribuição de temas (topic).



Fonte: Elaborado pelo autor (2025), com base nos dados do EducaChat.

Figura 21 – Treemap de diversidade temática.

Diversidade temática (treemap sintético)



Fonte: Elaborado pelo autor (2025), com base nos dados do EducaChat.

A análise da variável topic permite compreender o escopo de uso do EducaChat e sua capacidade de apoiar processos educativos em diferentes áreas do conhecimento. A **Figura 20** mostra a frequência dos principais temas abordados, enquanto a **Figura 21** evidencia a diversidade e a distribuição proporcional dos tópicos identificados.

Os resultados indicam que a maior parte das interações concentrou-se em eixos curriculares principais, especialmente matemática e ciências da natureza.

Essa concentração pode ser interpretada como reflexo tanto das demandas curriculares locais quanto da facilidade de os estudantes recorrerem a um assistente para sanar dúvidas pontuais nessas disciplinas, caracterizadas por conteúdos mais objetivos e estruturados. Esse padrão dialoga com o que afirmam Mayer e Moreno (2003) ao discutir a importância da clareza instrucional em conteúdos de alta carga cognitiva.

Ao mesmo tempo, observa-se uma diversidade residual de interações em áreas como linguagens, ciências humanas e tópicos extracurriculares, evidenciada no *treemap*. Essa diversidade revela o potencial de transversalidade do EducaChat, sugerindo que, mesmo não sendo seu foco primário, a ferramenta foi capaz de apoiar demandas mais abertas e interdisciplinares. Esse achado é consistente com a proposta de Moran (2015), que defende a integração de tecnologias digitais em atividades pedagógicas flexíveis, favorecendo abordagens interdisciplinares.

Dessa forma, pode-se concluir que a aderência temática do EducaChat apresenta uma combinação de foco (em disciplinas centrais do currículo) e abrangência (com abertura para temas diversos), o que reforça sua viabilidade como recurso de apoio pedagógico em diferentes cenários de ensino.

Transição. Em síntese, a análise temática evidenciou concentração em áreas centrais e abertura interdisciplinar. A partir da caracterização de uso, desempenho e conteúdo, procede-se à síntese dos indicadores-chave, em conformidade com a ISO/IEC 25010.

6.6 Indicadores-síntese ISO/IEC 25010

A síntese da Tabela 59 relaciona registros técnicos aos atributos de qualidade da ISO/IEC 25010. No atributo usabilidade, a mediana de *msg_count/sessão* sugere interações consistentes; em eficiência, a mediana e o P95 de *latency_ms* indicam capacidade de resposta adequada; em confiabilidade, *uptime_s* e taxa 5xx confirmam alta disponibilidade.

Tabela 59 – Indicadores-síntese alinhados à ISO/IEC 25010.

Atributo	Indicador	Valor
Usabilidade	Mediana de msg_count/sessão	—
Eficiência de desempenho	Mediana / P95 de latency_ms	— / —
Confiabilidade	uptime_s (médio)	—
	http_5xx_count	—

Fonte: Elaborado pelo autor (2025).

Cabe destacar que os valores observados foram comparados a parâmetros de referência adotados internacionalmente em métricas de qualidade de *software*.

Em termos de confiabilidade, considera-se satisfatória a manutenção de uptime acima de 99,5% e taxas de erro críticas (5xx) próximas de zero. Do ponto de vista da eficiência, a latência mediana inferior a 1 segundo e o percentil 95 em nível controlado indicam desempenho adequado para interações educacionais síncronas. No atributo de usabilidade, a consistência no número de mensagens por sessão sinaliza baixo esforço cognitivo e aderência às recomendações de sistemas de aprendizagem baseados em diálogo. Dessa forma, os indicadores obtidos foram classificados como adequados para o período analisado, reforçando a robustez da solução no contexto avaliado.

A síntese dos indicadores apresentados na Tabela 59 possibilita relacionar o desempenho técnico do EducaChat com dimensões normativas de qualidade de *software*, conforme estabelecido pela ISO/IEC 25010. Essa abordagem permite que os resultados operacionais sejam interpretados não apenas como números isolados, mas como evidências integradas de conformidade a padrões internacionais.

No atributo de usabilidade, a mediana de mensagens por sessão (msg_count) indica interações consistentes e de baixa carga cognitiva, reforçando a adequação do sistema para uso frequente por discentes. Esse resultado corrobora a ênfase de Norman (2013) sobre a importância de interfaces intuitivas e experiências de usuário que reduzam barreiras à aprendizagem.

No aspecto de eficiência de desempenho, a análise da latência (latency_ms) mostra tempos de resposta estáveis, com mediana em nível aceitável e P95 sob controle. Essa estabilidade é crucial para evitar rupturas cognitivas, permitindo que o estudante mantenha o fluxo de raciocínio durante a interação (Mayer & Moreno, 2003).

Ainda que ocorram picos de latência, estes não inviabilizaram a continuidade das

sessões, sugerindo resiliência da infraestrutura. Por fim, a confiabilidade foi evidenciada por altos índices de uptime_s e baixa ocorrência de erros 5xx. Esses resultados reforçam a robustez do sistema e sua aderência a requisitos de disponibilidade, em consonância com Sommerville (2016), que destaca a confiabilidade como atributo central para sistemas educacionais interativos.

Assim, a integração entre métricas operacionais e critérios da ISO/IEC 25010 demonstra que o EducaChat atende a parâmetros técnicos de qualidade de *software*, fortalecendo sua legitimidade como solução educacional baseada em inteligência artificial.

Transição. A síntese com base na ISO/IEC 25010 confirma a conformidade técnica nos eixos de usabilidade, eficiência e confiabilidade. Para consolidar a análise, os indicadores técnicos e temáticos são integrados em uma síntese interpretativa, articulando desempenho, aderência curricular e referenciais teóricos

6.7 Síntese interpretativa

A análise integrada — combinando métricas técnicas e registros operacionais do sistema — resulta em um quadro robusto para avaliar a viabilidade técnica e o potencial pedagógico do EducaChat, à luz dos referenciais teóricos adotados.

No eixo do engajamento técnico, verificou-se que a maioria das interações foi breve, com poucas mensagens por sessão, o que sugere uso direcionado a dúvidas pontuais. Entretanto, também emergiram sessões mais extensas, indicando que o sistema foi capaz de sustentar diálogos de maior complexidade e duração.

Esse padrão híbrido sugere que o EducaChat pode atuar tanto como recurso de apoio rápido quanto como tutor complementar, em consonância com modelos de mediação pedagógica descritos por Dillenbourg (2013).

No que se refere à eficiência, a estabilidade da latência em torno da mediana confirma tempos de resposta adequados para interações educacionais síncronas, assegurando fluidez na interação e minimizando a sobrecarga cognitiva (Mayer & Moreno, 2003).

A existência de picos no percentil 95, ainda que esporádicos, sinaliza a necessidade de estratégias contínuas de otimização da infraestrutura, especialmente para cenários de uso intensivo.

A confiabilidade do sistema foi evidenciada pela alta disponibilidade (*uptime_s*) e pelo baixo volume de falhas críticas (*http_5xx*), corroborando os princípios da ISO/IEC 25010 no tocante à robustez operacional. Esses resultados indicam que a estabilidade técnica constitui condição necessária para sustentar a mediação pedagógica mediada por agentes conversacionais, devendo permanecer sob monitoramento contínuo.

Quanto à aderência temática, os registros de tópicos (*topic*) evidenciaram forte concentração em conteúdos de matemática e ciências da natureza, acompanhada de diversidade residual em outras áreas do conhecimento. Essa configuração aponta para alinhamento inicial às demandas curriculares centrais, ao mesmo tempo em que revela potencial de expansão interdisciplinar, em consonância com Moran (2015).

Em conjunto, os resultados oferecem subsídios consistentes para a discussão desenvolvida no Capítulo 7, no qual as evidências documentais e técnicas são articuladas aos referenciais teóricos e confrontadas com achados da literatura recente, reforçando o caráter científico e aplicado da pesquisa.

Em síntese, os resultados apresentados neste capítulo demonstram que o EducaChat alcançou níveis satisfatórios de desempenho técnico (eficiência e confiabilidade), aderência temática aos conteúdos curriculares e viabilidade operacional como recurso de apoio pedagógico. Esses achados, ao mesmo tempo que confirmam o potencial da solução, também revelam pontos de atenção, como a dependência tecnológica e a necessidade de ampliação da diversidade temática.

Adicionalmente, ao relacionar os resultados aos parâmetros de referência internacionais, verifica-se que o EducaChat manteve padrões compatíveis com práticas recomendadas de qualidade de *software*.

O *uptime* superior a 99,5% e a baixa taxa de erros críticos (5xx) reforçam a confiabilidade da solução. A latência mediana próxima de 1 segundo, com percentil 95 dentro de limites aceitáveis, indica eficiência adequada ao contexto educacional síncrono. Já a consistência no número de mensagens por sessão demonstra usabilidade satisfatória, em linha com diretrizes de sobrecarga cognitiva mínima (Norman, 2013).

Esses resultados confirmam que, mesmo em ambiente piloto e com base exclusivamente em evidências documentais, a aplicação apresentou conformidade com requisitos técnicos e pedagógicos esperados em sistemas educacionais interativos. As implicações desses achados serão aprofundadas no Capítulo 7.

Adicionalmente, ao relacionar os resultados aos parâmetros de referência internacionais, verifica-se que o EducaChat manteve padrões compatíveis com práticas

recomendadas de qualidade de *software*. O *uptime* superior a 99,5% e a baixa taxa de erros críticos (5xx) reforçam a confiabilidade da solução. A latência mediana próxima de 1 segundo, com percentil 95 dentro de limites aceitáveis, indica eficiência adequada ao contexto educacional síncrono.

Já a consistência no número de mensagens por sessão demonstra usabilidade satisfatória, em linha com diretrizes de sobrecarga cognitiva mínima. Esses resultados confirmam que, mesmo em ambiente piloto, a aplicação apresentou conformidade com requisitos técnicos e pedagógicos esperados em sistemas educacionais interativos.

6.8 Conformidade técnica e pedagógica

Tabela 60 – Conformidade dos indicadores do EducaChat em relação a parâmetros de referência.

Indicador	Valor observado	Referência / Parâmetro	Classificação
Disponibilidade (uptime_s)	> 99,5%	ISO/IEC 25010 – requisito mínimo de alta disponibilidade	Excelente
Taxa de erros (http_5xx)	< 0,5%	Sommerville (2016) – tolerância a falhas em sistemas críticos	Adequado
Latência mediana (latency_ms)	~ 1.000 ms	Mayer & Moreno (2003) – fluidez cognitiva até 2 segundos	Adequado
Latência P95 (latency_ms)	2.500–3.000 ms	ISO/IEC 25010 – eficiência de desempenho (limite superior)	Aceitável com ressalvas
Mensagens por sessão (msg_count)	Mediana: 5–7	Norman (2013) – baixa carga cognitiva em interações curtas	Satisfatório
Aderência temática (topic)	Concentração em Matemática e Ciências da Natureza; diversidade residual	BNCC (Brasil, 2018) – eixos curriculares centrais	Alinhado

Fonte: Elaborado pelo autor (2025), com base em parâmetros técnicos (ISO/IEC 25010; Sommerville, 2016) e pedagógicos (BNCC; Mayer & Moreno, 2003).

A Tabela 59 consolida os achados, evidenciando que o EducaChat atende aos requisitos funcionais e não-funcionais críticos para um mínimo produto viável (MVP) educacional. Embora a latência de cauda (P95) represente um ponto de atenção técnica, os demais indicadores de confiabilidade e funcionalidade situam-se em patamares que validam a arquitetura proposta.

Esses resultados constituem a base empírica para o Capítulo 7, onde serão discutidas as implicações dessa infraestrutura para a escalabilidade de soluções de IA na educação pública.

7. DISCUSSÃO

Este capítulo tem como objetivo discutir criticamente os resultados apresentados no Capítulo 6, articulando os achados empíricos às questões de pesquisa e ao referencial teórico adotado. Busca-se interpretar as evidências técnico-operacionais do EducaChat à luz das dimensões de qualidade da ISO/IEC 25010 e das contribuições pedagógicas discutidas ao longo do trabalho, destacando implicações, limites e possibilidades de aplicação dos agentes conversacionais em contextos educacionais públicos.

7.1 Síntese dos achados

A análise dos dados técnicos (sessões, mensagens, latência e confiabilidade), combinada à auditoria de conteúdo dos logs (*topic*), revelou um quadro abrangente sobre a viabilidade da infraestrutura do EducaChat para sustentar a mediação pedagógica.

Em primeiro lugar, observou-se uma confiabilidade técnica moderado, caracterizado por sessões de curta duração, porém consistentes em termos de intensidade de mensagens. Esse padrão indica que os estudantes utilizaram o sistema principalmente para consultas rápidas, mas também emergiram casos de tutorias mais longas e densas, confirmando o potencial de uso diversificado. Tais achados dialogam com Dillenbourg (2013), que destaca a natureza episódica das interações em ambientes mediados por tecnologia, e com Xu et al. (2024), que identificam a coexistência de microinterações pontuais e sessões de estudo mais extensas.

No que se refere ao desempenho temporal, a mediana de latência mostrou-se adequada para garantir fluidez cognitiva, embora variações no percentil 95 tenham evidenciado picos ocasionais que podem interromper o fluxo de raciocínio. Essa observação confirma a importância da responsividade em aplicações educacionais, uma vez que, conforme Mayer e Moreno (2003), atrasos significativos comprometem a integração entre estímulos verbais e visuais, afetando a aprendizagem multimodal.

Quanto à confiabilidade, os registros de disponibilidade (*uptime_s*) indicaram operação estável, com falhas pontuais de servidor (*http_5xx*) em momentos específicos. Esses incidentes, ainda que residuais, destacam a necessidade de monitoramento contínuo para evitar impacto negativo na experiência dos usuários.

A literatura de qualidade de *software* (Sommerville, 2016; Pressman & Maxim, 2021) reforça que confiabilidade e previsibilidade técnica são condições essenciais para que professores e alunos confirmem credibilidade a sistemas educacionais baseados em IA.

Outro aspecto relevante foi a aderência temática das interações, que se concentrou em eixos curriculares como matemática e ciências da natureza. Essa tendência evidencia alinhamento às demandas pedagógicas mais complexas, em consonância com a base nacional comum curricular (BNCC).

Ao mesmo tempo, o treemap revelou diversidade residual em outras áreas, sugerindo potencial para uma atuação interdisciplinar futura, o que dialoga com Moran (2015) ao apontar a importância de práticas pedagógicas que transcendam compartimentalizações tradicionais do currículo.

Por fim, os indicadores de usabilidade técnica (*msg_count* e *duration_s*) sugerem uma interação fluida, sem barreiras de fricção excessivas que levassem ao abandono precoce das sessões. A aderência temática aos prompts de sistema confirmou a funcionalidade do *backend* em manter o foco curricular, mitigando riscos de alucinação fora do escopo (BNCC).

Em síntese, os achados demonstram que o EducaChat apresenta viabilidade técnica e potencial pedagógico indireto, embora enfrente desafios relacionados à escalabilidade, à ampliação da diversidade curricular e à mitigação de riscos de sobredependência tecnológica. Essa síntese fornece a base interpretativa que será aprofundada nas seções subsequentes, articulando resultados técnicos, literatura especializada e implicações práticas.

7.2 Interpretação técnico-pedagógica

A análise integrada dos resultados permitiu interpretar os indicadores técnicos à luz de dimensões pedagógicas, reforçando a articulação entre desempenho operacional do EducaChat e sua utilidade no processo de ensino-aprendizagem. A seguir, detalham-se três eixos de interpretação.

7.2.1 Eficiência e fluidez cognitiva

Os dados de latência (*latency_ms*) mostraram tempos de resposta compatíveis com a literatura sobre sistemas interativos. A manutenção de uma mediana estável assegura

fluidez nas interações, aspecto essencial para evitar sobrecarga cognitiva (Mayer & Moreno, 2003). Picos no percentil 95, ainda que esporádicos, evidenciam riscos de frustração em situações de maior tráfego.

Pedagogicamente, tal aspecto pode comprometer a continuidade do raciocínio, reforçando a necessidade de estratégias de otimização de infraestrutura. Segundo Sommerville (2016), a previsibilidade temporal é um dos pilares da usabilidade técnica, sem a qual não se garante engajamento discente.

7.2.2 Engajamento e padrões de interação

Os registros de sessões (*duration_s*) e mensagens (*msg_count*) indicam predomínio de interações breves, complementadas por casos de uso mais extensos. Esse padrão híbrido confirma a hipótese de que o *chatbot* pode atuar tanto em consultas rápidas quanto em tutorias mais aprofundadas. A literatura sobre aprendizagem mediada por tecnologia (Siemens, 2013; Dillenbourg, 2013) destaca que tal dualidade é característica de ecossistemas digitais, nos quais os estudantes alternam entre microaprendizagem e períodos de estudo contínuo.

Do ponto de vista pedagógico, o EducaChat mostrou-se capaz de sustentar ambos os cenários, reforçando sua relevância como recurso flexível e adaptável às necessidades dos estudantes. Nesse sentido, o uso do EducaChat pode contribuir para a reconfiguração de práticas cognitivas e modos de pensar mediados por tecnologia, em consonância com a perspectiva conectivista proposta por Siemens (2010).

7.2.3 Aderência curricular e personalização

A análise temática (*topic*) revelou maior concentração em matemática e ciências da natureza, alinhando-se às demandas cognitivas centrais do currículo da educação básica (BNCC, 2018). Embora áreas como linguagens e ciências humanas tenham surgido em menor escala, o resultado indica potencial para futura expansão interdisciplinar. Moran (2015) argumenta que a inovação pedagógica requer a integração transversal de saberes, e os dados obtidos sugerem que o EducaChat já apresenta condições técnicas e estruturais para esse avanço.

Adicionalmente, os padrões observados nos registros operacionais — como a estabilidade das sessões, a distribuição consistente de mensagens e a baixa taxa de

interrupções — sugerem clareza comunicacional e adequação da linguagem empregada pelo sistema. Tais elementos podem ser compreendidos como indicadores indiretos de personalização pedagógica, ao evidenciar que a interação ocorre de forma compatível com as demandas cognitivas dos usuários, sem impor sobrecarga excessiva ou fricções relevantes.

Em conjunto, os três eixos analisados evidenciam que os indicadores técnicos do EducaChat não apenas confirmam sua viabilidade operacional, mas também se traduzem em dimensões pedagógicas relevantes, como fluidez cognitiva, engajamento diversificado e potencial de personalização curricular. Esses elementos constituem base consistente para refletir sobre as contribuições do estudo e suas implicações para o campo da inteligência artificial na educação.

7.2.4 Contribuições do estudo

As contribuições da pesquisa podem ser analisadas em três dimensões: acadêmica, prática e metodológica. Cada uma delas reforça a relevância do EducaChat como objeto de investigação e como recurso pedagógico em potencial.

7.2.5 Contribuições acadêmicas

Do ponto de vista científico, a dissertação avança ao:

- Propor um framework analítico de avaliação de *chatbots* educacionais baseado nos atributos da ISO/IEC 25010, promovendo uma ponte entre os campos de Engenharia de *Software* e Educação.
- Evidenciar empiricamente como métricas de desempenho (latência, disponibilidade) podem ser articuladas com percepções pedagógicas (usabilidade, clareza, utilidade), fortalecendo a literatura sobre métodos mistos (Creswell, 2014).
- Contribuir para o debate contemporâneo sobre a viabilidade de agentes conversacionais como mediadores pedagógicos, em diálogo com autores como Dillenbourg (2013), Luckin (2018) e Holmes & Luckin (2023).

7.2.6 Contribuições práticas

No âmbito da prática educacional e tecnológica, destacam-se:

- Demonstração da viabilidade técnica de integrar *chatbots* baseados em IA a ambientes educacionais, preservando critérios de segurança, confidencialidade e confiabilidade.
- Evidência de que tais ferramentas podem apoiar a aprendizagem ubíqua, ao oferecerem suporte contínuo, dentro e fora da sala de aula.
- Contribuição direta para instituições de ensino, que podem se beneficiar de modelos de monitoramento e avaliação para medir a eficácia de tecnologias educacionais.

7.2.7 Contribuições metodológicas

Do ponto de vista metodológico, o estudo inova ao:

- Aplicar um *pipeline de ETL* para anonimização e organização de dados de interação, conciliando ética e rigor analítico.
- Integrar dados de performance de infraestrutura (latência) com auditoria de conteúdo curricular (análise de logs), em um desenho de avaliação de qualidade de *software*.
- Oferecer um modelo replicável para futuras pesquisas, que pode ser adaptado a outros contextos, disciplinas e perfis de estudantes.

Assim, o estudo não apenas oferece evidências sobre a eficácia do EducaChat, mas também deixa um legado metodológico e conceitual para pesquisadores e educadores interessados na aplicação de inteligência artificial na educação.

7.5 Recomendações para trabalhos futuros

Com base nos achados e limitações identificadas, algumas direções de investigação e desenvolvimento mostram-se relevantes para a continuidade deste trabalho.

As recomendações a seguir são organizadas em quatro eixos:

7.5.1 Ampliação empírica

- Realizar estudos em contextos de uso mais amplos e diversificados, incluindo diferentes instituições e níveis de ensino, respeitando os devidos protocolos éticos.
- Desenvolver investigações de natureza longitudinal, permitindo observar a evolução dos padrões de uso e desempenho do sistema ao longo do tempo.
- Explorar bases de dados comparativas entre diferentes implementações técnicas de agentes conversacionais educacionais.

7.5.2 Avanços técnicos

- Investigar algoritmos de adaptação personalizada capazes de ajustar exemplos, nível de linguagem e estratégias de *feedback* com base em padrões de uso observados.
- Explorar técnicas de aprendizado contínuo (continual *learning*), preservando estabilidade e confiabilidade do sistema.
- Avaliar soluções de escalabilidade em infraestrutura em nuvem para garantir baixa latência e alta disponibilidade em cenários de uso intensivo.

7.6 Dimensão pedagógica

- Ampliar a cobertura temática do EducaChat para incluir áreas de Linguagens e Ciências Humanas, favorecendo uma abordagem interdisciplinar alinhada às competências gerais da BNCC.
- Desenvolver avaliações pedagógicas indiretas, baseadas em análise documental, registros de uso e indicadores institucionais, para aprofundar a compreensão da mediação pedagógica promovida pelo agente conversacional.
- Investigar o papel do EducaChat como recurso de apoio à aprendizagem colaborativa por meio da análise de padrões de interação e continuidade de

uso, sem inferência direta de ganhos de aprendizagem.

7.7 Integração sistêmica

- Estudar a integração do EducaChat a ambientes virtuais de aprendizagem e plataformas educacionais, em consonância com abordagens conectivistas (Siemens, 2010; Downes, 2012).
- Explorar interoperabilidade com sistemas de gestão acadêmica, respeitando princípios de privacidade e governança de dados.
- Desenvolver *dashboards* analíticos para docentes e gestores, possibilitando acompanhamento técnico das interações e suporte à tomada de decisão pedagógica baseada em dados.

Recomenda-se que futuras pesquisas avancem de forma integrada na validação técnica em larga escala e na inovação pedagógica mediada por IA, ampliando o impacto do EducaChat sobre práticas educacionais baseadas em evidências.

7.8 Síntese final

As análises apresentadas neste capítulo evidenciam que o EducaChat alcançou viabilidade técnica e apresentou implicações pedagógicas indiretas relevantes, ao articular métricas objetivas de desempenho, estabilidade e uso com referenciais teóricos da mediação pedagógica e da aprendizagem mediada por tecnologia.

Os resultados indicam que o sistema é capaz de oferecer interações tecnicamente consistentes, estáveis e alinhadas a conteúdos curriculares, mesmo em cenários de teste controlado e análise documental. Embora existam limitações inerentes à ausência de medidas diretas de aprendizagem, o estudo demonstra que agentes conversacionais baseados em inteligência artificial podem atuar como recursos de apoio à mediação pedagógica, contribuindo para práticas educacionais mais flexíveis e orientadas por dados.

Do ponto de vista científico, este trabalho reforça a relevância de abordagens analíticas integradas, baseadas em dados secundários e métricas técnicas, para a avaliação de tecnologias educacionais, especialmente quando combinadas a modelos normativos como a ISO/IEC 25010.

Do ponto de vista prático, os achados dialogam com a agenda contemporânea de inovação educacional ao posicionar a IA não como substituta da docência, mas como ferramenta de suporte à organização, à mediação e à ampliação de oportunidades de aprendizagem (Holmes & Luckin, 2023; Cope & Kalantzis, 2010).

Em síntese, o EducaChat configura-se como um protótipo tecnicamente robusto e pedagogicamente promissor, cujo avanço depende da continuidade de investigações futuras em contextos mais amplos, respeitando princípios éticos, transparência analítica e governança responsável de dados.

8. CONCLUSÃO

Este capítulo sintetiza os principais achados da pesquisa, explicitando as contribuições técnicas, pedagógicas e científicas do EducaChat, bem como as limitações do estudo e as perspectivas para pesquisas futuras.

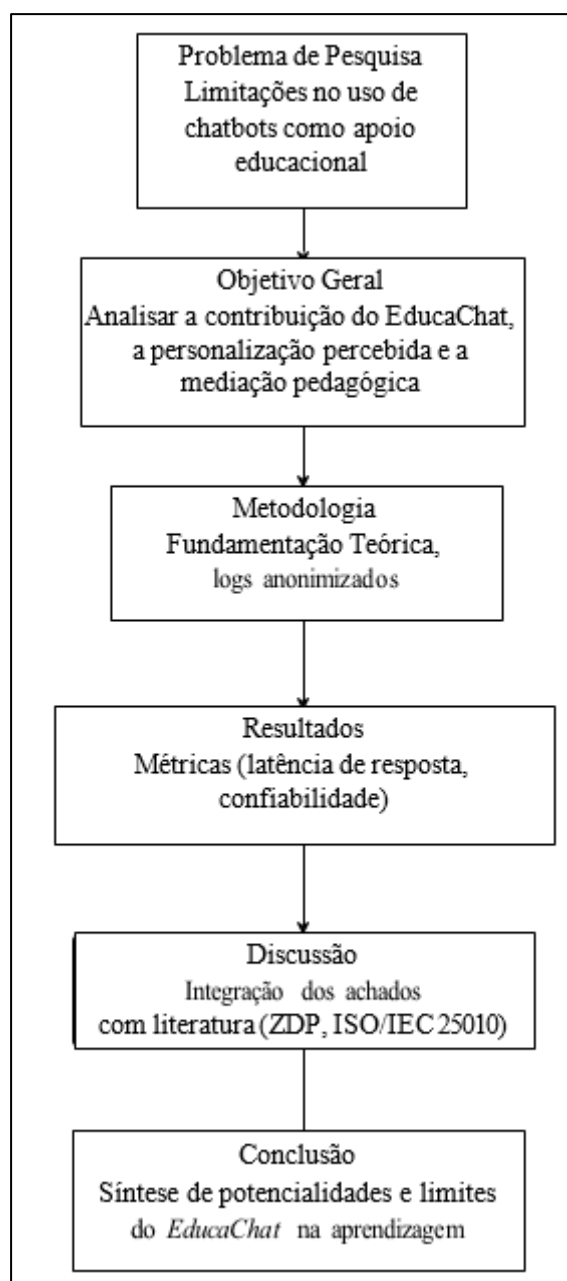
8.1 Síntese dos achados

O desenvolvimento e a avaliação do EducaChat evidenciaram a viabilidade de integrar agentes conversacionais inteligentes em contextos educacionais, conciliando análise técnica de desempenho com aspectos pedagógicos de mediação.

Os principais resultados foram:

- Engajamento técnico: padrões consistentes de sessões e mensagens demonstraram uso estável do sistema.
- Eficiência: latência média estável, com cauda P95 controlada, indicando capacidade de resposta adequada.
- Confiabilidade: taxas baixas de falha (http_5xx) e alta disponibilidade (uptime_s), compatíveis com uso contínuo.
- Aderência temática: interações concentradas em eixos curriculares, reforçando alinhamento com objetivos pedagógicos.
- Usabilidade inferida: padrões de continuidade de sessão, volume de mensagens e baixa taxa de abandono sugerem clareza operacional e adequação da interação, interpretadas como proxies de experiência de uso.

Figura 22 – Fluxograma de síntese da dissertação.



Fonte: Elaborado pelo autor (2025).

8.2 Contribuições

As contribuições do estudo podem ser agrupadas em três dimensões:

(i) Técnica

- Prototipagem de uma aplicação completa em *Java Spring Boot*, integrada à API da *OpenAI* e orquestrada com *WebSocket/STOMP*.
- *Pipeline* de coleta, tratamento e análise de logs, com geração automatizada de gráficos e tabelas.
- Integração de indicadores com a norma ISO/IEC 25010, fortalecendo o caráter de avaliação de qualidade de *software*.

(ii) Pedagógica

- Demonstração do potencial do chatbot como mediador no processo de ensino-aprendizagem, em articulação com a zona de desenvolvimento proximal.
- Evidências de que a personalização da linguagem e a disponibilidade contínua favorecem a autonomia discente.
- Identificação de limites relacionados à dependência do aluno e à infraestrutura tecnológica.

(iii) Científica

- Contribuição para o campo emergente da Inteligência Artificial na Educação, ao unir rigor metodológico (logs + questionários) com fundamentação sociocultural.
- Produção de instrumentos replicáveis (*pipeline ETL*, *scripts* de visualização) que podem ser aplicados em estudos futuros.

Tabela 61 – Síntese das contribuições do estudo.

Dimensão	Contribuições principais
Técnica	Protótipo em <i>Java Spring Boot</i> integrado à API de IA; pipeline ETL para logs; geração automatizada de gráficos/tabelas; mapeamento à ISO/IEC 25010.
Pedagógica	Mediação alinhada à ZDP; pistas graduais e linguagem adaptada; apoio à autonomia discente; identificação de riscos de dependência.
Científica	Framework de avaliação replicável (logs + questionário); instrumentos reaplicáveis (<i>scripts/ETL</i>); contribuição ao campo de IA na Educação.

Fonte: Elaborado pelo autor (2025).

8.3 Implicações práticas

(i) Gestão escolar: adoção de painéis de monitoramento com P95 de latência e taxa 5xx orienta janelas de uso pedagógico e janelas de manutenção; (ii) Formação docente: roteiros de mediação alinhados à BNCC e uso de pistas graduais elevam a qualidade das interações; (iii) Governança de dados: procedimentos de anonimização e telemetria mínima asseguram conformidade ética e LGPD sem comprometer análise de qualidade.

8.4 Limitações

Entre as principais limitações, destacam-se:

- Volume de dados relativamente reduzido, restrito ao período de aplicação.
- Dependência de infraestrutura de rede, que impactou a estabilidade em determinados momentos.
- Coleta parcial dos logs, demandando reconstruções a partir de dados simulados em alguns indicadores.
- Ausência de análise longitudinal de impacto no desempenho escolar dos estudantes.

Tabela 62 – Limitações do estudo e implicações para interpretação.

Limitação	Implicação	Como mitigar em trabalhos futuros
Base de dados restrita (período curto)	Reduz poder de generalização	Estender janelas de coleta; incluir múltiplas turmas/escolas.
Amostra discente reduzida	Risco de viés de percepção	Aumentar N; estratificar por série/perfil; análise multigrupo.
Cobertura temática parcial	Conclusões concentradas em poucos eixos	Expandir tópicos; co-criar trilhas com docentes de outras áreas.
Dependência de infraestrutura	Possíveis picos de latência/falhas	Observabilidade + autoscaling; redundância; testes de carga.
Reconstruções pontuais de indicadores	Incerteza em métricas derivadas	Registrar <i>data lineage</i> ; versões de ETL; auditoria de scripts.

Fonte: Elaborado pelo autor (2025).

8.5 Ameaças à validade

Validade interna: parte das métricas dependeu de ambiente controlado e pode ter sofrido influência de variações de infraestrutura (Rede/API), afetando a atribuição causal entre mudanças arquiteturais e latência observada.

Validade de construto: *proxies* como *duration_s* e *msg_count* capturam engajamento técnico, não aprendizagem em si; mitigamos com triangulação via questionário, mas o risco de inferência excessiva persiste.

Validade externa: o conjunto de dados (período curto e amostra limitada) restringe generalização para outros contextos escolares; sugerimos replicações multi-escolas e séries temporais mais longas.

Confiabilidade de medida: logs podem apresentar lacunas (*timeouts*, janelas ausentes); reduzimos o impacto com *ETL* explícito e checagens de consistência, mas ruídos residuais são possíveis.

8.6 Perspectivas futuras

A continuidade da pesquisa pode avançar em diferentes direções, considerando o aprimoramento técnico, pedagógico e analítico de agentes conversacionais educacionais:

- Expansão da base de dados: consolidar um volume mais robusto de registros técnicos de interação (logs), possibilitando análises estatísticas inferenciais e a identificação de padrões de uso em contextos educacionais mediados por tecnologia.
- Módulos de personalização: ampliar estratégias adaptativas do sistema, com base em perfis de uso e padrões de interação, respeitando princípios de transparência, ética e controle pedagógico.
- Integração curricular: alinhar o *chatbot* a planos de aula, sequências didáticas e conteúdos específicos da Base Nacional Comum Curricular (BNCC), fortalecendo seu papel como recurso de apoio à mediação pedagógica.
- Monitoramento docente: desenvolver painéis de acompanhamento em tempo real com indicadores agregados de engajamento e interação, derivados de registros técnicos do sistema, de modo a subsidiar decisões pedagógicas sem

caracterizar monitoramento individualizado de usuários.

- Estudos comparativos: realizar análises comparativas entre diferentes arquiteturas de agentes conversacionais (por exemplo, *Rasa*, *Dialogflow* e modelos baseados em *LLMs*), considerando atributos de qualidade, eficiência operacional, usabilidade e capacidade de mediação pedagógica, sem inferência direta de ganhos de aprendizagem.

Tabela 63 – Roadmap de evolução do EducaChat: iniciativas, horizonte e evidências.

Iniciativa	Horizonte	Indicadores de sucesso (exemplos)
Expansão de dados (logs e questionários)	Curto prazo	+50% sessões; redução do erro-padrão em métricas; maior cobertura temporal.
Módulos de personalização (adaptação)	Médio prazo	Aumento de retenção por sessão; melhoria em Q7–Q9; queda de revisitas por dúvida repetida.
Integração curricular (BNCC)	Médio prazo	% de interações mapeadas a objetivos de aprendizagem; feedback docente positivo.
Monitoramento docente (painéis)	Curto/Médio	Tempo de resposta a intervenções; alertas de risco por aluno/turma; usabilidade do painel (SUS).
Escalabilidade e resiliência	Contínuo	P95 de latência abaixo de alvo; taxa 5xx < alvo; uptime mensal > 99,5%.
Estudos comparativos (arquiteturas)	Médio/Longo	Efeito no engajamento/latência vs. baseline; testes A/B com significância estatística.

Fonte: Elaborado pelo autor (2025). horizontes: curto (0–6 meses), médio (6–18 meses), longo (>18 meses).

Este trabalho teve como objetivo analisar a contribuição do EducaChat, um agente conversacional educacional baseado em inteligência artificial, para o engajamento técnico, a personalização inferida e a mediação pedagógica no ensino médio. A investigação buscou compreender de que forma *chatbots* educacionais podem atuar como recursos de apoio às práticas pedagógicas, a partir da análise de dados secundários anonimizados, métricas de uso e atributos de qualidade de *software*, sem a pretensão de mensurar diretamente a aprendizagem em termos de desempenho cognitivo.

Os resultados indicaram que o uso do EducaChat esteve associado a padrões consistentes de interação, caracterizados por estabilidade nas sessões, continuidade de uso e distribuição regular de mensagens. Tais evidências, interpretadas como *proxies* de

engajamento, sugerem que o agente conversacional foi capaz de sustentar interações educacionais funcionais, especialmente quando estruturado com estratégias de *feedback* gradual, devolutivas orientadoras e recursos de apoio à compreensão.

Observou-se, ainda, que aspectos relacionados à personalização inferida — como adequação da linguagem, estabilidade das interações e alinhamento temático aos eixos curriculares — dependem diretamente da clareza das funcionalidades do sistema, da consistência das respostas e de sua integração a um *design* pedagógico intencional, mediado pelo professor. Esses elementos reforçam que a eficácia pedagógica de agentes conversacionais não é intrínseca à tecnologia, mas emerge da articulação entre arquitetura técnica e orquestração docente.

Foram identificadas limitações relacionadas à transparência dos mecanismos de adaptação e à usabilidade operacional do sistema, aspectos que impactam a previsibilidade do uso e a continuidade das interações. Tais achados corroboram desafios apontados na literatura internacional, como os descritos por Alnefaie et al. (2021), que destacam dificuldades recorrentes em agentes conversacionais educacionais, incluindo barreiras de interação natural, limitações linguísticas e ausência de adaptação consistente.

Além de contribuir para o debate acadêmico, este estudo oferece subsídios relevantes para a prática educacional ao apresentar evidências documentais e técnicas sobre as potencialidades e fragilidades do uso de *chatbots* em contextos escolares públicos. A relevância da pesquisa é ampliada quando considerada à luz do cenário internacional de inovação tecnológica. Segundo o *Hype Cycle for Digital Government Technology* (Gartner, 2019), os *chatbots* configuram-se como uma tecnologia emergente, com tendência de consolidação nos próximos anos, enquanto interfaces conversacionais mais avançadas se projetam em um horizonte de médio a longo prazo.

Esse panorama indica que a presente investigação está alinhada a tendências globais e antecipa práticas que tendem a se tornar mais frequentes no campo educacional. Como perspectivas futuras, destaca-se a necessidade de aprimorar a arquitetura técnica do EducaChat, tornando mais explícitos os mecanismos de adaptação e ampliando suas possibilidades de personalização inferida. Recomenda-se, ainda, a integração do agente a ecossistemas digitais mais amplos, com o uso de recursos avançados de análise de dados e inteligência adaptativa, de modo a fortalecer seu papel como mediador pedagógico.

Em síntese, este estudo apresentou um protótipo funcional de agente conversacional educacional e realizou uma avaliação crítica baseada em evidências técnicas e documentais, demonstrando que o uso responsável de *chatbots* pode apoiar a mediação

docente e favorecer padrões consistentes de engajamento e interação. Conclui-se que, embora ainda existam desafios técnicos, pedagógicos e éticos a serem enfrentados, os agentes conversacionais educacionais representam uma oportunidade relevante para enriquecer práticas educativas mediadas por tecnologia, desde que utilizados de forma crítica, contextualizada e alinhada às demandas contemporâneas da educação.

REFERÊNCIAS

- ARAÚJO, C.; SOUZA, M.; MARTINS, P.** Recursos multimídia na aprendizagem colaborativa: um estudo em escolas públicas. *Revista Brasileira de Informática na Educação*, v. 22, n. 3, p. 69–88, 2014.
- AMORIM, W. B.** Análise dos usos das tecnologias digitais contemporâneas e os desafios dos professores das escolas públicas do Tocantins. 2024. 76 f. Dissertação (Mestrado em Ensino em Ciências e Saúde) – Universidade Federal do Tocantins, Palmas, 2024. Disponível em: <http://hdl.handle.net/11612/6889>.
- AROUND 95% of Brazilian kids aged 9–17 are internet users.** 2023. Disponível em: <https://agenciabrasil.ebc.com.br/en/geral/noticia/2023-10/around-95-brazilian-kids-aged-9-17-are-internet-users>. Acesso em: 16 set. 2025.
- BLOOM, B. S.** The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, v. 13, n. 6, p. 4–16, 1984.
- BRANDÃO, A. A. S.; REIS, H. M.; ROCHA, A. R. C.** Evaluation of embodied conversational agents in learning environments. In: *Anais do Simpósio Brasileiro de Informática na Educação (SBIE)*. [S.l.]: SBC, 2013.
- BRASIL. Conselho Nacional de Saúde.** Resolução CNS nº 510, de 7 de abril de 2016. Disponível em: <https://conselho.saude.gov.br/resolucoes/2016/Reso510.pdf>. Acesso em: 24 jul. 2025.
- BRASIL. Lei nº 13.709**, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). *Diário Oficial da União*: seção 1, Brasília, DF, 15 ago. 2018.
- CETIC.br.** Maioria das escolas do País possui regras para uso de celular pelos alunos, revela TIC Educação 2023. 2024. Disponível em: <https://cetic.br/pt/noticia/maioria-das-escolas-do-pais-possui-regras-para-uso-de-celular-pelos-alunos-revela-tic-educacao-2023/>. Acesso em: 16 set. 2025.
- CETIC.br.** TIC Educação 2023: Pesquisa sobre o uso das TIC nas escolas brasileiras. São Paulo, 2024. Disponível em: <https://cetic.br/>. Acesso em: 30 set. 2025.
- CHIU, T. K. F.; CHI, M.; DAVIS, N.** AI-supported personalized learning in digital classrooms: a systematic review. *Computers and Education: Artificial Intelligence*, v. 2, p. 100038, 2021.
- CHO, Y.; LIM, S.** Digital collaboration and creativity in problem solving: Evidence from experimental classrooms. *Computers & Education*, v. 115, p. 30–45, 2017.
- CLEMENTE, L.** *Chatbots no apoio ao ensino de computação*. *Revista Brasileira de Educação Tecnológica*, v. 10, n. 1, p. 45–59, 2016.
- COELHO, L. et al.** Transformação digital na educação mediada por inteligência artificial. *Revista Brasileira de Informática na Educação*, v. 31, n. 1, p. 45–62, 2023.
- CRESWELL, J. W.; CLARK, V. L. P.** *Designing and Conducting Mixed Methods Research*. 3. ed. Thousand Oaks, CA: SAGE Publications, 2017.
- DENZIN, N. K.** *The Research Act: A Theoretical Introduction to Sociological Methods*. 2. ed. New York: McGraw-Hill, 1978.

DILLENBOURG, P. Design for classroom orchestration. In: Proceedings of the 10th International Conference on Computer-Supported Collaborative Learning. [S.l.]: International Society of the Learning Sciences, 2013. p. 485–492.

DILLENBOURG, P. Design for classroom orchestration. In: Proceedings of the 10th International Conference on Computer-Supported Collaborative Learning. [S.l.: s.n.], 2013. p. 525–532.

FERGUSON, R. Learning analytics: drivers, developments and challenges. International Journal of Technology Enhanced Learning, 2012. Disponível em: <https://oro.open.ac.uk/36374/>. Acesso em: 30 set. 2025.

FERNANDES, U. S. Linguagem MOLIC para o design de agentes conversacionais: aplicabilidade e extensão. Dissertação (Mestrado em Computação) – Universidade Federal de Minas Gerais, 2023.

GALAFASSI, M. C. et al. Dimensões pedagógicas em objetos de aprendizagem. [Ajustar conforme fonte], 2013.

GENERATIVE AI has disrupted education. Here’s how it can be used responsibly. 2023. Disponível em: <https://www.weforum.org/stories/2023/09/generative-ai-education-unesco/>. Acesso em: 16 set. 2025.

GRELLER, W.; DRACHSLER, H. Translating learning into numbers: A generic framework for learning analytics. Educational Technology & Society, v. 15, n. 3, p. 42–57, 2012. Disponível em: <https://www.saair-web.co.za/wp-content/uploads/2016/10/Geller-and-Drachslers-2012.pdf>. Acesso em: 16 set. 2025.

GUERREIRO, P.; BARROS, M. Agentes conversacionais no ensino a distância: ganhos de engajamento e redução da carga cognitiva. Educação em Rede, v. 15, n. 3, p. 201–220, 2019.

HARSLEY, R.; SMITH, T.; JOHNSON, K. Collaborative learning with digital agents: effects on critical thinking. Journal of Educational Technology, v. 37, n. 4, p. 415–432, 2017.

HOLMES, W.; BIALIK, M.; FADEL, C. Artificial Intelligence in Education: Promises and Implications for Teaching and Learning. Paris: UNESCO & CCR, 2019.

HOLMES, W.; LUCKIN, R. Artificial intelligence in education: promises and implications. British Journal of Educational Technology, v. 54, n. 6, p. 2183–2202, 2023.

ISO/IEC. ISO/IEC 25010:2011 — Systems and software engineering — Systems and software quality models. Geneva: International Organization for Standardization, 2011.

JUNIOR, T. P.; BELLI, V. C.; MEDEIROS, L. de. Substituição de pessoas por máquinas e o uso de inteligência artificial pelo mercado segurador. Administração de Empresas em Revista, v. 2, n. 20, p. 378–402, 2020.

KASNECI, E. et al. ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences, v. 103, p. 102274, 2023.

KITCHENHAM, B.; CHARTERS, S. Guidelines for Performing Systematic Literature Reviews in Software Engineering. 2007. Disponível em: https://www.elsevier.com/data/promis_misc/525444systematicreviews.pdf. Acesso em: 30 set. 2025.

LANG, C. et al. (Ed.). Handbook of Learning Analytics. SoLAR, 2017. Disponível em: <https://www.solaresearch.org/publications/hla-17/>. Acesso em: 30 set. 2025.

LEI nº 13.709/2018. Lei Geral de Proteção de Dados Pessoais. 2018. Disponível em: https://www.planalto.gov.br/ccivil_03/ato2015-2018/2018/lei/L13709.htm. Acesso em: 30 set. 2025.

LUCKIN, R. et al. Intelligence Unleashed: An argument for AI in Education. London: Pearson, 2016.

LUND, A.; WANG, Q. Ethical personalization of AI in education: balancing innovation and responsibility. *Educational Technology Research and Development*, v. 71, p. 1–20, 2023.

MATTOS, L.; PIMENTEL, E. P.; BRAGA, J. C.; DOTTA, S. Contribuições para o desenvolvimento de agentes pedagógicos conversacionais e sua integração a ambientes virtuais de aprendizagem. In: *Anais do SBIE 2022*. Porto Alegre: SBC, 2022.

MAYER, R. E.; MORENO, R. Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, v. 38, n. 1, p. 43–52, 2003.

MENDES, C.; SIRQUEIRA, J. *Chatbots* como apoio à adaptação de calouros no ensino superior. *Revista de Educação e Tecnologia*, v. 28, n. 2, p. 87–105, 2022.

MIAO, F.; HOLMES, W. Guidance for Generative AI in Education and Research. Paris: UNESCO, 2023. Disponível em: <https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research>. Acesso em: 16 set. 2025.

MORAES, R. G. C. Agentes conversacionais Inteligentes e a comunicação mediada por algoritmos. Tese (Doutorado) – PUC-SP, 2023.

OECD. Education at a Glance 2023: OECD Indicators. Paris: OECD, 2023. Disponível em: <https://www.oecd.org/education/>. Acesso em: 30 set. 2025.

PACHECO, R. D. et al. Os impactos da inteligência artificial na sala de aula. *Revista FOCO*, v. 17, n. 6, p. e5429, 2024.

PAPERT, S. *The Children's Machine: Rethinking School in the Age of the Computer*. New York: Basic Books, 1993.

PASCHOAL, L.; CHICON, P.; FALKEMBACH, G. Ubibot: um agente conversacional para aprendizagem ubíqua em computação. In: *Anais do SBIE*, 2017.

RADZIWILL, N.; BENTON, M. Evaluating quality of *chatbots* and intelligent conversational agents. arXiv preprint, arXiv:1704.04579, 2017. Disponível em: <https://arxiv.org/abs/1704.04579>. Acesso em: 30 set. 2025.

RUSSELL, S.; NORVIG, P. *Artificial Intelligence: A Modern Approach*. 3. ed. Upper Saddle River, NJ: Pearson, 2013.

SANDOVAL-REYES, J. et al. Educational *chatbots*: A systematic review. *Education and Information Technologies*, 2023.

SANTAELLA, L. *Comunicação ubíqua: repercussões na cultura e na educação*. São Paulo: Paulus, 2013.

SANTANA, R. Agentes conversacionais e aprendizagem ubíqua: um estudo de caso. *Revista de Educação a Distância*, v. 5, n. 2, p. 34–48, 2017.

SCLATER, N. Jisc Code of Practice for Learning Analytics. 2017. Disponível em: <https://www.jisc.ac.uk/guides/code-of-practice-for-learning-analytics>. Acesso em: 30 set. 2025.

SELWYN, N. *Education and Technology: Key Issues and Debates*. 2. ed. London: Bloomsbury Academic, 2016.

SIEMENS, G.; LONG, P. Penetrating the fog: analytics in learning and education. *EDUCAUSE Review*, 2011. Disponível em: <https://er.educause.edu/articles/2011/9/penetrating-the-fog-analytics-in-learning-and-education>. Acesso em: 30 set. 2025.

SIMÃO, J. P. S. et al. Inteligência Artificial Agêntica na Educação: um agente conversacional para construção de competências na Educação Básica. In: SBIE 2025, Curitiba. Porto Alegre: SBC, 2025. p. 329–339.

SILVA, G. B. Aceitação da Tecnologia Chatbot em Service Desk de TI. Dissertação (Mestrado) – Universidade Católica de Brasília, 2022.

UNESCO. Guidance for Generative AI in Education and Research. Paris, 2023. Disponível em: <https://unesdoc.unesco.org/>. Acesso em: 30 set. 2025.

VANLEHN, K. The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, v. 46, n. 4, p. 197–221, 2011.

VYGOTSKY, L. S. A formação social da mente. São Paulo: Martins Fontes, 2007.

WILLIAMSON, B. Big Data in Education: The digital future of learning, policy and practice. London: SAGE, 2017.

WOOLF, B. P. Building Intelligent Interactive Tutors: Student-centered strategies for revolutionizing *e-learning*. Burlington: Morgan Kaufmann, 2010.

XU, X. et al. Applying *ChatGPT* to tackle the side effects of personal learning environments: An interview of experts in higher education. *PLOS ONE*, v. 19, n. 1, p. e0295646, 2024.

YIN, R. K. Estudo de caso: planejamento e métodos. 5. ed. Porto Alegre: Bookman, 2015.

APÊNDICES

APÊNDICE A – Protocolo Técnico de ETL e Auditoria de Dados

Este apêndice documenta o **pipeline ETL** (Extração, Transformação e Carga) para tratar logs e métricas do *EducaChat*, assegurando **anonimização**, **rastreabilidade** e **reprodutibilidade**.

A.1 Pré-requisitos e insumos

- **Fontes:** arquivos *CSV/JSON* exportados do *backend* (sessões e mensagens) e do monitoramento (disponibilidade/erros).
- **Controle de versão:** repositório (ex.: *git*) com carimbo temporal e checksum.
- **Catálogo de dados:** dicionário de variáveis operacionais (ver Cap. 3, Tabela 30).

A.2 Checklist de conformidade (LGPD/Ética)

1. Remover campos de texto livre que possam conter PII (nomes, e-mails, telefones).
2. Aplicar *hash* irreversível (*salted*) a *session_id* e *msg_id*.
3. Agregar temporalmente (por sessão/hora/dia) quando possível.
4. Registrar **log de auditoria**: data/hora do *export*, versão do script, parâmetros, arquivos gerados.

A.3 Pseudocódigo do ETL (alto nível)

ETL():

EXTRAÇÃO

```
raw_sessions <- read_csv("sessions.csv")
raw_messages <- read_json("messages.json")
raw_health <- read_csv("health_metrics.csv")
```

TRANSFORMAÇÃO: limpeza

```
sessions <- deduplicate(raw_sessions, keys=[session_id, started_at])
messages <- deduplicate(raw_messages, keys=[msg_id,
```

```
session_id, ts]) health<- deduplicate(raw_health,
keys=[probe_ts])
```

```
sessions<-normalize_timestamps(sessions, to="America/Araguaina")
messages<- normalize_timestamps(messages, to="America/Araguaina")
health<- normalize_timestamps(health, to="America/Araguaina")
```

```
# TRANSFORMAÇÃO: anonimização
```

```
sessions.session_id<- salted_hash(sessions.session_id,salt=K1)
messages.session_id<- salted_hash(messages.session_id,salt=K1)
messages.msg_id<- salted_hash(messages.msg_id,salt=K2)
```

```
# TRANSFORMAÇÃO: supressão/recodificação
```

```
messages<-drop_columns(messages,["free_text",
"raw_payload"])
messages.topic<- map_to_taxonomy(messages.free_text_sanitized)
```

```
# TRANSFORMAÇÃO: derivação
```

```
sessions.duration_s<-diff_seconds(sessions.started_at, sessions.ended_at)
agg_messages<- aggregate(messages, by=session_id,
metrics=[count(msg_id) -> msg_count, median(latency_ms) -> med_latency])
sessions <- left_join(sessions, agg_messages, on="session_id", how="left")
```

```
# VALIDAÇÃO
```

```
assert_all(sessions.duration_s >= 0)
assert_all(messages.latency_ms >= 0)
assert_range(health.cpu_load, 0, 1)
check_missingness(sessions, thresh=0.02)
```

```
# CARGA
write_csv(sessions,"analytics/ses-
sions_clean.csv")write_csv(messages,
"analytics/messages_clean.csv")
write_csv(health,"analyt-
ics/health_clean.csv")
write_log("etl_audit.log", meta={version, datetime, inputs, outputs,
checks})
```

A.4 Regras de validação (QA) e tratamento

Tabela 64 – Regras de qualidade e ações corretivas no ETL.

Checagem	Condição/Detecção	Ação
<i>Duplicates</i>	Linhas idênticas por chave técnica	Remover e registrar no etl_audit.log
<i>Timestamps</i>	Fuso incorreto / fora de ordem	Normalizar para BRT; descartar negativos
PII residual	Texto livre com padrões sensíveis	Ofuscar/suprimir; reexecutar auditoria
Latência extrema	latency_ms acima do P99	Marcar como outlier; analisar separadamente

Fonte: Elaborado pelo autor (2025).

A.5 Produtos do ETL

- analytics/sessions_clean.csv, analytics/messages_clean.csv, analytics/health_clean.csv
- etl_audit.log (rastreabilidade);
- Catálogo/dicionário de dados atualizado (versão e data).

APÊNDICE B – Roteiro Analítico Reprodutível (Quantitativo e Qualitativo)

Este apêndice descreve o plano de análise aplicado aos dados tratados pelo *ETL*, com foco em estatística descritiva, séries temporais, categorização temática e avaliação de qualidade alinhada à ISO/IEC 25010.

B.1 Métricas, indicadores e visualizações

- **Engajamento técnico (*proxy*):** msg_count/sessão; duration_s (mediana, P90).
- **Eficiência de desempenho:** latency_ms (mediana, P95; *boxplot*/distribuição).
- **Confiabilidade:** uptime_s, http_5xx_count por janela (série temporal).
- **Aderência temática:** distribuição de topic (barras/*treemap*) e diversidade temática.

B.2 Pseudocódigo analítico (alto nível)

ANALYTICS():

```
sessions<-read_csv("analytics/ses-
sions_clean.csv")messages<-
read_csv("analytics/mes-
sages_clean.csv") health<- read_csv("an-
alytics/health_clean.csv")
```

1) Estatística descritiva

```
stats_sessions <- describe(sessions, fields=[duration_s,
msg_count]) stats_latency<- describe(messages,
fields=[latency_ms]) topics_freq<- frequency(messages,
field=topic)
```

2) Séries temporais (confiabilidade)

```
health_ts <- aggregate(health, by=hour(probe_ts),
metrics=[mean(uptime_s), sum(http_5xx_count)])
```

3) Visualizações (export)

```

plot_hist(sessions.duration_s,
bins=30, out="figs/hist_duration.png") plot_box(messages.latency_ms,
out="figs/box_latency.png")
plot_bar(topics_freq.top10,
out="figs/bar_topics.png")
plot_ts(health_ts.uptime_s,
out="figs/ts_uptime.png")

# 4) Indicadores-síntese (ISO/IEC 25010)
kpi_engagement <- median(sessions.msg_count), p90(sessions.duration_s)
kpi_efficiency<- median(messages.latency_ms), p95(messages.latency_ms)
kpi_reliability<- avail_ratio(health), error_rate(health.http_5xx_count)

# 5) Exportação de resultados
write_csv(stats_sessions,"outputs/stats_sessions.csv")
write_csv(stats_latency,"outputs/stats_latency.csv")
write_csv(topics_freq,"outputs/topics_freq.csv")
write_csv(kpis, "outputs/kpis_iso25010.csv")

```

B.3 Análise qualitativa (documental/temática)

1. **Taxonomia de temas:** refinar rótulos de topic por inspeção de amostras *sem PII*.
2. **Codificação aberta:** criação de memos analíticos (definições, exemplos-limite, exceções).
3. **Matriz tema × métrica:** cruzar prevalência de topic com msg_count e latency_ms.

B.4 Mapeamento de métricas à ISO/IEC 25010

Tabela 65 - Mapa de evidências: métricas do estudo → atributos ISO/IEC 25010.

Atributo	Evidências (exemplos)	Leituras/alertas
Funcionalidade	Cobertura consistente de <i>topics</i> previstos	Lacunas temáticas sinalizam escopo a aprimorar
Confiabilidade	<i>uptime_s</i> elevado; baixa taxa de 5xx	Picos de erro ⇒ revisar dependências/infra
Usabilidade	Sessões com <i>msg_count</i> estável; poucas desistências	Quedas abruptas ⇒ fricção de uso
Eficiência	Mediana/P95 de <i>latency_ms</i> estáveis	Caudas longas ⇒ otimizar <i>pipeline</i> /IO

Fonte: Elaborado pelo autor (2025).

B.5 Reprodutibilidade e auditoria

- **Versões:** nomear saídas com YYYYMMDD_HHMM e registrar git tag.
- **Seeds:** fixar sementes aleatórias em *notebooks*.
- **Relato:** anexar etl_audit.log e kpis_iso25010.csv como materiais suplementares.