

Chatbot para apoio inteligente à pesquisa de conteúdos na Internet

Susana Maria Germano Pereira

Dissertação para obtenção do Grau de Mestre em
Informação e Sistemas Empresariais

Orientadores: Prof. José Henrique Pereira São Mamede
Prof. Vítor Manuel Pereira Duarte dos Santos

Júri

Presidente: Prof. Miguel Leitão Bignolas Mira da Silva
Orientador: Prof. José Henrique Pereira São Mamede
Vogal: Prof. Hugo Alexandre Paredes Guedes da Silva

Novembro de 2021

Agradecimentos

À minha família, em especial à minha mãe e irmã Isabel, pelo apoio incondicional.

Aos meus orientadores pela sua disponibilidade, conselhos e incentivos.

A todos os professores do MISE pela dedicação e profissionalismo.

A todos os que, de uma forma ou de outra, contribuíram para a realização deste trabalho.

Resumo

Nas últimas décadas, a Internet tornou-se a principal ferramenta de comunicação e divulgação de informação em todo o mundo. O aumento do número de utilizadores levou a um aumento exponencial da quantidade de informação partilhada na *web*, fortemente impulsionado pela popularização das redes sociais. A informação tornou-se, por isso, mais imediata, volátil, e cada vez mais difícil de rastrear e validar.

Num mundo em constante mudança, as organizações devem ser capazes de acompanhar e prever tendências futuras de modo a poderem permanecer relevantes e, desta forma, garantir a sua sobrevivência. Nos últimos anos também tem havido um interesse crescente por *chatbots*, motivado por avanços significativos nas áreas da inteligência artificial e aprendizagem máquina.

O objetivo deste estudo é propor uma arquitetura de sistema de *chatbot* que permita um suporte inteligente para pesquisa de conteúdos na Internet. A utilização de soluções de inteligência artificial permite que esses agentes conversacionais simulem conversas humanas usando técnicas de processamento de linguagem natural, transformando-os em valiosos assistentes virtuais capazes de realizar uma variedade de tarefas de forma automatizada. Com essas funções, o *chatbot* torna-se uma ferramenta prática e personalizável capaz de recolher, de forma natural, os temas de interesse do utilizador e auxiliar não só na realização de pesquisas mais eficientes na *web* como também na recuperação de informação de sítios *web* de forma mais rápida.

Palavras-chave: Chatbots, Pesquisa de Informação na Internet, Processamento de Linguagem Natural, Recuperação de Informação, Gestão de Conteúdos Web.

Abstract

Over the last decades, the Internet has become the leading tool for communication and dissemination of information worldwide. The increase in the number of users has led to a colossal increase in the amount of information shared on the web, strongly driven by the popularization of social networks. Information has thus become more immediate, volatile, and increasingly difficult to track and validate.

In a constantly changing world, organizations must be able to track and predict future trends to be able to remain relevant and ensure their survival. In recent years there has also been a growing interest in *chatbots*, motivated by significant advances in the areas of Artificial Intelligence (AI) and Machine Learning.

The aim of this study is to design a chatbot that allows intelligent support for web content search. The use of AI solutions allows these conversational agents to simulate human conversations using Natural Language Processing, thus transforming them into valuable virtual assistants capable of performing a variety of tasks in an automated way. With these functions, the chatbot becomes a practical and customizable tool to collect, in a natural way, the topics of interest of the user and help conducting more efficient web searches and retrieve information from websites in a faster way.

Keywords: Chatbots, Web Search, Natural Language Processing, Information Retrieval, Web Content Management

Índice

Agradecimentos	ii
Resumo	iii
Abstract	iv
Índice	v
Índice de Figuras	vii
Índice de tabelas	ix
Lista de siglas	x
1 Introdução	1
1.1 <i>Problema de investigação</i>	2
1.2 <i>Objetivos</i>	3
1.3 <i>Estrutura do documento</i>	4
2 Metodologia	5
2.1 <i>Design Science Research</i>	5
2.2 <i>Aplicação da metodologia</i>	6
3 Revisão bibliográfica	7
3.1 <i>Gestão de conteúdos</i>	7
3.1.1 <i>Recuperação de Informação</i>	9
3.1.2 <i>Pesquisa de conteúdos na Internet</i>	9
3.1.3 <i>Desafios atuais na pesquisa de informação na web</i>	10
3.2 <i>Inteligência Artificial</i>	10
3.2.1 <i>Aprendizagem Máquina (Machine Learning)</i>	11
3.2.2 <i>Aprendizagem Profunda (Deep Learning)</i>	13
3.3 <i>Ferramentas inteligentes de apoio pessoal</i>	13
3.3.1 <i>Chatbots</i>	14
3.3.2 <i>Processamento de Linguagem Natural</i>	15
3.4 <i>Utilização de chatbots para pesquisar informação na Internet</i>	16
3.4.1 <i>Diferenças entre motores de busca e chatbots para a pesquisa de informação</i>	16
3.4.2 <i>Desafios no desenvolvimento de chatbots para pesquisa de informação na web</i>	17
3.4.3 <i>Frameworks e implementações de chatbots para pesquisa de conteúdos em sítios web</i> ..	18
3.5 <i>Discussão</i>	19
3.5.1 <i>Oportunidades</i>	19
3.5.2 <i>Desafios</i>	20
4 Proposta de modelo conceptual	21
4.1 <i>Análise de utilizadores e tarefas</i>	21
4.1.1 <i>Caracterização da amostra</i>	21
4.1.2 <i>Hábitos de pesquisa</i>	23

4.1.3 Utilização de <i>chatbots</i>	24
4.1.4 Discussão	27
4.2 <i>Análise de requisitos</i>	29
4.2.1 Requisitos funcionais.....	29
4.2.2 Requisitos não funcionais.....	30
4.3 <i>Casos de uso</i>	30
4.3.1 Caso de uso 1: Obter notícias	31
4.3.2 Caso de uso 2: Efetuar registo.....	32
4.3.3 Caso de uso 3: Efetuar login	33
4.3.4 Caso de uso 4: Conversa casual	34
4.4 <i>Diagrama de processos</i>	35
4.5 <i>História de utilizador e storyboard</i>	35
4.6 <i>Arquitetura</i>	37
5 Desenvolvimento e implementação	39
5.1 <i>Desenvolvimento do protótipo</i>	39
5.1.1 Contexto	39
5.1.2 Base de dados.....	40
5.1.3 API externa.....	40
5.1.4 Dialogflow	41
5.2 <i>Desenho e treino do agente de conversação</i>	42
5.3 <i>Implementação</i>	45
6 Avaliação	49
6.1 <i>Avaliação do protótipo</i>	49
6.1.1 Caracterização da amostra	49
6.1.2 Tarefas, medidas de usabilidade e critérios de usabilidade	51
6.1.3 Resultados e discussão.....	52
7 Conclusão	55
7.1 <i>Limitações</i>	56
7.2 <i>Trabalho futuro</i>	56
Bibliografia	57
Anexo I	67
Anexo II	77
Anexo III	79

Índice de Figuras

Figura 1 - Fases da Metodologia adaptada de Peffers et al. (2007).....	6
Figura 2 – Representação esquemática de um SGCW. Adaptado de Boiko (2005)	8
Figura 3 – Diagrama que ilustra a relação entre inteligência artificial, aprendizagem máquina e aprendizagem profunda	11
Figura 4 – Tipos de aprendizagem máquina.....	12
Figura 5 - Género, idade e habilitações literárias dos inquiridos	22
Figura 6 - Dispositivos utilizados para aceder à Internet	22
Figura 7 - Frequência e resultados da pesquisa à primeira tentativa	23
Figura 8 - Ferramentas utilizadas para pesquisar conteúdos na Internet	23
Figura 9 - Aspetos valorizados numa ferramenta de pesquisa	24
Figura 10 - Utilização e frequência de utilização de <i>chatbots</i>	25
Figura 11 - Contextos de utilização de <i>chatbots</i>	25
Figura 12 - Dados dispostos a partilhar com um <i>chatbot</i>	26
Figura 13 - Frequência de utilização de chatbots para pesquisa de conteúdos	26
Figura 14 - Forma de utilização de <i>chatbots</i> para pesquisa de conteúdos	27
Figura 15 - Futuro dos <i>chatbots</i> como principal ferramenta de pesquisa de conteúdos na Internet ..	27
Figura 16 - Diagrama de caso de uso 1: "Obter notícias"	31
Figura 17 - Diagrama do caso de uso 2: "Efetuar registo"	32
Figura 18 - Diagrama do caso de uso 3: "Efetuar login"	33
Figura 19 - Diagrama de caso de uso 4 - "Conversa casual"	34
Figura 20 - Diagrama de processos	35
Figura 21 - <i>Storyboard</i>	36
Figura 22 - Arquitetura do sistema de <i>chatbot</i> proposto	38
Figura 23 - Parte do código de <i>fulfillment</i> utilizado	41
Figura 24 - Fluxo básico para correspondência de intenções e resposta ao utilizador (Google Cloud, 2021a).....	42
Figura 25 - Exemplo de entidades anotadas nas frases de treino.....	43
Figura 26 - Relação entre entidades e parâmetros	44
Figura 27 - Exemplo de contextos de entrada e saída	44
Figura 28 - Exemplo de resposta estática.....	45
Figura 29 - Integração do <i>chatbot</i> no Dialogflow Web Demo.....	46
Figura 30 - Integração do <i>chatbot</i> no Telegram.....	46

Figura 31 - Integração do <i>chatbot</i> no Dialogflow Messenger	47
Figura 32 - Exemplos de ecrãs do <i>chatbot</i> implementado num sítio <i>web</i>	48
Figura 33 - Género, idade e habilitações literárias dos voluntários.....	50
Figura 34 - Utilização prévia de <i>chatbots</i> pelos voluntários	50
Figura 35 - Conclusão das tarefas propostas no teste de avaliação do <i>chatbot</i>	52
Figura 36 - Classificação da eficácia do <i>chatbot</i> pelos voluntários	53
Figura 37 - Classificação da eficiência do <i>chatbot</i>	53
Figura 38 - Percentagem de utilizadores que recomendariam o <i>chatbot</i> a outras pessoas.....	54

Índice de tabelas

Tabela 1 - Especificação do caso de uso 1: "Obter notícias"	31
Tabela 2 - Especificação do caso de uso 2: "Efetuar registo"	32
Tabela 3 - Especificação do caso de uso 3: "Efetuar login"	33
Tabela 4 - Especificação do caso de uso 4: "Conversa casual"	34
Tabela 5 - Comparação de diferentes plataformas de desenvolvimento de <i>chatbots</i>	37
Tabela 6 - Secções dos sítios web de onde foram extraídas as notícias	39
Tabela 7 - Parâmetros utilizados no desenvolvimento do <i>chatbot</i>	40
Tabela 8 - Entidades e sinónimos.....	43
Tabela 9 - Tarefas completadas pelos voluntários.....	52

Lista de siglas

AI - Artificial Intelligence

AM - Aprendizagem Máquina

API - Application Programming Interface

CLN - Compreensão de Linguagem Natural

CRM – Customer Relationship Management

DSR - Design Science Research

DSRM - Design Science Research Methodology

FAQ – Frequently Asked Questions

GLN - Geração de Linguagem Natural

HTTP – Hypertext Transfer Protocol

IA - Inteligência Artificial

IRIS - Serviço de Reconhecimento de Intenções e Informação

ISO - International Organization for Standardization

IU - Interface do utilizador

JSON – JavaScript Object Notation

ML - Machine Learning

NLP - Natural Language Processing

OCR - Optimal Character Recognition

PLN - Processamento de Linguagem Natural

REST – Representational State Transfer

RI - Recuperação de Informação

SGC - Sistema de Gestão de Conteúdos

SGCW - Sistema de Gestão de Conteúdos Web

SRC - Sistema de Recomendação de Conteúdos

SRI - Sistema de Recuperação de Informação

TIC - Tecnologias da Informação e Comunicação

WCAG - Web Content Accessibility Guidelines

1 Introdução

Nas últimas décadas, a Internet tornou-se numa das mais importantes ferramentas de comunicação e divulgação de informação em todo o mundo. O aumento do número de utilizadores, que em 2021 corresponde a 65,6% da população mundial (The Internet World Stats, 2021), provocou um crescimento colossal da quantidade de informação partilhada na *web*, fortemente impulsionado pela popularização das redes sociais (Heidemann, Klier & Probst, 2012).

A Internet mudou radicalmente a forma como atuamos perante nós próprios e o mundo. A comunicação em rede possibilitou a criação de uma sociedade global na qual a partilha de informação e do conhecimento atingiram proporções infinitas. Para Castells (2007, p.16), a Internet é “um meio de comunicação que permite, pela primeira vez, a comunicação de muitos para muitos, em tempo escolhido e a uma escala global”. De facto, basta haver acesso à Internet para termos, na nossa posse, todo um mundo de conhecimento e saber, que nos permite ultrapassar quase todas as barreiras do dia a dia. Como ferramenta de pesquisa de informação e de transmissão de conhecimento, é cada vez mais indispensável na vida moderna (Alguliyev & Mahmudov, 2016).

As vantagens da Internet são muitas: rapidez, capacidade ilimitada de armazenamento de informação e, acima de tudo, a diminuição do uso excessivo de recursos naturais, como o papel (através da utilização do *email*, do *ebook* e da fotografia digital, em detrimento dos seus correspondentes analógicos) ou mesmo a gasolina (deixa de ser necessário realizar deslocações para se obter a informação de que se precisa).

Langville & Meyer (2012) identificaram as quatro principais propriedades que caracterizam a *web* como repositório de informação: é enorme, dinâmica, auto-organizada e hiperligada. No entanto, são estas mesmas características que colocam alguns dos principais desafios no que diz respeito à pesquisa da informação.

Todos os sítios *web* são compostos por múltiplos conteúdos *web*, termo que se pode referir tanto ao tipo da informação como ao item individual usado para transmitir essa informação (Tatar, de Amorim, Fdida & Antoniadis, 2014). Existem vários tipos de conteúdos *web*, sendo os mais comuns apresentados na forma de texto, imagem, áudio ou vídeo, que são geridos através de aplicações de gestão de conteúdos.

Hoje em dia, há uma enorme quantidade e diversidade de dados disponíveis na *web*, tornando-se cada vez mais difícil fazer pesquisas sobre assuntos específicos ou encontrar informação útil na Internet (Chau & Chen, 2003). Por outro lado, há um interesse cada vez maior na utilização de ferramentas dotadas de inteligência artificial (IA) para aumentar a produtividade e automatizar tarefas. Entre essas ferramentas encontram-se os *chatbots*, que são aplicações de *software* utilizadas para interagir numa conversação humana de forma natural. Estes assistentes virtuais podem ser integrados em diversos canais, como aplicações de mensagens instantâneas ou sítios *web*, e são já utilizados em diversas áreas, como finanças, educação, comércio, cuidados de saúde e bem-estar, indústria, viagens, entre

outras (Mercieca, 2019). Há também uma crescente oferta de soluções de serviços de *chatbot* baseados na nuvem (Gupta, Vijayakumar & Hathwar, 2020), que permitem o seu desenvolvimento e implementação em larga escala, como o IBM Watson Assistant, o Dialogflow da Google, o Microsoft Bot Framework ou o Amazon Lex.

1.1 Problema de investigação

Nos dias de hoje, a Internet encontra-se sobrecarregada de informação, cada vez mais dispersa e difícil de encontrar. A pesquisa de conteúdos é uma das atividades mais comuns realizadas na *web* e os motores de busca, em geral, fazem um bom trabalho na pesquisa de palavras-chave concretas. No entanto, muitas vezes estão dependentes de *rankings* e apoios comerciais e podem apresentar respostas tendenciosas. Para além disso, a pesquisa de termos ou expressões sem indicação de contexto ainda está longe de ser eficaz.

Por sua vez, os *chatbots* têm sido uma das aplicações de IA que mais facilmente se consegue integrar em sítios *web*. São ferramentas versáteis que, para além de diálogo, são capazes de realizar várias tarefas, entre as quais a pesquisa e recuperação de informação (RI).

Num mundo em constante transformação, as organizações têm de ser capazes de acompanhar e prever futuras tendências para se conseguirem manter relevantes e garantirem a sua sobrevivência (Laudon & Laudon, 2017). Para tornar mais rápidos e eficientes os processos de negócio e de tomada de decisão, é necessário desenvolver soluções de pesquisa avançadas que permitam filtrar a informação disponível na *web* e, ao mesmo tempo, contribuir para a melhoria dos processos associados à gestão de conteúdos.

No entanto, apesar dos muitos avanços já realizados no desenvolvimento de motores de busca para a *web*, em particular no que diz respeito às tecnologias de *web* semântica (Kumari & Thakur, 2019), **nem sempre é fácil efetuar pesquisas de forma eficiente e os utilizadores ainda sentem dificuldades no processo de pesquisa e recuperação da informação** (Marx, Khalili & Valdestilhas, 2017), principalmente devido aos desafios derivados dos grandes dados (Fang, Jia, Li, Li & Wu, 2017).

Há vários fatores que contribuem para essa dificuldade, dos quais se destacam os seguintes:

- Os utilizadores podem não conseguir formular as suas complexas necessidades de informação numa única consulta (Aliannejadi, Zamani, Crestani & Croft, 2019);
- Os utilizadores podem não estar familiarizados com o domínio onde se enquadra a pesquisa (Teng, Wen & Liu, 2012), as palavras-chave necessárias para efetuar a consulta, o objetivo ou as formas de o alcançar (Fafalios, Holzmann, Kasturia & Nedji, 2017);
- A ambiguidade inerente à compreensão da língua humana e a ausência de contexto, sobretudo em quantidades de texto mais curtas, como é o caso das consultas de pesquisa (Zhu & Iglesias, 2018);

- Desafios relacionados com sistemas de recomendação de conteúdos (SRC) (Azmy & Emam, 2013; Mohamed, Khafagy & Ibrahim, 2019) e com os recursos de preenchimento automático dos motores de busca (Hazem, Olteanu, Kazai, Diaz & Golebiewski, 2020);
- A existência de resultados tendenciosos, deturpados ou viciados, que dificultam a capacidade de avaliar a sua credibilidade (Kattenbeck & Elsweiler, 2019).

Por outro lado, há um interesse cada vez maior na utilização de ferramentas dotadas de IA para aumentar a produtividade e automatizar tarefas, entre as quais se encontram os *chatbots*. Estes sistemas já são utilizados a nível empresarial para ajudar nas tarefas de apoio ao cliente, vendas, finanças, *marketing*, recursos humanos, entre outras.

Apesar da qualidade das interações entre os utilizadores e os agentes conversacionais ainda ser um problema na realização de tarefas mais complexas, há um manifesto interesse pela utilização de *chatbots* e a avaliação destes sistemas é uma área de investigação ativa na comunidade de IA (Ruane et al., 2018). Estes sistemas conversacionais inteligentes de pergunta-resposta podem ajudar a clarificar as necessidades de informação dos utilizadores ao pedir de forma proativa que clarifiquem as suas questões (Aliannejadi et al., 2019) e, desta forma, devolver resultados mais precisos. Por isso, seria muito oportuno existirem ferramentas de baixo custo e fácil utilização que facilitassem a pesquisa de conteúdos na Internet nas situações em que os utilizadores não dominam o léxico e o contexto onde a pesquisa se enquadra.

O desenvolvimento de um *chatbot* para apoio à pesquisa inteligente de conteúdos na *web*, que pudesse ser facilmente implementado em diversos tipos de plataformas, poderá ajudar a complementar as ferramentas já existentes e contribuir para novas pesquisas nesta área. Desta forma, com este trabalho pretende-se responder à seguinte pergunta de investigação: **será viável e possível arquitetar e implementar um *chatbot* para apoiar a pesquisa de informação no espaço contido de um determinado sítio *web*?**

1.2 Objetivos

Para podermos obter uma resposta ao problema de investigação identificado na secção 1.1 do presente capítulo, foram definidos seis objetivos:

- **Objetivo 1:** Investigar os problemas associados à gestão de conteúdos *web*;
- **Objetivo 2:** Investigar as tecnologias de *chatbot* e os diferentes *frameworks* disponíveis;
- **Objetivo 3:** Contextualizar o trabalho de pesquisa com os conceitos de processamento de linguagem natural (PLN), aprendizagem profunda (*deep learning*), aprendizagem máquina (AM), recuperação de informação e outros conceitos relacionados;
- **Objetivo 4:** Desenhar a arquitetura do sistema proposto;

- **Objetivo 5:** Desenhar e treinar um agente de conversação;
- **Objetivo 6:** Testar o protótipo e avaliar os resultados.

1.3 Estrutura do documento

A estrutura do estudo que aqui se apresenta está delimitada de forma a dar resposta aos objetivos previamente formulados na secção 1.2 do presente capítulo.

No **primeiro capítulo**, é feita uma introdução ao tema em estudo, explicando os problemas da pesquisa de conteúdos na Internet, a necessidade de se encontrar formas alternativas para realizar essa tarefa e o interesse em usar ferramentas inteligentes como os *chatbots*.

No **segundo capítulo**, é apresentada a metodologia usada para o desenvolvimento da investigação. Também são descritos os procedimentos e a metodologia aplicada, todo o processo para a aplicação do protótipo e a recolha de dados.

No **terceiro capítulo**, apresenta-se a revisão da literatura, onde é feito um enquadramento da temática deste trabalho com a gestão de conteúdos, inteligência artificial e sistemas de *chatbot*. Nestas fontes de recolha, foram analisados trabalhos com diferentes abordagens sobre a utilização de *chatbots* para pesquisa de informação, fazendo-se uma breve discussão no final acerca dos resultados observados.

No **quarto capítulo**, é explicada em detalhe a solução proposta para resolver o problema identificado na secção 1.1 do presente capítulo, com a elucidação dos objetivos que se pretendem alcançar e os detalhes da solução proposta.

No **quinto capítulo**, é mostrado como a arquitetura proposta no capítulo 4 foi implementada.

No **sexto capítulo**, apresenta-se a avaliação e a interpretação dos resultados obtidos no contexto do presente estudo, as métricas adotadas e a respetiva análise.

No **sétimo capítulo**, são enunciadas as conclusões que foram retiradas deste trabalho, assim como a análise das limitações encontradas e sugestões para trabalhos futuros.

2 Metodologia

Neste capítulo descreve-se a metodologia escolhida para a realização deste trabalho de investigação e a forma como foi aplicada.

2.1 Design Science Research

A *Design Science Research* (DSR), ou investigação de conceção, é um paradigma muito usado em sistemas de informação que consiste no desenvolvimento de artefactos inovadores para solucionar problemas específicos dentro de um determinado domínio e, desta forma, contribuir para a produção de novo conhecimento científico (Hevner, March, Park & Ram, 2004). Uma vez que se pretende desenhar e implementar um artefacto (um *chatbot*), considerou-se que esta era a metodologia mais adequada para a sua elaboração, concretização e análise crítica dos resultados.

O artefacto é o elemento central da DSR e deverá ser bem testado, compreendido e documentado, de modo a assegurar a sua validade pragmática. A avaliação é uma fase fulcral do processo, na qual se demonstra a utilidade, qualidade e eficácia do artefacto, utilizando métodos de avaliação bem executados (Hevner et al., 2004).

Peppers, Tuunanen, Rothenberger & Chatterjee (2007) apresentam um modelo mental que sintetiza os principais passos para se proceder à realização da investigação e respetiva avaliação através da metodologia DSR (DSRM – *Design Science Research Methodology*). Este é um modelo de produção e apresentação dos resultados da investigação de conceção comumente aceite em sistemas de informação (Hevner & Chatterjee, 2010) e inclui as seguintes seis atividades sequenciais:

1. **Identificação do problema e motivação** – Nesta fase define-se um problema de investigação específico e justifica-se a importância e o valor da solução;
2. **Definição de objetivos** – Nesta fase são definidos os objetivos para solucionar o problema identificado;
3. **Desenho e desenvolvimento** – Esta fase corresponde à criação de um artefacto de investigação, com a definição da sua arquitetura e das funcionalidades desejadas;
4. **Demonstração** – Nesta fase demonstra-se a solução de uma ou mais instanciações do problema utilizando o artefacto;
5. **Avaliação** – Esta fase corresponde à observação e avaliação da performance do artefacto no contexto do problema, comparando-a com os objetivos definidos anteriormente. No caso de os resultados não corresponderem ao esperado, é possível iterar a fase de desenho e desenvolvimento;
6. **Comunicação** – Na última fase comunica-se o problema, a importância do artefacto e a sua utilidade junto da comunidade académica e científica, profissionais da área e outras partes interessadas.

O processo iterativo inerente ao DSR permite melhorar o sistema através da identificação de riscos e incertezas em cada iteração, o que resulta na redução de erros e na criação de soluções mais eficientes.

A Figura 1 apresenta a metodologia DSR aplicada a este projeto, especificando o trabalho realizado em cada uma das fases.

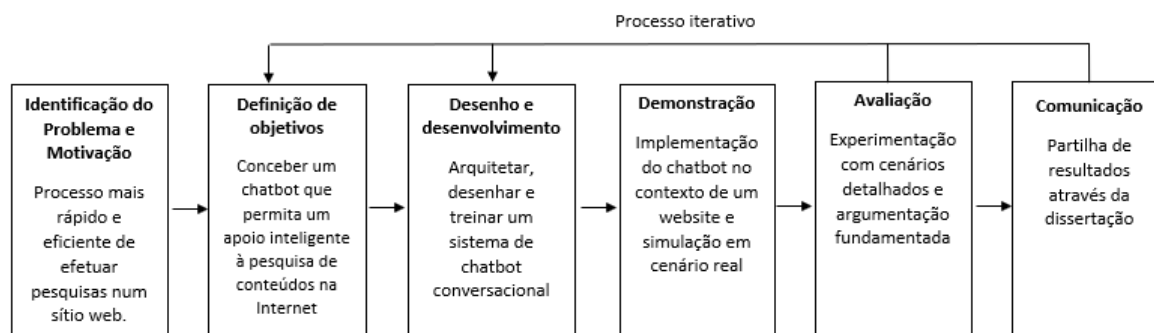


Figura 1 - Fases da Metodologia adaptada de Peffers et al. (2007)

2.2 Aplicação da metodologia

Com base no modelo DSRM, será aplicada a seguinte estratégia de investigação:

1. **Identificação do problema e motivação** – A primeira fase corresponde à contextualização e descrição do problema de investigação, já mencionados na secção 1.1 do capítulo 1, e na revisão da literatura, que será apresentada no capítulo 3. O problema identificado é a necessidade de existir um processo mais rápido e eficiente para efetuar pesquisas no espaço contido dentro de um determinado sítio *web*.
2. **Definir objetivos** – Esta fase corresponde à enumeração dos objetivos que deverão ser cumpridos para solucionar o problema identificado. Os objetivos foram definidos na secção 1.2 do capítulo 1.
3. **Desenho e desenvolvimento do artefacto** – Esta fase corresponde ao desenho do modelo conceptual proposto para o sistema de *chatbot* e posterior desenvolvimento e treino de um agente de conversação. Será apresentada nos capítulos 4 e 5.
4. **Demonstração** – Esta fase corresponde à implementação do protótipo do sistema de *chatbot* e à simulação do seu comportamento em cenários reais. Será apresentada no capítulo 5.
5. **Avaliação** – Esta fase corresponde à avaliação da performance do sistema e da sua viabilidade para solucionar o problema identificado, com base em dados recolhidos através da realização de testes com utilizadores e na análise de questionários de satisfação. Será apresentada no capítulo 6.
6. **Comunicação** – Esta fase corresponde à discussão dos resultados obtidos neste estudo, que serão divulgados através desta dissertação e será apresentada no capítulo 7.

3 Revisão bibliográfica

Neste capítulo é apresentada a revisão da literatura, com o objetivo fazer um enquadramento teórico e perceber o estado da arte dos temas que serão abordados durante a dissertação. Em primeiro lugar, serão exploradas as temáticas da gestão de conteúdos, recuperação de informação e pesquisa de conteúdos na Internet, incluindo os desafios subjacentes a estes procedimentos, seguindo-se uma análise sintética sobre inteligência artificial, aprendizagem máquina e aprendizagem profunda. Segue-se uma análise sobre as ferramentas inteligentes de apoio pessoal, em especial os *chatbots*, incluindo a forma como utilizam o processamento de linguagem natural. Em seguida, é investigado o potencial da utilização de *chatbots* como ferramentas de pesquisa, incluindo uma comparação com motores de busca tradicionais e o levantamento de *frameworks* e implementações de *chatbots* destinados particularmente à pesquisa e recuperação de conteúdos em sítios *web*. No final, é feita uma discussão onde são apresentadas as oportunidades e os desafios da utilização de *chatbots* para pesquisa de conteúdos na Internet.

3.1 Gestão de conteúdos

Barker (2016) define conteúdo como sendo a informação produzida através do processo editorial e destinada ao consumo humano através da publicação, salientando que esta definição aponta para a dicotomia base da gestão de conteúdos: a diferença entre a gestão e a distribuição.

Todos os sítios *web* são compostos por múltiplos conteúdos *web*, termo que se pode referir tanto ao tipo da informação como ao item individual usado para transmitir essa informação (Tatar et al., 2014). Os utilizadores visitam as páginas *web* com um ou mais objetivos em mente e, por isso, os conteúdos têm um papel muito importante na construção de um sítio *web* (Garrett, 2010). Existem vários tipos de conteúdos *web* divididos, essencialmente, em duas categorias: conteúdos de texto e conteúdos multimédia, que incluem imagens, ficheiros áudio, vídeos, entre outros (Baishya & Kakoty, 2019; Yao et al, 2017).

Hoje em dia, os sítios *web* cujo modelo de negócio é especialmente focado na divulgação de conteúdos utilizam sistemas de gestão de conteúdos *web* (SGCW) (Garret, 2010), um tipo de sistema de gestão de conteúdos (SGC) que permite a vários utilizadores criar, editar, publicar e gerir o conteúdo ao longo do seu ciclo de vida (Shivakumar, 2016). O conteúdo é normalmente armazenado numa base de dados e apresentado numa camada de apresentação baseada num conjunto de modelos (Rohilla, 2017). Existem centenas de SGC, proprietários ou de código-livre, sendo os mais populares o Wordpress, o Joomla e o Drupal.

Os SGC permitem a criação, gestão, distribuição e publicação de conteúdos, incluindo conteúdo corporativo. Eles asseguram que a informação certa vai para o utilizador certo, organizam e dão acesso a todos os tipos de conteúdos digitais, contêm informações sobre ficheiros e ligações para outros

ficheiros, facilitando a sua localização e acesso, e podem englobar todo o conteúdo criado dentro de uma organização (Jankulovski & Mitrevski, 2017).

Tipicamente, um SGC consiste em dois elementos: um sistema para gestão de conteúdos e um sistema de fornecimento de conteúdos (Srivastav & Nath, 2016). Por sua vez, um SGCW, como especificado na Figura 2, deverá incluir também uma componente de recolha de informação, que permite reunir e processar informações provenientes de diferentes fontes antes de serem formatadas e armazenadas no sistema de gestão como conteúdos. Posteriormente, o sistema de fornecimento extrai e envia os conteúdos para o sítio *web* (Boiko, 2005).

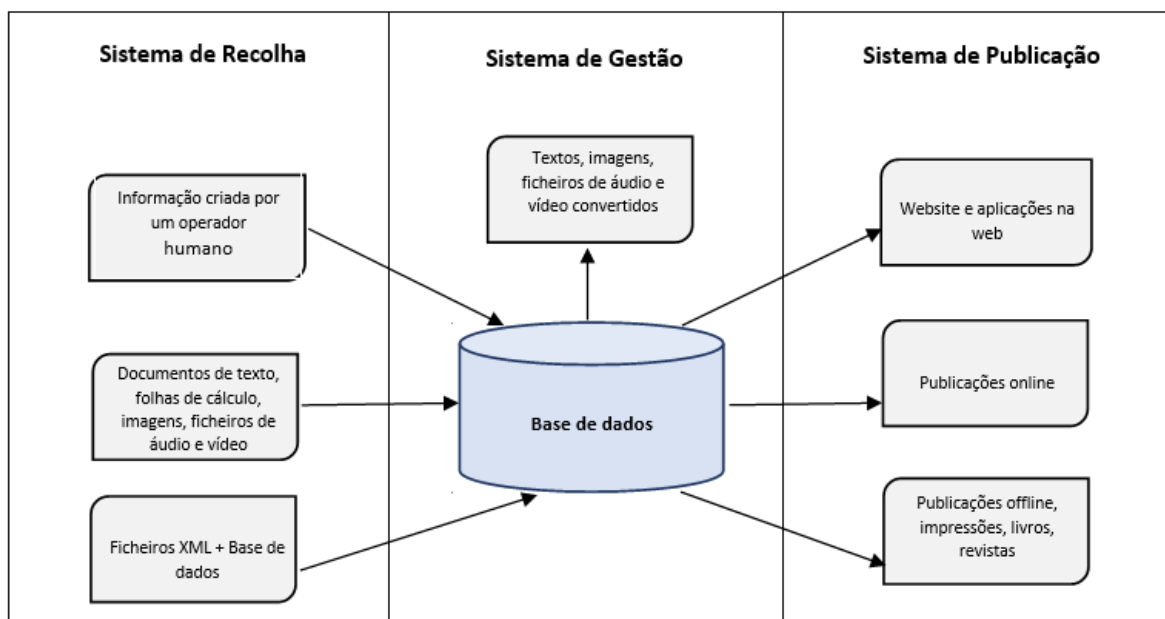


Figura 2 – Representação esquemática de um SGCW. Adaptado de Boiko (2005)

As características de um SGC variam, mas normalmente incluem soluções de publicação baseadas na *web*, gestão de formatos, controlo de revisão, indexação, pesquisa e recuperação de informação (Srivastav & Nath, 2016).

Com as suas capacidades, os SGC são vistos como uma das mais importantes tecnologias da informação e comunicação (TIC) no que diz respeito à gestão da informação e do conhecimento das organizações (Wan et al., 2016).

3.1.1 – Recuperação de Informação

A recuperação de informação é o processo de representar, armazenar, organizar e aceder a informação (Baez, Daniel & Casati, 2020) e pode ser pensada como uma forma de providenciar as ferramentas que permitam satisfazer as necessidades de informação do utilizador.

Dentro do contexto dos sistemas de recuperação de informação (SRI), deve ser feita uma distinção entre a recuperação de dados e a recuperação de informação, sendo a primeira usada para determinar os documentos que contêm as palavras-chave pesquisadas. No entanto, para tentar satisfazer as necessidades do utilizador, o SRI deve interpretar, de alguma forma, os conteúdos dos documentos, processo que inclui a extração de informação sintática e semântica, e compará-la com a pesquisa do utilizador (Baez et al., 2020).

Normalmente, um SRI efetua a sua pesquisa em coleções de dados estruturados ou semiestruturados, como por exemplo páginas *web*, documentos, imagens e vídeos, e normalmente é utilizado quando uma coleção alcança um tamanho que já não é suportado pelas técnicas tradicionais (Sanderson, 2012).

A área da RI foi uma das primeiras dentro das ciências da computação a reconhecer a importância da interação homem-máquina e que esta é necessária para garantir o acesso eficaz à informação (Croft, 2019).

3.1.2 Pesquisa de conteúdos na Internet

Os dois métodos utilizados com mais frequência para localizar informações disponíveis na Internet são a navegação e a pesquisa.

A navegação consiste no processo de seguir um caminho de hiperligações, que podem ou não estar organizadas em diretórios, até chegar à informação pretendida. Por outro lado, a pesquisa é realizada através de um sistema que tenta encontrar correspondências entre as palavras do utilizador e os documentos disponíveis na *web*.

Um motor de busca, ou de pesquisa, é um sistema de *software* projetado para encontrar informações armazenadas em sistemas computacionais a partir de palavras-chave indicadas pelo utilizador, reduzindo o tempo de pesquisa necessário para encontrar as informações. Surgiram logo após o aparecimento da Internet e permitem procurar qualquer tipo de informação na *web*, apresentando resultados de forma mais organizada, rápida e eficiente.

Um motor de busca é composto por três partes: um rastreador, que percorre a Internet em busca de páginas *web* para serem adicionadas à coleção de páginas do próprio motor de busca; um indexador, que é a principal estrutura usada para representar essas páginas, indexando cada uma das suas palavras numa enorme base de dados, e um gestor de consultas, o componente mais complexo do

motor de busca, que fornece as respostas às consultas do utilizador, comparando-as com o índice e devolvendo os documentos com melhor correspondência (Henzinger, 2004; Chau et al., 2002). Os resultados apresentados consistem em páginas da Internet onde a palavra-chave ou expressão pesquisada se encontra.

3.1.3 Desafios atuais na pesquisa de informação na web

A era digital e o crescimento exponencial da Internet nos últimos anos levaram as empresas e organizações a armazenarem cada vez mais informações na *web*, muitas vezes de forma pouco organizada. O enorme volume de dados dificulta a consulta de assuntos específicos, tornando-se evidente a necessidade de novas formas de aceder à informação. (Carlander-Reuterfelt et al., 2020; Eisman, López & Castro, 2012).

Apesar de estar cada vez mais integrada em sistemas operativos, aplicações de *software* e outros ambientes interativos que se estendem para além do navegador *web* tradicional (Boyd & Wilson, 2018), a pesquisa de informação na *web* é um processo extremamente exigente (Jenkins, Churchill, Cox & Smith, 2007). Quando efetuam uma pesquisa num sítio *web*, os utilizadores têm interesse em obter rapidamente as respostas corretas às suas consultas, de modo a não perderem o seu tempo a navegar por múltiplas páginas através dos tradicionais menus ou da pesquisa por palavras-chave em motores de busca otimizados para devolver os melhores resultados com base em estratégias de *marketing*.

Entretanto, a crescente popularidade de dispositivos inteligentes, assistentes pessoais e sistemas de gestão de relacionamento com o cliente (CRM - *Customer Relationship Management*) estimularam a comunidade científica a desenvolver novas metodologias para assistentes virtuais, entre os quais se encontram os *chatbots* (Jenkins et al., 2007).

3.2 Inteligência Artificial

A inteligência artificial é uma das áreas mais recentes na ciência e na engenharia e tem tido um papel muito importante na era da transformação digital. Embora nos últimos anos tenham sido publicados muitos artigos sobre IA, a verdade é que continua a ser difícil encontrar uma definição universal para descrever as suas capacidades (Kaplan, 2018; Wang, 2019; Bach, 2020, Sutton, 2020).

Russell & Norvig (2016) apresentam oito definições, organizadas em quatro categorias: pensamento humano, pensamento racional, comportamento humano e comportamento racional. Por sua vez, Monett et al (2020) consideram que o trabalho de Wang (2019) é uma das mais fortes tentativas para se chegar a um consenso sobre esta questão. Para Wang (2019), a essência da inteligência é o princípio de se adaptar ao ambiente enquanto se trabalha com conhecimento e recursos insuficientes. Da mesma forma, um sistema inteligente deve depender não só da capacidade de processamento finito

e de trabalho em tempo real como também ser aberto a tarefas inesperadas e aprender com a experiência. Esta definição operacional interpreta inteligência como uma forma de racionalidade relativa.

A IA incorpora várias ferramentas, técnicas e algoritmos, incluindo aprendizagem máquina, aprendizagem profunda, redes neurais e processamento de linguagem natural, tecnologias que, juntamente com o aumento de dados, conectividade constante e computação de alta *performance*, podem adicionar novos níveis de eficiência e sofisticação a tecnologias futuras (Kayid, 2020). Na Figura 3 é mostrada a forma como a aprendizagem profunda e a aprendizagem máquina se enquadram na categoria mais ampla de inteligência artificial.

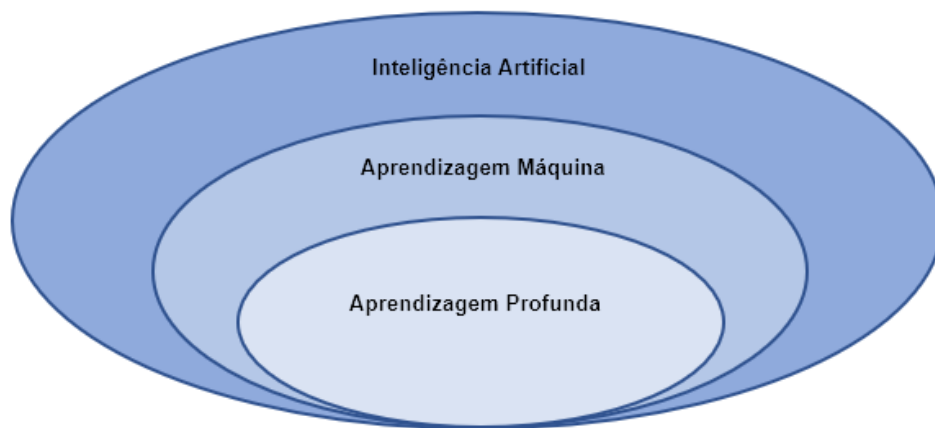


Figura 3 – Diagrama que ilustra a relação entre inteligência artificial, aprendizagem máquina e aprendizagem profunda

3.2.1 Aprendizagem Máquina (*Machine Learning*)

A aprendizagem máquina é uma área de estudo multidisciplinar que tem como principal objetivo fazer previsões através de computadores (Alzubi, Nayyar & Kumar, 2018). O conceito foi definido pela primeira vez por Samuel (1959) como sendo um campo de estudo que dá aos computadores a capacidade de aprender sem serem explicitamente programados: “um computador pode ser programado de forma a aprender a jogar um melhor jogo de damas do que a pessoa que escreveu o programa”.

Mais recentemente, Mitchell (1997) define "aprendizagem" como uma técnica que permite aos computadores aprenderem e melhorarem automaticamente a partir da experiência: “um programa aprende com a experiência E em relação a uma determinada tarefa T e uma medida de desempenho P, se o seu desempenho na tarefa T, medido por P, melhora com a experiência E” (Mitchell, 1997, p. 2).

Há duas grandes categorias que variam de acordo com o nível de intervenção humana na preparação e rotulação dos dados: aprendizagem supervisionada e aprendizagem não-supervisionada.

Há ainda técnicas de reforço, que aprendem e identificam padrões com o objetivo de reagir a um ambiente (Alloghani, Al-Jumeily, Mustafina, Hussain & Aljaaf, 2020). A Figura 4 mostra a relação entre estas técnicas de aprendizagem.

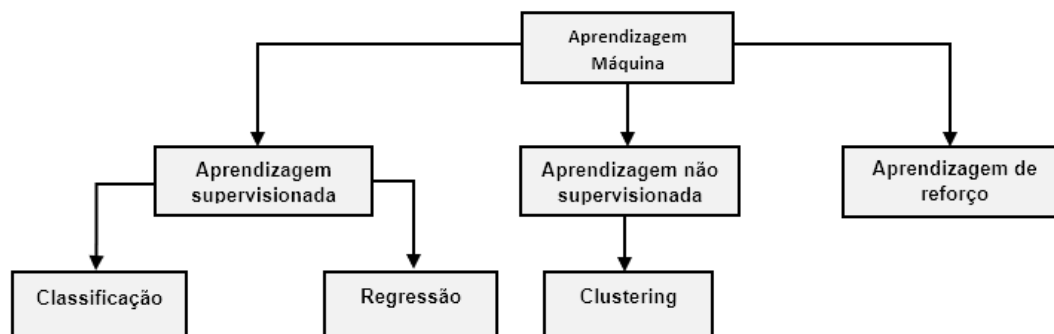


Figura 4 – Tipos de aprendizagem máquina

A escolha do algoritmo a ser usado depende do tipo de dados utilizados para treinar os modelos (Roh, Heo, & Whang, 2019) e de um conjunto de fatores, nomeadamente o tipo de algoritmo, a parametrização, o tamanho da memória, a tendência de sobreajuste, o tempo de aprendizagem e o tempo de previsão (Jatana, 2019).

A aprendizagem supervisionada consiste no mapeamento das variáveis de entrada para determinar as suas características e, desta forma, conseguir prever os resultados de saída. O algoritmo é alimentado com um conjunto de variáveis de entrada e categorias correspondentes a essas variáveis (conjunto de treino) com o objetivo de aprender a produzir os resultados corretos de novas variáveis, que serão introduzidas numa fase posterior (Sedkaoui, 2018). Esta versão é muito utilizada em problemas de classificação, onde o objetivo é encontrar uma classe dentro de um conjunto limitado de possibilidades, obtendo como resultado um valor discreto (Osisanwo et al, 2017) e de regressão, onde é necessário prever um valor numérico específico, tendo como resultado um valor contínuo (Salman & Kecman, 2012). É um método que pode ser utilizado em casos de uso como classificação de imagens, tecnologias de reconhecimento de caracteres OCR (*Optical Character Recognition*), reconhecimento facial (Younis & Alkhateeb, 2017), análise de sentimento (Samal Behera & Panda, 2017) e PLN (Khan, Daud, Nasir & Amjad (2016).

Por sua vez, a aprendizagem não-supervisionada consiste no reconhecimento de padrões sem que se conheça o atributo de destino (Alloghani et al, 2020). A técnica de *clustering* é um exemplo de aprendizagem não-supervisionada e consiste na organização de dados que não estão categorizados em grupos semelhantes, os *clusters* (Serra & Tagliaferri, 2019).

A aprendizagem de reforço é uma categoria da aprendizagem não-supervisionada que consiste em aprender através da interação com o ambiente por meio de diferentes ações, falhas e/ou sucessos, enquanto tenta maximizar as recompensas recebidas. Não é explicado ao agente qual é a ação a realizar. É uma forma similar ao processo de aprendizagem natural, onde a aprendizagem se

desenvolve através de tentativa e erro, relacionada com a forma como os agentes de aprendizagem máquina devem agir num determinado ambiente, de modo a maximizar um resultado específico. Retira variáveis do ambiente e faz decisões contínuas de modo a resolver o problema (Hammoudeh, 2018).

3.2.2 – Aprendizagem Profunda (*Deep Learning*)

A aprendizagem profunda é um ramo da aprendizagem máquina que se baseia no princípio básico de que existem várias camadas de processamento com estruturas complexas e onde cada etapa recebe como entrada o resultado da etapa anterior, criando assim uma ideia de profundidade (Goodfellow, Bengio & Courville, 2017). A aprendizagem utilizando este método tem um nível de intervenção humana muito baixo, já que não há necessidade de fazer grandes investimentos na preparação ou otimização dos dados. À medida que se avança para as camadas mais profundas, a extração de dependências, padrões e regularidades que permitem a abstração e a aprendizagem de descrições gerais é feita automaticamente durante o treino (Alpaydin, 2020).

As técnicas de aprendizagem profunda têm permitido alcançar importantes avanços na resolução de problemas aos quais a IA ainda não tinha conseguido dar resposta, nomeadamente ao nível da compreensão de linguagem natural, em tarefas como classificação de tópicos, análise de sentimentos, resposta a perguntas e tradução. Estas técnicas podem ser aplicadas em muitos domínios da ciência, bem como a nível empresarial e governamental (LeCun et al., 2015).

Apesar do seu enorme potencial, a aprendizagem profunda tem ainda alguns problemas, como o facto de necessitar de uma maior quantidade de memória e computação para fazer face a redes neurais muito grandes que, por sua vez, contêm uma maior quantidade de parâmetros livres, tornando o seu treino mais difícil (Alpaydin, 2020).

3.3 Ferramentas inteligentes de apoio pessoal

As ferramentas inteligentes de apoio pessoal são uma das principais aplicações da inteligência artificial. São agentes de *software* capazes de realizar tarefas ou serviços com base em comandos ou questões, que podem ser fornecidas por texto ou por voz.

Ao utilizar técnicas de IA como o PLN, reconhecimento de voz, tradução de máquina ou representação de conhecimento (Russell & Norvig, 2016), os assistentes digitais têm normalmente um elevado nível de interatividade e inteligência, que permite aumentar a sua performance na realização de tarefas (Maedche, Morana, Schacht, Werth & Krumeich, 2016).

Geralmente, a literatura utiliza vários termos como sinónimos para "*chatbots*". Estes termos incluem "assistentes virtuais", "assistentes digitais", "agentes de conversação", "*chatterbots*" ou "sistemas de

diálogo natural", entre outros, embora sejam todos usados de modo indistinto para descrever sistemas de conversas com recurso à linguagem natural.

3.3.1 Chatbots

Os *chatbots* são sistemas desenhados para simular conversas com humanos, especialmente através da Internet, capazes de realizar uma grande variedade de tarefas com níveis variados de complexidade. São um exemplo típico de sistemas de inteligência artificial e prometem revolucionar a área da interação homem-máquina.

Também conhecidos formalmente como agentes conversacionais, os *chatbots* são programas de computador que respondem como uma entidade inteligente quando se conversa por voz ou texto, sendo capazes de perceber uma ou mais línguas humanas através do PLN.

Os sistemas de *chatbot* tornaram-se populares nos últimos anos não só como forma de apoiar a comunicação externa de uma empresa com os clientes, mas também, e cada vez mais, para fins internos, especialmente para melhorar, automatizar e acelerar os fluxos de trabalho (Ding, Ranade & Cata, 2019; Johannsen, Schaller & Klus, 2020). Alguns trabalhos de investigação afirmam que a tecnologia de *chatbot* é tão disruptiva que eliminará a necessidade de sítios *web* e aplicações (Carlander-Reuterfelt et al., 2020).

Os agentes de conversação evoluíram de programas simples, baseados em reconhecimento de padrões, para sistemas bastante complexos que recorrem a tecnologias de IA. Usam uma combinação de PLN, compreensão de linguagem natural (CLN), geração de linguagem natural (GLN) e técnicas de aprendizagem automática, o que lhes permite ser mais flexíveis na manutenção de uma conversa (Eisman et al., 2012) e simplificam substancialmente a interação homem-máquina, graças ao seu paradigma de linguagem natural (Daniel, Matera, Zaccaria, & Dell'Orto, 2018). Estas características contribuem para o aumento da satisfação do utilizador, já que permitem ajudar a encontrar informações de uma forma mais confortável do que outras interfaces menos sofisticadas e mais demoradas (Carlander-Reuterfelt et al., 2020). Segundo Brandtzaeg e Følstad (2017), as investigações atuais estão mais focadas nas interfaces do utilizador (IU) de linguagem natural, onde a interação é feita através de cadeias de texto em línguas naturais em vez de métodos mais habituais como o clique, o *scrolling* ou o *swiping*.

Existem múltiplas formas de categorizar os *chatbots*, tais como o método que usam para classificar o texto, o seu papel para uma determinada empresa, a sua complexidade de interação ou ainda, de uma forma holística, pelo seu domínio e requisitos (Casas et al, 2020).

Nimavat e Champaneria (2017) identificam vários parâmetros para classificar os *chatbots*, como o domínio do conhecimento, serviço prestado, objetivos e método de geração de resposta. Os *chatbots*

não têm necessariamente de pertencer exclusivamente a uma ou outra categoria e existem em proporções variáveis.

Dentro do domínio do conhecimento, os *chatbots* podem pertencer a um domínio aberto ou fechado. Quanto aos serviços prestados, estes podem ser de carácter interpessoal, intrapessoal ou interagente. Os objetivos podem ser informativos, conversacionais ou baseado em tarefas. A geração de respostas, que processa o texto introduzido pelo utilizador, pode ser utilizada de formas diferentes. Os sistemas mais inteligentes usam aprendizagem máquina e conseguem compreender línguas naturais. Esta técnica é usada quando o domínio é limitado e, ao mesmo tempo, há uma grande quantidade de dados disponível para treinar o sistema. Já os sistemas baseados em regras usam combinações de padrões e são usados quando o número de resultados e cenários possíveis é fixo. Há ainda sistemas híbridos, que juntam as capacidades dos outros dois tipos (Nimavat & Champaneria, 2017).

Muitas plataformas comerciais têm *dashboards* com as ferramentas necessárias para definir, testar, modificar e implementar *chatbots* com recurso a poucos ou mesmo nenhuns conhecimentos de programação (Srivastava & Prabhakar, 2020).

O número de tecnologias relacionadas com o *chatbot* já é bastante significativo e cresce a cada dia. Os *chatbots* podem ser desenvolvidos de duas formas: usando qualquer linguagem de programação ou usando plataformas de última geração. Neste momento, há seis plataformas líderes que os desenvolvedores podem usar para criar aplicações capazes de entender línguas naturais: o Dialogflow da Google, o wit.ai do Facebook, o Microsoft LUIS, o IBM Watson Conversation, o Amazon Lex, e o SAP Conversation AI. Todas estas plataformas são suportadas por aprendizagem máquina e partilham algumas funcionalidades padrão (são baseadas em nuvem, suportam várias linguagens naturais e de programação), apesar de diferirem significativamente noutros aspetos. Outras plataformas conhecidas de desenvolvimento de *chatbots* são RASA, Botsify, Chafuel, Manychat, Flow XO, Chatterbot, Pandorobot, Botkit e Botlytics (Casas et al, 2020).

3.3.2 Processamento de Linguagem Natural

O processamento de linguagem natural é uma área de estudo transdisciplinar que reúne ciências da computação, linguística, estatística e inteligência artificial e tem como objetivo processar e analisar, de forma automática, as línguas naturais (Khurana, Koli, Khatter & Singh, 2017).

A compreensão e utilização das línguas humanas é utilizada para desenvolver técnicas que permitam aos computadores compreender e manipular expressões naturais para realizar determinadas tarefas. A maior parte das técnicas têm como base a aprendizagem máquina. A compreensão de linguagem natural é uma parte fundamental das tarefas do PLN e permite, por exemplo, implementar interfaces do utilizador como os *chatbots*.

Na CLN realizam-se dois tipos de análise: a análise sintática, que estuda a função de cada palavra numa oração e verifica se a expressão obedece às regras de gramática formais, e a análise semântica,

que está associada ao estudo do significado e contexto das palavras e orações (Aleedy, Shaiba & Bezbradica, 2019).

O objetivo da CLN consiste em identificar as intenções do utilizador, extrair entidades e responder de forma apropriada, de acordo com a intenção do utilizador. Desta forma, as intenções categorizam o mapeamento entre a intenção do utilizador e a forma como o *chatbot* deve dar continuidade à conversa, enquanto as entidades representam valores de parâmetros extraídos a partir da linguagem natural das orações introduzidas pelo utilizador. (Adamopoulou & Moussiades, 2020). Através das intenções, o agente é capaz de compreender a motivação por trás de uma oração introduzida pelo utilizador. Por sua vez, as entidades são usadas para selecionar pedaços específicos de informação mencionados pelo utilizador. Uma intenção pode ter múltiplas entidades.

3.4 Utilização de *chatbots* para pesquisar informação na Internet

Os *chatbots* estão a tornar-se cada vez mais abrangentes em múltiplos domínios como o *marketing*, sistemas de apoio, educação, cuidados de saúde, cultura, património e entretenimento (Augello, Gentile & Dignum, 2018; Casas, Tricot, Khaled, Mugellini & Cudré-Mauroux, 2020) e têm também várias aplicações nas áreas da recuperação de informação, negócios e *e-commerce* (Shawar & Atwell, 2007a). Há ainda soluções abertas, como as disponibilizadas pelos assistentes da Amazon ou da Google, que visam responder a qualquer tipo de pergunta em vez de uma área específica de conhecimento (Carlander-Reuterfelt et al., 2020), funcionando de forma semelhante aos sistemas de pesquisa tradicionais.

3.4.1 Diferenças entre motores de busca e *chatbots* para a pesquisa de informação

Nos últimos anos, tem-se assistido a um grande crescimento na investigação e na utilização de *chatbots*, com o objetivo de reduzir a carga cognitiva necessária para obter informações básicas a partir de um sítio *web* (Srivastava & Prabhakar, 2020; Yadav, Garg, & Mathur, 2019). A recuperação de informação através de *chatbots* é especialmente complexa, devido à interação com a linguagem natural, e especialmente quando o domínio é restrito (López De Luise, Pascal, Alvarez, Tournoud, Pankrac & Cruz, 2021). Para além disso, nos cenários de pesquisa de informação, estes sistemas ainda têm capacidades de conversação muito limitadas (Galitsky, 2019c), e servem principalmente como *proxies* de motores de pesquisa *web* existentes (Vtyurina, Savenkov, Agichtein & Clarke, 2017).

Os sítios *web* de pesquisa convencionais (ou motores de busca) utilizam um método unilateral que lista os resultados para uma palavra de pesquisa de entrada. Por outro lado, os *chatbots* atuais visam encontrar e fornecer os serviços e dados de que as pessoas precisam no momento certo sob a forma de conversação, sem executar aplicações ou sítios *web* separados (Jeong & Seo, 2019).

Os motores de pesquisa *web* atuais não são sistemas de conversação porque não respondem com linguagem natural, mesmo que o utilizador introduza consultas usando língua natural (Sa & Yuan, 2020). Ao contrário dos motores de busca e sistemas de recuperação de informação e de recomendação, os sistemas de conversação concebidos para pesquisa são caracterizados por uma sequência de interações entre o utilizador e o sistema, de modo a satisfazer as necessidades de informação do utilizador, realizando tarefas como procurar ficheiros, agendar lembretes, navegar na Internet ou mesmo conversar com o assistente (Galitsky, 2021a).

Outra diferença importante é que na ferramenta de pesquisa há uma ligação entre qualquer palavra e o documento onde ela se encontra, enquanto no sistema de conversação, normalmente, é fornecida uma ligação apenas para as palavras mais significativas (Shawar & Atwell, 2007b).

3.4.2 Desafios no desenvolvimento de *chatbots* para pesquisa de informação na *web*

Embora se tenha tornado muito mais fácil desenvolver um sistema de *chatbot*, o próprio sistema é complexo por natureza (Maroengsit et al., 2019). O trabalho de compor o *chatbot* a partir de cenários do mundo real não é uma atividade trivial e requer uma compreensão significativa do problema, bem como do domínio (Srivastava & Prabhakar, 2020).

Outro grande desafio consiste em tentar compreender de que forma o paradigma de conversação pode integrar ou mesmo substituir as interfaces visuais. A apresentação de grandes quantidades de dados numa interface de conversação é um desafio significativo, e proporcionar uma visão integrada sobre diferentes fontes de dados relacionadas é ainda mais complexo. Este não é apenas um problema de apresentação de dados, mas também de técnicas adequadas de análise de dados, tais como técnicas de resumo, que visam extrair ou gerar um resumo de um documento que responda diretamente ou seja relevante para a consulta de pesquisa (Daniel et al., 2018; Behere et al., 2020; Zhao et al., 2021).

Embora tenha havido um aumento acentuado do número de publicações relacionadas com *chatbots*, incluindo um aumento de publicações usando metodologias de avaliação, até à data não surgiu nenhuma norma bem reconhecida para a avaliação do *chatbot*, sendo que, durante algum tempo, eram apenas avaliados através da sua capacidade de passar no Teste de Turing (Casas et al, 2020).

Hernández-Orallo (2016) propõe um método de avaliação dos artefactos de IA onde um sistema é caracterizado pelas suas capacidades cognitivas e não pelos resultados da execução das tarefas que foi projetado para resolver. Por sua vez, Casas et al. (2020) examinam diversas metodologias de avaliação de *chatbot* e avaliam-nas de acordo com os conceitos de usabilidade ISO 9214: eficácia, eficiência e satisfação. Ding et al. (2019) fazem uma proposta heurística baseada na usabilidade, design de informação e padrões de acessibilidade para avaliar e projetar *chatbots* de suporte. As

avaliações são concretamente descritas como objetivos e são escritas em cenários bons, médios e pobres que fornecem uma ferramenta de medição simples, mas efetiva.

3.4.3 Frameworks e implementações de *chatbots* para pesquisa de conteúdos em sítios *web*

Vários dos estudos consultados no âmbito da revisão da literatura propõem *frameworks* ou implementam com sucesso sistemas ou módulos de apoio ao funcionamento de *chatbots*.

Shawar & Atwell (2007b) desenvolvem com sucesso uma ferramenta aplicada a um sítio *web* que avalia a capacidade de usar o *chatbot* como um sistema de recuperação de informação, comparando-o com um motor de busca. Os utilizadores que testaram o sistema preferiram a utilização do *chatbot* pela sua capacidade de dar respostas diretas em vez de uma lista de links, reduzindo assim o tempo de navegação e pesquisa. Eisman et al. (2012) apresentam um *framework* para projetar assistentes virtuais de domínio fechado que utilizam linguagem natural para ajudar a encontrar informações úteis sobre

um determinado tópico num site. Carlander-Reuterfelt et al. (2020) propõem o sistema Jaicob, projeto que conseguiu resultados muito favoráveis na avaliação da usabilidade e originalidade.

Kondylakis et al. (2020) apresentam o R2D2, um *chatbot* inteligente que se baseia em tecnologias semânticas da *web* e oferece uma interface de linguagem natural controlada e inteligente para aceder à informação disponível num sítio *web*. Goh, Fung & Wong (2007) propõem o AINI, um agente automático de extração de conhecimento para aquisição de conhecimentos a partir da Internet. Por sua vez, Zhao et al. (2021) descrevem um sistema integrado baseado em domínio aberto e conhecimentos específicos de domínio, com o objetivo de fornecer interações *web* inteligentes, baseadas em consultas. Já Anghelescu & Nicolaescu (2018) implementam um sistema de *chatbot* que utiliza motores de busca e técnicas de ensino.

Singh, Ramasubramanian & Shivam (2019) apresentam um *chatbot* personalizado, capaz de recuperar informações disponíveis em sítios *web*, chamado IRIS (Serviço de Reconhecimento de Intenções e Informação). Daswani, Desai, Patel, Vani & Eirinaki (2020) apresentam o CollegeBot, um agente de IA conversador que utiliza o PLN e aprendizagem automática para ajudar os visitantes do site de uma universidade a localizar facilmente informações relacionadas com as suas consultas.

Hariguna & Efendi (2020) projetam e constroem uma aplicação de *chatbot* como um serviço de informação baseado em aprendizagem automática, implementando o algoritmo de semelhança jaccard e a funcionalidade de algoritmo N-Gram, Bi gram, para obter respostas mais precisas e aumentar a eficácia do tempo da pesquisa dos utilizadores na procura de informação. Papenmeier et al. (2020) projetaram um sistema de pesquisa interativa a partir de uma visão centrada no utilizador que é capaz de lidar com a imprecisão e ambiguidade nas consultas dos utilizadores.

Ramesh, Nagaraju, Harish & Kumaraswamy (2021) constroem um *chatbot* que recupera conteúdos sob a forma de diálogos virtuais produzidos automaticamente a partir de documentos. Agrawal et al. (2020) apresentam o QnAMaker, que pretende simplificar o processo de criação de *bots* através da extração de pares de perguntas-respostas. Por sua vez, Galitsky (2019b) implementa um mecanismo de relevância baseado na semelhança das árvores de análise para uma série de componentes de *chatbot*, incluindo a pesquisa. Em Galitsky (2021b) é construído um *chatbot* que encontra documentos, extrai tópicos dos mesmos e organiza estes tópicos em *clusters*. Baez et.al. (2020) propõe uma arquitetura de referência para a inferência automatizada de *chatbots* específicos capazes de mediar entre a interação do utilizador com o site. Da Silva, Gao & Andreae (2014) criaram o Wallace para recuperar páginas da Wikipédia, analisá-las e selecionar frases adequadas para a conversas. Já Prajwal, Mamatha, Ravi, Manoj & Joisa (2019) propõem um sistema que visa a construção de um Assistente Virtual inteligente que tenha a capacidade de colocar prioridade no contexto da pesquisa.

3.5 Discussão

Com base na literatura, conclui-se que a utilização de *chatbots* para o apoio inteligente à pesquisa de conteúdos em sítios *web* tem sido estudada e implementada com sucesso por alguns investigadores. Existe, no entanto, um conjunto de oportunidades e desafios a ter em conta no desenvolvimento destas ferramentas.

3.5.1 Oportunidades

- Os *chatbots* podem ajudar os utilizadores a encontrar informações úteis para as suas necessidades de pesquisa;
- Tem existido um aumento no interesse da utilização dos *chatbots* dotados de inteligência artificial para realizar ações de pesquisa de conteúdos e recuperação de informação na *web*;
- Os resultados do estudo de Jain, Kumar, Kota & Patel (2018) revelam que os utilizadores preferem *chatbots* que proporcionem uma capacidade de conversação de linguagem natural "semelhante ao humano";
- O fator motivacional para a utilização de *chatbots* mais frequentemente relatado é a "produtividade" (Brandtzaeg e Følstad (2017);
- As recomendações personalizadas estão a tornar-se cada vez mais populares para permitir que as pessoas obtenham produtos e serviços (Galitsky, 2019a).

3.5.2 Desafios

- Os *chatbots* não podem ser aplicados a grandes domínios de informação (Da Silva et al., 2014);
- Os utilizadores gostam de usar *chatbots* para obter respostas rápidas e precisas. No entanto, consideram-nos limitados e ficam frustrados quando o sistema não os consegue compreender de imediato (Beriault-Poirier, Tep & Sénécal, 2019).
- A falta de eficácia na utilização dos *chatbots* pode estar relacionada com a falta de padrões para avaliar os *chatbots* (Casas et al, 2020).
- Seria interessante comparar o uso dos motores de busca com o do *chatbot*. Seria igualmente interessante comparar a facilidade de utilização do *chatbot* com um motor de busca convencional (Jenkins et al., 2007).
- É necessário repensar as IU associadas aos *chatbots* (Srivastava & Prabhakar, 2020).

4 Proposta de modelo conceptual

Neste capítulo será aplicada a terceira fase da metodologia DSR, com explicação detalhada da proposta para resolver o problema identificado na secção 1.1 do capítulo 1.

Em primeiro lugar, será feita a análise do público-alvo e análise de tarefas, com base nos resultados de um questionário realizado. Serão também apresentados os requisitos do sistema, casos de uso referentes às funcionalidades do *chatbot*, um diagrama de atividades que permite, de uma forma mais visual, entender o fluxo de interações do utilizador com o agente do *chatbot*, e ainda um digrama de arquitetura, que representa a estrutura do sistema.

4.1 Análise de utilizadores e tarefas

Com base na revisão da literatura, verifica-se que a utilização de *chatbots* para pesquisar conteúdos na Internet é vista como uma ideia interessante e com potencial. Para melhor se compreender o público-alvo, foi realizado um questionário a potenciais utilizadores do sistema de *chatbot* proposto (Anexo I), com o objetivo de avaliar se as funcionalidades pensadas inicialmente seriam realmente adequadas, aferir qual o nível de conhecimentos que tinham sobre *chatbots* e apurar o nível de interesse e os benefícios que um sistema de *chatbot* para apoio à pesquisa de informação na Internet poderia oferecer, face a outras formas de pesquisa.

Esta análise permitiu, também, identificar as tarefas que os utilizadores precisam de desempenhar e a forma como o fazem atualmente, utilizando os sistemas existentes.

O questionário, elaborado com recurso à ferramenta Google Forms, consiste em 25 perguntas de escolha múltipla e resposta obrigatória e foi desenhado de modo a poder ser respondido por qualquer tipo de utilizador, independentemente da sua idade ou formação, apenas com o requisito de ser falante de língua portuguesa. Foi divulgado em diversos canais *online*, nomeadamente nas redes sociais e em grupos de estudantes e entusiastas das áreas de inteligência artificial e ciências de dados. No total, foram contabilizadas 29 respostas válidas, que serão objeto de análise detalhada.

4.1.1 Caracterização da amostra

Do total de inquiridos, 62,1% identificam-se como sendo do género feminino e 37,9% do género masculino, mais de metade têm idades compreendidas entre os 25 e os 64 anos e a grande maioria tem habilitações literárias ao nível do ensino superior (Figura 5).

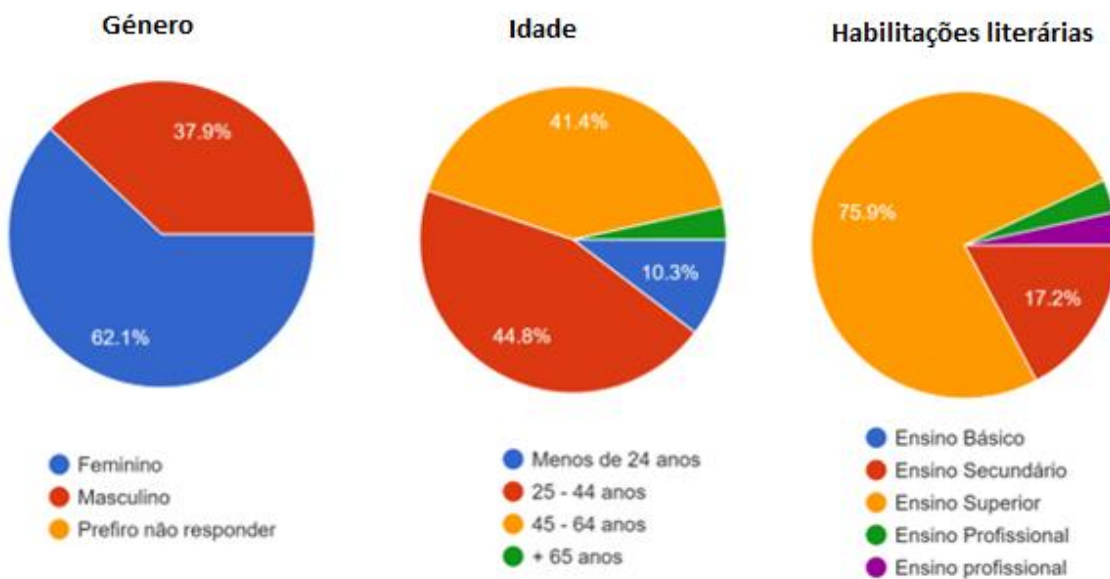


Figura 5 - Género, idade e habilitações literárias dos inquiridos

Mais de 65% dos indivíduos indicaram que os seus conhecimentos relativamente à utilização de TIC são bons ou muito bons. Os dispositivos mais usados para aceder à Internet são o computador, o *smartphone* e o *tablet* (Figura 6).

Qual (ou quais) do(s) seguinte(s) dispositivo(s) utiliza para aceder à Internet?

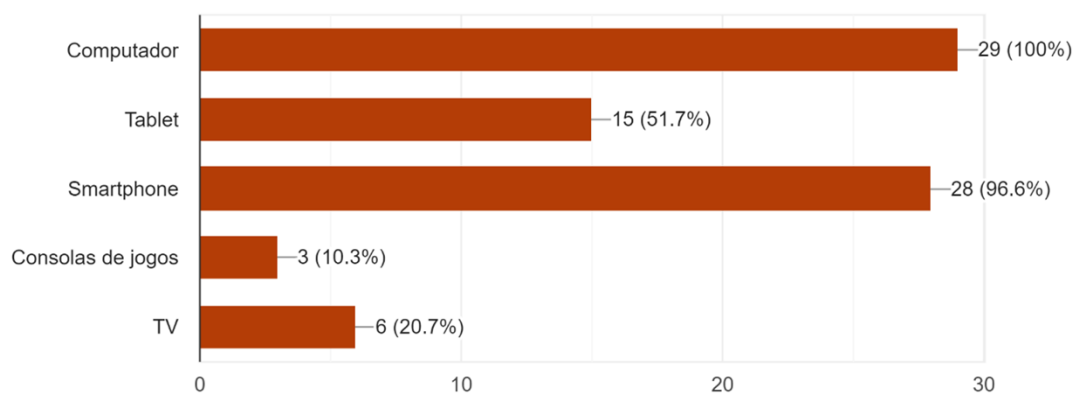


Figura 6 - Dispositivos utilizados para aceder à Internet

4.1.2 Hábitos de pesquisa

Sobre os hábitos de pesquisa na Internet, a grande maioria (89,7%) dos inquiridos indicou fazer pesquisas várias vezes por dia e apenas 6,9% afirmaram encontrar os resultados que procuram à primeira tentativa, com 89,6% dos utilizadores a indicar que só às vezes, raramente ou nunca o consegue fazer (Figura 7).

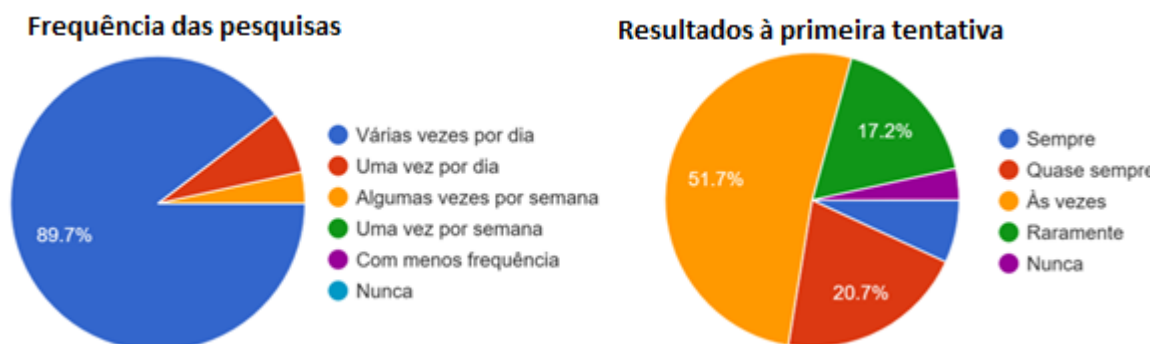


Figura 7 - Frequência e resultados da pesquisa à primeira tentativa

Para além de pesquisar conteúdos, os inquiridos utilizam a Internet sobretudo para trabalhar ou estudar, ler notícias, frequentar redes sociais e utilizar ferramentas de mensagens instantâneas (Figura 8).

Como costuma encontrar conteúdos na Internet? Que ferramentas utiliza?

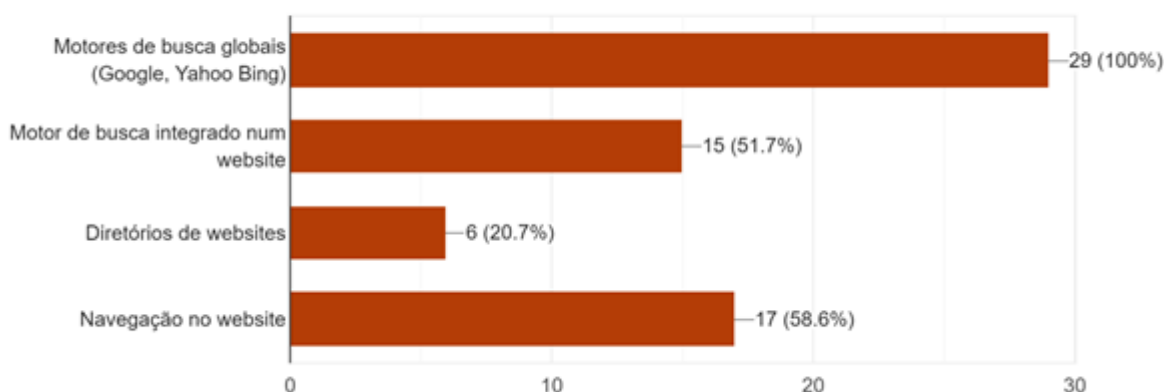


Figura 8 - Ferramentas utilizadas para pesquisar conteúdos na Internet

Todos os indivíduos que efetuam pesquisas costumam utilizar motores de busca globais, como o Google, o Yahoo ou o Bing, para fazer pesquisas na Internet. Mais de metade indica que também utiliza motores de busca integrados em sítios *web* ou que efetua a pesquisa através da navegação.

Os aspetos que mais valorizam numa ferramenta de pesquisa são o acesso rápido, a clareza da informação apresentada, os recursos de navegação (menus, ícones e botões) e a apresentação gráfica (Figura 9). Apenas 19,2% consideram importante a existência de FAQ (*Frequently Asked Questions*) ou outras ajudas para utilizar uma ferramenta de pesquisa.



Figura 9 - Aspetos valorizados numa ferramenta de pesquisa

Quando questionados sobre os aspetos que consideram mais importantes quando efetuam pesquisas na Internet, as três principais razões indicadas foram a rapidez em encontrar a resposta, a obtenção de resultados precisos e de fácil compreensão e a facilidade em encontrar o que procuram, ou seja, utilizando o menor número de passos possível. Mais de 44,8% dos indivíduos consideram que é importante poder filtrar a pesquisa e 31% consideram que é importante obter apenas resultados que correspondam ao contexto da pesquisa.

4.1.3 Utilização de *chatbots*

Relativamente aos *chatbots*, 72,4% dos inquiridos afirmam que já utilizaram este tipo de sistema, sobretudo em contextos de apoio ao cliente e vendas, embora não o façam com muita frequência, conforme se pode ver nos gráficos da Figura 10.

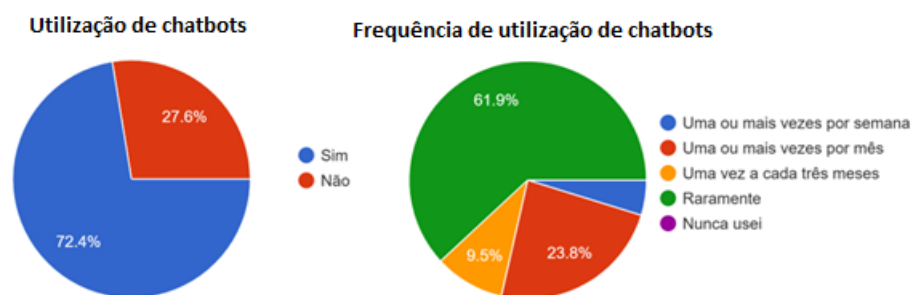


Figura 10 - Utilização e frequência de utilização de *chatbots*

Os contextos mais vezes referidos, conforme o gráfico representado na Figura 11, são o apoio ao cliente (76,2%) e vendas (52,4%).

Em que contextos já utilizou chatbots?

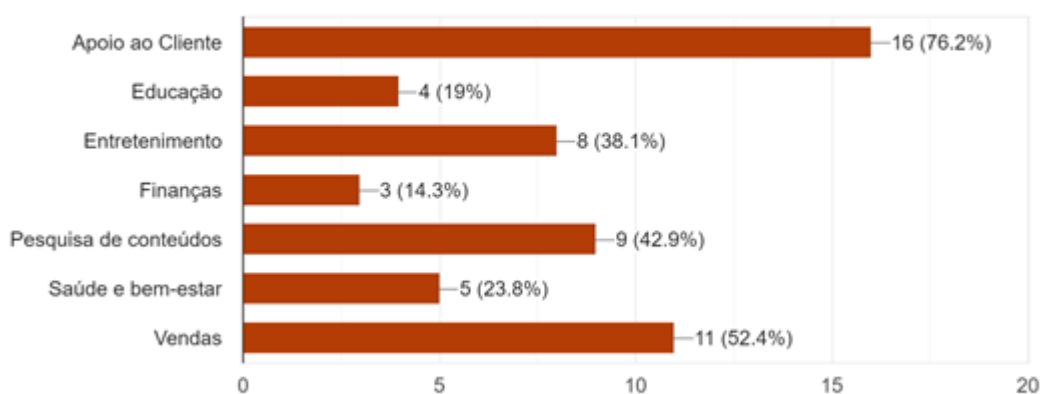


Figura 11 - Contextos de utilização de *chatbots*

Quando questionados sobre as razões pelas quais optaram por utilizar um *chatbot* em vez de outras opções disponíveis, 71,4% dos inquiridos indicaram que queriam experimentar o *chatbot*. Outras razões indicadas foram a facilidade em usar, o facto de ser uma forma mais rápida de encontrar a resposta pretendida, o rápido acesso ao *chatbot* e um histórico de experiência positiva com este tipo de sistema.

Sobre as características que gostariam de ver incluídas num *chatbot* de apoio à pesquisa de conteúdos na Internet, as opções mais requeridas foram a obtenção de resumos dos conteúdos sem ter necessidade de visitar o sítio *web* original, conversas personalizadas, capacidade de partilhar facilmente um conteúdo com outra pessoa e análise de sentimento e inteligência emocional.

A grande maioria dos utilizadores estaria disposta a partilhar dados pessoais com um *chatbot*. Tal como representado no gráfico da Figura 12, 82,8% dos inquiridos partilhariam o nome, 65,5% a idade e 62,1% o género, o *email* e a profissão.

Que dados pessoais estaria disposto a partilhar com um chatbot?

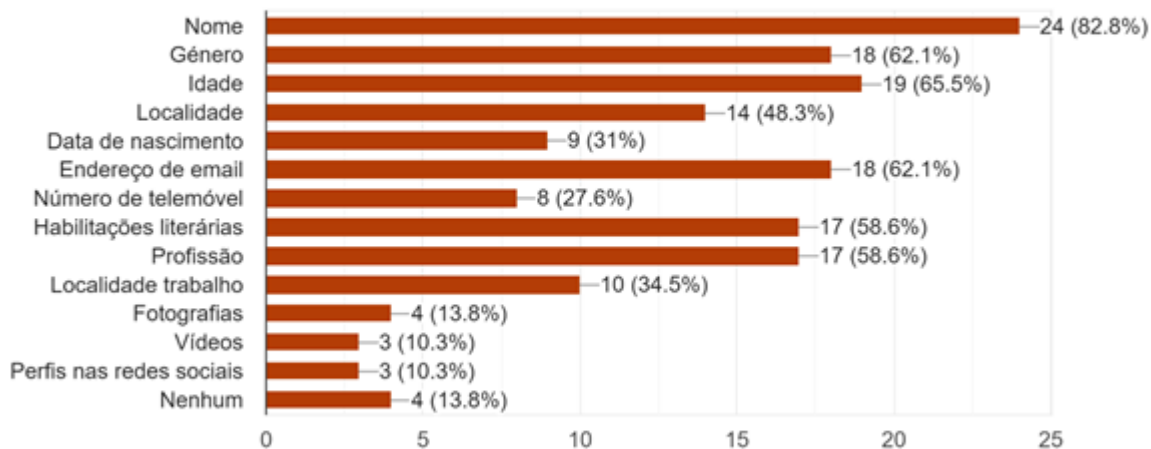


Figura 12 - Dados dispostos a partilhar com um *chatbot*

Quase todos os inquiridos (95,2%) recomendariam a utilização de *chatbots* para pesquisar conteúdos na Internet em vez dos motores de busca ou barras de pesquisa tradicionais e 86,2% dos indivíduos valorizam o facto de, ao contrário de um motor de busca tradicional, o *chatbot* lhes devolver apenas o resultado mais relevante, com base no contexto da conversa.

Quando questionados sobre a frequência com que utilizariam um *chatbot* para pesquisar conteúdos num determinado sítio *web*, a maioria (41,3%) indicou que utilizaria quase sempre ou sempre. 44,5% indicaram que utilizariam este sistema ocasionalmente e 24,1% responderam que o fariam raramente (Figura 13). Nenhum dos questionados respondeu que nunca utilizaria um *chatbot* para este fim.

Com que frequência utilizaria um chatbot para pesquisar conteúdos num determinado website?

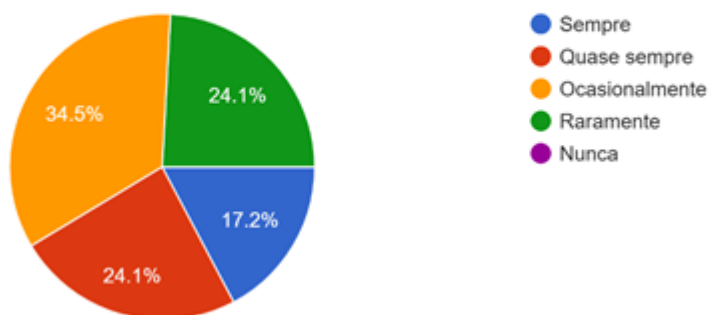


Figura 13 - Frequência de utilização de chatbots para pesquisa de conteúdos

Para 72,4% dos inquiridos, a utilização deste sistema seria feita em conjunto com outras formas de pesquisa, conforme indicado no gráfico da Figura 14.

Utilizaria o chatbot por si só ou seria mais uma ferramenta de apoio a outras formas de pesquisa?

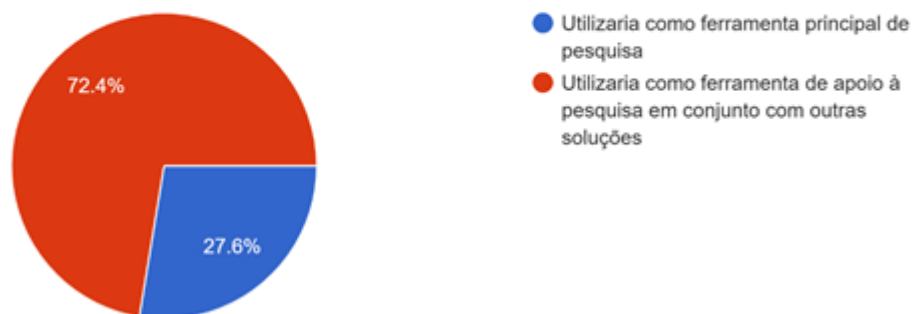


Figura 14 - Forma de utilização de *chatbots* para pesquisa de conteúdos

Por último, 72,4% dos inquiridos consideram que, no futuro próximo, os *chatbots* podem substituir os motores de busca e tornar-se na principal ferramenta de pesquisa de conteúdos na Internet (Figura 15).

Considera que, no futuro próximo, os chatbots podem substituir os motores de busca e tornar-se na principal ferramenta de pesquisa de conteúdos na Internet?

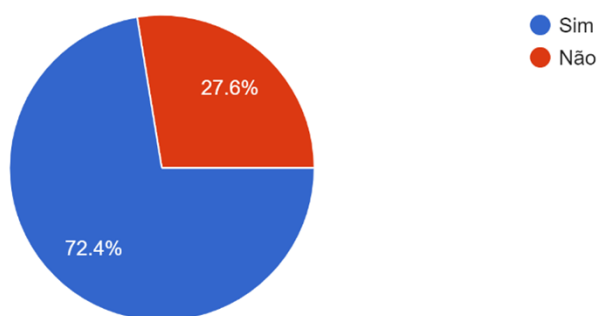


Figura 15 - Futuro dos *chatbots* como principal ferramenta de pesquisa de conteúdos na Internet

4.1.4 Discussão

Após a aplicação do questionário, é possível constatar que, apesar da maior parte dos utilizadores indicar ter conhecimentos médios a bons relativamente à utilização das TIC, há ainda uma percentagem

considerável de utilizadores que nunca utilizaram sistemas de *chatbot* e os que indicaram já ter usado não o fazem com frequência. Isto mostra que, por um lado, estes sistemas ainda não estão amplamente divulgados junto de utilizadores menos adeptos das TIC e, por outro, que ainda não são vistos como ferramentas úteis em alternativa aos sistemas já existentes. Ainda assim, é demonstrado interesse na utilização destas ferramentas.

Entre os inquiridos que indicaram já ter usado *chatbots*, os contextos mais referidos foram no âmbito do apoio ao cliente, vendas e pesquisa de conteúdos. Isto vai ao encontro das atividades que realizam na Internet, já que 41,4% afirmam realizar compras *online* pelo menos uma vez por mês e o sector das vendas é um dos que mais faz uso deste tipo de tecnologias.

A principal razão indicada pelos utilizadores para terem utilizado um *chatbot* em vez de outras opções disponíveis foi a vontade de querer experimentar este tipo de sistema. Isto confirma que os *chatbots* ainda são vistos principalmente como uma novidade, mas uma novidade que os utilizadores estão dispostos a experimentar devido às outras razões apresentadas, como a facilidade de uso e rapidez na obtenção da resposta a uma pergunta. Para além disso, 38,1% afirmaram que já tinham tido uma experiência positiva com o *chatbot*.

Quando questionados sobre o tipo de resposta que prefere obter do *chatbot*, a grande maioria indicou que prefere receber uma resposta avançada que combine vários elementos como texto, imagem e *link*. Este tipo de resposta é muito semelhante às respostas que são atualmente devolvidas pelos principais motores de busca, como o Google.

Mesmo nunca tendo utilizado um *chatbot*, a grande maioria dos utilizadores (96,6%) indicaram que seria interessante utilizar este tipo de sistema como ferramenta de apoio à pesquisa de conteúdos na Internet e 88,5% consideram que o facto de os *chatbots* devolverem apenas resultados com base no contexto da conversa é uma vantagem face aos motores de busca tradicionais. Verificou-se também que os utilizadores estão dispostos a partilhar alguns dados pessoais com um sistema de *chatbot*, o que mostra que teriam interesse em tirar partido de funcionalidades mais avançadas como conversas e sugestões personalizadas.

Todos os inquiridos demonstraram ainda interesse em utilizar *chatbot* para pesquisa de conteúdos num determinado site, mas sobretudo como ferramenta de apoio a outras formas de pesquisa. Apenas 27,6% afirmaram que utilizariam como ferramenta principal para a pesquisa. Isto mostra que a maioria dos utilizadores ainda não consideram que um *chatbot* seja a ferramenta mais adequada para efetuar esta tarefa de forma autónoma e que não estariam preparados para abandonar os processos de pesquisa atuais, que já estão enraizados nas suas rotinas diárias.

Sobre outras funcionalidades que gostariam de ver incluídas, as mais populares foram: conversas personalizadas (65,5%), possibilidade de obter resumos dos conteúdos gerados por inteligência artificial, sem necessidade de aceder ao sítio *web* original (62,1%), capacidade de partilhar facilmente conteúdos com outras pessoas (55,2%), análise de sentimento e inteligência emocional (44,8%) e possibilidade de reconhecer e traduzir diretamente os resultados apresentados para outras línguas (41,4%). Isto mostra que os utilizadores têm interesse em adicionar novas funcionalidades à pesquisa

que ainda não estão presentes nos motores de busca tradicionais. Por outro lado, 27,6% dos inquiridos indicaram que gostariam que um *chatbot* de apoio à pesquisa de conteúdos na Internet deveria poder transferir a conversa para um assistente em tempo real, o que sugere que, para alguns utilizadores, os *chatbots* podem ainda não ser suficientemente autónomos para completar esta tarefa.

Apesar de algumas ressalvas já identificadas, a grande maioria dos inquiridos (72,4%) considera que os *chatbots* têm potencial para substituir os motores de busca tradicionais e tornar-se na principal ferramenta de pesquisa de conteúdos na Internet, considerando-a uma tecnologia com um futuro promissor.

4.2 Análise de requisitos

Antes de se avançar com o desenho do protótipo, é importante fazer o levantamento dos requisitos necessários para o seu desenvolvimento e para determinar as funcionalidades que devem ser incluídas no *chatbot*. Desta forma, foram definidos um conjunto de requisitos funcionais e não funcionais que devem ser tidos em conta ao longo do seu ciclo de vida.

4.2.1 Requisitos funcionais

Os requisitos funcionais determinados para o *chatbot* a construir são os seguintes:

- O *chatbot* deve ser capaz de iniciar uma conversa e interagir com o utilizador;
- O *chatbot* deve conseguir compreender linguagem natural e ser capaz de identificar a intenção do utilizador, independentemente da forma como este construiu a oração;
- O *chatbot* deve ser capaz de ajudar os utilizadores a encontrarem o que procuram de forma rápida;
- O principal objetivo do *chatbot* deve ser devolver notícias;
- A funcionalidade de conversa casual será uma forma secundária de obter as preferências do utilizador;
- O *chatbot* deve ser capaz de corrigir erros ortográficos;
- As respostas devolvidas devem ser coerentes com o contexto;
- O sistema deve respeitar a privacidade do utilizador;
- O *chatbot* deve apresentar um layout intuitivo para melhorar a experiência do utilizador;
- O *chatbot* deve devolver respostas na língua do utilizador;
- O *chatbot* deve estar conectado a uma base de dados que contém a base de conhecimento das respostas a serem devolvidas ao utilizador;

- O sistema deve ser capaz de comunicar com uma base de dados externa e fazer pedidos GET a uma API (*Application Programming Interface*);
- As respostas do *chatbot* deverão ser apresentadas em formato de texto, imagem ou misto.

4.2.2 Requisitos não funcionais

Os requisitos não funcionais que foi possível identificar são os que se apresentam de seguida:

- Selecionar uma plataforma de criação de *chatbots* que permita implementações em diversos tipos de canais;
- Selecionar um canal que permita integrar o *chatbot* de forma simples e rápida;
- O *chatbot* deve poder ser utilizado pelo maior número de utilizadores possível;
- O utilizador deve conseguir ter acesso ao *chatbot* através de um canal adequado;
- As respostas às perguntas dos utilizadores devem ser rápidas;
- O código deve ser de fácil manutenção para desenvolvimentos futuros, tanto na sua base de conhecimento quanto na correção de pequenos problemas;
- O sistema deve estar sempre disponível;
- Deve existir documentação de suporte;
- Deve ser possível adicionar novas funcionalidades e customizações;
- O *chatbot* deve ser estável e resiliente;
- O *chatbot* deve ser transparente;
- O *chatbot* deve garantir a integridade e privacidade dos dados recolhidos;
- O *chatbot* deve obedecer aos critérios de usabilidade para conteúdo *web* WCAG 2.0.

4.3 Casos de uso

Nesta secção são apresentados os diagramas de caso de uso utilizados para descrever as ações que o utilizador irá desempenhar no sistema. Serão detalhados quatro casos de uso, nomeadamente:

- Caso de uso 1 - Obter notícias;
- Caso de uso 2 - Efetuar Registo;
- Caso de uso 3 - Efetuar Login;
- Caso de uso 4 - Conversa casual.

Nas secções seguintes detalham-se estes casos de uso.

4.3.1 Caso de Uso 1: Obter notícias

A obtenção de notícias é a principal funcionalidade do sistema. O caso de uso é ilustrado na Figura 16.

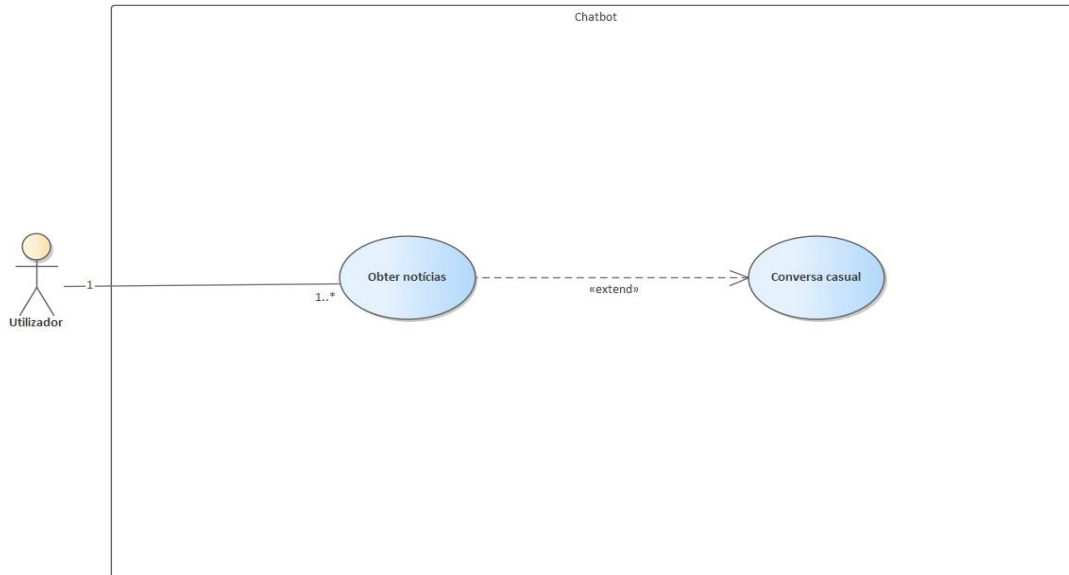


Figura 16 - Diagrama de caso de uso 1: "Obter notícias"

A especificação detalhada do caso de uso 1 é apresentada na Tabela 1.

Tabela 1 - Especificação do caso de uso 1: "Obter notícias"

Identificação	CU1
Nome	Obter notícias
Descrição	Este caso de uso tem por objetivo permitir a obtenção de notícias com base nas preferências indicadas pelo utilizador.
Condições prévias	-
Cenário	<ol style="list-style-type: none"> 1. O sistema dá as boas-vindas ao utilizador 2. O utilizador inicia uma conversa com o <i>chatbot</i> 3. O utilizador indica que não se quer registar 4. O sistema pede as preferências do utilizador 5. O utilizador indica as suas preferências 6. O sistema devolve uma notícia relacionada com as preferências do utilizador
Cenário alternativo	CA1t1. No passo 3, o utilizador indica que se quer registar. CA1t2. No passo 3, o utilizador indica que já está registado e faz o login
Exceções	E1. No passo 6, caso o sistema não encontre nenhuma notícia que corresponda à pesquisa, não devolve nenhuma notícia e pede novos dados ao utilizador para efetuar uma nova pesquisa.
Pós-condições	As preferências do utilizador são gravadas na base de dados e o utilizador pode continuar a usar o sistema ou terminar a conversa.

4.3.2 Caso de uso 2: Efetuar registo

O registo de um novo utilizador permite que o sistema guarde dados do utilizador para utilizar em conversas futuras. O caso de uso é ilustrado na Figura 17.

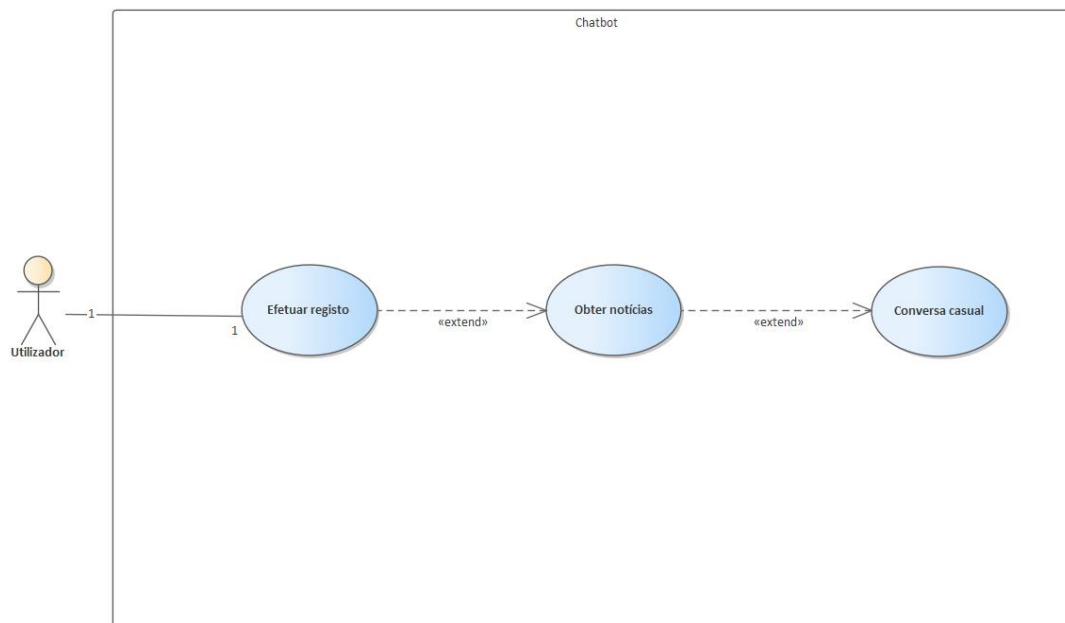


Figura 17 - Diagrama do caso de uso 2: “Efetuar registo”

Na Tabela 2 é feita a especificação detalhada do caso de uso 2.

Tabela 2 - Especificação do caso de uso 2: "Efetuar registo"

Identificação	CU2
Nome	Efetuar registo
Descrição	Este caso de utilização tem por objetivo registar o utilizador
Condições prévias	-
Cenário	<ol style="list-style-type: none"> 1- O sistema dá as boas-vindas ao utilizador 2- O utilizador inicia uma conversa com o <i>chatbot</i> 3- O utilizador indica que ainda não está registado 4- O sistema pede o nome e o email do utilizador 5- O sistema confirma o registo
Cenário alternativo	No passo 3, o utilizador pode indicar que já está registado, mas quando se procede à autenticação e o sistema não reconhece o utilizador, o utilizador pode optar por realizar um novo registo
Exceções	-
Pós-condições	Os dados do utilizador são gravados na base de dados e o utilizador pode continuar a usar o sistema.

4.3.3 Caso de uso 3: Efetuar login

O utilizador pode autenticar-se no sistema para obter uma experiência mais personalizada. O caso de uso 3 é ilustrado na Figura 18.

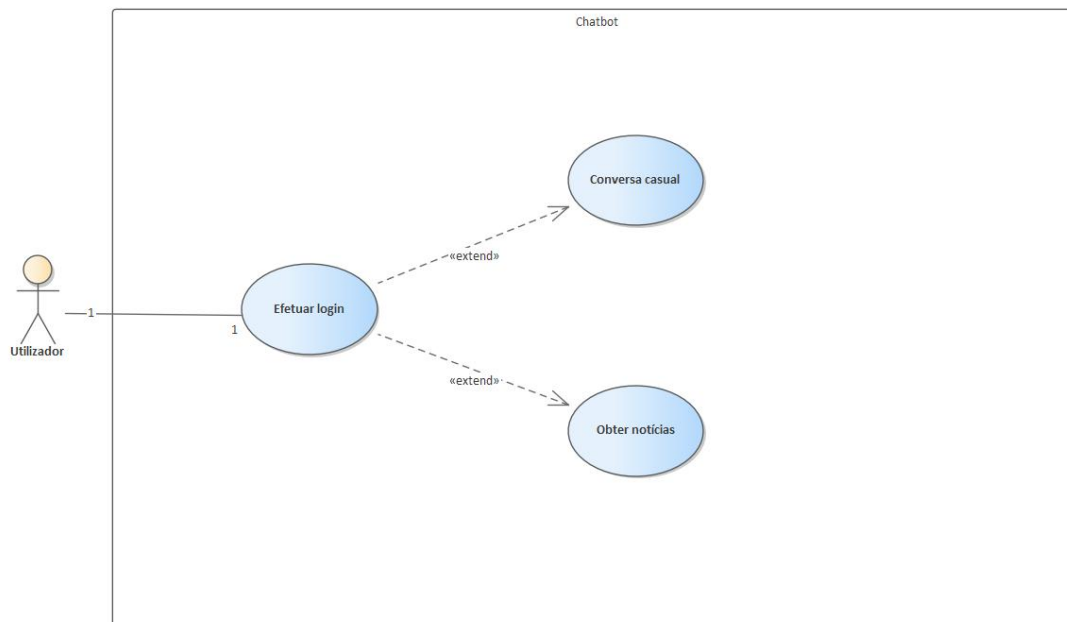


Figura 18 - Diagrama do caso de uso 3: "Efetuar login"

A especificação detalhada do caso de uso 3 é apresentada na Tabela 3.

Tabela 3 - Especificação do caso de uso 3: "Efetuar login"

Identificação	CU3
Nome	Efetuar login
Descrição	Este caso de uso tem por objetivo registar o utilizador
Condições prévias	-
Cenário	1- O sistema dá as boas-vindas ao utilizador 2- O utilizador inicia uma conversa com o <i>chatbot</i> 3- O utilizador indica que já está registado 4- O sistema pede o email do utilizador para autenticar 5- O sistema confirma que o utilizador está registado
Cenário alternativo	No passo 4, caso o sistema não reconheça o utilizador, o utilizador pode optar por realizar um novo registo
Exceções	-
Pós-condições	Os dados de início de sessão são gravados na base de dados e o utilizador pode continuar a usar o sistema.

4.3.4 Caso de uso 4: Conversa casual

O utilizador pode ter uma conversa casual com o *chatbot*. Este caso de uso é ilustrado na Figura 19.

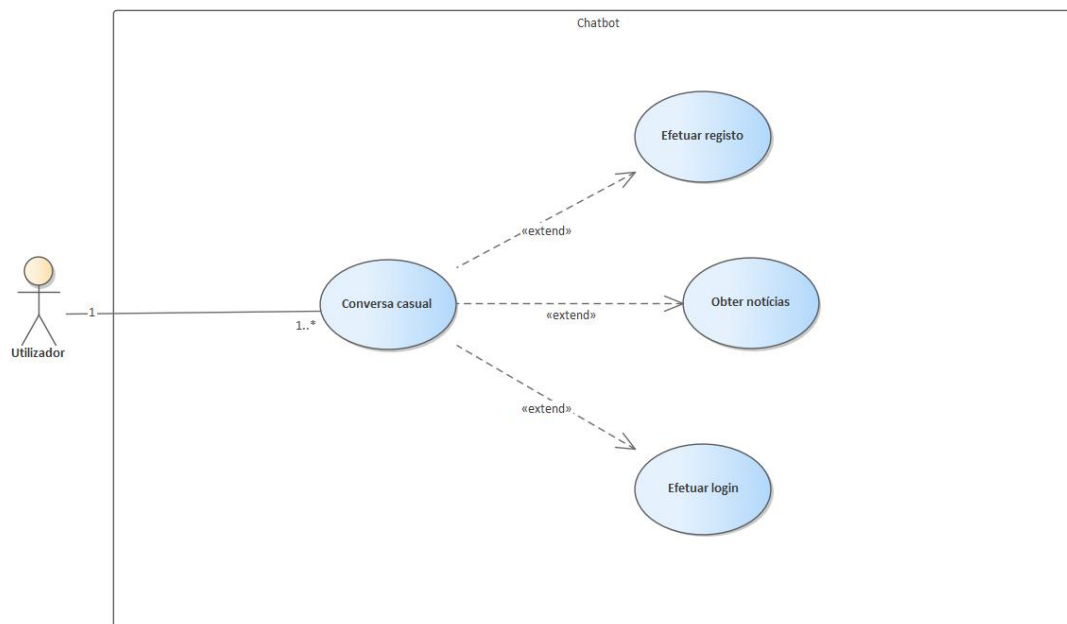


Figura 19 - Diagrama de caso de uso 4 - "Conversa casual"

Na Tabela 4 é feita a especificação detalhada do caso de uso 4.

Tabela 4 - Especificação do caso de uso 4: "Conversa casual"

Identificação	CU4
Nome	Conversa casual
Descrição	Este caso de uso tem por objetivo permitir uma conversa casual entre o sistema e o utilizador de modo a obter mais dados para pesquisa
Condições prévias	O utilizador tem de estar previamente registado e efetuar o login
Cenário	<ol style="list-style-type: none"> 1. O sistema dá as boas-vindas ao utilizador 2. O utilizador inicia uma conversa com o <i>chatbot</i> 3. O utilizador indica que já está registado 4. O sistema procede à autenticação do utilizador 5. O sistema inicia a conversa casual 6. O utilizador participa na conversa casual
Cenário alternativo	-
Exceções	E1. No passo 4, o sistema não consegue encontrar o registo desse utilizador. O utilizador pode tentar autenticar-se novamente, efetuar um novo registo ou prosseguir para a pesquisa sem se registar.
Pós-condições	As preferências do utilizador são gravadas na base de dados e o utilizador pode continuar a usar o sistema ou terminar a conversa.

4.4 Diagrama de processos

Para ilustrar o processo de interação a ser realizado, foi elaborado um diagrama de atividades, representado pela imagem da Figura 20, que representa a síntese do fluxo de interações entre o utilizador e o agente do *chatbot*.

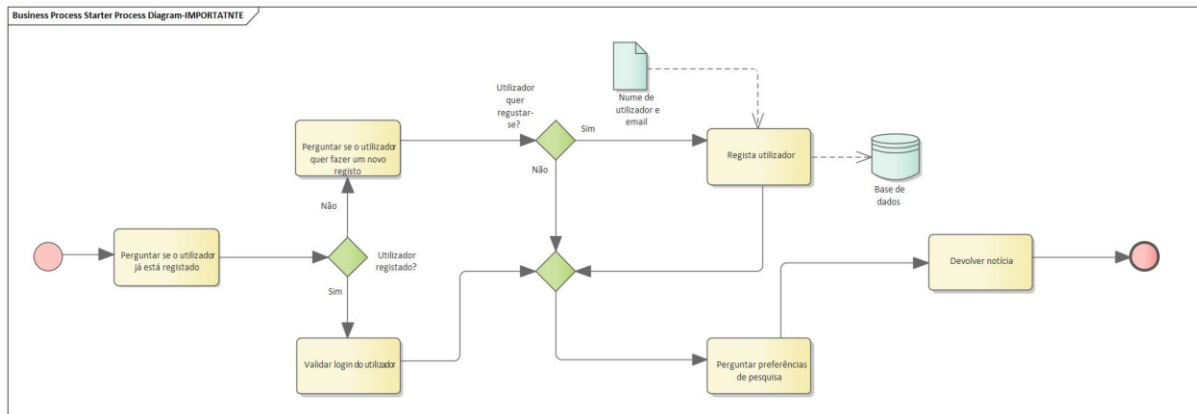


Figura 20 - Diagrama de processos

4.5 História de utilizador e storyboard

Para melhor entendimento do processo de interação a ser realizado, foi concebido um cenário baseado no caso de uso 1, "Obter notícias". Foi igualmente desenhado o *storyboard* que ilustra essa história de utilizador.

Cenário

O João é jornalista, está a trabalhar numa peça sobre vendas de motos e precisa de saber qual é a novidade mais recente das duas marcas de motos que mais veículos venderam no mês anterior. Para isso, dirige-se ao sítio *web* de notícias sobre motos e interage pela primeira vez com o *chatbot* "Rodas". O João é um utilizador frequente do sítio *web* e, por isso, aceita registar-se para obter resultados mais personalizados no futuro, indicando o seu nome e *email*. Em seguida, indica que tem interesse em notícias de motos e especifica uma marca, recebendo a notícia mais recente disponível no sítio *web* sobre essa marca. Tendo obtido a resposta sobre a primeira marca pesquisada, o João indica então a segunda marca que pretende pesquisar e obtém a notícia mais recente. Satisfeito com a resposta, o João resolve continuar a usar o sistema e indica que, desta vez, quer receber notícias de outros temas. Quando questionado sobre o interesse em desportos motorizados, o João rejeita o tema, mas aceita a sugestão do "Rodas" para saber a última novidade relacionada com veículos elétricos. Após receber a respetiva notícia, o João termina a conversa.

Na Figura 21 apresenta-se o storyboard que ilustra o cenário descrito.

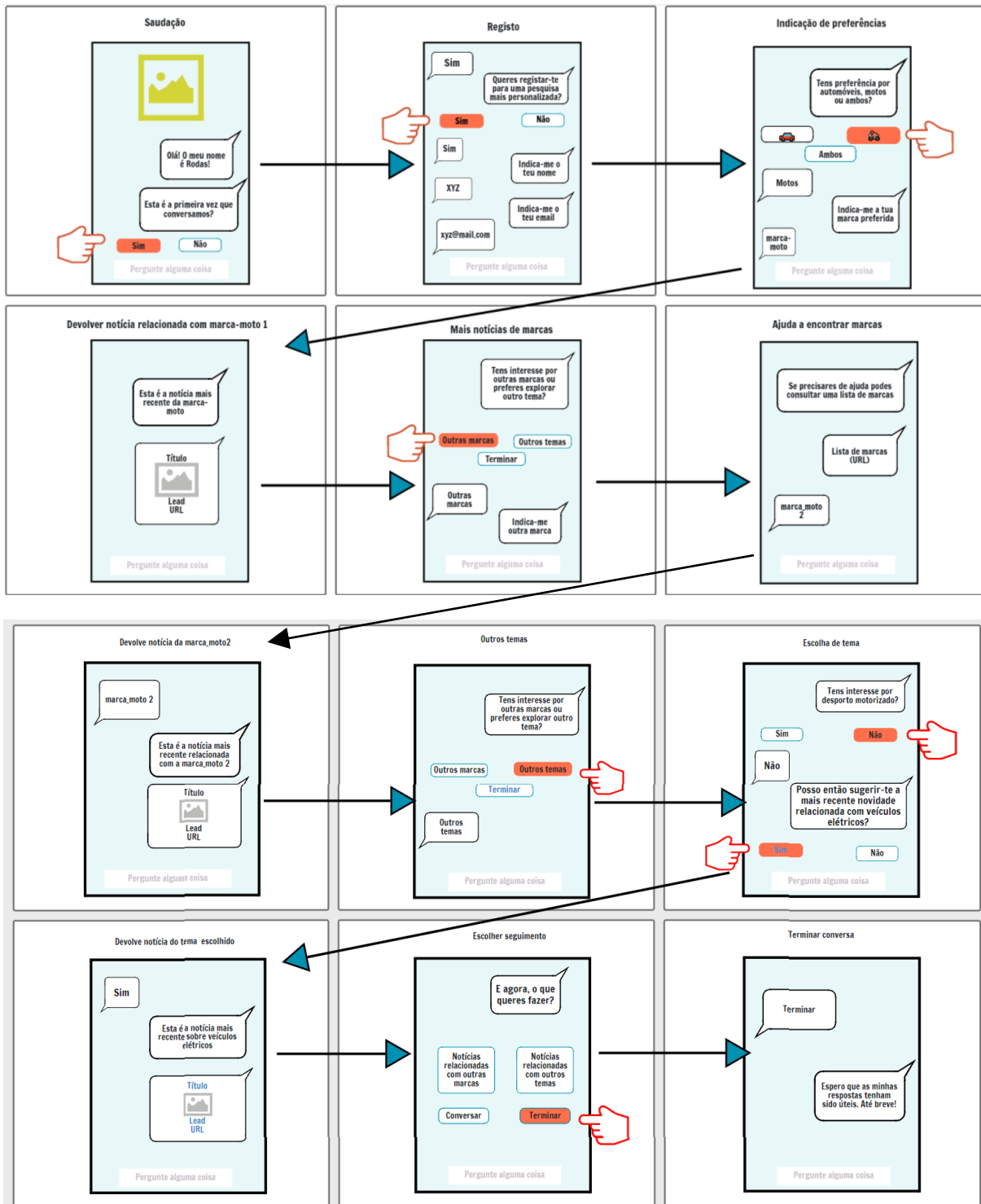


Figura 21 - Storyboard

4.6 Arquitetura

Nesta secção é apresentada a arquitetura do sistema de *chatbot* proposto, que será posteriormente implementada no capítulo 5.

A escolha do tipo de arquitetura para um sistema de *chatbot* depende de vários fatores como a audiência a que se destina, os recursos a que precisa de aceder e os modelos de comunicação que se pretende implementar.

Existem muitas plataformas que possibilitam o desenvolvimento de chatbots, como é mostrado na Tabela 5. Depois de uma análise inicial a algumas das propostas disponíveis, e tendo em consideração os requisitos apresentados previamente na secção 4.2, optou-se por utilizar o Dialogflow da Google.

Tabela 5 - Comparação de diferentes plataformas de desenvolvimento de *chatbots*

Plataforma	Tipo	Integrações	Facilidade de utilização	Suporte PLN	Preço
Amazon Lex	Texto, voz	Websites, Facebook, SMS	Intermédio	Sim	Pago
Botkit	Texto, voz	Websites, e-commerce	Intermédio	Integrado com MS LUIS	Grátis
Chatterbot	Texto	Websites	Difícil	Não	Grátis
Dialogflow	Texto, voz	Websites, Facebook, Google Assistant, Slack	Fácil	Sim	Grátis com limitações + Plano pago
IBM Watson	Texto, voz	Websites, Facebook, SMS, Slack	Intermédio	Sim	Grátis com limitações + Plano pago
Microsoft LUIS	Texto, voz	Websites, Facebook, Slack	Fácil	Sim	Grátis com limitações + Plano pago
Pandorobot	Texto voz	Websites, Facebook, SMS, Slack, ecommerce	Fácil	Não	Plano grátis mediante contacto com a empresa
Rasa	Texto, voz	Websites, Facebook, Slack, e-commerce	Intermédio	Sim	Grátis com limitações + Plano pago
SAP Conversational AI	Texto, voz	Soluções SAP, Backend	Fácil	Sim	Pago
Wit.AI	Texto voz	Facebook, Alexa	Fácil	Sim	Grátis

O Dialogflow é uma plataforma completa de PLN que inclui módulos de CLN e GLN e um gestor de diálogo responsável pelo fluxo da conversa entre um utilizador e o *chatbot*. Tem também um editor de código incorporado que permite configurar um serviço de *webhook* para executar ações e recuperar informações a partir de bases de dados externas.

Optou-se pela utilização do Google Sheets para criar uma base de dados simples com as informações da base de conhecimento que contêm as repostas do *chatbot* e que também armazena os dados que são introduzidos pelos utilizadores. Para aceder a esses dados, utiliza-se o serviço SheetDB,

que permite transformar as páginas do Google Sheets numa API JSON (*JavaScript Object Notation*) capaz de receber pedidos RESTful para ler, criar, atualizar e apagar registos.

A arquitetura final, que pode ser vista na Figura 22, fica completa com a IU, utilizada para fazer pedidos e receber respostas do *chatbot*.

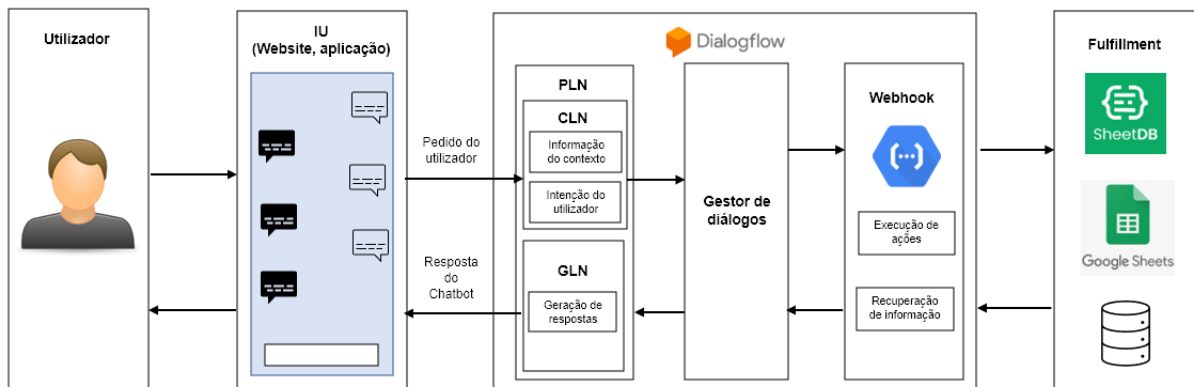


Figura 22 - Arquitetura do sistema de *chatbot* proposto

5 Desenvolvimento e implementação

Neste capítulo são descritos os passos realizados para o desenvolvimento e implementação do protótipo de sistema de *chatbot* proposto. Depois de se especificar o contexto e explicar em detalhe as tecnologias usadas, é descrito o processo de criação e treino do agente de conversação e, por último, a sua implementação.

5.1 Desenvolvimento do protótipo

Nesta secção é especificado com mais detalhe o desenvolvimento da arquitetura proposta na secção 4.6 do capítulo 4.

5.1.1 Contexto

O artefacto desenvolvido no âmbito deste trabalho consiste num sistema de *chatbot* que tem como objetivo ajudar os utilizadores a fazerem pesquisas num espaço contido dentro de um determinado sítio *web*. Nesse sentido, foram escolhidos dois sítios *web* de notícias de atualização diária, um sobre automóveis e outro sobre motos, dos quais foram selecionadas algumas secções para extrair os conteúdos a ser armazenados na base de conhecimento do sistema. No total, foram selecionadas seis secções, conforme indicado na Tabela 6.

Tabela 6 - Secções dos sítios web de onde foram extraídas as notícias

Secção	Descrição
Novidades Auto	Notícias relacionadas com marcas de automóveis e indústria automóvel
Desporto Auto	Notícias relacionadas com automobilismo
Automóveis elétricos	Notícias relacionadas com automóveis elétricos
Novidades Motos	Notícias relacionadas com marcas de motos e indústria das motos
Desporto Motos	Notícias relacionadas com motociclismo
Motos elétricas	Notícias relacionadas com motos elétricas

5.1.2 Base de dados

O sistema proposto utiliza uma base de dados simples que, por um lado, contém as informações relacionadas com as notícias que serão devolvidas pelo *chatbot* e, por outro, permite o armazenamento de determinados dados introduzidos pelo utilizador.

A solução encontrada para desenvolver e hospedar a base de dados foi o Google Sheets, uma ferramenta *online* e gratuita que permite criar e gerir folhas de cálculo de forma avançada e em tempo real.

Os conteúdos das notícias foram extraídos de cada sítio *web* através de técnicas de *web scraping*. Os elementos extraídos deram origem aos parâmetros utilizados posteriormente no Dialogflow e estão especificados na Tabela 7, que inclui também os parâmetros gerados a partir dos dados introduzidos pelos utilizadores.

Tabela 7 - Parâmetros utilizados no desenvolvimento do *chatbot*

Parâmetros dos conteúdos	Parâmetros do utilizador
Título	Nome
Lead	Email
URL	Tipo de Veículo
Conteúdo	Marca de Motos
Data	Marca de Automóveis
URL da imagem	Preferência Desporto
Secção	Preferência Eléctricos
Marca de Automóveis	Conversa casual
Marca de Motos	Data de registo

5.1.3 API externa

O *chatbot* requer uma integração externa para obter informação em tempo real. Apesar de existir uma API própria do Google Sheets, optou-se pela utilização do SheetDB, um *software* de gestão de APIs que permite usar folhas de cálculo criadas no Google Sheets como uma API JSON.

Esta solução oferece funcionalidades semelhantes a uma base de dados relacional e permite realizar solicitações RESTful. As solicitações e respostas são feitas com recurso ao Axios, um cliente HTTP (*Hypertext Transfer Protocol*) baseado em promessas para o Node.js, que é o ambiente utilizado para o desenvolvimento do código de *fulfillment*.

5.1.4 Dialogflow

O Dialogflow é uma plataforma de PLN que facilita o design da IU do agente conversacional e a sua integração com apps para dispositivos móveis, aplicações web ou outros sistemas interativos (Google Cloud, 2021b). É composto por módulos de CLN e GLN, um gestor de diálogo responsável pelo fluxo da conversa entre o utilizador e o *chatbot* e tem também um editor de código incorporado que pode ser usado para criar *scripts* em Node.js que permitem a execução de ações e recuperação de informação a partir da conexão a fontes de informação externas ou chamadas de API externas.

O Dialogflow permite utilizar configurações de aprendizagem máquina, nomeadamente algoritmos de classificação para remover resultados de falsos positivos e correção ortográfica automática.

Este *framework* permite dois tipos de *fulfillment*: via *webhook* e editor de código integrado (*inline editor*). O editor de código integrado é também ele próprio um *webhook*, mas hospedado pelo Google Cloud Functions, solução que permite criar funções autónomas capazes de responder a eventos na nuvem, sem necessidade de gerir um servidor ou um ambiente de execução. O editor de código integrado é compatível apenas com dois arquivos: o *index.js*, que contém todo o código de *fulfillment*, e o *package.json*, com todas as dependências especificadas na implementação.

Um *webhook* é um pedido HTTP que é enviado automaticamente sempre que um critério é satisfeito. No Dialogflow, o *webhook* pode ser usado para obter dados de um servidor externo sempre que é invocada uma intenção que tenha esta funcionalidade ativada.

Para este trabalho, utilizou-se o editor de código integrado, que se destina precisamente ao desenvolvimento de protótipos. O editor integrado é compatível apenas com Node.js e usa a biblioteca de *fulfillment* do Dialogflow por defeito que, apesar de não ser mais mantida, continua a ser uma opção válida para projetos menos complexos.

Quando se ativa o editor de código integrado, o código de *fulfillment* do Node.js aparece pré-preenchido com funções padrão e instruções comentadas, o que permite adicionar facilmente novas funções destinadas a gerir as intenções definidas para o agente. Na Figura 23 mostra-se um exemplo de uma destas funções.



```
index.js package.json
116 //5- grava dados do utilizador - Intent Boasvindas | Contexto: registosim | parametros: nomeutilizador; email
117 function BoasvindasHandler(agent) {
118     const nomeutilizador = agent.parameters.nomeutilizador;
119     const email = agent.parameters.email;
120     //axios.delete("https://sheetdb.io/api/v1/miyqsgp1zx9ai/all?sheet=Folha2");
121     axios.post("https://sheetdb.io/api/v1/miyqsgp1zx9ai?sheet=Folha2", {
122         "data": {
123             "Nome": nomeutilizador,
124             "Email": email,
125             "Created": new Date()
126         }
127     }).then(response => {
128         console.log(response.data);
129     });
```

Figura 23 - Parte do código de *fulfillment* utilizado

5.2 Desenho e treino do agente de conversação

O primeiro passo para o desenvolvimento de um *chatbot* no Dialogflow consiste na criação de um agente, um módulo de PLN capaz de traduzir textos introduzidos por um utilizador durante uma conversa para dados estruturados, de modo a serem entendidos por aplicações e serviços.

Há duas intenções que são criadas automaticamente para cada novo agente: uma intenção de boas-vindas, ativada quando um utilizador inicia uma conversa com uma saudação, e uma resposta padrão, acionada quando o agente não reconhece uma expressão do utilizador. Para este trabalho, foram criadas 22 novas intenções.

As intenções estabelecem a forma de mapear as expressões introduzidas pelo utilizador para uma resposta correspondente, a ser devolvida pelo sistema. Este mapeamento é feito com a ajuda de frases de treino, que são definidas dentro de cada intenção, e que correspondem a exemplos de possíveis expressões introduzidas pelo utilizador. Com esses exemplos, o agente pode avaliar quais as frases de treino mais aproximadas às frases introduzidas pelo utilizador, de modo a devolver a intenção mais apropriada. Na Figura 24 mostra-se como se realiza esse processo.

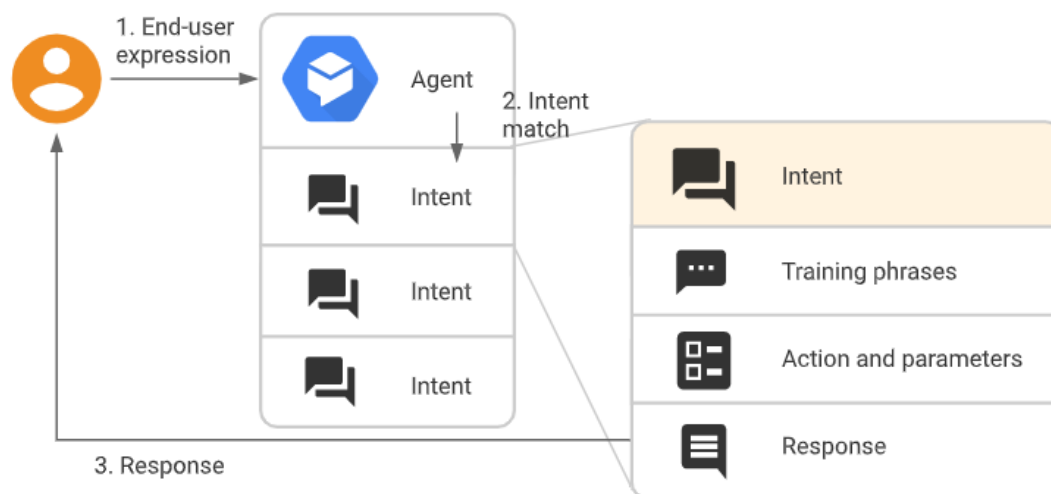


Figura 24 - Fluxo básico para correspondência de intenções e resposta ao utilizador (Google Cloud, 2021a)

As expressões introduzidas pelo utilizador podem conter informações específicas como um nome, um *email* ou uma data, que correspondem a entidades. As entidades podem igualmente ser especificadas nas frases de treino e podem ser de dois tipos: entidades do sistema ou entidades criadas pelo desenvolvedor (Figura 25).

When a user says something similar to a training phrase, Dialogflow matches it to the intent. You don't have to create an exhaustive list. Dialogflow will fill out the list with similar expressions. To extract parameter values, use [annotations](#) with available [system](#) or [custom](#) entity types.



Figura 25 - Exemplo de entidades anotadas nas frases de treino

Neste protótipo foram criadas 5 entidades, conforme especificado na Tabela 8. Nos sinónimos, para além de palavras, foram ainda introduzidos *emojicons* e *emojis*, símbolos visuais cada vez mais usados na comunicação em rede, que têm características semânticas e emocionais únicas. Estes símbolos ajudam a clarificar intenções em contextos ambíguos e melhoram a eficiência da comunicação (Bai, Danm Mu & Yang, 2019).

Tabela 8 - Entidades e sinónimos

Entidade	Sinónimos
@veiculos	Motos, automóveis, ambos, 🚗, 🏍️
@carros	Automóvel, carro, carrinha, jipe, descapotável, pick-up, 🚗
@motos	Moto, Mota, motorizada, bike, duas rodas, 🏍️
@marcaauto	Audi, BMW, Renault, Peugeot, Citroën, Honda, Ferrari, Mercedes-Benz (...)
@marcamoto	Honda, KTM, Vespa, Triumph, BMW, Ducati, Harley-Davidson, Kawasaki (...)

Na Figura 26 é mostrado um exemplo da relação entre algumas entidades e respetivos parâmetros na consola do Dialogflow.

Action and parameters ^

Boasvindas.Boasvindas-custom /

REQUIRED ?	PARAMETER NAME ?	ENTITY ?	VALUE	IS LIST ?	PROMPTS ?
<input checked="" type="checkbox"/>	marcaauto	@marcaauto	\$marcaauto	<input checked="" type="checkbox"/>	Boa! Eu também ...
<input type="checkbox"/>	carros	@carros	\$carros	<input type="checkbox"/>	–
<input type="checkbox"/>	Enter name	Enter entity	Enter value	<input type="checkbox"/>	–

[+ New parameter](#)

Figura 26 - Relação entre entidades e parâmetros

As intenções e as entidades são essenciais para construir respostas, mas para desenhar uma conversa é necessário um elemento que faça a ligação entre as diferentes intenções. Os contextos do Dialogflow são semelhantes ao contexto da língua natural e permitem controlar o fluxo de uma conversa. Desta forma, pode-se configurar contextos de entrada e saída em cada intenção e sempre que há uma correspondência num ou mais contextos de entrada, todos os contextos de saída configurados para essa intenção são ativados. Na Figura 27 é possível ver-se um exemplo da especificação de contextos para uma intenção.

Contexts ? ^

automoveis ⊗ Add input context

5 outrotema ⊗
5 outrasmarcascarros ⊗
5 terminar ⊗
✕

5 automoveis ⊗
2 Boasvindascarros-followup ⊗

Add output context

Figura 27 - Exemplo de contextos de entrada e saída

Quando uma intenção tem o *fulfillment* ativado, o Dialogflow chama o serviço de *webhook* que permite fazer solicitações à base de dados externa que contém as notícias, dados dos utilizadores e outras informações.

O *fulfillment* não está ativado, por defeito, na configuração das intenções. Quando é feita uma correspondência com uma intenção que não tem o *fulfillment* ativado, o Dialogflow utiliza uma resposta estática criada para essa intenção. A Figura 28 mostra um exemplo de uma resposta estática.

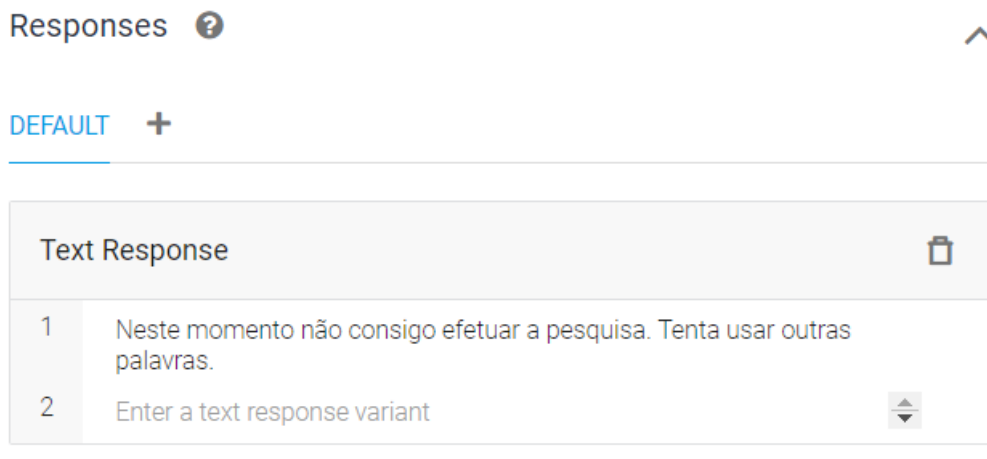


Figura 28 - Exemplo de resposta estática

5.3 Implementação

Como já foi referido na secção 5.2, o Dialogflow permite integrar os seus agentes em diversas plataformas e é totalmente compatível com o Google Assistente, o Slack, o Facebook Messenger, o Telegram, o LINE e as integrações do próprio Dialogflow (*Web Demo* e *Dialogflow Messenger*). Tendo em conta os requisitos previamente descritos na secção 3.4 do capítulo 3, era necessário desenvolver o *chatbot* de forma rápida e acessível, com o objetivo de alcançar o maior número de utilizadores possível.

Na primeira tentativa, optou-se por usar a solução de demonstração *Web* do próprio Dialogflow (Figura 29), que é a forma mais direta e rápida de integrar o *chatbot* num sítio *web*. No entanto, esta integração não oferece muitas opções de personalização e não é totalmente compatível com o *fulfillment*, devolvendo apenas mensagens estáticas.

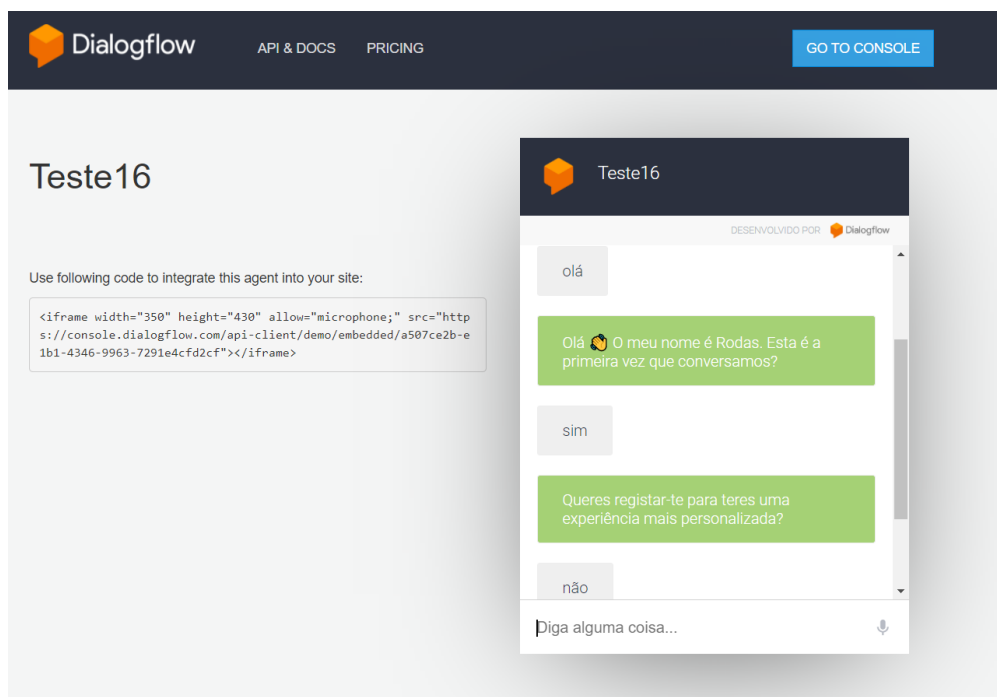


Figura 29 - Integração do chatbot no Dialogflow Web Demo

Em seguida, tentou-se integrar um agente de conversação no Facebook Messenger e no Telegram. Estas opções apresentavam mais possibilidades de customização e uma melhor interface gráfica face à apresentada no Dialogflow Web Demo, permitindo respostas dinâmicas com recurso ao *fulfillment*. Na Figura 30 pode ver-se um exemplo de interação entre um utilizador e o chatbot no Telegram.

No entanto, após o levantamento de requisitos e a análise de tarefas e utilizadores realizada no capítulo 3, verificou-se que o número de pessoas que utilizam estas aplicações de mensagens instantâneas é muito pequeno, sendo que atualmente a maior parte das pessoas utiliza o WhatsApp.

Como não havia nenhuma forma simples e direta de integrar o chatbot no WhatsApp, optou-se então por utilizar a integração Dialogflow Messenger (Figura 31).

Esta opção oferece uma caixa de diálogo de chat personalizável para o agente que pode ser incorporada num sítio web e, à semelhança das opções de integração em serviços de mensagens instantâneas, permite criar respostas dinâmicas com recurso ao *fulfillment*. A caixa de diálogo do chatbot é implementada como uma janela de diálogo que pode ser aberta e fechada pelo utilizador.

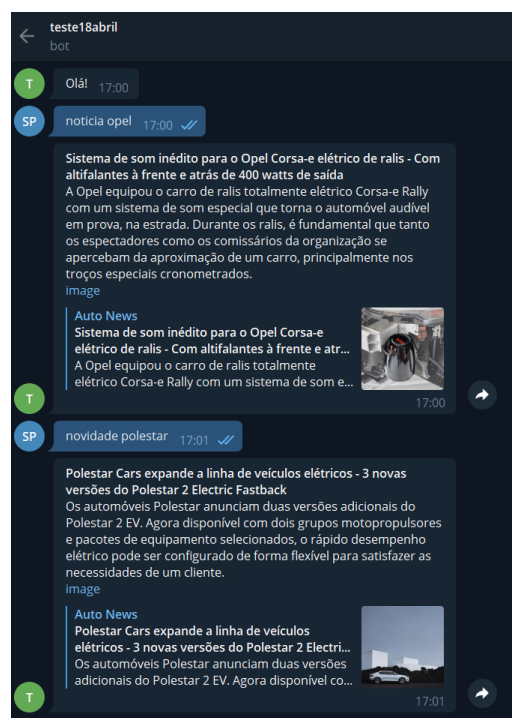


Figura 30 - Integração do chatbot no Telegram

Quando aberta, a caixa de diálogo de *chat* aparece acima do seu conteúdo, no canto inferior direito do ecrã.

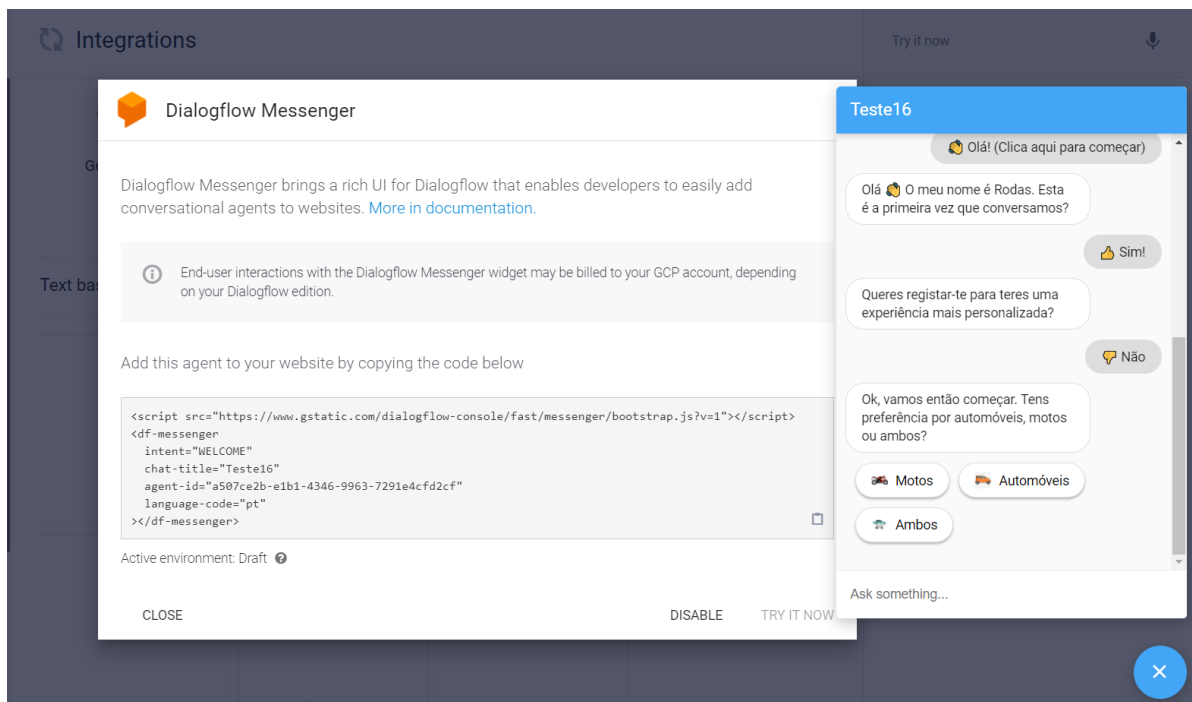


Figura 31 - Integração do *chatbot* no Dialogflow Messenger

Escolhida a integração, criou-se então um sítio *web* no qual foi incorporado o protótipo e que inclui também toda a documentação de apoio necessária para a utilização do *chatbot*.

O desenho da IU teve em consideração a simplicidade e o aspeto minimalista. O Dialogflow Messenger permite vários tipos de resposta avançada, nomeadamente no formato de cartões, que podem conter texto com *links* de ação, imagem, botão, listas, acordeão (um pequeno cartão que se pode expandir para revelar mais texto), ícones de sugestão ou uma combinação de todos estes tipos. Desta forma, o protótipo foi desenvolvido com uma combinação de vários elementos, tendo em conta a acessibilidade do utilizador.

Para facilitar a introdução de textos em ecrãs mais pequenos como os dos *smartphones*, optou-se, sempre que possível, pela utilização de botões para guiar a conversa. As cores foram escolhidas de modo a criar um contraste entre as falas do *chatbot* e as instruções do utilizador.

A Figura 32 mostra alguns exemplos dos ecrãs do *chatbot* implementado num sítio *web*.

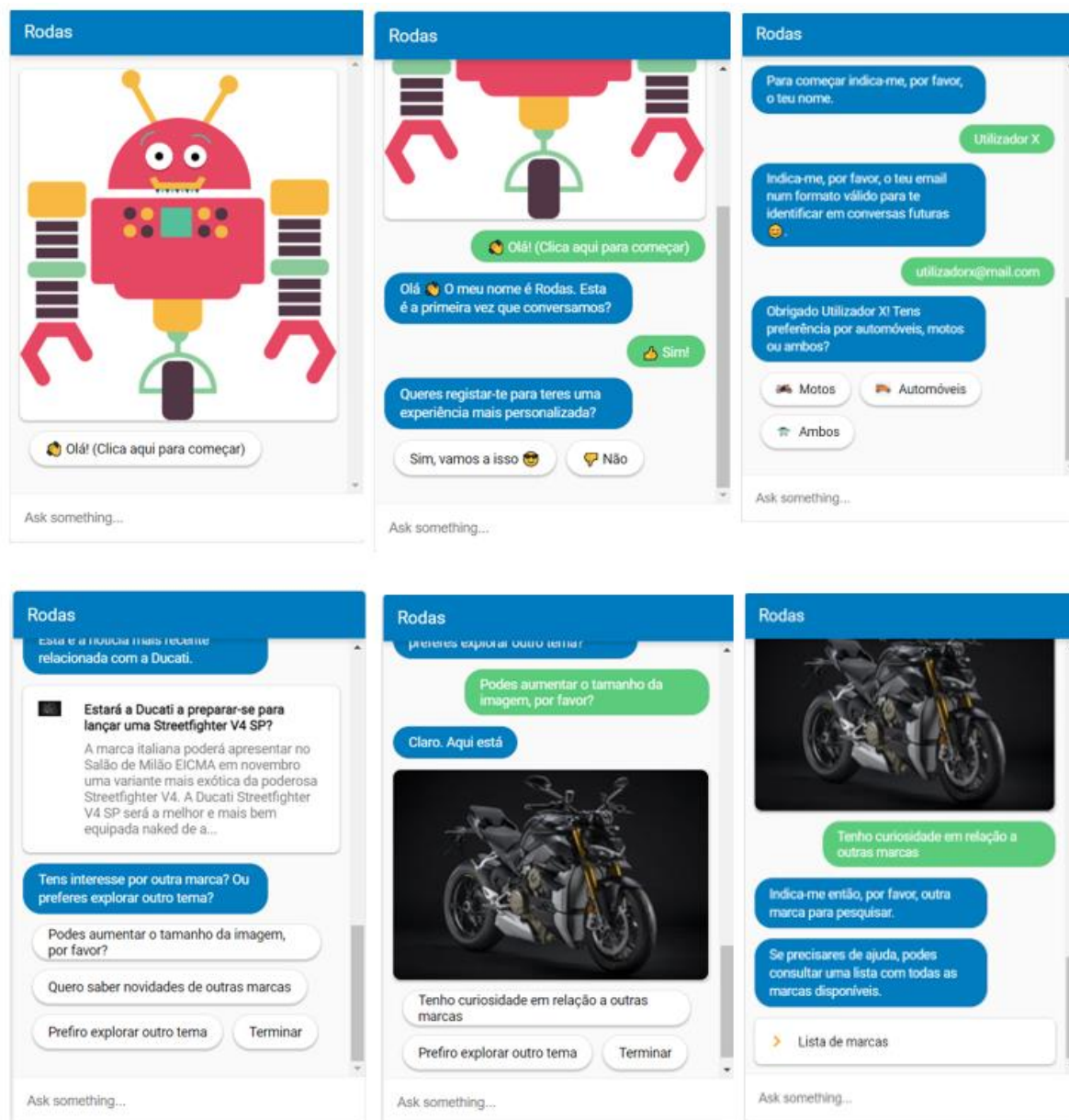


Figura 32 - Exemplos de ecrãs do *chatbot* implementado num sítio *web*

6 Avaliação

Neste capítulo, que corresponde à fase da avaliação da DSRM, são apresentados os resultados dos dados obtidos no âmbito da avaliação do sistema implementado, que foram recolhidos através de testes realizados com voluntários ao protótipo de *chatbot*.

6.1 Avaliação do protótipo

Apesar de, até à data, ainda não ter sido estabelecido nenhum padrão para a avaliação de *chatbots* (Maroengsit et al., 2019), a garantia do bom desenvolvimento do *chatbot*, tal como de outro qualquer *software*, depende de uma avaliação adequada.

Entre diversos métodos e métricas possíveis, neste estudo optou-se por uma sessão de testes com utilizadores, com o objetivo de aferir o seu grau de satisfação na utilização do sistema proposto. Desta forma, será possível avaliar a viabilidade da implementação de um sistema de *chatbot* para pesquisa inteligente de conteúdos no espaço contido de um determinado sítio *web*, através da avaliação empírica da forma como os utilizadores realizam as tarefas propostas.

Estes testes foram realizados num ambiente controlado, onde se recolheu dados relativos à forma como um utilizador interage com o sistema e os problemas com que se depara. O sistema foi avaliado em termos de eficácia, eficiência e realização dos objetivos.

O protótipo foi testado por um total de 9 voluntários, o que resultou numa amostra variada de utilizadores, com características distintas. Nenhum dos indivíduos teve contacto com o sistema antes do início dos testes.

Foi seguido um protocolo de testes de acordo com o guião disponível no Anexo II, tendo sido pedido aos voluntários que preenchessem um questionário inicial, seguido da realização de três tarefas no *chatbot* e, no final, o preenchimento de um questionário onde lhes era pedido para avaliar as funcionalidades do sistema testado e o grau de satisfação (Anexo III).

6.1.1 Caracterização da amostra

A amostra é constituída por nove indivíduos, de ambos os géneros, com idade superior a 25 anos, com habilitações literárias ao nível do ensino secundário ou ensino superior e diferentes níveis de conhecimentos sobre *chatbots*.

No que se refere ao género, 6 dos voluntários são do género feminino, o que equivale a 66,7%, e três utilizadores são do género masculino, correspondendo a 33,3% da totalidade dos inquiridos. Do

total de inquiridos, 77,8% têm habilitações literárias ao nível do ensino superior e 22,2% ao nível do ensino secundário (Figura 33).

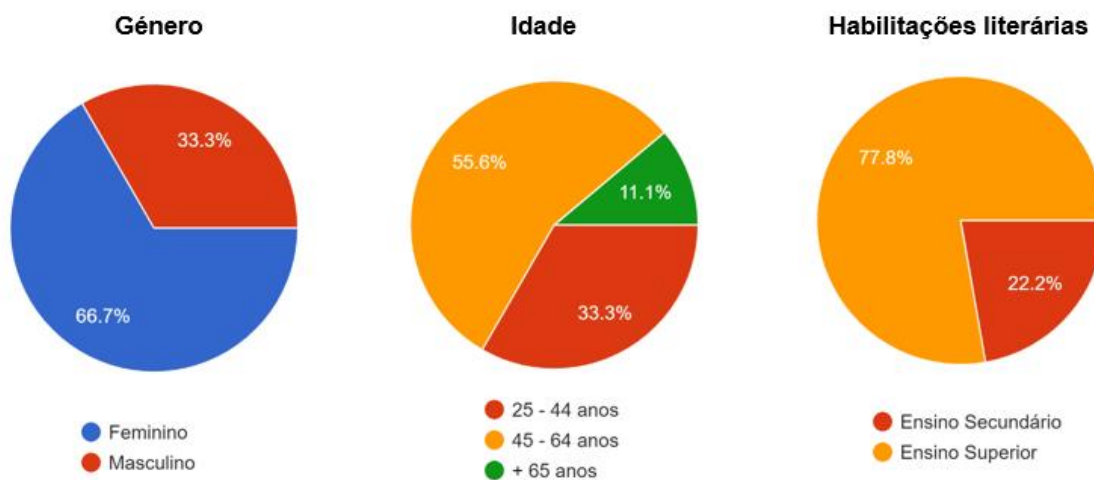


Figura 33 - Género, idade e habilitações literárias dos voluntários

Quando questionados sobre a utilização prévia de *chatbots*, 33,3% dos inquiridos admitem nunca ter utilizado este tipo de sistemas, enquanto 66,7% afirmam já ter usado anteriormente, conforme indicado no gráfico da Figura 34.

Já utilizou chatbots?

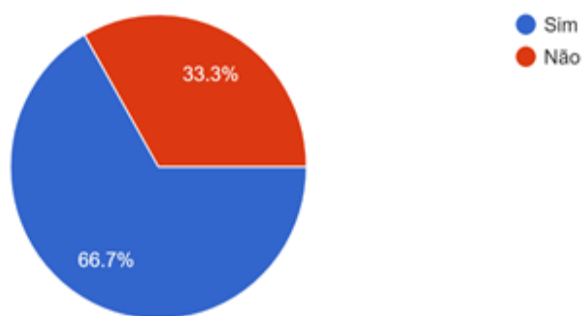


Figura 34 - Utilização prévia de *chatbots* pelos voluntários

6.1.2 Tarefas, medidas de usabilidade e critérios de usabilidade

Foi pedido aos voluntários que realizassem três tarefas no *chatbot* com o objetivo de aferir a eficácia e eficiência do sistema, bem como o grau de satisfação na utilização do mesmo. As três tarefas propostas foram as seguintes:

1ª Tarefa

Indicar que se trata de uma primeira interação e aceitar registrar-se para obter resultados mais personalizados; obter uma notícia relacionada com uma marca; obter uma notícia relacionada com desporto motorizado; obter uma notícia relacionada com veículos elétricos; terminar interação.

2ª Tarefa

Indicar que se trata de uma primeira interação e avançar sem aceitar registrar-se para obter dados mais personalizados; escolher um tipo de veículo; indicar a(s) marca(s) preferida(s); indicar que não tem interesse em notícias de desporto motorizado; indicar interesse em veículos elétricos; terminar interação.

3ª Tarefa

Indicar que já está registado; introduzir o *email* usado para o registo; conversar livremente; escolher um tipo de veículo; indicar que não tem interesse em notícias de desporto motorizado nem em notícias de veículos elétricos.

As medidas e critérios de usabilidade usadas nesta avaliação são as seguintes:

- **Eficácia:** Medida através do número de respostas satisfatórias e do número de utilizadores que terminaram com sucesso as tarefas propostas;
- **Eficiência:** Medida através da análise do funcionamento do sistema;
- **Satisfação:** Medida através da experiência de utilização, grau de dificuldade na utilização do sistema e recomendação do *chatbot* a outras pessoas.

6.1.3 Resultados e discussão

Depois de realizadas as três tarefas propostas, foi solicitado aos voluntários que preenchessem um questionário com base na interação com o protótipo. As respostas obtidas permitiram avaliar as funcionalidades do sistema ao nível da eficácia e eficiência, bem como o grau de satisfação do utilizador.

A eficácia é uma métrica que permite medir a relação entre as tarefas realizadas pelos utilizadores e a obtenção de resultados satisfatórios às suas pesquisas. Neste teste, a grande maioria dos voluntários (88,9%) conseguiram terminar todas as tarefas propostas, com apenas um utilizador a indicar que conseguiu terminar apenas uma tarefa (Figura 35).

Conseguiu terminar com sucesso todas as tarefas propostas?

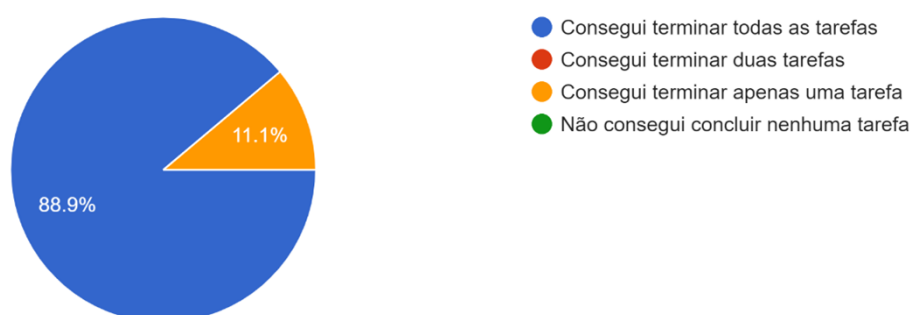


Figura 35 - Conclusão das tarefas propostas no teste de avaliação do *chatbot*

Esta situação em particular deveu-se à introdução de duas marcas ambíguas, que existem tanto para automóveis como para motos, e que não se encontrava bem mapeada na base de conhecimentos. Desta forma, este utilizador conseguiu cumprir com sucesso apenas a terceira tarefa, que não implicava a introdução de marcas, conforme especificado na Tabela 9, que regista quais foram as tarefas completadas por cada utilizador.

Tabela 9 - Tarefas completadas pelos voluntários

	U1	U2	U3	U4	U5	U6	U7	U8	U9
T1	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Não	Sim
T2	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Não	Sim
T3	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim

Tirando a exceção identificada, a maioria dos utilizadores indicou que obteve resultados que variam entre o medianamente satisfatório e o satisfatório.

Tendo em conta os dois elementos de avaliação usados, pode-se concluir que, no geral, os utilizadores consideram o sistema eficaz, como se pode ver na Figura 36, que mostra a classificação numa escala de 1 a 5, onde 1 é “Muito pouco eficaz” e 5 é “Muito eficaz”.

Conseguiu obter respostas satisfatórias com base nas tarefas propostas? Indique, por favor, a eficácia do chatbot.

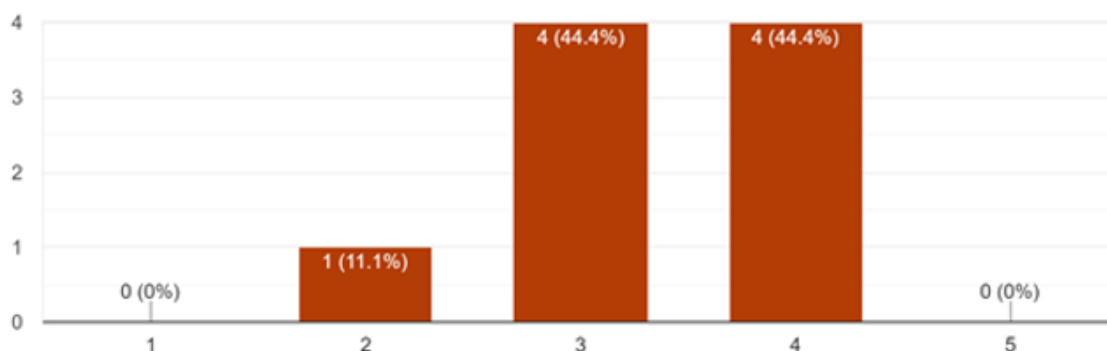


Figura 36 - Classificação da eficácia do *chatbot* pelos voluntários

Por sua vez, a eficiência é a métrica que permite verificar se o sistema funcionou conforme esperado. Esta métrica é medida através da averiguação da existência de erros que impossibilitaram o bom funcionamento do sistema.

Tendo já identificado previamente a situação do utilizador que indicou duas marcas ambíguas, houve ainda dois utilizadores que se depararam com uma falha no sistema e foram obrigados a recomeçar a sua interação com o *chatbot*. Apesar destes contratemplos, nenhum utilizador considerou que o *chatbot* não tivesse sido eficiente na execução das tarefas (Figura 37).

Durante a realização de cada tarefa, o chatbot deixou de funcionar corretamente ou foi necessário recomeçar a interação com o chatbot?

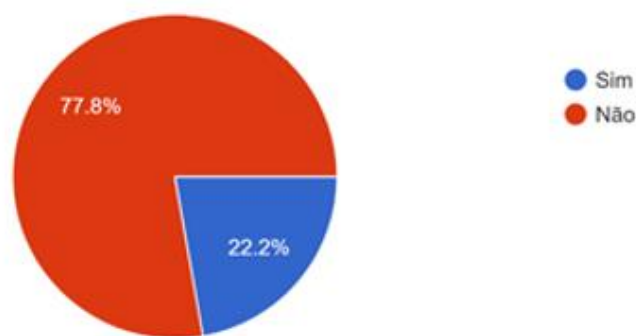


Figura 37 - Classificação da eficiência do *chatbot*

O grau de satisfação é uma das medidas do teste de usabilidade. Aferimos esta métrica através de perguntas relacionadas com a experiência de interação com o *chatbot*, grau de dificuldade na utilização do sistema e se estariam receptivos a recomendar o *chatbot* a outras pessoas.

Nenhum utilizador classificou de forma negativa a sua interação com o *chatbot* e a maior parte considerou que era de fácil utilização.

Os resultados obtidos indicam que, no geral, os utilizadores tiveram uma experiência positiva com o *chatbot*, o que pode ser comprovado pelo facto de a maioria dos utilizadores afirmar que recomendariam o *chatbot* testado a outras pessoas, como se pode constatar no gráfico da Figura 38.

Recomendaria o chatbot "Rodas" a outras pessoas?

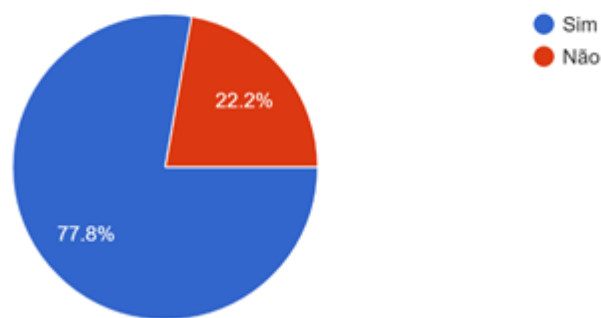


Figura 38 - Percentagem de utilizadores que recomendariam o *chatbot* a outras pessoas

7 Conclusão

Este trabalho foi concebido com o objetivo de verificar se seria possível e viável arquitetar e implementar um *chatbot* para apoiar a pesquisa de informação no espaço contido de um determinado sítio *web*. Para tal, foi utilizada a metodologia DSR.

Depois de uma breve introdução e da especificação do problema, foi realizada uma pesquisa exaustiva sobre gestão de conteúdos, *chatbots* e pesquisa de informação na Internet, bem como de outras tecnologias relacionadas, de modo a verificar o estado da arte e avaliar a viabilidade e o interesse na utilização destes sistemas.

Com base na revisão da literatura, verificou-se que a utilização de *chatbots* como ferramenta de pesquisa de informação na *web* é vista como uma ideia interessante e com potencial, tanto pela indústria como pela comunidade científica. No entanto, verificou-se que a recuperação de informação através destes sistemas é especialmente complexa, devido à linguagem natural, e que estes sistemas ainda têm capacidades de conversação muito limitadas. Foram ainda indicadas algumas oportunidades e desafios a ter em contra no desenvolvimento destas ferramentas

Para realizar a proposta do modelo conceptual de *chatbot*, começou-se por fazer a análise de utilizadores e de tarefas, com base num questionário. Esta análise permitiu, também, identificar as tarefas que os utilizadores precisam de desempenhar e a forma como o fazem atualmente, utilizando os sistemas existentes. Verificou-se que estes sistemas ainda não são utilizados com muita frequência, mas que há interesse na utilização destas ferramentas para a pesquisa de conteúdos na Internet, sobretudo devido às suas capacidades de inteligência artificial. Os *chatbots* são vistos pelos utilizadores como uma tecnologia com um futuro promissor e com potencial para se tornarem na principal ferramenta de pesquisa de conteúdos na Internet.

Ainda nesta fase, foi feito o levantamento de requisitos e apresentados quatro casos de uso. As funcionalidades do sistema foram demonstradas através de um cenário, acompanhado por um *storyboard*.

Depois de se detalhar e implementar a arquitetura do protótipo de *chatbot* num sítio *web*, procedeu-se à avaliação do sistema através de testes com utilizadores. O protótipo foi avaliado em termos de eficácia, eficiência e realização dos objetivos. Os resultados obtidos indicam que, no geral, os utilizadores tiveram uma experiência positiva com o *chatbot*.

Pode-se concluir que o estudo e desenvolvimento de *chatbots* dotados de inteligência artificial e com capacidade de processamento de linguagem natural é um tema atual e que há um interesse específico no desenvolvimento de sistemas que utilizem a funcionalidade de pesquisa para auxiliar a procura de informações em sítios *web*, sendo possível e viável arquitetar e treinar um sistema para essa finalidade.

7.1 Limitações

Durante o desenvolvimento do protótipo surgiram alguns desafios:

- Inicialmente pensou-se em usar uma base de dados em MySQL para armazenar os conteúdos dos sítios *web* e os dados dos utilizadores. Contudo, por um lado, não havia nenhuma forma direta de exportar essa base de dados para o Dialogflow e, por outro, não tinha os requisitos necessários para funcionar como serviço de *webhook*;
- Antes da escolha do SheetDb para a criação da API em JSON, considerou-se a utilização da API própria do Google Sheets. No entanto, esta apresentava um nível de complexidade demasiado elevado e era necessário encontrar uma solução mais simples e de rápida implementação;
- Ponderou-se a utilização de ferramentas de automação para facilitar a conexão entre o Dialogflow e a base de dados e chegou-se mesmo a fazer ensaios com o Microsoft Power Automate, o n8n.io, o Zapier, o Integrate.ly e o Integromat. Os dois primeiros foram descartados devido à sua complexidade. Já o Zapier, o Integrate.ly e o Integromat, apesar de até terem integrações diretas com diversas API, incluindo a API do Dialogflow, e de oferecerem ferramentas intuitivas e rápidas para visualizar, desenhar e criar *workflows* sem código, foram rejeitados pelas suas diversas limitações (preço, número de operações, intervalo entre operações e documentação escassa).
- A biblioteca de *fulfillment* do Dialogflow já não é mantida e a documentação sobre a sua utilização é escassa e obsoleta;
- A utilização do Facebook Messenger, Telegram, Slack, Line e outros canais possíveis para a implementação direta de um agente de um *chatbot* desenvolvido no Dialogflow implicam o registo nas respetivas plataformas. Desta forma, optou-se por implementar o protótipo num sítio *web*.

7.2 Trabalho futuro

Em termos de trabalho futuro, existem muitas linhas de desenvolvimento que podem ser seguidas:

- Comparar o uso dos motores de busca com o do *chatbot* em termos de eficácia e facilidade de utilização;
- Repensar a IU do *chatbot* e a forma como os conteúdos são apresentados, de modo a torná-lo mais acessível a partir de ecrãs mais pequenos como os dos *smartphones*. A criação automática de resumos gerados por inteligência artificial pode ser uma forma de diminuir a quantidade de informação apresentada e também pode facilitar a leitura e a compreensão do texto;
- Explorar a integração de funcionalidades de análise de sentimento e inteligência emocional para tentar identificar as emoções dos utilizadores;
- Expandir o modelo a outros casos de uso, nomeadamente dentro das áreas da gestão da informação, comunicação e *business intelligence*.

Bibliografia

- Adamopoulou, E., & Moussiades, L. (2020). An Overview of Chatbot Technology. *IFIP Advances in Information and Communication Technology*, 373–383. https://doi.org/10.1007/978-3-030-49186-4_31
- Agrawal, P., Menon, T., Kam, A., Naim, M., Chouragade, C., Singh, G., Kulkarni, R., Suri, A., Katakam, S., Pratik, V., Bansal, P., Kaur, S., Raj-put, N., Duggal, A., Chalabi, A., Choudhari, P., Satti, S. R., Nayak, N. & Rajput, N. (2020). QnAMaker: Data to Bot in 2 Minutes. *Companion Proceedings of the Web Conference 2020*, 131–134. <https://doi.org/10.1145/3366424.3383525>
- Aleedy, M., Shaiba, H. & Bezbradica, M. (2019). Generating and Analyzing Chatbot Responses using Natural Language Processing. *International Journal of Advanced Computer Science and Applications (IJACSA)*, 10(9). <http://dx.doi.org/10.14569/IJACSA.2019.0100910>
- Alguliev, R., & Mahmudov, R. (2016). The internet as a driving force of the development of civilization. *Problems of Information Society*, 07(1), 31–40. <https://doi.org/10.25045/jpis.v07.i1.04>
- Aliannejadi, M., Zamani, H., Crestani, F. & Croft, W. B. (2019). Asking clarifying questions in open-domain information seeking conversations. *Proceedings of the 42nd International ACM SIGIR Conference on Re-search and Development in Information Retrieval (SIGIR'19)*. 475-484. <https://doi.org/10.1145/3331184.3331265>
- Alloghani M., Al-Jumeily D., Mustafina J., Hussain A., Aljaaf A.J. (2020). A Systematic Review on Supervised and Unsupervised Machine Learning Algorithms for Data Science. In: Berry M., Mohamed A., Yap B. (eds) *Supervised and Unsupervised Learning for Data Science. Unsupervised and Semi-Supervised Learning*. Springer, Cham. https://doi.org/10.1007/978-3-030-22475-2_1
- Alpaydin, E. (2020). *Introduction to Machine Learning, fourth edition (Adaptive Computation and Machine Learning series)* (fourth edition). The MIT Press.
- Alzubi, J., Nayyar, A., & Kumar, A. (2018). Machine Learning from Theory to Algorithms: An Overview. *Journal of Physics: Conference Series*, 1142, 012012. <https://doi.org/10.1088/1742-6596/1142/1/012012>
- Angheliescu, P., & Nicolaescu, S.V. (2018). Chatbot Application using Search Engines and Teaching Methods. 2018 10th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), 1-6. <https://doi.org/10.1109/ECAI.2018.8678948>
- Augello A., Gentile M., Dignum F. (2018). An Overview of Open-Source Chatbots Social Skills. In: Diplaris S., Satsiou A., Følstad A., Vafopoulos M., Vilarinho T. (eds) *Internet Science. INSCI 2017. Lecture Notes in Computer Science, vol 10750*. Springer, Cham. https://doi.org/10.1007/978-3-319-77547-0_18
- Azmy, W. & Emam, O. (2013). Content-based recommendation system using search engine. *The Egyptian Society of Language Engineering, ESOLE'13, Cairo*, 13. <https://www.researchgate.net/publication/267631372>

- Bach, J. (2019). When Artificial Intelligence Becomes General Enough to Understand Itself. Commentary on Pei Wang's paper "On Defining Artificial Intelligence". *Journal of Artificial General Intelligence*, vol.11, no.2, 2019, pp.15-18. <https://doi.org/10.2478/jagi-2019-0002>
- Baez M., Daniel F., Casati F. (2020). Conversational Web Interaction: Proposal of a Dialog-Based Natural Language Interaction Paradigm for the Web. In: Følstad A. et al. (eds) *Chatbot Research and Design. CONVERSATIONS 2019. Lecture Notes in Computer Science*, vol 11970. Springer, Cham. https://doi.org/10.1007/978-3-030-39540-7_7
- Bai, Q., Dan, Q., Mu, Z. & Yang, M. (2019). A Systematic Review of Emoji: Current Research and Future Perspectives. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2019.02221>
- Baishya, A., & Kakoty, S. (2019). A Review on Web Content Filtering, Its Technique and Prospects. *International Journal of Computer Science Trends and Technology (IJCST) – Volume 7, Issue 3*. <https://doi.org/10.33144/23478578/IJCST-V7I3P5>
- Barker, D. (2016). *Web Content Management: Systems, Features, and Best Practices* (1st ed.). O'Reilly Media.
- Behere, T., Vaidya, A., Bihade, A., Shinde, K., Deshpande, P., & Jahira-badkar, S. (2020). Text Summarization and Classification of Conversation Data between Service Chatbot and Customer. *2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4)*, pp. 833-838. <https://doi.org/10.1109/worlds450073.2020.9210289>
- Berriault-Poirier A., Prom Tep S., Sénécal S. (2019). Putting Chatbots to the Test: Does the User Experience Score Higher with Chatbots Than Websites?. In: Ahram T., Karwowski W., Taiar R. (eds) *Human Systems Engineering and Design. IHSED 2018. Advances in Intelligent Systems and Computing*, vol 876. Springer, Cham. https://doi.org/10.1007/978-3-030-02053-8_32
- Boiko, B. (2005). *Content Management Bible*, 2nd Edition. Wiley
- Boyd, M., & Wilson, N. (2018). Just ask Siri? A pilot study comparing smartphone digital assistants and laptop Google searches for smoking cessation advice. *PLOS ONE*, 13(3), e0194811. <https://doi.org/10.1371/journal.pone.0194811>
- Brandtzaeg, P. B., & Følstad, A. (2017). Why People Use Chatbots. *Internet Science*, 377–392. https://doi.org/10.1007/978-3-319-70284-1_30
- Carlander-Reuterfelt, D., Carrera, Á., Iglesias, C., Araque, Ó., Rada, J.F., & Muñoz, S. (2020). JAICOB: A Data Science Chatbot. *IEEE Access*, 8, pp. 180672-180680. <https://doi.org/10.1109/ACCESS.2020.3024795>
- Casas, J., Tricot, M.O., Abou Khaled, O., Mugellini, E., & Cudré-Mauroux, P. (2020). Trends & Methods in Chatbot Evaluation. *Companion Publication of the 2020. International Conference on Multimodal Interaction*, 280–286. <https://doi.org/10.1145/3395035.3425319>
- Castells, Manuel (2007). *A galáxia Internet – Reflexões sobre Internet, negócios e sociedade*. Lisboa, Fundação Calouste Gulbenkian

- Chau M., Chen H. (2003). Personalized and Focused Web Spiders. In: Zhong N., Liu J., Yao Y. (eds) *Web Intelligence*. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-05320-1_10
- Chau, M., Chen, H., Qin, J., Zhou, Y., Qin, Y., Sung, W., & McDonald, D. (2002). Comparison of two approaches to building a vertical search tool: a case study in the nanotechnology domain. In *Proceedings of the 2nd ACM/IEEE-CS joint conference on Digital libraries (JCDL '02)*. Association for Computing Machinery, New York, NY, USA, 135–144. <https://doi.org/10.1145/544220.544246>
- Croft, W. B. (2019). The Importance of Interaction for Information Retrieval. In *SIGIR (Vol. 19)* pp. 1-2. <https://www.sigir.org/sigir2019/slides/10.1145-3331184.3331185.pdf>
- da Silva A.S., Gao X., Andreae P. (2014). Wallace: Incorporating Search into Chatting. In: Pham D. N., Park S. B. (eds) *PRICAI 2014: Trends in Artificial Intelligence. PRICAI 2014. Lecture Notes in Computer Science, vol 8862*. Springer, Cham. https://doi.org/10.1007/978-3-319-13560-1_68
- Daniel, F., Matera, M., Zaccaria, V., & Dell'Orto, A. (2018). Toward truly personal chatbots. *Proceedings of the 1st International Workshop on Software Engineering for Cognitive Services*, 31–36. <https://doi.org/10.1145/3195555.3195563>
- Daswani M., Desai K., Patel M., Vani R., Eirinaki M. (2020). CollegeBot: A Conversational AI Approach to Help Students Navigate College. In: Stephanidis C., Kurosu M., Degen H., Reinerman-Jones L. (eds) *HCI International 2020 - Late Breaking Papers: Multimodality and Intelligence. HCII 2020. Lecture Notes in Computer Science, vol 12424*. Springer, Cham. https://doi.org/10.1007/978-3-030-60117-1_4
- Ding, H., Ranade, N., & Cata, A. (2019). Boundary of content ecology. *Proceedings of the 37th ACM International Conference on the Design of Communication*, 1–2. <https://doi.org/10.1145/3328020.3353931>
- Eisman, E. M., López, V., & Castro, J. L. (2012). A framework for designing closed domain virtual assistants. *Expert Systems with Applications*, 39(3), 3135–3144. <https://doi.org/10.1016/j.eswa.2011.08.177>
- Fafalios, P., Holzmann, H., Kasturia, V. & Nedji, W. (2017). Building and querying semantic layers for web archives. *ACM/IEEE Joint Conference on Digital Libraries (JCDL)*, 1-10. <https://doi.org/10.1109/JCDL.2017.7991555>
- Fang, B., Jia, Y., Li, X., Li, A. & Wu, X. (2017). Big search in cyberspace. *IEEE Transactions on Knowledge and Data Engineering*, 29(9), 1793-1805. <https://doi.org/10.1109/TKDE.2017.2699675>
- Galitsky B. (2019a). A Content Management System for Chatbots. In: *Developing Enterprise Chatbots*. Springer, Cham. https://doi.org/10.1007/978-3-030-04299-8_9
- Galitsky B. (2019b). A Social Promotion Chatbot. In: *Developing Enterprise Chatbots*. Springer, Cham. https://doi.org/10.1007/978-3-030-04299-8_12
- Galitsky B. (2019c). Learning Discourse-Level Structures for Question Answering. In: *Developing Enterprise Chatbots*. Springer, Cham. https://doi.org/10.1007/978-3-030-04299-8_7

- Galitsky B. (2021a). A Virtual Social Promotion Chatbot with Persuasion and Rhetorical Coordination. In: *Artificial Intelligence for Customer Relationship Management. Human–Computer Interaction Series*. Springer, Cham. https://doi.org/10.1007/978-3-030-61641-0_4
- Galitsky B. (2021b). Chatbots for CRM and Dialogue Management. In: *Artificial Intelligence for Customer Relationship Management. Human–Computer Interaction Series*. Springer, Cham. https://doi.org/10.1007/978-3-030-61641-0_1
- Garrett, J. J. (2010). *The Elements of User Experience: User-Centered Design for the Web and Beyond* (2nd Edition) (Voices That Matter) (2nd ed.). New Riders.
- Goh, O. S., Fung, C. C., & Wong, K. W. (2007). Query Based Intelligent Web Interaction with Real World Knowledge. *New Generation Computing*, 26(1), 3–22. <https://doi.org/10.1007/s00354-007-0031-7>
- Goodfellow, Ian, Bengio, Yoshua, & Courville, Aaron. (2017). *Deep learning*. The MIT Press.
- Google Cloud (2021a). *Conceitos básicos do Dialogflow ES*. <https://cloud.google.com/dialogflow/es/docs/basics>
- Google Cloud (2021b). *Dialogflow*. <https://cloud.google.com/dialogflow/docs/>
- Gupta, A., Vijayakumar, A., & Hathwar, D. (2020). Introduction to AI Chatbots. *International Journal of Engineering Research And*, V9(07), 255–258. <https://doi.org/10.17577/ijertv9is070143>
- Hariguna, T. & Efendi, Y. (2020). Chatbot Smart Assistant Using N-Gram and Bi-Gram Algorithm. *International Journal of Advanced Trends in Computer Science and Engineering*, 9(4), 4788–4793. <https://doi.org/10.30534/ijatcse/2020/86942020>
- Hammoudeh, A. (2018). A concise introduction to reinforcement learning. <https://doi.org/10.13140/RG.2.2.31027.53285>
- Hazem, T. J., Olteanu, A., Kazai, G., Diaz, F. & Golebiewski, M. (2020). On the social and technical challenges of web search autosuggestion moderation. arXiv [cs.CY]. <https://arxiv.org/abs/2007.05039>
- Heidemann, J., Klier, M., & Probst, F. (2012). Online social networks: A survey of a global phenomenon. *Computer Networks*, 56(18), 3866–3878. <https://doi.org/10.1016/j.comnet.2012.08.009>
- Henzinger, M. R. (2004). Algorithmic Challenges in Web Search Engines. *Internet Mathematics*, 1(1), 115–123. <https://doi.org/10.1080/15427951.2004.10129079>
- Hernández-Orallo, J. (2016). Evaluation in artificial intelligence: from task-oriented to ability-oriented measurement. *Artificial Intelligence Review*, 48(3), 397–447. <https://doi.org/10.1007/s10462-016-9505-7>
- Hevner A. & Chatterjee S. (2010). Design science research in information systems. *Integrated Series in Information Systems (ISIS)*, 22, 9-22. Boston, MA: Springer. <https://doi.org/10.1007/978-1-4419-5653-8>

- Hevner, A. R., March, S. T., Park, J. & Ram, S. (2004). Design science in information systems research. *MIS Quarterly*, 28(1), 75-105. <https://doi.org/10.2307/25148625>
- Jain, M., Kumar, P., Kota, R., & Patel, S. (2018). Evaluating and Informing the Design of Chatbots. *Proceedings of the 2018 Designing Interactive Systems Conference*, pp 895-906. <https://doi.org/10.1145/3196709.3196735>
- Jankulovski, J., & Mitrevski, P. (2017). Cloud Computing and Content Management Systems : A Case Study in Macedonian Education. *International Journal on Cloud Computing: Services and Architecture*, 7(5), 01–13. <https://doi.org/10.5121/ijccsa.2017.7501>
- Jatana, Vansh. (2019). *Machine Learning Algorithms*. <https://doi.org/10.13140/RG.2.2.20559.92329>
- Jenkins MC., Churchill R., Cox S., Smith D. (2007). Analysis of User Interaction with Service Oriented Chatbot Systems. In: Jacko J.A. (eds) Human-Computer Interaction. *HCI Intelligent Multimodal Interaction Environments. HCI 2007. Lecture Notes in Computer Science, vol 4552*. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-73110-8_9
- Jeong, SS., Seo, YS. (2019). Improving response capability of chatbot using twitter. *Journal of Ambient Intelligence and Humanized Computing*. <https://doi.org/10.1007/s12652-019-01347-6>
- Johannsen, F., Schaller, D., & Klus, M. F. (2020). Value propositions of chatbots to support innovation management processes. *Information Systems and E-Business Management* 19, 205-246. <https://doi.org/10.1007/s10257-020-00487-z>
- Kaplan, A. & Haenle, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. In *Business Horizons, Volume 62*, Issue 1, pp 15-25. <https://doi.org/10.1016/j.bushor.2018.08.004>
- Kattenbeck, M. & Elsweler, D. (2019). Understanding credibility judgements for web search snippets. *Aslib Journal of Information Management* 71(3), 368-391. <https://doi.org/10.1108/AJIM-07-2018-0181>
- Kayid, A (2020). The role of Artificial Intelligence in future technology. *International Journal of Innovative Research in Advanced Engineering (IJIRAE)*, Issue 04, Volume 5, 146-148. <https://doi.org/10.26562/IJIRAE.2018.APAE10086>
- Khan, W., Daud, A., Nasir, J.A. & Amjad, T. (2016). 'A survey on the state-of-the-art machine learning models in the context of NLP'. *Kuwait Journal of Science*, 43 (4), pp. 95-113. <https://journalskuwait.org/kjs/index.php/KJS/article/view/946>
- Khurana, D., Koli, A., Khatter, K & Singh, S. (2017). Natural Language Processing: State of The Art, Current Trends and Challenges. <https://arxiv.org/abs/1708.05148>
- Kondylakis, H., Tsirigotakis, D., Fragkiadakis, G., Panteri, E., Papa-dakis, A., Fragkakis, A., Tzagkarakis, E., Rallis, I., Saridakis, Z., Trampas, A., Pirounakis, G., & Papadakis, N. (2020). R2D2: A Dbpedia Chatbot Using Triple-Pattern Like Queries. *Algorithms*, 13, 217. <http://doi.org/10.3390/a13090217>

- Kumari, A. & Thakur, J. (2019). Semantic web search engines : A comparative survey. *International Journal of Scientific Research in Computer Science Engineering and Information Technology*, 5(1). <https://doi.org/10.32628/CSEIT195115>
- Langville, A. N., & Meyer, C. D. (2012). *Google's PageRank and Beyond: The Science of Search Engine Rankings*. Princeton University Press.
- Laudon, K., & Laudon, J. (2017). *Management Information Systems: Managing the Digital Firm* (15th ed.). Pearson.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- López De Luise, D., Pascal A., Alvarez C., Tournoud M., Pankrac C., Cruz J.M.S. (2021) Information Retrieval in Restricted Domain for ChatterBots. In: Balas V., Jain L., Balas M., Shahbazova S. (eds) *Soft Computing Applications. SOFA 2018. Advances in Intelligent Systems and Computing, vol 1222*. Springer, Cham. https://doi.org/10.1007/978-3-030-52190-5_23
- Maedche, A., Morana, S., Schacht, S., Werth, D., & Krumeich, J. (2016). Advanced User Assistance Systems. *Business & Information Systems Engineering*, 58(5), 367–370. <https://doi.org/10.1007/s12599-016-0444-2>
- Maroengsit, W., Piyakulpinyo, T., Phonyiam, K., Pongnumkul, S., Chaovalit, P., & Theeramunkong, T. (2019). A Survey on Evaluation Methods for Chatbots. *Proceedings of the 2019 7th International Conference on Information and Education Technology - ICIET 2019*, 111–119. <https://doi.org/10.1145/3323771.3323824>
- Marx, E., Khalili, A., & Valdestilhas, A. (2017). Semantic Search User Interface Patterns: An Introduction. In *Human Computer Interaction (HCI-Europe). WSCG*. Plzen, Czech Republic. <http://hdl.handle.net/11025/29622>
- Mercieca, J. (2019). Human-chatbot interaction - Assessing technology acceptance, confidence and trust in chatbots within their application areas. (Master's thesis). City University of London, London, England. <https://doi.org/10.13140/RG.2.2.28452.09600>
- Mitchell, T. M. (1997). *Machine Learning (McGraw-Hill International Editions Computer Science Series)* (1st ed.). McGraw-Hill.
- Mohamed, M. H., Khafagy, M. & Ibrahim, M. H. (2019). Recommender systems challenges and solutions survey. *International Conference on Innovative Trends in Computer Engineering (ITCE)*, Aswan, Egypt, 149-155. <https://doi.org/10.1109/ITCE.2019.8646645>
- Monett, D. Lewis, C. & Thórisson, K. (2020). Introduction to the JAGI Special Issue “On Defining Artificial Intelligence” — Commentaries and Author’s Response. *Journal of Artificial General Intelligence*, 11(2) 1-100. <https://doi.org/10.2478/jagi-2020-0003>

- Nimavat, Ketakee & Champaneria, Tushar. (2017). Chatbots: An overview. Types, Architecture, Tools and Future Possibilities. *IJSRD - International Journal for Scientific Research & Development* Vol. 5, Issue 07, 1019-1024. <http://ijsrd.com/Article.php?manuscript=IJSRDV5170501>
- Osisanwo, F. Y., Akinsola, J. E. T., Awodele, O., Hinmikaiye, J. O., Olakanmi, O., & Akinjobi, J. (2017). Supervised Machine Learning Algorithms: Classification and Comparison. *International Journal of Computer Trends and Technology*, 48(3), 128–138. <https://doi.org/10.14445/22312803/ijctt-v48p126>
- Papenmeier, A., Sliwa, A., Kern, D., Hienert, D., Aker, A. and Fuhr, N. (2020). 'A Modern Up-To-Date Laptop' - Vagueness in Natural Language Queries for Product Search. *Proceedings of the 2020 ACM Designing Interactive Systems Conference*. Association for Computing Machinery, New York, NY, USA, 2077–2089. <https://doi.org/10.1145/3357236.3395489>
- Peffer, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems*, 24(3), 45–77. <https://doi.org/10.2753/mis0742-1222240302>
- Prajwal, S., Mamatha, G., Ravi, P., Manoj, D., & Joisa, S.K. (2019). Universal Semantic Web Assistant based on Sequence to Sequence Model and Natural Language Understanding. *2019 9th International Conference on Advances in Computing and Communication (ICACC)*, 110-115. <https://doi.org/10.1109/ICACC48162.2019.8986173>
- Ramesh G.S., Nagaraju G., Harish V., Kumaraswamy P. (2021). Chatbot for College Website. In: Kiran Mai C., Kiranmayee B.V., Favorskaya M.N., Chandra Satapathy S., Raju K.S. (eds) *Proceedings of International Conference on Advances in Computer Engineering and Communication Systems. Learning and Analytics in Intelligent Systems, vol 20*. Springer, Singapore. https://doi.org/10.1007/978-981-15-9293-5_47
- Roh, Y., Heo, G., & Whang, S. E. (2019). A Survey on Data Collection for Machine Learning: A Big Data - AI Integration Perspective. *IEEE Transactions on Knowledge and Data Engineering*, 1. <https://doi.org/10.1109/tkde.2019.2946162>
- Rohilla, N. (2017). Web Content Management System (WCMS): A Need. *International Journal Of Engineering And Computer Science*, 21729. <https://doi.org/10.18535/ijecs/v6i6.29>
- Ruane, E., Faure, T., Smith, R., Bean, D., Carson-Berndsen, J., & Ventresque, A. (2018). *BoTest. Proceedings of the 23rd International Conference on Intelligent User Interfaces Companion*. <https://doi.org/10.1145/3180308.3180373>
- Russell, S., & Norvig, P. (2016). *Artificial Intelligence: A Modern Approach (Pearson Series in Artificial Intelligence)* (3rd ed.). Pearson.
- Sa, N., Yuan, X. (2020). Challenges in conversational search: Improving the system capabilities and guiding the search process. *WMSCI 2020 - 24th World Multi-Conference on Systemics, Cybernetics and Informatics, Proceedings. Volume 3, 2020, Pages 37-42*. <https://www.iiis.org/CDs2020/CD2020Summer/PapersS3.htm#/>

- Salman, R., & Kecman, V. (2012). Regression as classification. *2012 Proceedings of IEEE Southeastcon*, 1. <https://doi.org/10.1109/secon.2012.6196887>
- Samal, B. R., Behera, A. K., & Panda, M. (2017). Performance analysis of supervised machine learning techniques for sentiment analysis. *2017 Third International Conference on Sensing, Signal Processing and Security (ICSSS)*, 1. <https://doi.org/10.1109/ssps.2017.8071579>
- Samuel, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, 3(3), 210–229. <https://doi.org/10.1147/rd.33.0210>
- Sanderson, M. & Croft, W. (2012). The History of Information Retrieval Research. in *Proceedings of the IEEE*, vol. 100, no. Special Centennial Issue, pp. 1444-1451. <https://doi.org/10.1109/JPROC.2012.2189916>
- Sedkaoui, S. (2018). Supervised versus Unsupervised Algorithms: a Guided Tour. *Data Analytics and Big Data*, 123–151. <https://doi.org/10.1002/9781119528043.ch7>
- Serra, A., & Tagliaferri, R. (2019). Unsupervised Learning: Clustering. *Encyclopedia of Bioinformatics and Computational Biology*, 350–357. <https://doi.org/10.1016/b978-0-12-809633-8.20487-1>
- Shawar, B.A., & Atwell, E. (2007a). Chatbots: Are they Really Useful? *LDV Forum*, 22, 29-49. https://ijcl.org/content/2-allissues/21-Heft1-2007/Bayan_Abu-Shawar_and_Eric_Atwell.pdf
- Shawar, B.A., & Atwell, E. (2007b). Different measurement metrics to evaluate a chatbot system. HLT-NAACL 2007. <https://aclanthology.org/W07-0313/>
- Shivakumar, (2019). *Build a Next-Generation Digital Workplace: Transform Legacy Intranets to Employee Experience Platforms*. 1st. ed. Apress
- Singh A., Ramasubramanian K., Shivam S. (2019). A Novel In-House Implementation of a Chatbot Framework. In: *Building an Enterprise Chatbot*. Apress, Berkeley, CA. https://doi.org/10.1007/978-1-4842-5034-1_6
- Srivastav, M. & Nath, A. (2016). Web Content Management System. *International Journal of Innovative Research in Advanced Engineering (IJIRAE)*. 3. 51-56. <https://www.ijirae.com/volumes/Vol3/iss3/10.MRAE10092.pdf>
- Srivastava, S., & Prabhakar, T. V. (2020). Intent Sets: Architectural Choices for Building Practical Chatbots. *Proceedings of the 2020 12th International Conference on Computer and Automation Engineering (ICCAE 2020)*, 194–199. <https://doi.org/10.1145/3384613.3384639>
- Sutton, R. (2019). John McCarthy's Definition of Intelligence. *Journal of Artificial General Intelligence*, vol.11, no.2, 2019, pp.66-67. <https://doi.org/10.2478/jagi-2020-0003>
- Tatar, A., de Amorim, M. D., Fdida, S., & Antoniadis, P. (2014). A survey on predicting the popularity of web content. *Journal of Internet Services and Applications*, 5(1), 1–20. <https://doi.org/10.1186/s13174-014-0008-y>

Teng, W., Wen, W. & Liu, Y. (2012). From experience to expertise: Digesting cumulative information for informational web search. *Journal of Information Science and Engineering*, 28(1):161-176.
<https://www.researchgate.net/publication/220587940>

The Internet World Stats (2021). *World Internet Users and 2021 Population Stats*.
<https://www.Internetworldstats.com/stats.htm>

Vtyurina, A., Savenkov, D., Agichtein, E., & Clarke, C. (2017). Exploring Conversational Search With Humans, Assistants, and Wizards. *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. <https://doi.org/10.1145/3027063.3053175>

Wan, S. Li, D. & Gao, J. (2016). Exploring the Advantages of Content Management Systems for Managing Engineering Knowledge in Product-service Systems. *Procedia CIRP*, Volume 56, Pp 446-450. <https://doi.org/10.1016/j.procir.2016.10.087>

Wang, P. (2019). "On Defining Artificial Intelligence". *Journal of Artificial General Intelligence*, vol.10, no.2, 2019, pp.1-37. <https://doi.org/10.2478/jagi-2019-0002>

Yadav, A. A., Garg, I., & Mathur, D. P. (2019). PACT - Programming Assistant ChaTbot. *2019 2nd International Conference on Intelligent Communication and Computational Techniques (ICCT)*, 1. <https://doi.org/10.1109/icct46177.2019.8969070>

Younis, K. & Alkhateeb, A. (2017). A new implementation of deep neural networks for optical character recognition and face recognition. *Proceedings of the New Trends in Information Technology (NTIT-2017)*, The University of Jordan, Amman, Jordan. 25-27 . p. 157–62.
https://www.researchgate.net/publication/324877673_A_New_Implementation_of_Deep_Neural_Networks_for_Optical_Character_Recognition_and_Face_Recognition

Zhao, M., Yan, S., Liu, B., Zhong, X., Hao, Q., Chen, H., Niu, D., Long, B., & Guo, W. (2021). QBSUM: a Large-Scale Query-Based Document Summarization Dataset from Real-world Applications. *Computer Speech & Language*, Volume 66, 101166. <https://doi.org/10.1016/j.csl.2020.101166>

Zhu, G. & Iglesias, C. A. (2018). Exploiting semantic similarity for named entity disambiguation in knowledge graphs. *Expert Systems with Applications*, 101, 8-24.
<https://doi.org/10.1016/j.eswa.2018.02.011>

Anexo I

Questionário para análise de utilizadores e tarefas

Questionário sobre a utilização de chatbots para apoio à pesquisa de conteúdos na Internet

No âmbito de um estudo que pretende investigar a viabilidade da utilização de sistemas de chatbot para auxiliar na tarefa de realizar pesquisas de conteúdos em websites, gostaria de contar com a sua ajuda para responder a um questionário que consiste em 26 perguntas de escolha múltipla. O tempo estimado para responder ao questionário é de 10 a 15 minutos.

Os dados recolhidos no questionário serão tratados de forma anónima e confidencial e as respostas serão utilizadas exclusivamente para fins científicos.

Agradeço, desde já, a sua disponibilidade e participação.

* Required

Declaração

Os dados recolhidos neste questionário serão tratados de forma anónima e confidencial.
As respostas serão utilizadas exclusivamente para fins científicos.

1. Confirmo que li e entendi as informações presentes neste formulário e aceito participar neste estudo. *

Selecione uma opção

Mark only one oval.

- Sim *Skip to question 2*
- Não

Caracterização demográfica

2. 1 - Indique o seu género com que mais se identifica. *

Selecione uma opção

Mark only one oval.

- Feminino
- Masculino
- Prefiro não responder

3. 2 - Indique a faixa etária em que se insere *

Selecione uma opção

Mark only one oval.

- Menos de 24 anos
- 25 - 44 anos
- 45 - 64 anos
- + 65 anos

4. 3 - Qual o nível de escolaridade mais elevado que completou? *

Selecione uma opção

Mark only one oval.

- Ensino Básico
- Ensino Secundário
- Ensino Superior
- Other: _____

Hábitos de pesquisa

5. 4 - Com que frequência efetua pesquisas na internet? *

Selecione uma opção

Mark only one oval.

- Várias vezes por dia
- Uma vez por dia
- Algumas vezes por semana
- Uma vez por semana
- Com menos frequência
- Nunca *Skip to question 11*

6. 5 - Costuma encontrar os resultados que procura à primeira tentativa? *

Selecione uma opção numa escala de 1 a 5, onde 1 é "Nunca" e 5 corresponde a "Sempre"

Mark only one oval.

- Sempre
- Quase sempre
- Às vezes
- Raramente
- Nunca

7. 6 - Como costuma encontrar conteúdos na Internet? Que ferramentas utiliza? *

Selecione uma ou mais opções

Check all that apply.

- Motores de busca globais (Google, Yahoo Bing)
- Motor de busca integrado num website
- Diretórios de websites
- Navegação no website

8. 7 - Que aspetos valoriza numa ferramenta de pesquisa? *

Selecione uma ou mais opções

Check all that apply.

- Apresentação gráfica
- Acesso rápido
- Recursos de navegação (menus, ícones, botões...)
- Clareza da informação
- Possuir FAQs ou ajudas
- Links para outras ferramentas ou aplicações

Other: _____

9. 8 - Onde costuma efetuar pesquisas? *

Selecione uma ou mais opções

Check all that apply.

- Em casa
- No trabalho/escola
- Em qualquer lado

Other: _____

10. 9 - Quais são os aspetos que considera mais importantes quando efetua uma pesquisa na internet? *

Selecione uma ou mais opções

Check all that apply.

- Conseguir encontrar rapidamente a resposta que procuro
- Conseguir encontrar facilmente a resposta que procuro (utilizando o menor número de passos possível)
- Obter resultados precisos e de fácil compreensão
- Obter apenas resultados que correspondam ao contexto da pesquisa
- Obter resultados completos e imparciais, baseados apenas na sua relevância e não em manipulações que favoreçam algum site, marca ou empresa
- Poder filtrar as minhas pesquisas para incluir ou excluir determinadas palavras para obter resultados mais precisos
- Poder ordenar os resultados obtidos por data de publicação

Other: _____

Perfil tecnológico

11. 10 - Como classifica o seu conhecimento relativamente à utilização das tecnologias de informação e comunicação? *

Selecione uma opção numa escala de 1 a 5, onde 1 é "Muito Mau" e 5 corresponde a "Muito bom"

Mark only one oval.

	1	2	3	4	5	
Muito Mau	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Muito bom

12. 11 - Qual (ou quais) do(s) seguinte(s) dispositivo(s) utiliza para aceder à Internet? *

Selecione uma ou mais opções

Check all that apply.

- Computador
- Tablet
- Smartphone
- Consolas de jogos
- TV

Other: _____

13. 12 - Com que frequência realiza as seguintes atividades na Internet? *

Selecione uma opção para cada atividade

Mark only one oval per row.

	Mais do que uma vez por dia	Uma vez por dia	Uma ou mais vezes por semana	Uma ou mais vezes por mês	Menos de uma vez por mês	Nunca
Trabalhar/estudar	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Pesquisar conteúdos	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ler notícias	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Redes Sociais	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Compras online	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Instant Messaging (ex: WhatsApp)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Jogos Online	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

14. 13 - Que redes sociais e ferramentas de comunicação utiliza na Internet? *

Selecione uma ou mais opções

Check all that apply.

- Facebook
 - Facebook Messenger
 - WhatsApp
 - Instagram
 - Twitter
 - LinkedIn
 - Tik Tok
 - Telegram
 - Reddit
 - Discord
 - Slack
- Other: _____

Chatbots

Um chatbot é um programa de computador programado para simular e processar conversas com utilizadores humanos em linguagem natural.

15. 14 - Considera que a utilização de chatbots como ferramenta de apoio à pesquisa de conteúdos na internet é uma ideia interessante? *

Selecione uma opção

Mark only one oval.

- Extremamente interessante
- Muito interessante
- Interessante
- Pouco interessante
- Nada interessante

16. 15 - Já alguma vez utilizou um chatbot? *

Mark only one oval.

- Sim *Skip to question 17*
- Não *Skip to question 22*

Utilização de chatbots

17. 16 - Em que contextos já utilizou chatbots? *

Selecione uma ou mais opções

Check all that apply.

- Apoio ao Cliente
- Educação
- Entretenimento
- Finanças
- Pesquisa de conteúdos
- Saúde e bem-estar
- Vendas

Other: _____

18. 17 - Com que frequência utiliza chatbots ou agentes virtuais como a Alexa, o Google Assistant, a Siri ou a Cortana? *

Selecione uma opção

Mark only one oval.

- Uma ou mais vezes por semana
- Uma ou mais vezes por mês
- Uma vez a cada três meses
- Raramente
- Nunca usei *Skip to question 22*

19. 18 - Porque optou por utilizar um chatbot em vez de outras opções disponíveis? *

Selecione uma ou mais opções

Check all that apply.

- Acesso mais rápido
- Fácil de usar
- Era uma forma mais rápida de obter uma resposta a uma pergunta
- Não fui capaz de encontrar por mim mesmo as respostas que procurava
- Não havia outras formas de obter a informação que procurava
- Já tinha tido experiências positivas ao utilizar o chatbot
- Queria experimentar o chatbot

Other: _____

20. 19 - Que tipo de resposta(s) prefere obter de um chatbot? *

Selecione uma ou mais opções

Check all that apply.

- Resposta simples de texto
- Resposta avançada que combine vários elementos (texto, imagem, link)
- Link para uma página onde possa encontrar a resposta
- Link para uma página onde possa encontrar mais informação sobre como obter a resposta

Other: _____

21. 20 - Recomendaria a um amigo a utilização de chatbots para pesquisar conteúdos na Internet em vez dos motores de busca ou barras de pesquisa tradicionais? *

Selecione uma opção

Mark only one oval.

Sim

Não

Chatbots para pesquisa de conteúdos na Internet

22. 21 - Ao contrário de um motor de busca tradicional, que devolve todos os resultados associados a uma palavra ou expressão pesquisada, muitas vezes fora de contexto, o chatbot devolve-lhe apenas o resultado mais relevante, com base no contexto da conversa. Considera que esta funcionalidade torna os chatbots mais apelativos como ferramenta de apoio à pesquisa de conteúdos na Internet? *

Selecione uma opção

Mark only one oval.

Sim

Não

23. 22 - Com que frequência utilizaria um chatbot para pesquisar conteúdos num determinado website? *

Selecione uma opção

Mark only one oval.

Sempre

Quase sempre

Ocasionalmente

Raramente

Nunca

24. 23 - Utilizaria o chatbot por si só ou seria mais uma ferramenta de apoio a outras formas de pesquisa? *

Selecione uma opção

Mark only one oval.

- Utilizaria como ferramenta principal de pesquisa
- Utilizaria como ferramenta de apoio à pesquisa em conjunto com outras soluções

25. 24 - Que características gostaria de ver incluídas num chatbot de apoio dedicado à pesquisa de conteúdos na Internet? *

Selecione uma ou mais opções

Check all that apply.

- Conversas personalizadas
- Análise de sentimento e inteligência emocional (tentar identificar a emoção por trás de um texto)
- Obter resumos dos conteúdos gerados por inteligência artificial sem ter necessidade de visitar o website original
- Possibilidade de transferir a conversa para um assistente em tempo real
- Possibilidade de transferir a conversa diretamente para outro chatbot (por exemplo, um chatbot de outro website)
- Capacidade de partilhar facilmente um conteúdo com outra pessoa
- Receber um relatório da conversa no seu email
- Possibilidade de integrar o chatbot com o software que utilizo diariamente (por exemplo, o software que uso no decorrer da minha atividade profissional)
- Possibilidade de reconhecer e traduzir diretamente os resultados apresentados para outras línguas
- Mais opções de acessibilidade

Other: _____

Anexo II

Guião de Testes com Utilizadores

Preparação

Onde se irão realizar os testes?

Os testes serão realizados ao ar livre, na zona da Grande Lisboa

Qual o equipamento necessário?

- Smartphone, computador ou tablet
- Browser (Chrome, Firefox, Safari, outro)
- Ligação à Internet para aceder à interface ao *website* do protótipo

O que é necessário completar?

É necessário completar os seguintes passos por ordem:

1. Responder ao questionário inicial;
2. Completar as três tarefas enunciadas;
3. Responder ao questionário de satisfação.

Introdução

Caro voluntário,

Agradeço desde já a sua disponibilidade para testar o protótipo.

Os testes que irá realizar em seguida servirão exclusivamente para avaliar o desempenho da interface.

Estes testes consistem em três tarefas, que terá de completar sem ajuda.

O tempo total estimado para sua é inferior a 10 minutos.

Antes da realização dos testes ser-lhe-á pedido para responder a um pequeno questionário inicial e no final deverá responder a um questionário de satisfação.

Todos os dados recolhidos nos questionários e nos testes serão tratados de forma anónima e serão apenas utilizados no âmbito deste estudo.

Pode optar por desistir dos testes a qualquer momento.

Tarefas a realizar

1ª Tarefa

Descrição:

Indicar que se trata de uma primeira interação e aceitar registar-se para obter resultados mais personalizados; obter uma notícia relacionada com uma marca; obter uma notícia relacionada com desporto motorizado; obter uma notícia relacionada com veículos elétricos; terminar interação

Medidas de Usabilidade:

Eficácia: Concluir a tarefas proposta

Eficiência: Concluir as tarefas sem encontrar erros no sistema

Satisfação: Escala de Likert

Critérios de Usabilidade:

Eficácia: Mais de 70% dos utilizadores serão capazes de concluir a tarefa proposta

Eficiência: Menos de 30% dos utilizadores ficarão confusos/insatisfeitos durante a realização da tarefa.

Satisfação Menos de 10% dos utilizadores ficarão insatisfeitos durante a realização da tarefa.

2ª Tarefa

Descrição:

Indicar que se trata de uma primeira interação e avançar sem aceitar registar-se para obter dados mais personalizados; escolher um tipo de veículo; indicar a(s) marca(s) preferida(s); indicar que não tem interesse em notícias de desporto motorizado; indicar interesse em veículos elétricos; terminar interação

Medidas de Usabilidade:

Eficácia: Concluir a tarefas proposta

Eficiência: Concluir as tarefas sem encontrar erros no sistema

Satisfação: Escala de Likert

Critérios de Usabilidade:

Eficácia: Mais de 70% dos utilizadores serão capazes de concluir a tarefa proposta

Eficiência: Menos de 30% dos utilizadores ficarão confusos/insatisfeitos durante a realização da tarefa.

Satisfação Menos de 10% dos utilizadores ficarão insatisfeitos durante a realização da tarefa.

3ª Tarefa

Descrição:

Indicar que já está registado; introduzir o email usado para o registo; conversar livremente; escolher tipo de veículo; indicar que não tem interesse em notícias de desporto motorizado nem em notícias de veículos elétricos. multibanco;

Medidas de Usabilidade:

Eficácia: Concluir a tarefas proposta

Eficiência: Concluir as tarefas sem encontrar erros no sistema

Satisfação: Escala de Likert

Critérios de Usabilidade:

Eficácia: Mais de 70% dos utilizadores serão capazes de concluir a tarefa proposta

Eficiência: Menos de 30% dos utilizadores ficarão confusos/insatisfeitos durante a realização da tarefa.

Satisfação Menos de 10% dos utilizadores ficarão insatisfeitos durante a realização da tarefa.

Anexo III

Questionário de avaliação do *chatbot*

Questionário sobre utilização de chatbots para apoio à pesquisa de conteúdos na Internet

TESTAR O PROTÓTIPO E RESPONDER AO QUESTIONÁRIO (TEMPO TOTAL ESTIMADO: 15-20 MINUTOS)

TESTAR O CHATBOT

O "Rodas" é um protótipo de chatbot desenvolvido no âmbito de um estudo que pretende investigar a viabilidade da utilização de sistemas de chatbot para auxiliar na tarefa de realizar pesquisas de conteúdos em websites. Trata-se de um agente conversacional que interage com o utilizador no sentido de devolver os conteúdos mais recentes relacionados com os interesses do mesmo.

Este protótipo, que se encontra numa fase muito embrionária, foi desenvolvido para funcionar em dois websites, um relacionado com notícias de automóveis, e outro relacionado com notícias de motos.

O teste do protótipo deverá ser feito através da realização de três tarefas, descritas na seguinte localização: Terminado o teste, deverá preencher o questionário na sua totalidade.

Este estudo pretende avaliar o chatbot através do grau de satisfação do utilizador e na capacidade de completar com sucesso as tarefas propostas.

Agradeço, desde já, a sua disponibilidade e participação.

* Required

Questionário pré-teste

1. 1 - Indique o seu género *

Selecione uma opção

Mark only one oval.

Feminino

Masculino

2. 2 - Indique a faixa etária em que se insere *

Selecione uma opção

Mark only one oval.

Até 24 anos

25 - 44 anos

45 - 64 anos

+ 65 anos

3. 3 - Qual o nível de escolaridade mais elevado que completou? *

Selecione uma opção

Mark only one oval.

Ensino Básico

Ensino Secundário

Ensino Superior

4. 4 - Já utilizou chatbots?

Mark only one oval.

Sim

Não

Questionário Pós-teste

5. Confirmando que realizei o teste ao protótipo *

Mark only one oval.

Sim

6. 1 - Como classifica a sua experiência de interação com o chatbot "Rodas"? *

Selecione uma opção numa escala de 1 a 5, onde 1 é "Muito negativa" e 5 corresponde a "Muito positiva"

Mark only one oval.

	1	2	3	4	5	
Muito negativa	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Muito positiva

7. 2 - Como classifica o nível de dificuldade na utilização do chatbot. *

Selecione uma opção numa escala de 1 a 5, onde 1 é "Muito difícil" e 5 corresponde a "Muito fácil"

Mark only one oval.

	1	2	3	4	5	
Muito difícil	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Muito fácil

8. 3 - Consegui terminar com sucesso todas as tarefas propostas? *

Mark only one oval.

- Consegui terminar todas as tarefas
- Consegui terminar duas tarefas
- Consegui terminar apenas uma tarefa
- Não consegui concluir nenhuma tarefa

9. 4 - Conseguiu obter respostas satisfatórias com base nas tarefas propostas?
Indique, por favor, a eficácia do chatbot. *

Selecione uma opção numa escala de 1 a 5, onde 1 é "Muito pouco eficaz" e 5 corresponde a "Muito eficaz"

Mark only one oval.

	1	2	3	4	5	
Muito pouco eficaz	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Muito eficaz

10. 5 - Considera que o chatbot foi eficiente na execução das tarefas? *

Selecione uma opção numa escala de 1 a 5, onde 1 é "Muito pouco eficiente" e 5 corresponde a "Muito eficiente"

Mark only one oval.

	1	2	3	4	5	
Muito pouco eficiente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Muito eficiente

11. 6 - Durante a realização de cada tarefa, o chatbot deixou de funcionar corretamente ou foi necessário recomeçar a interação com o chatbot?

Mark only one oval.

- Sim
 Não

12. 7 - Recomendaria o chatbot "Rodas" a outras pessoas? *

Mark only one oval.

- Sim
 Não