



Chatbots, para transformar a experiência e interação do utilizador em cenários imersivos 3D

Telmo Fernando de Oliveira Sampaio - a48561

Dissertação apresentada à Escola Superior de Tecnologia e de Gestão de Bragança para obtenção do Grau de Mestre em Informática.

Trabalho orientado por:
Pedro Filipe Fernandes Oliveira
Paulo Matos

11/2025

Bragança
2024-2025



Chatbots, para transformar a experiência e interação do utilizador em cenários imersivos 3D

Telmo Fernando de Oliveira Sampaio - a48561

Dissertação apresentada à Escola Superior de Tecnologia e de Gestão de Bragança para
obtenção do Grau de Mestre em Informática.

Trabalho orientado por:

Pedro Filipe Fernandes Oliveira

Paulo Matos

11/2025

Bragança

2024-2025

Agradecimentos

Em primeiro lugar, gostaria de agradecer aos meus pais pelo apoio incondicional, sacrifício e ajuda constante ao longo da minha jornada académica. Sem o vosso apoio, nada disso teria sido possível. Agradeço também ao meu irmão por estar sempre presente e me ajudar ao longo da minha vida.

Agradeço aos amigos que fiz durante o mestrado, que me ajudaram e facilitaram a minha jornada académica, tornando esta experiência mais enriquecedora. Aos meus amigos, obrigado por estarem presentes, tanto no dia a dia como nos momentos mais difíceis, e por me darem força quando precisei.

Gostaria de expressar o meu profundo agradecimento aos meus orientadores, Professor Dr. Pedro Filipe Fernandes Oliveira e Professor Dr. Paulo Matos, pela disponibilidade, dedicação e apoio durante a realização deste trabalho.

Agradeço à Escola Superior de Tecnologia e Gestão do Instituto Politécnico de Bragança pela minha formação académica, e à Escola Superior de Comunicação, Administração e Turismo do Instituto Politécnico de Bragança por ter disponibilizado o equipamento essencial para a realização deste trabalho.

Por fim, gostaria de agradecer ao Centro de Investigação em Digitalização e Robótica Inteligente (CeDRI) e ao SusTEC pelo apoio financeiro, que foi essencial para o desenvolvimento dos artigos que constituem a base deste trabalho.

Resumo

A crescente adoção de ambientes tridimensionais imersivos em contextos comerciais criou uma necessidade crítica por interfaces conversacionais inteligentes que transcendam as limitações dos chatbots bidimensionais tradicionais. Esta dissertação apresenta o desenvolvimento de um sistema integrado que combina agentes conversacionais baseados em Inteligência Artificial (IA) com ambientes de realidade virtual. O sistema implementa uma arquitetura Retrieval-Augmented Generation (RAG) que garante respostas factuais precisas por meio de um pipeline automatizado para aquisição e processamento de dados da web, combinando modelos de linguagem em grande escala (GPT-4o-mini) com bancos de dados vetoriais (Pinecone), processamento de voz em tempo real (Convai) e ambientes Tridimensional (3D) desenvolvidos no Unity, otimizados para dispositivos Meta Quest.

A avaliação empírica em dois cenários comerciais distintos (BNH e Rádio Popular) demonstrou uma taxa de sucesso geral de 94,6%, com tempos médios de resposta de 2,17 segundos, validando a eficácia da arquitetura proposta. O sistema inclui uma plataforma de gestão administrativa que permite monitorização em tempo real, análise de padrões de utilização e controlo dinâmico da base de conhecimento, garantindo a conformidade com o Regulamento Geral de Proteção de Dados (GDPR) através de mecanismos de anonimização automática. Os resultados confirmam a viabilidade técnica e a aplicabilidade comercial de agentes conversacionais em ambientes imersivos, demonstrando generalização em diferentes domínios e abrindo caminho para novas formas de interação humano-computador em contextos de comércio eletrónico e atendimento ao cliente.

Palavras-chave: Chatbots, Ambientes 3D, Realidade Virtual, Inteligência Artificial

Abstract

The growing adoption of immersive three-dimensional environments in commercial contexts has created a critical need for intelligent conversational interfaces that transcend the limitations of traditional two-dimensional chatbots. This dissertation presents the development of an integrated system that combines artificial intelligence based conversational agents with virtual reality environments. The system implements a RAG architecture that ensures accurate factual responses through an automated pipeline for web data acquisition and processing, combining large-scale language models (GPT-4o-mini) with vector databases (Pinecone), real-time voice processing (Convai), and 3D environments developed in Unity, optimised for Meta Quest devices.

Empirical evaluation in two distinct commercial scenarios (BNH and Rádio Popular) demonstrated an overall success rate of 94.6%, with average response times of 2.17 seconds, validating the effectiveness of the proposed architecture. The system includes an administrative management platform that allows real-time monitoring, analysis of usage patterns, and dynamic control of the knowledge base, ensuring compliance with the GDPR through automatic anonymisation mechanisms. The results confirm the technical feasibility and commercial applicability of conversational agents in immersive environments, demonstrating generalisation across different domains and paving the way for new forms of human-computer interaction in e-commerce and customer service contexts.

Keywords: Chatbots, 3D Environments, Virtual Reality, Artificial Intelligence

Conteúdo

1	Introdução	1
1.1	Enquadramento	1
1.2	Objetivos	3
1.2.1	Objetivos gerais	4
1.2.2	Objetivos Específicos	4
1.3	Estrutura do Documento	5
2	Estado da Arte	9
2.1	Revisão sistemática da literatura	9
2.1.1	Enquadramento Metodológico	9
2.1.2	Protocolo de Pesquisa	10
2.2	Paradigmas de Interface em Ambientes Imersivos	12
2.2.1	Interfaces de Voz vs. Interfaces Gráficas	12
2.2.2	Design e representação visual do avatar	13
2.3	Adaptação e Experiência do Utilizador	13
2.3.1	Agentes Contextuais e Consciência Ambiental	13
2.3.2	Fatores de Adoção e Aceitação Tecnológica	14
2.4	Integração de IA e Processamento de Linguagem Natural	16
2.4.1	Modelos de Linguagem de Grande Escala (LLMs)	16
2.4.2	Arquiteturas Cliente-Servidor e PLN Avançado	17
2.5	Implementação Responsável e Desempenho	18

2.5.1	Métricas de desempenho e satisfação	18
2.5.2	Antropomorfismo e Confiança do Utilizador	18
2.6	Análise Comparativa de Sistemas Conversacionais em Ambientes 3D	19
2.7	Privacidade, Ética e Regulação	21
2.7.1	Regulamento Geral de Proteção de Dados (GDPR)	21
2.7.2	Desafios Éticos em Agentes Conversacionais	22
2.7.3	Transparência e Explicabilidade (XAI)	23
2.8	Lacunas de Investigação	24
2.8.1	Integração Técnica e Arquitetural	24
2.8.2	Adaptabilidade e Generalização de Sistemas	25
2.8.3	Considerações Éticas e Privacidade	26
2.9	Tecnologias e Ferramentas	26
2.9.1	Plataformas de Desenvolvimento para Ambientes 3D Imersivos	26
2.9.2	Modelos Conversacionais	27
2.9.3	Arquiteturas de recuperação e Ampliação	28
2.9.4	Processamento de Dados e Infraestrutura	29
3	Arquitetura do Sistema	31
3.1	Descrição do Problema	31
3.2	Proposta de Solução	33
3.3	Ferramentas e Tecnologias Utilizadas	34
3.3.1	Ambiente virtual 3D	34
3.3.2	Inteligência Artificial e Processamento de Linguagem	35
3.3.3	Bases de Dados e Armazenamento	35
3.3.4	Backend e APIs	35
3.3.5	Backoffice	36
3.3.6	Infraestrutura	36
3.4	Análise dos Requisitos	37
3.4.1	Requisitos Funcionais	37

3.4.2	Requisitos Não Funcionais	38
3.5	Modelação do Sistema	38
3.5.1	Diagrama de Caso de Uso	39
3.5.2	Tarefas e Atores	40
3.5.3	Modelação de Dados	44
4	Implementação	47
4.1	Aquisição e Processamento de Dados	47
4.1.1	Pipeline de Aquisição e Transformação de Conteúdo Web	47
4.2	Desenvolvimento do Backend de IA	50
4.2.1	Pipeline de Processamento e Embedding de Documentos	50
4.2.2	Sistema RAG com Contextualização e Geração de Respostas	51
4.3	Desenvolvimento da Experiência em RV e Processamento de Voz	53
4.3.1	Ambiente Virtual e Integração Multimodal	53
4.3.2	Processamento de Voz Integrado e Canal de Encaminhamento de Consultas	56
4.4	Integração Completa e Fluxo de Interação	58
4.4.1	Arquitetura Integrada de Seis Passos	58
4.4.2	Otimizações de Desempenho e Arquitetura de Comunicação	60
4.5	Sistema de Gestão e Monitorização (Backoffice)	60
4.5.1	Arquitetura e Fundamentos do Sistema de Gestão	60
4.5.2	Dashboard Principal e Visualização de Métricas	62
4.5.3	Gestão de Sessões e Monitorização de Interações	63
4.5.4	Controlo da Base de Conhecimento e Gestão Documental	64
5	Avaliação e Resultados	67
5.1	Metodologia de Avaliação	67
5.1.1	Desenho Experimental	67
5.1.2	Categorização das Interações	68
5.1.3	Métricas de Avaliação	68

5.2	Resultados - Caso BNH	69
5.2.1	Análise Quantitativa	69
5.2.2	Análise de Falhas	70
5.3	Resultados - Rádio Popular	71
5.3.1	Análise Quantitativa	71
5.4	Artigos Científicos	72
6	Conclusões	75
6.1	Trabalho Futuro	76
A	Proposta Original do Projeto	A1
B	Figuras Complementares	B1
C	Código	C1
C.1	Crawler (GDPR)	C1
C.2	JWT.py	C10
C.3	Base.html	C13
D	Artigos	D1
D.1	ICTIS	D1
D.2	ARTIIS	D3
D.3	ICEET	D5
D.4	ICETC	D7

Lista de Tabelas

2.1	Critérios para a seleção	11
2.2	Análise comparativa de agentes conversacionais em ambientes imersivos	20
4.1	Mecanismos de Conformidade GDPR Implementados	49
5.1	Resultados da interação com o chatbot no cenário BNH (N=65)	69
5.2	Resultados da interação com o chatbot no cenário Rádio Popular (N=55)	71

Lista de Figuras

1.1	Modelo de interação multimodal em VR 3D (adaptado de [8])	3
2.1	Fluxograma da Revisão Sistemática da Literatura	10
3.1	Backoffice	39
3.2	Frontend RV	40
3.3	Modelo Mongo DB	44
4.1	Arquitetura do Web Crawler - Aquisição de dados	48
4.2	Pipeline de Processamento de Documentos	50
4.3	Arquitetura FastAPI	52
4.4	Meta Quest 3	53
4.5	Pipeline de Processamento de Voz	59
4.6	Página de Documentos	66
5.1	Resultados gerais	69
5.2	Pergunta de avaliação do chatbot	69
B.1	Tabela de Getão de Documentos	B1
B.2	Página de Gestão de Utilizadores	B2

Siglas

3D Tridimensional. vi, vii, x, 1–6, 10, 12, 14–17, 21, 22, 24, 26, 31, 32, 34, 37, 38, 40, 41, 43, 47, 54, 56, 67–69, 73, 75, 76

API Application Programming Interface. 28, 29, 34, 35, 61, 63, 76

BFS Breadth-first search. 48

CORS Cross-Origin Resource Sharing. 30, 51

CPU Central Processing Unit. 72

CSS Cascading Style Sheets. 36

DSR Design Science Research. 9

FAQS Frequently Asked Questions. 31

FPS Frames Per Second. 38

GDPR Regulamento Geral de Proteção de Dados. vi, vii, 6, 21, 22, 24, 26, 28, 49

GPU Graphics Processing Unit. 72

GPUs Graphics Processing Units. 25

GUI Interface gráfica do utilizador. 12

HTML HyperText Markup Language. 29, 36, 48, 49, 61, 63

HTTP Hypertext Transfer Protocol. 29, 35, 48

IA Inteligência Artificial. vi, 2, 3, 5, 6, 16–18, 20–24, 26, 27, 29, 35, 47, 58, 62

IBAN International Bank Account Number. 49

JSON JavaScript Object Notation. 44

JWT JSON Web Tokens. 36, 42, 61

KPI Key Performance Indicator. 43

LLMs Large Language Models. 16, 17, 20, 28, 33, 64, 75

LOD Level of Detail. 56

LTS Long Term Support. 53

MR Mixed Reality. 20

NIF Número de Identificação Fiscal. 49

NLP Processamento de Linguagem Natural. 4

NOSQL Not Only Structured Query Language. 35

NPCs Non-Player Characters. 20

PBR Physically Based Rendering. 56

PICAN Platform-Independent Context-Aware Onboarding Agent. 14, 19–21

PLN Processamento de Linguagem Natural. 17, 20

RA Realidade Aumentada. 13, 20, 21, 26

RAG Retrieval-Augmented Generation. vi, vii, 6, 24, 25, 28, 29, 33, 35, 41, 51, 52, 57, 58, 64, 66, 69, 73, 75

RAM Random Access Memory. 72

REST Representational State Transfer. 29, 35, 51, 61

RPM Ready Player Me. 55, 56

RSL Revisão Sistemática da Literatura. 5, 9, 10

RV Realidade Virtual. 2, 6, 13, 20, 21, 25–27, 29, 30, 34, 37, 38, 40–43, 45, 47, 53, 54, 58, 60, 68, 69

SSL Secure Sockets Layer. 49

STT Speech-to-Text. 35

TLS Transport Layer Security. 49

TTS Text-to-Speech. 35, 56

UE União Europeia. 21, 24

URL Uniform Resource Locator. 48, 49

URP Universal Render Pipeline. 53

VUI Interface de voz do utilizador. 12, 20

XAI Inteligência Artificial Explicável. 23

XR Extended Reality. 54, 55

Capítulo 1

Introdução

Esta dissertação apresenta uma proposta de desenvolvimento de chatbots orientados para transformar a experiência e interação do utilizador em cenários imersivos 3D. O trabalho enquadra-se no âmbito do Projeto de Mestrado em Informática. No presente capítulo, é introduzido o tema e destacada a sua relevância e aplicabilidade em diferentes contextos digitais. O capítulo conclui com uma apresentação da estrutura do documento.

1.1 Enquadramento

À medida que a transformação digital se acelera, a convergência da inteligência artificial e das interfaces tridimensionais está a redefinir os limites da interação online [1], impulsionando uma mudança fundamental na forma como as organizações interagem com os seus utilizadores e consolidando a inteligência artificial como uma tecnologia central nesta revolução. No contexto empresarial, a adoção de tecnologias de ponta tornou-se um fator crítico para a competitividade, acelerando a digitalização das pequenas e médias empresas nos mercados emergentes [2]. Essa pressão competitiva levou as organizações a procurar soluções inovadoras que não apenas melhorem a eficiência operacional, mas também criem experiências digitais diferenciadas e envolventes para seus consumidores. Nesse cenário, os chatbots baseados em inteligência artificial têm demonstrado um potencial significativo

em vários setores, especialmente onde a capacidade de compreender e processar a linguagem natural se tornou um requisito essencial para interações eficazes [3]. No campo do comércio eletrônico, estudos recentes mostram que os chatbots integrados com sistemas de recomendação baseados em IA podem melhorar significativamente a experiência de seleção de produtos, permitindo uma personalização mais eficaz das sugestões apresentadas aos consumidores [4]. Essa integração entre o processamento de linguagem natural e os sistemas de recomendação representa uma convergência tecnológica com implicações práticas relevantes para o mercado digital, ilustrando como a IA conversacional pode ser aplicada para criar valor tangível nas interações comerciais digitais.

Paralelamente a essa evolução dos chatbots tradicionais, a computação espacial e os ambientes tridimensionais surgiram como uma nova fronteira para a interação humano-computador, expandindo as possibilidades de como os utilizadores podem interagir com sistemas inteligentes. A integração da inteligência artificial generativa com tecnologias de Realidade Virtual (RV) imersivo permitiu o desenvolvimento de assistentes virtuais sofisticados, capazes de operar em espaços tridimensionais, mantendo a coerência conversacional e a precisão informativa [5]. Esta evolução marca uma transição das interfaces conversacionais bidimensionais baseadas em texto para agentes conversacionais incorporados em ambientes espaciais, onde a multi-modalidade se torna um fator crítico de sucesso [6].

A eficácia destes agentes conversacionais tridimensionais depende fundamentalmente da sua capacidade de integrar vários modos de interação de forma coerente e natural. A literatura contemporânea demonstra que agentes conversacionais incorporados eficazes devem combinar perfeitamente diálogo textual, reconhecimento de expressões faciais e processamento emocional em tempo real para criar experiências envolventes e naturais [7][8].

A Figura 1.1 ilustra o ciclo de interação multimodal adaptado para ambientes de RV 3D, destacando os processos de fusão de entrada (voz, gestos, olhar) e fissão de saída (avatar, áudio, elementos 3D) que caracterizam a comunicação com chatbots incorporados. Este ciclo de percepção, processamento e resposta permite ao chatbot manter consciência

espacial e promover interações naturais e contextualizadas no ambiente tridimensional.

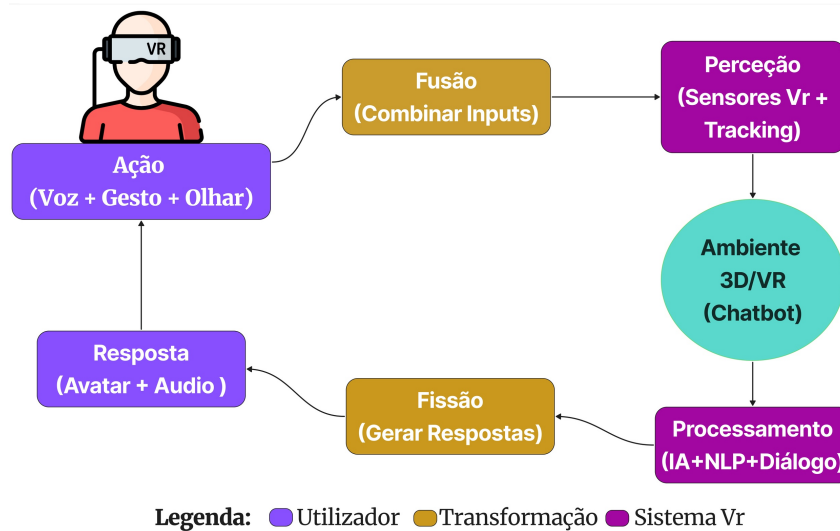


Figura 1.1: Modelo de interação multimodal em VR 3D (adaptado de [8])

No entanto, implementar grandes modelos de linguagem em chatbots baseados em avatares apresenta desafios arquitetônicos significativos que transcendem a simples integração tecnológica. Estudos recentes revelam a complexidade inerente à criação de chatbots de avatares de domínio aberto que possam manter conversas naturais e apropriadas, preservando a consistência visual e comportamental esperada de agentes incorporados em ambientes tridimensionais [9][10]. Esses desafios incluem não apenas questões técnicas de sincronização multimodal, mas também a manutenção de uma experiência de usuário fluida que aproveite as vantagens exclusivas dos espaços tridimensionais sem sacrificar a qualidade da interação conversacional.

1.2 Objetivos

Face ao enquadramento apresentado sobre a convergência entre chatbots baseados em IA e ambientes imersivos 3D, e reconhecendo o potencial transformador desta integração para a experiência do usuário, este trabalho apresenta os seguintes objetivos.

1.2.1 Objetivos gerais

Este trabalho visa desenvolver um chatbot genérico e adaptável, capaz de se integrar em cenários imersivos tridimensionais, transformando a forma como os utilizadores interagem com ambientes 3D através de interfaces conversacionais inteligentes e contextualizadas.

1.2.2 Objetivos Específicos

Desenvolvimento do Sistema de Chatbot:

- Criar um chatbot genérico capaz de se adaptar a diferentes tipos de cenários, fornecendo respostas relevantes e contextualizadas;
- Incorporar técnicas de Processamento de Linguagem Natural (NLP) e aprendizagem automática para melhorar as capacidades de compreensão e resposta do chatbot;
- Implementar mecanismos de personalização dinâmica que permitam configurar e parametrizar o comportamento do chatbot de acordo com as necessidades específicas de cada cenário.

Integração em Ambientes 3D:

- Investigar e implementar técnicas para incorporar chatbots interativos em cenários tridimensionais;
- Desenvolver soluções técnicas que permitam interações em tempo real entre o chatbot e ambientes 3D;
- Garantir a sincronização multimodal entre as respostas conversacionais e os elementos visuais tridimensionais.

Sistema de Alimentação de Dados:

- Implementar um sistema de recolha de informações baseado em web crawlers para alimentar o chatbot com dados dinâmicos;

- Desenvolver mecanismos autónomos que garantam a atualização contínua e a relevância dos dados recolhidos;

Avaliação e Monitorização:

- Testar o chatbot em diferentes tipos de cenários e contextos 3D, avaliando a sua flexibilidade e utilidade;
- Desenvolver ferramentas de monitorização para acompanhar as interações e avaliar o desempenho do sistema;
- Identificar métricas de usabilidade e eficiência relevantes para avaliar a qualidade da experiência do utilizador em ambientes imersivos.

1.3 Estrutura do Documento

Este documento está organizado em seis capítulos que refletem o caminho metodológico da pesquisa, desde o quadro teórico até a validação empírica da solução desenvolvida:

- O Capítulo 1 apresenta uma introdução ao tema, contextualizando a convergência entre chatbots baseados em inteligência artificial e ambientes tridimensionais imersivos. São definidos os objetivos gerais e específicos da pesquisa que orientaram o desenvolvimento deste trabalho.
- O Capítulo 2 apresenta o quadro teórico da investigação, começando com uma descrição da metodologia da Revisão Sistemática da Literatura (RSL) que sustentou a investigação bibliográfica. Segue-se o estado da arte, resumindo criticamente as contribuições identificadas através da RSL, organizadas em torno de domínios temáticos relevantes: paradigmas de interface em ambientes imersivos, adaptação e experiência do utilizador, integração de IA e processamento de linguagem natural, implementação e desempenho responsáveis, análise comparativa de sistemas conversacionais em ambientes 3D e questões de privacidade, ética e regulamentação.

O capítulo também identifica lacunas de investigação que justificam a abordagem proposta e apresenta as tecnologias e ferramentas relevantes para a implementação do sistema.

- O Capítulo 3 descreve a abordagem metodológica, começando com uma caracterização detalhada do problema que motivou a pesquisa. A solução proposta é apresentada com base em quatro pilares fundamentais: agente conversacional contextualizado, sistema RAG, experiência de realidade virtual imersiva e sistema de gestão e monitorização. As ferramentas e tecnologias selecionadas são descritas, organizadas por camadas funcionais, seguidas de uma análise dos requisitos funcionais e não funcionais. O capítulo termina com a modelagem do sistema por meio de diagramas de casos de uso que ilustram as interações nos subsistemas *Backoffice* e *Frontend* RV.
- O Capítulo 4 documenta a implementação do sistema 3D-Chatbot, estruturado em quatro fases principais: aquisição e processamento de dados através de rastreamento web assíncrono, em conformidade com o GDPR; desenvolvimento do *backend* de IA com um pipeline de processamento de documentos e sistema RAG orquestrado pelo FastAPI; desenvolvimento da experiência de RV com um ambiente Unity otimizado para Meta Quest 3 e integração multimodal via Convai; e integração completa com arquitetura de seis etapas e otimizações de desempenho. O sistema de gestão e monitorização baseado em modelos *Jinja2* e *Bootstrap* também é apresentado.
- O Capítulo 5 apresenta a avaliação sistemática do sistema através de testes realizados em dois estudos de caso complementares (BNH e Rádio Popular), totalizando 120 interações categorizadas em cinco tipos funcionais. São analisadas métricas de taxa de sucesso, tempo de resposta e qualidade da experiência, com uma discussão detalhada dos resultados quantitativos e análise qualitativa das falhas identificadas.
- O Capítulo 6 resume as conclusões da pesquisa, comparando os resultados obtidos com os objetivos estabelecidos, discutindo as contribuições científicas e práticas do

trabalho, identificando limitações e propondo direções para pesquisas futuras.

Capítulo 2

Estado da Arte

A integração de agentes conversacionais em ambientes tridimensionais representa um avanço significativo na interação humano-computador. Este capítulo apresenta o enquadramento teórico que sustenta esta investigação, começando com uma descrição da metodologia de RSL utilizada para identificar e analisar criticamente as contribuições científicas relevantes. Segue-se uma síntese estruturada do estado da arte, organizando as descobertas por domínios temáticos que fundamentam o desenvolvimento do sistema proposto.

2.1 Revisão sistemática da literatura

2.1.1 Enquadramento Metodológico

Para apoiar esta investigação, foi realizada uma RSL para identificar, avaliar e sintetizar estudos relevantes sobre a integração de chatbots em ambientes tridimensionais. Esta metodologia foi escolhida devido à necessidade de mapear o estado atual do conhecimento neste campo emergente de forma estruturada e replicável.

A RSL seguiu as diretrizes metodológicas de Design Science Research (DSR), que é uma teoria que investiga a geração de conhecimento no processo de conceção de artefactos, ou seja, como os métodos de design podem constituir investigação científica [11]. A DSR

ênfatiza a importância do design e do desenvolvimento de artefactos em vários contextos. Ela permite uma abordagem iterativa, na qual a construção e a avaliação de artefactos são fundamentais para a geração de conhecimento, abordando não apenas questões teóricas, mas também práticas, através da aplicação de um modelo claro e rigoroso [12][13].

Para operacionalizar o processo, foi utilizada a ferramenta Parsifal, uma plataforma web que facilita a gestão de revisões sistemáticas através de recursos para definição de protocolos de investigação, gestão de bases de dados, aplicação de critérios de seleção e síntese de resultados [14].

2.1.2 Protocolo de Pesquisa

O nosso RSL seguiu, conforme mostrado na Figura 2.1, uma abordagem estruturada com as seguintes fases:

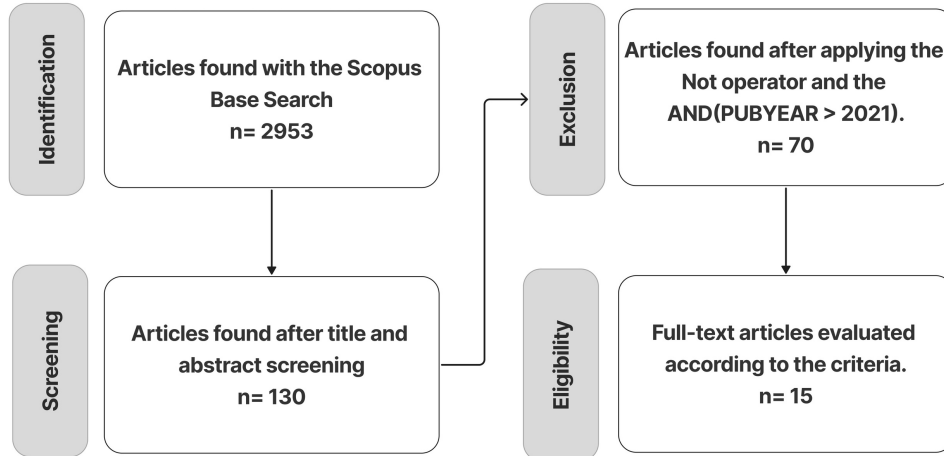


Figura 2.1: Fluxograma da Revisão Sistemática da Literatura

- Identificação: definimos conceitos amplos para capturar os temas centrais de interesse, incluindo chatbots, ambientes 3D, experiência do utilizador e áreas relacionadas. Esta abordagem abrangente visa maximizar a sensibilidade da pesquisa, garantindo que estudos potencialmente relevantes não sejam excluídos prematuramente.

- Triagem: refinamos a nossa pesquisa, limitando os resultados apenas aos artigos em que os nossos termos-chave apareciam nos títulos, resumos ou palavras-chave, utilizando o operador *TITLE-ABS-KEY*. Esta restrição garante que os conceitos centrais são os temas principais dos estudos e não apenas menções periféricas, reduzindo o ruído informativo.
- Exclusão: Para garantir a relevância dos resultados, limitamos a pesquisa a publicações recentes, incluindo apenas artigos publicados após 2021. Além disso, aplicamos um critério de exclusão utilizando o operador *NOT* para remover estudos pertencentes a domínios especializados fora do âmbito desta pesquisa. Especificamente, excluímos artigos cujos títulos ou resumos continham termos relacionados com Saúde ou Educação.
- Elegibilidade: após definirmos os critérios de inclusão e exclusão para filtrar o conjunto de artigos obtidos, os critérios utilizados são apresentados na Tabela 2.1.

Tabela 2.1: Critérios para a seleção

Critérios de inclusão	Critérios de exclusão
Artigos científicos	Menos de 2 citações Sem acesso

Implementação da pesquisa: A consulta de pesquisa final sintetizou todos os elementos de filtragem e é exibida na lista 2.1.

```

1 TITLE-ABS-KEY(("Chatbots" OR "Conversational Agents" OR "Virtual Assistants")
2 AND ("3D Environments" OR "Augmented Reality" OR "Virtual Reality")
3 AND ("User Experience" OR "E-commerce"))
4 AND NOT (TITLE-ABS-KEY("Healthcare" OR "Medical" OR "Education"
5 OR "E-learning" OR "Health Monitoring"))

```

Listing 2.1: Consulta de pesquisa

A pesquisa foi realizada em várias bases de dados acadêmicas, incluindo *IEEE Digital Library*, *Science@Direct*, *ACM Digital Library*, *Scopus* e *Springer Link*, para maximizar a cobertura da literatura técnica relevante.

Com os diferentes critérios definidos, a ferramenta online *Parsifal.al* foi utilizada para continuar a revisão sistemática da literatura. Utilizando a sequência de pesquisa, foram inicialmente encontrados 2.953 artigos e, após a aplicação dos critérios, os resultados foram reduzidos a 15 artigos, que foram utilizados ao longo deste capítulo para apoiar a análise do estado da arte.

2.2 Paradigmas de Interface em Ambientes Imersivos

2.2.1 Interfaces de Voz vs. Interfaces Gráficas

Evidências empíricas sugerem que as Interface de voz do utilizador (VUI) proporcionam experiências significativamente mais imersivas e acessíveis em comparação com as Interface gráfica do utilizador (GUI) em ambientes de realidade virtual. Estudos controlados demonstram que as VUI reduzem significativamente a carga cognitiva e o atrito de interface, especialmente para utilizadores com experiência técnica limitada [15]. Essa transição para modos naturais de interação estabelece uma base fundamental para a integração de chatbots em aplicações de computação espacial, sugerindo que os paradigmas conversacionais podem ser inerentemente mais adequados a contextos imersivos do que as interfaces tradicionais baseadas em menus.

A lógica por trás dessa preferência pode estar relacionada à redução da complexidade visual em ambientes já visualmente densos (ambientes 3D), onde a adição de elementos tradicionais da GUI pode competir pela atenção do utilizador e aumentar a sobrecarga cognitiva. Ao liberar o canal visual, as VUI permitem que o utilizador se concentre no ambiente tridimensional enquanto interage por meio de modalidades mais naturais e menos intrusivas.

2.2.2 Design e representação visual do avatar

Um trabalho sobre avatares conversacionais em realidade mista demonstra que a representação visual tem um impacto substancial nas dimensões cognitiva e afetiva da experiência do utilizador. Avatares semelhantes a humanos geraram métricas significativamente melhores para retenção de informação e envolvimento em comparação com representações simbólicas [16][17]. Além disso, o enquadramento psicológico das diferentes modalidades da realidade, com os ambientes de RV vistos como “espaço do agente” e a Realidade Aumentada (RA) como “espaço do utilizador”, sugere que as estratégias de implementação de chatbots devem ser adaptadas a contextos de realidade específicos. Estas conclusões sublinham a importância de um design de incorporação adequado ao utilizar agentes conversacionais em ambientes tridimensionais.

No entanto, é importante observar que pesquisas também revelam compromissos entre o realismo dos avatares e o desempenho técnico. Avatares altamente realistas exigem recursos computacionais substanciais, o que pode comprometer a fluidez da interação em dispositivos com capacidades limitadas. Essa tensão entre fidelidade visual e desempenho operacional representa um desafio de design que deve ser cuidadosamente equilibrado em implementações práticas. Pesquisas recentes sobre a incorporação de agentes em museus virtuais revelam que, embora objetos animados ofereçam novidade e potencial de envolvimento, a escolha deve equilibrar o envolvimento com a eficácia informativa [18]. Essa descoberta sugere que “mais realismo” nem sempre equivale a “melhor experiência”, exigindo uma abordagem contextual para o design de avatares.

2.3 Adaptação e Experiência do Utilizador

2.3.1 Agentes Contextuais e Consciência Ambiental

Inovações recentes em agentes de integração sensíveis ao contexto exemplificam a importância da consciência ambiental e histórica na criação de experiências conversacionais eficazes [19]. Ao modelar tanto o contexto espaço-temporal imediato quanto os padrões de

interação longitudinal, sistemas avançados demonstram como modelos de linguagem em grande escala podem fornecer respostas situacionalmente relevantes em ambientes virtuais complexos. O sistema Platform-Independent Context-Aware Onboarding Agent (PICAN) [20] representa um exemplo paradigmático dessa abordagem. Alimentado por modelos de linguagem de grande escala, o PICAN demonstra como a consciência contextual pode ser operacionalizada através da modelação de:

- **Contexto de curto prazo:** localização espacial atual, objetos próximos, ações recentes;
- **Contexto de longo prazo:** histórico de interações, preferências demonstradas, padrões comportamentais.

Estudos controlados com utilizadores revelam fortes preferências por interações sensíveis ao contexto, com melhorias significativas tanto no desempenho das tarefas como nas métricas de satisfação, estabelecendo assim a consciência contextual como um princípio fundamental de design para chatbots que operam em espaços 3D dinâmicos.

A capacidade de um agente virtual equilibrar orientação estruturada com liberdade exploratória surge como crucial neste contexto [21]. Os utilizadores demonstram preferências distintas em relação ao comportamento dos agentes virtuais, favorecendo particularmente sistemas que fornecem simultaneamente orientação especializada e capacidades de descoberta autónoma. Essa flexibilidade comportamental é fundamental para o desenvolvimento de chatbots que se adaptam a diferentes cenários e necessidades dos utilizadores, permitindo que sistemas mais sofisticados ajustem dinamicamente o seu nível de assistência com base nas habilidades demonstradas e nas preferências expressas pelo utilizador.

2.3.2 Fatores de Adoção e Aceitação Tecnológica

Uma análise aprofundada dos assistentes virtuais fornece informações valiosas sobre os fatores que impulsionam a adoção de assistentes de inteligência artificial em ambientes imersivos. Através da avaliação quantitativa de vários grupos de utilizadores, os investigadores identificaram cinco determinantes principais [22]:

1. **Utilidade Percebida** (Perceived Usefulness);
2. **Facilidade de Uso Percebida** (Perceived Ease of Use);
3. **Prazer Percebido** (Perceived Enjoyment);
4. **Intenção de Uso** (Intention to Use);
5. **Curiosidade** (Curiosity).

A análise estatística revela que a utilidade percebida é o indicador mais forte da intenção de adoção, uma descoberta particularmente relevante, pois desafia a suposição comum de que a “novidade tecnológica” ou o “entretenimento” são os principais impulsionadores da adoção. Em vez disso, os utilizadores priorizam a funcionalidade prática — o sistema deve realmente resolver problemas ou facilitar tarefas para justificar a sua adoção contínua.

Notavelmente, a curiosidade surge como particularmente influente para grupos demográficos específicos, enfatizando a importância de projetar sistemas que estimulem e mantenham o interesse do utilizador em ambientes 3D. Isso sugere que, para certos segmentos de utilizadores (possivelmente mais jovens ou mais aventureiros tecnologicamente), a exploração e a descoberta podem ser motivadores tão importantes quanto a utilidade funcional. Essas descobertas fornecem critérios de avaliação essenciais para a implementação de chatbots em espaços virtuais e ressaltam a importância do valor funcional para a aceitação do utilizador, indicando que qualquer sistema desenvolvido deve demonstrar valor tangível para os utilizadores, não apenas ser tecnologicamente impressionante.

Pesquisas sobre profundidade de conhecimento demonstram, por meio de estudos experimentais, que os utilizadores podem diferenciar claramente entre agentes com diferentes níveis de conhecimento (baixo, médio e alto), com a profundidade de conhecimento afetando significativamente as percepções de inteligência, relacionamento e factualidade [23]. Isso implica que os chatbots adaptativos requerem bases de conhecimento abrangentes que possam ser atualizadas dinamicamente, ao mesmo tempo em que adaptam o comportamento com base na experiência acumulada em diferentes cenários, criando um ciclo de

melhoria contínua que beneficia tanto o sistema quanto os utilizadores ao longo do tempo.

2.4 Integração de IA e Processamento de Linguagem Natural

2.4.1 Modelos de Linguagem de Grande Escala (LLMs)

Análises recentes documentam como modelos linguísticos avançados estão a remodelar as expectativas de interação natural em ambientes digitais. Os sistemas contemporâneos de IA generativa expandiram drasticamente as capacidades das interfaces conversacionais, permitindo uma compreensão linguística mais sofisticada, raciocínio contextual e geração de conteúdo dinâmico [24]. Esses avanços tecnológicos formam a base para chatbots mais capazes e responsivo em ambientes 3D. A implementação de grandes modelos linguísticos em assistentes virtuais para o ensino de anatomia em ambientes de realidade virtual imersivo [5] demonstra o potencial dos sistemas de IA generativa para fornecer assistência educacional sofisticada, conseguindo manter a coerência conversacional e, ao mesmo tempo, fornecer informações tecnicamente precisas em contextos espaciais tridimensionais.

No entanto, implementar Large Language Models (LLMs) em chatbots baseados em avatares apresenta desafios arquitetónicos significativos que transcendem a simples integração tecnológica. Estudos revelam a complexidade inerente à criação de chatbots de avatares de domínio aberto que possam manter conversas naturais e contextualmente apropriadas [10], preservando a consistência visual e comportamental esperada de agentes incorporados em ambientes tridimensionais. Os desafios documentados incluem alucinações de LLMs, em que os modelos podem gerar informações factualmente incorretas com aparente confiança [20], erros de transcrição em que os sistemas de reconhecimento de voz podem interpretar erroneamente comandos em ambientes ruidosos, dificuldades de sincronização multimodal ao alinhar respostas verbais com animações de avatares e *feedback* visual, e problemas de latência na manutenção de tempos de resposta aceitáveis em ambientes computacionalmente exigentes. Esses desafios ressaltam que a integração

bem-sucedida de LLMs em ambientes 3D requer não apenas recursos linguísticos avançados, mas também arquiteturas cuidadosamente otimizadas que levem em consideração as restrições específicas dos ambientes imersivos.

2.4.2 Arquiteturas Cliente-Servidor e PLN Avançado

Os assistentes industriais habilitados para linguagem natural demonstram um potencial sofisticado para a colaboração entre humanos e IA por meio de arquiteturas cliente-servidor com técnicas avançadas de Processamento de Linguagem Natural (PLN), facilitando interações intuitivas entre operadores e vários sistemas automatizados [4]. Esses princípios arquitetônicos são diretamente aplicáveis ao desenvolvimento de chatbots que se integram a rastreadores da web e várias fontes de informação, mantendo recursos de conversação natural. A integração de tecnologias de aprendizagem automática e personalização mostra melhorias mensuráveis na precisão e satisfação do utilizador, com sistemas de recomendação alimentados por IA a demonstrar a capacidade de melhorar significativamente a experiência de seleção de produtos no comércio eletrônico, permitindo uma personalização mais eficaz das sugestões apresentadas aos consumidores [25]. Esta convergência entre o processamento de linguagem natural e os sistemas de recomendação representa uma área com implicações práticas relevantes para as interações comerciais digitais.

Tecnologias emergentes, como assistentes virtuais e realidade aumentada, representam áreas promissoras para transformar a interação [22], embora a implementação bem-sucedida exija uma gestão cuidadosa dos desafios que incluem privacidade de dados, viés algorítmico e supervisão humana adequada [24]. Essa necessidade de equilibrar a inovação tecnológica com considerações éticas e práticas torna-se particularmente crítica em sistemas que operam em ambientes imersivos, onde a presença aumentada e o envolvimento emocional podem amplificar tanto os benefícios quanto os riscos potenciais da tecnologia.

2.5 Implementação Responsável e Desempenho

2.5.1 Métricas de desempenho e satisfação

Implementações práticas de IA em contextos conversacionais mostram melhorias mensuráveis no desempenho, alcançando uma redução de até 30% nos tempos de resposta e aumentando a satisfação do utilizador em 15% [26]. Essas métricas fornecem referências concretas para avaliar chatbots adaptativos e definir expectativas realistas para os ganhos de desempenho esperados. É importante contextualizar estes números: uma redução de 30% no tempo de resposta, embora significativa, deve ser avaliada em relação à experiência geral do utilizador, pois um sistema que responde rapidamente, mas fornece informações irrelevantes ou imprecisas, não representa uma melhoria genuína. Da mesma forma, um aumento de 15% na satisfação, embora positivo, deve ser validado por meio de estudos longitudinais para confirmar que o efeito persiste após a fase inicial de novidade tecnológica.

Sistemas avançados demonstram navegação dinâmica na interface enquanto aprendem com interações anteriores para melhorar a eficiência na conclusão de tarefas e a satisfação do utilizador [19]. Essa capacidade de aprendizagem adaptativa é crucial para chatbots que operam em vários cenários, permitindo a otimização contínua com base em padrões de uso reais e criando sistemas que se tornam progressivamente mais eficazes à medida que acumulam experiência com diferentes tipos de utilizadores e contextos de interação [27].

2.5.2 Antropomorfismo e Confiança do Utilizador

Características de interação semelhantes às humanas, particularmente vozes e simpatia percebida, aumentam significativamente a confiança e a aceitação do utilizador, correlacionando-se diretamente com uma maior adoção do sistema [28]. Essas características fornecem diretrizes de design para aplicações que exigem a confiança do utilizador. No entanto,

pesquisas também destacam a natureza dual dos sistemas inteligentes, que podem simultaneamente aumentar a satisfação do utilizador e causar consequências indesejadas [17]. A consciência desses mecanismos de influência é crucial para uma implementação responsável. Estudos defendem sistemas transparentes e éticos que incorporem mecanismos de explicabilidade e *feedback* do utilizador para garantir que a tecnologia sirva para enriquecer o utilizador, em vez de explorá-lo [29].

2.6 Análise Comparativa de Sistemas Conversacionais em Ambientes 3D

Para compreender a evolução das estratégias para melhorar as experiências de integração em ambientes metaversos imersivo, foram analisados e comparados três projetos distintos, mas tematicamente alinhados: PIKAN[20], Conversa Livre com Avatares Humanos e Simbólicos [17] e uma categoria mais ampla de Agentes de Integração Conceitual [26].

A seleção destes três projetos foi orientada pelo seu foco comum na integração do utilizador em ambientes imersivos, embora representem abordagens diferentes em termos de filosofia de design, integração tecnológica e envolvimento contextual.

A Tabela 2.2 apresenta uma análise comparativa estruturada de acordo com vários critérios relevantes que nos permitem identificar pontos fortes, limitações e oportunidades para a integração conceptual entre diferentes abordagens. Cada projeto relatou experiências positivas dos utilizadores, particularmente no que diz respeito à interação natural e à redução do atrito na integração. Os utilizadores do estudo Avatars preferiram avatares humanos pela sua naturalidade [16], enquanto os utilizadores do PIKAN apreciaram a capacidade de resposta sensível ao contexto [9]. O *feedback* dos projetos destacou o valor de referências claras, tratamento de erros e manutenção do envolvimento sem sobrecarregar o utilizador, elementos que emergem como requisitos transversais, independentemente da abordagem específica adotada.

Tabela 2.2: Análise comparativa de agentes conversacionais em ambientes imersivos

Critério	PICAN[20]	Avatares Humanos vs. Simbólicos[17]	Agentes de Integração Conceituais[26]
Uso de IA/PLN	Baseado em LLMs	Baseado em GPT-3	LLMs e Non-Player Characters (NPCs) generalizados
Consciência Contextual	Alta (curto e longo prazo)	Moderada: espacial, dialógica, consciente de ações	Presente: espacial, dialógica, curto e longo prazo
Tipo de Interface	Voz + avatar virtual	Voz + avatares de realidade mista	VUI
Design Centrado no Utilizador	Estudos de utilizadores incluídos	Design experimental intra-sujeito	Estudos formativos, inquéritos, <i>feedback</i> iterativo
Domínio de Aplicação	Integração no Meta-verso	Comunicação Mixed Reality (MR) e integração	Integração generalizada ao metaverso
Adaptabilidade	Moderada: focado em metaverso	Moderada: adaptável entre RA/RV	Moderada a alta: integração personalizada
Avaliação	Qualitativa (entrevistas)	Quantitativa: escala Likert	Qualitativa + Quantitativa
Desafios Reportados	Alucinações de LLMs, erros de transcrição	Realismo do avatar vs. compromissos de desempenho	Limites técnicos, variabilidade de conforto, manutenção do compromisso
Métricas de UX	Envolvimento, satisfação	Satisfação, prazer, naturalidade, recordação	Imersão, facilidade de uso, satisfação, envolvimento
Integração Tecnológica	LLMs + ambiente RV	GPT-3 + Meta Quest Pro (RV/RA)	Unity/Unreal + dispositivos RV/RA

A comparação revela uma ambição comum: tornar o metaverso mais acessível, centrado no ser humano e envolvente. No entanto, cada sistema tem pontos fortes e limitações distintas que justificam uma análise crítica. O PICAN exemplifica o poder da modelagem de contexto profundo, com sua capacidade de manter a memória de curto e longo prazo, permitindo interações verdadeiramente personalizadas. No entanto, a dependência de LLMs traz desafios de alucinação que podem comprometer a confiabilidade das informações [9]. Esse sistema seria particularmente adequado para cenários em que a personalização é crítica, mas menos apropriado para domínios em que a precisão factual é essencial (como sistemas médicos ou financeiros). O estudo Avatars [16] destaca o impacto da incorporação visual e do design de avatares na experiência do utilizador, com uma clara preferência por avatares humanóides, sugerindo que os utilizadores valorizam representações familiares em ambientes sociais virtuais. No entanto, as compensações entre realismo e desempenho indicam que este não é um problema resolvido, exigindo que

dispositivos com capacidades limitadas façam concessões no realismo visual.

As suas ideias reforçam-se mutuamente: o rico tratamento contextual do PICAN pode beneficiar do realismo dos avatares, enquanto as estruturas conceptuais podem adotar os métodos de avaliação do PICAN para amadurecer. Juntos, eles ilustram o desafio multi-dimensional da integração, abrangendo o design de IA, o desenvolvimento de interfaces, a psicologia do utilizador e as restrições técnicas. É importante reconhecer que esta comparação se baseia em estudos com diferentes metodologias de avaliação (qualitativa vs. quantitativa), populações de utilizadores potencialmente distintas e contextos de aplicação variados, pelo que uma comparação mais rigorosa exigiria estudos controlados com as mesmas métricas aplicadas a todos os sistemas em condições equivalentes.

2.7 Privacidade, Ética e Regulação

2.7.1 Regulamento Geral de Proteção de Dados (GDPR)

A implementação de chatbots conversacionais em ambientes 3D levanta desafios críticos de privacidade e proteção de dados, particularmente no contexto do GDPR da União Europeia [Regulamento (União Europeia (UE)) 2016/679] [30]. O GDPR estabelece requisitos rigorosos para o processamento de dados pessoais que são diretamente aplicáveis a sistemas conversacionais, especialmente quando estes recolhem dados biométricos (reconhecimento de voz, padrões de fala, dados de rastreamento ocular em ambientes RV/RA), processam históricos de interação (registos de conversas que podem revelar preferências, comportamentos ou informações sensíveis), ou implementam personalização através de perfis de utilizador construídos via aprendizagem automática que constituem processamento automatizado.

Princípios fundamentais do GDPR, como licitude, lealdade e transparência descritos no Artigo 5º [31], impõem a obrigação de informar clara e detalhadamente os usuários sobre os dados que são coletados e sua utilização, representando um desafio significativo em ambientes imersivos tridimensionais onde as interfaces devem equilibrar comunicação

eficaz com minimização de intrusão. Outro princípio fundamental é a limitação das finalidades: os dados recolhidos para melhorar a experiência conversacional não podem ser reutilizados para objetivos secundários, como marketing direcionado, sem o consentimento explícito e adicional do usuário. O princípio da minimização dos dados orienta que apenas informações estritamente necessárias sejam coletadas, evitando a captação excessiva ou “por precaução”. Por fim, o direito ao esquecimento, garantido pelo Artigo 17º [32], dá aos utilizadores o direito de exigir a eliminação de seus dados pessoais, o que levanta questões técnicas e éticas complexas em chatbots que utilizam o histórico para personalização, relacionadas com o equilíbrio entre a personalização contínua e a eliminação efetiva dos dados solicitados.

Em ambientes imersivos, a recolha de dados sensíveis atinge níveis mais profundos, incluindo dados comportamentais espaciais (movimentos, gestos, padrões de olhar), dados biométricos (características vocais e padrões linguísticos) e inferências emocionais baseadas em análise comportamental. Esses tipos de dados se enquadram nas “categorias especiais” do GDPR (Artigo 9º) [33], exigindo salvaguardas adicionais robustas. Dessa forma, a conformidade não deve ser vista como uma etapa posterior, mas implementada desde o design inicial do sistema, seguindo os preceitos de proteção de dados desde a conceção e por predefinição (Artigo 25º) [34]. Isso inclui técnicas como anonimização ou pseudonimização, processamento local de dados sensíveis para minimizar transmissões, mecanismos claros de consentimento e revogação, além de transparência algorítmica que esclareça o uso e o tratamento das informações pessoais.

2.7.2 Desafios Éticos em Agentes Conversacionais

Além dos requisitos legais, questões éticas substantivas surgem na implementação de chatbots em ambientes 3D que exigem uma análise cuidadosa por parte de designers, programadores e organizações implementadoras. Análises sobre a regulamentação ética global da IA identificam várias áreas de preocupação que são particularmente graves em contextos conversacionais imersivos [35].

O viés algorítmico é uma preocupação central, pois os modelos de linguagem treinados com dados históricos podem perpetuar ou amplificar preconceitos sociais existentes, com pesquisas que documentam preconceitos de gênero, raça e cultura em sistemas de IA conversacionais. Em ambientes imersivos, onde a presença social é amplificada, esses preconceitos podem ter impactos psicológicos mais profundos do que nas interfaces tradicionais.

Outro desafio ético significativo envolve a manipulação e a persuasão por meio de avatares antropomórficos, cuja semelhança com as interações humanas gera alta confiança por parte do utilizador. Essa confiança pode ser explorada para fins comerciais ou outros que não atendam aos interesses do utilizador, confundindo a linha entre “persuasão útil” e “manipulação exploradora”. Este risco é especialmente relevante em contextos comerciais, onde a maximização do lucro pode entrar em conflito com a proteção do bem-estar do utilizador, exigindo uma reflexão ética cuidadosa e uma regulamentação adequada [36].

Existe uma preocupação ética sobre o potencial de chatbots sofisticados substituírem interações humanas genuínas, particularmente em contextos sociais. Embora os chatbots possam proporcionar eficiência e disponibilidade 24 horas por dia, 7 dias por semana, eles também podem contribuir para o isolamento social ou a degradação das habilidades interpessoais se forem usados como substitutos sistemáticos da interação humana, em vez de complementos. A questão da responsabilidade e responsabilização torna-se crítica: quando um chatbot fornece informações incorretas ou causa danos (psicológicos, financeiros), quem é o responsável — o desenvolvedor, o operador da plataforma ou o modelo de IA subjacente? Essa questão de responsabilidade torna-se mais complexa em sistemas que aprendem e evoluem de forma autônoma, potencialmente desenvolvendo comportamentos não previstos pelos seus criadores originais.

2.7.3 Transparência e Explicabilidade (XAI)

A importância da Inteligência Artificial Explicável (XAI) nos chatbots está bem estabelecida na literatura, pois é essencial que os utilizadores compreendam as razões por

trás das recomendações, decisões e personalizações oferecidas pela IA. Em ambientes 3D, isso implica oferecer explicações que não comprometam a imersão nem sobrecarreguem cognitivamente o utilizador. Técnicas promissoras incluem explicações incorporadas espacialmente no ambiente virtual, destacando elementos relevantes para o contexto da conversa, variando os níveis de detalhe explicativo e o uso de metáforas visuais intuitivas para comunicar processos algorítmicos complexos de uma forma acessível.

Avanços recentes demonstram como modelos linguísticos em grande escala podem ser adaptados para tarefas específicas de conformidade jurídica, incluindo a classificação das regras do GDPR utilizando formatos estruturados, como *LegalDocML* e *LegalRuleML* [37].

A transparência é vista não apenas como conformidade regulatória, mas como um elemento ético essencial para construir confiança no sistema. O desenvolvimento de mecanismos de *feedback* e sistemas transparentes garante que a tecnologia beneficie e capacite os utilizadores, evitando a exploração e reforçando a responsabilidade. Várias estruturas éticas, como a Lei de IA da UE, o *IEEE Ethically Aligned Design* e os Princípios para Algoritmos Responsáveis, propõem diretrizes para a incorporação sistemática de transparência, explicabilidade e responsabilidade em sistemas de IA, incluindo chatbots 3D, enfatizando a necessidade de ir além da mera conformidade legal para garantir sistemas confiáveis e benéficos [38].

2.8 Lacunas de Investigação

A revisão sistemática da literatura revelou avanços significativos na integração de agentes conversacionais em ambientes imersivos, mas também identificou lacunas importantes que justificam mais investigação e desenvolvimento prático.

2.8.1 Integração Técnica e Arquitetural

Apesar do progresso documentado, ainda há uma escassez de estudos que abordem sistematicamente a integração completa dos sistemas RAG com ambientes 3D imersivos. A maioria das pesquisas concentra-se em componentes isolados, seja o design de avatares ou

o processamento de linguagem natural, mas raramente explora a arquitetura completa necessária para criar experiências conversacionais verdadeiramente funcionais em contextos comerciais reais. Essa fragmentação resulta numa lacuna entre protótipos acadêmicos promissores e implementações comercialmente viáveis que possam operar de forma confiável em ambientes de produção.

A literatura existente oferece orientações limitadas sobre otimização de desempenho em dispositivos móveis de RV com recursos computacionais limitados, como o Meta Quest. Embora os estudos explorem recursos avançados em sistemas *desktop* com Graphics Processing Units (GPUs) dedicadas, poucos abordam as compensações críticas necessárias para manter a qualidade da conversação e taxas de quadros adequadas em hardware autônomo. Essa lacuna é particularmente relevante, considerando que os dispositivos móveis representam o ponto de entrada mais acessível para experiências comerciais de RV.

2.8.2 Adaptabilidade e Generalização de Sistemas

Os sistemas descritos na literatura são predominantemente desenvolvidos para domínios específicos, como educação, saúde, museus virtuais, com profunda personalização para contextos específicos. No entanto, há uma notável falta de investigação sobre estruturas genéricas e adaptáveis que possam ser rapidamente configuradas para diferentes cenários comerciais sem exigir uma reengenharia fundamental. Esta limitação representa uma barreira significativa para uma adoção mais ampla, especialmente para pequenas e médias empresas que não dispõem de recursos para um desenvolvimento personalizado extensivo.

A questão da atualização dinâmica do conhecimento continua pouco explorada. Embora os sistemas RAG demonstrem capacidades de recuperação de informação, poucos estudos abordam como manter bases de conhecimento atualizadas de forma autônoma em ambientes onde os produtos, preços e especificações mudam frequentemente. A dependência de atualizações manuais compromete a escalabilidade e a precisão factual a longo prazo.

2.8.3 Considerações Éticas e Privacidade

Embora a literatura aborde extensivamente questões éticas relacionadas à IA conversacional e aos requisitos do GDPR em contextos gerais, há poucas pesquisas sobre os desafios específicos que surgem em ambientes imersivos. A recolha de dados biométricos em RV — padrões de movimento ocular, características vocais, comportamento espacial — levanta questões de privacidade que transcendem as preocupações tradicionais dos chatbots baseados em texto. A percepção intensificada da presença social em ambientes 3D também pode amplificar os riscos de manipulação ou impactos psicológicos negativos que exigem salvaguardas adicionais ainda não adequadamente conceituadas na literatura existente.

2.9 Tecnologias e Ferramentas

A implementação prática de agentes conversacionais em ambientes imersivos depende criticamente da seleção adequada das tecnologias subjacentes.

2.9.1 Plataformas de Desenvolvimento para Ambientes 3D Imersivos

O desenvolvimento de experiências imersivas de realidade virtual requer plataformas especializadas que forneçam ferramentas de renderização eficientes, sistemas físicos, gestão de áudio espacial e integração com dispositivos de realidade virtual. O Unity 2022.3 LTS (Long Term Support) surge como a escolha predominante para esse tipo de projeto, oferecendo um ecossistema de desenvolvimento maduro com amplo suporte para Meta Quest 3, *HTC Vive*, *Valve Index* e outras plataformas de RV/RA.

O Unreal Engine representa uma alternativa viável, especialmente para projetos que priorizam a máxima fidelidade gráfica e efeitos visuais sofisticados. O *Nanite* (virtualização geométrica) e o *Lumen* (iluminação global dinâmica) permitem a renderização de

ambientes extraordinariamente complexos em tempo real. No entanto, a curva de aprendizagem mais íngreme, os requisitos de hardware mais exigentes e a sobrecarga computacional significativamente maior tornam-no menos adequado para experiências em dispositivos de RV móveis, como o Meta Quest.

O *WebXR* e as plataformas web (*Babylon.js*, *Three.js*) oferecem uma alternativa de distribuição que elimina a necessidade de instalação nativa. Permitem o acesso através de navegadores web em dispositivos RV compatíveis, reduzindo o atrito na adoção. No entanto, ainda apresentam limitações em termos de desempenho máximo, acesso ao hardware e integração de processos complexos de IA em comparação com as soluções nativas. Esta abordagem é adequada para protótipos, demonstrações ou aplicações em que a acessibilidade supera os requisitos de desempenho extremos.

2.9.2 Modelos Conversacionais

A qualidade das interações conversacionais depende criticamente da seleção do modelo de linguagem utilizado. O *GPT-4o-mini* (OpenAI) oferece um equilíbrio ideal entre compreensão linguística sofisticada, velocidade de inferência (latência reduzida essencial para conversas naturais) e custo operacional, tornando-o a escolha preferida para sistemas de atendimento ao cliente em produção. O modelo demonstra uma compreensão contextual robusta, a capacidade de seguir instruções precisas através de instruções bem estruturadas e a geração de respostas coerentes e naturais em português, o que é importante para os mercados lusófonos.

O *GPT-4 Turbo* oferece capacidades superiores em raciocínio complexo, compreensão de nuances linguísticas e geração de conteúdo sofisticado. É adequado para cenários que exigem análise aprofundada de consultas ou respostas ambíguas envolvendo síntese de várias fontes de conhecimento. No entanto, latências significativamente mais altas (2 a 5 segundos contra 200 a 500 ms para o *GPT-4o-mini*) comprometem a naturalidade das conversas em tempo real, tornando-o menos adequado para chatbots imersivos, onde a resposta rápida é fundamental.

Modelos de código aberto, como *Llama 2* (Meta), *Mistral 7B* ou *Zephyr*, oferecem controlo total, sem dependência de Application Programming Interfaces (APIs) externas, e podem ser auto-hospedados em infraestruturas privadas (importante para conformidade com o GDPR e privacidade). As desvantagens incluem qualidade inferior na compreensão do português, recursos computacionais significativos necessários para uma inferência rápida e a necessidade de um ajuste fino extenso para domínios específicos. A decisão deve equilibrar a autonomia tecnológica com a qualidade do serviço e os recursos disponíveis.

O Convai representa uma solução especializada para processamento de voz em tempo real otimizada para conversas naturais. Ele integra reconhecimento de fala, compreensão de intenções, geração de respostas e síntese de fala em um pipeline unificado, reduzindo a latência e facilitando a sincronização multimodal com avatares. Ele oferece suporte nativo para vários idiomas, variações de sotaque e padrões de linguagem informal comuns em interações conversacionais reais.

2.9.3 Arquiteturas de recuperação e Ampliação

A RAG combina os recursos de geração de modelos de linguagem com acesso a bases de conhecimento estruturadas, permitindo respostas factualmente precisas com base em dados específicos da organização. O padrão RAG mitiga um dos principais desafios das LLMs, nomeadamente as alucinações (geração de informações factualmente incorretas), fornecendo um contexto factual verificado a partir do qual o modelo deve derivar respostas. O sistema segue um fluxo:

- a consulta do utilizador é primeiro usada para recuperar documentos relevantes de uma base de conhecimento;
- esses documentos são então fornecidos como contexto para o modelo de linguagem, que gera uma resposta com base nesse contexto.

O LangChain fornece uma estrutura que coordena componentes de processamento natural e integração com modelos de linguagem. Ele oferece abstrações de alto nível para

RAG, histórico de conversas, *prompts* avançados e integração com várias APIs e modelos, reduzindo significativamente a complexidade da implementação [39].

O Pinecone é uma base de dados vetorial sem servidor otimizado para operações de similaridade semântica. Cada documento é convertido em uma representação vetorial de alta dimensão (incorporação) que captura o significado semântico. Quando um utilizador faz uma consulta, ela também é convertida em uma incorporação e comparada com incorporações de documentos usando similaridade cosseno, identificando rapidamente documentos semanticamente relevantes. Ele mantém consistentemente tempos de resposta abaixo de 100 ms, mesmo com milhões de documentos, o que é fundamental para experiências em tempo real [40].

Text-embedding-3-small (OpenAI) é o modelo usado para converter texto em *embeddings* vetoriais de 1.536 dimensões. Ele oferece compreensão semântica superior em comparação com modelos anteriores, permitindo uma correspondência mais precisa entre a intenção do utilizador e o conteúdo armazenado, mesmo quando as palavras exatas diferem, mas o significado é semelhante.

2.9.4 Processamento de Dados e Infraestrutura

Beautiful Soup e *aiohttp* formam o pipeline de extração de dados da web. *Beautiful Soup* fornece análise HyperText Markup Language (HTML) estruturada e extração de conteúdo relevante de páginas da web desorganizadas. *Aiohttp* permite solicitações Hypertext Transfer Protocol (HTTP) assíncronas de alto desempenho, possibilitando o rastreamento paralelo de várias páginas sem bloqueio, acelerando significativamente a coleta de dados.

O FastAPI implementa a camada Representational State Transfer (REST) APIs que coordena as interações entre o cliente RV, os sistemas Convai remotos e os componentes de *Backend* de processamento de IA. O FastAPI oferece *async/await* nativo, geração automática de documentação APIs, validação de dados robusta e desempenho de produção, mantendo um código limpo e sustentável. Os padrões de *middleware* implementam

Cross-Origin Resource Sharing (CORS), limitação de taxa, registo estruturado e segurança adequada para ambientes de produção.

Meta Quest 3 e compatibilidade com RV em geral: a escolha dos dispositivos impõe restrições de desempenho que informam todas as decisões arquitetónicas anteriores. O Meta Quest 3 oferece processamento gráfico limitado em comparação com os computadores de secretária, exigindo otimizações agressivas na contagem de polígonos, texturas, sistemas de partículas e lógica em tempo real. Isso justifica a escolha do Unity em vez do Unreal (menor sobrecarga), do *GPT-4o-mini* em vez do *GPT-4 Turbo* (inferência mais rápida) e da comunicação assíncrona que minimiza a latência da rede.

Capítulo 3

Arquitetura do Sistema

Este capítulo apresenta a abordagem metodológica adotada no desenvolvimento do sistema 3D-Chatbot, detalhando desde a análise do problema até à modelação completa da solução proposta. A estrutura deste capítulo foi organizada de forma a proporcionar uma compreensão progressiva e aprofundada do projeto, começando pela identificação das limitações dos sistemas atuais de suporte ao utilizador em ambientes de realidade virtual.

3.1 Descrição do Problema

A crescente adoção de ambientes tridimensionais imersivos em contextos comerciais expôs limitações significativas nos modelos tradicionais de suporte ao utilizador. Sistemas baseados em interfaces bidimensionais, como chatbots na web, formulários e Frequently Asked Questions (FAQS) revelam-se inadequados para a realidade virtual, comprometendo tanto a experiência do utilizador quanto os objetivos comerciais.

A descontinuidade da experiência é uma das fraquezas mais evidentes dos sistemas atuais. Os utilizadores são frequentemente obrigados a sair do ambiente imersivo para aceder a informações ou suporte, quebrando a sensação de presença e reduzindo significativamente o envolvimento. Essa transição entre diferentes modalidades de interface cria atrito cognitivo e diminui a eficácia da experiência geral, resultando em frustração e abandono. Ao mesmo tempo, as limitações contextuais representam outro obstáculo

fundamental, uma vez que os chatbots convencionais carecem de consciência espacial do ambiente 3D no qual o utilizador está imerso, não conseguindo referenciar objetos específicos no espaço virtual, compreender a posição ou orientação do utilizador ou adaptar as suas respostas ao contexto imersivo específico. Essa desconexão entre o assistente virtual e o ambiente em que opera cria uma experiência artificial e pouco intuitiva, na qual os utilizadores precisam descrever verbalmente elementos visuais que estão diretamente à sua frente, comprometendo a naturalidade e a eficiência da comunicação.

O conteúdo desatualizado surge como um desafio operacional persistente que afeta particularmente as organizações com catálogos de produtos extensos ou informações que mudam com frequência. Atualizar manualmente o conteúdo dos chatbots é trabalhoso, propenso a erros e muitas vezes resulta em informações desatualizadas ou inconsistentes, um problema agravado pela proliferação de canais de comunicação e pela velocidade com que os produtos e serviços evoluem no mercado atual. Ao mesmo tempo, a falta de naturalidade na interação torna-se evidente em ambientes imersivos, onde os utilizadores esperam uma comunicação próxima da humana. Os chatbots tradicionais baseados em regras ou *scripts* predefinidos revelam-se inadequados, e a rigidez das respostas, combinada com a incapacidade de manter conversas fluidas e memória conversacional, compromete a credibilidade e a utilidade do sistema, obrigando os utilizadores a repetir informações já fornecidas.

A falta de monitorização e otimização representa uma lacuna estratégica significativa, uma vez que a maioria das implementações carece de sistemas robustos que permitam aos administradores compreender os padrões de utilização, identificar problemas de desempenho ou otimizar continuamente a experiência com base em dados de interação reais. Sem visibilidade sobre como os utilizadores interagem com o sistema, quais as perguntas mais frequentes ou onde ocorrem falhas, as organizações operam sem dados acionáveis que permitam a evolução orgânica do sistema. Por fim, os desafios técnicos da integração constituem uma barreira significativa à entrada, uma vez que a combinação das tecnologias necessárias processamento avançado de linguagem natural, reconhecimento e síntese de voz, animação de avatares, sincronização multimodal e acesso a bases de conhecimento

dinâmicas apresenta desafios arquitetônicos consideráveis que exigem conhecimentos multidisciplinares e investimentos substanciais. Essas limitações são particularmente críticas em contextos comerciais, onde a qualidade da experiência do cliente influencia diretamente as decisões de compra, a percepção da marca e a lealdade, representando uma lacuna que limita o potencial dos espaços virtuais comerciais e compromete o retorno do investimento em tecnologias imersivas.

3.2 Proposta de Solução

Para enfrentar os desafios identificados, este trabalho propõe o desenvolvimento de um sistema integrado, que combina tecnologias avançadas de inteligência artificial conversacional com ambientes imersivos de realidade virtual. A solução proposta baseia-se em quatro pilares fundamentais:

1. **Agente conversacional contextualizado:** Implementação de um chatbot baseado em LLMs capaz de manter a consciência do contexto espacial tridimensional do utilizador. Este agente preserva o histórico de conversas para interações mais naturais e coerentes, adaptando as respostas com base na posição e orientação do utilizador no ambiente virtual. O processamento avançado de linguagem natural permite-lhe compreender intenções e gerar respostas contextualizadas.
2. **Sistema RAG:** Para garantir respostas precisas e atualizadas, é integrado um pipeline automatizado para recolher informações (rastreamento da web) de sites corporativos. As informações são convertidas em *embeddings* semânticos, formando uma base de conhecimento vetorial que permite a recuperação inteligente de documentos relevantes com base na similaridade semântica. Assim, o sistema gera respostas bem fundamentadas e verificáveis, reduzindo significativamente a ocorrência de alucinações.

3. **Experiência imersiva de realidade virtual:** O ambiente tridimensional proposto integra naturalmente o avatar conversacional, permitindo a interação multimodal por meio de voz, gestos e controladores de RV. O sistema garante a sincronização precisa entre animações faciais e síntese de voz, bem como locomoção otimizada por meio de teletransporte, rotação ou movimento contínuo, proporcionando uma experiência fluida e imersivo.
4. **Sistema de gestão e monitorização:** Uma plataforma de *Backoffice* baseada na web fornece visualização em tempo real de métricas de desempenho e utilização, análise de padrões de interação e perguntas frequentes, e gestão dinâmica da base de conhecimento (adicionar, remover ou reindexar documentos). Além disso, permite monitorizar sessões ativas e aceder ao histórico de conversas, facilitando a manutenção e evolução contínuas do sistema.

3.3 Ferramentas e Tecnologias Utilizadas

A implementação do sistema 3D-Chatbot requer um conjunto diversificado de tecnologias, organizadas por camadas funcionais:

3.3.1 Ambiente virtual 3D

- Unity: Motor de desenvolvimento 3D principal utilizado para criar a experiência de realidade virtual.
- C#: Linguagem de programação utilizada no desenvolvimento Unity para comunicação com APIs externas.
- OpenXR: Camada de abstração que fornece compatibilidade com vários dispositivos de RVs.
- Meta Quest 3: Dispositivo de RV, o foco principal do projeto, oferecendo recursos autônomos com um processador *Snapdragon XR2 Gen 2* e telas de alta resolução.

3.3.2 Inteligência Artificial e Processamento de Linguagem

- GPT-4o-mini: modelo de linguagem OpenAI usado para gerar respostas conversacionais. Oferece um equilíbrio ideal entre qualidade de resposta e custo operacional.
- LangChain: estrutura Python para orquestrar componentes de IA, facilitando a implementação de pipelines complexos, como RAG.
- Convai: Plataforma especializada em processamento de voz em tempo real, fornecendo recursos de Speech-to-Text (STT), Text-to-Speech (TTS) e sincronização labial para avatares conversacionais.

3.3.3 Bases de Dados e Armazenamento

- Pinecone: Base de dados vetorial sem servidor otimizada para operações de similaridade semântica. Armazena as incorporações de documentos indexados e permite consultas rápidas (< 100 ms) essenciais para o sistema RAG.
- MongoDB: Base de dados Not Only Structured Query Language (NoSQL) orientada para documentos usada para armazenar sessões de utilizadores, histórico de conversas e metadados do sistema. A sua estrutura flexível é ideal para dados semiestruturados.

3.3.4 Backend e APIs

- FastAPI: Framework Python moderno para a criação de APIs REST assíncronas de alto desempenho e oferece validação automática de dados através do Pydantic .
- Python: principal linguagem de programação *Backend*, escolhida por seu rico ecossistema de bibliotecas de IA e processamento de dados.
- aiohttp: cliente HTTP assíncrono para Python, usado no rastreador da web para realizar várias solicitações simultâneas com eficiência.

- Beautiful Soup: biblioteca Python para analisar e extrair conteúdo HTML, essencial para processar páginas da web durante o rastreamento.

3.3.5 Backoffice

- Jinja2: Motor de modelos para Python que permite a geração dinâmica de páginas HTML no lado do servidor, mantendo a separação entre lógica e apresentação.
- Bootstrap: Estrutura Cascading Style Sheets (CSS) responsivo que fornece componentes pré-estilizados e um sistema de grelha flexível, garantindo uma interface administrativa consistente e profissional.
- JavaScript: linguagem de programação do lado do cliente usada no *Backoffice* para adicionar interatividade às páginas da web.

3.3.6 Infraestrutura

- Docker: Plataforma de contentorização que encapsula todos os serviços de *Backend* (FastAPI, MongoDB, Redis) em contentores isolados, garantindo a consistência entre os ambientes de desenvolvimento, teste e produção. Facilita a implementação do sistema e a escalabilidade horizontal.
- Git: Sistema de controlo de versão distribuído usado para gerenciamento de código-fonte, permitindo colaboração, rastreamento de alterações e reversão quando necessário.
- JSON Web Tokens (JWT): Padrão para autenticação e autorização sem estado. Os tokens JWT são gerados após o login bem-sucedido e incluídos em todas as solicitações subsequentes ao *Backoffice*, garantindo segurança sem a necessidade de sessões do lado do servidor.

3.4 Análise dos Requisitos

A análise dos requisitos envolve a identificação das necessidades dos utilizadores e as funcionalidades que o sistema 3D-Chatbot deve possuir. Estes requisitos foram divididos em duas categorias: funcionais e não funcionais.

3.4.1 Requisitos Funcionais

- Interação conversacional em RV: O sistema deve permitir que os utilizadores interajam com um avatar conversacional por meio de voz em um ambiente de RV.
- Recuperação de Informação Baseada em RAG: O chatbot deve responder a perguntas específicas sobre produtos/serviços usando informações recuperadas da base de conhecimento.
- Aquisição de Dados: O sistema deve recolher e processar automaticamente informação proveniente de websites previamente especificados.
- Gestão de Base de Conhecimento: Administradores devem poder gerir documentos indexados através de uma interface web.
- Monitorização de Sessões: O sistema deve registar e permitir a análise de todas as interações entre utilizadores e o avatar.
- Dashboard Analítico: Administradores devem ter acesso a métricas agregadas de utilização.
- Navegação em Ambiente 3D: Os utilizadores devem poder explorar livremente o espaço virtual.
- Autenticação e Autorização: O acesso ao *Backoffice* deve ser restrito a utilizadores devidamente autorizados.

3.4.2 Requisitos Não Funcionais

- Desempenho: A latência conversacional (desde a captura de voz até à resposta audível) deve ser inferior a 3 segundos. A taxa de frames no ambiente RV deve ser igual ou superior a 72 Frames Per Second (FPS) no Meta Quest 3.
- Usabilidade: A curva de aprendizagem deve ser reduzida, permitindo que novos utilizadores completem uma interação básica em menos de 2 minutos. O sistema deve oferecer suporte a múltiplos idiomas (inicialmente português e inglês).
- Compatibilidade: O sistema deve ser compatível com dispositivos RV como Meta Quest 2, 3 e Pro.
- Escalabilidade: O sistema deve ser capaz de lidar com um número crescente de utilizadores e consultas sem comprometer o desempenho.
- Segurança e Privacidade: As interações do Chatbot devem ser seguras, garantindo a proteção dos dados pessoais dos utilizadores.
- Feedback: O sistema deve permitir a coleta de *feedback* dos utilizadores sobre suas interações com o Chatbot. Esse *feedback* é crucial para identificar problemas, melhorar a precisão das respostas e adaptar o Chatbot às necessidades dos utilizadores ao longo do tempo.

3.5 Modelação do Sistema

O sistema 3D-Chatbot foi estruturado em dois subsistemas principais: o *Backoffice*, responsável pela gestão e monitorização do sistema, e o RV Frontend, que proporciona a experiência imersiva ao utilizador final. Esta separação reflete a arquitetura cliente-servidor descrita acima e facilita a compreensão das diferentes perspetivas de interação com o sistema.

3.5.1 Diagrama de Caso de Uso

O sistema de gestão administrativa identifica dois atores principais com diferentes níveis de privilégios:

- **Administrador:** Possui acesso completo ao sistema, herdando todas as permissões do Gestor da Empresa, com funcionalidades adicionais de gestão técnica.
- **Gestor da Empresa:** Focado em funcionalidades de análise e acompanhamento estratégico do desempenho da empresa.

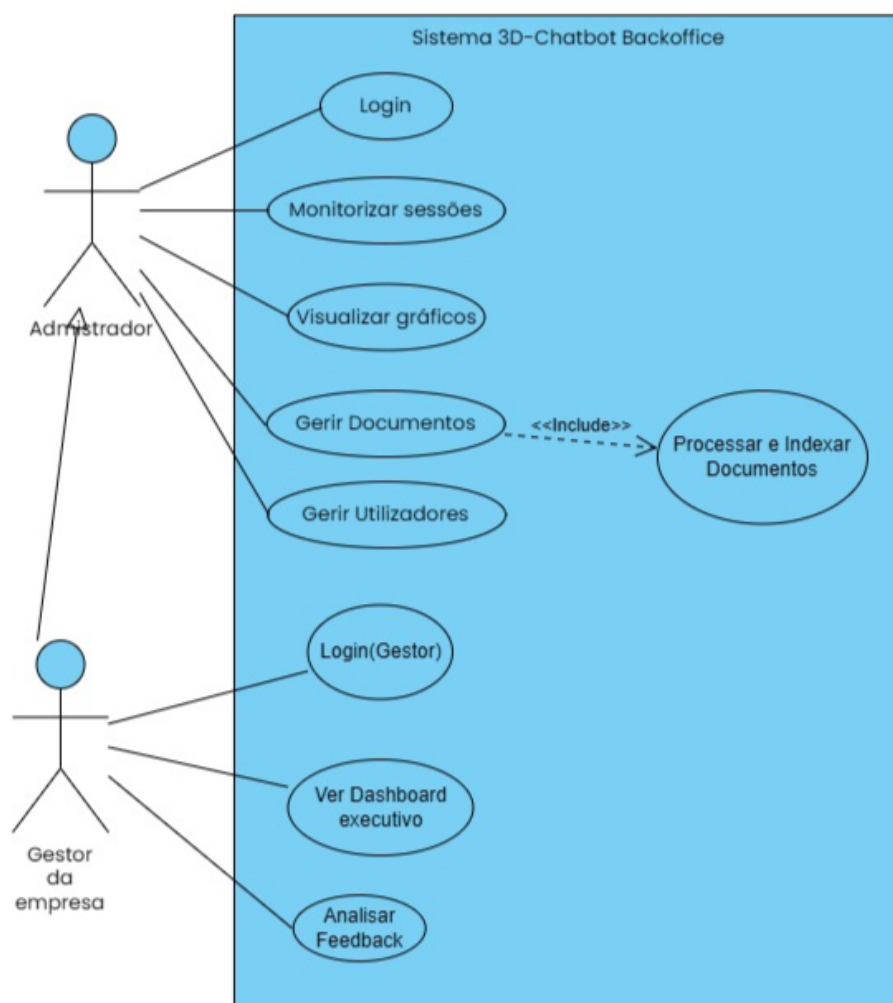


Figura 3.1: Backoffice

A interface imersiva foi concebida exclusivamente para o utilizador de RV, que interage com o sistema através de um ambiente tridimensional, utilizando interfaces naturais de voz e gestos.

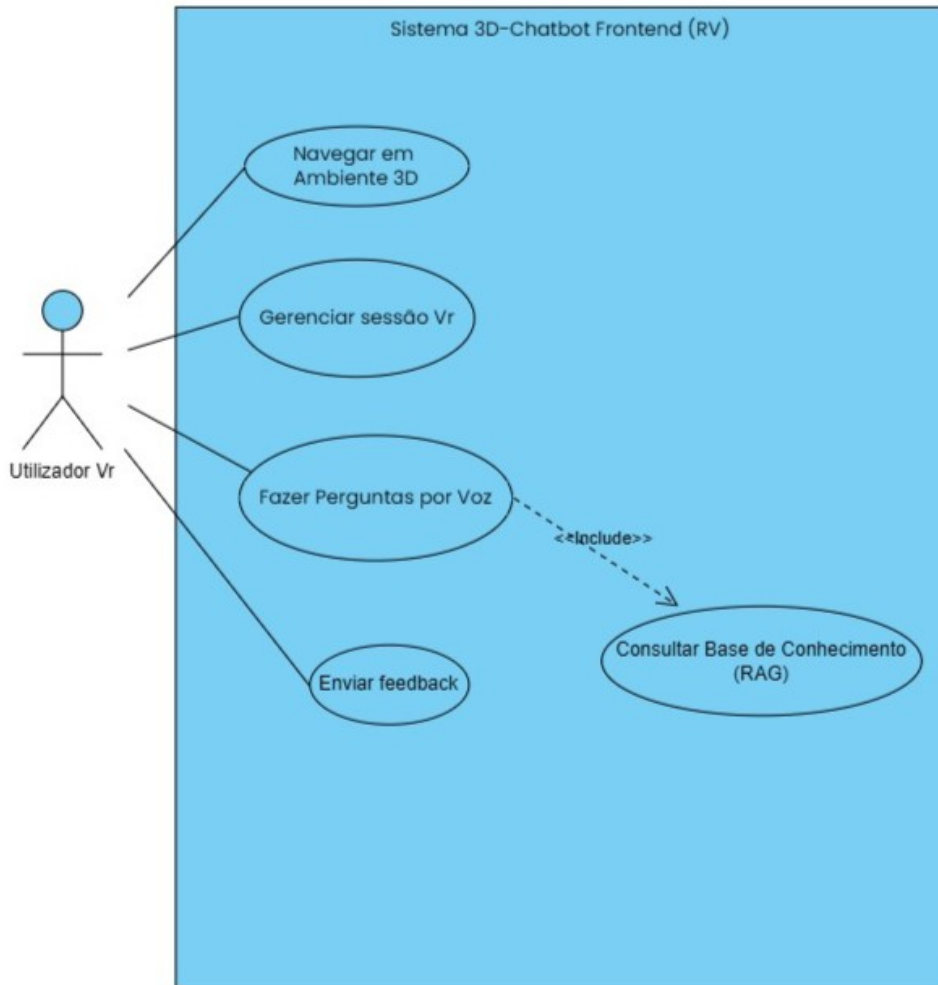


Figura 3.2: Frontend RV

3.5.2 Tarefas e Atores

As principais tarefas desempenhadas pelo sistema 3D-chatbot incluem:

Tarefas do Frontend RV

As principais tarefas desempenhadas pelo utilizador no ambiente de realidade virtual incluem:

- **Navegar em Ambiente 3D:** Permite ao utilizador explorar livremente o espaço virtual através de três modos de locomoção: teletransporte para saltos discretos que reduzem o enjoo, rotação ajustável (*snap turn* ou contínua) e movimento contínuo com controladores. Inclui também a interação com objetos *grabbable* e navegação por âncoras predefinidas para pontos de interesse específicos.
- **Fazer Perguntas por Voz:** Esta é a função principal do sistema, que consiste em o utilizador interagir conversacionalmente com o avatar através de reconhecimento de voz. O sistema captura o áudio, converte para texto via Convai, classifica a intenção da consulta e encaminha para o processamento adequado, seja conversação simples ou consulta à base de conhecimento especializada (RAG). O chatbot interpreta as perguntas, processa a linguagem natural e fornece respostas contextualizadas e fundamentadas.
- **Consultar Informações sobre Produtos:** Quando o utilizador solicita informações específicas, o sistema aciona automaticamente o processo RAG, que recupera documentos semanticamente relevantes do Pinecone, contextualiza com o histórico conversacional e gera respostas precisas via *GPT-4o-mini*, garantindo que as informações fornecidas são factualmente corretas e atualizadas.
- **Enviar Feedback:** O utilizador pode submeter uma avaliação da experiência através de uma interface dedicada no ambiente RV.

Tarefas do Backoffice

As principais tarefas realizadas no backoffice são:

- **Login:** Permite que o gestor e o administrador se autentiquem no sistema através de credenciais seguras (email/password). O sistema valida as credenciais, gera um

token JWT e estabelece a sessão do utilizador com as permissões apropriadas ao seu perfil, garantindo a segurança e o controlo de acesso ao *Backoffice*.

- **Monitorizar Sessões:** Apresenta visualização em tempo real e histórica de todas as sessões de utilizadores RV, permitindo ao administrador identificar sessões ativas, analisar durações, consultar históricos de conversação completos e detetar anomalias ou problemas de desempenho que possam afetar a experiência do utilizador.
- **Visualizar Gráficos:** Apresenta gráficos interativos no *Dashboard* principal com métricas agregadas sobre o desempenho do sistema, incluindo evolução temporal de sessões e mensagens, distribuição de tipos de consulta, taxas de sucesso de resposta e padrões de utilização. Estas visualizações ajudam a analisar o comportamento do utilizador e identificar tendências de uso.
- **Analisar Feedback:** Apresenta os comentários e avaliações enviados pelos utilizadores RV sobre a experiência com o chatbot, incluindo avaliações qualitativas, sugestões de melhoria e reportes de problemas. Permite ao gestor e administrador monitorizar a satisfação dos utilizadores e identificar oportunidades de otimização.
- **Gerir Documentos:** Permite ao administrador exercer controlo completo sobre a base de conhecimento do sistema: adicionar novos documentos para expandir o conhecimento disponível, remover conteúdo obsoleto ou irrelevante, visualizar estatísticas de indexação (número de documentos, *chunks* e vetores) e acionar o processo de reindexação completa quando necessário. Este processo automatizado segmenta documentos em *chunks*, gera *embeddings* vetoriais usando *text-embedding-3-small* e armazena os vetores no Pinecone, garantindo que as respostas fornecidas pelo chatbot são sempre precisas e atualizadas.
- **Gerir Utilizadores:** Permite ao administrador realizar a administração completa de contas de utilizadores do *Backoffice*, incluindo criar novos utilizadores, atribuir perfis adequados, modificar permissões de acesso, desativar contas quando necessário e alterar passwords, garantindo a gestão adequada do acesso ao sistema.

- **Ver Dashboard Executivo:** Apresenta ao gestor da empresa uma visualização consolidada de métricas estratégicas, incluindo volume total de interações, tendências de utilização ao longo do tempo, tópicos mais consultados pelos utilizadores e indicadores de satisfação. O foco está em Key Performance Indicator (KPI) de negócio em vez de métricas técnicas detalhadas, facilitando a tomada de decisões estratégicas.

Principais Atores

Os principais atores do sistema 3D-Chatbot são:

- **Utilizador RV:** Indivíduos que interagem com o sistema através do ambiente de realidade virtual imersivo. Estes utilizadores podem ser clientes potenciais explorando produtos em ambientes de e-commerce 3D, visitantes de museus ou exposições virtuais, ou utilizadores de serviços que necessitam de assistência conversacional em contextos imersivos. O seu objetivo principal é obter informações precisas de forma natural e envolvente.
- **Gestor da Empresa:** Responsável por monitorizar o desempenho estratégico do chatbot através da visualização de estatísticas consolidadas, análise dos conteúdos mais frequentemente consultados pelos utilizadores e avaliação do *feedback* recebido. O gestor utiliza estas informações para tomar decisões estratégicas sobre a evolução do sistema e identificar oportunidades de melhoria do serviço.
- **Administrador:** Pessoal técnico responsável pela manutenção operacional completa do sistema, incluindo gestão da base de conhecimento, monitorização técnica de sessões e desempenho, administração de utilizadores do *Backoffice* e garantia da qualidade e atualização do conteúdo disponibilizado pelo chatbot. O administrador herda todas as funcionalidades do Gestor da Empresa, podendo executar tanto tarefas estratégicas quanto operacionais e técnicas.

3.5.3 Modelação de Dados

A persistência e a gestão eficiente das informações no sistema 3D-Chatbot requerem uma estrutura de dados bem definida que suporte operações de leitura rápidas e gravações frequentes. O MongoDB foi selecionado como a solução de base de dados devido à sua flexibilidade esquemática e capacidade de armazenar documentos JavaScript Object Notation (JSON) complexos, recursos essenciais para um sistema conversacional que gera dados semiestruturados e em evolução.

The screenshot shows the MongoDB Atlas interface for a database named 'chatbot3d_db'. It displays five collections with their respective schemas:

Collection	Field	Type	Cardinality
chat_messages	_id	pk	old * (11.1)
	confidence	multi	*
	content	str	*
	language	str	*
	message_id	str	*
	metadata	doc	*
	sources	arr	*
	user_id	old	*
	response_time	multi	*
	role	str	* (15.1)
	session_id	str	* (12.1, 13.1)
	timestamp	date	* (13.2, 14.1)
	token_count	int32	*
	user_feedback_binary	bool	*
DocumentIndex	_id	pk	old * (11.1)
	category	multi	(110.1, 114.2)
	chunk_count	int32	*
	company_id	str	* (12.1, 111.1, 112.1, 113.1, 114.1)
	company_name	str	*
	document_id	str	*
	document_type	str	* (115.1, 111.2, 113.3, 115.1, 116.2)
	error_message	str	*
	file_hash	str	* (14.1)
	file_name	str	*
	file_path	str	* (13.1)
	file_size	int32	*
	indexed_at	date	* (17.1)
	is_active	bool	* (19.1, 113.2, 116.1)
	language	str	*
	last_updated	date	* (18.1)
	metadata	doc	*
	processing_time	double	*
status	str	* (16.1, 112.2, 115.2)	
tags	arr	*	
ChatSession	_id	pk	old * (11.1)
	context	doc	*
	created_at	date	* (15.1)
	duration_seconds	int32	*
	last_activity	date	* (16.1)
	message_count	int32	*
	session_id	str	* (12.1)
	settings	doc	*
	status	str	* (14.1)
	tags	arr	*
user_agent	str	*	
User	_id	pk	old * (11.1)
	created_at	date	*
	email	str	* (12.1)
	hashed_password	str	*
	is_active	bool	*
	last_login	date	*
	role	str	*
	username	str	* (13.1)
	updated_at	date	*

Figura 3.3: Modelo Mongo DB

A Figura 3.3 apresenta o modelo de dados completo implementado no MongoDB, organizado em cinco coleções principais que refletem os domínios funcionais do sistema:

1. Users: Armazena as credenciais e metadados dos utilizadores do Backoffice, suportando dois níveis de privilégios (Administrador e Gestor da Empresa).
2. ChatSession: Regista cada sessão de interação entre um utilizador RV e o chatbot, funcionando como o principal repositório para análise de comportamento e métricas de utilização.
3. Chat_Messages: Armazena o histórico completo de todas as interações conversacionais, com granularidade ao nível de cada pergunta e resposta individual.
4. DocumentIndex: Gerencia a base de conhecimento indexada no Pinecone, mantendo metadados essenciais para sincronização e auditoria.

Capítulo 4

Implementação

Este capítulo descreve a implementação do sistema 3D-Chatbot, com ênfase nos principais componentes técnicos e soluções desenvolvidas para integrar a inteligência artificial conversacional em ambientes de realidade virtual imersiva. O trabalho de implementação foi organizado em quatro fases principais: aquisição e processamento de dados, desenvolvimento de backend de IA, desenvolvimento de experiência de RV e integração completa do sistema.

4.1 Aquisição e Processamento de Dados

4.1.1 Pipeline de Aquisição e Transformação de Conteúdo Web

A fundação do sistema de atendimento ao cliente inteligente começou com o desenvolvimento de um sistema avançado de web crawling assíncrono capaz de adquirir dados abrangentes da presença web e catálogo de produtos de múltiplos domínios. Esta fase transforma conteúdo não estruturado dos websites em uma base de conhecimento estruturada e pronta para IA, capaz de suportar processamento sofisticado de consultas.

A Figura 4.1 ilustra a arquitetura completa do crawler multi-site utilizado para esta fase de aquisição de dados.

O sistema implementa uma arquitetura assíncrona robusta usando bibliotecas Python

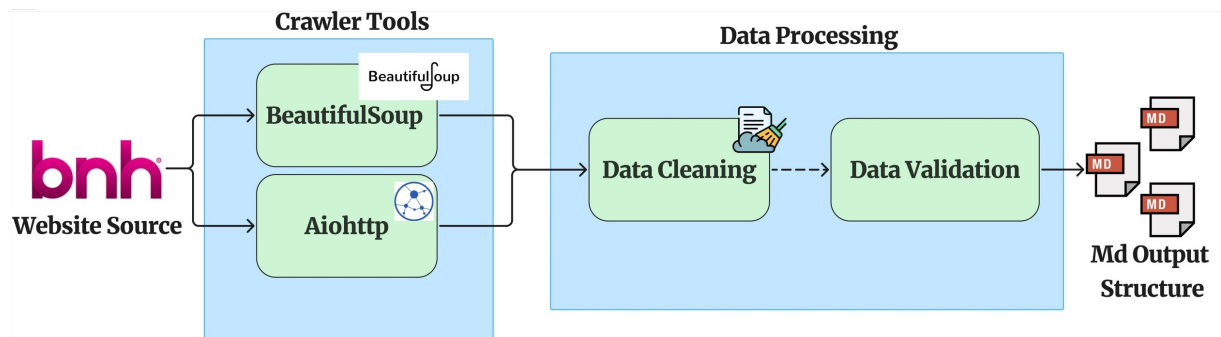


Figura 4.1: Arquitetura do Web Crawler - Aquisição de dados

especializadas: *Beautiful Soup* para análise e análise complexa de HTML, *aiohttp* para processamento HTTP assíncrono de alto desempenho com suporte para solicitações simultâneas e *asyncio* para orquestração eficiente de operações I/O. Esta arquitetura permite escalabilidade horizontal mantendo conformidade estrita com especificações *robots.txt* de cada domínio alvo.

A implementação inclui um sistema de fila de priorização multinível que organiza os Uniform Resource Locators (URLs) a serem processados de acordo com critérios estratégicos. O algoritmo de priorização classifica os URLs em três níveis hierárquicos: Alto (páginas de produtos individuais e categorias principais), Médio (subcategorias, páginas de serviços e perguntas frequentes) e Baixo (páginas institucionais, recursos estáticos e conteúdo auxiliar). O sistema implementa o controlo de profundidade com um limite de 5 níveis para evitar rastreamento excessivo em estruturas de navegação profundas, usando pesquisa modificada em largura (Breadth-first search (BFS)) para priorizar conteúdo comercialmente relevante e maximizar a densidade informacional da base de conhecimento resultante.

Um mecanismo de controlo de concorrência baseado em semáforos limita as solicitações simultâneas, evitando a sobrecarga do servidor de destino e mantendo um alto desempenho operacional. O sistema implementa limitação de taxa configurável (2,5 segundos entre solicitações) com *jitter* aleatório para simular o comportamento humano e evitar a deteção como um *bot* automatizado, bem como *backoff* exponencial (quatro tentativas com atrasos

progressivos) para repetir as solicitações que falham devido a tempos limite ou erros temporários de rede.

O conteúdo bruto extraído passa por uma limpeza e normalização extensivas através de um processo de três etapas:

1. Remoção de marcações HTML: excluir tags HTML preservando elementos estruturais importantes;
2. Padronização de conteúdo: atenção especial às especificações técnicas e formatação consistente
3. Validação e tratamento de erros: Identificação e tratamento de informações incompletas ou inconsistentes

O *crawler* implementa conformidade nativa com o GDPR através de três mecanismos principais, conforme ilustrado na Tabela 4.1:

Tabela 4.1: Mecanismos de Conformidade GDPR Implementados

Mecanismo	Implementação
Anonimização Automática	Expressões regulares substituem dados pessoais por <i>tokens</i> : emails, telefones, Número de Identificação Fiscal (NIF), International Bank Account Numbers (IBANs) e códigos postais → <code>[*_REDACTED]</code> , preservando o contexto.
Registo de Atividades (Art. 30)	Base legal (Art. 6(1)(f) — interesse legítimo), finalidade (base de conhecimento para chatbot), dados tratados (informações públicas) e medidas de segurança: Secure Sockets Layer (SSL)/Transport Layer Security (TLS), <i>rate limiting</i> , <i>logs</i> auditáveis.
Filtros de Exclusão	<i>Blacklist</i> de padrões de URLs sensíveis (<code>/user/</code> , <code>/contact/</code> , <code>/forum/</code> , <code>/login/</code>) e respeito a <code>robots.txt</code> .

Os mecanismos de deteção de conteúdo duplicado garantem a eficiência da base de conhecimento, mantendo uma cobertura abrangente. O resultado final consiste em mais de 300 ficheiros Markdown estruturados, organizados em categorias lógicas (informações

sobre produtos, categorias e empresas), cada um incluindo metadados abrangentes, seções de conteúdo estruturado e links de referência cruzada que permitem a descoberta sofisticada de relações durante o processamento de consultas.

4.2 Desenvolvimento do Backend de IA

4.2.1 Pipeline de Processamento e Embedding de Documentos

O pipeline de processamento de documentos transforma conteúdo estruturado em representações vetoriais de alta dimensão adequadas para operações de pesquisa semântica e recuperação inteligente. A Figura 4.2 ilustra a sequência de transformações do documento estruturado até a representação vetorial final.

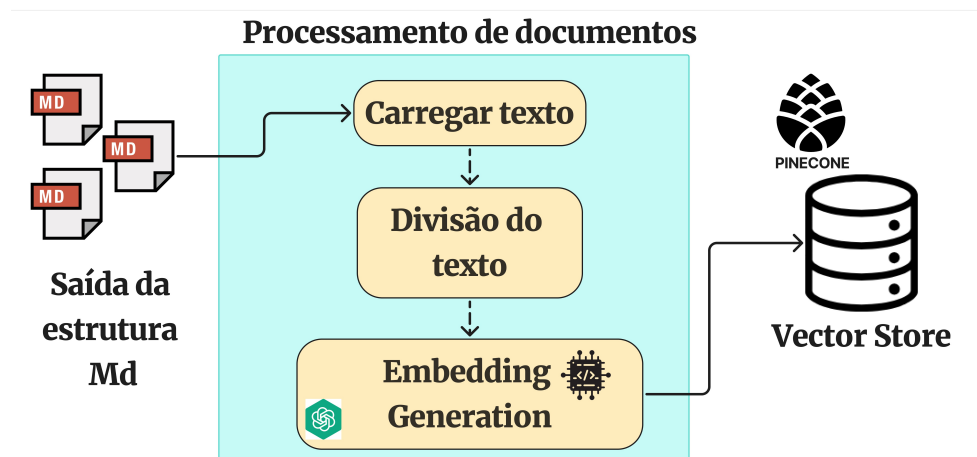


Figura 4.2: Pipeline de Processamento de Documentos

O processo começa com a ingestão de documentos pré-processados da fase anterior, mantendo os metadados dos documentos e as relações hierárquicas para preservar a integridade contextual durante toda a transformação. Os documentos passam por uma segmentação inteligente usando o *RecursiveCharacterTextSplitter* da LangChain, um componente crítico configurado para alcançar um equilíbrio ideal entre a preservação do contexto e a eficiência do processamento. O uso de blocos de 1.000 caracteres com sobreposição de 100 caracteres garante que informações importantes não sejam perdidas nas fronteiras dos

segmentos, permitindo operações de recuperação precisas que mantêm a coesão semântica mesmo em documentos fragmentados.

Para gerar representações vetoriais, o sistema usa o modelo *text-embedding-3-small* da OpenAI, que cria vetores de 1.536 dimensões para cada fragmento de texto. Esse modelo de incorporação de última geração oferece compreensão semântica superior em comparação com alternativas anteriores, permitindo uma correspondência de similaridade mais precisa entre as consultas dos utilizadores e o conteúdo armazenado, bem como uma recuperação contextual mais eficaz. A escolha do Pinecone como solução de base de dados vetorial forneceu uma infraestrutura sem servidor e com dimensionamento automático, otimizada especificamente para operações de pesquisa de similaridade. A configuração da base de dados usa medidas de similaridade cosseno para correspondência semântica ideal, mantendo consistentemente os tempos de resposta às consultas abaixo de 100 milissegundos, um requisito crítico para experiências de conversação fluidas em tempo real.

4.2.2 Sistema RAG com Contextualização e Geração de Respostas

A implementação do sistema RAG baseia-se em um servidor FastAPI que coordena todas as operações críticas: processamento de consultas, recuperação contextualizada de documentos e geração de respostas coerentes, mantendo elevados padrões de desempenho e confiabilidade sob cargas variáveis. A Figura 4.3 ilustra a arquitetura completa do servidor FastAPI e a orquestração dos diferentes componentes.

O servidor implementa padrões *sync/await* em todo o sistema para maximizar a capacidade de processar solicitações simultâneas sem bloqueio, permitindo que vários utilizadores interajam com o sistema ao mesmo tempo. A organização das rotas segue os princípios REST com *endpoints* dedicados para interações de chat, gestão de sessões, operações de documentos e monitorização da integridade do sistema. Os componentes de *middleware* lidam com a partilha de recursos entre origens CORS, registo estruturado de solicitações e erros, limitação de taxa para proteção contra abusos e implementação de

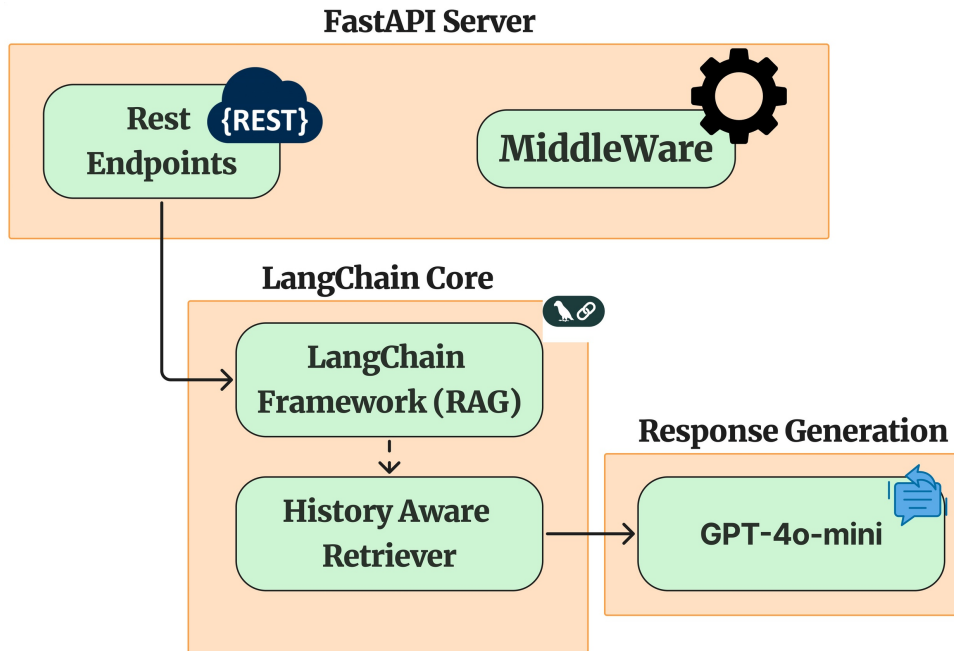


Figura 4.3: Arquitetura FastAPI

segurança apropriada para o ambiente de produção.

A estrutura LangChain fornece a implementação central do RAG, orquestrando o complexo fluxo de trabalho de recuperação e geração por meio de componentes altamente integrados. O sistema utiliza um recuperador sensível ao histórico para a recuperação de documentos sensível ao contexto, que considera explicitamente o histórico de conversas ao determinar quais informações são mais relevantes para responder à consulta atual, evitando respostas descontextualizadas. O *GPT-4o-mini* funciona como o motor de geração de respostas, estrategicamente selecionado pelo seu equilíbrio ideal entre qualidade de resposta, velocidade de processamento e relação custo-benefício para aplicações de atendimento ao cliente em produção. *Prompts* personalizados do sistema orientam o modelo a fornecer respostas úteis, precisas e profissionalmente adequadas, mantendo o fluxo natural da conversa e o envolvimento contínuo do cliente, evitando respostas genéricas ou mecanicamente estruturadas.

4.3 Desenvolvimento da Experiência em RV e Processamento de Voz

4.3.1 Ambiente Virtual e Integração Multimodal

O ambiente de RV foi projetado e implementado no Unity 2022.3 Long Term Support (LTS) usando uma arquitetura moderna baseada no OpenXR, o padrão aberto da indústria que oferece compatibilidade entre plataformas com vários headsets de RV, incluindo *Meta Quest*, *HTC Vive*, *Valve Index* e *Windows Mixed Reality*, eliminando a necessidade de implementações específicas para cada dispositivo e garantindo a portabilidade futura do aplicativo.



Figura 4.4: Meta Quest 3

A configuração fundamental do projeto Unity incorpora quatro componentes essenciais de renderização e interação:

1. Plugin OpenXR que fornece a camada de abstração de hardware, permitindo a comunicação direta com os recursos nativos do *headset* por meio de uma interface padronizada;
2. Universal Render Pipeline (URP) selecionado pelo seu desempenho otimizado em dispositivos móveis RV, como o Meta Quest, mantendo alta qualidade visual através de técnicas como sombreamento gráfico e pós-processamento configurável;

3. Gestão de *plug-ins* Extended Reality (XR), que centraliza a configuração e gestão de todos os subsistemas XR, incluindo entrada, exibição e rastreamento, fornecendo uma interface unificada;
4. *Kit* de ferramentas de interação XR que implementa um sistema completo de interação RV, incluindo interatores de raios para apontar à distância, interatores diretos para manipular objetos próximos e provedores de locomoção para movimentos naturais no espaço virtual.

A hierarquia da cena Unity é estruturada em torno de componentes fundamentais, começando com *Static Room*, que define o espaço físico estático da sala de exposições virtual, incluindo a geometria das paredes, piso, teto e elementos arquitetônicos fixos, todos marcados como estáticos para permitir otimizações de renderização, como *batching estático* e *occlusion culling*, reduzindo drasticamente as chamadas de desenho e melhorando a taxa de quadros. O XR Origin serve como a raiz de toda a hierarquia XR, funcionando como um ponto de referência para rastrear a posição e a rotação do *headset* e dos controladores no espaço virtual, encapsulando a *Main Camera*, que renderiza a visão estereoscópico para ambos os olhos do utilizador com planos de recorte próximos/-distantes otimizados e parâmetros de campo de visão configurados para conforto visual e desempenho. O Controlador Esquerdo e o Controlador Direito, que representam os controladores físicos no espaço virtual por meio de modelos 3D visualmente consistentes e detecção precisa de colisões, permitindo uma interação natural com objetos virtuais. Os controladores implementam três provedores de locomoção complementares: *Snap Turn Provider*, que permite rotação instantânea em incrementos fixos (45 graus) ativados por um *joystick* horizontal, reduzindo a sensação de enjoo em usuários sensíveis por meio de rotação discreta em vez de contínua; *Continuous Turn Provider*, que oferece rotação suave e contínua com controlo de velocidade para utilizadores experientes que preferem um controlo mais natural e fluido; e *Dynamic Move Provider*, que implementa movimento translacional no espaço virtual, respeitando a física básica com aceleração e desaceleração suaves, evitando mudanças abruptas de velocidade que causariam desconforto.

A iluminação e a navegação do ambiente incorporam Luz Direcional configurada como a luz solar principal com intensidade, cor e ângulo calibrados para criar uma atmosfera profissional e acolhedora, garantindo ao mesmo tempo a visibilidade adequada de todos os elementos importantes da cena, projetando sombras em tempo real através de cascatas de sombras otimizadas que equilibram a qualidade visual e o custo computacional. O sistema de teletransporte implementa dois mecanismos complementares: Áreas de Teletransporte que definem grandes regiões do piso onde os utilizadores podem se teletransportar livremente, proporcionando liberdade de exploração dentro de zonas seguras predefinidas, visualizadas através de uma trajetória em arco que mostra a parábola de destino e um retículo de aterragem que valida a posição antes de confirmar o teletransporte, e Âncoras de Teletransporte que marcam pontos de interesse específicos estrategicamente posicionados (por exemplo, em frente a expositores de produtos ou áreas de demonstração), permitindo uma navegação rápida para locais importantes através do *snap-to-position*, que garante uma orientação ideal para visualizar conteúdos específicos. Os objetos *grab-able* implementam um sofisticado sistema de manipulação de objetos virtuais através de componentes XR *Grab Interactable* que detetam a proximidade do utilizador e o gesto de agarrar, aplicam física realista durante a manipulação, incluindo peso, impulso e colisões com outros objetos, e permitem que os objetos sejam soltos, os quais continuam a sua trajetória com base na velocidade e rotação no momento da libertação, criando uma sensação tátil convincente. Os *soquetes* complementam os objetos agarráveis, definindo pontos de encaixe onde objetos específicos podem ser colocados com alinhamento automático (por exemplo, produto retornando à vitrine, ferramenta colocada no suporte), fornecendo *feedback* tátil e visual quando um objeto se aproxima de um *soquetes* compatível e travando o objeto no lugar quando colocado, validando a conclusão de tarefas ou quebra-cabeças de interação. A integração do avatar conversacional é implementada através do pacote *Convai Essentials*, que fornece uma ponte entre o Unity e os serviços em nuvem da Convai, incluindo componentes para *streaming* de áudio bidirecional, sincronização do estado conversacional e gestão de sessões persistentes.

O modelo de avatar foi criado utilizando a plataforma **Ready Player Me (RPM)**,

um sistema de geração de avatares 3D que permite a criação rápida de personagens humanoides personalizáveis e otimizados para aplicações interativas. O avatar RPM consiste numa malha humanoide totalmente equipada com um esqueleto humanoide padrão compatível com o sistema de animação Unity (especificação *Humanoid Rig*), permitindo o redirecionamento de animações genéricas, texturas Physically Based Rendering (PBR) de alta qualidade que proporcionam uma aparência realista sob diferentes condições de iluminação, e grupos Level of Detail (LOD) que otimizam a densidade da malha com base na distância do utilizador para manter o desempenho em tempo real. O *Animator Controller* orquestra uma máquina de estados complexa que gerência transições suaves entre animações de inatividade (*idle*), conversa (*talking*), escuta (*listening*) e gestos expressivos, utiliza árvores de mistura (*blend trees*) para interpolar entre variações de animação com base em parâmetros como intensidade emocional ou velocidade da fala, e sincroniza perfeitamente com o sistema de sincronização labial (*lip-sync*) fornecido pela Convai. Este sistema analisa fonemas de fala sintetizados em tempo real, aplicando *blend shapes* específicas aos vértices da malha facial para criar movimentos labiais convincentes e naturais que acompanham o áudio gerado pelo sistema de TTS.

4.3.2 Processamento de Voz Integrado e Canal de Encaminhamento de Consultas

A plataforma Convai fornece processamento de voz em tempo real otimizado para contextos conversacionais, com suporte para variações de linguagem natural e padrões de fala informais comuns em interações de atendimento ao cliente. O sistema lida com vários idiomas e sotaques, mantendo uma precisão de reconhecimento consistente. O pipeline completo de processamento de voz inclui redução de ruído, cancelamento de eco e detecção de atividade de voz otimizados para ambientes de RV. O pipeline segue uma sequência de quatro etapas:

1. Conversão de voz para texto em tempo real: processamento com suporte para vários idiomas e sotaques

2. Pipeline de processamento de voz: redução de ruído, cancelamento de eco e aplicação de detecção de atividade de voz
3. Roteamento inteligente de consultas: análise de intenção e determinação da estratégia de resposta adequada
4. Síntese de texto para voz com integração de sincronização labial: respostas naturais através de avatares

Essas melhorias técnicas garantem um reconhecimento de voz confiável, mesmo em condições de áudio desafiadoras com múltiplas fontes de ruído, mantendo o fluxo natural da conversa com latência mínima. O sistema completa ciclos completos de resposta de voz (captura, processamento, recuperação de informações e síntese) normalmente em menos de 3 segundos, um tempo crítico para manter o ritmo psicologicamente natural da conversa humana. Esse nível de desempenho permite que o utilizador experimente uma conversa autêntica, em vez de atrasos que criam uma sensação robótica ou artificial de comunicação.

O sistema implementa uma classificação sofisticada de intenções para determinar estratégias de resposta adequadas para diferentes categorias de consultas, evitando abordagens genéricas. Consultas conversacionais simples (saudações, perguntas vagas) são tratadas diretamente pela IA conversacional da Convai para respostas imediatas e naturais, enquanto perguntas específicas sobre produtos (especificações, preços, disponibilidade) são encaminhadas de forma inteligente para o backend RAG especializado para recuperação de informações detalhadas e factualmente precisas. Os recursos de conversão de texto em fala geram uma fala natural e emocionalmente apropriada com características de voz totalmente configuráveis, sincronizadas com precisão com as animações do avatar para fornecer respostas visualmente consistentes por meio do assistente virtual, criando experiências conversacionais que mantêm o envolvimento do cliente mesmo durante interações prolongadas.

4.4 Integração Completa e Fluxo de Interação

4.4.1 Arquitetura Integrada de Seis Passos

A fase final de integração combina todos os componentes do sistema numa experiência de utilizador coerente que combina perfeitamente a interação imersiva de RV com sofisticados sistemas de conhecimento baseados em IA. Esta integração aborda o desafio crítico de manter o fluxo natural da conversa enquanto acede a fontes de conhecimento externas em tempo real. Uma interação completa com o cliente segue um processo bem definido de seis etapas:

1. O utilizador entra no ambiente de realidade virtual usando óculos RV e entrando na sala de exposições virtual.
2. O assistente virtual oferece uma saudação natural usando a IA conversacional da Convai para estabelecer uma relação e compreender as necessidades do utilizador.
3. Quando os utilizadores fazem perguntas específicas sobre produtos, o sistema inicia o processamento natural da conversa.
4. O sistema faz uma transição perfeita para aceder ao backend do RAG para obter informações precisas e detalhadas.
5. A recuperação de informações acessa conhecimentos especializados, enquanto a síntese de respostas combina detalhes dos produtos recuperados com o contexto da conversa.
6. A resposta fornecida é coerente e útil, proporcionando uma experiência de consulta natural.

A integração mantém a continuidade do contexto da conversa ao longo dessas transições de múltiplas fases, assegurando que as informações especializadas sobre o produto

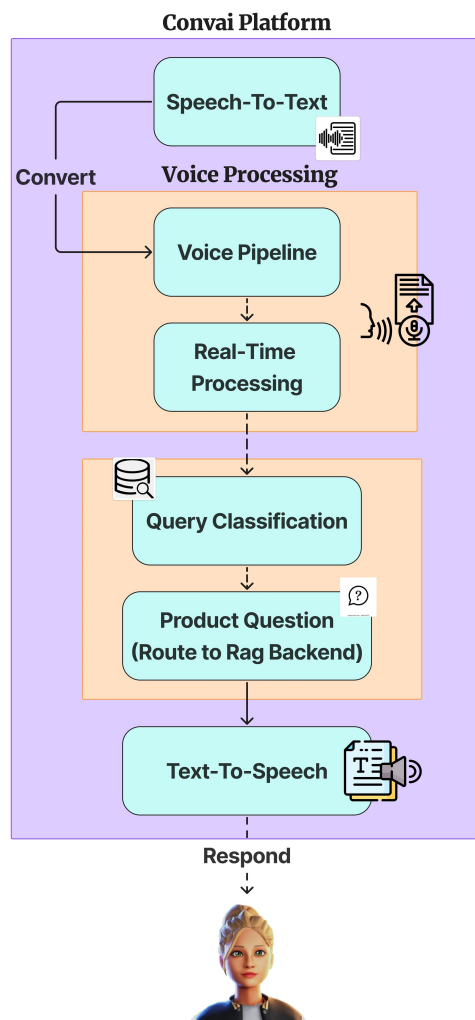


Figura 4.5: Pipeline de Processamento de Voz

sejam apresentadas naturalmente na conversa em andamento, em vez de parecerem fragmentadas ou interrompidas. A síntese da resposta combina de forma inteligente os detalhes do produto recuperados com o contexto completo da conversa para fornecer respostas coerentes e úteis que parecem surgir naturalmente do diálogo, em vez de serem geradas mecanicamente, preservando uma sensação de autenticidade na interação.

4.4.2 Otimizações de Desempenho e Arquitetura de Comunicação

O agrupamento de ligações e o processamento assíncrono minimizam a sobrecarga da rede entre clientes de RV, serviços Convai remotos e componentes backend personalizados, que são operações críticas para manter a capacidade de resposta. Estas otimizações garantem que todo o sistema mantenha um desempenho responsivo, mesmo durante operações complexas de recuperação de informações e processamento de linguagem natural que normalmente exigiriam altas latências. A arquitetura do sistema foi projetada para suportar um fluxo de conversação verdadeiramente contínuo, permitindo que os utilizadores façam perguntas naturais de acompanhamento e mantenham padrões de diálogo naturais durante toda a sessão de consulta, sem a necessidade de reiniciar ou reconfigurar.

4.5 Sistema de Gestão e Monitorização (Backoffice)

4.5.1 Arquitetura e Fundamentos do Sistema de Gestão

Durante o desenvolvimento do sistema 3D-Chatbot, foi identificada uma necessidade crítica de implementar uma plataforma de gestão e monitorização que permitisse aos administradores acompanhar o desempenho do chatbot em tempo real, analisar os padrões de utilização dos utilizadores e gerir eficientemente os recursos computacionais e de conhecimento do sistema. Esta necessidade surgiu da constatação de que sistemas conversacionais complexos baseados em IA requerem supervisão contínua para garantir a qualidade do serviço, identificar problemas precocemente e otimizar continuamente a experiência do utilizador final. O sistema de *Backoffice* foi desenvolvido como uma aplicação web dedicada usando uma arquitetura modular baseada em modelos *Jinja2*, permitindo fácil manutenção através da separação clara de responsabilidades e escalabilidade futura através de componentes bem encapsulados. A arquitetura implementa rigorosamente os princípios de separação de preocupações, com componentes claramente definidos para três camadas

distintas: a camada de apresentação responsável pela interface visual e interação do utilizador, a camada de lógica de negócios que processa e transforma dados para visualização adequada e a camada de acesso a dados que se comunica com o backend FastAPI por meio de uma APIs REST bem definida. A estrutura do modelo segue uma organização hierárquica com um modelo base (*base.html*) que encapsula todos os elementos comuns da interface, incluindo uma barra lateral de navegação com um menu lateral responsivo que fornece acesso a todos os recursos administrativos, um sistema de autenticação completo com gerenciamento seguro de *tokens* JWT e controle de acesso granular baseado em funções, componentes visuais consistentes usando *Bootstrap 5* para um sistema de grade responsivo e *Bootstrap Icons* para iconografia uniforme, uma biblioteca de funções JavaScript partilhadas que encapsula chamadas de APIs com tratamento de erros consistente e gestão do estado de carregamento com *feedback* visual por meio de sobreposições durante operações assíncronas. Cada página específica do sistema (*dashboard.html* para o painel principal com métricas agregadas, *sessions.html* para gestão detalhada das sessões dos utilizadores, *documents.html* para controlo da base de conhecimento indexada, *analytics.html* para análise avançada de padrões e *users.html* (ver B.2) para gestão de utilizadores e permissões) amplia o modelo base através do sistema de blocos *Jinja2*, permitindo uma reutilização massiva de código com elementos comuns definidos uma única vez e uma manutenção simplificada onde as alterações são feitas. HTML para gestão de utilizadores e permissões) amplia o modelo base através do sistema de blocos *Jinja2*, permitindo a reutilização massiva de código com elementos comuns definidos uma única vez, manutenção simplificada, onde as alterações na estrutura base se propagam automaticamente para todas as páginas derivadas, consistência visual absoluta, garantindo uma interface uniforme e profissional em toda a aplicação, e desenvolvimento ágil, permitindo a criação rápida de novas páginas administrativas sem duplicação de código estrutural ou lógica comum.

4.5.2 Dashboard Principal e Visualização de Métricas

O painel principal é o ponto central para monitorizar o sistema, fornecendo uma visão consolidada e atualizada do desempenho por meio de métricas-chave organizadas em cartões interativos com design material e animações suaves. As métricas estatísticas implementadas incluem o total acumulado de sessões de conversação, indicando o volume histórico de utilização, o total agregado de mensagens processadas, refletindo a intensidade da interação, o contador de sessões ativas nas últimas 24 horas, servindo como um indicador da atividade recente e do envolvimento atual, e a média calculada de mensagens por sessão, fornecendo métricas críticas sobre a profundidade do envolvimento e a satisfação do utilizador com o sistema. As visualizações gráficas implementam um gráfico de atividade diária com uma configuração otimizada de gráfico de linhas que mostra a evolução temporal das sessões e mensagens nos últimos 7 dias com suavização de curvas por meio de um parâmetro de tensão definido como 0,4, permitindo a identificação visual rápida de tendências, picos de uso e anomalias de desempenho; e um gráfico de distribuição de mensagens implementado como um gráfico, mostrando a proporção visual entre as mensagens originadas pelos utilizadores e as respostas geradas pelo sistema de IA, fornecendo uma indicação imediata da taxa de resposta e do equilíbrio conversacional.

A análise de conteúdo apresenta uma tabela classificada dinamicamente das consultas mais frequentes enviadas ao sistema pelos utilizadores, cada entrada incluindo o texto completo da pergunta, a contagem absoluta de ocorrências e a percentagem relativa calculada em relação ao volume total, permitindo aos administradores identificar rapidamente os tópicos de maior interesse, perguntas recorrentes que podem indicar lacunas na base de conhecimento ou oportunidades de melhoria na documentação e padrões de utilização emergentes que informam decisões estratégicas sobre funcionalidade ou expansão de conteúdo. A implementação técnica utiliza *Chart.js* com configuração otimizada para desempenho e capacidade de resposta, incluindo escalas configuradas para começar em zero, garantindo uma representação visual sem distorções, tensão da curva ajustada para uma suavização adequada sem comprometer a fidelidade aos dados reais e um sistema de

cores semântico consistente, onde o azul representa sessões, o verde representa mensagens processadas com sucesso, o amarelo destaca atividades recentes que requerem atenção e o azul claro informativo apresenta métricas calculadas derivadas. O painel implementa atualizações sob demanda por meio de um botão dedicado que aciona uma atualização completa de todas as métricas e gráficos, com *feedback* visual durante o processo de carregamento por meio de uma sobreposição semitransparente com um *spinner* animado e um sistema de alerta para notificar o sucesso ou a falha da operação. O painel implementa atualizações sob demanda por meio de um botão dedicado que aciona uma atualização completa de todas as métricas e gráficos, com *feedback* visual durante o processo de carregamento por meio de uma sobreposição semitransparente com um *spinner* animado e um sistema de alerta para notificar o sucesso ou a falha da operação de atualização.

4.5.3 Gestão de Sessões e Monitorização de Interações

O módulo de gestão de sessões fornece recursos essenciais para visualizar e monitorizar todas as interações históricas e atuais do utilizador com o sistema conversacional, implementando uma tabela responsivo com informações estruturadas, incluindo um identificador de sessão único apresentado em formato abreviado para facilitar a leitura, mantendo a capacidade de identificação inequívoca, identificador de utilizador associado que permite o rastreamento do comportamento individual em várias sessões, carimbo de data/hora preciso da criação da sessão que documenta a hora de início da interação, carimbo de data/hora da última atividade registada que identifica sessões potencialmente abandonadas ou inativas, contador em tempo real de mensagens trocadas indicando a profundidade e a duração da conversa e status da sessão atual exibido por meio de emblemas coloridos, fornecendo identificação visual imediata do status. A implementação técnica utiliza carregamento dinâmico via APIs com um *endpoint* dedicado que retorna até 50 sessões mais recentes ordenadas por carimbo de data/hora descendente, preenchimento assíncrono da tabela HTML através de literais de modelo JavaScript que constroem dinamicamente as

linhas, preservando o desempenho mesmo com altos volumes de dados, formatação localizada de carimbos de data/hora utilizando métodos JavaScript nativos configurados para a localidade portuguesa, garantindo uma apresentação familiar aos administradores.

As funcionalidades operacionais incluem a visualização detalhada de sessões individuais (atualmente implementada como um espaço reservado preparado para o desenvolvimento futuro de um modal ou página dedicada que exiba o histórico completo de mensagens e metadados da sessão), recursos de classificação que permitem a reorganização por qualquer coluna para facilitar a identificação de padrões específicos e um sistema de filtragem projetado para permitir a pesquisa por identificador, período de tempo ou status da atividade. O sistema mantém um histórico completo de todas as sessões para análise longitudinal de padrões de uso, identificação de usuários avançados que podem fornecer *feedback* valioso, detecção de sessões com comportamento suspeito ou padrões de erro e cálculo de métricas agregadas de engajamento, como duração média da sessão e taxa de abandono. Esta capacidade de monitorização granular das sessões fornece aos administradores visibilidade essencial sobre como os utilizadores interagem realmente com o sistema num ambiente de produção, informando decisões sobre otimização de *prompts* LLMs, ajustes na estratégia de recuperação de conhecimento e melhorias na experiência conversacional com base em dados de utilização reais, em vez de suposições teóricas.

4.5.4 Controlo da Base de Conhecimento e Gestão Documental

O módulo de gestão de documentos fornece controlo administrativo completo sobre a base de conhecimento que alimenta o sistema RAG, apresentando inicialmente estatísticas agregadas críticas, incluindo o número total de documentos indexados, indicando a cobertura completa da base de conhecimento disponível, o número total de blocos que representam a granularidade da segmentação aplicada aos documentos originais para otimização da recuperação semântica e o número de vetores armazenados no Pinecone, confirmando a sincronização entre documentos processados e incorporações disponíveis para pesquisa

vetorial. A lista detalhada de documentos apresenta uma tabela responsiva com informações estruturadas por documento, incluindo o nome do ficheiro original, preservando a rastreabilidade até à fonte; o caminho completo no sistema de ficheiros, facilitando a localização para manutenção ou resolução de problemas; o tamanho do ficheiro automaticamente convertido para unidades legíveis (KB, MB), fornecendo uma indicação do volume do conteúdo; o número de blocos gerados durante o processo de segmentação, refletindo a complexidade e o comprimento do documento; e um carimbo de data/hora de indexação preciso, documentando quando o documento foi processado e disponibilizado ao sistema de IA (ver B.1).

As operações administrativas críticas implementadas incluem a funcionalidade de reindexação completa, que reconstrói toda a base de conhecimento vetorial a partir do zero, uma operação essencial após atualizações do documento de origem, alterações nos parâmetros de segmentação ou migração de versões do modelo de incorporação. Esta operação potencialmente destrutiva implementa várias salvaguardas de segurança, incluindo uma caixa de diálogo de confirmação explícita que requer uma ação deliberada do administrador antes de prosseguir, *feedback* visual detalhado sobre o progresso através de uma barra de progresso ou sobreposição informativa indicando as etapas de reindexação (leitura de ficheiros, segmentação, geração de incorporação, upload para o Pinecone) e registo completo de todas as operações com carimbos de data/hora e resultados, permitindo auditoria e resolução de problemas. O sistema também implementa validação de integridade que compara contadores de documentos, fragmentos e vetores para identificar inconsistências que podem indicar falhas parciais de indexação ou problemas de sincronização com o Pinecone, alertando proativamente os administradores quando são detetadas discrepâncias e sugerindo ações corretivas apropriadas.

A gestão de documentos também oferece funcionalidades de análise, incluindo a identificação de documentos potencialmente obsoletos com base na idade da indexação em comparação com a política de atualização definida, documentos subutilizados que raramente aparecem nos resultados de pesquisa, sugerindo possível irrelevância ou necessidade

de otimização de conteúdo, e documentos com uma alta taxa de pesquisa, mas baixa utilidade percebida (com base no *feedback* implícito ou explícito do utilizador), indicando uma lacuna entre a relevância semântica e a utilidade prática. *feedback* dos utilizadores), indicando uma lacuna entre a relevância semântica e a utilidade prática. Essa visibilidade da base de conhecimento permite a otimização contínua, identificando documentos que devem ser atualizados, expandidos, simplificados ou removidos, garantindo que o sistema RAG mantenha respostas de alta qualidade e relevantes ao longo do tempo por meio da curadoria ativa, em vez do acúmulo passivo de conteúdo potencialmente desatualizado ou de baixa qualidade.

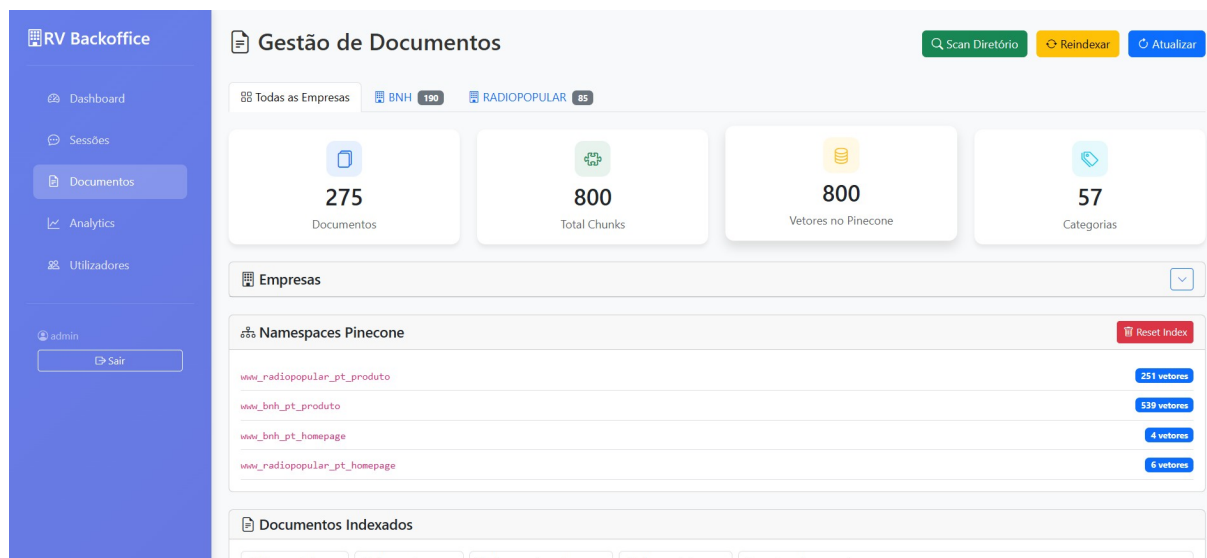


Figura 4.6: Pagina de Documentos

Capítulo 5

Avaliação e Resultados

Este capítulo apresenta a avaliação sistemática do sistema 3D-Chatbot, documentando testes para verificar o cumprimento dos objetivos e a eficácia da solução nas dimensões críticas de precisão das respostas, desempenho técnico, usabilidade da experiência imersiva e robustez operacional em cenários comerciais reais.

5.1 Metodologia de Avaliação

5.1.1 Desenho Experimental

A metodologia de avaliação foi desenhada em torno de dois casos de estudo complementares, representando diferentes contextos comerciais com características distintas:

1. **BNH**: Empresa portuguesa no Brigantia Ecopark (Bragança), contexto no mercado odontológico, com 65 interações testadas;
2. **Rádio Popular**: Focada em eletrónica de consumo e eletrodomésticos, com 55 interações avaliadas.

Esta abordagem multi-cenário permitiu avaliar a generalização e adaptabilidade do sistema, conforme especificado nos objetivos da investigação. Os testes foram realizados

num ambiente controlado, utilizando o Meta Quest 3 como dispositivo de RV principal, garantindo condições consistentes para todas as interações.

5.1.2 Categorização das Interações

As interações foram classificadas em cinco categorias funcionais, refletindo padrões típicos de utilização em contextos comerciais:

- **Consulta sobre produtos:** Perguntas gerais sobre categorias, disponibilidade ou características amplas;
- **Especificações técnicas:** Solicitações detalhadas sobre dimensões, materiais, capacidades ou características específicas;
- **Preço e disponibilidade:** Questões sobre valores, promoções, stock ou prazos de entrega;
- **Assistência à navegação:** Pedidos de orientação no ambiente 3D ou localização de produtos específicos;
- **Informações gerais:** Consultas sobre políticas da empresa, serviços, horários ou contactos.

5.1.3 Métricas de Avaliação

Para cada interação, foram registadas três métricas principais:

Taxa de sucesso: Classificação binária (sucesso/falha) baseada em critérios objetivos de precisão factual e relevância contextual;

Tempo de resposta: Latência total desde a captura de voz até à síntese da resposta audível;

Qualidade da experiência: Avaliação qualitativa da naturalidade conversacional e adequação ao contexto imersivo.

5.2 Resultados - Caso BNH

5.2.1 Análise Quantitativa

Para validar a eficácia do sistema de chatbot 3D, realizámos testes abrangentes envolvendo 65 interações distintas no cenário BNH. A avaliação centrou-se na precisão das respostas, na capacidade de resposta do sistema e na qualidade da experiência do utilizador no ambiente imersivo de RV. A Tabela 5.1 apresenta os resultados consolidados.

Tabela 5.1: Resultados da interação com o chatbot no cenário BNH (N=65)

Tipo de interação	Contagem	Taxa de sucesso (%)	Tempo médio (s)
Consulta sobre o produto	23	95,7	2,49
Especificações técnicas	15	93,3	2,18
Preço e disponibilidade	12	100,0	1,89
Assistência à navegação	8	87,5	1,24
Informações gerais	7	100,0	2,27
Total/Média	65	94,6	2,17

Os resultados demonstram uma taxa de sucesso global de **94,6%**, superando as expectativas iniciais e validando a eficácia da arquitetura RAG implementada. Particularmente notável é o desempenho perfeito (100%) nas categorias de “Preço e disponibilidade” e “Informações gerais”, indicando que o sistema consegue recuperar e apresentar informações factuais estruturadas com elevada precisão.



Figura 5.1: Resultados gerais

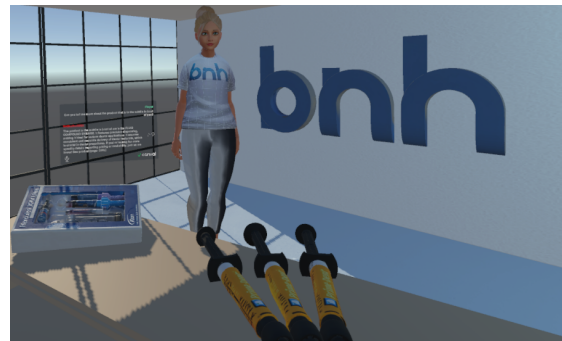


Figura 5.2: Pergunta de avaliação do chatbot

O tempo médio de resposta de **2,17 segundos** situa-se confortavelmente abaixo do limite de 3 segundos estabelecido como requisito não funcional, preservando a naturalidade conversacional. A variação nos tempos entre categorias reflete a complexidade diferenciada do processamento: consultas de navegação simples (1,24s) são significativamente mais rápidas que consultas de produtos complexas (2,49s), demonstrando a otimização adaptativa do sistema.

5.2.2 Análise de Falhas

As 4 interações mal sucedidas (5,4% do total) distribuíram-se da seguinte forma:

- **Consulta sobre produtos** (1 falha): Pergunta ambígua sobre “melhor produto” sem especificação de contexto de utilização;
- **Especificações técnicas** (1 falha): Solicitação de especificação não documentada na base de conhecimento;
- **Assistência à navegação** (1 falha): Comando de navegação mal interpretado devido a ruído ambiental.

A análise destas falhas revelou três causas principais:

1. **Ambiguidade linguística**: O sistema teve dificuldade com perguntas excessivamente vagas sem contexto suficiente;
2. **Lacunas na base de conhecimento**: Informações específicas ausentes dos documentos indexados;
3. **Erros de reconhecimento de voz**: Condições acústicas adversas comprometeram a transcrição inicial.

5.3 Resultados - Rádio Popular

5.3.1 Análise Quantitativa

A Tabela 5.2 apresenta os resultados consolidados das 55 interações realizadas no cenário Rádio Popular.

Tabela 5.2: Resultados da interação com o chatbot no cenário Rádio Popular (N=55)

Tipo de interação	Contagem	Taxa de sucesso (%)	Tempo médio (s)
Consulta sobre o produto	19	94,7	2,31
Especificações técnicas	14	92,9	2,45
Preço e disponibilidade	10	100,0	1,95
Assistência à navegação	7	85,7	1,18
Informações gerais	5	100,0	2,08
Total/Média	55	94,5	2,19

O cenário Rádio Popular apresentou resultados notavelmente consistentes com o caso BNH, com uma taxa de sucesso global de **94,5%** e tempo médio de resposta de **2,19 segundos**. Esta consistência entre cenários distintos valida empiricamente a generalização e adaptabilidade do sistema, um dos objetivos principais da investigação.

Diferenças Significativas

Apesar da consistência geral, algumas diferenças mensuráveis foram identificadas:

- **Especificações técnicas ligeiramente mais desafiadoras** na Rádio Popular (92,9% vs 93,3%), possivelmente devido à maior complexidade técnica de produtos eletrônicos, que envolvem terminologia mais especializada (processadores, resoluções, capacidades de armazenamento);
- **Tempo de processamento marginalmente superior** para especificações técnicas no contexto de eletrônica (2,45s vs 2,18s, um aumento de 12%), sugerindo que a recuperação de informações técnicas densas requer processamento adicional;
- **Taxa de sucesso em navegação ligeiramente inferior** na Rádio Popular (85,7% vs 87,5%), embora ambos os valores permaneçam acima de 85%;

- **Consultas sobre produtos mais rápidas** na Rádio Popular (2,31s vs 2,49s), potencialmente devido a uma estruturação mais uniforme do catálogo de produtos eletrônicos.

Interpretação das Diferenças

Essas diferenças, embora estatisticamente modestas, fornecem informações valiosas sobre os desafios específicos de cada domínio:

O domínio dos produtos eletrônicos de consumo (Rádio Popular) tem características que tornam as especificações técnicas particularmente complexas: maior densidade de informação técnica por produto (Central Processing Unit (CPU), Graphics Processing Unit (GPU), Random Access Memory (RAM), armazenamento, resolução, taxa de atualização), terminologia altamente especializada e em constante evolução (*OLED*, *QLED*, *HDR10+*, *Dolby Atmos*), e a necessidade de comparações multi-dimensionais entre produtos semelhantes e especificações interdependentes que requerem explicações contextualizadas.

5.4 Artigos Científicos

O desenvolvimento e a investigação subjacentes a esta dissertação resultaram em quatro publicações científicas em conferências internacionais, demonstrando a relevância acadêmica e a aplicabilidade prática do trabalho realizado:

- O primeiro artigo, “Chatbots, to transform the user experience at 3D environments: a systematic literature review”, foi aceite para apresentação na International Conference on Information and Communication Technology for Intelligent Systems (ICTIS), programada para 23 de maio de 2025, nos Estados Unidos. Esta publicação apresenta a revisão sistemática da literatura que fundamentou teoricamente o trabalho de investigação, seguindo rigorosamente as diretrizes metodológicas de Design Science Research. O artigo sintetiza as principais contribuições acadêmicas identificadas através da análise de 15 estudos selecionados após aplicação de critérios

rigorosos de inclusão e exclusão em bases de dados académicas reconhecidas. A revisão identifica lacunas críticas na investigação existente, particularmente no que respeita à integração técnica completa de sistemas RAG com ambientes 3D imersivos, adaptabilidade de sistemas conversacionais a múltiplos domínios comerciais, justificando assim a relevância e originalidade da presente dissertação (ver D.1).

- O segundo artigo, intitulado “3D-Chatbot: A Intelligent Conversational Agent Integrating RAG, NLP, And VR For Commercial Stores”, foi apresentado na 11th International Conference on Engineering and Emerging Technologies (ICEET), realizada em Kuala Lumpur, Malásia, nos dias 22-23 de outubro de 2025. Esta publicação documenta a arquitetura técnica completa do sistema desenvolvido, com ênfase particular na integração entre componentes RAG, processamento de linguagem natural e ambientes de realidade virtual. O artigo apresenta a implementação prática do sistema em contextos comerciais reais, detalhando as soluções arquitetónicas adotadas para garantir baixa latência, precisão factual das respostas e uma experiência imersiva fluida (ver D.3).
- Complementarmente, foram desenvolvidos dois trabalhos focados na aplicação de tecnologias conversacionais ao sector do turismo, demonstrando a generalização das competências técnicas adquiridas durante esta investigação. O terceiro artigo, “Smart Tourism through chatbots: A practical implementation for Northern Portugal”, foi apresentado na International Conference on Advanced Research in Technologies, Information, Innovation, and Sustainability (ARTIIS 2025), realizada entre 21-23 de outubro de 2025. Esta publicação explora a implementação de chatbots bidimensionais especificamente adaptados ao contexto turístico do Norte de Portugal, demonstrando como sistemas conversacionais podem enriquecer a experiência de visitantes através de recomendações personalizadas, informações culturais contextualizadas e assistência multilingue. O trabalho evidencia a aplicabilidade das arquiteturas RAG desenvolvidas nesta dissertação a domínios alternativos, validando a generalização da abordagem proposta (ver D.2).

- O quarto artigo, “ A chatbot to enhance the portuguese national tourism website ”, será apresentado na 17th International Conference on Education Technology and Computers (ICETC), programada para Barcelona, Espanha, entre 18-21 de setembro de 2025. Esta publicação documenta o desenvolvimento de um chatbot integrado no website oficial de turismo português, com o objetivo de melhorar a acessibilidade da informação turística e aumentar o envolvimento dos visitantes através de interfaces conversacionais naturais. O trabalho demonstra como tecnologias de processamento de linguagem natural podem transformar plataformas web tradicionais em experiências interativas mais envolventes, reduzindo significativamente o tempo necessário para os utilizadores encontrarem informações relevantes(ver D.4).

Estas publicações refletem a dupla contribuição do trabalho: por um lado, uma base teórica sólida baseada numa análise sistemática do estado da arte; por outro lado, uma implementação prática validada empiricamente que demonstra a aplicabilidade comercial das tecnologias investigadas. A aceitação em conferências internacionais revisadas por pares confirma a qualidade científica do trabalho realizado e a sua relevância para a comunidade académica e industrial.

Capítulo 6

Conclusões

Esta dissertação abordou o desafio de integrar chatbots baseados em inteligência artificial conversacional em ambientes tridimensionais imersivos, com o objetivo de transformar a experiência do utilizador em contextos comerciais digitais. O trabalho baseou-se na observação de que os sistemas tradicionais de apoio ao cliente, baseados em interfaces bidimensionais, são inadequados para a realidade virtual, comprometendo tanto a experiência do utilizador como os objetivos comerciais das organizações. Para resolver este problema, foi desenvolvido um sistema integrado que combina quatro componentes principais: um agente conversacional contextualizado baseado em LLMs, um sistema RAG para garantir respostas precisas e atualizadas, uma experiência de realidade virtual imersiva com interação multimodal e um sistema de gestão e monitorização que permite a evolução contínua do sistema.

A implementação técnica envolveu tecnologias de ponta, como *GPT-4o-mini* para geração de respostas, Pinecone para armazenamento vetorial, Unity para desenvolvimento de ambientes 3D e Convai para processamento de voz em tempo real. A validação empírica, realizada por meio de 120 interações distribuídas em dois cenários comerciais distintos (BNH e Rádio Popular), demonstrou a eficácia da solução proposta com resultados notáveis: uma taxa de sucesso geral de 94,6% e um tempo médio de resposta de 2,17 segundos, confirmando que o sistema atende aos objetivos estabelecidos, proporcionando uma experiência de conversação natural e eficaz em ambientes imersivos. A consistência

dos resultados entre diferentes cenários (BNH com 94,6% e Rádio Popular com 94,5%) valida a generalização e adaptabilidade do sistema, demonstrando que a arquitetura proposta é suficientemente robusta para ser aplicada em diferentes domínios comerciais sem a necessidade de reengenharia substancial, representando uma contribuição significativa para a democratização dos assistentes virtuais imersivos e reduzindo as barreiras técnicas e económicas que tradicionalmente limitam a sua adoção por pequenas e médias empresas.

6.1 Trabalho Futuro

Apesar dos resultados positivos alcançados, surgem várias direções promissoras para futuras pesquisas e desenvolvimentos:

- **Infraestrutura e implementação:** Atualmente, o sistema utiliza MongoDB e Redis em contentores Docker, proporcionando uma base sólida para o desenvolvimento. No entanto, a implementação em produção requer a implementação de uma orquestração robusta através do *Kubernetes* ou do *Docker Swarm*, permitindo escalabilidade horizontal, alta disponibilidade e gestão automatizada de falhas. A integração do **RabbitMQ** como um *broker* de mensagens assíncrono permitiria a dissociação de componentes críticos, melhorando significativamente a resiliência do sistema e permitindo o processamento paralelo de consultas durante períodos de alto tráfego.
- **Otimizações técnicas:** a implementação de *cache* distribuído via Redis para resultados de consultas frequentes reduziria a latência e os custos operacionais com APIs externas. O sistema de enfileiramento baseado em *RabbitMQ* permitiria que operações computacionalmente intensivas fossem processadas de forma assíncrona, sem comprometer a capacidade de resposta da interface do utilizador.
- **Expansão funcional:** integração de análise de sentimentos em tempo real para adaptar o tom de conversação do avatar, implementação de recomendações proativas com base no comportamento de navegação do utilizador no ambiente 3D e suporte multilingue nativo com deteção automática de idioma.

- Melhorias na base de conhecimento: Desenvolvimento de pipelines automatizados para atualizações incrementais de documentos, detecção de conteúdo obsoleto por meio de análise temporal e validação automática da qualidade da resposta por meio de modelos de avaliação especializados.

Bibliografia

- [1] B. S. Aswin, S. Vishnubala, D. Dhinakaran, N. J. Kumar, S. U. Sankar e A. M. Al Faisal, “A Research on Metaverse and its Application,” em *2023 World Conference on Communication & Computing (WCONF)*, IEEE, 2023, pp. 1–6.
- [2] I. J. Akpan, E. A. P. Udoh e B. Adebisi, “Small business awareness and adoption of state-of-the-art technologies in emerging and developing markets, and lessons from the COVID-19 pandemic,” *Journal of Small Business & Entrepreneurship*, vol. 34, n.º 2, pp. 123–140, 2022.
- [3] J. H. Al Shamsi, M. Al-Emran e K. Shaalan, “Understanding key drivers affecting students’ use of artificial intelligence-based voice assistants,” *Education and Information Technologies*, vol. 27, n.º 6, pp. 8071–8091, 2022.
- [4] M. Rahevar e S. Darji, “The adoption of AI-driven Chatbots into a recommendation for e-commerce systems to targeted customer in the selection of product,” *International Journal of Management, Economics and Commerce*, vol. 1, n.º 2, pp. 128–137, 2024.
- [5] V. Chheang et al., “Towards anatomy education with generative AI-based virtual assistants in immersive virtual reality environments,” em *2024 IEEE international conference on artificial intelligence and eXtended and virtual reality (AIxVR)*, IEEE, 2024, pp. 21–30.
- [6] F. de Almeida Barros e P. A. Tedesco, “Agentes inteligentes conversacionais: conceitos básicos e desenvolvimento,” *JAI-Jornada de Atualização em Informática (2016)*, pp. 169–218, 2016.

- [7] S. Jolibois, A. Ito e T. Nose, “Multimodal expressive embodied conversational agent design,” em *International Conference on Human-Computer Interaction*, Springer, 2023, pp. 244–249.
- [8] B. Dumas, D. Lalanne e S. Oviatt, “Multimodal interfaces: A survey of principles, models and frameworks,” em *Human machine interaction: Research results of the mmi program*, Springer, 2009, pp. 3–26.
- [9] T. Araujo, “Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions,” *Computers in human behavior*, vol. 85, pp. 183–189, 2018.
- [10] T. Yamazaki, T. Mizumoto, K. Yoshikawa, M. Ohagi, T. Kawamoto e T. Sato, “An open-domain avatar chatbot by exploiting a large language model,” em *Proceedings of the 24th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, 2023, pp. 428–432.
- [11] S. Gregor e A. R. Hevner, “Positioning and presenting design science research for maximum impact,” *MIS quarterly*, pp. 337–355, 2013.
- [12] M. Majdalawieh e S. Khan, “Building an integrated digital transformation system framework: a design science research, the case of feduni,” *Sustainability*, vol. 14, n.º 10, p. 6121, 2022.
- [13] S. J. Miah, H. Genemo et al., “A design science research methodology for expert systems development,” *Australasian Journal of Information Systems*, vol. 20, 2016.
- [14] A. R. C. P. Sena, F. Thiago e A. M. de Vasconcelos, “Innovation Culture in Public Organizations: A Systematic Literature Review,” *Future Studies Research Journal: Trends and Strategies*, vol. 16, n.º 1, e821–e821, 2024.
- [15] K. Buchta et al., “NUX IVE—a research tool for comparing voice user interface and graphical user interface in VR,” em *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, IEEE, 2022, pp. 982–983.

- [16] G. Calahorra-Candao e M. J. Martín-de Hoyos, “The effect of anthropomorphism of virtual voice assistants on perceived safety as an antecedent to voice shopping,” *Computers in Human Behavior*, vol. 153, p. 108 124, 2024.
- [17] J. Zhu, R. Kumaran, C. Xu e T. Höllerer, “Free-form conversation with human and symbolic avatars in mixed reality,” em *2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, IEEE, 2023, pp. 751–760.
- [18] I. L. Garcia, E. Schott, M. Gohsen, V. Bernhard, B. Stein e B. Froehlich, “Speaking with objects: Conversational agents’ embodiment in virtual museums,” em *2024 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, IEEE, 2024, pp. 279–288.
- [19] M. D. Vu et al., “Gptvoicetasker: Advancing multi-step mobile task efficiency through dynamic interface exploration and learning,” em *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*, 2024, pp. 1–17.
- [20] J. Hong et al., “A Context-Aware Onboarding Agent for Metaverse Powered by Large Language Models,” em *Proceedings of the 2024 ACM Designing Interactive Systems Conference*, 2024, pp. 1857–1874.
- [21] A. Bönsch, D. Hashem, J. Ehret e T. W. Kuhlen, “Being Guided or Having Exploratory Freedom: User Preferences of a Virtual Agent’s Behavior in a Museum,” em *Proceedings of the 21st ACM international conference on intelligent virtual agents*, 2021, pp. 33–40.
- [22] P. Safadel, S. N. Hwang e J. M. Perrin, “User acceptance of a virtual librarian chatbot: an implementation method using ibm watson natural language processing in virtual immersive environment,” *TechTrends*, vol. 67, n.º 6, pp. 891–902, 2023.
- [23] R. Yang, M. Fu, C. Tantithamthavorn, C. Arora, L. Vandenhurk e J. Chua, “RAGVA: Engineering retrieval augmented generation-based virtual assistants in practice,” *arXiv preprint arXiv:2502.14930*, 2025.

- [24] V. Chamola, S. Sai, R. Sai, A. Hussain e B. Sikdar, “Generative ai for consumer electronics: Enhancing user experience with cognitive and semantic computing,” *IEEE Consumer Electronics Magazine*, vol. 14, n.º 2, pp. 10–19, 2024.
- [25] S.-C. Necula e V.-D. Păvăloaia, “AI-driven recommendations: A systematic review of the state of the art in e-commerce,” *Applied Sciences*, vol. 13, n.º 9, p. 5531, 2023.
- [26] R. Sharma, S. Srivastva e S. Fatima, “E-commerce and digital transformation: Trends, challenges, and implications,” *Int. J. Multidiscip. Res.(IJFMR)*, vol. 5, pp. 1–9, 2023.
- [27] S. Zhou, W. Zheng, Y. Xu e Y. Liu, “Enhancing user experience in VR environments through AI-driven adaptive UI design,” *Journal of Artificial Intelligence General Science (JAIGS) ISSN: 3006-4023*, vol. 6, n.º 1, pp. 59–82, 2024.
- [28] S. Singh e T. S. Vijay, “Technology roadmapping for the e-commerce sector: A text-mining approach,” *Journal of Retailing and Consumer Services*, vol. 81, p. 103977, 2024.
- [29] P. Kazienko e E. Cambria, “Toward responsible recommender systems,” *IEEE Intelligent Systems*, vol. 39, n.º 3, pp. 5–12, 2024.
- [30] *General Data Protection Regulation (GDPR) – Legal Text*, <https://gdpr-info.eu/>, 2016.
- [31] *Art. 5 GDPR – Principles relating to processing of personal data - General Data Protection Regulation*, <https://gdpr-info.eu/art-5-gdpr/>.
- [32] *Art. 17 GDPR – Right to erasure (‘right to be forgotten’) - General Data Protection Regulation*, <https://gdpr-info.eu/art-17-gdpr/>.
- [33] *Art. 9 GDPR – Processing of special categories of personal data - General Data Protection Regulation*, <https://gdpr-info.eu/art-9-gdpr/>.
- [34] *Art. 25 GDPR – Data protection by design and by default - General Data Protection Regulation*, <https://gdpr-info.eu/>.

- [35] P. Kashefi, Y. Kashefi e A. Ghafouri Mirsarai, “Shaping the future of AI: balancing innovation and ethics in global regulation,” *Uniform Law Review*, vol. 29, n.º 3, pp. 524–548, 2024.
- [36] A. J. Wulf e O. Seizov, ““Please understand we cannot provide further information”: evaluating content and transparency of GDPR-mandated AI disclosures,” *AI & SOCIETY*, vol. 39, n.º 1, pp. 235–256, 2024.
- [37] D. Liga e L. Robaldo, “Fine-tuning GPT-3 for legal rule classification,” *Computer Law & Security Review*, vol. 51, p. 105 864, 2023.
- [38] S. Brdnik e B. Šumak, “Current trends, challenges and techniques in XAI field; A tertiary study of XAI research,” em *2024 47th MIPRO ICT and Electronics Convention (MIPRO)*, IEEE, 2024, pp. 2032–2038.
- [39] *O que é o LangChain? — Explicação sobre o LangChain — AWS*. URL: <https://aws.amazon.com/pt/what-is/langchain/>.
- [40] *Chatbots with Pinecone*, jun. de 2023. URL: <https://www.pinecone.io/learn/chatbots-with-pinecone/>.

Apêndice A

Proposta Original do Projeto



Mestrado em Informática

Tese do 2º ano – Ano escolar de 2024/2025

Chatbots, para transformar a experiência e interação do utilizador em cenários imersivos 3D

Tipo de Orientador: Interno

Orientador: Pedro Filipe Fernandes Oliveira

Coorientador: Paulo Jorge Teixeira Matos

Instituição do coorientador: Escola Superior de Tecnologia e Gestão de Bragança

1 Objetivos:

No atual cenário de transformação digital, os chatbots e o comércio eletrónico 3D emergem como tecnologias revolucionárias. A inteligência artificial, com as suas aplicações comprovadas em diversos setores, tem desempenhado um papel essencial no comércio eletrónico.

Enquanto os chatbots transformam a interação entre empresas e consumidores com processamento de linguagem natural, o comércio eletrónico 3D oferece experiências imersivas e realistas. A integração dessas tecnologias tem o potencial de criar experiências de compra inovadoras e interativas. Os objetivos desta investigação concentram-se no desenvolvimento de um chatbot genérico e dinâmico, integrado em cenários tridimensionais.

Os objetivos principais desta pesquisa incluem:

- **Desenvolver um chatbot genérico e adaptável:** Criar um chatbot capaz de se adaptar a diferentes tipos de cenários, fornecendo respostas relevantes e contextualizadas.
- **Explorar a integração de chatbots em ambientes 3D:** Investigar e implementar técnicas para incorporar chatbots interativos em cenários tridimensionais, melhorando a imersão e a interação.

- Fornecer personalização dinâmica: Permitir que o chatbot seja configurado e parametrizado de forma dinâmica, ajustando o comportamento e os dados de acordo com as necessidades específicas para cada cenário.
- Implementar um sistema de recolha de informação baseado em web crawlers: Desenvolver um mecanismo autónomo e dinâmico para alimentar o chatbot com dados provenientes de crawlers, garantindo que estes se mantêm atualizados e relevantes.
- Integrar o chatbot em cenários 3D: Desenvolver soluções técnicas que liguem o chatbot a ambientes tridimensionais, permitindo interações em tempo real nesses espaços.
- Avaliar a usabilidade e a eficiência em diferentes cenários: Testar o chatbot em diferentes tipos de cenários e contextos, garantindo que é flexível e útil para uma variedade de casos de utilização.
- Incorporar técnicas de inteligência artificial: Utilizar métodos de aprendizagem automática e de processamento de linguagem natural (NLP) para melhorar a capacidade do chatbot para compreender e responder.
- Desenvolver ferramentas de monitorização e análise: Criar ferramentas para acompanhar as interações, avaliar o desempenho do chatbot e identificar possíveis áreas de melhoria.

2 **Calendarização:**

O planeamento começa com uma extensa revisão do estado de arte, seguida de ciclos iterativos de implementação, avaliação e documentação científica, culminando na elaboração e revisão da tese final.

O calendário foi concebido para permitir uma abordagem metodológica rigorosa, em que cada fase alimenta e sustenta as fases seguintes:

- Revisão do estado de arte (8 semanas).
- Escrita de artigo (1 semana).
- Design e implementação de protótipos (12 semanas).
- Escrita de artigo (2 semanas).
- Implementação de ambiente de testes (10 semanas).
- Avaliação dos resultados (5 semanas).
- Escrita de artigo (2 semanas).

- Escrita da tese (10 semanas).
- Revisão final (2 semanas).

Tarefa/Etapa	Descrição	Duração(semanas)
T1	Revisão do estado de arte	8
T2	Escrita de artigo	1
E1	Publicação do artigo	0
T3	Design e implementação de protótipos	12
T4	Escrita de artigo	2
E2	Publicação do artigo	0
T5	Implementação de ambiente de testes	10
T6	Avaliação dos resultados	5
E3	Proposta de Arquitetura	0
T7	Escrita de artigo	2
E4	Publicação do artigo	0
T8	Escrita da tese	10
T9	Revisão final	2
E5	Apresentação da tese	0
Total (semanas)		52

Figura 1 - Calendário

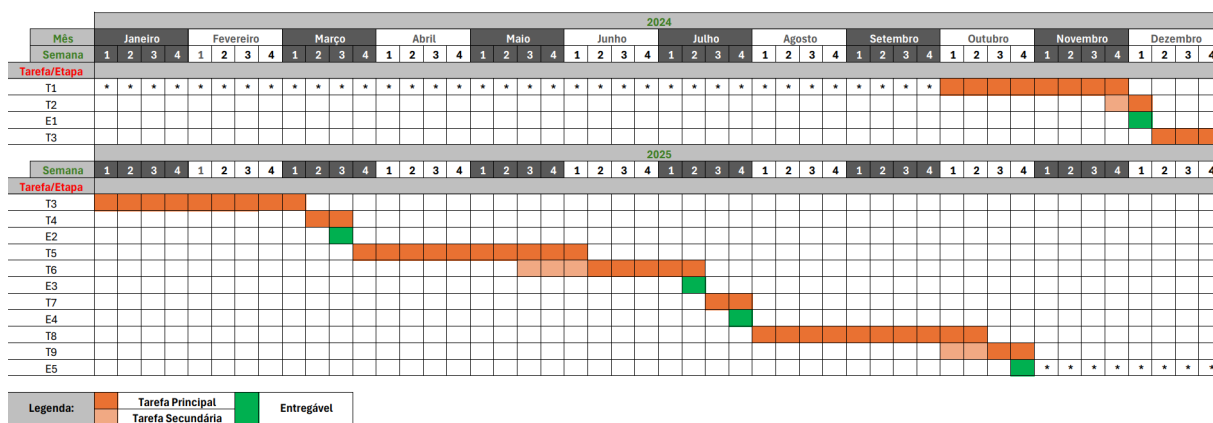


Figura 2 - Cronograma

Apêndice B

Figuras Complementares

Nome do Ficheiro	Empresa	Tipo	Categoria	Status	Chunks	Tamanho	Indexado em	Ações
PORTA-MATRIZES-BNH-af9a32c2.md www_bnh_pt\produtos\PORTA-MATRIZES-BNH-af9a32c2.md	BNH	produto	Clear	Completo	2	2.85 KB	30/10/2025, 18:21:48	
PROFEX-CÁPSULAS-BNH-2c8884f7.md www_bnh_pt\produtos\PROFEX-CÁPSULAS-BNH-2c8884f7.md	BNH	produto	Clear	Completo	2	2.79 KB	30/10/2025, 18:21:48	
PROFEX-PASTA-BNH-048f0b08.md www_bnh_pt\produtos\PROFEX-PASTA-BNH-048f0b08.md	BNH	produto	Profilaxia	Completo	2	2.82 KB	30/10/2025, 18:21:49	
PROFEX-PASTA-BNH-13941ee2.md www_bnh_pt\produtos\PROFEX-PASTA-BNH-13941ee2.md	BNH	produto	Profilaxia	Completo	2	2.68 KB	30/10/2025, 18:21:50	
PROFEX-PASTA-BNH-afc49cbc.md www_bnh_pt\produtos\PROFEX-PASTA-BNH-afc49cbc.md	BNH	produto	Profilaxia	Completo	2	3.31 KB	30/10/2025, 18:21:51	
PROFEX-PASTA-BNH-da336a54.md www_bnh_pt\produtos\PROFEX-PASTA-BNH-da336a54.md	BNH	produto	Profilaxia	Completo	2	2.85 KB	30/10/2025, 18:21:52	
Profilaxia-BNH-2ce4e2de.md www_bnh_pt\produtos\Profilaxia-BNH-2ce4e2de.md	BNH	produto	Profilaxia	Completo	4	4.6 KB	30/10/2025, 18:21:53	
PROPHY-PASTE-CCS-BNH-89db2b02.md www_bnh_pt\produtos\PROPHY-PASTE-CCS-BNH-89db2b02.md	BNH	produto	Clear	Completo	3	3.76 KB	30/10/2025, 18:21:53	
PURE-PROTECT-BNH-d33bf4a7.md www_bnh_pt\produtos\PURE-PROTECT-BNH-d33bf4a7.md	BNH	produto	Branqueamento	Completo	3	3.59 KB	30/10/2025, 18:21:54	
...	

Figura B.1: Tabela de Getão de Documentos









RV Backoffice

Gestão de Utilizadores

Adicionar Utilizador

Lista de Utilizadores

Atualizar

Username	Email	Role	Último Login	Estado	Ações
admin	admin@bnh.com	Admin	31/10/2025, 17:31:14	Ativo	  
user	user@example.com	Viewer	01/01/1970, 01:00:00	Ativo	  
ceo	ceo@ceo.com	Manager	01/01/1970, 01:00:00	Ativo	  

admin

Sair

Figura B.2: Pagina de Gestão de Utilizadores

Apêndice C

Código

C.1 Crawler (GDPR)

```
1 class Priority(Enum):
2     """Prioridades para URLs"""
3     HIGH = 1
4     MEDIUM = 2
5     LOW = 3
6
7
8 @dataclass
9 class CrawlConfig:
10     """Configuracao do crawler"""
11     base_urls: List[str]
12     output_dir: str = "crawled_data_gdpr"
13     max_pages_per_domain: int = 50
14     max_total_pages: int = 100
15     delay_between_requests: float = 1.5
16     max_concurrent_requests: int = 10
17     respect_robots_txt: bool = True
18     max_retries: int = 3
19     timeout: int = 30
```

```

20     max_depth: int = 5
21     user_agent: str = "AdvancedCrawler/2.0 (+http://example.com/bot)"
22
23     # GDPR
24     anonymize_personal_data: bool = True
25
26     # Filtros
27     exclude_extensions: List[str] = field(default_factory=lambda: [
28         '.pdf', '.zip', '.exe', '.jpg', '.jpeg', '.png', '.gif',
29         '.mp4', '.mp3', '.doc', '.docx', '.xls', '.xlsx'
30     ])
31     exclude_patterns: List[str] = field(default_factory=lambda: [
32         '/login', '/logout', '/admin', '/cart', '/checkout',
33         '/contacto', '/contact', '/testemunhos', '/reviews',
34         '/perfil', '/profile', '/conta', '/account',
35         '/forum', '/comunidade', '/comments', '/user', '/cliente'
36     ])
37
38
39
40 @dataclass
41 class PageContent:
42     """Estrutura para armazenar conteudo de uma pagina"""
43     url: str
44     domain: str
45     title: str
46     content: str
47     metadata: Dict
48     links: List[str]
49     images: List[str]
50     last_updated: str
51     category: str
52     content_hash: str
53     depth: int
54     word_count: int

```

```

55     crawl_timestamp: str
56     anonymized: bool = False
57
58
59 @dataclass
60 class CrawlStats:
61     """Estatísticas do crawling"""
62     total_pages: int = 0
63     successful_pages: int = 0
64     failed_pages: int = 0
65     anonymized_items: int = 0
66     pages_per_domain: Dict[str, int] = field(default_factory=dict)
67     total_links: int = 0
68     total_images: int = 0
69     start_time: str = ""
70     end_time: str = ""
71     duration_seconds: float = 0.0
72
73 class PersonalDataAnonymizer:
74     """Anonimizador de dados pessoais conforme GDPR"""
75
76     def __init__(self):
77         self.patterns = {
78             'email': (
79                 r'\b[A-Za-z0-9._%+-]+@[A-Za-z0-9.-]+\.[A-Z|a-z]{2,}\b',
80                 '[EMAIL_REDACTED]'
81             ),
82             'phone_pt': (
83                 r'\b(\+351|00351)?\s*[29]\d{8}\b',
84                 '[PHONE_REDACTED]'
85             ),
86             'phone_international': (
87                 r'\b\+\d{1,3}\s*\d{6,14}\b',
88                 '[PHONE_REDACTED]'
89             ),

```

```

90         'vat_pt': (
91             r'\bPT\s*\d{9}\b',
92             '[VAT_REDACTED]',
93         ),
94         'nif': (
95             r'\b\d{9}\b(?:.*NIF|.*fiscal)',
96             '[NIF_REDACTED]',
97         ),
98         'cc': (
99             r'\b\d{8}\s*\d\s*[A-Z]{2}\d\b',
100            '[CC_REDACTED]',
101        ),
102        'iban': (
103            r'\b[A-Z]{2}\d{2}\s*\d{4}\s*\d{4}\s*\d{4}\s*\d{4}\s*\d
104                {1,2}\b',
105            '[IBAN_REDACTED]',
106        ),
107        'postal_code': (
108            r'\b\d{4}-\d{3}\b',
109            '[POSTAL_CODE_REDACTED]',
110        ),
111    }
112
113    self.name_indicators = [
114        'nome completo', 'full name', 'nome:', 'name:',
115        'autor:', 'author:', 'escrito por', 'written by'
116    ]
117
118    self.sensitive_terms = [
119        'senha', 'password', 'token', 'api key', 'secret',
120        'cartao', 'credit card', 'cvv', 'pin'
121    ]
122
123    def anonymize_text(self, text: str) -> Tuple[str, int]:
124        """Anonimizar texto removendo dados pessoais"""

```

```

124     if not text:
125         return text, 0
126
127     anonymized = text
128     redactions_count = 0
129
130     # Aplicar padroes de regex
131     for data_type, (pattern, replacement) in self.patterns.items():
132         matches = re.findall(pattern, anonymized, re.IGNORECASE)
133         if matches:
134             anonymized = re.sub(pattern, replacement, anonymized,
135                                 flags=re.IGNORECASE)
136             redactions_count += len(matches)
137             logger.debug(f"Anonimizado {len(matches)} instancias de
138                          {data_type}")
139
140     # Alertar sobre termos sensiveis
141     for term in self.sensitive_terms:
142         if term.lower() in anonymized.lower():
143             logger.warning(f"Termo sensivel encontrado: {term}")
144
145     return anonymized, redactions_count
146
147 def anonymize_metadata(self, metadata: Dict) -> Tuple[Dict, int]:
148     """Anonimizar metadata"""
149     anonymized_metadata = {}
150     redactions_count = 0
151
152     for key, value in metadata.items():
153         if isinstance(value, str):
154             anonymized_value, count = self.anonymize_text(value)
155             anonymized_metadata[key] = anonymized_value
156             redactions_count += count
157         elif isinstance(value, dict):
158             anonymized_value, count = self.anonymize_metadata(value)

```

```

157         anonymized_metadata[key] = anonymized_value
158         redactions_count += count
159     elif isinstance(value, list):
160         anonymized_list = []
161         for item in value:
162             if isinstance(item, str):
163                 anon_item, count = self.anonymize_text(item)
164                 anonymized_list.append(anon_item)
165                 redactions_count += count
166             else:
167                 anonymized_list.append(item)
168         anonymized_metadata[key] = anonymized_list
169     else:
170         anonymized_metadata[key] = value
171
172     return anonymized_metadata, redactions_count
173
174
175 class ProcessingRecord:
176     """Registro de atividades de tratamento (Art. 30 GDPR)"""
177
178     def __init__(self, output_dir: str):
179         self.output_dir = Path(output_dir)
180         self.record_file = self.output_dir / "processing_record.json"
181         self.record = {
182             'controller': {
183                 'name': 'Instituto Politecnico de Braganca',
184                 'contact': 'dpo@ipb.pt',
185                 'address': 'Campus de Santa Apolonia, 5300-253 Braganca'
186             },
187             'dpo': {
188                 'name': 'Data Protection Officer',
189                 'contact': 'dpo@ipb.pt'
190             },
191             'processing_activity': {

```

```
192         'name': 'Web Crawling para Base de Conhecimento Chatbot'
193         ,
194         'purpose': 'Criacao de base de conhecimento para sistema
195         chatbot de atendimento ao cliente',
196         'legal_basis': 'Interesse legitimo (Art. 6(1)(f) GDPR)',
197         'legitimate_interest': 'Melhoria do servico ao cliente
198         atraves de chatbot inteligente'
199     },
200     'data_categories': [
201         'Informacoes publicas de produtos (nomes, descricoes,
202         especificacoes)',
203         'Precos e disponibilidade',
204         'Imagens de produtos',
205         'Estrutura de website (links, categorias)'
206     ],
207     'data_subjects': [
208         'Nao aplicavel - apenas dados publicos de produtos'
209     ],
210     'recipients': [
211         'Sistema interno de chatbot',
212         'Equipa de desenvolvimento e manutencao'
213     ],
214     'security_measures': [
215         'Anonimizacao automatica de dados pessoais detetados',
216         'Controle de acesso aos dados crawleados',
217         'Logs de auditoria de acesso',
218         'Backups encriptados',
219         'Rate limiting para evitar sobrecarga'
220     ],
221     'data_protection_impact_assessment': {
222         'required': False,
223         'reason': 'Tratamento limitado a dados publicos com
224         anonimizacao'
225     },
226 }
```

```

221         'international_transfers': 'Nao aplicavel - dados
222             processados localmente',
223         'created_at': datetime.now().isoformat(),
224         'last_updated': datetime.now().isoformat()
225     }
226
227     def save_record(self):
228         """Salvar registro de atividades"""
229         self.output_dir.mkdir(exist_ok=True)
230         self.record['last_updated'] = datetime.now().isoformat()
231
232         with open(self.record_file, 'w', encoding='utf-8') as f:
233             json.dump(self.record, f, indent=2, ensure_ascii=False)
234
235         logger.info(f"Registro de atividades GDPR salvo em {self.
236             record_file}")
237
238     def update_statistics(self, stats: CrawlStats):
239         """Atualizar estatisticas no registro"""
240         self.record['statistics'] = {
241             'total_pages_crawled': stats.successful_pages,
242             'total_anonymizations': stats.anonymized_items,
243             'domains_processed': list(stats.pages_per_domain.keys()),
244             'last_crawl': datetime.now().isoformat()
245         }
246         self.save_record()
247
248     class RobotsChecker:
249         """Gerenciador de robots.txt"""
250
251         def __init__(self, user_agent: str):
252             self.user_agent = user_agent
253             self.parsers: Dict[str, RobotFileParser] = {}
254             self.cache: Dict[str, bool] = {}

```

```

254     async def can_fetch(self, url: str, session: aiohttp.ClientSession)
        -> bool:
255         """Verificar se pode fazer fetch da URL"""
256         parsed = urlparse(url)
257         base_url = f"{parsed.scheme}://{parsed.netloc}"
258
259         cache_key = f"{base_url}:{url}"
260         if cache_key in self.cache:
261             return self.cache[cache_key]
262
263         if base_url not in self.parsers:
264             parser = RobotFileParser()
265             robots_url = urljoin(base_url, '/robots.txt')
266
267             try:
268                 async with session.get(robots_url, timeout=aiohttp.
                    ClientTimeout(total=10)) as response:
269                     if response.status == 200:
270                         content = await response.text()
271                         parser.parse(content.splitlines())
272                     else:
273                         parser.parse([])
274             except Exception as e:
275                 logger.warning(f"Erro ao buscar robots.txt de {base_url}
                    ): {e}")
276                 parser.parse([])
277
278             self.parsers[base_url] = parser
279
280         can_fetch = self.parsers[base_url].can_fetch(self.user_agent,
            url)
281         self.cache[cache_key] = can_fetch
282         return can_fetch
283
284     #...resto do codigo

```

Listing C.1: Crawler.py

C.2 JWT.py

```
1
2 from datetime import datetime, timedelta
3 from typing import Optional
4
5 from jose import JWTError, jwt
6 from passlib.context import CryptContext
7 from fastapi import Depends, HTTPException, status, Request
8 from fastapi.security import HTTPBearer, HTTPAuthorizationCredentials
9 from server.models.user_models import User, UserRole
10 from server.config.settings import get_settings
11 import hashlib
12
13 settings = get_settings()
14
15 # Password
16 pwd_context = CryptContext(schemes=["bcrypt"], deprecated="auto",
17                             bcrypt__rounds=12)
18
19 # Security
20 security = HTTPBearer()
21
22 def _prepare_password(password: str) -> str:
23     """Palavra-Passe com SHA256"""s
24     return hashlib.sha256(password.encode('utf-8')).hexdigest()
25
26 def verify_password(plain_password: str, hashed_password: str) -> bool:
```

```

27     prepared_password = _prepare_password(plain_password)
28     return pwd_context.verify(prepared_password, hashed_password)
29
30 def get_password_hash(password: str) -> str:
31     prepared_password = _prepare_password(password)
32     return pwd_context.hash(prepared_password)
33
34 def create_access_token(data: dict, expires_delta: Optional[timedelta] =
    None):
35     to_encode = data.copy()
36     if expires_delta:
37         expire = datetime.utcnow() + expires_delta
38     else:
39         expire = datetime.utcnow() + timedelta(minutes=settings.
            jwt_access_token_expire_minutes)
40
41     to_encode.update({"exp": expire})
42     encoded_jwt = jwt.encode(to_encode, settings.jwt_secret_key,
            algorithm=settings.jwt_algorithm)
43     return encoded_jwt
44
45
46 async def get_token_from_request(request: Request) -> Optional[str]:
47     """Extrair token do cabeçalho de autorizacao ou cookie"""
48
49     auth_header = request.headers.get("Authorization")
50     if auth_header and auth_header.startswith("Bearer "):
51         return auth_header.replace("Bearer ", "")
52
53     token = request.cookies.get("access_token")
54     if token:
55         return token
56
57     return None
58

```

```

59 async def get_current_user(request: Request, credentials: Optional[
    HTTPAuthorizationCredentials] = Depends(security)) -> User:
60     credentials_exception = HTTPException(
61         status_code=status.HTTP_401_UNAUTHORIZED,
62         detail="Could not validate credentials",
63         headers={"WWW-Authenticate": "Bearer"},
64     )
65
66     token = None
67     if credentials:
68         token = credentials.credentials
69     else:
70         token = await get_token_from_request(request)
71
72     if not token:
73         raise credentials_exception
74
75     try:
76         payload = jwt.decode(token, settings.jwt_secret_key, algorithms
            =[settings.jwt_algorithm])
77         username: str = payload.get("sub")
78         if username is None:
79             raise credentials_exception
80     except JWTErrror as e:
81         print(f"JWT Error: {e}")
82         raise credentials_exception
83
84     user = await User.find_one(User.username == username)
85     if user is None:
86         raise credentials_exception
87
88     return user
89
90 async def get_current_active_user(request: Request, credentials:
    Optional[HTTPAuthorizationCredentials] = Depends(security)) -> User:

```

```

91     current_user = await get_current_user(request, credentials)
92     if not current_user.is_active:
93         raise HTTPException(status_code=400, detail="Inactive user")
94     return current_user
95
96 def require_role(*allowed_roles: UserRole):
97     """Verificar se o utilizador tem funcao"""
98     async def role_checker(request: Request, credentials: Optional[
99         HTTPAuthorizationCredentials] = Depends(security)) -> User:
100         current_user = await get_current_active_user(request,
101             credentials)
102         if current_user.role not in allowed_roles:
103             raise HTTPException(
104                 status_code=status.HTTP_403_FORBIDDEN,
105                 detail=f"Not enough permissions. Required role: {
106                     allowed_roles}")
107         return current_user
108     return role_checker

```

Listing C.2: Crawler.py

C.3 Base.html

```

1 <!DOCTYPE html>
2 <html lang="pt">
3 <head>
4     <meta charset="UTF-8">
5     <meta name="viewport" content="width=device-width, initial-scale=1.0">
6     <title>{% block title %}Backoffice{% endblock %}</title>
7
8

```

```

9     <link href="https://cdn.jsdelivr.net/npm/bootstrap@5.3.0/dist/css/bootstrap
↳ .min.css" rel="stylesheet">
10
11
12     <link rel="stylesheet" href="https://cdn.jsdelivr.net/npm/bootstrap-icons@1
↳ .11.0/font/bootstrap-icons.css">
13
14
15     <script src="https://cdn.jsdelivr.net/npm/chart.js@4.4.0/dist/chart.umd.js"
↳ ></script>
16
17     <!-- Styles -->
18     <link rel="stylesheet" href="{ url_for('static', path='/styles.css') }">
19
20
21     {% block styles %}{% endblock %}
22 </head>
23 <body>
24     <!-- Loading Overlay -->
25     <div id="loadingOverlay">
26         <div class="text-center">
27             <div class="spinner-border text-light" role="status" style="width:
↳ 3rem; height: 3rem;">
28                 <span class="visually-hidden">Loading...</span>
29             </div>
30             <p class="text-light mt-3">A carregar...</p>
31         </div>
32     </div>
33
34     <div class="container-fluid">

```

```

35     <div class="row">
36         <!-- Sidebar -->
37         <nav class="col-md-2 d-md-block sidebar p-0">
38             <div class="p-4">
39                 <h4 class="text-white mb-4">
40                     <i class="bi bi-building"></i>RV Backoffice
41                 </h4>
42                 <hr class="text-white-50">
43                 <ul class="nav flex-column">
44                     <li class="nav-item">
45                         <a class="nav-link {% block nav_dashboard %}{%
↳ endblock %}" href="/backoffice/dashboard">
46                             <i class="bi bi-speedometer2"></i> Dashboard
47                         </a>
48                     </li>
49                     <li class="nav-item">
50                         <a class="nav-link {% block nav_sessions %}{%
↳ endblock %}" href="/backoffice/sessions">
51                             <i class="bi bi-chat-dots"></i> Sessoes
52                         </a>
53                     </li>
54                     <li class="nav-item">
55                         <a class="nav-link {% block nav_documents %}{%
↳ endblock %}" href="/backoffice/documents">
56                             <i class="bi bi-file-earmark-text"></i>
↳ Documentos
57                         </a>
58                     </li>
59                     <li class="nav-item">

```

```

60         <a class="nav-link {% block nav_analytics %}{%
↳ endblock %}" href="/backoffice/analytics">
61             <i class="bi bi-graph-up"></i> Analytics
62         </a>
63     </li>
64     <li class="nav-item">
65         <a class="nav-link {% block nav_users %}{% endblock
↳ %}" href="/backoffice/users">
66             <i class="bi bi-people"></i> Utilizadores
67         </a>
68     </li>
69 </ul>
70
71 <hr class="text-white-50 mt-4">
72
73 <div class="text-white-50 p-3">
74     <small>
75         <i class="bi bi-person-circle"></i>
76         <span id="currentUser">{{ user.username if user
↳ else 'Admin' }}</span>
77     </small>
78     <br>
79     <button class="btn btn-sm btn-outline-light mt-2 w-100"
↳ onclick="logout()">
80         <i class="bi bi-box-arrow-right"></i> Sair
81     </button>
82 </div>
83 </div>
84 </nav>
85

```

```

86     <!-- Main Content -->
87     <main class="col-md-10 ms-sm-auto px-md-4 py-4">
88         {% block content %}{% endblock %}
89     </main>
90 </div>
91 </div>
92
93
94 <script src="https://cdn.jsdelivr.net/npm/bootstrap@5.3.0/
95 dist/js/bootstrap.bundle.min.js"></script>
96
97 <!-- Scripts -->
98 <script>
99     const API_BASE_URL = 'http://localhost:8000';
100     let authToken = localStorage.getItem('authToken');
101
102     // Check authentication
103     if (!authToken) {
104         window.location.href = '/backoffice/login';
105     }
106
107     // API Call Helper
108     async function apiCall(endpoint, method = 'GET', data = null) {
109         showLoading();
110
111         const options = {
112             method: method,
113             headers: {
114                 'Content-Type': 'application/json',
115                 'Authorization': 'Bearer ${authToken}'

```

```

116         }
117     };
118
119     if (data) {
120         options.body = JSON.stringify(data);
121     }
122
123     try {
124         const response = await fetch(`${API_BASE_URL}${endpoint}`,
↪ options);
125
126         if (response.status === 401) {
127             logout();
128             return null;
129         }
130
131         return await response.json();
132     } catch (error) {
133         console.error('API Error:', error);
134         showAlert('Erro ao comunicar com o servidor', 'danger');
135         return null;
136     } finally {
137         hideLoading();
138     }
139 }
140
141
142 function showLoading() {
143     document.getElementById('loadingOverlay').style.display = 'flex';
144 }

```

```

145
146     function hideLoading() {
147         document.getElementById('loadingOverlay').style.display = 'none';
148     }
149
150     function showAlert(message, type = 'info') {
151         const alertDiv = document.createElement('div');
152         alertDiv.className = 'alert alert-${type} alert-dismissible fade
↪ show position-fixed top-0 start-50 translate-middle-x mt-3';
153         alertDiv.style.zIndex = '10000';
154         alertDiv.innerHTML = '
155             ${message}
156             <button type="button" class="btn-close" data-bs-dismiss="alert"
↪ ></button>
157         ';
158         document.body.appendChild(alertDiv);
159
160         setTimeout(() => alertDiv.remove(), 5000);
161     }
162
163     function logout() {
164         localStorage.removeItem('authToken');
165         window.location.href = '/backoffice/login';
166     }
167
168     function formatDate(dateString) {
169         return new Date(dateString).toLocaleString('pt-PT');
170     }
171
172     function formatBytes(bytes) {

```

```
173     if (bytes === 0) return '0 Bytes';
174     const k = 1024;
175     const sizes = ['Bytes', 'KB', 'MB', 'GB'];
176     const i = Math.floor(Math.log(bytes) / Math.log(k));
177     return Math.round(bytes / Math.pow(k, i) * 100) / 100 + ' ' + sizes
    ↪ [i];
178     }
179     </script>
180
181     <!-- Page Specific Scripts -->
182     {% block scripts %}{% endblock %}
183 </body>
184 </html>
```

Listing C.3: base.html

Apêndice D

Artigos

D.1 ICTIS

Chatbots, to transform the user experience at 3D environments: a systematic literature review

Telmo Sampaio², Pedro Filipe Oliveira¹, and Paulo Matos¹

¹ CeDRI, SusTEC, Instituto Politécnico de Bragança, 5300-253 Bragança, Portugal
{poliveira, pmatos}@ipb.pt

² Polytechnic University of Bragança, 5300-253 Bragança, Portugal
a48561@alunos.ipb.pt

Abstract. This paper proposes the implementation and evaluation of an intelligent environment system designed to enhance the management of comfort preferences in a residence setting on campus. With the growing importance of personalized comfort in shared living spaces, the integration of smart technologies offers promising solutions to meet individual needs while optimizing energy efficiency. Leveraging sensors, actuators, and machine learning algorithms, the proposed system aims to dynamically adapt environmental conditions such as temperature, lighting, and ventilation based on occupants' preferences. Through a combination of user-centric design, data analytics, and automation, the intelligent environment offers a seamless and intuitive interface for residents to interact with and customize their living environment. Furthermore, the paper discusses the practical challenges and opportunities associated with deploying such a system in a campus residence, including privacy concerns, user acceptance, and scalability. The effectiveness of the proposed solution is evaluated through energy consumption analysis, and feedback mechanisms, highlighting its potential to enhance comfort, well-being, and sustainability in residential settings. Ultimately, this research contributes to the advancement of smart living technologies and informs the design of future intelligent environments tailored to the needs of campus residences and similar shared living spaces.

Keywords: chatbot, 3d-environments, user-experience, nlp

1 Introduction

As digital transformation accelerates, the convergence of artificial intelligence and three-dimensional interfaces redefines the limits of online interaction [4]. Chatbots, once limited to basic text-based assistance, are evolving into sophisticated virtual entities capable of guiding users through immersive spatial experiences. This integration of conversational AI with 3D visualization marks a significant change in the way companies interact with consumers and how users experience digital products [1].

The integration of natural language processing with three-dimensional environments opens up new opportunities for intuitive navigation, personalized

D.2 ARTIIS

Smart Tourism through chatbots: A practical implementation for Northern Portugal

Telmo Sampaio¹, Pedro Filipe Oliveira²[0000-0002-2848-1606], and Carlos R. Cunha³[0000-0003-3085-1562]

¹ Polytechnic University of Bragança, Campus de Santa Apolónia, 5300-253 Bragança, Portugal
a48561@alunos.ipb.pt

² CeDRI, SusTEC, Instituto Politécnico de Bragança, 5300-253 Bragança, Portugal
poliveira@ipb.pt

³ CeDRI, SusTEC, Instituto Politécnico de Bragança, 5300-253 Bragança, Portugal
crc@ipb.pt

Abstract. The integration of artificial intelligence in tourism has revolutionized visitor engagement by offering interactive, personalized, and real-time information services. This paper presents the development of an intelligent tourist chatbot system tailored for Mirandela, a culturally rich destination in northern Portugal. The system combines advanced web scraping techniques and natural language processing to deliver relevant and context-aware recommendations to tourists. A specialized web crawler was developed to extract up-to-date content from the official VisitMirandela website, which is then processed and stored using a hybrid architecture integrating markdown and vector databases. User queries are handled through a structured response pipeline utilizing OpenAI's GPT-4o Mini model to generate semantically accurate and sourced answers. The system was evaluated through user-based question testing, achieving high accuracy in most scenarios while revealing opportunities for enhancement, particularly in providing named recommendations. The results highlight the potential of AI chatbots to enrich the visitor experience, improve access to local information, and reduce the dependence on multiple digital platforms. Furthermore, the paper addresses key challenges such as preserving human interaction, ensuring data reliability, and maintaining cultural authenticity. The proposed architecture demonstrates a scalable and effective approach for implementing AI-driven conversational systems in local tourism strategies, contributing to the digital transformation of the sector. Future work includes expanding the knowledge base, incorporating multimodal interactions, and enhancing emotional intelligence in chatbot responses.

Keywords: chatbot · artificial-intelligence · tourism · web-scraping · natural language processing

D.3 ICEET

3D-Chatbot: A Intelligent Conversational Agent Integrating RAG, NLP, And VR For Commercial Stores

Telmo Sampaio*, Pedro Filipe Oliveira†, Paulo Matos‡

Instituto Politécnico de Bragança, Campus de Santa Apolónia, 5300-253 Bragança, Portugal,
Research Centre in Digitalization and Intelligent Robotics (CeDRI),

Instituto Politécnico de Bragança, Campus de Santa Apolónia, 5300-253 Bragança, Portugal

Email:*a48561@alunos.ipb.pt, †poliveira@ipb.pt, ‡pmatos@ipb.pt

†Corresponding author

Abstract—This paper presents a comprehensive 3D chatbot system that combines advanced AI capabilities with practical commercial applications in immersive virtual reality environments. The system integrates natural language processing, retrieval-augmented generation (RAG) architectures, and real-time speech processing to create seamless customer inquiry experiences. The implementation combines four main components: an intelligent data acquisition pipeline that transforms over 1,000 unstructured web documents into AI-ready knowledge bases; an advanced RAG system utilizing OpenAI’s text-embedding-3-small model and Pinecone vector database for semantic search; multimodal speech processing with real-time speech-to-text conversion and lip-sync animation; and a Unity-based VR environment with professional avatar interactions. The results demonstrate exceptional performance with a 94.6% success rate across 65 test interactions, achieving response times under 100 milliseconds for knowledge retrieval and completing full voice response cycles in less than 3 seconds. The system maintains conversation context through a six-step interaction flow while accessing external knowledge sources, with 96.9% contextually appropriate responses and 92.3% conversation continuity in multi-turn dialogues. Performance optimizations ensure responsive operation under variable load conditions, demonstrating the viability of intelligent virtual assistants in three-dimensional commercial spaces.

Index Terms—3D-chatbot, virtual reality, RAG, NLP, conversational AI

I. INTRODUCTION

The evolution of conversational interfaces has reached a turning point with the emergence of three-dimensional chatbots, representing a fundamental paradigm shift in human-computer interaction. Recent advances in generative artificial intelligence have demonstrated remarkable potential for creating sophisticated virtual assistants capable of supporting complex interactions in immersive virtual reality environments [1]. The integration of large language models with spatial computing technologies enables the development of intelligent agents that can understand context, maintain conversational

coherence, and provide accurate information while operating in three-dimensional spaces.

The integration of multimodal interaction capabilities has emerged as a critical factor for the success of 3D conversational agents. Contemporary research shows that effective embedded conversational agents must seamlessly combine textual dialogue, facial expression recognition, and real-time emotional processing to create compelling user experiences [2]. The architectural challenges of implementing large language models in avatar-based chatbots have been thoroughly investigated, revealing both opportunities and technical limitations. Highlighting the complexity of creating open-domain avatar chatbots that can engage users in natural and contextually appropriate conversations while maintaining the visual and behavioral consistency expected of embodied agents [3].

This work addresses these limitations by presenting a comprehensive approach to developing an intelligent 3D chatbot system designed specifically for commercial applications. Our implementation integrates advanced natural language processing, retrieval-augmented generation (RAG) architectures, immersive virtual reality environments, and real-time speech processing to create a seamless customer consultation experience.

A. Research Objectives

The primary objective of this research is to develop and validate a comprehensive 3D chatbot system that bridges the gap between advanced AI capabilities and practical commercial applications.

Our specific research objectives:

- 1) **Intelligent Knowledge Integration:** Develop an efficient data acquisition and processing pipeline that transforms unstructured web content into AI-ready structured knowledge bases capable of supporting sophisticated query processing.
- 2) **Advanced RAG system implementation:** Design and implement a retrieval-augmented generation system that combines vector-based semantic search with large language models to provide accurate and contextually relevant responses.

D.4 ICETC

A Chatbot to Enhance the Portuguese National Tourism Website

Telmo Sampaio*, Pedro Filipe Oliveira†, Paulo Matos‡

Research Centre in Digitalization and Intelligent Robotics (CeDRI),

Laboratório Associado para a Sustentabilidade e Tecnologia em Regiões de Montanha (SusTEC),

Instituto Politécnico de Bragança, Campus de Santa Apolónia, 5300-253 Bragança, Portugal

Email: *a48561@alunos.ipb.pt, †poliveira@ipb.pt, ‡pmatos@ipb.pt

†Corresponding author

Abstract—The digital transformation of tourism in Portugal has led to the integration of AI-driven solutions to improve visitor engagement. This paper presents the development and implementation of a chatbot for the Portuguese national tourism website, VisitPortugal. The chatbot utilizes natural language processing to provide real-time, personalized assistance across multiple languages, including European languages (Portuguese, English, French, Spanish) and East Asian languages (Traditional Chinese), demonstrating robust cross-linguistic semantic understanding. The architecture incorporates a web crawler to collect and process tourism-related data, a vector database for efficient content retrieval, and GPT-4o Mini for generating contextually relevant responses. By filling gaps in tourism services, the chatbot improves the user experience, increases accessibility, and reinforces Portugal’s position as a technologically advanced tourist destination. The results demonstrate high accuracy in the responses, reinforcing the effectiveness of the chatbot in facilitating seamless travel planning.

Index Terms—smart-tourism, chatbot, llm, web-crawler

I. INTRODUCTION

Portugal’s tourism landscape is undergoing a digital transformation, with conversational artificial intelligence emerging as a cornerstone of modern visitor engagement. The implementation of an AI (Artificial Intelligence)-powered chatbot for VisitPortugal represents a strategic reinforcement of the country’s tourism infrastructure, responding to the evolving expectations of contemporary travelers, and maintaining Portugal’s competitive edge in the global tourism market.

This innovative virtual assistant acts as a digital ambassador for Portugal, providing round-the-clock support that transcends the traditional limitations of customer service. Using natural language processing capabilities, the chatbot provides personalized recommendations and instant access to information, facilitating seamless travel planning experiences for visitors exploring Portugal’s diverse attractions, from historic sites and cultural landmarks to coastal retreats and urban destinations [1].

This AI-powered chatbot addresses the growing demand for instant, multilingual tourist information. As travelers increasingly rely on mobile technology for travel planning and

destination orientation, the VisitPortugal chatbot bridges a critical service gap, immediate assistance regardless of time zones or staff availability [2]. This capability is particularly valuable during peak tourist periods, when human resources face significant constraints. In addition to operational efficiency, the implementation of the chatbot demonstrates Portugal’s commitment to embracing technological advances in the provision of tourism services. The system’s ability to process natural language queries allows visitors to obtain contextually relevant information on accommodation options, local customs, transportation logistics, and regulatory requirements through intuitive conversational interfaces [3].

As the global tourism sector continues its technological evolution, Portugal’s investment in conversational AI positions the country as a forward-thinking destination that balances traditional hospitality with digital innovation. The chatbot not only simplifies the operational aspects of tourism management, but also reinforces Portugal’s brand identity as an accessible, visitor-centric destination in an increasingly competitive international market.

A. Research Objectives

To bridge the gap between generic AI systems and domain-specific tourism information needs, this article presents a comprehensive chatbot architecture specifically designed for the Portuguese national tourism website. Our solution addresses the limitations of general-purpose AI by integrating official tourism data with culturally aware conversational capabilities. The proposed system consists of four main research contributions:

- 1) A specialized data acquisition system that automatically extracts and processes Portuguese tourism content through an intelligent web crawler.
- 2) A vector-based real-time retrieval system, optimized specifically for queries related to tourism and preservation of cultural context.
- 3) A culturally adaptable response generation framework that provides specific recommendations for Portugal with citations from official sources.
- 4) A comprehensive evaluation methodology to evaluate the accuracy of the tourism chatbot in various types of queries.

The authors are grateful to the Foundation for Science and Technology (FCT, Portugal) for financial support through national funds FCT/MCTES (PIDDAC) to CeDRI (UIDB/05757/2020 and UIDP/05757/2020) and SusTEC (LA/P/0007/2021).